

一种基于标准化人眼图像的视线估计算法

任冬淳¹, 张 鹏², 王 敏¹, 古楠楠³

(1. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室, 北京 100190; 2. 民政部国家减灾中心, 北京 100124;
3. 首都经济贸易大学统计学院, 北京 100070)

摘 要: 本文提出了一种基于标准化人眼图像的视线估计算法, 能有效解决头部运动对视线估计的影响. 本算法在一个红外摄像机和三个红外光源的条件下, 利用三个光斑(光源在角膜表面的反射点)的图像信息, 将人眼图像变换到标准化人眼图像. 通过引入标准化人眼图像, 使屏幕平面到标准化人眼图像平面的射影变换为定值, 从而使标准化人眼图像的瞳孔中心和屏幕上的视点的变换关系为定值, 由此解决了头部运动对视线估计的影响. 实验结果验证了该算法在头部运动情况下的有效性.

关键词: 视线估计; 视线追踪; 标准化人眼图像

中图分类号: TP391.41

文献标识码: A

文章编号: 0372-2112 (2015)03-0536-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2015.03.019

A Gaze Estimation Method Based on Normalized Eye Image

REN Dong-chun¹, ZHANG Peng², WANG Min¹, GU Nan-nan³

(1. *State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China*; 2. *National Disaster Reduction Center of China, Beijing 100124, China*;
3. *School of Statistics, Capital University of Economics and Business, Beijing 100070, China*)

Abstract: In order to improve the effects of head movements on gaze estimation, this paper proposes a gaze estimation method based on normalized eye image. The proposed method is built on the installation of one infrared camera and three infrared light sources. We utilize the image of three light spots, which are the reflections of lights on cornea surface, to transform original eye image into the normalized one. By introducing normalized eye image, the projective transformation from screen plane to normalized eye image plane is constant. Consequently, the transformation from the pupil center in a normalized eye image to the viewpoint on the screen is fixed. Therefore, the effects caused by head movements can be eliminated. Experimental results demonstrated the effectiveness of the proposed method under various conditions of head movements.

Key words: gaze estimation; gaze tracking; normalized eye image

1 引言

视线追踪是利用眼动测量设备测量眼睛的运动情况估计视线方向或者视线落点位置的技术. 通过对视线直接、准确的估计, 视线追踪技术在人机交互、医学诊断和残疾人康复等领域有广泛的应用^[1], 在虚拟现实和智能感知领域也有很好的应用前景.

现有的测量眼睛运动情况的方法主要有: 机械记录法、电磁记录法、电流记录法和光学记录法^[1]. 其中机械记录法和电磁记录法需要在眼睛上附着辅助设备用于测量眼睛运动; 电流记录法需要在眼睛周围放置电极用于测量眼睛运动; 而光学记录法只需要用照相机或摄像

机记录用户的眼睛运动情况. 这种基于眼睛图像信息的视线追踪方法(光学记录法)对硬件复杂度要求相对较低, 而且用户体验更好, 因此被广泛使用.

2 相关方法综述

现有的基于眼睛图像信息的视线追踪方法主要包括两个步骤: 相关特征提取和视线估计. 常用的特征主要包括: 瞳孔轮廓、瞳孔中心、虹膜轮廓、虹膜中心、眼角以及外界光源(一般为红外光源)在角膜上的反射点(普金野图像). 现有的视线估计算法主要包括: (1) 瞳孔-角膜反射法^[2~8]、(2) 虹膜-巩膜边缘法^[9]和(3) 基于外观的方法^[10,11].

(1) 瞳孔-角膜反射法是利用瞳孔和红外光源(选取红外光源是因为人类眼睛对红外光不敏感)在角膜上的反射点的图像信息进行视线估计的方法.该方法首先需要检测瞳孔中心的位置以及光源在角膜上反射点(光斑)的位置.为了准确检测瞳孔的位置,一些算法使用了亮暗瞳技术^[3].基于瞳孔中心和角膜反射点的视线估计方法主要有:瞳孔中心与角膜反射向量法^[2-6]、交比不变性方法^[7]、基于眼球三维模型的视线估计方法^[8]和双普金野法.

(a) 基于瞳孔中心与角膜反射向量的视线估计方法. Gale 提出当眼球的旋转中心静止,光源和摄像机位置也静止的时候,随着视点的变化,眼球发生转动,瞳孔中心与光源在角膜的反射点构成的向量也发生变化.瞳孔中心与光源在角膜的反射点构成的向量与注视平面上的视点一一对应^[2].因此,可以构建瞳孔中心角膜反射向量到注视平面上视点的映射函数,从而进行视线估计.文献[3,4]使用了多项式映射函数,其中文献[3]采用了线性映射函数,文献[4]采用了二次多项式映射函数.文献[5]采用了神经网络的方法构建映射函数.此外, Broly 和 Mulligan 尝试在头动的情况下,用构建映射函数的方法进行视线估计^[6].

(b) 基于交比不变性的视线估计方法. Yoo 等人提出利用射影几何中交比不变的性质进行视线估计^[7].该方法在摄像机光轴附近设置一个红外光源用于检测瞳孔,在显示屏四角设置四个红外光源用于视线估计,用于视线估计的四个光源的位置需要提前测量得到.由于显示屏四个光源和视点与眼睛图像上四个光斑和瞳孔中心的对应交比相等,于是显示屏四个光源和视点相应的交比可以通过眼睛图像中四个光斑和瞳孔中心的位置计算.视点可以通过显示屏四个光源和视点相应的交比和四个光源的位置计算得到.交比不变性方法的优点是不需要用户标定,但需测量四个光源点的位置.此外,该算法没有考虑视轴和光轴的夹角,也忽略了瞳孔中心与四个角膜反射点不共面的问题,因此有一定的误差.

(c) 基于眼球三维模型的视线估计方法.该方法是人对人眼的物理结构进行建模,根据模型计算视线的三维向量从而进行视线估计.文献[8]提出了一种基于眼球三维模型的视线估计方法,该方法通过已知光源的三维位置信息、摄像机的内外参数及三维位置信息、显示屏的大小和三维位置信息以及眼睛参数(眼球半径、角膜半径、眼球中心到角膜中心的距离、房水与角膜之间的折射率等),结合眼睛的图像信息,计算出角膜中心和瞳孔中心的三维坐标,从而得到光轴的直线方程,计算出视点.该方法实时计算角膜中心和瞳孔中心的三维坐标,因此对头动是鲁棒的,然而算法需要已

经测量的环境信息和已经标定的摄像机,并且精度要求较高,对不同用户眼睛参数的标定过程也较为复杂.

(d) 基于双普金野图像的视线估计方法.光线进入眼睛,在空气与角膜交界面、角膜与房水交界面、房水与晶状体交界面和晶状体和玻璃体交界面分别形成普金野图像(反射点).当视线发生变化时,第一和第四普金野图像均发生移动,同时他们之间的距离发生了变化,由此可以进行视线估计.双普金野法就是通过两个普金野图像的位置关系进行视线估计的方法.该方法准确度较高,但是第四普金野图像难以检测,需要专业的摄像机才能进行捕捉,价格昂贵.

(2) 虹膜-巩膜边缘法是通过检测虹膜和巩膜交界的圆在相机拍摄的图片中呈现的大小和椭圆形状,来估计视线方向.由于虹膜和巩膜交界的圆的上半部分和下半部分分别被上眼皮和下眼皮遮挡,眼睫毛也会对虹膜和巩膜交界圆造成部分遮挡,因此对图像中虹膜和巩膜交界圆的准确检测较为困难.这对视线估计的准确性会造成较大影响.

(3) 基于外观的方法是对眼睛的整体图片进行处理并估计视线方向的方法.该算法并不提取眼睛图像的瞳孔、光源的角膜反射点等特征,而是通过训练得到整个眼睛图像与视点的对应关系.文献[10]用 2000 张眼睛图像和对应的视点作为训练样本,提出一种基于神经网络的视线估计方法,文献[11]采用流形学习的方法,对高维的眼睛图像学习了一个低维流形,对 252 张眼睛图像和对应的视点进行训练,达到较好的视线估计效果.基于外观的视线估计方法运算复杂度低,实时效果非常好,但需要大量的训练样本,也就是需要大量的标定点,这给用户使用带来不便.

现有的基于眼睛图像信息的视线追踪算法适用环境的对比如表 1 所示,现有方法的主要问题有:(1)头部运动(头部相对照相机运动)带来的视线追踪不准确;(2)需要标定相机;(3)需要测量环境距离;(4)需要较多的标定点;(5)需要专业、昂贵的硬件装置.基于上述问题,本文提出了一种基于标准化人眼图像的视线估计

表 1 6 种主要视线估计方法适用环境的比较

视线估计方法	头动影响	相机标定	环境距离测量	大量标定点	专业相机
瞳孔角膜反射向量法	有影响	不需要	不需要	不需要	不需要
交比不变性法	前后运动有影响	不需要	需要	不需要	不需要
眼球三维模型法	没有影响	需要	需要	不需要	不需要
双普金野图像法	有影响	不需要	不需要	不需要	需要
虹膜-巩膜边缘法	有影响	不需要	不需要	不需要	不需要
基于外观的方法	有影响	不需要	不需要	需要	不需要

方法,该方法在头部运动的情况下,仍能准确的估计视线,而且对硬件要求简单(一个摄像机和三个红外光源),不需要对摄像机进行标定,也不需要光源、摄像机、屏幕位置进行测量,并且最少只需要四个标定就可完成标定过程.

3 基于标准化人眼图像的视线估计算法

头部运动情况下的视线估计,是视线追踪的关键问题之一.本章提出了一种基于标准化人眼图像的视线估计算法,通过引入标准化眼睛图像,使屏幕平面到标准化眼睛图像平面的射影变换为定值,从而使标准化眼睛图像的瞳孔中心和屏幕上的视点的变换关系为定值,由此解决了头部运动对视线估计的影响.同时,基于标准化人眼图像的视线估计算法不需要标定摄像机,也不需要测量系统中光源和摄像机的位置,便于系统搭建.下面将从眼睛的光学原理、系统描述、图像特征提取和视线估计算法四个方面详细介绍.

3.1 眼睛的光学原理

眼睛是一个非常复杂的光学系统,光线要经过角膜、房水、晶状体和玻璃体的几次折射后,才能到达视网膜,同时在它们之间的每个交接面上形成反射点,包括角膜的前表面(空气和角膜的接触面)、角膜的后表面(角膜和房水的接触面)、晶状体的前表面(房水和晶状体的接触面)和晶状体的后表面(晶状体和玻璃体的接触面).角膜的主要作用是保护眼睛,同时具有很强的折射能力.晶状体通过放松和收缩睫状肌来改变自身的形状,进而将影像聚焦在视网膜上.光线在到达晶状体之前要先经过瞳孔,瞳孔用于控制到达视网膜的光线强度.由于角膜和晶状体均有折射能力,因此在眼睛这个光学系统的光轴上,实际上存在两个节点,由于此二者的距离只有 0.3mm ,且它们的位置又非常接近角膜中心,因此有关视线估计的研究^[8,12]将其近似为一个点.当一个视力正常的人用眼睛检测一个物体时,物体的图像会落在其每只眼睛的视网膜中央窝上,眼睛就是通过这种方式来对物体加以定位的.视网膜中央窝是视网膜最为敏锐的区域,角膜中心与视网膜中央窝形成的视轴与角膜中心与瞳孔中心形成的光轴在角膜中心处有 $0.6^\circ \sim 1^\circ$ 的视觉角度差^[12].由于光轴容易计算,一般视线估计算法往往先计算光轴,通过光轴与视轴的夹角补偿计算得到视轴,进而计算视点.

3.2 系统描述

本文提出的系统包含三个红外光源和一个红外摄像机,如图1所示.三个光源用于在角膜上形成反射点,摄像机位于显示屏下方,捕捉眼睛图像.本文提出的视线估计算法不需要对光源、摄像机位置进行测量,对光源的位置也不需要特别设定,只要光源可以在角膜上

形成反射点即可.因此,本文提出的算法便于进行系统搭建.

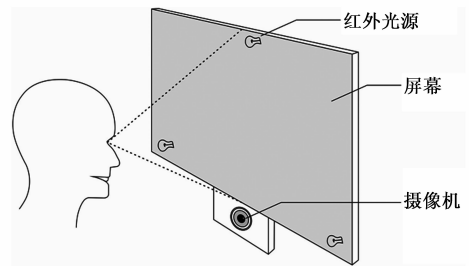


图1 硬件系统结构图

3.3 图像特征提取

为了进行视线估计,首先要对眼睛图像进行特征提取.本文提出的基于标准化人眼图像的视线估计算法需要提取眼睛图像中的三个角膜反射点和瞳孔中心的位置信息.由于本文提出的算法是可以解决头部运动情况下的视线估计不准确的问题,因此要求摄像机在头部发生运动后仍能捕捉眼睛图像.为此摄像机的监控范围相对较广,摄像机捕捉到的图像如图3(a)所示.图像特征提取的过程包括:(1)眼睛区域选取;(2)角膜反射点检测;(3)瞳孔中心检测.图像特征提取算法流程图如图2所示.眼睛区域选择的过程如图3所示,角膜反射点及瞳孔中心检测的过程如图4所示.

3.3.1 眼睛区域的选取

眼睛区域的选取是从摄像机捕捉到的图像中选出眼睛所在的区域,以便对眼睛图像进行特征提取.由于两只眼睛的视轴相交于视点,为了减少算法运算量,只选取一只眼睛的区域便可进行视线估计.本文采用检测和跟踪相结合的方法选取眼睛区域.这是由于对整个摄像机捕捉到的图像进行人眼区域检测的运算量较

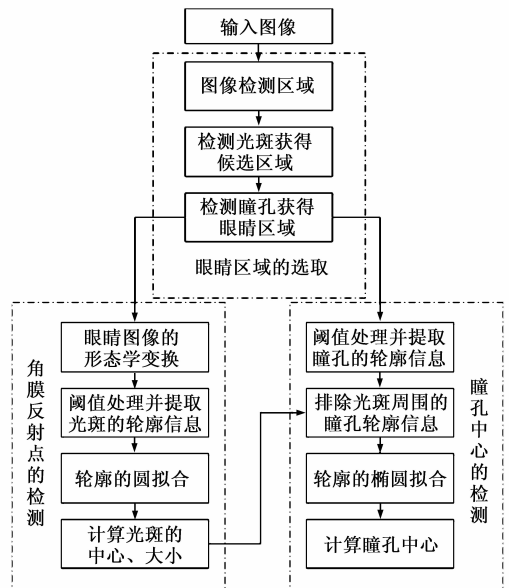
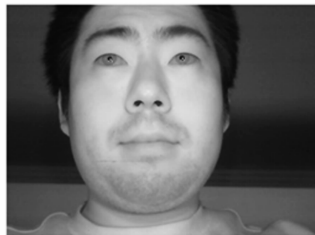
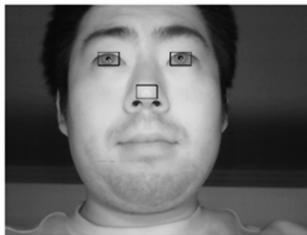


图2 图像特征提取算法流程图

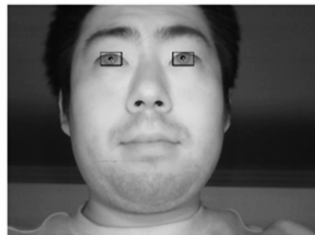
大,而即便眼睛移动,相邻帧的眼睛在图像中的位移也不会很大,在上一帧检测到的眼睛区域周围对眼睛进行搜索,将有效地减小运算量.此外,眨眼等情况的发生会影响跟踪方法对眼睛区域的选择,此时可以重新使用检测的方法在整个摄像机捕捉到的图像中检测眼



(a) 摄像机捕捉的灰度图像



(b) 根据光斑获得候选区域



(c) 根据瞳孔获得眼睛区域

图3 眼睛区域选择过程图

3.3.2 角膜反射点的检测

红外光源在角膜进行反射,在摄像机捕捉的图像中呈现白色,与周围像素差异较大,因此相对容易检测,使用阈值的方法便可检测到眼睛图像中的角膜反射点.但是由于角膜表面的不平整,一个光源在角膜反射后可能形成两个距离非常近的白色光斑,因此首先对眼睛图像进行形态学变换(腐蚀变换和膨胀变换各一次),再对图像进行阈值处理,便可检测到光斑区域.提取每个光斑的轮廓信息,对轮廓进行圆的拟合,拟合结果如图4所示.圆的中心即为光斑位置,圆的半径用来表示光斑大小.

3.3.3 瞳孔中心的检测

瞳孔的检测比角膜反射点的检测复杂,主要是由于角膜反射点的存在影响瞳孔中心的检测.为了排除角膜反射点对瞳孔中心计算的影响,本算法首先对眼睛图像进行阈值处理,其次提取瞳孔的轮廓信息,再次剔除光斑范围内的瞳孔轮廓信息,最后用剩余的轮廓进行椭圆拟合,瞳孔拟合结果如图4所示,椭圆中心即为瞳孔中心.

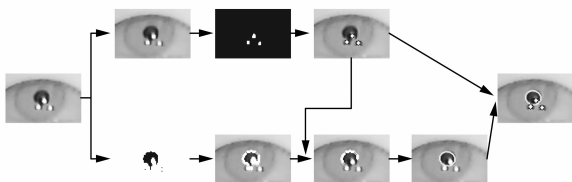


图4 眼角膜反射点及瞳孔中心检测的过程图

3.4 视线估计算法

本文提出了一种基于标准化人眼图像的视线估计算法.通过引入标准化人眼图像,使屏幕平面到标准化眼睛图像平面的射影变换为定值,从而使标准化眼睛图像的瞳孔中心和屏幕上的视点的变换关系为定值,由此解决了头部运动对视线估计的影响.

睛区域.眼睛区域的特点是有白色的光斑和黑色的瞳孔,其中白色光斑的强度较高.首先通过动态阈值来选取指定区域中具有白色的候选区域,然后在候选区域中动态阈值检测是否有黑色像素点,如果存在黑色像素点则认为该候选区域是眼睛区域.

算法原理如图5所示,红外光源 L_1, L_2, L_3 发出光线,在角膜表面形成反射点 V_1, V_2, V_3 并构成平面.通过摄像机捕捉到的角膜反射点 V_1, V_2, V_3 是红外光源 L_1, L_2, L_3 通过角膜表面的虚像.由于角膜表面的焦距约为 3.9mm ,角膜半径约为 7.8mm ,因此角膜中心到虚像平面 $V_1V_2V_3$ 的距离为 3.9mm ^[2].同时,角膜中心到瞳孔的距离约为 4.2mm ^[2],并且角膜中心与瞳孔中心的连线垂直于虚像平面,因此可以近似认为瞳孔中心 P 也位于平面 $V_1V_2V_3$ 上.摄像机捕捉眼睛图像,角膜反射点 $V_1V_2V_3$ 和瞳孔中心 P 在图像平面上分别形成 $U_{V_1}, U_{V_2}, U_{V_3}$ 和 U_P .摄像机平面经过一个仿射变换 N ,将其映射到标准化图像上形成 $N_{V_1}, N_{V_2}, N_{V_3}$ 和 N_P .这里仿射变换 N 是使 $U_{V_1}, U_{V_2}, U_{V_3}$ 映射到 $N_{V_1}, N_{V_2}, N_{V_3}$ 的变换.设 U_{V_i} ($i=1,2,3$) 的坐标为 (x_{ui}, y_{ui}) , N_{V_i} ($i=1,2,3$) 的坐标为 (x_{ni}, y_{ni}) ,则仿射变换 N 满足:

$$\begin{pmatrix} x_{ni} \\ y_{ni} \\ 1 \end{pmatrix} = \begin{bmatrix} n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} x_{ui} \\ y_{ui} \\ 1 \end{pmatrix} \quad (1)$$

由于图像平面上的点 U_{V_i} ($i=1,2,3$) 的坐标通过第3.3节介绍的方法求得,标准化图像上点 N_{V_i} ($i=1,2,3$) 的坐标是确定的.因此通过3个对应点,可以建立6个方程6个未知数的线性方程组,求解该线性方程组得到仿射变换 N .于是,像平面上瞳孔中心 U_P 经过 N 的变换,可得标准化图像上的瞳孔中心 N_P .设 U_P 的坐标为 (x_{up}, y_{up}) , N_P 的坐标为 (x_{np}, y_{np}) ,则有:

$$\begin{pmatrix} x_{np} \\ y_{np} \\ 1 \end{pmatrix} = \begin{bmatrix} n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} x_{up} \\ y_{up} \\ 1 \end{pmatrix} \quad (2)$$

如图5所示,屏幕平面 $L_1L_2L_3$ 经过射影变换形成平面 $V_1V_2V_3$,假设该射影变换为 H_1 .平面 $V_1V_2V_3$ 经过射影变换形成平面 $U_{V_1}U_{V_2}U_{V_3}$,假设该射影变换为 H_2 ,

成的向量与注视平面上的视点一一对应的原理.为了更有效地对比两种算法的准确性,基于瞳孔角膜反射向量算法与基于标准化人眼图像的视线估计算法采用相同的图像处理结果,并采用三个光斑中的上方光斑与瞳孔中心构建瞳孔角膜反射向量(v_x, v_y).映射函数按照文献[4]选择二次多项式函数.

$$J_x = 1 + a_1 v_x + a_2 v_y + a_3 v_x v_y + a_4 v_x^2 + a_5 v_y^2 \quad (6)$$

$$J_y = 1 + b_1 v_x + b_2 v_y + b_3 v_x v_y + b_4 v_x^2 + b_5 v_y^2 \quad (7)$$

本文提出的基于标准化人眼图像的视线估计算法最少需要 4 个标定.采用二次多项式拟合的瞳孔角膜反射向量算法最少标定个数随着二次多项式的形式的变化而有所不同,根据文献[4]提出的式(6)和式(7)的二次多项式进行拟合,最少需要 5 个标定.为此,本文在两种算法最少标定和 9 个标定两种情况下进行实验.在两种情况下,针对头部不运动和头部运动(实验人员在 10cm 以内进行前后运动,在 15cm 以内进行上下左右运动)分别进行了实验.实验 1 是在头部不运动的情况下进行的,在该实验中实验人员距离屏幕的距离 D 为 50cm.实验 2 是头部在标定过程不运动,在测试过程运动,实验人员头部与屏幕的距离范围为 40cm ~ 50cm,并且头部可以在该范围内自由运动.实验 3 是头部在标定和测试过程均运动,实验人员头部与屏幕的距离范围为 40cm ~ 60cm,并且头部可以在该范围内自由运动.在实验 2 中,实验人 1 在测试点 13 记录的数据发生异常(有可能实验人员正在快速运动造成摄像机捕捉图像不清晰所致),实验人 2 在测试点 15 数据发生异常.因此实验 2 中测试点 13 和 15 的平均视线偏差角分别是由其他 5 名实验人采集到的正常数据取均值后得到的.

4.1 4 点标定下标准化人眼图像法和 5 点标定下瞳孔角膜反射向量法的对比

本节对基于标准化人眼图像方法和瞳孔角膜反射向量法在最少标定下进行实验.该实验用于标准化人眼图像视线估计算法的 4 个标定点为矩形的四个顶点,用于瞳孔角膜反射向量法的 5 个标定点为上述矩形的四个顶点及矩形对角线交点.两算法分别在(1)头部不运动,(2)头部在标定过程不运动在测试过程运动,(3)头部在标定、测试过程均运动三种情况下分别进行实验,在 16 个测试点的平均视线偏差角如表 2 和表 3 所示.在头部不运动情况下(实验 1),本文提出的标准化人眼图像方法在 16 个测试点中的 13 个优于瞳孔角膜反射向量法;在头部在测试过程运动的情况下(实验 2),本文提出的标准化人眼图像方法在 16 个测试点中的 15 个优于瞳孔角膜反射向量法;在头部在标定、测试过程均运动的情况下(实验 3),本文提出的标准化人眼

图像方法在 16 个测试点中全部优于瞳孔角膜反射向量法.可见本文提出算法的有效性,在头部运动情况下算法准确度优势显著.

表 2 4 个标定下标准化人眼图像法在 16 个测试点的平均视线偏差角(单位:°)

测试点	1	2	3	4	5	6	7	8
实验 1	1.97	1.53	1.37	1.87	0.81	0.87	0.82	1.57
实验 2	1.97	1.24	1.19	2.47	1.20	0.52	1.11	1.24
实验 3	2.05	1.22	1.23	1.89	1.04	1.23	1.90	1.22
测试点	9	10	11	12	13	14	15	16
实验 1	0.74	0.52	0.65	1.66	0.95	0.64	0.87	0.77
实验 2	1.00	0.57	1.17	1.05	1.54	0.76	0.84	1.15
实验 3	0.76	1.70	1.29	1.55	0.95	1.20	1.55	1.28

表 3 5 个标定下瞳孔角膜反射向量法在 16 个测试点的平均视线偏差角(单位:°)

测试点	1	2	3	4	5	6	7	8
实验 1	3.51	1.24	1.18	3.64	2.02	1.61	1.92	2.56
实验 2	3.42	3.25	6.46	2.00	9.04	2.68	8.16	6.55
实验 3	16.53	11.37	20.25	20.15	9.92	6.37	15.84	9.92
测试点	9	10	11	12	13	14	15	16
实验 1	1.84	1.03	1.26	1.62	2.52	1.14	1.75	2.90
实验 2	2.90	7.42	10.01	3.74	2.85	1.75	1.84	9.43
实验 3	17.91	9.85	10.37	11.14	7.13	21.34	21.40	6.72

4.2 9 点标定下标准化人眼图像法和瞳孔角膜反射向量法的对比

本节对基于标准化人眼图像方法和瞳孔角膜反射向量法均在 9 个标定下进行实验,9 个标定按照 3 × 3 的分布排列.两算法分别在(1)头部不运动,(2)头部在标定过程不运动在测试过程运动,(3)头部在标定、测试过程均运动三种情况下分别进行实验,6 名实验人员在 16 个测试点的平均视线偏差角如表 4 和表 5 所示.在头部不运动情况下(实验 1),本文提出的标准化人眼图像方法在 16 个测试点中的 13 个优于瞳孔角膜反射向量法;在头部在测试过程运动的情况下(实验 2),本文提出的标准化人眼图像方法在 16 个测试点中的 15 个优于瞳孔角膜反射向量法;在头部在标定、测试过程均运动的情况下(实验 3),本文提出的标准化人眼图像方法在 16 个测试点中全部优于瞳孔角膜反射向量法.可见本文提出算法的有效性,在头部运动情况下算法准确度优势显著.图 6 ~ 8 绘制了 6 名实验人员的在两种算法下计算的视点(图中用 '+' 表示)和目标注视点(图中用实心圆表示).图 6 绘制了 6 名实验人员在头部不运动的情况下,两种算法计算的视点和目标注视

点.图7绘制的是头部在标定过程不运动在测试过程运动情况下的6名实验人员在两种算法计算的视点和目标注视点.图8绘制的是头部在标定和测试过程均运动情况下的6名实验人员在两种算法计算的视点和目标注视点.图6~8可以更加直观的反应本文提出的标准化人眼图像方法优于瞳孔角膜反射向量法.

表4 9个标定点下标准化人眼图像法在16个测试点的平均视线偏差角(单位:°)

测试点	1	2	3	4	5	6	7	8
实验1	1.87	1.29	1.47	1.81	0.81	0.80	0.88	1.21
实验2	2.23	1.72	1.20	2.16	0.94	0.75	0.60	0.89
实验3	2.04	1.19	0.88	2.34	0.74	0.76	1.15	1.44
测试点	9	10	11	12	13	14	15	16
实验1	0.58	0.52	0.54	1.38	0.82	0.56	0.67	0.47
实验2	1.05	0.76	0.83	0.72	1.20	0.64	0.62	0.83
实验3	0.69	1.12	0.65	0.93	1.07	0.60	1.24	1.23

表5 9个标定点下瞳孔角膜反射向量法在16个测试点的平均视线偏差角(单位:°)

测试点	1	2	3	4	5	6	7	8
实验1	3.03	1.81	1.36	1.88	1.31	1.39	0.84	1.52
实验2	1.91	2.77	4.34	2.86	4.42	1.73	9.05	2.24
实验3	6.12	3.15	3.76	2.70	4.31	2.93	2.92	2.72
测试点	9	10	11	12	13	14	15	16
实验1	1.08	0.76	1.58	1.28	1.52	1.19	1.42	2.71
实验2	2.58	4.03	3.80	2.30	3.19	1.71	1.19	7.91
实验3	6.93	3.01	1.63	1.75	4.38	6.83	4.84	5.69

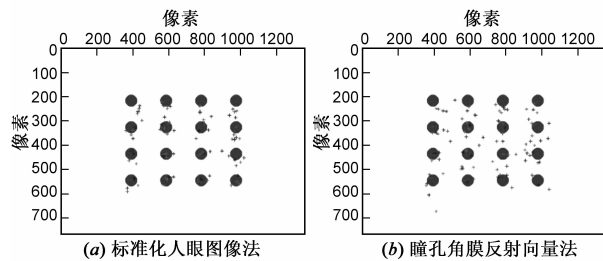


图6 头部不运动情况下两种算法对比

两种算法在最少标定点和9个标定点两种情况下,在(1)头部不运动,(2)头部在标定过程不运动在测试过程运动,(3)头部在标定、测试过程均运动的平均视线偏差角如表6所示.基于标准化人眼图像算法在采用4点标定时算法的平均视线偏差角为1.10°(头部不动),1.19°(头部在测试过程运动)和1.38°(头部在标定、测试过程均运动).当眼睛位于屏幕前方50cm处对应的视点偏差分别为0.96cm,1.04cm和1.20cm.在采用9点

标定时算法的平均视线偏差角为0.98°(头部不动),1.07°(头部在测试过程运动)和1.13°(头部在标定、测试过程均运动).当眼睛位于屏幕前方50cm处对应的视点偏差分别为0.86cm,0.93cm和0.99cm.上述实验结果表明,本文提出的基于标准化人眼图像的视线估计算法在头部运动和不运动两种情况下均优于瞳孔角膜反射向量法.在头部运动的情况下,基于标准化人眼图像算法的准确率变化不大,验证了该算法在头部运动下视线估计的有效性.

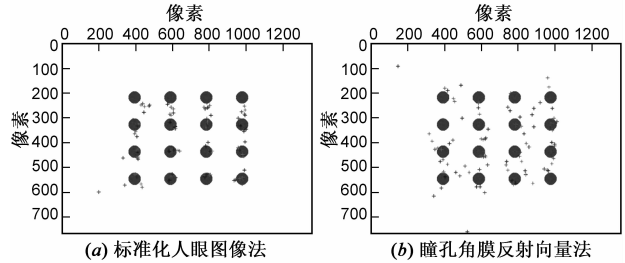


图7 测试过程中头部运动情况下两种算法对比

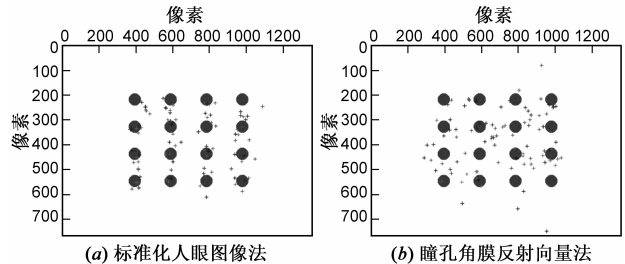


图8 标定和测试过程中头部均运动情况下两种算法对比

表6 两种算法的平均视线偏差角(单位:°)

	标准化人眼图像法		瞳孔角膜反射向量法	
	4点标定	9点标定	5点标定	9点标定
头部不运动	1.10	0.98	1.99	1.54
头部运动(测试)	1.19	1.07	5.09	3.50
头部运动(标定、测试)	1.38	1.13	13.51	3.98

5 讨论

本文提出了一种对头部运动鲁棒的视线估计算法.该算法只需一个摄像机和三个红外光源,并且不需要测量环境信息,也不需要标定摄像机.通过三个红外光源在角膜反射点的图像信息,构建了标准化人眼图像,进而进行视线估计.通过构建标准化人眼图像,使得屏幕平面与标准化人眼图像平面之间的射影变换为定值,从而解决头部运动对视线估计带来的干扰.在图像处理阶段,本文提取了三个光斑中心和瞳孔中心,在瞳孔椭圆的拟合过程中,排除了光斑中心附近的瞳孔轮廓,从而使瞳孔拟合更加准确.本文提出的基于标准化人眼图像的视线估计算法同样适用于其他眼睛特征

(如:虹膜中心)和其他图像处理方法(文献[13,14]关于眼睛图像的处理做了大量的研究,为不同任务下提取不同图像特征用于本文算法进行视线估计提供了帮助).对比实验的结果也表明了本文提出的基于标准化人眼图像的视线估计算法在头部不运动和(标定以及测试过程中)头部运动情况下视线估计的有效性.

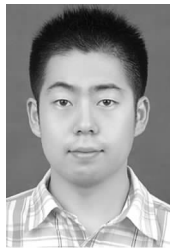
基于人眼图像法是通过一系列射影变换(如图5所示),将屏幕平面变换到标准化人眼图像平面.该算法还有一些可以改进的地方:(1)瞳孔中心在角膜表面形成的折射点与3个光源在角膜表面形成的反射点并不共面,瞳孔中心的折射点距离3个光源的反射点构成平面的距离约为0.3mm.该距离对视线估计的影响以及补偿方法将是下一阶段的工作.解决这一问题后算法的准确性还能得到一定的提高.(2)视线估计时视点的抖动问题,这主要是由于图像处理时,对瞳孔拟合的椭圆的波动造成的,通过提高摄像机像素和改进图像处理算法将改善这一问题.

参考文献

- [1] Duchowski A. Eye Tracking Methodology: Theory and Practice [M]. Second Edition. London: Springer, 2007.
- [2] Gale A. A note on the remote oculometer technique for recording eye movements[J]. Vision Research, 1982, 22(1): 201 - 202.
- [3] Hutchinson T, White K, Martin W, Reichert K, Frey L. Human-computer interaction using eye-gaze input[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1989, 19(6): 1527 - 1534.
- [4] Morimoto C, Mimica M. Eye gaze tracking techniques for interactive applications[J]. Computer Vision and Image Understanding, 2005, 98: 4 - 24.
- [5] Zhu Z, Ji Q. Eye and gaze tracking for interactive graphic display[J]. Machine Vision and Applications, 2004, 15(3): 139 - 148.
- [6] Brolly X, Mulligan J. Implicit calibration of a remote gaze tracker[A]. Proceedings of Conference on Computer Vision and Pattern Recognition Workshop[C]. Washington, DC: IEEE Computer Society, 2004. 134.
- [7] Yoo D, Kim J, Lee B, Chung M. Non-contact eye gaze tracking system by mapping of corneal reflections[A]. Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition[C]. Washington, DC: IEEE Computer Society, 2002. 94 - 99.

- [8] Guestrin E, Eizenman M. General theory of remote gaze estimation using the pupil center and corneal reflections[J]. IEEE Transactions on Biomedical Engineering, 2006, 53(6): 1124 - 1133.
- [9] Wang J, Sung E, Venkateswarlu R. Estimating the eye gaze from one eye[J]. Computer Vision and Image Understanding, 2005, 98(1): 83 - 103.
- [10] Baluja S, Pomerleau D. Non-intrusive Gaze Tracking Using Artificial Neural Networks [R]. Pittsburgh: Department of Computer Science, Carnegie Mellon University, 1994.
- [11] Tan K, Kriegman D, Ahuja N. Appearance-based eye gaze estimation[A]. Proceedings of IEEE Workshop on Applications of Computer Vision[C]. Orlando: IEEE Computer Society, 2002. 191 - 195.
- [12] Young L, Sheena D. Methods and designs—Survey of eye movement recording methods[J]. Behavioral Research Methods and Instructions, 1975, 7(5): 397 - 429.
- [13] 张闯,迟健男,张朝晖,王志良. 视线追踪系统中眼睛跟踪方法研究[J]. 自动化学报, 2010, 36(8): 1051 - 1061. Zhang C, Chi J, Zhang Z, Wang Z. The research on eye tracking for gaze tracking system [J]. Acta Automatica Sinica, 2010, 36(8): 1051 - 1061. (in Chinese)
- [14] Hansen D, Ji Q. In the eye of the beholder: A survey of models for eyes and gaze[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(3): 478 - 500.

作者简介



任冬淳 男, 1985 年生于天津. 现为中国科学院自动化研究所博士研究生. 主要研究方向为视线追踪、计算机视觉和智能机器人.

E-mail: dongchun.ren@gmail.com



张鹏 男, 1984 年出生于山东济南. 现为民政部国家减灾中心助理研究员. 主要研究方向为机器学习和信息系统.

E-mail: zhangpeng@ndrcc.gov.cn