A Method of Gaze Direction Estimation Considering Head Posture

Wan-zhi Zhang, Zeng-cai Wang, Jun-kai Xu and Xiao-yan Cong

School of Mechanical Engineering, Shandong University, Jinan, China
wangzc@sdu.edu.cn, zhangwanzhi@163.com

Abstract

A new method for head posture and gaze direction estimation is proposed. Firstly, three models of head position are established and postures are judged based on triangle attribute constituted by eyes and mouth. Then pupil is located using Hough transform in eye area. With the method of horizontal and vertical projection and eye prior knowledge, the normal eye outline is fitted. Finally, gaze direction is estimated according to the position of pupil in normal state eye and head posture. The experimental results demonstrate the proposed method can accurately detect head posture and gaze direction. For considering head posture, the method has more accuracy in gaze estimation.

Keywords: *gaze estimation; head posture; pupil; Hough transform; gray projection*

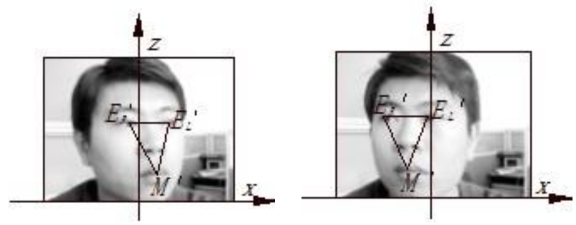
1. Introduction

Eyes usually contain more available information, in which gaze direction is one of very important information. For example, long time unchanged gaze direction indicates that the driver has distracted or been in fatigue. With the development of computer, gaze direction tracking has become hot. Head position judgment is the premise of gaze direction. Recently, the methods of gaze direction research mainly contain: neural network, feature extraction, iris reflection flare and mapping function, *etc.* SeWell [1-3] has proposed neural network which had higher precision and real-time, but the method is more easily affected by illumination and has poor robustness. Yamazoe [4-5] has analyzed the relationship between pupil and eye socket to estimate direction, but he doesn't consider head posture. Iwata [6-9] has used Purkinje flare to detect direction in the basis of pupil detection, but the method uses infrared light source which has a certain invasive. Nguyen [10-11] has used single mapping between facial information and gaze points to estimate direction, but the method has lower adaptability and accuracy. Besides, the above algorithms have a common fault which does not consider head posture [12-13]. At present, the methods of head posture analysis generally are divided into two classes: geometric analysis based on 2D facial features and appearance method. Wu [14] has analyzed the geometric relationships of five facial points to estimate head posture, but the algorithm needs to locate five points which is difficult and vulnerable to facial expressions change. Chen [15] has used adaboost method to locate eyes and mouth and analyzed head posture with the characteristic triangle, but he has only considered two simple postures. Lu [16] has proposed nonlinear nuclear transformation which has overcome traditional linear estimation shortcomings, but the algorithm needs a large number of training samples.

We propose a new method for head posture and gaze direction estimation. Firstly, head posture is judged based on triangle attribute constituted by eyes and mouth. Then pupil is located using Hough transform in eye area. With the method of horizontal and vertical projection and eye prior knowledge, the normal eye outline is fitted. Finally, gaze direction is estimated according to the position of pupil in normal state eye and head posture.

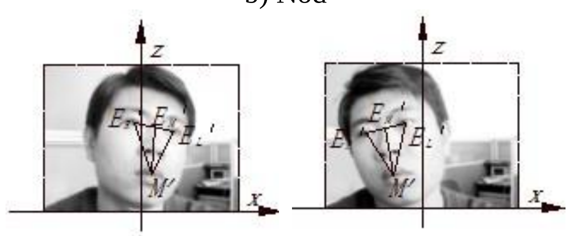
2. Head Posture Analysis

In videos, the relative position of eyes and mouth will change when head posture changes and the change has some geometric features. Therefore, we can analyze head posture based on the geometric features. Research shows that any head posture can be divided into three basic postures (turn, nod, swing), as Figure 1 indicates. *X*, *z* axis is shown in picture and *y* axis is perpendicular to the *x*-*z* plane.



a) Turn

b) Nod



c) Swing

Figure 1. Head Posture

E_L, E_R, M respectively represents the center of eyes and mouth. According to facial geometric relationships, when head faces camera, $\Delta E_L E_R M$ is an isosceles triangle; when head position changes, triangle attributes will change. Therefore, we can analyze head posture with triangle attributes. Such as turn, when head rotates α around **z** axis, $\Delta E_L E_R M$ changes into $\Delta E_L' E_R' M'$. For convenience of calculation, two triangles are projected into *x*-*y* plane, as shown in Figure 2(a).

According to geometric relationships, we can deduce:

$$\alpha = \arccos(\frac{L_{EE}'}{L_{EE}}) \qquad (1)$$

Among them, L_{EE} is the distance of two eyes in positive head image; L_{EE}' is the distance of two eyes in rotating head image. When $M_x > 0$, α is positive.

Such as nod, which head rotates β around **z** axis, $\Delta E_L E_R M$ changes into $\Delta E_L' E_R' M'$.

The triangle attributes and distance between two eyes remain unchanged. H_{EM} changes into $H_{EM}{}'$. H_{EM} represents the distance between mouth and two eyes attachment. Two triangles are projected into **y-z** plane, as shown in Figure 2 (b).

According to geometric relationships, we can deduce:

$$\beta = \arccos(\frac{H_{EM}{}'}{H_{EM}}) \qquad (2)$$

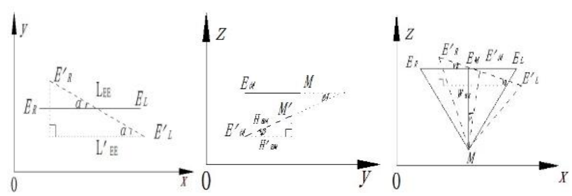
When $M_z{}' < M_z$, β is positive.

Such as swing, which head rotates γ around ***y*** axis, the only change is that triangular high has a certain angle with **z** axis. Two triangles are projected into **x-z** plane, as shown in Figure 2 (c).

According to geometric relationships, we can deduce:

$$\gamma = \arccos(\frac{W_{EE}}{L_{EE}}) \qquad (3)$$

W_{EE} represents horizontal distance of two eyes in image. When $E_{Lz} < E_{Rz}$, γ is positive.

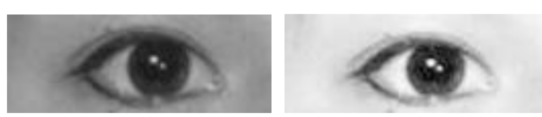


a) Turn b) Nod c) Swing

Figure 2. Projection Drawing

3. Pupil Detection

Due to the influence of cab light changing, the signal-to-noise and contrast of collected driver video image are lower and they affect detection results. Illumination compensation is necessary and the result is shown in Figure 3.



a) Original image b) Compensation image

Figure 3. Illumination Compensation

3.1. Edge detection

Because pupil's gray scope is different with other eye part, we use 'canny' operator to detect edge in binary image. Result is shown in Figure 4.

Figure 4. Image Binarization

3.2. Hough transform

Hough transform is a computer vision algorithm which can detect any object with known shape. The standard template of Hough transform can be defined as:

$$H(\Omega)=\sum_{j=1}^{N} p(X_i,\Omega) \quad (4)$$

Among them,

$$p(X,\Omega)=\begin{cases}1, \forall(X,\Omega):\{\Lambda: f(X,\Lambda)=0 \bigcap C_\Omega \neq \Phi\} \\ 0, \text{otherwise}\end{cases}$$

, $\{X_1, X_2, \cdots, X_N\}$ represents the feature point in image space. C_Ω represents accumulators unit in parameters space which center is Ω. $H(\Omega)$ represents the value of Hough transform in Ω. $f(X,\Lambda)=0$ represents parametric constraint equations. The key of Hough transform is to find a reasonable parametric expression and appropriate discrete space.

As pupil is round, the analytical expression is:

$$(x-a)^2+(y-b)^2=r^2 \quad (5)$$

(*a*, *b*) is center and *r* is radius. So accumulators' data structure must be 3D. Although Hough transform can detect any curve with known shape, accumulators' data structure presents exponential growth along with the growth of curve parameters. There will be larger amount of calculation. If we use the priori information of edge direction, we can significantly reduce the amount of calculation. Because the pupil size is fixed, the number of unit accumulators greatly decrease in parameter space. Pupil detection is shown in Figure 5.

Figure 5. Pupil Detection

4. Eye Contour Extraction

The outline of normal eyes is constitutes by two parabolas which are corresponding to the eyelids. Two parabolic intersections are corresponding to the canthus. Two parabolic centers are corresponding to the limit point of eyelid. Therefore, the key of eye contour extraction is locating the canthus and eyelid points. We use the algorithm of integral projection to determine the positions. In horizontal integral projection, peaks are the vertical coordinates of eyelid; troughs are vertical coordinates of canthus. In vertical integral projection, peaks are the horizontal coordinates of eyelid; troughs are horizontal coordinates of canthus. Locating result is shown in Figure 6. The method is not sensitive to head position and does not need high resolution images.

Figure 6. Eye Contour Extraction

5. Gaze Direction Estimation

We detect eye gaze direction according to different pupil positions. Generally, gaze directions can be divided into nine directions, as shown in Table 1.

Table 1. Eye Gaze Direction

Upper left	Upper middle	Upper right
Middle left	Middle	Middle right
Lower left	Lower middle	Lower right

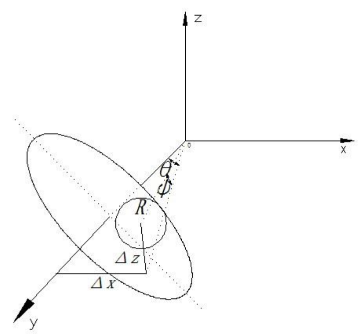


Figure 7. Direction Estimation

When we don't consider head position, the position of pupil relative to eye contour is shown in Figure 7. Ψ is up or down angle; θ is right or left angle. So, we can deduce:

$$\psi = \arcsin(\frac{\Delta z}{R}) \quad (6)$$

$$\theta = \arcsin(\frac{\Delta x}{R\cos\psi}) \quad (7)$$

Among them, R is eyeball radius; Δx and Δz respectively is the distance between pupil and eye contour center.

When we consider head position, eye gaze direction formula can be expressed as:

$$\psi^{'} = \arcsin(\frac{\Delta z\cos\gamma}{R}) + \beta \quad (8)$$

$$\theta^{'} = \arcsin(\frac{\Delta x\cos\gamma}{R\cos\psi}) + \alpha \quad (9)$$

6. Algorithm Process

The algorithm flow chart is shown in Figure 8.

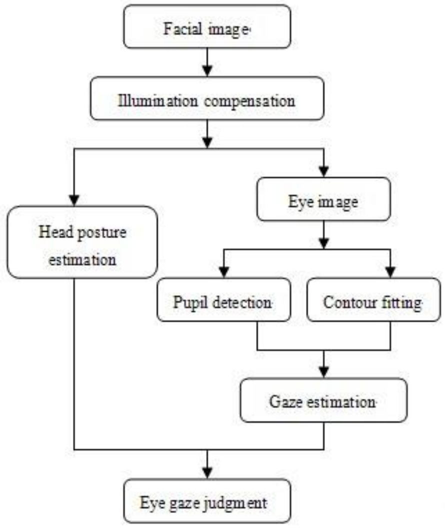


Figure 8. Algorithm Flow Chart

7. Experiments and Discussions

The experiment is based on face database which we have built. Part of experimental results is shown below.

Images in Figure 9 are from any head position. In experiment, the distance between face and camera is 40cm. In image, when head posture is normal, the distance between two eyes is 4.9cm and the distance between mouth and two eyes attachment is 5.7cm. Head posture detection results are shown in Table 2 and we can see that any head position can be described by three basic postures.

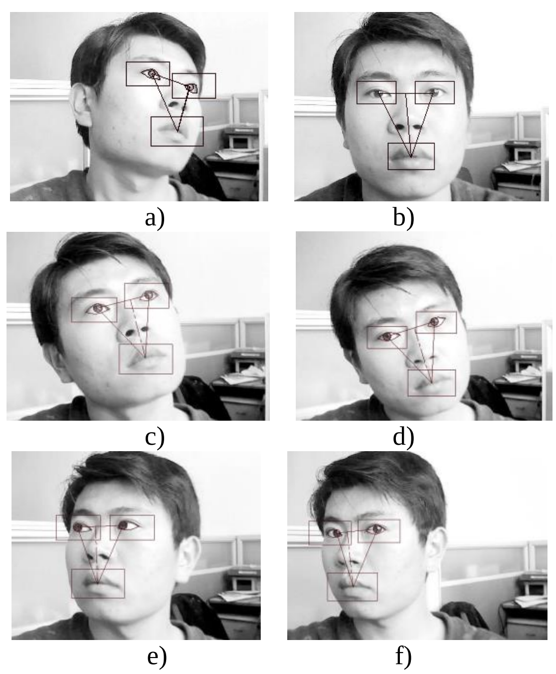


Figure 9. Head Posture Estimation

Table 2. Results of Head Posture Estimation

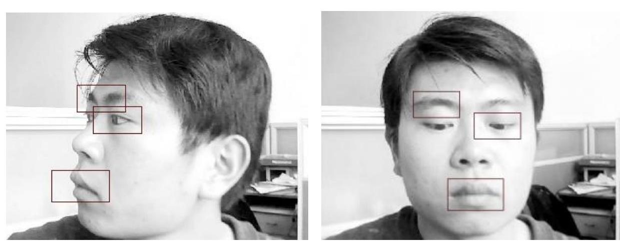
Image	Parameters					
	L_{EE}'	H_{EE}'	W_{EE}	α	β	γ
a	3.3	3.8	3.1	47.6	-48.2	50.7
b	4.9	5.7	4.9	0	0	0
c	4.2	4.8	3.9	31.0	-35.6	-37.3
d	4.4	5.1	4.2	26.1	26.5	-31.0
e	3.5	4.6	4.4	-44.4	-36.2	-26.1
f	3.4	5.5	4.3	-46.1	-15.2	-28.6

After head posture calculation, results of eye gaze estimation are shown in Table 3 according to formula (8, 9). We can deduce that considering head posture is very necessary for gaze estimation.

Table 3. Results of Gaze Estimation

Image	ψ /(°)	θ /(°)	ψ' /(°)	θ' /(°)	Result
a	11	4	-37.2	51.6	Upper left
b	0	0	0	0	Middle
c	12	5	-23.6	36.0	Upper left
d	-14	-21	12.5	4.1	Lower right
e	2	-42	-34.2	-86.4	Upper right
f	15	46	0	0	Middle

Two types that the proposed algorithm can't estimate gaze direction correctly are shown in Figure 10. One type is that head rotation angle is too big; the other is that eye open is smaller or there is much pixels noise.



a) b)

Figure 10. Failure Image Estimation

8. Conclusion

We propose a new method for head posture and gaze direction estimation. Three models of head posture are established and head posture is judged based on triangle attribute constituted by eyes and mouth. Then pupil is located using Hough transform in eye area. With the method of horizontal and vertical projection and eye prior knowledge, the normal eye outline is fitted. Finally, gaze direction is estimated according to the position of pupil in normal state eye and head posture. The experimental results demonstrate the proposed method can accurately detect head posture and gaze direction.

References

[1] S. Weston and K. Oleg, "Real-time eye gaze tracking with an unmodified commodity webcam employing a neural network", Conference on Human Factors in Computing Systems-Proceedings, no. 28, **(2010)**, pp. 3739-3744.

[2] Z. Zhiyu, Z. Jianxin and Z. Sixuan, "Eye tracking based on rule and grey prediction", Journal of Computational Information Systems, vol. 4, no. 6, **(2008)**, pp. 2947-2953.

[3] J. Chi, C. Zhang, T. Hu, *et al.*, "Eye characteristic detection and gaze direction calculation in gaze tracking system", Control and Decision, vol. 24, no. 9, **(2009)**, pp. 1345-1350.

[4] Y. Hirotake, U. Akira and Y. Tomoko, *et al.*, "Remote gaze estimation with a single camera based on facial-feature tracking without special calibration actions", Eye Tracking Research and Application Symposium, **(2008)**, pp. 245-250.

[5] Y. Hirotake, U. Akira and A. Shinji, "Gaze estimation with a single camera using facial feature tracking", Journal of the Institute of Image Information and Television Engineering, vol. 61, no. 12, **(2007)**, pp. 1750-1755.

[6] M. Iwata and Y. Ebisawa, "Pupil mouse supported by head pose detection", Proc VECIMS 2008[C], **(2008)**, pp. 178-183.

[7] J. W. Lee, C. W. Cho and K. Y. Shin, *et al.*, "3D gaze tracking method using Purkinje image on eye optical model and pupil", Optics and Lasers in Engineering, vol. 50, no. 5, **(2012)**, pp. 736-751.

[8] W. Zhang, T. Zhang and S. Chang, "Gaze estimation based on extracted parameter of one iris", Journal of Optoelectronics Laser, vol. 22, no. 6, **(2011)**, pp. 916-920.

[9] A. Villanueva and R. Cabeza, "A novel gaze estimation system with one calibration point", IEEE Transactions on Systems, vol. 28, no. 4, **(2008)**, pp. 1123-1138.

[10] B. L. Nguyen, Y. Chahir and F. Jouen, "Free eye gaze tracking using Gaussian processes", IPCV 2009[C], **(2009)**, pp. 137-141.

[11] C. Zhang, J. Chi and Z. Zhang, *et al.*, "A novel eye gaze tracking technique based on pupil center cornea reflection technique", Chinese Journal of Computers, vol. 33, no. 7, **(2010)**, pp. 1272-1285.

[12] R. Valenti, N. Sebe and T. Gevers, "Combining head pose and eye location information for gaze estimation", IEEE Transactions on Image Processing, vol. 21, no. 2, **(2012)**, pp. 802-815.

[13] S. Zhu, Z. Wang, S. Wang, *et al.*, "Eye localization and state analysis of head multi-position in color image", Journal of Chongqing University, vol.33, no. 1, **(2010)**, pp. 20-26.

[14] J. Wu and M. M. Trivedi, "A two-state head pose estimation framework and evaluation", Pattern Recognition, vol. 41, no. 3, **(2008)**, pp. 1138-1158.

[15] Z. Chen, F. Chang and C. Liu, "Pose parameters estimate based on adaboost algorithm and facial feature triangle", Geomatics and Information Science of Wuhan University, vol. 36, no. 10, **(2011)**, pp. 1164-1168.

[16] Y. Lu, Z. Wand and X. Li, "Head Pose Estimation Based on Kernel Principal Component Analysis", Opto-Electronic Engineering, vol. 35, no. 8, **(2008)**, pp. 63-66.

Authors

ZHANG Wan-zhi is currently Ph.D candidate in Shandong University. His research interests includes automotive safety and man-machine engineering, image processing. He has published three papers at Journal of Optoelectronics • Laser and Journal of Chongqing University. All of them have been indexed by EI Compendex. Currently, he is doing the research of Study on modeling and early-warning critical technologies of driver visual distraction and fatigue.

Wang Zengcai is currently the professor and the doctoral supervisor of vehicle engineering at Shandong University, the deputy director of Shandong mechanical engineering society hydraulic and pneumatic professional committee. His research includes image processing, vehicle engineering, Machinery monitoring and control, hydraulic drive and control. He has published a book and a number of journal articles, in which there are more than twenty papers indexed by EI. He also undertakes a number of scientific research projects and one of them won the national safety production science scientific or technical award.

Xu Junkai is currently the master of mechatronic engineering in Shandong University. He is study on the information processing and mechanical control. He knows well several computer languages and has basic hardware design ability. He also takes part in a national project supported by national natural science foundation. Now he is interested in pattern recognition and working on the improvement of recognition accuracy.

CONG Xiao-yan is currently Ph. D candidate in Shandong University. Her research interests include automotive safety, man-machine engineering and driving fatigue detection. During her working for Master degree, she researched cab comfort and driving fatigue detection with applied ergonomics.