



# 基于瞳孔角膜反射法的视线估计系统设计与实现

曹奕铨, 蒋路茸, 童基均

(浙江理工大学信息学院, 杭州 310018)

**摘要:** 设计并实现了一种低成本硬件条件下的视线估计系统, 该系统的视线估计模型采用瞳孔角膜反射法, 对用户干扰小, 允许用户头部移动。系统由视频采集、瞳孔中心和反射光斑中心提取、视线估计三部分组成。针对系统中低成本硬件导致的图像质量不高、瞳孔图像分辨率低以及光照变化等情况, 设计了实时、精确的瞳孔中心、反射光斑中心提取方法。该方法首先将粗定位与精确定位相结合, 得到人眼瞳孔图像; 再利用灰度膨胀、卷积操作增强反射光斑, 提取出反射光斑中心; 然后通过一定的预处理, 融合边缘提取、边缘过滤以及椭圆拟合, 得到瞳孔中心; 最后经过后处理滤去异常值。对视线估计系统进行测试, 结果发现: 视线估计平均均方根误差为 12 mm, 满足某些应用的基本需求。该系统降低了对硬件的要求, 有利于实际应用。

**关键词:** 视线估计; 瞳孔角膜反射; 瞳孔中心提取; 反射光斑中心提取; 瞳孔定位

**中图分类号:** TS195.644

**文献标志码:** A

**文章编号:** 1673-3851(2019)03-0201-07

## Design and implementation of gaze estimation system based on pupil and corneal reflection

CAO Yicheng, JIANG Lurong, TONG Jijun

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

**Abstract:** A gaze estimation system under low-cost hardware condition was designed and implemented in this paper. The gaze estimation of this system is based on pupil corneal reflection. The system brings about small interference to users, and user's head movement is allowed. The system is composed of three parts, including video acquisition, pupil center and corneal reflection center extraction, and gaze estimation. Under the conditions of low image quality and low resolution of pupil image caused by low-cost hardware as well as the change of illumination, a real-time and accurate pupil center and corneal reflection center extraction method was designed. Firstly, coarse positioning and precise positioning were combined to gain human pupil images. The gray scale expansion and convolution operations were used to enhance the corneal reflection and help extract the center of the reflection. Then, the pupil center was obtained through preprocessing, edge extraction, edge filtering and ellipse fitting. Finally, exceptional value was removed during postprocessing. The test result for the gaze estimation system showed that, the average root mean square error of gaze estimation was 12 mm, which satisfies the basic requirements of some applications. The system reduces the hardware requirements and has certain reference value in practical applications.

**Key words:** gaze estimation; pupil corneal reflection; extraction of pupil center; extraction of corneal reflection; pupil position

收稿日期: 2018-09-04 网络出版日期: 2018-12-03

基金项目: 浙江省重点研发计划项目 (2015C03023); 浙江理工大学“521 人才培养计划”

作者简介: 曹奕铨 (1994-), 男, 江苏江阴, 硕士研究生, 主要从事计算机视觉方面的研究。

通信作者: 童基均, E-mail: jijuntong@zstu.edu.cn

0 引言

视觉是人类信息获取的重要来源,人眼视线的方向反映了观察者的兴趣范围与关注信息。视线估计技术通过光学、电子等手段获取观察者的注视点,在认知障碍诊断、心理学和人机交互领域有着广泛的应用,如:在医疗上用来分析精神疾病的生物学指标<sup>[1]</sup>,在心理学研究中用以分析用户习惯<sup>[2]</sup>,在虚拟现实作为人机交互的手段<sup>[3]</sup>。

视线估计可以粗略地划分为干扰式和非干扰式,当前视线估计系统大多采用基于视频处理的非干扰式视线估计技术。视线估计系统由视线估计模型建立和视线参数获取两个关键技术组成。基于瞳孔-角膜反射法的视线估计方法是目前广泛使用的视线估计模型,该方法精度较高,对用户无干扰,允许用户头部自由移动。Hansen 等<sup>[4]</sup>将它细分为基于回归与基于模型两类方法。前者假设视线方向向量是瞳孔中心-反射光斑中心向量的映射,通过数值方法进行拟合,使用大量数据针对性地学习训练,如使用多项式拟合<sup>[5]</sup>、神经网络<sup>[6-7]</sup>进行拟合。后者通过几何模型计算出角膜曲率中心,而后得到注视点。Guestrin 等<sup>[8-9]</sup>建立的广义模型是这一类方法的代表,基于模型的方法建立在几何模型之上,通用性与可解释性更强。

视线参数包括眼球的生理参数以及模型参

数,在瞳孔-角膜反射法中主要是瞳孔中心以及角膜反射光斑中心两个中心点,中心点精确地提取是研究的重点。Valenti 等<sup>[10]</sup>提出了等照度线曲率以定位瞳孔中心,陈建等<sup>[11]</sup>提出了一种椭圆拟合与参数传递的瞳孔检测方法,王静等<sup>[12]</sup>提出了非极大抑制圆目标亚像素中心定位。Fuhl 等<sup>[13-14]</sup>针对户外场景提出了 ElSe 与 ExCuSe 两种算法提取中心点。这些提取方法均得到了较高精度的结果,但是对图像分辨率要求很高,需要高分辨率、高解析度的工业相机,在实际应用中受到限制。

本文设计并实现了一种低成本硬件条件下的视线估计系统,该系统使用双相机双光源,使用基于模型的瞳孔角膜反射法,对用户干扰小,允许用户头部移动。针对系统中低成本的硬件条件导致的图像质量不高、瞳孔分辨率低以及光照变化等问题,本文设计了一种实时、精确的瞳孔中心、反射光斑中心提取方法,完成视线参数提取。最后构建了一个完整的视线估计系统,并进行测试。

1 系统设计

本文提出的视线估计系统采用两个红外 LED 点光源、两个低成本的普通红外相机。系统由视频采集、反射光斑中心和瞳孔中心提取以及视线估计 3 个模块组成,系统架构如图 1 所示。

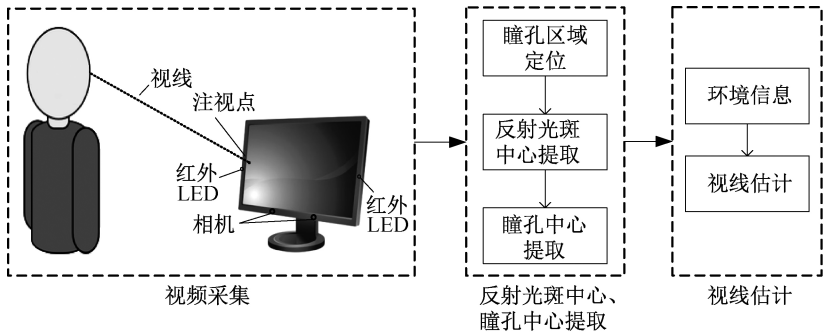


图1 系统架构

视频采集模块负责捕捉用户脸部图像。图像中重点关注瞳孔成像以及用户注视显示屏时,屏幕两侧红外 LED 点光源在角膜上形成的反射光斑成像。

反射光斑中心、瞳孔中心提取模块负责分析用户脸部图像以得到中心点的图像坐标。该子系统是整个系统的核心。针对系统中低成本的硬件条件,本文在反射光斑中心、瞳孔中心提取系统中设计了一种实时且精确的中心提取方法,确保在图像质量

不高、瞳孔图像分辨率低的情况下准确提取瞳孔中心、反射光斑中心同时有效应对光照条件变化导致的光斑区域灰度值变化。

视线估计模块负责计算用户注视点。系统采用瞳孔角膜反射法<sup>[8]</sup>,结合中心点的图像坐标转换得到的世界坐标、相机位置以及红外 LED 位置等环境信息得到用户注视点。该系统需要解决相机标定<sup>[15]</sup>、相机姿态估计<sup>[16]</sup>以及坐标系转换等问题。

## 2 瞳孔中心以及反射光斑中心提取

### 2.1 瞳孔区域定位

用户脸部图像采集完成后,首先定位瞳孔区域。瞳孔区域的精确定位可以减少无关区域的影响,提高算法鲁棒性。本系统中瞳孔定位由两部分组成:粗定位与精确定位。

粗定位由人脸检测与人眼检测组成,二者均使用基于 Haar 特征的级联分类器。人脸检测可以有效减小后续人眼检测搜索区域,将人脸检测结果窗口下采样到同一基准以保证人眼检测能够以一致的尺度进行搜索。

精确定位的目的在于移除粗定位结果中眉毛、眼睑、皮肤等非瞳孔区域。分别计算人眼图像每行每列像素点灰度值,瞳孔区域所在的行或列灰度值和相对更大。计算图像水平方向与垂直方向像素积分图  $I_h$  与  $I_v$  :

$$I_h = \int_{I(x,y) < t} (255 - I(x,y)) dx \quad (1)$$

$$I_v = \int_{I(x,y) < t} (255 - I(x,y)) dy \quad (2)$$

其中:  $I(x,y)$  表示图像  $I$  上坐标为  $(x,y)$  的像素点的像素值,  $t$  为阈值。  $I_h$  与  $I_v$  中非 0 区域就是瞳孔区域所在的行或列,略微偏移修正即可得到瞳孔区域的精确位置。

### 2.2 反射光斑中心提取

反射光斑即相机捕捉的红外光源在角膜形成的虚像,复杂的实际环境以及瞳孔与光源距离的变化导致反射光斑灰度值不断变化。本系统中的反射光斑中心提取方法能够有效应对这种情况,流程如图 2 所示。

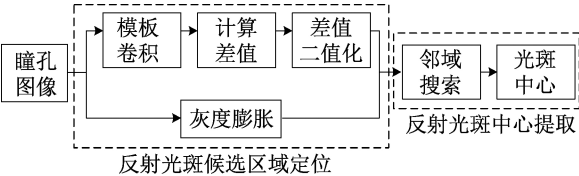


图 2 反射光斑提取流程

#### a) 反射光斑候选区域定位

首先对瞳孔图像执行灰度膨胀操作,即每个像素点的灰度值由包括它在内的四周 5 个像素点灰度值的极大值代替,得到  $I_{dilate}$ 。当反射光斑成像小而且暗时,膨胀可以有效扩大并且增亮光斑区域。

然后使用如图 3 所示的模板进行卷积操作得到

$I_{conv}$ 。

$$I_{conv} = I * h \quad (3)$$

其中  $h$  为卷积模板。

	$\frac{1}{8}$	
$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$
	$\frac{1}{8}$	

图 3 卷积模板

模板  $h$  中中心点处灰度值对卷积结果影响最大。卷积后高亮区域像素点灰度值会显著减少,其余点基本不变。

最后,计算原始图像  $I$  与卷积后图像  $I_{conv}$  的差值  $I_{diff}$ ,对其二值化得到  $I_{binary}$ ,二值化所用阈值由  $I_{diff}$  的统计直方图动态决定。从差值等于 0 开始,累加为当前差值的像素点数量,当数量占比达到预设值时,最后一个像素差值即是二值化所用的阈值。

$I_{binary}$  与  $I_{dilate}$  的逐点相乘得到  $I_{response}$ ,  $I_{response}$  中留下的非 0 区域即是反射光斑候选区域。

#### b) 反射光斑中心提取

求取  $I_{response}$  中每个连通域灰度值的和,值最大的两个区域即为光斑区域。然后在光斑区域的邻域再次进行搜索。设定一个阈值,达到这个阈值的未被包含的点也认为属于光斑区域。本文设定的阈值为上述光斑区域像素均值的 75%。

最终,反射光斑中心由每片区域中像素点坐标按照原始灰度值加权得到。

### 2.3 瞳孔中心提取

图像质量不佳、瞳孔图像分辨率低给瞳孔中心提取带来了一定的难度,本系统瞳孔中心提取由预处理、瞳孔边缘检测与过滤、椭圆拟合三部分组成,具体流程如图 4 所示。

#### 2.3.1 预处理

预处理包括灰度腐蚀和中值滤波两个部分。首先对瞳孔图像执行灰度腐蚀操作,即每个像素点的灰度值由包括它在内的四周 5 个像素点灰度值的极小值代替。然后执行中值滤波,平滑图像。

#### 2.3.2 边缘检测与过滤

##### a) 边缘检测

本系统中使用 Canny 算子对图像进行边缘检测。Canny 检测出的边缘含有一定噪声,并且含有眼睑边缘等非瞳孔边缘线,需进一步进行形态学处理与过滤。

b)形态学处理

形态学处理主要目的进行瞳孔边缘与非瞳孔边缘分割。形态学处理模板如图 5 所示,模板中灰色部分表示匹配成功后移除该像素点,白色表示

示匹配成功后添加该像素点。图 5(a)中模板主要目的在于细化边缘线。图 5(b)分割三岔线,分离瞳孔边缘与眼睑边缘的连接点。图 5(c)补全误删点。

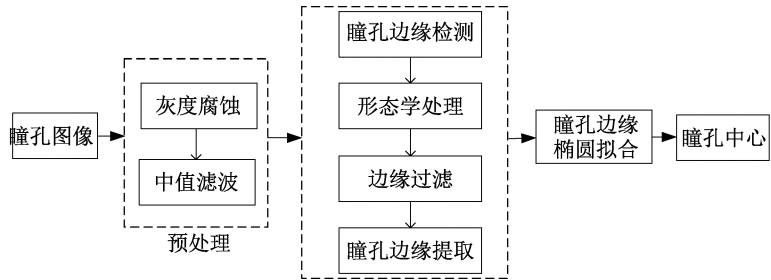


图 4 瞳孔中心提取流程

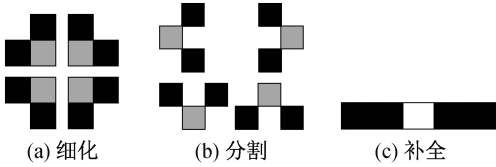


图 5 形态学处理模板

c)非瞳孔边缘过滤

形态学处理后,进一步对离散噪点、直线以及上眼睑等非瞳孔边缘进行过滤。

根据连通域的面积移除离散噪点。直线的判断需要定义一个阈值,当某一线段上所有的点到该线段质心的最短距离小于这个阈值,就认为该线段是直线而后移除。最后移除上眼睑,使用区域生长法得到瞳孔区域点集,其质心为  $g$ ,整条线段都位于质点上方就认为是上眼睑。

d)提取瞳孔边缘

非瞳孔边缘过滤后,余下线段中最长的那条就是瞳孔边缘,但是该边缘两侧部分属于下眼睑,如图 6(a)。计算边缘上相邻两点  $x$  轴坐标差  $\Delta_x$ ,从边缘两侧同时开始搜索,第一个  $\Delta_x=0$  的点就是瞳孔边缘的起始点,如图 6(b)所示。两点之间的线段就是最终的瞳孔边缘线。

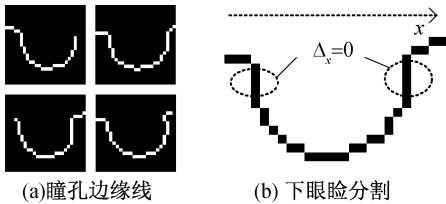


图 6 瞳孔边缘线及下眼睑分割示意

2.3.3 椭圆拟合

相机成像质量以及人眼与相机距离的限制导致多数情况下无法得到完整的瞳孔边缘线。受此限制,本系统中选择了基于近似均方的椭圆拟合算法。

椭圆方程定义为

$$A_{xx}x^2 + A_{xy}xy + A_{yy}y^2 + A_x x + A_y y + A_0 = 0 \tag{4}$$

其中  $A^T = [A_{xx}, A_{xy}, A_{yy}, A_x, A_y, A_0]$  为表达式的系数。

$D$  为设计矩阵,其第  $i$  行为

$$D(i,:) = \{x_i^2, x_i y_i, y_i^2, x_i, y_i, 1\} \tag{5}$$

其中  $x_i$  与  $y_i$  为第  $i$  个边缘点的坐标。

椭圆满足约束条件

$$A^T (D_x^T D_x + D_y^T D_y) A = 1 \tag{6}$$

其中  $D_x$  与  $D_y$  为矩阵  $D$  的偏导数。

损失函数为

$$\epsilon^2 = \frac{A^T D^T D A}{A^T (D_x^T D_x + D_y^T D_y) A} \tag{7}$$

最小化损失函数得到椭圆方程的系数,完成椭圆拟合。椭圆中心坐标即为瞳孔中心图像坐标。

2.4 后处理

取每 30 帧数据的均值作为视线估计所使用的图像坐标。反射光斑中心提取错误或者瞳孔中心提取时由于眨眼等导致瞳孔边缘拟合错误,均会得到错误的坐标,求均值前首先需要剔除这些异常值。设样本均值为  $\mu$ ,标准差为  $\sigma$ ,数据分布在区间  $(\mu - 2\sigma, \mu + 2\sigma)$  中为正常值,否则为异常值。

一组数据包括瞳孔中心在两个相机中的图像坐标,两个反射光斑在两个相机中的图像坐标。每组数据中,对每个数据的每个维度均应用这一原则,任意一个维度不满足,整组数据均被剔除。

3 系统测试

3.1 测试环境

本文点光源采用波长为 940 nm 的红外 LED。

相机为普通红外摄像机,图像分辨率为  $1280 \times 1024$ ,相机仅接收波长  $940\text{ nm}$  附近的光线。电脑硬件配置为:CPU: Intel(R) Core(TM) i5 -4460 CPU @  $3.20\text{ GHz}$ ; 内存:  $16\text{ GB}$ 。

视线估计采用双相机双光源视线瞳孔角膜反射法,测试用户左眼。两个光源固定在显示器两侧,约定为左光源与右光源。两个相机固定在显示器底部,约定为左相机与右相机。

3.2 瞳孔中心以及反射光斑中心提取系统测试结果

a) 单帧测试结果

测试人员距离屏幕约  $45\text{ cm}$ , 注视屏幕上标记点。使用本系统中的反射光斑中心、角膜中心提取方法进行实验。图 7 中选取了部分反射光斑模糊或瞳孔模糊时的结果。图 7(a)中左侧图像为原始瞳孔图像,原始瞳孔图像分辨率低,右侧图像中反射光斑中黑色点为中心点。图 7(b)中第一层图像为原始图像,第二层为瞳孔边缘提取结果,第三层为椭圆拟合结果,椭圆中白色点为瞳孔中心。提取方法平均百次耗时约  $1.61\text{ s}$ , 处理速度可以达到  $60\text{ fps}$  以上,满足实时性要求。

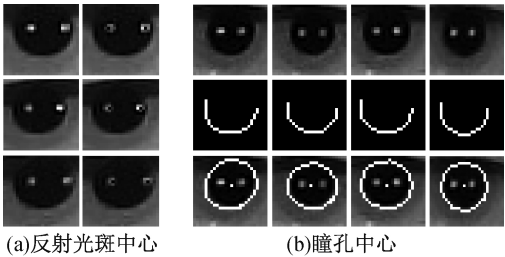


图 7 反射光斑中心以及瞳孔中心提取结果

b) 序列测试结果

测试人员距离屏幕约  $45\text{ cm}$ , 正常注视屏幕上标记点大约  $5\text{ s}$ , 共采集  $70$  组数据。其中两个光源形成的反射光斑中心和瞳孔中心在左右相机中成像的图像坐标如图 8—图 9 所示, 表 1、表 2 为对应的均值与均方差, 以像素为单位。从图表中可以看出图像坐标  $X$  轴均方差不超过  $0.94$  像素,  $Y$  轴均方差不超过  $1.37$  像素, 瞳孔中心、反射光斑中心提取算法精确稳定, 能有效应对光斑区域灰度值变化以及眨眼等导致瞳孔被遮挡的情况。

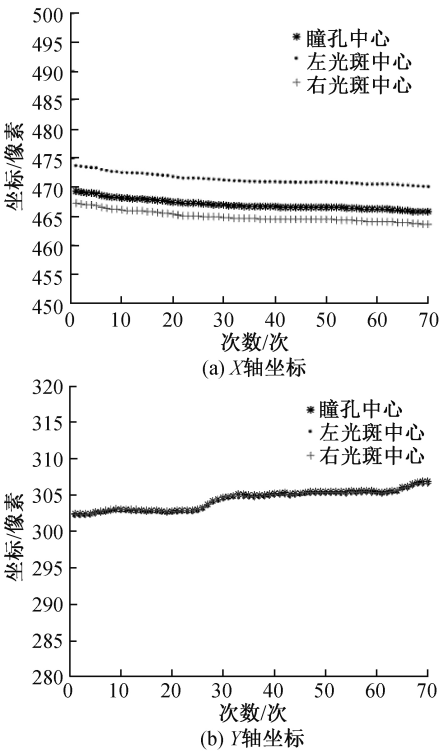


图 8 反射光斑中心、瞳孔中心在左相机中图像坐标

表 1 左相机中 70 次结果两个中心图像坐标均值与均方差

坐标类型	瞳孔中心 X 轴	左光斑中心 X 轴	右光斑中心 X 轴	瞳孔中心 Y 轴	左光斑中心 Y 轴	右光斑中心 Y 轴
均值/像素	467.0514	471.4127	464.9539	304.4405	304.1136	304.3783
均方差/像素	0.8952	0.9340	0.9230	1.3542	1.3157	1.2829

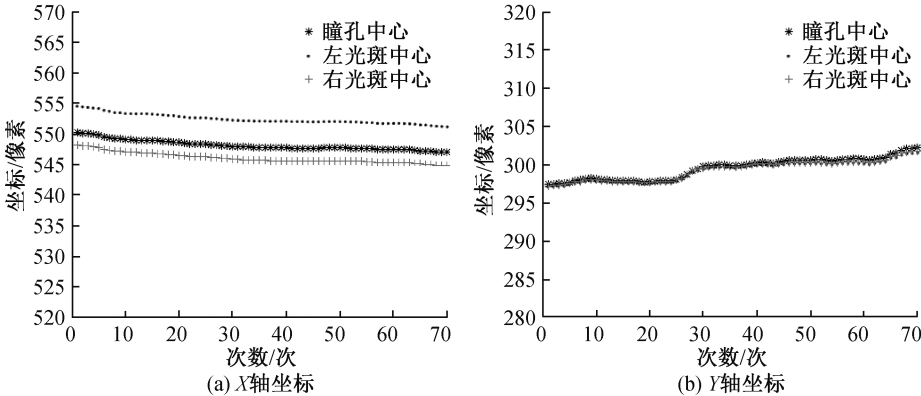


图 9 反射光斑中心、瞳孔中心在右相机中图像坐标

表2 右相机中 70 次结果两个中心图像坐标均值与均方差

坐标类型	瞳孔中心 X 轴	左光斑中心 X 轴	右光斑中心 X 轴	瞳孔中心 Y 轴	左光斑中心 Y 轴	右光斑中心 Y 轴
均值/像素	548.1559	552.4335	546.0622	299.5328	299.1154	299.3206
均方差/像素	0.8174	0.8271	0.8500	1.3685	1.3157	1.3105

3.3 视线估计系统测试结果

测试人员距离屏幕约 45 cm。完成对相机内参标定,通过相机姿态估计得到相机位置以及相机坐标系与世界坐标系之间的旋转平移矩阵后,将本文提取到的中心点图像坐标转换到世界坐标系。利用瞳孔角膜反射算法,估计出注视点坐标。

实验结果如图 10 所示,十字中心为注视点,半径 15 mm。测试人员依次注视 6 个目标点,整体均方根误差约 12 mm,即 1.5°。视线估计系统由视频采集,反射光斑中心、瞳孔中心提取以及视线估计 3 个模块成。误差主要来自中心提取模块以及视线估计模块,误差一方面由瞳孔中心、反射光斑中心提取时的精度导致,另一方面由视线估计方法,即瞳孔角膜反射法导致,该方法对眼球建模简单理想,是误差的主要来源。

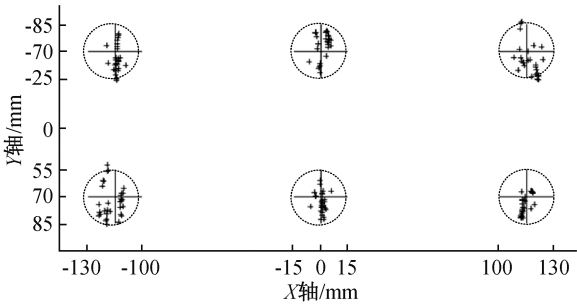


图 10 视线估计系统测试结果

4 结 语

本文设计并实现了低硬件成本条件下基于瞳孔角膜反射法的视线估计系统,系统由视频采集、反射光斑中心和瞳孔中心提取以及视线估计 3 个子系统组成。低成本硬件组成的系统采集的瞳孔区域图像分辨率低、质量差,同时光照条件变化导致了光斑区域灰度值变化。本系统针对采集图像的特点设计了瞳孔中心、反射光斑中心提取方法,利用灰度膨胀腐蚀、形态学处理等技术,实时、准确提取出两个中心点,并且在后处理中滤去异常值,提高方法鲁棒性。对构建的视线估计系统进行测试,结果表明视线估计整体均方根误差为 12 mm,该误差主要由视线估计系统中瞳孔角膜反射法简单理想的眼球模型导致,但是仍然满足某些应用的基本需求。本系统有效降低了对硬件的要求,在实际应用中有一定的参

考价值。

在后续工作中,一方面需要完善硬件配置以得到分辨率更高的瞳孔区域图像,另一方面需要改进视线估计算法,在满足实时处理下引入误差补偿机制以提高精度。

参考文献:

[1] 李瑜, 张天宏, 王俊杰, 等. 精神病超高危人群自由视图眼动特征分析[J]. 中华精神科杂志, 2017, 50(1): 67-71.

[2] 王秋惠, 杨爱慧, 任成元. 基于视线追踪产品设计概念可用性评价方法[J]. 人类工效学, 2011, 17(1):55-59.

[3] 吴艳繁. 基于视线追踪的人机交互系统[D]. 长春: 吉林大学, 2014:5-6.

[4] Hansen D W, Ji Q. In the eye of the beholder; A survey of models for eyes and gaze[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(3):478-500.

[5] Stampe D M. Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems [J]. Behavior Research Methods Instruments & Computers, 1993, 25(2):137-142.

[6] Manabe H, Fukumoto M, Yagi T. Direct gaze estimation based on non-linearity of EOG [J]. IEEE Transactions on Biomedical Engineering, 2015, 62(6): 1553-1562.

[7] Ji Q, Yang X. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance[J]. Real-Time Imaging, 2002, 8(5):357-377.

[8] Guestrin E D, Eizenman M. General theory of remote gaze estimation using the pupil center and corneal reflections [J]. IEEE Transactions on Bio-Medical Engineering, 2006, 53(6):1124-33.

[9] Guestrin E D, Eizenman M. Remote point-of-gaze estimation requiring a single-point calibration for applications with infants[C]//Etra Proceedings of the 2008 Symposium on Eye Tracking Research & Applications, Savannah, Georgia. ACM, 2008:267-274.

[10] Valenti R, Gevers T. Accurate eye center location and tracking using isophote curvature [C]// 2008 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2008:1-8.

[11] 陈健, 郑绍华, 潘林, 等. 结合椭圆拟合与参数传递的瞳孔检测方法[J]. 仪器仪表学报, 2014(8):

1890-1899.

[12] 王静,王海亮,向茂生,等. 基于非极大值抑制的圆目标亚像素中心定位[J]. 仪器仪表学报, 2012, 33(7): 1460-1468.

[13] Fuhl W, Santini T C, Kasneci E. ElSe: Ellipse selection for robust pupil detection in real-world environments[C]//Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications. ACM, 2016:123-130.

[14] Fuhl W, Kübler T, Sippel K, et al. ExCuSe: Robust pupil detection in real-world scenarios[C]// Computer Analysis of Images and Patterns. 2015:39-51.

[15] Zhang Z. A flexible new technique for camera calibration[J]. Microsoft Research, 2000, 22(11): 1330-1334.

[16] Gao X S, Hou X R, Tang J, et al. Complete solution classification for the perspective-three-point problem [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2003, 25(8):930-943.

(责任编辑:康 锋)