

Menschenzentrierte Robotik

Assistenzsystem für Verpressung

Banujan Baskaran Esat Sefa Bayhan Erim Erel Caliskan
Christopher Kamphausen Abinash Selvarajah
Simon Spettmann Mara Fiona Stange

18. Mai 2023

Inhaltsverzeichnis

1	Sicherheitskonzept	2
2	Software Architektur	3
3	Sprachsteuerung	4
3.1	Anforderungen	4
3.2	Recherche: Mögliche Technologien	5
3.2.1	Rhasspy Raven	5
3.3	Probleme von Online-Lösungen	6
3.4	Ansatz: Speech to Text	6
3.5	Ansatz: Hot word recognition	6
3.6	Gewählte Lösung: XYZ	6
4	Greifen	7
5	Bewegung	8
6	3D Druck	9

Kapitel 1

Sicherheitskonzept

Kapitel 2

Software Architektur

Kapitel 3

Sprachsteuerung

Einführung Sprachsteuerung Wofür Weshalb Warum Sinn und Zweck für Kathrin Psychologie

3.1 Anforderungen

Bei der Entwicklung einer Sprachsteuerung gibt es eine Vielzahl von Anforderungen, die berücksichtigt werden müssen. Eine dieser Anforderungen ist die Offline-Fähigkeit, die besonders wichtig ist, wenn keine zuverlässige Internetverbindung vorhanden ist. Dies ist bei uns der Fall, denn die Sprachsteuerung wird in einer Werkstatt mit potentiellen vielen Störfaktoren verwendet. In diesem Fall muss die Sprachsteuerung in der Lage sein, ohne Verbindung zum Internet zu funktionieren.

Eine weitere wichtige Anforderung an eine Sprachsteuerung ist eine hohe Genauigkeit. Dies bedeutet, dass die Sprachsteuerung in der Lage sein muss, Befehle von Kathrin mit hoher Präzision zu erkennen und umzusetzen. Ein weiterer wichtiger Faktor ist die Generalisierung, da die Sprachsteuerung in verschiedenen Umgebungen eingesetzt wird, wie z.B. in einer Werkstatt mit lauten Werkzeugen und vielen Personen. Es ist daher wichtig, dass die Sprachsteuerung in der Lage ist, die Stimme von Kathrin in verschiedenen Umgebungen zu erkennen.

Ein weiterer wichtiger Faktor ist, dass die Sprachsteuerung kostenlos sein sollte, da die Verfügbarkeit von Mitteln zur Finanzierung des Projekts eventuell begrenzt ist. Darüber hinaus muss die Sprachsteuerung effizient sein, da die eingesetzte Hardware-Plattform unbekannt ist, deren Leistung unterschiedlich sein kann. Die Befehle, die von Kathrin gegeben werden, sind kritisch und müssen mit niedriger Latenz erkannt werden.

Schließlich ist es auch wichtig, dass die Sprachsteuerung in Python entwickelt wird, da dies die bevorzugte Sprache im Team ist. Die Sprachsteuerung wird in einem großen Software-Projekt eingesetzt und von verschiedenen Komponenten verwendet. Entsprechend ist die Integration und Interaktion wichtige Aspekte

die berücksichtigt werden müssen.

Insgesamt müssen die Anforderungen der Sprachsteuerung sorgfältig abgewogen und umgesetzt werden, um sicherzustellen, dass die Sprachsteuerung den Bedürfnissen von Kathrin und den Anforderungen der Umgebung entspricht.

3.2 Recherche: Mögliche Technologien

- Künstliche Neuronale Netze (Wieso, Weshalb, Warum)
 - Speech To Text
 - Hot Word recognition
- SpeechRecognition
- OpenWhisper
- Custom Model
- Mycroft Precise (Aktuell)
- Snowboy
- Picovoice Porcupine

3.2.1 Rhasspy Raven

- Keine Nutzung von neuronalen Netzen
- Sehr effizient
- Nutzung von Audio Algorithmen
- Sehr schnell
- Umfangreiche leicht zu bedienende API
- Schlechte Generalisierung
- Aufnahme von Person die mehrmals die Befehle sagt
- Es wird nach ähnlichkeiten gesucht
- Da die Stimme von Kathrin unbekannt können wir entsprechend keine Ähnlichkeiten suchen

3.3 Probleme von Online-Lösungen

Eine der potentiellen Fehlerquellen bei einer sprachgesteuerten Anwendung ist die Verwendung einer Online-Lösung. Eine Online-Lösung erfordert eine stabile Internetverbindung, was jedoch nicht immer gewährleistet werden kann. Schwache Internetverbindungen oder Internetabbrüche können dazu führen, dass die Anwendung nicht mehr ordnungsgemäß funktioniert oder sogar vollständig ausfällt.

Dartüber hinaus ist die Anwendung auch von der Erreichbarkeit des Servers oder der API abhängig. Falls der Server oder die API nicht erreichbar sind, kann die Anwendung nicht mehr auf die benötigten Ressourcen zugreifen, was zu Fehlfunktionen oder einem vollständigen Ausfall der Anwendung führen kann. Diese Abhängigkeit kann zu einem kritischen Problem werden und die Arbeit von Kathrin massiv behindern.

Ein weiterer wichtiger Aspekt ist die des Datenschutzes. Bei der Nutzung von Drittanbieter-Online-Diensten kann die Sicherheit der Aufnahmen der Spracheingabe nicht mehr gewährleistet werden. Schließlich ist es wichtig zu beachten, dass Kathrin möglicherweise nicht in einem isolierten Arbeitsbereich arbeitet. Es besteht die Möglichkeit, dass unfreiwillige Aufnahmen von Personen in der Nähe aufgezeichnet und an den Online-Dienst gesendet werden. Dies kann ein schwerwiegendes Problem sein und die Vertraulichkeit von Gesprächen und möglicherweise auch von personenbezogenen Daten beeinträchtigen.

3.4 Ansatz: Speech to Text

- Leistungshungrig
- Hohe false negative rate (zumindest offline Modelle)

3.5 Ansatz: Hot word recognition

- Passende Bibliothek muss noch gefunden werden
- potentiell Effizienter
- Läuft eher offline

3.6 Gewählte Lösung: XYZ

Kapitel 4

Greifen

Kapitel 5

Bewegung

Kapitel 6

3D Druck