

基于双目立体视觉的视线跟踪系统标定

温晴川 达飞鹏 方旭

(东南大学自动化学院, 江苏 南京 210096)

摘要 针对使用主动红外光的视线跟踪系统存在的限制, 构建了在自然光条件下允许头部自然移动的视线跟踪系统, 该系统通过双目立体视觉系统获取眼睛特征参数, 在三维空间中重构眼睛视线向量来估计眼睛的视线。通过头戴式标定板来估计头部姿态, 减少了视线跟踪过程中头部移动带来的误差影响; 提出了一种视线跟踪系统中个人参数的标定方法, 该方法先假定眼睛视轴与光轴重合, 由此计算出眼球的中心、半径、眼睛视轴与光轴的夹角的初始估计值, 用非线性优化算法对这些初始值进行优化。通过实验表明, 提出的视线跟踪算法对使用者限制较少, 同时具有较高的视线估计精度, 能满足很多视线跟踪应用场合的需要。

关键词 机器视觉; 视线跟踪; 立体视觉; 个人参数标定; 摄像机标定

中图分类号 TP391 **文献标识码** A **doi**: 10.3788/AOS201232.1015002

Calibration for Gaze Tracking System Based on Binocular Stereo Vision

Wen Qingchuan Da Feipeng Fang Xu

(School of Automation, Southeast University, Nanjing, Jiangsu 210096, China)

Abstract To solve the limitation in the gaze tracking system using active infrared light, a gaze tracking system under the condition of natural light allowing the human head moves freely is constructed. Binocular stereo vision method is employed to obtain the characteristic parameters of eyes, and then the sight line can be estimated. The head pose can be estimated by using headset calibration board and the error caused by head moving can be reduced. Then a calibration method of individual parameters in gaze tracking system is proposed. This method assumes the optical axis is coincided with the line of sight. Then the center and radius of eyeball, the angle between eye optical axis and visual axis can be calculated. The optimal parameters can be obtained by employing the nonlinear optimization algorithm to process the initial estimate values. Experiment results verify that the proposed method is almost unconstrained, and the estimation accuracy of sight line is high. So this method can satisfy the need of gaze tracking system in many cases.

Key words machine vision; gaze tracking; stereo vision; individual parameters calibration; camera calibration

OCIS codes 100.4999; 330.1400; 150.0155; 150.1488

1 引言

视线跟踪是判断使用者视线方向的技术, 在人机交互和疾病诊断等领域有着广泛的应用^[1], 如辅助残疾人、车辆安全驾驶、市场研究与广告分析、认知障碍诊断和虚拟现实等。随着电子技术、计算机技术、图像处理技术的迅速发展, 基于眼睛视频图像

处理的视线跟踪技术逐渐成为当前的研究热点, 该技术具有对使用者干扰少、精度高和操用简单等优点, 其基本工作原理为: 先利用摄像机获取人眼或脸部图像, 然后用软件实现图像中人脸和人眼的定位与跟踪, 从而估算用户在屏幕上的注视点位置^[2]。

在基于眼睛视频图像处理的视线跟踪技术中,

收稿日期: 2012-3-29; **收到修改稿日期**: 2012-04-28

基金项目: 国家自然科学基金(51175081, 61107001)和江苏省自然科学基金(BK2010058)资助课题。

作者简介: 温晴川(1982—), 男, 研究生, 主要从事视线跟踪及图像处理等方面的研究。

E-mail: wenqingchuan2010@163.com

导师简介: 达飞鹏(1968—), 男, 博士, 教授, 主要从事三维测量及三维人脸识别等方面的研究。

E-mail: dafp@seu.edu.cn(通信联系人)

使用瞳孔-角膜反射(PCCR)方法较为普遍,即通过一个或多个红外光源照射眼部,由于眼睛瞳孔对红外光的反射和吸收与眼睛其他部分不同,因此很容易在眼睛图像中定位到瞳孔中心,同时眼睛角膜在红外光照射下会在角膜表面产生亮斑,根据瞳孔中心与角膜亮斑的相对位置关系,使用基于二维映射的视线估计方法^[3,4]或直接的三维视线估计方法^[5,6],可以计算出视线方向。然而,基于主动红外光的视线跟踪方法使用起来有很多限制^[7,8]:1)红外光会受到室外光线的影响,因而在室外无法使用;2)红外光源与摄像机之间的相对位置需要复杂的标定过程;3)瞳孔和角膜反射亮斑通常在图像中较小,因而需要分辨率较高的摄像机;4)对于戴有眼镜的使用者,眼镜镜面对红外光会产生反射亮斑,会对角膜亮斑的提取产生干扰而导致视线跟踪算法的失效。

目前,也有很多在自然光条件下的视线跟踪方法。如文献[8]使用双摄像机,其中一个摄像机用来估计使用者的头部姿态,另一个摄像机用来跟踪眼睛,先通过标定算法获得眼睛视线模型的参数,然后结合头部姿态,得到眼睛在空间中的视线方向;文献[9]利用眼睛虹膜在摄像机拍摄的图像中投影映射为椭圆,然后根据椭圆的形状特征来估计人眼视线;文献[10]直接利用眼睛虹膜中心与眼角的相对位置关系来估计人眼视线。但是以上的方法还存在一定的缺点:1)标定过程对使用者限制较多,如文献[8]视线跟踪标定过程中需要使用者保持头部静止不动,不够人性化;2)采用的眼睛视线模型没有考虑到眼睛视轴与光轴之间存在的夹角,精度不高,如文献[9,10]。

针对以上的问题,通过双目立体视觉系统获取眼睛特征和头部姿态参数,构建了一种在自然光条件下允许头部在较小范围内自然移动的视线跟踪系统。对双目立体视觉系统进行标定,获得摄像机内外参数,同时确定双摄像机的相对位置,建立统一的世界坐标系,对于视线跟踪系统中的特征点使用双目立体视觉系统来获取其空间三维坐标。提出了一种视线跟踪系统中的个人参数标定方法和头部姿态估计方法,先对使用者进行个人参数标定,获得使用者眼球中心在头部坐标系中的坐标、眼睛视轴与光轴的夹角,然后结合头部姿态估计结果,获得人眼视线在空间中的视线向量;最后,对屏幕空间位置进行标定,获得屏幕平面在世界坐标系下的平面方程,通过求取眼睛视线向量与屏幕平面的交点,得到最终的视线估计值。

2 视线跟踪系统

2.1 系统结构

构建的视线跟踪系统如图1所示,具体的硬件设备为:两台UNIQ公司生产的分辨率为1380 pixel×1030 pixel的UP-1800/UP-1800CL摄像机、Matrox Solios图像采集卡、处理器为Core2的计算机、19 inch (48.26 cm)的Dell液晶显示器。使用者在使用时,戴着头戴式标定板注视屏幕,由设立在屏幕前方的双摄像机分别获取人脸图像,然后通过图像采集卡传到电脑主机,通过对提取到的特征参数进行处理,获得眼睛视线与屏幕的交点来得到眼睛视线估计值。



图1 视线跟踪系统

Fig. 1 Gaze tracking system

2.2 视线估计模型

根据眼睛的结构,采用图2中简化的眼睛视线模型^[8],定义眼球为一个几何中心为 Q ,半径为 r 的圆球体,眼睛虹膜中心 P 和中央凹 F 为圆球体表面上的点。眼球的光轴为眼球中心到虹膜中心的连线,眼睛的视轴为中央凹到虹膜中心的连线,眼睛的视线估计点为眼睛视轴与空间平面的交点。眼睛以眼球中心为旋转中心转动眼球来改变视线方向,因此当头部移

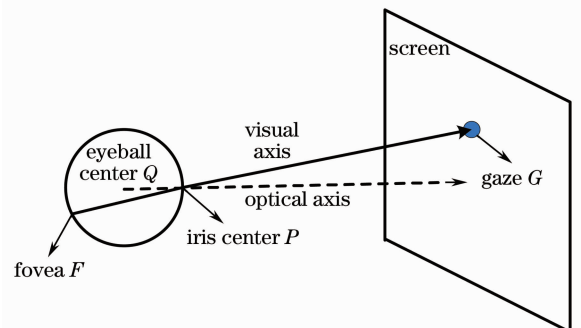


图2 眼睛视线模型

Fig. 2 Model of eye visual

动时,可以认为眼球中心与头部相对静止。

由于中央凹并不在人眼光轴轴线上,所以眼睛视线方向与眼睛光轴存在一定的夹角,也称之为Kappa角。眼睛视轴与光轴在水平方向上的夹角大约为 5° ,其中左眼 5° 左右,右眼 -5° 左右,垂直方向上的夹角大约为 $2^\circ\sim 3^\circ$,Kappa角因人而异,并没有固定值^[11]。由于眼睛的视线结构较为复杂,眼睛视轴方向一般不能直接得到,通常需要在估计到眼睛光轴方向后,加入眼睛视轴与光轴之间的补偿夹角,即Kappa角,进而得到眼睛视线方向。

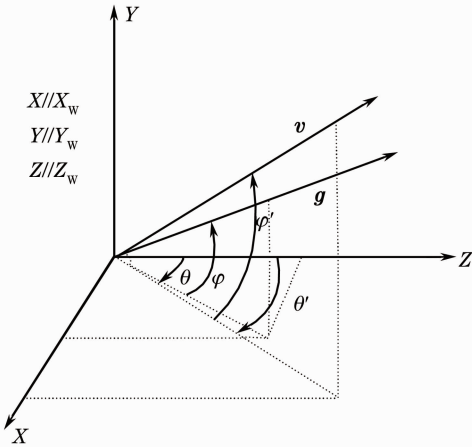


图3 视轴与光轴夹角

Fig. 3 Angle between visual axis and optical axis

为了更好地描述眼睛视轴与光轴之间的空间位置关系,建立如图3所示的空间坐标系,坐标系中的X轴、Y轴、Z轴分别与世界坐标系的 X_w 轴、 Y_w 轴、 Z_w 轴平行,眼睛光轴与坐标系水平方向和垂直方向的夹角分别为 θ 和 φ ,则眼睛光轴在世界坐标系下的单位向量可表示为^[12]

$$\mathbf{g} = \begin{bmatrix} -\sin\theta\cos\varphi \\ \sin\theta \\ \cos\theta\cos\varphi \end{bmatrix}, \quad (1)$$

且 θ 和 φ 可以通过光轴在空间中的单位向量 \mathbf{g} 得到

$$\begin{cases} \theta = \arctan(g_x/g_z) \\ \varphi = \arcsin(g_y) \end{cases}. \quad (2)$$

设视轴与光轴在水平和垂直方向的夹角分别为 α_{eye} 、 β_{eye} ,则眼睛视轴在水平方向和垂直方向上的夹角分别为 $\theta+\alpha_{eye}$ 、 $\varphi+\beta_{eye}$,则空间中的视线向量 \mathbf{v} 可以表示为

$$\mathbf{v} = \begin{bmatrix} -\sin(\theta+\alpha_{eye})\cos(\varphi+\beta_{eye}) \\ \sin(\theta+\alpha_{eye}) \\ \cos(\theta+\alpha_{eye})\cos(\varphi+\beta_{eye}) \end{bmatrix}. \quad (3)$$

通过(2)式、(3)式可以得到

$$\begin{cases} \alpha_{eye} = \arctan(v_x/v_z) - \arctan(g_x/g_z) \\ \beta_{eye} = \arcsin(v_y) - \arcsin(g_y) \end{cases}. \quad (4)$$

当使用者看着空间中的视线点G时,G点位于眼睛视线方向上,且满足

$$\mathbf{G} = \mathbf{P} + t\mathbf{v}, \quad (5)$$

式中 \mathbf{P} 为虹膜中心坐标, t 为参数, \mathbf{v} 为视线在空间中的单位向量,(5)式写成参数方程形式为

$$\begin{cases} G_x = Q_x + tv_x \\ G_y = Q_y + tv_y \\ G_z = Q_z + tv_z \end{cases}. \quad (6)$$

眼睛的视线估计点为眼睛视线与空间平面的交点,如果已知空间的平面方程为

$$Ax + By + Cz + D = 0, \quad (7)$$

联立(6)式、(7)式可得参数 $t = \frac{-AQ_x + BQ_y + CQ_z + D}{Av_x + Bv_y + Cv_z}$,把 t 代入(5)式即可得到视线估计点。

根据上面的讨论,要得到眼睛视线点的估计值,需要知道眼球中心Q与虹膜中心P的空间三维坐标、眼睛视轴与光轴在水平和垂直方向上的夹角 α_{eye} 与 β_{eye} 、屏幕的空间平面方程。首先为了补偿头部移动带来的视线估计误差,需要建立头部坐标系,对摄像机获取的每帧图像都进行头部姿态估计,得到头部坐标系相对于世界坐标系的变换关系,通过个人参数标定可以得到眼球中心在头部坐标系下的坐标,然后经过坐标变换可以得到各帧图像中眼球中心在世界坐标系下的坐标Q;虹膜中心P的空间三维坐标可以通过在左右摄像机图像中分别定位虹膜中心,然后结合双摄像机标定结果计算得到;眼睛视轴与光轴之间的夹角 α_{eye} 与 β_{eye} 可以通过对使用者进行个人参数标定得到;屏幕的空间平面方程可以通过屏幕标定得到。

3 视线跟踪系统的标定

3.1 双目立体视觉系统的标定

3.1.1 摄像机的标定

摄像机标定是计算机视觉和三维测量中最基本、最重要的工作^[13],现有的摄像机标定方法可分为传统标定方法、主动视觉标定方法^[14]和摄像机自标定方法^[15,16],其主要目的是确定摄像机的内、外部参数,在基于双目立体视觉的视线跟踪系统中,还需要确定双摄像机之间的相对位置关系,从而建立统一的世界坐标系,使整个视线跟踪系统处于同一个坐标系统中。

对单摄像机进行标定,采用 Zhang^[17]提出的一种基于平面模板的摄像机标定方法,这种方法计算简单,稳定性好,精度高,获得的摄像机标定参数较为准确。通过分别对左右摄像机参数进行标定,可以得到左右摄像机的内、外部参数。

确定双摄像机的相对位置关系,设空间中一点 I 在世界坐标系下非齐次坐标为 X_w ,在左摄像机坐标系中的非齐次坐标为 X_c^l ,在右摄像机坐标系中的非齐次坐标为 X_c^r ,对任意点 I 从世界坐标转换到摄像机坐标,左右摄像机分别有

$$X_c^l = R_l X_w + T_l, \quad (8)$$

$$X_c^r = R_r X_w + T_r, \quad (9)$$

则可以得出两者的相对关系

$$X_c^l = R X_c^r + T, \quad (10)$$

式中 $R=R_r R_l^{-1}$, $T=T_r - R_r R_l^{-1} T_l$ 。 R, T 就是右摄像机相对左摄像机的姿态和位置。通过确立双摄像机的相对位置关系,最终建立统一的世界坐标系。

3.1.2 双目立体视觉原理

双目视觉系统中空间点的三维坐标通过空间交汇法来获取,具体测量原理如图 4 所示。设空间中任意一点 I 在左右摄像机中成的像点分别为 I_l, I_r ,左右两个摄像机坐标系的原点分别位于其光心 O_l, O_r 处,根据摄像机的成像原理,点 I 同时位于 $O_l I_l$ 与 $O_r I_r$ 两条直线上。在理想情况下, $O_l I_l$ 与 $O_r I_r$ 的交点应为物点 I 的真实空间位置。但是由于在实际中存在噪声等问题, $O_l I_l$ 与 $O_r I_r$ 往往构成异面直线,此时取 $O_l I_l$ 与 $O_r I_r$ 构成的异面直线的公垂线中点 W 的三维坐标,认为是 I 点的三维坐标。

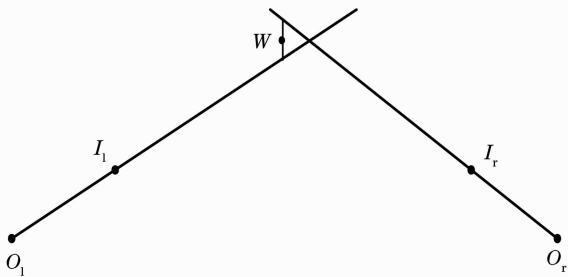


图 4 空间交汇法

Fig. 4 Method of space intersection

3.2 个人参数标定

个人参数标定的目的是求得眼球中心 Q 、眼睛视轴与光轴在水平和垂直方向上的夹角 α_{eye} 与 β_{eye} ,其标定过程如图 5 所示。首先在屏幕上设置一系列已知的视线点,然后让使用者带着头戴式标定板,在自然光照和允许头部自然移动的情况下,分别注视

屏幕上这些已知视线点,然后用摄像机拍摄图像,获取眼睛虹膜中心和头戴式标定板上标志圆的圆心等特征信息,最后使用本文提出的个人参数标定方法来获取使用者的个人参数信息。

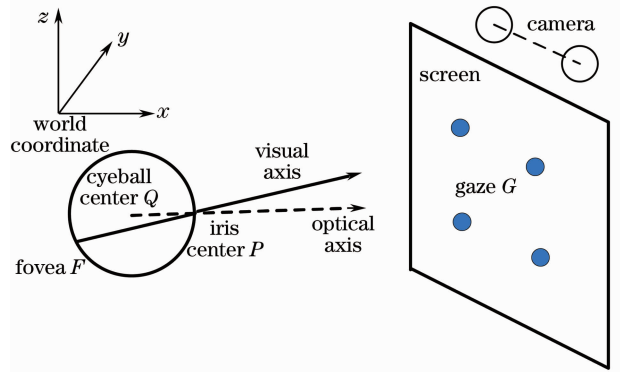


图 5 个人参数标定示意图

Fig. 5 Image of individual parameter calibration

3.2.1 头部姿态估计

人的视线由头部姿态和眼球方向共同决定,当头部静止时可通过旋转眼球来改变视线方向,当眼球相对头部静止时可通过改变头部的姿态来改变视线方向,因此,在视线估计中首先需要知道头部姿态。

头部姿态可以通过提取脸部特征点来估计,如鼻孔、眼角和嘴角等特征点^[18]。为了检验本文提出的个人参数标定方法的有效性,在使用者头上戴一个标定板,然后提取上面的特征标志点来估计头部姿态。标定板如图 6(a)所示,由于圆形标识点易于识别、圆心提取的精度高,因此选取用圆形标识点制成的圆阵列标定板。

对摄像机获取的头戴式标定板图像用自适应阈值分割可以得到图 6(b),经过边缘检测和圆形目标滤波可以得到图 6(c),用最小二乘法进行椭圆拟合可以得到每个标志圆的圆心、长轴半径和短轴半径,如图 6(d),最后取椭圆长轴和短轴半径的平均值作为圆形标志点的半径特征,根据半径大小可以把圆形标志点区分为大圆和小圆。

计算 5 个大圆圆心相互之间的距离,可以得出一对相互距离最小的大圆,以及一对相互距离最大的大圆,除去这 4 个大圆,还得到 1 个唯一的大圆,根据其他大圆圆心与这个唯一的大圆圆心之间连线的斜率,可以确定标定板上 5 个大圆的圆心在图像中坐标的匹配关系,然后可以得到标定板上其他小圆圆心的匹配关系。这样建立了左摄像机中图像上的圆心与右摄像机中图像上的圆心的匹配关系,然

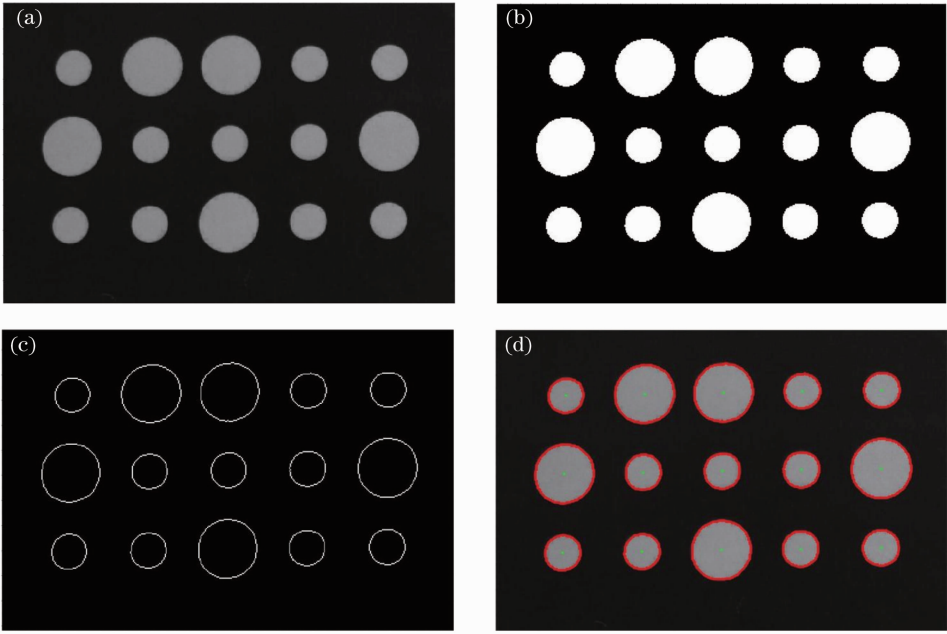


图 6 特征点提取。(a)标定板图像；(b)阈值分割结果；(c)边缘检测结果；(d)圆心定位结果

Fig. 6 Feature point extraction. (a) Image of calibration plane; (b) result of binarization; (c) result of edge detection; (d) result of center detection

后结合摄像机标定结果,可以求得其在世界坐标系下的三维坐标。

取相互距离最大的 2 个大圆的中间点为头部坐标系原点,按如图 7 所示方向建立头部坐标系,由于标定板是平面,所以标定板上的点在头部坐标系中 Z 轴方向的分量都为 0,且标定板上大圆与小圆为阵列分布,因此,标定板上 5 个大圆的圆心在头部坐标系中的非齐次坐标从上至下、从左至右依次为 $(-1, 1, 0)$, $(0, 1, 0)$, $(-2, 0, 0)$, $(2, 0, 0)$,

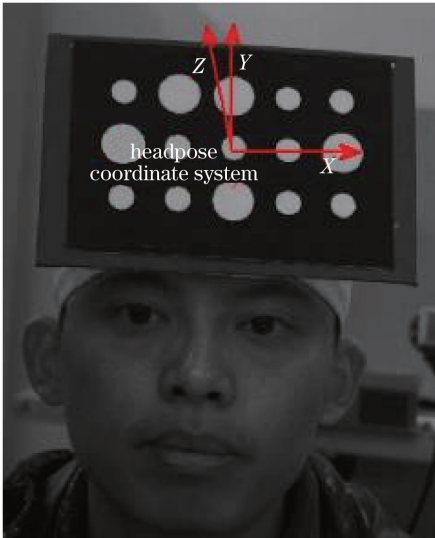


图 7 头部坐标系定义

Fig. 7 Definition of head coordinate system

$(0, -1, 0)$,其他小圆圆心在头部坐标系中的坐标也可以依次确定。

设头部坐标系与世界坐标系之间的旋转矩阵为 \mathbf{R} ,平移矩阵为 \mathbf{T} ,尺度因子为 s ,则头部姿态可用旋转矩阵 \mathbf{R} 、平移向量 \mathbf{T} 、尺度因子 s 来描述。假设头部坐标系中某一点 I 在世界坐标系与头部坐标系下的齐次坐标分别是 $(X_w, Y_w, Z_w, 1)^T$ 、 $(X_H, Y_H, Z_H, 1)^T$,存在如下的坐标变换关系

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} s\mathbf{R} & \mathbf{T} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} X_H \\ Y_H \\ Z_H \\ 1 \end{bmatrix} = \mathbf{M} \begin{bmatrix} X_H \\ Y_H \\ Z_H \\ 1 \end{bmatrix}, \quad (11)$$

式中 \mathbf{R} 为 3×3 正交单位矩阵, \mathbf{T} 为三维平移向量,因此设

$$\mathbf{R} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}, \quad (12)$$

$$\mathbf{T} = (T_x, T_y, T_z)^T. \quad (13)$$

由于 \mathbf{R} 为正交单位矩阵,则 \mathbf{R} 中的参数必定满足条件

$$\begin{cases} a_1^2 + a_2^2 + a_3^2 = 1 \\ b_1^2 + b_2^2 + b_3^2 = 1 \\ c_1^2 + c_2^2 + c_3^2 = 1 \\ a_1 a_2 + b_1 b_2 + c_1 c_2 = 0 \\ a_1 a_3 + b_1 b_3 + c_1 c_3 = 0 \\ a_2 a_3 + b_2 b_3 + c_2 c_3 = 0 \end{cases}. \quad (14)$$

(11)式中共有 13 个未知参数,对于 N 对匹配的坐标点,共有 $3N + 6$ 个表达式,所以理论上只要 $N > 3$ 就可以解得头部坐标系与世界坐标系之间的坐标变换矩阵 \mathbf{M} 。头戴式标定板上共有 15 个特征标志点,经过双目立体视觉系统三维重构可以得到 15 对匹配的坐标点,通过采用单位四元数^[19]的方法可以解得头部坐标系与世界坐标系之间的坐标变换矩阵 \mathbf{M} 。

3.2.2 眼球中心与半径的估计

设眼球的中心为 \mathbf{Q} ,当眼睛改变视线方向时,以眼球中心为旋转中心转动睛球来改变视线方向,因此可以认为眼球中心在头部坐标系中的坐标保持不变,眼球中心在世界坐标系中的位置由头部姿态决定。对于摄像机获取的每一帧图像,设眼球中心在头部坐标系中的坐标为 \mathbf{Q}^H ,在世界坐标系下的坐标为 \mathbf{Q}^Q ,经过头部姿态估计可以得到每帧图像中头部坐标系相对于世界坐标系的旋转矩阵 $\mathbf{R}_{H,i}^Q$,平移矩阵 $\mathbf{t}_{H,i}^Q$,尺度因子 s ,则有

$$\mathbf{Q}^Q = s\mathbf{R}_{H,i}^Q\mathbf{Q}^H + \mathbf{t}_{H,i}^Q. \quad (15)$$

为了估计眼球中心在头部坐标系中的坐标 \mathbf{Q}^H ,先假设眼睛视轴与光轴重合,则视线点 G 位于眼睛光轴方向上。设在每一帧图像中通过虹膜定位得到的虹膜中心的空间坐标为 P_i ,对于每一个空间中已知坐标的视线点 G_i ,眼睛虹膜中心 P_i 满足

$$P_i = Q_i + r \frac{G_i - Q_i}{|G_i - Q_i|}, \quad (16)$$

式中 r 为眼球半径,设 $b'_i = r/|G_i - Q_i|$,则有

$$(1 - b'_i)Q_i + b'_iG_i - P_i = 0. \quad (17)$$

设 $b = b'/(1 - b')$, $a = 1/(1 - b')$,则有

$$s\mathbf{R}_{H,i}^Q\mathbf{Q}^H + b_iG_i - a_iP_i = -\mathbf{t}_{H,i}^Q. \quad (18)$$

设用来进行个人参数的标定的图像共有 n frame,则可以得到方程组

$$\begin{bmatrix} s\mathbf{R}_{H,1}^Q & G_1 & -P_1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s\mathbf{R}_{H,n}^Q & 0 & 0 & \cdots & 0 & G_n & -P_n \end{bmatrix} \times \begin{bmatrix} \mathbf{Q}^H \\ b_1 \\ a_1 \\ \vdots \\ b_n \\ a_n \end{bmatrix} = \begin{bmatrix} -\mathbf{t}_{H,1}^Q \\ \vdots \\ -\mathbf{t}_{H,n}^Q \end{bmatrix}. \quad (19)$$

通过解(19)式的方程组,可以得到眼球中心在头部坐标系中的坐标 \mathbf{Q}^H 、参数 a_i, b_i ,则眼球半径 r 可以得出

$$r = \frac{1}{N} \sum_i \left(\frac{b_i}{1 + b_i} |G_i - Q_i| \right). \quad (20)$$

得到的眼球半径值可以用来判断解得的参数是否正确,因为眼球的平均半径在 12 mm 左右,因此求得的眼球半径值应该在 [10 mm, 14mm] 范围内,如果超出此范围,说明参数求解失败。

3.2.3 眼睛光轴与视轴夹角的估计

对于摄像机获得的每一帧图像,在得到眼球中心在头部坐标系中的空间坐标 \mathbf{Q}^H 后,根据头部坐标系相对于世界坐标系之间的坐标转换关系,即旋转矩阵为 $\mathbf{R}_{H,i}^Q$,平移矩阵 $\mathbf{t}_{H,i}^Q$,尺度因子 s ,由(15)式经过坐标转换后,可以得到眼球中心在世界坐标系下的坐标 \mathbf{Q}^Q ,经过虹膜定位可以得到虹膜中心 P_i ,则眼睛光轴的空间单位向量 \mathbf{g}_i 为

$$\mathbf{g}_i = \frac{P_i - Q_i^Q}{|P_i - Q_i^Q|}. \quad (21)$$

眼睛视轴的空间单位向量 \mathbf{v}_i 为

$$\mathbf{v}_i = \frac{G_i - P_i}{|G_i - P_i|}. \quad (22)$$

由(4)式可以求出每帧图像中眼睛视轴与光轴之间的夹角

$$\begin{cases} \alpha_{eye,i} = \arctan(v_{x,i}/v_{z,i}) - \arctan(g_{x,i}/g_{z,i}) \\ \beta_{eye,i} = \arcsin(v_{y,i}) - \arcsin(g_{y,i}) \end{cases}. \quad (23)$$

眼睛视轴与光轴之间的夹角 $\alpha_{eye}, \beta_{eye}$ 的最后估计值为各帧图像中计算得到的夹角 $\alpha_{eye,i}, \beta_{eye,i}$ 的平均值,即

$$\begin{cases} \alpha_{eye} = \frac{1}{N} \sum \alpha_{eye,i} \\ \beta_{eye} = \frac{1}{N} \sum \beta_{eye,i} \end{cases}. \quad (24)$$

3.2.4 个人参数优化

由于眼睛视轴与光轴之间存在一定的夹角,同时在个人参数标定过程中会存在一定的噪声,因此标定得到的个人参数结果会与真实值存在一定的误差。为了得到更高精度的结果,在得到眼球中心、眼睛视轴与光轴夹角的初始估计值后,可以用 Levenberg-Marquart 非线性优化算法进行优化来求取最佳值,优化的目标函数为

$$f = \min \sum_i \|G_i - \hat{G}_i\|^2, \quad (25)$$

式中 G_i 为空间中已知的视线点, \hat{G}_i 为经过计算得到的视线估计点,由(5)式可知视线点 \hat{G}_i 的估计为关于眼睛模型中眼球中心 \mathbf{Q} 、视轴与光轴的空间夹角 α_{eye} 与 β_{eye} 、虹膜中心 P 、参数 t 的函数。而在标定过程

中,虹膜中心 P 可以通过三维重构得到,同时由于视线点坐标已知,因此参数 t 也可以得到,视线点 \hat{G}_i 为关于 $Q, \alpha_{eye}, \beta_{eye}$ 的函数,故目标函数可以表示为

$$f = \min_{Q, \alpha, \beta} \sum_i \|G_i - \hat{G}_i(Q, \alpha_{eye}, \beta_{eye})\|^2. \quad (26)$$

经过优化,最后可以求得最佳的用户个人参数。

3.3 屏幕平面的标定

根据图 2 所示的视线估计模型,视线估计点为眼睛视线与屏幕空间平面的交点,因而需要确定屏幕的空间位置,但屏幕并不在双摄像机的可视范围内,因而不能直接确定其空间位置。使用平面镜反射^[11]的方法可以确定屏幕的空间位置,首先在屏幕上显示一幅标定图像,用贴有标定图像的平面镜反射屏幕上的图像到摄像机的视野之内,双摄像机可以获得平面镜上的标定点和平面镜反射屏幕上的标

定点,用平面镜上的标定点进行平面拟合可以得到平面镜的空间位置,然后根据平面镜反射原理,可以得到屏幕上的标定点的空间坐标,进而得到屏幕平面的空间位置。

4 实 验

4.1 视线跟踪系统的标定与结果

使用平面标定板在双摄像机前转动多个角度,左右摄像机各拍摄 4 幅不同姿态的标定板图像,获得如图 8 所示的标定板图像。

然后采用张正友等^[17]提出的一种基于平面模板图像的标定方法,经过摄像机标定后得到的双摄像机内、外部参数如表 1 所示。

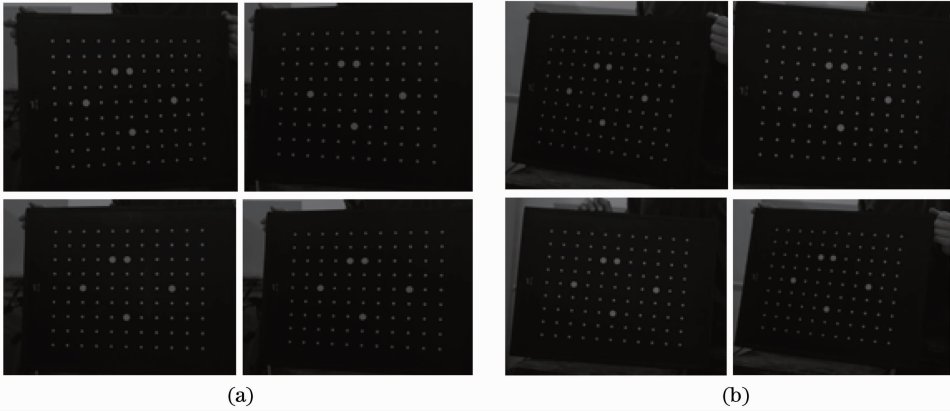


图 8 摄像机标定图像。(a)左摄像机; (b)右摄像机

Fig. 8 Calibration images of camera. (a) Left camera; (b) right camera

表 1 摄像机标定结果

Table 1 Result of camera calibration

Parameter matrix Camera	Intrinsic parameter matrix A	Rotation matrix R	Translation matrix T
Left camera	$\begin{bmatrix} 2477.2341 & 0 & 688.4926 \\ 0 & 2476.5832 & 560.8563 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0.9764 & 0.0918 & 0.1956 \\ -0.0712 & 0.9914 & -0.1096 \\ -0.2040 & 0.0930 & 0.9745 \end{bmatrix}$	$\begin{bmatrix} -60.0733 \\ -82.2704 \\ 1356.3387 \end{bmatrix}$
Right camera	$\begin{bmatrix} 2484.7839 & 0 & 667.5085 \\ 0 & 2487.7346 & 530.5238 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0.9983 & -0.0014 & 0.0572 \\ 0.0052 & 0.9977 & 0.0663 \\ -0.0569 & 0.0665 & 0.9962 \end{bmatrix}$	$\begin{bmatrix} 136.2447 \\ -58.8493 \\ 1374.4699 \end{bmatrix}$

图 9 是头戴式标定板上的特征点经过三维重构得到的空间点,根据这些标志点在头部坐标系与世界坐标系中的对应关系,可以求得头部坐标系到世界坐标系的旋转变换矩阵。为了评价头部姿态的估计精度,设圆形标志点 S_i 在头部坐标系中的坐标

为 $(X_{H,i}, Y_{H,i}, Z_{H,i})^T$, 在世界坐标系中的坐标为 $(X_{W,i}, Y_{W,i}, Z_{W,i})^T$, 经过计算得到的坐标之间的旋转矩阵为 R , 平移向量为 T , 尺度因子为 s , 那么头部坐标系中的点 $(X_{H,i}, Y_{H,i}, Z_{H,i})$ 转换到世界坐标系下坐标为

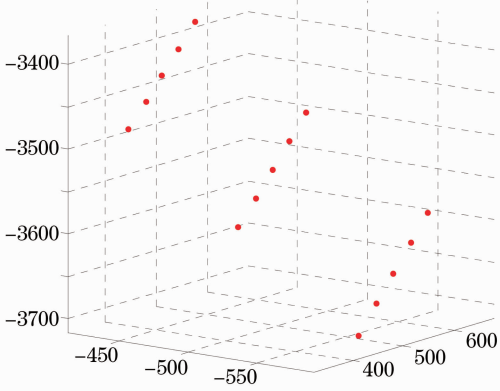


图 9 头戴式标定板上特征点的三维重构结果
Fig. 9 3D reconstruction result of headset calibration board feature point

$$\begin{bmatrix} X'_{w,i} \\ Y'_{w,i} \\ Z'_{w,i} \end{bmatrix} = s\mathbf{R} \begin{bmatrix} X_{H,i} \\ Y_{H,i} \\ Z_{H,i} \end{bmatrix} + \mathbf{T}. \quad (27)$$

头部姿态的估计精度 $\epsilon(i)$ 为

$$\epsilon(i) = \sqrt{(X_{w,i} - X'_{w,i})^2 + (Y_{w,i} - Y'_{w,i})^2 + (Z_{w,i} - Z'_{w,i})^2}. \quad (28)$$

得到的误差曲线如图 10 所示,图中横坐标为标

定板上的坐标点的序号,纵坐标为误差,误差值以像素为单位。从图 10 中可以看出,头部姿态误差控制在 8 pixel 以内,因此头部姿态估计的精度可以满足视线跟踪系统的要求。

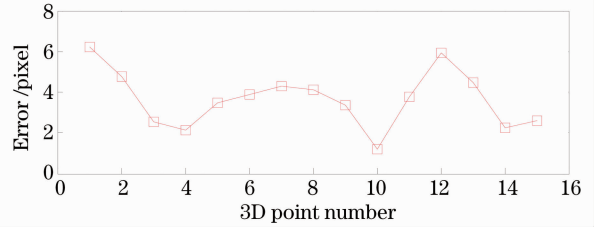


图 10 头部姿态估计误差

Fig. 10 Error of head pose estimation

图 11 为摄像机获取的带有平面镜上标定板和屏幕上标定板的图像,首先对得到图像进行阈值分割、边缘检测、圆形目标滤波、轮廓跟踪、椭圆拟合与中心提取,可以分别得到左右摄像机图像上特征点在图像中的坐标,接着结合双摄像机标定数据,使用空间交汇法可以得到这些点的空间三维坐标,然后对这些点进行平面拟合可以得到平面镜平面及屏幕镜像平面,最后根据镜面反射原理可以求得屏幕像平面关于平面镜的对称平面,即显示器屏幕在空间坐标中的位置。

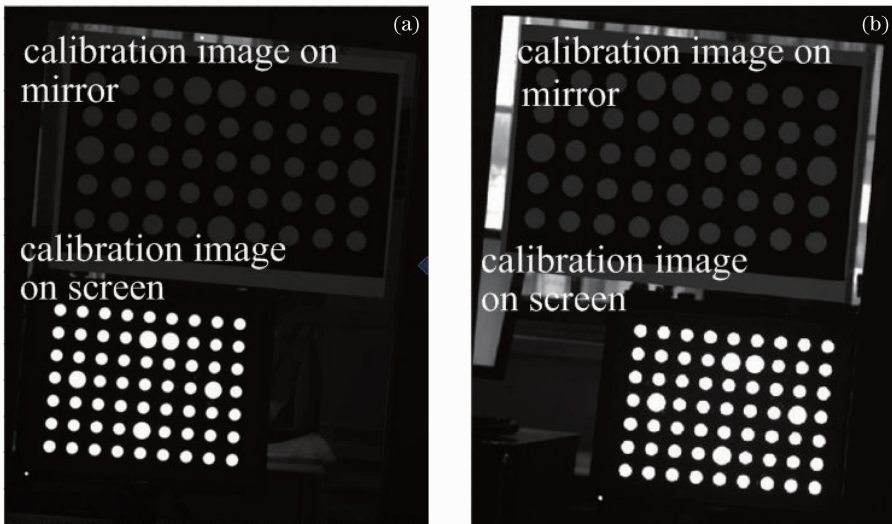


图 11 屏幕标定图像。(a)左摄像机图像;(b)右摄像机图像

Fig. 11 Calibration image of screen. (a) Image taken by left camera; (b) image taken by right camera

标定的过程如图 12 所示,首先通过在屏幕上设置 4 个坐标已知的视线点,然后让使用者带着头戴式标定板,在自然光条件和头部自然移动的情况下,分别看着这些坐标已知的视线点,然后通过摄像机获取眼睛虹膜中心和头部标定板上标志圆的圆心等特征信息,每个标定点视线停留 3 s 以上,左右摄像机分别拍摄 4 幅图像,4 个点共 16 对图像,接着用

提出的个人参数标定方法来获取使用者个人参数信息。

图 13 是通过虹膜定位算法分别在左右摄像机图像中定位到的眼睛虹膜,然后结合双摄像机标定结果,可以获得双眼虹膜中心的空间三维坐标,根据个人参数标定得到的眼球中心的三维坐标点以及眼睛光轴与视轴之间的补偿夹角,可以得到眼睛视线



图 12 个人参数标定场景图

Fig. 12 Scene graph of individual parameter calibration

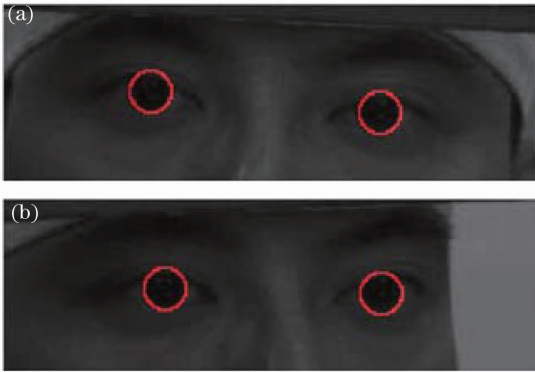

 图 13 虹膜定位。(a)左摄像机定位结果；
(b)右摄像机定位结果

Fig. 13 Iris location. (a) Iris location result of left camera image; (b) iris location result of right camera image

的空间向量,然后求取眼睛视线向量与屏幕空间平面方程的交点,最终取左右两眼视线估计的平均值作为视线估计值。

4.2 视线估计的结果与分析

在经过摄像机标定与屏幕空间位置标定后,选取 20 位测试者,距离显示器平面 800 mm 左右,显示器屏幕 19 inch (48.26 cm),分辨率为 1280 pixel × 1024 pixel。先使用提出的个人参数标定方法,求得个人参数后,在显示器屏幕上预先设定 9 个注视点,然后让每位测试者依次注视屏幕上的注视点,其中每个注视点至少注视 2 s 以上时间,双摄像机采集此时图像。数据采集完成后,在 Matlab2010a 平台上进行数据离线处理,根据系统标定结果,逆向求得眼睛视线估计值。图 14 给出视线估计结果,图中圆形表示屏幕上预先设定的视线点,十字形代表着经过计算得到的视线估计点,从图中可以看出,提出的视线跟踪算法精度较高,视线的估计值比较接近视

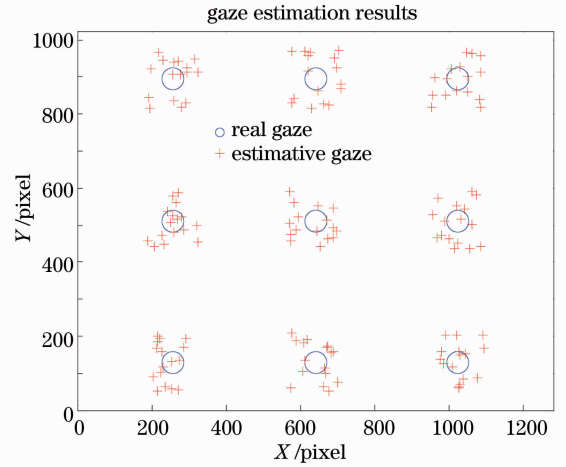


图 14 视线估计值与视线真实值比较

Fig. 14 Comparison between estimative gaze and real gaze

线真实值。

表 2 列出了部分测试者的视线估计误差,其中 ϵ_x, ϵ_y 分别代表视线估计值与真实视线值在屏幕平面水平和垂直方向上的误差,其计算公式为

$$\epsilon_x = \sqrt{\frac{1}{m} \sum_{i=1}^m (x_i - x'_i)^2}, \quad (29)$$

$$\epsilon_y = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - y'_i)^2}, \quad (30)$$

式中点 (x_i, y_i) 为视线估计点在屏幕平面中的坐标,点 (x'_i, y'_i) 为真实视线点在屏幕平面中的坐标。表 2 中括号内的数值为换算成角度来表示的视线跟踪误差,其计算公式为

$$\theta = \arcsin(\epsilon/d), \quad (31)$$

式中 d 为使用者到屏幕的距离, ϵ 为屏幕平面上水平或垂直方向上的误差。

表 2 使用者视线估计误差

Table 2 Gaze tracking errors of five users

	ϵ_x /pixel	ϵ_y /pixel
User 1	57.4 (1.22°)	63.8 (1.35°)
User 2	72.8 (1.54°)	65.2 (1.38°)
User 3	48.3 (1.02°)	55.2 (1.17°)
User 4	60.4 (1.28°)	69.6 (1.48°)
User 5	62.1 (1.32°)	59.4 (1.29°)

从表 2 可以看出,视线估计的精度在不同使用者之间存在着细微的差异,如第 3 个使用者的误差精度就比其他使用者的精度要高,其原因主要是双摄像机的焦距设定在 800 mm 左右,如果使用者向前或向后偏离摄像机范围过大,摄像机获得的图像会变得模糊,这会给系统特征参数的提取带来较大

的误差,因而影响到视线估计的精度,因此,在视线跟踪的使用过程中,为获取较清晰图像,头部活动范围为在摄像机光轴方向上 700~900 mm 之间。另外,双目视觉系统存在的测量误差^[20]和视线估计模型本身的误差也会给视线估计带来误差,同时,由于人眼视觉特性的复杂性,如眼球会有轻微的眼震和漂移,也会给视线估计带来不可避免的误差。最终的实验结果表明,构建的视线跟踪系统的视线估计精度约为 1.5°,能够满足很多视线跟踪应用场合的需要。

所构建的视线跟踪系统与其他视线跟踪系统的比较如表 3 所示,从表中可以看出,与基于自然光条件下的视线跟踪系统相比,提出的算法精度较高,视线估计平均精度接近于基于红外光的视线跟踪系统;与基于红外光的视线跟踪系统相比,此系统不需要红外光源,减少了硬件的限制,同时不受室外光线的影响,在室内室外都能使用。

表 3 视线跟踪系统比较

Table 3 Comparison of gaze tracking system

Method	Average gaze estimation accuracy / (°)	Amount of camera	Amount of IR lights
Reference[4]	1.5	2	1
Reference[6]	0.5	1	4
Reference[7]	3	1	0
Reference[10]	2	1	0
Reference[11]	1	2	2
Ours	1.5	2	0

5 结 论

对视线跟踪系统中的相关问题进行了研究,提出一种个人参数标定方法和一种头部姿态估计方法,基于上述算法,在双目立体视觉系统下,构建了一种自然光条件下,允许头部自然移动的视线跟踪系统。最后通过实验表明,本文提出的方法对使用者限制较少,可以允许头部在一定范围内自然移动,系统具有较高的视线估计精度,能满足很多视线跟踪应用场合的需要,对于自然光条件下视线跟踪的发展具有重要的参考价值。但同时系统还存在一定限制,如标定过程比较复杂,使用过程中使用者需要戴着头戴式标定板等。

未来还需要在以下几方面作进一步工作:1)使用云台承载具有自动变焦功能的摄像机,从而持续跟踪拍摄使用者,在进一步扩大使用者活动范围同时采集到具有较高清晰度的图像;2)为减少对使用

者的限制,研究人脸特征点的提取技术,通过人脸特征点的空间坐标变化来估计头部姿态;3)为进一步提高视线跟踪系统的精度,在眼睛特征参数的获取中使用亚像素边缘检测与中心定位方法。

参 考 文 献

- Z. Zeng, M. Pantic, G. Roisman *et al.*. A survey of affect recognition methods: audio, visual, and spontaneous expressions [J]. *IEEE Trans. on PAMI*, 2009, **31**(1): 39~58
- Liu Ruian. Study on Eye Gaze Tracking Using One Camera[D]. Tianjin: Tianjin University, 2007. 16~17
刘瑞安. 单摄像机视线跟踪技术研究[D]. 天津: 天津大学, 2007. 16~17
- Guojian Shao, Ming Che, Bingyi Zhang *et al.*. A novel simple 2D model of eye gaze estimation [C]. 2010 2nd International Conference on Intelligent Human-Machine Systems and Cybernetics(IHMSC), 2010, 300~304
- Zhiwei Zhu, Qiang Ji. Novel eye gaze tracking techniques under natural head movement [J]. *IEEE Trans. Biomedical Engineering*, 2007, **54**(12): 2246~2260
- J. J. Kang, M. Eizenman, E. D. Guestrin *et al.*. Investigation of the cross-ratios method for point-of-gaze estimation[J]. *IEEE Trans. Biomedical Engineering*, 2008, **55**(9): 2293~2302
- E. D. Guestrin, M. Eizenman. Remote point-of-gaze estimation with single-point personal calibration based on the pupil boundary and corneal reflections[C]. 2011 24th Canadian Conference on Electrical and Computer Engineering, 2011, 000971
- Jixu Chen, Qiang Ji. 3D gaze estimation with a single camera without IR illumination[C]. 19th International Conference on Pattern recognition, 2008, 1~4
- M. Reale, T. Hung, L. Yin. Viewing direction estimation based on 3D eyeball vonstruction for HRI[C]. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2010, 24~31
- Wen Zhang, Taining Zhang, Shengjiang Chang. Eye gaze estimation from the elliptical features of one iris[J]. *Opt. Eng.*, 2011, **50**(4): 0470031
- Jing Xie, Xueyin Lin. Gaze direction estimation based on natural head movements[C]. Fourth International Conference on Image and Graphics, 2007. 672~677
- Zhang Qiong, Wang Zhiliang, Chi Jiannan *et al.*. Calibration for gaze tracking system based on stereo cameras and plane mirror [J]. *Acta Optica Sinica*, 2011, **31**(4): 0415002
张琼, 王志良, 迟健男等. 基于平面镜面向双摄像机视线追踪系统的标定方法[J]. *光学学报*, 2011, **31**(4): 0415002
- E. D. Guestrin, M. Eizenman. General theory of remote gaze estimation using the pupil center and corneal reflections[C]. *IEEE Trans. Biomedical Engineering*, 2006, **53**(6): 1124~1132
- Jiang Guangwen, Chao Zhichao, Fu Sihua *et al.*. High-accurate camera calibration technique based on controllable rotation[J]. *Acta Optica Sinica*, 2010, **30**(5): 1308~1314
姜广文, 晁志超, 伏兴华等. 基于可控旋转的像机高精度标定技术[J]. *光学学报*, 2010, **30**(5): 1308~1314
- Zhu Jia, Li Xingfei, Xu Yingxin. Camera calibration technique based on active vision[J]. *Acta Optica Sinica*, 2010, **30**(5): 1297~1303
朱嘉, 李醒飞, 徐颖欣. 摄像机的一种主动视觉标定方法[J]. *光学学报*, 2010, **30**(5): 1297~1303
- Huo Ju, Yang Wei, Yang Ming. A self-calibration technique based on the geometry property of the vanish point[J]. *Acta Optica Sinica*, 2010, **30**(2): 465~472
霍炬, 杨卫, 杨明. 基于消隐点几何特性的摄像机自标定

- 方法[J]. 光学学报, 2010, **30**(2): 465~472
- 16 Xiang Zhiyu, Sun Bo. Mirror pose self-calibration method for catadioptric camera system[J]. *Acta Optica Sinica*, 2012, **32**(4): 0415001
- 项志宇, 孙 波. 折反射全向相机镜面位姿的自标定方法[J]. 光学学报, 2012, **32**(4): 0415001
- 17 Zhengyou Zhang. A flexible new technique for camera calibration [J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2000, **22**(11): 1330~1334
- 18 H. Kubota, M. Takeshi, H. Saito. 3D head pose tracking using a particle filter with nose template matching[J]. *IEEE Trans. Electronics, Information and Systems*, 2008, **128** (9): 1447~1454
- 19 K. Berthold, P. Horn. Closed-form solution of absolute orientation using unit quaternions[J]. *J. Opt. Soc. Am. A*, 1987, **4**(4): 629~642
- 20 Sun Nan, Zhang Liyan, Ye Nan *et al.*. Calibration of binocular stereo measurement system based on calibration cross[J]. *Acta Optica Sinica*, 2012, **32**(4): 0412001
- 孙 楠, 张丽艳, 叶 南等. 基于十字靶标的双目立体测量系统标定[J]. 光学学报, 2012, **32**(4): 0412001

栏目编辑: 李文喆