

多相机多光源系统三维视线估计方法研究

王丰

北京科技大学

密 级：_____
加密论文编号：_____

论文题目：多相机多光源系统三维视线估计方法研究

学 号： S20170567

作 者： 王丰

专 业 名 称： 仪器科学与技术

2019 年 12 月 18 日

多相机多光源系统三维视线估计方法研究

Research on multi-camera-multi-light-source system
3D gaze estimation method

研究生姓名：王丰

指导教师姓名：迟健男

北京科技大学自动化学院

北京 100083，中国

Master Degree Candidate: Feng Wang

Supervisor: Jian-nan Chi

School of Automation and Electrical Engineering

University of Science and Technology Beijing

30 Xueyuan Road, Haidian District

Beijing 100083, P.R.CHINA

分类号: TN911.73

密 级: _____

U D C: _____

单位代码: 1 0 0 0 8

北京科技大学硕士学位论文

论文题目: 多相机多光源系统三维视线估计方法研究

作者: 王丰

指 导 教 师: 迟健男 副教授 单位: 北京科技大学

指导小组成员: _____ 单位: _____

_____ 单位: _____

论文提交日期: 2019 年 12 月 18 日

学位授予单位: 北 京 科 技 大 学

致 谢

两年半的研究生生涯转瞬即逝，回顾我在北科的六年半美好时光，真的觉得很幸运很感激。感谢学校对我的培养，感谢老师们和同学们对我的帮助，使我不仅掌握了扎实的专业知识，更树立了正确的人生观、价值观。祝福母校桃李满天下。

我要感谢我的硕士研究生导师迟健男老师，在迟老师的悉心指导下，我逐步深入到视线追踪这个领域，不仅顺利完成了课题的选题、研究以及实验，同时也培养了独立分析问题、解决问题的能力。迟老师对工作的认真负责，对学术的钻研精神和严谨的治学态度，都对我产生了深远的影响，是值得我终身学习的。再次感谢两年半以来迟健男老师对我的悉心指导，祝福迟健男老师工作顺利，身体健康。

另外，我要感谢两位博士前辈——刘佳惠师姐和杨作云师兄，感谢他们对我课题的指导和我生活上的帮助，在两年半的时光建立了深厚的感情，亦师亦友。祝福他们顺利毕业，前程似锦。还要感谢实验室其他同学的帮助和鼓励，我们相互陪伴，共同进步，共同成长。

最后，感谢我的父母，在我漫长求学路上一直无条件支持我鼓励我，使我顺利完成学业。同时，也要感谢各位老师百忙之中抽出时间对我的毕业论文进行审核。

摘 要

视线追踪技术作为一种新型的人机交互技术，应用广泛，前景广阔。如何优化视线追踪算法是提高视线追踪精度的核心问题。在视线追踪多相机多光源系统中，传统的三维视线估计方法通常遵循一个原理，即瞳孔中心的成像投影是图像中瞳孔椭圆的中心。但是根据计算机视觉原理，空间圆形目标的中心的像点不是其椭圆像的中心，即空间圆形目标存在成像畸变。因此，传统方法有其固有的缺陷，会导致视线估计的误差，特别是在眼球相对于相机光轴以大角度斜视的情况下。针对传统方法的不足，本文提出了一种基于瞳孔边缘匹配的双相机双光源系统三维视线估计方法。

在本文中，我们首先对传统方法进行了描述，并对传统方法模型的误差进行了分析，由此提出了本文的方法。之后对本文的相机标定、系统标定以及用户标定进行了介绍，重点介绍了系统标定所使用的单相机辅助标定法，该方法使用单相机借助一个半径已知的标准球，就可以标定出光源和屏幕的位置。接下来详细地论述了本文提出的三维视线估计模型，首先进行角膜曲率中心的检测，之后通过瞳孔边缘匹配的极平面法得到匹配点，由匹配点和其它特征参数可以进行光轴重建，本文提出了三种光轴重建方法。通过 Kappa 角标定得到的转换矩阵，可以实现光轴到视轴的转换，进而估计出视线落点。最后，通过仿真实验和真实实验对本文所提出的方法进行了详细的验证。实验结果证明了本文方法的有效性。

关键词： 三维视线估计，匹配点，光轴，视轴

Research on multi-camera-multi-light-source system 3D gaze estimation method

Abstract

Gaze tracking technology, as a new type of human-computer interaction technology, is widely used and has a bright future. How to optimize the gaze tracking algorithm is the core issue to improve the accuracy of gaze tracking. The traditional 3D gaze direction estimation method generally follow a principle, which is the imaging projection of the pupil center is the center of elliptical pupil in image. But according to the principle of computer vision, the center of the circular object is not the center of its elliptical image, which is called imaging distortion. Therefore, the traditional method has its inherent disadvantages, which can lead to errors in the estimation of the line of sight, especially when the eyeball is squinting at a large angle relative to the optical axis. Aiming at overcoming the deficiency of the traditional method, a 3D gaze estimation method is proposed based on pupil edge matching in the multi-camera-multi -light-source system.

In this paper, we first describe the traditional method and analyze the errors of the traditional method model, and propose the method of this paper. Then we introduce the camera calibration, system calibration and user calibration, and focus on the single camera-assisted calibration method used in system calibration. This method uses a single camera with a standard sphere with a known radius to calibrate the light source and the position of the screen. Next, the 3D line-of-sight estimation model proposed is discussed in detail. First, the corneal curvature center is detected, and then the matching points are obtained by the polar plane method of pupil edge matching. The matching points and other characteristic parameters can be used to reconstruct the optical axis. Three optical axis reconstruction methods are presented. Through the conversion matrix obtained by Kappa angle calibration, the conversion from the optical axis to the visual axis can be realized, and the point of fall of the line of sight can be estimated. Finally, the method proposed in this paper is verified in detail through simulation experiments and real experiments. Experimental results prove the effectiveness of the method.

Key Words: 3D Gaze Estimation, Matching Point, Line of Gaze, Line of Sight

目 录

致 谢.....	I
摘 要.....	III
Abstract.....	V
目 录.....	VII
1 引言.....	1
1.1 课题背景.....	1
1.2 研究目的及意义.....	3
1.3 视线追踪技术概述.....	5
1.3.1 国内外发展状况.....	5
1.3.2 二维视线估计技术概述.....	7
1.3.3 三维视线估计技术概述.....	9
1.4 论文结构及内容.....	11
2 问题描述.....	12
2.1 传统三维视线估计方法概述.....	12
2.2 传统方法模型误差分析及问题提出.....	13
2.3 本文三维视线估计方法概述.....	17
2.4 本章小结.....	17
3 三维视线追踪系统标定.....	18
3.1 相机标定方法.....	18
3.1.1 张正友标定法.....	19
3.1.2 试验场标定法.....	20
3.2 系统标定方法.....	20
3.2.1 系统标定方法概述.....	20
3.2.2 单相机辅助标定法.....	22
3.3 用户标定方法.....	27
3.3.1 角膜半径标定.....	27
3.3.2 Kappa 角标定.....	29
3.4 本章小结.....	31
4 三维视线估计模型.....	32

4.1 三维视线估计方法流程.....	32
4.2 角膜曲率中心估计.....	33
4.3 瞳孔边缘匹配的极平面法.....	34
4.4 光轴估计.....	37
4.4.1 极平面圆锥高线法.....	37
4.4.2 极平面椭圆长轴法.....	38
4.4.3 极平面长轴交点法.....	39
4.5 本章小结.....	40
5 三维视线估计实验分析.....	41
5.1 仿真实验.....	41
5.2 真实实验设备.....	45
5.3 真实实验过程.....	45
5.3.1 用户标定实验.....	45
5.3.2 本文三维视线估计方法实验.....	46
5.3.3 与传统三维视线估计方法结果比较.....	53
5.3.4 光源数量对视线估计精度的影响.....	55
5.3.5 匹配点数量对视线估计精度的影响.....	56
5.4 本章小结.....	57
6 结论.....	59
参考文献.....	61
作者简历及在学研究成果.....	67
独创性说明.....	69
关于论文使用授权的说明.....	69
学位论文数据集.....	71

1 引言

1.1 课题背景

从计算机诞生的那一刻起，就产生了操作者与计算机之间密切的交流互动关系。可以简单的总结为，使用者需要不断地给计算机发出指令，同时也接收计算机对指令执行结果的反馈。这样就形成了一个概念——人机交互。自计算机开始大规模应用以来，人机交互的效率提升始终是计算机应用中一个热门的问题。所谓人机交互技术，是指通过计算机输入、输出设备，以有效的方式实现人与计算机对话的技术^[1]。人机交互技术包括计算机通过输出设备给人提供有关信息，人通过输入设备给计算机输入有关信息。人机交互技术是计算机用户界面设计中的重要内容之一。它与认知学、人机工程学、心理学等学科领域有密切的联系。作为计算机系统的一个重要组成部分，人机交互技术一直伴随着计算机的发展而发展。随着高速计算设备的普及，很多情况下，制约生产力发展的已经不再是设备的硬件性能，而是以交互方式为代表的软件设计。

人机交互技术的发展过程，也是从人适应计算机到计算机不断适应人的发展过程。总体来说，它经历了如下几个阶段^[1]：第一，早期的手工作业阶段。计算机发明初期，由计算机设计者或者参与设计者来使用计算机，他们采用手工操作和依赖机器的方法去适应十分笨拙的计算机；第二，交互命令和作业控制语言阶段。程序员取代了设计者成为计算机的主要使用者。最初，程序通常直接采用机器语言指令或汇编语言编写，通过纸带传输入机或卡读卡输入，通过打印机输出计算结果，人与计算机的交互一般是采用控制键或控制台直接手工操纵。这种形式很不符合人们的习惯，既耗费时间，又容易出错，只有专业的计算机管理员才能做到应用自如。后来，出现了 ALGOL 60、FORTRAN、COBOL、PASCAL 等高级语言，使人们可以用比较习惯的符号形式描述计算过程，交互操作由受过一定训练的程序员即可完成，命令行界面开始出现。这一时期，程序员可采用批量处理作业语言或交互命令语言方式和计算机打交道，虽然要急用许多命令和熟练地敲击键盘，但已经可以用较方便的手段来调试程序、了解计算机执行的情况。通过命令行界面，人们可以通过问答式对话、文本菜单或命令语言等方式来进行人机交互；第三，图形用户界面（GUI）阶段。GUI 是指采用图形方式显示的计算机操作用户界面，图形用户界面的出现使人机交互方式发生了巨大变化。GUI 的主要特点是桌面隐喻、WIMP 技术和直接操纵。GUI 简单易学、减少了键盘操作，

使得普通用户也可以熟练地使用，从而拓展了用户群，使计算机技术得到了普及；第四，网络用户界面阶段。此阶段的实现基础为超文本标记语言 HTML 及超文本传输协议 HTTP；第五，多通道、多媒体的智能人机交互阶段。20 世纪 90 年代，美国麻省理工学院 Negroponte 教授的媒体实验室在多通道的用户界面人机交互，例如手势、语音等方面做了大量开创性的工作，从此多通道、多媒体的交互方式开始被提及。目前，人类常用的自然交互方式——基于语音和笔的交互技术，包括手写识别、笔迹交互、语音识别、语音合成、数字墨水等，已经有了很大的发展。



图 1-1 车载设备进行人机交互

近年来，强大的社会需求产生了各种各样的应用场景。以 AR、VR、MR 为代表的计算机系统的拟人化和以笔记本电脑、智能手机、智能手表为代表的计算设备的微型化、随身化、嵌入化，成为了计算机交互发展的两个重要方向。在此背景下，以传统的鼠标和键盘为代表的 GUI 技术，成为了制约新思路的瓶颈。要想实现自然和谐的人机交互关系，需要在进行交互设计时考虑物理、社会等不同的计算环境，理解人机交互的复杂本质，探索与之相关的社会的、自然的和认知的环境以及人们使用计算机的原因，将领域知识应用到系统设计中，并在此过程中逐步形成人机交互新方法。多通道、多媒体的本质是利用人的多种感觉通道（触觉、听觉、视觉等）、多种动作通道（语音、手写、姿势、视线等），以并行、非精确的方式（可见或不可见的）与计算机环境进行交互，可以提高人机交互的自然性和高效性，多通道、多媒体的智能人机交互对我们既是一个挑战，也是一个极好的机遇^[1]。

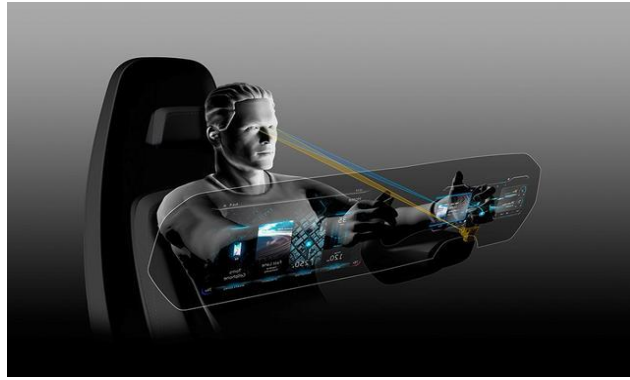


图 1-2 视线追踪在车辆辅助驾驶中的应用

视线追踪是一个综合性的研究课题，它的研究方向包括生理学，心理学，光学，计算机科学领域；技术手段包括人脸检测，人眼特征检测与跟踪，头部姿态检测，视线建模等多种识别技术；目的是要寻找能够客观、准确的反映用户实时的注视方向与视线落在空间位置的有效方法。视线追踪系统就是将视线落点实时的输入到计算机中来完成用户的命令的系统，这样不仅避免了传统输入设备需要手参与的冗余，而且增加了设备使用的灵活性与趣味性，当与计算机操作结合时可以模拟鼠标键盘操作；当与机械操作结合时可以操作各种机器等。因其易用性，正在被广泛应用于各种交互场合^[2]，例如军事，游戏，体育研究，虚拟现实，助残，车辆辅助驾驶，人因分析等领域，用于提升传统的交互方式，该方式已被证实比传统输入设备更加的高效，前景十分广阔^[3]。视线追踪技术作为一种新型的人机交互方式，通过视线追踪技术实现视线控制的新兴技术，已成为人机交互领域内活跃的讨论话题。

1.2 研究目的及意义

正是由于视线追踪技术包括了众多的研究方向，它的应用领域也十分广泛。当将视线追踪技术应用于计算机领域创建视线追踪系统模拟鼠标操作时，需要对人脸图像进行处理、提取人眼的特征参数，人眼建模，建立人眼空间坐标与计算机屏幕坐标点之间的映射关系，并进行实时的反馈。

传统的人机交互方式局限于鼠标、键盘和触摸屏等输入设备，它使得人们按固定的模式与计算机之间进行通信和交互，这是一种以计算机为中心的交互方式。现在最迫切的是提高人机交互技术的灵活性、健壮性以及类似于人与人之间交流中的自然性。而视线追踪技术就能满足这样的交互特性，在现有的众多视线追踪技术中，瞳孔角膜反射法具有对被试者无干扰，操作简单方便，远控、非接触等优点，相对于以往的接触式（头戴式）视线追踪系

统最大限度的减少了对被试者的干扰和限制。这样的视线追踪技术将具有更大的应用潜力。

在另一方面，作为一种自然人机交互手段，视线追踪技术一直在追求头部能够自由运动条件下的视线精确估计。对这一问题的解决，目前都是通过改进硬件系统来解决，包括采用多相机三维视线估计方法和单相机多光源视线估计系统等。而硬件系统的复杂化仍然给系统的实现和应用带来障碍。因此如何以最简化的硬件配置实现自由头动状态下的视线估计成为视线追踪技术发展的瓶颈问题，阻碍了视线追踪技术的推广和应用。

对于三维视线追踪系统，一般而言需要多个相机或者多个光源实现^[4]，虽然每种方法在组成上有很大的不同，但视线追踪的过程大同小异，首先都是通过图像处理的方法得到某些特征参数，然后通过一定的模型对眼球进行重建并解算得到注视方向。为了达到交互的目的，需要把眼球坐标系下的视线方向转换到系统相机坐标系下，因此，需要知道光源和屏幕在系统相机坐标系下的三维坐标，在得到这些信息之后，就可以通过计算得到视线在屏幕上的落点，从而实现交互的目的。

在三维视线追踪系统的使用中，标定是不可或缺的过程。标定分为用户标定和系统标定。视线追踪系统的最终目的是要定位出视线方向和视线落点，而视线方向取决于眼睛的视轴，视轴是指眼睛中心凹与角膜曲率中心的连线。视轴不可被直接检测，往往通过另外两个量：光轴和“卡帕角”（Kappa 角）来进行确定^[5]，如下图 1-3 所示。对于不同的眼球，其 Kappa 角是不一样的，也就是说每个人眼球的 Kappa 角都不同，因此在视线追踪系统中，Kappa 角是需要进行标定的值，标定这个量的过程叫做用户标定。

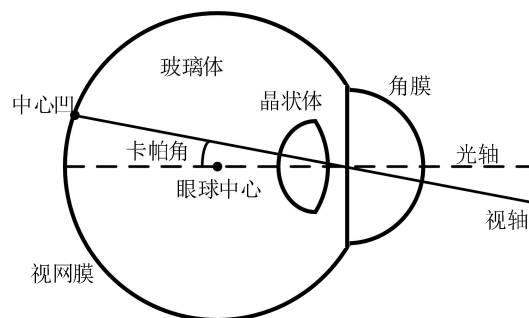


图 1-3 眼球简易结构示意图

对于某个固定的眼球，Kappa 角和光轴之间的夹角固定，所以，为了知道视轴方向，还需要知道光轴方向，光轴方向的求解依赖于光源照射到瞳孔之后的反射成像，所以在求解视线方向的过程中，光源的位置必须已知，而

光源位置是通过系统标定得到的。视线落点是指屏幕和眼球视线方向的交点，若视线方向已知，要求解视线落点，就必须要知道屏幕在空间中的位置，而这个位置需要标定得到，光源和屏幕的标定合起来称为系统标定。用图像的方式实现一个非侵入式的视线追踪系统，相机的参数是需要知道的，求解相机参数的过程称为相机标定。因此，视线追踪系统的标定，包括用户标定、系统标定和相机标定是进行视线追踪的基础。

本文主要研究的是一种基于双相机双光源的新的三维视线追踪模型。由于空间圆形目标的成像失真，三维圆形目标中心的图像不是其成像椭圆的中心，基于双相机双光源系统的传统三维视线估计方法是用瞳孔中心的虚拟图像替换瞳孔中心的投影，来进行三维视线估计。当眼睛的光轴与相机的光轴之间的倾斜角较大时，就会引起较大的视线估计误差。为了解决这个问题，本文提出了一种新的三维视线估计方法，由于不需要使用瞳孔中心的虚拟图像来估计空间瞳孔中心或折射平面，因此该方法减轻了传统方法中由光轴重建引起的固有误差，提高了视线估计精度。

1.3 视线追踪技术概述

1.3.1 国内外发展状况

国际上对于视线追踪技术的研究起步较早，在 20 世纪初，就已经有机构开始研制眼动相关的仪器，到目前为止，在算法和实用方面都取得了较大的进展。文献^[4]证明了视线追踪系统若想获得三维的视线方向，至少需要使用一个相机和两个以上光源或者两个以上的相机。Morimoto 等人最早提出一种基于瞳孔角膜矢量的视线跟踪算法^[6]，输入参数为瞳孔中心和普尔钦斑坐标，通过多项式拟合方式建立特征参数和视线方向的映射模型，来确定视线落点的位置，但此系统对头部限制较大，头部需要静止，抗干扰能力差。Guestri 等人^[7]讨论了基于瞳孔中心和角膜反射点进行的视线估计方法的一般原理，对于单相机双光源视线追踪系统，给出了针对人眼参数的用户标定模型，需要通过用户盯视屏幕上的 9 个点来标定，需要标定的参数有角膜曲率中心、瞳孔中心、角膜曲率半径等。Witzner H D 和 Zhu Z 以及 Q Ji^[8-10]等人引入了新的方法，采用广义回归神经网络或者支持向量回归机的学习型模型，通过对头部不同位置、眼球不同位置的样本进行学习来建立特征和视线落点之前的映射关系。这种方法无需标定，且允许头部在一定范围内运动，但在使用前需要大量的特征量训练模型，通用性较差。Ohno 和 Beymer^[11-12]等人将视线追踪分为两步，首先采用立体相机检测头部位置，粗定位之后再由高分辨率

的窄角相机对人眼进行成像，通过两步方法来解决头动问题，在此基础上提出了一种使用多个相机和多个光源的三维视线估计算法，但这种计算三维视线方向的算法，需要知道角膜曲率、折射率以及角膜反射光斑中心和瞳孔中心之间的距离等参数，通过这些参数构建一个理想的眼球模型，因此在使用中存在由个体差异引起的误差，且这些系统都需要对光源和屏幕的位置进行标定。近些年，德国 SMI 公司，美国 ASL 公司，尤其瑞典 Tobii 眼部追踪技术公司，在计算机视觉领域较为领先。



图 1-4 瑞典 Tobii 眼动仪

国内对于视线追踪技术的研究开展时间和投入规模都比较小，与国际先进水平有一定的差距，但很多高校和科研机构，都对此技术表现出了很大的兴趣，如天津大学、中国科学技术大学、南京航空航天大学、南开大学、北京科技大学、东南大学、武汉理工大学等^[13]。上世纪八十年代末，中科院上海生理研究所率先开展了国内的研究，张名魁、孙复川等研制了基于红外光电反射的视线测量系统^[14]，通过量化人眼运动时角膜与巩膜的反射光线的变化来进行眼动的测量。上世纪九十年代初开始，西安电子科技大学陈怀琛教授及其课题组着手研发国内第一个头盔式眼动仪^[15]，眼动仪的研究填补了国内空白，推动了国内视线追踪技术的发展。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，对视线追踪技术的要求越加迫切，相关技术和产品的发展有加速的趋势。武汉理工大学采用了基于头部几何模型的视线估计方法，将测试者的头部状态和视线向量结合，实现注视点的估计，方法的缺点在于对头动非常敏感。北京的七鑫易维公司开发了基于二维映射模型的商用眼动仪。上海青研公司开发了单相机实用化产品。华南理工大学的秦华标^[16]和哈尔滨工程大学的高迪^[17]有关于头动下视线估计的相关工作。中科院自动化所熊春水博士有关于单相机系统单点标定的研究工作^[18]。

综上所述，国内视线追踪技术由于开展时间较晚，应用场景较为单一，虽然取得了一定的成绩，开发出了一些产品，但相比国外，总体研究水平不高，无论实用产品的研发还是关键技术与算法的研究都与国外有一定的差距。从检索的文献和专利来看，视线追踪的理论方法和创新研究方面差距更大。随着应用场景的扩展，视线追踪系统向着小型化、简易化发展，成熟、精确、不受制于头动的解决方案的开发还需要各国研究者共同努力。

1.3.2 二维视线估计技术概述

从整体结构上讲，视线追踪技术一般包括两大部分：眼部特征检测与视线参数提取、视线方向与视线落点判别^[19]，如图 1-5 所示。

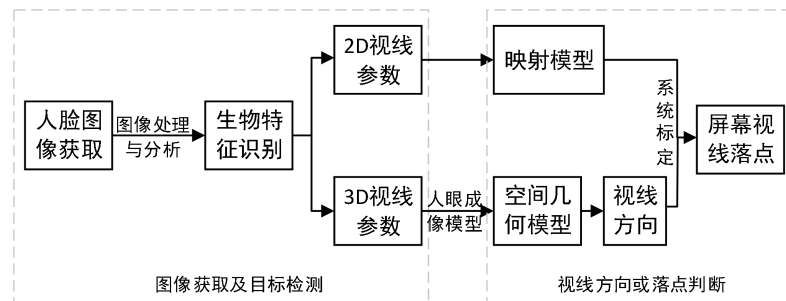


图 1-5 视线追踪技术原理流程图

视线追踪技术作为计算机视觉用作人机交互的核心技术，可以有多种分类，按照人机关系来说，分为侵入式^[6]和非侵入式，非侵入式的视线追踪技术在应用上有着天然的优势；按照系统的组成，分为单相机系统和多相机系统；按照相机离眼睛的远近，分为近眼系统和远眼系统；按照系统的光照环境，可分为自然光照系统和红外光照系统^[21]。自然光照系统在不借助任何辅助光源的情况下，只通过计算机视觉获取人脸图像，进而估计视线方向或视线在屏幕上的落点。无光源系统虽然有明确的应用场合和背景，但很难获得精确的视线估计结果。红外光源系统可以分为主动光源系统和被动光源系统。所谓的主动光源是指通过光源的变化来产生亮瞳和暗瞳现象，进而突出瞳孔特征的系统，其中研究最广泛、最典型的是双环形主动红外光源系统。红外光源系统在视线追踪系统中占有重要地位，红外光源是视线估计的基准，是获得高精度视线方向的必备条件。按照光源的数量，红外系统又可分为单光源系统^[22-24]和多光源系统^[25-27]。

总体上说，视线追踪技术通常可以分为两类：基于二维映射的视线估计方法^[28-29]和基于三维模型的视线估计方法^[9-10,30]。由于很少受相机自身畸变的

影响，三维视线估计方法精度高于二维映射模型。

二维视线估计系统通常由单相机单光源组成，它利用二阶多项式来表示瞳孔光斑向量到视线落点（POR）的直接映射。在该系统中，需要用户盯视屏幕上多个特定点来进行用户标定，从而求解出二阶多项式所需的多项式系数。由于该二维模型使用映射函数来表示眼睛参数与屏幕上的视线落点之间的关系，由一个相机采集人脸图像并提取眼睛的视线参数，只能获得单目视觉图像，通过单目成像无法获得物体的三维位置信息，因此这类系统只能通过图像的视线特征参数到在屏幕上落点之间的映射模型求取视线在屏幕上的视线落点，无法获得视线在系统相机坐标系下真实的空间方向，因此这种系统在自然头部运动方面受到很大的限制。实际标定中，用户的头动对系统的影响较大，因此标定过程需要用户的头部静止不动，并且标定过程较为繁琐，需要用户盯视屏幕上多个标定点才能解算出视线映射模型中的个性化参数。在单相机单光源系统中，为了克服自然头部运动的影响并提高视线估计的精度，一些研究人员基于传统视线映射方法提出了许多的改进方法^[30-33]。另一种二维视线估计方法是交比不变性法，该方法利用投影变换中交叉比率的不变性来估计自然头部运动下的视线落点^[34-37]，如图 1-6 所示。由于瞳孔位于与角膜反射确定的平面不同的平面上，因此交叉比率的模型是一个近似值。尽管交比不变性法的视线追踪系统通常由单相机和四个光源组成，但交比不变性法仍然是二维视线估计方法。基于传统的交比不变性方法，Coutinho^[35]等人提出了用单相机和五个光源来补偿 Kappa 角。Yoo^[37]等人使用了两个摄像头和五个光源，即使在较大的头部运动下，也可以实现交比不变性法的高精度结果。交比不变性法的优点是它不需要标定相机，它仅需要知道相对于屏幕的光源的位置。但是这种方法实际上对肉眼非常敏感，而且，将光源正确放置在计算机屏幕的角上并不容易。

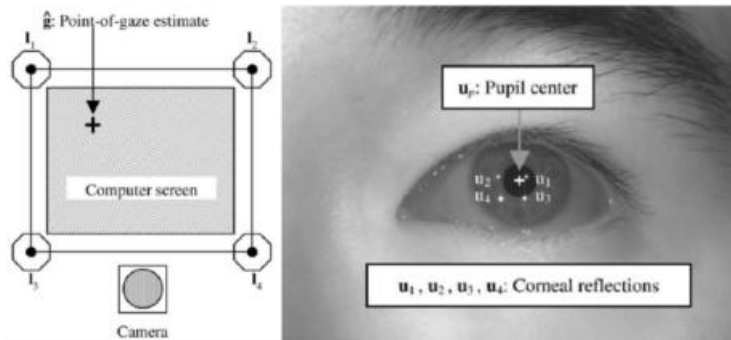


图 1-6 单相机四光源交比值不变法视线落点估计系统示意图

1.3.3 三维视线估计技术概述

三维视线追踪系统通常指单相机多光源和多相机多光源视线追踪系统。一般情况下，单相机单光源系统只能采用二维回归方法估计视线在屏幕上的视线落点，相比之下，单相机多光源系统和多相机系统则可以检测出在系统相机坐标系下的三维的视线方向。只有检测到三维的视线方向，视线方向才能不受头动变化的影响，并且仅需要对估计用户眼球个体参数的用户标定程序进行简化即可。文献证明了视线追踪系统如果要获得三维的视线方向，至少需要一个相机配两个以上光源或者两个以上的相机构成立体视觉系统^[38]。三维视线追踪系统能够通过眼球结构和成像模型估计出在系统相机坐标系下视线的空间方向，因此这类系统可以获得自然头部运动下的三维视线方向，不受使用者头部运动的影响。特别地，对于多相机多光源系统，可以通过立体视觉获得眼球的三维特征信息，不需要通过用户标定程序确定过多的眼球结构参数，如角膜半径、角膜曲率中心到瞳孔中心的距离等，因此三维视线系统能够简化用户标定程序。由此可见，三维视线追踪系统相比二维视线追踪系统有很大优势，三维视线方向估计是视线追踪技术所追求的目标。

Shih^[39]等人证明，若角膜曲率已知，则使用至少两个相机和至少两个光源就可以唯一确定角膜中心三维坐标；若角膜半径已知，则使用至少一个相机和至少两个光源就可以唯一确定角膜中心三维坐标。当仅使用单相机时，通常需要检测瞳孔或虹膜特征以及光源在角膜表面形成的光斑，以便在用户标定过程中确定一些眼睛的固有参数，例如角膜半径，虹膜半径，视轴和光轴之间的角度（Kappa 角），或者瞳孔中心和角膜中心之间的距离等。通过使用这些已知的眼睛参数，大多数估计视线的三维模型方法首先是估计角膜中心三维坐标，然后重建光轴，就可以通过光轴和 Kappa 角估算视线方向。视线和计算机屏幕的交点就是视线落点^[40-43]。Villanueva 和 Cabeza^[39]提出了一种基于单相机多光源系统的三维视线估计的典型方法。首先，通过光源的反射平面和角膜表面的反射点与角膜中心之间距离的相交来求解角膜中心。由于从每个瞳孔边缘点到瞳孔中心的距离相等，因此可以通过基于瞳孔边缘的折射来优化瞳孔中心。最终通过角膜中心和瞳孔中心来重建眼球的光轴，从而确定视线落点。为了简化用户标定程序，基于单相机多光源系统，通过使用一些眼睛参数的经验值，提出了一些视线追踪方法^[44-45]。Morimoto^[44]等人基于球面光学表面的理论，使用球面 Gullstrand 模型的一些已知量（角膜半径，角膜折射率，瞳孔到角膜中心的距离）求解了角膜中心和瞳孔中心的位置，来处理自然头部运动。Hennessey^[45]等人开发了一个系统，该系统包含

一个高分辨率相机和两个光源，以实现低于 1° 视角的精度。单相机多光源系统的视线估计方法主要分为以下两类：（1）眼睛的光轴由两个或多个折射平面直接确定，这些折射平面包括角膜中心，相机的结点，瞳孔中心和瞳孔中心的图像^[39,46]；（2）角膜中心可以从捕获的面部图像中的两个或多个角膜反射来确定，并且瞳孔中心可以通过两个或多个相机坐标系中的瞳孔中心图像来近似估计，因此，可以通过角膜中心和虚拟瞳孔中心来重建眼睛的光轴^[47-49]。因此，一旦通过简单的标定过程估计出了特定对象的 κ 角，就可以通过光轴估计视线的方向。当基于角膜反射确定角膜中心后，可以获得视线。根据是否考虑角膜屈光，可以将这些方法分为两组。Gastrin 和 Eisenman^[47]提出了一种视线落点估计方法，该方法可以实现自然头部运动，需要使用两个相机和四个光源进行单点标定，并分析了视线落点估计误差的主要来源。他们指出，在用户标定过程中使用更多的标定点将更加稳定。Beymer 和 Flicker^[50]使用广角立体音响系统来检测自然头部运动下的面部，并操纵了主动式窄视角 FOV 立体音响系统以高分辨率跟踪眼睛。他们讨论了立体声系统的标定，眼睛检测和跟踪，并估计了监视器上的注视点的准确性在 0.6° 。结果表明，即使不进行用户标定，多相机多光源系统也可以实现高精度的视线落点，并将用户标定的过程简化为单点标定甚至不需要标定。然而，它在系统标定过程中相对复杂，并且面临立体视觉系统的常见问题，例如点匹配，更复杂的图像处理和数据处理。

显然，通过三维眼球结构和眼睛成像模型在相机坐标系中估计实际三维视线方向的三维视线追踪器，通常由多个相机和多个光源组成。三维视线估计只能基于单相机多光源系统或多相机多光源系统来实现。但是，一些现代几何方法仍尝试使用单相机单光源系统估计三维视线，由于其简单性，该方法依赖于人类平均值或某些眼睛固有参数的经验值^[46,51]。Ohno 和 Mukawa^[51]使用角膜曲率、瞳孔中心和角膜中心之间距离的人类平均值以及恒定的折射率来估计光轴。Guestrin 和 Eizenmen^[46]证明，如果已知眼睛和计算机屏幕之间的距离或头部固定，则单相机单光源系统可以确定视线落点的估算值。由于不同用户之间的某些眼睛差异以及头部运动的影响会降低单相机单光源系统中视线估计的准确性，因此这种方法不是真正的三维视线估计方法。总之，尽管单相机单光源系统具有相对简单的系统配置，但其视线落点估算的准确性会受到复杂的用户标定或使用总体平均值作为特定对象参数的影响，并且不允许自由头部运动。随着光源或相机数量的增加，会提高自由头部运动下视线追踪系统的精度，但代价是复杂的系统构造和系统标定。

基于摄像机的视线追踪技术具有对用户干扰性小，精度相对理想且便于

产品化的特点。本文采用多相机多光源系统来进行视线估计以及落点的计算。采用多相机多光源进行视线估计方法具有下列优点：在标定过程中只需确定视轴与光轴的夹角，不需要额外标定以获取用户眼部特征，减少了标定点的数量，减少了用户标定的时间成本；因为采用立体视觉，可以检测使用者三维眼睛特征，进而获取三维视线方向；对于用户的头部运动无要求，可以使用户有一个更舒适自然的使用体验。多相机多光源系统有如下缺点：一般需要两台相机配两个以上光源的复杂的硬件配置，才能检测到眼球光学模型所需的特征参数，不但增加了系统的硬件成本还使得系统笨重不易使用；需要对相机、多相机系统和视线追踪系统（标定出光源、相机和屏幕之间的位置关系）进行标定，加大了系统标定的难度，高精度的系统标定具有挑战性；多相机使视野变窄，限制了用户头部的活动范围。

1.4 论文结构及内容

本文共分为六章，具体如下：

- 1) 第一章 引言。本章介绍了本文的课题背景，研究目的及意义，视线追踪技术概述以及论文整体结构安排。
- 2) 第二章 问题描述。本章介绍了传统三维视线估计方法，并对传统方法模型误差进行了分析，由此提出本文的方法。
- 3) 第三章 三维视线追踪系统标定。本章介绍了相机标定方法，系统标定方法，以及用户标定方法。在系统标定方法中，本文采用单相机辅助标定法对光源和屏幕位置进行标定。
- 4) 第四章 三维视线估计模型。本章详细介绍了本文提出的基于瞳孔边缘匹配的多相机多光源系统三维视线估计方法。
- 5) 第五章 三维视线估计实验分析。本章介绍了仿真实验、真实实验设备、真实实验过程，并分析了实验结果。
- 6) 第六章 结论。本章总结了本文。

2 问题描述

2.1 传统三维视线估计方法概述

从以上分析可以知道，基于多相机多光源系统的传统三维视线估计可以分为两种。一种方法是通过检测角膜曲率中心和瞳孔中心来构建眼球光轴^[52-53]，这种方法在求取瞳孔中心的过程中，需要考虑角膜折射的影响。当瞳孔的位置和姿态固定时，畸变误差随着瞳孔离相机的距离增加而减小，瞳孔越偏离相机的光轴，畸变误差越大；瞳孔平面与像平面之间的夹角越大，畸变误差越大。由于光在角膜中的折射指数与在空气中的折射指数不同，因此当光从空气进入角膜时，光的方向将发生变化。在不考虑光折射变化的情况下进行瞳孔中心检测会导致视线方向估计误差超过 2° 。另一种方法是“折射平面”法，根据入射光和折射光之间的几何关系，来构造包含角膜中心和瞳孔中心的折射平面，不同相机建立的折射平面的交线即为眼球的光轴^[40,54-55]。如图 2-1 所示，在“折射平面”法中，假设已知角膜曲率中心 C ，由折射定律可知，入射光线、折射光线、法线在同一平面上，由于角膜曲率中心 C 在角膜外表面的法线上，假设瞳孔中心发出光线，瞳孔中心 E 在入射光线上，相机光心 O_1 和瞳孔像的中心点 E_{img} 在折射光线上，因此，点 C 、 E 、 O 、 E_{img} 共面，这个“折射平面”可以表示为 $C \times E_{img}$ ，包含光轴 EC 。用两个相机分别对人眼进行拍照，可以构建两个包含人眼光轴的折射平面，这两个平面的交线即为光轴 EC 。

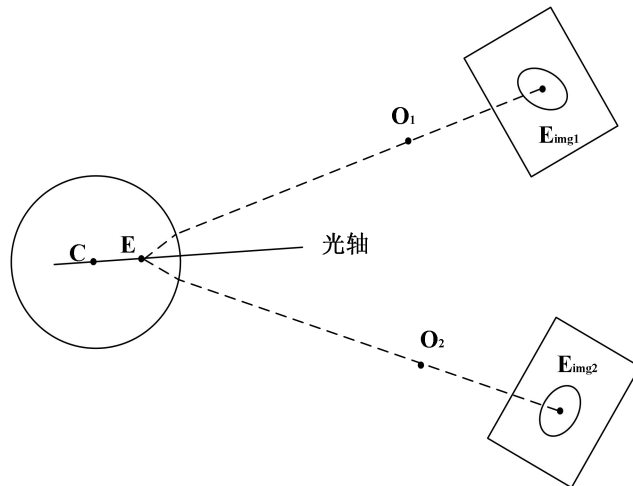


图 2-1 传统三维视线估计方法示意图

2.2 传统方法模型误差分析及问题提出

由上述分析可见，通过“折射平面”法重建眼球光轴的方法，基于如下假设：即图像上的瞳孔中心就是三维空间中实际瞳孔中心在像平面投影的像。基于这个假设，则通过两个相机图像中的瞳孔中心，分别与相机镜头的光学中心和角膜曲率中心构成的两个折射平面的交线就是眼球光轴。重建光轴后，再通过用户标定得到的 $Kappa$ 角，就可以重构眼球的视轴。但根据机器视觉的基本原理，空间的瞳孔中心在相机中像平面的投影，并不是图像上瞳孔的中心^[54,56]。也就是说，空间的圆形目标在相机中成椭圆像，而空间圆的中心的像并不是图像中椭圆的中心，以下给出详细的证明。

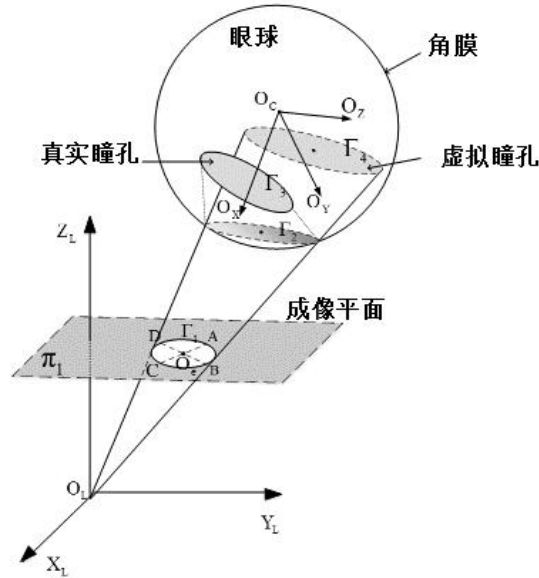


图 2-2 瞳孔成像示意图

如图 2-2 所示，以相机镜头的光学中心 O_s 为原点，相机镜头的光轴为 Z 轴建立相机坐标系 $O_s - X_s Y_s Z_s$ ， π 为相机的像平面。人眼的瞳孔位于由角膜和房水组成的凸透镜内部。 Γ_0 是实际的瞳孔， p_0 为其圆心， C 是角膜中心， U 是角膜曲面，其形状假定为假想球体的一部分。空间瞳孔的边缘发出的光线与角膜外表面的交点仍然在一个平面上，这些交点构成了一个空间的圆 Γ_2 。同时在相机的成像面上形成瞳孔的椭圆像 Γ_1 。因此，从相机观察到的是虚拟瞳孔，而不是瞳孔本身。也就是说，瞳孔在相机中所成的椭圆像为瞳孔的虚像，是由空间的虚拟瞳孔 Γ_3 所成的像。辅助眼球三维坐标系定义如下：原点由角膜中心 C 定义， X 轴定义为与 $O_c O_s$ 平行。定义 Y 轴，使得 X - Y 平面包含眼球的光轴 EC 。假设瞳孔法线和 X 轴之间的角度是 φ ，瞳孔的半径

为 r_p ，角膜中心与瞳孔中心的距离为 d 。瞳孔圆形边缘的三维坐标由下式给出： $(d \cos \varphi - r_p \cos \theta \sin \varphi \quad d \sin \varphi + r_p \cos \theta \cos \varphi \quad r_p \sin \theta)$ ，其中 $\theta \in [0, 2\pi]$ 。瞳孔中心的三维坐标由下式给出：

$$(p_{0x} \quad p_{0y} \quad p_{0z}) = (d \cos \varphi \quad d \sin \varphi \quad 0) \quad (2-1)$$

空间圆 Γ_2 的派生坐标由下式给出： $[\rho - \frac{(\rho - p_x)}{S \cdot l + n} \quad m \cdot p_{0y} \quad m \cdot p_{0z}]$ 。其中 $l = \rho - p_{0x}$ ， $S = |(n' - n) / \rho|$ ， $m = n / (R \cdot (\rho - p_{0x}) + n)$ 。空间圆 Γ_2 的中心可由下式给出：

$$\mathbf{p}_2^{o_c} = [\rho - \frac{n'(\rho - d \cos \varphi)}{n + \frac{(n - n')(\rho - d \cos \varphi)}{\rho}} \quad \frac{dn \sin \varphi}{n + \frac{(n - n')(\rho - d \cos \varphi)}{\rho}} \quad 0] \quad (2-2)$$

其中 n 是角膜的折射率， $n' \approx 1$ 是空气的折射率。因此，在坐标系 $O_c - X_c Y_c Z_c$ 下，瞳孔 Γ_0 的中心坐标与空间圆 Γ_2 的中心坐标之间的关系为：

$$\mathbf{p}_2^{o_c} = \left[\rho - \frac{\rho(\rho - p_{0x})}{\rho n + (n - 1)(\rho - p_{0x})} \quad \frac{\rho n p_{0y}}{\rho n + (n - 1)(\rho - p_{0x})} \quad 0 \right] \quad (2-3)$$

这样由瞳孔中心发出的光线经角膜折射后，在相机中形成像点 \mathbf{p}'_1 ，则 \mathbf{p}'_1 点也是空间圆 Γ_2 的圆心的像点。假设眼球坐标系和相机坐标系之间的转换矩阵为 \mathbf{M} 。则空间圆 Γ_2 的圆心在相机坐标系下的坐标可以表示为：

$$\mathbf{p}_2^{o_s} = \mathbf{M} \cdot \mathbf{p}_2^{o_c} \quad (2-4)$$

Γ_1 为真实瞳孔经角膜折射后在相机中所成的椭圆像，也是虚拟瞳孔 Γ_3 的像。在坐标系 $O_s - X_s Y_s Z_s$ 下，由相机镜头的光学中心和瞳孔的虚像构成了空间的一个斜的椭圆锥 $O_s - ABCD$ 可表示为：

$$Ax^2 + By^2 + Cxy + Dxz + Eyz + Fz^2 = 0 \quad (2-5)$$

其与球形角膜外表面的交线就是空间圆 Γ_2 ， Γ_2 和虚拟瞳孔 Γ_3 都是由空间的一个平面截取斜椭圆锥 $O_s - ABCD$ 所形成的圆，它们在相机中的像都是 Γ_1 。虚拟瞳孔 Γ_3 和空间圆 Γ_2 的法向量都为：

$$\left[e_{1x} \pm e_{3x}, e_{1y} \pm e_{3y}, e_{1z} \pm e_{3z} \right] \cdot \sqrt{\frac{|\lambda_2| + |\lambda_3|}{|\lambda_1| + |\lambda_3|}} \quad (2-6)$$

其中， $\lambda_1, \lambda_2, \lambda_3$ 为二次型 $Q = [A \quad C/2 \quad D/2; C/2 \quad B \quad E/2; D/2 \quad E/2 \quad F]$ 的特征值。其对应的特征向量分别为 $e_1 = [e_{1x} \quad e_{1y} \quad e_{1z}]^T$ 、

$e_2 = [e_{2x} \ e_{2y} \ e_{2z}]^T$ 、 $e_3 = [e_{3x} \ e_{3y} \ e_{3z}]^T$ 。因此，空间圆 Γ_2 和虚拟瞳孔 Γ_3 的法向量相同，则这两个圆所在的平面是平行的。设虚拟瞳孔 Γ_3 的半径也为 r_p ，则 Γ_3 的圆心坐标为：

$$[x_{03}, y_{03}, z_{03}] = [\pm r_p \sqrt{\frac{|\lambda_3|(|\lambda_1| - |\lambda_2|)}{|\lambda_1|(|\lambda_1| + |\lambda_3|)}}, 0, r_p \sqrt{\frac{|\lambda_1|(|\lambda_2| - |\lambda_3|)}{|\lambda_3|(|\lambda_1| + |\lambda_3|)}}] \quad (2-7)$$

连接 O_s 和虚拟瞳孔 Γ_3 的圆心 p_3 的直线方程为：

$$\frac{x}{x_{01}} = \frac{y}{y_{01}} = \frac{z}{z_{01}} \quad (2-8)$$

设空间圆 Γ_2 的半径为 r_T ，将式(2-7)中的 r_p 替换为 r_T ，可得 Γ_2 的圆心 p_2 坐标 $[x_{02}, y_{02}, z_{02}]$ 。将 $[x_{02}, y_{02}, z_{02}]$ 的值带入直线方程(2-8)，则可以证明 Γ_2 的圆心 p_2 位于直线 $O_s p_3$ 上，说明斜椭圆锥与角膜外表面相交形成的空间圆 Γ_2 的圆心与虚拟瞳孔 Γ_3 的圆心在一条直线上。

由上述分析，我们可以得出如下结论。如图 2-1 所示，真实的瞳孔 Γ_0 发出的光线，经角膜折射后在相机的成像面上成椭圆像 Γ_1 。角膜为近似的球形，瞳孔发出的光线与角膜外表面的交点构成一个空间圆 Γ_2 ，由于角膜折射的影响，椭圆像 Γ_1 可以看作是虚拟瞳孔 Γ_3 所成的像。空间圆 Γ_2 所在平面与虚拟瞳孔所在的平面平行，虚拟瞳孔 Γ_3 的圆心和空间圆 Γ_2 的圆心有同一个像点，这个像点也是真实瞳孔中心的像点。

在 $O_s - X_s Y_s Z_s$ 坐标系下，可推导出斜圆锥 $O_s - ABCD$ 的方程为：

$$(nx + py + qz)^2 + (kx + ly + mz)^2 = (rx + sy + tz)^2 \quad (2-9)$$

其中 $k = \sin \omega(r_1 \cos \gamma - r_7 \cos \alpha) - \cos \omega(r_4 \cos \gamma - r_7 \cos \beta)$ ，

$$r = r_7 \tan \nu \quad s = r_8 \tan \nu, \quad t = (r_9 \tan \nu)f,$$

$$l = \sin \omega(r_2 \cos \gamma - r_8 \cos \alpha) - \cos \omega(r_5 \cos \gamma - r_8 \cos \beta),$$

$$m = [\sin \omega(r_3 \cos \gamma - r_9 \cos \alpha) - \cos \omega(r_6 \cos \gamma - r_9 \cos \beta)]f,$$

$$n = \cos \omega(r_1 \cos \gamma - r_7 \cos \alpha) + \sin \omega(r_4 \cos \gamma - r_7 \cos \beta),$$

$$p = \cos \omega(r_2 \cos \gamma - r_8 \cos \alpha) + \sin \omega(r_5 \cos \gamma - r_8 \cos \beta),$$

$$q = [\cos \omega(r_3 \cos \gamma - r_9 \cos \alpha) + \sin \omega(r_6 \cos \gamma - r_9 \cos \beta)]f,$$

α, β, γ 分别为在辅助坐标系 $O_s - X'Y'Z'$ 下（在此坐标系下， $O_s - ABCD$ 为标准椭圆锥），坐标系的原点 O_s 和虚拟瞳孔中心 p_3 的连线 $O_s p_3$ 的方向角。

$\mathbf{N}=[r_1 \ r_2 \ r_3; \ r_4 \ r_5 \ r_6; \ r_7 \ r_8 \ r_9]$ 为坐标系 $O_s-X_sY_sZ_s$ 到 $O_s-X'Y'Z'$ 之间的转换矩阵, ω, ν 为椭圆像 Γ_1 在 $O_s-X_sY_sZ_s$ 坐标系下的角度参数。则得到在坐标系 $O_s-X_sY_sZ_s$ 下, 虚拟瞳孔在像平面 π 上的透视投影曲线 Γ_1 的方程为:

$$\begin{aligned} & (k^2 + n^2 - r^2)x^2 + 2(kl + np - rs)xy + (l^2 + p^2 - s^2)y^2 \\ & + 2(mk + nq - rt)x + 2(mq + pq - st)y + m^2 + q^2 - t^2 = 0 \end{aligned} \quad (2-10)$$

式 (2-10) 为一个二次的闭合的曲线表达式, 为椭圆或者圆。因此, 虚拟瞳孔在像平面上的椭圆曲线 Γ_1 的中心 p_1 的坐标 (a, b) 由下式给出:

$$\begin{cases} a = \frac{(l^2 + p^2 - s^2)(mk + nq - rt) - (kl + np - rs)(ml + pq - st)}{(kl + np - rs)^2 - (k^2 + n^2 - r^2)(l^2 + p^2 - s^2)} \\ b = \frac{(k^2 + n^2 - r^2)(ml + pq - st) - (kl + np - rs)(mk + nq - rt)}{(kl + np - rs)^2 - (k^2 + n^2 - r^2)(l^2 + p^2 - s^2)} \end{cases} \quad (2-11)$$

虚拟瞳孔的中心 p_3 在像平面 π 上的实际像点的坐标, 可以利用直线 $O_s p_3$ 与像平面 π 的交点来求取。直线 $O_s p_3$ 在坐标系 $O_s-X_sY_sZ_s$ 下的方程为:

$$\frac{(r_1 x + r_2 y + r_3 z) \cos \gamma}{\cos \alpha} = \frac{(r_4 x + r_5 y + r_6 z) \cos \gamma}{\cos \beta} = r_7 x + r_8 y + r_9 z \quad (2-12)$$

将 $z = f$ 带入式 (2-12) 可求得 p_3 点的实际像点 p'_1 的坐标 (a', b') 的表达式为:

$$[a' \ b'] = \begin{bmatrix} \frac{q_1 l_1 - m_1 p_1}{k_1 p_1 - n_1 l_1} & \frac{m_1 n_1 - k_1 q_1}{k_1 p_1 - n_1 l_1} \end{bmatrix} \quad (2-13)$$

其中: $k_1 = r_1 \cos \gamma - r_7 \cos \alpha$, $l_1 = r_2 \cos \gamma - r_8 \cos \alpha$, $m_1 = (r_3 \cos \gamma - r_9 \cos \alpha)f$, $n_1 = r_4 \cos \gamma - r_7 \cos \beta$, $p_1 = r_5 \cos \gamma - r_8 \cos \beta$, $q_1 = (r_6 \cos \gamma - r_9 \cos \beta)f$ 。式 (2-11) 和 (2-13) 只有在虚拟瞳孔所在的平面与像平面 π 平行时才相等, 当二者之间存在一定夹角的时候, 它们并不相等。用这两个点之间的距离 d 来描述它们之间偏差值, 即透视投影变换中椭圆中心的畸变误差 (瞳孔的椭圆像中心与空间瞳孔 Γ_3 的中心的像点之间的误差) 为:

$$d = \sqrt{(a - a')^2 + (b - b')^2} \quad (2-14)$$

本文对式 (2-14) 进行了仿真计算, 椭圆中心畸变误差 d 有如下变化规律: (1) 当瞳孔的位置和姿态固定, 畸变误差 d 随着瞳孔离相机的距离增加而减小; (2) 当瞳孔的姿态固定, 瞳孔越偏离相机的光轴, 畸变误差 d 越大; (3) 当瞳孔平面与像平面之间的夹角越大, 则畸变误差越大。因此, 目前普遍采用的多相机多光源系统的眼球光轴重建方法存在原理上的误差, 或者说

存在着固有的光轴重建误差，特别是当检测用户大视角的视线方向，即瞳孔的朝向较大地偏离相机光轴时，光轴重建的误差会更大。

2.3 本文三维视线估计方法概述

由上可知，空间瞳孔在相机中成椭圆像，而空间瞳孔中心的像并不是图像中椭圆的中心，并且给出了详细的证明。因此，针对传统三维视线估计方法的缺陷，本文提出了一种基于瞳孔边缘匹配的三维视线估计方法。在双相机双光源系统中，找到两个相机成像面上多对瞳孔边缘的匹配点，通过多对匹配点信息及其它信息数据进行光轴估计，并且提出了三种重建光轴的方法。在本文提出的基于瞳孔边缘匹配的方法中，不需要检测瞳孔中心，不必考虑角膜折射带来的影响，因此避免了由瞳孔中心光的折射引起的误差。本文在第五章进行了真实实验，验证了本文方法的有效性。

2.4 本章小结

本章对传统三维视线估计方法进行了介绍，提出了传统方法的缺陷并给予了详细的证明，进而引出本文对于传统方法缺陷进行的改进，即本文提出了一种基于瞳孔边缘匹配的三维视线估计方法，能避免传统方法带来的误差，提高视线估计的精度。

3 三维视线追踪系统标定

本文硬件系统主要由双摄像机、计算机、光源三部分组成。两个摄像机主要用于同步采集用户脸部图像，然后将采集到的图像输入计算机进行图像处理并提取所需数据。光源使用了波长为 880nm 的近红外发光二极管，在视线追踪系统中，光源都被当做点光源，主要用途是用来计算角膜曲率中心的位置。

从算法的角度讲，视线追踪技术应用于人机交互，必须至少有一个已知位置的光源和一个已知参数的相机。相机的参数（包括焦距、主点、畸变等）需要通过标定得到，光源相对于相机的位置也需要进行标定；在求解出视线方向之后，要想完成交互，还要求解视线在屏幕上的落点，即求屏幕和视线方向的交点。而且对于二维或三维的视线追踪系统而言，光轴和视轴之间的 Kappa 角，在使用之前也需要标定，这个标定需要借助在屏幕上出现的标定点，因此，屏幕的位置也需要通过标定得到。综上，本文的标定主要分为三部分：

(1) 在进行视线估计前，需要获得一系列的参数，首先是双摄像机的内外参数，一般通过标定的方法获得参数。先进行内部标定获得摄像机的内部参数，再进行外部标定确定系统结构参数 \mathbf{R} 和 \mathbf{T} 。通常由标定靶标对相机进行标定。

(2) 标定系统中光源的位置，以便求取角膜曲率中心；标定屏幕的位置，求屏幕所在平面的方程，以便求取视线在屏幕上的落点。

(3) 标定用户的 Kappa 角，从而实现由光轴到视轴的转换。

3.1 相机标定方法

某个物体被相机成像时，其表面可成像点的物理坐标和该点在图像中的像素坐标之间有一定的关系，相机标定就是根据相机成像系统的数学模型建立这种对应关系。在标定结果中，这种对应关系用相机的内外参数表示。其中内参数主要包括焦距、主点坐标、畸变率等，外参数主要是与世界坐标系之间的转换关系^[57]。相机的标定在视线追踪系统的标定中不可或缺，相机参数标定准确与否直接关系着图像处理的精度，因此在整个系统的标定中，相机的标定是最为基础的也是很重要的。

标定方法根据原理和实现方式的不同分为多种，就目前使用情况而言，

Tsai 提出的“两步法”及其改进技术依然是采用最多的标定方法，尤其是以张正友法^[58]为代表的标定方法，因为其简单易于操作的特性，应用在很多的标定场合。

3.1.1 张正友标定法

理想情况下，数字相机成像遵循小孔成像模型，实际中，各种因素对成像的影响可以归结为：相机镜头的径向畸变和偏心畸变，以及像平面内比例及正交畸变^[59]。图像径向畸变是图像像素点以畸变中心为中心点，沿着径向产生的位置偏差，这个畸变中心通常为图像的主点，具体的，将此项写到 X 与 Y 轴：

$$\begin{cases} \Delta x_r = K_1 \bar{x} r^2 + K_2 \bar{x} r^4 + K_3 \bar{x} r^6 + \dots \\ \Delta y_r = K_1 \bar{y} r^2 + K_2 \bar{y} r^4 + K_3 \bar{y} r^6 + \dots \end{cases} \quad (3-1)$$

上式中 Δ 表示偏差量， \bar{x} 、 \bar{y} 表示图像原始坐标，一般的，常取 2、4、6 阶量为径向畸变系数。偏心畸变是因为相机的透镜组中心偏离相机主光轴而产生，分解到 X 坐标轴和 Y 坐标轴之后如下：

$$\begin{cases} \Delta x_d = P_1(r^2 + 2\bar{x}^2) + 2P_2\bar{x}\bar{y} \\ \Delta y_d = P_2(r^2 + 2\bar{y}^2) + 2P_1\bar{x}\bar{y} \end{cases} \quad (3-2)$$

像素长宽比例因子与像平面轴不正交产生畸变为：

$$\begin{cases} \Delta x_m = b_1 \bar{x} + b_2 \bar{y} \\ \Delta y_m = 0 \end{cases} \quad (3-3)$$

相机主点指相机光轴和成像平面的交点，将主点和像平面中点的误差作为主点误差，记为 x_0 ， y_0 ，焦距误差记为 Δf ，对像素的影响为：

$$\begin{cases} \Delta x_m = -x_0 - \frac{\bar{x}}{f} \Delta f \\ \Delta y_m = -y_0 - \frac{\bar{y}}{f} \Delta f \end{cases} \quad (3-4)$$

总体偏差

$$\begin{cases} \Delta x' = \Delta x_r + \Delta x_d + \Delta x_m + \Delta x_n \\ \Delta y' = \Delta y_r + \Delta y_d + \Delta y_m + \Delta y_n \end{cases} \quad (3-5)$$

上述的 10 个参数 x_0 、 y_0 、 Δf 、 K_1 、 K_2 、 K_3 、 P_1 、 P_2 、 b_1 、 b_2 可以构成相机的十参数模型^[60-61]。根据此模型，可以对相机拍摄到的图片进行矫正，在本文中，由于从图像中提取到的都是点的信息或轮廓点信息，若对图像进行整体矫正，会在插值过程中丢失边缘点的信息，从而造成误差，因此本文中使用对点矫正而不是对图像矫正的方法。

3.1.2 试验场标定法

我们采用试验场标定法标定相机的内参数，具体做法为：一些空间坐标已知的点构成试验场，待标定的相机对其进行成像，根据单相片空间后方交会或者多相片后方交会原理，可以求解出相机的内外参数^[59]。

两个相机坐标系之间的转换关系标定时，由于相机外参数表示的是相机坐标系和标定场坐标系的关系，设两个相机坐标系分别为 S 和 G，标定场坐标系为 K，则 S 和 G 相机的外参为：

$$\begin{cases} \mathbf{S}^T = \mathbf{R}_{sk} \mathbf{K} + \mathbf{T}_{sk}^T \\ \mathbf{G}^T = \mathbf{R}_{gk} \mathbf{K} + \mathbf{T}_{gk}^T \end{cases} \quad (3-6)$$

\mathbf{R} 代表旋转矩阵， \mathbf{T} 代表平移矩阵，消去上式中的 K，可得到：

$$\mathbf{S}^T = \mathbf{R}_{sk} (\mathbf{R}_{gk}^{-1} (\mathbf{G}^T - \mathbf{T}_{gk}^T)) + \mathbf{T}_{sk}^T \quad (3-7)$$

变换之后有：

$$\begin{cases} \mathbf{R}_{sg} = \mathbf{R}_{sk} \mathbf{R}_{gk}^{-1} \\ \mathbf{T}_{sg}^T = -\mathbf{R}_{sk} \mathbf{R}_{gk}^{-1} \mathbf{T}_{gk}^T + \mathbf{T}_{sk}^T \end{cases} \quad (3-8)$$

因此由两相机同时拍摄标定场，求解出外参之后通过上式进行转换，可以得到两个相机坐标系之间的转换矩阵。

本文实验所用相机的标定工作由北京普达迪泰公司完成，采用试验场标定法，对于单相机，标定得到 10 个参数，对于双相机，标定出两相机之间按照三个轴旋转的旋转角和平移矩阵，标定达到的精度见表 3-1。

表 3-1 相机标定精度

单相机标定误差					双相机标定误差	
焦距 (mm)	主点 (像素)	径向畸 变系数	偏心畸 变系数	比例畸 变系数	旋转角误差 (°)	平移矩阵误差 (mm)
0.001	0.25	1e-6	1e-6	1e-6	0.005	0.01

3.2 系统标定方法

3.2.1 系统标定方法概述

对于视线追踪系统，系统标定是基础，并且有着极其重要的意义。首先，对于视线追踪成品，光源屏幕的相对位置可以固化，但固化之前必须进行标

定；其次，对于三维视线追踪系统的研究，必须在构建系统的过程中，通过标定程序来精确定位屏幕和光源在系统相机坐标系中的三维位置。在现有研究中，三维眼球视线模型是在光源和屏幕位置已知的情况下建立的，光源用来估计眼球参数中角膜曲率中心和角膜半径，屏幕用来估计视线的落点。在大部分三维视线模型中，只有知道光源和屏幕的精确位置后才可以准确地求解出眼球视线方向和视线落点，当系统中的光源位置不够精确时，求解出的角膜曲率中心和半径会产生一个较大的偏差，而当屏幕位置不精确时，视线落点的计算也会产生一个较大的偏差。因此，一个三维视线模型的可行性，关键在于能否精确标定光源和屏幕在系统相机坐标系下的位置。

在三维视线追踪系统中，为了估计三维视线方向，需要确定光源和屏幕平面在系统相机坐标系中的位置，然而存在一个问题，即系统相机往往位于屏幕的前方，光源一般分布在摄像机的侧面，使屏幕和光源不在摄像机的视野内，因此系统的各部分的相对位置不能由系统本身的视觉系统来标定。这一矛盾的解决目前有以下几种方法：人工测量法，平面镜反射法，单相机辅助标定法等^[62-63]。

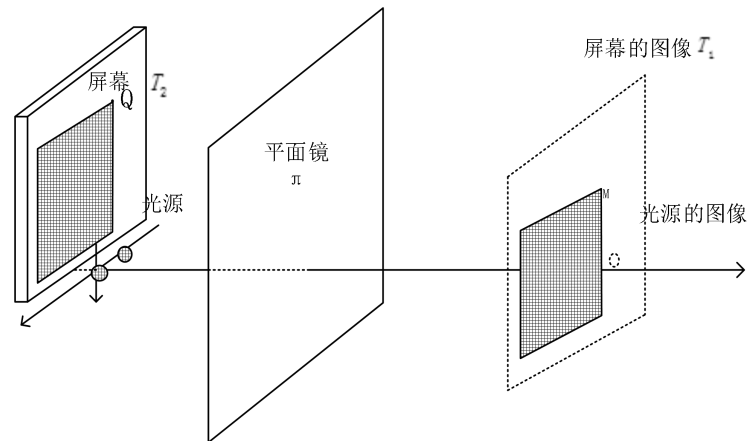


图 3-1 反射镜标定方法示意图

人工测量法，顾名思义，即设定系统相机为坐标原点，之后手工测量光源和屏幕的三维位置。平面镜反射法，如图 3-1 所示，即将一个平面镜置于系统相机成像面之前，在镜子表面贴一个平面标定模板，另一标定模板附在屏幕上，要求左右相机图像中均有镜子、光源及完整的屏幕，之后用系统相机拍摄图像，这样经过摄像机标定后便可求得屏幕和光源在镜子中像的三维坐标，通过镜面空间点拟合得到镜子平面方程，再根据离差得到光源及屏幕的像的关于镜子平面的对称点，即光源的真实三维坐标和屏幕平面方程^[64]。

3.2.2 单相机辅助标定法

单相机辅助标定方法借助辅助标定物，对光源和屏幕进行标定。本文借助一个半径已知的标准球对光源和屏幕进行标定。标定系统由系统相机、待标定的光源、标定球组成，如下图 3-2 所示：

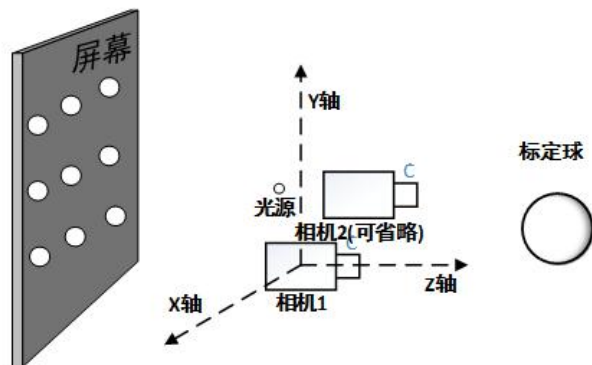


图 3-2 相机辅助标定示意图



(a) 位置 1

(b) 位置 2

(c) 位置 3

图 3-3 标定球标定实验图

图中可以看到，参数已知的标定球放置在相机对面，对于单相机，只需借助一个标定球，通过移动球的位置得到多个反射平面，从而进行标定，如图 3-3 所示。首先，光源发出的光在标定球表面产生反射，相机拍摄标定球的图像。移动标定球的位置，小球在不同的位置下成像，从而标定得到光源的位置。同理，为标定屏幕，可以在屏幕上显示几个高亮的点，其在标定球表面的反射点同样被相机所成像，移动标定球，使其多次成像，最后解算得到屏幕上点的位置，进而拟合出屏幕平面方程。原理图如图 3-4 所示：

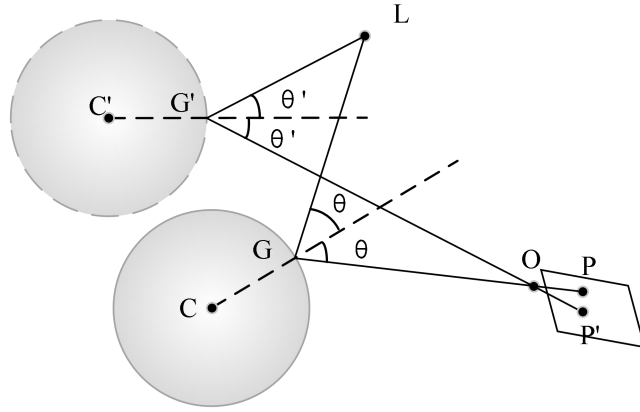


图 3-4 单相机辅助标定示意图

上图中L为待标定的点（如光源或屏幕上的点），C为标定球的中心，O为相机光心，P为标定点在球面上的反射点在相机成像面上的像点，G为待标定点在球面上的实际反射点。

根据光学知识，我们知道，L,G,C,O,P五个点一定共面，根据反射定律，可以求得直线LG的方向，对于标定球的一个位置，将会得到一条直线LG，在直线LG上的任何一点都满足标定点的条件，即根据标定球的一个位置，无法得到标定点L的三维坐标，只能确定L所在的直线。为解决这一矛盾，我们引入了第二个条件，即移动标定球的位置，得到另外一组点，即L,G',C',O,P'，两个标定球位置下，可以得到两条过L点的直线，两条直线的交点即为真正的L点的坐标，其求解过程如下。

根据共面向量的性质，可以写出方程：

$$\begin{cases} (\mathbf{L}-\mathbf{C}) \times (\mathbf{O}-\mathbf{C}) \times (\mathbf{L}-\mathbf{P}) = 0 \\ (\mathbf{L}-\mathbf{C}') \times (\mathbf{O}-\mathbf{C}') \times (\mathbf{L}-\mathbf{P}') = 0 \end{cases} \quad (3-9)$$

入射角等于反射角：

$$\begin{cases} \frac{(\mathbf{P}-\mathbf{G}) \cdot (\mathbf{C}-\mathbf{G})}{|\mathbf{P}-\mathbf{G}| \times |\mathbf{C}-\mathbf{G}|} - \frac{(\mathbf{L}-\mathbf{G}) \cdot (\mathbf{C}-\mathbf{G})}{|\mathbf{L}-\mathbf{G}| \times |\mathbf{C}-\mathbf{G}|} = 0 \\ \frac{(\mathbf{P}'-\mathbf{G}') \cdot (\mathbf{C}'-\mathbf{G}')}{|\mathbf{P}'-\mathbf{G}'| \times |\mathbf{C}'-\mathbf{G}'|} - \frac{(\mathbf{L}-\mathbf{G}') \cdot (\mathbf{C}'-\mathbf{G}')}{|\mathbf{L}-\mathbf{G}'| \times |\mathbf{C}'-\mathbf{G}'|} = 0 \end{cases} \quad (3-10)$$

G,G'点是光线反射点，以G为例，它是直线OP和球C表面的交点，通过求解直线和球面的交点可得到G和G'点的位置，设G点位置为 (G_x, G_y, G_z) ，O点为参考坐标系的原点，坐标为 $(0, 0, 0)$ ，P点的坐标假设为 $P(P_x, P_y, P_z)$ ，则直线OP的方程为：

$$\frac{x}{P_x} = \frac{y}{P_y} = \frac{z}{P_z} \quad (3-11)$$

根据文献^[65]中球的中心求解的方法，可以算出标定球的中心 $C(C_x, C_y, C_z)$ ，标定球的半径已知为 R ，则其空间方程为：

$$(x-C_x)^2 + (y-C_y)^2 + (z-C_z)^2 = R^2 \quad (3-12)$$

联立方程 (3-11) 和 (3-12)，即可解得交点 G 。由于直线穿过球时，会两次和球面相交，因此 G 一般会有两个解，当 $x_1 \neq x_2$ 时，可得到交点坐标方程：

$$\begin{cases} y = \frac{P_y}{P_x} x \\ z = \frac{P_z}{P_x} x \\ (x-C_x)^2 + (y-C_y)^2 + (z-C_z)^2 = R^2 \end{cases} \quad (3-13)$$

如下图所示，直线 OP 和球 C 交于两点： G 和 G' ，求解得到两个坐标。

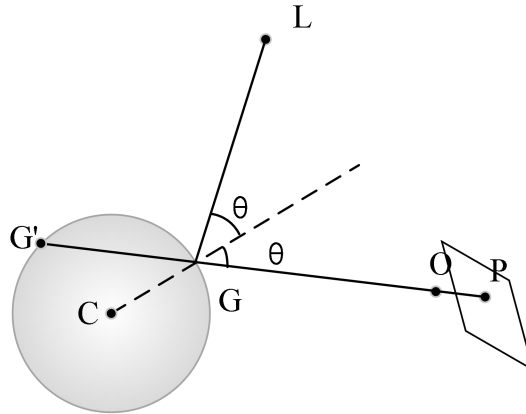


图 3-5 G 点多解示意

根据上图 3-5 可知， G 点是真正的反射点，而 G' 是需要排除的解，可以根据反射定律和空间位置关系进行排除，对于 G 点，有：

$$\angle LGC = \angle OGC \quad (3-14)$$

而对于 G' ，一般而言不存在这样的关系，因此可以通过这一条件对 G' 进行排除。特殊的，若 $\angle LGC = \angle OGC$ 的同时 $\angle LG'C = \angle OG'C$ 也成立，则可以根据反射点的空间位置关系确定唯一解：由于世界坐标系即以 O 为原点的相机坐标系，因此在空间上， G 点离 O 点的距离一定要小于 G' 点离 O 点的距离，通过这一约束，可以唯一的确定 G 点的位置。

得到反射点的坐标之后，设直线 LG 和 LG' 的方程为：

$$\begin{cases} \frac{x-G_x}{a} = \frac{y-G_y}{b} = \frac{z-G_z}{1} \\ \frac{x-G'_x}{a'} = \frac{y-G'_y}{b'} = \frac{z-G'_z}{1} \end{cases} \quad (3-15)$$

其中 LG 和 LG' 的方向向量分别为 $(a,b,1)$ 、 $(a',b',1)$ ，联立公式 (3-13)，可解得两直线的方程，得到直线方程之后，通过公式 (3-15) 即可得到 L 点的坐标。

在实际标定光源的过程中，需要提前获取标定球球心的空间三维坐标。球心的求解由于边缘误差、图像畸变等的影响会存在一定的误差，而在上述求解的过程中，需要标定球球心 C 和 C' 具有精确的三维位置，由于约束条件中包含标定球半径，因此当球心不准会导致其光源反射点位置出现偏差，进而导致入射光线不相交于真实的 L 点，而是得到一个错误的交点 L_{err} ，甚至不会交于一点，这样求得的 L 点坐标是没有意义的，如下图 3-6 所示。

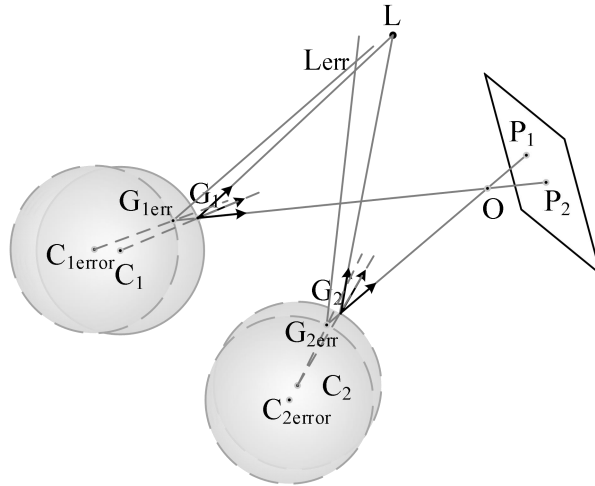


图 3-6 L 点误差示意图

产生这个问题的本质在于两个球位置产生的约束不够，引入更多的约束可以有效的避免得到虚假的标定点。而上述的问题产生的根本因素在于确定 LG 的两个平面在同一平面内，因此，增加更多球的位置（至少三个位置），保证至少有两个 O,P,C,L 构成的平面不在同一平面内，最终针对多个标定球位置求解入射光线的方程，消除两个求解平面位于同一平面带来的误差。在得到经过光源点的多条直线后，再利用最小二乘的方式求解多条直线的交点，从而提升精度。

下图 3-7 为实验光环设备，具体介绍见第 5.2 节。实验采用了如下五个

光源，为方便后续的图像处理，其余光源被遮盖。求解得到的光源坐标见表 3-2。同样的方法可以标定屏幕上各点的位置，如图 3-8 所示，得到屏幕四个角点的坐标见表 3-3。

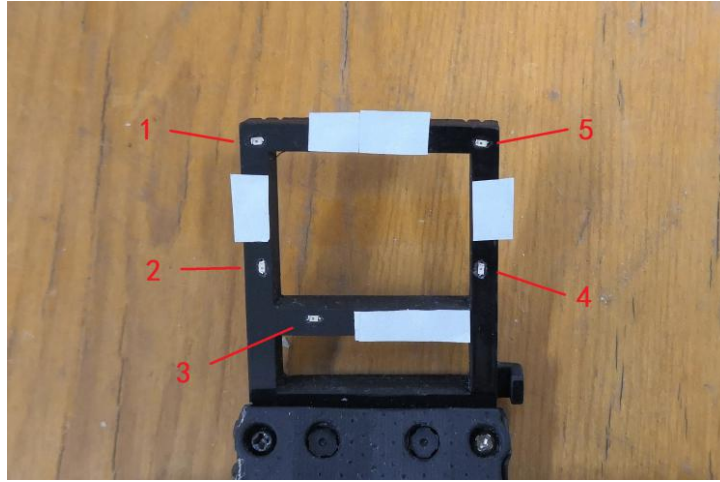


图 3-7 实验所用光环设备

表 3-2 标定得到的光源坐标

光源坐标（以左相机建立坐标系）/(mm)	
光源 1	(-11.6099,37.0562,20.6900)
光源 2	(-9.8681,21.4391,11.6299)
光源 3	(-2.7513,16.1333,7.7191)
光源 4	(20.7769,23.2294,12.7226)
光源 5	(19.3910,39.3170,19.9647)

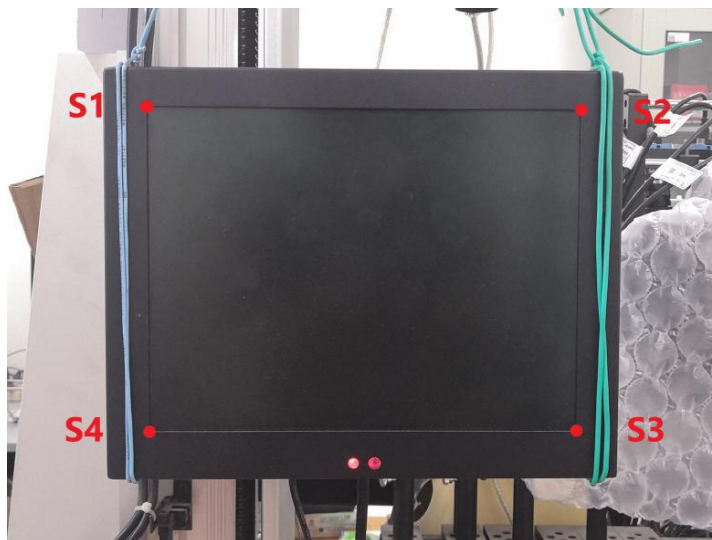


图 3-8 屏幕上四个标定角点

表 3-3 标定得到的屏幕四个角点坐标

屏幕角点坐标（以左相机建立坐标系）/(mm)	
角点 1	(-81.6143,315.3657,-286.2023)
角点 2	(130.1649,327.2562,-285.1138)
角点 3	(145.4129,193.2083,-364.4289)
角点 4	(-66.3663,181.3177,-365.5175)

3.3 用户标定方法

三维视线模型的建立，需要先确定眼球的一些参数。这些参数是眼球的不变参数，由于用户之间的个体差异，这些用户特有的参数无法直接估算，所以，需要通过用户标定程序来获得，这些参数包括（a）角膜半径 R ，（b）视轴与光轴之间的偏离角，即 $Kappa$ 角。在统一坐标系法中，角膜半径标定的目的是统一两个相机成像平面的坐标系。实际上，对于本文提出的新的视线估计方法，不要求解角膜半径，所以角膜半径标定可以省略。尽管如此，接下来还是会对角膜半径的标定进行表述。在用户标定中，要求用户注视屏幕上的一个标定点。因为标定点在屏幕平面上的二维位置是已知的，所以可以通过系统标定恢复系统相机空间中相应的三维位置^[56]。

3.3.1 角膜半径标定

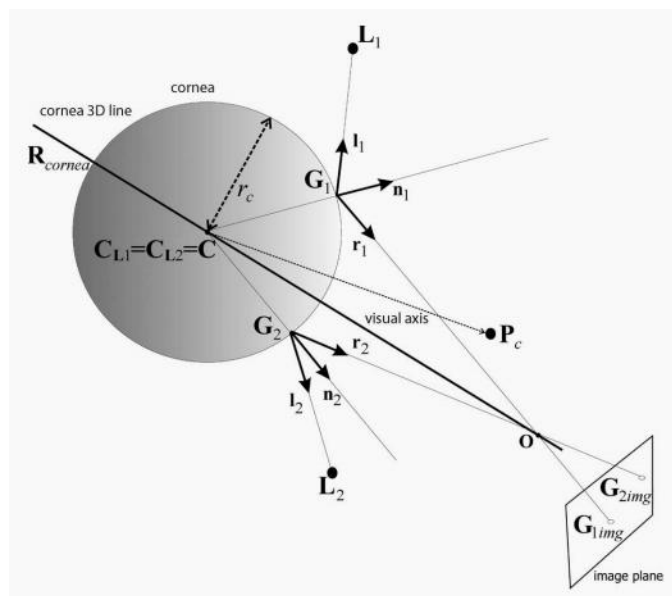


图 3-9 角膜表面反射原理图

在双相机双光源系统中，角膜表面反射原理如图 3-9 所示，已知角膜曲率中心为 C ，角膜外表面半径为 r_c 。相机光心为 O_i ($i=1,2$ ，代表两个系统相机)， L_j ($j=1,2$ ，代表两个光源) 为近红外光源，光源照射眼球时在角膜表面产生反射，反射点为 G_{ij} ($i,j=1,2$ ，分别代表分别在两个相机中成像时的两个光源在角膜上的反射点)， g_{ij} 为 G_{ij} 光源 j ($j=1,2$) 在相机 i ($i=1,2$) 中的成像点，即普尔钦斑。

根据光的反射原理，光的反射满足两个条件：(1) 入射光线，反射光线和法线三线共面。由于光源 L_j 的反射光线分别经过 G_{ij}, O_i, g_{ij} 三点(三点共线)，设： $G_{ij} = O_i + k_{ij}(O_i - g_{ij})$ ，其中 k_{ij} 为常数。(2) 入射角等于反射角。根据反射定律，光源 L_j 满足：

$$\mathbf{r}_{ij} + \mathbf{I}_{ij} = 2\mathbf{n}_{ij}(\mathbf{n}_{ij} \cdot \mathbf{g}_{ij} \cdot \mathbf{I}_{ij}), \quad i=1,2; j=1,2 \quad (3-16)$$

其中， \mathbf{I}_{ij} 为入射光线的单位方向向量， \mathbf{n}_{ij} 为法线的单位方向向量， \mathbf{r}_{ij} 为反射光线的单位方向向量。它们分别可以表示为：

$$\mathbf{I}_{ij} = \frac{(\mathbf{L}_j - \mathbf{G}_{ij})}{\|\mathbf{L}_j - \mathbf{G}_{ij}\|}; \quad \mathbf{n}_{ij} = \frac{(\mathbf{G}_{ij} - \mathbf{C})}{\|\mathbf{G}_{ij} - \mathbf{C}\|}; \quad \mathbf{r}_{ij} = \frac{(\mathbf{O}_i - \mathbf{G}_{ij})}{\|\mathbf{O}_i - \mathbf{G}_{ij}\|} \quad (3-17)$$

将式 (3-17) 代入式 (3-16)，可得：

$$\frac{(\mathbf{O}_i - \mathbf{G}_{ij})}{\|\mathbf{O}_i - \mathbf{G}_{ij}\|} + \frac{(\mathbf{L}_j - \mathbf{G}_{ij})}{\|\mathbf{L}_j - \mathbf{G}_{ij}\|} = 2 * \frac{(\mathbf{G}_{ij} - \mathbf{C})}{\|\mathbf{G}_{ij} - \mathbf{C}\|} * \left(\frac{(\mathbf{G}_{ij} - \mathbf{C})}{\|\mathbf{G}_{ij} - \mathbf{C}\|} \cdot \frac{(\mathbf{L}_j - \mathbf{G}_{ij})}{\|\mathbf{L}_j - \mathbf{G}_{ij}\|} \right) \quad (3-18)$$

由于角膜外表面上的反射点 G_{ij} 到角膜曲率中心 C 点的距离为角膜曲率半径 r_c ，故存在：

$$\|\mathbf{G}_{ij} - \mathbf{C}\| = r_c \quad (3-19)$$

在已知角膜曲率中心 C 的条件下，代入式 (3-19)，则 r_c 满足下式：

$$\|\mathbf{O}_i - \mathbf{g}_{ij}\|^2 k_{ij}^2 + 2(\mathbf{O}_i - \mathbf{g}_{ij}) \cdot (\mathbf{O}_i - \mathbf{C}) k_{ij} + \|\mathbf{O}_i - \mathbf{C}\|^2 - r_c^2 = 0 \quad (3-20)$$

对于上式方程， k_{ij} 可以通过式 (3-18) 解出，然后将其代入 G_{ij} 的表达式中可以求得 G_{ij} 点的坐标，那么式 (3-20) 可以组成一个包含未知量 r_c 的一元二次方程组。为了得到更好的求解结果，可以将其转换为优化问题。方程组为：

$$f_t = \|\mathbf{O}_i - \mathbf{g}_{ij}\|^2 k_{ij}^2 + 2(\mathbf{O}_i - \mathbf{g}_{ij}) \cdot (\mathbf{O}_i - \mathbf{C}) k_{ij} + \|\mathbf{O}_i - \mathbf{C}\|^2 - r_c^2, \quad i,j=1,2; t=1,2,3,4 \quad (3-21)$$

转为无约束优化问题，得：

$$\begin{cases} \min F(r_c) = \sum_{t=1,2,3,4} f_t^2 \\ s.t. (r_c \in R) \end{cases} \quad (3-22)$$

由此，利用最优化原理，当 $F(r_c)$ 的最小值为 0 时， r_c 即为角膜曲率半径的解。

3.3.2 Kappa 角标定

根据人眼的 Le-Grand 模型，视网膜的中心并不完全位于眼球的轴上，所以视线和眼球光轴之间有一定的夹角。在水平方向上，光轴与视轴的夹角大约为 5 度，左眼为 5 度，右眼为 -5 度，竖直角度大约为 1.5 度，个体之间的最大差异不超过 3 度，而某个人的 Kappa 角通常为固定值，所以在进行视线跟踪之前需要标定用户 Kappa 角，只有获得 Kappa 角才可通过光轴获得视轴方向，进而获得视轴与屏幕的交点即人眼视线落点。事实上，Kappa 角的标定过程就是求取视轴与光轴之间转换矩阵的过程。

本文的 Kappa 角标定包含两个部分：

- (1) Kappa 角标定：在用户标定过程中，标定出 Kappa 角的值。
- (2) Kappa 角解算：在视线估计过程中，根据标定的 Kappa 角，实时计算当时的空间 Kappa 角的值。

如图 3-10 所示，在用户标定过程中，以角膜曲率中心 C 为原点，以眼球光轴为 Z 轴，以与 Z 轴垂直并相互垂直的任意两条直线分别为 X 轴和 Y 轴，设立眼球坐标系。使用者盯视屏幕上的标定点 P_c ，根据 4.3 的方法构建眼球光轴，而角膜曲率中心和标定点坐标可以得到视轴，这样就可以确定空间 Kappa 角的值。

$$\cos Kappa = \frac{(E - C) \cdot (P_c - C)}{|E - C| |P_c - C|} \quad (3-23)$$

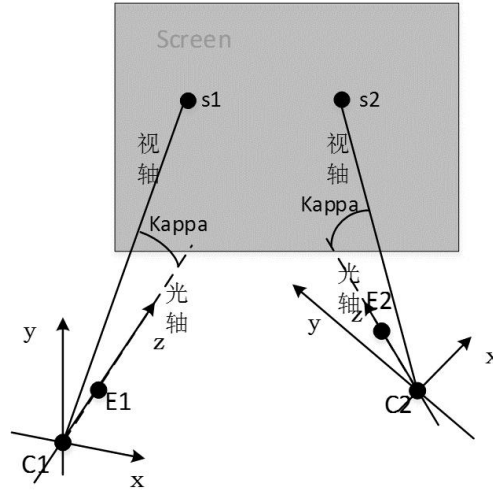


图 3-10 Kappa 角示意图

在视线估计过程中，眼球从用户标定的位置变化到另一个位置，其位置和姿态的改变，可以用坐标系 OXYZ 的旋转和平移来表征。由于 Z 轴为光轴，本文只考虑眼球坐标系的平移和绕 X、Y 轴转动的情况下的 Kappa 角的计算，即计算出眼球在空间任意位置时的空间 Kappa 角。假设眼球在标定位置处其光轴方向向量为 $\mathbf{O}_0 = [x_0 \ y_0 \ z_0]^T$ ，视轴的方向向量为 $\mathbf{V}_0 = [x'_0 \ y'_0 \ z'_0]$ ，角膜曲率中心为 $\mathbf{C}_0 = [c_{x0} \ c_{y0} \ c_{z0}]$ 。在任意位置时眼球光轴向量为 $\mathbf{O}_1 = [x_1 \ y_1 \ z_1]^T$ ，视轴的方向向量为 $\mathbf{V}_1 = [x'_1 \ y'_1 \ z'_1]$ ，角膜曲率中心为 $\mathbf{C}_1 = [c_{x1} \ c_{y1} \ c_{z1}]$ 。在系统相机坐标系下，眼球从标定位置旋转和平移到任意位置时，眼球坐标系原点平移量为： $[c_{x1} - c_{x0} \ c_{y1} - c_{y0} \ c_{z1} - c_{z0}]^T$ ，绕 X 轴旋转 α 角，绕 Y 轴旋转 β 角，绕 Z 轴无转动。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：

$$\mathbf{M}\mathbf{O}_0 = \mathbf{O}_1 \quad (3-24)$$

即：

$$\begin{bmatrix} \cos \beta & 0 & -\sin \beta & c_{x1} - c_{x0} \\ -\sin \alpha \sin \beta & \cos \alpha & -\sin \alpha \cos \beta & c_{y1} - c_{y0} \\ \cos \alpha \sin \beta & \sin \alpha & \cos \alpha \cos \beta & c_{z1} - c_{z0} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ 1 \end{bmatrix} \quad (3-25)$$

上述方程组可以得到两组解，根据用户标定中得到的 Kappa 角，可以确定唯一的 α 角和 β 角。在视线估计过程中，眼球在空间任意位置处重建眼球

光轴后，根据式（3-23）和（3-25）实时计算出眼球坐标系的旋转角 α 和 β ，再根据眼球的平移量 $\begin{bmatrix} c_{x1} - c_{x0} & c_{y1} - c_{y0} & c_{z1} - c_{z0} \end{bmatrix}^T$ ，得到眼球坐标系的变换矩阵 \mathbf{M} 。由于在眼球坐标系中，眼球光轴和视轴之间具有确定的位置关系，因此眼球在标定位置处的视轴和在任意位置处的视轴之间的转换矩阵也是 \mathbf{M} ，则有：

$$\mathbf{V}_1 = \mathbf{M}\mathbf{V}_0 \quad (3-26)$$

这样，就可以由眼球在标定位置处的视轴方向计算出眼球处于任意位置处的视轴方向上述 Kappa 角标定方法可以解决头部平移、转动和俯仰情况下的 Kappa 角计算问题。

3.4 本章小结

本章主要对本文相机标定、系统标定和用户标定的方法进行了介绍，重点介绍了在系统标定中，标定光源和屏幕位置所采用的单相机辅助标定法。标定准确与否将直接影响实验的结果。

4 三维视线估计模型

4.1 三维视线估计方法流程

针对文献^[39,46,66]中描述的传统多相机多光源系统中光轴重建方法的不足,本文提出了一种基于双相机双光源系统的三维视线估计新方法。在双相机双光源系统中,两个光源发出的光线分别在角膜表面发生反射。已知两个相机和两个光源的位置,由光线反射的几何关系可以构建四个反射平面,这些平面两两相交可获得两条交线,两条交线的交点即为所求角膜曲率中心。在此基础上,寻找两个相机中瞳孔边缘上的匹配点,即瞳孔边缘上的空间点分别在两个相机图像上的对应点。这样由瞳孔上的边缘点、角膜曲率中心和相机的光学中心可以构成一个折射平面,在两个相机中得到的两个折射平面的交线,就是瞳孔上的某一空间点与角膜曲率中心的连线。这样根据一个相机中的瞳孔图像的多个边缘点与其在另一个相机中的匹配点,就可以求得多条瞳孔边缘点与角膜曲率中心的连线,这些连线构成了一个空间圆锥,则过角膜曲率中心的圆锥的对称轴即为眼球的光轴。在这种方法中,不需要检测瞳孔中心,不必考虑角膜折射带来的影响,因此避免了由瞳孔中心光的折射引起的误差。该方法更加准确,且简化为单点用户标定。该方法的流程图如图 4-1 所示。

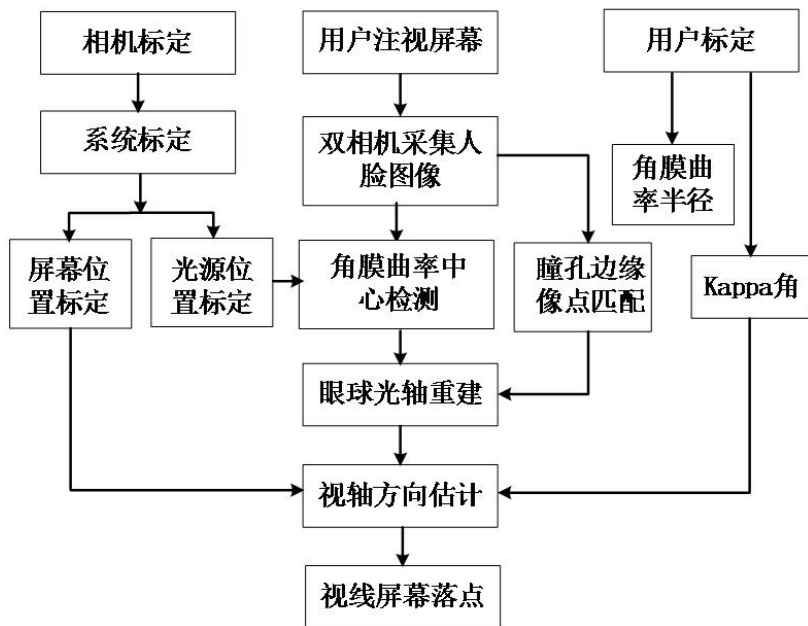


图 4-1 本文方法的流程图

(1) 系统标定。系统标定的目的是确定光源和屏幕在系统相机空间中的位置。系统标定还包括确定两个系统相机之间的相对位置关系，该相对位置关系用两个相机的坐标变换矩阵来表示。本文的系统标定已经在 3.2 节进行了说明。

(2) 用户标定。用户标定的目的是标定角膜半径和 $Kappa$ 角。在用户标定过程中，当用户盯视屏幕上的标定点时，可以基于 (3) 的方法来计算角膜曲率中心。然后通过 (4) 的方法确定眼球的光轴。因此，根据用户标定中检测到的光轴和视线方向，可以标定 $Kappa$ 角。本文的 $Kappa$ 角标定在 3.3 节进行了说明。在本文方法中，不需要标定角膜半径。由于角膜曲率半径具有个体差异，所以在统一坐标系法中，角膜半径标定的目的是统一两个相机成像平面的坐标系。

(3) 求取角膜曲率中心。在双相机双光源系统中，两个光源发出的光线分别在角膜表面发生反射。已知两个相机和两个光源的位置，由光线反射的几何关系可以构建四个反射平面，这些平面两两相交可获得两条交线，两条交线的交点即为所求角膜曲率中心。但由于光源位置、图像分割等存在误差的影响，求取的角膜曲率中心往往不够精确，还需对角膜曲率中心做进一步的精确估计。

(4) 光轴重建。寻找两个相机中瞳孔边缘上的匹配点，即瞳孔边缘上的空间点分别在两个相机图像上的对应点。这样由瞳孔上的边缘点、角膜曲率中心和相机的光学中心可以构成一个折射平面，在两个相机中得到的两个折射平面的交线，就是瞳孔上的某一空间点与角膜曲率中心的连线。这样根据一个相机中的瞳孔图像的多个边缘点与其在另一个相机中的匹配点，就可以求得多条瞳孔边缘点与角膜曲率中心的连线，这些连线构成了一个空间圆锥，则过角膜曲率中心的圆锥的对称轴即为眼球的光轴。基于以上原理，4.4 节介绍了三种光轴重建方法。

(5) 视线方向估计。根据 (4) 中重建的光轴和 (2) 中标定的 $Kappa$ 角，最终确定视线方向和视线落点。

4.2 角膜曲率中心估计

由文献^[46,66]可知，角膜被看作是以 C 为球心， r_c 为半径的球面。点光源发出的光线照射到角膜上，形成反射亮斑，亮斑在相机中成像。由反射定律可知：(1) 入射角与反射角相等；(2) 点光源、入射光线、反射光线及法线共面。

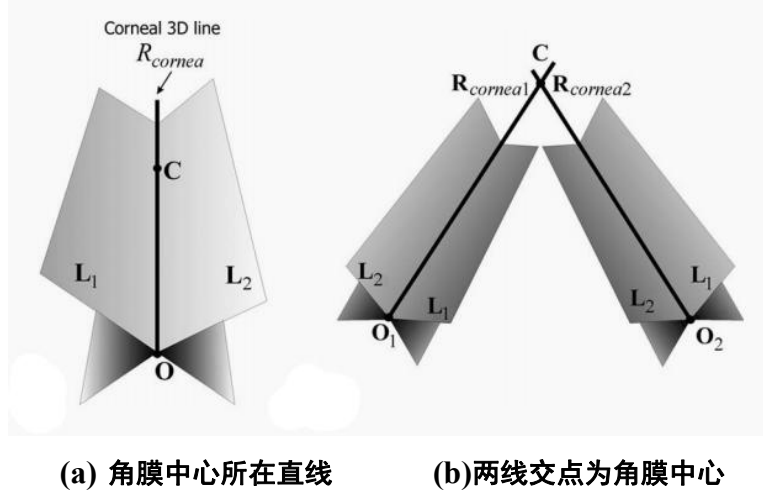


图 4-2 求取角膜曲率中心原理图

如图 4-2 (a) 所示, 设光源 L 、入射光线、反射光线和图像中亮斑 G_{img} 构成的平面为 Π_L 。角膜曲率中心 C 在角膜球面入射点的法线上, 因此, C 在平面 Π_L 上, 如果有两个点光源 L_1 和 L_2 , 在角膜表面发生两次反射, 则形成两个普尔钦斑, 在相机像平面产生两个亮斑图像, 分别记作 G_{img1} 和 G_{img2} 。每个光源和它产生的图像上的亮斑可确定一个反射平面, 该平面包含角膜曲率中心 C 和相机光心 O 。

$$\Pi_{L1} = L_1 \times G_{img1}, \Pi_{L2} = L_2 \times G_{img2} \quad (4-1)$$

角膜曲率中心 C 在两个平面的交线上。同理, 在双相机系统中, 如图 4-2 (b) 所示, 可以得到两条经过角膜曲率中心 C 的直线, 这两条直线的交点即角膜曲率中心 C 。

4.3 瞳孔边缘匹配的极平面法

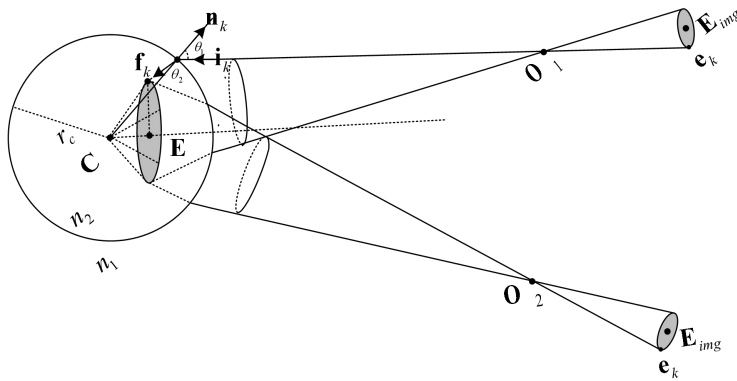


图 4-3 眼球光轴重建示意图

如图 4-3 所示, 本文注意到角膜曲率中心和空间中的瞳孔边缘点的连线, 构成一个空间圆锥体 Ω , 而这个圆锥体的中心轴线就是眼球的光轴。本文方法通过两个相机的图像上瞳孔的边缘点, 求取圆锥 Ω 上的多个连接角膜曲率中心和瞳孔边缘点的直线, 则通过这些直线就能够构造出空间的圆锥体 Ω , 再求得这个圆锥体 Ω 的中心轴线, 就确定了眼球的光轴 CE。在双相机双光源系统中, 空间瞳孔的边缘点经过角膜折射后, 分别在两个相机中成像, 在两个相机的瞳孔图像中, 需要确定两个相机中瞳孔边缘上的对应的匹配点, 并根据这些成对的匹配点, 重建空间圆锥表面上的通过角膜曲率中心的直线, 因此, 确定两个相机中瞳孔边缘上的对应的匹配点是本文方法的关键。

如图 4-4 所示, C_1 、 C_2 点分别为角膜中心 C 在相机 1 和相机 2 的像平面上的投影点, C 、 O_1 、 O_2 、 C_1 、 C_2 构成了极平面, M_1N_1 、 M_2N_2 分别为角膜中心 C 在相机 1 的成像面和相机 2 的成像面上的极线, 且 M_1 、 N_1 为相机 1 的像平面上的极线与瞳孔的椭圆像的交点, M_2 、 N_2 为相机 2 的像平面上的极线与瞳孔的椭圆像的交点, 则 M_1 和 N_1 、 M_2 和 N_2 分别为相机 1 的像平面和相机 2 的像平面上的匹配点。即瞳孔边缘在两个相机成像面上的匹配点。

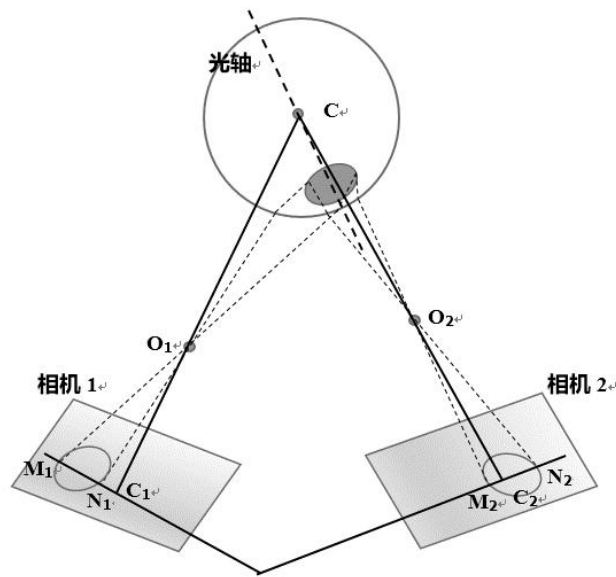


图 4-4 瞳孔及角膜中心成像示意图

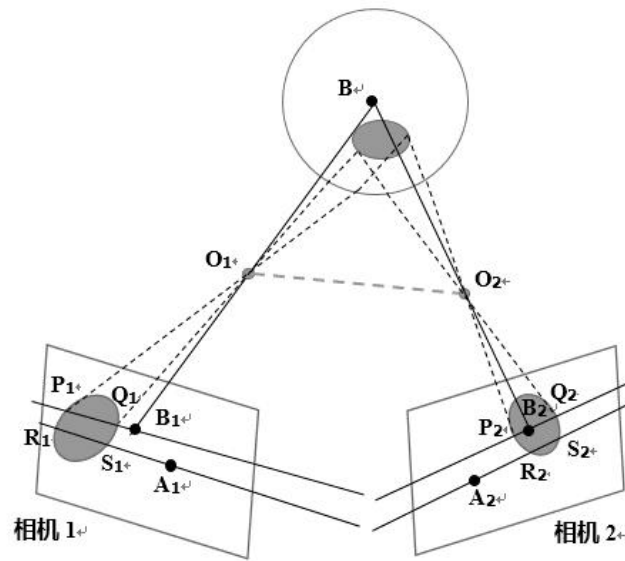


图 4-5 图像瞳孔边缘的匹配点示意图

同理，如图 4-5 所示，设 B_1 、 B_2 分别为空间点 B 在两个相机像平面上的投影，而 B 点的极平面与两个像平面所交的极线，与两个相机的瞳孔像分别相交于 P_1 、 Q_1 和 P_2 、 Q_2 ，则 P_1 和 Q_1 、 P_2 和 Q_2 互为匹配点。进一步，设 A_1 为在相机 1 的像平面上的一点，其为空间中的某一点 A 在像平面的投影点，则 A_1 、 O_1 和 O_2 构成的平面即为 A 点的极平面，该平面与两个相机中的椭圆瞳孔像的边缘分别交于 R_1 、 S_1 和 R_2 、 S_2 。则 R_1 和 S_1 、 R_2 和 S_2 一定分别互为匹配点。

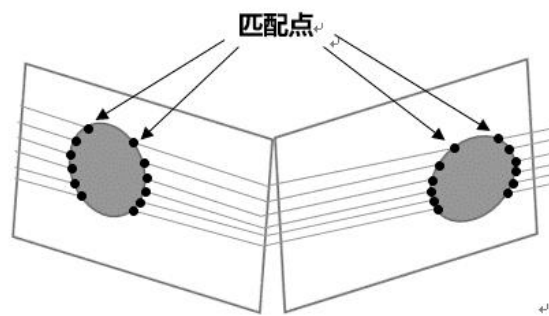


图 4-6 两个相机中瞳孔椭圆边缘上匹配点示意图

可见，在相机 1 的像平面上任取一点，做出其极平面后，则该极平面在两个相机像平面上所形成的极线与椭圆瞳孔像的边缘的交点一定互为匹配点。这样就可以得到瞳孔边缘上的多对匹配点，如图 4-6 所示。

4.4 光轴估计

4.4.1 极平面圆锥高线法

由折射定律可知，入射光线、折射光线、法线在同一平面上。已知角膜曲率中心 C ，两个相机的光心 O_1 、 O_2 ，图像瞳孔边缘点 A_1 和 A_2 、 B_1 和 B_2 。 C 在眼球光轴上， A_0 在入射光线上， O_1 、 A_1 在折射光线上，因此， C 、 A_0 、 O_1 、 A_1 共面。同理可得 C 、 A_0 、 O_2 、 A_2 共面。两个平面分别记作 Π_{A1} 、 Π_{A2} ，可表示为：

$$\Pi_{A1} = CO_1 \times O_1A_1 \quad (4-2)$$

$$\Pi_{A2} = CO_2 \times O_2A_2 \quad (4-3)$$

平面 Π_{A1} 、 Π_{A2} 的交线为 CA_0 。同理， C 、 B_0 、 O_1 、 B_1 共面， C 、 B_0 、 O_2 、 B_2 共面，两个平面分别记作 Π_{B1} 、 Π_{B2} ，可表示为：

$$\Pi_{B1} = CO_1 \times O_1B_1 \quad (4-4)$$

$$\Pi_{B2} = CO_2 \times O_2B_2 \quad (4-5)$$

平面 Π_{B1} 、 Π_{B2} 的交线为 CB_0 。

这样，对于两个相机图像中所有的瞳孔边缘匹配点（4.3 节中求得） D_1 、 E_1 、 F_1 ... 和 D_2 、 E_2 、 F_2 ...，都可以求得相应的空间直线 CD_0 、 CE_0 、 CF_0 ...。则 CA_0 、 CB_0 、 CD_0 、 CE_0 、 CF_0 ... 等直线为母线构成了以 C 点为顶点、瞳孔为底面的空间圆锥体 Ω ，而圆锥体 Ω 的高线即为光轴。对于任意的一对瞳孔边缘上的匹配点确定圆锥体上的一条母线，例如 CA_0 、 CB_0 ，对于 CA_0 ，其方向向量 \hat{n} 可表示为：

$$\hat{n} = \frac{(M_1 C \times M_1 O_1) \times (M_2 C \times M_2 O_2)}{|(M_1 C \times M_1 O_1) \times (M_2 C \times M_2 O_2)|} \quad (4-6)$$

对于瞳孔边缘上任意一点 P_i ，均可由上式得到一条母线 CP_i ，其单位方向向量记作 \hat{n}_i ，则有：

$$P_i = C + k_i \hat{n}_i \quad (4-7)$$

如图 4-3 所示，角膜曲率中心 C 与瞳孔边缘构成一个正圆锥， P_i 为正圆锥底面边缘上一点，因此比例系数 k_i 为常数，记作 k 。多个 P_i 所在平面法线向量即为圆锥中心轴的方向向量，因此光轴可表示为：

$$\mathbf{op} = \frac{1}{n-2} \sum_{i=3}^n \frac{P_{i-1}P_i \times P_iP_{i+1}}{|P_{i-1}P_i \times P_iP_{i+1}|} (i \geq 3) \quad (4-8)$$

带入化简可得，

$$\mathbf{op} = \frac{1}{n-2} \sum_{i=3}^n \frac{\hat{\mathbf{n}}_{i-1} \hat{\mathbf{n}}_i \times \hat{\mathbf{n}}_i \hat{\mathbf{n}}_{i+1}}{|\hat{\mathbf{n}}_{i-1} \hat{\mathbf{n}}_i \times \hat{\mathbf{n}}_i \hat{\mathbf{n}}_{i+1}|} (i \geq 3) \quad (4-9)$$

式（4-9）即为所求光轴，且与比例系数 k 无关。由式（4-9）可知，需至少三对匹配点才能计算空间圆形瞳孔的法线。当匹配点数量足够多时，为提升算法鲁棒性和快速性，可采用带剔除的最小二乘法进行优化。

4.4.2 极平面椭圆长轴法

空间圆的投影是一个椭圆，并且图像椭圆的中心不对应于 3D 圆的中心。如图 4-7 所示，在圆的平面图中有一些构造线。我们可以看到，圆心是两个正方形对角线的交点，并且该关系具有仿射不变性。如图 4-7 的右侧所示，我们可以清楚地看到椭圆的中心任意地远离圆的真实中心的投影。椭圆必须穿过外接四边形的边缘点，并且椭圆的长轴也必须对应于将正方形的相切点与圆连接起来的线。简而言之，投影圆的真实中心位于图像椭圆的长轴上。

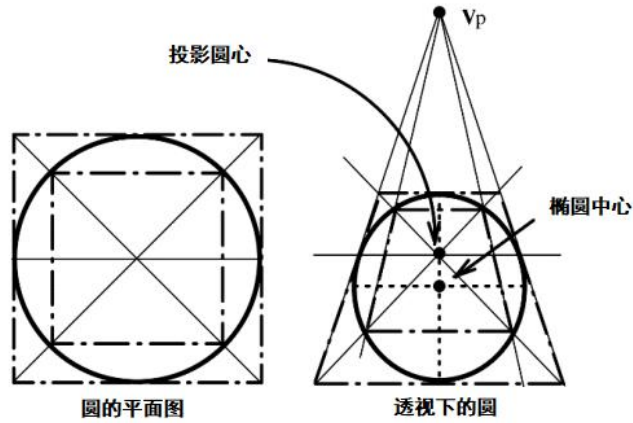


图 4-7 圆的透视构造图

可以证明，对于空间瞳孔来说，经过角膜的折射后在相机中成像，则图像的瞳孔椭圆的长轴仍然对应瞳孔的一个直径。

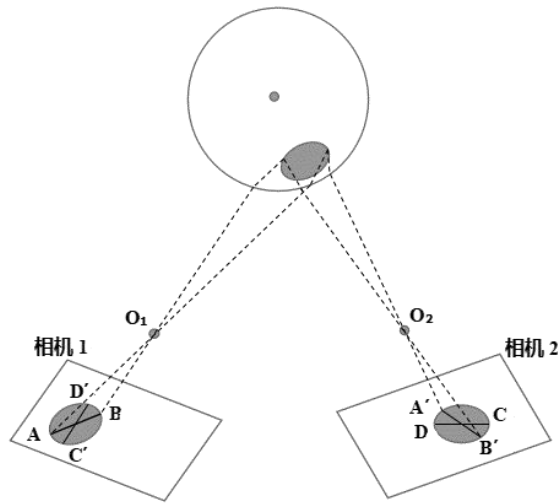


图 4-8 极平面椭圆长轴法示意图

如图 4-8 所示，瞳孔在相机 1 和相机 2 中分别成两个椭圆像，设相机 1 中的椭圆为 T_1 ，相机 2 中的椭圆为 T_2 ， T_1 的长轴为 AB ， T_2 的长轴为 CD 。根据 4.3 节中所提出的求取匹配点的方法，可以确定相机 1 中的 A ， B 在相机 2 中的匹配点为 A' ， B' 。同理，可以求得相机 2 中的 C ， D 在相机 1 中的匹配点 C' ， D' 。则由 A 和 A' ， B 和 B' 可以重构出空间圆锥体 Ω 的两条母线，而这两条母线对应着圆锥底面的直径，因此这两条母线的角平分线就是光轴。同理， C 和 C' ， D 和 D' 也可以重构出空间圆锥体 Ω 的两条母线，并求出光轴。

4.4.3 极平面长轴交点法

在上述极平面椭圆长轴法的基础上更进一步，如图 4-9 所示，在相机 1 中，长轴 AB 和直线 $C'D'$ 交点 W 为空间瞳孔圆心在相机 1 中的真实投影点。而在相机 2 中，长轴 CD 和直线 $A'B'$ 的交点 V 为空间瞳孔圆心在相机 2 中的真实投影点，则 W 和 V 互为匹配点，则根据 W 和 V 可以直接重构出眼球的光轴。

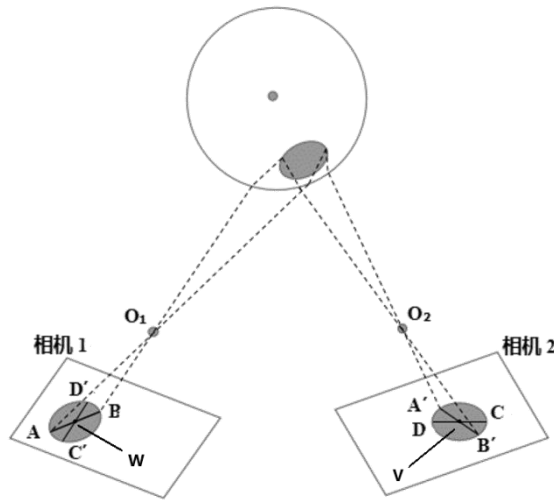


图 4-9 极平面长轴交点法示意图

根据上述三种构建光轴的方法和 3.3 标定的 Kappa 角，设 \mathbf{M} 为视轴与光轴之间的转换矩阵，可以实时计算出三维视线方向，即视轴方向：

$$\text{LoS} = \mathbf{M} \cdot \text{LoG} \quad (4-10)$$

之后，由系统标定的参数计算视轴和屏幕的交点，即人眼在屏幕上的视线落点。

4.5 本章小结

本章详细介绍了本文提出的三维视线估计方法。首先提出了本文三维视线估计方法的流程，先进行了相机标定、系统标定和用户标定，之后用户注视屏幕，双相机采集人脸信息，提取特征参数进行角膜曲率中心估计，接着采用瞳孔边缘匹配的极平面法寻找匹配点，然后介绍了光轴估计的三种方法，通过标定的转换矩阵得到视轴方向，进而得到视线落点。

5 三维视线估计实验分析

5.1 仿真实验

通过仿真实验验证了该算法的可行性和性能，本文使用 3D 绘图软件 Rhino5.0 和 Matlab 进行了仿真实验。

通过使用基于 Le Grand 眼睛模型的 3D 软件 Rhino5.0，建立了具有两个相机和两个光源的三维视线跟踪系统。光源、屏幕和系统相机的相对位置是根据实际系统的比例确定的。从 Rhino5.0 开始，该视线跟踪系统的所有参数都是已知的，包括光源和标定点的位置，眼睛的特征参数，相机的焦距等。因此，根据第四章的算法模型，可以在 Matlab 中用这些参数来标定眼睛的不变参数，并根据 4.4 节计算视线落点。如图 5-1 所示，C 点为角膜曲率中心，E 点为瞳孔中心，角膜曲率半径为 8mm，瞳孔离角膜中心距离为 5mm， O_1 、 O_2 为两个相机光学中心，离角膜曲率中心距离 33mm， L_1 、 L_2 为光源，两个相机的成像面与相机光学中心距离（即焦距）为 2.7/5.4mm，像元大小设为 2.7 μm ，相机成像的放大倍率为 2.7/33 和 5.4/33，将计算的视线落点与预定义的标定点进行比较，以验证所提出方法的可行性。分析了光源位置对标定和三维视线估计结果的影响，此外，在 Matlab 程序中将不同的噪声添加到瞳孔轮廓参数和光源位置参数中，以测试所提出方法的性能。

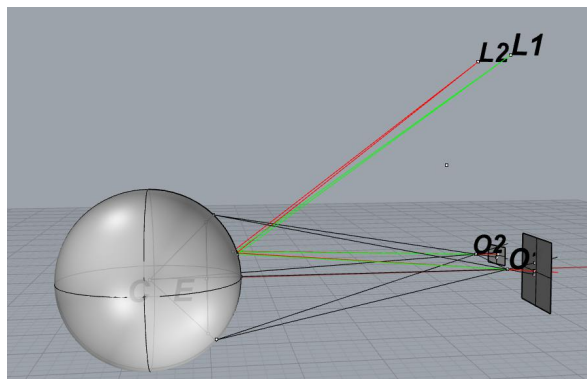


图 5-1 视线估计方法仿真示意图

表 5-1 中所列数据为在 2.7/33 的放大倍率下，普尔钦斑、匹配点、光源三者误差对光轴重建的影响。对普尔钦斑的中心的坐标和图像匹配点坐标分别加均值为 0，方差为以 0.25 个像素（1 像素等效于 2.7 μm ）为步长的递增的高斯噪声，对光源加均值为 0，方差为以 0.5mm 为步长的递增的高斯分布噪声，每种变化条件下取 100 个计算结果进行平均得到仿真结果。

表 5-1 普尔钦斑、匹配点、光源误差对光轴的影响

普尔钦斑误差/像素	光轴误差/ $^{\circ}$	匹配点误差/像素	光轴误差/ $^{\circ}$	光源误差/mm	光轴误差/ $^{\circ}$
0.25	0.205999	0.25	0.09041	0.5	1.523979
0.5	0.305576	0.5	0.233952	1	2.162092
0.75	0.35034	0.75	0.189797	1.5	2.748579
1	0.441151	1	0.199051	2	3.30997
1.25	0.490752	1.25	0.631669	2.5	3.689918
1.5	0.500322	1.5	0.293053	3	4.120048
1.75	0.564998	1.75	0.132646	3.5	4.378683
2	0.569984	2	0.307887	4	5.254727
2.25	0.618123	2.25	0.379035	4.5	5.531509
2.5	0.680151	2.5	0.092488	5	5.686981

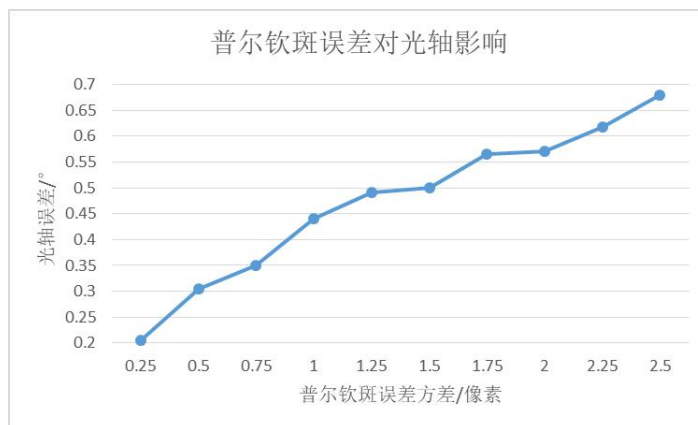


图 5-2 普尔钦斑检测误差对光轴重建精度的影响

可以看出，普尔钦斑误差越大，光轴误差越大，两者近似线性相关，在普尔钦斑存在 2 个像素的误差时，眼球光轴的误差只有不到 0.6° ，普尔钦斑存在 1 个像素误差时眼球光轴误差不到 0.5° ，而在实际实验中，采用亚像素方法，普尔钦斑的检测误差往往不会高于一个像素，即普尔钦斑对于光轴的影响不会高于 0.5° ，因此普尔钦斑检测误差对光轴重建结果的影响较小。

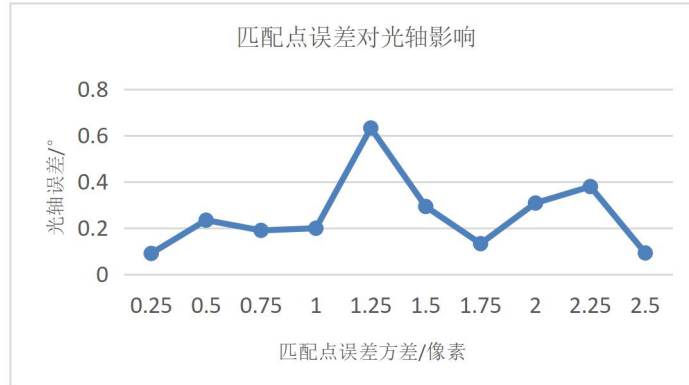


图 5-3 匹配点误差对光轴重建的精度影响

考虑在两个相机图像中的一对瞳孔边缘上的匹配点，对其中的一个点的坐标值加均值为 0，不同方差的高斯噪声，然后计算不同匹配点误差情况下光轴重建的误差。两个相机图像中瞳孔边缘点错误匹配对光轴重建精度的影响如图 5-3 所示，可以看出在 2.5 个像素以内，基本是随机分布，且对光轴的影响小于 0.7° ，而在实际系统中，匹配点误差一般不会超过若干个像素。因此可以认为，匹配点选取的误差对于结果的影响不大。

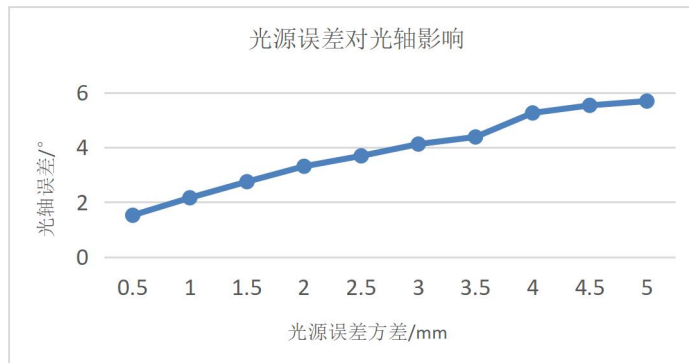


图 5-4 光源误差对光轴重建精度的影响

由图 5-4 可知，对于光源来说，其对光轴的影响也是近似线性的，在光源误差方差为 1mm 时，光轴的重建误差将会达到将近 3.5° ，而在 5mm 误差时光轴的偏差超过了 5.5° ，而 Kappa 角的大小一般为 5° ，因此光源对于眼球光轴的影响非常大，在标定时至少需要精确到 0.5mm 以下。

考虑上述各个因素的综合影响，光轴重建的结果如表 5-2 所示。由于普尔钦斑和匹配点是在同一幅图像中，通过类似的图像处理方法进行提取，因此在仿真时对二者加入相同的，均值为 0，方差按 0.25 个像素（1 像素等效于 $2.7\mu\text{m}$ ）递增的高斯分布的噪声。对光源来说，其标定误差一般可达到

0.5mm 以内，因此对光源加均值为 0，方差按 0.5mm 递增的高斯分布噪声，以仿真真实情况。图 5-5 中，每条线分别代表不同的光源误差，一条线下，光源误差的值标记在折线后方。

表 5-2 2.7/33 放大倍率下各因素对光轴重建的综合影响

光轴误差/°		光源误差/mm									
		0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
图像	0.25	1.599	2.247	2.759	3.397	3.728	3.925	4.531	5.062	5.495	6.208
检测	0.5	1.63	2.305	2.851	3.147	3.658	4.108	4.647	5.315	6.001	6.384
误差/	0.75	1.617	2.24	2.779	3.214	3.8	4.126	4.392	4.732	5.765	5.995
像素	1	1.706	2.31	2.885	3.261	3.694	4.227	4.818	5.497	5.73	6.404
	1.25	1.626	2.378	2.767	3.268	3.572	4.316	4.715	5.142	5.492	6.211
	1.5	1.708	2.325	2.887	3.206	3.95	4.151	4.709	5.375	5.672	6.026
	1.75	1.654	2.275	3.018	3.372	3.807	4.292	4.186	5.023	5.661	6.197
	2	1.704	2.262	2.902	3.209	3.811	4.594	4.88	5.218	5.631	6.586
	2.25	1.705	2.288	3.182	3.217	3.65	4.048	4.532	5.343	5.768	6.389
	2.5	1.768	2.281	2.822	3.379	3.707	4.101	4.488	5.061	5.637	5.872

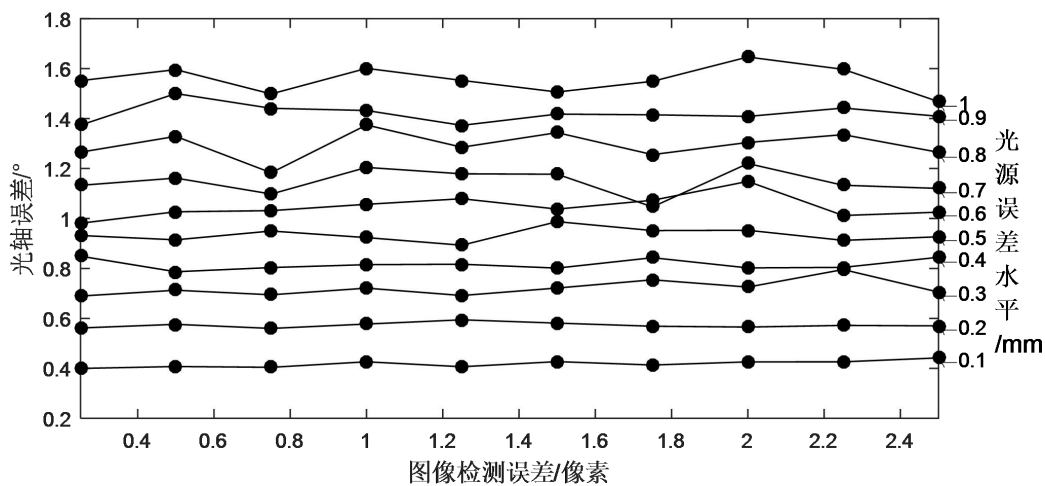


图 5-5 普尔钦斑、匹配点和光源位置对结果的综合影响

由上图可以看出，光源对眼球光轴重建的影响很大，普尔钦斑、匹配点、光源三者对眼球光轴重建的影响的关联性不大，误差主要来源于光源位置的误差。总误差可看作是普尔钦斑、匹配点、光源三者误差的叠加。

5.2 真实实验设备

本文采用的实验系统为作者设计的头盔式视线跟踪系统的视线检测组件，通过 USB 接口与计算机连接，如图 5-6 所示。它是一个双相机多光源系统。系统相机的镜头焦距为 2.8mm，相机分辨率为 640×480 pixels，像元大小为 3 μ m。预先通过坐标变换矩阵对两个系统相机之间的关系进行了标定。光源为红外 LED 构成，均匀分布在环形方框上，光源波长为 850nm，光源的亮度可调。先前已经使用 3.2.2 节中所述的方法进行了系统标定，以便确定两个系统相机的坐标系中的光源和屏幕平面的位置。人眼距离相机、光源的距离约为 35mm。

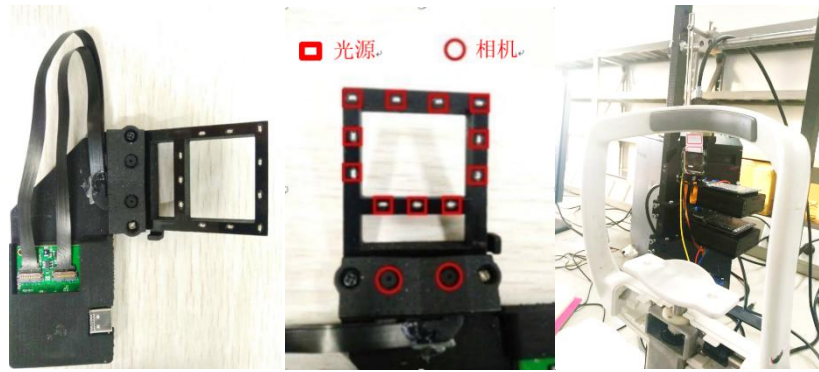


图 5-6 头盔式系统视线检测组件示意图

5.3 真实实验过程

本文的视线估计实验分为两个部分：用户标定实验和三维视线估计实验。根据 3.3 节的方法标定用户眼球的不变参数。先采用 4.4 节的方法重建眼球光轴，并根据标定出的 Kappa 角，重建眼球的视轴。之后将本文方法的视线估计结果与文献^[39,46,66]中传统方法的视线估计结果进行对比，验证本文方法的优越性。

5.3.1 用户标定实验

在用户标定中，我们采用单点标定。受试者进行用户标定过程的示意图如图 5-7 所示。受试者的下巴置于系统的下巴托上，眼睛盯视屏幕上的一点，即我们预设的标定点。注视时间约为 1s。在这一过程中，两个系统相机均拍摄人眼图像，定位眼睛，分割瞳孔和普尔钦斑，检测用户标定所需要的图像眼睛参数。光轴与视轴之间的夹角就是 Kappa 角。八个用户的 Kappa 角的标

定结果如表 5-3 所示。

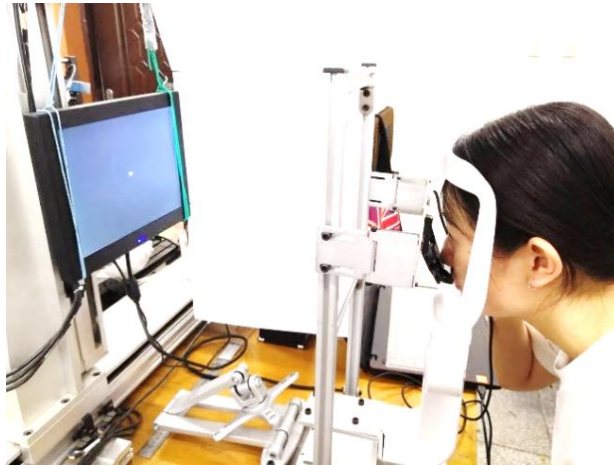


图 5-7 用户标定过程

表 5-3 各用户标定结果

用户序号	1	2	3	4	5	6	7	8
Kappa 角/°	5.6416	5.8068	5.0829	5.1668	4.7614	4.5889	5.4767	5.2320

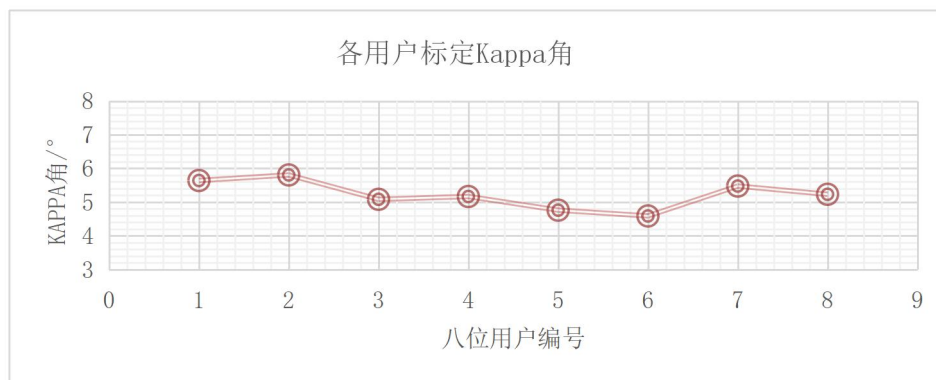


图 5-8 各用户标定 Kappa 角

由于 Kappa 角的大小约为 5° ，因此由图 5-8 可以看出，采用角膜中心极平面法进行图像瞳孔边缘点的匹配计算所标定的 Kappa 角的范围，基本接近于人类 Kappa 角的平均值。

5.3.2 本文三维视线估计方法实验

如图 5-9 所示，屏幕在系统相机中的位置已经由系统标定完成，因此测试点在系统相机坐标系中的坐标是已知的。在三维视线估计实验中，屏幕上设置了均匀分布的 25 个测试点，用于测试计算视线落点与实际注视点之间的

偏差，以确定该系统的视线估计精度。上述标定了 Kappa 角的用户同样需要将头部置于系统的下巴托上，将眼睛放置于光环前面，使光源在用户的角膜上成像，用户依次注视屏幕上的每一个测试点，注视每一个点的时间约为 1s。在注视过程中，受试者的头部可以在平移、俯仰和偏视三个方向进行轻微运动，使头部与相机之间发生小的位移变化。与此同时，两个系统相机同步采集受试者的人眼图像，采样频率仍为 10 帧/s。

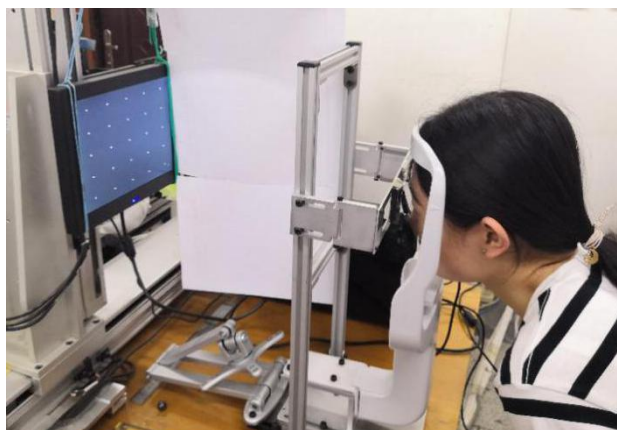


图 5-9 三维视线估计实验过程

对于较大视野的相机，首先需要进行人眼检测，在 Visual Studio 2013 环境下采用 OpenCV 自带的人眼检测库函数，检测人眼位置。输入为拍摄的人眼图像，如图 5-10 所示，输出参数为人眼的位置坐标，包括起始点位置和人眼在图片中所占像素的长宽。得到人眼的位置参数之后，将参数对应人眼区域图像截取出来，对这部分小区域进行处理。在实验中，由于通过眼球图像提取的参数为瞳孔边缘和普尔钦斑，为了降低后续图像处理的难度，在得到上述检测结果之后，往往对输出的参数进行调整，具体做法为减小区域大小，使得区域能够尽可能只包含瞳孔和普尔钦斑。

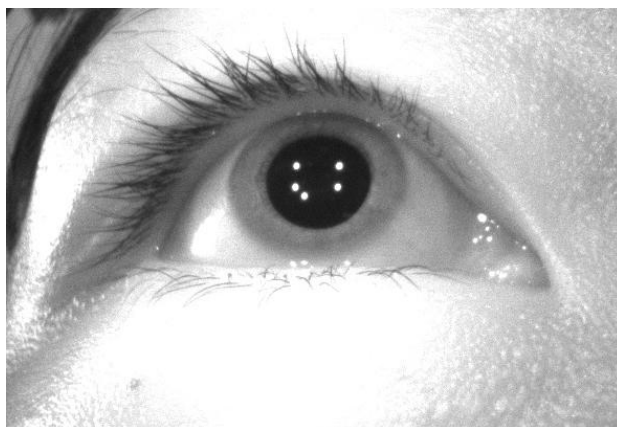
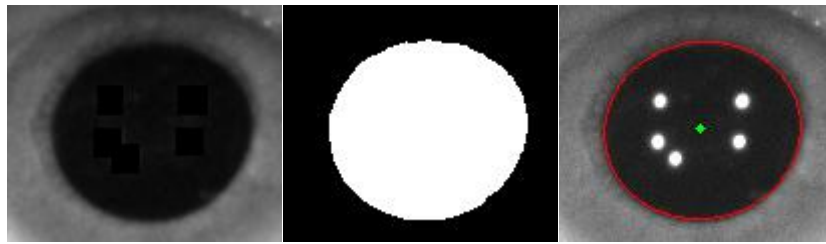


图 5-10 实验中右相机拍摄人眼图

如图 5-11 所示，首先对普尔钦斑所在位置进行检测，具体检测方法为以一个较大的阈值，对图像进行二值化，二值化后的图像中将只存在普尔钦斑的信息，轮廓检测即可得到普尔钦斑的位置，将普尔钦斑所在位置的图像坐标以及普尔钦斑的图像保存。根据质心法，求解出普尔钦斑的坐标，得到一组参数。之后对原始的瞳孔图像进行处理，将普尔钦斑的所在位置区域的像素灰度值进行更改，即从图像上消除普尔钦斑，使得瞳孔边缘的拟合不受普尔钦斑影响。消除普尔钦斑的影响之后，可以对瞳孔的图像进行图像分割，以确定瞳孔边缘，在这里采用拟合的方式，对得到的瞳孔边缘进行拟合，得到拟合后瞳孔椭圆的参数。



(1) 原始瞳孔图像 (2) 二值化后的瞳孔图像 (3) 定位到的普尔钦斑



(4) 消除普尔钦斑图像 (5) 分割后的瞳孔图像 (6) 拟合后的瞳孔轮廓

图 5-11 图像处理过程

提取到了瞳孔边缘点坐标和普尔钦斑中心坐标后，采用 4.3 节中的角膜中心极平面法求取两个相机中瞳孔边缘点的匹配点。这里选用了 20 对匹配点。

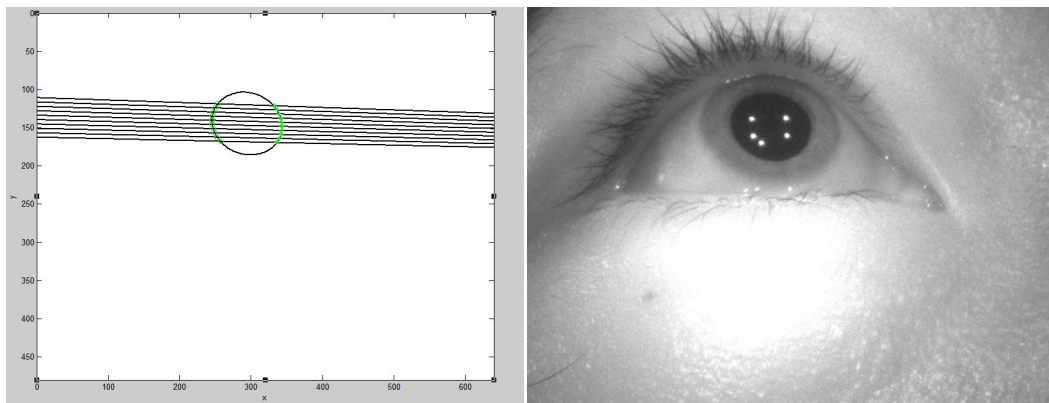


图 5-12 左相机匹配点示意图

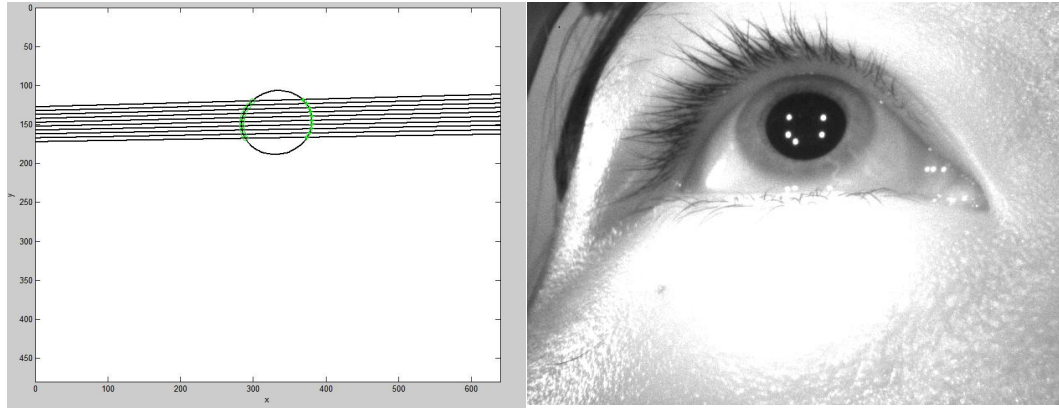


图 5-13 右相机匹配点示意图

图 5-12、图 5-13 分别为用户 4 在左相机、右相机成像面上匹配点的示意图，匹配点在图像坐标系下的二维坐标以及世界坐标系下的三维坐标见表 5-4。

表 5-4 左右相机匹配点坐标

序号	左相机(图像坐标系下) / (像素)	左相机(世界坐标系下) / (mm)	右相机(图像坐标系下) / (像素)	右相机(世界坐标系下) / (mm)
1	(332.1788,121.5278)	(-0.128,-0.367,-2.008)	(368.3506,118.2410)	(14.840,-0.424,0.306)
2	(253.3580,118.9766)	(0.109,-0.375,-2.008)	(298.6175,119.9952)	(15.041,-0.419,0.366)
3	(336.5469,126.9988)	(-0.141,-0.351,-2.008)	(373.3586,123.5067)	(14.826,-0.408,0.302)
4	(249.7218,124.2975)	(0.120,-0.359,-2.008)	(293.3479,125.4309)	(15.056,-0.403,0.371)
5	(339.7039,132.4227)	(-0.150,-0.335,-2.008)	(376.8524,128.8187)	(14.816,-0.392,0.299)
6	(247.3353,129.6650)	(0.127,-0.343,-2.008)	(289.5593,130.8213)	(15.067,-0.387,0.374)
7	(341.8646,137.8101)	(-0.157,-0.318,-2.008)	(379.1880,134.1642)	(14.809,-0.376,0.297)
8	(245.9837,135.0680)	(0.131,-0.327,-2.008)	(286.8958,136.1794)	(15.075,-0.371,0.376)
9	(343.1397,143.1680)	(-0.161,-0.302,-2.008)	(380.5453,139.5357)	(14.806,-0.360,0.295)
10	(245.5561,140.4997)	(0.132,-0.310,-2.008)	(285.1773,141.5124)	(15.080,-0.355,0.377)
11	(343.5777,148.5008)	(-0.162,-0.286,-2.008)	(381.0164,144.9280)	(14.804,-0.344,0.295)
12	(246.0041,145.9553)	(0.131,-0.294,-2.008)	(284.3119,146.8253)	(15.082,-0.339,0.378)
13	(343.1793,153.8117)	(-0.161,-0.270,-2.008)	(380.6376,150.3376)	(14.806,-0.327,0.295)
14	(247.3269,151.4315)	(0.127,-0.278,-2.008)	(284.2632,152.1217)	(15.082,-0.323,0.378)
15	(341.8979,159.1028)	(-0.157,-0.255,-2.008)	(379.4005,155.7618)	(14.809,-0.311,0.296)
16	(249.5712,156.9261)	(0.120,-0.261,-2.008)	(285.0397,157.4041)	(15.080,-0.307,0.377)
17	(339.6247,164.3749)	(-0.150,-0.239,-2.008)	(377.2491,161.1986)	(14.815,-0.295,0.298)
18	(252.8455,162.4380)	(0.110,-0.245,-2.008)	(286.6973,162.6743)	(15.076,-0.291,0.376)
19	(336.1485,169.6275)	(-0.140,-0.223,-2.008)	(374.0628,166.6469)	(14.825,-0.278,0.301)
20	(257.3613,167.9679)	(0.097,-0.228,-2.008)	(289.3565,167.9335)	(15.068,-0.275,0.374)

基于两个相机中的这些匹配点，采用 4.4 节中的三种眼球光轴构建方法，即圆锥高线法、椭圆长轴法和长轴交点法来进行光轴重建。根据 4.2 节中的

方法检测角膜曲率中心，再结合用户标定中每个用户的 $Kappa$ 角，确定屏幕上标定点在系统相机坐标系中的坐标（由系统标定来确定），重建每个人的视轴及其在屏幕上的视线落点。本文邀请八位用户进行实验，用户 4 的实验结果如表 5-5 所示。其中，极平面圆锥高线法求取光轴采用 20 对匹配点。

表 5-5 用户 4 视线估计实验结果

（法一为圆锥高线法，法二为椭圆长轴法，法三为长轴交点法）

序号		视线落点	标定点	光轴	视轴
1	法一	(25.05,135.07)		(-0.032,0.612,-0.790)	(-0.128,0.642,-0.756)
	法二	(27.81,130.10)	(21.20,140.85)	(-0.033,0.620,-0.784)	(-0.118,0.660,-0.742)
	法三	(24.68,137.05)		(-0.040,0.623,-0.781)	(-0.127,0.654,-0.746)
2	法一	(69.57,144.48)		(0.035,0.629,-0.776)	(-0.059,0.659,-0.750)
	法二	(62.26,134.68)	(63.60,140.85)	(0.026,0.628,-0.778)	(-0.058,0.668,-0.742)
	法三	(68.47,130.94)		(0.031,0.636,-0.771)	(-0.054,0.666,-0.744)
3	法一	(113.81,140.29)		(0.131,0.651,-0.747)	(0.040,0.680,-0.732)
	法二	(110.96,146.37)	(106.00,140.85)	(0.114,0.639,-0.761)	(0.031,0.678,-0.734)
	法三	(111.57,134.12)		(0.133,0.660,-0.739)	(0.050,0.689,-0.723)
4	法一	(150.30,143.10)		(0.308,0.559,-0.770)	(0.212,0.588,-0.780)
	法二	(153.17,147.55)	(148.40,140.85)	(0.299,0.488,-0.820)	(0.207,0.530,-0.822)
	法三	(152.35,149.67)		(0.346,0.502,-0.793)	(0.255,0.533,-0.807)
5	法一	(197.83,148.75)		(0.337,0.646,-0.685)	(0.251,0.672,-0.697)
	法二	(197.40,142.54)	(190.80,140.85)	(0.318,0.618,-0.719)	(0.238,0.655,-0.717)
	法三	(198.24,145.69)		(0.333,0.641,-0.692)	(0.254,0.668,-0.699)
6	法一	(27.37,112.33)		(-0.038,0.577,-0.816)	(-0.136,0.608,-0.782)
	法二	(21.77,109.35)	(21.20,109.55)	(-0.037,0.583,-0.811)	(-0.125,0.625,-0.770)
	法三	(25.62,103.49)		(-0.045,0.585,-0.809)	(-0.135,0.617,-0.775)
7	法一	(68.98,104.18)		(0.051,0.609,-0.792)	(-0.046,0.639,-0.768)
	法二	(61.17,109.17)	(63.60,109.55)	(0.038,0.594,-0.803)	(-0.049,0.635,-0.771)
	法三	(72.62,109.56)		(0.040,0.594,-0.803)	(-0.050,0.626,-0.778)
8	法一	(111.39,107.97)		(0.149,0.608,-0.780)	(0.054,0.638,-0.769)
	法二	(108.60,112.11)	(106.00,109.55)	(0.131,0.590,-0.797)	(0.044,0.631,-0.775)
	法三	(115.72,103.77)		(0.146,0.604,-0.783)	(0.058,0.635,-0.770)

续表 5-5 用户 4 视线估计实验结果

	法一	(147.11,105.91)		(0.254,0.600,-0.758)	(0.161,0.629,-0.761)
9	法二	(146.74,116.77)	(148.40,109.55)	(0.240,0.600,-0.763)	(0.156,0.639,-0.753)
	法三	(153.84,100.26)		(0.240,0.611,-0.755)	(0.154,0.641,-0.752)
	法一	(182.19,104.94)		(0.336,0.605,-0.722)	(0.246,0.632,-0.734)
10	法二	(195.13,119.09)	(190.80,109.55)	(0.314,0.613,-0.725)	(0.233,0.650,-0.723)
	法三	(185.39,122.70)		(0.328,0.634,-0.700)	(0.248,0.662,-0.707)
...					

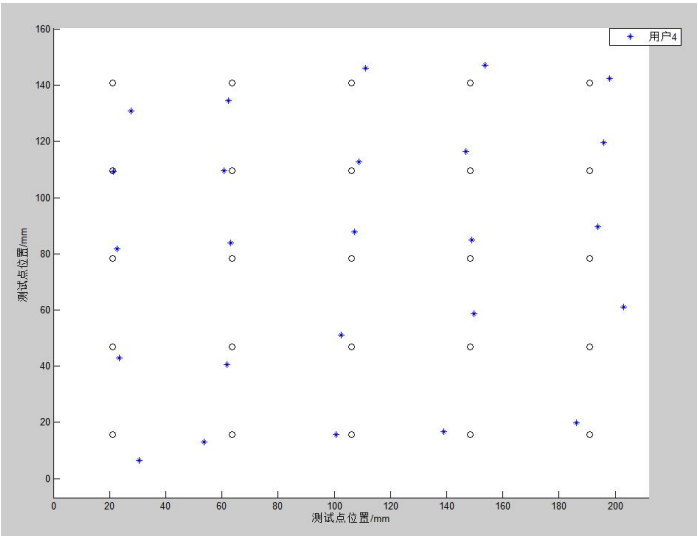


图 5-14 用户 4 圆锥高线法的实际视线落点和标定点偏差图

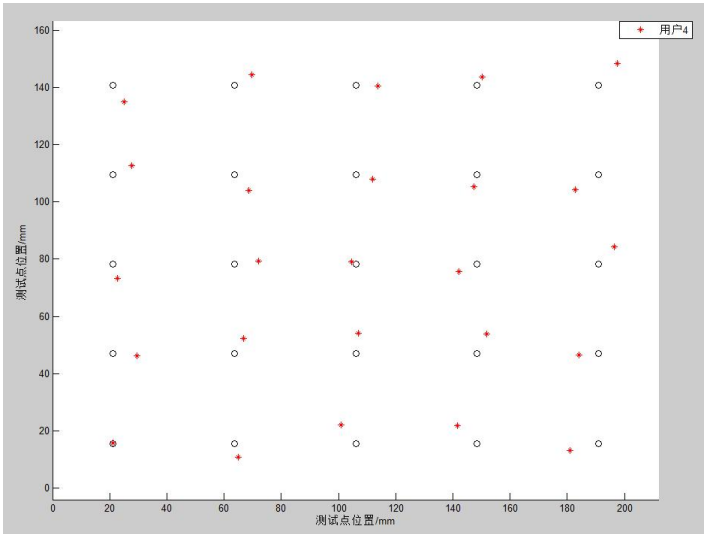


图 5-15 用户 4 椭圆长轴法的实际视线落点和标定点偏差图

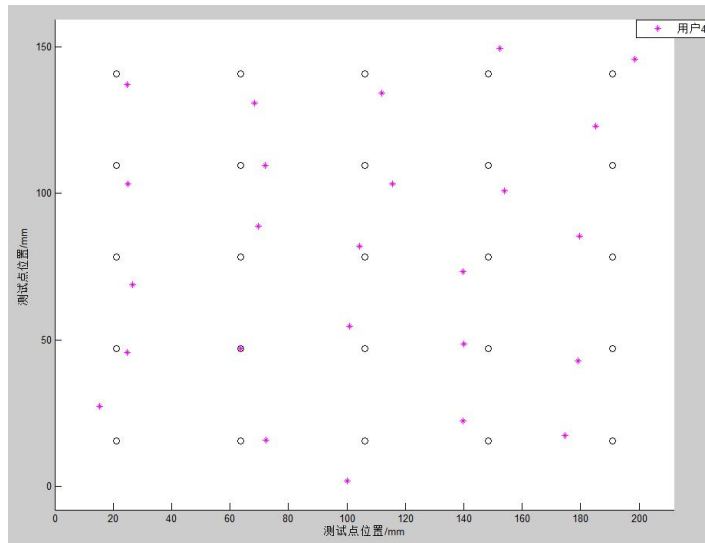


图 5-16 用户 4 长轴交点法的实际视线落点和标定点偏差图

图 5-14、图 5-15、图 5-16 分别显示了用户 4 采用极平面圆锥高线法、极平面椭圆长轴法、极平面长轴交点法的实际视线落点和标定点之间的偏差。为了更直观地显示三种三维视线估计的结果，我们将视线估计的结果与预设的测试点的真实坐标相比较，计算出视线落点与真实注视点之间的偏差，分别计算了三种方法中，八位用户的视线落点在 X 方向和 Y 方向的误差，如表 5-6 和图 5-17 所示。

表 5-6 各用户在 X 轴和 Y 轴方向的误差

方向/°		1	2	3	4	5	6	7	8
法一	X	0.82	0.94	1.20	0.68	1.20	1.30	1.06	1.12
	Y	1.04	1.08	1.25	0.92	0.99	1.40	1.32	1.27
法二	X	0.97	1.08	1.32	0.76	1.26	1.35	1.18	1.26
	Y	1.08	1.31	1.42	1.02	1.03	1.25	1.37	1.47
法三	X	1.06	1.15	1.37	0.84	1.29	1.37	1.20	1.33
	Y	1.13	1.34	1.48	1.05	1.07	1.26	1.39	1.49

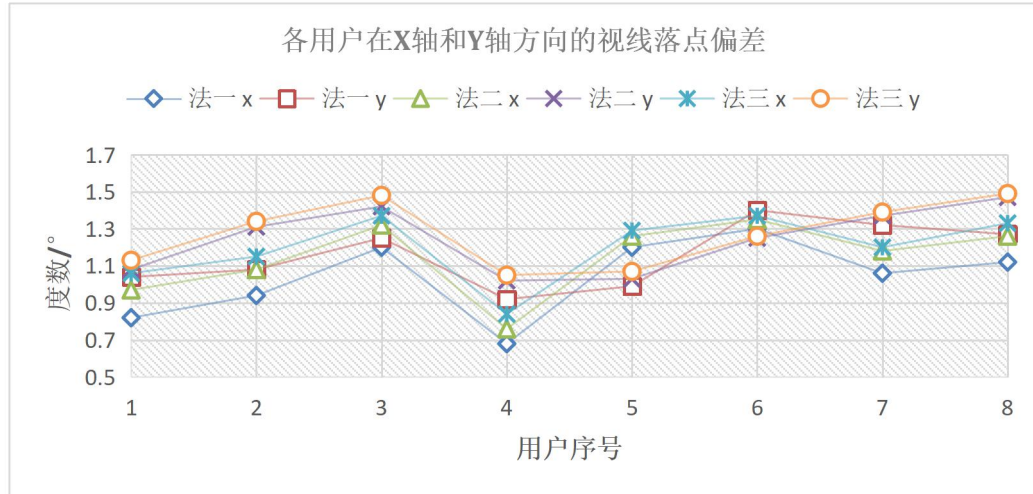


图 5-17 各用户在 X 轴和 Y 轴方向的视线落点偏差

从图 5-17 中各用户的视线估计结果可以看出，当采用圆锥高线法计算光轴方向时，视线落点在 X 方向的最大偏差为 1.30 度，在 X 方向的平均偏差为 1.04 度；在 Y 方向的最大偏差为 1.40 度，在 Y 方向的平均偏差为 1.16 度。采用椭圆长轴法计算光轴方向时，视线落点在 X 方向的最大偏差为 1.35 度，在 X 方向的平均偏差为 1.15 度；在 Y 方向的最大偏差为 1.47 度，在 Y 方向的平均偏差为 1.24 度。采用长轴交点法计算光轴方向时，视线落点在 X 方向的最大偏差为 1.37 度，在 X 方向的平均偏差为 1.20 度；在 Y 方向的最大偏差为 1.49 度，在 Y 方向的平均偏差为 1.28 度。

因此，由上述数据可以看出，圆锥高线法相对于椭圆长轴法和长轴交点法具有更高的眼球光轴重建精度，且光轴数据更加稳定。

5.3.3 与传统三维视线估计方法结果比较

本文方法的提出，主要是针对传统的双相机多光源方法的固有缺点而提出的，基于提取出的瞳孔和普尔钦斑特征，我们采用文献^[39,46,66]中的传统方法进行了三维视线估计，得到的视线估计结果如表 5-7 所示。

表 5-7 不同方法精度比较

方法	X 方向/°	Y 方向/°
文献 ^[39,46,66]	1.12	1.23
极平面圆锥高线法	1.04	1.16
极平面椭圆长轴法	1.15	1.24
极平面长轴交点法	1.20	1.28

从表 5-7 可以看出,本文所提出的极平面圆锥高线法比文献^[39,46,66]中的传统方法精度高。

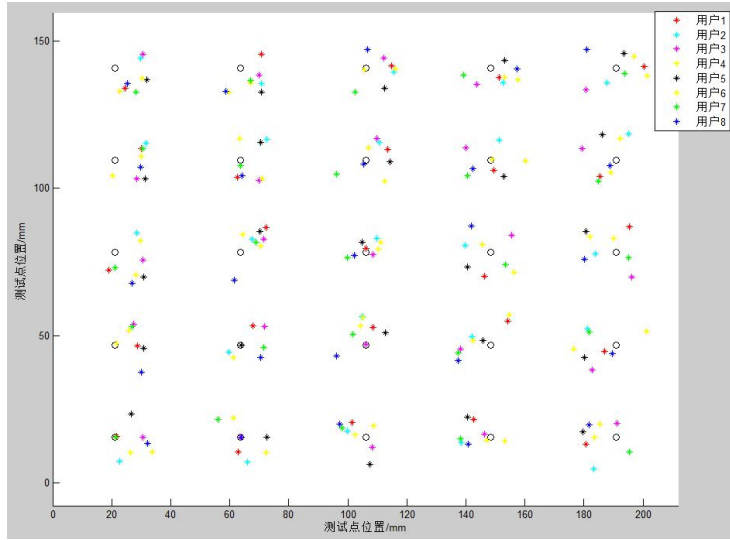


图 5-18 传统方法实际视线落点图

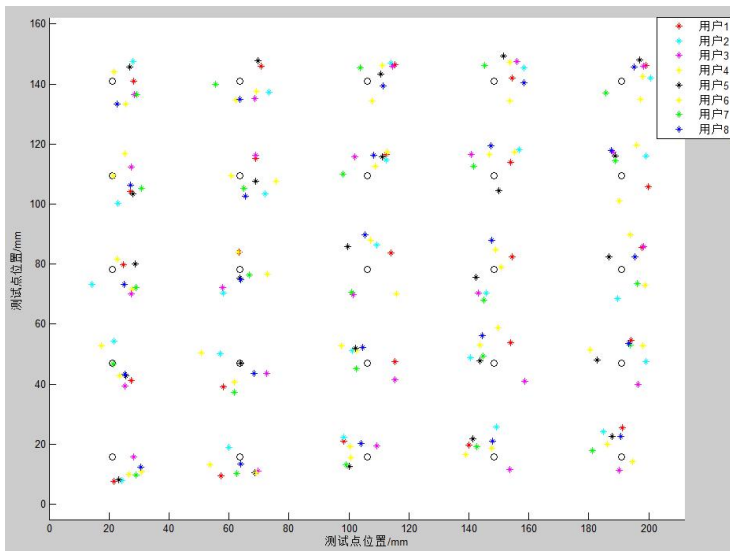


图 5-19 圆锥高线法实际视线落点图

图 5-18 显示了八位用户采用传统方法的实际视线落点和标定点的图,图 5-19 显示了八位用户采用圆锥高线法的实际视线落点和标定点的图。

当我们改变视线斜视角进行实验时,假设左斜视角为负,右斜视角为正,则在斜视范围 -13.2° 到 $+13.2^{\circ}$ (受实验设备限制无法从 -45° 到 $+45^{\circ}$)范围内,根据本文方法和文献^[39,46,66]中的方法计算出的视线落点的误差曲线如图 5-20 所示,可以看出,特别是在视线斜视角度较大(即图像上瞳孔椭圆变形大)的时候,文献^[39,46,66]方法的误差较大,这充分地验证了本文方法是对传统方

法的缺陷的有效改进，证明了本文方法的有效性。

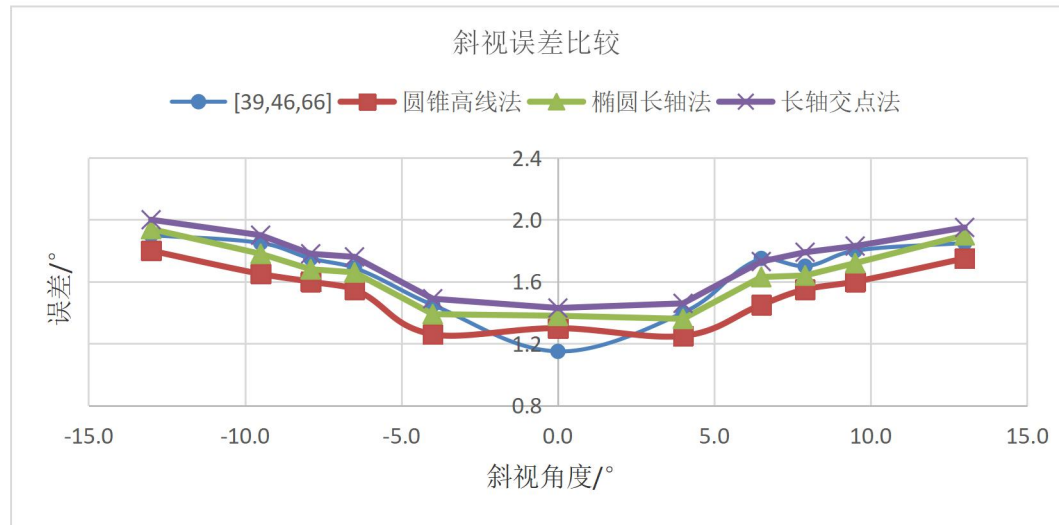
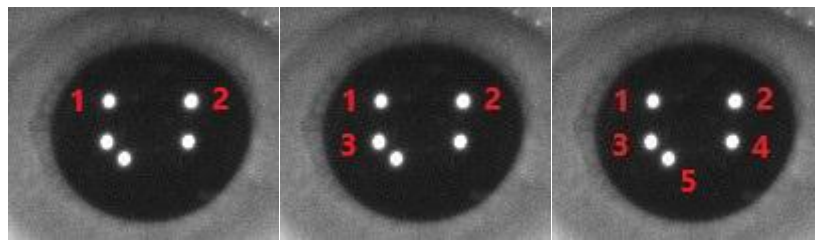


图 5-20 不同方法斜视误差比较

5.3.4 光源数量对视线估计精度的影响

以上结果均基于双相机双光源系统，接下来对光源数量的影响进行讨论。由 4.2 节可以知道，确定角膜曲率中心 C 的三维位置至少需要两个光源，当然也可以使用更多光源。如图 5-21 所示，以下将对使用双光源、三个光源以及五个光源得到的视线估计结果进行比较。



(1) 两个普尔钦斑 (2) 三个普尔钦斑 (3) 五个普尔钦斑

图 5-21 普尔钦斑图像

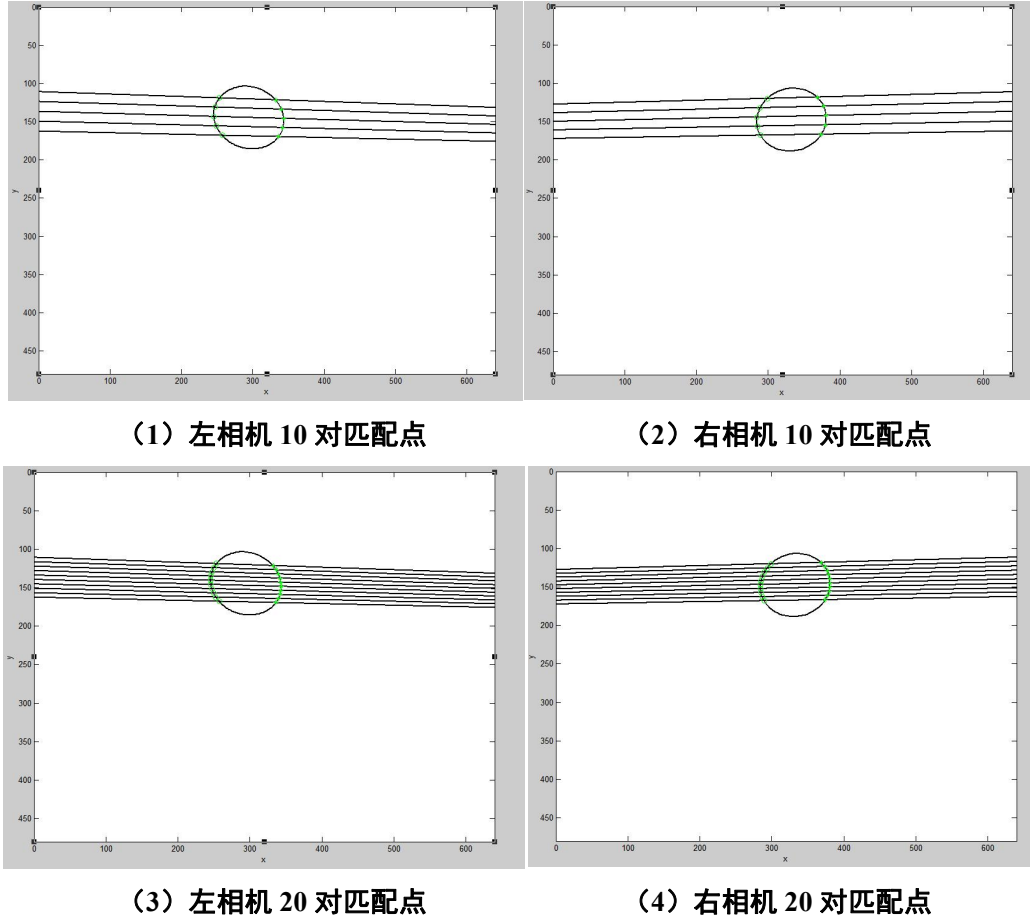
参与实验的八个用户采用圆锥高线法的误差如表 5-8 所示。对于每个用户，针对不同数量的光源给出水平 (X) 方向和垂直 (Y) 方向的误差。由表 5-8 可以看出，使用更多的光源点可以产生更低的平均误差，光源数量的增加通过平均化减少了模型的误差。但在某些情况下（如用户 4），两个光源就可以达到较好的实验效果。如果正确确定了普尔钦斑的坐标，则具有两个光源的系统足以进行视线估计。

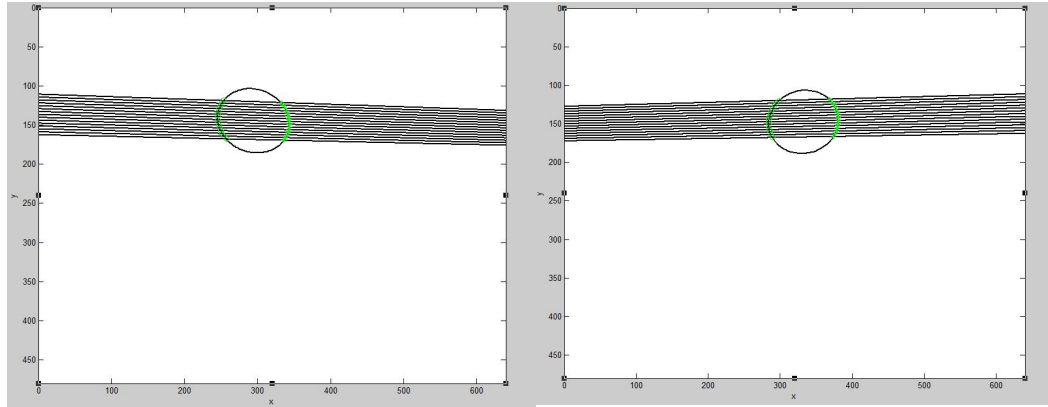
表 5-8 不同数量光源在 x 方向和 y 方向的精度

	两个光源		三个光源		五个光源	
用户	X	Y	X	Y	X	Y
用户 1	0.82°	1.04°	0.80°	1.07°	0.79°	1.07°
用户 2	0.94°	1.08°	0.92°	1.11°	0.87°	1.10°
用户 3	1.20°	1.25°	1.17°	1.26°	1.16°	1.27°
用户 4	0.68°	0.92°	0.67°	0.93°	0.67°	0.93°
用户 5	1.20°	0.99°	1.20°	1.02°	1.16°	1.02°
用户 6	1.30°	1.40°	1.28°	1.40°	1.25°	1.39°
用户 7	1.06°	1.32°	1.04°	1.33°	1.03°	1.33°
用户 8	1.12°	1.27°	1.11°	1.29°	1.10°	1.28°
均值	1.04°	1.16°	1.02°	1.18°	1.00°	1.17°

5.3.5 匹配点数量对视线估计精度的影响

以上极平面圆锥高线法均选用 20 对匹配点,接下来针对匹配点数量对极平面圆锥高线法精度的影响进行讨论。如图 5-22 所示,对使用多对匹配点得到的视线估计结果进行比较。





(5) 左相机 30 对匹配点

(6) 右相机 30 对匹配点

图 5-22 多对匹配点示意图

八个用户采用不同对匹配点的极平面圆锥高线法误差如表 5-9 所示。对于每个用户，针对不同对匹配点给出水平（X）方向和垂直（Y）方向的误差。由表 5-9 可以看出，使用更多的匹配点可以产生更低的平均误差，相比较而言，10 对匹配点并不能产生很好的效果，30 对匹配点的效果略好于 20 对匹配点。但在某些情况下（如用户 1），20 对匹配点就可以达到较好的实验效果。在实际实验中，尝试了多对匹配点，由于匹配点运算量较大，最终选用 20 对匹配点进行视线估计就可以得到较好的结果。

表 5-9 不同对匹配点在 X 方向和 Y 方向的精度

匹配点数量	10 对匹配点		20 对匹配点		30 对匹配点	
用户	X	Y	X	Y	X	Y
用户 1	0.94°	1.12°	0.82°	1.04°	0.82°	1.04°
用户 2	1.06°	1.20°	0.94°	1.08°	0.91°	1.08°
用户 3	1.28°	1.33°	1.20°	1.25°	1.19°	1.24°
用户 4	0.80°	1.06°	0.68°	0.92°	0.67°	0.91°
用户 5	1.27°	1.11°	1.20°	0.99°	1.17°	0.99°
用户 6	1.34°	1.47°	1.30°	1.40°	1.28°	1.39°
用户 7	1.15°	1.39°	1.06°	1.32°	1.05°	1.31°
用户 8	1.22°	1.33°	1.12°	1.27°	1.10°	1.27°
均值	1.13°	1.25°	1.04°	1.16°	1.02°	1.15°

5.4 本章小结

本章首先进行了仿真实验，从仿真实验结果可以看出，本文方法具有可

行性，从原理上可以克服传统方法的缺陷，重建更准确的光轴。但光源的偏差对眼球光轴重建的影响较大，因此对于光源标定的精度要求较高。

在真实实验中，我们分别采取了三种光轴重建方法：极平面圆锥高线法、极平面椭圆长轴法和极平面长轴交点法，从原理上说，三种光源重建方法的本质是一样的，都是可行的。但是从实验结果看出，只有圆锥高线法具有较高的三维视线估计精度，能够克服传统方法的不足，特别是在眼球的大角度斜视的情况下，圆锥高线法的精度明显高于传统方法。这证明了采用圆锥高线法重建光轴能够克服传统方法的固有缺陷，也证明了本文改进传统方法的出发点是正确的。但圆锥高线法的运算量比较大，缺乏实用价值。椭圆长轴法运算简单，具有实际应用价值，但是在实验中精度低于传统方法，原因可能是，视线估计的精度十分依赖于瞳孔椭圆拟合的准确性，即图像上瞳孔椭圆长轴的方位和角度，这很容易造成随机误差的干扰。长轴交点法是最具有实用价值的眼球光轴重建方法，但很遗憾，在我们的实验中，这种方法的精度也低于传统方法，其原因应该是多方面的，除了瞳孔椭圆的拟合精度影响以外，双目相机系统的结构不合理等因素也严重地影响了长轴交点法的精度，如双相机的光轴夹角太小，导致瞳孔在双目相机中的成像非常近似，椭圆长轴交点的误差被放大；双目相机的基线距离不够远所导致的测量误差等。

在本文方法中，还进行了光源数量以及匹配点数量对视线估计影响的实验。结果表明，如果准确获得了普尔钦斑的坐标，则具有两个光源的系统足以进行视线估计。瞳孔边缘匹配点的数量对光轴重建的影响较大，因此实际应用中需要确定合理的匹配点数量。

6 结论

视线追踪技术作为一种新型的人机交互方式，广泛应用于虚拟现实、车辆辅助驾驶和军事技术等领域，具有非常广阔的前景。本文对多相机多光源系统三维视线估计方法进行了研究。

本文主要研究工作如下：

1) 本文在三维视线估计模型的研究方面做了大量工作，针对传统多相机多光源系统三维视线估计方法的不足，提出了一种基于瞳孔边缘匹配的双相机双光源系统三维视线估计方法。与传统的三维视线估计方法相比，本文提出的方法是通过两个相机图像中的瞳孔边缘匹配点构造的折射平面来获得光轴，克服了传统方法采用图像瞳孔中心来重建眼球光轴的固有误差。仿真实验证明了本文方法的有效性。

2) 在光轴估计中，提出了三种方法：极平面圆锥高线法、极平面椭圆长轴法以及极平面长轴交点法。在实验中，极平面圆锥高线法的精度高于传统方法，这也验证了传统方法的不足和本文方法的有效性。在本文方法中，还进行了光源数量对视线估计影响的实验。结果表明，如果准确获得了普尔钦斑的坐标，则具有两个光源的系统足以进行视线估计。

由于时间和相关设备的原因，本文工作还不够完善，有所欠缺，需要在以下方面作进一步工作：

1) 理论上讲，光轴估计的三种方法都可以提高视线估计精度，但实验结果表明，只有极平面圆锥高线法的精度高于传统方法，其它两种方法容易受到各种误差因素的干扰。因此，需要提高计算过程中用到的信息数据的精度。

2) 对于模型来说，为了进一步提高其鲁棒性，需要加入更多的匹配点进行计算，但是圆锥高线法重建光轴使用的匹配点数量已经较多，计算用时较大，无法满足实时性要求。并且目前采用的确定匹配点的方法，对匹配点的精度要求较高。因此，能否准确地确定瞳孔边缘的匹配点，尽可能使用较少的高精度的匹配点来达到光轴重建的要求，是本文的下一步工作。

综上所述，本文的实验结果很好地验证了本文工作的出发点的合理性和本文方法的有效性，但实验结果还不够满意，未来需要在模型和系统优化等方面作进一步的研究。

参考文献

- [1] 董士海. 人机交互的进展及面临的挑战[J]. 计算机辅助设计与图形学学报, 2004, 16(1): 1-13.
- [2] Jacob R J K, Karn K S. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises[M]. The Mind's Eye, 2003: 573-605.
- [3] Coetzer R C, Hancke G P. Development of a robust active infrared-based eye tracker[J]. IET Computer Vision, 2014, 8(6): 523-534.
- [4] Jacob R J K. The use of eye movements in human-computer interaction techniques: what you look at is what you get[J]. ACM Transactions on Information Systems (TOIS), 1991, 9(2): 152-169.
- [5] Model D, Eizenman M. User-calibration-free remote eye-gaze tracking system with extended tracking range[C]. 24th Canadian Conference on IEEE(CCECE), 2011: 1268-1271.
- [6] Morimoto C H, Koons D, Amir A, et al. Pupil detection and tracking using multiple light sources[J]. Image&Vision Computing, 2000, 18(4): 331-335.
- [7] Guestrin E D, Eizenman M. General theory of remote gaze estimation using the pupil center and corneal reflections[J]. IEEE Transactions on Bio-medical Engineering, 2006, 53(6): 1124-1133.
- [8] Witzner H D, Qiang J. In the eye of the beholder: a survey of models for eyes and gaze[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2010, 32(3): 478-500.
- [9] Zhu Z, Ji Q. Eye and gaze tracking for interactive graphic display[J]. Machine Vision and Applications, 2004, 15(3): 139-148.
- [10] Zhu Z, Ji Q, Bennett K P. Nonlinear eye gaze mapping function estimation via support vector regression[C]. 18th International Conference on Pattern Recognition (ICPR'06). IEEE, 2006, 1: 1132-1135.
- [11] Ohno T, Mukawa N, Kawato S. Just blink your eyes: A head-free gaze tracking system[C]. CHI'03 Extended Abstracts on Human Factors in Computing Systems. ACM, 2003: 950-957.
- [12] Beymer D, Flickner M. Eye gaze tracking using an active stereo head[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2003: 443-451.
- [13] 李娅萍. 基于瞳孔角膜反射法的视线追踪技术研究[D]. 重庆: 重庆邮电大学, 2016.

- [14] 张名魁, 孙复川. 高抗干扰性的红外光电反射眼动测量仪[J]. 生物医学工程学杂志, 1989(4): 282-286.
- [15] 武媛媛. 眼动跟踪技术研究[D]. 西安: 西安电子科技大学, 2009.
- [16] 秦华标, 严伟洪, 王信亮, 余翔宇. 一种可克服头动影响的视线追踪系统. 电子学报, 2013, 41(12): 2403-2409.
- [17] 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[J]. 通信学报, 2012(12): 133-139.
- [18] 熊春水, 黄磊, 刘昌平, 等. 一种新的单点标定视线估计方法[J]. Acta Automatica Sinica, 2014, 40(3): 459-470.
- [19] Alexander Toet. Gaze directed displays as an enabling technology for attention aware systems[J]. Computers in Human Behavior, 2006, 22: 615-647.
- [20] Morimoto C H, Koons D, Amir A, et al. Pupil detection and tracking using multipule light sources[J]. Image&Vision Computing, 2000, 18(4): 331-335.
- [21] Duchowski A T. Eye Tracking Methodology: Theory and Practice[M]. Berlin: Springer London, 2003: 131-132.
- [22] Colombo C, Andronico S, Dario P. Prototype of a vision-based gaze-driven man-machine interface[C]. IEEE/RSJ International Conference on Intelligent Robots & Systems 95 Human Robot Interaction & Cooperative Robots. IEEE, 1995: 188-192.
- [23] Remmel R S. An inexpensive eye movement monitor using the scleral search coil technique[J]. IEEE Transactions on Biomedical Engineering, 2007, BME-31(4): 388-390.
- [24] Amir A, Zimet L, Sangiovanni-Vincentelli A, et al. An embedded system for an eye-detection sensor[J]. Computer Vision and Image Understanding, 2005, 98(1): 104-123.
- [25] Cornsweet T N, Crane H D. Accurate two-dimensional eye tracker using first and fourth Purkinje images[J]. Journal of the Optical Society of America, 1973, 63(8): 921-928.
- [26] Young L R, Sheena D. Survey of eye movement recording methods[J]. Behavior Research Methods & Instrumentation, 1975, 7(5): 397-429.
- [27] Sliney D, Aronrosa D, Delori F, et al. Adjustment of guidelines for exposure of the eye to optical radiation from ocular instruments: statement from a task group of the International Commission on Non-Ionizing Radiation Protection (ICNIRP)[J]. Applied Optics, 2005, 44(11): 2162-2178.
- [28] Morimoto C H, Mimica M R M. Eye gaze tracking techniques for interactive applications[J]. Computer Vision and Image Understanding,

- 2005, 98(1): 4-24.
- [29] Sigut J, Sidha S A. Iris center corneal reflection method for gaze tracking using visible light[J]. IEEE Transactions on Biomedical Engineering, 2011, 58(2): 411-419.
- [30] Zhang C , Chi J N , Zhang Z , et al. Gaze estimation in a gaze tracking system[J]. Science China(Information Sciences), 2011, 54(11): 2295-2306.
- [31] Zhu Z, Ji Q. Eye gaze tracking under natural head movements[C]. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, 1: 918-923.
- [32] Kim S C, Sked M, Ji Q. Non-intrusive eye gaze tracking under natural head movements[C]. The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2004, 1: 2271-2274.
- [33] K.P. White Jr., T.E. Hutchinson, and J.M. Carley, Spatially Dynamic Calibration of an Eye-Tracking System[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1993 , 23(4): 1162-1168.
- [34] Kang Jeffrey J, Eizenman Moshe, Guestrin Elias D, Eizenman Erez. Investigation of the Cross-Ratios Method for Point-of-Gaze Estimation[J]. IEEE TRANSACTIONS ON BIO-MEDICAL ENGINEERING, 2008, 55(9): 2293-2302.
- [35] Coutinho F.L., Morimoto C.H. Free head motion eye gaze tracking using a single camera and multiple light sources[C]. 19th Brazilian Symposium on Computer Graphics and Image Processing, 2006: 171-178.
- [36] Dong H Y, Kim J H, Bang R L, et al. Non-contact eye gaze tracking system by mapping of corneal reflections[C]. Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, 2002: 101-106.
- [37] Yoo D H, Chung M J. A novel non-intrusive eye gaze estimation using cross-ratio under large head motion[J]. Computer Vision and Image Understanding, 2005, 98(1): 25-51.
- [38] Piratla N M, Jayasumana A P. A neural network based real-time gaze tracker[J]. Journal of Network and Computer Applications, 2002, 25(3): 179-196.
- [39] Sheng-Wen Shih, Jin Liu. A Novel Approach to 3D Gaze Tracking Using Stereo Cameras[J]. IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS-PART B: CYBERNETICS, 2004, 34(1): 234-245.
- [40] Xie X, Sudhakar R, Zhuang H. A cascaded scheme for eye tracking and head movement compensation[J]. IEEE Transactions on System Man Cybern Part A-System Human, 1998, 28: 487-490.

- [41] J. Wang, E. Sung, and R. Venkateswarlu. Eye gaze estimation from a single image of one eye[C]. Proceedings of International Conference of Computer Vision, 2003, 136-143.
- [42] Jonathon B. Hiley, Andrew H. Redekopp and Reza Fazel-Rezai. A Low Cost Human Computer Interfaze based on Eye Tracking[C]. Proceedings of 28th IEEE EMBS Annual International Conference, 2006: 3226-3230.
- [43] Dong Hyun Yoo and Myung Jin Chung. Eye-mouse under Large Head Movement for Human-Compter Interface[C]. Proceedings of the 2004 IEEE international conference on robotics and automation, 2004, 237-242.
- [44] H Morimoto, A. Amir, M. Flickner. Detecting eye position and gaze from a single camera and 2 light sources[C]. Object recognition supported by user interaction for service robots, 2002, 4: 314-317.
- [45] C. Hennessey, B. Nouredin, and P. Lawrence. A Single Camera Eye-Gaze Tracking System with Free Head Motion[C]. Processing of Symposium. Eye Tracking Research and Applications, 2006: 87-94.
- [46] E. D. Guestrin and M. Eizenmen. General theory of remote gaze estimation using the pupil center and corneas reflections[J]. IEEE Transaction on Bio-medical Engineering, 2006, 53(6): 1124-1133.
- [47] ED Guestrin , M Eizenman. Remote point-of-gaze estimation free head movements requiring a single-point calibration[C]. International Conference of the IEEE Engineering in Medicine & Biology Society, 2007: 4556-4560.
- [48] Hennessey C, Nouredin B, Lawrence P. Fixation precision in high-speed noncontact eye-gaze tracking[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2008, 38(2):289-98.
- [49] Zhang K, Zhao X, Ma Z, et al. A Simplified 3D Gaze Tracking Technology with Stereo Vision[C]. 2010 International Conference on Optoelectronics and Image Processing, 2010: 131-134.
- [50] David Beymer, Myron Flickner. Eye Gaze Tracking Using an Active Stereo Head[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003, 443-451.
- [51] T Ohno, N. Mukawa, A Yoshikawa. Freegaze: A gaze tracking system for everyday gaze interaction[J]. In Proceedings of Symposium. ETRA. New Orleans, USA, ACM Press: 2002.125-132.
- [52] Craig Hennessey and Peter Lawrence. Noncontact Binocular Eye-Gaze Tracking for Point-of-gaze Estimation in Three Dimensions[J]. IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, 2009, 56(3): 790-800.
- [53] Nicole M. Bakker, Boris A. J. Lenseigne, Sander Schutte, Elsbeth B. M.

- Geukers, Pieter P. Jonker, Frans C. T. van der Helm, and Huibert J. Simonsz. Accurate Gaze Direction Measurements with Free Head Movement for Strabismus Angle Estimation[J]. IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, 2013, 60(11): 3028-3036.
- [54] Wei Zhengzhong, Zhang Guangjun. A Distortion Error Model of the Perspective Projection of Ellipse Center and Its Simulation[J]. Chinese Journal of Science Instrument, 2003, 24(2): 160-164.
- [55] Jose Sigut and Sid-Ahmed Sidha. Iris Center Corneal Reflection Method for Gaze Tracking Using Visible Light[J]. IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, 2011, 58(2): 411-421.
- [56] Xing Dequan, Da Feipeng, Zhang Hu. Research and application of location of circular target with high accuracy[J]. Chinese Journal of Science Instrument, 2009, 30(12): 2593-2598.
- [57] Tsai R. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses[J]. IEEE Journal on Robotics and Automation, 1987, 3(4): 323-344.
- [58] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22: 1330-1334.
- [59] Fraser C S. Photogrammetric camera component calibration: A review of analytical techniques[M]. Calibration and Orientation of Cameras in Computer Vision. Springer, Berlin, Heidelberg, 2001: 95-121.
- [60] Fraser C S. Digital camera self-calibration[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 1997, 52(4): 149-159.
- [61] 黄桂平. 数字近景工业摄影测量关键技术研究与应用[D]. 天津: 天津大学, 2005.
- [62] 张飞. 相机标定方法综述[C]. 全国通信安全学术会议. 2011: 36-39.
- [63] 金欢. 视线跟踪系统中的标定技术和瞳孔定位问题研究[D]. 合肥: 中国科学技术大学, 2013.
- [64] 张琼, 王志良, 迟健男, 等. 基于平面镜面向双摄像机视线追踪系统的标定方法[J]. 光学学报, 2011, 31(4): 192-198.
- [65] Shiu Y C, Ahmad S. 3D location of circular and spherical features by monocular model-based vision[C]. IEEE International Conference on Systems, Man and Cybernetics, 1989, 2: 576-581.
- [66] Arantxa Villanueva and Rafael Cabeza. A Novel Gaze Estimation System with One Calibration Point [J]. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, 2008, 28(4): 1123-1139.

作者简历及在学研究成果

一、 作者入学前简历

起止年月	学习或工作单位	备注
2013 年 09 月至 2017 年 06 月	在北京科技大学测控技术与仪器专业 攻读学士学位	

二、 在学期间发表的论文

[1] 王丰, 张国胜, 迟英凯, 迟健男. 一种新的多相机多光源三维视线估计方法[C]. 第九届全国可穿戴计算学术会议(CWCC 2019). 已录用

独创性说明

本人郑重声明：所呈交的论文是我个人在导师指导下进行的研究工作及取得研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京科技大学或其他教育机构的学位或证书所使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中做了明确的说明并表示了谢意。

签名：_____ 日期：_____

关于论文使用授权的说明

本人完全了解北京科技大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

（保密的论文在解密后应遵循此规定）

签名：_____ 导师签名：_____ 日期：_____

学位论文数据集

关键词*	密级*	中图分类号*	UDC	论文资助
三维视线估计， 匹配点，光轴， 视轴		TN911.73		
学位授予单位名称*		学位授予单位 代码*	学位类别*	学位级别*
北京科技大学		10008	工学	硕士
论文题名*		并列题名		论文语种*
多相机多光源系统三维视线估计 方法研究				中文
作者姓名*	王丰		学号*	S20170567
培养单位名称*		培养单位代码*	培养单位地址	邮编
北京科技大学		10008	北京市海淀区 学院路 30 号	100083
学科专业*		研究方向*	学制*	学位授予年*
仪器科学与技术		计算机视觉	2.5 年制	2020
论文提交日期*	2019 年 12 月 18 日			
导师姓名*	迟健男		职称*	副教授
评阅人	答辩委员会主席*		答辩委员会成员	
	蓝金辉			
电子版论文提交格式 文本 (<input checked="" type="checkbox"/>) 图像 (<input type="checkbox"/>) 视频 (<input type="checkbox"/>) 音频 (<input type="checkbox"/>) 多媒体 (<input type="checkbox"/>) 其他 (<input type="checkbox"/>) 推荐格式: application/msword; application/pdf				
电子版论文出版 (发布) 者		电子版论文出版 (发布) 地		权限声明
论文总页数*	71			
共 33 项，其中带*为必填数据，为 22 项。				

