

一种快速的非侵入式眼动跟踪方法^①

徐洪飏, 陈灵坚, 吴曦华

(温州医学院 信息与工程学院, 温州 325035)

摘要: 针对眼控人机交互系统对实时性要求较高的特点, 提出了一种快速、非侵入式的眼动跟踪方法。首先融合 CamShift 跟踪算法和 AdaBoost 型目标检测方法快速地分割出眼睛所在矩形区域; 然后在第一帧视频图像中精确定位虹膜并建立虹膜模板; 最后通过放大或缩小虹膜模板自适应地在后续各帧视频图像中搜索匹配以预测虹膜位置。虹膜中心相对眼角或鼻孔的位移可用于控制鼠标指针在屏幕上移动。在分辨率为 640×480 的视频上, 取得了 32 帧/秒的跟踪速度, 正确率达到 95.5%。实验结果表明方法是有效的, 可直接应用于眼控人机交互系统。
关键词: 无障碍技术; 非侵入式眼动跟踪; 虹膜定位; CamShift 算法; AdaBoost 算法

Rapid and Non-Intrusive Eye-Tracking Method

XU Hong-Biao, CHEN Ling-Jian, WU Tun-Hua

(School of Information & Engineering, Wenzhou Medical College, Wenzhou 325035, China)

Abstract: To meet the real-time requirement of eye-controlled human-computer interaction systems, a fast and non-intrusive eye-tracking method was proposed. Firstly, CamShift algorithm and AdaBoost algorithm were combined to extract eye regions. Secondly, iris was located from the eye region in the first video frame. And the iris region was used as a template. Finally, the iris template was scaled according to new size of eye region in the subsequent video frames. Then the scaled template was used for matching in the eye region. The relative displacement between iris center and canthus or nostril can be used to control the movement of mouse pointer on computer screen. The experiments on a video of resolution at 640*480 result a tracking speed of 32 frames per second, and the correct rate is 95.5%. Results prove the effectiveness and efficiency of this method. So, this method can be directly applied to eye-controlled human-computer interaction systems.

Key words: accessible technology; Non-Intrusive eye tracking; iris location; CamShift algorithm; AdaBoost algorithm

1 引言

眼睛是人类从外界获取信息的重要器官, 对于存在上肢运动障碍的残疾人群(无法用手操控鼠标和键盘), 眼睛的作用显得尤为重要, 因为手势识别、手指活动识别等都无法使用, 而语音识别又不适合于控制鼠标。Gorodnichy^[1]等人提出的基于鼻头跟踪的控制方法虽然也无需手部操作, 但极易引起颈部疲劳, 且不适用于患有肌萎缩性脊髓侧索硬化症(Amyotrophic Lateral Sclerosis)的残疾人(因为他们只能通过眼睛移动或轻微的头部摆动来表达信息)。因此, 研究如何

利用眼睛运动信息来帮助残疾人实现其与计算机交互信息具有非常重要的意义。

眼动跟踪(Eye Tracking)^[2,3]是指通过跟踪眼球的运动获取眼球注视的方向、注视时间以及注视点的移动轨迹。目前主要的眼动跟踪方法有红外线跟踪法、脑机接口(BCI, Brain-Computer Interface)法、眼电图(EOG)法、视线跟踪法。红外线跟踪法^[4]先用红外或近红外线照射眼睛并将反射图像用带红外滤镜的摄像头记录下来, 再从记录下来的图像中快速检测出瞳孔位置并对其跟踪。这种方法具有准确、快速的优点,

① 基金项目:国家自然科学基金(60873179,11005081);浙江省自然科学基金(Y1110322);浙江省教育厅科技项目(Y201016244);浙江省青年教师资助计划(2010);校科研启动项目(QTJ09004,QTJ09009)

收稿时间:2011-05-03;收到修改稿时间:2011-06-01

但由于红外线发射器和眼睛的距离较近, 需要一个头上装置将发射器和摄像头一起固定于眼睛附近。如图 1 左侧所示, 这种解决方案属于侵入式 (Intrusive), 用户可能因为佩戴这种装置而感到不适。另外, 红外线跟踪法易受可见光影响, 红外或近红外线长时间近距离照射眼睛是否会对眼睛造成伤害亟待进一步研究。脑机接口法^[5]首先通过紧贴头皮的一组传感器采集脑电信号, 再进行实时处理并转为控制信号控制外部应用程序。脑机接口尽管处理速度快, 但对用户的干扰较大, 且需要长时间的训练才能使用, 大部分处于临床阶段。眼电图法^[6]与脑机接口法类似, 需要通过传感器采集眼睛附近的电信号, 眼动跟踪精度较高, 且对用户的干扰降低了许多, 但紧贴眼部的传感器对于多数用户来说还是比较难以适应的。与上述三类方法不同, 视线跟踪法^[7]属于非侵入式 (Non-Intrusive), 采用普通摄像头作为输入设备, 不需要佩戴任何装置, 如图 1 右侧所示, 用户可以较自然地使用跟踪器, 并且降低了成本。有关视线跟踪的研究已有很多, 粒子滤波、卡尔曼滤波、3D 模型匹配等复杂算法纷纷被引入该领域, 但目前仍然无法彻底解决速度、准确性和鲁棒性三者之间的矛盾。这主要是因为人类视觉内在机制尚未探明, 此外, 环境光照变化、眼镜反光、头部姿态变化 (各方向旋转)、距离变化等干扰因素都会影响跟踪的速度和准确性。

由于眼动控制鼠标操作对跟踪速度要求较高但不需要太高的跟踪精度, 我们提出了一种视线跟踪类型的眼动跟踪方法, 仅用单摄像头捕捉头部视频。为了实现高速跟踪, 我们融合 CamShift 跟踪算法、AdaBoost 型目标检测方法快速地分割出眼睛所在矩形区域, 再通过缩放虹膜模板自适应地在上述眼睛区域中搜索匹配。为了建立初始虹膜模板, 我们提出了一种准确定位虹膜的方法。文献[8, 9]介绍了我们先前的相关工作, 描述了基于 AdaBoost 算法和 C-V 模型的多姿态人脸检测与人脸特征提取方法。



图 1 侵入式(左)与非侵入式(右)眼动跟踪设备

2 基于 CamShift 算法和 AdaBoost 算法的眼睛区域跟踪

CamShift (Continuously Adaptive MeanShift, 连续自适应均值漂移算法) 建立在 MeanShift 算法基础上, 利用人脸颜色概率分布在各帧视频图像中预测目标窗口的位置、大小和旋转角度。与面向目标跟踪的 MeanShift 算法不同, CamShift 可自动调整搜索窗口大小。CamShift 算法的复杂度很低, 因此可以做到实时跟踪。CamShift 算法描述如下:

1) 初始化搜索窗的大小 s 和中心位置 (x_c, y_c) 。

2) 计算以 (x_c, y_c) 为中心的边长为 $s \times 1.1$ 的区域内颜色概率分布。色调量化为 256 级。

3) 利用 Mean-Shift 算法获得搜索窗新位置, 并作为下一帧图像中搜索窗的初始位置。

4) 重新设置搜索窗的大小 $s = 2\sqrt{M_{00}/256}$ 。

5) 估计跟踪目标的方向和尺度:

$$a) \quad M_{20} = \sum_x \sum_y x^2 I(x, y), \quad M_{02} = \sum_x \sum_y y^2 I(x, y), \\ M_{11} = \sum_x \sum_y xy I(x, y)$$

$$b) \quad a = M_{20} / M_{00} - x_0^2, \quad b = 2(M_{11} / M_{00} - x_0 y_0), \\ c = M_{02} / M_{00} - y_0^2$$

c) 包含目标的椭圆长轴的方向角为: $\theta = \frac{1}{2} a \tan\left(\frac{b}{a-c}\right)$

d) 包含目标的椭圆区域的长轴和短轴分别为:

$$l = \sqrt{\frac{(a+c) + \sqrt{b^2 + (a-c)^2}}{2}}, \quad w = \sqrt{\frac{(a+c) - \sqrt{b^2 + (a-c)^2}}{2}}$$

6) 重复步骤 2) 到 5), 直到用户中断跟踪或者视频结束。

基于 AdaBoost 算法的目标检测器是目前检测速度最高、鲁棒性最好的目标检测器。由 Viola 等人提出的层叠型物体检测器的构成方法^[10]: 首先利用积分图计算样本中包含的所有 Haar 型特征; 然后根据 AdaBoost 算法训练这些 Haar 型特征, 得到一组强分类器, 每个强分类器由若干弱分类器组成, 弱分类器与筛选出的分类能力较强的 Haar 型特征一一对应; 最后将这组强分类器串联为一个层叠分类器。这种构造层叠型检测器的方法具有通用性, 可以根据样本的不同得到各种物体检测器, 如人脸检测器、嘴巴检测器等。文献[8, 9]中介绍了我们之前关于 AdaBoost 型人脸、眼睛检测器的训练学习方法以及实验结果。因此,

这里直接利用之前训练好的 AdaBoost 型人脸检测器在第一帧视频图像中定位人脸区域, 然后以区域边长的 0.5 倍缩小得到一个同心矩形, 将该矩形作为 CamShift 的初始搜索窗并计算其内部的颜色分布图作为参考颜色分布。在定位人脸区域之后, 再利用 AdaBoost 型眼睛检测器分割出眼睛所在的矩形区域, 后续的虹膜定位就在该区域中进行。眼睛区域跟踪结果如图 2 所示。

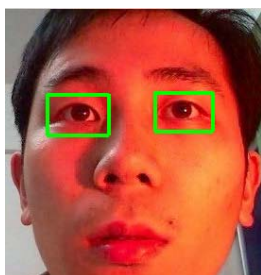


图 2 基于 CamShift 和 AdaBoost 算法的眼睛区域跟踪

3 虹膜定位与跟踪

本节的目标是从首帧人脸视频图像的眼睛区域中提取虹膜边缘并将虹膜内的灰度和分辨率信息保存为虹膜模板。在后续视频帧中, 首先根据第 2 节介绍的方法定位眼睛区域, 然后利用虹膜模板在眼睛区域中自适应地模板匹配并最终获取虹膜的新位置。为了确保虹膜定位成功, 开始跟踪前眼睛须正视摄像头。虹膜边缘为圆形并且与巩膜区(眼白)有显著的颜色和灰度差, 根据这个特点我们提出一种精确定位虹膜的方法。本方法按以下 4 个步骤展开:

1) 图像预处理: 应用高斯低通滤波平滑眼睛区域, 实验中窗口大小为 4、方差为 5;

2) 筛选圆心(即虹膜中心)候选点: 快速排除灰度非局部最小的点, 这些点不可能是圆心, 因为虹膜中心必定是灰度较低的点; 假设眼睛区域的高为 h , 根据实验测定, 虹膜半径 r 在 $r_{\min} = \text{Int}(0.1 * h)$ 到 $r_{\max} = \text{Int}(0.3 * h)$ 之间 (Int 为取整运算), 因此还要把与眼睛区域边缘距离小于 r_{\min} 的候选圆心排除掉。

3) 虹膜粗定位: 对于每个候选圆心, 半径 $r \in [r_{\min}, r_{\max}]$, 假设存在 $n = r_{\max} - r_{\min} + 1$ 个圆可能与虹膜边缘匹配。首先分别计算这 n 个圆上所有像素的平均灰度(圆用多边形近似, 实验中边数达 600 条), 记为 $\{I_1, I_2, I_3, \dots, I_n\}$; 然后计算这些平均灰度的差分 $\text{diff} = \{d_1, d_2, \dots, d_n\} = \text{Abs}\{0, I_2 - I_1, I_3 - I_2, \dots, I_n - I_{n-1}\}$ (Abs 为绝

对值运算), 并对 diff 做高斯平滑; 最后将具有最高灰度差分值的候选圆作为虹膜粗定位的结果。

4) 虹膜细定位: 假设第 3 步粗定位的虹膜中心为 (x_0, y_0) , 以 (x_0, y_0) 为中心的 10×10 的邻域内每个点为候选圆心, 在 (r_{\min}, r_{\max}) 范围内寻找具有最大差分值的圆作为细定位的结果。

虹膜跟踪的方法是: 首先利用上述虹膜定位方法从首帧视频图像中分割出虹膜图像区域(实验中采用正方形区域), 并将该图像区域保存为虹膜模板; 然后根据后续帧中眼睛区域大小与虹膜模板所在眼睛区域大小的比值放大或缩小虹膜模板, 并将缩放后的虹膜模板在眼睛区域中搜索匹配。

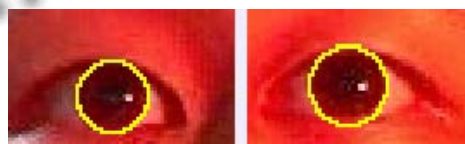


图 3 虹膜定位示例, 从眼睛区域中分割出虹膜图像区域并将其作为虹膜模板

4 实验与分析

4.1 实验设计

实验程序采用微软 Visual Studio 2008 结合 Intel OpenCV Library 开发而成, 操作系统为 Windows XP, CPU 为 Intel Core2 E8400 @4.0G (双核, 超频), 4GB DDR2 800 双通道内存。通过训练获得的 AdaBoost 型眼睛检测器由 17 层强分类器层叠而成, 包含 414 个 Haar 型特征。

摄像头为罗技快看太空版, CCD 感光元件。摄像头位于眼睛的斜下方, 分辨率为 640×480 , 帧率为 30 帧/秒, 环境光照均匀。首先随机截取 200 帧头部轻微晃动、跟踪距离前后变化的视频图像用于测试跟踪速度和准确性; 再截取 200 帧头部大角度旋转的视频图像用于测试跟踪方法对头部旋转的适应性。

4.2 实验结果

在评价跟踪性能时, 只要由跟踪预测所得的虹膜中心没有偏离实际虹膜区域就判定为跟踪成功。图 4 给出了部分实验结果。具体结果如下:

1) 头部小角度旋转下的跟踪速度和性能测试结果: 对于 200 帧 640×480 的视频图像, 跟踪速度达到 32 帧/秒, 正确率达到 95.5%, 即 9 帧跟踪失败。跟踪准确性不受跟踪距离变化的影响。

2) 头部大角度旋转(平面旋转、左右侧转、抬头俯视)下的跟踪结果: 对于 200 帧 640×480 的视频图像, 跟踪速度依然为 32 帧/秒, 但跟踪性能波动较大。当头部平面旋转角度大于 60°、侧转角度大于 30°、向上旋转大于 45°或向下旋转大于 30°时, 准确性急剧下降。

4.3 实验分析

1) 跟踪速度: 由于 CamShift 跟踪算法和 AdaBoost 型检测器的运算量很小, 加上仅在眼睛区域中自适应地模板匹配, 因此跟踪速度达到了实时的要求。另外, 视频分辨率较低也是速度较快的一个原因。

2) 跟踪准确性和鲁棒性: 头部小角度旋转、跟踪距离变化对准确性几乎没有影响, 因为 CamShift 算法本身就可以较准确地预测头部的平面旋转角度, 且 AdaBoost 算法对于头部左右侧转和上下旋转的适应角度较大^[8]。对于眼镜反光问题, 本方法没有提出解决方案, 所做的实验都是在佩戴眼镜的情况下完成的。另外, 在眼睛闭合时(眨眼), 跟踪也会失败, 因为本方法没有闭眼时虹膜位置的预测机制。因此, 在不明显降低跟踪速度的前提下, 准确判断眼睛状态、处理眼镜反光问题都是今后亟需解决的问题。



图 4 眼动跟踪示例, 虹膜中心用星号标识

在分辨率为 640×480 的视频上, 取得了 32 帧/秒的跟踪速度

5 结语

针对眼控人机交互系统对速度要求较高的特点, 提出了一种计算复杂度较低的眼动跟踪方法。首先融

合 CamShift 跟踪算法和 AdaBoost 型目标检测方法快速地分割出眼睛所在矩形区域; 然后在第一帧视频图像中精确定位虹膜并建立虹膜模板; 最后通过放大或缩小虹膜模板自适应地在后续各帧视频图像中搜索匹配以预测虹膜位置, 虹膜中心相对眼角或鼻孔的位移可用于控制鼠标指针在屏幕上移动。在分辨率为 640×480 的视频上, 取得了 32 帧/秒的跟踪速度, 实现了实时跟踪。所提出的虹膜定位方法速度上较慢, 但精度较高, 适于建立虹膜模板。本方法对头部旋转、跟踪距离变化具有较好的适应性。但无法处理眼镜反光和闭眼问题, 这些是今后本方法亟需改进的地方。

参考文献

- 1 Gorodnichy DO, Roth G. Nouse use your nose as a mouse perceptual vision technology for hands-free games and interfaces. *Image and Vision Computing*, 2004,22:931-942.
- 2 Hansen DW, Ji Q. In the Eye of the Beholder: A Survey of Models for Eyes and Gaze. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2010,32(3):478-500.
- 3 Daunys G. Report on new approaches to eye tracking. Technical Report, Lithuania: Siauliai University, 2006. <http://www.cogain.org/results/reports/COGAIN-D5.2.pdf>
- 4 Li DH, Winfield D, Parkhurst DJ. Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. *Proc. of CVPR 2005*. USA: IEEE Computer Society San Diego, 2005:79-79.
- 5 Coyle D, Prasad G, McGinnity TM. GiFaster self-organizing fuzzy neural network training and a hyperparameter analysis for a brain-computer interface. *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2009, 39(6): 1458-1471.
- 6 Lv Z, Wu XP, Li M, et al. A novel eye movement detection algorithm for EOG driven human computer interface. *Pattern Recognition Letters*, 2010,31(9):1041-1047.
- 7 Moriyama T, Kanade T, Xiao J, Cohn JF. Meticulously Detailed Eye Region Model and Its Application to Analysis of Facial Images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2006,28(5):738-752.
- 8 吴敬华,周昌乐.平面旋转人脸检测与特征定位方法研究. *电子学报*,2007,35(9):1714-1718.
- 9 吴敬华.面向中医面诊诊断信息提取的若干关键技术研究. 厦门:厦门大学,2008.
- 10 Paul V, Michael J. Robust real-time object detection. *International Journal of Computer Vision*, 2004,57(2):137-154.