

Hybrid Eye Tracking: Combining Iris Contour and Corneal Imaging

Alexander Plopski^{†1}, Christian Nitschke², Kiyoshi Kiyokawa¹, Dieter Schmalstieg³, and Haruo Takemura¹

¹Osaka University, Japan

²Kyoto University, Japan

³Graz University of Technology, Austria

Abstract

Passive eye-pose estimation methods that recover the eye-pose from natural images generally suffer from low accuracy, the result of a static eye model, and the recovery of the eye model from the estimated iris contour. Active eye-pose estimation methods use precisely calibrated light sources to estimate a user specific eye-model. These methods recover an accurate eye-pose at the cost of complex setups and additional hardware. A common application of eye-pose estimation is the recovery of the point-of-gaze (PoG) given a 3D model of the scene. We propose a novel method that exploits this 3D model to recover the eye-pose and the corresponding PoG from natural images. Our hybrid approach combines active and passive eye-pose estimation methods to recover an accurate eye-pose from natural images. We track the corneal reflection of the scene to estimate an accurate position of the eye and then determine its orientation. The positional constraint allows us to estimate user specific eye-model parameters and improve the orientation estimation. We compare our method with standard iris-contour tracking and show that our method is more robust and accurate than eye-pose estimation from the detected iris with a static iris size. Accurate passive eye-pose and PoG estimation allows users to naturally interact with the scene, e.g., augmented reality content, without the use of infra-red light sources.

Categories and Subject Descriptors (according to ACM CCS): I.4.8 [Computer Graphics]: Scene Analysis—Shape, Object recognition

keywords: eye-pose estimation, corneal imaging, 3D interaction, gaze interaction

1. Introduction

Gaze estimation is beneficial for a wide range of applications, such as gaze sensing, user attention visualization, automatization, passive data collection, gaze reactive and first person view applications, as well as point-of-regard (PoR) estimation. Detecting the user's PoR offers interesting opportunities for interaction with content in augmented reality (AR) [IR11, PLC08] and virtual reality (VR) [SGRM14, TJ00] environments. Especially when wearing a head-mounted display (HMD), a non-intrusive interaction solution is desirable. Current interaction through a built-in or attached touchpad requires the user to divide the focus between the interaction with the input device and the content. Eye-gaze interaction

allows the user to naturally interact with the content and will, thus, likely improve the overall experience. Another application scenario is projector-based AR, where the user triggers augmentations with his or her gaze.

A common solution to acquire the PoR is the geometric approach. Hereby, the PoR is acquired through estimation of the eye-pose relative to the eye-tracking camera and intersecting the estimated eye-gaze with the scene model.

Eye Pose Estimation A variety of solutions that recover the geometrical eye-pose have been developed in more than 30 years of research on eye-tracking [HJ10]. They can be divided into *active methods*, which require extensive additional hardware and complex calibration, and *passive methods* that are used with natural images. Active eye-pose estimation requires multiple light-sources (commonly IR LEDs) whose position

[†] E-mail: plopski@lab.ime.cmc.osaka-u.ac.jp

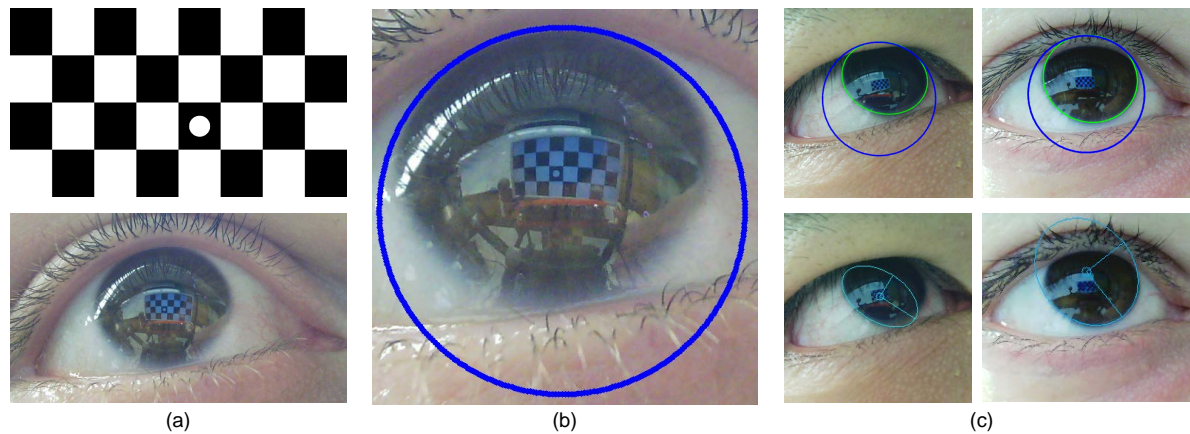


Figure 1: Proposed hybrid eye-pose estimation. ((a) top) The pattern shown on a monitor reflects on the user's corneal surface ((a) bottom). (b) To determine the eye-pose, we first estimate the position of the cornea. We recover the orientation of the eye from the contour of the iris. ((c) top) The estimated position of the cornea allows us to estimate the gaze direction from edges detected in the image. ((c) bottom) We can successfully determine the correct eye-pose even in scenarios where the iris cannot be recovered by standard iris-reconstruction methods.

relative to the camera is calibrated. The position of the eye is determined from the reflection of the light sources on the cornea surface, the outer layer of the eye. The orientation is recovered by computing the center of the pupil, segmented under IR light. This approach is known as pupil center cornea reflection (PCCR) [GE06] and is widely used in commercial eye trackers, such as the Eye Tribe Tracker [eye]. Mobile versions of such trackers can be used in combination with HMDs, e.g., the FOVE HMD [fov], or the work of Ishiguro et al. [IMMR10]. Active eye-pose estimation solutions recover the eye-pose with an accuracy of up to 0.5 deg. To preserve this accuracy, a complicated and time-consuming recalibration must be performed, whenever the position of the LEDs changes. Additionally, although eye trackers pass safety standard requirements, long-term deployment in everyday devices is undesirable, especially for elderly and children [NKB11].

The requirement for active-light image capture also makes it impossible to apply active methods in unknown environments, outdoor in sunlight, or with natural images. Eye-pose estimation methods which do not require active manipulation of the environment and, thus, can be applied to natural images are referred to as *passive*. While active methods recover the eye-pose in a two-stage approach, passive methods compute all 6 DOF at the same time. Passive methods can be divided into appearance based methods [WBZ*15] and reconstruction solutions that use the detected eye features, commonly the iris contour and eye lids, to reconstruct the pose of the 3D eye model [IK14, WKW*07]. Hereby, the region of interest (ROI) around the iris is necessary to reliably detect eye features.

Passive methods usually assume static eye model parameters. This means that the geometrical eye model is identical for all users. In recent years, a number of solutions which try to actively estimate the parameters of the eye model have been proposed. It has been shown that it is possible to accurately estimate the eye-pose from natural images. These methods reconstruct the eye model from the features detected in the image and require extensive calibration sessions [TK12] or estimate a large number of parameters at runtime [WKW*07], which makes them more prone to errors. Compared to active eye-pose estimation, passive methods achieve lower accuracy as a result of ambiguous or falsely detected features, e.g., noise from the eye-lashes, gradual transition of the iris and modeling errors.

Active as well as passive solutions recover the eye-pose up to the optical axis of the eye. To determine the user-specific parameters which align the optical axis with the visual axis, the actual gaze direction, a one-time calibration is necessary [HJ10]. In practice, passive methods often ignore this parameter.

PoR Estimation The geometric approach reconstructs the PoR as the intersection of the estimated gaze direction with the scene model. The PoR is often estimated at a predefined plane in front of the user. This approximation is sufficient to generate the visualization of the PoR, e.g., in a heat map. However, this leads to parallax issues, if the visual depth changes. If the user is interacting with a planar surface, e.g., a monitor or handheld device that is rigidly attached to the tracker, a mapping of the estimated eye-pose to points on the

plane can be learned in a calibration session, in which the user is looking at multiple points displayed on the surface.

To enable interaction with a 3D scene, for example triggering distributed AR markers, an accurate scene model is necessary. It can be acquired through a number of approaches, e.g., manual reconstruction, single camera SfM, multi-camera stereo, or KinectFusion. This model is continuously aligned with the eye tracking camera by using an outwards pointing camera [IK14].

Alternatively, the PoR can be recovered from the gaze reflection point (GRP) [Nit11] that is estimated on the surface of the cornea [NN12].

This work In AR and VR environments, the model of the environment is known, be it the object that will be augmented, the plane of the optical-see-through head-mounted-display (OST-HMD), or the screen of a handheld device. Our goal is to enable accurate eye-pose and PoR estimation in such environments by passive eye-pose estimation, as we strongly believe that AR, especially in combination with HMDs, will become a commodity in the future.

Plopski et. al. [PIN*15] use corneal imaging (CI) [NN06] to determine the position of the cornea from the reflection of the OST-HMD screen. They combine multiple observations to recover the center of the eye and suggest that the gaze can be recovered as the ray through the centers of the eyeball and the cornea. This geometric approach requires accurate estimation of both centers, because small inaccuracies will result in large errors. As such, it is difficult to apply in more general scenarios in which the tracker is not rigidly mounted onto the head. Furthermore, their solution requires a number of distinctive eye-gaze directions to recover the eye center. Thus, it is not applicable if the user’s gaze remains fixated at the point of interest.

We expand this work and introduce a *hybrid eye-pose estimation* approach (Figure 1) — a combination of active and passive eye-pose estimation methods. Our method does not require interference with the environment, and uses a tracked model of the environment to recover the eye-pose. We propose to separate the eye-pose estimation approach into the estimation of the position and the orientation of the eye, the state-of-the-art in active methods, instead of recovering all 6 DOF from the iris contour. We also use the accurately estimated eye-pose to determine personal parameters at runtime. We show that the size of the iris recovered from natural images varies for each user, as a result of changing illumination or camera exposure time.

Our method is designed for natural images. This allows it to be applied in indoor as well as outdoor scenarios, where strong illumination may interfere with active trackers. We compare the requirements of our method to passive and active estimation in Table 1.

Table 1: Comparison of eye-pose estimation strategies.

Features	Active	Passive	Hybrid
Accurate eye-pose	yes	no	yes
Eye-pose from natural images	no	yes	yes
Personal parameter calibration	beneficial	beneficial	beneficial
Eye-model parameter estimation	yes	no	yes
Geometrical estimation	yes	no	yes
Restrictions			
ROI required	no	yes	no
IR light: limited outdoor use	yes	no	no
IR light: long-term exposure	yes	no	no
Complex setup	yes	no	no
Scene model required	no	no	yes
Parallax issues	yes	yes	no

Contribution Our eye-pose estimation recovers an accurate eye-pose that can be easily used to determine the PoR. The main contributions of our paper are:

- We show how the two-stage approach used in active eye-pose estimation can be used with passive eye-pose estimation in environments with a known scene model.
- We propose a novel iris estimation approach, which uses the accurately estimated 3D position of the cornea to determine a user dependent iris size and, thus, increases the robustness of the method and improves the recovered eye-pose.
- Our approach does not require an estimated ROI of the eye and recovers the eye-pose from the detected reflection of the scene on the cornea.

2. Eye Model

When looking at a human eye (Figure 2a), we see two distinct parts of the eyeball — the iris and the white sclera. Additionally, the pupil and the iris pattern can be seen as parts of the iris. The cross-section of the eye (Figure 2b) shows that the iris is covered by a transparent, protective layer — the cornea. The sclera and the cornea have an approximately spherical shape; thus, the eye can be seen as two spheres of different radii and displaced centers. The essential elements of these spheres are shown in Figure 2c.

The two-sphere model is a very simple, but not an ideal, representation of the human eye; neither the sclera nor the cornea are ideal spheres. Nonetheless, we use this model, as it is a simple and good representation of the eye that is often used in eye related studies [IK14, NN06, NN12, PIN*15].

In this model, the cornea C is the smaller sphere with a radius r_C , located at C . The center of the sclera corresponds to the eye’s center of the rotation, E . When the eye rotates to observe various objects in the surroundings, C moves around E on a spherical orbit. The intersection of the cornea and the sclera is the limbus L . It is commonly assumed that L corresponds to the circular circumference of the iris.

E , C and L , the center of the limbus, lie on the optical axis of the eye, \mathbf{o} . The visual axis \mathbf{v} is described by the fovea and the nodal point of the eye. The position of the nodal point

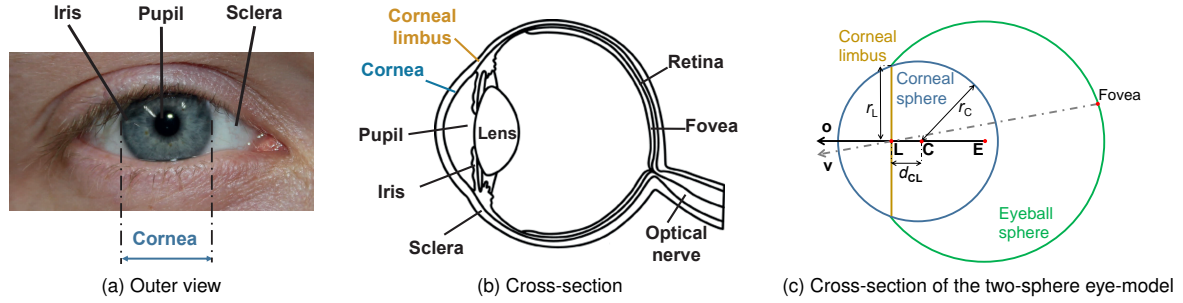


Figure 2: The parts of the eye and the corresponding geometric eye model used in our method. [PIN*15]

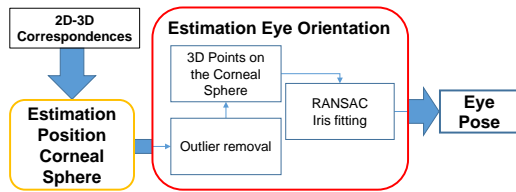


Figure 3: Our method estimates the eye-pose in two steps. First we use detected 2D-3D correspondences to estimate the position of the corneal sphere. With this information we can recover the orientation of the eye. The result is an accurate eye pose.

changes whenever the user focuses at a different distance, however, it is generally assumed to coincide with L . Usually \mathbf{o} and \mathbf{v} are approximately 5° apart [SF72]. This offset can be described by horizontal and vertical offset angles (α, β) .

According to Nitschke et al. [NNT13] the model parameters are as follows: $r_C = 7.8$ mm, the radius of the corneal limbus $r_L = 5.5$ mm, and the distance between C and L $d_{CL} = 5.7$ mm. Although it is possible to compute a rough eye-pose using static model parameters, accurate eye-pose estimation requires the user specific offset, (α, β) .

It has been shown that estimating personal parameters instead of using static values can improve the eye-pose estimation results [TK12, WKW*07].

3. Method

Passive methods recover the eye-pose by fitting the eye model to the extracted iris contour. We use the inverse approach (Figure 3). We first recover an accurate position of the corneal sphere that provides the translational parameters of the eye. The accurately estimated position of the corneal sphere is used to determine the rotational parameters of the eye. In the following section, we describe how we use corneal imaging to recover the position of the corneal sphere, followed by the proposed recovery of the gaze direction. As we assume a

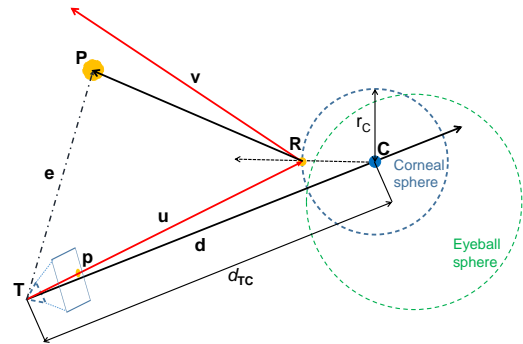


Figure 4: At a given distance d_{TC} , the ray \mathbf{u} reflects at the cornea C in point R as \mathbf{v} . At the correct distance \mathbf{v} will intersect P . [PIN*15]

known scene model, the recovered eye-pose can be used to compute the PoG by intersecting the estimated gaze with the scene model.

3.1. Corneal Position Estimation

The estimation of the position of the corneal sphere has been covered in Plopski et al. [PIN*15]. We describe the approach here for convenience.

As shown in Figure 4, let \mathbf{p} be the reflection of a 3D-Point P on the corneal sphere captured by the eye-tracking camera T . The backprojection ray \mathbf{u} through \mathbf{p} and \mathbf{e} , the ray from T towards P , lie in the plane π with a normal $\mathbf{n} = \mathbf{u} \times \mathbf{e}$. The center of the corneal sphere C also lies in the plane π . Given two planes π_1 and π_2 , the ray \mathbf{d} , the ray from T towards C , is lying in both planes. Thus, it is the intersection of the planes, $\mathbf{d} = \hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2$.

For any given distance d_{TC} along \mathbf{d} , the ray \mathbf{u} intersects the corneal sphere in R , where it reflects as \mathbf{v} . If d_{TC} is correct, \mathbf{v} will intersect P . For N correspondence pairs, the correct distance \hat{d}_{TC} is computed as

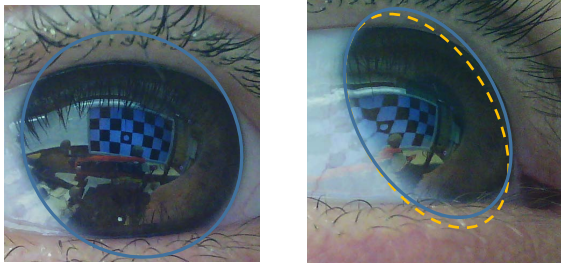


Figure 5: (left) When the user looks towards the camera the contour of the iris is clearly visible and the correct iris contour is easily recovered. We show the estimated iris contour in blue. (right) As the eye rotates sideways, reflections on the corneal sphere occlude the iris contour shown in orange. Naïve ellipse fitting assumes that the occluding contour of the cornea is part of the iris.

$$\tilde{d}_{\text{TC}} = \operatorname{argmin}_{d_{\text{TC}}} \frac{1}{N} \sum_{i=1}^N \|\mathbf{v}_i \times (\mathbf{P}_i - \mathbf{R}_i)\|. \quad (1)$$

3.2. Eye-Pose Estimation

In PCCR, the orientation of the eye is computed as the ray through the centers of the cornea and the pupil. In images captured under visible light, the pupil cannot be detected reliably, e.g., Figure 1. Therefore, we estimate \mathbf{L} as the center of the iris. Iris detection in the camera image suffers from erroneously detected edges. While the iris contour is clearly visible in Figure 5a, reflections on the cornea occlude a portion of it in Figure 5b. These situations are indistinguishable without 3D constraints. We use the 3D corneal sphere to improve the fitting results and recover a closer representation, which accounts for both cases. Additionally, our approach is robust against other detected edges, such as eye-lids, eyelashes, sclera and iris patterns, and reflections on the corneal surface. In this section, we describe how we use the accurately estimated corneal sphere to determine the orientation of the eye from edge points detected in the captured image.

Given N edge points \mathbf{p}_i , $i = 1 \dots N$, detected in the image, we remove all obvious outliers by intersecting the backprojected rays \mathbf{u}_i with the corneal sphere. For an inlier point \mathbf{p}_i , \mathbf{u}_i intersects the corneal sphere in \mathbf{R}_i .

For M points on the 3D sphere, we determine \mathbf{o} and d_{CL} through a RANSAC approach. From the M 3D points, we select $L \geq 3$ candidate points and fit the limbal plane to them. The estimated limbal plane intersects the corneal sphere in the corneal limbus. Therefore, the normal of the limbal plane will correspond to \mathbf{o} and $d_{\text{CL}} = \mathbf{o}^T (\mathbf{R}_k - \mathbf{C})$, where \mathbf{R}_k is one of the candidate points. We determine the support of the estimated limbal plane by counting the number of inlier

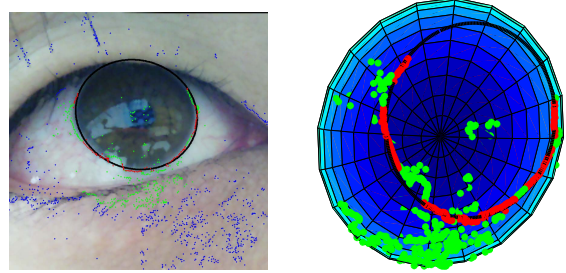


Figure 6: Points which do not lie within the projection of the corneal sphere into the image are removed as outliers (blue). From all points on the corneal sphere (green), the corneal limbus (black) is the ring which is supported by the highest number of points on the corneal sphere surface (red). The fitting results are shown in the image on the left, and the corresponding 3D sphere on the right.

points \mathbf{R}_i , $i \in M$. An inlier of the fitted limbal plane satisfies one of the following conditions:

$$\|\mathbf{R}_i - \mathcal{P}_L\| < t_1, \text{ or} \quad (2)$$

$$|\mathbf{u}_i^T \frac{\mathbf{R}_i - \mathbf{C}}{\|\mathbf{R}_i - \mathbf{C}\|}| < t_2, \quad (3)$$

where t_1 and t_2 are user-defined inlier thresholds, and \mathcal{P}_L is a set of points on the 3D limbus, evenly distributed in 1° steps. If \mathbf{R}_i satisfies (2), the point is lying at most t_1 away from the limbus contour. \mathbf{R}_i will satisfy (3), if the eye is oriented so that the cornea is occluding a portion of the iris, as in Figure 5b. In this case, \mathbf{p}_i lies at the edge of the projection of the corneal sphere into the image. After the best inlier subset has been selected, we perform the fitting step again with all inlier points. We use the following empirically selected thresholds: $t_1 = 0.3$ mm, and $t_2 = 5^\circ$. We show a sample result of the fitting process in Figure 6.

4. Experiment Environment

We have implemented our hybrid method in C++ on an Intel i7-7000 with 32 GB RAM. Our implementation recovers the eye-pose in less than 0.6 s/frame (0.1-0.3 s for checkerboard detection and matching, 0.1-0.25 s for estimation of the position of the corneal sphere and 0.05 s for estimation of the orientation).

We have prepared a simple environment shown in Figure 7(a) to evaluate the accuracy of the eye-pose recovered by our method. The users were shown a 8×4 checkerboard pattern on an LCD monitor S (293.2×521.3 mm) that was positioned at a variable distance in front of the user. We use a Delock USB 2.0 camera with a 64° lens focused at a 5-7 cm distance as the eye-tracking camera T (Figure 7(b)).

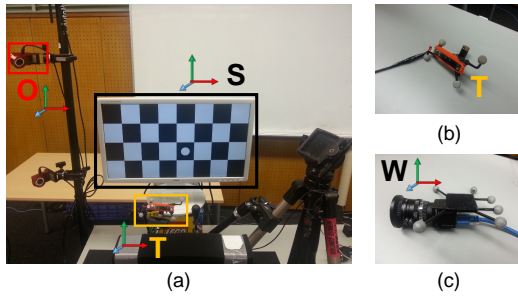


Figure 7: (a) Our experiment environment consists of an OptiTrack system, an eye-tracking camera and an LCD monitor: (b) The camera T is mounted onto a tripod, and its position can be continuously adjusted. IR markers attached to T allow us to continuously track the pose of the environment S relative to T. (c) We compute a precise position of S with a second camera, which is tracked by the OptiTrack system.

We mount the camera onto an adjustable mount and adjust its position for each user. To track the camera pose, we have attached IR-reflective markers that can be tracked by an OptiTrack tracking system to T. We use Ubitrack [HPK*07] to calibrate the transformation ${}^T_O T$ which transforms a point ${}^O P$ in the OptiTrack coordinate systems to ${}^T P = {}^T_O T {}^O P$, the point P in the coordinate system of T.

We reconstruct ${}^O P_S$, the position of the checkerboard corners relative to O, with a PointGrey FL3-U3-13S2C-CS camera W with IR-markers attached to it (Figure 7(c)). We use Ubitrack again to compute ${}^W T$. We show the checkerboard on the monitor screen and detect the corners in images taken by W. We repeat this step for different camera poses. ${}^O P_S$ is the intersection of the backprojected rays from all images taken by W.

5. Results

We compare our hybrid approach with the method presented by Itoh and Klinker [IK14] (IK), as both methods are designed to recover the eye-pose from extracted ellipse edge points. We acquire 2D-3D correspondences of points in the camera image and the scene [PIN*15]. The iris contour is recovered from edge points detected by IK. Our method recovers the eye-pose from all detected edge points. For IK, we manually select a ROI slightly larger than the iris contour.

We conduct our evaluation on four male participants (two Asians, two Europeans; 22-31 years old with no vision impairments (participant 2 underwent a laser surgery)). The participants were asked to look at each inner corner of the checkerboard. For each participant, we recorded two sessions. Between the sessions, we changed the distance to the monitor, the position of the user's head and the eye-tracking camera. The distance to the monitor was 40 and 90 cm.

5.1. Personal Parameter Estimation

Although our method can estimate all relevant parameters of the model, we found that imprecisions in the corner detection and the fact that the cornea is not an ideal sphere, resulted in ambiguous solutions for r_C . Additionally, as the eye is located very closely to the camera T, changes in the cornea size did not impact the results of our method. Therefore, we use $r_C = 7.8$ mm and estimate d_{CL} . Note, that the estimated d_{CL} is up to scale of r_C . We show the results of the estimation in Table 2, where $r_L = \sqrt{r_C^2 - d_{CL}^2}$. For all participants, our method estimates that the size of the iris is as large or larger than the values assumed in previous work [NNT13]. We believe that this is a result of the gradual transition of the cornea into the sclera and the assumption that the iris and limbus are identical. This signifies the importance of the estimation of personal parameters. Our method estimates a stable radius d_{CL} for each recorded session. However, in the case of participants 1 and 4, this distance varied by more than 0.2 mm between the sessions. This suggests that the size has to be reestimated for the conditions present during the eye-pose estimation.

5.2. Eye-Pose Estimation

We compare three different methods to estimate the eye-pose: IK, our hybrid approach with a per-frame estimated size of the iris (HC) and a fixed iris size estimated for each session separately (HF).

We distinguish between HF and HC, because it may be necessary to reestimate the size of the iris to account for illumination changes. Our method achieves an accuracy of 3.63° with a standard deviation (stddev) of 1.37° for HC and 3.44° (stddev 1.23°) for HF. IK performs worse with an accuracy of 9.57° (stddev 6.16°).

For each session, we perform a calibration of (α, β) to determine the accuracy after alignment with the visual axis. We perform outlier removal for each session. Given the gaze errors e_i , $i = 1 \dots N$ for N frames, we determine the first quartile Q_1 and the third quartile Q_3 . The eye-pose estimated for frame i is an outlier, if $e_i < Q_1 - 1.5(Q_3 - Q_1)$ or $e_i > Q_3 + 1.5(Q_3 - Q_1)$. Out of 160 evaluated frames, three were removed as outliers for HF and six for IK. We estimate (α, β) for each session and user separately for HF and IK, and apply the values computed for HF to HC as well.

For the estimated visual axis, the eye-pose error is reduced for IK to 6.73° (stddev 8.15°), HA to 2.09° (stddev 1.49°), and HF to 1.74° (stddev 1.35°). We show the results for each session after outlier removal in Table 2 and display some of the estimation results in Figure 9. Overall, HF performs the best, followed by HA. KI falls short for all, but one sequence.

We have estimated a different offset of the visual and optical axes for the two session for each user. According to the two-sphere model, this value should be similar or identical.

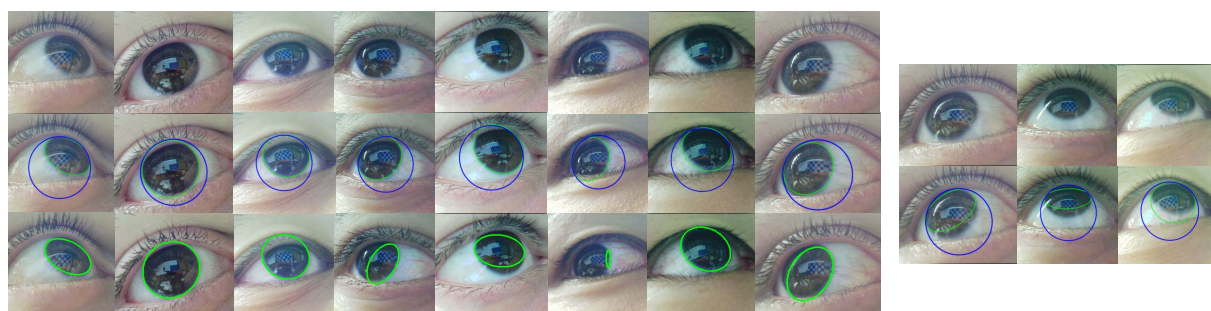


Figure 8: Results of the iris estimation. (top row) We show the cropped eye region within the captured images. We show the recovered iris contour with our method (middle row) and by [IK14] (bottom row). Our method successfully recovers the iris boundary for most cases. We show some of failure cases to the right.

We suspect that the difference is caused by our eye model, which does not perfectly represent the human eye. Another explanation could be that the camera had to be positioned at a much steeper angle, when the display was at a 40 cm distance to prevent it from occluding the screen. This is supported by the fact that for participants 2 and 4, the difference of the estimated angles is primarily along the vertical axis.

6. Conclusion

We have presented a novel method for gaze estimation from natural images. Our method assumes that a model of the environment relative to the eye camera is available and uses a hybrid approach, an adaptation of active gaze estimation methods for natural images, to first recover the eye position parameters. Based on this accurate estimation, we use detected candidates of the iris contour to reconstruct the iris on the corneal sphere. Our method robustly recovers the eye-pose even under extreme orientations of the camera and the user's gaze. Our solution uses the estimated corneal sphere to determine the iris size and can account for variances due to lighting conditions. We believe that by using our approach, it will be possible to recover all personal parameters, including a precise model of the eye. This could lead to eye-pose estimation results which are on par with active eye estimation methods.

Our current implementation recovers the orientation of the eye from edges detected in the image. To improve the robustness of our approach against incorrectly estimated edges, we want to include pose refinement from the image gradients as well as continuous tracking. Finally, to apply our system with AR or VR environments, we will explore how the reflection of a natural scene can be robustly detected within the corneal image.









Acknowledgements

This research was funded in part by Grant-in-Aid for Scientific Research (B), #15H02738 from Japan Society for the Promotion of Science (JSPS), Japan.

References

- [eye] URL: <http://theeyetribe.com>. 2
- [fov] URL: <http://www.getfove.com>. 2
- [GE06] GUESTRIN E. D., EIZENMAN M.: General theory of remote gaze estimation using the pupil center and corneal reflections. *IEEE Transactions on Biomedical Engineering* 53, 6 (2006), 1124–1133. 2
- [HJ10] HANSEN D., JI Q.: In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 3 (2010), 478–500. 1, 2
- [HPK*07] HUBER M., ET AL.: A system architecture for ubiquitous tracking environments. In *Proc. IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR)* (2007), pp. 211–214. 6
- [IK14] ITOH Y., KLINKER G.: Interaction-free calibration for optical see-through head-mounted displays based on 3d eye localization. In *Proc. IEEE Symposium on 3D User Interfaces (3DUI)* (2014), pp. 75–82. 2, 3, 6, 7
- [IMMR10] ISHIGURO Y., ET AL.: Aided eyes: Eye activity sensing for daily life. In *Proc. ACM Augmented Human International Conference (AH)* (2010), pp. 25:1–25:7. 2
- [IR11] ISHIGURO Y., REKIMOTO J.: Peripheral vision annotation: Noninterference information presentation method for mobile augmented reality. In *Proc. ACM Augmented Human International Conference (AH)* (2011), pp. 8:1–8:5. 1
- [Nit11] NITSCHKE C.: *Image-based Eye Pose and Reflection Analysis for Advanced Interaction Techniques and Scene Understanding*. PhD thesis, Graduate School of Information Science and Technology, Osaka University, Japan, January 2011. 3
- [NKB11] NORIS B., ET AL.: A wearable gaze tracking system for children in unconstrained environments. *Computer Vision and Image Understanding* 115, 4 (2011), 476–486. 2
- [NN06] NISHINO K., NAYAR S. K.: Corneal imaging system: Environment from eyes. *International Journal of Computer Vision* 70, 1 (2006), 23–40. 3

Table 2: Estimated personal parameters and eye-pose error:

Sample eye images								
Participant	1	1	2	2	3	3	4	4
Distance to the monitor[cm]	90	40	90	40	90	40	90	40
Personal Parameters								
d_{CL}, r_L [mm]	(4.60, 6.29)	(5.06, 5.92)	(5.03, 5.96)	(5.11, 5.87)	(5.14, 5.85)	(5.08, 5.91)	(4.97, 6.00)	(5.45, 5.56)
stddev d_{CL}, r_L	(0.11, 0.8)	(0.37, 0.34)	(0.14, 0.12)	(0.38, 0.35)	(0.29, 0.27)	(0.26, 0.23)	(0.21, 0.18)	(0.19, 0.18)
IK(α, β)	(0.09, 10.35)	(1.66, 7.47)	(6.31, 4.48)	(1.38, 4.73)	(0.64, 1.23)	(2.65, 12.93)	(1.88, 4.91)	(0.57, 3.53)
HF(α, β)	(1.98, 3.84)	(3.04, 2.07)	(2.42, 1.99)	(2.49, 0.29)	(1.32, 0.79)	(0.72, 0.83)	(2.65, 1.77)	(2.40, 0.10)
Error optical axis								
IK (mean, std)	(11.65, 3.67)	(8.97, 3.55)	(8.01, 1.40)	(6.89, 3.80)	(2.98, 1.44)	(12.57, 5.02)	(8.32, 3.45)	(6.05, 4.01)
HA (mean, std)	(4.24, 1.06)	(4.64, 1.21)	(3.46, 0.78)	(3.55, 1.50)	(1.89, 1.04)	(3.09, 1.83)	(3.39, 0.75)	(3.88, 1.40)
HF (mean, std)	(4.43, 1.03)	(4.32, 0.68)	(3.72, 1.19)	(3.23, 1.03)	(1.80, 0.69)	(2.68, 1.37)	(3.48, 0.66)	(3.21, 1.42)
Error visual axis								
IK (mean, std)	(4.03, 2.39)	(5.87, 4.30)	(2.76, 1.15)	(4.31, 3.07)	(2.70, 1.41)	(6.14, 5.07)	(7.29, 3.47)	(5.10, 3.48)
HA (mean, std)	(2.07, 0.79)	(3.07, 1.42)	(3.06, 1.25)	(2.24, 1.37)	(1.45, 0.67)	(2.69, 1.66)	(1.65, 0.69)	(3.23, 1.55)
HF (mean, std)	(1.42, 0.71)	(1.65, 1.08)	(2.00, 1.54)	(1.84, 1.06)	(1.28, 0.62)	(2.22, 1.02)	(1.20, 0.72)	(1.43, 0.88)

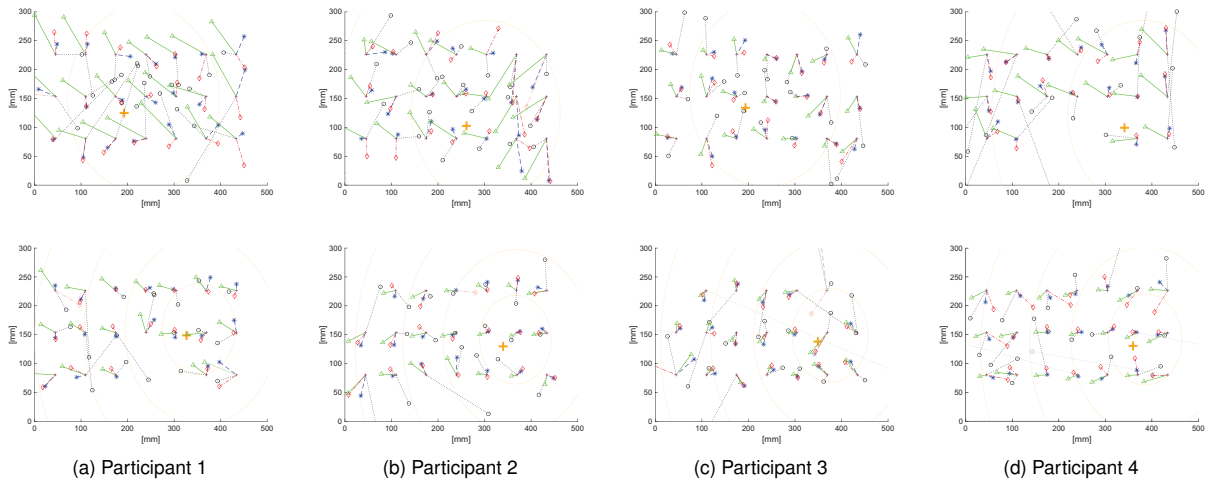


Figure 9: The PoG on the screen estimated by HF (green triangles), after applying the calibrated offset angles (α, β) to the estimation of HF (blue stars), and HA (red diamonds), as well as KI with the corresponding correction angles (black circle). Values assumed to be outliers are grayed out. The ground truth is shown as magenta crosses. Additionally, we show the PoG when the user is looking straight forward as the orange cross and draw contours around it in 10 deg increments as grayed out orange lines. (top row) The monitor is positioned 90 cm and (bottom row) 40 cm away from the participant.

[NN12] NAKAZAWA A., NITSCHKE C.: Point of gaze estimation through corneal surface reflection in an active illumination environment. In *Proc. European Conference on Computer Vision (ECCV)* (2012), pp. 159–172. 3

[NNT13] NITSCHKE C., ET AL.: Corneal imaging revisited: An overview of corneal reflection analysis and applications. *IPSJ Transactions on Computer Vision and Applications* 5 (2013), 1–18. 4, 6

[PIN*15] PLOPSKI A., ET AL.: Corneal-imaging calibration for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Virtual Reality 2015)* 21, 4 (April 2015), 481–490. 3, 4, 6

[PLC08] PARK H. M., ET AL.: Wearable augmented reality system using gaze interaction. In *Proc. IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR)* (2008), pp. 175–176. 1

[SF72] SLATER A., FINDLAY J.: The measurement of fixation position in the newborn baby. *Journal of Experimental Child Psychology* 14, 3 (1972), 349–364. 4

[SGRM14] STENGEL M., ET AL.: A Nonobscuring Eye Tracking Solution for Wide Field-of-View Head-mounted Displays. In *Eurographics 2014 - Posters* (2014), Paulin M., Dachsbacher C., (Eds.), The Eurographics Association. 1

[TJ00] TANRIVERDI V., JACOB R. J. K.: Interacting with eye movements in virtual environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2000), CHI '00, ACM, pp. 265–272. 1

[TK12] TSUKADA A., KANADE T.: Automatic acquisition of a 3d eye model for a wearable first-person vision device. In *Proc. ACM Symposium on Eye Tracking Research and Applications (ETRA)* (2012), pp. 213–216. 2, 4

[WBZ*15] WOOD E., ET AL.: Rendering of eyes for eye-shape registration and gaze estimation. *CoRR abs/1505.05916* (2015). 2

[WKW*07] WU H., ET AL.: Tracking iris contour with a 3d eye-model for gaze estimation. In *ACCV (1)* (2007), vol. 4843 of *Lecture Notes in Computer Science*, Springer, pp. 688–697. 2, 4