Manual Control Behavior in Stereoscopic Vision-Enhanced Depth Control Tasks

Kemna, Maarten; Pool, Daan; Wentink, Mark; Mulder, M.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Manual Control Behavior in Stereoscopic Vision-Enhanced Depth Control Tasks

Maarten H.H. Kemna*, Daan M. Pool†
Delft University of Technology
Delft, The Netherlands

Mark Wentink‡
multiSIM B.V.,
The Netherlands

Max Mulder§
Delft University of Technology
Delft, The Netherlands

Numerous critical manual teleoperation tasks, such the control of the refueling boom during aerial refueling, require human controllers to accurately manipulate objects in the depth dimension, i.e., aligned with the viewing direction. To better understand the intricacies of depth control tasks and to be able to better support human controllers in such tasks, a cybernetic analysis of human control behavior in stereoscopic vision-enhanced depth control tasks would be a valuable extension of the current state-of-the-art in manual control research. This paper presents the initial findings of a human-in-the-loop experiment in which participants performed an abstract pursuit tracking task in which multisine target and disturbance forcing functions were used to facilitate cybernetic analysis of the measured control behavior. In terms of depth perception (i.e., perspective, viewing distance), the task was modeled after an aerial refueling scenario. Participants performed the pursuit tracking task for a reference "flat-plane" condition (task axis aligned with vertical screen axis) and depth tracking tasks either without stereoscopic cues, with natural stereoscopic vision, and with amplified hyperstereoscopic vision. Overall, the results of the experiment showed that participants achieved degraded task performance and less effective control dynamics in depth tracking tasks compared the the reference "flat-plane" condition. However, in line with earlier research on aerial refueling operator support systems, increased strength of the stereoscopic vision enhancements is found to enable much improved performance and increased human control gains.

I. Introduction

For a variety of reasons, human operators have been more frequently distanced from the actual position of operation. Such *teleoperation* may be advantageous for various reasons, e.g., in environments that form a substantial health risk or places that are simply problematic to physically reach, such as space or undersea operations [1, 2]. However, separation from the actual stage of action generally results in considerable complications for performing manual control, as, for instance, certain key aspects of perception through human senses may be altered, deteriorated or even missing. As human operators must still be able to perform their tasks up to a certain standard, this requires careful design and implementation of the system interface between the operator and the environment of operation.

One of the critical scenarios where this is currently difficult to achieve is for tasks where human operators are required to control an element of *depth* from a remote location without a natural or unobstructed view, such as the task of the refueling boom operator in an aerial refueling scenario [3–5]. The natural human capability to perceive depth cues through binocular vision is absent during such teleoperation, yet may be fundamental to adequately control an object when heavily reliant on depth information. Similar depth control tasks are also encountered in, among others, medical, industrial, and military applications [6, 7]. Although technologies such as stereoscopic three-dimensional displays may help overcome the absence of natural stereoscopic vision of the environment, a better understanding of how human controllers actually control in stereoscopic-vision enhanced depth control tasks is sorely needed.

Manual control cybernetics has been a topic of research for many years and numerous quantitative models have been constructed subsequently to understand and predict human control behavior [8–10]. In a recent review paper,

---

*M.Sc. student, Control & Simulation Section, Faculty of Aerospace Engineering, P.O. Box 5058, 2600GB Delft, The Netherlands; m.h.h.kemna@student.tudelft.nl.

†Assistant Professor, Control & Simulation Section, Faculty of Aerospace Engineering, P.O. Box 5058, 2600GB Delft, The Netherlands; d.m.pool@tudelft.nl. Senior Member AIAA.

‡Director, multiSIM B.V., Kampweg 55, 3769 DE Soesterberg, The Netherlands.

§Professor, section Control and Simulation, Faculty of Aerospace Engineering, P.O. Box 5058, 2600GB Delft, The Netherlands; m.mulder@tudelft.nl. Member AIAA.

Mulder et al. [10] propose that our current cybernetic models and knowledge are insufficient to fully exploit the benefits of modern technologies such as three-dimensional displays. While generally lacking an explicit cybernetic analysis of the adopted control behavior, the effects of stereoscopic three-dimensional displays on human control performance have been reviewed by McIntire et al. [6, 7]. For a total of six categories of experimental tasks, McIntire et al. report the effects on task performance (i.e., accuracy, response time, time to completion) after addition of stereoscopic vision through such a display. Most importantly, the sort of task that seemed to benefit most from implementation of stereoscopic vision is defined as "*spatial manipulation of real or virtual objects*" [7]. For about two thirds of the total of 80 performed experiments, McIntire et al. report a significant performance increase with stereoscopic displays. Additionally, probable causes are given for cases where such displays showed no significant improvement, i.e., when monoscopic depth cues were sufficiently available or the task was less challenging [6, 7]. Similar findings were reported by Kim et al. [11], who report an optimal perspective view of the scenario with only monoscopic view and visual enhancements could sufficiently assist the human operator. The effect of perspective by rotation of the view was also investigated in a recent study by Karasinski et al. [12], which again confirmed the potential positive effect of perspective or rotated view.

It is thus evident that stereoscopic vision through stereoscopic three-dimensional displays affects manual control performance in depth control tasks. However, at this moment it is not clear to what extent the control behavior of human operators in depth control tasks (with or without stereoscopic vision) matches what is known from well-studied "flat-plane" target tracking tasks, with controlled visual display axis perpendicular to the viewing direction, as considered in most earlier cybernetic studies [8, 10]. In addition, it would be beneficial to also have a model that can adequately describe human control behavior in depth control tasks with stereoscopic vision [10, 13–15]. While there is agreement in literature regarding the potential for stereoscopic displays in spatial manipulation tasks, we still lack a thorough understanding of *how* the stereoscopic view affects human manual control. Insight into such manual control behavior obtained through a "*cybernetic approach*" should ultimately yield a first step towards understanding human adaption with regard to stereoscopic displays to possibly support future innovations on three-dimensional system interfaces [10].

The goal of this paper is to identify human manual control behavior in stereoscopic vision-enhanced depth control tasks from human-in-the-loop experiment data using a cybernetic approach. As part of this study, an experiment is performed using a custom in-house developed experiment setup that includes a stereoscopic three-dimensional display and a passive control stick. Measurements of participants' control behavior were collected for a pursuit tracking task [16, 17] modeled after an aerial refueling scenario [4] and with matching stereoscopic depth cues. The task was implemented as a combined target-following and disturbance-rejection task, to facilitate multi-channel human operator identification [10, 18] and explicitly verify participants' possibly multi-channel control organization. In the experiment, human behavior is measured in four display conditions: a reference flat-plane control task and a depth control task performed with monoscopic, a stereoscopic and hyper-stereoscopic displays. At the time of writing of this paper, the data from only the first eight participants in the experiment were analyzed and will be presented here.

This paper is structured as follows. First, Section II describes the details of the human-in-the-loop experiment and performed data analysis. The experimental results are presented and discussed in Section III. The paper ends with conclusions in Section IV.

II. Methods

A. Control Task

The control task selected for this experiment was a single-axis *pursuit* tracking task, as considered in many earlier investigations [16, 17, 19, 20]. This task was considered a reasonable simplification of an aerial refueling scenario, where the refueling boom operator controls the extension of the refueling boom ("follower") with respect to a moving target aircraft ("target"). Fig. 1 shows a block diagram of a pursuit tracking task, where a Human Controller (HC) controls a system with dynamics defined by the transfer function H_c. In a pursuit task, HCs can explicitly observe the target signal f_t, the controlled system output x, and the tracking error e from the presented visual display. In essence, this also implies HCs can mechanize feedback (x and e) and/or feedforward (f_t) control responses with respect all three signals [16, 20]. However, in this study, rate-control system dynamics – i.e., $H_c(s) = K/s$ – are considered, for which it is known that HCs in fact utilize a "compensatory" error-reducing control strategy based on H_{p_e} and in fact do not control with additional H_{p_x} or H_{p_t} responses [16, 17, 19]. Hence, these optional responses are indicated with dashed lines in Fig. 1.
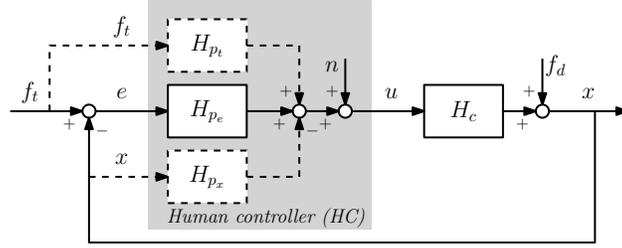
2

Fig. 1. A schematic representation of a pursuit tracking task.

As shown in Fig. 1, both a target and a disturbance forcing function – f_t and f_d, respectively – were present in the considered control task. The target signal indicates the reference position for the controlled system, while f_d represents a disturbance on the controlled system output (e.g., turbulence). Both forcing functions were constructed as quasi-random multisine signals. Two forcing functions were used to allow for verification of HCs' control organization from the collected experiment data. As it was thought unlikely, but not impossible, that HCs would still adopt additional H_{p_x} and H_{p_t} responses in the (previously unstudied) depth control tasks, we elected for ensuring that multi-channel frequency-domain identification methods [18, 19, 21] could be applied to our experiment data, to verify whether a two-channel control strategy was perhaps still adopted by human operators. Table 1 lists all forcing function signal properties (frequencies, amplitudes and phases). Three different sets of phases were generated for both the target and disturbance signals, to be able to mix-up the time-domain realization of f_t and f_d and thereby minimise the risk of signal recognition and/or anticipation by participants.

Table 1. Multisine forcing function data.

		Target signal f_t							Disturbance signal f_d				
					ϕ_t, rad							ϕ_d, rad		
k	n_t	ω_t, rad/s	A_t, m	I	II	III	k	n_d	ω_d, rad/s	A_d, m	I	II	III
1	5	0.383	0.920	3.042	0.217	6.150	1	6	0.460	0.956	1.293	5.882	2.323
2	11	0.844	0.695	4.499	2.851	3.298	2	13	0.997	0.669	4.454	5.624	5.493
3	23	1.764	0.346	5.980	0.911	0.323	3	27	2.071	0.302	2.632	3.489	6.010
4	37	2.838	0.176	1.269	1.129	0.336	4	41	3.145	0.162	4.173	0.654	2.639
5	51	3.912	0.107	2.110	0.877	2.899	5	53	4.065	0.109	1.037	0.820	3.771
6	71	5.446	0.065	0.645	3.996	2.674	6	73	5.599	0.068	5.492	1.370	5.449
7	101	7.747	0.041	0.435	5.621	1.399	7	103	7.900	0.043	4.946	4.811	4.877
8	137	10.508	0.029	0.073	1.516	2.049	8	139	10.661	0.031	2.961	1.520	6.079
9	171	13.116	0.025	3.318	3.352	1.274	9	194	14.880	0.024	3.068	2.955	5.615
10	226	17.334	0.021	3.082	4.610	2.080	10	229	17.564	0.022	2.833	3.515	6.273

B. Experiment Setup

The human-in-the-loop experiment was performed at the Faculty of Aerospace Engineering at TU Delft using a custom-built setup developed by multiSIM B.V. as an aerial refueling training simulator. The experimental setup comprises the 3D PluraView 28" 4K Monitor, a stereoscopic 3D display which consists of two 28-inch monitors with 60 Hz frame rates and resolutions of 3840 by 2160 pixels each (see Fig. 2(a)). A semi-transparent mirror allows the operator to see both monitors at a distance of approximately 80 cm. With passive polarized glasses, the image of only one monitor is visible to each eye, enabling 3D depth cues. While the setup included both left- and right-hand control sticks (CLS-E Active Force Feedback Joystick by Brunner Elektronik AG), as are needed for realistic aerial refueling tasks (see Fig. 2(b)), for this experiment only the right-hand control stick was used for the pursuit tracking task. The stick had no active force feedback, it was passive with mass-spring-damper dynamics.

Fig. 2(c) shows the visual display that was used for the current experiment. A traditional pursuit display was implemented, with a target marker (red crosshair) that indicates the current value of f_t and a green follower crosshair depicting the current controlled element output x. As can be verified from comparison with Fig. 2(b), the red and green crosshairs correspond to the target aircraft and the refueling boom nozzle in the corresponding realistic aerial refueling task.
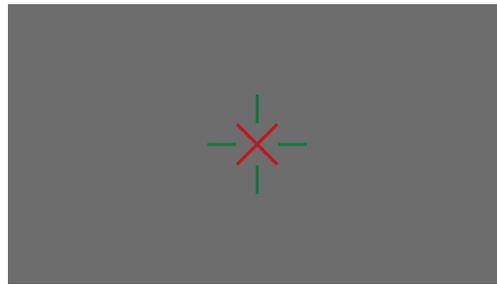
3

(a) Experiment hardware setup.

(b) Screenshot of realistic aerial refueling task.

(c) Screenshot of pursuit tracking task.

Fig. 2. Experimental setup to do measurements on human control behavior in depth control tasks. (a) The setup comprises a stereoscopic 3D display with two 28-inch monitors, with a semi-transparent mirror in between, and two control sticks. (b) The visual display for a realistic aerial refueling scenario. Here an augmented view is shown, where a target element (red crosshair) is slaved to the receiver aircraft and the controlled element (green crosshair) to the refueling boom in an aerial refueling environment with according information overlay [3, 4]. (c) The visual display used for the pursuit tracking task, where also the target element (red crosshair) and follower marker (green crosshair) are shown on a grey background.

The setup accommodates for various scenarios that can be displayed to the participant. Establishing an understanding of possible differences between human control behavior on flat plane tasks and depth tasks, and the significance of stereopsis in such tasks is essential to this research. The pursuit tracking task in Fig. 2(c) is therefore designed to have a most simplistic form possible, while still facilitating the required control signals from Fig. 1. It only shows the target and follower crosshairs and no background or further visual environment. The crosshairs' shapes are made such that interposition as a depth cue is almost fully eliminated (except in extreme cases where the tracking error e is exceptionally high). On the other hand, the visual display for the realistic refueling scenario in Fig. 2(b) shows the same control task signals, yet with a more realistic visual scene and background. In the research project that this paper is the first result from, it is our intention to verify if realistic manual depth control behavior matches what is measured in an "abstract" pursuit tracking task as shown in Fig. 2(c).

C. Axis of Control
This paper's research aims to identify, contrary to more extensive researched Flat-Plane (FP) manual control behaviour [8, 10], the capability of a human controller to perform a target following and disturbance rejection control task along the depth axis. Fig. 3(a) illustrates a typical FP control task where the movement of a target element is *perpendicular* to the line of sight of the observer, i.e., horizontal, vertical or a combination of both. In this paper, a vertical FP pursuit tracking task will be used as a reference condition (see Section II.D).
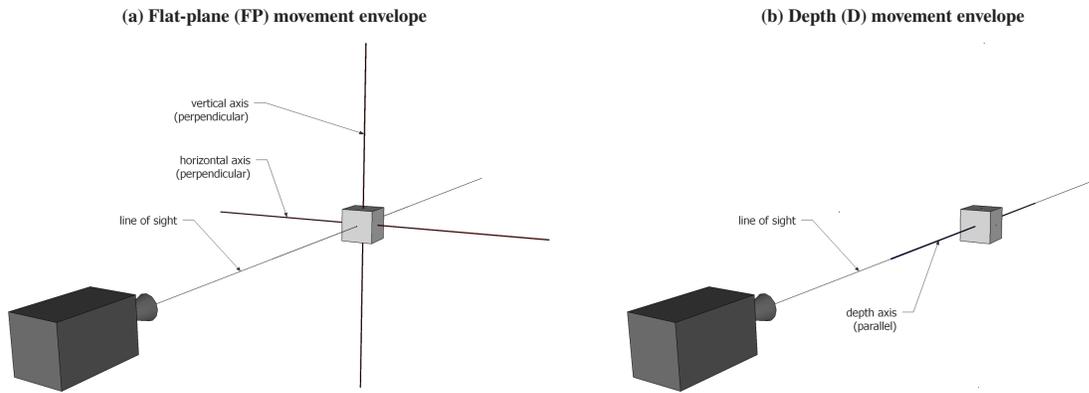
4

Fig. 3. Overview of the distinction between flat-plane (FP) versus depth (D) control tasks. (a) Movement of a target element is horizontal and/or vertical and perpendicular to the observer. (b) Movement of a target element is strictly along the line of sight of the observer.

On the other hand, as illustrated in Fig. 3(b), control tasks along the Depth (D) axis as performed in the current experiment focus on a controlled movement *along* the camera viewing axis. When we assume the participant's eyes to be close to the camera center of projection, this viewing axis is aligned with the line of sight of the observer. Fig. 4 illustrates the perceived depth cues for the pursuit tracking task of Fig. 1 for binocular observation of a pursuit display's target and follower markers. The observable signals f_t, e and x (see the control diagram of Fig. 1) are in this case (generally) at offset positions along the depth axis. The follower and target elements are restricted to moving towards or away from the eyes of the observer, as represented by the camera icons in Fig. 4. As perception of relative differences in the depth dimension is generally less accurate than offsets in the plane orthogonal to the viewing direction (FP task) [22], such a visual depth control task is more difficult.
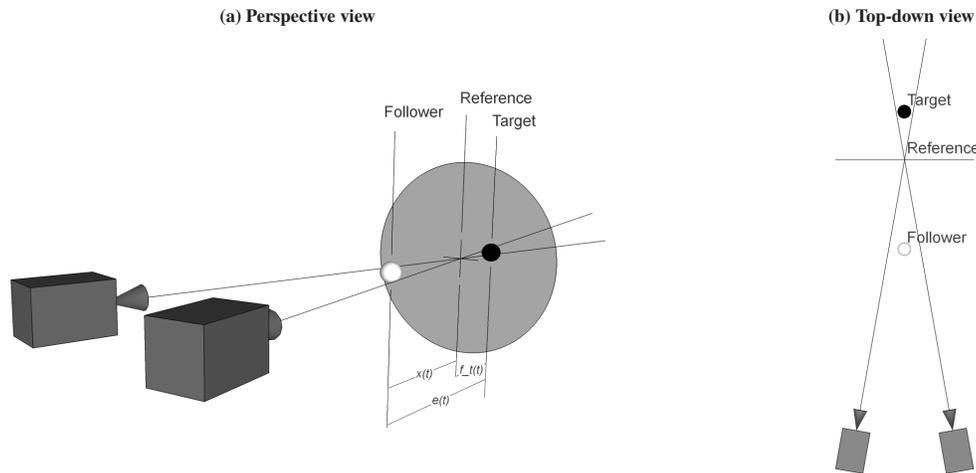

Fig. 4. Schematic representation of a depth control task from (a) a perspective view and (b) a top-down view. The two cameras represent the human observer's two eyes granting stereoscopic vision.

D. Experiment Conditions

In the experiment, the control task described in Section II.A was performed for *four* different experiment conditions, each with a different display configuration. First of all, a reference flat-plane (FP) condition was included, where the red and green markers of the pursuit display shown in Fig. 2(c) moved only in the vertical screen direction, as illustrated in Fig. 3(a). The remaining three display configurations all presented the pursuit tracking task along the depth axis, as illustrated in Fig. 3(b). While holding the convergence point fixed on the reference plane, the three depth conditions had camera offsets of 0 m (D-Mono, with monoscopic vision), 0.06 m (D-Stereo, with natural stereoscopic vision), and

0.5 m (D-Hyper, with hyper-stereoscopic vision). In all conditions, the reference plane for the control task was placed 18.0 m away from the observer's position, matching this offset position in the refueling training scenario of Fig. 2(b). These depth cue settings were based on earlier research into an improved "*Tanker Remote Vision System*" at TNO in the Netherlands [3, 4]. The target and disturbance forcing functions in Table 1 induced a maximum marker displacement of approximately 2 m with respect to this reference plane.

Table 2. Experiment conditions.

Condition	Axis	Description	Inter-eye distance
FP	Vertical	Isometric (flat-plane) view	–
D-Mono	Depth	Monoscopic view	0 m
D-Stereo	Depth	Natural stereoscopic view	0.06 m
D-Hyper	Depth	Hyperstereoscopic view	0.5 m

E. Participants and Experiment Procedures

At the time of writing of this paper, only the data of the first eight participants that performed the experiment were analyzed. All participants were students or staff at TU Delft, and had some prior experience with pursuit tracking tasks, but could be considered as unfamiliar with stereoscopic 3D displays. Prior to the experiment, all participants were tested for their capacity in depth perception due to stereopsis (i.e., their stereoscopic acuity) using both the "*TNO Stereopsis Test*" and the "*Titmus Stereo Fly Test*" [23]. Participants should be able to demonstrate roughly the indicated normal stereoscopic acuity (50 - 100 seconds of arc) on the tests to be allowed in the experiment [24].

After meeting the inclusion criteria, participants were briefed on their task and the experiment. In addition, they were familiarized with both the realistic aerial refueling scenario (Fig. 2(b)) and the pursuit tracking task (Fig. 2(c)), including the different stereoscopic depth cue settings. The measurement data were collected after this familiarization phase, with each participant performing the four experiment conditions in a different randomized order, as generated with 4x4 Latin square designs, see Table 3. In total, each participant performed eight repeated trials of each condition, of which the last three were used as the measurement data set.

Table 3. Experiment Latin Square design.

Participant	Condition			
	1	2	3	4
1	D-Mono	FP	D-Stereo	D-Hyper
2	D-Hyper	D-Stereo	FP	D-Mono
3	D-Stereo	D-Mono	D-Hyper	FP
4	FP	D-Hyper	D-Mono	D-Stereo
5	FP	D-Stereo	D-Mono	D-Hyper
6	D-Mono	D-Hyper	D-Stereo	FP
7	D-Hyper	D-Mono	FP	D-Stereo
8	D-Stereo	FP	D-Hyper	D-Mono

All participants provided written informed consent prior to their participation in the experiment. The study was approved by the TU Delft Human Research Ethics Committee as application number 862.

F. Data Analysis

A number of different dependent variables were calculated from the collected tracking data, to quantify changes in HC control behavior across the tested experiment conditions. The considered control behavior metrics were:

- *Task performance and control effort.* Variations in task performance and exerted control effort were measured in terms of the variances of the error and control (e and u, see Fig. 1) signals, respectively. In addition, using spectral methods [25], the contributions of the target signal (f_t), disturbance signal (f_d), and human control remnant (n) to $\sigma^2(e)$ and $\sigma^2(u)$ were separated.
- *Control signal correlation coefficient.* As a direct measure of the linearity of HCs' control input signal u with respect to the applied forcing function signals, the "correlation coefficient" as defined by McRuer et al. [26] was calculated, at both the frequencies of f_t and f_d separately:

$$\rho_{u_{t,d}}^2(j\omega_{t,d}) = 1 - \frac{\tilde{S}_{uu,n}(j\omega_{t,d})}{S_{uu}(j\omega_{t,d})} \tag{1}$$

In Eq. (1), $\tilde{S}_{uu,n}(j\omega_{t,d})$ is the remnant noise spectrum contribution estimated at each forcing function frequency $\omega_{t,d}$, while $S_{uu}(j\omega_{t,d})$ indicates the peak spectrum value at this same frequency.

- *Human controller modeling.* To explicitly compare participants' control behavior for the different test conditions, the collected human-in-the-loop data were analyzed with HC identification and modeling methods. First, a frequency response function (FRF) estimate of HCs' error response (H_{p_e}) dynamics was estimated:

$$\hat{H}_{p_e}(j\omega_{t,d}) = \frac{U(j\omega_{t,d})}{E(j\omega_{t,d})} \tag{2}$$

As is clear from Eq. (2), these FRFs were directly calculated from the Fourier transformed error (e) and control (u) signals at both the frequencies of the target (ω_t) and the disturbance signal (ω_d), see Table 1. Equivalent to the "effective open-loop describing function" used in earlier research on pursuit tracking [16, 17], here any systematic discrepancies in the dynamics of the estimated $\hat{H}_{p_e}(j\omega_{t,d})$ at ω_t and ω_d would reveal that HCs in fact made use of additional feedforward/feedback control responses in addition to H_{p_e}. The experimental findings presented in Section III will show that this was not found to be the case for the current experiment data.

To model participants H_{p_e} control dynamics, an HC model appropriate for control of K/s system dynamics [8, 10, 26] was used:

$$H_{p_e}(j\omega) = K_p e^{-j\omega\tau_p} \frac{\omega_{nm}^2}{(j\omega)^2 + 2\zeta_{nm}\omega_{nm}j\omega + \omega_{nm}^2} \tag{3}$$

In Eq. (3), K_p and τ_p represent the HC's control gain and time delay, respectively. The characteristics of HCs' neuromuscular system (NMS) dynamics are quantified with the NMS natural frequency ω_{nm} and damping ratio ζ_{nm}. In total, the HC model of Eq. (3) thus has four free parameters, which were estimated by minimizing a sum-of-squared-errors cost function with respect to the (complex) HC FRF data obtained from Eq. (2).

Finally, the quality-of-fit of each fitted HC model was verified by calculating the coefficient of determination, a well-known metric for the correspondence of two time-domain signals, for the control signal predicted by the model ($\hat{u}$) compared to the measured control signal (u):

$$R^2(\hat{\theta}) = \max\left(1 - \frac{\sum_{k=1}^N \|u[k] - \hat{u}[k|\hat{\theta}]\|^2}{\sum_{k=1}^N \|u[k]\|^2}, 0\right) \tag{4}$$

In Eq. (4), the symbol $\hat{\theta}$ indicates a set of estimated parameters of the HC model defined by Eq. (3), i.e., $\theta = \begin{bmatrix} K_p & \tau_p & \omega_{nm} & \zeta_{nm} \end{bmatrix}$. Note from Eq. (4) that here a bounded minimum value of R^2 of 0 is defined for very inaccurate model fits.

G. Hypotheses

For the current experiment, three hypotheses were formulated to test for the effects of varying depth cues in the considered pursuit tracking task:

H1: Tracking performance in depth control tasks is degraded compared to flat-plane tasks, especially at low frequencies (difficult to see) and high frequencies (difficult to track).

H2: The human control dynamics adopted in depth control tasks with stereoscopic 3D displays have an equivalent dynamical shape (but different control behavior parameters) as those used in flat-plane control tasks.

H3: With increasing stereopsis and strength of the available depth cues, task performance and control activity will increase in depth tracking tasks.

III. Results

A. Tracking Performance and Control Effort

This section presents the preliminary results of the human-in-the-loop experiment. Fig. 5 shows the tracking performance and control effort data measured for the eight participants tested so far. Fig. 5 shows the average measured

error signal variance ($\sigma^2(e)$) and control signal variance ($\sigma^2(u)$) data, with the stacked bar charts indicating the contributions of the target signal f_t (white), the disturbance signal f_d (black), and human controller remnant (grey). These different contributions were isolated using spectral methods [25]. The variance bars show the 95% confidence intervals for the total $\sigma^2(e)$ and $\sigma^2(u)$.
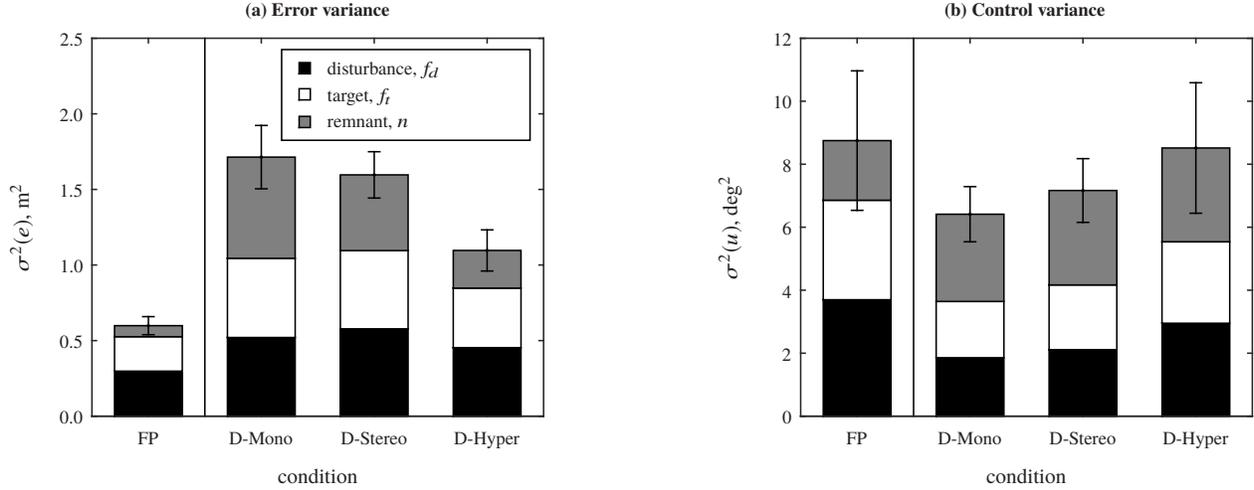

Fig. 5. Measured task performance (error signal variance) and control effort (control signal variance) for all conditions.

Fig. 5 shows that, as expected (Hypothesis **H1**), the best tracking performance (lowest $\sigma^2(e)$) was measured for the Flat Plane (FP) condition. When performing exactly the same pursuit tracking task in the depth dimension with a monoscopic visual (D-Mono), $\sigma^2(e)$ is found to increase by a factor 3, see Fig. 5(a). In addition, Fig. 5(b) shows a drop in control effort for D-Mono compared to FP, due to a reduction in the f_t and f_d components of $\sigma^2(u)$ of around 50%. In addition, the remnant contribution in both $\sigma^2(e)$ and $\sigma^2(u)$ increases notably for the depth tracking conditions. A natural stereoscopic view (D-Stereo) is seen to result an equivalent level of task performance and control effort as measured for a monoscopic visual (D-Mono). For the hyper-stereoscopic display condition (D-Hyper), a clear improvement in tracking performance and increased control effort is measured, as can be verified from Fig. 5. While still not achieving the same level of performance as measured in an FP task, this does indicate that the strong depth cues available with D-Hyper are used for performing the tracking task, which is in agreement with Hypothesis **H3**.

B. Control Correlation Coefficient

Fig. 6 shows the correlation coefficients calculated from the spectra of the measured control signals u. Average $\rho^2_{u_{t,d}}$ data at the target and disturbance frequencies, indicative of the signal-to-noise ratio of HCs' responses to both signals, are presented in Fig. 6(a) and (b), respectively. The dashed lines in Fig. 6, with colors matching those of the mean data, show the minimum and maximum measured $\rho^2_{u_{t,d}}$ for each condition.

For the reference FP condition, Fig. 6 shows $\rho^2_{u_{t,d}}$ above 0.8 for all frequencies, as expected for HCs that actively control the full range of frequencies included in f_t and f_d. For the depth tracking conditions (D-Mono to D-Hyper) the correlation coefficient is clearly lower than for the FP task, especially at high frequencies where $\rho^2_{u_{t,d}} \approx 0.5$, as expected (Hypothesis **H1**). Consistent with the task performance and control effort data of Fig. 5 and Hypothesis **H3**, for the D-Hyper condition the control correlation is found to be consistently improved compared to D-Mono and D-Stereo, especially at low and mid-frequencies, indicative of more effective tracking behavior.

C. Human Controller Modeling

As explained in Section II.F, the collected tracking data were analyzed with a frequency-domain HC modeling approach. Figs. 7 to 10 show example results of the model fitting approach for all four experiment conditions. The FRFs of the HC's control dynamics estimated at the frequencies of f_t and f_d are indicated in these figures with red asterisks and blue circular markers, respectively. The obtained fits of the HC model of Eq. (3) to the FRF data
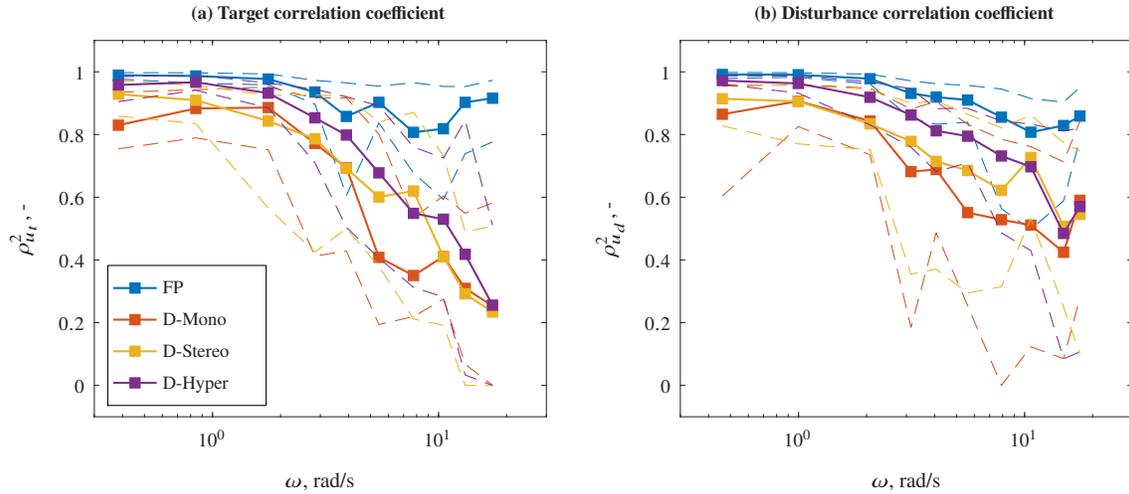
8

Fig. 6. Control correlation coefficient data at both the target and disturbance signal frequencies compared for all experiment conditions. Dashed lines of color that matches the mean condition data indicate the minimum and maximum $\rho^2_{u_{t,d}}$ measured at each frequency.

are indicated in each figure with a solid black line. In addition, each figure legend indicates the corresponding R^2 (coefficient of determination) value for each model fit.

Figs. 7 to 10 show that the HC model of Eq. (3) provided an accurate fit to the FRF data for all experiment conditions. This was expected (Hypothesis **H2**), as independent of the display configuration, participants were controlling an integrator system. In addition, consistent with earlier research on pursuit tracking [16, 17, 19, 20], the presented HC FRFs show no signs of participants' adopting a pursuit control organization for this single integrator control task. The estimated FRFs do not show the increased low-frequency phase lead that is normally associated with pursuit [16, 17]. Furthermore, the FRFs estimated at the frequencies of f_t and f_d show the same $H_{p_e}(j\omega)$ dynamics, which is an indication that participants used no additional feedback or feedforward control responses in addition to the estimated response to the tracking error e, and thus applied pure compensatory control [10, 19].

For conditions D-Mono to D-Hyper, Figs. 8 to 10 show increased spread in the estimated FRFs and lower R^2 values compared to the reference FP case (Fig. 7). These results are indicative of reduced accuracy in applying our HC modeling techniques to the depth tracking task data. This effect was highly consistent over all participants for whom data are analyzed in this paper. Fig. 11 shows the average obtained R^2 values for the HC model fits. As is clear from Fig. 11, with an average $R^2 \approx 0.7$ the highest accuracy in modeling $H_{p_e}(j\omega)$ was obtained for the reference FP condition. On the other hand, for conditions D-Mono and D-Stereo the HC model fits in general did not accurately fit the measured time-domain control signals u. For both conditions, for five out of the eight participants the average R^2 calculated over the three repeated measurement runs was even equal to 0 (minimum value, see Eq. (4)). Consistent with Fig. 10, for condition D-Hyper this only occurred for one participant, while on average R^2 values of around 0.5 were obtained.

In the same format as used for the R^2 data in Fig. 11, Fig. 12 shows the corresponding estimated HC model parameters: the HC gain K_p, the HC delay τ_p, and the NMS natural frequency and damping ratio, ω_{nm} and ζ_{nm}, respectively. For all parameters except ζ_{nm}, which is found to be around 0.25 on average independent of the display setting, a clear variation across experiment conditions is observed.

Fig. 12(a) shows that participants consistently controlled with a gain K_p that was around 50% lower in the depth tracking tasks than for the reference FP condition. This finding is consistent with degraded performance and reduced control activity observed for conditions D-Mono to D-Hyper in Fig. 5. With increasing depth cues, K_p is seen to increase notably from a median of 1.82 for D-Mono to 2.77 for D-Hyper. While this does indicate participants were able to control with a higher gain (and crossover frequency) with improved depth cues, a clear gap with the more effective high-gain strategy used for the FP condition remains.

As can be verified from Fig. 12(b), for the FP condition participants adopted an HC delay of on average 0.26 s, a value that is highly consistent with earlier experiments [19, 26, 27]. Independent of the strength of the provided stereoscopic depth cues, performing the tracking task in the depth dimension is found to result in an increase in τ_p of around 60 ms for all participants. Thus, the fact that equivalent HC delay values were measured for conditions
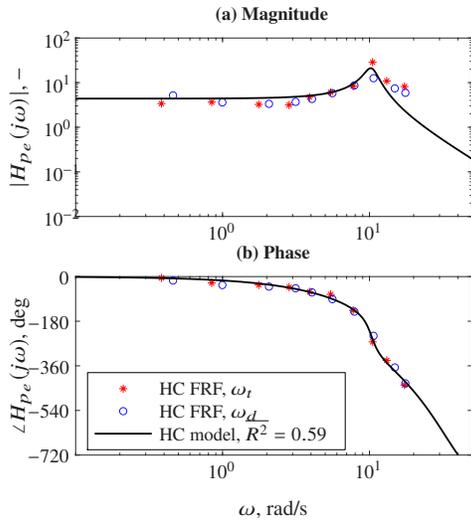
9

replacements

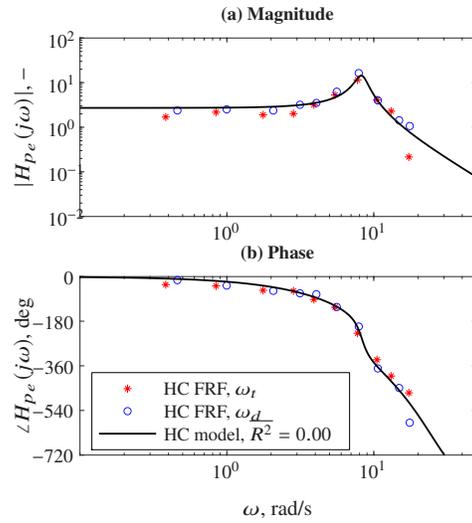**Fig. 7. HC FRFs and fitted model (Subject 1, condition FP).**

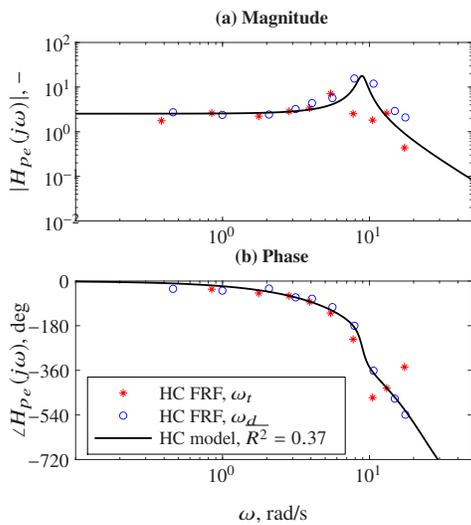**Fig. 8. HC FRFs and fitted model (Subject 1, condition D-Mono).**

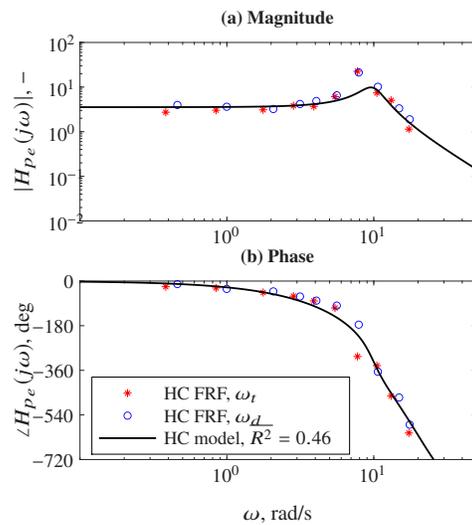**Fig. 9. HC FRFs and fitted model (Subject 1, condition D-Stereo).**

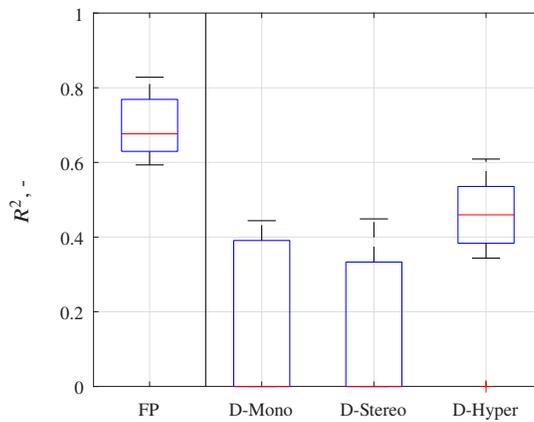**Fig. 10. HC FRFs and fitted model (Subject 1, condition D-Hyper).**

Fig. 11. HC model coefficient of determination for all conditions.
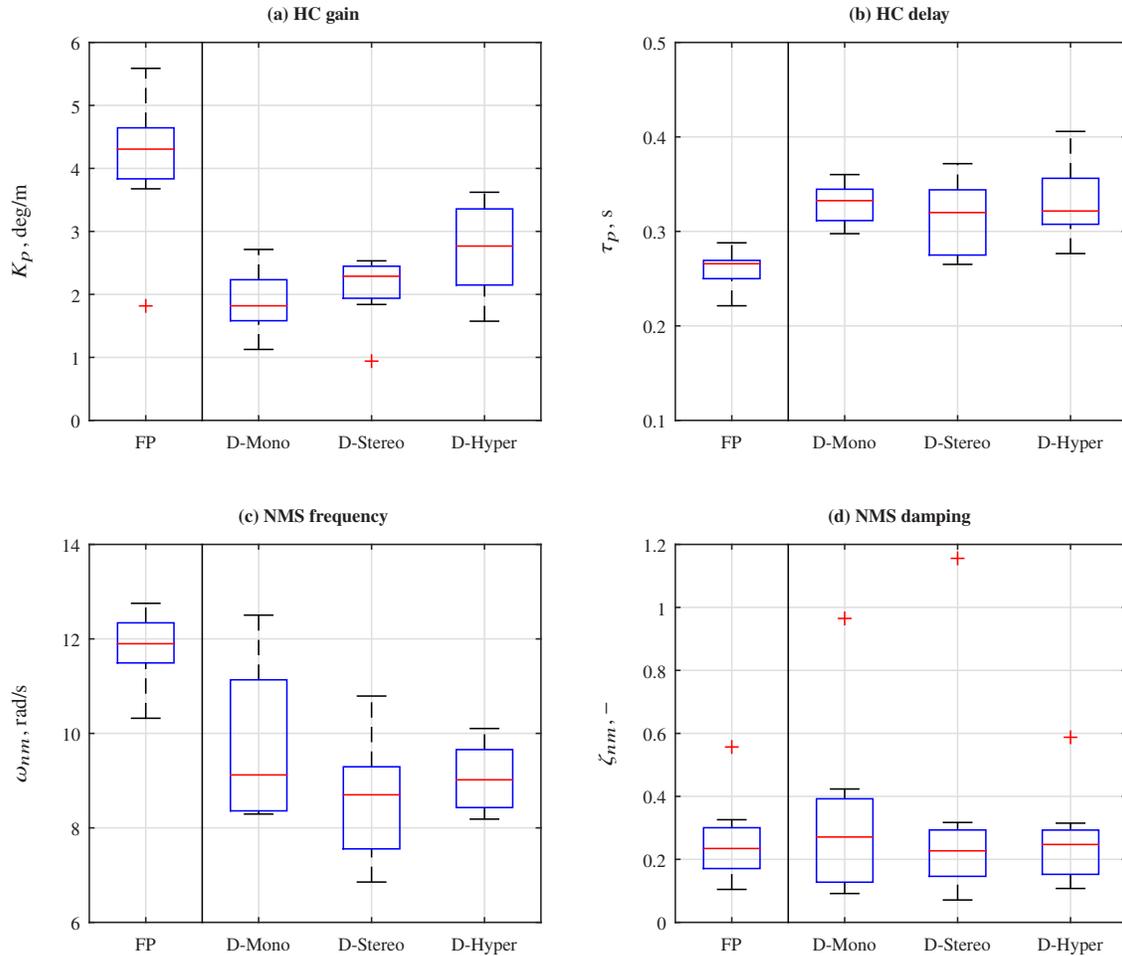
10

replacements

Fig. 12. Measured HC model parameters for all experiment conditions.

D-Mono to D-Hyper suggests that it simply takes longer to perceive and process the depth tracking task, equivalent to observations from multi-axis compared to single-axis manual control [27, 28].

For the NMS natural frequency, a similar variation across conditions as observed for τ_p was found, see Fig. 12(c). For the FP condition, ω_{nm} is highest with a median value of 11.9 rad/s. Consistent with a lower-gain control strategy, for the depth tracking conditions ω_{nm} is almost 3 rad/s lower on average (average median of 8.95 rad/s). No clear trend between different depth cue settings was found, so again this result is explained by an increased difficulty of performing a control task along the depth dimension.

IV. Conclusions

This paper presented the preliminary results (initial eight participants) of a human-in-the-loop experiment performed to collect measurements on human control behavior depth control tasks with varying levels stereoscopic vision enhancement. In addition to pursuit-style depth control tasks with monoscopic, natural stereoscopic, and hyperstereoscopic depth cues, all participants performed a matching reference "flat-plane" control task where the target and follower markers moved in the vertical screen direction.

As expected, manual control performance in depth control tasks was inferior compared to the reference "flat-plane" task, with error variances increased by a factor 2-3. A cybernetic analysis of participants' control dynamics showed that this degraded performance was caused by an approximately 50% reduction in human controller gain and an increase in response delay of 60 ms on average for the depth tasks. In line with earlier research on aerial refueling scenarios [3, 4],

11

increasing the strength of the stereoscopic depth cues to beyond-natural levels (i.e., condition D-Hyper) was found to result in improved task performance, increased control effort and increased estimated human controller model gains. While these results indicate that participants were able to use a more effective control strategy with improved depth cues, overall the level of task performance and the linearity of human control behavior still does not reach that of the "flat-plane" tracking tasks generally considered in cybernetic studies of human manual control behavior.

Currently, we are finalizing the complete human-in-the-loop experiment and data analysis. A larger group of 16+ participants has in the end performed the four-condition experiment focused on the abstract pursuit tracking task. In addition, a second experiment was performed with aerial refueling operators of the Royal Netherlands Air Force, who performed the pursuit tracking task with both the abstract and realistic displays (see Fig. 3) and variations in stereoscopic vision enhancement. Overall, these experiments are expected to provide key insights into how human manual control in realistic depth control tasks (such as aerial refueling) corresponds to the state-or-the-art in cybernetics research and explicitly show the implications for human operators of going from natural human stereoscopic vision to visual displays with or without artificial stereoscopic vision.

References

[1] Sheridan, T. B., "Teleoperation, Telerobotics and Telepresence: A Progress Report," *Control Engineering Practice*, Vol. 3, No. 2, 1995, pp. 205–214. https://doi.org/10.1016/0967-0661(94)00078-U.

[2] Cui, J., Tosunoglu, S., Roberts, R., Moore, C., and Repperger, D. W., "A Review of Teleoperation System Control," *Proceedings of the Florida Conference on Recent Advances in Robotics, Boca Raton (FL)*, 2003.

[3] Bol, J., Van Breda, L., and Ringers, M., "Tanker Remote Visual System (TRVS)," Tech. Rep. DenV S070047, TNO, 2007. URL https://www.tno.nl/en/focus-areas/defence-safety-security/roadmaps/information-sensor-systems/tanker-remote-vision-system/.

[4] Bijl, P., De Vries, S., and Kalthof, J., "Upgrading the KDC-10 RARO Refueling Vision System: WP3b," Confidential Internal Report TNO-DV3 2005-A 161, TNO, 2005.

[5] Winterbottom, M., Lloyd, C., Gaska, J., Wright, S., and Hadley, S., "Stereoscopic Remote Vision System Aerial Refueling Visual Performance," *Electronic Imaging*, Vol. 2016, 2016, pp. 1–10. https://doi.org/10.2352/ISSN.2470-1173.2016.5.SDA-437.

[6] McIntire, J. P., Havig, P. R., and Geiselman, E. E., "What is 3D Good For? A Review of Human Performance on Stereoscopic 3D Displays," *Proceedings of SPIE - The International Society for Optical Engineering*, Vol. 8383, 2012. https://doi.org/10.1117/12.920017.

[7] McIntire, J. P., Havig, P. R., and Geiselman, E. E., "Stereoscopic 3D Displays and Human Performance: A Comprehensive Review," *Displays*, Vol. 35, No. 1, 2014, pp. 18–26. https://doi.org/10.1016/j.displa.2013.10.004.

[8] McRuer, D. T., and Jex, H. R., "A Review of Quasi-Linear Pilot Models," *IEEE Transactions on Human Factors in Electronics*, Vol. HFE-8, No. 3, 1967, pp. 231–249. https://doi.org/10.1109/THFE.1967.234304, URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1698271.

[9] McRuer, D. T., and Krendel, E. S., "Mathematical Models of Human Pilot Behavior," AGARDograph AGARD-AG-188, Advisory Group for Aerospace Research and Development, Jan. 1974.

[10] Mulder, M., Pool, D. M., Abbink, D. A., Boer, E. R., Zaal, P. M. T., Drop, F. M., van der El, K., and van Paassen, M. M., "Manual Control Cybernetics: State-of-the-Art and Current Trends," *IEEE Transactions on Human-Machine Systems*, Vol. 48, No. 5, 2018, pp. 468–485. https://doi.org/10.1109/THMS.2017.2761342.

[11] Kim, W. S., Ellis, S. R., Tyler, M. E., Hannaford, B., and Stark, L. W., "Quantitative Evaluation of Perspective and Stereoscopic Displays in Three-Axis Manual Tracking Tasks," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 17, No. 1, 1987, pp. 61–72. https://doi.org/10.1109/TSMC.1987.289333.

[12] Karasinski, J. A., and Robinson, S. K., "Evaluating Augmented Reality in a Three-Axis Manual Tracking Task," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, San Diego (SA)*, 2019. https://doi.org/10.2514/6.2019-1227.

[13] Mulder, M., van Paassen, M. M., and Boer, E. R., "Exploring the Roles of Information in the Control of Vehicular Locomotion: From Kinematics and Dynamics to Cybernetics," *Presence: Teleoperators and Virtual Environments*, Vol. 13, No. 5, 2004, pp. 535–548. https://doi.org/10.1162/1054746042545256.

[14] Grunwald, A. J., and Merhav, S. J., "Vehicular Control by Visual Field Cues," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 6, No. 12, 1976, pp. 835–845.

[15] Mulder, M., and Mulder, J. A., "Cybernetic Analysis of Perspective Flight-Path Display Dimensions," *Journal of Guidance, Control, and Dynamics*, Vol. 28, No. 3, 2005, pp. 398–411. https://doi.org/10.2514/1.6646.

[16] Wasicko, R. J., McRuer, D. T., and Magdaleno, R. E., "Human Pilot Dynamic Response in Single-loop Systems with Compensatory and Pursuit Displays," Tech. Rep. AFFDL-TR-66-137, Air Force Flight Dynamics Laboratory, Dec. 1966.

[17] Hess, R. A., "Pursuit Tracking and Higher Levels of Skill Development in the Human Pilot," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-11, No. 4, 1981, pp. 262–273. https://doi.org/10.1109/TSMC.1981.4308673.

[18] van Paassen, M. M., and Mulder, M., "Identification of Human Operator Control Behaviour in Multiple-Loop Tracking Tasks," *Proceedings of the Seventh IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems, Kyoto Japan*, Pergamon, 1998, pp. 515–520.

[19] Vos, M. C., Pool, D. M., Damveld, H. J., van Paassen, M. M., and Mulder, M., "Identification of Multimodal Control Behavior in Pursuit Tracking Tasks," *Proceedings of the 2014 IEEE International Conference on Systems, Man, and Cybernetics, San Diego (CA)*, 2014, pp. 69–74.

[20] Mulder, M., Pool, D. M., van der El, K., Drop, F. M., and van Paassen, M. M., "Manual Control with Pursuit Displays: New Insights, New Models, New Issues," *Proceedings of the 14th IFAC Symposium on Analysis Design and Evaluation of Human Machine Systems Tallinn, Estonia*, 2019, pp. 139–144.

[21] van Lunteren, A., "Identification of Human Operator Describing Function Models with One or Two Inputs in Closed Loop Systems," Ph.D. thesis, Delft University of Technology, Faculty of Mechanical Engineering, Jan. 1979.

[22] Wickens, C. D., "Three-Dimensional Stereoscopic Display Implementation: Guidelines Derived from Human Visual Capabilities," *Proceedings of SPIE - The International Society for Optical Engineering*, Vol. 1256, 1990. https://doi.org/10.1117/12.19883.

[23] Lee, J., and McIntyre, A., "Clinical Tests for Binocular Vision," *Eye*, Vol. 10, 1996, p. 282–285. https://doi.org/10.1038/eye.1996.59.

[24] Lee, S., and Koo, N., "Change of Stereoacuity with Aging in Normal Eyes," *Korean journal of ophthalmology : KJO*, Vol. 19, 2005, pp. 136–9. https://doi.org/10.3341/kjo.2005.19.2.136.

[25] Jex, H. R., Magdaleno, R. E., and Junker, A. M., "Roll Tracking Effects of G-vector Tilt and Various Types of Motion Washout," *Proceedings of the Fourteenth Annual Conference on Manual Control*, 1978, pp. 463–502.

[26] McRuer, D. T., Graham, D., Krendel, E. S., and Reisener, W. J., "Human Pilot Dynamics in Compensatory Systems: Theory, Models, and Experiments with Controlled Element and Forcing Function Variations," Tech. Rep. AFFDL-TR-65-15, Air Force Flight Dynamics Laboratory, Wright-Patterson Air Force Base (OH), 1965.

[27] Barendswaard, S., Pool, D. M., van Paassen, M. M., and Mulder, M., "Dual-Axis Manual Control: Performance Degradation, Axis Asymmetry, Crossfeed, and Intermittency," *IEEE Transactions on Human-Machine Systems*, Vol. 49, No. 2, 2019, pp. 113–125. https://doi.org/10.1109/THMS.2019.2890856.

[28] Levison, W. H., and Elkind, J. I., "Two-Dimensional Manual Control Systems with Separated Displays," *IEEE Transactions on Human Factors in Electronics*, Vol. 8, No. 3, 1967, pp. 202–209. https://doi.org/10.1109/THFE.1967.233969.