TrackLine: Refining Touch-to-Track Interaction for Camera Motion Control on Mobile Devices

Axel Hoesl, Sarah Aragon Bartsch, Philipp Burgdorf, Andreas Butz
{axel.hoesl,andreas.butz}@ifi.lmu.de,{aragonbartsch,burgdorf}@cip.ifi.lmu.de
LMU Munich
Munich, Germany

New motion control technologies such as drones, gimbals or industrial robots provide smooth repeatable camera motion and image stabilization. These tools are usually motor driven and therefore operated remotely. The user interfaces for this remote operation are often implemented on mobile touch-screen devices. With touch-screens however, a decrease in precision can be generally observed [2]. One way to address the issue by system-design is providing assistance on a higher control level. In cinematography this is implemented as assistance through computer vision and visual servoing and allows image-based motion control directly on the video-stream[1]. For such image-based control, systems often use a touch-to-track (TTT) approach for the selection of an object or person to be tracked and followed. With TTT, users tap on the object or person they see in the video-stream and the system then continuously adjusts its position to keep the selected object in the same position within the frame. As TTT is based on manual selection of a moving object, performance in such goal-directed aiming can be expected to decrease for faster moving targets [1] as in car commercials, sports-broadcasting or high-speed recordings. We developed an alternative concept addressing the issue that lets operators define the desired tracking position in advance and delegates the correct timing to an assisting system. TrackLine allows operators to define a motion trigger within the image frame that is represented by a line and displayed on top of the video-stream (Figure 1).

To determine whether TrackLine helps to address the mentioned human performance issues and to estimate its efficiency, easy of use and comfort in use, we conducted a user study. For data collection early in the design process, we used a virtual environment for prototyping. The concept was implemented in Unity 5.3.1f1 running on an off-the-shelf Android tablet. Besides TTT, systems used in-the-wild often provide a software-joystick for operation. To be able to compare our approach to established interaction techniques, we also implemented software-joystick and TTT.

For the study a within-subjects design with the independent variables user interface (3 levels) and task (3 levels) was chosen. The tasks to be carried out were *direction change* (object changes movement direction), *fast object* (object moves at high velocity) and *track&pan* (transition from tracking shot to panning shot). To avoid learning effects both variables were counter-balanced by using a Latin-Square design.

Our measurements consisted of the *number of trials*, the *quality of control* and *user feedback*. We recruited 12 participants (9 male, 3 female). The average age was 24, with ages ranging from 21 to 27. All were familiar with touch-screen devices. Three reported to be acquainted with camera operation.

In the following data analysis we used non-parametric tests (Friedman's ANOVA, Wilcoxon Signed-Rank) to test for statistical significance. Our alpha level was 0.05. Bonferroni correction was applied in post-hoc tests to make up for pairwise comparisons. All presented pairwise comparisons were conducted only after a significant main effect was found.

To determine *number of trials*, the participants were asked to carry out each task until a further performance increase seemed unlikely for them. The *fast object* and *track&pan* tasks showed noteworthy effects. For *fast object*, TTT needed the most trials, whereas for *track&pan* the software-joystick. Despite this effect due to the different tasks, TrackLine outperformed both interfaces in both tasks. With software joystick, participants did 5.08 trials in average (SD=2.07), while it were 7.08 (SD=3.45) with TTT and only 3.17 with TrackLine (SD=1.95). Post-hoc pairwise comparisons showed that using TrackLine resulted in significantly fewer trials than software-joystick ($p<0.01$) and TTT ($p<0.01$). For *track&pan*, the participants did 5.50 tries on average with software-joystick (SD=2.15), with TTT 3.17 (SD=1.34) and only 2.17 (SD=1.27) with TrackLine. Comparing the conditions pairwise, we found TrackLine resulted in fewer trials than software-joystick ($p<0.01$) and TTT ($p<0.01$).

For estimating the *quality of control*, we adapted the idea of measuring the standard deviation of lateral position [3] to the Rule of Thirds. In

[1]Example system: `http://www.vertical.ai/studio`

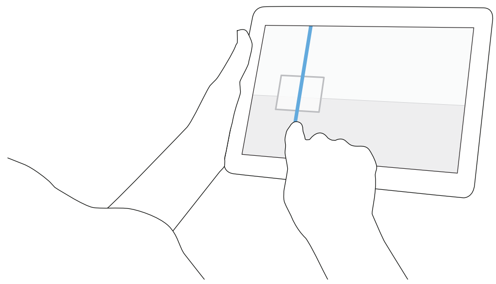

Figure 1: TrackLine (blue) serves as trigger and determines the framing.

detail, a moving object should be followed and framed at the first horizontal third in movement direction. The distance of the objects screen position to this third was continuously logged (Figure 2). For *fast object*, the software-joystick resulted in a 55.12 px distance on average (SD=27.91), TTT in 63.91 px (SD=95.11), whereas TrackLine only in 15.52 px (SD=5.42). In pairwise comparison we found participants to be closer the ideal position with TrackLine than with software-joystick ($p<0.01$) and TTT ($p<0.01$). For the *track&pan* task, the average distances were 28.98 px with software-joystick (SD=12.35), 29.41 px with TTT (SD=62.91) and only 1.78 px with TrackLine (SD=0.44). Pairwise comparison indicated that TrackLine led to a smaller distance to the ideal position than software-joystick ($p<0.01$) and TTT ($p<0.01$). Additionally, it helped avoiding misses that occurred especially with TTT resulting in a large variance in the measured data (see error bars in Figure 2).

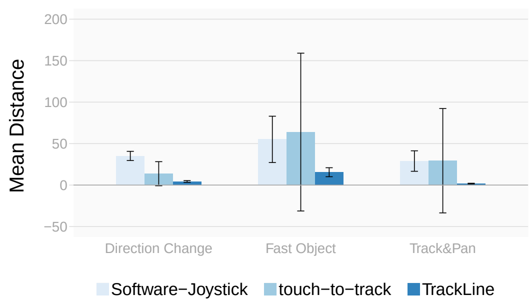

Figure 2: Mean distance from ideal position for each task and interface.

Self-reported data on efficiency, ease of use and comfort was collected via 5-item rating scales. The software-joystick was rated low in efficiency (Mdn=2) and ease of use (Mdn=2), but as comfortable to hold (Mdn=4.5). TTT was rated to be efficient (Mdn=4) and easy to use (Mdn=4), but lowest in comfort (Mdn=3.5). TrackLine was perceived as very efficient (Mdn=5), easy to use (Mdn=5) and comfortable to hold (Mdn=5).

Our results suggest that with TrackLine operators need fewer trials while being more precise compared to established techniques. It is perceived as efficient and easy to use. Participants also pointed out that extension in future work should include fostering exploration and expressiveness of the technique.

[1] Tien-Ting Chiu, Kuu-Young Young, Shang-Hwa Hsu, Chun-Ling Lin, Chin-Teng Lin, Bing-Shiang Yang, and Zong-Ren Huang. A Study of Fitts' Law on Goal-Directed Aiming Task with Moving Targets. *Perceptual and Motor Skills*, 113(1):339–352, 2011.

[2] Andrew Sears and Ben Shneiderman. High Precision Touchscreens: Design Strategies and Comparisons with a Mouse. *International Journal of Man-Machine Studies*, 34(4):593–613, April 1991.

[3] Joris C Verster and Thomas Roth. Standard Operation Procedures for Conducting the on-the-Road Driving Test, and Measurement of the Standard Deviation of Lateral Position (SDLP). *International Journal of General Medicine*, 4:359, 2011.