Twist and Pulse: Ephemeral Adaptation to Improve Icon Selection on Smartphones

Abstract

The concept of ephemeral adaptation was introduced to reduce visual search time in GUI menus, while preserving spatial consistency and minimizing distraction. We extend this concept to the visual search of app icons on smartphones in order to speed up launching apps from a homescreen. We created ephemeral highlighting effects based on preattentive visual properties including size, orientation, color, opacity and blur. We then conducted informal rapid design and evaluation cycles from which Twist (icon rotates back and forth) and Pulse (icon grows and shrinks) were the most promising effects. An experiment comparing these two effects to a control condition showed that these effects improve search time performance by 8-10% and Pulse is subjectively preferred to the control condition.

Keywords: Adaptive interfaces, interaction techniques, visual search, preattentive property, icons, smartphones.

Index Terms: H.5.2 [User Interfaces]: Evaluation/methodology, interaction styles.

1 Introduction

The growing variety and number of mobile applications, or *apps*, have dramatically expanded the functionality of smartphones. Not surprisingly, the number of apps installed by users has rapidly increased in recent years [2]. This growth has resulted in several usability problems, one of which is searching visually for the desired icon in order to launch an app [2]. This problem is exacerbated because the icons commonly used for these apps are often complex and lack adequate discriminatory features [3].

There is, however, some regularity to the patterns of app usage that can be leveraged in order to speed up access: only a small proportion of apps are used frequently, and different apps are used in different contexts [2,7,9]. These patterns have enabled successful prediction of the most probable app selections [7,9].

To address the problem of accessing app icons, researchers have tried to leverage these predictive models by designing adaptive interfaces that speed up access to the predicted icons. Two common techniques for visually distinguishing the predicted icons from other distractor icons have been studied: static highlighting, for instance with a bright surrounding halo [9], and copying the likely-to-be-used apps to a separate part of the interface [9]. None of these techniques have been shown to significantly improve icon selection performance. The main problem with static highlighting is that its permanence can negatively impact performance—by distracting the user—if the algorithm does not predict the target icon. In contrast, copying icons results in spatial inconsistency, which has been shown to confuse users [5,9].

In the context of menu selection, Findlater et al. addressed the problem of visual search by introducing a technique called *ephemeral adaptation* [4], which consists in adapting the interface and reverting it back to its normal state after a short period of time. Findlater et al. implemented this idea by first displaying the predicted items then gradually fading in the non-predicted items, maintaining spatial consistency throughout. The temporal dimension is thus used to attract attention to predicted items. This technique was shown to result in faster selection times than static highlighting.

Inspired by this idea, we extend and evaluate the concept of ephemeral adaptation to the visual search of icons on smartphones. This task is significantly more complex than searching items on a menu: while menus are vertical lists of text labels, app icons are often organized on a grid and vary considerably in visual complexity. Additionally, users typically have to optimize their search strategy across several pages of apps.

Our proposed solution is to leverage the human visual system by using preattentive visual properties [8], including size, orientation, color, opacity, blur, and motion to increase the saliency of predicted icons—those deemed likely to be used. However, speeding up icon search is more than just attracting attention to the predicted icons: it requires balancing attention and distraction. Drawing attention to the predicted icons is necessary to speed up the visual search, but must be balanced with minimizing the distraction caused when the prediction is wrong. Otherwise the speedup gained for correct predictions will be cancelled out, on average, by incorrect predictions—which typically result in much longer selection times. Ephemeral adaptation is intended to achieve this balance, but it has never been extended beyond its original formulation of abrupt onset and gradual fade-in. No other visual dimensions were explored.

While previous work has proposed applying visual transformations to icons, and in particular adding simple motions, these visual transformations were designed for very different purposes than ephemeral adaptation. Moticons [1] support a dual task, namely they attract the user's attention away from her primary task. In our approach, the transformations are used in the primary search task itself. Besides, their work did not take into consideration an *incorrect attentional draw,* since the animated notifications were assumed to always be meaningful. Kineticons are another example of adding motions to icons, but they were designed as an "iconographic scheme based on motion" [6]. While Kineticons resemble some of our ephemeral adaptation effects, their purpose is to convey semantics—their impact on attention and distraction have not been studied.

The contributions of this paper are as follows. To the best of our knowledge, we are the first to extend the ephemeral technique to visual icons. We explore the design space of ephemeral highlighting for icons to identify visual effects that carefully balance attentional draw and distraction. In doing so, we

* email address

Figure 1: Direct and indirect ephemeral highlighting effects applied to the predicted weather icon (top right icon). The other three icons are non-predicted. Each effect is labelled with its corresponding preattentive visual property (top), and the name we used to describe it to participants (bottom, in quotes). The arrows indicate how the direct highlighted icons change over time.

document which preattentive visual properties work well for ephemeral highlighting and which do not. The two most promising visual effects that emerged are Twist (icon rotates back and forth) and Pulse (icon grows and shrinks), as shown in Figure 1. Our experiment shows that these two effects significantly improve search performance over a control condition, independent of the number of pages of icons and the accuracy of the predication algorithm used. Pulse is also preferred subjectively to the control condition with no highlighting.

2 Ephemeral Highlighting Effects

The design goals for our ephemeral highlighting effects were threefold: to attract attention to predicted icons, to maintain spatial consistency, and to minimize distraction so as to be suitable for everyday use. Preattentive visual properties are a natural candidate for attracting attention to predicted icons, by increasing their saliency. Through rapid iterative design and evaluation cycles, we found the effects that best match these design goals.

2.1 Design Space

Our exploration of the design space was intended to be thoughtful but also opportunistic, starting with a reasonable set of preattentive properties and varying their parameters, rather than exhaustive of all possibilities. Figure 1 shows a sample set of designs. The nature of the individual properties means that we either *directly* transform predicted icons to make them more salient, or we "dim" all the non-predicted icons to make the predicted ones stand out—an approach we refer to as *indirect* highlighting.

Each preattentive property includes a number of parameters that further expand the design space. For instance, we considered a range of angles for Rotate, from 10° to 180°, and explored different sizes for Shrink, from 10% larger to 50% larger. Additionally, we experimented with the opposite effect (Grow, not shown in Figure 1) with sizes ranging from 10% to 50% smaller. For indirect highlighting effects, we varied the alpha value (Transparency) of the non-predicted icons from .1 to .5, and the Gaussian blur radius from 2 to 10.

The ephemeral nature of these highlighting effects requires the interface to return to its normal state after a certain period of time. The total duration of the effects is therefore a critical parameter. Findlater et al. explicitly compared 250ms and 500ms durations, and found the latter to be the most effective. However, 500ms was often not enough for indirect highlighting effects to be perceived; for these effects, we increased the total duration to 1000ms.

To make the effects more visually appealing we added smooth transitions between the highlighted state and the normal state. In the case of indirect highlighting like Greyscale, the colors of the non-predicted icons simply fade in slowly. For Size and Orientation, the predicted icons are animated back to their normal appearance (hence the names Shrink and Rotate). We also tried adding smooth transitions at the beginning of the effects to reach the highlighted state. For example, Twist and Pulse are variants of Size and Orientation with a continuous animation from normal to highlighted then back to normal. There is therefore no abrupt visual change when a new page of icons is loaded. We explored different durations for these transitions between states, from 100 to 1000ms. We found that a delay before transitioning was useful for the indirect highlighting effects as it gave more time to find the highlighted icons; however, this pause was found superfluous and even annoying for the direct effects.

We note that the use of transitions with the direct effects meant that highlighted icons were perceived as being in motion. While motion is one of the best ways to attract attention [1], it was not clear if it would be appropriate for our task, or whether the indirect techniques would provide a better balance between salience and distraction. We iteratively refined the parameters of the motion effects so that they would not be overly annoying (given that there were multiple icons highlighted per page) nor overly distracting (in the case of wrong prediction).

2.2 Preliminary Feedback

We iteratively refined our design via informal evaluation with five smartphone users, and then conducted a pilot study with 10 participants. Participants were asked to select a target icon among three pages of apps on a smartphone, with 20 icons per page. In total, 9 icons were highlighted, but only one type of highlighting effect was used at a time. The goal was to identify the most promising effects, and to adjust their parameters, such as: total duration, transition duration, size, angle, etc. The key findings are summarized below.

Blur was discarded early: the in-focus predicted icons could hardly be told apart from non-predicted blurred ones, unless the latter were degraded to unrecognizable color blobs.

Color and Opacity were more effective than Blur, but participants sometimes had trouble identifying which icons were highlighted, especially when several were highlighted on the same page at once. Color was particularly problematic: as one of our participants pointed out, *"some of the icons don't have a color to begin with!"* Additionally, all indirect highlighting effects share the same weakness: they degrade the appearance of non-predicted icons, making them less recognizable. If the algorithm's prediction is wrong, it is harder for users to find the icon they are looking for: the user must wait for the ephemeral effect to end and the non-predicted icons to revert to their normal appearance, thus diminishing the benefits of adaptive highlighting.

The effects involving some form of motion were found to be the most effective: Shrink, Rotate, Twist and Pulse. We learned that increasing the size of an icon is more likely to attract attention than decreasing it, likely because it visually corresponds to an object coming towards the observer. Most participants expressed a preference for Pulse over Shrink, finding it "*more aesthetic*" and having *"an organic feel"*. Twist was unanimously well received, and described as *"cute"*, *"playful"* and even *"chic"*. We therefore selected Twist and Pulse for further study. The key final parameters were as follows: a maximum size of +25% for Pulse, a maximum angle of ±15° for Twist, and a transition duration of 450ms plus a short 100ms delay at the outset for both.

3 EXPERIMENT

The main goal of this experiment was to determine if Twist and Pulse could improve performance for selecting icons, by reducing the selection time and/or the number of errors. We also investigated two factors: (1) the number of pages of apps users have to search through, in order to assess the impact of app volume; and (2) the accuracy of the underlying prediction algorithm, as adaptive interfaces are known to cause frustration when the accuracy of the prediction is inadequate [5].

3.1 Methodology

3.1.1 Participants & Apparatus

The study included 12 participants (mean age 25, 4 females), all frequent smartphone users. On average, participants had 50 apps across 3.8 pages on their own phones (1 min, 8 max pages). The custom experiment application ran on two Google Nexus 5 devices. App icons were retrieved from Icon100, an online database (www.icon100.com), while labels for the icons were sampled from a corpus of neutral nouns.

3.1.2 Task

The experiment consisted of a sequence of icon selection trials. In each trial, participants were shown a target icon consisting of an image and a label. Tapping once on the screen brought users to the first page of 20 icons, arranged in a 5x4 grid. Participants could navigate through the pages by swiping. Once they clicked on an icon, the target for the next trial was displayed. If an icon other than the target was selected, the application provided feedback and proceeded to the next trial. The system reported a timeout when a trial hit 20 seconds, and moved on to the next trial. To generate a sequence of icon selections, target icons were randomly sampled from a Zipfian distribution, as in Findlater et al [4]. For 40 trials, we sampled 20 unique target icons with relative frequencies 8, 5, 3, 3, 2, 2, 2, 2, 2, 1 … 1 (Zipfian $R^2 = .93$).

3.1.3 Pages, Accuracy & Prediction Algorithm

To represent a range of possible configurations, we chose to try both 3 and 6 pages of apps, roughly equivalent to an average and a large number of apps on a smartphone. To determine realistic accuracy levels, we used Shin et al.'s comparison of various algorithms for predicting app usage on smartphones [9], which showed that the accuracy of prediction algorithms depends on the number of apps predicted. Based on our pilot study, we decided to highlight on average 3 icons per page, to ensure that participants could perceive all of them. Therefore, 9 and 18 icons were highlighted for 3 and 6 pages, respectively. According to Shin et al., this corresponds to accuracy levels of 80% to 95%. We used these two values in our experiment.

Predicted icons were highlighted by a custom prediction algorithm, similar to the one used by Findlater et al [4]. It combines three prediction strategies: most frequently used (MFU), most recently used (MRU), and random selection. The predictions of the algorithm were randomly adjusted before the experiment to reach the desired accuracy levels, defined as the percentage of trials in which the target icon is among the ones predicted by the algorithm (and therefore highlighted with one of our effects).

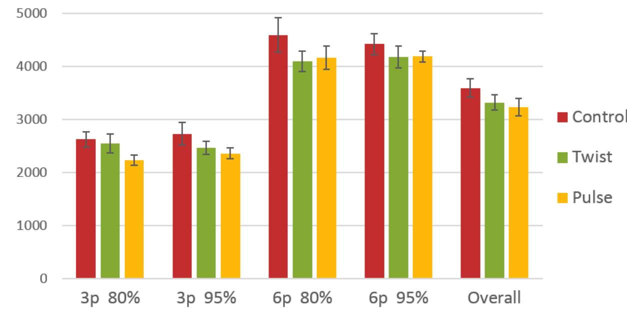

Figure 2: Selection times (ms) in each combination of accuracy and number of pages. Overall aggregates across all the data.

3.1.4 Design

A 3-factor within-subject design was used: 3 highlighting effects (Twist, Pulse, Control—no effect) x 2 prediction accuracies (80%, 95%) x 2 number of pages (3, 6). Prediction accuracy and number of pages were fully counterbalanced across participants, while the order of highlighting effects was determined by a Latin Square. Both icons and labels were randomly sampled for each combination for each participant. Each participant completed 480 trials (12 combinations x 40 trials) for a total of 5760 trials across all 12 participants.

3.1.5 Procedure

The experiment fit into a single 90-minute session. Participants first completed a demographic questionnaire before doing all 12 combinations of the three factors. In each combination, participants had a chance to practice, followed by 40 timed trials. Both completion time and error rate were recorded. Participants first saw the 6 combinations at one accuracy level before moving onto the combinations with the other accuracy level. Participants were informed that the prediction may occasionally be wrong, and they were told at the beginning of their second assigned accuracy level that the algorithm had changed and it could be better or worse than before. After each accuracy level, participants were given a questionnaire with subjective Likert scale questions. At the end of the study, they were asked to compare and rank their overall preference and perceived performance for the three effects.

3.2 Results

Prior to the analysis, we removed all the trials that timed out. Out of 1920 total trials per highlighting effect, we removed 32, 30 and 27 timeouts corresponding to Control, Twist and Pulse, respectively. Additionally, participants selected the wrong icon 39, 33 and 22 times for Control, Twist and Pulse, respectively. These trials were excluded as well, to avoid disadvantaging Control.

3.2.1 Speed

Selection time was measured as the time from tapping on the displayed target icon to selecting an icon. We ran a 3-way RM ANOVA on selection time. As Figure 2 shows, there was a significant main effect of Highlighting Condition ($F_{2,22}$=10.60, p=.001, η^2=.491). Twist and Pulse were both significantly faster than Control (p<.01 and p<.05, Bonferroni corrected), but not significantly different from each other (p=.87). Twist was 8% faster than Control, while Pulse was 10% faster. As expected, there was a main effect of the number of pages ($F_{1,11}$=456.32, p<.001, η^2=.976), with trials for 6 pages being much longer than trials for 3 pages. However, there was no main effect of accuracy (p=.90), nor any interaction effects.

3.2.2 Subjective Finding

We ran non-parametric Friedman tests on the Likert scale questions and Chi square on the comparative ranking questions. While we asked participants for their preference after each level of prediction accuracy, there was no significant difference based on accuracy. The comparative ranking at the end, however, did show a significant difference in users' preference and their perceived performance of the highlighting effects ($\chi^2_{(2, 12)}$= 18.19, p=.001, and $\chi^2_{(2, 12)}$=22.82, p<.001, respectively). 11/12 participants preferred Pulse to Control, perceiving it as being faster; only 6/12 participants preferred Twist to Control despite 9/12 perceiving it as being faster.

4 DISCUSSION

The results of our proof-of-concept experiment demonstrate the potential of extending ephemeral adaptation for selecting app icons. Both Pulse and Twist were faster than Control by 8 to 10%—a speedup comparable to previous work in adaptive interfaces [4].

Additionally, while we expected better performance for 95% accuracy than 80% accuracy, we found no effect of accuracy. Even though the algorithm made 4 times fewer errors in the 95% case, these errors resulted in longer selection times for wrong predictions (4.9s on average at 95% accuracy, vs 4.2s at 80% accuracy). The fact that prediction errors are more costly when they occur less frequently is perhaps not surprising, but that this additional cost fully balanced out the effect of accuracy is surprising.

Because the two levels of accuracy did not impact the performance of the participants, our experiment shows that performance is not sensitive to having an imperfect prediction algorithm—users benefit from this technique even at 80% accuracy, which is achievable with 9 highlighted icons or more simply by using MRU or MFU icons [9]. A common concern with adaptive systems is that they will not be effective unless the accuracy is very high. Here we show that even at 80% it is effective. In addition, there was no interaction between highlighting effect and the number of pages, which suggests likewise that our ephemeral techniques are not particularly sensitive to the number of pages of apps.

Although Twist and Pulse fared similarly in terms of selection time, participants expressed a clear preference for Pulse, judging it more effective for finding icons and less distracting than Twist. This difference was not expected, as both effects involve a similar type of motion. One possible explanation is that rotating the icon creates a motion blur stronger than scaling it up and down. Thus, even if both effects are equally effective at signaling which icons are highlighted, it may be harder to recognize an icon that twists than one that pulses.

While our results are promising, our experiment has some limitations, which come mostly from the simplified homescreen that we used. Recent mobile operating systems offer several mechanisms to help users launch apps, such as sorting them alphabetically, or allowing users to customize their homescreen by repositioning icons or grouping them into folders. Our technique should still provide benefits if apps are sorted alphabetically, or grouped into folders (the highlighting would follow the folder hierarchy). However, if users move all their frequently used apps to the first page of their homescreen, the prediction algorithm should be modified to avoid highlighting too many icons—perhaps at most five icons per page. A possible way to reduce the number of highlighted icons per page would be to not highlight the top most frequently used apps, as users already know their positions well and rely on muscle memory to launch them. But that approach would require further investigation. Regardless, past work has found that a large proportion of users do not customize their homescreen [2], and could therefore benefit directly from our highlighting effects.

5 CONCLUSION

Ephemeral adaptation is a promising technique to improve selection of app icons on smartphones. Among the many preattentive visual properties we tried to draw attention to predicted app icons, those based on motion were the most effective. Our ephemeral effects use smooth motion to highlight several predicted icons, and revert back to their original state afterwards. A controlled experiment found that Twist and Pulse significantly improved icon selection time, and Pulse was preferred over Control by all but one participant. These results are robust across medium and large number of apps, and are surprisingly robust to prediction accuracies as low as 80%.

While we evaluated ephemeral visual highlighting for icon selection on smartphones, it is likely to provide benefits in other contexts as well, when the task involves selecting an element among a large set of images. For example: photo galleries, album covers in music libraries, and the app panels featured in recent desktop operating systems.

REFERENCES

[1] Bartram, L., Ware, C., and Calvert, T. Moticons: detection, distraction and task. *International Journal of Human-Computer Studies 58*, 5 (2003), 515–545.

[2] Böhmer, M. and Krüger, A. A Study on Icon Arrangement by Smartphone Users. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2013), 2137–2146.

[3] Byrne, M.D. Using icons to find documents: simplicity is critical. *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems*, (1993), 446–453.

[4] Findlater, L., Moffatt, K., McGrenere, J., and Dawson, J. Ephemeral Adaptation: The Use of Gradual Onset to Improve Menu Selection Performance. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2009), 1655–1664.

[5] Gajos, K.Z., Czerwinski, M., Tan, D.S., and Weld, D.S. Exploring the Design Space for Adaptive Graphical User Interfaces. *Proceedings of the Working Conference on Advanced Visual Interfaces*, ACM (2006), 201–208.

[6] Harrison, C., Hsieh, G., Willis, K.D.D., Forlizzi, J., and Hudson, S.E. Kineticons: using iconographic motion in graphical user interface design. *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, ACM Press (2011), 1999.

[7] Huang, K., Zhang, C., Ma, X., and Chen, G. Predicting mobile application usage using contextual information. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, (2012), 1059–1065.

[8] Palmer, S.E. *Vision science: Photons to phenomenology.* MIT press Cambridge, MA, 1999.

[9] Shin, C., Hong, J.-H., and Dey, A.K. Understanding and Prediction of Mobile Application Usage for Smart Phones. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ACM (2012), 173–182.