Published in final edited form as:

Proc SPIE Int Soc Opt Eng. 2016 May ; 9859: . doi:10.1117/12.2224382.

Human-like object tracking and gaze estimation with PKD android

Indika B. Wijayasinghe^a, Haylie L. Miller^b, Sumit K Das^a, Nicoleta L. Bugnariu^b, and Dan O. Popa^a

^aUniversity of Louisville, Louisville, KY 40292, USA

^bUniversity of North Texas Health Science Center, Fort Worth, TX 76107, USA

Abstract

As the use of robots increases for tasks that require human-robot interactions, it is vital that robots exhibit and understand human-like cues for effective communication. In this paper, we describe the implementation of object tracking capability on Philip K. Dick (PKD) android and a gaze tracking algorithm, both of which further robot capabilities with regard to human communication. PKD's ability to track objects with human-like head postures is achieved with visual feedback from a Kinect system and an eye camera. The goal of object tracking with human-like gestures is twofold : to facilitate better human-robot interactions and to enable PKD as a human gaze emulator for future studies. The gaze tracking system employs a mobile eye tracking system (ETG; SensoMotoric Instruments) and a motion capture system (Cortex; Motion Analysis Corp.) for tracking the head orientations. Objects to be tracked are displayed by a virtual reality system, the Computer Assisted Rehabilitation Environment (CAREN; MotekForce Link). The gaze tracking algorithm converts eye tracking data and head orientations to gaze information facilitating two objectives: to evaluate the performance of the object tracking system for PKD and to use the gaze information to predict the intentions of the user, enabling the robot to understand physical cues by humans.

Keywords

Gaze estimation; Object tracking; Humanoid robot

1. Introduction

Human gaze is an important cue during social interaction, and plays a significant role in human-robot interaction as well. For instance, gaze tracking can be used to anticipate human intent prior to and during physical contact with a robot. It can also be used to program a robot to maintain eye contact during conversational dialogue with humans. The realism of a robot's gaze is especially relevant for humanoid robots, from which an interactive human user anticipates human-like cues and gestures.

Further author information: (Send correspondence to I. B. Wijayasinghe): I. B. Wijayasinghe: indika.wijayasinghe@louisville.edu, H. L. Miller: Haylie.Miller@unthsc.edu, S. K. Das: sumitkumar.das@louisville.edu, N. L. Bugnariu: nicoleta.bugnariu@unthsc.edu, D. O. Popa: dan.popa@louisville.edu.

The history of humanoid social robots can be traced back to the early 1970's, when WABOT 1 [1] was developed by Waseda University, Japan and succeeded by WABOT 2 [2] in the early 1980's. Examples of early work to create humanoid social robots include WABOT 1 [1] and 2 [2], Greenman [3], Saika [4], ASIMO [5], Nexi [6], P1, P2[7][8], that have been constructed with a head and a neck system with cameras in place of eyes. More recently, a new generation of robot androids such as Saya [9], Albert Hubo [10] and Philip K. Dick (PKD) [11] were designed with a realistic appearance and a construction capable of human-like cues and gestures with the head-eye system. These robots have movable eyes with respect to the head and are also complemented by the capability of human-like facial expressions. In the implementation of effective communication capabilities with these androids, gaze estimation of both the human and the robot are significant factors. This paper discusses gaze tracking of the PKD android and human subjects currently under development in our labs.

Virtual reality (VR) technology offers a unique opportunity to stimulate and measure the reaction of humans to a combination of external stimuli, approximating real-world demands in a repeatable and controlled task. In related research, we collected full-body motion data from typically-developing children and children with Autism Spectrum Disorder (ASD) during a variety of social and nonsocial visuomotor tasks using one such VR motion capture system [12][13]. During the course of this work, we noticed that many tasks require the use of visual information that also include head pose and eye gaze to guide appropriate motor responses. There are numerous eye tracking systems commercially available today, including Tobii glasses [14], Applied Science Laboratories's Mobile Eye XG [15], and SensoMotoric Instruments's (SMI) Eye Tracking Glasses (ETG) [16] which are wearable systems employing stereo cameras. Other systems such as Tobii EyeX Controller [17] are non-wearable and are intended for peripheral control. The options for body motion capture include systems from OptiTrack [18], Vicon [19], and MotionAnalysis [20], which are marker based, and software such as iPi Soft software [21] which uses markerless motion capture techniques.

In this paper, we describe a mixed system composed of an eye tracker, a body motion capture system, and an immersive VR environment that have been integrated to allow measurements of human as well as Android robot gazes. The ETG mobile eye-tracking system (SensoMotoric Instruments; Boston, MA) allows computing of visual events such as blinks, saccadic movement, and fixation on central and peripheral targets concurrent with full-body kinematic data from a motion capture system (Cortex, MotionAnalysis Corporation). A virtual environment with objects to be tracked was generated using the CAREN system (Motek ForceLink)[22] as shown in Figure 1.

We describe a methodology for integrating these three systems—ETG, Cortex and CAREN—in a way that allows for quantitative analysis of gaze performance. The algorithms developed transform gaze vectors from the eye tracking system and object coordinates in the virtual reality environment to a global coordinate space used by the motion capture system. This makes it possible to precisely identify how accurate a participant's pursuit and fixation eye movements are for a variety of targets. The ability to use these systems with both human subjects and PKD provides a means of comparing the gaze performance of the android with

that of humans. These algorithms may advance the field of robotics by enhancing our ability to program more human-like movements and facilitate greater realism in human-robot interaction. In addition, by combining performance data from social movement tasks (imitative gesturing) with data from nonsocial movement, we can also quantitatively study the clinical significance of eye movement differences between individuals with autism and typically-developing persons.

The contents of the paper are organized as follows: Section 2 contains the descriptions of different sub systems used for the gaze estimation; Section 3 describes in detail the algorithms developed; Section 4 presents the results; Section 5 provides a discussion and conclusion of the results.

2. System Description

The gaze tracking system employed in this work consists of three technologies: Eye Tracking Glasses (ETG), motion capture system and Computer Assisted Rehabilitation Environment (CAREN). These are briefly described in this section.

2.1 Eye Tracking Glasses

ETG is a mobile eye tracking system that employs two cameras and image processing techniques to obtain the binocular gaze vector information. Two cameras—one for each eye—are embedded in the inner rim of the glasses. The cameras track pupillary movement and a proprietary algorithm uses data from the eye cameras and the scene camera to convert this motion to a 3D gaze vector. These data are sampled at a rate of 60 Hz and recorded on a customized Samsung Galaxy S4 that runs the iView software for the system. Prior to usage, the ETG system was calibrated with either a one-point or three-point procedure provided by the manufacturer.

The ETG system includes a third camera embedded above the outer nosebridge of the glasses, facing forward, which records the scene in front of the subject. This video can be merged with the detected gaze, which is marked by the circle on the video, to visualize the gaze of the subject in context of the scene being observed.

2.2 Motion Capture System: Cortex and Eagle-4 IR Cameras

Cortex and 12 Eagle-4 infrared cameras (IR) make up the marker-based real-time motion capture system used in our work. PKD Android was only marked with 18 upper body markers due to its normal sitting posture. Markers were tracked in 3D space using the Eagle-4 IR cameras (see Figure 3 for camera placement), and marker position data were analyzed with the Cortex software.

Cortex records the 3D position of each marker at a sampling rate of 120 Hz. A pulse signal is sent from the CAREN system, to sync the eyetracker and body movement recordings in time. Cortex provides a 3D simulation of the captured markers streaming in real-time, as well as additional tools for motion analysis. The coordinate frame used by the motion capture system is the base coordinate frame for the work described in this paper.

2.3 CAREN System

CAREN is a virtual reality (VR) system developed for treatment and rehabilitation purposes. In the work described in this paper, we used the CAREN system to generate objects in a virtual environment that was projected on a 180° wrap-around screen (Figure 3). The CAREN coordinate system for the virtual environment is in the same frame of reference as the motion capture system. For the tasks described here, the height of projections was adjusted based on the marker at the back of the neck ('C7' marker), captured by the motion capture system before each task, so that the projected images were at the same vertical position relative to each subject's head.

CAREN is a versatile system that can generate many different virtual environments and games, which can be controlled by multi-sensory inputs. These include static or moving balls, arcade-style shooting games, and navigation tasks. For the work described in this paper, only the moving ball scenario was used.

2.4 Philip K. Dick Android

PKD android (Figure 5), was created in the image of the science fiction writer Philip K. Dick by Hanson Robotics. Its skin is made out of an artificial skin material, which gives it a realistic appearance and feel of human skin. It has 28 degrees of freedom: 7 controlling the neck-eye system and 21 driving facial expressions. Of the 7 actuators that control the neck-eye system, 4 actuators control the roll-pitch-yaw of head and 3 actuators control the pan-and-tilt of the eyes. The neck controls have 2 actuators for the roll of the head. The eyes have one shared tilt control and individual pan controls. For this work, 6 of these neck-eye actuators were used to implement head-eye control for object tracking. In order to keep the roll-pitch-yaw to be about approximately the same point, only the upper neck roll actuator was used. The actuators were calibrated and the range of each actuator was measured to stay within the mechanical limits. The right eye of PKD has a built-in camera which provides feedback to measure tracking accuracy.

The Microsoft Kinect, which is a RGB-D based camera, provides an additional image-stream and a depth-stream for PKD. It comprises an IR emitter and a detector, a RGB camera, microphones and tilt motors. The field of view of Kinect is 43° in the vertical direction and 57° in the horizontal direction. In this work, the Kinect was used (as shown in Figure 5) as an external camera to increase the field of view of PKD so that the Kinect could detect the object and pass the coordinates to the controller when the object was outside the field of view of the eye camera. Microsoft Kinect SDK was used to obtain the image and depth streams at a rate of 30 fps with a resolution of 640×480 pixels. The z-coordinate (depth) of the object was calculated by the SDK and the x and y coordinates were calculated using a pinhole camera model and the pixel position of the object.

3 Algorithms for Gaze Estimation

3.1 Description of coordinate frames

While the CAREN and motion capture systems operate in the same coordinate frame, the ETG system uses its own coordinate frame to localize objects. Figure 7 shows the coordinate

frames used by Cortex and CAREN whereas Figure 8 describes the coordinate frame used by the ETG system. Cortex coordinate system is where the marker coordinates are measured and in this work was considered to be the base coordinate system. It has to be noted that the CAREN coordinate system, where the virtual objects live, is a virtual 3D coordinate system which is projected on to a 2D screen. Although the axes of CAREN coordinate system and Cortex coordinate system approximately align, for a quantitative analysis of the gaze performance, a transformation was calculated due to the distortions that arise in the process of projecting the virtual coordinate system on to the screen.

The ETG coordinate system is attached to the frame of the glasses, with the origin located approximately at the center of the right eye. We only consider the right eye in this paper, due to the fact that PKD only has a right-eye camera. However, for the left eye gaze, the equivalent would be to fix the origin of the coordinate system at the approximate center of the left eye.

3.2 Object Tracking with Philip K. Dick Humanoid Robot

The object tracking algorithm implemented on PKD is based on the work described in [23]. The controller directs PKD's gaze by computing the redundant head and eye motion required to fixate PKD's eyes onto a target. Our controller is based on a hypothetical potential function and a friction term, and can be identified with threshold control for muscle movements studied in literature [24]. This controller is abstract enough to be used for both the head and the eyes without significant modifications. It also has the advantage of being able to easily implement constraints, so that it is possible to make PKD obey the same orientation constraints obeyed by humans; namely Donder's [25] and Listing's [26] constraints. Because PKD's eyes only have two degrees of freedom (see Section 2.4), prohibiting Listings' constraint, only Donder's constraint was implemented.

3.3 Gaze Tracking with ETG Mobile Eye Tracking System

In order to assess the accuracy of tracking, the three previously-described systems had to be integrated so that the data obtained from each was transformed to the same base coordinate system. In this work, the base coordinate system was the one used by the motion capture system. This transformation has two steps: converting ETG gaze vectors to the base coordinate system and converting CAREN coordinates to the base coordinate system.

3.3.1 Converting ETG gaze vectors to Cortex coordinate system—The conversion of ETG gaze vectors to the base coordinate system could be reduced to a standard problem of coordinate transformation with one unknown parameter. However, this requires identifying the rotation and translation of the ETG coordinate system with respect to the base coordinate system. This can be done by calculating $(\hat{i}_g, \hat{j}_g, \hat{k}_g)$, which are the unit vectors along x, y and z respectively, using the coordinates of the markers m_f , m_r and m_l given by the motion capture system. Assignment of $(\hat{i}_g, \hat{j}_g, \hat{k}_g)$ were as given in Figure 8. In this context, the ETG coordinate system and the base coordinate system are represented by the subscripts g and b respectively.

The unknown in this calculation is the angle that the plane created by m_f , m_r and m_l makes with the straight-ahead direction. The angle θ was measured to be approximately 60° . The calculation of $(\hat{i}_g, \hat{j}_g, \hat{k}_g)$ from the marker coordinates are given in (1)-(5).

$$\hat{i}_g = \frac{h_h}{\|h_h\|_2} \quad (1)$$

$$\hat{k}_g = R_{(-60^\circ, \hat{i}_g)} \left(h_f - \frac{h_h \cdot h_f}{\|h_h\|_2} \frac{h_h}{\|h_h\|_2} \right) \quad (2)$$

$$\hat{j}_g = \hat{k}_g \times \hat{i}_g \quad (3)$$

where

$$h_h = m_l - m_r \quad (4)$$

$$h_f = m_f - m_l \quad (5)$$

and $R(-60^\circ, \hat{i}_g)$ represents a rotation about $\hat{i}_g$ by an angle 60° clockwise. Then the rotational transformation can be given as

$$R_g^b = \begin{pmatrix} \hat{i}_g & \hat{j}_g & \hat{k}_g \end{pmatrix} \quad (6)$$

and the translation $P_{g, \text{origin}}^b$ is the vector from the origin to the center of the eye in the base coordinate system. It can be seen that for the left and the right eyes, $P_{g, \text{origin}}^b$ changes while R_g^b remains the same. The transformation matrix can then be written as

$$T_g^b = \begin{pmatrix} R_g^b & P_{g, \text{origin}}^b \\ 0 & 1 \end{pmatrix}. \quad (7)$$

If v^g and v^b represent the eye gaze in ETG coordinate system and the base coordinate system respectively, then

$$\begin{pmatrix} v^b \\ 1 \end{pmatrix} = T_g^b \begin{pmatrix} v^g \\ 1 \end{pmatrix} \quad (8)$$

3.3.2 Converting CAREN coordinates to Cortex coordinate system—

Transforming the CAREN coordinates to the base coordinate system presented a greater challenge due to potential distortions or unknown transformations applied during projection of the virtual environment on to the screen. Since the axes of the CAREN system and the base coordinate system were designed to align, and appeared to align approximately, several different models that slightly alter this alignment were attempted. As such, the following models were evaluated as potential transformations. (P^c - object coordinates in CAREN coordinate system, P^b -object coordinates in base coordinate system, λ - scaling factor). In the following models, λ , R_c^b and $P_{c,origin}^b$ were considered to be unknowns.

- Model I: Translation and scaling:

$$P^b = \lambda P^c + P_{c,origin}^b \quad (9)$$

- Model II: Rotation, translation and scaling:

$$P^b = \lambda R_c^b P^c + P_{c,origin}^b \quad (10)$$

- Model III: Perspective transformation, translation and scaling:

$$P^b = \lambda \frac{P^c}{P_z^c} + P_{c,origin}^b \quad (11)$$

- Model IV: Perspective transformation, rotation, translation and scaling:

$$P^b = \lambda R_c^b \frac{P^c}{P_z^c} + P_{c,origin}^b \quad (12)$$

For the human subject data, evaluation of the above models were performed with a set of 6 calibration points. These calibration points were selected such that the eye gaze is on the object at the corresponding time as seen in the ETG video (Figure 10). The performance of the model was measured by the RMS error of the angle between the object vector from the eye and the eye gaze vector in degrees as given by (11).

$$\text{RMSE}_{\text{RMSE}} = \frac{180}{\pi} \sqrt{\sum_{i=1}^6 \arccos \left(\frac{(P^b - P_{g,\text{origin}}^b) \cdot v^b}{\|P^b - P_{g,\text{origin}}^b\|_2 \|v^b\|_2} \right)^2} \quad (13)$$

However, due to the unknowns λ , R_c^b and $P_{c,\text{origin}}^b$, each model had to be evaluated under an optimization algorithm that minimizes (11). For this work, only a rotation about the x-axis (θ) and a translation in the y-axis (d) were considered. Therefore, for this work, the unknowns were λ , θ and d. With these parameters, (11) can be written as $C(\lambda, \theta, d) = \text{RMSE}$ and a gradient descent algorithm was used for this optimization purpose.

Let $\Theta = [\lambda \ \theta \ d]^T$ and $\Theta(n)$ be the n^{th} iteration of Θ . Then the gradient decent algorithm can be given as an iterative calculation as

$$\Theta(n+1) = \Theta(n) - \gamma \nabla C(\Theta) \quad (14)$$

where ∇ represents the gradient with respect to Θ and γ being a parameter that governs the descent rate. When converged, Θ holds the parameter values that minimizes (11). The results of this process are given in Table 1.

Model II was then used for the same type of experiment with PKD android. Except for the parameter d, the other two parameters were unchanged from Table 1 values. Parameter d was selected differently from that for the human subject data due to the fact that the height of the projection from the CANREN system changes with the height of the marker at the back of the neck named 'C7'. Hence d is expected to change for different subjects. Also, unlike for the human data, the x-values of PKD required a scaling by a factor of 1.4.

4 Results

The gradient descent algorithm applied on the cost function (11) with P^b given by the corresponding model out of (7)-(10) produced the results given in Table 1.

Figure 11 plots the x, y and z components of the object vector with respect to the right eye (red) and the gaze vector (blue) for Model II for a human subject. Both vectors were normalized to be unit vectors for this plot.

For PKD data, adjustment for height based on the 'C7' marker (other two parameters unchanged from Table 1) obtained d = -500. The RMSE for the 4 calibration points for PKD data was 2.82 degrees. Figure 12 plots the x, y, z components of the object vector with respect to the right eye (red) and the gaze vector (blue) for Model II for PKD. Both vectors were normalized to be unit vectors for this plot. (This is analogous to Figure 11 for human data).

Conclusion

In this paper we described the implementation of an object tracking algorithm on PKD android and the implementation of an algorithm that quantifies the tracking performance by estimating the gaze of the human and robot.

The gaze estimation algorithm was developed to integrate the three component systems—ETG, Cortex, and CAREN—that make up the gaze tracking system. The integration of ETG and motion capture systems could be simplified down to a standard problem of coordinate transformation. However, integration of CAREN system required a more complex model due its virtual existence. Four types of models considered here were subjected to an optimization algorithm to find the best parameter values that minimize the root mean square error which is the cost function.

For this work, only a rotation about the x-axis and a translation in the y-direction were used together with a scaling parameter. This was due to the fact that rotations about the other two axes and translations in the other two directions appeared to be negligible. However, the same optimization algorithm can be modified to find a transformation with all three rotations and translations.

The results from the optimization algorithm implies that the best fit model is one that relates the virtual coordinate system of CAREN to the base coordinate system by a rotation about the x-axis, a translation and a scaling. Except for the different value of d and the x-coordinate scaling for PKD, the same parameters obtained for the human subject were used for PKD. The change in d value is expected as the height of the projection changes depending on the marker at the back of the neck. However, the scaling in x was not expected. This could be related to how the ETG system was calibrated. For the human subject, the 3-point calibration method on ETG software was used whereas for PKD, the 1-point calibration had to be used due to failed attempts to use the 3-point calibration. This has been one of the challenges encountered with using the ETG system on PKD.

Although PKD has a very realistic appearance to a human, which is apparent by the ability to use the ETG system on it, the imperfections resulted in several difficulties in obtaining valid data. Gloss of the PKD eye surface is different from those of humans which could be one reason for the eye tracker to have difficulties. One other major factor could be the lack of proper binocular vision on PKD. With only one eye camera on the right eye, it was not possible to implement a feedback controller on the left eye. Therefore the left eye lacked an independent controller which resulted in lack of realistic binocular control. This could be one major factor that could confuse the gaze estimation algorithms in the ETG software. The results of PKD object tracking are also heavily dependent on the effectiveness of the object recognition algorithm. Imperfect object recognition has been a major factor for the noise in PKD data.

Despite these difficulties, it can be seen from the results that the algorithm performs reasonably well in quantifying the object tracking performance of both the human subject and PKD android. In the future, the resulting metrology system will be used to

comparatively study human and robot object tracking performance and also to develop human intent models enabling improved human robots interactions.

Acknowledgments

This research was supported by National Science Foundation (NRI-1208623; SMA-1514495), the National Institutes of Health–National Center for Advancing Translational Sciences (KL2TR001103), and IIP-1534124. We wish to thank Hanson Robotics staff for their assistance with PKD setup, and Speed School of Engineering at the University of Louisville for additional support.

References

1. Kato I, Ohteru S, Kobayashi H, Shirai K, Uchiyama A. Information-Power Machine with Senses and Limbs (WABOT-1). First CISM - IFToMM Symp Theory Pract Robot Manip. 1974; 1:11–24.
2. Sugano S, Tanaka Y, Ohoka T, Kato I. Autonomic Limb Control of the Information Processing Robot - Movement control system of robot musician 'WABOT-2'. J Robot Soc Japan. 1985; 3(no. 4):81–94.
3. Rosheim ME. Robot Evolution: The Development of Anthrobotics. 1994
4. Konno A, Nagashima K, Furukawa R, Nishiwaki K, Noda T, Inaba M, Inoue H. Development of a humanoid robot Saika. Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robot and Systems Innovative Robotics for Real-World Applications IROS '97. 2:805–810.
5. Sakagami Y, Watanabe R, Aoyama C, Matsunaga S, Higaki N, Fujimura K. The intelligent asimo: System overview and integration. IEEE IROS, Switz. 2002:2478–2483.
6. Siegel M, Breazeal C, Norton MI. Persuasive Robotics : the influence of robot gender on human behavior Citation Accessed Citable Link Detailed Terms Persuasive Robotics : The Influence of Robot Gender on Human Behavior. IEEE/RSJ Int Conf Intell Robot Syst. 2009:2563–2568.
7. Ceccarelli M. Technology Developments: the Role of Mechanism and Machine Science and IFToMM. 2011
8. Hirai, K., Hirose, M., Haikawa, Y., Takenaka, T. The Development of Honda Humanoid Robot; Proc 1998 IEEE International Conference on Robotics and Automation; 1998. p. 1321-1326.
9. Kobayashi H. Realization of realistic and rich facial expressions by face robot. IEEE IROS, Las Vegas, Nevada, USA. 2003:1123–1128.
10. Oh, JH., Hanson, D., Kim, WS., Han, Y., Kim, JY., Park, IW. Design of Android type Humanoid Robot Albert HUBO; Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on; 2006. p. 1428-1433.
11. Hanson D. Exploring the aesthetic range for humanoid robots. ICCS CogSci. 2006:39–42.
12. Singhal R, Bugnariu K, Patterson N. Methodological consideration for application of Multi-Scale Entropy on Center of Pressure data from children with Autism Spectrum Disorder-A Pilot Study. South Central American Society of Biomechanics Regional Meeting. 2014
13. Miller, HL., Patterson, R., Popa, D., Garver, C., de Weerd, C., Bugnariu, N. Quantifying imitative behavior deficits in children with Autism Spectrum Disorder; International Meeting for Autism Research; 2014.
14. Fehring B, Bulling A, Krüger A. Analysing the potential of adapting head-mounted eye tracker calibration to a new user. Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA'. 2012; 12:245–248.
15. Hurwitz D, Monsere C, Marnell P, Paulsen K. Three- or Four-Section Displays for Permissive Left Turns? Some Evidence from a Simulator-Based Analysis of Driver Performance. Transp Res Rec J Transp Res Board. 2014; 2463:1–9.
16. Paletta L, Neuschmied H, Schwarz M, Lodron G, Pszeida M, Ladstätter S, Luley P. Smartphone eye tracking toolbox: accurate gaze recovery on mobile displays. Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA'. 2014; 14:367–368.
17. Stenström A. Clicking using the eyes, a machine learning approach. Linköping University. 2015

18. Millard, AG., Hilder, JA., Timmis, J., Winfield, AFT. A Low-cost Real-time Tracking Infrastructure for Ground-based Robot Swarms; Swarm Intelligence: 9th International Conference (ANTS 2014); 2014.
19. Kapur, A., Tzanetakis, G., Virji-Babul, N., Wang, G., Cook, PR. A Framework for Sonification of Vicon Motion Capture Data; Proc of the 8th Int Conference on Digital Audio Effects (DAFX-05); 2005.
20. Ranatunga, I., Torres, NA., Patterson, R., Bugnariu, N., Stevenson, M., Popa, D. RoDiCA: a Human-Robot Interaction System for Treatment of Childhood Autism Spectrum Disorders; Proceedings of the 5th International Conference on PErvasive Technologies Related to Assistive Environments; 2012.
21. Gabbasov B, Danilov I, Afanasyev I, Magid E. Toward a human-like biped robot gait: Biomechanical analysis of human locomotion recorded by Kinect-based Motion Capture system. 2015 10th International Symposium on Mechatronics and its Applications (ISMA). 2015:1–6.
22. Collins JD, Markham A, Service K, Reini S, Wolf E, Sessoms P. A systematic literature review of the use and effectiveness of the Computer Assisted Rehabilitation Environment for research and rehabilitation as it relates to the wounded warrior. Work. 2015; 50:121–129. [PubMed: 25167904]
23. Wijayasinghe I, Aulisa E, Büttner U, Ghosh BK, Glasauer S, Kremmyda O. Potential and Optimal Target Fixating Control of the Human Head/Eye Complex. IEEE TCST. 2014; 23(no. 2):796–804.
24. Feldman AG, Levin MF. The Equilibrium-Point Hypothesis ? Past, Present and Future. Prog Mot Control. 2009:699–726.
25. Ceylan M, Henriques DYP, Tweed DB, Crawford JD. Task-Dependent Constraints in Motor Control: Pinhole Goggles Make the Head Move Like an Eye. J Neurosci. 2000; 20(7):2719–2730. [PubMed: 10729353]
26. Tweed D, Vilis T. Geometric Relations of Eye Position and Velocity Vectors During Saccades. Vis Res. 1990; 30:111–127. [PubMed: 2321357]

Figure 1.
Participant completing a nonsocial visuomotor integration task while wearing markers and the ETG system at UNT Health Science Center.

Figure 2. Eye Tracking Glasses (ETG)

Figure 3. Eagle-4 and CANREN system setup

Figure 4. 3D point cloud of markers worn by the subject shown in Cortex software

Figure 5. Experimental setup with a Kinect placed behind Phillip K. Dick Android head, covered in artificial skin. Once eye gaze models are developed, this setup can be used for coarse gaze estimation and intent modeling during pHRI.

Figure 6. Block diagram of PKD and Kinect system

Figure 7. Coordinate systems used by Cortex and CAREN

Figure 8. ETG coordinate system

Figure 9. Markers placed on the eye glasses

Figure 10. Video frame from the ETG system showing object (white ball) and the PKD gaze (black ring)

Figure 11.

Components of the eye gaze vector (v^b) and the normalized object vector with respect to the eye (P^b) for the human subject. The object vector was calculated using Model II in Table 1. Red - object vector components, blue - gaze vector components.

Figure 12.

Components of the eye gaze vector (v^b) and the normalized object vector with respect to the eye (P^b) for PKD. The object vector was calculated using Model II in Table 1. Red - object vector components, blue - gaze vector components.

Table 1**Results of model fitting**

Model #	λ	θ	d	RMSE
I	322	0	-690	4.67
II	680	-15.6	-720	1.33
III	1499	0	-803	4.11
IV	8000	-24.5	-700	3.80