

基于暗瞳图像的人眼视线估计*

张太宁 孟春宁 刘润蓓 常胜江†

(南开大学现代光学研究所, 天津 300071)

(2013年2月6日收到; 2013年3月20日收到修改稿)

虹膜外边缘受眼睑遮挡较为严重时, 会给虹膜中心的准确提取造成很大的困难. 为此, 提出利用放置在相机轴外的红外光源产生的暗瞳图像估计瞳孔中心, 该方法避免了提取虹膜外边缘遇到的遮挡问题. 首先利用角膜反射光斑在相机像面中的位置估计角膜所在球体中心的三维空间坐标, 作为眼球的平动信息; 然后考察瞳孔中心与角膜球体中心在相机成像面投影位置的相对偏移, 作为眼球的转动信息; 最后利用神经网络完成视线特征向量与注视点坐标间的映射. 在人眼区域定位的问题上, 利用两部大视场相机, 采用自适应增强算法和主动表观模型算法实现眼部区域的精确定位, 该步骤可以将提取反射光斑和瞳孔中心需要考虑的图像区域限定在较小的范围内. 实验结果表明, 本文视线估计方法在水平方向上的平均误差为 0.62° , 在竖直方向上的平均误差为 1.05° , 是解决视线点估计的有效方法.

关键词: 暗瞳, 人工神经网络, 自适应增强, 主动表观模型

PACS: 42.30.Sy, 42.30.Tz, 42.30.Va

DOI: 10.7498/aps.62.134204

1 引言

对人眼注视点的估计一直都是人们感兴趣的课题, 高精度注视点估计技术愈加成熟. 在心理学领域, 注视点估计技术已逐渐成为心理评估的重要辅助方式, 在广告心理学、体育心理学、交通心理学、航空心理学等领域中发挥着重要的作用. 除此之外, 注视点估计技术还可作为一种新型的人机交互 (human-computer interaction, HCI) 手段, 为高位截瘫患者或“渐冻人”患者^[1]带来福音. 更广泛的应用还包括: 驾驶员疲劳安全警告系统、战斗机飞行员火控系统、虚拟现实系统 (virtual reality system, VRS)^[2] 和游戏操作系统^[3] 等.

早期的注视点估计方法大多带有一定的侵入性, 例如将隐形眼镜、电极等辅助设备直接与使用者的眼角膜或眼部周围皮肤相接触, 限制了这类方法的使用范围. 随着数码成像技术的飞速发展, 基于计算机视觉 (computer vision, CV) 的非侵入式注视点估计技术受到了越来越多的关注. 其中, 反射光斑法是目前最为常用, 同时也是最为有效的方法

之一. 这种方法利用角膜反射光斑与虹膜/瞳孔中心的相对偏移量进行注视点的估计^[4]. 但虹膜外边缘受眼睑遮挡较为严重, 会给外边缘的椭圆拟合造成较大困难, 因此考察虹膜内边缘, 即瞳孔边缘成为了更加理想的选择. 文献^[5]指出, 利用人眼生物组织对近红外光的反射特性, 通过合理摆放相机与红外光源的相对位置, 能够产生十分清晰的亮/暗瞳图像. 本文利用摆放在远离相机光轴的近红外光源所产生的较为理想的暗瞳图像来对注视点进行估计.

由于眼球在平动和转动过程中均会引起注视点的改变, 因此有必要对眼角膜所在球体中心 A 的三维空间坐标和眼球光轴方向进行估计. 在估计角膜球体中心的问题上, 尽管文献^[6]使用了较为简洁的单相机、双光源系统, 但需要事先提供角膜球体半径 R_A 的尺寸. 在估计眼球光轴方向的问题上, 由于光轴方向由瞳孔中心和角膜球体中心的连线决定, 因此估计光轴方向的问题即等效于瞳孔中心在三维空间中的定位问题. 文献^[7]利用瞳孔中心在两个相机成像面上的投影坐标对真实瞳孔的三维空间位置进行估计, 但考虑到相机拍摄到的瞳孔

* 教育部博士点基金 (批准号: 20090031110033) 和国家自然科学基金 (批准号: 61171027) 资助的课题.

† 通讯作者. E-mail: sjchang@nankai.edu.cn

图像实际上是真实瞳孔经角膜房水折射后的虚像,因此需要将房水的有效折射率 ρ 作为已知参数. 文献 [8] 同样指出,对瞳孔中心的三维坐标估计需要事先知道角膜房水的有效折射率 ρ , 或者已知角膜球体中心与瞳孔中心的距离 K . 尽管文献 [9] 中没有用到 ρ 和 K , 却采用了傍轴近似的策略, 认为瞳孔虚像位置不受观测位置的影响而是永远处于眼球光轴的某一位置上, 忽略了球差的存在.

由于不同使用者的眼球参数 (R_A , ρ 和 K) 各不相同, 且不易测量, 这给视线估计带来了较大误差. 本文采用眼球三维模型和 BP 神经网络相结合的方法来解决眼球参数未知时视线点的估算问题. 首先使用双相机、双反射光斑估计角膜球体中心的三维坐标, 然后考察相机成像面上瞳孔中心与角膜球体中心投影位置的相对偏移, 最后将得到的角膜球体三维坐标和相对偏移量整体作为视线特征向量, 利用 BP 神经网络完成特征向量与注视点在屏幕上坐标之间的映射. 从单人训练、单人测试的实验结果来看, 无论在水平方向还是竖直方向均实现了良好的注视点估计.

2 系统组成

本文的硬件系统如图 1 所示, 其中包括两个小视场相机 C_1 , C_2 , 两个大视场摄像头 C_3 , C_4 , 三个可见光 LED 光源 L_1 , L_2 , L_3 和一个红外光源 L_4 . 两个大视场摄像头用于定位测试者眼部区域在三维空间中的位置; 两个小视场相机用来提取角膜反射光斑和瞳孔中心; 光源 L_1 , L_2 和 L_3 用来产生角膜反射光斑; 放置在相机 C_1 , C_2 轴外位置的红外光源 L_4 用来产生暗瞳图像.

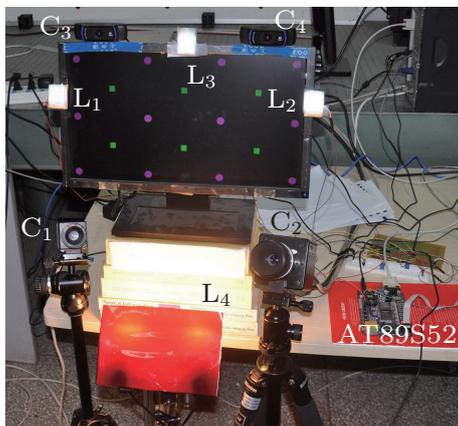


图 1 硬件系统照片

尽管在理论上, 对角膜球体中心 A 的三维坐标估计仅需要两个角膜反射光斑, 但在实际应用中, 当眼球产生大角度转动时, 会发生反射光斑过于靠近虹膜外边缘, 甚至进入巩膜区域而导致光斑丢失的情况. 因此, 在本文的实验系统中, 我们将产生角膜反射光斑的光源数量增加至三个 (即 L_1 , L_2 和 L_3), 使得无论眼球处于哪个转动角度, 至少会有两个反射光斑处于角膜区域内. 此外, 由于靠近巩膜区域的角膜表面与理想球面的偏差较大^[10], 因此当三个反射光斑全部处于角膜区域内时, 选择最靠近瞳孔中心的两个反射光斑来估计角膜球体中心 A .

为排除环境中其他光源在角膜产生的干扰光斑, 本文将辅助光源 L_1 , L_2 和 L_3 设置为闪烁光源. 以 MOXA 公司的串口开发软件包 Pcomm 为基础, 完成上位机 (PC 机) 与下位机 (AT89S52 单片机开发板) 的串行通信, 实现辅助光源在奇数帧为亮, 偶数帧为暗的效果.

在使用本文系统时, 使用者头部距离显示器约为 60 cm. 由于小视场相机 C_1 , C_2 景深和视场的限制, 本系统允许头部运动范围是: 水平 ± 4 cm、竖直 ± 4 cm、前后 ± 2.5 cm. 系统作如下近似: 辅助光源 L_1 , L_2 和 L_3 以点光源近似; 角膜表面以球面近似. 此外, 由于测试者头部的运动范围较小, 相机仅对小范围空间成像, 因此可将相机以针孔相机近似.

3 系统标定

系统标定是本文视线估计算法的必要步骤, 共包含两个部分: 1) 相机焦距以及相机之间相对位置的确定; 2) 辅助光源与相机之间相对位置的确定. 在介绍本文采用的标定方法之前, 对坐标系做如下说明: 将屏幕 (显示器) 的水平方向、竖直方向和垂直于屏幕的方向分别设定为世界坐标系 (world coordinate system, WCS) 的 x 轴、 y 轴和 z 轴, 将屏幕左上角设为 WCS 的原点. 相机坐标系 (camera coordinate system, CCS) 与 WCS 之间的相对位置用旋转矩阵 R 与平移向量 T 来描述.

3.1 相机之间相对位置的确定

本文使用 MATLAB 工具箱^[11] 对相机焦距和相机间的相对位置进行标定. 使用的棋盘格标定板如图 2 所示, 格子尺寸为 $9\text{ mm} \times 9\text{ mm}$, 并在标定板中给出每个格子的编号 (图 2(b)), 用来方便识

别由不同相机采集到的部分棋盘格图像角点间的对应关系. 图 2(c) 和 (d) 分别为相机 C_1 和相机 C_2 在某一时刻对固定在头部允许运动区域内的标定板采集到的图像, 标定板角点位置由工具箱提供的 Harris 角点检测算子得到. 借助标定板中提供的格子编号, 可以得到一组一一对应的标定点. 使用相机对不同姿态的标定板进行多次图像采集, 利用棋盘格角点在各图像中的位置信息, 即可计算出各个相机的内部参数和它们之间相对位置.

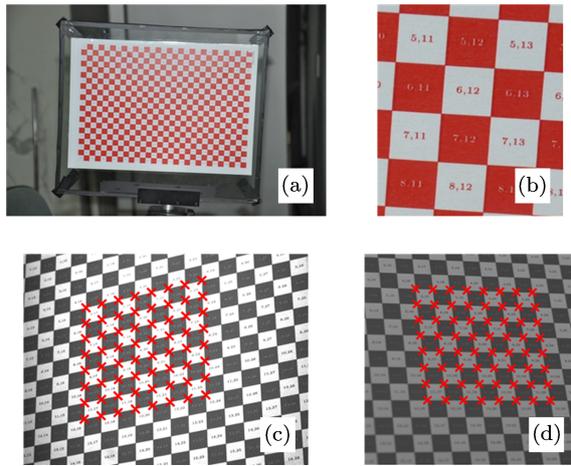


图 2 (a) 棋盘格标定板照片; (b) 标定板的一部分; (c) 和 (d) 标定板角点定位结果

3.2 辅助光源与相机相对位置的确定

由于辅助光源 L_1, L_2 和 L_3 不在相机视场内, 因此无法直接采用双目立体视觉技术定位其在 CCS 下的坐标. 但由于辅助光源 L_1, L_2 和 L_3 被放置在屏幕平面内, 我们可以直接采用刻度尺测量的方法获得 L_1, L_2 和 L_3 在 WCS 下的坐标. 倘若我们知道 WCS 与 CCS 之间的旋转矩阵 R 和平移向量 T , 即可从辅助光源的 WCS 坐标计算出其在 CCS 下的坐标.

假设已知三维空间中的一组点 $\{p_i\} (i = 1, \dots, N)$ 在 WCS 和 CCS 下的坐标分别为 $\{a_i\} (i = 1, \dots, N)$ 和 $\{b_i\} (i = 1, \dots, N)$, 则存在 3×3 的旋转变换矩阵 R 和 3×1 的平移向量 T 满足

$$a_i = R \times b_i + T. \quad (1)$$

考察 $\{a_i\}$ 和 $\{b_i\}$ 的质心 $C_a = (1/N) \sum_{i=1}^N a_i$ 和 $C_b = (1/N) \sum_{i=1}^N b_i$, 二者显然满足

$$C_a = R \times C_b + T. \quad (2)$$

联立 (1) 式和 (2) 式, 得

$$\tilde{a}_i = R \times \tilde{b}_i, \quad (3)$$

其中, $\tilde{a}_i = a_i - C_a; \tilde{b}_i = b_i - C_b$.

对矩阵 $H = \sum_{i=1}^N \tilde{a}_i \tilde{b}_i^T$ 进行奇异值分解 (singular value decomposition, SVD), $H = USV^T$, 即可得到旋转变换矩阵 $R = UV^T$ [12]. 将 R 代入 (2) 式可得平移向量 $T = C_a - R \times C_b$.

文献 [13] 借助一块平面镜来估计一组公共点在 WCS 和 CCS 下的坐标, 但要求成像系统有足够大的景深以便可以对平面镜上的标记点和平面镜对屏幕所成的虚像同时清晰成像, 但本文系统中相机 C_1 和 C_2 的景深有限, 因此无法使用这种方法. 下面简述本文估计一组公共点在 WCS 和 CCS 下坐标的方法.

图 3 为相机 C_1 和 C_2 对固定在头部可运动范围内的一把量程为 50 cm 的刻度尺采集的图像. 在利用 3.1 小节的方法对 C_1 与 C_2 标定完毕之后, 可以得到在图 3(a) 和 (b) 重合区域内的刻度尺刻度在 C_1 (或 C_2) 坐标系下的三维坐标. 尽管相机 C_1 和 C_2 视场较小, 无法采集到刻度尺两个端点的图像, 但可以由刻度尺两个以上任意刻度的延长线推算出两个端点在 CCS 下的三维坐标. 利用一根皮尺测量刻度尺端点与屏幕上三个以上任意标记点之间的距离 (图 4), 利用距离作为约束条件求解方程组即可得到刻度尺端点在 WCS 下的三维坐标. 改变刻度尺的位置, 重复上面的步骤, 从而获得一组用于联系 WCS 与 CCS 的公共点.

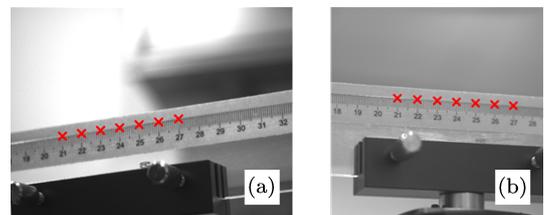


图 3 计算刻度尺端点在 CCS 下的坐标

4 算法流程

4.1 人眼区域定位

本文利用大视场相机 C_3 和 C_4 对三维空间中的人眼区域进行定位. 主动外观模型 (active appearance model, AAM)[14] 是一种在主动形状模型 [15] (active shape model, ASM) 的基础上发展起来的特

征点提取方法. 由于 AAM 解决了 ASM 无法反映目标物体整体纹理信息的问题, 因此在人脸特征点提取的领域应用很广. 利用 AAM 对人脸的整体纹理 (texture) 和整体形状 (shape) 进行建模, 比单纯考虑局部人眼图像信息的匹配方法^[16] 具有更高的准确性. 首先利用自适应增强 (adaptive boosting, AdaBoost) 检测算法^[17,18] 给出人脸形状的初始位置 (图 5(a), (b)), 再通过循环迭代得到最终的收敛结果 (图 5(c)). 在提取出人眼区域在两幅不同视角图像中的二维坐标后, 即可得到人眼在三维空间中的坐标.

4.2 瞳孔中心定位

将眼部区域三维坐标投影至相机 C_1 , 获得如图 6(a) 所示的感兴趣区域 (region of interest, ROI) 图像. 通过合适的阈值对 ROI 的灰度直方图进行处理, 可以得到瞳孔区域 (图 6(b)). 然后从 Canny 边缘图像 (图 6(c)) 中确定瞳孔区域边缘 (图 6(d)). 再利用高亮区域 (图 6(e)) 和角点区域 (图 6(f)) 对瞳孔区域边缘 (图 6(d)) 做进一步筛选, 得到最终的瞳孔边缘 (图 6(g)). 最后采用基于随机采样一致性 (random sample consensus, RANSAC) 的椭圆拟合算法定位瞳孔中心 (图 6(h)).

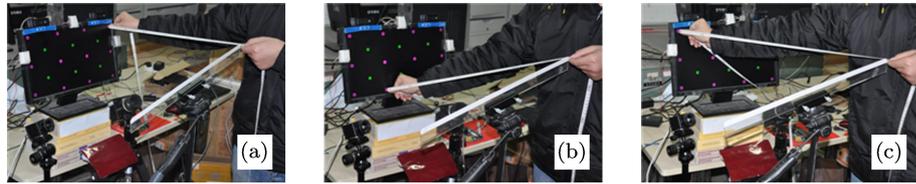


图 4 计算刻度尺端点在 WCS 下的坐标

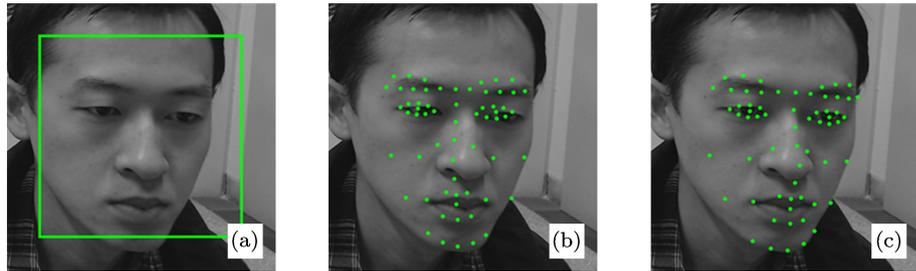


图 5 相机 C_4 对人眼区域的定位 (a) AdaBoost 人脸定位; (b) AAM 的形状初始位置; (c) AAM 的收敛位置

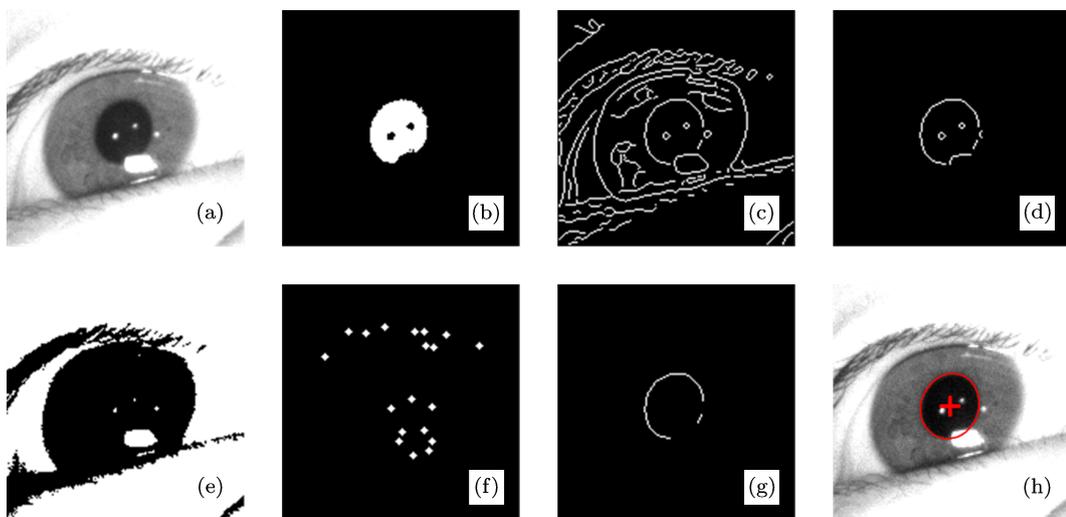


图 6 瞳孔中心定位 (a) ROI 图像尺寸为 151×151 , 瞳孔尺寸约为 40×40 ; (b) 瞳孔区域; (c) Canny 边缘检测; (d) 靠近瞳孔区域的边缘; (e) 高亮区域; (f) 角点区域; (g) 瞳孔边缘; (h) 利用 RANSAC 椭圆拟合定位瞳孔中心

4.3 反射光斑定位

首先利用相邻两帧图像(图 7(a) 和 (b))相减得到差分图像(图 7(c)). 然后使用 Harris 角点检测算子确定角膜反射光斑的大致位置. 最后利用期望最大化(expectation maximization, EM) 算法对每个光斑以一个二维高斯曲面进行拟合, 将高斯曲面的顶点位置作为反射光斑的最终估计位置, 实现反射光斑亚像素(Sub-pixel)精度的定位(图 7(d)).

4.4 基于眼球模型的角膜球体中心定位

相机和辅助光源在 WCS 下的位置固定后, 辅助光源经过角膜反射后投影于相机成像面的光线轨迹如图 8 所示. 以光源 L_2 经过角膜反射后对相机 C_2 成像为例, 从 L_2 发出的一条光线在角膜外表面 G_{22} 处反射, 反射光线通过 C_2 与像平面 P_2 相交于普尔钦(Purkinje)斑点 G'_{22} . 由反射定律可知, 入射光线 $\overline{G_{22}L_2}$ 、反射光线 $\overline{G'_{22}C_2}$ 和法线 $\overline{G_{22}A}$ 三线共面. 若将此面记为 $\Pi_{22} = (L_2 - C_2) \times (G'_{22} - C_2)$, 则角膜所在球体中心 A 满足

$$\Pi_{22} \cdot (A - C_2) = 0, \quad (4)$$

类似的, 图 8 中还存在另外三个包含角膜球体中心 A 的平面 Π_{11} , Π_{12} 和 Π_{21} , 因此可以通过求解如下方程组得到 A 在三维空间中的坐标

$$\begin{aligned} \Pi_{11} \cdot (A - C_1) &= 0, \\ \Pi_{12} \cdot (A - C_2) &= 0, \\ \Pi_{21} \cdot (A - C_1) &= 0, \\ \Pi_{22} \cdot (A - C_2) &= 0, \end{aligned} \quad (5)$$

上面方程组中 $\{L_k\}$, $\{C_l\}$ 和 $\{G'_{kl}\}$ ($k, l = 1, 2$) 的坐标由第 3 小节系统标定过程获得. 然后将 A 投影到相机成像面上, 作为计算瞳孔中心相对偏移量的基准点.

最后提取的视线特征向量为 $G_1 = (T_x, T_y, T_z, S_{H1}, S_{V1})$ 和 $G_2 = (T_x, T_y, T_z, S_{H2}, S_{V2})$. 其中 (T_x, T_y, T_z) 表示角膜球体中心 A 在 WCS 下三维坐标, 描述了眼球的平动; $(S_{H1}, S_{V1}) = P_{C1} - I_{C1}$, 表示在相机 C_1 的成像面上瞳孔中心 P_{C1} 与角膜球体中心的投影位置 I_{C1} 之间的相对偏移(图 8(a)), 描述了眼球的转动. 同理, $(S_{H2}, S_{V2}) = P_{C2} - I_{C2}$, 表示在相机 C_2 的成像面上瞳孔中心 P_{C2} 与角膜球体中心的投影位置 I_{C2} 之间的相对偏移(图 8(b)).

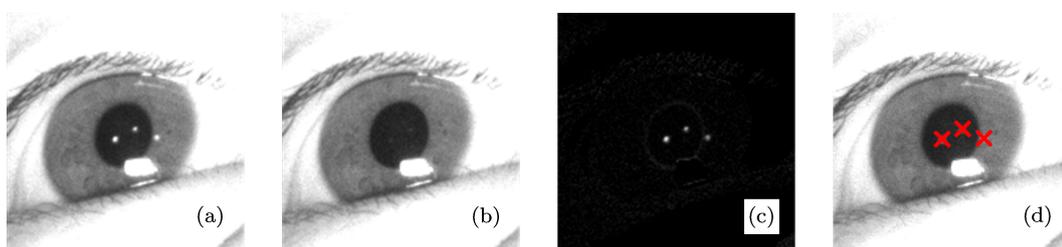


图 7 (a) 奇数帧图像; (b) 偶数帧图像; (c) 差分图像; (d) 角膜反射光斑定位

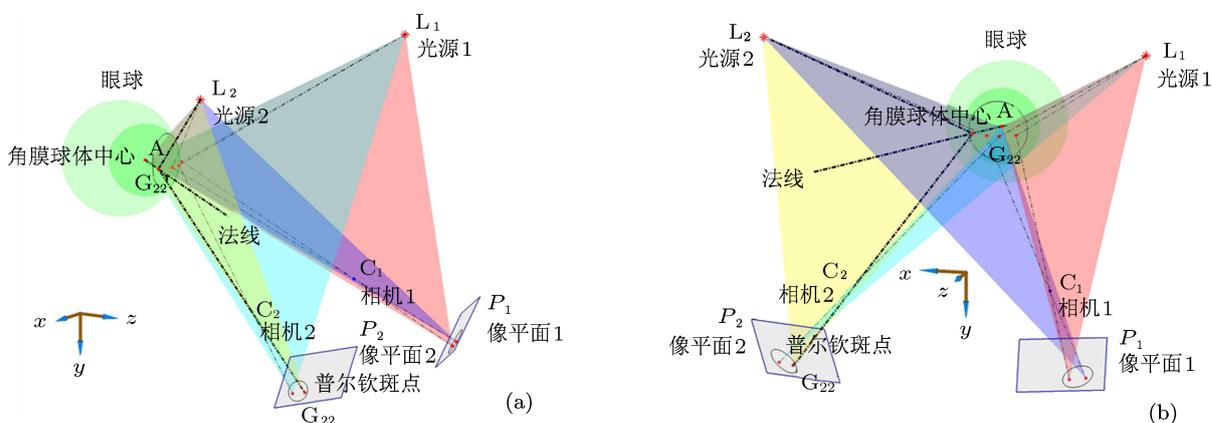


图 8 角膜、相机、光源相对位置示意图 (a) 视角 1; (b) 视角 2

利用 BP 前馈神经网络建立视线特征向量 G_1 (或 G_2) 和注视点 (X, Y) 之间的非线性映射关系, 实现对注视点的估计. 本文采用 5-6-2 的网络结构, 隐含层采用双曲正切 S 形传输函数, 输出层采用线性传输函数. 使用 G_1 和 G_2 分别进行注视点的估计, 将两个估计结果的平均值作为最终的注视点估计结果.

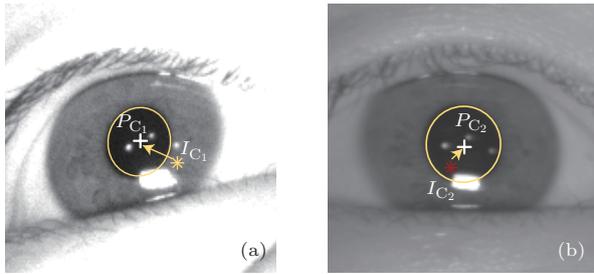


图9 角膜球体中心在相机成像面的投影位置与瞳孔中心的相对偏移向量

5 实验结果

实验中测试屏位于使用者正前方约 60 cm 处, 物理尺寸为: 宽 408 mm、高 229.5 mm. 测试屏中共包含 18 个标记点, 均匀分布在测试屏中 (图 10), 其中的 12 个圆形标记点用于训练, 在测试阶段除了再次使用 12 个圆形标记点外又新增加 6 个方

形标记点. 系统标定后得到的相机参数如表 1 所示. 其中, 相机 C_1 和 C_2 的仰角约 20° ; 相机 C_3 和 C_4 的俯角约 -15° .

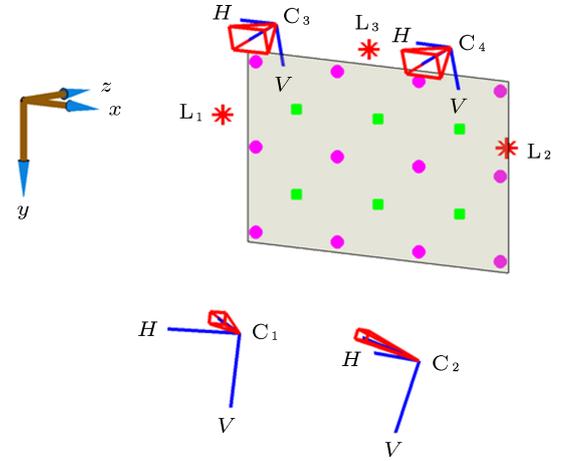


图 10 实验布局斜视图

本文采用单人训练、单人测试的策略, 共有 3 名使用者. 每名使用者的训练样本和测试样本均由 10 组图像组成, 训练样本数为 $120 (= 12 \times 10)$ 幅, 测试样本数为 $180 (= 18 \times 10)$ 幅. 对 1 号使用者的视线估计结果如图 11 所示. 表 2 给出了对 3 名使用者的视线估计误差均值.

表 1 由系统标定得到的相机参数. 俯仰角指相机光轴与 WCS 的 XOZ 平面之间的夹角; 方位角指相机光轴在平面 XOZ 上的投影与 Z 轴负方向的夹角

	C_1	C_2	C_3	C_4
相机型号	Microvision MV-VE120SM	DALSA CA-D7-1024T	Logitech C920	Logitech C920
分辨率/pixel	1280×960	1024×1024	1280×720	1280×720
镜头焦距/像素点中心距	6034	5216	1303	1276
水平方向视场角/ $^\circ$	12.1	11.2	52.3	53.3
竖直方向视场角/ $^\circ$	9.1	11.2	30.9	31.5
俯仰角/ $^\circ$	19.2	22.9	-14.5	-15.8
方位角/ $^\circ$	-23.1	13.0	-2.6	0.5

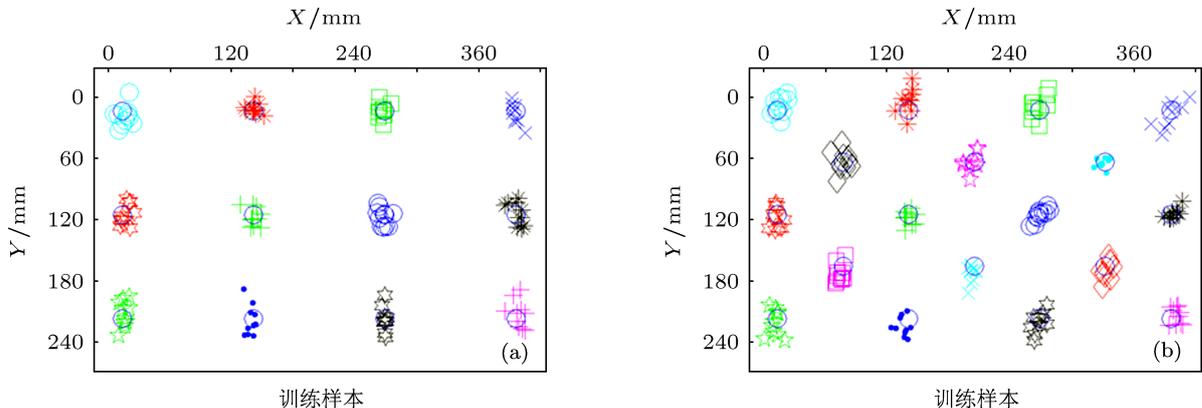


图 11 使用者 1 的目标注视点与估计注视点的标记图 (a) 训练样本误差分布图; (b) 测试样本误差分布图

眼球的转动信息除了可以利用如本文所使用的瞳孔中心相对偏移外, 还可以利用拟合椭圆的畸变来描述^[19]. 由于椭圆畸变由长短轴之比 $a_{\text{minor}}/a_{\text{major}}$ 和长轴与竖直方向的夹角 θ 描述, 因此在 4.4 节出现的视线特征向量 G_1 可以用 $(T_X, T_Y, T_Z, a_{\text{minor}1}/a_{\text{major}1}, \theta_1)$ 替代, 同理 G_2 可替换为 $(T_X, T_Y, T_Z, a_{\text{minor}2}/a_{\text{major}2}, \theta_2)$. 采用与 4.4 节中相同的 5-6-2 网络结构, 并使用此处的 G_1 和 G_2 分别进行注视点的估计, 再将两个估计结果求平均后得

到的结果如表 3 所示.

文献 [20] 同时利用瞳孔中心的相对偏移和拟合椭圆的畸变来描述眼球的转动. 当我们利用带有冗余信息的视线特征向量进行视线估计时, 得出的结果如表 4 所示. 信息冗余的视线特征向量如下:

$$G_1 = (T_X, T_Y, T_Z, S_{H1}, S_{V1}, a_{\text{minor}1}/a_{\text{major}1}, \theta_1), \quad (6a)$$

$$G_2 = (T_X, T_Y, T_Z, S_{H2}, S_{V2}, a_{\text{minor}2}/a_{\text{major}2}, \theta_2), \quad (6b)$$

对应的神经网络结构为 7-6-2.

表 2 当眼球的转动信息由瞳孔中心相对偏移描述时, 三名使用者视线估计的均方根角度误差

	训练样本		测试样本	
	水平方向/(°)	竖直方向/(°)	水平方向/(°)	竖直方向/(°)
使用者 1	0.47	1.01	0.55	0.95
使用者 2	0.72	0.77	0.64	1.12
使用者 3	0.62	0.90	0.68	1.08
平均注视精度	0.60	0.89	0.62	1.05

表 3 当眼球的转动信息由拟合椭圆的畸变描述时, 三名使用者视线估计的均方根角度误差

	训练样本		测试样本	
	水平方向/(°)	竖直方向/(°)	水平方向/(°)	竖直方向/(°)
使用者 1	1.66	1.85	2.17	1.81
使用者 2	1.99	2.23	2.84	2.61
使用者 3	2.37	2.69	2.21	2.68
平均注视精度	2.00	2.26	2.41	2.37

表 4 当眼球的转动信息由瞳孔中心的相对偏移和拟合椭圆的畸变共同描述时, 三名使用者视线估计的均方根角度误差

	训练样本		测试样本	
	水平方向/(°)	竖直方向/(°)	水平方向/(°)	竖直方向/(°)
使用者 1	0.80	0.87	0.90	1.06
使用者 2	0.78	0.56	1.10	0.88
使用者 3	0.78	0.65	1.06	0.96
平均注视精度	0.79	0.69	1.02	0.97

从表 2 与表 3 的对比可以看出, 当改用椭圆畸变来描述眼球转动时, 视线估计精度明显下降. 我们认为, 精度的明显下降是由于椭圆拟合得到的长轴与竖直方向的夹角 θ 精度不高导致的. 尤其在瞳孔边缘接近正圆的情况下, 更加难以对倾角 θ 进行准确的估计. 从表 2 与表 4 的对比可以看出, 当改用信息冗余的视线特征向量进行视线估计时, 估计

精度并没有改善, 却加大了神经网络的结构, 延长了网络训练的时间. 综上所述, 仅使用瞳孔中心的相对偏移描述眼球的转动是最佳的选择.

6 结论

本文提出了一种基于暗瞳图像的视线估计方法, 克服了边缘提取过程中的眼皮遮挡问题, 并采

用三维模型与神经网络相结合的方法避免了需要事先知道眼球参数的问题. 其次, 在系统标定过程中, 利用皮尺测量的方法巧妙的解决了辅助光源在 CCS 下的定位问题. 最后, 采用两部大视场相

机实现了人眼区域的准确定位, 为后续的视线特征向量提取提供方便. 实验结果表明, 本文视线估计的水平方向平均误差为 0.62° , 垂直方向平均误差为 1.05° , 是一种有效的解决方案.

-
- [1] Liu S S, Rawicz A, Rezaei S, Ma T, Zhang C, Lin K, Wu E 2012 *J. Med. Biol. Eng.* **32** 111
- [2] Lim C J, Kim D 2012 *Sens. Actuators A: Phys.* **185** 151
- [3] Corcoran P M, Nanu F, Petrescu S, Bigioi P 2012 *IEEE Trans. Consum. Electr.* **58** 347
- [4] Laura S S, Villanueva A, Cabeza R 2012 *IEEE Trans. Biomed. Eng.* **59** 2235
- [5] Ebisawa Y 1998 *IEEE Trans. Instrum. Meas.* **47** 948
- [6] Villanueva A, Cabeza R 2008 *IEEE Trans. Biomed. Eng.* **55** 2812.
- [7] Nagamatsu T, Kamahara J, Iko T, Tanaka N 2008 *Proceedings of the 2008 symposium on Eye tracking research & applications Savannah, Georgia, March 26–28, 2008* p95
- [8] Guestrin E D, Eizenman M 2006 *IEEE Trans. Biomed. Eng.* **53** 1124
- [9] Zhu Z W, Ji Q 2007 *IEEE Trans. Biomed. Eng.* **54** 2246
- [10] Nagamatsu T, Iwamoto Y, Kamahara J, Tanaka N, Yamamoto M 2010 *Proceedings of the 2010 symposium on Eye tracking research & applications Austin, Texas, March 22–4, 2010* p255
- [11] Bouquet J Y http://www.vision.caltech.edu/bouquetj/calib_doc/ 2012
- [12] Arun K S, Huang T S, Blostein S D 1987 *IEEE Trans. Pattern Anal.* **9** 698
- [13] Pogalin E, Redert A, Patras I, Hendriks E A 2006 *Proceedings of the 3rd International Symposium on 3D Data Processing, Visualization, and Transmission Chapel Hill, North Carolina, June 14–16, 2006* p57
- [14] Matthews I, Baker S 2004 *Int. J. Comput. Vision* **60** 135
- [15] Zhang W, Cheng B, Zhang B 2012 *Acta Phys. Sin.* **61** 060701 (in Chinese) [张伟, 成波, 张波 2012 物理学报 **61** 060701]
- [16] Zhang Z T, Zhang J S 2010 *Chin. Phys. B* **19** 104601
- [17] Viola P, Jones M J 2004 *Int. J. Comput. Vision* **57** 137
- [18] Zhang L, Chu R F, Xiang S M, Lian S C, Li S Z 2007 *Proceedings of 2007 International Conference on Biometrics Seoul, Korea, August 27–29, 2007* p11
- [19] Wang J G, Sung E, Venkateswarlu R 2005 *Comput. Vis. Image Und.* **98** 83
- [20] Zhang C, Chi J N, Zhang Z H, Wang Z L 2010 *Acta Automatica Sinica* **36** 1051 (in Chinese) [张闯, 迟健男, 张朝晖, 王志良 2010 自动化学报 **36** 1051]

Eye gaze tracking based on dark pupil image^{*}

Zhang Tai-Ning Meng Chun-Ning Liu Run-Bei Chang Sheng-Jiang[†]

(*Institute of Modern Optics, Nankai University, Tianjin 300071, China*)

(Received 6 February 2013; revised manuscript received 20 March 2013)

Abstract

The accurate localization of iris center is difficult since the outer boundary of iris is often occluded significantly by the eyelids. In order to solve this problem, an infrared light source un-coaxial with the camera is used to produce dark pupil image for pupil center estimation. Firstly, the 3D position of the center of cornea curvature, which is used as translational movement information of eyeball, is computed using two cameras and the coordinates of two cornea reflections on the cameras' imaging planes. Then, the relative displacement of pupil center from the projection of the cornea curvature center on 2D image is extracted, describing the rotational movement of the eyeball. Finally, the feature vector is mapped into coordinates of gazing point on the screen using artificial neural network. As for the eye region detection problem, two wide-view webcams are used, and adaptive boosting + active appearance model algorithm is adopted to limit the region of interest within a small area. The result of our experiment shows that the average root-mean-square error is 0.62° in horizontal direction and 1.05° in vertical direction, which demonstrates the effectiveness of our solution in eye gaze tracking.

Keywords: dark pupil, artificial neural network, adaptive boosting, active appearance model

PACS: 42.30.Sy, 42.30.Tz, 42.30.Va

DOI: 10.7498/aps.62.134204

^{*} Project supported by the Ph.D. Programs Foundation of Ministry of Education of China (Grant No. 20090031110033), and the National Natural Science Foundation of China (Grant No. 61171027).

[†] Corresponding author. E-mail: sjchang@nankai.edu.cn