Effect of Visuo-Haptic Co-location on 3D Fitts' Task Performance

Michael J. Fu, Andrew D. Hershberger, Kumiko Sano, and M. Cenk Çavuşoğlu

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 haptic workspaces using immersive virtual reality (VR) modalities. Significant understanding has been gained by previous work comparing physical tasks against VR tasks, visuo-haptic co-location versus non-colocation, and even visuo-haptic rotational misalignments in VR. However, few studies have explored all of these paradigms in context with each other and it is difficult to do inter-study comparisons because of the variation in tested motor tasks. Therefore, the goal for the current study was to characterize human performance of Fitts' point-to-point reaching task – an established test of manual performance – in the physical, co-located/non-colocated VR, and rotated VR visualization conditions.

A key finding was the significant decrease observed in end-point error for tasks performed in a co-located virtual reality environment. The results also showed cyclic performance degradations due to rotational visuo-haptic misalignments that were consistent with trends reported by the literature.

I. INTRODUCTION

Optimization of human task performance in virtual reality (VR) 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 VR immersion 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 haptic workspace are not aligned (as in a common computer display and mouse interface). Investigations into the effect of visuo-haptic misalignments on task performance are rooted in motor control studies regarding the physiological processes behind adaptations to optical prisms[7], [17]. Since then, the increased accessibility of computers and VR brought the field into intersections with the study of human-computer interaction, where the focus is on optimization of human performance in virtual environments – which is also the focus of the current work.

There are two general causes of 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 haptic workspace scales, origins, and orientations are aligned (similar to our

Michael J. Fu (mjfu@case.edu), Kumiko Sano (kxs305@case.edu) and M. C. Çavuşoğlu (cavusoglu@case.edu) are with the Dept. of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, OH, USA. Andrew D. Hershberger was with CWRU. This work was supported in part by NSF under grants CNS-0423253, IIS-0805495, IIS-0905344, and CNS-1035602, NIH under grant R21HL096941, and EC under grant FP7-IRG-2009-247470.

physical hand-eye interactions). Therefore, translational misalignment refers to the condition where only the scales and orientations are aligned, and rotational misalignment refers to the case where only the scales and origins are aligned. In this paper, the co-located condition was the $0°$ rotation condition.

Previous findings regarding the effect of rotational misalignment on virtual task performance are surprisingly consistent, despite the notable variation in the type of tasks that were tested. The tasks include 2D point-to-point targeting with a joystick interface [1], 3D pick-and-place and tracking using two joysticks [9], [8], whole-arm 3D point-to-point reaching [2], and 3D object orientation matching [24]. 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 $\pm 90°$, and decreased to a local minimum at $180°$. In short, the literature found that task performance was best when there was no visuo-haptic rotational dislocation and poorest when the dislocation was $\pm 90°$ about the azimuth.

In contrast, previous findings on the effect of translational misalignments on task performance are conflicting. For instance, Swapp *et al.* reported that co-location significantly improved performance metrics for a set of 3D tasks [21]. 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 *et al.* reported that a virtual endoscopic surgery suturing task was performed significantly faster using a stereographic, co-located fish tank VR modality than a monoscopic, non-colocated monitor [10]. Their co-located modality placed the mirrored display between the user and the haptic device used for input. In contrast, Teather *et al.* tested the effect of co-location using a 3D Fitts' task and did not find a significant improvement in task completion time or end-point error [22]. 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.

The cited works have provided useful information regarding human performance in VR, but several gaps in the literature are apparent. 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 two studies have taken this into consideration, [21], [22]. Of those two, only one specified how difficulty was defined [22]. Most importantly, although rotation and translation misalignments have been shown to impact task performance with respect to the same task in physical reality – only one attempt has been made to investigate all three factors using the same task [2]. The attempt by Blackmon *et al.* was a small study of 4 subjects that examined 0, 45, and 90 rotations.

Therefore, the goal for the current study was to characterize human performance of a point-to-point reaching task in the physical, co-located/non-colocated VR, and rotated VR visualization conditions. Also, the reaching tasks should span a range of difficulties, but still facilitate inter-study comparison and repeatability. 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 [4]. 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 it takes for a human to move and point to a target is

$$MT = a + b \log_2 \left(1 + \frac{D}{W} \right),$$ (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 $\log_2(1 + \frac{D}{W})$ is called the index of difficulty (ID). ID is a measure of the difficulty of the motor task, and carries the unit of 'bits,' in reference to an information theoretic interpretation of Fitts' Law.

To use Fitts' task to evaluate human performance in different experimental conditions, Soukoreff and Mackenzie recommended the use of 'throughput' (TP) [19]. Throughput was defined as

$$TP = \frac{1}{y} \sum_{i=1}^{y} \left(\frac{1}{x} \sum_{j=1}^{x} \frac{\mathrm{ID_{eij}}}{MT_{ij}} \right),$$ (2)

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

$$\mathrm{ID_e} = \log_2 \left(1 + \frac{\mathrm{D_e}}{\mathrm{W_e}} \right),$$ (3)

where D_e is the average distance traveled for multiple repetitions of the same movement condition and $W_e = \sigma\sqrt{2\pi e} = 4.133\sigma$, where σ was the standard deviation of the end point locations. This formulation for $\mathrm{ID_e}$, detailed in [19], assumes that the end-point error has a normal random distribution since it is due to human error.

Fitts' law was originally formulated for 1D motion, but has since been extended into 2D for use in evaluating computer input devices and graphic user interfaces. The applicability of Fitts' Law to human-computer interface research is generally accepted, as evidenced by it's adoption as ISO standard 9241-9 in 2000. For 3D tasks, there is strong evidence that Fitts' law applies and can even be adapted for the complications involved with 3D reaching.

Murata and Iwase proposed that a third parameter can be added to Fitts' Law to better account for the target's angle with respect to the horizon [15]. Grossman and Balakrishnan, proposed a version of Fitts' Law that was parameterized for 3D rectangular-shaped targets and the azimuth angle of reach, which they studied using a volumetric virtual display [6]. Liu *et al.* proposed to add horizon angles, azimuth angles, and path curvature parameters to Fitts' Law. They verified this using a virtual 3D tunnel steering task requiring subjects to move a cursor through paths with constant curvature on a stereoscopic display [11].

These extensions to Fitts Law have been shown to accurately predict task completion time. However, with so many parameters, it is not clear how to define the concept of task throughput using these new models. Therefore, the current work maintains the use of Shannon's formulation of Fitts' law and Mackenzie's throughput, from (1) and (2), for comparing experimental conditions.

A. Study Objectives

This work examined human performance of a 3D variation of Fitts' point-to-point reaching task performed using a stylus-based haptic interface device under various experimental conditions. A total of ten conditions of the reaching task were considered: physical targets (real), non-colocated (NC) virtual targets, co-located virtual targets ($0°$) using a stereographic fish tank VR display modality, and virtual targets using fish tank VR with azimuth perspective rotations of 45, 90, 135, 180, 225, 270, and $315°$. Also, a range of task difficulties, ID 2–6, were used.

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

Objective 2: Examine the effect of visual rotations (0–$315°$) on task performance measures.

II. PERFORMANCE MEASURES FOR ANALYSIS

The following quantitative measures were 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 [16]. 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`).

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.

A. Throughput

Throughput, also referred to as task completion rate, is inversely proportional to the task completion times measured across a range of target difficulties. 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. In order to perform statistical analyses, TP was computed for each subject over each condition and had units of 'bits/s'.

B. End-point Error

End-point error was defined as the Euclidean distance from the location of movement termination to the target's central location, without regard for the width of the target. Since healthy, unimpaired individuals were tested for this study, increased end-point error was equated with decreased task performance. While direction of the error is also worth examining and may be affected by the visual rotations, the current work focuses only on the magnitude of end-point error.

III. METHODS

The following methods have been reviewed and provided exemption status by the institution's Internal Review Board.

A. Equipment

The fish tank VR modality was selected for its ability to provide a high-fidelity virtual environment that aligns the visual and haptic workspaces while minimizing user fatigue. The visual and haptic workspaces were co-located by placing a haptic device behind the image plane of the calibrated fish tank VR setup. In this way, the haptic device's representation in the virtual environment will appear to match both the motion and location of the physical device.

A custom fish tank VR setup (Fig. 1) was designed, which was reconfigurable for the physical task, co-located VR, and non-colocated VR 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 the experiment conditions.

OpenGL was used to develop a VR 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.
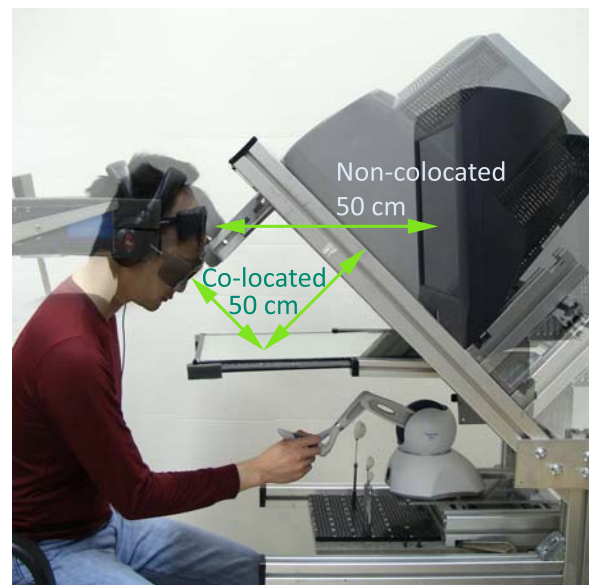

Fig. 1. User positioning for the fish tank VR setup used for the 0–315° conditions (tilted monitor with user facing downward) and non-colocated condition (upright monitor with subject looking forward and head stabilized by a custom headrest). The setup for the physical condition required removing the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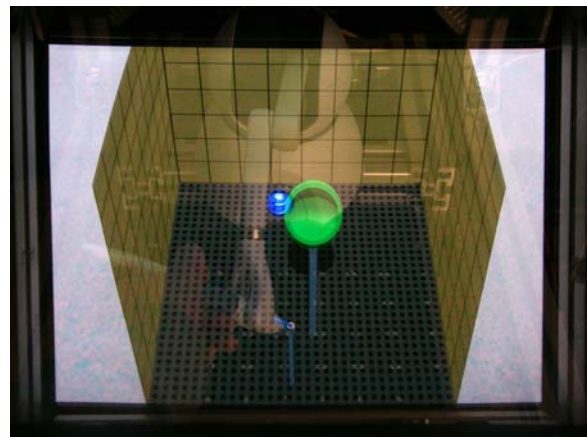

Fig. 2. First person view of the co-located fish tank VR setup through a semi-transparent mirror. The semi-transparent mirror was used for calibration only.

Stereographic calibration between the virtual and physical workspaces was performed manually by matching the appearance of several physical objects seen through a half-mirror with their virtual representations with respect to a fixed forehead rest used by all subjects (Fig. 2). The calibration objects included a cylinder (2 in. diameter x 1 in. height), a cuboid (6 in. x 4 in. x 1 in.), the pegboard used for setting physical targets, and several targets placed at the extremes of the workspace. However, 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 20" (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 approximately 26 cm wide, 15 cm deep, and 15 cm high.

B. 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' participation spanned two separate days – one day consisting of the rotated conditions and another day consisting of the physical and non-collocated conditions.

C. Experiment Paradigms

For each of the ten conditions, subjects were asked to sit at a fish tank VR station, grip the haptic device stylus like a pen, and perform the following task quickly, but as accurately as possible. 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, a home position and one target were displayed simultaneously to the subject. Each subject was then 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, and press again when the tip of the stylus was within the target volume. Each button press triggered a chime sound effect. Ample rest was provided to subjects between different paradigms, but no rest was provided between the practice and actual test runs in order to maintain the subject's familiarity with the specific paradigm.

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 (Table I), 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. Performance measures for all four repetitions were averaged for the analyses.

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 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

TABLE I
TARGET LIST

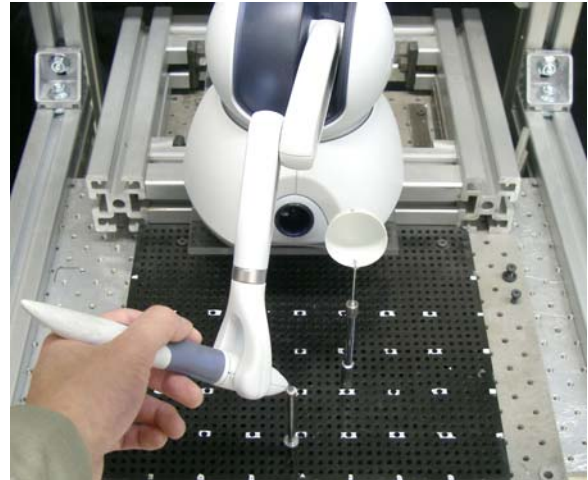| ID (bits)	Distance (cm)	Diameter(mm)
1.9	5.6	20
2.3	8.0	20
2.7	6.7	12
3.2	9.6	12
3.6	9.2	8.0
4.1	13	8.0
4.6	18	8.0
5.1	9.7	3.0
5.6	14	3.0
6.0	19	3.0

Fig. 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.

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 user so that the stylus tip becomes visible to the user. This alteration to the targets was not needed for the VR 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 screen shots from all eight rotation conditions. The 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 or error.

It is important to note that force feedback was not provided for the virtual targets in order to evoke a vision-based motor control from the subject and record error rates that are not affected by contact-based strategies where subjects might search around for haptic contact with the target before deciding to register the end point click.
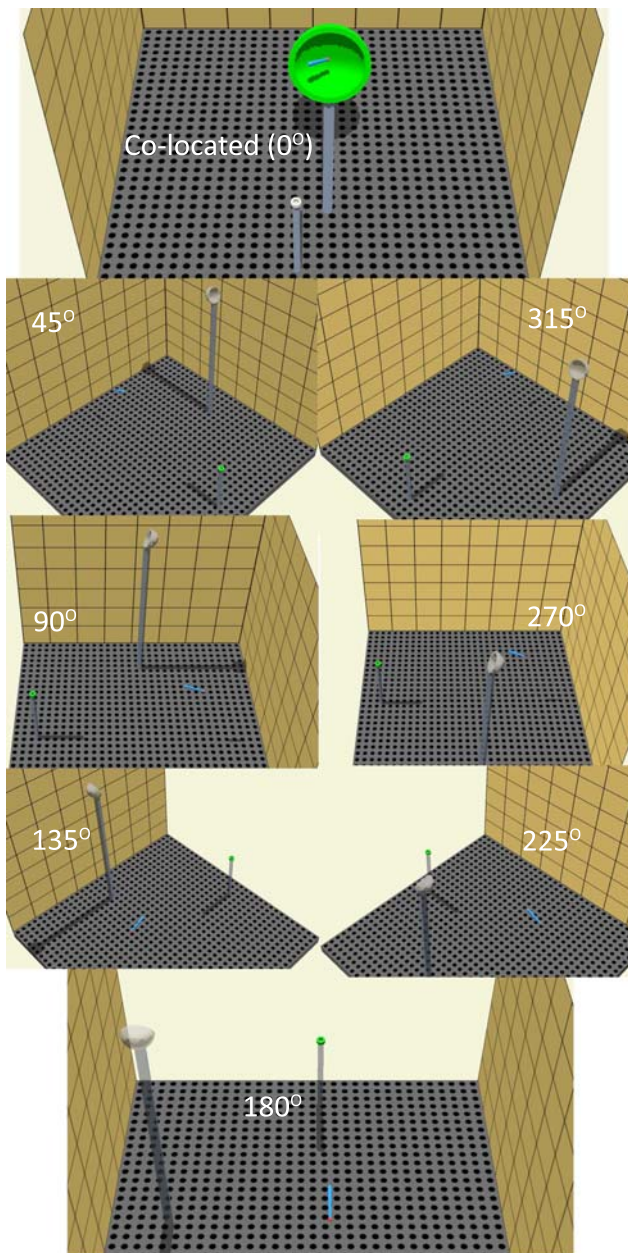
Fig. 4. The co-located and rotated conditions, rotated about the azimuth at 45, 90, 135, 180, 225, 270, and 315°.

Non-colocated: The NC experiment condition, used the GUI for the 0° condition, but was viewed directly on the computer display as shown in Fig. 5.

IV. RESULTS

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).
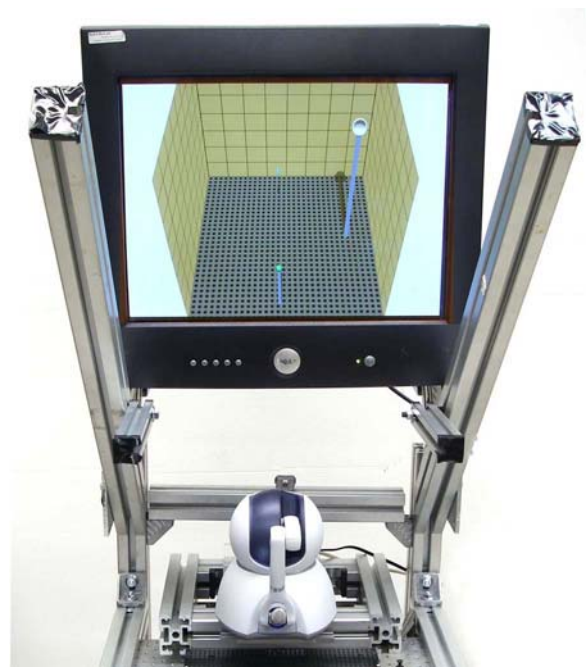

Fig. 5. Non-colocated condition first person view.

Statistical power for the visualization paradigms was computed to be 0.69 (as calculated by G*Power 3.1 [3], 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).

A. 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 ($p \leq 0.001$, $F(9,189) = 253.1$) and rotations ($p \leq 0.001$, $F(9,189) = 169.6$).

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 (Fig. 6). A histogram of all linear regression R^2 values are shown in Fig. 7. Also, the TP for all subjects were reported as boxplots in Fig. 8.

In addition, mean differences in TP for each condition were tested using ANOVA with experiment condition as the within-subjects factor and throughput as the dependent variable. The paradigm was found to have a significant effect on TP ($p=0.0016$, $F(2,42)=15.96$). Highest mean throughput 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 throughput values were not found to be significantly different.

Rotations were also found to significantly affect throughput ($p=0.00$, $F(7,147)=81.66$). Significant multiple comparison results for the rotations are reported in Table II. At 0°, TP was significantly higher than all the other rotations. Also, TP between the 0, 45 (2.73 b/s), and 90° (1.55 b/s) conditions
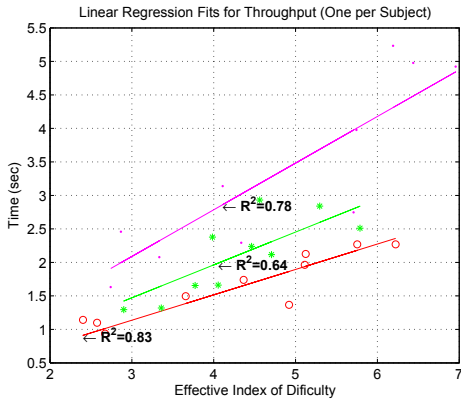
Fig. 6. For conciseness, several examples of a one-parameter slope fit to the movement time data. Each line and set of points denoted by a marker type represents the movement time data for a subject at a particular adjusted index of difficulty (ID_e).
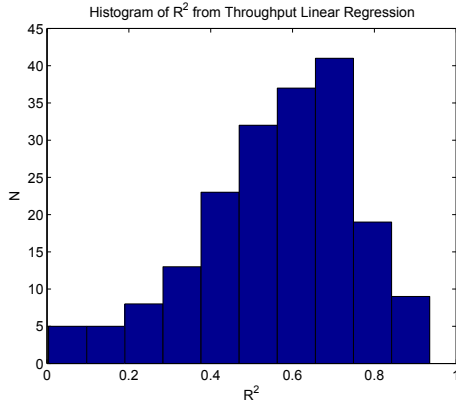

Fig. 7. Histogram of R^2 results from all TP linear regressions as a function of ID_e.

TABLE II
SIGNIFICANT MULTIPLE COMPARISONS – THROUGHPUT

$0° - 45°$	$45° - 90°$	$90° - 315°$	$135° - 180°$
$0° - 90°$	$45° - 135°$		$135° - 270°$
$0° - 135°$	$45° - 180°$		$135° - 315°$
$0° - 180°$	$45° - 225°$		
$0° - 225°$	$45° - 270°$		
$0° - 270°$	$45° - 315°$		
$0-315°$			

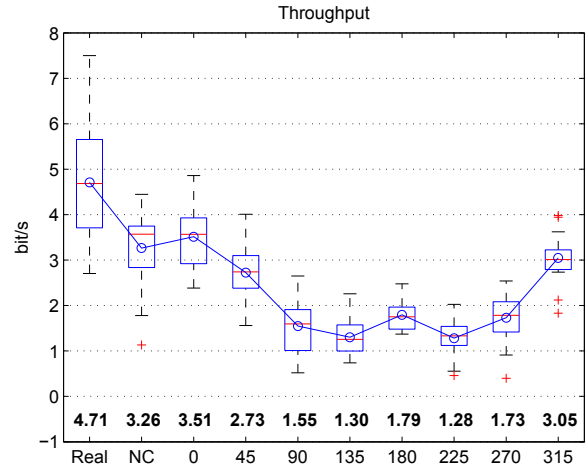| $180° - 225°$	$225° - 270°$	$270° - 315°$	
$180° - 315°$	$225° - 315°$		

Fig. 8. Boxplots of throughput (ID_e/Movement Time) computed for each experimental condition (real targets, non-colocated VR, and fish tank VR with rotations 0–315°). Higher throughput infers better performance. The bold number below each boxplot is the mean value denoted by the blue, circle markers. The red line inside each box is the median, the lower and upper edges of the box mark the 25% and 75% quartiles, respectively, and the lower and upper horizontal bars represent 1.5x less than the 25% quartile and 1.5x greater than the 75% quartile, respectively. The red cross markers represent data outside the 1.5x quartile ranges.

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 throughput 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), throughput increased almost to the 0° level. It is noteworthy that throughput at 180° (1.79 b/s) was significantly higher than both the 135 and 225° conditions, which is where the lowest throughput values occurred. Mean throughput for 135 and 225° were not significantly different from each other.

B. End-Point Error

End-point error was not observed to be affected by target ID, so it was plotted in (Fig. 9) without distinguishing each target by difficulty. Statistical tests for significant mean differences were computed using experiment condition as the within-subjects factor and end-point error as the dependent variable.

Paradigm had a significant effect on end-point error (p $\leq$ 0.001, F(2,42) = 54.9). End-point error was significantly higher for the NC condition versus the co-located case. From Fig. 9, it would appear that error was highest for the real targets, but this is in fact not true and was possibly due to a calibration limitation with the Phantom Omni haptic device. 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 can result in slight calibration errors between the physical joint angles and those reported by the Phantom's application programming interface. Therefore, the end-point error for the real target condition is was not considered reliable and was not included in the analysis.

Rotations were also found to significantly impact end-point error (corrected p = 0.025, F(4,147) = 2.905). Signifi-

TABLE III
SIGNIFICANT MULTIPLE COMPARISONS – END-POINT ERROR

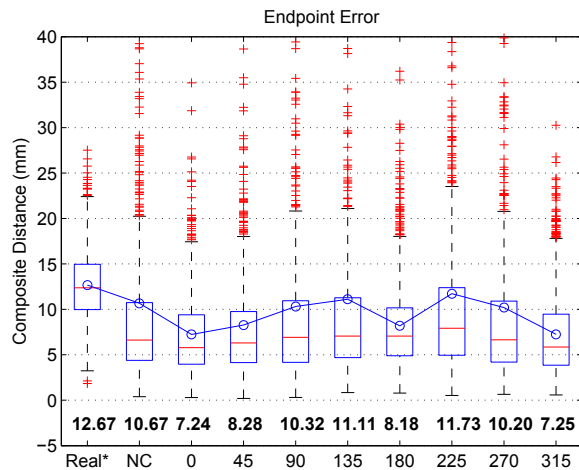|			
0°– 90°	45°– 90°	90°– 180°	135°– 180°
0°– 135°	45°– 135°	90°–315°	135°– 315°
0°– 225°	45°– 225°		
	45°– 270°		
180°– 225°	225°– 270°	270°–315°	
180°– 270°			

Fig. 9. Boxplots for the composite end-point error (distance from end location to the target) for each experimental condition (real targets, non-colocated VR, and fish tank VR with rotations 0–315°). Lower error infers better performance. The bold number below each boxplot is the mean value denoted by the blue, circle markers. The red line inside each box is the median, the lower and upper edges of the box mark the 25% and 75% quartiles, respectively, and the lower and upper horizontal bars represent 1.5x less than the 25% quartile and 1.5x greater than the 75% quartile, respectively. The red cross markers represent data outside the 1.5x quartile ranges. ∗ The real condition was excluded for analysis due to calibration issues.

cant multiple comparison results for the rotations are reported in Table III. Mean end-point error was observed to increase from 0° (7.24 mm) to a local maximum at 135° (11.11 mm). After 135°, end-point error significantly decreased to 8.18 mm at 180° before rising to another significantly higher local maximum of 11.73 mm at 225°. From 225–315°, mean end-point error decreased down to 7.25 mm, which was not significantly different from the 0° condition.

The highest mean end-point error (not including the real target case) occurred in the 225° condition and the lowest was in the 0° condition.

V. DISCUSSION

A. Real vs. Non-colocated VR vs. Co-located VR

Unsurprisingly, the results indicated that performance was significantly better for the physical task than the NC and co-located conditions. Subjects were able to accomplish the physical reaching task at a faster rate and with less end-point error than the virtual conditions. Mean throughput for the real task was 1.4x higher than the NC condition and 1.3x higher

than the co-located condition. This indicates that across target ID 2–6, task completion times were significantly lower for the real targets versus the virtual ones. This was in line with previous reports of 1.5–2x decreases in completion time for 2D physical reaching tasks versus virtual by [5], [13] and [20].

Less expected, however, was that co-location exhibited significantly lower end-point error than the NC condition, while TP was not. End-point error was 1.5x lower for the co-located condition than the NC condition. Mean values for TP were higher for the co-located condition than for the NC condition, but did not reach statistical significance. The significant effect of co-location on end-point error differs from the findings of Sprague *et al.* , which tested a 2D Fitts' task and did not find an effect of co-location on error, [20]. However, they compared a co-located condition against three levels of poor co-location that translated the VR stylus and targets farther away from the user (visually scaled smaller).

In contrast, the current study's co-located and NC conditions minimized scaling effects by ensuring a 50 cm distance between the eyes and the image plane in both conditions. Therefore, the detected difference in end-point error likely cannot be attributed to viewpoint scaling.

It is possible that changes in visuo-motor processes are responsible for the difference in end-effector error. Humans spend years unconsciously tuning their motor control strategies to the ideal condition where the visual field and haptic workspace are aligned. So, it is likely that reaching tasks in the NC condition requires some sort of cognitive re-mapping. The NC condition may require a re-mapping of the perceptual processes responsible for converting visual differences between the hand and the target into muscle forces necessary to make a movement to close the gap [18]. This re-mapping may be minimal for the co-located condition, assuming that the main new mapping is to convert hand movements to virtual cursor movements. However, re-mapping for the NC condition may require a visual field-to-haptic workspace transform that is unusual and may add to cognitive load.

Assuming this is true and noting the fact that task completion rates did not suffer, it is plausible that accuracy does. This theory was echoed in [14], that suggests VR interfaces should take advantage of the highly accurate body-relative proprioception by keeping manipulated virtual objects within arm's reach. They found that, when objects were outside of the arm's workspace, completion times significantly increased for a virtual object docking task where error was not possible. Similarly, this type of effect has also been reported in physical reality for a pick-and-place task (which also does not allow for error) performed with endoscopic graspers. Tendick *et al.* measured faster task completion times when surgeons viewed the workspace directly through one eye (and one eye patched) versus through a monoscopic video feed from an endoscopic camera [23].

Although verification of the above theory is beyond the scope of the current work, evidence from psychophysics work suggests that error rates caused by artificially im-

posed visual and haptic misalignments can be overcome by practice. It has been demonstrated through optical prism experiments that subjects who practice for several days can overcome initially significant errors and throw balls at targets with the same accuracy regardless of whether they are wearing prisms or not [18], [12].

B. Effect of Azimuth Rotations

Azimuth rotations impacted TP and end-point error in a cyclic manner similar to the trends reported by the literature.

Compared to the lowest error condition at $0°$, rotations did not significantly cause increased end-point error until $90°$. Also, end-point error means appeared to be symmetric about $180°$. There was no significantly difference between 135 and $225°$, 90 and $270°$, and 45 and $315°$ conditions. The cyclic trend can be seen in Fig. 9 to increase from $0°$ to a peak at $135°$, and return to a level not significantly different from $0°$ at $180°$. The same behavior, but mirrored about $180°$, occured from 180–315°.

In contrast, throughput was significantly decreased from $45°$ onward and was not symmetric about any rotation angle, as significant differences between means were detected between 135 and $225°$, 90 and $270°$, and 45 and $315°$ conditions.

The major difference between the current findings and the previous work was that poorest task performance occurred for the $90°$, $135°$, and $225°$ conditions, compared to 90 and $270°$ in previous findings. However, this does not contradict psychophysics literature, which reported that poorest manual performance occurs for visual rotations in the range of 90–$135°$ or 225–$270°$ for physical tasks under camera rotations [18].

VI. Conclusion – System Design Implications

Several system-design recommendations can be gathered from the results of this study. First, if end-point error is of major concern, then a co-located VR configuration is recommended over a non-colocated modality. Second, the desire is to optimize task throughput for rotated perspectives in fish tank VR, then haptic and visual perspectives should be aligned, as visual rotations of $45°$ in either direction significantly impacted these measures. However, if only end-point error is of concern, then visual rotations of up to $\pm 90°$ may be acceptable before significant effects on performance might be detected.

References

[1] R. K. Bernotat, "Rotation of visual reference systems and its influence on control quality," *IEEE Transactions on Man-Machine Systems*, vol. 11, no. 2, pp. 129–131, 1970.

[2] T. T. Blackmon, M. C. Çavuşoğlu, F. Lai, and L. W. Stark, "Human Hand Trajectory Analysis in Point-and-Direct Telerobotics," in *Proceedings of the 8th International Conference on Advanced Robotics (ICAR'97)*, July 1997, pp. 927–932.

[3] F. Faul, E. Erdfelder, A.-G. Lang, and B. Axel, "G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences." *Behavior Research Methods*, vol. 39, no. 2, pp. 175–191, 2007.

[4] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *Journal of Experimental Psychology*, vol. 47, pp. 381–391, 1954.

[5] E. D. Graham and C. L. MacKenzie, "Physical versis virthal pointing," *Proc. of the 2001 ACM Conference on Human Factors in Computing Systems*, pp. 292–299, April 1996.

[6] T. Grossman and R. Balakrishnan, "Pointing at trivariate targets in 3D environments," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ser. CHI '04. ACM, 2004, pp. 447–454.

[7] R. Held, "Plasticity in sensory-motor systems." *Scientific American*, vol. 213, no. 5, pp. 84–94, 1965.

[8] W. S. Kim, s. R. Ellis, M. E. Tyler, B. Hannaford, and L. W. Stark, "Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks," *IEEE Transactions on Man-Machine Systems*, vol. 17, no. 1, pp. 61–72, 1987.

[9] W. S. Kim, F. Tendick, and L. W. Stark, "Visual enhancements in pick-and-place tasks: Human operators controlling a simulated cylindrical manipulator," *IEEE Journal of Robotics and Automation*, vol. 3, no. 5, pp. 418–425, 1987.

[10] D. D. Lev, R. Rozengurt, T. Gelfeld, A. Tarkhnishvili, and M. Reiner, "The effects of 3d collocated presentation of visuo-haptic information," in *Proc. of EuroHaptics 2010*, 2010, pp. 432–437.

[11] L. Liu, J.-B. Martens, and R. v. Liere, "Revisiting Path Sterring for 3D Manipulation Tasks," in *Proc. of the IEEE Symposium on 3D User Interfaces 2010*, 2010, pp. 39–46. [Online]. Available: http://dx.doi.org/10.1109/VRAIS.1993.380802

[12] T. A. Martin, J. G. Keating, H. P. Goodkin, A. J. Bastian, and T. W. T., "Throwing while looking through prisms." *Brain*, no. 119, pp. 1199–1211, 1996.

[13] A. H. Mason, M. A. Walji, E. J. Lee, and C. L. MacKenzie, "Reaching movements to augmented and graphic objects in virtual environments," *Proc. of the 2001 ACM Conference on Human Factors in Computing Systems*, pp. 426–433, March 2001.

[14] M. R. Mine, F. P. Brooks, Jr., and C. H. Sequin, "Moving objects in space: exploiting proprioception in virtual-environment interaction," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, ser. SIGGRAPH '97. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1997, pp. 19–26.

[15] I. Murata, Atsuo and H. Iwase, "Extending fitts' law to a three-dimensional pointing task," *Human Movement Science*, vol. 20, pp. 791–805, 2001.

[16] C. N. Riviere, W. T. Ang, and P. K. Khosla, "Toward active tremor canceling in handheld microsurgical instruments." *IEEE Transactions on Robotics*, vol. 19, no. 5, pp. 793–800, 2003.

[17] R. Shadmehr and S. P. Wise, *The Computational Neurobiology of Reaching and Pointing, A Foundation for Motor Learning*, ser. Computational Neuroscience Series, T. J. Sejnowski and T. A. Poggio, Eds. Cambridge, MA: The MIT Press, 2005.

[18] ——, *The Computational Neurobiology of Reaching and Pointing, A Foundation for Motor Learning*, ser. Computational Neuroscience Series, T. J. Sejnowski and T. A. Poggio, Eds. Cambridge, MA: The MIT Press, 2005.

[19] R. W. Soukoreff and I. S. MacKenzie, "Towards a standard for pointing device evaluation: Perspectives on 27 years of fitts law research in hci." *International Journal of Human-Computer Studies*, no. 61, pp. 751–789, 2004.

[20] D. W. Sprague, B. A. Po, and K. S. Booth, "The importance of accurate vr head registration on skilled motor performance," in *Proc. of Graphics Interface 2006*, 2006, pp. 131–137.

[21] D. Swapp, V. Pawar, and C. Loscos, "Interaction with co-located haptic feedback in virtual reality." *Virtual Reality*, no. 10, pp. 24–30, 2006.

[22] R. J. Teather, R. S. Allison, and W. Steurzlinger, "Evaluating visual/motor co-location in fish-tank virtual reality," in *Proceedings of the IEEE Toronto International Conference - Science & Technology for Humanity*. IEEE, 2009, pp. 624–629.

[23] F. Tendick, R. W. Jennings, G. Tharp, and L. Stark, "Sensing and manipulation problems in endoscopic surgery: Experiment, analysis, and observation," *Presence*, vol. 2, no. 1, pp. 66–79, 1993.

[24] C. Ware and R. Arsenault, "Frames of reference in virtual object rotation," in *Proceedings of the 1st Symposium on Applied perception in graphics and visualization*. New York, NY, USA: ACM, 2004, pp. 135–141.