

DOI: 10.3969/j.issn.1001-4551.2016.07.025

# 面向人机交互的眼动跟踪方法研究\*

杨庆华, 张达磊, 荀 一, 鲍官军

(浙江工业大学 机械工程学院, 浙江 杭州 310014)

**摘要:** 针对患有严重身体残疾和行动不便的特殊人群难以与计算机进行交互的问题, 对基于眼动跟踪的人机交互系统进行了研究, 采用可穿戴式的系统硬件方案: 将连接个人电脑的微型 CCD 摄像头通过 3D 打印成型的载具安装在普通眼镜上实现了眼图采集功能, 运用基于梯度值和圆特性的改进的椭圆拟合方法检测定位瞳孔中心, 并进行了基于二阶模型的注视点的标定, 最终实现了对眼睛视线的跟踪检测。实验测试结果表明, 该系统的水平方向视线检测误差和垂直方向视线检测误差分别在 3.74% 和 5.34%, 验证了所提出的眼动跟踪人机交互系统的可行性和有效性。

**关键词:** 眼动跟踪; 人机交互; 视线检测

中图分类号: TP314.7

文献标志码: A

文章编号: 1001-4551(2016)07-0904-05

## Eye-gaze tracking method for HCI system

YANG Qing-hua, ZHANG Da-lei, XUN Yi, BAO Guan-jun

(School of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310014, China)

**Abstract:** Aiming at the difficulty of interaction with computer for people with severe physical disabilities and special people with mobility impairments, study on an interactive system based on eye-tracking was carried out. Wearable systems hardware is introduced as the solution: the CCD camera that connected with the PC is mounted on the spectacles via the carrier made through 3D printing. Through eye diagram capturing, utilizing the improved ellipse fitting method to detect the center of the pupil, and calibrating the gaze point based on second-order model, eventually, the eye gaze tracking and detection in natural light is achieved. The user tests show that the error of eye-gaze detection on horizontal and vertical is 3.74% and 5.34% respectively, verifying the feasibility and effectiveness of the proposed interactive eye-tracking system.

**Key words:** eye tracking; human computer interaction (HCI); eye-gaze detection

## 0 引 言

在各种领域的眼动跟踪研究有助于了解人类的行为和对视觉刺激的反应<sup>[1]</sup>。此外, 特别针对患有残疾的人们, 眼动跟踪还被研究运用于电脑界面的输入设备上。早期, 基于机械式和光学的直接安装在巩膜接触透镜或者眼睛上的一些眼睛运动检测方法(如 Duchowski) 已经能达到非常高的精度, 然而其强烈的侵入性会对用户造成不良的使用感。<sup>[2]</sup> 后来, 由 Young 和 Sheena<sup>[3]</sup> 提出的眼电图法( Electro-Oculogram ), 同

样具有侵入性, 且精度较低, 噪声较大。

目前大多数商业视线检测使用红外( Infrared Radiation ) 相机和主动红外照明的方法来实现高精度的视线方向评估。然而, 阳光和用户所佩戴的眼镜会一定程度上降低红外图像跟踪方法的可用性和准确性, 同时长时间的红外光照射也会对眼睛产生一定程度上的伤害。为了解决这个问题, 可见图像跟踪方法在过去的几年中得到了大量的研究。不同于基于眼图中瞳孔和角膜反射坐标的红外图像跟踪方法, 可见图像法直接映射眼睛虹膜中心( Iris Center ) 的位置到目标平

收稿日期: 2016-03-03

基金项目: 国家科技支撑计划资助项目( 2015BAF01B02 )

作者简介: 杨庆华( 1964- ) , 男, 浙江义乌人, 博士, 教授, 博士生导师, 主要从事塑性成形与机器人技术方向的研究。E-mail: king@zjut.edu.cn

面显示器屏幕。Daugman<sup>[4]</sup>提出一种使用积分-微分算子(Integro-Differential Operator)在目标形状模型下提取眼图像中圆形虹膜边界(Iris Boundary)的方法,其缺点在于模型的参数需要迭代或手动调整,不适合实际应用。Valenti and Gevers<sup>[5]</sup>提出一种基于特征的跟踪方法,使用等照度线特征(比如同等强度下的曲线连接点)来确定虹膜的中心,这种方法的不足之处在于当眼睛区域的特征数目不够多时经常会检测到眉毛或者眼角而不是虹膜中心。Takeshi Takegami等人<sup>[6]</sup>提出了利用Hough变换拟合椭圆的算法来定位瞳孔,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。后来,Lin等人<sup>[7]</sup>提出基于最小二乘法的椭圆拟合方法,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。

本研究以椭圆法线法为基础,引入瞳孔边缘像素点的灰度梯度方向特性,利用最小二乘法,运用改进的椭圆拟合方法来检测视线点,开发一套便携移动式眼动跟踪系统。

## 1 眼动跟踪方法框架

标准的实时眼动跟踪的算法流程图如图1所示。

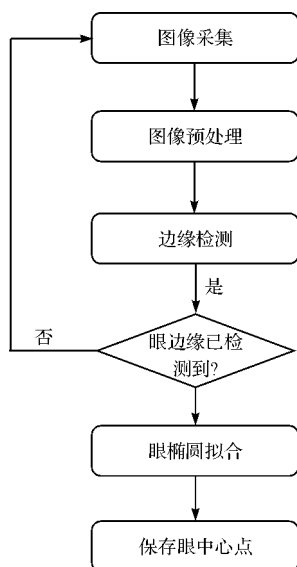


图1 标准眼动跟踪方法框架

## 2 眼动图像处理

通过视屏摄像头采集用户眼睛图像后,需对图像进行滤波、二值化等图像预处理,为后续眼动跟踪提供基础。

### 2.1 图像滤波

选取合适的滤波方法对后续的图像处理至关重要。经过多次对比测试,在均值滤波、高斯滤波、中值滤

波和双边滤波这四种主流滤波方法中,本研究选取中值滤波进行图像预处理,选取参数为 $9 \times 9$ 的核大小。

滤波后的瞳孔图像如图2所示。



图2 中值滤波后的瞳孔图像

### 2.2 图像二值化

因为自然光的条件下,瞳孔与虹膜、眼白等其他眼组织反射的光线有比较明显的色差,在采集的图像中,瞳孔呈现黑色,因此本研究可以运用一个比较合适的低阈值对图像进行二值化处理,即可得到灰度值比较低的瞳孔区域的图像。但自然光线与环境的变化会对采集到的图像中瞳孔的灰度值有较大的影响,所以不适宜设置固定不变的阈值来进行二值化处理,在本研究中,本研究运用最大类间方差法来选择最佳阈值<sup>[8]</sup>,具体算法如下:

设图像中灰度为 $i$ 的像素数为 $n_i$ ,灰度范围为 $[0, L-1]$ ,则总的像素数为:

$$N = \sum_{i=0}^{L-1} n_i \quad (1)$$

则各灰度值出现概率为:

$$p_i = \frac{n_i}{N} \quad (2)$$

对于 $p_i$ ,有:

$$\sum_{i=0}^{L-1} p_i = 1 \quad (3)$$

把图中的像素用阈值 $T$ 分成两类 $C_0$ 和 $C_1$ , $C_0$ 由灰度值在 $[0, T-1]$ 的像素组成, $C_1$ 由灰度值在 $[T, L-1]$ 的像素组成,则区域 $C_0$ 和 $C_1$ 概率分别为:

$$P_0 = \sum_{i=0}^{T-1} p_i \quad (4)$$

$$P_1 = \sum_{i=T}^{L-1} p_i = 1 - P_0 \quad (5)$$

区域 $C_0$ 和 $C_1$ 的平均灰度分别为:

$$\mu_0 = \frac{1}{P_0} \sum_{i=0}^{T-1} i p_i = \frac{\mu(T)}{P_0} \quad (6)$$

$$\mu_1 = \frac{1}{P_1} \sum_{i=T}^{L-1} i p_i = \frac{\mu - \mu(T)}{1 - P_0} \quad (7)$$

其中: $\mu$ —整幅图像的平均灰度,且:

$$\mu = \sum_{i=0}^{L-1} i p_i = \sum_{i=0}^{T-1} i p_i + \sum_{i=T}^{L-1} i p_i = P_0 \mu_0 + P_1 \mu_1 \quad (8)$$

两个区域的总方差为:

$$\sigma_B^2 = P_0(\mu_0 - \mu)^2 + P_1(\mu_1 - \mu)^2 = P_0P_2(\mu_0 - \mu_2)^2 \quad (9)$$

让  $T$  在  $[0, L-1]$  范围内依次取值, 使总方差获得最大的值的  $T$  便是最佳区域分割阈值。

二值化后的瞳孔图像如图 3 所示。



图 3 二值化后的瞳孔图像

### 3 瞳孔中心定位

在普通自然光线下, 不同人的瞳孔大小在一个特定的范围内变化, 因此本研究采用了一种基于霍夫圆变换的算法进行瞳孔中心的定位, 该算法在噪声、遮挡和变化照明的不利情况下有较好的稳定性。

#### 3.1 累加阵列计算

高梯度的前景像素点被选定作为候选像素点, 并且可以在累加阵列中进行“表决”。在经典霍夫圆变换执行中, 候选像素在周围形成的固定半径的完整圆周围进行投票。描述了一个候选像素坐落在实际圆(实心圆)和经典霍夫圆变换对该候选像素进行表决的模式如图 4(a) 所示。

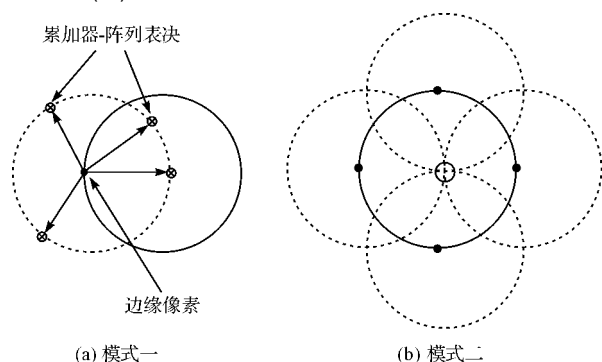


图 4 经典霍夫圆变换表决模式

#### 3.2 圆心定位

属于一个图像中圈的候选像素的表决往往积聚在相应于圆的圆心累加器阵列中。因此, 圆的圆心通过检测累加器阵列中的峰值得到定位。候选像素们(实心点)坐落在实际圆上(实心圆圈)和它们的表决模式(虚线圆圈) - 与实际圆的中心重合如图 4(b) 所示。

本研究运用以上方法对瞳孔图像逐帧进行图像处理, 以确定瞳孔中心。检测到的瞳孔中心示意图如图 5 所示。



图 5 检测到的瞳孔中心

## 4 注视点标定

本研究使用均匀分布在屏幕上的 25 个圆点来完成标定的程序。每一次只有一个点以随机的方式出现, 每个点每次显示 2 s, 从而提供给研究人员足够的标定数据。并且, 为了最大程度上增加测试者的注意力, 屏幕上显示的圆点的半径依次逐渐减小。屏幕上的标定圆点如图 6 所示。

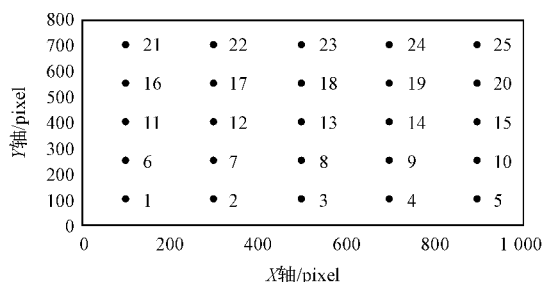


图 6 屏幕上的标定点

在完成标定程序之后, 测试者要观察一组数量为 25 个的在屏幕上随机出现的圆点, 这样每一次实验都能拥有一组既随机又唯一的测试点。这些测试点将用来检测标定技术是否达到功能要求。

本研究采用基于二阶模型的标准标定方法来进行验证。一个标准标定算法通常会设置一组 5 个、9 个或者 25 个的标定圆点。最简单的线性标定模型只考虑 5 个标定圆点(即图中的点: 1, 5, 13, 21, 25)。本研究用如下的方程来取得屏幕坐标和测量结果之间的映像:

$$S_x = a_0 + a_2 \cdot x \quad (10)$$

$$S_y = b_0 + b_1 \cdot y \quad (11)$$

式中:  $(S_x, S_y)$  — 屏幕坐标  $(x, y)$  — 瞳孔-眼的角向量,  $a_0, a_1$  和  $b_0, b_1$  — 未知系数并能通过最小二乘法求得。

二阶多项式标定方程需要用到 9 个标定圆点(图中点: 1, 3, 5, 11, 13, 15, 21, 25, 23) 或者 25 个点(图中所有点)。该多项式如下所示:

$$S_x = a_0 + a_1 \cdot x + a_2 \cdot y + a_3 \cdot xy + a_4 \cdot x^2 + a_5 \cdot y^2 \quad (12)$$

$$S_y = b_0 + b_1 \cdot x + b_2 \cdot y + b_3 \cdot xy + b_4 \cdot x^2 + b_5 \cdot y^2 \quad (13)$$

式中:  $a_0, a_1, b_0, b_1$  — 未知系数, 并能通过最小二乘

法求得。

标定完成之后,注视点的坐标可通过插值算法求得。比如,假设屏幕坐标和用于标定的眼角到瞳孔角向量在点集  $P1$  和  $P2$  中是相对独立的,即  $\{(sx_1, sy_1), (x_1, y_1)\}$  和  $\{(sx_2, sy_2), (x_2, y_2)\}$ 。那么在测量眼角到瞳孔的角向量  $(x, y)$  之后,屏幕坐标可由以下方程式求得<sup>[9]</sup>:

$$S_x = S_{x1} + \frac{x - x_1}{x_2 - x_1} (S_{x2} - S_{x1}) \quad (14)$$

$$S_y = S_{y1} + \frac{y - y_1}{y_2 - y_1} (S_{y2} - S_{y1}) \quad (15)$$

标定误差的平均映射误差可由以下公式计算得出:

$$E = \frac{\sum_{i=1}^{25} e_i}{25} \quad (16)$$

$$e_i = \sqrt{(x_{ei} - x_{mi})^2 + (y_{ei} - y_{mi})^2} \quad (17)$$

式中:  $E$ —标定误差;  $i$ —测试点代号;  $e_i$ —单个测试点的映射误差;  $(x_{ei}, y_{ei})$ —测试点的实际坐标;  $(x_{mi}, y_{mi})$ —对应映射瞳孔//眼角向量的坐标。

## 5 实验评估

为了评估系统的准确性,本研究招募了15名测试者来完成最终的测试,各个测试者视力情况均正常。笔者分别测量了每名测试者的注视点,并在水平和垂直两个方向上评估了视线检测误差。为保证实验测试质量,本研究测试环境固定为白天光线充足的室内,避免了光线过于强烈的室外和光线组成较复杂的夜晚。

### 5.1 测试系统的组成

为了验证本研究提出的眼动跟踪系统,本研究设计开发了一个原型系统,其硬件如图7所示。



图7 系统原型

基于树脂镜架的载具方便用户进行佩戴,并且与眼睛保持相对固定的位置,在镜架右眼的位置,本研究运用3D打印技术设计了一个CCD摄像头采集工具,通过镜面反射,可以拍摄到整个眼球区域。CCD摄像头品牌为昌斯特,最高分辨率为1200\*800,图像采样频率为30帧/秒,焦距为0~10cm,数据输出口为USB2.0,可与电脑直接连接。整个系统轻便体积小,不会给用户带来大的干扰。



图8 标定测试示例

### 5.2 试验结果和评估

本研究分别测量了每名测试者的注视点,并在水平和垂直两个方向上评估了视线检测误差。运用上文提出的标定算法后,本研究得到了其中一名测试者的初始视线数据和标定后的矫正数据,如图9和图10所示。

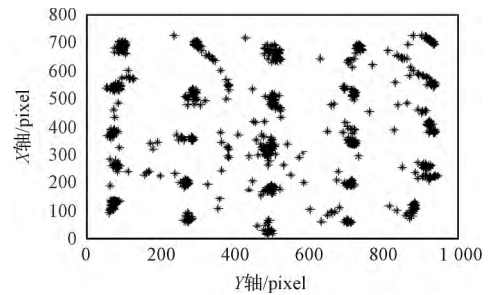


图9 标定前的视线点数据

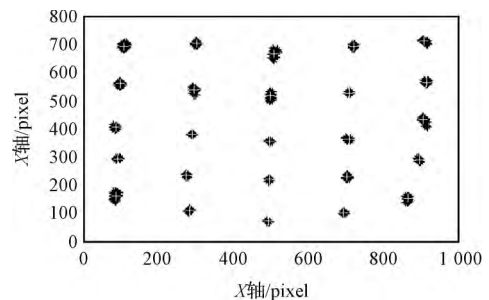


图10 标定后的视线点数据

该测试者的映射测试数据如图11所示。

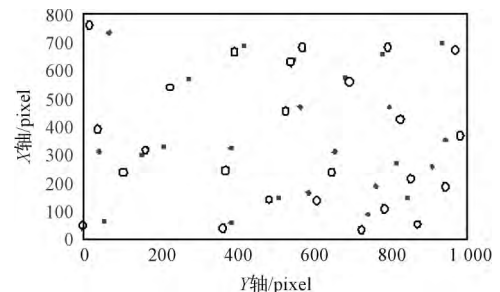


图11 测试点映射数据

空心圆点—实际的测试点; 空心圆点—视线数据

### 5.3 视线点的检测误差

本研究将15名测试者随机分成5组,测试结果取每组五名测试者的平均值,表1和表2即为5组测试者的视线检测误差和垂直与水平方向的眼睛注视检测的结果。研究人员可以明显地看出,水平方向的检测误差

明显小于垂直方向的检测误差。在实际测试过程中, A, B, D 和 E 这 4 组测试者的数据显示, 垂直方向检测的误差超过了 30(像素), 大于水平方向检测的误差。

表 1 水平方向的检测误差(单位: pixel)

	TesterA	TesterB	TesterC	TesterD	TesterE
1st test	8.7	3.3	21.6	9.9	15.9
2st test	24.6	15.9	15.6	18.6	26.4
3st test	2.7	20.1	24.3	39.9	10.5
4st test	5.7	22.8	36.9	11.1	9.6
5st test	15	0.3	9.9	24.6	23.4
Average	11.4	12.6	21.6	20.7	17.1
variance	8.7	10.2	10.2	12.3	7.5

表 2 垂直方向的检测误差(单位: pixel)

	TesterA	TesterB	TesterC	TesterD	TesterE
1st test	75.9	6	18.6	45.9	61.8
2st test	11.4	83.7	18	18.9	31.8
3st test	52.8	44.4	7.5	41.1	18.3
4st test	0.9	14.7	19.5	1.2	60.6
5st test	2.1	4.8	39.3	87.9	53.4
Average	28.5	30.6	20.7	39	45.3
variance	33.9	33.6	11.7	32.7	19.2

对 5 组测试人员的测试数据进行对比分析发现, 水平方向误差明显大于垂直方向的误差, 而同一测试者的每次测试数据也有不同程度的差异, 但是在合理范围内。原因总结如下: 在日常生活中, 瞳孔(眼球)的水平方向运动范围明显大于垂直方向的运动范围, 造成水平方向的测试精度大于垂直方向的测试精度。

本研究中 CCD 摄像头的分辨率设置在 640\*480 的中等分辨率以增加系统的检测跟踪速度, 图像跟踪帧率达到 12 fps, 基本满足了系统速度的要求。误差除以摄像头水平与垂直方向的总像素后即得到系统的水平方向视线检测误差和垂直方向视线检测误差, 分别是 3.74% 和 5.34%, 具有较高的精度, 并且与其他的单摄像头视线检测方法有同等级别的表现。考虑到本研究提出的系统硬件配置简单, 便于携带, 具有更高的性价比。本研究提出的方法与其他眼动跟踪方法(VE<sup>[10]</sup>、IP<sup>[5]</sup>)的精度对比如表 3 所示。

表 3 与其他眼动跟踪方法的对比

方法	平均精度	技术特征
VE	5.2%	垂直矢量法
IP	4.59%	基础等直线法
本研究	4.54%	改进的椭圆拟合方法

本文引用格式:

杨庆华, 张达磊, 荀一, 等. 面向人机交互的眼动跟踪方法研究[J]. 机电工程, 2016, 33(7): 904-908.

YANG Qing-hua, ZHANG Da-lei, XUN Yi, et al. Eye-gaze Tracking Method Designed for HCI System[J]. Journal of Mechanical & Electrical Engineering, 2016, 33(7): 904-908.

《机电工程》杂志: <http://www.meem.com.cn>

## 6 结束语

本研究提出基于梯度值和圆特性的改进椭圆拟合算法进行对眼球瞳孔中心进行检测跟踪, 通过基于二阶模型的标定对注视点进行标定。实验测试结果表明, 该系统在水平与垂直方向的检测误差分别为 3.74% 和 5.34%, 精度较高, 满足了人机交互系统功能性的需求。同时本研究运用 3D 打印技术辅助设计了一套便携移动式眼动跟踪装置, 通过个人电脑和普通 CCD 摄像头在自然光下获取眼动图像, 结构简单、便于佩戴, 有较大应用前景。

在下一步的工作中, 笔者将改进眼动跟踪算法, 提高检测精度, 完善系统硬件, 将这一系统投入真正实际的眼动输入应用中。

### 参考文献(References):

- [1] AUN K A H, JEFF P, JUAN C. Lightweight, low-cost, side-mounted mobile eye tracking system[C]. Image processing workshop eWNYIPN, New York: IEEE, 2012: 1-4.
- [2] DUCHOWSKI A T. Eye tracking methodology: theory and practice[M]. London: Springer, 2003.
- [3] LAURENCE R, YOUNG, DAVID S. Survey of eye movement recording methods[J]. *Behavior Res. Methods Instrument*, 1975, 7(5): 397-429.
- [4] JOHN D. How iris recognition works[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2004, 14(1): 21-30.
- [5] VALENTI, R. GEVERS. T. Accurate eye center location through invariant isocentric patterns[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2012, 34(9): 1785-1798.
- [6] TAKEGAMI T, GOTOH T, 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.
- [7] XIANG L, JENNIFER C, SIMON D et al. Accurately measuring the size of the pupil of the eye[D]. Auckland: University of Auckland, 2003.
- [8] OTSU N. A threshold selection method from gray-level histogram[J]. *Systems Man & Cybernetics*, 1979, 9(1): 62-66.
- [9] MEHRUBE M, LINH M P, HUNG T L, et al. Real-time eye tracking using a smart camera[J]. *Applied Imagery Pattern Recognition Workshop (AIPR)*, 2011(10): 1-7.
- [10] WANG J G, SUNG E, VENKATEWARLU R. "Estimating the eye gaze from one eye," [J]. *Computer Visual & Image Understanding*, 2005, 98(1): 83-103.

[编辑: 周昱晨]