**Proceedings of Virtual Concept 2005**
**Biarritz, France, November 8th – November 10th, 2005**

# Semi-Automated Ultrasound Interpretation System Using Anatomical Knowledge Representation

Michael S. Downes [1], Brian A. Barsky [2]

**(1) :** University of California, Berkeley
EECS Computer Science Division
535 Soda Hall
Berkeley CA 94720
U.S.A.
*(510) 642-9716*
*E-mail* : msdownes@yahoo.com

**(2) :** University of California, Berkeley
EECS Computer Science Division
387 Soda Hall # 1776
Berkeley CA 94720-1776
U.S.A.
*(510) 642-9838*
*E-mail* : barsky@cs.berkeley.edu

**Abstract:** Interpreting ultrasound data presents a significant challenge to medical personnel, which limits the clinical applications of the technology. We have addressed this issue by developing a prototype computer-based system designed to aid non-expert medical practitioners in using ultrasound devices in a variety of different diagnostic situations. Essentially, the system treats the collection of images generated during an ultrasound examination as an ordered sequence of views of the anatomical environment and picks out key views in which the contents of the scan image changes. It stores descriptions of expected key views and matches incoming images to this key view sequence during an orientation phase of an examination. The prototype can guide a novice user through an examination of a patient's abdomen and automatically identify anatomical structures within the region. Overall, the design represents a novel approach to processing and augmenting ultrasound data and to representing spatial knowledge.

**Key words**: ultrasound; anatomical knowledge representation; anatomical structure labelling; 3D anatomical models; spatial knowledge representation; automated medical imaging systems

## 1- Introduction

A change in perspective can often make even a well-known scene unrecognizable. Perhaps you've seen your old, familiar city from the top of a brand new building. Maybe you've walked through a neighborhood a hundred times before one day finally seeing it through the window of a bus. In these situations, as in a myriad of others, you've probably had that strange sensation that the familiar has suddenly and at least momentarily become incomprehensible.

Doctors must often confront the problems associated with looking at familiar objects from unusual perspectives. In particular, such problems frequently arise when doctors attempt to make use of ultrasound technology. Almost all medical imaging, including ultrasound, consists of 2D cross-sectional pictures, or slices, of anatomical structures. As a result, a medical professional who wishes to make use of most imaging data must learn to adapt his or her knowledge of anatomy to accommodate the cross-sectional view afforded by imaging technologies. Specialists devote years to focused training in order to develop an expertise for interpreting such data. If ultrasound devices could be made easier to use for non-experts, then medical practitioners, like battlefield medics, doctors on a space craft or space station, aid workers in remote locations, and others who have no easy access to trained experts could still reap some of the diagnostic benefits of using ultrasound devices.

With this in mind, our goal is to help such professionals to overcome these obstacles. Specifically, we have developed a novel, flexible, and efficient view-based high-level representation for anatomical knowledge. The model incorporates aspects of human expert ultrasound interpretation strategies. We have used this model to create a prototype computer-based ultrasound interpretation system that augments ultrasound data in real-time in order to make such data accessible to medical professionals who are not ultrasound experts.

In a departure from most previous work in medical augmented reality, the current system focuses on identifying anatomical structures and labeling them in 2D images rather than on reconstructing highly accurate 3D models of structures. In this case, the system generates simple 3D models with knowledge of their related anatomical structures and uses these models to create image labels.

The prototype demonstrates a unique method for representing anatomical knowledge as sequences of descriptions of 2D views. The concept has analogues in the autonomous navigation literature, and it also allows the system to learn to recognize variations on standard anatomical configurations.

The following section presents a description of related earlier work. Section three describes key object identification techniques used by human sonography experts, and section four provides details of the design of the ultrasound interpretation system. The final two sections report the results of running the system on real scans and give plans for future work.

## 2- Previous Work

Our work represents a synthesis of concepts from a variety of different sources. In particular, the design has been heavily influenced by techniques from the ultrasound literature and by work in the fields of autonomous navigation and medical augmented reality. For example, we have conducted a study of the ultrasound examination techniques taught to sonography students. This study included a review of relevant literature along with conversations with sonography experts and students and observations of students during training.[1] The next section presents a summary of various exam techniques, but overall, the study made clear that many of the object identification strategies taught to sonography students could be adapted for use in computer-based systems designed to interpret ultrasound data.

This realization of the connection between human strategies for accomplishing what are essentially spatial tasks and possible computer-based approaches to performing the same functions came from an examination of part of the literature on autonomous navigation. On the face of it, there might seem to be little relationship between navigation in a large-scale environment and using an ultrasound device. However, ultrasound devices do provide a user with a limited view of a complex space that is not unlike the limited view one has when walking through a city. In addition, sonography students learn to recognize certain important, distinctive structures during exams and to use consistent structures as guides to more complex anatomy just as people learn to recognize landmarks in cities and to use maps to find their way. A very large amount of research has focused on developing computational models for large-scale navigation, because most autonomous robots must be capable of at least some degree of successful navigation. In particular, Kuiper's work on representing spatial knowledge has had wide influence.[2,3] He has developed a number of computational frameworks for modeling the type of spatial knowledge that many psychologists theorize is contained within a person's cognitive map, or their internal representation of environmental spaces. In essence, these computational models consist of techniques for representing spatial information and methods for applying this information when choosing to take actions, like moving through a space. Reviewing this work led us to consider the possibility of modeling the types of spatial knowledge and actions necessary to perform an ultrasound exam. Together, the object identification techniques used in ultrasound and previous work on autonomous navigation provided the motivation for the underlying techniques used in our system.

The motivation for the actual application came in large part from previous work in using augmented reality for biomedical applications. For instance, projects at the University of North Carolina, Chapel Hill (UNC) have focused on augmented reality visualizations for ultrasound data and for laparoscopic surgery.[4,5,6,7] The UNC ultrasound augmented reality system displays volume rendered 3D ultrasound data accurately registered with a patient during a breast needle biopsy procedure and during fetal examinations. The system seeks to make ultrasound data more useful to a surgeon by presenting it in a way that is intended to be more intuitive than the standard display. The work represents a significant achievement in motion tracking, rendering, interface design, and other areas, but it requires costly specialized hardware and still relies to a large extent on the expertise of the ultrasound user to interpret the displayed information. We set out to try to augment ultrasound data in a manner that focuses less on visualization and more on structure identification. As mentioned above, our goal is to help non-experts to derive benefits from using an ultrasound device for diagnostic purposes.

## 3- Object Identification Techniques in Ultrasound

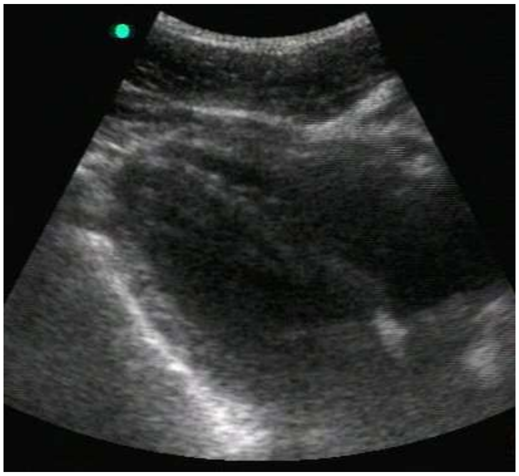


**Figure 1** An ultrasound scan image of a heart. The image shows the left and right ventricles.

Figure 1 presents an example of an ultrasound scan image. To the untrained eye, the image appears to contain nothing more than visual noise. Even an extensive knowledge of human anatomy probably would not be enough to allow one to accurately interpret this image. How, then, do sonography experts make sense of the data they must analyze? They use a variety of spatially and visually oriented techniques to identify anatomical structures in the scans they examine. As it happens, many of these techniques can be adapted to serve as the foundation for our novel computer-based scanning system. A few such techniques are highlighted below. For a full discussion of the topic, the reader should refer to the relevant section in [8].

When examining some anatomical regions, sonography experts use a particular structure, or collection of structures, as a type of map or guide with which they can orient themselves to a particular patient's anatomy. Generally, one chooses a guide structure based on its consistency of appearance and location from one patient to another and on its ease of visual identification in sonographic images. For example, when examining the abdomen, sonographers often use the portal venous system as a guide. In this case, the structure consists of a network of blood vessels that retains essentially the same topography in most patients. In ultrasound scan images like the one in figure 2, the tubes that compose the portal venous system show up as clear black regions, so they are relatively easy to pick out.

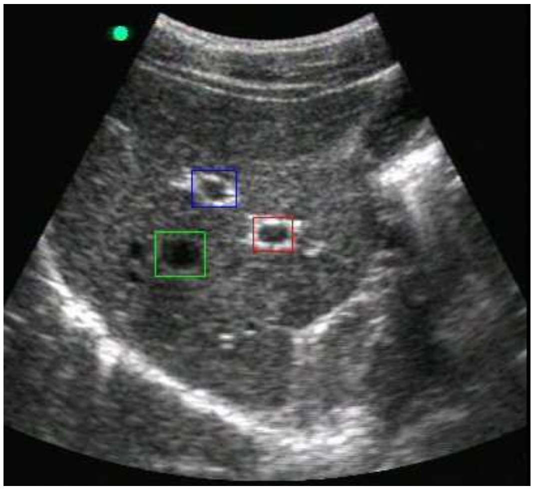
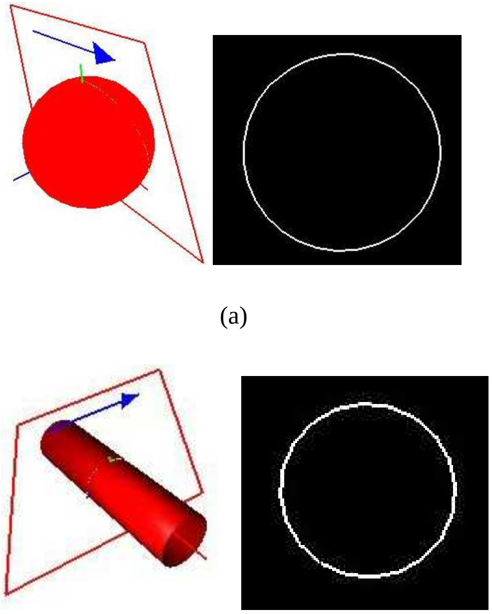


**Figure 2** Ultrasound scan of a section of the portal venous system. Elements of the PVS are highlighted with colored rectangles.

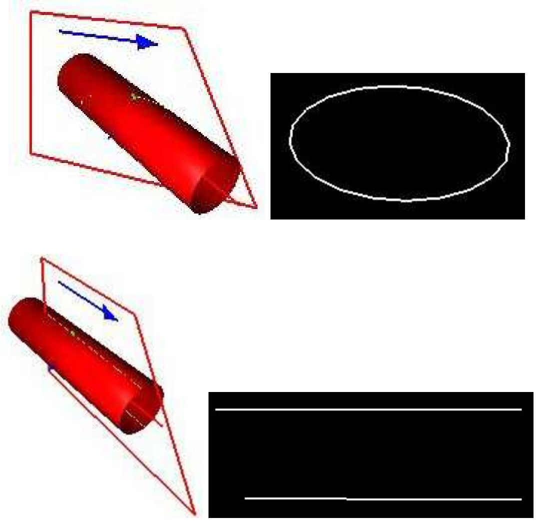
Echographers use a guide structure to match their knowledge of standard anatomy to the patient at hand. By finding and examining such a structure, an ultrasound user can determine the configurations of other harder-to-identify structures whose positions can be found relative to the guide structure. This concept brings to light an obvious, but important, point. Sonography experts approach each new exam with expectations as to what they will see. They have prior knowledge of and experience with the anatomy to be examined, and they use this information to interpret what they see during an exam. In the case of guide structures, the structure helps to orient a general mental representation to fit a specific situation.

Ultrasound users also impose an ordering on most sonographic exams to further aid in image interpretation. In most exams scans are usually collected in the same order for each patient. The examiner places the ultrasound probe in specified locations on a patient's body and sweeps the probe in particular directions. For example, an exam of the abdomen generally begins on the right side of the body with the probe oriented parallel to the long axis of the body. The examiner sweeps the probe to the left until the entire abdomen has been examined and then places the probe at the bottom of the abdomen and sweeps in the direction of the head. In echocardiographic exams, the probe is moved from one sonographic window to another in a specific sequence in order to collect a complete set of images of the heart's component structures. (In this context, a window is a region on the body in which an ultrasound probe can be placed to get an image of a particular structure. Since ultrasound waves cannot penetrate bone or air-filled spaces, an echocardiographer must find scanning positions in which the ribs and lungs do not block the heart from view.) These scanning sequences help to ensure that the ultrasound user does not miss imaging any important structures. They also provide the examiner with yet another source of expectations. A scanning sequence essentially defines an order in which one can expect to encounter various anatomical structures, so an expert could view a recording of an exam conducted by someone else and still be able to interpret the resulting image sequence.

Ultrasound images, like the one shown in figure 1, often seem to include extremely complex shapes. However, many complex anatomical structures can be approximated by collections of simple shapes that make deciphering such images easier. For example, Weyman suggests a simple approximation of the heart that consists of cylinders, spheres, a tetrahedron, and a pear.[9] Each one of these objects has a limited set of possible cross-sectional images that can be generated by intersecting the object with a scan plane. For example, figure 3 shows the different cross sections possible for a cylinder and a sphere. An ultrasound user can learn these simple cross sections and the approximate heart model and then use all of this information to interpret ultrasound scans of the heart. Knowing the simple shapes to expect in a particular view of the heart gives one a basis on which to pick out and identify the more complex shapes present in a real scan.



(a)



(b)

**Figure 3** The possible cross-sectional shapes generated by cutting (a) a sphere and (b) a cylinder

Human sonography experts spend years perfecting their abilities to use the techniques described above. Such expertise will always be necessary to interpret complex and highly unusual cases, but similar techniques can be adapted to allow a computer-based system to interpret ultrasound examinations and assist non-experts in using ultrasound devices. Our semi-automated computer-based scanning system design incorporates many of these object identification approaches. For instance, the system stores an internal representation of a guide structure approximated by simple objects and uses knowledge of the scanning order to organize its stored information. The scanner also incorporates knowledge of the expected locations of various structures into its exam strategy and identifies some structures based on their spatial relationships to other structures.

The prototype we have constructed focuses on assisting in examinations of the upper abdomen. As mentioned above, such an examination generally proceeds from right to left and from feet to head across the abdomen. In addition, sonography experts often use the portal venous system as a guide to examining the region. Our system uses the same strategies.

## 4- Abdominal Scanning System

### 4.1 – Overview

Our prototype ultrasound scanning assistant guides a user through an initial ultrasound examination of a patient's abdomen and uses the information acquired during the exam to identify anatomical structures for the user during subsequent free scans. After orienting itself to the patient's anatomy during the initial directed scanning phase, the program can label the elements of the portal venous system, or PVS, and other structures as they appear in the ultrasound image during user-directed scanning. The program can also generate a rough 3D representation of the patient's PVS.

The system hardware consists of a portable ultrasound device whose image output is fed to a standard pc running Linux. A six-degree of freedom optical motion tracker tracks the position and orientation of both the ultrasound probe and the patient's body. Figure 4 shows the prototype system.

**Figure 4** The components of the prototype computer-based scanning system.

From a high-level perspective, the system software operates as follows. The program first creates an initial representation of the standard anatomy for the region it will examine. This process occurs off-line and need only happen once. During the orientation phase of an examination, the program uses image processing techniques to isolate and classify shapes in an input scan image, and then it uses its stored knowledge representation to identify the anatomical component that relates to each shape in the image. After the orientation phase, the system uses a registered 3D model of the patient-specific anatomy to label structures in scan images.

The subsections that follow cover the details of the knowledge representation used in the system. Subsection 4.2 introduces the concepts of view descriptions and key views and describes how they can be used to represent anatomical knowledge in a flexible and extendable manner. Subsection 4.3 details the approach used to allow the system to learn simple anatomical variations. Finally, the last subsection describes the techniques used to construct 3D models of a patient's anatomy and to label 2D ultrasound images during an examination.

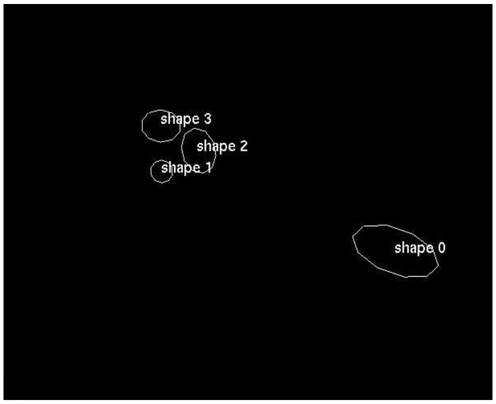
### 4.2 – View Descriptions and Key Views

In order to relieve the user of some of the burden of recognizing structures during an ultrasound exam, the computer-based scanner must maintain some sort of representation of the anatomy being examined. As mentioned in section 3, human beings approach an ultrasound examination with at least some concept of what they might see during the examination. Based on training and past experience, they have a set of expectations for how anatomical structures will appear in scans. For novices, these expectations might be based entirely on their knowledge of the standard configuration of the anatomy to be examined. In this case, patient-specific variations from the standard anatomy could easily lead to misidentification of structures. More experienced sonography experts develop the ability to incorporate variations into their mental representations of anatomical structures. The computer-based scanning system must make use of a representation scheme that gives it some of this same flexibility.

The scanning system design uses a scheme that stems from advice given to almost every new sonography student. Instructors frequently tell their students to focus during an exam on what they see on the ultrasound screen, rather than on the patient's body or on their own scanning hand's position in space. One could interpret this advice as an indication that one should concentrate primarily on understanding the anatomy as it appears in 2D cross-sectional images, as opposed to worrying about exact positioning in 3D space. The computer-based scanning system reflects this advice by representing learned anatomy as an ordered sequence of key views.

As discussed in section 3, many ultrasound exams proceed according to a clearly defined scanning order. For example, in the abdominal exam on which the prototype system focuses, scans are normally collected from right to left and from feet to head across a patient's abdomen. The system stores a sequence of key views in this order.

The key views stand as the core of the program's knowledge representation. They are stored not as images but as text descriptions of the contents of a scan. A view description lists the simple shapes contained within the view and their associated anatomical structures, and it identifies the 2D spatial relationships between all of these shapes. Figure 5 shows a variety of examples of view descriptions. For the sake of clarity, the spatial relationships in the figure are listed as above/below and right/left. However, for each shape the system actually stores a bearing angle toward each other shape in the view. Thus, a shape A that is directly above another shape B would have a bearing of 0 degrees from shape B. These bearings allow the system to define spatial relationships accurately.



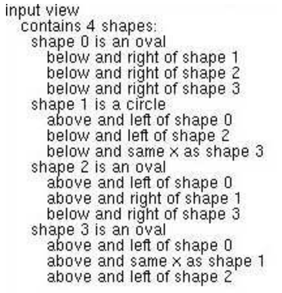

```
input view
  contains 4 shapes:
    shape 0 is an oval
      below and right of shape 1
      below and right of shape 2
      below and right of shape 3
    shape 1 is a circle
      above and left of shape 0
      below and left of shape 2
      below and same x as shape 3
    shape 2 is an oval
      above and left of shape 0
      above and right of shape 1
      below and right of shape 3
    shape 3 is an oval
      above and left of shape 0
      above and same x as shape 1
      above and left of shape 2
```

**Figure 5** Sample view description.

The key view sequence describes the configuration of the standard anatomy by identifying the essential components of the scan images one would expect to encounter as one performs an exam of the anatomical region of interest. Each key view marks a point in the scan sequence at which something changes in the ultrasound scan images of the standard anatomy. These changes include the appearance and disappearance of shapes in the 2D image along with changes in the type of a shape, like a change from a circle to an ellipse. A new key view can also be generated when the 2D spatial relationships between shapes in the scan image change.

The use of text-based view descriptions provides a number of advantages over using images or 3D models. For instance, the view descriptions eliminate metrical information, so a single description can be used for many patients. The text descriptions also retain the essential topological relationships between structures that are key to successfully identifying them. In addition, by storing an ordered sequence of key view descriptions, the program essentially maintains a representation of the three dimensional structure of the anatomy being examined. The order of the key views gives this added dimension by ensuring that each 2D slice occurs at some predefined point in the sequence.

Before it can be used to examine a patient, the computer-based scanning system must be loaded with a key view sequence for the anatomical region to be examined. A human expert can generate this sequence by hand by simply writing a list of view descriptions for the key cross-sectional views encountered during an exam. However, if even a simple 3D computer graphics model of the relevant standard anatomy exists or can be created, then the program can generate a sequence of key views automatically. The process consists of performing a virtual ultrasound exam of the 3D model. In order to perform the exam, the program must know how to move the ultrasound probe to acquire the proper sequence of scan images. In a case like the abdominal exam used in the prototype system, the scanning procedure can be easily translated into straightforward instructions. For example, to scan a virtual model of the portal venous system, the program positions the probe on the far right of the space with the scan plane parallel to the long axis of the body and then translates the probe to the left. In other cases, like an examination of the heart, the probe motions necessary to generate the desired sequence can be somewhat more difficult to translate into transformation instructions, so the program can also read in recorded probe motions. In this case, a human expert moves a tracked virtual probe through an exam of the 3D model, and the program uses the recorded probe motion to control its own exam.

Once the program knows how to move the virtual ultrasound probe to generate the desired scan sequence, it can begin the process of generating key views. During the exam, the program moves the virtual probe incrementally through the desired motion sequence. The size of the incremental motion can be adjusted to ensure an adequate level of accuracy. After each move, the program computes the intersection of the scanning plane with the 3D model and generates a 2D image of this intersection. Figure 6 shows an example of a probe position and the related cross-sectional image generated during key view creation for the abdominal scanner. Once it has created a scan image, the program uses image-processing techniques to determine the types of all of the 2D shapes in the image. When computing the intersection of the scan plane with the 3D model, for each shape in the resulting image, the program stores the identity of the 3D structure for which the shape is a cross section. Thus, the program has all the information it needs to generate a view description.

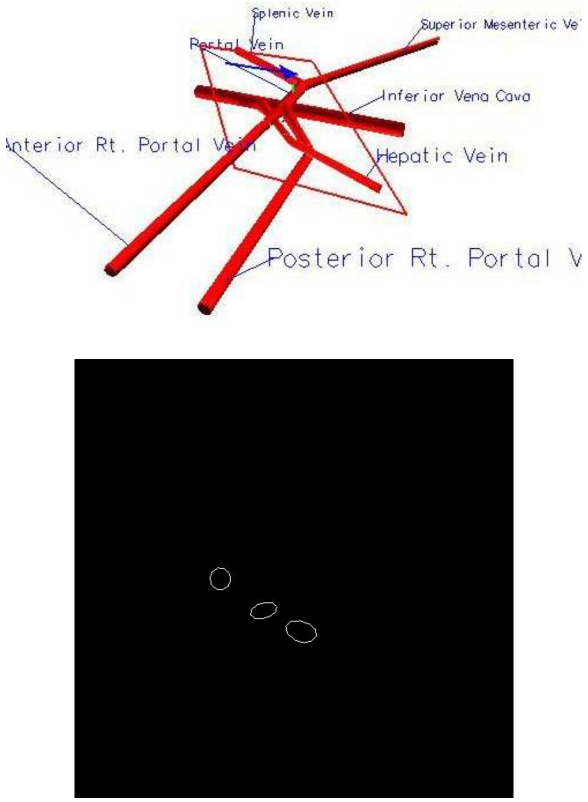


**Figure 6** Part (a) shows the probe position, indicated by the red outline with a blue arrow on top, that generates the 2D cross-sectional image shown in (b)

After creating a view description, the program must determine whether or not it should generate a new key view based on the current scan. The system compares the current view description to the most recently generated key view description. If the two differ in any of the ways mentioned above, the program stores the current view description as the next key view in the ordered key view sequence. Once the virtual exam has come to an end, a human expert can review the generated key view sequence and add or remove view descriptions if desired.

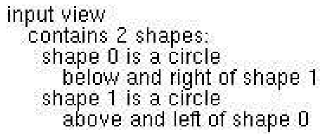
With a key view sequence in place, the scanning system can be used to examine patients. In essence, a real exam proceeds very much like the virtual exam used to generate key views. However, in a real exam the input scans come not from intersections with a virtual model but from a real ultrasound machine. The system generates a view description for an input image and compares this description to the expected key views in order to determine the anatomical structures that relate to the shapes in the input image. An image processing module segments out target shapes in the input scan images and associates them with simple shapes. The current shape recognition code uses a principal component analysis-based technique to compare input shapes to a training set.[10,11] The program passes each input shape in turn to the recognition code which outputs the simple shape that most closely matches the input. Details of the image processing techniques used in the prototype can be found in [8]. As will be described in subsection 4.4, once the ordered exam has been completed, the system also has enough information to help the user to identify structures in arbitrary scans taken from any vantage point in the region of interest.

The key view sequence also provides another important benefit. By representing the standard anatomy as an ordered sequence of views, the program can easily add to its knowledge by including multiple options for different views in the sequence in order to deal with anatomical variations that alter some of the spatial relationships that ordinarily remain consistent from patient to patient. The process of adapting to variations is examined at length in the next subsection.

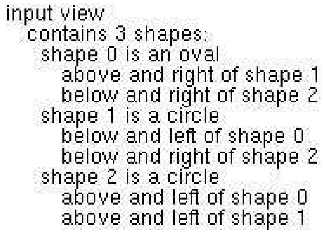
### 4.3 – Learning New Anatomical Variations

No two human beings have identical anatomies. Even structures, like the portal venous system, that remain relatively consistent from patient to patient do in fact vary to some degree on occasion. The flexibility imparted by abstracting away much of the metrical information associated with a representation of the standard version of the anatomy to be examined allows the scanning system to remain robust in the face of simple variations from the standard. However, larger variations that alter some of the key spatial relationships between components of the standard structure would lead to identification failures if the system were not able to adapt to new situations. The scanning system uses a collection of simple techniques to update its anatomical knowledge as it performs exams. As it encounters variations, it incorporates them into its representation of the relevant anatomy and it becomes more adept at successfully scanning subsequent patients.

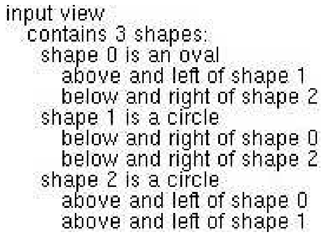
Rather than storing key views as a simple linear sequence, the scanning system actually stores them as a tree. After performing the initial virtual scan of a 3D model of the relevant anatomy, the tree does in fact look like a linear sequence with only one transition leading out from each node. Figure 7 shows a simple initial key view sequence. The system could then begin running on patients. However, if an expert user knows of some common anatomical variations that could be encountered during an exam, then the expert can add key view options to the tree to describe these variations. During normal operation, the system incorporates these key view options into its exam by searching when appropriate for all of the possible key view options present at its current location in the tree.



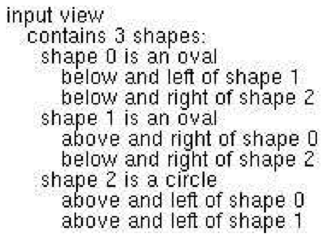
```
input view
   contains 2 shapes:
     shape 0 is a circle
        below and right of shape 1
     shape 1 is a circle
        above and left of shape 0
```



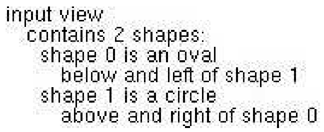
```
input view
   contains 3 shapes:
     shape 0 is an oval
        above and right of shape 1
        below and right of shape 2
     shape 1 is a circle
        below and left of shape 0
        below and right of shape 2
     shape 2 is a circle
        above and left of shape 0
        above and left of shape 1
```



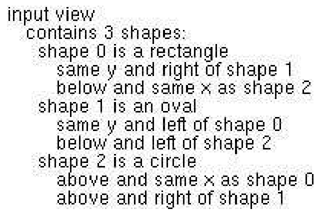
```
input view
   contains 3 shapes:
     shape 0 is an oval
        above and left of shape 1
        below and right of shape 2
     shape 1 is a circle
        below and right of shape 0
        below and right of shape 2
     shape 2 is a circle
        above and left of shape 0
        above and left of shape 1
```



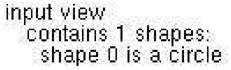
```
input view
   contains 3 shapes:
     shape 0 is an oval
        below and left of shape 1
        below and right of shape 2
     shape 1 is an oval
        above and right of shape 0
        below and right of shape 2
     shape 2 is a circle
        above and left of shape 0
        above and left of shape 1
```



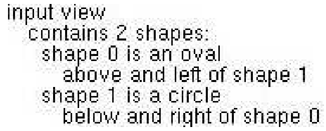
```
input view
   contains 2 shapes:
     shape 0 is an oval
        below and left of shape 1
     shape 1 is a circle
        above and right of shape 0
```



```
input view
   contains 3 shapes:
     shape 0 is a rectangle
        same y and right of shape 1
        below and same x as shape 2
     shape 1 is an oval
        same y and left of shape 0
        below and left of shape 2
     shape 2 is a circle
        above and same x as shape 0
        above and right of shape 1
```



```
input view
   contains 1 shapes:
     shape 0 is a circle
```



```
input view
   contains 2 shapes:
     shape 0 is an oval
        above and left of shape 1
     shape 1 is a circle
        below and right of shape 0
```

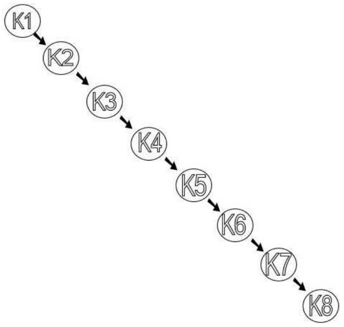


**Figure 7** A sequence of key views generated for the 3D model shown in Figure 6.

The scanning program can also create key view options on its own if it has access to 3D models of the relevant anatomy that include structural variations on the standard configuration. In this case, the program can scan each alternative 3D model in turn and add the key view options necessary to correctly describe the structure in each case. For instance, figure 8 shows a 3D anatomical model of the same structure shown in figure 6. In this case, though, the structure differs somewhat from the one shown in figure 7 in that the upper tube section is completely above the lower section, so the two tubes that are intersected by the scanning plane in figure 6 do not surround the lower tube segment in the model in figure 8. Instead, they are both above the segment. The initial exam of the new 3D model proceeds as shown with the first key view matching as expected. However, at scan 2 the program encounters an unexpected view that does not match the current key view or the next one in the key view sequence. The program stores the view description of the unexpected view as a second child of the previously matched key view. After finding the new view, the program adds all subsequent views as new key views that descend from the first unknown. Figure 9 shows the key view tree after adding the necessary options to deal with the model from figure 8. During an exam, either real or virtual, when the system matches a key view that has multiple children, it compares each new input to all of the child options until it matches the next appropriate key view.

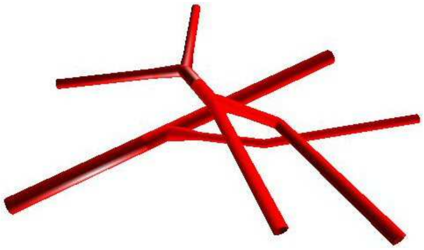
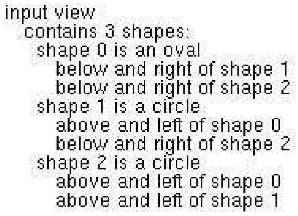


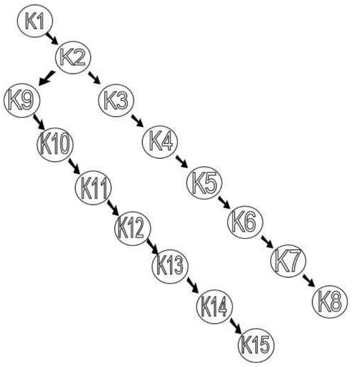
**Figure 8** A variation on the object shown in figure 6. In this case

the upper structure is completely above the lower structure.



```
input view
   contains 3 shapes:
     shape 0 is an oval
        below and right of shape 1
        below and right of shape 2
     shape 1 is a circle
        above and left of shape 0
        below and right of shape 2
     shape 2 is a circle
        above and left of shape 0
        above and left of shape 1
```

(a)



(b)

**Figure 9** (a) The new second key view (K9) describing the model shown in figure 8 (b) The same key view tree from figure 7 with extra options added to reflect the variation shown in figure 8. Key views 9 and 10 are new. Key views 11 – 15 match key views 4 – 8.

The system not only learns key view options from 3D models of anatomical variations, but can also use a similar simple procedure to learn new variations dynamically during a real exam. When the program encounters an unexpected view during normal operation, the view is added to the key view tree as an option under the previously matched key view. The system compares the number of shapes in the new view to the number present in the last matched key view. If the new view has the same number of shapes as does the last known view, then the system uses an assumption of coherence between adjacent scans to posit an identification of the anatomical structures that relate to each of the shapes in the new view. For each shape in the new view, the program associates it with the structure that generated the shape in the previous view that is closest to the current shape in the scan image.

If the new, unexpected view contains a different number of shapes than does the last known view, then structure identification becomes more complex. The system cannot simply propagate the structure identities from the previous view, because at least one structure has just appeared in or disappeared from the view. The program must defer identification until it can find a known view later in the exam sequence.

Whether or not the shape count for the unexpected view matches that of the last known view, after adding the new view to the key view tree, the program proceeds to continue trying to match known key views based on the assumption that the patient's anatomy should not vary too greatly from known configurations. When the system finds a matching key view, it tries to propagate the structure identifications in this view back through any preceding unknown views in the same way as was described above for forward propagation. However, if the system has encountered a number of unexpected views in a row, and the shape counts change during the sequence of unknown views, then it will not be possible to identify the structures in every new view. In this case, an expert would need to examine the scan sequence and identify unknown structures. Whenever the program encounters an unexpected view, it alerts the user and flags all subsequent views as uncertain even if structure identifications have been posited. When automatic structure identification fails, the program warns the user that it has encountered not only unexpected but unidentifiable scans and directs the user to complete the guided scanning phase and then terminate the exam. For any case in which the program finds unexpected scans, an expert should be consulted to verify the program's identifications for structures in the new scans. The system also allows a user to back out any options added to the key view tree during a real exam in order to avoid including false information in the tree. However, if the program correctly identifies structures in new key views, then it can use these new views to recognize similar variations in future exams.

As mentioned above, the scanning system could face a situation in which it must process multiple unknown views consecutively. The program still tries to find a match to a known key view, but it must also maintain some sense of the original sequence or risk making incorrect matches. Each key view can be identified not only by its contents but also by its position in the overall sequence. Two key views might contain the same simple shapes in the same spatial configuration, but the structures associated with the shapes in each image might be different. The views are delineated by their different positions in the key view sequence. Thus, even if it encounters a string of unexpected key views, the program must advance through the tree of key views as the exam progresses in order to avoid matching a key view that might contain the same shapes as the input but cannot occur in the region currently being scanned. In order to maintain the proper scan sequence, each key view includes a notation indicating the 3D region in which it was matched in the training model. The program scales and positions these regions based on tracked markers placed on the patient's body and on the locations of any successfully matched key views. When it encounters a sequence of unknown views, the program uses the regions assigned to key views to know when to move on from searching for one set of views to searching for a subsequent set in order to try to pick up the thread of recognition, as it were. Of course, as mentioned in previous sections, general metrical information can be notoriously unreliable when translated to a specific patient, but the system treats the key view regions as being only approximations and allows neighboring regions to overlap at the transition boundaries. In addition, the regions are only used when no other information is available to continue an exam.

### 4.4 – Building 3D Models and Labeling

After guiding the user through an ordered examination of a patient's abdomen, the prototype system has enough information to label anatomical structures in arbitrary ultrasound scans of the region. The initial exam essentially serves to translate the program's general background knowledge of the expected anatomy into a form that reflects the specific details of the patient being examined. Once the system has this patient-specific information, the user can freely explore the anatomical region and concentrate on any structures that might be relevant for the particular diagnostic situation. For example, the user might be interested primarily in determining the damage done by a gunshot to a specific part of the abdomen.

The program represents its patient-specific knowledge by generating a simple approximate 3D model of the anatomy being examined. For instance, in the case of the abdominal scanning prototype, the system creates a simple model of the portal venous system. The patient-specific 3D model is generated based on a combination of the input scan data and the anatomical knowledge stored in the system's key view tree. For each input image, the program first finds the simple shapes in the image. It then finds a matching key view and uses this view to identify the anatomical structures whose cross sections are the shapes in the image. The program also stores the 3D position and orientation of each simple shape along with the related anatomical structure for each one. Since the system tracks the ultrasound probe during the exam, the program can determine the approximate 3D locations of the shapes in the image based on the dimensions of the image. In order to create an approximate 3D model of the patient's anatomy, the program uses the simple shapes as cross sections for 3D objects and stitches together adjacent shapes in 3D. Figure 10 demonstrates the process for a small number of scans. By the end of the guided exam, the program has created a simple 3D model in which all of the components are correctly identified. The model also lines up with the patient's body and matches the patient's anatomy.

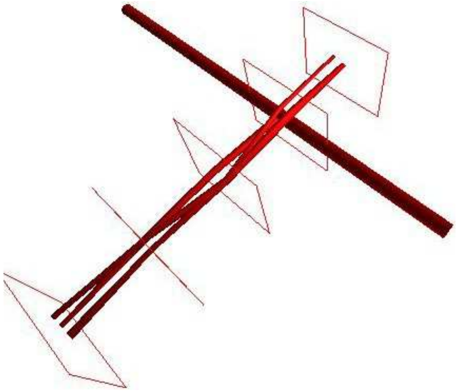


**Figure 10** 3D patient-specific model constructed by stitching together adjacent cross sections. Rectangles indicate the positions and orientations of scans.

When the user begins to scan freely, the program uses the tracking data for the ultrasound probe to compute the intersection of the scanning plane with the patient-specific 3D model. The intersection produces a virtual scan image that shows a cross section of the 3D model. Since the program stores the names of the anatomical structures represented by the various components of the 3D model, it can print the appropriate names next to the cross-sectional shapes in the virtual scan. The program overlays this image on top of the real ultrasound image and attaches text labels to the shapes in the virtual scan. The simple 3D model approximates the patient's anatomy, so the outlines in the virtual scan image, and thus the labels, align with their associated real shapes in the actual scan image. Of course, the 3D model does not align perfectly with the real anatomy, so the labels do not sit at the exact centers of the shapes to which they refer. However, they align well enough that the user can easily recognize which shapes are being labeled. This is in fact an important strength of the scanning system. It does not need to construct extremely accurate, registered reconstructions of a patient's anatomy in order to provide the user with the benefits of its stored anatomical knowledge.

## 5- Results

Figure 11 shows four labeled frames from an ultrasound examination of a subject's abdomen. The labels were generated in real time by our system. The rectangles indicate the simple shapes found in the images.

In order to examine the impact of our system on non-experts' performances on representative ultrasound tasks, we have conducted a small pilot user study. In the study, we asked participants to find tumors in ultrasound scans of an abdomen and to identify the component of the portal venous system nearest to each tumor. We used an image manipulation program to create virtual tumors in normal ultrasound scans by adding bright, blurred circular regions to the images.

The experiment consisted of two groups. Participants in the control group viewed standard ultrasound images, while those in the experimental group viewed the same images but with labels generated by our system for some components of the portal venous system. Results from the study indicate that our system helps participants to accomplish the task more quickly and more accurately than without labels. Based on these results, we plan to conduct a larger study.

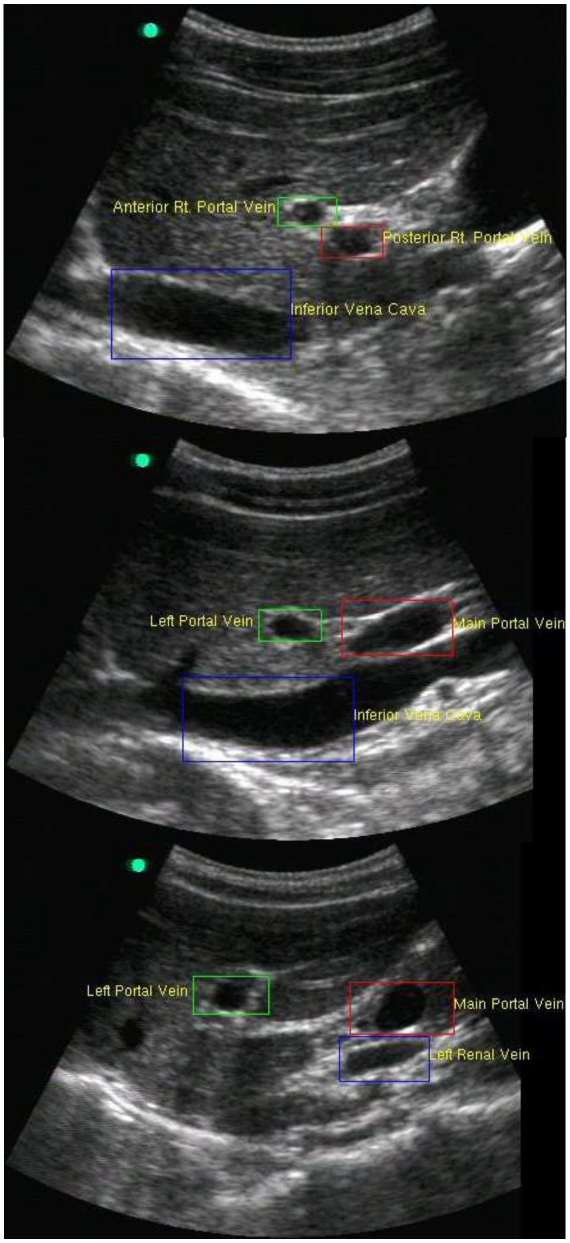


**Figure 11**. Sample images from an assisted ultrasound examination.

## 6- Conclusion

The design concepts used in the prototype can be extended to create systems that operate in other anatomical regions. Along this line, we plan to extend the prototype to work with scans of the heart. This will require improvements to the image-processing component of the system, but the underlying framework for knowledge representation will remain unchanged. In addition, we intend to optimise the key view tree construction and traversal processes, and we are also exploring the possibility of using the scanning system to create a searchable, indexed library of ultrasound scans.

## 7- References

**[1]** Tacy, Theresa. Personal communications, 2003.

**[2]** Kuipers, B. The skeleton in the cognitive map: a computational hypothesis. In J. Peponis, J. Wineman and S. Bafna (Eds.), Space Syntax: Proceedings of the 3rd International Symposium, Ann Arbor: A. Alfred Taubman College of Architecture and Urban Planning, University of Michigan, 10.1–10.7, 2001

**[3]** Kuipers, B. Modeling spatial knowledge. Cognitive Science 2: 129-153, 1978

**[4]** Ohbuchi, R., D. Chen, and H. Fuchs. Incremental Volume Reconstruction and Rendering for 3D Ultrasound Imaging. In Visualization in Biomedical Computing, SPIE Proceedings, vol. 1808: 312-323, 1992

**[5]** State, A., D. Chen, C. Tector, A. Brandt, H. Chen, R. Ohbuchi, M. Bajura, and H. Fuchs. Case Study: Observing a Volume-Rendered Fetus within a Pregnant Patient. In Proceedings of IEEE Visualization '94, Los Alamitos, Calif., 1994

**[6]** State, A., G. Hirota, D. Chen, W. Garrett, and M. Livingston. Superior Augmented-Reality Registration by Integrating Landmark Tracking and Magnetic Tracking. In Proceedings of SIGGRAPH 96, New Orleans, LA, 1996

**[7]** State, A., M. Livingston, G. Hirota, W. Garrett, M. Whitton, H. Fuchs, and E. Pisano. Technologies for Augmented-Reality Systems: realizing Ultrasound-Guided Needle Biopsies. In Proceedings of SIGGRAPH 96, New Orleans, LA, 1996

**[8]** M. Downes. Augmenting Ultrasound Data. Doctoral thesis, University of California, Berkeley, 2005.

**[9]** Weyman, A. Cross-Sectional Echocardiography. Philadelphia, PA: Lea & Febiger, 1982

**[10]** Black, M. and P. Anandan. A framework for the robust estimation of optical flow. In Proc. Int Conf. On Computer Vision, ICCV-93, Berlin, Germany, 1993

**[11]** Black, M. and A. Jepson. EigenTracking: Robust Matching and Tracking of Articulated Objects Using a View-Based Representation. Tech. Report T95-00515, Xerox PARC, Dec. 1995