EYEDIAP: a database for the development and evaluation of gaze estimation algorithms from RGB and RGB-D cameras

Conference Paper · March 2014

DOI: 10.1145/2578153.2578190

3 authors:

Kenneth Funes Mora
Eyeware Tech SA

Florent Monay
Idiap Research Institute

Jean-Marc Odobez
Idiap Research Institute

EYEDIAP: A Database for the Development and Evaluation of Gaze Estimation Algorithms from RGB and RGB-D Cameras

Kenneth Alberto Funes Mora , Florent Monay and Jean-Marc Odobez
Idiap Research Institute and École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
{kfunes, monay, odobez}@idiap.ch

Abstract

The lack of a common benchmark for the evaluation of the gaze estimation task from RGB and RGB-D data is a serious limitation for distinguishing the advantages and disadvantages of the many proposed algorithms found in the literature. This paper intends to overcome this limitation by introducing a novel database along with a common framework for the training and evaluation of gaze estimation approaches. In particular, we have designed this database to enable the evaluation of the robustness of algorithms with respect to the main challenges associated to this task: i) Head pose variations; ii) Person variation; iii) Changes in ambient and sensing condition and iv) Types of target: screen or 3D object.

CR Categories: I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Tracking; H.1.2 [Models and Principles]: User/Machine Systems—Human Information Processing

Keywords: gaze estimation; RGB; RGB-D; remote sensing; natural-light; depth; head pose; database

1 Introduction

In recent years there has been a growing interest in accurate and reliable gaze estimation systems, due to their potential importance in the development of diverse applications related to human computer interfaces, entertainment, marketing, assistance of people with disabilities, etc. Moreover, gaze is also of high interest in the sociology and psychology research where it is considered to be one of the most important cues in non-verbal behavior analysis.

Significant efforts have been devoted to the design of automatic gaze tracking solutions, leading to methods which differ according to their sensing technique and principles: from -highly intrusive- electro-oculography to more flexible video-oculography [Hansen and Ji 2010] (i.e. gaze tracking relying on video input). Solutions are available in the market, but most of them rely on costly and specialized hardware such as calibrated setups of infra-red (IR) light sources and IR cameras [Guestrin and Eizenman 2006], limiting the applications and conditions under which they properly function.

Natural light based methods are the best candidates in terms of availability, cost and potential applications. Two main approach categories can be identified in the literature: model-based, that leverage parametric and geometric description of the gaze observations but often require high-resolution images to extract the gaze features through for instance, the iris and pupil fitting techniques [Ishikawa et al. 2004; Winfield and Parkhurst 2005; Yamazoe et al. 2008]; and appearance based methods avoiding this fitting by inferring a mapping from the high-dimensional image data to the low-dimensional space of gaze parameters that can be learned from data [Baluja and Pomerleau 1994; Noris et al. 2010; Lu et al. 2011], making them adequate to handle low-resolution images. Nevertheless, gaze estimation from remote standard (RGB) cameras remains a very difficult task. The challenges are numerous: person variability, head pose variations, eyelid movements, illumination conditions, specular reflections, image resolution and contrast.

Even though the evaluations made by researchers have clearly advanced the development of gaze tracking technologies, one seldom finds evaluations done on the same data and conditions. This makes it difficult to clearly compare algorithms and identify their advantages and disadvantages. The main reason is the lack of a standard benchmark dataset.

We intend to fill this gap by releasing a database for gaze estimation from remote RGB, and RGB-D (standard vision and depth), cameras. We have designed the recording methodology in order to systematically include, and isolate, most of the variables which affect the remote gaze estimation algorithms: i) Head pose variations; ii) Person variation; iii) Changes in ambient and sensing condition and iv) Types of target: screen or 3D object. We have also defined a set of benchmarks which are intended to evaluate each one of these aspects in an independent manner, and pre-processed the data to extract and provide complementary observations (e.g. head pose) helping researchers to focus on only a subset of the problem if wanted. To our knowledge, this is the first dataset to be made publicly available for this task. We believe this is an important contribution to the community, and we therefore encourage researchers to develop gaze estimation algorithms and to report results using this data. Please visit our website to obtain this dataset[1].

We summarize below the main aspects of the dataset. Section 2 describes the recording methodology and sessions of the dataset. Section 3 summarizes the pre-processed information provided along with the data. Section 4 provides a description on how to use this dataset, including the definition of different benchmark protocols. Section 5 illustrate the usage of this dataset by evaluating the performance of a gaze estimation method under one of the defined protocols. Finally Section 6 raises some conclusions and perspectives. Further details (protocols, methodology and results) are provided in the technical report [Funes Mora et al. 2014].

2 Data

Set-up. The recording setup is as shown in Fig. 1, and comprises the elements described below along with their purpose or function:

- Kinect: this consumer device provides standard (RGB) and Depth video streams at VGA resolution (640×480) and 30fps.
- HD camera: the Kinect was designed with a large field of view imposing less restriction on user mobility but this is problematic for eye tracking based on VGA resolution. Therefore, we also recorded the scene with a full HD camera (1920x1080) at 25fps.
- LEDs: 5 LEDs visible to both cameras were used to synchronize the RGB-D and HD streams.

[1] EYEDIAP database: `www.idiap.ch/dataset/eyediap`

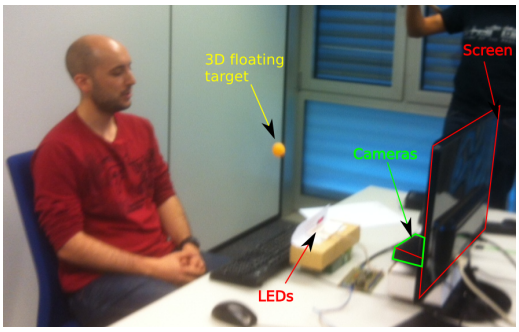

Figure 1: *Recording setup*

- Flat screen: we used a 24" screen to display a visual target.
- Small ball: we used a 4cm diameter ball as a visual target with a double purpose: to serve as a visual target in a 3D environment and be discriminative in both RGB and depth data such that its 3D position could be precisely tracked (see Section 3)

As shown in Fig. 1, the cameras are right below the computer screen, such that the eyes of the participant are observed from below and minimize eyelid occlusions. Participants were asked to sit in front of the setup at a distance depending on the type of visual target (see next paragraphs), and to gaze the specified visual target. No instructions were given in terms of speaking activity, facial expression, etc.

Recording sessions. In order to evaluate different aspects of gaze estimation algorithms, we designed a set of recording sessions, each one characterized by a combination of the four main variables that can affect gaze estimation accuracy: visual target, head pose, participant and recording conditions. These are described below:

Visual Target. It is the object which the participant was requested to gaze at. To be representative of different applications, we included the following cases: *Discrete screen target (DS)*, where a small circle was uniformly drawn every 1.1 seconds on random locations in the computer screen; *Continuous screen target (CS)*, in which the circle was programmed to move along a random trajectory for 2s, to obtain examples with smoother gaze movement; *3D floating target (FT)*: a ball with a 4cm diameter hanging from a thin thread attached to a stick that was moved within a 3D region between the camera and the participant. In contrast to the screen target, the participant was at a larger distance (1.2m instead of 80-90cm) from the camera to allow sufficient space for the target to move.

Head pose. To evaluate methods in terms of robustness to head pose, we asked participants to keep gazing at the visual target while (i) keeping an approximately static head pose facing towards the screen (*Static case, S*); or (ii) performing head movements (translation and rotation) to introduce head pose variations (*Mobile case, M*). Sample distributions are visible in the Technical report [Funes Mora et al. 2014].

Participants. We have recorded 16 people: 12 male and 4 female.

Recording conditions. For participant 12, 13 and 14, some sessions were recorded twice, in two different conditions (denoted A or B): different day, illumination and distance to the camera.

Sessions summary. We recorded 94 sessions of 2 to 3 minutes, obtaining a total of more than 4 hours of data. Each session is denoted by the string "P-C-T-H" which refers to the participant id P=(1-16), the recording conditions C=(A or B), the used target T=(DS, CS or FT) and the head pose H=(S or M) respectively. Examples of the recordings can be seen in Fig. 2.

3 Data processing

Besides the raw data itself, we also provide additional information that is essential for deriving ground truth measures or simply useful to exploit the dataset and run experiments. More details on how we estimated them can be found in [Funes Mora et al. 2014].

RGB-D sensor calibration. We provide the calibration parameters for the RGB-D stereo ensemble, which were obtained using an open source calibration toolbox [Herrera C. et al. 2012]. This allows to combine the RGB-D data into a textured 3D surface.

RGB-D to screen calibration. We provide the calibration between the camera coordinate system (3D) and the 2D screen coordinates.

RGB-D and HD camera synchrony and calibration. The HD data was synchronized with the RGB-D video stream thanks to the use of the 5 LEDs [Funes Mora et al. 2014]. In addition, standard stereo calibration between the two cameras was achieved.

Head pose and eyes tracking. For each participant we created a 3D mesh corresponding to his/her specific facial shape by fitting a 3D Morphable Model [Paysan et al. 2009] to depth data using the method described in [Funes Mora and Odobez 2012]. Furthermore, given this template, we tracked the 3D head pose using the Iterative Closest Points (ICP) algorithm, from which an approximate location of the eyeballs within the camera 3D space was then derived.

Floating target tracking. For the recording sessions using a ball as visual target, we provide the 3D center of the ball at every time step t, computed using chromatic filtering and ICP fitting.

Manual annotations. Further manual annotations of when the data is considered as being unreliable for gaze estimation (and evaluation) are provided. This corresponds to moments of eye blinking or when the person is distracted (not looking at the target). The manual annotations were done for the frontal sessions involving screen targets. Given the low occurrence of these cases, we currently did not label it in the other sessions (it will be considered as noise). Nevertheless, further annotations could be made if needed.

4 Considered tasks

In this section we describe different evaluation benchmarks that can be used to evaluate the accuracy of a gaze estimation algorithm and its robustness to different variants such as head pose, illumination conditions, etc. We first summarize the main elements of the evaluation protocol framework, including the performance measures, and then list a set of benchmarks.

4.1 Evaluation protocol and measures

This section introduces notions involved in the description of experiments: what is understood as gaze estimation algorithm; definitions of train, test and evaluation sets; and performance measures.

Gaze estimation algorithm. It is denoted as a function $\mathcal{G}$ which, provided a training set $\mathcal{V}$ and test data $\mathcal{T} = \{\mathbf{I}_i, i = 1 \dots T\}$, outputs a set of gaze estimates $\mathcal{G} = \mathcal{H}(\mathcal{T}|\mathcal{V})$ with $\mathcal{G} = \{\mathbf{g}_i, i = 1 \dots T\}$.

The output of the gaze estimation depends on the application. Here we consider two very common cases: A *3D gaze ray* $\mathbf{g} = \{\mathbf{o}, \mathbf{v}\}$ defined by its origin $\mathbf{o} \in \mathbb{R}^3$ and a *unitary* vector $\mathbf{v} \in \mathbb{R}^3$; or *Screen coordinates* $\mathbf{g} = \mathbf{s} \in \mathbb{R}^2$ (pixels) which are often used for screen based applications in HCI. Note that provided our additional information (camera-screen calibration) it is possible to infer screen coordinates from the 3D gaze ray [Funes Mora et al. 2014].

Training data. The training data consist of pairs of data (images) and associated ground truth information leading to the training set $\mathcal{V} = \{(\hat{\mathbf{I}}, \hat{\mathbf{g}})_{\hat{t}}, \hat{t} = 1, \dots, N\}$. We consider different ways to collect these samples: *Temporal*, i.e. data corresponds to section(s) of a larger video; or *Structured*: the training data is collected in a structured manner in order to fulfill a specific requirement of the gaze estimation algorithm (e.g. to obtain the closest samples to a predefined number of points in a screen with specific $\hat{\mathbf{g}}$ values).

As ground truth data, we used the 3D location $\hat{\mathbf{g}} := \hat{\mathbf{p}} \in \mathbb{R}^3$ of the

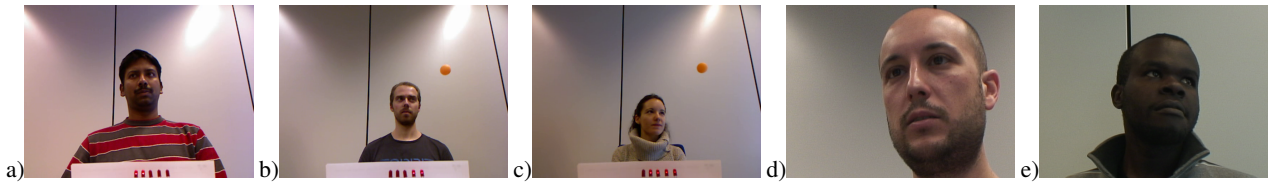

Figure 2: *Recorded data samples using: a–c) the RGB-D camera; and d–e) the HD camera, for which a patch of* 640×480 *pixels of the original images is shown for comparison to the VGA resolution data. In these examples the participant is gazing at: a,d) the screen target with a static head pose; b) the floating target with a static head pose; c,e) the floating target while moving the head.*

3D visual target, or the $\hat{\mathbf{g}} := \hat{\mathbf{s}} \in \mathbb{R}^2$ screen coordinates (note that given our calibration $\hat{\mathbf{p}}$ can be computed from $\hat{\mathbf{s}}$).

Test data $\mathcal{T}$. It is assumed to be a temporal section of a larger video within the frame index range $[t_0, t_1]$.

Evaluation data. We define as evaluation set $\hat{\mathcal{E}}$ the samples used to compute an algorithm's performance. This is a subset of the test data obtained by removing the samples where either the data or the ground truth is corrupted, due to blinking and distractions, extreme head poses which compromise the eye visibility (e.g. occlusion by the nose), and train and test set intersection.

Interpolation based methods and convex-hull. We considered the issue of algorithms needing test data that match the training conditions, like many appearance based methods that are mainly capable of estimating gaze only within the convex hull of the training data gaze directions. In our dataset, such condition cannot always be guaranteed, particularly when considering the floating target cases. To allow fair comparison for such methods, we will provide an evaluation set for which the test samples outside of the convex hull of eye-in-head gaze directions are taken out from the evaluation set. Users will need to report which of the evaluation sets they used for evaluation. Note that comparing performances on both the convex hull and in the full range will be interesting to distinguish algorithms which are capable of extrapolating gaze estimation (typically, those that are model-based) and those that cannot.

Performance measures. For an index t in the evaluation set $\hat{\mathcal{E}}$, with estimated gaze direction $(\mathbf{o}_t, \mathbf{v}_t)$ or screen coordinates $\mathbf{s}_t$, we considered the following performance measures to compare algorithms (formulas can be found in [Funes Mora et al. 2014]:

The 3D distance error ${\epsilon^d}_t$, useful for the 3D gaze estimation tasks, and that conveys how close the estimated 3D gaze ray passes near the visual target 3D position $\hat{\mathbf{p}}_t$.

The Angular error ${\epsilon^\circ}_t$, which is a normalization alternative to ${\epsilon^d}_t$, measuring the error in terms of directional error.

The Screen pixel error ${\epsilon^{\mathbf{s}}}_t$, used for the screen pixel coordinate prediction task. Note that by using the provided calibration, we can compute an angular error from screen coordinates predictions.

The above errors allow to compute statistics (usually the mean) on the evaluation set $\hat{\mathcal{E}}$, like the mean distance error $\epsilon^d = \frac{1}{|\hat{\mathcal{E}}|}\sum_{t\in\hat{\mathcal{E}}} {\epsilon^d}_t$, the mean angular error $\epsilon^\circ = \frac{1}{|\hat{\mathcal{E}}|}\sum_{t\in\hat{\mathcal{E}}} {\epsilon^\circ}_t$, and the mean screen pixel error $\epsilon^{\mathbf{s}} = \frac{1}{|\hat{\mathcal{E}}|}\sum_{t\in\hat{\mathcal{E}}} {\epsilon^{\mathbf{s}}}_t$.

Sensitivity. In addition to accuracy, we can report additional performance measures, like sensitivity that can be used to measure algorithm robustness under different experimental conditions.

4.2 Predefined experimental protocols

To allow comparisons between algorithms and their merit under different experimental conditions, we have defined a set of protocols that differ mainly in the recordings of our database that are used for training and testing the algorithms[2]. Notice this dataset has two main types of visual targets: 3D floating target (FT) and screen target (CS or DS). Therefore, the defined evaluation protocols can have variations according to the preferred visual target. We summarize them below. More details are given in [Funes Mora et al. 2014].

Protocol 1: Gaze estimation accuracy. In this protocol, we evaluate the accuracy of an algorithm $\mathcal{H}$ under minimal variation of all parameters which are not gaze variations. More precisely, for a session S, where the only variation is in the gaze itself, we define the training set $\mathcal{V}$ as the first half of S. The test set $\mathcal{T}$ is defined as the second half of S. The result of such experiment is the mean angular error ϵ°. The relevant sessions can be derived according to the type of visual target.

Protocol 2: Robustness to head pose variations. The objective here is to measure how much the gaze accuracy decays due to head pose variations. Experiments can be conducted with the static head pose (S), and then with head pose variations (M). The average errors under both conditions, as well as the sensitivity of the algorithm to head pose variations are then reported.

Protocol 3: Person dependence. The goal is to evaluate how well a method $\mathcal{H}$ generalize to unseen users. This can be conducted using a leave-one-person-out experimental set-up.

Protocol 4: Condition variations. Finally, in this case the goal is to study the generalization properties of a method $\mathcal{H}$ to different conditions. To this end, experiments can be conducted for participants 12, 13 and 14 for which recording sessions under different set-up and illumination conditions are available.

In the next section we show an example of use.

5 Evaluation protocol example

To illustrate the usage of the dataset, here we describe in detail the data related to one of the benchmarks: the gaze estimation accuracy protocol (Protocol 1) for the 3D floating target and using the RGB-D stream as input data for the gaze estimation algorithm. Note that this configuration is one of the most challenging cases in our dataset, for which the typical eye image size is $\approx 14 \times 10$ pixels.

Head pose and 3D target tracking. In Table 1 we show the total number of frames corresponding to each session, together with the number of frames for which we were able to successfully estimate the head pose or the visual target position. Note that the head pose recall is high (99.9%) as expected since in this protocol the recordings involving people with a near frontal and static head pose. The recall of visual target location is lower (69%) due to the target being outside the camera's field of view, or too close to the sensor causing missing depth data. Nevertheless, these numbers show that a large quantity of gaze labeled data is available for experimentation (around 100 seconds per recording).

[2]To allow comparison with others, researchers are strongly encouraged to (i) use the benchmarks defined here; or (ii) publish their protocol details (e.g. people id, frames numbers) to allow evaluation comparison if they define their own evaluation protocols using the provided dataset.

Table 1: *First to third row: number of frames of the recorded video (Total), and on which the head pose or floating target position estimation succeeded. Fourth to sixth row: size of the training, test and evaluation sets for Protocol 1.*

	"P-C" (Participant-Conditions) for Session: "P"- "C"-FT-S																			
	1-A	2-A	3-A	4-A	5-A	6-A	7-A	8-A	9-A	10-A	11-A	12-B	13-B	14-A	14-B	15-A	15-B	16-A	16-B	Avg.
Total frames	4231	4441	4201	4291	4171	3571	4381	4381	4351	4201	4321	5769	5907	4411	5255	4411	4037	4261	5634	4538
Head Pose	4231	4441	4189	4291	4169	3568	4332	4372	4345	4186	4311	5769	5907	4411	5255	4411	4037	4257	5634	4532
Ball Target	2012	2323	2274	1752	2616	3064	2505	3471	3842	2561	2673	4429	4717	3924	3309	4014	3038	2383	4371	3119
Training	833	890	1164	581	1253	1319	924	1640	1868	1133	975	2097	2590	1827	1642	1922	1482	1272	2171	1451
Test	1179	1433	1100	1171	1361	1742	1550	1822	1968	1416	1688	2332	2127	2097	1667	2092	1556	1107	2200	1663
Evaluation	1049	1216	994	941	1229	1302	765	1778	1429	1063	1258	1967	1383	1658	1091	1566	890	1020	1495	1268

Table 2: *Gaze angular error comparison for Protocol 1. The 'Head' case only uses the estimated head orientation as gaze prediction.*

	"P-C" (Participant-Conditions) for Session: "P"- "C"-FT-S																			
	1-A	2-A	3-A	4-A	5-A	6-A	7-A	8-A	9-A	10-A	11-A	12-B	13-B	14-A	14-B	15-A	15-B	16-A	16-B	Avg.
ϵ°(Head)	25.7	24.4	24.1	27.2	25.7	26.0	27.7	24.6	29.0	26.1	25.0	26.5	26.9	27.1	26.7	25.9	31.3	22.9	26.1	26.3
ϵ°(PR-ALR)	6.8	6.8	6.8	11.3	12.6	7.2	16.3	5.9	10.4	7.4	8.0	8.9	4.8	5.7	6.6	6.3	8.0	6.9	6.8	8.1

Protocol sets. Each session is divided equally into two temporal sections: the first half is the *training* set and the second half is the *test* set. The test set was filtered to define the *evaluation* set (the criteria is described in the next section). Table 1 shows the number of samples for each set (only considering samples with known head pose and target location).

Gaze estimation method. We implemented an RGB-D based method [Funes Mora and Odobez 2012] that relies on RGB-D data to rectify the eye images viewpoint into a canonical head pose. We refer to this method as pose-rectified adaptive linear regression (PR-ALR). For each participant, a gaze appearance model of 42 samples was extracted from the training set. These samples are regularly distributed with gaze yaw values between $\pm 40^\circ$ and gaze elevation values between $\pm 30^\circ$. Since PR-ALR is an interpolation based approach, we only considered the test samples within the convex-hull of the data used in the gaze appearance model, i.e. the evaluation set consisted of only the test samples for which the ground truth measures respected the same yaw and elevation gaze criteria.

Estimation accuracy. To demonstrate the data variability, we computed the gaze angular errors obtained when assuming the participant is gazing towards the front, that is, assuming the gaze direction is given by the head pose direction. The results of the experiments are shown in Table 2. Note that both methods output a 3D gaze ray and we used the same evaluation set such that these results are directly comparable. The angular errors shown by the "Head" case provides evidence of the large gaze variability within the data. The gaze estimation accuracy is drastically improved once the PR-ALR gaze estimation algorithm is used.

Still, the errors are high in comparison to results reported in the literature, which is mainly due to the low resolution ($\sim 14 \times 10$ pixels per eye) and poor contrast (e.g. participant 7-A has black skin). In addition, outliers (blinks, distractions, etc) were not yet taken out from the evaluation set[3]. Notice this is one of the most challenging scenarios in our data and this experiment is adequate to demonstrate how to use the data, the defined protocols, and how to characterize an experiment.

6 Conclusion

We have described a novel dataset for the development and evaluation of gaze estimation algorithms from RGB or RGB-D data that addresses the need of the community for standarized benchmarks.

The database is rich and diverse as it is representative of the main challenges of this task. Most variables (head pose, person, conditions and type of target) have been systematically isolated, with the goal of properly characterizing a gaze estimator in terms of accuracy and robustness to adverse conditions.

The recording methodology, and a summary of the recorded data, have been described. We have also listed the additional information provided to the users, such as the setup calibration, target position, and head and eye tracking information. We have described in more detail an experiment which serves as usage example.

We believe this database is of high value to researchers as it will help to advance the development of gaze estimation technologies under less constrained conditions.

Acknowledgements. Authors gratefully acknowledge the support from the Swiss National Science Foundation (Project: FNS-203, TRACOME) `www.snf.ch`, and express their gratitude to the people who kindly participated in the data recordings.

[3] Annotations are provided to remove outlier samples for some sessions.

References

BALUJA, S., AND POMERLEAU, D. 1994. Non-Intrusive Gaze Tracking Using Artificial Neural Networks. Tech. rep., CMU.

FUNES MORA, K. A., AND ODOBEZ, J.-M. 2012. Gaze Estimation From Multimodal Kinect Data. In *Computer Vision and Pattern Recognition Workshops*, 25–30.

FUNES MORA, K. A., MONAY, F., AND ODOBEZ, J.-M. 2014. Eyediap database: Data description and gaze tracking evaluation benchmarks. Tech. Rep. RR-05-2014, Idiap, Feb.

GUESTRIN, E. D., AND EIZENMAN, M. 2006. General theory of remote gaze estimation using the pupil center and corneal reflections. *Trans. on bio-medical engineering* (June).

HANSEN, D. W., AND JI, Q. 2010. In the eye of the beholder: a survey of models for eyes and gaze. *IEEE Tr. Patt. Anal. and Machine Intelligence 32*, 3, 478–500.

HERRERA C., D., KANNALA, J., AND HEIKKILÄ, J. 2012. Joint Depth and Color Camera Calibration with Distortion Correction. *TPAMI 34*, 10, 2058–2064.

ISHIKAWA, T., BAKER, S., MATTHEWS, I., AND KANADE, T. 2004. Passive Driver Gaze Tracking with Active Appearance Models. In *Proceedings of the 11th World Congress on Intelligent Transportation Systems*, 1–12.

LU, F., SUGANO, Y., TAKAHIRO, O., AND SATO, Y. 2011. Inferring Human Gaze from Appearance via Adaptive Linear Regression. In *ICCV*.

NORIS, B., KELLER, J., AND BILLARD, A. 2010. A wearable gaze tracking system for children in unconstrained environments. *Comp. Vis. and Image Understanding*.

PAYSAN, P., KNOTHE, R., AMBERG, B., ROMDHANI, S., AND VETTER, T. 2009. A 3D Face Model for Pose and Illumination Invariant Face Recognition. In *AVSS*.

WINFIELD, D., AND PARKHURST, D. 2005. Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. *Computer Vision and Pattern Recognition Workshops*.

YAMAZOE, H., UTSUMI, A., YONEZAWA, T., AND ABE, S. 2008. Remote gaze estimation with a single camera based on facial-feature tracking without special calibration actions. In *Proc. of ETRA*, ACM, New York, NY, USA, ETRA '08.