

基于人工神经网络的视线跟踪系统

李 兵,戴尔哈

(南京邮电大学 自动化学院,江苏 南京 210046)

摘要:利用一种新的健壮的低像素的虹膜中心定位算法和眼睛角点检测算法,在摄像头像素不高的情况下,实时地提取出了注视屏幕时的眼睛特征,即虹膜中心和眼睛角点。然后利用人工神经网络(ANN),将人眼注视屏幕不同点时的眼睛特征进行分类,确定出人眼特征和屏幕上点的映射关系。这种方法只需要一个普通的商业摄像头和一般的光照条件,并且允许小范围内的头部运动,很大程度上降低了系统的硬件成本,减少了使用者的限制条件,增强了系统的实用性。实验结果表明,在普通的实验室光照条件下,这种方法在视线跟踪中达到了良好的效果。

关键词:视线跟踪;人工神经网络;虹膜中心检测;眼睛角点检测

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2015)04-0098-04

doi:10.3969/j.issn.1673-629X.2015.04.023

Eye Gaze Tracking System Based on Artificial Neural Network

LI Bing,DAI Er-han

(College of Automation,Nanjing University of Posts and Telecommunications,Nanjing 210046,China)

Abstract:Use a new robust and low-resolution iris center localization algorithm and eye corners detection algorithm to extract the eye features of looking at the screen in real time under low-resolution camera conditions. Then classify the eye features of looking at different positions through the Artificial Neural Network (ANN) to determine the mapping relationship of eye features and the screen points. This way only needs a commercial camera under the general illumination conditions and allows a small range of head motion, further more, it has reduced the hardware cost of the system and the restrictions of users, enhancing the system's practicability. Experimental results show that in the ordinary laboratory illumination conditions, this method achieves a good accuracy in the gaze tracking.

Key words:gaze tracking;artificial neural network;iris center detection;eye corners detection

0 引言

人眼视线的跟踪在很多领域具有重要的作用,比如人机交互、广告分析、心理分析等。实现视线跟踪的方法有很多种,总体上可分为2D视线跟踪法和3D视线跟踪法^[1-3]。

2D视线跟踪法使用比较多的是瞳孔-角膜反射法(Cor-neal-Pupil Reflection)^[4-5],这种方法使用红外光源照射人眼,在角膜上形成亮斑,亮斑中心和瞳孔中心组成一个向量,这个向量和人眼注视点具有对应的关系,然后通过这个向量就可以求出人眼注视点的位置;3D视线跟踪法^[6-7]通过建立眼球模型,使用两个或两个以上的红外光源,在眼球上形成多个亮斑,然后根据这些亮斑通过一些复杂的公式计算出眼睛的注视点。这些方法都达到了一定的视线跟踪精度,但是

都不可避免地要使用红外光源或者多个摄像头,红外光源不仅增加了系统硬件成本,而且可能会对人体产生一定的伤害。因此文中采用了一种基于神经网络的2D视线跟踪方法,不需要使用红外光源,并且在单个摄像头像素不高的情况下取得了较好的精度。并且对系统做了下面几点改进:

(1)采用了一种根据梯度向量的方法来定位瞳孔中心,这种方法能在低像素的情况下准确地检测出瞳孔的中心,而且具有良好的健壮性,增强了系统的性能,降低了对硬件的要求。

(2)使用了一种人工神经网络的方法来映射人眼图像和注视点,并对人工神经网络的结构、特征向量等进行改进和优化,使系统不需要使用红外光源和多个摄像头,降低了系统的硬件要求,减少了硬件成本,提

收稿日期:2014-06-10

修回日期:2014-09-15

网络出版时间:2015-02-23

基金项目:国家自然科学基金资助项目(61203213)

作者简介:李 兵(1989-),男,硕士研究生,研究方向为图像处理、视线跟踪系统等;戴尔哈,硕士生导师,副教授,研究方向为测试计量技术及仪器、图像处理等。

网络出版地址:<http://www.cnki.net/kcms/detail/61.1450.TP.20150223.1239.032.html>

高了系统跟踪精度,增强了系统的实用性。

1 系统介绍

文中使用分级的方式来检测瞳孔中心和眼角点,即先从视频流检测出人脸,然后在人脸中检测人眼,再在人眼图像中定位瞳孔中心和眼睛角点。使用分级的方式不仅提高了眼睛参数检测的精度,还增强了系统在使用过程中的稳定性,从而提高了系统的实用性。系统的总体流程如图1所示。

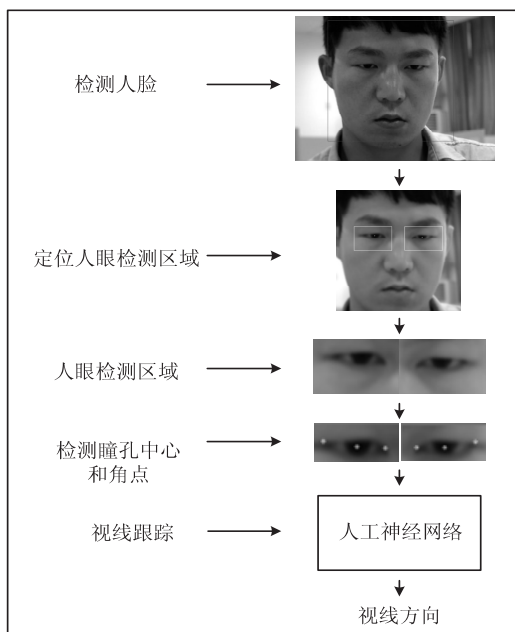


图1 系统的总体流程

具体的过程是:首先通过人脸检测算法从摄像头拍摄的视频帧中检测出人脸,然后在检测出的人脸当中大致地定位出人的眼睛区域,然后在这个人眼区域中准确地定位出人眼虹膜中心,然后在此基础上,以人眼虹膜中心点为中心,取像素大小为 128×64 的眼睛图像,在这个图像中进行眼睛的角点检测,提取出眼睛的各个特征参数。最后先将眼睛观看屏幕不同位置时的特征参数作为人工神经网络的训练、学习样本进行训练,训练完成之后,再将眼睛特征参数作为已经训练好的人工神经网络的输入特征向量,然后人工神经网络预测出人眼的视线方向。这里关键的步骤是人眼虹膜中心的定位、眼睛角点的检测和人工神经网络的特征向量的选取、人工神经网络结构、训练方法的设置,这些都直接影响整个系统的精度。

2 人脸检测

人脸检测在计算机视觉中具有很重要的作用。人脸检测的方法有很多种^[8],有基于相关匹配的方法、基于几何特征的方法、基于统计理论的方法等。在基于统计理论的人脸检测方法当中,基于 AdaBoost 学习算

法是使用最为广泛的一种人脸检测算法,这种人脸检测算法是 Viola 等于 2001 年提出的,因此本系统也采用这种方法检测人脸。使用这种算法检测人脸分为两步,第一步是基于某些特征训练分类器,第二步就是用训练好的分类器进行人脸检测。使用比较多的人脸检测分类器大多是基于 Haar-like 特征然后通过 AdaBoost 学习算法训练的,具体过程是:利用样本的 Haar-like 特征进行分类器训练,得到一个级联的层叠分类器。训练样本分为正例样本和反例样本,其中正例样本是指待检目标样本,反例样本指其他任意图片。分类器训练完成以后,就可以通过输入图像中的感兴趣区域进行检测。文中系统是基于 OpenCV 平台实现,OpenCV 带有训练好的 AdaBoost 人脸检测分类器数据,并保存在 XML 文件当中,在系统启动的时候加载这个文件生成分类器,然后输入待检测图像,分类器就能检测出人脸。

3 虹膜中心检测

虹膜中心的检测在计算机视觉当中具有非常重要的作用。在本系统中,虹膜中心定位的精度对后面视线跟踪的精度有很大影响。虹膜中心检测有多种方法,一般都使用阈值技术和椭圆拟合的方法,但这种方法对摄像头像素要求较高,稳定性不好,在像素较低的图像中对眼睛瞳孔中心的检测效果不理想。因此文中采用了另外一种在低像素情况下检测效果较好的方法,基于梯度向量的瞳孔中心检测方法^[9]。这种方法是通过求出眼球附近图像中各个方向的梯度向量来确定瞳孔中心,这些梯度向量会有交点,经过瞳孔中心点的梯度向量数目最多,也就是在瞳孔中心点上相交的次数最多,然后求出各个方向的梯度相交次数最多的点,这个点就是虹膜中心。具体的计算方法是:设 c 为一个可能的瞳孔中心点, x_i 是像素点位置, $i \in \{1, 2, \dots, N\}$, N 是图像的像素大小, g_i 是在 x_i 这点的梯度矢量, d_i 是从 c 到 x_i 之间的所有的位移矢量(方向相同大小不同),那么矢量 d_i 和梯度矢量 g_i 有相同的方向(图2左边的 d_i 和 g_i 方向不一致,因为 c 不是瞳孔中心),因此可以通过求出各个可能的瞳孔中心点 c 的所有位移向量 d_i 和对应的梯度向量 g_i 的点积之和来确定瞳孔中心点(两个相同方向的矢量点积最大,点积和最大的点就是瞳孔中心点),数学公式如下:

$$c^* = \max_c \left\{ \frac{1}{N} \sum_{i=1}^N (d_i^T g_i)^2 \right\} \quad (1)$$

$$d_i = \frac{x_i \cdot c}{\|x_i \cdot c\|_2}, \forall i: \|g_i\|_2 = 1 \quad (2)$$

其中,为了提高在光照条件变化情况下的健壮性,需要将位移向量 d_i 和梯度向量 g_i 都单位化。

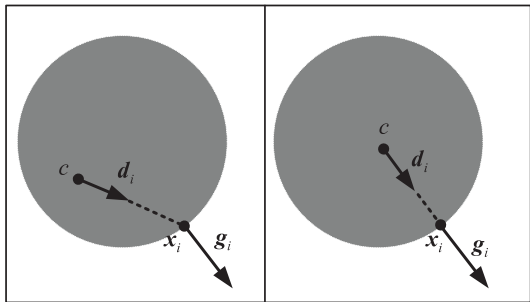


图 2 瞳孔检测模型(灰色区域表示眼睛虹膜区域)

4 眼睛角点检测

在本系统中,角点检测的精度很大程度上决定了视线估计的精度。图像处理中角点简单地理解就是图像在水平和垂直方向上变化都比较大的点。角点检测算法总的来说可以分为三类:基于灰度图像的角点检测、基于二值图像的角点检测、基于轮廓曲线的角点检测。基于灰度图像的角点检测又分为基于梯度、基于模板和基于模板梯度结合的方法。使用比较多的基于模板的角点检测算法有 Kitchen-Rosenfeld 角点检测算法、Harris 角点检测算法、KLT 角点检测算法及 SUSAN 角点检测算法。这里采用经典的 Harris 角点检测方法^[10-11]。为了准确地检测出眼睛的角点,在进行角点检测之前需要对检测出的人脸图像进行高斯滤波,去除各种噪声的影响,再根据检测出的虹膜中心,获取以这个点为中心、大小为 128×64 的眼睛图像,再将眼睛图像左右各 $1/3$ 宽的部分作为眼睛两个角点的检测的感兴趣区域,去除眼睛虹膜部分等对眼睛角点检测的影响,最后在这个区域内进行 Harris 角点检测。

5 视线跟踪

视线跟踪的方法有很多种,主要分为 2D 跟踪法和 3D 跟踪法两大类。文中使用 2D 视线跟踪方法中基于神经网络的方法,即使用图像处理中的人工神经网络来映射人眼特征参数和人眼注视点。采用这种方法需要将注视点人眼特征参数作为人工神经网络的训练样本,并且根据样本设定神经网络的输出,根据设定的样本输出对这些样本进行学习训练,再将实时获取的人眼特征作为已训练好的神经网络的输入,然后神经网络的预测输出就是人眼特征所对应的注视点。这里合适的人眼特征参数的选取、神经网络参数的设置和训练方法等很大程度上决定了视线估计的精度。

5.1 神经网络输入特征向量的选取

特征向量的选取是整个系统的关键。人工神经网络的特征向量需要具有一定的差异性,否则神经网络的预测输出将不准确。根据 Baluja 和 Pomerleau 的方法^[12],先将图像归一化到 32×16 大小,然后将整个 32

$\times 16$ 大小眼睛图像的所有像素作为一个 512 大小的特征向量,经实验证明这种方法效果不理想。这种方法将整个眼睛图像像素作为神经网络的特征向量,特征向量的差异性不明显,且特征向量维数较大,增加了神经网络的训练和预测时间。因此在本系统中将两个眼睛图像的瞳孔中心和眼睛角点的位置信息^[13]作为神经网络特征向量,这种特征向量具有较好的差异性,适合作为神经网络的输入,并且大大减小了特征向量的维数,提高了系统的速度。这里,眼睛瞳孔中心和眼睛角点的位置信息组成了一个大小为 16 的特征向量:其中的 8 个值由瞳孔中心和眼睛角点的距离和夹角组成,剩下的 8 个值由两个眼睛的 4 个角点的坐标位置组成。图 3 中比较直观地描述了这些向量。

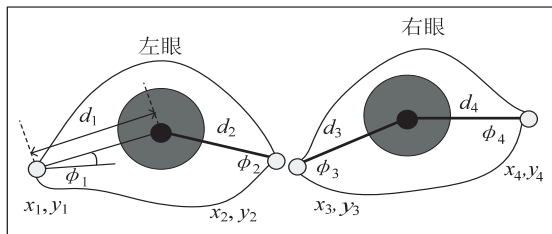


图 3 眼睛的几何特征

5.2 神经网络的结构

人工神经网络的种类有多种,文中采用 BP 神经网络^[14]。BP 神经网络是一种按误差逆传播算法训练的多层前馈网络,是目前应用最广泛的神经网络模型之一,BP 网络能学习和存贮大量的输入-输出模式映射关系,而无需事前揭示描述这种映射关系的数学方程。它的学习规则是使用最速下降法,通过反向传播来不断调整网络的权值和阈值,使网络的误差平方和最小。本系统中需要创建两个单独的神经网络 bp_x 和 bp_y ,分别用于 x 方向和 y 方向的坐标计算,训练方法选用反向传播法。 bp_x 采用包含一个隐藏层的三层结构,输入层包含 16 个节点单元,隐藏层 8 个节点,输出层 5 个节点; bp_y 输入层和隐藏层的结构与 bp_x 相同,输出层包含 3 个节点。两个网络的详细结构如图 4 所示。

5.3 神经网络的训练

文中将电脑屏幕分成 5×3 大小的 15 个区域,为了方便获取样本数据训练神经网络,这里采用的方法是:当眼睛注视屏幕上某个区域时,把这时的眼睛特征的数据保存下来,每个区域获取 10 个样本数据,这样便得到 15×10 组样本数据,最后将这些眼睛特征参数数据写入到本地文件当中,然后在程序启动时读取存入在本地文件当中的数据并进行训练,训练完成后将训练好的神经网络数据保存在文件中,以便下次使用时直接加载已经训练好的神经网络的参数数据而不是每次都需要获取样本进行训练。

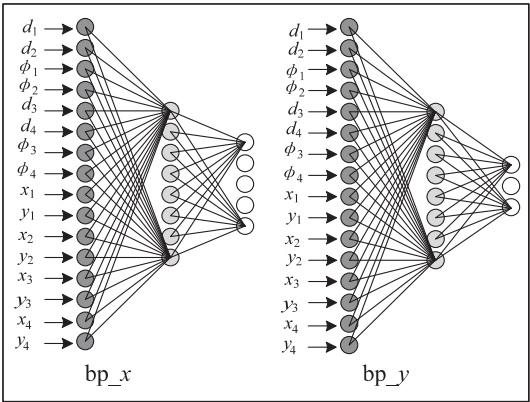


图4 bp_x、bp_y 结构图

6 实验结果

实验在一般光照条件下的教研室中进行,200万像素的工业摄像头,人脸距离电脑屏幕50 cm左右,电脑的屏幕大小是15.6英寸,长约34.5 cm,宽约19.5 cm,像素大小是1366×768,在头部保持不变或者轻微变化情况下进行10次实验。实验结果见表1~表3。

表1 x方向的识别率

x方向	平均识别率/%	x方向	平均识别率/%
1	91	4	91.3
2	82.3	5	97
3	94.8	平均值	91.3

表2 y方向的识别率

y方向	平均识别率/%	y方向	平均识别率/%
1	98	-	-
2	87.3	-	-
3	93.8	平均值	93

表3 15个位置的识别率

位置	平均识别率/%	位置	平均识别率/%
1	92.5	9	90
2	86	10	95
3	97.3	11	61.3
4	97.3	12	67.3
5	99.8	13	84.5
6	87.3	14	93.5
7	90	15	99.5
8	83.5	平均值	88.3

实验结果证明,本系统对视线的跟踪效果良好,达到了不错的跟踪精度,其中x方向的平均识别率达到了91.3%,y方向的平均识别率达到了93%,15个位置的平均识别率达到了88.3%,除了个别位置的识别率不高外,一些位置的识别率更是达到了99%。经计算实验的全局误差:x方向是3.9°,y方向是3.5°。

7 结束语

文中采用一种新的基于神经网络的视线跟踪方法,并加以优化了的准确的和健壮的眼睛特征检测算法,使头部在比较自然的情况下获得较好的跟踪精度,并且只需要一个普通的摄像头,在很大程度上降低了系统的硬件要求和成本,减少了对使用者的限制,增强了系统的实用性。未来的工作主要是针对大的范围内的头部自由移动的视线跟踪方法的研究,以实现真正的高精度,非侵入的,高实用性的,能应用到多个领域的视线跟踪系统。

参考文献:

[1] 张 昀,赵荣椿,赵歆波,等. 视线跟踪技术的2D和3D方法综述[J]. 信号处理,2007,23(4A):218-224.

[2] 赵歆波,邹晓春,张 昀. 视线跟踪关键技术原理及其实现[J]. 测控技术,2007,26(S):327-329.

[3] 迟健男,张 闯,胡 涛,等. 视线追踪系统眼部特征检测及视线方向计算方法研究[J]. 控制与决策,2009,24(9):1345-1350.

[4] 张 闯,迟健男,张朝晖,等. 一种新的基于瞳孔-角膜反射技术的视线追踪方法[J]. 计算机学报,2010,33(7):1272-1285.

[5] Ebisawa Y, Fukumoto K. Head-free, remote eye-gaze detection system based on pupil-corneal reflection method with easy calibration using two stereo-calibrated video cameras[J]. IEEE Transactions on Biomedical Engineering, 2013, 60(10):2952-2960.

[6] 刘瑞安,靳世久,宋 维,等. 单摄像机视线跟踪[J]. 计算机应用,2006,26(9):2101-2104.

[7] 赵新灿,卢朝阳. 头部自由运动的视线跟踪方法[J]. 南京航空航天大学学报,2010,42(4):435-439.

[8] 刘 敏,李 晖. 基于AdaBoost分类器和特征脸的人脸检测与识别[J]. 漳州师范学院学报:自然科学版,2010,23(2):43-47.

[9] Timm F, Barth E. Accurate eye centre localisation by means of gradients[C]//Proceedings of the Sixth international conference on computer vision theory and applications. Vilamoura, Algarve, Portugal:[s. n.],2012.

[10] 张 静. Opencv2 计算机视觉编程手册[M]. 北京:科学出版社,2013.

[11] 高 飞,张宪民. 视线跟踪系统中的眼角点精确定位方法[J]. 计算机工程,2007,33(18):199-201.

[12] 冯成志,沈模卫. 视线跟踪技术及其在人机交互中的应用[J]. 浙江大学学报:理学,2002,29(2):225-232.

[13] Sesma L, Villanueva A, Cabeza R. Evaluation of pupil center-eye corner vector for gaze estimation using a web cam[C]//Proceedings of the 2012 symposium on eye tracking research & applications. [s. l.]:[s. n.],2012:217-220.

[14] 侯 勇. 基于最速下降算法BP神经网络及应用探讨[J]. 现代商贸工业,2011,23(8):245-245.

基于人工神经网络的视线跟踪系统

作者：[李兵](#)，[戴尔哈](#)，[LI Bing](#)，[DAI Er-han](#)
作者单位：[南京邮电大学 自动化学院, 江苏 南京, 210046](#)
刊名：[计算机技术与发展](#)[ISTIC](#)
英文刊名：[Computer Technology and Development](#)
年，卷(期)：2015(4)

引用本文格式：[李兵](#). [戴尔哈](#). [LI Bing](#). [DAI Er-han](#) [基于人工神经网络的视线跟踪系统](#)[期刊论文]-[计算机技术与发](#)
[展](#) 2015(4)