Text reading performance on small circular display

Xiaorui Dong[1], Yupeng Zhang[2], Maxwell Yang[3]

{1,2,3}@University of Toronto, {dongx;maxwell.yang;Yupeng.zhang}@utoronto.ca

Abstract

Advance in manufacturing technology have enables the display shapes with arbitrary shapes, and circular display one of the form received most of the focus from the industry. To addressing the issue of layout text and other elements properly in the circular display, we proposed a novel text layout design to improve screen utilize and reduce distortion effect. We also conduct a formal study to compare our technique with two kind of traditional optimized layout technique in quantitative manner. Results shows a significant improvement in both reading time and user satisfaction. We also drives a general design principle to layout text or other elements in non-rectangle display.

Author Keyword

Circular display, Text layout, Reading

Introduction

The industry is becoming more active on designing displays with different shapes. Sharp Corporation recently released a prototype display with a free-form boundary shape, which leaves designers and manufacturers with more design freedom and challenges. Circular displays might be one of the most popular designs. For instance, the smart watches, Moto 360 and LG G Watch R, show a promising future on the potential popularity of circular displays. Unlike conventional displays, the design guideline for the circular display would be quite different. There have been many researches done on the readability of rectangular displays, but very few on circular displays. Therefore, we conducted a research on the performance of text reading on circular displays.

In recent years, technology is shifting from desktops to smaller devices such as mobile phones, tablets, and watches. The display space has always been a limiting factor on how we can view information as clear as possible while presenting the content effectively through various methods such as panning, scrolling, zooming, and rotating. This becomes more of a challenge as we begin to fit more information onto smaller devices such as circular-display watches. There have been implementations of frameworks, such as the Elastic Presentation Framework (EPF), that unifies distortion and non-distortion presentation paradigms to effectively use display spaces [1]. The focus area in EFP is varied through a series of different boundary shapes and research papers show that users prefer shapes with minimized distortion. However, the research conducted is focused on graphic content (e.g. maps, figures) but the special case of rendering text is still controversial. From the classic research on Eye movements [2], we understand the characteristics of users when performing reading tasks. This knowledge serves as a guidance in designing a text layout user-interface where we addressed the "undershoot of a return sweep." Many researches on reading speed tend to focus on helping users read faster. In the paper [3], the researcher provides us methods to quantify both reading speed and comprehension. Reading speed is often measured in words per minute while reading comprehension is measured by test scores. As circular displays become more prevalent in our daily devices, and the main medium of information displayed is still text, different layouts will surely have the most impact on user's readability. The paper [4] introduces several methods to conduct experiments on text reading, which may help us in conducting our tests. Also, the paper compares the speed of text reading among various methods of visualization including vertical scrolling, dividing texts into several screens, horizontal scrolling and so on.

The vertical scrolling and dividing text into several screens are two of the most efficient reading methods. In the paper [5], the researcher proposes a dynamic text presentation on a small screen, Rapid Serial Visual Presentation (RSVP). The RSVP lets the text be dynamically shown on the display. The paper compares the reading speed and comprehension

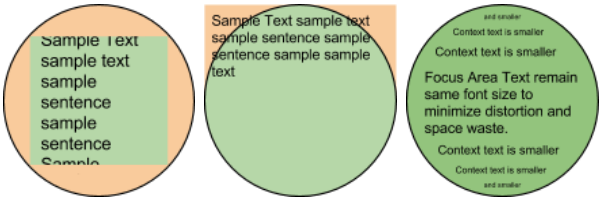
rate between RSVP and traditional text presentation. The paper demonstrated that the reading speed for short texts gets increased by RSVP formats by about 33%, and no sacrifice in comprehension or workload. In contrast, for long texts, RSVP significantly increases the workload while there is no improvement in either reading speed or comprehension. Although there are efficient adaptive algorithms which can be used for decreasing the workload, there is no reason to use RSVP if traditional text presentation can be used efficiently. With the increasing popularity of wearable devices, fitting more content to the limited screen size of wearable devices are becoming more important. In the paper [6], they introduce several document rewriting methods.

We applied some of those methods in our experiment, which helped us a lot in shortening the articles. The paper claimed those rewriting methods not only reduced the amount of text up to 13.3%, also improved the readability of documents. The number of people who choose to read on mobile devices is growing. Also, traditional keyboard and mouse interactions are slowly being replaced by small, touch-sensitive reading devices. In the research paper [7], the authors categorize readers into three types. They are full screen, linewise and blockwise. The experiment they did takes gaze into account, and puts it into relation with the displayed content. According to the experiment result, we decided to control our participants' reading behavior by asking them to read line by line without any skips.

Question and Hypothesis

A commonly occurring problem is the rectangle text flow getting chopped off on circular displays. Our project compared the following three scrollable text layouts presented below, and address the underlying issues that may affect user readability. In the experiment, we measured the readability by testing the time it takes for participants to locate specific words as well as the total reading time. When testing the reading time, we ask participants to do a detailed reading task.

For measuring the word locating time, participants need to quickly scan and locate keywords which answer questions from the reading materials. Compared to the overflow and cropped layout, the proposed adaptive layout should optimize users' readability. The adaptive layout should require less time and make reading more efficient as well as help users browse through the content more effortlessly. Layout is a general term to describe the aggregation of layout parameters. "Lines per page" and "Words per line" are two significant factors of one layout with a fixed screen size and a chosen font. In a circular display setup, finding the optimal traditional layouts is equivalent to finding a chopped rectangle within the bounding of the screen. "Lines per page"



A. Cropped B. Overflow C. Adaptive

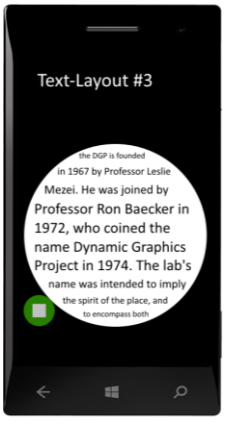
and "words per line" in the chopped rectangle could vary from height and width. No matter how the chopped rectangle gets chosen, there is a tradeoff between these two parameters. The overflow layout maximize the words per line while sacrificing the lines in focus, and the Overflow layout maximizes "words per page", product of words per line and Lines in focus. Ideally, the proposed adaptive layout should avoid this tradeoff, and maximum both of those two factors.

Methodology

In this section, we will describe the methodology we followed to test our research question and related hypothesis.

Apparatus

We implemented the whole apparatus in a smart phone with a study application. It would be better to run the experiment on a real device with a circular display such as Moto 360 but due to accessibility issues, the apparatus used in our study was a Nokia Lumia 920 with a simulated circular display. To simulate the display, experimental software was built to mock up the layout shown in the figure below. The participants could only interact with the circular area where they performed actions pertaining to the reading task. The outside area war blocked from user's interaction in order to prevent accidental input. All three layouts could be scrolled vertically but not horizontally. In the user study, we asked participant to interact in bi-manual style to simulate the real world usage pattern in wear devices.



Testing Application

Device Spec:
Nokia Lumia 920
Screen size:
4.5 inch (Diagonal)
Pixel density:
332 pixel per inch
Dimension:
130.3 x 70.8 x 10.7 mm
Processor:
Qualcomm Snapdragon™ S4
Dual-core 1.5 GHz
Simulated Display Spec:
3.2 inch (diameter size)
Circular display
332 pixel per inch

Participants

There were 20 native speaking participants and 20 non-native speaking participants. We had four layouts: overflow, cropped, paged adaptive and non-paged adaptive. Each group had 10 participants. Within each group, there are 5 native and 5 non-native speaker. The result analysis is based on 36 participants from the whole participants since we kick out the outliers from each groups. They performed some reading tasks on one article in one layout. All of them were volunteers with similar education level (BSc candidate or above), and we recruited them from class, graduate laboratory, and student lounges. All participants needed to have some experience in using smart phones and touch input. Therefore, they were capable of using the experiment software, and we did not had to spend too much time on training. The participants' age varied between 20 to 27. They should be healthy enough to sit for at least half an hour, and were able to concentrate on the experiment.

Experimental Design

The experiment has one independent variable which is the text layout. The experimental software manipulated the text layout directly. There are three levels for the independent variable since we have three different text layouts. The experiment followed the "between participants" design pattern to eliminate the significant learning curve of the reading task. If participants took part in all layouts, they would get familiar with the experiment and all the reading materials since the same reading materials were used for all the layouts. Also, this design reduced the time cost for each participant and the noise introduced by concentration differences. All participants were volunteers, so they probably did not want to spend too much time on the experiment. We think the result would be affected if participants got tired and annoyed.

Tasks and Procedures

Our study application followed the following procedures to measure the detail reading and locating time:

1. The researcher gives a brief introduction to the participant about the experiment.
2. The researcher asks the participant to sign a consent form.
3. The researcher collects the participant's basic information such as age, English expertise, smartphone experiences, and fills in the questionnaire.
4. The participant is asked to perform a baseline test by detail reading article #1 in the traditional rectangular layout.
5. The participant reads article #2 (with similar difficulty and length to material #1) in detail.
6. The participant gets asked for 3 questions based on the article #2. The participant needs to locate and tap on the specific word which answers the question (like a name or number shown in the article).
7. The participant rates the layout and gives feedback, if any.

READING MATERIAL

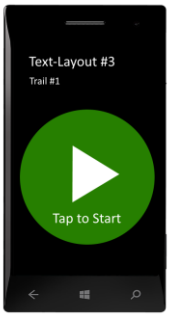
We uses two different material in the whole process of user study. The first one is for baseline testing (201 words, 1351 characters), and the second one for detail reading and search reading (237 words, 1436 characters). Both of the article is cropped from the Wikipedia article of Alan Turing [link], and have similar difficulty (in the meaning of words length, words level, sentence length, clause per sentence), see our supplemental material for the article itself.

READING TASK IN STEP 4 AND 5:

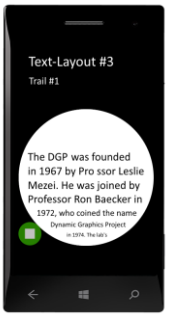
The participant clicks the start button on the screen, and then the reading task begins on the screen. At the same time, the timer starts in the background. The participant finishes the reading task by clicking the stop button on the software which stops the stopwatch.

LOCATING TASK IN STEP 6:

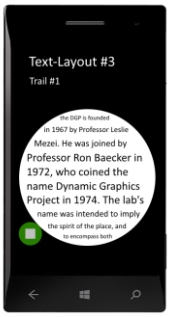
The participant reads the question first, and then taps on the start button to show the material, article #2. For each question, the circular screen restores back into the initial state which is the beginning of the article. The participant has to locate and tap on the specific answer keyword and then the timer stops.



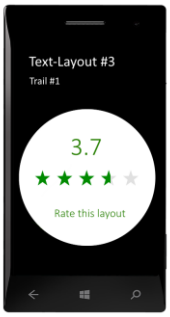
Welcome



Begin



Scroll & Reading



Rating

Measures

In our experimental design, there are two dependent variables which are the reading time and the word locating time. Both of them will be "normalized" by the baseline reading time to narrow the difference between individuals.

The reading time stands for the total time and effort spent on going through the entire reading material and understanding the article in a specific layout.

The word locating time is the time it takes for a participant to find a keyword which answers a question. The reading material should take approximately one minute. Its difficulty should be similar to that of an article on the daily news with the content being more general and avoiding scientific jargon so that we do not bring the participant's expertise into play. A more suitable layout on a circular screen should make reader spend less time on going through reading materials. Therefore, comparing the reading time among layouts can find us the best one. In addition to that, a better text layout tends to make searching easier for readers. Hence, the time spent on scanning and locating a specific word should be less on a better text layout. By testing the two dependent variables among the three layouts, we can understand how readability gets affected.

Data Collection

The experimental software was set up on a cell phone, Nokia Lumia 920. All participants involved in the reading task interacted with the cell phone and experimental software. Through the software itself, we collected the quantitative data automatically. The time for each task was recorded in the background silently without any notifications to the participants.

To measure the reading time, participants were asked to perform a detailed reading on the article instead of a quick scanning. For the detailed reading, we asked the participants to read the article line by line without skipping any words. The article appeared immediately after tapping the start button. To complete the task, the participants tapped the stop button shown above which prompted the software to record the time it took between tapping the start and stop button.

For measuring the word locating time, the participants were asked to scan the article as fast as they could to locate the specific words which answer the questions. Each participant was asked 3

questions based on the article. For each question, the participant scrolled up/down, scanned, and tapped on the keyword to signal his completion. For each question, there was only one keyword needed to be tapped on. The same reading interface was shown for this task, and the initial state was the beginning of the article. The word locating time also was measured by the software. In case of failure, a skip button would be shown after 30 seconds.

Data analysis and results

Data model and error distribution

We modeled the reading time in a particular layout in proportion to the reading time of other layout, (with a factor α_L to the "Standard Layout").

$$\begin{cases} T_L(A,I) = \alpha_L(\hat{T}_S(A,I) + N(0,\sigma_A)Length(A)) \\ \hat{T}_S(A,I) = \dfrac{Difficulty(A)Length(A)}{Profession(I)} \\ * \text{ A: Article, L: Layout, I: Individual} \end{cases}$$

Assumptions

The model could be derived from the foundational assumptions below:

1. Layout's affection is localized. (The time spent in reading "one page" for a particular layout does not statistically change among pages).
2. Participant's mental state is constant along the reading process.
3. The noise of reading time is an additive zero, meaning Gaussian random proportion to the article length.
4. The reading of the tested articles is not affected by the profession of an Individual.
5. The index of difficulty of baseline-testing article and detail reading article is identical. The noise distribution parameter (σ_A) is identical. (*this is a strong assumption based on study setup)

Please note that the layout factor α_L may be dependent on the testing article and individual in real world scenario. E.g. There may be different layouts suitable for articles with different difficulty and different structures. Also, there is no evidence that shows that the Index of difficulty and the individual's profession is independent of each other and contribute to the reading time by proportion relation; this is an assumption for simplifying the model.

Under these assumptions, we can derive

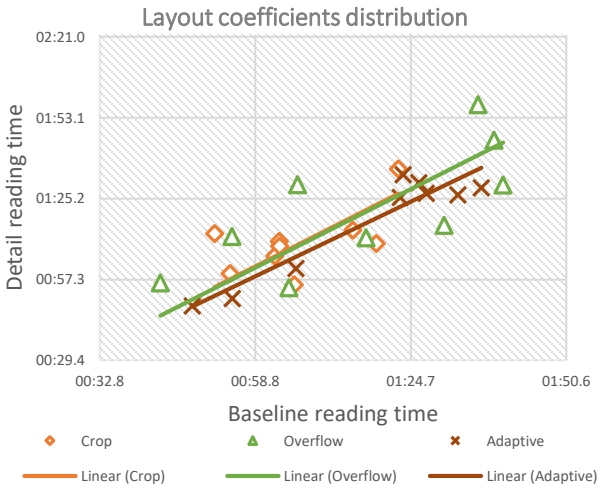
$$\begin{cases} \alpha'_L = \alpha_L \dfrac{Length(A_{test})}{Length(A_{base})} \\ T_L(A_{test},I) = \alpha'_L T_S(A_{base},I) \end{cases}$$

Where $T_L(A_{test},I)$ is the detail reading time for participant I (with Article A_{test} under layout L), and $T_S(A_{base},I)$ stands for the baseline testing time of participant.

Thus, we extracted the linear regression coefficient of detail reading to baseline reading test. Under the ***classic regression assumption***, the coefficient $\boldsymbol{\alpha'_L}$ **follows normal distribution**.

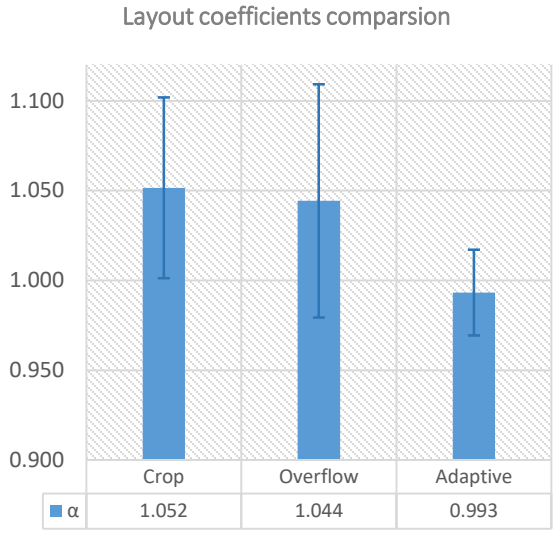
Since the factor of length index ratio stays constant across our study, we will directly use α'_L as the measurement of $\alpha_L$to compare between results to avoid extra assumptions on the function form of $Length(A)$ (e.g. The index of length depends on the word count and character count, and it's internal structure. Blindly using word count or characters does not make sense.)

Result for Layout coefficient



The data points fit into the linear model well, but more sample is needed for stronger support.

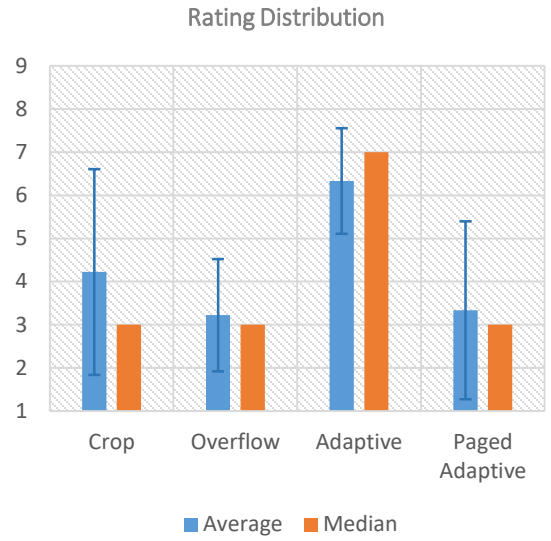
A note here is our **outlier policy**, since the least-square line regression problem is extremely sensitive to outliers, we manually exclude data points which gives more offset than **3 times of standard deviation** from computation. There is totally 36=9x4 data points fits into 4 layout groups.



As the layout coefficient (in reading time) is modeled by normal distribution, and the result **between groups is independent**, we conduct three pairs of **student's t-test** among our result to see the detailed difference brought by the layout. The **overall F-test** (or one way ANOVA) is **not** conducted because there is **no physical connection between levels** for that test. Any interpretation from this statistic result in this general F-test is **invalid**. (In analog, you cannot deduct there is no difference between student A and student B by reasoning from there is no difference between student B and student C.)

The **t-test** shows the result between Cropped-Overflow layout is not significant **(P=0.796),** while the difference between Overflow-Adaptive **(P=0.044)** and Cropped-Adaptive (**P=0.007)** is both significant in the 0.05 alpha level.

Result for User rating



The distribution of user's rating is usually comes with a significant variance, and the sample scale is not large enough, thus we also compute the median to avoid extreme value's affection. The t-test between groups show the same result in reading time; no difference for Cropped-Overflow **(P=0.291)**, significant difference for Overflow-Adaptive **(P=0.0004)** and Cropped-Adaptive **(P=0.036)**.

Discussion

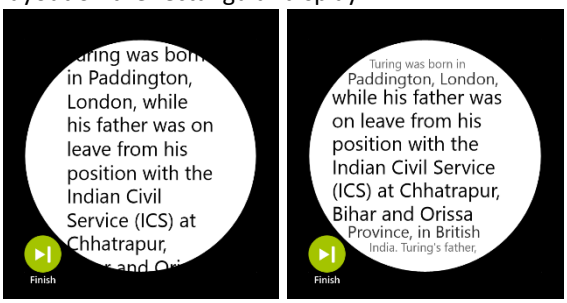
Result & hypothesis

Our study shows a quantitative result on the reading performance among different text layouts in circular display. Just like what we mentioned in our hypothesis, the adaptive layout provides both the best reading performance and user satisfaction. The main reason of this result is that the adaptive layout avoids the cropping distortion and maximizes screen space utilization. We will discuss the details in the following sections.

Detail reading

This result is just like what we predicted. The adaptive layout maximizes the screen space utilization by providing 10 lines in a single page and about 4-5 words per line, which outscores the cropped layout that provides 7 lines per page and 2-3 words per line. Fewer words per line causes users to perform line switching and scrolling much more frequently, and results in lower speed. In contrast, it is surprising that the adaptive layout make

participants perform very close to the traditional layout on the rectangular display.



Cropped Layout Adaptive Layout

Keyword Searching

The data in the keyword searching task are too sparse to perform a quantified study, but we can look at the trend of the data to do some analysis. The overflow layout has a better performance in keyword searching. It is expected since searching a keyword is kind of context independent. Participants perform this task by just scanning the article. The overflow layout provides the maximum number of words in a single page, which lets participants see more word during scrolling. Also, the lines do not get shrunk at the top and bottom, so it is easier to tap on the keyword. With maximized words per line, user also gets the highest “Scroll Gains” in the overflow layout, which makes them cover the article quicker.

Rating

The result from our rating data shows a significant margin between the adaptive layout and overflow/cropped layout. The adaptive layout provides the best experience., and the cropped layout has slightly advantage over the overflow layout regardless their similar performance.

Finger problem

Participants' fingers covered the underlying content on the screen, which slowed down the reading speed. We found that lots of participants used the unused blank edge in the cropped layout as a scrolling bar to avoid covering reading materials. The fat finger problem was more serious on a smaller screen. Tapping on the small screen became even harder.

Distortion effect

It is expected that users were hesitant to new experiences that were not commonly used. The resizing effect in the adaptive layout gave some users motion-sickness. One participant in our pilot study reported this issue. After we adjusted the parameters, we got less complaints in the large scale study. Another participant in pilot study considered the distortion effect in animation as a distracting factor for her although her performance was not bad. It is notable that we did not add an extra training section to the new adaptive layout. It is most likely that users’ performance would get better if they got some training on the layout in advance.

Limitation & future Work

The most significant limitation of our study was the hardware we used. Since we could not access the actual watches, we had to mock up the circular displays on a cell phone. The cell phone was still different in weight and other factors. The other limitation was the article topic. Different participants shared diverse interests, which was hard to control. The interest also affected the reading speed a lot.

Conclusion

The difference is not significant when comparing all the three layouts, however, since we applied the "Between participants" design, we can just compare any two layouts separately. Overflow vs. Adaptive (P=0.044) and Cropped vs. Adaptive (P=0.007) are both significant. By analyzing the current data, the adaptive layout provides both the best reading performance and user satisfaction. This result is just like what we predicted before. However, there are still some future works that need to be done. The first one is that we need to keep collecting data in order to make our research more precise. Another notable future work is to do quantitative analysis on “lines per page" and "words per line" respectively. These two terms, as we discussed in the research question and hypothesis section, are two main factors in the layout. The reason why the adaptive layout provided the best performance is that the layout maximize both “lines per page" and "words per line".

References

1. Carpendale, M. Sheelagh T., and Catherine Montagnese. A framework for unifying presentation space. In *Proceedings of the 14th annual ACM symposium on User interface software and technology*, UIST '01. ACM, 2001, 61-70.

2. Keith Rayner, Barbara J. Juhasz and Alexander Pollatsek. Eye Movements During Reading. *The science of reading: A handbook*. Chapter 5. 79–97. 2005.
3. Timothy Bell. Extensive Reading: Speed and Comprehension. *The Reading Matrix*, Vol. 1, No. 1, April 2001.
4. Tsvetozar Georgiev, Investigation of the user's text reading speed on mobile devices. In *Proceedings of the 13th International Conference on Computer Systems and Technologies*, CompSysTech '12, ACM, 2012, 329-336.
5. Gustav Oquist, Mikael Goldstein. Towards an improved readability on mobile devices: evaluating adaptive rapid serial visual presentation. *Interacting with Computers,* Volume 15, Issue 4, 1 August 2003.
6. Johannes Leveling, Gareth J. F. Jones. Making Results Fit Into 40 Characters: A Study in Document Rewriting. In *Proceeding SIGIR '12 Proceedings of the 35th international ACM SIGIR conference*, ACM New York, NY, USA, 2012, 1107-1108
7. Ralf Biedert, Georg Buscher, Arman Vartan. Reading and Estimating Gaze on Smart Phones. In *Proceeding ETRA '12 Proceedings of the Symposium on Eye Tracking Research and Applications*, New York, NY, USA, 2012, 385-388.