Tracking and Visualizing Visual Attention in Real 3D Space

Thies Pfeiffer

A.I. Group, Faculty of Technology, Bielefeld University, Germany
Thies.Pfeiffer@Uni-Bielefeld.de

Keywords: visual attention, eye tracking, 3D tracking, spatiality, visualization.

Motivation

Humans perceive, reason and act within a 3D environment. In empirical methods, however, researchers often restrict themselves to 2D, either in using 2D content or relying on 2D recordings for analysis, such as videos or 2D eye movements.

Regarding, e.g., multimodal deixis, we address the open question of the morphology of the referential space (Butterworth and Itakura, 2000), For modeling the referential space of gaze pointing, precise knowledge about the target of our participants' visual attention is crucial. To this ends, we developed methods to assess the location of the point of regard, which are outlined below.

3D Tracking and Visualization

Recent technological advances allow researchers to record body and eye movements of participants at an unmatched precision. We combined body tracking with binocular eye tracking and developed machine learning methods to estimate in real-time the location of the
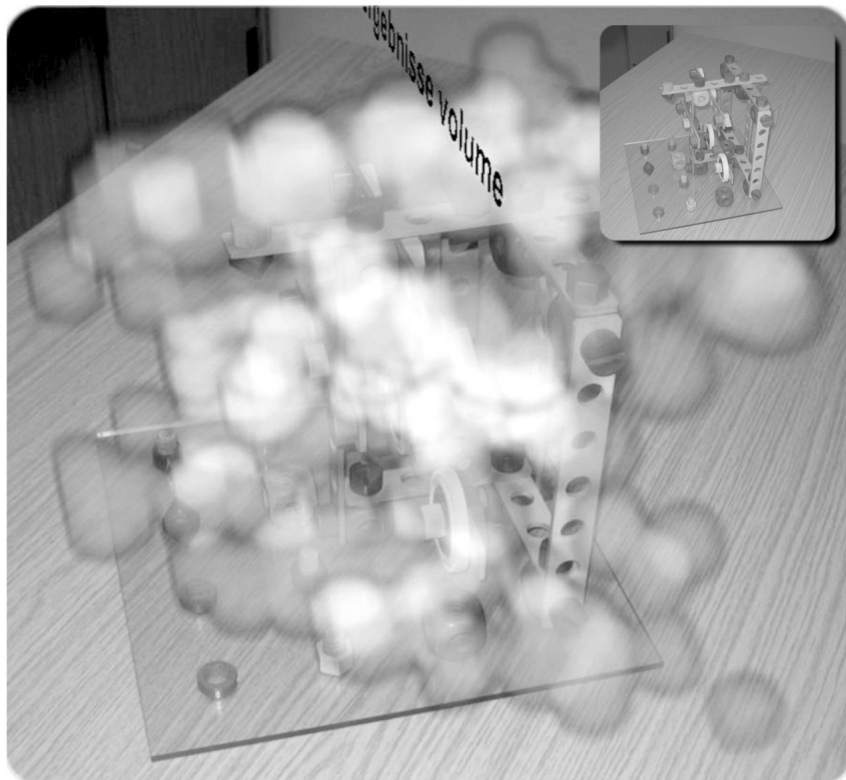

Figure 1: 3D Attention Volume depicting the visual attention on an aggregate of real objects. The 3D locations of the fixations were estimated by measuring convergence.

point of regard in space from positions, orientations and convergence of the eyes (Pfeiffer, Latoschik and Wachsmuth, 2009), These methods can be used in virtual as well as in real worlds and do not require computer-models of the targets. Analysis of the data requires little manual effort. To support visual analysis, we developed *3D Attention Volumes* (Fig. 1), transferring the idea of 2D heatmaps to 3D space, which allow for an interactive exploration of the distribution of attention from any perspective. We expect that these methods can support research in other areas as well.

References

Butterworth, G. and Itakura, S. (2000), How the Eyes, Head and Hand Serve Definite Reference, *British Journal of Developmental Psychology*, 18: 25-50.

Pfeiffer, T., Latoschik, M. E. and Wachsmuth, I. (2009), Evaluation of Binocular Eye Trackers and Algorithms for 3D Gaze Interaction in VR Environments, *Journal of Virtual Reality and Broadcasting*, 5 (16).

Object Deixis: Interaction Between Verbal Expressions and Manual Pointing Gestures

Thies Pfeiffer

A.I. Group, Faculty of Technology, Bielefeld University, Germany
Thies.Pfeiffer@Uni-Bielefeld.de

Keywords: deixis, manual pointing, interaction between speech and gesture, 3D gesture space

Motivation

Object deixis is at the core of language and an ideal example of multimodality. Speech, gaze and manual gestures are used by interlocutors to refer to objects in their 3D environment. The interplay of verbal expressions and gestures during deixis is an active research topic in linguistics as well as in human-computer interaction.

Previously, we conducted a study on manual pointing during dialogue games using state-of-theart tracking technologies to record gestures with high spatial precision (Kranstedt, Lücking, Pfeiffer, Rieser and Wachsmuth, 2006), To reveal strategies in manual pointing gestures, we present an analysis of this data with a new visualization technique.
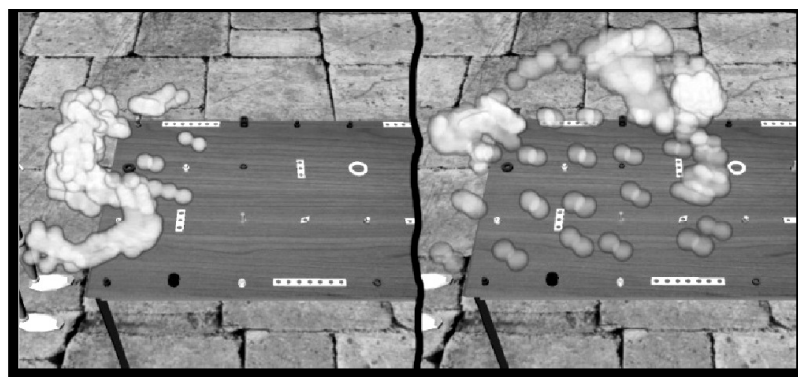

Fig. 1: The *Gesture Space Volumes* of the end positions of manual pointing gestures (spheres, left: co-verbal, right: without speech) reveal different strategies for manual pointing depending on the availability of speech.

Method
During the study with 50 participants, about 1000 manual pointing gestures to objects located on a real table (Fig. 1 shows a virtual model) were recorded in two settings. In one setting the participants were allowed to speak, in the other setting they had to remain silent.

Results
Previously, we identified an increase in words when referring to distant objects, but were not able to identify strategies used in pointing *without* speech. Therefore we recently developed the new *Gesture Space Volume (GSV)* visualization. GSVs visualize spatial and temporal properties of multiple gesture acts, similar as heatmaps do for gaze, but in 3D. We found that over 60% of the participants used a successful *leaning-forward* strategy when pointing without speech (Fig. 1),

References

Kranstedt, A., Lücking, A., Pfeiffer, T., Rieser, H. and Wachsmuth, I. (2006), Deixis: How to Determine Demonstrated Objects Using a Pointing Cone, In: S. Gibet, N. Courty and J.-F. Kamp (Eds.), *Gesture in Human-Computer Interaction and Simulation: 6th International Gesture Workshop*, Springer-Verlag GmbH, Heidelberg, p. 300-311.

Is Verbal Short-Term Memory Linguistically Structured?

Kamila Polišenská, Shula Chiat, Penny Roy

City University London, United Kingdom
kamila.polisenska.1@city.ac.uk

Keywords: short-term memory, linguistic structure, cross-linguistic

This study set out to explore cognitive processes involved in verbal repetition by investigating the effects of linguistic factors on memory span. In most studies investigating verbal memory, items are presented in a list and the focus is on properties of these items such as imageability, lexical status or word-likeness; investigation beyond the word level is rare. The present study compares performance on sequences of items which are manipulated regarding syntactic, semantic and prosodic properties at a sentence level, while holding constant the nonexperimental characteristics of constituent words (frequency, familiarity, imageability and phonological complexity), Performance on immediate repetition of sequences of items in 7 conditions, with well-formed sentences at one end of the spectrum and lists of nonwords at the other, is compared in two typologically different languages (see the examples from English below):