Semi-Autonomous Robotic Arm Reaching With Hybrid Gaze–Brain Machine Interface

Front. Neurorobot., 24 January 2020

https://doi.org/10.3389/fnbot.2019.00111

Assistive robotic systems have demonstrated high potential in enabling people with **upper limb physical disabilities**, such as traumatic spinal cord injuries (SCI), amyotrophic lateral sclerosis (ALS), and tetraplegic patients, to achieve greater independence and thereby increase quality of life (Vogel et al., 2015; Beckerle et al., 2017; Muelling et al., 2017).

For people with no residual movement or muscular activity, previous studies have focused on two key aspects for facilitating the interaction between patients and the assistive robot. One is the design of **human–robot interfaces (HRI)**. The other is the devising of **human–robot coordination strategies** tailored to the interface.

To provide HRI for individuals with severe upper extremity impairment, the brain signals or gaze signals have been largely explored through brain–machine interfaces (BMI) and gaze-trackers, respectively.

With the advent of invasive BMI technology, the **invasively recorded brain signals have facilitated successful manipulation of dexterous robotic arms** (Collinger et al., 2013; Wodlinger et al., 2015) due to their high bandwidth and signal-to-noise ratio (SNR). Nevertheless, the benefit of effective robotic arm control may be outweighed by **the medical risks** associated with the current electrode implantation techniques.

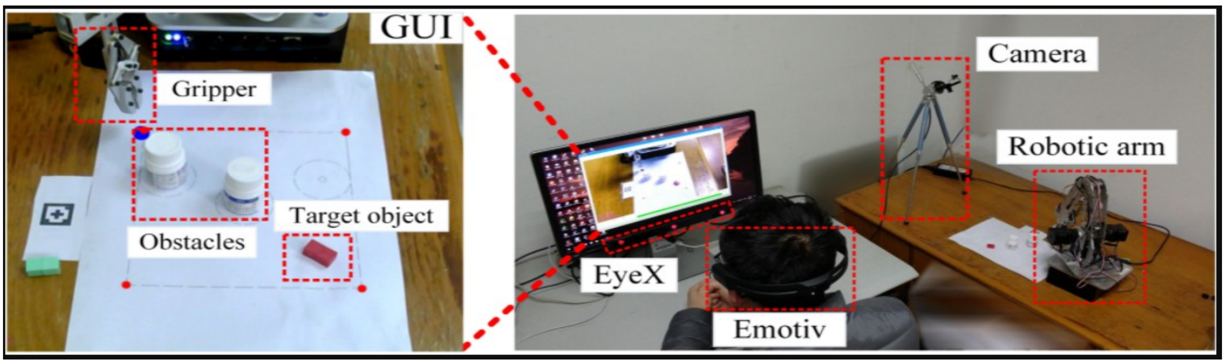
Non-invasive BMI, in particular the widely accepted **electroencephalogram (EEG)**-based BMI, provides a desirable alternative, and it is thus adopted in this study.

(1) Hybrid Gaze-BMI, which combines gaze tracking and BMI.

(2) Camera and Graphical User Interface (GUI), which provide the live scene of the robotic arm workspace for the normal and enhanced visual feedback as well as the coordinate transformations from camera coordinate system to the robot coordinate system for all the three stages.

(3) Shared Controller, which fuses the user commands from the hybrid gaze-BMI and the robot autonomy commands to form a new one, for directing the end-effector toward the target horizontally while avoiding obstacles in stage 2;

(4) Actuated System and Control, where the resulting end-effector commands are converted into reaching and grasping motions with a 5-Dof robotic arm.



Studies on Grounding with Gaze and Pointing Gestures in Human-Robot-Interaction

- 10.1007/978-3-642-34103-8_38
- Conference: International Conference on Social Robotics, June 2013

Field Task. The Field Task is analog to the Piece Task as can be seen in figure 1. Except that, in this case, the robot points to the position of the puzzle field, where the participant has to drag and drop the previously selected piece. The verbal instructions describe the target position in reference to a nearby piece that is already on the field, such as "Now place this piece left to the black piece with the red circle in it".

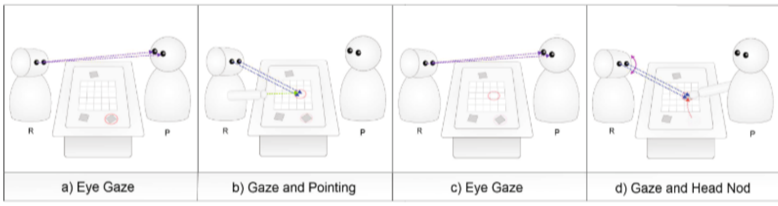


Fig. 1. Structure of a Field Task using speech, gaze and pointing gestures; Robot (R) and Participant (P)

Spatial References with Gaze and Pointing in Shared Space of Humans and Robots

Patrick Renner1 , Thies Pfeiffer2 , and Ipke Wachsmuth1

https://core.ac.uk/download/pdf/211813368.pdf

A **saccade** is a rapid, conjugate, eye movement that shifts the center of gaze from one part of the visual field to another.

When conducting manual interactions, we need to have a spatial representation of the target and thus gaze is quite naturally linked to hand movements, such as pointing gestures. **Once a spatial representation has been built, gaze might no longer be needed to control movements towards a target**, but it is still relevant for fine controlled end positioning [1].

In our own work, we found that pointing directions can be determined most precisely when considering t**he dominant eye aiming over the pointing finger tip** [21]. There is evidence for a temporal coordination between gaze and gestures. For example, Prablanc et al. [24] found that hand movements are initiated about 100 ms after the first saccade to the target in a pointing task.

In their study, participants had to point at a first target and then as soon as possible look at a second one, which was presented a bit later. Their participants **were not able to fixate the second target while still reaching towards the firs**t and they explain that by a neural inhibition of saccades during manual pointing

Fig. 5. SMI Eye Tracking Glasses, a binocular mobile eye tracker.

Fig. 6. Gloves with tracking markers attached.

1. Given the onset of a hand movement, the target of a pointing gesture can be predicted by looking at the location of preceding fixations.
2. How accurate and how precise can the target area be predicted?
3. How large is the advantage in timing that can be gained?
4. What influence does the distance of the target have on the performance of the pointing gesture?

Robot reading human gaze: Why eye tracking is better than head tracking for human-robot collaboration

- 10.1109/IROS.2016.7759741
- Conference: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 2016

Figure 2. Example output of our eye tacking system. White lines from eyes – eye gaze, black line from top of the nose – head gaze.

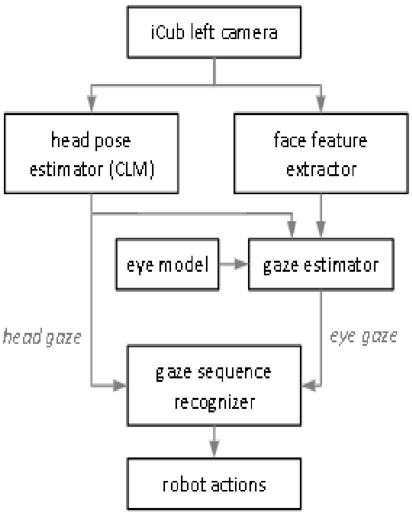


Figure 1. Eye tracking system diagram.

We implemented a model-based, visual light, monocular, calibration-free, remote gaze tracking algorithm, using existing head pose and face feature tracking algorithms (see Figure 1). Head pose was calculated using the Constrained Local Models (CLM) approach implemented by Baltrusaitis [12]. It provided us with the head orientation that is directly used in the *head gaze* experimental condition as well as in calculating the eye gaze. Face features were found using the approach described in [13] and implemented by King in [14]. This provided us with robust tracking of locations like the corners of the eyes and mouth. Once the eye region was located we used averaging methods to find the center of the darkest area of the eye, which approximates the center of the pupil, due to the light color of the sclera. Once we found the locations of the eye corners, pupil center and head orientation, we applied these points to an eye model with the goal of calculating gaze angles. The parameters of this model were estimated in a least squares approach on the Columbia gaze dataset. The approach was adopted from [15]. This allowed us to create an eye model for a "generic subject", thus eliminating calibration for each new user. We verified the newly obtained eye gaze model by assembling our own gaze dataset using the iCub's eyes. We found that the mean absolute error in horizontal gaze was around 5 degrees, while for the vertical gaze it was 9 degrees. The larger error in vertical gaze detection stems from the fact that when people look down their upper eyelids covered most of their eyes, which in turn causes imprecisions in detecting the eye corners and pupil center. None the less the robot equipped with this system can reliably understand which objects are gazed upon by its interacting partner [11].

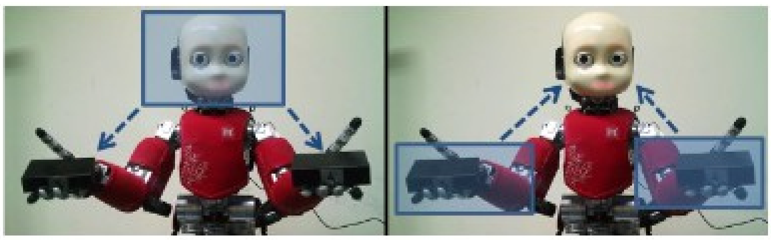


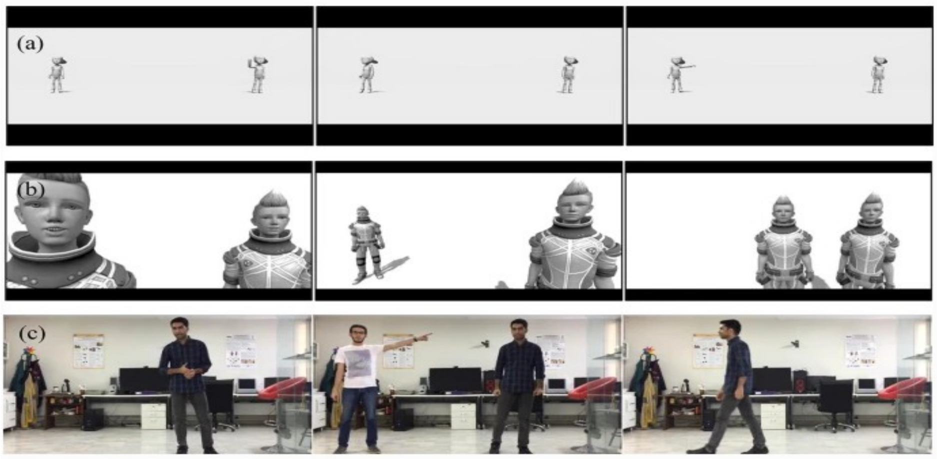
Figure 4. Gaze actions triggering robot reaction: a) gazing at face then one of the hands and b) gazing at a hand then at face.

When asked how the robot knew which object to hand over, four out of ten subjects didn't mention gaze in any way, four thought it was some combination of voice commands, gaze and gestures, while two were correct in thinking that it based its decision on gaze only.

Aliasghari, P., Taheri, A., Meghdari, A. *et al.*
Implementing a gaze control system on a social robot in multi-person interactions

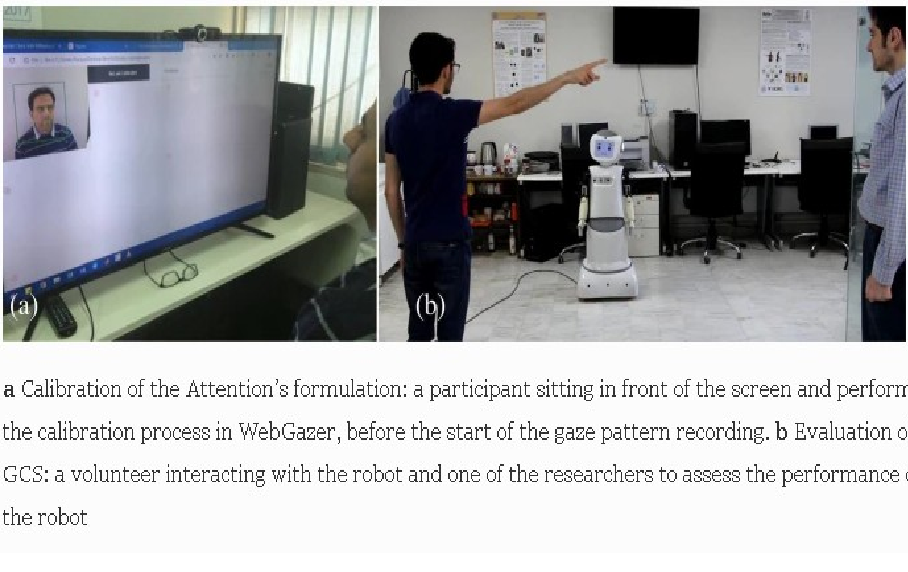
SN Appl. Sci. **2,** 1135 (2020).
https://doi.org/10.1007/s42452-020-2911-0

In this paper, implementing a social gaze control system suitable for multi-person interactions with a RASA social robot is discussed. This system takes some important verbal and non-verbal social cues into account, and at each moment enables the robot to **decide socially at which human it should direct its gaze**. The algorithm for the target selection has been enhanced, compared to past studies, by quantitating the effects of distance and orientation on grabbing humans' attention in addition to the inherent importance of each cue in communications based on the gaze behavior of a group of human participants. After this was completed, another group of volunteers were employed to examining the performance of the RASA robot equipped with this system. Their average gaze pattern was compared with the targets selected by the robot in a real situation, and their opinions on the sociability and intelligence of the system were recorded. We indicated that **the gaze generated by the robotic system matched the average gaze pattern of the participants 76.9% in an 80-s real-life scenario**. Moreover, the results of the questionnaire showed us that ~ 90% of the subjects felt that at times RASA was really looked at them with a quite high average score of 4.33 out of 5.



a Screen-shots from the video used to evaluate social cues to attracting humans' attention. Speaking is compared to the other cues: (from left to right) hand-waving, engagement to the viewer, and pointing. **b** From left to right, the subscenes number 1, 2, and 4 in the animation played to measure the effect of proxemics and field of view. **c** Screen-shots from the real-life video recorded to evaluate the performance of the robotic system in gaze-shifting

Fig. 6



a Calibration of the Attention's formulation: a participant sitting in front of the screen and performing the calibration process in WebGazer, before the start of the gaze pattern recording. **b** Evaluation of the GCS: a volunteer interacting with the robot and one of the researchers to assess the performance of the robot