



论 文

视线追踪系统中视线估计方法研究

张闯*, 迟健男*, 张朝晖, 高小亮, 胡涛, 王志良

北京科技大学信息工程学院, 北京 100083

* 通信作者. E-mail: zhangchuangtc@163.com, sy-jnchi@126.com

收稿日期: 2010-05-13; 接受日期: 2010-08-30

国家自然科学基金 (批准号: 60574090) 和国家高技术研究发展计划 (批准号: 2007AA01Z160) 资助项目

摘要 文中提出一种基于瞳孔—角膜反射 (PCCR) 的视线估计方法 (GEMHSSO). 针对现有 PCCR 存在的主要问题: 限制使用者头部运动和个体标定问题, 提出了一种单相机单光源条件下头部位置的补偿方法, 实现了头部位置变化对瞳孔角膜向量影响的解析补偿, 并建立一种个体差异的转化模型, 进而简化标定过程为单点标定. 以此为基础形成一种新的视线估计方法, 本方法使精确视线估计的最小硬件要求降低到单相机 (未标定) 单光源, 既不需要繁杂的系统标定, 又实现了自然头动视线估计, 并且简化用户标定为单点标定. 该方法的各个环节都满足实时性要求, 为面向人机交互的视线追踪系统提供了有效的解决方案.

关键词 视线追踪 视线估计 瞳孔—角膜反射 (PCCR) 技术 头动补偿 个体标定

1 引言

视线追踪^[1,2]是利用机械、电子、光学等各种检测手段获取受试者当前“注视方向”的技术. 按照系统构成和采用的检测方法可以粗略划分为侵入式和非侵入式两种. 随着数字化技术、计算机视觉、人工智能技术的迅速发展, 基于数字视频分析 (video oculographic, VOG) 的非侵入式视线追踪技术成为当前热点研究方向. 基于 VOG 的视线追踪技术分为视线特征参数提取和视线估计模型建立两个组成部分. 随着系统硬件和视线估计模型不同, 基于 VOG 的视线追踪方法可分为如下几类:

(a) 瞳孔—角膜反射法 (the pupil center cornea reflection technique, PCCR): 采用红外光源产生角膜反射, 通过计算瞳孔中心与角膜反射之间的向量来估计视线方向^[3-6]. 视线估计模型可以采用映射模型 (通常为非线性多项式模型)^[7], 也可以采用几何模型^[8-10].

(b) 角膜反射矩阵法 (cross-ratios method): 多红外光源在角膜上产生多个角膜反射形成角膜反射矩阵. 通过计算角膜反射与瞳孔中心的位置特征估计视线方向, 视线估计模型使用投影平面不变的交比 (cross-ratios) 值^[11,12].

(c) 椭圆法线方向法 (“one-circle” algorithm): 通过记录高分辨率眼睛图像, 计算虹膜边界圈的法线方向估计视线方向. 在可见光下追踪高分辨率的眼睛图像并通过基本的投影几何关系计算椭圆虹膜圈的法线方向并视其为视线方向^[13-16].

VOG 系统普遍使用瞳孔—角膜反射方法. 基于瞳孔角膜反射的视线追踪系统, 其精确性与稳定性可以通过设置与摄像机镜头同轴的近红外光源进一步提高, 其可以产生“亮瞳”现象. 大部分基于瞳

孔—角膜反射技术的视线估计可分为两类: 基于二维映射模型的视线估计方法^[7,17]和直接的三维视线估计方法^[8-10,18-21].

对于直接的三维视线估计方法, 首先要估计三维的视线方向, 然后通过视线方向和屏幕的交点即可得到盯视点. 文献 [18] 提出一种用以估计三维视线方向的方法, 使用一个单校准摄像机和至少两个光源. 首先, 使用至少 3 个光源测量每个使用者的眼球角膜半径. 一系列高阶多项式方程用来计算角膜的半径和中心, 但是它们的解不是唯一的. 因此, 如何从这些解中找到正确的答案仍然是个问题. 此外, 还没有用这种技术建立起相应的实用系统. 文献 [19] 提出了一种类似的方法来估计三维视线方向. 限定了一些初始条件, 首先角膜半径、瞳孔和角膜中心的距离对所有用户是一样的, 当然事实上并不是如此的. 其次, 用来计算角膜中心的公式是基于红外光源虚图像在角膜表面上的假设, 事实上, 红外光源的虚图像是在角膜里面而非表面. 因此, 计算角膜中心的公式是不准确的. 文献 [20] 提出了另一个系统和至少有 7 个参数的复杂眼球模型来估计三维视线方向. 首先, 三维眼球模型会针对每个使用者进行建模, 它将一系列的图像特征通过非线性估计技术适配至眼球模型中以达到建模的目的. 用以适配模型的图像特征仅仅包括由红外光源产生的角膜反射和瞳孔边缘. 但是角膜反射是由红外光源照射角膜产生的虚图像的投影, 并不在角膜表面而在角膜内部. 同样, 瞳孔边缘也不是在三维眼球模型的表面. 因此, 角膜半径不能基于上述方法得到. 此外, 对这样复杂的三维模型只使用很少的特征点, 方程的解对噪声会很敏感, 鲁棒性也较差. 文献 [21] 提出了一种使用多摄像机和多光源估计三维视线方向的方法. 在这个方法中, 虽然不需要知道使用者眼睛的个体差异参数, 但是立体摄像机使其系统有明显的局限. 特别是, 当用户盯视点在两个摄像机的光学中心的连线上时, 三维视线方向不能唯一的得到. 文献 [8] 使用一个或两个摄像机配合多个光源来估计三维视线. 首先重建眼睛的光轴. 随后通过一个校准过程用已知的光轴得到视轴. 通过比较分析了多个系统的不同, 理论上得出了三维视线估计的系统最少配置, 一个摄像机和两个光源配以一些使用者眼睛的独立参数, 或是不用参数情况下使用两摄像机和两个光源. 文中用一个摄像机和两个光源实现了一个视线估计系统, 并实验验证了准确性. 与双摄像机直接估计视线的系统相比, 单摄像机代表了一种重要的简化, 但是这种方法不能适应头动较大的情况.

综上所述, 三维视线估计方法有以下几个优点: (a) 因为采用立体视觉, 可以检测使用者头部的空间位置, 所以允许使用者头部自然运动. (b) 因为只要标定视线与眼球光轴的夹角, 不需要依靠标定确定眼睛特征参数与视线的映射函数, 所以标定点数较少, 在双相机双光源的条件下可以做到单点标定. (c) 三维视线估计方法估计的是空间视线方向, 而不是屏幕视线落点, 所以视线估计不依赖于屏幕位置、大小、形状, 可以计算视线在任何物体上的注视点, 并且不需要重新标定视线估计函数.

同时, 三维视线估计方法有以下几个缺点: (a) 在现存的三维视线估计方法中, 要么需要关于用户眼球参数的独立信息, 要么需要至少两摄像机和两个光源的复杂硬件配置. 否则就不能解决头动问题. 但是使用者眼球的独立信息, 如角膜半径和瞳孔与角膜中心的距离, 是非常小的 (一般小于 10 mm). 因此, 在不借助其他仪器的情况下, 准确的间接估计眼球独立参数是很难实现的. (b) 即使采用一个摄像机和两个光源配以使用者眼睛独立参数的方案, 在系统标定过程中也要使用至少两个摄像机. (c) 因为需要估计角膜球面中心的空间位置, 而角膜球面中心不可见, 需要至少两个光源的像去估计角膜球面中心的空间位置, 所以至少需要两个光源. (d) 需要进行相机标定, 光源位置标定和屏幕位置标定, 并且三维视线估计对以上位置非常敏感, 相机发生变化不仅要对自身重新进行标定, 而且要对光源和屏幕重新进行标定, 光源和屏幕发生变化要对自己重新进行标定.

基于二维映射模型的视线估计方法, 通过一个经过校准的视线映射函数来估计视线方向, 映射函

数的输入是从眼睛图像提取的一系列二维眼动特征, 输出是视线方向或注视点. 二维映射模型不需要估计三维的视线方向, 所以不需要使用立体视觉系统, 不需要进行相机的标定, 不需要进行光源和屏幕三维位置的标定, 为低硬件配置条件下的视线估计提供了有效的解决方案.

提取的二维眼动特征随视线而变化, 使它们之间的关系可以由一个视线映射函数来表示. 为了得到视线映射函数, 需要对每个使用者进行在线校准. 但是, 二维眼动特征随着头部位置的变化而显著变化, 因此, 校准的视线映射函数对头部位置非常的敏感^[17]. 因此, 为了得到准确的注视点, 使用者需要保持头部静止. 如果使用者保持其头部固定, 或通过支架限制其头部活动, 眼睛注视点跟踪的结果可以达到非常高的精度. 平均误差可以小于 1° (对应在计算机屏幕上小于 10 mm). 但是如果头部离开使用者校准时的位置, 视线跟踪系统的准确性将显著下降, 文献 [17] 报告了详细的数据, 数据显示, 视线映射函数的准确性将严重降低. 可以通过使用者局部手工重新校准解决这个问题^[22], 但是这给使用者带来许多麻烦.

综上所述, 二维视线估计方法有以下几个优点: (a) 因为二维视线估计仅需要使用平面眼睛特征参数, 眼睛特征识别和参数提取简单快速. (b) 因为不需要采用立体视觉, 所以可以使用单相机. (c) 不需要估计角膜球面中心的空间位置, 仅需要知道平面图像上的一个 Purkinje 斑与瞳孔中心的相对位置就可以进行二维视线估计, 所以可以采用单光源. (d) 二维视线估计的精度依赖于标定算法而不依赖于立体视觉, 不需要进行三维重构, 对硬件配置要求较低. (e) 不需要进行相机标定, 光源位置标定和屏幕位置标定, 仅需进行用户标定.

同时, 二维视线估计方法有以下几个缺点: (a) 二维视线估计依赖于标定位置, 视线估计的精度随着使用者头部远离标定位置而迅速下降, 所以使用者需要保持头部静止. (b) 需要依靠标定确定眼睛特征参数与视线的映射函数, 所以标定点数较多.

总括来说, 理想的视线估计方法需要满足以下几点要求: (a) 准确的, 例如, 精确到分; (b) 可靠的, 结果可重复; (c) 鲁棒的, 可以在各种条件下正常工作, 例如室内/室外, 带眼镜, 带隐形眼镜; (d) 非接触式的, 对用户无害, 舒适的; (e) 允许自由的头部运动; (f) 尽可能少的初始校准; (g) 实时的; (h) 硬件配置简单低廉. 而现有的视线估计方法都不能完全满足上述要求. 直接的三维视线估计方法都存在两个共同的缺点: 第一, 至少需要两摄像机 (已标定) 和两个光源的复杂硬件配置; 第二, 繁杂的系统标定过程. 现有的基于二维映射模型的估计方法都存在两个共同的缺点: 第一, 用户使用前都需要进行多点个体标定; 第二, 用户需要保持头部静止.

为满足上述要求, 并针对现有视线估计方法的不足, 本文提出一种可适应自然头动的视线估计方法 (GEMHSSO, gaze estimation method with head compensation, single camera, single light source, one-point calibration). 在单相机单光源条件下实现了头动对视线参数影响的解析补偿, 使精确视线估计的最小硬件要求降低到单相机 (未标定) 单光源, 既不需要繁杂的系统标定, 又实现了自然头动视线估计, 并且简化用户标定为单点标定.

2 视线追踪系统

本文立足于开发一种具有普适应用价值的低成本单摄像机单光源视线追踪系统, 系统由光源、滤光片、镜头、采集卡、CCD (charge coupled device)、GPIO (general purpose input/output) 卡、单片机、主机和屏幕组成. 使用者注视屏幕, 由 CCD 摄像机获取人脸图像, 通过图像采集卡传到主机, 主机通过特征参数提取和视线映射函数来得到视线落点, 显示在屏幕上. 本文视线估计基于瞳孔—角膜反

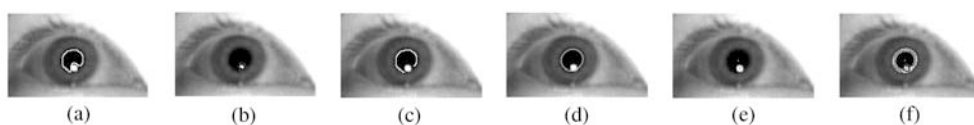


图 1 人眼特征提取过程

Figure 1 Feature extraction process

(a) Pupil edge; (b) corneal reflex contour; (c) pupil edge (after filtering); (d) fitted pupil ellipse; (e) the pupil center and corneal reflex center; (f) results

射技术.

另一方面, 为了得到方便处理的人脸图像, 采用 GPIO 卡获取 CCD 视频图像的帧同步信号, 再通过单片机控制光源的开关. 通过内外环光源的交替亮灭产生亮瞳和暗瞳图像. 本文采用的近红外发光二极管的功率在安全阈值以下, 对人体无害且对使用者不可见.

本文视线追踪系统原理流程包括特征参数提取, 视线映射函数和头部位置及个体差异补偿算法. 其中本文特征提取过程如下: 首先, 亮瞳与暗瞳图像相减得到差分图像, 对差分图像做滤波, 得到瞳孔区域. 检测瞳孔区域的边缘 (图 1(a)) 并在眼睛区域附近基于灰度搜索角膜反射 (图 1(b)). 求质心定位角膜反射中心 (图 1(e)), 并对瞳孔边缘做滤波消除角膜反射对瞳孔边缘轮廓的影响 (图 1(c)), 椭圆拟合定位瞳孔中心, 得到亚像素的中心坐标 (图 1(e)). 最后提取的视线特征向量 L_t (图 1(f)) 如下:

$$L_t = \left(\Delta x, \Delta y, \frac{a_{\text{major}}}{a_{\text{minor}}}, \theta, i_c, j_c \right).$$

其中 $(\Delta x, \Delta y) = p_{i_p j_p} - p_{i_c j_c}$ 为瞳孔中心到角膜反射的向量, $\frac{a_{\text{major}}}{a_{\text{minor}}}$ 为瞳孔椭圆长短轴的比, θ 为瞳孔椭圆长轴与垂直方向的角度, (i_p, j_p) 为瞳孔中心在图像中的位置, (i_c, j_c) 为角膜反射在图像中的位置.

3 视线估计

本文提出一种基于 PCCR 的视线估计方法, 包括个体差异补偿, 视线映射函数和头部位置补偿, 如图 2 所示.

(1) 个体差异补偿:

本文采用的视线特征向量 L_t 如下:

$$L_t = \left(\Delta x, \Delta y, \frac{a_{\text{major}}}{a_{\text{minor}}}, \theta, i_c, j_c \right). \quad (1)$$

为了补偿个体差异对视线估计的影响, 本文经过大量实验总结出以下事实: (a) 不同的使用者在同一位置注视同一点的情况下, 眼球的方位是大致相同的, 因此 $(\Delta x, \Delta y)$ 的方向是大致相同的, 其方向的差异是由于视线和眼球光轴夹角的个体差异造成的, 如图 3(a) 所示, 这个差异可以通过单点用户标定进行捕捉. (b) 不同的使用者在同一位置注视同一点的情况下, $(\Delta x, \Delta y)$ 的长度差异往往大于方向的差异, 其差异是由于角膜球面半径大小的个体差异造成的, 这个差异也可以通过单点用户标定进行捕捉. 如图 3(b) 所示, 当角膜球面半径发生变化时, 瞳孔中心 P_1 发生线性变化, 图像中的瞳孔中心

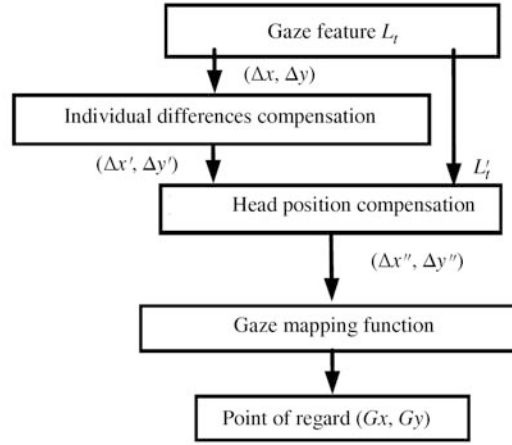


图 2 视线估计框架

Figure 2 Gaze estimation flowchart

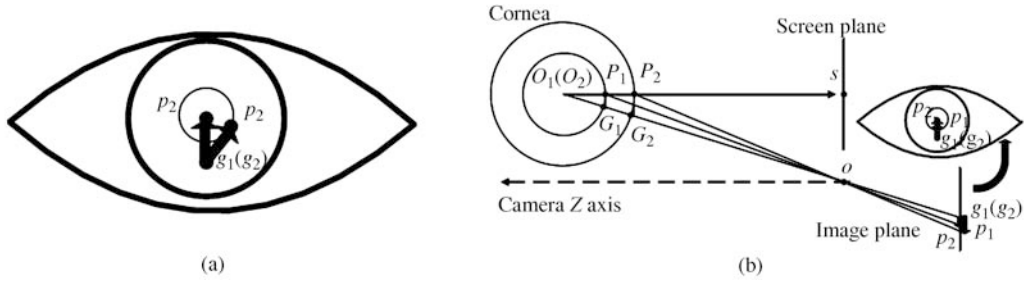


图 3 角膜个体差异示意图

Figure 3 A schematic diagram of the individual differences of cornea

(a) The angle of optical axis and gaze; (b) individual difference of corneal spherical radius

p_1 也随之线性变化, 而图像中的 Purkinje 斑 g_1 不变. 即角膜球面半径增加 k 倍, g_1 到 p_1 的向量 $(\Delta x, \Delta y)$ 长度随之增加 k 倍. (c) 不同的使用者在同一位置注视同一点的情况下, $(\frac{a_{\text{major}}}{a_{\text{minor}}}, \theta, i_c, j_c)$ 可以认为是相同的.

本文提出一种视线特征向量个体差异补偿方法, 通过单点标定确定使用者的角膜半径比例系数 k 和视线偏角 λ , 补偿方法如下:

$$(\Delta x', \Delta y') = k \cdot (\Delta x \cos \lambda - \Delta y \sin \lambda, \Delta x \sin \lambda + \Delta y \cos \lambda), \quad (2)$$

$(\Delta x', \Delta y')$ 为补偿以后的瞳孔中心到 Purkinje 斑向量. 补偿以后的视线特征向量 L'_t 为

$$L'_t = (\Delta x', \Delta y', \frac{a_{\text{major}}}{a_{\text{minor}}}, \theta, i_c, j_c).$$

(2) 多项式模型:

完整的视线映射函数 $(Gx, Gy) = F(\Delta x, \Delta y)$ 是一个复杂的非线性函数, 本步骤讨论的是使用者和头部位置固定情况下的视线映射函数 $(Gx, Gy) = f(\Delta x, \Delta y)$. 在本文的系统配置下, 其函数规律总

结如下:

$$\begin{aligned} Gx &= f_x(\Delta x, \Delta y) \approx a_1(\Delta y) + a_2(\Delta y)\Delta x, \\ a_1(\Delta y) &\approx a_3 + a_4\Delta y, \\ a_2(\Delta y) &\approx a_5 + a_6\Delta y, \\ Gy &= f_y(\Delta x, \Delta y) \approx b_1(\Delta x) + b_2\Delta y + b_3\Delta y^2, \\ b_1(\Delta x) &\approx b_4 + b_5\Delta x. \end{aligned}$$

可以看到, $(Gx, Gy) = f(\Delta x, \Delta y)$ 中有 8 个未知数, 通过 4 个以上的标定就可以确定其关系, 为了覆盖屏幕落点的各个区域, 本文通过 9 点标定对应的 18 个等式对 8 个未知数进行多项式回归. 这个回归过程只在建立模型的时候进行一次, 一旦 8 个未知数已经确定, 以后不同使用者使用系统时, 因为已经经过上个步骤的个体差异补偿, 所以可以直接使用这个视线映射函数, 无需重新进行回归过程.

(3) 头部位置补偿:

为了解决使用者头部位置改变对视线估计精度的影响, 本文通过分析头动后眼部特征和头部位置变化的关系, 提出了相应的补偿方法. 当用户在 O_1 和 O_2 两个位置盯视同一屏幕点 S 时, 如图 4 所示, T 为屏幕盯视点 S 到镜头中心 o 的距离, P_1G_1 , P_2G_2 分别为 O_1 和 O_2 位置瞳孔到 Purkinje 斑的矢量值, p_1g_1 , p_2g_2 分别为 O_1 和 O_2 位置瞳孔到 Purkinje 斑矢量的图像值, r 为角膜曲率半径, D_1 , D_2 为角膜曲率中心到视线落点 S 的距离.

通过三角形 oSO_1 , oSO_2 的几何比例关系可得到

$$\frac{P_1G_1}{T} = \frac{r}{D_1}, \quad \frac{P_2G_2}{T} = \frac{r}{D_2}, \quad \text{即} \quad \frac{P_1G_1}{P_2G_2} = \frac{D_2}{D_1}.$$

d_1, d_2 分别为 O_1 和 O_2 位置角膜曲率中心到屏幕的距离, f 为镜头焦距, 即焦点到 CCD 面板的距离, 根据成像原理有

$$\frac{P_1G_1}{p_1g_1} = \frac{d_1}{f}, \quad \frac{P_2G_2}{p_2g_2} = \frac{d_2}{f}.$$

由上式可以得到

$$\frac{p_1g_1}{p_2g_2} = \frac{d_2}{d_1} \cdot \frac{P_1G_1}{P_2G_2} = \frac{d_2}{d_1} \cdot \frac{D_2}{D_1}, \quad \text{即} \quad \frac{p_1g_1}{p_2g_2} = \frac{d_2}{d_1} \cdot \frac{D_2}{D_1}.$$

由于单摄像机视线追踪系统无法测量角膜曲率中心到屏幕的距离 d 和角膜曲率中心到视线落点的距离 D , 所以本文采用使用者在不同位置时摄像机测量的瞳距值进行代替.

如图 5(a) 所示, 使用者的瞳距 L 是固定值, 但当使用者处于不同位置时, 瞳距在图像上的值 L_1 , L_2 与用户距摄像机距离成反比, 如下式所示:

$$\frac{L_1}{L} = \frac{f}{d_1}, \quad \frac{L_2}{L} = \frac{f}{d_2}.$$

由此可得 $\frac{d_1}{d_2} = \frac{L_2}{L_1}$.

如图 5(b) 所示, 可以通过瞳孔椭圆的曲率估计角膜曲率中心到视线落点的距离 D .

$$D_1 = \frac{d_1}{\cos \gamma_1}, \quad D_2 = \frac{d_2}{\cos \gamma_2}.$$

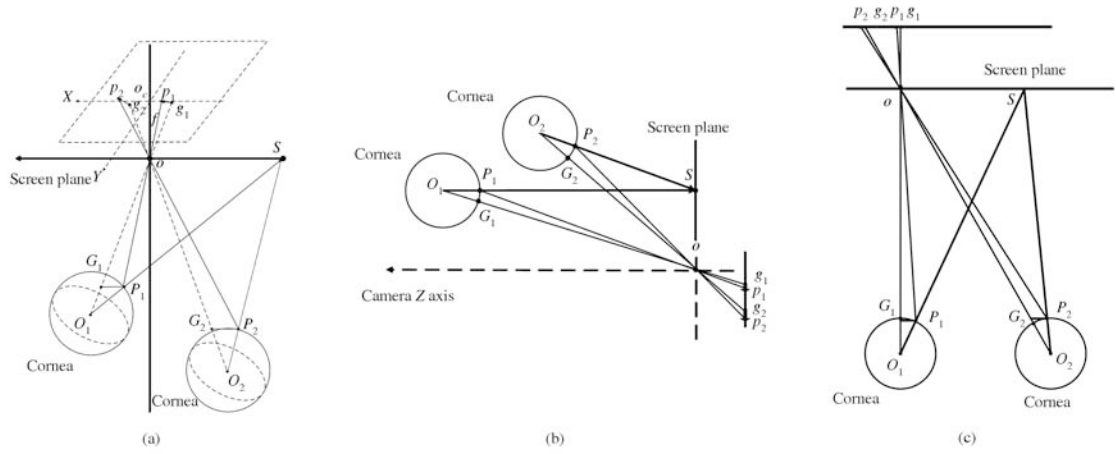


图 4 头部位置对视线特征向量的影响

Figure 4 The impact of head position on gaze feature vectors

(a) Space diagram; (b) side view; (c) top view

其中 $\cos \gamma_1 = \frac{a_{\text{minor}1}}{a_{\text{major}1}}$, $\cos \gamma_2 = \frac{a_{\text{minor}2}}{a_{\text{major}2}}$, 所以 $\cos \gamma_2$ 由特征向量可知. $\cos \gamma_1$ 为标定位置瞳孔角度, 如图 5(c) 所示, $\frac{T_1}{d_1} = \tan \gamma_1$, $\frac{T_2}{d_2} = \tan \gamma_2$, 即 $\frac{T_1 + \Delta T}{d_2} = \tan \gamma_2$, 可得, $\frac{d_1}{d_2} \tan \gamma_1 + \frac{\Delta T}{d_2} = \tan \gamma_2$. 另外, 根据成像关系^[23] $\Delta T = \frac{d_1}{f} \Delta t$, 其中 Δt 为图像距离, f 为相机焦距. 由此, $\tan \gamma_1$ 可知.

如图 5(d) 所示, 头部转动会造成图像中瞳距的变化, 需要计算消除头部转动的瞳距 L_2 . 设左右瞳孔在图像中的长轴分别为 a_L , a_R , 左右角膜曲率中心到屏幕的距离分别为 z_L , z_R , L_1 为标定位置的图像瞳距, z_1 为标定位置到屏幕的距离, 根据成像关系和几何关系有

$$\begin{cases} \frac{z_L}{z_R} = \frac{a_R}{a_L}, \\ \frac{z_L - z_R}{L} = \frac{o_L M}{o_L o_R}, \\ \frac{L_2}{L_1 \frac{z_1}{\left(\frac{z_L + z_R}{2}\right)}} = \frac{o_R M}{o_L o_R}, \\ \left(\frac{o_L M}{o_L o_R}\right)^2 + \left(\frac{o_R M}{o_L o_R}\right)^2 = 1. \end{cases}$$

即

$$\begin{cases} \frac{z_L}{z_R} = \frac{a_R}{a_L}, \\ \left(\frac{z_L - z_R}{L}\right)^2 + \left(L_2 / L_1 \frac{z_1}{\left(\frac{z_L + z_R}{2}\right)}\right)^2 = 1. \end{cases}$$

解得

$$z_R = a_L L L_1 \sqrt{1 / [L_1^2 (a_R - a_L)^2 + L^2 (L_2 (a_R + a_L) / 2 z_1)^2]},$$

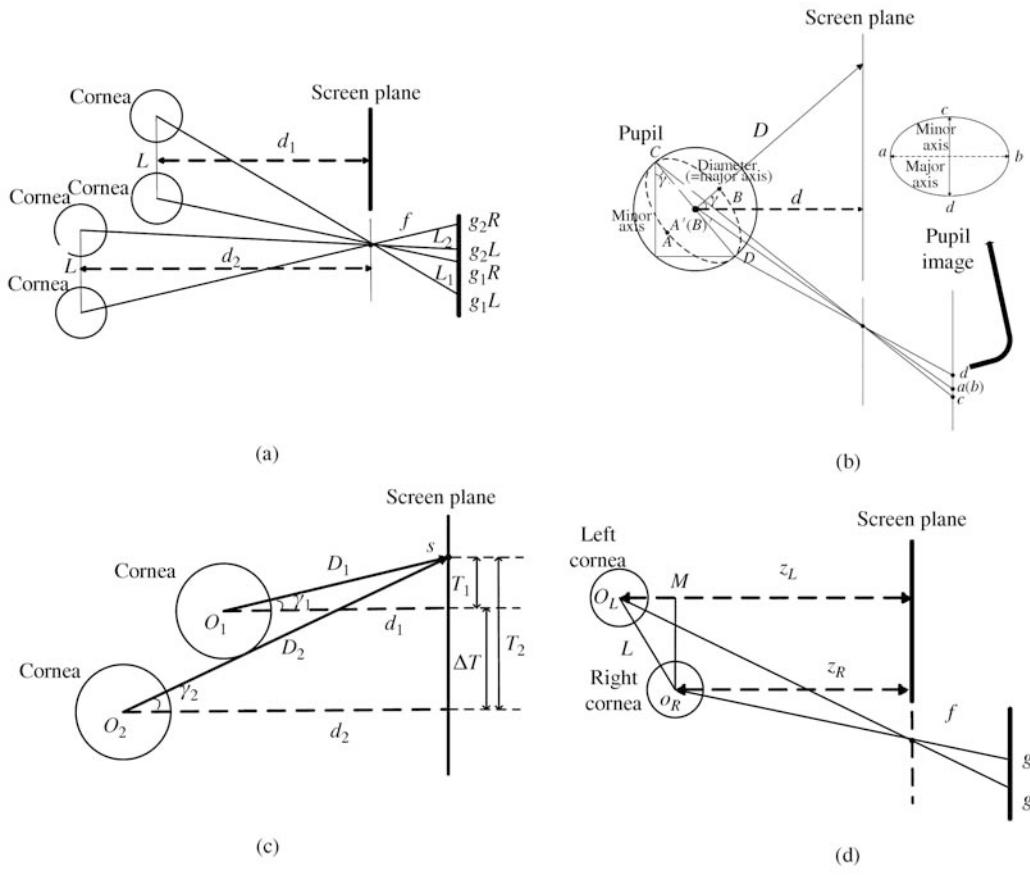


图 5 头部位置对瞳距值的影响

Figure 5 The impact of head position on interpupillary distance

- (a) The impact of head distance on interpupillary distance; (b) the relation between the pupil angle and distance; (c) the impact of head position on the pupil angle; (d) the impact of head rotation on interpupillary distance

$$z_L = a_R L L_1 \sqrt{1/[L_1^2(a_R - a_L)^2 + L^2(L_2(a_R + a_L)/2z_1)^2]},$$

$$L'_2 = L_1 \frac{z_1}{\left(\frac{z_L + z_R}{2}\right)} = \frac{2z_1}{(a_R + a_L)L \sqrt{1/[L_1^2(a_R - a_L)^2 + L^2(L_2(a_R + a_L)/2z_1)^2]}}.$$

所以根据 $\frac{p_1 g_1}{p_2 g_2} = \frac{d_2}{d_1} \cdot \frac{D_2}{D_1}$ 有

$$\frac{p_1 g_1}{p_2 g_2} = \frac{d_2}{d_1} \cdot \frac{d_2 \cos \gamma_1}{d_1 \cos \gamma_2} = \left(\frac{d_2}{d_1}\right)^2 \cdot \frac{\cos \gamma_1}{\cos \gamma_2} = \left(\frac{L_1}{\frac{2z_1}{(a_R + a_L)L \sqrt{1/[L_1^2(a_R - a_L)^2 + L^2(L_2(a_R + a_L)/2z_1)^2]}}}\right)^2 \cdot \frac{\cos \gamma_1}{\cos \gamma_2}$$

$$= \left(\frac{L L_1(a_R + a_L) \sqrt{1/[L_1^2(a_R - a_L)^2 + L^2(L_2(a_R + a_L)/2z_1)^2]}}{2z_1}\right)^2 \cdot \frac{\cos \gamma_1}{\cos \gamma_2},$$

即

$$\frac{p_1 g_1}{p_2 g_2} = \left(\frac{L L_1(a_R + a_L) \sqrt{1/[L_1^2(a_R - a_L)^2 + L^2(L_2(a_R + a_L)/2z_1)^2]}}{2z_1}\right)^2 \cdot \frac{\cos \gamma_1}{\cos \gamma_2}. \quad (3)$$

当用户在使用系统时, 若头部偏离了标定位置, 此时图像中瞳距为 L_2 , 根据前述的瞳距和图像矢量差的关系, 将 L_2 位置的瞳孔与 Purkinje 斑矢量差 p_2g_2 自动转化为标定位置的矢量差 p_1g_1 :

$$p_1g_1 = p_2g_2 \left(\frac{LL_1(a_R + a_L)\sqrt{1/[L_1^2(a_R - a_L)^2 + L^2(L_2(a_R + a_L)/2z_1)^2]}}{2z_1} \right)^2 \cdot \frac{\cos \gamma_1}{\cos \gamma_2}.$$

这样, 用户在瞳距为 L_2 位置上时, 仍然可以得到标定位置对应的矢量值, 由此消除了头动带来的视线估计误差.

4 实验结果

本文通过一系列实验来说明本文方法的效果.

4.1 不同使用者的视线估计结果

在本文开发的视线追踪系统中, 对不同的使用者的视线估计效果进行了测试, 测试过程中系统将实时记录下所有落点信息, 并计算出在分别注视每个目标对象的一定时长内所有落点的位置均值. 考虑到人相对于相机的位置不同, 追踪精度也不尽相同, 因此在测试过程中选取了几个有重要意义的位置. 在标定位置为原点的坐标系下 (单位: mm), 首先在 (0, 0, 0) 处即标定位置处. 选取此位置的目的是测试由标定信息带入的原始误差; 然后分别到 (100, 0, 0) 和 (0, 50, 0) 处, 它们分别为水平和竖直方向正常头动位置; 接着到 (0, 0, -100) 处, 此位置算是离相机较近的位置; 最后到 (0, 0, 100) 处, 这是离相机较远的位置.

对数据进行离线处理, 得到 7 位使用者在几个位置上对盯视屏上所有对象的追踪精度平均水平, 如表 1 所示, 水平方向平均精度 1.4° 左右, 竖直方向平均精度 1.6° 左右. 图 6 为使用者 3 在 (100, 0, 0) 点的视线估计结果.

4.2 与其他视线估计方法的比较

对几种视线估计方法进行了对比. 图 7 给出使用者视线估计结果与真实视线的角度误差. 由图 7 可见, (a) 在标定位置, 几种算法都能正确估计视线; (b) 随着使用者头部远离标定位置, 文献 [9, 17] 方法虽仍能估计视线, 但结果已明显不再准确, 而本文方法在整个头部运动过程中, 都能一直保持对视线的稳定估计能力; (c) 相对于文献 [17] 中二维映射模型视线估计方法和文献 [9] 中直接三维视线估计方法, 本文提出的方法能更为准确和稳定地跟踪目标, 其跟踪结果与目标真实位置的误差最小, 完全能满足实时视线追踪对精度要求.

4.3 人机交互系统

本文基于视线追踪的残疾人环境控制人机交互系统如图 8 所示. 可以使残疾人通过视线追踪系统提供的交互环境得到更加丰富的服务资源. 系统根据主人的要求向服务器发出指令, 来控制灯光控制器、电视机、空调器、热水器、电饭煲控制器、录像机等家电设备的电源控制系统. 实验中准确率在 95% 以上.

表 1 不同使用者的视线估计结果

Table 1 Average gaze estimation accuracy for seven subjects

User	Horizontal accuracy (λ, σ)	Vertical accuracy (λ, σ)
1	(12.00±7.54) mm ($0.86^\circ \pm 0.54^\circ$)	(14.24±8.65) mm ($1.02^\circ \pm 0.62^\circ$)
2	(26.11±10.19) mm ($1.87^\circ \pm 0.73^\circ$)	(32.83±12.42) mm ($2.35^\circ \pm 0.89^\circ$)
3	(26.81±5.58) mm ($1.92^\circ \pm 0.40^\circ$)	(19.83±6.28) mm ($1.42^\circ \pm 0.45^\circ$)
4	(16.89±6.42) mm ($1.21^\circ \pm 0.46^\circ$)	(23.46±7.40) mm ($1.68^\circ \pm 0.53^\circ$)
5	(18.43±4.18) mm ($1.32^\circ \pm 0.30^\circ$)	(24.02±5.72) mm ($1.72^\circ \pm 0.41^\circ$)
6	(21.78±9.21) mm ($1.56^\circ \pm 0.66^\circ$)	(37.44±12.98) mm ($2.68^\circ \pm 0.93^\circ$)
7	(20.11±4.74) mm ($1.44^\circ \pm 0.34^\circ$)	(26.81±7.26) mm ($1.92^\circ \pm 0.52^\circ$)

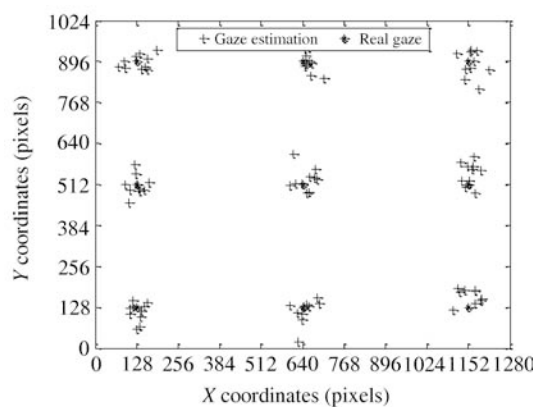


图 6 估计视线与真实视线比较

Figure 6 Gaze estimation results

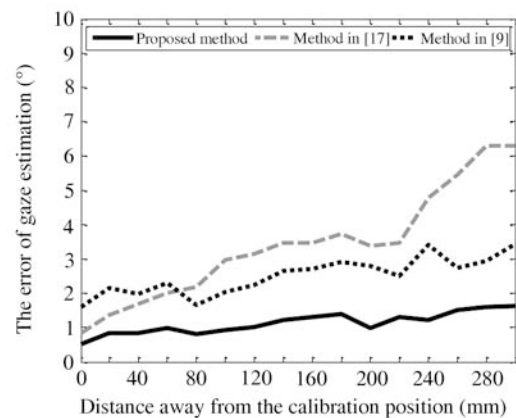


图 7 视线估计结果与真实视线的误差

Figure 7 The angle error of the user's gaze estimation result and the real gaze



图 8 基于视线追踪的人机交互系统

Figure 8 A human-machine interaction system for disabled individuals based on gaze tracking

4.4 与其他系统的比较

与其他视线追踪系统比较结果见表 2, 可以看到, 本文提出了一种单相机单光源条件下的, 单点用户标定的, 可以允许头部较大范围运动的, 高性价比的视线追踪解决方案.

表 2 与其他系统的比较
Table 2 Comparison with other systems^{a),b)}

Methods	Head movement volume (Camera Z axis)	Average accuracy	Configuration	Calibration
[21]	70 mm	1.0°	One stereo camera and multiple lights	One-point calibration
[20]	N/A, but >70 mm	0.6 (only one person)	Two stereo cameras and multiple lights	One-point calibration
[9]	200 mm	1.6°	One stereo cameras and two lights	One-point calibration
[7]	500 mm	5°	Single camera and single light	Multiple-point calibration
[8]	40 mm	0.9°	Single camera and two lights	Multiple-point calibration
[12]	100 mm	1.0°	Single camera and four lights	One-point calibration
Ours	400 mm	1.5°	Single camera and single light	One-point calibration

a) 文献 [8, 12] 需要另外配置立体相机, 用来标定相机, 光源和屏幕的空间三维位置.
b) 本文方法在头部运动小于 50 mm 时, 平均精度在 1.0° 左右.

5 结论与展望

现有单相机单光源条件下的视线追踪系统存在以下几个问题: (a) 精度不高. (b) 头动受限. (c) 标定复杂. 本文提出一种单相机单光源条件下视线追踪的高性价比解决方案, 本方法具有以下几个特点: (a) 通过高质量的特征检测与跟踪, 为高精度视线估计提供了参数依据. (b) 在相机视野范围内, 头部可以自由移动, 头部移动不会带来视线估计精度的明显下降. (c) 通过个体差异补偿机制, 实现了单相机单光源条件下的单点标定. 本文通过高质量的视线估计方法, 使开发的视线追踪系统的精度达到了满足实时人机交互的水平.

参考文献

1 Duchowski T. Eye Tracking Methodology: Theory and Practice. New York: Springer-Verlag, 2003.
2 Jacob R J K. The use of eye movements in human computer interaction techniques: what you look at is what you get. ACM Trans Inf Syst, 1991, 9: 152–169
3 LC Technologies Inc. Eyegaze Systems. Available: <http://www.eyegaze.com/>
4 Eyetech Digital Systems Ind. Eyetech Digital Systems. Available: <http://www.eyetechds.com/>
5 Tobii Technology. Available: <http://www.tobii.se/>
6 Applied Science Laboratories. Available: <http://www.a-s-l.com/>
7 Zhu Z, Ji Q. Eye and gaze tracking for interactive graphic display. Mach Vis Appl, 2004, 15: 139–148
8 Guestrin D, Eizenman M. General theory of remote gaze estimation using the pupil center and corneal reflections. IEEE Trans Biomed Eng, 2006, 53: 1124–1133
9 Zhu Z, Ji Q. Novel eye gaze tracking techniques under natural head movement. IEEE Trans Biomed Eng, 2007, 54: 2246–2260
10 Villanueva A, Cabeza R. A novel gaze estimation system with one calibration point. IEEE Trans Syst Man Cybern Part B-Cybern, 2008, 38: 1123–1138
11 Yoo D H, Chung M J. A novel non-intrusive eye gaze estimation using cross-ratio under large head motion. Comput Vis Image Underst, 2005, 98: 25–51
12 Kang J J, Eizenman M, Guestrin E D, et al. Investigation of the cross-ratios method for point-of-gaze estimation. IEEE Trans Biomed Eng, 2008, 55: 2293–302
13 Hansen D W, Pece A E C. Eye tracking in the wild. Comput Vis Image Underst, 2005, 98: 155–181

- 14 Wang J G, Sung E, Venkateswarlu R. Eye gaze estimation from a single image of one eye. In: Proceedings of the 9 th International Conference on Computer Vision. Beijing, 2003. 136–143
- 15 Wang J, Sung E. Study on eye gaze estimation. IEEE Trans Syst Man Cybern Part B-Cybern, 2002, 32: 332–350
- 16 Xie X, Sudhakar R, Zhuang H. A cascaded scheme for eye tracking and head movement compensation. IEEE Trans Syst Man Cybern Part A-Syst Hum, 1998, 28: 487–490
- 17 Morimoto H, Mimica M. Eye gaze tracking techniques for interactive applications. Comput Vis Image Underst, Special Issue on Eye Detection and Tracking, 2005, 98: 4–24
- 18 Ohno T, Mukawa N, Yoshikawa A. Freegaze: a gaze tracking system for everyday gaze interaction. In: Proceedings of the Symposium on Eye Tracking Research ang Applications. New Orleans, 2002. 125–132
- 19 Beymer D, Flickner M. Eye gaze tracking using an active stereo head. In: Proceedings of the Computer Vision and Pattern Recognition. Madison, 2003. 451–458
- 20 Shih S W, Liu J. A novel approach to 3-D gaze tracking using stereo cameras. IEEE Trans Syst Man Cybern Part B-Cybern, 2004, 34: 234–245
- 21 Morimoto H, Amir A, Flickner M. Detecting eye position and gaze from a single camera and 2 light sources. In: Proceedings of the 16th International Conference on Pattern Recognition. Quebec City, 2002. 314–317
- 22 Jacob R J K. Eye-movement-based human-computer interaction techniques: towards non-command interfaces. Adv Hum-Comput Interact, 1993, 4: 151–190
- 23 Zhang S J, Cao X B, Zhang F, et al. Monocular vision-based iterative pose estimation algorithm from corresponding feature points. Sci China Inf Sci, 2010, 53: 1682–1696

Gaze estimation in a gaze tracking system

ZHANG Chuang*, CHI JianNan*, ZHANG ZhaoHui, GAO XiaoLiang, HU Tao & WANG ZhiLiang

School of Information Engineering, University of Science and Technology Beijing, Beijing 100083, China

*E-mail: zhangchuangtc@163.com, sy_jnchi@126.com

Abstract This article presents a gaze estimation method (GEMHSSO) based on the pupil center cornea reflection (PCCR) technique. Existing PCCR systems suffer from several problems, including the restriction of users' head movement, and the requirement of individual calibration. This paper presents a head position compensation model using a single camera and a single light source, which realizes the analytic compensation of head motion effects on pupil-glint vectors. We then present a transformation model for individual differences to simplify the calibration process to a one-point calibration. On this basis, we establish a novel gaze estimation method that reduces the minimum hardware requirements for accurate estimation to a single camera (not calibrated) and a single light source. Without the need for a complex system, our method has the ability to estimate gaze during natural head movement, and to substantially simplify user calibration. Each step of the proposed method in this paper makes real-time implementation possible, which provides an effective solution for eye-gaze tracking in human-computer interaction systems.

Keywords gaze tracking, gaze estimation, pupil center cornea reflection (PCCR), head compensation, individual calibration