

# Potenzial von MMK-Technologien für zukünftige Fahrzeug-Bedienkonzepte

Klaus Bengler

BMW Group Forschung und Technik GmbH, Hanauerstr. 46, D-80992 München

## Einleitung

Die mittlerweile mehr als hundertjährige Entwicklung des Automobils ist seit ihrem Beginn auch mit Fragen der Bedienung und der geeigneten Darstellung von Informationen und deren Wahrnehmung durch den Fahrer verknüpft. Somit ergibt sich eine interessante Überschneidung mit den aktuellen Forschungsarbeiten am Lehrstuhl für Mensch-Maschine-Kommunikation.

In den letzten Jahrzehnten der Automobilentwicklung konnte eine deutliche Zunahme von Anzeigen- und Bedienelementen im Fahrzeug festgestellt werden. Diese Entwicklung war unter anderem getrieben von den durch die Mikroelektronik ermöglichten Funktionsausweitungen.

Um trotz dieses Anstiegs eine optimale Bedienbarkeit für den Fahrer zu gewährleisten wurde 2001 bei BMW das Konzept "iDrive" eingeführt. Das Konzept zeichnet sich aus durch die gezielte Unterscheidung der Funktionalität in direkt bedienbare Funktionen und Menüfunktionen.

Fahrbezogene Funktionen sind im direkten Zugriff zum großen Teil in der Umgebung des Lenkrades angeordnet. Häufig genutzte Funktionen können direkt im Bereich des Lenkrads und der Instrumententafel erreicht werden. Vor allem Informations- und Kommunikationsfunktionen sind in einem Bildschirmmenü dargestellt und werden mittels eines zentralen Controllors genutzt. Die Integration der Sprachverarbeitung stellt einen weiteren wesentlichen Bestandteil des iDrive-Konzeptes dar. Sie ermöglicht direkte Funktionsaufrufe ohne einen Menüpfad zu durchlaufen und trägt damit zur Intuitivität und Bediensicherheit bei. Das Sprachverarbeitungssystem ist sehr zuverlässig und erlaubt eine redundante Bedienung von 60 % des gesamten Funktionsumfangs.

## Erkennungstechnologien und Fusionsansätze

Es wird deutlich, daß mittlerweile neben der manuell-visuellen Bedienung über Taster, Schalter, Hebel und andere "klassische" Bedienelemente zunehmend Erkennungstechnologien eine ausschlaggebende Rolle spielen. Sie werden zudem häufig unter schwierigen Umgebungsbedingungen neben der eigentlichen Fahraufgabe von einer Vielzahl unterschiedlicher Fahrer genutzt. [1][2] Hinzu kommt aber, daß neben den bekannten Erfolgsfaktoren wie z.B. der Erkennungsleistung zunehmend die Dialogqualität eine ausschlaggebende Rolle spielen wird [3].

Zukünftige Herausforderungen liegen also darin, aktuelle Forschungsansätze der Mensch-Maschine-Kommunikation auf ihren Einsatz im Automobil hin zu untersuchen.

Prinzipiell zeigt sich, daß die alleinige Integration weiterer Bedienmöglichkeiten, die zusätzlich zur Verfügung stehen würden, nur einen Ansatz darstellt. Ebenfalls sinnvoll scheint die Fusion von multimodalen Kontextinformationen, die begleitend zur Benutzung der gewählten Leitmodalität im Fahrzeug zur Verfügung stehen. Als ein Beispiel seien hier die Arbeiten zur Gestikerkennung zur direkten Bedienung oder als Kontextinformation in einem Fusionsansatz genannt, die in Kooperation mit dem Lehrstuhl MMK durchgeführt wurden. Es zeigt sich, daß bei einem zielgerichteten Einsatz von Handgesten die Ablenkung des Fahrers merklich gesenkt werden kann, da die Suche nach einem Bedienelement entfällt [4] [5]. Den in diesem Zusammenhang auftretenden Fragestellungen der Mustererkennung und Segmentierung stehen vielversprechende Lösungsansätze gegenüber, wie Abbildung 1 verdeutlicht.



Abbildung 1: Erkennung der Hand als relevante Informationseinheit vor verwaschenem Hintergrund

Im Sinn eines Fusionsansatzes kann auch durch die Berücksichtigung von Kopfgesten als unbewusster Co-Artikulation im Zusammenhang mit Spracheingabe die Erkennungsleistung der Spracherkennung im Fall von JA/NEIN Antworten des Nutzers deutlich erhöht werden. Um dieses Ziel zu erreichen, ist es ausschlaggebend, robuste Algorithmen zur Detektion des Kopfes bzw. Gesichtes und dessen Bewegung zu entwickeln. Die eingeschränkten Freiheitsgrade des Fahrers begünstigen diesen Ansatz, allerdings stellt die Robustheit gegen stark variierende Umgebungsbeleuchtung wiederum eine große Herausforderung dar. Abbildung 2 veranschaulicht wiederum den Lösungsansatz und den Reifegrad der Lösung in einer automobilen Umgebung.



**Abbildung 2:** Erkennung des Kopfes, der Gesichtsregion und relevanter Gesichtsm Merkmale in einer Fahrzeugumgebung

Ein weiterer – sehr vielversprechender - Fusionsansatz besteht in der Auswertung prosodischer Merkmale im Zusammenhang mit Spracheingaben. Wiederum am Beispiel von JA/NEIN Eingaben läßt sich zeigen, daß die Berücksichtigung mehrerer Basisklassifikatoren die Robustheit erhöht werden kann. Dies gilt gerade unter geräuschvollen Umgebungen und bei undeutlich gesprochenen Eingaben. Derzeit werden die Äußerungen als positiv – neutral – ärgerlich klassifiziert.

Für den Erfolg dieses Ansatzes ist ausschlaggebend, welche und wieviele Modalitäten nach welchen Verfahren fusioniert werden sollten [6] [7]. Für den Einsatz im Automobil sind gerade Fusionsansätze interessant, die auf bereits existierender Sensorik z.B. Freisprechmikrofon realisiert werden können.

Es wird deutlich, daß über die Berücksichtigung von Kontextinformationen die Mensch-Maschine Kommunikation robuster und damit natürlicher gestaltet werden kann. Dabei soll die Maschine nicht „vermenschlicht“ werden, sondern vielmehr bekannte Phänomene der menschlichen Kommunikation berücksichtigt werden, um die Erwartungskonformität zu erhöhen.

## Zusammenfassung

Zusammenfassend läßt sich sagen, daß die Mensch-Maschine-Kommunikation im Automobil anspruchsvolle Fragestellungen im Zusammenhang mit multimodaler Kommunikation aufwirft:

- Wie kann generell die Robustheit der Interaktion durch die Fusion verschiedener Interaktionskanäle erhöht werden ?
- Welches Potenzial stellt die Ausnutzung von Kontextinformationen (z.B. Emotionserkennung) dar, um die Robustheit eines Erkennungsprozesses zu erhöhen ?
- Welche Erkennungsansätze ermöglichen in Zukunft die zielgerichtete Nutzung umfangreicher und multilingualer Vokabulare für z.B. Zieleingabe oder MP3 ?

## Literatur

- [1] Zobl, Martin, M. Geiger, K. Bengler, M. Lang: A Usability Study on Hand Gesture Controlled Operation of In-Car Devices. HCI 2001 New Orleans, 2001.
- [2] Bengler, Klaus: Aspekte der multimodalen Bedienung und Anzeige im Automobil. In Th. Jürgensohn & K.-P. Timpe (Hrsg.) *Kraftfahrzeugführung*. (S. 195-206). Berlin, Heidelberg, New York: Springer, 2001.
- [3] Bengler, Klaus: Automotive Speech-Recognition - Success Conditions Beyond Recognition Rate. Proc. LREC 2000, Athens, (pp. 1357-1360), 2000.
- [4] Geiger, M., M. Zobl, K. Bengler, M. Lang: Intermodal differences in distraction effects while controlling automotive user interfaces. HCI 2001 New Orleans, 2001.
- [5] Akyol Saat, U. Canzler, K. Bengler, W. Hahn: Gesture Control for Use in Automobiles. In: Proceedings of the IAPR MVA 2000 Workshop on Machine Vision Applications , (pp. 349-352), 2000.
- [6] Althoff, Frank: *Ein generischer Ansatz zur Integration multimodaler Benutzereingaben*. Dissertation, TU-München, Lehrstuhl für Mensch-Maschine-Kommunikation, 2004.
- [7] Hoch, Stefan, F. Althoff, G. McGlaun, G. Rigoll: Bimodal fusion of emotional data in an automotive environment, IN: Proceedings ICASSP 2005, IEEE Signal Processing Society, March 2005, Philadelphia, USA, 2005.