Designing a Radiology Workstation: A Focus on Navigation During the Interpretation Task
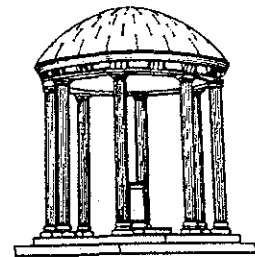
TR91-004

August, 1990

David Beard

Medical Image Display Group
Department of Radiology
The University of North Carolina
Chapel Hill, NC 27599-7510

Designing a Radiology Workstation: A Focus on Navigation During the Interpretation Task

David Beard

The potential cost and logistic improvements of picture archive and communication system (PACS) over film-based medical image management awaits the development of viable radiology workstations (RWS) targeted for the primary interpretation task. While the quality of electronically displayed images has been highly investigated, only recently have design and experimental work been devoted to the other critical aspect of workstation design, mainly its computer human interaction, and, in particular, its navigation. By RWS navigation we include its underlying mental model or metaphor, and the commands and hand motions used to access patient folders and to display images. For the last 5 years, the University of North Carolina (UNC) Medical Image Display Research Group has analyzed the primary interpretation task and designed, developed, and evaluated the FilmPlane series of RWS prototypes. This work has helped us understand both RWS requirements and viable design approaches. In this paper, we present our workstation design strategy and our observations and understanding of the issues and problems with RWS navigation. To illustrate the discussion, we describe FilmPlane2, the UNC radiology workstation. We also briefly detail three rapid-evaluation techniques (including two observer experiments) for quickly gaining feedback on a design. These observations and evaluation techniques may aid other RWS designers in producing superior tools for the clinicians.
© 1990 by W.B. Saunders Company.

KEY WORDS: computer human interaction, radiology workstations, human factors, medical image display.

DESIGNING AN acceptable computer-human interaction (CHI) is art rather than science. Nevertheless, a general methodology has been widely adapted that allows for the rapid evolution toward a viable product.[1] First, a well designed CHI starts with a complete requirements analysis including an understanding of the users and their tasks. Second, a clear mental model or metaphor must be developed that allows the user to understand quickly the workstation operation. Third, the complete interaction must be carefully designed and implemented. Fourth, various subject experiments, observation sessions, and time-motion analysis techniques must be used to refine the workstation into a viable tool. Designers must remember that as toolsmiths, their objective is to develop a productive tool for the user. In the final analysis, "the quality of a swordsmith is measured by the longevity of his customers."[2]

First, the radiology interpretation task is described. Second, workstation design issues are discussed. Third, the design of FilmPlane is presented including objectives, mental models, direct-manipulation interaction, image navigation, display-area requirements, and response-time requirements. Fourth, a series of experiment protocols are described that were used to evolve FilmPlane; these protocols may prove useful to other workstation designers.

RADIOLOGY WORKSTATION REQUIREMENTS

The design of a radiology workstation (RWS) must begin with a clear understanding of tasks performed with film and viewbox, including both projection radiographs and multiple cross section digital studies such as computed tomography (CT).[3,4] During the interpretation task, the radiologist views the images resulting from the radiology procedure, and together with other information (patient history, images from previous procedures, input from the referring physician, etc), interprets the images for the referring physician. The primary interpretation task is typically performed on either a "four over four" viewbox array, or an "alternator," a device which stores the films on a large horizontal strip of clear plastic that is mechanically moved over a row of four viewboxes, alternating which four films are displayed (Figure 1). Smaller viewbox configurations are used for viewing single studies.

A successful interpretation requires text as well as image information. The requisition form details the medical questions the referring physi-
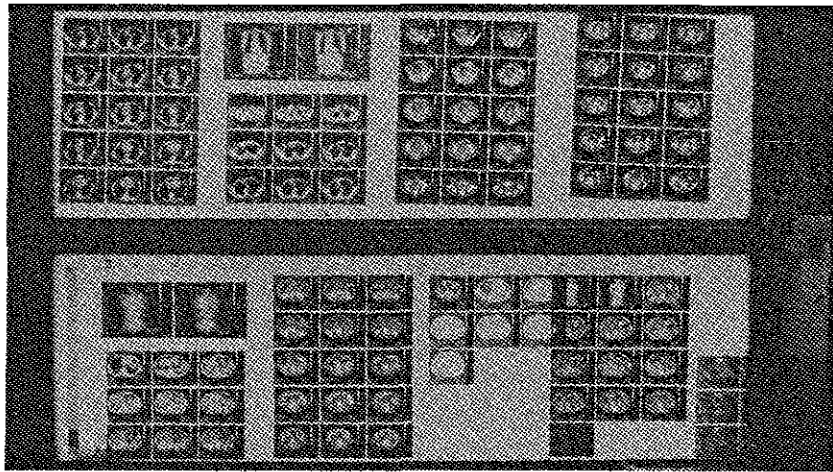
Fig 1. Alternator with multiple CT studies.

cian wants the radiology procedure to answer; these medical questions greatly affect the search patterns the radiologist uses when viewing the images.[5] The radiology procedure date, slice thickness, type of radiologic device, the name of the technician that performed the procedure, whether contrast was administered, etc, is provided. The radiology information system (RIS) provides patient information. Finally, the interpretation reports generated from previous radiology studies are (ideally) available.

Image Scan Patterns

An Eyemark EMR-V eyetracker was used to study interpretation of one class of critical patient folders: those containing single and multiple CT chest studies.[6] This eyetracker, which allows head and body movement, determines the position of a subject's eyes, and superimposes corresponding markers onto a video tape. The following search pattern was generally observed. First, images were removed from the patient folder, sorted, and some of the films are placed onto the viewbox. Images were viewed during this viewbox loading process. Second, a systematic search pattern was performed over all the images, sometimes by organ, but most often as a systematic sequential scan. Third, critical slices showing important anatomy are reviewed in detail. Finally, an interpretation report is dictated, often while continuing to view the images.

A great deal of time is spent locating and accessing small clusters of images that show important anatomy. An image "cluster" is identified when a radiologist repeatedly fixates on the same proximate images. Based on our eye movement studies, radiologists only seem to view small localities of reference, that is, small clusters of typically three to six CT images. However, the remaining images appeared to be used as a pictorial image index, allowing the radiologist to locate quickly any particular image. With patient folders containing multiple studies, radiologists often compare clusters of images from the current study—typically from two to four images—to clusters from previous ones to determine whether an anatomical object is abnormal, or whether it had increased or decreased in size.

Others have recorded eye movements of radiologists reading medical images,[5,7,8,9] though most of this research was for determining source of interpretation errors rather than developing task analyses. Experienced radiologists employ a variable, though basically circumferential, scan pattern when reading radiographs.[10] Search patterns develop with experience,[11] are affected by prior knowledge,[5] and deviate from textbook recommendations.[9] While some studies have suggested that misreadings (false negatives which range around 30%[12,13,14] may occur because large areas of film are not foveally viewed (viewed with the fovea area of the retina which affords acute vision),[15] or because there is nonuniform coverage of the film,[16] eyetracking experiments indicate that only about 30% of missed lung nodules can be attributed to the lesion not having been foveally viewed.[17] More recently, studies have been initiated to determine workstation requirements for interpreting single small radiographs.[18,19]

Image Quality and Contrast Adjustment

Any viable radiology workstation must have adequate image quality. To display an image, the system must have adequate values for screen size, number of pixels, gray scale dynamic range, number of digital intensity levels, and should be standardized so that a gray scale value produces the same luminance regardless of the screen on which it is displayed. Theoretical calculations[20] indicate that large radiographs require a sampling of 4000^2 16-bit pixels to capture their full resolution. Anecdotal reports by radiologists revealed that pneumothorax pleural lines on chest x-rays that have been digitized at 2000 lines were seen well only when the image was displayed at that resolution, and that these same pleural lines were discernible at lesser resolution only if edge enhancement processing were performed. Controlled studies,[21-23] have confirmed the need for at least 2000^2 pixel display of radiographs, and in some cases, 4000^2.

Contrast enhancement, or intensity windowing, creates an image focused on presenting a particular anatomy; it is almost essential for the electronic display of medical images.[24-26] Such contrast adjustment allows details of the very dense bone, and details of soft tissue to be available from the same slice. When intensity windowing an image, the radiologist chooses a window of a certain width from the entire total range of contrast in the CT or other digitally produced image and then places this window at a certain level in the total contrast range. If a given pixel value falls within this window, it is translated into a corresponding pixel value in the (typically) eight-bit framebuffer. Pixel values outside the window are set at zero or one, respectively. Image processing such as contrast-limited adaptive histogram equalization (CLAHE) and unsharp masking are also very useful. Pizer et al[4] and others[27] have provided overviews of image processing. Discussions of detailed topics are also found in Horii et al[28] and others.[29-31]

Multiple film images frequently must be made from the same digital slice in order to allow for viewing the image at multiple contrast levels. One critical advantage of a workstation over film is dynamic contrast adjustment, that is, the ability to change dynamically the window width and level (WWL) used to display an image.

While radiologists need the ability to select all possible intensity windows, most often a selection from a small standard set of WWLs refined for bone, lung, soft tissue, brain, etc is preferred. With a press of a button, the radiologist can rapidly change the contrast range of a displayed image. Since multiple images, displaying intensity windows for bone, lung, etc required for film, are not necessary, fewer separate images are needed to convey the same information.[4]

Display Area, Size, and Response Time

Display area refers to the number of pixels that can be displayed simultaneously. It is difficult to determine how much display area is sufficient for a primary interpretation workstation. The four-over-four viewbox, with 8000 × 4000 pixels (eight 2000^2 radiographs) might be taken as a starting point, but even its large display area is often insufficient for very large patient folders containing many films. Eye-movement studies[6,7] indicates that most of the images placed on the viewbox serve as a pictorial index into the patient folder, and that far fewer images are actually needed for simultaneous viewing. But how much display area for a workstation is sufficient?

While roaming and zooming is possible, full resolution display of entire images is superior, based on the reduced number of hand motions required for viewing. Therefore, roughly 2048^2 monitors (the minimum resolution for large radiographs[4]) are necessary, and at least two monitors are needed if multiple films are to be compared. Larger monitors, if available, would be superior. For tasks other than primary interpretation, or when dealing with smaller images, 1024^2 monitors are required. Based on the findings of Beard et al,[6] we believe that a display area capable of showing at least 12 CT images is the minimum required for multiple CT study interpretations. Carboni et al[19] considers the possibility of comparing images displayed in sequence rather than side by side. Sequential comparison does reduce by 50% the amount of required display area, but may produce an unacceptable cognitive load on the radiologist.

Display size refers to the total physical area of all the display screens. It is limited by the maximum distance at which a radiologist can adequately view an image, and by radiologist

head and body movement limitations. We believe the four-over-four viewbox at about 3 × 5 inch represents a reasonable bound on the total square footage than can be used for display without strain.

Our observations[32,33] indicate that for common operations such as roaming or moving to the next slice in a CT study, a response time of 1 second is acceptable, with improvements down to a half second or so having an effect on total interpretation time. Consistency in response time may be more critical than absolute speed.[34]

FILMPLANE

Design Objectives

Radiology workstation design can affect interpretation time and quality.[32,35] But producing a workstation with acceptable navigation is difficult. The typical viewbox array can easily display over 64 Mbyte of images which the physician can quickly access using techniques—the movement of eyes and head—which have been practiced for an entire professional career. Even the best workstations barely have this much real memory, let alone display area. Further, the cognitive motor skill requirements of the viewing task create additional technical complexities for workstation design.

We concentrated on several critical navigational tasks when developing and evaluating FilmPlane[32,33]. Image location for locating any image in the patient folder, arbitrary access for displaying and viewing the images, and sequential access for moving through slices and for roaming within a radiograph. Side-by-side comparison is also needed to compare various images from one study with images from another, or for comparing one side of a chest radiograph with the other. Our goal is to develop a workstation prototype upon which a radiologist can conduct an interpretation with quality equal to that of film and viewbox, and in equal or less time. We feel that the radiologist should be able to conduct a simple interpretation after only 5 to 10 minutes of training. Developing a radiology workstation viable for the primary interpretation is very difficult, and while the FilmPlane design looks promising, it may still prove insufficient for the task, particularly for viewing and comparing images from many different multiple cross-section studies. Controlled subject experiments are underway to determine FilmPlane's viability.

The original version of FilmPlane used a single 1024 × 900 pixel display. FilmPlane2 operates on a Sun 3/60 with a single greyscale monitor, and we are modifying it to operate on a Stellar high performance graphics workstation using three 1280 × 1024 displays. The current design can be adapted quickly to operate with a varying number of 1024^2, 2048^2, or even larger displays.

Mental Model

A key objective of CHI is for the user to obtain an understanding of the computer tool by developing a mental model of its operation.[36-42] Such a mental model is critical, for it not only allows users to quickly perform tasks, but it also helps them feel more "in control" of the computer. Developing and teaching an arbitrary mental model of a complex tool is difficult and frustrating. As an alternative, a mental model of a real-world tool, such as a desk top or the radiologist's viewbox, can be used as a metaphor for the operation of the computer tool. Thus, as long as the computer tool stays within the bounds of the metaphor, the user will, a priori, understand its operation. Basing computer-tool operation on a metaphor is particularly useful when designing for noncomputer specialists, especially when that tool must be used after only minimal training. Note that multiple metaphors may be viable for a given application, and that the presence of a metaphor is no guarantee of a successful interface.

FilmPlane's operation is based on the metaphor of a two-dimensional navigation view upon which all the images in the patient folder are systematically arrayed in two, four, or more column strips (Figure 2).[32] In the navigation
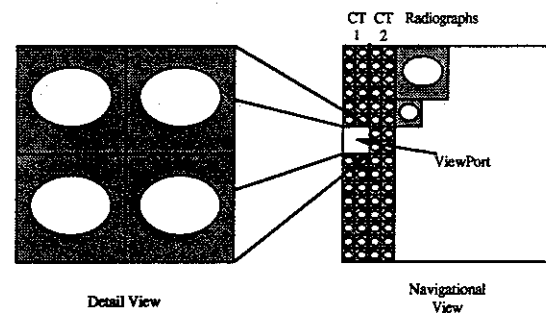

Fig 2. Navigation view, detail view, and viewport.

view, these columns are arranged horizontally and ordered by time. A full-screen image of this navigation view is provided to the user to reinforce the metaphor, with the medical images displayed at a fraction of their full resolution. Times and dates for each study are provided. The navigation view can either be permanently displayed or only appear on command. The permanent display of the navigation view reinforces its metaphor to the user and eliminates the cognitive load and handmotions required to make it appear and disappear. On the other hand, permanent display takes up limited screen space. Contrast enhancement, such as CLAHE,[4] appears to improve navigation-view or pictorial-index viability.

Direct-Manipulation Interfaces

Interfaces with a direct manipulation style,[43] are based on the mental model of a desk top and use a very simple one-place verb grammar to simplify the interaction and improve user performance. With such a grammer, most common statements are of the form ⟨subject⟩⟨one-place-verb⟩, in which the subject is indicated by being "grabbed" or selected by the mouse, and the one-place-verb to be applied to the subject is often selected from the pull-down menu. With such an interface, one would, for example, delete an object by selecting it with the mouse and applying the one-place verb "cut."

Computer interfaces with a large number of states can be confusing. Direct manipulation interfaces such as the Macintosh[44] are often said to be mode-less or state-less. This is incorrect because the Macintosh has a large number of state variables that control the font, size of the text, etc. But state control on the Macintosh and similar interfaces is not a problem. First, critical states, eg which object on the screen is selected, are represented visually. Second, state variables

are orthogonal, that is, one state variable can be changed without affecting another. Thus, with the Macintosh, one can modify font without affecting text size. Since state variables are typically represented visually, direct manipulation interfaces can be said to have orthogonal-visual state control. We have adapted this state control strategy to FilmPlane2.

Navigation

Superimposed on the FilmPlane navigational view are one or more wire-framed boxes called viewports, each corresponding to one of the workstation's display screens (Fig 2). Each viewport is sized to contain enough miniature images to fill the corresponding display screen when displayed at full resolution. (One small radiograph, four CT, or 16 MR images are needed to fill the current 1024^2 display screen.) The position of the viewports can be controlled by the mouse (or some other pointing device) and any display screen can be toggled to either display the navigation view or the full resolution images, which when displayed are called the detail view.

Taken together, the navigation view, the movable viewports, and the ability to display the images in the viewports at full resolution provide two-dimensional pictorial image-location and arbitrary access to all the images in the patient folder. Radiologists can quickly locate any image in the patient folder by scanning the navigation view. They can quickly display any arbitrarily selected image by moving the cursor to a viewport, selecting the viewport by depressing the mouse button, dragging the viewport to a new location, and releasing the mouse button. These actions cause the images in the viewport to be displayed at full resolution on the corresponding display (Fig 3).

Two alternatives to a pictorial image index have been constructed by others. The Arizona
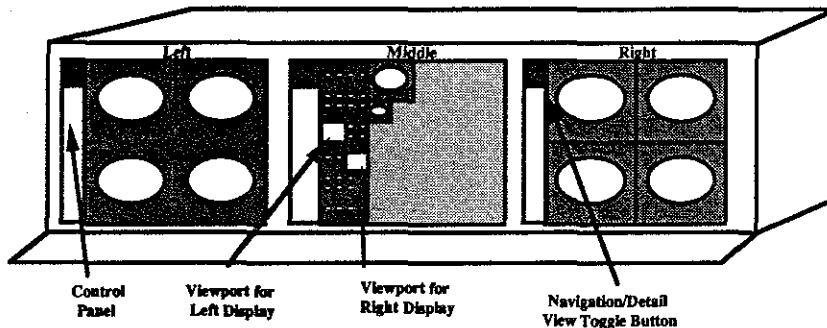

Control Panel Viewport for Left Display Viewport for Right Display Navigation/Detail View Toggle Button

Fig 3. Any of the three displays can be toggled to display either the detail or the navigation view.

workstation[45] provides image location using a text listing of all the studies in the patient folder and images in each study. Text image indexing is useful for locating studies consisting of a few large images, each having clear text labels such as "chest, 20Jul89." However, we suspect that for experienced radiologists accessing several multi-slice CT studies containing 40 or more slices, no text label will work as well as the images themselves in indicating contents.

A second alternative to a pictorial index is an icon index[18] in which a simple icon of the body is used to represent and access the contents of the patient folder. An experiment described in the above citation indicates that icon indexing may be superior to text indexing. While no comparison between pictorial and icon indexing has been undertaken, we speculate that an icon index would be superior to a pictorial index, such as FilmPlane's navigation view, for inexperienced users who can not instantly determine the modality and represented anatomy of a medical image. But for typical experienced radiologists, we would anticipate the navigational view to be advantageous (Fig 4). Not only do the miniature medical images themselves serve as icons denoting general represented anatomy, but relative time ordering and the exact size and coverage of each study are available at a glance.

Radiologists hold their fingers on viewbox-displayed images to provide a visual and tactile reference[6] thus speeding access. FilmPlane allows radiologists to quickly mark critical images, causing a bright border to be displayed. Such markers not only allow the radiologist to locate critical images more quickly, but also tell the computer system which images are important to the radiologist allowing improved memory management. Some radiologists have also suggested adding a tool for moving viewports quickly to one of several standard locations within a typical multi-slice study.

Radiologists often wish to view images near the ones currently being viewed, either by roaming around on a radiograph that is too large to completely display at full resolution, or by sequentially moving up or down through a multiple-slice CT or MR study. In particular, it is important that the interaction be optimized for long strings of sequential movement in the same direction; this often occurs when the radiologist is sequentially viewing a multiple-slice study. To facilitate sequential movement, FilmPlane2 provides the user with a simple mouse-controlled mechanism for moving the viewport sequentially up, down, left or right through a multi-slice study or radiograph.[33]

Sequential movement in a radiology workstation can be either continuous or discrete, and either aligned or unaligned. With continuous sequential movement, either the viewport, or the underlying FilmPlane appears to scroll smoothly in the direction of sequential movement. Continuous sequential movement or scrolling enforces the mental model or metaphor with the user by animating the interaction to show direction. With discrete sequential movement, the next set of images are simply displayed with no apparent motion; some other direction indicator, such as an arrow, is required to reinforce the metaphor. Discrete movement is easier to implement and may require less costly hardware for the same level of performance. With aligned movement, the viewport always "snaps to a grid" on the navigation view corresponding to image borders. Aligned movement eliminates the need for the radiologist to carefully align the viewport with the borders of an image after a move operation. FilmPlane uses aligned discrete sequential movement with CT and MR studies because users want to see the maximum number of whole images. With FilmPlane, a sequential operation moves only one-half the width or height of the viewport, allowing any image to be simultaneously viewed with its neighbor. It remains to be seen whether aligned discrete sequential movement is viable for viewing of large radiographs.

Zooming, the ability to minify or magnify the images in the patient folder, provides two functions: first, minification allows more images to be displayed simultaneously, though at lower resolution. Second, magnification allows an image to be viewed larger, and possibly at higher resolution, if it originally was displayed at less than full resolution. While zooming may be occasionally useful, it adds complexity to the interaction. FilmPlane is designed to minimize the need for zooming. Images in the detail view are displayed at full resolution, and display screens are large enough to eliminate the need for zooming merely to increase the size of an image. FilmPlane does provide a tool that allows the images in a study to be displayed at half resolution.
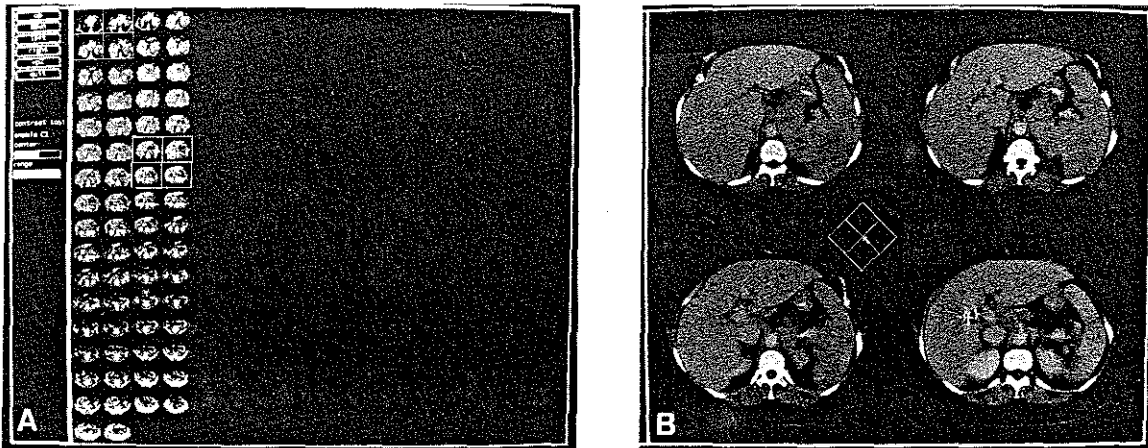
Fig 4.   FilmPlane alternate display screens. (A) Navigational view, (B) detail view.

Comparing Two Studies

Radiologists often need to compare two or more studies in a patient folder to complete an interpretation. With multiple slice studies such as CT, the radiologist typically views two studies side by side comparing corresponding slices of the studies, sequentially moving through both studies at the same time. This is one of the most difficult tasks performed on a radiology workstation. Not only are many handmotions involved, but a great deal of judgement and effort is required to insure that the same anatomy from the two studies is compared. Multiple cross-section studies may have slices taken at different intervals requiring, for example, 10 slices from one study to be compared with six from another.

To do a side by side comparison between two studies with FilmPlane, the radiologist places a viewport for one display screen at the beginning of one study, and the viewport of a second display screen over the second study. Then by sequentially moving each viewport downward, the corresponding anatomical regions of the studies are compared. A viewport may be split in two and each half independently navigated. While this method does allow side by side comparison of two multiple slice studies, it requires a considerable number of hand motions; the radiologist must first move the cursor to one viewpoint and sequentially move it down, then move it to the second and move it, and then back to the first. To compare two studies each containing 40 slices would require almost 3 minutes just to complete the required handmotions.[33] To reduce the re-

quired number of handmotions, FilmPlane2 allows the radiologist to connect viewports; the radiologist need only move one viewport to move the other. Synchronization is indicated by a white line connecting the two viewports in the navigation view. The synchronized viewports will automatically track the same anatomy because FilmPlane takes into account the slice intervals of the studies.

Visually comparing images from different studies is difficult and time consuming. Tools that quickly and accurately measure the length, area, and volume of anatomical objects in multiple-slice studies would be a very beneficial alternative. Several new techniques allow reconstruction of multiple slice CT or MR studies into three dimensional pictures of anatomical features.[46,47] These three-dimensional pictures appear to be particularly useful for understanding complex spatial structures such as intricate bone fractures. While current three-dimensional volume rendering techniques do not produce sufficient image quality to serve as the only view of a study's images during the interpretation, such renderings can serve as an important auxiliary role, and eventually should be incorporated into a viable two-dimensional image-viewing workstation.

Implementation

FilmPlane2's interaction design, allowing any screen to access any portion of the image space, presents two difficult to achieve requirements: a very large image memory, and the ability to

move rapidly any image from that memory to any screen's framebuffer. Radiology workstation memory often needs to contain over 60 Mbytes of images; cost prohibits using this much main memory. We have found[32] that virtual memory can be an effective alternative when three requirements are met: first, there must be sufficient main memory to store a minimum of 16 Mbyte for a single screen workstation. Second, data transfer between secondary storage and the framebuffer must be very fast. Third, there must be function to anticipate which images will be diplayed next, and to insure these images are prefetched into main memory from disk. We suspect that in 5 or more years, PACS image management will be implemented with very high speed networks such as HSC (High Speed Channel) or BISDN (Broadband Integrated Services Digital Network) providing a sustained 50 Mpbs to each framebuffer in the workstation from a very high-speed central archive.[48]

For the above reasons, many current commercial radiology workstations use separate image memories for each display screen; thus a multi-screen workstation is simple concatenation of several single screen consoles. While this approach does reduce costs, users must think of each screen as a separately manipulated image memory, adding to the computer human interaction complexity.

RAPID EVALUATION OF RWS NAVIGATION DURING DEVELOPMENT

We have utilized two response-time experiment protocols (described below) as well as time-motion analysis to quickly compare and evaluate workstation navigation. They are designed for a very rapid evaluation of a navigation technique or tool and only require a minimal amount of radiologist time. Both experiment protocols use patient folders containing several multiple-slice studies, a very challenging navigation task. These experiments are described in more detail elsewhere.[32,33] They are presented here briefly to illustrate how fairly simple, but still formal experiments, can speed the development of viable interactions.

Mental Model Experiment Protocol

The first experiment protocol determines the effectiveness of a workstation's arbitrary move-

ment and pictorial image index versus film and viewbox, by comparing response times for radiologists locating anatomical targets or landmarks in four patient folders, each containing two abdomen CT studies.

METHODS AND MATERIALS

FilmPlane1, an early single-screen prototype workstation was evaluated with this rapid experimental protocol. The single viewport could be split in two and each half independently navigated allowing side by side comparison of images. FilmPlane1 was implemented on a Sun 3/180 with a 1024² greyscale monitor and 16 Mbyte of main memory. This hardware was able to complete a sequential movement operation in about 0.75 seconds. Dynamic contrast window adjustment was available. Films were reviewed with a four-over-four viewbox in an environment controlled for light and sound. The single contrast window setting, typically used with abdomen CT, was picked by the technologist for each study and used for both film and workstation display. The CT images displayed on the Sun greyscale CRT were determined acceptable during a pilot study. Three radiologists participated, each having several years experience with CT. They all had limited experience with wordprocessors and two had previously used a mouse. The radiologists were given about 10 minutes training with FilmPlane1 as well as sufficient time to become familiar with the layout of the viewbox controls. The radiologist without mouse experience was given an additional 5 minutes training with the mouse. After training, the radiologists appeared to both understand and feel comfortable with the workstation.

For each trial, radiologists were asked to locate a CT slice containing a particular small body feature or "landmark," such as the pancreatic duct, in one of the two abdomen CT studies in each patient folder. The experimenter stated the landmark to be located, which study was to be searched, and then gave the start signal. Trials always began when all needed films were mounted on the viewbox or loaded into FilmPlane. The radiologists were instructed to work as quickly as possible without making errors. Time for each trial ended when the radiologist displayed an image and pointed to the "landmark." Each subject repeatedly located targets in each study of each patient folder using both film and FilmPlane. Sessions with film and FilmPlane were separated by time and alternated in order to control for bias and learning. The independent variable was the use of film and FilmPlane. The dependent variable was error-free response time. Very few errors were noted.

RESULTS AND CONCLUSIONS

Radiologists averaged 19.1 seconds (SD 16.6) to locate a target slice using FilmPlane1, and 9.4 seconds (SD 6.5) with film and viewbox. Film and viewbox was significantly faster (0.975 confidence). Several points were noted: first, the FilmPlane1 mental model worked, in that it allowed the radiologist to locate the required images. Analysis of the search patterns (automat-

ically recorded by FilmPlane) indicated that for the most part, radiologists moved directly to the target slice. This indicates that the tiny images in the navigation view (50^2 pixels) were sufficient for locating and accessing images. Analysis of the video tape showed very few incorrect actions and the radiologists' verbal protocol, (the comments they made during the trials) indicated that they were comfortable with the interaction.

Failure to locate an object, particularly due to a confusing or slow radiology workstation, would be very frustrating, so the fact that radiologists were consistently able to locate and successfully display the landmarks shows a fundamental soundness in the design. The workstation sequential operation was felt by radiologists to be too slow. Based on these results, subsequent Film-Plane prototypes were modified to reduce response time for landmark location. While landmark location was not a problem for experienced radiologists, it is possible that referring physicians, beginning residents, and medical students might have difficulty using a pictorial index such as FilmPlane's navigation view, if they cannot determine quickly the anatomy represented in a slice. Either another type of image index should be optionally available, or some tool should be provided to move automatically the viewports to one of several standard anatomical locations.

Because successful slice location requires the user to comprehend both a workstation's mental model and navigation methods, this rapid experiment protocol may be appropriate for evaluating a radiologist's general comprehension of a workstation. While the only formal dependent variable is response time, the other observations mentioned above must be taken into consideration before one can assume that workstation design is viable.

Task Time Experiment Protocol

The second experiment protocol determines the time to interpret four patient folders, each containing a current and a previous abdomen CT study, using film and FilmPlane. No prior interpretation report, patient history, or requisition form is provided.

MATERIAL AND METHODS

We used this protocol to evaluate FilmPlane1, using the same viewbox, subjects, and the experimental environment.

For each trial, the radiologists were asked to look through both abdomen CT studies and "indicate anything they could about the patient." This task utilizes image access patterns similar to a standard diagnostic task without prior knowledge. Radiologists were instructed to "insure their interpretation quality was up to typical professional standards." Tasks were timed from the moment the films were loaded onto the viewbox, or displayed on the navigation view, until the end of the interpretation. The time to load and unload the films on the viewbox was also measured. In addition to verbal protocol, observation data were gathered manually, by computer and with video tape. Radiologists were asked to describe how confident they were in their interpretations. During two separate sessions, each radiologist read two studies using FilmPlane, and two studies using film and viewbox. Presentation method and patient folder orders were systematically varied to control for learning and bias.

RESULTS

FilmPlane1 averaged 482 seconds (SD 131) for interpretation of these abdomen studies. An interpretation with film and viewbox averaged 264 seconds (SD 94). Film and viewbox were significantly faster (0.995 confidence). In general, the following pattern was used to review the two CT studies: first, the radiologist sequentially scanned all the images in the older study, and second a differential comparison of the old and the new study was made by splitting the screen into two viewports, moving the halves to the top of each study, and then moving the half viewports down through both studies, comparing the corresponding images side by side. During this comparison, only two slices from each study could be simultaneously viewed. This is insufficient, and either more display area is required, or the images should be viewed at lower resolution. We suspect that had the requisition form and/or the interpretation report for the previous study been present, the initial sequential scan through the older study would have been replaced with a highly directed viewing of the anatomical area in question.

The radiologists felt that they were able to conduct viable interpretations with the Film-Plane1 workstation for patient folders containing two abdomen CT studied. However, they also felt the system had "imposed itself" on them during the interaction. That is, they had found themselves thinking about manipulating the workstation rather than thinking about interpreting the images. The radiologists had only the training and experience from experiment one, so a certain unfamiliarity was expected. Nevertheless, this

initial version of FilmPlane had too much "friction" in the interaction. To correct this, the workstation interaction has been refined using the time-motion calculations described in the next section. In addition, the single 1024^2 available for the experiment did not appear to have sufficient display area for comparing two CT studies at full resolution. On the average, each radiologist made 3.3 arbitrary access moves and 69 sequential moves when viewing each patient folder, indicating the importance of reducing the time for sequential movement operations.

Time-Motion Analysis Method

Once a general workstation design has been accepted, the design must be "lapped" and improved to reduce "friction." Besides observation and user comments, one of the best methods for rapidly polishing an interaction is with the use of time-motion analysis calculations. In brief, time motion analysis is used to estimate how long a user will require to perform a specific task using a specific tool.[49] Time-motion analysis of a small aspect of a computer tool can be as simple as counting keystrokes or as complex as a detailed keystroke analysis[50] that takes into account the mental pauses a user makes, even when doing a complex but well learned task.

After completing the above experiments, we used the keystroke model to compare several alternative interactions for sequential movement in FilmPlane2.[33] While the keystroke model was designed for analysis of experts and our users are novices, it is still valid for comparing two design alternatives. To verify the accuracy of the time-motion analysis model, we used it to calculate the duration of a sequential movement operation with FilmPlane1, and then conducted several simple timing studies of this operation. Generally, the keystroke estimates were within 0.1 second of the actual sequential-movement duration. These keystroke estimates showed that the old sequential operation took 2.7 seconds while the improved sequential operation would take about 1.6 seconds. Additional analysis showed that only 47 sequential operations would be needed with the new interface to complete the same task requiring 69 operations with the old. These changes would result in a total potential improvement of about 110 seconds out of the 482 average for the task in experiment 2. Together

with several other improvements, this new sequential access method should cut interpretation time by almost one third for patient folders containing two CT abdomen studies and no previous information. Considering the weeks of time required to conduct even a simple experiment, time-motion analysis, even with its limited accuracy, is as an important tool in CHI development. Further, the analysis comes before the prototype has been implemented, further reducing the time for obtaining feedback on a design idea.

DISCUSSION

In this paper, we have described the issues and problems with radiology workstation navigation design and described the objectives and design of the FilmPlane2 radiology workstation as well as our human computer interaction strategy. We have also presented three rapid evaluation techniques for quickly gaining feedback on a radiology workstation navigation design. Radiology workstations are the essential components to the acceptence of electronic medical image management, and adequate workstation computer-human interaction is essential to workstation acceptance.

In general, radiologists are currently not willing to work without film using a computer workstation, even if all modalities are available on-line. This is understandable for the information bandwidth of even the best worksations, including FilmPlane, is still insufficient. This seems particularly true for interpretations involving multiple studies. Analogously, as scientists, engineers, and toolsmiths, we are still unwilling to work without paper when using our computer workstations. Radiologists have become understandably skeptical of success claims after repeatedly hearing of new workstations and constantly finding them unacceptable.

Nevertheless, there is a light at the end of the PACS tunnel. Relatively low cost 2048^2 display monitors of acceptable quality are available , as are workstation busses that can move 60 million 32 bit pixels a second. All this, in our opinion, adds up to sufficient hardware to construct viable radiology workstations. Now can begin the process of developing, evaluating, evolving, and lapping the surfaces of a viable tool.

ACKNOWLEDGMENTS

REFERENCES

1. Mantei MM, Teorey TJ: Cost/benefit analysis for incorporating human factors in the software life cycle," CACM, (31)4:428-439, April 1988

2. Brooks FP: The computer "scientist" as toolsmith: Studies in interactive computer graphics, in Gilchrist B (ed): Information Processing, North-Holland, 1977, pp 625-634

3. Pizer SM, Johnston RE, Rogers DC, et al: "Effective presentation of medical images on an electronic display station." Radiographics 76:1267-1274, 1987

4. Pizer SM, Beard DV: Medial image workstation: State of science & technology. J Digital Imag 4:185-193, 1989

5. Kundel HL, Wright DJ: The influence of prior knowledge on visual search strategies during the viewing of chest radiographs. Radiology 93:315-320, 1969

6. Beard DV, Johnston RE, Toki O, et al: Search patterns for viewing multiple CT chest images: An eyetracking study. J Digital Imag 3:1990 (in press)

7. Rogers DC, Johnston RE, Hemminger BM, et al: "Development of and experience with a prototype medical image display," Farwest Image Perception Conference, University of New Mexico, Department of Radiology, 1986

8. Beard DV, Pizer SM, Rogers DC, et al: A prototype single-screen PACs console development using human computer interaction techniques. SPIE Med Imag 767:646-653, 1987

9. Gale AS, Worthington BS: The utility of scanning strategies in radiology, in Groner R, Menz C, Fisher DF, et al (eds): Eye Movements and Psychological Functions: International Views. Erlbaum, Hillsdale, NJ, 1983

10. Carmody DP, Nodine CF, Kundel HL: Global and segmented search for lung nodules of different edge gradients. Invest Radiol 15:224-233, 1980

11. Kundel HL, LaFollette PS: Visual search patterns and experience with radiological images. Radiology 103:523-528, 1972

12. Yerushalmy J: The statistical assessment of the variability in observer perception and description of roentgenographic pulmonary shadows. Radiol Clin North Am 1:381-392, 1969

13. Smith MJ: Error and Variation in Diagnostic Radiology, Springfield, IL, Thomas, 1967, p 148

14. Tuddenham WJ: Visual search, image organization, and reader error in roentgen diagnosis: Studies of the psychophysiology of roentgen image perception. Radiology 78:694-704, 1962

15. Llewellyn-Thomas E: Search behavior. Radiol Clin North Am 1:403-417, 1969

16. Tuddenham WJ, Calvert WP: Visual search patterns in roentgen diagnosis. Radiology 76:225-256, 1961

17. Kundel HL, Nodine CF, Carmody D: Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. Invest Radiol 13:1978

18. Rogers E, Goldberg M: Image organization and investigation strategies for a radiological workstation. SPIE Med Imaging III 1091:315-324, 1989

19. Carboni N, Tombaugh J, Dillon R: The effect of simultaneous vs sequential viewing of digital images on comparative judgement performance. SPIE Med Imaging III 1091:252-261, 1989

20. Pizer SM: Psychovisual issues in the display of medical images, in Hoehne, KH (ed): Pictorial Information Systems in Medicine, New York, NY, Springer-Verlag, pp 235-250

21. MacMahon H, Vyborny C, Metz C, et al: Digital radiography of subtle pulmonary abnormalities: An ROC study of the effect of pixel size on observer performance. Radiology 158:21-26, 1986

22. Chakraborty DP, Breatnach ES, et al: Digital and conventional chest imaging: A modified ROC study of observer performance using simulated nodules. Radiology 158:35-39, 1986

23. Goodman LR, Foley WD, et al: Digital and conventional chest images: Observer performance with film digital radiography system. Radiology 158:27-38, 1986

24. Judy PF, Swensson RG, Twible D, et al: Effects of display level on detectability of small lesions on CT images. Med Phys 10:525, 1982

25. Judy PF, Swensson RG: Detection of small focal lesions in CT images: Effects of reconstruction filters and visual display windows. Br J Radiol 58:137-145, 1985

26. Seeley G, Robles-Sotelo E, Cannon G, et al: The use of psychophysics as a system design aid: Comparison of film-screen to an electronic review console. SPIE Med Imaging 767:639-643, 1987

27. American Association of Radiology: American College of Radiology Engineering Research in Visual Perception. 1986

28. Horii S, Horii H, Kowalski P: Ecletric look at viewing station design, SPIE Med Imaging II 914:920-928, 1988

29. Pizer SM, Amburn EP, Austin JD, et al: Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing 4:335-368, 1987

30. Zimmerman JB: Effectiveness of adaptive contrast enhancement. Doctoral thesis, University of North Carolina, Chapel Hill, NC, 1985

31. Zimmerman JB, Pizer SM, Staab E, et al: An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement. IEEE Trans Med Imaging 7:304-312, 1988

32. Beard DV, Cromartie R, Creasy J, et al: Experiment comparing image locating on film and the FILM-PLANE radiology workstation. SPIE Med Imaging II 914:933-937, 1988

33. Beard DV, Walker JGF, Bell I, et al: Evolved design of a radiology workstation using time-motion analysis and the keystroke model. SPIE Med Imaging III 1091:121-131, 1989

34. Shniederman B: Designing the User Interface, Reading, MA, Addison-Wesley, 1987

35. Foley WD, Jacobson DR, Taylor AJ, et al: Display of CT studies on a two-screen electronic workstation versus a film panel alternator: Sensitivity and efficiency among radiologists. Radiology 174:769-773, March 1990

36. Rumelhart D, Norman D: Analogical processes in learning, in Anderson JR (ed): Cognitive Skills and Their Acquisition, Hillsdale, NJ, Erlbaum, 1981, pp 335-359

37. Young R: The machine inside the machines: User's models of pocket calculators. Int J Man-Machine Stud 15:51-85, 1981

38. Foss D, Rosson M, Smith P: Reducing manual labor: An experimental analysis of learning aids for a text editor. Proceedings of the Gaithersburg Conference on Human Factors in Computing Systems, Washington, DC, National Bureau of Standards, 1982, pp 332-336

39. Halasz F, Moran TP: Analogy considered harmful. Proceedings of the Gaithersburg Conference on Human Factors in Computing Systems, Washington, DC, National Bureau of Standards, 1982

40. Douglas SK, Moran TP: Learning text editor semantics by analogy. Proceedings of the CHI '83 Conference on Human Factors in Computing Systems, Ann Arbor, MI, 1983

41. Dumas ST, Jones WP: A comparison of symbolic and spatial filing. Proceedings of the CHI '85 Conference on Human Factors in Computing Systems, Association for Computing Machinery, San Francisco, CA, 1985, 127-130

42. Olson JR, Carrol JM, Anderson N: What does the user of software know? Mental model and software human factors. Proceedings of the Workshop on Software Human Factors: Users Mental Models, 1987. Available from the Committee on Human Factors, National Research Console, 2101 Constitution Ave, NW, Washington DC 20418

43. Shneiderman B: Direct manipulation: A step beyond programming. IEEE Computer 16:57-69, 1983

44. Apple Computer, Inc, Macintosh System Software User's Guide, Cupertino, CA, 1988

45. McNeill K, Fisher H: A model for radiologic workstation user interface design. SPIE Med Imaging, 767:713-716, 1987

46. Fuchs H, Levoy M, Pizer SM: Interactive visualization of 3D medical data. IEEE Computer, August 1989

47. Levoy M: Display of surfaces from volume data. IEEE Computer Graphics and Applications, 8:29-37, 1988

48. Beard DV, Parrish D, Stevenson D: A cost analysis of film image management and four PACS systems based on four network protocols. J Digital Imag 3:108-118, 1990

49. Card SK, Moran TP, Newell A: The Psychology of Human-Computer Interaction. Hillsdale, NJ, Erlbaum, 1983

50. Card SK, Moran TP, Newell A: The keystroke-level model for user performance time with interactive systems. CACM, 23:296-410, 1980