9th International Workshop on Human
Human Friendly Robotics, 29. & 30. September
2016, Genoa, Italy

Social Gaze Behavior for Face to Face Human-Robot Interaction

Hagen Lehmann[1], Frank Broz[2], Reza Ahmadzadeh[3], Alessio Del Bue[1], Lorenzo Natale[1] and Giorgio Metta[1]

Abstract— **This short paper discusses the importance of human-like gaze behaviors for humanoid robots with physical eyes. It gives a brief overview of the functions gaze fullfills in human-human interactions from a social evolutionary perspective. In the second part we describe how human-like gaze has been implemented in robot in the research field of social robotics in the last years. The last part of the paper briefly introduces an architecture for a conversational gaze controller (CGC). The parameters of this gaze controller are based on the analysis of a large corpus of gaze tracking data collected during human-human conversations. We describe our experiemental approach for obtaining this data and discuss the implications of endowing robots with humanlike behaviors and enabling them to engage in face to face social interactions.**

I. INTRODUCTION

The eyes are the most important nonverbal communication channel humans use in face to face interactions. The role of the human eye in communication has a long evolutionary history in human phylogentic development, and has most likely shaped its physiological struture. No other primate species has visible as pupils like as humans [9], [10]. It has been argued that the high visibility of the human sclera has evolved as an additional source of information about the intention of social interaction partners within a group [19]. On one hand communicating with glances instead of pointing gestures can be very advantages in situation in which stealth is needed, on the toher hand the increased visibility of the sclera made involuntary and subconscious eye movements visible to other members of human groups. The ability to recognize and correctly interpret these movements and the relevant emotional states [1] facilitated the human capacity to live in large and complex social groups and to understand each other as intentional agents. This enabled us to interact empathically with one another. Without the appropriate gaze information we feel uncomfortable during social interactions with others - being starred at makes us feel uncomfortable very quickly, not being looked in the eyes while being spoken to makes us feel nervous. During conversations humans jointly regulate their eye contact. This is called mutual gaze and has a variety of social functions. It helps to regulate turn taking in conversations [8], transmits social dominance [6], and is an expression of the interaction of different personality traits of the persons involved in a conversation [8]. The ability to interpret, follow and use the gaze of the others can be found in humans at a very young age. It has been shown in many studies that human infants can follow eye gaze to discover a hidden reward from the age of 12 - 18 months (e.g. [5]).

Robots are increasingly moving into the public domain. This will require them to be able to communicate intuitively and naturalistic with their human users. One of the main intended functions of this new generation of service robots will be social interaction. Since humans rely heavily both on head and on eye gaze information from their social partners during social interactions, we propose that is necessary for an intuitive and comfortable face to face interaction with a robot to take the social functions of the human eye into consideration. We will briefly discuss how eye gaze is used in human-robot interaction at the moment and describe our approach to implement gaze and our conversational gaze controller (CGC) in more detail. In the final part of the paper we will discuss the implications of the use of a naturalistic CGC, as well as of our future plans of integrating our CGC into a more holistic and reactive social interaction architecture.

II. HUMAN-LIKE GAZE IN SOCIAL ROBOTICS

The importance of incorporating gaze into robotic behavior has been recognized in the research field of social robotics early on. Different attempts have been made to simulate humanlike gaze (e.g. [3], [14]). There are different obstacles that need to be overcome in order to succeed in this endeavor. The existence of human-like features in a robot creates an expectation about their movements in the user that can, if not fulfilled, make the robot awkward and uncomfortable to interact with [12]. For parts that are crucial for naturalistic human robot communication, like robot eyes, this is specifically true. There are different variables that have to be taken into consideration in order to achieve natural looking robotic gaze. The moving speed of the pupils is as important as the frequency in which the robot is switching from one focal point to the next. In case of face to face interaction these focal points need to be the different facial features of the interlocutor. A question that needs to be answered is whether it is necessary to simulate saccadic movements to achieve a naturalistic appearance of the robotic eye movement. It is also important to explore how important the exact simulation of the vergence is for direct face to face interaction. The movement of the head during and in between the gaze changes is another variable that needs to be taken

[1]Hagen Lehmann, Alessio Del Bue, Lorenzo Natale and Giorgio Metta are with the iCub Facility, Istituto Italiano di Tecnologia, Via Morego 30, 16163 Genoa, Italia hagen.lehmann@iit.it, alessio.delbue@iit.it, lorenzo.natale@iit.it, giorgio.metta@iit.it

[2]Frank Broz is with the Department of Mathmatical and Computer Science, Heriot Watt University, Edinburgh, EH14 4AS, UK f.broz@hw.ac.uk

[2] Reza Ahmadzadeh is with the Institute for Robotics and Intelligent Machines, College of Computing, Georgia Institute of Technology reza.ahmadzadeh@gatech.edu

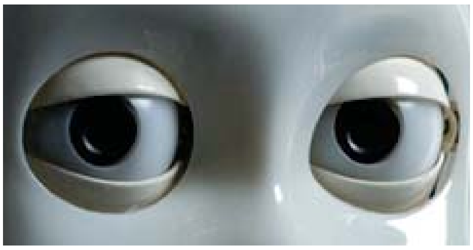

Fig. 1. The diffferent structures of iCub's eyes

into consideration. The impact of these variables needs to be systematically explored in HRI experiments during which the robot is reactive to the behaviors of its human counterpart.

The design of physical robotic eyes has been approached from different directions. Due to the difficulties to build and control physical eyes, many solutions feature projected eyes (e.g. [7]). Nevertheless there are robots that mimic or have real humanlike eyes. Some solutions use dark plastic disk or differently colored lights as indicators of where the robots eyes are (e.g. [17]), others try to mimic certain aspects of the human eye (e.g. [2]), and in some cases very authentic looking artificial eyes are used to create the illusion human likeness (e.g. [15]).

There has been a number of studies that investigated how people respond to different gaze behaviors of robots during conversational tasks, but most of the models used to generate the robot behaviors were either not based on human gaze data (e.g. [20]) or were not reactive (e.g. [13]). Nevertheless the results of these experiments show that humans can use robot gaze information to solve tasks and that human gaze information can be used by robotic systems. But for creating a situation in which the robot is simulating supposedly subconscious gaze behaviors that have the potential to influence a humans comfort level during a conversation it is necessary to make robot gaze reactive.

We decided to use the iCub robot for the implementation of our conversational gaze controller (CGC). Its eyes have all the features needed to simulate human eye gaze (see Figure 1). Each of iCubs eyes incorporates a camera within a black central structure. This black center is surrounded by a white eyeball. The eyes are integrated in eye sockets. The robot has also moveable eyelids, which enable it to blink. The camera lens and the black center of each eye resembles the pupil and the iris of a human eye, whereas the white eyeball resembles the sclera. These structures make the eye gaze of iCubs eyes visible in a similar way human eye gaze is.

III. Conversational Gaze Controller Architecture

The specific goal of our architecture is to generate appropriate reactive gaze behaviors for the iCub robot. We will integrate the different variables necessary to achieve this goal into our architecture (see Figure 2). In a first step the robot will switch its gaze during a conversation in such a way that it will successively focus on different facial features of its human interlocutor. We are modelling these gaze switches

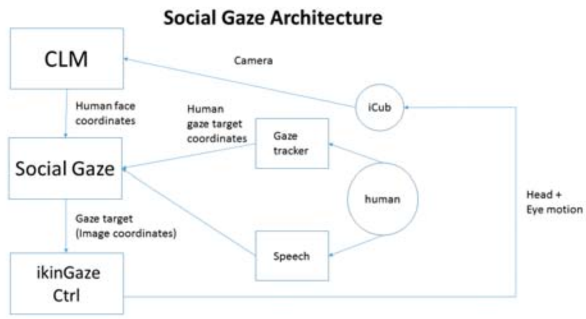

Fig. 2. Illustration of the architecture of our CGC. The outer loop is what we have realized at the moment. The gaze and blinking information will be integrated in a second step. In a third step we will integrate also the speech onset information from the human.

based on data collected during human-human conversational interaction (see section Data Colletion).

Currently, the model only incorporates visual data and does not take the face-oriented gaze direction of the conversation partner into account. The next step for this model will be to integrate in realtime the human partners displayed gaze direction, so that the robots gaze is also dependent on its partners gaze behavior. We plan to experimentally evaluate the impact of modelling and responding to partner gaze direction. A further extension of this model is to include the modality of speech. Adding information about which partner is currently speaking should further improve the naturalism of the gaze. Use of speech data would also enable the generation of a model that incorporates the dynamics of turn-taking as hidden state. While the partner who holds the conversational floor is not directly observable, the turn-taking state can be inferred from the gaze and speech signals of the participants, leading to sophisticated control for conversational gaze.

A. Data Collection

The human-human gaze data was collected from 32 pairs of participants. At the beginning of the experiment the participants were instructed that they were allowed to discuss any topics they liked during their conversation. In case they could not think of a topic, a list of ice breakers was provided. These suggested conversation topics included: hobbies, a recent vacation, restaurants, television shows, or movies. The pairs were seated approximately 1 meter apart with a desk between them. At the beginning of the session, the participants were asked to complete a survey to collect their demographic information and level of familiarity with their partner. The data collection for this experiment and different results from the analysis of the human-human interactions are described in detail in [4]. The gaze was measured by using automated face tracking to locate the participants faces in the video streams and fusing this information with the partners gaze location on a frame-by-frame basis.

The data collected was used to create a Markov chain model of facial feature oriented gaze direction in order to produce naturalistic mutual gazing behavior. Rather than solely modelling mutual gaze as looking toward or away from the partners face, the model also captures the facial features that are looked at and the direction in which gaze is averted. We hypothesize that this greater fidelity in mutual gaze behavior will help humanoid robots with sophisticated eyes to appear significantly more natural. The face tracker used for data collection obtains an outline of the eyes and mouth. For simplicity, the model classifies face-directed gaze to each feature based on a coarse segmentation of the face bounding box into regions. This course segmentation allows us to classify feature-directed gaze to the correct facial feature in the presence of small gaze tracker calibration errors. The states of the Markov chain are as follows:

- Left eye region
- Right eye region
- Mouth
- Away up
- Away left
- Away down
- Away right

The gaze state for the robot controller is produced by simulating from this Markov chain and choosing a target gaze point in the selected region in relation to the partners tracked face at each time step.

IV. Discussion

We have successfully implemented the first step of our social gaze architecture, which allows the robot to switch its gaze between the different facial features of its conversational partner in a human-like way. While implementing the first step we have encountered a variety of interesting problems concerning the appearance of the gaze behavior. It seems critical to choose the correct ratio between head and eye movements when shifting gaze between different facial features. The participants of our preliminary tests were very sensitive to the amount of head movements during gaze shifts. One of the outcome of our first pilot study was that according to our participants the robot should only use head movements when shifting its gaze away from the human's face. When the robot was only shifting the gaze between facial features, the participants felt more comfortable when its head stayed fixed and it only moved its eyes. This however poses another challenge. When the robots head is completely fixed, the behaviors start to appear mechanic. Humans use a lot of different small and subconscious movements during conversations to signal their engagement level. These signals include nodding or shaking of the head as well as minimal hand gestures. To integrate these movements into the overall behavior of the robot is an additional issue and for the moment we are planning to use small head movements to improve the animacy of the robot.

Even though the movements of our CGC are right now not reactive, they already induced strong reactions in the participants in our preliminary test. The issues we have encountered illustrate that implementing naturalistic conversational gaze in social robots is important, but far from easy. It is not enough just to copy human behavior, it is necessary to adapt it to the affordances of the robotic embodiment.

When adding communication gestures, blinking and other nonverbal behaviors to our CGC we will undoubtedly find new unforeseen problems that will help us to further improve the quality of human-robot interactions. There are still many open questions on how and to which extent we should implement human-like conversational gaze in robots. In our approach we see gaze as part of a holistic nonverbal communication strategy that is intertwined with many other social behaviors. On the basis of human phylogenetic evolution and ontogenetic development, we argue that gaze behavior has a strong impact on the impressions we form about a robot and on the characteristics we attribute to it. Together with the necessity to create intuitive and pleasant interaction policies for robots, this requires a structured and inclusive approach to robot-human gaze interactions based on typical human-human behavior. The implementation of the architecture introduced in this short paper will allow a detailed exploration of the human perception of different aspects of conversational robotic gaze.

Acknowledgment

We would like to thank Ali Paikan for his help with the facial feature recognition module. This research was funded by the EU via the Marie-Sklodowska Curie Action Program (SICSAR - grant agreement no 627688).

References

[1] S. Baron-Cohen, S. Wheelwright, and T. Jolliffe. Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger Syndrome. Vis. Cogn., 4:311331, 1997.

[2] C. Breazeal and B. Scassellati. How to build robots that make friends and influence people. In Intelligent Robots and Systems, 1999. IROS'99. Proceedings. 1999 IEEE/RSJ International Conference on (Vol. 2, pp. 858-863). IEEE.

[3] C. Breazeal. Toward sociable robots. Robotics and autonomous systems, 42(3), pp.167-175, 2003.

[4] [24] F. Broz, H. Lehmann, C.L. Nehaniv, and K. Dautenhahn. Mutual gaze, personality, and familiarity: Dual eye-tracking during conversation. In RO-MAN, 2012 IEEE. IEEE, pp. 858864, 2012

[5] V. Corkum and C. Moore. Development of joint visual attention in infants. In C. Moore and P. Dunham, editors, Joint Attention: Its Origins and Role in Development. Erlbaum, Hillsdale, NJ, 1995.

[6] N.J. Emery. The eyes have it: the neuroethology, function and evolution of social gaze. Neuroscience & Biobehavioral Reviews 24, no. 6 (2000): 581-604.

[7] C.C. Ford, G. Bugmann and P. Culverhouse. Modeling the human blink: A computational model for use within humanrobot interaction. International Journal of Humanoid Robotics 10, 1350006. 2013.

[8] C. Kleinke. Gaze and eye contact: A research review. Psychological Bulletin, 100(1):78100, 1986.

[9] H. Kobayashi and S. Kohshima. Unique morphology of the human eye. Nature, 387:767768, 1997.

[10] H. Kobayashi and S. Kohshima. "Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye." Journal of Human Evolution, 40:419435, 2001.

[11] M. Mori. Bukimi no tani. Energy 7, 2235, 1970

[12] M. Mori, K. MacDorman and N. Kageki. The Uncanny Valley [From the Field]. IEEE Robotics & Automation Magazine 19, 98100. 2012. doi:10.1109/MRA.2012.2192811

[13] B. Mutlu, J. Forlizzi, and J. Hodgins. A storytelling robot: Modeling and evaluation of human-like gaze behavior. In Humanoids, pages 518523, 2006.

[14] B. Mutlu, T. Kanda, J. Forlizzi, J. Hodgins and H. Ishiguro. Conversational gaze mechanisms for humanlike robots. ACM Transactions on Interactive Intelligent Systems (TiiS), 1(2). 2012
[15] S. Nishio, H. Ishiguro and N. Hagita. Geminoid: Teleoperated android of an existing person (pp. 343-352). Vienna: INTECH Open Access Publisher. 2007.
[16] O. Palinko, F. Rea, G. Sandini, and A. Sciutti. Eye gaze tracking for a humanoid robot. In Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on (pp. 318-324). IEEE, 2015
[17] https://www.ald.softbankrobotics.com/en/cool-robots/pepper
[18] L. Schillingmann and Y. Nagai. Yet another gaze detector: An embodied calibration free system for the iCub robot, in: Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on. IEEE, pp. 813, 2015
[19] M. Tomasello, B. Hare, H. Lehmann, and J. Call. Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. Journal of Human Evolution, 52:314320, 2007.
[20] A.L. Vollmer, K.S. Lohan, K. Fischer, Y. Nagai, K. Pitsch, J. Fritsch, K.J. Rohlng and B. Wrede. People modify their tutoring behavior in robot-directed interaction for action learning. In DEVLRN 09: Proceedings of the 2009 IEEE 8th International Conference on Development and Learning, pages 16, Washington, DC, USA, 2009. IEEE Computer Society.