Smartwatch-based Pointing Interaction

Javid Abbasov[1], Tom Horak[2], Raimund Dachselt[2]

T-Systems Multimedia Solution[1]
Interactive Media Lab, Technische Universität Dresden[2]

javid.abbasov@t-systems.com, tom.horak@tu-dresden.de,
raimund.dachselt@tu-dresden.de

Abstract

In this work, we present the design and evaluation of a smartwatch-based mid-air pointing and clicking interaction technique called *Twist, Point, and Tap*, or short *TPT*. Incorporating only commodity devices, we aim to provide a fast and error-prone pointing approach that can easily be deployed to existing environments with a shared display, e.g., meeting rooms or public info points. Detected by internal sensors, TPT maps horizontal forearm movements as well as wrist rotation to relative cursor movements on a nearby large display. Left and right-click interactions are supported through tapping on the smartwatch's touchscreen. By running a Fitts's law study, we compared our TPT concept against an existing smartwatch-based pointing technique called Watchpoint (Katsuragawa et al., 2016). The study revealed that the TPT concept has a smaller error rate while maintaining a comparable performance.

1 Introduction & Background

Large high-resolution displays can increasingly be found in everyday environments, such as meeting rooms, shopping centres, study halls, or even homes. These displays can effectively provide both a higher density of information and improved visibility from a distance (Andrews et al., 2011), which, however, lead to the need of remotely accessing content, e.g., by pointing interaction. Existing work on mid-air pointing is extensive, but often involve specialized hardware (e.g., Wiimote used by Campbell et al., 2008; Myo strap by Haque et al., 2015) or instrumentation of the environment (e.g., camera-detected laser pointer by Olsen and Nielsen, 2001; Vicon system by Vogel and Balakrishnan, 2005), resulting in a high deployment effort. As an alternative, personal mobile devices (e.g., smartwatches, smartphones, tablets) can be incorporated. These devices are especially promising for pointing, as they are (a) cost-effective, (b) multi-purpose devices and thus more likely to be available, and (c) more and more equipped with high-precision built-in sensors (Ballagas et al., 2006; Pietroszek et al., 2014; Siddhpuria et al., 2018). As Smartwatches also come with the advantage of being always accessible at

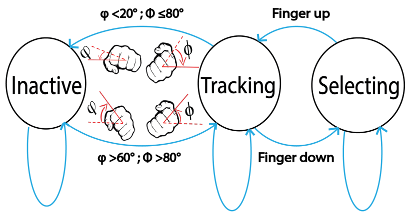

Figure 1: This figure summarizes empirically determined thresholds used to support the TPT concept transitions from Inactive to Tracking to Selecting states.

the wrist of the user, they were already incorporated for pointing interaction in multiple research projects (e.g., Horak et al., 2018; Katsuragawa et al., 2016). As one of the most recent works presenting a technical implementation using only smartwatch's built-in sensors, Watchpoint (Katsuragawa et al., 2016) allows to control cursor movements on a nearby display through left-right-up-down forearm movements, while click interaction is performed via wrist rotation gestures. However, the technique still has to face issues regarding error-rate and speed, especially when performing clicks. Also, in certain situations, the proposed movements are not feasible, e.g., while sitting.

We contribute the *Twist, Point, and Tap (TPT)* concept that alters the incorporate movements for cursor manipulation to horizontal movements and wrist rotation as well as tap interaction to trigger left and right clicks. Further, we contribute an empirical evaluation showing that TPT is less error-prone than Watchpoint while maintaining similar performance.

2 The Twist, Point, and Tap Concept

As the basic idea, we map horizontal arm movements to horizontal cursor movements and the angle of wrist rotation to vertical cursor movements (*Twist and Point*). Limiting the arm movements to one dimension comes with multiple advantages. Firstly, it is easier to perform horizontal movements instead of vertical ones while sitting at a table, as it is often the case in meeting rooms. Secondly, when only focusing on one movement direction, it can get easier for a user to perform steady arm movement, thus reducing jitter while pointing. Similar, using a tap on the smartwatch to trigger clicks also reduces accidental movements compared to triggering clicks by wrist rotation, as proposed in the Watchpoint concept (Katsuragawa et al., 2016).

In order to avoid unintentional pointing interaction, we apply a 3-state model distinguishing between *Inactive*, *Tracking*, and *Selecting* states (Figure 1). Initially, the system is in the *Inactive* state when the face of the watch is slightly down or up and can be put into the *Tracking* state by turning the watch face towards an orthogonal position. To prevent unintentional switches between the *Inactive* and *Tracking* state, we utilize asymmetric wrist angles: For activation, the wrist must be rotated outwards (watch face down) by 60°, and then back to less than 20° to return to the *Inactive* state. However, a symmetric angle is utilized when the face of the watch is down (80°). Both the thresholds and the decision on asymmetric/symmetric transition are based on an empirical test and represent the movement patterns that are most convenient.

In the *Tracking* mode, a cursor appears on the nearby display allowing the user to control the cursor position as described above. From the *Tracking* state, the user can perform a left-click event by tapping on the left side of the watch's screen, and a right-click by tapping on the right. Similar to a regular mouse, drag events are supported by keeping the finger down.

We implemented the TPT concept as an Android Wear application. For horizontal arm movements, we utilized the azimuth (yaw) value of the game rotation vector as rotation angle ϕ_i (Equation 1), while the z-value of the gravity sensor was used for detecting wrist rotations (Equation 2). The detected relative movements were sent via WiFi to the system running on the large display, where a Java application then manipulates the cursor position of the host system. Furthermore, to reduce the impact of common pointing challenges, such as jitter or Heisenberg effect (Bowman et al., 2001), we applied a low-pass filter (Casiez et al., 2012) as well as a bookkeeping method (Bowman et al., 2001).

$$x_i = w - \frac{3w\phi_i}{\pi} \tag{1}$$

$$y_i = h - \frac{(9\beta_i - 4\pi)h}{4\pi}, \text{ where: } \quad \beta_i = \cos^{-1}\frac{Gravity_z}{9.8} \tag{2}$$

3 Evaluation: Fitts's Law Pointing Task

We implemented our TPT concept and compared it against a Watchpoint implementation, hypothesizing that our concept will perform equally.

Participants and Setup. We recruited 10 participants from the T-Systems Multimedia Solutions company (7 male, 3 female; age 20–29). The study took place in a meeting room featuring a 75-inch display. While executing the experimental tasks, the participants stood 2m from the display. We used the implementation as described in the section above.

Procedure. We ran a multi-directional Fitts's Law pointing task (Fitts, 1954), with independent variables for target width and distance, and dependent variables for selection time and error rate. We included three-level target widths (16, 48, and 144mm) as well as three target distances

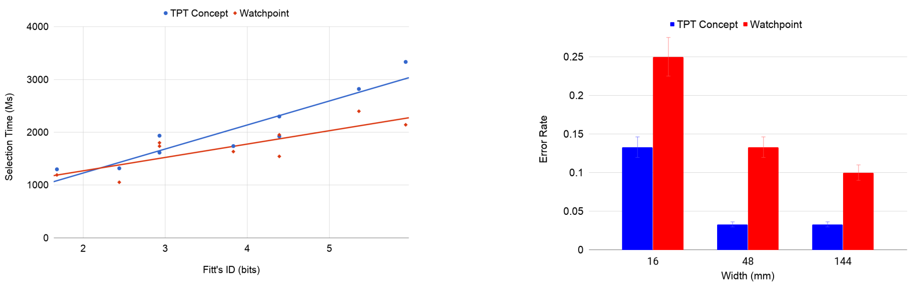

Figure 2: Selection time by Fitt's ID (left) and error rate by target width (right)

(320, 640, and 960mm). For each width-distance combination, participants performed two trials per condition (TPT and Watchpoint). In each trial, we logged the error-rate as well as the time of completion. At the end of the study, a total of 360 trials were collected.

Results. We analyzed the measured times and errors using an RM-ANOVA test ($\alpha = .05$). For the *selection time*, it was observed that the TPT concept was slightly slower than Watchpoint ($F(1,8) = 6.241$, $p = .037174$). Notably, this difference occurs for medium and small targets ($F(1,9) = 6.535$, $p = .030867$), while the required time for large targets (144 mm) is roughly the same (Figure 2). The equations and R^2 values were as follow: $R^2 = .86$, $MT = 318 + 455 \times ID$ for the TPT concept and $R^2 = .68$, $MT = 763 + 253 \times ID$ for Watchpoint.

The overall *error rate* of the TPT concept was 6.3%, while it was 13.8% in Watchpoint. The TPT concept mostly outperformed Watchpoint in terms of large targets (Figure 2), 48 mm target ($F(1,9) = 3.857$, $p = .081126$) and 144 mm target ($F(1,9) = 3.273$, $p = .103888$). For small targets (16 mm), the post-hoc test found a slight difference between the concepts ($F(1,9) = 1.995$, $p = .191403$).

4 Discussion & Conclusion

The results of our evaluation show that our TPT concept performs on a par or better with existing implementations like Watchpoint. The RM-ANOVA revealed that user required more time for small and medium targets with our concept, which we believe is mostly due to the limited training phase. The participants were sometimes confused with controlling cursor's vertical movement through wrist rotations, resulting in occasionally twisting their wrist too fast and producing overshooting. However, with users getting more used to the interaction style, we expect a notable improvement in completion time.

Regarding error rate, the TPT concept outperformed Watchpoint. These differences were particularly evident for medium-sized targets, where the error rate for our concept was nearly three times less than for the Watchpoint technique. We expect this difference to be caused by the way we implemented the selection: as the tapping was performed with the second hand, the arm used for pointing could stay steady. In contrast, Watchpoint requires a wrist movement for performing a click, which also can cause unintentional arm movements along the x- and y-axis. However, requiring a second hand as in TPT can be felt as additional effort, as commented by some participants during the evaluation.

All in all, our work provide multiple insights that can inform the design of future interaction concepts in everyday environments. Firstly, alongside existing work, our TPT concept shows that pointing interaction can be efficiently realized with commodity devices and, thus, also integrate the Bring-Your-Own-Technology (BYOT) paradigm. Secondly, in the context of large displays, users are enabled to interact with the large displays regardless of the distance, their posture, or possibly existing obstacles between them and the display. Finally, we were able to show that also indirect mappings, like wrist rotation to vertical cursor movements, are feasible and perform on par with more common mappings, while bringing additional benefits that allow us to provide a more error-prone interaction concept.

References

Andrews, C., Endert, A., Yost, B., & North, C. (2011). Information visualization on large, high-resolution displays: Issues, challenges, and opportunities. *Inform. Visual. 10*(4), 341–355.

Ballagas, R., Borchers, J. O., Rohs, M., & Sheridan, J. G. (2006). The smart phone: A ubiquitous input device. *IEEE Pervasive Comput. 5*(1), 70–77.

Bowman, D. A., Wingrave, C. A., Campbell, J. M., & Ly, V. Q. (2001). Using pinch gloves™for both natural and abstract interaction techniques in virtual environments. *Proceedings of HCI internationall*, 205–211.

Campbell, B. A., O'Brien, K. R., Byrne, M. D., & Bachman, B. J. (2008). Fitts' law predictions with an alternative pointing device (wiimote). *Proc. HFES '08*, *52*(19), 1321–1325.

Casiez, G., Roussel, N., & Vogel, D. (2012). 1 € filter: A simple speed-based low-pass filter for noisy input in interactive systems. In *CHI conference on human factors in computing systems, CHI '12, austin, tx, USA - may 05 - 10, 2012* (pp. 2527–2530).

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. In *The journal of experimental psychology* (pp. 262–269).

Haque, F., Nancel, M., & Vogel, D. (2015). Myopoint: Pointing and clicking using forearm mounted electromyography and inertial motion sensors. In *Proc. CHI '15* (pp. 3653–3656). ACM.

Horak, T., Badam, S. K., Elmqvist, N., & Dachselt, R. (2018). When david meets goliath: Combining smartwatches with a large vertical display for visual data exploration. In *Pro. CHI '18* (19:1–19:13). ACM.

Katsuragawa, K., Pietroszek, K., Wallace, J. R., & Lank, E. (2016). Watchpoint: Freehand pointing with a smartwatch in a ubiquitous display environment. In *Proc. AVI '16* (pp. 128–135). ACM.

Olsen, D. R. & Nielsen, S. T. (2001). Laser pointer interaction. In *Proc. CHI '01* (pp. 17–22).

Pietroszek, K., Kuzminykh, A., Wallace, J. R., & Lank, E. (2014). Smartcasting: A discount 3d interaction technique for public displays. In *Proc. OZCHI '14* (pp. 119–128). ACM.

Siddhpuria, S., Malacria, S., Nancel, M., & Lank, E. (2018). Pointing at a distance with everyday smart devices. In *Proc. CHI '18* (173:1–173:11). ACM.

Vogel, D. & Balakrishnan, R. (2005). Distant freehand pointing and clicking on very large, high resolution displays. In *Proc. UIST '05* (pp. 33–42). ACM.