3D Ocular Ultrasound Using Gaze Tracking on the Contralateral Eye: A Feasibility Study

Narges Afsham, Mohammad Najafi, Purang Abolmaesumi, and Robert Rohling

University of British Columbia, Vancouver, BC, Canada
rohling@ece.ubc.ca

Abstract. A gaze-deviated examination of the eye with a 2D ultrasound transducer is a common and informative ophthalmic test; however, the complex task of the pose estimation of the ultrasound images relative to the eye affects 3D interpretation. To tackle this challenge, a novel system for 3D image reconstruction based on gaze tracking of the contralateral eye has been proposed. The gaze fixates on several target points and, for each fixation, the pose of the examined eye is inferred from the gaze tracking. A single camera system has been developed for pose estimation combined with subject-specific parameter identification. The ultrasound images are then transformed to the coordinate system of the examined eye to create a 3D volume. Accuracy of the proposed gaze tracking system and the pose estimation of the eye have been validated in a set of experiments. Overall system error, including pose estimation and calibration, are 3.12 mm and 4.68˚.

1 Introduction

Ultrasound has become an indispensible diagnostic tool for many ocular and orbital diseases [1-3]. Many studies show the ability of ultrasound to perform a pathological evaluation of the eyeball in its posterior segment [4, 5]. For example, in vitreo-retinal surgery, preoperative planning with ultrasound can be used [5]. Recent studies have also shown the benefit of ocular ultrasound in emergency medicine, such as diagnosing retinal detachment [6, 7]. There are some pathological characteristics that differentiate retinal detachment from other abnormalities such as vitreous hemorrhages [7]. Clinical manifestation of retina and vitreous detachments varies depending on where the adhesion is strongest [8]. These studies suggest that diagnosis, guidance and treatment would all benefit from imaging the entire posterior segment instead of a small portion.

Conventional ophthalmologic 2D ultrasound transducers have a small footprint and can image a portion of the posterior segment of the eye through the sclera. Since ultrasound has poor penetration of the cornea and lens, images are taken through the sclera when the gaze is deviated toward different positions [1]. However, the evaluation of pathological features using such 2D ultrasound makes the examination more complex and may contribute to false diagnoses [7]. In order to obtain views of the entire posterior segment, multiple ultrasound images must be captured for a variety of gaze directions. Moreover, the use of different gazes and the spherical geometry of the eyeball make it difficult to interpret the pose of each 2D ultrasound

relative to the eye's coordinate system. The complexity of such examinations comes from the fact that the 3D pose of the eye should be known for each image acquisition [2]. In practice, the examiner has to visualize the intersection of the ultrasound plane with a mental 3D model of the eye and guess the eye pose to estimate where the absolute position of the image is located in the coordinate system of the eye [2]. There is a need for a new solution for this problem, especially since the eyelid is closed during the examination, so the eye pose is difficult to discern visually.

It has been suggested that a dedicated 3D ultrasound transducer facilitates the assessment of certain posterior ocular abnormalities [9]. However, even in the case of a specialized 3D ultrasound transducer, only a portion of the posterior surface would be imaged from each pose of the eye and it does not eliminate the localization problem associated with a gaze-deviated examination. Moreover, this challenge cannot be solved by tracking the ultrasound transducer relative to a fixed base because the eye changes its pose during the examination.

We propose a novel system using a combination of a camera, an ultrasound transducer, and a subject-specific eye pose estimation method to tackle this problem. The pose of the examined eye is estimated from the pose estimation of the contralateral eye during the imaging process. Eye parameters for each subject and in different lightening conditions are modeled together with pose estimation in an attempt to achieve high localization accuracy. The method of model-based gaze tracking proposed in [10] is adopted and modified to replace the parameters based on population averages, such as cornea radius and pupil radius, with parameters specific to the subject. The proposed system aims to produce accurate 3D ultrasound reconstruction of the posterior segment of the eye from a set of 2D images acquired from different eye gazes.

2 Methods and Materials

2.1 Overview

As mentioned, the goal is to find the transformation from the ultrasound image to the examined eye's coordinate system. This transform, $^{E2}T_{US}$, can be defined as:

$$^{E2}T_{US} = {}^{E2}T_{E1}\,{}^{E1}T_C\,{}^{C}T_{US} \tag{1}$$

where $^{C}T_{US}$, $^{E1}T_C$ and $^{E2}T_{E1}$ are the transformations from ultrasound image to camera, camera to Eye1, and Eye1 to Eye2, respectively (Fig. 1(a)). Eye2 is the examined eye and Eye1 is the contralateral eye. $^{C}T_{US}$ is a fixed transformation that is found by a separate calibration process. $^{E1}T_C$ is determined for each pose of Eye1 by gaze tracking together with estimating subject-specific eye parameters. These parameters are the cornea center, c, cornea radius, r_c, pupil center, p, pupil radius, r_p, and cornea-center-to-pupil-center distance, r_{cp}. The optical axis is a vector that joins c and p, and the visual axis passes through c and hits the fovea (Fig. 2(a)). To find $^{E2}T_{E1}$ some parameters should first be calibrated in an experiment with two open eyes for each individual. One is the transform from the rest pose of Eye1 to Eye2, $^{Er2}T_{Er1}$. The rest pose is defined as the eye pose when it is looking at a distant target and head is upright and static [11]. The other is the angular offset between the visual axis and the optical axis, β_v [12] (Fig. 2(a)). The visual axes of the two eyes intersect at
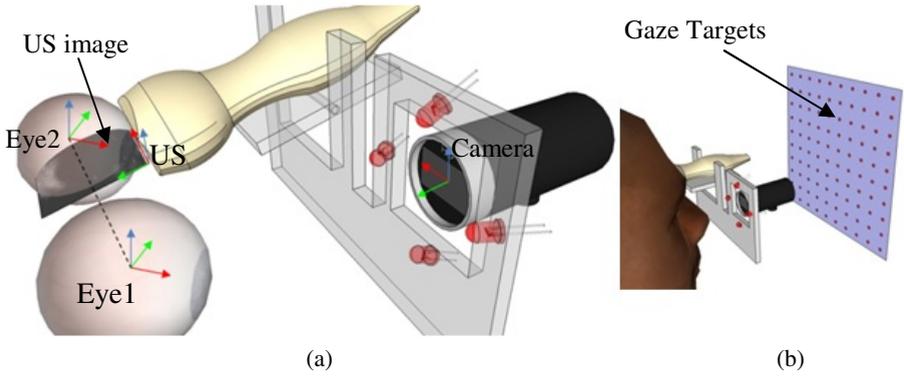
Fig. 1. System overview. (a) Each US image should be transformed into the coordinate system of Eye2. The pose of Eye1 is tracked in the coordinate system of the camera using glint tracking from four LEDs. The US to camera transform is fixed during the experiment. The transform from Eye1 to Eye2 is determined by a subject-specific model. (b) A subject fixates at several target points to create a wide range of eye movements during an examination.

the fixation points. Considering this fact for each target point, and having $^{Er2}T_{Er1}$ and the angular offset for both eyes, β_{v1}, β_{v2}, the Eye1 to Eye2 transformation, $^{E2}T_{E1}$, can be determined. The detailed calculations of these stages are described in the following sections. Given space limitations only the key equations are described in full.

2.2   Calibration

The subject is asked to look at a distant point with both eyes open. The rest position of both eyes is found from gaze tracking of the two eyes with a single camera placed farther from the face and $^{Er2}T_{Er1}$ is measured using gaze tracking on both eyes as a subject is guided to a distant target. Then the subject is asked to look at several known target points on a board and the optical axis is found for each target (Fig. 1(b)). Then the method as described in [13] is used to calculate β_{v1} and β_{v2}.

For ultrasound to camera calibration, $^{C}T_{US}$, the single-wall calibration method [14] is used with the modification of using a checker-board pattern on the portion of the wall surface outside the water bath so that it can be seen by the camera. The camera's intrinsic parameters are calibrated with the *Camera Calibration Toolbox* [15], which uses images of a checker-board pattern.

2.3   Combined Pose Estimation and Eye Model Parameter Identification

The primary application of gaze tracking is in the field of human-computer interfaces. In the video-based gaze tracking literature, those that give the full 3D pose of the eye, instead of simply a gaze intersection with a monitor, are of interest here. We have implemented a gaze tracking system based on a single camera and multiple IR light sources (LEDs) [10]. In this method, the glints are used to estimate the position of cornea center. The circumference of the pupil in the image is refracted into the eye according to Snell's law. With a known position of the cornea, it is possible to
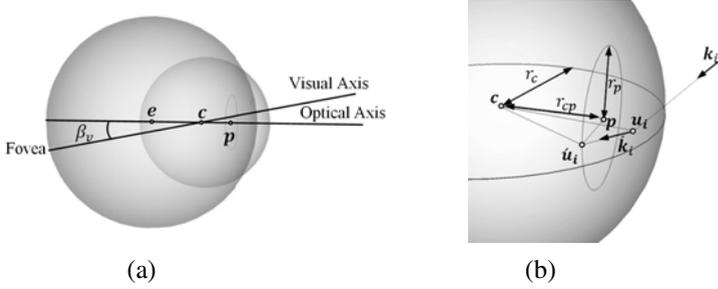
Fig. 2. (a) 3D model of the eye. e, c, and p are eye, cornea, and pupil centers (b) The refraction at the cornea surface is shown. k_i is the incident ray and k_i' is the refracted ray.

calculate the center of the pupil. The calculations of the cornea center and pupil center are the same as explained in [10] with some modifications to make it subject-specific by including identification of r_p, r_c and r_{cp} in the calculations.

Geometric Eye Model. Fig. 2 (a) shows the geometry of the eye modeled as two spheres, the eyeball and the cornea [12]. As a result of a fixed head position, the center of the eyeball, e, remains fixed during the experiments. This confines the transformation of the eye from one position to another as a pure rotation. By knowing the optical axis and applying Listing's law, it is possible to determine the full pose of the eye [11]. Listing's law states that the eye's axis of rotation from the rest pose to any arbitrary pose is restricted to a single head-fixed plane, known as Listing's plane [16]. To determine the optical axis for each pose, the cornea and pupil centers should be estimated. Finally $^{E1}T_C$ and $^{E2}T_{E1}$ are calculated based on the equations in [11].

Cornea Parameter Estimation. There exists an auxiliary coordinate system in which the xz plane includes the camera origin, o, light source, l_i, cornea center, c, its corresponding glint in the image, g_i, and its 3D position, g_i' [10] (see Fig. 3(b)). The subscript i ranges from 1 to 4 light sources. Each light source vector resides along the x direction of its auxiliary coordinate. The transform from the camera coordinate to the auxiliary coordinate is defined by a rotation matrix, R_i, in terms of l_i and c_i' (cornea center in the auxiliary coordinate) such that $c = R_i c_i'$.

$g_i' d_i$ is the bisector of the $og_i' l_i$ angle based on the reflection theorem. As a result of the bisector equation, one may write:

$$\frac{\sqrt{g_{ix}'^2 + g_{iz}'^2}}{d_i} = \frac{\sqrt{(g_{ix}'-l_i)^2 + g_{iz}'^2}}{l_i - d_i} , \quad \xrightarrow{yields} \quad d_i = \frac{l_i \sqrt{g_{ix}'^2 + g_{iz}'^2}}{\sqrt{g_{ix}'^2 + g_{iz}'^2} + \sqrt{(g_{ix}'-l_i)^2 + g_{iz}'^2}} . \tag{2}$$

The triangulation for $d_i g_i' h_i$ and $g_i' c_i' f_i$ results in:

$$\frac{r_c}{r_c + \sqrt{(g_{ix}'-d_i)^2 + g_{iz}'^2}} = \frac{c_{iz}'-g_{iz}'}{c_{iz}'} , \text{ and} \tag{3}$$

$$\frac{r_c}{r_c + \sqrt{(g_{ix}'-d_i)^2 + g_{iz}'^2}} = \frac{c_{ix}'-g_{ix}'}{c_{ix}'-d_i} . \tag{4}$$

where r_c is the cornea radius. Since the angle between ol_i and og_i is known from the pinhole model of the camera, it gives another equation:

$$\frac{g'_{iz}}{g'_{ix}} = tan\,\alpha_i .$$
(5)

Eqs. (2) to (5) can be written for each light source and they result in $4n$ equations (n = number of light sources). Since the auxiliary coordinate system has the same origin as the camera, the transform from the camera coordinate system to the auxiliary system is defined by the rotation matrix, $\boldsymbol{R}_i$. Another set of equations is produced by the constraint that the transformation of auxiliary cornea centers to the camera coordinate system are all the same:

$$\boldsymbol{R}_i^{-1}\, \boldsymbol{c}'_i = \boldsymbol{R}_i^{-1}\, \boldsymbol{c}'_j .$$
(6)

In summary, by considering n light sources, the number $(4n + 3\binom{n}{2})$ of equations equals or exceeds the number $(5n + 1)$ of unknowns $(d_i,\ g'_{ix},\ g'_{iz}, c'_{ix}, c'_{iz}, r_c)$. Although two light sources are sufficient to solve for the cornea center, we used four sources to get more accurate results [13].
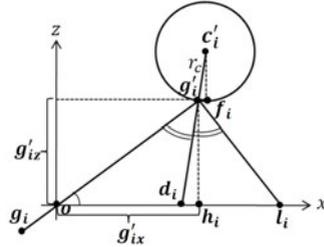
(a)   (b)

Fig. 3. (a) An example of camera image with four glints (blue) and extracted pupil perimeter (red) (b) Auxiliary coordinate system. $\boldsymbol{c}'_i$ is the auxiliary cornea center. $\boldsymbol{g}'_i$ is the glint of light source l_i in the image, and $\boldsymbol{g}_i$ is its 3D position.

Pupil Parameter Estimation. Fig. 2(b) depicts the geometry of the pupil center estimation problem. Instead of fitting an ellipse on the image of the pupil and estimating the pupil center from its refracted center in the image, we directly solve for pupil center as well as pupil radius, r_p, and the distance between pupil-cornea centers, r_{cp}. The direction of the vector from a point residing on the circumference of the pupil to the point of refraction on the surface of the cornea (k_i) is known from the pinhole camera model, so the intersection of the ray with the cornea sphere surface (u_i) is easily determined. $\boldsymbol{k}'_i$ is the direction of the refracted ray and it is known by considering Snell's law and the cornea's refractive index ($\boldsymbol{n_{cornea}}$ =1.336). The following gives three equations for each point on the pupil circumference, where $\boldsymbol{s}_i$ is an unknown parameter:

$$\boldsymbol{u}'_i = \boldsymbol{u}_i + s_i \boldsymbol{k}'_i.$$
(7)

u_i' is the intersection of the refracted ray and pupil perimeter , so we have:

$$\|u_i' - p\|^2 = r_p^2. \tag{8}$$

The constraint between the pupil perimeter and the cornea center gives:

$$\|u_i' - c\|^2 = r_{cp}^2 + r_p^2. \tag{9}$$

By adding the constraint of the distance between cornea center and pupil center:

$$\|c - p\|^2 = r_{cp}^2. \tag{10}$$

By considering Eqs. (7) to (10) for at least four points ($n \geq 4$) on the pupil, the number ($5n + 1$) of equations equals or exceeds the number ($4n + 5$) of unknowns (u_i', s_i, p, r_p, r_{cp}). Since the camera is close to the eye, the pupil perimeter points are easily labeled using Canny edge detection and if some portion of the pupil perimeter are missing in the image, a solution can still be found. In practice, at some far deviated gazes, the eyelid partly hides the perimeter of the pupil.

2.4 Experiments

The experimental setup consists of ultrasound imaging and gaze tracking components. The subject is positioned with a standard chin and forehead support to keep the head fixed related to the camera and the ultrasound transducer. The ultrasound system used is a Sonix MDP (Ultrasonix Medical Corporation, Richmond, BC, Canada) equipped with an L14-5/38 10 MHz linear 2D ultrasound transducer.

The ultrasound transducer is positioned on the eyelid using coupling gel, as standard practice. The transducer can be tilted around the elevation axis to obtain different imaging views. Then the transducer is fixed to the suitable position with respect to the chin rest. The gaze tracking part includes a high definition USB camera[1] (LifeCam Cinema, Microsoft Co., USA) and four infrared LEDs to make four bright glints in the image of the eye.

From Eq. 1, it is clear that the accuracy of the camera to ultrasound calibration, $^C T_{US}$, can be easily separated from the gaze tracking results. To evaluate our overall system accuracy, we therefore give these as three separate components of the total error. In this way, the errors from our eye pose estimation $^{E2} T_{E_1}$ and $^{E1} T_C$ can be compared to a gold standard measurement of $^{E2} T_C$ using eye tracking directly on the examined eye with the eyelid open. It also allows direct comparison of our novel combined modeling and tracking method with standard gaze tracking methods using parameters based on population averages.

Ultrasound to camera calibration is therefore performed first. Then in an experiment with two eyes open, the subject pursues target points on a board. The target points are evenly distributed (2 cm spacing) and the plane is 25 cm away from the camera. As the subject fixates at each target point, the gaze tracking system tracks the optical axis of the eyes and β_{v1} and β_{v2} are calibrated. Also, the rest pose of the eyes and $^{Er2} T_{Er1}$ are determined using a distant gaze.

To evaluate $^{E2} T_{E1}$ estimation, first it is directly calculated from the optical axis of both eyes. Then, it is estimated based on the contralateral eye from the fixation point

---

[1] The infra-red filter was removed.

(a)   (b)
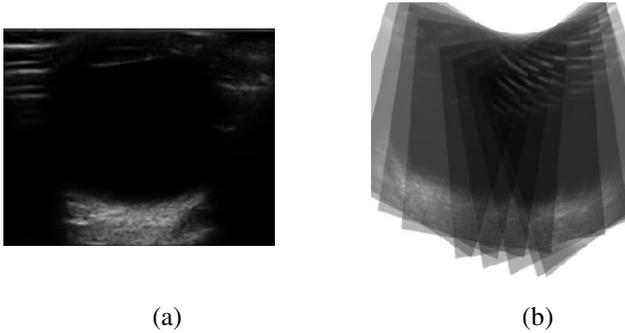
Fig. 4. (a) An example 2D ultrasound image of the posterior segment of the eye. (b) Ultrasound frames transformed in the coordinate system of the eye for a set (n=15) of eye gazes.

and the calibrated parameters. The fixation point is found from the intersection of Eye1's visual axis with the target plane. The unknown visual axis of Eye2 passes through its cornea center and the fixation point. The cornea center, the optical and the visual axes can be related to their rest pose with a single transform. This transform is found so that the visual axis passes through the target point. The difference in the eye orientation for $^{E2}T_{E1}$ between our method and the gold standard are then calculated. The equations are solved with nonlinear least squares using the Levenberg-Marquardt method. As a feasibility study, one subject is examined and the errors are calculated (Table 1). Since the head is fixed, the eye center, e, remains fixed. The standard deviation in estimation of e is also calculated as a measure of precision of the gaze tracking method.

3   Results

The residual errors (the minimized error) from the ultrasound to camera calibration are 0.2, 0.3, 0.5 mm, 0.2˚, 0.2˚, and 0.6˚. The accuracy of a single-wall calibration technique based on minimizing the residual errors is well documented [14]. Table.1 summarizes the differences in $^{E2}T_{E1}$ from the gold standard. The standard deviation of the error in β_{v1} and β_{v2} estimations are 1.28˚ and 1.58˚, respectively. RMS error in eye center estimation for n=30 target points with our method is 2.91 mm compared to 3.72 mm error with model-based method using parameters based on population averages. Adding the errors of independent sources in Eq.1 gives an overall accuracy of 3.12 mm and 4.68˚.

By positioning each ultrasound image in the coordinates system of the eye for each pose, a 3D image of the eye is constructed. Fig. 4 demonstrates that the angular view of the posterior segment of the eye has been increased.

Table 1. The differences in $^{E2}T_{E1}$ from the gold standard in deg

Estimation errors of rotation parameters of $^{E2}T_{E1}$	α (Roll)	β (Pitch)	γ (Yaw)
Mean	2.04	-0.32	-1.42
STD	3.15	1.91	1.61

4   Discussion and Conclusion

In this work a novel system has been proposed to determine the absolute pose of a 2D ultrasound image in the 3D coordinate system of the eye in a gaze-deviated eye examination. To this end, a subject-specific eye pose estimation method has been developed. The method is based on single camera eye gaze tracking.

The accuracy of the overall system is mainly governed by the accuracy of the camera to ultrasound calibration and the accuracy of eye pose estimation from the contralateral eye. Experimental results show errors 3.12 mm and 4.68˚compared to the gold standard. The gold standard is eye gaze tracking directly on the examined eye, and itself has an error of approximately 1˚[10]. The main sources of error are related to camera quality, position of the light sources and eye modeling.

Accurate 3D reconstruction of the eye's posterior segment and the assessment of dynamic ultrasound images of the eye are the other future discussion for this research.

References

1. DiBernardo, C.W., Schachat, A.P., Sharon, Fekrat, M.D.: Ophthalmic Ultrasound: A Diagnostic Atlas. Thieme Medical Publishers (1998).
2. Garcia Jr, J.P.S., Finger, P.T., Rosen, R.B.: Dynamic Ophthalmic Ultrasonography: A Video Atlas for Ophthalmologists and Imaging Technicians. Lipp. Williams & Wilkins, Baltimore (2009)
3. Fledelius, H.C.: Ultrasound in ophthalmology. Ultrasound Med. Biol. 23, 365–375 (1997)
4. Fisher, Y., Hanutsaha, P., Tong, S., Fenster, A., Mazarin, G., Mandava, N.: Three-dimensional ophthalmic contact B-scan ultrasonography of the posterior segment. Retina 18, 251–256 (1998)
5. Muldoon, T.O., Will, D.V., Garcia, J.P., Walsh, J.B., Rosen, R.B., Rosenthal, J.L., Gentile, R.C.: 3D B-mode Ultrasound of the Vitreous Prior to Vitrectomy Surgery. Invest. Ophthalmol. Vis. Sci. 44, 3625 (2003)
6. Shinar, Z., Chan, L., Orlinsky, M.: Use of Ocular Ultrasound for the Evaluation of Retinal Detachment. The Journal of Emergency Medicine 40, 53–57 (2011)
7. Yoonessi, R., Hussain, A., Jang, T.B.: Bedside Ocular Ultrasound for the Detection of Retinal Detachment in the Emergency Dept. Acad. Emerg. Med. 17, 913–917 (2010)
8. Sebag, J.: Anomalous Posterior Vitreous Detachment: A Unifying Concept in Vitreo-Retinal Disease. Graefes Arch. Clin. Exp. Ophthalmol. 242, 690–698 (2004)
9. Downey, D.B., Nicolle, D.A., Levin, M.F., Fenster, A.: Three-dimensional Ultrasound Imaging of the Eye. Eye (Lond) 10, 75–81 (1996)
10. Hennessey, C.: A Single Camera Eye-Gaze Tracking System with Free Head Motion. In: Proceedings of etra: Etra Symposium, pp. 87–94 (2006)
11. Horaud, R., Hansard, M.: Cyclorotation Models for Eyes and Cameras. IEEE Transaction on Systems Man and Sybernetics 40, 151–161
12. Villanueva, A., Daunys, G., Hansen, D.W., Böhme, M., Cabeza, R., Meyer, A., Barth, E.: A Geometric Approach to Remote Eye Tracking. Univ. Access Inf. Soc. 8, 241–257 (2009)
13. Villanueva, A., Cabeza, R.: Models for Gaze Tracking Systems. In: EURASIP, pp. 1–16 (2007)
14. Prager, R.W., Rohling, R.N., Gee, A.H., Berman, L.: Rapid Calibration for 3-D Freehand Ultrasound. Ultrasound in Med. & Bio. 24, 855–869 (1998)
15. Bouguet, J.-Y.: Visual Methods for Three-dimensional Modeling, Phd Thesis (1999)
16. Hepp, K.: Theoretical Explanations of Listing's law and their Implication for Binocular Vision. Vision Res. 35, 3237–3241 (1995)