



Gesten- und Sprachsteuerung für einen mobilen Roboter mittels Kinect for Windows unter Java

Studienarbeit

für die Prüfung zum

Bachelor of Science

der Angewandte Informatik

an der Dualen Hochschule Baden-Württemberg Karlsruhe

von

Simon Ebner, Volker Werling

Januar 2013

Bearbeitungszeitraum
Matrikelnummer, Kurs
Gutachter

12 Wochen
5837963 7012192, TAI10B2
Prof. Hans-Jörg Haubner

Erklärung

Wir erklären hiermit ehrenwörtlich:

1. dass wir unsere Studienarbeit mit dem Thema *Gesten- und Sprachsteuerung für einen mobilen Roboter mittels Kinect for Windows unter Java* ohne fremde Hilfe angefertigt haben;
2. dass wir die Übernahme wörtlicher Zitate aus der Literatur sowie die Verwendung der Gedanken anderer Autoren an den entsprechenden Stellen innerhalb der Arbeit gekennzeichnet haben;
3. dass wir unsere Studienarbeit bei keiner anderen Prüfung vorgelegt haben;
4. dass die eingereichte elektronische Fassung exakt mit der eingereichten schriftlichen Fassung übereinstimmt.

Wir sind uns bewusst, dass eine falsche Erklärung rechtliche Folgen haben wird.

Karlsruhe, Januar 2013

Simon Ebner, Volker Werling

Zusammenfassung

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Einleitung | 6 |
| 2 | Aufgabenstellung | 7 |
| 3 | Stand der Technik | 8 |
| 3.1 | Hardware | 8 |
| 3.1.1 | Kinect | 8 |
| 3.1.2 | Roboter - Ausblick | 8 |
| 3.2 | Software | 8 |
| 3.2.1 | Microsoft Kinect SDK | 8 |
| 3.2.2 | Java | 8 |
| 3.2.3 | GUI | 8 |
| 3.3 | Mustererkennung | 8 |
| 4 | Kinect-Framework | 9 |
| 4.1 | Microsoft Kinect SDK | 9 |
| 4.2 | Java Frameworks | 9 |
| 4.2.1 | OpenNI | 9 |
| 4.2.2 | OpenKinect | 10 |
| 4.2.3 | jnect | 10 |
| 4.2.4 | Scoring-Modell | 10 |
| 4.2.5 | Analyseergebnis | 10 |
| 5 | Konzeption | 11 |
| 5.1 | Technische Vorgaben | 11 |
| 5.1.1 | Vorgaben durch Kinect | 11 |
| 5.1.2 | Vorgaben durch jnect-Framework | 11 |

| | | |
|----------|--|-----------|
| 5.1.3 | Service-Architektur | 11 |
| 5.1.4 | Abhängigkeit zu Robotertechnik | 11 |
| 5.2 | Fachliche Vorgaben | 11 |
| 5.2.1 | Vorgaben durch Mensch-Computer-Interaktion | 11 |
| 5.2.2 | Klassifikation der Anwendung | 11 |
| 5.2.3 | Analyse der Gesten- und Sprachsteuerung | 11 |
| 6 | Modelle zur Gesten- und Spracherkennung | 12 |
| 6.1 | Hidden-Markov-Modell | 12 |
| 6.1.1 | ToDo | 12 |
| 6.2 | Baum-Weich-Algorithmus | 12 |
| 6.2.1 | ToDo | 12 |
| 6.3 | ToDo | 12 |
| 7 | Gesten | 13 |
| 7.1 | Kreisbewegung | 13 |
| 7.1.1 | ToDo | 13 |
| 7.2 | Vorwärtsbewegung | 13 |
| 7.2.1 | ToDo | 13 |
| 7.3 | Haltesignal | 13 |
| 7.3.1 | ToDo | 13 |
| 8 | Sprachbefehle | 14 |
| 8.1 | ToDo | 14 |
| 9 | Implementierung | 15 |
| 9.1 | OSGI-Bundles | 15 |
| 9.1.1 | Gesten | 15 |
| 9.1.2 | Sprachbefehle | 15 |
| 9.1.3 | Roboterschnittstelle | 15 |
| 9.1.4 | Ausführbare Aktionen | 15 |
| 9.2 | Oberfläche | 15 |
| 9.2.1 | RCP-Anwendung | 15 |
| 9.2.2 | Grafische Darstellung | 15 |
| 9.3 | Ablaufbeschreibung | 15 |

| | |
|--------------------------------|----|
| 10 Ausblick - weitere Arbeiten | 16 |
| Abbildungsverzeichnis | 17 |
| Tabellenverzeichnis | 18 |
| Quellcodeverzeichnis | 19 |
| Literaturverzeichnis | 20 |
| Glossar | 21 |

1 Einleitung

Gestik im Sinne von kommunikativen Bewegungen stellt einen wesentlichen Teil der nonverbalen Kommunikation dar. Die Definition nach Kurtenbach und Hulteen (1990) besagt:

A gesture is a motion of the body that contains information. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key on a keyboard is neither observed nor significant. All that matters is which key was pressed.

Als solches bildet eine Geste abstrakte Strukturen und bildhafte Vorstellungen unmittelbar ab. Sie ist somit eine Körperbewegung, die Informationen enthält. Informationen, die herkömmliche Eingabegeräte, darunter Tastatur und Maus, nicht in dieser Form wiedergeben können. Eine Problematik, die sich besonders im Bereich der Mensch-Computer-Interaktion ergibt. Das Ziel der Mensch-Computer-Interaktion ist es nämlich die Kommunikation so intuitiv und unmittelbar wie möglich zu gestalten. Genau an dieser Stelle setzen Gesten an, in dem sie eine bessere Schnittstelle zwischen Mensch und Maschine darstellen.

Frühe Ansätze, wie das **Theremin** - ein 1919 erfundenes elektornisches Musikinstrument, das mit Händen berührungslos durch Beeinflussung eines elektromagnetischen Feldes gesteuert wird und dadurch Töne erzeugt - oder Datenhandschuhe, die erstmals 1983 von AT& T Bell Labs entwickelt wurden zeigen, die große Spannweite solcher Mensch-Maschine-Schnittstelle

2 Aufgabenstellung

3 Stand der Technik

3.1 Hardware

3.1.1 Kinect

3.1.2 Roboter - Ausblick

3.2 Software

3.2.1 Microsoft Kinect SDK

3.2.2 Java

3.2.3 GUI

3.3 Mustererkennung

4 Kinect–Framework

4.1 Microsoft Kinect SDK

asdf

4.2 Java Frameworks

adsf

4.2.1 OpenNI

Beschreibung

adsf

Technische Daten

SWOT–Analyse

4.2.2 OpenKinect

Beschreibung

Technische Daten

SWOT–Analyse

4.2.3 jnect

Beschreibung

Technische Daten

SWOT–Analyse

4.2.4 Scoring–Modell

4.2.5 Analyseergebnis

adsasfd

5 Konzeption

5.1 Technische Vorgaben

5.1.1 Vorgaben durch Kinect

5.1.2 Vorgaben durch jnect-Framework

5.1.3 Service-Architektur

5.1.4 Abhängigkeit zu Robotertechnik

5.2 Fachliche Vorgaben

5.2.1 Vorgaben durch Mensch-Computer-Interaktion

5.2.2 Klassifikation der Anwendung

5.2.3 Analyse der Gesten- und Sprachsteuerung

6 Modelle zur Gesten- und Spracherkennung

6.1 Hidden-Markov-Modell

6.1.1 ToDo

6.2 Baum-Weich-Algorithmus

6.2.1 ToDo

6.3 ToDo

7 Gesten

7.1 Kreisbewegung

7.1.1 ToDo

7.2 Vorwärtsbewegung

7.2.1 ToDo

7.3 Haltesignal

7.3.1 ToDo

8 Sprachbefehle

8.1 ToDo

9 Implementierung

9.1 OSGI–Bundles

9.1.1 Gesten

9.1.2 Sprachbefehle

9.1.3 Roboterschnittstelle

9.1.4 Ausführbare Aktionen

9.2 Oberfläche

9.2.1 RCP–Anwendung

9.2.2 Grafische Darstellung

9.3 Ablaufbeschreibung

10 Ausblick - weitere Arbeiten

Abbildungsverzeichnis

Tabellenverzeichnis

Quellcodeverzeichnis

Literaturverzeichnis

Glossar

Datenhandschuh

Der Datenhandschuh ist ein Eingabegerät in Form eines Handschuhs. Durch Bewegungen der Hand und Finger erfolgt eine Orientierung im virtuellen Raum. 6

Mensch-Computer-Interaktion

Die Mensch-Computer-Interaktion (englisch *Human-Computer Interaction, HCI*) als Teilgebiet der Informatik beschäftigt sich mit der benutzergerechten Gestaltung von interaktiven Systemen und ihren Mensch-Maschine-Schnittstellen. 5, 6, 11

Mensch-Maschine-Schnittstelle

Die Benutzerschnittstelle (nach Gesellschaft für Informatik, Fachbereich Mensch-Computer-Interaktion auch Benutzungsschnittstelle) ist die Stelle oder Handlung, mit der ein Mensch mit einer Maschine in Kontakt tritt. 6