

视线跟踪过程中变形瞳孔的定位

张文聪 李斌 邓宏平 谭文明 庄镇泉

(中国科学技术大学多媒体计算与通信教育部-微软重点实验室 合肥 230027)

摘要: 在视线跟踪过程中, 变形瞳孔的定位至关重要。针对眼部图像出现干扰情况下的变形瞳孔定位, 该文提出了一种基于3点的随机采样一致性定位算法 RANSAC_3, 即利用随机采样到的2点(以及它们的梯度方向)和搜索获得的1个点来确定椭圆参数。由于该算法在确定参数时只需随机采样两个点, 从而大大降低了采样到干扰点的几率, 利用搜索到的第3点来决定是否对当前点进行参数计算, 有效地解决了标准随机采样一致性定位算法中的无效采样和误差累积问题, 提高了椭圆拟合效率和瞳孔定位精度。实验证明, 该算法对变形瞳孔具有很好的定位效果, 对光斑、睫毛、头发、眼镜框以及眼球运动模糊的引起的干扰具有较强的鲁棒性, 并且定位速度快, 可以达到实时要求。

关键词: 视线跟踪; 变形瞳孔定位; 随机采样一致性; 最小二乘法

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2010)02-0416-06

DOI: 10.3724/SP.J.1146.2008.01461

Distorted Pupil Localization in Eye Tracking

Zhang Wen-cong Li Bin Deng Hong-ping Tan Wen-ming Zhuang Zhen-quan

(Department of Electronic Science and Technology, University of Science and Technology of China, Hefei 230027, China)

Abstract: It is important to locate the center of distorted pupil in eye tracking process. To tackle the issues induced by distortion and disturbance, a new Random Sample Consensus algorithm based on only 3 points, named RANSAC_3, is proposed in this paper. Ellipse model is adopted to match the distorted pupil, its parameters are calculated from the two randomly sampled points and one searched point. Only two randomly sampled points are needed to estimate the ellipse parameters in this method, so the probability of sampling disturbing points is reduced. The point derived by searching can be used to determine whether to calculate the ellipse parameters, so the efficiency of ellipse fitting and the accuracy of localization are improved. Experimental results show that the method has pretty good performance on the localization of distorted pupils, and is effective and robust to various disturbances, i.e. spectrum reflection, eyelash, hair, glasses and motion blur. Moreover, its high speed ensures that it can be used in a real-time eye tracking system.

Key words: Eye tracking; Distorted pupil localization; Random sample consensus; Least Square Method(LSM)

1 引言

在日常生活中, 人们获取到的外界信息有80%~90%是通过眼睛完成的。由于视线具有直接性、自然性和双向性等特点, 使得视线跟踪技术在许多领域有着广泛的应用, 如广告分析研究、动态分析、场景研究和人机交互等^[1-3]。目前常用的视线跟踪技术都是基于视频图像处理的方法, 基本原理是先利用摄像机获取人眼图像, 然后准确地定位和跟踪图像中的眼睛位置, 最后根据平面映射技术估算出用户在屏幕上的注视位置。其核心是如何精确定位出瞳孔中心, 定位的准确程度直接影响到视

线跟踪的精度。

在视线跟踪过程中, 眼球会随着注视点的变化而发生相应的转动。当转动角度较大时, 摄像机拍摄到的瞳孔图像将呈现出椭圆形状, 而不是正常情况下的圆形, 如图1(a)所示。另外, 在拍摄过程中, 眼部图像还容易受到不均光斑、睫毛、头发、眼镜框以及眼球的运动模糊等因素的干扰, 给定位造成一定的影响, 如图1(b)所示。视线跟踪过程中的瞳孔定位也因此转化为干扰情况下的椭圆定位。

目前瞳孔定位的算法很多, 如Daugman的圆周差分算法^[4], 径向对称算法^[5], Hough变换拟合圆的算法^[6], 灰度积分投影算法^[7]等。这些算法大都仅仅针对瞳孔未变形的情况, 一旦瞳孔出现变形就不再适用。灰度积分投影算法虽然不受瞳孔形状的限制, 但其定位精度太低, 且易受到光照的影响,

2008-11-07 收到, 2009-11-23 改回

多媒体计算与通信教育部-微软重点实验室科研基金(07122808)和安徽省科技攻关项目(07010202046)资助课题

通信作者: 张文聪 zwcg@mail.ustc.edu.cn

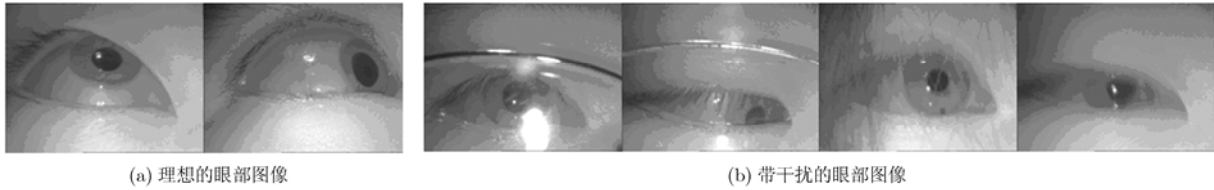


图1 部分典型的眼部图像

因此不能满足视线跟踪的精度要求。针对变形瞳孔的定位, Takeshi Takegami等人提出了利用Haugh变换拟合椭圆的算法来定位瞳孔^[8],但由于椭圆的参数有5个,参数空间过于庞大,因此拟合效率很低,满足不了实时要求。另外, Lin等人提出基于最小二乘法的椭圆拟合方法^[9],但是由于瞳孔在边缘检测过程中容易受到干扰点的影响,拟合出来的椭圆与真实结果相比往往具有较大偏差。

近年来, Li 等人提出了基于5点的随机采样一致性(RANSAC_5)思想的瞳孔定位算法^[10]。该算法首先通过在瞳孔边缘上随机采样5个点计算出相应的椭圆参数,然后计算出此参数空间对应的一致集。如果一致集的规模小于某个阈值,则重新采样5个点,继续上述过程直到找到满足条件的一致集为止,最后通过去一致性完成最终的椭圆定位。虽然该算法大大提高了椭圆的拟合速度,但由于每次仍需要采样5个点,导致在有较大干扰的图像中容易引入无效采样,使得计算量和内存需求增大,而且还可能导致算法在可接受的时间或内存占有情况下定位失败。Colombo针对标准随机采样一致性定位算法的不足^[11],引入了椭圆约束条件,使得无用的随机采样大为降低,提高了椭圆拟合的效率和精度。

针对当前算法的不足,本文引入瞳孔边缘像素点的灰度梯度方向特性,提出了基于3点的随机采样一致性瞳孔定位算法RANSAC_3,利用随机采样到的2点(包括它们的梯度方向)和搜索获得的第3点来确定椭圆参数,最后利用最小二乘法进行去一致集处理。该算法大大降低了采样到干扰点的几率,有效地解决了无效采样和误差累积问题,从而提高了效率和精度。

在本文以下各部分中,第2节简要介绍RANSAC_5算法,第3节详细讨论RANSAC_3算法,最后是实验结果及讨论。

2 基于 RANSAC_5 的变形瞳孔定位算法

2次曲线的一般方程可以表示为

$$x^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0 \quad (1)$$

其中 (x, y) 为图像空间坐标, B 、 C 、 D 、 E 、 F 是二次曲线的参数,若满足

$$B^2 - C < 0 \quad (2)$$

则2次曲线为椭圆。

RANSAC^[10]的基本思想是:通过迭代地随机抽取最小点集来找出能够使得内点(非干扰点)相对于总样本点所占比例最高的最小点集,然后利用这些内点估计出样本点对应的最佳模型。将RANSAC直接用于瞳孔定位时,需每次从边缘点集合中随机采样5个点,代入式(1)可得5个方程,构成关于椭圆参数的线性方程组;求解该方程组,若解满足式(2),则得到一组椭圆参数,否则重新随机采样;然后在一定误差容许范围内,统计边缘点集合中满足该组椭圆参数的边缘点个数,作为此参数对应的一致集;判断一致集是否大于指定阈值,若满足条件,则此时一致集中的点称为瞳孔的内点,其余点称为野值(干扰点),否则,继续重复上述过程,直到找到满足条件的一致集;最后根据求出的内点估计瞳孔椭圆的参数。

相比Hough变换,RANSAC_5在运算速度上有了很大的提高。

3 基于 RANSAC_3 的变形瞳孔定位算法

由于瞳孔的边缘提取容易受到不均匀光斑、眼睑和睫毛等的影响,提取出的边缘点集合中将包含大量的干扰点。在RANSAC_5的随机采样过程中,即使每次仅仅引入了一个干扰点,也会造成此次定位的失败。当干扰点的比例较大时,就会造成大量的无效采样从而使得算法的计算量和内存需求增大,还可能导致定位的失败。

假设图像经过边缘提取后得到的边缘点集 S 的大小为 N ,其中存在 M 个干扰点,又设每次随机采样的点数为 n ,那么 n 个随机采样点同时落在瞳孔真正边缘上的概率为

$$\text{Prob} = C_{N-M}^n / C_N^n = ((N-M) \cdot (N-M-1) \cdots$$

$$(N-M-n+1)) / (N \cdot (N-1) \cdots (N-n+1)) \quad (3)$$

从式(3)可以看出,要提高随机采样点同时落在瞳孔边缘上的概率有两种方法:第一种就是减少干扰点 M 的数目,提高边缘检测的效果;另一种方法是减少每次随机采样点的个数。本文从这两个方面同时着手,减少无效采样次数,提高瞳孔的定位效



图 2 采用射线发散法提取瞳孔边缘的结果

率。在瞳孔边缘提取方面,采用射线发散法,而边缘的拟合采用基于 RANSAC_3 的瞳孔定位法。

3.1 瞳孔边缘的提取

传统的边缘提取方法通常采用 sobel 算子或 canny 算子实现^[6,8,9,12]。但由于不均匀光斑、眼睑和睫毛等因素的影响,提取出的边缘点集中包含大量的干扰点,如图 2(b)所示。

为提高边缘检测的质量,本文利用射线发散法实现瞳孔边缘提取。该算法的主要思想是由瞳孔的大致中心向四周发出射线,当射线遇到灰度变化较大的点时,就停止前进,并将此时的位置作为瞳孔的边缘点。其中瞳孔中心的粗定位采用径向对称算法实现^[5]。边缘的提取结果如图 2(c)所示。

3.2 瞳孔边缘的精确定位

采样过程中,采样点数目越少,所有采样点同时落在椭圆边缘上的概率就越大。本文利用 2 次曲线的几何性质,从曲线的极和极线理论出发,同时引入边缘点的梯度信息来减少每次随机采样的数目。

2 次曲线的极和极线定义^[13,14]表明:一条 2 次曲线可由 3 个位于曲线上的点来确定,其中 2 点需要知道切线方向,如图 3 所示。2 次曲线上的 2 点 P_1, P_2 处的切线为 l_1, l_2 ,若二者不平行,则它们的交点 T 称为 2 次曲线的极,线段 $\overline{P_1P_2}$ 所在的直线 l_3 称为 2 次曲线的极线。设直线 l_1, l_2 和 l_3 的参数方程为

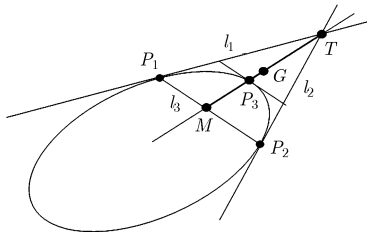


图 3 椭圆的极和极线的定义

$$\left. \begin{aligned} l_1(x, y) &= a_1x + b_1y + c_1 = 0 \\ l_2(x, y) &= a_2x + b_2y + c_2 = 0 \\ l_3(x, y) &= a_3x + b_3y + c_3 = 0 \end{aligned} \right\} \quad (4)$$

其中 a, b 和 c 为直线的参数,则过 p_1, p_2 两点的 2 次曲线簇方程为

$$\lambda \cdot l_3^2(x, y) + l_1(x, y) \cdot l_2(x, y) = 0 \quad (5)$$

其中 λ 是任意常数, λ 的确定还需在边缘样本库中采样第 3 点。图中线段 \overline{TM} (M 是线段 $\overline{P_1P_2}$ 的中点)与 2 次曲线的交点 P_3 控制着曲线的形状。根据椭圆的形状控制点约束性质^[13],只有当 P_3 靠近点 M 时,曲线为椭圆。而且,交点 P_3 处的切线与直线 l_3 平行。

利用上述椭圆的 2 次曲线性质,本文提出的 RANSAC_3 算法首先随机从边缘样本集中采样两点,作为 P_1, P_2 ,如果 P_1 和 P_2 处的切线 l_1, l_2 不平行,则可求出 T 点。然后在线段 \overline{MG} (G 是线段 \overline{TM} 的中点)上随机采样第 3 点作为 P_3 ,若 P_3 处的切线与直线 l_3 近似平行,则依据此 3 点求出瞳孔椭圆的参数方程。其中, P_1, P_2 点的切线方向求取采用这两点的灰度梯度方向负倒数来近似得到。假设 k_i 为 P_i 点切线 l_i 的斜率, g_i 为 P_i 点的梯度方向,则有

$$k_i = -1/g_i \quad (6)$$

RANSAC_3 算法的具体步骤:

(1)从集合 S 中随机采样 2 点为 P_1, P_2 , 计算这 2 点所对应的切线方程 l_1, l_2 , 若 P_1 和 P_2 处的切线 l_1, l_2 不平行,则可求出 T 点和极线方程 l_3 ; 否则重新采样;

(2)计算线段 $\overline{P_1P_2}$ 的中点坐标 M , 然后计算出线段 \overline{TM} 的中点 G ;

(3)在线段 \overline{MG} 附近随机采样一个样本点,若存在一点 P_3 ,且该点处的切线方向与直线 l_3 近似平行,便可根据式(5)得到椭圆的参数 PO ; 若线段 \overline{MG} 附近不存在满足此条件的点,则返回第(1)步重新采样

P_1, P_2 两点;

(4) 根据得到的椭圆参数, 计算出椭圆的焦点坐标与长轴的大小;

(5) 计算 S 中所有点到两个焦点的距离之和, 若在容许误差范围内, 则对满足条件的点进行统计, 求得参数 PO 所对应的一致集 CO ;

(6) 如果 PO 的一致集 CO 大于指定域值 T_0 时, 则根据求出的内点, 利用最小二乘法重新估计瞳孔椭圆的参数; 若不满足条件, 继续重复上述步骤, 直到找到满足条件的一致集;

(7) 如果上述迭代次数大于一定的值 N_{times} , 则从 N_{times} 次迭代中挑选出最大的一致集 CO , 并且根据 CO , 利用最小二乘法重新估计瞳孔椭圆的参数。

由于去除了大部分干扰点, 利用最小二乘法^[9]对瞳孔边缘点的重新拟合就不会因为干扰点的存在而出现较大的偏差, 因此能够取得较好的定位精度, 如图 4 所示。

4 实验结果与分析

由于目前尚没有能够测试变形瞳孔定位算法的公开数据库, 本文实验中所用的数据库是中国科学技术大学智能信息处理实验室^[15]拍摄的眼部图像数据库—USTCEyeImg 的两个子库, 即理想的眼部图像库 Ea 和带干扰眼部图像库 Ec。其中每个子库均包含 15 个人共 750 幅眼部图像, 图像尺寸为 756×576 。干扰条件包括光照变化、不均匀光斑、睫毛、

眼镜框以及眼球的运动模糊等因素的影响, 如图 1 所示。

实验结果主要从瞳孔定位的正确率、算法的平均迭代次数以及定位时间 3 个指标进行衡量。另外, 选择最小二乘拟合法和 RANSAC_5(文献[9,10])进行实验对比(边缘提取都采用射线发散法)。实验编程环境为 VC6.0, 计算机的配置为 P4 3.0G, 512 M 内存。实验结果如表 1 所示。

从表 1 中可以看出, 在同等条件下, 本文算法的定位正确率是最高的。尤其在眼部图像受到干扰较大的情况下, 文献[9]和文献[10]的方法都有明显下降时, 本文的方法仍能保持 98.5%的正确率。这充分说明了本文算法的鲁棒性。在干扰情况下, 文献[9]的方法性能下降最快, 这是因为该方法使用最小二乘法进行椭圆拟合, 因而对干扰点的影响很敏感, 拟合出来的椭圆与真实结果往往偏差较大。文献[10]的方法效果略有改善, 这是因为: 采用随机采样一致性原理能将具有明显偏差的干扰点去除, 因而受干扰点的影响降低了。本文算法在文献[10]的基础上进行改进, 利用了椭圆的极和极线理论使得用于计算椭圆参数的点仅需 3 个, 大大降低了采样到干扰点的概率。因此, 在眼部图像出现复杂的干扰情况, 如光斑、头发、睫毛、镜框和运动模糊时, 本文的算法仍具有较好的定位效果。3 种方法的定位结果比较如图 5 所示。

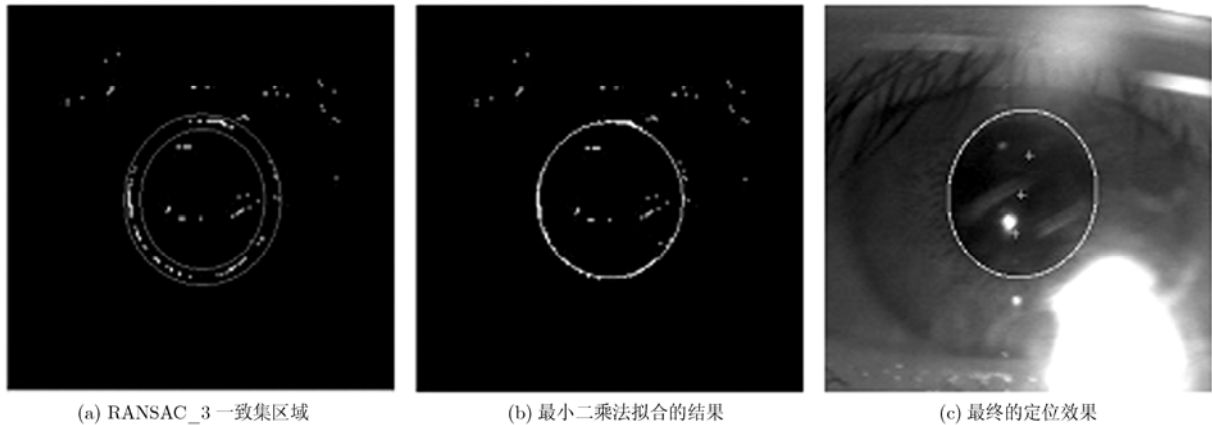


图 4 RANSAC_3 算法定位瞳孔中心

表 1 实验结果

方法	理想的眼部图像 Ea			带干扰的眼部图像 Ec		
	定位正确率(%)	平均迭代次数	平均定位时间(s)	定位正确率(%)	平均迭代次数	平均定位时间(s)
文献[9]	95.3		0.07	84.6		0.07
文献[10]	98.1	49	0.16	93.7	97	0.25
本文算法	99.6	15	0.08	98.5	21	0.08

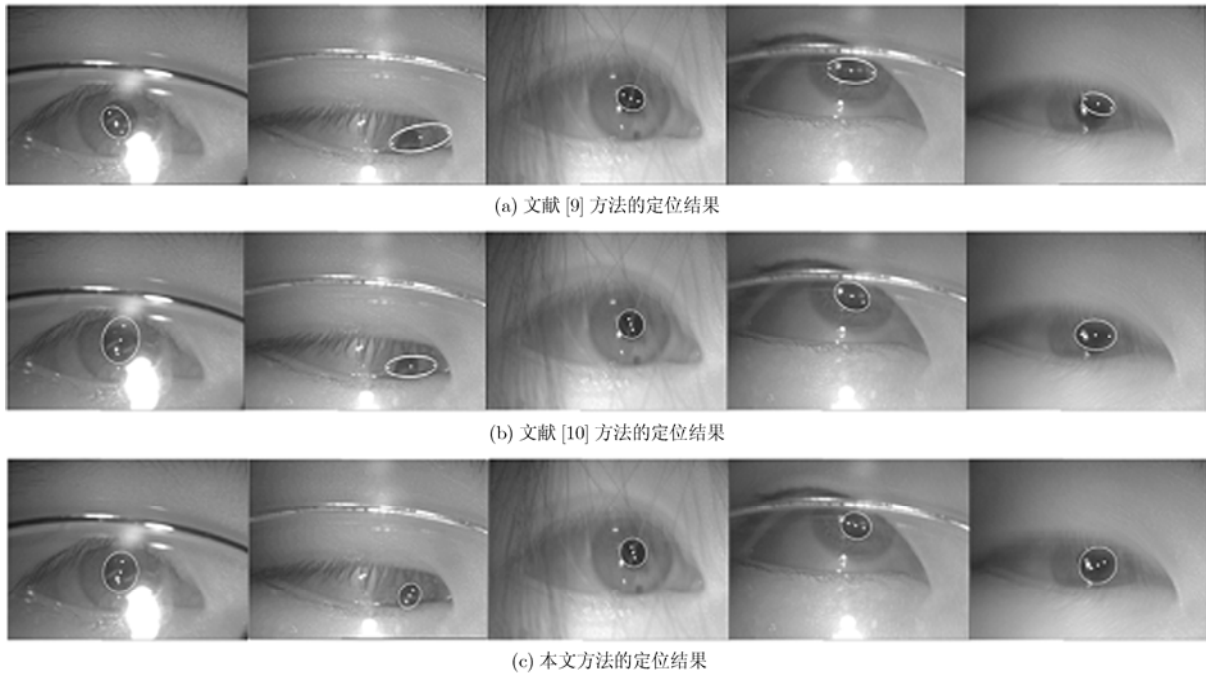


图 5 三种不同方法对于各种干扰情况下的眼部图像的定位结果比较

本文算法的另一个优点是，在保证高准确率的同时，还保证了高速度。如表 1 所示，本文算法的平均定位时间与文献[9]相差无几，但与文献[10]相比较，则占有明显的优势。算法速度的提高得益于迭代次数的明显降低，如表 1 所示。由于文献[10]每次需要采样 5 个点，因干扰点造成的无效采样的几率大大增加，从而增加了迭代次数。如表 1 所示，在理想情况下，文献[10]的平均迭代次数为 53 次，而在干扰情况下，平均迭代次数急剧上升至 97 次，这是造成该算法低效的主要原因。本文算法对文献[10]进行了针对性地改进，减少了无效采样的次数，使每次采样都有利于定位结果的收敛，从而使迭代次数明显降低。即便在干扰比较严重的情况下，迭代次数也不会明显增加。这是保证本文算法高效性的主要原因。

5 结束语

针对标准随机采样一致性(RANSAC_5)算法在瞳孔定位过程中存在的大量无效采样问题，本文根据采样概率对其进行分析，并从两个方面进行了改进，提出了基于 3 点的随机采样一致性定位算法 RANSAC_3。该算法首先在瞳孔边缘检测过程中利用射线发散法，减少边缘点集中干扰点的数目；然后在边缘拟合过程中每次仅需采样 3 个点，从而大大降低了采样到干扰点的几率，并且在采样过程中，利用像素的梯度方向特性对采样点进行筛选；最后利用最小二乘法进行去一致集处理，从而进一步提

高了瞳孔定位的精度。实验效果表明，本文的算法不仅鲁棒性强，而且速度快，完全能够达到实时要求。

参考文献

- [1] Jacob R J K and Karn K S. Eye tracking in human computer interaction and usability research: Ready to deliver the promises [J]. *Computer Vision and Image Understanding*, 2003, 24: 682-701.
- [2] Majaranta P and Raiha K. Twenty years of eye typing: systems and design issues [C]. Proceedings of the 2005 symposium on eye tracking research and applications, New Orleans, Louisiana, USA, March 2002: 15-22.
- [3] Kumar M, Garfinkel T, and Boneh D, *et al.* Reducing shoulder-surfing by using gaze-based password entry [C]. Proceedings of the 3rd Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania, 2007(7): 13-19.
- [4] Daugman J. How iris recognition works [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004, 14(1): 21-30.
- [5] Zhang Wencong, Li Bin Li, and Ye Xueyi, *et al.* A Robust Algorithm for Iris Localization Based on Radial Symmetry [C]. Proceedings of the 2007 International Conference on Computational Intelligence and Security, Haerbin, 2007: 324-327.
- [6] Ma Li and Tan Tieniu, *et al.* Personal identification based on iris texture analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(12): 1519-1533.
- [7] 耿新, 周志华, 陈世福. 基于混合投影函数的眼睛定位[J]. 软

- 件学报, 2003, 14(8): 1394-1400.
- Geng Xin, Zhou Zhi-hua, and Chen Shi-fu. Eye location based on hybrid projection function [J]. *Journal of Software*, 2003, 14(8): 1394-1400.
- [8] Takegami T, Gotoh T, and Ohyama G. An algorithm for model-based stable pupil detection for eye tracking system [J]. *Systems and Computers in Japan*, 2004, 35(13): 4, 21-31.
- [9] Lin Xiang, Craig J, and Dean S, *et al.* Accurately measuring the size of the pupil of the eye [R]. Communication and information technology research technical report, 2003, 127.
- [10] Li Dongheng, David W, and Derrick J P. Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches [C]. Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, June 2005: 79-87.
- [11] Colombo C, Comanducci D, and Bimbo A D. Robust tracking and remapping of eye appearance with passive computer vision [J]. *ACM Transactions on Multimedia Computing, Communications and Applications*, 2007, 3(4): 1-20.
- [12] 闫蓓, 王斌, 李媛. 基于最小二乘法的椭圆拟合改进算法. 北京航空航天大学学报, 2008, 34(3): 295-298.
- Yan Bei, Wang Bin, and Li Yuan. Eye location based on hybrid projection function [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2008, 34(3): 295-298.
- [13] Yoo J H and Sethi I K. An ellipse detection method from the polar and pole definition of conics [J]. *Pattern Recognition Letters*, 1993, 26(2): 307-315.
- [14] 陈燕新, 戚飞虎. 一种新的基于随机Hough变换的椭圆检测方法. 红外与毫米波学报, 2000, 19(1): 43-47.
- Chen Yan-xin and Qi Fei-hu. A new ellipse detection method using randomized hough transform. *Journal of Infrared Millimeter and Waves*, 2000, 19(1): 43-47.
- [15] 中国科学技术大学电子科学与技术系智能信息处理研究所 http://est.ustc.edu.cn/lab_ai.html.
- 张文聪: 男, 1982年生, 博士生, 研究方向为图像处理、模式识别和视线跟踪.
- 李斌: 男, 1970年生, 副教授, 研究方向为智能信息处理、生物特征识别和进化计算.
- 邓宏平: 男, 1982年生, 博士生, 研究方向为图像处理、生物特征识别.
- 庄镇泉: 男, 1938年生, 教授, 博士生导师, 研究方向为智能信息处理、模式识别和多媒体技术.