

Sprach-Gestik Experimente mit IADE, dem Interactive Augmented Data Explorer

Thies Pfeiffer*, Alfred Kranstedt*, Andy Lücking†

*AG Wissensbasierte Systeme & VR, Technische Fakultät

†Fakultät für Linguistik und Literaturwissenschaft

Universität Bielefeld

33501 Bielefeld

Tel.: +49 (0)521 106 2919 Tel.: +49 (0)521 106 3680

E-Mail: {tpfeiffe,akranste}@ {andy.luecking}@

techfak.uni-bielefeld.de Uni-Bielefeld.de

Zusammenfassung: Für die empirische Erforschung natürlicher menschlicher Kommunikation sind wir auf die Akquise und Auswertung umfangreicher Daten angewiesen. Die Modalitäten, über die sich Menschen ausdrücken können, sind sehr unterschiedlich - und genauso verschieden sind die Repräsentationen, mit denen sie für die Empirie verfügbar gemacht werden können. Für eine Untersuchung des Zeigeverhaltens bei der Referenzierung von Objekten haben wir mit IADE ein Framework für die *Aufzeichnung, Analyse und Resimulation* von Sprach-Gestik Daten entwickelt. Mit dessen Hilfe können wir für unsere Forschung entscheidende Fortschritte in der linguistischen Experimentalmethodik machen.

Stichworte: Empirie, Experimentalmethodik, Tracking, Visualisierung, Anwendung

1 Einleitung

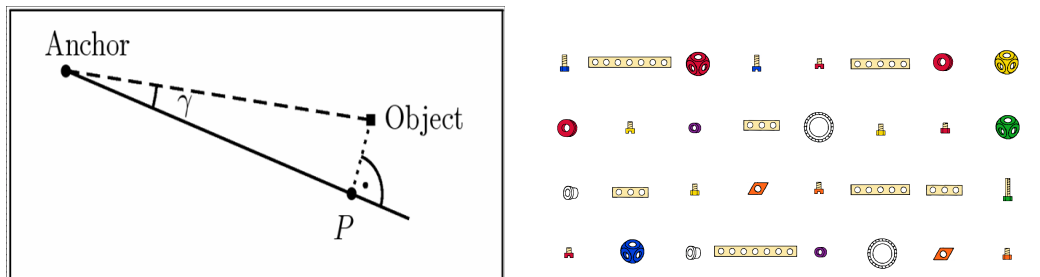
In einer Kooperation von Linguistik und Informatik untersuchen wir die multimodale Interaktion zwischen Menschen bzw. zwischen Menschen und Maschinen. Dabei kombinieren wir linguistische Grundlagenforschung und Theoriebildung mit der Entwicklung und Exploration intelligenter Mensch-Maschine Schnittstellen z.B. für die Anwendung in der Virtuellen Realität. Zur Zeit konzentrieren wir uns primär auf die beiden Modalitäten Sprache und Gestik. In der Zukunft ist jedoch eine Ausweitung auf weitere Modalitäten geplant, wie z.B. Blicke (Eyetracking) oder Mimik (Facial Feature Tracking), damit z.B. Turn-Taking Signale, sowie der aktuelle Fokus der visuellen Aufmerksamkeit erfasst und ausgewertet werden können. In einem neuen Projekt sollen auch sozial-emotionale Aspekte berücksichtigt und z.B. die Herzfrequenz und die Hautleitfähigkeit als basale Indikatoren für den emotionalen Zustand in die Auswertung mit einbezogen werden.

1.1 Herausforderungen für die Empirie

Die Komplexität menschlichen Verhaltens und die Vielfalt der ihm zur Verfügung stehenden Modalitäten stellt den empirischen Forscher vor einige Schwierigkeiten:

- Menschliche Kommunikation benutzt viele Kanäle, die sehr unterschiedlich in ihren zeitlichen und räumlichen Ausprägungen sind: Sprache, Gestik, Mimik, Blicke oder z.B. Körperhaltung.
- Menschliche Kommunikation ist stark kontextabhängig. Dies bedeutet, dass durch die Gestaltung eines Experimentes einen starken Einfluß auf die Interaktion genommen wird. Insbesondere ist dies dann kritisch, wenn unterschiedlich stark eingreifende Meßtechniken verwendet werden
- Erfasste Daten unterschiedlicher Kanäle liegen in der Regel als sehr heterogener Datenstrom vor, der sich nur mit großem Aufwand einer Analyse erschließen lässt.
- Technisch gemessene Daten (z.B. Video- oder Sprachaufzeichnungen) liefern basale Daten auf niedriger Abstraktionsebene. Höhere Bedeutungsebenen erschließen sich erst über eine Interpretation durch Fachkräfte (*Annotation*). Dieser Prozess ist sehr aufwändig und fehleranfällig.
- In Studien natürlicher Kommunikationssituationen sind die erhobenen Daten über verschiedene Versuchspersonen hinweg sehr unterschiedlich und erst nach gezielter Abstraktion und Reduktion einer sinnvollen statistischen Auswertung zugänglich.
- Für die Bearbeitung der Daten werden unterschiedliche Spezialprogramme eingesetzt. Die Wissenschaftler müssen sich daher mit Problemen der Datenkonsistenz, Datentransformation und Datenversionierung auseinandersetzen, die nicht im eigentlichen Fokus ihres Interesses liegen.
- Studien entwickeln sich iterativ, da die Sichtung des Datenmaterials neue Hypothesen und Interpretationen aufwirft, die zu einer Verfeinerung des Annotationsschemas führen können.

Im Folgenden möchten wir an einem Beispiel zeigen, wie durch den Einsatz von Techniken aus der interaktiven Virtuellen Realität die Experimentalmethodik für natürliche Kommunikationssituationen entscheidend vorangetrieben werden konnte. Dazu werden wir zuerst kurz eine Serie von Studien und die bislang verwendete Methodik vorstellen. Danach beschreiben wir, wie das ursprüngliche Setting der Studie für den Einsatz von VR Technologie vorbereitet wurde und stellen im Anschluß den *Interactive Augmented Data Explorer* (IADE) vor.



(a) Die Verankerung und der Öffnungswinkel sind die entscheidenden Parameter des Zeigekegels.

(b) Objektkonfigurationen wie die abgebildete bilden die Domäne unserer Dialogspiele.

Abbildung 1: Um die Parameter für das Modell eines Zeigekegels (a) bestimmen zu können, wurden Dialogspiele durchgeführt, in denen Objekte auf einem Tisch (b) sprachlich und gestisch vermittelt werden mussten.

2 Studien über Deixis

Um genauere Informationen zur Semantik und Pragmatik von Zeigegesten zu erhalten, haben wir Experimente und Studien in Form von Objektidentifikations-Spielen in multimodalen Dialogen durchgeführt (siehe [KLP⁺06b]). Die dabei gewonnenen empirischen Ergebnisse bilden nun die Grundlage für die Modellierung einer Schnittstelle zwischen Sprache und Gestik, für die glaubwürdige Produktion von Zeigegesten bei anthropomorphen Agenten [KW05] und für die Optimierung eines Systems zur Interpretation von Objektreferenzen [PL04]. Den theoretischen Hintergrund und die Einbettung in linguistische Theorie und informatische Arbeiten beschreiben wir ausführlich in [KLP⁺06a]. Über einen Teil unserer Ergebnisse wird in [KLP⁺06c] berichtet. An dieser Stelle wollen wir kurz die Fragestellung und den Aufbau der Studien skizzieren, soweit diese Informationen zum Verständnis des Methodenteils notwendig erscheinen.

2.1 Die Fragestellung

Eine konkrete Fragestellung in zwei unserer Studien ist die Frage nach der Morphologie des Zeigekegels (*Pointing Cone*), einer Konzeption des Referenzbereiches einer Zeigegeste. Der Zeigekegel ist wie ein Zeigestrahl verankert, z.B. im Zeigefinger, und gibt durch seinen Öffnungswinkel die Unschärfe des Zeigens an (Abb. 1). Wäre Zeigen objektgenau, dann könnte ein Öffnungswinkel von Null angenommen werden und der Zeigekegel würde mit einem Zeigestrahl zusammenfallen. In den Studien wurden Daten erhoben, die eine Modellierung des Zeigekegels und eine Prüfung des Modells erlauben.



Abbildung 2: Links der Demonstrierende, rechts der Identifizierende, in der Mitte der Tisch mit den Objekten. Das Bild zeigt die beiden Perspektiven der Videokameras.

2.2 Der Aufbau des Experimentes

Das Grundszenario unserer Studien ist in Abb. 2 zu sehen: eine Person übernimmt die Rolle des Demonstrierenden (hier sitzend), die andere die des Identifizierenden. Der Demonstrierende hat die Aufgabe, dem Identifizierenden ein bestimmtes Objekt über Sprache und/oder Gestik, je nach Szenario, zu vermitteln.

3 Bisherige Methodik

In den ersten Studien wurden die Experimente mit zwei Kameras über ein Mischpult aufgezeichnet, so dass pro Experiment ein Video mit zwei Teilbildern aus den beiden unterschiedlichen Kameraperspektiven gewonnen wurde (Abb. 2). Die wesentlichen Faktoren sind für eine statistische Auswertung damit allerdings noch nicht zugänglich. Sie müssen erst durch eine aufwändige manuelle Transkription und Annotation (Abb. 3) erschlossen werden. Die Transkription des Materials ist zwar relativ einfach und schnell zu erledigen, da sich die Äußerungen auf kurze Phrasen des Demonstrierenden beschränken. Bei der Annotation gibt es jedoch einige Probleme:

1. *Hoher zeitlicher Aufwand*, da Gesten sehr schnell ablaufen und Start und Ende des Gestenstrokes, dem für uns interessanten Teil der Geste, nur schwer zu bestimmen sind. Oft muß das Material Bild für Bild angeschaut und beurteilt werden.

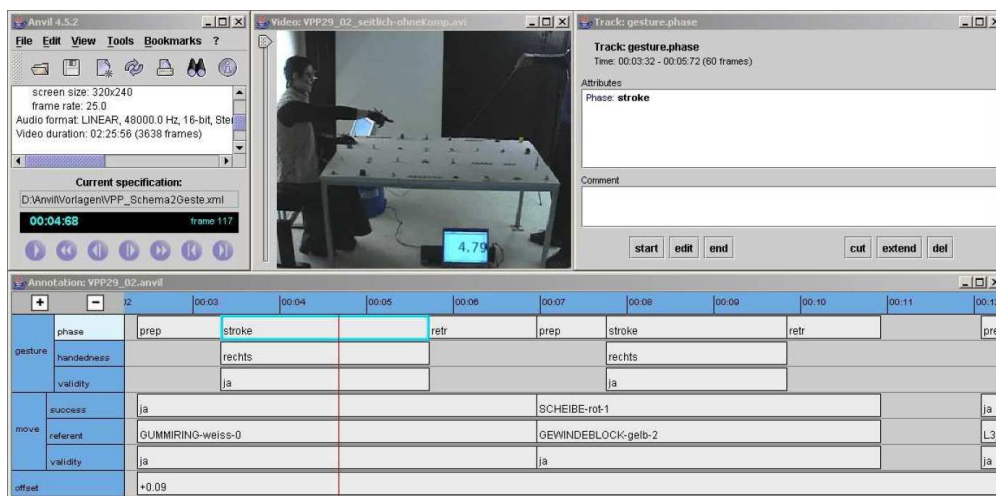


Abbildung 3: Die aufgezeichneten Videos werden mit dem Tool *Anvil* manuell transkribiert und markante Stellen, wie z.B. Gestenanfang und -ende annotiert.

2. *Iteratives Vorgehensmodell sehr zeitaufwändig*. Bei dieser Art von Studien über neuem Material wird durch die Diskussion der gesichteten Daten das Annotationsschema oft noch während der Annotation verfeinert. Damit müssen aber auch die bereits annotierten Videos überarbeitet werden.
3. *Uneinstimmigkeiten* zwischen den Annotatoren müssen kontrolliert werden, dazu werden ein *Annotationsschema* erstellt und Tests auf *Inter-Rater-Agreement* durchgeführt.
4. *Verdeckungen* der Arme und Hände des Demonstrierenden können Teile der Daten unbrauchbar machen. Diese könnten durch den Einsatz von weiteren Kameraperspektiven zwar reduziert werden, dadurch würde aber der Bearbeitungsaufwand weiter ansteigen.
5. *Ungenauere Positions- und Orientierungsbestimmung* bei der manuellen Ausmessung an den perspektivisch verzerrten Bildern der Videoaufnahme. Speziell für die Bestimmung der Parameter des Zeigekegels sind diese Werte jedoch besonders wichtig.

Der in den Punkten (1) und (2) beschriebene Zeitaufwand lässt sich nur bedingt, wie in (3) beschrieben, durch den Einsatz weiterer Annotatoren reduzieren. Eine kurze Übersicht zur Qualitätssicherung bei Annotationen gibt [SL05]. Eine entsprechende Evaluation unseres Annotationsschemas beschreibt [LS05].

Die bislang eingesetzte Methode ist also wegen (4) und (5) in dieser Form für unsere Zwecke nicht mehr geeignet. Wir brauchen eine Lösung, die uns bei einer hohen räumlichen Genauigkeit mehrere Perspektiven auf den Ablauf bietet - und das ohne einen weiteren Anstieg des Zeitaufwandes für die Annotation.

4 VR Technologie für die Grundlagenforschung

Für die Studie zur Parametrisierung des Zeigekegels haben wir eine neue Methode entwickelt, die aktuelle VR Technik einsetzt, um die beschriebenen Probleme (4) und (5) zu lösen. Dabei koppeln wir die Messung der Position und Orientierung der Hände des Demonstrierenden von der entsprechenden Annotation der Videos durch den Einsatz eines Trackingsystem. Mit diesem werden relevante Daten der Postur des Demonstrierenden während des Experimentdurchlaufs aufgezeichnet.

Für Analyse und Auswertung haben wir ein Simulationstool, den *Interactive Augmented Data Explorer* (IADE), entwickelt, mit dem die gemessenen Daten der unterschiedlichen Modalitäten und Aufzeichnungsverfahren zusammengeführt und in integrierter Form interaktiv dargestellt werden können. Dabei können in einem iterativen Verfahren auch neue Daten, die durch Annotation oder aus Berechnungen gewonnen wurden, hinzugefügt und im Ensemble mit den Originaldaten dargestellt und evaluiert werden.

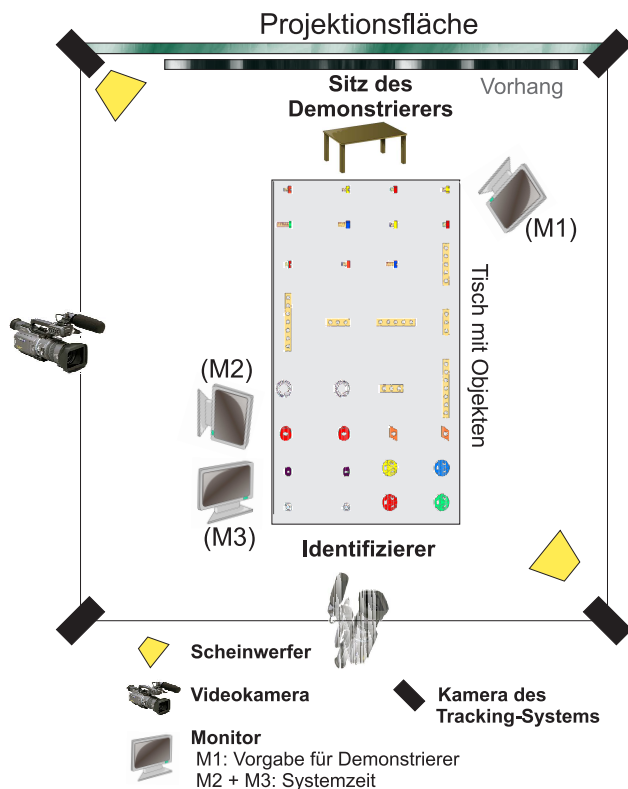
4.1 Aufbau des Experimentalsettings

Der grundsätzliche Aufbau des Settings konnte erhalten bleiben (Anordnung von Personen, Tisch und Objekten) und leicht in die existierende VR Installation eingepasst werden. Das für die Interaktion montierte Trackingsystem konnte ohne Umbau für die Experimente verwendet werden (Abb. 4). Durch diese Einpassung ergibt sich aus Sicht des Experimentators ein faszinierender Vorteil: das reale Setting kann exakt an der gleichen Stelle in der Virtuellen Realität repliziert werden.

4.2 Hardware

Zu den zwei Videokameras kommen im erweiterten Setting neun Kameras eines optischen Trackingsystems der *advanced realtime tracking GmbH*. Die schwarzen Kästen in Abb. 4 zeigen die ungefähre Positionierung. Pro Position sind jeweils zwei Kameras befestigt, eine oben und eine unten. Die neunte Kamera schaut über den Kopf des Demonstrierenden.

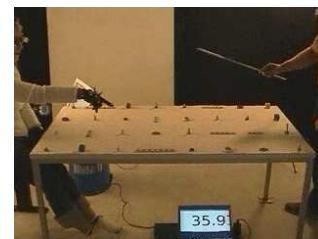
In einem Vorexperiment haben wir für Kopf und Arme das optische Trackingsystem eingesetzt. Die Hände wurden jedoch über spezielle Datenhandschuhe getrackt, da die Bewegungen der Finger nur schwer optisch aufzulösen sind. Wir haben dabei jedoch beobachtet, dass die kabelgebundenen, in den Fingern relativ starren Handschuhe das natürliche Zeigeverhalten der Teilnehmer behinderten. Daher ersetzten wir die Datenhandschuhe später durch die in Abb. 4 d) gezeigten Spezialanfertigungen, die ebenfalls das optische System nutzen. Da der Zeigefinger in der kritischen Phase des Gestenstrokes in der Regel exprimiert ist, kann er leicht von den Kameras erfasst werden. Der neue Handschuh basiert auf angepassten Golfhandschuhen, ist leichter und in den Fingern flexibler als die komplexeren Datenhandschuhe. Zudem ist der Demonstrierende durch



(a) Schematischer Aufbau des Experimentalsettings.



(b) Oberes Kamerabild



(c) Seitliches Kamerabild



(d) Handschuhe mit optischen Markern

Abbildung 4: Der Aufbau des Experimentalsettings unter der neuen Methode. Die Anordnung wurde in die Gegebenheiten der VR Installation mit montierten optischen Trackern eingepasst.

die fehlenden Kabel weniger eingeschränkt in der Bewegungsfreiheit. Starre Handhaltungen, wie im Vorexperiment beobachtet, traten damit in späteren Studien nicht mehr auf.

Neben den Videoaufzeichnungen und dem Tracking wurden über ein Funkmikrofon auch die sprachlichen Äußerungen des Demonstrierenden und über ein Szenenmikrofon die Akustik des Gesamtsettings aufgezeichnet. Die Audioaufzeichnungen wurden dabei direkt in die beiden Videos eingespielt (das Funkmikrofon in die seitliche, das Szenenmikrofon in die obere Szenenkamera), so dass keine nachträgliche Synchronisation zwischen Audio und Video notwendig wurde. Die Videodaten wurden mit der Systemzeit des Trackingsystems synchronisiert. Dazu wurde diese, wie in Abb. 4 b) bzw. c) ersichtlich, auf zwei Monitoren angezeigt, die jeweils von einer Kamera erfasst wurden.

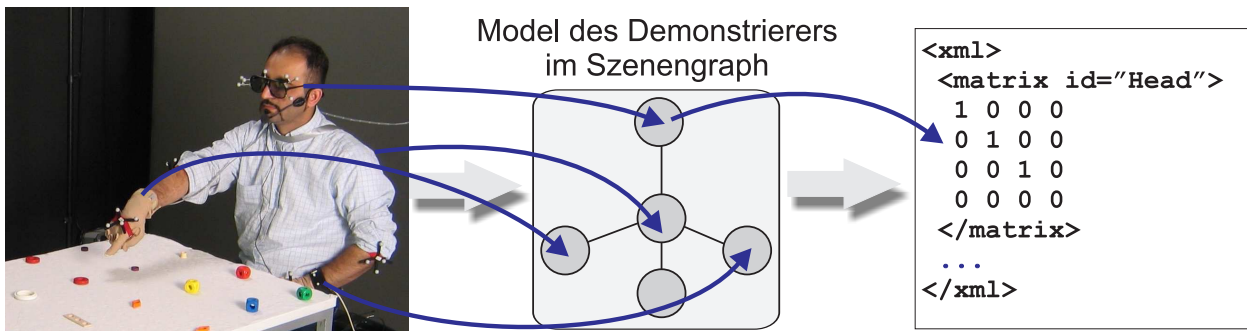


Abbildung 5: Während der Aufzeichnung eines Experimentes baut IADE an Hand der Daten der optischen Tracker ein Modell des Demonstrierenden in einem Szenengraph auf und speichert das fertige Modell dann in einer XML-Datei.

5 IADE: Der Interactive Augmented Data Explorer

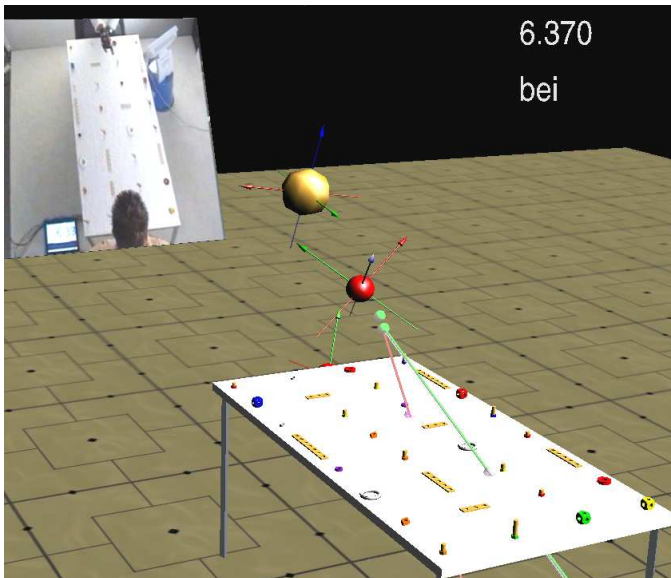
IADE basiert auf dem VR Framework AVANGO [Tra01] und dem PrOSA Framework für multimodale Interaktion [Lat01]. Die Software kann in zwei unterschiedlichen Modi operieren: im *Experimentatormodus* koordiniert sie die Online-Datenerfassung und speichert sie für die Offline-Auswertung im *Explorermodus*. IADE lässt sich über die Skriptsprache *Scheme* steuern und besitzt eine 2D GUI, die z.B. für die Steuerung der Aufzeichnung verwendet werden kann, wenn die Visualisierung in der VR Installation abgeschaltet bleibt.

5.1 Experimentatormodus

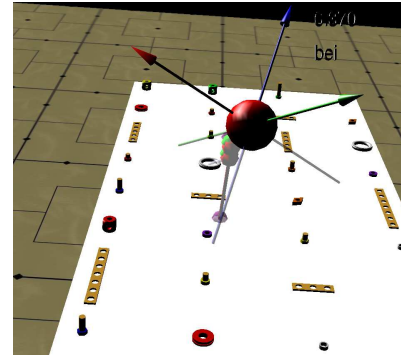
Im *Experimentatormodus* koordiniert IADE die Aufzeichnung der Trackerdaten (Abb. 5) und stellt die Systemzeit auf externen Displays zur Verfügung. Die Videoaufzeichnungen werden momentan manuell gestartet, lassen sich bei Bedarf aber ebenfalls über IADE und eine direkte Firewire-Verbindung automatisieren. Die Videodaten werden dabei auf DV-Kassetten aufgezeichnet und später auf den Rechner übertragen. Dies hat zum einen technische Gründe, da die parallele Erfassung der Videos die Aufzeichnung der Trackingdaten beeinflussen könnte, zum anderen aber auch Sicherheitsgründe: im Eifer der Experimentdurchführung zahlt es sich aus, möglichst nichtflüchtige Medien einzusetzen.

5.2 Explorermodus

Im *Explorermodus* kann IADE die aufgezeichneten Experimente in 3D simulieren. Dabei läuft IADE sowohl auf dem Desktop, als auch in der begehbaren immersiven VR Installation. Für die beschriebene Studie haben wir für die Simulation die wesentlichen Elemente des Settings, d.h. den Tisch mit den Objekten, in der Virtuellen Realität modelliert (Abb. 6). Die aufgezeichneten



(a) Links oben wird über eine Videotextur das von oben aufgezeichnete Video synchron zur Simulation angezeigt. Das seitlich aufgenommene Video ist rechts angebracht und momentan nicht im Bild.



(b) In IADE können beliebige Perspektiven eingenommen werden. Zu sehen ist die Perspektive des Demonstrierenden, die automatisch den aufgezeichneten Daten folgt. Die Pfeile in den Kugeln von Kopf und Handrücken geben die Orientierung an.

Abbildung 6: Die Bilder zeigen Aufnahmen der Desktopversion von IADE. Das Setting auf dem Tisch und der Demonstrierende wurden modelliert. Rechts oben ist die Systemzeit zu sehen, direkt darunter wird aus den Annotationsdaten das aktuelle Transkript der Sprache angezeigt. Die große gelbe Kugel ist der Kopf, die mittlere rote der Handrücken und die kleinen grünen Kugeln sitzen auf dem Zeigefinger (siehe auch 4 d). Die Strahlen aus den Kugeln des Zeigefingers sind Visualisierungen von Zeigestrahlen, die auf Grund der Daten berechnet wurden.

Daten der Tracker werden durch einfache Kugelrepräsentationen visualisiert um die Simulation möglichst neutral beurteilbar zu machen. Ein komplexerer Avatar ist zwar zuschaltbar, vermittelt jedoch einen falschen Eindruck, da z.B. die Blickrichtung nicht aufgezeichnet wurde und somit nicht exakt rekonstruiert werden kann.

Sowohl auf dem Desktop als auch insbesondere in der VR Installation kann die Resimulation wie ein Video abgespielt, angehalten, sowie zu bestimmten Zeitpunkten gesprungen werden. Für die Annotation ist dies ein großer Vorteil, da kritische Stellen nun aus mehreren Blickwinkeln betrachtet werden können und keine Probleme mehr mit Verdeckungen existieren. Zudem können zu jedem Zeitpunkt automatisch Positionen und Orientierungen für die statistische Analyse ausgegeben werden; eine manuelle Annotation entfällt.

Neben den ursprünglichen Daten der Tracker kann IADE auch Annotationen einlesen und synchron abspielen. In Abb. 6 ist oben rechts unter der Systemzeit mit dem Wort „bei“ das aktuell vom Demonstrierenden gesprochene Wort zu lesen. Die aufgezeichneten Videos können über Vi-

deotexturen frei in der Simulation positioniert und synchron abgespielt werden (Abb. 6 a).

IADE präsentiert damit dem Wissenschaftler in der VR Installation oder am Desktop eine integrierte Sicht auf die Daten. Der Forscher kann für die Annotation und Auswertung ergänzende Informationen aus verschiedenen Quellen nutzen und Defekte in einzelnen Kanälen erkennen.

Mittels der Skriptsprache *Scheme* können Berechnungen über den vorhandenen Datensätzen durchgeführt und in die Visualisierung der aktuellen Simulation eingefügt werden. In Abb. 6 a) sind Visualisierungen unterschiedlicher Zeigestrahlen dargestellt: Kopf-Finger-Zeigen (in der Abb. 6 a) rechts, grün), der ideale Zeigestrahl direkt auf das Objekt (mittig, weiß) und der nur über den Zeigefinger definierte Zeigestrahl (links, rot). Im konkreten Beispiel wird für die Berechnungen auch annotierte Informationen über Beginn und Ende eines Gestenstrokes aus einer vorhergehenden Iteration der Annotation verwendet. Die erzeugten Zeigestrahlen wiederum bilden als Hilfslinien die Grundlage für eine neue Iteration, da durch die Schnittpunkte zwischen Strahl und Tisch die Abweichung der Zeigegeste von der Ideallinie gut ermittelt werden kann. IADE kann also effektiv und effizient das iterative Vorgehen bei der Annotation unterstützen.

Komplette Analysesitzungen können von IADE in einem Video mitgeschnitten werden, so dass Interaktionen mit bestimmten Simulationen auch für spätere Arbeitsschritte, zur Archivierung oder z.B. für Tagungsvorträge eingefroren werden können. Dabei können Einstellungen für einen Experimentaldurchgang im Batch-Verfahren auf andere angewendet werden.

6 Diskussion

IADE wurde erst im Rahmen der beschriebenen Experimente konzipiert und entwickelt, konnte aber schnell produktiv eingesetzt werden. Dies war nur durch den hohen Abstraktions- und Reifegrad der verwendeten Frameworks und insbesondere die Rapid Prototyping Fähigkeit von AVANGO möglich.

6.1 Fortschritt in Bezug auf die Studie

Für die Studie zur Parametrisierung des Zeigekegels konnte durch die entwickelte Experimental-methode die gewünschte Genauigkeit erreicht werden. Das optische Trackingsystem ist mit neun Kameraperspektiven mehr als ausreichend dimensioniert und Verdeckungen spielen damit in den kritischen Zeitabschnitten keine Rolle mehr. Diese kommen nur noch dann vor, wenn der Demonstrierende die Hände in der Ruhestellung zwischen zwei Durchgängen unter den Tisch nimmt.

Während der iterativen Konsolidierungsphase des Annotationsschemas konnte IADE einen entscheidenden Beitrag zur Qualitätssicherung leisten, da Unstimmigkeiten zwischen den einzelnen Datenkanälen durch die integrierte Visualisierung frühzeitig erkannt werden konnten.

6.2 Ausblick

IADE unterstützt momentan mit Sprache und Gestik die wesentlichen inhaltstragenden Modalitäten der menschlichen Kommunikation. Für zukünftige Studien sollen mit Hilfe von moderner mobiler Eyetracking Technologie zusätzlich noch die Augenbewegungen und damit Blickrichtungen erfasst und mittels IADE integriert werden. Neben der visuellen Aufmerksamkeit, die über die Augen genauer als über die Kopforientierung ermittelt werden kann, spielen die Blicke auch speziell beim Turn-Taking in natürlichen Dialogsituationen eine Rolle.

Während die Virtuelle Umgebung bei IADE momentan primär zur Exploration verwendet wird und die Konfiguration der Daten per Skript geschieht, soll dies in Zukunft durch eine interaktive Schnittstelle direkt in der Umgebung geschehen. Diese Schnittstelle wird auf Vorarbeiten zur Visuellen Programmierung in AVANGO aufsetzen [BW03].

IADE wird über den Sonderforschungsbereich 360 hinaus im neu eingerichteten Sonderforschungsbereich 673, „Alignment in Communication“, und im EU Projekt PASION, „Psychologically Augmented Social Interaction Over Networks“, eingesetzt.

6.3 Danksagung

Die beschriebenen Experimente wurden an der Universität Bielefeld im Rahmen des Sonderforschungsbereiches 360, „Situierete Künstliche Kommunikatoren“, durchgeführt. Die Autoren danken Hannes Rieser, Marc Staudacher, Tobias Gövert und Dominique Westerhues für die fruchtbare Zusammenarbeit und der Arbeitsgruppe Wissensbasierte Systeme für Unterstützung und Technik.

Literatur

- [BW03] BIERMANN, P. und I. WACHSMUTH: *An Implemented Approach for a Visual Programming Environment in VR*. In: *Proceedings Fifth Virtual Reality International Conference (VRIC 2003)*, Seiten 229–234, Laval, France, May 2003.
- [KLP⁺06a] KRANSTEDT, A., A. LÜCKING, T. PFEIFFER, H. RIESER und I. WACHSMUTH: *Deictic object reference in task-oriented dialogue*. In: RICKHEIT, G. und I. WACHSMUTH (Herausgeber): *Situated Communication*. Mouton de Gruyter, Berlin, 2006.
- [KLP⁺06b] KRANSTEDT, A., A. LÜCKING, T. PFEIFFER, H. RIESER und I. WACHSMUTH: *Deictic Object Reference in Task-oriented Dialogue*. In: GIBET, SYLVIE, NICOLAS COURTY und JEAN-FRANÇOIS KAMP (Herausgeber): *Gesture in Human-Computer Interaction and Simulation: 6th International Gesture Workshop, GW 2005, Berder Island, France, May 18-20, 2005, Revised Selected Papers*, LNAI 3881, Seiten 300 – 311, Berlin, 2006. Springer-Verlag GmbH.

- [KLP⁺06c] KRANSTEDT, A., A. LÜCKING, T. PFEIFFER, H. RIESER und I. WACHSMUTH: *Measuring and Reconstructing Pointing in Visual Contexts*, submitted. In: UNKNOWN (Herausgeber): *Proceedings of BRANDIAL 2006*, unknown, Seite unknown, unknown, 2006. unknown.
- [KW05] KRANSTEDT, ALFRED und IPKE WACHSMUTH: *Incremental Generation of Multimodal Deixis Referring to Objects*. In: *Proceedings 10th European Workshop on Natural Language Generation (ENLG 2005)*, Aberdeen, UK, 2005.
- [Lat01] LATOSCHIK, MARC ERICH: *A General Framework for Multimodal Interaction in Virtual Reality Systems: PrOSA*. In: BROLL, W. und L. SCHÄFER (Herausgeber): *The Future of VR and AR Interfaces - Multimodal, Humanoid, Adaptive and Intelligent. Proceedings of the workshop at IEEE Virtual Reality 2001, Yokohama, Japan*, Seiten 21–25, Sankt Augustin, March 2001. GMD Report No. 138, GMD-Forschungszentrum Informationstechnik GmbH.
- [LS05] LÜCKING, ANDY und JENS STEGMANN: *Assessing Reliability on Annotations - Part 2: Statistical Results for the DEIKON Scheme*. Technischer Bericht 3/2005, Sonderforschungsbereich 360, Universität Bielefeld, 2005.
- [PL04] PFEIFFER, T. und M. E. LATOSCHIK: *Resolving Object References in Multimodal Dialogues for Immersive Virtual Environments*. In: IKEI, YASUSHI, MARTIN GÖBEL und JIM CHEN (Herausgeber): *Proceedings of the IEEE Virtual Reality 2004*, Seiten 35–42. IEEE, 2004.
- [SL05] STEGMANN, JENS und ANDY LÜCKING: *Assessing Reliability on Annotations - Part 1: Theoretical Considerations*. Technischer Bericht 2/2005, Sonderforschungsbereich 360, Universität Bielefeld, 2005.
- [Tra01] TRAMBEREND, HENRIK: *Avango: A Distributed Virtual Reality Framework*. In: *Proceedings of Afrigraph '01*. ACM, 2001.