ABSTRACT

Alternatives using the Leap Motion to extend Mid-Air Word-Gesture Keyboards

Garrett Benoit, M.S.

Faculty Adviser: Michael G. Poor, Ph.D.

Abstract goes here

Alternatives using the Leap Motion to extend Mid-Air Word-Gesture Keyboards

by

Garrett Benoit, B.S.

A Thesis

Approved by the Department of Computer Science

_____
Gregory D. Speegle, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree of

Master of Science

Approved by the Thesis Committee

_____
Michael G. Poor, Ph.D., Chairperson

_____
Bill Poucher, Ph.D.

_____
Brian Garner, Ph.D.

Accepted by the Graduate School
July 2015 (DATE TBD)

_____
J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I want to express my great appreciation to several people. Acknowledgements.

Pour la rose qui fait important le temps qu'on perd

CHAPTER ONE

Introduction — purpose for study, significance of study, original contribution

In the first chapter, clearly state what the purpose of the study is and explain the study's significance. The significance is addressed by discussing how the study adds to the theoretical body of knowledge in the field and the study's practical significance for communication professionals in the field being examined.

Ph.D. students also must explain how their research makes an original contribution to the body of knowledge in their discipline. They also should address the significance of the study for mass communication education.

It is especially critical that this chapter be well developed. Without a clearly defined purpose and strong theoretical grounding, the thesis or dissertation is fundamentally flawed from the outset.

1.1 Motivation

With the increase in gesture-controlled interfaces for touch screen and other modern devices, gesture-controls have started to see a transition for use in mid-air. Mid-air, gesture-controlled content has seen it's emergence in large displays [10], smart phones [4], augmented reality [12], and desktop computers [15]. Mid-air pointing has been a common approach to many of these gesture-controlled interactions and is used to select and manipulate on-screen objects [1, 2, 5, 14, 18]; however, means of reasonable mid-air text entry are fairly new. Past technologies allowed for mid-air text entry, but those approaches have fallen short of any meaningful text entry rates [7]. More modern approaches of mid-air text entry have seen improved results but still low, around 13 [8] to 18.9 [13] words per minute. These approaches were limited to selection of individual characters and also lacked the multi-tap feature of touch-based entry [8, 11, 13]. Last year, the largest improvement was seen in mid-air text

entry when Markussen et al. [9] transitioned word-gesture keyboards for mid-air use with the development of Vulture, reaching a text entry rate of 20.6 words per minute for their initial study. They achieved an even better text entry rate of 28.1 words per minute in a second study, with training and repeated measures, indicating learning the new techniques will help bring mid-air text entry closer to touch-based text entry. Vulture reached 59

The purpose of this study is to use the Leap Motion, a new and emerging technology [15], for interpreting mid-air gestural inputs for text entry [16]. The only previous attempt for text entry with the Leap Motion was in mid-air handwriting [17]; however, even regular hand-writing is slow, and confined around 15 words per minute [3]. Instead, this study aims to follow the path of Markussen et al. [9] and use the Leap Motion to extend word-gesture keyboards to mid-air text entry. Word-gesture keyboards have garnered popularity with the advent of smart phones and tablets and have been proven to perform well on touch screens [6, 19, 20]. This study intends to use the Leap Motion to find alternatives to mid-air, word-gesture keyboards and find a better approach than wearing a glove or detecting pinching [9, 11] for the mid-air equivalent of tapping and releasing for delimiters of words. This study will explore the option of using the extra degrees of freedom available in mid-air (e.g., depth) which was decided against by Markussen et al. [9] to create virtual keyboards in mid-air, and it will also use several techniques of simulating touch for the mid-air equivalent of tapping and releasing for delimiters of words. The goal is to make it feel as similar to using a touch-based device as possible while still allowing for common gestures.

The rationale behind this research is to improve mid-air text entry using techniques that will still allow for other gesture-controls when working with gesture-interfaces [1, 2, 5, 14, 18]. Touch-less gesture-controllers with mid-air text entry will benefit augmented reality (e.g., Google Glass, Microsoft HoloLens) as well as benefit

the medical world (e.g., operating rooms) when it comes to sanitation and sterile environments to reduce the spread of pathogens.

1.2 Definition

1.3 Gestures/Text Entry: Past, Present and Future — rename this and subsections (Related Works)

1.3.1 A Brief History

1.3.2 Current Trends

1.3.3 Future Plans

1.4 Problems, Challenges, Limitations — rename this section

1.5 Solution Scope — rename

1. Explore alternatives using the Leap Motion for mid-air, word-gesture keyboards to find a better approach than pinching for the mid-air equivalent of tapping and releasing for delimiters of words and increasing text entry rates [9].

2. Explore alternatives using the Leap Motion to touch-based, word-gesture keyboards. Alternatives, without training or repeated measures, are not expected to surpass touch-based, word-gesture keyboards in text entry rates; however, some alternatives are expected to be similar in error rates, precision, and usability.

3. Look for evidence of a correlation between having previous word-gesture keyboard experience and the text entry rates, error rates, precision, and usability of the various word-gesture keyboards.

CHAPTER TWO

Literature Review — related works, describe and analyize previous works, what is known/not known, leads to hypotheses

The purpose of the study should suggest some theoretical framework to be explained further in this chapter. The literature review thus describes and analyzes previous research on the topic.

This chapter, however, should not merely string together what other researchers have found. Rather, you should discuss and analyze the body of knowledge with the ultimate goal of determining what is known and is not known about the topic. This determination leads to your research questions and/or hypotheses. In some cases, of course, you may determine that replicating previous research is needed.

2.1   3 dimensions beyond desktop — rename

2.2   Interaction

2.2.1   Gesture-based Interaction — similar to introduction materials

2.2.1.1   Mid-Air Gestures.

CHAPTER THREE

Methodology — describes/justifies data gathering methods, how data is analyzed

Pilot 1 – Changes for study: Fixed: - Constructed dictionary so inputs will all type very similar words for best comparision - Simplified keyboard used (removed enter/space/period/shift/comma/numbers, changed location of backspace) - keyboard mod - disallow backspace of correct letters - keyboard mod - Removed having to press a key for confirmation of word (now confirmation on "release") - extended keyboard view for better view of 3 dimensions for users - limit interaction space to onscreen keyboard (works more like a phone keyboard now) - Fix calibration order - New Keyboard Types - Bimodal, Dynamic - Reduced number of experiment words from 30 to 10 per input Attempted but Unfixed: - slight corrections to help with pinch - Attempted Augmented Reality with Meta Glass, could not get rid of blurriness with glasses, so not implemented Observed: - sphere problem with touching plane - light matters

Pilot 2 – Changes for study: Fixed: - Narrowed objsectives/statistical analysis data - Change to intermittent exit surveys - UP IN THE AIR - remove stylus for mid air keyboards and use finger instead - UP IN THE AIR - change controller to WGK?? - UP IN THE AIR - change dictionary builder from custom algorithm to frechet distance - not doing - UP IN THE AIR - don't force person to be correct?? — No need to modify (justification of why I did this) - Made calibration faster (different depending on computer) - Had to remove lamement surface and only use paper, "bubbling" was interferring with detection Observed: - leap surface, utensil has to be perpendicular to surface for it to work properly - sphere problem with touching plane - light matters

This chapter describes and justifies the data gathering method used. This chapter also outlines how you analyzed your data.

Begin by describing the method you chose and why this method was the most appropriate. In doing so, you should cite reference literature about the method.

Next, detail every step of the data gathering and analysis process. Although this section varies depending on method and analysis technique chosen, many of the following areas typically are addressed:

–description of research design internal validity external validity

–description of population and description of and justification for type of sample used or method for selecting units of observation

–development of instrument or method for making observations (e.g., question guide, categories for content analysis) pre-test reliability and validity of instrument or method

–administration of instrument or method for making observations (e.g., interviews, observation, content analysis)

–coding of data

–description of data analysis statistical analysis and tests performed identification of themes/categories (qualitative or historical research)

3.1   Study Design === Task???

— This should really be a description of the exact task performed (the number of trials, the number of words, the number of input devices) and what they were trying to do, the procedure doesn't include this, it just talks about the experiment ordering as a whole.

The task performed by the subjects consisted of 10 trials for each of the 7 keyboard input devices for a total of 70 trials per subject. In each trial, a predetermined word appears on the screen. Subjects were required to correctly type in the displayed word with the active keyboard, using the active keyboard's backspace key to remove errors. Correct letters in a word are protected from the backspace key

3.1.1   Pilot Study 1

3.1.2   Pilot Study 2

3.1.3   Final Study

3.2   Study Setup === Input Devices and Interaction

Computer Used, Leap, Tech etc, the table touch screen.

3.3   Participants

How many subjects we got for each phase, some identifiers, male/female/hours used/experience etc. Describe how all subjects participated in every keyboard and all used the same dictionary to more accurartely depict each of the keyboard inputs.

A sample size of 14 was used in this study. The justification for this sample size comes from the formula to calculate the sample size for two independent group means using a pooled standard deviation:

$$N = \frac{2(z_{\frac{\alpha}{2}} + z_{1-\beta})^2}{(\frac{\mu_1 - \mu_2}{\sigma_{pooled}})^2} \tag{3.1}$$

Power of $1 - \beta = 0.80$ and a significance level of $\alpha = 0.05$ were used when calculating the sample size. See Appendix G to see what variables and keyboard comparisons were used in the sample size calculation. The derived sample size was the average sample size for all relevant variable comparisons based on the study objectives. Outliers requiring a sample size greater than 100 were removed. Furthermore, a sample size of 14 justifies the Latin Squares design for 7 input methods. The Latin Squares design was chosen for counterbalancing the experimental design and to reduce the effect of participation in one condition affecting performance of other conditions. Further detaisl are explained in Section 3.5, *Experimental Design*.

3.3.1 Pilot Study 1

3.3.2 Pilot Study 2

3.4 Procedure

The process of how we went through the steps (this is the same procedure as in the protocol, so this shoudl be an easy copy and paste). Just make it more beefy and change the vocabulary to reflect the trails/tasks.

There will only be a single study visit for each subject in which all experiments will be performed and will end with an exit survey. The study visit will take no more than 60 minutes to complete. The subject will be asked to do the following procedures:

- Complete a set of tasks on the computer for each of the 7 virtual keyboard inputs. These tasks, for all of the virtual keyboard inputs, are expected to take an upward bound of 54 minutes to complete. For each of the 7 virtual keyboard inputs, the subject will complete the following steps:
  * Given a brief explanation of the current input. This explanation will take a total of about 30 seconds. The dialog will be similar to: This is the ABC keyboard (e.g., standard, controller, leap-air, leap-surface, etc.). It is a JKL keyboard (e.g., mid-air, controller-based, or touch-based) and you will use XYZ (e.g., stylus, hand, or controller) to interact with it. The subject will then be given the interaction object and allowed to interact with the virtual keyboard input.
  * The subject will be given multiple opportunities to optionally recalibrate the keyboard interaction-space (if applicable to the current input) as many times as needed. Some inputs do not have an interaction-space and therefore do not require calibration of any kind. This task will take a total of about 10 seconds for each calibration and is expected to take no more than 6 calibrations (1 minute) for each applicable input.

* The subject will then be given the opportunity to use each virtual keyboard input to type in a variety of practice words. Practice words are randomly selected from a large dictionary but filtered to remove swear words and words used in the experiments. The subject will be able to attempt as many practice words as needed until they feel comfortable with the current virtual keyboard input. At any time during this phase, the participant can opt to recalibrate the interaction-space if applicable. This task is expected to take a maximum time of 3 minutes.

* Next, the subject will be going through the experiment. They will type in a total of 10 words for the current virtual keyboard input. These words are preselected words for each input before the experiment begins and have been selected based on a similarity calculation and filtered for swear words. This task is expected to take a maximum time of 3 minutes.

* Finally, there will be a small survey section after using the current keyboard input to rate each one on the Likert scale relating to difficulty, discomfort and fatigue experienced when using the devices. This task is expected to take a maximum time of 30 seconds. See Appendix D for the keyboard input survey example.

- After all experiments are completed for each input device, the subject will be asked to fill out an exit survey. This exit survey will obtain basic data such as age, gender, major, and handedness as well as several questions about any prior experience or impairments that might relate to the study. Finally, the exit survey will have a section to rank each device on a numerical scale. This exit survey will take a total of about 5 minutes. See Appendix E for the full exit survey.

Table 3.1. Schedule of Assessments for a single study visit (in minutes).

	Controller	Touch Screen	Leap-Surface	Leap-Air Static	Leap-Air Dynamic	Leap-Air Pinch	Leap-Air Bimodal	**total**
explain	.5	.5	.5	.5	.5	.5	.5	**3.5**
calibrate	0	0	1	1	1	1	1	**5**
practice	3	3	3	3	3	3	3	**21**
task	3	3	3	3	3	3	3	**21**
survey	.5	.5	.5	.5	.5	.5	.5	**3.5**
total	**7**	**7**	**8**	**8**	**8**	**8**	**8**	**54**

3.4.1 Pilot Study 1

The number of words that were used in the initial pilot study were 30 words per keyboard input totalling 210 trials per subject. This lead to complaints about fatigue, one subject completely giving up on the Leap-Air Static keyboard input. The number of words per keyboard input was reduced to 10 words.

3.4.2 Pilot Study 2

During the second pilot study subjects were asked to fill out the Likert scale, keyboard evaluations for discomfort, ease of use and fatigue at the end of the experiment during the final exit survey. Subjects were having trouble recalling which keyboard was which so this prompted a change to the procedure to include intermittent keyboard evaluation surveys after each keyboard input task was completed rather than during the final exit survey.

3.5 Experimental Design

TODO: get reference for latin squares design, Alivin has a good one in 4.4.6 Design section. Also get a reference for within-subjects design so that we can have a reference for these statements in the next paragraph.

A Within-Subjects design was used for this study. The strength of a Within-Subjects design is that the overall power will increase and there will be a reduction in error variance associated with individual differences. The weakness of using the Within-Subjects design is that it suffers from carryover effects between each keyboard

input device. The participation in one condition may affect performance in other conditions. To account for this weakness, the study was supplemented with a Latin Squares design for counterbalancing. Table 3.2 and Table 3.3 show how the Replicated Latin Squares design was utilized for 7 different keyboard inputs and a sample size of 14.

Table 3.2. First rep of a Replicated Latin Squares design for 14 subjects and 7 conditions.

subjects	Rep 1 conditions						
1	A	B	C	D	E	F	G
2	B	C	D	E	F	G	A
3	C	D	E	F	G	A	B
4	D	E	F	G	A	B	C
5	E	F	G	A	B	C	D
6	F	G	A	B	C	D	E
7	G	A	B	C	D	E	F

Table 3.3. Second rep of a Replicated Latin Squares design for 14 subjects and 7 conditions.

subjects	Rep 2 conditions						
8	G	A	B	C	D	E	F
9	A	B	C	D	E	F	G
10	B	C	D	E	F	G	A
11	C	D	E	F	G	A	B
12	D	E	F	G	A	B	C
13	E	F	G	A	B	C	D
14	F	G	A	B	C	D	E

3.5.1 Pilot Study 1

The initial pilot study used a Within-Subjects design without any counterbalancing.

12

3.5.2 Pilot Study 2

The second pilot study also used the Within-Subjects design and introduced the Latin Squares design for counterbalancing. There were only 7 subjects in the second pilot study; therefore, a Replicated Latin Squares design was not used. See Table 3.4 for a standard Latin Squares design.

Table 3.4. Latin Squares design for 7 subjects and 7 conditions.

subjects	conditions						
1	A	B	C	D	E	F	G
2	B	C	D	E	F	G	A
3	C	D	E	F	G	A	B
4	D	E	F	G	A	B	C
5	E	F	G	A	B	C	D
6	F	G	A	B	C	D	E
7	G	A	B	C	D	E	F

3.6 Dependent Measures === Breaking the Standard Evaluation Method

Why did I make the changes to the variables I used, or why did I use the task that I used. (Forcing backspace/correctness and not allowing correct stuff to be backspaced etc. Should be justification from Pilot study + andy sources on this kind of stuff).

The statistical methods being used will consist of One-Way ANOVAs for each set of the recorded and calculated variables for the quantitative data. There will be 7 conditions for each which are representative of the 7 separate inputs methods being tested. Then, Tukey's HSD (honest significant difference) for multiple-compare will be used in conjunction with the ANOVAs using a post-hoc analysis. In addition, a Two-Way ANOVA will be used on each set of variables in conjunction with the previous swipe experience variable.

The qualitative approaches being analyzed are the Likert scale and a ranking system. The Likert scale will use a One-Way ANOVA for analysis because responses

13

to several Likert questions may be summed up providing that all the questions use the same Likert scale and that the scale is a defensible approximation to an interval scale. The ranked system will use the Friedman's test for analysis, and will be subject to a post-hoc analysis using Tukey's HSD.

3.7   Word Dictionary

3.7.1   Why Similar Words are Important

3.7.2   Word Dissimilarity Algorithm

3.8   Word-Gesture Keyboard Construction

3.8.1   Leap Motion — Needs to be referenced in intro

3.8.2   Writing the Code — rename

3.8.3   Designing various WGKs

3.8.3.1   Separating words — rename (figure out ordering).

3.8.3.2   Separation of motor space and display space — rename (figure out ordering).

3.8.3.3   Size of motor space / User calibrated — rename (figure out ordering)

.

3.8.4   Word-Gesture Recognition

Lack of recognition and reason for it. Simulated recognition.

3.9 WGKs

3.9.1 Controller WGK — Need to see if we want to change the implementation to a Controller WGK from std controller

3.9.2 Surface WGK

3.9.2.1 Touch screen.

3.9.2.2 leap surface emulation.

3.9.3 Mid-Air WGKs

3.9.3.1 static.

3.9.3.2 dynamic.

3.9.3.3 bimodal.

3.9.3.4 pinch gesture emulation.

3.10 Study/Experimental Design

CHAPTER FOUR

Results and Analysis

This chapter addresses the results from your data analysis only. This chapter does not include discussing other research literature or the implications of your findings.

Usually you begin by outlining any descriptive or exploratory/confirmatory analyses (e.g., reliability tests, factor analysis) that were conducted. You next address the results of the tests of hypotheses. You then discuss any ex post facto analysis. Tables and/or figures should be used to illustrate and summarize all numeric information.

For qualitative and historical research, this chapter usually is organized by the themes or categories uncovered in your research. If you have conducted focus groups or interviews, it is often appropriate to provide a brief descriptive (e.g., demographic) profile of the participants first. Direct quotation and paraphrasing of data from focus groups, interviews, or historical artifacts then are used to support the generalizations made. In some cases, this analysis also includes information from field notes or other interpretative data (e.g., life history information).

4.1 Text-Entry Rate

4.2 Errors

4.2.1 MWD

4.2.1.1 not modified.

4.2.1.2 modified.

4.2.2 KSPC

4.2.2.1 not modified.

4.2.2.2   modified.

4.2.3   Total Error Rate

4.3   Frchet Distance

4.3.1   modified vs nonmodified

4.4   Interaction with Inputs

Use table to show the various times, distances, accuracy etc

4.5   User preferences

CHAPTER FIVE

Discussion, Future Work and Conclusion

The purpose of this chapter is not just to reiterate what you found but rather to discuss what your findings mean in relation to the theoretical body of knowledge on the topic and your profession. Typically, students skimp on this chapter even though it may be the most important one because it answers the "So what?" question.

Begin by discussing your findings in relation to the theoretical framework introduced in the literature review. In some cases, you may need to introduce new literature (particularly with qualitative research). This chapter also should address what your findings mean for communication professionals in the field being examined. In other words, what are the study's practical implications?

Doctoral students also should discuss the pedagogical implications of the study. What does the study suggest for mass communication education?

This chapter next outlines the limitations of the study. Areas for future research then are proposed. Obviously, the thesis or dissertation ends with a brief conclusion that provides closure. A strong final sentence should be written.

5.1   Discussion

5.1.1   Calibration

5.1.2   Reasons stylus was used instead of hands

5.1.3   Reasons that I didn't implement a word-prediction algorithm for word-gesture keyboard

5.1.4   issues with entrance/exit of touch plane

Solve the problem of entering/leaving the touch plane (we move along a sphere) so when moving away from the touch plane, you often hit the key above what you're leaving on.

5.2   Future Work

5.2.1   image processing, any keyboard

5.2.2   augmented reality adaption

5.2.3   using sphere techniques to better determine the leaps position on a surface

5.2.4   with better leap detection, use hand instead of stylus

5.2.5   would word-gesture controller-based keyboard be better for consoles

5.2.5.1   Use WKG with controller.

5.2.5.2   Use gesture controller (x-box kinect) for text entry.

5.2.6   *since we use 3 degrees, if above intersect plane, use gestures for other functions,*
        aka controlling the mouse or other gesture related things

5.2.7   *Accessibility*

5.3   *Conclusion*

APPENDICES

APPENDIX   A

Code

This is an example of an appendix. Perhaps for code listings or large tables or notation guides.

APPENDIX B

Vita

This is another example of an appendix. Perhaps for listing your CV. Or giving examples of having multiple appendices.

BIBLIOGRAPHY

Banerjee, A., J. Burstyn, A. Girouard, and R. Vertegaal (2012). Multipoint: Comparing laser and manual pointing as remote input in large display interactions. In *International Journal of Human - Computer Studies*, Volume 70(10) of *CHI '11*, pp. 690–702. Elsevier Ltd.

Cockburn, A., P. Quinn, C. Gutwin, G. Ramos, and J. Looser (2011). Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. In *International Journal of Human-Computer Studies*, Volume 69(6), pp. 401–414. Elsevier Ltd.

Devoe, D. (1967). Alternatives to handprinting in the manual entry of data. In *IEEE Transactions on Human Factors in Electronics*, Volume HFE-8(1), pp. 21–32. IEEE.