Eye Gaze Estimation from a Single Image of One Eye

Jian-Gang Wang
Institute for Infocomm Research
21 Heng Mui Keng Terrace
Singapore 119613
jgwang@i2r.a-star.edu.sg

Eric Sung
School of EEE
Nanyang Technological University
Singapore 639798
eericsung@ntu.edu.sg

Ronda Venkateswarlu
Institute for Infocomm Research
21 Heng Mui Keng Terrace
Singapore 119613
vronda@i2r.a-star.edu.sg

Abstract

In this paper, we present a novel approach, called the "one-circle" algorithm, for measuring the eye gaze using a monocular image that zooms in on only one eye of a person. Observing that the iris contour is a circle, we estimate the normal direction of this iris circle, considered as the eye gaze, from its elliptical image. From basic projective geometry, an ellipse can be back-projected into space onto two circles of different orientations. However, by using an anthropometric property of the eyeball, the correct solution can be disambiguated. This allows us to obtain a higher resolution image of the iris with a zoom-in camera and thereby achieving higher accuracies in the estimation. The robustness of our gaze determination approach was verified statistically by the extensive experiments on synthetic and real image data. The two key contributions in this paper are that we show the possibility of finding the unique eye gaze direction from a single image of one eye and that one can obtain better accuracy as a consequence of this.

1. Introduction

There are two components to the human visual line-of-sight: pose of human head and the orientation of the eye within their sockets. We have investigated these two aspects but will concentrate on the eye gaze estimation in this paper. We found that the domain knowledge of the human face is important and essential for determining the head pose and eye gaze utilizing only minimal robust features and under real-time requirement. In our work, the domain knowledge used is not merely from facial features but from anatomical properties. For instance, we found that the eye gaze can be estimated using the normal to the iris contour, which has an approximate fixed angle with the true gaze. Hence, we have developed a novel approach, called the "one-circle" algorithm, for measuring eye gaze using monocular image that zooms in on only one eye of a person. In addition, we make an observation that the eye-lines (connecting the two far eye corners and the two neighboring eye corners respectively) are parallel to the mouth-line (connecting the two mouth corners). This domain knowledge led us to develop a new method for determining head pose fast and robustly using the vanishing point in the image formed by the intersection of the eye-lines and mouth-line. The details of this pose determination paradigm and two alternative proofs of this approach can be found in [19]. In this paper, we will "focus" on the eye gaze determination.

A good survey of earlier eye-gaze (eye-movement) tracking techniques, most of them being active methods, is provided in [21]. The techniques include Electro-Oculography [8], Limbus, Pupil and Eyelid tracking [11, 7, 17], Contact Lens Method, Corneal and Pupil Reflection Relationship [7], Purkinje Image Tracking [2], Artificial Neural Networks [17], Morphable Models [14] and geometry [6, 10, 12].

In most existing approaches [20, 3, 10, 9], the iris contours on the image plane are simplified to be circles, so the felicitous circular geometry is utilized and iris outer boundaries (limbus) are detected using a circle edge operator. For instance, the center of the iris is detected using the circular Hough Transform in [10]. In [9], the iris is located by matching the left and right curvatures of the iris (circle) candidate with those of the iris to be detected in the edge image.

Zelinsky et al [10] presented an eye gaze estimation in which the eye corners are located using a stereo vision system. Then the eyeball position can be calculated from the pose of the head and a 3D "offset" vector from the mid-point of the corners of an eye to the center of the eye. Consequently the radius of the eyeball can be obtained. However the "offset vector" and the radius of the iris have to be manually adjusted through a training sequence where the gaze point of the person is known.

It is difficult to determine the eye gaze by analyzing the eyeball rotations from a typical image with low resolution for the eye region [18, 6]. The iris is partially occluded by the upper and lower eyelids so it will be difficult to fit its contour consistently and reliably. For example, in [9] the field of the view of the camera is set to capture the whole face in the image, the width of an eye is only about 30 (pixels) and the radius of the iris in the image plane is only 5 (pixels) in a typical situation. Therefore, it is hard

to determine the gaze in a 3D scene accurately with such low resolution.

In this paper, we make use of the observation that the iris contour (not the iris) while being a circle in 3D is perspectively an ellipse in the image. The gaze, defined as the normal to the plane of the iris circle, can be estimated from the ellipse/circle correspondence. However, it will result in two possible solutions of the normal. We propose a "one-circle" algorithm to disambiguate the solutions. The unique supporting plane was obtained based on a geometric constraint, namely that the distances between the eyeball′s center and the two eye corners should be equal to each other. We will refer to this as the "distance constraint".

Domain knowledge played a key role in the success of the one-circle approach. Although the eyeball center cannot be seen, its location can be inferred from the pose of the head. This is because its average 3D location relative to the observed features is very close to a generic constant and can be fixed during model acquisition [12]. In this paper, the ratio of the radius of a person's iris and the radius of his/her eyeball in 3D space is found to possess very low ensemble variance and consequently we can fix the ratio as the generic average. The small variation from person to person of this ratio thus had no significant effect on the results. Hence, the eyeball center can be located once the radius of the iris has been calibrated.

In summary, our method differs from others in the following respects. We treat the image of the iris contour not as a circle but correctly as an ellipse in the image. Hence, our approach is more realistic than the existing approaches. The other difference is that our method is more accurate since our method only needs an image of one eye we can zoom in on it and thereby achieving higher resolution. However, in doing so, we need to address an issue that emerges. Zooming and tracking a single eye poses a problem. The problem of having possible out-of-field views can be settled by guiding the "gaze" camera to locate the eye by the head pose estimation results. Considering that, we have developed a general approach that combines head pose determination with eye gaze estimation. The pose of the human head, including the 3D location of the eye corners, mouth corners and the orientation of the face, can be obtained from the head pose estimation subsystem. The details of this integrated approach can be found in [19].

The approach to estimate the eye gaze is discussed in section 2 and the iris detection in section 3. Experimental results are given on simulated data as well as on real images in section 4. The conclusion is discussed in Section 5.

2. Gaze positioning

Thus our gaze determination methodology relies on estimating the ellipse of the iris contour and by projective geometry, this leads to two normal solutions. In order to discard the false normal, we made use of the anthropomorphic ratio of the eyeball radius to the iris circle radius.

2.1 Eye model

A simple eye model is defined (see Fig. 1). The eyeball is assumed to be a sphere with radius R. Actually, it is not quite a sphere but this discrepancy does not affect our methodology. The iris is located at the front of eyeball and its contour is a circular ring of radius r (see Fig. 1(a)). The distance from the center of the eyeball to the iris plane is d. The relation between R, r and d (see Fig. 1(c)) is:
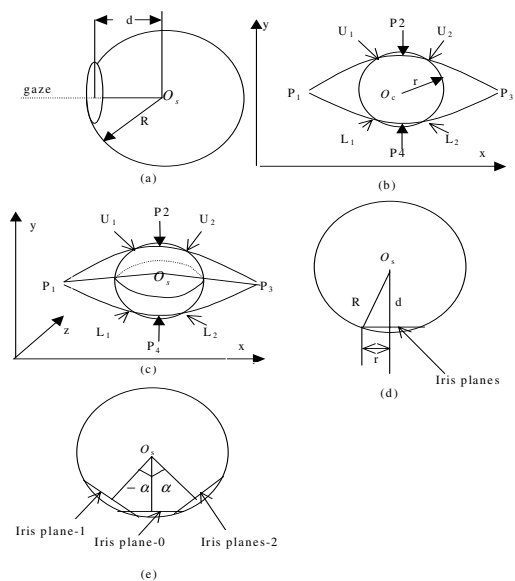
$$R^2 = r^2 + d^2 \qquad (1)$$

Figure 1. **Eye model.**

The optical axis of the eye is the line passing through the center of the eyeball and the center of the iris, and we take it to be the *eye gaze*. When changing the eye gaze, the eyeball rotates around its center (see Fig 1(e)). Our defined eye gaze keeps a nearly fixed angle (k, the angle between the visual and the anatomical axis of the eye) with the central gaze vector that is determined by the eye lens. By a simple calibration, the central gaze vector can be obtained from our estimated eye gaze.

The radius of the iris is very close to an anatomical constant (around 7 mm) [12], and the radius of the eyeball ranges from 12 mm to 13 mm according to the

anthropometric data [9]. Hence, the ratio of the radius of a person's iris and the radius of his/her eyeball in 3D space is stable with a small standard derivation. In our eye model, the average ratio of the radius of a person's iris and the radius of his/her eyeball in 3D space, i.e. R/r, is assumed to be a generic constant. This assumption is justified because of the very small deviations over the ensemble. Hence, we use the anatomical average assumed in [12] in our algorithm. Once r has been found, both R and d can be obtained consequently. The computation cost involved in our calibration is lighter than the method presented in [10]. In [10], both r and R are determined manually from a training sequence. We will discuss the sensitivity of our gaze determination algorithm to the ratio R/r in section 2.4.

In our method, we need to estimate the iris ellipse in the un-occluded parts of the iris. Thus, we have to detect the upper and lower eyelids. The upper and lower eyelids are modeled as parabolas (in Fig. 1(b)). Let the upper eyelid pass through points $P_1(x_1, y_1)$, $P_2(x_2, y_2)$ and $P_3(x_3, y_3)$ and the lower eyelid pass through points $P_1(x_1, y_1)$, $P_4(x_4, y_4)$ and $P_3(x_3, y_3)$. The equation of an eyelid is of the form:

$$y = a(x\text{-}b)^2 + c \tag{2}$$

which for the upper eyelid yields:

$$a = \frac{-y_2}{(x_1 - x_2)^2}, \ b = x_2, c = y_2 \tag{3}$$

while for the lower eyelid yields:

$$a = \frac{-y_4}{(x_1 - x_4)^2}, \ b = x_4, c = y_4 \tag{4}$$

In our simulations, the iris contour edges that lie between the upper and lower eyelids are consistently located and used to fit the elliptical contour. For instance, curves U_1L_1 and U_2L_2 shown in Fig. 1(b) are the fitting edges that we want. These fitting edges can be located by using the above equations of the iris and eyelids.

2.2 The "one-circle" algorithm

From the observed perspective projection of a circle having known radius, it is possible to infer analytically the supporting plane on which the circle lies, as well as where the center of the circle lies. The problem has been extensively investigated and there are many papers concentrating on 3D location of circular objects [5, 15, 19, 16]. We adopted the monocular camera-positioning algorithm proposed in [16, 19]. It was noted that two solutions of the 3D position of the iris plane will be obtained from circle/ellipse correspondent. Our paper shows how we disambiguate the solution using the "one-circle" algorithm to be discussed below.

Define the origin of the camera coordinate system to be at the lens center and let the Z-axis coincide with the optical axis of the camera. Then for a right-hand system, take Y-axis to be vertical and the X-axis to be horizontal.

Now, as noted before, the distance between the two corners of an eye and the center of the eyeball should be equal to each other (see Fig.1(c)). Hence:

$$O_s P_1 = O_s P_3 \tag{5}$$

Consider the iris contour $\mathbf{Q}$. Call the two solutions of the normal of the plane of $\mathbf{Q}$ as $\mathbf{n}_1 = (\cos\alpha_1, \cos\beta_1, \cos\gamma_1)^T$, $\mathbf{n}_2 = (\cos\alpha_2, \cos\beta_2, \cos\gamma_2)^T$ and the corresponding solutions of the center of the iris contour $O_{c1}(x_{01}, y_{01}, z_{01})$ and $O_{c2}(x_{02}, y_{02}, z_{02})$ respectively. Using the eye model defined in Section 2.1, the center of the eyeball O_{si} can be calculated by:

$$x_{si} = x_{0i} + d\cos\alpha_i, \ y_{si} = y_{0i} + d\cos\beta_i, \ z_{si} = z_{0i} + d\cos\gamma \tag{6}$$

where i =1, 2, d is the distance from center of the eyeball to the iris plane (see Fig. 1 (a) or Fig. 1(c)):

$$d = \sqrt{R^2 - r^2} = r\sqrt{c^2 - 1} \tag{7}$$

where c is a constant.

After that, the solutions of the two eye corners are projected to the gaze camera coordinate system. The distances between the center of the eyeball and the two eye corners are compared. Due to image noise, the unique solution of the iris plane should be the one that satisfies:

$$O_s P_1 \approx O_s P_3 \tag{8}$$

In our algorithm, we calculate $O_{s1}P_1$, $O_{s1}P_3$, $O_{s2}P_1$ and $O_{s2}P_3$. If

$$|O_{s1}P_1\text{-}O_{s1}P_3| \leq |O_{s2}P_1\text{-}O_{s2}P_3| \tag{9}$$

then $(\mathbf{n}_1, O_{c1})$ is the solution what we want, else $(\mathbf{n}_2, O_{c2})$ is the solution.

2.3 Degenerate cases of "one-circle" algorithm

The "one-circle" algorithm degenerates when the following condition is met: the iris contour is symmetrical about the Y-Z plane of the camera, where Y- and Z- axis is the vertical and the optical axis of the camera, respectively.

Actually, this condition corresponds to the case that user is facing front whereby the iris is symmetrical about the optical axis. Thus, it is impossible to distinguish, from the ellipse, whether the person is looking upwards or downwards. Fortunately, this degenerate case can be prevented by comparing the Y-position (vertical) of the iris centers with the one of the eye corners. If the Y coordinate of the iris center is greater, then the person is looking upwards, else the person is looking downwards.

2.4 Sensitivity of the "distance constraint" to the ratio

In our method, the average ratio of the radius of a person's iris and the radius of his/her eyeball in 3D space is assumed to be a generic constant, refer to the eye model presented in section 2.1. The observed iris radius is

calibrated and the radius of the eyeball is calculated using the ratio. The center of the eyeball is located along the line-of-sight. The angle between the two normal solutions is large if the head pose is not close to the frontal view (degenerate case). Consequently, the separation of the two resulting eyeball centers is large enough to disambiguate the solutions based on the "distance constraint". In our experiments, we found that the algorithm based on the "distance constraint" is robust to the ratio. The same unique solution can be obtained for different gazes even when the ratio is varied by ±50%.

3. Iris detection

The most prominent and reliable features within the eye region are edges of the iris. Obviously for our algorithm, the existing iris detection methods using circular edge operators cannot be used. Instead, we detect the iris edge (bright-to-dark and dark-to-bright step edge) using a (3×3) vertical edge operator, which detects and emphasizes vertical edges, and a (3×3) morphological "open" operation. Since the 3D position of the corners of the eyes are already known in our pose determination [19], the location of the eye corners in the gaze image are subsequently known. Hence, the iris detection search can be executed on a small region between them. Because of the high contrast between the eyeball and the eye white, the eye image is easily segmented based on a threshold that was automatically selected from the histogram [13] of the eye region. It is known that "opening" an image breaks narrow isthmuses. Hence, the morphological "open" operation is applied to separate the iris from the eyelid. In our experiments, we found that just one "open" operation is sufficient. After that, the Canny edge operator is used to detect the edges of the iris and the edges with directions $90^0 \pm 5^0$ were retained. All of the vertical edge segments are then tracked respectively using an "edge following" technique. The lengths of the edges are obtained. The two sides of the iris are the ones with the two longest vertical edges.

Once the iris edges are obtained reliably, the iris contour is fitted to an ellipse. The accuracy of the ellipse fitting affects the gaze computing, hence a robust ellipse fitting method is needed. There has been continued interest in the fitting of ellipse to image data because the ellipse, being the perspective projection of the circle, is of great importance in pattern recognition and computer vision. We adopt the method [4], which is specific to ellipse and direct at the same time.

An example of the iris detection process is shown in Fig. 2. An original eye image is shown in Fig. 2(a). In the segmented eye image, some of the eyelids are connected with the iris, (see Fig. 2(b)). Applying an "open" operation cuts the connection (narrow isthmuses), see Fig. 2(c). A vertical edge operator is applied to the Fig. 2(c),

resulting in Fig. 2(d). the two longest edges are deemed the iris edges to be used for fitting the iris contour, (see Fig. 2(e) or 2(f)). Actually, the spurious edges, e.g. the edges in the horizontal direction in Fig. 2(d), are not connected each other although it looks like being so. Hence the two sides of the iris, Fig. 2(e), can be found from Fig. 2(d). The fitted ellipse can be seen in Fig. 2(g) or 2(h).
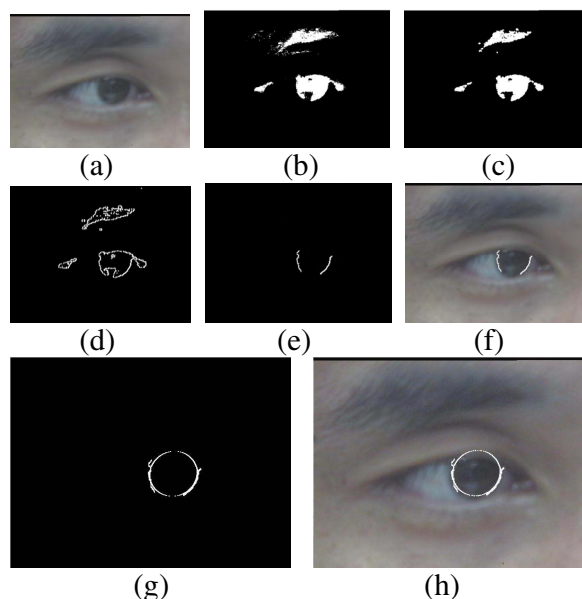

(a)	(b)	(c)
(d)	(e)	(f)
(g)	(h)	

Figure 2. **Iris detection.**

4. Experimental investigation and results

The experiments on synthetic and real data have been done. The good results obtained verified the accuracy and robustness of the method. These experiments using synthetic data and real images are now presented below.

4.1 Experimental results on synthetic data

We test the algorithm on two sets of synthetic data: exact and noisy. By observing the pre-defined targets, the accuracy of the algorithm can be measured.

4.1.1. Simulations on exact data

We consider the possible gaze using the synthetic images of a model eye in which the iris is rotated in every possible direction. Let the head face directly at the camera. Then, without moving the head, the "eye" is made to gaze in various directions. For each gaze, the eyeball is made to rotate accordingly about its center (x_s, y_s, z_s). Consequently, the iris contour will be rotated to the position which corresponds to the expected gaze,

following the eyeball movement. We project the two eyelids and the rotated iris onto the image plane according the preset camera parameters, and then estimate the gaze from the projections. Only the edges between the two eyelids are considered for computing the eye gaze.

In our simulations, the initial contour of the iris (circle) lies parallel to the image plane at a distance of 60-100cm. The radius of the iris is set as 0.65 cm. The ratio of R/r is fixed at 2. The size of the image is set to be 640×480. The intrinsic parameters of the camera are set as:

$$u_0 = 320, \ v_0 = 240, \ f_x = f_y = 5500$$

where (u_0, v_0) are the coordinates of the principle point, f_x and f_y are the scale factors of the camera along the X- and Y-axis respectively. The settings of f_x and f_y here imply that the camera (zoom-in) requires a larger focal length in order for the eyes to appear big enough on the image. The eyeball are rotated about the vertical axis from -50^0 to 50^0 in steps of 1^0 (azimuth) and rotated about the horizontal axis from -10^0 to 10^0 in steps of step 1^0 (elevation) to form a set of synthetic images. The performances of the "one-circle" algorithm are tested on these synthetic images.
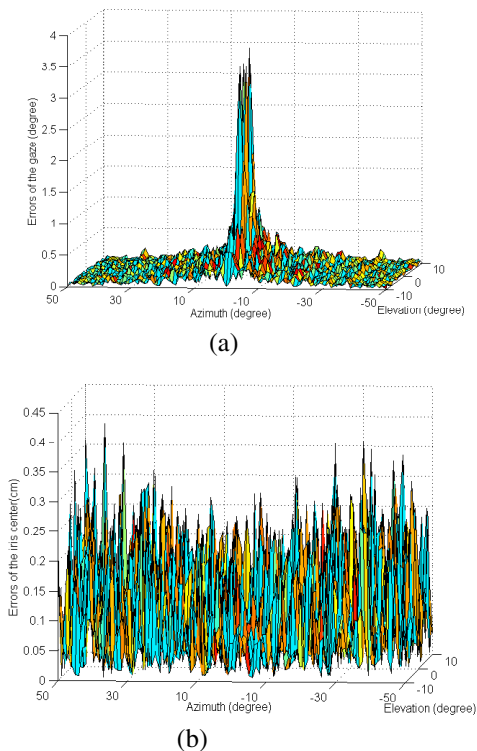

(a)

(b)

Figure 3. **The errors for different poses for the case of noiseless synthetic data. (a) the normal of the iris (b) the center of the iris.**

The errors of the eye gaze over the synthetic test images are illustrate in Fig. 3(a). We can see that the accuracy tends to fall greatly when face is around the frontal view, and falls lesser as the gaze turns away from the fronto-

parallel position. This is because the iris contour becomes nearly a circle instead of an ellipse when the camera direction is approximately fronto-parallel. Fortunately, this degenerate case can be prevented in our application by simply putting the camera slightly skewed to the face. Hence, we will consider the errors of the eye gaze in the following excluding this degenerate case. The errors of the iris center are shown in Fig. 3(b).

The experimental results (for noiseless data) show that using the "one-circle" method, the maximum error of the gaze due to the eyelids' occlusion is 0.3^0, while the maximum error of the center of the iris is 0.1 cm, this can be seen in Fig. 3(a) and Fig. 3(b) respectively.

4.1.2 Robustness to geometrical disturbances

We tested the performance of the algorithm when the data are subjected to geometric disturbances. Corrupting the locations of the imaged features by Gaussian noise can do this. This simulates the condition of imprecise location of the features caused by the feature extraction algorithm, noise and other processes.

The mean error from 100 trials for a gaze is computed to indicate the robustness of the gaze under geometric disturbances. Different levels of noise with standard deviations of 1, 1.4, 2.0 and 2.8 pixels are tried. The results showed that the algorithm is robust. The algorithm can work even the noise is increased significantly, with the resulting accuracy degrading gracefully. When we disturb the iris edges with a standard Gaussian noise (zero mean and standard deviation one pixels), the errors of the estimated gaze over the testing range are less than 1^0, and the errors of the estimated iris center are less than 0.3 cm.

4.1.3 Accuracy of the eye gaze versus point-of-regard

The "one-circle" method proposed here would be able to determine the eye gaze, but it cannot detect point-of-regard from one image. To do the latter, two views would be needed. However, the point-of-regard can be determined by one view with the "one-circle" method if we assume that the person is looking at a point on a fixed plane.

In the simulation, we assume a plane-of-regard positioned in front of the subject. The position of the plane-of-regard with respect to the gaze camera is known. This assumption simplifies the computing because the point-of-regard can be located by intersecting the line-of-sight with the known plane-of-regard which in some HCI applications could be the monitor screen. The simulation setup is shown in Fig. 4. The settings of the camera and the subject are the same as those in Section 4.1.1. The focus plane is defined in the gaze camera coordinate system, and set to be parallel to the image plane and with a distance 1.2 m to the camera. The distance from the subject to the plane-of-regard is thus 1.8 m.
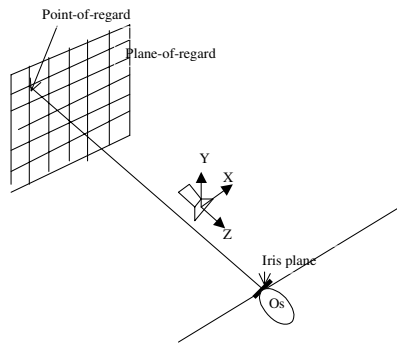
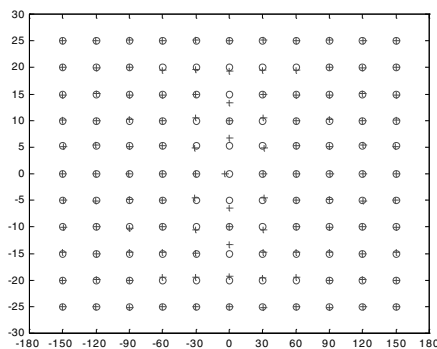Figure 4. **Experiments on the accuracy of the eye gaze versus the focus points.**
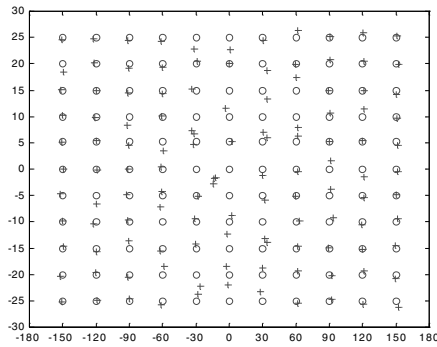

Figure 5. **The errors of the point-of-regard.**

Figure 6. **The errors of the point-of-regard for the synthetic images when the image is perturbed by a Gaussian noise of the 1 pixel standard deviation.**

For a point-of-regard P on the plane-of-regard, the position of the iris plane is changed in order to focus on the point-of-regard (see Fig. 4): the iris plane is rotated around the center of the eyeball O_s until the line O_sP becomes the normal of the iris plane. From that, the eye gaze and consequently the point-of-regard are estimated from the synthetic image.

The results are shown in Fig. 5, where the true and estimated point-of-regard are marked as "O" and "+" respectively. The testing range on the plane-of-regard is set to be: -150 cm to 150 cm in 30 cm steps for horizontal displacement; -25 cm to 25 cm in 5 cm steps for vertical displacement. The results on the noisy data are shown in Fig. 6. Similarly to the tests done in Section 4.1.2, noise levels of standard derivations of 1, 1.4, 2.0 and 2.8 pixels are tried. The algorithm can work even when the noise level is increased significantly; the results degrades gracefully.

The accuracy of the point-of-regard is found to be satisfactory. Corrupting the image with Gaussian noise (zero mean, standard deviation one pixel), the error is less than 1cm within the range of 1.8 m (Fig. 6).

The simulations have verified the accuracy and robustness of the algorithm with quantitative performance measures. The next step is naturally to test the performance of the algorithm on the real images.

4.2 Experiments on real images

We demonstrate our gaze estimation approach on real images of ten subjects. As we shall see in following sections, the experimental results on the real images results are satisfactory. The errors of the point-of-regard are less than 1.5cm within a 1.5m range and consequently the errors of the gaze are less than 1^0. Higher accuracy of the eye gaze determination has been attained because of the higher image resolution made possible and that we used reliable facial features such as eye corners and iris contour. The result is found to better than the existing non-intrusive approaches, such as Zelinsky [12]. We will deal with some examples to discuss the problems encountered that include iris detection, accuracy of the gaze and integration as follows.

In the following examples, the radius of the iris contour is 0.63 cm; the ratio between the radius of the eyeball and the radius of the iris is set to be 2.

Using the same method presented in Section 4.1.3, the accuracy of the eye gaze respect to the point-of-regard is evaluated on real images.

The experimental setup is shown in Fig. 7. The person gazes in four pre-specified targets including two top corners of a whiteboard, V_1 and V_4, and two points, V_2 and V_3, that divides V_1V_4 into three equal segments. The subject maintains his/her head stationary relative to the gaze camera while he/she gazes at each of the four targets in turn. Hence, the coordinates of his/her eye corners remain almost fixed for the four line-of-sight. In Fig. 7 (b), $\mathbf{n}_1$, $\mathbf{n}_2$, $\mathbf{n}_3$ and $\mathbf{n}_4$ correspond to the true eye gaze at points V_1, V_2, V_3 and V_4 respectively.

A camera is placed between the subject and the whiteboard. The 3D coordinates of the whiteboard with respect to the gaze camera are known. Consequently the

coordinates of V_1-V_4 in the gaze camera system are known. We take the coordinates of the four targets as the reference coordinates for the true point-of-regard. The errors of the algorithm are computed by comparing the estimated point-of-regard and the reference coordinates. The results of the gaze estimation are shown in Fig. 8.
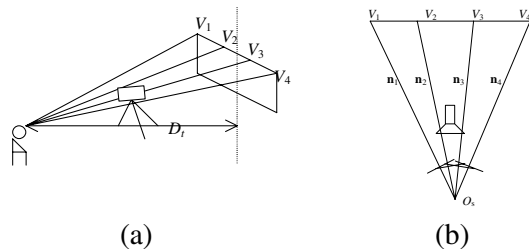

Figure 7. **Testing the accuracy of the eye gaze versus point-of-regard (a) side view (b) top view.**

The image coordinates of the eye corners, which are obtained by projecting the head pose results [19] to the gaze camera, are listed in Table 1. The eye gaze (**n**), the center of the iris (O_c), and consequently the center of the eyeball (O_s) are listed in Table 2. D_L and D_R are the distance from the O_s to the two eye corners (R and L in Table 1) respectively.

We can see the errors of point-of-regard are less than 1.5 cm within 1.5 m (D_t) range, and consequently the errors of the gaze are all less than 1^0.

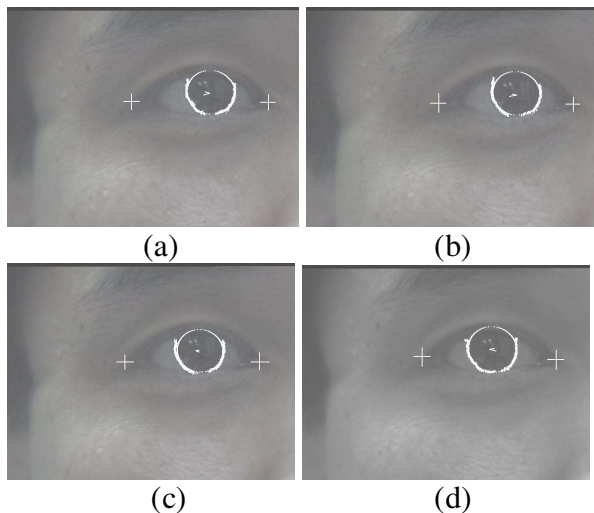$$D = |D_L - D_R| \qquad (10)$$

(a)   (b)

(c)   (d)

Figure 8. **Computed eye gaze for the four target points (a) V_1 (b) V_2 (c) V_3 (d) V_4.**

Applying the "distance constraint" to disambiguate the solutions of the normal and the center, the results are listed in Table 3, where the unique solution is marked as

"T". The errors, E_p, of the point-of-regard are then estimated and listed in Table 2. The errors of the eye gaze are also estimated by

$$E_g = \tan^{-1}(E_p/D_t) \qquad (11)$$

where D_t is the distance from the subject to the plane-of-regard, see Fig. 7 (a).

Table 1. **Eye corners with respect to the gaze camera.**

	Eye corners (cm)
$V_1, V_2,$ V_3, V_4	R: (-4.010, 0.030, 59.100), L: (-2.501, 0028, 59.301)

Table 2. **Determining the gaze by applying "distant constraint".**

	True (cm)	Estimated (cm)	E_p (cm)	E_g
V_1	(60, 15, -90)	(61.056, 14.992, -90)	1.056	0.40^0
V_2	(20, 15, -90)	(19.277, 16.290, -90)	1.479	0.56^0
V_3	(-20, 15, -90)	(19.249, 16.294, -90)	1.496	0.57^0
V_4	(-60, 15, -90)	(-61.020, 14.991, -90)	1.025	0.39^0

Table 3 **Errors of the point-of-regard and the eye gaze.**

	n	O_c (cm)	O_s (cm)	D_R (cm)	D_L (cm)	D (cm)	
V_1	(-0.392, -0.092, 0.915)	(-2.854, 0.097, 58.814)	(-3.297, -0.007, 59.844)	0.964	1.024	0.059	T
	(0.260, 0.096, 0.961)	(-2.858, -0.095, 58.813)	(-2.565, 0.204, 59.895)	0.632	1.653	1.020	F
V_2	(-0.148, -0.107, 0.983)	(-3.123, 0.101, 58.662)	(-3.290, -0.020, 59.769)	0.919	0.976	0.057	T
	(0.001, 0.119, 0.993)	(-3.124, -0.099, 58.062)	(-3.123, 0.233, 59.780)	0.820	1.134	0.314	F
V_3	(-0.260, 0.103, 0.960)	(-3.406, 0.010, 58.477)	(-3.699, 0.215, 59.558)	1.246	0.588	0.657	F
	(0.105, -0.108, 0.989)	(-3.409, -0.101, 58.477)	(-3.290, -0.021, 59.590)	0.842	0.862	0.020	T
V_4	(-0.510, 0.094, 0.855)	(-3.666, 0.094, 58.567)	(-4.240, 0.200, 59.529)	1.767	0.531	1.236	F
	(-0.359, 0.093, 0.929)	(-3.671, 0.095, 58.566)	(-3.268, -0.010, 59.612)	0.829	0.893	0.064	T

In Zelinsky's approach [12], the eye gaze is determined using stereo vision. A total of four eyes in the stereo image pair were used to compute the eye gaze. However, the resolution of the images is low, since the width of an eye is only 30 pixels. Each measurement is not sufficiently accurate to determine the gaze point. Hence, four gaze vectors were averaged to generate a single gaze. Accuracies of $\pm 3.5^0$ were reported. Nonetheless, we acknowledge that Zelinsky´s work is a fully operational system, a rare achievement in computer vision.

5. Conclusion

In this paper, we presented a non-intrusive method of robustly estimating the eye gaze by zoom-in iris imaging. The motivation of our approach is to estimate the eye gaze robustly in real-time and with satisfactory accuracy from a single image frame. The use of the domain knowledge of the human face is crucial and this makes our paradigm original and novel. Actually, an ellipse can be back-projected into the space onto two possible circles. The principle is applied to the eye-gaze by observing that the (3D) contour of the iris is circular and hence it is the circle that we are looking for. By using the "distance constraint" described, we were able to discard the ambiguous solution and retain the correct gaze. This constraint is based on anthropometric properties of the eye, a novel idea. We improve on current eye gaze determination methods by achieving higher resolutions. This comes about mainly because our algorithm is designed to use only a single image of the eye which subsequently means that we can zoom in on it.

In conclusion, eye gaze estimation is important in applications such as virtual reality, video conferencing and human-machine interface/controls. The eye gaze method above is integrated with a head pose estimation module and together will offer great potential especially in these mentioned applications. Of importance to note is that our method is non-invasive, fast and robust. Robust because the segmentation of the iris contour is one of the simplest and most robust facial feature to extract.

References

[1] S. Barattelli, L. Sichelschmidt and G. Rickheit, Eye-movements as an input in human computer interaction: exploiting natural behavior, *Annual Conference of IEEE Industrial Electronic Society*, Vol. 4, pp. 2000-2005, 31 Aug – 4 Sept. 1998.

[2] T. N. Cornsweet, H. D. Crane, Accurate two-dimensional eye tracker using first and fourth Purkinje images, *Journal of the Optical Society of America*, Vol. 63, No. 8, pp.921-928, Aug. 1973.

[3] J. G. Daugman, High confidence visual recognition of persons by a test of statistical independence, *IEEE Transactions on PAMI*, Vol. 15, No. 11, pp.1148-1161. November 1993.

[4] A. Fitzgibbon, M. Pilu and R. B. Fisher, Direct least square fitting ellipse, *IEEE Transactions on PAMI*, Vol. 21, No.5, pp. 476-480, 1999.

[5] D. Forsyth, J. L. Mundy, A. Zisserman, C. Coelho, A. Heller and C. Rothwell, Invariant descriptors for 3-D object recognition and pose, *IEEE Transactions on PAMI*, Vol. 13, No. 10, pp. 971-991, 1991.

[6] A. Gee and R. Cipolla, Determining the gaze of faces in images, *Image and Vision Computing*, Vol. 12, No. 10, pp. 639-647, December 1994.

[7] T. E. Hutchinson, K. P. White, Jr., W. N. Martin, K. C. Reichert and L. A. Frey, Human-computer interaction using eye-gaze input, *IEEE Transactions on System, Man and Cybernetics*, Vol. 19, No. 6, pp.1527-1533, 1989.

[8] A. E. Kaufman, A. Bandopadhay and B. D. Shaviv, An eye tracking computer user interface, *Proceedings* of *Research Frontier in virtual reality workshop*, IEEE Computer Society Press, October 1993, pp. 78-84.

[9] K.-N. Kim and R. S. Ramakrishna, Vision-based eye-gaze tracking for human computer interface, *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, 12-15, Oct. 1999, Vol. 2, pp.324-329.

[10] Y. Matsumoto and A. Zelinsky, An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement, *Proceedings of Fourth International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 499-504.

[11] G. A. Myers, K. R. Sherman and L. Stark, Eye Monitor, *IEEE Computer Magazine*, Vol. 3, pp. 14-21, 1991.

[12] R. Newman, Y. Matsumoto, S. Rougeaux and A. Zelinsky, Real-time stereo tracking for head pose and gaze estimation, *Proceedings of Fourth International Conference on Automatic Face and Gesture Recognition*, pp. 122-128, 2000.

[13] N. Otsu, A threshold selection method from gray-level histogram, *IEEE Transactions on Systems, man and Cybernetics*, Vol. SMC-9, pp.62-66, Jan. 1979.

[14] T. Rikert, M. Jones, Gaze estimation using morphable models, *Proceedings of International Conference on Automatic Face and Gesture Recognition*, 1998, pp.436-441.

[15] R. Safaee-Rad, I. Tchoukanov, K. C. Smith, B. Benhabib, Three-dimensional location estimation of circular features for machine vision, *IEEE Transactions on Robotics and Automation*, Vol. 8, No. 5, pp. 624-639, 1992.

[16] H. S. Sawhney, J. Oliensis and A. R. Hanson, Description and reconstruction from image trajectories of rotational motion, *Proceedings of IEEE International Conference on Computer Vision*, 1990, pp. 494-498.

[17] R. Stiefelhagen and J. Yang, Gaze tracking for multimodal human-computer interaction, Proceedings of *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-97*, 1997, Vol. 4, pp. 2617 -2620.

[18] K. Talmi and J. Liu, Eye and gaze tracking for visually controlled interactive stereoscopic displays, *Signal Processing: Image communication* Vol.14, pp. 799-810, 1999.

[19] Jian-Gang Wang, *Head-pose and eye-gaze determination for human-machine interaction*, PhD thesis, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, 2001.

[20] X. Xie, R. Sudhakar, and H. Azhang, On improving eye feature extraction using deformable templates, *Pattern Recognition* Vol. 17, pp.791-799, 1994.

[21] L. R. Young and D. Sheena, Survey of eye movement recording methods, *Behavior research methods and instrumentation*, Vol. 7, No. 5, pp.397-429, 1975.