International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2016)
D. Reiners, D. Iwai, and F. Steinicke (Editors)

An efficient interpolation approach for low cost unrestrained gaze tracking in 3D space

Christian Scheel[† 1], ABM Tariqul Islam[‡ 1] and Oliver Staadt[§ 1]

[1] Institute for Computer Science, University of Rostock, Germany

Abstract

We present a first attempt to use interpolation based approach to combine a mobile eye tracker with an external tracking system to obtain a 3D gaze vector for a freely moving user. Our method captures calibration points of varying distances, pupil positions and head positions/orientations while the user can move freely within the range of the external tracking system. For this approach, it is not necessary to know the position of the eye or the orientation of the eye coordinate system. In addition to the calibration of the external tracking system, we can calibrate the head-tracked eye tracker in a one-step process which only requires the user to look at the calibration points. Here, we don't need any extra calibration of the eye tracker, because the raw pupil position from the eye tracker can be used. Moreover, we use low cost tracking hardware which might be affordable to a wide range of application setups. Our experiment and evaluation show that the average accuracy of the visual angle is better than 0.85 degree under unrestrained head movement with a relatively low cost system.

Categories and Subject Descriptors (according to ACM CCS): I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Tracking,Sensor fusion I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques

1. Introduction

Eye gaze information plays an important role in estimating user's point-of-interest on a given scene or object. Hence, there have been growing interest to support gaze awareness in many application fields, such as human-computer interaction, usability advertising, remote collaboration (telepresence) or industrial design. Recent advances in telepresence systems, such as integration of large high-resolution display (LHRD), aim at supporting gaze awareness for the user, at remote locations, who can move independently.

However, most eye gaze tracking systems do not allow the user to move freely, because they require a static head position and orientation. Thus, they are not suitable for many human-computer interaction systems, for example a telepresence system with LHRD setup. Our goal is to devise a system for eye gaze tracking that supports free movement of the user by combining a mobile eye tracker with an additional external six-degree-of-freedom (6-DOF) tracking system.

We propose a new and simple calibration procedure for such a system based on observing multiple calibration points at varying distance to the user. In our novel calibration system, the user is allowed to move freely, while and after the calibration procedure, within the tracking volume of the external tracking system (e.g., maker-based infrared camera tracker). We can estimate the eye position and the 3D gaze vector in real time without requiring manual measurements, purkinje images, or glints. Moreover, our proposed approach does not require the assumptions about the anatomy/geometry of the eye; which portraits our system, to the best of our knowledge, as the first attempt of using interpolation-only (non geometrical) approach in unrestrained gaze tracking in 3D space.

Our contributions, in this paper, towards developing a novel eye gaze tracking approach are:

- Introducing a new interpolation based eye gaze tracking approach for freely moving users.
- Adapting the Radial Basic Function (RBF) interpolation with a Thin Plate Spline (TPS) kernel for the first time in eye tracking. RBF with TPS has been used in other fields, but this is the first attempt to apply it in eye tracking area.
- Achieving a high accuracy of gaze tracking; the accuracy is nearly as high as current geometrical approaches.
- Using a low cost eye tracker and a low cost optical tracking system.

† cscheel@gmx.de
‡ tariqul.islam@uni-rostock.de
§ oliver.staadt@uni-rostock.de

DOI: 10.2312/egve.20161427

2. Related work

Existing eye gaze tracking methods can be divided primarily into two approaches – interpolation approach and geometrical approach [ZJ07, CVC08, HJ10]. They are also often referred to as mapping (interpolation) and model-based (geometrical) approaches. Based on [ZJ07, CVC08, HJ10], these approaches are defined as follows. The interpolation or mapping approaches try to map eye image features to a 2D or 3D gaze point. They use, for example, polynomials or neural networks for the interpolation. Geometrical approaches, instead, calculate the gaze point from the eye features based on a geometrical model of the eye.

Interpolation approaches, in general, model the optical properties, geometry and the eye physiology indirectly and hence, they do not require a calibration of the whole system and, moreover, they are relatively easy to implement [CVC08, HJ10]. There exist quite a few eye tracking methods (e.g. [BM04, HHN*02, HP05, JY02, TKA02, PCG*03, KR04, ZJ04, WBC06, EPR06, ZJ07]) which use interpolation approach, but none of them supports user's free movement in terms of walking around and, as far as we know, no interpolation approaches possess the attribute of supporting user's free movement.

There also exist research works, on eye gaze tracking based on geometrical approach, where most of them (e.g., [VSG12, KBB*08, NSI*10, GE06, WSV05, VCP06]) do not support user's free movement except only a few (e.g., [AEC96, DMG*01, HAJ04, RWL07, HCT*08, EDP*12, CvdLLd13]). Moreover, current geometrical approaches have the limitation that the eye position is estimated based on anatomical data. In other words, in case of geometrical approaches, the relation between head and eye coordinate systems is determined with the help of optimization algorithms which need initial values or other means of assumptions are used for the eye geometry. Besides this, they are quite difficult to construct.

Therefore, because of the advantages in calibration process and easy implementation feature, eye gaze tracking by interpolation approach is gaining much popularity. Nevertheless, we did not see any work for eye gaze tracking by interpolation based approach that supports user's free movement. This is, perhaps, due to the fact that, until now, interpolation based approaches have provided lower accuracy than the geometrical approaches. However, Cerrolaza et al., in [CVC08], state that the accuracy of geometrical and interpolation approaches can be very similar. Hence, to build a eye gaze tracking system which have superior calibration procedure and easier implementation process, we present here the first interpolation approach that supports free movement and aims to achieve an accuracy that is comparable to geometrical approaches.

Similar to our approach, systems with geometrical approach which support free user movement, estimate the 3D gaze vector after a calibration process. In those systems, often a combination of a mobile eye tracker and a motion capturing system is used. To support free movement, these system often use a 6-DOF tracking system to obtain the geometry. Despite that, the challenge remains in establishing the relationship between the head and the eye coordinate system. The head is tracked with a motion capturing system, but the eye position and orientation remains unknown.

In our proposed interpolation based approach, the eye position is estimated directly from the calibrated system and there is no need to know the eye position or the orientation of the eye coordinate system. Furthermore, no additional parameters have to be measured or specified by the user.

3. Problem description

Zhu et al. [ZJ04] define gaze tracking as: "Gaze tracking is the procedure of determining the point-of-gaze in the space, or the visual axis of the eye"; here, the *point-of-gaze* is the point which a user have fixeated with his eyes. However, getting this point and accordingly, the visual axis is quite challenging. This is due to the limiting attribute of today's image based eye trackers where these trackers are not so invasive than the electro-oculography or magnetic scleral search coils, and they are not able to look into the eye and see the fovea and other things inside [Mod11]. So, these eye trackers are only able to see the eye from the outside and deliver, for example, the position of the pupil center in the image of the eye tracker.

As a result, with this information from these trackers, it's only possible to get the pupil axis but not the optical or visual axis. Fig. 1 depicts and illustrates a more precise description of the axis based on Nowakowski et al. [NSNG12]. To get the visual axis, which is the aim of gaze tracking, a calibration is required. This is done by looking at known positions and then extracting the information about the difference between the pupil and the visual axis.

Geometrical approaches, for eye tracking, often make simplified assumptions about various components of the complex structure of the eye, for example, about the spherical cornea or coincidence of the axes [PPR11]. Moreover, geometrical approaches usually also assume that the eye position relative to the head is fixed when the eye is looking in different directions, but in reality, the eyes move in relation to the head [PP33]. Moreover, neither the geometrical approaches nor the interpolation based approaches support user's free movement, in terms of walking around, while and after the calibration is pursued for tracking the eyes.

Another issue, usually seen with existing approaches of unrestrained gaze tracking, is the usage of very expensive tracking hardware. Cesqui et al. [CvdLLd13], for example, provide very good results, but they use a tracking system and an eye tracker which cost approximately 230,000$ in total. Such an expensive hardware setup might not be affordable to a variety of application scenario. In the proposed approach, we use a tracking setup whose cost lies below 10,000$ (eye tracker for approx. 1,100$ and a tracking system for approx. 8,000$). Although, 10,000$ still seems much, but comparing to other approaches it's relatively cheaper (23 times cheaper than Cesqui et al. [CvdLLd13]) and we get results which are close to the top state of the art approaches like Cesqui et al. [CvdLLd13].

In this paper, we address these issues by proposing an unrestricted eye gaze tracking method which is based on an interpolation based approach by using low cost hardware.

4. Our approach

Our aims, in this work, is to obtain a 3D gaze vector without restricting the user's movement and to use low cost hardware compared to other approaches. We propose to use an interpolation based

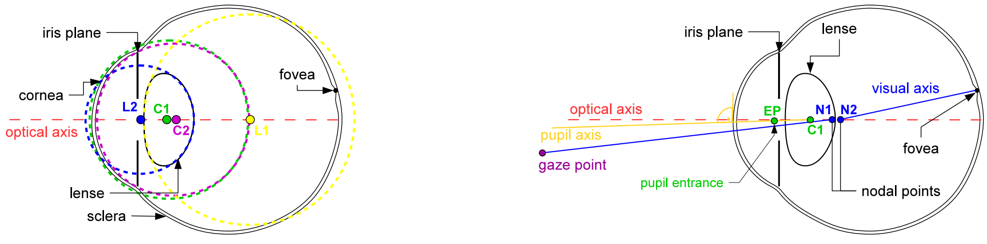


Figure 1: *Structure of the eye according to [NSNG12]. (a) Circles and middle points from the outer (green,C1) and inner (purple,C2) curvature of the cornea and from the outer (yellow, L1) and inner (blue, L2) curvature of the lense. Also shown is the optical axis which is defined as a line, or more precisely as a line of best fit, which goes through C1, C2, L1 and L2. (b) pupil axis: is the line defined through the pupil entrance EP and C1 (see a) and perpendicular in relation to the outer cornea; visuel axis: the connection between the gaze point and the fovea that go to the both nodal points N1 and N2.*

approach, which has not been explored before for supporting user's free movement, to get the 3D gaze vector.

As a first step, a combination of eye tracking and external head tracking is required to obtain a 3D gaze vector. To pursue that, we conduct a new calibration procedure for a monocular eye tracker and a 6-DOF head tracking system. Calibration points with known positions relative to the tracked head are captured with their corresponding pupil position from the eye tracker. The calibration points are captured at different distances. Based on these data, we estimate a gaze vector relative to the head position. The advantage of our proposed approach is that we use at least two points at the gaze vector, rather than using only one point, for the interpolation. With at least two points on the same gaze vector or respectively on the same line of sight we can calculate a 3D gaze vector.

To allow free movement and to get this 3D gaze vector in a world coordinate system, we use a combination of a 6-DOF tracking system (for tracking the head) and the eyetracker. We use the data from the head tracking to translate the gaze vector into a world coordinate system. We obtain the pupil position from the eye tracker. With the pupil position, the data from the calibration points and the head tracker, we calculate a 3D gaze vector in real time.

It is worth to mention that we do not require the orientation of the eye coordinate system. Moreover, in our proposed method, we make no assumptions about the eye geometry or position. So, all the errors that generate through assumptions about the eye geometry and position, can be ignored. However, errors might also originate from the interpolation method we choose to use. Therefore, to keep the error as low as possible level, we focus on attaining an interpolation method to approximate a good gaze point. Here, we use a radial basis function (RBF) with a thin plate spline (TPS) kernel for the interpolation.

In the following subsections, we describe the coordinate systems we use, explain our proposed calibration procedure, outline the calculation details to obtain the gaze vector, and illustrate the procedure to calculate the eye position respectively.

4.1. Coordinate systems

We use two different coordinate systems in our approach – a world and a head coordinate system. The world coordinate system is originated through the calibration of the 6 DoF tracking system. The head coordinate system is defined through the markers at the eyetracker and lay inside the world coordinate system.

In our approach, it is not necessary to calibrate or align the head coordinate system in any way to the world coordinate system or to any other part of the real world. We only need the information about the position and orientation inside the world coordinate system and the eyetracker with the markers should not be moved relative to the head.

4.2. Calibration procedure

For the calibration process, the user have to stand in front of a calibration point grid which consists of 3×3 calibration points, and fixate on those calibration points. The user needs to repeat this process a total of n times, each time standing at a different distance from the calibration point grid. For our procedure, we need the user to repeat the process for at least 2 times.

For each distance a grid with another size is used. This is to get the angle between the left and right column of the grid more or less the same for the user for each distance. So at the biggest distance the biggest grid is used and at the smallest distance the smallest grid is used. The approach will also work if grid size would stay the same. But it's more practical to use different sizes because if the user stands very close to the grid it's very large and if he stands very far it's become very small.

The user is advised to look straight to the middle of the calibration points and keep his head still while he is looking at the calibration points. We recommend to keep the head still for getting a wider range of eye movement for the calibration. However, we do not use any external apparatus to keep the user's head still and hence, the user can relax and make small movement's of his head and body during the calibration. Since the head is tracked and each calibration point is captured in the head coordinate system, small movement of user's head doesn't affect our procedure.

(a) (b)

Figure 2: *(a) For each layer a grid of* 3×3 *calibration points is taken. In reality we have the problem with measurement errors, that the eye position is not fixed relativ to head and that belonging calibration points not at one line of sight. Because of the problems in the reality it's necessary to do interpolation to get calibration points at the same line of sight. (b) Gaze vector interpolation: For each calibration point layer a gaze vector point (*$\mathbf{v}_p$*) will be calculated. The gaze vector is represented by a line through all these gaze points.*

We also tried the calibration procedure with natural head movement. In most of the cases natural head movement produce better results. We think this occurs due to the fact that if the user looks at the calibration points by moving the head naturally during both the calibration process and in the application, the pupil position value are more similar. And, consequently, if values are similar a interpolation provides better results. However, if a user moves his head in a strange way, in one of our experiments a user moved his head in such a way that the pupil position was the same for each calibration point, there is a high probability to obtain a bad and numerically unstable calibration / gaze vector.

4.3. Gaze vector calculation

In our approach, after a successful calibration, we are able to determine the current gaze vector in realtime. For the calculation of the current gaze vector, we use the current pupil position from the eye tracker, the head position and orientation, and the calibration data. We use a three-step process to determine the gaze vector. In the first step, the gaze vector is calculated in the head coordinate system, in the second step, this gaze vector is transferred to the world coordinate system and finally, in the third step we calculate the current gaze point.

In the first step, we use the current pupil position to interpolate **n** points at the gaze vector with the help of the calibration points in the head coordinate system. For each calibration point layer, one gaze vector point ($\mathbf{v}_p$) is interpolated for the current pupil position, see Figure 2. For the gaze vector calculation, we need at least two gaze vector points which should lie on the line of sight of the user. Here, we call the line of sight of the user as *gaze vector* which goes through all the gaze vector points, see Figure 2.

The gaze vector points are normaly not lying at one line. So we have to find the line of best fit for the gaze vector points. To get the line of best fit we calculating the average of all gaze vector points. Then we subtract this average from each gaze vector point and put this points in a matrix **P**, where each point is put into one column. After that we calculating the covariance matrix $\mathbf{C} = PP^t$. Now we calculating the eigenvector with the the biggest eigenvalue from **C**, which is then the gaze vector $\mathbf{g}^h$ in the head coordinate system. To calculate the eigenvector we using the Von Mises iteration.

For the interpolation, one can use any method from the existing interpolation methods. Often higher order polynomials are used for the interpolation in gaze tracking techniques. But polynomial interpolation have the disadvantage that it causes the appearance of strong oscillation. To overcome this problem, we use a radial basis function (RBF) with a thin plate spline (TPS) kernel for the interpolation. TPS interpolation is commonly used for dense scattered data interpolation which fits perfectly in our case.

In the second step, the translation of the gaze vector $\mathbf{g}^h$ from the head into the world coordinate system is performed with the transformation Matrix $\mathbf{M}^h$ (see Equation 1). $\mathbf{M}^h$ consists of rotation matrix $\mathbf{R}_h$ and translation vector $\mathbf{t}^h$. Here, $\mathbf{R}_h$ corresponds to the orientation and $\mathbf{t}^h$ corresponds to position of the head coordinate system.

$$\mathbf{g} = \mathbf{M}^h \cdot \mathbf{g}^h \quad (1)$$

$$\mathbf{M}^h = \mathbf{t}^h \cdot \mathbf{R}_h \quad (2)$$

So, after this step, we obtain the gaze vector **g**. However, we still do not know at which point of this vector the user is looking at i.e. the gaze point is not known yet. Hence, we approach for the third and final step.

In the third step, we can obtain the gaze point in two possible ways. The first way is to have two gaze vectors and calculate the intersection point between both of these vectors. In this case, if there is a gaze vector for each eye then the intersection point of this two gaze vectors is the gaze point. The second way is to use the known room geometry. Here, if only one gaze vector is available, we need to calculate the intersection point of the gaze vector with the given room geometry. This intersection point is considered as the gaze point. Here, we choose to work with the second option and a large high-resolution display (LHRD) is used for the calculation of the intersection point.

The position and dimension of the LHRD is known in the world coordinate system. The LHRD **L** is defined as a plane:

$$\mathbf{L} = \mathbf{p}^L + s \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + u \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad (3)$$

Where $\mathbf{p}^L$ is the position of the LHRD and **s** and **u** are two scalar parameters. To get a gaze point $\mathbf{g}^p$ from the gaze vector $\mathbf{g}$, we use the following formula in Equation 4:

$$\mathbf{g}^p = \mathbf{a} + \lambda \cdot \mathbf{g} \tag{4}$$

Here, $\mathbf{a}$ is the average value of all calculated gaze points and λ is calculated as in Equation 5. We need λ to obtain $\mathbf{g}^p$ at the LHRD. Because the z-axis of the LHRD is perpendicular to the z-axis of the world coordinate system the calculation of λ is done as follows:

$$\lambda = \frac{\mathbf{p}_z^L - \mathbf{a}_z}{\mathbf{g}_z} \tag{5}$$

Placing this λ in (4), we get $\mathbf{g}^p$ at the LHRD. This known point at the LHRD can, for example, be used for interaction purposes or gaze contingent rendering.

4.4. Eye position

The position of the eye, in relation to the head tracking markers, can be calculated with the help of the calibration data. But in our approach, to calculate the gaze vector, we do not need the information about the eye position.

For each calibration point on the calibration point grid, the corresponding point to the same pupil position is calculated in each calibration point layer. All points which belong to the same pupil position are on one line of sight of the user. Therefore, we get, for each calibration point, a line that go straight into the eye.

Theoretically, all lines should intersect in the middle point of the eye. But, in practice this does not occur due to the anatomy of the eye and the measurement inaccuracies. Between each possible line pair, the point with the closest distance is calculated. The average of these points is then taken as the eye position $\mathbf{e}^h$ in head coordinates. It is worth to mention that, we are able to get the eye position in this way only because all points are in the head coordinate system. Fig. 5a depicts the result of a calibration run along with the calculated eye position.

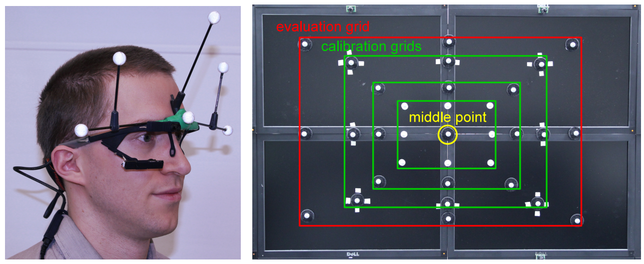
5. Implementation and Results

5.1. Prototype setup

Our prototype system consists of an Optitrack optical tracking system, a Pupil Labs monocular mobile eye tracker and a large high-resolution display (LHRD).

The tracking system is based on passive infrared reflective markers and runs with a frequency of 100Hz. The tracking volume is approximately 3.0 $m \times 3.4$ $m \times 3.0$ m (H×W×D). Our tracking system, utilized to track user's head, uses 12 infrared cameras with infrared LEDs to recognize the passive infrared markers. The tracking system is calibrated once before the evaluation; we calibrate it until it gives a mean error below 1 mm.

The eye tracker works with a frequency of up to 120Hz and has one camera to track a user's eye and another to record the scene. In our approach, we use a resolution of 640×480 from the eyecamera. The information about the eye tracker can be found at [https://pupil-labs.com/pupil/. Accessed on: 04.08.2016]. To track the position



(a) (b)

Figure 3: *(a) Pupil Labs eyetracker with attached tracking system markers and attached eyewear strap. (b) Four 3×3 grids of markers attached at the LHRD. The middle marker is the same for each grid. The three inner grids (green) are for the calibration and the outer grid (red) is for the evaluation.*

and orientation of the eye tracker, some markers from the tracking system are attached, see Fig. 3a.

The LHRD in the lab consist of 24 displays, which are arranged in a 6×4 (row×column) array. Each display has a resolution of 1920×1200. The LHRD is used to show the current gaze point. For the evaluation, 33 tracking system markers are attached to the LHRD, see Fig. 3b.

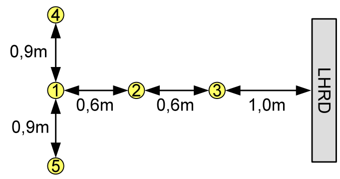


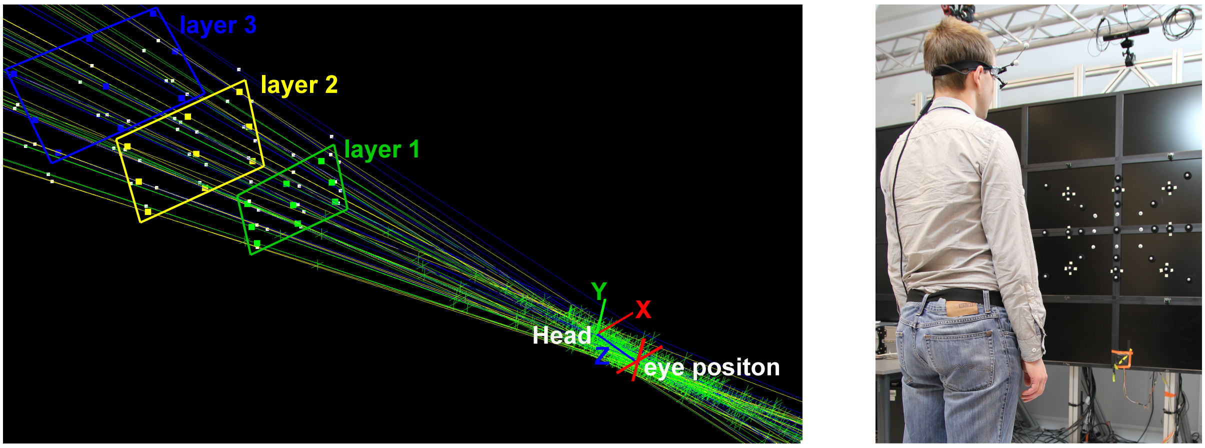
Figure 4: *Calibration positions: 1 to 3; Evaluation positions: 1 to 5.*

The implementation consists of mainly two programs – one for the calibration procedure and the other for showing the current gaze point of a user on the LHRD, which is also used for evaluation. The implementation of these programs is done on Linux with C++ and with the help of the Vrui VR Toolkit (http://idav.ucdavis.edu/ okreylos/ResDev/Vrui/). Vrui is based on OpenGL and used to show the gaze point on the LHRD. We also use VRPN (Virtual-Reality Peripheral Network, http://www.cs.unc.edu/Research/vrpn/) to stream all tracking data to one PC.

5.2. Evaluation

Once the calibration is done, we begin the evaluation of the proposed approach. The evaluation consists of two parts – at first, we evaluate the calculated eye position and then, we evaluate the accuracy of the gaze tracking. The evaluation of the approach is done with five users.

To avoid slippage as much as possible we attach a eyewear strap to the eye tracker and fix the eye tracker with double-faced adhesive tape at the forehead.



(a) (b)

Figure 5: *(a) The result of a calibration run with the calculated eye position. The green, yellow and blue dots are the calibration points of a layer. The white points are the points which are interpolated for the calibration points to get the eye position. For each calibration point two points are calculated, for each other layer one. The lines are the connection from the calibration points with their interpolated points. The small green crosses are the intersection points of the lines. The eye position is the average of the intersection points. (b) System setup with user, eye tracker and LHRD with calibration and evaluation markers.*

Table 1: *The median and the standard deviation (SD) of the Euclidean distance from the ground truth to the calculated eye position are shown in centimeter in the x-, y-, z-direction and in 3D space.*

user	1	2	3	4	5	Median ± SD
x	0.15	3.91	1.35	0.07	1.86	1.46 ± 1.56
y	1.11	2.54	0.28	0.55	0.34	0.96 ± 0.93
z	3.44	6.43	14.93	8.7	3.32	7.36 ± 4.78
3D	3.61	7.94	14.99	8.7	3.82	7.81 ± 4.63

According to the standard procedure specified by the manufacturer of our eye tracker, it is able to deliver the raw pupil position and the gaze position in pixel coordinates in the scene camera image. Our approach is designed to work with either of these information. However, we use only raw pupil position to obtain the gaze vector. In this case, the advantage is that we do not need to calibrate the eye tracker.

Then for each user a calibration is done with our approach. The calibration process, for each user, takes an approximate of two minutes. Once the calibration is done, the calculation of the current gaze vector takes less then 0.35 ms with our approach.

5.2.1. Eye position

For evaluating the calculated eye position from the calibration, the real eye position is measured with a tracking tool wand. A tracked marker at the wand is placed, at first, on the right side of the closed eyelid and then, on the left side. We take the average of these two values as eye position.

We get an accuracy of 1.46 cm in the x, 0.96 cm in the y and of 7.36 cm in z-direction. Their is a relative high scattering in the z-direction, see Fig. 5a. This leads to a standard deviation of 4.78 cm for the z-value. Because of the scattering of the z-value, also the Euclidean distance between the ground truth and the result of our approach for the eye position have a high standard deviation, see Table 1.

5.2.2. Gaze tracking

For evaluating the gaze vector, we asked the user to look naturally at the nine evaluation markers which are placed on the surface of the LHRD, see Fig. 3b. The whole evaluation process is done from five different positions in the room, see Fig. 4. For each marker, we calculate the the angle between the vector from the ground truth eye position, see section 5.2.1, to the calibration point and to the calculated gaze point. Then we calculate the average angle for each user, see Table 2.

We receive an overall accuracy, for unrestrained head movement, of 0.85° with a precision of ± 0.30°.

5.3. Discussion

The calibration process in our approach does not require any parameter to be measured manually, for example we estimate the eye position and the gaze vector in realtime without any manual measurement. In our approach, the eye position is calculated in relation to the head coordinate system and hence, the position and orientation of the eye coordinate system is not required. In our initial experiments for determining the eye positions, we get better and stable values in x and y-directions, but relatively unstable values in

Table 2: *Visual angular error with median and standard deviation (SD), three calibration point layers. Calibrated from the positions 1, 2 and 3 in Fig. 4*

user	1	2	3	4	5	Median ± SD
position 1	0.62°	0.59°	0.66°	0.55°	0.68°	0.62° ± 0.05°
position 2	0.68°	0.88°	0.65°	0.39°	0.43°	0.60° ± 0.20°
position 3	0.76°	1.03°	1.26°	1.12°	0.50°	0.93° ± 0.30°
position 4	0.60°	0.44°	1.19°	0.37°	1.13°	0.74° ± 0.38°
position 5	1.08°	0.74°	2.01°	1.20°	1.69°	1.34° ± 0.50°
average	0.74°	0.73°	1.15°	0.72°	0.88°	0.85° ± 0.30°

Table 3: *Visual angular error with median and standard deviation (SD), two calibration point layers. For this data only two of three calibration layers from the same calibration as in Table 2 are used for the interpolation. As expected the results are more worse when removing one calibration layer.*

user	1	2	3	4	5	Median ± SD
layer 1 & 2	0.64°	0.66°	1.67°	0.75°	1.02°	0.95° ± 0.30°
layer 1 & 3	0.92°	0.96°	1.89°	6.49°	1.29°	2.31° ± 2.42°
layer 2 & 3	0.84°	0.94°	0.97°	0.89°	1.08°	0.95° ± 0.34°

z-direction. We can see, in Fig. 5a, that there is a relative high scattering in z-direction which leads to a standard deviation of 4.78 cm for the z-value.

In our proposed approach, the user is allowed to move freely in the range of the tracking volume both during and after the calibration. The first evaluation test shows that a natural head movement during the calibration gives, in most cases, better results. However, in some cases, we obtained strange results due to strange head movement of the user. Due to this reason, in our evaluation, we asked the user to keep his head in a steady position during the calibration. However, the user's head is not fixated by any external apparatus during the calibration.

Another solution would be to show head fixed calibration points at a display so that the user can move his head freely but is also forced to move his eyes. In this case, it will also be possible to show the calibration points in a way that the same calibration point in each calibration point layer lies at the same pupil position. We believe that this will improve the calibration results.

The limitation of our proposed approach, in case of not achieving more accurate gaze vector, occurs due to the occasional imprecise data delivered by the eye tracker and the head tracking system. Besides, we also experience issues with the slippage of the eye tracker when the tracker is placed on the users. Although we were able to reduce the slippage issue by using a eyewear strap and double-faced adhesive tape, the slippage still occurs to some extent. If we look at the results in Table 2, we can see that the first point in the evaluation is the most stable one and then, the results tend to be slightly worse. We think the main reason for this is the slippage issue that is introduced during the movement of the users. Table 3 shows the result of removing one layer from the calibration data and we can clearly see that the results are getting worse in comparing with using all the three layers in the evaluation. For example, we can see for user 4, on second row, we get an outlier which perhaps occurs due to removing second layer .

Moreover, we use nine calibration points for each distance in our approach which could have been increased to a bigger number for obtaining even more accurate gaze vector. Besides this, the limitation of our approach occurs also due to the usage of interpolation method which have its own limitations like every other interpolation methods.

We achieve an average visual angle error of approximately 0.85° under unrestrained head movement. Although this value does not seem as good as some of the methods described in literature, we use a simpler, automatic and low cost approach whereas, the calibration procedures of the methods in literature are more complex, not completely automatic and they use more expensive hardware. Moreover, if we would have considered only one evaluation position (which normally gives stable and less slippage-affected result), like Cesqui et al. [CvdLLd13], we would also get very good result in terms of accuracy. In fact, by looking at the results in Table 2, we can see that for the first position, we get an average visual angle error of approximately 0.62° with a standard deviation of 0.05°. We think that this, or even slightly better accuracy, is the real limit of our approach if we could remove the slippage.

6. Conclusion

We have introduced a new calibration method which delivers the eye position, in relation to the head coordinate system, and a current gaze vector in real time with support for free movement of the user. We also deliver a relatively low cost unrestrained gaze tracking system with a high accuracy. Although, it is still a bit expensive for the use at home, but it is quite affordable and will open gaze tracking possibilities for research and industry.

Our approach is the first attempt to use interpolation based method for gaze tracking of freely moving users. It is also easily adoptable since it does not require any specific or special version of hardware that we use. It is designed to work with any available eye and head tracker. In our approach, after performing the calibration, the calculation of the gaze vector can be done only with the position information of the pupil and the head, which are obtained from the eye and head trackers respectively. With our proposed approach, we achieve an overall gaze vector accuracy of 0.85°. We plan to do more evaluation for the gaze tracking, especially to improve the interpolation method used in our approach.

7. Acknowledgements

This work was supported by the German Research Foundation (DFG) within the research training group GRK 1424 MuSAMA.

References

[AEC96] ALLISON R. S., EIZENMAN E., CHEUNG B. S.: Combined head and eye tracking system for dynamic testing of the vestibular system. *Biomedical Engineering, IEEE Transactions on 43*, 11 (1996), 1073–1082. 2

[BM04] BROLLY X. L., MULLIGAN J. B.: Implicit calibration of a remote gaze tracker. In *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on* (2004), IEEE, pp. 134–134. 2

[CVC08] CERROLAZA J. J., VILLANUEVA A., CABEZA R.: Taxonomic study of polynomial regressions applied to the calibration of video-oculographic systems. In *Proceedings of the 2008 symposium on Eye tracking research & applications* (2008), ACM, pp. 259–266. 2

[CvdLLd13] CESQUI B., VAN DE LANGENBERG R., LACQUANITI F., D'AVELLA A.: A novel method for measuring gaze orientation in space in unrestrained head conditions. *Journal of vision 13*, 8 (2013), 28. 2, 7

[DMG*01] DUCHOWSKI A. T., MEDLIN E., GRAMOPADHYE A., MELLOY B., NAIR S.: Binocular eye tracking in vr for visual inspection training. In *Proceedings of the ACM symposium on Virtual reality software and technology* (2001), ACM, pp. 1–8. 2

[EDP*12] ESSIG K., DORNBUSCH D., PRINZHORN D., RITTER H., MAYCOCK J., SCHACK T.: Automatic analysis of 3d gaze coordinates on scene objects using data from eye-tracking and motion-capture systems. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (2012), ACM, pp. 37–44. 2

[EPR06] ESSIG K., POMPLUN M., RITTER H.: A neural network for 3d gaze recording with binocular eye trackers. *The International Journal of Parallel, Emergent and Distributed Systems 21*, 2 (2006), 79–95. 2

[GE06] GUESTRIN E. D., EIZENMAN M.: General theory of remote gaze estimation using the pupil center and corneal reflections. *Biomedical Engineering, IEEE Transactions on 53*, 6 (2006), 1124–1133. 2

[HAJ04] HUANG H., ALLISON R. S., JENKIN M.: Combined head-eye tracking for immersive virtual reality. In *ICAT'2004 14th International Conference on Artificial Reality and Telexistance Seoul, Korea* (2004), Citeseer. 2

[HCT*08] HERHOLZ S., CHUANG L. L., TANNER T. G., BÜLTHOFF H. H., FLEMING R. W.: Libgaze: Real-time gaze-tracking of freely moving observers for wall-sized displays vision, modeling, and visualization. *Vision, Modeling, and Visualization 2008: Proceedings, October 8-10, 2008, Konstanz, Germany* (2008), 101. 2

[HHN*02] HANSEN D. W., HANSEN J. P., NIELSEN M., JOHANSEN A. S., STEGMANN M. B.: Eye typing using markov and active appearance models. In *Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on* (2002), IEEE, pp. 132–136. 2

[HJ10] HANSEN D. W., JI Q.: In the eye of the beholder: A survey of models for eyes and gaze. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 32*, 3 (2010), 478–500. 2

[HP05] HANSEN D. W., PECE A. E.: Eye tracking in the wild. *Computer Vision and Image Understanding 98*, 1 (2005), 155–181. 2

[JY02] JI Q., YANG X.: Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging 8*, 5 (2002), 357–377. 2

[KBB*08] KOHLBECHER S., BARDINST S., BARTL K., SCHNEIDER E., POITSCHKE T., ABLASSMEIER M.: Calibration-free eye tracking by reconstruction of the pupil ellipse in 3d space. In *Proceedings of the 2008 symposium on Eye tracking research & applications* (2008), ACM, pp. 135–138. 2

[KR04] KIAT L. C., RANGANATH S.: One-time calibration eye gaze detection system. In *Image Processing, 2004. ICIP'04. 2004 International Conference on* (2004), vol. 2, IEEE, pp. 873–876. 2

[Mod11] MODEL D.: *A Calibration Free Estimation of the Point of Gaze and Objective Measurement of Ocular Alignment in Adults and Infants*. PhD thesis, University of Toronto, 2011. 2

[NSI*10] NAGAMATSU T., SUGANO R., IWAMOTO Y., KAMAHARA J., TANAKA N.: User-calibration-free gaze tracking with estimation of the horizontal angles between the visual and the optical axes of both eyes. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (2010), ACM, pp. 251–254. 2

[NSNG12] NOWAKOWSKI M., SHEEHAN M., NEAL D., GONCHAROV A. V.: Investigation of the isoplanatic patch and wavefront aberration along the pupillary axis compared to the line of sight in the eye. *Biomedical optics express 3*, 2 (2012), 240–258. 2, 3

[PCG*03] PERÉZ A., CORDOBA M., GARCIA A., MÉNDEZ R., MUNOZ M., PEDRAZA J. L., SANCHEZ F.: A precise eye-gaze detection and tracking system. 2

[PP33] PARK R. S., PARK G. E.: The center of ocular rotation in the horizontal plane. *American Journal of Physiology–Legacy Content 104*, 3 (1933), 545–552. 2

[PPR11] PIRRI F., PIZZOLI M., RUDI A.: A general method for the point of regard estimation in 3d space. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on* (2011), IEEE, pp. 921–928. 2

[RWL07] RONSSE R., WHITE O., LEFEVRE P.: Computation of gaze orientation under unrestrained head movements. *Journal of neuroscience methods 159*, 1 (2007), 158–169. 2

[TKA02] TAN K.-H., KRIEGMAN D. J., AHUJA N.: Appearance-based eye gaze estimation. In *Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on* (2002), IEEE, pp. 191–195. 2

[VCP06] VILLANUEVA A., CABEZA R., PORTA S.: Eye tracking: Pupil orientation geometrical modeling. *Image and Vision Computing 24*, 7 (2006), 663–679. 2

[VSG12] VALENTI R., SEBE N., GEVERS T.: Combining head pose and eye location information for gaze estimation. *Image Processing, IEEE Transactions on 21*, 2 (2012), 802–815. 2

[WBC06] WILLIAMS O., BLAKE A., CIPOLLA R.: Sparse and semi-supervised visual mapping with the sˆ 3gp. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on* (2006), vol. 1, IEEE, pp. 230–237. 2

[WSV05] WANG J.-G., SUNG E., VENKATESWARLU R.: Estimating the eye gaze from one eye. *Computer Vision and Image Understanding 98*, 1 (2005), 83–103. 2

[ZJ04] ZHU Z., JI Q.: Eye and gaze tracking for interactive graphic display. *Machine Vision and Applications 15*, 3 (2004), 139–148. 2

[ZJ07] ZHU Z., JI Q.: Novel eye gaze tracking techniques under natural head movement. *Biomedical Engineering, IEEE Transactions on 54*, 12 (2007), 2246–2260. 2