

Orientierungstraining durch bild- und sprachbasierte Szenenanalyse mit Interaktion

Beschreibung des Themas:

Am Fachgebiet Neuroinformatik und Kognitive Robotik wird in einem Forschungsschwerpunkt roboterassistiertes Gangtraining in der klinischen Rehabilitation bearbeitet. Ein Anwendungsfeld betrifft die im Eigentaining durchgeführte Schulung von Orientierung und Mobilität von blinden und sehingeschränkten Personen. Zu dieser Schulung soll eine auf dem Roboter on-board durchgeführte bild- und sprachbasierte Szenenanalyse mit Interaktion erfolgen, wie z. B. zur Art von Objekten im Nahfeld des Patienten auch unter Berücksichtigung von für Außenumgebungen typische Situationen, wie z. B. Bordsteine oder Treppenstufen. Ziel dieser Arbeit ist die Untersuchung, wie basierend auf aktuellen Verfahren zur bild- und sprachbasierten Szenenanalyse mit Interaktion ein gemeinsames Szenenverständnis von Patient und Roboter aufbaubar ist.



Detaillierte Aufgabenstellung:

- Ermittlung relevanter Use Cases für gegenseitige Anfragen/Aufgaben für Mensch und Roboter nach Absprache mit klinischen O&M-Rehalehrern (bspw.: „Welches Objekt können Sie vor sich mit dem Blindenstock ertasten?“ oder „Welches der sich vor Ihnen befindenden Objekte ist ein Sofa?“)
- Recherche & Aufarbeitung relevanter Modelle zur bild- und sprachbasierten Szenenanalyse im Nahfeld einer Person, auch unter Nutzung eines Blindenstocks
- Vergleich der möglichen Lösungsansätze hinsichtlich Leistungsfähigkeit und benötigter on-board Rechentechnik des Roboters sowie ggf. notwendiger Edge Computer
- Aufbau einer modularen Verarbeitungspipeline zur bild- und sprachbasierten Szenenanalyse mit wechselseitiger Interaktion
- Auswahl & Kombination vorgefertigter Teillösungen entsprechend der auf den verwendeten Plattformen zur Verfügung stehenden Rechenressourcen (MetraLabs TORY, Unitree Go1, Unitree Go2)
- Inbetriebnahme des Systems, bestehend aus
 - Spracheingabe (STT) & Sprachausgabe (TTS)
 - Bildverarbeitung mit Objekt- bzw. Szenenanalyse
 - Erkennung der Person in 3D-Roboterweltkoordinaten
 - Beschreibung der Objektposition relativ zur Nutzerposition
- Nutzung manuell gesteuerter Beobachtungspositionen des Roboters in Funktionstests zur Evaluation des Systems hinsichtlich
 - Leistungsfähigkeit verschieden kombinierter Ansätze
 - korrekt erfasster Objekte und der Verbindung von Objekt und Text
 - Interaktionsstrategien mit dem Nutzer
 - Berechnungszeit und Latenzen
- Anfertigen der Ausarbeitung und Präsentation der Ergebnisse nach den Vorgaben des FG NIKR

Voraussetzungen und Vorkenntnisse:

- Erfolgreicher Abschluss der Vorlesung Neuroinformatik
- Programmiererfahrung in Python

Betreuer: M.Sc. Nicholas Heyer (nicholas.heyer@tu-ilmenau.de)
M.Sc. Söhnke Fishedick (soehnke.fishedick@tu-ilmenau.de)
Betr. Hochschullehrer: Prof. Dr. H.M. Groß
Bearbeiter: Moritz Drescher