

Studienarbeit

Entwicklung eines Messsystems zur Bestimmung von Zungenbewegungen- und Gesten auf Basis von Infrarot-Abstandssensorik

Hamza Ben Hassen

Geboren am: 9. Juli 1993 in Duisburg

Matrikelnummer: 4114500

Immatrikulationsjahr: 2014

Betreuer

MSc. Christoph Wagner

Betreuernder Hochschullehrer

Jun.-Prof. Dr.-Ing. Peter Birkholz

Eingereicht am: 10. Januar 2020

Aufgabenstellung für die Anfertigung einer Studienarbeit

Studiengang: Elektrotechnik
Studienrichtung: Studienrichtung
Name: **Hamza Ben-Hassen**
Matrikelnummer: 4114500
Immatrikulationsjahr: WS2014
Titel: **Entwicklung eines Messsystems zur Bestimmung von Zungenbewegungen- und Gesten auf Basis von Infrarot-Abstandssensorik.**

Ziele der Arbeit

Gelähmten oder anderweitig motorisch beeinträchtigten Personen bleibt die Interaktion mit einem Computer oft verwehrt, da sie konventionelle Eingabemethoden nicht nutzen können. Eine aussichtsreiche Alternative zu ihren Händen ist die Zunge, welche eine ähnliche Bewegungsflexibilität aufweist und darüber hinaus ermüdungsarm ist. Dazu ist es nötig, die dreidimensionalen Zungenbewegungen innerhalb des Mundraumes zu erfassen und auf zweidimensionale Steuersignale abzubilden. Sensorseitig kann die Bewegungsmessung sowohl kontaktbehaftet als auch kontaktlos erfolgen. Eine Möglichkeit ist die Verwendung von mehreren Abstandssensoren im Infrarotbereich, über welche die Bewegungsrichtung aus der Intensitätsverteilung des reflektierenden Lichts von der Zunge zu den verschiedenen Sensoren ermittelt werden kann. Ein solches Sensorsystem soll in dieser Studienarbeit entwickelt und als Hardwareprototyp aufgebaut werden. Im Zuge dessen soll außerdem untersucht werden, inwieweit Eingabe- bzw. Bestätigungsbefehle mit reiner optischer Abstandssensorik umgesetzt werden können.

Schwerpunkte der Arbeit

- Einarbeitung in den Stand der Technik und der Forschung
- Auslegung des Systems und Auswahl der beteiligten Komponenten
- Aufbau eines Hardwareprototypen
- Systematische Evaluation der Funktionalität des Systems
- Ausführliche schriftliche Dokumentation

Die Arbeit wird in deutscher Sprache verfasst.

Gutachter: Jun.-Prof. Dr.-Ing. Peter Birkholz
Betreuer: MSc. Christoph Wagner
Ausgehändigt am: 7. Oktober 2019
Einzureichen am: 6. Januar 2020

Jun.-Prof. Dr.-Ing. Peter Birkholz
Betreuer Hochschullehrer

Selbstständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts habe ich Unterstützungsleistungen von folgenden Personen erhalten: MSc.Christoph Wagner, Dipl.-Ing. Kathleen Große.

Weitere Personen waren an der geistigen Herstellung der vorliegenden Arbeit nicht beteiligt. Mir ist bekannt, dass die Nichteinhaltung dieser Erklärung zum nachträglichen Aberkennen der Prüfungsleistung führen kann.

Dresden, den 10.01.2020

Unterschrift:

Inhaltsverzeichnis

Abstract	2
1. Einleitung	3
1.1. Motivation	3
1.2. Zielstellung	3
2. Stand der Technik	4
3. Methoden und Materialien	7
3.1. Signalerfassung	7
3.2. Signalverarbeitung	11
4. Evaluation	16
5. Zusammenfassung	18
Literaturverzeichnis	19
A. Pinbelegung	21

Nomenklatur

AT Assistive Technology

TCI Tongue Computer Interface

Kurzfassung

Die Interaktion mit einem intelligenten Gerät durch traditionelle Eingabemethoden kann für Menschen mit eingeschränkter Fingerfertigkeit eine Herausforderung darstellen. Der Zungenmuskel zeigt vielversprechende dreidimensionale Flexibilität und Ermüdungsbeständigkeit, die ihn für die Verwendung als Eingabemodalität geeignet macht.

Ziel dieser Arbeit ist es, den Einsatz von Infrarot-Näherungssensoren zu untersuchen, um ein ergonomisches zungenbasiertes Eingabesystem zu implementieren. Dieses Eingabesystem implementiert einen Gestenerkenner sowie eine Mauszeigertsteuerung.

Abstract

Interacting with a smart device through traditional input methods can be challenging for people with reduced dexterity. The tongue muscle shows promising three dimensional flexibility and resistance to fatigue that makes it suitable to be used as an input modality.

The aim of this work is investigate the usage of infrared proximity sensors to implement an ergonomic toungue based input system. This input system will implement a gesture recognizer as well as a mouse cursor controller.

1. Einleitung

1.1. Motivation

Menschen mit schweren Behinderungen als Folge von Hirn- und Rückenmarkverletzungen, Schlaganfall oder Parkinson sind für ihre täglichen Kommunikations- und Mobilitätsbedürfnisse auf assistive Technologien (AT) angewiesen. AT Geräte verlassen sich auf die Verwendung von zugänglichen Eingabemodalitäten, um mit externer Hardware zu interagieren.

Die Zunge zeigt einige Eigenschaften, die sie in diesem Zusammenhang als Eingabemodalität geeignet macht. Insbesondere die Zungenkontrolle geht selten verloren, da sie über den *nervus hypoglossus* mit dem Gehirn verbunden ist. Dieser Nerv ist robust und wird selten verletzt. [CB18]. Eine weitere relevante Eigenschaft der Zunge ist ihre Widerstandsfähigkeit gegen Muskelermüdung bei freier Bewegung im Mund. Dadurch kann sie über längere Zeiträume eingesetzt werden. .[Sol06]

Zur Bestimmung der Zungenposition können verschiedene Sensorsysteme eingesetzt werden, wie z.B. kapazitive, induktive oder optische Näherungssensoren.

Aktuelle AT Geräte neigen dazu, sperrig zu sein, externe Headsets zu benötigen oder in ihren Funktionalitäten eingeschränkt zu sein. In dieser Arbeit versucht eine prototypische Implementierung einer Zungencomputerschnittstelle auf Basis von optischer Infrarot-Näherungssensorik diese Einschränkungen zu überwinden.

1.2. Zielstellung

In dieser Arbeit wird eine einer Zungencomputerschnittstelle (TCI) basierend auf optischer Infrarot-Näherungssensorik entwickelt und evaluiert. Diese Schnittstelle soll die Interaktion mit externer Hardware wie z.B. einem Computer oder einem motorisierten Rollstuhl durch Zungengesten und Bewegungen ermöglichen.

2. Stand der Technik

Zur Erkennung der Zungenposition werden verschiedene Sensortechniken eingesetzt.

Direkte Kontaktaufnahme :

Der physikalische Kontakt zwischen Zunge und Sensorfläche ermöglicht die Ableitung der Zungenposition.

Die Elektropalatographie macht sich dieses Konzept zunutze, indem ein Elektrodenarray in einen künstlichen harten Gaumen gelegt wird. Dieser wurde zur Bedienung eines Computers und zur Steuerung eines Mauszeigers verwendet. 16 Elektroden waren dafür bereits ausreichend. [HK15].



Abbildung 2.1.: Electropalatograph[HK15]

Induktive Näherungssensorik :

Ein induktiver Näherungssensor arbeitet durch die Erzeugung eines elektromagnetischen Feldes mit Hilfe eines Oszillators. Nähert sich ein metallisches Objekt diesem elektromagnetischen Feld, werden Wirbelströme im metallischen Objekt induziert. Dies reduziert die Amplitude der Schwingung.

Figure2.2 veranschaulicht das Funktionsprinzip für induktive Sensoren.

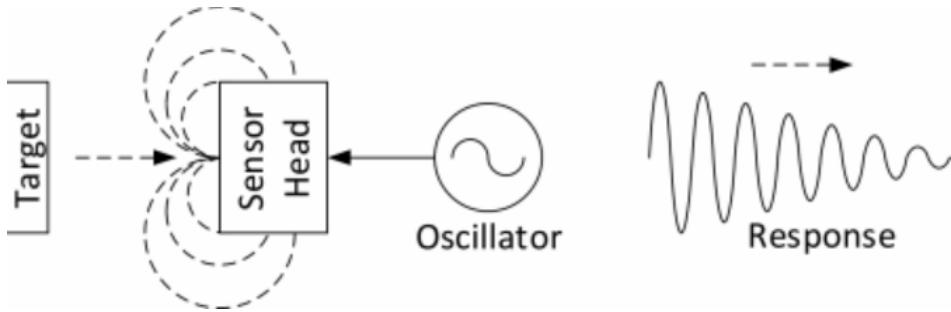


Abbildung 2.2.: Induktive Näherungssensorik[LYC⁺15].

Magnetische Näherungssensorik:

Diese Technik beruht auf der Erfassung der Veränderungen der Magnetfelder, die durch die Bewegung des Zielobjekts verursacht werden. Magnetische Sensoren erkennen Magnetfelder und liefern die Position, den Winkel und die Stärke der Veränderungen.

Eine Umsetzung dieses Prinzips ist das "Tongue Drive System"[HWG08]. Diese Implementierung verwendet mehrere magnetische Sensoren und einen Permanentmagneten, der an der Zunge befestigt ist, um ihre Bewegung zu erkennen. Die Abbildung 2.3 veranschaulicht das Sensorprinzip.

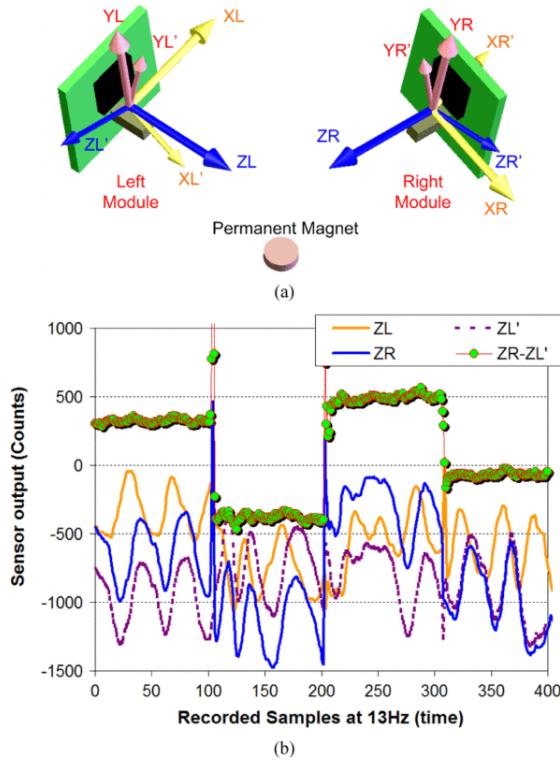


Abbildung 2.3.: TDS Magnetisches Näherungssensorprinzip[HWG08].

Infrarot-Näherungssensorik :

Das Grundprinzip der optischen Abstandserfassung besteht darin, dass das Lumen (Helligkeit) einer von einer Punktlichtquelle beleuchteten Fläche proportional zum inversen Quadrat des Abstands ist.[CW78]

Zur Anwendung dieses Prinzips sendet ein Lichtsender einen Lichtstrahl aus, der an der Oberfläche des Zielobjektes reflektiert wird. Diese Reflexionen werden von einem optischen Sensor, z.B. einer Photodiode, erfasst. Je näher das Zielobjekt, desto höher ist die Intensität nach Gleichung 2.1

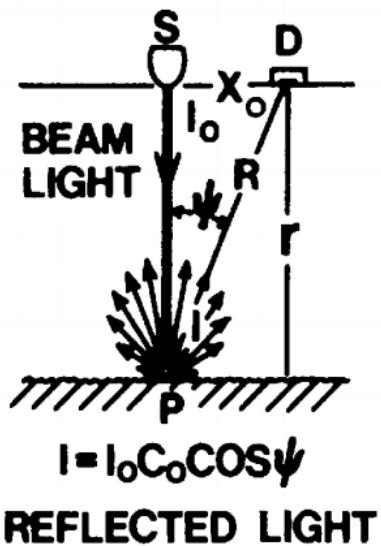


Abbildung 2.4.: Optical sensing principle[CW78]

$$I = \frac{I_0 r}{2\pi(r^2 + x_0^2)^{\frac{3}{2}}} \quad (2.1)$$

Eine Implementierung mit dieser Technik hat gezeigt, dass diese Einschätzung für eine Reihe von Interaktionen möglich ist [SKPT09].

3. Methoden und Materialien

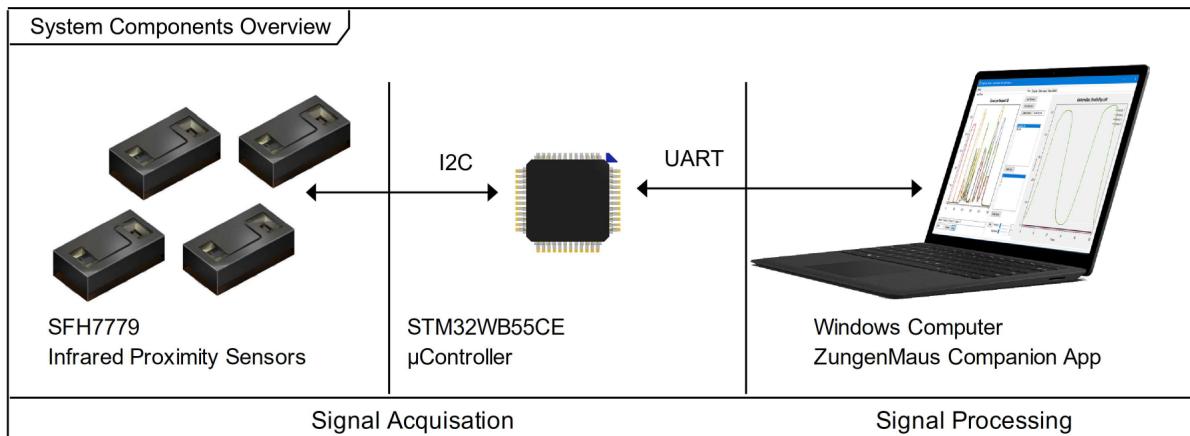


Abbildung 3.1.: Übersicht der Systemkomponenten

3.1. Signalerfassung

Die Signalanpassung basiert auf 4 Infrarot-Näherungssensoren des Herstellers OS-RAM SFH7779 .

Diese Sensoren werden mit einem Mikrocontroller im i2c-Fastmodus gemäß Abbildung 3.3 verbunden und liefern die Distanz als Zählwerte nach dem Empfindlichkeitsdiagramm 3.2.

Der verwendete Mikrocontroller ist der ultra low power STM32WB55 . Für die Evaluierung wurde das Entwicklungsboard P-NUCLEO-WB55 verwendet.

3.1 Signalerfassung

PS sensitivity f(R = reflectivity)
 $V_{DD}=2.5V$; $I_{LED}=200mA$; $T_{rep}=100ms$; $A_{Reflector} = 10 \times 13 \text{ cm}^2$

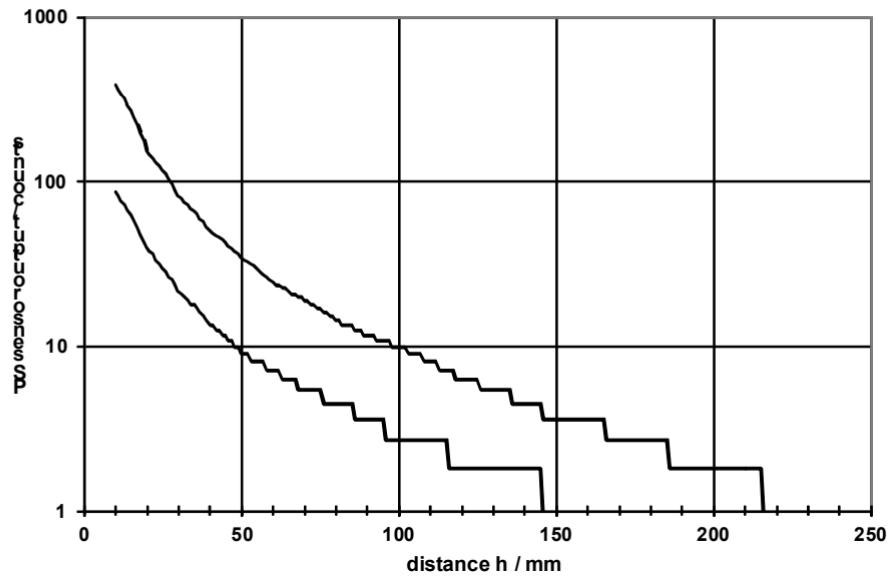


Abbildung 3.2.: Empfindlichkeitsdiagramm SFH7779 [SFH]

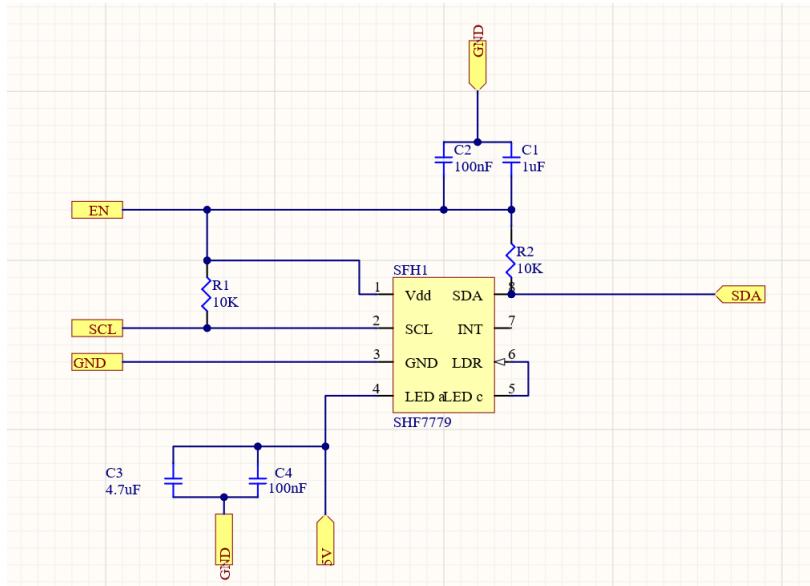


Abbildung 3.3.: Schaltplan der Sensoreinheit

Multiplexing

Die Sensoren teilen sich eine feste I2C-Adresse. Um alle verschiedenen Sensoren anzusprechen, war Multiplexing erforderlich.

Das implementierte Multiplexing funktioniert durch sequentielles Abfragen einzelner Sensoren. Ein GPIO-Ausgang des Mikrocontrollers ist mit dem VDD verbunden, wodurch der Sensor ein- und ausgeschaltet wird nach dem in Abbildung dargestellten Abfrage-Algorithmus.

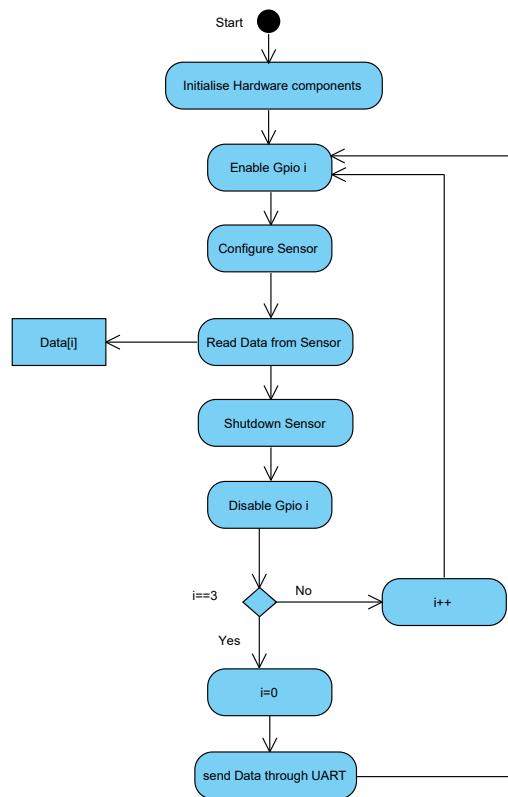


Abbildung 3.4.: Abfragezyklus

Hardware-Prototyp

Die Näherungssensoren sind paarweise gruppiert und haben jeweils einen separaten I₂C-Anschluss. 3.6 zeigt den Evaluierungsprototyp. Die Pinbelegung ist im Anhang beigelegt.

Die Daten werden über das UART-Protokoll in strukturierten Frames geliefert.

Ein Frame besteht aus einer Messung von jedem Sensor, wobei dieser Formatierung gefolgt wird:

xxx;xxx;xxx;xxx\n\r

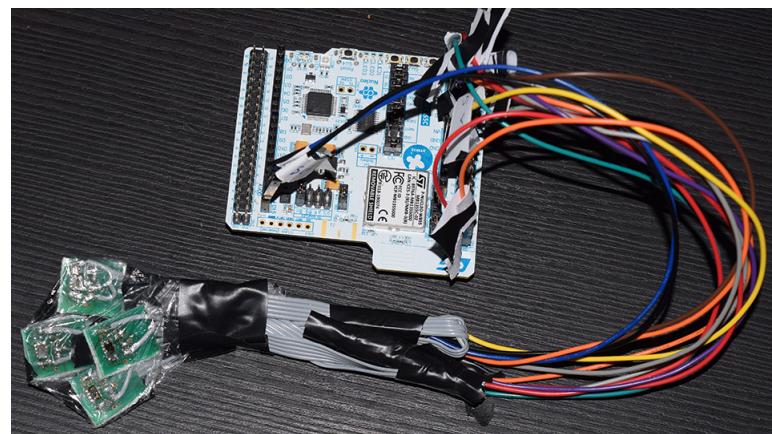


Abbildung 3.5.: Evaluierungs-Hardware



Abbildung 3.6.: Sensorenanordnung

3.2. Signalverarbeitung

Ein in C# implementiertes Programm erhält die seriellen Daten und implements the following functionalities :

Aufbau der Gestensammlung

Die Software bietet eine GUI, die das Einlernen neuer Gesten unterstützt. Gestensammlungen können importiert oder in einer externen Datei gespeichert werden.

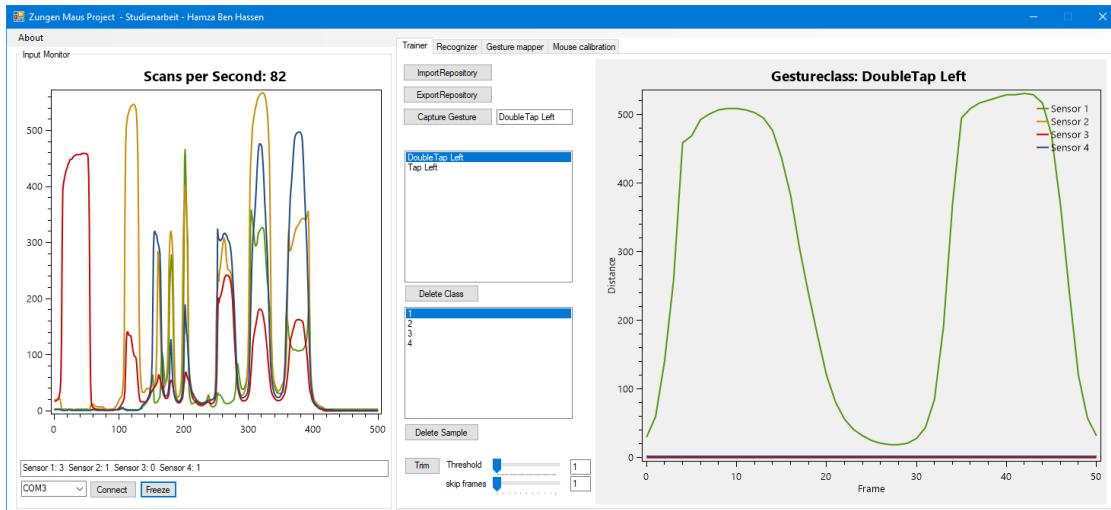


Abbildung 3.7.: Gestensammlung GUI

Gestenerkennung

Die in dieser Arbeit implementierte Gestenerkennung funktioniert wie folgt:

- Es wird ein Trainingsmodell mit Referenzgesten mit mehrfachen Wiederholungen in unterschiedlichen Geschwindigkeiten und Positionen aufgebaut.
- Ein Erkenner vergleicht Eingabesequenzen mit Referenzgesten. Dieser Vergleich beinhaltet eine Vorverarbeitung der Eingabesequenz, gefolgt von einer Distanzberechnung zur Referenzgeste.
- Auf Basis des Erkennungsergebnisses kann die Software entscheiden, bei Erreichen bestimmter Schwellenwerte ein Erkennungsergebnis auszulösen.

Die Vorverarbeitung aligns the signals by trimming the idle state from the signal. Der Ruhezustand kann durch kleine Änderungen in der Geschwindigkeit des Eingangssignals charakterisiert werden. Der Vorverarbeitungsalgorithmus sucht nach einer vorgegebenen minimalen Geschwindigkeitsänderung in der Eingangssequenz,

3.2 Signalverarbeitung

welche den Beginn der Geste markiert. The same principle applied to the reversed input sequence marks the gesture end.

Abbildung 3.8 veranschaulicht die Ergebnisse dieses Algorithmus bei einer minimalen Geschwindigkeitsänderung von *10 Counts/Frame*.

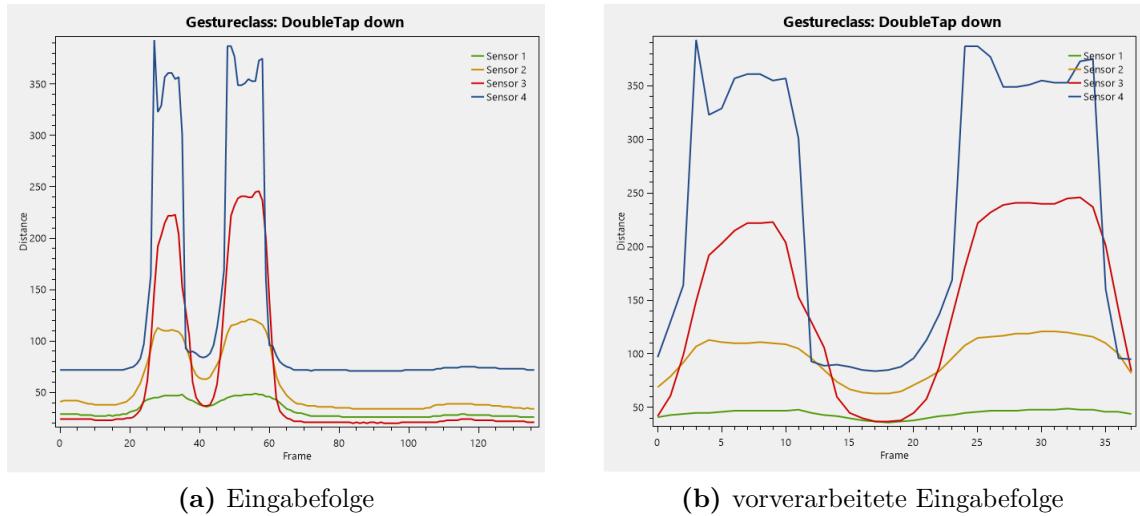


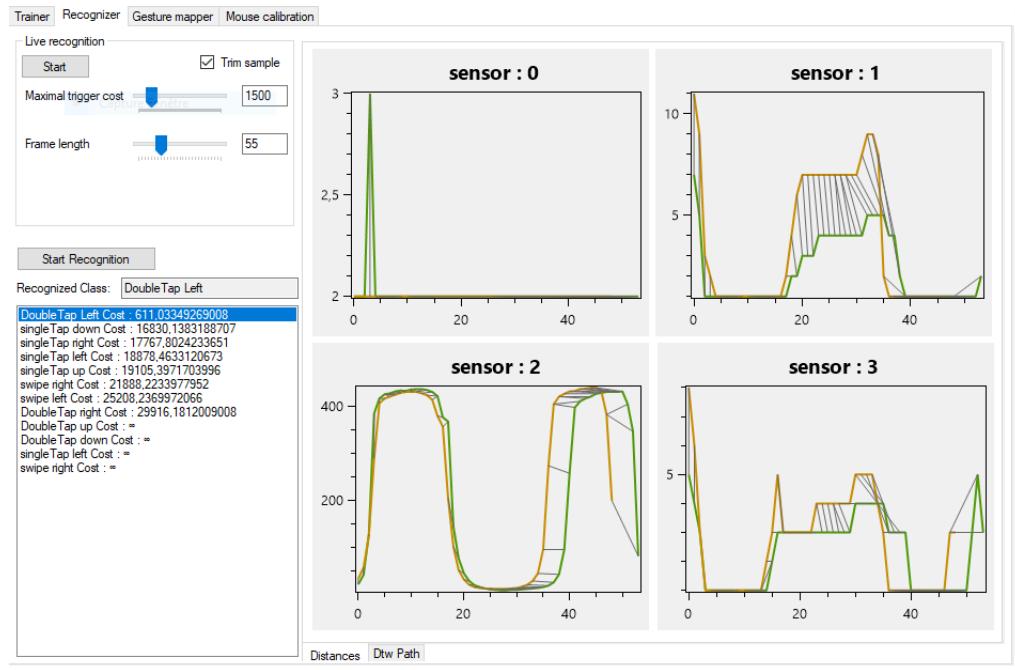
Abbildung 3.8.: Vorverarbeitung Beispiel

Die Abstandsberechnung bietet ein Maß an Ähnlichkeit zwischen der Eingabesequenz und der Referenzgeste. Dabei wird Dynamic Time Warping (DTW) verwendet.

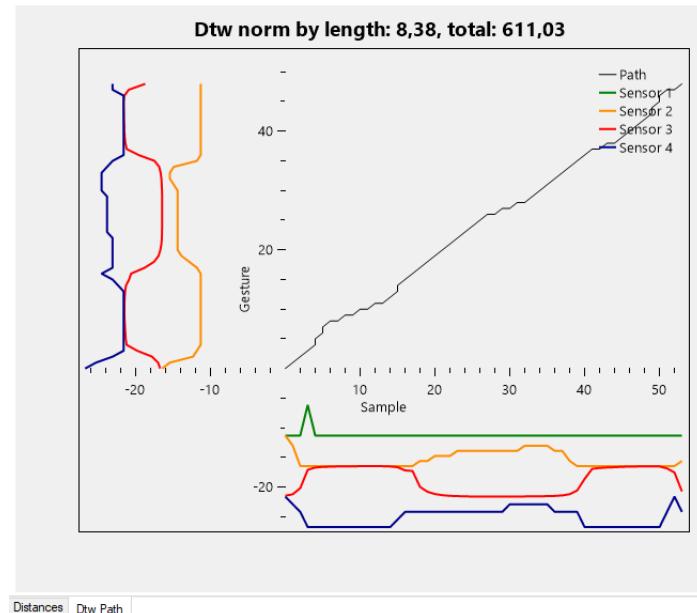
Eine Open-Source-Implementierung wurde in dieser Arbeit verwendet [Obl13].

Die Erkennung liefert die am besten passende Geste und einen Vertrauenswert.

3.2 Signalverarbeitung



(a) Distanzen Visualisierung



(b) Visualisierung des DTW-Pfades

Abbildung 3.9.: Erkennung-GUI

Abbildung von Gesten

Die erlernten Gesten können verschiedenen Aktionen zugeordnet werden:

- Das Drücken einer Taste oder Tastenkombination auf der Tastatur.
- Simulation eines Mausklicks oder einer Radaktion.
- Starten eines bestimmten Programms oder Ausführen eines Skripts.

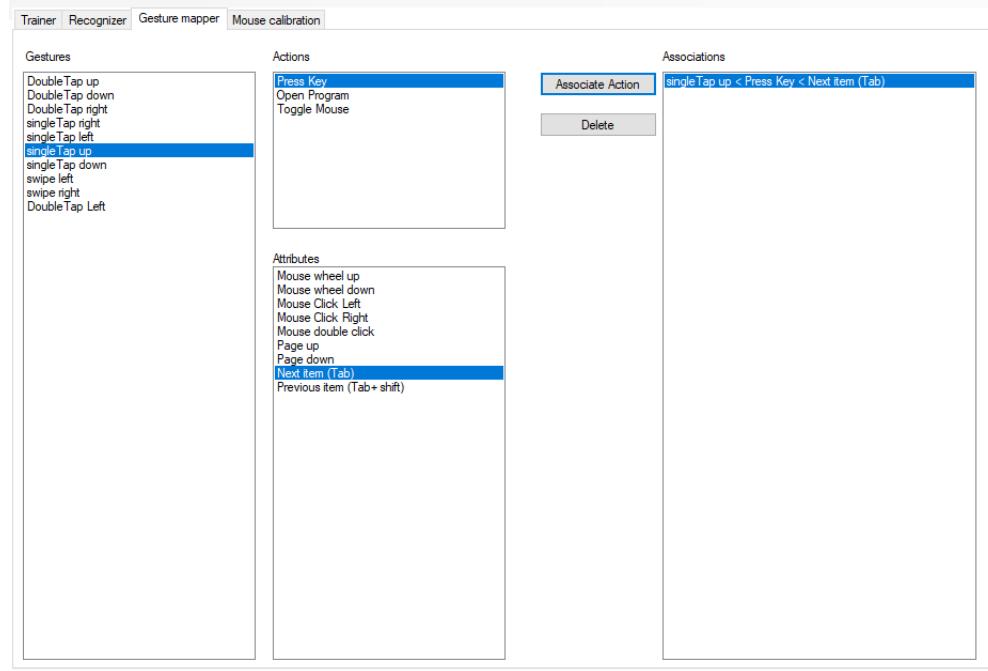


Abbildung 3.10.: Gestenzuordnung GUI

An example association can be the following :

- *Tap UP : ENTER*
- *Double Tap UP : BACK*
- *Tap LEFT : TAB+SHIFT*
- *Tap RIGHT : TAB*

Dieses Belegungskonzept basiert auf der reinen Tastatur-Navigation im Windows-Betriebssystem . In den meisten Anwendungen wird durch Drücken der Tab-Taste zum nächsten GUI-Element gesprungen. SHIFT+TAB findet das vorherige Element . Ein Mausklick setzt die GUI-Auswahl auf das nächstgelegene Element zurück.

Dies kann eine Navigationsmethode für viele bestehende Anwendungen ohne zusätzliche Software bieten.

Maussteuerung

Der implementierte Mauscontroller behandelt jeden Sensor als ein eindimensionales Joystick. Jedem Sensor ist eine Bewegungsrichtung zugeordnet: $+x$, $+y$, $-x$, $-y$

Überschreitet der von einem Sensor gemeldete Abstand einen bestimmten Schwellenwert, bewegt sich der Mauszeiger in die aktivierte Richtung. Die Geschwindigkeit des Mauszeigers ist proportional zur Zungenbewegungsgeschwindigkeit.

Es werden nur positive Geschwindigkeitsänderungen berücksichtigt. Dies ermöglicht dem Benutzer, die Zungenposition zurückzusetzen, ohne den Mauszeiger fälschlicherweise zu bewegen.

Die Software bietet eine GUI zur Kalibrierung der Empfindlichkeit und der Aktivierungsschwellen der Sensoren.

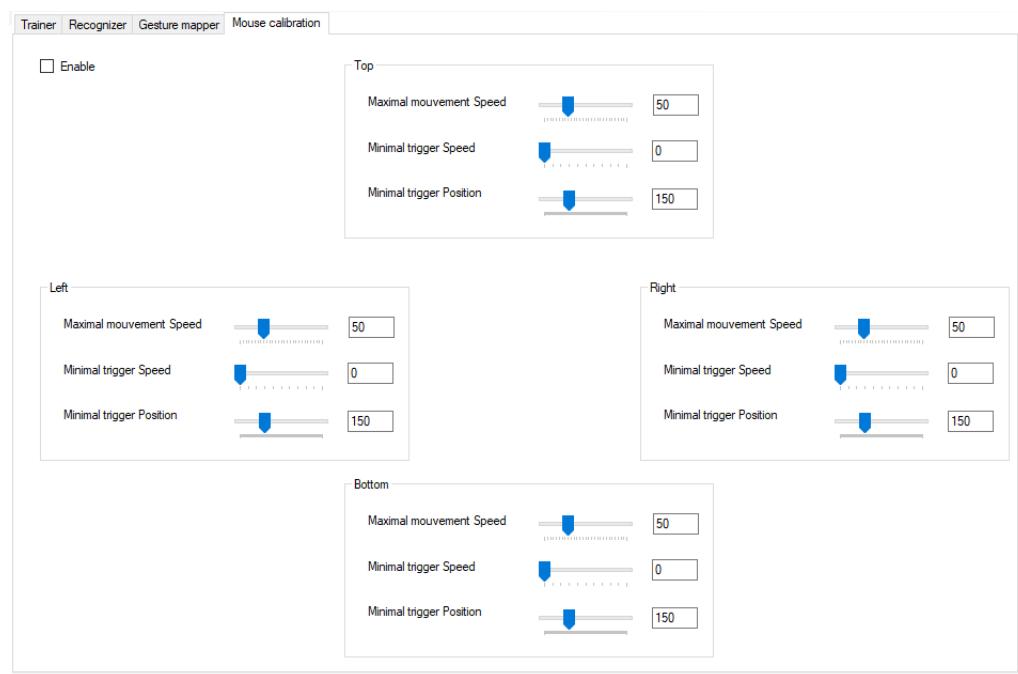


Abbildung 3.11.: Maus-Kalibrierung GUI

4. Evaluation

Ergebnisse

Es wurden 3 verschiedene Verdrahtungskonfigurationen getestet. Die Tabelle 4.1 zeigt die Abtastraten, die bei jeder Variante erreicht wurden.

Anzahl der I2C-Schnittstellen	Anzahl der Sensoren	Abtastrate
1	1	488 Hz
2	2	488 Hz
2	4	84 Hz

Tabelle 4.1.: Abtastraten

Unter Verwendung von 4 Sensoren, die mit 84 Hz abgefragt wurden, war es erfolgreich, zwischen den folgenden Gesten zu unterscheiden:

Short Taping , Double Taping , Long Taping , left swiping , right swiping.

Abbildung 4.1 illustriert Beispielaufnahmen bei 84 Hz.

Der Mauscontroller ermöglichte eine Unterscheidung zwischen langsamem, präzisen Cursorbewegungen und schnelleren Bewegungen zur Abdeckung größerer Strecken.

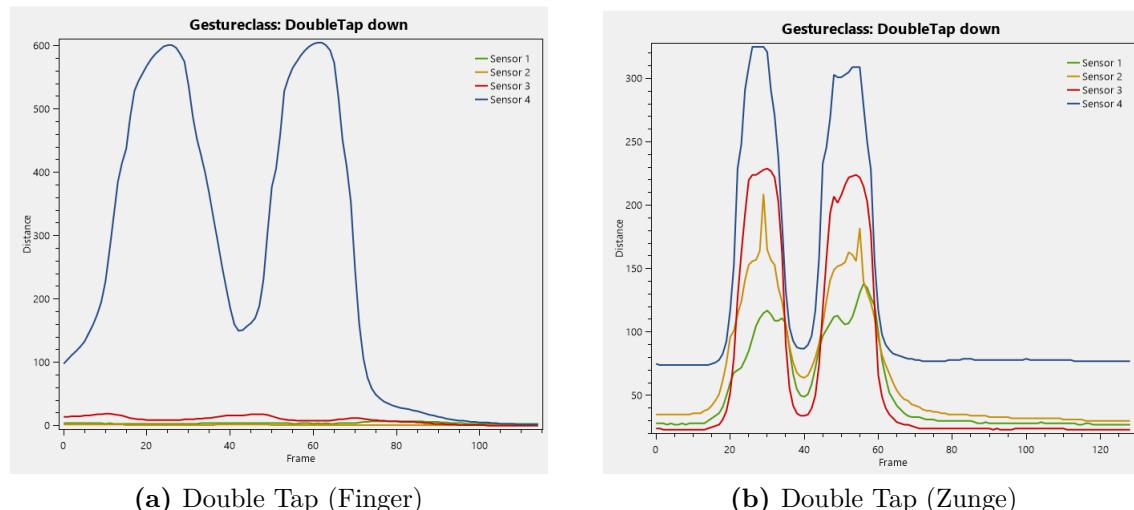


Abbildung 4.1.: Beispiel Aufnahme

Diskussion

Die Erkennung war genauer für Gesten, die durch Mundstrukturen wie z. B. die *Alveolarkammlinie* geführt wurden.

Das implementierte zeitliche Multiplexing ist langsam und für Sensorkonfigurationen mit mehr als 4 Einheiten nicht geeignet. Ein externer I2C-Multiplexer kann hier verwendet werden.

Die Verwendung einer adaptiven Scanrate kann eine kontextbasierte Rekonfiguration der Multiplexing-Konfiguration ermöglichen. Dies kann die aktuelle Implementierung in die Lage versetzen, bei Bedarf zwei Sensoren mit ihrer nativen Abtastrate abzufragen.

Die Implementierung der Gestenerkennung vergleicht eine Eingabesequenz mit allen Elementen aus der Gestensammlung. Dies führt zu einer langsameren Reaktionszeit, wenn mehrere Gesten und Wiederholungen eingegeben sind. Dies kann durch die Gruppierung von Gestenwiederholungen in einem repräsentativen Modell optimiert werden.

Es wurde ein Leiterplattenentwurf 4.2 unter Verwendung von 4- und 6-Sensor-Anordnungen versucht.

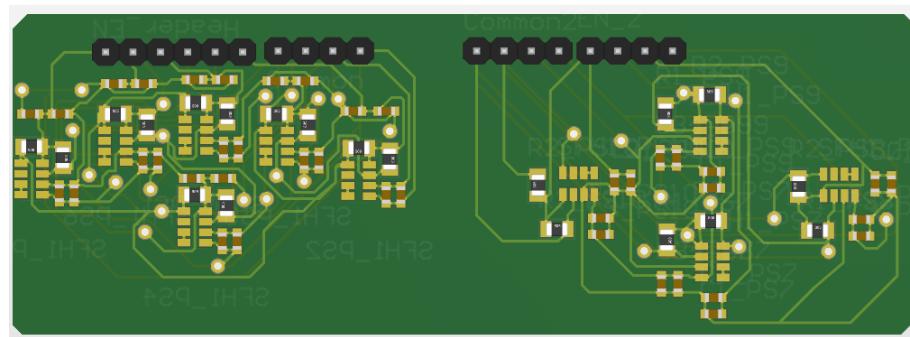


Abbildung 4.2.: Leiterplattenentwurf

Die 6-Sensor-Version enthielt aufgrund eines Versehens eine umgekehrte Bestückung und war nicht verwendbar. Die 4-Sensor-Version konnte nur einen Sensor ansprechen, da die Pullup-Widerstände des i2c-Busses nicht korrekt eingestellt waren. Ein besseres Leiterplattendesign hätte mehrere einzelne, kleinere Sensormodule beinhalten können, um die Zeit bei der Wiederherstellung der Module aus den ausgefallenen Leiterplatten zu sparen.

5. Zusammenfassung

In dieser Arbeit wurde ein Computer-Zungen-Interface entwickelt. Eine Signalerfassungseinheit wurde mit 4 Infrarot-Näherungssensoren aufgebaut. Eine Software wurde programmiert, um die Sensoren anzuschließen und ein Gestenerkennungssystem zu implementieren.

In einer ersten Evaluierung des Systems war es möglich, einfache, wiederholbare Zungengesten zu erkennen und diese auf verschiedene Funktionalitäten abzubilden. Auch die zuverlässige Steuerung einer Computermaus war ebenfalls möglich.

Diese Implementierung kann in ein kleines, billiges und drahtloses Gerät minituarisiert werden.

Literaturverzeichnis

- [CB18] Allen Cheng and R. Bryan Bell. Oral, head and neck oncology and reconstructive surgery. 2018, <https://www.sciencedirect.com/topics/medicine-and-dentistry/hypoglossal-nerve-injury>.
- [CW78] C Chuang and William Wang. Use of optical distance sensing to track tongue motion. *Journal of speech and hearing research*, 21: 484, 10 1978.
- [HK15] Robert Horne and Stephen Kelly. Electropalatography as a human input device for patients suffering from spinal cord injuries. 11 2015.
- [HWG08] X. Huo, J. Wang, and M. Ghovanloo. A magneto-inductive sensor based wireless tongue-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(5): 497–504, Oct 2008.
- [LYC⁺15] P. Leons, A. Yaghoubian, G. Cowan, J. Trajkovic, Y. Nazon, and S. Abdi. On improving the range of inductive proximity sensors for avionic applications. In *Sixteenth International Symposium on Quality Electronic Design*, pages 547–551, March 2015.
- [Obl13] Darjan Oblak. Dynamic Time Warping (DTW) algorithm implementation for .NET, January 2013.
- [SFH] Sfh7779 datasheet. Source: https://dammedia.osram.info/media/resource/hires/osram-dam2496468/SFH%207779_EN.pdf.
- [SKPT09] Scott Saponas, Daniel Kelly, Babak A. Parviz, and Desney Tan. Optically sensing tongue gestures for computer input. In *Proceedings ACM Symposium on User Interface Software and Technology*, pages 177–180. ACM, October 2009, <https://www.microsoft.com/en-us/research/publication/optically-sensing-tongue-gestures-computer-input/>.
- [Sol06] Nancy Solomon. What is orofacial fatigue and how does it affect function for swallowing and speech? *Seminars in speech and language*, 27: 268–82, 12 2006.

Abbildungsverzeichnis

2.1.	Electropalatograph[HK15]	4
2.2.	Induktive Näherungssensorik[LYC ⁺ 15].	5
2.3.	TDS Magnetisches Näherungssensorprinzip[HWG08].	5
2.4.	Optical sensing principle[CW78]	6
3.1.	Übersicht der Systemkomponenten	7
3.2.	Empfindlichkeitsdiagramm SFH7779 [SFH]	8
3.3.	Schaltplan der Sensoreinheit	8
3.4.	Abfragezyklus	9
3.5.	Evaluierungs-Hardware	10
3.6.	Sensorenanordnung	10
3.7.	Gestensammlung GUI	11
3.8.	Vorverarbeitung Beispiel	12
3.9.	Erkennung-GUI	13
3.10.	Gestenzuordnung GUI	14
3.11.	Maus-Kalibrierung GUI	15
4.1.	Beispiel Aufnahme	16
4.2.	Leiterplattenentwurf	17

A. Pinbelegung

2. Pinout Configuration

