Universität Koblenz - Landau
Department für Informatik,
Institute for Web Science and Technologies
Prof. Dr. Steffen Staab
Dr. Chandan Kumar

7. August 2020

Master Thesis of Ramin Hedeshy

TGSBoard: Touch and Gaze Sensitive On-screen Keyboard for Effective and Efficient Text Entry

Abstract

Gaze-based text entry systems have been an important means of communication for people with motor disabilities. Although several dwell-time and dwell-free tools have been developed to facilitate the process of gaze-based text entry, still the typing speed is quite slow and the cognitive load is rather high. Moreover, most previous methods are developed only based on gaze and fixations sequence. However, these methods can result in lengthy amounts of time for typing. Besides, users cannot always perfectly gaze at every key in many cases. The proposed approach will combine the simplicity and accuracy of touch inputs with the speed of eye typing by gaze swiping to provide efficient and comfortable dwell-free text entry. We aim to use the gaze path information to compute candidate words, and to allow explicit activation of common text entry commands, such as selection, deletion, and revision by using touch gestures. The system will also be able to automatically correct the word to some extent even if the user misses few letters. Moreover, we propose candidate words augmented on top of the keys which will reduce the visual search and present more comfort for the end user.

Keywords: Text Entry, Eye-typing, Dwell-free, Swipe typing, Gaze and touch

1 Introduction

Communication, known as the foundation of all human relationship, plays a vital role in human life. However, people with motor disabilities may be unable or face severe difficulties in communication, e.g., Motor neuron diseases (MND), such as Amyotrophic lateral sclerosis (ALS) or Multiple sclerosis (MS). The crude annual incidence rate of ALS in the general European population is reported to be about 2.16 per 100000 person-years population [LTH+10]. Even though eye-tracking techniques have provided new methods of communication for them, there is still room for improvement in the usability, speed, and comfort of these methods.

Eye-tracking systems provide only a single continuous input, hence the ability to distinguish gaze intended to gather visual information and gaze intended to interact is ambiguous. This problem is known as Midas Touch problem [Jac90]. Dwell-time method, the period of time when a user fixates or focuses on an element to select it or to activate an interaction, is the most commonly used technique to tackle this problem. Although there are techniques to adjust the duration of the dwell time [Maj12], text entry systems based on these methods are still slow. The use of eye gestures is another technique that has been used by researchers in recent years. For instance, the inside-outside-inside fashion in EyeK [SPC13] or the reverse crossing in EyeSwipe [KFJ+16].

Several dwell-free tools have been developed to enhance the traditional solutions. EyeSwipe is one of the latest approaches in this topic, has achieved 11.7 WPM [KFJ+16], which is still far behind the physical keyboards. One of the reason is that users still need to accurately specify the first and the last letter. Another well-known eye-typing method called Dasher [WBM00] has introduced a continuous typing mechanism which allows users to select the letters dynamically. The performance of Dasher is determined by the language model and the user interface. Filteryedping [PPT15] finds the intended word by filtering extra letters from the sequence of letters gazed at by the user. Users have reached an average of 14.75 WPM typing speed with Filteryedping after about 2 hours of practicing.

Most previous methods are developed only based on gaze and fixations sequence. However, these methods can result in lengthy amounts of time for typing, mainly because of determining the start or end of the typing action with eye movements. In addition, users cannot always perfectly gaze at every key in many cases. For instance, some patients may have difficulty in keeping the cursor steady on any key. Besides, eye-tracker calibration is not perfect. Consequently, for some users, it is not easy to gaze at all the letters of a word or accurately specify the intention of typing by gaze.

Although many physically challenged people could no longer isolate their index fingers for typing, they could still move their arms [Rom13]. Moreover, majority of patients are using switch inputs so the applicability of integrating them as secondary sort of input beside eye gaze is very feasible. As a result, combination of gaze and simple touch gestures can provide opportunities for efficient, faster and more usable inputs.

Therefore, in this thesis, we propose TGSBoard, a new approach to help people with motor impairments achieve text communication using eye movement embedded with simple touch gestures. Providing a more efficient and faster system aiming at minimizing the efforts of eye typing is the main objective of our study through the following contributions:

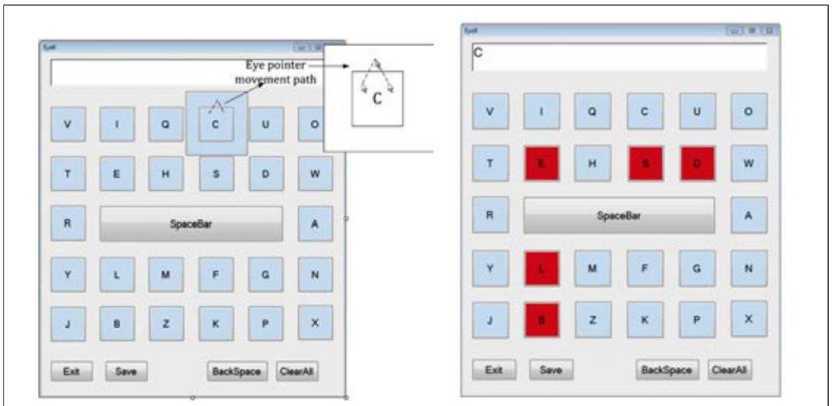
- TGSBoard will integrate simple touch gestures with gaze swiping for enhanced accuracy and speed.
- TGSBoard will not require users to look at each letter of a word to be typed.
- TGSBoard will suggest candidate words dynamically augmented on the keys, with an optimized method of selection, when users pause while swiping over the keys.
- TGSBoard support a detachable trigger or touch device that can be easily integrated into the system via the application web services.

The rest of the proposal is organized as follows. The next section discusses related work. Section 3 explains the TGSBoard technique and presents the proposed solution in details. Section 4 details the evaluation phase. Organizational matters, preliminary outline and time schedule are presented in Sections 5, 6 and 7, respectively.

2 Related work

In this chapter, we present the related studies in regard to usage of gaze for dwell free text entry.

Sarcar et al.[SPC13] have proposed the EyeK eye typing interface which replaced dwell operation with moving the eye pointer through the key in an inside-outside-inside fashion, figure 1. They have achieved on an average 15% higher text entry rate over dwell free interfaces and 6.03 WPM.



(a) Eye movement on a key (b) The key gets selected

Figure 1: Dwell time free eye typing in EyeK interface Minimizing visual [SPC13]

EyeSwipe [KFJ+16] have scored 11.7 WPM. It requires users to accurately select the first and the last characters using a novel method called reverse crossing, and glimpse through or even the vicinity of the middle characters in sequence, figure 2. Although the reverse crossing showed better efficiency than dwelling [SPC13], still the effort of accurately specifying the first and the last character is rather significant, especially for words with few characters. Moreover, since users continually need to pick the suggested word with reverse crossing, the typing flow is disrupted.

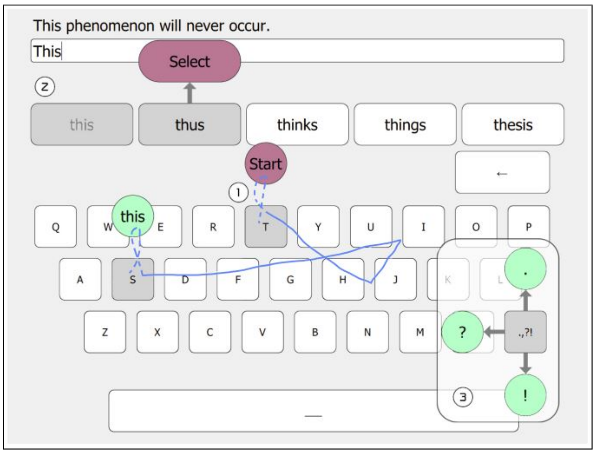


Figure 2: Entering the word "This" in EyeSwipe interface using reverse crossing [KFJ+16]

Another significant approach in this field is Filteryedping [PPT15]. It can suggest all the potential words by filtering extra letters from the sequence of letters gazed at by the user. Users have reached an average of 14.75 WPM typing speed. They have claimed that their proposed solution has achieved 19.28WPM after about 2h of typing. Filteryedping does not require users to specify the first and the last character. However, if users do not gaze one or two letters of the intended word, the system fails to recommend the correct word. The interface and the study setup of this approach are illustrated in the figure 3.

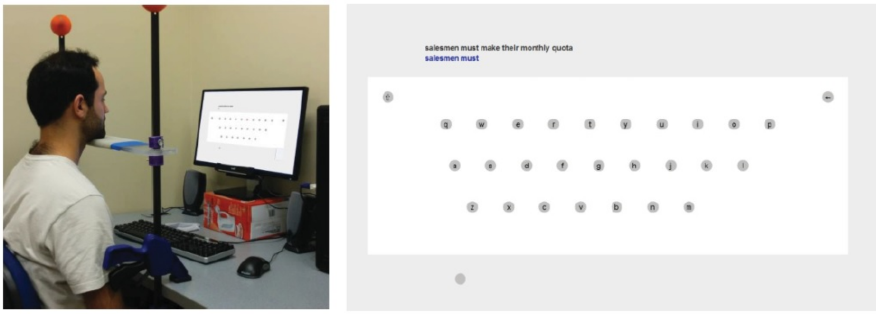


Figure 3: Study setup and the main interface of Filterping. [PPT15]

Combination of gaze and touch has been used in different studies. In a study [PAG16] authors have combined gaze and touch in tablets to allow users to select by gaze and to manipulate by touching. The technique allows users to zoom into map locations they look at. Another study [PG16] has used the touch+gaze technique in tablets to provide the possibility for users to interact with on hand, finger, and grip. They have delivered the possibility of typing with one thumb while holding the tablet with the same hand, figure 4. However, their method does not allow users to swipe, so letters ought to be selected one by one.

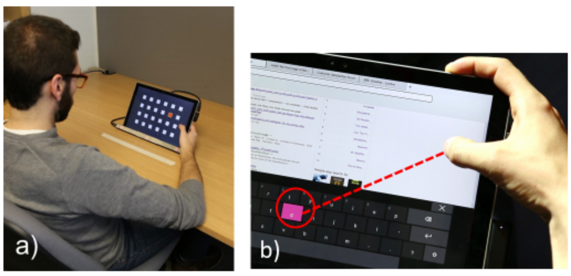


Figure 4: Combination of touch and gaze interaction, user interface and the home screen (a), one hand text entry (b). [PG16]

Recently, this technique has been also used in similar study [Akb18] conducted at the University of Koblenz-Landau. The study revealed a new method for PIN entry with a remarkably higher security level. They have used gaze path for selecting the target keys while touching the corner of the screen to confirm the selection. Their method has addressed shoulder surfing problem. The interface can be seen in figure 5.

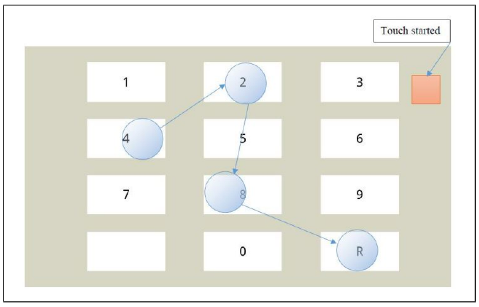


Figure 5: TouchGaze interface: The user touches any part of the screen and holds the touch, then looks at 4, 2, 8 and R to enter 4288 [Akb18]

Overall, the most important aspect that differentiates this master thesis from the related work is that in our solution users are allowed to type more freely since it is not required to accurately specify all the letters or the first and the last character of the word. Therefore, we believe our approach can increase the accuracy and reduce the workload during eye typing. Our main goal is to show how gaze swipe typing embedded with touch can make eye typing effective and efficient. Additionally, we propose a new way of suggesting the candidate words on the keys while users' gaze is swiping through the virtual keyboard, which can address the problem of interruption in typing flow by suggestions.

3 Methodology

The proposed method consolidates touch- and gaze-based swipe for text entry, i.e., a word can be typed by completing the following actions. Firstly the user fixates on the first letter of the desired word and confirm the typing action by touch and then glance through the intermediate letters. The first and last letters are used by the system to filter the lexicon. The corresponding empirical gaze path is used to compute the candidate words. At this point, the user will have two possible options. The first option will be to release the touch and the most probable word will be added to the written text. The other option would be to pause/dwell shortly on the last character to see the candidate words list shown above the key. It can tackle the problem of ambiguity in swipe typing caused by words having similar paths. The user can select the correct suggestion by swiping to right or left. Swipe down can be used to delete a typed word or to cancel a gaze path. Figure 6 showcase the TGSBoard concept design. For the realization of this concept following are the essential components.

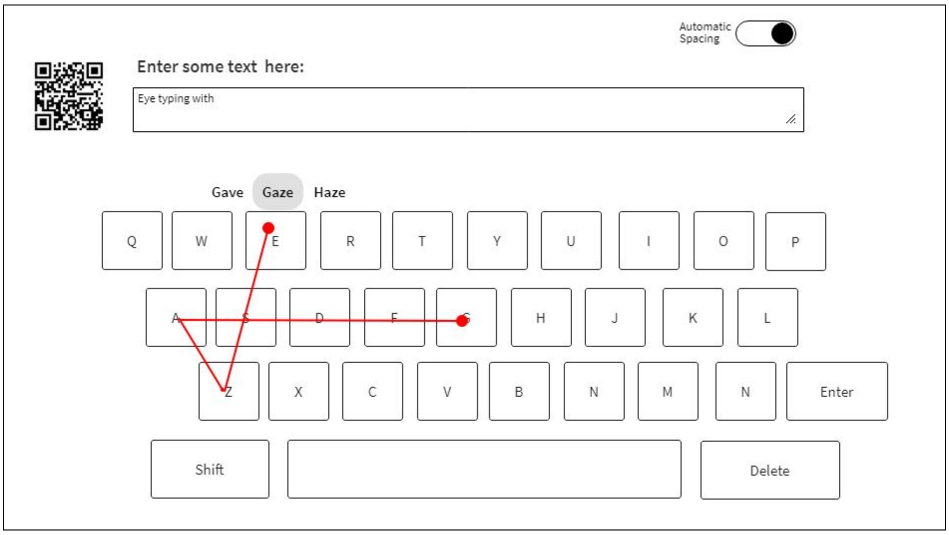


Figure 6: TGSBoard User Interface Wireframe

3.1 Candidate word generation

There are several methods available that can be used for computing gaze paths and generating a set of candidate words. Dynamic Time Warping (DTW) is widely used in gesture and input pattern recognition for comparing two time sequences [BC]. This genetic model can be used in our approach also to compare the distance between the user's gaze path and candidate words. However, this algorithm has a time and a space complexity of O (X*Y) for input lengths X and Y, where X is the lengths of the gaze path, and Y is the length of the candidate word. There are alternative techniques that we can employ instead of DTW, e.g. Fréchet distance [Fré06]. Moreover, using a language model we can reorder and sort the set of candidate words that we get from DTW and make a set of likely candidate tuples (word, probability) based on their distance to an empirical path. Besides, by observing many sentences from a training set and using n-gram model we can find the most probable candidate.

3.2 Apparatus

Figure 7 (a) illustrates a simple tool that makes it possible for many people with finger and hand weakness to continue to use their mobile device touchscreen, a keyboard or a paper communication board. There are also tools available for people with severe physical disabilities, e.g., wheelchair joysticks, foot switches, figure 7 (b). In our proposed application users can differ eye deliberate interaction with the normal visual gaze by a simple touch gesture or trigger. The touch action could be on a mobile phone or any other detachable touch device. TGSBoard will permit authenticating any device for touch interaction by simply scanning a QR code. We will utilize AngularJS two-way data binding to have instant interplay between the main system and the connected device. TGSBoard will also allow users to rebind/remap the default touch gestures with their favored input. To achieve this goal, we will use Javascript input listeners to capture user input and set it instead of the default value.

Figure 7: A simple tool to maintain the ability to touch [Rom13] (a), Mini proportional joysticks.[Lan] (b)

3.3 Interface

The proposed work is to implement and develop a web-based on-screen keyboard using Python[1] and Angular JS[2]. The reasons for choosing to use Python and Angular JS are as follows. Firstly, Python has a huge collection of libraries which support and advance the process of working with eye trackers. Secondly, as far as the application needs to instantly synchronize and compare the fixation points between clients, Angular JS will be employed as the client-side language of the application because of its exceptional run-time performance. It will be possible to embed TGSBoard within the browser window which will pop up when a specified entry field is focused. The typing interface will compose of a text box, a QR code, a virtual keyboard, and settings (figure 6).

Although the keyboard layout will be QWERTY (to follow the conventional layout), it will be possible to extend the algorithms used in TGSBoard to support other layouts. Furthermore, the default language of the application will be English but supporting other languages will be easily possible. As far as the programming layers of the proposed application are planned to be decoupled, it will be also possible to easily program different UI for the keyboard on other platforms.

[1] Python: http://www.python.org/
[2] Angular JS: https://angularjs.org/

Once the application is ready, we will invite around 20 participants to test the TGSBoard. We will analyze and compare the average eye typing rate as well as the error rate. We will also draw a comparison between the previous methods and our proposed approach. This study will be conducted between December 2018 and April 2019. The evaluation phase is described in detail in the evaluation section.

4 Evaluation

The evaluation of the proposed approach will be a mixture of qualitative and quantitative methods. The evaluation will be conducted by around 20 participants. The contributors will try a dwell-time keyboard and Eyeswipe which is one of the latest dwell-free eye typing tools, alongside our approach, as baselines. Eventually, we will draw a comparison between our approach and the previous methods. The following are the steps of the planned evaluation:

Firstly, we will give an introduction regarding the evaluation process to the participants, and they will have time to test and get familiar with the eye tracker and the application. Then, they will receive a series of texts to enter into the system, all the data during the practice will be collected and logged for the future analysis. All the subjects will type using TGSboard, Eyeswipe and a dwell-time keyboard. In order to address errors due to fatigue, subjects will have a short break between every experiment. In addition, we will use the counter-balancing experimental design to control the variation. Therefore, participants will be divided into different groups and every group will try the applications in a different order. In the last phase, we will conduct a survey to collect information about the subject's experience, preference and also their recommendation for enhancement.

A quantitative analysis of the experiment will be assessed by drawing a comparison between the achieved performance result and the previous methods. The average eye typing rate, words per minute (WPM), characters per minute (CPM) and the total error rate are all the examples of assessments that will be taken into consideration in the evaluation phase.

As far as the qualitative evaluation is concerned, the focus will be on inquiring the user satisfaction of subjects in using the application for parameters of speed, accuracy, comfort, and learnability. For the purpose, qualitative analysis will be assessed by having the contributors answering the following normative question:

- How would you rate the speed of typing with current method?
- How would you rate the accuracy of the method?
- How comfortable did you find the text entry method?
- How easy was to learn the text entry method?

  Furthermore, we will ask specific question about word suggestions.

- Have you found the suggestions distracting?
- How easy was to select the suggested words?

A more detailed questionnaire will be designed based on the experimental variables. The output of the analysis report will determine whether the approach leads to a higher efficiency and could increase eye typing speed or not. Furthermore, it indicates the effectiveness of using the trigger in minimizing the effort of initializing the text entry.

5 Organizational matters

Duration of work:	15.12.2018 – 30.04.2019
Candidate:	Ramin Hedeshy
E-Mail:	Hedeshy@uni-koblenz.de
Student number:	216100869
Primary supervisor:	Prof. Dr. Steffen Staab
Supervisor:	Dr. Chandan Kumar

6 Preliminary outline

1. Introduction
2. Related work
3. Methodology
   (a) Requirements Analysis
   (b) Design
   (c) Implementation
   (d) Testing
4. Evaluation
   (a) Methods
   (b) Results
   (c) Discussion
5. Conclusion
6. Bibliography

7 Time schedule

- Introduction: 15/12/2018 – 10/01/2019
  - Literature review: 15/12/2018 – 29/12/2019
  - Related work: 30/12/2018 – 05/01/2019
- Methodology: 05/01/2019 – 28/02/2019
  - Requirements: 05/01/2019 – 10/01/2019
  - Design: 11/01/2019 – 25/01/2019
  - Implementation: 26/01/2019 – 15/02/2019
  - Testing: 15/02/2019 – 28/02/2019
- Evaluation: 01/03/2019 – 30/03/2019
  - Preparing evaluation: 01/03/2019 – 09/03/2019
  - Conducting evaluation: 10/03/2019 – 25/03/2019
  - Result analysis : 26/03/2019 – 31/03/2019
- Revision: 01/04/2019 – 30/04/2019

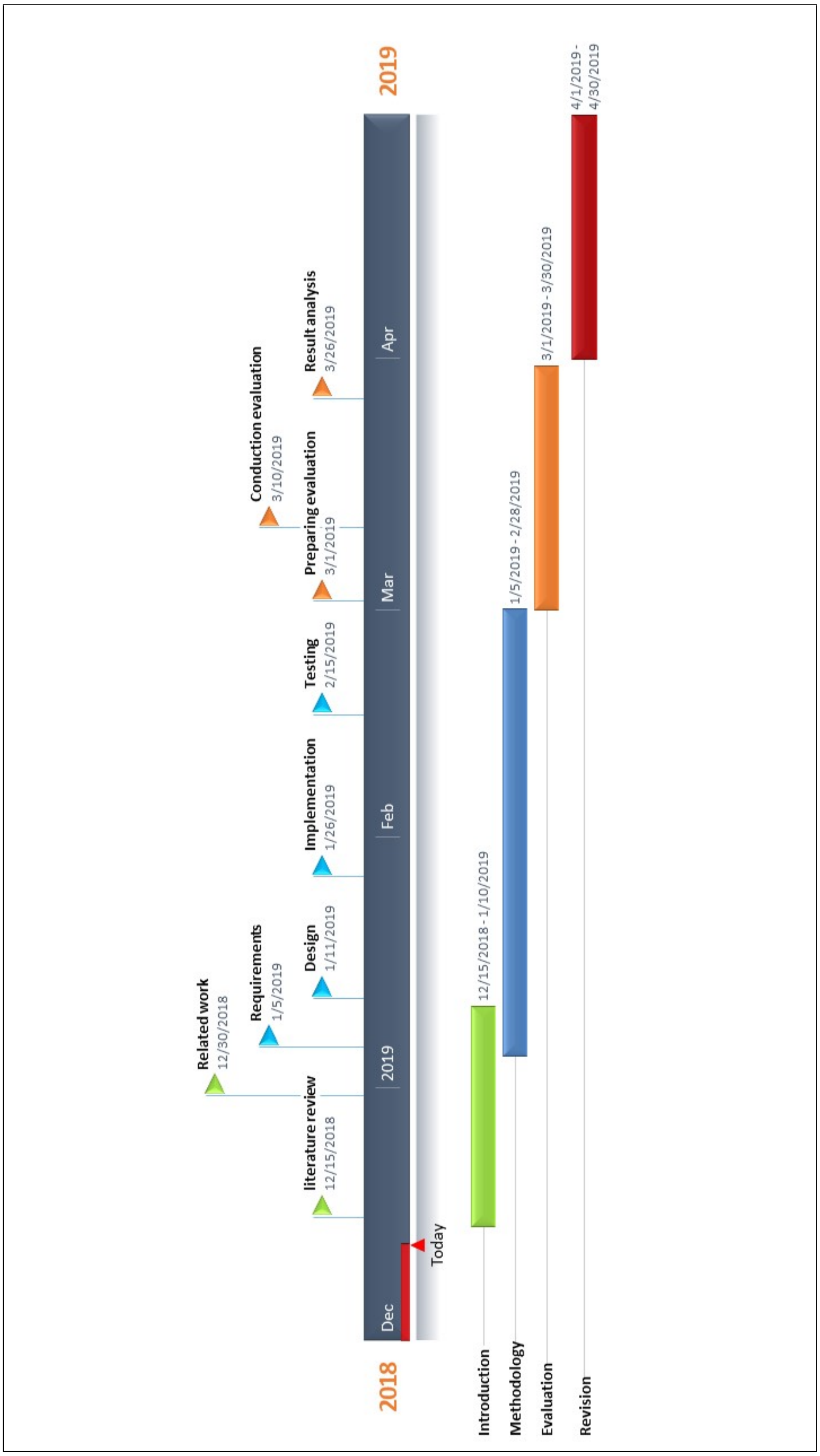


Figure 8: Gantt chart of time schedule

References

[Akb18] Daniyal Akbari. Multi-model optimization of gaze and touch-based pin entry. 2018.

[BC] Donald J Berndt and James Clifford. Using dynamic time warping to find patterns in time series.

[Fré06] M Maurice Fréchet. Sur quelques points du calcul fonctionnel. *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, 22(1):1–72, 1906.

[Jac90] Robert J. K. Jacob. What you look at is what you get: Eye movement-based interaction techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '90, pages 11–18, New York, NY, USA, 1990. ACM.

[KFJ+16] Andrew Kurauchi, Wenxin Feng, Ajjen Joshi, Carlos Morimoto, and Margrit Betke. Eyeswipe: Dwell-free text entry using gaze paths. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 1952–1956, New York, NY, USA, 2016. ACM.

[Lan] Power wheelchair driving methods.

[LTH+10] Giancarlo Logroscino, Bryan J Traynor, Orla Hardiman, Adriano Chiò, Douglas Mitchell, Robert J Swingler, Andrea Millul, Emma Benn, Ettore Beghi, et al. Incidence of amyotrophic lateral sclerosis in europe. *Journal of Neurology, Neurosurgery & Psychiatry*, 81(4):385–390, 2010.

[Maj12] Päivi Majaranta. Communication and text entry by gaze. In *Gaze interaction and applications of eye tracking: Advances in assistive technologies*, pages 63–77. IGI Global, 2012.

[PAG16] Ken Pfeuffer, Jason Alexander, and Hans Gellersen. Partially-indirect bimanual input with gaze, pen, and touch for pan, zoom, and ink interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2845–2856, New York, NY, USA, 2016. ACM.

[PG16] Ken Pfeuffer and Hans Gellersen. Gaze and touch interaction on tablets. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, pages 301–311, New York, NY, USA, 2016. ACM.

[PPT15] Diogo Pedrosa, Maria da Graça Pimentel, and Khai N. Truong. Filteryedping: A dwell-free eye typing technique. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, pages 303–306, New York, NY, USA, 2015. ACM.

[Rom13] Amy Roman. Maintain ability to type, swipe point with hand weakness in als, Nov 2013.

[SPC13] Sayan Sarcar, Prateek Panwar, and Tuhin Chakraborty. Eyek: an efficient dwell-free eye gaze-based text entry system. In *Proceedings of the 11th asia pacific conference on computer human interaction*, pages 215–220. ACM, 2013.

[WBM00] David J. Ward, Alan F. Blackwell, and David J. C. MacKay. Dasher—a data entry interface using continuous gestures and language models. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology*, UIST '00, pages 129–137, New York, NY, USA, 2000. ACM.

8 Signatures

Ramin Hedeshy

Prof. Dr. Steffen Staab

Dr. Chandan Kumar

9 Declaration of Authorship

I hereby declare that the thesis submitted is my own unaided work. All direct or indirect sources used are acknowledged as references.

I am aware that the thesis in digital form can be examined for the use of unauthorized aid and in order to determine whether the thesis as a whole or parts incorporated in it may be deemed as plagiarism. For the comparison of my work with existing sources I agree that it shall be entered in a database where it shall also remain after examination, to enable comparison with future theses submitted. Further rights of reproduction and usage, however, are not granted here.

This paper was not previously presented to another examination board and has not been published.

Koblenz, on August 7, 2020

Ramin Hedeshy