Eye Tracking for Human Robot Interaction

Oskar Palinko, PhD Fellow
Robotics, Brain and Cognitive Sciences
Istituto Italiano di Tecnologia
Genova, Italy

Abstract

Humans use eye gaze in their daily interaction with other humans. Humanoid robots, on the other hand, have not yet taken full advantage of this form of implicit communication. We designed a passive monocular gaze tracking system implemented on the iCub humanoid robot [17]. The validation of the system proved that it is a viable low-cost, calibration-free gaze tracking solution for humanoid platforms, with a mean absolute error of about 5 degrees on horizontal angle estimates. We also demonstrated the applicability of our system to human-robot collaborative tasks, showing that the eye gaze reading ability can enable successful implicit communication between humans and the robot.

Keywords: eye tracking, human robot interaction

1 Introduction

Humans use a number of communication cues in their daily interaction with other humans: primarily speech but also gestures, pointing and gaze [1]. The main purpose of gaze is to provide visual information to the subject, but at the same time a person's gaze implicitly provides information to an outside observer about what the subjects are focusing their attention on. There is a number of ways how eye gaze is implicitly used during communication: gaze aversion, mutual gaze, gaze pointing, join attention, etc.

Humans are very good at reading other people's gaze, but robots are less so. This ability would be especially important for humanoid robots to be able to mimic human abilities. However, most human robot interaction experiments today use head pose as a proxy for real eye gaze because it's easier to extract than eye gaze [2] [3] [4] [5]. But "head gaze" does not provide all the information that eye gaze does [6], thus enabling robots to perform eye tracking could significantly improve its abilities and also its acceptance by humans. Moreover most interaction experiments focusing on gaze use external eye tracking systems [18]. Although this choice guarantees high spatial and temporal precision of the measure, it might in some contexts interfere with natural interactive behavior, inducing participants to be constantly aware of their own gaze motions. A proof of concept gaze tracker was realized by Matsumoto and Zelinsky [7] and implemented on the HRP2 humanoid [8]. More recently Sciutti et al. [9] implemented a mutual gaze detection system on the iCub which facilitated a teacher/student scenario. Still, so far no extensive use of eye gaze tracking has been done in human-robot interaction.

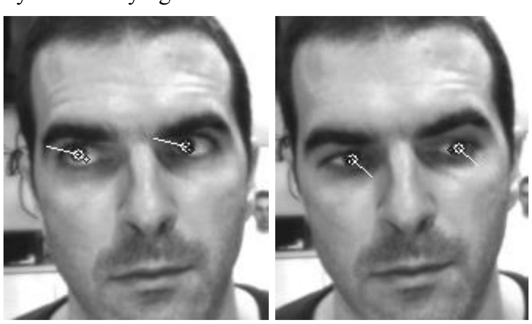

Figure 1. Example performance of our eye tracking system on the iCub robot.

2 Approach and Results

We implemented a monocular feature-based passive gaze tracking algorithm on the iCub platform with the goal of facilitating human robot interaction. The first step in eye tracking is detecting faces and finding face features. For this purpose we used King's implementation [10] of Khazemi and Sullivan's approach for finding features like the corners of the eyes and mouth [11]. We also used Baltrusaitis implementation of the Constrained Local Models approach for tracking head pose [12]. Once these measures were found we proceeded to apply an eye model to the detected center of the pupil similarly as in [13]. The model finally provided the estimate of the gaze angle of the subject, see Figure 1. We then performed a validation experiment in which we found the gaze estimates to be quite acceptable for our setup: the absolute error in the horizontal plane was 5 degrees on average. The accuracy of our system was limited by the cameras used in the iCub setup. We employed PointGrey Dragonfly2 cameras in 1024x768 resolution with fixed-focus 4mm lenses, which produce images of the iris with 20 pixels in diameter when the subject is at 60cm. Knowing that the average diameter of the iris [14] is similar in size to the average eye radius (12mm) [15], then one

Figure 2. Human robot interaction example where the subject needs to ask for taking toy building blocks from the robot.

pixel difference in the middle of the iris corresponds to about 3 degrees difference in gaze. Thus our accuracy is greatly influenced by the hardware used. It is foreseeable that the progressive development of cheaper and smaller cameras will allow future robotic platforms to have higher resolution sensors, with a consequent improvement of the accuracy of our system. In the meanwhile, the current hardware already enables a gaze estimation from the iCub robot that it can exploit to manage human-robot collaboration tasks.

We also conducted a proof of concept human robot interaction experiment in which subjects were seated opposite of the robot and experimenter, who held toy building blocks in their hands, see Figure 2. The subject's role was to ask for the blocks in specific order, but we did not provide information on how to communicate with the robot. Participants used a combination of speech, pointing and gaze to achieve the task, but the robot really only reacted to gaze. More precisely, the robot handed over pieces of toy building blocks when it detected a succession of mutual gaze and gazing at the requested object. The subjects were not aware of the robot's gaze reading ability, but could still complete the task of building a pillar out of these blocks just by using natural eye behavior in less than 30 seconds. Hence, the robot succeeded in exploiting naturally occurring human gaze behavior to control its helping actions in a collaborative manner.

3 Discussion

Future benefits of a built-in gaze tracker in a humanoid robot can be manifold: it could improve turn taking, joint attention and in general the processing of all the communicative gaze cues typical of human interaction. Furthermore, the robot could potentially be used for diagnosing early behavioral problems associated with gaze processing as Autism Spectrum Disorders, by monitoring subjects' gaze in real time. The monitoring of gaze and post hoc analysis has proven to be helpful for ASD diagnosis [16]. Our system would allow for appropriate contingent reaction to gaze by the robot, something that now is achieved only through remote control by the therapist.

4 Future Work

Within the scope of the PhD research, it is planned to improve the quality of the eye tracking system as well as conducting more experiments which would prove the usefulness of eye tracking in human robot interaction. The quality of tracking could be increased by applying more precise tracking of the pupil, face features and head orientation. The planned experiments would emphasize the benefits of real eye tracking as opposed to using head tracking only. We would focus on exploring phenomena as joint attention, mutual gaze and gaze aversion in both in dyadic and group scenarios.

Acknowledgements

The research presented here has been supported by the European CODEFROR project (PIRSES-2013-612555).

References

[1] N. George and L. Conty, "Facing the gaze of others," *Neurophysiol. Clin.*, vol. 38, no. 3, pp. 197–207, Jun. 2008.

[2] M. W. Doniec, G. Sun, and B. Scassellati, "Active Learning of Joint Attention," in *Humanoid Robots, 6th IEEE-RAS International Conference on*, 2006, pp. 34–39.

[3] H. Kim, H. Jasso, G. Deák, and J. Triesch, "A robotic model of the development of gaze following," in *7th IEEE International Conference on Development and Learning, ICDL*, 2008, pp. 238–243.

[4] S. Ivaldi, S. M. Anzalone, W. Rousseau, O. Sigaud, and M. Chetouani, "Robot initiative in a team learning task increases the rhythm of interaction but not the perceived engagement," *Front. Neurorobot.*, vol. 8, 2014.

[5] S. Sheikhi and J.-M. Odobez, "Combining dynamic head pose–gaze mapping with the robot conversational state for attention recognition in human–robot interactions," *Pattern Recognition Letters, 2014.*

[6] A. Borji, D. Parks, and L. Itti, "Complementary effects of gaze direction and early saliency in guiding fixations during free viewing.," *J. Vis.*, vol. 14, no. 13, Jan. 2014.

[7] Y. Matsumoto and A. Zelinsky, "An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement," in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, 2000, pp. 499–504.

[8] J. Ido, Y. Matsumoto, T. Ogasawara, and R. Nisimura, "Humanoid with interaction ability using vision and speech information," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2006.

[9] and G. S. A. Sciutti, L. Schillingmann, O. Palinko, Y. Nagai, "A Gaze-contingent Dictating Robot to Study Turn-taking," in *Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction*, 2015.

[10] D. E. King, "Dlib-ml : A Machine Learning Toolkit," *J. Mach. Learn. Res.*, vol. 10, pp. 1755–1758, 2009.

[11] V. Kazemi and J. Sullivan, "One Millisecond Face Alignment with an Ensemble of Regression Trees," in *Computer Vision and Pattern Recognition (CVPR), 2014*.

[12] T. Baltrusaitis, P. Robinson, and L. P. Morency, "3D Constrained Local Model for rigid and non-rigid facial tracking," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2012, pp. 2610–2617.

[13] T. Ishikawa, S. Baker, I. Matthews, and T. Kanade, "Passive Driver Gaze Tracking with Active Appearance Models," in *Proc. World Congress on Intelligent Transportation Systems*, 2004.

[14] S. Thainimit, L. A. Alexandre, and V. M. N. de Almeida, "Iris surface deformation and normalization," *Communications and Information Technologies (ISCIT), 2013 13th International Symposium on*. pp. 501–506, 2013.

[15] I. Bekerman, P. Gottlieb, and M. Vaiman, "Variations in Eyeball Diameters of the Healthy Adults," *J. Ophthalmol.*, vol. 2014.

[16] S. M. Mavadati, H. Feng, A. Gutierrez, and M. H. Mahoor, "Comparing the gaze responses of children with autism and typically developed individuals in human-robot interaction," in *Humanoid Robots (Humanoids), 14th IEEE-RAS International Conference on*, 2014, pp. 1128–1133.

[17] G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori, "The iCub humanoid robot : an open platform for research in embodied cognition.," in *Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems*, 2008, pp. 50–56.

[18] F. Broz and H. Lehmann, "Mutual gaze, personality, and familiarity: Dual eye-tracking during conversation," *RO-MAN*, 2012.