AD_____

Award Number: W81XWH-05-1-0101

TITLE: Development and Evaluation of Stereographic Display for Lung Cancer Screening

PRINCIPAL INVESTIGATOR: Xiao Hui Wang, M.D., Ph.D.
                                              Walter F. Good
                                              Carl R. Fuhrman
                                              Howard E. Rockett
                                              David Gur

CONTRACTING ORGANIZATION: University of Pittsburgh
                                                     Pittsburgh, PA, 15260

REPORT DATE: December 2006

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
                           Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
                                           Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01-12-2006	Annual	8 Nov 2005 – 7 Nov 2006

4. TITLE AND SUBTITLE

Development and Evaluation of Stereographic Display for Lung Cancer Screening

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-05-1-0101

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Xiao Hui Wang, M.D., Ph.D.   Carl R. Fuhrman

Walter F. Good   Howard E. Rockett
David Gur

E-Mail: xwang@mail.magee.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Pittsburgh
Pittsburgh, PA, 15260

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
Original contains colored plates: ALL DTIC reproductions will be in black and white.

14. ABSTRACT

SEE ATTACHED PAGE

15. SUBJECT TERMS
lung cancer screening, stereo display, volumetric rendering, observer performance study

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			
U	U	U	UU	34	**19b. TELEPHONE NUMBER** *(include area code)*

ABSTRACT

The main purpose of this project is to investigate the feasibility and efficacy of using a stereo display workstation for lung cancer screening on CT images. The tasks included in this project are development and evaluation of stereo image projection and display for chest CT images, observer performance evaluation for the stereo display, and stereo feature analysis and comparison to the conventionally used display methods for lung cancer detection. In the previous report period, we have built a stereo display workstation for chest CT images and conducted a pilot observer performance study. In this annual report period, we have continued the study based on the projected tasks as listed below. 1. Analyzing the results from the pilot study: we applied Free-response Receiver Operating Characteristic (FROC) statistic method to analyze the data from the pilot study for lung nodule detection and classification. Results indicate that the stereo display achieved the best performance followed by the slice-by-slice display, and the conventional MIP display gave the worst performance, although there is no statistically significant difference between the three display modes. Subjective assessment indicates that the stereo display was well accepted by the radiologists. Efficiency measurement indicates that the radiologists spent the least interpretation time with the stereo display when compared to the other two display modes. Further analysis of the radiologists' interpretation patterns indicates that novelty and training effect substantially influenced the radiologists' interpretation behavior and performance. The conclusion from the preliminary results is that we have observed a potential role of stereo display for improving radiologists' performance in medical detection and diagnosis, and also observed some factors likely affecting the performance with new display, such as novelty, training effect and confidence with the new technology, including the stereo display. Appropriate training and practice is necessary for achieving optimal performance with 3D display device and new display technology. 2. Implementing advanced features for the stereo display: we have tested the feasibility and efficacy of performing 3D rendering on GPUs (Graphics Processing Units) for stereo display of medical images. Our GPU-based program achieved real-time rendering, real-time displaying and real-time interactive controls by radiologists, which is desirable and necessary for prompt and accurate medical diagnosis. 3. Conducting main observer performance study: we have started the main observer performance study that uses larger database and improved study design based on the feedbacks from the pilot study.

Table of Contents

INTRODUCTION

Lung cancer is a leading cause of death in the United States [1,2]. The results from several large lung cancer screening studies indicate that early detection and treatment can reduce mortality rate in most types of lung cancer cases [3-6]. Currently, low-dose CT scanner is a primary tool used for lung cancer screening. For each screening case, a set of image slices covering entire lung area is generated and viewed on display workstations. Despite of 3D format of CT datasets, the conventional reading method for lung CT image interpretation is to read images slice-by-slice. This reading method requires radiologists to mentally reconstruct images in 3D space from a set of 2D images to differentiate normal tubular structures from nodules. Furthermore, with improved technology for CT scanner, higher resolution imaging techniques produce more images per scan, which eventually will exceed radiologists' ability to read cases in slice-by-slice mode. The need of 3D data presentation of CT images has become crucial for ever-increasing numbers of images generated from CT scanner and for improvement of radiologists' performance on image data interpretation. We have proposed to develop a stereo display workstation for reading lung CT images. Stereopsis is the mechanism used in human vision system to perceive objects in our three dimensional space. The 3D display using stereoscopic projection should produce a natural and efficient solution for 3D data presentation. In this proposal, we hypothesized that the efficacy of lung cancer screening using CT scanned images can be increased by use of a suitable designed stereoscopic display. Specifically, we expect that both efficiency, and accuracy for the detection of lung nodules, will be increased significantly over what can be achieved when reading cases in currently used display modes. To achieve the goals in this proposal, we have specified our aims as followings:

1) Develop and integrate the hardware and software required to implement a stereoscopic display tailored to chest CT images.
2) Use a subset of lung cancer cases, verified either by pathology or by followup, to evaluate the display system.
3) Perform a retrospective study to measure relative accuracy and reading efficiency, for detection and classification of lung nodules, between three display modes including stereoscopic 3D mode from this project, and other two commonly used modes, slice-by-slice and maximum-intensity-projection (MIP) thick slice.

BODY

Implement Secondary Features (Task 5)

As we proposed in the project, in this report period we continued development of certain display features we consider to be of value in this application. These features will become important in broadening the applicability of stereo displays for chest CT. The main secondary features implemented were volume projection and rotation in real-time with GPU (Graphics Processing Unit) card.

Recent advanced commodity GPUs are very efficient at manipulating and displaying computer graphics for a range of complex algorithms. The advanced features of GPUs are especially useful to medical practice, in which data interpretation is timely dependent, extensive interactions are required, and multiple format of data presentation in real time is desired for different diagnostic purposes.

Applying programmable GPUs is likely a solution for real-time stereo image compositing and display. In this particular application, the tasks that the GPUs can facilitate for lung CT stereo display include stereo pair compositing from CT data set at desired viewing position and viewing volume, multiple rendering algorithms, brightness/contrast adjustment, and image rotation.

The GPU-based program has achieved real-time rendering and real-time display effect without any perceptive delay in each successive frame rendering and display following a user controlled frame switch command. We found no difference in frame rate between MIP and average renderings. To test and demonstrate the real-time stereo rendering process on GPU card, we used lung CT images for quantitative measurements of the performance shown in table 1 and 2.

4

Table 1. Frame rates measured as stereo pairs per second for rendering on GPU card and CPU (Central Processing Unit) card at different number of interpolated slices.

Number of interpolated slices	GPU (stereo pairs per second)	CPU (stereo pairs per second)
1	103.3	-
9	20.1	1.3
15	13.2	0.8
21	10.1	0.5
33	6.6	0.3
45	5.0	0.2

Table 2. Frame rates measured as stereo pairs per second for rendering on GPU card with and without rotation implementation.

Number of interpolated slices	Rotation implemented (stereo pairs per second)	Without rotation (stereo pairs per second)
1	103.3	103.3
3	44.4	44.4
5	33.7	33.7
9	20.1	20.1
15	13.2	13.2
21	10.1	10.1
33	6.6	6.6
45	5.0	5.0

Our results indicate that programming on GPUs can not only avoid lengthy process of precalculation and overloaded disc space, but also provide some functionalities that would be virtually impossible for the prestaged process, such as rotations and real-time interpolations (for example, changing image resolution). The GPUs solution has shown to be efficient for real-time stereo pair renderings and display. We have submitted our results for peer-reviewed publication (Appendix A).

Analysis of Pilot Study (Task 10)
There were a total of 286 nodule-like features found in the pilot study. Since each study case was interpreted in 3 display modes by 8 radiologists, any nodule-like feature in the study case could be found 24 times if the feature were detected by all of the radiologists in all of the display modes. Figure 1 shows the distribution of number of times features were detected, for example the leftmost bin represents the features that were found only by one radiologist in one display mode, so these features were the least agreeable ones; whereas the rightmost bin represents the features were found by all of the radiologists in all of the display modes, so these features were the most agreeable ones. This distribution gives general idea of the variability of inter-readers and inter-modes for lung nodule detection.
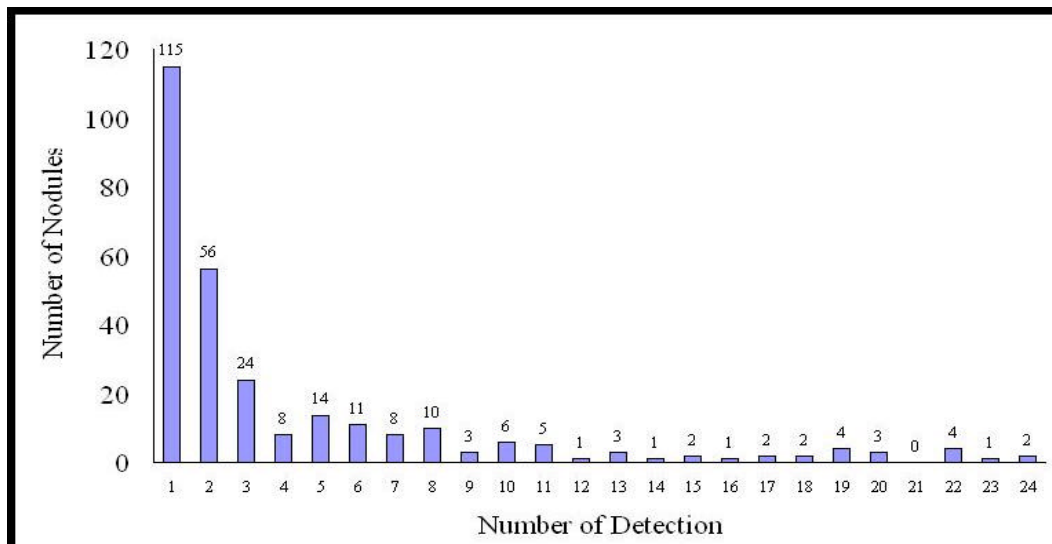

Figure 1. Distribution of number of times features were detected.

To reach a consensus result for nodule verification and nodule truth profile, we pooled features detected from the eight radiologists' interpretation in the three display modes. These features were reviewed and verified by an experienced chest radiologist, who did not participate the study but had read and discussed the cases with other radiologists multiple times.

We have also analyzed size distribution of the features found in the study, including the true positives and the false positives. Most of the features found in the pilot study are less than 10-mm as shown in figure 2.
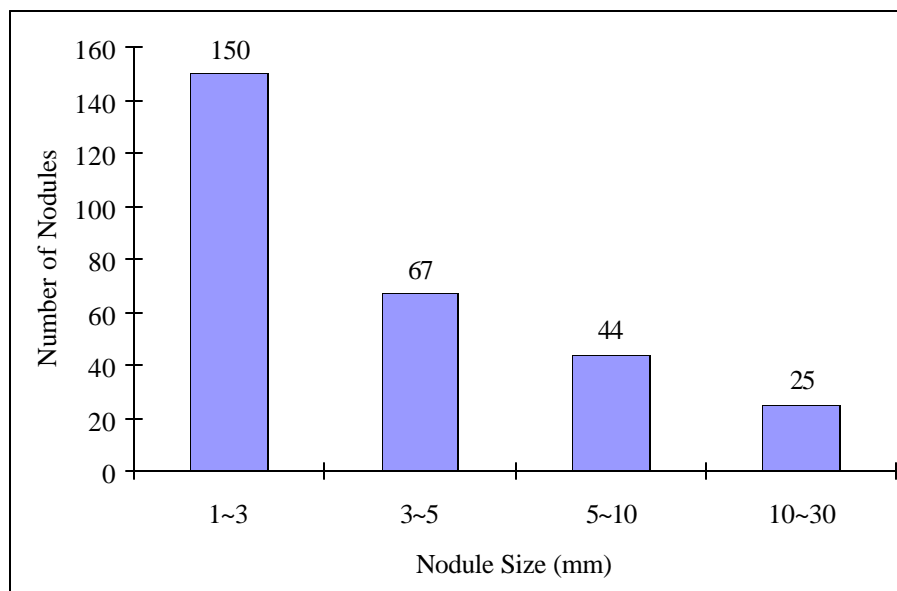

Figure 2. Distribution of nodule-like features by size.

Free-response Receiver Operating Characteristic (FROC) analysis suggests that the stereo display resulted the performance that was better than the orthogonal MIP display, but was equivalent to (or slightly better than) slice-based display, although no statistically significant difference was shown between the three display modes. The figure-of-merit (FOM) from the outputs of the JAFROC software (JAFROC, Chakraborty and Berbaum) for the three FROC curves were 0.57 (stereo display), 0.56 (slice-by-slice display) and 0.52 (orthogonal MIP display), respectively.

One of the efficiency measurements is interpretation time on each tested display mode. We have recorded interpretation time as well as navigation patterns from 4 participating radiologists randomly selected to anonymize attributes associated with each individual. By averaging the time over the 4 radiologists on each display mode, we have shown that the average interpretation time was significantly less with the stereo display (3.5 minutes per case) than with the slice-by-slice display (4.5 minutes per case). There is, however, not much difference in interpretation time between in the stereo display and the orthogonal MIP display (3.7 minutes per case).

Even though the stereo display resulted generally less interpretation time and less falsely claimed nodules among the three tested display modes, the overall performance from the stereo display did not surpass the one from the slice-by-slice display. Subjective opinions and objective observations suggest that training effects significantly influence radiologists' search behavior and interpretation results. Of the three display modes tested in the pilot study, the stereo display has never been used or tried by the participating radiologists and the orthogonal MIP display has been experienced to a very limited extent. We observed from the navigation patterns recorded from 4 participating radiologists that, at the beginning of the study, radiologists were vigorously tuning their search patterns to try to find the optimal search pattern and optimal viewing volume with the two 3D displays. Towards the end of the study, the navigation patterns in either the orthogonal MIP display or the stereo display were getting much easier and smoother, and stayed in a more stable and controllable manner (see figure 1, 2 in Appendix B). In contrast, the navigation patterns in the slice-by-slice display were more randomized and undifferentiated between the beginning and the end of the study (see figure

3 in appendix II). Further, we observed that when interpreting cases with the 3D displays, radiologists tended to adjust the viewing volume from initial thick slab to single slice during early stage of the study. As single slice was the subset of the viewing volume in these volumetric display modes, the preference for the single slice suggested strong influence of training effect to radiologists' interpretation behavior.

Further analysis from the search patterns revealed that some of the missed nodules were actually received extra attention from the radiologists despite of no report being filed. There were about 25% of missed detections that received extra attention in the slice-based mode, 15% in the orthogonal MIP mode and 16% in the stereo mode. Radiologists have been trained conventionally in single projected radiographic image interpretation and maintained consistent practice manner for scrutinizing this kind of images carefully. But they are not extensively exposed to 3D display for volumetric data and, meantime, lacking of systematical training for volumetric data interpretation. Furthermore, since volumetric display can show more information in one view and clear geometrical relationship for easy understanding compared to single slice based display, radiologists may be over-confident for their observation and tend to neglect some subtle structures needed for more attention and/or different skills in 3D view. Appropriate training and practice, therefore, is necessary for achieving optimal performance with 3D display device and new display technology.

While novelty seemed to substantially affect navigation patterns and the performance, other factors associated with our 3D displays may also influence the results. Despite similarity in the navigation patterns and in the use of thickness information, the orthogonal MIP rendered display and the stereo display showed some differences in nodule detection. Vessel-like structures were much easier to be mistakenly recognized as nodules in the orthogonal MIP display as compared to that in the stereo display. Overall, the orthogonal MIP resulted more false positive findings than the stereo display (table 3) and the lowest performance score among the three display modes, although with no statistically significant difference. The low performance and high false positive rate of orthogonal MIP are most likely attributed to superimposed structures of monoscopic thick slab. Despite high contrast volumetric images, orthogonal MIP rendering may not produce correct geometric representation of volumetric objects due to that the algorithm takes the highest intensity along each projection ray, which may very well not preserve structural continuity between adjacent pixels in the rendered image. The stereoscopic rendering, on the other hand, was implemented with perspective transformation and transparency mechanism so that no superimposition was introduced and local geometric information was better preserved, especially with averaging method.

Table 3. Distribution of false positive findings in different structural groups.

	Stereo	Orthogonal MIP	Slice by slice	Total
Vessel	11	27	18	56 (32%)
Scar	23	32	33	88 (51%)
Other	10	7	12	29 (17%)
Total	44 (26%)	66 (38%)	63 (36%)	173 (100)

Subjective evaluation of the study program and data analysis indicate that the study design, including the number of readers, sample size, the scope of data collection and the study software, was satisfactory and did not need any major changes when applied to the main study. One minor adjustment for study software was to record time point when a nodule-like feature is detected and characterized, so that we can locate the detection activity during interpretation course and study search patterns for improving display design in the future.

Perform Main Study (Task 11)
The main study was organized as a retrospective study of 576 nodules in 100 cases. Eight radiologists with various specialties have participated the study. Among the 8 radiologists, 4 of them were in the pilot study. Those who have not been in the pilot study, had a short training for the study program before the actual study. The study is in the progress and will be finished before the mid next year.

KEY RESEARCH ACCOMPLISHMENTS:
- Implemented rendering and display software on programmable graphics card that has achieved real-time volume rendering/displaying and user manipulation/interaction.
- Analyzed the data collected from the pilot study and gained some insight of interpretation behavior on volumetric displays that can be valuable for future guide of more efficient medical image rendering and display.
- Started a main study.

REPORTABLE OUTCOMES

Peer reviewed paper
Real-Time Stereographic Rendering and Display of Medical Images With Programmable GPUs. Submitted to Investigative Radiology.

Characteristics of CT Image Interpretation for Lung Nodule Detection Between Slice-Based and Volumetric Displays. Submitted to AJR, American Journal of Roentgenology.

Abstract
Stereo Display of CT Images for Lung Cancer Screening: A Pilot Study. SPIE, Medical Imaging, 2007 (accepted)

Application of the 3D Tensor Voting Paradigm for Segmenting Cylindrical Segments and Bifurcations from Volumetric Datasets. SPIE, Medical Imaging, 2007 (accepted)

Presentation
Stereoscopic Display Workstation for Visualization of Medical 3D Dataset. Lecture Series at the Department of Biomedical Informatics, University of Pittsburgh, Sept 29, 2006

Grant application
Immersive "Wall-of-Images" Display for Radiology - Preliminary Assessment. Submitted to NIH, September 2006.

Real-Time Interactive Stereo Display of Breast Tomosynthesis. Submitted to Susan G. Komen Breast Cancer Foundation, October 2006.

Immersive Stereographic Display for Real-Time Navigation through 3D Datasets. Submitted to NIH, October 2006.

CONCLUSION
Our primary objective is to determine whether a stereoscopic display concept has potential for improving the efficiency and accuracy of chest CT interpretation for lung cancer screening. In this report period, our main tasks were to implement secondary features involving volume rotation and real-time projection, analyze the data collected from the pilot study, and conduct a main study extended from the pilot study.

During this report period, we have implemented rendering and displaying on a programmable graphics card to achieve real-time user controlled image presentation (including volume rotation) and real-time user interaction. In the meantime, we have started the main observer performance study for further evaluating clinical value of stereo display workstation. The most important results, however, came from the analysis of the data from the pilot study. The preliminary data showed that the stereo display overall resulted better and more efficient performance for lung nodule detection and classification with less interpretation time compared with the other display modes tested in the pilot study, and also showed the factors, such as novelty and training effect, possibly affecting efficiency of using the 3D displays and other new technology related to medical imaging applications.

Our preliminary results strongly suggest that systematic training and practice is necessary for achieving optimal performance with 3D display device and new display technology in medical image diagnosis. The reader interpretation patterns revealed in the study as well as possible improvement of display software design based on the observations from the study can be generally applied for improving performance in medical image interpretation.

REFERENCES
1. Greenlee RT, Murray T, Bolder S, et al. Cancer statistics, 2000. CA Cancer J Clin 2000; 50:7-33.
2. Ries LAC, Elsner MP, Kosary CL, et al. SEER: cancer statistics review, 1973-1997. Bethesda, MD: National Cancer Institute, 2000.
3. van Klaveren RJ, Habbema JDF, Pedersen JH, et al. Lung cancer screening by low-dose spiral computed tomography. Eur Respir J 2001; 18:857-866.
4. Flehinger BJ, Kimmel M, Melamed M. The effect of surgical treatment on survival from early lung cancer. Chest 1992; 101:1013-1018.
5. Sobue T, Suzuki T, Matsuda M, et al. Survival for clinical stage I lung cancer not surgically treated. Comparison between screen-detected and symptom-detected cases. Cancer 1992; 69:685-692.
6. Miettinen OS. Screening for lung cancer. Radiol Clin North Am 2000; 38:479-486.

APPENDICES
Appendix A
Peer reviewed paper: **Real-time stereographic rendering and display of medical images with programmable GPUs**.

Appendix B
Peer reviewed paper: **Characteristics of CT image interpretation for lung nodule detection between slice-based and volumetric displays**.

Appendix C
Abstract: **Stereo display of CT images for lung cancer screening: a pilot study.**

Appendix D
Abstract: **Application of the 3D Tensor Voting Paradigm for Segmenting Cylindrical Segments and Bifurcations from Volumetric Datasets.**

SUPPORTING DATA
See APPENDICES, and the tables and the figures in the BODY section.

Appendix A

Real time stereographic rendering and display of medical images with programmable GPUs

Xiao Hui Wang and Walter F. Good

Department of Radiology, University of Pittsburgh, School of Medicine, Pittsburgh, PA

Abstract

The amount of volumetric data being acquired in radiology is rapidly increasing. To maintain performance and efficiency in reading this data, it is desirable to be able to display the data as 3-D monoscopic- or stereoscopic-renderings, with real-time interactive control by radiologists. This paradigm has not been widely adopted because of the difficulty and expense of providing the required computational resources. With the availability of newer commodity graphics processing units (GPUs) for personal computers, it may be possible to overcome the computational impediments to interactive 3-D displays. This study compared the frame rates that can be achieved on CPUs to those that can be achieved by exploiting GPUs, and finds that GPUs are capable of rendering large 3-D datasets at real-time interactive rates.

Introduction

Inherently 3-D medical imaging modalities, such as Computerized Tomography (CT) and Magnetic Resonance (MR) imaging systems, are generating an ever increasing volume of image data that must be reviewed by radiologists. This trend will almost certainly continue into the future as radiologists, in an effort to increase spatial resolution, depict 3-D volumes by using thinner, but more numerous, slices.

Current Display Paradigms -- By far, the most common method used for viewing inherently 3-D data has been reading 2-D slices sequentially, from the 3-D dataset, in a slice-by-slice mode, a laborious and error prone process, or viewing the data as projections of thicker sections comprised of multiple adjacent slices.

It is known that the visibility of certain kinds of subtle features can be increased by presenting the data as a thicker 3-D volume [1,2], rendered with an appropriate projection algorithm. This is normally achieved by combining the thin slices directly acquired during volumetric imaging to form a thicker slab, and then projecting this slab onto a 2-D display, but increasing the thickness of projected volumes can cause ambiguities due to the superposition of tissues. Also, use of an averaging process to combine slices can reduce the contrast of features that are small relative to the thickness of the resulting slab. As slabs become thicker by adding more of the originally acquired thin slices, the contrast of smaller features, which often are visible on only one or two thin slices, may be reduced by averaging with the remaining thin slices [1]. As slices become thinner, the signal-to-noise ratio in individual slices is reduced making it more difficult to detect certain kinds of features, and at the same time, the number of slices that must be read increases. Furthermore, the process of reading individual slices sequentially forces viewers to reconstruct mentally the 3-dimensional structure, and does not permit the reader's visual system to take full advantage of correlations between adjacent slices to improve apparent signal-to-noise ratios.

Various methods for 3-D display of volumetric radiographic datasets have been devised to make the reading process more efficient, but they have not been widely adopted because of certain performance limitations. Specifically, the task of rendering 3-D datasets in a form that is suitable for radiological applications is computationally intensive and it has not been possible to perform these calculations sufficiently fast to be able to provide radiologists with real-time interactive displays, except on superpremium computers. There is a consensus that, without real-time interactivity, volumetric display (monoscopic or stereoscopic) is often not justified by the added complexity.

Potential Role of Stereographic Displays – Stereographic display of 3-D radiographic datasets, which takes full advantage of readers' binocular vision, may provide benefits beyond those attributed to monoscopic 3-D display [3]. Certain kinds of objects can be detected in a stereo 3-D display of data, which cannot be detected when the data is viewed in a slice-by-slice manner. Stereo projection can improve the visibility of objects by enhancing features that are correlated between slices, while reducing noise in a manner analogous to the signal-to-noise improvements obtained by averaging slices or MIPs – but stereo projection does not introduce tissue superposition ambiguities that would be caused by these methods [4]. Nevertheless, stereographic presentation has received even less attention than monoscopic 3-D because it further increases the computational burden.

Application of GPUs – With the evolution of commodity graphics processing units (GPUs) for accelerating games on personal computers, over the past couple of years, the amount of computing power that is available for rendering complex scenes has been rapidly increasing. GPUs may be capable of performing a wide range of reconstruction, volume reformatting and stereo projection in real-time under user control. In particular, the most recent GPUs are approaching a performance level where real-time interactivity with stereographic displays is feasible.

GPUs are organized as pipelined parallel processors. They differ from general purpose processors, that basically perform one instruction at a time and need to have the result returned immediately, in that they process parallel streams of independent data and can wait for an individual result as long as the entire dataset is processed quickly [5]. In this sense, they are ideal for tasks that are computationally intensive in volumetric rendering of 3-D datasets. Dietrich, et al, report that they were able to achieve real-time rendering of a 512 x 512 x 100 liver CT dataset on a 2 GHz Pentium 4, with a ATI 9800 GPU, though they were primarily concerned with only the volume clipping component of the rendering algorithm [6].

Several researchers, including our own, have shown the potential benefit of GPUs for efficient image manipulation and visualization within medical applications [7-12]. For example, Briggs, et al, have demonstrated a display for volumetric electrical impedance tomography [7]. While their datasets are smaller than many that occur in radiology, they were able to achieve real-time performance. A Doppler-ultrasound display was implemented by Heid, et al, by exploiting the performance of a GPU [8]. GPU-based programming has been implemented for interactive 4-D motion segmentation and volume rendering of cardiac data and has resulted efficient data processing and visualizing with high quality and at real-time speeds [9]. GPUs were also demonstrated to be efficient in generating high quality reconstructed radiographs from portal images and CT volume data for radiation therapy [10]. Sorensen, et al, have also applied the technique to surgical simulation of the liver to achieve a real-time performance, where surface rendering involving dynamic geometric transformations and texture manipulations were implemented on a GPU [11]. These specialized applications can often achieve significant levels of performance by optimizing their systems for the application, but these systems do not necessarily retain that performance when used in a different context.

We have tested the feasibility and efficacy of performing renderings on GPUs for stereo display of medical 3-D dataset. Previously, we prestaged and prerendered stereo pair renderings of lung CT images for display. Because of different viewing positions and viewing volumes, rendering a complete set of image pairs for a case took a substantial amount of time and consumed vast storage space. Such a practice may work within certain research environments, but is not practical for the general clinical settings, where real-time rendering and manipulation are necessary for prompt and accurate diagnosis.

While GPUs have been applied to a number of radiological imaging tasks, their potential performance characteristics are not well understood. This study is an attempt to measure frame rates that can be achieved for stereographic rendering on a GPU and compares these to rates that can be achieved on CPUs alone.

Methods

Data set – Images used for developing GPU-based rendering and display were obtained from a 4-detector CT scanner (LightSpeed Plus, GE medical Systems, Milwaukee, WI) for lung cancer screening program. The CT

images were acquired in the axial plane and reconstructed to a thickness of 2.5 mm/slice with lung kernel reconstruction algorithm provided by GE standard software. The pixel size on each slice ranges from 0.63-mm × 0.63-mm to 0.92-mm × 0.92-mm. There are approximately 512×512×100 data voxels for a typical lung CT case in our dataset.

Hardware -- The study was run on an off-the-shelf personal computer with a 2.0 GHz AMD Athlon 64 3200+ processor and 512 MB of RAM. The computer is equipped with a 128 MB NVIDIA Quadro FX 1100 graphics card, which has build-in support for stereographic buffering system to hold left- and right-eye images in separate frame buffers and to swap frame buffers for a frame-swapped display. The stereo image pairs are viewed either on CRT monitors via shutterglasses controlled by frame-swapping signals or on superimposed cross-polarized displays via passive polarizing eyeglasses.

Volume rendering for stereo display – Two rendering methods, Maximum Intensity Projection (MIP) and averaging, have been implemented to generate stereo pairs of the lung CT images. Because lesions must be detected before they can be evaluated, high contrast MIP images were preferable for lesion detection while images rendered by averaging, which preserves local geometry, were preferable for lesion evaluation [12,13]. The rendering process for both MIP and averaging in this application involves perspective transformation [14], transparency modeling based on optical occlusion/distance characteristics, and ray casting [12-13,15-18]. All the rendering processes that we have previously performed on CPU card can be now processed on a programmable GPU card.

Rendering on GPUs – Stereographic compositing and display was implemented and compiled in the OpenGL and Cg languages on NVIDIA programmable GPUs. A flowchart, shown in Figure 1, illustrates the operations performed on GPU card.

For a given slab thickness, a vertex block with dimensions of 512×512×thickness was generated to include all vertices for perspective transformation and texture-coordinates. The dimensions of each vertex were approximated so as to be isotropic in all three axes (x, y and z) based on acquired x and y dimensions. Vertex-coordinates and texture-coordinates were then specified and interpolated during rasterization before being input to the vertex and fragment programs.

A sufficient number of interpolated slices were generated to provide continuity of display in the axial direction. Typically, for a dataset such as the one employed in this project, 3 interpolated slices are generated for every real slice.

Perspective projection in ray casting was performed in vertex program for each input vertex. The matrices for perspective transformation were determined by a presetting of eye-offsets and viewing distance. In the case of stereo compositing, the projection centers for the left- and right-eye images are offset laterally relative to each other. The parallax value for each eye-offset is set close to 1° to achieve stereo depth perception while avoiding excessive eyestrain. The rotation transform was also performed in vertex program. Transformed vertices that were out of the clip volume were not used for display. An example of Cg vertex program for vertex transformations is shown in Code 1.

Code 1.

```
vertOutput main (                    float4 Position : POSITION,
                        float4 Texcoord : TEXCOORD0,
                        uniform float4x4 rotate_x,
                        uniform float4x4 rotate_y,
                        uniform float4x4 translate_matrix,
                        uniform float4x4 perspective_matrix
            )
{
      vertOutput OUT;
```

```
        float4 rot_xP, rot_yP, tPosition, pPosition;
        rot_xP=mul(rotate_x,Position);
        rot_yP=mul(rotate_y,rot_xP);
        tPosition=mul(translate_matrix,rot_yP);
        pPosition=mul(perspective_matrix, tPosition);
        OUT.Position=pPosition;
        OUT.texcoord=Texcoord;
        return OUT;
}
```

Once a vertex has been geometrically transformed to a proper position, texture mapping for the vertex takes place in a fragment program. The 16-bit lung CT volume data (approximately 512×512×100) was loaded into the graphics memory to serve as a 3-D texture map. Texture values were automatically interpolated in the texture map with the OpenGL linear filter function for a given texture-coordinate. Occlusion/distance based transparency and window-level settings were also implemented in the fragment program. A Cg code fragment implementation is shown in Code 2.

Code 2.
```
float4 main (          vertOutput IN,
                       uniform sampler3-D testTexture,
                       uniform float window_level,
                       uniform float transparency_coef;
        ) : COLOR
{
        float4 color;
        float temp, tt;

        temp=tex3-D(testTexture, IN.texcoord);
        tt=(temp-window_level) * transparency_coef;
        color=tt;

        return color;
}
```

The final rendering process for displayed pixels was actualized by implementing the openGL blending functions. For MIP rendering the display value of each pixel was rendered by taking the maximum value among the points of a projection ray using MAX blending function, while for averaging rendering the display value was rendered by adding distance-weighted fractions of each fragment along a projection ray to the pixel using the ADD blending function.

Rendering on CPU card – The stereo image pairs for a lung CT case were prestaged and precalculated for all volume sizes between 1 up to 45 interpolated slices (see column 1 in table 1 and table 2) at all axial viewing positions. The detailed methods can be found in ref. In brief, we used trilinear interpolation to resample the data for a given volume of CT images to achieve final pixel dimension close to isotropic. Perspective transformation and ray casting based on compositing methods were performed for each pixel on stereo images. For MIP rendering the highest voxel value along a projection ray was used for projection value, while for averaging rendering each voxel value along a projection ray contributed a fraction to the final projection value.

Other functionality – An OpenGL based window display was built for displaying both GPU- and CPU-based stereo images. Specifically, window-level adjustment, viewing volume and viewing position selection and choice between MIP rendering and averaging rendering, were implemented. Image rotations were only performed by GPU-based rendering.

Results

The GPU-based program achieved real-time rendering and real-time display rates without any perceptible delay in the display of successive frames, following a user controlled frame switch command. We found no difference in frame rates between renderings by MIP and by averaging. A comparison of the rendering rates between GPU- and CPU-rendering, for our lung CT dataset, is shown in Tables 1 and 2. Table 1 lists the frame rate measurements of stereo compositing on GPU card as well as on CPU card at various volume sizes. The highest volume we rendered for lung CT images is 45 interpolated slices, which is about the thickness of 15 real CT slices at 2.5-mm. When we reviewed various stereo images with several experienced radiologists, we found that the preferred viewing volume for detection and diagnosis ranged from 3 to 7 real slices (i.e., 9 to 21 interpolated slices), and 15 slices (i.e., 45 interpolated slices) contained too much information to be useful for detection and diagnosis. Even with volume of 15 slices, which has more than 23 million vertex rendering processings ($512\times512\times45\times2$ stereo images), we still achieved a rate at 5-frames per second. Rendering performed on the CPU card resulted in much slower frame rates and would not give the impression of real-time interactivity. If we precalculate all of these stereo image pairs for a case, it would take less than a minute on the GPU card versus more then 20 minutes on the CPU. Implementing rotation on the GPU card did not measurably reduce frame rates for the data volumes used in this study, as shown in Table 2.

Discussion

Traditionally, 3-D medical image datasets are rendered predominantly on CPUs to generate precalculated images that can be prestaged for reading by radiologists. This preprocessing procedure puts many constraints on the review process and, at the same time, consumes a substantial amount of storage space and CPU time. These CPU-based processes most likely will be replaced in the near future by processes performed on the advanced graphics cards, due to the fact that these cards are becoming readily available and their real-time processing speed and improved arithmetic precision is makes them suitable for the processing of many types of radiological images. The study presented in this paper shows that GPU-based rendering can achieve real-time interactive stereo display rates for lung CT images up to volumes larger than the optimal volume used for diagnosis. Monoscopic rendering rates, though not measured in this study, would likely be nearly double the stereoscopic rates for a given volume.

The benefit of using GPUs processing power can be widely appreciated in medical image detection and diagnosis. As show in Table 1, GPU-based programming renders stereo pairs in real-time for as many as 45 slices (more than 23 million vertices) and gives no perceptional delay between frame changes. The capability of real-time process eliminates the constraints from prestaged paradigms. Viewing angles, for example, can be important for detection and differentiation of an object. It is, however, impractical and impossible to prestage and precalculate all viewing angles for a set of images, or to perform smooth rotations. Whereas programmable GPUs perform real-time renderings, rotation functionality can be seamlessly and smoothly implemented during rendering process and consumes negligible GPU processing time compared to the overall processing time, as shown in Table 2.

From research conducted by others and our previous studies, we have observed that no single algorithm can meet all the requirements of clinical tasks. We have demonstrated that for stereo display, MIP rendering is the best for detection owing to the high contrast of rendered images, but not optimal for classification because of lack of local geometric fidelity in the rendered images. On the other hand, rendering by averaging will preserves local geometry despite providing low contrast of the rendered images. The two renderings can be used for different tasks during medical image interpretations. We have implemented this mechanism in CPU-based prestaged calculations and display, and the results were satisfactory at the expense of longer processing time and much more storage space. The GPU-based programming not only naturally solved this problem of dynamically switching between MIP and averaging renderings, but can also, in general, implement any algorithms, whichever needed, specific to the task in real time. This will dramatically improve efficacy of image presentation and diagnostic performance.

Acknowledgement

References

1   Brown DG, Riederer SJ. Contrast-to-noise ratios in maximum intensity projection images. Magn Reson Med 1992; 23:130-137.

2   Keller PJ, Drayer BP, Fram EK, Williams KD, Dumoulin CL, Souza SP. MR angiography with two-dimensional acquisition and three-dimensional display. Work in progress. Radiology 1989; 173(2):527-32.

3   Roberts JW, Slattery OT. Display characteristics and the impact on usability for stereo. *Proc. SPIE, Stereoscopic Displays and Virtual Reality Systems VIII* 2000, 3957:128-138.

4   Maidment ADA, Bakic PR, Albert M. Effects of quantum noise and binocular summation on dose requirements in stereomammogry. Med Phys 2003; 30(12):3061-3071.

5   Ujval J. Kapasi UJ, Rixner S, Dally WJ, Khailany B, Ahn JH, Mattson P, Owens JD. Programmable Stream Processors. Computer. August 2003; (Vol. 36, No. 8) 54-62.

6   Dietrich CA, Nedel LP, Olabarriaga SD, Comba JLD, Zanchet DJ, da Silva AMM, Montero EF. Real-time interactive visualization and manipulation of the volumetric data using GPU-based methods. SPIE Medical Imag 2004; 5367:181-191.

7   Briggs NM, Avis NJ, Kleinermann F. A real-time volumetric visualization system for electrical impedance tomography. Physiol. Meas. 21(2000)27-33.

8   Heid V, Evers H, Henn C, Glombitza G, Meinzer HP. Interactive real-time Doppler-ultrasound visualization of the heart. Stud Health Technol Inform. 2000;70:119-25.

9   Levin D, Aladl U, Germano G, Slomka P Techniques for efficient, real-time, 3-D visualization of multi-modality cardiac data using consumer graphics hardware. 2005, Computerized Medical Imaging and Graphics 29: 463-475.

10 Khamene A, Bloch P, Wein W, Svatos M, Sauer F Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy. 2006, Medical Image Analysis 10: 96-112.

11 Sørensen TS, Mosegaard J Haptic feedback for the GPU-based surgical simulator. 2006, Medicine Meets Virtual Reality 14:523-528.

12 Wang XH, Good WF, Fuhrman CR, Sumkin JH, Britton CA, Warfel TE, Gur D. Stereo Display for Chest CT. Proc SPIE EI2004 2004; In Press.

13 Wang XH, Good WF, Fuhrman CR, Sumkin JH, Britton CA, Warfel TE, Gur D. Projection Models for Stereo Display of Chest CT. Proc SPIE MI2004 2004; In Press.

14 Mortenson, ME. Computer Graphics Handbook: Geometry and Mathematics, 1990

15 Drebin RA, Carpenter L, Hanrahan P. Volume rendering. Comput Graph 1988; 22:65-74.

16 Levoy M. Display of surfaces from volume data. IEEE Comput Graph Applicat 1988; 8:29-37.

17 Ney DR, Drebin RA, Fishman EK, Magid D. Volumetric rendering of computed tomographic data: principles and techniques. IEEE Comput Graph Applicat 1990; 10:24-32.

18 Ney DR, Fishman EK, Magid D, Robertson DD, Kawashima A. Three-dimensional volumetric display of CT data: effect of scan parameters upon image quality. J Comput Assist Tomogr 1991; 15:875-885.
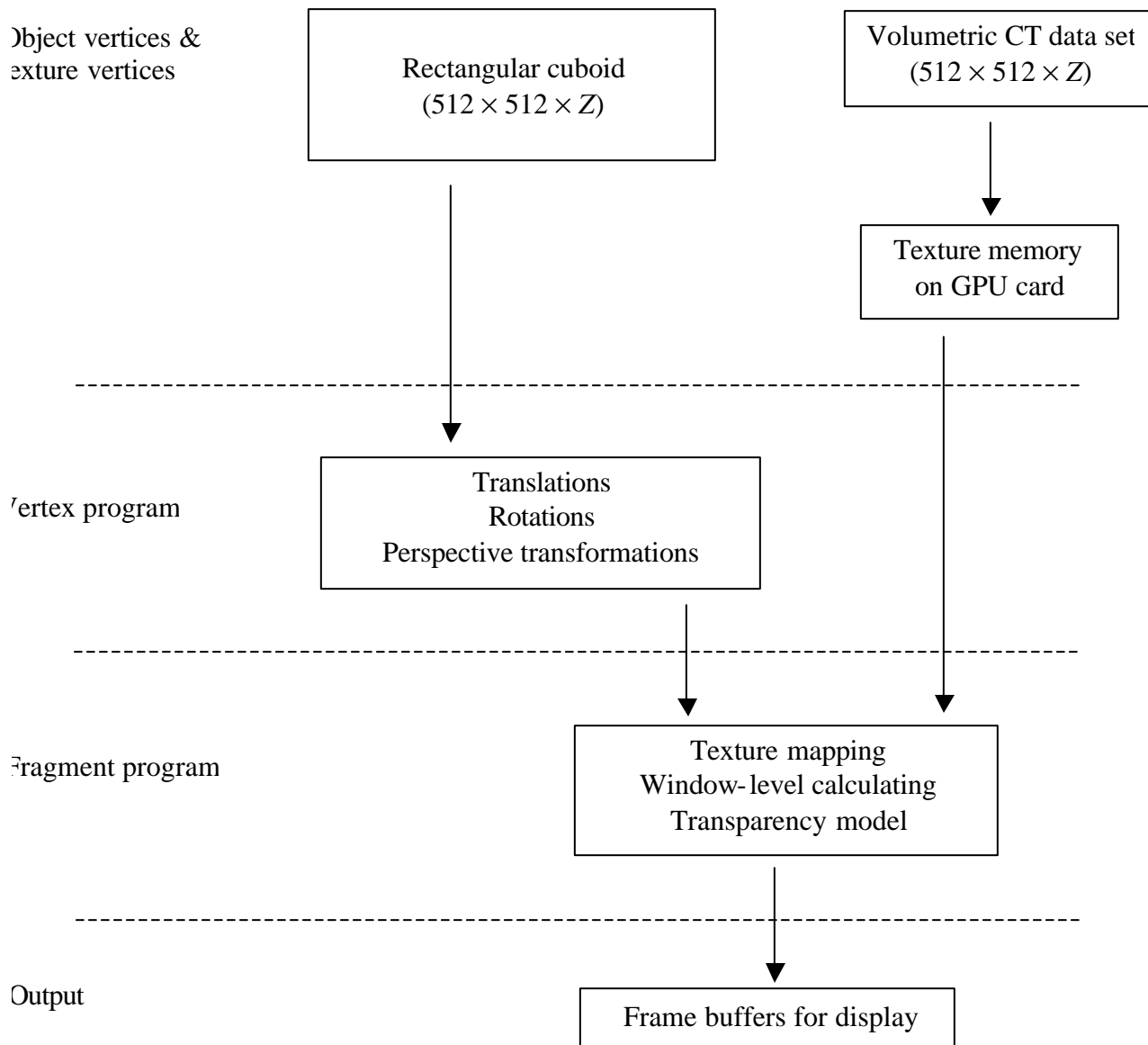
Object vertices &
texture vertices

| Rectangular cuboid $(512 \times 512 \times Z)$ | | Volumetric CT data set $(512 \times 512 \times Z)$ |

Texture memory
on GPU card

Vertex program

Translations
Rotations
Perspective transformations

Fragment program

Texture mapping
Window-level calculating
Transparency model

Output

Frame buffers for display

Figure 1. A diagram of stereo image rendering process on GPU card. *Z* is the depth measure of a given rendering volume.

Number of interpolated slices	GPU (stereo pairs per second)	CPU (stereo pairs per second)
1	103.3	-
9	20.1	1.3
15	13.2	0.8
21	10.1	0.5
33	6.6	0.3
45	5.0	0.2

Table 1. Frame rates measured as stereo pairs per second for rendering on GPU card and CPU card at different number of interpolated slices.

Number of interpolated slices	Rotation implemented (stereo pairs per second)	Without rotation (stereo pairs per second)
1	103.3	103.3
3	44.4	44.4
5	33.7	33.7
9	20.1	20.1
15	13.2	13.2
21	10.1	10.1
33	6.6	6.6
45	5.0	5.0

Table 2. Frame rates measured as stereo pairs per second for rendering on GPU card with and without rotation implementation.

Appendix B

Characteristics of CT Image Interpretation for Lung Nodule Detection Between Slice-Based and Volumetric Displays

Xiao Hui Wang, Walter F. Good, Janet E. Durick, Amy Lu, David L. Herbert, Saraswathi K. Golla, Kristin Foley, C, Samia Piracha, Dilip D.Shinde, Shindel , Carl R. Fuhrman, Cynthia A. Britton, Sherry S. Shang, Christoper Deible and Joan M. Lacomis

Department of Radiology, University of Pittsburgh, School of Medicine, Pittsburgh, PA

Abstract

OBJECTIVES. The purpose of this study was to investigate characteristics of radiologists' search patterns and search results in lung nodule detection on CT images with different rendering and display schemes for improving medical volumetric image visualization and diagnostic performance.
MATERIALS AND METHODS. Retrospective lung nodule detection with computerized tomographic images was conducted on three display modes, including slice-by-slice display, orthogonal maximum intensity projection display and stereoscopic display. Thirty lung-cancer-screening CT cases containing 91 nodules were used in the study, and eight radiologists interpreted the cases. Radiologists' search course within the volumetric data was recorded along with the probability of a nodule, location, size and shape for each detected feature. Characteristics of detected features and radiologists' search patterns were compared for the three display modes. The nodule detection performance was analyzed with Free-response Receiver Operating Characteristic method.
RESULTS. The stereo display provided better visual effect of 3D representation and produced better detection and classification performance with less interpretation time compared with other display modes tested in the study. However, the difference between the stereo display and the other displays was not statistically significant. Further analysis of the navigation patters showed that novelty and training effect were associated with the nodule detection performed on the volumetric displays. Among the tree display modes, the orthogonal maximum intensity projection display resulted the highest number of false positives, in which most were vessel structures. Scar tissue was the most common structure falsely recognized as lung nodule in all three display modes.
CONCLUSION. Our preliminary results indicate a potential role of stereo display for improving radiologists' performance in medical detection and diagnosis, and also strongly suggest that systematic training and practice is necessary for achieving optimal performance with volumetric displays or any new display technology in medical image diagnosis.

Keywords: volumetric dataset, navigation, stereoscopic display, lung nodule screening

Introduction

Medical image interpretation involves heavily human-image interaction. Extensive studies have been conducted to investigate eye search patterns on projected radiographic images for lesions [1, 2, 3, 4, 5, 6, 7]. The results indicate that the eye search characteristics are more on experience bases, are influenced by image quality, and can be correlated to the performance of detection and diagnosis [1, 3, 4, 6, 7, 8, 9]. These studies have helped to improve image quality, image representation and visual inspection technique. However, most of these works have been focused on 2-dimensional (2D) radiographic images and very little research by far has been done on human-computer interaction and searching behavior associated with 3-dimentional (3D) medical datasets.

Medical imaging is rapidly evolving into 3D representation [10, 11, 12]. In the near future, it is very likely that 3D datasets from various imaging modalities will dominate medical imaging format for diagnosis, treatment and image-guided surgery. Radiologists will have to adopt new search or navigation strategy to interpret image datasets. One big difference between 2D projected image and 3D image dataset is that resolution on each single 2D image in a 3D dataset is much lower than that on a single 2D projected image. Because of reduced resolution on each image and expanded information into one more dimension, radiologist needs to rely more on the information between images, which introduces information exploring in additional dimension and changes drastically the behavior of gazing and searching during image interpretation.

The features unique to 3D datasets are all likely to affect radiologists' interpreting behavior. For example, ever-increasing image volume forces radiologists to adopt computerized display (soft copy display) and to be involved more and more actively in computer-based procedures and operations for image interpretation. Furthermore, in order to optimally utilize information captured in 3D datasets, various computer algorithms have been developed to render 3D images into 2D images for display. Such practice has changed traditional radiographic presentation that radiologists have learnt and accustomed, and may require different approaches to perceive and interpret the images.

While radiologists are experiencing the transition from 2D projected radiography to 3D image datasets, image research has also to face the question of how this change would likely challenge previous observations and derive comprehensive conclusions from radiologists' practice with images from new imaging modalities. Despite the knowledge we have obtained from eye tracking system on 2D radiographic image, it is likely that 3D images interpretation is more relying on the combination of the characteristics of 2D and 3D images, which including human-computer interaction, 3D rendering presentation and navigation in third dimension. It is more important that we can understand the impact of new imaging modalities and image formats on radiologists' interpretation and therefore help radiologists to adapt and develop more efficient interpretation methods to improve their performance.

To the best of our knowledge there are very few, if any, researches in navigation and search patterns for medical 3D image dataset interpretation. Nevertheless, considerable efforts have been devoted to developing or designing 3D rendering and display methods to make image presentation more effective for medical detection diagnosis [13, 14]. Surface rendering, for example, is commonly used rendering method for displaying external structures and object shapes [15, 16]. Volumetric rendering methods, on the other hand, are more diagnostic relevant for revealing internal anatomical structures [17, 18]. One of the most commonly used volumetric rendering methods is maximum intensity projection (MIP), which maximizes contrast on a rendered image by taking brightest voxel on a projected voxel ray [19]. Researchers have also tried to combine different rendering algorithms into one application to manage different anatomical structures and different diagnostic purposes. Various display workstations and user interface have been developed in order to achieve better image perception, ease reader-computer interaction and improve the efficacy of interpretation process [20, 21, 22, 23, 24, 25]. Our stereo display tested in this study is one of those attempts to improve radiologists' performance in lesion detection and classification. [22, 23]

All the works on 3D image data manipulation and presentation have significantly facilitated medical 3D data visualization. It is, however, unclear how well radiologists would adapt to the new technology and, more importantly, what kind of features or functionalities will likely improve radiologists' performance and maximize the utility value of the new technology. To understand radiologists' search pattern during interpretation of 3D image datasets, we have collected navigation data from a pilot study designed for ROC (Receiver Operating Characteristic)-type analysis for lung nodule detection on CT images and characterized the patterns that are related to the nodule detection and classification. The search patterns obtained from different display modes provided useful information on 3D image interpretation and possible improvement of display design.

Materials and Methods

A pilot study of lung nodule detection and classification on CT images was used for studying interpretation and navigation patterns. The detection and classification task was performed on three display modes (conventional slice-by-slice display, orthogonal MIP display and stereoscopic display), and radiologists' search patterns were collected and the performance was compared between the three display modes.

Data specification

Low dose lung CT images for lung cancer screening were acquired from multislice CT scanner (LightSpeed, GE, Milwakee), at a reconstructed thickness of 2.5-mm per slice and pixel resolution ranged from 0.69×0.69 mm^2 to 0.94×0.94 mm^2. There are about 100 axial images for each case, and a total of 30 cases were randomly selected from the lung cancer screening cases.

We have recruited six experienced staffed radiologists and two fellow radiologists to interpret the images. The primary task of interpretation was to detect and then classify any nodules equal to or larger than 3 mm in diameter with three distinct computer display modes, which are described in followings.

Image rendering

The display modes used in this study included slice-by-slice, orthogonal MIP rendering and stereoscopic view. Raw CT images were first processed with the convolution kernel provided by GE standard reconstruction software to form reconstructed images that are optimal for viewing lung tissues. The reconstructed images were then rendered based on the specification of each display mode. All renderings were precalculated and stored on hard disk for real-time display.

Slice-by-slice -- This is the most common display method adopted by radiologists for CT image interpretation. As images are read one at a time in sequence, no further rendering process was applied after the raw images were reconstructed with the lung kernel filtration. This set of single images was also included as a subset in the next two display modes.

Orthogonal MIP -- A stack of various number of CT slices were used to form MIP images. In this study, we implemented MIP images at thickness of 3, 5, 7, 9, 13 and 15 CT slices, respectively. Thickness of single slice was included in this display mode. For a given number of CT slices (thickness), a serial MIP images were rendered along the axial direction.

Voxel resampling was performed at axial direction (z-direction) to approximate isotropic voxel before performing 3D rendering. For orthogonal MIP rendering, the maximum value on an orthogonally projected voxel ray was selected as the final display value for each pixel on the MIP image.

Stereo perspective projection -- Slab thickness selection and voxel resampling used in orthogonal MIP rendering were also applied for stereo rendering. Linear perspective projection was applied to a stack of images to form horizontally shifted transformations of left- and right-eye images. Interocular distance (6.5-cm) and viewing distance between a viewer and computer screen (45-cm) were used to determine the angles of both eyes for perspective transformations.

Two rendering methods were employed for the stereo images [22, 23]. One was distance-weighted MIP rendering to produce high contrast images for nodule detection; and the other was distance-weighted averaging rendering to produce images of highly preserved local geometry for nodule classification. Distance-weighted algorithm incorporated in the stereo renderings provided transparency mechanism to adjust light transmission according to voxel locations. The detailed methods were included in references 22 and 23.

Display interface

A desktop personal computer was used for three display modes to display lung CT images. The computer has a central processor of 2.0 GHz ADM Athlon 64 3200+, 512 MB RAM, and a 128 MB NVIDIA Quadro FX 1100 graphics card. During stereo display, stereo effect was achieved through a shutterglasses

(Stereo3D) controlled by frame-swap signals of displaying left-eye and right-eye images on the graphics card. A 21.0" (20.0" viewable) PerfectFlat CRT monitor, ViewSonic® Graphics Series G220f, was used in the display workstation. The monitor refresh rate was set to 144 Hz to produce stereo view without flickering effect.

A user interface was implemented using Microsoft Visual C++ API combined with OpenGL for image display and user interaction tools. Interactive operations during case interpretation basically involved navigation/search activity for lung nodules by moving along the axial direction throughout the lung area, and nodule assessment for any detected nodules. All the navigation/search related activities were conducted on a programmable keypad, which was dedicated to the specific needs for this study. The function keys on the programmable keypad can be used for selecting image axial viewing position and viewing volume (slab thickness), changing window/level settings, switching between MIP rendering and averaging rendering during stereo display, and toggling detected nodules.

An onscreen scoring form was designed and implemented for lung nodule classification. When radiologist clicks on a detected nodule, the scoring form with questionnaire related to the detected nodule would pop up for nodule assessment. We have also implemented mouse cursor as an onscreen ruler that can be used for nodule size estimation.

Study design
This study was designed for Free-response Receiver Operating Characteristic (FROC) type of detection. The task of the study was to detect any nodule-like feature and characterize it. Randomization of case order and display modes was applied to each interpretation session to avoid bias caused by predictability from case order or particular mode. To avoid bias of familiarity, there were at least 14 days apart between two studies of same case with different display modes.

Data collection
We have recorded navigation patterns from four participating radiologists randomly selected to anonymize attributes associated with each individual. Navigation pattern during interpretation was collected by recording viewing volume (slab thickness) and viewing position at a 250 millisecond interval. For a detected nodule, the position of x, y, and z dimensions were recoded along with the other parameters, such as the likelihood probability of a nodule, the likelihood probability of malignancy, nodule shape, calcification and nodule size.

Data analysis
Interpretation time for each case was computed and compared between the modes. The navigation and nodule detection patters were visually analyzed and compared between the modes. Viewing volumes were analyzed from slab thickness recorded during case interpretation.

The nodule detection performance was determined by FROC analysis using JAFROC software (JAFROC, Chakraborty and Berbaum, http://www.devchakraborty.com/). The Figures of Merit from FROC were presented on a per-nodule basis. To verify the nodules, we used consensus results as the truth profile. The nodule-like features pooled from eight radiologists' interpretation in the three display modes were reviewed and verified by an experienced chest radiologist, who did not participate the study but had read and discussed the cases with other radiologists multiple times.

Results

Performance of nodule detection
The performance was evaluated based on consensus results of nodule location and likelihood probability. Total of 174 nodule-like features at the size of equal to or larger than 3-mm in diameter have been found in the 30 cases and 91 of them are true nodules. FROC analysis suggests that the stereo display resulted the performance that was better than the orthogonal MIP display, but was equivalent to the slice-based display,

although no statistically significant difference was shown between the tree display modes. The Figures of Merits from the JAFROC software were, 0.57 (stereo display), 0.56 (slice-by-slice display) and 0.52 (orthogonal MIP display) for 8 radiologists, and 0.59 (stereo display), 0.61 (slice-by-slice) and 0.53 (orthogonal MIP display) for 4 radiologists whose navigation courses were recorded.

One of the efficiency measurements is interpretation time on each tested display mode. By averaging the time over 4 radiologists on each display mode, we have shown that the average interpretation time was significantly less with the stereo display (3.5 minutes) than with the slice-by-slice display (4.5 minutes), but was not much difference between the stereo display and the orthogonal MIP display (3.7 minutes).

Navigation pattern
The average viewing volume for the 3D displays was between 3 and 5 CT slices. There was no apparent difference in the preference of viewing volume between the stereo display and the orthogonal MIP display. When a region or a feature was in suspicious, a quick back-and-forth navigating across several slices was observed. This distinctive navigation pattern was more typically seen in the slice-by-slice display mode and in the nodules described as non-solid or semi-solid features. To interpret the case, the radiologists typically navigated through the dataset axially between the top and the base of the lung several times. The average number of such navigation rounds for the stereo display, the orthogonal MIP display and the slice-by-slice display were 3.4±4.3, 4.3±2.2 and 3.5±1.7, respectively.

The learning curve related to the 3D displays (the stereo display and the orthogonal MIP display) was demonstrated by the comparison of the navigation patterns at the beginning and the end of this study in figures 1, 2 and 3. The data was recorded from four radiologists' interpretations on each display mode. Comparing to the search course at the end of the study, the navigation patterns and viewing volume with the stereo (figure 3) and the orthogonal MIP (figure 2) displays were more complicated and dynamic at the initial stage of the study. Toward the end of the study, the navigation patterns became much smoother and more stabilized in both the stereo display and the orthogonal MIP display. The navigation patterns from the slice-by-slice display were, however, more like random search manner than a learning process when comparing the navigation patterens at the beginning and the end of the study (figure 1). Since case order was randomized at each interpretation session for each radiologist, the navigation patterns between radiologists shown in figures 1, 2 and 3 were not taken from the same cases.

Characteristic of missed nodules
We have compared missed nodules at apical lung area as well as the area close to diaphragm between the three display modes. Since a nodule could be detected 8 times (8 participating radiologists) in each display mode, it would be more appropriate to use number of detections, instead of number of nodules, for comparison. The total number of detections in the apical area and diaphragm area should be 128 and 112, respectively. In the apical area, there was a higher missed detection rate either in the stereo display (55%) or the orthogonal MIP display (55%) than that in the slice-by-slice display (42%). However, the difference was not such obvious in the lung area close to diaphragm, in which the missed detection rates were 36% for the stereo display, 38% for the orthogonal MIP display and 35% for the slice-by-slice display.

Further analysis from the search patterns revealed that some of the missed nodules were actually received extra attention from radiologists despite of no report being filed. Typical search pattern of the area, where a missed nodule resides and radiologist paid extra attention, are shown in figure 4A and 4B. There were about 25% of missed detections that received extra attention in the slice-by-slice mode, 15% in the orthogonal MIP mode and 16% in the stereo mode.

Structural characteristics of false detections
Most of falsely claimed nodules were various forms of scar tissues and vessels (table 1), in which scar tissues occurred more than vessels. Other structures that falsely recognized as nodules include bronchiectasis, atelectasis, and soft tissues. In the vessel group, more false detections were found in the orthogonal MIP display mode than in the slice based or the stereo display mode as shown in table 1.

Discussion

When medical imaging is rapidly evolving from 2D radiography to volumetric datasets, information presentation is also being changed. The main difference between 2D data and 3D data is that spacial information is not compressed in the 3rd dimension and therefore each image within a volume shares partial information that is much less than information in a projected radiographic image. To help radiologists more efficiently handle the volumetric data such as data from CT and MR, many programs were implemented to render and display the data to be visually comprehensible. The results and feedbacks seemed very positive regarding to the performance [13, 16, 26]. The actual clinical utility and impact, however, are not well documented and demonstrated. It is our main interest in this paper to present our preliminary results of interpretation patterns with volumetric datasets and to understand the impact of different display schemes on the search characteristics, for lung nodule detection and classification.

Even though the stereo display resulted generally less interpretation time and less falsely claimed nodules among the three tested display modes, the overall performance from the stereo display did not surpass the one from the slice-by-slice display. Subjective opinions and objective observations suggest that training effects significantly influence radiologists' search behavior and interpretation results. Of the three display modes tested in the pilot study, the stereo display has never been used or tried by the participating radiologists and the orthogonal MIP display has been experienced to a very limited extent. We observed from the navigation patterns that, at the beginning of the study, radiologists were vigorously tuning their search patterns to try to find the optimal search pattern and optimal viewing volume with the two 3D displays, suggesting active involvement of learning and improving, which could including familiarization to the 3D displays and optimization of search strategy. In contrast, the navigation patterns in the slice-by-slice display were more randomized and undifferentiated between the beginning and the end of the study (figure 1). Further, we observed that when interpreting cases with the 3D displays, radiologists tended to adjust the viewing volume from initial thick slab to single slice during early stage of the study. As single slice was the subset of the viewing volume in these volumetric display modes, the preference for the single slice suggested strong influence of training effect to radiologists' interpretation behavior.

Evidence of training effect further comes from the observation that there were more attention-paid missed nodules associated with the slice-by-slice display than either with the orthogonal MIP or with the stereo display. Radiologists have been trained conventionally in single projected radiographic image interpretation and maintained consistent practice manner for scrutinizing this kind of images carefully. But they are not extensively exposed to volumetric display, and meantime lacking of systematical training for volumetric data interpretation. Furthermore, since volumetric display can show more information in one view and clear geometrical relationship for easy understanding compared to single slice based display, radiologists may be over-confident for their observation and tend to neglect some subtle structures needed for more attention and/or different skills in 3D view. Appropriate training and practice, therefore, is necessary for achieving optimal performance with 3D display device and new display technology.

While novelty seemed to substantially affect navigation patterns and the performance, other factors associated with our 3D displays may also influence the results. Despite similarity in the navigation patterns and in the use of thickness information, the orthogonal MIP rendering and the stereo view showed some differences in nodule detection. Vessel-like structures were much easier to be mistakenly recognized as nodules in the orthogonal MIP display as compared to that in the stereo display. Overall, the orthogonal MIP resulted more false positive findings than stereo display (table 1) and the lowest performance score among the three display modes, although with no statistically significant difference. The low performance and high false positive rate of the orthogonal MIP rendering are most likely attributed to superimposed structures of monoscopic thick slab. Despite high contrast volumetric images, orthogonal MIP rendering may not produce correct geometric representation of volumetric objects due to that the algorithm takes the highest intensity along each projection ray, which may very well not preserve structural continuity between adjacent pixels in the rendered image. The

stereoscopic rendering, on the other hand, was implemented with perspective transformation and transparency mechanism so that superimposition was not introduced and local geometric information was better preserved, especially with averaging method.

When lung nodules are neighbored with similar intensity of non-lung tissues in a thick viewing volume, they are likely to be missed due to camouflage effect. We have examined two places where lung tissue could be obscured by surrounding structures. One was apical lung area, where lung tissues are closely surrounded by rib cage. The other one was the area close to diaphragm. The results indicate that there were more missed detections with either the stereo or the orthogonal MIP display than with the slice-by-slice display in the apical area, while no such difference showed in the diaphragm area between the three displays. As obscuration can lower the conspicuity of the nodules, other factors, such as structure density and shape relationship, may also have effect on the detection as suggested by the different results from the two areas.

In this study, we have not implement multiple reformations for different viewing angles because of the complexity of preparing prestaged multiple reformations. Results from other researches and our current project of real-time rendering on programmable graphics units indicate that volumetric displays that allow multiple reformatted viewing angles by rotating images can help reduce ambiguity caused by some poorly differentiated spacial relationships including tissue superimposition [26, 27, 28]. The advantage of multiple reformations can be more appreciated by volumetric displays than single slice based display. Multiple views for single slice are geometrically discontinuously transformed because they lack the information of the third dimension and require intensive mental work on geometrical correlations between two viewing angles. When viewing volumetric data, volume can be smoothly transformed between two viewing angles by rotating objects in 3D space to produce natural continuation of views of the objects. The improvement of structure differentiation may be further enhanced by making non-interested tissue transparent to reduce the ambiguity.

Although more 3D imaging modalities are being employed for medical screening and diagnosis, slice-by-slice display is still predominantly being used as a primary viewing method for interpretation. Adopting volumetric displays, therefore, involves learning process that extents and transforms current 2D understanding of medical images to the knowledge of volumetric information discovery. Effective utilization of 3D display for medical volumetric data relies both on software design and user training. Our preliminary data from a pilot study for lung nodule detection on CT images indicate that current 3D displays can be further improved by understanding radiologists' interpretation behavior and diagnostic performance. In addition, stereoscopic display produced more efficient interpretations and lower false position detections comparing to other displays, and has promising potential for improving radiologists' performance and efficiency of 3D dataset interpretation.

References

1. Krupinski EA. Visual search of mammographic images: influence of lesion subtlety. *Acad Radiol*. 2005; 12:965-969

2. Krupinski EA, Berger WG, Dallas WJ, et al. Searching for nodules: what features attract attention and influence detection? *Acad Radiol.* 2003; 10:861-868

3. Nodine CF, Mello-Thoms C, Kundel HL, at el. Time course of perception and decision making during mammographic interpretation. *AJR Am J Roentgenol*. 2002; 179:917-923

4. Kundel HL, Nodine CF, Toto L. Searching for lung nodules. The guidance of visual scanning. *Invest Radiol*. 1991; 26:777-781

5. Kundel HL. Visual cues in the interpretation of medical images. *J Clin Neurophysiol*. 1990; 7:472-483

6. Kundel HL, Nodine CF, Krupinski EA. Searching for lung nodules. Visual dwell indicates locations of false-positive and false-negative decisions. *Invest Radiol*. 1989; 24:472-478

7. Kundel HL, Nodine CF, Thickman D, et al. Searching for lung nodules. A comparison of human performance with random and systematic scanning models. *Invest Radiol*. 1987; 22:417-422

8. Krupinski EA, Roehrig H. The influence of a perceptually linearized display on observer performance and visual search. *Acad Radiol.* 2000; 7:8-13

9. Nodine CF, Kundel HL, Lauver SC, et al. Nature of expertise in searching mammograms for breast masses. *Acad Radiol.* 1996; 3:1000-1006

10. Barish MA, Rocha TC. Multislice CT colonography: current status and limitations. *Radiol Clin North Am*. 2005; 43:1049-1062

11. Rowe VL, Tucker SW Jr. Advances in vascular imaging. *Surg Clin North Am*. 2004; 84:1189-1202

12. Israel GM, Bosniak MA. Renal imaging for diagnosis and staging of renal cell carcinoma. *Urol Clin North Am*. 2003; 30:499-514

13. Aufort S, Charra L, Lesnik A, et al. Multidetector CT of bowel obstruction: value of post-processing. *Eur Radiol.* 2005; 15:2323-2329

14. Remy J, Remy-Jardin M, Artaud D, et al. Multiplanar and three-dimensional reconstruction techniques in CT: impact on chest diseases. *Eur Radiol.* 1998; 8:335-351

15. Bomans M, Hohne KH, Tiede U, et al. 3-D segmentation of MR images of the head for 3-D display. *IEEE Transactions on Medical Imaging*. 1990; 2:177-183

16. Hoehne TU, Bomans KH, Pommert M, et al. Investigation of medical 3D-rendering algorithms. *IEEE Comp Graphics & Applications*. 1990; 2:41-53

17. Calhoun PS, Kuszyk BS, Heath DG, et al. Three-dimensional Volume Rendering of Spiral CT Data: Theory and Method. *Radiographics*. 1999; 19:745-764

18. Kuszyk BS, Heath DG, Ney DR, et al. CT Angiography with Volume Rendering: Imaging Findings. *AJR, American Journal of Roentgenology*, 1995; 165:445-448

19. Fishman EK, Ney DR, Heath DG, et al. Volume Rendering versus Maximum Intensity Projection in CT Angiography: What Works Best, When, and Why. *RadioGraphics* 2006; 26:905-922

20. Demiralp C, Jackson CD, Karelitz DB, Zhang S, Laidlaw DH. CAVE and fishtank virtual-reality displays: a qualitative and quantitative comparison. IEEE Trans Vis Comput Graph. 2006 May-Jun;12(3):323-30.

21. Levin D, Aladl U, Germano G, et al. Techniques for efficient, real-time, 3D visualization of multi-modality cardiac data using consumer graphics hardware. *Comput Med Imaging Graph*. 2005; 29:463-475

22. Wang XH, Maitz GS, Leader JK, et al. Real-time stereographic display of volumetric datasets in radiology. 2006, SPIE, Electronic Imaging vol 6055, 1A-1 - 1A-6.

23. Wang XH, Walter FG, Fuhrman CR, et al. Stereo CT Image Compositing Methods for Lung Nodule Detection and Characterization. Academic Radiology 2005, 12:1512-1520.

24. Yee DK, Lee W, Kim D, et al. RadGSP: a medical image display and user interface for UWGSP3. *Proc. SPIE* 1991; 1444:292-305

25. Fuchs H, Levoy M, Pizer SM. Interactive visualization of 3D medical data. *Computer*. 1989; 22:46-51

26. Salvolini L, Bichi Secchi E, Costarelli L, et al. Clinical applications of 2D and 3D CT imaging of the airways--a review. *Eur J Radiol.* 2000; 34:9-25

27. Ou P, Celermajer DS, Calcagni G, et al. Three-Dimensional CT Scanning: A Novel Diagnostic Modality in Congenital Heart Disease. *Heart*. 2006; Epub ahead of print:
http://heart.bmjjournals.com/cgi/content/abstract/hrt.2006.101352v1
28. Durkee NJ, Jacobson J, Jamadar D, et al. Classification of common acetabular fractures: radiographic and CT appearances. *AJR Am J Roentgenol*. 2006; 187:915-925

Table 1. Distribution of false positive findings in different structural groups.

	Stereo	Orthogonal MIP	Slice by slice	Total
Vessel	11	27	18	56 (32%)
Scar	23	32	33	88 (51%)
Other	10	7	12	29 (17%)
Total	44 (26%)	66 (38%)	63 (36%)	173 (100)

Figure 1. Navigation patterns from four radiologists in **stereo mode**. Each graph is the navigation recorded from one case interpretation. Two graphs in each row are taken from one radiologist's interpretations in the beginning of the study (left) and the end of the study (right). The dark solid line represents viewing positions referenced to the scale on left axis and grey solid line represents viewing thickness referenced to the scale on right axis, with time.
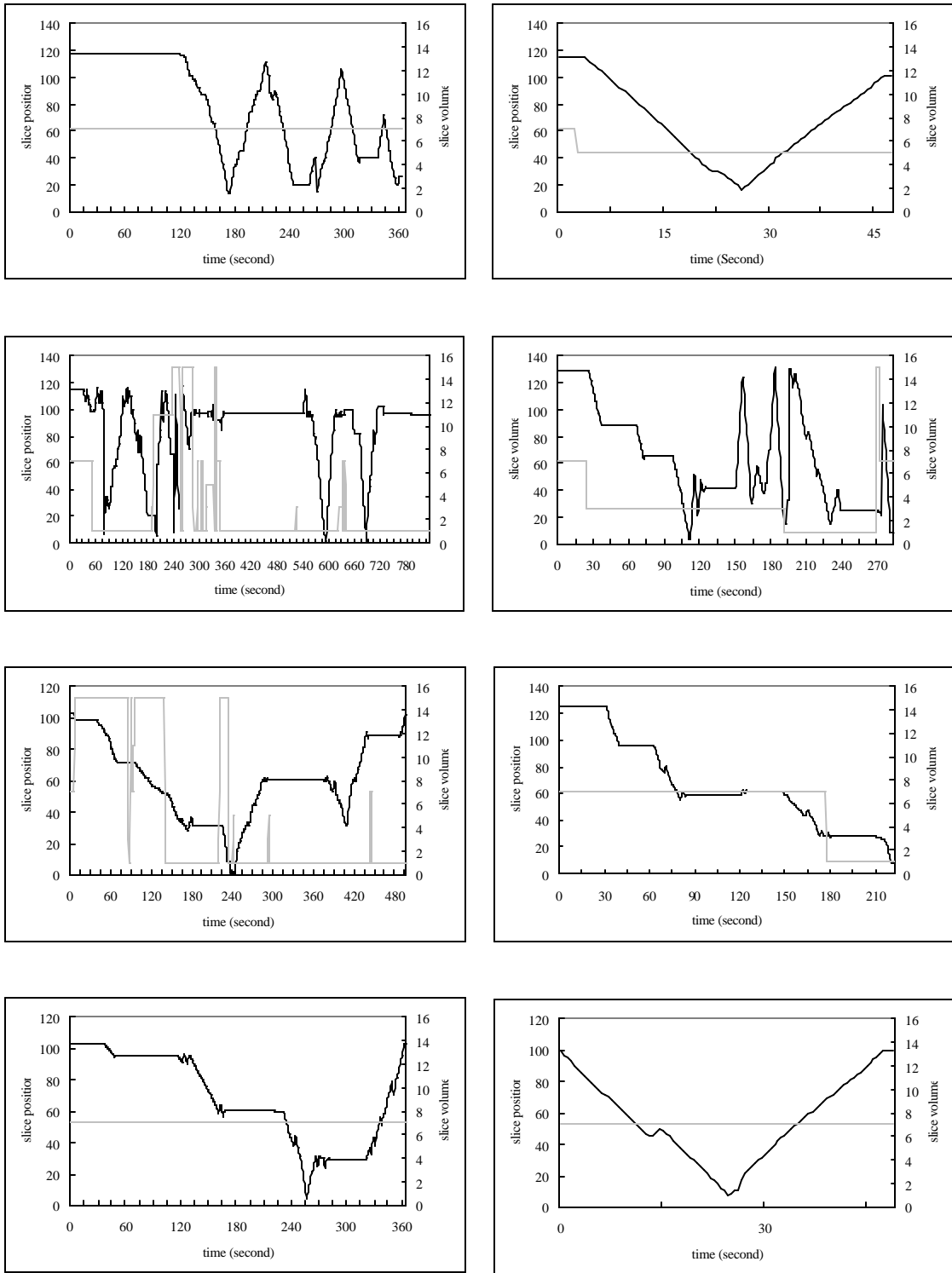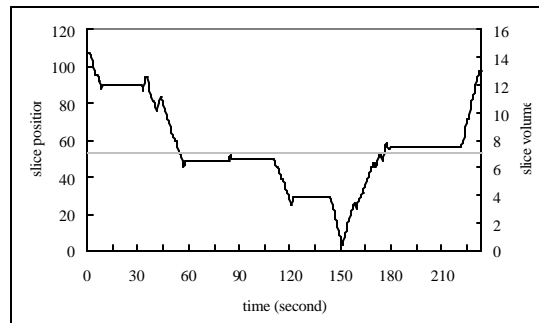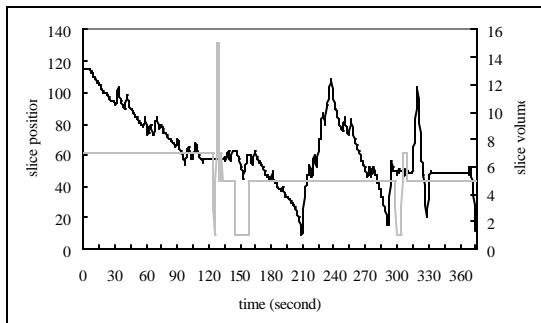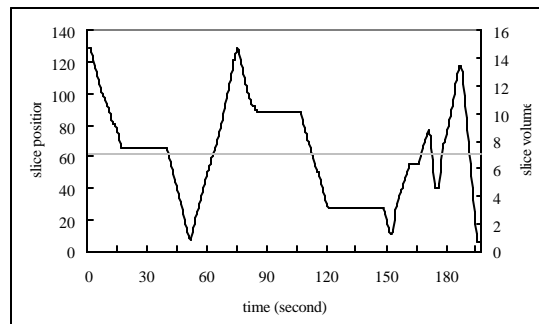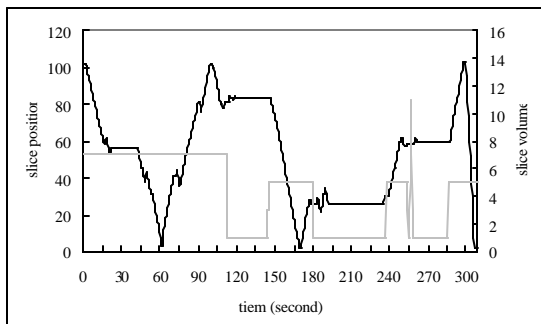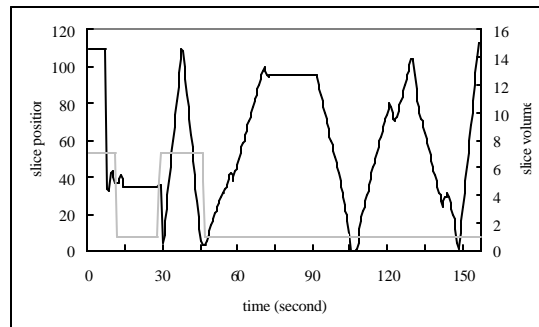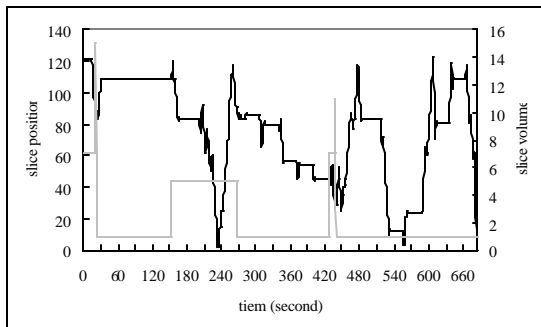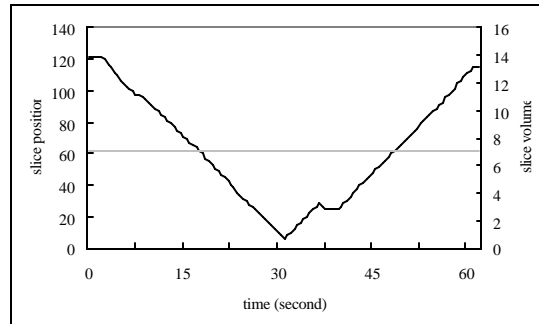
Figure 2. Navigation patterns from four radiologists in **orthogonal MIP mode**. Each graph is the navigation recorded from one case interpretation. Two graphs in each row are taken from one radiologist's interpretations in the beginning of the study (left) and the end of the study (right). The dark solid line represents viewing positions referenced to the scale on left axis and grey solid line represents viewing thickness referenced to the scale on right axis, with time.
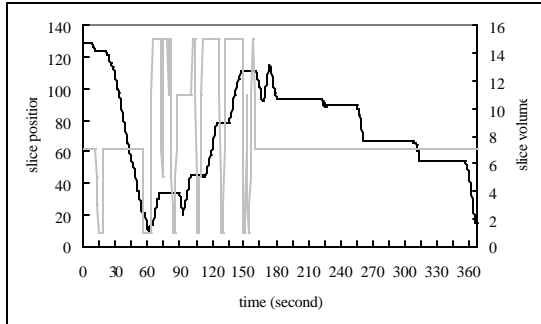
Figure 3. Navigation patterns from four radiologists in **slice-by-slice mode**. Each graph is the navigation recorded from one case interpretation. Two graphs in each row are taken from one radiologist's interpretations in the beginning of the study (left) and the end of the study (right). The dark solid line represents viewing positions with the time.
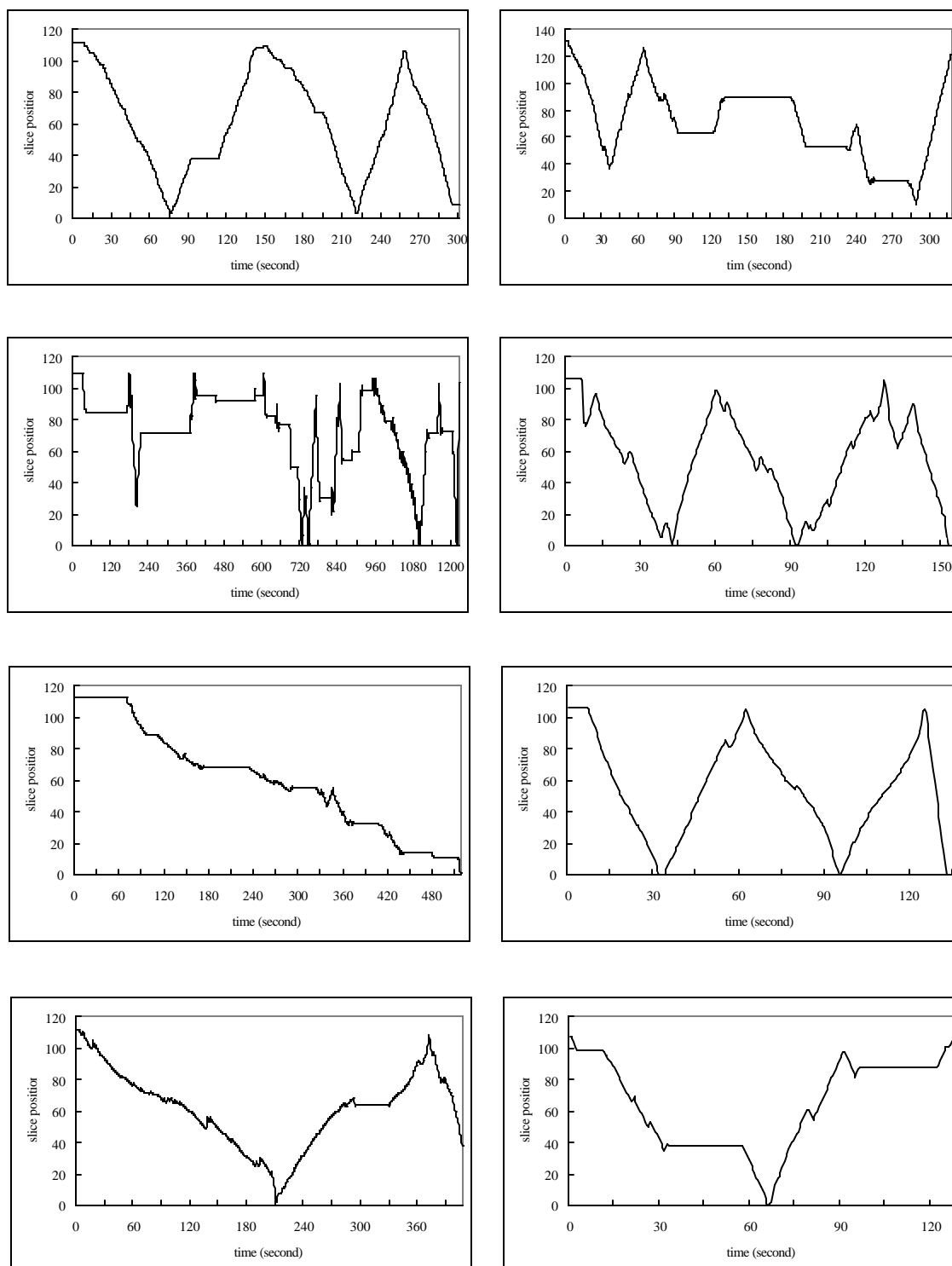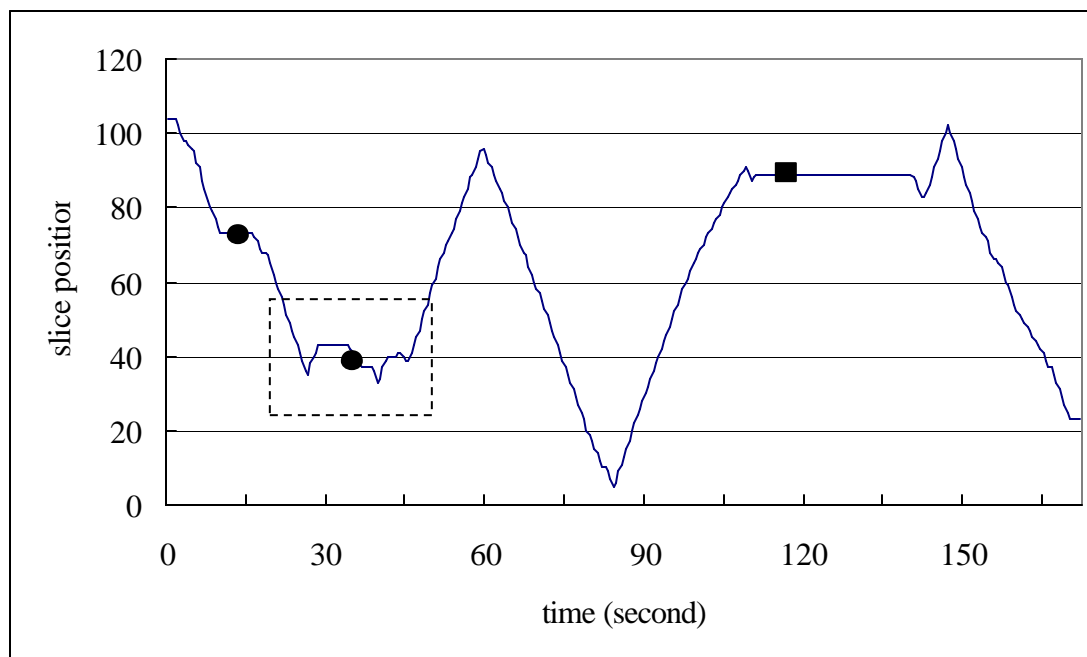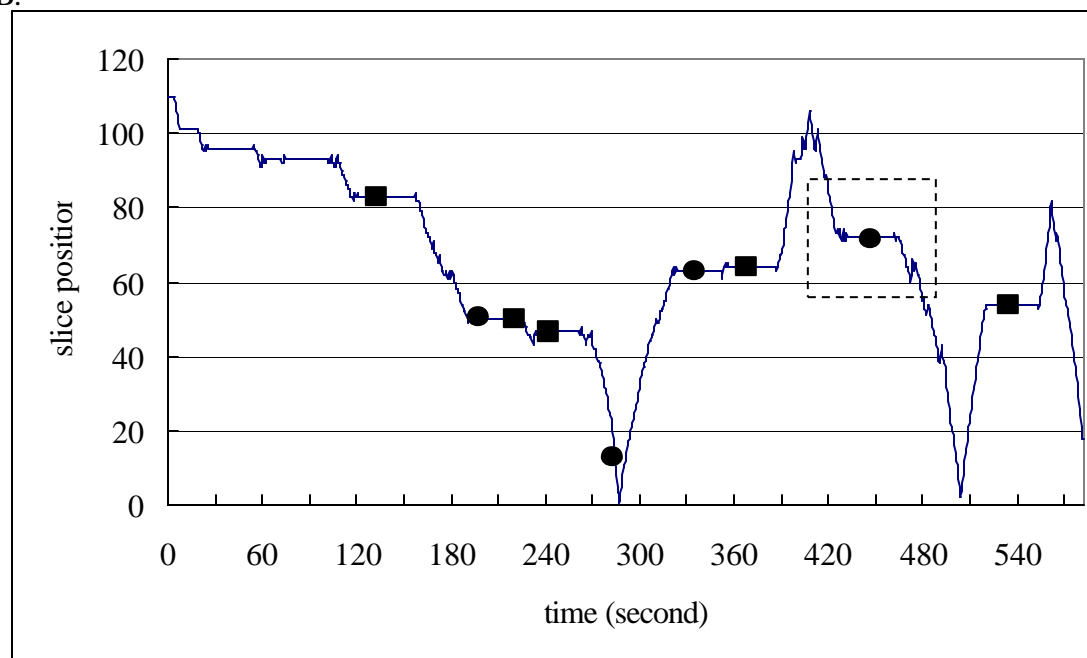
Figure 4. The diagrams A and B illustrate the missed nodules that received extra attention.
— viewing position ⌐ ¬ location of missed nodules ● true nodule ■ false nodule

A.

B.

Appendix C

Stereo display of CT images for lung cancer screening: a pilot study

Xiao Hui Wang, Janet E. Durick, David L. Herbert, Amy Lu, Saraswathi K. Golla, Dilip D. Shinde, Samaia Piracha, Kristin Foley, Carl R. Fuhrman, Betty E. Shindel, J. Ken Leader, Walter F. Good

Department of Radiology, University of Pittsburgh

Purpose: To improve radiologist's performance in lesion detection and diagnosis on 3D medical image dataset, we have developed a stereo display workstation and then conducted a pilot study to test viability and efficiency of the stereo display for lung nodule detection and classification.

Methods: Using our previously developed stereo compositing methods, stereo image pairs were prestaged and precalculated from various number of CT slices for real-time interactive display. A computerized study program has been built and organized as FROC study to compare three display modes (i.e., stereoscopic 3D, orthogonal MIP and slice-by-slice). Cases used in the study were randomly selected from lung cancer screening population, and total of eight radiologists have participated this pilot study to interpret the images in all display modes. The performance of lung nodule detection was analyzed and compared between the modes using Free-Response Receiver Operating Characteristic (FROC) analysis.

Results: Subjective assessment indicates that stereo display was well accepted by the radiologists, despite some uncertainty of beneficial results due to the novelty of the display. The FROC analysis indicates a trend that, among the three display modes, stereo display resulted the best performance of nodule detection followed by slice-based display, although no statistically significant difference was shown between the three modes.

New or breakthrough work to be presented: The stereo display of a stack of thin CT slices has the potential to clarify three-dimensional structures, while avoiding ambiguities due to tissue superposition, commonly associated with thicker slices. Few studies, however, have addressed actual utility of stereo display for medical diagnosis.

Conclusions: Our preliminary results suggest a potential role of stereo display for improving radiologists' performance in medical detection and diagnosis, and also indicate some factors likely affect the performance with new display, such as novelty of the display, training effect from projected radiography interpretation and confidence with the new technology.

Keywords:
    Stereo display, Lung CT, Visualization, Validation, FROC

Appendix D

Application of the 3D Tensor Voting Paradigm for Segmenting Cylindrical Segments and Bifurcations from Volumetric Datasets

Walter F. Good, Xiao Hui Wang, Carl Fuhrman, Jules H. Sumkin, Glenn S. Maitz, Joseph K. Leader, Cynthia Britton, David Gur

Department of Radiology, University of Pittsburgh

1) Description of purpose – Many diagnostic problems involve the assessment of vascular structures or bronchial trees depicted in volumetric datasets. Analytical methods for segmenting these structures have been applied, in combination with *ad hoc* heuristic techniques, for the past couple of decades, but previous algorithms are not sufficiently robust for them to be widely applied clinically because they generally do not utilize all of the relevant information available in datasets, and do not employ methods based on an understanding of psychophysical aspects of human perception of salient structures.

2) Methods – This study attempts to improve the segmentation process and the depiction of small arteries and bifurcations by exploiting all information in structure tensor fields derived from 3D pulmonary datasets, along with context dependent information about expected structures. In particular, tensor voting, which attempts to replicate the way human observers perceptually organize features in images or in spaces of arbitrary dimension, was employed to identify voxels comprising surfaces of vessels and bifurcations. The justification for this is a belief that perceptual methods learned by humans, for organizing visual scenes, may be very nearly optimal for organizing spatial information and detecting salient objects. While this is not known, and much research remains to be done to characterize human vision, it is well known that current computerized vision systems cannot come close to the performance of human vision systems.

Gradients, which provide surface normals, and Hessian matrices, which provide eigenvalues and eigenvectors, are calculated and preserved. Prior expectations about diameters and branching angles were employed to constrain filtration processes, which defined the scale of the segmentation process. Subsets of locations are identified as token locations and structures are continued between tokens by performing voting between the tensors in their immediate neighborhoods. This mechanism is ideal for identifying coherent structures, such as cylindrical objects and bifurcations.

These methods have been tested on simulated datasets at varying noise levels, as well as on actual volumetric pulmonary CD studies. Simulated arterial segments were generated with a range of diameters and with various levels of added Gaussian noise. Copies of these images also had gaps of various sizes artificially inserted. Both the current algorithm and a traditional segmentation algorithm, which uses the ratio between eigenvalues as a measure of cylindricity, were used to segment the synthetic data.

3) Results – The algorithms performed similarly for cases where there were no gaps, cylinder diameters were relatively large, and the signal-to-noise ratio was high. The performance of the traditional algorithm deteriorated more rapidly than the performance of the newer algorithm, as any of these conditions were degraded. Specifically, the tensor based methods have been shown to be relatively insensitive to noise and outliers, and can bridge discontinuities in structures in a manner that is perceptually plausible.

4) New or breakthrough work to be presented – This is one of the first attempts to segment arterial and bronchial structures using methods that exploit all relevant information contained in the image in combination with context sensitive expectations and methods based on an understanding of psychophysical physical characteristics of human perception.

5) Conclusions – Tensor voting techniques seem to provide many advantages over traditional methods for segmenting arterial or bronchial structures. Their main disadvantage is in their computational complexity but it seems likely that more efficient methods can be developed.