

# 激光与光电子学进展

## 基于视线水平偏移特征的眼动向量计算方法

田会娟<sup>1,2\*</sup>, 王思琦<sup>1,2</sup>

<sup>1</sup>天津工业大学电气与电子工程学院天津市光电检测技术与系统重点实验室, 天津 300387;

<sup>2</sup>天津工业大学大功率半导体照明应用系统教育部工程研究中心, 天津 300387

**摘要** 视线追踪是一种人机交互技术,在学术界和工业界倍受关注。针对基于多项式回归的非侵入式系统注视点的精度平均误差较大的问题,提出一种基于视线水平偏移特征的眼动向量计算方法。该方法在准确提取人脸图像上虹膜中心和内眼角坐标的基础上,以内眼角坐标为基准点计算虹膜中心相对于内眼角坐标的偏移量,再通过偏移量计算水平眼动向量的权值,然后利用权值对左右眼的水平眼动向量分别加权合成新的水平眼动向量。采用哥伦比亚注视数据集的 20 名受试人员头部水平角度为 0° 的数据,在 4 个回归模型上对该方法进行实验验证。结果表明,与双眼平均眼动向量计算方法相比,该方法能进一步减小注视点水平视角误差、提升系统精度。

**关键词** 视觉光学; 眼动向量; 视线追踪; 多项式回归; 虹膜中心

中图分类号 TP391 文献标志码 A

doi: 10.3788/LOP202158.2333001

## Calculation Method of Eye Vectors Based on Line-of-Sight Horizontal Offset Feature

Tian Huijuan<sup>1,2\*</sup>, Wang Siqi<sup>1,2</sup>

<sup>1</sup>Tianjin Key Laboratory of Optoelectronic Detection and System, School of Electrical and Electronic Engineering, Tiangong University, Tianjin, 300387, China;

<sup>2</sup>Engineering Research Center of Ministry of Education on High Power Solid State Lighting Application System, Tiangong University, Tianjin, 300387, China

**Abstract** Gaze tracking is a human-computer interaction technology, which has received extensive attention in academia and industry. Aiming at the problem of the large average error of the gaze points accuracy in non-invasive systems based on polynomial regression, a method of eye vectors calculation based on line-of-sight horizontal offset feature is proposed. On the basis of accurately extracting the coordinates of the iris center and inner corner of the eye on the face image, this method uses the inner corner coordinates as the reference point to calculate the offset of the iris center relative to the inner corner coordinates, calculates the weight of the horizontal eye movement vector through the offset, and then uses the weight to weight the horizontal eye movement vectors of the left and right eyes to synthesize new horizontal eye movement vectors. Using the data of 20 subjects with the head horizontal angle 0° from the Columbia Gaze database, this method was experimentally verified on 4 regression models. The results show that compared with the calculation method of binocular average eye vectors, this method can further reduce the horizontal of viewing angle error of the gaze points and improve the system accuracy.

**Key words** visual optics; eye vectors; gaze tracking; polynomial regression; iris center

**OCIS codes** 330.2210; 150.0150; 080.1753; 100.0100

收稿日期: 2021-03-29; 修回日期: 2021-05-10; 录用日期: 2021-07-13

基金项目: 国家自然科学基金(61504095)、天津市科技计划项目(18ZXCLGX00090)

通信作者: \*tianhjx@126.com

## 1 引言

注视估计是通过电子与光学设备等采集眼球运动所对应的人眼特征图像信息,并利用分析算法估计参与者注视方向或注视点的技术,它是观察和分析人类视觉注意的一个重要课题<sup>[1-4]</sup>。目前该技术已广泛应用于各个领域,例如人机交互、视觉行为分析和心理学研究。近年来,基于单个普通相机采集图像信息建立注视估计系统的方法成为当下研究的热点<sup>[5-7]</sup>。

该方法采用多项式回归函数拟合从注视特征图像提取的眼动特征与屏幕注视点坐标的映射关系,其系统精度主要受到眼动特征点选取、眼动向量计算和回归函数构造的影响。1) 针对眼动特征点的选取, Sesma 等<sup>[8]</sup>对瞳孔(虹膜)中心眼角向量作为基于多项式回归方法的眼动特征进行了研究,通过实验发现,即使受试人员的头部处于静止状态,当注视屏幕的不同位置时,眼角位置也会轻微移动,取内外眼角特征点作为眼动向量的参考点会使精度降低到 $2^\circ \sim 3^\circ$ 。Xiao 等<sup>[9]</sup>提出了一种在自然光下的单摄像头注视跟踪系统,该系统选取平均人脸特征点作为眼动向量的参考点。考虑到人脸的微表情变化,如嘴角抽动、脸部轻微转动等,选取平均人脸特征点作为参考点,其眼动向量的稳定性会受到一定的影响。2) 针对眼动向量的计算, Cui 等<sup>[10]</sup>研究表明相对于根据主视眼理论的单眼眼动向量法,根据中间点假说的双眼眼动向量法得到的精度平均误差值更小。Xiao 等<sup>[9]</sup>综合考虑了主视眼理论和中间点假说,在 EyeDiap 数据集上对双眼眼动向量的权值选取做了实验评估,经过实验数据对比,在该数据集上选取 0.5 作为权值,得到的双眼眼动向量平均值输入系统能够得到最优的注视估计结果。3) 针对回归函数的构造, Hennessey 等<sup>[11]</sup>提出两个对称的一次多项式回归函数。然而,眼动向量与注视点坐标实际呈非线性关系,一阶多项式回归函数无法表示这种映射关系,考虑到非线性补偿,许多研究者在此基础上增加了多次项<sup>[12-14]</sup>。Cerrolaza 等<sup>[15]</sup>实验发现注视估计的结果没有随着多项式项数的增加与阶数的增长而变得更准确。

基于以上研究分析,针对单个普通相机和多项式回归的注视估计系统的精度问题,本文选取眼动时脸部相对稳定的鼻子平均特征点作为眼动向量的参考点、高精度的回归函数作为眼动特征到注视

点位置的映射关系,提出了一种双眼的水平眼动向量( $\mathbf{x}$ )的权值计算方法,用于构建基于视线水平偏移特征的眼动向量( $\mathbf{x}_w$ ),并利用哥伦比亚注视数据集对该方法进行实验验证。

## 2 基于多项式回归的注视估计原理

### 2.1 眼动参考点的选取

眼动向量是由与眼动相关的动点和与眼动无关的参考点计算得到。选取虹膜中心特征点作为眼动向量的动点,面部稳定的特征点作为眼动向量的参考点,由此计算的眼动向量可以准确地表征眼睛的注视方向。本文针对每张注视特征图像,选取该图像的鼻子平均特征点作为其眼动向量的参考点( $u_r, v_r$ ),

$$\begin{cases} u_r = \frac{\sum_{m=a}^{m=b} u_m}{b-a+1}, \\ v_r = \frac{\sum_{m=a}^{m=b} v_m}{b-a+1} \end{cases}, \quad (1)$$

式中: $u_r$ 和 $v_r$ 为眼动向量的参考点在图像的 $u$ - $v$ 像素坐标系中 $u$ 和 $v$ 的坐标值; $u_m$ 和 $v_m$ 为第 $m$ 个人脸特征点在图像的 $u$ - $v$ 像素坐标系中 $u$ 和 $v$ 的坐标值; $a$ 为鼻子特征点的起始编号; $b$ 为鼻子特征点的终点编号。 $(u_a, v_a), \dots, (u_b, v_b)$ 均为鼻子特征点坐标, $(u_a, v_a)$ 为 $b-a+1$ 个鼻子特征点的起始特征点坐标, $(u_b, v_b)$ 为鼻子的最后一个特征点坐标。

### 2.2 平均眼动向量的计算

图 1 为左右眼的眼动向量示意图,可以表示为

$$\begin{cases} \mathbf{x}_L = u_r - u_{lc} \\ \mathbf{y}_L = v_r - v_{lc} \\ \mathbf{x}_R = u_r - u_{rc} \\ \mathbf{y}_R = v_r - v_{rc} \end{cases}, \quad (2)$$

式中: $\mathbf{x}_L$ 和 $\mathbf{y}_L$ 分别为左眼的水平和垂直眼动向量; $\mathbf{x}_R$ 和 $\mathbf{y}_R$ 分别为右眼的水平和垂直眼动向量; $(u_{lc}, v_{lc})$ 和 $(u_{rc}, v_{rc})$ 分别为左眼和右眼的虹膜中心点在图像的像素坐标系中的坐标。

Sesma-Sanchez 等<sup>[16]</sup>曾将单眼注视估计方法扩展为结合双眼数据完成注视估计的方法,与使用单眼数据参与注视估计相比,使用双眼数据计算平均注视估计点的这种方法能达到更高的精度。同时,使用双眼平均眼动向量替代单眼眼动向量,得到的注视估计结果更精确<sup>[9]</sup>。双眼平均眼动向量可表示为

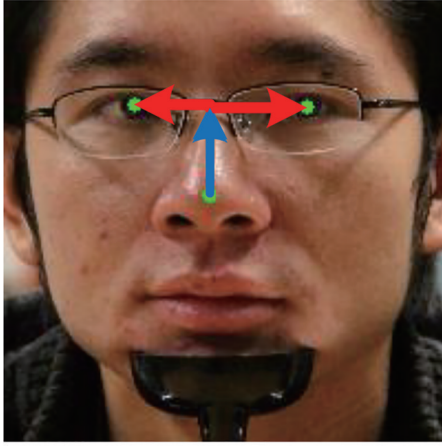


图1 眼动向量示意图

Fig. 1 Schematic of eye vectors

$$\begin{cases} \mathbf{x}_a = 0.5 \times (\mathbf{x}_L + \mathbf{x}_R) \\ \mathbf{y}_a = 0.5 \times (\mathbf{y}_L + \mathbf{y}_R) \end{cases}, \quad (3)$$

式中： $\mathbf{x}_a$ 和 $\mathbf{y}_a$ 分别为双眼水平和垂直眼动向量的平均值。

### 2.3 多项式回归函数

在注视估计系统中,多项式回归函数将 $\mathbf{x}$ 和垂直眼动向量( $\mathbf{y}$ )作为输入,输出视线在屏幕落点的水平坐标值( $X$ )和垂直坐标值( $Y$ )。高性能的回归函数可以提供输入到输出最合适的映射关系,具有平均绝对误差小、精度高,误差值集中低误差区、稳定性好的特性。特定注视估计场景下的最佳多项式映射关系,需要通过实验检验模型做最优选择得到。通用的多项式回归函数表达式为

$$\begin{cases} X = \sum_{k=0}^{n-1} a_k \mathbf{x}^i \mathbf{y}^j, \\ Y = \sum_{k=0}^{n-1} b_k \mathbf{x}^i \mathbf{y}^j, \end{cases} \quad (4)$$

式中： $j$ 的取值范围为 $[0, k]$ ； $k$ 为函数各项系数的编号； $n$ 为标定点的数量； $a_k$ 和 $b_k$ 分别为 $X$ 和 $Y$ 的回归函数的系数； $\mathbf{x}^i$ 表示 $i$ 个 $\mathbf{x}$ 的数量积； $\mathbf{y}^j$ 表示 $j$ 个 $\mathbf{y}$ 的数量积。标定过程是用户在使用基于多项式回归的注视估计系统之前需要完成的,要求用户依次观看视点板上的标定集注视点,系统同步从采集到的注视特征图像中提取眼动特征,计算实时眼动向量,根据已知的眼动向量和与之对应的标定坐标,通过最小二乘法计算得到 $a_k$ 和 $b_k$ 。

## 3 $\mathbf{x}_w$ 计算方法

根据研究,当水平注视角度偏离双目视野中心

11°时,眼优势会发生改变,即左侧视野目标倾向于左眼作为主视眼,右侧视野目标倾向于右眼作为主视眼,而中间视野目标倾向于双目视觉<sup>[17]</sup>。在记录眼球运动、分析注视方向时,这种眼优势的变化可能会影响基于单眼或双眼追踪注视方向结果的准确性<sup>[10]</sup>。基于对该问题的研究,Elbaum等<sup>[18]</sup>通过实验对比了使用双眼、主视眼、非主视眼的眼动数据对注视估计精度的影响,证明了双眼精度优于主视眼或非主视眼单独预测的精度结论,同时也提出交替采用双眼数据与优势眼数据跟踪注视的结果更精确。

对 $\mathbf{x}_L$ 和 $\mathbf{x}_R$ 取权值为0.5得到 $\mathbf{x}_a$ 和 $\mathbf{y}_a$ 的计算方法仅考虑双目视觉,而未考虑实际用眼情况有侧重地提取优势眼的眼动特征,以达到更高精度的注视估计效果。基于双目视觉和主视眼随着注视视线水平角度的变化而切换的用眼规律,本文提出了 $\mathbf{x}_w$ 的计算方法,该方法利用权值对优势眼的眼动特征有侧重地选取,能够得到更加准确的特征向量。

设 $C$ 为内眼角点在图像的 $u-v$ 像素坐标系中的坐标值, $P$ 为虹膜中心点在图像的 $u-v$ 像素坐标系中 $u$ 的坐标值, $d = |C - P|$ 为内眼角点 $C$ 与 $P$ 的水平欧氏距离, $d_L$ 表示左眼, $d_R$ 表示右眼。 $d_L$ 和 $d_R$ 如图2所示。

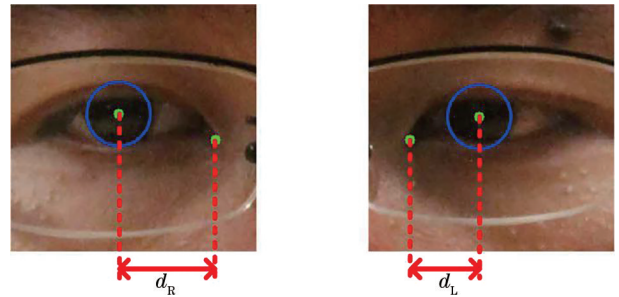


图2  $d_L$ 和 $d_R$ 示意图

Fig. 2 Schematic of  $d_L$  and  $d_R$

对 $\mathbf{x}_L$ 加权值 $w_L = d_L / (d_L + d_R)$ ,对 $\mathbf{x}_R$ 加权值 $w_R = d_R / (d_L + d_R)$ ,权值表明了单眼的水平偏移程度,则 $\mathbf{x}_w$ 可以描述为

$$\begin{aligned} \mathbf{x}_w &= w_L \times \mathbf{x}_L + w_R \times \mathbf{x}_R = \\ &= \frac{d_L}{d_L + d_R} \times \mathbf{x}_L + \frac{d_R}{d_L + d_R} \times \mathbf{x}_R, \end{aligned} \quad (5)$$

同时,对于 $\mathbf{y}$ 的计算按照 $\mathbf{y}_a$ 的计算方法。

上述加权方法考虑了双目视觉和优势眼切换的情况,当视线方向在视野中心附近时,对注视目标的成像是双目视觉,各单眼的瞳孔中心距离内眼

角的偏移量基本相等,有  $d_L \approx d_R, w_L \approx w_R \approx 0.5$ , 合成的  $\mathbf{x}$  是对  $\mathbf{x}_L$  和  $\mathbf{x}_R$  取均值的结果。当视线方向向视野中心左侧偏移时,  $d_L > d_R, w_L > 0.5 > w_R$ , 随偏移视角增大,左眼优势百分比也增大,双目视觉逐渐转为左眼作为主视眼,作为量化用眼情况的偏移度值  $w_L$  和  $w_R$  相应表现为  $w_L$  增大、 $w_R$  减小,其合成的  $\mathbf{x}$  中左眼的水平眼动特征向量比重更大。当视线方向向视野中心右侧偏移时,  $d_L < d_R, w_L < 0.5 < w_R$ , 随偏移视角增大,右眼优势百分比也增大,双目视觉逐渐转为右眼作为主视眼,作为量化用眼情况的偏移度值的  $w_L$  和  $w_R$  相应表现为  $w_R$  增大、 $w_L$  减小,其合成的  $\mathbf{x}$  中右眼的水平眼动特征向量比重更大。利用  $w_L$  和  $w_R$  分别对  $\mathbf{x}_L$  和  $\mathbf{x}_R$  加权,最后合成的眼动向量  $\mathbf{x}_w = w_L \times \mathbf{x}_L + w_R \times \mathbf{x}_R$  能有效表示人眼的水平注视方向特征。

## 4 实验及结果分析

### 4.1 实验设置

选取哥伦比亚注视数据集<sup>[19]</sup>中头部水平角度为  $0^\circ$  的 56 名受试人员注视特征图像进行虹膜分割实验。由于人眼虹膜分割模型<sup>[20]</sup>可能受到眼睑遮挡、眼部区域亮度不均匀的影响<sup>[21]</sup>,导致虹膜分割

不全、虹膜中心定位出错。因这种局限性,实验结果显示部分受试人员的注视特征图像虹膜数据不完整,图 3 为虹膜分割失败图例,圆圈标注处为通过模型分割处理的虹膜位置。为确保参与本实验的受试人员的虹膜数据完整性和准确性,对参与虹膜分割实验的受试人员进行了筛选。最后保留了通过虹膜分割模型能得到完整且准确虹膜中心位置的 20 名受试人员的数据,即 420 张  $5184 \text{ pixel} \times 3456 \text{ pixel}$  的注视特征图像。每个受试人员的 21 张图像对应垂直距离为 250 cm 的视点板上的 21 个注视点,这些注视点分别在  $0^\circ, \pm 5^\circ, \pm 10^\circ, \pm 15^\circ$  的水平注视方向和  $0^\circ, \pm 10^\circ$  的垂直注视方向上。图 4 为某一受试人员在不同注视方向的左眼眼睛外观和成功的虹膜分割标注。

为验证该方法的有效性,构建了多项式回归的注视估计系统,其系统框图如图 5 所示,包括个人标定阶段和测试阶段。在个人标定阶段,使用受试人员的 9 张注视特征图像作为标定图像(图像对应的水平注视方向分别为  $0^\circ, \pm 15^\circ$ ,垂直注视方向分别为  $0^\circ, \pm 10^\circ$ ),提取标定图像与注视相关的眼动特征,并通过多项式回归函数,构建眼动特征与在屏幕的视线落点坐标[以视角角度( $^\circ$ )为单位]之间的

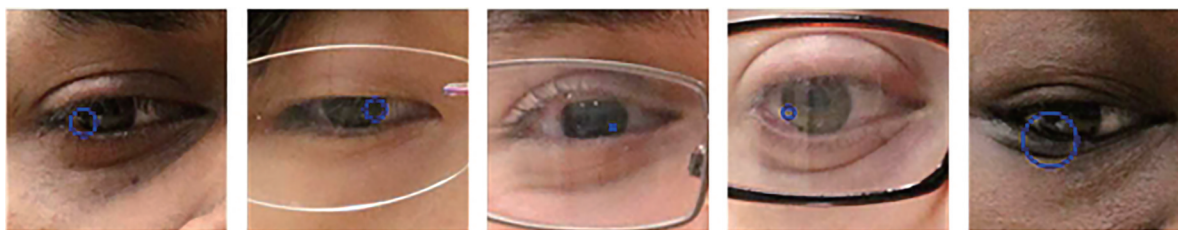


图 3 虹膜分割失败图例

Fig. 3 Failure illustrations of iris segmentation

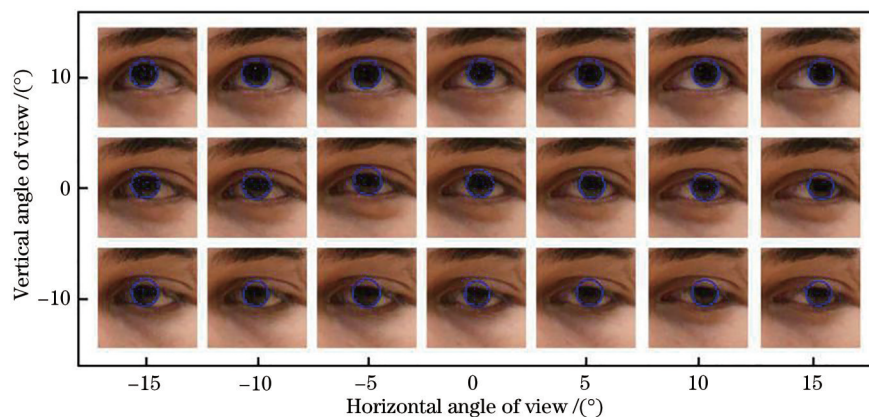


图 4 不同注视方向的左眼眼睛外观及虹膜分割

Fig. 4 Left eye appearance and iris segmentation with different gaze directions

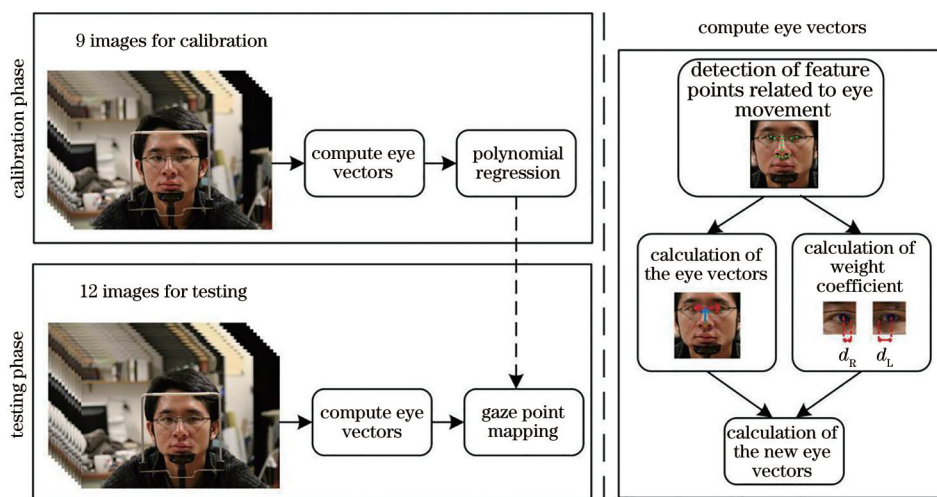


图 5 注视估计系统

Fig. 5 Gaze estimation system

映射关系。在测试阶段,使用受试人员的剩余 12 张注视特征图像作为测试图像,提取测试图像的眼动特征并计算眼动向量,作为标定阶段映射关系的输入,输出视线在屏幕落点的坐标[以视角角度(°)为单位]。系统采用“shape\_predictor\_68\_face\_landmarks.dat”68 个标记点的 Dlib 官方人脸识别模型和基于图像语义分割网络(U-net)的人眼虹膜分割模型<sup>[19]</sup>提取图像与注视相关的眼动特征。先使用人脸特征点检测模型对图像进行特征点检测,得到眼睛特征点(左右眼特征点均为 6 个,不包括虹膜中心)、鼻子特征点(9 个)的图像坐标信息。确定每只眼睛特征点的最小包围矩形,根据该矩形中心,裁剪为高宽均为该只眼睛内外眼角点欧氏距离的 1.5 倍值的眼睛区域。再使用人眼虹膜分割模型完成对归一化尺寸的眼睛区域图像的虹膜分割,并根据椭圆拟合方法提取虹膜轮廓,确定虹膜中心点位置。基于该注视估计系统框架,本文提出的眼动向量计算方法和平均眼动向量法分别在 4 个回归模型中进行实验结果对比,表 1 为 4 个回归模型。

表 1 4 个模型

Table 1 4 models

Model	Polynomial for X	Polynomial for Y
1	$1, x, x^2, y, y^2, xy, x^2y^2$	$1, x, x^2, y, y^2, xy, x^2y^2$
2	$1, x, x^2, y$	$1, x^2, y, xy, x^2y$
3	$1, x, x^3, y^2, xy$	$1, x, x^2, y, y^2, xy, x^2y$
4	$1, x, x^2, x^3, y, xy, x^2y, x^3y$	$1, x, x^2, y, y^2, xy, x^2y$

#### 4.2 精度评估方法

本文采用总体测试点的注视视线视角误差的平均值评估精度。受试人员的估计视线和实际视

线的夹角值为注视视线视角误差,如图 6 所示, $\alpha$  是估计视线与实际视线的注视视线视角误差。

注视视线视角误差表示为

$$\begin{cases} \alpha(\theta) = \arccos \frac{D_1^2 + D_2^2 - d_{error}^2}{2 \times D_1 \times D_2} \\ D_1 = \sqrt{d_1^2 + L^2} \\ D_2 = \sqrt{d_2^2 + L^2} \\ d_{error} = \sqrt{(X_{Gaze} - X_{GT})^2 + (Y_{Gaze} - Y_{GT})^2} \end{cases}, (6)$$

式中: $\alpha$  为估计视线与实际视线的夹角值; $D_1$  为实际注视点( $X_{GT}, Y_{GT}$ )到受试人员的距离; $D_2$  为估计注视点( $X_{Gaze}, Y_{Gaze}$ )到受试人员的距离; $d_{error}$  为估计注视点( $X_{Gaze}, Y_{Gaze}$ )与实际注视点( $X_{GT}, Y_{GT}$ )的欧氏距离; $X_{Gaze}$  和  $Y_{Gaze}$  为水平视角、垂直视角的估计视线映射在视点板上估计注视点的坐标值; $X_{GT}$  和  $Y_{GT}$  为水平视角、垂直视角的实际视线映射在视点板上真实注视点的坐标值。

精度表示为

$$A_{accuracy} = \frac{1}{N} \sum_{i=1}^N \alpha_i, (7)$$

式中: $\alpha_i$  为第  $i$  个测试注视点的注视视线视角误差,共  $N$  个测试注视点。

本文采用误差降低(精度提高)百分比说明提出方法的有效性,其计算公式为

$$E_{reduction} = \frac{|P_{proposed} - O_{original}|}{O_{original}} \times 100\%, (8)$$

式中: $P_{proposed}$  为使用本文提出的眼动向量计算方法得到的估计视线水平视角(垂直视角); $O_{original}$  为使用平均眼动向量方法得到的估计视线水平视角(垂直视

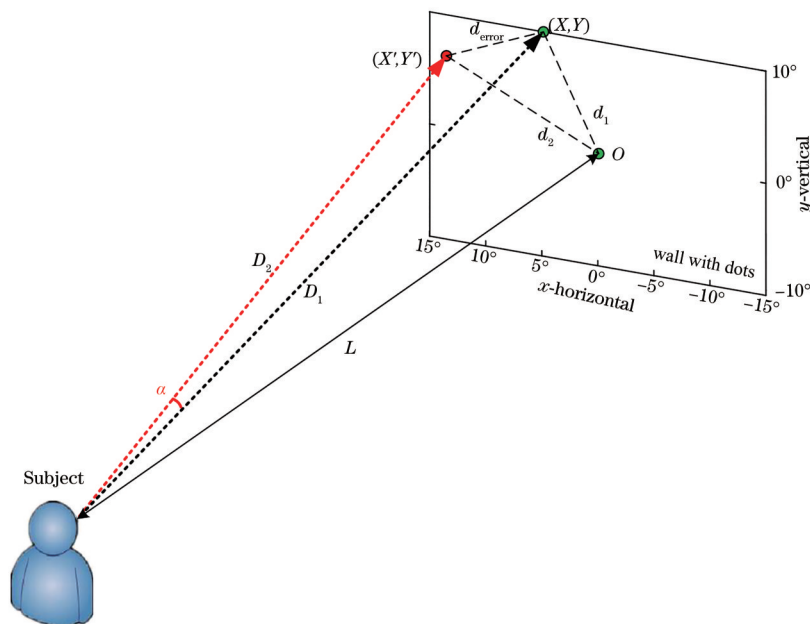


图 6 估计视线与实际视线的视角误差

Fig. 6 Viewing angle error between the estimated line of sight and the actual line of sight

角);  $E_{reduction}$  为水平视角(垂直视角)误差降低百分比。

$$I_{accuracy} = \frac{|A_{proposed} - A_{original}|}{A_{original}} \times 100\%, \quad (9)$$

式中:  $A_{proposed}$  为使用本文提出的眼动向量计算方法得到的精度;  $A_{original}$  为使用平均眼动向量方法得到的精度;  $I_{accuracy}$  为精度提升百分比。

### 4.3 结果分析

为验证本文提出的眼动向量计算方法的有效性,将对从哥伦比亚注视数据集中选取的受试人员数据进行实验,分别利用 4 个多项式回归函数构建眼动向量到注视点坐标的映射关系,对比采用本文提出方法和平均眼动向量法在每种回归模型的水平、垂直视角误差和精度值,同时横向对比该 4 个回归模型的注视估计性能,从水平、垂直视角以及注视视线视角的平均绝对误差值和误差分布直方图进行对比分析。

先从本文提出的方法和平均眼动向量方法计算的  $x$  数据分布对比来说明本文方法的优势。以一位受试人员的注视特征图像数据为例进行说明,将图像按照水平视角信息分类,分别计算每个水平视角下所有注视特征图像的两种  $x$  的平均值和标准差,得到如图 7 所示的眼动向量数据分布图。经过对比发现,不同水平视角下基于视线水平偏移特征方法计算的  $x$  的差值更大,视线的水平偏移特征区别更明显。

表 2 给出了 4 个回归模型采用两种眼动向量计

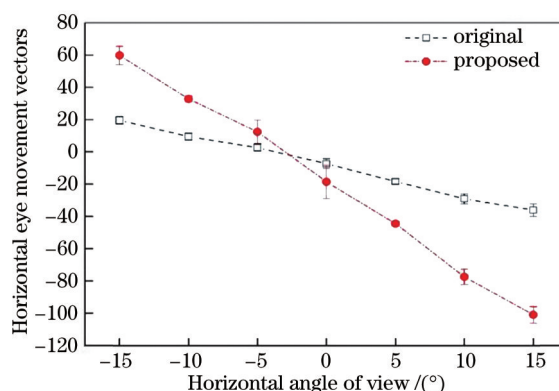


图 7 不同视角下的  $x$

Fig. 7  $x$  under different viewing angles

算方法的水平和垂直视角平均绝对误差结果对比,其中误差降低百分比通过(8)式计算得到。从表中可以看出,本文眼动向量计算方法能在一定程度上降低估计注视点的水平视角误差,在 4 个回归模型上,误差分别降低了 34.50%、17.39%、10.88%、21.67%,其中在回归模型 1 上,误差降低百分比值最大。同时,本文算法在一定程度上增加了垂直视角误差。通过对注视视角误差的结果分析,垂直视角误差的增加在一定程度上降低了精度提升百分比。表 3 给出了不同回归模型的两方法精度结果,其中精度值、精度提升百分比分别通过(7)式、(9)式计算得到,本文算法在 4 个回归模型上,精度分别提升了 8.16%、4.79%、7.96%、18.90%。

表 2 不同模型的水平视角平均绝对误差对比

Table 2 Comparison of the average absolute error of horizontal and vertical viewing angles of different models

Model	Horizontal viewing angle error / (°)		Error reduction percentage / %	Vertical viewing angle error / (°)		Error reduction percentage / %
	Original	Proposed		Original	Proposed	
1	2.22	1.45	34.50	3.29	3.43	-4.29
2	1.69	1.40	17.39	2.93	2.96	-0.96
3	1.71	1.53	10.88	3.25	3.37	-3.86
4	4.73	3.70	21.67	3.25	3.37	-3.86

表 3 不同模型的精度对比

Table 3 Accuracy comparison of different models

Model	Error / (°)		Percentage improvement in accuracy / %
	Original	Proposed	
1	3.14	2.89	8.16
2	2.93	2.79	4.79
3	3.18	2.92	7.96
4	5.26	4.26	18.90

针对 420 个测试注视点样本的水平视角误差、注视视角误差进行统计分析。图 8 所示为 4 个回归模型的水平视角误差分布直方图,本文眼动向量法和平均眼动向量法在每个回归模型中表现的数据分布类似,但本文方法的样本主要集中在低误差区,在水平视角的准确度上表现更好。通过 4 个模

型的统计直方图比较,模型 4 的大误差数据在 10°~40°不等分布,模型 1 存在少量数据分布在 30°~35°,模型 2、3 的数据量主要集中在 0°~10°;模型 4 的水平视角误差 0°~2°的数据量在 85 左右,低误差数据量最少,模型 2 的水平视角误差 0°~2°的数据量在 180 左右,低误差数据量居多。

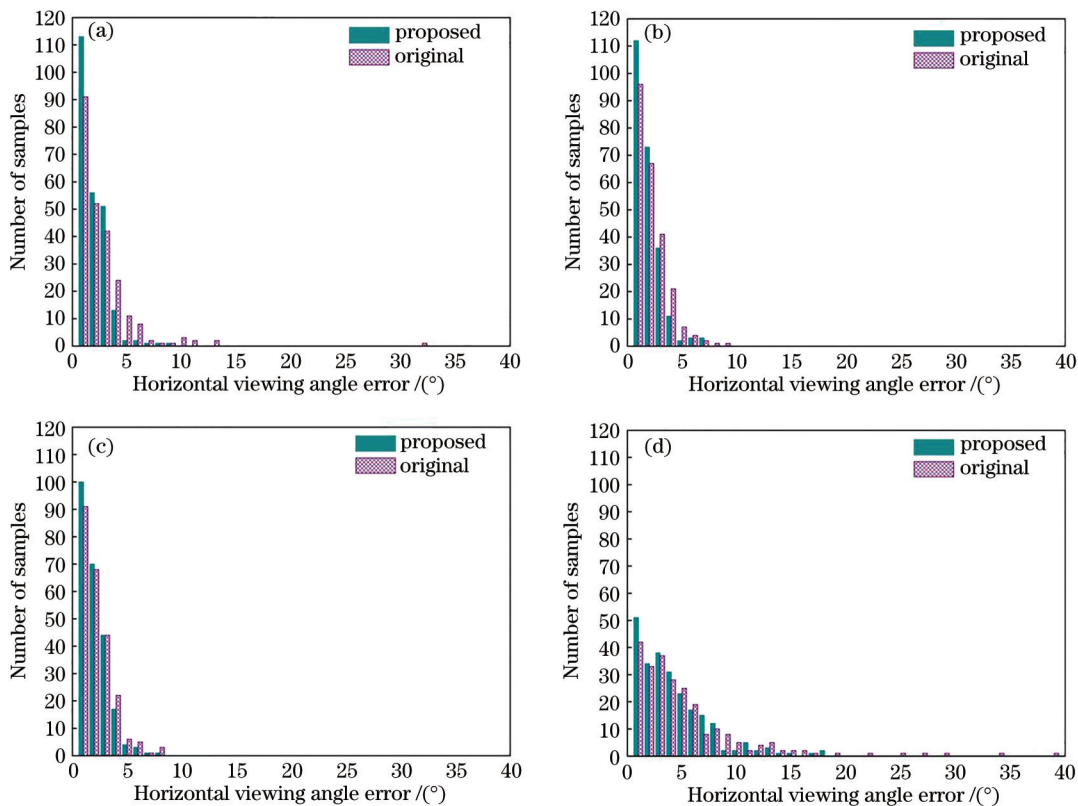


图 8 不同模型的水平视角误差分布直方图。(a)模型 1;(b)模型 2;(c)模型 3;(d)模型 4

Fig. 8 Histograms of horizontal viewing angle error distribution of different models. (a) Model 1; (b) model 2; (c) model 3; (d) model 4

图 9 为 4 个回归模型的注视视角误差分布直方图。模型 2 的注视视角误差数据量在  $0^{\circ}\sim 18^{\circ}$  不等分

布,且主要集中在  $0^{\circ}\sim 3^{\circ}$ ,对比其他 3 个回归模型,测试点的注视估计性能明显优于其他回归模型。

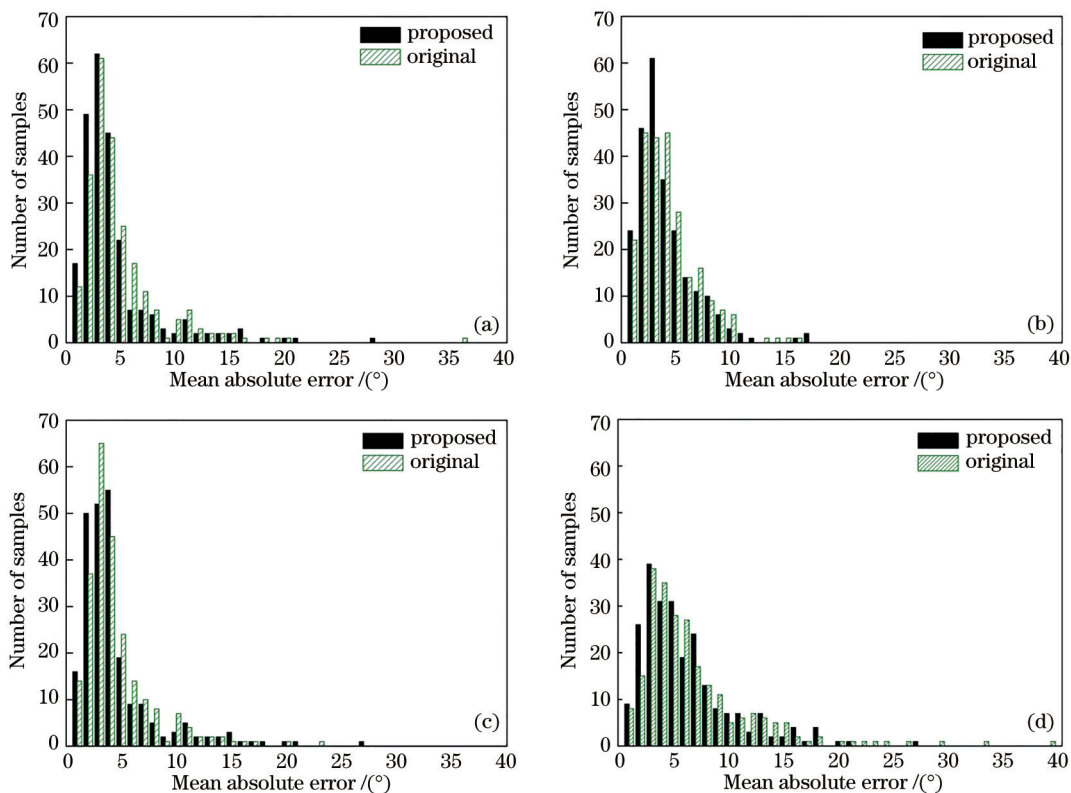


图 9 不同模型的注视视角误差分布直方图。(a)模型 1;(b)模型 2;(c)模型 3;(d)模型 4

Fig. 9 Histograms of gaze angle error distribution of different models. (a) Model 1; (b) model 2; (c) model 3; (d) model 4

## 5 结 论

本文提出了  $x_w$  的计算方法。该方法通过计算单眼的虹膜中心点相对于内眼角特征点的水平偏移量在双眼水平偏移量的比值,量化单眼的水平偏移程度,在此基础上,分别对  $x_L$  和  $x_R$  加权,最后合成的眼动向量能更好地表示人眼的水平注视方向特征。采用哥伦比亚注视数据集的 20 名受试人员头部水平角度为  $0^{\circ}$  的数据,在 4 个注视估计回归模型上评估和对比了  $x_w$ 、双眼平均水平眼动向量计算方法的精度。结果表明,本文眼动向量计算方法相较于双眼平均眼动向量法,在 4 个回归模型上(按照 Model 1~4 的顺序),水平视角误差分别降低了 34.50%、17.39%、10.88%、21.67%,精度分别提升了 8.16%、4.79%、7.96%、18.90%。通过水平视角误差分布直方图、注视视角误差分布直方图对比了 4 个回归模型的注视预测表现,其中模型 2 误差分布集中,稳定性好,低误差值样本量大,性能最优。

本方法根据双目视觉和主视眼切换的用眼规律,有侧重地提取优势眼注视特征,既利用了双目注视特征数据,又同时结合了主视眼注视特征数据,在一定程度上降低了注视点的水平坐标误差、提升了精度。由于头部运动直接影响注视点精度,基于单个普通相机和多项式回归的注视估计系统的精度不高。实际的应用场景中,头部运动无时不在,要获得精度水平良好的注视估计系统必须要克服头部运动对精度影响。因此,提高本系统对于头部运动的抗干扰性仍需继续探索,以推动低成本、精度良好的注视估计系统的发展和应用。

## 参 考 文 献

- [1] Wang X J, Bai H Y, Ni Y B. Two-eye gaze tracking based on pupil shape in space[J]. Laser & Optoelectronics Progress, 2019, 56(2): 023301.  
王向军, 白皓月, 倪育博. 基于瞳孔空间形态的双眼视线跟踪方法[J]. 激光与光电子学进展, 2019, 56(2): 023301.
- [2] Ji C, Huang X B, Cao W, et al. Salient region



- detection of images based on deep learning[J]. *Laser & Optoelectronics Progress*, 2019, 56(9): 091007.
- 纪超, 黄新波, 曹雯, 等. 基于深度学习的图像显著区域检测[J]. *激光与光电子学进展*, 2019, 56(9): 091007.
- [3] Morimoto C H, Mimica M R M. Eye gaze tracking techniques for interactive applications[J]. *Computer Vision and Image Understanding*, 2005, 98(1): 4-24.
- [4] Zhang Q Z, Wang Y. Binocular stereo vision calibration accuracy evaluation using epipolar constraint [J]. *Laser & Optoelectronics Progress*, 2019, 56(23): 231504.
- 张青哲, 王勇. 基于对极约束的双目立体视觉标定精度评价方法[J]. *激光与光电子学进展*, 2019, 56(23): 231504.
- [5] Cheung Y M, Peng Q M. Eye gaze tracking with a web camera in a desktop environment[J]. *IEEE Transactions on Human-Machine Systems*, 2015, 45(4): 419-430.
- [6] Skodras E, Kanas V G, Fakotakis N. On visual gaze tracking based on a single low cost camera[J]. *Signal Processing: Image Communication*, 2015, 36: 29-42.
- [7] Martinikorena I, Larumbe-Bergera A, Ariz M, et al. Low cost gaze estimation: Knowledge-based solutions [J]. *IEEE Transactions on Image Processing*, 2020, 29: 2328-2343.
- [8] Sesma L, Villanueva A, Cabeza R. Evaluation of pupil center-eye corner vector for gaze estimation using a web cam[C]//ETRA '12: Proceedings of the Symposium on Eye Tracking Research and Applications, March 28-30, 2012, Santa Barbara, CA. New York: ACM, 2012: 217-220.
- [9] Xiao F, Zheng D D, Huang K J, et al. A single-camera gaze tracking system under natural light[J]. *Journal of Eye Movement Research*, 2018, 11(4): 5.
- [10] Cui Y Q, Hondzinski J M. Gaze tracking accuracy in humans: two eyes are better than one[J]. *Neuroscience Letters*, 2006, 396(3): 257-262.
- [11] Hennessey C, Noureddin B, Lawrence P. Fixation precision in high-speed noncontact eye-gaze tracking [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2008, 38(2): 289-298
- [12] Morimoto C H, Mimica M R M. Eye gaze tracking techniques for interactive applications[J]. *Computer Vision and Image Understanding*, 2005, 98(1): 4-24.
- [13] Blignaut P. A new mapping function to improve the accuracy of a video-based eye tracker[C]//Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference on - SAICSIT '13, October 7-9, 2013, East London, South Africa. New York: ACM Press, 2013: 56-59.
- [14] Blignaut P. Mapping the pupil-glint vector to gaze coordinates in a simple video-based eye tracker[J]. *Journal of Eye Movement Research*, 2013, 7(1): 4, 1-11.
- [15] Cerrolaza J J, Villanueva A, Cabeza R. Taxonomic study of polynomial regressions applied to the calibration of video-oculographic systems[C]//Proceedings of the 2008 symposium on Eye tracking research & applications-ETRA '08, March 26-28, 2008, Savannah, Georgia. New York: ACM Press, 2008: 259-266.
- [16] Sesma-Sanchez L, Villanueva A, Cabeza R. Gaze estimation interpolation methods based on binocular data[J]. *IEEE Transactions on Biomedical Engineering*, 2012, 59(8): 2235-2243.
- [17] Khan A Z, Crawford J D. Ocular dominance reverses as a function of horizontal gaze angle[J]. *Vision Research*, 2001, 41(14): 1743-1748.
- [18] Elbaum T, Wagner M, Botzer A. Cyclopean, dominant, and non-dominant gaze tracking for smooth pursuit gaze interaction[J]. *Journal of Eye Movement Research*, 2017, 10(1): 1-14.
- [19] Smith B A, Yin Q, Feiner S K, et al. Gaze locking: passive eye contact detection for human-object interaction[C]//Proceedings of the 26th annual ACM symposium on User interface software and technology, October 8-11, 2013, St. Andrews Scotland, United Kingdom. New York: ACM, 2013: 271-280.
- [20] Lozej J, Meden B, Struc V, et al. End-to-end iris segmentation using U-net[C]//2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI), July 18-20, 2018, San Carlos, Costa Rica. New York: IEEE Press, 2018: 1-6.
- [21] Wu X Q, Zhao L. Study on iris segmentation algorithm based on dense U-net[J]. *IEEE Access*, 2019, 7: 123959-123968.