

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/249968908>

Detecting eye contact using wearable eye-tracking glasses

Conference Paper · January 2012

DOI: 10.1145/2370216.2370368

CITATIONS

17

READS

357

7 authors, including:



[Yin Li](#)

Georgia Institute of Technology

17 PUBLICATIONS 266 CITATIONS

[SEE PROFILE](#)



[Agata Rozga](#)

Georgia Institute of Technology

62 PUBLICATIONS 704 CITATIONS

[SEE PROFILE](#)



[Gregory Abowd](#)

Georgia Institute of Technology

320 PUBLICATIONS 22,169 CITATIONS

[SEE PROFILE](#)

Detecting Eye Contact using Wearable Eye-Tracking Glasses

Zhefan Ye Yin Li Alireza Fathi Yi Han
Agata Rozga Gregory D. Abowd James M. Rehg
Georgia Institute of Technology
{frankye, yli440, afathi3, yihan, agata, abowd, rehg}@gatech.edu

ABSTRACT

We describe a system for detecting moments of eye contact between an adult and a child, based on a single pair of gaze-tracking glasses which are worn by the adult. Our method utilizes commercial gaze tracking technology to determine the adult's point of gaze, and combines this with computer vision analysis of video of the child's face to determine their gaze direction. Eye contact is then detected as the event of simultaneous, mutual looking at faces by the dyad. We report encouraging findings from an initial implementation and evaluation of this approach.

Author Keywords

Wearable Eye-Tracking, Eye Contact, Developmental Disorders, Autism, Attention

ACM Classification Keywords

H.5.2 Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

INTRODUCTION

In this paper, we describe a system for detecting eye contacts between two individuals, which is based on the use of a single wearable gaze-tracking system. Eye contact is an important aspect of face-to-face social interactions, and measurements of eye contacts are important in a variety of contexts. In particular, atypical patterns of gaze and eye contacts have been identified as potential early signs of autism, and they remain important behaviors to measure in tracking the social development of young children. Our focus in this paper is on the detection of eye contact events between an adult and a child. These gaze events arise frequently in clinical settings, for example when a child is being examined by a clinician or is receiving therapy. Social interactions in naturalistic settings, such as classrooms or homes, are another source of important face-to-face interactions between a child and their teacher or care-giver.

In spite of the developmental importance of gaze behavior in general, and eye contact in particular, there currently exist no good methods for collecting large-scale measurements of these behaviors. Classical methods for gaze studies require either labor-intensive manual annotation of gaze behavior from video, or the use of screen-based eye tracking technologies, which require a child to examine content on a monitor screen. Manual annotation can produce valuable data, but it is very difficult to collect such data over long intervals of time or across a large number of subjects. Furthermore, it is difficult for an examiner to record instances of eye contact at the same time that they are interacting with a child, and it is often difficult for an external examiner to assess the occurrence of eye contacts. In contrast, monitor-based gaze systems are both accurate and easily scalable to large numbers of subjects, but they are generally not suitable for naturalistic, face-to-face interactions. We present preliminary findings from a wearable system for automatically detecting eye contact events that has the potential to address these limitations.

The existence of wearable gaze tracking glasses, produced by manufacturers such as Tobii or SMI, have created the possibility to measure gaze behaviors under naturalistic conditions. However, the majority of these glasses are currently being marketed for use by adults, and their form-factors preclude them being worn by children. It remains a technical challenge to adapt wearable eye-tracking technology for successful use by child subjects, and there always remains the possibility that a subject will simply refuse to wear the glasses, regardless of how light-weight or comfortable they may be. It is therefore useful to consider an alternative, noninvasive approach in which a single pair of gaze tracking glasses are worn by a cooperative adult subject, such as a clinician or teacher, and used to detect eye contact events. This is possible because, in addition to capturing the gaze behavior of the adult subject, these glasses also capture an outward-facing video of the scene in front of the subject. This video will contain the child's face in the case of a face-to-face interaction, and analysis of this video can be used to infer the child's gaze direction.

We present a system for eye contact detection which employs a single pair of gaze tracking glasses to simultaneously measure the adult's point of gaze (via standard gaze estimation) and the child's gaze direction (via computer vision analysis of video captured from the glasses). We describe the design and preliminary findings from an initial research

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

UbiComp '12, Sep 5-Sep 8, 2012, Pittsburgh, USA.

Copyright 2012 ACM 978-1-4503-1224-0/12/09...\$10.00.

study to test this solution approach. We believe this is the first work to explore this particular approach to eye contact detection.

System Setup

In order for an automatic system to detect eye contacts in a dyadic interaction, it needs to be capable of measuring each individual's gaze and determining if they are looking at each other. Here we describe some of the possible hardware and software options for building a system with such capabilities, and discuss the motivations for our system design.

Multiple Static Cameras: The most straightforward option is to instrument the environment with multiple cameras that are simultaneously capturing the faces in the scene. After synchronization of the cameras, we can then use available software for face detection and analysis, including face orientation estimation, eye detection, and gaze direction tracking, in order to identify eye contacts. However, this approach is fraught with practical difficulties: (1) the size of the interaction space is limited to the area that can be covered by a fixed set of cameras, (2) people might occlude each other in the cameras' views, (3) faces might be distant from the cameras and appear with low resolution, making the analysis of gaze extremely difficult, and (4) faces might not appear in frontal view, which makes eye detection and gaze estimation impractical. In addition to these issues, the cameras must be calibrated and the locations of the faces in the scene must be known in order to correctly interpret the computed gaze directions.

Mutual Eye Tracking: It is possible to track both the examiner's and the child's eyes using electrooculography based eye tracking systems or video based eye tracking systems. Electrooculography (EOG) based eye tracking systems place several electrodes on the skin around the eyes, and use the measured potentials to compute eye movement. In video based systems, infrared light is used to illuminate the eye, producing glints that can be used for gaze direction estimation. Unfortunately, it is probably not realistic to expect children to tolerate wearing either of these eye-tracking systems. This suggests a system based on a wearable eye-tracking device that can be worn by the adult examiner.

Examiner Eye Tracking (Our System): If only the adult examiner is able to wear the eye-tracking device, we need to record the first-person view video, i.e. egocentric video of the examiner in order to estimate the gaze of the child. We propose to detect eye contacts between the child and the adult examiner using the data captured from a wearable eye-tracking device that is worn by the adult examiner. We use SMI's wearable eye-tracking glasses for this purpose. The device is similar in appearance to regular glasses, with an outward looking camera that captures a video of the scene in front of the examiner. They further have two infrared cameras that look at wearer's eyes and estimate their gaze location in video from the outward-facing camera. Our system uses a state of the art face detection system to detect the child's face in the egocentric video, and further estimate the child's gaze orientation in 3D space. Eye contact is de-

tected if the child's gaze direction faces towards the camera, and examiner's gaze location falls on the child's face in the video.

PREVIOUS WORK

We divide the previous works into four groups: (1) first-person (egocentric) vision systems, (2) ubiquitous eye-tracking, (3) eye-tracking for identifying developmental disorders, (4) eye-tracking for interaction with children.

First-Person (Egocentric) Vision: The idea of using wearable cameras is not new [27], however, recently there has been a growing interest in using them in the computer vision community, motivated by advances in hardware technology [31, 8, 15, 1, 26, 4, 35, 9]. Ren and Gu [26] show that figure-ground segmentation improves object detection results in egocentric setting. Kitani et al. [15] recognize atomic actions such as turn left, turn right, etc. from the first-person camera movement. Aghazadeh et al. [1] detect novel scenarios from everyday activities. Pirsivash and Ramanan [25] and Fathi et al. [10, 8] use wearable cameras for detecting activities of daily living. Lee et al. [18] discover important people and subjects in first-person footage for video summarization. Fathi et al. [9] use wearable cameras to detect social interactions in a trip to an amusement park. In contrast to these methods, we use first-person vision for detecting eye contacts between two individuals in the scene.

Ubiquitous Eye-Tracking: There is a rich literature on using eye movement to analyze behaviors. Pelz and Consa [23] show that humans fixate on objects that will be manipulated several seconds in the future. Tatler et al. [33] state that high acuity foveal vision must be directed to locations that provide the necessary information for completion of behavioral goals. Einhauser et al. [7] observe that object-level information can better predict fixation locations than low-level saliency models. Bulling et al. [4] look at eye-movement patterns for recognizing reading. Liu and Salvucci [20] use gaze analysis for human driver behavior modeling. Land and Hayhoe [17] study gaze patterns in daily activities such as making peanut-butter sandwich and making tea. Researchers have shown that visual behavior is a reliable measure of cognitive load [32], visual engagement [30] and drowsiness [28]. Bulling and Roggen [3] analyze gaze patterns to identify whether individuals remember faces or other images. Different than these works, our method uses mobile eye-tracking for detection of eye contacts.

Eye-Tracking for Identifying Developmental Disorders: A large body of behavioral research indicates that individuals with diagnoses on the autism spectrum disorder (ASD) have atypical patterns of eye gaze and attention, particularly in the context of social interactions [6, 19, 29]. Eye tracking studies using monitor-based technologies suggest that individuals with autism, both adults [16] as well as toddlers and preschool-age children [5, 14], show more fixations to the mouth and fewer fixations to the eyes when viewing scenes of dynamic social interactions as compared to typically developing and developmentally delayed individuals. Importantly, atypical patterns of social gaze may already be evi-



Figure 1. Experiment setup of our method. The examiner wears the SMI glasses and interact with the child (bottom left). The glasses are able to record the egocentric video with the gaze data (bottom right).

dent by 12 to 24 months of age in children who go on to receive a diagnosis of autism (e.g. [36, 34]).

Eye-Tracking for Interaction with Children: Several eye tracking methods for infants in daily interactions have been proposed in [24, 22, 11, 12]. In particular, Noris et. al [22] presented a wearable eye tracking system for infants and compared the statistics of gaze behavior between typical developed children and children with ASD in a dyadic interaction. Guo and Feng [12] measured the joint attention during a storybook reading, by showing the same book on two different screens and simultaneously track the eye gaze of the parent and his child by two eye trackers. However, these previous work either need a specially designed wearable eye trackers for infants [22, 24, 11, 21], or limit the eye tracking on a computer screen [12]. Our method, instead, only utilizes one commercial eye tracker for adult and is able to detect eye contacts between a child and an adult in natural dyadic interactions.

WEARABLE EYE TRACKING

We use the SMI eye tracking glasses ¹ in this work. To track the eye gaze, the glasses use active infrared lighting sources. The surface of a cornea can be considered as a mirror. When light falls on the curved cornea of the eye, the corneal reflection, also known as a glint occurs. The gaze point can thus be uniquely determined by tracking the glints using a camera [13]. The SMI glasses track both of the two eyes with automatic parallax compensation at a sample rate of 30Hz, and record the high definite (HD) egocentric video at a resolution of 1280×960 for 24 frames per second. The field of view of the scene camera is 60 degree (horizontal) and 46 degree (vertical). The output of the eye tracking is the 2D gaze point on the image plane of the egocentric video. The accuracy of gaze point is within 0.5 degree. Fig.1 demonstrates the configuration of the SMI glasses in our experiment.

¹www.eyetracking-glasses.com

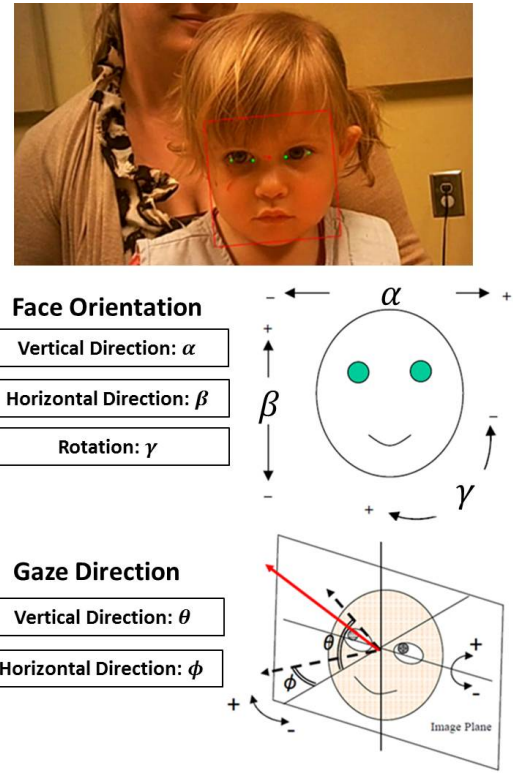


Figure 2. Face analysis results by the OKAO vision library, including the bounding box, facial parts, head pose and gaze directions. Red bounding box shows the the position and 3D orientation of the face. The green dots are the four corners of the eyes. The red line demonstrates the eye gaze direction.

FACE ANALYSIS

The problem of finding and analyzing faces in the video is a well established topic in computer vision. The state-of-the-art face detection and analysis algorithms are able to localize the face and facial parts (eyebrow, eye, nose, mouth and face contour) for real world scenarios. Recently, gaze estimation using the 2D appearance of the eye has been proposed [13]. Now, we can estimate a rough 3D gaze direction based on a single image of an eye with sufficient resolution. The core idea is learning the appearance model of the eye for different gaze directions from a large number of training samples.

We rely on a commercial software, the OMRON OKAO vision library ², to detect and analyze the child’s faces in the adult’s egocentric video. The software takes the video as the input, localizes all faces and facial parts in the video, and estimates 3D head pose and gaze direction if the eyes can be found in the current face. As we observed from the experiments, though far from perfect, the software provides promising results, especially for the near frontally-presented faces. A illustration of the detection results can be found in Fig.2. The average processing time for the HD video is 15 frames per second.

EXPERIMENTAL SETUP

²www.omron.com/r_d/coretech/vision/okao.html

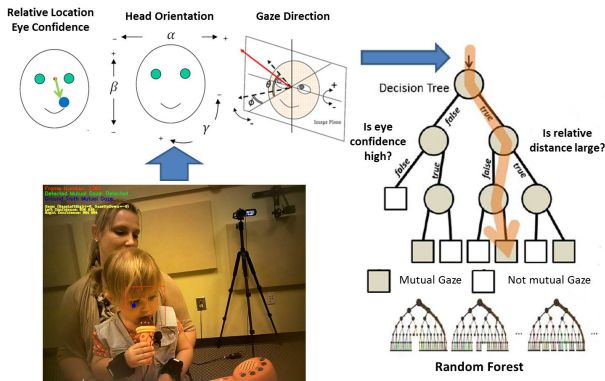


Figure 3. Overview of our approach. We combine the gaze data from the SMI glasses and face information from OKAO vision library. Features as the relative location, gaze direction, head pose and eye confidence are extracted and fit into a random forest.

Our experiment is designed for two objectives: 1) to record the video and gaze data with minimum obtrusiveness for the children; 2) to allow the analysis of the data for eye contact detection.

Protocol

We designed an interactive session (5-8 minutes) for this purpose. In our setting, the examiner would wear the SMI eye tracking glasses and interact with the child, who was sitting across from her at a small table. A number of toys were also provided for casual play. We made sure that the examiner wears the glasses at the beginning of the session, such that it would not be a distraction for the child. In addition, the examiner was required to provide online annotation of each occurrence of eye contact by pressing a foot pedal.

The gaze was tracked and the egocentric video was recorded during the interaction. We would expect a high quality image of the child’s face in the egocentric video in this setting. The OKAO vision library was further applied to the video to obtain face information of the child, including the location and orientation of the face and the 3D direction of the gaze. The adult gaze information provided by the SMI glasses and the face information by the OKAO library was then used to determine moments of eye contact.

Participants

We report the result of a preliminary study based on a typically developing female subject, age 16 months. The recorded session lasts for roughly 7 minutes. Meanwhile, we continue to collect data and expect a larger sample in the future.

METHOD

Our eye contact detection algorithm combines the eye gaze of the examiner (given by the SMI glasses) and the face information of the child, including eye gaze (given by the face analysis on the egocentric video). We extract features from both the gaze and face information for each frame of the video and train a classifier to detect the existence of an eye contact in this particular frame.

Feature Extraction

The extracted feature set includes the relative location (RL) of the examiner gaze point with respect to the child’s eye center, the 3D gaze direction (GD) of the child with respect to the image plane (up and down/left and right), the 3D head pose of the child, i.e. head orientation (HO) and confidence of the eye detection (CE). The final feature is an 8 dimension vector for each frame of the video, as shown in Fig.3.

Eye Contact Detection

The problem of eye contact detection can be considered as a binary classification problem with the human annotation as the ground truth label. For each frame in the video and given the feature, our method needs to decide whether there is an eye contact between the examiner and the child. We observe that even a simple rule would lead to reasonable results. For example, we threshold over the RL and GD, so that the eye contact is detected when the examiner’s gaze point is close to the child’s eyes and the child’s gaze is facing toward the examiner. The rule can thus be encoded by a decision tree, as shown in Fig.3. However, both the adult’s gaze data and the child’s face information contain some errors. For example, the gaze estimation by the OKAO library includes substantial frame-to-frame variation. And it is inaccurate when the child is not facing toward the examiner or the facial part is not correctly located (See the second row of Fig.7).

To deal with all these problems, we train a random forest for regression [2] on the feature vectors using human annotations. A random forest is essentially an ensemble of single decision trees, as illustrated in Fig.3. It captures different models of the data, with each model a simple decision tree, and allows us to analyze the importance of different features (See [2] for the details of random forest). We train the model on the training set (See Results for more details). The model can then detect eye contact in each frame. We leave the further temporal integration, such as an Hidden Markov Model, as future work.

RESULTS

The human annotation by the foot pedal is considered as the ground truth for training and testing. We randomly select a subset (60%) of the data as the training set and the rest for testing, and train a random forest with 5 trees with the maximum height of 6. All results are averaged over 20 runs.

Features

We first analyze the importance of features with respect to eye contact detection. The random forest algorithm outputs the importance score of each feature based on its discriminative power. The result is shown in Fig.4. We find the three most important features are relative locations (both vertical and horizontal) of adult’s gaze and vertical gaze direction of the child. The ranking yields an intuitive explanation: 1) the examiner gaze given by the SMI glasses is more reliable than the child’s gaze given by OKAO vision library; 2) vertical gaze shifts have higher scores than horizontal ones in our experiments, since the former one is more frequent than the latter one when the participants play with the toys on the desk.

Relative Location		Gaze Direction	
Vertical	Horizontal	Vertical	Horizontal
1.00	0.78	0.66	0.42
Face Orientation		Eye Confidence	
Vertical	Horizontal	Rotation	Confidence
0.40	0.31	0.35	0.51

Figure 4. Ranking of the features by the random forest. The scores are normalized, such that 1 indicates the most important feature.

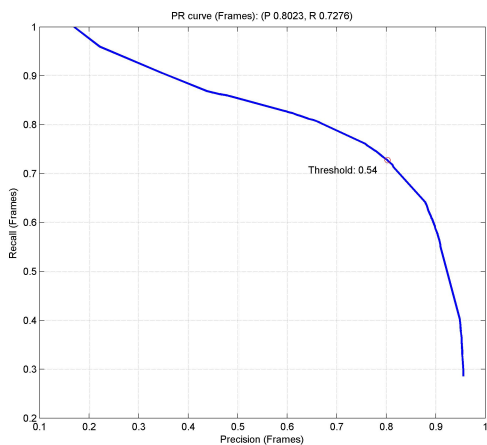


Figure 5. Precision recall curve of our eye contact detection algorithm.

Detection Performance

We consider eye contact detection as a binary classification problem, where the positive samples occupy a small portion of data. Therefore, the detection performance can be measured by precision and recall, defined as

$$\begin{aligned}
 Precision &= \frac{\# \text{ correct mutual gaze}}{\# \text{ detected mutual gaze}}, \\
 Recall &= \frac{\# \text{ correct mutual gaze}}{\# \text{ real mutual gaze}}.
 \end{aligned} \tag{1}$$

Precision measures the accuracy of the detected eye contacts by the algorithm. Recall describes how well the algorithm is able to find all ground truth eye contacts. Please note that the human annotation is not perfect. This is reported by the examiner as not able to capture every eye contact due to heavy cognitive loading. Another possible problem is the reaction delay during the boundary of eye contacts. A second rater for the annotation of eye contacts would help to disambiguate these errors, which would be considered in future work.

The precision recall curve of our eye contact detection algorithm is shown in Fig.5. Each point on the curve is a pair of precision and recall by selecting different threshold on the regression results. We choose the threshold with the highest F1 score (the harmonic mean of precision and recall) in Fig.5. The optimal threshold is 0.54 that best balances between two different type of errors. For this threshold, the overall performance is reasonably good with the precision

Algorithm \ Ground Truth	True	False
	True	1277(72.8%)
False	478(27.2%)	8225(95.6%)

Figure 6. Confusion matrix of the frame level eye contact detection results.

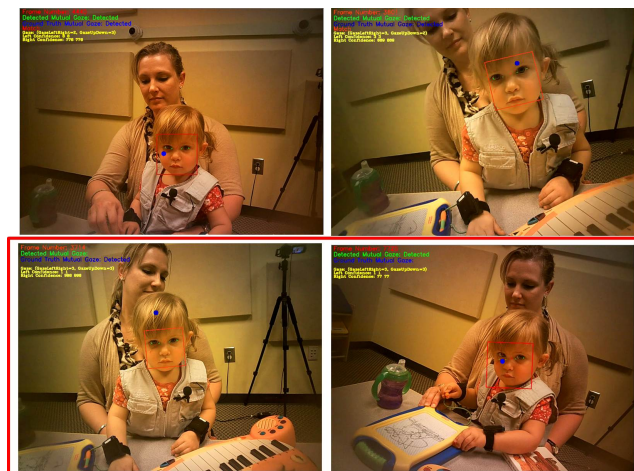


Figure 7. Example of successful (first row) and failure (second row) cases of our algorithm.

80% and recall 72%. The confusion matrix for the optimal threshold is also shown in Fig.6. Our algorithm has more false negative than false positive. The main reason of the errors, as we find in the data, is that the OKAO vision library fails to estimate the correct gaze direction in the video and thus the algorithm fails to detect eye contacts. Some of the successful and failure cases of our algorithm is demonstrated in Fig.7, which displays preliminary results from a second subject.

DISCUSSION AND CONCLUSION

We have described a system for detecting eye contact events based on the analysis of gaze data and video collected by a single pair of wearable gaze tracking glasses. Our system can be used to monitor eye contact events between an adult clinician, therapist, teacher, or care-giver and a child subject. We present encouraging preliminary experimental findings based on an initial laboratory evaluation.

REFERENCES

- O. Aghazadeh, J. Sullivan, and S. Carlsson. Novelty detection from an ego-centric perspective. In *CVPR*, 2011.
- L. Breiman. Random forests. *Mach. Learn.*, 45(1):5–32, Oct. 2001.
- A. Bulling and D. Roggen. Recognition of visual memory recall processes using eye movement analysis. In *UbiComp*, 2011.
- A. Bulling, J. A. Ward, H. Gellersen, and G. Troster. Robust recognition of reading activity in transit using wearable electrooculography. In *Pervasive Computing*, 2008.
- K. Chawarska and F. Shic. Looking but not seeing: Atypical visual scanning and recognition of faces in 2 and 4-year-old children with

- autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 39:1663–1672, 2009.
6. G. Dawson, K. Toth, R. Abbott, J. Osterling, J. Munson, A. Estes, and J. Liaw. Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology*, 40(2):271–283, 2004.
 7. W. Einhauser, M. Spain, and P. Perona. Objects predict fixations better than early saliency. In *Journal of Vision*, 2008.
 8. A. Fathi, A. Farhadi, and J. M. Rehg. Understanding egocentric activities. In *ICCV*, 2011.
 9. A. Fathi, J. K. Hodgins, and J. M. Rehg. Social interactions: A first-person perspective. In *CVPR*, 2012.
 10. A. Fathi, X. Ren, and J. M. Rehg. Learning to recognize objects in egocentric activities. In *CVPR*, 2011.
 11. J. M. Franchak, K. S. Kretch, K. C. Soska, J. S. Babcock, and K. E. Adolph. Head-mounted eye-tracking of infants' natural interactions: a new method. In *Proceedings of the 2010 Symposium on Eye-Tracking Research and Applications*, ETRA '10, pages 21–27, 2010.
 12. J. Guo and G. Feng. How eye gaze feedback changes parent-child joint attention in shared storybook reading? an eye-tracking intervention study. In *2nd Workshop on Eye Gaze in Intelligent Human Machine Interaction*, 2011.
 13. D. W. Hansen and Q. Ji. In the eye of the beholder: A survey of models for eyes and gaze. *PAMI*, 32(3):478–500, Mar. 2010.
 14. W. Jones, K. Carr, and A. Klin. Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry*, 65(8):946–954, 2008.
 15. K. M. Kitani, T. Okabe, Y. Sato, and A. Sugimoto. Fast unsupervised ego-action learning for first-person sports videos. In *CVPR*, 2011.
 16. A. Klin, W. Jones, R. Schultz, F. Volkmar, and D. Cohen. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. In *Archives of General Psychiatry*, 2002.
 17. M. F. Land and M. Hayhoe. In what ways do eye movements contribute to everyday activities? *Vision Research*, 41:3559–3565, 2001.
 18. Y. J. Lee, J. Ghosh, and K. Grauman. Discovering important people and objects for egocentric video summarization. In *CVPR*, 2012.
 19. S. R. Leekam, B. Lopez, and C. Moore. Attention and joint attention in preschool children with autism. *Developmental Psychology*, 36(2):261–273, 2000.
 20. A. Liu and D. Salvucci. Modeling and prediction of human driver behavior. In *HCI*, 2001.
 21. D. Model and M. Eizenman. A probabilistic approach for the estimation of angle kappa in infants. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, pages 53–58, 2012.
 22. B. Noris, M. Barker, J. Nadel, F. Hentsch, F. Ansermet, and A. Billard. Measuring gaze of children with autism spectrum disorders in naturalistic interactions. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 5356–5359, 30 2011-sept. 3 2011.
 23. J. B. Pelz and R. Consa. Oculomotor behavior and perceptual strategies in complex tasks. In *Vision Research*, 2001.
 24. L. Piccardi, B. Noris, O. Barbey, A. Billard, F. Keller, and et al. Wearcam: A head mounted wireless camera for monitoring gaze attention and for the diagnosis of developmental disorders in young children. In *16th IEEE International Symposium on Robot and Human Interactive Communication*, 2007.
 25. H. Pirsiavash and D. Ramanan. Detecting activities of daily living in first-person camera views. In *CVPR*, 2012.
 26. X. Ren and C. Gu. Figure-ground segmentation improves handled object recognition in egocentric video. In *CVPR*, 2010.
 27. B. Schiele, N. Oliver, T. Jebara, and A. Pentland. An interactive computer vision system - dypers: dynamic personal enhanced reality system. In *ICVS*, 1999.
 28. R. Schleicher, N. Galley, S. Briest, and L. Galley. Blinks and saccades are indicators of fatigue in sleepiness warnings: looking tired? In *Ergonomics*, 2008.
 29. A. Senju and M. H. Johnson. Atypical eye contact in autism: models, mechanisms and development. *Neuroscience and biobehavioral reviews*, 33(8):1204–1214, 2009.
 30. J. Skotte, J. Nojgaard, L. Jorgensen, K. Christensen, and G. Sjogaard. Eye blink frequency during different computer tasks quantified by electrooculography. In *European Journal of Applied Physiology*, 2007.
 31. E. H. Spriggs, F. D. L. Torre, and M. Hebert. Temporal segmentation and activity classification from first-person sensing. In *Egovision Workshop*, 2009.
 32. E. Stuyven, K. V. der Goten, A. Vandierendonck, K. Claeys, and L. Crevits. The effect of cognitive load on saccadic eye movements. In *Acta Psychologica*, 2000.
 33. B. W. Tatler, M. M. Hayhoe, M. F. Land, and D. H. Ballard. Eye guidance in natural vision: reinterpreting saliency. In *Journal of Vision*, 2011.
 34. A. M. Wetherby, N. Watt, L. Morgan, and S. Shumway. Social communication profiles of children with autism spectrum disorders late in the second year of life. In *Journal of Autism and Developmental Disorders*, 2007.
 35. W. Yi and D. Ballard. Recognizing behavior in hand-eye coordination patterns. In *International Journal of Humanoid Robots*, 2009.
 36. L. Zwaigenbaum, S. E. Bryson, T. Rogers, W. Roberts, J. Brian, and P. Szatmari. Behavioral manifestations of autism in the first year of life. In *International Journal of Developmental Neuroscience*, 2005.