Virtual Keyboard App Using Computer Vision

Project Plan

COMP4801 Final Year Project [2023-24]
Supervisor: Dr. Pan Jia (jpan@cs.hku.hk)
Student: Mak Tsz Shing (3035685914) (u3568591@connect.hku.hk)
Date of submission: 1st October 2023

Table of Contents

List of Figures

1. Introduction

1.1 Background

Keyboard has been the most common input device for computers since the 1970s [1]. With the advanced touchscreen technology, virtual onscreen keyboard has become the most common typing method for mobile devices like smartphones and tablet. Although virtual onscreen keyboards are highly available and convenient, typing on a mobile phone with small screen size can be cumbersome. As mobile devices have become an integral part of our lives which serves as communication tools and productivity companions, there is a growing demand for intuitive and user-friendly text input methods on mobile devices.

1.2 Virtual Keyboard

Virtual keyboards are keyboards without physical keystrokes, such as onscreen keyboard and optically projected keyboard (see section 1.3.2 for detailed introduction). Comparing to physical keyboards, virtual keyboards are more flexible in terms of customization and portability [2]. In addition, virtual keyboards are more secure for sensitive information inputs like entering passwords as hardware-based keyloggers can be prevented [2]. In this project, the type of virtual keyboard to be developed is similar to optically projected keyboard using mobile devices as the physical sensor. The main scope of this project will be the alphabetical keys on the QWERTY keyboard (see Figure 1).

Figure 1. Scope of keys on virtual keyboard to be developed. Keys highlighted in yellow will be focused. Keys highlighted in blue will be considered throughout the project period.

1.3 Related Works

1.3.1 SelfieType

Figure 2. SelfieType [3]

SepfieType is an inside project of Samsung's C-Lab. The project was introduced in the Consumer Electronics Show 2020 (CES 2020). SelfieType is an optical virtual keyboard app which analyses finger movements by the front camera of mobile devices and converts them into QWERTY keyboard inputs. [3]

1.3.2 Laser Projection Keyboard

Figure 3. A laser projection keyboard [4]

Laser Projection Keyboard was invented by IBM engineers in 1992 [5]. It is a physical device with a laser beam projecting the keyboard onto a flat surface (see Figure 3). When a user types on the projected keys, the finger movements are recorded by a camera or an infrared sensor.

2. Project Objectives

The project aims to provide an intuitive and flexible typing experience for mobile devices by developing an interactive and innovative Android app.

The intermediate objective for the project is to develop a demo application, in which the users can try to control their mobile devices by predefined simple hand gestures. The demo application also explores the possibility of implementing vision-based controlling into different scenarios in using electronic devices like gaming and virtual reality. The effectiveness and accuracy of the demo application will be evaluated.

The ultimate objective is to develop an Android keyboard app which generates text input by recognizing the finger actions of users when typing on a flat surface. The front camera will be leveraged as a sensor to detect finger gestures in real time.

3. Proposed Methodology

3.1 Architecture

Computer Vision Algorithms will be used to recognize fingertips coordinates (see Figure 4). The coordinates will determine whether the finger has touched the surface. When touching is detected, a neural network will be used to determine the most likely keystroke entered by the following features: 1) which finger is typed, 2) the distance of the typed finger from the camera, and 3) the possibility of keystroke entered given the previous entered key. The training dataset for the neural network is proposed to be provided by the developer.

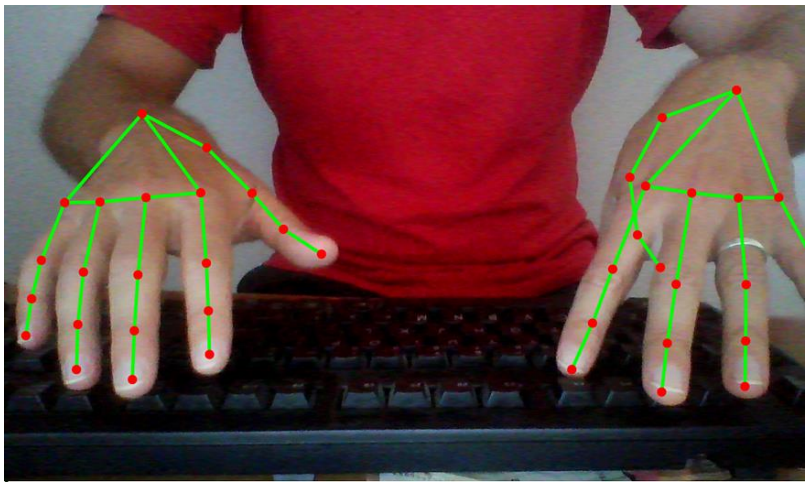

Figure 4. Visualization of finger recognition. The coordinates of the fingertips will be calculated by computer vision algorithms [6]

3.2 Computer Vision

OpenCV is the most well-known Computer Vision library for image processing and real-time computer vision. It would be deployed for background subtraction and real time finger motion tracking. Other computer vision libraries like FastCV would also be considered for supported functions and performance.

3.3 Android App Development

Android Studio would be used for the app development. Kotlin is preferred as the programming language. For the keyboard app, InputMethodService from the Android SDK would be utilized for implementation of the input method. To improve the typing experience, the text service APIs offered by the Android platform would be considered for spell checking and word suggestions.

4. Proposed Schedule

Oct 1	Deliverables of Phase 1 • Finalized Project Plan • Project Webpage
Oct	Study Computer Vision and the related libraries Study research on related works
Nov	Development of demo application • Evaluate the effectiveness of visual-based controlling system • Investigate the algorithm to be used in the keyboard app
Jan 21	Deliverables of Phase 2 • Demo application • Interim report
Dec - Feb	Development of the keyboard app
Feb - April	Optimization and Testing
April 23	Deliverables of Phase 3 • Finalized implementation • Final Report

5. References

[1] *"Input & output devices"*. Khan Academy. Accessed: Sept. 30, 2023 [Online]. Available: https://www.khanacademy.org/computing/computers-and-internet/xcae6f4a7ff015e7d:computers/xcae6f4a7ff015e7d:computer-components/a/input-output-devices

[2] *"Discover the Benefits of Virtual Keyboards"*. Lenovo US. Accessed: Sep. 30, 2023 [Online]. Available: https://www.lenovo.com/us/en/glossary/virtual-keyboard/?orgRef=https%253A%252F%252Fwww.google.com%252F

[3] "*Samsung Electronics to Showcase Successful 'C-Lab Inside' Projects and 'C-Lab Outside' Start-ups at CES 2020"*. Samsung Newsroom. Accessed: Sep. 30, 2023. [Online]. Available: https://news.samsung.com/global/samsung-electronics-to-showcase-successful-c-lab-inside-projects-and-c-lab-outside-start-ups-at-ces-2020

[4] *"Projection Keyboard"*. Wikipedia. Accessed: Sept. 30, 2023 [Online]. Available: https://en.wikipedia.org/wiki/Projection_keyboard

[5] "*Method and device for optical input of commands or data"*. IBM [US]. (1993, Aug. 11). EP0554492A1. Accessed: Sep. 30, 2023 [Online]. Available: https://worldwide.espacenet.com/patent/search/family/008209311/publication/EP0554492A1?q=pn%3DEP0554492

[6] Luciano Sphere. *"Exquisite hand and finger tracking in web browsers with MediaPipe's machine learning models"*. Medium. Accessed Sep. 30, 2023 [Online]. Available: https://towardsdatascience.com/exquisite-hand-and-finger-tracking-in-web-browsers-with-mediapipes-machine-learning-models-2c4c2beee5df