

眼动跟踪仪的光源布局优化方法

徐泽夫^{1,2}, 王晶晶¹, 曾冠中¹, 谢迪^{1*}, 冯华君²

(1. 杭州海康威视数字技术股份有限公司研究院, 浙江 杭州 310052;
2. 浙江大学光电科学与工程学院, 浙江 杭州 310027)

摘要: 使用角膜瞳孔反射法的眼动跟踪仪相比基于表观的方法具有较高的视线估计精度, 在人机交互、AR/VR 等领域具有较大的实际应用价值。但是角膜瞳孔反射法受制于角膜非球面的影响, 需要光源距离较近, 从而限制普尔钦斑处于逼近球面的角膜区域。但光源距离太近又会放大计算的误差, 所以针对眼动跟踪仪的硬件结构缺乏量化设计指导的问题, 提出了一种光源布局优化方法。通过计算机仿真的手段, 建立了角膜的非球面几何模型, 理论分析了不同光源布局下的平均视线估计精度, 然后通过由粗到精的方式, 优化求解得到最优光源布局, 抑制了非球面角膜带来的影响。最后以两类典型的设备为例, 在典型参数下分别讨论了头戴式和遥测式的最优双光源布局, 结果可作为典型硬件设计的理论参考。

关键词: 视线估计; 光源布局; 瞳孔角膜反射法; 非球面角膜

中图分类号: TP391.9 **文献标志码:** A **DOI:** 10.3788/IRLA20240209

0 引言

视线追踪技术是一种追踪人眼视觉方向、眼视觉注意力的技术, 被广泛应用于人机交互^[1-2]、虚拟现实^[3]、疲劳驾驶监控、心理分析^[4]等领域。现有的视线跟踪技术从原理上可分为两大类, 一种为基于表观的视线估计技术, 另一种为基于瞳孔角膜反射 (Pupil Center Corneal Reflection, PCCR) 法的视线估计技术。

基于表观的视线估计技术无需特定硬件, 仅使用单个相机即可实现视线估计^[5-8]。目前利用多模态学习以提高视线估计的泛化能力也是研究热点之一, 该方法适用的视线范围较广, 但是该类方法的视线精度相对 PCCR 法较低, 一般被应用于智能车舱的驾驶员监控系统 (Driver Monitor System, DMS), 通过车载相机拍摄驾驶员面部图像, 估计驾驶员的注视方向及注视区域, 从而判断驾驶员是否分心或疲劳。

基于 PCCR 法的视线估计技术则需要特定的硬件, 采用近红外光源与近红外相机集成为眼动仪类产品。利用红外光源在角膜表面的反射斑点 (普尔钦斑) 和瞳孔边缘分别计算出角膜中心和视线矢量, 从而较为准确地计算得到使用者当前注视的区域。该

类方法适用的视线范围小, 但是视线精度高, 一般被应用于人机交互、虚拟现实产品中。对于基于瞳孔角膜反射法的视线估计技术, 可分为 2D 回归的视线追踪方法^[9]和基于 3D 模型的视线追踪方法, 基于 3D 模型的视线追踪方法允许头部运动, 目前是 PCCR 法中的主要研究领域。现有的研究集中在视线的光轴重建方法、硬件配置、眼部特征的提取方法方面。尽管已有的研究已经对单相机单光源^[10-11]、单相机多光源^[12]、多相机多光源^[13]等硬件结构进行了详尽的讨论, 但是对于光源布局这种更为量化的硬件优化还缺乏研究。与此同时, PCCR 法把角膜认为是球面, 制约了该方法的精度上限, 其实当光源距离较远时, 普尔钦斑处于角膜边缘, 不符合球面模型。目前的主要抑制手段是在眼动跟踪仪中, 控制普尔钦斑处于角膜中心区域, 较为逼近球面模型, 但光源之间距离太近, 则又会产生较大的计算误差。因此, 对眼动跟踪仪的光源布局进行优化具有很强的应用意义。在仿真分析方面, 已有研究^[14]通过仿真手段分析了瞳孔透视变形对光轴重建的影响, 但其只将角膜考虑为球面, 未考虑角膜的非球面影响因素。

为了解决光源布局优化的问题, 文中建立了非球

收稿日期: 2024-05-16; 修订日期: 2024-07-22

作者简介: 徐泽夫, 男, 博士, 主要从事视线估计、成像光学建模方面的研究。

通讯作者: 谢迪, 男, 高级工程师, 博士, 主要从事计算机视觉、深度网络优化方面的研究。

面角膜模型,通过数值计算的方式得到不同光源布局下的视线估计精度水平,并构建了评价函数,通过评价函数优化即可得到最优光源布局。为眼动仪设计及 VR 设备的视线追踪硬件设计提供了理论依据和技术途径。文中的贡献有以下三方面:

第一,提出了一种适用于 PCCR 视线估计的由粗到精的光源布局优化方法,从而得到量化的最优光源布局。

第二,考虑了角膜非球面对光轴重建的影响,更符合实际人眼角膜,而已有研究的仿真工作大多将角膜作为球面处理。

第三,对典型参数的最优光源布局进行了详细讨论,分析的结果可作为相关设备的理论参考。

1 方 法

为了量化不同光源布局下的视线精度,文中建立了一套视线仿真系统,模拟了视线估计过程中的成像及光轴重建步骤。通过计算不同光源布局的理论视线估计精度,得到最优的光源布局。整个方法流程可分为眼球成像几何模型、视线估计光轴重建算法及评价函数构建三部分,如图 1 所示。在不同的光源设置下,可以计算得到不同光源布局的误差统计结果,由于误差函数不一定是凸函数,因此文中采用了由粗到

精的求解方法。先计算多组离散位置得到最优的位置,然后在该位置附近通过最优化迭代求解得到该位置附近的局部极值,该位置附近的局部极值可作为更为精细的最优光源布局。

1.1 眼球成像几何模型

1.1.1 眼球建模

人眼包含角膜、晶状体、虹膜、巩膜、瞳孔、房水、视网膜等部位,如图 2 所示。在视线估计中,与视线估计相关的部位主要是角膜、瞳孔,有些方法也利用了虹膜^[15]。

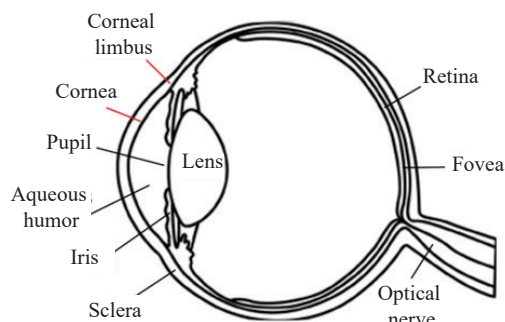


图 2 人眼组成示意图^[16]

Fig.2 Schematic diagram of human eye^[16]

对于瞳孔,其处于眼球内部,几何形状为圆形。对于角膜,现有的建模方式普遍近似为球面,但人眼角膜并不是理想球面,只有角膜中心部分可以近似为球面。为了表征非球面角膜,现有文献^[17]采用绕光轴旋转对称的椭球面来表征,其数学表达式为:

$$\frac{z_{eye}^2}{a^2} + \frac{x_{eye}^2 + y_{eye}^2}{b^2} = 1 \quad (1)$$

式中: $(x_{eye}, y_{eye}, z_{eye})$ 为椭球表面三维坐标,沿着光轴方向为 z 方向,水平方向为 X 方向,垂直方向为 Y 方向,坐标系原点为角膜中心; a 为椭球 Z 方向的轴长; b 为椭球 X/Y 方向的轴长。

进一步的,考虑角膜的非旋转对称结构,例如使用者的散光,则 XOY 平面内的角膜边缘不是圆形,用公式 (2) 来表征^[18]:

$$\frac{z_{eye}^2}{a^2} + \frac{\gamma_1 \cdot x_{eye}^2 + \gamma_2 \cdot x_{eye} y_{eye} + \gamma_3 \cdot y_{eye}^2}{b^2} = 1 \quad (2)$$

其中, $\gamma_1, \gamma_2, \gamma_3$ 由公式 (3) 计算:

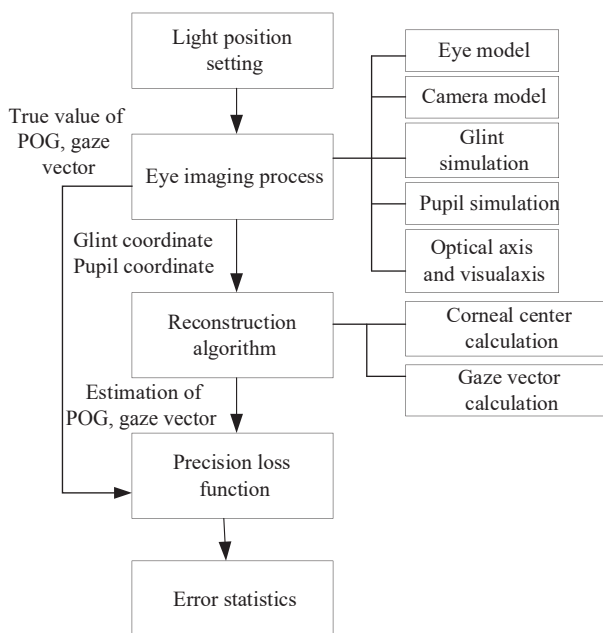


图 1 光源布局优化流程图

Fig.1 Flow chart for optimizing light source layout

$$\begin{aligned}\gamma_1 &= \frac{\cos^2 \delta}{\xi} + \xi \sin^2 \delta \\ \gamma_2 &= 2 \cos \delta \sin \delta \left(\frac{1}{\xi} - \xi \right) \\ \gamma_3 &= \frac{\sin^2 \delta}{\xi} + \xi \cos^2 \delta\end{aligned}\quad (3)$$

式中: ξ 为 XOY 平面内椭圆长轴与短轴之比, 即 $\xi = a_{cs}/b_{cs}$; δ 为 XOY 平面内椭圆长轴与水平 X 轴的夹角, 如图 3 所示。

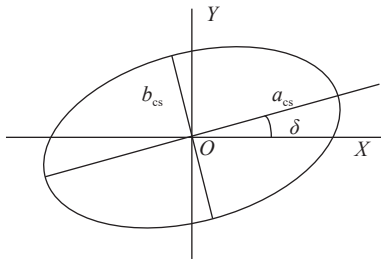

 图 3 XOY 平面内的椭圆参数

 Fig.3 Elliptic parameters in the XOY plane

1.1.2 相机成像建模

理想相机成像过程符合针孔模型, 将角膜坐标系的坐标经过坐标转换得到相机坐标系下的三维坐标, 再通过透视变换即可得到图像坐标系下的二维坐标。该过程表示为:

$$\begin{aligned}s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= \begin{bmatrix} \frac{1}{dX} & 0 & u_0 \\ 0 & \frac{1}{dY} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \\ \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} x_{eye} \\ y_{eye} \\ z_{eye} \\ 1 \end{bmatrix} &\end{aligned}\quad (4)$$

式中: (u, v) 为图像中的像素点; $\frac{1}{dX}, \frac{1}{dY}$ 分别为 u 方向和 v 方向的缩放倍数; u_0, v_0 分别为 u 方向和 v 方向的平移量; f 为相机焦距; $\mathbf{R}$ 和 $\mathbf{t}$ 分别为眼球坐标系转至相机坐标系的旋转矩阵与平移矩阵。由于实际相机镜头并非理想镜头, 且存在畸变的影响, 使得空间点所成的像并不在线性模型所描述的位置, 而是受到镜头失真影响而偏移的实际平面坐标, 一般采用多项式来表征这种畸变的映射^[19], 如公式 (5) 所示:

$$\begin{cases} x' = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 xy + p_2(r^2 + 2x^2) \\ y' = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2y^2) + 2p_2 xy \\ r^2 = x^2 + y^2 \end{cases}\quad (5)$$

式中: (x', y') 为畸变后的图像坐标系坐标; (x, y) 为畸变前的图像坐标系坐标。 x, y 与 u, v 的关系为 $x = (u - u_0)/f, y = (v - v_0)/f$ 。

1.1.3 普尔钦斑位置计算

对于近红外光线, 角膜可以近似认为是全反射镜, 普尔钦是近红外光源在角膜上的反射光斑。普尔钦斑的位置与光源、相机光心满足反射定理, 即入射光线、出射光线与法线共面且入射角等于反射角, 如图 4 所示。此外, 普尔钦斑位于角膜表面, 它与角膜曲率中心的距离为角膜半径。

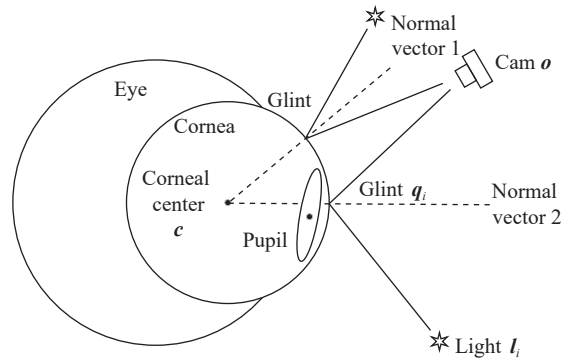


图 4 普尔钦斑几何关系图

Fig.4 The geometric relationship diagram of glints

在典型的单相机双光源硬件结构中, 两个光源会在角膜上产生两个普尔钦斑位置, 光源用 I_i 来表示, 光源对应的普尔钦斑位置用 q_i 表示, 相机光心位置表示为 o , 角膜中心位置表示为 c 。将角膜表面上的点表示为 $E(x_{eye}, y_{eye}, z_{eye}) = 0$, 其中椭圆函数为:

$$E(x_{eye}, y_{eye}, z_{eye}) = \frac{z_{eye}^2}{a^2} + \frac{\gamma_1 \cdot x_{eye}^2 + \gamma_2 \cdot x_{eye} y_{eye} + \gamma_3 \cdot y_{eye}^2}{b^2} - 1 \quad (6)$$

光源至普尔钦斑的入射单位矢量可以表示为公式 (7):

$$\frac{I_i - q_i}{\|I_i - q_i\|} \quad (7)$$

同理, 相机至光源的单位出射矢量为:

$$\frac{o - q_i}{\|o - q_i\|} \quad (8)$$

则入射角与出射角的角平分线可以表示为公式 (9):

$$\frac{(l_i - q_i)/\|l_i - q_i\| + (o - q_i)/\|o - q_i\|}{\|(l_i - q_i)/\|l_i - q_i\| + (o - q_i)/\|o - q_i\|} \quad (9)$$

对于椭球表面的单位法向量则可以用该点的梯

$$\frac{\nabla E(x_{eye}, y_{eye}, z_{eye})}{\|\nabla E(x_{eye}, y_{eye}, z_{eye})\|} - \frac{(l_i - q_i)/\|l_i - q_i\| + (o - q_i)/\|o - q_i\|}{\|(l_i - q_i)/\|l_i - q_i\| + (o - q_i)/\|o - q_i\|} = 0 \quad (11)$$

式中的法向量可以由公式 (6) 推导得到, 椭球上某点的梯度如公式 (12) 所示:

$$\nabla E(x_{eye}, y_{eye}, z_{eye}) = \begin{bmatrix} \frac{2\gamma_1 x_{eye} + \gamma_2 y_{eye}}{b^2} \\ \frac{\gamma_2 x_{eye} + 2\gamma_3 y_{eye}}{b^2} \\ \frac{2z_{eye}}{a^2} \end{bmatrix} \quad (12)$$

则普尔钦斑位置需要联立公式 (6) 和公式 (11), 通过最优化最小二乘求解得到普尔钦斑的理论位置。

1.1.4 瞳孔边缘计算

对于瞳孔边缘, 相机获取的像是瞳孔的折射点, 并不是实际瞳孔。实际瞳孔会经过房水和角膜的折射, 然后从角膜表面的折射点进入空气中, 如图 5 所示。

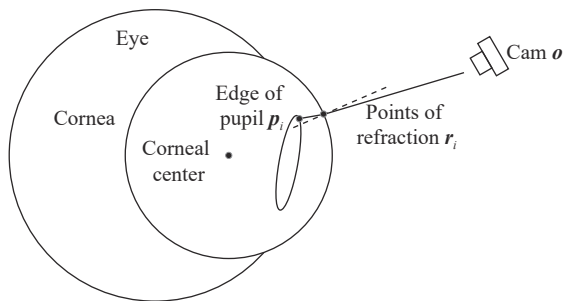


图 5 瞳孔边缘成像过程

Fig.5 Imaging process of pupil edge

在瞳孔边缘取多个点表示为 p_i , 对应的瞳孔边缘折射点则表示为 r_i 。对于瞳孔边缘折射点, 其满足在角膜表面及折射定理的约束, 可联立方程如下:

$$\begin{cases} E(x_{eye}, y_{eye}, z_{eye}) = 0 \\ \left(n_1 \frac{p_i - r_i}{\|p_i - r_i\|} + n_2 \frac{o - r_i}{\|o - r_i\|} \right) \times \nabla E(x_{eye}, y_{eye}, z_{eye}) = 0 \end{cases} \quad (13)$$

1.1.5 视轴与光轴转换

人眼光轴为瞳孔中心与角膜中心的连线, PCCR

度除以梯度的模表示, 如公式 (10) 所示:

$$\frac{\nabla E(x_{eye}, y_{eye}, z_{eye})}{\|\nabla E(x_{eye}, y_{eye}, z_{eye})\|} \quad (10)$$

则满足反射定理的普尔钦斑位置符合数学约束,

如公式 (11) 所示:

法估计的是人眼光轴, 在极坐标下可以用 θ , φ 两个角度表征, 如图 6 所示。

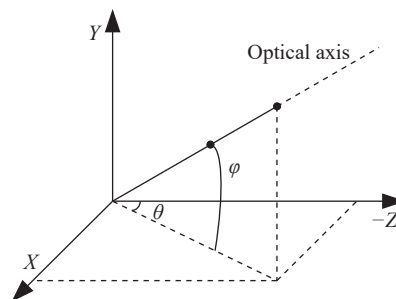


图 6 极坐标系中的光轴

Fig.6 Optical axis in polar coordinate system

在极坐标系下, 光轴可以表示为公式 (14):

$$OA = \begin{bmatrix} \cos \varphi \sin \theta \\ \sin \varphi \\ -\cos \varphi \cos \theta \end{bmatrix} \quad (14)$$

在人眼结构中, 中央凹并不在光轴上, 而是偏离光轴一点, 将角膜中心与中央凹之间的连线定义为视轴, 人眼的注视点在视轴上, 而非在光轴上。每个人视轴和光轴之间存在固定的夹角, 这个夹角称为 Kappa 角, Kappa 角水平分量记为 α_{eye} , 垂直分量记为 β_{eye} 。则视轴方向可以通过 $(\alpha_{eye} + \theta)$ 和 $(\beta_{eye} + \varphi)$ 来确定, 表示为公式 (15)^[19]:

$$VA = \begin{bmatrix} \cos(\varphi + \beta_{eye}) \sin(\theta + \alpha_{eye}) \\ \sin(\varphi + \beta_{eye}) \\ -\cos(\varphi + \beta_{eye}) \cos(\theta + \alpha_{eye}) \end{bmatrix} \quad (15)$$

在视线仿真中, 设置到注视点和角膜中心位置后可以获取视轴的方向向量, 从而可以计算得到视轴的极坐标表示, 即 $(\alpha_{eye} + \theta)$ 和 $(\beta_{eye} + \varphi)$, 再根据设置的 Kappa 角, 可以计算得到光轴的单位矢量。

1.2 光轴重建算法

现有的光轴重建方法均将角膜近似为一个球体,

主流的硬件结构为单相机双光源、双相机多光源。不同的硬件结构可采用的光轴重建算法不同,光轴重建的方法有多种计算途径,已有较多系统的研究,文中不展开赘述。文中以单相机双光源结构的光轴重建方法为例,分别介绍角膜中心求解方法和光轴矢量计算方法。

1.2.1 角膜中心求解

角膜中心位置求解是根据图像中的普尔钦斑位置计算得到的。需要事先已知光源、相机位置、相机内参、角膜曲率半径。在该方法中,角膜被认为是球面,其具备球面法线必经过角膜中心的特点。在该近似条件下,角膜中心与普尔钦斑、相机、光源共面。因此,利用两个普尔钦斑与相机、光源、角膜中心共面,可以得到两个平面(相机与双光源非共线情况),如图 7 所示。

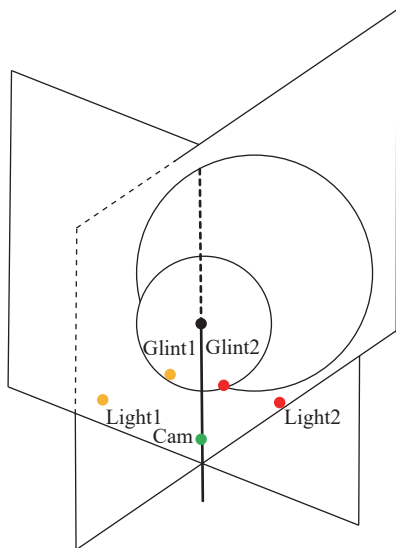


图 7 共面平面示意图

Fig.7 Schematic diagram of coplanar plane

联立两个共面的约束,如公式(16)所示:

$$\begin{cases} (l_1 - o) \times (q_1 - c) \cdot (c - o) = 0 \\ (l_2 - o) \times (q_2 - c) \cdot (c - o) = 0 \end{cases} \quad (16)$$

平面方程可以表示为 $A_i x + B_i y + C_i z + D_i = 0$, 则平面法向量为 (A_i, B_i, C_i) , 由于两平面交线与两个平面的法向量都垂直, 所以交线向量 $L_{\text{cross}} = (A_1, B_1, C_1) \times (A_2, B_2, C_2)$, 由于该交线经过相机光心与角膜中心, 角膜中心可以表示为 $c = k \cdot L_{\text{cross}} + o$, 其中 k 为待求系数。同时, 根据图像中普尔钦斑坐标推导出普尔钦斑

所在的直线方程为:

$$\frac{x - x_o}{m} = \frac{y - y_o}{n} = \frac{z - z_o}{l} \quad (17)$$

式中: (x_o, y_o, z_o) 为相机光心三维坐标; (m, n, l) 为普尔钦斑所在直线的单位向量。角膜表面可以用空间球面方程表示:

$$(x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2 = R_c^2 \quad (18)$$

式中: (x_c, y_c, z_c) 为角膜中心坐标; R_c 为角膜曲率半径。普尔钦斑所在直线与以角膜球面的交点可以通过直线与圆球的求交点公式计算得到普尔钦斑的空间坐标。不同的待定系数 k 可计算得到不同的普尔钦斑坐标, 但其还需要满足反射定理, 如公式(19)所示:

$$\frac{(l_i - q_i) \cdot (q_i - c)}{\|l_i - q_i\|} - \frac{(o - q_i) \cdot (q_i - c)}{\|o - q_i\|} = 0 \quad (19)$$

通过最优化求解 k , 即可计算得到角膜中心。

1.2.2 光轴矢量计算

光轴矢量计算主要依赖于图像中的瞳孔边缘坐标或瞳孔中心坐标。实际的瞳孔中心经过房水、角膜折射后, 在相机成像过程中存在透视变形, 导致二维图像中的椭圆几何中心并非严格意义上的瞳孔中心。根据是否将椭圆几何中心近似作为瞳孔中心像点可以将光轴重建方法分为两种。第一种为利用图像中瞳孔椭圆的几何中心坐标作为瞳孔中心, 再通过光线追迹可以计算得到瞳孔中心位置, 如图 8 所示, 最后连接瞳孔中心与角膜中心作为光轴。第二种为利用瞳孔椭圆边缘进行光线追迹, 计算角膜内的瞳孔边缘折射光线, 加入瞳孔是圆形的约束条件, 最优化求解光轴矢量方向。

对于第一种方法, 在二维图像中拟合瞳孔椭圆得到几何中心作为瞳孔中心的近似, 则瞳孔中心折射点

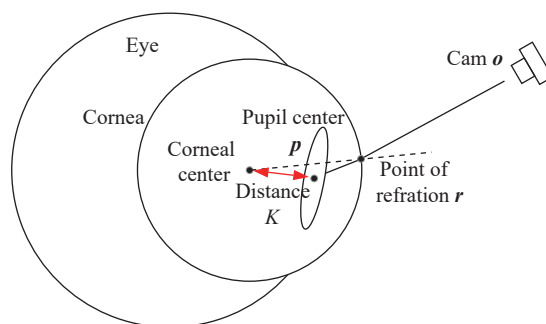


图 8 光线追迹求解瞳孔中心示意图

Fig.8 Ray tracing for solving the center of the pupil

可以表示为 $\mathbf{r} = \mathbf{o} + \chi(\mathbf{o} - \mathbf{v})$, 其中, $\mathbf{v}$ 为图像椭圆几何中心的三维空间坐标, χ 为待定系数, 当角膜中心 $\mathbf{c}$ 已知后, 可以通过射线与圆的交点计算得到 χ 。然后, 瞳孔折射点到相机的光线单位向量可以表示为:

$$\boldsymbol{\tau} = \frac{\mathbf{o} - \mathbf{r}}{\|\mathbf{o} - \mathbf{r}\|} \quad (20)$$

折射点 $\mathbf{r}$ 处的法向量为:

$$\boldsymbol{\eta} = \frac{\mathbf{r} - \mathbf{c}}{\|\mathbf{r} - \mathbf{c}\|} = \frac{\mathbf{r} - \mathbf{c}}{R} \quad (21)$$

根据折射定理可以计算得到角膜内的折射光线向量为公式 (22):

$$\boldsymbol{\iota} = \frac{n_2}{n_1} \left\{ \left[\boldsymbol{\eta} \cdot \boldsymbol{\tau} - \sqrt{\left(\frac{n_1}{n_2} \right)^2 - 1 + (\boldsymbol{\eta} \cdot \boldsymbol{\tau})^2} \right] \boldsymbol{\eta} - \boldsymbol{\tau} \right\} \quad (22)$$

式中: n_1 为房水角膜的等效折射率; n_2 为空气的折射率。再利用角膜中心与瞳孔中心的距离为 K 的条件, 可以计算折射光线 $\boldsymbol{\iota}$ 和半径为 K 的圆球的交点即为瞳孔中心位置。

对于第二种方法, 也需要根据折射定理计算角膜内折射光线, 但其不计算瞳孔中心折射向量, 而是计算多个瞳孔边缘点的折射向量。其计算公式为:

$$\boldsymbol{\iota}_b = \frac{n_2}{n_1} \left\{ \left[\boldsymbol{\eta}_b \cdot \boldsymbol{\tau}_b - \sqrt{\left(\frac{n_1}{n_2} \right)^2 - 1 + (\boldsymbol{\eta}_b \cdot \boldsymbol{\tau}_b)^2} \right] \boldsymbol{\eta}_b - \boldsymbol{\tau}_b \right\} \quad (23)$$

则角膜内瞳孔边缘点可以表示为公式 (24):

$$\mathbf{p}_b = \mathbf{r}_b + \chi_b \boldsymbol{\iota}_b \quad (24)$$

这些瞳孔边缘点在瞳孔平面上, 假设瞳孔中心为 $\mathbf{p}$, 则瞳孔边缘满足公式 (25) 的约束。

$$\frac{(\mathbf{p}_b - \mathbf{c}) \cdot (\mathbf{p} - \mathbf{c})}{K} = K \quad (25)$$

根据公式 (25) 可以计算得到瞳孔边缘的坐标。最后, 由于瞳孔是个圆, 其满足约束:

$$\|\mathbf{p}_1 - \mathbf{p}\| = \|\mathbf{p}_2 - \mathbf{p}\| = \cdots = \|\mathbf{p}_b - \mathbf{p}\| \quad (26)$$

$$Loss = \frac{1}{N} \left(1 + \sqrt{\frac{C}{\|\mathbf{POG}_{GT} - \mathbf{POG}_{center}\|}} \right) \|\mathbf{POG}_{estimation} - \mathbf{POG}_{GT}\| \quad (29)$$

式中: C 为常数项; $\mathbf{POG}_{center}$ 为注视屏幕的中心位置。通过加权可以让屏幕中心的权重大于屏幕周边的权重, 因为屏幕中心被注视的频率更高。

$$Loss = \frac{1}{N} \left(1 + \sqrt{\frac{C}{\|\mathbf{gaze}_{GT} - \mathbf{gaze}_{center}\|}} \right) \arccos(\mathbf{gaze}_{estimation} \cdot \mathbf{gaze}_{GT}) \quad (30)$$

则求解瞳孔中心位置变为使得瞳孔边缘离瞳孔中心方差最小的最优化问题, 如公式 (27) 所示:

$$\min_{\theta, \varphi} \sum (\|\mathbf{p}_b - \mathbf{p}\| - r_{pupil})^2 \quad (27)$$

其中, r_{pupil} 为平均瞳孔半径, 计算方式如公式 (28) 所示:

$$r_{pupil} = \frac{1}{N_{boundary_points}} \sum \|\mathbf{p}_b - \mathbf{p}\| \quad (28)$$

1.3 评价函数构建

通常有两种指标用来评估视线估计的精度, 一类为视线方向的精度, 另一类为注视点 (Point of Gaze, POG) 的精度。同时, 在较大视线角度下, 普尔钦斑处于角膜的边缘, 角膜近似球面的计算会引入更大的估计误差, 所以不同视线范围的视线估计精度不一样。对于实际应用, 无论是遥测式还是头戴式, 均有一个视线范围的限制。为了综合评估视线精度, 文中设置了一个注视的虚拟平面, 在该虚拟平面上均匀采样取测试点, 如图 9 所示。

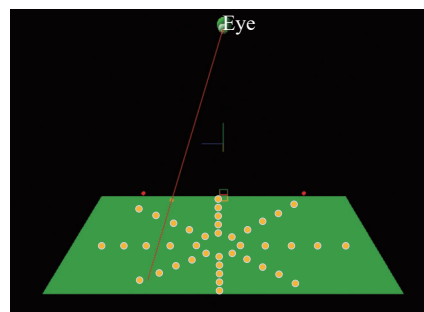


图 9 注视平面上的测试点分布

Fig.9 Distribution of the test points on the plane

通过仿真计算, 可以得到 N 个注视点的视线矢量及视线落点估计值, 如图 10 所示, 红色点为真值, 蓝色点为计算得到的预测值。

取 N 个注视点的加权平均误差作为评价函数, 即:

或者取 N 个注视点的加权平均视线角度误差作为评价函数, 即:

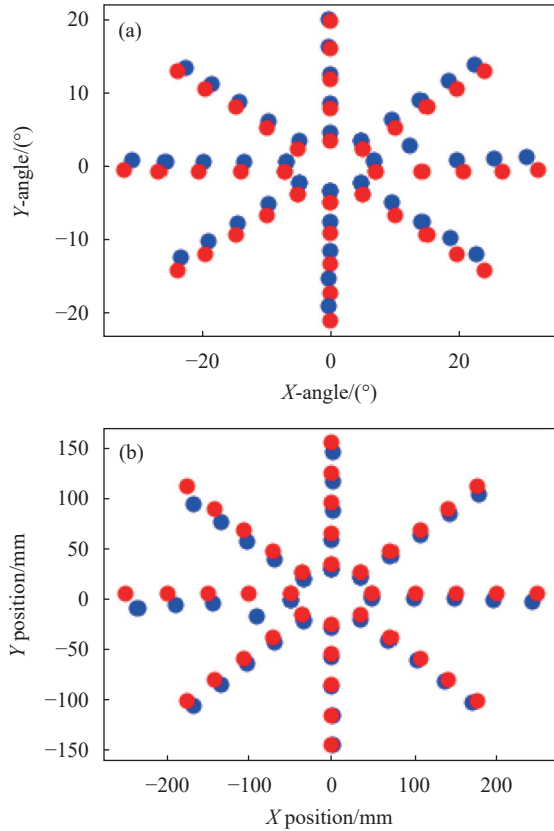


图 10 仿真计算的估计值与真值分布。(a) 视线矢量的估计值与真值; (b) 视线落点的估计值与真值

Fig.10 Distribution of estimated and true values in simulation. (a) The estimated and true values of the gaze vector; (b) The estimated and true values of the POG

式中: $\mathbf{gaze}_{center}$ 为注视屏幕的中心视线向量。

评价函数的值代表了该光源布局下的平均误差水平,其值越低越好。当注视点确定后,该函数与相机位置 $\mathbf{o}$ 、光源位置 $\mathbf{l}_i$ 、人眼参数 $\mathbf{para}_{eye}$ 都有关,可以表示为 $Loss(\mathbf{o}, \mathbf{l}_i, \mathbf{para}_{eye})$ 。对于固定的相机位置,单一人眼参数可以直接采用该评价函数作为优化目标。当眼球参数存在多种时,则可以统计多种典型人眼参数下的 $Loss$, 统计 $Loss$ 之和作为评价函数。

2 典型光源布局优化结果

2.1 典型设置

在现有的 PCCR 法硬件中,主要分为头戴式和遥测式两大类。头戴式的视线估计设备中,相机与人眼位置关系较为固定,典型的应用有 VR 设备、AR 眼镜等,相机固定于眼球下方,光源则要布置于眼眶四周空间,避免遮挡视线,如图 11 所示。

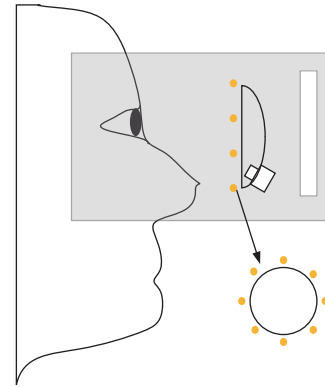


图 11 头戴式设备硬件布局示意图

Fig.11 Hardware layout of head mounted devices

而对于遥测式的设备,其相机往往架设于注视屏幕的正下方,光源一般布置于相机两侧,如图 12 所示。

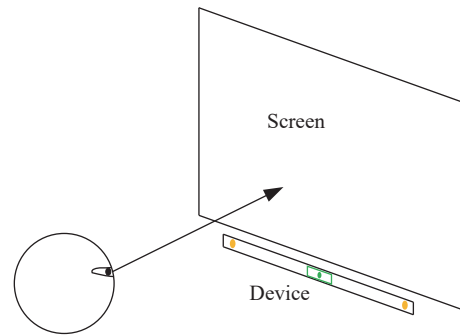


图 12 遥测式设备硬件布局示意图

Fig.12 Hardware layout of desktop devices

由于不同的人眼参数、不同的相机参数所导致的优化结果不同,文中以典型参数为例,对光源布局进行优化。典型参数设置如表 1。

表 1 典型参数设置

Tab.1 Typical parameter setting

Parameter	Value
Corneal elliptical long axis a/mm	7.8
Corneal elliptical short axis b/mm	7.6
Corneal ellipse parameter ξ	1
Corneal ellipse parameter δ	0
Corneal radius in solver R/mm	7.8
Kappa angles	$\alpha = -5^\circ$, $\beta = 1.5^\circ$
Distance between cornea and pupil center K/mm	4.2
Effective relative index of refraction of the cornea and aqueous humor combined	1.3375 ^[20]
Radius of pupil/mm	2

续表 1

Continued Tab.1

Parameter	Value
Camera focal length	16 mm desktop 4 mm head mounted
Distance between screen and eye/cm	40
Camera position @ Corneal coordinate system/mm	(0, 170, 400) desktop (0, 20, 35) head mounted
Camera resolution	1280×720
Pixel size/μm	4.8

2.2 双光源的头戴式光源布局优化

对于头戴式的光源布局,光源可布置的范围只有眼眶周围,设置光源位置在距离角膜中心 20 mm 的 XOY 平面上,先将可布置的光源位置离散化。在眼眶周围取半径为 20 mm 的圆周,从 1 号位置开始,在该圆周上每隔 45°弧度取一个离散位置,眼眶正下方要布置相机,并不能安装光源,因此文中总共取 7 个离散化位置,如图 13 所示。

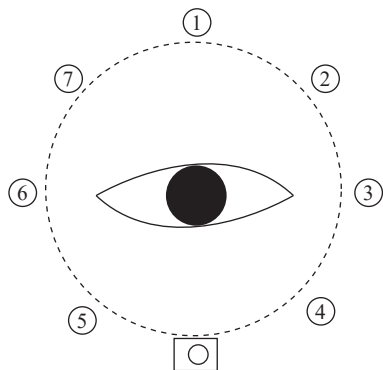


图 13 头戴式设备光源位置索引

Fig.13 Position index of light in head mounted device

先固定其中一个光源,然后另一个光源在剩下的光源遍历,总共有 42 种组合。先统计这 42 种组合的误差评价函数,如图 14 所示。遍历 42 组光源布局组合,发现误差最小的组合是位置 3 和位置 5。

在获取误差最小的离散化位置后,平均视线误差的函数在这附近会是一个凸函数,因此可以再用最优化求解的方式计算更精细的光源布局位置,即求解问题:

$$\min Loss(o^*, l_i, para_{eye}^*) \quad (31)$$

迭代初值为 90°和 225°,经过最优化计算得到结

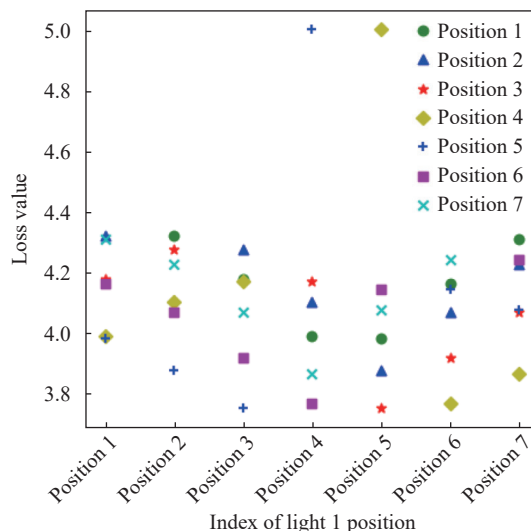


图 14 42 组光源布局的视线精度结果(头戴)

Fig.14 MAE for 42 sets of light source layouts (head mounted)

果为 111.99°和 249.65°。由于视线估计设备可能被不同的人群使用,文中除了典型人眼参数外,还对列举了其他一些非球面人眼参数进行仿真,仿真的方法与典型参数保持一致,文中不再赘述,结果如表 2 所示。

表 2 不同人眼参数下的最优布局(头戴)

Tab.2 Optimal layout of different eye parameters (head mounted)

Corneal parameters	Corneal radius in solver	Optimal layout
7.8, 7.6	7.8	111.99°, 249.65°
7.8, 7.6	7.7	113.00°, 247.90°
7.8, 7.6	7.6	116.75°, 249.50°
7.6, 7.4	7.6	113.83°, 252.47°
7.6, 7.4	7.5	114.61°, 250.72°
7.6, 7.4	7.4	114.24°, 247.70°
8.0, 7.8	8.0	113.82°, 251.65°
8.0, 7.8	7.9	114.35°, 248.83°
8.0, 7.8	7.8	114.98°, 247.38°

由表 2 可知,不同人眼参数下的最优布局也不同,在文中的头戴式设备中,±2 mm 的角膜半径会引起光源位置大约变化 5°。可考虑人群的角膜参数分布权重再进行最优布局的取舍。

2.3 双光源的遥测式光源布局优化

对于遥测式的光源布局,其一般固定在被注视平面的边缘,而且由于设备体积的限制,其一般与相机集成成为长条状设备,架设于屏幕下方。与头戴式的光

源布局优化一样, 首先设置光源布局的离散位置, 然后再进行最优化的精细位置求解。考虑设备的体积限制, 以及光源与相机非共线的要求, 设置的离散化光源位置如图 15 所示, 两个光源布置于相机的两侧。光源和相机都布置于距离角膜中心 400 mm 的 XOY 平面内, 而且光源布置于相机 Y 轴方向上方 10 mm 距离, 在 X 轴方向上每隔 20 mm 取一个离散位置, 相机正上方跳过, 两个光源总共有 64 种组合的离

散位置。

从图 16 可知, 最优的离散位置为 X 轴坐标 -140 和 140 的光源布局, 然后以此作为优化初始值, 最优化代价函数可以得到精细的最优位置为 (-132.98, 132.63)。该结果表明, 对于遥测式的光源布局并不是越远越好或者越近越好, 其存在最优的距离。与头戴式设备一样, 对于遥测式设备也对其他一些非球面人眼参数进行仿真, 仿真结果如表 3 所示。

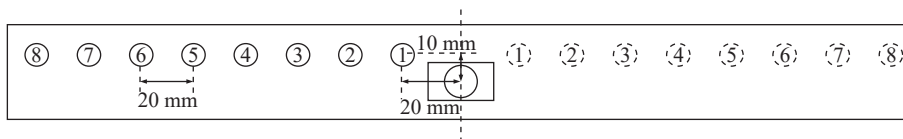


图 15 遥测式设备光源位置索引

Fig.15 Position index of light in desktop device

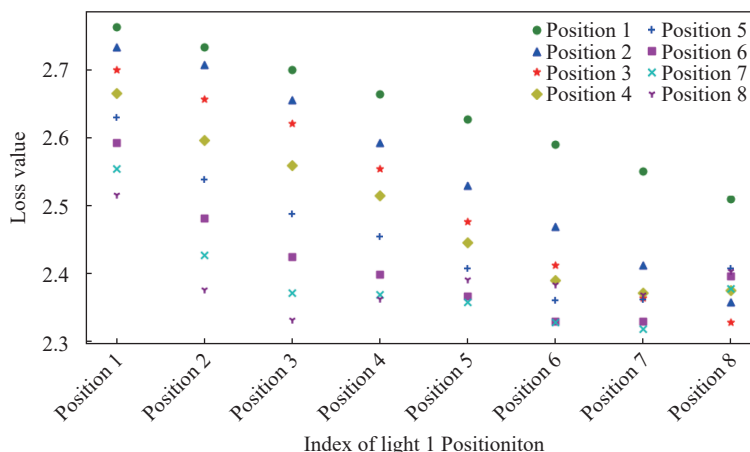


图 16 64 组光源布局的视线精度结果 (遥测)

Fig.16 MAE for 64 sets of light source layouts (desktop)

表 3 不同人眼参数下的最优布局 (遥测)

Tab.3 Optimal layout of different eye parameters (desktop)

Corneal parameters	Corneal radius in solver	Optimal layout
7.8, 7.6	7.8	-132.98, 132.63
7.8, 7.6	7.7	-136.90, 136.73
7.8, 7.6	7.6	-143.17, 142.66
7.6, 7.4	7.6	-131.27, 130.68
7.6, 7.4	7.5	-137.24, 136.79
7.6, 7.4	7.4	-141.07, 141.43
8.0, 7.8	8.0	-133.96, 133.68
8.0, 7.8	7.9	-138.09, 137.78
8.0, 7.8	7.8	-144.01, 143.85

由表 3 可知, 对于文中的遥测式设备, 当角膜参数上下浮动 ± 0.2 mm 时, 最优光源布局也会随之改变大约 10 mm 量级的位置。与头戴式设备一样, 可考虑人群的角膜参数分布权重再进行最优布局的取舍。

3 结 论

瞳孔角膜反射法是目前视线追踪中精度较高的一类方法, 其依赖于专用的设备, 在对精度有要求的人机交互应用中, 该方法是目前的主流技术手段。但目前对于硬件结构布局的讨论更多是定性的研究, 对于量化的硬件优化还缺乏指导, 而且角膜非球面也是制约该方法精度的重要因素之一。文中提出了一种光源布局优化方法, 可以减少角膜非球面对光轴重

建的影响,从而提高视线估计精度。文中的工作主要分析了典型参数下双光源的布局优化问题,讨论了头戴式和遥测式的最优布局,也可以作为 PCCR 设备的理论参考。此外,对于更多数量光源的布局优化以及更多场景(例如智能座舱的 DMS 系统)下的光源布局优化也是未来值得进一步分析的工作。

参考文献:

- [1] CHI Jiannan, WANG Zhiliang, ZHANG Chuang. Eye Gaze Tracking [M]. Beijing: China Machine Press, 2011. (in Chinese)
- [2] LIU Jiahui. Research on 3D gaze estimation methods based on simplified hardware system[D]. Beijing: University of Science and Technology Beijing, 2020. (in Chinese)
- [3] ROBERT K, ANASTASIOS A, AND G W. Gaze-contingent ocular parallax rendering for virtual reality [J]. *ACM Transactions on Graphics (TOG)*, 2020, 39(2): 1-12.
- [4] ZHOU Xiaolong, TANG Fanyang, GUAN Qiu, et al. A survey of 3D eye model based gaze tracking [J]. *Journal of Computer-aided Design & Computer Graphics*, 2017, 29(9): 1579-1589. (in Chinese)
- [5] YIN Pengwei, WANG Jingjing, DAI Jiawu, et al. NERF-GAZE: A head-eye redirection parametric model for gaze estimation [C]//ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2024: 2760-2764.
- [6] BAO Yiwei, LIU Yunfei, WANG Haofei, et al. Generalizing gaze estimation with rotation consistency [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 4197-4206.
- [7] WANG Yaoming, JIANG Yangzhou, LI Jin, et al. Contrastive regression for domain adaptation on gaze estimation [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 19376-19385.
- [8] YIN Pengwei, ZENG Guanzhong, WANG Jingjing, et al. CLIP-Gaze: Towards general gaze estimation via visual-linguistic model [C]//Proceeding of the AAAI Conference on Artificial Intelligence, 2024, 38(7): 6729-6737.
- [9] ZHANG Chuang, CHI Jiannan, ZHANG Zhaohui, et al. A novel eye gaze tracking technique based on pupil center cornea reflection technique [J]. *Chinese Journal of Computers*, 2010, 33(7): 1272-1285. (in Chinese)
- [10] LIU Jiahui, CHI Jiannan, et al. Iris feature-based 3-D gaze estimation method using a one-camera-one-light-source system [J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(7): 4940-4954.
- [11] LIU Jiahui, CHI Jiannan, HU Wenxue, et al. 3D model-based gaze tracking via iris features with a single camera and a single light source [J]. *IEEE Transactions on Human-machine Systems*, 2021, 51(2): 75-86.
- [12] VILLANUEVA A, CABEZA R. A novel gaze estimation system with one calibration point [J]. *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, 2008, 38(4): 1123-1138.
- [13] SHIH S, LIU J. 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.
- [14] VILLANUEVA A, CABEZA R. Evaluation of corneal refraction in a model of a gaze tracking system [J]. *IEEE Transactions on Biomedical Engineering*, 2008, 55(12): 2812-2822.
- [15] LAI C, SHIH S, HUNG Y. Hybrid method for 3-D gaze tracking using glint and contour features [J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2015, 25(1): 24-37.
- [16] NLTSCHKE C, NAKAZAWA A, TAKEMURA H. Corneal imaging revisited: an overview of corneal reflection analysis and applications [C]//IPSJ Transactions on Computer Vision and Application, 2013, 5: 1-18.
- [17] LUO Lili. Research on gaze estimation based on eye model[D]. Xi'an: Xidian University, 2019. (in Chinese)
- [18] GUESTRIN E D. Remote, non-contact gaze estimation with minimal subject cooperation[D]. Canada: University of Toronto, 2010.
- [19] BROWN D C. Close-range camera calibration [J]. *Photogrammetric Engineering*, 1971, 37(8): 855-866.
- [20] CORBETT M C, ROSEN E S, O'BRART D P S. Corneal Topography: Principles and Application [M]. London: BMJ Books, 1999: 6.

Optimization method for light layout of eye tracking device

XU Zefu^{1,2}, WANG Jingjing¹, ZENG Guanzhong¹, XIE Di^{1*}, FENG Huajun²

(1. Hikvision Research Institute, Hangzhou 310052, China;

2. College of Optical Science and Engineering, Zhejiang University, Hangzhou 310027, China)

Abstract:

Objective Compared to the appearance-based method, Pupil Center Corneal Reflection (PCCR) method has higher accuracy, and has great significant application value in fields such as human-computer interaction and AR/VR. However, the existing researches mainly discuss the structural forms of eye tracking devices such as single camera single light, single camera multiple light, and multi camera multiple light, lacking quantitative hardware layout optimization methods. At the same time, in the PCCR method, the cornea is considered as a spherical surface, while human cornea is not ideal spherical, only the central part of it can be approximated as spherical. Especially when the distance between the light sources is too far, it can cause glints at the edge of the cornea, where the cornea is not spherical. The main solution is to make the glints as close as possible to the spherical area of the cornea, but if the distance between the light sources is too close, significant calculation errors will emerge. Therefore, optimizing the light layout of eye tracking devices has strong practical significance, when the distance between light sources is mutually constrained.

Methods The method of computer simulation is used to theoretically calculate the average error of gaze estimation under different light layout, thereby deriving the optimal light source position. The main simulation process is divided into three parts, including eye imaging, gaze estimation based on simulation images, and construction of loss functions. The non spherical surface of the human cornea is considered in the content of eye modeling and imaging, which is closer to the actual human cornea. A typical pinhole model with superimposed distortion is adopted in the camera imaging model. For the optical axis reconstruction in gaze estimation, a typical optical axis reconstruction method suitable for single camera multi light devices has been adopted, which first estimates the corneal center position and then calculates the gaze optical axis. In the loss function, the error statistics of uniformly distributed test points are used as comprehensive evaluation indicators to calculate the iteration step of the light source position.

Results and Discussions Different devices have different optimal light layouts, due to different camera focal lengths, installation positions, and user eye parameters. As an example, typical human eye parameters are used, and the optimal light layout for different forms of devices is discussed for both head mounted and desktop devices. In head mounted devices, the light source can be distributed in a circular pattern, and 7 discrete positions around the eye are taken as rough positions. After obtaining the optimal discrete position sets from 42 sets (Fig.14), as the initial value for optimization, the optimal light position can be iteratively optimized. Similarly, the light source of desktop devices is distributed in a linear pattern at the bottom of the screen. 8 discrete positions are taken and the optimal set of discrete positions is obtained from 64 sets (Fig.16). Then, the optimal light position can be iteratively obtained. The quantitative results indicate that there is an optimal position for the light source, and it can improve the accuracy of gaze estimation. Compared to existing studies, the results obtained by this method are more quantitative, and the ellipsoidal cornea is closer to the human cornea.

Conclusions The PCCR method is an effective method for gaze estimation, but the cornea is considered as a sphere, which limits the accuracy. A quantitative optimization method of light layout has been proposed, and a framework for light layout optimization has been established, which can optimize different optimal light layouts according to actual application settings. At the end, typical system parameters are simulated, and the results show that the distance between the light sources is not necessarily better as it is larger or smaller. There is an optimal light source position that can suppress the error increase caused by ellipsoidal corneal. The simulation results can also serve as a theoretical reference for corresponding device design.

Key words: gaze estimation; light layout; pupil center-corneal reflection; non-spherical cornea