

视线追踪系统中特征参数提取方法研究

张 闯, 迟健男, 邱亚飞, 张朝晖

¹⁾(北京科技大学信息工程学院, 北京 100083) ²⁾(北京科技大学钢铁流程先进控制教育部重点实验室, 北京 100083)

摘 要: 为了提取人眼的高精度亚像素特征参数, 利用亮瞳现象, 提出了一种基于多通道图像的高精度亚像素特征参数提取方法。该方法首先通过差分图像滤波获得瞳孔区域, 进而检测瞳孔区域的边缘, 并在眼睛区域附近基于灰度, 搜索角膜反射区域; 然后求取其质心用于定位角膜反射区域中心, 并对瞳孔边缘做滤波, 以消除角膜反射对瞳孔边缘轮廓的影响, 进而利用椭圆拟合来定位瞳孔中心; 最后提取包括人眼特征和人脸位置的多个参数, 另外, 还建立了一个多特征参数提取的流程, 为下一步的视线估计提供了参数依据。实验结果及视线追踪系统最终的视线估计结果证明, 该方法是有用的。

关键词: 视线追踪; 特征参数提取; 瞳孔定位; 视线估计; 瞳孔-角膜反射(PCCR)技术

中图分类号: TP391 **文献标志码:** A **文章编号:** 1006-8961(2010)09-1299-10

The research on feature extraction for gaze tracking system

ZHANG Chuang, CHI Jiannan, QIU Yafei, ZHANG Zhaohui

¹⁾(School of Information Engineering, University of Science and Technology Beijing, Beijing 100083)

²⁾(Key Laboratory for Advanced Control of Iron and Steel Process (Ministry of Education), University of Science and Technology Beijing, Beijing 100083)

Abstract: In order to extract the high-precision sub-pixel feature parameters, this paper presents a high-precision sub-pixel feature extraction method based on the multichannel image by use of the phenomenon of the bright pupil. Firstly, the pupil area is obtained by filter on a difference image. Then the edge of the pupil area is detected and cornea reflection area is searched nearby the eye area based on gray level. Then corneal reflex center is located by centroid and the edge of the pupil is filtered to eliminate the impact of corneal reflex to the contour of the pupil. Ellipse fitting is used to locate the pupil center. Finally, multiple parameters of facial features and face position are extracted. We establish a multi-feature extraction process, which provides parameter basis for gaze estimation which is in the next step. The results of feature extraction and the ultimate gaze estimation results made by gaze tracking system verify the effectiveness of the method.

Keywords: gaze tracking; feature extraction; pupil location; gaze estimation; pupil center cornea reflection (PCCR) technique

0 引 言

视线追踪是利用机械、电子、光学等各种检测手段来获取受试者当前“注视方向”的技术。在人机交互和疾病诊断^[1-2]两个领域有着广泛的应用, 如

助残、虚拟现实、认知障碍诊断、车辆辅助驾驶、人因分析等。用于诊断的视线追踪系统可以采用侵入式视线跟踪技术, 以获得更高的精度。用于交互的视线追踪系统除了对精度、鲁棒性、实时性有要求以外, 还需要最大程度地减少或消除对使用者的干扰。随着数字化技术、计算机视觉、人工智能技术的迅速

基金项目: 国家高技术研究发展计划(863)项目(2007AA01Z160); 国家自然科学基金项目(60775042); 北京市教育委员会重点学科共建项目(XK100080537)。

收稿日期: 2010-03-22; **改回日期:** 2010-04-20

第一作者简介: 张 闯(1982—), 男, 北京科技大学控制科学与工程专业博士研究生。主要研究方向为图像处理与分析、视线追踪、人工智能、模式识别等。E-mail: zhangchuangtc@163.com。

发展,基于数字视频分析(VOG)的非侵入式视线追踪技术成为当前热点研究方向。

VOG 系统普遍使用瞳孔-角膜反射方法(PCCR)和采用红外光源产生角膜反射,并通过计算瞳孔中心与角膜反射之间的向量来估计视线方向^[3-6]。视线估计模型可以采用非线性多项式模型^[7],也可以采用几何模型^[8-10]。

人眼特征提取是视线追踪的必要过程和前提,且特征提取的质量直接关系到视线估计的精度。可将用瞳孔-角膜反射方法提取的瞳孔中心到角膜反射的向量作为视线估计的依据。为了提高系统的精确性与稳定性,瞳孔-角膜反射方法一般采用亮暗瞳差分方案,以便于在整幅脸部图像中快速捕捉眼睛和在局部眼睛图像中精细准确的分割瞳孔^[3-6]。亮暗瞳差分方案一般过程如下:先设置与摄像机镜头同轴的内外环近红外(IR)光源分别产生“亮瞳”“暗瞳”现象(如图 1 所示);然后控制光源交替亮暗,产生亮瞳和暗瞳隔帧交替出现的视频序列;最后利用相邻亮暗瞳图像的差分使瞳孔信息相对突出,以消除背景影响。亮暗瞳差分方案一般采用阈值法分割瞳孔^[11-13]。为了提高检测的稳定性,文献[14]通过设置动态阈值来消除不同光照条件下的背景影响。文献[15]结合眼睛外观,进行眼睛检测。文献[16]对差分图像进行形状滤波,以消除外部光线干扰。文献[17-18]在阈值化后的差分图像中,通过支持向量机对瞳孔候选区域进行分类,以排除由头部随机运动产生的干扰区域。但由于这种方法的识别效果过于依赖样本,致使对新的使用者的识别效果明显下降。

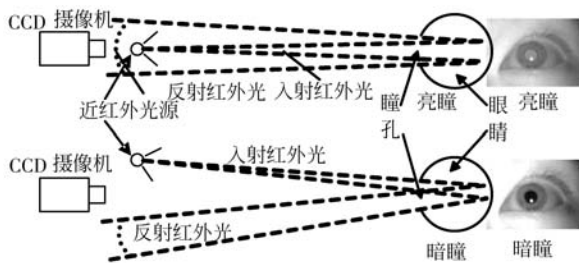


图 1 亮瞳和暗瞳

Fig. 1 The bright pupil and dark pupil

瞳孔及瞳孔中心与角膜反射定位是该类视线特征提取的共性问题,特征参数提取需满足以下几点

要求:1) 眼睛检测与跟踪的特征参数提取过程是自动的,不需要使用者干预;2) 整个特征参数提取过程要快速准确,以满足视线估计对参数的精度和实时性要求;3) 能够提取用于视线估计的所有特征参数,包括估计头部位置所需要的参数;4) 特征参数提取过程能适应个体差异,且不会因为眼睛大小、形态的个体差异而影响算法的性能。

上述特征提取过程中,影响最终视线估计精度的原因主要有以下几点:1) 因角膜反射造成瞳孔轮廓改变而产生瞳孔定位误差;2) 因瞳孔运动造成的差分图像瞳孔区域不全而产生运动瞳孔定位误差;3) 因角膜反射区域形状的不规则和边界的不确定性而产生的角膜反射中心定位误差;4) 由于头部的随机运动,致使亮暗瞳图像差分后,图像还残留其他面部信息,这是造成阈值分割法瞳孔定位错误的主要原因;5) 作为头动补偿的其他特征参数产生的误差。

本文的视线追踪系统是采用单摄像机单光源,通过瞳孔-角膜反射方法获取视线方向,立足于开发一种具有普适应用价值的低成本视线追踪系统。瞳孔及角膜反射中心定位是本文视线追踪系统实现的必要步骤,为满足上述要求,并针对影响特征提取精度的主要原因,本文提出了一种基于多通道图像的特征参数提取方法,并在追求高精度亚像素特征参数提取及鲁棒性方面做了如下研究:1) 角膜反射是造成瞳孔轮廓改变的主要因素,是影响瞳孔中心定位精度的主要因素之一,因此针对角膜反射对瞳孔边缘的影响,提出了一种基于径向距离的瞳孔边缘专用滤波算法——RDPEF 算法,提高了瞳孔定位的精度;2) 瞳孔运动造成差分图像瞳孔区域不全是造成运动瞳孔定位误差的主要因素,为此通过迭代扩展的方式还原运动瞳孔区域,提高了运动瞳孔定位的精度;3) 角膜反射区域形状的不规则和边界的不确定性是造成角膜反射区域定位误差的主要因素,为此通过多阈值分割角膜反射的模糊区域,从而减小了模糊的边界区域对质心定位结果的影响;4) 头部的随机运动是造成阈值分割法产生瞳孔定位错误的主要原因,为此利用瞳孔的历史信息,即在 $t+1$ 时刻首先在 t 时刻瞳孔位置的附近进行特征参数提取,从而提高了瞳孔定位鲁棒性;5) 建立了一个多特征参数提取的流程,用来提取包括人眼特征和人脸位置的多个参数,从而为下一步的视线估计提供了参数依据。本文提出的检测方法优于通常

采用的阈值分割方法,实验结果及系统实际应用证明,本文方法是有效的。

1 视线追踪系统

本文硬件系统由光源、滤光片、镜头、采集卡、CCD、GPIO 卡、单片机、主机和屏幕组成。

工作时,使用者注视屏幕,由 CCD 摄像机获取人脸图像,通过图像采集卡传到主机,主机通过特征参数提取和视线映射函数来得到视线落点,并显示在屏幕上。

另一方面,为了得到方便处理的人脸图像,先采用 GPIO 卡 (general purpose input/output) 获取 CCD 视频图像的帧同步信号,再通过单片机控制光源的开关,然后通过内外环光源的交替开关产生交替的亮瞳和暗瞳图像;最后采用主动近红外光源进行照明,并且加了对应波长的滤光片,使拍摄的人脸图像受外部光照影响很小,图像灰度稳定。本文采用的近红外发光二极管的功率在安全阈值以下,对人体无害,而且对于使用者是不可见的。亮瞳和暗瞳的产生原理见图 1,本文硬件系统结构如图 2 所示。

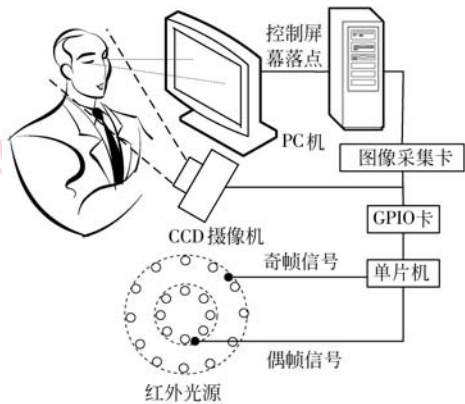


图 2 本文硬件系统结构

Fig. 2 The structure of the hardware system

2 特征参数提取

本文提出了一种满足视线估计要求的特征参数提取方法,该方法流程如图 3 所示。

2.1 定位瞳孔区域

定位瞳孔的原理为:亮瞳(图 4(a))与暗瞳图像(图 4(b))相减得到差分图像(图 4(c)),然后对差分图像做滤波,得到瞳孔区域。

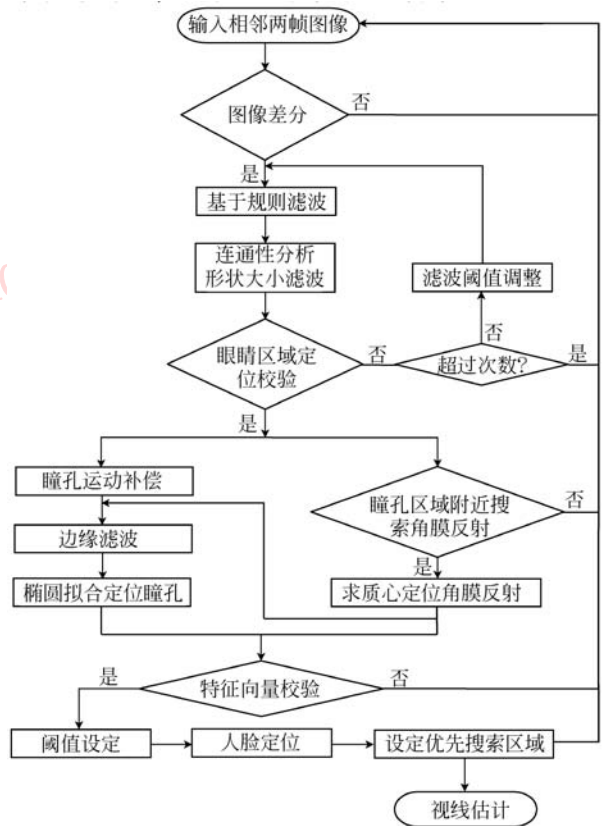


图 3 特征参数提取框架

Fig. 3 Framework of feature extraction



(a) 亮瞳图像 (b) 暗瞳图像



(c) 差分图像

图 4 亮瞳、暗瞳和差分图像

Fig. 4 Bright pupil image, dark pupil image and differential image

控制光源,使交替亮暗,产生亮瞳和暗瞳隔帧交替出现的视频序列。对相邻两帧亮暗瞳图像进行差分即可得到差分图像。为了在差分图像中找到瞳孔区域,本文经过大量实验得到以下规律:1) 暗瞳区域

的灰度值个体差异很小,是一个比较稳定的瞳孔特征;2) 尽管亮瞳区域的灰度值个体差异较大,但其与暗瞳区域的灰度差值在光源固定的前提下,不会小于一个固定阈值;3) 在光源固定的前提下,尽管角膜反射区域的大小个体差异较大,但角膜反射区域的灰度值个体差异很小;4) 在室内使用的前提下,室内照明亮度的变化对亮瞳区域的灰度值和脸部其他区域的灰度值有影响,但对暗瞳区域的灰度值和角膜反射区域的灰度值几乎没有影响。本文对 50 位使用者进行了特征检测,其亮瞳、暗瞳和角膜反射的灰度值如图 5 所示。

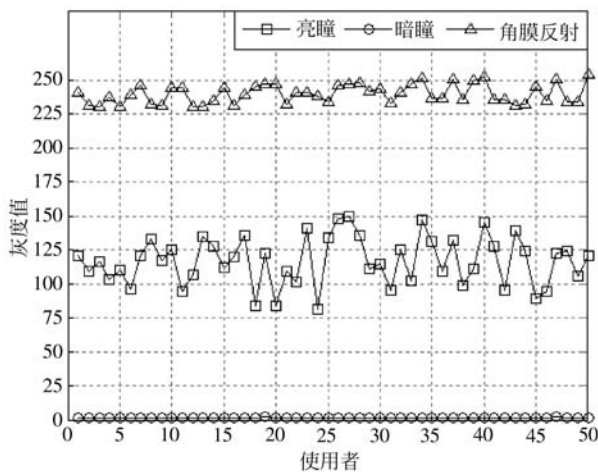


图 5 亮瞳、暗瞳和角膜反射灰度值的个体差异

Fig. 5 Gray value differences of bright pupil, dark pupil and cornea reflection between different individuals

可以根据亮瞳图像、暗瞳图像、差分图像 3 幅图像中瞳孔和角膜反射特征的先验知识对差分图像做滤波,以得到瞳孔区域。根据上述的实验规律,可以发现瞳孔区域主要有以下 3 个特征:1) 暗瞳很暗,像素的灰度接近于 0,暗瞳现象是由于瞳孔与光源的特定位置造成的,其他脸部区域在红外光源照射下都没有这么低的灰度值,即使最暗的眉毛和鼻孔,由于对红外光线有漫反射,所以其灰度通常不为 0;2) 亮暗瞳区域具有明显的较大的灰度差,这是其他区域所不具备的特征,也是亮暗瞳差分方案定位瞳孔的主要依据;3) 瞳孔区域附近必然有角膜反射,这也是基于红外光源帮助定位瞳孔的原因之一,因此可以依据上述瞳孔区域特征来进行瞳孔区域定位,即在可能的角膜反射区域附近搜索暗瞳很暗并且亮暗瞳灰度差较大的瞳孔区域。由于角膜反射区域的特征是“突变的亮点”,即呈现亮度很高、梯度

很大的区域,所以在暗瞳图像中,从灰度和梯度两个方面即可以判断出可能的角膜反射区域。

在初步判定瞳孔区域后,还需要精确定位瞳孔。由于瞳孔区域内亮瞳的灰度值不是均匀一致的(使用者 1 的实验数据如图 6 所示),所以使用单一的阈值不能有效识别整个瞳孔区域,本文采用双阈值加区域连通的两步方法来进行瞳孔区域判断。

定位时,首先识别具有较强亮暗瞳特征的像素点,差分图像中,同时满足暗瞳像素灰度 $p_{i,j}^d$ 小于暗瞳灰度阈值 T_1 和差分图像像素灰度 $(p_{i,j}^b - p_{i,j}^d)$ 大于差分图像瞳孔灰度严阈值 T_2 两个条件的像素 (i,j) 的集合即构成差分图像中判定的瞳孔区域 R_1 。差分图像中,同时满足暗瞳像素灰度 $p_{i,j}^d$ 小于暗瞳灰度阈值 T_1 和差分图像像素灰度 $(p_{i,j}^b - p_{i,j}^d)$ 大于差分图像瞳孔灰度宽阈值 T_3 两个条件的像素 (i,j) 的集合即构成差分图像中可能的瞳孔区域 R_2 。容易知道, $R_1 \subseteq R_2$ 。

接下来,在 R_2 中搜索与 R_1 连通的像素,对于 R_2 中的像素 (i,j) ,如果存在 R_1 中的像素 (p,q) 与其相邻,则将这些像素与 R_1 一起加入规则滤波后的差分图像瞳孔区域 R_3 ,即将那些与强亮暗瞳特征点连通的较弱特征点也作为瞳孔区域像素,将迭代搜索到的 R_3 附近属于 R_2 的像素加入 R_3 ,直至 R_3 收敛,即

$$R_3 = R_1$$

搜索

$$R_3 = R_3 \cup \left\{ (i,j) \mid (i,j) \in R_2 \wedge (\exists (p,q)) \left((p,q) \in R_3 \wedge \|(i,j) - (p,q)\| \leq 1 \right) \right\}$$

直到 R_3 收敛

在以上基于规则的滤波基础上,对差分图像做进一步几何滤波来得到瞳孔区域。文献[17-18]的连通区域分析只是先分块,然后用支持向量机(SVM)来做瞳孔识别。本文经过上面的基于规则的滤波已经过滤了多数的干扰点,在连通性分析以后,再进行几何滤波就可以达到定位瞳孔的要求,从而避免了 SVM 的训练过程和识别误差。最后将面积、宽度、高度、宽高比满足条件的团块判断为瞳孔。 B (图 7)为滤波后的瞳孔区域。

搜索眼睛区域后,再校验眼睛区域的合理性。如果团块集合 B 中的团块数量 $|B|$ 等于 2,也就是正好得到左右 2 个瞳孔区域,即数量满足条件,则认为定位正确。如果不满足条件,则调整差分图像瞳

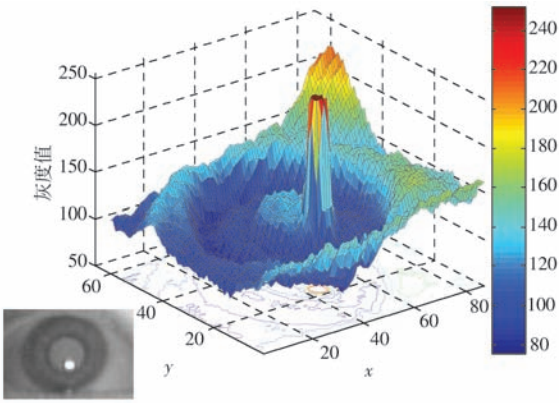


图6 瞳孔区域内亮瞳的灰度值分布

Fig.6 The gray value distribution in bright pupil region



图7 滤波后的瞳孔区域

Fig.7 The pupil region after filtering

孔灰度严阈值 T_2 和差分图像瞳孔灰度宽阈值 T_3 再次进行规则滤波和几何滤波,当得到的瞳孔区域多于两个,即当 $|B|$ 大于2时,则说明对于这个使用者来说,差分图像瞳孔阈值设置过低,“差出了”一些伪瞳孔区域,此时需提高差分图像瞳孔阈值 T_2, T_3 再次定位,相反,当得到的瞳孔区域少于两个,即当 $|B|$ 小于2时,则说明对于这个使用者来说,差分图像瞳孔阈值设置过高,“差掉了”真正的瞳孔区域,此时需降低差分图像瞳孔阈值 T_2, T_3 再次定位,循环这个过程就可以搜索到最符合瞳孔特点的两个区域作为左右瞳孔区域。当重复滤波次数大于3时,则返回,不进行下面的处理。

如果数量条件满足,则继续判断左右瞳孔距离 $\|b_L - b_R\|$ 是否在瞳孔距离的下阈值与瞳孔距离的上阈值之间,若不满足,则返回,不进行下面的处理。

在确定瞳孔区域后,还需要消除瞳孔运动对差分图像瞳孔区域的影响,瞳孔运动会造成亮暗瞳图像瞳孔区域不完全重合,以至差分图像瞳孔区域不完整是造成运动瞳孔定位误差的主要因素,本步骤通过迭代扩展的方式还原运动瞳孔区

域,规则如下:

$$R_4 = B$$

搜索

$$R_4 = R_4 \cup \{ (i,j) \mid (\exists (p,q)) ((p,q) \in R_4 \wedge p_{i,j}^d \leq T_1 \wedge \| (i,j) - (p,q) \| = 1) \}$$

直到 R_4 收敛

对得到的瞳孔区域 B 进行扩展,在与 B 边界相邻的像素中,迭代搜索暗瞳像素灰度 $p_{i,j}^d$ 小于暗瞳灰度阈值 T_1 的像素,并将其加入扩展后的瞳孔区域 R_4 ,直至 R_4 收敛。至此,即得到精确的瞳孔区域。

2.2 定位角膜反射区域及瞳孔中心

在暗瞳图像中,在瞳孔区域 R_4 附近搜索角膜反射,在 R_4 附近小于距离阈值的像素集合中,搜索到的满足灰度 $p_{i,j}^d$ 大于角膜反射灰度阈值条件的像素作为初步判定的角膜反射区域 R_5 的像素。由于外部光源有时会在瞳孔附近产生其他一些亮斑,因此还需要对 R_5 进行连通性分析,以排除干扰,将找到的角膜反射特征(面积、宽高比)满足条件的团块判断为角膜反射区域。角膜反射区域 B_{pnr} 如图8(a)所示。

定位角膜反射后,再计算角膜反射区域中心。求角膜反射中心的方法一般有两种:一种是先求角膜反射的轮廓,然后对角膜反射边缘进行椭圆拟合,再通过拟合椭圆的中心来估计角膜反射中心^[9];另一种是通过质心来估计角膜反射中心。文献[19]在二值图像中求角膜反射区域的质心,并且通过500个样本比较了两种方法的效果,结果是两种方法的精度相当。由于角膜反射区域边界的不确定性,致使图像二值化时,采用的阈值对最后的角膜反射中心的估计结果有较大影响。本文通过以下两种方法来减小角膜反射中心的估计误差:1)采用多个阈值先分别分割出多个角膜反射区域,然后将分别求取的各个区域质心的平均结果作为角膜反射中心;2)分割出多个角膜反射区域后,不在二值图像中,而是在暗瞳图像中求质心,因为角膜反射中心是角膜反射区域中最亮的部分,越往边界,亮度越低,这样就可以减小模糊的边界区域对质心定位的影响。具体方法如下:

首先,采用多个阈值分别分割出多个角膜反射区域:

$$b_1^{\text{pnr}} \subseteq b_2^{\text{pnr}} \subseteq \dots \subseteq b_k^{\text{pnr}} \quad (1)$$

然后,分别求取各个区域的质心,并将各个区域

质心的平均值作为角膜反射中心,计算公式如下:

$$i^c = \frac{\sum_{k=1}^K i_k^c}{K} = \frac{\sum_{k=1}^K \frac{\sum_{(i,j) \in b_k^{\text{pur}}} p_{i,j} \cdot i}{\sum_{(i,j) \in b_k^{\text{pur}}} p_{i,j}}}{K} \quad (2)$$

$$j^c = \frac{\sum_{k=1}^K j_k^c}{K} = \frac{\sum_{k=1}^K \frac{\sum_{(i,j) \in b_k^{\text{pur}}} p_{i,j} \cdot j}{\sum_{(i,j) \in b_k^{\text{pur}}} p_{i,j}}}{K} \quad (3)$$

(i_k^c, j_k^c) 分别为区域 b_k^{pur} 的质心, (i^c, j^c) 为最后估计的角膜反射区域质心(如图 8(b)所示)。

在确定瞳孔区域后,再检测差分图像瞳孔区域的边缘,即对 R_4 进行二值图像边缘检测, R_6 (图 8(c))为边缘检测的结果。

角膜反射是造成瞳孔轮廓改变的主要因素,是影响瞳孔中心定位精度的主要因素之一。文献[20]通过迭代逼近的方法来定位瞳孔中心,可以在一部分情况下减小角膜反射对瞳孔椭圆拟合的影响,但是该方法存在以下几点问题:1)当角膜反射在瞳孔内部而不与瞳孔轮廓相交的情况下,会造成错误滤波,进而造成瞳孔中心定位的不准确;2)有些使用者的角膜反射较大,而瞳孔较小,这种情况下该滤波算法不能有效地消除由角膜反射产生的边缘;3)迭代方法效率较低,会增加处理时间。

为了有效消除角膜反射对瞳孔椭圆拟合的影响,首先应判断角膜反射区域与瞳孔椭圆的相对位置,当角膜反射与瞳孔椭圆相离时,角膜反射对瞳孔轮廓无影响,不需要进行滤波;当角膜反射区域与瞳孔椭圆相交或包含时,角膜反射改变瞳孔轮廓,需要进行滤波。理想的实心椭圆区域是一个有界严凸集,其边界是一个具有严格凸性的全局闭凸曲线,即使存在采样误差和离散误差,图像中瞳孔的边界也应该是—个全局闭凸曲线,正是由于角膜反射的存在,致使对瞳孔边界产生了较大的影响。本文先通过对瞳孔边界求凸壳来还原瞳孔区域,然后用还原后的瞳孔区域去判断与角膜反射的位置关系。本文采用一种简单快速的凸壳检测算法进行检测,其时间复杂度为 $1 \log N^{[21]}$ 。

$$R_7 = \text{convex hull}(R_6) \quad (4)$$

R_7 是 R_6 的凸壳。 convex hull 为凸壳检测函数。

在此基础上,本文提出了一种基于径向距离的瞳孔边缘专用滤波算法——RDPEF 算法。规则如下:

如果 $R_7 \cap b^{\text{pur}} \neq \emptyset$

$$R_7 = \left\{ (i,j) \mid (i,j) \in R_6 \wedge \left(\forall (p,q) \left(\frac{j-j^c}{i-i^c} = \frac{q-j^c}{p-i^c} \rightarrow \|(i,j) - (i^c, j^c)\| \geq \|(p,q) - (i^c, j^c)\| \right) \right) \right\}$$

否则 $R_7 = R_6$

首先,判断通过凸壳算法还原后的瞳孔区域 R_7 与角膜反射区域 b_k^{pur} 的位置关系,当角膜反射区域与瞳孔椭圆相离时,则不进行滤波,当角膜反射区域与瞳孔椭圆相交或包含时,则进行滤波。

在瞳孔边缘 R_6 中,以角膜反射中心 (i^c, j^c) 为圆心,在每一个半径方向上,如果存在两点 (i,j) 和 (p,q) ,则点 (i,j) 到角膜反射中心 (i^c, j^c) 的直线斜率 $\frac{j-j^c}{i-i^c}$ 与点 (p,q) 到角膜反射中心 (i^c, j^c) 的直线斜率 $\frac{q-j^c}{p-i^c}$ 相同,即点 (i,j) 和点 (p,q) 在同一半径方向上,远端边缘点为瞳孔形成的边缘,近端边缘点为由角膜反射形成的边缘,过滤掉角膜反射形成的边缘后的眼睛区域边缘如图 8(d)所示。RDPEF 算法滤波过程如图 9 所示。

接下来,利用椭圆拟合定位瞳孔中心。文献[22]在亮瞳图像中用椭圆拟合定位瞳孔中心,文献[20]在暗瞳图像中用椭圆拟合定位瞳孔中心。文献[23]采用等照度线的方法定位眼睛中心。本文在差分图像中用椭圆拟合定位瞳孔中心,使噪声边缘更少。本文采用文献[24]中的最小二乘非迭代椭圆拟合算法进行定位。椭圆拟合的另一个好处是可以得到亚像素的中心坐标^[25]。拟合得到的瞳孔椭圆如图 8(e)所示,瞳孔中心如图 8(b)所示。 t 时刻的视线特征向量 L_t 如下:

$$L_t = (\Delta x, \Delta y, \frac{a_{\text{major}}}{a_{\text{minor}}}, \theta, i^c, j^c) \quad (5)$$

其中 $(\Delta x, \Delta y) = (i^p, j^p) - (i^c, j^c)$ 为瞳孔中心到角膜反射中心的向量, $\frac{a_{\text{major}}}{a_{\text{minor}}}$ 为瞳孔椭圆长短轴的比, θ 为瞳孔椭圆长轴与垂直方向的夹角, (i^c, j^c) 为角膜反射中心在图像中的位置。使用者 1 的部分实验数据见表 1, 其中 L_t 为注视屏幕像素点为 (G_x, G_y) 时对应的特征向量, $(\Delta x, \Delta y)$ 的单位为 pixel, $\frac{a_{\text{major}}}{a_{\text{minor}}}$ 为比例,无量纲, θ 的单位为 $^\circ$, (i^c, j^c) 的单位为 pixel。

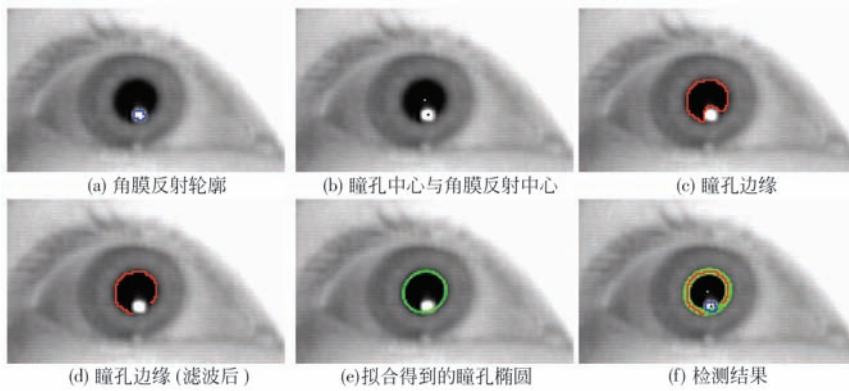


图 8 人眼特征提取过程

Fig. 8 Feature extraction process of the eye

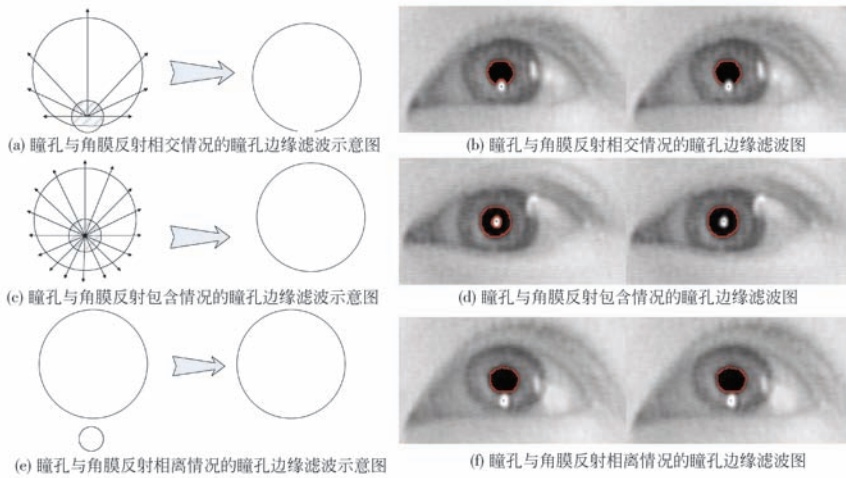


图 9 边缘滤波

Fig. 9 Edge filter

表 1 使用者 1 的部分实验数据
Tab. 1 Part of the user1 datas

特征向量 L_i	屏幕点 (G_x, G_y)
(4. 870, 11. 207, 1. 213, 117. 552, 287. 637, 437. 932)	(88, 86)
(-0. 880, 11. 357, 1. 058, 65. 256, 231. 91, 443. 770)	(726, 34)
(-6. 269, 10. 866, 1. 139, 54. 755, 185. 966, 453. 136)	(1216, 75)
(5. 159, 7. 824, 1. 057, 147. 106, 213. 819, 459. 712)	(42, 469)
(-0. 901, 7. 718, 1. 066, 12. 618, 199. 004, 467. 304)	(707, 387)
(-7. 231, 7. 710, 1. 142, 42. 529, 177. 404, 471. 379)	(1205, 384)
(5. 063, 3. 783, 1. 182, 256. 705, 208. 628, 475. 365)	(26, 860)
(-0. 979, 3. 287, 1. 140, 3. 635, 196. 633, 477. 365)	(581, 800)
(-6. 922, 3. 617, 1. 104, 13. 373, 173. 485, 485. 351)	(1201, 874)

接下来,校验瞳孔中心到角膜反射中心向量的合理性。校验时,首先判断特征向量长度 $\| (i^p, j^p) - (i^c, j^c) \|$ 是否满足条件,如果不满足条件,则返回,不进行下面的处理;如果满足条件,则用本次使用者

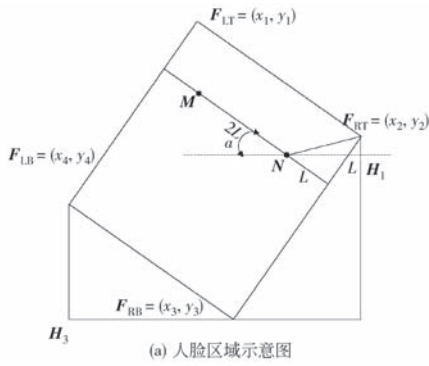
实际定位后的瞳孔灰度去设定差分图像瞳孔的灰度阈值 T_2, T_3 ,最后的检测结果如图 8(f) 所示。

2.3 人脸定位

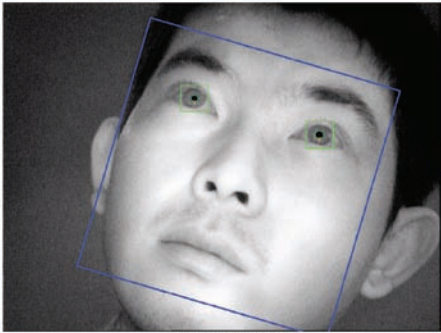
本文根据双眼位置确定人脸位置,左眼坐标记为 $M = (x_L, y_L)$,右眼坐标记为 $N = (x_R, y_R)$,两眼之间距离记为 $2L$,人脸位置记为 $F_{LT}, F_{RT}, F_{RB}, F_{LB}$ (如图 10 所示)。推导过程如下:

如图 10(a) 所示,其中 $\cos \alpha = \frac{x_R - x_L}{2L}$, $\sin \alpha = \frac{y_R - y_L}{2L}$,连接 $\vec{F_{RT}N}$,作 $\vec{NH_1} \perp \vec{F_{RT}H_1}$,则 $\angle F_{RT}NH_1 = \beta = 45^\circ - \alpha$,则

$$\sin \beta = \sin (45^\circ - \alpha) = \frac{\sqrt{2}}{2} (\cos \alpha - \sin \alpha) = \frac{\sqrt{2}}{2} \cdot \frac{x_R - x_L - y_R + y_L}{2L}$$



(a) 人脸区域示意图



(b) 人脸区域图像

图 10 人脸区域

Fig. 10 Face region

则

$$\sin \beta = \frac{y_R - y_2}{\sqrt{2}L} = \frac{\sqrt{2}}{2} \cdot \frac{x_R - x_L - y_R + y_L}{2L}$$

整理得 $y_2 = (3y_R - x_R + x_L - y_L)/2$

同理可得,由

$$\cos \beta = \cos (45^\circ - \alpha) = \frac{\sqrt{2}}{2}(\cos \alpha + \sin \alpha) = \frac{\sqrt{2}}{2} \cdot \frac{x_R - x_L + y_R - y_L}{2L}$$

则

$$\cos \alpha = \frac{x_2 - x_R}{\sqrt{2}L} = \frac{\sqrt{2}}{2} \cdot \frac{x_R - x_L + y_R - y_L}{2L}$$

整理得 $x_2 = (3x_R - y_L + y_R - x_L)/2$

其他点类似可得,整理 4 点坐标为

$$F_{LT} = \left(\frac{3x_L + y_R - x_R - y_L}{2}, \frac{3y_L - y_R - x_R + x_L}{2} \right)$$

$$F_{RT} = \left(\frac{3x_R - y_L + y_R - x_L}{2}, \frac{3y_R - x_R + x_L - y_L}{2} \right)$$

$$F_{RB} = \left(\frac{-y_L + 3y_R + 3x_R - 3x_L}{2}, \frac{3y_L - 3y_R + 3x_R - x_L}{2} \right)$$

$$F_{LB} = \left(\frac{3y_L - 3y_R - x_R + 3x_L}{2}, \frac{3x_R - y_R - 3x_L + 3y_L}{2} \right)$$

以上参数提取过程是在相邻两帧中完成的,尽

管在接下来的视频序列中可以通过重复进行以上的参数提取过程,来得到实时的视线参数,但是这样做有下面两点问题: 1) 每次在整幅图像中搜索瞳孔,效率低,时间长,影响系统的实时性;2) 由于没有使用瞳孔的历史信息,所以使瞳孔定位的鲁棒性差,而且对外部光照、眼睛角度、眨眼等情况的适应性差,不能满足系统的鲁棒性要求。本文在 $t-1$ 时刻成功进行特征参数提取后,再在 t 时刻首先在 $t-1$ 时刻瞳孔位置附近进行特征参数提取,如果不成功,则加大搜索范围;然后通过设定优先搜索区域步骤,可以减少平均特征参数提取的时间,因为人在正常注视屏幕信息的情况下,眼睛运动幅度较小,绝大多数情况下都可以在上一时刻眼睛的附近找到这一时刻的眼睛。

3 实验结果

由于特征提取是视线追踪系统的共性问题,其要求在不同的头部位置都能有效地进行,为了验证本文特征提取方法的效果,进行了如下实验。本文首先对 50 位使用者的数据进行统计,得到阈值,然后对另外 150 位使用者进行了检测实验,实验过程在室内进行,但不同使用者的测试环境不相同,包括白天晚上、晴天阴天、室内照明、灯开关等各种室内光线情况,图 11 为使用者 1,2,3,4 在各种头部位置下的检测结果,由图 11 可以看到,即使在头部转动明显的情况下,本文方法也能有效地提取视线特征参数。

与其他用于视线追踪系统的眼睛检测方法的比较结果见表 2,表中是几种方法的检测成功率,由表 2 可以看到,本文方法的检测效果好于文献[18]中的组合亮暗瞳差分方法与 SVM 方法,也好于文献[26]中的组合亮暗瞳差分方法与 Haar 分类器方法。这两种方法是在亮暗瞳差分方案基础上,加上分类器来对眼睛进行定位。由于训练分类器用的是图像的原始像素信息,是现有方法在瞳孔定位上的应用,没有充分考虑瞳孔的特点,所以难免有一定的漏检和错检,这也是因为瞳孔区域的图像排他性较差,很容易在图像的其他位置有类似瞳孔像素的区域所致。本文方法在亮暗瞳差分方案的基础上,“基于规则”进行瞳孔定位,用的信息是处理后的一

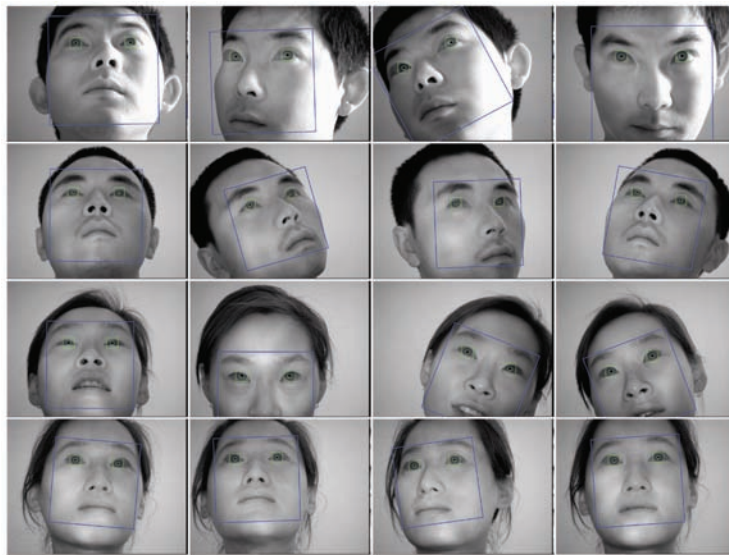


图 11 不同头部位置的参数提取

Fig. 11 Parameter extraction on different head positions

系列“特征”,从而降低了对样本的依赖性,因为这些特征要比图像的原像素信息具有更好的稳定性和排他性。

表 2 眼睛检测方法比较(成功率/%)

Tab. 2 Comparison of eye detection methods(success rate/%)

图像	文献[18]方法	文献[26]方法	本文方法
3 200 帧	91.8	93.6	96.9

在本文开发的视线追踪系统中,对不同使用者的视线估计效果进行了测试,部分测试结果见表 3,水平方向的角度 σ 平均精度在 1.5° 左右,垂直方向的角度 σ 平均精度在 1.9° 左右。图 12 为使用者 3 的视线估计结果。

表 3 不同使用者的视线估计结果

Tab. 3 Average gaze estimation accuracy for seven subjects

使用者	精度($\lambda/\text{mm}(\sigma/^\circ)$)	
	水平方向	垂直方向
1	12.00 \pm 7.54 (0.86 \pm 0.54)	14.24 \pm 8.65 (1.02 \pm 0.62)
2	26.11 \pm 10.19 (1.87 \pm 0.73)	32.83 \pm 12.42 (2.35 \pm 0.89)
3	26.81 \pm 5.58 (1.92 \pm 0.40)	33.80 \pm 6.28 (2.42 \pm 0.45)
4	16.89 \pm 6.42 (1.21 \pm 0.46)	23.46 \pm 7.40 (1.68 \pm 0.53)
5	18.43 \pm 4.18 (1.32 \pm 0.30)	24.02 \pm 5.72 (1.72 \pm 0.41)
6	28.49 \pm 9.21 (2.04 \pm 0.66)	37.44 \pm 12.98 (2.68 \pm 0.93)
7	24.30 \pm 4.74 (1.74 \pm 0.34)	26.81 \pm 7.26 (1.92 \pm 0.52)

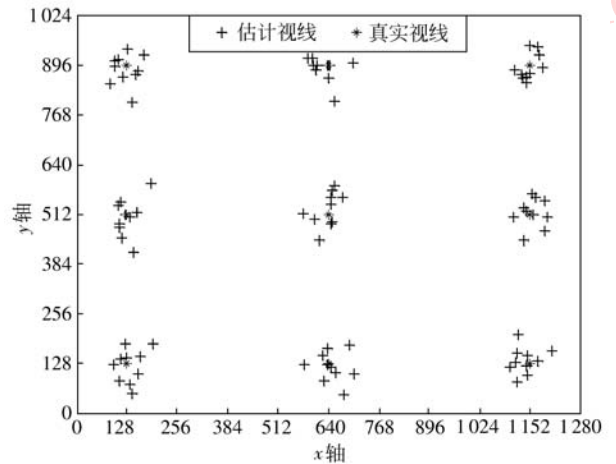


图 12 估计视线与真实视线比较

Fig. 12 Plot of estimated gaze points and true gaze points

4 结 论

特征提取的质量是影响视线追踪系统精度的主要因素之一,为此本文在高精度亚像素特征参数提取方面进行了研究,针对角膜反射对瞳孔边缘的影响,提出了一种基于径向距离的瞳孔边缘专用滤波算法——RDPEF 算法,从而提高了瞳孔的定位精度。针对因瞳孔运动造成的差分图像瞳孔区域不全的问题,通过迭代扩展的方式还原运动瞳孔区域,提高了运动瞳孔定位精度;针对角膜反射区域形状

的不规则和边界的不确定性问题,通过多个阈值多次分割的方式计算瞳孔质心,减小了模糊的边界区域对瞳孔质心定位结果的影响。利用瞳孔的历史信息,即在 $t+1$ 时刻,首先在 t 时刻瞳孔位置的附近进行特征参数提取,提高了瞳孔定位的鲁棒性。本文建立了一个多特征参数提取的流程,用来提取包括人眼特征和人脸位置的多个参数,为下一步的视线估计提供了参数依据。本文通过高质量的特征检测,使开发的视线追踪系统的精度达到了满足实时人机交互的水平。

参考文献 (References)

- [1] Duchowski T. Eye Tracking Methodology: Theory and Practice [M]. New York, USA: Springer-verlag, 2003.
- [2] Jacob R J K. The use of eye movements in human computer interaction techniques: What you look at is what you get [J]. ACM Transactions on Information Systems, 1991, 9(3):152-169.
- [3] LC Technologies Inc. Eyegaze Systems [EB/OL]. <http://www.eyegaze.com/>. [2010-03-01].
- [4] Eyetechn Digital Systems Inc. Eye Tracker [EB/OL]. <http://www.eyetechns.com/>. [2010-03-01].
- [5] Tobii Technology, Eye Tracking technology [EB/OL]. <http://www.tobii.com/>. [2010-03-01].
- [6] Applied Science Laboratories, New to Eye Tracking [EB/OL]. <http://www.a-s-l.com/>. [2010-03-01].
- [7] Zhu Z, Ji Q. Eye and gaze tracking for interactive graphic display [J]. Machine Vision and Applications, 2004, 15(3):139-148.
- [8] Guestrin D, Eizenman M. General theory of remote gaze estimation using the pupil center and corneal reflections [J]. IEEE Transactions on Biomedical Engineering, 2006, 53(6): 1124-1133.
- [9] Zhu Z, Ji Q. Novel eye gaze tracking techniques under natural head movement [J]. IEEE Transactions on Biomedical Engineering, 2007, 54(12):2246-2260.
- [10] Villanueva Arantxa, Cabeza Rafael. A novel gaze estimation system with one calibration point [J]. IEEE Transactions on System Man and Cybernetics System, Man Cybernetics Part B-Cybernetics, 2008, 38(4):1123-1138.
- [11] Ebisawa Y, Satoh S. Effectiveness of pupil area detection technique using two light sources and image difference method [C]//Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Los Alamitos, CA, USA: IEEE Computer Society Press, 1993:1268-1269.
- [12] Ebisawa Y. Improved video-based eye-gaze detection method [J]. IEEE Transactions on Instrumentation and Measurement, 1998, 47(2): 948-955.
- [13] Morimoto C H, Koons D, Amir A, et al. Pupil detection and tracking using multiple light sources [J]. Image and Vision Computing, 2000, 18(4):331-335.
- [14] Morimoto C, Flickner M. Real-time multiple face detection using active illumination [C]//Proceedings of fourth IEEE International Conference on Automatic Face and Gesture Recognition. Los Alamitos, CA, USA: IEEE Computer Society Press, 2000: 8-13.
- [15] Haro A, Flickner M, Essa I. Detecting and tracking eyes by using their physiological properties, dynamics, and appearance [C]// Proceedings of International Conference on Computer Vision and Pattern Recognition. Los Alamitos, CA, USA: IEEE Computer Society Press, 2000: 163-168.
- [16] Ji Q, Yang X. Real time visual cues extraction for monitoring driver vigilance [C]//Proceedings of International Conference on Computer Vision Systems. Los Alamitos, CA, USA: IEEE Computer Society Press, 2001: 107-124.
- [17] Zhu Zhiwei, Fuj Imura K, Iqiang J. Real-time eye detection and tracking under various light conditions [C]//Proceedings of ACM SIGCHI Symposium on Eye Tracking Research and Applications. New Orleans, LA, USA: ACM SIGGRAPH, 2002: 139-144.
- [18] Zhu Zhiwei, Ji Qiang. Robust real-time eye detection and tracking under variable lighting conditions and various face orientations [J]. Computer Vision and Image Understanding, 2005, 38(1): 124-154.
- [19] Wang J G, Sung E, Venkateswarlu R. Eye gaze estimation from a single image of one eye [C]//Proceedings of International Conference on Computer Vision. Los Alamitos, CA, USA: IEEE Computer Society Press, 2003:136-143.
- [20] Talmi Kay, Liu Jin. Eye and gaze tracking for visually controlled interactive stereoscopic displays [J]. Signal Processing: Image Communication, 1999, 14(10):799-810.
- [21] Lin Jachen, Lin Jennyih, A 1 logN parallel algorithm for detecting convex hulls on imageboards [J]. IEEE Transactions on Image Processing, 1998, 7(6): 922-925.
- [22] Hennessey C, Noureddin B, Lawrence P. Fixation precision in high-speed noncontact eye-gaze tracking [J]. IEEE Transactions on System Man and Cybernetics Systems. Man Cybernetics Part B-Cybernetics, 2008, 38(2):289-298.
- [23] Valenti R, Gevers T. Accurate eye center location and tracking using isophote curvature [C]//Proceedings of International Conference on Computer Vision and Pattern Recognition. Los Alamitos, CA, USA: IEEE Computer Society Press, 2008: 1-8.
- [24] Fitzgibbon A, Pilu M, Fisher R. Direct least square fitting of ellipses [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(5):476-480.
- [25] Matsumoto Y, Ogasawara T, Zelinsky A. Behavior recognition based on head pose and gaze direction measurement [C]// Proceedings of IEEE/RSJ International Conference on Intelligence Robots and Systems. Takamatsu, Japan: IEEE Consumer Electronics Society, 2000: 2127-2132.
- [26] Hansen Dan Witzner, Hammoud Riad I. An improved likelihood model for eye tracking [J]. Computer Vision and Image Understanding, 2007, 106(2-3): 220-230.