基于特征的视线跟踪方法研究综述

刘佳惠 [1] 迟健男 [1, 2] 尹怡欣 [1]

摘 要 针对基于特征的视线跟踪方法进行了综述. 首先对视线跟踪技术的发展、相关研究工作和研究现状进行了阐述; 然后将基于特征的视线跟踪方法分成了两大类: 二维视线跟踪方法和三维视线跟踪方法, 从硬件系统配置、误差主要来源、头部运动影响、优缺点等多个方面重点分析了这两类视线跟踪方法, 对近五年现有的部分基于特征的视线跟踪方法进行了对比分析, 并对二维视线跟踪系统和三维视线跟踪系统中的几个关键问题进行了探讨; 此外, 介绍了视线跟踪技术在人机交互、医学、军事、智能交通等多个领域的应用; 最后对基于特征的视线跟踪方法的发展趋势和研究热点进行了总结与展望.

关键词 基于特征, 视线跟踪, 二维映射模型, 三维视线估计

引用格式 刘佳惠, 迟健男, 尹怡欣. 基于特征的视线跟踪方法研究综述. 自动化学报, 2021, **47**(2): 252−277

DOI 10.16383/j.aas.c180844

A Review of Feature-based Gaze Tracking Methods

LIU Jia-Hui[1] CHI Jian-Nan[1, 2] YIN Yi-Xin[1]

Abstract In this paper, the feature-based gaze tracking methods are reviewed. Firstly, the development, the related works and the research status of gaze tracking technology are introduced. Then, the feature-based gaze tracking methods are classified into two categories: Two-dimensional gaze tracking methods and three-dimensional gaze tracking methods. These two kinds of gaze tracking methods are emphatically analyzed in terms of hardware system configuration, main error sources, head movement, advantages and disadvantages, etc. Some of the existing feature-based gaze tracking methods in recent five years are compared, and several key issues in two-dimensional gaze tracking systems and three-dimensional gaze tracking systems are discussed. Moreover, the applications of gaze tracking technology in many fields, such as human-computer interaction, medicine, military, intelligent transportation, etc. are introduced. Finally, the development trends and research topics of feature-based gaze tracking methods are summarized.

Key words Feature-based, gaze tracking, two-dimensional mapping model, three-dimensional gaze estimation

Citation Liu Jia-Hui, Chi Jian-Nan, Yin Yi-Xin. A review of feature-based gaze tracking methods. *Acta Automatica Sinica*, 2021, **47**(2): 252−277

视线跟踪是利用机械、电子、光学等现有的各种检测手段获取受试者当前 “注视方向” 的技术[1]. 视线跟踪的研究起源于有关眼睛运动规律的研究, 早期心理学家是用肉眼或借助简单的仪器来观察眼动情况, 如: 观察法、机械记录法、胶片记录法等, 后来逐渐探索出了电流记录法、电磁线圈法、反光记录法、双普尔钦象法、角膜反射法等多种眼动分析方法. 随着眼动技术研究的不断深入和发展, 其应用的领域也越来越广泛. 人们不再是单纯地分析眼球的运动, 而是将它用于视线的分析与跟踪, 从人的眼睛特征乃至人的脸部特征中理解他们的欲望和需求. 如今视线跟踪技术已广泛应用于人机交互、虚拟现实、车辆辅助驾驶、人因分析和心理研究等多个领域, 有很好的应用前景.

视线跟踪系统按照系统构成和采用的检测方法可以粗略划分为侵入式和非侵入式[1], 又称为穿戴式和遥测式. 穿戴式系统需要用户佩戴特制的配置有光学系统的设备, 采用近眼的相机获取眼睛图像, 因此, 视线方向检测精度较高, 且其应用不受头部运动 (头部与穿戴装置之间的相对滑动除外) 的干扰, 主要应用于虚拟现实头盔、军用头盔和智能眼镜等. 遥测式系统则利用外部的摄像机获取人脸和人眼图像, 通过对人脸和人眼图像进行处理与分析, 估算出视线方向或视线关注点. 遥测式方法不需要佩戴任何的设备, 对人的干扰小, 有着广泛的研究和应用

收稿日期 2018-12-20 录用日期 2019-04-19
Manuscript received December 20, 2018; accepted April 19, 2019
国家重点研发计划 (2018YFC2001700), 北京市自然科学基金 (4172040), 北京科技大学顺德研究生院科技创新专项资金, 中央高校基本科研业务费专项资金 (FRF-GF-20-04A) 资助
Supported by National Key R&D Program of China (2018YFC2001700), Beijing Municipal Natural Science Foundation (4172040), Scientific and Technological Innovation Foundation of Shunde Graduate School, USTB, the Fundamental Research Funds for the Central Universities (FRF-GF-20-04A)
本文责任编委 黄庆明
Recommended by Associate Editor HUANG Qing-Ming
1. 北京科技大学自动化学院 北京 100083 2. 北京科技大学顺德研究生院 佛山 528399
1. School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083 2.Shunde Graduate School of University Of Science and Technology Beijing, Foshan 528399

空间.

目前, 对视线跟踪技术的研究主要分为两大类: 基于特征的视线跟踪方法和基于外观的视线估计方法. 基于特征的视线跟踪方法一般是从采集的图像中检测眼睛的一些可视特征, 如: 瞳孔、虹膜、眼角和角膜反射点等, 并提取相关视线参数, 通过映射模型估计视线落点或通过几何模型估计三维视线方向. 基于特征的视线跟踪方法在提高视线精度和处理头部自由运动问题上有很大的优势[2]. 而基于外观的视线估计方法是直接将人眼图像信息作为输入, 通过一组大的训练样本将人眼图像中反映的人眼外观与屏幕视线落点之间的映射模型训练出来, 从而通过采集到的新图像中的人眼外观, 找出对应的屏幕视线落点. 基于外观的视线估计方法只需要单目相机来获取人眼的外观, 采用的统计样本信息量大, 故鲁棒性比基于特征的视线跟踪方法强, 但是不能有效地解决头部自由运动和光照条件变化的问题[2−3]. 相比于基于特征的方法, 基于外观的视线估计方法距离实用化更远一些.

本文主要针对基于特征的视线跟踪方法进行了详细的研究和分析. 结构如下: 首先, 第 1 节对视线跟踪技术进行了阐述, 包括: 视线跟踪技术的发展、国内外相关研究工作以及视线跟踪技术的研究现状三部分; 然后, 将基于特征的视线跟踪方法分成了二维视线跟踪方法和三维视线跟踪方法两大类, 分别论述了二维视线跟踪方法和三维视线跟踪方法中的一些经典方法, 并从硬件系统配置、误差主要来源、头部运动影响、优缺点等多个方面对这些方法进行了比较; 此外, 从系统标定、用户标定、视线精度、头部运动及应用场景等方面对二维视线跟踪方法和三维视线跟踪方法进行了总结, 并对比分析了近五年现有的部分基于特征的视线跟踪方法; 接着, 第 3 节对基于特征的视线跟踪系统中的几个关键问题进行了探讨; 第 4 节介绍了视线跟踪技术在多个领域的应用; 最后, 对基于特征的视线跟踪方法的研究趋势进行了总结与展望.

1 视线跟踪技术

1.1 发展概述

视线跟踪技术在获取视线方向或视线关注点的过程中, 需要研究眼睛的生理结构和人眼运动特点、提取人眼特征并获得视线参数, 根据人眼成像原理建立视线方向模型或视线落点映射模型, 因此视线跟踪技术的研究内容涉及生理学与心理学、生物特征识别、计算机视觉等多个学科, 技术手段需融合人脸和人眼特征检测与跟踪、计算机视觉三维重建、眼光学等多种技术, 是一种典型的多学科交叉性研究课题.

早期的视线跟踪技术主要基于眼动检测技术. 观察法简单方便且易操作, 但只能对眼动有一个粗略的了解. 随后出现的机械记录法、胶片记录法等操作过程复杂, 且准确度低. 1958 年, Norman 设计了基于一个电影摄像机的视线追踪器, 这为当今视线跟踪技术的发展走向奠定了基础.

20 世纪以后, 研究者寻找新的能进行精确眼动测量的方法, 逐渐探索出了通过眼球转动时眼睛周围皮肤的电位差的变化来获取眼动情况的电流记录法, 该方法在 20 世纪 70 年代中期是应用最广泛的一种眼动研究方法. 1963 年, 出现了基于电磁感应原理来分析眼球的运动的电磁线圈法, Robinson 设计了一种装有探测线圈的隐形镜片, 将隐形镜片吸附在受试者眼睛上, 线圈中产生感应电压, 通过对线圈中感应电压的检测可精确的测量受试者眼睛运动方向. 这些方法都需要人眼与装置有直接的接触, 会给用户带来干扰和不适.

随着计算机视觉、图像处理、机器学习等技术的发展, 视线跟踪技术也有了很大的进步[4−8]. 目前, 眼动和视线检测方法主要是基于数字视频分析 (Video oculographic, VOG) 的视线跟踪技术, 通过相机拍摄人眼图像, 根据人眼图像信息来分析眼睛的注视方向. 基于 VOG 的视线跟踪技术具有非侵入性, 操作方便, 用户体验较好.

1.2 相关研究工作

国外对视线跟踪技术的研究有着较长的研究历史, 在 2005 年前后处于一个高峰期, 主要开展了基于特征的视线跟踪方法研究. 基于不同硬件系统配置的视线估计方法大量出现, 满足在不同应用场合下进行的视线估计的需求. 目前针对自由头动、标定复杂等问题的研究也在不断深入[9−11], 促进了视线跟踪技术的快速发展. 在国外, 基于外观的视线跟踪方法[12−17] 的研究开始于 2009 年, 目前主要致力于解决该方法在头部自由运动方面存在的局限性[14−17].

国内对视线跟踪技术的研究起步较晚, 近年来, 随着计算机视觉、人机交互、模式识别等相关学科的发展, 国内对视线跟踪技术的研究有上升的趋势. 如天津大学、西安电子科技大学、北京理工大学、中国科学技术大学、哈尔滨工业大学、东南大学等都开展了视线跟踪技术的研究[18−24]. 在基于特征的视线跟踪方法的研究领域, 华南理工大学的秦华标等[25] 和哈尔滨工程大学的高迪等[26] 对头动下视线估计的相关工作进行了研究. 中科院自动化所熊春水等针对单相机系统单点标定进行了研究[27]. 浙江大学的 Sun 等采用 Kinect 深度传感器进行视线估计[28],

为近年国内较有特色的研究. 可见这些研究分别针对头部自由运动、简化用户标定过程和简化硬件系统配置三大研究热点进行了深入研究. 近几年国内对基于外观的视线跟踪方法的研究比较多, 如: 北京航空航天大学、南京大学、清华大学等近几年都有相应的成果[29−30].

总体来说, 国内视线跟踪技术起步晚, 无论是应用基础研究还是产业化水平, 都与国外存在差距. 但相比于学术方面的研究, 国内的产业化水平上升的比较快, 在视线跟踪技术的产业化应用拓展方面, 北京七鑫易维信息技术有限公司和上海青研科技有限公司在融资的支持下, 在某些类型的产品方面正在逐步缩小与国外公司的差距.

1.3 研究现状

理想的视线跟踪系统应该具有如下特点: 1) "即拿即用", 无需个体标定; 2) 大范围全自由头动下的高精度视线估计; 3) 简化的硬件系统配置; 4) 适合各种应用环境.

但目前, 视线跟踪系统还属于一种在特定场合应用的专用设备, 一些瓶颈技术限制了其普适应用. 在物联网、大数据、云计算等技术高速发展的今天, 对视线跟踪技术提出了更高的需求. 例如: 简化用户标定 (甚至在保证一定的视线估计精度的情况下, 免除用户标定) 和实现真正的全自由头动下的视线估计, 可以使基于眼动跟踪的人因分析、认知疾病筛查与诊断等网络和大数据行为真正成为可能. 再例如: 简化硬件系统配置, 直至与基于外观的方法一样, 采用单一相机就可以实现简化的用户标定、全自由头动下的视线估计, 在移动式、便携式设备广泛应用的今天, 有着非常重要的意义, 这也对视线跟踪技术提出了新的挑战.

2 基于特征的视线跟踪算法

基于特征的视线跟踪技术是基于 VOG 的视线跟踪技术的重要组成部分, 其基本原理是利用某些在眼球转动时相对位置不变的眼部特征, 将其作为参照, 然后根据相对位置变化的特征和这些不变特征之间的关系, 提取变化的特征参数, 作为判别视线方向的主要依据. 因此视线跟踪技术一般包括两个部分: 图像处理与特征提取、视线方向或视线落点估计, 如图 1 所示. 图像处理与特征提取过程对相机采集的人脸图像进行图像处理与分析, 提取出估计视线所需要的视线参数特征; 视线方向或视线落点估计过程根据视线估计模型来估计屏幕视线落点.

基于特征的视线跟踪方法一般分为二维视线跟踪方法和三维视线跟踪方法, 二维视线跟踪方法从图像中提取眼睛的二维视线参数, 将其代入由用户标定确定的视线参数到屏幕注视点的映射函数, 换算成视线的屏幕落点. 而三维视线跟踪方法是通过研究人眼的成像模型和空间几何模型, 由图像的视线特征参数和光源位置坐标来计算三维视线方向, 在系统标定出屏幕位置的情况下, 视线与屏幕相交得到屏幕视线落点. 下面本文分别对二维视线跟踪方法和三维视线跟踪方法进行了论述.

2.1 二维视线跟踪方法

二维视线跟踪方法是将从采集的图像中提取的二维人眼特征参数与采集图像时刻用户眼睛的注视方向或注视点建立映射模型, 通过获取图像上用户的二维人眼特征参数, 即可根据人眼特征参数与注视点之间的模型估计出用户的视线关注点[31−74]. 二维视线跟踪方法主要有: 瞳孔 (或虹膜) – 眼角法[35−42]、瞳孔 (或虹膜) – 角膜反射法[1, 3, 43−56]、交比值法[57−66]、单应性归一化 (Homography normalization, HN) 法[67−74] 等.

2.1.1 瞳孔 (或虹膜) – 眼角法

瞳孔 (或虹膜) – 眼角法通常基于单相机无光源系统获取用户的注视点, 在自然光照的条件下, 虽不能将瞳孔与虹膜分开, 但可以认定二者是同心圆, 且在眼动过程中是同步变化的. 如图 2 所示, 当头

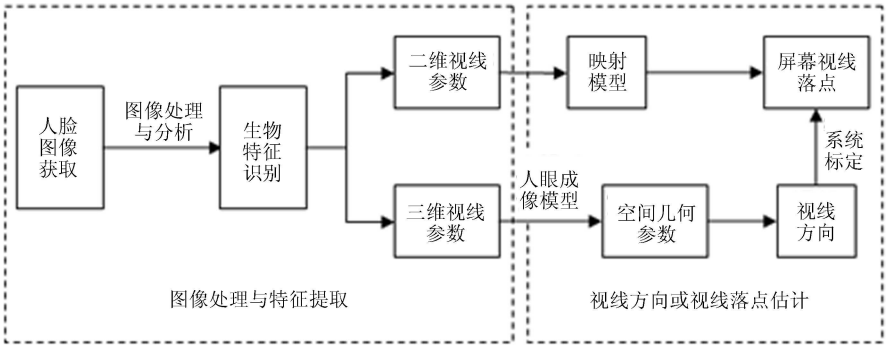


图 1 基于特征的视线跟踪方法原理流程图

Fig. 1 The flow diagram for feature-based gaze tracking methods

部固定不动时, 眼角的位置不动, 由眼球的转动改变视线方向. 当眼球转动时, 瞳孔 (或虹膜) 中心的位置会发生变化, 则以眼角为基准, 视线方向改变时由瞳孔 (或虹膜) 中心指向眼角的矢量会相应发生变化. 因此, 通过从相机采集的图像中检测瞳孔 (或虹膜) 中心的位置, 以及眼角的位置, 得到瞳孔 (或虹膜) 中心–眼角的矢量, 设定一个二维的映射模型, 建立瞳孔 (或虹膜) 中心–眼角矢量与视线之间的对应关系, 常用的二维映射模型如式 (1) 所示.

$$\begin{cases} s_x = a_0 + a_1 v_x + a_2 v_y + a_3 v_x v_y + a_4 {v_x}^2 + a_5 {v_y}^2 \\ s_y = a_6 + a_7 v_x + a_8 v_y + a_9 v_x v_y + a_{10} {v_x}^2 + a_{11} {v_y}^2 \end{cases} \tag{1}$$

其中, s_x, s_y 为屏幕上视线落点的横纵坐标, v_x, v_y 为图像上对应的瞳孔 (或虹膜) 中心–眼角矢量的横纵坐标, $a_0 \sim a_{11}$ 为映射模型的系数. 用户需要通过用户标定过程求取映射模型中未知的固定系数, 这些系数包含用户的个体差异信息. 下次使用该系统时, 即可根据新采集的图像中提取出的瞳孔 (或虹膜) 中心–眼角矢量估计出当时的视线信息.

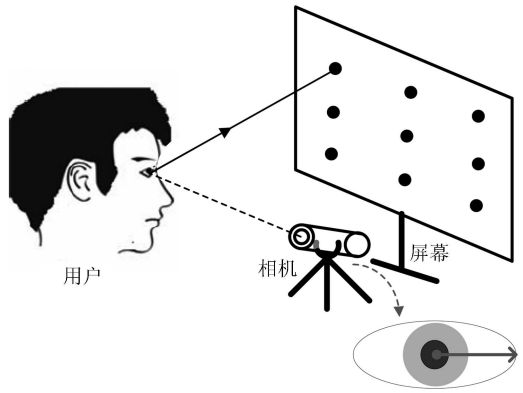


图 2　瞳孔 (或虹膜)–眼角法视线跟踪系统示意图

Fig. 2　The gaze tracking system of pupil (or iris)-corner technique

Zhu 等[35] 建立了一个虹膜中心–眼角矢量与注视角之间的二维线性映射模型, 通过线性插值计算连续帧的视线方向, 在眼睛图像的分辨率为 320×240 的情况下, 视线精度可达到 $1.4°$. 该方法的特点在于: 采用的模型比较简单, 提出了使虹膜中心和眼角检测达到亚像素精度的方法, 从而在简单的线性模型中达到较高的精度, 但需要保持头部静止. Shao 等[36] 先提出了传统的瞳孔中心–眼角矢量与屏幕上的视线关注点之间的映射模型, 然后提出了一种根据人眼与屏幕的二维几何模型估计视线的方法, 从所需的系统输入、标定点个数、视线精度和对头动的影响四个方面对这两种方法进行了对比分析, 结果表明, 后者比前者具有更高的精度, 但后者利用了一个自变量为瞳孔中心–眼角矢量的横坐标的三次函数作为视线落点的误差补偿函数, 导致所需要的人眼特征更多, 且标定点更多. 与式 (1) 中映射模型不同的, Xia 等[40] 将映射模型简化为图像上瞳孔中心的坐标与屏幕注视点之间的线性映射模型, 结果证明视线精度并未因模型简单受到影响. 并基于伽柏小波变换和结构张量检测眼角点, 确定感兴趣的区域用于检测瞳孔中心, 以此增强光照条件发生变化时眼睛特征检测结果的鲁棒性, 但该系统限制了头部的自由运动.

瞳孔 (或虹膜)–眼角法所需的硬件配置简单, 只需要一个普通相机, 便于操作与实现, 特征提取相对简单, 只需要提取瞳孔 (或虹膜) 和眼角在图像上的坐标, 但用户标定过程复杂, 需要通过多个标定点计算人眼特征与视线之间的映射模型, 耗时较长, 并且为了保证视线的精度, 在标定过程中需要保持头部固定, 会给使用者带来不便. 更重要的是, 由于在角膜上没有光源形成的亮斑, 为了构建眼球运动的基准点, 只能采用眼角等固定点作为表述眼球运动的基准点, 通过瞳孔 (或虹膜) 中心与眼角之间的矢量来描述眼球运动. 而在自然光照条件下, 准确稳定地分割眼角点也具有一定的挑战性.

2.1.2　瞳孔 (或虹膜) – 角膜反射法

针对瞳孔 (或虹膜)–眼角法应用于单相机无光源系统中存在的精确检测眼角点较难的问题, 瞳孔 (或虹膜)–角膜反射法在系统中增加了光源, 用光源通过角膜反射形成的普尔钦斑代替眼角点, 作为眼球运动的基准点. 当头部固定不动时, 眼球的转动改变视线的方向. 由于角膜近似球形, 当眼球转动时, 假设光源形成的普尔钦斑在角膜上的位置不动, 以普尔钦斑为基准, 则视线方向改变时由瞳孔 (或虹膜) 中心指向普尔钦斑中心的矢量会相应地发生变化. 因此, 将图像中提取出的瞳孔 (或虹膜) 中心和普尔钦斑中心构成矢量, 作为视线参数, 与屏幕上的视线落点建立映射模型 (与式 (1) 类似), 将该参数代入映射模型即可得到视线在屏幕上的落点.

采用瞳孔–角膜反射法和虹膜–角膜反射法一般选取的光源不同. 由于在可见光条件下, 瞳孔的边缘与虹膜的灰度值较接近, 准确地定位瞳孔中心较难, 而在红外光源照射下, 瞳孔呈黑色特征, 虹膜较白, 瞳孔特征比较突出, 分割难度较低, 因此在使用瞳孔特征进行视线跟踪的方法中, 通常采用的是红外光源[1, 43−45], 尤其是双环形红外光源[1, 43, 45]. 在双环形红外光源系统中, 内环光源在光轴附近, 主要的作用是用于检测瞳孔, 而不对视线估计起作用, 外环光源为远光轴光源, 在角膜上形成普尔钦斑. 双环形主动红外光源系统进行视线估计采用的是内环光

源和外环光源交替亮灭, 如图 3 所示. 当内环光源点亮时, 图像中会产生明亮的瞳孔, 即 "亮瞳" 现象, 便于精准地定位瞳孔. 当外环光源点亮时, 角膜表面形成的普尔钦斑会在相机中成像, 用于检测普尔钦斑的中心位置. 对亮瞳图像和暗瞳图像进行差分, 可以更完整地保留瞳孔区域, 更精确地检测瞳孔并确定瞳孔的中心位置. 相反地, 虹膜外侧的巩膜呈白色, 在可见条件下可以较清晰地检测出虹膜的轮廓, 与在红外光源条件下相比, 对比度更高, 因此在使用虹膜特征进行视线跟踪的方法中一般采用可见光源对虹膜 – 角膜反射法进行分析[3].

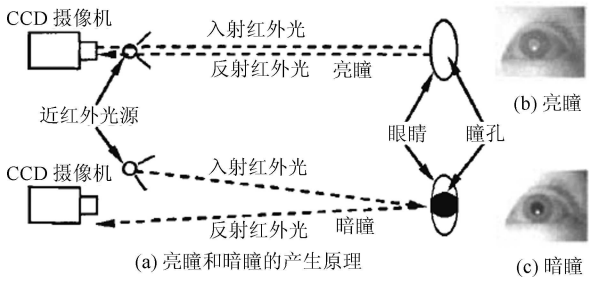


图 3 "亮瞳" 现象与 "暗瞳" 现象[43]

Fig. 3 Bright pupil and dark pupil[43]

1) 单相机单光源系统

采用瞳孔 (或虹膜) – 角膜反射法进行视线估计的最小硬件系统配置为单相机单光源系统. 在现有的文献中, 单相机单光源系统的人眼视线参数到视线落点的映射关系都是采用一组多项式映射模型来描述[1, 3, 43−48]. Blignaut[46] 在一个相机和一个红外光源的系统配置下对比分析了已有文献中的不同的二维映射模型, 得出标定点个数为 9 时的最佳的二维映射模型如式 (1) 所示, 而标定点个数大于等于 14 时, 最佳的二维映射模型如式 (2) 所示[48], 视线精度可以达到 0.5° 以内, 最佳的用户标定过程是每个用户盯视屏幕上的 14 个标定点, 大约耗时 20 秒.

$$\begin{cases} s_x = a_0 + a_1 v_x + a_2 {v_x}^2 + a_3 {v_x}^3 + a_4 v_y + \\ \qquad a_5 v_x v_y + a_6 {v_x}^2 v_y + a_7 {v_x}^3 v_y \\ s_y = b_0 + b_1 v_x + b_2 {v_x}^2 + b_3 v_y + b_4 v_x v_y + b_5 {v_x}^2 v_y \end{cases} \tag{2}$$

其中, s_x, s_y 为屏幕上视线落点的横纵坐标, v_x, v_y 为图像上对应的瞳孔 (或虹膜) 中心 – 普尔钦斑矢量的横纵坐标, $a_0 \sim a_7$, $b_0 \sim b_5$ 为映射模型的系数. Morimoto 等[44] 指出当使用者的头部位于用户标定位置时, 单相机单光源系统能够获得精确的视线在屏幕上的落点; 当使用者头部偏离标定位置时, 单相机单光源系统的视线落点估计精度会显著下降. Sigut 等[3] 通过实验详细研究了当采用典型的非线性多项式作为映射模型时, 头部偏离标定位置后对屏幕视线落点精度的影响, 结论是: 当使用者在上下和左右方向移动头部时, 视线落点的精度变化不大; 当头部远离和靠近相机, 即相对于屏幕前后移动时, 对视线落点的精度影响非常大. 因此, 利用二维映射模型在单相机单光源系统下进行视线估计通常需要用户头部保持静止状态.

为了解决瞳孔 (或虹膜) – 角膜反射法在单相机单光源系统配置下对头部偏离标定位置的影响, Zhang 等[1] 在单相机双环形光源系统中推导了头动状态下的视线参数的补偿模型, 用来补偿头部运动产生的视线参数的变化, 将视线参数修正回标定位置的相应的参数, 进而获得准确的视线落点信息, 但从补偿模型中可以看出, 需要已知用户的真实瞳距. Kim 等[45] 通过调整两个反光镜使人眼在窄视场角的相机中成像, 并将两个反光镜的转角与人眼特征参数一起作为模型的输入, 基于广义回归神经网络预测出来, 即可将头部的运动通过反光镜的转角反馈到视线模型中, 用于实现自由头动状态下的视线估计. 该方法解决了单相机单光源系统的自由头动问题, 鲁棒性强, 但在系统的布局 (如: 反光镜的放置) 和操作方面要求较高, 也使系统过于复杂, 缺少实用价值.

单相机单光源系统是目前应用最广的眼动仪形式, 其系统配置相对简单, 对硬件配置要求较低, 视线参数只涉及到瞳孔 (或虹膜) 以及普尔钦斑中心的检测, 且不需要对光源和屏幕的位置进行标定, 利用补偿方法可以实现自由头动下的视线估计, 但在一般的方法中, 仍需要用户头部尽可能地位于标定位置, 且尽可能地保持静止, 标定过程相对复杂, 映射模型中未知的系数较多, 造成所需的标定点个数较多.

2) 单相机多光源系统

当基于单相机多光源系统改进瞳孔 (或虹膜) – 角膜反射法[49−52] 时, 由于两个光源在同一只眼睛上会形成两个普尔钦斑, 将两个普尔钦斑的位置关系引入映射模型中, 可以有效地改善单相机单光源系统二维映射模型对头部运动的敏感程度, 在保证视线精度的情况下, 允许用户在一定范围内头动. Cerrolaza 等[53] 强调对瞳孔 – 角膜矢量进行归一化有利于保证用户在头动状态下视线估计的鲁棒性, 因此, Sesma-Sanchez 等[50] 提出了三种归一化因子的计算方法, 分别为: 同一只眼睛两个普尔钦斑之间的距离、左右眼瞳孔之间的距离、以及两个光源分别产生的左右眼普尔钦斑之间的距离的平均值, 并在三个映射模型中进行了实验, 结果表明, 从头动状态下的系统精度及鲁棒性分析, 同一只眼睛两个普尔钦斑之间的距离是最佳归一化因子, 这决定了系统需要包含两个光源, 此外, Sesma-Sanchez 等指出

映射模型的复杂程度对用户头动状态下的视线精度无显著影响, 也就是说, 可以采用相对简单的映射模型来进行视线估计. 张太宁等[51] 提出的基于单相机双光源系统的二维映射模型中, 根据两个光源在同一只眼睛中产生的两个普尔钦斑, 计算两个普尔钦斑的连线中心坐标以及两个普尔钦斑之间的水平距离, 用虹膜中心与两个普尔钦斑的连线中心构成的矢量、两个普尔钦斑的连线中心以及两个普尔钦斑之间的水平距离这五个特征量与屏幕视线落点构成了一个一次多项式模型, 如式 (3) 所示.

$$\begin{cases} s_x = c_1 \cdot G_x + c_2 \cdot M_x + \frac{c_3 \cdot G_x}{L} \\ s_y = d_0 + d_1 \cdot G_y + d_2 \cdot M_y + \frac{d_3 \cdot G_y}{L} \end{cases} \tag{3}$$

其中, s_x, s_y 为屏幕视线落点的横纵坐标, G_x, G_y 为用虹膜中心与两个普尔钦斑的连线中心构成的矢量的横纵坐标, M_x, M_y 为两个普尔钦斑的连线中心坐标, L 为两个普尔钦斑之间的水平距离, $c_1 \sim c_3, d_0 \sim d_3$ 为映射模型系数. 该方法允许头部在左右约 15 cm、上下约 8 cm、转动约 10° 范围内运动, 简化了映射模型, 进而减少了所需用户标定点的个数.

与基于单相机单光源系统的瞳孔 (或虹膜) – 角膜反射法相比, 基于单相机多光源系统的改进的瞳孔 (或虹膜) – 角膜反射法通常具有以下优点: 1) 降低了映射模型对头部运动的敏感度, 可实现一定范围内的头部自由运动; 2) 在保证视线精度的情况下, 映射模型较简单, 所需标定点个数较少, 运算量较小.

3) 多相机系统

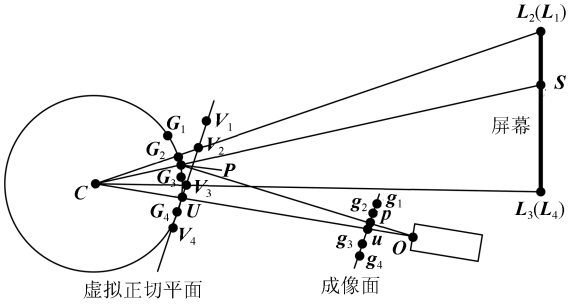
针对基于单相机系统的映射模型中存在的自由头动影响、光照不均匀对结果的影响等, 也有一些研究者在多相机系统中采用瞳孔 (或虹膜) – 普尔钦斑矢量与视线落点的映射模型[54−56]. Zhu 等[54−55] 利用双相机系统对头部运动前后瞳孔中心 – 普尔钦斑中心矢量的变化进行了推导, 将自由头动下的瞳孔中心 – 普尔钦斑中心矢量转换到标定位置下对应的瞳孔中心 – 普尔钦斑中心矢量, 代入标定位置下的二维映射模型中计算视线落点, 以补偿头动对视线估计的影响. Kim 等[56] 从实用化产品的角度, 设计了基于三个相机的能够补偿光照不均匀的视线跟踪头盔, 改善了光照不均匀对瞳孔检测的影响, 并建立了图像上的瞳孔中心坐标与屏幕视线落点之间的二维多项式映射模型, 用于估计视线落点.

利用多相机系统进行二维视线估计的研究较少, 主要原因有: 1) 利用多相机系统可以高精度的估计三维视线, 与二维视线估计相比, 用户标定过程更简单, 能更充分地利用多相机的优势, 更便捷地实现视线估计中的头部自由运动、提高精度等. 2) 针对单相机系统中的二维映射模型存在的对头部运动敏感等问题, 目前已经有很多研究者在单相机系统中进行相关的研究, 简化系统的硬件配置无论从理论层面还是实用角度都具有更深远的意义.

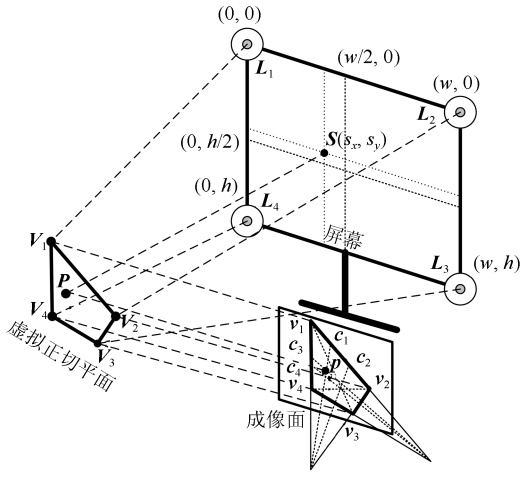
2.1.3　交比值法

基于交比不变性的屏幕视线落点估计方法[57−66] 应用了射影几何学中的交比射影不变性, 即: 空间中一条直线经空间几何变换后保持了直线上对应点列的交比不变, 对于空间平面多边形, 投影图像上某一条边上的交比值与几何变换前平面多边形上对应边的交比值相等.

如图 4 所示, 使用交比值法的视线跟踪系统一般至少包含四个光源, 分别放置在屏幕的四个角上, 当用户看屏幕上一点 $\boldsymbol{S}$ 时, 四个光源在角膜上形成反射, 反射点为 $\boldsymbol{G}_i$ $(i = 1, 2, 3, 4)$, 则在角膜图像上形成四个普尔钦斑 $\boldsymbol{g}_i$. 将角膜近似看作一个球体, 故过 $\boldsymbol{U}$ 点存在一个虚拟正切平面, 如图 4 (a) 所示. 该虚拟正切平面与四个光源 $\boldsymbol{L}_i$ 和角膜中心 $\boldsymbol{C}$ 之间的连线会有交点 $\boldsymbol{V}_i$. 由于反射点为 $\boldsymbol{G}_i$ 与虚拟正切平面上的 $\boldsymbol{V}_i$ 不相同, 在成像面上, $\boldsymbol{V}_i$ 的像点用



(a) 基本原理
(a) Basic principles



(b) 屏幕和成像面的交比不变性
(b) Cross-ratio invariance between the screen and the image

图 4　交比值法视线估计模型
Fig. 4　Cross-ratio based gaze estimation model

$\boldsymbol{v}_i = \boldsymbol{u} + \alpha(\boldsymbol{g}_i - \boldsymbol{u})$ 表示, 其中的比例因子 α 可通过标定获得. 以虚拟正切平面为媒介, 根据交比不变性, 屏幕上 LED 光源构成的多边形的某个边上的交比值与成像面上 $\boldsymbol{v}_i$ 构成的多边形对应边上的点的交比值相等, 如图 4 (b) 所示, 屏幕的宽为 w, 高为 h, 屏幕注视点为 (s_x, s_y), 则屏幕上的交比值为:

$$\begin{cases} CR^x_{screen} = \dfrac{s_x}{w - s_x} \\ CR^y_{screen} = \dfrac{s_y}{h - s_y} \end{cases} \tag{4}$$

成像面上, 如果 $\boldsymbol{v}_i = (x_i^v, y_i^v)$, 根据瞳孔像点和 $\boldsymbol{v}_i$ 构成的多边形的对角线交点确定的辅助点 $\boldsymbol{c}_i = (x_i^c, y_i^c)$, 则图像上 $\boldsymbol{v}_i$ 构成的多边形对应边上的点的交比值为:

$$\begin{cases} CR^x_{image} = \dfrac{(x_1^v y_1^c - x_1^c y_1^v)(x_2^c y_2^v - x_2^v y_2^c)}{(x_1^v y_2^c - x_2^c y_1^v)(x_1^c y_2^v - x_2^v y_1^c)} \\ CR^y_{image} = \dfrac{(x_2^v y_3^c - x_3^c y_2^v)(x_4^c y_3^v - x_3^v y_4^c)}{(x_2^v y_4^c - x_4^c y_2^v)(x_3^c y_3^v - x_3^v y_3^c)} \end{cases} \tag{5}$$

根据图像上与屏幕上的交比值对应相等, 可以解出 s_x 和 s_y, 如式 (6) 所示, 即得到了视线在屏幕上的关注点.

$$\begin{cases} s_x = \dfrac{w \cdot CR^x_{image}}{1 + CR^x_{image}} \\ s_y = \dfrac{h \cdot CR^y_{image}}{1 + CR^y_{image}} \end{cases} \tag{6}$$

Kang 等[58] 着重分析了利用投影几何交比值不变的性质进行视线估计的两个主要误差来源: 一是视轴与光轴不共线, 存在着一个水平方向的偏差角和一个垂直方向的偏差角; 二是真实的瞳孔中心不在角膜反射平面上, 这一偏差与瞳孔中心到角膜中心的距离、角膜曲率半径等参数有关. Coutinho 等[59] 对基于交比不变性的方法进行了大量的实验, 在已知角膜曲率半径、角膜中心位置、光源位置、相机位置的情况下, 对比分析了三种已有的利用投影不变性估计视线落点的方法, 在这三种方法中, 均分析了考虑光轴与视轴的夹角以及不考虑光轴与视轴的夹角两种情况. 并在这些方法通过标定比例因子 α 使普尔钦斑位于虚拟正切平面内的基础上, 分别利用 4 点标定和 9 点标定计算比例因子 α, 用于投影计算, 也通过这个标定过程计算了一个偏移向量来补偿光轴与视轴的夹角. 通过比较分析, 体现了该方法的优越性. 为了改善传统交比值法中瞳孔中心不在虚拟正切平面内引入的计算偏差, Cheng 等[62] 提出了使用动态虚拟正切平面的思想, 通过动态矩阵描述光源反射点与虚拟点之间的关系, 视线精度可达到 0.7°. Yoo 等[61] 提出了在两个相机和五个红外 LED 光源下实现的基于交比值法的视线估计方法, 一个可变焦相机用于捕捉人眼特征, 另一个广角相机用于检测人脸, 以实现大范围头动下的视线估计, 除放置在屏幕四个角的光源之外, 还有一个光源置于可变焦相机上方, 用于照亮人眼, 以便得到清晰的人眼图像, 提取出更准确的瞳孔特征.

交比值法的优点在于它只需要已知光源位置及屏幕的大小, 而不需要一个标定了的相机, 就可以实现自由头动状态下的较精确的视线落点估计, 模型简单. 但由于实际的瞳孔中心不在虚拟正切平面上[58], 因此交比值法是一种近似的模型, 且对光源位置标定和图像处理精度的要求较高. 为了利用交比不变性计算视线落点, 系统中的四个光源均需在人眼中形成反射, 并在相机中成像, 这将限制头部运动的范围. 此外, 为了补偿光轴与视轴之间的夹角, 用户标定过程需要多个标定点, 增加了用户标定的复杂度.

2.1.4 单应性归一化法

基于单应性归一化的方法[67−74] 的硬件系统配置与交比值法类似, 通常包含一个相机和置于屏幕四个角上的四个光源, 视线估计模型如图 5 所示. 该方法不使用单一的映射函数, 而是基于瞳孔中心与角膜反射平面共面的假设, 通过两次投影的转换关系估计视线落点, 可以补偿瞳孔 (或虹膜) – 角膜反射法和交比值法的误差. 基本原理是: 根据成像面 Π_{I} 到角膜反射平面 Π_{C}, 以及角膜反射平面 Π_{C} 到屏幕平面 Π_{S} 之间的两个投影映射关系, 计算出对应的单应性矩阵 $\boldsymbol{H}_{\boldsymbol{I}}^{\boldsymbol{C}}$ 以及 $\boldsymbol{H}_{\boldsymbol{C}}^{\boldsymbol{S}}$. 估计视线落点时, 根据成像面上瞳孔中心的位置 $\boldsymbol{p}$, 利用成像面 Π_{I} 与角膜反射平面 Π_{C} 的单应性矩阵 $\boldsymbol{H}_{\boldsymbol{I}}^{\boldsymbol{C}}$, 先转换到角膜反射平面 Π_{C} 中的点 $\boldsymbol{P}$, 再利用角膜反射平面 Π_{C} 与屏幕平面 Π_{S} 的单应性矩阵 $\boldsymbol{H}_{\boldsymbol{C}}^{\boldsymbol{S}}$, 转换到屏幕平面 Π_{S} 上的一点 $\boldsymbol{S}$, 该点即为注视点.

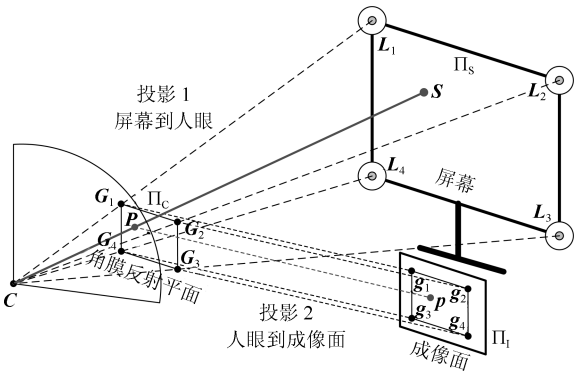


图 5 单应性归一化法视线估计模型

Fig. 5 HN based gaze estimation model

由于未对系统进行标定, 角膜反射平面 Π_{C} 的

三维坐标是未知的, 因此假设角膜反射平面为一个归一化平面 Π_{N}, 即一个具有预定义坐标的单位正方形, 则 $\boldsymbol{H}_{\boldsymbol{I}}^{\boldsymbol{N}}$ 可以通过角膜反射产生的四个普尔钦斑和预定义的归一化平面 Π_{N} 的四个角点计算出来. 而通过用户标定过程可以确定归一化平面 Π_{N} 与屏幕平面 Π_{S} 之间的投影矩阵 $\boldsymbol{H}_{\boldsymbol{N}}^{\boldsymbol{S}}$, 标定点个数不少于 4 个. 最终得到从成像面到屏幕的投影转换矩阵为:

$$\boldsymbol{H}_{\boldsymbol{I}}^{\boldsymbol{S}} = \boldsymbol{H}_{\boldsymbol{I}}^{\boldsymbol{N}} \boldsymbol{H}_{\boldsymbol{N}}^{\boldsymbol{S}} \tag{7}$$

基于单应性归一化视线估计方法在四个红外光源的基础上, 不需要增加额外的红外光源, 且由于引入了归一化空间的特性, 不需要知道由红外光源形成的屏幕矩形的尺寸, 便于系统的实现[70]. 由于这两个单应性矩阵在估计视线之前对头部运动进行了标准化, 且补偿了光轴与视轴之间的角度偏差, 因此, 该方法比瞳孔 (或虹膜) – 角膜反射法、交比值法鲁棒性更强, 对头部姿态变化的敏感性较低[69]. 但为了求解两个投影转换矩阵, 系统中的四个光源均需在人眼中形成反射, 并在相机中成像, 这将导致头部运动的范围受到限制. 且投影模型中近似瞳孔中心在四个普尔钦斑构成的平面内, 引入了误差. 此外, 为了标定出归一化平面与屏幕平面之间的投影矩阵, 用户标定过程需要至少 4 个标定点, 增加了用户标定的复杂度.

2.1.5　小结

二维视线跟踪方法利用人眼的可视特征, 如: 瞳孔、虹膜、眼角、光源反射形成的普尔钦斑等, 通过模拟眼动变化特征与视线落点之间的二维模型来进行视线跟踪. 根据上述二维视线跟踪方法, 表 1 简单总结了这些方法的系统配置、误差主要来源、头动影响、优缺点等.

由表 1 可知, 在单相机系统中, 当不采用光源或只采用一个光源时, 映射模型无法保证头动不变性, 故要求用户头部保持固定. 而采用两个及两个以上光源时, 可以根据不同光源在角膜外表面产生反射形成的普尔钦斑之间的相对位置关系, 将头动因素考虑在内, 实现头部在一定范围内的自由运动. 基于单相机系统的二维映射模型所需系统配置简单, 不需要标定相机, 且便于实现, 在头部固定和小范围头动的情况下可以达到较高的精度, 故在含有头枕、腮托等固定装置的设备以及头盔、智能眼镜中应用较广泛. 但在遥测式系统中, 头部自由运动会使得视线精度明显降低. 在多相机系统中, 头部自由运动引入的眼动变化特征的变化量可以计算出来, 用于补偿头动对视线落点估计的影响, 但需要标定相机之间的位置关系, 增加了系统标定的复杂度. 二维视线跟踪方法不可避免的需要多个标定点来完成用户标定过程, 以确定眼部特征与视线落点的转换关系, 过程相对复杂, 与实现自然人机交互相违背.

2.2　三维视线跟踪方法

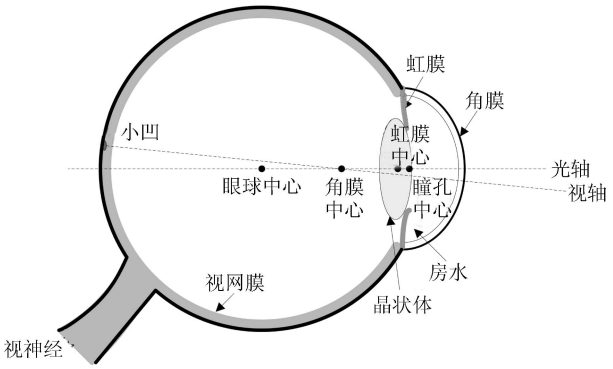
三维视线跟踪方法[28, 75−116] 是根据眼球的几何模型来估计三维视线. 人眼模型如图 6 所示, 眼睛近似于球状, 半径约为 12 mm. 从眼球的外层向内, 依次是巩膜, 虹膜和瞳孔. 角膜覆在虹膜外面, 虹膜中心有一个环状体, 称作瞳孔, 用来调节进入眼睛的光量. 虹膜后面是晶状体, 用来调节光线聚集在视网膜上形成图像. 视网膜位于眼球后面, 蕴含大量感光细胞. 角膜和晶状体之间是水样液体, 晶状体和视网膜之间是玻璃体. 光线进入眼睛后经过这一系列的光介质并在各个层面上反射和折射后才到达视网膜.

表 1　二维视线跟踪方法的特点对比

Table 1　Comparison of the characteristics of two-dimensional gaze tracking methods

方法	相机	光源	误差主要来源	头动影响	优点	缺点
瞳孔 (虹膜) – 眼角法	1	0	眼角检测 头部偏离标定位置	头部固定	系统简单 便于操作与实现	用户标定复杂 头部固定带来的不适感
瞳孔 (虹膜) – 角膜反射法	1	1	头部偏离标定位置	头部固定	系统简单 无需标定光源和屏幕位置	用户标定复杂 头部固定带来的不适感
	1	≥2	普尔钦斑检测	头部一定范围运动	无需标定光源和屏幕位置 模型相对简单	用户标定复杂
	≥2	≥1	多个相机的转换关系	头部自由运动	允许头部自由运动 精度较高	用户标定复杂、系统标定复杂 相机标定复杂
交比值法	≥1	≥4	虚拟正切平面的构建	头部一定范围运动	免相机标定 模型简单	四个光源均需成像限制头部运动范围 需多点用户标定补偿光轴与视轴的夹角
HN 法	1	≥4	近似瞳孔在普尔钦斑构成的平面中	头部一定范围运动	无需已知光源构成的矩形尺寸 模型简单 补偿了光轴与视轴的夹角	四个光源均需成像限制头部运动范围 用户标定复杂

(a) 眼睛图像
(a) Eye image



(b) 眼球构造
(b) Eyeball structure

图 6 人眼模型
Fig. 6 Eye model

在视网膜上有一个特殊的小区域, 称为小凹. 上面集中了绝大部分颜色敏感的细胞, 用于感知景物细节. 小凹并不准确位于眼球的对称轴上, 而是与角膜中心构成视轴. 眼球的对称轴是光轴, 眼球的光轴与视轴之间存在一个因人而异的固定的偏角, 这个角被称为 kappa 角, 大小通常在 5° 左右. 视线跟踪指的是估计视轴的方向, 并根据所观察景物的信息确定注视点. 视轴上有小凹和角膜中心, 角膜中心是光轴与视轴的唯一交点, 但小凹从眼球外侧无法获取其相关信息, 所以直接计算视轴是无法实现的. 目前常用的三维视线估计方法通常先重建出眼睛的光轴, 然后根据人眼特定的 kappa 角, 从光轴换算到视轴, 进而得到注视点的位置. 因此, 三维视线估计的基本任务是光轴的求解.

眼球中心 $\boldsymbol{E}$、角膜中心 $\boldsymbol{C}$、虹膜中心 $\boldsymbol{I}$ 和瞳孔中心 $\boldsymbol{P}$ 都位于光轴上, 此外, 由于在眼球的转动过程中, 眼球的光轴始终垂直于瞳孔边缘所在的平面, 也同时垂直于虹膜边缘所在的平面, 故如果将瞳孔或虹膜看作是空间的圆形目标, 则光轴的方向向量可以用瞳孔或虹膜的空间法向量表示. 综上, 在视线跟踪系统中, 利用眼球的可视特征重建光轴的方法如图 7 所示. 在瞳孔或虹膜的空间法向量已知的情况下, 如果已知眼球中心 $\boldsymbol{E}$、角膜中心 $\boldsymbol{C}$、虹膜中心 $\boldsymbol{I}$、瞳孔中心 $\boldsymbol{P}$ 中的任意一点的坐标, 则可以利用点法式方程重建光轴; 而在瞳孔或虹膜的空间法向量未知的情况下, 则需要已知眼球中心 $\boldsymbol{E}$、角膜中心 $\boldsymbol{C}$、虹膜中心 $\boldsymbol{I}$、瞳孔中心 $\boldsymbol{P}$ 中的任意两点的坐标, 才能重建光轴.

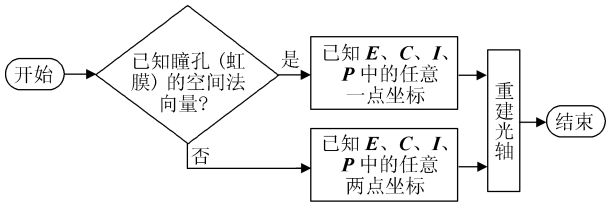


图 7 光轴重建方法示意图
Fig. 7 The methods of optical axis reconstruction

在上述光轴重建方法中, 由于眼球中心很难通过眼睛的外部可视特征计算出来, 除非眼球转动, 盯视不同点时才可以反映眼球中心的信息, 因此在眼球光轴重建模型中, 一般不采用眼球中心作为重建光轴的条件.

三维视线估计关键技术如图 8 所示, 主要包括三个部分: 系统标定技术、用户标定技术、三维视线估计技术, 而用户标定和三维视线估计都包含图像

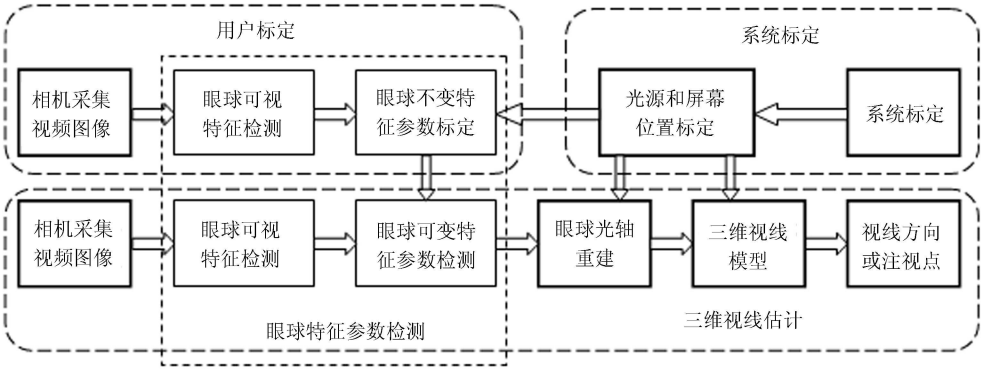


图 8 三维视线估计关键技术构成图
Fig. 8 Key technologies of three-dimensional gaze estimation

的人眼特征检测部分. 三维视线估计需要在用户标定阶段确定用户的眼球不变参数, 然后在视线跟踪阶段基于标定的眼球不变参数和系统参数估计眼球的变化参数, 进而重建眼球光轴和视轴. 眼球的不变参数是每个用户固定不变、但具有个体差异的眼球参数, 如角膜曲率半径、虹膜半径、kappa 角等, 而眼球的变化参数是眼动过程中与视线方向密切相关的眼球参数, 如角膜中心、瞳孔中心、虹膜中心等.

根据采用的相机种类的不同, 将基于特征的三维视线跟踪方法分为以下两类: 基于深度相机的三维视线估计方法和基于普通相机的三维视线估计方法.

2.2.1　基于深度相机的三维视线估计方法

基于深度相机的三维视线估计方法[28, 88−97] 通常只需要一个深度相机, 如: Kinect 等, 在获取图像信息基础上还能获取深度信息. 通过结合图像信息和深度信息来对人眼进行三维建模, 进而实现三维视线估计. 以利用 Kinect 的三维视线跟踪方法为例, 如图 9 所示.

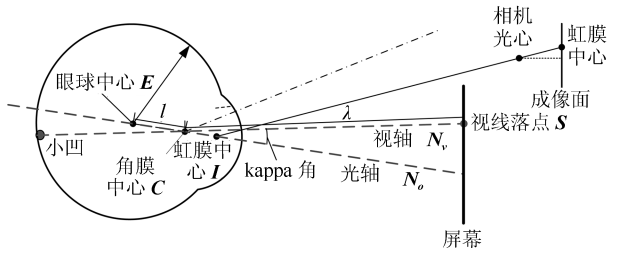


图 9　基于 Kinect 的三维视线估计模型

Fig. 9　3D gaze estimation model based on a kinect

其基本思路为: 首先, 根据头部姿态跟踪得到头部坐标系与相机坐标系之间的转换关系, 包括: 头部坐标系原点在相机坐标系下的坐标以及头部坐标系与相机坐标系之间的转换矩阵. 根据眼球中心相对于头部位置保持固定这一性质, 通过标定过程计算出头部坐标系下的眼球中心坐标, 从而可以根据转换矩阵得到相机坐标系下三维空间中的眼球中心坐标 $\boldsymbol{E}$; 然后根据虹膜中心的成像关系, 利用小孔成像模型计算出三维空间中虹膜中心的坐标 $\boldsymbol{I}$. 利用眼球中心和虹膜中心建立光轴单位方向向量 $\boldsymbol{N_o} = \frac{\boldsymbol{I}-\boldsymbol{E}}{|\boldsymbol{I}-\boldsymbol{E}|}$, 通过用户标定计算出光轴与视轴之间的转换关系 (转换矩阵 $\boldsymbol{M}$ 或水平与竖直方向的旋转角 α, β), 即可得到视轴单位方向向量 $\boldsymbol{N_v} = \boldsymbol{M}\boldsymbol{N_o}$ (或 $f(\boldsymbol{N_o}, \alpha, \beta)$). 最后根据眼球中心 $\boldsymbol{E}$、眼球中心与角膜中心的矢量、以及角膜中心与视线落点的矢量, 可以直接表示出视线落点 $\boldsymbol{S}$ 的三维坐标为:

$$\boldsymbol{S} = \boldsymbol{E} + l\boldsymbol{N_o} + \lambda \boldsymbol{N_v} \tag{8}$$

其中, l 为眼球中心与角膜中心的距离, 一般采用 5.3 mm[98] 或忽略不计; λ 为角膜中心到视线落点之间的距离, 可以用眼球中心 $\boldsymbol{E}$、光轴单位方向向量 $\boldsymbol{N_o}$、视轴单位方向向量 $\boldsymbol{N_v}$ 以及系统标定出的屏幕在相机坐标系的位置关系表示为: $\lambda = -\frac{(\boldsymbol{E}+l\boldsymbol{N_o})\cdot\boldsymbol{N_s}+n}{\boldsymbol{N_v}\cdot\boldsymbol{N_s}}$, 对于屏幕上的任意一点 $\boldsymbol{S}$, $\boldsymbol{N_s}$ 和 n 满足 $\boldsymbol{S}\cdot\boldsymbol{N_s} = -n$, $\boldsymbol{N_s}$ 和 n 可以通过系统标定获得[28].

基于深度相机的三维视线估计方法需要系统标定和用户标定, 系统标定包括相机参数标定及屏幕位置标定, 用户标定包括眼球中心标定以及光轴与视轴之间的转换关系标定. Sun 等[28] 为了提高头部姿态跟踪的精度, 借助内眼角点来计算眼球中心坐标, 用户标定过程还需要标定眼球半径以及眼球中心与内眼角点矢量, 最终视线估计的精度为 1.38° ∼ 2.71°. 该方法在只使用一个 Kinect 的简单系统中实现了在自由头动状态下的视线估计, 但系统标定中的屏幕标定太繁琐, 需要用一个长为 1 800 mm 的棍子, 一头指向屏幕角点, 另一头指向相机视线范围内的某一点, 用来标定屏幕角点, 还需要换屏幕四个不同的角点进行多次实验; 并且用户标定的参数较多, 建立的方程较多, 增加了标定过程的复杂性. Zhou 等[89] 利用双眼模型来提高视线估计的精度, 并提出了一种只需要一个标定点的用户标定方法, 最终视线精度为 1.99°. 该方法中将视轴用虹膜中心与视线落点的连线近似表示, 导致视线精度不高.

基于深度相机的三维视线估计方法系统配置简单, 且允许头部自由运动, 但其运用的基本原理是以眼球中心为起点, 根据已知长度的方向向量的矢量叠加, 直接换算到空间中的视线落点, 这非常依赖于每一个量的精确检测和计算, 如果眼球中心、虹膜中心、光轴方向向量或视轴方向向量中的任意一个量偏离了正确值, 都会造成最终的视线落点不准确, 这也是该方法精度不高的一个主要原因. 实际上, 基于深度相机的三维视线估计方法需要采集和检测人眼或人脸的若干三维特征, 例如: 眼角点的三维坐标等. 因此, 高精度的三维视线估计十分依赖三维测量设备的检测精度, 如果三维成像设备的检测精度足够高, 根据人眼的成像模型, 可以高精度重构人眼光轴.

2.2.2　基于普通相机的三维视线估计方法

基于普通相机的三维视线估计方法通常利用光源反射以及瞳孔折射等几何成像关系, 求解与视线相关的空间三维信息, 如: 角膜中心、光轴方向向量、光轴与视轴的转换关系等, 估计出三维视线, 与屏幕相交得到视线落点. 根据采用的硬件系统的复杂程度可分为: 单相机单光源系统、单相机多光源系统和

多相机多光源系统, 下面分别对这三种系统配置下的三维视线估计方法进行论述.

1) 单相机单光源系统

单相机单光源系统三维视线估计模型如图 10 所示, 光源 $\boldsymbol{L}$ 在角膜外表面形成反射, 反射点为 $\boldsymbol{G}$, 在相机成像面上成像为普尔钦斑 $\boldsymbol{g}$, 法线为角膜中心 $\boldsymbol{C}$ 与反射点 $\boldsymbol{G}$ 的连线. 设角膜曲率半径为 R, 则存在: $|\boldsymbol{G}-\boldsymbol{C}|=R$. 根据反射平面可知, 光源 $\boldsymbol{L}$、反射点 $\boldsymbol{G}$、相机光心 $\boldsymbol{O}$ 以及角膜中心 $\boldsymbol{C}$ 四点共面, 表示为:

$$(\boldsymbol{L}-\boldsymbol{O})\times(\boldsymbol{G}-\boldsymbol{O})\cdot(\boldsymbol{C}-\boldsymbol{O})=0 \tag{9}$$

其中, 反射点 $\boldsymbol{G}$ 用成像面上的普尔钦斑 $\boldsymbol{g}$ 表示为: $\boldsymbol{G}=\boldsymbol{O}+u(\boldsymbol{O}-\boldsymbol{g})$, u 为比例系数.

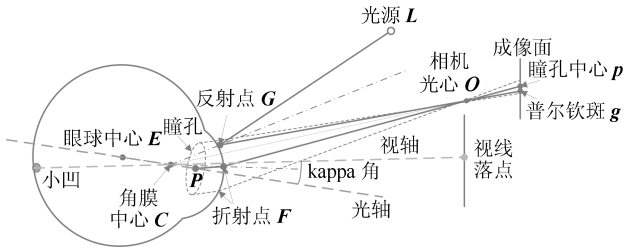


图 10 单相机单光源系统三维视线估计模型

Fig. 10 3D gaze estimation model based on one-camera-one-light-source system

根据入射角与反射角相等, 可得:

$$\frac{(\boldsymbol{L}-\boldsymbol{G})\cdot(\boldsymbol{G}-\boldsymbol{C})}{|\boldsymbol{L}-\boldsymbol{G}|}=\frac{(\boldsymbol{O}-\boldsymbol{G})\cdot(\boldsymbol{G}-\boldsymbol{C})}{|\boldsymbol{O}-\boldsymbol{G}|} \tag{10}$$

瞳孔在角膜外表面存在折射, 则存在: $|\boldsymbol{F}-\boldsymbol{C}|=R$. 根据折射平面可知, 瞳孔中心 $\boldsymbol{P}$、折射点 $\boldsymbol{F}$、相机光心 $\boldsymbol{O}$ 以及角膜中心 $\boldsymbol{C}$ 四点共面, 表示为:

$$(\boldsymbol{F}-\boldsymbol{O})\times(\boldsymbol{C}-\boldsymbol{O})\cdot(\boldsymbol{P}-\boldsymbol{O})=0 \tag{11}$$

其中, 折射点 $\boldsymbol{F}$ 用成像面上的瞳孔中心 $\boldsymbol{p}$ 表示为: $\boldsymbol{F}=\boldsymbol{O}+m(\boldsymbol{O}-\boldsymbol{p})$, m 为比例系数.

根据折射定律可得:

$$n_1\cdot\frac{|(\boldsymbol{F}-\boldsymbol{C})\times(\boldsymbol{P}-\boldsymbol{F})|}{|\boldsymbol{P}-\boldsymbol{F}|}= n_2\cdot\frac{|(\boldsymbol{F}-\boldsymbol{C})\times(\boldsymbol{O}-\boldsymbol{F})|}{|\boldsymbol{O}-\boldsymbol{F}|} \tag{12}$$

其中, n_1 为房水和角膜的有效折射率, n_2 为空气折射率.

利用上述关系式计算出角膜中心 $\boldsymbol{C}$ 和瞳孔中心 $\boldsymbol{P}$ 的坐标, 则眼球的光轴可以根据求得的角膜中心 $\boldsymbol{C}$ 和瞳孔中心 $\boldsymbol{P}$ 重建. 为了得到视线方向, 需要获取光轴与视轴之间的转换关系, 一般通过用户标定过程计算. 在用户标定过程中, 用户依次盯视屏幕上的多个已知标定点, 则根据上述公式求解出的角膜中心 $\boldsymbol{C}$ 和标定点坐标 $\boldsymbol{S}$ 表示出视轴单位方向向量. 根据盯视多个标定点对应的多组光轴单位方向向量与视轴单位方向向量, 可以利用最小二乘法求解光轴与视轴的转换矩阵 $\boldsymbol{M}$, 如式 (13) 所示[54].

$$\boldsymbol{V_p}=\boldsymbol{M}\boldsymbol{V_o} \tag{13}$$

则在视线估计过程中, 根据得到的此刻的角膜中心 $\boldsymbol{C}$ 和瞳孔中心 $\boldsymbol{P}$ 计算光轴单位方向向量, 通过式 (13) 计算出视轴单位方向向量, 那么根据此刻的角膜中心 $\boldsymbol{C}$ 和视轴单位方向向量可以建立视线方向, 与系统标定出的屏幕方程相交, 得到视线落点.

而 Shih 等[99] 证明了当人眼参数未知的情况下, 用空间三维的角膜中心计算三维视线需要至少两个相机和至少两个光源. 因此, 基于单相机单光源系统的三维视线跟踪方法通常依赖于根据人类平均值或人眼模型参数设定人眼的一些不变参数, 如: 角膜曲率半径、瞳孔中心与角膜中心之间的距离等, 将这些眼球参数作为已知量, 用于三维视线估计[98, 100]. Guestrin 等[98] 指出, 如果角膜曲率半径、瞳孔中心到角膜中心之间的距离以及房水的折射率已知的情况下, 在包含一个相机和一个光源的最简化的系统配置下, 需要满足头部位置相对于系统保持静止或者已知眼睛与相机之间的距离. Ohno 等[100] 使用一个网络相机和一个红外 LED 阵列对三维视线进行了估计, 其中将因人而异的角膜曲率半径、瞳孔中心到角膜中心之间的距离以及眼球内外的折射率, 用文献中的固定值表示, 最终得到的视线用瞳孔中心和角膜中心确定的光轴代替表示, 未考虑人眼光轴与视轴之间的偏角.

利用单相机单光源系统的三维视线估计方法需要已知一些眼球不变参数来辅助计算三维空间中的点坐标, 进而重建光轴和视轴, 得到视线落点. 然而每一个用户的眼球参数往往都是不同的, 利用某一固定值进行视线估计会引入系统误差, 视线落点的精度会因此降低. 近几年基于单相机单光源系统的三维视线跟踪方法的研究成果较少, 需要解决的问题是如何在该系统配置下利用简单的用户标定过程标定出这些人眼不变参数, 用于三维视线估计. 这样既可以规避因人眼不变参数引入的误差, 提高视线精度, 也可以突破已有文献对三维视线估计所需硬件系统配置的限制, 实现简化硬件系统配置下的三维视线估计, 促进视线跟踪技术在手机、平板电脑等移动便携式设备的广泛应用.

北京科技大学的迟健男课题组对单相机单光源系统的三维视线估计进行了深入研究. 针对瞳孔检测对光源的需求以及瞳孔在角膜外表面产生折射引出的复杂的折射方程, 课题组选取虹膜作为眼球的

研究特征, 在单相机单光源系统配置下, 根据空间圆形目标的三维重建方法[117] 以及光轴与视轴之间固定的位置关系, 将三维视线估计所需的人眼不变参数, 即: 虹膜半径、kappa 角, 通过用户标定过程标定出来. 三维视线估计过程中, 利用标定的虹膜半径计算光轴的方向向量及角膜中心, 然后根据光轴与视轴之间的转换关系计算视轴的方向向量, 结合角膜中心建立三维视线, 并与屏幕相交得视线落点. 该方法解决了在单相机单光源系统中进行三维视线估计需要预先设定人眼不变参数的问题, 通过简单的用户标定过程计算人眼不变参数, 将个体差异考虑在内, 提高了视线的精度, 且虹膜特征可以在自然光照条件下进行检测, 基于虹膜特征的三维视线估计为今后用于单相机无光源系统奠定了基础.

2) 单相机多光源系统

在单相机多光源系统中, 设光源个数为 N, 光源 $\boldsymbol{L}_i$ $(i = 1, 2, \cdots, N)$ 在角膜外表面产生反射, 反射点为 $\boldsymbol{G}_i$, 在相机中成像为普尔钦斑 $\boldsymbol{g}_i$. 设 $\boldsymbol{G}_i = \boldsymbol{O} + k_i(\boldsymbol{O} - \boldsymbol{g}_i)$, 其中, k_i 为比例系数, 则存在:

$$|\boldsymbol{G}_i - \boldsymbol{C}| = R \tag{14}$$

根据入射光线、法线、反射光线共面可知, 反射平面中 $\boldsymbol{L}_i, \boldsymbol{G}_i, \boldsymbol{C}, \boldsymbol{O}, \boldsymbol{g}_i$ 共面, 则在 $N \geq 2$ 时, 反射平面的交线即为角膜中心 $\boldsymbol{C}$ 和相机光心 $\boldsymbol{O}$ 构成的直线[98]. 因此, 角膜中心 $\boldsymbol{C}$ 的坐标表示为: $\boldsymbol{C} = \boldsymbol{O} + t\boldsymbol{e}_{\boldsymbol{cross}}$, 其中, $\boldsymbol{e}_{\boldsymbol{cross}}$ 用两个反射平面交线的方向向量表示为: $\boldsymbol{e}_{\boldsymbol{cross}} = \frac{[(\boldsymbol{L}_1-\boldsymbol{O})\times(\boldsymbol{g}_1-\boldsymbol{O})]\times[(\boldsymbol{L}_2-\boldsymbol{O})\times(\boldsymbol{g}_2-\boldsymbol{O})]}{|[(\boldsymbol{L}_1-\boldsymbol{O})\times(\boldsymbol{g}_1-\boldsymbol{O})]\times[(\boldsymbol{L}_2-\boldsymbol{O})\times(\boldsymbol{g}_2-\boldsymbol{O})]|}$, t 为比例系数. 根据反射原理可知: 入射角等于反射角, 则满足下式:

$$\frac{\boldsymbol{O} - \boldsymbol{G}_i}{|\boldsymbol{O} - \boldsymbol{G}_i|} + \frac{\boldsymbol{L}_i - \boldsymbol{G}_i}{|\boldsymbol{L}_i - \boldsymbol{G}_i|} = 2\left(\frac{\boldsymbol{G}_i - \boldsymbol{C}}{|\boldsymbol{G}_i - \boldsymbol{C}|} \cdot \frac{\boldsymbol{L}_i - \boldsymbol{G}_i}{|\boldsymbol{L}_i - \boldsymbol{G}_i|}\right) \times \frac{\boldsymbol{G}_i - \boldsymbol{C}}{|\boldsymbol{G}_i - \boldsymbol{C}|} \tag{15}$$

式 (14) 可以列出 N 个方程, 根据式 (15) 可以列出 $3N$ 个方程, 共有 $4N$ 个方程, 未知数有 $N + 2$ 个, 即: k_i, t, R. 故通过求解非线性方程组可以解出角膜中心 $\boldsymbol{C}$ 的坐标和角膜曲率半径 R[76, 80].

然后根据瞳孔的折射作用, 瞳孔边缘点在角膜上的几何成像模型满足:

$$n_1 \cdot \frac{|(\boldsymbol{F}_j - \boldsymbol{C}) \times (\boldsymbol{P}_j - \boldsymbol{F}_j)|}{|\boldsymbol{P}_j - \boldsymbol{F}_j|} = n_2 \cdot \frac{|(\boldsymbol{F}_j - \boldsymbol{C}) \times (\boldsymbol{O} - \boldsymbol{p}_j)|}{|\boldsymbol{O} - \boldsymbol{p}_j|} \tag{16}$$

$$|\boldsymbol{F}_j - \boldsymbol{C}| = R \tag{17}$$

其中, n_1 为房水和角膜的有效折射率, n_2 为空气折射率, $j = 1, 2, \cdots, T$ 为 T 个瞳孔边缘点, 由于空间各瞳孔边缘点到空间瞳孔中心之间的距离为瞳孔半径, 可得表达式:

$$|\boldsymbol{P}_j - \boldsymbol{P}| = |\boldsymbol{P}_k - \boldsymbol{P}| \tag{18}$$

其中 $j \neq k$, $j, k = 1, 2, \cdots, T$.

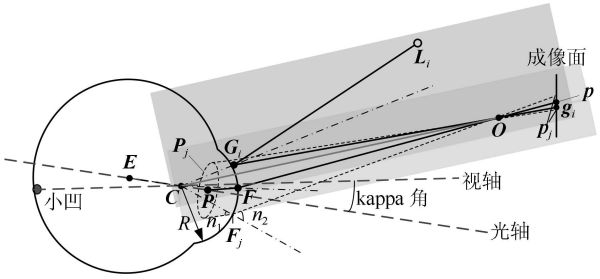


图 11　单相机多光源系统三维视线估计模型

Fig. 11　3D gaze estimation model based on one-camera-multi-light-source system

基于上述求解出来的角膜中心 $\boldsymbol{C}$、角膜曲率半径 R, 求解式 (16) ∼ (18) 可得到瞳孔中心 $\boldsymbol{P}$ 的坐标[98, 101]. 则利用角膜中心 $\boldsymbol{C}$ 和瞳孔中心 $\boldsymbol{P}$ 重建眼球的光轴, 并根据式 (13) 标定出的光轴与视轴之间的转换关系, 确定视线方向及视线落点.

与此经典方法不同的是, Morimoto 等[102] 提出的基于单相机双光源系统的三维视线估计方法是以球面光学表面理论为基础, 利用 Gullstrand 眼球模型中的一些人眼参数, 包括: 角膜曲率半径、角膜折射率、瞳孔中心与角膜中心之间的距离, 来计算空间中的角膜中心和瞳孔中心, 这两点的连线用于计算视线. 该方法不需要进行用户标定, 且允许自由头动, 但没有考虑光轴与视轴之间的偏角, 以及不同用户的眼睛之间存在的个体差异.

基于单相机多光源的三维视线跟踪方法[98, 101−105] 从视线估计的几何模型角度是可行的, 但如果不根据人类平均值去定义一些眼球不变参数, 则需要求解大量的非线性方程组, 求解速度较慢, 与视线跟踪所追求的实时性相违背, 并且, 得到的结果是数值解, 精度与人为设定的求解非线性方程组的终止条件密切相关.

3) 多相机多光源系统

多相机多光源系统在不需要已知任何眼睛参数的情况下, 就可以实现三维视线估计. 目前多相机多光源系统的三维视线估计方法从本质上与单相机多光源系统的类似, 需要求解出角膜中心, 然后重建光轴, 利用光轴与视轴之间的转换关系, 确定视线方向, 最后根据求解得到的角膜中心和视线方向确定三维视线. 但由于增加了相机个数, 与单相机多光源

系统相比, 三维空间点及眼球参数的计算更加简单.

角膜中心是根据角膜反射进行求解, 一般原理如图 12 所示. 两个及两个以上的光源在角膜外表面形成反射, 每一个光源形成的反射平面相交可以得到一条经过角膜中心和相机光心的直线, 设光源为 $\boldsymbol{L}_i$ $(i = 1, 2, \cdots)$, 相机光心为 $\boldsymbol{O}_t$ $(t = 1, 2, \cdots)$, $\boldsymbol{L}_i$ 产生的角膜反射点在相机 t 中的成像点为 $\boldsymbol{g}_{it}$, 则角膜中心 $\boldsymbol{C}$ 满足:

$$(\boldsymbol{L}_i - \boldsymbol{O}_t) \times (\boldsymbol{g}_{it} - \boldsymbol{O}_t) \cdot (\boldsymbol{C} - \boldsymbol{O}_t) = 0 \tag{19}$$

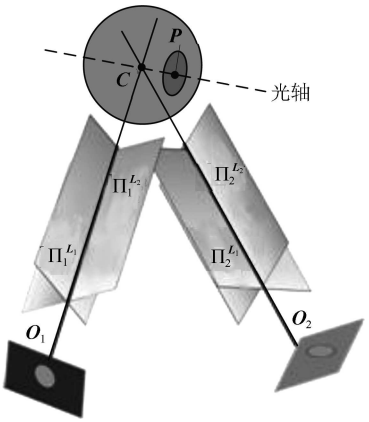


图 12 多相机多光源系统中角膜中心的求解

Fig. 12 Solution of the cornea center in multi-camera-multi-light-source system

在双相机系统中, 经过角膜中心和第一个相机光心的直线与经过角膜中心和第二个相机光心的直线相交, 可以得到角膜中心 $\boldsymbol{C}$ 的坐标[98−99, 106−114].

为了得到视线, 除了角膜中心坐标之外, 还需要重建光轴来转换出视轴的方向. 光轴经过角膜中心、瞳孔中心、虹膜中心、眼球中心, 一般利用角膜中心和瞳孔中心进行求解. 基于多相机系统的光轴重建方法可以归纳为以下两种方法: 折射平面法和直接求解法.

折射平面法是根据角膜中心和瞳孔中心都位于瞳孔中心折射平面内, 因此利用每个由角膜中心、成像平面上的瞳孔中心和相机光学中心确定的折射平面相交, 交线经过角膜中心和瞳孔中心, 用来表示眼球的光轴[98, 107, 109, 112], 如图 13 所示.

设相机光心分别为 $\boldsymbol{O}_1$, $\boldsymbol{O}_2$, 瞳孔中心在两个相机中的成像点分别为 $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, 则光轴的单位方向向量为:

$$\boldsymbol{V_o} = \frac{[(\boldsymbol{O}_1 - \boldsymbol{p}_1) \times (\boldsymbol{C} - \boldsymbol{O}_1)] \times [(\boldsymbol{O}_2 - \boldsymbol{p}_2) \times (\boldsymbol{C} - \boldsymbol{O}_2)]}{|[(\boldsymbol{O}_1 - \boldsymbol{p}_1) \times (\boldsymbol{C} - \boldsymbol{O}_1)] \times [(\boldsymbol{O}_2 - \boldsymbol{p}_2) \times (\boldsymbol{C} - \boldsymbol{O}_2)]|} \tag{20}$$

直接求解法是根据瞳孔成像几何模型, 利用多个相机采集的人眼图像中的瞳孔特征, 直接计算出瞳孔中心的坐标, 分为考虑瞳孔折射的方法[99, 110]和不考虑瞳孔折射的方法[108, 111, 114], 其中, 不考虑瞳孔折射的方法是利用小孔成像原理计算出虚拟的瞳孔中心, 作为人眼的瞳孔中心. 然后利用求解得到的角膜中心和瞳孔中心重建光轴[99, 108, 110−111, 114].

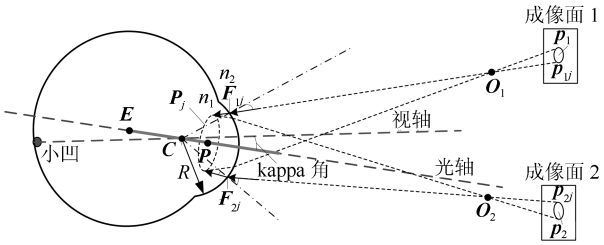


图 13 多相机多光源系统中折射平面法求解光轴

Fig. 13 Refraction plane method for solving the optical axis in multi-camera-multi-light-source system

根据上述计算得到的光轴, 与标定出的光轴与视轴之间的转换关系, 根据式 (13) 计算得到视轴方向向量, 结合计算得到的角膜中心, 即可估计出三维视线, 与屏幕平面相交可得视线落点.

多相机多光源视线跟踪系统可以简化用户标定过程, 实现单点用户标定. Guestrin 等[108] 指出, 在至少两个相机和至少两个光源的视线跟踪系统中, 如果光轴与视轴之间的转换关系未知, 则标定时只需要盯视某一固定点即可实现三维视线估计[98], 但用多个标定点进行标定获得的结果鲁棒性更高[108]. 多相机视线跟踪系统可以将窄视场角相机和宽视场角相机组合起来, 既能跟踪头部姿态, 解决自由头动的影响, 又能高分辨率拍摄人眼, 提升视线估计精度. Beymer 等[115] 采用一个广角的立体视觉系统检测头部运动, 根据检测到的自由头动下的人脸, 用一个主动的窄视场立体视觉系统跟踪人眼并拍摄人眼的高分辨率图像, 文献对立体视觉系统标定、眼睛特征检测方法等作了详细的阐述, 屏幕上视线落点的精度可以达到 0.6°.

多相机多光源系统采用立体视觉, 直接获得使用者的三维眼睛特征, 进而获取三维视线方向, 用户标定过程简单, 能够更容易地获得更精确的视线估计结果, 并且可以计算自由头动状态下视线在任何物体上的注视点. 但硬件系统配置相对复杂, 造成系统标定和相机标定相对复杂.

2.2.3 小结

三维视线跟踪方法根据眼球、相机、屏幕之间的几何关系计算眼球的不变特征, 如: 角膜曲率半径、kappa 角等, 以及眼球的可变特征, 如: 眼球中心、角膜中心、虹膜中心、瞳孔中心等, 然后计算眼睛的光轴方向向量和视轴方向向量, 最后估计三维视线和计算视线落点. 不同的三维视线跟踪方法各具特点, 如表 2 所示.

基于深度相机的三维视线估计方法需要借助深度相机获取深度信息, 从而在无光源系统中获取眼球中心、虹膜中心等三维空间点的坐标, 用于计算视线落点. 该方法系统配置简单, 模型简单, 且允许头部在一定范围内运动, 但用户标定过程复杂, 且视线精度非常依赖于人眼特征的检测精度. 基于普通相机的三维视线估计方法根据光源反射及瞳孔折射等几何关系建立方程, 用于求解角膜中心、瞳孔中心等三维空间点的坐标, 进而重建光轴并估计三维视线, 与系统标定出的屏幕方程相交得到视线落点. 根据采用的硬件系统配置的不同, 分为单相机单光源系统、单相机多光源系统和多相机多光源系统. 单相机单光源系统的系统配置简单, 通过预先设定人眼不变参数的值, 求解估计三维视线所需的眼球参数, 未考虑个体差异性, 故精度较低. 单相机多光源系统理论上可以根据几何关系计算空间三维的眼球参数信息, 如: 角膜中心、瞳孔中心等, 但需求解复杂的非线性方程组, 运算速度较慢, 且得到的是数值解, 或者通过预先设定人眼不变参数的值, 精度较低. 多相机多光源系统基于立体视觉系统可以直接获得空间三维信息, 简化用户标定过程, 视线精度高, 且利用宽视场角相机可以跟踪头部姿态, 解决自由头动的问题, 但复杂的硬件系统配置造成系统标定过程复杂.

2.3　总结

基于特征的视线跟踪方法利用瞳孔、虹膜、眼角、面部标志等人眼特征或人脸特征建立视线估计模型, 从而计算注视方向或注视点. 基于特征的视线跟踪方法包括二维视线跟踪方法和三维视线跟踪方法, 下面分别从系统标定、用户标定、视线精度、头部运动及应用场景等方面对二维视线跟踪方法和三维视线跟踪方法进行分析, 如表 3 所示.

二维视线跟踪方法是通过人眼特征参数直接与

表 2　三维视线跟踪方法的特点对比

Table 2　Comparison of the characteristics of three-dimensional gaze tracking methods

方法	相机	光源	误差主要来源	头动影响	视线信息	优点	缺点
基于深度相机的三维视线估计方法	1	0	眼球中心、虹膜中心等中间量的精确计算	头部一定范围运动	视线落点	系统简单、模型简单	用户标定复杂、视线精度非常依赖于人眼特征检测的精度
基于普通相机的三维视线估计方法	1	1	预先设定的人眼不变参数	头部一定范围运动	三维视线	系统简单、便于实现	未考虑个体差异性, 精度较低
	1	≥2	普尔钦斑检测、人眼不变参数的求解或预先设定	头部一定范围运动	三维视线	系统相对简单、可计算空间三维信息	解非线性方程组求取人眼不变参数, 运算速度较慢, 精度较低
	≥2	≥1	系统标定	头部自由运动	三维视线	用户标定简单、视线精度高	系统复杂、系统标定复杂、成本较高

表 3　二维视线跟踪方法与三维视线跟踪方法的比较

Table 3　Comparison of two-dimensional gaze tracking methods and three-dimensional gaze tracking methods

类型	系统配置	系统标定	用户标定	视线精度	头部运动	应用场景
二维视线跟踪方法	单相机无光源系统	–	复杂	较低	×	系统简单, 精度较低, 应用场合有限
	单相机单光源系统	–	复杂	较高 (头部固定)	×	头部固定或微小运动, 适用于穿戴式或遥测式设备, 适合人因分析、认知疾病诊断等应用
	单相机多光源系统	简单	复杂	较高	√	头部较小运动范围, 适用于穿戴式或遥测式设备, 适合心理分析、助残等应用
	多相机多光源系统	复杂	复杂	高	√	头部较大运动范围, 精度要求较高, 适合虚拟现实等应用
三维视线跟踪方法	单相机无光源系统	复杂	复杂	较低	√	需获取真实彩色图像, 精度要求低, 适用于移动便携式设备、网络终端设备
	单相机单光源系统	复杂	较复杂	较低	√	在三维空间物体上的注视点, 精度要求较低, 适用于移动便携式设备, 如手机、平板电脑等
	单相机多光源系统	复杂	较复杂	较高	√	在三维空间物体上的注视点, 精度要求一般, 适用于穿戴式或遥测式设备, 适合虚拟现实等应用
	多相机多光源系统	复杂	简单	高	√	在三维空间物体上的注视点, 精度要求较高, 适用于穿戴式或遥测式设备, 适合虚拟现实等应用

视线落点建立二维模型, 其中未涉及到相机的内参数, 故不需要标定相机和屏幕位置. 在单相机多光源系统中需要标定光源之间的相对位置关系, 而在多相机系统中需要标定不同相机之间的转换关系. 为了得到人眼特征参数与视线落点之间的关系, 都需要通过复杂的用户标定过程求解中间参量. 在单相机无光源系统中, 由于眼角点检测较难, 故视线精度较低; 在单相机单光源系统中, 头部保持固定的情况下可以保证较高的精度, 但是当头部偏离标定位置时, 视线落点的精度显著下降; 在单相机多光源系统中, 借助多个光源可以适当简化模型, 提高视线精度; 而在多相机系统中, 可以实现头部运动状态下的高精度的视线落点估计. 二维视线跟踪系统比较简单, 目前所能实现的视线落点的精度较高, 由于采用的是视线参数和视线落点之间的映射模型, 类似于"黑箱"结构, 在某些应用场合, 如虚拟屏幕的视线落点估计时, 二维视线跟踪是不可替代的.

三维视线跟踪方法与二维视线跟踪方法相比, 需要对系统进行标定, 且标定过程相对较复杂, 系统标定包括: 相机的内参数标定、多相机之间关系标定以及光源、屏幕的位置标定. 基于单相机系统的三维视线估计方法的用户标定过程需要标定眼球中心或者人眼的一些不变参数, 如: 角膜曲率半径、光轴与视轴之间的转换关系等, 过程较复杂; 而在多相机系统中, 基于立体视觉可以直接计算出三维空间点坐标, 用于求取人眼不变参数, 故用户标定过程简单. 单相机无光源系统三维视线估计根据三维空间点与矢量的叠加直接计算视线落点, 精度与每一个中间量的精确计算有关; 单相机单光源系统需要通过预先设定人眼不变参数才能求得三维视线, 未考虑个体差异性, 故视线精度较低且普适性差; 单相机多光源系统可以利用多个光源的反射计算三维视线, 计算过程相对复杂, 实时性较差, 但与单相机单光源系统相比, 精度较高; 而基于多相机多光源系统可以有效地提高视线的精度, 但系统标定复杂. 三维视线跟踪技术表征了视线的本质特征, 能够简化用户标定并具备实现全自由头动的可能.

随着移动式设备的发展, 以及大数据、云计算、移动互联网等技术的兴起, 视线跟踪技术的发展具有其必要性. 表 4 对比了近五年提出的部分基于

表 4 近五年部分基于特征的视线跟踪方法对比

Table 4 Comparison of feature-based gaze tracking methods in recent five years

文献	相机	光源	类型	特征	头动	精度	特殊说明
Yu 等[37]	1	0	二维映射	虹膜、眼角	头部固定	水平 0.99°, 竖直 1.33°	
Skodras 等[39]	1	0	二维映射	眼角、眼睑、脸部定位点	头部固定	1.42° ± 1.35°	免相机标定
Xia 等[40]	1	0	二维映射	眼角、瞳孔	头部固定	水平 0.84°, 竖直 0.56°	简单映射模型
Cheung 等[41]	1	0	二维映射	虹膜、眼角	运动范围: 15.5° × 15.5° × 5°	2.27°	
熊春水等[27]	1	1	二维映射	瞳孔、普尔钦斑	运动范围: 100 mm × 100 mm × 100 mm	1.05°	引入增量学习, 单点标定
Blignaut 等[48]	1	1	二维映射	瞳孔、普尔钦斑	头部固定	0.5°	标定点个数 ≥12
Banaeeyan 等[52]	1	2	二维映射	虹膜、眼角	头部固定	水平 0.928°, 竖直 1.282°	
Cheng 等[62]	1	5	二维交比值	瞳孔、普尔钦斑	自由运动	0.70°	动态正切平面
Arar 等[63]	1	5	二维交比值	瞳孔、普尔钦斑	自由运动	1.15°(5 点标定)	图像分辨率低, 标定点少
Zhang 等[67]	1	8	二维 HN	瞳孔、普尔钦斑	运动范围: 80 mm × 100 mm × 100 mm	0.56°	双目视觉
Ma 等[69]	1	4	二维 HN	瞳孔、普尔钦斑	自由运动	1.23° ± 0.54°	
Choi 等[70]	1	4	二维 HN	瞳孔、普尔钦斑	头部固定	1.03° ∼ 1.17°	单点标定
Kim 等[73]	1	4	二维 HN	瞳孔、普尔钦斑	仅平转	0.76°	
Zhou 等[89]	1 (Kinect)	0	三维	虹膜、眼角	运动范围: 11.63° × 13.44° × 9.43°	1.99°	
Sun 等[28]	1 (Kinect)	0	三维	虹膜、眼角	自由运动	1.28° ∼ 2.71°	
Li 等[92]	1 (Kinect)	0	三维	虹膜、眼角	自由运动	水平 3.0°, 竖直 4.5°	
Cristina 等[11]	1	0	三维	虹膜、眼角	运动范围 (5° 内)	1.46° ∼ 1.93°	
Wang 等[77]	1	4	三维	瞳孔、普尔钦斑	自由运动	1.3°	免个体标定
Lai 等[107]	2	2	三维	瞳孔、普尔钦斑	运动范围: 50 mm × 25 mm × 50 mm	1.18°	
Bakker 等[113]	3	3	三维	瞳孔、普尔钦斑	自由运动	0.5°	
Lidegaard 等[116]	4	4	三维	瞳孔、普尔钦斑	自由运动	1.5°	头戴式

特征的视线跟踪方法, 可以看出目前视线跟踪技术在实现头部自由运动、提高视线精度及简化标定过程三个方面的研究不断深入, 大部分新提出的视线跟踪方法可以在头动状态下实现高精度的视线估计[11,27−28, 62−63, 69, 77, 107], 并且对于减少光源数量以实现无光源系统的自由头动下的三维视线估计有所突破, 借助一个深度相机来获取人脸特征及人眼特征, 不仅解决了单相机系统下的头部运动问题, 而且简化了进行三维视线估计的硬件系统配置[28, 89].

3 基于特征的视线跟踪系统关键技术讨论

3.1 二维视线跟踪系统的头动问题

头动问题是二维视线估计的重要问题, 如前所述, 二维视线估计通过视线参数到视线落点之间的映射模型估计使用者的屏幕视线落点, 一般映射模型为二次多项式. 当使用者的头部位于标定位置时可以获得高精度的视线落点, 当头部离开标定位置时, 视线精度显著下降. 研究人员对二维视线估计的头动问题进行了大量研究, 提出了各类解决问题的方法, 如采用不同头部位置标定和选择不同的映射模型等, 目前二维视线估计可以在小范围的头动下实现较精确的视线落点估计, 但仍然无法实现大范围全自由头动下的视线估计, 这也是制约其推广应用的原因之一. 二维视线跟踪系统虽然从原理上与三维视线跟踪系统不同, 但本质上仍然是视线方向的估计问题, 因此二维视线估计问题的完全解决, 还需采用二维和三维相结合的方法, 才能实现真正的大范围全自由头动下的视线估计.

3.2 二维视线跟踪系统的简化用户标定问题

用户标定问题是二维视线估计的另一个重要问题, 也是制约其广泛应用的主要问题. 一般情况下, 二维映射模型为包含 12 个未知数的两个二次多项式, 往往需要使用者盯视屏幕上 9 个及 9 个以上的标定点, 或跟踪盯视移动的覆盖更大屏幕区域的标定点, 才能获得高精度的视线映射模型. 很多文献也都在尝试提出新的映射模型以减少标定点的个数, 但效果不够理想, 如将标定点简化为 2 个时, 无论标定点如何分布, 视线高精度都有所下降, 而且普适性差. 目前对于二维视线跟踪系统, 视线估计的精度还是需要靠用户盯视多个标定点来保障. 根据视线跟踪技术的本质原理来说, 简化二维视线跟踪系统用户标定过程还需要采用二维和三维相结合的方法.

3.3 三维视线跟踪系统的 kappa 角标定问题

kappa 角的标定问题是目前三维视线跟踪系统还没有完全解决全自由头动问题的根本原因. 三维视线估计的一般思路是: 首先在用户标定过程中确定眼球的 kappa 角; 然后在视线估计过程中, 根据眼球的可视特征估计眼球光轴的方向; 最后根据检测到的光轴方向和标定的 kappa 角, 将眼球光轴的方向换算成眼球视轴的方向. 可见, 对于三维视线估计系统, kappa 角的标定是用户标定中重要的且不可回避的内容.

假设头部自由移动和转动, 眼球围绕眼球中心自由转动, 这样头部和眼球的运动的叠加构成了眼球的三维自由运动. 由上述眼球的结构特点可知, 如图 6 (a) 所示拍摄眼睛图像提取眼球的可视化特征 (瞳孔、虹膜), 根据可视化特征推导出的计算特征 (空间的角膜中心坐标、空间的瞳孔中心坐标) 都位于眼球光轴上, 无任何眼球光轴以外的特征可以用来表述眼球的姿态, 因此对眼球空间状态的描述缺乏足够的信息, 即在头部运动和眼球运动的叠加下, 眼球的空间运动如果包含绕眼球光轴的转动的分量, 那么这个转动分量是无法通过眼球的特征 (可视特征和计算特征) 来表征的. 即使在配置红外光源的视线跟踪系统中, 光源在角膜上形成了反射亮斑, 由于角膜近似球形, 亮斑的位置仍然无法反映眼球绕其光轴转动的信息. 而 kappa 角是眼球光轴和视轴之间的空间角, 在确定了眼球光轴的空间位置后, 如果无法确定眼球的空间状态, 特别是眼球绕其光轴的转动量, 则无法确定在眼球任何空间位置时的空间 kappa 角.

因此, 在三维视线跟踪系统中, 既使在每个时刻都能检测出眼球光轴的方位, 也不一定能准确推算出眼球视轴的方位. 这也是目前的三维视线跟踪系统无法做到在全自由头动下进行视线估计的原因. 具体证明如下:

根据眼球的结构可知, 眼球除了其光轴和光轴上的点以外, 没有其他可以由可视特征计算出来的特征点或线用于建立眼球坐标系. 因此, 以眼球的角膜中心为坐标系原点, 眼球的光轴为 z 轴, 以与 z 轴垂直的平面内的任意两条相互垂直的直线作为 x 轴和 y 轴, 建立眼球坐标系. 如图 14 所示, 假设在用户标定过程中, 已经标定出了眼球在标定位置处的眼球光轴和视轴之间的 kappa 角. 在视线估计过程中, 眼球从用户标定位置 $\boldsymbol{C}_0$ 变化到另一个空间位置 $\boldsymbol{C}_1$, 其位置和姿态的改变, 可以用坐标系 $\boldsymbol{C}_0$-xyz 的旋转和平移来表征. 眼球从标定位置 $\boldsymbol{C}_0$ 变化到任意空间位置 $\boldsymbol{C}_1$ 的过程, 就是眼球坐标系 $\boldsymbol{C}_0-xyz$ 平移和旋转到坐标系 $\boldsymbol{C}_1-x'y'z'$ 的过程.

假设在系统相机坐标系下, 眼球在标定位置 $\boldsymbol{C}_0$ 处光轴方向向量为 $\boldsymbol{V}_{\boldsymbol{o}0}=[x_{o0}, y_{o0}, z_{o0}]^{\mathrm{T}}$, 角膜中心为 $\boldsymbol{C}_0=[x_{c0}, y_{c0}, z_{c0}]^{\mathrm{T}}$. 在任意视线估计位置 $\boldsymbol{C}_1$ 时的光轴方向向量为 $\boldsymbol{V}_{\boldsymbol{o}1}=[x_{o1}, y_{o1}, z_{o1}]^{\mathrm{T}}$, 角膜中心为 $\boldsymbol{C}_1=[x_{c1}, y_{c1}, z_{c1}]^{\mathrm{T}}$. 当眼球坐标系从标定位置

$\boldsymbol{C}_0$ 旋转和平移到任意位置 $\boldsymbol{C}_1$ 时, 其平移量为:

$$\boldsymbol{T} = [x_{c1} - x_{c0}, y_{c1} - y_{c0}, z_{c1} - z_{c0}]^{\mathrm{T}} \tag{21}$$

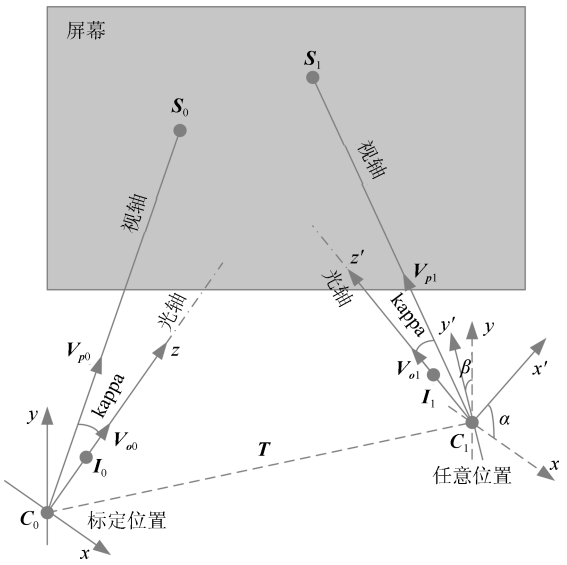


图 14 Kappa 角示意图

Fig. 14 Diagram of kappa angle

其旋转矩阵 $\boldsymbol{R}$ 可以设为具有 3 个自由度的正交旋转矩阵, 则:

$$\boldsymbol{R} = (\boldsymbol{I} + \boldsymbol{B})(\boldsymbol{I} - \boldsymbol{B})^{-1}, \boldsymbol{B} = \begin{bmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{bmatrix} \tag{22}$$

其中, a, b, c 为三个独立的未知数. 则标定位置的眼球光轴与任意位置的眼球光轴的之间的变换矩阵即为两个位置处眼球坐标系之间的变换矩阵:

$$\boldsymbol{V_{o}}_1 = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{T} \\ \boldsymbol{O} & 1 \end{bmatrix} \boldsymbol{V_{o}}_0 \tag{23}$$

由式 (23) 可得:

$$\begin{bmatrix} x_{o1} - x_{o0} \\ y_{o1} - y_{o0} \\ z_{o1} - z_{o0} \end{bmatrix} = \begin{bmatrix} 0 & -z_{o0} - z_{o1} & -y_{o0} - y_{o1} \\ -z_{o0} - z_{o1} & 0 & x_{o0} + x_{o1} \\ y_{o0} + y_{o1} & x_{o0} + x_{o1} & 0 \end{bmatrix} \begin{bmatrix} c \\ b \\ a \end{bmatrix} \tag{24}$$

由于式 (24) 中参数矩阵的秩为 2, 所以无法求出 a, b, c 三个独立变量. 此时旋转矩阵 $\boldsymbol{R}$ 无解.

总之, kappa 角是一个空间角, 由于眼球的特殊结构, 无法建立描述眼球姿态的坐标系, 那么在头部运动时, 特别是头部做滚转运动时, 将无法确定相对于光轴的 kappa 角方位, 也就无法从光轴转换到视轴.

3.4 三维视线跟踪的系统配置和标定问题

三维视线跟踪系统通过眼球的成像模型和空间几何模型重构眼球的光轴和视轴, 就目前的技术现状而言, 三维视线跟踪系统存在系统配置和系统标定两个问题.

1) 系统配置. Shih 等[99] 证明, 三维视线跟踪系统在无预先设定人眼参数的情况下需要单相机配两个及两个以上的光源或双相机配两个及两个以上的光源, 即构成单相机多光源系统或多相机多光源系统, 才能由眼球的可视特征确定角膜中心. 而目前三维视线估计的一般思路都是通过检测角膜中心, 然后重建光轴和视轴. 这样的视线估计思路, 不可避免地造成系统的配置比较复杂, 阻碍了三维视线估计的应用, 特别是在移动式和便携式设备高速发展的今天.

2) 系统标定. 三维视线跟踪系统不止系统复杂, 还涉及系统标定, 即需要标定出光源在系统相机坐标系下的坐标和屏幕 (或关注对象) 在系统相机坐标系下的方程, 无论是遥测式系统, 还是穿戴式系统. 这也是一个比较有挑战性的问题, 因为一般情况下系统的光源并不在系统相机的视野范围内, 无法通过系统相机直接测量系统光源的坐标. 目前对于有些特定的系统可能通过加工过程的尺寸控制来实现系统标定, 但对于自由布设的系统, 或进行三维视线估计研究的实验系统, 系统标定仍然是一个具有挑战性和无法回避的问题.

4 视线跟踪技术的应用

视线跟踪技术在人机交互、医学诊断、国防军事、交通安全、人因分析、虚拟现实等多个领域有着广泛的应用.

在人机交互中, 视线跟踪技术可以使用户通过视线与机器进行交流, 用视线代替手动操作[118−124], 形成眼控鼠标, 如图 15 (a) 所示. 视线跟踪技术是一种自然人机交互手段, 是多模自然人机交互的重要通道之一. 视线跟踪技术也是残障人士的人机交互方式[123−124], 如图 15 (b) 所示的眼控智能轮椅系统, 老年人以及行动存在障碍的残疾人可以用眼睛操控面前的计算机界面, 控制轮椅出行和辅助患者的日常生活.

在医学应用中, 视线跟踪技术是精神和认知障碍型疾病筛查和诊断的重要途径, 为多动症、自闭症、精神分裂、抑郁症、阿尔茨海默症等疾病的早期识别与治疗提供了有力的技术支持[125−131]. 图

16 (a) 所示为 NovaSight 公司推出的儿童视力筛查诊断设备, 结合最新的眼球跟踪技术和智能眼镜技术, 只需要儿童观看一段几分钟的视频, 就能够直接诊断出其斜视、弱视、融像以及阅读障碍等问题, 并给出定量的分析结果. 如图 16 (b) 所示的便携式眼球交互眼镜 EyeControl 可以使闭锁综合征患者、重症病房患者等通过眼动控制, 与医生和家人交流沟通, 使其得到更好的治疗. 通过人的眼动特征, 如眼动轨迹、眼跳次数等特征, 还可以分析人的心理活动, 视线跟踪最早、最广泛的应用领域就是心理学分析.

(a) EyeAssist 眼动辅助交互页面[123]
(a) EyeAssist eye movement assistant interaction[123]

(b) 眼控智能轮椅系统
(b) Eye control intelligent wheelchair system

图 15　视线跟踪技术在人机交互中的应用

Fig. 15　Applications of gaze tracking technology in human-computer interaction

(a) 儿童视力筛查诊断
(a) Screening and diagnosis of children's vision

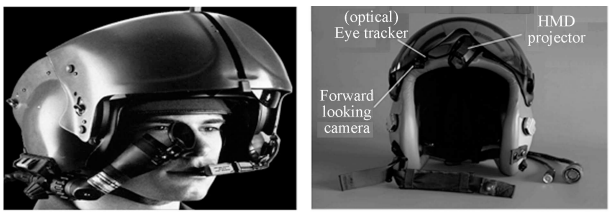
(b) EyeControl
(b) EyeControl

图 16　视线跟踪技术在医学诊断中的应用

Fig. 16　Applications of gaze tracking technology in medical diagnosis

近年来, 随着视线跟踪及微型头显、姿态捕获等相关技术的发展, 视线跟踪技术在军事领域的应用越来越受到关注[132], 可应用于多智能体集群协同指挥、机载武器系统观瞄和人机融合武器控制等方面, 提高武器的智能化水平, 图 17 (a) 为加拿大空军提出的集成眼动探测单元的新型飞行员头盔, 用于机载武器的控制. 如图 17 (b) 所示, 澳大利亚已研制出内置眼动跟踪瞄准的 HMD 系统的样机, NLR HMD 为澳大利亚研制的一款内置眼动跟踪瞄准功能的 HMD 工程样机.

在智能交通领域, 视线跟踪技术可用于驾驶分心和驾驶疲劳监测, 如图 18 (a) 所示, 通过在线检测驾驶员视线的扫视频率、视线范围等指标, 可以判别驾驶员是否处于驾驶疲劳和驾驶分心状态[133−134]. 图 18 (b) 所示为通用汽车 Super Cruise 驾驶辅助系统, 其中通过车内方向盘附近的微型摄像头与红外传感器实时监测驾驶员视线是否保持在路面上, 在车辆自动巡航过程中, 当监测到驾驶员有分心状态时, 系统会提示驾驶员马上接管车辆. 视线跟踪技术还可用于驾驶意图的判别, 尤其在特殊路况行驶过程中, 通过对驾驶员的视线进行统计分析, 可以获得驾驶员在弯道驾驶、换道驾驶时的行为特征.



(a) 飞行员头盔
(a) Pilot helmet

(b) NLR 内置眼动跟踪 HMD
(b) NLR built-in eye tracking HMD

图 17　视线跟踪技术在军事领域的应用

Fig. 17　Applications of gaze tracking technology in military

(a) 驾驶过程中分心与疲劳检测[133]
(a) Distraction and fatigue detection in driving[133]

(b) 通用汽车 Super Cruise 驾驶辅助系统
(b) Super Cruise driving assistance system

图 18　视线跟踪技术在智能交通领域的应用

Fig. 18　Applications of gaze tracking technology in intelligent transportation

人因分析也是视线跟踪技术应用的一个领域[135−145], 随着视线跟踪技术的进步, 基于眼动跟踪的人因分析必将获得更广泛的应用, 例如: 在网上购物中, 可以通过统计人的视线落点, 判别人们对哪种商品更感兴趣, 如图 19 (a) 所示; 在理财产品销售终端, 通过对视线落点的统计, 可以很精确地分析人们更倾向于哪种交易. 在人机功效学设计中, 视线及其关注点检测是页面设计和布局、大型设备仪表盘布局等的重要依据[141, 143, 145], 如图 19 (b) 所示, 可以使产品更符合人的注视和操作习惯.

视线跟踪设备是虚拟现实系统的重要的组成部分[146−148]. 视线跟踪技术可以通过分析眼球运动对场景更精确地瞄准和定位, 并且可以根据注视点的位置对注视点周边进行渲染, 使得画面与自然环境下看到的体验一样, 防止虚拟与现实进行切换时造

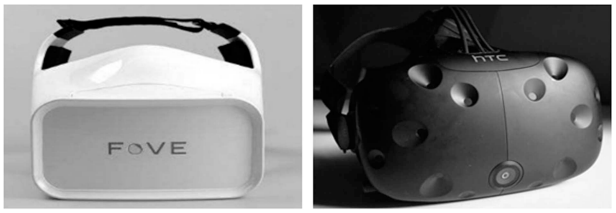
成的晕眩感. 如图 20 (a) 所示的 FOVE 虚拟现实头盔, 其拥有全球首创的眼球追踪技术, 是世界上第一款支持红外眼球追踪技术的 VR 头盔, 用户可以用眼睛来控制显示器, 也可以用眼球追踪打字. 如图 20 (b) 所示的 HTC ViveVR 虚拟现实设备安装了专用的视线追踪装置 aGlass, 通过瞳孔与眼睑的移动来进行动作追踪, 增强了体验感和便捷度, 并且眼睛直视的地方将呈现较高的分辨率, 而非直视的画面将呈现低分辨率, 保证体验的流畅度. 还有一些虚拟现实设备也将加入眼球追踪功能, 如: Oculus Rift、StarVR 等.

(a) 网购中的人因分析
(a) Human factors analysis in online shopping

(b) 右转弯角度指导页面设计[141]
(b) Map perspectives for right-turn guidance[141]

图 19 视线跟踪技术在人因分析中的应用

Fig. 19 Applications of gaze tracking technology in human factors analysis



(a) FOVE 虚拟现实头盔
(a) FOVE virtual reality helmet

(b) HTC ViveVR 内置 aGlass
(b) HTC ViveVR built-in aGlass

图 20 视线跟踪技术在虚拟现实中的应用

Fig. 20 Applications of gaze tracking technology in virtual reality

5 总结与展望

视线跟踪技术通过检测人眼的视觉注意方向, 实现人机交互和对人的神经和行为特征进行分析. 从狭义分类来说, 视线跟踪技术主要是通过近眼设备来精确地检测眼动信息. 从广义范围来说, 从远距离进行人的注意力判断、群体交流分析等技术都属于视线跟踪范畴. 视线跟踪技术有着广泛的应用背景和较长时间的研究历史, 其中, 基于特征的视线跟踪技术更具有实用价值.

但由于许多瓶颈技术的制约, 其发展和应用远未达到普适的水平, 仍然作为一种专用的技术应用于人机交互、虚拟现实、人因分析等领域. 在未来的研究中, 如何突破目前的各种技术瓶颈, 研制理想的视线跟踪设备, 使其真正成为 “即拿即用” 的人机交互通道, “无约束和意识” 的人因分析手段、“人机融合” 的控制技术等, 将是未来视线跟踪技术的发展方向. 基于上述目标, 如下研究方向将成为研究的热点:

1) 简化直至免除用户标定. 作为人机交互、人因分析技术, 用户标定一直是阻碍视线跟踪系统应用的重要因素, 特别是广泛应用的二维视线跟踪系统, 需要多个标定点才能计算出视线落点和视线参数之间映射模型的系数, 即使其视线落点精度很高, 也严重地制约了其使用的便捷性. 三维视线跟踪系统可以将标定点减少为单点标定, 但 kappa 角的初始标定似乎是无法回避的. 目前有文献探讨自动标定问题[149], 即在使用视线跟踪设备时, 使视线精度逐渐收敛到较高精度, 但如果使用初期的视线估计精度较低, 可能无法提供准确的视线信息, 这种无意识标定方式比主动标定优越. 因此, 如何采用新技术, 免除用户标定或减少自动标定的时间, 使视线估计在较高的精度要求下快速收敛, 将是未来重要的研究方向.

2) 大范围全自由头动下的高精度视线估计. 目前视线跟踪的技术状态需要将用户的头部限制在一定的范围内, 才能采集高质量的眼睛图像, 进而实现高精度视线估计. 毫无疑问, 这将极大地限制视线跟踪技术的应用领域, 使其局限于专用设备的范畴. 二维视线跟踪系统从技术原理上不支持头部偏离标定位置, 虽经各种改进, 仍然不适合大范围全自由度头部运动下的视线估计. 虽然三维视线跟踪系统从技术原理上支持头部运动, 但由于无法实现全自由头动状态下空间 kappa 角的实时计算, 仍然无法实现在头部各种姿态下精确的视线估计. 受限于眼球独特的结构特点, 全自由头动问题将是视线跟踪技术亟待解决并富于挑战性的瓶颈技术.

3) 简化硬件系统配置下的视线估计. 视线跟踪技术一直在追求允许头部全自由运动和免除用户标定下的视线精确估计. 对这一问题的解决, 目前都是通过采用单相机多光源系统或多相机多光源系统实现三维视线估计来解决. 但硬件系统的复杂化仍然给系统的实现和应用带来障碍. 如何以最简化的硬件配置实现视线估计和简化用户标定, 使视线跟踪技术成为真正的自然人机交互或人因分析手段, 在移动式和便携式设备快速发展的今天有着重要的意义.

4) 适合不同领域和不同环境的视线跟踪技术. 目前, 随着光学探测、机器视觉、人工智能等相关技术的发展, 移动通信、大数据、物联网等新技术应用的驱动, 很多视线跟踪技术潜在的应用领域亟待开展, 例如: 在军事领域, 可以用于飞行员的眼控机载

火控系统, 用视线指挥武器观瞄. 在医学领域, 可以用于认知和精神类疾病的筛查和诊断, 对阿尔兹海默症进行早期识别. 不同领域的应用对视线跟踪系统的各项关键技术提出了不同的要求. 因此, 针对领域需求进行研究, 可以推动视线跟踪技术在不同领域的应用.

References

1 Zhang C, Chi J N, Zhang Z H, Gao X L, Hu T, Wang Z L. Gaze estimation in a gaze tracking system. *Science China Information Sciences*, 2011, **54**(11): 2295−2306

2 Wang K, Wang S, Ji Q. Deep eye fixation map learning for calibration-free eye gaze tracking. In: Proceedings of the 9th Biennial ACM Symposium on Eye Tracking Research & Applications. Charleston, USA: ACM, 2016. 47−55

3 Sigut J, Sidha S A. Iris center corneal reflection method for gaze tracking using visible light. *IEEE Transactions on Biomedical Engineering*, 2011, **58**(2): 411−419

4 Hansen D W, Ji Q. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(3): 478−500

5 Kar A, Corcoran P. A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms. *IEEE Access*, 2017, **5**(1): 16495−16519

6 Hou Zhi-Qiang, Han Chong-Zhao. A survey of visual tracking. *Acta Automatica Sinica*, 2006, **32**(4): 603−617
(侯志强, 韩崇昭. 视觉跟踪技术综述. 自动化学报, 2006, **32**(4): 603−617)

7 Yin Hong-Peng, Chen Bo, Chai Yi, Liu Zhao-Dong. Vision-based object detection and tracking: A review. *Acta Automatica Sinica*, 2016, **42**(10): 1466−1489
(尹宏鹏, 陈波, 柴毅, 刘兆栋. 基于视觉的目标检测与跟踪综述. 自动化学报, 2016, **42**(10): 1466−1489)

8 Zhang X C, Sugano Y, Bulling A. Evaluation of appearance-based methods and implications for gaze-based applications. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Glasgow, UK: ACM, 2019. Article No. 416

9 Jung D, Lee J M, Gwon S Y, Pan W Y, Lee H C, Park K R, et al. Compensation method of natural head movement for gaze tracking system using an ultrasonic sensor for distance measurement. *Sensors*, 2016, **16**(1): Article No. 110

10 Stefanov K. Webcam-based eye gaze tracking under natural head movement. arXiv: 1803.11088, 2018.

11 Cristina S, Camilleri K P. Model-based head pose-free gaze estimation for assistive communication. *Computer Vision and Image Understanding*, 2016, **149**: 157−170

12 Jeni L A, Cohn J F. Person-independent 3D gaze estimation using face frontalization. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Las Vegas, USA: IEEE, 2016. 792−800

13 Mora K A F, Odobez J M. Person independent 3D gaze estimation from remote RGB-D cameras. In: Proceedings of the 2013 IEEE International Conference on Image Processing. Melbourne, Australia: IEEE, 2013. 2787−2791

14 Lu F, Sugano Y, Okabe T, Sato Y. Adaptive linear regression for appearance-based gaze estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**(10): 2033−2046

15 Lu F, Sugano Y, Okabe T, Sato Y. Head pose-free appearance-based gaze sensing via eye image synthesis. In: Proceedings of the 21st International Conference on Pattern Recognition. Tsukuba, Japan: IEEE, 2012. 1008−1011

16 Lu F, Okabe T, Sugano Y, Sato Y. Learning gaze biases with head motion for head pose-free gaze estimation. *Image and Vision Computing*, 2014, **32**(3): 169−179

17 Zhang X C, Sugano Y, Fritz M, Bulling A. Appearance-based gaze estimation in the wild. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015. 4511−4520

18 Wang Xin-Liang. Study of Gaze Tracking Algorithm in Natural Light [Master thesis], South China University of Technology, China, 2014
(王信亮. 自然光下视线跟踪算法研究 [硕士学位论文], 华南理工大学, 中国, 2014)

19 Liu Shuo-Shuo. The Research on Head-Mounted Gaze Tracking Technology [Master thesis], Southeast University, China, 2016
(刘硕硕. 头盔式视线跟踪技术研究 [硕士学位论文], 东南大学, 中国, 2016)

20 Wang Lin. Head-Mounted Gaze Tracking System With Head Motion [Master thesis], University of Science and Technology of China, China, 2014
(王林. 头部可运动的头戴式视线跟踪系统关键技术研究 [硕士学位论文], 中国科学技术大学, 中国, 2014)

21 Zhang Bao-Yu. Algorithms and Robustness Analysis for Solving Nonlinear Equations in 3D Gaze Tracking System [Master thesis], Xidian University, China, 2015
(张宝玉. 3D 视线跟踪系统中的非线性方程组算法与鲁棒性分析 [硕士学位论文], 西安电子科技大学, 中国, 2015)

22 Lu Ya-Nan. Structure Design of the Head-mounted Eye Tracking System and the Acquiring Method of Eye Information [Master thesis], Tianjin University, China, 2017
(鲁亚楠. 头戴式视线跟踪系统结构设计与视线信息获取方法研究 [硕士学位论文], 天津大学, 中国, 2017)

23 Gong De-Lin. Research on Eye Gaze Tracking Technology Based on Monocular Vision [Master thesis], Beijing Institute of Technology, China, 2016
(宫德麟. 基于单目视觉的视线跟踪技术研究 [硕士学位论文], 北京理工大学, 中国, 2016)

24 Lan Shan-Wei. The Research of Eye Gaze Tracking Technology Based on Pupil Corneal Reflection [Master thesis], Harbin Institute of Technology, China, 2014
(兰善伟. 基于瞳孔角膜反射视线追踪技术的研究 [硕士学位论文], 哈尔滨工业大学, 中国, 2014)

25 Qin Hua-Biao, Yan Wei-Hong, Wang Xin-Liang, Yu Xiang-Yu. A gaze tracking system overcoming influences of head movements. *Acta Electronica Sinica*, 2013, **41**(12): 2403−2408
(秦华标, 严伟洪, 王信亮, 余翔宇. 一种可克服头动影响的视线跟踪系统. 电子学报, 2013, **41**(12): 2403−2408)

26 Gao Di, Yin Gui-Sheng, Ma Chun-Guang. Non-invasive eye tracking technology based on corneal reflex. *Journal on Communications*, 2012, **33**(12): 133−139
(高迪, 印桂生, 马春光. 基于角膜反射的非侵入式视线跟踪技术. 通信学报, 2012, **33**(12): 133−139)

27 Xiong Chun-Shui, Huang Lei, Liu Chang-Ping. A novel gaze estimation method with one-point calibration. *Acta Automatica Sinica*, 2014, **40**(3): 459−470
(熊春水, 黄磊, 刘昌平. 一种新的单点标定视线估计方法. 自动化学报, 2014, **40**(3): 459−470)

28 Sun L, Liu Z C, Sun M T. Real time gaze estimation with a consumer depth camera. *Information Sciences*, 2015, **320**: 346−360

29 Lu F, Chen X W, Sato Y. Appearance-based gaze estimation via uncalibrated gaze pattern recovery. *IEEE Transactions on Image Processing*, 2017, **26**(4): 1543−1553

30 Deng H P, Zhu W J. Monocular free-head 3D gaze tracking with deep learning and geometry constraints. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 3162−3171

31 Li B, Fu H, Wen D S, Lo W. Etracker: A mobile gaze-tracking system with near-eye display based on a combined gaze-tracking algorithm. *Sensors*, 2018, **18**(5): Article No. 1626

32 He G, Oueida S, Ward T. Gazing into the abyss: Real-time gaze estimation. arXiv: 1711.06918, 2017.

33 Wojciechowski A, Fornalczyk K. Single web camera robust interactive eye-gaze tracking method. *Bulletin of the Polish Academy of Sciences Technical Sciences*, 2015, **63**(4): 879−886

34 Topal C, Gunal S, Koçdeviren O, Doğan A, Gerek Ö N. A low-computational approach on gaze estimation with eye touch system. *IEEE Transactions on Cybernetics*, 2014, **44**(2): 228−239

35 Zhu J, Yang J. Subpixel eye gaze tracking. In: Proceedings of the 5th IEEE International Conference on Automatic Face Gesture Recognition. Washington, USA: IEEE, 2002. 131−136

36 Shao G J, Che M, Zhang B Y, Cen K F, Gao W. A novel simple 2D model of eye gaze estimation. In: Proceedings of the 2nd International Conference on Intelligent Human-Machine Systems and Cybernetics. Nanjing, China: IEEE, 2010. 300−304

37 Yu M X, Lin Y Z, Tang X Y, Xu J, Schmidt D, Wang X Z, et al. An easy iris center detection method for eye gaze tracking system. *Journal of Eye Movement Research*, 2015, **8**(3): 1−20

38 George A, Routray A. Fast and accurate algorithm for eye localisation for gaze tracking in low-resolution images. *IET Computer Vision*, 2016, **10**(7): 660−669

39 Skodras E, Kanas V G, Fakotakis N. On visual gaze tracking based on a single low cost camera. *Signal Processing: Image Communication*, 2015, **36**: 29−42

40 Xia L, Sheng B, Wu W, Ma L Z, Li P. Accurate gaze tracking from single camera using gabor corner detector. *Multimedia Tools and Applications*, 2016, **75**(1): 221−239

41 Cheung Y M, Peng Q M. Eye gaze tracking with a web camera in a desktop environment. *IEEE Transactions on Human-Machine Systems*, 2015, **45**(4): 419−430

42 Jen C L, Chen Y L, Lin Y J, Lee C H, Tsai A, Li M T. Vision based wearable eye-gaze tracking system. In: Proceedings of the 2016 IEEE International Conference on Consumer Electronics. Las Vegas, USA: IEEE, 2016. 202−203

43 Zhang Chuang, Chi Jian-Nan, Zhang Zhao-Hui, Wang Zhi-Liang. A novel eye gaze tracking technique based on pupil center cornea reflection technique. *Chinese Journal of Computers*, 2010, **33**(7): 1272−1285
(张闯, 迟健男, 张朝晖, 王志良. 一种新的基于瞳孔 – 角膜反射技术的视线追踪方法. 计算机学报, 2010, **33**(7): 1272−1285)

44 Morimoto C H, Mimica M R M. Eye gaze tracking techniques for interactive applications. *Computer Vision and Image Understanding*, 2005, **98**(1): 4−24

45 Kim S M, Sked M, Ji Q. Non-intrusive eye gaze tracking under natural head movements. In: Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. San Francisco, USA: IEEE, 2004. 2271−2274

46 Blignaut P. Mapping the pupil-glint vector to gaze coordinates in a simple video-based eye tracker. *Journal of Eye Movement Research*, 2013, **7**(1): 1−11

47 Blignaut P, Wium D. The effect of mapping function on the accuracy of a video-based eye tracker. In: Proceedings of the 2013 Conference on Eye Tracking South Africa. Cape Town, South Africa: ACM, 2013. 39−46

48 Blignaut P. A new mapping function to improve the accuracy of a video-based eye tracker. In: Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference. East London, South Africa: ACM, 2013. 56−59

49 Hennessey C, Noureddin B, Lawrence P. Fixation precision in high-speed noncontact eye-gaze tracking. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2008, **38**(2): 289−298

50 Sesma-Sanchez L, Villanueva A, Cabeza R. Gaze estimation interpolation methods based on binocular data. *IEEE Transactions on Biomedical Engineering*, 2012, **59**(8): 2235−2243

51 Zhang Tai-Ning, Bai Jin-Jun, Meng Chun-Ning, Chang Sheng-Jiang. Eye-gaze tracking based on one camera and two light sources. *Journal of Optoelectronics · Laser*, 2012,

23(10): 1990−1995
(张太宁, 白晋军, 孟春宁, 常胜江. 基于单相机双光源的视线估计. 光电子 · 激光, 2012, **23**(10): 1990−1995)

52 Banaeeyan R, Halin A A, Bahari M. Nonintrusive eye gaze tracking using a single eye image. In: Proceedings of the 2015 IEEE International Conference on Signal and Image Processing Applications. Kuala Lumpur, Malaysia: IEEE, 2015. 139−144

53 Cerrolaza J J, Villanueva A, Cabeza R. Taxonomic study of polynomial regressions applied to the calibration of video-oculographic systems. In: Proceedings of the 2008 Symposium on Eye Tracking Research & Applications. Savannah, USA: ACM, 2008. 259−266

54 Zhu Z W, Ji Q. Novel eye gaze tracking techniques under natural head movement. *IEEE Transactions on Biomedical Engineering*, 2007, **54**(12): 2246−2260

55 Zhu Z W, Ji Q. Eye gaze tracking under natural head movements. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 918−923

56 Kim J H, Han D I, Min C O, Lee D W, Eom W. IR vision-based LOS tracking using non-uniform illumination compensation. *International Journal of Precision Engineering and Manufacturing*, 2013, **14**(8): 1355−1360

57 Arar N M, Thiran J P. Robust real-time multi-view eye tracking. arXiv: 1711.05444, 2017.

58 Kang J J, Eizenman M, Guestrin E D, Eizenman E. Investigation of the cross-ratios method for point-of-gaze estimation. *IEEE Transactions on Biomedical Engineering*, 2008, **55**(9): 2293−2302

59 Coutinho F L, Morimoto C H. Free head motion eye gaze tracking using a single camera and multiple light sources. In: Proceedings of the 2006 19th Brazilian Symposium on Computer Graphics and Image Processing. Manaus, Brazil: IEEE, 2006. 171−178

60 Yoo D H, Kim J H, Lee B R, Chung M J. Non-contact eye gaze tracking system by mapping of corneal reflections. In: Proceedings of the 5th IEEE International Conference on Automatic Face Gesture Recognition. Washington, USA: IEEE, 2002. 101−106

61 Yoo D H, Chung M J. A novel non-intrusive eye gaze estimation using cross-ratio under large head motion. *Computer Vision and Image Understanding*, 2005, **98**(1): 25−51

62 Cheng H, Liu Y Q, Fu W H, Ji Y L, Yang L, Zhao Y, et al. Gazing point dependent eye gaze estimation. *Pattern Recognition*, 2017, **71**: 36−44

63 Arar N M, Gao H, Thiran J P. Towards convenient calibration for cross-ratio based gaze estimation. In: Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision. Waikoloa, USA: IEEE, 2015. 642−648

64 Huang J B, Cai Q, Liu Z C, Ahuja N, Zhang Z Y. Towards accurate and robust cross-ratio based gaze trackers through learning from simulation. In: Proceedings of the 2014 Symposium on Eye Tracking Research and Applications. Safety Harbor, USA: ACM, 2014. 75−82

65 Coutinho F L, Morimoto C H. Augmenting the robustness of cross-ratio gaze tracking methods to head movement. In: Proceedings of the 2012 Symposium on Eye Tracking Research and Applications. Santa Barbara, USA: ACM, 2012. 59−66

66 Arar N M, Gao H, Thiran J P. A regression-based user calibration framework for real-time gaze estimation. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, **27**(12): 2623−2638

67 Zhang Z Y, Cai Q. Improving cross-ratio-based eye tracking techniques by leveraging the binocular fixation constraint. In: Proceedings of the 2014 Symposium on Eye Tracking Research and Applications. Safety Harbor, USA: ACM, 2014. 267−270

68 Kang J J, Guestrin E D, MacLean W J, Eizenman M. Simplifying the cross-ratios method of point-of-gaze estimation. In: Proceedings of the 30th Canadian Medical and Biological Engineering Conference (CMBEC30). Toronto, ON, Canada: Oxford Univ Press, Jun 2007. 14-17

69 Ma C F, Baek S J, Choi K A, Ko S J. Improved remote gaze estimation using corneal reflection-adaptive geometric transforms. *Optical Engineering*, 2014, **53**(5): Article No. 053112

70 Choi K A, Ma C F, Ko S J. Improving the usability of remote eye gaze tracking for human-device interaction. *IEEE Transactions on Consumer Electronics*, 2014, **60**(3): 493−498

71 Hansen D W, Agustin J S, Villanueva A. Homography normalization for robust gaze estimation in uncalibrated setups. In: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications. Austin, USA: ACM, 2010. 13−20

72 Shin Y G, Choi K A, Kim S T, Ko S J. A novel single IR light based gaze estimation method using virtual glints. *IEEE Transactions on Consumer Electronics*, 2015, **61**(2): 254−260

73 Kim S T, Choi K A, Shin Y G, Kang M C, Ko S J. Remote eye-gaze tracking method robust to the device rotation. *Optical Engineering*, 2016, **55**(8): Article No. 083108

74 Shin Y G, Choi K A, Kim S T, Yoo C H, Ko S J. A novel 2-D mapping-based remote eye gaze tracking method using two IR light sources. In: Proceedings of the 2015 IEEE International Conference on Consumer Electronics. Las Vegas, USA: IEEE, 2015. 190−191

75 Wan Z J, Wang X J, Zhou K, Chen X Y, Wang X Q. A novel method for estimating free space 3D point-of-regard using pupillary reflex and line-of-sight convergence points. *Sensors*, 2018, **18**(7): Article No. 2292

76 Li J F, Li S G. Eye-model-based gaze estimation by RGB-D camera. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Columbus, USA: IEEE, 2014. 606−610

77 Wang K, Ji Q. 3D gaze estimation without explicit personal calibration. *Pattern Recognition*, 2018, **79**: 216−227

78 Park K R. A real-time gaze position estimation method based on a 3-D eye model. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2007, **37**(1): 199−212

79 Villanueva A, Cabeza R. Evaluation of corneal refraction in a model of a gaze tracking system. *IEEE Transactions on Biomedical Engineering*, 2008, **55**(12): 2812−2822

80 Villanueva A, Cabeza R. Models for gaze tracking systems. *Eurasip Journal on Image and Video Processing*, 2007, **2007**(1): Article No. 023570

81 Nakazawa A, Nitschke C. Point of gaze estimation through corneal surface reflection in an active illumination environment. In: Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: Springer, 2012. 159−172

82 Villanueva A, Cabeza R, Porta S. Gaze tracking system model based on physical parameters. *International Journal of Pattern Recognition and Artificial Intelligence*, 2007, **21**(5): 855−877

83 Mansouryar M, Steil J, Sugano Y, Bulling A. 3D gaze estimation from 2D pupil positions on monocular head-mounted eye trackers. In: Proceedings of the 9th Biennial ACM Symposium on Eye Tracking Research & Applications. Charleston, USA: ACM, 2016. 197−200

84 Kacete A, Séguier R, Collobert M, Royan J. Head pose free 3D gaze estimation using RGB-D camera. In: Proceedings of the 8th International Conference on Graphic and Image Processing. Tokyo, Japan: SPIE, 2016. Article No. 102251S

85 Wang X, Lindlbauer D, Lessig C, Alexa M. Accuracy of monocular gaze tracking on 3D geometry. In: Proceedings Springer International Publishing of the 1st Workshop on Eye Tracking and Visualization. Springer, 2017. 169−184

86 Sun L, Song M L, Liu Z C, Sun M T. Real-time gaze estimation with online calibration. *IEEE Multimedia*, 2014, **21**(4): 28−37

87 Cen K F, Che M. Research on eye gaze estimation technique base on 3D model. In: Proceedings of the 2011 International Conference on Electronics, Communications and Control. Ningbo, China: IEEE, 2011. 1623−1626

88 Tamura K, Choi R, Aoki Y. Unconstrained and calibration-free gaze estimation in a room-scale area using a monocular camera. *IEEE Access*, 2017, **6**: 10896−10908

89 Zhou X L, Cai H B, Li Y F, Liu H H. Two-eye model-based gaze estimation from a Kinect sensor. In: Proceedings of the 2017 IEEE International Conference on Robotics and Automation. Singapore: IEEE, 2017. 1646−1653

90 Zhou X L, Cai H B, Shao Z P, Yu H, Liu H H. 3D eye model-based gaze estimation from a depth sensor. In: Proceedings of the 2016 IEEE International Conference on Robotics and Biomimetics. Qingdao, China: IEEE, 2016. 369−374

91 Wang K, Ji Q. Real time eye gaze tracking with Kinect. In: Proceedings of the 2016 23rd International Conference on Pattern Recognition. Cancun, Mexico: IEEE, 2016. 2752−2757

92 Li J F, Li S G. Gaze estimation from color image based on the eye model with known head pose. *IEEE Transactions on Human-Machine Systems*, 2016, **46**(3): 414−423

93 Mora K A F, Odobez J M. Gaze estimation from multimodal Kinect data. In: Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Providence, USA: IEEE, 2012. 25−30

94 Jafari R, Ziou D. Gaze estimation using Kinect/PTZ camera. In: Proceedings of the 2012 IEEE International Symposium on Robotic and Sensors Environments. Magdeburg, Germany: IEEE, 2012. 13−18

95 Xiong X H, Liu Z C, Cai Q, Zhang Z Y. Eye gaze tracking using an RGBD camera: A comparison with a RGB solution. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. Seattle, USA: ACM, 2014. 1113−1121

96 Jafari R, Ziou D. Eye-gaze estimation under various head positions and iris states. *Expert Systems with Applications*, 2015, **42**(1): 510−518

97 Ghiass R S, Arandjelovic O. Highly accurate gaze estimation using a consumer RGB-D sensor. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence. New York, USA: AAAI Press, 2016. 3368−3374

98 Guestrin E D, Eizenman M. General theory of remote gaze estimation using the pupil center and corneal reflections. *IEEE Transactions on Biomedical Engineering*, 2006, **53**(6): 1124−1133

99 Shih S W, Liu J. A novel approach to 3-D gaze tracking using stereo cameras. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2004, **34**(1): 234−245

100 Ohno T, Mukawa N, Yoshikawa A. FreeGaze: A gaze tracking system for everyday gaze interaction. In: Proceedings of the 2002 Symposium on Eye Tracking Research & Applications. New Orleans, USA: ACM, 2002. 125−132

101 Hennessey C, Noureddin B, Lawrence P. A single camera eye-gaze tracking system with free head motion. In: Proceedings of the 2006 Symposium on Eye Tracking Research & Applications. San Diego, USA: ACM, 2006. 87−94

102 Morimoto C H, Amir A, Flickner M. Detecting eye position and gaze from a single camera and 2 light sources. *Object Recognition Supported by User Interaction for Service Robots*, 2002, **4**: 314−317. doi: 10.1109/ICPR.2002.1047459

103 Hennessey C, Lawrence P. Noncontact binocular eye-gaze tracking for point-of-gaze estimation in three dimensions. *IEEE Transactions on Biomedical Engineering*, 2009, **56**(3): 790−799

104 Mujahidin S, Wibirama S, Nugroho H A, Hamamoto K. 3D gaze tracking in real world environment using orthographic projection. *AIP Conference Proceedings*, 2016, **1746**(1): Article No. 020072

105 Meyer A, Bohme M, Martinetz T, Barth E. A single-camera remote eye tracker. *Perception and Interactive Technologies*. Berlin, Heidelberg: Springer, 2006. 208−211

106 Villanueva A, Cabeza R. A novel gaze estimation system with one calibration point. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2008, **38**(4): 1123−1138

107 Lai C C, Shih S W, Hung Y P. Hybrid method for 3-D gaze tracking using glint and contour features. *IEEE Transactions on Circuits and Systems for Video Technology*, 2015, **25**(1): 24−37

108 Guestrin E D, Eizenman M. Remote point-of-gaze estimation with free head movements requiring a single-point calibration. In: Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Lyon, France: IEEE, 2007. 4556−4560

109 Lai C C, Shih S W, Tsai H R, Hung Y P. 3-D gaze tracking using pupil contour features. In: Proceedings of the 22nd International Conference on Pattern Recognition. Stockholm, Sweden: IEEE, 2014. 1162−1166

110 Shih S W, Wu Y T, Liu J. A calibration-free gaze tracking technique. In: Proceedings of the 15th International Conference on Pattern Recognition. Barcelona, Spain: IEEE, 2000. 201−204

111 Zhang K, Zhao X B, Ma Z, Man Y. A simplified 3D gaze tracking technology with stereo vision. In: Proceedings of the 2010 International Conference on Optoelectronics and Image Processing. Haikou, China: IEEE, 2010. 131−134

112 Nagamatsu T, Kamahara J, Tanaka N. Calibration-free gaze tracking using a binocular 3D eye model. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems. Boston, USA: ACM, 2009. 3613−3618

113 Bakker N M, Lenseigne B A J, Schutte S, Geukers E B M, Jonker P P, Van Der Helm F C T, et al. Accurate gaze direction measurements with free head movement for strabismus angle estimation. *IEEE Transactions on Biomedical Engineering*, 2013, **60**(11): 3028−3035

114 Nagamatsu T, Sugano R, Iwamoto Y, Kamahara J, Tanaka N. User-calibration-free gaze tracking with estimation of the horizontal angles between the visual and the optical axes of both eyes. In: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications. Austin, USA: ACM, 2010. 251−254

115 Beymer D, Flickner M. Eye gaze tracking using an active stereo head. In: Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Madison, USA: IEEE, 2003. vol. 2, pp. II-451−II-458

116 Lidegaard M, Hansen D W, Kruger N. Head mounted device for point-of-gaze estimation in three dimensions. In: Proceedings of the 2014 Symposium on Eye Tracking Research and Applications. Safety Harbor, USA: ACM, 2014. 83−86

117 Shiu Y C, Ahmad S. 3D location of circular and spherical features by monocular model-based vision. In: Proceedings of the 1989 IEEE International Conference on Systems, Man and Cybernetics. Cambridge, USA: IEEE, 1989. 576−581

118 Meena Y K, Cecotti H, Wong-Lin K, Dutta A, Prasad G. Toward optimization of gaze-controlled human-computer interaction: Application to hindi virtual keyboard for stroke patients. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018, **26**(4): 911−922

119 Rajanna V, Hammond T. A gaze-assisted multimodal approach to rich and accessible human-computer interaction. arXiv: 1803.04713, 2018.

120 Mazeika D, Carbone A, Pissaloux E E. Design of a generic head-mounted gaze tracker for human-computer interaction. In: Proceedings of the 2012 Joint Conference New Trends in Audio & Video and Signal Processing: Algorithms, Architectures, Arrangements and Applications. Lodz, Poland: IEEE, 2012. 127−131

121 Rantanen V, Vanhala T, Tuisku O, Niemenlehto P H, Verho J, Surakka V, et al. A wearable, wireless gaze tracker with integrated selection command source for human-computer interaction. *IEEE Transactions on Information Technology in Biomedicine*, 2011, **15**(5): 795−801

122 Cristanti R Y, Sigit R, Harsono T, Adelina D C, Nabilah A, Anggraeni N P. Eye gaze tracking to operate android-based communication helper application. In: Proceedings of the 2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing. Surabaya, Indonesia: IEEE, 2017. 89−94

123 Khasnobish A, Gavas R, Chatterjee D, Raj V, Naitam S. EyeAssist: A communication aid through gaze tracking for patients with neuro-motor disabilities. In: Proceedings of the 2017 IEEE International Conference on Pervasive Computing and Communications Workshops. Kona, USA: IEEE, 2017. 382−387

124 Hyder R, Chowdhury S S, Fattah S A. Real-time non-intrusive eye-gaze tracking based wheelchair control for the physically challenged. In: Proceedings of the 2016 IEEE EMBS Conference on Biomedical Engineering and Sciences. Kuala Lumpur, Malaysia: IEEE, 2016. 784−787

125 Bovery M D M J, Dawson G, Hashemi J, Sapiro G. A scalable off-the-shelf framework for measuring patterns of attention in young children and its application in autism spectrum disorder. *IEEE Transactions on Affective Computing*, doi: 10.1109/TAFFC.2018.2890610, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8598852

126 Han J X, Kang J N, Ouyang G X, Tong Z, Ding M, Zhang D, et al. A study of EEG and eye tracking in children with autism. *Chinese Science Bulletin*, 2018, **63**(15): 1464−1473

127 Harezlak K, Kasprowski P. Application of eye tracking in medicine: A survey, research issues and challenges. *Computerized Medical Imaging and Graphics*, 2018, **65**: 176−190

128 Florea L. Future trends in early diagnosis for cognition impairments in children based on eye measurements [Trends in Future I&M]. *IEEE Instrumentation & Measurement Magazine*, 2018, **21**(3): 41−42

129 Tsutsumi S, Tamashiro W, Sato M, Okajima M, Ogura T, Doi K. Frequency analysis of gaze points with CT colonography interpretation using eye gaze tracking system. In: Proceedings of the 2017 SPIE Medical Imaging. Orlando, United States: SPIE, 2017. Article No. 1014011

130 Kattoulas E, Evdokimidis I, Stefanis N C, Avramopoulos D, Stefanis C N, Smyrnis N. Predictive smooth eye pursuit in a population of young men: II. Effects of schizotypy, anxiety and depression. *Experimental Brain Research*, 2011, **215**(3−4): 219−226

131 Wyder S, Hennings F, Pezold S, Hrbacek J, Cattin P C. With gaze tracking toward noninvasive eye cancer treatment. *IEEE Transactions on Biomedical Engineering*, 2016, **63**(9): 1914−1924

132 Lu Shan-Shan, Wang Qing-Min, Yao Yong-Jie, Li Ke-Hua, Liu Qiu-Hong. Application of eye tracking technology in foreign military aviation. *Aircraft Design*, 2015, **35**(3): 61−64
(卢姗姗, 王庆敏, 姚永杰, 李科华, 刘秋红. 视觉追踪技术在国外军事航空领域中的应用. 飞机设计, 2015, **35**(3): 61−64)

133 Rani P S, Subhashree P, Devi N S. Computer vision based gaze tracking for accident prevention. In: Proceedings of the 2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare. Coimbatore, India: IEEE, 2016. 1−6

134 Craye C, Karray F. Driver distraction detection and recognition using RGB-D sensor. CoRR, vol. abs/1502.00250, 2015 http://arxiv.org/abs/1502.00250

135 Jianu R, Alam S S. A data model and task space for data of interest (DOI) eye-tracking analyses. *IEEE Transactions on Visualization and Computer Graphics*, 2018, **24**(3): 1232−1245

136 Takahashi R, Suzuki H, Chew J Y, Ohtake Y, Nagai Y, Ohtomi K. A system for three-dimensional gaze fixation analysis using eye tracking glasses. *Journal of Computational Design and Engineering*, 2018, **5**(4): 449−457

137 O′Dwyer J, Flynn R, Murray N. Continuous affect prediction using eye gaze and speech. In: Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine. Kansas City, USA: IEEE, 2017. 2001−2007

138 Steil J, Müller P, Sugano Y, Bulling A. Forecasting user attention during everyday mobile interactions using device-integrated and wearable sensors. In: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services. Barcelona, Spain: ACM, 2018. Article No. 1

139 Mikalef P, Sharma K, Pappas I O, Giannakos M N. Online reviews or marketer information? An eye-tracking study on social commerce consumers. In: Proceedings of the 16th Conference on e-Business, e-Services and e-Society. Delhi, India: Springer, 2017. 388−399

140 George A, Routray A. Recognition of activities from eye gaze and egocentric video. arXiv: 1805.07253, 2018.

141 Zheng R C, Nakano K, Ishiko H, Hagita K, Kihira M, Yokozeki T. Eye-gaze tracking analysis of driver behavior while interacting with navigation systems in an urban area. *IEEE Transactions on Human-Machine Systems*, 2016, **46**(4): 546−556

142 Low T, Bubalo N, Gossen T, Kotzyba M, Brechmann A, Huckauf A, et al. Towards identifying user intentions in exploratory search using gaze and pupil tracking. In: Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval. Oslo, Norway: ACM, 2017. 273−276

143 Peterson J, Pardos Z, Rau M, Swigart A, Gerber C, McKinsey J. Understanding student success in chemistry using gaze tracking and pupillometry. In: Proceedings of the 2015 International Conference on Artificial Intelligence in Education. Madrid, Spain: Springer, 2015. 358−366

144 Jin Rui-Ming. Research on Application of Eye Tracking Technology in Extracting Web Page Concerns [Master thesis], Xi′an Technological University, China, 2016
(金瑞铭. 视线追踪技术在网页关注点提取中的应用研究 [硕士学位论文], 西安工业大学, 中国, 2016)

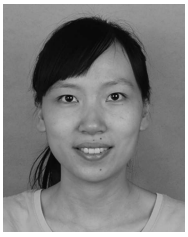
145 Li W C, Braithwaite G, Greaves M, Hsu C K, Lin S C. The evaluation of military pilots attention distributions on the flight deck. In: Proceedings of the 2016 International Conference on Human-computer Interaction in Aerospace. Paris, France: ACM, 2016. Article No. 15

146 Chen S Y, Gao L, Lai Y K, Rosin P L, Xia S H. Real-time 3D face reconstruction and gaze tracking for virtual reality. In: Proceedings of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces. Reutlingen, Germany: IEEE, 2018. 525−526

147 Tripathi S, Guenter B. A statistical approach to continuous self-calibrating eye gaze tracking for head-mounted virtual reality systems. In: Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision. Santa Rosa, USA: IEEE, 2017. 862−870

148 Wibirama S, Nugroho H A, Hamamoto K. Evaluating 3D gaze tracking in virtual space: A computer graphics approach. *Entertainment Computing*, 2017, **21**: 11−17

149 Alnajar F, Gevers T, Valenti R, Ghebreab S. Auto-calibrated gaze estimation using human gaze patterns. *International Journal of Computer Vision*, 2017, **124**(2): 223−236



刘佳惠 北京科技大学自动化学院博士研究生. 主要研究方向为计算机视觉与模式识别. E-mail: ustbljh@163.com
(**LIU Jia-Hui** Ph. D. candidate at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. Her research interest covers computer vision and pattern recognition.)

迟健男 北京科技大学自动化学院副教授. 2005 年获得东北大学模式识别与智能系统系博士学位. 主要研究方向为计算机视觉, 人机交互, 光学测量. 本文通信作者. E-mail: ustbjnc@ustb.edu.cn
(**CHI Jian-Nan** Associate professor at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. In 2005, he received his Ph. D. degree from the Department of Pattern Recognition and Intelligent Systems, Northeastern University. His research interest covers computer vision, human-computer interaction, and optical measurement. Corresponding author of this paper.)

尹怡欣 北京科技大学自动化学院教授. 2002 年获得北京科技大学控制理论与控制工程系博士学位. 主要研究方向为复杂系统的建模与控制, 网络控制, 流程工业自动化. E-mail: yyx@ies.ustb.edu.cn
(**YIN Yi-Xin** Professor at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. In 2002, he received his Ph. D. degree from the Department of Control Theory and Control Engineering, University of Science and Technology Beijing. His research interest covers modeling and control of complex systems, network control, and process industry automation.)