Available online at www.sciencedirect.com
ScienceDirect
Progress in Natural Science 18 (2008) 1155–1159
Progress in Natural Science
www.elsevier.com/locate/pnsc

Minimization of influence by the drooping eyelid in extracting pupil center for eye tracking systems

Xiaopeng Wei [a,b,*], Dongsheng Zhou [a], Qiang Zhang [b,*], Boxiang Xiao [a]

[a] School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China
[b] Liaoning Key Laboratory of Intelligent Information Processing, Dalian University, Dalian 116622, China

Received 4 March 2008; received in revised form 24 March 2008; accepted 24 March 2008

Abstract

Precise pupil center detection is an important factor for gaze tracking in video-oculography (VOG) systems. Existing methods perform well to extract the features when the area of pupil in eye image is clear, whereas, interferences, such as eyelashes, corneal reflection etc., will lead to a low success rate. One main reason is the closure of eyelids. In this paper, a systemic 3D transformation algorithm is proposed to accurately ascertain the pupil center, in spite of the complicating factors mentioned above. Experiments show the good performance of our method. And the pupil center could be extracted accurately, even though only 25% of the pupil is visible.

Keywords: Pupil center; VOG; 3D transformation; Projection transformation; Least squares

1. Introduction

As a reality, eye gaze plays an important role in human communication, and it has the following specialties: directness, naturalness, intercommunication etc. So people exhibit keen interest and have conducted a great deal of research on it. Following the progress of research, there appear three main techniques for detecting eye movement and eye gaze: Sclera search coil (SSC), electro-oculography (EOG) and infra-red video-oculography (VOG) [1–3]. Among them, SSC and EOG are the intrusive techniques, i.e., they require some special equipment to put physical contact with the users. These equipments include, for example, coils, electrodes, contact lens etc. [4]. Those are inconvenient to be used and will produce some discomfort after a long wearing. Contrasting with the above two methods, VOG is a non-intrusive technique. It is mainly vision based, i.e., it is trying to determine the position of the eye using the information provided by a camera that captures eye images and employing computer vision techniques together with digital image processing algorithms [4,5]. Due to its insignificant artifacts, noninvasive feature, and freedom of head movement, VOG is becoming more and more attractive. Although there are so many advantages, VOG also has the shortage that it cannot determine the precise center and radius of the pupil when the upper eyelids droop [3,6,7], and this shortage will raise the error rate for the whole VOG system.

In order to solve the problem mentioned above, many researchers have carried out a great deal of work in this area, and have presented different techniques. The conventional ways are averaging the coordinates of the edge points of the pupil [8], fitting a circle or ellipse to the pupil boundary [9,10], calculating the center of gravity of the pupil (COG) [11,12]. Some researchers employed new techniques and improved those traditional methods. Ebisawa [13] suggested an accurate image-based eye gaze estimation

* Corresponding authors. Tel.: +86 411 87403728; fax: +86 411 87403733.
E-mail addresses: xpwei@dlu.edu.cn (X. Wei), zhangq@dlu.edu.cn (Q. Zhang).

1002-0071/$ - see front matter © 2008 National Natural Science Foundation of China and Chinese Academy of Sciences. Published by Elsevier Limited and Science in China Press. All rights reserved.
doi:10.1016/j.pnsc.2008.03.019

system using two IR light sources placed in different locations: near and away from the optical axis of the camera. This greatly enhanced the contrast between the pupil and the other parts of the eye, and will be conducive to extract the pupil contour. Yoo and Chung [14] utilized the method of Ebisawa and presented a boundary detection method similar to the active shape model (ASM). He first used a circle as an initial contour to locate the rough position of the pupil area. Then a simple temple [000111] was used for edge detection. Finally, the boundary of the pupil was fitted by using the direct ellipse algorithm. This method is robust and good at pupil detection, but it is not good at pupil feature extraction. Some interference, especially dropping eyelid will lead to inaccurate results. Zhu et al. [15] utilized the curvature characteristics of the pupil boundary to eliminate the deterring factors caused by the eyelids, eyelashes and so on. This technique is robust and accurately estimates pupil center with less than 40% of the pupil boundary points visible. However, there still is a shortage that the algorithm could lead to false results when the obscuration of the pupil by an eyelid is small and the curvature transition from the visible pupil to the eyelid is smooth. Moore et al. [16] used the polar cross correlation method for the measurement of ocular torsion. He assumed the eyeball as a sphere and constructed a three-dimensional eyeball model by using the general anatomic knowledge of the human eye. Recently, Kim and Park [17] developed a precise pupil size calculation technique by the Moore's method. Moore and Kim mainly used a rotation method, i.e., they transformed the coordinates of boundary of the pupil from eccentric eye position into reference position which paralleled to the image plane of the camera, and then the shape of the pupil could be constructed.

In this paper, a new technique for extracting the features of pupil, especially obtaining the center of pupil is developed. It is assumed that the shape of the human eyeball is a sphere, and the contour of the pupil is a circle which belongs to a smaller virtual sphere concentric with the eyeball. A 3D eyeball model is reconstructed. And then, the iterative least square (ILS) algorithm, projection transformation and circle fitting techniques are adopted to extract the features of the pupil. The main differences between ours and Kim's are that we used a spatial plan fitting algorithm to replace the ration algorithm which is unstable and complicated to be calculated.

2. Image acquisition and pretreatment

In our experiment, a goggle equipped with a CCD camera and infra-red luminescent diodes were mounted on the head of a subject to acquire the eye images. The camera worked at 50 frames/s, and the size of captured pictures was 640 × 480 pixels (see Fig. 1). The picture that the eye looked straight ahead at the camera lens was selected as the reference picture. Because of the IR light illumination, the pupil is very outstanding from the whole image. It is easy to ascertain the reference pupil center by using the hybrid projection functions algorithm (HPF) [12]. The HPF is defined as:

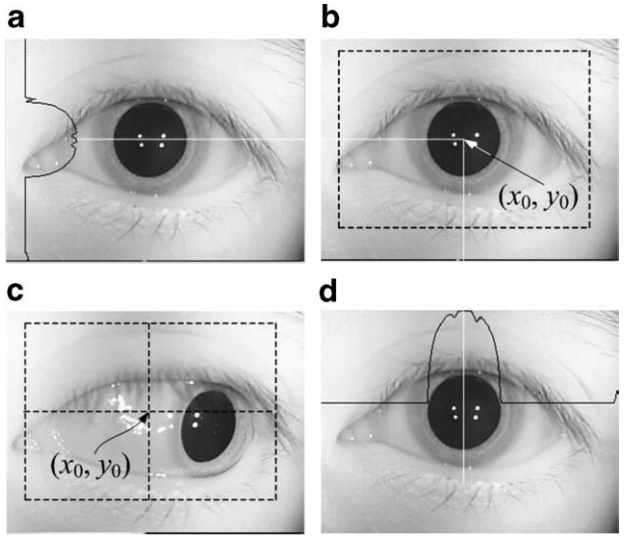

Fig. 1. The captured eye pictures. (b) The reference picture which eye looked straight to the camera lens. The size of a, b, c and d is 640 × 480 pixels. (a) Illustrates the vertical of *HPF*, (b) illustrates the horizontal of *HPF*, (c) illustrates the area of normalization. (x_0, y_0) is the center point of the pupil of the reference picture (b), and it is set as the reference point which is used to normalize the sequence pictures.

$$\begin{cases} HPF_v(x) = \alpha IPF_v(x) + \beta VPF_v(x) \\ HPF_h(x) = \alpha IPF_h(x) + \beta VPF_h(x) \end{cases} \tag{1}$$

where v is the vertical direction in the images, h is the horizontal direction. α and β are constants, and satisfied $\alpha + \beta = 1$. They are determined by the experiment. *IPF* is the integral projection function, *VPF* is variance projection function, and they can be defined as follows:

$$\begin{cases} IPF_v = \frac{1}{y_2 - y_1} \int_{y_1}^{y_2} I(x, y)\mathrm{d}y \\ IPF_h = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} I(x, y)\mathrm{d}x \end{cases} \tag{2}$$

$$\begin{cases} VPF_v(x) = \frac{1}{y_2 - y_1} \sum_{y_i = y_1}^{y_2} [I(x, y_i) - IPF_v(x)] \\ VPF_h(y) = \frac{1}{x_2 - x_1} \sum_{x_i = x_1}^{x_2} [I(x_i, y) - IPF_h(y)] \end{cases} \tag{3}$$

where $I(x, y)$ is the intensity of a pixel at location (x, y). $[y_1, y_2]$ and $[x_1, x_2]$ are the intervals of vertical and horizontal of an image, respectively. The detection of reference pupil center is illustrated in Fig. 1.

Normalization followed data collection in order to better locate the interested region. Set the center point as the origin point, and use a $l \times h$ rectangle to standardize the sequence images (see Fig. 1(c)). The value of l is approximately equal to the length of a general human eye, and the value of h is approximately equal to 1.5 times the width of a general human eye.

3. Pupil center extraction algorithm

3.1. Human eye structure

Fig. 2 shows the major structure of a human eye. It can be seen that the shape of the eyeball is approximately a sphere with a radius of about 11.5 mm [18,19]. The iris looks like a plane section and the boundary of it is similar to a circle on the sphere. The pupil is just like a circular aperture on the plane which has the same center as the iris.

3.2. 3D eye model

We first used the technique in Ref. [20] to achieve the conversion factor δ which can be used for transforming values from world coordinate system (WCS) to image coordinate system (ICS). Then the value of R_v was calculated according to the geometry of the eyeball as shown in Fig. 2.

$$R_v = \sqrt{R_e^2 - R_i^2 + R_p^2} \tag{4}$$

in which R_v is the radius of the virtual ball in ICS. $R_e = \delta \times 11.5$ is the radius of the eyeball. R_i and R_p are the radius of iris and pupil in ICS, respectively, and can be acquired through a sophisticated image processing algorithm [21]. The 3D eye model was developed by the right-handed Cartesian coordinate system and the above parameters. In Fig. 3, the pupil is at an eccentric eye position and the apparent shape of it is an ellipse. P is the space plane which contains the edge points of the pupil. Arrow d denotes the direction that parallels to the vector of the plane P.

3.3. Extraction algorithm

The relative position of the camera and eye is immovable. So, as mentioned above, the apparent shape of the pupil is an ellipse. Whereas, if we use an active camera where the optical of camera lens is always vertical to the vector of plane P, i.e., the camera image plane is always parallel to the plane P, the shape will always be a circle

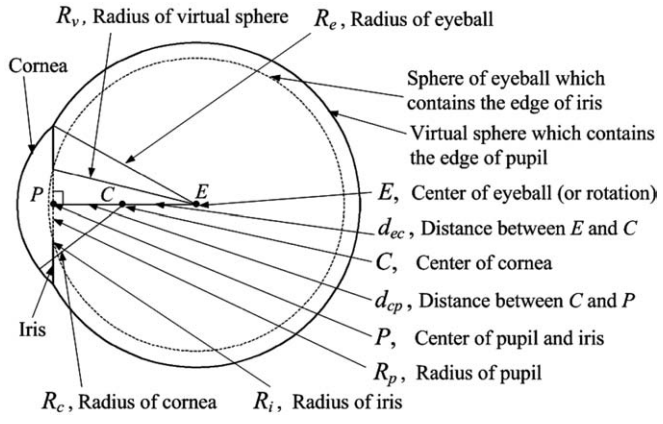

Fig. 2. Structure of a human eye.

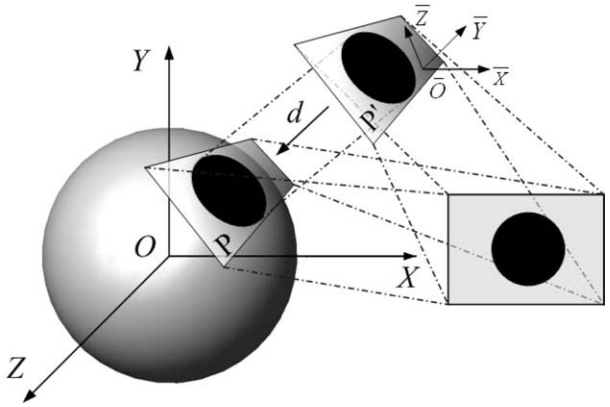

Fig. 3. The 3D eyeball model. P is the space plane which contains the edge points of the pupil, P' is a plane parallel to P.

which is simpler and more accurately fitted. So, considering those advantages, an algorithm of space plane fitting based on ILS algorithm was firstly used to acquire the plane P. Then, space projection and circle fitting algorithms were used to extract the features of the pupil. The detail is as follows.

The first step is the determination of the projection plane. From Section 3.2, a part of 3D coordinates of the edge points of the pupil has been achieved. Then, the fitting space plane P (shown in Fig. 3) can be calculated by using the ILS algorithm. A general space plane is described as the following equation:

$$Ax + By + Cz + D = 0 \tag{5}$$

The sum of vertical distance from all the discrete points to a plane can be defined as:

$$\sum_{n=1}^{N} d_n^2 = \sum_{n=1}^{N} \frac{(Ax + By + Cz + D)^2}{A^2 + B^2 + C^2} \tag{6}$$

According to the principle of least squares and the necessary condition of multivariate function extreme value, the following equations can be obtained:

$$\begin{cases} H\sum_{n=1}^{N}(Ax_n+By_n+Cz_n+D)\cdot x_n = A\sum_{n=1}^{N}(Ax_n+By_n+Cz_n+D)^2 \\ H\sum_{n=1}^{N}(Ax_n+By_n+Cz_n+D)\cdot y_n = B\sum_{n=1}^{N}(Ax_n+By_n+Cz_n+D)^2 \\ H\sum_{n=1}^{N}(Ax_n+By_n+Cz_n+D)\cdot z_n = C\sum_{n=1}^{N}(Ax_n+By_n+Cz_n+D)^2 \\ \sum_{n=1}^{N}(Ax_n+By_n+Cz_n+D)=0 \end{cases} \tag{7}$$

where

$$H = (A^2 + B^2 + C^2) \tag{8}$$

N is the number of scatter red space points of the pupil edge; D is a constant; (x_n, y_n, z_n) are the coordinates of those points; and $n = 1,2,\cdots,N$. After solving (8), (A, B,

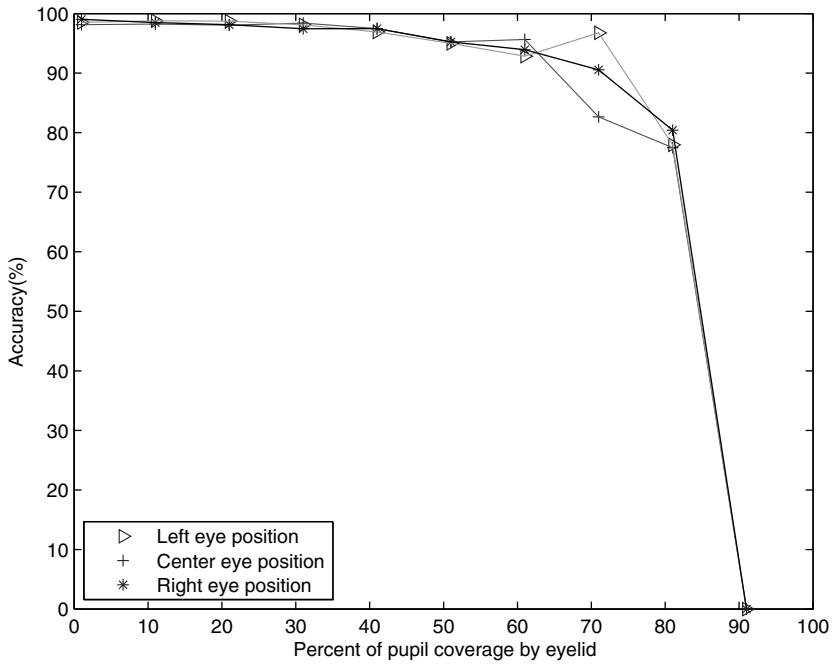

Fig. 4. Accurate rate under the effect of the upper eyelid drooping.

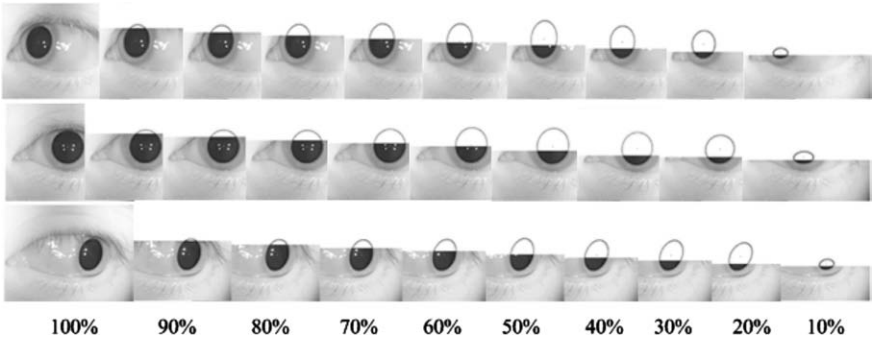

Fig. 5. Experimental results of pupil fitting. From left to right, the area of pupil was reduced. When the area of the pupil was reduced to 20%, the algorithm began to fail.

C) which is represented by D can be obtained, i.e., we get a series of parallel space planes which in different position determined by D. Once the value of D is fixed, the projection plane P' (shown in Fig. 3) can be determined.

The next step is the space projection transformation. In Fig. 3 the transformation from coordinate system $OXYZ$ to $\overline{OXYZ}$ can be represented as:

$$\begin{pmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{9}$$

where $[a_{11}, a_{22}, a_{33}]$, $[a_{21}, a_{22}, a_{23}]$, $[a_{31}, a_{32}, a_{33}]$ are vectors of axis $\overline{X}$, $\overline{Y}$ and $\overline{Z}$, respectively. And the others can be decided easily by the plane P'. In the projection transformation, homogeneous equations are used and the projection function is as follows:

$$\begin{pmatrix} \bar{x}_q \\ \bar{y}_q \\ \bar{z}_q \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\frac{a_{21}}{a_{23}} & 0 \\ 0 & 1 & -\frac{a_{22}}{a_{23}} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{M} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{10}$$

where

$$\mathbf{M} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{11}$$

$$\bar{x} = \bar{x}_q/q, \bar{y} = \bar{y}_q/q, \bar{z} = \bar{z}_q/q \tag{12}$$

$$a_{i4} = -(a_{i1}x_0 + a_{i2}y_0 + a_{i3}z_0), i = 1, 2, 3 \tag{13}$$

(x_0, y_0, z_0) is the coordinate of the origin $\overline{O}$ in the coordinate system $OXYZ$.

Next, a circle can be fitted by the traditional principle that three points which are not on the same line can define

a circle. And then, the anti-projection is employed to finish the whole pupil feature extracting algorithm.

4. Experiments

To evaluate the performance of our algorithm, three pictures where pupils are on the right, left and in the middle of the eye, respectively were selected. We simulated the degree of eyelid droop as the pupil was covered gradually from 10% to 90%. The detection performance was evaluated based on subjective judgment since no ground truth was available. We adopted and modified the evaluation criteria in Ref. [22]. The hand-localized pupil centers served as ground truth. Let $H(i, j)$ denote hand-localized pixel and $E(i, j)$ the detected pixel. The distance Dis of H and E is defined as $\|H - E\|_2$. The accurate rate of the algorithm is defined as:

$$Accurate = \left(1 - \frac{Dis}{\max\{Dis\}}\right) \times 100\% \quad (14)$$

The experiments are illustrated in Figs. 4 and 5.

5. Conclusion

In the eye tracking system, estimation of the pupil radius and center using the traditional circle or ellipse fitting algorithm in 2D image would result in glaring error, mainly because the pupil is not an ideal circle or the pupil area is distributed by the drooping eyelid. In this paper, an accurate eye feature extraction algorithm is presented. It overcomes those shortages effectively by using a 3D eye model and a projection transformation algorithm. However, there still exist some shortages. Our experimental results showed that the following situations may lead to inaccuracy.

(1) If the coverage of drooping eyelid is more than 75%, the algorithm will be inaccurate.
(2) The rotation center of eyeball is varying. Too large pupil eccentric angle will result in inaccurate results.

To overcome the limitations above, future work will be concentrated on improving the accuracy and robustness of the algorithm.

Acknowledgements

This work was supported by the Program for New Century Excellent Talents in University (NCET-06-0298), the Program for Liaoning Excellent Talents in University (No. RC-05-07), the Program for Study of Science of the Educational Department of Liaoning Province (No. 05L020), the Program for Dalian Science and Technology (No. 2005A10GX106), and the Open Fund of Liaoning Key Lab of Intelligent Information Processing, Dalian University (No. 2006-2, 2005-6).

References

[1] Young L, Sheena D. Methods & designs: survey of eye movement recording methods. Behav Res Methods Instrum 1975;7(5):397–429.
[2] van der Geest JN, Frens MA. Recording eye movements with video-oculography and scleral search coils: a direct comparison of two methods. J Neurosci Methods 2002;114(2):185–95.
[3] Kim SC, Nam KC, Lee WS, et al. A new method for accurate and fast measurement of 3D eye movements. Med Eng Phys 2006;28(1):82–9.
[4] Morimoto CH, Mimica MRM. Eye gaze tracking techniques for interactive applications. Comput Vis Image Und 2005;98(1):4–24.
[5] Villanueva A, Cabeza R, Porta S. Eye tracking: pupil orientation geometrical modeling. Image Vision Comput 2006;24(7):633–79.
[6] Fetter M. 3D-analysis of eye movements. Clin Neurophysiol 2000;3(4):247–50.
[7] Schreiber K, Haslwanter T. Improving calibration of 3-D video oculography systems. IEEE Trans Biomed Eng 2004;51(4):676–9.
[8] Hatamian M, Anderson DJ. Design considerations for a real-time ocular counterroll instrument. IEEE Trans Biomed Eng 1983;30:278–88.
[9] Sung K, Anderson DJ. Analysis of two video eye tracking algorithms. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society 1991;13(5):1949–50.
[10] Daniela I, Matteo L. Parametric characterization of the form of the human pupil from blurred noisy images. Comput Methods Programs Biomed 2005;77(1):39–48.
[11] Kim SI, Cho JM, Jung JY, et al. A fast center of pupil detection algorithm for VOG-based eye movement tracking. In: Proceedings of the IEEE Engineering in Medicine and Biology 27th Annual Conference 2005;3188–91.
[12] Zhou ZH, Geng X. Projection functions for eye detection. Pattern Recogn 2004;37(5):1049–56.
[13] Ebisawa Y. Improved video-based eye-gaze detection method. IEEE Trans Instrum Meas 1998;47(4):948–55.
[14] Yoo DH, Chung MJ. RA novel non-intrusive eye gaze estimation using cross-ratio under large head motion. Comput Vision Image Und 2005;98(1):25–51.
[15] Zhu DJ, Steven TM, Theodore R. Robust pupil center detection using a curvature algorithm. Comput Methods Programs Biomed 1999;59(3):145–57.
[16] Moore ST, Haslwanter T, Curthoys IS, et al. A geometric basis for measurement of three-dimensional eye position using image processing. Vision Res 1996;36(3):445–59.
[17] Kim J, Park K. An image processing method for improved pupil size estimation accuracy. In: Proceedings of the 25th Annual Imitational Conference of the IEEE EMBS 2003;9:17–21.
[18] Ge J, Cui H. Ophthalmology. 2nd ed. Beijing: People's Medical Publishing House; 2002, p. 332–3.
[19] Guestrin ED, Eizenman M. General theory of remote gaze estimation using the pupil center and corneal reflections. IEEE Trans Biomed Eng 2006;53(6):1124–33.
[20] Strobl K, Sepp W. Camera Calibration Toolbox for Matlab. http://www.vision.caltech.edu/bouguetj.
[21] Yuan WQ, Xu L, Lin ZH. An accurate and fast iris location method based on the features of human eyes. Fuzzy Syst Knowledge Discov, LNAI 2005;3614:306–15.
[22] Zheng ZL, Yang J, Yang LM. A robust method for eye features extraction on color image. Pattern Recogn Lett 2005;26(14):2252–61.