

Gestenbasierte Google-Earth-Bedienung: Implikationen für ein natürliches Gesten-Set am Beispiel einer 3D-Topographieanwendung

Thomas SEELING¹, Ellen FRICKE², Ulrike LYNN², Daniel SCHÖLLER²,
Angelika C. BULLINGER¹

¹ *Professur Arbeitswissenschaft und Innovationsmanagement,
Technische Universität Chemnitz
Erfenschlager Str. 73, D-09125 Chemnitz*

² *Professur Germanistische Sprachwissenschaft,
Semiotik und Multimodale Kommunikation
Technische Universität Chemnitz
Thüringer Weg 11, D-09107 Chemnitz*

Kurzfassung: Am Beispiel der 3D-Topographieanwendung Google Earth, die mittels des gestenbasierten Eingabegerätes Leap Motion steuerbar ist, wird untersucht, welche Gesten als Bedienstrategien durch Nutzer typischerweise intuitiv hervorgebracht werden. Ziel ist es, Erkenntnisse für ein Gesten-Repertoire zu gewinnen, mit dessen Hilfe die gestenbasierte Freihandbedienung in 3D-Umgebungen „natürlicher“, das heißt konform zu vorhandenen allgemeinen Designrichtlinien wie z.B. Komfort, leichte Erlernbarkeit, gute Abbildung von Gesten auf Funktionen, hohe Freiheitsgrade gestaltet werden kann.

Schlüsselwörter: Gestensteuerung, Usability-Test, Gesten-Interface, Natural User Interfaces, Freihandgesten, redebegleitende Gesten

1. Einleitung

Interaktive Systeme, die durch innovative gestenbasierte Eingaben zu bedienen sind, sind oftmals nach der technischen Machbarkeit konzipiert und umgesetzt. Fähigkeiten, Fertigkeiten und Bedarfe des Nutzers werden nicht strukturiert berücksichtigt (Morency et al. 2006). Die implementierten Gesten sind selten intuitiv und deshalb durch den Nutzer schwer zu erlernen, weshalb die Benutzung als ineffizient, ineffektiv und nicht intuitiv wahrgenommen wird. Nur wenn die Belange der Nutzer von Post-WIMP-Technologien von Beginn an in den Entwicklungsprozess einbezogen werden, kann ein intuitiver Gebrauch dieser als Natural User Interfaces bezeichneten Schnittstellen tatsächlich gewährleistet werden (Liu 2010).

Am Beispiel der 3D-Topographieanwendung Google Earth, die mittels des gestenbasierten Eingabegerätes Leap Motion steuerbar ist, wird untersucht, welche Gesten als Bedienstrategien durch Nutzer typischerweise intuitiv hervorgebracht werden. Dabei steht die Leitfrage im Fokus, wie die vorbestimmten Interaktionsmöglichkeiten der Anwendung („patterns“) hinsichtlich eines intuitiven Nutzungsverhaltens von den Anwendern eingeschätzt werden. Ziel ist es, Konsequenzen für ein Gestenrepertoire abzuleiten, mit dessen Hilfe die gestenbasierte Freihandbedienung in 3D-Umgebungen „natürlicher“, das heißt konform zu vorhandenen allgemeinen Designrichtlinien für Gestensteuerung wie Komfort, leichter Erlernbarkeit und Vermittlung, guter Abbildung der Gesten auf

Funktionen, hohe Freiheitsgrade, ausreichender Affordancecharakter und geringer Ermüdung gestaltet werden kann (Microsoft 2013; Baudel et al. 1993).

2. Methodik

Zur Bewertung der implementierten Patterns der 3D-Topographieanwendung durchlaufen Probanden in einem Usability-Test typische Aufgaben wie etwa das Auffinden verschiedener Points-Of-Interests (POI) durch gestengesteuerte Navigation des digitalen 3D-Globus. Zuvor besteht die Aufgabe der Probanden jedoch darin, zunächst zu zwei Orten auf einem analogen Globus zu navigieren, um in der Auswertung einen Vergleich zwischen den in „realer“ und digitaler Welt hervorgebrachten Gesten durchführen zu können.

Daran anschließend haben die Probanden die Möglichkeit, sich mit der Funktionalität des Eingabecontrollers am Anwendungsfall der 3D-Topographieanwendung Google Earth vertraut zu machen. Wird dieser Prozess vom Probanden als abgeschlossen bewertet, startet die eigentliche Aufgabe: das Auffinden dreier POI und das virtuelle Fotografieren der virtuellen Sehenswürdigkeiten mittels Gestensteuerung. Dabei wird der Proband im Rahmen der Methode des Think-Aloud gebeten, seine Gedanken hinsichtlich des Bedienkonzepts zu verbalisieren (Häder 2006).

Der Test wird mittels HD-Video- und Audiotechnik synchron erfasst, um auszuwerten, welche Gesten durch Nutzer, die mit der Technologie nicht oder nur wenig vertraut sind, für die grundlegenden Patterns Zooming, Panning, Rotation, Homing usw. typischerweise dargeboten werden. Insgesamt sind vier Feld-Kameras auf den Gesten-/Eingaberaum des Probanden gerichtet. Ein Mikrofon sowie ein Screencapturing-Programm zeichnen das Nutzungsverhalten des Probanden zusätzlich auf (Abbildung 1). Nach Absolvierung der Aufgaben wird der Proband mittels eines leitfadengestützten und problemzentrierten Interviews gebeten die Bedienung hinsichtlich der empfundenen Intuition einzuschätzen. Zusätzlich werden mit dem QUESI und INTUI Fragebogeninstrumente mit standardisierten Fragen zur intuitiven Benutzung eingesetzt, um zu prüfen, inwieweit die unbewusste Anwendung von Vorwissen des Benutzers, zu einer effektiven Interaktion mit dem System führt und die Vorteile der als intuitiv empfundenen Benutzung, wie eine reduzierte kognitive Beanspruchung, geringere Lernaufwände, der Wahrnehmung von Vertrautheit im Umgang sowie der Wahrnehmung, Arbeitsaufgaben fehlerfrei erreichen zu können, mit dem in der Anwendung implementierten Gestenrepertoire eingelöst werden können (Prümper et al. 2010, Ullrich & Diefenbach, 2010).



Abbildung 1: Multiperspektivische Erfassung der zur Navigation dargebotenen Gesten des Probanden

Daran anschließend wird der Proband, der den Testdurchlauf absolviert hat, instruiert, die Bedienung von Google Earth einem weiteren Probanden so genau zu beschreiben, so dass auch dieser in die Lage versetzt wird, ein bestimmtes topographisches Ziel mittels der Gestensteuerung in Google Earth aufzufinden (Fricke 2002, 2007). Etwaige Probleme bei der Bedienung werden dabei zusätzlich retrospektiv erfasst: Sprachliche Instruktionen sowie die dazu verwendeten redebegleitenden Gesten werden nach Parametern der Gestenforschung (für einen Überblick zum internationalen Forschungsstand siehe Müller, Cienki, Fricke et al. 2013 und 2014) segmentiert und annotiert und hinsichtlich der von Saffer (2008) und Karam et al. (2005) vorgeschlagenen Attribute und Freiheitsgrade, die es zum Entwerfen eines „natürlichen“ Gesten-Vokabulars für Mensch-Maschine-Schnittstellen zu berücksichtigen gilt klassifiziert.



Abbildung 2: Der Proband erklärt seinen Umgang mit dem gestenbasierten Eingabegerät unter Rückgriff auf redebegleitende Gesten

Im Testdurchlauf benutzt der Proband spontan das Konzept eines Flugzeugs, das um die Erde fliegt, um mit dem virtuellen Globus aus der Vogelperspektive zu interagieren. Dabei repräsentiert die rechte flache Hand gesamthaft das Flugzeug (Abbildung 2). Dass beim Probanden das Konzept eines Flugzeugs während der Bedienung aktiviert wird, zeigt sich insbesondere in dem sich anschließenden Gespräch, in dem der Proband seinen Umgang mit dem gestenbasierten Eingabegerät beschreibt. Begleitend zur verbalen Äußerung „also dann fliegst du wie über die die Oberfläche kannst das auch son bisschen schräg machen (...)“ repräsentiert wiederum die rechte Hand ein Flugzeug. Die flache Hand als Flugzeug ist jedoch eine Geste, die in der Alltagskommunikation über verschiedene Sprecher und Kontexte hinweg konventionalisiert ist (Fricke 2015).

3. Diskussion

Aus dem explorativen Testdurchlauf zur Google-Earth-Bedienung leiten sich die folgenden Annahmen ab (Fricke 2015): 1. Die Bedienung eines virtuellen Globus

kann durch den Rückgriff auf das metaphorische Konzept des Fliegens eines Flugzeugs verbessert und „intuitiver“ gestaltet werden. 2. Konventionalisierte Gesten der Alltagskommunikation sind Indikatoren für solche Konzepte. Eine Berücksichtigung von Gesten in Usability-Studien erlaubt eine Erweiterung des bisherigen Methodenspektrums. 3. Menschliche Gesten der Alltagskommunikation sind ein vielversprechender Ausgangspunkt, um „künstliche Gesten“ für die Gestensteuerung von Robotern und Mensch-Maschine-Schnittstellen zu gestalten. Eine Zusammenarbeit zwischen Arbeitswissenschaft und linguistisch-semiotischer Gestenforschung lässt insgesamt eine Erweiterung des arbeitswissenschaftlichen Methodenspektrums sowie innovative Impulse für die Gestaltung „natürlicher“ Mensch-Maschine-Schnittstellen erwarten.

4. Literatur

- Baudel T, Beaudouin-Lafon M (1993) Charade: Remote Control of Objects Using Free-hand Gestures. *CACM*, 36(7):2 8–35.
- Fricke E (2002) Origo, pointing, and speech: The impact of co-speech gestures on linguistic deixis theory. In: *Gesture*, 2(2), 207–226.
- Fricke E (2007) Origo, Geste und Raum – Lokaldeixis im Deutschen. Berlin und New York: De Gruyter.
- Fricke E (2015) Hands and Objects in Language, Culture, and Technology: Manual Actions at Workplaces between Gesture, Robotics, and Product Design (MANUACT). Projektpräsentation auf der BMBF-Statustagung „Die Sprache der Objekte“, Bonn, 26. Oktober 2015.
- Häder M (2006) Empirische Sozialforschung. VS Verlag für Sozialforschung.
- Hurtienne J, Naumann A (2010) QUESI – A questionnaire for measuring the subjective consequences of intuitive use. In R. Porzel, N. Sebanz & M. Spitzer (Eds.), *Interdisciplinary College 2010*.
- Karam M, Schraefel M (2005). A taxonomy of gestures in human computer interactions. Univ. of Southampton.
- Liu W (2010) Natural user interface-next mainstream product user interface, In: *IEEE 11th International Conference on Computer-Aided Industrial Design & Conceptual Design (CAIDCD)*, Vol. 1, IEEE, 2010, pp. 203–205.
- Microsoft (2013) KINECT for Windows. Human Interface Guidelines. Version 1.8.
- Müller C, Cienki A, Fricke E, Ladewig S, McNeill D, Bressen J (Hrsg.) (2014). *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction (Handbooks of Linguistics and Communication Science 38.2)*. Berlin und Boston: De Gruyter Mouton.
- Müller C, Cienki A, Fricke E, Ladewig S, McNeill D, Teßendorf, S (Hrsg.) (2013). *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction (Handbooks of Linguistics and Communication Science 38.1)*. Berlin und Boston: De Gruyter Mouton.
- Morency L, Darrell T (2006) Head Gesture Recognition in Intelligent Interfaces. *The Role of Context in Improving Recognition*.
- Saffer D (2008). *Designing Gestural Interfaces*. Beijing. Köln. Taipei. Cambridge. Tokyo: O’Reilly.
- Ullrich D, Diefenbach S (2010). INTUI. Exploring the Facets of Intuitive Interaction. In Ziegler J & Schmidt A (Eds.) *Mensch & Computer 2010* (pp. 251-260). München: Oldenbourg.
- Quek F, McNeill D, Bryll R. et al. (2002). Multimodal Human Discourse: Gesture and Speech. *ACM Trans. on CHI*, 9(3):171–193.

Danksagung: Gefördert vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen des Forschungsprojekts “Hands and Objects in Language, Culture, and Technology: Manual Actions at Workplaces between Robotics, Gesture and Product Design“ (MANUACT). Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.