

# **Interaktion Hands-Free. Eine Analyse alternativer Interaktionsmöglichkeiten**

Hanna Wagner



**BACHELORARBEIT**

Nr. S1410456033

eingereicht am  
Fachhochschul-Bachelorstudiengang  
Kommunikation, Wissen, Medien  
in Hagenberg

im Juni 2017

Diese Arbeit entstand im Rahmen des Gegenstands

**Course not specified!**

im

Sommersemester 2017

Betreuung:

FH-Prof. DI (FH) Dr. Mirjam Augstein

# Erklärung

Ich erkläre eidesstattlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benutzt und die den benutzten Quellen entnommenen Stellen als solche gekennzeichnet habe. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt.

Hagenberg, am 1. Juni 2017

Hanna Wagner

# Inhaltsverzeichnis

<b>Erklärung</b>	<b>iii</b>
<b>Kurzfassung</b>	<b>vii</b>
<b>Abstract</b>	<b>viii</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Zielsetzung . . . . .	1
<b>2 Eingabemethoden</b>	<b>2</b>
2.1 Sprachsteuerung . . . . .	2
2.2 Augensteuerung . . . . .	4
2.3 Gestensteuerung . . . . .	7
2.3.1 Kinnsteuerung . . . . .	7
2.3.2 Mundsteuerung . . . . .	9
2.3.3 Kopfsteuerung . . . . .	10
2.3.4 Fußsteuerung . . . . .	11
2.4 Muskelsteuerung . . . . .	12
2.5 Steuerung durch Gehirnaktivität . . . . .	15
<b>3 Ausgabemethoden</b>	<b>18</b>
3.1 Auditive Ausgabe . . . . .	18
3.2 Haptische Ausgabe . . . . .	20
3.3 Visuelle Ausgabe . . . . .	21
<b>4 Gegenüberstellung der Methoden</b>	<b>22</b>
4.1 Fördernde und hemmende Faktoren . . . . .	22
4.2 Einsatzgebiete . . . . .	28
<b>5 Fazit und Ausblick</b>	<b>29</b>
<b>Quellenverzeichnis</b>	<b>31</b>
Literatur . . . . .	31
Online-Quellen . . . . .	32

# Abbildungsverzeichnis

2.1	Hierarchie von Eye-Tracking-Systemen [3]. . . . .	5
2.2	Übersicht über die Eye-Tracking-Systeme von Tobii und SMI (a) Tobii Pro Glasses 2 [21], (b) Tobii Pro Spectrum [22] und (c) SMI Red250mobile [23]. . . . .	6
2.3	(a) Moso® Kinnsteuerung [18], (b) Smile Smart Kinnsteuerung [25] und (c) Sensory Kinnsteuerung [26]. . . . .	8
2.4	Sechs Freiheitsgrade [28]. . . . .	9
2.5	(a) IntegraMouse Plus [30] und (b) QuadStick Gamecontroller [29] . . .	10
2.6	(a) Frühwald Joystick [31], (b) Frühwald Fußsteuerung [31] und (c) Hidrex Fußsteuerung [32]. . . . .	12
2.7	EMG Signalverarbeitung [2] . . . . .	13
2.8	Trainingsoberflächen: 1D-Steuerung [15] (a) und 2D-Steuerung [15] (b) . . . . .	14
2.9	Myo Armband [34]. . . . .	15
2.10	BCI Signalverarbeitung [6]. . . . .	16
3.1	Unterteilung der Sprachausgabe [4]. . . . .	19

# Tabellenverzeichnis

4.1	Verschiedene Faktoren der Eingabemethoden . . . . .	24
-----	---	----

# Kurzfassung

# Abstract



# Kapitel 1

## Einleitung

1.1 Motivation

1.2 Zielsetzung

## Kapitel 2

# Eingabemethoden

Die meisten Interaktionen mit elektronischen Geräten jeder Art finden durch den Einsatz unserer Hände statt. Das reicht von der einfachen Bedienung von Geräten (z.B. Einschalten eines Computers) bis hin zur Steuerung von Abläufen mittels Gestik (z.B. Liederwechsel durch eine Wischgestik, die von einer Software erkannt und in einen Befehl umgewandelt wird). Doch nicht immer gibt es die Möglichkeit mit Hilfe der Hände zu interagieren. Beispielsweise sollen beim Autofahren beide Hände am Lenkrad bleiben, bei anstrengenden und präzisen Arbeiten an Maschinen müssen oft beide Hände benutzt werden oder aber auch für Menschen mit Tetraplegie bzw. Tetraparese, die ihre Hände und Arme nicht oder oft nur sehr eingeschränkt benutzen können, sind alternative Interaktionsmöglichkeiten äußerst nützlich. Aus diesem Grund beschäftigt sich dieses Kapitel mit einzelnen alternativen Eingabemethoden, die es auch ohne den Einsatz von den Händen ermöglichen, mit einem System interagieren zu können. Vor- und Nachteile, gesonderte Rahmenbedingungen und Einsatzgebiete werden in Kapitel 4 näher erläutert.

### 2.1 Sprachsteuerung

Eine Alternative zur Interaktion mit den Händen stellt die Steuerung mit Hilfe der menschlichen Stimme dar. Ein Spracherkennungssystem soll im optimalen Fall das Gesprochene genau so gut erkennen und verstehen können wie ein Mensch. Da jeder andere Sprachangewohnheiten hat, einen unterschiedlichen Dialekt spricht und es auch syntaktische Unterschiede gibt, kann ein System an das menschliche Verstehen allerdings nur angenähert werden. Oft werden daher in der Praxis die Systeme speziell für einzelne Szenarien entwickelt bzw. angepasst.

Wird der Inhalt des Gesprochen analysiert und zur Interaktion verwendet, wird dies als Spracherkennung bezeichnet. Hier gibt es zwei Ansätze, damit das System das Gesprochene versteht:

- Mustervergleich
- Statistische Spracherkennung

Bei einem Mustervergleich werden die gesprochenen Wörter mit den bereits abgespeicherten Mustern und Begriffen verglichen. Jenes Wort, das dem Eingabewort am ähn-

lichsten ist, wird verwendet. Hingegen bei der statistischen Spracherkennung werden Beschreibungen von Lauten und Wörtern verwendet. Darüber hinaus werden Statistiken verwendet, die Wahrscheinlichkeiten von Wortfolgen beinhalten [7].

Bevor das Signal im gesamten Prozess bei dem Analyseschritt angelangt, laufen vorher noch einige Schritte ab. Wenn eine Person etwas spricht werden die analogen Wellen durch elektroakustische Wandler (zum Beispiel ein Mikrophon) in ein elektrisches Signal (Sprachsignal) umgewandelt, damit der Computer diese interpretieren und weiterverarbeiten kann. Anschließend werden auf den gespeicherten Daten verschiedenste Filter angewendet, um störende Signale, wie zum Beispiel Hintergrundgeräusche oder Pausen während des Sprechens zu entfernen. Erst dann kann im nächsten Schritt das System das Gesprochene mit Hilfe eines Mustervergleichs oder einer statistischen Spracherkennung analysieren und so den Sinn erfassen. Ist der Inhalt des Gesprochenen bekannt, kann dieser in einen Text und somit in einen Befehl umgewandelt und dadurch der Computer bzw. das Programm gesteuert werden [7].

Es gibt viele Faktoren, die einen Einfluss auf das Sprachsignal haben. So hat jeder Mensch neben seiner einzigartigen Stimme einen unterschiedlichen Dialekt, Sprachgewohnheiten, Emotionen in der Sprache, Pausengewohnheiten und eine verschiedene Sprechgeschwindigkeit [7]. Diese Komponenten sind nicht nur wichtig für die Erkennung, was gesprochen wird, sondern viel mehr dafür, wer spricht (Sprechererkennung)

Um festzustellen zu können welche Person gerade spricht, muss es Referenzdaten im System geben, die mit dem aktuellen Signal verglichen werden können. Dies geschieht entweder, wenn die Person ein bestimmtes Signalwort spricht, das mit dem abgespeicherten Wort verglichen wird, oder wenn das Sprachsignal lange genug ist, kann auf Grund von statistischen Merkmalen eine Übereinstimmung gefunden werden [7].

Kommunikation durch Sprache ist eines der leichtesten, wenn nicht sogar das intuitivste Mittel zur Verständigung. Daher sollte auch die Interaktion mit einem Computer diesem hohen Standard entsprechen. Allerdings ist im Gegensatz zur gewohnten Face-to-Face-Kommunikation das Gegenüber im Bezug auf Spracherkennung kein intelligenter Gesprächspartner. Ein Computer kann nur schwer lange und komplizierte Sätze und Gedankengänge mitverfolgen und daraus schlecht eine Kernaussage zu ziehen. Daher muss dem System verständlich gemacht werden, wann es zuhören soll [14]. Zwei gängige Methoden für den Start zur Interaktion sind:

- das Drücken einer Starttaste
- das Aussprechen eines Aktivierungswortes

Weil das Drücken eines Startknopfes die Hilfe der Hände benötigt, steht in diesem Zusammenhang das Aussprechen eines Aktivierungswortes zum Start der Interaktion im Vordergrund. Durch das Wort 'Alexa' wird beispielsweise der Amazon Echo <sup>1</sup> aktiviert. Bei Amazon Echo handelt es sich um ein Sprachinteraktionssystem, das beispielsweise Musik abspielen, den Einkauf bestellen oder auch Nachrichten und Wetterbeiträge zur

---

<sup>1</sup><https://www.amazon.com/Amazon-Echo-Bluetooth-Speaker-with-WiFi-Alexa/dp/B00X4WHP5E>

Verfügung stellen kann. Im Gegensatz zur Benutzung von Amazon Echo, muss bei der Sprachsteuerung von Apple (Siri)<sup>2</sup> das System erst auf die eigene Stimme eingestellt, also kalibriert werden. Das Aktivierungswort 'Hey Siri' muss drei Mal einzeln und zwei Mal in einem Satz gesprochen werden, damit diese Funktion beispielsweise am iPhone freigeschaltet wird. Anschließend können ebenfalls Wetter oder Nachrichtendaten abgefragt werden, aber auch Anrufe oder Nachrichten können mit Hilfe der Stimme getätigt werden.

Sobald das System bereit ist den Input aufzunehmen, muss folgendes beachtet werden:

- keine zu große Distanz zum Mikrophon
- laut und deutlich zu sprechen
- nicht zu schnell zu sprechen

Zusammenfassend lässt sich sagen, dass die Sprachinteraktion mit dem Computer aus Benutzersicht noch einige Verbesserung braucht, damit die Bedienung natürlicher und intuitiver wird und der einer Face-to-Face-Kommunikation näher kommt.

## 2.2 Augensteuerung

Eine weitere alternative Interaktionsmöglichkeit stellt die Steuerung mit Hilfe der Augen dar. Eye-Tracking-Systeme werden verwendet, um die Augenbewegungen zu verfolgen. Parallel dazu wandelt eine Software die Augenbewegungen in einen Mauszeiger um, um so mit dem System interagieren zu können.

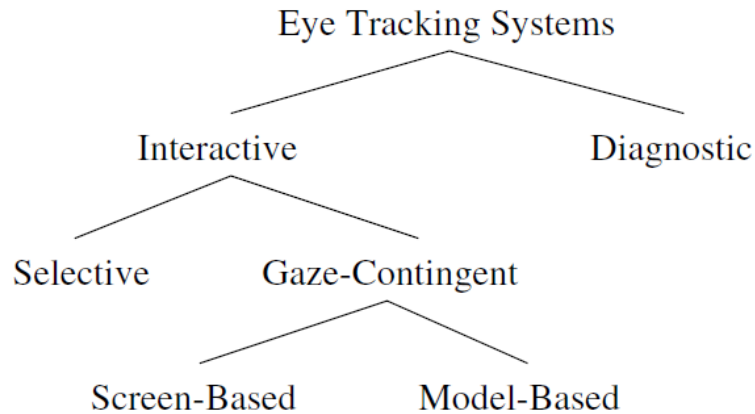
Um die Augenbewegungen nachvollziehen zu können, werden sowohl ein Infrarotlichtstrahl, als auch eine Videokamera auf das Auge fokussiert. Eine Software analysiert die Bewegungen im Hintergrund und erkennt, wo der Fokus auf dem Bildschirm liegt. Für die Umrechnung ist außerdem wichtig, dass die Position des Kopfes vom Interaktionssystem erkannt wird [11].

Die Teilbereiche des Auges bzw. des menschlichen Sehens, die für die Augensteuerung relevant sind, sind das periphere und foveale Sehen. Das foveale Sehen beinhaltet jenen Teil, den das menschliche Auge als scharfen Bildausschnitt wahrnimmt, also auf welchen Ausschnitt es fokussiert. Hingegen enthält das periphere Sehen den Großteil unserer Seh Wahrnehmung. Dies beinhaltet jene Bereiche, die nicht fokussiert sind [11]. Daher ist das foveale Sehen für die Augensteuerung relevant.

Nielsen und Pernice [11] unterscheiden beim menschlichen Sehen zusätzlich zwischen Fixation und Sakkaden. Das menschliche Auge bewegt sich nicht kontinuierlich, sondern ein Blick setzt sich aus schnellen Bewegungen mit Pausen dazwischen zusammen. Dies geschieht so schnell, dass es im Alltag nicht bewusst wahrgenommen wird. Wenn das Auge auf einem bestimmten Punkt ruht spricht man von Fixation. Die schnellen Zwischenbewegungen von einer Fixation zur nächsten werden Sakkaden genannt. Die

---

<sup>2</sup><https://www.apple.com/ios/siri/>



**Abbildung 2.1:** Hierarchie von Eye-Tracking-Systemen [3].

Schwierigkeiten der Eye-Tracking-Systeme und der Augensteuerungssoftware bestehen darin, diese Fixationen zu erkennen, da diese meist zwischen einer hundertstel und einer zehntel Sekunde liegen [11].

Wie in Abb. 2.1 dargestellt, unterscheidet Duchowski [3] zwischen interaktiven und diagnostischen Systemen. Da bei diagnostischen Systemen keine direkte Interaktion stattfindet, sondern die Augenbewegungen aufgezeichnet und im Nachhinein ausgewertet werden, sind für die Analyse von Interaktionsmöglichkeiten ausschließlich interaktive Systeme relevant.

Interaktive Systeme lassen sich wiederum in zwei Bereiche teilen. Selektive Systeme benutzen den Blickpunkt als analoges Eingabegerät, während Blickkontingentsysteme für die Darstellung von komplexen Displays verwendet werden.

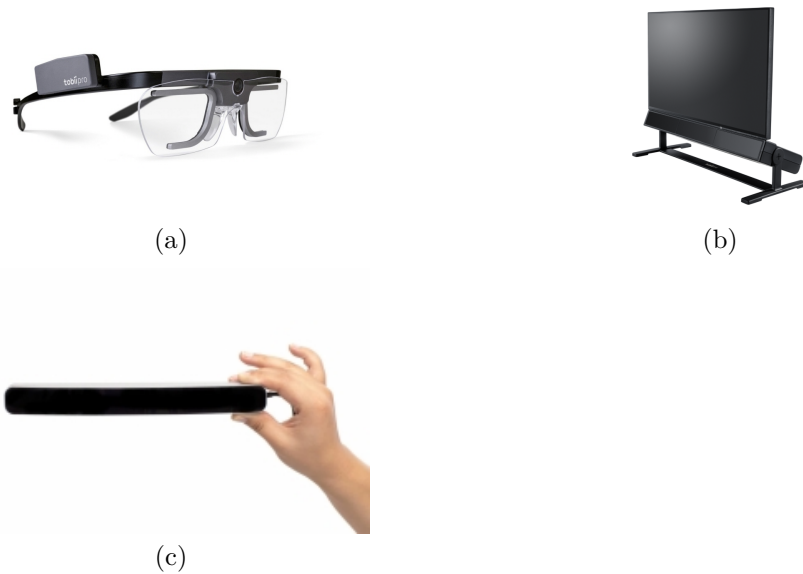
Abgesehen davon, wie die Daten aufgezeichnet werden bzw. was aufgezeichnet wird, ist die Unterscheidung zwischen den Eye-Tracking-Systeme relevant. Diese können in intrusive und nicht-intrusive Systeme unterteilt werden. Intrusive Eye-Tracker haben einen direkten Kontakt mit den Anwendern (z.B. durch Kontaktlinsen). Nicht-intrusive Eye-Tracking-Systeme hingegen messen die Blicke mit Hilfe einer oder mehrerer Kameras.

In der Praxis werden eine Vielzahl verschiedener Eye-Tracking-Systeme als Basis für die Augensteuerung verwendet. Die Unternehmen Tobii <sup>3</sup> und SensoMotoricInsturmens (SMI) <sup>4</sup> sind Spitzenreiter auf dem Gebiet des Eye-Trackings und verfügen über eine umfangreiche Produktpalette bzw. Anwendungsgebiete.

In Abb. 2.2(a) ist die tragbare Eye-Tracking-Brille Tobii Pro Glasses 2 zu sehen. Dieses ist so gestaltet worden, um ein möglichst natürliches Nutzungsverhalten zu erzielen.

<sup>3</sup><http://www.tobii.com/group/>

<sup>4</sup><https://www.smivision.com/>



**Abbildung 2.2:** Übersicht über die Eye-Tracking-Systeme von Tobii und SMI  
(a) Tobii Pro Glasses 2 [21], (b) Tobii Pro Spectrum [22] und (c) SMI Red250mobile [23].

Es müssen keine Voreinstellungen vorgenommen werden, die Brille kann sofort benutzt werden. Bei Tobii Pro Spectrum (Abb. 2.2(b)) handelt es sich um ein nicht-intrusives System, das die Blickdaten mit einer Geschwindigkeit von bis zu 600 Hz aufnimmt. Das Unternehmen SMI hat, wie in Abb. 2.2(c) zu sehen ist, ein mobiles Eye-Tracking-System entwickelt, das die Bewegungen mit bis zu 250 Hz aufzeichnen kann. Alle diese Systeme verwenden das Prinzip von Fixation und Sakkaden des Auges, um die Augenbewegungen nachvollziehen zu können. Damit nicht nur die Augenbewegungen aufgezeichnet werden, müssen diese Systeme in Kombination mit einer Software eingesetzt werden, die die Blicke in Mausbewegungen o.ä. umwandelt.

Bevor ein Benutzer mit Hilfe der Augen beispielsweise eine Bildschirmtastatur bedienen kann, muss das System bzw. die Software erst auf den Anwender eingestellt werden. Bei der Kalibrierung tauchen am Bildschirm nacheinander Symbole auf, die fixiert werden müssen. Dadurch kann der Blickwinkel und der Augenabstand vermessen werden um eine optimale Interaktion zu erreichen. Bei der eigentlichen Anwendung erkennt das System dann, wenn der Blick auf einem Punkt länger verharret, also scharf stellt. Verharret der Blick also länger auf einem Buchstaben der Tastatur, wird dies als Mausklick gewertet und der Buchstabe oder das Symbol ausgewählt. Am Beispiel von Seetech von Humanelektronik <sup>5</sup> beträgt die Fixationszeit etwa 0.5 bis 1.5 Sekunden [24].

Zusammenfassend kann gesagt werden, dass für die Interaktion mit einer Anwendung bzw. einem System interaktive Eye-Tracking-Systeme für die Nachvollziehbarkeit der Augenbewegungen und eine Software für die Umrechnung in Mausbewegungen relevant

---

<sup>5</sup><http://humanelektronik.de/>

sind. Nach einer kurzen Kalibrierung können durch Fixation einzelne Elemente auf dem Bildschirm ausgewählt und auf Grund der Fixationszeit in einen Mausklick umgewandelt werden.

## 2.3 Gestensteuerung

Eine weitere Alternative zur Interaktion mit den Händen stellt die Gestensteuerung<sup>6</sup> dar. Gesten bzw. Gestik werden unterschiedlich definiert, in diesem Zusammenhang versteht man unter einer Geste [13]:

...die Bewegung von Fingern, Händen und Armen - oder auch weiterer Körperteile, wie Kopf, Augen und Lippen - aufgrund einer kommunikativen Absicht. Damit enthält die Bewegung als solche signifikante Informationen, die an den Computer übermittelt werden sollen.

In diesem Abschnitt werden Kinn-, Mund-, Kopf- sowie Fußsteuerung näher erläutert.

### 2.3.1 Kinnsteuerung

Bei der Kinnsteuerung werden mit Hilfe des Kinns Bewegungen ausgeführt, um mit einem System interagieren zu können. Die unterschiedlichen Möglichkeiten werden anhand der nachfolgenden Beispiele erklärt.

Wie in Abb. 2.3(a) zu sehen ist verwendet das Unternehmen moso®<sup>7</sup> einen Joystick und zwei Tasten für die Interaktion. Um einen angenehmen Umgang zu gewährleisten, ist der Joystick mit einem kleinen weichen Ball ausgestattet. Zwei Tasten sind links und rechts des Joysticks angebracht. Der Joystick kann sich in alle Himmelsrichtungen bewegen und kann so entweder als Mauszeiger, oder zur Steuerung eines Rollstuhls genutzt werden. Die beiden Tasten können als linke und rechte Maustaste verwendet werden [18].

Abb. 2.3(b) zeigt die Version der Firma Smile Smart Technology<sup>8</sup>. Diese besteht aus einem Mini Rollstuhl Joystick, der auf einer schwenkbaren Halterung montiert ist [25]. Die Kinnsteuerung des Unternehmens Sensory Guru<sup>9</sup> weist im Vergleich zu den anderen beiden keinen Ball auf. Wie in Abb. 2.3(c) dargestellt, weist der Joystick eine Mulde für das Kinn auf. Mit einer Bewegung nach links oder rechts können die beiden äußeren Sensoren aktiviert werden. Das Kinn muss dabei im ständigen Kontakt mit dem Ball bzw. Joystick stehen [26].

Durch eine starke Kinnbewegung nach links bzw. Rechts oder über zusätzlich angebrachte Tasten an der Steuerung, kann durch ein Menü navigiert werden oder auch andere Symbole auf einem Bildschirm ausgewählt werden. Durch die Bedienung des

---

<sup>6</sup>Mit Hilfe von Gesten kann eine Handlung mit verschiedenen Teilen des Körpers (z.B: den Händen, Füßen, Kopf) zum Ausdruck gebracht werden. In diesem Abschnitt bezieht sich die Gestensteuerung auf alle Körperteile ausschließlich der Hände.

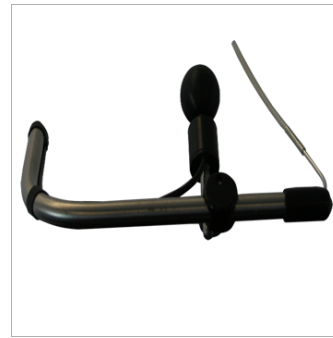
<sup>7</sup><http://www.moso-gmbh.de/>

<sup>8</sup><https://smilesmart-tech.com/>

<sup>9</sup><http://www.sensoryguru.com/>



(a)



(b)



(c)

**Abbildung 2.3:** (a) Moso® Kinnsteuerung [18], (b) Smile Smart Kinnsteuerung [25] und (c) Sensory Kinnsteuerung [26].

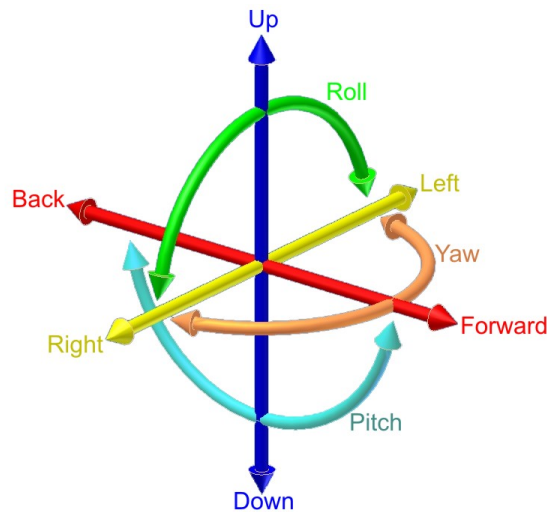
Joystickelementes kann der Benutzer beispielsweise die Bewegungen einer Computermaus nachahmen [18] [26].

Dies kann je nach Ausführung der Steuerung verschieden viele Freiheitsgrade beinhalten. Freiheitsgrade bzw. Degrees of Freedom (DoF) bezeichnen die Bewegungsfreiheit eines Körpers. Abb. 2.4 zeigt die sechs Freiheitsgrade bzw. sechs verschiedene Bewegungsrichtungen [27]:

- Rauf und runter (Up and Down)
- Links und rechts (Left and Right)
- Vorwärts und Rückwärts (Forward and Back)
- Kippen entlang der X-Achse (Roll)
- Kippen entlang der Y-Achse (Pitch)
- Kippen entlang der Z-Achse (Yaw)

Bei der Kinnsteuerung von Mosco werden fünf Freiheitsgrade ermöglicht, dass bedeutet dass der Joystick in alle Himmelsrichtungen und Neigungen bewegt werden kann, das Kippen entlang der Z-Achse ist allerdings nicht möglich. Bei der Ausführung des Unternehmens Sensory sind nur vier Freiheitsgrade möglich, da der Joystick nicht in Z-Richtung bewegt werden kann. Wie stark der Joystick in die bestimmte Richtung





**Abbildung 2.4:** Sechs Freiheitsgrade [28].

gedrückt werden muss, um eine gewisse Bewegung zu erzeugen, kann an den jeweiligen Benutzer angepasst werden, um eine optimale Interaktion zu erreichen.

### 2.3.2 Mundsteuerung

Das Prinzip der Mundsteuerung funktioniert ähnlich wie die Kinnsteuerung. Durch die verschiedenen Interaktionsmöglichkeiten, die durch den Einsatz des Mundes möglich sind, wird versucht, eine Interaktion, beispielsweise mit einem Computer, möglich zu machen.

Bei der in Abb. 2.5(a) gezeigten Mundsteuerung handelt es sich um die IntegraMouse Plus, die das Unternehmen LIFEtool<sup>10</sup> entwickelt hat. Die IntegraMouse Plus wird durch Nippen, Pusten und Bewegen des Mundes gesteuert. Wird die Maus mit einem Computer verbunden, so kann durch ein Nippen ein Linksklick erzeugt werden, durch Pusten ein Rechtsklick und durch das Bewegen des Mundstückes kann die Richtung bzw. der Cursor verändert werden [19].

Ähnlich wie bei der IntegraMouse Plus wird bei dem mundgesteuerten Gamecontroller der Firma QuadStick<sup>11</sup>, der in Abb. 2.5(b) zu sehen ist, ebenfalls nippen, pusten und die Bewegungen des Mundes zur Interaktion verwendet. Dieses Modell besitzt allerdings vier Nipp- bzw. Pustsensoren, um einen regulären Gamecontroller besser imitieren zu können. Die drei zusätzlichen Sensoren können individuell angepasst und beispielsweise als Shortcuts für bestimmte Computerspiele verwendet werden. Ein Shortcut bzw. ein Sensor beinhaltet eine Reihenfolge von Klick- und Bewegungsabläufen, die als Kombination dargestellt werden können [29]. In einem gewöhnlichen Computersetting könnten die zusätzlichen Sensoren beispielsweise als verschiedene Wischgesten mit einem oder

<sup>10</sup><http://www.lifetool.at/startseite/>

<sup>11</sup><http://www.quadstick.com/>



**Abbildung 2.5:** (a) IntegraMouse Plus [30] und (b) QuadStick Gamecontroller [29]

mehreren Fingern verwendet werden.

Bei beiden Steuerungen ist vor der Benutzung keine Kalibrierung notwendig. Lediglich muss ein USB-Stick am Computer angeschlossen werden, der die Mundsteuerung per Bluetooth-Verbindung ermöglicht.

### 2.3.3 Kopfsteuerung

Es gibt zwei verschiedene Arten, wie man nur durch den Einsatz des Kopfes mit einem System interagieren kann. Zum einen kann eine Kamera verwendet werden, um die Bewegungen des gesamten Gesichtes oder einen bestimmten Teil davon (z.B. den Mund) zu verfolgen und diese dann anschließend von einem System in die gewünschten Befehle umzuwandeln. Andererseits kann auch eine Vorrichtung am Kopf befestigt werden, dann läuft die Interaktion haptisch ab.

Bei der haptischen Variante könnte beispielsweise ein Stirnband am Kopf befestigt werden, das mit je einem Sensor links und rechts ausgestattet ist. Wird der Kopf in eine Richtung geneigt, wird einer der beiden Sensoren aktiviert. Das Stirnband müsste dafür über Bluetooth oder WLAN mit einem Computer verbunden sein, damit eine Software die Bewegungen bzw. Sensoren erkennen und auswerten kann. Damit das Stirnband benutzt werden kann, müsste es individuell angepasst werden, da die seitliche Neigung des Kopfes je nach Benutzer unterschiedlich ist.

Bei der optischen Variante werden grundlegend zwei Arten der Kopfsteuerung bzw. der Nachvollziehbarkeit der Bewegungen unterschieden:

- Gesichtserkennung
- Punkteverfolgung

Bei der Gesichtserkennung wird versucht das Gesicht als Gesamtheit zu erfassen und

dieses als Berechnungsgrundlage für die Bewegungen heranzuziehen. Es gibt verschiedene Arten, um die Position des Kopfes zu erkennen. So kann beispielsweise die Hautfarbe dazu verwendet werden. Da diese von Mensch zu Mensch verschieden ist, ist es abzurufen sie als Erkennungsgrundlage heranzuziehen. Des Weiteren sind bei der Berechnung bzw. Filterung viele Schritte notwendig, was mit großem Rechenaufwand verbunden ist. Daher ist eine Mischung aus der haarähnlichen-Gesichtserkennung und dem Camshift-Gesichtstrackingalgorithmus effizienter.

Für letztere sind grob vier Schritte notwendig [12]:

- Als Ausgangslage wird der Unterschied vom Haaransatz zum Gesicht verwendet. Hierbei wird das Gesicht in verschiedene Rechtecke aufgeteilt.
- Zusätzlich wird ein Farbgesichtsdarstellungshistogramm erzeugt.
- Danach wird eine Gesichtswahrscheinlichkeitsberechnung für jedes Pixel durchgeführt.
- Abschließend werden die Größe und die Winkel berechnet.

Bei der Punkteverfolgung wird ein bestimmter Teilbereich des Gesichtes fokussiert, der die Bewegungsgrundlage des gesamten Kopfes darstellt. Naizhong und Jing [10] haben einen Algorithmus entwickelt, der die Mundbewegungen als Grundlage für die Kopfbewegungen verwendet. Dafür wird eine Kamera als Videoinput verwendet. In der Analyse können die verschiedenen Eingabebefehle unterschieden werden. So ergibt beispielsweise ein Kopfschütteln nach Links einen Linksklick oder ein zweimaliges Kopfschütteln nach Rechts einen rechten Doppelklick [10].

Für den Benutzer gibt es einige Punkte, die vor und während der Interaktion ausschlaggebend für eine optimale Verwendung sind. Ähnlich wie bei der Augensteuerung in Abschnitt 4.1 muss auch hier das System vorab kalibriert werden, um eine reibungslose Interaktion zu ermöglichen. Des Weiteren gilt es zu beachten, dass keine störenden Lichtquellen auf das Gesicht fallen und ein möglichst frontaler Blick in die Kamera erfolgt [10].

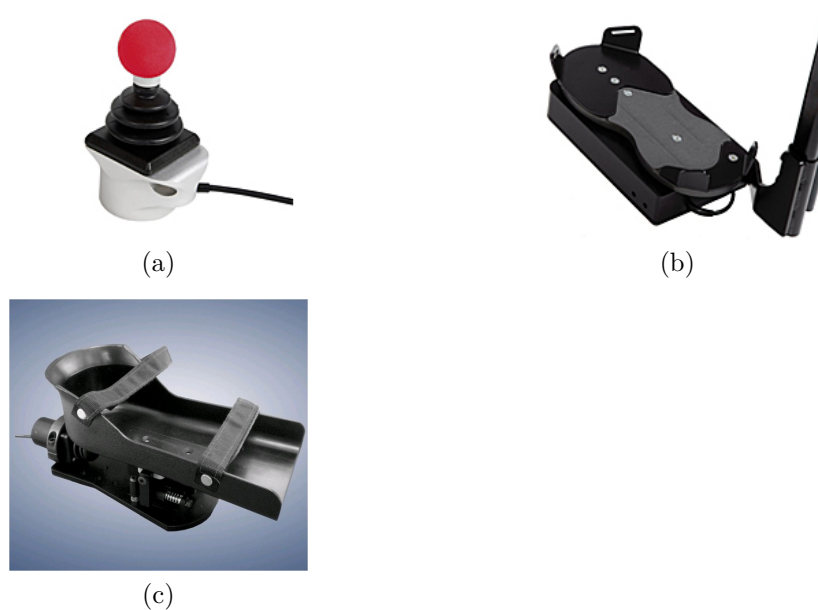
### 2.3.4 Fußsteuerung

Es gibt eine Vielzahl an verschiedenen Fußsteuerungen. Manche ähneln dem System der Kinnsteuerung und verwenden einen Joystick, um beispielsweise Mausbewegungen zu erzeugen. Darüber hinaus werden verschiedene Tasten für die Auswahl bestimmter Elemente genutzt. Andere wiederum verwenden einen zusätzlichen Schuh bzw. eine Halterung um den Fuß, um die Fußbewegungen zu messen.

Wie in Abb. 2.6(a) zu sehen ist, kann ein gewöhnlicher Joystick, der auch mit dem Kinn bedienbar ist, ebenfalls mit dem Fuß bedient werden. Der Benutzer bewegt den Joystick in die gewünschte Richtung und kann damit beispielsweise die Fahrtrichtung eines Rollstuhls oder die Mausbewegung am Computer steuern. Die Fußsteuerung bzw. das Fußpedal (Abb. 2.6(b)) des Unternehmens Frühwald<sup>12</sup> kann durch unterschiedlich starke Druckausübung zur Steuerung verwendet werden [31].

---

<sup>12</sup><https://www.fruehwald.net/>



**Abbildung 2.6:** (a) Früwald Joystick [31], (b) Früwald Fußsteuerung [31] und (c) Hidrex Fußsteuerung [32].

Bei der in Abb. 2.6(c) dargestellten Fußsteuerung der Firma Hidrex <sup>13</sup>, handelt es sich um einen Schuhhalter, der als Joystick funktioniert. Die Bedienung ist ähnlich der eines herkömmlichen Joysticks. Der Fuß muss in die gewünschte Richtung gezogen bzw. bewegt werden, um einen Bewegungsablauf nachvollziehen zu können.

Wie die einzelnen Systeme die verschiedenen Gestiken in eine Bewegung umwandeln, ist von System zu System verschieden. Häufig findet ein joystickähnliches Element für die Bewegungsrichtung Verwendung. Zusätzliche Tasten helfen bei der Auswahl von gewünschten Symbolen oder Elementen.

## 2.4 Muskelsteuerung

Bei der Muskelsteuerung wird das Zusammenziehen eines Muskels oder bestimmter Muskelpartien verwendet, um so mit einem Computer interagieren zu können.

Der Großteil der Muskelsteuerungssysteme nützt zur Messung der Signale das Prinzip der Elektromyografie (EMG). Bei einer EMG-Schnittstelle handelt es sich um die Messung der Muskelkontraktionen und verknüpft diese mit der gewünschten Interaktion. Die Voraussetzung an den Benutzer besteht lediglich darin, die ausgewählten Muskelpartien selbstständig anzuspannen bzw. zu aktivieren. Im Zusammenhang mit der Muskelsteuerung kommt die Oberflächenelektromyografie (sEMG, englisch für Surface Electromyography) zum Einsatz. Hierbei werden die Elektroden, die für die Messung der Signale ausschlaggebend sind, an der Hautoberfläche angebracht. Mit Hilfe der Elek-

<sup>13</sup><http://www.hidrex.de/>

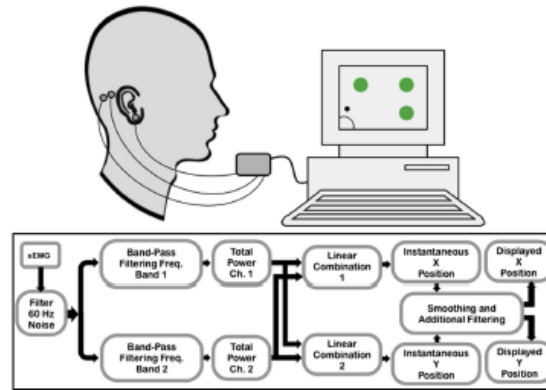


Abbildung 2.7: EMG Signalverarbeitung [2]

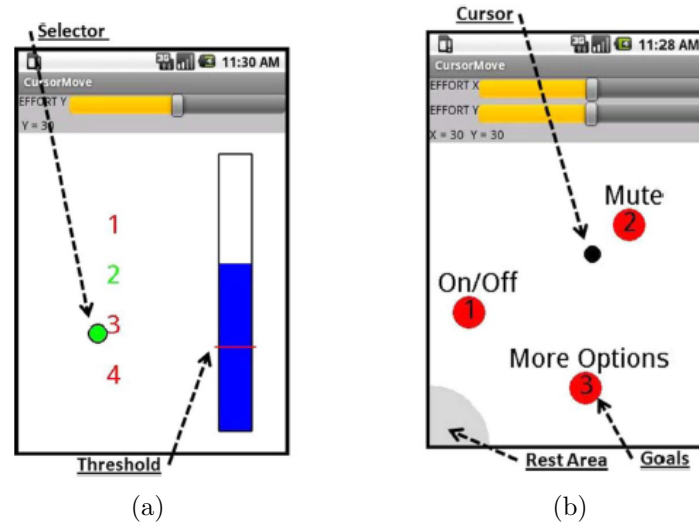
troden können die myoelektrischen Signale, die die Muskeln produzieren, gemessen und in einem späteren Schritt als bestimmte Interaktion interpretiert werden [5].

Abb. 2.7 zeigt die einzelnen Schritte, die das Signal im System durchläuft, um am Ende die Interaktion mit einem Computer zu ermöglichen. In der Studie von Barszap, Skavhaug und Joshi [2] werden zwei Frequenzbänder verwendet, um eine X- und Y-Richtung des Mauszeigers zu ermöglichen. Zieht sich der Muskel zusammen, so werden alle messbaren Signale erfasst und die relevanten herausgefiltert. Anschließend werden die Signale in zwei Partien aufgeteilt und durch zwei separate Bandpassfilter geschickt. Dadurch werden die gewünschten Signale bzw. Frequenzen erneut gefiltert und angepasst. Danach werden die beiden Signale in einen Wert zwischen 0 und 100 umgerechnet, der die Ausprägungsstärke in Prozent der jeweiligen Koordinate repräsentiert [2].

Damit ein Benutzer durch den Einsatz von myoelektrischen Signalen mit einem Computer interagieren kann, müssen zu Beginn die richtigen Positionen für die Anbringungen der Elektroden gefunden werden.

Es gibt verschiedene Ansätze, wie die Kalibrierung im Detail aussehen kann. Folgendes Beispiel bezieht sich auf eine Studie von Vernon und Joshi [15].

Die Elektroden werden an der vermuteten Stelle der zu messenden Muskeln angebracht und mit dem Computer verbunden. Auf dem Bildschirm ist das aktuelle sEMG-Signal zu sehen. Beinhaltet das Signal viele Störfaktoren, das sog. Rauschen, müssen die Elektroden gereinigt oder neu positioniert werden. Ist die optimale Position gefunden worden, muss das System an die individuellen Muskelkontraktionen des Anwenders angepasst werden. Bei diesem Schritt muss der Benutzer die ausgewählten Muskeln über einen Zeitraum von ein paar Sekunden so fest wie möglich anspannen, um die maximale Frequenz bzw. Leistung erfassen zu können. Als nächstes werden beim Benutzer zwei verschiedene Testphasen durchgeführt. Der erste Test (Abb. 2.8(a)) dient dazu, ein Bewusstsein zu entwickeln, wann und wie stark der gewünschte Muskel angespannt werden sollen. Der grüne Punkt auf dem Display bewegt sich je nach Anspannung durch die verschiedenen Stufen und soll auch durch Entspannen wieder zur Ausgangslage gebracht werden. Bei



**Abbildung 2.8:** Trainingsoberflächen:  
1D-Steuerung [15] (a) und 2D-Steuerung [15] (b)

dem zweiten Test, der in Abb. 2.8(b) dargestellt ist, handelt es sich um eine zweidimensionale Überprüfung. Hier soll der Cursor in X- und Y-Richtung bewegt werden. Dies geschieht ebenfalls durch An- und Entspannung der einzelnen Muskeln. Beide Versuche bauen auf dem Trial und Error Prinzip auf. Der Benutzer muss also durch mehrere Versuche zur richtigen bzw. besten Lösung kommen. Ist der Anwender mit beiden Trainings vertraut, kann mit der gewünschten Interaktion mit dem Computer begonnen werden. Diese ist auf dem selben Prinzip aufgebaut, allerdings gibt es unterschiedliche Oberflächen, Anordnungen der Elemente und Unterscheidungen in der Aufgabenstellung selbst [15].

Abb. 2.9 zeigt das Armband des Unternehmens Myo<sup>14</sup>. Dieses verwendet die myoelektrischen Signale des Armes, um durch verschiedene Gesten mit dem Computer interagieren zu können. Zu den fünf vordefinierten Gesten zählen z.B. das Machen einer Faust oder Wischgesten. Es können aber auch eigene Gesten definiert bzw. programmiert werden. Zum einen können die vordefinierten Gesten durch die Einberechnung von Armbewegungen erweitert werden, zum anderen gibt es die Möglichkeit neue Gesten zu erstellen, da der Zugriff auf die rohen EMG-Daten möglich ist [33].

Damit ein Computer mit Hilfe von Muskelkontraktionen gesteuert werden kann, ist ein aufwendiger und gut überlegter Algorithmus notwendig, um die erfassten Signale in das gewünschte Ergebnis zu transferieren. Vor der Interaktion müssen einige Schritte

<sup>14</sup><https://www.myo.com/>



**Abbildung 2.9:** Myo Armband [34].

durchlaufen werden, um das System an das Individuum anzupassen und um ein Bewusstsein für die Muskelsteuerung zu erlangen.

## 2.5 Steuerung durch Gehirnaktivität

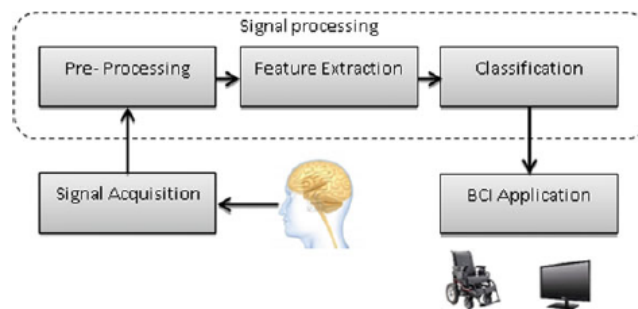
Dieses Kapitel beschäftigt sich mit der Schnittstelle zwischen dem Gehirn und dem Computer. In der Literatur wird in diesem Zusammenhang von Brain-Computer-Interface (BCI) gesprochen. Ziel ist es, eine direkte Verbindung der beiden Komponenten für eine Interaktion herzustellen. Das BCI liest die Wellen, die vom Gehirn an den verschiedensten Stellen produziert werden ein und übersetzt diese Signale in Befehle, damit mit einem Computer interagiert werden kann [6].

Es gibt verschiedene Aufnahmetechniken von BCIs:

- Invasiv
- Teilweise invasiv
- Nicht invasiv

Invasive Gehirn-Computer-Schnittstellen werden direkt im menschlichen Gehirn durch eine Operation eingesetzt. Es können hier einzelne oder mehrere Einheiten in verschiedenen Bereichen des Gehirns platziert und anschließend angesprochen werden. Diese Systeme weisen die höchste Qualität auf, bringen aber durch die Operation gewisse Risiken mit sich. Bei teilweise invasiven BCI-Systemen werden die Einheiten nicht direkt in den Gehirnzellen sondern im Schädel eingesetzt, allerdings können daher die Signale etwas schlechter empfangen. Bei nicht invasiven Systemen handelt es sich um die sicherste und kostengünstigste Variante auf Kosten von schwächerem Gehirnsignalempfang. Die Signale werden über Elektroden empfangen, die am Kopf befestigt werden. Für diese Technik werden meist mit Hilfe eines Elektroenzephalografen (EEG) die neuronalen Wellen des Gehirns gemessen [6].

Die neuronale Aktivität des Gehirns ist ausschlaggebend für die Messung. Daher ist es wichtig, die verschiedenen Signale, die ein EEG misst, kurz darzustellen [6]:



**Abbildung 2.10:** BCI Signalverarbeitung [6].

- Delta Signal: Hierbei handelt es sich um die langsamsten Wellen mit einer Frequenz zwischen 0.5 und 3.5 Hz, die während des Schlafens oder im komaähnlichen Zustand auftreten.
- Theta Signal: Es handelt sich hierbei um Wellen, die beispielsweise während des Träumens vorkommen und liegen in einem Bereich zwischen 3.5 und 7.5 Hz gemessen.
- Alpha Signal: Dies sind langsamere Wellen, die beim Problemlösen oder Entspannen auftreten (Frequenzen zwischen 7.5 und 12 Hz).
- Beta Signal: Diese Wellen sind oft mit Anstrengung, wie beispielsweise des Lösen einer mathematischen Aufgabe, verbunden und liegen in einem Bereich zwischen 12 und 30 Hz.
- Gamma Signal: Je größer hier die Amplitude ist, desto mehr Stress, Angst oder Panik erlebt die Personen (alle Frequenzen ab 31 Hz).

Je nach Anwendungsbereich können verschiedene Signale verwendet werden. Für eine Interaktion mit einer Person, die sich nicht in einem komaähnlichen Zustand befindet, sind alle Signale ab einer Stärke von 3.5 Hz relevant.

Für den gesamte Ablauf, der nötig ist, um eine Gehirn-Computer-Schnittstelle zu erstellen sind, wie in Abb. 2.10 veranschaulicht, sind grob drei Schritte notwendig [6]:

- Im ersten Schritt werden die elektrischen Signale von der Kopfhaut oder der Oberfläche des Gehirns erfasst (Signalerfassung).
- Da diese Signale in der Regel sehr schwach sind, werden sie als nächstes verstärkt und von Störvariablen gereinigt. Danach durchlaufen die Signale einen Übersetzungsalgorithmus, damit die Absichten des Benutzers erkannt werden können (Signalverarbeitung).
- Im letzten Schritt müssen die Daten von einem Programm bzw. einer Software so interpretiert werden, damit diese für eine Computerinteraktion verwendet werden können.

Um als Benutzer einen Computer mit Hilfe von Gehirnaktivität steuern zu können,



muss das System vor der Interaktion individuell angepasst werden. Als mögliche Kalibrierungsmethode kann ein Mauszeiger gedanklich mitverfolgt werden, während dieser sich bewegt. Der Benutzer soll sich hierbei vorstellen, er würde die Bewegung mit der Hand oder dem Arm ausführen. Je nach Person zeigt sich eine Anstrengung in dem motorischen Teil des Gehirns unterschiedlich stark und das System wird dahingehend angepasst. Im nächsten Schritt wird erneut ein Mauskursor gedanklich mitverfolgt. Allerdings erscheint ein zweiter Cursor, der auf Grundlage der bereits erhobenen Daten den momentanen Stand aufzeigt. Diese zwei Schritte werden mehrmals durchgeführt, um ein optimales Ergebnis zu erzielen. Neben der Mausbewegung kann auch ein Mausklick simuliert werden. Die Auswahl eines Menüpunktes kann beispielsweise dann so funktionieren, dass ein Menüpunkt hervorgehoben wird. Hat der Anwender sich für einen bestimmten entschieden, muss er sich besonders darauf konzentrieren bzw. die Gehirnbereiche müssen eine erhöhte Aktivität vorweisen [8].

Die Interaktion mit Hilfe der Messung der Gehirnaktivität ist im Gegensatz zu den zuvor beschriebenen Methoden eine komplexere und aufwendigere Eingabemethode. Auch die Interaktion ist mit viel Aufwand hinsichtlich der Kalibrierung verbunden.

## Kapitel 3

# Ausgabemethoden

Im Kapitel 2 wurden alternative Eingabemethoden zur Interaktion Hands-Free näher erläutert. Dieses Kapitel befasst sich mit Ausgabemethoden, die aber nicht zwingend an bestimmte Eingabemethoden gebunden sind. So gibt es nicht nur bei Systemen, die beispielsweise über die Sprache gesteuert werden, ein akustisches Feedback, sondern auch bei Systemen, die beispielsweise über die Augen oder den Mund gesteuert werden.

### 3.1 Auditive Ausgabe

Eine mögliche Form, wie ein Computer ein Feedback geben kann, geschieht durch eine auditive Ausgabe. Dies kann entweder durch ein Geräusch, ein Signal oder durch Sprachausgabe funktionieren.

Eine Möglichkeit auditives Feedback zu erhalten sind vom Computer erzeugte Signale oder bestimmte Geräusche. Bei diesen Signalen bzw. Geräuschen handelt es sich beispielsweise um einen hohen angenehmen Ton bei korrektem Verhalten oder um einen dumpfen unangenehmen Ton bei einer Fehlereingabe. Darüber hinaus können auch Klick-Geräusche als Ausgabe fungieren. So gibt ein Lautsprecher bei einer Fußgängerampel beispielsweise ein kontinuierliches Klick-Geräusch von sich, wenn es rot ist. Schaltet die Ampel auf grün verdoppelt sich die Geschwindigkeit des Geräusches.

Weitaus komplexer funktioniert allerdings die Sprachausgabe. Während bei der Spracheingabe die gesprochene Sprache von einem System in einen Befehl umgewandelt wird, gibt bei der Sprachausgabe der Computer die Informationen in Form von Sprache wieder. Da die Sprachausgabe von Menschen gehört und automatisch bewertet wird, ist es wichtig, dass die Sprache eine hohe Qualität aufweist, also möglichst real erscheint.

Wie in Abb. 3.1 dargestellt gibt es zwei unterschiedliche Ansätze, wie der Computer den Text in Sprache umwandeln kann:

- Sprachwiedergabe
- Sprachsynthese

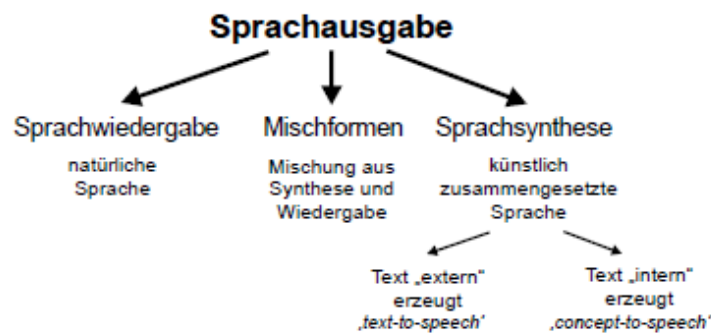


Abbildung 3.1: Unterteilung der Sprachausgabe [4].

Bei der Sprachwiedergabe werden Sprachsignale ausgegeben, die zuvor aufgenommen und abgespeichert worden sind. Der Umfang an möglichen akustischen Meldungen, also an Wörtern, ist aus diesem Grund an die Anzahl der Aufnahmen gebunden und daher begrenzt. Da die Signale bzw. Wörter von einer realen Person eingesprochen wurden, weist diese Methode eine sehr hohe Sprachqualität auf [7].

Es gibt unterschiedliche Methoden bzw. Verfahren, wie die Sprachwiedergabe im Detail funktioniert. Zusammenfassend werden die Wörter zu Beginn von einer Person eingesprochen, bearbeitet und abgespeichert. Die einzelnen Sprachelemente werden anschließend adressiert, sodass sie vom System später leicht gefunden werden können. Wenn der Computer nun einen Text vorlesen soll, werden im System anhand der gespeicherten Adressen die einzelnen Wörter bzw. Phrasen herausgesucht und verwendet [4].

Im Gegensatz zur Sprachwiedergabe wird bei der Sprachsynthese das Ziel verfolgt, alle möglichen Wortkombinationen in ein Sprachsignal umwandeln zu können. Da die Wörter, die als Grundlage für die Ausgabe dienen, nicht von einer Person eingesprochen werden, sondern die einzelnen Wörter mit ihrer Betonung künstlich erzeugt werden müssen, sinkt die Sprachqualität und das Endergebnis gleicht oft mehr einer Computerstimme als einer menschlichen.

Bei der Sprachsynthese werden Silben aus verschiedenen Wörtern einzeln abgespeichert und so alle beliebigen Kombinationen zusammengesetzt. Im Anschluss werden Betonungen, Lautstärke und Geschwindigkeit hinzugefügt bzw. angepasst. Bei der sogenannten textgesteuerte Sprachsynthese (text-to-speech) ist dem System der eigentliche Inhalt des Textes nicht bekannt. Nicht nur einzelne Sätze bzw. ganze Texte können verarbeitet, sondern auch Konzepte können vom System in Signale umgewandelt werden (concept-to-speech). Das System kreiert einen eigenständigen Text, wenn das Konzept bzw. der Inhalt bekannt ist [4].

Laut Fellbaum [4] besteht die größte Herausforderung der Sprachsynthese darin, die Stimme noch natürlicher und nicht roboterähnlich wirken zu lassen. In der Praxis sind daher Mischformen zwischen Sprachwiedergabe und Sprachsynthese üblich.

## 3.2 Haptische Ausgabe

Haptische Ausgabe beinhaltet jenes Feedback, bei dem etwas durch den menschlichen Tastsinn wahrgenommen werden kann. Hierbei gibt es verschiedene Arten:

- Thermales Feedback (Wärme)
- Vibration
- Feedback durch ein haptisches Steuerungselement

Bei einer thermalen Ausgabe handelt es sich um ein Feedback in Form von Wärme. Der Computer bzw. das Gerät, mit dem interagiert wird, erhitzt sich auf eine vordefinierte Temperatur und kühlt sich wieder ab. Lee & Lim [16] stellten fest, dass warme Temperaturen mit etwas Positivem und kalte Temperaturen mit etwas Negativem verbunden werden. Des Weiteren konnten die Teilnehmer Abstufungen der Wärmeunterschiede wahrnehmen, die eine strikte Trennung von warm oder kalt überschreiten. Die Veränderungen von warm auf kalt oder umgekehrt sollen nicht zu plötzlich geschehen, da dies ein Unwohlbefinden bei der Interaktion auslöst [16].

Laut Jones Berri [9] soll sich das Ausgabegeräte um höchstens 20°C pro Sekunde abkühlen und um max. 10°C pro Sekunde aufwärmen. Die Spanne der Temperatur sollte sich im Bereich zwischen 5°C und 45° befinden, da dieser Bereich am besten von den Rezeptoren in der Haut wahrgenommen werden kann.

Eine haptische Ausgabe ist auch in Form von Vibration möglich. Die Vibration unterscheidet sich in ihrer Intensität und Dauer. Diese Art von Feedback kann unter anderem zur Bestätigung einer Eingabe erfolgen [17]. Will ein Benutzer beispielsweise auf einer virtuellen Tastatur einen Buchstaben auswählen, wird bei Erfolg eine leichte Vibration durch ein Armband o.ä. erfolgen.

Neben der thermalen Ausgabe und der Ausgabe durch Vibration ist das Feedback auch über das Eingabeelement möglich. Handelt es sich bei dem Eingabegerät um einen Joystick, kann die Ausgabe durch verschiedene Vibrationsmuster, aber auch durch eine Gegenbewegung des Joysticks in X-, Y- oder Z-Richtung erfolgen, je nachdem wie viele Freiheitsgrade das Eingabegerät erlaubt (vgl. Abschnitt 2.3.1) [1].

Es gibt unterschiedliche Möglichkeiten, wie haptische Ausgabe geschehen kann. Soll eine Eingabe besätigt werden, ist ein Feedback zu Vibration meist üblich. Soll ein möglichst realistisches Feedback bzw. eine Interaktion mit dem System gewährleistet werden, ist darüber hinaus ein Feedback vom Eingabeelement durch Bewegungsrichtung empfehlenswert.

### 3.3 Visuelle Ausgabe

Um Informationen visuell darzustellen und diese als Ausgabe des Systems zu nutzen, sind ein oder mehrere Displays oder Projektoren notwendig. Die Ausgabe kann entweder auf dem selben Bildschirm wie die Eingabe zu sehen sein, oder auf einem separaten Bildschirm.

Es gibt sehr viele unterschiedliche Formen und Ausprägungen, wie eine visuelle Ausgabe erfolgen kann. Dies beginnt bei der simplen Ausgabe eines Textes. Wird eine bestimmte Interaktion richtig oder falsch ausgeführt, kann daraufhin ein Text bzw. ein Wort erscheinen, das den Erfolg bzw. Misserfolg mitteilt. Des Weiteren können auch Bilder oder andere Mediaelemente in einem ähnlichen Kontext gezeigt. Bei der bloßen Unterscheidung zwischen korrekt und inkorrekt erscheint beispielsweise ein grünes Häkchen bzw. rotes Kreuz. Visuelle Ausgabe erfolgt aber nicht nur über unterschiedliche Symbole, auch die bloße Änderung der Farbe eines Elementes kann ausschlaggebend für dessen Zustand sein.

Abb. 2.8(a) beschreibt einen Kalibrierungsschritt vor der Verwendung eines Systems, das durch die Muskelkontraktionen gesteuert werden kann. Hier gibt der Computer laufend ein visuelles Feedback über den aktuellen Zustand bzw. die Anspannung der Muskeln.

Bei dem in Abschnitt 4.1 erwähnten Augensteuerungssystem des Unternehmens<sup>1</sup> kann eine virtuelle Tastatur mit Hilfe der Augen bedient werden. Wird ein Buchstabe über einen längeren Zeitraum fokussiert, wird dieser auch ausgewählt. Während der Fixation erscheint eine Sanduhr, die dem Benutzer mitteilt, wie lange der gewünschte Buchstabe noch fixiert werden muss [24].

Zusammengefasst kann gesagt werden, dass es sehr viele verschiedene Möglichkeiten gibt, wie ein Computer bzw. das Interaktionsgerät selbst visuell Feedback geben kann. Je nach Anwendungsfall, ob es sich beispielsweise um eine richtig oder falsch Unterscheidung oder um eine komplexere Ausgabe handelt, können Texte, Symbole, Farben oder auch Kombinationen ausgewählt werden.

Alle beschriebenen Ausgabemethoden können einem menschlichen Sinn zugeordnet werden (auditive Ausgabe dem Hörsinn, haptische Ausgabe dem Tastsinn und visuelle Ausgabe dem Sehsinn). Erwähnenswert wäre auch, dass es eine Ausgaben gibt, die dem Geschmackssinn und dem Geruchssinn zugeordnet werden können. Allerdings kommen diese in der Praxis nur vereinzelt zum Einsatz und werden daher im Zuge dieser Arbeit nicht vertiefend erläutert.

---

<sup>1</sup><http://humanelektronik.de/>

## Kapitel 4

# Gegenüberstellung der Methoden

In Kapitel 3 wurden die verschiedenen Eingabemethoden näher erläutert. Dieses Kapitel beschäftigt sich damit, welche Vor- und Nachteile, gesonderte Rahmenbedingungen und welche Einsatzgebiete es zu den einzelnen Methoden gibt.

### 4.1 Fördernde und hemmende Faktoren

Jede einzelne Eingabemethode weist gewisse Pro- und Contraseiten auf. Daher wurde, wie in Tab. 4.1 dargesellt, eine Übersicht bzw. ein Vergleich erstellt.

Ein wichtiger Punkt, den es vorab zu klären gibt, ob ein System verwendet werden soll, bezieht sich auf die Kosten. Bei der Sprachsteuerung gibt es hierfür keine Pauschalaussage, sondern es ist je nach Anwendungsfall unterschiedlich. Es gibt einige Sprachassistenten wie z.B. Siri<sup>1</sup> oder Cortana<sup>2</sup>, mit denen ein Computer zum Teil mit Hilfe der Sprache gesteuert werden kann. Wie in Abschnitt erklärt, werden für die Augensteuerung sowohl ein Eye-Tracking-System, als auch eine Software benötigt. Die in Abb. 2.2(a) belaufen sich beispielsweise auf rund 13.500€, wobei hier noch die Kosten für die Software hinzukommen [35]. Kinn- und Mundsteuerungen sind um einiges billiger und kosten zwischen 400€ und 2000€ [26] [30]. Das Myo-Armband als Referenz für die Muskelsteuerung beläuft sich auf ca. 250€, wobei komplexere Systeme wesentlich teurer sind [34].

Auch die Interaktionsdauer mit dem System ist sehr unterschiedlich. Da es die Sprachsteuerung und die Gestensteuerung nur eine minimale bis geringe kognitive und körperliche Anstrengung erfordert, gibt es bei der Benutzungsdauer kaum Einschränkungen. Im Gegensatz findet man bei der Augensteuerung sowohl eine körperliche, als auch eine kognitiven Anstrengung und sollte daher nicht zu lange benutzt werden, damit die Augen langfristig nicht geschädigt werden. Ähnlich wie bei der gewöhnlichen Benutzung eines Computers sollte alle 20 Minuten eine Pause eingelegt werden [20]. Das Benutzen einer Steuerung, die durch die Muskeln oder durch die Gehirnaktivität funktioniert, ist mit hoher kognitiven Anstrengungen verbunden. Zusätzlich erfordert das Anspannen

---

<sup>1</sup><https://www.apple.com/ios/siri/>

<sup>2</sup><https://support.microsoft.com/en-us/help/17214>

und Entspannen der Muskelpartien eine sehr hohe körperliche Anstrengung. Aus diesen Gründen ist die Interaktionsdauer der beiden Methoden sehr beschränkt. Daher muss während der Interaktion überprüft werden, ab welchem Zeitpunkt eine Überanstrengung erfolgt und danach Pausen einteilen.

Ausschließlich der Gestensteuerung müssen alle Eingabegeräte vor der ersten Benutzung individuell angepasst (kalibriert) werden. Bei der Sprachsteuerung ist dies von System zu System verschieden, allerdings ist der Zeitaufwand hier sehr gering. Da das Verwenden einer Muskelsteuerung oder einer Steuerung durch die Gehirnaktivität durch Trial und Error geprägt ist, ist auch die Kalibrierung zu Beginn und vor jeder Benutzung sehr aufwendig.

Bei den meisten Systemen gibt es keine räumliche Einschränkung und sind daher nicht an ein gewisses Setting gebunden. Bei der Benutzung einer Augensteuerung gibt es allerdings zu beachten, dass keine störenden Lichtquellen die Kamera bzw. die Interaktion behindern. Des Weiteren ist man bei der Verwendung von Systemen, die durch die Gehirn- oder Muskelaktivität gesteuert werden, oft an notwendige Geräte und somit an gewisse Räumlichkeiten gebunden.

Für die Benutzung jeder der verschiedenen Eingabemethoden gelten die selben Grundvoraussetzungen. Der Benutzer muss eine eigenständige Kontrolle über das jeweilige Eingabemittel (Sprache, Muskelanspannung, Augenbewegungen etc.) besitzen, damit eine Interaktion möglich ist. Zusätzlich gibt es spezifische Faktoren wie z.B. ist bei der Sprachsteuerung eine laute und deutliche Sprache ein fördernder Faktor, hingegen erweisen sich verstaubte spiegelnde Brillengläser bei der Augensteuerung als ein hemmender Faktor.

Tabelle 4.1: Verschiedene Faktoren der Eingabemethoden

	<b>Sprachsteuerung</b>	<b>Augensteuerung</b>	<b>Gestensteuerung</b>	<b>Muskelsteuerung</b>	<b>Steuerung durch Gehirnaktivität</b>
Kosten	je nach Anwendungsfall, für Desktopsteuerung einige kostenlose Anwendungen	Kosten für Eye-Tracking-System und Software, Tobii Pro Glasses 2 ca. 13.500€	Kosten für gewöhnlichen Joystick gering, Sensory Kinnsteuerung: ca. 400€, QuadStick Gamecontroller: ca. 500€, IntegraMouse Plus: ca. 2000€	Myo-Armband: ca. 250€	keine Produkte am Markt
Dauer	keine Einschränkung	beschränkt, 20-20-20-Methode	keine Einschränkung, bei optischer Methode der Kopfsteuerung ähnl. Augensteuerung	beschränkt	beschränkt
Kognitive Anstrengung	gering	mittel	gering	hoch	sehr hoch
Körperliche Anstrengung	keine	hoch	gering	sehr hoch	gering

*Wird auf der nächsten Seite weitergeführt*



Tabelle 4.1 – *Verschiedene Faktoren der Eingabemethoden*

	<b>Sprachsteuerung</b>	<b>Augensteuerung</b>	<b>Gestensteuerung</b>	<b>Muskelsteuerung</b>	<b>Steuerung durch Gehirnaktivität</b>
UX	intuitiv, Sprache ist häufigstes mittel zur Kommunikation/Interaktion	gewöhnungsbedürftig, aber dennoch intuitiv (Auswahl funktioniert über Fixation)	intuitive bewegung mit dem Kinn, Füßen und Kopf, kurze Eingewöhnungsphase bei Mundsteuerung (nippen, pusten)	gewöhnungsbedürftig (Trial and Error)	gewöhnungsbedürftig (Trial and Error)
Kalibrierung	teils	ja	nein, ja bei optischer Variante der Kopfsteuerung	ja, aufwendig	ja, sehr aufwendig
Wartung/Reinigung	Wartung, wenn sich die Stimme verändert (z.B. Stimmbruch)	nein	Wartung nein, Reinigung ja z.B. bei Mundsteuerung	nein	nein
Räumlichkeiten	keine Einschränkung	drinnen (Lichteinflüsse)	keine Einschränkung, ja bei optischer Kopfsteuerung	Laborsetting, keine Einschränkung bei Benutzung des Myo Armbandes	Laborsetting

*Wird auf der nächsten Seite weitergeführt*

Tabelle 4.1 – *Verschiedene Faktoren der Eingabemethoden*

	<b>Sprachsteuerung</b>	<b>Augensteuerung</b>	<b>Gestensteuerung</b>	<b>Muskelsteuerung</b>	<b>Steuerung durch Gehirnaktivität</b>
Größe	meist kein zusätzliches Gerät, ansonsten richtet sich die nach Mikrophon und Interaktionssystem	SMI Red250mobil: 24 x 2.7 x 3 cm; Tobii Pro Spectrum 55 cm x 18 cm x 6 cm; Tobii Pro Glasses 2: 130 x 85 x 27 mm;	Mundsteuerungen bzw. gewöhnliche Joysticks ungefähr Faustgroß, andere Systeme etwas größer	Myo-Armband: 11,9 x 7,4 x 10,4 cm	sehr klein (Größer der Elektroden)
Gewicht	kein zusätzliches Gewicht, Mikrophon meist im System integriert	Tobii Pro Glasses 2: 312 g, Tobii Pro Spectrum: 3.8 kg	gering (Gewicht des Joysticks)	Myo-Armband: 254g	sehr gering (Gewicht der Elektroden)
Grundvoraussetzung	klare Sprache	kontrollierte Augenbewegungen	kontrollierte Bewegungen der ausgewählten Körperteile	Kontrolle über Muskelan- und entspannungen	Selbstständige Aktivierung der Gehirnsignale in den zu messenden Arealen

*Wird auf der nächsten Seite weitergeführt*

Tabelle 4.1 – *Verschiedene Faktoren der Eingabemethoden*

	<b>Sprachsteuerung</b>	<b>Augensteuerung</b>	<b>Gestensteuerung</b>	<b>Muskelsteuerung</b>	<b>Steuerung durch Gehirnaktivität</b>
Spezifische Faktoren	laut und deutlich sprechen, normale Sprechgeschwindigkeit, keine zu große Distanz zum Mikrophon	verstaubte spiegelnde Brillengläser sind ein hemmender Faktor, zwei unterschiedliche Profile wenn ein Benutzer zwischen Kontaktlinsen und Brille wechselt	System muss auf Benutzer genau angepasst werden, wenn die Bewegungen sehr rucksartig sind	Richtige Positionierung der Elektroden, damit genug Muskelaktivität gemessen werden kann	messbare Gehirnsignale über 3.5 Hz

## 4.2 Einsatzgebiete

Es gibt viele verschiedene Bereiche in denen die in Kapitel 3 beschriebene alternativen Eingabemethoden eingesetzt werden.

Zum einen werden viele Steuerungen als assistive Technologie für Menschen mit Beeinträchtigung eingesetzt. Besonders für Menschen mit Tetraplegie bzw. Tetraparese, die ihre Hände und Arme nicht oder oft nur sehr eingeschränkt benutzen können, sind die Einsatzgebiete relevant. Abb. 2.5(a) zeigt die IntegraMouse Plus, die die Produktion von Musik, das Computerspielen, die alltägliche Arbeit, aber vor allem eine selbständige und unabhängige Benutzung des Computers ermöglicht [36]. Die Augensteuerung der Firma Humanelektronik verschafft Menschen, die Probleme haben sich sprachlich auszudrücken, ein leistungsfähiges Kommunikationswerkzeug [24]. Die in Abb. 2.3(a)(b) dargestellten Kinnsteuerungen werden in der Praxis dazu verwendet, einen Rollstuhl selbstständig zu bedienen. Es können so die Fahrtrichtung gesteuert und durch das Menü navigiert werden.

Es gibt aber auch viele alltägliche Situationen, in denen Systeme eingesetzt werden, die ohne die Hände zu bedienen sind. Beispielsweise funktionieren Freisprechanlage oder Navigationssysteme im Auto bereits über Spracheingabe. Momentaner Trend ist allerdings Alexa bzw. der Amazon Echo Dot. Dabei handelt es sich um einen Lautsprecher inklusive Software, mit dem über Sprache kommuniziert wird. Das System kann Musik von verschiedenen Plattformen abspielen, Telefonate starten, Nachrichten schicken, Nachrichten beantworten oder vorlesen und kann Lampen bzw. Lichter im Haus steuern [37].

Aber nicht nur im privaten, sondern auch im beruflichen Bereich werden oft alternative Eingabemethoden eingesetzt. Das in Abb. 2.9 dargestellte Myo-Armband wird durch den Einsatz der verschiedenen Gesten als Präsentationswerkzeug verwendet, um beispielsweise Folien vor oder zurückzuschalten [34].

Zusammenfassend lässt sich sagen, dass es viele Faktoren bei der Verwendung der verschiedenen Eingabemethoden zu beachten gilt. Je nach Anwendungsbereich, Räumlichkeiten und vor allem Benutzer sollten daher die alternativen Systeme vorher gründlich überprüft werden, bevor sie ausgewählt werden. Darüber hinaus gibt es ausschließlich der Steuerung durch Gehirnaktivität bereits viele verschiedene Produkte bzw. Einsatzgebiete.

## Kapitel 5

# Fazit und Ausblick

Alternative Eingabemethoden gewinnen immer mehr an Bedeutung. Weil die verschiedenen Systeme sowohl im Alltag, als assistive Technologie oder im Arbeitsbereich eingesetzt werden, wird der Bedarf erkannt und dahingehend entwickelt.

In den letzten Monaten gewannen Spracheingabesysteme immer mehr an Beliebtheit, wie an dem Beispiel von Amazon Echo zu sehen ist. Weiters gibt es bereits auch einen Sprachassistenten, der von Google entwickelt wurde. Der Google Assistent<sup>1</sup> hat die selben Ziele, wie das System von Amazon - Fragen zu beantworten, Aufgaben zu übernehmen und technische Geräte im Haus zu steuern. Laut Futurezone [38] soll auch das Unternehmen Apple<sup>2</sup> im nächsten Jahr einen Lautsprecher mit digitaler Assistenz herausbringen. Das Grundkonzept baut auf dem bereits existierenden Sprachassistenten Siri auf.

Absehen von der Möglichkeit einen Computer mit Hilfe der Stimme zu steuern, gibt es in den Bereichen der Steuerung durch Muskel- und Gehirnaktivität die meisten Potentiale. Im Vergleich sind bei den anderen beschriebenen Eingabesystemen die Fragen von zentraler Bedeutung, wie gewisse Fehler ausgebessert (z.B. Lichtprobleme bei der Augensteuerung) bzw. die Leistungsfähigkeit der Systeme gesteigert werden können. Hingegen bei der Steuerung durch Muskel- und Gehirnaktivität liegt der Fokus auf der Herstellung von marktreifen Geräten, die im Alltag eingesetzt werden können. Das Unternehmen Facebook<sup>3</sup> arbeitet beispielsweise an einem System, dass Benutzern ermöglichen soll 100 Wörter pro Minuten durch reine Gedankensteuerung tippen zu können [39].

Die Frage, die man sich in Zukunft, vor allem bei der Steuerung durch Gehirnaktivität, stellen muss, ist, wie die rechtlichen Rahmenbedingungen ausschauen. Gibt es ausreichend großes Zielpublikum, dass sich von einer Software die Gedanken lesen lässt. Wer entscheidet welche Informationen für das Generieren eines Textes relevant sind und wer besitzt aller Zugang zu diesen Daten? Sollten diese Fragen ausreichend und transparent geklärt werden, gibt es in diesem Bereich das meiste Zukunftspotenzial.

---

<sup>1</sup><https://assistant.google.com/>

<sup>2</sup><https://www.apple.com/>

<sup>3</sup>[www.facebook.com](http://www.facebook.com)

# Quellenverzeichnis

## Literatur

- [1] B. An und W.P. Stiles. *Haptic feedback joystick*. US Patent 6,429,849. Aug. 2002. URL: <http://www.google.com/patents/US6429849> (siehe S. 20).
- [2] Alexander G. Barszap u.a. „Effects of muscle fatigue on the usability of a myoelectric human-computer interface“. *Human Movement Science* 49.5 (Okt. 2016), S. 225–238 (siehe S. 13).
- [3] Andrew Duchowski. *Eye Tracking Methodology. Theory and Practice*. 2. Aufl. London: Springer, 2010 (siehe S. 5).
- [4] Klaus Fellbaum. *Sprachverarbeitung und Sprachübertragung*. 2. Aufl. Heidelberg: Springer, 2012 (siehe S. 19).
- [5] Maria Hakonen u.a. „Current state of digital signal processing in myoelectric interfaces and related applications“. *Biomedical Signal Processing and Control* 18.4 (Apr. 2015), S. 334–259 (siehe S. 13).
- [6] Ahmad Taher Hassanien Aboul Ella und Azar. *Brain-Computer Interfaces. Current Trends and Applications*. Schweiz: Springer, 2015 (siehe S. 15, 16).
- [7] Tobias Kaufmann und Beat Pfister. *Sprachverarbeitung. Grundlagen und Methoden der Sprachsynthese und Spracherkennung*. Heidelberg: Springer, 2008 (siehe S. 3, 19).
- [8] Sung-Phil Kim u.a. „Point-and-Click Cursor Control With an Intracortical Neural Interface System by Humans With Tetraplegia“. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 19.2 (Apr. 2011), S. 193–203 (siehe S. 17).
- [9] Berris L.A. Jones und M. „The psychophysics of temperature perception and thermal-interface design“. In: IEEE, 2002 (siehe S. 20).
- [10] Zhang Naizhong und Wen Jing. „Hand-Free Head Mouse Control Based on Mouth Tracking“. In: IEEE, 2015 (siehe S. 11).
- [11] Kara Nielsen Jakob und Pernice. *Eyetracking Web Usability*. Berkeley, CA: New Riders, 2010 (siehe S. 4, 5).
- [12] Christina Pahl, Tan Yong Oon und Eko Supriyanto. „Fuzzy System Based Face Tracking for Head Movement Control in Progressive Health Care“. In: IEEE, 2015 (siehe S. 11).

- [13] Raimund Preim Bernhard und Dachzelt. *Interaktive Systeme: Band 2. User Interface Engineering, 3D-Interaktion, Natural User Interfaces*. 2. Aufl. Heidelberg: Springer, 2015 (siehe S. 7).
- [14] Ryszard Tadeusiewicz. „Speech in Human System Interaction“. In: IEEE, 2010 (siehe S. 3).
- [15] Sanjay S. Vernon Scott und Joshi. „Brain–Muscle–Computer Interface: Mobile-Phone Prototype Development and Testing“. *IEEE Transactions on Information Technology in Biomedicine* 15.4 (Juni 2011), S. 531–538 (siehe S. 13, 14).
- [16] Lim Wonjun Lee und Youn-kyung. „Thermo-message: exploring the potential of heat as a modality of peripheral expression“. In: ACM New York, 2010, S. 4231–4236 (siehe S. 20).

## Online-Quellen

- [17] 2011. URL: <http://th-wildau.de/sbruntha/Material/VR/Websites-T09/Feedback/> (besucht am 14.05.2017) (siehe S. 20).
- [18] Sep. 2012. URL: <http://www.moso-gmbh.de/de/motion-concepts/sondersteuerungen/kinnsteuerung-moso.html> (besucht am 15.04.2017) (siehe S. 7, 8).
- [19] Sep. 2013. URL: <https://www.youtube.com/watch?v=rpA2y5IPY0A> (besucht am 15.04.2017) (siehe S. 9).
- [20] 2016. URL: <https://www.aao.org/eye-health/tips-prevention/computer-usage> (besucht am 27.05.2017) (siehe S. 22).
- [21] Apr. 2017. URL: <http://www.tobiipro.com/product-listing/tobii-pro-glasses-2/> (besucht am 03.04.2017) (siehe S. 6).
- [22] Apr. 2017. URL: <http://www.tobiipro.com/product-listing/tobii-pro-spectrum/> (besucht am 03.04.2017) (siehe S. 6).
- [23] Apr. 2017. URL: <https://www.smivision.com/eye-tracking/product/eye-trackingproducteye-trackingproductred250mobile-eye-tracker/> (besucht am 05.04.2017) (siehe S. 6).
- [24] Apr. 2017. URL: <http://humanelektronik.de/produkte/seetech/> (besucht am 29.04.2017) (siehe S. 6, 21, 28).
- [25] Apr. 2017. URL: <https://smilemart-tech.com/assistive-technology-products/wheelchair-controls/wheelchair-chin-control/> (besucht am 15.04.2017) (siehe S. 7, 8).
- [26] Apr. 2017. URL: <http://www.sensoryguru.com/product/chin-mouse-controller/> (besucht am 15.04.2017) (siehe S. 7, 8, 22).
- [27] Mai 2017. URL: <https://www.techopedia.com/definition/12702/six-degrees-of-freedom-6dof> (besucht am 09.05.2017) (siehe S. 8).
- [28] Mai 2017. URL: [https://en.wikipedia.org/wiki/Six\\_degrees\\_of\\_freedom](https://en.wikipedia.org/wiki/Six_degrees_of_freedom) (besucht am 09.05.2017) (siehe S. 9).
- [29] Apr. 2017. URL: <http://www.quadstick.com/> (besucht am 16.04.2017) (siehe S. 9, 10).

- [30] Apr. 2017. URL: [http://www.integramouse.com/en/home/highlights-details/?tx\\_ttnews%5Btt\\_news%5D=470&cHash=6457ff0d57ee2de8b5c41ac23b53d9b1](http://www.integramouse.com/en/home/highlights-details/?tx_ttnews%5Btt_news%5D=470&cHash=6457ff0d57ee2de8b5c41ac23b53d9b1) (besucht am 15.04.2017) (siehe S. 10, 22).
- [31] Apr. 2017. URL: <https://www.fruehwald.net/mobilitaet/rollstuehle-u-elektrollstuehle/sondersteuerungen> (besucht am 29.04.2017) (siehe S. 11, 12).
- [32] Apr. 2017. URL: <http://www.hidrex.de/reha/rollstuhlsondersteuerung/r-net-system-penny-u-gilles/1-fuss-steuerung-mit-schuhhalter-inkl-joystick> (besucht am 29.04.2017) (siehe S. 12).
- [33] Mai 2017. URL: <https://support.getmyo.com/hc/en-us/categories/200176946-Frequently-Asked-Questions> (besucht am 13.05.2017) (siehe S. 14).
- [34] Mai 2017. URL: <https://www.myo.com/techspecs> (besucht am 13.05.2017) (siehe S. 15, 22, 28).
- [35] 2017. URL: <http://vandrico.com/wearables/device/tobii-pro-glasses-2> (besucht am 27.05.2017) (siehe S. 22).
- [36] Mai 2017. URL: <http://www.integramouse.com/en/user-stories> (besucht am 27.05.2017) (siehe S. 28).
- [37] 2017. URL: <https://www.amazon.com/Amazon-Echo-Bluetooth-Speaker-with-WiFi-Alexa/dp/B00X4WHP5E> (besucht am 27.05.2017) (siehe S. 28).
- [38] Juni 2017. URL: <https://futurezone.at/produkte/apples-smarte-siri-lautsprecher-werden-bereits-produziert/267.277.617> (besucht am 01.06.2017) (siehe S. 29).
- [39] Apr. 2017. URL: <https://www.researchsnipers.com/facebook-mind-reading-technology-unveiled-event/> (besucht am 01.06.2017) (siehe S. 29).