

DOI:10.19651/j.cnki.emt.2005429

基于圆形滑动窗口的瞳孔定位算法

童言 穆平安 林玉莹

(上海理工大学 光电信息与计算机工程学院 上海 200082)

摘要: 瞳孔定位是眼动识别中的一个关键步骤。针对目前瞳孔定位算法存在精度低、耗时高、人脸偏转时的定位误差较大等问题,提出了一种通过创建滑动窗口来精确定位瞳孔中心的算法。首先利用 CLM 模型提取人脸的眼部区域,对提取后的图片进行预处理,然后通过创建圆形滑动窗口对图片进行扫描并计算每个窗口内灰度值的和,利用人眼瞳孔的灰度值相对较低这一特性选取灰度值相对较低的窗口为人眼瞳孔区域,计算出其中心点坐标即可定位出瞳孔中心。实验结果表明,该算法针对不同测试对象可以相对精确地定位瞳孔中心,并且具有较高的鲁棒性。

关键词: 计算机视觉;图片处理;瞳孔定位;滑动窗口

中图分类号: TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Pupil location algorithm based on circular sliding window

Tong Yan Mu Ping'an Lin Yuying

(School of Optoelectronic Information and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200082, China)

Abstract: Pupil localization is a key step in pupil recognition. In order to solve the problems of low accuracy, missing detection and false detection in the current pupil location algorithm, this paper proposes an algorithm to accurately locate the pupil center by creating a sliding window. Firstly, CLM model is used to extract the eye region of human face, and the extracted image is preprocessed. Then, by creating a circular sliding window to scan the image and calculate the sum of the gray values of pixels in each window. Taking advantage of the fact that the gray value of the pupil is very low, the window with the lowest gray value is selected as the pupil region, and the coordinates of its center point are calculated as the coordinates of the pupil center. The experimental results show that the algorithm can accurately locate the pupil center for different test objects, and can enhance the robustness of the system.

Keywords: computer vision; image processing; pupil localization; sliding window

0 引言

瞳孔定位在身份识别,视线追踪与医疗诊断等领域起着极其重要的作用^[1]。通过定位瞳孔并分析其运动轨迹可判断视线的移动状态与落点。临床上某些异常的眼动指标可作为诊断精神分裂症的重要依据^[2]。然而眼部遮挡、眨眼、脸部偏转等外界因素的干扰,为精确识别瞳孔中心增加了难度^[3]。

目前使用较为广泛的瞳孔定位方法有 Hough 变换法、椭圆拟合法、梯度向量法与微积分法等。其中 Hough 变换法使用最为广泛,但由于其耗时较高,故不能满足定位的实时性^[4];椭圆拟合法在速度上有一定的优势,但其抗干扰能力很弱并且定位精度一般^[5];梯度向量法定位速度快,并且对图片的质量要求较低,但在光照不均匀的场景下定位精度较低,易受光斑与噪声的影响,鲁棒性较差^[6]。微积分法

定位准确,抗干扰能力强,但存在运算量过大,耗时高的问题^[7]。此外,闫蓓等^[8]提出一种利用射线法提取瞳孔边缘,通过最小二乘法定位瞳孔中心的算法,该算法可以在低分辨率图像中定位瞳孔中心;李擎等^[9]提出一种由粗到精的瞳孔定位算法,采用改进的大津算法(OTSU法)自适应分割瞳孔区域,亚像素定位的方法获取瞳孔边缘,再通过椭圆拟合定位瞳孔中心,可在一定程度上提高定位精度;毛云丰等^[10]通过深度神经网络实现了瞳孔中心的准确定位。尽管瞳孔定位已经有了一定的发展,但针对不同场景下的检测精度与速度仍有待提高。

针对传统算法中存在的定位精度低,抗干扰能力较差等问题,本文提出了一种通过建立滑动窗口检测瞳孔中心的算法:首先通过特征点检测提取图片中的人眼区域,粗定位瞳孔中心。将采集到的眼部图片进行去噪处理,以虹膜为半径建立圆形滑动窗口对图片进行扫描并计算每个窗口

收稿日期:2020-11-25

• 99 •

的灰度值之和,根据人眼瞳孔区域灰度值相对较低的特性找出检测过程中灰度值和最低的窗口即为瞳孔所在区域,由此求出瞳孔的中心位置。该算法可在不同检测角度下准确定位瞳孔中心。

1 人眼区域的粗定位

瞳孔定位的第一步是将人眼区域从一幅图片或一帧视频中提取出来。传统的人眼区域提取方法是根据人脸的结构特性和灰度值的差异结合“三庭五眼”的脸部特征对眼部区域进行提取^[11],该方法在人脸出现偏转时的检测效果并不理想。考虑到人脸的偏转、眉毛、眼镜和睁眼状态对检测结果的影响,首先利用约束局部模型(constrained local model,CLM)对人脸进行特征点检测,再根据眼部特征点提取出眼部区域。检测的过程可分为两步:第 1 步建立模型,构建人脸形状模型并建立图片块模型;第 2 步根据建立好的 Patch 模型搜索出最有可能会出现人眼的位置,再根据形状模型缩小搜索范围,确定出搜索区域^[12],最后在搜索出的人脸位置中提取出眼部区域。

其中 Patch 模型建立旨在寻找一个最佳特征点匹配准则,通过该准则判断当前位置是否为最合理的特征点位置,对每一个特征点的邻域建模从而搜索出最有可能是人脸的点位。在一个训练集中,每一个特征点都存在正例样本和负例样本,样本总数为一个固定值 m ,对这些特征点进行训练从而获得一个如式(1)所示的线性支持向量机。

$$y^{(i)} = \mathbf{w}^T \cdot \mathbf{x}^{(i)} + \theta \tag{1}$$

式中:训练集样本表示为 $\mathbf{x}^{(i)}$; \mathbf{w}^T 为样本的权重且 $\mathbf{w}^T = [\omega_1, \omega_2, \dots, \omega_n]$; θ 表示样本的偏移量; $y^{(i)}$ 为根据线性关系求出的样本的输出。Patch 模型即可由式(1)所示的线性支持向量机训练得到。检测流程如图 1 所示。

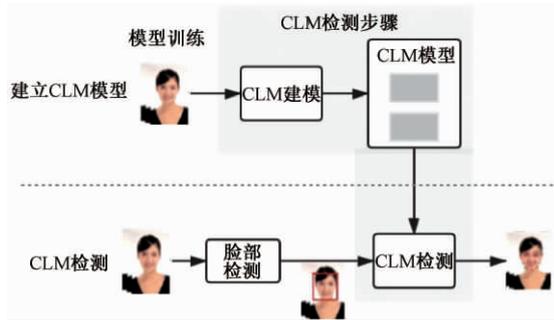


图 1 CLM 模型检测流程

为了解决模型建立过程中的脸部转动、缩放、平移对眼部区域提取造成的影响,采用普氏分析法对数据集中的样本进行形状对齐,再利用主成分分析法(principal component analysis,PCA)对形状对齐后的样本进行降维操作,如式(2)所示。

$$\mathbf{X} = \overline{\mathbf{X}} + \mathbf{PB} \tag{2}$$

式中: \mathbf{P} 代表降维操作后的特征向量矩阵; \mathbf{B} 为权重系数的

集合,且 $\mathbf{B} = [a_1, a_2, a_3, \dots, a_n]$; $\overline{\mathbf{X}}$ 为脸型经平均化处理后的模型。将上式变形,用其他变量表示 \mathbf{B} ,可得式(3):

$$\mathbf{B} = \mathbf{P}^T(\mathbf{X} - \overline{\mathbf{X}}) \tag{3}$$

在 CLM 模型建立完成后,用训练好的模型对人脸进行搜索找出全部的眼部特征点,具体步骤如下:

- 1) 利用平均形状法根据建立好的形状模型粗略定位人脸的特征点位置;
- 2) 在步骤 1) 中的特征点附近的局部区域利用线性支持向量机判断该特征点的可靠性,并为每个特征点生成一张 8×8 大小的响应图;
- 3) 用二次函数拟合法对生成的响应图进行函数拟合;
- 4) 根据 CLM 模型对二次函数优化,通过限制搜索范围得到最可能是特征点的位置;
- 5) 重复上述步骤直至函数收敛。

通过该方法可在人脸图片上得到表示眼部的 22 个特征点,如图 2 所示。

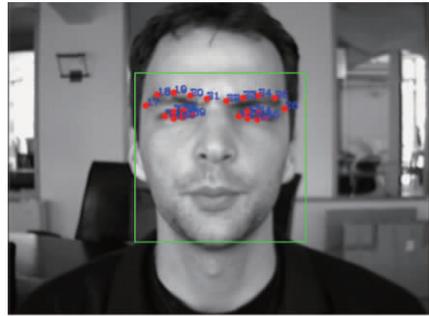


图 2 人脸特征点检测结果

图 2 中的点为检测到的特征点,所有 22 个特征点将具有显著特征的眼部区域表示出来。在检测到的特征点中,单个人眼部分用 6 个特征点表示。点 36~41 为图中左眼的边缘点(实际为人的右眼);点 42~47 为图中右眼的边缘点(实际为人的左眼,本文统一以图中左右眼为准)。点 17~21 构成左眼眉毛部分的边缘点;点 22~26 为右眼眉毛的特征点。

2 眼部区域的提取算法

在获得表示人眼的特征点后,须根据特征点的坐标进一步将眼部区域提取出来。操作思路是以一个矩形框将眼部单独选取并去除眼镜、眉毛等部位的影响。所以这一步的关键是如何确定矩形框的大小以及具体位置。由几何知识可知,只要确定了矩形的宽 W 、高 H 以及左上角的顶点 (X, Y) 便可以根据几何特性分离出眼部区域的图片。

左眼区域和左半眉毛的特征点如图 3 所示,以最左边的点为起点,按照顺时针方向将眼部 6 个特征点依次表示为 $M_1(X_{m1}, Y_{m1}), M_2(X_{m2}, Y_{m2}), M_3(X_{m3}, Y_{m3}), M_4(X_{m4}, Y_{m4}), M_5(X_{m5}, Y_{m5}), M_6(X_{m6}, Y_{m6})$, 眉毛的 5 个特征点依次表示为 $N_1(X_{n1}, Y_{n1}), N_2(X_{n2}, Y_{n2}), N_3(X_{n3}, Y_{n3}),$

$N_4(X_{n4}, Y_{n4}), N_5(X_{n5}, Y_{n5})$ 。

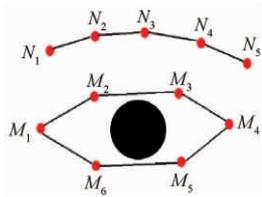


图3 左眼与眉毛的特征点分布

利用特征点的坐标可将眼部矩形区域的左上角顶点表示出来,其横坐标 X 为:

$$X = x_{m1} - \alpha\omega_e \quad (4)$$

$$\omega_e = x_{m4} - x_{m1} \quad (5)$$

式中: α 的范围为 $0 \sim 0.5$; ω_e 表示眼睛的宽度。通过眉毛和眼睛的特征点可确定矩形左上角顶点的 Y 坐标。本文通过计算眉毛的最低特征点与眼睛最高特征点 Y 坐标的算术平均值来确定矩形框左上角顶点的 Y 坐标,如式(6)所示。

$$Y = \frac{1}{2}[\max(y_m) + \min(y_n)] \quad (6)$$

在确定矩形框的左上顶点后,可根据顶点坐标计算出矩形框的高度 H 以及宽度 W ,如式(7)、(8)所示。

$$H = \beta(y_{m1} - Y) \quad (7)$$

$$W = \lambda\omega_e \quad (8)$$

式中: β 的取值范围在 $2 \sim 2.5$ 之间; λ 的取值范围在 $1.5 \sim 2$ 之间。由于各个对象的眼睛眉毛之间存在一定的差异性^[13],所以眼部矩形框的大小没有固定标准,只要参数在给定范围内都能满足要求。在确定矩形框的左上角顶点与宽和高后,便可以根据这3个要素从人脸区域将眼部提取出来并剔除眼镜和眉毛等因素的影响,实验结果如图4所示,白色矩形框内为提取出的人眼区域。

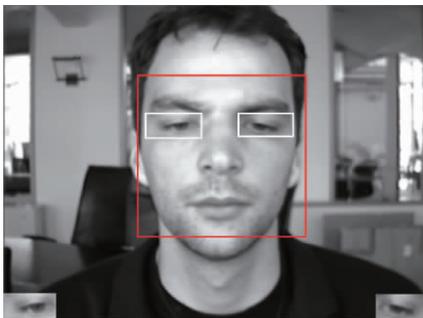


图4 人眼区域提取结果

实验表明,对于正面的人脸图片,本文提到的算法可以准确地将眼部区域提取出来,而对于脸部具有一定倾斜角度的图片,其眉毛部位的特征点并不总位于眼睛特征点的上方,因此按照本文方法得出的矩形框往往不能分割出完整的眼部区域,并且会将眼镜之类的干扰也包含进去。如图5所示,当图片中的人脸发生偏转时,提取出的人眼图片不完整。

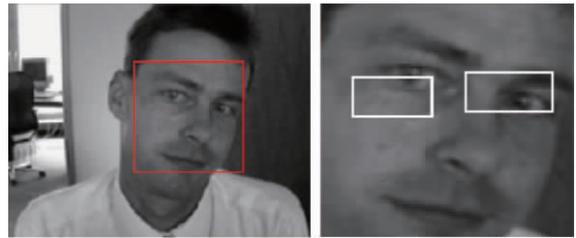


图5 人脸偏转下的眼部提取效果

为了解决这个问题,首先对倾斜的人脸图片进行校正,再将人眼区域从校正后的图片中提取出来。校正的方法是选取两条眉毛之间距离最近的两个特征点 N_1 与 N_2 ,定义它们的连线与水平面的夹角为旋转角 θ ,则有:

$$\theta = \frac{\arctan(y_{n2} - y_{n1})}{x_{n2} - x_{n1}} \quad (9)$$

在确定了角度 θ 之后,分3个步骤将原图片旋转:

1)建立新的数学坐标系,将原图片的坐标系转换为新的坐标系;

2)根据公式转换原图片中的坐标,旋转后的图片可以在原图片中找到每一点所对应的坐标,通过复制对应坐标点的像素值实现图片的旋转,再用本文提到的方法提取眼部区域;

3)再将转换后的数学坐标系转换为原图片的坐标系。坐标的转换模板如式(10)所示。

$$[x \ y \ 1] = [x' \ y' \ 1] \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ -\frac{W'}{2} & \frac{H'}{2} & 1 \end{bmatrix} \quad (10)$$

$$\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ -\frac{W}{2} & \frac{H}{2} & 1 \end{bmatrix}$$

其中, x, y 分别表示原图片中的任意一个点的横坐标与纵坐标; x', y' 为该点旋转后所对应的横纵坐标; W, H 与 W', H' 分别表示旋转前原图片的宽和高和旋转后的图片的宽与高。通过该公式可将原图片旋转进而得到理想的眼部矩形区域。用该方法提取图片的效果如图6所示,即使人脸发生了偏转,依然准确地将眼部完整分离出来。

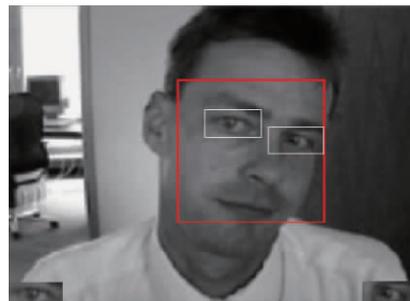


图6 旋转校正后提取出的人眼区域

3 瞳孔中心定位算法

由眼部解剖学理论可知,人眼虹膜是眼部灰度值最低的区域,它的形状近似于圆形^[14]。虹膜上的圆形开口便为瞳孔,开口区域的中心点的位置就是瞳孔中心。根据本文中得到的人眼特征点与几何知识可以计算出图片中的人眼虹膜大小,如式(11)所示。

$$R_i = \frac{|M_2 M_5|}{2} \quad (11)$$

式中: R_i 表示虹膜半径,当外界光照条件发生变化时,虹膜通过改变瞳孔直径的大小可控制进入人眼的光亮,而其形状与大小整体始终保持不变。所以不管眼球的转动方向如何,虹膜形状为固定大小的圆形与它是眼部灰度值最低区域的特性依然显著。尽管它的位置发生了变化,但只要按照其特征在眼部区域搜索便能实现瞳孔的定位,再根据瞳孔的圆形特性找出圆心便可定位瞳孔中心。因此,本文设计了一种基于圆形滑动窗口根据灰度值的和来确定瞳孔中心的定位算法,具体实现方法分为以下步骤:

1) 提取眼部区域,根据前文中提到的算法从检测到的人脸特征点中分离出眼部区域;

2) 平滑图片,按照步骤 1) 得出的眼部区域图片往往包含噪声,通过实验对比发现,利用高斯滤波模板对图片进行去噪相比较于其他滤波器能起到更好的去噪效果,所以采取高斯滤波器对眼部图片进行平滑;

3) 创建圆形模板对眼部图片进行逐行提取,依据本文中求出的虹膜半径 R_i , 利用大小为 $R_i \times R_i$ 的圆形窗口,设置步长为 $0.5R_i$ 从图片的左上角出发对眼部区域扫描后得到若干的圆形子区域图片;

4) 计算子区域内灰度值的和,将其表示为 v 并进行归一化处理,图片背景区域的灰度值较高,所以其对应圆形窗口灰度值的和 v 较大。反之,虹膜区域的灰度值较低,所以其对应的 v 相对较小;

5) 选取步骤 4) 中 v 最小的窗口作为人眼的瞳孔,计算出窗口的中心点即可作为瞳孔中心的定位结果。

用上述方法创建眼部滑动窗口对眼部区域进行灰度值计算,针对不同情况下的瞳孔中心定位如脸部发生偏转,被测者佩戴眼镜等情形下的部分测试结果如图 7 所示。左侧为检测到的人脸图片,右侧为根据本文算法定位图片中的人眼瞳孔的结果展示。由结果可知,在眼睑挡住大部分虹膜的情况下本文算法依然能够较为准确地定位瞳孔中心,出于对虹膜与瞳孔的形状,灰度特性的综合考虑,本算法在针对不同脸部姿势下的不同被测对象进行瞳孔定位测试时都能取得很好的结果,证明本文算法的准确性高,并具有良好的鲁棒性。

4 实验与分析

为了更加客观地比较本文算法与其他算法对于瞳孔中

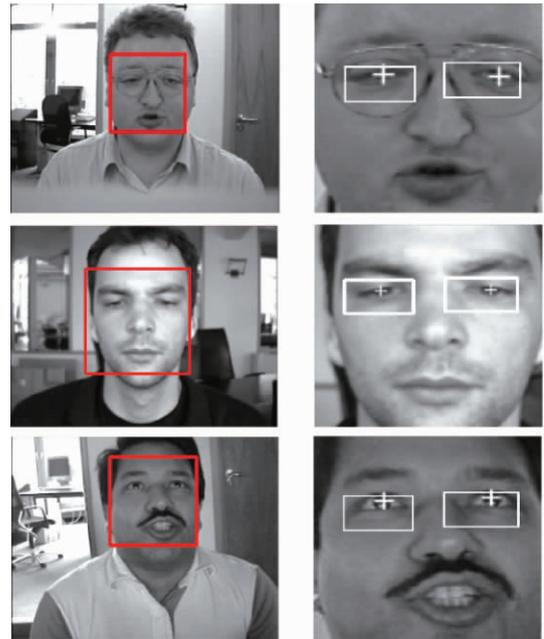


图 7 本文算法在不同情况下的检测结果

心定位的准确性,本文根据 Jesorsky 等所提出的一种误差评价方法,定义瞳孔中心的定位误差为 ϵ , ϵ 可用式(12)来表示。

$$\epsilon = \frac{\max(d_l, d_r)}{\|P_l - P_r\|} \quad (12)$$

式中: d_l 表示真实的左眼瞳孔中心 P_l 和算法得出的左眼瞳孔中心 P_l 之间的欧氏距离,右眼的欧氏距离表示为 P_r ; 左右眼瞳孔的相对距离表示为 $\|P_l - P_r\|$, 由于一般情况下人的两眼运动方向与幅度都具有有一致性,所以两眼之间的相对距离可以看作为一个固定值^[15]。以灰度值为单位,通过计算瞳孔中心的期望位置与估测位置的相对距离,采用欧几里得距离度量法,定义瞳孔中心定位的误差大小。

式(11)对瞳孔定位的误差大小做出了定义,通过分析该式可知, ϵ 的值越小则代表实验结果越接近期望值,瞳孔中心的定位越准确。本文规定,若通过算法定位出的瞳孔落在虹膜区域则视为定位成功,反之则表示定位失败。由生理学知识可知,人眼的虹膜半径通常为左右眼瞳孔的相对距离 $\|P_l - P_r\|$ 的 $1/8$, 因此规定若定位误差 ϵ 小于 0.125 则视为定位成功,反之视为定位失败。

根据该误差评判方法,本文从 BioID 数据库中选取了 20 位测试人员的 200 张不同角度、姿态的正脸灰度图片,并分别用 Hough 变换法,梯度向量法和本文算法对其进行瞳孔中心定位。由于该数据集的每张图片都带有手工标注的人眼位置文件,因此可以方便地对算法误差进行比较。本文实验设备的 CPU 为 Intel core i7-4720HQ@2.6GHz, 8 G 运行内存,在 Python3.7 的环境下通过调用 openCV3.4.2 进行本次实验,实验结果如表 1 所示。

表1 不同算法之间的实验对比

算法	Hough变换法	梯度向量法	本文算法
平均误差 ϵ	0.043	0.105	0.027
平均耗时/ms	202.39	164.28	178.84

试验结果表明,在检测精度上本文算法优于其他两种算法,并且在运行耗时上远远低于 Hough 变换法,虽然速度略逊于梯度向量法,但是更高的精度足以弥补这一方面的不足。为了进一步比较 3 种算法鲁棒性的优劣,选取不同偏转角度下的人脸图片再次进行实验比较,实验结果如表 2 所示。

表2 不同偏转角度下的实验结果

误差	Hough变换法	梯度向量法	本文算法
左偏 10°	0.046	0.061	0.033
右偏 10°	0.081	0.065	0.051
左偏 20°	0.081	0.087	0.045
右偏 20°	0.067	0.117	0.052
左偏 30°	0.144	0.167	0.102
右偏 30°	0.136	0.136	0.122

根据表 2 中的实验结果可以得到图 8 所示的误差曲线,其中横坐标为偏转角的方向以及偏转大小,纵坐标为相应人脸图片的平均定位误差大小。

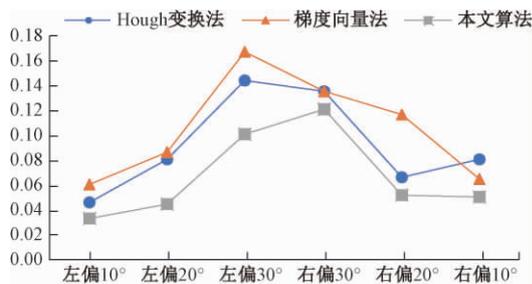


图8 不同角度下各个算法的实验结果

通过分析图表中的数据可知,当人脸偏移角度在 10° 以内时,3 种算法均能较准确的定位出瞳孔中心,平均误差 ϵ 总体保持在 0.08 以内。在偏右侧 10° 的实验结果中, Hough 变换法的误差超过了其他两种算法,分析后发现有几张图片在用 Hough 变换法检测时出现了错误导致误差变大。可以发现,当偏转角度变大时,3 种算法的 ϵ 也随之变大,但总体来说,本文算法相比较于其他两种算法的定位误差更小,鲁棒性更高。

5 结 论

本文提出了一种基于计算机视觉的瞳孔中心定位算法,该算法首先对偏转后的人脸进行校正以便分离出完整的眼部图片,在提取出眼部区域后利用人眼虹膜的灰度与形状特

性对瞳孔中心进行定位。与其他算法相比,本文算法定位精度高,计算复杂程度小,能够有效克服脸部偏转和眼睛部分被遮挡的影响,可在不同姿态下对瞳孔进行准确定位。

参考文献

- [1] 胡艳红,魏江.红外图片中瞳孔定位算法[J].电子设计工程,2019,27(1):189-193.
- [2] FRANCO J G, PABLO D J, GAVIRIA A M, et al. Smooth pursuit eye movements and schizophrenia: Literature review [J]. Archivos De La Sociedad Española De Oftalmología, 2014, 89(9):361-367.
- [3] ABBASI M, KHOSRAVI M R. A robust and accurate particle filter-based pupil detection method for big datasets of eye video [J]. Journal of Grid Computing, 2020, 18(2): 305-325.
- [4] NUGROHO R H, NASRUN M, SETIANINGSIH C. Lie detector with pupil dilation and eye blinks using Hough transform and frame difference method with fuzzy logic [C]. 2017 International Conference on Control, Electronics, Renewable Energy and Communication, 2017:40-45.
- [5] 王娟,刘嘉润,李瑞瑞.基于深度学习的红外相机视力检测算法[J].电子测量与仪器学报,2019,33(11):36-43.
- [6] 王剑.自由立体显示系统中的人眼跟踪定位方法[J].电子测量技术,2015,38(9):52-54.
- [7] 周小龙,姜嘉祺,林家宁,等.一种基于三维眼球模型和Snakuscle的虹膜中心定位算法[J].计算机科学,2019,46(9):284-290.
- [8] 闫蓓,吴梦瑶.低分辨率图像中的瞳孔中心精确定位方法[J].电子测量技术,2018,41(16):74-78.
- [9] 李擎,胡京尧,迟健男,等.视线追踪中一种新的由粗及精的瞳孔定位方法[J].工程科学学报,2019,41(11):1484-1492.
- [10] 毛云丰,沈文忠,滕童.基于深度神经网络的视线跟踪技术研究[J].现代电子技术,2020,43(16):39-43.
- [11] 孟春宁.人眼检测与跟踪的方法及应用研究[D].天津:南开大学,2013.
- [12] 曾启飞,姚剑.基于 CLM 的人脸特征点检测[J].黑龙江科技信息,2017(9):92-92.
- [13] 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,2014,35(8):1890-1899.
- [14] 吕慷,张旭秀,李卫东.基于人眼识别原理的运动目标检测方法[J].电子测量技术,2019,42(4):65-69.
- [15] 李锐,蔡兵,刘琳,等.基于模型的驾驶员眼睛状态识别[J].仪器仪表学报,2016,37(1):184-191.

作者简介

童言,硕士研究生,主要研究方向为图像处理与视觉检测。
E-mail:1956685582@qq.com

穆平安,博士,教授,硕士生导师,主要研究方向为图像处理、测试信息获取与处理、在线检测技术与装置。
E-mail:mupingan@163.com

林玉莹,硕士研究生,主要研究方向为测试计量技术与仪器。
E-mail:lyyusst@163.com