Novel Metrics for 3D Remote Pointing

Steven J. Castellucci
York University
4700 Keele St.
Toronto, ON, Canada
stevenc@cse.yorku.ca

Robert J. Teather
York University
4700 Keele St.
Toronto, ON, Canada
rteather@cse.yorku.ca

Andriy Pavlovych
York University
4700 Keele St.
Toronto, ON, Canada
andriyp@cse.yorku.ca

ABSTRACT
We introduce new metrics to help explain 3D pointing device movement characteristics. We present a study to assess these by comparing two cursor control modes using a Sony *PS Move*. "Laser" mode used ray casting, while "position" mode mapped absolute device movement to cursor position. Mouse pointing was also included, and all techniques were also analyzed with existing 2D accuracy measures. Results suggest that position mode shows promise due to its accurate and smooth pointer movements. Our 3D movement metrics do not correlate well with performance, but may be beneficial in understanding how devices are used.

Categories and Subject Descriptors
H.5.2 Information interfaces and presentation (e.g., HCI): User Interfaces—evaluation/methodology.

General Terms
Measurement, Performance, Human Factors.

Keywords
3D measures, remote pointing, evaluation.

1. INTRODUCTION
Remote pointing is becoming more common, largely due to the recent availability of inexpensive multi-DOF game controllers. Many researchers are using remote pointing in both 2D user interfaces [5, 13] and in virtual reality systems [7, 16]. Thus there is interest in determining how effective these devices are for pointing. Existing measures such as throughput indicate device performance, but do not explain device movement characteristics.

We conducted a study to evaluate the pointing efficiency of the *PlayStation Move* as a representative remote pointing device. The task required controlling a 2D cursor to select targets on a large display. The study included two distinct cursor control modes using the *Move*. The first technique, "laser mode" positions the cursor where the *Move* is pointed. The second technique maps absolute device movement to cursor control. Mouse pointing was included as a known benchmark of pointing performance. The techniques were compared using the ISO 9241-9 standard [3].

We chose a simple pointing task because it is common to many interaction methods. In contrast, a 3D docking task would also gauged participants' spatial and problem-solving skills, rather than focus on device characteristics. Further benefits of the standard are presented in the Related Work section.

We include a detailed analysis of device motion to help explain performance differences. Our analysis includes 2D measures developed by MacKenzie et al. [9]. Moreover, we propose and validate similar measures for use with 3–6DOF input devices. Our goal is to supplement existing 2D measures with tools specifically designed to investigate higher dimensional input devices.

2. RELATED WORK
Pointing interfaces are often evaluated in the context of Fitts' law [1], an empirical model of the well-known tradeoff between speed and accuracy in pointing tasks. The model is given as follows:

$$MT = a + b \cdot ID, \quad \text{where} \quad ID = \log_2(A/W + 1) \qquad (1)$$

MT is movement time, *A* is target amplitude (distance), and *W* is target size, while *a* and *b* are empirically derived. The log term is the index of difficulty (*ID*) and indicates pointing task difficulty.

The ISO 9241-9 standard suggests using "effective" measures, a post-experiment correction to adjust the error rate to 4%. This enables the computation of throughput, a measure that incorporates both speed and accuracy by "normalizing" the accuracy. Throughput is computed as follows:

$$TP = \frac{\log_2(A_e/W_e + 1)}{MT}, \quad \text{where} \quad W_e = 4.133 \cdot SD_x \qquad (2)$$

MT is the average movement time. A_e, effective distance, is the average movement distance. Effective width, W_e, is computed by projecting the cursor onto the task axis (the line between subsequent targets) and taking the standard deviation (SD_x) of these distances multiplied by 4.133. This assumes that movement endpoints are normally distributed around the target center and 4.133 (±2.066) standard deviations (i.e., 96%) of clicks hit the target [4]. W_e corrects error rate to 4%, and allows comparison between studies with differing error rates [8]. Throughput exhibits low variability for the same condition between studies [14, 18], improving comparability. For example, previous work [15] found mouse throughput was consistent across three different 3D pointing tasks. Exclusively measuring movement time can be unreliable as it varies at the expense of accuracy.

Mouse pointing throughput is typically higher than remote pointing throughput [10, 12, 15]. Still, there is interest in using remote pointing in both 2D [5, 10, 11, 13] and 3D [2, 6, 15, 17] user interface research. Our primary goal is not to re-establish the performance differences between the mouse and remote pointing. Instead, we propose and validate metrics to characterize 3D movements. This should provide better insight into *why* performance differences occur.

Metrics to evaluate 2D pointing movement characteristics were proposed by MacKenzie [9]. These measures are taken relative to the task axis and reported as per-trial averages. The first four metrics are discrete measures. *Target Re-Entry* (TRE) is the number of times the cursor re-entered the target. *Task Axis Crossing* (TAC) represents the number of times the cursor crosses the task axis. *Movement Direction Change* (MDC) quantifies the number of direction changes that occur relative to the task axis, while *Orthogonal Direction Change* (ODC) counts direction changes orthogonal to the task axis. The final three metrics are continuous measures measured in pixels. *Movement Variability* (MV) represents how parallel the traversed path is to the task axis. *Movement Error* (ME) indicates the scalar deviation from the task axis, while *Movement Offset* (MO) is the non-scalar deviation from the task axis.

These metrics provide additional insight into why performance (throughput) scores vary between 2D pointing techniques. We propose similar measures to characterize 3D movements, and then experimentally assess the value of these. In previous studies [2, 15, 17] researchers provide qualitative explanations for observed performance differences. Our metrics provide an additional quantitative tool to enrich the evaluation of 3D input devices.

3. 3D ACCURACY MEASURES

Motivated by the aforementioned 2D metrics [9], we propose three measures to help characterize 3D motions. Previous research [2, 5, 6, 15] use a motion tracked stylus or wand for 2D selection. However, none of these report how users moved the device when performing the selection task. Characterizing users' free-space 3D motions can reveal inefficiencies in movement and/or possible sources of arm or wrist fatigue. Once identified, the pointing technique can then be improved.

There are two main benefits to our 3D accuracy measures. First, high-DOF input devices can control 2D cursors; this is common in games (e.g., on Nintendo *Wii*). We use similar pointing modes in our study. While 2D metrics help explain differences in such 2D cursor techniques, they fail to capture some usage behavior. For example, one can point the device from different positions and/or orientations yielding the same cursor position (Figure 1).

Second, high-DOF devices can be used to directly select remote 3D objects via ray casting. In these cases, it is likely infeasible to use the existing 2D accuracy measures [9]. Measures that consider the higher-dimensional nature of 3D pointing are required. The following sections propose the new measures.

3.1 Depth Variability (DV)

Most trackers provide at least 3DOF of movement detection. Device depth direction may not change the 2D cursor position, but may still demonstrate some pointing inefficiency (Figure 2). *Depth variability* is the standard deviation from the average device depth during a trial. This is based on the *movement variability* measure. [9]. *DV* is computed as follows:

$$DV = \sqrt{\frac{1}{n-1}\sum (z_i - \bar{z})^2} \qquad (3)$$

where z_i is the sample distance from $z = 0$ plane, and $\bar{z}$ is the average distance of all samples for the trial from the $z = 0$ plane. For depth-insensitive pointing techniques, *DV* should ideally be 0, as a higher number would represent unnecessary depth motion.

3.2 Rotation/Movement Ratio (RMR)

Our second measure relates to the amount of rotation vs. movement used to control a pointing technique. For example, one can select the same target by pointing at from very different positions (Figure 1), i.e., the device position "trades off" with the device orientation. It is also possible to move the device great distances, while rotating it by the same amount (Figure 3).

Rotational control seems less fatiguing than arm movement. Consequently, we propose to use the ratio between device movement and rotation as another measure of pointing efficiency. We ignore roll, as this usually will not affect selection. For pitch and yaw, we first find the difference between the maximum and minimum rotation angle for a trial. These extrema are then projected onto the display surface to find the distance between them, D_r. Next, D_m, the "movement distance" is computed as the distance between the minimum and maximum device position in the specified axis. This measure is then computed as:

$$D_r = 2 \times dist \times \tan(\Delta\Theta/2), \text{ and } D_m = \Delta x, \text{ then}$$

$$RMR_{axis} = D_r / (D_r + D_m) \qquad (4)$$

Note that *dist* is the average distance from the device to the screen and $\Delta\Theta$ is the difference between the min and max rotation angles. The difference between the min and max position is Δx. This measure indicates how much rotational control contributed to the entire pointer movement in a given trial.

3.3 Rotation Direction Change (RDC)

We measure rotation direction change frequency in each axis. For example, increasing the device pitch would reflect an increase in the cursor y-coordinate. An inefficient pointing technique may yield alternating increases and decreases of device pitch, i.e., rotation direction changes (Figure 4). This metric is computed as the count of such rotation direction changes, greater than a threshold, in each axis, averaged per trial. We use a 1° threshold.

4. METHOD

4.1 Participants

Twelve paid, right-handed participants (7 males, 5 females) were recruited from our university campus. Ages ranged from 20 to 31 years (mean = 23.8; SD = 3.8). Participants were frequent mouse users, but had limited experience using the *Move* controller.

4.2 Apparatus

Participants used a mouse and a Sony *PlayStation Move* to perform pointing tasks on a PC. The *Move* was connected to the PC via a *PlayStation 3* (*PS3*) gaming console using Sony's

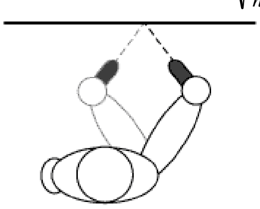

Figure 1. Two device positions and orientations that result in the same cursor position on the screen.

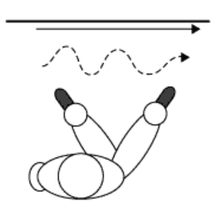

Figure 2. The solid arrow is the intended path, but the dashed arrow is the actual device movement.

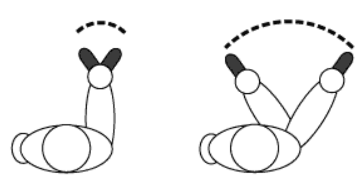

Figure 3. Wrist movement (left) and arm movement (right) traverse different rotational distances but yield the same pointer movement.

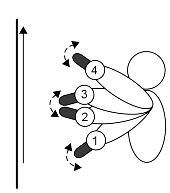

Figure 4. While moving the device to change cursor position (solid), the user often alternates rotation (dashed).

Move.Me server software, which captured buttons events and mapped *Move* position and orientation to cursor movement. *Move* latency was 78 ± 3 ms. Participants stood 2.5 m from a projected display (1.4 m diagonal at 1024 × 768 resolution). A height-adjustable podium provided a surface for the mouse.

4.3 Procedure

The experiment had three conditions: the mouse (baseline), and two using the *Move*. In "position mode", the *Move*'s *x*/*y* motion moved the cursor, while depth was ignored. This used absolute mapping of a small tracked rectangle to the screen. "Laser mode" placed the cursor where the *Move's* selection ray intersected the display. Each condition was preceded by a practice session. During experimental sessions, participants were instructed to select the highlighted target "as quickly and as accurately as possible". Circular targets were arranged in a circle, with six width–distance combinations per block. Each trial concluded upon clicking (whether the target was hit or missed).

4.4 Design

The experiment used a within-subjects design with the following factors and levels:

Technique:	mouse, laser mode, position mode
Target Width:	20, 35, 60 px (22, 39, 67 mm)
Target Distance:	450, 550, 650 px (500, 611, 722 mm)

At 2 blocks of 15 selection trials each, there were 9,720 experiment trials over all 12 participants. The target width and distance combinations represent nine *ID*s (per Equation 1) ranging from 3.09 to 5.07 bits. These were presented in random order without replacement for each block and technique. The technique ordering was counterbalanced using a Latin Square. The experiment took about 30 minutes to complete.

The dependent variables were error rate and throughput. We also report motion both MacKenzie's 2D accuracy measures [9] and the our new proposed 3D measures.

5. RESULTS AND DISCUSSION

5.1 Throughput

Position and laser throughput was 43.3% and 62.6% lower than the mouse, respectively (Figure 5). Technique had a significant main effect on throughput ($F_{2,18} = 293.4$, $p < 0.0001$). A Scheffé test revealed each technique was significantly different ($p < .05$).

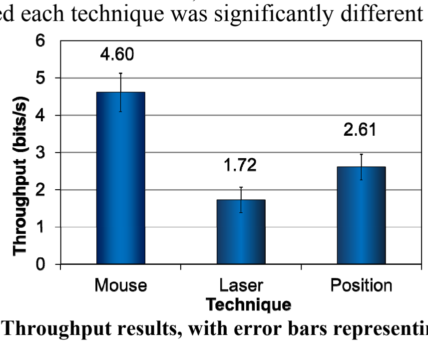

Figure 5. Throughput results, with error bars representing ±1 SD.

Jota et al. [5] report laser throughput of 3.82 bits/s, but did not use effective measures. Our results are thus better compared to Teather and Stuerzlinger's [15] "pen ray" throughput of 1.5 bits/s.

5.2 Error Rate

Error rates are summarized in Figure 6. Technique had a significant main effect on error rate ($F_{2,18} = 96.41$, $p < 0.0001$), and each technique was significantly different ($p < .05$).

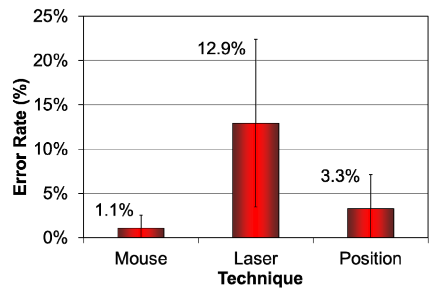

Figure 6. Error rate, with error bars representing ±1 SD.

Jota's laser mode yielded an error rate of only 3.4% [5], but used a 1D task requiring less precision. Teather's pen ray error rate of 13.6% [15] is more consistent with our results.

5.3 2D Movement Fidelity

Here, we analyze each technique according to MacKenzie's 2D accuracy measures [9]. Each of these is summarized in Table 1.

Table 1. The per-trial mean (and SD) for each metric (best result highlighted). Post hoc significance at 5% level, where non-significant groups are shown in brackets, and "<" signs indicate differences. * $p < 0.0001$. ** $p < 0.0005$. *** $p < 0.005$.

Metric	<u>M</u>ouse	<u>L</u>aser	<u>P</u>osition	*F*-value	Post Hoc
TRE	0.06 (0.06)	0.51 (0.17)	0.09 (0.07)	81.68*	(M, P) < L
TAC	2.39 (0.12)	4.93 (0.61)	2.28 (0.31)	162.63*	(P, M) < L
MDC	5.38 (0.64)	12.64 (2.50)	2.91 (0.38)	137.89*	P < M < L
ODC	0.74 (0.49)	9.42 (1.74)	0.82 (0.42)	260.74	(M, P) < L
MV	16.89 (5.22)	17.92 (3.83)	13.69 (2.97)	6.89***	P < L
ME	18.41 (5.27)	15.60 (2.77)	14.15 (2.40)	6.98***	P < M
MO	3.96 (3.01)	−0.18 (2.37)	2.37 (1.83)	11.71**	L < (P, M)

Surprisingly, the *Move* modes yielded the best result in five of the metrics (Table 1). Its low TRE and ODC values signify definitive target selection and consistent cursor movement towards the target. The position mode had the best scores for most metrics. The low TAC, MDC, MV, and ME values imply straight pointer paths parallel to the task axis. The fact that position mode throughput was lower than the mouse may be because the technique required more device movement. The laser condition yielded the lowest MO, indicating pointer paths close to the task axis. The high MDC and ODC for the laser condition quantitatively illustrate the propensity for hand tremors.

5.4 3D Movement Fidelity

Our new measures (Section 3) were computed for the laser and position control modes, averaged per trial, and compared using an independent samples *t*-test assuming unequal variance. The laser was significantly better than position mode in several measures (Table 2). This suggests that while these measures help characterize device motion, they may not relate to performance.

Table 2. The per-trial mean (and SD) for each metric (best result highlighted). *DV* is measured in mm, all other are count/ratios without units. * $p < 0.001$. ** $p < 0.05$.

Metric	Laser	Position	*t*-value
Depth Variability (DV)	9.06 (3.02)	28.18 (9.85)	−13.16*
x Rotation/Movement Ratio (RMR*x*)	0.84 (0.04)	0.82 (0.03)	0.95
y Rotation/Movement Ratio(RMR*y*)	0.82 (0.04)	0.80 (0.03)	1.13
Roll Direction Change (RDC_{roll})	0.29 (0.06)	0.63 (0.12)	−17.03*
Yaw Direction Change (RDC_{yaw})	0.42 (0.06)	0.48 (0.12)	−3.14**
Pitch Direction Change (RDC_{pitch})	0.43 (0.08)	0.49 (0.09)	−3.07**

Depth variability was significantly lower for the laser than position mode. This is likely because the *z* coordinate did not affect position mode, but would change the ray origin in laser mode, affecting the cursor position. Thus, in laser mode, participants scrutinized their depth motion. Depth motion in position mode was inefficient, but not necessarily detrimental. Similarly, it would not impact any of the 2D metrics where position mode performed significantly better, as depth movement would not result in cursor position changes.

Laser mode had significantly fewer rotational direction changes than position mode in all three axes, suggesting that rotational control is more important in laser mode. This makes sense, given that the ray direction controls the cursor position. In position mode, any device rotation would only affect the cursor position insofar as it changed the device position. Thus, participants were more careful with device orientation in laser mode.

There was no significant difference in RMR between modes. Both modes had relatively high ratios, suggesting that both modes were primarily controlled by wrist rotation, rather than sweeping arm motions. This makes sense, as participants would quickly find that rotating the device is easier than moving it large distances. However, it also highlights a propensity for wrist fatigue.

Ultimately, this analysis suggests the influence of rotational degrees of freedom may be stronger than positional degrees. It is well known (see e.g., [15]) that higher-DOF techniques generally perform worse. Our results suggest that the rotation-based laser mode not only performed worse, but yielded more erratic 2D cursor trails as well. Conversely, the position mode was not affected by device rotation (as reflected by our 3D measures), yet produced more efficient cursor trails, and higher performance.

6. CONCLUSION AND FUTURE WORK

We introduced three novel metrics to characterize 3D pointing and used them to characterize the PlayStation *Move* in a standardized selection task. Although there was no correlation between device movement and pointing performance, the metrics revealed and quantified how the *Move* was used during pointing. For example, the high Depth Variability of the position mode (versus laser mode) could indicate an area for improvement if movement efficiency were paramount. Alternatively, the higher DV and throughput values for position mode could illustrate a robust technique that performs well, despite user inefficiency.

The laser mode exhibited high Movement Direction Change and Orthogonal Direction Change, but also significantly better Rotational Direction Change on all three axes. Thus, one could quantitatively support using linear movement for coarse pointer control and rotational movement for fine pointer control.

For both modes, the Rotation/Movement Ratios show not all degrees of freedom are equally used – rotational motion is primarily used. This preference could forecast localized fatigue or strain after extended use. By combining our 3D metrics with existing 2D and performance metrics, designers of 3D pointing techniques can characterize and quantify device usage to show strengths and identify weaknesses in their techniques.

7. REFERENCES

[1] Fitts, P.M. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47* (6). 381-391.

[2] Grossman, T. and Balakrishnan, R. 2006. The design and evaluation of selection techniques for 3D volumetric displays. In *Proc.of ACM UIST 2006*, 3-12.

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

[4] Jagacinski, R.J. and Flach, J.M. 2003. Chapter 4 - Information theory and Fitts' law. In *Control theory for humans: Quantitative approaches to modeling performance*.

[5] Jota, R., Nacenta, M.A., Jorge, J.A., Carpendale, S. and Greenberg, S. 2010. A comparison of ray pointing techniques for very large displays. In *Proc. of Graphics Interface 2010*, 269-276.

[6] Kopper, R., Bowman, D.A., Silva, M.G. and McMahan, R.P. 2010. A human motor behavior model for distal pointing tasks. *Int. J. of Human-Computer Studies*, *68* (10). 603-615.

[7] LaViola, J.J. 2008. Bringing VR and Spatial 3D Interaction to the Masses through Video Games. *IEEE Computer Graphics and Applications*, *28* (5). 10-15.

[8] MacKenzie, I.S. 1992. Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, *7*. 91-139.

[9] MacKenzie, I.S., Kauppinen, T. and Silfverberg, M. 2001. Accuracy measures for evaluating computer pointing devices. In *Proc. of ACM CHI 2001*, 9-16.

[10] McArthur, V., Castellucci, S.J. and MacKenzie, I.S. 2009. An empirical comparison of "Wiimote" gun attachments for pointing tasks. In *Proc. EICS 2009*, 203-208.

[11] Myers, B.A., Bhatnagar, R., Nichols, J., Peck, C.H., Kong, D., Miller, R. and Long, A.C. 2002. Interacting at a distance: Measuring the performance of laser pointers and other devices. In *Proc. of ACM CHI 2002*, 33-40.

[12] Natapov, D., Castellucci, S.J. and MacKenzie, I.S. 2009. ISO 9241-9 evaluation of video game controllers. In *Proc. of Graphics Interface 2009*, 223-230.

[13] Oh, J.-Y. and Stuerzlinger, W. 2002. Laser pointers as collaborative pointing devices. In *Proc. of Graphics Interface 2002*, 315-320.

[14] Soukoreff, R.W. and MacKenzie, I.S. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *Int. J. of Human-Computer Studies*, *61* (6). 751-789.

[15] Teather, R.J. and Stuerzlinger, W. 2011. Pointing at 3D targets in a stereo head-tracked virtual environment. In *Proc. of IEEE 3DUI 2011*. IEEE Press, 87-94.

[16] Wingrave, C.A. 2010. The Wiimote and beyond: spatially convenient devices for 3D user interfaces. *IEEE Computer Graphics and Applications*, *30*. 71-85.

[17] Wingrave, C.A., Tintner, R., Walker, B.N., Bowman, D.A. and Hodges, L.F. 2005. Exploring Individual Differences in Raybased Selection; Strategies and Traits. In *Proc.of IEEE VR 2005*, 163-170.

[18] Wobbrock, J.O., Shinohara, K. and Jansen, A. 2011. The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models. In *Proc. ACM CHI 2011*, 1639-1648.