Neuronale Verfahren zur Lokalisation und Bewertung von Handgesten in Real-World-Szenen

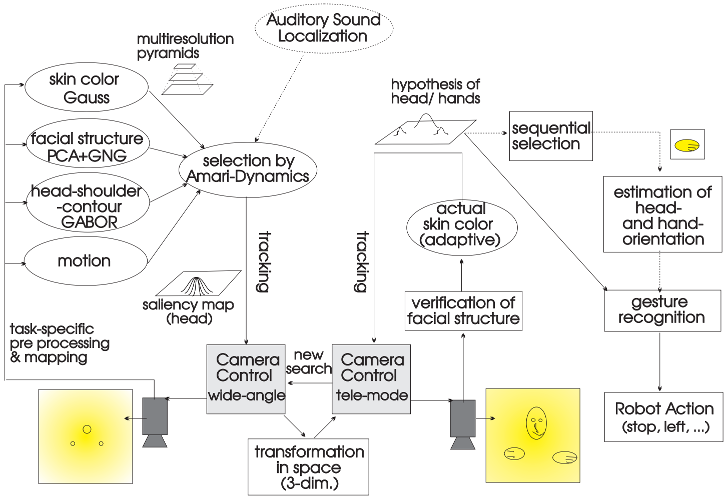
Fachgebiet Neuroinformatik
Technische Universität Ilmenau
PF 100565
Tel. +49 3677 / 691307

Anja Brakensiek, Ulf-Dietrich Braumann, Andrea Corradini, Hans-Joachim Böhme, Horst-Michael Groß
E-Mail: anja@informatik.tu-ilmenau.de

GESTIK-Projekt
gefördert durch das TMWFK

Zusammenfassung:

Für eine natürliche gestenbasierte *Mensch-Maschine-Kommunikation* in Real-World-Umgebungen ist zum einen die robuste Lokalisation eines potentiellen Nutzers (Gesicht) und zum anderen die Detektion der Hände eine elementare Voraussetzung.
Es werden 4 sich ergänzende Merkmale (Hautfarbe, Gesichtsstruktur, Kopf-Schulter-Silhouette, Bewegung) zur Nutzerlokalisation verwendet.
Erst die Kombination der Verfahren liefert unter beliebigen Bedingungen (Beleuchtung, Szene) die erforderliche Robustheit, auf die die Handdetektion und Gesteninterpretation aufbauen kann.
Das *Ziel* ist eine intuitive gestenbasierte Kommunikation (hier: Navigationsanweisungen) mit unserem mobilen Robotersystem MILVA, welches beispielsweise als Transportassistent agieren könnte.



Modellarchitektur

Nutzerlokalisation

Originalbild

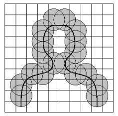
Farbadaptation:
- Manipulation der Punktwolke (Farbtripel) des Bildes derart, daß der Schwerpunkt unbunt ist
- Ziel: Realisierung stabiler Farbempfindungen

Berechnung auf verschiedenen **Auflösungsebenen** (Pyramide)

Durchschnitts-gesicht

farbadaptiertes Bild mit Ergebnissen der **Gesichtsstrukturdetektion**

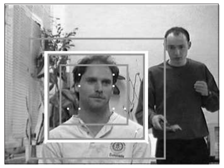
Merkmalsextraktion:
PCA (Eigenfaces)
Klassifikation:
supervised **GNG**

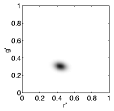


KSP-Modell

die am sichersten detektierte **Kopf-Schulter-Kontur**

Merkmalsextraktion:
Gaborfilterarrangement
Klassifikation:
spezielles Vergleichsmaß mit Prototypen





Hautfarbmodell (chromatisch)

Hautfarbklassifikationsergebnis

die Klassifikation erfolgt über eine 2-dim. Gaußfunktion, die die Verteilung einer intensitäts-normierten Hautdatenbank approximiert

Integration der Differenzbilder einer Sequenz
-> Bewegungsregionen
Ausblick: zeitl. Abfolge -> dynamische Gesten

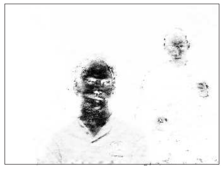
Bewegungsinformation

Verknüpfung über **dynamische Neuronale Felder**

Bestimmung der **Position**
(x- und y-Koordinate, Entfernung)

nutzerspezifisches Hautfarbmodell
wird erstellt:
Anpassung der Gaußfunktion an die aktuelle Farbverteilung

Hautfarbklassifikation
(adaptiertes Modell)



Growing-Neural-Gas-Netzwerk (GNG)

- Kombination aus Growing-Cell-Structures und Neural-Gas
- topologiebeschreibendes Netzwerk
- im überwachten Modus starke Ähnlichkeit mit RBF-Netzwerken
- Initialisierung des Netzwerkes mit 2 Neuronen; automatisches Einfügen von Neuronen und Kanten in Abhängigkeit von der Statistik der Daten

Handdetektion und Gestenerkennung

- z.Z. erfolgt die Detektion der Hände rein hautfarbbasiert
- die Erkennung der (statischen) Geste beruht vorerst auf der relativen Lage von Kopf und Händen

- die Geste (z.B. 'halt', 'links rum') soll in ein Kommando zur Roboternavigation umgesetzt werden (s. MILVA)

MILVA

Präsentiert auf dem 3. Internationalen Workshop „Neuronale Netze in der Anwendung" (NN'98), 12./13. Februar 1998, Magdeburg