Studying Gaze-based Human Robot Interaction: An Experimental Platform

Stefan Kohlbecher
Institute of Clinical Neurosciences
University Hospital Munich

Eva Wiese
Department of Experimental Psychology,
LMU München

Klaus Bartl
Institute of Clinical Neurosciences
University Hospital Munich

Jürgen Blume
Human-Machine Communication
Electrical Engineering and Information Technologies
Technische Universität München

Alexander Bannat
Human-Machine Communication
Electrical Engineering and Information Technologies
Technische Universität München

Erich Schneider
Institute of Clinical Neurosciences
University Hospital Munich

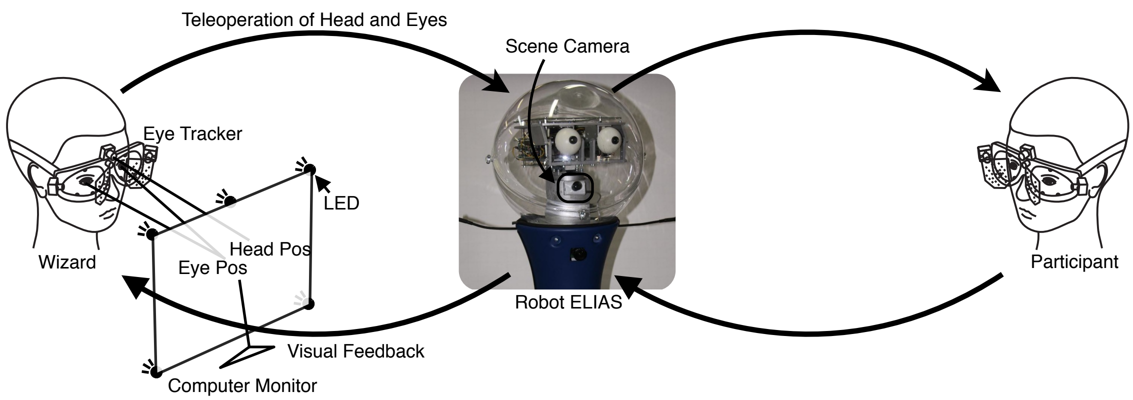

Figure 1. Wizard-of-Oz scenario: A human (Wizard) controls the Robot ELIAS with his own head and eye movements. The wizard‘s monitor displays a video stream of the robot‘s scene camera. Seeing the world from the robot‘s perspective allows him to interact virtually with the Participant.

ABSTRACT

In this paper, a platform for studying human-robot interaction is presented. Apart from presenting a robotic eye module, which exceeds human performance in terms of velocity and acceleration, we introduce a novel method of controlling a robot‘s head and eye movements. In a Wizard-of-Oz scenario, a human operator (the wizard) perceives the world from the perspective of a robot through a video link. His eye and head movements are recorded and, in turn, fed back to the robot. This allows for very natural eye and head movements when the wizard virtually interacts through the robot with one or more participants. Additionally the robot can carry out scripted movements to conduct repeatable experiments in a controlled manner. It is also possible to record the gaze behaviors of the participant(s).

Keywords

Head-mounted, eye tracking, teleoperation, wizard-of-oz.

Proceedings of the „Gaze in Human-Robot Interaction“ Workshop held at the 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2012). March 5, 2012, Boston, MA, USA.

1. INTRODUCTION

Recently, user interfaces are becoming more and more natural, the trend is moving away from artificial paradigms like mouse and keyboard towards more intuitive methods like multitouch or speech interfaces. Applying this trend to humanoid robotics also suggests to use natural speech, but one can easily think of further characteristics to use; for example, gaze, body gestures and mimics. Especially gaze plays a vital role in communication between humans [11]. During a conversation of more than two people, there is a positive correlation between gaze activity and amount of turn taking [12]. Gaze can be used to convey emotions such as rolling one's eyes when the topic becomes boring.

To evaluate how well gaze-based human-robot interaction can work, an experimental platform is indispensable. In this work, we present a solution which can be used to exactly reproduce human eye movements on a robotic platform. This is achieved by a Wizard-of-Oz setting, where a human operator can see the world from a robot‘s perspective and, in turn, can control the robot‘s head and eyes with his own head and eyes. The participant, whose eyes can also be tracked optionally, interacts with the wizard only through the robot and – depending on the experiment – has not to be made aware of the fact that a real human is behind it.

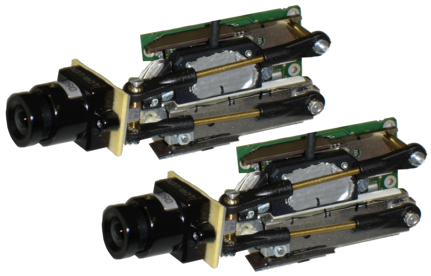

Figure 2. Robotic eyes. The robot‘s eyes are moved by piezo actuators which exceed the human performance in terms of velocity and acceleration.

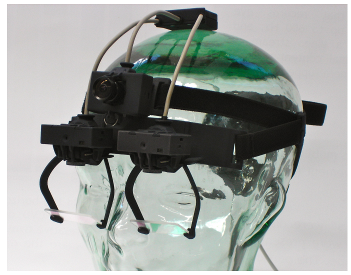

Figure 3. Eye tracker. The eye tracking cameras are mounted on the head band facing downwards. A wide angle scene camera is mounted in the center.

Figure 4. View from the wizard‘s perspective. The wizard gets feedback where he is looking/pointing his head at to ensure correct calibration.

2. TECHNIQUES

Several key techniques have been used in this platform.

2.1 Robotic Eyes

At the core are the robotic eyes, which not only match, but exceed human performance in terms of velocity and acceleration, due to their ultrasonic motors. On the basis of earlier work [4, 9, 8, 7, 6] we have developed a new 2-DOF camera motion device that was considerably reduced in size and weight (see Figure 2). At the center of its design there is a parallel kinematics setup with two piezo actuators (Physik Instrumente, Karlsruhe, EU) that rotate the camera platform around a cardanic gimbal joint (Mädler, Stuttgart, EU) by means of two push rods with spherical joints (Conrad, Hirschau, EU). The push rods are attached to linear slide units (THK, Tokyo, Japan), each of which is driven by its own piezo actuator. The position of each slide unit is measured by a combination of a magnetic position sensor (Sensitec, Lahnau, EU) and a linear magnetic encoder scale with a resolution of 1 µm (Sensitec, EU) that is glued to the slide unit. The microprocessor (PIC, Microchip, Chandler, AZ, USA) on the PCB implements two separate PID controllers, one for each actuator-sensor unit. The controller outputs are fed into analog driver stages (Physik Instrumente, EU) which deliver the appropriate signals to the piezo actuators. The transformation of desired angular orientations to linear slide unit positions is performed by means of an inverse kinematic which has been presented before [8].

2.2 Robot

The robot head itself (ELIAS) was taken from a commercially available robot (MetraLabs, Ilmenau, EU), which consists of a mobile platform and a 2-DOF head unit (pan/tilt). Originally, the head was also equipped with 2-DOF pivotable eyes which, however, did not meet the requirements for velocity and acceleration and were therefore replaced. We chose this particular robot head since, in contrast to a more human-like android platform, it possesses stylized facial features and is therefore perceived as artificial. This allowed us to leave aside ”uncanny valley” effects that appear as soon as the robot closely resembles its human model [2]. Furthermore, this choice also reduced human likeness even more to eye and head movement behavior, as opposed to comparing the overall appearance.

2.3 Eye Tracking

For the eye tracking, we use our own EyeSeeCam system [6]. It is based on a head band with two adjustable, downward-facing infrared cameras (see Figure 3). These cameras observe the wearer‘s eyes via a hot mirror that only reflects infrared light but lets visible light pass straight trough, thus not obstructing the user‘s field of view. The eyes are illuminated by IR LEDs, making the system independent of the ambient lighting. The eyes appear in the camera images in greyscale with the darkest areas being the pupils. The center of the pupils are detected by applying image processing algorithms. To determine the actual line-of-sight in head coordinates, we use an eye model which is calibrated by fixating five known points (the primary position as well as 8.5° to the left, right, up and down). These points are generated by splitting up a laser beam originating from the user‘s forehead by a special diffraction grating [3]. A third, wide angle camera is used to capture the scene. All cameras are fully synchronized and record images at the exact same point in time at a frame rate of 220 Hz. The images are transferred to an Apple MacBook running Mac OS X, on which the image processing and tracking algorithms for all three cameras are implemented.

2.4 Head Tracking

In this application, we use an infrared scene camera to detect the head pose of the wizard who sits in front of a computer monitor (see Figure 1). Infrared markers are placed in each corner of the display. These markers can easily be detected in the scene camera image. In 3D space, a plane is defined by three points. Since we only observe the projection of the markers on the image plane, a fourth point is needed to determine the position and orientation of the monitor with respect to the camera [1]. Given the eye tracker geometry, and an additional calibration step to determine the eye ball positions, both the head position and line-of-sight can be calculated with respect to the wizard‘s monitor.

2.5 Teleoperation Link

The robot has a scene camera mounted to its neck, which is fixed to its body and does not move with the head. The image of this camera is displayed on the wizard‘s monitor. His eye and head movements are recorded and fed back to the robot which moves accordingly. Since both the robot‘s scene camera and the wizard‘s monitor are fixed, positive feedback is of no concern. In previous work, we measured the overall latency between the wizard‘s and robot‘s eye movement, which was on the order of 12 ms, and thus physiologically irrelevant [5].

3. OPERATIONAL AREAS

The platform as a whole can be used to conduct several types of studies:

3.1 Wizard-of-Oz Control

In this setup, one or more participants interact with a robot who is controlled by a real human, the wizard. This guarantees eye and head movements to be as natural as possible. An additional audio link can be easily set up to support natural spoken conversation.

We have already experimented with a wireless variant of this setup on trade fairs, where the wizard was even able to control the robot‘s locomotion (see Figure 4). The overall reaction suggests a high acceptance. Even though we did not implement an audio channel, visitors of all ages could be easily lured into interaction by means of gaze alone.

In upcoming experiments, we plan on alternating the eye movement characteristics. For example, artificial latency can be introduced, or saccades can be completely removed. This will enable us to determine which properties of human-human gaze are vital to communication and which are not necessary and can therefore be omitted from the transfer to robotic systems.

3.2 Script Control

The robotic platform can also be used without teleoperation. In this case, we can use a MATLAB script to control the robot. This allows us to conduct well-defined and repeatable gaze-based human-robot experiments. One such experiment is presented in detail in another workshop contribution: Participants were asked to estimate the gaze direction of human and robotic counterparts. Remarkably, the participants did overall equally well in estimating the gaze directions. Results are presented in the companion paper [10].

3.3 Eye Movement Model Validation

Ultimately, gaze should be modeled for autonomous robots. Then, our platform opens up the possibility of evaluating such a model. Its performance can then be compared to that of a real human by switching between model and wizard controlled gaze.

3.4 Synchronous Eye Tracking

By using an additional eye tracker for each participant, we can also measure the gaze behavior of the humans interacting with the robot synchronously to its eye movements. This, in turn, can also be used to compare the gaze behavior of the humans depending on their counterpart being a robot or another human.

4. CONCLUSION

We presented a novel platform for studying gaze-based human-robot interaction. With the Wizard-of-Oz scenario, we created a unique facility to operate a humanoid robot in a very natural way. Together with the possibility to build and validate autonomous gaze models, and to measure the participants‘ reaction, a foundation has been laid which can help to eventually develop a humanoid robot with human-like gaze behavior. Interacting with such a robot would be a very natural experience and therefore easy to learn.

5. ACKNOWLEDGMENTS

This work is supported by Bundesministerium für Bildung und Forschung (*IFB*[LMU], 01EO0901), in part within the DFG excellence initiative research cluster “Cognition for Technical Systems – CoTeSys”, see also www.cotesys.org and also in part by the project CustomPacker funded by the European Commission under the 7th Framework Programme, see www.custompacker.eu.

6. REFERENCES

[1] S. Kohlbecher, K. Bartl, S. Bardins, and E. Schneider. Low-latency combined eye and head tracking system for teleoperating a robotic head in real-time. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, pages 117–120. ACM, 2010.

[2] M. Mori. The uncanny valley. *Energy*, 7(4):33–35, 1970.

[3] J. Pelz and R. Canosa. Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, 41(25-26):3587–3596, 2001.

[4] E. Schneider, K. Bartl, S. Bardins, T. Dera, G. Boening, and T. Brandt. Eye movement driven head-mounted camera: It looks where the eyes look. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pages 2437–2442, 2005.

[5] E. Schneider, S. Kohlbecher, K. Bartl, F. Wallhoff, and T. Brandt. Experimental platform for wizard-of-oz evaluations of biomimetic active vision in robots. In *Robotics and Biomimetics (ROBIO), 2009 IEEE International Conference on*, pages 1484–1489. IEEE, 2009.

[6] E. Schneider, T. Villgrattner, J. Vockeroth, K. Bartl, S. Kohlbecher, S. Bardins, H. Ulbrich, and T. Brandt. EyeSeeCam: An eye movement-driven head camera for the examination of natural visual exploration. *Ann N Y Acad Sci*, 1164:461–467, 2009.

[7] T. Villgrattner and H. Ulbrich. Control of a piezo-actuated pan / tilt camera motion device. In *11th Int. Conf. on New Actuators, Actuator 2008*, pages 129–132, Bremen, Germany, Jun 2008.

[8] T. Villgrattner and H. Ulbrich. Piezo-driven two-degree-of-freedom camera orientation system. In *Industrial Technology, 2008. ICIT 2008. IEEE Int. Conf. on*, pages 1–6, 21-24 April 2008.

[9] P. Wagner, W. Günthner, and H. Ulbrich. Design and implementation of a parallel three-degree-of-freedom camera motion device. Joint Conference of the 37th International Symposium on Robotics ISR 2006 and the German Conference on Robotics Robotik 2006, May 2006.

[10] E. Wiese, S. Kohlbecher, and E. Schneider. Acuity in estimating gaze direction in human-human and human-robot interaction. In *Proceedings of the „Gaze in Human-Robot Interaction“ Workshop (HRI 2012), Boston*, 2012.

[11] M. Argyle and M. Cook. Gaze and mutual gaze. 1976.

[12] R. Vertegaal, G. van der Veer, and H. Vons. Effects of gaze on multiparty mediated communication. In Graphics Interface, pages 95–102, 2000.