Effects of lag and frame rate on various tracking tasks

Steve Bryson
Computer Sciences Corporation
Applied Research Branch, Numerical Aerodynamics Simulation Systems Division
NASA Ames Research Center
MS T045-1
Moffett Field, Ca. 94035-1000
bryson@nas.nasa.gov

ABSTRACT

Virtual environments involve the user in an interactive three-dimensional computer generated environment. The methods of interaction typically involve direct manipulation of virtual objects via three-dimensional trackers. The tracking signal may be degraded in various ways, impacting the ability of the user to perform various tasks. This presentation will address the impact of two types of degradation in the tracking signal, lag (transport delay) and low frame rate. These degradations are common in existing virtual reality systems. While the impact of lag on human performance is comparatively well studied, the impact of low frame rate has not been widely studied. The impact of lag and low frame rate on two tasks will be compared and studied: Pursuit tracking and placing. The tasks will be studied in a two-dimensional context, eliminating ambiguities due to three-dimensional perception and display. Simple conclusions will be drawn that can serve as guidelines for developers designing interactive virtual environments. The relationship between these conclusions and theories of human performance will be briefly addressed.

1. INTRODUCTION

1.1. Purpose and motivation

Tracking tasks are critical to human performance in real-time interactive computer graphics applications. In realistic computer systems and applications, however, various distortions occur in both the tracking and the presentation of the resulting graphics [1][2][3][4]. The tracking devices may be inherently inaccurate, as well as subject to filtering and processing delays [1][2]. It is in the nature of computer graphics systems to introduce lags (or delays) and discrete sampling through animation frame rate into the visual feedback presented to the person performing the tracking task. It is well known that good visual feedback is critical in the performance of tracking [3], as well as picking and placing tasks. In this paper we will study two types of distortion that are introduced in this feedback: delays and sampling due to low frame rates. We shall study the effects of these distortions on human performance, arriving at simple rules stating what one can expect in particular situations. We shall be not be concerned in this paper with inaccuracies, filtering or other sources of distortion, though they will certainly have an impact on performance.

Tracking and picking and placing tasks are common in many applications, particularly in so-called virtual reality or virtual environment applications [5][6]. In virtual reality, a dominant interaction paradigm is “direct manipulation”, in which the user uses a three-dimensional tracking device, such as a Polhemus 3Space tracker, to “pick up” and move objects in the virtual environment directly to a desired location. In an ideal system the perceived position of the user’s hand in the virtual environment exactly matches the actual position of the user’s hand relative to the user’s head. This is a generalization to three dimensions of the common paradigm of using a mouse to pick up and move objects on a two-dimensional screen.

One can decompose the basic direct manipulation task described above into the placing of the hand cursor over the position of the object to be picked up, and then moving of the hand cursor to where the object is to be released. The ease with which the user can accurately place the hand cursor at the desired location will determine the ease with which the user can perform direct manipulation tasks in the virtual environment.

Another task which is often encountered in virtual reality applications is the task of tracking and sometimes catching a moving target. Sometimes the motion of the target will be predictable, sometimes it will be unpredictable. The accuracy with which the user can track the target is an important determinant of the ease with which the tracking task can be performed.

The results in this paper are to be considered as preliminary. The series of experiments run on each subject is rather thorough and time consuming. Only two subjects have completed the full experimental sequence. Though most of the results are consistent between the two subjects, testing on more subjects is required before the results in this paper can be considered confirmed.

In this paper the tracking and placing tasks will be studied independently. Those elements common to both experiments will be described in the remainder of this section. In both experiments, existing work relevant to the study will be sketched, the experimental protocol will be described, and the results of the experiments presented. Finally some brief comments will be made.

1.2. Definitions of lag and frame rate

As mentioned above, lag and frame rate are two common types of dynamic distortion in real time computer graphics systems. In this paper, lag refers to a time offset in the data stream from the tracker to the graphical image. There are several sources of lag in a typical computer system:

- **delays in the tracker signal** [1]
- **delays in communication between the tracker and the computer system**
- **delays due to computations required to process the tracker data**
- **delays due to graphical rendering**

The total lag will be the result of the combination of all these sources of lag. Lag in mouse-based real-time interactive systems can range from 0.01667 to 0.5 seconds. Lag in typical nontrivial virtual reality applications, which have additional lag introduced by three-dimensional tracking devices, range from 0.1 to 0.5 seconds. Systems with lags longer than about 0.5 seconds will not be considered real-time interactive.

The frame rate of a real-time interactive graphics system is the rate at which new frames of the computer graphical scene can be computed, rendered and displayed. This rate is also known as the animation frame rate. The frame rate is determined by the **frame time**, the time necessary to compute, render, and display a scene. The frame time will, of course, depend heavily on the nature of the application and the scene being rendered. Some applications require essentially no computation and very simple rendering, while others require considerable computation and rendering in response to user movement. There are several sources of low frame rate in typical computer graphics systems:

- **computation required to process tracker data**
- **computation of the effects of the user's movement**
- **graphical rendering time**
- **communication overhead in distributed systems**

The total frame time will result from a combination of these frame times. Typical frame times in current computer graphics systems range from 0.016667 seconds = frame rate of 60 Hz, to 0.5 seconds = frame rate of 2 Hz. As is the case with lag, we shall not consider systems with frame times greater than 0.5 seconds = 2 Hz.

1.3. The experimental environment

The studies of the impact of lag and frame rate in this paper are intended to shed insight on the performance of tasks in three-dimensional virtual environments. It is difficult, however, to isolate the effects of lag and frame rate from the effects of other distortions that are common in virtual environment systems. Three-dimensional trackers typically suffer from significant inaccuracies and limited range. The perception of three-dimensional relationships is critical for placing and tracking in three dimensions. Though virtual environments have made significant progress in providing three-dimensional displays, there are still significant problems in these displays and a great deal to be studied as to how effective they are. For these reasons we have chosen to study the impacts of lag and frame rate in a highly controlled two-dimensional environment taking advantage of mature tracking tech-

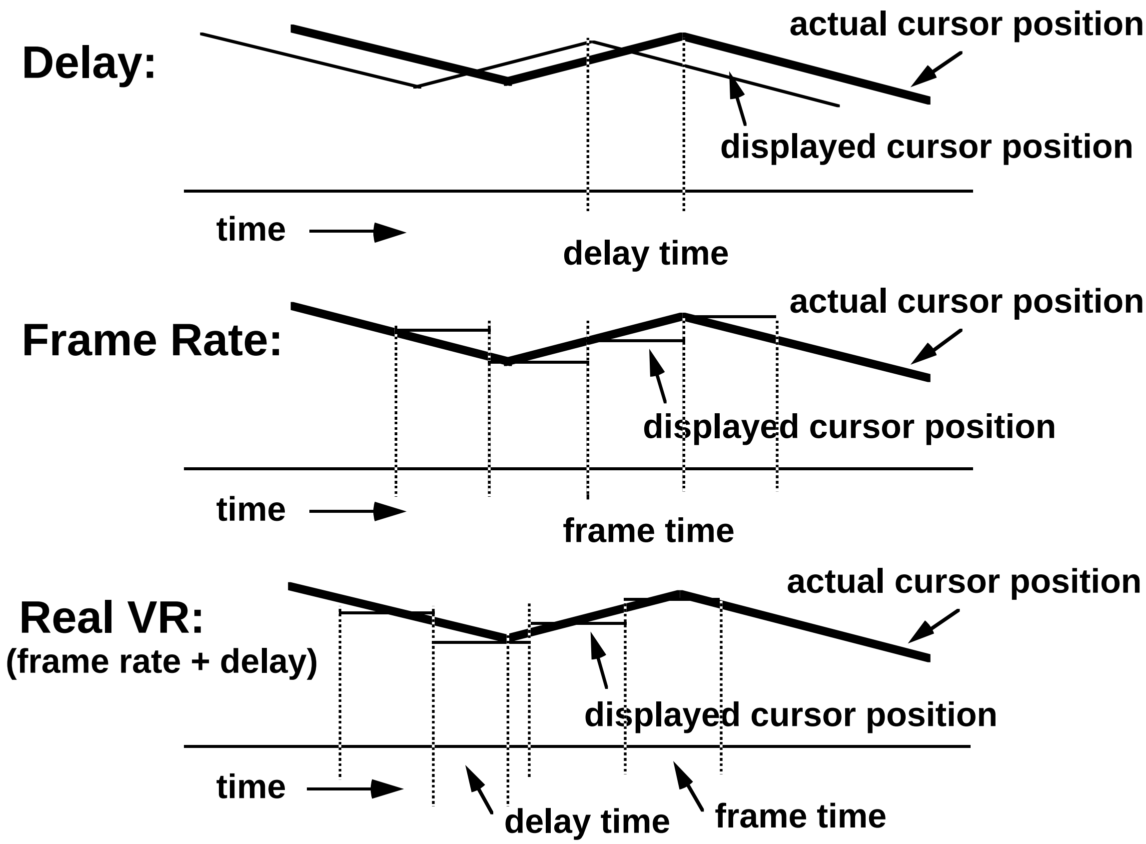

Figure 1: Timing diagram for cursor display with delay, frame rate, and combination.

nology in the form of a conventional mouse. Working in two dimensions eliminates uncertainties about issues of three-dimensional perception, while tracking with a mouse provides essentially perfect tracking which can be used for baseline studies and artificially degraded by introduction of lag.

The graphics platform used is a Macintosh IIci computer with conventional mouse input. This system was chosen for convenience, and because it provides a reliable scheduling mechanism which allows a routine to be executed at exactly 60 Hz. This Macintosh provides a graphics interface, and sufficient computational power to perform the required computations of both the target and the user's cursor within 0.01667 (=1/60) seconds. In all experiments, the background color of the screen was black. The experimental area of the screen is 24.13 x 17.145 cm, with a resolution of 640 x 440 pixels. The pitch of the monitor used is about 0.0377 cm/pixel. The user's cursor was a white 3x3 (0.113 x 0.113 cm) pixel square. The target was green, with a shape and size that depended on the experiment performed.

The experimental subjects in this experiment were experienced mouse users. One of the subjects was experienced with systems which had long delays and low frame rates, while the other was not. We wish to reiterate that the results reported in this paper are preliminary only, as only two subjects have been tested.

The 60 Hz interrupt cue was used to drive the experiment and artificially introduce lags and frame rates into the system. During each 60 Hz call, the following steps were performed:

- The target was processed in ways dependant on the experiment
- The current mouse position is read and placed into a user position history buffer
- For delay, an offset into the user position history buffer is used to provide a delayed user cursor position
- For frame rate, if a preset number of interrupt cycles have not passed since the last presentation of graphics, routine exits

- The old user cursor (and target when the target is moving) is erased from the screen
- The new user cursor (and target when the target is moving) is drawn on the screen

Thus at time t, the user's mouse movements at time t-Δt_D were displayed on the screen, where Δt_D is the lag and is equal to (delay offset/60) seconds. Every Δt_F seconds the screen is redrawn, where Δt_F is equal to (the number of interrupt cycles between frames)/60 seconds. The resulting timing with respect to user cursor position is shown in figure 1. A similar timing diagram is applies to the target motion in the tracking experiment. Most results quoted in this paper apply to the conditions of pure delay (Δt_F = 1) or pure frame rate (Δt_D = 0).

2. TRACKING

2.1. Background

The ability to track a continuously moving human target depends on both the amplitude and frequency content of the target motion. It is most convenient to describe the target motion in terms of its Fourier components, decomposing the motion into sinusoidal waves of fixed frequency, amplitude and phase. Assuming that the subject can position herself optimally, it can generally be stated that for sinusoidal oscillations with a fixed amplitude higher frequencies are more difficult to track than low frequencies [3]. The difficulty increases as the amplitude increases due to the accelerations required. Generally, the upper limit of the frequencies that can be tracked is roughly five Hz.

The impact of delays on tracking performance has been widely studied both in the context of computer graphics and telerobotics [3][4][7], and has been found to increase the tracking error at a fixed frequency. In particular, the falloff of accuracy with frequency has been shown to increase in the presence of lags. Tharp et al found a linear relationship between tracking error and time delays for delays in the 0 to 0.5 second range [4]. Measurements of the ability to keep a cursor on a target in the presence of a time delay indicate that the amount of time the cursor can be kept on the target decreases linearly with an increase in the time delay [8].

2.2. Tracking experiment design

In the tracking experiment, a target is moved in a path defined as sum of 10 sinusoids. The phases of the sinusoids are randomized from experiment to experiment so that the path of the target will be unpredictable. The target is drawn as a green 5x5 pixel (0.1885 x 0.1885 cm) square on the white background. The base frequencies of the sinusoidal components of the target motion range from 0.005 Hz to about 2.39 Hz. The amplitude of each component decreases rapidly with frequency as shown in figure 2. These base frequencies are multiplied by a factor ranging from 0.5 to 3. The amplitudes of each component is scaled so that the target motion fills the entire screen.

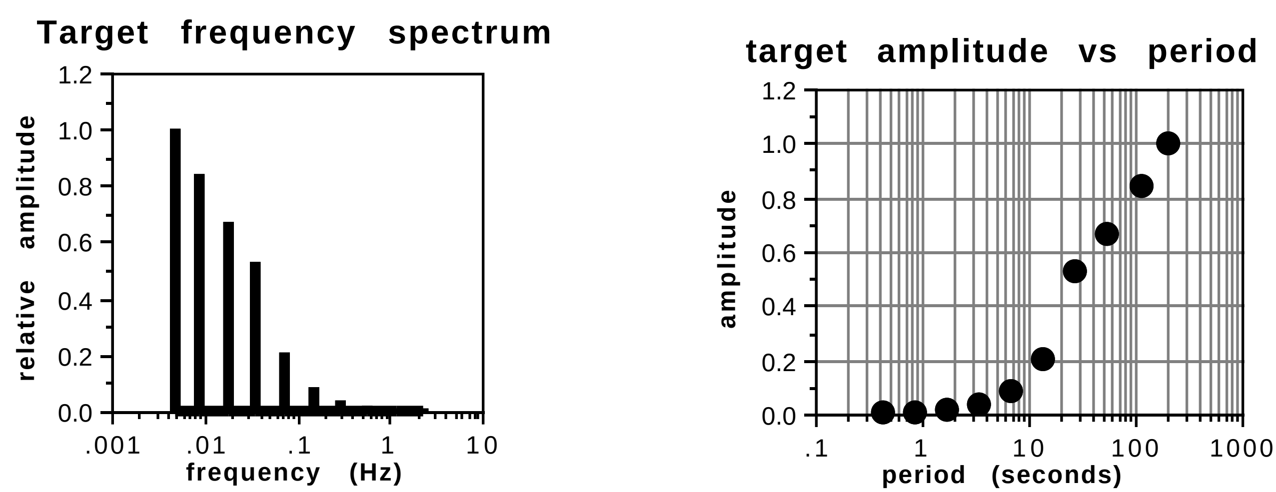

Figure 2: The basic power spectrum of the target motion in terms of frequency (left) and period (right).

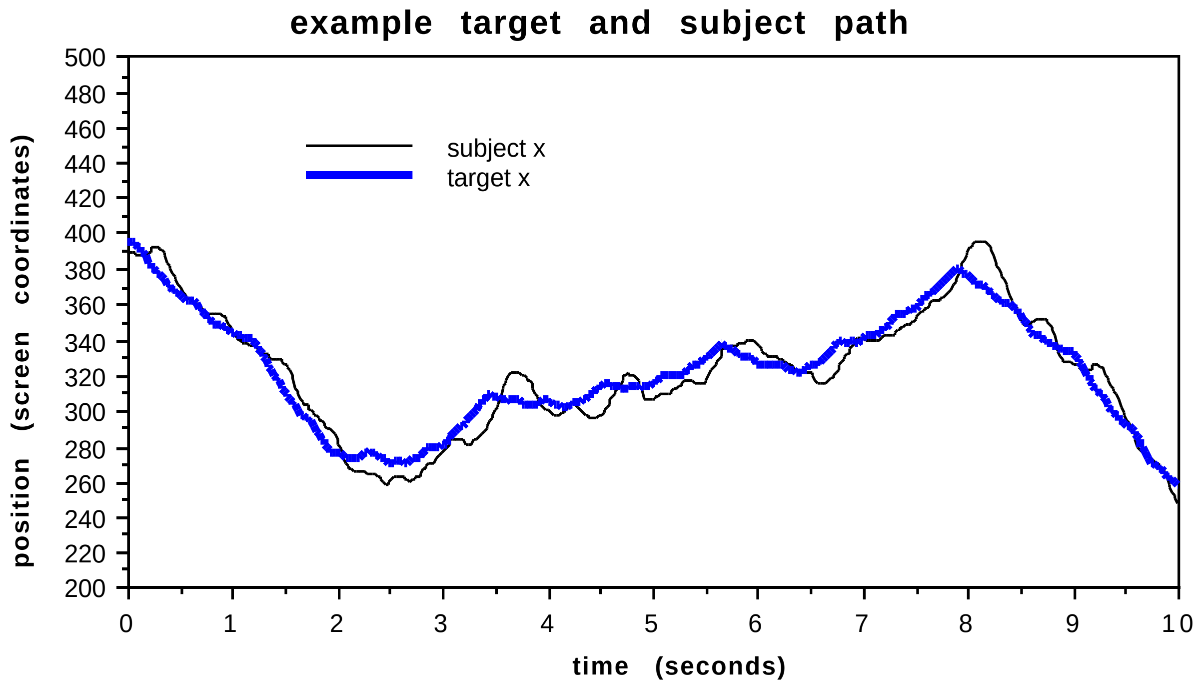

Figure 3: An excerpt from the recorded target and cursor positions over time for delay = frame time = 1/60 of a second.

Each experiment is run with a fixed value of either delay or frame time. The experimental subject tracks the target for 44.13 seconds. The first ten seconds of this tracking time are not recorded, allowing the subject time to acquire the target. Both the target position vector $\mathbf{T}_i$ and (delayed) cursor position vector $\mathbf{C}_i$ are recorded. In the framerate experiment the continuous motion of the target and cursor are recorded. A total of 2048 values of $\mathbf{T}_i$ and $\mathbf{C}_i$ are recorded. An excerpt of the target and cursor motion is shown in figure 3. The subject initiates each experiment, and is encouraged not to start if fatigued. The tracking error vector $\mathbf{E}_i$ is defined as the difference $\mathbf{E}_i = \mathbf{C}_i - \mathbf{T}_i$. The measure of tracking performance $\mathbf{E}_{rms}$ is defined as the root mean square value of the length of $\mathbf{E}_i$ over the time:

$$\mathbf{E}_{rms} = \sqrt{\sum_{i=1}^{N} \frac{|\mathbf{E}_i|^2}{N}}$$

For each run, $\mathbf{E}_{rms}$ is normalized by dividing by the amplitude of the target motion. This normalization is preferred to dividing by the rms distance from the center of motion often used [4], as randomization of the phases in the motion can change this normalization factor while $\mathbf{E}_{rms}$ is note observed to vary when the phases are changed. Using non-constant normalization factors mask the dependence of $\mathbf{E}_{rms}$ on delay or frame times.

A complete experimental run is defined as a tracking experiment for the following eight values of delay or frame time in the following order: Δt_F or Δt_D = (0.0167, 0.1, 0.05, 0.167, 0.25, 0.333, 0.033). In the case of the frame rate experiment, these values correspond to frame rates of (60, 10, 20, 2, 6, 4, 3, 30) frames per second. A complete experimental session consists of four runs, with the base frequencies of the target motion multiplied by the factor s =(0.5, 1, 2, 3). Three complete sessions were each performed on two subjects for both the delay and the frame rate experiment.

2.3 Tracking experiment results

The dependance of the error measure $\mathbf{E}_{rms}$ on the delay times for subject 1 is shown in figure 4 for all four values of the frequency multiplier s. It is readily apparent that $\mathbf{E}_{rms}$ is linearly dependant on the delay time in agreement with previously known results [4][8], and the slope of the linear dependance is in turn dependant on the frequency factor s. A similar plot is shown for the frame rate experiment in figure 5, and quantitatively similar results are apparent. Similar results were found for the second subject. Figure 6 shows all values of $\mathbf{E}_{rms}$ for s = 1 and s = 3 for both subjects, giving some indication of the amount of variation between subjects. The dependance of the slope of the least-squares linear fit to the data on the frequency factor for both types of experiments and both subjects is shown in figure 7. Within individuals, there is a strong linear correlation between the dependance of $\mathbf{E}_{rms}$ on the time delay or frame time and the frequency factor. When this dependence for both subjects and both experiments is considered together, there is still a fairly strong linear fit, implying that the phenomenon is stable across individuals and applies to both time delays and frame times.

Based on these observations, we conjecture that

- Frame time mimics delay in tracking tasks.
- Tracking error is linearly related to delay time or frame time.
- The slope of the linear dependance of tracking error on frequency factor is linearly related to that factor.

The implications of these observation include: In a system in which contains a task which requires manual position tracking of unpredictable targets, if the delay or frame time is increased by a factor D, then performance can be regained by dividing all target frequencies by the same factor D.

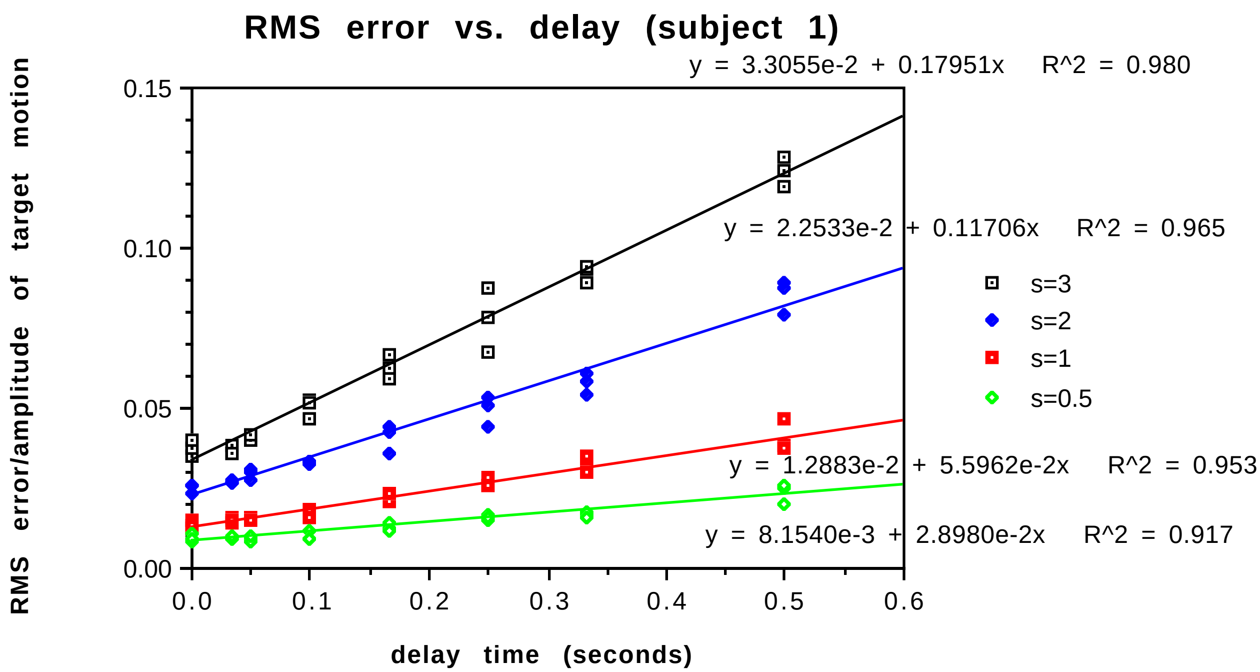

Figure 4: $\mathbf{E}_{rms}$ for various delay times and for various values of the frequency factor s for the first subject.

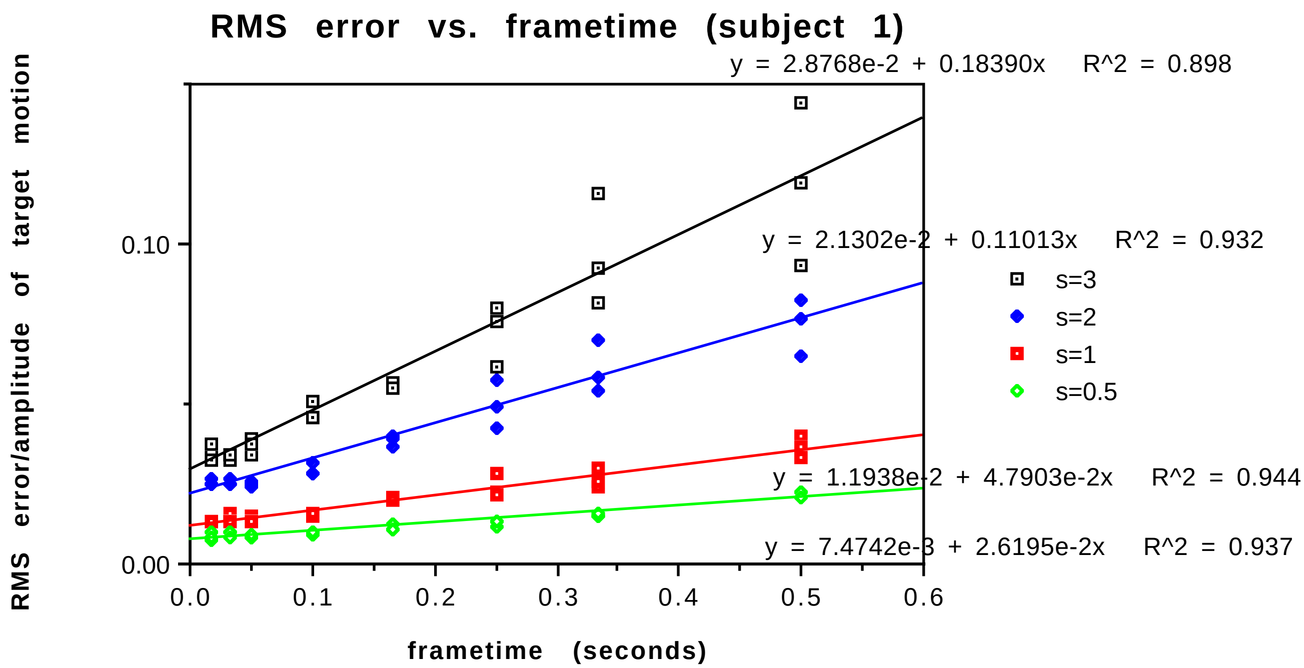

Figure 5: $\mathbf{E}_{rms}$ for various frame times and for various values of the frequency factor s for the first subject.

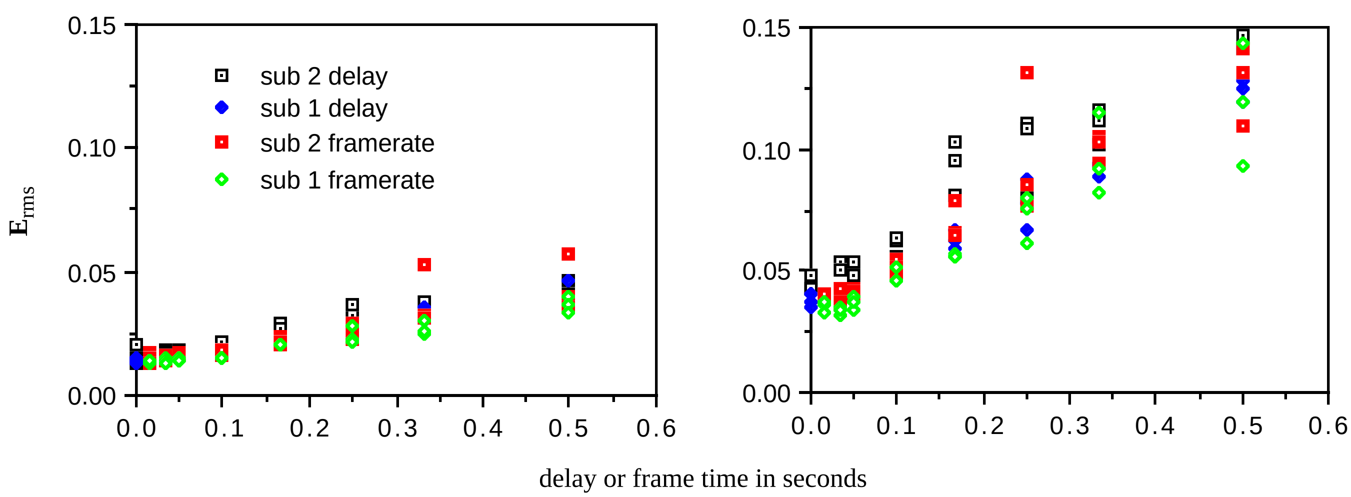

Figure 6: $\mathbf{E}_{rms}$ for both delay and frame times and for the values of the frequency factors (left) s=1 and (right) s=3 for the both subjects.

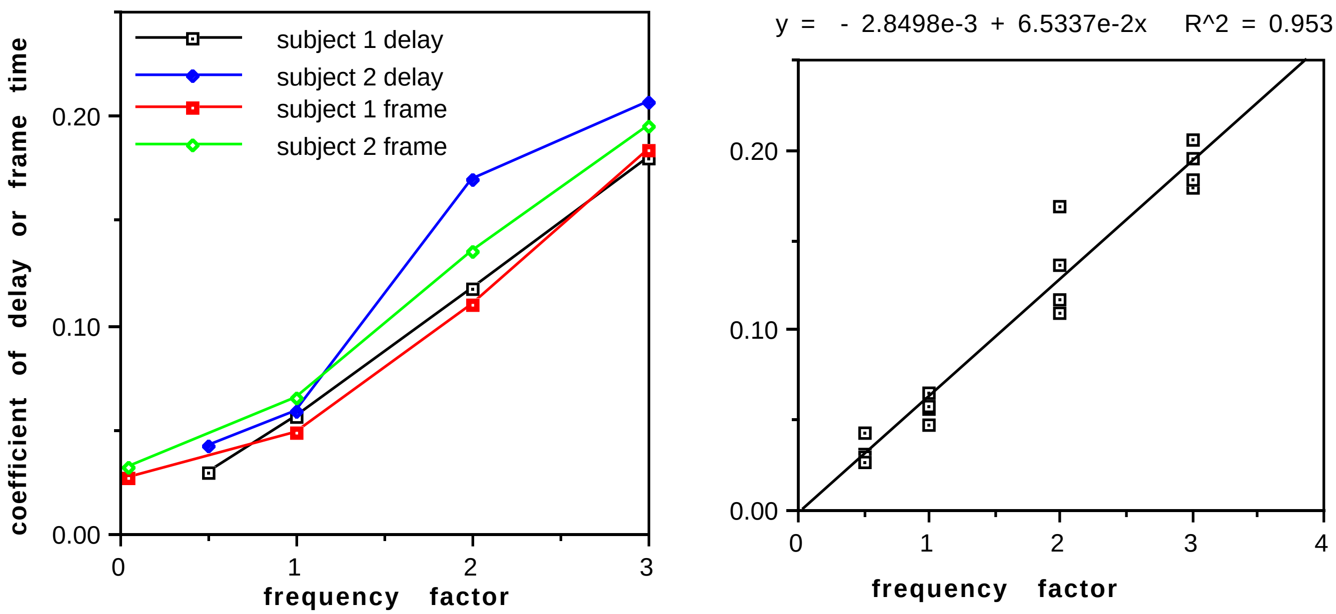

Figure 7: Plots of the slope of dependence of $\mathbf{E}_{rms}$ on delay or frame time vs. frequency factor s. Left: plots for both subjects for both delay and frame times. Right: linear fit of data for both experiments for both subjects.

3. Placing

3.1 Background

The basic manual position placing task is that of moving a cursor from an initial position to a final position. Performance of placing tasks is measured in the time required to perform the task and is determined by the distance from the initial to the final position and the accuracy with which the cursor must be placed. The classical placing experiment is to present a target and to require a subject to move a cursor from a standard initial position to the target [9][10]. Fitts has determined that human performance of this task is characterized by a simple formula. The time T required to move a cursor from an initial position to inside target at a distance D and width W is given by the formula $T = \mathbf{a} + \mathbf{b} \log_2(2D/W)$. This formula is known as Fitts' law. The coefficient **b** is observed to be about 0.1 when the cursor is controlled anthropocentrically [10], and the constant **a** measures reaction time.

3.2 Placing experiment design

The effect of lag and frame rate on placing tasks was measured by observing the effects of lag and frame rate on Fitts' law. The display timing and screen and cursor appearance were as described in section 1.3. The target is a 20x20 (0.754 x 0.754 cm) pixel square rendered in a green outline with a hollow center. The subject is instructed to place the cursor in a small 10 x 10 pixel (0.377 x 0.377 cm) square in the lower left corner of the screen and press the mouse button when ready. After between 0.8 to 1.5 seconds, the target is presented at a random location on the screen. The subject then places the cursor in the target. The cursor is considered placed in the target when it remains in the target for a full two seconds. After the task is completed, the distance of the target from the starting position is logged, along with the time from when the subject cursor left the starting square (after the target was presented) to the time the cursor entered the target square for the last time. This cycle was repeated 50 times for each value of delay or frame time for each of the two subjects. Thus in this experiment the coefficient **a** in Fitts' law is not measured, as it presumably is not affected by the experimental condition.

3.3 Placing experiment results

Two example data sets are shown in figures 8 and 9. Figure 8 shows the plot of the time required to perform the task for two values of the delay, 1/60 and 1/2 seconds. Figure 9 shows the same data for frame times = 1/60 and 1/2 seconds. It is readily seen that delay and frame time have similar effects on performance: they tend to increase the coefficient **b** in fitts' law and to introduce significant scatter into the data. This scatter implies that it is difficult to predict performance in the presence of long lag times or long frame times, except to say that performance will be worse. This scatter can be interpreted by the fact that for long lag or frame times the subject plans the placing motion, makes the motion, and sees feedback as to the success of the move. Occasionally the subject may make the initial move accurately, making the move very quickly, but most times the initial move will miss, requiring feedback to plan and perform the next move to the target. Sometimes several moves will be required to acquire the target, typically when the distance **D** is large.

These results are shown in more detail in figures 10 through 12. Figure 10 shows the dependence of the coefficient **b** on delay or frame time for both subjects. For delay, the effect is similar for both subjects: **b** stays around 0.1 to 0.2 until delay or frame times of greater than 0.25 seconds, when it increases somewhat dramatically. For frame time, one of the subjects showed similar behavior while the other subject showed no meaningful dependence of **b** on the delay or frame time. The latter subject is highly experienced with low frame rate systems, indicating that for the placing task people may adapt to low frame rates.

Figure 11 expresses the scatter of the performance data as the standard deviation of the data from the least-squares fit to Fitts' law for each of the subjects for both delay and frame time experiments. As expected from figures 9 and 10, the scatter increases with delay or frame time. The effects for delay are very similar for both subjects, while the effect for frame time is somewhat less severe for the subject experienced with low frame rate systems.

Figure 12 shows the predicted average time to reach the target vs. delay or frame time. This average is obtained by integrating the least-squares fit to Fitts' law over the distance on the screen for each value of delay or frame time for each subject. As expected, the average time increases with delay time for both subjects, while a similar effect is seen for frame rate for the subject less experiences with low frame rate systems. A less severe rise is observed for the subject experienced in low frame rate systems.

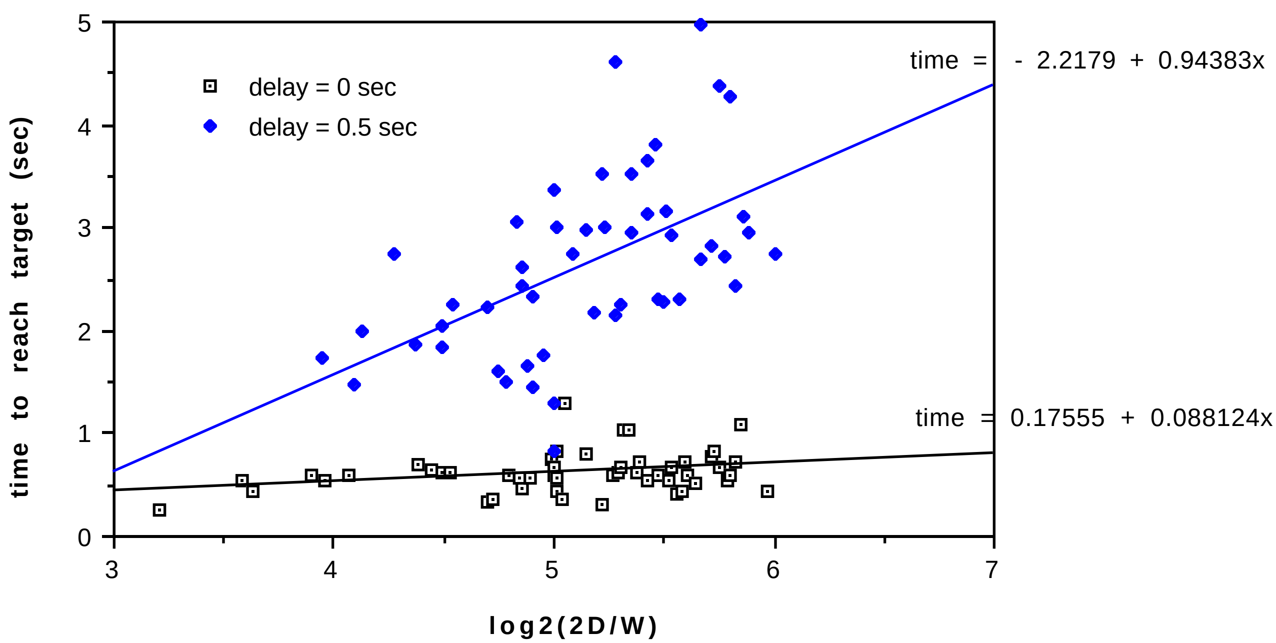

Figure 8: Plots of time to place cursor in target vs. $\log_2(2D/W)$ in placing experiment for delays = 0 and 0.5 seconds. Linear least-squares fits to the data show Fitts' law.

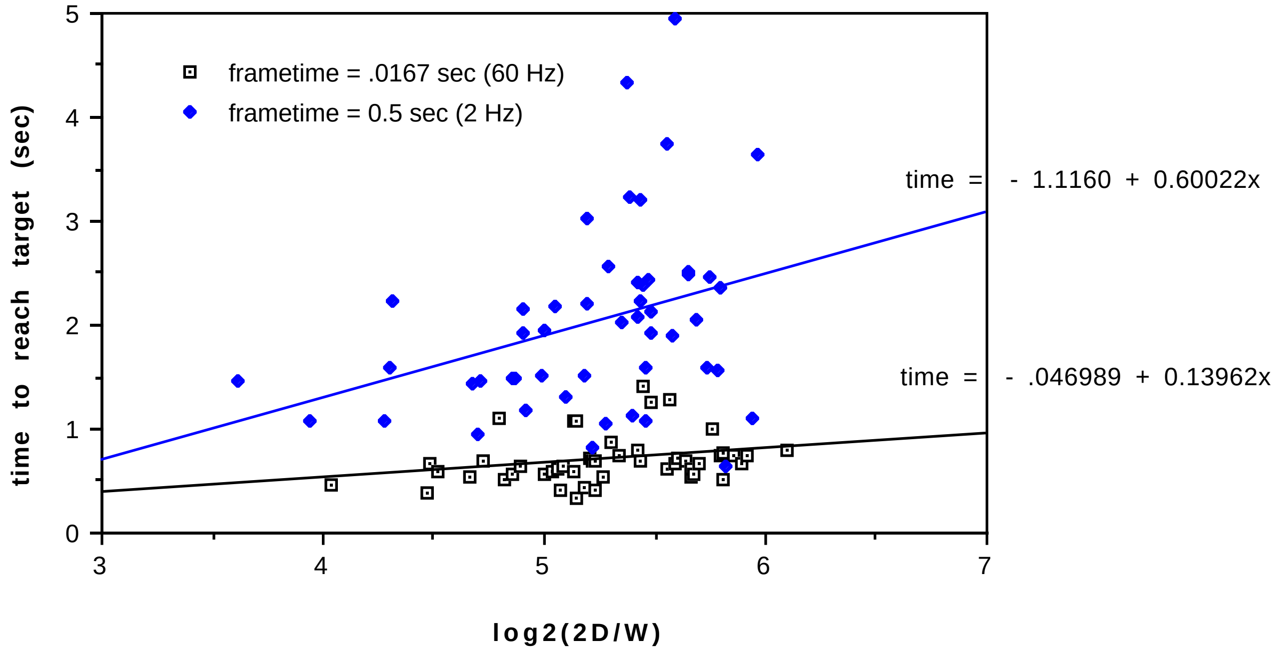

Figure 9: Plots of time to place cursor in target vs. $\log_2$(2D/W) in placing experiment for frame times = 0.01667 and 0.5 seconds. Linear least-squares fits to the data show Fitts' law.

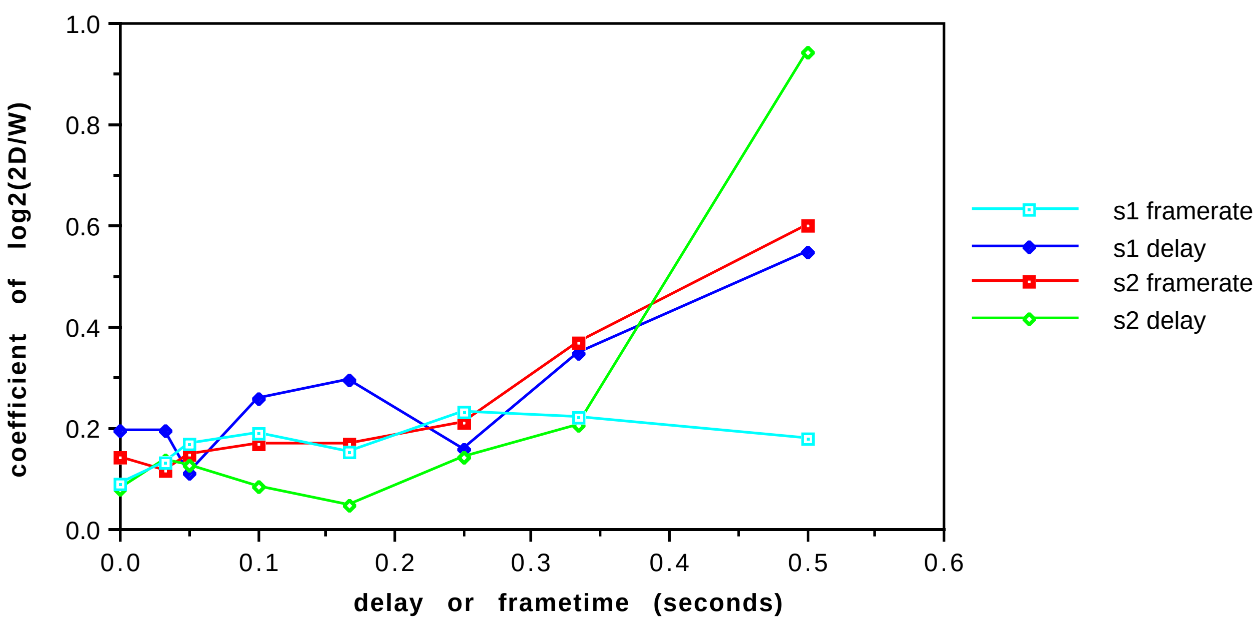

Figure 10: Plot of coefficient b in Fitts' law vs. delay or frame times for both subjects. Subject 1 is labeled s1 and subject 2 is labeled s2.

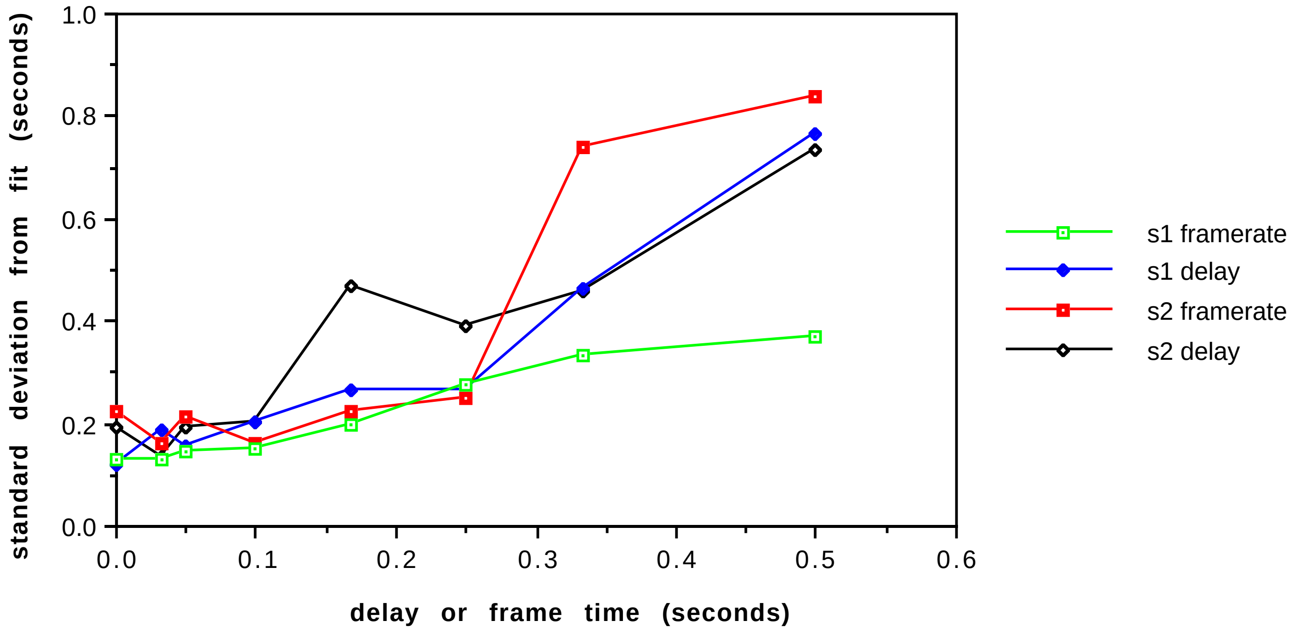

Figure 11: Scatter in placing experiment data measured as standard deviation of difference between data and linear least-squares fit to Fitts' law. Subject 1 is labeled s1 and subject 2 is labeled s2.

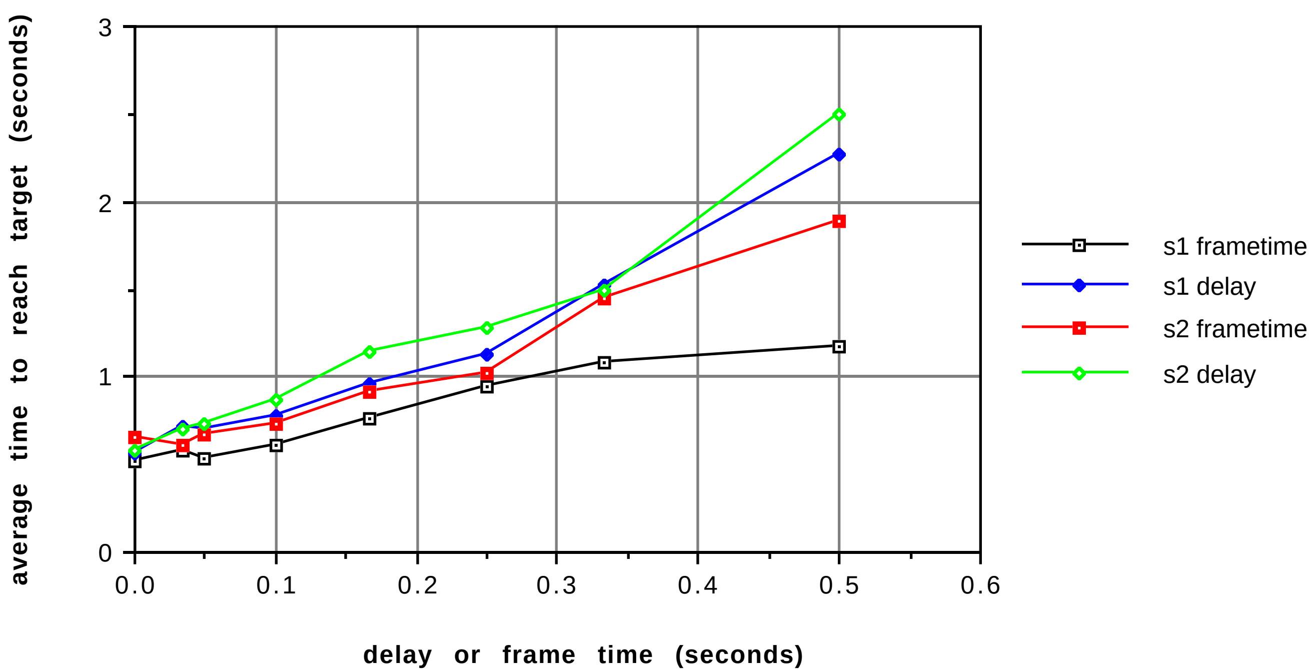

Figure 12: Average time to place cursor on target as predicted from Fitts' law for both delay and frame time for both subjects. The prediction is based on integration of least-squares fit of data to Fitts law for each experimental condition over the distance range.

4. Conclusions

The preliminary results reported in this paper indicate that with respect to tracking tasks, low frame rates induced by a frame time Δt_F has an impact quantitatively like the impact of a time delay or transport delay of the same time value. Tracking errors are linearly dependant on the delay or frame time for times in the 0 to 0.5 second range, and scaling the frequency linearly increases the slope of that linear dependence. We conjecture that the effects of delay and frame time will combine linearly when both are present. Delay and frame time also effect Fitts' law describing placing tasks, by increasing the scatter of that law. For inexperienced users, delay or frame times increase the slope of the linear relationship in Fitts' law, but there is some evidence of long-term adaptation.

Further data collection to verify these results is necessary, both for more individuals and more data for each individuals.

Acknowledgments

The author would like to thank Dov Adelstein and Steve Ellis for useful discussions and comments. He would also like to thank Vivian Tong and Jeff McIver for performing the experiments.

References

1 Adelstein, B. D., Johnson, E. R, and Ellis, S. R., A Testbed for Characterizing Dynamic Response of Virtual Environment Spatial Sensors, *Proceedings of Fifth Annual ACM Symposium on User Interface Software and Technology*, Monterrey, Ca., Nov. 1992

2 Bryson, S., and Fisher, S. Defining, Modeling, and Measuring System Lag in Virtual Environments, *Proceedings of the 1990 SPIE Conference on Stereoscopic Displays and Applications*, Santa Clara, Ca. 1990

3 Sheridan, T. B. and Ferrill, W. R., *Man Machine Systems*, MIT Press, Cambridge, Ma. 1974

4 Tharp, G., Liu, A., French, L., Lai, S., and Stark, L., Timing Considerations of Helmet Mounted Display Performance, *Proceedings of the 1992 SPIE Conference on Human Vision, Visual Processing, and Digital Displays*, Santa Jose, Ca. Feb. 1992

5 Fisher, S. et. al., Virtual Environment Interface Workstations, *Proceedings of the Human Factors Society 32nd Annual Meeting*, Anaheim, Ca. 1988

6 Bryson, S., et. al. *Implementation of Immersive Virtual Environments*, SIGGRAPH '92 course notes #9, SIGGRAPH '92, Chicago, Ill, July 1992

7 Held, R., and Durlach, N., *Telepresense, time delay, and adaptation* in Ellis, S. R., Kaiser, M. K., and Grunwald, A., C., *Pictorial Communication in Real and Virtual Environments*, Taylor and Francis, Bristol Pa, 1991

8 Warrick, M. J., Effect of Transmission-Type Control Lags on Tracking Accuracy, AF-TR-5916, Wright-Patterson AFB, OH, Air Materiel Command DTIC No. AD630292, 1949

9 Fitts, P. M., The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement, *Journal of Experimental Psychology*, 47, 381-391, 1954

10 Card, S. K., English, W. K., and Burr, B., J., Evaluation of Mouse, Rate-Controlled Isometric Joystick, Step Keys, and Text Keys for Text Selection on a CRT, *Ergonomics*, vol. 21, no. 8, 601-613 1978