

基于人脸运动信息和改进保局投影的疲劳识别

张 伟 夏利民 罗大庸

(中南大学信息科学与工程学院 长沙 410075)

摘 要 提出了一种基于人脸运动信息和改进保局投影的疲劳识别方法。利用光流技术计算人脸皮层的运动速度,并以此作为疲劳特征;为了有效地进行疲劳特征降维,提出了改进的保局投影方法,该方法很好地保留了数据的局部流形结构和全局结构;采用加权k近邻的方法进行疲劳识别。实验结果表明该方法具有很好的识别效果。

关键词 疲劳识别,保局投影,加权k近邻

中图分类号 TP391.41 **文献标识码** A

Fatigue Recognition Based on Facial Motion and Improved Locality Preserving Projections

ZHANG Wei XIA Li-min LUO Da-yong

(College of Information Science and Engineering, Central South University, Changsha 410075, China)

Abstract In this paper, we presented a new method for fatigue recognition based on facial motion and improved locality preserving projections. Facial velocity information, which is determined using Optical flow techniques, was used to characterize fatigue. The improved locality preserving projections, which preserve the local and global structure of the data manifold, was proposed to extract the effective fatigue feature. Weighted K-nearest neighbor (WKNN) was instructed to recognize fatigue. A set of experiments of fatigue recognition were presented, and the experiment results show that the proposed method is effective and attains a satisfactory effect.

Keywords Fatigue recognition, Improved locality preserving projections, WKNN

研究表明,驾驶疲劳是引起交通事故的主要原因之一,因此驾驶人疲劳状态的自动识别是一个迫切需要解决的问题,也是新世纪交通安全性的重点研究领域,其对预防和控制交通事故、减少事故中的人员伤亡具有十分重要的理论意义和应用价值。

目前,有多种适应不同工作环境下的疲劳检测方法。根据检测原理的不同,这些方法可粗略地分为以下几类:1)基于生理学的方法:人体疲劳会引起脑电波、心率的变化,这些可以依靠 EEG、ECG^[1]等医疗设备来检测,但是这种方法会对人体正常活动造成影响,不能应用到真实的环境中去。2)基于行为的方法:通过监视驾驶员和车辆的行为来检测疲劳,如方向盘监视装置 S. A. M,方向盘正常运动时传感器装置不报警,若方向盘 4S 不运动,意味着驾驶员处于疲劳状态, S. A. M 就会发出报警声直到方向盘继续正常运动为止;或者利用一组传感器来测量驾驶员的头部位置,检测驾驶员的点头频率来检测疲劳。这两种方法均需要给人体附加检测设备,使用不方便。3)基于机器视觉的方法^[2-6],这是目前最流行的疲劳检测方法,它通过视频图像测量人眼的运动(眨眼频率和持续时间)、点头频率、嘴巴的张合等局部特征来判断人的疲劳状态,这是一种与人非接触的疲劳检测方法,便于使用,但是

由于不同的人的习惯不同,这些特征往往不能正确反映人的疲劳状态,并且人眼的一些特征如视线移动频率很难准确地获得。

根据生理学显示,人的疲劳特征往往反映在人的面部表情中,一方面人脸包含比眼睛、嘴巴更丰富的疲劳信息;另一方面,人脸定位与跟踪比人眼更容易,获得的特征信息更准确可靠,所以可以通过面部表情来进行疲劳识别。本文采用光流模型获取人脸皮层的运动速度,并以此作为疲劳特征;提出了改进的保局投影方法,该方法很好地保留了数据的局部流形结构和全局结构,能有效进行特征降维,提取有效的疲劳特征;最后采用加权k近邻方法进行疲劳识别。为了说明文中方法的有效性,我们利用文中方法与已有几种方法进行了对比实验。

1 表征疲劳的脸部运动信息

根据生理学,人的表情能反映人的疲劳状态,而表情是动态事件,需要用人脸的运动信息来表示,目前人脸运动信息已被用于性别识别、表情识别^[7]。本文利用脸部运动信息进行疲劳识别。为此,我们采用光流技术计算人脸各点的速度。

设 t 时刻图像中物体点 (x, y) 的灰度为 $I(x, y, t)$, 光流速

到稿日期:2009-12-10 返修日期:2010-02-11 本文受国家自然科学基金(69975003),国家博士点基金(20090162110057),湖南省自然科学基金项目(05JJ30121)资助。

张 伟(1972-),男,博士生,讲师,主要研究方向为交通信息与控制,E-mail: xlm@mail. csu. edu. cn;夏利民(1963-),男,博士,教授,博士生导师,主要研究方向为模式识别与图像处理;罗大庸(1944-),男,教授,博士生导师,主要研究方向为智能系统。

度为 $u=(u_x, u_y)$, 经过时间间隔 dt 后, 物体点由 (x, y) 点移到点 $(x+u_x dt, y+u_y dt)$, 根据物体点在图像中的灰度不变, 有:

$$I(x+u_x dt, y+u_y dt, t+dt) = I(x, y, t) \quad (1)$$

将上式左边进行泰勒级数展开, 得到:

$$I(x+u_x dt, y+u_y dt, t+dt) = I(x, y, t) + \frac{\partial I}{\partial x} u_x + \frac{\partial I}{\partial y} u_y + \frac{\partial I}{\partial t} dt + O(dt^2)$$

忽略上式中的 $O(dt^2)$ 项, 且将式(1)代入得到光流模型:

$$I_x u_x + I_y u_y + I_t = 0 \quad (2)$$

式中,

$$I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y}, I_t = \frac{\partial I}{\partial t}$$

I_x, I_y, I_t 可从图像序列直接求得。为了求解 $u=(u_x, u_y)$, 需附加约束条件: 光流场相邻点具有连续性, 即光流速度矢量 (u_x, u_y) 是局部平滑的。

$$\|\nabla u\|^2 = \left(\frac{\partial u_x}{\partial x}\right)^2 + \left(\frac{\partial u_x}{\partial y}\right)^2 + \left(\frac{\partial u_y}{\partial x}\right)^2 + \left(\frac{\partial u_y}{\partial y}\right)^2$$

通过极小化下列代价函数可求得 (u_x, u_y) :

$$C = \int_D [\lambda^2 \|\nabla u\|^2 + (\nabla I \cdot u + I_t)^2] dx dy$$

利用迭代运算^[8]求得 $u=(u_x, u_y)$:

$$\begin{aligned} u_x^{k+1} &= \bar{u}_x^k - \frac{I_x [I_x \bar{u}_x^k + I_y \bar{u}_y^k + I_t]}{\lambda + I_x^2 + I_y^2} \\ u_y^{k+1} &= \bar{u}_y^k - \frac{I_y [I_x \bar{u}_x^k + I_y \bar{u}_y^k + I_t]}{\lambda + I_x^2 + I_y^2} \end{aligned} \quad (3)$$

式中, k 表示迭代次数, 速度初始值 $u_x^0 = u_y^0 = 0$, $\bar{u}_x^k, \bar{u}_y^k$ 表示点 (x, y) 邻域的速度平均值。

对于含有 n 帧的视频图像, 可求得点 (x, y) 的平均速度 $\bar{u}=(\bar{u}_x, \bar{u}_y)$:

$$\bar{u}_x = \sum_{i=1}^n u_{ix}, \bar{u}_y = \sum_{i=1}^n u_{iy} \quad (4)$$

式中, $u_i=(u_{ix}, u_{iy})$ 表示第 i 帧图像中点 (x, y) 的速度。

对于人脸图像, 我们用上述方法可求得人脸区域各点的平均速度, 这些速度反映了人脸皮层的运动信息, 我们以此信息作为疲劳的原始特征。

2 基于改进保局投影的疲劳特征提取

根据上节方法我们可以得到疲劳的原始特征, 如果人脸区域图像大小为 $N \times M$, 则原始特征是维数为 $2(N \times M)$ 的高维特征, 需要降维处理(特征提取)。PCA 是常用的特征降维方法, 但利用 PCA 进行降维经常使得数据中的非线性结构无法保留, 而图像往往含有大量的非线性特性, 针对高维数据集的非线性特性, 目前提出了保局投影算法^[9](Locality Preserving Projections, LPP), 该算法具有较强的流形学习能力, 当数据存在非线性结构时, 利用该算法能很好地保留图像的本质特点, 而且具有线性方法计算简单、直观快捷的优点。但是 LPP 无法保证非邻近点经过投影后在低维空间也能够尽量散开, 为此我们提出了一种改进的保局投影, 该方法不仅保持样本的局部流形结构不变, 同时还能保持样本的全局流形结

构不变。这种改进的保局投影能有效地提取疲劳特征。

2.1 保局投影

LPP 算法本质上是一种线性降维方法, 根据最近邻图来建立映射, 设在高维欧式空间 R^D 中有数据集 $X=\{x_1, x_2, \dots, x_N\}$, 寻求一个投影矩阵 W , 将这些数据映射到一个相对低维的特征空间 $R^d (d \leq D)$ 中。数据集在 R^d 中的表述为 $Y=\{y_1, y_2, \dots, y_N\}$, 且 $Y=W^T X$ 。

LPP 算法的目的是在降维的同时, 保持样本固有的局部流形结构不变。LPP 的准则函数为:

$$\begin{aligned} J_1 &= \frac{1}{2} \sum_{ij} (y_i - y_j)^2 S_{ij}^n = \frac{1}{2} \sum_{ij} (W^T x_i - W^T x_j)^2 S_{ij}^n \\ &= \sum_i W^T x_i D_{ii}^n x_i^T W - \sum_{ij} W^T x_i S_{ij}^n x_j^T W \\ &= W^T (D^n - S^n) X^T W = W^T X L^n X^T W \end{aligned} \quad (5)$$

式中,

$$S_{ij}^n = \begin{cases} 1, & \text{if } x_i \in N_k(x_j) \text{ or } x_j \in N_k(x_i) \\ 0, & \text{otherwise} \end{cases}$$

式中, $N_k(x_j)$ 表示点 x_j 的 k -最近邻点的集合。

约束条件为:

$$W^T X D^n X^T W = 1$$

矩阵 D^n 为对角阵, $D_{ii}^n = \sum_j S_{ij}^n$; $L^n = D^n - S^n$ 为拉普拉斯矩阵。 D_{ij}^n 越大, 表示 y_i 越重要。

投影矩阵 W 可以通过求解约束条件(6)下最小化准则函数 J_1 获得。即求解下面的特征值问题:

$$X L^n X^T w = \lambda X D^n X^T w \quad (6)$$

很显然, 高维空间两点 x_i 和 x_j 距离很近, 那么经过保局投影, 它们的低维投影点 y_i 与 y_j 之间的距离也应该很近。为了有利于数据的分类, 我们同时还希望非邻近点经过投影后在低维空间也能够尽量散开, 但保局投影却无法保证。为此, 我们提出了改进的保局投影。

2.2 改进的保局投影

对于非邻近点定义:

$$S_{ij}^f = \begin{cases} 1, & \text{if } x_i \notin N_k(x_j) \text{ and } x_j \notin N_k(x_i) \\ 0, & \text{otherwise} \end{cases}$$

要使非邻近点的低维投影能够尽量散开, 下列目标函数应最大。

$$J_2 = \frac{1}{2} \sum_{ij} (y_i - y_j)^2 S_{ij}^f = W^T X L^f X^T W \quad (7)$$

$$D^f \text{ 为对角矩阵, } D_{ii}^f = \sum_j S_{ij}^f, L^f = D^f - S^f.$$

在此基础上, 我们定义改进保局投影的目标函数为:

$$J = \frac{J_2}{J_1} = \frac{W^T X L^f X^T W}{W^T X L^n X^T W} \quad (8)$$

使目标函数最大的 W 即为投影矩阵。如果 $X L^n X^T$ 是非奇异的, 求目标函数最大的问题转化为求下列广义特征值问题:

$$X L^f X^T W = \lambda X L^n X^T W \quad (9)$$

显然, 改进的保局投影不仅使得邻近点的低维投影点很近, 而且非邻近点的低维投影点离得远。这对后面的分类非常有效。

3 基于加权 k 近邻分类疲劳识别

k 近邻分类法(k-NN)是模式识别中一种非常有效的非参数化分类器,被广泛地用于模式识别和信号的检测,本文采用下列加权 k-NN 方法。

设有 M 个类别 $C_1, C_2, \dots, C_M$, 每类有 N_i 个样本, 若 $K_1, K_2, \dots, K_M$ 分别是未知样本 y 的 K 个最近邻(最相似)中属于 $C_1, C_2, \dots, C_M$ 类的样本数; 设 y_j^i 表示第 j 类第 i 个样本矢量, y 与 y_j^i 之间的相似性 $\text{sim}(y, y_j^i)$ 定义为:

$$\text{sim}(y, y_j^i) = \frac{y \cdot y_j^i}{\|y\| \|y_j^i\|} \tag{10}$$

判决函数定义为:

$$g_j(y) = \sum_i^{K_j} \text{sim}(y, y_j^i) \quad j=1, 2, \dots, M \tag{11}$$

若

$$g_m(y) = \max(g_j(y)) \tag{12}$$

则决策 $y \in C_j$ 。

在本文疲劳识别中有两类, 分别为: 正常类 C_1 , 疲劳类 C_2 。

4 实验与分析

我们创建了一个疲劳表情数据库。使用网络摄像头获取 40 个人的 人脸视频图像, 20 个男性, 20 个女性, 每个人的图像库包括 10 组正常情况下的视频图像、10 组疲劳情况下的视频图像, 总共包括了 800 组视频图像。用 400 组视频图像(正常、疲劳各 200)作训练样本, 其余 400 组视频图像作测试样本。

我们利用粒子滤波器进行人脸跟踪^[10], 提取视频图像中的人脸区域, 利用光流技术计算出人脸皮层各点的速度, 构成 $2 \times 64 \times 64$ 维的疲劳特征, 采用改进的 LPP 算法进行特征降维, 降维后的特征维数为 64, LPP 算法中紧邻个数取 3。k-NN 分类的参数 k 取 5。识别正确率为 93.3%。

为了说明文中方法的优越性, 我们与其他几种方法进行了对比实验。实验结果如表 1 所列。

表 1 4 种不同方法的比较结果

方法	疲劳识别率	(降维后)特征维数
方法一	83.5%	9
方法二	90.8%	64
方法三	92.0%	64
方法四	93.3%	64

其中, 方法一分别以眼睛、嘴巴和头部的运动信息作为疲劳特征, 包括: 眼睛闭合时间占特定时间的百分比 PERCLOS、眼皮的平均闭合速度 AECS、视线分布状态 GAZE-DIS、视线快速移动的发生频率 PERSAC、点头频率 NodFreq、哈欠的频率 YawnFreq。方法二、方法三、方法四(文中方法)是以人脸皮层运动速度为特征, 并分别用 PCA、LPP、改进的 LPP 进行降维。

实验结果表明, 方法一的识别率较低, 这是因为眼睛、嘴巴、头部运动信息不能充分反映人的疲劳状态, 并且由于眼睛、嘴巴很难准确定位, 导致提取的眼睛、嘴巴的运动信息不太准确, 尤其是视线分布状态。方法二、方法三和方法四识别率均高于 90%, 说明文中所采用的人脸皮层的运动速度能很好地反映人的疲劳状态, 同时方法四由于采用的是改进的 LPP, 在降维的同时, 很好地保留了数据的局部流形结构和全局结构, 使得识别率高于 PCA 和 LPP。

结束语 驾驶人疲劳识别对预防和控制交通事故, 减少事故中的人员伤亡具有十分重要的理论意义和应用价值。本文采用光流模型获取人脸皮层的运动速度, 并以此作为疲劳特征; 提出了改进的保局投影方法, 该方法很好地保留了数据的局部流形结构和全局结构, 能有效进行特征降维, 提取有效的疲劳特征; 最后采用加权 k 近邻的方法进行疲劳识别。实验结果表明文中方法具有好的识别效果。

参 考 文 献

[1] Redmond S, Heneghan C. Electrocardiogram - based automatic sleep staging in sleep disordered breathing[C]// IEEE International Computer in Cardiogram Conference Proceedings. 2003; 609-612

[2] Norimatsu Y, Mita S, Kozuka K, et al. Detection of the gaze direction using the time-varying image processing[C]// Proceedings of IEEE International Conference on Intelligent Transportation Systems. Vol. 1, 2003; 74-79

[3] Yang G, Lin Y, Bhattacharya P. A driver fatigue recognition model using fusion of multiple features[C]// Systems, Man and Cybernetics, 2005 IEEE International Conference. Volume 2, Oct. 2005; 1777-1784

[4] Eriksson M, Papanikotopoulos N P. Eye-tracking for detection of driver fatigue [C]// Intelligent Transportation System, 1997. ITSC 97, IEEE Conference. Nov. 1997; 314-319

[5] 李衡峰, 夏利民. 一种新颖的眼部状态识别方法[J]. 计算机工程, 2005, 31(6): 166-167

[6] Bao-Caiyinetc, et al. Multiscale dynamic features based driver fatigue detection[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2009, 23(3): 3575-589

[7] Anderson K, McOwan P W. A Real-Time Automated System for the Recognition of Human Facial Expressions[J]. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, 2006, 36(1): 96-105

[8] Horn B K P, Schunck B G. Determining optical flow[J]. Artificial Intelligence, 1981, 17: 185-204

[9] He Xiaofei, Niyogi P. Locality preserving projections[C]// Proc of the Conference on Neural Information Processing Systems. Canada: Vancouver, 2003; 153-160

[10] 夏利民, 张良春. 基于自适应粒子滤波器的物体跟踪[J]. 中国图像图形学报, 2009, 14(1): 112-117