

基于人脸特征点的疲劳检测方法研究*

朱名流, 李顶根†

(华中科技大学 能源与动力工程学院, 武汉 430074)

摘要: 疲劳驾驶是造成交通事故的主要原因之一, 由于通过传感器等穿戴设备所需要的传感器太多或设备结构太复杂, 且深度学习模型对设备配置要求较高, 设计了一种基于普通摄像头视频进行疲劳检测的方法。该方法通过提取视频图像中的 HOG 特征进行人脸识别, 使用人脸特征点模型进行人眼和嘴部的定位, 通过计算人眼纵横比来判断睁眼和闭眼状态, 从而获得 PERCLOS 值。通过计算嘴部张度来判断是否处于哈欠状态, 并通过计算精确率 (precision)、召回率 (recall) 和 F_1 -measure 指标来评估哈欠检测的效果, 最后进行多特征融合来对驾驶员的驾驶状态进行综合判定。不依赖深层网络, 可以很好地移植到嵌入式设备。实验结果表明, 提出方法在准确率和实用性方面具有很好的效果, 对光照和佩戴物也具有较好的鲁棒性。

关键词: 人脸识别; 人脸特征点; PERCLOS; 疲劳检测

0 引言

驾驶员疲劳是造成机动车交通事故的主要原因之一。国家公路交通安全管理局的报告中记录了 846 起与疲劳驾驶司机有关的死亡事故^[1]。这些死亡人数在过去十年中基本上保持一致。据统计, 2005—2009 年, 平均每年有 83 000 起与疲劳驾驶有关的车祸。基于这些原因, 强烈建议为驾驶员提供风险警报系统, 该系统使用可确定睡意的探测器。报警系统可以唤醒处于疲劳或者有睡意的驾驶员, 也可以将控制权交给自动驾驶车辆。

目前已经有多种技术用于测量驾驶员的睡意。这些技术大致可分为车辆行驶模式、驾驶员的心理生理特征、驾驶员监控的计算机视觉技术三类。在第一类技术中, 最新技术是基于监视方向盘运动来实现对司机驾驶行为的检测^[2,3]。这一类的技术主要通过计算在加速或刹车时间、车道偏离以确定^[4,5]司机的困倦程度。第二类技术通过生物电信号, 如 EEG (脑电图)、ECG (心电图) 和 EOG (眼电图)^[6]。但是, 前面提到的两个类中的技术有严重的限制。前者只能在特定的驾驶条件下使用, 本质上鲁棒性不强, 而后者由于驾驶员在车身上佩戴各种信号测量工具而难以达到实际目的。所以, 最后一类基于计算机视觉的驾驶员监控技术正变得越来越流行^[7,8]。计算机视觉技术主要集中在眼睛闭合、打呵欠模式、面部表情和头部运动等方面。

间接的车辆行为如方向盘、横向位置也可以表征驾驶员的警惕性, 例如在短时间的空位后突然修正可能是警惕性丧失的迹象。例如, 文献[9]使用混沌理论分析转向运动, 以推断驾驶员的疲劳程度。此外, 由于昏昏欲睡的驾驶员无法精确操纵方向盘, 驾驶员很可能将车辆驶出车道中心。所以, 通过监控车道和行驶路径的偏差, 也可以测量驾驶员的警惕状态。

当人们感到疲劳时, 可以从头部、面部和眼睛观察到各种明显的视觉行为, 如频繁地点头、较小程度地睁眼 (甚至闭眼)、眨眼、打哈欠等^[10]。随着计算机视觉技术的发展, 这种非侵入性的办法能有效地提取视觉特征, 通过分析从正面采取的图像中提取到的特征, 可以依此推断司机的状态。具体来说, 文献[11]设计的头部姿势检测方法可以用来测量睡意。估计驾驶员头部的方位, 并将头部姿势信息应用于驾驶员辅助系统。眼睛状态通常可用于测量驾驶员的困倦程度, 方法类似于计算闭眼百分比 (PERCLOS)^[12]、闭眼频率 (FEC) 和闭眼持续时间 (ECD)。文献[13]设计了一种基于打哈欠的方法, 通常是先定位驾驶员的嘴, 然后将嘴的状态分类为正常或打哈欠, 通过打哈欠的频率来判断驾驶员的精神状态。而文献[14]设计了一种基于面部表情的方法, 主要是将打哈欠、眨眼、扬眉等面部暗示结合起来, 分析表面上的突出点以确定疲劳程度。

以上方法虽然能很好地提取驾驶员的脸部特征, 并对驾驶员的疲劳状态进行检测, 但需要先对人脸进行检测, 然后再训练网络对人眼和嘴巴的张开和闭合状态进行识别。这样不仅需要更多的关于眼睛和嘴巴的数据集进行训练, 而且识别眼睛和嘴巴状态的过程中会因为光照等其他原因进一步产生误差。本文基于人脸特征点模型设计了一种计算规则来对眼睛和嘴巴的状态进行判断, 无须额外的深度学习网络训练来识别眼睛和嘴巴的状态。再使用 PERCLOS^[14]算法和采集到的打哈欠频率对驾驶员是否疲劳进行综

合判断。

1 人脸检测

本文人脸检测方法使用的是基于 HOG 特征和利用 SVM 训练高度精确的目标分类器方法^[15], 再在人脸检测的基础上使用面部标志算法^[16]检测出眼睛和嘴巴的特征点。

1.1 HOG 特征提取

在减小光照和背景因素对目标检测的影响这一方面, HOG 描述算子相比于其他的特征集表现得更好。HOG 特征利用边缘梯度和拐角的方向来描述物体的形状, 而梯度对光照有很好的鲁棒性, 因此, HOG 特征常用于目标检测当中。提取 HOG 特征的主要步骤如下:

a) 首先需要对获取的图像进行归一化和 Gamma 校正的预处理。Gamma 变换能通过调节图像对比度, 对光照或阴影的影响有很好的调节作用。Gamma 正则化公式如下, 其中 $H(x, y)$ 表示像素点 (x, y) 的像素值。

$$H(x, y) = H(x, y)^{\text{Gamma}} \quad (1)$$

b) 计算每个像素点在水平方向和垂直方向的梯度, 分别用 $G_x(x, y)$ 、 $G_y(x, y)$ 表示。其中每个像素水平方向和垂直方向梯度的公式为

$$G_x(x, y) = H(x+1, y) - H(x-1, y) \quad (2)$$

$$G_y(x, y) = H(x, y+1) - H(x, y-1) \quad (3)$$

然后通过水平方向梯度和垂直方向梯度计算该像素点的梯度大小 $G(x, y)$ 和方向 $\theta(x, y)$, 计算公式为

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (4)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{G_y(x, y)}{G_x(x, y)} \right) \quad (5)$$

c) 将获取的图像分割成若干个单元, 每 8×8 个像素构成一个单元, 在每个单元内进行梯度大小和方向的统计, 并将结果做成直方图。将 $0 \sim 360^\circ$ 划分成九个区间的直方图通道, 并将九个区间作为直方图的横坐标。然后加权计算每个方向上的梯度大小, 将累加的梯度值作为直方图纵轴。

d) 由于光照和背景的变化具有多样性, 而每次光照或者背景发生较大变化时, 梯度强度的变化也会非常大, 这会影响到目标检测的准确率。为了增强模型对光照强度的鲁棒性, 将相邻的 M 个单元合并成区块, 对区块进行对比度归一化处理。再将区块内单元的特征向量进行组合, 形成 HOG 特征, 用于表征获取的整张图像。

1.2 训练 SVM 分类器

可以使用 HOG 描述算子和 SVM 来进行高精度的目标分类器训练, 其训练方法 1 如下:

方法 1 高精度目标分类器训练方法

a) 从训练数据集中采样 P 个含有人脸的正样本, 采样 N 个不含有人脸的负样本, 并提取正、负样本的 HOG 特征, 其中 $N \gg P$ 。

b) 用采集的正、负样本训练线性支持向量机。

c) 在任意负样本图像以及任意尺寸负样本图像训练集上, 进行窗口滑动。计算每个窗口的 HOG 特征并用步骤 b) 训练好的分类器进行分类。记录错误分类负样本的特征向量以及分类的

收稿日期: 2019-12-09; 修回日期: 2020-02-04 基金项目: 国家自然科学基金联合基金重点资助项目 (U1713203)

作者简介: 朱名流 (1994-), 男, 湖北武汉人, 硕士研究生, 主要研究方向为图像处理、深度学习; 李顶根 (1977-), 男 (通信作者), 安徽桐城人, 副教授, 硕士, 主要研究方向为智能驾驶、图像处理、深度学习 (lidinggen@hust.edu.cn)。

概率。

d) 用错误分类的负样本重新对步骤 b) 训练好的分类器再次进行训练,最后优化好的人脸检测模型。

最终人脸检测器检测到的人脸示意图如图 1 所示。

2 基于人脸特征点的疲劳检测

为了方便后面计算 PERCLOS 以及设计嘴部张度计算规则,本文选用文献[17]提出的具有回归树集合的人脸对齐方法,人脸特征点检测示意图如图 2 所示。然后通过设定眼睛高度与宽度距离比值的阈值和嘴巴张度的大小以及张开频率的阈值来综合判断驾驶员是否处于疲劳。

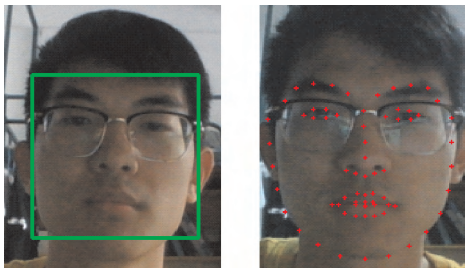


图 1 人脸检测示例

图 2 人脸特征点检测示例

2.1 基于 PERCLOS 的人眼疲劳检测

PERCLOS 被认为是最可靠和有效的驾驶员警觉性水平的测定指标,PERCLOS 是眼脸闭合的百分比,是指单位时间内测得的眼脸闭合的帧数占总帧数的比值,如图 3 所示^[18]。它的具体计算公式为

$$\text{PERCLOS} = \frac{N_{\text{close}}}{N} \times 100\% \quad (6)$$

其中: N_{close} 表示的 $t_3 - t_2$, 即眼睛闭合的时间; N 表示的 $t_4 - t_1$, 即眼睛从闭到最大开度 80% 到完全闭合,然后再睁开到最大开度 80% 的时间。而 PERCLOS 会根据眼睛闭合程度的不同而设定不同的阈值,对人眼闭合的程度分为三种情况: a) EM 为瞳孔面积的 50% 被眼脸遮住就定义为闭合; b) P70 为瞳孔面积的 70% 被眼脸遮住就定义为闭合; c) P80 为瞳孔面积的 80% 被眼脸遮住就定义为闭合。

不同阈值对应的 PERCLOS 值也不相同,在大多数的研究认为 P80 指标的效果最好,能很好地反映驾驶员是否处于疲劳状态,因此被广泛运用于疲劳检测的关键指标参数。由于摄像头采集图像的帧率是固定的,所以在单位时间内所采集到的图像数量也是不变的。因此要计算单位时间内的 PERCLOS 值,可以转换为该时间段内所采集图像中眼睛闭合时的连续帧数总和占单位时间内视频图像帧数总和的比值,计算 PERCLOS 值的公式如下:

$$\text{PERCLOS} = \frac{\text{单位时间内闭眼帧数总和}}{\text{单位时间内图像帧数总和}} \times 100\% \quad (7)$$

根据人脸特征点定位模型输出的人眼的六个特征点,如图 4 所示。本文通过计算人眼纵横比 EAR 来代替瞳孔面积被眼脸遮挡的程度,人眼一次眨眼过程中 EAR 值的变化如图 5 所示。具体的人眼纵横比的公式为

$$\text{EAR} = \frac{\|P_6 - P_2\| + \|P_5 - P_3\|}{2 \|P_4 - P_1\|} \quad (8)$$

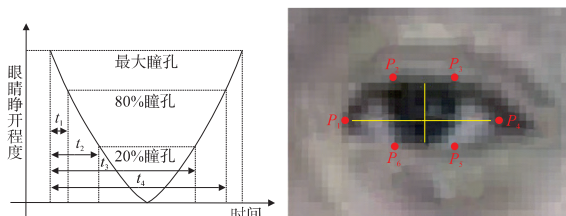


图 3 PERCLOS 原理图

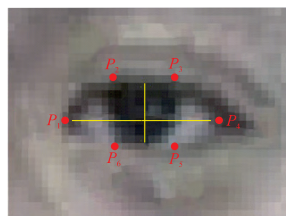


图 4 人眼特征点示例

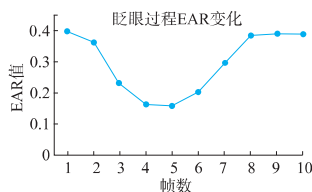


图 5 人眼特征点示例

根据不同的人眼闭合程度,设定合适的 EAR 阈值,取左、右眼分别计算 EAR,然后取平均值作为最终的纵横比值,通过与阈值的比较,进行人眼的眨眼检测,再配合 PERCLOS 算法进行眼部疲劳检测。

2.2 嘴部哈欠检测

嘴部特征点的检测示例图如图 6 所示。由于嘴部内部轮廓相比外部轮廓能更好地反映嘴部张度变化,所以,可以选取 61 ~ 68 这八个点来进行嘴部张度的计算。计算公式为

$$M = \frac{\|P_{63} - P_{67}\|}{\|P_{65} - P_{61}\|} \quad (9)$$



图 6 嘴部特征点示例

由于打哈欠和平时说话以及惊讶时,嘴巴都会有一定张度变化,所以要区分打哈欠和其他的情况下嘴巴张度的变化。为了能对打哈欠和其他情况进行区分,可以将嘴巴张开的程度和张开的时间作为区分的依据。设定超过张度 M 的连续帧数为 T ,当连续的帧数 T 超过一定的阈值时,则认为在接下来的 N 秒处于哈欠状态。一次正常哈欠的嘴部张度变化如图 7 所示。

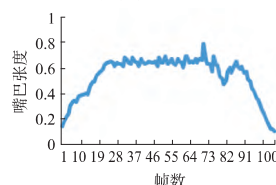


图 7 打哈欠时嘴部张度变化

结合 PERCLOS 算法,在单位时间内,如果检测到 PERCLOS 的值大于阈值 X ,未检测到哈欠状态,则认为处于轻度疲劳状态;如果 PERCLOS 值未超过阈值 X ,检测到哈欠状态,则认为处于中度疲劳状态;如果 PERCLOS 值超过了 X ,而且同时检测到了哈欠状态,则认为处于重度疲劳状态,需要提醒驾驶员提车休息,或者采取强制措施。

3 实验与分析

本文所有实验均是基于 Python 3.5 和 TensorFlow 1.1 实现的。摄像头使用的是普通摄像头, FPS 为 30, 处理器为 CPU (3.2 GHz), 内存大小为 4 GB。为了验证本文疲劳检测方法能否达到效果,分别进行了以下实验。

3.1 EAR 阈值确定实验

为了确定正常情况下,人睁开时 EAR 值的范围,总共选取六位测试人员,依次获取每个人正常眨眼的图片 300 帧,并统计 EAR 的最大值和最小值范围。经过多次实验,确定最大张度的 EAR 均值为 0.4,闭眼时的 EAR 均值为 0.15。再经过下面的公式来计算睁眼和闭眼的 EAR 阈值,计算公式为

$$\text{EAR}_{\text{open}} = (\text{EAR}_{\text{max}} - \text{EAR}_{\text{min}}) \times (1 - X_1) + \text{EAR}_{\text{min}} \quad (10)$$

$$\text{EAR}_{\text{close}} = (\text{EAR}_{\text{max}} - \text{EAR}_{\text{min}}) \times (1 - X_2) + \text{EAR}_{\text{min}} \quad (11)$$

经过大量实验,取 X_1 为 0.3,即人眼 EAR 的值达到最大张度的 70% 就认为处于睁开状态,带入式 (10),计算得到人眼睁开状态的 EAR 阈值为 0.35。取 X_2 为 0.7,即人眼 EAR 值达到最大张度的 30% 就认为处于闭合状态,带入式 (11),计算得到人眼闭合状态的 EAR 阈值为 0.225。

3.2 眨眼检测实验

本文眨眼检测实验在 Talking-Face 数据集以及自己采集数据集上进行。Talking-Face 数据集包含 5 000 张图片,总共眨眼 68 次。自己采集的数据包含 1 878 张图片,总共眨眼 33 次。实验结果如表 1 所示。本文实验结果与黑色像素值方法在 Talking-Face 上的对比结果如表 2 所示。由此可以得出本文所设定的 EAR 阈值以及所使用的眨眼检测方法具有很好的效果。

表 1 眨眼检测实验结果

数据集	实际眨眼次数	检测眨眼次数	准确率/%
Talking-Face	61	68	89
自制数据集	33	38	85

表 2 不同方法在 Talking-Face 上的实验结果对比

使用方法	准确率/%
本文方法	89
黑色像素值	77.94

3.3 疲劳检测实验

本文疲劳检测实验在数据集 YawDD 数据集上进行, YawDD 数据集包含亚洲和欧美驾驶员在不同光照下模拟驾驶视频,还包括不同性别以及佩戴眼镜的驾驶视频。

为区分哈欠和张嘴说话的情况,由于这两种情况的嘴部张度大小和持续的帧数存在差异,正常打哈欠的时间为 2~3 s,而张嘴说话或者笑的时间不超过 1 s。所以本文取连续帧数 T 为 30,即 1 s。超过阈值 M 的连续帧数超过 30 帧时,则判定为哈欠状态。由于在检测过程中存在误检和漏检,所以引入精确率 (precision)、召回率 (recall) 和 F_1 -measure 来对检测结果进行评估。具体的计算公式为

$$P = \frac{TP}{TP + FP} \quad (12)$$

$$R = \frac{TP}{TP + FN} \quad (13)$$

$$F_1 = \frac{2PR}{P + R} \quad (14)$$

其中: P 表示精确率; R 表示召回率; F_1 表示 F_1 -measure, 是对精确率和召回率的加权调和平均能更好地反映检测结果, F_1 -measure 越高,说明实验结果越理想; TP 表示正类判定为正类,即打哈欠被检测为打哈欠的次数; FP 表示负类判定为正类,即非打哈欠状态被检测为打哈欠的次数; FN 表示正类判定为负类,即漏检次数。

针对嘴部张度阈值 M 确定进行了多次实验,实验结果如表 3 所示。由表 3 可以得出当张度阈值 M 取值为 0.45,连续帧数取值为 30 时,哈欠的检测效果最为理想。

表 3 张度阈值确定实验

M 取值	实际打哈欠次数	总共检测到的打哈欠次数	漏检次数	误检次数	P (精确率)	R (召回率)	F_1 -measure
0.4	27	31	2	6	0.81	0.93	0.87
0.45	27	24	4	1	0.96	0.85	0.90
0.5	27	23	6	2	0.91	0.78	0.84
0.6	27	18	10	1	0.94	0.63	0.75

疲劳检测的测试集选取 YawDD 数据集中配戴/未配戴眼镜的男性和女性驾驶员视频进行测试。以往的疲劳检测方法多采取给予 PERCLOS 值和哈欠次数一定权重的办法,由于两项特征所赋予的权重严重失衡,会使得两种特征中的一项直接决定疲劳检测的结果。

所以,本文疲劳检测会对多特征的融合进行改进,从而对疲劳检测的结果进行综合判断。由文献[19]中的研究结论可以知道:当驾驶员的 PERCLOS 值小于 0.1 时,处于非疲劳状态;当驾驶员的 PERCLOS 值在 0.1~0.3 时,处于轻度疲劳状态;当驾驶员的 PERCLOS 值在 0.3~0.5 时,处于中度疲劳状态;当 PERCLOS 值大于 0.5 时,则处于重度疲劳状态。在本文的检测方法中,取 PERCLOS 的阈值为 0.4,当一定时间内 PERCLOS 值大于 0.4 时,则认为是处于疲劳状态。在获取的视频中,每 10 s 检测一次 PERCLOS 值,同时检测这 10 s 内驾驶员是否有打哈欠的现象。如果 PERCLOS 值低于 0.4,而且未检测到打哈欠,则认为处于非疲劳状态;如果检测到 PERCLOS 值大于 0.4,但未检测到打哈欠,则认为处于轻度疲劳状态;当 PERCLOS 值低于 0.4,但检测到了打哈欠时,则认为处于中度疲劳状态;当 PERCLOS 值大于 0.4,而且同时检测到了打哈欠,则认为处于重度疲劳状态。该判断决策的流程如图 8 所示。

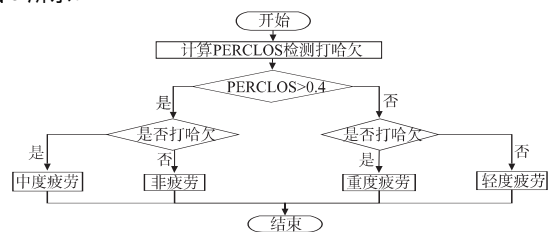


图 8 多特征融合决策流程

本文选取了 YawDD 数据集中配戴/未配戴眼镜的男女驾驶员视频各一段进行疲劳检测。检测结果如表 4 所示。由表 4 可以看出在打哈欠或者 PERCLOS 值大于 0.4 时,本文方法能准确发出预警,且对性别、种族、光照和是否佩戴眼镜具有很好的鲁棒性。

表 4 疲劳检测结果

视频	PERCLOS 均值	实际打哈欠次数	检测到的打哈欠次数	预警数
no glass female	0.28	2	2	2
no glass male	0.23	2	2	3
glass female	0.24	4	4	4
glass male	0.25	3	3	3

4 结束语

本文方法通过利用 HOG 特征检测视频中的人脸,再利用人脸特征点模型对人眼和嘴部进行定位。通过计算人眼纵横比 EAR 进行眨眼检测并在 Talking-Face 数据及上进行测试,由实验得出对眨眼具有好的检测效果。通过计算嘴部张度进行哈欠检测并在 YawDD 数据及上进行哈欠检测实验,由实验结果得出,本文方法对哈欠检测的 F_1 -measure 指标为 0.9,能较好地识别打哈欠的状态。最后通过 PERCLOS 和打哈欠次数进行多特征融合对疲劳驾驶进行综合检测。通过实验表明,本文方法在疲劳驾驶检测方面性能良好。

参考文献:

- [1] NHTSA. Drowsy driving [R/OL]. (2017-07-02). <https://www.nhtsa.gov/risky-driving/drowsy-driving>.
- [2] Navarro J, Mars F, Young M S. Lateral control assistance in car driving: classification, review and future prospects [J]. IET Intelligent Transport Systems, 2011, 5(3): 207-220.
- [3] Dai Jiangpeng, Teng Jin, Bai Xiaole, et al. Mobile phone based drunk driving detection [C] // Proc of the 4th International Conference on Pervasive Computing Technologies for Healthcare. Piscataway, NJ: IEEE Press, 2010: 1-8.
- [4] Malik H, Naem F, Zuberi Z, et al. Vision based driving simulation [C] // Proc of International Conference on Cyberworlds. Piscataway, NJ: IEEE Press, 2004: 255-259.
- [5] Mardi Z, Ashtiani S N M, Mikaili M. EEG-based drowsiness detection for safe driving using chaotic features and statistical tests [J]. Journal of medical signals and sensors, 2011, 1(2): 130-137.
- [6] Danisman T, Bilasco I M, Djeraba C, et al. Drowsy driver detection system using eye blink patterns [C] // Proc of International Conference on Machine and Web Intelligence. Piscataway, NJ: IEEE Press, 2010: 230-233.
- [7] Kaplan S, Guvansan M A, Yavuz A G, et al. Driver behavior analysis for safe driving: a survey [J]. IEEE Trans on Intelligent Transportation Systems, 2015, 16(6): 3017-3032.
- [8] Takei Y, Furukawa Y. Estimate of driver's fatigue through steering motion [C] // Proc of IEEE International Conference on Systems, Man and Cybernetics, 2005: 1765-1770.
- [9] Bergasa L M, Nuevo J, Sotelo M A, et al. Real-time system for monitoring driver vigilance [J]. IEEE Trans on Intelligent Transportation Systems, 2006, 7(1): 63-77.
- [10] Murphy-Chutorian E, Doshi A, Trivedi M M. Head pose estimation for driver assistance systems: a robust algorithm and experimental evaluation [C] // Proc of IEEE Intelligent Transportation Systems Conference. Piscataway, NJ: IEEE Press, 2007: 709-714.
- [11] Banks S. Behavioral and physiological consequences of sleep restriction [J]. Journal of Clinical Sleep Medicine, 2007, 3(5): 519-528.
- [12] D' Orazio T, Leo M, Guaragnella C, et al. A visual approach for driver inattention detection [J]. Pattern Recognition, 2007, 40(8): 2341-2355.
- [13] Saradadevi M, Bajaj P. Driver fatigue detection using mouth and yawning analysis [J]. International Journal of Computer Science and Network Security, 2008, 8(6): 183-188.
- [14] Jiménez-Pinto J, Torres-Torriti M. Driver alert state and fatigue detection by salient points analysis [C] // Proc of IEEE International Conference on Systems, Man and Cybernetics. Piscataway, NJ: IEEE Press, 2009: 455-461.
- [15] Bergasa L M, Nuevo J, Sotelo M A, et al. Real-time system for monitoring driver vigilance [J]. IEEE Trans on Intelligent Transportation Systems, 2006, 7(1): 63-77.
- [16] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C] // Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005: 886-893.
- [17] Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees [C] // Proc of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2014: 1867-1874.
- [18] 郭永彩, 李文涛, 高潮. 基于 PERCLOS 的驾驶员疲劳检测算法 [J]. 计算机系统应用, 2009, 18(8): 54-57.
- [19] 汪磊, 孙瑞山. 基于面部特征识别的管制员疲劳监测方法研究 [J]. 中国安全科学学报, 2012, 22(7): 66-71.