Kinematic/Dynamic Characteristics for Visual and Kinesthetic Virtual Environments

FINAL REPORT
Cooperative Agreement NCC 2-797

National Aeronautics and Space Administration
NASA-Ames Research Center
Moffett Field, California 94035

Compiled by
Michael R. Bortolussi

Report Authors
B. D. Adelstein
Miriam Gold

Submitted by:
Western Aerospace Laboratories, Inc.
16111 Mays Avenue
Monte Sereno, California 95030-4212
FINAL REPORT to Western Aerospace Laboratories

Bernard D. Adelstein

Summary

Work was carried out on two topics of principal importance to current progress in virtual environment research at NASA Ames and elsewhere.

The first topic was directed at maximizing the temporal dynamic response of visually presented virtual environments (VEs) through reorganization and optimization of system hardware and software. The final results of this portion of the work was a VE system in the Advanced Display and Spatial Perception Laboratory at NASA Ames capable of updating at 60 Hz (the maximum hardware refresh rate) with latencies approaching 30 msec. In the course of achieving this system performance, specialized hardware and software tools for measurement of VE latency and analytic models correlating update rate and latency for different system configurations were developed.

The second area of activity was the preliminary development and analysis of a novel kinematic architecture for three degree of freedom (dof) haptic interfaces—devices that provide force feedback for manipulative interaction with virtual and remote environments. An invention disclosure was filed on this work and a patent application is being pursued by NASA Ames.

Activities in these two areas are expanded upon below.


Inadequate dynamic response due to excessive latency and insufficient update rates remains a stumbling block to the implementation of effective virtual environment (VE) applications. While enhancements to computing capability have offered steady significant increases in specified polygon fill rates, improvements in model computation and drawing speed affect only a portion of the factors that determine overall VE update rates and latencies. Thus, these computer hardware improvements can have only limited impact on the VE dynamic response characteristics experienced by the user.

One approach to ameliorating poor VE dynamic response has been the introduction of prediction algorithms to compensate for the latency (i.e., delay) between input action and rendered output. While reductions in apparent time delay have been reported for a variety of prediction schemes, predictors increase noise and generate overshoot not originally present in the sensed input to the VE, and thus may degrade human operator performance. Furthermore, prediction does not benefit VE update rates—in fact, the added computational burden of a poorly implemented predictor can result in slower rates.

In this work, we have been able both to reduce latency and increase update rates as a consequence of examining the detailed timing characteristics of individual VE components and intercomponent communication. Significant improvements in dynamic response were achieved through a reorganization of system software and hardware that eliminates a number of “inefficiencies” typical to the configuration of most VE systems. The resulting performance enhancements do not entail heightened noise or overshoot. Though one cannot attain zero latency through the reorganization measures we describe below, the results of this work—shortening the actual time delay to be
compensated and making data available more frequently through quicker update rates—are essential to improved predictor implementations.

End-to-end latency for a given VE system pathway is the time elapsed from transduction of an event or action by the input device until the consequences of that action are first made available in the display. It is the sum of the series of latencies, or delay between input and output, for the individual elements comprising the data pathway. As noted earlier these elements include hardware components and their associated software as well as intercomponent communication. We further delineate internal latency to include only those elements housed physically inside the computer. In the case of the VE system described in this paper, internal latency is the time between completed receipt of tracker data at the computer’s input port and the instant the data begins to be issued at the video port for scanout on the display CRT.

Because individual hardware and software components can each update according to their own cycle timing and since there may be no synchronization between particular VE system components, the effective update rate could be considered nominally as the rate of the slowest component in the pathway of interest. We also recognize that the update rates associated with an individual component, or for the complete VE system, may vary with time. For example the computer polling of an external sensor may occur at a nonuniform rate. Furthermore, under the scheduling structure of the UNIX environment, software cycle times can be highly nondeterministic. Thus, we consider the average effective update rate to be a suitable measure of performance.

A series of hardware and software reconfigurations were considered that in the final implementation has significantly improved end-to-end latency and update rate in our UNIX based VE system. The significant steps (and the stage at which they were achieved) in the development of our current VE configuration (SPI/SPI+) were:

- Separation into separate tracker driver and application processes using shared memory data transfer (FOFT).
- Continuous mode tracker data transmission (FOFT).
- Multiple concurrent tracker driver processes (AST).
- Shared memory data age thresholding (SPI+).

Throughout each stage of this work we used objective, detailed performance measurements to guide subsequent development. Quantitative results for the best end-to-end latencies and effective update rates measured for each configuration are summarized in Figure 1.

Many of the results and techniques developed in this work are transferable to other VE systems. The measurement apparatus and methods can be easily replicated and applied to quantify VE behavior in any system with a spatial displacement tracker and available CRT display. The series of hardware and software configurations that we studied can be implemented on most UNIX based VE systems.
2. Three degree-of-freedom haptic interface for VEs and teleoperation

Haptic perception is the process through which we explore and evaluate the physical characteristics (e.g., size, weight, shape, stiffness, viscosity, temperature, etc.) of objects or fields (e.g., gravity) in our immediate surroundings. Haptic perception of mechanical characteristics involves the cognitive integration of sensory input from strain (length), strain rate (velocity), and force sensors in the muscles and joints, as well as normal and shear pressures from tactile sensors in the skin induced by direct interaction between our limbs and the environment. In the work described here, we are concerned with the type of mechanical inputs intended specifically to stimulate the muscle and joint sensors.

Haptic displays that are designed to stimulate the muscle senses combine both force-reflecting interface hardware and a computation engine. The interface hardware typically consists of a mechanical linkage in the form of a joystick or exoskeleton that joins the human operator to a source of mechanical power—either electromagnetic, electrohydraulic, or electropneumatic actuators. Because of this external coupling at the skin, a force reflecting interface that stimulates the human limb and muscle sense of physical dynamics, either by default or by design, also affects the tactile sensors in the skin. For virtual environments (VE), the computation engine governs the behavior of the actuators and linkage as a function of kinematic and force measurements from interface transducers, according to algorithms and equations that describe the models to be simulated. In telemanipulation, the computer still modulates interface behavior, but the model is now either replaced or augmented by transducer information from the remote site.
The growing interest in force reflecting interfaces—as display channels for VEs, as master hand controllers for full and micro scale telemanipulation, and as research tools for studying passive arm dynamics, haptic perception, and manual control—has prompted a broad array of haptic display designs. In nearly all current multi-dof haptic interfaces, the weight and inertia problems associated with serial direct drive configurations are alleviated to some degree by incorporating transmission elements between one or more motors and the point of attachment to (or grasp by) the human operator. Transmissions serve to transform rotary into rectilinear motion, reduce speed and multiply force (or vice versa), and transfer motion and force from one location in the linkage to another. By introducing transmissions of the various forms described below, actuators may be located remotely from the joints that they drive. In some joystick interfaces, transmission elements allow all actuators to be mounted on a common base link or ground, thereby reducing significantly the weight and inertia that must be carried and thus decreasing the power requirements and size for the actuators.

The haptic interface we developed is intended initially for manual control research in a three dimensional coordinated haptic-visual virtual environment. For work in a three dof haptic environment, in which the purpose of the interface is the display of mechanical dynamics explicitly for muscle sensory organs, we chose to develop a device capable of the minimum necessary number of dofs. These three dofs correspond to translational displacements (as seen in the immersing VE visual display) and forces at the human-machine interface and will not include orientation angles and torques. Thus, all interactions at the human-machine interface will occur through a single point where only forces, as opposed to arbitrary rigid body moments or couples, can be applied. By restricting the interface to three dofs, we reduce design and implementation complexity associated with higher dof devices.

Our haptic interface is a mechanical linkage that couples three degree-of-freedom (dof) translational displacements at the linkage endpoint to link rotations about three axes that are fixed with respect to a common base or ground link. By mounting rotary actuators at each of the three base axes, forces can be generated at the linkage endpoint causing the endpoint to move. Conversely, endpoint motion can be measured by transducers that directly sense rotation about each of the linkage’s three base axes. The haptic interface is sized so that its workspace conforms with the displacement region over which binocular stereo vision is a significant contributor to human visual depth perception. This depth range coincides with the normal extent of human arm motion in the midsagittal plane—i.e., the 60 cm (24 in) or so beginning from ~15 cm (6 in) in front of the nose up to full arm reach. Our intent is to have at minimum a 15 cm spherical well conditioned manual workspace centered in this region.

The mechanical linkage is a three dof, 10-link, 12 revolute-joint parallel mechanism, depicted as a kinematic model in Figure 2. The link lengths, link angles, and joint placements shown in Figure 2 are, in general, chosen for convenience of illustration and are not necessarily those that optimize the device’s workspace characteristics, structural properties, or manufacturability. In this illustration, rotary actuators and/or sensors, labeled A, B, and C, are attached to a common ground link (link 1, not labeled). These actuators can drive, and be backdriven, by the spherical grip D in three spatial degrees of freedom. Each rigid link (numbered 2 through 10, plus ground link 1) is paired to its neighbors by single dof rotary joints (represented by the wheels and hubs). In addition, links 5, 2, and 8 are extensions to the shafts of the motors A, B, and C respectively, and thus can use the motor shaft bearings for their individual joints with ground link 1. Thus all mechanism force and motion is transferred through transmissions made of rigid links and—ideally frictionless and backlash free—single dof rotary ball bearing joints.

Since this linkage consists solely of rigid link and revolute joint pairs and does not require gears, belts, cable, screw or other types of transmission elements, it is useful in applications requiring full backdrivability. Thus, this invention can serve as the mechanical linkage for actively powered
devices such as compliant robotic manipulators and force-reflecting handcontrollers, and passive devices such as manual input devices for computers and other systems.

Figure 2. Three degree of freedom parallel haptic interface hand controller linkage.
The Single Channel Processing Bottleneck in the Cockpit: Can It Be Bypassed?

Miri Gold, James C. Johnston and Robert S. McCann
NASA Ames Research Center, Moffett Field, CA

ABSTRACT

Recently, Johnston & Delgado (1993) showed that the single channel processing bottleneck in human information processing can be bypassed if one task is zero-order visual position tracking and the other requires classifying an auditory stimulus. Since most information in the cockpit is presented visually, we attempted to determine if such bypassing is also possible when the auditory stimulus is replaced by a visual stimulus, such as a HUD gauge. Results show partial success in achieving the desired level of parallel processing.

INTRODUCTION

The performance of a joint human-machine system, such as an aircraft, is limited by the information-processing capabilities of both partners. Since it is within our power to improve the design of machines but not of humans, over the long run the human limitations will be the most constraining. One particular imbalance in the cockpit is that onboard computers can carry out many tasks simultaneously, while pilots find it difficult to do so. If human multi-tasking could be facilitated, it is reasonable to expect 1) an increase in pilot work throughput, 2) a decrease in mean response time to stimuli that require quick action, and 3) a reduction in the likelihood that critical stimuli will not be noticed. One particular example has motivated our research: helping pilots landing an aircraft simultaneously adjust the yaw of the aircraft to align it with the runway, and adjust the airspeed when it is too high or too low.

Until the availability of Head-Up-Display (HUD) technology, there was little possibility of achieving this goal. Panel-mounted gauges are located so far off the center of the retina that limited visual acuity typically prevents them from being processed in parallel with out-the-window stimuli. HUDs ameliorate this obstacle by bringing gauge stimuli closer to the center of vision. While the use of HUDs may be a necessary condition for cockpit multi-task parallel processing, it is by no means sufficient. Limited acuity outside of central vision is only the first of numerous obstacles to parallel task performance. Recent research suggests numerous further obstacles including perceptual grouping, spatial attention, categorization and perceptual decision-making, response selection, motor programming, response initiation and response execution. Since the obstacles to parallel task performance are numerous and complex, such processing may, at best, be possible under limited circumstances.

In order to investigate parallel processing, it is necessary to measure it. Both accuracy and response time (RT) have been used as dependent measures. Accuracy measures have been popular, especially in investigations of relatively continuous tasks. In the aviation domain, accuracy measures such as RMSE (root-mean-square-error) have often been used with tracking tasks. Although these measures provide a useful practical performance metric, they provide little insight into parallel processing (Pashler & Johnston, 1989). Even if two continuous tasks are performed over the same extended time interval without loss of accuracy on either, it is not possible to tell how much parallel processing has occurred. Both tasks could be handled sequentially by a processor that switches back and forth between them while buffering stimuli and responses. Such a "buffer and switch" strategy need not produce any decrements in accuracy measures, but it would necessarily produce increases RTs to stimuli on one or both tasks. Hence RT measures are better suited for studying parallel processing. Particularly useful is the Psychological Refractory Period (PRP) paradigm (Pashler & Johnston, 1989; Welford, 1952). Two discrete stimulus-response mapping tasks are employed, with a variable Stimulus-Onset-Asynchrony (SOA) interval between the stimuli for the two tasks. At long SOA intervals, the tasks can be processed sequentially with no requirement for parallel processing. The critical data are provided by short SOA trials. RTs as short as those obtained at long SOAs can only be produced with high levels of parallel processing. Failure of parallel processing, reflecting interference between the two tasks, will show up as a slowing of RT on one or both tasks.

Such RT slowing at short SOAs, usually concentrated on the second task, has been repeatedly observed by hundreds of studies over five decades of research with the PRP paradigm (Pashler, 1994). PRP interference
has been documented across an unusually broad range of conditions, including the use of extremely easy tasks, and careful separation of stimulus and/or response modalities. Only a handful of exceptions have been reported (e.g. Greenwald, 1972; Johnston & Delgado, 1993; McLeod & Posner, 1984). This evidence for interference in dual-task processing led to conceptions of the human operator as a "single-channel" device (Welford, 1952), capable of working on only one task at a time. Alternatively, it has been proposed (e.g. Kahneman, 1973) that humans have "limited processing capacity" which can be flexibly allocated among tasks. According to this model, dual-task slowing occurs because reductions in the processing capacity available to a task reduce processing rate on that task. More recently, strong evidence has been advanced for a modified single-channel theory hypothesizing that some early and late stages of processing can be carried out in parallel on more than one task, but one or more central stages constitute a single-channel bottleneck (De Jong, 1993; McCann & Johnston, 1992; Johnston, McCann & Remington, 1994; Pashler & Johnston, 1989). Evidence has also been provided for several distinct bottlenecks at different processing loci (De Jong, 1993; Johnston, McCann & Remington, in press).

Since finding interference-free parallel processing in the PRP paradigm is rare, it is advantageous to build on a known exception. Johnston & Delgado (1993) found only very slight interference in a PRP design with a first task of tone discrimination (with a verbal response) and a second task that required using a joystick to keep a circle over a horizontally moving cross stimulus. Note that the visual tracking task provides a rough analog to controlling the yaw of a plane. Low interference was found using simple, zero-order (position) control dynamics and a step-tracking approximation of normal continuous tracking. This approximation was used to ensure that discrete momentary stimuli-- jumps in cross position-- could be related to discrete responses, allowing RT measurement. Since our goal is to promote parallel processing of tasks that require visual instrument monitoring, we wanted to see if low interference could be observed with two visual tasks. Specifically, we attempt here to measure whether there are gauges that can be responded to in parallel with stick tracking. In Experiment 1, to provide a baseline when visual-visual interference is absent, we first replicate the Johnston & Delgado (1993) paradigm, but reversing the task priority. It appears that using tracking as the primary task more closely corresponds to the priority pilots give to controlling the attitude of the plane. In Experiments 2-4 the stimulus in the secondary task was a visual gauge, using a different gauge design in each experiment.

In selecting designs for gauges to be tested, we anticipated that one major problem would be incompatible demands of spatial attention for the two tasks. Our first decision was to discard alphanumeric gauges as candidates. Although used extensively in modern glass cockpits, they are unlikely to satisfy our needs. Character perception falls off rapidly from the center of vision, especially with adjacent characters, so acuity appears to be a barrier to processing digital stimuli while the pilot is fixated elsewhere. In addition, the spatial attention demands of characters appear to be severe (cf. Yantis & Johnston, 1990; McCann, Folk & Johnston, 1992). Hence we settled on analog gauges as the most likely to facilitate parallel processing (in practical applications redundant digital display of variables like speed would be advisable). Two different tasks offer some promise in avoiding attentional obstacles to parallel gauge processing. The first tack, followed in Experiments 2 and 4, is to try to make use of simple stimulus properties for which there is already evidence for parallel processing from other paradigms. Work in visual search has found evidence for parallel processing of properties such as line orientation, color, and line terminations (cf. Julesz, 1981; Treisman, 1985). Although this evidence was obtained with rather different tasks than those used here, it would seem worthwhile to try to design gauges using properties that under other conditions seem to be processed (at least to some level) in parallel. The second tack, followed in Experiment 3, is to take advantage of the finding that there is often parallel processing of several properties of the same object. That suggests trying to integrate properties of the two tasks into the same object. This task is followed in a variety of current HUD interfaces. The drawback of this task is that if it works, applications would be restricted to flight modes where the pilot is following a flight-director stimulus; in out-the-window flying, cockpit systems generally do not know to what object the pilot is attending.

**EXPERIMENT 1: AUDITORY CONTROL**

**Method.** Subjects were 20 college students, ages 18 to 34, who received course credit for participation. The experiment ran on a Compaq IBM-compatible personal computer. Stimuli for task 1 (tracking) were presented on a Compaq color monitor, driven by a VGA graphic card (640x480 pixels, or 25.6x19.2 degrees of visual angle at 63 cm viewing distance). These stimuli consisted of a cross shaped red Target, and a blue Tracking Circle. The Target and the Tracking Circle were vertically centered. The Target horizontal position...
changed once per trial. The Tracking Circle's horizontal position was determined by the position of a joystick (zero-order control). The stimulus for task 2 consisted of one of two spoken words: "house" and "luck" (digitally recorded male human voice). A stimulus word was presented at a controlled time (SOA ms) after a task 1 Target shift. A microphone, Schmidt trigger and a VOTAN software recognizer were used to acquire, time and recognize the subject's voice response. Joystick position was continuously recorded to permit extraction of timing and accuracy information. The experiment was divided into 10 blocks of 48 trials each. The first 2 blocks were treated as practice and not included in the analysis. The 48 trials in each block consisted of six repetitions of all combinations of 4 SOA levels (50, 150, 400 and 700 ms) and two spoken words ("house", "luck"), in random order. Each new target position was constrained to be at least 1.6 and at most 4.8 degrees of visual angle away from its previous position and at most 4.8 degrees away from either side of the screen. Values were randomly selected until within the constraints (thus shift direction probability deviated from 0.5 near either side of the screen). To reduce any fixed temporal response patterns, the time delay from the beginning of a trial to Target shift was selected quasi-randomly. Subjects were given written and verbal instructions. They were instructed to move the Tracking Circle by means of the joystick to keep the Tracking Circle centered over the Target cross, and to respond verbally to the auditory stimulus (saying "house" as a response to "one" and "luck" as a response to "two"). They were instructed to respond accurately and rapidly to both stimuli, but were told to give priority to the tracking task. At the end of each block subjects were given performance feedback.

Results & Discussion. The results of Experiment 1 are shown in Figure 1. For the higher priority tracking task, RT1 (measured from stimulus change to crossing the midpoint for the required movement) was very close to flat over SOA, F(3,57) = 0.51, P = 0.7, confirming that interference effects were successfully limited to only the second task. For the tones task, RT2 increased significantly as SOA shortened, F(3,57) = 46.9, P < .001. However, the total increase of 107 ms between SOA 700 and SOA 50 was substantially less than the 200-300 ms typical in PRP experiments. Over the initial SOA 50 - SOA 150 interval, the slope of RT2 was -0.32, again much lower than usual for PRP results. Thus the results of Experiment 1 provide a comparison baseline for the relatively low level of interference obtainable with tracking as task 1 and an auditory task 2. For comparison, Johnston & Delgado (1993), data with the reverse task order (task 1 auditory, task 2 tracking), found total interference of 67 ms and a slope of the initial curve segment of -0.22. The probable reason that the present data show more interference is that with the present task order tracking continues beyond the time of the nominal response. With the present task order, the resource demands of this additional performance are more likely to coincide with demands of the other task.

EXPERIMENT 2: SPLIT-BAR GAUGE

Method. Subjects were 19 additional students from the same subject pool as in experiment 1. The method was the same as for experiment 1, except the stimulus for task 2 consisted of a gauge made of three green horizontal line segments. In the Normal reading all three segments were at the same level, forming a continuous horizontal line centered above the tracking axis. In the High reading, the middle segment was elevated with respect to the side (reference) segments, and in the Low reading the center segment was lowered. The gauge reading changed from Normal to either High or Low at a controlled time (SOA ms) after target shift. It returned to Normal either after subject's voice response, or after a predetermined amount of time (4 sec) after the trial began, whichever occurred first. The design was the same as in Experiment 1. The procedure was also the same, except that on task 2 subjects were instructed to respond verbally (saying "High" or "Low") to the corresponding gauge deviation. They were instructed to use their peripheral vision for reading the gauge, if possible.

Results & Discussion. The results of Experiment 2 are shown in Figure 2. RT1 was very close to flat over SOA, F(3,54) = 6.17, P = .001. For the split-bar gauge task, RT2 increased significantly as SOA shortened F(3,54) = 324.09, P < .001. The total increase of 302 ms between SOA 700 and SOA 50 was substantially higher than the 107 ms for the auditory control data from Experiment 1, F(1,37) = 90, P < .001. Over the initial SOA 50 - SOA 150 interval, the slope of RT2 was -0.66, vs. -0.32 for the auditory control. Note that the total interference of 302 ms, is not only several times higher than the auditory control values, but is in the normal range of typical PRP findings. Hence we must conclude that the split bar gauge, unlike auditory stimuli, fails to take advantage of the special parallel processing advantages provided by the tracking task. Thus visual-visual interference is, as expected, a serious additional problem undermining parallel task performance.
The results of Experiment 4 show some progress toward parallel processing, although not to the level achieved in the control auditory Experiment 1. Two of our three well-practiced subjects showed only about 120-130 msec slow-down at low SOAs, well below usual PRP results, and only about one third of the interference with the digital gauge control that required fixation. The study showed that a considerable amount of parallel processing is possible when one task is zero-order tracking, even when the stimuli for both tasks are in the same (visual) modality. This result has practical implications for HUD symbology design. Since different gauge designs produced different amounts of interference, further research may lead to a design with even less interference. While further control studies are needed, we hypothesize that the ingredients critical to the success of the bent gauge include: 1) use of a quantity isomorphism principle, where more of the quantity represented corresponds to a larger display 2) use of multiple cues that are easily processed in the periphery, including orientation, size and color, and 3) presenting information substantially above the single-task threshold. Following Treisman’s filter attenuation theory, it may be necessary to present stimuli outside of spatial attention well-above threshold to overcome partial signal attenuation.

Future extensions of this work are planned in several directions. We will continue to test a variety of designs. However, the design space is indefinitely large and very heterogeneous. For guidance, one high priority is to carry out locus-of-slack experiments (McCann & Johnston, 1992) with our displays that will permit us to pinpoint the processing locus of residual interference. This should permit us to make principled improvements in gauge designs. We also plan to use a continuous tracking paradigm, with some superimposed large shifts that will still produce discrete responses traceable to particular stimulus changes. Nettick and Klapp (1994), using a different paradigm, but a continuous version of the zero-order tracking task, found “stick-freezing” they attributed to a single-channel bottleneck. Whether their results were due to the use of continuous tracking or to other task differences needs to be determined. Finally, we need to test the best designs in a more realistic experiment with a more complete HUD image. We can then determine whether parallel processing degrades with a more cluttered screen and a more variable task environment.

REFERENCES

Figure 1. Dual-task performance with auditory task 2 stimulus.

Figure 2. Dual-task performance with split-bar gauge task 2 stimulus (solid lines). Dotted line shows Expt. 1 RT2 for comparison.

Figure 3. Dual-task performance with integrated icon task 2 stimulus (solid lines). Dotted line shows Expt. 1 RT2 for comparison.

Figure 4. Dual-task performance with bent gauge task 2 stimulus (solid lines). Dotted line shows Expt. 1 RT2 for comparison.
Experiment 1

Figure 1. Dual-task performance with auditory task 2 stimulus.

Experiment 2

Figure 2. Dual-task performance with split-bar gauge task 2 stimulus (solid lines). Dotted line shows Expt. 1 RT2 for comparison.

Experiment 3

Figure 3. Dual-task performance with integrated icon task 2 stimulus (solid lines). Dotted line shows Expt. 1 RT2 for comparison.

Experiment 4

Figure 4. Dual-task performance with bent gauge task 2 stimulus (solid lines). Dotted line shows Expt. 1 RT2 for comparison.