

Integration of Size and Binocular Disparity Visual Cues in Manual Depth-Control Tasks

Barbara T. Sweet* and Mary K. Kaiser†
 NASA Ames Research Center, Moffett Field, CA 94035

Humans rely on a plethora of visual cues to inform judgments of the depth or range of a particular object or feature. One source of information is binocular (or stereo) vision, enabled by the slight differences in the images between the two eyes. Another source of information is the size, in visual angle, of the feature or object. Although it is possible to provide binocular disparity in a display, it typically implies increased system cost or lower update rates; therefore it is useful to determine how critical it is to provide binocular disparity to the user. This paper describes the results of several experiments investigating the integration of stereo and size cues in performing manual control tasks. In the first experiment, a visual cue integration model was developed for two types of manual control tasks (rate-control and acceleration-control) and different levels of cue salience. The results of this experiment were that stereo disparity dominated judgments of depth position, while size dominated judgment of depth rate. From this experiment, it was hypothesized that stereo disparity would do more to improve performance on rate-control tasks more than on acceleration-control tasks. Two additional experiments were conducted, with and without stereo display, to test this hypothesis at different update rates. The results confirmed that while stereo disparity improved performance on rate-control tasks, it did not improve performance on acceleration-control tasks; in fact, performance was reduced with stereo disparity when the display method reduced update rate below a particular threshold.

Nomenclature

d	=	displayed depth, inches
H_{RS}	=	modeled operator describing function, control output to relative size
H_{SD}	=	modeled operator describing function, control output to stereo disparity
H_T	=	modeled operator describing function, control output to displayed depth
\hat{H}_{RS}	=	measured operator describing function, control output to relative size
\hat{H}_{SD}	=	measured operator describing function, control output to stereo disparity
\hat{H}_T	=	measured operator describing function, control output to depth
K_P	=	position control gain
K_V	=	velocity control gain
p	=	perceived depth position, inches
r	=	remnant
s	=	laplace transform variable
v	=	perceived depth velocity, inches
W_{PS}	=	weight of stereo disparity in position perception
W_{VS}	=	weight of stereo disparity in velocity perception
x	=	disturbance in displayed depth, inches
y	=	disturbance in relative size, inches
Y_C	=	controlled element dynamics
Y_P	=	human operator compensation
Y_N	=	neuromuscular dynamics

* AST Man-Machine Systems, Human Systems Integration Division, MS 262-2, member.

† Research Psychologist, Human Systems Integration Division, MS 262-2, non-member.

z	=	disturbance in stereo disparity, inches
δ	=	operator control output, unitless, range from -1.0 to 1.0
ϕ_M	=	phase margin, deg
τ	=	time delay, seconds
τ_E	=	effective time delay, seconds
τ_L	=	lead time constant, seconds
ω_C	=	crossover frequency, rad/sec
ω_N	=	neuromuscular dynamics natural frequency, rad/sec
ζ_N	=	neuromuscular dynamics damping ratio

I. Introduction

PSYCHOLOGISTS have long recognized that the human visual system has access to multiple sources of information specifying depth. These depth cues are usually grouped in terms of their information “type” (that is, physiological, pictorial, or motion), or by the level to which they specify depth (ordinal, relative, or absolute). Any good textbook on visual perception¹ can provide an overview of these cues and taxonomies. More thorough treatments can be found in the relevant chapters of Reference 2; see especially the Volume I chapters by Sedgwick³, Hochberg⁴, and Arditi⁵. We shall provide a very cursory summary.

A. Depth Cue Taxonomies

The British philosopher Berkeley⁶ provided an early taxonomy of depth cues. Berkeley was most concerned with what he termed “primary” depth cues (now more commonly called physiological cues): accommodation, convergence, and binocular stereopsis. Accommodation refers to the degree to which ocular muscles tense or relax to adjust the thickness of the eye's lens to focus on an object. Convergence is the degree to which the eyes angle toward one another to look at the object. In principle, both accommodation and convergence can provide absolute depth information, although most research suggests that, in practice, these cues play a minor role⁷. The third primary (or physiological) depth cue, binocular stereopsis, exploits the disparity information resulting from the displacement of our two eyes. If eyepoint separation and vergence angle is known, binocular disparity can, in principle, specify absolute depth. At a minimum, it provides compelling relative depth cues to people with functional stereopsis. (Approximately 5% of the population lack this ability and are “stereo blind”⁸.)

What Berkeley termed “secondary” depth cues are now more commonly called pictorial cues. As might be expected, these refer to the cues resulting from linear perspective, and have been exploited since the Renaissance by artists to convey an impression of depth in two-dimensional depictions. A partial list of these cues include: occlusion (the occluding object is closer); image size (larger images appear closer); height in the visual field (images closer to the horizon appear more distant); and atmospheric perspective (distant objects lose brightness and contrast due to atmospheric attenuation). Occlusion is a good example of an ordinal depth cue: the fact that Object A occludes Object B tells us only that Object A is closer. If one knew the contrast/brightness fall-off function for distance, atmospheric attenuation could, in principle, specify absolute depth. In practice, it too functions as an ordinal cue. Relative image size and height-in-field generally provide relative depth information (although given additional knowledge, such as absolute object size and eye height respectively, absolute depth could, in theory, be recovered).

The final class of depth cues results from motion. While this can be object motion (e.g., the image velocity of an object falling is inversely proportional to its distance from the observer), most depth-from-motion results from the motion of the observer through the environment. Psychologists typically describe this information in terms of “motion parallax” (i.e., motion lateral to a pair of objects results in greater image velocity for the nearer object) or “optical expansion” (as an observer approaches a pair of objects, the closer one's image will have a greater radial flow rate). Just as knowing inter-ocular separation and vergence allows one to recover absolute distance from disparity, knowledge of ego-speed allows recovery of absolute distance from motion parallax. Even without such knowledge, motion parallax is a compelling relative depth cue.

B. An Alternate Depth Cue Taxonomy

More recently, Cutting and Vishton⁹ proposed an alternative analysis of depth cues by examining which cues are more or less useful as a function of context. Obviously, motion depth cues are only informative in situations where the observer (or objects) are moving. But the functional utility of all cues vary depending on situational specifics. For example, accommodation, convergence, and stereopsis are only useful at relatively near distances; beyond

fifteen feet, all of these cues become sub-threshold (i.e., imperceptible to human observers). Conversely, atmospheric perspective is sub-threshold at close distance, and only becomes a meaningful cue when objects are thousands of meters distant (unless one is in San Francisco on a foggy day).

Cutting and Vishton⁹ categorize depth cues into those whose utility is invariant with distance (e.g., occlusion and relative size), those whose utility diminishes with distance (e.g., the physiological depth cues); and those whose utility increases with distance (e.g., atmospheric perspective). They then divide the space surrounding an individual into three functional regions: Personal space (0-2 meters -- generally, the region in which a person manipulates objects); Action space (2 - 30 meter -- the region in which a person moves quickly to act upon the environment); and Vista space (beyond 30 meters -- basically the region in which a person plans future navigation). Cutting and Vishton argue that, because the relative utility of depth cues vary as a function of region, the relative importance (or weighting) the observer places on those cues will likewise vary. This raises the more general question of how observers integrate depth cues.

C. Depth-Cue Integration

Since the late 1980's, there has been a concerted effort to model how human observers integrate depth cues. In most natural viewing situation, of course, the various cues are consistent with one another. In fact, information provided by one depth cue can “promote” another's information by providing additional constraints¹⁰. For example, if convergence specifies the absolute distance to the nearer of two objects, then it can “promote” stereopsis such that metric layout is recoverable. Thus, the visual system need only determine a “depth map” that satisfies the multiple constraints of the contributing cues.

In contrast, an artificial spatial display (be it a painting or a simulator screen) typically contains conflicting depth cues. For example, in a depiction of an outdoor scene, linear perspective might specify Tree B's depth to be twice as great as Tree A's, but accommodation and convergence suggest they're at the same depth. How does our visual system resolve this paradox?

1. Depth-Cue Integration Models

Generally, models posit similar depth-cue integration mechanisms for both situations: an algebraic combination of depth cues. Three classes of combinatory rules have been proposed: Selection; Additive combinations; and Multiplicative combinations. Bruno and Cutting¹¹ provide a useful overview of these classes.

Selection occurs when only one depth cue is used to determine an object's depth and the other available cues are disregarded. This mechanism is mathematically equivalent to Bulthoff and Mallot's veto process¹². In such a scheme, one cue is utilized, unchallenged by other cues (despite possible discrepancies in their depth information). The equation describing this process is:

$$d = f(s_1) \tag{1}$$

where d is the distance perceived, s_1 and s_2 are two candidate sources of information (i.e., the two depth cues) with s_1 being the only one that is utilized for the determination of depth. (For simplicity, we will use only two depth cues in our examples.)

In an *Additive* scheme of cue integration, observers process all available cues, weight them, then add the results to determine the depth. This model can be described by the following equation:

$$d = f(w_1 s_1 + w_2 s_2) \tag{2}$$

where d is the perceived distance, s_1 and s_2 are sources of information, and w_1 and w_2 are the weights assigned to each source depth. Note, of course, that Selection is simply a special case of the Additive model, in which the weights for all but one cue are set to zero.

The third possible rule class involves the *Multiplicative* combination of depth cues. In these models, observers use some cues to modify information from other cues. A plausible equation for Multiplicative integration is:

$$d = f(w_1 s_1 w_2 s_2) \tag{3}$$

As Bruno and Cutting acknowledge, hybrid combinatory rules may prove viable, combining Addition and Multiplication in various way, such as where a particular depth cue (s_1) is weighted independently and also influences the weighting of a second depth cue (s_2), as in:

$$d = f(w_1 s_1 + s_1 w_2 s_2) \quad (4)$$

Or cues could be weighted both independently and in the context of other cues simultaneously, as in:

$$d = f(w_1 s_1 + w_2 s_2 + w_1 s_1 w_2 s_2) \quad (5)$$

2. Cue Integration Findings

While Selection is seldom proposed as the primary mechanism for depth cue integration, instances can be found in which Selection appears to operate, particularly in the case of cue conflict. For example, Bulthoff and Mallot¹² found that if edge information (i.e., occlusion) is present, it overrides both shape-from-shading and disparate shading-depth information.

More commonly, empirical studies suggest Additive combination rules. Bruno and Cutting¹¹ performed three experiments testing perceived exocentric distances as a function of both static and motion cues (including relative size, height in the projection plane, occlusion, and motion parallax) and found the greatest support for the Additive combination rule. Similarly, linear combination rules provide good fits for the combination of stereo disparity and texture gradient¹³, texture gradient and motion parallax¹⁴, and stereo disparity and linear perspective¹⁵.

A number of researchers have reported findings consistent with Multiplicative combination rules. Massaro's Fuzzy Logical Model of Perception (FLMP) used a specific Multiplicative model of cue integration based on fuzzy logic^{16,17} to fit depth judgment data and reported a fit superior to that obtained with linear models. Others have reported superior fits with non-linear models, especially in cases of recovering surface structure from multiple depth cues^{18,19}.

A study by Johnston, Cumming, and Landy²⁰ lends empirical credence to Cutting and Viston's proposal of contextual cue weighting. Johnston, et al. pitted stereo disparity against motion parallax cues in their task, and varied both the number of frames of animation (to vary the utility of the motion cue) and the observer's viewing distance (to vary the utility of the disparity cue); they found that observers' weighting of the two cues varied as a function of condition, with greater weight assigned to the stronger cue.

D. Extending Cue Integration to an Active Control Task

Both the Cutting and Viston chapter and the Johnston, et al. study recognize that depth-cue integration is unlikely to be a fixed, inflexible process. Rather, our perceptual system is sufficiently intelligent to consider the quality and reliability of the various sources of information when deriving an estimate of depth. The Modified Weak Fusion model proposed by Landy, et al. and Massaro's FLMP likewise recognizes that the weighting of cues should be dynamic (i.e., adjusting to accommodate changes in viewing circumstances, and resulting changes in the various cues' utility).

However, all of this work has examined depth-cue integration in the context of “passive” perception – that is, observers are asked to view displays and make verbal or keyboard responses concerning scene layout or surface curvature. Our goal is to study depth-cue integration in the context of active control, and to model depth perception as one component of the manual control task. In this way, we build upon previous models of depth-cue integration, and expand their application to a dynamic, closed-loop control model.

As we will show, current formulations of depth-cue integration are amenable to inclusion as modules in larger control models. Once the cue-integration module is integrated into the control model, we can examine whether people's depth-cue integration is impacted, not only by changes in the “quality” of the cue, but also by the utility that information holds for the control task they must perform. Thus, we can investigate whether people's depth cue integration strategies are merely clever enough to adjust to changes in cue “quality,” or sufficiently intelligent to utilize the cues best suited for the task at hand.

II. Depth Cue Control Model

In this report, a model is developed that describes the control strategy the human operator adopts in performing a depth control task when two depth cues are available to the operator. It is an extension of a modeling technique that was developed to examine manual control in perspective scene viewing situations²¹. This modeling technique relies heavily on the discipline of manual control, and the Crossover Model^{22,23}. In this section, a brief background on the Crossover Model is presented. Then, a model of depth-cue integration and control is presented that is based upon the characteristics of the Crossover Model.

A. Crossover Model of Manual Control

Manual control is described as a control situation in which a human operator is required to make nearly continuous control adjustments to achieve a desired outcome. This situation has been extensively studied and modeled; a summary of the various approaches can be found in Ref. 24. One of the first models developed for simple control situations is termed the Crossover Model^{22,23}. This model is descriptive of the control compensation a human operator provides in a variety of circumstances. A simplified block diagram of this model is shown in Fig. 1.

1. McRuer and his colleagues found that, over a large variation in the dynamics of the controlled element Y_C , the operator adjusted his compensation Y_P in such a way that $Y_P Y_C = \exp(-s\tau_E) \omega_C/s$. In words, the operator adjusts his compensation so that the product of his compensation and the controlled element will yield an integrator with a time delay. The crossover frequency ω_C determines the bandwidth of the closed-loop system, or the input frequencies above which effective tracking cannot be accomplished. Typical values²² for ω_C range from 1.0 to 6.0 rad/sec, effective time delays τ_E range from 0.2 to 0.5 seconds. The crossover frequency ω_C is defined as the frequency at which the open-loop system transfer function has a magnitude of unity, $|Y_P(j\omega_C)Y_C(j\omega_C)| = 1$.

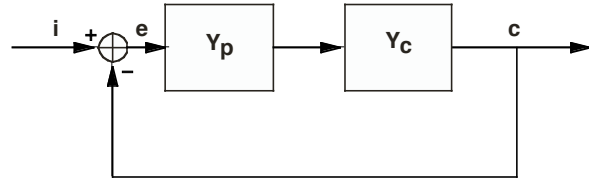


Figure 1. Simplified block diagram of Crossover Model of manual control.

The effects of changing controlled element dynamics can be plainly seen with this model. Consider the case of rate-control (first-order) dynamics ($Y_C = 1/s$). For this case, the operator would apply the approximate compensation $Y_P = \omega_C \exp(-s\tau_E)$. This type of compensation on the part of the operator is termed proportional compensation; the output of the operator is simply a scaled (ω_C) and time-delayed ($\exp(-s\tau_E)$) version of the input. When presented with acceleration-control dynamics ($Y_C = 1/s^2$), the operator needs to provide compensation of the approximate form $Y_P = s \omega_C \exp(-s\tau_E)$. This type of compensation is called derivative compensation because of the s term; instead of feeding back position, the operator is feeding back a time-delayed derivative of the input, which can also be termed velocity. When using rate-control dynamics, the operator needs to supply only proportional or position information. When the dynamics become acceleration control, the operator must feed back velocity information instead.

B. The Depth-Cue Integration and Control Model

1. An Idealized Depth Control Paradigm

Because the controlled element dynamics affect what type of information the operator is trying to use as feedback, these dynamics might also affect the way in which the operator obtains information from a display. A conceptual block diagram of the human performing a depth-control task is shown in Fig. 2.

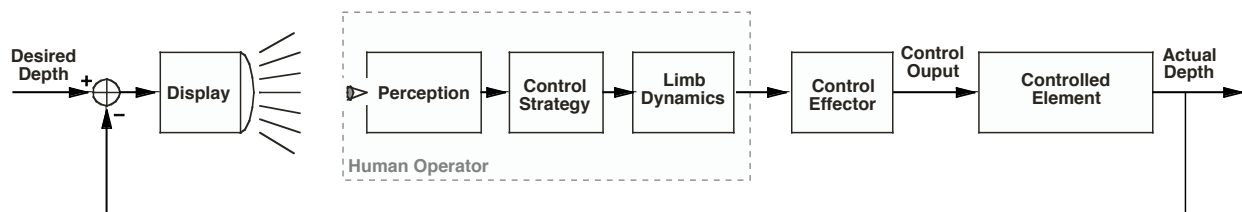


Figure 2. Conceptual block diagram of a manual depth-control task.

In this idealized system, the human operator is presented with both desired depth and actual depth via a perspective display. The operator perceives some characteristic(s) of the display, and formulates a desired control response. The perception and control process includes integration of all available cues into a single intended control response. This intended control response is used to drive the limb of the operator, which is manipulating the control effector (such as a joystick). The displacement of the control effector drives the dynamics of the controlled element, which affects the actual depth being displayed.

Although this figure is useful for conceptualizing the depth-cue integration and control task, it is not a particularly useful model for experimental validation. This is because many of the states and characteristics shown are unmeasurable. Specifically, states internal to the operator, or even direct force output of the operator that cause

the displacement of the control manipulator, are not generally available. Thus, when examining this situation in an experiment, it is the input-output relationships between the controlled states (in this case depth) and the control effector displacement that can be characterized. In the next section, such a model is described.

2. Describing Function Model and Measurement

A block diagram of the model used for experimental measurement and validation is shown in Fig. 3. As can be seen, the display, human operator, and control effector are effectively lumped into one system. This is necessary because the direct inputs to the human operator from the display are not known, and the direct force output of the human is not known. However, the inputs to the display, and the output of the control effector, are known.

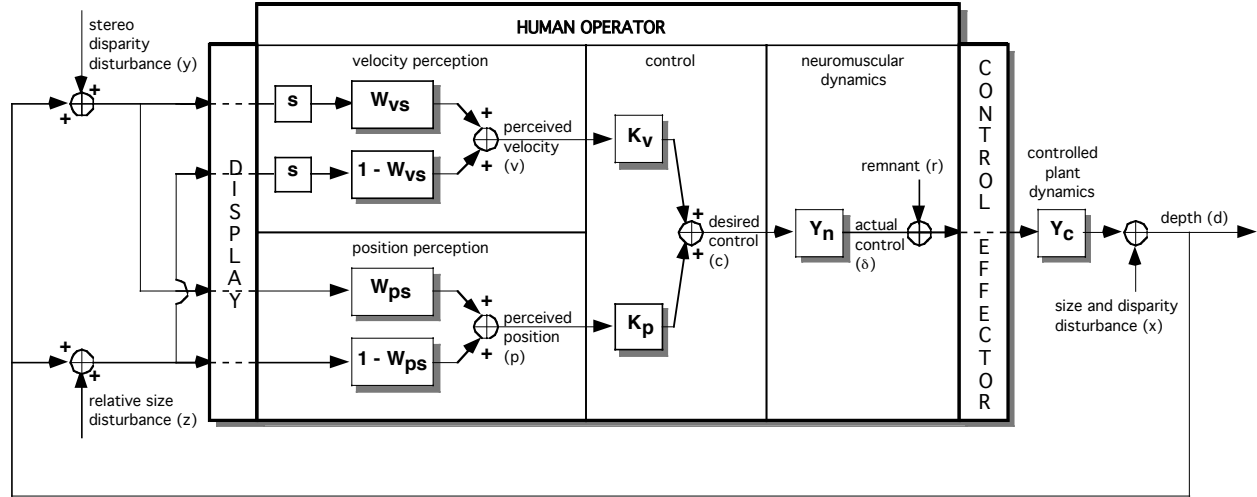


Figure 3. Block diagram of depth-cue integration model used for describing-function identification, assuming neuromuscular dynamics process

The human elements of the model include perceptual, control, and neuromuscular dynamics processes. The input to the perceptual process is the display, and the output of the neuromuscular dynamics goes directly into the control effector. The perceptual process includes both position and motion perception. It is assumed that the operator uses both relative size and stereo disparity to form perceptions of position (depth) and velocity (depth motion). The outputs of the perception, both position and velocity, are weighted and combined to form a desired control response. This desired control response then goes through the neuromuscular system before acting upon the control effector.

Several parameters and dynamic elements define the model shown in Fig. 3. The contribution of the perceptual processes are defined by two weights, W_{VS} and W_{PS} , which can each have values between zero and one. W_{VS} is the weighting put on stereo disparity in the velocity perception process; W_{PS} is the weighting of stereo disparity in the position perception process. By constraining the weights to vary between zero and one, the outputs of the perception process are simply weighted sums of the inputs from stereo disparity and relative size, without any gain factors. In the control process, the control strategy of the operator is created by applying gains to velocity K_V and position K_P , then summing these to form a control input. It has long been established in manual control that an operator's control strategy resembles the above control block: a linear combination of velocity and position feedback. The neuromuscular process contains a neuromuscular transfer function, Y_N , which will be elaborated upon in Section III, as well as a source of internal noise r , called remnant.

Characteristics of this model related to the perception, control, and neuromuscular dynamics can be measured through careful selection and manipulation of the disturbances affecting the displayed depth. An overall disturbance, x , is used to perturb both stereo disparity and relative size simultaneously. At the same time, disturbances y and z are used to independently perturb the stereo disparity and relative size, respectively. By examining the interrelationships between the disturbances and the control output of the operator, experimental measurements related to the model parameters and functions (W_{VS} , W_{PS} , K_V , K_P , Y_N) can be obtained.

From the block diagram in Fig. 3, we can write the following relationships:

$$\delta = -Y_N(K_V v + K_P p) + r \quad (6)$$

$$d = -Y_C \delta + x \quad (7)$$

$$v = s[(W_{VS}(d+y)+(1-W_{VS})(d+z)] \quad (8)$$

$$p = [(W_{PS}(d+y)+(1-W_{PS})(d+z)] \quad (9)$$

By substituting Eqs. (7), (8), and (9) into Eq. (6), we obtain the following expression for δ which is only a function of the model parameters and model inputs:

$$\begin{aligned} \delta[1 + Y_N Y_C (sK_V + K_P)] = & -Y_N \{ (sK_V + K_P)x + (sK_V W_{VS} + K_P W_{PS})y \\ & + [sK_V(1 - W_{VS}) + (K_P(1 - W_{PS}))z] + r \end{aligned} \quad (10)$$

Similarly, substituting Eqs. (6), (8) and (9) into Eq. (7) will yield an expression for d which is only a function of the model parameters and model inputs:

$$\begin{aligned} d[1 + Y_C Y_N (sK_V + K_P)] = & Y_C \{ -Y_N [(sK_V W_{VS} + K_P W_{PS})y \\ & + (sK_V(1 - W_{VS}) + K_P(1 - W_{PS}))z] + r \} + x \end{aligned} \quad (11)$$

The term $[1 + Y_C Y_N (sK_V + K_P)]$ in the previous equations appears repeatedly in the following derivations. A simplifying term will now be defined for ease of interpretation:

$$\Delta = 1 + Y_C Y_N (sK_V + K_P) \quad (12)$$

Taking the cross-spectral densities of Eqs. (10) and (11) with respect to x will yield:

$$\begin{aligned} \Phi_{\delta X} = \frac{1}{\Delta} [-Y_N \{ (sK_V + K_P) \Phi_{XX} + (sK_V W_{VS} + K_P W_{PS}) \Phi_{YX} + \\ s(K_V(1 - W_{VS}) + K_P(1 - W_{PS})) \Phi_{ZX} \} + \Phi_{RX}] \end{aligned} \quad (13)$$

$$\begin{aligned} \Phi_{DX} = \frac{1}{\Delta} [Y_C \{ -Y_N [(sK_V W_{VS} + K_P W_{PS}) \Phi_{YX} \\ + (sK_V(1 - W_{VS}) + K_P(1 - W_{PS})) \Phi_{ZX}] + \Phi_{RX} \} + \Phi_{XX}] \end{aligned} \quad (14)$$

Taking into account the fact that the disturbances x , y , and z are not correlated with each other ($\Phi_{YX} = \Phi_{ZX} = 0$), Eqs. (13) and (14) become:

$$\Phi_{\delta X} = \frac{1}{\Delta} [-Y_N (sK_V + K_P) \Phi_{XX} + \Phi_{RX}] \quad (15)$$

$$\Phi_{DX} = \frac{1}{\Delta} [Y_C \Phi_{RX} + \Phi_{XX}] \quad (16)$$

Further assuming that the noise signal r is uncorrelated with x ($\Phi_{RX} = 0$) and taking the ratios between the two expressions, we get:

$$\Phi_{\delta X} / \Phi_{DX} = -Y_N (sK_V + K_P) \quad (17)$$

The cross-spectral density of δ can also be derived relative to y and z :

$$\begin{aligned} \Phi_{\delta Y} = & \frac{1}{\Delta} [-Y_N \{ (sK_V + K_P)\Phi_{XY} + (sK_V W_{VS} + K_P W_{PS})\Phi_{YY} \\ & + (sK_V(1 - W_{VS}) + K_P(1 - W_{PS}))\Phi_{ZY} \} + \Phi_{RY}] \end{aligned} \quad (18)$$

$$\begin{aligned} \Phi_{\delta Z} = & \frac{1}{\Delta} [-Y_N \{ (sK_V + K_P)\Phi_{XZ} + (sK_V W_{VS} + K_P W_{PS})\Phi_{YZ} \\ & + (sK_V(1 - W_{VS}) + K_P(1 - W_{PS}))\Phi_{ZZ} \} + \Phi_{RZ}] \end{aligned} \quad (19)$$

Accounting for the uncorrelated disturbances ($\Phi_{XY} = \Phi_{XZ} = \Phi_{YZ} = \Phi_{RX} = \Phi_{RY} = \Phi_{RZ} = 0$), Eqs. 11 and 12 become:

$$\Phi_{\delta Y} = \frac{1}{\Delta} [-Y_N \{ sK_V W_{VS} + K_P W_{PS} \} \Phi_{YY}] \quad (20)$$

$$\Phi_{\delta Z} = \frac{1}{\Delta} [-Y_N \{ sK_V(1 - W_{VS}) + K_P(1 - W_{PS}) \} \Phi_{ZZ}] \quad (21)$$

We can use these relationships as the basis for empirical modeling based upon experimental measurements of operator response. Time histories of the variables d , x , y , z , and δ (see Fig. 3) can be used to estimate these power- and cross-spectral densities^{25,26}. The describing functions derived from the cross-spectral density measurements are defined as:

$$\hat{H}_T = \Phi_{\delta X} / \Phi_{DX} \quad (22)$$

$$\hat{H}_{SD} = \Phi_{\delta Y} / \Phi_{YY} \quad (23)$$

$$\hat{H}_{RS} = \Phi_{\delta Z} / \Phi_{ZZ} \quad (24)$$

The next section describes an experiment designed to experimentally determine these describing functions, and develop corresponding parameterized models.

III. Experiment 1

An experiment was conducted to estimate the describing functions defined in the previous section, and to develop parameterized models corresponding to those measurements. Both the controlled element dynamics and the viewing distance were manipulated to determine what effect (if any) these variables had on the operator characteristics. Section A contains a description of the experiment methods, Section B contains the results, and discussion of the results is contained in Section C.

A. Method

1. Participants

Eight male, general-aviation pilots participated in the study. They were recruited from a paid contractor pool at Ames Research Center. All had normal or corrected-to-normal visual acuity and good stereo vision (40 seconds of arc or better). Their flight experience ranged from 100 to 4500 logged hours.

2. Apparatus

The experimental control program was run on a Silicon Graphics Octane computer with an R10000 processor. Control inputs were made via a B&G Systems JF3 3-axis joystick. (Only the longitudinal degree of freedom of the stick was used; the lateral and yaw inputs of the stick were disregarded.) Stereo images were viewed through Crystal Eyes polarizing shutter glasses. The monitor displayed the views for the left and right eye on alternating

refreshes at a rate of 96 Hz, yielding an effective update rate to each eye of 48 Hz. Control data from the joystick was updated at 48 Hz. The images were displayed on a 19-inch diagonal monitor, with a resolution of 1024 (width) by 768 (height) pixels.

3. Stimuli and Control Tasks

In all conditions, participants viewed two horizontally aligned geometric forms. The background color was grey, with an RGB (red, green, blue) value of [0.7, 0.7, 0.7] (where [0, 0, 0] is black and [1.0, 1.0, 1.0] is white). The faces of the cubes were blue [0.5, 0.5, 1.0], and the edges of each face were outlined with a darker blue [0.4, 0.4, 0.8]. (These colors were chosen to minimize the stereo “ghosting” that can result from persistence of the monitor phosphors, while still providing adequate levels of contrast.)

The left-hand object served as the “standard” and was rendered at a constant depth. The object on the right was the control target. Participants were instructed to move the joystick longitudinally (fore and aft) to maintain the target at the same apparent depth as the standard.

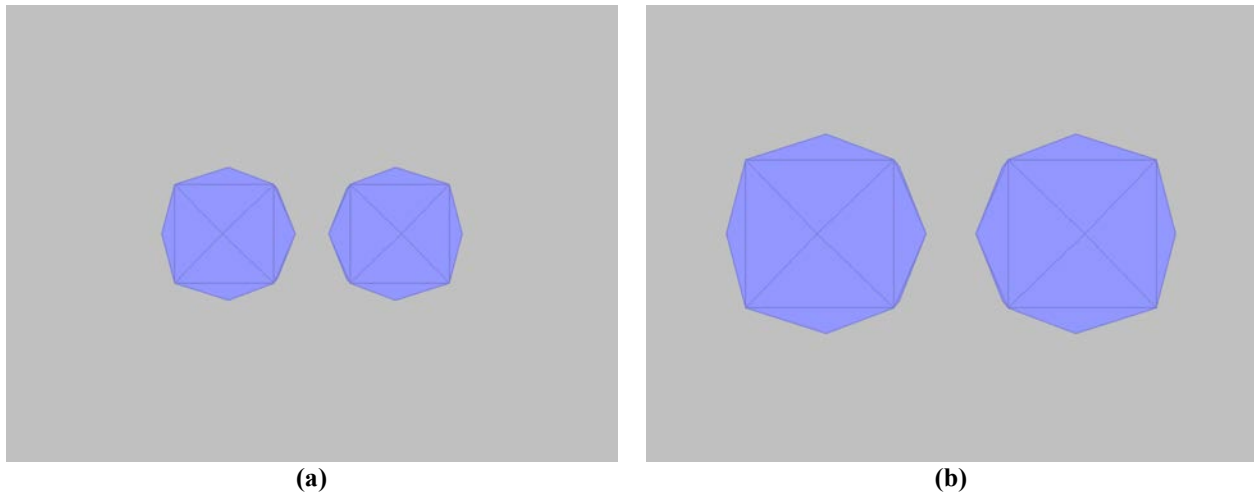


Figure 4. Screen image for near (a) and far (b) viewing conditions. These images were rendered using zero stereo disparity; in the experiment, the Standard object was rendered with a set disparity and the Target object with a variable disparity.

Other aspects of the displays and control task were varied as a function of experimental condition. The three experimental factors were: viewing distance; control task dynamics; and disturbance function. We discuss each of these in turn.

Viewing Distance - Participants were seated at two different viewing distances: “near” (22 inches from the screen); and “far” (33 inches from the screen). In the near condition, the display subtended approximately 35 (horizontal) by 26 (vertical) degrees. In the far viewing condition, the display subtended approximately 24 (horizontal) by 18 (vertical) degrees. In the near condition, the objects (at standard depth) were scaled to have a screen image size of 3.0 inches, and were spaced 4.5 inches apart (center to center). The near-viewing scene is shown in Fig. 4a. In the far condition, the scene was magnified to maintain visual angles equivalent to the near condition. Thus, because the distance was increased by a factor of 1.5, all of the scene features were also scaled by a factor of 1.5. However, the eyepoint separation used to render the views of the two eyes was held fixed at 3.0 inches. These manipulations ensured that the visual angle subtended by the object would remain consistent between the two conditions, while the stereo disparity was diminished because of the greater viewing distance. The far-viewing scene is shown in Fig. 4b.

Control Task Dynamics - Two different sets of control task dynamics were simulated: rate control and acceleration control. In the rate-control condition, the velocity (in depth) of the target was proportional to the amount of joystick displacement. In the acceleration-control condition, acceleration was proportional to stick displacement. The equations describing these dynamics, and difference equations used for digital simulation of the dynamics, are provided in the appendix.

Disturbances - Three disturbance sources were generated to perturb the target object's depth:

- i) A sum-of-sines disturbance of both depth cues (image size and stereo disparity). Thus, the disturbance of the two cues is correlated.
- ii) A sum-of-sines disturbance of the image size cue alone.

iii) A sum-of-sines disturbance of the stereo disparity cue alone.

In the Baseline Disturbance condition, only the correlated disturbance source is present (referring to Fig. 3, the disturbances y and z were set to zero, and x was non-zero). In the Multiple Disturbance condition, all three disturbance sources are present (specifically, x , y and z in Fig. 3 were all non-zero). The baseline condition was included to compare with operators' strategies in response to multiple, uncorrelated disturbance sources. The appendix contains additional information regarding the disturbances.

4. Design

As stated above, our three experimental factors were: 1) Viewing Distance (Near versus Far); 2) Control Task Dynamics (Rate versus Acceleration); and 3) Disturbance Type (Baseline versus Multiple). Limitations in our test apparatus (specifically, display size) made it infeasible to conduct a full factorial design. Specifically, the tight spacing of the Standard and Target objects in the Far Viewing condition was not compatible with Acceleration-Control Dynamics -- participants could not achieve sufficiently precise control to avoid images overlapping one another or the screen edge (which completely compromises the fidelity of the rendered depth cues). Thus, a partial factorial design (within-participant) was employed, wherein the five conditions presented to each participant were:

- i) Rate Control, Baseline Disturbance, Near Distance
- ii) Rate Control, Multiple Disturbances, Near Distance
- iii) Rate Control, Multiple Disturbances, Far Distance
- iv) Acceleration Control, Baseline Disturbance, Near Distance
- v) Acceleration Control, Multiple Disturbances, Near Distance

5. Procedure

Participants were given written instructions describing their task and its discernable variations (i.e., Near and Far viewing, Rate and Acceleration control). Participants were then given an opportunity to ask questions. Once started, the task was entirely self-paced. The experimenter intervened only to assist with changes in viewing distance as required between blocks of trials. The experiment took two days for participants to complete; participants experienced only one type of control dynamics (Rate or Acceleration) per day. Each day's session began with a brief session of training trials for each of the conditions the participant would see that day. Training trials were of 1 minute duration, data trials lasted four minutes, five seconds. Participants completed a total of eight data runs in each condition. Participants were not given feedback on their performance on either training or data trials.

6. Dependent Measures

Three types of dependent measures are considered. The first is a task performance metric, the root mean square (RMS) of the displayed depth error. The second type was based on the frequency response of the operator; it includes RMS of the percent of operator control activity correlated with the two independent disturbance sources, and the closed-loop crossover frequency and phase margin of the measured operator describing function. The third type of measure is the model parameters identified to match the operator describing function measurements.

B. Results

1. Task Performance

An ANOVA was performed on displayed depth error RMS to determine the effects of Viewing Distance and Control Task Dynamics. This analysis was conducted for only the Multiple Disturbance trial data (i.e., those trials that contained independent disturbances of the stereo disparity and relative size cues). A significant main effect on depth error RMS was found for Viewing Distance ($F[1,7] = 64.06$, $p < 0.001$), with a smaller error associated with the Near condition (Fig. 5a). A significant main effect was found for Control Task Dynamics ($F[1,7] = 32.09$, $p < 0.001$), with a smaller error associated with the Rate condition (Fig. 5b).

2. Frequency Response

ANOVA analysis was performed on the percent of control activity correlated with the relative size and stereo disparity disturbances, and on the crossover frequency and phase margin of the combined operator-controlled element describing function.

Percent of Control Activity – There was no significant main effect for Viewing Distance or Disturbance Source (Fig. 6a), but there was a significant interaction ($F[1,7] = 14.45$, $p < 0.01$); the percent of control activity associated with the two cues is approximately equal in the Near condition, whereas relative size dominates in the Far condition. There was a main effect for Control Task Dynamics ($F[1,7] = 6.20$, $p < 0.05$), with a higher percent of correlated control power in the Acceleration Condition (Fig. 6b). The effect of Disturbance Source did not reach significance ($F[1,7] = 4.55$, $p < 0.07$), but relative size demonstrated slightly greater correlated control power. Likewise, a trend towards an interaction of Control Task Dynamics and Disturbance Source was noted ($F[1,7] = 3.74$, $p < 0.094$).

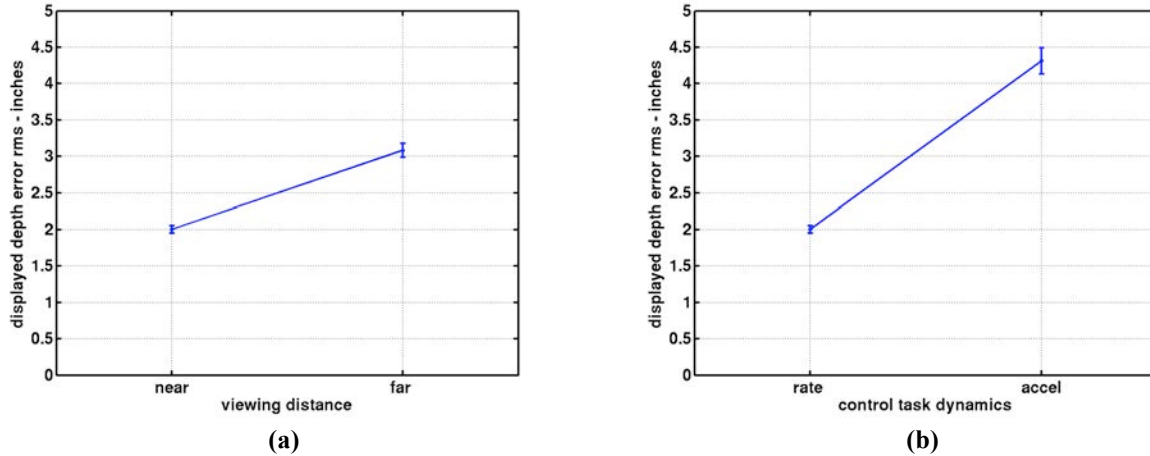


Figure 5. Effect of viewing distance (a) and control task dynamics (b) on depth error RMS. Means and standard error bars are shown.

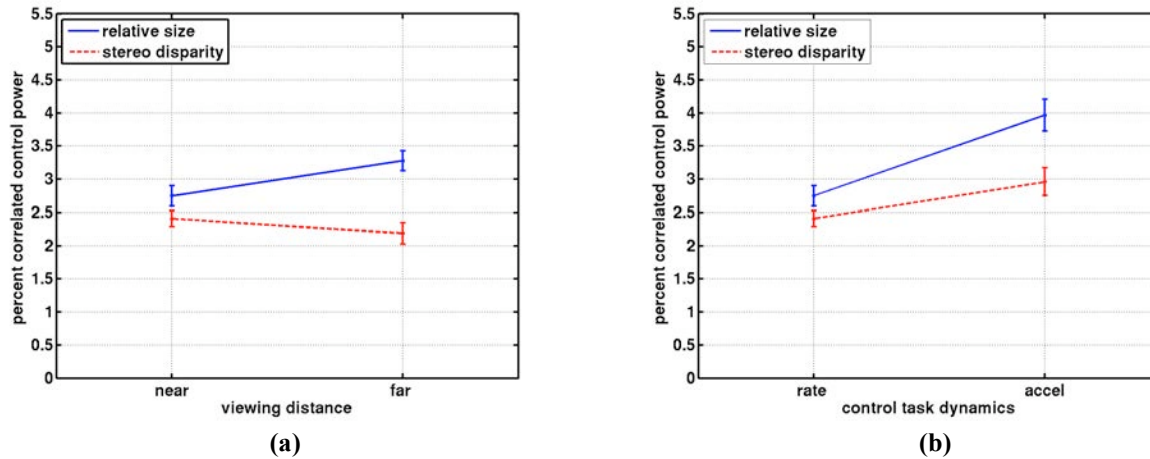


Figure 6. Effect of viewing distance (a) and control task dynamics (b) on the percent of control power correlated with the independent disturbances to relative size and stereo disparity. Means and standard error bars are shown.

Crossover Frequency and Phase Margin – The measured describing function defining the operator’s response to depth, \hat{H}_T , was used to determine the crossover frequency and phase margin of the operator-controlled element combination. The mean crossover frequencies and phase margins as a function of condition are shown in Figs. 7 and 8. There is no significant variation in crossover frequency as a function of either viewing distance or control task dynamics. Similarly, there is no significant difference in phase margin as a function of viewing condition. There was a significant main effect of control task dynamics on phase margin; the acceleration control case was associated with a significantly lower phase margin.

The observed range of crossover frequencies, phase margins, and trends as a function of controlled element dynamics are all consistent with the previous manual control findings^{22,23}. These parameters do little to examine the perceptual and control processes of the operator, which are considered in detail in the next section.

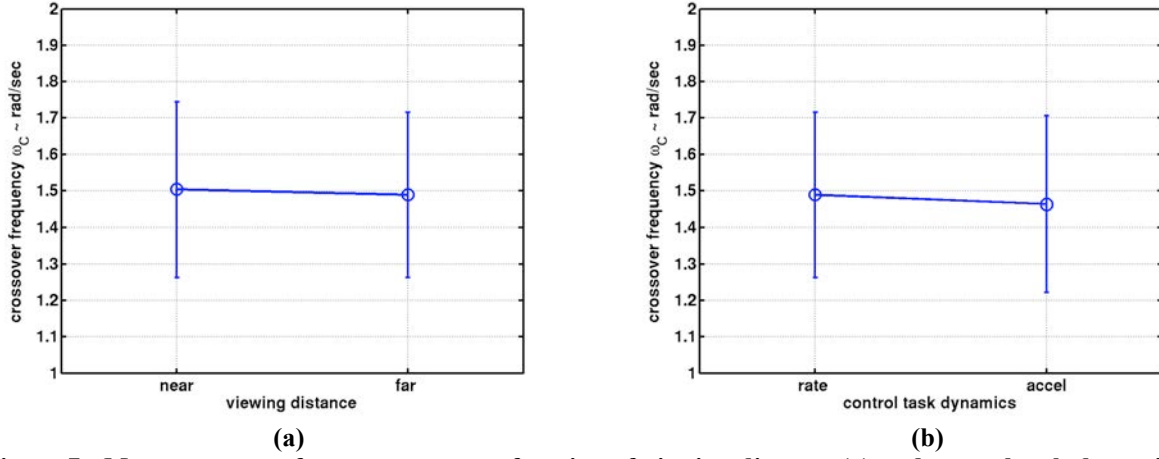


Figure 7. Mean crossover frequency ω_C as a function of viewing distance (a) and control task dynamics (b). Means and standard error bars are shown.

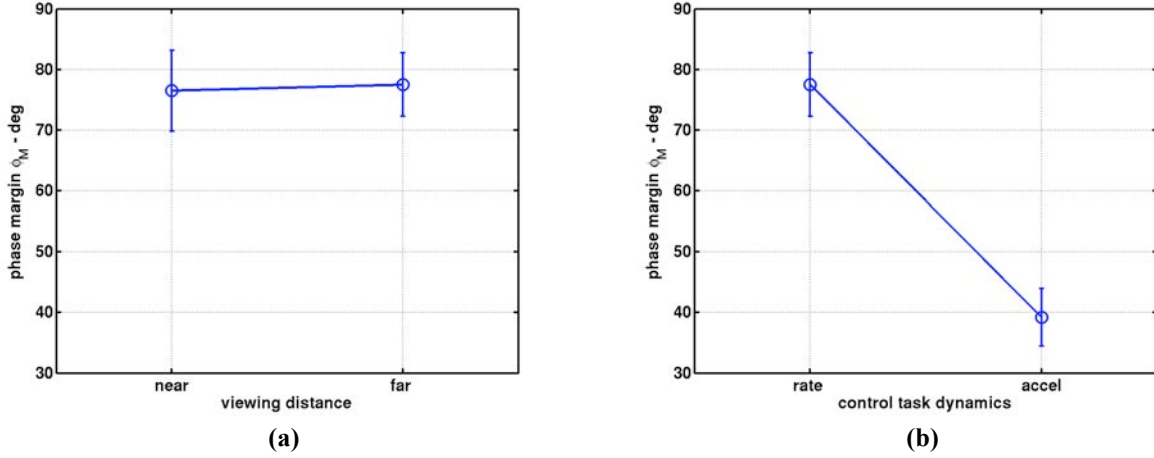


Figure 8. Phase margin ϕ_M as a function of viewing distance (a) and control task dynamics (b). Means and standard error bars are shown.

3. Individual Models

Parameters of the model shown in Fig. 3 can be chosen to best correspond with the describing function measurements, \hat{H}_T , \hat{H}_{RS} , and \hat{H}_{SD} (Eqs. 22-24). Using Eqs. (12), (17), (20) and (21), modeled describing functions H_T , H_{SD} , and H_{RS} can be related to the model parameters as follows:

$$H_T = -Y_N(s K_V + K_P) \quad (22)$$

$$H_{SD} = - (Y_N/\Delta)(s K_V W_{VS} + K_P W_{PS}) \quad (23)$$

$$H_{RS} = - (Y_N/\Delta)\{s K_V (1 - W_{VS}) + K_P (1 - W_{PS})\} \quad (24)$$

$$\text{where } \Delta = 1 + Y_C Y_N (s K_V + K_P) \quad (25)$$

In the proposed model, W_{VS} , W_{PS} , K_V , and K_P are all scalar elements. Y_N is a parameterized transfer function of the form:

$$Y_N = \exp(-s \tau) / (s^2 / \omega_N^2 + 2s\zeta_N / \omega_N + 1) \quad (29)$$

The term Y_N represents the combination of the neuromotor limb dynamics and control effector. This form was chosen because it generally provided good correspondence with the data. Note that the only time delay present in the model is captured by Y_N ; this time delay is meant to represent the sum of the perceptual and motor delays present in the system. This representation is mathematically equivalent to putting a separate perceptual delay directly “downstream” of the display, and was done to simplify the model identification.

The parameters of these models, specifically τ , ω_N , ζ_N , W_{VS} , W_{PS} , K_V , and K_P , were determined to best fit the measurements for each operator and condition; see Fig. 9 for one particular example. The resulting parameters are discussed in the following sections.

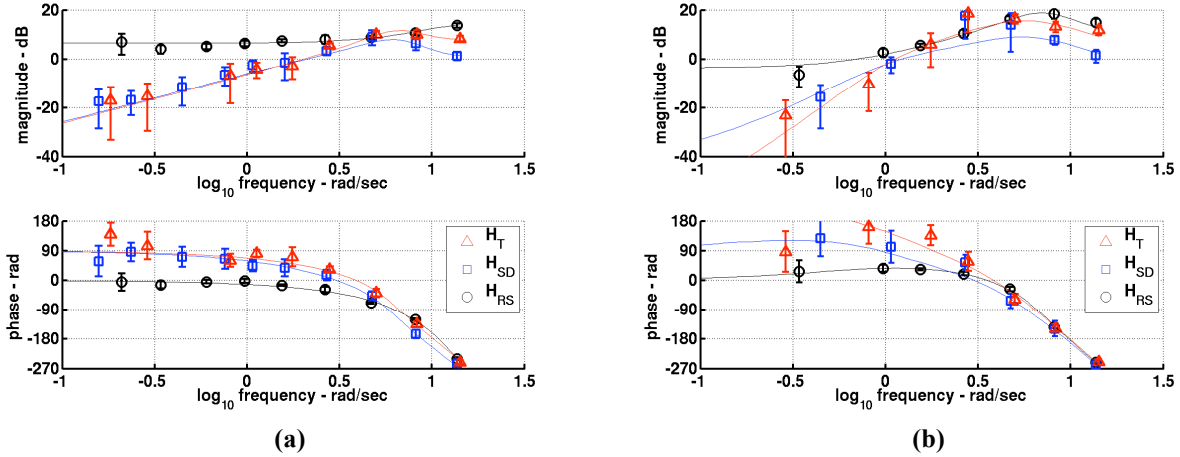


Figure 9. Example plots showing measured describing functions (plotting symbols) and parameterized models (solid line) for one operator. Rate control (a) and acceleration control (b) cases are shown, both at the near viewing distance. Means and standard error bars are shown.

Velocity and Position Perception Parameters - The parameters in the model shown in Fig. 3 relating to velocity and position perception are W_{VS} and W_{PS} , respectively. W_{VS} is the weighting factor on stereo disparity for velocity perception, and W_{PS} is the weighting factor on stereo disparity for position perception. Because the model structure assumes that the sum of the inputs will produce a unity gain, the weightings of relative size are determined by the weightings on stereo disparity. Figure 10 shows the mean values derived for W_{PS} (position weight) and W_{VS} (velocity weight). It is clear that position estimation is more dependent upon stereo disparity (i.e., $W_{PS} > 0.5$), and velocity estimation is more dependent upon relative size ($W_{VS} < 0.5$). Statistical analysis showed this to be a significant effect ($F[1,7] = 37.023$, $p < 0.0005$) across all conditions. In the rate-control conditions, the weighting on stereo disparity decreased for the far viewing condition, for both velocity and position perception. This main effect of viewing condition was shown to be significant ($F[1,7] = 16.999$, $p < 0.004$). In comparing the acceleration-control and rate-control conditions, the only significant effect seen in the perception parameters was that the weights of stereo disparity associated with position perception (W_{PS}) are greater than the weights associated with velocity perception (W_{VS}). There was no significant correlation between the participant’s measured static stereo acuity and their W_{VS} or W_{PS} .

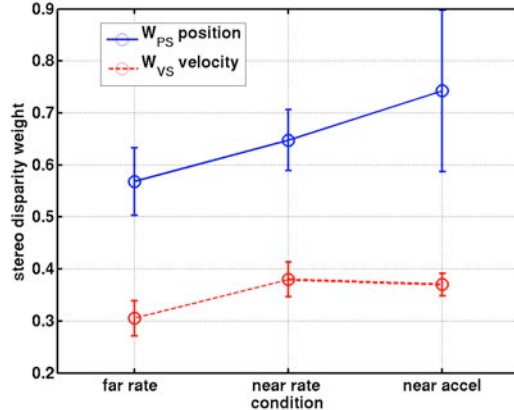


Figure 10. Weightings on stereo disparity for position and velocity. Means and standard error bars are shown.

Control Parameters - The parameters that specify the control strategy of the operator are the gain on velocity, K_V , and the gain on position, K_P . Fig. 11 shows these values, as well as the ratio between them, for all operators and conditions. The ratio K_V / K_P has a special meaning in control engineering, and is referred to as “lead.” A term of the form $sK_V + K_P$ is labeled a “lead network.” This is because the output of a circuit with this transfer function would “lead” the input, in phase, because its output is proportional to not only the input, but also to the input velocity. For large values of K_V ($K_V > K_P$), lead is high, and the output largely proportional to the velocity of the input. For the converse case ($K_V < K_P$), lead is low, and the output is largely proportional to the input. Previous work in manual control has shown that for acceleration-control dynamics, the operator needs to generate lead to achieve acceptable levels of closed-loop performance. This is clearly demonstrated in Fig. (11b); for all operators and conditions, lead is dramatically higher for the near acceleration condition.

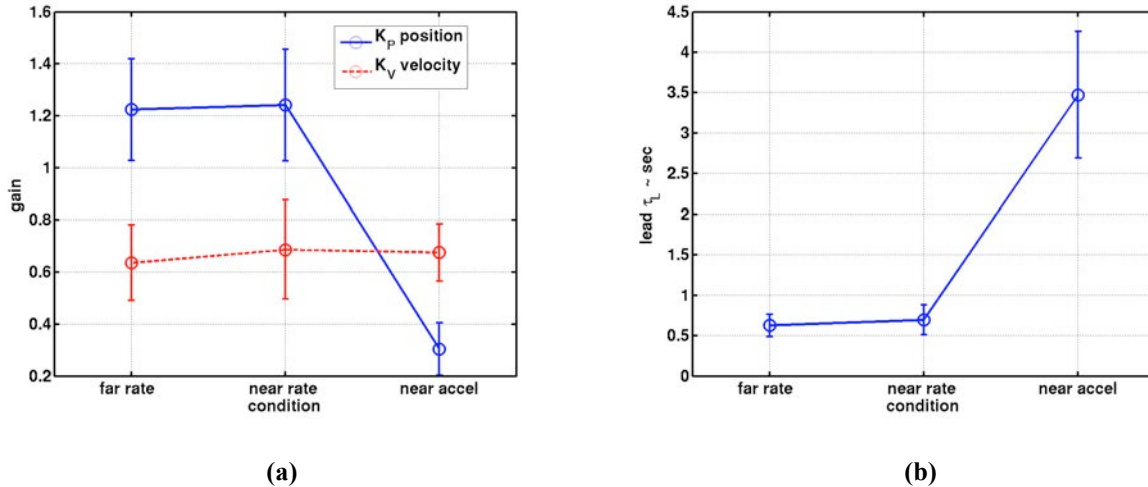


Figure 11. Control gain parameters. Control gain on position (K_P) and velocity (K_V) are shown in (a). The ratio K_V/K_P which specifies lead is shown in (b). Means and standard error bars are shown.

Neuromuscular Parameters - The neuromuscular dynamics function Y_N as defined in Eq. (29), consists of a second-order system in the numerator (assumed to be related to the neuromuscular dynamics of the operator) and a pure time delay. The mean values derived for time delay τ are shown in Fig. 12; the mean values for natural frequency (ω_N) and damping (ζ_N) are shown in Fig. 13. The observed range of natural frequencies are below the typically observed range of 15 to 20 rad/sec, but similar results have been obtained before²⁷. It has been demonstrated that identified values in this range can result from a “pulsive” control strategy, which some operators are known to adopt²⁸. Although the second-order neuromuscular system is required to obtain accurate fits to the measurements, it has little direct effect on the closed-loop system performance. This is because the frequency range in which it operates was above the observed crossover frequency in all cases (the highest observed crossover frequency was 2.45 rad/sec). Time delays clustered in the range of 0.23 to 0.31 seconds, which is well within the expected range from previous manual control studies²³.

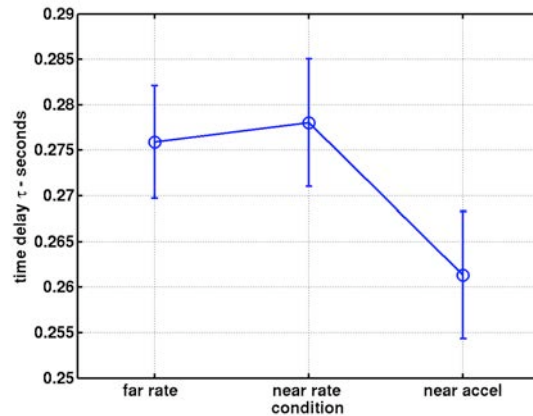


Figure 12. Time delay parameter τ . Means and standard error bars are shown.

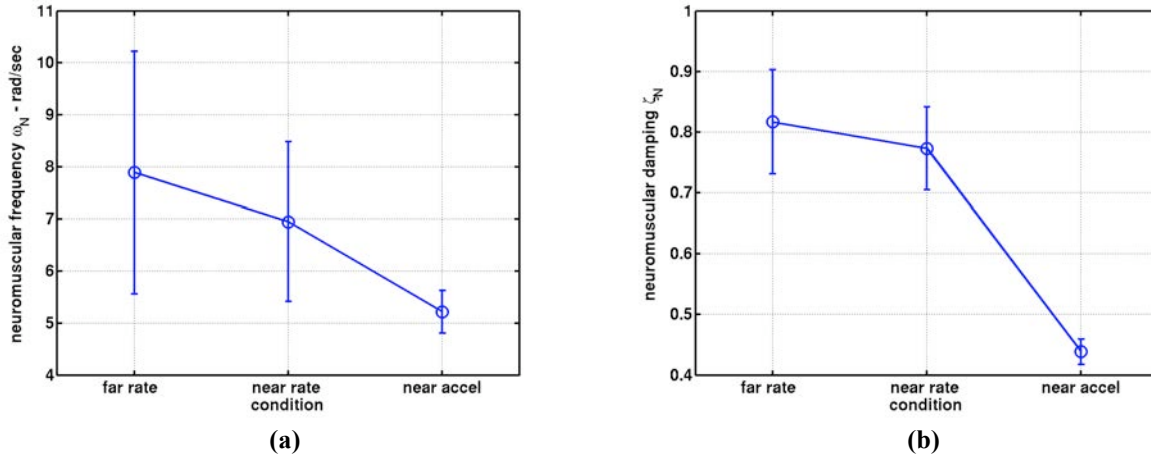


Figure 13. Parameters for neuromuscular natural frequency ω_N (a) and damping ζ_N (b). Means and standard error bars are shown.

Overall Fit Quality - In most cases this simple model could be adjusted to achieve a relatively high quality of fit to the measurements. For each operator and condition, the magnitude and phase of the model corresponding to the frequencies of the measurements was evaluated. Figure 14 shows the absolute values of the differences between the models and measurements, in magnitude (Fig. 14a) and phase (Fig. 14b). As can be seen the overall fit quality is quite good. For each condition and operator, a relatively low number of parameters (seven total) were used to fit up to 60 measurements (3 describing functions \times 10 frequencies \times 2 values per complex measurement). The correspondence of the combined size and disparity signal measurement H_T is clearly superior to the other measurements; this is likely because the measurement was made with a much higher amplitude of input signal (x , as opposed to y and z). This yields a higher signal-to-noise ratio, resulting in smaller variances in measurement.

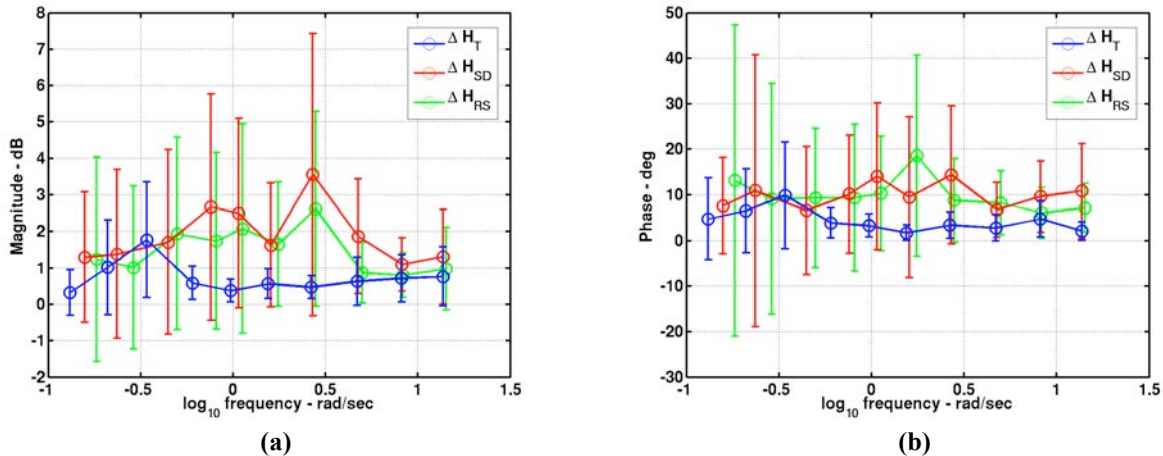


Figure 14. Mean absolute differences in magnitudes (a) and phases (b) of the ratio of the measurements to models for the three functions H_T , H_{RS} and H_{SD} , for all operators and conditions. Means and standard deviation bars are shown.

Effect of Disturbances – Before discussing modeling parameters, the effect of the independent disturbances will be addressed. In order to derive the effects of relative size and stereo disparity independently, it was necessary to input independent disturbances on these variables. However, varying stereo disparity cues and size cues independently without correlation between them would yield the perception of an object that is not only varying in depth, but also varying in its true size. The goal in designing the study was to include independent disturbances that would be large enough to measure the independent effects of the disturbances, but not so large that the operator failed to perceive the Target as a constant-size object varying in depth. Measurements of describing functions for a baseline disturbance case (i.e. no independent disturbances to size and disparity) were made and compared with the

multiple disturbance cases. Figure 15 shows the difference between the measured describing function \hat{H}_T between the Multiple Disturbance and Baseline Disturbance conditions, for both the rate-control and acceleration-control dynamics. Overall, on average operators exhibited higher magnitude and phase for the multiple disturbance condition. This slight difference is probably due to the perceived higher bandwidth of the input disturbances in the multiple disturbance condition. However, these changes are relatively small, indicating that the control task was not perceptually different with the addition of the independent disturbances to relative size and stereo disparity.

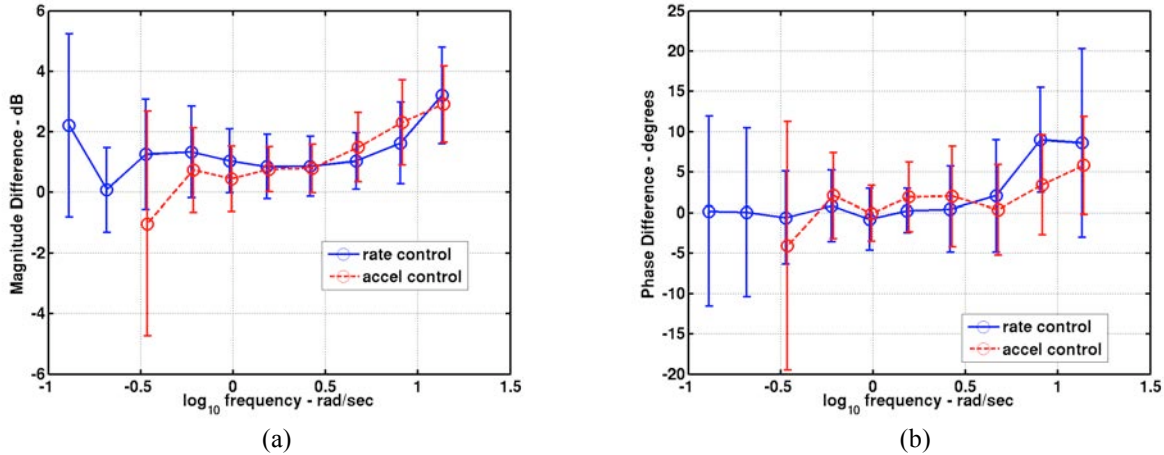


Figure 15. Difference in magnitude (a) and phase (b) of measured describing function H_T between multiple disturbance and baseline disturbance cases. Means and standard deviation bars are shown.

C. Discussion

The purpose of this study was to develop a model of depth-cue integration for a closed-loop manual control task. The model consists of three basic components: perception, control, and neuromuscular dynamics. The control and neuromuscular dynamics portions of our model are derived from manual control research. The perception component is similar to the Additive models previously advanced by Bruno and Cutting¹¹, and Clark and Yuille²⁹. However, the current model is somewhat more complex than the Additive models in that position and velocity perception are considered to be different processes. Both velocity and position perception are modeled as Additive systems, but these two systems are allowed to operate independently. This model is highly effective at describing the input/output relationships of the human operator.

When the model was tested with our data, both the neuromuscular dynamics and control portions of the model behaved in ways consistent with the existing body of manual control research. The neuromuscular dynamics were generally represented with a second-order system and a time delay; because the frequency of the second-order system typically was well above the crossover frequency of the closed-loop system, it does not particularly impact the closed-loop system performance, and thus will not be discussed further. The control portion of the model consisted of a lead element, specifically a weighted summation of position and velocity signals. As previous manual control findings would predict, in rate-control tasks, the control output was dominated by position feedback. In acceleration-control tasks, the output was dominated by velocity feedback (see Fig. 11a). There was effectively no change in the model control parameters due to the manipulation of viewing distance; only the manipulation of control task type affected these parameters of the model.

The perception parameter fits of the model revealed some interesting characteristics. First, the perception of position was more dependent upon stereo disparity, and perception of velocity was more dependent upon relative size (refer to Fig. 10). This effect was seen in all conditions, both near and far viewing distance, and both rate and acceleration control. Second, the perception parameters changed significantly when the viewing distance changed; both position and velocity perception became more reliant upon relative size. Because the viewing distance manipulation did not affect the magnitude of the relative size cue (in visual angle), and did diminish the stereo disparity cue, it follows that operators modified their depth cue integration strategy when the stereo disparity cue became less salient. The perception parameters were *not* affected significantly by the task dynamics (rate-control vs acceleration-control). Additionally, the perception parameters showed no correlation with static stereo acuity scores of the participants; good static stereo acuity did not imply more reliance upon stereo disparity as a cue.

The ANOVA analyses on the outcome variables (depth error rms, percent of control activity correlated with stereo disparity and relative size disturbances) are consistent with the modeling results. Depth error rms increased

significantly when the viewing distance was increased, due to the fact that the stereo disparity cue becomes less useful. Depth error rms also increased significantly for the acceleration-control task; this is expected from manual control findings, because the acceleration-control task is more difficult to do than the rate-control task. The other two dependent measures were the percent of control activity correlated with the stereo disparity disturbance, and the percent of control activity correlated with the relative size disturbance. At the far viewing distance, the percent of control activity correlated with relative size increased, and the percent of control activity correlated with stereo disparity decreased. This result is completely consistent with our modeling results, which showed that the weighting on relative size increased, and weighting on stereo disparity decreased, in the far viewing condition.

IV. Experiment 2

In the previous experiment, human operator characteristics were measured in an active depth-control task. Both relative size and stereo disparity were available as depth cues. Both viewing distance and control task were varied. The results of this study indicated that depth position perception was more dependent upon stereo disparity, and depth motion perception was more dependent upon relative size. Results of the previous experiment implied that the utility of stereo disparity as a cue was potentially affected by the task to be performed. Position perception was more dependent upon stereo disparity; from manual control findings, we know that position information is necessary for rate-control tasks. Motion perception was less dependent upon stereo disparity; from manual control findings, we know that motion information is necessary to perform acceleration-control tasks. Together, these facts imply that stereo disparity would provide a larger performance benefit for rate-control tasks than acceleration-control tasks. Additionally, because providing stereo disparity typically reduces update rate, there was a potential for this to compromise the motion perception. This experiment was designed to examine the performance trade-offs of stereo displays and update rates with different active-control tasks.

A. Method

This experiment was conducted in two phases, termed A and B. The major difference between the phases was the update rate. In Phase A, effective stereo update rates of 48 and 24 Hz were used; in Phase B, effective update rates of 24 and 12 Hz were used.

1. Participants

In each phase, eight male, general aviation pilots participated. They were recruited from a paid contractor pool at Ames Research Center. All had normal or corrected-to-normal visual acuity and good stereo vision (40 seconds of arc or better). Participants in Phase A ranged from 25 to 45 years of age, and their flight experience ranged from 100 to 4500 hours. Participants in Phase B ranged from 23 to 43 years of age, and their flight experience ranged from 210 to 2200 hours. Because of the time period over which the phases were run (Phases A and B were conducted approximately a month apart), some of the participants in the Phase A study were not available for Phase B. Rather than run Phase B with a mixed-experience group of participants, we chose a naïve group to participate in Phase B.

2. Apparatus

The apparatus was identical to that used in Experiment 1, as described in Section III.A.2.

3. Stimuli and Control Tasks

The stimulus and task were identical to those used in Experiment 1. Other aspects of the displays, simulation and control task were varied as a function of experimental condition. The three experimental factors were: stereo; control task dynamics; and update rate. We discuss each of these in turn.

Stereo - In the stereo conditions, participants were instructed to put on stereo glasses, and the computer scene was rendered for each eye alternately, based on a viewing distance from the screen of 22 inches, and an eyepoint separation of 3 inches. In the non-stereo conditions, the participants did not wear the shutter glasses, and the scene was rendered without stereo.

Control Task Dynamics - The control task dynamics were identical to those used in Experiment 1, as described in Section III.A.3.

Update Rate - Three update rates were tested, termed Fast, Medium, and Slow. In Phase A, the Fast and Medium conditions were tested. In Phase B, the Medium and Slow conditions were tested. In the Fast condition, the scenes were computed and updated at a rate of 96 Hz. In the non-stereo conditions, this corresponds to an update rate of 96 Hz. In the stereo condition, because two frames are necessary to complete an update cycle (frames for each eye are independently rendered), the corresponding update rate was 48 Hz. In the Medium condition, the scenes were computed and rendered at a rate of 48 Hz; the corresponding update rates were 48 Hz in the non-stereo condition, and 24 Hz in the stereo condition. In the Slow condition, the scenes were computed and rendered at a rate of 24 Hz; the corresponding update rates were 24 Hz in the non-stereo condition, and 12 Hz in the stereo condition.

The control information was input to the simulation at the update rate of the condition. In ALL cases, the monitor refreshed at a rate of 96 Hz.

4. Design

As stated above, our three experimental factors were: 1) Stereo Disparity (disparity versus no disparity); 2) Control Task Dynamics (Rate versus Acceleration); and 3) Update Rate (Fast versus Medium in Phase A, Medium versus Slow in Phase B). Multiple trials were run within each condition. Each block of trials consisted of a training trial and four data trials. The order of trial blocks was pseudo-randomized across participants.

5. Procedure

Participants were given written instructions describing their task and its discernable variations (i.e., Rate and Acceleration control dynamics). Participants were then given an opportunity to ask questions. Once started, the task was entirely self-paced. The experiment took one day for participants to complete. Each subject ran a total of four conditions. For each condition, the participant had eight one-minute training trials, a four-minute break, and four four-minute data trials. Participants were administered a stereo vision acuity test during one of their breaks. Each data trial lasted a total of four minutes, five seconds. Training trials lasted one minute. During training trials, operators were given feedback on their own performance. During data trials, operators were given feedback on their own performance, as well as the “best” scores of other operators.

6. Dependent Measures

Two types of dependent measures are considered. The first type consisted of performance metrics, the root mean square (RMS) of the displayed depth error and operator control activity. The second type was based on the frequency response of the operator, specifically the closed-loop crossover frequency and phase margin of the measured operator describing function.

B. Results

1. Task Performance

An ANOVA was performed with two dependent measures: 1) displayed depth error rms; and 2) control rms. This was an $8 \times 2 \times 2 \times 2$ factorial analysis, with repetitions, stereo, control task, and update rate as factors.

Phase A - The mean values of displayed depth error rms for the rate-control and acceleration-control tasks are shown in Figs. 16a and 16b), respectively. For this dependent measure, there was a main effect of control task ($F[1,7] = 16.723, p < 0.005$); displayed depth error rms was higher in the acceleration-control condition. There was a significant interaction between control task and stereo ($F[1,7] = 5.786, p < 0.05$); in the rate-control task, displayed depth error rms was significantly lower in the stereo condition than in the non-stereo condition, while in the acceleration-control task, there was no discernable difference between the stereo and non-stereo conditions. There was no significant effect of update rate.

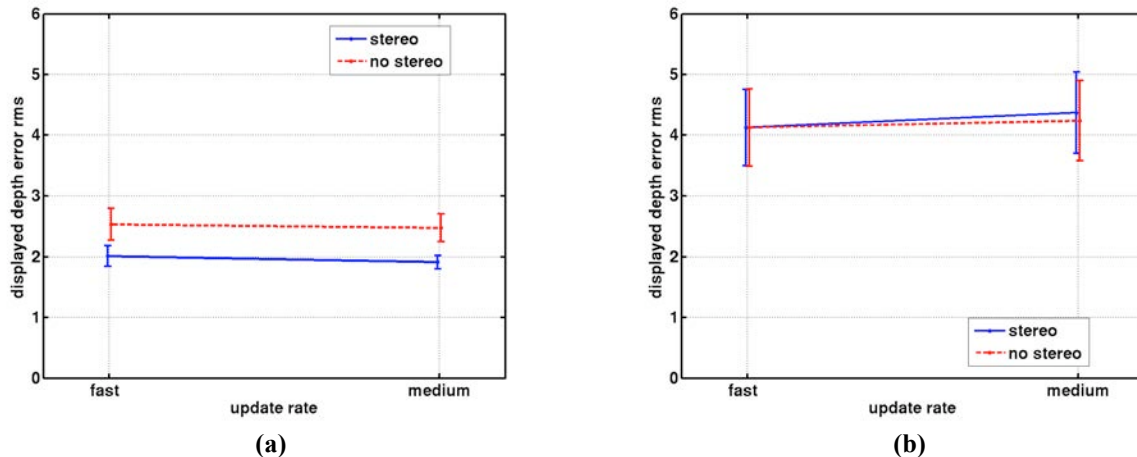


Figure 16. Displayed depth error rms with rate control (a) and acceleration control (b) for Phase A. Means and standard error bars are shown.

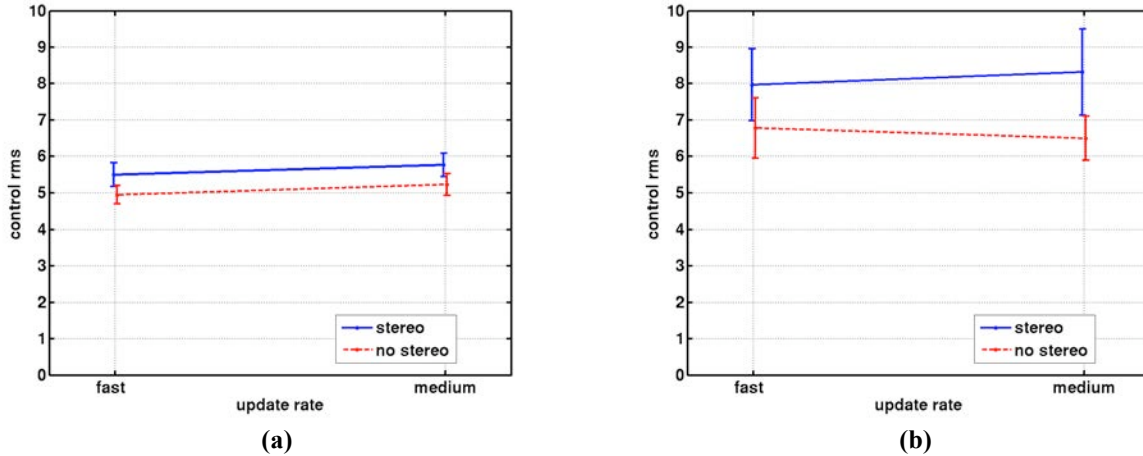


Figure 17. Control rms with rate control (a) and acceleration control (b) for Phase A. Means and standard error bars are shown.

Phase B - The mean values, with standard error bars, of displayed depth error rms are shown in Fig. 18. For this dependent measure, there was a main effect of control task ($F[1,7] = 48.161, p < 0.001$); displayed depth error rms was higher in the acceleration control condition. There was a significant interaction between control task and stereo ($F[1,7] = 17.120, p < 0.004$); in the rate-control task, displayed depth error rms was lower in the stereo condition than in the non-stereo condition, while in the acceleration control task, the error was higher in the stereo condition than in the non-stereo condition.

There was no significant effect of update rate in this analysis, although some trends appeared to be present. Two subanalyses were done, one examining only the stereo conditions, another examining only the acceleration control conditions. In the stereo-only analysis, there was a main effect of control task ($F[1,3] = 42.174, p < 0.001$), and a trend towards a main effect of update rate ($F[1,3] = 4.897, p < 0.063$). The effect of control task was the same as the main factorial analysis: acceleration control was associated with higher displayed depth error rms than rate control. The effect of update rate was that the lower update rate (slow) had higher depth error rms than the medium update rate. In the acceleration control only analysis, there was a trend toward a main effect of stereo ($F[1,3] = 5.354, p < 0.054$); the depth error rms was higher in the stereo condition than in the non-stereo condition. There was a trend towards a main effect of update rate ($F[1,3] = 4.242, p < 0.078$); displayed depth error rms was higher at the slower update rate.

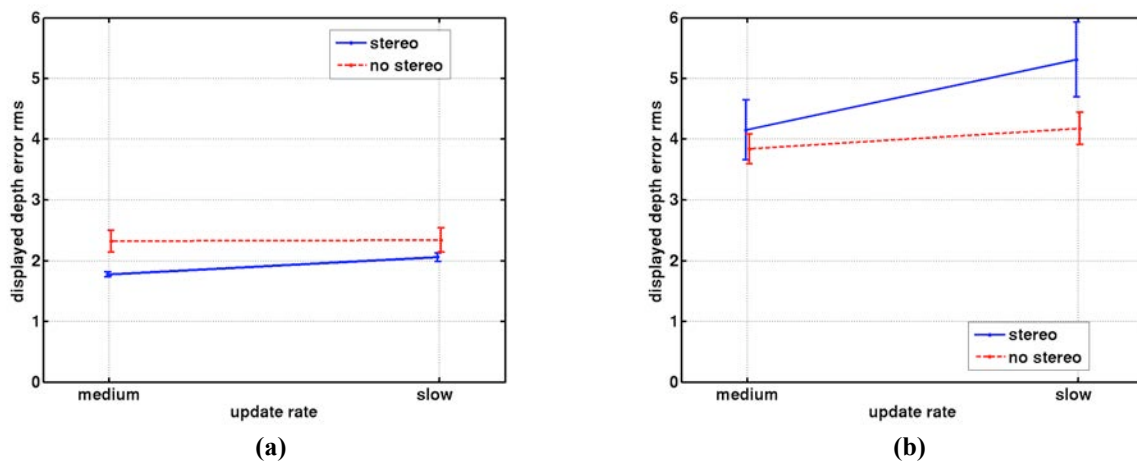


Figure 18. Displayed depth error rms with rate control (a) and acceleration control (b) for Phase B. Means and standard error bars are shown.

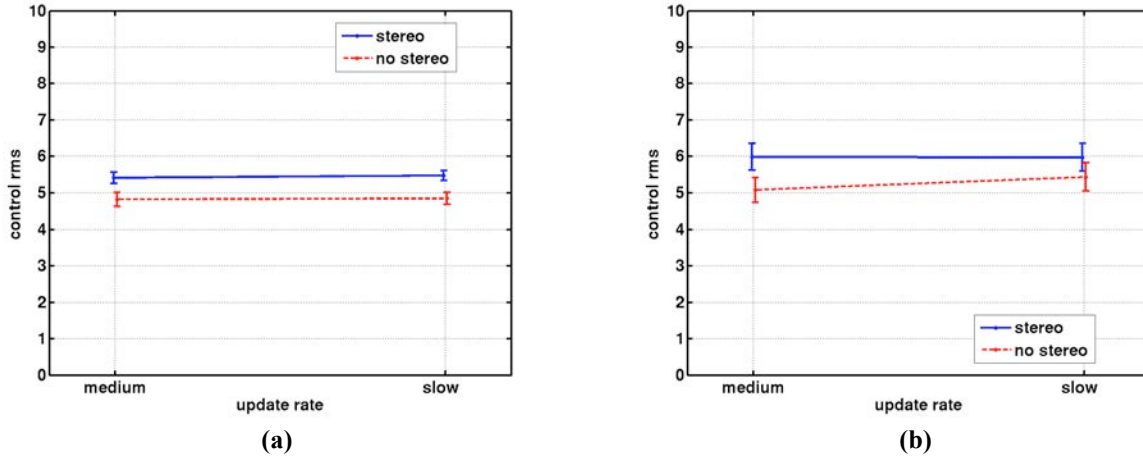


Figure 19. Control rms with rate control (a) and acceleration control (b) for Phase B. Means and standard error bars are shown.

2. Frequency Response

Crossover frequency and phase margin of the describing function for the operator and controlled element were estimated. An ANOVA was performed with crossover frequency and phase margin as dependent measures. This was an 8 x 2 x 2 x 2 factorial analysis, with repetitions, stereo, control task, and update rate as factors.

Phase A - For crossover frequency (Fig. 20), there was a main effect of control task ($F[1,7] = 26.27, p < 0.001$), with higher crossover frequencies associated with the rate-control task. There was also a main effect of stereo ($F[1,7] = 18.19, p < 0.004$), with higher crossover frequencies occurring in the stereo conditions. There was a significant interaction between update rate and control task ($F[1,7] = 10.804, p < .02$); in the rate control condition, the slower update rate was associated with higher crossover frequency than the faster update rate, while in the acceleration control condition, this pattern was reversed. A trend towards an effect of update rate was observed ($F[1,7] = 4.25, p < .078$), with higher crossover frequencies observed at the slower of the update rates. A trend towards an interaction between stereo and control task was also observed ($F[1,7] = 3.73, p < .095$); stereo was associated with a much greater increase in crossover frequency in the rate-control condition than in the acceleration-control condition.

For phase margin (Fig. 21), there was a significant main effect of control task ($F[1,7] = 102.426, p < 0.001$), with lower phase margins associated with the acceleration-control conditions. There was also a significant main effect of stereo ($F[1,7] = 25.267, p < .002$); lower phase margins were observed in the stereo conditions.

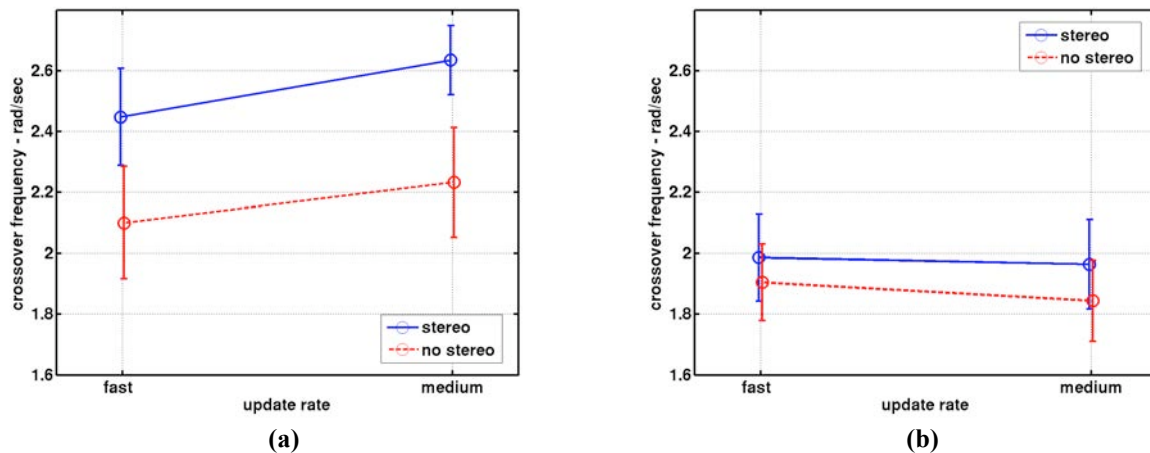


Figure 20. Crossover frequency with rate control (a) and acceleration control (b) for Phase A. Means and standard error bars are shown.

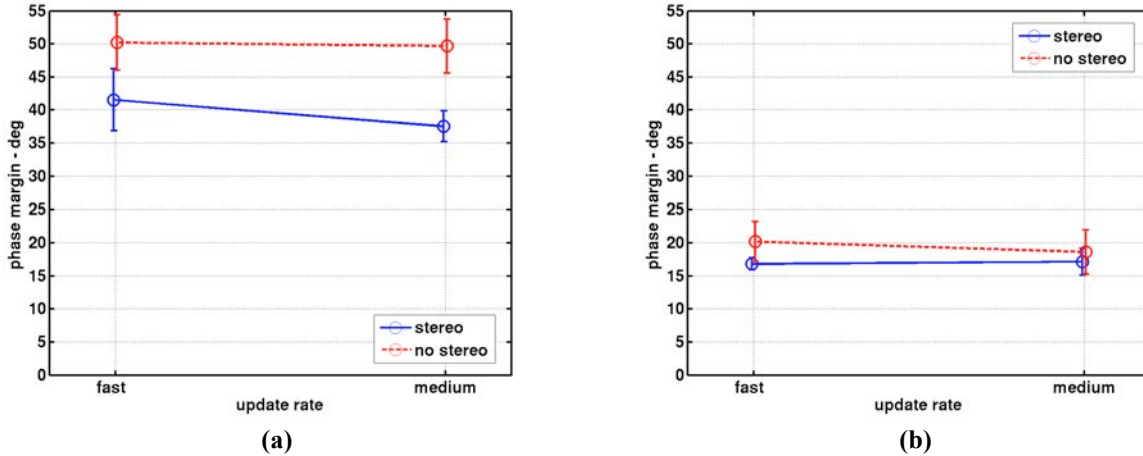


Figure 21. Phase margin with rate control (a) and acceleration control (b) for Phase A. Means and standard error bars are shown.

Phase B - For crossover frequency (Fig. 22), there was a main effect of control task ($F[1,7] = 31.857, p < 0.001$); crossover frequencies were higher in the rate-control condition. There was a main effect of update rate ($F[1,7] = 10.175, p < 0.02$); overall, the slower update rate was associated with lower crossover frequencies. There was a significant interaction between stereo and control task ($F[1,7] = 8.226, p < 0.03$); stereo had a large effect on crossover frequency in the rate control condition (higher with stereo), and no effect in the acceleration-control condition. There was also a significant interaction between update rate and stereo ($F[1,7] = 15.053, p < 0.006$); in stereo conditions the slower update rate was associated with a decrease in crossover frequency; in non-stereo conditions, the opposite trend is present.

For phase margin (Fig. 23), there was a main effect of control task ($F[1,7] = 58.153, p < 0.001$); lower phase margins were associated with the acceleration-control condition. There was a main effect of stereo ($F[1,7] = 14.841, p < 0.006$); lower phase margins were observed in the stereo condition. There was a significant interaction between update rate and stereo ($F = 24.921, p < 0.002$); in stereo conditions, the slower update rate was associated with increased phase margin, while in the non-stereo conditions this pattern was reversed. There was also a significant interaction between stereo and control task ($F[1,7] = 12.219, p < 0.010$); in the rate-control task, the non-stereo conditions had a much higher phase margin than the stereo-conditions; in the acceleration-control task, there was no difference.

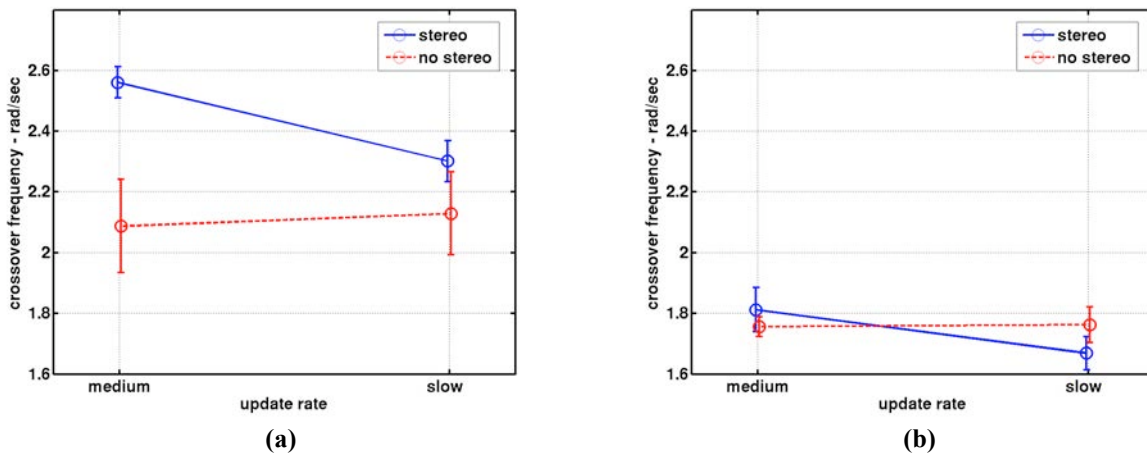


Figure 22. Crossover frequency with rate control (a) and acceleration control (b) for Phase B. Means and standard error bars are shown.

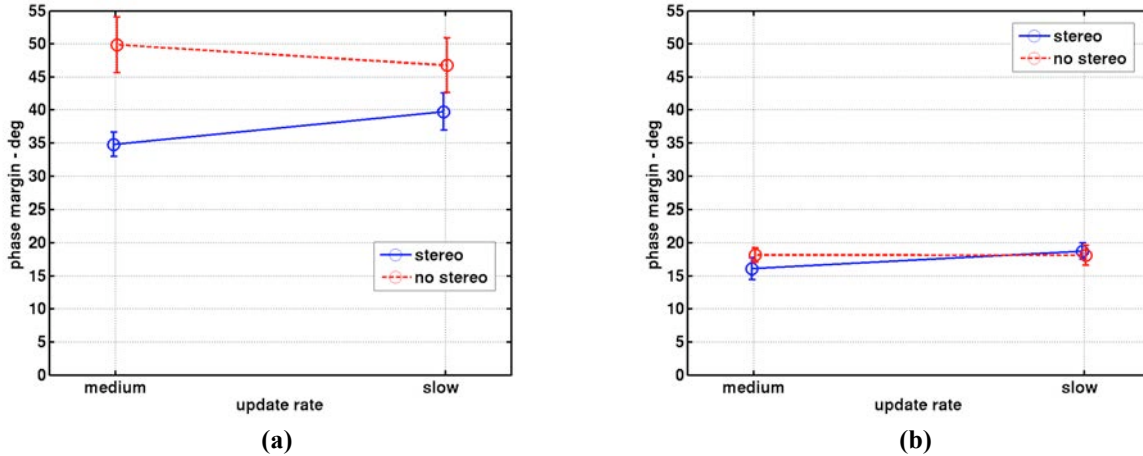


Figure 23. Phase margin with rate control (a) and acceleration control (b) for Phase B. Means and standard error bars are shown.

C. Discussion

As had been predicted as a result of previous modeling of depth-cue integration, the presence of stereo has a much greater benefit in the rate-control task than in the acceleration-control task. It is interesting to note that the presence of stereo in the display led to higher crossover frequencies, lower phase margins, and higher control activity for *both* control task types. These three parameters are all interrelated; higher crossover frequencies are the result of higher gains from the operator, which directly results in increased control activity. Higher gains also lead to reductions in phase margin. In the rate-control condition, the higher operator gains were associated with an improvement in performance (displayed depth error rms). However, in the acceleration-control condition, higher gains did not yield an improvement in performance. This is likely due to the fact that the operators were working in a region where the closed-loop (man-machine) system is close enough to its stability margin that increases in gain do not improve performance.

One somewhat surprising result from the crossover frequency analysis of Phase A was the interaction between update rate and control task. In the rate-control condition, the Medium update rate was associated with slightly higher crossover frequencies than the Fast update rate, while in the acceleration-control task, the Medium update rate was associated with lower crossover frequencies. While this effect was small and did not express itself with significant effects on the other measures, it is interesting to consider why it might be present. Because the operators need good position sensing in the rate-control task, the Fast update rate might have made the display appear overly active and the operators might have felt more comfortable adjusting their gains downward. In the acceleration-control condition, the operator needs to sense and feed back velocity, and the Fast update rate might have produced better perception of velocity.

Although no effect from update rate was observed in Phase A, effects became apparent in Phase B. In this phase (with acceleration-control dynamics), stereo was associated with a decrement in performance, the greatest decrement occurring at the Slow update rate. In Phase A, stereo was associated with higher control activities in all the conditions. In Phase B, for the rate-control task, this increased control activity (and higher crossover frequency) produced improved performance; in the acceleration-control task, the higher control activity with stereo actually led to worse performance.

V. Conclusion

Our depth-cue integration and control model accurately characterizes the activity of the operator over a range of tasks. This model incorporates control and neuromuscular dynamics from previous manual control work with perceptual models suggested by depth-cue integration paradigms. The modeling results suggest that the depth-cue integration strategy of the operator changes as a function of the saliency of the available cues, but does not change as a function of the control task dynamics. The modeling also suggests that the operator depends more on stereo disparity than relative size for position perception, and more on relative size than stereo disparity for velocity perception. As predicted by manual control theory, the operator uses more velocity information with acceleration-control dynamics than with rate-control dynamics.

Because the operator uses more velocity feedback in acceleration-control tasks, and because velocity

perception is more dependent upon relative size than stereo disparity, the results of Experiment 1 imply that stereo disparity could be a much less useful cue in acceleration-control tasks. Conversely, because accurate position information is necessary for rate-control tasks, and stereo disparity dominates position perception, stereo disparity is probably a highly useful cue for rate-control tasks. Experiment 2 was conducted to test these predictions. This experiment confirmed the prediction that the utility of stereo disparity in a display was a function of the type of task. On a task that requires good depth-position information, performance benefits from the addition of stereo disparity. On a task that also requires good depth-velocity information, performance does not benefit from stereo disparity. When the method used to generate stereo disparity also reduces effective update rate of the display, performance can suffer if the update rate drops too low; for this experiment, performance decrements were noted when update rate dropped to 12 Hz. Additionally, workload (as evidenced by control activity) was increased when stereo disparity was provided in the task that required good velocity information. For manual control tasks, information requirements should be carefully considered before adding stereo disparity as a cue source.

Appendix

A. Vehicle Dynamic Simulation

The following equation was used for the rate-control dynamics:

$$\dot{w} = K_{\delta} \delta \quad (\text{A1})$$

For the acceleration-control dynamics, the equation was:

$$\ddot{w} = K_{\delta} \delta \quad (\text{A2})$$

In both cases, the disturbances affected the stereo disparity and relative size as follows:

$$d = w + x \quad (\text{A3})$$

$$d_{SD} = d + y \quad (\text{A4})$$

$$d_{RS} = d + z \quad (\text{A5})$$

The position, d , is in units of inches. The control input of the operator is δ ; the maximum range achievable was from -1 to 1. The scaling factor on the control was adjusted depending upon the condition. For the rate-control conditions in the near position, K_{δ} was set to a value of 20. For the rate-control, far position condition, the value was 30. This was done to keep the sensitivity to the changes in visual angle constant. For the acceleration-control condition, the value was set to 10. d_{SD} and d_{RS} are the depths used to graphically render stereo disparity and relative size, respectively. In implementation, the object was drawn at a depth corresponding to d_{SD} ; then the size of the object was scaled to be consistent with the relative size depth, d_{RS} .

These state-space equations were converted to discrete form for real-time simulation with a sampling interval of $T = 1/48$ seconds³⁰. The resulting discrete state-space equations were:

Rate Control: $w_{k+1} = w_k + T K_{\delta} \delta_k \quad (\text{A6})$

Acceleration Control: $w_{k+1} = 2 w_k - w_{k-1} + (T/2) K_{\delta} (\delta_k + \delta_{k-1}) \quad (\text{A7})$

$$d_k = w_k + x_k \quad (\text{A8})$$

$$d_{SDk} = d_k + y_k \quad (\text{A9})$$

$$d_{RSk} = d_k + z_k \quad (\text{A10})$$

B. Disturbances

The disturbances x , y , and z had the following form as a function of time (t):

$$x(t) = \sum_{i=1}^{10} a_{xi} \sin\left(\frac{2\pi k_{xi}}{240} t + \rho_{xi}\right) \quad (\text{A11})$$

$$y(t) = \sum_{i=1}^{10} a_{yi} \sin\left(\frac{2\pi k_{yi}}{240} t + \rho_{yi}\right) \quad (\text{A12})$$

$$z(t) = \sum_{i=1}^{10} a_{zi} \sin\left(\frac{2\pi k_{zi}}{240} t + \rho_{zi}\right) \quad (\text{A13})$$

The disturbance spectra were designed to conform to guidelines for pilot frequency response identification^{22,23}. Values for k were chosen to create frequency values that were logarithmically spaced between approximately 0.15 and 15 rad/sec. The actual values of a , k , and resulting frequencies ($\omega = 2\pi k/240$) are shown in Table 1 for the three disturbance spectra.

i	k_{xi}	a_{xi}	ω_{xi}	k_{yi}	a_{yi}	ω_{yi}	k_{zi}	a_{zi}	ω_{zi}
1	5	3.0	0.13	6	0.3	0.16	7	0.3	0.18
2	8	3.0	0.21	9	0.3	0.24	11	0.3	0.29
3	13	3.0	0.34	17	0.3	0.45	19	0.3	0.50
4	23	3.0	0.60	29	0.3	0.76	31	0.3	0.81
5	37	3.0	0.97	41	0.3	1.07	43	0.3	1.13
6	59	3.0	1.54	61	0.3	1.60	67	0.3	1.75
7	101	0.3	2.64	103	0.3	2.70	107	0.3	2.80
8	179	0.3	4.69	181	0.3	4.74	191	0.3	5.00
9	311	0.3	8.14	313	0.3	8.19	317	0.3	8.30
10	521	0.3	13.64	523	0.3	13.69	541	0.3	14.16

Table A1. Disturbance spectra magnitudes and frequencies. The frequencies (ω_{xi} , ω_{yi} , ω_{zi}) are expressed in radians/sec.

The phase offsets (ρ) for each repetition and disturbance (x , y and z) were precomputed with a random number generator, randomly distributed from $-\pi$ to π .

Acknowledgments

This research was supported by the Human Measurement and Performance Project of NASA's Airspace Operations Program. The authors thank Michael Logan and Wendy Davis for programming and logistical support of the experiments.

References

- ¹Bruce, V., Green, P. R., and Georgeson, M. A., *Visual Perception: Physiology, Psychology, and Ecology*, John Wiley and Sons, New York, 1996.
- ²Boff, K. R., and Thomas, J. P. (eds.), *Handbook of Perception and Human Performance*, Vol. 1., John Wiley and Sons, New York, 1986.
- ³Sedgwick, H.A., "Space Perception," *Handbook of Perception and Human Performance*, Vol. 1, edited by K. R. Boff and J. P. Thomas, John Wiley and Sons, New York, 1986, Chap. X.
- ⁴Hochberg, J., "Representation of motion and space in video and cinematic displays," *Handbook of Perception and Human Performance*, Vol. 1, edited by K. R. Boff and J. P. Thomas, John Wiley and Sons, New York, 1986, Chap. XX.

- ⁵Arditi, J., "Binocular vision," *Handbook of Perception and Human Performance*, Vol. 1, edited by K. R. Boff and J. P. Thomas, John Wiley and Sons, New York, 1986, Chap. XXX.
- ⁶Berkeley, G., (1709/1910). "An essay toward a new theory of vision", *A New Theory of Vision and Other Writings*. London: J.M. Dent and Sons, London, 1709/1910.
- ⁷Foley, J. M., "Binocular distance perception," *Psychological Review*, Vol. 87, 411-435, 1980.
- ⁸Richards, W., "Anomalous stereoscopic depth perception," *Journal of the Optical Society of America*, Vol. 61, pp. 410-414, 1971.
- ⁹Cutting, J. E., and Vishton, P. M., "Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth," *Handbook of Perception and Cognition: Volume 5, Perception of Space and Motion*, edited by W. Epstein & S. Rogers, Academic Press, New York, 1995.
- ¹⁰Landy, M. S., Maloney, L. T., Johnston, E. B., Young, "Measurement and modeling of depth cue combination: In defense of weak fusion," *Vision Research*, Vol. 35, No. 3, pp. 389-412, 1995.
- ¹¹Bruno, N., and Cutting J. E., "Minimodularity and the perception of layout," *Journal of Experimental Psychology: General*, Vol. 117, No. 2, pp. 161-170, 1988.
- ¹²Bulthoff, H. H., and Mallot, H. A., "Integration of depth modules: Stereo and shading," *Journal of the Optical Society of American A: Optics, Image Science, and Vision*, Vol. 5, No. 10, 1749-1758, 1988.
- ¹³Johnston, E. B., Cumming, B. G., and Parker, A. J., "Integration of depth modules: Stereopsis and texture," *Vision Research*, Vol. 33, No. 5/6, pp. 813-826, 1993.
- ¹⁴Young, M. J., Landy, M. S., and Maloney, L. T., "A perturbation analysis of depth perception from combinations of texture and motion cues," *Vision Research*, Vol. 33, No. 18, pp. 2685-2696, 1993.
- ¹⁵Stevens, K. and Brooks, A., "Integrating stereopsis with monocular interpretations of planar surfaces," *Vision Research*, Vol. 28, pp. 371-386, 1988.
- ¹⁶Massaro, D. W., "Ambiguity in perception and experimentation," *Journal of Experimental Psychology: General*, Vol. 117, No. 4, pp. 417-421, 1988.
- ¹⁷Massaro, D. W., and Cohen, M. M., "The paradigm and the Fuzzy Logic Model of Perception are alive and well," *Journal of Experimental Psychology: General*, Vol. 122, No. 1, pp. 115-124, 1993.
- ¹⁸Bradshaw, M. F., and Rogers, B. J., "The interaction of binocular disparity and motion parallax in the computation of depth," *Vision Research*, Vol. 36, No. 21, pp. 3457-3468, 1996.
- ¹⁹Curran, W., and Johnston, A., "Integration of shading and texture cues: Testing the linear model," *Vision Research*, Vol. 34, No. 14, 1863-1874, 1994.
- ²⁰Johnston, E. B., Cumming, B. G., and Landy, M. S., "Integration of stereopsis and motion shape cues," *Vision Research*, Vol. 34, No. 17, pp. 2259-2275, 1994.
- ²¹Sweet, B. T., "The Identification and Modeling of Visual Cue Usage in Manual Control Task Experiments," NASA TM-1999-208798, 1999.
- ²²McRuer, D., E. Krendel, E., and Reisener, W. Jr., "Human Pilot Dynamics in Compensatory Systems," AFFDL-TR-65-15, 1965.
- ²³McRuer, D. T., and E. S. Krendel, "Mathematical Models of Human Pilot Behavior," AGARD-AG-188, 1974.
- ²⁴Hess, R. A. (1997). "Feedback Control Models - Manual Control and Tracking," *Handbook of Human Factors and Ergonomics*, 2nd ed., edited by G. Salvendy, John Wiley & Sons, New York, 1997.
- ²⁵Levison, W. H., "Measurement of Human Operator Response Behavior," NASA CR-166038, 1980.
- ²⁶Levison, W. H., "Some computational techniques for estimating human operator describing functions," *Proceedings of the Twenty-First Annual Conference on Manual Control*, NASA CP 2428, 1986.
- ²⁷Stapleford, R.L., Craig, S. J., and Tennant, J. A., "Measurement of Pilot Describing Functions in Single-Control Multiloop Tasks," NASA CR-1238, January 1969.
- ²⁸Hess, R. A., "A Rationale for Human Operator Pulsive Control Behavior," *Journal of Guidance and Control*, Vol. 2, No. 3, pp. 221-227, 1979.
- ²⁹Clark, J. J., and Yuille, A. L., *Data Fusion for Sensory Information Processing Systems*, Kluwer, Boston, 1990.
- ³⁰Franklin, G. F., Powell, J. D., and Workman, M. L., *Digital Control of Dynamic Systems*, Addison-Wesley, Reading, 1998.