8—29 On eye gaze determination via iris contour

Jian-Gang Wang[1], Eric Sung[1] and Ronda Venkateswarlu[2]

[1]Division of Control and Instrumentation, [2] Center for Signal Processing
School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore 639798
Email: eericsung@ntu.edu.sg, evronda@ntu.edu.sg

Abstract

We present a new approach for measuring the eye gaze using images that contains only one eye of a person. The model of the eye is investigated and gaze can be estimated by tracking the eyeball rotations. Iris contour is modeled as one planar circle and we estimate the ellipses of their projections onto a retinal plane. The unique solution of the orientations was obtained based on a geometric constraint, namely that the distance between the eyeball's center and the two eye corners should be equal to each other. The robustness of this approach was verified by experiments on synthetic and real image data. In order to obtain a higher resolution of the iris, a zoom-in camera is used. A general approach that combines head pose determination with gaze estimation is proposed.

1 Introduction

There are two components to gaze orientation: pose of human head and the orientation of the eye within their sockets. We have investigated these two aspects and will concentrate on the second part in this paper.

A method to determine the gaze of a person by using his/her one iris contour is demonstrated in this paper. 3D location of the iris contour can be determined besides the gaze orientation. Determining the gaze by analyzing the eyeball rotations is restricted due to low image resolution [2, 1, 3]. We propose tracking gaze using zoom-in camera while combining the pose result obtained from the second pose camera. We have ever proposed a method to determine the gaze based on two iris-contours –"two-circles" algorithm [6]. Two iris contours are modeled as circles and their projections are ellipses. The gaze can be estimated from the ellipse/circle correspondent. Unfortunately, there are two solutions of the gaze from an iris contour. We disambiguate the gaze results based on the knowledge namely that left and right iris boundaries are reasonably circular and that the normals of the two iris planes are parallel irrespective of eyeball rotations and head movement. In this paper, a "one-circle" algorithm has been proposed that only one iris is required for determining the gaze. Consequently the field of view of the camera can be narrowed further focus on one eye. The higher precision and robustness can be obtained due to the improvement of the resolution. The problem of having possible out-of-field views can be settled by guiding the gaze camera using the head pose information obtained by pose camera. We disambiguate the normals solution based on the prior knowledge of the eye model, namely that the distances between eyeball center and two eye corners should be equal to each other, we call it "distance constraint". The eyeball center can be calculated from the resulting iris center and the normal of the iris plane. Then the eyeball center is transferred to the pose camera coordinate system and the unique solution can be obtained by comparing the distance between the distance from the eyeball center to the two corners under the pose camera coordinate system.

Different from existing eye models (e.g. [10, 9]), the iris contour is modeled as one circle and its projection is so fitted as ellipse rather than circle. So our approach is more realism than the existing approaches.

2 Positioning

With the contour of the iris on the image, it will be shown that the prior knowledge of the eye model and the equation of the ellipse lead to a unique solution of the eye gaze.

2.1 Eye model

The eyeball is assumed to be a sphere with radius R. The iris is located at the front of eyeball and its contour is modeled as a circular ring of radius r (see Fig. 1(a)). The distance from the center of the eyeball to the iris plane is d. From Fig. 1(c), we can see:

$$R^2 = r^2 + d^2 \tag{1}$$

The optical axis of the eye is the line passing through the center of the eyeball and the center of the iris, it is also defined as the ***gaze direction***. By changing the gaze direction, the eyeball rotates around its center (see Fig 1(d)).

The upper and lower eyelids are modeled as two parabolas (Fig. 1(b)). The equation of the eyelid is:

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

which for the upper eyelid yields:

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

while for the lower eyelid yields:

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

In general, the upper and lower eyelids in real face images occlude parts of the iris contours and only the unoccluded iris edges can be used to fit the iris contour on the image plane. Consistently, the synthetic iris contour edges that lie between the upper and lower eyelids are located and used to fit the elliptical contour in our simulations. For instance, curves U_1L_1 and U_2L_2 shown in Fig. 1(b) are the fitting edges we want. These fitting edges can be located by using the equations of iris and eyelids. We will investigate the performances of our method using the eye model of dimensions close to that for human in the following sections.

2.2 "one-circle" algorithm

In order to improve resolution, as well as to solve the possible occlusion in "two-circles" algorithm, we develop new method, i.e. "one-circle", to determine gaze from one iris contour image. We use eye model to disambiguate the solutions of the orientation and center of the iris circle.

The distance between the two corners of an eye and the center of the eyeball should be equal to each other (see Fig.1 (e)):

$$O_sP_1 = O_sP_3 \qquad (5)$$

Call the two solutions of the normal of iris contour Q as $n_1(\cos\alpha_1, \cos\beta_1, \cos\gamma_1)$, $n_2(\cos\alpha_2, \cos\beta_2, \cos\gamma_2)$, and the correspondent solutions of the center of the iris contour are $O_{c1}(x_{01}, y_{01}, z_{01})$ and $O_{c2}(x_{02}, y_{02}, z_{02})$ respectively. Using the eye model, the center of the eyeball O_{si} can be calculated:

$$x_{si} = x_{0i} - d\cos\alpha_i,\ y_{si} = y_{0i} - d\cos\beta_i,\ z_{si} = z_{0i} - d\cos\gamma_i \qquad (6)$$

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

$$d = \sqrt{(R^2 - r^2)} \qquad (7)$$

The solutions of the center of the eyeball are translated to the pose camera coordinate system. The distance between the center of the eyeball and the two corners can be compared consequently. Due to the image noise, the unique solution should be the one that satisfies

$$O_sP_1 \approx O_sP_3 \qquad (8)$$

We calculate $O_{s1}P_1$, $O_{s1}P_3$, $O_{s2}P_1$ and $O_{s2}P_3$, If

$$|O_{s1}P_1 - O_{s1}P_3| \le |O_{s2}P_1 - O_{s2}P_3| \qquad (9)$$

then $(n_1\ O_{c1})$ is the solution what we want, else (n_2, O_{c2}) is the solution.

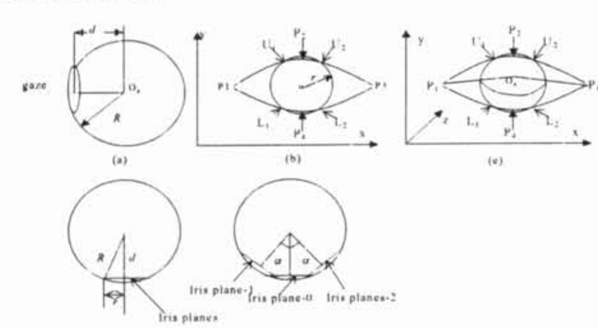

Fig. 1 Eye model

From the observed perspective projection of a circle having known radius, it is possible to infer analytically the 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, e.g. [7, 5].

2.3 Degenerate cases of "one-circle" algorithm

In our human-machine applications, the "one-circle" algorithm degenerates when the following condition is satisfied: the two iris contours are 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 two irises are symmetrical about the optical axis. Thus it is impossible to distinguish, from the ellipse, whether the person is looking upwards or downwards. We can see this in Fig. 4 (a) of section 3.1. Fortunately, this degenerate case can be prevented in our application by simply putting the camera slightly skewed to the face, See Fig. 2.

Fig. 2 Integration of pose and gaze, Right: zoom-in gaze camera; Left: pose camera.

3 Experimental investigation and results

We have tested gaze determination method using a Pentium450 PC, 128MB RAM (see Fig. 2). The algorithm runs on Image-Pro Plus software/SDK with Matrox Meteor-II imaging board. The pose camera is mounted on a fixed tripod while the gaze camera is mounted on a computer controlled pan-tilt unit (PTU-46-17.5). The distance between camera and operator is about half a meter.

3.1 Simulations on synthetic data

The coordinate system of the camera is defined: the image plane is defined as the X-Y plane and the Z-axis is along the optical axis of the camera and pointing toward the frontal object. The initial plane is where the iris is positioned parallel to the X-Y plane of the camera. The coordinate system of eyeball is defined: center of the eyeball is set to be the origin, the X- Y- and Z-axis set parallel to X- Y- and Z-axis of the gaze camera coordinate system respectively. The projective ellipse can be fitted by least-squares method. The 3D location (six degrees of freedom, three for translation and three for rotation) can be calculated from ellipse/circle correspondent. The errors between the calculated results and corresponding original synthetic data can be obtained. These sets of errors are obtained under different poses. An example is given as follows. The size of the image is 640 × 480. The intrinsic parameters of the camera are described as follows.

$$u_0 = 320, v_0 = 240, f_x = f_y = 5500 \quad (10)$$

where (u_0, v_0) are 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 coordinate (represented in the coordinate system of camera) of the center of the circle given in *cm*, is:

$$(x_c, y_c, z_c) = (0, 0, 60) \quad (11)$$

The radius of the iris is set as: $r = 0.65$ *cm*.

The 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. Once the radius has been calibrated, both the radius of the eyeball and the distance between the eyeball's center to the iris plane can be obtained consequently. The distance between two extreme corners of the parabola, see Fig. 1, is assumed: $P_1P_3 =$ 3.5 *cm*. The ratio of the radius of a person's iris and the radius of his/her eyeball in 3D space is assumed to be 0.5. So the radius of the eyeball is: $R = 1.3$ *cm*. The distance from eyeball center to the iris plane is calculated from equation (7): $d = 1.1.3$ *cm*.

In our experiments, we found that "distant constraint" is robust. The same unique solution can be obtained for different poses even though the ratio is varied ±50%.

The top and bottom points of parabola are symmetry about initial iris center, see Fig. 1, the distance:

$$P_2O = P_4O = 0.6\ cm \quad (12)$$

We consider the possible poses of eye iris as follows. Let the head look towards the camera. The eyeball is rotated around its center about their own *Y*- (azimuth) and X-axis (elevation) respectively. Consequently the iris contour will rotate following the eyeball. We project the rotated eyeball and the iris onto the image plane and re-estimate the poses and compute the error with the actual poses. The errors here are measured using the angle between the actual and estimated gaze directions. The iris contour, being an ellipse, is fitted [4] and its 3D location is recalculated. The points that lie within the upper and lower parabolas are found and used to fit the iris elliptical contour.

Rotating the eyeball about vertical axis from -50^0 to 50^0 in steps of 1^0 (azimuth) and about horizontal axis from -10^0 to 10^0 step 1^0 (elevation) forms a set of synthetic images. The error of the normal and center for different poses are shown in Fig. 4. The image when the rotation angle around *Y*-axis -30^0 and X-axis -15^0 is shown in Fig. 3. Two orientations of the iris plane can be obtained using least-squared fitted ellipse (dotted ellipse). Two eyeball centers are calculated using the normals, centers of the iris contour respectively and the known distance from eyeball center to the iris plane. Two arrows from the resulting eyeball centers to the iris centers, i.e. gaze according to our definition, are shown respectively. For every eyeball center, the distances between the eyeball center to the two eye corners are calculated, see Fig. 3. The eyeball center that satisfied the distance constraint, i.e. the difference of above two distances is smaller, will be selected. Consequently the unique solution of the normal and center of the are obtained.

The errors of the gaze, i.e. errors between calculated results and actual synthetic values, for different poses are shown in Fig. 4 (a). The errors of the center of the circle are shown in Fig. 4 (b). The results are satisfied. The maximum error of the gaze due to eyelids' occlusion is 1.5^0 while the maximum error of the center of the iris is 0.8 *cm*. However using "two-circles" algorithm, the maximum error of the gaze is 3^0 while the maximum error of the center of the iris is 1.2 *cm*.

Fig.3 Disambiguating the gaze solutions using distance constraint when the rotation angle around *Y*-axis -30^0 and X-axis -15^0 (a) Cross points are the eyeball's center calculated using the two solutions of the iris's center and the normals (b) The eyeball's center (cross point) obtained using "distance constraint" in Fig 3(a).

In section 2.3, we have discussed the degenerate cases of our algorithm, i.e. the normal direction constraint we used is ill conditioned for nearly front-parallel position, and this can be seen here from Fig 4. This can be overcome by using the distance constraint for continuous series of frames. We expect the solution of the orientations should not flip quickly from frame to frame. For our application, namely human-machine interaction, the camera is mounted on top of computer monitor and slightly skewed to the operator's face skewed operator (see Fig. 2). So the degenerate case is avoided in our application.

The experiments show that the algorithm we developed is robust enough. When we disturb 2D position of fitting points of the iris with a standard Gaussian noise (zero mean, standard variance equals 1), the mean errors (100 times of simulation) of the gaze of different poses are less than 4^0, the error of the centers of two circle is less than 1.2 *cm*.

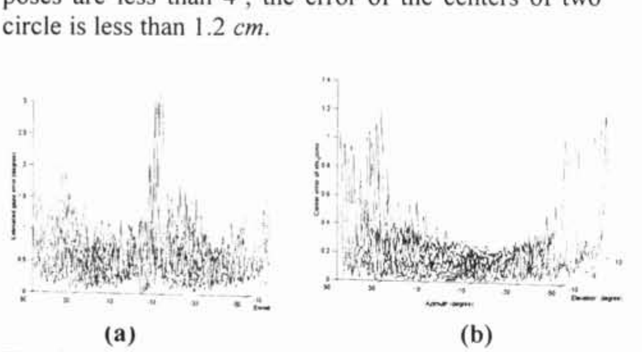

Fig. 4 (a) The error of the normal for different poses (b) The error of the center for different poses

3.2 Experiments on real images

The most prominent and reliable features within the eye region are edges of the iris. Because we wish to model iris contours on the image plane as an ellipse, obviously the existing iris detection methods using circular edge operators [10, 9] cannot be used. Instead, we detect the iris edge (bright-to-dark and dark-to-bright step edge) using a (3 x 3) vertical edge operator and morphological "open" operations. First, the eye-region is located using our pose-determination algorithm [8]. Since the corners of the eyes are already known, the iris detection is executed on a small region between them. A threshold is selected based on the histogram of the eye region, and then the eye is segmented using the threshold. The the morphological "open" operation, with a 3 x 3 structuring element, is applied to separate the iris from the eyelid. Then a vertical edge operator is used to detect the edges of the iris. Once the reliable iris edges are obtained, iris contour is least squares fitted to an ellipse [4]. An example of iris detection is shown in Fig. 5.

In the experiments on real scene, a zoom lens 30-55 *mm* is used in our real image experiments that guarantee the eyes are big enough when the distance between the human face and the monitor is within 70-100*cm*. The distance here is suitable for our application, which is for human-machine interaction.

The ratio of the radius of the iris is calibrated. Some of the gaze determination results are shown in Fig. 6. We can see the results are consistent with our observations.

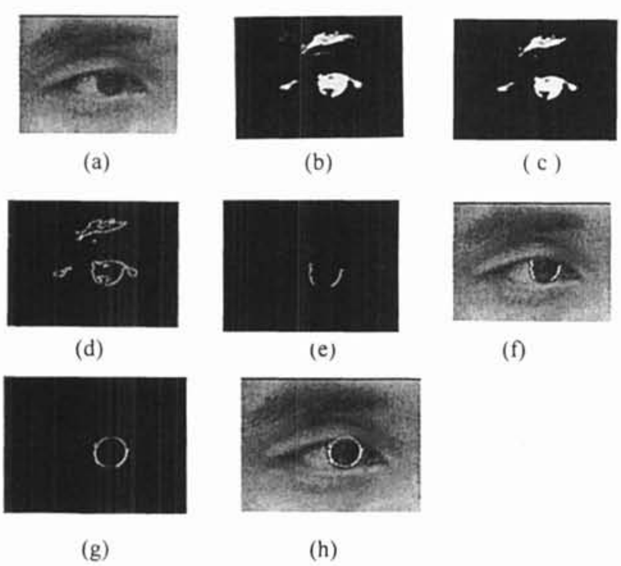

Fig. 5 Iris detection
(a) Original image (lens 55 mm) (b) Thresholding results (c) Morphological "open" operation (d) Vertical edges (e) Two longest edges by region-following (f) Overlay edges onto the original image (g) Edge result and Least-squares fitted ellipse (h) Overlay edge result and least-squares fitted ellipse onto the original image

Fig. 6 Eye gaze determination results

4 Conclusion

In this paper, we present a non-invasive method of robustly estimating the eye gaze by iris imaging. This method is based on the human eye model and a simple distance principle that we proposed. This principle states that the distance between the eyeball center and the two eye corners should be equal to each other. The orientation of an iris contour can be uniquely deduced from an image of one circle lying on the iris plane. This principle is applied to the eye-gaze by observing that the contour of the iris is circular and hence it is a circle that we are looking for. Others have never tried the use of iris contour in this way before.

The eye gaze method above is integrated with a head pose estimation module and together will offer great potential especially in applications such as virtual reality, video conferencing, and human-machine interface/controls. 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] Kay Talmi, Jin Liu, Eye and gaze tracking for visually controlled interactive stereoscopic displays, *Signal Processing: Image communication*, Vol. 14, pp.799-810, 1999.

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

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

[4] F.L.Bookstein, Fitting conic sections to scattered data, *Computer Graphics and Image Processing*, Vol. 9, pp.56-71, 1979.

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

[6] J.G.Wang and E.Sung, Gaze determination via images of iris, to appear in *the 11th British Machine Vision Conference*, September 11-14, 2000, Bristol, UK.

[7] H.S.Sawhney, J. Oliensis and A.R.Hanson, Description and reconstruction from image trajectories of rotational motion, Proceedings of *ICCV'95*, pp. 494-498, 1995.

[8] E. Sung and J.G. Wang, Head pose and eye gaze for human-computer interface, to appear in *the sixth International Conference on Control, Automation, Robotics and Vision*, 5-8 December 2000, Singapore.

[9] J-Y Deng and F. Lai, Region-based template deformation and masking for eye-feature extraction and description, *Pattern Recognition*, Vol. 30, No. 3, pp. 403-419, 1997.

[10] Yuille, A. L., Hallinan, P. W., Cohen, D. S., Feature extraction from faces using deformable temples, *International Journal of Computer Vision*, Vol. 8, No. 2, pp.99-111, 1992.