

基于计算机视觉的驾驶员低头行为检测

杨晓峰¹ 邓红霞² 李海芳²

(山西建筑职业技术学院 榆次 030600)¹ (太原理工大学 榆次 030600)²

摘要 在驾驶过程中使用手机会引起驾驶员的注意力分散,为了对这种行为进行监督和提醒(在公共交通中检测更有意义),提出了一种基于脸部特征提取的驾驶员低头行为的检测方法。该方法使用主动形状模型(Active Shape Model,ASM)算法得到脸部特征点,在此基础上通过脸部特征点的位置信息计算出头部姿势描述信息,最后通过SVM将上述信息分类进而得出头部姿势,其可用于判断驾驶员是否在驾驶过程中低头看手机行为,该方法能够有效检测出驾驶员在驾驶过程中低头使用手机的行为。实验结果表明,该方法的平均检出率在94%以上。

关键词 ASM,脸部特征,驾驶,手机,注意力分散

中图分类号 TP391.9 **文献标识码** A

Detection of Driver's Head-dipping Based on Computer Vision

YANG Xiao-feng¹ DENG Hong-xia² LI Hai-fang²

(Shanxi Architectural College, Yuci 030600, China)¹ (Taiyuan University of Technology, Yuci 030600, China)²

Abstract In order to monitor and alert the distraction by using a mobile phone during driving, the method based on the facial feature extraction to detect the driver's head behavior was proposed. This method uses ASM (Active Shape Model) to obtain the facial feature points, calculates the head posture description on the position information of the face feature points, and draws the head posture classified by SVM from the above information finally. Experimental results show that the method can effectively detect the driver's head-dipping during driving, and the average detection rate is above 94%.

Keywords ASM, Facial feature, Driving, Phone, Distraction

随着智能手机的普及,微博、微信等手机应用越来越丰富,不少司机开车时养成了看手机的习惯。驾驶员在行车时使用手机引起驾驶绩效下降是导致交通事故的重要原因^[1],尤其在公共交通运输中危害最大。针对这种情况可以通过加装固定检测设备,提醒驾驶员减少低头看手机的行为。

目前通过视频检测驾驶行为的方法主要是针对驾驶员头部或者面部的特征进行检测。判别疲劳驾驶的方案较多,而对驾驶员低头看手机的行为检测较少。采用ASM^[2]算法的驾驶状态检测方法有:基于ASM算法眼睛和嘴巴的特征提取,结合疲劳模糊推理判断疲劳驾驶^[3];基于ASM跟踪人眼检测疲劳^[4];基于ASM算法的肤色区跟踪定位耳部和手部进行开车打电话行为的判别^[5]。

与本文相关的还有对于头部姿势的检测方法,此类检测方法都需要确定头部姿势的3个偏移角度(3个维度),计算量大而且实时性不够强,如王侃等使用基于AAM算法的双摄像机方案^[6];张博文等提出基于深度图结合ICP的头部姿态分析^[7];施华等提出头部姿势估计的方法中,没有使用迭代法得出头部姿势,但在仰视角度大于30°时检出率最高只有43.33%^[8],并不适合对驾驶状态下的头部行为的实时判断。

国外相关的方法如下,Breitenstein提出了基于红外摄像头和双目摄像头的算法^[12],基本实现实时头部姿态估计,但是该方案需要较高的硬件条件。KoiChi和Yasue提出了使用简化AAM算法结合ZNCC算法的头部跟踪算法^[11]。该算法中使用相邻两帧进行头部姿态计算,对静态图像无效,而且该算法使用局部方形纹理匹配,计算量大于ASM算法使用的法线纹理匹配。Yan Wang等^[13]提出了通过手机APP应用程序的方法检测驾驶的安全性,Yantao Li等^[14]也提出类似的方法,这些方法基于手机中的多种传感器与本思路不同。

本文提出了基于ASM算法的驾驶员行为检测方法,利用面部特征点的相对位置与头部行为的相关性进行头部姿势的判定,对静态图像和动态图像均有效,并采用头部姿势判定参数冗余设计克服了ASM算法的精确性欠佳的缺点。本文所涉及到头像图片均取自MUCT。该方法对驾驶员低头行为检出率高,实验表明平均检出率在94%以上,最高检出率在97%以上,而且不会对驾驶员正常驾驶行为造成任何干扰。

1 检测方案

在驾驶员低头使用手机行为的检测方案中,摄像头的安

本文受山西省自然科学(青年科技研究)基金(2014021022-5),国家电网公司科技项目(52053015000W)资助。

杨晓峰(1979—),男,硕士,讲师,主要研究方向为智能信息处理、图像识别,E-mail:18817710@qq.com;邓红霞(1976—),女,博士,副教授,CCF会员,主要研究方向为智能信息处理、图像识别;李海芳(1963—),女,教授,博士生导师,CCF会员,主要研究方向为智能信息处理、脑认知模型与脑机接口技术、视听觉信息计算和情感计算,E-mail:lihaifang@tyut.edu.cn(通信作者)。

放位置如图 1 所示。检测过程中,通过 Adaboost 算法获取图像中的面部位置,再由 ASM 算法分析面部图像获得面部特征点,对面部特征点坐标归一化处理后计算头部姿势判定参数,最后进行 SVM 分类判断驾驶员是否有低头姿势,如果低头时间或频率大于一定阈值时,则发出报警提示。检测流程如图 2 所示。

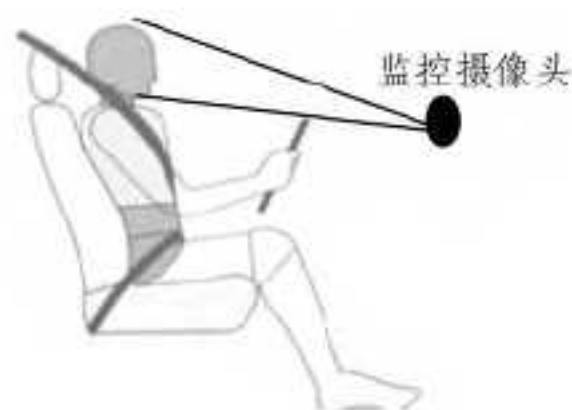


图 1 摄像头的安放位置

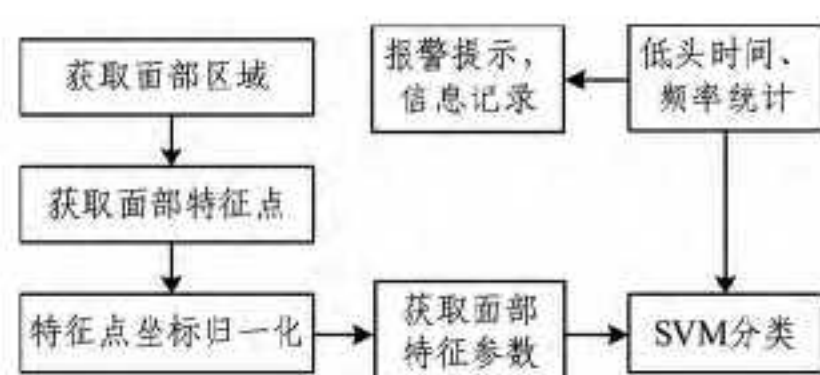


图 2 驾驶行为检测流程

由于摄像头在头部水平线以下摆放,因此正常开车时,脸部相对于摄像头是仰视状态。而摄像头中显示平视和俯视都属于非正常开车状态,在本文中均认为是低头玩手机。正常开车面部图像如图 3 所示,非正常开车面部图像如图 4、图 5 所示。通过本文的方法能有效检测出图 3 为正常驾车行为,图 4、图 5 为非正常驾车行为。



图 3 正常开车面部图像



图 4 非正常开车面部图像



图 5 非正常开车面部图像

检测过程中,出于实时性的考虑,采用以下方法:1)在进行 ASM 算法提取特征点之前,先缩小 ASM 搜索范围,减少 ASM 搜索时间;2)没有采用精度更高的 AAM 算法,因为 AAM 算法所花费的时间和空间要大于 ASM 算法。

2 获取面部特征点信息

2.1 获取面部区域

ASM 算法性能对初始搜索位置非常敏感,适合的搜索起

始位置可以大大减少算法迭代次数。为了提高 ASM 算法的收敛速度,避免得到局部最优解而导致的搜索失败,实验中采用了 Viola 等人提出的一种基于简单类 Haar 特征的级联 Adaboost 的快速人脸检测算法[6]。人脸搜索示意图如图 6 所示。至于文献[5]中提出“驾驶人头部产生偏转的情况下 Adaboost 算法无法有效检测出侧面人脸”,在本文涉及的环境下,可以简单认为是非正常驾驶行为,不影响本文检测方案的准确度。



图 6 人脸搜索示意图

2.2 获取面部特征点

传统的 ASM 算法和改进的 ASM 算法[10]都是基于点分布模型(Point Distribution Model,PDM)的算法。在 PDM 算法中,人脸、人手、心脏等的同类外形相似的物体可以通过若干关键特征点的坐标信息依次串联形成一个形状向量。ASM 利用形状向量进行图像识别,包括本文中使用的脸部特征的识别。

训练时,随机选择 N 幅人脸图像作为训练数据,每一幅图像都需要手工标定特征点。要求特征点在每幅图像中要基本处于相同位置,分布均匀,一般标定在能够代表目标轮廓的地方,如下巴、额头、眼部、嘴部等。把标定点记为: $S_i = (x_1^i, y_1^i, x_2^i, y_2^i, \dots, x_n^i, y_n^i)$, $i=1, 2, \dots, n$, 其中 (x_j^i, y_j^i) 表示第 i 个训练样本中第 j 个特征点的坐标。

在对样本图像的形状向量平移、旋转、缩放对齐之后,就可以对形状向量进行 PCA 处理,得到形状模型 $S \approx S_0 + P_s b$, S 表示目标形状, S_0 表示由样本产生的平均脸, b 表示形状参数, P_s 表示存放特征向量组成的矩阵。

ASM 局部灰度模型的建立。局部特征一般用灰度差分作为特征,以防光照变化。在第 i 幅 $(i=1, \dots, N)$ 图像上,对于每个特征点 $j(j=1, \dots, n)$,通过该特征点做一条直线,并与邻近的两点连线垂直,分别在该垂线上特征点的两边取 k 个像素的灰度值(包括特征点的灰度,一共 $2k+1$ 个灰度值),于是得到一个 $2k$ 个灰度值的差分向量 $g_{ij} = (g_{ij(-k+1)}, g_{ij(-k+2)}, \dots, g_{ijk})^T$ 。然后对 g_{ij} 进行标准化, $g_{ij}' = \frac{g_{ij,k}}{\sum g_{ij,k}}$ 。

于是得到对应特征点的局部灰度值模板,包括平均值 $\bar{g}$ 和协方差 S^{-1} 。在形状匹配过程中,使用马氏距离为代价函数,

$$f(g^*) = \sqrt{(g^* - \bar{g})^T S^{-1} (g^* - \bar{g})}.$$

3 特征点的选取与处理

利用 ASM 算法拟合任意一个图像可以得到点集为 $S = ((x_1, y_1)^T, (x_2, y_2)^T, \dots, (x_n, y_n)^T)^T$ 。

首先对点集进行预处理得到点集 S' :

$$\bar{S} = (\bar{x}, \bar{y}) \quad (1)$$

$$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} \quad (2)$$

$$\sigma_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n}} \quad (3)$$

$$\text{设 } \sigma^{-1} = \left(\frac{1}{\sigma_x}, \frac{1}{\sigma_y} \right)^T$$

$$S' = (S - \bar{S}) \sigma^{-1} = ((x_1', y_1')^T, (x_2', y_2')^T, \dots, (x_n', y_n')^T)^T \quad (4)$$

然后从点集 S' 提取特征量, 本实验使用 14 个特征点坐标, 生成 7 个点距 ($L1, L2, L3, L4, L5, L6, L7$), 最后得到 4 个头部姿势判定参数 (a, b, c, d)。这 14 个点分别为两鬓两点、两眼内角点、两眼外角点、鼻子中线三点、下巴两点、脸颊下部两点、位置如图 7、图 8 所示。4 个参数分别设置为: $a = \frac{L1^2}{L2^2}$, $b = \frac{L3^2}{L4^2}$, $c = \frac{L5^2}{L6^2}$, $d = \frac{L7^2}{L2^2}$ 。选取特征点时考虑三方面的因素: 1) 仰视、平视和俯视时, 由于透视原理点距会发生变化, 变化显著的点对应该被优先选取; 2) 受到表情影响小的点距应该优先考虑; 3) 冗余性。

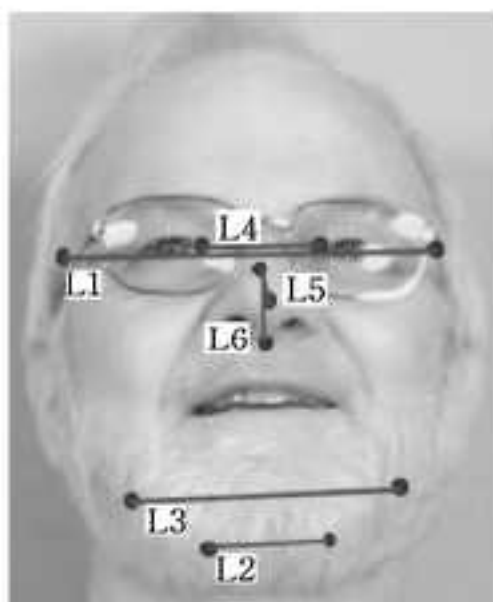


图 7 脸部特征点位置

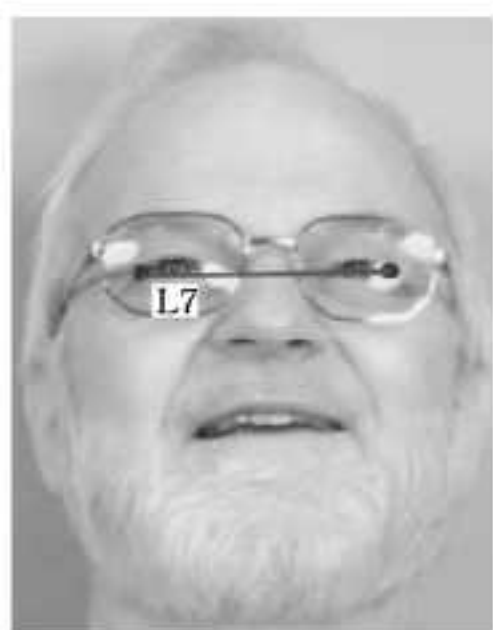


图 8 脸部特征点位置

预处理后, 点集中任意两点 (x_i', y_i') , (x_j', y_j') 的距离的平方为:

$$\begin{aligned} L^2 &= (x_i' - x_j')^2 + (y_i' - y_j')^2 \\ &= \left(\frac{x_i - \bar{x}}{\sigma_x} - \frac{x_j - \bar{x}}{\sigma_x} \right)^2 + \left(\frac{y_i - \bar{y}}{\sigma_y} - \frac{y_j - \bar{y}}{\sigma_y} \right)^2 \\ &= \frac{(x_i - x_j)^2}{\sigma_x^2} + \frac{(y_i - y_j)^2}{\sigma_y^2} \end{aligned}$$

任意两点的距离的平方比值如下, 其中 L_{ij}^2 表示点 i 到点 j 的距离平方, L_{nm}^2 表示点 n 到点 m 的距离平方。

$$\begin{aligned} \frac{L_{ij}^2}{L_{nm}^2} &= \frac{\frac{(x_i - x_j)^2}{\sigma_x^2} + \frac{(y_i - y_j)^2}{\sigma_y^2}}{\frac{(x_n - x_m)^2}{\sigma_x^2} + \frac{(y_n - y_m)^2}{\sigma_y^2}} \\ &= \frac{\sigma_y^2 (x_i - x_j)^2 + \sigma_x^2 (y_i - y_j)^2}{\sigma_y^2 (x_n - x_m)^2 + \sigma_x^2 (y_n - y_m)^2} \end{aligned}$$

从上式可以看出头部姿势判定参数 (a, b, c, d) 只与点的坐标和方差有关, 与脸部在图像中的绝对位置无关, 与脸部与图像的比例无关。

4 行为识别及预警

行为识别本次实验选用 SVM 支持向量机, 通过 SVM 判定当前的状态。由于 SVM 二分类的分类器, 因此首先选取正性(正常驾驶)、负性(非正常驾驶)两组图像并且进行标注, 如下所示。

正性图片:



负性图片:



通过上述过程得到图像的头部姿势判定参数以及正负性标注, 即可作为 SVM 支持向量机的训练和测试数据。通过测试结果的比对, 可以计算本方法的准确性。

在“检测方案”部分提到的时间频率统计和报警提示过程可以采用如下方案:

. 统计 t (判定为非正常驾驶姿势的总时间长度), n (判定为非正常驾驶姿势的次数), T (行车总时间, 停车时停止统计);

. 如果单位时间内的非正常行为时间过长 ($t/T > \phi$), 或者次数过多 ($n/T > \psi$) 都会引发报警提示, 其中阈值 ϕ, ψ 可根据实际情况设定。

5 实验结果及分析

由于 MUCT 是从不同角度拍摄的人脸库, 非常适合给本次实验提供丰富的测试用例。但是由于本次实验是固定仰视拍摄, 因此需要人工地从中选取测试图片。本次实验采用了 MUCT 人脸库中抽取的 1037 张照片, 其中正常驾驶(正性)图片 161 张、非正常驾驶(负性)图片 876 张。

训练数据选取其中正性图片 50 张, 负性图片 100 张。测试数据选取其余正性图片 111 张, 负性图片 776 张。测试结果如表 1 所列, 平均预测正确率为 94.29%, 证明本文的检测方法是有效的。实际应用中可以选取一组预测结果最优的训练集进行训练, 从而达到最高的预测准确率。

表 1 分组测试结果

测试组	训练集(图片编号)		测试集(图片编号)		预测正确率 (%)
	正性	负性	正性	负性	
1	1-50	1-100	51-161	101-876	97.07
2	51-100	101-200	1-50	1-100	94.59
3	101-151	201-300	1-100	1-200	91.21
			201-161	301-876	

与其他相似的研究结果^[13,14]对比,可知本文提出方法已经达到较好的水平,比较结果如表2所列。

表2 实验结果对比

	参考文献 ^[17]	参考文献 ^[16]	本文
准确率(%)	87	90(最高97)	94.29

基于第1组测试,本实验还做了头部姿势判定参数分类性能测试。测试方案,4个参数两两组合,三三组合,然后通过SVM训练测试结果如下。图9为ab组合预测结果,图10为ac组合预测结果,图11为ad组合预测结果,图12为bc组合预测结果,图13为bd组合预测结果,图14为cd组合预测结果。预测正确率如表3所列。

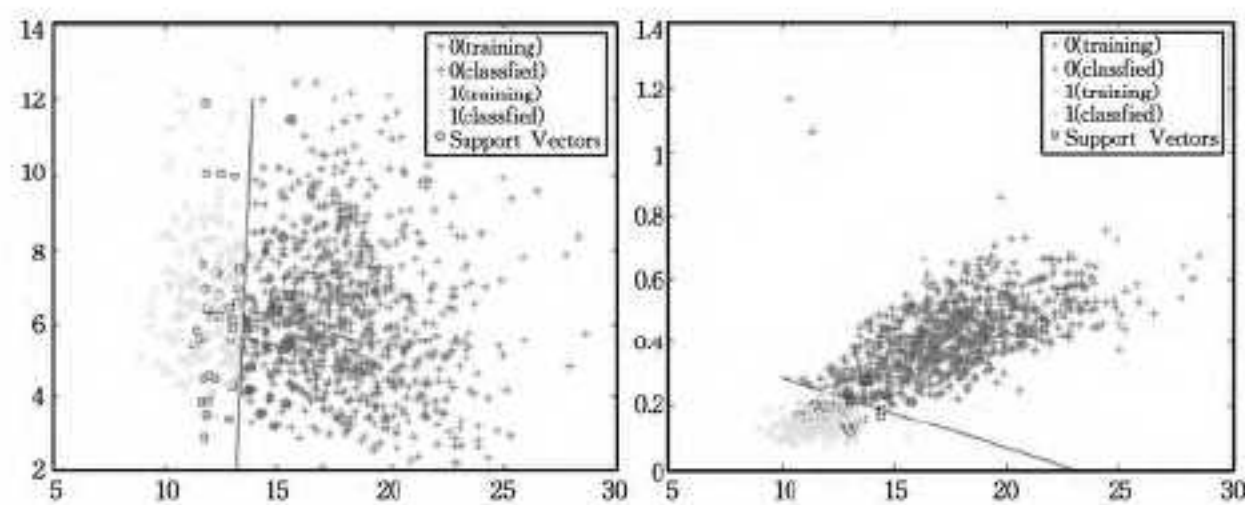


图9 ab组合预测结果

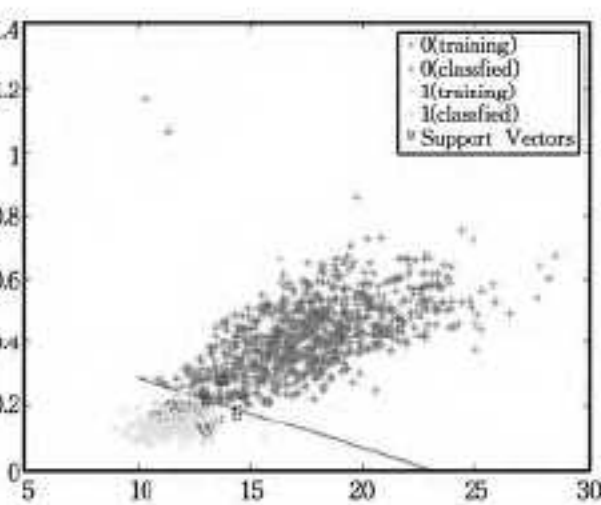


图10 ac组合预测结果

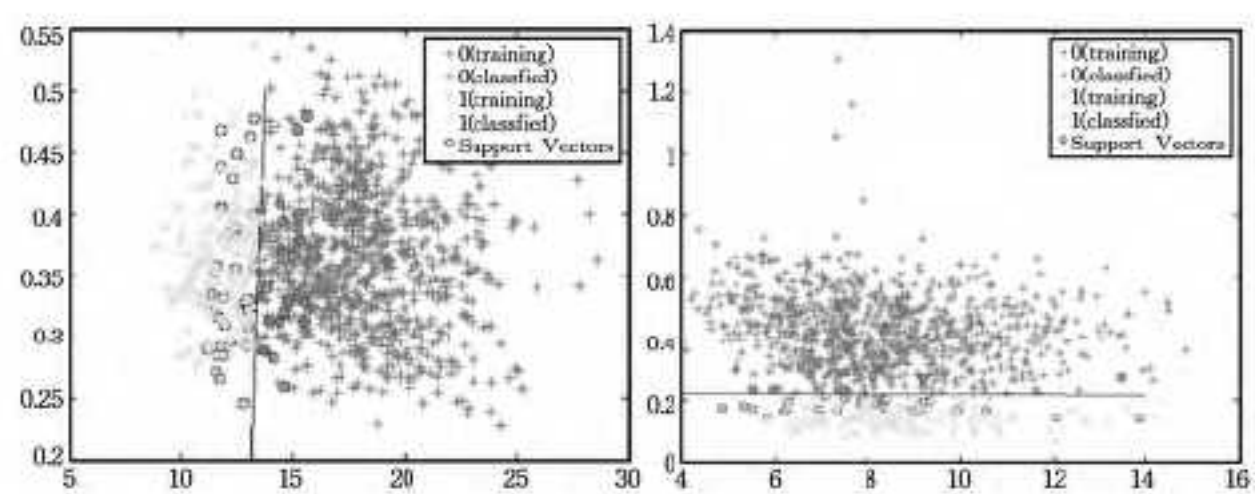


图11 ad组合预测结果

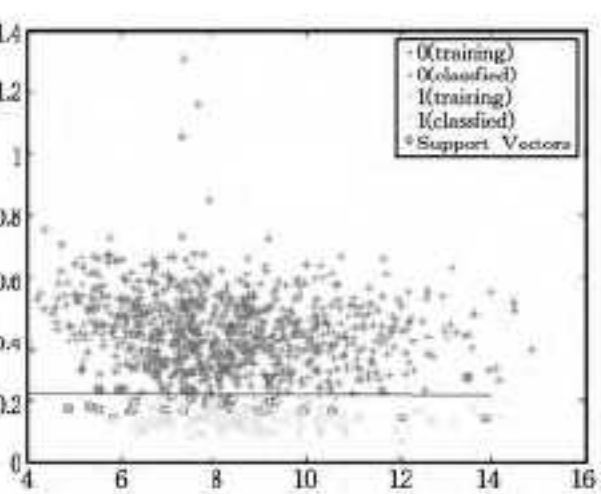


图12 bc组合预测结果

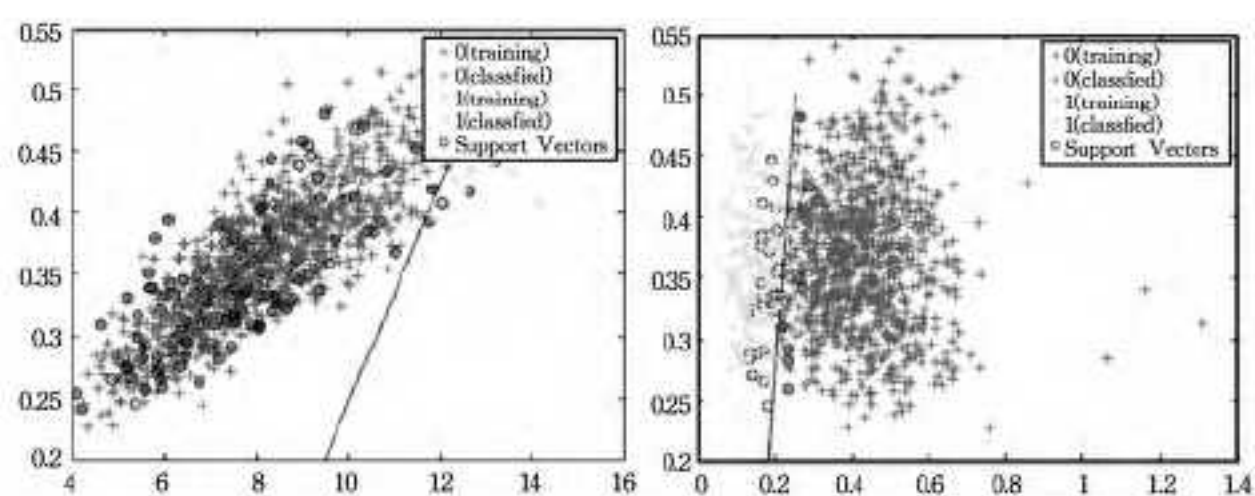


图13 bd组合预测结果

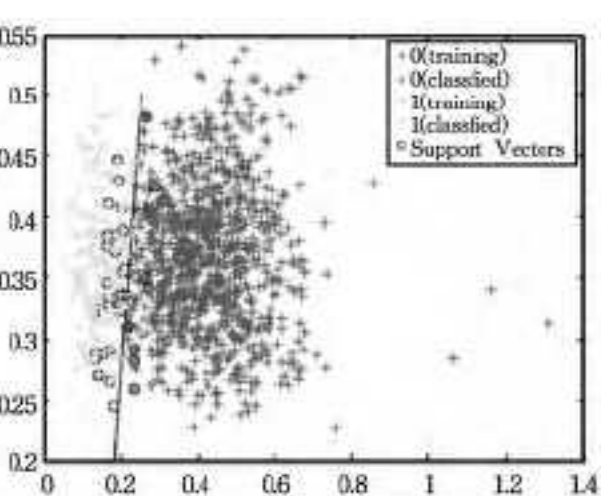


图14 cd组合预测结果

表3 头部姿势判定参数分析(%)

ab组合	ac组合	ad组合	bc组合	bd组合	cd组合
92.11	96.39	91.88	96.05	84.89	95.83
abc组合	abd组合	acd组合	bcd组合	综合(abcd)正确率	
96.39	92.00	96.39	96.51	97.07	

从表2可以看出多特征对于增加SVM的预测准确率有一定的作用,但是只有有效的特征量选取才会达到上述效果。此外,虽然多特征对于提高正确率有限,但是可以增加算法的

鲁棒性,尤其是脸部有遮挡的情况。本次实验的主要工作量在于选取有效的特征量得出测试结果,并对结果分析综合得出最优解,本实验最终选取的4参数方案,其他应用情况随后深入讨论。表2的结果只是针对第1组实验的分析,其他组的参数分析从略。

结束语 基于ASM算法得出人脸局部特征,利用透视原理通过脸部特征点可以有效判定头部姿势。本次实验中选取14个面部特征点,建立4个头部姿势判定参数,通过SVM进行分类可以有效地判定是否抬头。本次实验是一个成功的尝试,在后续的实验中可以基于ASM算法对其他行为进行检测。

参考文献

- [1] Tornros J, Bolling A. Mobile Phone Use-Effects of Conversation on Mental Workload and Driving Speed in Rural and Urban Environments[J]. Transportation Research Part F: Traffic Psychology and Behaviour, 2006, 9(4): 298-306
- [2] Cootes T, Taylor C, Cooper D, et al. Active shape models-their training and application[J]. Computer Vision and Image Understanding, 1995, 61(1): 38-59
- [3] 白中浩, 焦英豪, 白芳华. 基于主动形状模型及模糊推理的驾驶疲劳检测[J]. 仪器仪表学报, 2015, 36(4): 768-775
- [4] 杨智勇, 蒲亦非, 巩晓倩, 等. 疲劳驾驶检测中人眼实时定位与跟踪[J]. 微电子学与计算机, 2014, 31(8): 115-122
- [5] 张波, 王文军, 魏民国, 等. 基于机器视觉的驾驶人使用手持电话行为检测[J]. 吉林大学学报(工学版), 2015, 45(5): 1688-1695
- [6] 王佩. 人体头部姿态参数测量[D]. 西安: 西安工业大学, 2015
- [7] 张博文. 基于深度图的驾驶员头部姿态分析[D]. 大连: 大连海事大学, 2015
- [8] 施华. 头部姿态估计与跟踪系统研究与实现[D]. 上海: 华东师范大学, 2015
- [9] Viola P, Jones M J. Robust real-time face detection[J]. International Journal of Computer Vision, 2004, 57(2): 137-154
- [10] 田华, 蒲天银. 一种改进的ASM人脸特征点定位方法[J]. 重庆邮电大学学报(自然科学), 2014, 26(1): 124-130
- [11] Takahashi K, Mitsukura Y. Head Pose Tracking Based on Optimizing Normalized Cross-Correlation[C] // Nagoya, Japan: SICE, 2013: 2513-2517
- [12] Breitenstein M D, Daniel, et al. Real-time face pose estimation from single range images[C] // Anchorage, AK: CVPR 2008, 2008: 1-2
- [13] Wang Yan, Chen Ying-ying, Yang Jie, et al. Determining Driver Phone Use by Exploiting Smartphone Integrated Sensors[J]. IEEE Transactions on Mobile Computing (TMC), 2015, (9): 1
- [14] Li Yan-tao, Zhou Gang, Li Yue, et al. Determining driver phone use leveraging smartphone[J/OL]. [2015-08-09]. <http://link.springer.com/content/pdf/10.1007%2Fs11042-015-2969-7.pdf>