Toward a Systematic Understanding of Children's Touchscreen Gestures

Alex Shaw
Dept of CISE
University of Florida
Gainesville, FL 32611, USA
alexshaw@ufl.edu

Lisa Anthony
Dept of CISE
University of Florida
Gainesville, FL 32611, USA
lanthony@cise.ufl.edu

CHI'16 Extended Abstracts, May 07-12, 2016, San Jose, CA, USA
ACM 978-1-4503-4082-3/16/05.
http://dx.doi.org/10.1145/2851581.2892425

Abstract

This paper presents ongoing work on characterizing children's touchscreen surface gesture interactions. We report a preliminary analysis of gestures elicited from children ages 5 to 10. We focus specifically on the differences among gestures made by children of various ages. Continuation of this project will consist of further in-depth analysis and a systematic characterization of these gestures. This work will help create a deeper understanding of the way children make gestures and, in turn, motivate new ways in which we can design better touchscreen interactions and gesture recognition algorithms for children.

Author Keywords

Children; touchscreen; gesture interaction; recognition.

ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

General Terms

Design, Human Factors.

Introduction

Although an increasingly large number of children regularly use touchscreen devices [9,14], applications for these devices are rarely tailored toward children's

Figure 1. The 20 gestures used in the study, developed by Anthony et al. [1].

Figure 2. Screenshots of the apps used to collect the gestures [18].

unique interaction patterns [7,8]. In particular, children's gestures are quite different from those of adults [8]. For example, children tend to use more strokes to make a gesture [8]. Current algorithms are not able to recognize children's gestures as well as adults' [7,8,18]. Read et al. showed that children consider a recognition rate of 91% (in the domain of handwriting input) to be satisfactory [12], yet Anthony et al. showed that current recognizers only achieve 84% accuracy on 7- to 10-year-old children's gestures [3]. More work is needed in this area to develop better gesture recognition algorithms for children.

To achieve better recognition rates for children, we must first gain a deeper understanding of the ways they make gestures and how their gestures differ from those made by adults. Prior work investigating qualitative and quantitative features of children's gestures has been limited. Anthony et al. examined features of children's gestures, but focused on how they were affected by the presence of visual feedback [2]. Several gesture visualization and analysis tools, such as the Gesture Heatmaps Toolkit (GHoST) [17] and the Gesture Clustering Toolkit (GeCKO) [4], have been used to show interesting and useful patterns in adult gestures. By using these tools, as well as others we develop, we can systematically examine gestures elicited from children to find patterns that can help researchers design better recognizers. This paper offers a first step in ongoing work to address this topic. Though the gestures used in this work were collected in the study reported previously [18], the results presented in this paper are distinct in that we focus on examining the *features* of the children's gestures (instead of just overall recognition rates) to create a deeper understanding of the way kids make gestures.

Study Design

The dataset we are examining was collected from a study designed to explore the effect of interface complexity on children's touchscreen interactions [18]. We provide a brief description of the study design and data collection procedures, but more information can be found in the full paper [18].

Participants

The participants were students at a local elementary school. A total of 30 children participated, including three 5-year-olds, six 6-year-olds, four 7-year-olds, seven 8-year-olds, five 9-year-olds, and five 10-year-olds. The 5- to 10-year-old range was selected because of the important physical and psychological development which children undergo during this time [11]. We also collected an analogous gesture dataset from 30 adults over age 18.

Gesture Set

The gestures in this study are identical to those used by Anthony et al. [1], which were chosen based on a survey of gestures commonly used by adults as well as an investigation of psychological literature and handwriting apps for kids [5]. They include letters, numbers, and simple shapes (Figure 1).

Data Collection Apps

To collect the gestures used in this study, we used two Android apps, shown in Figure 2 [18]. In one of these applications, referred to as the abstract app, the user is asked to draw a series of gestures on a white canvas. There is feedback showing the strokes made, and the user cannot erase. In the other application, designed to be a more visually complex interface, an animated bird asks the user to draw the gesture on a canvas with a

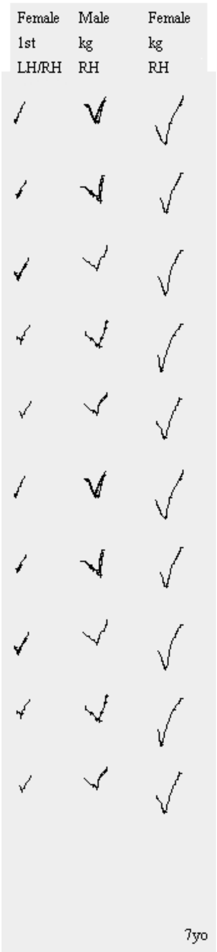

Figure 3. Drawings of checkmark gestures from three 7-year-old participants.

forest scene. The more complex app also contains animations which are not found in the abstract app. These apps were designed to examine the effect of additional interface complexity on children's gesture interactions, the results of which we describe in detail in another paper [18]. In both apps, the user was given points for each gesture to motivate and encourage completion, as recommended by Brewer et al. [6].

In total, we aimed to collect 7,200 gestures (2 apps x 30 participants x 20 gesture types x 6 repetitions per gesture type). However, six participants did not completely finish the study. Because some of our cross-user analyses could be affected by imbalance in number of gestures per participant, the gestures of these six participants were removed, leaving 5,760 gestures for us to analyze.

Analyzing the Gestures

To examine the gestures collected, we have begun by using several analytical tools, including a simple plotter to draw each gesture (which we have created for this work) and the Gesture Heatmaps Toolkit (GHoST) [17].

Plotting the Gestures

Our first step in examining the gestures we collected was to reproduce the gestures visually from the log files, so we could see if any patterns or trends emerged. We also noted the age, grade level, handedness, and reported level of expertise of each participant. Figure 3 shows the result of plotting the checkmark gesture for three of the 7-year-old participants in the study. The first row shows the gender of the participant, the second shows the grade level (kg = kindergarten), and the third shows the participant's preferred writing hand (LH = left hand, RH = right hand). The gestures in each column are plotted in the order in which they were made, top to bottom, so that we can examine each participant's consistency over time and with repeated practice.

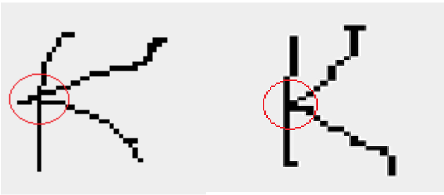

Figure 4. A drawing of the "K" gesture by a 6-year-old (left) and a 10-year-old (right). The red circle shows the point of interest.

Some observations from plotting the gestures include:

- Younger children (ages 5-7) scratched out gestures more frequently than older children (ages 8-10).
- Younger children had more stray marks (that is, marks that are not intended to be part of the gesture) than did older children.
- Younger children tended to have more difficulty joining lines together than older children. Figure 4 shows an example of this type of behavior.

It should be noted that we did not remove gestures that were scratched out or illegible, as we felt doing so would not truly capture the interaction patterns exhibited by children when making gestures.

Plotting the gestures allows us to gain some idea of general trends in the data. We will continue to analyze the gesture dataset for other trends. In future work, we will test the validity of our observations by calculating some existing metrics on the gestures such as the gesture articulation features from Anthony et al. [4]

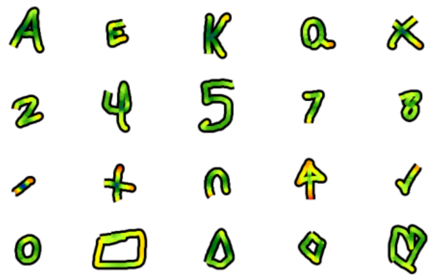

Figure 5. Using GHoST to visualize 5- to 7-year-olds' gestures. The yellow and red indicate lower consistency.

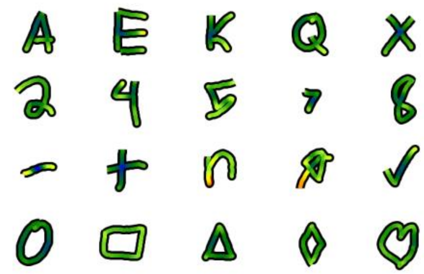

Figure 6. Using GHoST to visualize 8- to 10-year-olds' gestures. The increased prevalence of green indicates a higher level of consistency in many of their gestures.

and others from existing literature [10,13,15,19], as well as new ones which we will develop.

Applying GHoST

GHoST provides a visualization of how a number of articulation features vary over a set of gestures in the form of a heatmap [17]. Previous work has shown GHoST is an effective tool for understanding adults' gestures [17], and in our work we will use heatmaps to compare children's gestures to those of adults. Figures 5 and 6 show the result of using GHoST to visualize the shape error of the gestures of the 5- to 7-year-old and 8- to 10-year-old participants in our study, respectively. Shape error is the average absolute deviation of a set of gestures from a reference gesture (in our case the reference gesture is one of the user's gestures chosen at random) [15]. Studying the gesture heatmaps of the children's shape error further supports the hypothesis that children have trouble at points where lines join. The "arrow" gesture in Figure 5 is a good example: the top point, where lines meet, is red. The "plus" and "rectangle" are also good exemplars of the lower consistency seen in younger children's gestures, as evidenced by the increase in yellow and red in their heatmaps compared to older children.

Recognition Rates

A major motivating factor of this work is to distinguish the characteristics of children's gestures to help explain why conventional recognizers perform so poorly on children's gestures, and how recognition rates can be improved. We performed a number of recognition experiments using the gesture data we collected for this study [18]. We used the $P recognition algorithm [16], which we selected due to its popularity, versatility, and ease of implementation. The results confirmed previous work that gesture recognition rates show a significant positive correlation with age [1–3], and extended this finding to even younger children. Gestures by 5-year-olds had a recognition rate of just over 60% with $P, and those by 10-year-olds had a rate of just over 90%. Full details of those experiments can be found in the associated paper [18]. To further explore the recognition results here, we analyzed the output by creating confusion matrices for each age.

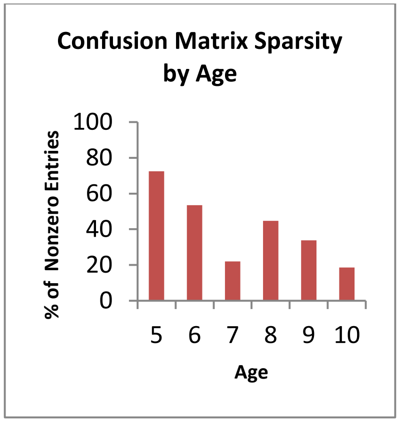

Figure 8. Younger children tended to have denser confusion matrices than older children.

A confusion matrix has one row and one column per gesture. The values in the cells represent the percentage of times the gesture in the row label was recognized as the one in the column label, including true matches. Figure 7 shows a partial confusion matrix for the 5-year-olds in our study.

The most interesting discovery in studying the confusion matrices so far is the variation in sparsity across the age groups. In the matrices, a nonzero value in a given cell indicates that the given input and output pair were seen at least once. Younger children tended to have very dense matrices, with relatively few nonzero entries, while older children had sparser matrices, with mostly zeros. This pattern indicates that $P tends to confuse a smaller number of pairs of gestures across gestures for older children. This trend (Figure 8) indicates that the errors $P makes on younger children's gestures are less predictable.

Our future work will systematically investigate the differences in consistency across age groups in within-user and between-user tests.

Conclusions and Results

The results of this work show that there are clear differences in the way older and younger children make gestures, even between the ages of 5 to 10 years.

	2	4	5	7	8
2	0.58	0.03	0.02	0.01	0.04
4	0	0.66	0.01	0.01	0.04
5	0.04	0.02	0.62	0.02	0.06
7	0.01	0.03	0.02	0.6	0.03
8	0.01	0.01	0.06	0.03	0.73

Figure 7. Partial confusion matrix showing % of tests for 5-year-olds. The highlighted cells represent correct recognitions.

Younger children are less consistent than older children in their gestures, and there are some common patterns of behavior, such as scratching out gestures, which are more prevalent in younger children. Our first look into applying GHoST on children's gestures led to several important observations. Children have higher degrees of error throughout the articulation of their gestures, and they struggle at points where lines meet. Our work shows the value of applying GHoST to children's gestures, and points to potential benefits of applying analytic tools to them.

Our work also helps demonstrate the importance of examining children's gestures across different ages, rather than lumping the ages together to compare against adults, as in prior work [2,3]. There are clearly differences in the ways kids of different ages make gestures, and future work should take this into account.

Future Work

Our study paves the way for future work in a number of areas, including development of new metrics and more accurate gesture recognizers for young children.

Examining Other Demographic Information

In most of the work on children's gestures done until this point, the primary independent variable of concern has been age [1–3,7,8,18]. In Figure 3, we showed demographic information when plotting the gestures, but we so far have not examined how this information is related to gesture recognition rates or other metrics. This analysis could be a rich source of data to help us understand how children of different backgrounds make gestures, allowing the creation of better touchscreen applications for diverse children. In particular, examining recognition accuracy in children by age group may not be the best way to study the data. Children of the same age may not always be in the same grade, and may have different levels of touchscreen experience. It is reasonable to expect that two children of the same age but in different grades may not have the same motor skills, and as such the child in the lower grade may not be able to make gestures as well as the child in the higher grade. Further work is needed to examine and understand the effect of these demographics on children's gestures.

Calculating Existing Metrics & Developing New Metrics
As discussed previously, there are a variety of metrics that have been proposed for describing surface gestures. However, few of these metrics have been systematically calculated on large sets of gestures elicited from children. A primary focus of our continuing work will be calculating these metrics on our dataset and examining whether they are affected by age or other demographic factors.

That said, based on our preliminary examinations of the dataset, it is clear to us that we will need to develop new metrics to help quantify the differences that exist among the gestures from various age groups. For example, one metric worth developing might examine the stray marking rate of a gesture (e.g., what percentage of a gesture sample is actually stray markings based on some locality threshold). Another potential metric might be to measure the degree to which strokes fail to meet (such as the gesture on the left in Figure 4, which should have a high value for this metric). Future work will examine new potential metrics which could help analyze children's gestures.

Designing Better Recognizers for Children
As discussed earlier, low recognition rates for children's gestures is one of the main motivations for this work. To be able to achieve higher recognition rates for children's gestures, we must first better understand the ways in which they make gestures. New recognition algorithms may build heavily on some of the new metrics discussed in the previous section.

Another possibility for improving recognition rates for children's gestures is to design adaptive algorithms that adjust based on the user's behavior. For example, if a user tends to make jagged lines instead of straight ones, the recognizer could realize this and adjust to it. An adaptive recognizer could have a number of benefits in educational applications [10], as it could provide personalized help for children when learning to draw certain shapes or gestures. A more personalized recognizer could also be used to design better and more responsive applications in a variety of contexts for touchscreen devices for children. Our work explores several ways in which we could potentially gain a deeper understanding of children's touchscreen gestures to help achieve these goals.

Acknowledgments
This work is partially supported by National Science Foundation Grant Awards #IIS-1218395 / IIS-1433228 and IIS-1218664. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect these agencies' views.

References

1. Lisa Anthony, Quincy Brown, Jaye Nias, Berthel Tate, and Shreya Mohan. 2012. Interaction and Recognition Challenges in Interpreting Children's Touch and Gesture Input on Mobile Devices. *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ACM Press, 225–234. http://doi.org/10.1145/2396636.2396671

2. Lisa Anthony, Quincy Brown, Jaye Nias, and Berthel Tate. 2013. Examining the need for visual feedback during gesture interaction on mobile touchscreen devices for kids. *Proceedings of the International Conference on Interaction Design and Children*, ACM Press, 157–164. http://doi.org/10.1145/2485760.2485775

3. Lisa Anthony, Quincy Brown, Berthel Tate, Jaye Nias, Robin Brewer, and Germaine Irwin. 2014. Designing smarter touch-based interfaces for educational contexts. *Personal and Ubiquitous Computing* 18, 6, 1471–1483. http://doi.org/10.1007/s00779-013-0749-9

4. Lisa Anthony, Radu-Daniel Vatavu, and Jacob O Wobbrock. 2013. Understanding the Consistency of Users' Pen and Finger Stroke Gesture Articulation. *Proceedings of Graphics Interface*, Canadian Information Processing Society, 87–94.

5. K. Beery, N. Buktenica, and N. A. Beery. 2004. *The Beery-Buktenica Developmental Test of Visual-Motor Integration, 5th Edition*. Modern Curriculum Press, New Jersey.

6. Robin Brewer, Lisa Anthony, Quincy Brown, Germaine Irwin, Jaye Nias, and Berthel Tate. 2013. Using gamification to motivate children to complete empirical studies in lab environments. *Proceedings of the International Conference on Interaction Design and Children*, ACM Press, 388–391. http://doi.org/10.1145/2485760.2485816

7. Quincy Brown, Lisa Anthony, Jaye Nias, Berthel Tate, Robin Brewer, and Germaine Irwin. 2013. Towards Designing Adaptive Touch-Based Interfaces. *Proceedings of the ACM SIGCHI Mobile Accessibility Workshop*, 4pp.

8. Quincy Brown and Lisa Anthony. 2012. Toward comparing the touchscreen interaction patterns of kids and adults. *Proceedings of the SIGCHI Workshop on Educational Software, Interfaces and Technology*, 4pp.

9. Cynthia Chiong and Carly Shuler. 2010. *Learning: Is there an app for that? Investigations of young children's usage and learning with mobile devices and apps*. The Joan Ganz Cooney Center at Sesame Workshop, New York, NY. Retrieved from http://dmlcentral.net/resources/4496

10. Hong-hoe Kim, Paul Taele, Stephanie Valentine, Erin McTigue, and Tracy Hammond. 2013. KimCHI: a sketch-based developmental skill classifier to enhance pen-driven educational interfaces for children. *Proceedings of the International Symposium on Sketch-Based Interfaces and Modelin*, ACM Press, 33–42. http://doi.org/10.1145/2487381.2487389

11. Jean Piaget. 1983. *Piaget's Theory*. Wiley & Sons, New York, New York, USA.

12. Janet C Read, Stuart MacFarlane, and Chris Casey. 2003. "Good enough for what?": acceptance of handwriting recognition errors by child users. *Proceedings of the International Conference on*

Interaction Design and Children, ACM Press, 155. http://doi.org/10.1145/953536.953565

13. Tevfik Metin Sezgin and Randall Davis. 2005. HMM-based efficient sketch recognition. *Proceedings of the International Conference on Intelligent User Interfaces*, ACM Press, 281–283. http://doi.org/10.1145/1040830.1040899

14. Carly Shuler. 2009. *Pockets of Potential: Using Mobile Technologies to Promote Children's Learning*. Joan Ganz Cooney Center at Sesame Workshop, New York, NY. Retrieved from http://www.joanganzcooneycenter.org/publication/industry-brief-pockets-of-potential-using-mobile-technologies-to-promote-childrens-learning/

15. Radu-Daniel Vatavu, Lisa Anthony, and Jacob O Wobbrock. 2013. Relative accuracy measures for stroke gestures. *Proceedings of the ACM International Conference on Multimodal Interaction*, ACM Press, 279–286. http://doi.org/10.1145/2522848.2522875

16. Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2012. Gestures as point clouds: a $P recognizer for user interface prototypes. *Proceedings of the ACM International Conference on Multimodal Interaction*, ACM Press, 273–280. http://doi.org/10.1145/2388676.2388732

17. Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2014. Gesture Heatmaps: Understanding Gesture Performance with Colorful Visualizations. *Proceedings of International Conference on Multimodal Interaction*, ACM Press, 172–179. http://doi.org/10.1145/2663204.2663256

18. Julia Woodward, Alex Shaw, Annie Luc, Brittany Craig, Juthika Das, Phillip Hall Jr., Akshay Holla, Germaine Irwin, Danielle Sikich, Quincy Brown, and Lisa Anthony. 2016. Characterizing How Interface Complexity Affects Children's Touchscreen Interactions. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, ACM Press, to appear.

19. Bo Yu and Shijie Cai. 2003. A domain-independent system for sketch recognition. *Proceedings of the International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia*, 141–146. http://doi.org/10.1145/604471.604499