

A Time-Motion Comparison of Several Radiology Workstation Designs and Film

David V. Beard, Brad M. Hemminger, Kevin M. Denelsbeck,
and Peter H. Brown

Departments of Radiology and Computer Science
University of North Carolina, Chapel Hill, NC, 27599-7510

1. INTRODUCTION

What are the relative effects of screen space and system response time on the speed of an accurate CT chest study interpretation? Screen space refers to the number of full resolution images of a given size that can be simultaneously displayed on a radiology workstation's screens. System response time refers to that portion of the total interpretation time during which the radiologist has to wait for the computer to complete an operation such as filling a monitor with images. This question is critical to the radiology workstation designer who must decide whether to add more or larger monitors, use a faster computer, or redesign the interface, and critical to the radiologist deciding whether a given workstation design will be viable in the clinic. Prototypes and observer studies are the ideal method of comparison. However, prototypes and observer studies can be very time consuming, and often even crude time motion models are sufficient to make reasonable design decisions.

We first describe time-motion analysis methods focusing on Card, Moran, and Newell's Keystroke and GOMS models. Second, we present a generic CT study and scenario for single CT study interpretation. Third, we describe the nine different image viewing systems, and fourth, we present the results of our models for these nine systems. Several systems are compared with observer study data. Fifth and last, we discuss the implementation of the model results for workstation design.

2. BACKGROUND

While time motion analysis techniques¹ have been around since the turn of the century for analyzing repetitive tasks requiring minimal mental effort, only recently have these techniques been expanded to include tasks that require choice in methodology. Card, Moran, and Newell's Keystroke² and GOMS³ (goals, operators, methods, and selection rules) models have proven to be fairly effective in describing experts using computer tools such as text editors to conduct decision-intensive tasks such as correcting a document. These models estimate the completion time for a task by summing the completion times of its subtasks which are in turn recursively analyzed until aprior tasks such as button presses are encountered. Task times for these atomic tasks can either be located in a table or determined by observation of users in-

teracting with miniature, highly focused interface mockups. Table 1 gives some times for several typical atomic tasks.

Atomic Task	Average
Press button	0.1 seconds
Type, per character	0.2 seconds
Point cursor at object with mouse	1.1 seconds
Move hand from keyboard to mouse	0.4 seconds

Table 1
Typical Task Times

An example may be helpful. Suppose a workstation can display a single CT image at a time and has an interface consisting of two buttons, one for scrolling forward and one for scrolling backwards through the study. We will assume a system response time of 0.5 seconds and an average of 41 slices in the CT study. If users wished to locate a slice containing a particular anatomical object, they would have to scroll backwards or forwards to move to the appropriate slice. We will assume that on average, the search will start at the middle of the study, and that on average, the users must scroll 10 times to locate the desired slice.

To scroll once, the user must press the appropriate button (0.1 seconds) and wait for the system to respond (0.5 seconds) resulting in 0.6 seconds per scroll for a total of 6 seconds for this ten-scroll task. What If the system response time for a single scroll was 3.5 seconds or 0.1 seconds? then the task time would be 36 seconds and 2 seconds respectively. (Those of you who have experience with time-motion analysis realize that the above discussion is highly simplified. For example, real users mentally pause to decide what operations to perform and how to perform them, and both mental pauses and physical movements like button presses can be concurrent with system response time. While the models used later in the paper included these considerations, we will avoid their discussion for the sake of simplicity and clarity.)

These time motion models are neither designed for, nor easily model information retrieval-and-analysis tasks – including medical image interpretation – for they cannot account for the time the radiologist will spend viewing the various images. Nevertheless, these models can account for the image manipulation time of a given workstation design and compare it against the image manipulation time of another. By assuming the the image viewing time will be constant across image viewing systems, we can use the models to roughly estimate task time for a system.

3. SINGLE CT CHEST INTERPRETATION

3.1 Generic Single CT Interpretation scenario

Comparisons of CT workstations are highly dependent on the task used in the analysis. For example if the CT study only contained eight slices, the analysis might un-

realistically favor a workstation with 2 1Kx1K monitors (holding eight full resolution CT images) and a 45 second scroll operation, as compared with a workstation with a single 1Kx1K monitor and a 0.1 second scroll operation.

For our analysis we have developed a generic 40 slice CT chest study with 30 slices containing lung, 30 containing soft tissue, and 10 containing liver. When printed on 14"x17" film, the lung is on three sheets of film, the soft tissue is on three, and the liver on a single sheet. This generic study is the average of 20 randomly selected CT chest studies from UNC Hospitals.

In addition to a generic CT study, we also need a generic image interpretation scenario detailing the exact number of scroll and other operations a typical radiologist will follow while interpreting our generic single CT study. We assume the following generic scenario: First, the radiologists load the films onto the workstation or alternator (times for pre loaded alternator included). Second, they overview the entire study starting with the lung, then the soft tissue, and finally the liver. Third, they focus on two interesting anatomical areas (each area involved 10 slices). Fourth, they dictate the report again viewing these two areas. Fifth and finally, they unload the films and place them back into the patient's folder (film only). This generic scenario was developed from observation of 25 CT chest interpretations involving 8 different radiologists. Obviously there are hundreds of different ways to read a CT chest study and different scenarios could result in different estimates for image manipulation times.

3.2 Image Viewing Systems Modeled

We constructed GOMS time-motion models for nine different real and hypothetical CT viewing systems using the above generic CT study and generic CT interpretation scenarios.

System #1 was the UNC FilmPlane radiology workstation, configured using a single 1Kx1K monitor with a 1.5 second scroll. For a "down" scroll operation the bottom two 512x512 images were moved to the top of the monitor and two new images displayed in their place. Details of the workstations design and observer evaluations are given in previous papers^{4,5}. FilmPlane uses a navigation view pictorial index to allow the radiologist to randomly access any image in the patient folder without having to sequentially scroll through a CT study. Pictorial indexes are useful for workstations with small screens and slow response times.

Radiologists often need to move their eyes back and forth over six to eight CT images in order to form an understanding of the 3D anatomy visualized⁴. With only a single 1Kx1K monitor displaying only four 512x512 images at the same time, there is a considerable amount of back-and-forth scrolling while the radiologist tries to visualize the anatomy while focusing on interesting anatomical areas. Our generic CT interpretation scenario does not take this additional scrolling into account. Therefore it may somewhat underestimate the image manipulation for this configuration.

Due to the X windowing system, FilmPlane can be configured with any number of 1Kx1K monitors. However as the number of monitors increase, there are more images that must be displayed during a scroll, and therefore the system response time increases. *System #2* used two 1Kx1K monitors with a response time of 2 seconds.

Foley⁶ conducted an observer experiment analyzing a commercial workstation with two 1Kx1K monitors and either a 7.5 second response time with eight 512x512 images or a 26 second response time with 32 256x256 images. *System #3* models such a workstation by modifying the FilmPlane model with two screens and an 8 second scroll time.

System #4 models a fast and simple CT display system called FilmStrip II we have recently developed with one 2.5Kx2K monitor (a Megascan) and a 0.11 second scroll time⁷. Filmstrip uses a very simple mental model of a three-wide vertical strip of CT images. In practice, the radiologist uses "up" and "down" buttons to scroll through the study. "Lung", "softtissue," and "liver" preset buttons are provided for fixed intensity windowing. No provision is made for dynamic windowing; as with film, we assume that a technologist has pre-windowed the study. Since the radiologist can scroll from one end of the study to the other in only a few seconds, a navigation view or pictorial index is not necessary.

FilmStrip II is quite expensive with the hardware alone costing at least \$40,000 in 1992. As a low cost alternative, Brad Hemminger has developed *System #5* called FilmStriplet, which uses two 1Kx1K monitors attached to a Sun Sparc II workstation with 64-128MB of memory. FilmStriplet uses a horizontal FilmStrip mental model similar to FilmStrip II. Because we only provide fixed preset intensity windows, we are able to scroll the workstation in under 0.1 seconds.

Systems #6-#9 involve the use of film and a horizontal alternator. Few radiologists use the upper four fixed lightboxes with reading a single CT study, so we assume that only the bottom row of four lightboxes is available. Most radiologists place their seat between either lightbox #1 and #2, or #3 and #4 and can therefore only view two boxes without having to move their chair so we include systems #6 and #7 which use only 2 boxes, as well as systems #8 and #9 with four boxes. Since moving the chair to switch to the other two boxes takes about the same amount of time as scrolling the alternator, there is almost no difference between using all four boxes on the bottom panel and keeping only two boxes lit.

Some radiology departments have their technologists pre-load the alternator with the day's cases making it difficult to decide whether a speed comparison of radiology workstation and film interpretation times should include the time to take the films out of the patient folder and mount them on the alternator, and the time to take the films off the alternator and place them back into the patient's folder. This is an important consideration because timing studies (see section 3.X) show the load/unload operation to take about 70 seconds or roughly 23% of the film interpretation time for

a single ordered CT chest study placed as the sole study in a patient folder. Therefore, film-viewing *Systems #6-#9* have modeled film interpretation with both two and four box alternators and with and without film load/unload time.

3.3 Model Results

1 Tool	2 # Screens	3 Screen	4 Scroll	5 Image Manip- ulation	6 Estimated Interpret. Time	7 Actual Time
1) FilmPlane	1	1Kx1K	1.5 sec	2:23	7:36	8:02
2) FilmPlane	2	1Kx1K	2.1 sec	1:58	6:48	-
3) SlowFilmPlane	2	1Kx1K	7.5 sec	3:21	8:11	-
4) FilmStripII	1	2Kx2K	0.12sec	0:28	5:18	5:39
5) FilmStriplet	2	1Kx1K	0.1sec	0:48	5:38	-
6) Film/Altw/oLoadUnld	2	-	1 sec	0:16	5:06	-
7) Film/Altw LoadUnld	2	-	1 sec	1:26	6:16	-
8) Film/Altw/oLoadUnld	4	-	1 sec	0:09	4:59	5:01
9) Film/Altw LoadUnld	4	-	1 sec	1:19	6:09	6:13

Table 2
Single CT study image manipulation time estimates
for workstations and film with estimated interpretation
times and actual interpretation times where available.

Table 2 contains the GOMS time-motion analysis results for the various image viewing methods described in section 3.2. System response time in the fourth column is for the scroll operations described in 3.2. It is important to remember that the image manipulation time given in the fifth column is not the time for a complete interpretation. For example, the time to interpret a CT study on film with an alternator using four boxes not counting the time to load and unload the film (#8 in table 2) is actually about five minutes. The time motion analysis (backed up by observation sessions) indicate that only about 10 seconds of that time is actually spend manipulating the alternator; the other 4 minutes and 50 seconds is spent looking at the images, dictating the report, and manipulating the dictation machine. If we make the very rough assumption that the same amount of time will be spent viewing images with an electronic workstation as with film and an alternator, we can add four minutes and 50 seconds to each of the image manipulation times given in table four column five and generate the estimated interpretation time given in column six.

In order to be able to judge the accuracy of the modeling data, column seven contains average interpretation times for those image viewing systems with available observer studies^{4,5,7}. Generally, the modeled times are roughly within 30 seconds of the ac-

tual times. While the slowFilmPlane (system #3) is similar to the Foley workstation⁶, it cannot be compared easily because Foley's workstation allowed the user to view either 8 512x512 images with a 7.5 second scroll or 32 256x256 images with a 24 second scroll and we do not know how often the radiologists chose each viewing method.

3.4 Discussion

It is important to understand the relationship between screen space (the number of CT images that can be simultaneously displayed) and system response time for a scroll operation. The more images that are displayed simultaneously, the fewer the number of scroll operations that will be necessary, but the more information that must be moved across the workstation's bus and into the various framebuffers, increasing the time for a scroll operation. Since there is a limit to the number of monitors radiologists can actually use without having to move their chairs, there comes a point after say three or so monitors, where it would be better to develop a faster system response time than to add additional monitors.

In the case of FilmPlane, having two monitors is clearly better because eyetracker studies⁴ indicate that radiologists often view clusters of six or more CT images; a single 1Kx1K monitor containing only four full-resolution CT images would require the radiologist to scroll back and forth to visualize a cluster, greatly increasing the number of scroll operations. On the other hand, adding a third monitor to FilmPlane would be counterproductive because two monitors can already show eight full resolution images (more images than in most clusters noted during the eyetracker study) the increase in system response time would not be offset by the additional screen space, and the third monitor would require radiologists to move their chairs.

The estimated interpretation times of systems #4 and #5 are of particular interest. We conducted an observer study⁷ in which four radiologists interpreted 10 single CT chest studies using film/alternator and FilmStrip II (system #4) with a single 2Kx2K monitor. As table 2 column 7 shows (systems #4, #8, and #9), the interpretation times for FilmStrip were about 30 seconds faster than film/alternator if load and unload time are included, and about 40 seconds slower if load and unload are ignored. Further, there was no detectable difference in interpretation accuracy of all dictated reports from both systems, and radiologists were confident in the image manipulation with the workstation and in the quality of their dictated reports. FilmStrip would make an acceptable CT workstation for primary interpretation of single studies. Unfortunately, FilmStrip requires \$40,000 in hardware, let alone software and marketing costs.

However, FilmStriplet's interpretation time estimate is almost the same as FilmStrip and FilmStriplet can be constructed with a \$15,000 commodity priced workstation containing 64MB - 128 MB of main memory and a 1Kx1K monitor. While we have not conducted formal observer studies of FilmStriplet, preliminary

observation indicates that a two screen FilmStriplet with two 1Kx1K monitors and 0.1 second scroll times has promise as a viable single CT interpretation system.

FilmStriplet does not facilitate dynamic intensity windowing, because such operations can not be performed quickly with 8-bit framebuffers and 12-bit image data. However the lack of dynamic intensity windowing may be an advantage; when using a radiology workstation, radiologists have a tendency to spend a lot of time (over 10 seconds per window setting) locating the perfect intensity window, while they almost always accept the intensity window the technologists set when a CT image was printed on film. We suspect that it will be more efficient for the technologist to continue to preset the intensity windows even if the radiologist plans to read from softcopy.

4. REFERENCES

1. Gilbreth FB *Motion Study*, New York, D. Van Nostrand, 1911.
2. Card SK, Moran TP, Newell A, The Keystroke-Level Model for user performance time with interactive systems. *Communications of the ACM*, 23, 396-410
3. Card SK, Moran TP, Newell A, *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum Associates, Hillsdale, NJ, London, 1983.
4. Beard DV Designing a Radiology Workstation: A Focus on Navigation During the Interpretation Task, *J. Digital Imaging*, 3(3) (Aug) 1990 152-163.
5. Beard DV, Perry JR, Muller K, Misra R, Brown PH, Hemminger BM, Johnston RE, Mauro M, Jaques P, and Schiebler M, Evaluation of Total Workstation CT Interpretation Quality: A Single-Screen Pilot Study, *SPIE Vol. 1446 Medical Imaging V: PACS Design and Evaluation*, San Jose, 1991.
6. Foley DW, 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*, Vol. 174, March 1990, 769-773.
7. Beard DV Speed and Accuracy Comparison of Film vs. Electronic Display, Poster presentation at the Association of University Radiologists, Chicago, 1992.