



Evaluating Tap-and-Drag: A Single-Handed Zooming Method

Manoel Farhad^(✉) and I. Scott MacKenzie

Department of Electrical Engineering and Computer Science, York University,
Toronto, ON, Canada
manoelfd@my.yorku.ca, mack@cse.yorku.ca

Abstract. We conducted a user study comparing the accuracy and speed of two zooming methods for touch-screen devices: tap-and-drag (a single-handed zooming method) and the traditional pinch-to-zoom (performed with one hand). The study involved 12 participants and employed a Google *Pixel 2* mobile phone. The results for task completion time favored tap-and-drag, which was about 18% faster than pinch-to-zoom. The accuracy results for tap-and-drag were slightly lower with an average accuracy of 84.8% compared to 87.9% for pinch-to-zoom. This was attributed to users being unfamiliar with tap-and-drag. As users became more familiar and comfortable with tap-and-drag, accuracy improved. Tap-and-drag was about 47% more efficient than pinch-to-zoom, with efficiency measured as the number of gestures to complete a trial. Participants indicated a preference for tap-and-drag for one-handed zoom gestures.

Keywords: Zoom methods · Gesture input · Mobile interaction
Touch-screen input

1 Background

Most of today's mobile devices offer multi-touch capabilities that allow user interactions with more than one finger simultaneously. For example, a user can zoom in and out of content by "pinching" two fingers on the display. However, two-finger gestures generally require two hands: One hand holds the device and the other performs the gesture. This is problematic when the user is on the move, as the free hand (the one not holding the device) might be occupied due to demands in the real world (e.g., carrying items).

Pinch-to-zoom is the most common zoom method employed on mobile devices. The most common way to perform pinch-to-zoom is shown in Fig. 1a. The user holds the device in one hand while performing the pinch-to-zoom gesture with the other hand. There are two variations on this. In Fig. 1b, the user holds the device in two hands and uses a thumb on each hand for the pinch gesture. Alternatively, in Fig. 1c only one hand is used to both hold the device and perform the zoom gesture. However, this latter technique causes an awkward and uneasy hand posture. Users are often forced to perform the pinch-to-zoom gesture via one hand when their other hand is occupied, as this may be the only way of zooming.

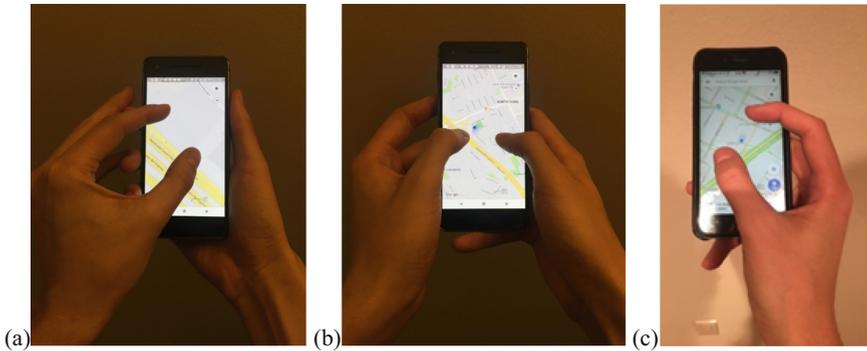


Fig. 1. Three ways to perform the pinch-to-zoom gesture. See text for discussion.

Most mobile devices also employ a double-tap gesture which allows for zooming in or out by double-tapping an area of interest on the display. While this method works well with one hand, it only allows for zooming by a discrete amount (e.g., $\times 3$) as opposed to a continuous amount. Thus, users must still resort to the pinch-to-zoom gesture even when there is an alternative that works with one hand.

Karlson et al. [2] conducted a survey on user habits and preferences on hand usage in mobile device interaction. Of 228 users surveyed, 45% use one hand for nearly all device interactions, as opposed to only 19% who use both hands. The majority of two-handed use occurs when it is the only way to perform the task, given the interface. However, on preference, 66% of participants stated they prefer to use one hand for the majority of interactions, as opposed to only 9% who prefer to use both hands. Twenty-three percent did not express a preference.

Tap-and-drag is an alternative zoom method that allows for one-handed zooming. The user can zoom in and out of content by double-tapping a thumb (or any finger) on the display and then dragging the finger up to zoom out or down to zoom in. Tap-and-drag has not been widely adopted; however, a few apps, such as Google Maps, support the gesture.

In this paper, the user performance of tap-and-drag is evaluated. The results are compared to the conventional pinch-to-zoom method when performed using one hand.

2 Related Work

Ti and Tjondronegoro [7] evaluated a collection of tilt-based input methods for easy single-handed zooming. Their methods used rate-of-rotation readings from an accelerometer – a sensor commonly found in contemporary mobile devices. They compared the results against conventional touch-based zooming and found that tilt-based gestures perform significantly better than touch-based gestures when performed with one hand. All participants found the conventional touch-based methods inferior due to the awkward and uncomfortable hand posture when performed with one hand. However, they still preferred using pinch-to-zoom when two hands were available.

Miyaki and Rekimoto [5] evaluated GraspZoom, a pressure-based single-handed method for zooming. GraspZoom uses the fingers positioned behind the mobile device. The back of the device is equipped with a pressure sensor. Users press the sensor to temporarily switch from panning to zooming. The requirement of additional hardware is a practical limitation.

Like the method just described, Silfverberg et al. [6] present a system that uses additional hardware while engaging the user's fingers on the back of the device. The method positions two touch sensors underneath a mobile device. Content is viewed on the front of the device while fingers below the device pan and zoom the content via touch gestures. See Fig. 2.

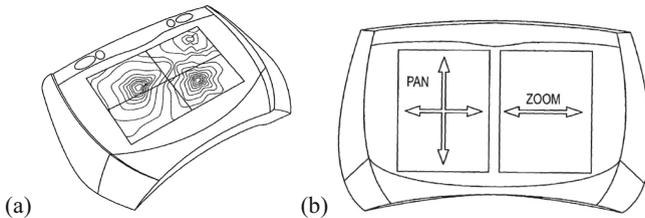


Fig. 2. Method for panning and zooming on a mobile device [6]. (a) top view with display. (b) bottom view with touch sensors.

Boring et al. [1] proposed Fat Thumb, a single-handed method that uses the thumb's contact size as a form of simulated pressure. The contact size allows for switching between panning and zooming depending on the contact area. The participants generally had positive impressions of the method.

Malacria et al. [4] evaluated CycloStar, a pan-zoom control for touch-sensitive surfaces. CycloStar encompassed two implementations, CycloPan and CycloZoom+. The general idea is to use fingers to perform oscillating movements. The user makes closed-loops around a point of interest to zoom in or out. The results were promising due to the multiple geometrical and kinematic parameters used as controls.

Lai et al. [3] evaluated a single-handed partial zooming technique. ContextZoom allows users to specify any place on a display as the zooming center (i.e., point of interest) by long-pressing the position on the display. Once the location is set, the user moves their thumb on the display to zoom in or out. Panning is disabled while zooming. The results were good, with the completion time and number of discrete actions generally low. Participants also reported higher levels of perceived effectiveness and overall satisfaction.

3 Tap-and-Drag

Tap-and-drag attempts to address the limitations of existing methods employed in touch-screen devices (as described previously).

Tap-and-drag is performed by double tapping anywhere on the display. See Fig. 3. On the second tap, the user keeps their finger on the display, thus advancing to drag

mode. Then, the user slides their finger up to zoom out (Fig. 3a) or down to zoom in (Fig. 3b). Panning is disabled during the tap-and-drag gesture. Once the user lifts their finger from the display, tap-and-drag is disabled and panning is re-enabled.

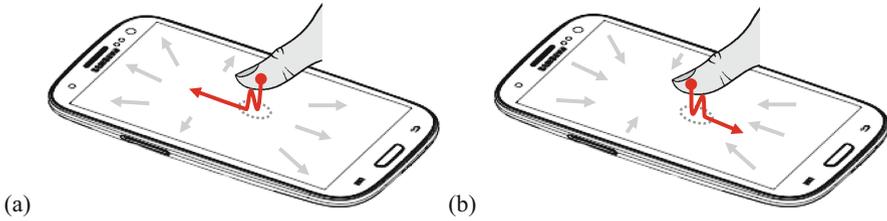


Fig. 3. Tap-and-drag zoom gesture. (a) zoom out (b) zoom in.

The most notable difference between tap-and-drag and the double-tap zoom method described previously is that tap-and-drag allows for zooming by a continuous amount. Another difference is that the focus point of zooming is always the center of display, regardless of where the finger is placed.

The next section describes a user study to compare pinch-to-zoom and tap-and-drag. Each method was tested for its speed, accuracy, and efficiency. Pinch-to-zoom is used as a reference point as it is the most commonly used zooming method on touch-screen devices. To make the comparison fair, both methods are performed using one hand only.

4 Method

4.1 Participants

Twelve participants were recruited from the local university campus. Six were male, six were female. Ages ranged from 21 to 28 years. All were daily users of smartphones with touch-screen displays. Participants were compensated \$10 for their assistance.

4.2 Apparatus

The experiment used a Google *Pixel 2* running the Android (8.1.0) operating system. The device had a 5.0-in. display with a resolution of 1920×1080 pixels and a density of 441 ppi. See Fig. 4.

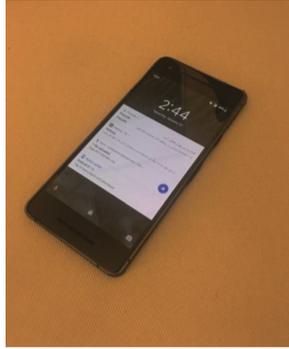


Fig. 4. Google *Pixel 2* employed in the user study.

The software was developed in Java using the Android SDK. The experiment application was developed specifically for this research. Two modes of operation were implemented, one for each zooming method.

The application began with a setup activity which prompted for the participant's initials and experiment parameters, such as group code (for counterbalancing) and zooming method. Once completed, the application transitioned to a map activity to perform the experiment.

An example trial for tap-and-drag zooming is shown in Fig. 5. A trial began upon pressing a *START* button (Fig. 5a). An audible tone indicated that a trial and data collection had begun. For each trial, a red target was placed on a map (Fig. 5b). Six target locations and zoom levels were generated for each of the six trials per block. For each trial, the target's location and zoom level were chosen from this set using a random-without-replacement process. The target was always placed in a location on the map that was within the field of view at the start of a trial.

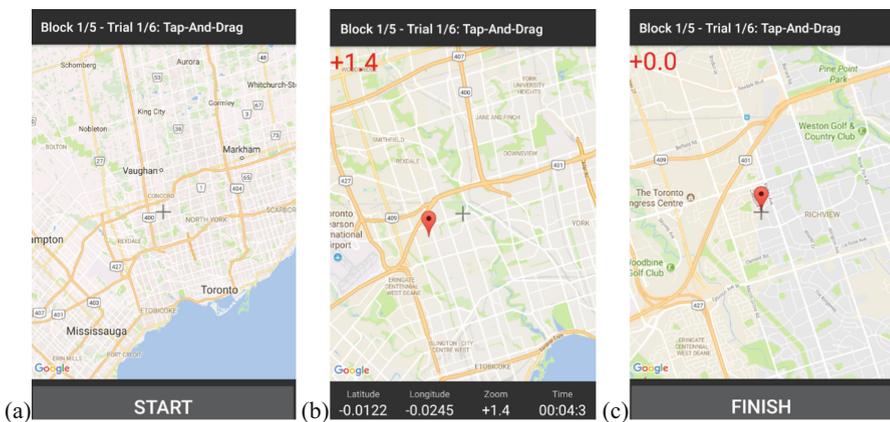


Fig. 5. Example trial. Participants zoom and pan the map so that red target on the map is aligned with the plus symbol marker located in the center of the display. (Color figure online)

The participant was required to place the red target on the marker and zoom to the required amount until the differences were less than a threshold (± 20 pixels for position and $\pm 0.1\%$ for zoom). The marker was always located at the center of the display. When a trial was complete, an audible tone was produced to indicate the completion of the trial and that data collection had terminated (Fig. 5c). The participant then pressed a FINISH button to advance to the next trial. A status area was shown at the bottom of the display which indicated the differences in the present location and zoom of the target with respect to the marker. The difference in zoom was also shown at the top-left of the map, as seen in Fig. 5b. A positive difference indicated that the participant must zoom in, and a negative difference indicated that the participant must zoom out. The method of panning remained the same regardless of the zooming method. By swiping their finger on the display, the participant could pan the map in the direction of the swipe.

The time to complete each trial and the number of discrete gestures performed were recorded. The differences in the position and zoom amount of the target with respect to the marker were logged every 50 ms.

4.3 Procedure

Participants were informed of the purpose and procedures of the experiment, and a short demonstration of each zooming method was given. No practice trials were used. Participants were seated and instructed to hold the device with whichever hand they felt most comfortable with. They were instructed to strictly interact with the device using one hand and to not use their other hand. They were not allowed to rest their arms or hands on a surface. Figure 6 shows a participant performing an experiment trial using each zoom method.



Fig. 6. A participant doing the experiment task (a) using pinch-to-zoom (b) using tap-and-drag.

With this introduction, testing began. Participants completed 5 blocks of 6 trials for each zooming method. Each participant took around 15 min to complete the experiment. Their performance was measured with data logged in the device's internal memory. The data were later transferred to a computer for analysis.

After all blocks and testing were complete, a questionnaire was given in which the participant described his/her impression of the zoom methods and which one they preferred.

4.4 Design

The user study employed a 2×5 within-subjects design. The independent variables and levels were as follows:

- Zoom method (pinch-to-zoom, tap-and-drag)
- Block (1, 2, 3, 4, 5)

Each participant completed five blocks of six trials for each zoom method. As such, the total number of trials for the experiment was $12 \text{ Participants} \times 2 \text{ Zoom Methods} \times 5 \text{ Blocks} \times 6 \text{ Trials} = 720$.

To offset learning effects, participants were divided into two groups to counterbalance the order of testing. One group used tap-and-drag first, then pinch-to-zoom. The order was reversed for the other group.

The dependent variables were the completion time per trial, accuracy, and efficiency. Full descriptions of the accuracy and efficiency dependent variables are provided below.

5 Results and Discussion

All of the trials completed successfully. The data from the experiment were imported into a Microsoft *Excel* spreadsheet where summaries of various measures were calculated and charts were created. The analysis of variance test was performed using the GoStats application.¹

As participants were divided into two groups for counterbalancing, we performed an analysis of variance to test for a possible group effect on each dependent variable. The group effect was not statistically significant ($p > .05$) for completion time and accuracy. Thus, counterbalancing had the desired effect of offsetting order effects. However, the group effect was marginally significant for accuracy ($p = .0435$). The accuracy results were slightly better for group 1 which performed tap-and-drag first, then pinch-to-zoom. Perhaps the participants who were first exposed to tap-and-drag were extra cautious, as they were simultaneously learning the experiment task and using a zoom method with which they had little experience. However, the differences were minor.

5.1 Completion Time

The grand mean for completion time per trial was 13.8 s. The mean time per trial for tap-and-drag zooming was 12.4 s, whereas the mean time for pinch-to-zoom was 15.1 s.

¹ Available as a free download at <http://www.yorku.ca/mack/HCIbook/>.

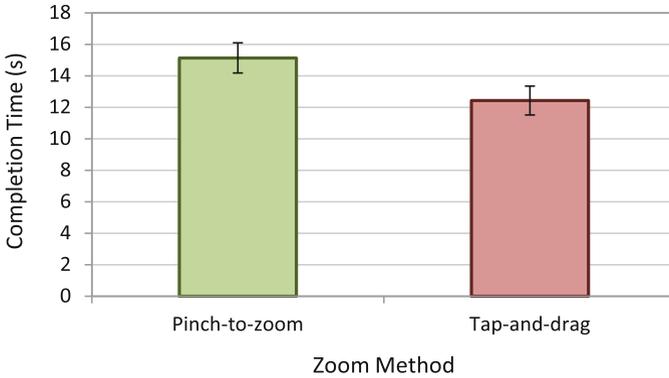


Fig. 7. Completion time (s) by zoom method. Error bars show ± 1 SE.

As such, trials using tap-and-drag were 17.9% faster than the pinch-to-zoom trials. See Fig. 7. The difference was statistically significant ($F_{1,10} = 11.78, p < .01$).

It was observed that the completion time for both zooming methods decreased during the later blocks, with tap-and-drag decreasing a greater amount than pinch-to-zoom. See Fig. 8. The effect of block on completion time was statistically significant ($F_{4,40} = 25.43, p < .0001$). However, the Zoom Method \times Block interaction effect was not statistically significant ($F_{4,40} = 1.35, p > .05$).

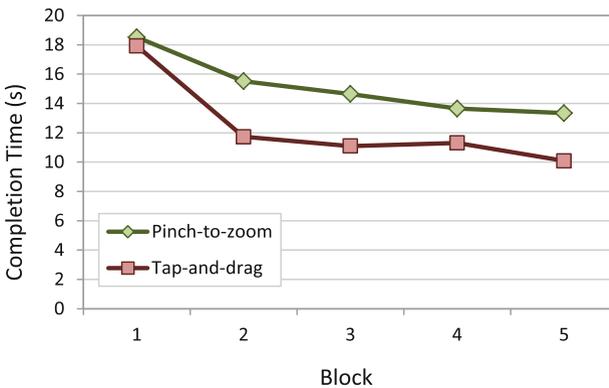


Fig. 8. Completion time (s) by zoom method and block.

As evident in Fig. 8, the improvement with practice was more dramatic for tap-and-drag than pinch-to-zoom. From the first to fifth block for tap-and-drag, the improvement in completion time was 43.7%. Over the same interval for pinch-to-zoom, the improvement was 27.9%. The greater improvement for tap-and-drag is likely due to participants having less experience performing zoom operations using tap-and-drag compared to pinch-to-zoom. Less experience increases the opportunity for improvement.

5.2 Accuracy

Accuracy was calculated for each trial as the ratio of time the participant was zooming in the correct direction to the total time. Idle time (discussed below) is the time in which the participant was not zooming. Idle time was not factored into the computation of accuracy.

The grand mean for accuracy was 87.1%. The interpretation here is that 87.1% of the time during which the participant was zooming, the zoom direction was correct; that is, the participant was correctly zooming in or zooming out. Conversely, 12.9% of the time the participant was zooming, they were zooming in the wrong direction.

The accuracy for tap-and-drag was slightly less than for pinch-to-zoom, with means of 86.0% and 88.3%, respectively. See Fig. 9. However, the difference was not statistically significant ($F_{1,10} = 3.70, p > .05$).

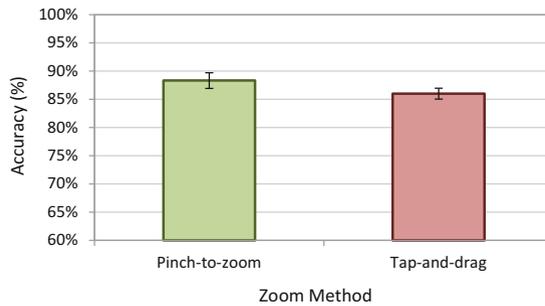


Fig. 9. Accuracy (%) by zoom method. Error bars shows $\pm 1 SE$.

It was observed that when performing the tap-and-drag gesture, participants would often drag their finger too far down the display, resulting in a zoom level that surpassed the required zoom level (thus decreasing accuracy). Once they noticed that the current zoom level was incorrect, they readjusted the zoom direction and zoom amount to achieve the correct level.

For tap-and-drag, participants also would sometimes drag their finger in the wrong direction when zooming (e.g., up instead of down). This confusion also contributed to the decreased accuracy.

Over the five blocks, tap-and-drag accuracy increased from 83.2% in the first block to 85.6% in the last block. The largest increase occurred between the first and second blocks. For pinch-to-zoom, accuracy increased from 88.1% to 89.8%. The progression in accuracy by block is shown in Fig. 10. However, the effect of block on accuracy was not statistically significant ($F_{4,40} = 1.31, p > .05$). The Zoom Method \times Block interaction effect on accuracy was also not statistically significant ($F_{4,40} = 1.68, p > .05$).

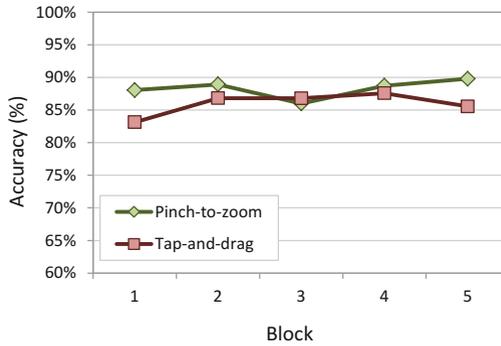


Fig. 10. Accuracy (%) by zoom method and block.

5.3 Efficiency

Efficiency was calculated by counting the number of panning or zooming gestures performed during a trial. Obviously, lower scores are better.

The grand mean for efficiency was 10.1 gestures per trial. The overall mean efficiency for the tap-and-drag trials was 7.0 gestures, while the overall mean for pinch-to-zoom trials was 13.3 gestures. The difference demonstrates a considerable efficiency advantage for tap-and-drag, as it was 47.2% more efficient than pinch-to-zoom. The difference was statistically significant ($F_{1,10} = 135.2, p < .0001$). This is clearly seen in Fig. 11.

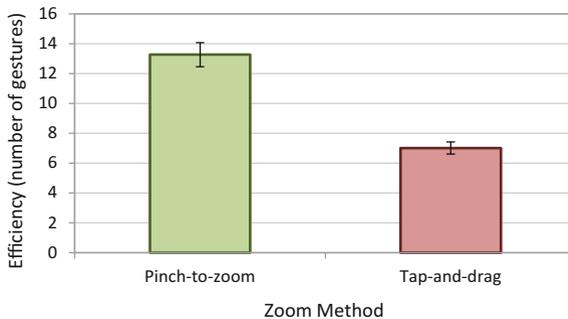


Fig. 11. Efficiency by zoom method. Lower scores are better. Error bars show $\pm 1 SE$.

The improvement in efficiency over the five blocks is shown in Fig. 12. The story here is slightly different. The progression is fairly flat for tap-and-drag with efficiency measures only varying from 6.72 to 7.56 gestures per trial. For pinch-to-zoom, however, there was a clear improvement over the five blocks. The efficiency measure was 14.8 in the first block and 12.3 in the fifth block, an improvement of 16.7%. Despite participants' prior experience with pinch-to-zoom, they clearly demonstrated a greater learning trend on the efficiency measure than with tap-and-drag. This we attribute to the nuances of

the task – panning and zooming a map image – as implemented in our experiment software.

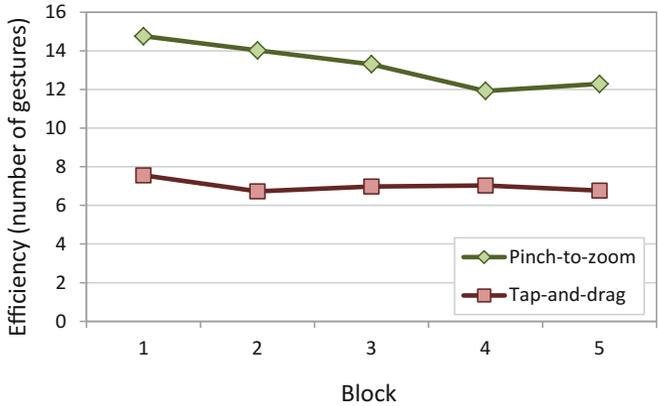


Fig. 12. Efficiency by zoom method and block.

Another reason for the poorer efficiency with pinch-to-zoom is clutching. When a participant’s fingers reach the edge of the display, they must lift their fingers, reposition them, and then “pinch” the display again to continue zooming. This is the most probable factor as to why the number of gestures (efficiency) is higher for pinch-to-zoom than for tap-and-drag. Clutching is required less frequently with tap-and-drag. This is examined in further detail below.

5.4 Idle Time and Zoom Patterns

The idle time was a significant portion of the total time per trial. On average, idle time accounted for 67.1% of the total time for tap-and-drag and 68.1% of the total time for pinch-to-zoom. This is mostly attributed to the participant panning the display in order to position the target on the marker, as opposed to zooming. The method of panning remained the same for both methods of zooming, and as such was irrelevant to our accuracy calculation. To be clear, idle time includes the time for panning operations.

Observing the difference between the participant’s zoom level and the target zoom level over time for each trial reveals why accuracy increased during the later blocks. Figure 13 shows the zoom difference over time for an example trial in the first block (left) and last block (right) for both tap-and-drag (top) and pinch-to-zoom (bottom) for a particular participant. For most participants, the trials followed a similar pattern.

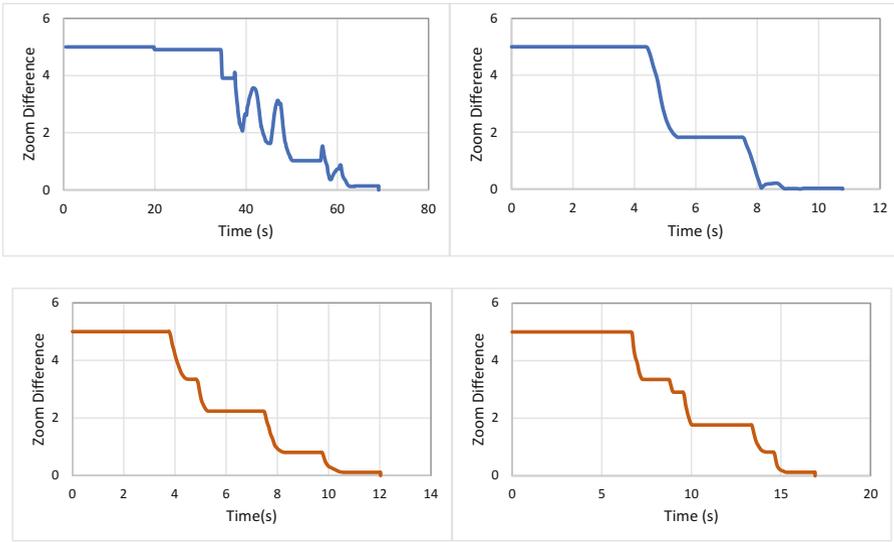


Fig. 13. Difference in zoom over time for a particular participant using tap-and-drag (top) and pinch-to-zoom (bottom) with example trials from the first block (left) and last block (right).

The trial in the first block for tap-and-drag (top-left) shows a high level of fluctuation in the zoom level. The most likely reason is that the participant was unaccustomed to the tap-and-drag zooming method, which resulted in difficulty controlling the zoom amount. These early trials contributed to the lower accuracy of tap-and-drag, and resulted in poorer accuracy than pinch-to-zoom. In hindsight, it may be beneficial to give the participants practice trials before data collection – to familiarize themselves with the tap-and-drag zooming method.

Observing a trial in the final block of tap-and-drag (top-right), however, shows significantly less fluctuation as the zoom level is more controlled and stabilized in comparison to the first trial. This is most likely attributed to learning effects.

The pinch-to-zoom trials generally followed the same pattern. Both the first and last trials follow a similar path, whereas the completion time slightly increased during the final trial. Participants were well accustomed to the pinch-to-zoom method, and as such performed consistently throughout the experiment.

5.5 Participant Feedback

Based on the questionnaire given at the end of the experiment, participants expressed positive feedback regarding tap-and-drag. All twelve participants indicated a preference to using tap-and-drag as opposed to pinch-to-zoom when using only one hand.

Most participants stated they were pleased by the comfortable hand posture afforded by tap-and-drag compared to pinch-and-zoom. They found it easier to hold the phone, and some expressed a worry that they might drop the phone when using pinch-to-zoom. Some participants noted hand fatigue and cramps when performing pinch-to-zoom. They

also noted that tap-and-drag was quick and easy to use, albeit having a slight learning curve.

One participant noted:

At first, I didn't like tap-and-drag. However, once I got the hang of it I liked it a lot – much more than pinch-to-zoom. It's hard for me to both hold the phone and pinch with one hand as my hand is small.

Overall, participants praised tap-and-drag for its comfortable hand posture and ease of use. When asked for a rating on a Likert scale from 1 (least likely) to 10 (most likely) on how likely they would use tap-and-drag when only one hand is available, the average rating was 9.5. That's a very favorable response for tap-and-drag.

6 Conclusion

An experiment was conducted comparing the performance of two zooming methods when performed with one hand. The results for tap-and-drag, a zooming method specifically designed for one-handed interaction, were generally good. The method was compared with the conventional pinch-to-zoom interaction, performed one-handed. Tap-and-drag performed 17.9% better in terms of speed and 47.2% better in terms of efficiency. The differences were statistically significant. Tap-and-drag was slightly worse in accuracy, although, in this case, the difference was not statistically significant.

The study also revealed that, as participants became more familiar with tap-and-drag, performance results improved. Participants performed better on later trials as they became more accustomed to tap-and-drag. Overall, participants gave a highly favorably and preferential rating for tap-and-drag for one-handed zoom gestures.

References

1. Boring, S., Ledo, D., Chen, X.A., Marquardt, N., Tang, A., Greenberg, S.: The fat thumb: using the thumb's contact size for single-handed mobile interaction. In: Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI 2014, pp. 39–48. ACM, New York (2012)
2. Karlson, A.K., Bederson, B.B., Contreras-Vidal, J.L.: Understanding one-handed use of mobile devices. In: Lumsden, J. (ed.) Handbook of Research on User Interface Design and Evaluation for Mobile Technology, pp. 86–101. IGI Global, Hershey (2008)
3. Lai, J., Zhang, D., Wang, S.: ContextZoom: a single-handed partial zooming technique for touch-screen mobile devices. *Int. J. Hum. Comput. Interact.* **33**, 475–485 (2016)
4. Malacria, S., Lecolinet, E., Guiard, Y.: Clutch-free panning and integrated pan-zoom control on touch-sensitive surfaces: the CycloStar approach, In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems - CHI 2010, pp. 2615–2624. ACM, New York (2010)
5. Miyaki, T., Rekimoto, J.: Zooming and scrolling control model for single-handed mobile interaction. In: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI 2009, Article No. 11. ACM, New York (2009)

6. Silverberg, M., Korhonen, P., MacKenzie, I.S.: Zooming and panning content on a display screen. In: International Patent Number WO 03/021568 A1. Assignee: Nokia Corp., Helsinki (2003)
7. Ti, J., Tjondronegoro, D.: TiltZoom: tilt-based zooming control for easy one-handed mobile interactions. In: Proceedings of Internet of Things Workshop at OZCHI 2012, pp. 1–4. ACM, New York (2012)