



Teil einer Vision

Auch in den Projekten SIPaDIM und SIKOWO werden, wie in SIPBILD, adaptive Aktionssysteme entwickelt, die die Umwelt erfassen und darauf reagieren. Durch das Zusammenspiel dieser Systeme können gleichzeitig Bilder und Sprache verarbeitet werden. Daraus ergeben sich innovative Anwendungsmöglichkeiten, die Ihr Leben erleichtern und komfortabler gestalten sollen. Eine adaptive und intuitive Kommunikation ist zum Beispiel unerlässlich zur Steuerung von Servicerobotern. Sie möchten mehr über SIPBILD erfahren? Weitere Informationen finden Sie unter www.forsip.de.

Das SIPBILD-Team

PROJEKTLEITER

Prof. Dr. Bernd Radig

MITARBEITER

Dipl.Inf.(FH), M. Sc. Simone Hämmerle

Dipl.-Inf. Matthias Wimmer

TU MÜNCHEN

Lehrstuhl für

Bildverstehen und

Wissensbasierte Systeme

Boltzmannstraße 3

85748 Garching

KONTAKT

Tel +49 (0)89-28917754

Fax +49 (0)89-28917757

sipbild@forsip.de

www.forsip.de/sipbild



In nahezu jeder Zukunftsvision trifft man auf intelligente Computer, die mit dem Menschen reden, auf ihn eingehen, seine Bedürfnisse erkennen und seine Wünsche erfüllen. Der Bayerische Forschungsverbund für Situierung, Individualisierung und Personalisierung in der Mensch-Maschine-Interaktion (FORSIP) bietet Forschern aus verschiedensten Wissenschaftsbereichen eine Plattform zur Verwirklichung dieser Vision. Hier arbeiten Experten aus der Sprach- und Bildverarbeitung, der Systemtechnologie, der künstlichen Intelligenz und aus dem Datenbankbereich sowie Spezialisten für betriebliche Anwendungen zusammen an der Zukunft der Mensch-Maschine-Interaktion.

Gefördert durch
das Bayerische
Staatsministerium
für Wissenschaft,
Forschung und Kunst

SPRECHER

Prof. Dr. Werner Kießling
Universität Augsburg

REFERENTIN

Dr. Sibylle Peuker
abayfor

Arcisstr. 21
80333 München

KONTAKT

Tel +49 (0)89-289 225 38

Fax +49 (0)89-289 225 89

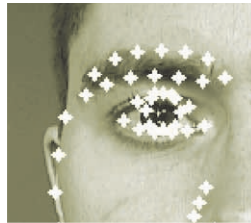
info@forsip.de

www.forsip.de



SIPBILD

Mimik- und Gestikerkennung
in Videobildfolgen



Jürgen Heile Kommunikationstraining

SIPBILD



Leichtigkeit, Bewegung und Individualität sind Eigenschaften, die Menschen beim Tanz auszeichnen. Harmonisch, spielerisch und mit allen Sinnen sollte auch die Interaktion zwischen Mensch und Maschine funktionieren. Eine Vision, an der zahlreiche Wissenschaftler bei FORSIP arbeiten.

Sie werden Augen machen!

Leicht verständlich, flexibel und situationsbezogen – so sollte die Mensch-Rechner-Kommunikation sein? Mit SIPBILD erfüllen wir Ihnen diesen Traum: Die hier entwickelten Verfahren gestalten die Kommunikation zwischen Mensch und Rechner adaptiv und intuitiv. Statt einer mühsamen Anpassung Ihrerseits, liest Ihnen der Computer Wünsche vom Gesicht ab oder reagiert auf Ihre Anweisungsgesten.



Mensch-Maschine-Kommunikation im Blickpunkt

Immer noch ist für viele Menschen der Umgang mit Rechnern schwierig: Die Kommandos erfolgen über einige wenige Eingabegeräte wie Tastatur und Maus und zudem in einer festgelegten Art und Weise. Mit normalen menschlichen Kommunikationsverhalten hat diese Verständigung wenig zu tun. Die Kommunikation ist keineswegs intuitiv oder an die gegenwärtige Situation bzw. den Benutzer angepasst. Wenn man die aktuelle Umgebungssituation eines Rechnersystems erfasst, kann man den Dialog mit dem Anwender adaptiv und intuitiv gestalten und dadurch den Komfort des Anwenders erhöhen.



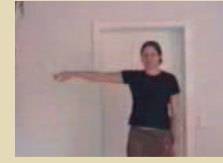
Ausschnitte aus der Bildfolge einer Geste: Die Dynamik der Geste wird erfasst. Arm anheben bedeutet zum Beispiel „Helligkeit erhöhen“.

Mit dem Finger schnippen und das Licht geht an

Der Benutzer soll dem Computer auf einfache Weise mitteilen können, was er möchte. SIPBILD hat es sich deshalb zur Aufgabe gemacht, die Umgebung eines Rechnersystems – einschließlich des Benutzers – anhand von Videoaufnahmen visuell zu erfassen. Diese Informationen werden anschließend ausgewertet und liefern dem Computer so das nötige Wissen, um die Wünsche des Benutzers angemessen erfüllen zu können.

Gestikerkennung

So kann der Benutzer den Computer beispielsweise durch Gesten dazu veranlassen, bestimmte Aufgaben zu erledigen. Gestikerkennung kann zum Beispiel zur der Steuerung des Wohnkomforts in High-Tech-Häusern angewendet werden. Der Benutzer steuert die Beleuchtung im Wohnraum, indem er Anweisungen in Form von Armbewegungen gibt. Das Bewegen eines Armes nach oben signalisiert dem Rechner „Helligkeit erhöhen“.



SIPBILD „sieht“, wie sich der Arm nach oben bewegt und definiert die Bewegung als „erhöhen“.

Hierzu wird wie folgt vorgegangen:

- Erfassen der Umgebung eines Rechnersystems durch das Aufnehmen von Videosequenzen
- Berechnung der Differenzen von Bild zu Bild
- Generierung von räumlichen und zeitlichen Bewegungsmasken
- Berechnung von Bewegungsmerkmalen
- Klassifizierung der beobachteten Bewegungen in Anweisungsgesten
- Ausführen des jeweiligen Befehls

Machen Sie doch nicht so ein Gesicht

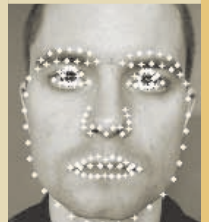
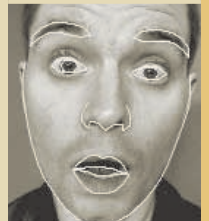
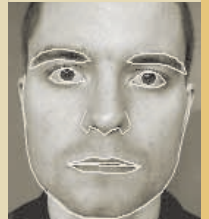
Mimikerkennung

Auch die Erfassung der Mimik eines Benutzers kann von großem Interesse sein. Wissen über die emotionale Verfassung des Anwenders hilft beispielsweise dem Verkaufsgenossen aus dem Teilprojekt COSIMA, die Verkaufsverhandlung zu optimieren.

Wie geht SIPBILD hierzu vor?

Zunächst wird ein Punktemodell der wesentlichen Gesichtskonturen bei neutralem, lachendem und überraschtem Ausdruck erstellt. Dieses Modell wird dann mit den jeweiligen Konturen aus den aufgezeichneten Videosequenzen abgeglichen. Der Rechner erhält so Informationen über den emotionalen Zustand des Benutzers. So kann der Verkaufsgenosse aus dem Teilprojekt COSIMA nach Auswertung der Daten erkennen, ob ein Benutzer mit dem getätigten Kauf bzw. dem erstandenen Produkt zufrieden ist und gegebenenfalls die Verkaufsverhandlung optimieren.

Mimikerkennung bei SIPBILD: Das Gesichtsmodell deformiert sich und passt sich an die Gesichtszüge an.



Anhand der wesentlichen Gesichtskonturen wird ein Punktemodell erstellt. Dieses Modell dient für den Abgleich mit den aktuell aufgenommenen Gesichtsbildern.