

Handgestenerkennung in der Mixed Reality

Anton R. Skrabal, Jeronimo Dzaack

Fachgebiet Mensch-Maschine-Systeme
Technische Universität Berlin
Franklinstraße 28/29
D-10587 Berlin
{ask, jdz}@mms.tu-berlin.de

Abstract: Handgesten sind ein möglicher Ansatz für die berührungslose Interaktion in Mixed Reality (MR) Umgebungen. Durch Schwierigkeiten bei der Navigation reicht diese Methode alleine jedoch nicht aus, um MR-System effizient und *natürlich* zu steuern. Es ist notwendig weitere Modalitäten für die umfassende Interaktion zu ergänzen. In diesem Beitrag wird gezeigt, wie die Handgestenerkennung und die Blickverfolgung genutzt werden können, um die Vorteile beider Methoden für die Interaktion in MR Umgebungen zu vereinen. Es wird deutlich, dass mit der Kombination der zwei Interaktionsmodalitäten die Steuerung von MR Umgebungen intuitiver und natürlicher wird.

1 Einführung

Gesten sind ein wesentlicher Bestandteil der menschlichen Kommunikation, wobei Handgesten insbesondere bei zwischenmenschlicher Kommunikation das gesprochene Wort verstärken und somit den Standpunkt der Gesprächsteilnehmer unterstreichen [WBJ02]. Wenn eine natürliche Umgangsweise mit dem Computer gefördert werden soll, sind Gesten eine Methode, um Wünsche und damit Befehle auf natürliche und menschliche Art und Weise an das System zu übertragen. Handgesten sind definiert, als eine in der Umwelt bedeutsame Körperbewegung, also als eine räumliche Bewegung der ganzen Hand. Dazu zählen auch das Ausstrecken und Einziehen einzelner oder auch aller Finger. Diese Bewegungen sind informationstragend. Dies verdeutlicht, dass statische Zustände nicht zu den Gesten gehören [Tu98]. Zusammenfassend sind Handgesten Bewegungen der Hand an sich, als auch von Fingern aller Art, nicht aber der ruhende Zustand einer Hand.

Mixed Reality (MR) oder deutsch *vermischte Realität* ist nahe verwandt mit der virtuellen Realität (VR), deckt aber gleichzeitig ein Kontinuum zwischen *augmented reality* (AR), also einer durch elektronischen Daten erweiterten Realität und *augmented virtuality* (AV), also einer durch Daten der realen Welt erweiterten virtuellen Umgebung ab [TYK01].

In Mixed Reality Umgebungen werden dementsprechend Daten aus der echten Welt und Daten aus der virtuellen Welt miteinander verknüpft. Dies geschieht zum Beispiel durch Projektion von virtuellen Daten auf reale Oberflächen oder auf halbdurchlässige Displays, die zum Beispiel wie Brillen aufgesetzt werden können. Diese Art der Darstellung eignet sich besonders gut für hochinteraktive Systeme, die auf Benutzereingaben reagieren und verstehen sollen, was die Absichten des Benutzers zu jedem Zeitpunkt der Interaktion sein könnte. Aus diesem Grund ist eine Kombination von Mixed Reality und Gestenerkennung – als Methode der direkten Manipulation von Objekten – überaus geeignet für ein solches System.

2 Methodischer Ansatz

2.1 Gestenerkennung

Die Hauptidee des am Fachgebiet für Mensch Maschine Systeme (FG MMS) derzeit in Entwicklung befindliche Systems ist es, ein *one-for-all* Gerät zu entwickeln, bei dem weder eine umfassende Kalibrierung notwendig ist, noch weitere Hilfsmittel wie Elektroden oder ähnliches verwendet werden. Das soll den Einsatz von gewissen Grundgesten für jedermann zu jeder Zeit ermöglichen. Die Gestenerkennung selbst erfolgt dabei mithilfe einer einfachen Kamera (Auflösung 320x240 px), die den Arbeitsbereich erfasst. Die so erhaltenen Bilder werden mittels Methoden aus der Bildverarbeitung vorverarbeitet, wobei die Hand aufgrund der chromatischen Eigenschaften der Hautfarbe mithilfe einer Umwandlung der Bilder aus dem RGB Farbraum in den LUV Farbraum und einer anschließenden Analyse der Chrominanz, bestehend aus den Farbkomponenten U und V (unter Nichtberücksichtigung der Luminanz L), extrahiert wird [Go00]. Anschließend wird eine morphologische Operation *Open* angewandt, um Artefakte zu bereinigen. Die auf diese Weise extrahierte Hand wird bei Bewegungen getrackt (d.h. erfasst). Die dadurch entstehenden Bewegungssequenzen werden nach einer Datenkompression durch ein neuronales Netzwerk klassifiziert. Erkannte Befehle werden an das Betriebssystem weitergeleitet und von diesem ausgeführt.

2.2 Anwendung in der Mixed Reality

Die Gestenerkennung ist eine gute Methode zur Manipulation (also Befehlsabsetzung), die Navigation allerdings gestaltet sich ausgesprochen schwierig, da Absolutpositionen im Vergleich zur Kamera verwendet werden müssen. Dafür wäre eine aufwändige und zeitintensive Kalibrierung des Systems notwendig. Auch hat die Handgröße in diesem Verfahren einen Einfluss auf die Absolutpositionen, was der Idee des *one-for-all* Systemgedanken widersprechen würde. Umgekehrt verhält es sich mit Blickverfolgungssystemen und den damit erfassten Blickdaten, die für die Navigation gut einzusetzen sind, für die Manipulation aufgrund der binären Information der Fixationen (z.B. Auswahl – lange Fixation, keine Auswahl – keine/kurze Fixation) aber eher ungeeignet.

Daher wäre es von großem Nutzen, die oben beschriebene Gestenerkennung mit vorhandenen Blickverfolgungssystemen zu kombinieren, um so eine Lösung sowohl für die Navigation, als auch für die Manipulation in MR Umgebungen bereitzustellen. Diese Kombination würde die Bedienung von MR Umgebungen deutlich erleichtern, da keine herkömmlichen Eingabemodalitäten wie Maus und Tastatur mehr benötigt würden.

3 Fazit

Ausgehend von dem Ansatz der Mixed Reality ist konzeptionell ein System entwickelt worden, mit dem eine MR Umgebung mittels Gestenerkennung intuitiv gesteuert werden kann. Aufgrund der Schwierigkeiten bei der Navigation durch Handgesten muss eine weitere Eingabemodalität hinzugefügt werden, um den Bereich der Navigation vollständig abdecken zu können. Dementsprechend wird als nächster Arbeitsschritt am FG MMS ein bereits bestehendes Blickverfolgungssystem für die Blickinteraktion [DTN+09] mit dem hier beschriebenen System zur Handgestenerkennung und -steuerung verknüpft. Es wird davon ausgegangen, dass diese Kombination die bisher unzureichende Navigation und somit die Nutzung von MR Umgebungen vereinfacht.

Literaturverzeichnis

- [DTN+09] Dzaack, J., Trösterer, S., Nicolai, T., Rötting, M. (2009). iCOMMIC: Multimodal Interaction in Computing Systems. *Proceedings of the 17th World Congress of the International Ergonomics Association*, Peking.
- [Go00] Gomez, G. (2000). On selecting colour components for skin detection. In: *Proceedings of the ICPR*, 2, S. 961–964.
- [Tu98] Turk, M. (1998). Gesture recognition. In: *Handbook of Virtual Environments: Design, Implementation, and Applications (Human Factors and Ergonomics)*, Microsoft Research, S. 223-239.
- [TYK01] Tamura, H., Yamamoto, H., Katayama, A. (2001). Mixed reality: future dreams seen at the border between real and virtual worlds, Mixed Reality Systems-Laboratory, Canon Inc. In: *Computer Graphics and Applications, IEEE*, 21(6); S. 64-70.
- [WBJ02] Watzlawick, P., Beavin, J.H., Jackson, D.D. (2002). *Menschliche Kommunikation: Formen, Störungen, Paradoxien*, Verlag Hans Huber, Bern.