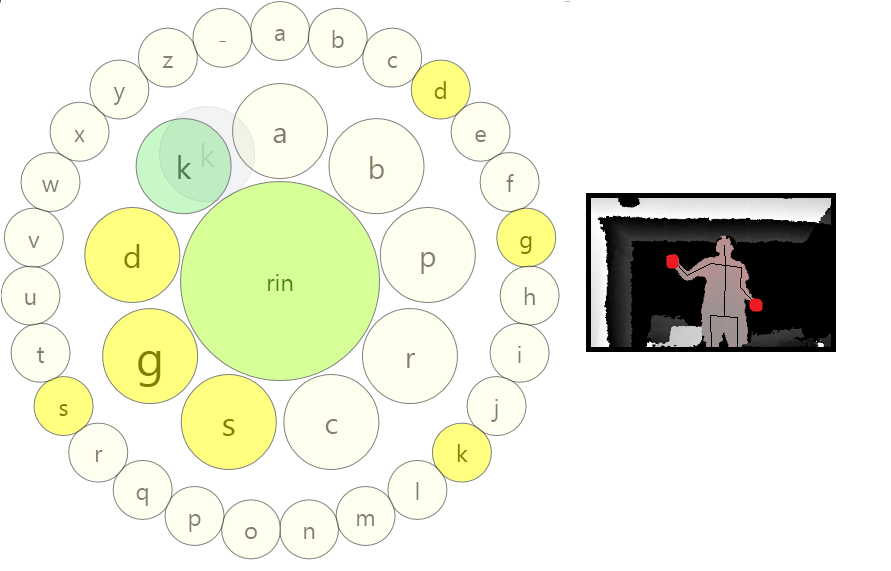
Bubble Keyboard: A Gesture-Based Mid-Air Text Entry Method Designed for Simple Selection and Increased Speed Utilizing Microsoft Kinect

Figure 1. A user typing the word ‘rink’ with the Bubble Keyboard. Basic Skeleton and Tracked Hands are displayed.



# ABSTRACT

## First Page Copyright Notice

Leave 3 cm (1.25 in.) of blank space for the copyright notice at the bottom of the left column of the first page. In this template a floating text box will automatically generate the required space. Note however that the text box is anchored to the **ABSTRACT** heading, so if that heading is deleted the text box will disappear as well.

This paper presents the Bubble Keyboard, a text entry system that utilizes the Microsoft Kinect for Gesture-Based Mid-Air Text Entry. This system incorporates gesture recognition with predictive analysis to simplify selection and increase overall speed and accuracy. In contrast to existing Mid-Air Text Entry systems in the industry, this system reduces the necessity for precise motion and thus improves accuracy as well as provides predictive assistance for simplified selection. This research demonstrates proof for the potential this technology has to improve further Mid-Air Text Entry Systems.

## Author Keywords

Bubble Keyboard, Text Entry, Mid-Air, Gesture, Kinect, Trigram, Trie.

## ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

# INTRODUCTION

Recent progression of Mid-Air Platforms such as Microsoft Kinect and Leap 3D has brought forth a need for text entry systems that function on these platforms. Currently these Text Entry programs are utilized for typing small phrases or keywords for search in applications from social media networking such as Twitter to entertainment platforms such as XBOX Live or Netflix as these applications are commonly used in a living room setting without an actual controller. Text Entry on these platforms currently employs basic gesture techniques for character selection. The general text entry structure used in the industry is best described by that of Evoluce, whose keyboard includes a display QWERTY keyboard and a pointer which is controlled by the position of your hand and a forward gesture to select letters. However, other non-standard keyboards have been created by members of the public, many of which can be seen on YouTube.

However, the systems currently used in the industry do not use the greater potential of gesture based analysis. Little research has been done on the potential of more gesture dependent text entry systems in contrast with more position dependent text entry systems. Also, a combination of prediction algorithms when used in conjunction with these more gesture dependent systems can further improve their potential. This research demonstrates the possible improvements that could be made for mid-air text entry systems if we utilized more gesture dependent text entry systems. The potential of more gesture dependent text entry systems could yield inputs with minimalized, faster, and simpler motion. Furthermore, current methods offer difficulty for users with fine motor control impairments as locating and maintaining a hand position is exceedingly difficult.

This paper presents the Bubble Keyboard, a new keyboard designed, by us, with the specific intention of using these gesture based systems and prediction algorithms to increase speed and simplify motion, in contrast to a model of the industry standard. The Bubble Keyboard consists of two concentric rings of letters equidistant from a center point where the outer ring is a static representation of the entire alphabet and the inner is a variable representation of the most probable letters. Selection of letters is done by passing the cursor, which is controlled by the user’s hand, over the desired bubble in a wave rather than placing the cursor over the letter and making a selection gesture as done in the industry model. Other functions, such as Backspace and Shift, are done through gestures rather than button selection. The letters displayed in the inner ring are calculated through a trigram algorithm composed with a Trie (or prefix tree) auto-completion algorithm. The end result being a fluid system that feels natural to the user and minimalizes motion and increases speed.

Our research demonstrates empirical results showing that the Bubble Keyboard produces text faster with a lower error rate than the conventional on-screen QWERTY keyboard. Users find using the Bubble Keyboard requires less effort and produces less frustration and rate their performance with the Bubble Keyboard as significantly better. Furthermore, users have an 18% improvement in adjusted WPM when accounting for error rate.

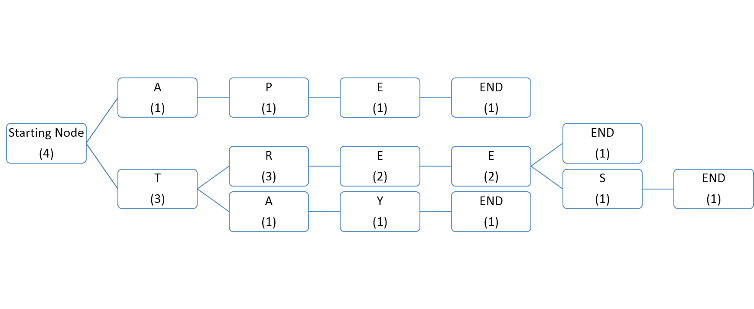
# Related Work

[FILL BY WOBBROCK

# Description

The Bubble Keyboard is a mid-air gesture based text entry method currently designed for use with the Microsoft Kinect, but can expanded to other platforms such as using a stylus, laser pointer, or cell phone. Letter selection is done by moving a virtual pointer, whose position is based on your dominant-hand position, over a letter-bubble and dragging back the holographic representation of the letter to the center of the keyboard. Keystrokes such as backspace and space are done with your non-dominant hand by gesturing swipe in one of the six specifically designated gestures (up, down, right, left, in, out). The Bubble Keyboard is a non-static keyboard which adapts to optimize performance based on previous words typed and currently typed letters in the word by placing more probable letters in the inner ring of the keyboard.

Figure 2. A Trie dictionary with the words ‘ape’, ‘tree’, ‘trees’, and ‘tray’ as well as the descendant *end* node count.

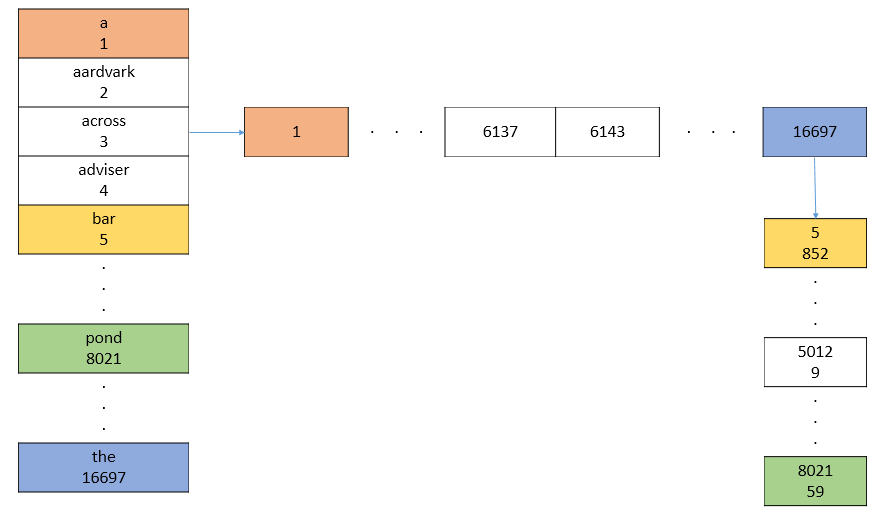


The principal difference between the Bubble Keyboard and that of the conventional on-screen keyboard is the method by which letters are selected. Rather than requiring the user to maintain their cursor above a button and creating a selection gesture, the Bubble Keyboard selection process works by calculating the hand’s velocity and selecting the letter where the cursor is when the velocity reaches near zero. The Bubble Keyboard is designed to not select anything unless the cursor is still. The user needs to simply drag the bubble back over the center bubble and pause; the dragging action can be seen in Figure 1.

Furthermore, the predictive aspect of the Bubble Keyboard simplifies the user’s difficulty in selecting a small target by suggesting probable letters in the center ring. Due to the nature of the English language and the spelling of words, it is unlikely that a user will have to select any more than two letters from the outside ring and typically those letters will be the preliminary two letters of the word. The advantage of which is minimalized motion and reduced search time.

The predicted letters are the result of two prediction processes. The first functions to ensure that probable letters are in the inner ring once the preliminary few characters have been typed and the second attempts to place the preliminary character in the inner ring. The first is an adaptation of a Trie dictionary such that at any iteration, the children characters of the current node representing the word thus typed are in the inner ring. If there are more children then spaces, the most probable ones are placed. The second is a Trigram prediction which places the initial characters of probable words in the inner ring based on the previous two words typed. This helps when the initial letter is an ‘f’ or ‘v’, for example, which would not normally be in the inner ring when only considering the Trie dictionary.

Figure 3. A Trigram dictionary demonstrating how the phrases “across the bar” and “across the pond” are stored. Similarly colored nodes represent equivalent elements.



The conventional on-screen keyboard has the visual search advantage over the Bubble Keyboard in that most users are accustomed to the layout of a QWERTY keyboard and know where the next letter will be. However, with the Bubble Keyboard, users have an idea of where a letter will appear in the outside due to the alphabetical distribution, but have little knowledge of the positioning of the letters in the inside ring. For that reason there are two features in place for reduces search time. Firstly, letters which lead to words in the dictionary are highlighted yellow in contrast to the white of the other letters so that they are visually caught. Secondly, the letters have a tendency to remain in the same position as they were in the last couple iterations of the keyboard if they were in the inner ring in either of the previous two iterations. As a user is searching for a letter they subconsciously remember the positions of the other letters in the inner ring. By attempting to preserve the general position (not all positions can be maintained), we simplify search time. Also, font size of the letters is variable and contingent on their probability of being selected; more probable bubbles have bigger fonts.

# Study

To determine the validity of our assertions and to generate a comparison of our keyboard to that of a model of the current industry standard, we conducted an exploratory user study which anonymously tested user performance in both settings for comparison of the two keyboards.

## Procedure: Subjects, Tasks, Apparatus, Measures

The user was placed in a situation similar to that of when a user would need to type in the name of a movie or phrase they want to search. The user was positioned in front of a television set and asked to type a series of phrases (all lowercase) of approximately 50 characters including spaces with both keyboards. Phrases and order varied from user to user but were all form a single bank of test sentences. Users were placed in one of two testing situations: a 1 hour testing session which involved 2 practice phrases and 8 test phrases with each of the two keyboards or a 1.5 hour testing session which involved 5 practice phrases and 15 test phrases with each of the two keyboards. Subjects would complete all testing with one keyboard and then continue to the next. We tested 3 users in the 1 hour setting and 9 in the 1.5 hour setting.

Figure 4. The display for the conventional on-screen keyboard used in comparison to the Bubble Keyboard.



The practice phrases were identical to the test phrases and were solely to give the user time to experience using the keyboard and understanding the sensitivity of the keyboard. Before the practice session, users were shown how to use the functions such as Backspace and Space in the bubble keyboard and also how to select letters in either keyboard. During the practice session, any questions about how to use the keyboard were answered. For testing purposes, subjects were asked to perfect their practice phrases with use of the backspace and to attempt to perfect their test phrases unless the mistake is too difficult or time consuming to adjust; it was left to their discretion.

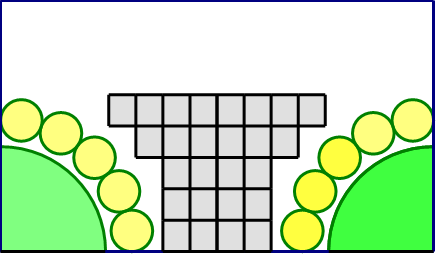
During the session, position data on the hands of the user were collected through the Kinect motion tracker. Following the testing, users were asked to fill out a NASA TLX form rating the difficulty of the various portions of the task.

## Results

# Future Work

The results of this study imply that the predictive nature of the Bubble Keyboard is a principal factor in its successes; we want to research further methods of positioning predicted letters in the inner ring that seem move natural and reduce user motion. Also experimentation with animation and variably sized bubbles based on probability to reduce search time are being tried. The idea of the Bubble Keyboard can be used in other platforms, not necessarily mid-air: Laser Pointers, Eye Tracking, and Cell Phone. We plan on continuing research in which we reduce the number of times the keyboard recognizes accidental involuntary movement by the user as a function. Further research is necessary in other environments as well such as movie titles when the trigram dictionary has the collection of movie titles built in and common names not built into the keyboard’s dictionary. Also further research is required to verify the results of this study across all sets of users, for example motor impaired users. Finally, we would like to generate models of natural mid-air motion as conducted by users to better create more natural feeling text entry systems.

Figure 5. An abstract design of the Bubble Keyboard adapted for a cell phone, with selection done on either side with either finger.



# Conclusion

Although mid-air gesture based text entry is a relatively new system of text entry, the principle of improving efficiency and reducing difficulty is an old research topic whose interest has been recently renewed. This research proposes the Bubble Keyboard as an alternate to the conventional on-screen QWERTY keyboard for this text entry platform. The Bubble Keyboard is a combination of gesture recognition for selection combined with predictive analysis where letter selection is conducted by dragging letters like placing tiles. Our conducted user study showed that users found the frustration and effort required considerably less than the conventional on-screen QWERTY keyboard as well as considered themselves more successful with the Bubble Keyboard. Further analyses of the study showed an improvement of 18% in Adjusted WPM when using the Bubble Keyboard over the conventional keyboard.

The Bubble Keyboard demonstrates that selection of letters does not require specific gestures or button presses, but rather can be completed through gesture analysis. Furthermore, the predictive features that reduces the total distance that the user needs to move, have been shown to be effective, rather than a nuisance. Thus, we have shown that a mid-air gesture based text entry system is more reliable when its dynamic nature assists the user in selecting letters and the selection of letters does not required specific gestures.

We hope that the Bubble Keyboard will be useful in more domains than just the one presented. Also, we hope to seem implementations of the Bubble Keyboard for text entry in the growing field of mid-air gesture based text entry.

# AckNOwLEDGMents

The authors thank the Information School at the University of Washington, Dr Lauren Bricker, Professor Richard Ladner, Alexander W. Glassford, Dr Madhura Nirkhe, and Dr Vivek Nirkhe.

# REFERENCES

1. Adobe Acrobat Reader 5. http://www.adobe.com/products/acrobat/.
2. Anderson, R.E. Social impacts of computing: Codes of professional ethics. *Social Science Computing Review 10*, 2 (1992), 453-469.
3. CHI Conference Publications Format. http://www.acm.org/sigchi/chipubform/.
4. How to Classify Works Using ACM’s Computing Classification System.  
   <http://www.acm.org/class/how_to_use.html>.
5. Klemmer, R.S., Thomsen, M., Phelps-Goodman, E., Lee, R. and Landay, J.A. Where do web sites come from? Capturing and interacting with design history. In *Proc. CHI 2002*, ACM Press (2002), 1-8.
6. Mather, B.D. Making up titles for conference papers. *Ext. Abstracts CHI 2000*, ACM Press (2000), 1-2.
7. Schwartz, M., and Task Force on Bias-Free Language. *Guidelines for Bias-Free Writing*. Indiana University Press, Bloomington, IN, USA, 1995.
8. Zellweger, P.T., Bouvin, N.O., Jehøj, H., and Mackinlay, J.D. Fluid Annotations in an Open World. *Proc. Hypertext 2001*, ACM Press (2001), 9-18.

**The columns on the last page should be of approximately equal length.**