

Multimodale Interaktion in der Produktionsstätte der Zukunft

Hermann Fürntratt¹, Ferdinand Fuhrmann²

Intelligente Informationssysteme, DIGITAL, JOANNEUM RESEARCH¹
Intelligente Akustische Lösungen, DIGITAL, JOANNEUM RESEARCH²

Zusammenfassung

In diesem Beitrag stellen wir ein neuartiges multimodales Interaktionskonzept für die Produktionsstätte der Zukunft vor. Ziel ist es den Beschäftigten in seiner Interaktion mit Menschen und Geräten sowohl physisch als auch kognitiv bestmöglich zu unterstützen. Ausgehend von Kontrollräumen, mit ihren besonderen Anforderungen an die Operatoren, erweitern wir unser Interaktionskonzept auf die gesamte Produktionshalle und zeigen anhand eines umgesetzten Prototyps die technische Machbarkeit. Probleme und Herausforderungen bei der Umsetzung, sowie Verbesserungsmöglichkeiten werden im Anschluss diskutiert.

1 Einleitung

Multimodale Interaktion gewinnt im Produktionsumfeld immer mehr an Bedeutung. Sie ermöglicht den wechselseitigen Informationsaustausch zwischen Mensch und Computer über natürliche Arten der menschlichen Kommunikation (Ratzka 2008). Die klassische Mensch-Maschine Schnittstelle – Maus und Tastatur – stößt hingegen immer mehr an ihre Grenzen. Wir sitzen zu lange tippend vor den Bildschirmen und obwohl es Versuche gibt, dem entgegen zu wirken (z.B. durch den Einsatz höhenverstellbarer Möbel), ist vorerst keine Besserung in Sicht. Mediziner warnen inzwischen eindringlich vor den gesundheitlichen Folgen dieser Entwicklung (Buckley et al. 2015).

Zusätzlich nimmt auch die kognitive Belastung der Arbeiter immer mehr zu (Spath et al. 2013). Am Beispiel eines industriellen Leitstandes etwa zeigt sich, dass die Datenmenge, die von den Operatoren überwacht wird und auf die im Störfall in kürzest möglicher Zeit reagiert werden muss, rapide gestiegen ist. Dazu kommt, dass die Benutzerschnittstellen für die Beobachtung und Bearbeitung dieser Daten in den seltensten Fällen aufeinander abgestimmt sind. Da der Datenstrom meist aus unterschiedlichsten Teilsystemen stammt, die aufgrund wirtschaftlicher und technischer Überlegungen meist stufenweise dem laufenden Gesamtsystem hinzugefügt wurden, verliert das Gesamtsystem seine ursprünglich

angedachte Intuitivität und wird über die Maßen komplex. Der Mensch steht nicht mehr im Mittelpunkt, sondern der Wettstreit der Systeme. Diese Tatsache lässt sich auf den Schreibtischen unschwer anhand der Vielzahl von Tastaturen und Mäusen – je ein Paar pro Teilsystem – nachvollziehen.

In der vorliegenden Arbeit versuchen wir daher Antworten auf folgende Fragen zu geben:

1. Wie kann ich den Arbeiter im Produktionsprozess wieder in den Mittelpunkt rücken, ihn körperlich und kognitiv entlasten?
2. Wie kann ich von einem bestehenden zu einem moderneren und intuitiveren Produktionssystem übergehen, ohne dass die Kosten für den Umbau den wirtschaftlichen Rahmen sprengen?

Im Folgenden wird das entwickelte Interaktionskonzept anhand des Beispiels Industrielitstand vorgestellt und mithilfe eines implementierten Prototyps die technischen Möglichkeiten sowie Usability Aspekte beleuchtet (Abschnitt 2). In Abschnitt 3 erweitern wir anschließend dieses Konzept auf die gesamte Produktion. In Abschnitt 4 werden schließlich die wichtigsten Erkenntnisse zusammengefasst und die nächsten Schritte diskutiert.

2 Interaktionskonzept

Ausgangspunkt für unsere Überlegungen ist der Leitstand, in dem die Kommunikation mit anderen Menschen, die Abfrage von Statusinformationen, das Steuern von Prozesszuständen, sowie das schnelle Reagieren in Alarmsituationen die wichtigsten Aufgaben sind.

Das Beispiel des Leitstandes

Das entwickelte Gesamtinteraktionskonzept ist von der Vision geprägt, dass der Mensch mehr Freiraum in der Bedienung der Vielzahl an technischen Geräten im Leitstand erhalten soll. Durch die daraus resultierende Multimodalität und Intuitivität der Interaktion wird er sowohl physisch als auch kognitiv entlastet. In Zukunft agiert der Anwender (Operator) nicht mehr lokal über Maus, Tastatur und Telefon vor dem Bildschirm, sondern mobil und erweitert seine Interaktionsmöglichkeiten durch den Einsatz technischer Systeme, die ihm akustisch und visuell im Raum folgen.

Das hier entwickelte Konzept basiert auf den folgenden Interaktionsmechanismen:

- Gesteninteraktion – berührungsbasierte und berührungslose Sensoren ermöglichen eine intuitive und unmittelbare Steuerung des Systems und schränken den Menschen in seiner Mobilität nicht ein.
- Visuelle Interaktion – die dem Menschen übermittelte visuelle Information wird kontextsensitiv aufgearbeitet und auf das wesentliche reduziert.
- Akustische Interaktion – spezielle Technologien erlauben die Fokussierung des akustischen Signals auf bestimmte Zielpersonen sowie eine räumliche Kodierung der Information.



Abbildung 1: Gesamtübersicht eines möglichen zukünftigen Leitstandes.



Abbildung 2: Demonstrator-Setup mit gerichteten Lautsprechern und Operatorsessel.

Abbildung 1 zeigt die Gesamtübersicht des Leitstandes der Zukunft, in dem die Operatoren den Großteil ihrer Routinearbeit unterstützt von mehreren unterschiedlichen Sensoren erledigen. Alle wesentlichen Systemkomponenten sind in der Abbildung grafisch hervorgehoben. Die Gesteninteraktion folgt dem Grundprinzip: *Zeige mit einer Hand auf das, was du tun willst, und steure mit der anderen Hand, wie du es tun willst*. Dieses Prinzip wird sowohl für Armgesten (Makrogestik), welche über weitreichende Tiefensensoren eine Interaktion mit externen Geräten ermöglichen, als auch für Handgesten (Mikrogestik), die über im Sessel eingebaute Sensoren verarbeitet werden, angewandt. Ziel der Mikrogestik ist, dass der Operator in möglichst entspannter Haltung mittels kleinster Bewegungen der Finger das System steuert. Durch einen Drehsensor (IMU) im Operatorsessel lässt sich die aktuelle Zeigerichtung ermitteln und mit dem gespeicherten geometrischen Modell des Raumes und seiner darin enthaltenen Objekte abgleichen. Um Geräte steuern zu können, die keine visuelle oder akustische Rückmeldung geben können, sind Vibrationselemente als taktile Feedbackgeber in den Operatorsesseln eingebaut. Für das Manipulieren von Prozessvariablen steht dem Operator außerdem ein Multi-Touch Ein-/Ausgabegerät zur Verfügung, das auch als mobile Anzeige und Steuerung verwendbar ist.

Ein Lautsprechersystem, welches stark gerichteten Schall abstrahlen kann, versorgt den Operator mit akustischen Informationen. Um andere sich im Raum befindliche Personen nicht zu stören, werden die Lautsprecher den Ohren des Operators automatisch nachgeführt. Dafür wird das Head-tracking eines Tiefensensors zur Ansteuerung von zwei-motorigen Nachführungseinheiten verwendet. Weiters wird die abgestrahlte akustische Information über virtuelle Schallquellen räumlich positioniert. Dadurch ist es möglich, dem aus den Lautsprechern wiedergegebenen Schall eine räumliche Information zu geben. Daraus ergibt sich der für die Wahrnehmungspsychologie wichtiger Effekt, dass bei örtlicher Übereinstimmung von visueller und akustischer Information die menschliche Aufmerksamkeit ungleich höher ist als ohne solche Übereinstimmung (Schönhammer 2013). Analog dazu ist es dem Operator möglich, über ein Mikrofon-Array, das den Standort des Operators ermittelt und sein Sprachsignal automatisch aus anderen Geräuschquellen herausfiltert, mit anderen Gesprächspartnern zu kommunizieren.

Die Werkzeuge zum Manipulieren von Prozessvariablen werden kontextabhängig, je nach Anwendungsfall und Berechtigungsstufe des Operators, zur Verfügung gestellt. Dadurch

lässt die visuelle Information kompakt und intuitiv darstellen. Alle eingehenden Sensorinformationen durchlaufen außerdem einen intelligenten Vorverarbeitungsschritt. Dabei können kontextabhängig irrelevante Informationen gefiltert, unwichtige Meldungen verzögert oder zusammenhängende Hinweise zusammengefasst werden. Diese Reduktion an Daten führt zu einer weiteren kognitiven Entlastung des Operators.

Implementierung

Für Demonstrations- und Evaluationszwecke wurde das entwickelte Interaktionskonzept für den Anwendungsfall Verkehrsleitstand in einem Prototypen implementiert. Dabei nimmt eine zentrale Serverkomponente alle Sensorströme in Echtzeit über LAN/WLAN entgegen. Die weiteren Steuermodule, etwa für die Maustreiber, die 3D-Raumkonfiguration, die grafische Anzeige von Daten, oder für eine Klanguausgabe sind verteilt installiert und bilden ein lose gekoppeltes Client-Server-Netzwerk. Abbildung 2 zeigt ein Bild des implementierten Arbeitsplatzes und eines sensorbestückten Operatorsessels.

Evaluation

Die Interaktionsmechanismen sowie das gesamte Interaktionssystem wurden in zwei Benutzerstudien mit 24 bzw. 6 Teilnehmern evaluiert. Ziel war die Bewertung der einzelnen Ein-/Ausgabegeräte sowie des Interaktionskonzeptes für den Einsatz in einem Leitstand.

Im ersten Teil wurde über quantitative (Bearbeitungszeit) und qualitative (Interviews) Methoden die Bedienfreundlichkeit der einzelnen Benutzerschnittstellen evaluiert. Dabei erzielten vor allem die mobilen und touch-basierten Sensoren (Tablet und TrackPads) sehr gute Ergebnisse. Grund dafür war einerseits die direkte und intuitive Bedienung dieser Geräte, andererseits gaben fast alle Teilnehmer der Studie an, mit diesen Technologien bereits viel Erfahrung zu haben. Die berührungslosen Sensoren dagegen schnitten in den Bewertungen schlechter ab, obwohl die Teilnehmer durchwegs beeindruckt von der Unmittelbarkeit und gewonnenen Bewegungsfreiheit der Interaktion waren. Die Gründe dafür lagen einerseits in einer gewissen Berührungsangst vor der neuen Technologie, andererseits lies auch die Robustheit in der Gestenerkennung noch zu wünschen übrig.

Der zweite Teil dieser Studie umfasste die Evaluation des gesamten Interaktionssystems für den Einsatz in einem Leitstand. Hierfür wurden ausführliche Tests und Befragungen mit Operatoren einer Verkehrsleitzentrale durchgeführt. Besonders hervorgehoben wurde hier die mobile Variante des Leitstandes über den Tablet Computer, welche den Operatoren Bewegungsfreiheit während ihrer langen Schichtzeiten ermöglicht. Weiters wurde der Sensor-bestückte Operatorsessel sehr positiv bewertet. Mit ihm ist es den Operatoren möglich, verschiedene Komponenten ihres Leitstandes mit einer einzigen multimodalen Benutzerschnittstelle zu steuern. Besonders gut schnitt auch die Integration verschiedener Systeme in ein ganzheitliches Interaktionssystem ab. Dadurch wird mühsames Wechseln zwischen den einzelnen Systemen obsolet, was zur Optimierung von Arbeitsabläufen und zur Steigerung der Ergonomie führt. Schließlich wurde auch das akustische Interface mit seinen nachgeführten Lautsprechern und dem Mikrofonarray als sehr nützlich für den Leitstand eingestuft, da dadurch Mobilität und Ergonomie erhöht werden.

3 Die Produktion der Zukunft

War die konzeptionelle Umsetzung im Leitstand hauptsächlich auf das Steuern von Computeranwendungen und die freie Kommunikation ohne Headset ausgerichtet, lässt sich das vorgestellte Konzept der multimodalen und mobilen Interaktion auf die gesamte Produktion ausdehnen. In einer vernetzten und mit Sensoren verschiedenster Art bestückten Produktionsstätte kann sich der Mensch frei bewegen und mit den Geräten und Menschen in seiner Umgebung direkt interagieren.

Der Arbeiter kann über das Prinzip des Zeigens und Steuerns via Makro- und Mikrogestik eine Vielzahl verschiedenartigster Geräte im gesamten Produktionsumfeld bedienen. Er zeigt beispielsweise mit einer Hand auf die Beleuchtung und reguliert mit der anderen die Helligkeit in seiner Umgebung. In der Produktionshalle zeigt der Monteur auf den Industrieroboter und versetzt ihn mit einer Geste in den Wartungsmodus.

Auch die visuelle Interaktion kann hier angewandt werden. So wird auf dem mobilen Anzeigergerät (z.B. Tablet) immer die Information angezeigt die für den Menschen relevant ist – unnötige Informationen werden ausgefiltert. Befindet er sich direkt vor der Maschine, erhält er Manipulationswerkzeuge zum Steuern oder Warten. Demgegenüber lassen sich Statusinformationen von der gesamten Produktionsstrecke anzeigen, sollte er sich nicht in der Nähe einer Maschine befinden. Die Gesteninteraktion lässt sich mit der visuellen Interaktion in Form von *Augmented Reality* kombinieren: der Stahlarbeiter zeigt auf einen Kessel und bekommt durch eine Touchgeste die Statusinformation über diesen übersichtlich auf seinem Tablet präsentiert.

Schließlich lässt sich auch die akustische Schnittstelle in die Produktion der Zukunft integrieren. Über mehrere nachführbare Lautsprecher kann akustische Information zielgerichtet im gesamten Produktionsbereich an bestimmte Personen abgegeben werden. Weiters können über in der Produktionsstätte installierte Mikrofonarrays Sprachsignale von Menschen aufgefangen werden und so eine Kommunikation ohne tragbare Geräte ermöglicht werden. Da in vielen Produktionsstätten der Lärm von Maschinen diese Art der Kommunikation allerdings verhindert, ist ein Rückgriff auf kabellose Headsets problemlos möglich. So können auch in diesem Fall akustische Warnsignale räumlich kodiert werden sodass Richtungsinformation an den Menschen übermittelt wird. Eine Verbindung zwischen Headset und Gesteninteraktion ist ebenfalls sinnvoll: Der Techniker zeigt während der Störungsbehebung auf das rote Telefon an der Wand und baut mit einer Geste eine Telefonverbindung – ohne das bestehende Telefon zu verwenden – zum Kontrollzentrum auf.

4 Diskussion und Ausblick

Aus der Evaluation der Interaktionsmechanismen zeigt sich, dass Touch-basierte Geräte von den Testkandidaten anstandslos akzeptiert werden, während für die berührungslose Gesteninteraktion über Tiefensensoren für die Akzeptanz bei den Arbeitern zwei Voraussetzungen erfüllt sein müssen:

1. Ausreichende Robustheit der Gestenerkennung
2. Energieeffiziente intuitive Interaktion

Die Hardware-Entwicklung der Sensoren (z.B. höchstgenauer Tiefensensoren) schreitet zügig voran (Soli 2015). Die softwareseitige Algorithmik zur robusten Erkennung von Körperbewegungen hinkt derzeit noch hinterher. Auch werden menschliche Eigenheiten im Zusammenhang mit HCI, wie etwa motorische Schwächen, Tremor, oder die Tatsache, dass wir eigentlich nicht dorthin zeigen, wo wir glauben hinzuzeigen (Mayer 2015), noch nicht ausreichend berücksichtigt.

Für den Einsatz in der Produktionsstätte jedoch, hat das vorgestellte Interaktionskonzept unserer Meinung nach großes Potential. Alle Funktionalitäten lassen sich mit geringem Aufwand in bestehende Infrastrukturen integrieren. Darüber hinaus ergibt sich durch den Aufbau eines Sensornetzwerkes, mit dem Arbeiter als menschliches Eingabemedium im Mittelpunkt die Möglichkeit, obsoleete Geräte sanft aus dem Betrieb herauszunehmen und generationsbedingt neue Arbeitsweisen zu etablieren.

Literaturverzeichnis

Buckley, J. P., Hedge, A., Yates, T., Copeland, R. J., Loosemore, M., Hamer, M., Bradley, G. & Dunstan, D. W. (2015). *The sedentary office: a growing case for change towards better health and productivity. Expert statement commissioned by Public Health England and the Active Working Community Interest Company*. Br J Sports Med doi: 10.1136/bjsports-2015-094618.

Google ATAP: Project Soli (2015). Radar based interaction sensor for gesture recognition <https://plus.google.com/+GoogleATAP/videos> (zuletzt abgerufen am 22. 6. 2015)

Mayer, S., Wolf, K., Schneegass, S. & Henze, N. (2015). Modeling Distant Pointing for Compensating Systematic Displacements. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 4165-4168.

Ratzka, A. (2008). *Patterns for Robust and Flexible Multimodal Interaktion*. EuroPLoP, volume 610 of CEUR Workshop Proceedings, CEUR-WS.org.

Schönhammer, R. (Hrsg.) (2013). *Einführung in die Wahrnehmungspsychologie: Sinne, Körper, Bewegung. 2. Auflage*. Wien: facultas wuv, Universitätsverlag.

Spath, D. (Hrsg.), Ganschar, O., Gerlach, S., Hämmerle, M., Krause, T. & Schlund, S. (2013). *Produktionsarbeit der Zukunft – Industrie 4.0*. Stuttgart: Fraunhofer IAO.

Kontaktinformationen

Hermann Fürntratt

Ferdinand Fuhrmann

JOANNEUM RESEARCH
Intelligente Informationssysteme
Leonhardstrasse 59
8010 Graz

JOANNEUM RESEARCH
Intelligente Akustische Lösungen
Leonhardstrasse 59
8010 Graz

hermann.fuerntratt@joanneum.at

ferdinand.fuhrmann@joanneum.at