Virtual Buttons for Eyes-Free Interaction: A Study

Jens Bauer, Achim Ebert

HAL Id: hal-01599662

https://inria.hal.science/hal-01599662

Submitted on 2 Oct 2017

Virtual Buttons for Eyes-Free Interaction: A Study

Jens Bauer and Achim Ebert

Computer Graphics & HCI Group, TU Kaiserslautern
j_bauer@cs.uni-kl.de, ebert@cs.uni-kl.de

Abstract. The touch screen of mobile devices, such as smart phones and tablets, is their primary input mechanism. While designed to be used in conjunction with its output capabilities, eyes-free interaction is also possible and useful on touch screens. One of the several possible techniques for eyes-free interaction is the virtual button method, where the screen is divided into a regular grid of buttons that can be pressed even without looking at the screen.

This paper contains an exploratory study about influence factors on this interaction method. Results indicate, that not only the size of the buttons matter, but also the device orientation and user dependent factors, such as the age or general experience with touch screens. By involving small children in the evaluation we can see the validity of this approach even for the youngest users.

Keywords. Eyes-Free; Evaluation; Virtual Buttons; Mobile Devices

1   Introduction

In most cases, interaction with mobile phones and tablets is done via the touch screen of those devices. Users will naturally look at the screen while using it. However, there are some cases, where looking at the screen might not be possible or viable, e.g. while driving or when interacting with another system via the mobile device. This is called eyes-free interaction. A comprehensive list of such cases is presented by Yi et.al. [7]. Consumer-level devices do not provide haptic feedback on their screens, thus making this style of interaction difficult. One possibility of dealing with this is to divide the screen into a grid of virtual buttons. Other common approaches are described in the Related Work section later. There are studies showing the impact of different grid sizes on the accuracy (i.e., the hit-miss-ration of the user) for this kind of interaction. The paper from Wang et.al [6] is a prime example of this. More are covered later in the Related Work. No research has yet been done on the impact of different devices and their sizes on the accuracy and the duration of actually interacting, since the device size impacts those values inversely.

In this paper, we describe an exploratory study of eyes-free virtual button presses. Unlike previous studies, we look for the impact of not only the number of virtual buttons on the touch screen, but also the size of the device itself, attributes of the user, etc. To get further insight into the general usability of this method our participants are not only adults, but also small children. Children are getting their own mobile phones

(with touch screens) in early age, thus it is important to know, if methods working well for adults are actually viable for them.

In the next section we will give a brief overview of related literature. Then we will present the design of our evaluation and the process, followed by the actual evaluation of the results. We finish with the conclusions of the results.

2   Related Work

Yi et.al [7] presented a classification of motives for eyes-free interaction and identified several actual reasons. The four main categories are *Environmental*, *Social*, *Device Features* and *Personal*. The number and variety of reasons shows that eyes-free interaction is an important topic.

Alternatives for eyes-free interaction to the virtual button approach include Bezel Swipe [5] and Touch Gestures [2]. Bezel Swipe requires the user to touch one of the sides of the screen and swipe inwards. The tactile feedback with this technique makes it viable for eyes-free interaction. Touch gestures, which do not require any special position on the screen, such as a simple pinch gesture are also easy to use eyes-free.

Azenkot and Zhai [1] evaluated the effect of using a finger, a thumb or two thumbs on text input on touch capable smart phones.

An evaluation about virtual buttons on smart phones was done by Wang et.al. [6]. They found that grid sizes of bigger than 3x3 are not feasible to use for almost any user. Their evaluation was only done with one touch screen capable phone and only using one handed-interaction.

A huge evaluation done by Henze et.al. [3] focuses on the impact of several hardware design factors, such as device size, screen resolution, target sizes and position. They gathered data through a game distributed in an official app store. However their study was by design not abled to measure those effects on eyes-free interaction.

3   Evaluation

3.1   Design

The main goal of the study was to find out how different factors impact on the eyes-free usability of virtual buttons. Those factors are the size of the virtual buttons, as given by the touch screen size and the layout of the button grid. According to the results of Wang et. al. [6], button grids with more than 3 buttons in any direction were excluded, as were trivial layouts with only one or two buttons in total. The remaining grid layouts were : 1x3, 2x3, 3x3, 3x2 and 3x1. To gain insight about the effect of device size, we also used three different devices, an iPad 2 (241.2mm x 185.7mm, 9.7 inch display), iPad Mini (200mm x 134.7mm, 7.9 inch display) and an iPhone 4S (115.2mm x 58.6mm, 3.5 inch display). We chose to use only this device family to minimise hardware influences on the tests. For example the resolution of touch input (*not* screen pixels) is similar on all devices. This assures the App to behave the same on all devices, too. Since the device orientation might impact user performance as

well, all tests were to be done in both landscape and portrait orientation. As shown by Azenkot and Zhai [1], usage of thumb vs. finger can have an impact, too, and according to Karlson et.al. [4] this is the primary way of mobile touch interaction. Therefore each user was asked to complete all tests a second time using only one thumb in portrait orientation and two thumbs in landscape. In the other test runs, user had the freedom of choice of how to interact with the device, the only constraint being, the device has to be held with at least one hand and may not be put down on a table or anywhere else. Because of the physical size of the devices, this was only done for the phone. For additional data, users were asked for their experience with touch devices (self perceived), age and gender.

The evaluation design is straightforward. Every user is presented a grid of virtual buttons, displayed not on the mobile device, but on a separate computer screen, as we focus on eyes-free interaction. The mobile device screen stays completely blank, so if a user accidentally looks at the mobile device, they get no additional insight, and thus no advantage over other participants not looking at the device.

Depending on the complexity of the grid, 20, 30, or 40 touches were needed per grid layout and device orientation (more complex layouts needing more touches). Considering that every participant was asked to perform the test with three different devices and one device having an additional test run under special conditions described later, every user had to perform a total of 1520 touches. We used three different mobile device sizes, large tablet, small tablet and mobile phone to measure the impact of the device/screen size on the test.

To have a wider age range in the evaluation, we decided to not only have adult participants. To see how age actually affects the results of our study, we created another design, targeted at small children. This will help to understand if virtual buttons is a concept simple enough for even children to understand. We can also get an insight on the effect the grid layout and device size has on their performance, given their smaller hand sizes.

To get results across all possible grid setups, we reduced the number of touches needed per grid layout to ten. Also to help the children understand the task they were supposed to accomplish, their test started with a short example of the test with the grid visible on both the mobile device and the computer screen. These touches were, of course, not counted in the evaluation. The Thumb-Only run on the mobile phone was also not done by the children, since their hands were not big enough to accomplish it. To not negatively influence the study due to the children getting bored by the tasks, they had the option of stopping the test after each device. Thus, less data is available from the children's test, but still enough for an evaluation.

For the evaluation we developed a simple mobile application connected to a desktop application. The desktop application will display the current grid of virtual buttons and highlight the button the user is supposed to be pressing. To avoid confusion, the system guarantees that no button will be highlighted twice in a row, giving reliable feedback to the participants. The system automatically randomizes the test order (order of device, grid layout and button to press) to avoid learning effects affecting the results. The grid was displayed on a standard LCD computer display with 21 inches of

size, connected to a desktop PC. A dedicated WiFi connected the PC and the mobile device.

Each touch was recorded individually. Accuracy and duration was measured by aggregating the gathered data. Accuracy is simply the ration of hitting the correct button and the total number of button requests. Duration is the total time between a button prompt being displayed and it being pressed, summed up for each test run.

For the test involving adults, we had 44 participants (30 of them male, 14 female) in the age range of 21 to 62 years. All test candidates were asked to rate their own experience with touch screen interaction (on a Likert-scale of 1-5) and to state if they own a touch device themselves (which the majority of 34 users did). The average experience level was 3.57, unsurprisingly due to the high number of touch device owners in the test. As a side note, the high number of touch device owners shows the ubiquity of these devices and does not affect the viability of the test as we still have 10 participants not owning such a device.

The 29 children (18 male, 11 female) were between 4 and 6 years old. They were not asked to rate their experience, as children of that age are unreliable of their self-assessment. But they were asked, if anybody in their family owns a touch device that they are allowed to use on a regular basis. 18 of the 29 children confirmed that, while 9 had no prior access to touch devices (2 children opted to not to answer that question).

3.2   Results

To get basic results from the study, we tested for influence of four basic factors on the accuracy. The factors are the device orientation (Landscape/Portrait), a combined factor of the device used and if the user was asked to use only their thumb, if the user actually owns a touch device, and the self-rated experience value.

The mean values are given in Table 1 below. As supposed from the related work, the accuracy results are above 90% for adults. Children's accuracy is significantly lower at values between 70% and up to more than 90%. This shows, that children are actually able to understand the concept of virtual buttons, even for eyes-free interaction. But due to the lower hit-ratio, care must be taken to allow undo-operations or similar mechanisms when designing applications for children.

Portrait orientation supports higher accuracy for both adults and children. It also influences the duration, with adults being faster in portrait orientation. Interestingly children were faster using the (less accurate) landscape orientation. More evaluation is needed in order to gain insight into this.

As expected, bigger devices allow for higher accuracy. It also seems, that the bigger device size does not cause the duration to go up (due to longer ways to move fingers), but instead seem to lower the duration. This is probably caused by users being more confident in hitting the desired area.

Using the thumbs only for interaction with the phone seems to actually increase accuracy and interaction speed. Since the alternative involves moving the whole hand to reposition the finger, this is not surprising.

Experience levels do also influence both accuracy and duration, mostly in the expected way. There is a outlier in the data, as the low experience level of 2 performed faster than levels 1, 3 and 4. The cause of this is unclear at this point.

Gender does not seem to influence accuracy at all for the adults. Though male candidates performed the requested tasks quicker than females. The cause of this needs more evaluation. Even more interesting is the fact, that with children this is actually reversed. Girls were more accurate hitting the virtual buttons and even performed fasters.

Table 1. Base Results. Values are divided into children and adult values. Duration is average duration per touch in seconds.

Classification	Accuracy		Duration	
	Adults	Children	Adults	Children
Orientation				
Portrait	0.9638	0.8222	0.7840	3.008
Landscape	0.9538	0.7928	0.8534	2.972
Device				
iPad 2	0.9736	0.8571	0.7689	1.629
iPad Mini	0.9687	0.8115	0.7736	1.434
iPhone 4S	0.9300	0.7566	0.8863	5.887
iPhone 4S (Thumb)	0.9500	N/A	0.8440	N/A
Owns Touch Device				
No	0.9515	0.8595	0.9207	6.175
Yes	0.9631	0.7756	0.7947	1.579
Experience Level				
1	0.9203	N/A	0.9637	N/A
2	0.9391	N/A	0.7848	N/A
3	0.9717	N/A	0.8681	N/A
4	0.9798	N/A	0.8277	N/A
5	0.9695	N/A	0.7405	N/A
Gender				
Male	0.9557	0.7637	0.7640	3.967
Female	0.9597	0.8948	0.9341	1.648

Looking at the actual target button sizes, as defined by touch screen size and the grid layout yields the results seen in Table 2 and 3. For Table 2 the results were grouped by the minimal dimension of the button. For example, a button size of 10x15 mm would then have a minimal dimension of 10mm. In Table 3 the whole square size is used to group the results. Roughly the results reinforce the hypothesis, that bigger target sizes will lead to better accuracy and execution speed. Not all bigger sizes perform better than smaller sizes. More research is needed here to find the cause of this. Generally the standard deviation (and therefore variance) of accuracy and duration decreases with increasing button size (in both tables). This might be a sign of users being more confident with bigger sizes or simply more users can reliably use the vir-

tual buttons as their size increases. There also seems to be a point till where user performance increases faster with increasing size. Unfortunately our data is not abled to support such a hypothesis statistically, so this is also a point needing further research.

Table 2. Accuracy and duration by minimal dimension of target button (i.e., a button size of 10x15mm has a minimal dimension of 10mm). Columns contain mean values and standard deviation. Size is given in milimeters and duration in seconds.

min.	Accuracy		Duration	
Size	Adults	Children	Adults	Children
19.5	0.9349 ± 0.1103	0.7448 ± 0.2805	0.9191 ± 1.2987	5.647 ± 16.3594
29.3	0.9569 ± 0.1009	0.8139 ± 0.2489	0.8223 ± 0.6596	6.238 ± 23.7814
38.4	0.9464 ± 0.1071	0.7595 ± 0.2443	0.7452 ± 0.2164	5.681 ±17.0854
44.9	0.9645 ± 0.0907	0.8050 ± 0.2595	0.7843 ± 0.1672	1.524 ± 0.6258
61.9	0.9683 ± 0.0679	0.8464 ± 0.1685	0.7896 ± 0.2002	1.660 ± 0.6454
66.7	0.9665 ± 0.0749	0.7970 ± 0.2514	0.7764 ± 0.2366	1.344 ± 0.6166
67.4	0.9790 ± 0.1026	0.8735 ± 0.2013	0.7333 ± 0.1874	1.385 ± 0.6058
80.4	0.9692 ± 0.0666	0.8711 ± 0.1413	0.7591 ± 0.1591	1.678 ± 0.7999
92.9	0.9938 ± 0.0166	0.8965 ± 0.1803	0.7190 ± 0.1111	1.503 ± 0.5611

Table 3. Accuracy and duration by size of the button. Columns contain mean values and standard deviation. Size is given in square millimeters and duration in seconds.

Sq.	Accuracy		Duration	
Size	Adults	Children	Adults	Children
750	0.9223 ± 0.1048	0.7192 ± 0.3213	0.9038 ± 0.2660	5.8143 ± 17.620
1125	0.9397 ± 0.1203	0.7679 ± 0.2690	0.8764 ± 0.9219	6.4241 ± 24.571
1688	0.9777 ± 0.0805	0.8415 ± 0.2279	0.8240 ± 0.9643	5.6302 ± 16.959
2250	0.9538 ± 0.0918	0.7573 ± 0.2474	0.8348 ± 1.5992	5.5466 ± 16.877
2993	0.9611 ± 0.0781	0.7941 ± 0.2644	0.8260 ± 0.1703	1.6114 ± 0.5039
4490	0.9645 ± 0.0992	0.8012 ± 0.2779	0.7965 ± 0.2172	1.4944 ± 0.6702
4977	0.9625 ± 0.0836	0.8297 ± 0.1682	0.8210 ± 0.1668	1.7047 ± 0.4986
6735	0.9790 ± 0.1026	0.8735 ± 0.2013	0.7333 ± 0.1874	1.3853 ± 0.6058
7465	0.9700 ± 0.0625	0.8734 ± 0.1535	0.7905 ± 0.1977	1.8205 ± 0.8531
8980	0.9710 ± 0.0569	0.8058 ± 0.2277	0.7063 ± 0.1505	1.3203 ± 0.6133
11198	0.9938 ± 0.0166	0.8965 ± 0.1803	0.7190 ± 0.1111	1.5032 ± 0.5611
14930	0.9718 ± 0.0583	0.8527 ± 0.1573	0.7108 ± 0.1632	1.4917 ± 0.6016

We did a Pearson's test of correlation between the accuracy, user's age, the virtual button size, the minimal length or width of a button (called MinSize for the remainder of this paper) and the duration of a test. We got a significant correlation between accuracy and age, button size and MinSize for both test groups. The same goes for duration, respectively. Since Pearson's test of correlation is well known to be impacted by

outliers, we confirmed our results with an additional Spearman-correlation test. All correlations were confirmed, except for the correlation of duration and MinSize in the children group, so correlation here is improbable.

Analyses of Variance (i.e. ANOVE) was then done to test the influence of the virtual button position on the accuracy. Separate ANOVAs were performed for each combination of device and grid layout. One of the interesting findings here is the fact, that for all devices the middle button in the 1x3 (one single row of 3 buttons) layout was significantly (on a 5% level) harder to hit, while this does not hold true for the 3x1 layout. Here no significant differences could be found. Other layouts with statistically significant inaccurate buttons are 2x3, 3x2 and 3x3.

For the 3x3 layout, the buttons in the middle column had least accuracy, with the top and bottom button in this column having even lesser accuracy than the middle one. This holds true on all devices. For this layout the middle device size has the highest accuracy for all individual buttons. In the 3x2 and 2x3 layouts, also the middle buttons had lower accuracy than the other buttons, again on all devices. This is probably due to the fact that buttons near a corner are easier to find without looking on the device.

4   Conclusion

Our evaluation shows, that there are several influence factors confirmed for eyes-free virtual button interaction on smart phones and tablets. The device size is important as one might assume as is the actual size of the target. As the improvement of accuracy when using the iPad over the iPad Mini is not significant, this hints, that there is a point, where enlarging the mobile device does not yield better results for accuracy. Since grid sizes bigger than 3x3 virtual buttons are strongly discouraged by Yi et.al. [7], this is understandable.

For the design of virtual buttons we can conclude, that it is important to use the minimal possible amount of buttons and to move secondary functionality into other interaction mechanisms, like gestures or an on-screen menu. Especially since application designers cannot influence the factors of age and touch screen experience of their potential users, they are well-advised to keep the user interface as simple as possible. Especially when developing for platforms that target several different device sizes (such as Android or iOS), the designers have to keep in mind, that they might not know the actual device size their application is used and thus not the actual target sizes of virtual buttons they use.

When using only three different buttons, it is advised to arrange them in column order, as there is no button with lower accuracy in that layout as found out above. In general the 3x1 and 2x2 layouts are most suited for eyes-free interaction. If more functions are needed, applications designers should take care to identify critical functionality, that causes most problems if invoked by mistake. Important functions should be assigned to a button in a corner of the device, unimportant ones can be connected with the middle buttons. The centre button in the 3x3 layout has a special role, as it has higher accuracy than the buttons to its sides, but less than the corner buttons.

We can also conclude, that virtual buttons is a concept, that is simple enough for even some small children to understand and use properly. As the number of children owning their own touch device is steadily increasing, this is an important point for application design. While the actual accuracy numbers are fairly low for the children test group, one has to keep in mind that eyes-free interaction is not easy for children at all. As future work the results mentioned in this paper need thorough evaluation to statistically prove their validity. Especially looking for a perfect size and the influence of grid layout and device size is very interesting for further evaluation. Still the results presented can be used as a hint for designing apps using the virtual button approach.

5   References

1. Azenkot, S., and Zhai, S. Touch behavior with different postures on soft smartphone keyboards. In Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services, ACM (2012), 251–260.
2. Bragdon, A., Nelson, E., Li, Y., and Hinckley, K. Experimental analysis of touch-screen gesture designs in mobile environments. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM (2011), 403–412.
3. Henze, N., Rukzio, E., and Boll, S. 100,000,000 taps: analysis and improvement of touch performance in the large. In Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, ACM (2011), 133–142.
4. Karlson, A., Bederson, B., and Contreras-Vidal, J. Understanding single-handed mobile device interaction. Handbook of research on user interface design and evaluation for mobile technology (2006), 86–101.
5. Roth, V., and Turner, T. Bezel swipe: conflict-free scrolling and multiple selection on mobile touch screen devices. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM (2009), 1523–1526.
6. Wang, Y., Yu, C., Liu, J., and Shi, Y. Understanding performance of eyes-free, absolute position control on touchable mobile phones. In Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services, MobileHCI '13, ACM (New York, NY, USA, 2013), 79–88.
7. Yi, B., Cao, X., Fjeld, M., and Zhao, S. Exploring user motivations for eyes-free interaction on mobile devices. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12, ACM (New York, NY, USA, 2012), 2789–2792.