



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2013
08
VOL.18

ISSN1006-8961
CN11-3758/TB



北魏



隋



初唐

数字文化遗产研究 P968

综述

尺度不变特征变换算子综述

刘立, 詹茵茵, 罗扬, 刘朝晖, 彭复员 885

图像处理和编码

改进的H.264帧间模式选择算法

任克强, 张旭光, 罗会兰 893

均质雾天下摄像机动态标定的车速估计

宋洪军, 陈阳舟, 郜园园 901

基于人脸五官特征的空域差错掩盖算法

张江鑫, 谢晋, 邝万坤 913

采用圆谐-傅里叶矩的图像区域复制粘贴篡改检测

秦娟, 李峰, 向凌云, 殷蓓蓓 919

图像分析和识别

基于视频感知哈希的视频篡改检测与多粒度定位

朱映映, 文振焜, 杜以华, 邓良太 924

视觉注意模型的道路监控视频关键帧提取

刘云鹏, 张三元, 王仁芳, 张引 933

格拉斯曼流形上的半监督判别分析

姜伟, 陆瑶, 杨炳儒 944

精神疲劳实时监测中多面部特征时序分类模型

陈云华, 张灵, 丁伍洋, 严明玉 953

图像分类中的概率乘积核函数

杨赛, 赵春霞 961

轮廓整体结构约束的壁画图像相似性度量

唐大伟, 鲁东明, 杨冰, 许端清 968

应用色彩空间聚类方法实现道路建模

向宸薇, 王拓, 于舰 976

图像理解和计算机视觉

杂乱背景和摄像机移动下的时空兴趣点检测

刘长红, 陈勇, 王明文 982

形状和颜色混合不变量在图像检索中的应用

公明, 曹伟国, 李华 990

计算机图形学

结合统计分类和边缘特征的最优视图提取

潘翔, 陈敖, 章国栋 1004

医学图像处理

局部熵驱动下的脑MR图像分割与偏移场恢复耦合模型

张建伟, 杨红, 陈允杰, 方林, 詹天明 1011

改进的模糊聚类实质肺结节3维分割

李翠芳, 聂生东, 王远军, 孙希文, 郑斌 1019

遥感图像处理

高光谱海岸带区域分割的活动轮廓模型

王相海, 金弋博 1031

改进的基于非局部均值的极化SAR相干斑抑制

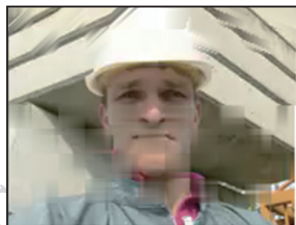
赵忠民, 赵拥军, 牛朝阳 1038

小波变换和稀疏表示相结合的遥感图像融合

刘婷, 程建 1045

《中国图象图形学报》论文摘要写作要求

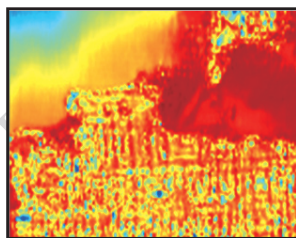
1054



基于人脸五官特征的空域差错掩盖算法 (第913页)



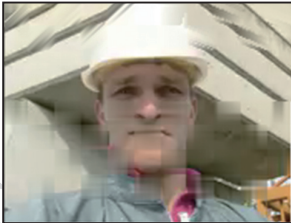
应用色彩空间聚类方法实现道路建模 (第976页)



改进的基于非局部均值的极化SAR相干斑抑制 (第1038页)

CONTENTS

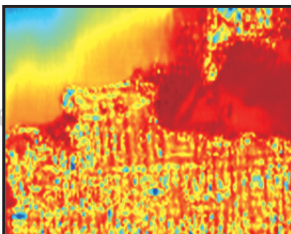
JOURNAL OF IMAGE AND GRAPHICS



Spatial error concealment algorithm based on the facial features (P913)



Road modeling using color spatial clustering (P976)



Improved polarimetric SAR speckle filter based on non-local means (P1038)

Review

- Summarization of the scale invariant feature transform
Liu Li , Zhan Yinyin, Luo Yang, Liu Chaohui, Peng Fuyuan 885

Image Processing and Coding

- Improved inter frame mode selection algorithm for H.264
Ren Keqiang, Zhang Xuguang, Luo Huilan 893
- Vehicle velocity estimation based on camera calibration under homogenous fog
Song Hongjun, Chen Yangzhou, Gao Yuanyuan 901
- Spatial error concealment algorithm based on the facial features
Zhang Jiangxin, Xie Jin, Kuang Wankun 913
- Detection of image region copy-move forgery using radial harmonic Fourier moments
Qin Juan, Li Feng, Xiang Lingyun, Yin Changming 919

Image Analysis and Recognition

- Video forgery detection and multi-granularity location based on video perceptual hashing
Zhu Yingying, Wen Zhenkun, Du Yihua, Deng Liangtai 924
- Key frame extraction based on the visual attention model for lane surveillance video
Liu Yunpeng, Zhang Sanyuan, Wang Renfang, Zhang Yin 933
- Semi-supervised discriminant analysis on Grassmannian manifold
Jiang Wei, Lu Yao, Yang Bingru 944
- Time-series classification model based on multiple facial feature for real-time mental fatigue monitoring
Chen Yunhua, Zhang Ling, Ding Wuyang, Yan Mingyu 953
- Probability product kernels in image classification
Yang Sai, Zhao Chunxia 961
- Similarity metrics between mural images with constraints of the overall structure of contours
Tang Dawei, Lu Dongming, Yang Bing, Xu Duanqing 968
- Road modeling using color spatial clustering
Xiang Chenwei, Wang Tuo, Yu Jian 976

Image Understanding and Computer Vision

- Spatio-temporal interest point detection in cluttered backgrounds with camera movements
Liu Changhong, Chen Yong, Wang Mingwen 982
- Application of combined shape and color invariants to image retrieval
Gong Ming, Cao Weiguo, Li Hua 990

Computer Graphics

- Best view selection based on statistical classification and edge features
Pan Xiang, Chen Ao, Zhang Guodong 1004

Medical Image Processing

- Brain MR image segmentation and bias correction model based on local entropy
Zhang Jianwei, Yang Hong, Chen Yunjie, Fang Ling, Zhan Tianming 1011
- Segmentation of sub-solid pulmonary nodules based on improved fuzzy C-means clustering
Li Cuifang, Nie Shengdong, Wang Yuanjun, Sun Xiwen, Zheng Bin 1019

Remote Sensing Image Processing

- The active contour model for segmentation of coastal hyperspectral remote sensing image
Wang Xianghai, Yin Yibo 1031
- Improved polarimetric SAR speckle filter based on non-local means
Zhao Zhongmin, Zhao Yongjun, Niu Chaoyang 1038
- Remote sensing image fusion with wavelet transform and sparse representation
Liu Ting, Cheng Jian 1045

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2013)08-0953-08

论文引用格式: 陈云华, 张灵, 丁伍洋, 严明玉. 精神疲劳实时监测中多面部特征时序分类模型[J]. 中国图象图形学报, 2013, 18(8): 953-960.
 [DOI:10.11843/jig.20130809]

精神疲劳实时监测中多面部特征时序分类模型

陈云华, 张灵, 丁伍洋, 严明玉

广东工业大学计算机学院计算机工程系, 广州 510006

摘要: 针对现有疲劳监测方法仅根据单帧图像嘴巴形态进行哈欠识别准确率低, 采用阈值法分析眨眼参数适应性较差, 无法对疲劳状态的过渡进行实时监测等问题, 提出一种新的进行精神疲劳实时监测的多面部特征时序分类模型。首先, 通过面部视觉特征提取张口度曲线与虹膜似圆比曲线; 然后, 采用滑动窗口分段、隐马尔可夫模型(HMM)建模等方法在张口度曲线的基础上构建哈欠特征时序并进行类别标记, 在虹膜似圆比曲线的基础上构建眨眼持续时间时序并进行类别标记; 最后, 在 HMM 的基础上增加时间戳, 以便自适应地选取时序初始时刻点并进行多个特征时序的同步与标记结果的融合。实验结果表明, 本文模型可降低哈欠误判率, 对不同年龄的人群眨眼具有很好的适应性, 并可实现对精神疲劳过渡状态的实时监测。

关键词: 实时精神疲劳监测; 虹膜似圆比; 张口度; 时间序列; 滑动窗口法; 隐马尔可夫模型

Time-series classification model based on multiple facial feature for real-time mental fatigue monitoring

Chen Yunhua, Zhang Ling, Ding Wuyang, Yan Mingyu

Department of Computer Engineering, Guangdong University of Technology, Guangzhou 510006, China

Abstract: In computer vision based fatigue monitoring, there are still some unresolved issues remained, including low recognition accuracy in yawn detection based on a single-frame; poor adaptability in blink analysis because of the required threshold, the inability to monitor the transition stages of fatigue in real-time. Attempted to solve these problems, we propose a new classification model in this paper, which is based on two feature time-series for real-time mental fatigue monitoring. First, the mouth opening degree and iris circularity ratio are calculated through facial visual feature extraction. Based on this, we can generate a corresponding time-series called α (the proportion of the time during which mouth opening exceeds a given threshold) time series and eye blink time (EBT) time series. Then, using sliding window to partition and annotate the two kinds of time series and build hidden markov model (HMM) for EBT time series. Finally, add a time stamp on HMM to adaptively calculate the initial time point of the next time series, in addition, we can use it to perform the synchronization and fusion of the two time series. Experimental results show that the promoted model can improve yawn detection rate, have good adaptability for blink features of different age groups, and can monitor the transition stage of mental fatigue in real-time.

Key words: real time mental fatigue monitoring; iris circularity ratio; mouth opening degree; time series; sliding window; hidden Markov model (HMM)

收稿日期: 2012-10-30; 修回日期: 2013-01-15

基金项目: 国家自然科学基金项目(60272089); 广东省科技计划国际合作项目(2010B050400007)

第一作者简介: 陈云华(1977—), 女, 讲师, 广东工业大学控制理论与控制工程专业在读博士研究生, 主要研究方向为计算机视觉、模式识别、可拓检测。E-mail: chen_yh_winy@163.com

0 引言

对人体精神疲劳状态的检测包括主观评价法、生理信号检测法、基于计算机视觉技术的疲劳检测等,其中,基于计算机视觉技术的疲劳检测由于其具有非侵扰性而成为疲劳检测中的重要研究方向。该方法主要通过实时采集被测者视频图像,综合运用计算机视觉、图像处理、模式识别等技术,获取能够反映被测者疲劳状态的头部运动信息、眼睛与嘴巴的活动信息等,进行疲劳识别。反映疲劳状态的眼部信息包括 PERCLOS (percentage of eyelid closure over the pupil, over time), 即单位时间内眼睛闭合一定比例所占的时间; AECS (average eye closure speed), 即眼皮的平均闭合速度,一般用眼皮从完全睁开到完全闭合所经历的时间来表示; 眨眼频率; 瞳孔直径变化; 视线方向^[1-5]等。其中,瞳孔直径变化、视线方向估计等方法对视频采集设备的要求很高; 而基于 PERCLOS、AECS、眨眼频率等的疲劳识别则需要给出相应的阈值(区间), 研究表明,不同年龄层次的人在同样的精神状态下眨眼的次数和眼皮闭合速度有明显区别^[6], 因此,采用阈值分析眨眼参数进行疲劳判定适应性较差。

反映疲劳状态的嘴部信息主要是哈欠^[7-8]。文献[7]先从单幅图像中分割出嘴巴,根据嘴巴区域大小判断哈欠的产生。文献[8]则先定位两边嘴角,然后检测嘴唇的上下边界获取嘴巴的高宽比,并通过高宽比判断哈欠的产生。上述方法在判别哈欠时均未对张口的持续时间进行分析,导致瞬时的嘴巴张大如惊讶时张大嘴巴会被误判为哈欠。

人体精神疲劳是一个随时间发展的过程,在现有面部特征的基础上增加时间维度将有利于分析出眼、嘴等特征随时间变化的规律,提高精神疲劳辨识的准确性。文献[9]将隐马尔可夫模型(HMM)用于对驾驶员疲劳状态进行建模,在模拟驾驶环境中拍摄视频,以 PERCLOS 值为观测对象,每隔 6 min 生成一个观测节点,共 30 个观测节点,训练一个 HMM 模型,由此得到整个驾驶时间段内的驾驶疲劳状态序列。然而,疲劳监测的目的是当监测到人体处于疲劳状态时能够及时给予警告,而不是等整个过程结束后,评价该过程是一个怎样的状态序列过程。因此,该方法对于智能汽车空间的上下文感知具有一定的意义,但却不适用于人体疲劳的实时

监测。

综上,以时间序列和 HMM 理论为基础,建立一种面向精神疲劳实时监测的面部特征时序模型。首先,获取面部视觉特征,以视觉特征为基础计算能够反映精神疲劳状态的参数,以参数为基础生成时间序列。然后,对时间序列进行分段和索引标记;利用 HMM 对各种疲劳状态序列进行建模,为了实现精神疲劳的实时监测,在现有 HMM 模型的基础上增加时间戳,使得眼部序列与嘴部序列的识别能够同步,并且可根据当前疲劳状态自适应地选择下一个序列的起始时刻点。最后,在实时监测框架下进行单个特征时序的识别和多个特征时序识别结果的融合。实验结果表明,本文模型可降低哈欠误判率,对不同年龄的人群的眨眼具有很好的适应性,并可实现精神疲劳的准确实时监测。

1 面部视觉特征提取

1.1 虹膜分割与似圆比计算

首先,采用 harr-like 级联分类器^[10]进行人眼检测。将图像灰度化,对检测到的人眼区域进行直方图统计,为消除噪声数据的影响,对直方图进行均值滤波,滤波器的窗口大小可根据实验进行选取。由于眼部灰度图中虹膜区域的亮度值偏低,而且占较大比例,因此滤波后直方图中左边第 1 个峰值处对应的亮度值应为虹膜亮度值,将其作为种子,进行区域生长操作,区域生长的阈值根据实验结果进行选取。根据区域生长的结果将图像二值化,在二值化图中查找轮廓。采用文献[11]中的方法检测并存储轮廓,找到长度最长的轮廓,对该轮廓上的点利用最小二乘法进行椭圆曲线拟合,即可求得露出眼皮外的虹膜的拟合椭圆,如图 1 所示。



图 1 露出眼皮外的虹膜轮廓拟合结果

Fig. 1 Contour curve fitting results of iris exposed outside the eyelids

设椭圆的水平方向的半轴长为 a_E , 垂直方向的半轴长为 b_E (由于露出眼皮的虹膜水平宽度大于垂

直宽度,故 $a_E > b_E$),则虹膜似圆比

$$r = \frac{\pi a_E b_E}{\pi a_E^2} = \frac{b_E}{a_E} \quad (1)$$

1.2 嘴唇内轮廓角点检测与张口度计算

采用文献[12]中的唇色模型,得到嘴巴区域的位置,采用文献[13]中方法对嘴巴区域进行角点检测,得到嘴唇内轮廓上的点,利用最小二乘法对这些点进行椭圆拟合,得到嘴唇内轮廓曲线如图2所示。

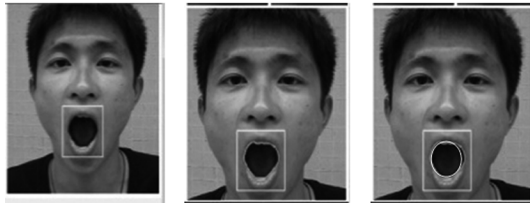


图2 嘴唇内轮廓角点曲线拟合结果

Fig. 2 Curve fitting results for corner points on the inside mouth contour

设 b_M 表示嘴唇内轮廓椭圆的垂直方向的半轴长, a_M 表示嘴唇内轮廓椭圆的水平方向的半轴长,则张口度

$$o = \frac{b_M}{a_M} \quad (2)$$

2 疲劳特征的时序构建与分类

2.1 α -时间序列建模与类别标记

2.1.1 张口度 o -曲线

对在讲话过程中分别产生了一次不同深度哈欠的两段典型视频求取张口度 o ,其曲线如图3所示。

图3中实线曲线超过阈值 T_2 的部分对应的是—次开口度大,且大开口度持续时间较长的深哈欠,

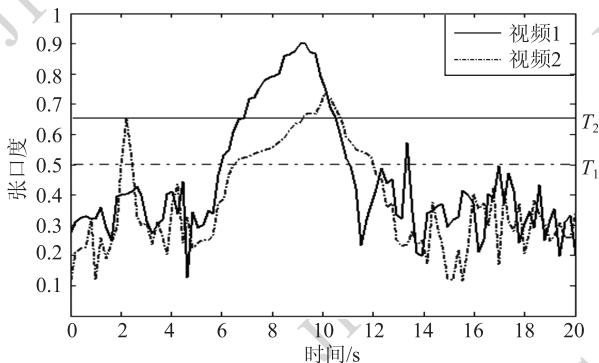


图3 两种典型哈欠的张口度曲线

Fig. 3 Mouth opening degree curves of two typical yawning

虚线曲线中,左边超过阈值 T_1 的部分对应的是一次瞬时嘴巴张大过程,右边超过阈值 T_1 的部分对应的是一次开口度较大,且较大开口度持续时间较长的浅哈欠。由此可见,哈欠不仅与张口度大小有关,还与张口度持续时间有关。以张口度 o -曲线为基础,采用阈值法与加时间窗的方式提取张口度的 α 特征,并以其为基础构建时间序列。

2.1.2 α -时序构建

在长度为 Lt s (Lt 值通过实验的方式获取,取普通人哈欠持续时间的均值)的时间窗内,利用阈值 o_T 对张口度值进行二值化,即

$$o_b = \begin{cases} 1 & o \geq o_T \\ 0 & o < o_T \end{cases} \quad (3)$$

时间窗内连续1的个数占二进制总位数的比例达到某个值 $\alpha_T \in (0.5, 1]$,才判断为哈欠, α_T 值可根据实验进行选取。设 l_i^i 表示 t_i 时刻时间窗内连续1的个数, $l_{\max}^i$ 表示最长的连续1的个数,则 $l_{\max}^i = \max \{l_i^i\}$,将时间窗内连续1的最大个数占总二进制位数的比例 α 沿时间轴构成时间序列,记为 $Y = \langle y_1 = (t_1, \alpha(t_1)), y_2 = (t_2, \alpha(t_2)), \dots, y_n = (t_n, \alpha(t_n)) \rangle$,且

$$\alpha(t_i) = \begin{cases} l_{\max}^i / L & l_{\max}^i \geq \alpha_T L \\ 0 & l_{\max}^i < \alpha_T L \end{cases} \quad (4)$$

L 为时间窗内二进制总位数,是一个常数,由此可得到 α -时序中的3种模式如图4所示。

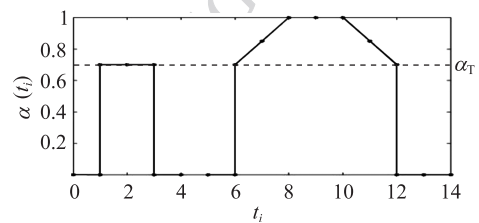


图4 张口度时序中的3种模式

Fig. 4 Three patterns in mouth opening degree time series

图4中左边矩形形状的连线段表示的是一个持续时间小于时间窗长度 Lt ,大于 αLt 的呵欠;右边的多边形连线段表示的是一个持续时间大于时间窗长度 Lt 的哈欠;横轴上点的连线表示未检测到哈欠。显然,这3种模式分别对应浅哈欠、深哈欠和无哈欠3种状态。

2.1.3 α -时序分类

1) α -时序分段与分段标记

由于 α -时间序列的模式简单,只有3种,如图4所示,在采用滑动窗口^[14]进行分段的过程中,只需

要按如下步骤进行,即可在分段的同时进行分段标记。

(1) 判断滑动窗口中第 1 个数据 $\alpha(t_1)$ 的值,开始一个分段 $p_j (1 \leq j \leq n)$ 表示分段的编号,即

$$v(p_j) = \begin{cases} 0 & \alpha(t_1) = 0 \\ 1 & \alpha(t_1) = \alpha_T \\ 2 & \alpha(t_1) > \alpha_T \end{cases} \quad (5)$$

(2) 从第 2 个数据开始,按照下面的规则判断每一个数据 $\alpha(t_i)$:

① 如果 $\alpha(t_i) \leq \alpha_T$, 且 $\alpha(t_i) = \alpha(t_{i-1})$, 则分段继续,窗口长度加 1 ($i = i + 1$), 读取下一个数据 $\alpha(t_i)$, 转到步骤(2);

② 如果 $\alpha(t_i) > \alpha_T$, 且 $\alpha(t_{i-1}) > \alpha_T$, 则分段继续,窗口长度加 1 ($i = i + 1$), 读取下一个数据 $\alpha(t_i)$, 转到步骤(2);

③ 否则,开始新分段, $i = i + 1$, 转到步骤(1)。

按照上述算法,每个分段对应 1 种哈欠模式,总共有 3 种哈欠模式:无哈欠、浅哈欠、深哈欠,对应的索引值分别为 0、1、2。

2) α -时序的类别标记

时长 T_{best} 的视频所对应的 α -时序的分段结果可表示为

$$Y_{\text{best}} = \{(v(p_1), t_1), (v(p_2), t_2), \dots, (v(p_K), t_K)\} \quad (6)$$

式中, K 表示分段的个数。 α -时序在分段过程中已经对每个分段进行了标记,标记值 0、1、2 分别代表该分段所对应的 3 种哈欠模式。 α -时序的类别索引为

$$\text{Index}(Y_{\text{best}}) = \max \{v(p_j)\} \quad (7)$$

式中, $j = 1, 2, \dots, K$ 表示分段编号,即某个 α -时序的标记值取该子序列中各个分段标记的最大值,因为一段时间内的精神疲劳具有随时间单调递增的特点。

2.2 EBT-时间序列建模与类别标记

2.2.1 虹膜似圆比 r -曲线

对一段包含了 5 次眨眼的视频计算虹膜似圆比 r , 得到如图 5 所示的曲线。

由图 5 可看出,虹膜似圆比 r 与眼睛睁开程度成正比,当眼睛睁开到最大时, r 值达到最大;当眼睛闭合时, r 值达到最小。可见,通过 r 值的变化可计算眨眼(睁到最大—闭合—睁到最大)持续时间(EBT),而 EBT 正是反映精神状态差异的首要特征,所以,以 EBT 沿时间轴方向的值构成时间序列,

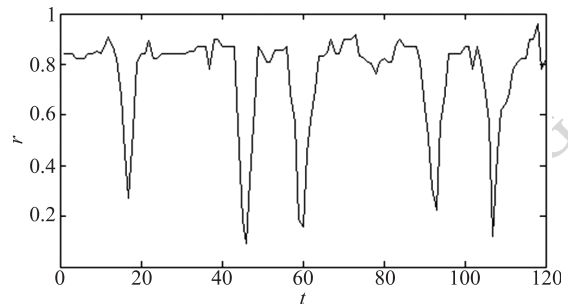


图 5 虹膜似圆比曲线

Fig. 5 Iris circularity ratio curve

对其建模。

2.2.2 EBT-时序表示

记 EBT 沿时间轴构成的时间序列为 $X = \langle x_1 = (bn_1, \text{EBT}(bn_1)), x_2 = (bn_2, \text{EBT}(bn_2)), \dots, x_n = (bn_n, \text{EBT}(bn_n)) \rangle$, 其中, bn_i 表示第 i 次眨眼, $\text{EBT}(bn_i)$ 表示第 i ($1 \leq i \leq N$) 次眨眼的持续时间, N 表示眨眼的总次数。

设 $t(x)_j$ 表示在 r -曲线中第 j 次出现 $r = x$ 的时刻点, 则

$$\text{EBT}(bn_i) = t(r_T)_{2i} - t(r_T)_{2i-1} \quad (8)$$

某个典型样本在不同精神状态下的 EBT-时序如图 6 所示。

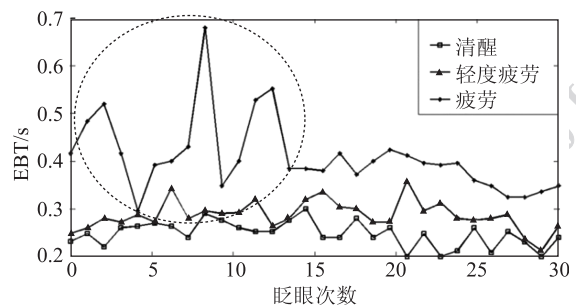


图 6 3 种精神状态对应的 EBT-时间时序

Fig. 6 EBT-time series corresponding to three mental states

从图 6 可看出:1) 每种精神状态时间序列中的点,各自呈现出较好的聚类特性,因此,可以估计出各种精神状态时序中点的统计分布模型;2) 图 6 中虚线圈处对应从轻度疲劳过渡到疲劳时出现的一种特定模式,同样可以估计出这种特定模式中点的分布模型。

2.2.3 EBT-时序建模与分类

1) EBT-时间序列分段与分段特征量提取

采用滑动窗口法^[14]对 EBT 时间序列进行分段。提取每个分段的均值和方差作为子序列的特征向量。每个 EBT 子序列对应的特征量为一个 1×2

的矢量, 整个 EBT 序列的特征量为一个 $K \times 4$ 的矩阵, 子序列 EBT_i 的特征量可以表示为 $\mathbf{O}_i = (\mu_i, \sigma_i)$, 整个子序列的特征量为 $\mathbf{O} = \{\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_K\}$, K 为序列分段个数。

2) EBT-时序的 HMM

HMM 具有很强的时序建模能力, 在语音识别、文字识别、步态身份识别等领域取得了非常广泛的应用^[15]。

记 HMM 的 N 个状态为 $\theta_1, \theta_2, \dots, \theta_N$, 每个状态对应的可能观察值为 $V_1, V_2, \dots, V_M$ (N 表示状态数; M 表示可能的观测值的个数); t 时刻马尔可夫链所处的状态记为 $\theta(t)$, 观测值计为 $V(t)$, $\theta^L = (\theta(1), \theta(2), \dots, \theta(L))$ 表示不能被直接观测到的隐状态序列, $V^L = (V(1), V(2), \dots, V(L))$ 表示能被直接观测到的序列, L 表示马尔可夫链的长度。则 HMM 可用 $\lambda = (\pi, A, B)$ 来表示, π 为初始状态概率, A 为状态转移分布矩阵, B 为观察值概率分布矩阵。在建立 HMM 时, 需要先根据样本的 V^L 对系统进行训练得到模型参数 λ , 再根据模型参数对隐状态序列 θ^L 所属的类别进行识别。

选取两个状态无跨越左右型的 HMM 结构, 如图 7 所示。在训练阶段, 针对精神疲劳状态序列“警醒→警醒”(模式编号 $PN=0$)、“警醒→轻度疲劳”($PN=1$)、“轻度疲劳→轻度疲劳”($PN=2$)、“轻度疲劳→疲劳”($PN=3$)、“疲劳→疲劳”($PN=4$) 分别建立 HMM。首先设定模型的初始参数, 然后通过样本训练得到重估后的 HMM 参数, 所有的模型参数构成系统的 HMM 参数库; 在识别阶段, EBT-序列经过分段与特征提取后, 输入到每种 HMM 中, 分别计算出各模型 λ 的参数匹配下的输出概率 $p(\lambda)$, 比较各输出概率的大小, 进而完成决策, 给出模式分类结果。其中, 模型参数训练采用 Baum-welch 算法^[16], 模型识别采用 Viterbi 算法^[16]。

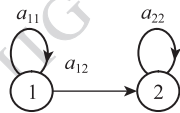


图 7 HMM 模型的结构

Fig. 7 Structure of HMM model

图 7 中, a_{ij} 表示从状态 i 转移到状态 j 的概率, 即结构模型的状态转移概率矩阵为

$$A = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} \quad (9)$$

设 $\alpha_i^h(t)$ 表示 HMM 模型 h ($1 \leq h \leq H$) 在 t 时刻位于状态 θ_i , 并且已经产生了 t 时刻为止的目标序列的概率, H 表示 HMM 模型的个数。

模型 h 产生观测序列 V^L 的概率的迭代公式为

$$\alpha_j^h(t) = b_{jk}^h V(t) \sum_{i=1}^N \alpha_i^h(t-1) a_{ij}^h \quad (10)$$

式中, $1 \leq j \leq N, 1 \leq t \leq L$ 。模型 h 产生观测序列 V^L 的概率为

$$P(V^L | \lambda_h) = \max_{1 \leq j \leq N} (\alpha_j^h(L)) \quad (11)$$

最有可能产生观测序列 V^L 的模型为

$$h^* = \arg \max_{1 \leq h \leq H} (P(V^L | \lambda_h)) \quad (12)$$

3) EBT-时序类别标记

时长为 T_{best} 的 EBT-时序的隐状态序列模式可表示为

$$\text{Index}(EBT_{\text{best}}) = \{i | p(\lambda_i) = \max_k p(\lambda_j)\} \quad (13)$$

式中, $k=0, 1, \dots, 4$ 表示隐状态序列模式的编号。

3 两种时序的同步与标记结果融合

3.1 时间戳与自适应初始时刻点求取

为了对 EBT-时序进行实时识别, 在现有 HMM 的基础上, 增加时间戳 t_{stamp} , 记录待识别的 EBT-时序的起始时刻点, 根据当前隐状态序列类别、当前隐状态序列与前一隐状态序列间的关系, 自适应地选取下一个初始时刻点到当前 EBT 起始时刻点的时间间隔。设待识别的第 i 段 EBT-时序的起始时刻点为 t_{stamp}^i , 待识别的第 $i+1$ 段与第 i 段 EBT-时序起始时刻的间隔时间为 T_{inter}^{i+1} , 则

$$t_{\text{stamp}}^{i+1} = t_{\text{stamp}}^i + T_{\text{inter}}^{i+1} \quad (14)$$

$$T_{\text{inter}}^{i+1} = 2^{4-\text{val}(\lambda_i)} P(j \cdot 2^{1+\text{val}(\lambda_i)-\text{val}(\lambda_{i-1})}) T_{\text{inter}}^i \quad (15)$$

$$P(j \cdot x) = \begin{cases} x & 1 \leq j \leq 4 \\ 1 & j > 4 \end{cases} \quad (16)$$

$$T_{\text{inter}}^0 = T_{\text{best}} \quad (17)$$

式中, $\text{val}(\lambda)$ 表示 λ 所对应的隐状态序列模式的编号 $PN=0, 1, \dots, 4, i \geq 1$ 表示处理的 EBT-时序的段数, $j \geq 1$ 表示连续乘 2 的次数。 T_{best} 为获取原始特征值序列的固定时长, 其值的选取依据文献^[17]的研究成果, 文献^[17]对各种与驾驶疲劳相关的特征参数统计的时间间隔进行了研究, 结果表明, 取 60 s 连续视频进行特征参数统计, 疲劳识别效果最好。

由式(14)(15)可知,随着检测到的隐状态序列模式的编号越来越大,选取的时刻间隔越来越小;当相邻两段时序模式相同时,间隔变为当前间隔的2倍,否则保持间隔不变,当间隔增加到前一个间隔的16倍后停止增加。

3.2 α -时序与 EBT-时序的同步

由于眨眼特征相对于哈欠特征能够更精确地反映精神疲劳及其动态发展过程,以 EBT-时序的时间戳进行二者的同步:每次以时间戳为起始时刻点,同步获取时长为 T_{best} 的 EBT-时序和 α -时序。

3.3 两种时序判别结果的融合

采用模糊神经网络对两种时序的判别结果进行融合,采用文献[18]中的网络结构。各层节点个数为:第1层2个节点,分别接收 EBT-时序和 α -时序的标记结果作为输入;第2层8(7+3)个节点,分别代表 EBT-时序的7种模式和 α -时序的3种模式;第3层21(7×3)个节点,分别代表21条规则;第4层21个节点代表规则后件的归一化结果;第5层1个节点为输出层,输出当前序列的模式编号,取值为{0,1,...,4}中的某个值。采用带动量项的BP算法学习网络的权系数。训练好的网络用于进行两种时序判别结果的融合。

4 实验结果与分析

4.1 实验数据准备

由于目前国内外还没有公用的精神疲劳视频数据库,因此本文实验在自建的视频数据库上进行。为了能采集到各种不同程度的精神疲劳视频,在咨询有关医生的意见后,采取每天固定时间段采集的方式,分为3个时间段:10:00—10:45、14:00—14:45、23:00—23:45。参与采集视频的志愿者主要为长期从事脑力活动者,包括教师、学生、工程师,年龄跨度从20岁到65岁,每人采集15段视频,样本的年龄分布情况如表1所示。

表1 样本的年龄分布

Table 1 Age distribution of volunteers

	年龄/岁				
	20~29	30~39	40~49	50~59	60~69
人数	20	15	7	3	2

以 $T_{best}/2$ 时间间隔为单位进行样本的标注,采用被测者主观打分、专家打分、PVT(Psychomotor Vigilance Task)测试值^[19]三者结合的方式进行标注。将

主观打分、专家打分与 PVT 测试结果量化为3级,0:警醒、1:轻度疲劳、2:疲劳。取0、1、2中与三者平均值最接近的值作为当前时间间隔视频类别标注。

4.2 哈欠检测对比实验

在705段视频中,有228段视频含有深浅不一的哈欠,对应于29人,为了测试本文算法对哈欠检测的准确性,录制哈欠者的讲话视频、突然张大嘴的视频各87段,用于进行实验。对402段视频的哈欠检测结果见表2。

表2 哈欠检测率对比

Table 2 Yawning detection rate comparison of existing methods and our method

方法	哈欠判为哈欠	非哈欠判为哈欠	哈欠漏检 /%
文献[7]	80	5	20
文献[8]	95	5	5
α -时序分类	100	0	0

4.3 EBT-时序标记与 AECS 结果比较

在进行 HMM 训练时,首先根据标注结果将视频分割成小段,分别得到 $PN=0,1,\dots,4$ 的5类样本视频,然后,将每一类样本均匀分为2组,在划分时,使每组样本尽可能覆盖所有年龄。每类的两组视频中一组作为训练样本,另一组作为测试样本。各种疲劳状态(0:警醒;2:轻度疲劳;4:比较疲劳;)的 AECS 区间根据测试组中的0、2、4类样本求均值确定。各年龄组的各种疲劳状态测试样本的识别正确率如表3所示。

由表3可看出:1)AECS 阈值法的识别结果随着样本年龄的增加而急剧降低,而 HMM 对不同年龄的人群识别率差别不大,后者较前者更为稳定;2)HMM 建模法对不同年龄人群的同一种疲劳状态的平均识别率高于 AECS 阈值法;3)HMM 建模法能够对两种疲劳状态的过渡阶段,对应于 PN 值为1.3进行识别,而 AECS 阈值法只能对某种疲劳状态进行识别。

4.4 实时监测结果与 PVT 测试结果相关性

从每个年龄组中挑选出约一半人,共22人,进行为期5天的实时精神疲劳监测,录制视频的时间仍然是4.1节所述的3个时间点。在录制的过程中,同时启动两个程序运行,PVT测试程序、视频录制与疲劳监测程序。以 T_{best} 为时长分别获取 EBT-时序与 α -时序模式的类别编号,将2个类别编号作为神经网络的输入,得到网络输出,将输出结果进行

均值平移;将测试者在 T_{best} 时间内鼠标点击特定图案的次数归一化后的值进行均值平移后作为该时间间隔的 PVT 测试值。计算每个样本疲劳监测程序的输出值与 PVT 测试值的皮尔逊相关系数 r , 取平均值得到 $r=0.871$, 证明本文模型所得结果与 PVT 测试结果具有很强的正相关性。

表 3 EBT-时序模型与 AECS 识别率对比
Table 3 Mental fatigue recognition rate comparison of EBT-time series model and AECS

方法	PN	年龄/岁					平均
		20 ~ 29	30 ~ 39	40 ~ 49	50 ~ 59	60 ~ 69	
EBT-时序分类法	0	98.1	98.0	98.3	97.8	97.7	98.0
	1	97.6	97.4	97.3	97.0	96.9	97.2
	2	97.1	97.0	97.2	97.0	96.9	97.0
	3	97.4	97.3	97.2	97.1	97.0	97.2
	4	98.2	98.5	98.4	98.2	98.2	98.3
AECS 阈值法	0	98.1	94.5	87.9	79.4	77.4	87.5
	2	97.7	93.8	79.5	79.5	76.9	85.5
	4	98.3	94.6	89.3	80.0	79.0	88.2

4.5 实验平台及处理时间

采用 Logitech C270 720p USB 摄像头进行视频采集,图像大小为 640×480 ,使用 i5 2.5 GHz CPU、4 G 内存 PC 机进行处理,软件编程环境为 Win 7 旗舰版,Visual Studio 2008 平台,OpenCV 2.1 库。搭建的实验平台如图 8 所示,平均处理速度为 23 帧/s,基本能够满足实时检测的要求。

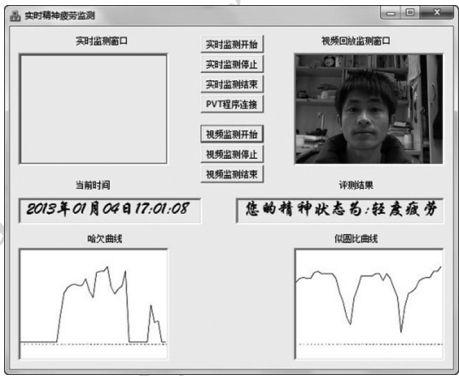


图 8 实验平台界面

Fig. 8 Experimental platform interface

5 结 论

提出了一种进行精神疲劳实时监测的多面部特征时序分类模型,主要贡献包括:1) 提出了随精神

疲劳状态动态变化的眼部特征的时序及其分类方法,由于 HMM 的参数可反映不同年龄人群的差异,克服了眼部特征阈值进行疲劳判定适应性较差的问题;由于 HMM 可对相邻两种疲劳状态的过渡序列进行识别,相较于只能对某种疲劳状态进行识别的方法更适用于精神疲劳的实时监测;2) 提出了能够反映精神疲劳状态的嘴部特征的时序及其分类方法,该方法不仅能够识别哈欠,还能够将深度哈欠与较浅哈欠和瞬时张大嘴巴等区别开来,提高了哈欠识别的准确率和精度;3) 提出在 HMM 基础上增加时间戳,用于进行两种特征时序的同步和自适应地求取下一时序的初始时刻,从而实现精神疲劳的实时监测。

不足之处在于进行疲劳测试的志愿者还不能覆盖所有的脑力工作者,采集更多工种的脑力工作者的面部视频进行实验是下一步工作的重点。

参考文献 (References)

[1] Ji Q, Zhu Z W, Lan P L. Real-time nonintrusive monitoring and prediction of driver fatigue[J]. IEEE Transactions on Vehicular Technology, 2004, 53(4): 1052-1068.
[2] Ji Q, Yang X. Real time visual cues extraction for monitoring driver vigilance[C]//Proceedings of Lecture Notes Presented at the Proc. Int. Workshop Computer Vision Systems. Vancouver, Canada: Springer, 2001: 107-124.
[3] Miyakawa T, Takano H, Nakamura K. Development of noncon-

- tact real-time blink detection system for doze alarm[C] //Proceedings of SICE Annual Conference. Sapporo, Japan: IEEE, 2004:1626-1631.
- [4] Horak K. Fatigue features based on eye tracking for driver inattention system[C]//Proceedings of the 34th International Conference on Telecommunications and Signal Processing. Budapest, Hungary: IEEE, 2011: 593-597.
- [5] Wang Q, Wang H, Zhao C X, et al. Driver fatigue monitoring based on eye state recognition[J]. Journal of Nanjing University of Science and Technology: Natural Science, 2010, 34(4):448-453. [王琼, 王欢, 赵春霞, 等. 基于眼睛状态识别的驾驶员疲劳监测[J]. 南京理工大学学报: 自然科学版, 2010, 34(4):448-453.]
- [6] Pan X D, Li J X, Xu X D. Threshold value of indices of eye states to monitor driver fatigue[J]. Journal of Tongji University: Natural Science, 2011, 39(12):1811-1815. [潘晓东, 李君羨, 徐小冬. 基于眼部行为的驾驶疲劳评价指标的阈值[J]. 同济大学学报: 自然科学版, 2011, 39(12):1811-1815.]
- [7] Abtahi S, Hariri B, Shirmohammadi S. Driver drowsiness monitoring based on yawning detection[C]//Proceedings of International Instrumentation and Measurement Technology Conference. Hangzhou, China: IEEE, 2011:1606-1610.
- [8] Wang T S, Shi P F. Yawning detection for determining driver drowsiness[C]//Proceedings of International Workshop on VLSI Design & Video Technology. Suzhou, China: IEEE, 2005:373-376.
- [9] Yu W W, Wu Q. Applying HMM-based driver fatigue recognition in smart vehicle space[J]. Computer Applications and Software, 2011, 28(10):43-46. [郁伟伟, 吴卿. 基于 HMM 的驾驶员疲劳识别在智能汽车空间的应用[J]. 计算机应用与软件, 2011, 28(10):43-46.]
- [10] Viola P, Jones M J. Rapid object detection using a boosted cascade of simple features[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE, 2001:511-518.
- [11] Van Kreveland M, Van Ostrum R, Bajaj C L. Contour trees and small seed sets for isosurface traversal[C]//The 13th ACM Symposium on Computational Geometry. Nice, France: ACM, 1997: 212-220.
- [12] Liew A W, Leung S H, Lau W H. Segmentation of color lip image by spatial fuzzy clustering[J]. IEEE Transactions on Fuzzy System, 2003, 11(4):542-549.
- [13] Rabi G, Si W L. Energy minimization for extracting mouth curves in facial image[C]//Proceedings of Intelligent Information Systems. Grand Bahama Island: IEEE, 1997: 381-385.
- [14] Yang S B. Research of dimensionality reduction method for time series[D]. Lanzhou: Lanzhou University, 2008. [杨仕博. 时序数据维归约方法的研究[D]. 兰州: 兰州大学, 2008.]
- [15] Luo Z Z, Yan T F. Classification of surface electromyography patterns by using HMM[J]. Journal of Huazhong Univ. of Sci. & Tech: Natural Science Edition, 2008, 36(4):72-75. [罗志增, 严庭芳. 基于 HMM 的表面肌电信号模式分类[J]. 华中科技大学学报: 自然科学版, 2008, 36(4):72-75.]
- [16] Duda R O, Hart P E, Stork D G. Pattern Classification[M]. Beijing: China Machine Press, 2003:105-114. [Duda R O, Hart P E, Stork D G. 模式分类[M]. 李宏东等译. 北京: 机械工业出版社, 2003: 105-114.]
- [17] Sandberg D. The performance of driver sleepiness indicators as a function of interval length[C]//Proceedings of the 14th International Conference on Intelligent Transportation Systems. Washington DC USA: IEEE, 2011:1735-1740.
- [18] Li Z C, He R, He C Q. Application of fuzzy neural network in driver fatigue detection[J]. Journal of Jiangsu University: Natural Science Edition, 2008, 29(2):123-126. [李志春, 何仁, 何翠群. 模糊神经网络在驾驶员疲劳检测中的应用[J]. 江苏大学学报: 自然科学版, 2008, 29(2):123-126.]
- [19] Graw P, Krauchi K, Knoblauch V, et al. Circadian and wake-dependent modulation of fastest and slowest reaction times during the psycho-motor vigilance task[J]. Physiology & Behavior, 2004, 80:695-701.