

文章编号 1004-924X(2017)11-2947-11

## 人体下肢应激微反应自动识别

王昊鹏<sup>1,2</sup>, 冯显英<sup>2\*</sup>, 张明亮<sup>3</sup>

- (1. 山东警察学院, 山东 济南 250014;
2. 山东大学高效洁净机械制造教育部重点实验室, 山东 济南 250061;
3. 山东行政学院, 山东 济南 250014)

**摘要:** 由于现有的动作识别方法不能直接用于人体微反应动作识别, 本文基于人体下肢微反应动作特点, 构建了一种时空金字塔韦伯局部描述子并设计了基于字典学习的人体下肢微反应自动识别算法。该方法利用时空金字塔韦伯局部描述子提取每一类人体下肢微反应动作特征, 使用主成分分析法对特征降维; 然后, 建立每一类动作子字典并将子字典串联形成总的动作字典; 最后, 通过实验分析了金字塔级数  $L$ , 降维后每类动作特征维数  $d_{\text{PCA}}$ , 每类动作子字典原子个数  $n_{\text{Atom}}$ , 以及稀疏阈值  $C$  等参数对识别结果的影响, 并确定最优参数值  $L=3$ ,  $d_{\text{PCA}}=30$ ,  $n_{\text{Atom}}=40$ ,  $C=10$ 。实验结果表明, 提出的算法对 10 种人体下肢微反应动作的识别率均在 0.83~0.91 之间, 平均识别率达到 0.86, 高于其他动作识别算法。设计的算法更适用于人体下肢微反应动作分类, 并可有效提高分类识别率。

**关键词:** 人体下肢; 应激微反应; 动作识别; 自动识别; 时空金字塔韦伯局部描述子  
**中图分类号:** TP391.4   **文献标识码:** A   **doi:** 10.3788/OPE.20172511.2947

## Automatic recognition of micro-expressions action for human lower limb

WANG Hao-peng<sup>1,2</sup>, FENG Xian-ying<sup>2\*</sup>, ZHANG Ming-liang<sup>3</sup>

- (1. Shandong Police College, Jinan 250014, China;
  2. Key Laboratory of High Efficiency and Clean Mechanical Manufacture, Ministry of Education, Shandong University, Jinan 250061, China;
  3. Shandong Academy of Governance, Jinan 250014, China)
- \* Corresponding author, E-mail: fxying@sdu.edu.cn

**Abstract:** Because the existing motion recognition method couldnot be directly used in human micro-expression action recognition. A spatio-temporal pyramid Weber Local Descriptor (STPWLD) was constructed and an automatic recognition algorithm of human lower limb micro-expression action based on dictionary learning according to characterize human lower limb micro-expression action was designed. With the method, the features of human lower limb micro-expression action was extracted by the STPWLD. Then, the dimensions of STPWLD feature were reduced by the principal component analysis. Furthermore, the sub-dictionaries of human lower limb micro-expressions action was

收稿日期: 2017-02-08; 修订日期: 2017-05-12.

基金项目: 国家自然科学基金资助项目 (No. 51405194); 山东省自然科学基金资助项目 (No. ZR2016EEP13, No. ZR2013CQ025); 山东省高等学校科技计划资助项目 (No. J16LN16)

established and these sub-dictionaries were connected in series to construct a general action dictionary. Finally, the influence of the parameters of the algorithm on the recognition results was analyzed, and the optimal value of these parameters was determined. It shows that the optimal value of pyramid scales is 3, the optimal feature dimension of each action after dimension reduction is 30, the optimal number of atoms in each action dictionary is 40 and the optimal value of sparse threshold is 10. The experimental results indicate that the recognition rates of the proposed algorithm for 10 kinds of human lower limb micro-expression actions are all between 0.83~0.91, and the average recognition rate is 0.86, higher than that of other algorithms. The algorithm is suitable for the classification of human lower limb micro-expression actions and improves the classification recognition rate effectively.

**Key words:** human lower limb; micro-expression; action recognition; spatio-temporal pyramid weber local descriptor; dictionary learning

## 1 引 言

心理应激微反应(Micro-expressions)<sup>[1]</sup>是指在受到有效刺激的瞬间表现出的不受思维控制的真实反应,简称微反应。微反应是人类经过长期进化而遗传继承下来的一种生物本能反应,无法通过人主观地表演“装”出来,因此通过分析人的这些细微行为痕迹,可以逆推人的精神系统状态和情绪,从而得出人的真实心理。微反应识别可广泛应用于公安机关破案、商业谈判和心理咨询等领域<sup>[2]</sup>。

由于微反应具有发生频率低、持续时间短和动作幅度小等特点,通过人工来识别微反应不仅识别难度大,误识别、漏识别也较多,而且需要对识别人进行较长时间地识别训练,因此利用图像处理技术进行微反应的自动识别来有效提高微反应识别的准确性,已成为当前研究的热点。人的微反应主要体现在面部表情、四肢动作和躯干姿态等 3 方面<sup>[3]</sup>。其中对面部微表情识别的研究<sup>[4]</sup>开展得最广泛,成果较多且识别精度较高,对四肢动作和躯干姿态微反应识别的研究相对较少,这是因为人体是非刚性物体,其四肢和躯干具有复杂多变的动作和姿势,尤其是四肢和躯干的微反应动作非常隐蔽且不同类型的微反应本身具有较高的相似度,再加上背景、光照、着装、遮挡等因素的干扰,使得对四肢动作和躯干姿态的识别成为目前计算机视觉领域中最难且最具有挑战性的问题之一<sup>[5]</sup>。

Doan<sup>[6]</sup>等利用 Kinect 传感器获取手部 RGB-

D 图像,构造三层金字塔特征描述符,并使用支持向量机进行手部动作的分类;Marcos<sup>[7]</sup>等对面试中的人体进行视频监控,用光流法描述被监控人的躯干和上肢运动,并进行动作识别,从而实现对被监控人的心理状态的推断;Yang<sup>[8]</sup>等利用人体步行时重心的变化规律对下肢关节点进行特征描述及精确定位,实现了对人体步态的识别;Kwak<sup>[9]</sup>等设计了一种利用累计运动信息进行人体动作识别的方法,首先逐帧提取人体目标轮廓,然后计算相邻帧间人体轮廓的差分图像,再将差分图像进行累积,最后用这种累积信息来描述人体的动作特征。这种方法能有效识别小幅度抬手和提腿等动作;蔡加欣<sup>[10]</sup>等利用 Procrustes 形状分析表达人体轮廓特征,建立人体姿态字典,实现了对人体行为的识别。这些动作识别方法虽然都取得了较高的识别率,也进行了成功地应用。但人体四肢和躯干的微反应动作比这些识别方法所能识别的动作幅度更小且持续时间更短,故现有的动作识别方法不能直接用于人体微反应动作识别。

针对这些问题,本文设计了一种针对坐姿状态下人体下肢微反应动作的识别方法,在分析坐姿状态下人体下肢微反应动作特点的基础上,使用一种时空韦伯局部描述子提取下肢微反应动作特征,构造动作特征字典,求解测试视频在动作特征字典下的稀疏表示,利用最大聚集(Max-pooling)计算测试视频对应每一类动作的重构误差,将测试视频归类到具有最小重构误差的类,从而实现微反应动作的识别。

## 2 人体下肢微反应动作描述

### 2.1 人体下肢微反应动作分析

本文所讨论的人体下肢微反应动作主要是指双腿和双脚的动作,有些手脚共同作出的微反应

动作如双手抱膝等不在本文讨论范围内。坐姿状态下,人体下肢常见微反应动作可归纳为 7 种,分别为晃脚、转向脚、翘脚、踮脚、锁脚、双腿交叉和双腿叉开,其动作特征和情绪含义如图 1 和表 1 所示。

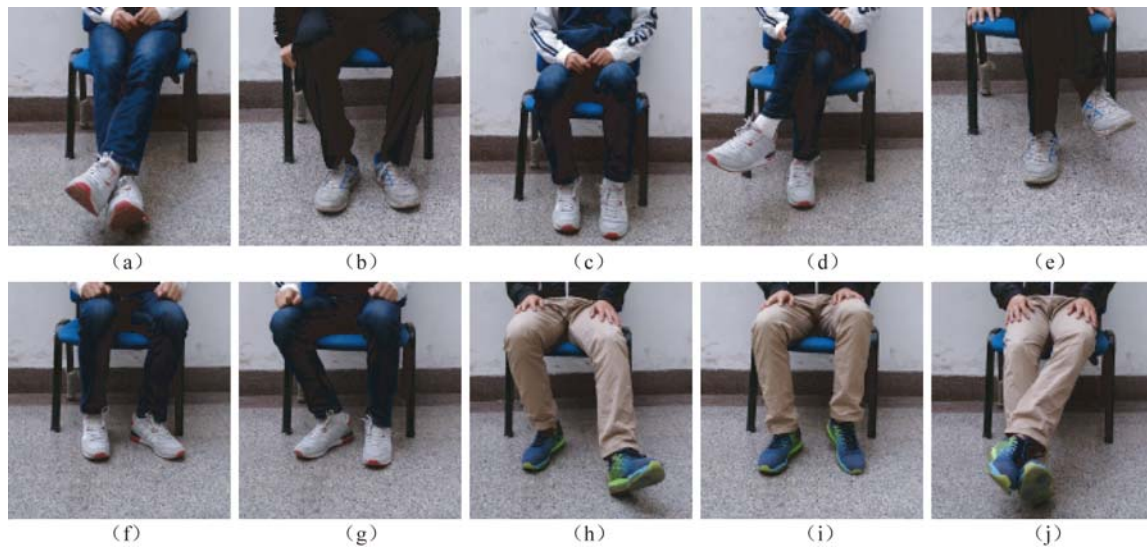


图 1 各种下肢微反应动作

Fig. 1 All kinds of human lower limb micro-expression action

表 1 人体下肢微反应动作特征

Tab. 1 Features of human lower limb micro-expression action

序号	动作名称	动作特征	情绪含义
1	双腿交叉	双腿伸直交叉在一起	舒适、自信、放松
2	双腿叉开	腿和脚叉开	烦乱、对峙、紧张
3	晃脚 1	单脚或双脚以脚跟或脚掌支撑上下颤动	积极、兴奋、快乐
4	晃脚 2	跷二郎腿摇晃脚	愉快、喜悦、积极
5	晃脚 3	脚从左右摇晃变为踢动	紧张、反感、消极
6	转向脚 1	转向某一人或事物	逃离、消极、反对
7	转向脚 2	转离某一人或事物	亲密、积极、赞同
8	翘脚	脚跟着地,脚趾指向天空	积极、兴奋、快乐
9	踮脚	脚跟翘起,脚掌着地	逃离、消极、反对
10	锁脚	双脚交叉在一起,脚趾转向内侧	焦虑、恐惧、消极

通过图 1 和表 1 的动作特征描述可知,坐姿状态下人体下肢的微反应动作复杂多变。有些微反应动作是一种姿态的保持,例如双腿交叉,动作过程是双腿双脚伸直上下叠着交叉起来,动作做完后,下肢会保持这种姿态一段时间;有些微反应

动作是一种周期性重复动作,例如单脚晃脚动作,以脚跟(或脚掌)支撑,脚掌(或脚跟)一下下点地,每次点地加抬起是一次动作过程,只点地一次不算单脚晃脚微反应,只有周期性地重复点多下才算。

## 2.2 基于时空金字塔韦伯局部描述子的人体下肢微反应动作特征提取

韦伯定律源于心理学研究,认为大脑受到刺激时所感知的刺激强度与原刺激值的比值是一个常数。受韦伯定律的启发,陈杰<sup>[11]</sup>等提出了韦伯局部描述子(Weber Local Descriptor, WLD),将其用于描述图像局部纹理特征并取得了良好的效果。

WLD除了被用于图像纹理识别外,还被广泛应用于人脸识别、表情识别、性别识别和种族识别等方面<sup>[12-13]</sup>。在这些应用中都是将WLD用于单幅图像,但微反应动作由于其动作幅度小,很难通过关键帧的姿态进行判定识别。

2007年,Zhao<sup>[14]</sup>等为了克服局部二值模式(Local Binary Patterns, LBP)在序列图像处理上的不足,提出了LBP-TOP(Local Binary Patterns from Three Orthogonal Planes)方法,将LBP扩展到时空3个正交平面,获取到序列图像在时空域的运动特征。张轩阁<sup>[15]</sup>等将LBP-TOP用于微表情识别,取得了较高的识别率。与LBP算子一样,WLD算子也是用于描述图像局部细节,能够捕捉细微的显著特征,对图像较大区域的宏观特征描述能力较差,不同的是WLD算子对图像局部强度变化的描述能力要优于LBP算子。

受此启发本文提出一种时空金字塔韦伯局部描述子(Spatio-temporal Pyramid Weber Local Descriptor, STPWLD),将视频图像 $I(x, y)$ 序列视为时空域的3个正交平面XY、XZ和YZ,如图2所示。

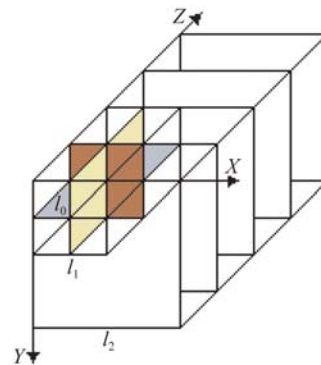


图2 图像序列三正交平面时空金字塔

Fig. 2 Three orthogonal planes space-time Pyramid of image sequence

在每一个平面内,利用金字塔将平面划分为 $L$ 级大小不同的区域,计算第 $l_i$  ( $i=0, 1, \dots, L-1$ )级每个区域的WLD特征,对直方图进行归一化(设归一化系数为 $\alpha=1/4^i$ )处理后,将每一级每一个区域的直方图串联在一起得到该平面的PWLD直方图,将3个平面的PWLD直方图串联在一起得到最终的STPWLD特征,如图3所示(以第8类动作为例),图中为第8类动作(翘脚)视频样本中的部分视频帧及其STPWLD特征直方图,金字塔级数 $L=3$ 。特征维数为:

$$d_{\text{STPWLD}} = 3 \sum_{i=0}^{L-1} 4^i M_i T_i S_i, \quad (1)$$

其中: $T_i$ 为WLD算子主要梯度方向的个数, $M_i$ 为WLD直方图的段数, $S_i$ 为WLD直方图每段的子区间个数。

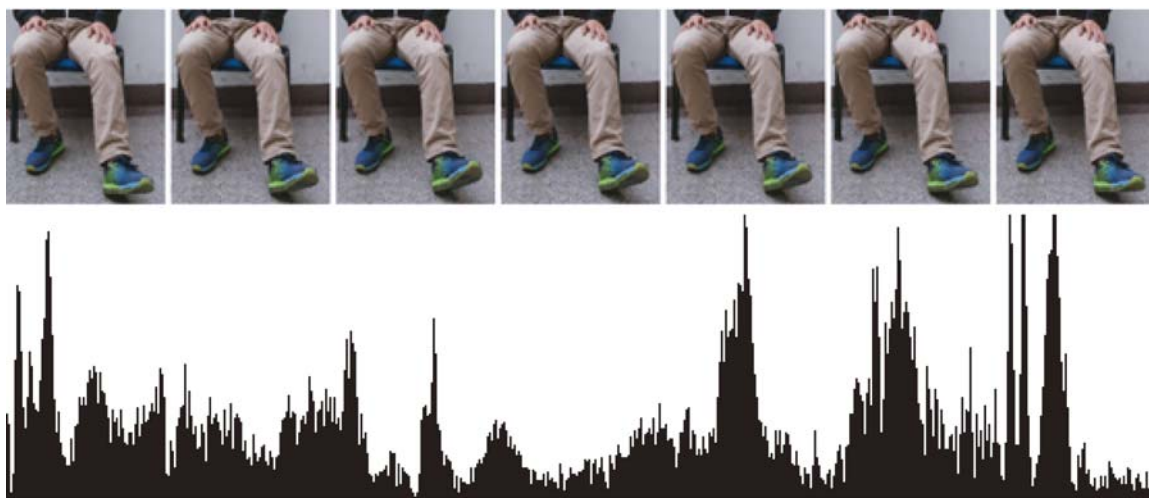


图3 第8类动作及其STPWLD特征直方图

Fig. 3 The eighth type action and STPWLD feature histogram

STPWLD 特征的计算复杂度为:

$$O_{\text{STPWLD}} = C_0 \sum_{i=0}^{L-1} (m_i n_i + m_i t_i + n_i t_i), \quad (2)$$

其中:  $L$  为金字塔级数,  $m_i$ 、 $n_i$  和  $t_i$  ( $i=0, 1, \dots, L-1$ ) 分别为每一级金字塔  $X$ 、 $Y$ 、和  $Z$  边中需计算 WLD 特征的区域个数,  $C_0$  是一个常量, 表示计算每个 WLD 特征所进行的加、除以及求反正切等运算的个数。

WLD 特征, 由差分激励  $\xi$  和方向  $\theta$  两部分构成, 计算公式为:

$$\xi(x_c) = \arctan \left[ \frac{v_{s00}^{00}}{v_{s01}^{01}} \right] = \arctan \left[ \sum_{i=0}^{n-1} \left[ \frac{x_i - x_c}{x_c} \right] \right], \quad (3)$$

其中:  $v_{s00}$  和  $v_{s01}$  分别为滤波窗口  $f_{00}$  和  $f_{01}$  与像素点  $x_c$  及其周围  $3 \times 3$  邻域的卷积,  $x_i$  表示中心像素点  $x_c$  的第  $i$  个邻域像素,  $n$  表示邻域像素个数, 取值为 8。像素点  $x_c$  的方向  $\theta(x_c)$  是指  $x_c$  的梯度方向, 为:

$$\theta(x_c) = \arctan \left[ \frac{v_{s11}^{11}}{v_{s10}^{10}} \right], \quad (4)$$

其中:  $v_{s11}$  和  $v_{s10}$  分别为滤波窗口  $f_{11}$  和  $f_{10}$  与像素点  $x_c$  及其周围  $3 \times 3$  邻域的卷积。

### 3 基于字典学习的人体下肢微反应动作识别

#### 3.1 字典学习与稀疏表示

字典学习的任务是对于给定的某一类测试数据, 设计一个字典, 使得该类数据能在该字典上被稀疏表示。所谓稀疏表示是一种通过用少量原子线性组合来表示一个信号的方法。对于给定的测

试数据  $Y$  可以利用一个稀疏的系数矩阵  $A$  和一个字典矩阵  $D$  表示, 并要求  $A$  尽可能稀疏, 则称  $A$  是  $Y$  的稀疏表示。K-SVD 算法是最常用的字典学习算法之一<sup>[16]</sup>, 其目标方程描述如下:

$$\begin{cases} \operatorname{argmin}_{D, A} \| Y - DA \|_F^2 \\ \text{s. t. } \| a_i \|_0 \leq C, \forall i \end{cases} \quad (5)$$

从误差逼近角度, 式(5)可改写为:

$$\begin{cases} \operatorname{argmin} \| a_i \|_0 \\ \text{s. t. } \| y'_i - Da_i \|_F^2 \leq C \end{cases}, \quad (6)$$

其中:  $Y \in \mathbf{R}^{m \times l}$  为训练样本矩阵, 其每一列是一个训练样本;  $D \in \mathbf{R}^{m \times n}$  为要学习的字典;  $A \in \mathbf{R}^{n \times l}$  为稀疏系数矩阵;  $a_i$  为  $A$  的第  $i$  列;  $\| \cdot \|_F$  表示矩阵的 Frobenius 范数; 约束项中的  $\| \cdot \|_0$  表示系数矩阵中非零元素个数, 代表了稀疏程度;  $C$  为稀疏阈值。

K-SVD 算法采用迭代的方式求解式(5)中的优化问题, 每次迭代时首先固定字典求解稀疏系数, 然后固定稀疏系数求解字典。其中求解稀疏系数最常用的算法是正交匹配追踪 (Orthogonal Matching Pursuit, OMP)<sup>[17]</sup> 算法。

#### 3.2 基于字典学习的人体下肢微反应动作识别算法

假设人体下肢微反应动作共有  $K$  类, 对于第  $k$  类动作, 从其所有测试视频中提取时空金字塔韦伯局部特征  $Y$ , 利用 K-SVD 算法学习  $Y$  的字典  $D_k$ , 对每一类动作进行字典学习, 则整个人体下肢微反应动作字典  $D$  可通过串联每一类动作字典的方式获得:

$$D = [D_1 | D_2 | \dots | D_k]. \quad (7)$$

10 类动作总的字典如图 4 所示。

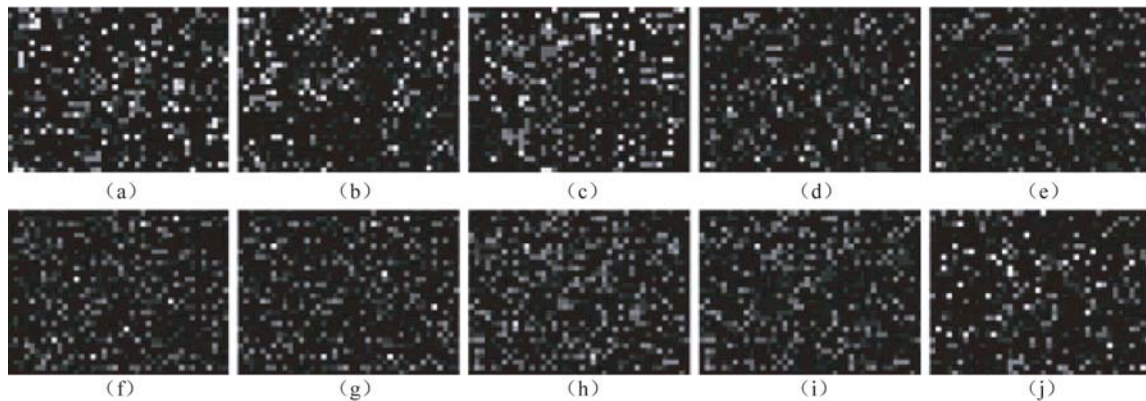


图 4 全部 10 类动作的动作字典

Fig. 4 Action dictionary of all ten types of action

构建好人体下肢微反应动作字典  $\mathbf{D}$  后,对给定测试视频  $\mathbf{Y}' = [y'_i]_{i=1}^T$  ( $T$  为视频帧数) 中的任意一帧  $y'_i$  利用式(6)求解其在字典  $\mathbf{D}$  下的稀疏表示。得到  $y'_i$  的稀疏表示后,其对应的第  $k$  类重构误差  $r_k(y'_i)$  可表示为:

$$r_k(y'_i) = \|y'_i - \mathbf{D}\delta_k(\mathbf{a}_i^*)\|_2, \quad (8)$$

其中:  $\delta_k(\mathbf{a}_i^*)$  用于将稀疏矩阵的向量  $\mathbf{a}_i^*$  中所有不与第  $k$  类对应的行置 0。

利用重构误差进行动作分类,首先对全部  $K$  类动作计算测试视频所有帧对每一类动作的重构误差的和,将其标记为测试视频对该类动作的重构误差;然后在全部  $K$  类动作中寻找重构误差最小的类,将测试视频标记为该动作,即:

$$k' = \arg \min_k \sum_{i=1}^T r_k(y'_i). \quad (9)$$

整个识别过程总结如下:提取视频的 STPWLD 特征;利用式(1)建立每一类动作的 STPWLD 特征子字典;利用式(6)将子字典串联成总的动作字典;利用式(5)计算视频在字典下的稀疏表示;利用式(7)计算视频对每类动作的重构误差;利用式(8)对视频进行分类。

## 4 实验结果与分析

### 4.1 实验样本与实验设计

目前在心理应激微反应自动识别研究中较常用的数据库主要有 METT 数据库、SMIC 数据库和 CASME 数据库,这些数据库均为各研究团队自行设计开发,且均为微表情数据库。由于尚未发现公开的、较为权威的人体肢体微反应数据库。因此,本文实验所用视频样本均为自行拍摄采集,共有 10 类微反应动作,动作名称、特征及编号如表 1 所示。选择 10 名被测试人,每名被测试人分别做表 1 中的 10 种微反应动作,每种动作做 10 遍,

整个视频样本集共 1 000 段视频,采用德国 Basler acA1280-60gc 摄像机(视频分辨率为  $1\ 280 \times 1\ 024$  pixel,帧速率为 60 f/s,时长为 3 s)拍摄。

本文算法中,对识别精度产生影响的参数主要有生成 STPWLD 特征时的金字塔尺度、特征的特征维数、字典学习时字典原子的个数和稀疏阈值的大小。本文将对这 4 个参数进行实验比较,并采用留一法交叉验证对样本进行识别并统计识别率 (Recognition Rate, RR),识别率定义如下:

$$RR_k = \frac{\text{第 } k \text{ 类动作分类正确的数量}}{\text{第 } k \text{ 类动作测试视频总数}}, \quad (10)$$

$$RR_{\text{avg}} = \frac{\sum_{k=1}^K RR_k}{K}, \quad (11)$$

其中:  $RR_k$  为第  $k$  类动作 ( $k=1, 2, \dots, K$ ) 的识别率,  $RR_{\text{avg}}$  为全部  $K$  类动作的平均识别率。

### 4.2 金字塔级数对识别率的影响

在计算 STPWLD 特征阶段,将视频时空域 3 个平面每个平面均分为  $L$  级金字塔,每级金字塔的 ( $M_i, T_i, S_i$ ) 均设为 (6, 8, 10), 特征维数可由式(8)计算得到;在字典学习阶段,设置动作字典  $\mathbf{D}$  中每个子字典的原子个数为 40, 稀疏阈值  $C$  为 10。金字塔级数  $L$  与 10 类动作的平均识别率关系如图 5 所示。

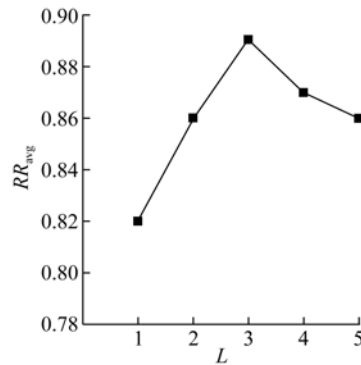


图 5 金字塔级数  $L$  与平均识别率  $RR_{\text{avg}}$  的关系

Fig. 5 Relationship between scale of pyramid and average recognition rate

从图 5 中可以看出,当  $L=1$  (即单尺度) 时  $RR_{\text{avg}}$  仅为 0.82,这是因为单尺度所提取的特征是整个视频的时空韦伯特征,无法很精确地表达视频的局部时空韦伯特征;随着  $L$  值的增大  $RR_{\text{avg}}$  有了明显提高,并在  $L=3$  时达到了最高值 0.89,这说明三级金字塔尺度最能合理表达视频的局部时空韦伯特征;当  $L$  继续增大,  $RR_{\text{avg}}$  反而降低,这是因为尺度划分过小,无法正确表达视频的局部时空结构。因此本文取  $L=3$ 。

### 4.3 特征维数对识别率的影响

根据式(1)可知随着金字塔级数  $L$  的增加,特征维数指数级增长,大大增加了计算的复杂性,因此需要对特征进行降维。本文采用主成分分析法<sup>[18]</sup> (Principal Component Analysis, PCA) 对特征进行降维。首先设置金字塔级数  $L=3$ , 提取出 STPWLD 特征,然后使用 PCA 对特征降维,最后

进行字典学习, 算法的相关参数设置与 4.2 节相同, 降维后特征维数与 10 类动作的平均识别率关系如图 6 所示。

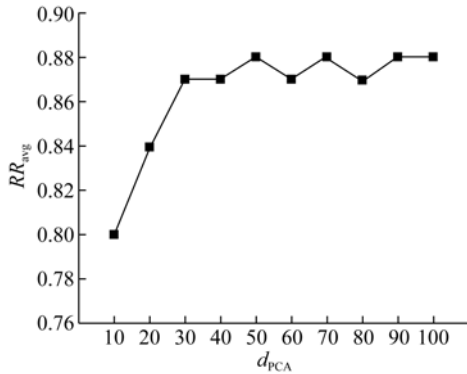


图 6 特征维数  $d_{PCA}$  与平均识别率  $RR_{avg}$  的关系

Fig. 6 Relationship between feature dimension and average recognition rate

从图 6 中可以看出, 在  $[10, 30]$  区间内, 随着  $d_{PCA}$  的增加  $RR_{avg}$  显著提高, 这是因为降维后特征过少, 无法准确表达动作的全部特征; 当  $d_{PCA}$  超过 30 后,  $RR_{avg}$  稳定在  $0.87 \sim 0.88$  之间, 这说明使用 30 维以上的特征就可以较为准确地表达动作的全部特征, 因此本文取  $d_{PCA} = 30$ 。

#### 4.4 字典原子个数对识别率的影响

设置金字塔级数  $L = 3$ ,  $d_{PCA} = 30$ , 算法的其他相关参数设置与 4.2 节相同, 每类动作字典原子个数与 10 类动作的平均识别率关系如图 7 所示。

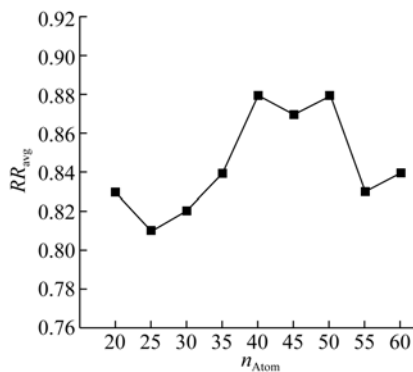


图 7 字典原子个数  $n_{Atom}$  与平均识别率  $RR_{avg}$  的关系

Fig. 7 Relationship between dictionary atomic number and average recognition rate

从图 7 中可以看出, 在  $n_{Atom}$  在  $[20, 60]$  区间内  $RR_{avg}$  有一定波动, 其中  $n_{Atom} = 40$  和  $n_{Atom} = 50$  时  $RR_{avg}$  最高, 达到 0.88, 本文取  $n_{Atom} = 40$ 。

#### 4.5 稀疏阈值对识别率的影响

设置金字塔级数  $L = 3$ ,  $d_{PCA} = 30$ ,  $n_{Atom} = 40$ , 算法的其他相关参数设置与 4.2 节相同, 稀疏阈值与 10 类动作的平均识别率关系如图 8 所示。

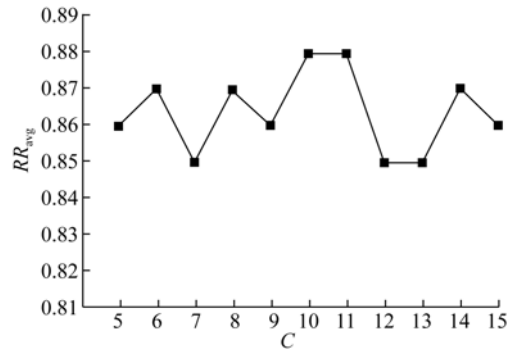


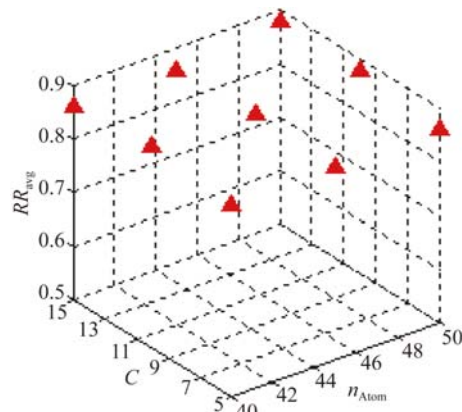
图 8 稀疏阈值  $C$  与平均识别率  $RR_{avg}$  的关系

Fig. 8 Relationship between sparse threshold and average recognition rate

从图 8 中可以看出, 在  $C$  在  $[5, 15]$  区间内  $RR_{avg}$  虽然有一定波动, 但波动幅度很小, 最低 0.85, 最高 0.88, 这说明在  $[5, 15]$  内识别率对参数  $C$  不敏感,  $C$  可任意取值, 因此本文取  $C = 10$ 。

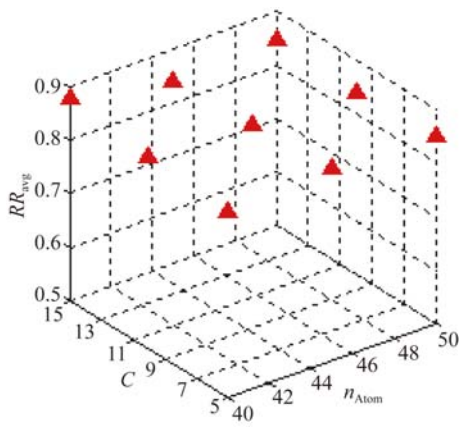
#### 4.6 $d_{PCA}$ 、 $n_{Atom}$ 和 $C$ 交互作用对识别率的影响

各参数中,  $L$  反映了视频的时空结构, 不适合过大过小, 在  $L = 3$  的前提下,  $d_{PCA}$ 、字典  $n_{Atom}$  和  $C$  在取值范围内具有一定的相互影响, 因此需分析各参数间交互作用对识别结果的影响。通过图 6~8 可以看出,  $d_{PCA}$  在  $[30, 100]$  范围内、 $n_{Atom}$  在  $[40, 50]$  范围内、 $C$  在  $[5, 15]$  范围内时, 平均识别率  $RR_{avg}$  较稳定, 因此设置  $d_{PCA}$  分别为 30、65 和 100,  $n_{Atom}$  和  $C$  对识别结果的影响如图 9 所示。



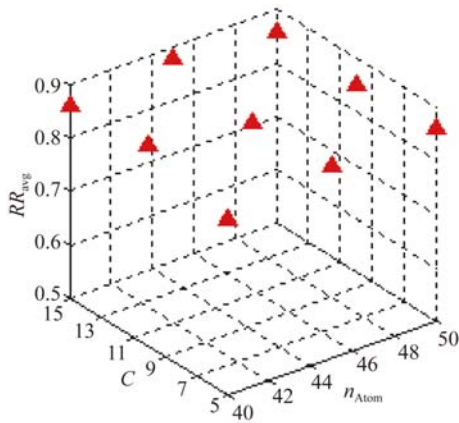
(a)  $d_{PCA} = 30$  时  $n_{Atom}$  和  $C$  对识别率的影响

(a) Impact of  $n_{Atom}$  and  $C$  on recognition rate at  $d_{PCA} = 30$



(b)  $d_{PCA} = 65$  时  $n_{Atom}$  和  $C$  对识别率的影响

(b) Impact of  $n_{Atom}$  and  $C$  on recognition rate at  $d_{PCA} = 65$



(c)  $d_{PCA} = 100$  时  $n_{Atom}$  和  $C$  对识别率的影响

(c) Impact of  $n_{Atom}$  and  $C$  on recognition rate at  $d_{PCA} = 100$

图 9  $d_{PCA}$ 、 $n_{Atom}$  和  $C$  对识别结果的影响

Fig. 9 Impact of  $d_{PCA}$ ,  $n_{Atom}$  and  $C$  on recognition rate

#### 4.7 不同被测试人对识别率的影响

考虑到不同被测试人做微反应动作的差异,因此需分析被测试人的差异对识别率的影响,设置参数  $L = 3$ ,  $M_i = 6$ ,  $T_i = 8$ ,  $S_i = 10$ ,  $d_{PCA} = 30$ ,  $n_{Atom} = 40$ ,  $C = 10$ 。不同被测试人做不同微反应动作的识别率结果如表 2 所示。

表 2 的每一个单元格为 1 个被测试人做 1 种微反应动作 10 遍的平均识别率,每一行合计为同一被测试人分别做 10 种微反应动作的平均识别率。从表 2 中可以看出,每一个被测试人做每一种微反应动作 10 遍的平均识别率几乎均在 0.8 以上,只有极个别的为 0.7,同一被测试人分别做 10 种微反应动作的平均识别率在 0.84~0.89 之间,这说明本文设计算法对不同被测试人微反应动作的识别率不敏感,一致性较好,可用于对不同被测试人微反应动作的检测。

表 2 不同被测试人对识别率的影响

Tab. 2 Effect of different subjects on recognition rate

被测试人编号	动作编号										合计
	1	2	3	4	5	6	7	8	9	10	
1	0.9	0.8	0.9	1	0.9	0.9	0.8	0.8	0.8	0.9	0.87
2	0.8	0.9	0.8	0.9	0.8	1	0.9	0.8	0.7	0.8	0.84
3	0.9	0.9	0.9	0.8	0.9	0.9	0.9	0.9	0.8	0.8	0.87
4	0.9	1	0.8	0.8	0.7	0.8	1	0.9	0.9	0.9	0.87
5	0.8	0.8	0.8	0.9	0.8	0.9	0.9	0.8	0.8	0.9	0.84
6	0.7	0.9	0.9	0.8	0.9	0.8	1	0.9	0.9	0.8	0.86
7	0.7	1	1	0.8	0.9	0.8	0.8	0.8	0.9	0.9	0.86
8	0.9	0.9	0.8	0.9	0.8	0.8	0.9	0.8	0.9	0.8	0.85
9	0.9	1	0.9	0.9	0.9	0.9	0.8	0.9	0.8	0.9	0.89
10	0.8	0.9	0.8	0.9	0.9	0.9	0.9	0.8	0.8	0.8	0.85

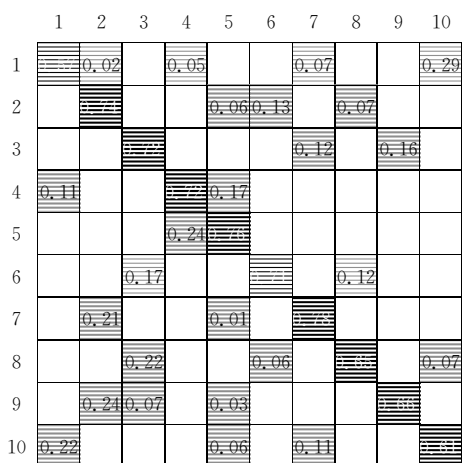
#### 4.8 本文提出算法与其他算法的识别对比

为进一步验证本文提出算法的有效性,分别使用文献[10]、[15]设计的算法和本文提出的算法对 3.1 节采集的视频样本集进行分类识别。文献[10]提出了一种基于姿势字典学习的算法,通过人体的姿态特征实现动作识别,文献[15]提出了一种光流与 LBP-TOP 特征结合的算法,用于识别人体面部微表情,这两种算法分别在 Weizmann、MuHAVi-MAS14 和 CASMEII 等数据库中取得了不错的识别率<sup>[10,15]</sup>(由于文献[10]使用的数据库 Weizmann、MuHAVi-MAS14 和文献[15]使用的数据库 CASMEII 与本文自建数据库的动作特点不同,对文献[10]和[15]中算法的参数进行了优化,文献[10]算法中的参数优化为轮廓点序列长度  $L = 100$ ,稀疏阈值  $C = 5$ ,字典原子个数  $n = 18$ ,高斯函数参数  $\epsilon = 5$ ,LPP 维数 20;文献[15]算法中的参数优化为 LBP 半径  $R_x = R_y = 5$ , $R_T = 1$ ,邻域点数  $P = 8$ )。3 种算法的识别结果如表 3 和图 10 所示。 $T_{avg}$  为算法完整运算时间,包括特征计算及学习的时间。

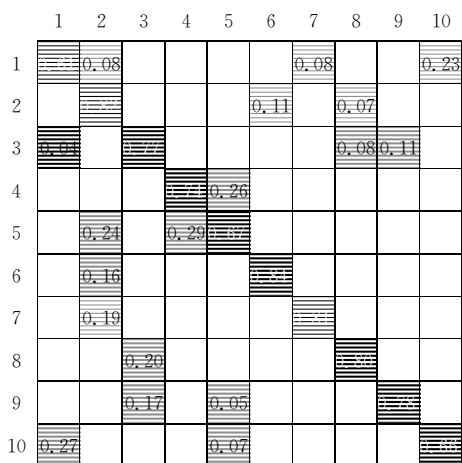
表 3 各算法识别结果比较

Tab. 3 Comparison among proposed method and methods of Ret. [10], Ret. [15]

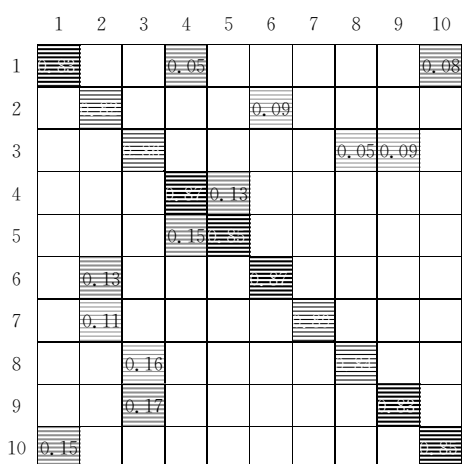
Methods	$RR_{avg}$	$T_{avg}/s$
Ref. [10]	0.69	47.3
Ref. [15]	0.75	58.5
Proposed method	0.86	63.1



(a)文献[10]提出算法识别结果的混淆矩阵  
(a)Confusion matrix of Ref. [10]



(b)文献[15]提出算法识别结果的混淆矩阵  
(b)Confusion matrix of Ref. [15]



(c)本文提出算法识别结果的混淆矩阵  
(c)Confusion matrix of proposed algorithms

图 10 3 种算法识别结果的混淆矩阵

Fig. 10 Confusion matrix of three algorithms

从表 3 中可以看出,文献[10]和[15]算法平均识别率分别为 0.69 和 0.75,远低于本文提出算法。这是因为,文献[10]算法以人体剪影的 Procrustes 形状为特征实现对人体跑、跳、弯腰等行为的识别,与这些动作相比,不同种类的微反应动作之间姿态、形状差异较小,例如双腿交叉和锁脚动作从姿态形状上来看几乎一样,用剪影形状无法区分。文献[15]算法利用光流与 LBP-TOP 的结合特征识别人体面部微表情,该特征适合描述图像的局部细节,但容易忽略图像的整体特征。人体下肢包含胯、膝、踝等关节,不仅不同被测试人做出的同一种微反应动作之间具有一定差异,同一被测试人多次做出的同一种微反应动作之间也存在一些微小差异,这些微小差异对识别结果会产生一定的影响。

在识别速度方面,由于不同类型的微反应动作持续时间不同,为方便比较,测试视频统一使用 3 s 时长,  $T_{avg}$  是所有样本的平均识别时间。从表 3 中可看出本文算法与文献[15]算法的平均识别时间相差不大,分别为 63.1 s 和 58.5 s,文献[10]算法的平均识别时间较短,只有 47.3 s,这是因为文献[10]算法的特征计算相对简单。

本文提出算法对 10 类微反应动作的识别率均在 0.83~0.91 之间,平均识别率达到 0.86,说明本文提出的算法对人体下肢的各种微反应动作都有较好的识别效果。这是因为本文使用的时空金字塔韦伯特征,对韦伯特征进行了时空方向的拓展,提高了其对一帧图像内细节和相邻图像帧间变化的区分能力,同时利用多尺度金字塔思想,不仅关注下肢微动作细节的变化,同时关注下肢微动作的整体姿态。

## 5 结 论

本文在分析人体下肢微反应动作特点的基础上,设计了一种时空金字塔韦伯局部描述子,利用该描述子提取人体下肢微反应动作特征,建立人体下肢微反应动作字典,实现动作的分类识别。通过实验分析了  $L$ 、 $d_{PCA}$ 、 $n_{Atom}$ 、 $C$  等参数对识别结果的影响并确定了最优参数值  $L=3$ ,  $d_{PCA}=30$ ,  $n_{Atom}=40$ ,  $C=10$ ,使得本文提出的算法对人体下肢微反应动作的平均识别率达到 0.86,与其他算法相比,本文算法有着更强的适用性与更高的分类识别率。

## 参考文献:

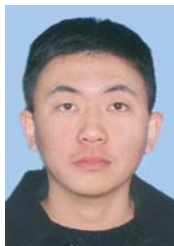
- [1] 姜振宇. 审讯中的应激微反应的应用[J]. 中国检察官, 2014(18):59-64.  
JIANG ZH Y. The application of micro-expressions in the trial [J]. *The Chinese Procurators*, 2014 (18):59-64. (in Chinese)
- [2] 吴奇, 申寻兵, 傅小兰. 微表情研究及其应用[J]. 心理科学进展, 2010, 18(9):1359-1368.  
WU Q, SHENG X B, FU X L. Micro-expression and its applications [J]. *Advances in Psychological Science*, 2010, 18(9):1359-1368. (in Chinese)
- [3] SHEN X B, WU Q, FU X L. Effects of duration of expressions on the recognition of microexpressions [J]. *Journal of Zhejiang University Science B*, 2012, (13):221-230.
- [4] XIE L, LIU X, YANG X J, *et al.*. Cognitive regulation and emotion modeling for micro-expression [J]. *International Journal of Control and Automation*, 2016, 9(2):361-372.
- [5] TANG A, LU K, WANG Y F, *et al.*. A real-time hand posture recognition system using deep neural networks [J]. *ACM Transactions on Intelligent Systems and Technology*, 2015, 6(2):21.
- [6] DOAN H G, NGUYEN V T, VU H, *et al.*. A combination of user-guide scheme and kernel descriptor on RGB-D data for robust and realtime hand posture recognition [J]. *Engineering Applications of Artificial Intelligence*, 2016, 49:103-113.
- [7] MARCOS-RAMIRO A, PIZARRO-PEREZ D, MARRON-ROMERA M, *et al.*. Body communicative cue extraction for conversational analysis[C]. *IEEE Proceedings of 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, 2013.
- [8] YANG NG, LI J T, WANG R. A method of lower limb joint points extraction based on pendulum model under arbitrary gesture walk[J]. *Applied Mechanics and Materials*, 2014, 556-562:4347-4351.
- [9] KWAK N J, SONG T S. Human action recognition using accumulated moving information [J]. *International Journal of Multimedia and Ubiquitous Engineering*, 2015, 10(10):211-222.
- [10] 蔡加欣, 冯国灿, 汤鑫, 等. 基于姿势字典学习的人体行为识别[J]. 光学学报, 2014, 34(12):1215002.  
CAI J X, FENG G C, TANG X, *et al.*. Human action recognition by leaning pose dictionary [J]. *Acta Optica Sinica*, 2014, 34 (12): 1215002. (in Chinese)
- [11] CHEN J, SHAN S G, HE C, *et al.*. WLD:a robust local image descriptor [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(9):1705-1720.
- [12] ZHANG ZH, WANG L, ZHU Q, *et al.*. Pose-invariant face recognition using facial landmarks and Weber local descriptor [J]. *Knowledge-Based Systems*, 2015, 84:78-88.
- [13] ALHUSSEIN M. Automatic facial emotion recognition using weber local descriptor for e-Healthcare system [J]. *Cluster Computing*, 2016, 19(1):99-108.
- [14] ZHAO G Y, PIETIKAINEN M. Dynamic texture recognition using local binary patterns with an application to facial expressions [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(6):915-928.
- [15] 张轩阁, 田彦涛, 郭艳君, 等. 基于光流与 LBP-TOP 特征结合的微表情识别[J]. 吉林大学学报(信息科学版), 2015, 33(5):516-523.  
ZHANG X G, TIAN Y T, GUO Y J, *et al.*. Micro-expression recognition based on feature combination of optical flow and LBP-TOP [J]. *Journal of Jilin University (Information Science Edition)*, 2015, 33(5):516-523. (in Chinese)
- [16] 王雪虎, 杨健, 艾丹妮, 等. 结合先验稀疏字典和空洞填充的 CT 图像肝脏分割[J]. 光学精密工程, 2015, 23(9):2687-2697.  
WANG X H, YANG J, AI D N, *et al.*. Liver segmentation in CT image based on priori sparse dictionary and hole filling [J]. *Opt. Precision Eng.*, 2015, 23(9):2687-2697. (in Chinese)
- [17] 郝云胜, 叶艺山, 邓振森, 等. FEKO 稀疏微多普勒建模及 CS 重构方法[J]. 光学精密工程, 2016, 24(6):1482-1489.  
HAO Y SH, YE Y SH, DENG ZH M, *et al.*. FEKO sparse micro-doppler modeling and CS re-

construction method [J]. *Opt. Precision Eng.*, 2015, 24(6):1482-1489. (in Chinese)

- [18] 高向东,李国华,萧振林,等.焊接缺陷的磁光成像小波多尺度识别及分类[J].*光学精密工程*, 2016, 24(4):930-936.

GAO X D, LI G H, XIAO ZH L, *et al.*. Detection and classification of welded defects by magneto-optical imaging based on multi-scale wavelet [J]. *Opt. Precision Eng.*, 2015, 24(4):930-936. (in Chinese)

#### 作者简介:



王昊鹏(1981—),男,山东济南人,博士后,讲师,2004年、2008年、2014年于山东大学分别获得学士、硕士、博士学位,主要从事智能检测与控制方面的研究。E-mail:whp-whp-whp@163.com



冯显英(1965—),男,山东济宁人,博士,教授,1987年、1992年、1998年于山东工业大学分别获得学士、硕士、博士学位,主要从事数字化制造技术、智能检测与控制方面的研究。E-mail:fxying@sdu.edu.cn