

# 基于双目相机的眼动仪头部姿态估计方法

韩俊杰<sup>1</sup>, 骆开庆<sup>1,2\*</sup>, 邱健<sup>1,2</sup>, 刘冬梅<sup>1,2</sup>, 彭力<sup>1,2</sup>, 韩鹏<sup>1,2\*\*</sup>

<sup>1</sup>华南师范大学物理与电信工程学院, 广东 广州 510006;

<sup>2</sup>广东省光电检测工程技术研究中心, 广东 广州 510006

**摘要** 头部姿态估计是视线追踪的重要组成部分,但相关研究大多利用人脸的二维(2D)图像进行计算,因此,提出了一种基于双目相机的眼动仪三维(3D)头部姿态估计方法。首先,建立空间坐标系与人脸本征坐标系;然后,利用Dlib库检测双目图像的人脸特征点,并求取对应特征点的视差及3D坐标;最后,结合人脸对称性计算3D头部姿态的欧拉角,并提出了一种俯仰角修正算法。为了验证本方法的精确性,利用模特人头和高精度光学仪器构建了双目头部姿态数据集。实验结果表明,本方法测量欧拉角的平均误差为 $1.08^\circ$ ,均方根误差为1.25。相比基于2D图像的OpenFace方法,本方法的精度更高、鲁棒性更强。

**关键词** 图像处理; 双目视觉; 视线追踪; 头部姿态; 欧拉角

中图分类号 O436

文献标志码 A

doi: 10.3788/LOP202158.1410022

## Head Pose Estimation Method of Eye Tracker Based on Binocular Camera

Han Junjie<sup>1</sup>, Luo Kaiqing<sup>1,2\*</sup>, Qiu Jian<sup>1,2</sup>, Liu Dongmei<sup>1,2</sup>, Peng Li<sup>1,2</sup>, Han Peng<sup>1,2\*\*</sup>

<sup>1</sup> School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou, Guangdong 510006, China;

<sup>2</sup> Guangdong Provincial Engineering Research Center for Optoelectronic Instrument, Guangzhou, Guangdong 510006, China

**Abstract** Head pose estimation is an important part of gaze tracking but is usually calculated from two-dimensional (2D) images of human faces in most studies. Therefore, a three-dimensional (3D) head pose estimation method of eye tracker based on binocular camera is proposed in this paper. First, the spatial coordinate system and the facial intrinsic coordinate system are established. Then, the facial landmarks of the binocular image are detected by referencing the Dlib and the disparity and 3D coordinates of the corresponding facial landmarks are calculated. Finally, the Euler angle of the 3D head pose is calculated based on the symmetry of the face, and a pitch-angle correction algorithm is applied. To verify the accuracy of the method, a binocular head pose dataset is constructed using model heads and high-precision optical instruments. The experimental results show that the average error of the Euler angle by the method is  $1.08^\circ$  and the root mean square error is 1.25. The accuracy and robustness of the method exceeded those of the OpenFace method based on 2D images.

**Key words** image processing; binocular vision; eye tracking; head pose; Euler angle

**OCIS codes** 100.2000; 100.5010; 100.6890

## 1 引言

实现头部全自由、大活动范围的高精度视线估

计是理想视线跟踪系统的目标<sup>[1]</sup>,基于特征的视线跟踪技术从采集图像中提取眼睛特征的相关参数,如瞳孔、虹膜、眼角和角膜反射点特征,并通过建立

收稿日期: 2020-08-06; 修回日期: 2020-09-10; 录用日期: 2020-10-12

基金项目: 国家自然科学基金(61975058)、广东省自然科学基金(2019A1515011401)、广州市科技计划(201707010485, 201704020137)、华南师大物电学院蓝盾信息安全技术联合实验室课题(LD20170208)

通信作者: \*kqluo@sncu.edu.com; \*\*hanp@sncu.edu.cn

视线映射模型以及头部姿态运动模型,估计视线方向和落点<sup>[2-3]</sup>,实现头部自由活动的视线追踪。因此,需要进行头部模型姿态的补偿校正<sup>[1-3]</sup>。目前,对于头部姿态的检测大多是基于二维(2D)图像进行估计,这类方法往往只能反映头部姿态的基本朝向,当转动角度较大时误差偏大,且对光照变化也十分敏感<sup>[4]</sup>。

头部姿态包括三个旋转欧拉角,即偏航角(Yaw)、俯仰角(Pitch)和翻滚角(Roll)<sup>[5-6]</sup>。根据数据源可将头部姿态估计方法分为基于 2D 彩色图像、基于深度图像和基于三维(3D)图像的方法<sup>[5-6]</sup>。基于 2D 图像的方法只需一个摄像头,且大多数方法都是通过 2D 人脸姿态库结合深度学习方法实现<sup>[7-8]</sup>,如 OpenFace 方法可以反映头部的基本朝向;基于深度图像的头部姿态估计方法通过各类深度传感器实现,获取头部稠密深度图像后用随机森林等方法估计头部姿态<sup>[9-11]</sup>,但该方法的复杂度较高,不易推广;基于 3D 图像的头部姿态估计方法使用的传感器价格较贵,且应用范围窄、3D 图像重构技术不成熟,相关研究较少<sup>[12]</sup>。

基于双目立体视觉的 3D 重建技术具有速度快、精度高等优点<sup>[13]</sup>,在 3D 稠密、稀疏重建等领域具有很大的优势;且该技术仅需两个摄像头,成本低、设计更简单,得到了广泛的应用<sup>[14-15]</sup>。因此,本文提出了一种基于双目相机的眼动仪 3D 头部姿态估计方法。首先,建立了空间和人脸本征直角坐标系;然后,结合 Dlib 库人脸特征点检测和双目立体感知原理<sup>[15]</sup>得到特征点的 3D 空间位置坐标;最后,根据 3D 几何方法计算出头部姿态的欧拉角,并对俯仰角进行欧拉修正。为了验证本方法的准确性,利用精密光学仪器构建了双目头部姿态数据集,并与目前应用最广泛的 OpenFace 方法<sup>[16]</sup>进行了对比,验证了本方法的高精度和强鲁棒性。相比其他基于深度图像和 3D 图像的方法,本方法简化了硬件成本和实验复杂度,并获得了更精确的测量结果,为眼动仪的头部姿态估计提供了一种新思路。

## 2 人脸特征点三维坐标及头部姿态的估计

### 2.1 人脸特征的三维重建

本方法利用 Dlib 库进行人脸特征点检测<sup>[17]</sup>,Dlib 库是一个现代 C++ 跨平台通用框架,通过用方向梯度直方图(HOG)特征与级联回归树算法训

练的模型<sup>[18]</sup>进行人脸检测与定位,是目前人脸特征点检测定位中最常用的方法。通过 Dlib 人脸特征点检测可得到左右图像中人脸特征点的像素坐标,经过双目相机标定后,可认为左右图像各对应特征点的  $v$  坐标值在同一极线上。由于标定时标定板的精度较低,会产生微小的偏差。因此,取左右图像中特征点的  $v$  坐标平均值作为计算的 3D 特征点  $v$  坐标值,可表示为

$$v_d^i = (v_R^i + v_T^i) / 2, \quad (1)$$

式中,上标  $i$  为人脸特征点的标号,取值范围为 1~68,  $v_R^i$  和  $v_T^i$  分别为参考相机(Reference)和目标相机(Target)获取的像素  $v$  坐标值。实验中参考相机为左相机,目标相机为右相机。图 1 为人脸特征点的视差示意图,其中,特征点  $u$  坐标值的差即为对应特征点的视差  $d_d$ ,可表示为

$$d_d^i = u_R^i - u_T^i, \quad (2)$$

对于参考坐标系,有

$$u_d^i = u_R^i. \quad (3)$$

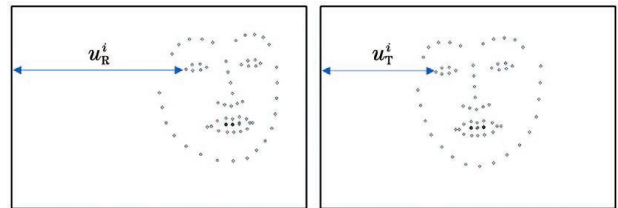


图 1 人脸特征点的视差

Fig. 1 Disparity of facial landmarks

记所求 68 个人脸特征点的相关坐标信息为  $(u_d^i, v_d^i, d_d^i)$ ,以参考相机的相机坐标系建立世界坐标系,基于双目立体视觉原理以及相机成像模型,得到人脸特征点在空间中的真实 3D 坐标  $(X^i, Y^i, Z^i)$  为

$$\begin{bmatrix} X^i \\ Y^i \\ Z^i \end{bmatrix} = \begin{bmatrix} \frac{fb}{u_R^i - u_T^i} \times \frac{u_d^i - u_{R_0}}{f_x} - \gamma \frac{fb}{u_R^i - u_T^i} \times \frac{v_d^i - v_{R_0}}{f_x f_y} \\ \frac{fb}{u_R^i - u_T^i} \times \frac{v_d^i - v_{R_0}}{f_y} \\ \frac{fb}{u_R^i - u_T^i} \end{bmatrix}, \quad (4)$$

式中,  $b$  为双目相机的基线距离,  $f$  为相机的焦距,  $(u_{R_0}, v_{R_0})$  为参考图像的像素中心坐标。由于左右相机为相同型号,且 CCD 感光板的制造精度较高,  $f_x, f_y$  几乎相等,可统一用  $f$  代替。因此,通过输入的人脸双目图像可计算出特征点在空间中的 3D 坐标。

### 2.2 头部姿态的估计

在获取人脸特征点的 3D 位置信息后,可进行头部姿态估计。本方法基于人脸的对称性及特征点的 3D 位置估计头部姿态的欧拉角,具体流程如图 2 所示。以参考相机坐标系建立世界坐标系  $O-XYZ$ ,在人头部建立人脸本征坐标系,如图 3 所示。以人脊柱顶部作为坐标原点,取平行于两眼角连线方向且过坐标原点的直线为  $a$  轴,取正脸姿态时,过原点且垂直水平面方向为  $b$  轴;取垂直于  $a$ 、 $b$  轴且过原点的垂线为  $c$  轴。其中, $\alpha$ 、 $\beta$ 、 $\gamma$  分别为头部姿态的偏航角、翻滚角和俯仰角。

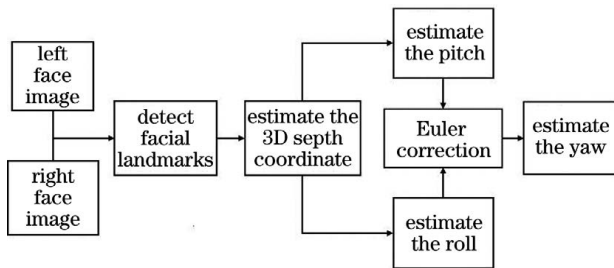


图 2 3D 头部姿态的估计流程

Fig. 2 Flow chart of 3D head pose estimation

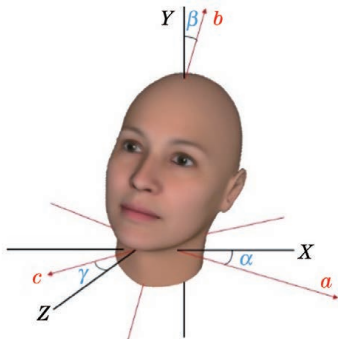


图 3 人脸的本征坐标系

Fig. 3 Facial intrinsic coordinate system

偏航角为  $a$  轴与世界坐标系  $O-XY$  平面的夹角,可通过人脸对称点的连线在  $O-XZ$  平面的投影与  $X$  轴的夹角计算,图 4 为头部与世界坐标系之间的位置关系。其中, $\Delta x$ 、 $\Delta z$  为所用对称点在该方向的 3D 坐标差值。偏航角可表示为

$$\alpha = \arctan \frac{Z^l - Z^r}{X^l - X^r}, \quad (5)$$

式中, $X$ 、 $Y$ 、 $Z$  为特征点的 3D 坐标,上标  $l$ 、 $r$  分别表示人脸左右对称点, $\alpha$  为负值时,表示头朝右偏转。

翻滚角为  $b$  轴与世界坐标系  $O-YZ$  平面的夹角,可通过人脸对称点的连线在  $O-XY$  平面的投影与  $X$  轴的夹角计算。图 5 为世界坐标系沿  $Z$  轴方向观察到的人脸, $\beta$  为负值时,表示头部向右翻滚,

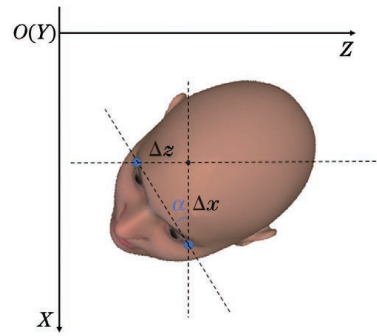


图 4 偏航角的平面图

Fig. 4 Plan view of the yaw

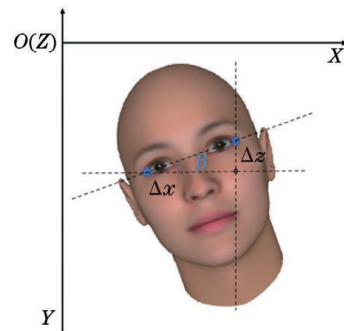


图 5 翻滚角的平面图

Fig. 5 Plan view of the roll

可表示为

$$\beta = \arctan \frac{Y^l - Y^r}{X^l - X^r}. \quad (6)$$

俯仰角为  $c$  轴与世界坐标系  $O-XZ$  平面的夹角,选取不同高度的人脸特征点连线并在  $O-YZ$  平面投影,计算上下两点在  $Y$  轴和  $Z$  轴上的坐标差  $\Delta y$ 、 $\Delta z$ ,就能得到俯仰角的变化量  $\Delta \gamma$ 。图 6 为沿  $X$  轴方向观察到的人脸侧视图,在翻滚角、偏航角都为  $0^\circ$  时, $\Delta \gamma$  可表示为

$$\Delta \gamma = \arctan \frac{\frac{Z^l + Z^r}{2} - \frac{Z^l + Z^r}{2}}{\frac{Y^l + Y^r}{2} - \frac{Y^l + Y^r}{2}}, \quad (7)$$

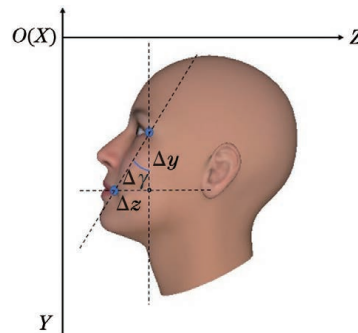


图 6 俯仰角的平面图

Fig. 6 Plan view of the pitch

式中,  $l', r'$  表示与  $l, r$  不同高度的左右对称点。

(7) 式中的  $\Delta\gamma$  为俯仰角的变化量, 预先测量初始的俯仰角  $\gamma_0$  后, 计算得到真正的俯仰角  $\gamma$  为

$$\gamma = \Delta\gamma + \gamma_0. \quad (8)$$

上述方法只能计算偏航角和翻滚角为  $0^\circ$  时的俯仰角, 当偏航角和翻滚角发生变化时, 会导致高度不同的一组特征点在  $Z$  和  $Y$  方向的坐标差  $\Delta z$ 、 $\Delta y$  发生变化, 进而导致俯仰角的计算偏差。因此, 计算任意姿态的俯仰角时, 要对不同大小偏航角与翻滚角导致的  $\Delta z$ 、 $\Delta y$  偏差进行修正, 即欧拉修正。翻滚角的欧拉修正原理如图 7 所示, 存在翻滚变化时, 根据所求 3D 坐标可直接得到实际  $y$  方向的坐标差  $\Delta y'$ , 与  $\Delta\gamma = 0^\circ$  时  $y$  方向坐标差  $\Delta y$  的偏差为  $y^*$ 。根据翻滚角  $\beta$  以及三角关系得到无翻滚角时的  $\Delta y$  为

$$\gamma\Delta y = \Delta y' / \cos \beta. \quad (9)$$

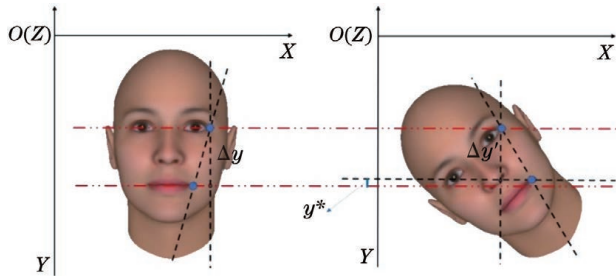


图 7 翻滚角的欧拉修正

Fig. 7 Euler correction of the roll

偏航角的欧拉修正原理如图 8 所示, 与翻滚角的欧拉修正类似, 存在偏航角变化时, 通过特征点 3D 坐标可直接计算  $Z$  方向的坐标差  $\Delta z'$ , 与  $\Delta\gamma = 0^\circ$  时的偏差为  $z^*$ 。根据三角关系及  $\Delta z'$ , 修正拟合出的真实的  $\Delta z$  为

$$\Delta z = \Delta z' / \cos \alpha. \quad (10)$$

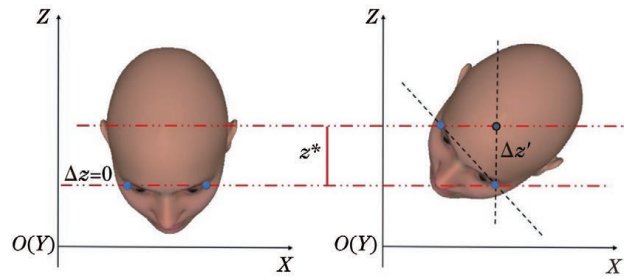


图 8 偏航角的欧拉修正

Fig. 8 Euler correction of the yaw

综上所述, 真实空间中一般状态的俯仰角可表示为

$$\gamma = \arctan\left(\frac{\Delta y / \cos \beta}{\Delta z / \cos \alpha}\right) + \gamma_0. \quad (11)$$

$\gamma$  为正值时, 表示在俯仰角初始值  $\gamma_0$  的基础上进行了仰视; 反之, 表示俯视。

### 3 实验结果与分析

实验搭建了双目视觉测量系统并构建了双目头部姿态数据集, 首先, 对双目相机系统进行立体标定, 得到相机的内外参数; 然后, 根据双目视觉 3D 感知原理得到人脸特征点的 3D 坐标; 最后, 利用特征点的 3D 空间位置结合 3D 几何法计算头部姿态。

#### 3.1 双目相机的立体标定

实际中双目相机及镜头不可能做到理想对齐, 为了修正相机安装偏差和镜头的偏差, 同时获取相机的参数, 需对双目相机进行标定和立体校正。采用 Zhang<sup>[19]</sup> 提出的棋盘格相机标定方法进行标定, 标定模板为  $9 \times 7$  的棋盘图案, 大小为  $25 \text{ mm} \times 25 \text{ mm}$ 。Matlab、OpenCV 等工具的相机标定都可以实现调用, 且标定精度较高。本相机标定得到的参数为

$$\mathbf{M}_R = \begin{bmatrix} 649.406 & 0.53374 & 335.497 \\ 0 & 649.903 & 233.844 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{M}_T = \begin{bmatrix} 652.709 & 0.66963 & 331.625 \\ 0 & 653.573 & 257.568 \\ 0 & 0 & 1 \end{bmatrix}, \quad (12)$$

$$\mathbf{D}_R = [0.0190002 \quad 0.528514 \quad -0.000827586 \quad -0.00350118 \quad -2.40424] \\ \mathbf{D}_T = [-0.0336197 \quad 0.538227 \quad -0.00230793 \quad -0.00292669 \quad -1.3883], \quad (13)$$

$$\mathbf{R} = \begin{bmatrix} 0.99992 & 0.00046615 & 0.012652 \\ 0.00039611 & 0.99998 & -0.0055382 \\ 0.012654 & 0.0055327 & 0.9990 \end{bmatrix}, \quad (14)$$

$$\mathbf{T} = [-60.54 \quad 0.4938 \quad -0.1906]^\top, \quad (15)$$

式中,  $\mathbf{M}_R$  和  $\mathbf{M}_T$  分别为左右相机的内参矩阵,  $\mathbf{D}_R$  和  $\mathbf{D}_T$  分别为左右相机的畸变参数,  $\mathbf{R}$  和  $\mathbf{T}$

分别为右相机相对于左相机的旋转矩阵和平移矩阵。

### 3.2 实验平台及双目头部姿态数据集

搭建的实验平台如图 9(a)所示,相机与头部模型的距离约为 55 cm,符合眼动仪的一般应用场景距离。人头模型装配精密的光学测量组件,两个摆角器与一个光学旋转平台可模拟人头模型在三个方向的转动,且测量精度高,可作为欧拉角测量的实际标准值,如图 9(b)所示。双目相机使用全瑞视

讯 70°USB 双目摄像头模组,如图 9(c)所示。实验建立了一个双目头部姿态数据集,包括不同背景、光照条件下的图像,如图 10 所示。考虑到仪器的量程,将偏航角、俯仰角、翻滚角的范围均设置为  $-15^{\circ}\sim 15^{\circ}$ 。实验时,三个欧拉角每次旋转  $1^{\circ}$ ,共旋转 31 次,每组欧拉角包含 93 种头部姿态的图像,共 25 组、2325 张( $25\times 93$ )头部姿态图像。

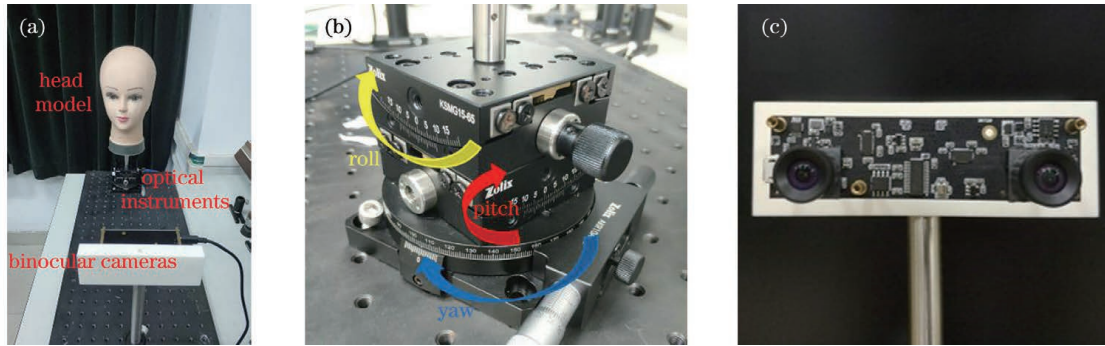


图 9 实验平台。(a)实验装置;(b) 3D 旋转光学仪器;(c)双目相机模组

Fig. 9 Experiment platform. (a) Experiment system; (b) 3D rotating optical instruments; (c) binocular cameras

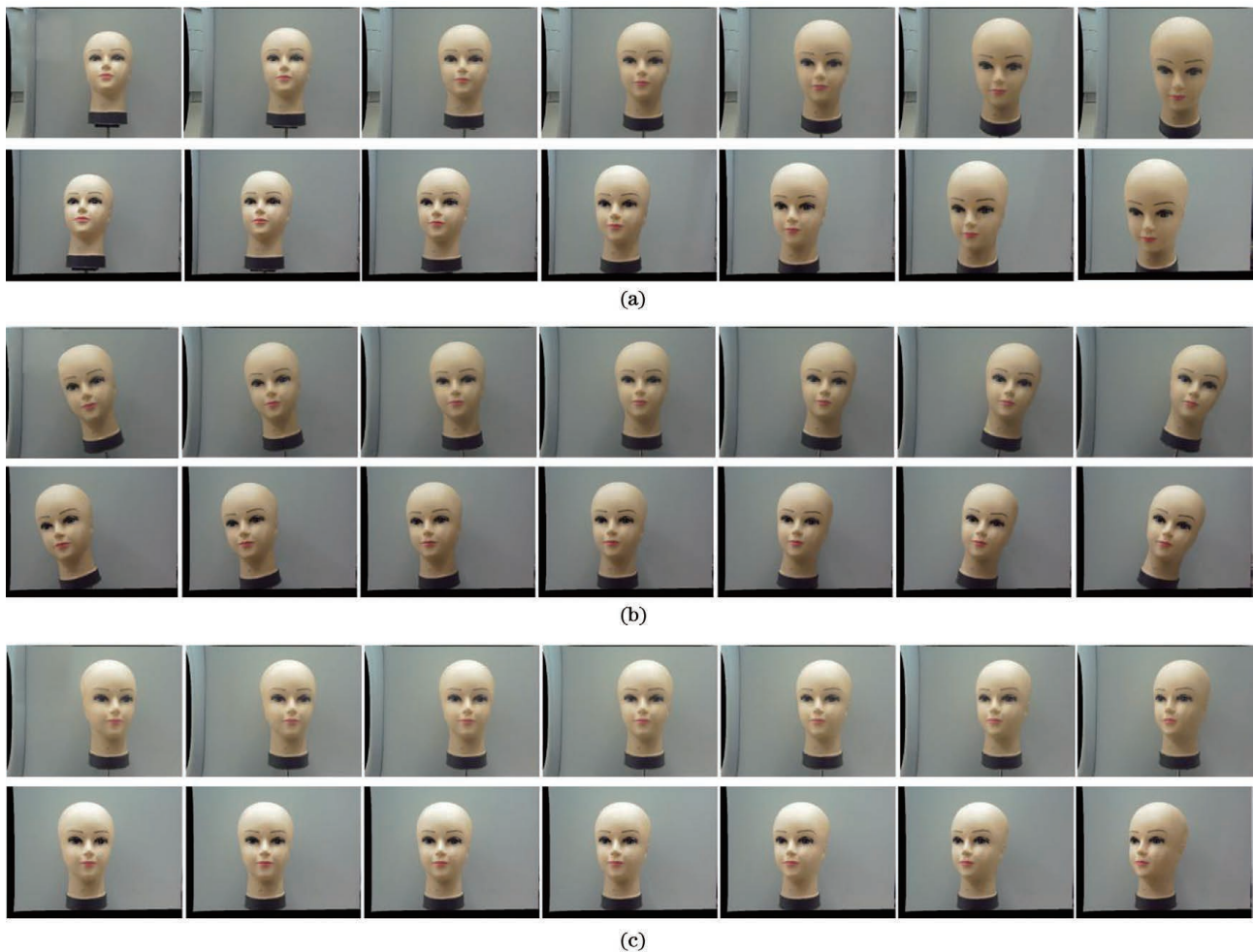


图 10 头部姿态数据集中的图像。(a)俯仰角;(b)翻滚角;(c)偏航角

Fig. 10 Images in the head pose dataset. (a) Pitch; (b) roll; (c) yaw

### 3.3 实验结果

#### 3.3.1 人脸特征点的 3D 坐标

检测并匹配左右相机人脸各 68 个特征点后,根据双目视觉视差计算原理,可计算世界坐标系中特征点的 3D 坐标,并进行特征点 3D 绘图。图 11 为某一姿态下的人脸 3D 特征点轮廓信息的可视化结果。

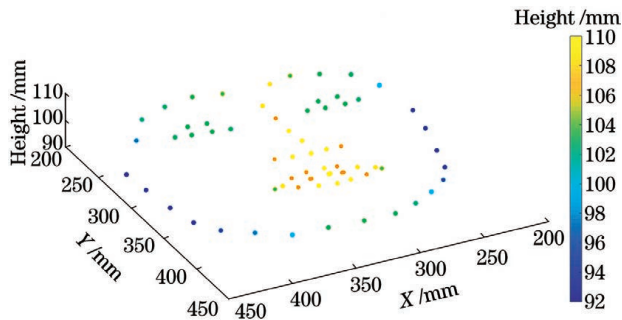


图 11 3D 人脸特征点的可视化

Fig. 11 Visualization of 3D facial landmarks

#### 3.3.2 头部姿态估计结果

用本方法和基于 2D 图像中广泛应用的 OpenFace 方法进行欧拉角测量,结果如图 12~图 14 所示。可以发现,在  $-15^{\circ} \sim 15^{\circ}$  范围测量三个头部姿态的欧拉角时,本方法的平均误差明显小于 OpenFace 方法,且对于翻滚角和俯仰角的测量精度更高。

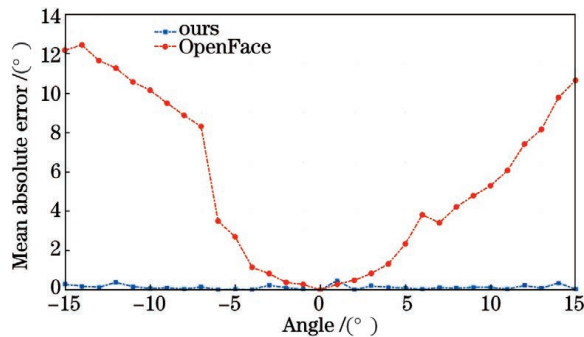


图 12 翻滚角的平均绝对误差

Fig. 12 Average absolute error of the roll

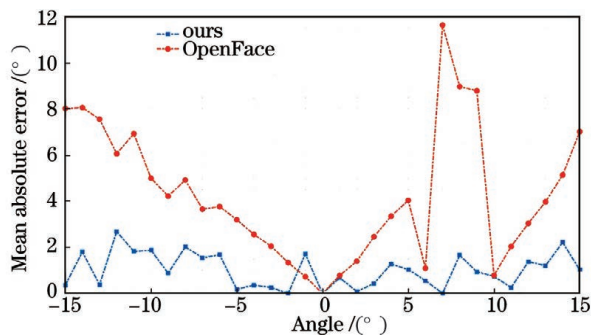


图 13 俯仰角的平均绝对误差

Fig. 13 Average absolute error of the pitch

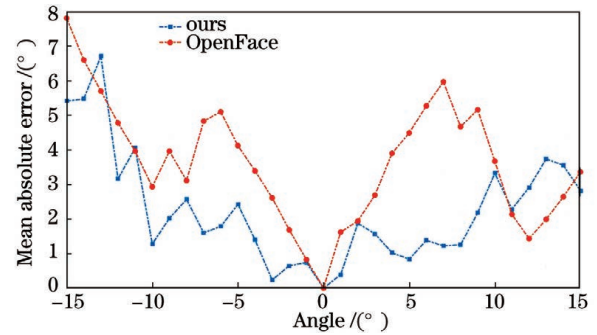


图 14 偏航角的平均绝对误差

Fig. 14 Mean absolute error of the yaw

表 1~表 3 为本方法和 OpenFace 方法测量三个欧拉角时的平均绝对误差、均方根误差(RMS)和大角度旋转( $>7^{\circ}$ )的平均绝对误差。可以发现,本方法测量翻滚角、俯仰角、偏航角的绝对误差平均值分别为  $0.14^{\circ}$ 、 $1.00^{\circ}$ 、 $2.09^{\circ}$ ,绝对误差的 RMS 分别为  $0.18$ 、 $1.11$ 、 $2.47$ ; OpenFace 方法测量翻滚角、俯仰角、偏航角的绝对误差平均值分别为  $5.25^{\circ}$ 、 $4.28^{\circ}$ 、 $3.62^{\circ}$ ,绝对误差的 RMS 分别为  $7.00$ 、 $5.24$ 、 $4.07$ 。对于大角度头部旋转欧拉角的测量,本方法的准确性、鲁棒性也明显优于 OpenFace 方法。原因是传统 OpenFace 方法受算法本身的限制,测量误差随头部转动角度的增大而增大,只能反映头部的基本朝向,若将其应用于眼动仪等高精度场合中,头部姿态估计的误差偏大,难以进行精确追踪。

表 1 欧拉角测量的绝对误差

Table 1 Absolute error of Euler angle measurement

unit: ( $^{\circ}$ )				
Absolute error	Roll	Pitch	Yaw	Mean
OpenFace	5.25	4.28	3.62	4.38
Ours	0.14	1.00	2.09	1.08

表 2 欧拉角测量的 RMS

Table 2 RMS of Euler angle measurement

RMS	Roll	Pitch	Yaw	Mean
OpenFace	7.00	5.24	4.07	5.44
Ours	0.18	1.11	2.47	1.25

表 3 大角度欧拉角测量的绝对误差

Table 3 Absolute error of large-angle Euler angle measurement

unit: ( $^{\circ}$ )				
Absolute error	Roll	Pitch	Yaw	Mean
OpenFace	8.60	5.88	4.21	6.23
Ours	0.16	1.27	1.93	1.12

## 4 结 论

将基于双目立体视觉的人脸特征 3D 重建与头部姿态估计相结合,提取双目图像的人脸特征点,并计算对应特征点的视差和 3D 空间坐标,以建立世界坐标系和人脸本征坐标系,获得精确的头部姿态估计。为了衡量本方法的稳定性和准确性,制作了符合眼动仪使用环境的双目头部姿态数据集。实验结果表明,相比基于 2D 图像的 OpenFace 方法,本方法估计的头部姿态欧拉角精度更高。原因是本方法包含头部姿态的 3D 信息,大大提高了头部姿态估计的精确度和准确度,在眼动仪及视线追踪场景中具有广泛的应用前景。但本方法对于偏航角的测量精度略低,下一步还需在人脸特征的匹配算法层面继续优化。此外,采用更高分辨率的相机会得到精度更高的视差值,进一步提高本方法的测量精度。

### 参 考 文 献

- [1] Liu J H, Chi J N, Yin Y X. A review of feature-based gaze tracking methods [J]. *Acta Automatica Sinica*, 2021, 47(2): 252-277.  
刘佳惠, 迟健男, 尹怡欣. 基于特征的视线跟踪方法研究综述 [J]. *自动化学报*, 2021, 47(2): 252-277.
- [2] 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.
- [3] Qin H B, Yan W H, Wang X L, et al. Agaze tracking system overcoming influences of head movements [J]. *Acta Electronica Sinica*, 2013, 41(12): 2403-2408.  
秦华标, 严伟洪, 王信亮, 等. 一种可克服头动影响的视线跟踪系统 [J]. *电子学报*, 2013, 41(12): 2403-2408.
- [4] Alioua N, Amine A, Rogozan A, et al. Driver head pose estimation using efficient descriptor fusion [J]. *EURASIP Journal on Image and Video Processing*, 2016, 2016(1): 2.
- [5] Tang Y Q, Sun Z N, Tan T N. A survey on head pose estimation [J]. *Pattern Recognition and Artificial Intelligence*, 2014, 27(3): 213-225.  
唐云祁, 孙哲南, 谭铁牛. 头部姿势估计研究综述 [J]. *模式识别与人工智能*, 2014, 27(3): 213-225.
- [6] Murphy-Chutorian E, Trivedi M M. Head pose estimation in computer vision: a survey [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(4): 607-626.
- [7] Wang W W, Chen X Y, Zheng S W, et al. Fast head pose estimation via rotation-adaptive facial landmark detection for video edge computation [J]. *IEEE Access*, 2020, 8: 45023-45032.
- [8] Liang L Y, Zhang T T, He W. Head pose estimation based on multi-scale convolutional neural network [J]. *Laser & Optoelectronics Progress*, 2019, 56(13): 131003.  
梁令羽, 张天天, 何为. 多尺度卷积神经网络的头部姿态估计 [J]. *激光与光电子学进展*, 2019, 56(13): 131003.
- [9] Fanelli G, Dantone M, Gall J, et al. Random forests for real time 3D face analysis [J]. *International Journal of Computer Vision*, 2013, 101(3): 437-458.
- [10] Fanelli G, Weise T, Gall J, et al. Real time head pose estimation from consumer depth cameras [M] // Mester R, Felsberg M. *Pattern recognition. Lecture notes in computer science*. Heidelberg: Springer, 2011, 6835: 101-110.
- [11] Fanelli G, Gall J, van Gool L. Real time head pose estimation with random regression forests [C] // *CVPR 2011*, June 20-25, 2011, Colorado Springs, CO, USA. New York: IEEE Press, 2011: 617-624.
- [12] Bruno F, Bianco G, Muzzupappa M, et al. Experimentation of structured light and stereo vision for underwater 3D reconstruction [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2011, 66(4): 508-518.
- [13] Ding S W, Zhang X H, Yu Q F, et al. Overview of non-contact 3D reconstruction measurement methods [J]. *Laser & Optoelectronics Progress*, 2017, 54(7): 070003.  
丁少闻, 张小虎, 于起峰, 等. 非接触式三维重建测量方法综述 [J]. *激光与光电子学进展*, 2017, 54(7): 070003.
- [14] Xing J G, Wang W L, Zhao W M, et al. A novel multi-touch human-computer-interface based on binocular stereo vision [C] // *2009 International Symposium on Intelligent Ubiquitous Computing and Education*, May 15-16, 2009, Chengdu, China. New York: IEEE Press, 2009: 319-323.
- [15] Dahl T R, Spector G. Binocular stereovision: US4982278 [P]. 1991-01-01.
- [16] Baltrušaitis T, Robinson P, Morency L P. OpenFace: an open source facial behavior analysis toolkit [C] // *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, March 7-10, 2016, Lake

- Placid, NY, USA. New York: IEEE Press, 2016: 1-10.
- [17] Dlib C++ library[EB/OL]. [2020-08-01]. <http://dlib.net/>.
- [18] Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 1867-1874.
- [19] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.