



GAL

Niedersächsischer Forschungsverbund Gestaltung altersgerechter Lebenswelten

Ambientes Interface zur multimodalen Ein- und Ausgabe für einen persönlichen Haushaltsassistenten (PAHA)

Christian Rollwage¹, Tobias Herzke¹, Dr. Rainer Huber¹, Prof. Dr. Jörg Bitzer¹

¹ HörTech gGmbH, Oldenburg

Hintergrund

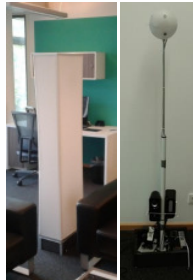
Grundideen von AP1

Entwicklung eines **persönlichen Haushaltsassistenten zur Erinnerung an anstehende Termine** unter den Gesichtspunkten:

- **Multimodal**
 - Eingabemöglichkeiten: *Sprache, Tablet, PC*
 - Ausgabemöglichkeiten: *optische und akustische Reize*
- **Ambient**
 - Integration in bestehende Möblierung
- **Ökonomisch tragbar**
 - Keine Umbaumaßnahmen notwendig (proprietäres System)
- **Individuell**
 - Anpassung der Erinnerungsmechanismen
- **Stationär**
 - Verwendung ausschließlich innerhalb eines Raumes

» Idee:

- **Stehlampe**
 - Unterbringung einiger notwendiger Komponenten zur Gewährleistung der Ein- und Ausgabe der Modalitäten
 - Versorgt über ein Daten- und ein Stromkabel
- **PC zur Steuerung der Lampe und Kalenderfunktion**



Software

Master Hearing Aid

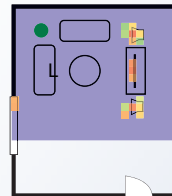
Software-Plattform zur Echtzeit-Audiosignalverarbeitung mit **niedriger Input-Output-Latenz**

- **Implementierung in C++**
 - Echtzeitfähige Implementierung der Algorithmen
- **Plugin-Konzept**
 - Unabhängige Algorithmen werden zu einer Signalverarbeitungskette zusammengefasst:
 - **Störquellenerkennung**
 - Erkennt räumlich stationäre Quellen (s. unten)
 - **Ereigniserkennung**
 - Erkennt Klatschen zum Start des Eingabedialogs
 - **Spracherkennung**
 - Erkennt Eingaben des Benutzers während des Eingabedialogs
 - **Dialogmodul**
 - Kombiniert Sprachausgabefragmente zu vollständigen Sätzen und gibt diese aus
 - Speichert die akustische Terminbeschreibung des Benutzers
 - **Kommunikationsmodul**
 - Sendet erkannte Ereignisse und Spracheingaben asynchron an den GAL-Controller
- **Einheitliche Konfigurationsschnittstelle**
 - Alle Algorithmen werden durch den GAL-Controller über dieselbe Schnittstelle einheitlich gesteuert

Technische Umsetzung

Akustische Aufnahme und Wiedergabe

- **achtkanaliges Mikrofonarray in Kugelform**
 - Verstärkung und Digitalisierung (ADAT elektrisch) im Array
 - Sehr gutes Signal-Rausch-Verhältnis
- **zwei leistungsfähige Lautsprecher mit vorgeschaltetem DA-Wandler**
 - Gespeist durch SPDIF-Stream



Beispiel einer Langzeitkarte mit unterschiedlichen Quellen

Erkennung stationärer Quellen

Spracherkennung funktioniert nur bei ruhiger Umgebung

→ System zur Detektion von Fernseher, Radio und weiterer Geräuschquellen implementiert:

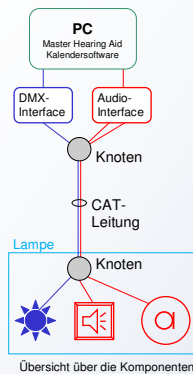
- Positionsschätzung mittels Lokalisationsalgorithmus: Global Coherence Field (GCF)
- Zusammenfassen der ermittelten Positionen über einen definierten Zeitraum (fuzzy c-means-Clustering)
- Abhängig von Sprachabwesenheit Übertragung der Cluster in eine Langzeitkarte
- Mittelung der aktuellen Karte mit der Langzeitkarte, um etwaige Verschiebungen der Lampe zu detektieren

Optische Reizdarbietung

- **LED-Scheinwerfer**
 - Ca. 150 Einfarbige 1 Watt LEDs aus der Bühnen-Lichttechnik

Weitere Technik

- **Knoten**
 - Verteilung aller Ein- und Ausgangssignale in Richtung der Lampe und vice versa auf die Adern eines CAT-Kabels (ermöglicht weite Distanzen zum Steuer-PC zu überbrücken)



Erkenntnisse

- **Testphase zeigte gutes Zusammenspiel aller Hardware-Komponenten**
- **Softwareoptimierung führte zu einem funktionierenden Gesamtsystem**
- **Erkennung stationärer Quellen scheint sicher**