

# Framework für multimodale Bediensysteme in der Automatisierungstechnik

Dipl.-Ing. Thomas Stiedl

Institut für Automatisierungs- und Softwaretechnik  
Universität Stuttgart  
Pfaffenwaldring 47  
70550 Stuttgart  
thomas.stiedl@ias.uni-stuttgart.de

**Abstract:** The usability of technical systems is moving more and more into focus, since increasing functionality is not sufficient to acquire customers on a long-term basis. With constantly growing functionality an elementary requirement for the development of human-machine interfaces (HMI) consists of enabling users to cope with the complexity of industrial automation systems. Since existing concepts are reaching their limits, modern user control concepts and a stronger consideration of the users tasks are necessary in the field of industrial automation. Technical systems that are intuitively and easy to use lead to an improvement of quality and productivity. Regarding usability and intuitive control the use of several senses of humans promises advantages for the human-machine interaction. However realizing multimodal user interfaces results in a higher complexity for developers of those systems. Within this paper, a conceptual framework for the development of multimodal user interfaces in the context of industrial automation is presented.

## 1 Mensch-Maschine-Schnittstellen in der Automatisierungstechnik

In der Automatisierungstechnik ist ein deutlicher Trend zu PC-basierten Steuerungen zu erkennen. Stand bei älteren Automatisierungssystemen die Bedienung der Maschine über elektronische Bedienelemente im Vordergrund, werden heute überwiegend Dialogsysteme mit grafischen Bedienapplikationen verwendet. Diese nutzen in der Regel die aus Büroanwendungen bekannten Bedienelemente, um Informationen zu strukturieren und die Interaktion mit dem Benutzer zu steuern. Analog dazu werden entsprechende Bediengeräte wie Tastatur und Maus für den Industrieinsatz ausgelegt. Diese können aufgrund ihrer flexiblen Einsetzbarkeit für unterschiedliche Anwendungen verwendet werden. Folglich ist die Funktionalität der Benutzungsschnittstelle nicht mehr durch die Anzahl mechanischer Taster oder Schalter begrenzt. Neue Funktionalität lässt sich mit wenig Aufwand in Software implementieren [Da06] und nahezu kostenfrei vervielfältigen [Zü05]. Aufgrund des technischen Fortschritts und wegfallender Einschränkungen durch elektronische Bedienelemente ist ein stetiger Anstieg des Funktionsumfangs bei Automatisierungssystemen zu beobachten. Damit wird der Mensch als Nutzer vor vollkommen neue Herausforderungen gestellt, welche die Grenzen seiner mentalen Ressourcen in Zukunft überfordern werden [Kr01].

Dies führt zu einer immer größeren Diskrepanz zwischen dem technisch Machbaren und dem durch den Menschen Bedienbaren. Infolgedessen wird es für Benutzer dieser Systeme immer schwerer, sie zu beherrschen. Andererseits bleiben bei einer derartigen Gestaltung der Benutzungsschnittstellen Fähigkeiten des Menschen ungenutzt. Im Gegensatz zur zwischenmenschlichen Kommunikation, bei der unterschiedliche Sinne des Menschen zur Wahrnehmung und Handlung eingesetzt werden, berücksichtigt die Mensch-Maschine-Interaktion bei Automatisierungssystemen nur wenige Sinne. Die Wahrnehmung erfolgt in der Regel durch visuelle Abbildungen und die Aktionen des Benutzers werden in Form taktiler Eingaben getätigt. Diese Einschränkung ist für die Automatisierungstechnik, in der neben der indirekten Interaktion über Benutzungsoberflächen auch Eingriffe direkt am Automatisierungssystem oder am technischen Prozess notwendig sind, von Nachteil. Bei Wartungsarbeiten werden beispielsweise zum Austausch defekter Bauteile sowohl die Hände des Technikers, als auch dessen visuelle Aufmerksamkeit gefordert. Wenn zudem Sensordaten des Automatisierungssystems für die Kalibrierung abgefragt werden müssen, ist ein Wechsel an das entsprechende Bediengerät erforderlich. Dieses muss ebenfalls manuell betätigt werden und liefert seine Rückmeldung über eine visuelle Darstellung. Die dadurch erforderliche sequenzielle Ausführung dieser Arbeitsschritte verzögert den Wartungsvorgang und erfordert ein häufiges Umdenken des Technikers, wodurch dessen Arbeitsbelastung erhöht wird [ALK01].

## **2 Konzept zur Entwicklung multimodaler Bediensysteme**

Um eine Parallelisierung von Arbeitsabläufen zu ermöglichen, wie es bei dem oben genannten Beispiel der Wartung hilfreich wäre, ist die Nutzung mehrerer Sinne des Menschen bei der Interaktion mit dem Automatisierungssystem erforderlich. Durch die Erweiterung der Kommunikation zwischen Mensch und Maschine um zusätzliche Modalitäten kann das verwendete Eingabemedium auf die Umgebungssituation und die jeweilige Aufgabe des Benutzers abgestimmt werden. Ebenso wird, entsprechend der zwischenmenschlichen Kommunikation, durch Sprache oder Gestik eine Verbesserung der Intuitivität und Bedienbarkeit des technischen Systems ermöglicht [Ov00], wodurch auch die Sicherheit bei der Bedienung erhöht werden kann [Ab03]. Der Einsatz multimodaler Bediensysteme bietet daher Vorteile für die immer komplexeren Systeme der Automatisierungstechnik.

Bei der Entwicklung multimodaler Bediensysteme ist es notwendig, die spezifischen Randbedingungen der Automatisierungstechnik zu berücksichtigen. Aufgrund der Komplexität der Automatisierungssysteme ist eine flexible Anpassung der Bediensysteme an die jeweiligen Aufgaben und Arbeitsabläufe der Benutzer erforderlich. So kann durch eine benutzerspezifische Auswahl der Funktionen die steigende Funktionalität der Automatisierungssysteme beherrscht werden. Die Benutzer eines Automatisierungssystems können in der Regel anhand ihrer jeweiligen Aufgaben zu Gruppen zusammengefasst werden. Hier ist die Einführung von Rollen sinnvoll, welche spezifische Benutzergruppen mit definierten Aufgaben repräsentieren. Zusätzlich müssen unterschiedliche Umgebungsbedingungen, wie beispielsweise schwer zugängliche Anlagenteile, Umgebungslärm etc., berücksichtigt werden.

Um multimodale Bediensysteme mit entsprechender Flexibilität entwickeln zu können, ist eine direkte Verknüpfung zwischen Bedienelement und Funktionalität des Automatisierungssystems nicht mehr möglich, wodurch neue Konzepte für die Entwicklung von Bediensoftware notwendig werden. Aus diesem Grund wurde am IAS ein dreistufiges Verfahren, zur multimodalen Repräsentation der Informationen eines Automatisierungssystems, entwickelt:

1. Die Funktionalität des Automatisierungssystems wird mithilfe einer Beschreibung der Funktionen und Daten spezifiziert und über eine Diensteschnittstelle zur Verfügung gestellt.
2. Durch Spezifikation von Dialogen werden die Dienste in Abhängigkeit ihrer Semantik strukturiert. Für die jeweiligen Benutzerrollen werden Sichten auf die Funktionalität definiert.
3. In Abhängigkeit der Benutzerrolle werden für die Dienste geeignete Modalitäten und Bedienelemente zur Interaktion ausgewählt.

Bei der Entwicklung einer multimodalen Bediensoftware mithilfe des oben beschriebenen Verfahrens muss zunächst der Zugriff auf die Funktionen und Daten des Automatisierungssystems realisiert werden. Damit die Funktionsaufrufe über das Bediensystem unabhängig von den jeweiligen Kommunikationsparametern des Automatisierungssystems erfolgen können, müssen die spezifischen Funktionsaufrufe gekapselt werden. Der Zugriff auf die Funktionen und Daten des Automatisierungssystems erfolgt daher über die Diensteschnittstelle eines entsprechenden Adapters.

Im zweiten Schritt werden alle verfügbaren Dienste in Form von Dialogen strukturiert. Zur Verringerung der Komplexität bei der Bedienung werden Sichten für die unterschiedlichen Benutzerrollen definiert, wodurch für die jeweiligen Aufgaben relevante Dienste ausgewählt werden können. Um die Arbeitsabläufe der Benutzer bei der Gestaltung der Interaktion zu berücksichtigen, werden den Diensten zusätzlich Interaktionsformen zugeordnet. Auf diese Weise kann z. B. benutzerspezifisch die Sprachsteuerung für bestimmte Funktionsaufrufe aktiviert werden.

Für die Benutzerinteraktion muss im dritten Schritt eine geeignete Repräsentationsform erstellt werden. Hierzu gehört zum Einen die Transformation der Informationen des Automatisierungssystems in eine für den Menschen verständliche Form, beispielsweise in grafische Bedienelemente oder Sprachbefehle, zum Anderen müssen die Eigenschaften der eingesetzten Bediengeräte berücksichtigt werden. Die multimodale Informationspräsentation muss für die verschiedenen Interaktionskanäle jeweils geeignete Bedienelemente bereitstellen. Da eine parallele Interaktion über mehrere Sinne ermöglicht werden soll, ist zudem eine Synchronisation der jeweiligen Darstellungen erforderlich.

Ein aus den genannten Systemelementen bestehendes multimodales Bediensystem lässt sich in die drei Schichten, Informationspräsentation, Dialogsteuerung und Schnittstelle des Automatisierungssystems, unterteilen (vgl. Abbildung 1).

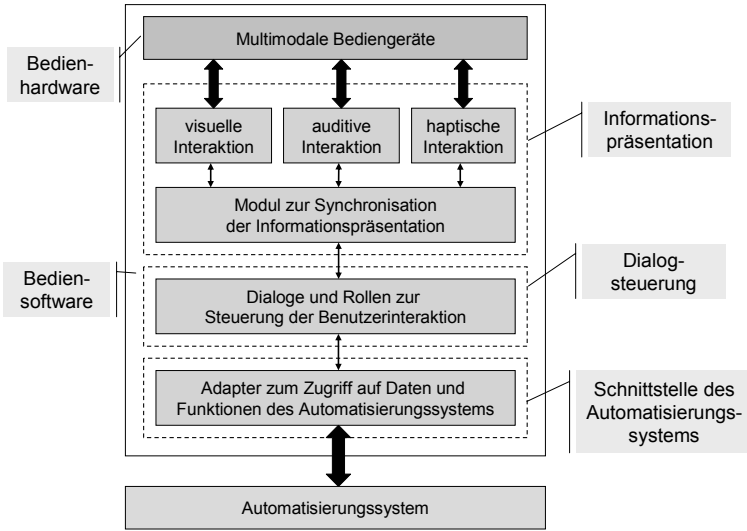


Abbildung 1: Architektur eines multimodalen Bediensystems

Basierend auf dem vorgestellten Ansatz wurden beispielsweise multimodale Bediensysteme zur Bedienung, Diagnose und Wartung eines industriellen Kaffeeautomaten [St05] sowie einer Sortieranlage entwickelt. Dabei wird ein Head-Mounted-Display verwendet, um dem Techniker Prozessdaten und Funktionen einer Sortieranlage anzuzeigen, zusätzlich können über eine entsprechende Sprachsynthese unterstützende Hinweise zur Bedienung gegeben werden. Eingaben erfolgen wahlweise per Stift oder durch Sprachbefehle. Definierte Funktionen der Anlage, wie z. B. das Ausfahren eines Schiebers, werden durch Spracheingabe aufgerufen, wodurch während dieses Arbeitsschritts keine Handeingabe notwendig ist. So kann der Benutzer beispielsweise die zugehörige Lichtschranke justieren und gleichzeitig auf notwendige Informationen der Diagnoseapplikation zugreifen. Abbildung 2 zeigt beispielhaft die Realisierung eines multimodalen Bediensystems mithilfe eines Head-Mounted-Displays.

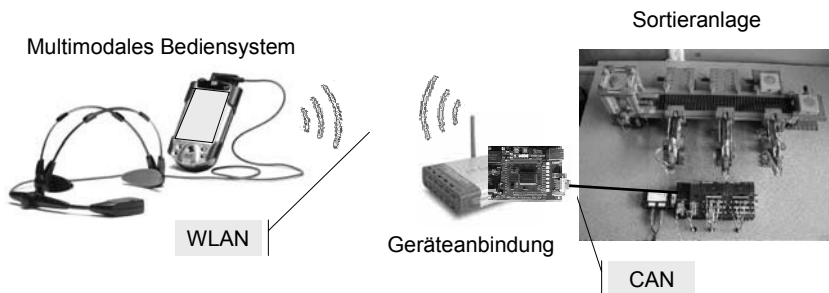


Abbildung 2: Multimodales Bediensystem zur Steuerung einer Sortieranlage

### 3 Framework für multimodale Bediensysteme

Die Erfahrungen bei der Realisierung multimodaler Bediensysteme für Automatisierungsgeräte haben gezeigt, dass durch die Vielfalt der verfügbaren Bediengeräte mit jeweils unterschiedlichen Schnittstellen ein großer Aufwand darin besteht, die Informationen spezifisch für das jeweilige Bediengerät darzustellen. Zudem steigt durch die Verwendung verschiedener Interaktionsmedien bei multimodalen Bediensystemen die Komplexität für die Erstellung derartiger Systeme. So müssen Entwickler neben der herkömmlichen GUI-Programmierung auch über Kenntnisse zur Spracherkennung und Sprachsynthese verfügen. Da dieses Wissen im Bereich der Automatisierungstechnik nicht zu den Kernkompetenzen der Entwickler zählt, ist eine entsprechende Unterstützung bei der Realisierung multimodaler Bediensysteme zur Bedienung, Diagnose und Wartung von Automatisierungssystemen notwendig.

Werden multimodale Bediensysteme mithilfe des in Kapitel 2 vorgestellten Verfahrens entwickelt, weisen diese eine gleichartige Systemfunktionalität auf. Neben der Anbindung von Automatisierungssystemen sind die rollenbasierte Einschränkung der Automatisierungsfunktionalität sowie die multimodale Informationspräsentation elementare Funktionen dieser Bediensysteme. Um die Gemeinsamkeiten bei der Entwicklung zu nutzen und eine einheitliche Systemarchitektur zu erhalten, wurde das Framework zur Erstellung multimodaler Bediensysteme (FMBS) entwickelt. Mithilfe von FMBS kann multimodale Bediensoftware durch entsprechende Konfiguration von Komponenten individuell für die jeweilige Bedienhardware, wie z. B. Head-Mounted-Display, Head-Set oder PDA, generiert werden.

Um eine Trennung zwischen Funktionalität und Informationspräsentation zu erreichen, ist eine Kapselung des Zugriffs auf die Automatisierungssysteme erforderlich. Das FMBS enthält vordefinierte Adapter, die zur Anbindung von Automatisierungssystemen parametrierbar sind. Auf diese Weise wird eine einheitliche Schnittstelle für verschiedene Kommunikationssysteme (z. B. CAN, PROFIBUS, Ethernet) realisiert. Abbildung 3 zeigt die Generierung eines Adapters zum Zugriff auf die Funktionen und Daten eines Automatisierungssystems am Beispiel der IAS-Sortieranlage.

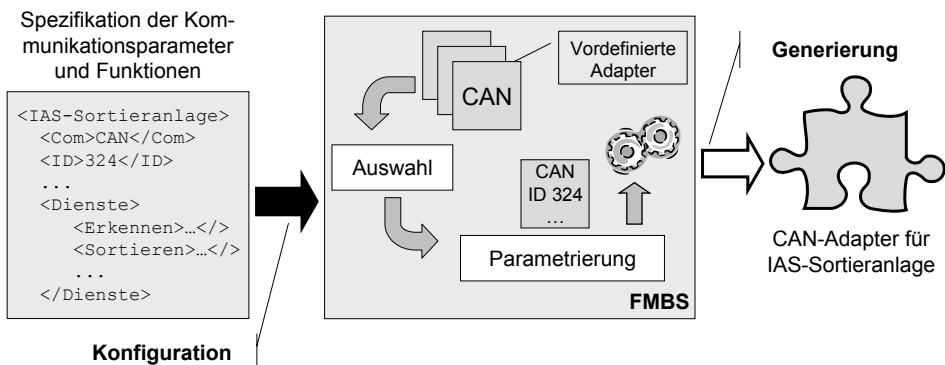


Abbildung 3: Generierung eines CAN-Adapters für die IAS-Sortieranlage mit dem FMBS

Die Kommunikationsparameter geben, den verwendeten Feldbus und die entsprechenden Zugriffsdaten (Netzwerkadresse, Port, ID etc.) an. Beispielsweise kann über einen CAN-Adapter auf einzelne CAN-Knoten einer Sortieranlage zugegriffen werden, um so aktuelle Prozessdaten auszulesen und definierte Funktionen auslösen zu können. Zur Gliederung der Funktionalität von Automatisierungssystemen existieren bereits Lösungen, wie z. B. OPC oder NOAH [Di00]. Da die Definition einer standardisierten Beschreibung nicht Kern des Konzepts ist, wird die bestehende XML-basierte Beschreibungssprache SDML erweitert [Ja05]. Mit dieser wird ein allgemeiner Dienst wie beispielsweise „Werkstück erkennen“ auf die syntaktische Beschreibung des Funktionsaufrufs für das jeweilige Automatisierungssystem abgebildet, z. B. als CAN-Nachricht.

Bei der Erstellung von Rollen dienen die Funktionsbeschreibungen als Basis für die Zuordnung von Diensten zu den jeweiligen Benutzergruppen. Mit diesem Mechanismus kann beispielsweise festgelegt werden, dass Dienste, die durch das Bedienpersonal genutzt werden dürfen, für die entsprechende Rolle *Anlagenbediener* sichtbar sind. Durch eine derartige Einschränkung der Funktionalität wird die allgemeine Diensteschnittstelle mit allen verfügbaren Diensten eines Automatisierungssystems benutzerspezifisch zugeschnitten (vgl. Abbildung 4). Die mithilfe des FMBS implementierte Bediensoftware beinhaltet nur noch Dienste, welche von den konkreten Benutzergruppen benötigt werden.

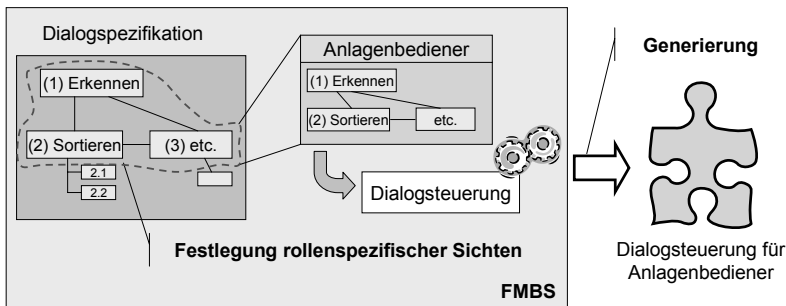


Abbildung 4: Benutzerspezifische Auswahl der Funktionalität

Um Informationen für mehrere Sinne geeignet darstellen zu können, enthält das Framework für jede Interaktionsform konfigurierbare Module, welche multimodale Bedienelemente bereitstellen. Diese können entsprechend der verwendeten Bediengeräte parametrisiert werden. Die Parametrierung der Informationsdarstellung erfolgt ausgehend von einer Beschreibung der Eigenschaften des jeweiligen Bediengeräts, diese enthält unter anderem die angebotenen Ein- und Ausgabemedien (z. B. grafische Ausgabe und Stifteingabe) sowie die konkreten Ausprägungen einer bestimmten Bediengeräteklasse. Ausprägungen eines grafischen Displays sind beispielsweise Seitenverhältnis, Bildauflösung, Farbtiefe und Bildwiederholfrequenz. Anhand dieser Eigenschaften wird durch das Framework eine Auswahl möglicher Bedienelemente getroffen. Anschließend werden dem Entwickler aufgrund der Eigenschaften der darzustellenden Dienste geeignete Bedienelemente vorgeschlagen.

Zu den Eigenschaften eines Dienstes gehören beispielsweise die Anzahl der Parameter sowie der Datentyp des Rückgabewerts. Der Entwickler wählt daraufhin das gewünschte Bedienelement aus und konfiguriert die Bedienoberfläche. Bei einer grafischen Ausgabe werden hierbei GUI-Widgets ausgewählt und die Gruppierung von Bedienelementen zu Container-Elementen festgelegt. Die Konfiguration eines Sprachdialogs legt unter anderem die Zuordnung von Sprachbefehlen (wie z. B. „Werkstück sortieren“) zu Menüpunkten sowie die Menütiefe fest. Nachdem die Informationsdarstellungen konfiguriert wurden, kann durch das Framework für multimodale Bediensysteme die entsprechende Bediensoftware generiert werden (vgl. Abbildung 5).

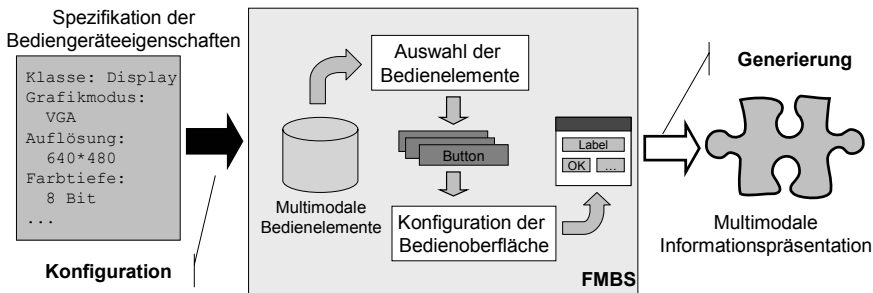


Abbildung 5: Bedienelemente für die multimodale Informationspräsentation

Die generierte Bediensoftware weist die in Kapitel 2 hergeleitete Architektur auf und enthält eine für das Automatisierungssystem parametrisierte Instanz eines Adapters, der über die Dienste-Schnittstelle alle notwendigen Funktionen des Automatisierungssystems bereitstellt. Zudem verfügt die multimodale Bediensoftware über eine Dialogsteuerung, welche auf die konfigurierten Rollen zugeschnitten ist. Für jeden Kommunikationskanal kapselt ein entsprechendes Modul die Transformationen der Ein- und Ausgabesignale. Dadurch wird eine Trennung der Informationsschicht von den übrigen Schichten der Bediensoftware erreicht. So wird beispielsweise die Sprachsteuerung komplett im Modul für auditive Interaktion gekapselt. Gibt der Benutzer z. B. einen Befehl durch Sprache ein, so müssen unter Umständen auch die übrigen Darstellungsformen aktualisiert werden. Hierfür ist die Synchronisation der Kommunikationskanäle erforderlich.

Das Framework für multimodale Bediensysteme sieht eine Werkzeugunterstützung für Entwickler vor, um die genannten Konfigurationsaufgaben durchführen zu können. Über einen grafischen Editor können beispielsweise Rollen angelegt werden. Anschließend werden die benötigten Funktionen per Drag-and-Drop den jeweiligen Rollen zugeordnet. Zusätzlich kann in den Eigenschaften eines Dienstes festgelegt werden, über welchen Kommunikationskanal dieser vom Benutzer aufgerufen werden kann. Für die unterschiedlichen Interaktionsformen kann zudem die Informationspräsentation konfiguriert werden. Bei einer auditiven Interaktion können unter anderem Sprachmuster der Benutzer angelegt werden, um die Spracherkennung möglichst fehlerfrei gewährleisten zu können. Das Softwarewerkzeug ermöglicht dem Entwickler mithilfe entsprechender Editoren, die Anordnung der Bedienelemente sowie die Navigationsstruktur der unterschiedlichen multimodalen Darstellungen festzulegen.

## 4 Zusammenfassung

Durch die Nutzung multimodaler Bediensysteme kann die Mensch-Maschine-Interaktion besser auf den Benutzer und dessen jeweilige Aufgabe abgestimmt werden, wodurch eine Steigerung der Effizienz bei der Benutzung erreicht werden kann. Zur Entwicklung multimodaler Bediensysteme für unterschiedliche Automatisierungssysteme, wurde ein dreistufiges Verfahren entwickelt, welches die direkte Verknüpfung zwischen Informationspräsentation und Funktionalität auflöst. Bei der Realisierung derartiger Bediensysteme spielen moderne Displaysysteme wie Head-Mounted-Displays in Verbindung mit drahtlosen Bediengeräten und Head-Sets zur Sprachsteuerung eine wichtige Rolle. Sie bieten eine große Flexibilität ohne Einschränkung der Handhabung. Die erforderlichen Informationen lassen sich dabei auf unterschiedliche Kommunikationskanäle aufteilen und können dadurch besser vom Benutzer aufgenommen werden. Das FMBS ermöglicht Entwicklern von Automatisierungssystemen eine effiziente Realisierung multimodaler Bediensysteme auf Basis konfigurierbarer Komponenten, ohne detaillierte Kenntnisse über die zugrunde liegenden Technologien der multimodalen Informationspräsentation, wie beispielsweise Spracherkennungs- oder Sprachsyntheverfahren, zu fordern. Es unterstützt darüber hinaus den Einsatz moderner Bediengeräte, die aufgrund der konfigurierbaren Informationspräsentation einfach in multimodale Bediensysteme integriert werden können.

## Literaturverzeichnis

- [Da06] Dahm, M.: Grundlagen der Mensch-Computer-Interaktion. Pearson Studium Verlag, München, 2006.
- [Zü05] Zühlke, D.: Der Intelligente Versager – Das Mensch-Technik-Dilemma. Primus Verlag, Darmstadt, 2005.
- [Kr01] Krauß, L.: Mensch-Maschine-Interaktion im Jahr 2010: eine Vision. In MMI Interaktiv - Online Journal zu Fragen der Mensch-Maschine-Interaktion, Ausgabe Nr. 5, 2001.
- [ALK01] Akyol, S., Libuda, L., Kraiss, K.-F.: Multimodale Benutzung adaptiver Kfz-Bordsysteme. Erschienen in: Jürgensohn, T., Timpe, K.-P. (Hrsg.) Kraftfahrzeugführung, Springer-Verlag, Berlin, 2001; S. 137-154.
- [Ov00] Oviatt, S. L.: Designing the User Interface for multimodal Speech and Pen-based Gesture Applications: State-of-the-Art Systems and Future Directions. Human Computer Interaction, 15(4):263–322, 2000.
- [Ab03] Abel, D. et. al.: Automatisierungstechnik 2010. Thesen der Task-Force „Automatisierungstechnik 2010“ der VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik (GMA), 2003.
- [St05] Stiedl, T.: Multimodale Benutzungsschnittstellen in der Automatisierungstechnik, 36. Jahrestagung des Siemens-Automatisierungskreis (SAK), Berlin, 2005.
- [Di00] Diedrich, C., Bangemann, T., Simon, R.: Das EU-Projekt NOAH. In atp - Automatisierungstechnische Praxis 12 (2000), Oldenbourg Verlag, München, 2000.
- [Ja05] Jazdi, N.: Universelle Fernzugriff-Infrastruktur für eingebettete Systeme. In atp - Automatisierungstechnische Praxis 10 (2005), Oldenbourg Verlag, München, 2005.