Published in final edited form as:
Presence (Camb). 2012 ; 21(3): 305–320.

Effect of Visuo-Motor Co-location on 3D Fitts' Task Performance in Physical and Virtual Environments

Michael J. Fu[1], **Andrew D. Hershberger**[2], **Kumiko Sano**[2], and **M. Cenk Çavuşoğlu**[2]

[1]Cleveland FES Center of Excellence, Louis Stokes Cleveland Dept. of Veterans Affairs Medical Center, Cleveland, OH, USA

[2]Dept. of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, OH, USA

Abstract

Given the ease that humans have with using a keyboard and mouse in typical, non-colocated computer interaction, many studies have investigated the value of co-locating the visual field and motor workspaces using immersive display modalities. Significant understanding has been gained by previous work comparing physical tasks against virtual tasks, visuo-motor co-location versus non-colocation, and even visuo-motor rotational misalignments in virtual environments (VEs). However, few studies have explored all of these paradigms in context with each other and it is difficult to perform inter-study comparisons because of the variation in tested motor tasks. Therefore, using a stereoscopic fish tank display setup, the goal for the current study was to characterize human performance of a 3D Fitts' point-to-point reaching task using a stylus-based haptic interface in the physical, co-located/non-colocated, and rotated VE visualization conditions.Five performance measures – throughput, initial movement error, corrective movements, and peak velocity – were measured and used to evaluate task performance. These measures were studied in 22 subjects (11 male, 11 female, ages 20–32) performing a 3D variant of Fitts' serial task under 10 task conditions: physical, co-located VE, non-colocated VE, and rotated VEs from 45–315° in 45° increments.

Hypotheses—All performance measures in the co-located VE were expected to reflect significantly reduced task performance over the real condition, but also reflect increased performance over the non-colocated VE condition. For rotational misalignments, all performance measures were expected to reflect highest performance at 0°, reduce to lowest performance at 90° and rise again to a local maximum at 180° (symmetric about 0°).

Results—All performance measures showed that the co-located VE condition resulted in significantly lower task performance than the physical condition and higher mean performance than the non-colocated VE condition, but the difference was not statistically significant. Also, rotation misalignments showed that task performance were mostly reduced to minimums at 90°, 135°, and 225°. We conclude that co-located VEs may not significantly improve point-to-point reaching performance over non-colocated VEs. Also, visual rotations of ±45° affected throughput, efficiency, peak velocity, and initial movement error, but the number of corrective movements were not affected until ±90°.

1 Introduction

Optimization of human task performance in virtual environments (VEs) is desirable from an engineering standpoint, but it is also crucial in applications such as surgical robot control, where impaired performance can lead to costly consequences. The effect of immersive display modalities on task performance is a well studied area, but there are still gaps in the literature that can be filled.

Many have investigated the value of immersive technologies over typical, non-colocated (NC) computer interaction where the visual field and motor workspace are not aligned (as in a common computer display and mouse interface). Investigations into the effect of visuo-motor misalignments on task performance are rooted in motor control studies regarding the physiological processes behind adaptations to optical prisms (Held, 1965; Shadmehr & Wise, 2005). Since then, the increased accessibility of computers brought the field into intersections with the study of human-computer interaction, where the focus is on optimization of human task performance in VEs – which is also the focus of the current work.

There are two general causes of visuo-motor misalignment: rotational and translational dislocation of the visual display from the input device. For the current study, a co-located interface was defined as the condition when visual and motor workspace scales, origins, and orientations are aligned (similar to human hand-eye interactions). Therefore, translational misalignment refers to the condition where only the scales and orientations are aligned, while rotational misalignment refers to the case where only the scales and origins are aligned. In this paper, the co-located condition was equivalent to the 0° rotation condition.

Previous findings regarding the effect of rotational misalignment on virtual task performance are surprisingly consistent despite the large variation of tasks that were tested. Investigated tasks include 2D point-to-point targeting with a joystick interface (Bernotat, 1970), 3D pick-and-place and tracking using two joysticks (Kim, Tendick, & Stark, 1987; Kim, Ellis, Tyler, Hannaford, & Stark, 1987), whole-arm 3D point-to-point reaching (Blackmon, Çavuşoğlu, Lai, & Stark, 1997), and 3D object orientation matching (Ware & Arsenault, 2004). Task completion times and error rates for visual rotations about the azimuth (direction perpendicular to the ground) were found to have a quasi-symmetric trend about the 0° condition. Specifically, both measures were lowest for the 0° condition, increased to maximums at ± 90°, and decreased to a local minimum at 180°. In short, literature showed highest task performance with no visuo-motor rotational dislocation and lowest when the dislocation was ±90° about the azimuth.

In contrast, previous findings on the effect of translational misalignments (visuomotor co-location) on task performance are conflicting. For instance, Swapp, Pawar, and Loscos (2006) reported that co-location significantly improved performance metrics for a set of 3D tasks. Their method of co-location was to physically align and stereographically calibrate a haptic device located at eye level between the user and the computer display. Three virtual tasks (3D reaching, 3D maze navigation, and object juggling) were tested, each over three arbitrarily defined difficulty levels. Similarly, Lev, Rozengurt, Gelfeld, Tarkhnishvili, and Reiner (2010) reported that a virtual endoscopic surgery suturing task was performed significantly faster using a stereographic, co-located fish tank display modality than a monoscopic, non-colocated monitor. Their co-located modality placed the fish tank display between the user and the haptic device used for input. In addition, Tendick, Jennings, Tharp, and Stark (1993) tested surgeons and showed that a pick-and-place task with laparoscopic tools took up to 1.5x as long when a non-colocated, monoscopic endoscope video image was provided, compared to directly viewing the workspace monoscopically with one human eye. In contrast, Teather, Allison, and Steurzlinger (2009) tested the effect of co-location using a 3D Fitts' task and did not find a significant improvement in task completion time. They used an optically-tracked stylus that was operated directly over a stereoscopic display for the co-located condition. For the non-colocated condition, the stylus workspace was shifted to the right of the display by a distance equal to the width of the display. Additionally, Sprague, Po, and Booth (2006) also did not find a significant difference in completion rate of a virtual 2D Fitts' Task when the visual workspace was translated up to 30% farther from the co-located orientation.

The cited works have answered many questions regarding human performance in VEs, but several gaps in understanding still exist. First, the large variation of task paradigms in the literature make repeatability and inter-study comparisons difficult. Second, task difficulty is known to affect task performance, but only a few studies have taken this into consideration (Swapp et al., 2006; Teather et al., 2009; Teather & Steurzlinger, 2011). Third, it is not clear from the literature whether task completion rates are improved by the use of a co-located VE display. Lastly, although there is evidence that reduced task performance occurs in VEs (vs. physical environments) and under translational and rotational misalignments – only one attempt has been made to investigate all three factors using the same manual task. That attempt by Blackmon et al. (1997) was a small study of 4 subjects that only examined 0, 45, and 90 rotations.

Therefore, the goal for the current study was to comprehensively investigate human performance of a point-to-point reaching task in the physical, co-located/non-colocated VE, and rotated VE visualization conditions. Also, the reaching task should span a range of difficulties in a well-defined manner, but still be general and well-established in order to facilitate inter-study comparison and repeatability. Consistent with these requirements, Fitts' point-to-point reaching task stands out as an appropriate motor task for this goal.

Fitts' task is an established motor task for testing manual performance that has well-defined parameters for adjusting task difficulty (Fitts, 1954). The basic Fitts' task involves a user using a stylus to start at rest at a specific location, and then moving the stylus to rest within a designated target area. Fitts' law formally models the speed/accuracy trade-offs in rapid, aimed movement. According to the modern, Shannon formulation of Fitts' law, the time required for a human to move and point to a target is

$$\mathrm{MT}=a+b\log_2\left(1+\frac{D}{W}\right), \quad (1)$$

where MT is the movement time, *D* is the distance from the starting point to the center of the target, *W* is the width of the target, and constant parameters *a* and *b* are identified by linear regression. The term $\log_2\left(1+\frac{D}{W}\right)$ is called the index of difficulty (ID). ID is commonly interpreted as a measure of the difficulty of the motor task, and carries the unit of `bits,' in reference to an information theory interpretation of Fitts' Law. One example of an application for Fitts' task with modulated difficulty levels is its use by Guiard and Olafsdottir (2011) to assess motor impairment in the reaching performance of hemiparetic stroke survivors.

Although the basic interpretation of Fitts' law is one-dimensional, Fitts' task is applicable to and can be executed in one, two, and three spatial dimensions (I. S. MacKenzie & Buxton, 1992). Fitts' task has also been used to study a variety of computer input devices, including digital pointers, computer mouse inputs, and haptic devices (Ware & Balakrishnan, 1994; Ware & Arsenault, 2004). In fact, Fitts' task has been adapted into an ISO standard to evaluate 2D, non-keyboard computer interfaces (ISO/DIS 9241-9, 2000). Recently, Teather and Steurzlinger (2011) utilized Fitts' law and the ISO 9241-9 task to investigate human pointing performance using various 2D and 3D cursor control and visual display techniques.

In order to apply Fitts' task to evaluate human performance in different experimental conditions, Soukoreff and Mackenzie recommended the use of `throughput' (TP) (Soukoreff & MacKenzie, 2004). Throughput was defined as

$$\mathrm{TP}=\frac{1}{y}\sum_{i=1}^{y}\left(\frac{1}{x}\sum_{j=1}^{x}\frac{\mathrm{ID}_{\mathrm{e}ij}}{\mathrm{MT}_{ij}}\right), \quad (2)$$

where *x* is the number of unique movement conditions, *y* is the number of subjects, and ID_e is the effective ID calculated from the actual distance traveled and end-point distribution measured from human experiment. Since human subjects tend to miss the target or move to the edges of a wide target, ID_e is defined for each unique movement condition as

$$\mathrm{ID}_{\mathrm{e}}=\log_2\left(1+\frac{D_{\mathrm{e}}}{W_{\mathrm{e}}}\right), \quad (3)$$

where D_e is the average distance traveled for multiple repetitions of the same movement condition and $W_{\mathrm{e}}=\sigma\sqrt{2\pi e}=4.133\sigma$, where σ was the standard deviation of the end point locations. This formulation for ID_e, detailed by Soukoreff and MacKenzie (2004), assumes that the end-points have a normal random distribution since it is due to human variability.

1.1 Study Objectives and Hypotheses

This work investigated human performance of a 3D variation of Fitts' point-to-point reaching task (ID 2–6 bits) performed using a stylus-based haptic interface device under various experimental conditions. A total of ten conditions were studied: physical targets (real), non-colocated (NC) virtual targets, co-located virtual targets (0°) using a stereographic fish tank display modality, and virtual targets using fish tank display with rotations about the azimuth of 45, 90, 135, 180, 225, 270, and 315°. Task performance was quantified in five ways (all adapted from literature and detailed in Sec. 2): throughput, efficiency, initial movement error, number of corrective movements, and peak velocity.

Objective 1: Investigate the effect of visualization paradigms (real, NC, and 0°) on task performance measures.

Hypothesis 1: All performance measures in the co-located VE were expected to show significantly reduced task performance over the real condition due to imperfect depth perception in a VE display - even with stereoscopy, due to the lack of accommodation cues to the eyes. This hypothesis was supported by evidence in literature that have shown 1.5–2x decreases in completion rate for virtual tasks versus physical tasks (Graham & MacKenzie, 1996; Mason, Walji, Lee, & MacKenzie, 2001; Sprague et al., 2006; Blackmon et al., 1997). Additionally, the NC condition was expected to exhibit the lowest performance of these three conditions since the visual field was translated away from the motor workspace and may require the most change to typical human visuo-motor mappings. Evidence for this was shown by (Tendick et al., 1993), where endoscopic surgeons were up to 1.5x as slow performing pick-and-place tasks while viewing the workspace and tools through a non-colocated, monoscopic video feed compared to directly viewing the workspace monoscopically (one eye).

Objective 2: Investigate the effect of visual rotation misalignments (0–315° about the azimuth) of an immersive VE on task performance measures.

Hypothesis 2: For all performance measures, it was hypothesized that the effect of rotation misalignments would be consistent with the trends reported in the literature. Specifically, performance should be highest at 0°, reduce toward the lowest performance at 90° and rise again to a local maximum at 180° (which is analogous to performing the task through a mirror image). This behavior should also be symmetric about 0°, as reported in literature.

2 Performance Measures for Analysis

The following quantitative measures were adapted from literature in order to capture a broad range of motor-execution behaviors that reflect different aspects of task performance. They were all calculated from the reaching trajectories after removing data unrelated to movement. To eliminate dwell-time (time between movement termination and computer registration of end-point position) and movement onset delays from interfering with the analysis, only data with velocity greater than 1.5 mm/s were analyzed. This threshold was based on the hand tremor frequency response of retinal surgeons using a stylus grip, which were measured to have an amplitude of 0.03 mm at a fundamental frequency of 9 Hz (Riviere, Ang, & Khosla, 2003). Velocity profiles were estimated from the first difference of the trajectory data after it had been low-pass filtered at 5 Hz with a 3rd-order Butterworth filter (Matlab's filtfilt.m). In addition, all the following performance measures were scalars calculated from composite position, velocity, and acceleration signals. The composite was defined as the square root of the sum of squares of the data at each axis.

Throughput, also referred to as task completion rate, is inversely proportional to the task completion times measured across a range of target difficulties(Soukoreff & MacKenzie, 2004). Therefore, increased values of throughput indicate increased task performance. Throughput was derived as the inverse of the slope from one-parameter linear regression of movement times with respect to ID_e, as in (2). In order to perform statistical analysis, TP was computed for each subject over each condition and had units of `bits/s'.

Number of Corrective Movements was defined as the number of local maxima of the acceleration signal during each trial and indicates the smoothness of a reaching motion. Since each corrective movement (CM) signifies a direction change, if the reaching motion was ideally smooth, then the number of corrections should be 0. This measure was used by Blackmon et al. (1997) to quantify human reaching performance in virtual and real environments. They found that CMs were minimized for real environments with physical targets and was increased for the virtual environment cases. Therefore, an increased number of corrections was equated with decreased task performance.

Efficiency is a measure of how far a subject's trajectory deviated from the shortest, straight line path to the target. A form of it was first defined by Zhai and Milgram (1998) for use in quantifying the ability of subjects to perform a 6 DOF orientation-matching task. In the current work, efficiency was defined as

$$\text{Efficiency}=\frac{D_{\text{endpoint}}}{D_{path}-D_{\text{endpoint}}}, \quad (4)$$

where D_{endpoint} is the Euclidean distance from the location of movement onset to the endpoint position and D_{path} is the length of the actual reaching motion. Therefore, efficiency is higher if a subject reaches in a straight line from starting point to endpoint, versus in a curved-motion. Efficiency equals infinity if the path taken is exactly a straight line, however this is not expected to happen for human reaching motions. The current analysis assumed that increased efficiency implies increased performance. This is consistent with the use of efficiency in Zhai and Milgram (1998), where efficiency was used as a benchmark for two 6 DOF input devices. The input device that facilitated lower task completion times was shown to also have increased efficiency compared to an alternative device that facilitated higher task completion times.

Initial Movement Error was defined as the magnitude of difference between two normalized vectors: the target vector and the initial movement vector. The target vector points from the location of motion onset to the target location, while the initial movement vector points from the location of motion onset to the location where the first CM occurred (the first local maximum of acceleration). Increased initial movement error was considered to indicate degraded performance. This was based on findings of Blackmon et al. (1997) that reported an increase in initial movement error of over 4.5x for virtual reaching tasks compared to an identical physical task.

Peak velocity was defined as the highest magnitude of velocity that was measured during each reaching motion. It has been found that higher peak velocities resulted from reaching in physical environments versus in virtual environments (Graham & MacKenzie, 1996; Blackmon et al., 1997). Therefore, higher peak velocity in an experiment condition was considered to indicate motor control confidence and higher performance. It was also reported in C. L. MacKenzie, Marteniuk, Dugas, Liske, and Eickmeier (1987); Mason (2007); Mason et al. (2001) that peak velocity was positively correlated with target distance. This means that the farther away a target is, the more likely peak velocity will increase. Since target difficulty is a function of target distance, the current study used the method described in Sec. 2.1 to account for the possible effect of parameters such as target distance or ID on this or any other performance measure.

2.1 Accounting for Effect of ID on Performance Measures

If target ID was observed to have a significant effect on a performance measure, linear regression was performed on the values as a function of ID for each experiment condition. As a result, instead of one parameter for each target ID (per experiment condition) – each experiment condition will have two: regression slope and offset. The slope was considered to indicate the sensitivity of the performance measure the tested range of IDs. Meanwhile, the offset represented the performance measure for the minimum tested ID and was derived from the regression slope, m, and zero intercept, $y_\circ$, as

$$\text{offset}=\text{y}_\circ+\text{m}\left(\text{ID}_{\text{min}}\right), \quad (5)$$

where ID_{min} was the lowest tested target ID of 1.9 bits (Table 1). This definition of offset was used in order to maintain a physical relationship between the numerical regression and the empirical results. In fact, Guiard and Olafsdottir (2011) cautions against the use of the regression intercept at zero ID for estimating performance due to the impracticality of constructing a physical target of zero ID.

3 Methods

The following human experiment protocol has been approved by the Institutional Review Board of Case Western Reserve University, where these experiments were performed.

3.1 Equipment

The fish tank display modality was selected for this study due to its ability to provide high-fidelity VEs that align the visual and motor workspaces. They were co-located by placing a haptic device behind the image plane of the calibrated fish tank display. In this way, the haptic device's representation in the VE will appear to match the motion and location of the physical device.

A custom fish tank display (Fig. 1) was designed to be reconfigurable for the physical task, co-located VE, and non-colocated VE configurations. It supports both a 22” CRT monitor (Dell Corp., Round Rock, TX) and a Phantom Omni haptic device (Sensable Technologies Corp., Woburn, MA). The same haptic device and workspace were used for all of the experiment conditions.

OpenGL was used to develop a graphic user interface (GUI) that was rendered on a dual-core workstation computer (Dell Corp.) running Windows XP (Microsoft Corp, Redmond, WA). Data was sampled at 1 kHz using the OpenHaptics API (Sensable Technologies Corp.). Stereographic images were rendered using non-symmetric frustums and viewed using Crystal Eyes 3 active shutter glasses and transmitter (RealD Corp., Beverly Hills, CA). The physical targets used in the experiments were custom fabricated using hollow half-spheres mounted on telescoping stems.

Stereographic calibration between the virtual and physical workspaces was performed manually using physical reference objects (solid blocks and reaching task targets placed at the extremes of the workspace) seen through a half-mirror as reference with respect to a fixed forehead rest used by all subjects (Fig. 2). The actual experiments were conducted using a full mirror in order to maintain occlusion depth cues (close objects visually obstruct farther ones) that are important for visual depth perception.

Distance between the eyes and the image plane (computer display screen for the NC configuration and mirror reflected image for the co-located condition) was approximately 50 cm for both the co-located and non-colocated conditions (Fig. 1). A custom headrest was used for the non-colocated condition in order to restrict head motion and prevent `stereo swim', the effect when the fused image appears to move due to head motion. For the physical task, the mirror was removed, but the forehead rest was still used in order to maintain a consistent viewpoint across all experiment conditions. The physical targets were placed at various locations within a workspace measuring 26 cm W × 15 cm L × 15 cm H.

3.2 Subjects

Twenty-two subjects (11 male and 11 female, ages 20–32) were recruited and compensated for their participation in this study. All subjects were right handed and tested using their dominant hand. The experiments used a repeated-measures design in which each subject performed each of the ten experiment conditions once. The entire experiment took approximately 3 hours per subject. Due to time constraints, all but two subjects performed the entire set of tests over 2 separate days – one day consisting of randomly presented rotated conditions and another day consisting of the physical and non-colocated conditions. At least 1 week of time separated the first and second day sessions.

3.3 Experiment Paradigms

Each of the ten conditions were presented to participants in random order. During each condition, subjects were asked to sit before the fish tank display, grip the haptic device stylus like a pen, and perform the experiment task. Each session tested one experimental condition, consisting of a set of 40 practice trials (40 targets) followed immediately by a set of 40 recorded trials. During each trial, the home position and one target were displayed simultaneously to the subject. Each subject was instructed to first set the tip of the stylus at the home position, press a button on the stylus when ready to move to the target, move the stylus tip as quickly as possible to the target, and press again when they were sure that the tip of the stylus was anywhere within the target volume. Each button press triggered a chime sound effect. Also, ample rest was provided to subjects between each condition, but no rest was provided between the practice and actual test runs in order to maintain the subject's familiarity with the specific condition.

The home position was laterally centered near the edge of the workspace closest to the subject. Each set of 40 targets were randomly constructed from 10 unique targets spanning IDs of 2–6 bits (Table 1), each repeated four times. Two of the repetitions were placed on the opposite lateral side of the other two with respect to home position in order to minimize the effect of direction bias. Data from all four repetitions were averaged for analysis.

Real Task—For the real task (Fig. 3), subjects reached to physical targets that the experimenter manually changed. Subjects were additionally instructed to judge the accuracy by vision and not by contacting the stylus tip with the target. One peculiarity with the Phantom Omni haptic device was that the gimbal attached to the stylus can obstruct the view of the stylus tip when a right-handed user points toward a left-sided target (and vice versa). In order to account for this, during the real condition, the hollow face of the right-sided targets were rotated toward the subject, while left-sided targets were rotated 45° about the target stem to face just right of the operator so that the stylus tip becomes visible to the operator. This alteration was not needed for the VE conditions. Also, separate trial runs were analyzed to ensure that the rotation of targets did not result in significantly different completion times between targets located on opposite sides.

Co-located and Rotations—Figure 4 shows all eight rotation conditions. Virtual targets were generated to be hollow half-spheres in order to match the appearance of the physical targets. Additionally, virtual targets were made semi-transparent for the rotated conditions so that the cursor would not be obstructed. Separate trial runs were made to ensure that transparent targets did not result in significantly different completion times. Also, force feedback was not provided for virtual targets in order to evoke vision-based motor control from participants and to record performance metrics that are not affected by contact-based strategies where subjects might hunt for haptic contact with the target before deciding to register the end point click.

Non-colocated—The NC experiment condition also used the GUI for the 0° condition, but was viewed by participants directly on the computer display as shown in Fig. 1.

Statistical Analysis—Statistical testing for mean differences were performed using repeated measures analysis of variance (ANOVA) with Greenhouse-Geisser epsilon corrections and Holm-Sidak multiple comparisons (performed in OriginPro 8.5, OriginLab Corp., Northampton, MA). The performance measures were statistically tested in two tests, one consisting of the real, NC, and co-located conditions (referred to as the visualization paradigms) and another with only the 0–315° conditions (referred to as the rotations). The statistical power for visualization paradigms was computed to be 0.69 (as calculated by G*Power 3.1 by Faul, Erdfelder, Lang, and Axel (2007), sample size of 22, 3 repeated measurements, $\alpha = 0.05$, Cohen's f medium effect size of 0.25, 1 group of subjects) and 0.99 for the rotations analysis (8 repeated measurements). Statistical tests for significant mean differences were computed using experiment condition as the within-subjects factor and performance measures as the dependent variable.

4 Results

Boxplots for all performance measures are in Fig. 5. The boxplots serve to illustrate the statistical distribution and means of each performance measure across all subjects, grouped by experimental condition.

Throughput

Task ID was observed to have an effect on mean task completion time, which was confirmed via ANOVA. ID was found to have a significant effect on task completion time for visualization paradigms ($p < 0.001$, $F(9,189) = 253.1$) and rotations ($p < 0.001$, $F(9,189) = 169.6$). Thus, TP was calculated by fitting a one-parameter linear slope to each subject's movement time data as a function of ID_e and taking the inverse of the fitted slope. A histogram of all linear regression coefficient of determination, R^2, values are shown in Fig. 6.a. Also, the TP for all subjects were reported as boxplots in Fig. 5.a.

Task paradigm was found to have a significant effect on TP (p=0.0016, F(2,42)=15.96). Highest mean TP was observed for real targets (4.71 b/s), which was found to be significantly greater than both the NC (3.26 b/s) and co-located (3.51 b/s) cases. The NC and co-located mean TP values were not found to be significantly different. Rotations were also found to significantly affect TP (p=0.00, F(7,147)=81.66). At 0°, TP was significantly higher than all the other rotations. Also, TP between the 0 (3.51 b/s), 45 (2.73 b/s), and 90° (1.55 b/s) conditions were found to be significantly different from each other. In addition, TP at 45° was significantly different from all the other rotations. There was no significant difference between TP for the 90° case versus 135–270°, but it was significantly lower than TP at 315°.

Throughput decreased from 0° to a local minimum at 135° (1.3 b/s), peaked at a local maximum at 180° (1.70 b/s) before decreasing to another local minimum at 225° (1.28 b/s). From 225–315° (3.05 b/s), TP increased almost to the 0° mean. It is noteworthy that TP at 180° (1.79 b/s) was significantly higher than both the 135 and 225° conditions, which is where the lowest TP values occurred. Mean TP for 135 and 225° were not significantly different from each other.

Efficiency

Efficiency measures, in Fig. 5.b, were calculated for each subject without distinguishing each target by its difficulty. Paradigm significantly impacted efficiency (corrected p = 0.0065, F(1.7,42)=6.24). Efficiency was highest for the real target condition, which was significantly greater than both the NC and 0° cases. A significant difference was not detected between the NC and 0° conditions.

Rotations were also found to exert a significant effect on efficiency (corrected $p < 0.001$, F(3.34,147) = 21.11). Efficiency was highest for the 0° condition (10.66) compared to any other rotation. Like other performance measures, efficiency decreased from 0° to a local minimum at 135° (2.71) before peaking at the local maximum observed at the 180° condition (3.91). Efficiency again decreased at 225° (2.84) before increasing from 270–315° to a mean of 6.75, which was not significantly different from the 45° mean of 6.11.

Initial Movement Error

Initial movement error measures, boxplotted in Fig. 5.c, were calculated for each subject without distinguishing each target by its difficulty. This was done because there was no apparent trend in the individual initial movement errors with respect to target difficulty. The range of initial movement errors was [0–2] because the target and initial movement vectors were normalized prior to taking their difference.

Paradigm did not significantly affect mean initial movement error (corrected p = 0.07, F(2,42) = 4.19). Due to possible calibration errors, the real target condition was not included in the analysis. The Phantom Omni calibration is hard-coded into the haptic device based on a well on the base that serves as both a holder for the stylus and a calibration point every time the pen is inserted. However, the joints of the haptic device can shift slightly even when the stylus tip is within the well, which may lead to calibration errors between the actual and estimated joint angles. Therefore, the initial movement error for the real target condition was unreliable and not included in the analysis.

Rotations, however, did significantly impact initial movement error ($p < 0.001$, $F(7,147) = 20.38$). The lowest mean initial movement error occurred for the 0° condition (0.86, which was significantly different from all other rotations except 45 and 315°) and increased up to a maximum mean error of 1.18 for 135°. However, the initial movement error at 135° was only significantly different from the 0 and 315° conditions. Similar to other measures, a local minimum occurred for 180° (1.09), but it was not significantly different from 135° or 225° as in the other measures. Also, initial movement error decreased from 1.15 at 225° down to 0.91, which was not significantly different from the initial movement error at 0°.

Number of Corrective Movements

Linear regressions were performed on each subject's number of CMs as a function of target difficulty and a histogram of the coefficients of determination are reported in Figure 6.b. This was done because the ID had a significant effect on the number of CMs for paradigms ($p < 0.001$, $F(9, 189) = 145.6$) and for rotations ($p < 0.001$, $F(9,189) = 138.4$). Linear regression parameters for each subject are reported as boxplots in Fig. 5.d.

Paradigm did not have a significant effect on the linear regression offset, which represents the number of CMs for the target with lowest ID ($p = 0.063$, $F(2,42) = 2.96$). However, a significant effect was found for paradigm on the linear regression slope ($p < 0.01$, $F(2,42) = 22.83$). Multiple comparisons revealed that the real condition (0.92) had significantly lower mean slope than both the NC (1.78) and 0° (1.94) cases. In addition, the NC and 0° conditions did not have significantly different mean slopes.

Rotation was found to have a significant effect on linear regression offset ($p < 0.001$, $F(7,147) = 12.95$). The lowest mean offset occurred for the 0° condition (3.47 CMs). Local maximum offsets occurred at 90 (8.82 CMs) and 225° (8.36 CMs). A local minimum offset of 6.58 CMs occurred at 180°. Rotation also had a significant effect on linear regression slope ($p < 0.001$ $F(7,147) = 22.51$). Similar to the offsets, the 0° exhibited a minimum slope of 1.94. Unlike the offsets, local maximum slopes occurred at 135° (6.29) and 225° (6.49). A local minimum also occurred for the 180° condition (4.36), which was significantly different from only the 0 and 315° conditions.

Peak Velocity

Figure 5.e reports the coefficient of determination results for linear regressions performed for each subject's peak velocity as a function of target difficulty. This was done because ID had a significant effect on peak velocity for paradigms ($p < 0.001$, $F(9,189) = 320.4$) and for rotations ($p < 0.001$, $F(9,189) = 166.3$). Linear regressions were also performed as a function of target distance, but this produced lower correlations. Also, a histogram of all linear regression R^2 values are shown in Fig. 6.c.

Paradigm exhibited a significant effect on both offsets ($p = 0.0013$, $F(2,42) = 14.48$) and slopes ($p < 0.001$, $F(2,42) = 29.04$). In both analyses, the real target condition (0.29 m/s and slope 0.06) exhibited a significantly higher mean peak velocity than the NC (0.23 m/s and slope 0.04) and 0° (0.23 m/s and slope 0.05) cases. In both paradigms and rotations, there was no significant difference between the NC and 0° conditions.

Rotations exhibited a significant effect for offsets ($p < 0.001$, $F(7,147) = 14.63$). A maximum mean offset of 0.23 m/s occurred in the 0° case. Mean offsets decreased from 0° to a minimum of 0.15 for the 180° condition (0.15 m/s) and then increased up to 0.21 m/s for 315°. Unlike other measures, there was no local maximum at 180°. Rotations also had significant effect on slopes ($p < 0.001$, $F(7,147) = 11.18$). Maximum slope occurred for the 0° (0.047) case and decreased to a local minimum of 0.028 at the 90 and 135° conditions, which both had mean slopes of 0.028. A local maximum mean slope occurred at 180°

(0.033), but it was not significantly different than the mean slopes at 135 and 225°. Another local minimum occurred at 225° (0.024), after which mean slope increased to significantly higher values at 270 (0.03) and 315° (0.04).

5 Discussion

For convenience, Table 2 summarizes the mean and standard deviations (in parentheses) for all performance measures relative to a baseline (either the real or co-located condition). Measures from the real, co-located, and non-colocated conditions in the top three rows were normalized to the real condition as baseline, while the rotated conditions (bottom eight rows) were normalized to the 0° condition as baseline. This shows the multiplicative effect that each condition had on the performance measures relative to baseline.

5.1 Real vs. Non-colocated and Co-located

As hypothesized and consistent with literature, all performance measures indicated that performance was better for the physical task than the NC and co-located conditions. Subjects performed the physical tasks at a higher rate than the virtual tasks. As seen in Table 2, mean throughput for the NC and co-located tasks were 0.69x and 0.75x those for the real task, respectively. This indicates that across the tested target ID range of 2–6 bits, task completion times were significantly lower for the real targets versus the virtual ones.

These findings were in line with previous reports that completion times for a physical reaching task were 0.67–0.5x of those for the identical virtual task. Specifically, studies using 2D reaching by Graham and MacKenzie (1996); Mason et al. (2001); Sprague et al. (2006) reported this, as did a study of 3D stylus-based reaching by Liu, Liere, Nieuwenhuizen, and Martens (2009). Blackmon et al. (1997) reported that completion time for a physical whole-arm reaching task was 0.18x that of the virtual task, but the authors noted that a head-mounted display (HMD) used for the co-located condition caused excessively long movement times because subjects had to search for the target by moving their head around the virtual environment.

Similarly, subjects performed the real task with smoother and straighter trajectories than the virtual tasks. Straighter motion was evidenced by efficiencies for the NC and co-located conditions, which were 0.77 and 0.83x that of real task. In addition, smoother motion was indicated by corrective movement regression slopes for both the co-located and NC conditions, which were 1.97x higher than the real task. This increase was similar to the 2.5x increase in the number of submovements for virtual 3D reaching over an identical physical task reported by Liu et al. (2009), but was lower than the 4.2–11x increase in corrective movements for the virtual reaching task reported by Blackmon et al. (1997). However, stereoscopic display was not used in the non-colocated conditions tested by Blackmon et al. (1997) and the 11x result was for the co-located HMD condition, which was likely influenced by the previously mentioned, undesired target searching behavior. It has been well established that stereo displays reduced completion times (Kim, Tendick, & Stark, 1987; Tendick et al., 1993; Arthur, Booth, & Ware, 1993; Aresenault & Ware, 2004). Corrective movement regression offsets for the real task were also lower than both NC and co-located conditions, but the result did not reach statistical significance.

In addition, subjects also reached higher velocities of hand movement during the real task than the virtual ones. This was indicated by lower peak velocity regression slopes (approximately 0.8x) and offsets (approximately 0.7x) for the NC and 0° than the real task. The current results agreed with previous work that found peak velocity to be significantly affected by target distance and extends the literature by finding and analyzing the linear regression parameters of the trend.

Co-located vs. Non-colocated—Contrary to Hypothesis 1, the co-located VE did not exhibit a significant difference in any of the performance measures compared to the non-colocated condition. Although all but one of the mean performance measures (number of corrective movements) indicated better performance in the co-located condition, the differences did not reach statistical significance. The lack of a significant difference in task completion rate supports and extends the literature that investigated Fitts' task performed in VEs (Teather et al., 2009; Sprague et al., 2006). Teather et al. (2009) found no significant difference in the completion rate of a 3D Fitts' Task between the co-located VE and a non-colocated VE where the motor workspace was shifted adjacent to the right of the visual display. Meanwhile, Sprague et al. (2006) found no significant difference in the completion rate of a 2D Fitts task by translating the virtual targets farther away from the participant by up to 30% in distance. The current results provide strong evidence that task completion rate, efficiency, initial movement error, and peak velocity were not significantly affected by a non-colocated VE created by rotating the visual display approximately 45° about the horizon (while maintaining a constant distance between the eyes and the viewing plane) – compared to a co-located VE.

The mechanisms behind the current results is not clear. However, Teather et al. (2009) suspected that the difficulty that humans have with depth perception within VEs leads to longer completion times that may mask the effect that visuo-motor misalignments have on completion time. Sprague et al. (2006) also theorized that faster-than-expected human sensorimotor adaptation may have resulted in the lack of significant differences in completion time between co-located and non-colocated conditions. Although these are both possible explanations, their verification will require further investigation in properly designed experiments.

5.2 Effect of Azimuth Rotations

In support of Hypothesis 2, azimuth rotations impacted performance measures in a cyclic manner that was symmetric about 180°, similar to the trends reported by the literature. Specifically, all performance measures, except for throughput, did not exhibit statistically significant differences between means for the 135 and 225°, 90 and 270°, and 45 and 315° conditions. This symmetry was also observed in the completion times reported by Bernotat (1970); Kim, Tendick, and Stark (1987); Kim, Ellis, et al. (1987); Ware and Arsenault (2004). It is not apparent why, unlike the other measures, throughput was not symmetric about 180°.

In addition, the sensitivity of performance measures to rotations appeared to be separated into two groups. Compared to the 0° condition, rotations did not cause significantly increased peak velocity offset or corrective movement offset and slope parameters until 90°. In contrast, throughput, efficiency, peak velocity slope, and initial movement error exhibited significant performance decreases between 0 and 45°. These findings were in line with the literature, which reported the smallest differences in performance measures to exist between 0 and 45° (Bernotat, 1970; Kim, Tendick, & Stark, 1987; Kim, Ellis, et al., 1987; Blackmon et al., 1997; Ware & Arsenault, 2004).

The major difference between the current findings and the previous work was the fact that poorest task performance occurred for the 90°, 135°, and 225° conditions, compared to 90 and 270° in previous findings (Bernotat, 1970; Kim, Tendick, & Stark, 1987; Kim, Ellis, et al., 1987; Blackmon et al., 1997; Ware & Arsenault, 2004). However, the current results do not contradict psychophysics literature, which reported that poorest manual performance occured for visual rotations in the range of 90–135° or 225–270° for physical tasks under camera rotations (Shadmehr & Wise, 2005).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

6 Conclusion – System Design Implications

Several VR system design principles may be gathered from the results. First, for point-to-point reaching-type tasks, if the goal is optimize task throughput, efficiency, peak velocity, or initial movement error for fish tank display modalities, then motor and visual perspectives should be aligned – since visual rotations of 45° in either direction significantly impacted these measures. However, if only trajectory smoothness is of concern, then visual rotations of up to ±90° may be acceptable before significant effects on performance might be detected. Second, a co-located VE display may not provide significant improvement over a non-colocated VE display in terms of throughput, efficiency, trajectory smoothness, peak velocity, and initial movement error for virtual point-to-point reaching tasks.

References

Aresenault R, Ware C. The importance of stereo and eye-coupled perspective for eye-hand coordination in fish tank vr. Presence. Oct; 2004 13(5):549–559.

Arthur KW, Booth KS, Ware C. Evaluating 3D task performance for fish tank virtual worlds. ACM Transactions on Information Systems. Jul; 1993 11(3):239–265.

Bernotat RK. Rotation of visual reference systems and its influence on control quality. IEEE Transactions on Man-Machine Systems. 1970; 11(2):129–131.

Blackmon, TT.; Çavuşoğlu, MC.; Lai, F.; Stark, LW. Human Hand Trajectory Analysis in Point-and-Direct Telerobotics. Proc. of the 8th International Conference on Advanced Robotics (ICAR'97); Jul. 1997 p. 927-932.

Faul F, Erdfelder E, Lang A-G, Axel B. G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods. 2007; 39(2):175–191. [PubMed: 17695343]

Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology. 1954; 47:381–391. [PubMed: 13174710]

Graham, ED.; MacKenzie, CL. Physical versis virthal pointing. Proc. of the 2001 ACM Conference on Human Factors in Computing Systems; Apr. 1996 p. 292-299.

Guiard Y, Olafsdottir HB. On the measurement of movement difficulty in the standard approach to fitts law. PLoS ONE. Oct.2011 6(10):e24389. [PubMed: 22053175]

Held R. Plasticity in sensory-motor systems. Scientific American. 1965; 213(5):84–94. [PubMed: 5828465]

ISO/DIS 9241-9. (Ergonomic requirements for office work with visual display terminals - Part 9: Requirements for non-keyboard input devices). 2000. International Organization for Standardization.

Kim WS, Ellis s. R. Tyler ME, Hannaford B, Stark LW. Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks. IEEE Transactions on Man-Machine Systems. 1987; 17(1):61–72.

Kim WS, Tendick F, Stark LW. Visual enhancements in pick-and-place tasks: Human operators controlling a simulated cylindrical manipulator. IEEE Journal of Robotics and Automation. 1987; 3(5):418–425.

Lev, DD.; Rozengurt, R.; Gelfeld, T.; Tarkhnishvili, A.; Reiner, M. Proc. of EuroHaptics 2010. 2010. The effects of 3d collocated presentation of visuo-haptic information; p. 432-437.

Liu, L.; Liere, R. v.; Nieuwenhuizen, C.; Martens, J-B. Comparing virtual vs. real world performance by fitt's law, aimed movements in a virtual environment are less efficient than those performed in the real world. Proc. of the 2009 IEEE Virtual Reality Conf.; 2009. p. 219-222.

MacKenzie CL, Marteniuk RG, Dugas C, Liske D, Eickmeier B. Three-idmensional movement trajectories in fitts' task: Implicaitons for control. The Quarterly Journal of Experimental Psychology. 1987; (39A):629–647.

MacKenzie, IS.; Buxton, W. Extending fitts' law to two-dimensional tasks. CHI'92: Proc. of the SIGCHI Conf. on Human Factors in Computing Systems; 1992. p. 219-226.

Mason AH. An experimental study on the role of graphical information about hand movement when interacting with objects in virtual reality environments. Interacting with Computers. 2007; 19:370–381.

Mason, AH.; Walji, MA.; Lee, EJ.; MacKenzie, CL. Reaching movements to augmented and graphic objects in virtual environments. Proc. of the 2001 ACM Conference on Human Factors in Computing Systems; Mar. 2001 p. 426-433.

Riviere CN, Ang WT, Khosla PK. Torward active tremor canceling in handheld microsurgical instruments. IEEE Transactions on Robotics. 2003; 19(5):793–800.

Shadmehr, R.; Wise, SP. The computational neurobiology of reaching and pointing, a foundation for motor learning. Sejnowski, TJ.; Poggio, TA., editors. The MIT Press; Cambridge, MA: 2005.

Soukoreff RW, MacKenzie IS. Towards a standard for pointing device evaluation: Perspectives on 27 years of fitts law research in hci. International Journal of Human-Computer Studies. 2004; (61): 751–789.

Sprague, DW.; Po, BA.; Booth, KS. Proc. of Graphics Interface 2006. 2006. The importance of accurate vr head registration on skilled motor performance; p. 131-137.

Swapp D, Pawar V, Loscos C. Interaction with co-located haptic feedback in virtual reality. Virtual Reality. 2006; (10):24–30.

Teather, RJ.; Allison, RS.; Steurzlinger, W. Evaluating visual/motor co-location in fish-tank virtual reality. IEEE Toronto International Conference on Science and Technology for Humanity Symposium on Human Factors and Ergonomics; IEEE; Sep. 2009 p. 624-629.

Teather, RJ.; Steurzlinger, W. IEEE Symposium on 3D User Interfaces. IEEE; Mar. 2011 Pointing at 3d targets in a stereo head-tracked virtual environment; p. 87-94.

Tendick F, Jennings RW, Tharp G, Stark L. Sensing and manipulation problems in endoscopic surgery: Experiment, analysis, and observation. Presence. 1993; 2(1):66–79.

Ware, C.; Arsenault, R. Proc. of the 1st Symposium on Applied Perception in Graphics and Visualization. ACM; New York, NY, USA: 2004. Frames of reference in virtual object rotation; p. 135-141.

Ware C, Balakrishnan R. Reaching for objects in vr displays: lag and frame rate. ACM Transactions on Computer-Human Interaction. Dec; 1994 1(4):331–356.

Zhai, S.; Milgram, P. Quantifying coordination in multiple dof movement and its application to evaluating 6 dof input devices. Proc. of the 1998 ACM Conference on Human Factors in Computing Systems; Apr. 1998 p. 320-327.

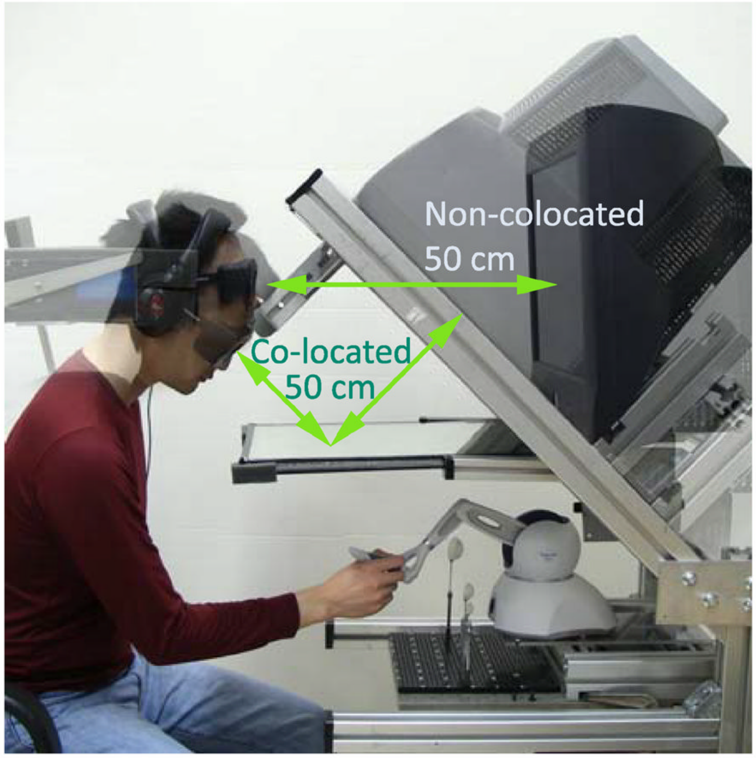


Figure 1.
User positioning at the fish tank display used for the 0–315° conditions (tilted monitor with user facing downward) and non-colocated condition (upright monitor with subject looking forward and head stabilized by a custom headrest). The setup for the physical condition required removing the mirror, but the forehead rest was still used.

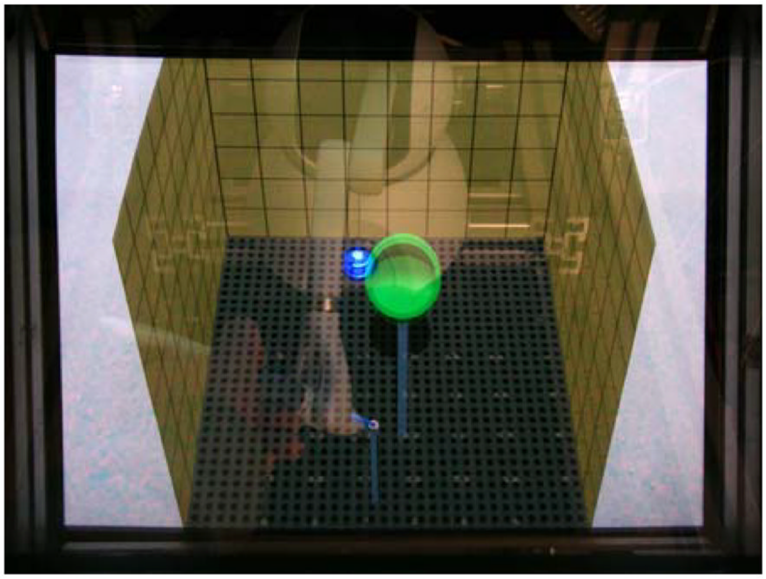


Figure 2.
First person view of the co-located fish tank display through a semi-transparent mirror. The semi-transparent mirror was used for calibration only.

Figure 3.
The experimental setup for the physical target experiment condition. The home target is the small stem centered farthest away from the haptic device's base and an example target is the hollow half-sphere resting on a stem.

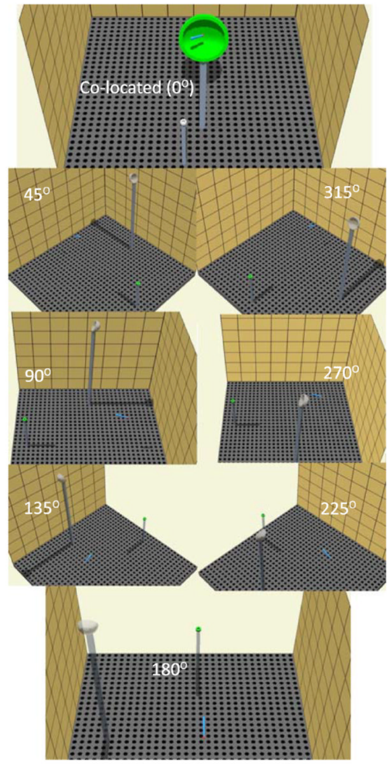


Figure 4.
The co-located and rotated conditions, rotated about the azimuth.

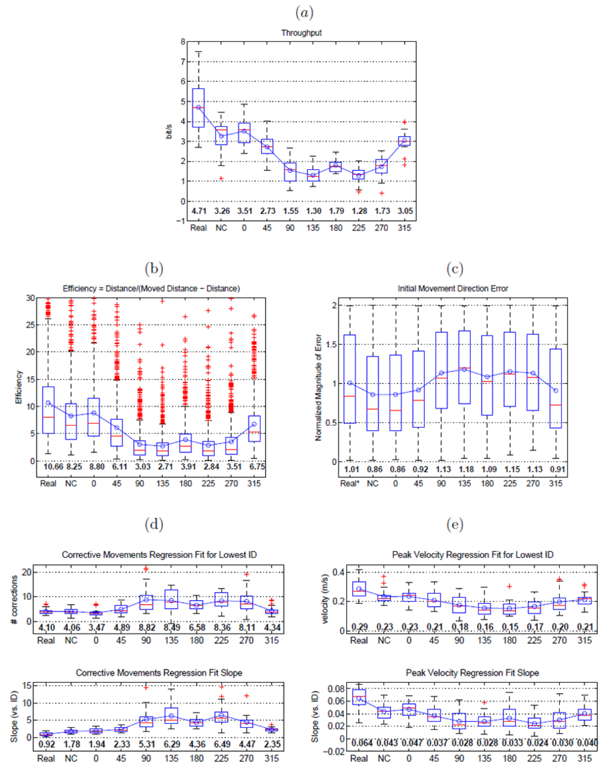
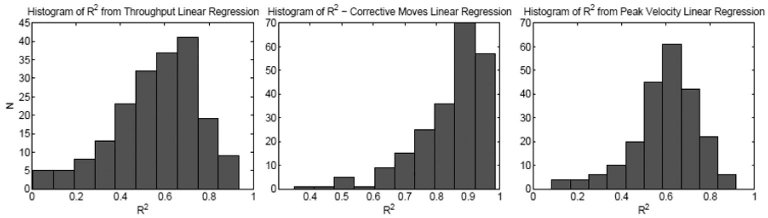


Figure 5.

Boxplots for performance measures along with mean values (bold under each box, plotted with blue circles). From the lower boundary of each box, moving upward, are the lower quartile, median (red line), and upper quartile. Lower and upper whiskers of each box represent 1.5× the lower and upper quartiles. Red cross symbols that fall beyond the whisker boundaries are datapoints that exceed 1.5× the upper and lower quartiles. *Denotes data that was not analyzed due to calibration issues.



(a) (b) (c)

Figure 6.

Histograms of the coefficient of determination (R^2) for linear regressions on a) throughput, b) corrective movements, and c) peak velocity as a function of ID.

Table 1

Target List

ID (bits)	1.9	2.3	2.7	3.2	3.6	4.1	4.6	5.1	5.6	6.0
Distance (cm)	5.6	8.0	6.7	9.6	9.2	13	18	9.7	14	19
Diameter (mm)	20	20	12	8.0	8.0	8.0	3.0	3.0	3.0	3.0

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 2

Performance Means and (Std. Dev.) vs Real and 0° Baselines

	TP	Eff.	Init Err.	CM offset	CM slope	PV offset	PV slope
Real	*1.0 (1.0)*	*1.0 (1.0)*	N/A	*1.0 (1.0)*	*1.0 (1.0)*	*1.0 (1.0)*	*1.0 (1.0)*
NC	0.69 (0.68)	0.77 (0.80)	*1.0 (1.0)*	0.99 (1.12)	1.94 (1.08)	0.81 (0.75)	0.67 (0.67)
0 °	0.75 (0.52)	0.83 (0.79)	1.00 (0.99)	0.85 (1.21)	2.11 (1.48)	0.82 (0.75)	0.73 (0.72)
0 °	*1.0 (1.0)*	*1.0 (1.0)*	*1.0 (1.0)*	*1.0 (1.0)*	*1.0 (1.0)*	*1.0 (1.0)*	*1.0 (1.0)*
45 °	0.78 (0.83)	0.70 (0.84)	1.07 (1.00)	1.28 (1.48)	1.25 (1.18)	0.78 (1.05)	0.88 (1.17)
90 °	0.44 (0.99)	0.34 (0.61)	1.32 (0.94)	2.12 (3.90)	2.72 (5.20)	0.58 (1.15)	0.74 (1.31)
135 °	0.37 (0.60)	0.31 (0.44)	1.37 (0.92)	2.09 (2.98)	3.23 (5.48)	0.58 (1.02)	0.66 (1.14)
180 °	0.51 (0.54)	0.44 (0.62)	1.26 (0.98)	1.67 (1.49)	2.29 (2.46)	0.69 (1.53)	0.65 (1.06)
225 °	0.37 (0.65)	0.32 (0.51)	1.34 (0.94)	2.07 (2.48)	3.28 (5.71)	0.52 (1.38)	0.69 (1.49)
270 °	0.49 (0.85)	0.40 (0.72)	1.32 (0.96)	1.98 (3.71)	2.27 (3.53)	0.66 (1.37)	0.82 (1.49)
315 °	0.87 (0.78)	0.77 (0.83)	1.06 (1.02)	1.15 (1.33)	1.20 (1.07)	0.87 (1.00)	0.90 (0.91)