Searching in Axial and 3D CT Visualisations

Peter Phillips[a], David Manning[a], Trevor Crawford[b],
David Burling[c], Chi-Leung Tam[d], Alasdair Taylor[d]

[a]University of Cumbria, UK
[b]Lancaster University, UK
[c]Intestinal Imaging Centre, St Mark's Hospital, Northwick Park, UK
[d]Department of Radiology, Lancaster Royal Infirmary, UK

ABSTRACT

Traditional diagnostic modalities have been, for the most part, static two-dimensional images displayed on film or computer screen. More recent diagnostic modalities are solely computer-based and consist of large data-sets of multiple images. Image perception and visual search using these new modalities are complicated by the need to interact with the computer in order to navigate through the data. This paper reports the late-breaking results from two small studies into visual search within two types of CT Colonography (CTC) visualisations. The twelve novice observers in the study were taking part in a week-long course in CTC and were tested at the beginning and end of the course. A number of expert observers were also recorded. The two visualisations used in the study were 2D axial view and 3D colon fly-through. In both cases, searching was performed by inspecting the colon wall, but by two distinct mechanisms. The first study recorded observer eye-gaze and image navigation in a CTC axial view. The search strategy was to follow the lumen of the colon and detect abnormalities in the colon wall. The observer used the physical computer interface to navigate through the set of axial images to perform this task. The 3D fly-through study recorded observer eye-gaze whilst watching a recording of a computed flight through the colon lumen. Unlike the axial view there was no computer control, so inspection of the colon surface was dictated by the speed of flight through the colon.

Keywords: Eye Tracking, Visual Search, Human Computer Interaction, Virtual Colonoscopy

1. INTRODUCTION

Computers have now become well established in the radiology workplace. Technologies such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) are now extremely well established, and their continuing evolution has reached the boundaries of traditional film-based working methods. Each new generation of modern CT equipment can produce thinner slices and larger image sizes. Equipment can now produce gigabytes of data per acquisition, representing hundreds of images, more than can be printed out onto film or displayed simultaneously on a workstation. As the volume of data increases, new methods of handling and interpreting this data are required.

A popular method for viewing a CT or MRI data-set on screen is known as *stack mode* viewing. In this mode images are logically stacked in sequence on top of each other, with a single image being displayed on screen at any one time. Each image in the stack can be thought of as a frame in a movie, but unlike a movie the speed of playback is under the control of the observer. What differentiates stack mode viewing from traditional hard or soft copy 2D viewing, is the need for physical interaction with the computer so as to see the whole data-set. In static 2D viewing, all the data is available to see at once. Stack mode viewing requires observer interaction with the computer, in addition to visually searching what is currently displayed on screen.

The evolution of imaging technology has seen image size and quality increase. Modern multi-slice CT can now produce hundreds of images per individual scan. Stack mode interpretation, by necessity has become longer. Increases in speed and memory within computers have been able to accommodate the increase in stack sizes, ensuring that images in the stack can be computed and rendered to the screen with little or no delay or artifact. The relentless improvement of 3D graphics technology has also opened up another method of viewing large sliced data-sets. The performance and affordability of 3D-capable workstations now permit sliced data-sets to be computed and presented as interactive 3D reconstructions. Observers no longer need to navigate the data

Medical Imaging 2008: Image Perception, Observer Performance, and Technology Assessment,
edited by Berkman Sahiner, David J. Manning, Proc. of SPIE Vol. 6917, 69171J, (2008)
1605-7422/08/$18 · doi: 10.1117/12.790424

from a scan in a serial manner, but can explore with up to six degrees of freedom. The continuing trend for increasing data-set size is likely to push radiology further towards 3D interpretation.[1]

Both stack and 3D visualisations of large data-sets have a necessary navigation component, in addition to the perception and interpretation tasks that are required for the diagnostic process. What impact does this navigation requirement have on visual perception? How do eye movements differ in these two visualisation methods compared to traditional static 2D viewing?

This paper presents some late-breaking results from an eye tracking study into visual perception using stack mode and 3D visualisation methods. Eye tracking multiple image modalities such as CT is by no means new; there have been several studies into the benefits of stack mode versus tile mode viewing.[2–5] The tile mode method arranges images from a CT or MRI scan across a surface, be it film or a computer screen, so that all of the data-set is visible at once. Tiling works well with smaller numbers of images from a scan, but with the advent of multi-slice CT scanners, and the increasing digitisation of radiology, tile mode viewing of a single scan has increasingly been replaced. Digital tiling now occurs between computed visual planes of a scan or to compare two or more scans side by side. Multi-slice CT scanners are also opening up new fields of diagnostic radiology,[6] one of which is CT, or Virtual, Colonography (CTC), which utilises both stack and 3D methods.

2. CT COLONOGRAPHY

CT colonography uses a modern multi-slice CT scanner to obtain and compute a set of images of the patient's colon. A day or two prior to the scan, the patient takes a laxative so as to clean the bowel of feces. As the patient is on the scanner, their bowel is inflated with gas to distend the colon. Both these actions improve the visibility of the colon wall where polyps or other lesions might be present. The patient is usually scanned in prone and supine positions, so as not to miss a part of the bowel wall which might have been submerged in any residual fluid on the first scan. Abnormalities are discovered by viewing the images from the patient's scan using a suite of methods. Computed axial, saggital, and coronal stack views of both prone and supine scans allow a slice by slice investigation of the bowel wall. Specialist software can identify the colon boundaries from the scan, mathematically smooth the slices together, and produce a 3D visualisation of the bowel wall. The primary aim of the observer is to check all of the bowel wall for any abnormality for both methods.

2.1 Overview of Studies

The initial results presented in this paper come from a pair of small studies conducted around a week-long CTC training course at St Marks Hospital, England. The first study measured participant eye and navigation behavior when viewing a single scan presented using the stack method, at the beginning and end of the training week. The second study measured participant eye movements whilst watching a video playback of a 3D colon fly-through, and was composed of both expert and novice viewers. The use of the term novice in this paper is a catch-all term for participants who have no familiarity with CTC. Participants taking the training courses came from a number of hospitals around England, and had a variety of backgrounds. Some were familiar with the physiology of the bowel, being barium radiographers. Some had CT experience, so had an idea of how to navigate through a CT stack, but had no previous bowel experience. None of them had any direct CTC experience. Those termed experts in these studies are participants who perform some sort of CTC reporting on a regular basis.

Both these studies were designed to collect data on eye movements and patterns of navigation. They are not yet large enough to make statistical judgments on diagnostic accuracy or performance. The purpose of this paper is to highlight the difference in eye movement and navigation style within observers in training and those who have day-to-day experience of using the these modalities.

3. AXIAL STACK MODE STUDY

The first study investigated eye movements and navigation in axial stack mode viewing, and any changes that occurred in these two properties due to an intense week long training course. Participants were asked to view either a prone or supine scan of approximately 450 images in an axial stack presentation. Windowing was fixed at the recommended colon windowing level and not changed during the course of viewing. Participants had to follow the path of the colon, starting at the rectum, and mark points they thought were abnormal.

Navigation through the stack was performed using a USB mouse as an input device; utilising either the scroll-wheel between the left and right buttons, or by pressing the right mouse button and dragging. The visualisation software used was MedicRead Colon by MedicSight,[7] which had been modified to record the time and image number of the currently visible stack image. The computer screen was a 1600 pixel by 1200 pixel Dell widescreen LCD monitor. A Tobii X50[8] 50Hz eye tracker was positioned under the screen to record movement in both left and right eyes of the participant, and analysed using in-house software.[9] The threshold for a fixation was set at 100ms.

Figure 1. A single image from a CTC stack. The two dark regions are the lumen of the colon. The larger region has 500ms of eye data, investigating some Haustral folds.

3.1 Initial Findings

The total study time was an indication of the complexity of this task to those observers with no experience of it. The task of navigating through the stack whilst simultaneously following the seemingly convoluted route of the colon maked this task very difficult. In its most simple form, this is task was an exercise in tube following. Two previous studies into stack mode tube following, both comparing with tile mode, had used simple entangled tubes,[4] or porcine intestines.[5] The porcine intestine study commented on the convoluted nature of the intestine tube, compared with the previous study. For a CTC observer the problem increases in complexity further due to several factors. The appearance of the colon can be affected by the natural Haustral folds in the wall, fecal residue, and in some cases diverticular disease. Figure 1 shows a single image from a stack with overlaid eye movement data. The colon wall is not round or regular in shape. The incursions into the centre of the lumen are due, in this image, to Haustral folds.

Table 1 shows that the average time for viewing a single prone or supine stack, after undergoing a week of training, was reduced by nearly 50% to around 6 minutes. Naturally, the number of fixations fell from pre to post-training, but taking into account the cumulative duration of the fixations, the percentage of time spent *in fixation* rose slightly.

	Pre-Training	Post-Training
Total study time	712.27 sec	368.90 sec
Number of fixations	809	461
Percentage of time in fixation	41%	50%

Table 1. Stack mode study results.

3.1.1 Navigation

Perception, and therefore correct decision making, are dependent on successful navigation of the stack of images. In this study, the two possible controls for moving within the stack were on the workstation mouse. The mouse scroll wheel, placed in between the left and right mouse buttons, is a familiar control for many computer users for scrolling a page up and down the screen. Unlike the other navigation method, pressing the right mouse button and dragging, it allows the mouse pointer to also be used as a marker on screen. Three out of the twelve novice participants began their training week using this method, and produced three of the longest study times. The drawback of mouse-wheel navigation is the stop-start mechanics required for turning the wheel. The index finger can only turn the wheel so far before it must be lifted and repositioned on the wheel so scrolling can continue. This makes navigation through a stack very jerky and slow, making it difficult to build up a sense of smooth motion within the changing images. All three participants used the mouse pointer as a movable marker on the stack, placing it in the centre of the section of colon they were trying to follow. Mouse-drag navigation, where the right button is pressed to signal movement and the calculated vertical offset of the mouse controls the position in the stack, does not require resetting of the finger or hand. The display software computes the offset and updates the stack view. By the end of the week all three mouse-wheel users had adapted their navigation technique to use mouse-dragging, and their study times had dropped. For this type of task it is recommended that input methods and their associated physical devices allow smooth and continuos changes in stack speed. A previous study into the effectiveness of other input devices for CT navigation, such as jog-wheels and trackballs,[10] also had similar findings.

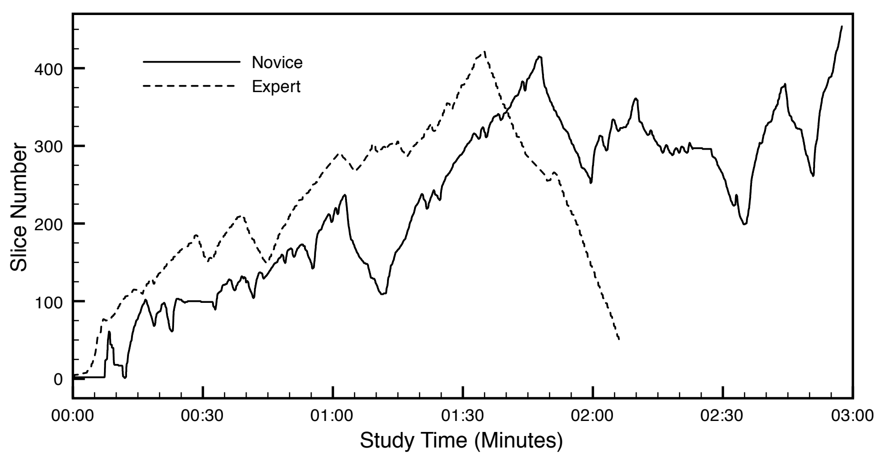

Figure 2. Novice and expert navigation paths through a prone view image stack.

Figure 2 shows the difference in stack navigation for an expert and a novice participant (post-training). The first image in the stack, numbered 1, is always at the rectum. Large changes in direction, switch-backs, of the colon are shown as major peaks on the graph. Clusters of smaller peaks and troughs represent parts of the colon that an observer needed to *rock back and forth* over, repeatedly changing the direction of navigation, so as to examine a localised section of colon. The expert examined the colon in just over two minutes, whilst the novice was still examining after three minutes. An example of *rocking* around the same section of colon wall (around image 300) can be seen in the expert at 1:10 minutes, and in the novice at 2:20 minutes. Another notable feature of figure 2 is the smooth change of speed of the expert, compared to the novice. The novice plot shows several places where the observer lost the path of the colon and had to backtrack to a previously known point; 10 seconds into their study being a good example.

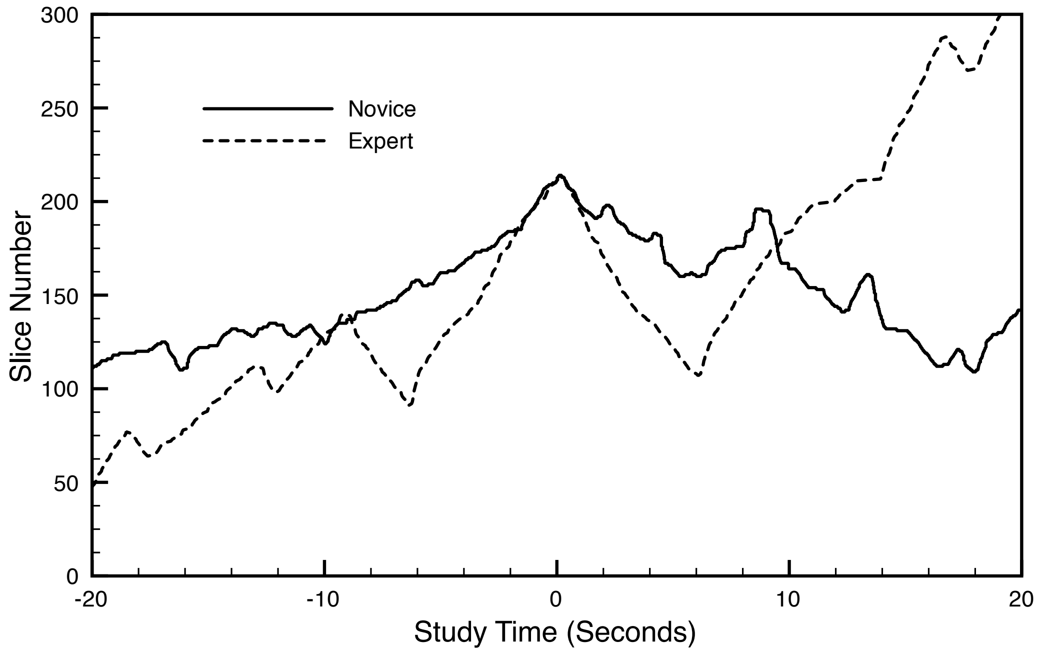

Figure 3. Part of two stack navigation paths in a 40 second window, centred around a specific image number.

The two lines in figure 2 share several key *peaks* at certain image numbers, where the colon made a marked change in direction. Figure 3 takes one of these peaks and aligns a novice and expert scan around it. Setting the study time to zero at the point of the peak, and drawing a 20 second window before and after the peak, magnifies the navigation in the local region and highlights further the differences between the expert and novice techniques. The speed of navigation through the stack given by the gradient of the line. It is very smooth for the expert, and the change of the gradient is also smooth compared to the novice. In the 40 second window they have navigated from image 50 to beyond image 300. In the same time period the novice follows the colon from image 112 to 140, a volume of the colon covered by the expert from -6 seconds to +7 seconds (a total of 13 seconds). The gradient of the novice line is shallower than the expert, indicating a slower navigation through that particular part of the stack. The novice line is also punctuated by several changes in direction, using the mouse to rock back and forth to review several images (-15 to -10 seconds being an example). The expert shows no such activity in this section of their study.

4. CTC FLY-THROUGH STUDY

The second study recorded observer fixations during a fixed speed 3D endoluminal fly-through. A single CTC scan fly-through was rendered and exported from a Viatronix[11] workstation to a digital video file at 15 frames a second. The fly-through started at the rectum and travelled at a fixed speed through the colon. The duration of the fly-through was 133 seconds, and participants were asked to audibly remark on any suspicious areas as they came into view. The participants in this study were made up of 9 of our previously defined CTC experts and 9 novices on the CTC training course. Novices viewed the fly-through towards the end of the training week. Once again a Tobii X50[8] 50Hz eye tracker was used to record eye movement. The threshold for a fixation was set at 80ms.

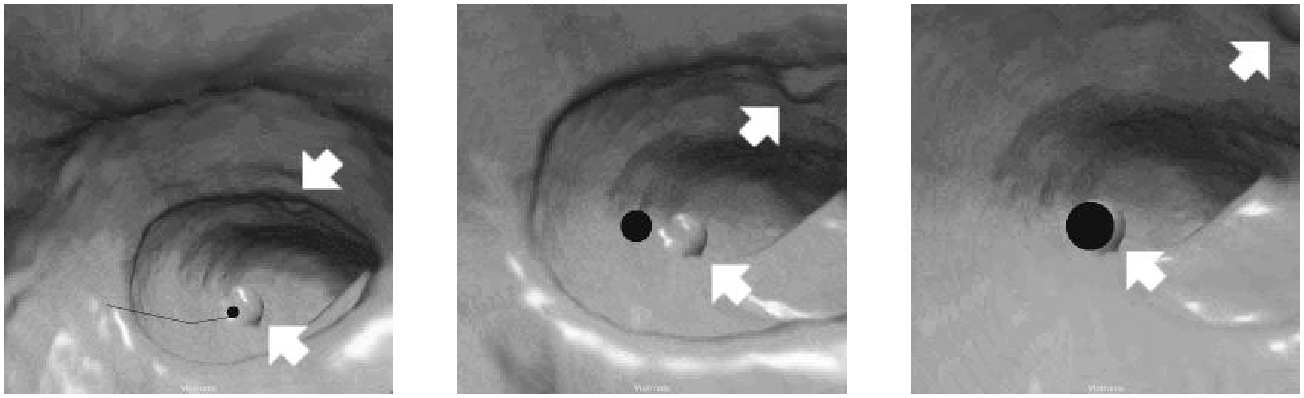

Figure 4. Three frames taken from a endoluminal fly-through, showing two optically confirmed polyps (arrowed) and observer eye gaze. As the fly-through progresses, reading from left to right, the observer fixates on the lower polyp. The upper polyp is not fixated, and moves out of frame as the fly-through proceeds.

Unlike the first stack mode study, the fly-through had was no navigation component for a participant to contend with. As the fly-through had been rendered to a linear digital video file, it was not possible to stop, turn round and otherwise navigate within the colon.

4.1 Initial Findings

In a study of this kind an observer has a fixed time in which to view the whole fly-through, and the frames within it. Compare this to a time-limited but free search of a static 2D image. In a static image, an observer has a global impression of the image and can re-fixate on areas of interest if needed. Fixation times can also vary. For the video task, the view is constantly changing as the fly-through travels along the colon. An observer cannot afford the luxury of fixating an area for a long period of time as it will move out of view and other areas of the wall will be missed. The speed of fly-throughs also makes re-fixations difficult and costly. Consequently the average fixation time must, by necessity, be less than a static 2D image, to sufficiently examine the colon wall during the fly-through.

Figure 4 shows three frames from a fly-through, containing two polyps (arrowed)confimed by optical colonoscopy. A novice observer's fixation is shown on each frame, as a dark circle. The larger the circle, the longer the fixation. The first frame shows the observer saccading to the lower polyp and fixating near by. As the polyp draws nearer, the observer re-fixates near its position. The third frame shows the observer performing a long fixation directly over the lower polyp. During this time, the upper polyp has passed by, in view, unfixated. The observer audibly confirmed only the lower polyp.

Figure 5 shows the spread of fixation durations for each participant in the fly-through study. Experts, those with some familiarity with CTC, are grouped on the left while novices or grouped on the right. Although the difference between average fixation times for these two groups is not statistically significant enough, there are some features worthy of comment. For nearly all participants, 75% of their fixations fall under 500ms, giving some indication of the strategy required for the task. The standout experts, towards the middle of figure 5, have all of their fixation durations around 1000ms or less. Visual inspection of the eye data indicated that some short duration fixations would be missed if the fixation threshold of 100ms was used. For the analysis in this paper, the fixation level was dropped to 80ms to capture more of these short fixations.

5. CONCLUSIONS

The late breaking results from this study have indicated that the fixation characteristics of the stack and 3D fly-through presentation methods are markedly different to those usually found in analysis of static 2D images. Shorter, faster fixations are an indication of familiarity with the presentation method, if not an indication of expertise. Unlike 2D static images where fewer, longer fixations are indicative of expertise, long fixations in these

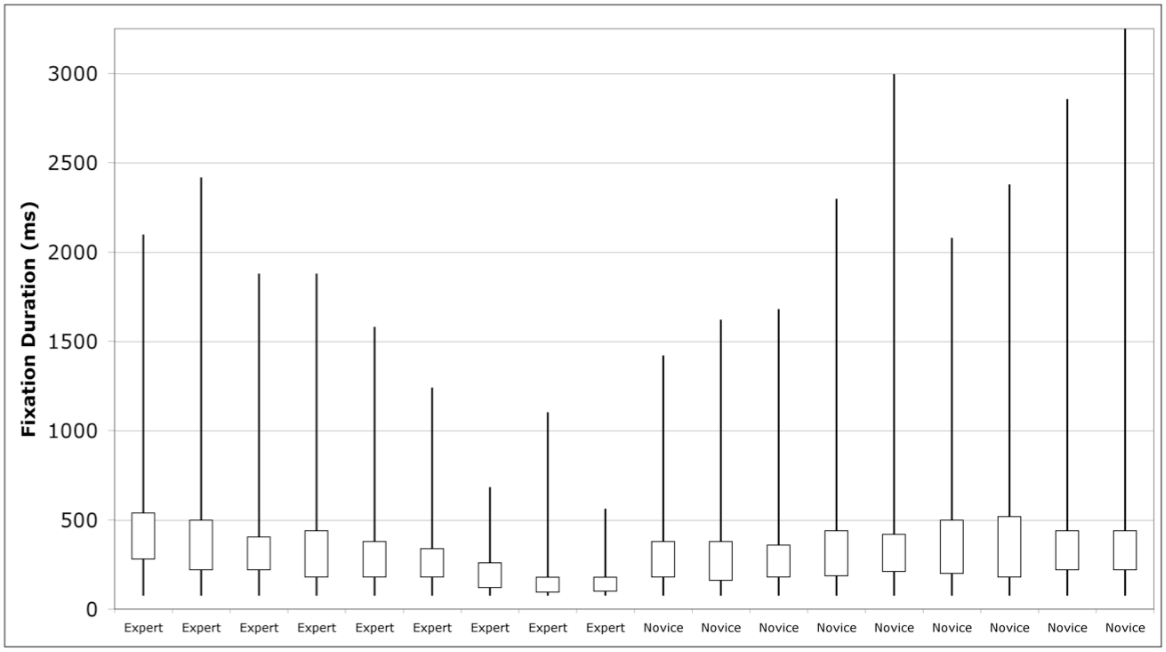

Figure 5. Fixation spread between expert and novice observers in the 3D fly-through study.

studies are costly to the overall search strategy. Short fixations to enable greater coverage of the colon wall in the limited time that it appears on screen.

CTC itself is a completely new way of looking at the colon. The novel visual presentation and physical manipulation of the scans force even those with previous bowel imaging experience, to re-learn their way around the colon.[12] As experience increases, study times are reduced. Eye movements and search strategy are intimately tied to the observer's navigation behaviour. The use of the wrong input device can lead to longer study times, and a stop-start progression through the colon. The appearance of movement, the visual flow of over a steady progression of images, is an important visual channel for the observer. A physical device that gives *bursty* input to the computer interrupts this visual flow. Three users modified their behaviour to improve navigation quality and avoid this bursty feature.

Further analysis is needed on the threshold of a fixation. Very short fixation durations of experienced observers, particularly in the 3D fly-through study, may be more akin to the fixation types seen in reading, rather than classic image interpretation. It might be possible to break down fixation characteristics further, and identify more patterns of fixation. Future work will also attempt to identify points in a stack mode study where the observer becomes lost in the colon. When navigating through the colon, inexperienced users lost their place or became confused about which direction they were traveling through the stack. An error such as this is positioned at the boundary between visual perception and navigation. Within the collected data there could be some indication of what causes this breakdown.

REFERENCES

1. K. Siddiqui, E. Siegel, and B. Reiner, "Interpretation strategies for large dataset multislice ct studies." Web page: www.imagingeconomics.com, July 2004.
2. D. V. Beard, R. E. Johnston, O. Toki, and C. Wilcox, "A study of radiologists viewing multiple computed tomography examinations using an eyetracking device.," *J Digit Imaging* **3**, pp. 230–7, Nov 1990.
3. S. E. Seltzer, P. F. Judy, D. F. Adams, F. L. Jacobson, P. Stark, R. Kikinis, R. G. Swensson, S. Hooton, B. Head, and U. Feldman, "Spiral CT of the chest: comparison of cine and film-based viewing.," *Radiology* **197**, pp. 73–8, Oct 1995.

4. A. G. Mathie and N. H. Strickland, "Interpretation of CT scans with PACS image display in stack mode.," *Radiology* **203**, pp. 207–9, Apr 1997.
5. Y. J. Kim, J. K. Han, S. H. Kim, J. Y. Jeong, S. K. An, C. J. Han, K.-R. Son, K. H. Lee, J. M. Lee, and B. I. Choi, "Small-bowel obstruction in a phantom model of ex vivo porcine intestine: Comparison of pacs stack and tile modes for ct interpretation," *Radiology* **236**(3), pp. 867–871, 2005.
6. J. Rydberg, K. A. Buckwalter, K. S. Caldemeyer, M. D. Phillips, J. Conces, Dewey J., A. M. Aisen, S. A. Persohn, and K. K. Kopecky, "Multisection CT: Scanning Techniques and Clinical Applications," *Radiographics* **20**(6), pp. 1787–1806, 2000.
7. MedicSight. http://www.medicsight.com.
8. Tobii Technology. http://www.tobii.se.
9. P. W. Phillips, D. J. Manning, T. Donovan, T. Crawford, and S. Higham, "A software framework for diagnostic medical image perception with feedback, and a novel perception visualization technique," in *Medical Imaging 2005: PACS and Imaging Informatics.*, S. C. Ratib, Osman M.; Horii, ed., **5749**, pp. 572–580, SPIE, Apr. 2005.
10. A. J. Sherbondy, D. Holmlund, G. D. Rubin, P. K. Schraedley, T. Winograd, and S. Napel, "Alternative input devices for efficient navigation of large CT angiography data sets.," *Radiology* **234**, pp. 391–398, Feb 2005.
11. Viatronix. http://www.viatronix.com.
12. S. Taylor, S. Halligan, and C. Bartram, "Ct colonography: Methods, pathology and pitfalls," *Clinical Radiology* **58**, pp. 179–190, March 2003.