

# 基于判别协作表征分类器的人体行为识别

咎宝锋<sup>1</sup>, 孔军<sup>1,2</sup>, 蒋敏<sup>1</sup>

<sup>1</sup>江南大学物联网工程学院, 江苏 无锡 214122;

<sup>2</sup>新疆大学电气工程学院, 新疆 乌鲁木齐 830047

**摘要** 为了解决协作表征分类器(CRC)对相似样本误分类概率高的问题,提出一种判别协作表征分类器(DCRC)。该分类器考虑了所有训练样本和每一类样本对协作表征系数的影响,得到判别性强的协作表征系数,提升了对相似样本的判别性。基于DCRC进行人体行为识别研究。首先用深度运动映射图(DMMs)提取深度动作序列特征,得到DMMs特征描述子,然后利用DCRC对特征描述子进行协作表征编码,最后利用新的判别规则进行分类识别。在人体行为识别数据集上的实验结果表明,DCRC对相似动作具有一定的判别性,且识别精度优于现有的方法。

**关键词** 图像处理; 人体行为识别; 判别协作表征分类器; 协作表征分类器; 深度图; 深度运动映射图

中图分类号 TP391

文献标识码 A

doi: 10.3788/LOP55.011010

## Human Action Recognition Based on Discriminative Collaborative Representation Classifier

Zan Baofeng<sup>1</sup>, Kong Jun<sup>1,2</sup>, Jiang Min<sup>1</sup>

<sup>1</sup>School of Internet of Things Engineering, Jiangnan University, Wuxi, Jiangsu 214122, China;

<sup>2</sup>College of Electrical Engineering, Xinjiang University, Urumqi, Xinjiang 830047, China

**Abstract** In order to solve the problem of high probability of misclassification for similar samples of the collaborative representation classifier (CRC), we propose a discriminative CRC (DCRC), which takes the effect of all training samples and each class of training samples on the collaborative representation coefficient into account. The coefficient obtained has strong discrimination and can improve the discriminability of the similar samples. Human action recognition is conducted based on DCRC. We first extract the features of depth action sequence via depth motion maps (DMMs). Then, we use DCRC to encode the DMMs features and perform classification and recognition by new classification rules. Experimental results on the human action recognition datasets show that the DCRC has certain discriminative properties for similar actions, and the recognition accuracy is superior to some existed methods.

**Key words** image processing; human action recognition; discriminative collaborative representation classifier; collaborative representation classifier; depth maps; depth motion maps

**OCIS codes** 100.4993; 100.4994; 100.5010

## 1 引言

人体行为识别是模式识别、机器学习、计算机视

觉领域的研究热点。近年来,人体行为识别技术被应用于多个领域,如群体行为识别已应用于智能监控系统 and 人机交互游戏等<sup>[1-3]</sup>。当前,人体行为识别

收稿日期: 2017-07-21; 收到修改稿日期: 2017-08-20

基金项目: 国家自然科学基金(61362030,61201429)、中国博士后科学基金(2015M571720,2016M606360)、江苏省博士后科学基金(1601416C)、中国公安部技术科研项目(2014JSYJB007)

作者简介: 咎宝锋(1992—),男,硕士研究生,主要从事人体行为识别、机器学习方面的研究。

E-mail: zanbf1123@163.com

导师简介: 孔军(1974—),男,博士,副教授,主要从事目标跟踪、人体行为识别方面的研究。

E-mail: kongjun@jiangnan.edu.cn

主要针对 4 种对象:基于视频流序列<sup>[4-5]</sup>,静态图像序列,深度图像序列和可穿戴传感器。基于静态图像序列的行为识别关注度很高,但是由于二维图像容易受到环境变化以及光照变化的影响,在实际应用中效果并不理想。由于深度信息具有背景简单、受环境影响小等特点,基于深度信息序列的行为识别逐渐成为新兴热点。

人体行为识别最重要的两个步骤为特征提取和行为分类。特征提取方面,文献[6]首先将深度运动图像投影到 3 个视角上,通过差分、阈值化和累加计算得到每个视角下的深度运动映射图(DMMs),然后从 DMMs 提取方向直方图特征(HOG)作为人体行为特征描述子。虽然,DMMs-HOG 特征描述子可以有效地表达出运动对象的形状和运动信息,但是高维的特征增加了计算复杂度,难以保证实时性。文献[7]对 DMMs-HOG 特征描述子进行了改进,省略了 HOG 特征提取,降低了计算复杂度,提高了人体行为识别的实时性。

行为分类方面,最常用的分类器为支持向量机(SVM),但是需要先进行训练,再进行测试。当数据量大时,SVM 训练的时间过长。大量实验证明,稀疏表示分类器(SRC)可以有效解决人脸识别和图像分类等问题<sup>[7-12]</sup>,且 SRC 所需训练时间短,甚至不需要训练。Wright 等<sup>[8]</sup>最早提出 SRC,对系数进行了 1 范数约束,最终优化得到的系数向量具有稀疏性,但是 SRC 的优化求解较为困难。Zhang 等<sup>[9]</sup>认为,2 范数约束不仅使优化过程变得简单,而且识别率相较于 1 范数约束没有发生下降,基于此,提出了协作表征分类器(CRC)。文献[7]提出了 2 范数正则化 CRC(L2-CRC),并将其应用到人体行为识别领域。首先对深度动作序列提取 DMMs 特征描述子,然后利用提出的分类器进行稀疏编码及分类,从而实现对人体行为的实时识别。文献[10]提出了稀疏增强 CRC,结合稀疏特征和密集特征对测试样本进行分类,在人脸库等公开的数据集上进行验证,得到很好的分类结果。文献[11]将常见的线性回归分类器和协作表征分类器结合起来,在人脸库得到了良好的识别效果。文献[12]将稀疏表示分类器和 k 近邻结合起来,提出一种稳健性强的高光谱图像分类方法。

虽然,SRC/CRC 通过对表征系数进行约束增加了稀疏性,降低了过拟合的风险,然而当训练集中有 2 类动作非常相似时,这 2 类样本对重构测试样本的贡献力度相近,那么利用测试样本和重构样本

之间残差的分类规则产生的误分类概率就会变大。为此,本文基于 L2-CRC<sup>[9]</sup> 提出判别协作 CRC(DCRC),通过对协作表征系数的两次 2 范数约束,得到重构能力较强的协作表征系数,并使用新的分类规则来判定测试样本的类别,提高分类器的分类精度。

## 2 特征提取

DMMs 特征描述子<sup>[9]</sup>具有实时性高、计算简单的特点,因此,参照文献[7]中的方法提取 DMMs 特征描述子。

首先将  $N$  帧动作深度序列的每一帧深度图都投影到笛卡尔直角坐标系的三个方向上,得到三幅视图(前视图、侧视图、上视图)。这一操作可以将一个三维的深度图像映射成为三个二维图像,分别用  $m_f$ 、 $m_s$ 、 $m_t$  表示三个二维图。然后分别计算相同视角下连续两帧投影图之差的绝对值,累加得到 DMMs 特征描述子:

$$C_{\text{DMMs}} = \sum_{i=a}^b |m_{\{f,s,t\}}^{i+1} - m_{\{f,s,t\}}^i|, \quad (1)$$

式中  $m_f^i$  表示第  $i$  帧深度图在前视角下的投影图像。在一个动作深度序列中,前几帧和最后几帧中的对象实际上是处于静止状态或者只有微小的移动,这对提取 DMMs 特征描述子没有任何贡献或者贡献非常小,因此在实际操作中起始帧和终止帧分别设置为  $a \in \{2, 3, \dots, N\}$ 、 $b \in \{2, 3, \dots, N\}$ 。

提取到的 DMMs 特征描述子通过限位框来提取非 0 区域,即感兴趣区域,从而更加精确地捕捉运动动作的变化区域。最后,经过 PCA 降维处理,得到最终的 DMMs 特征描述子。由(1)式可以看出,DMMs 特征描述子的计算非常简单,可有效提高特征提取阶段的实时性,并且该特征将三维图像映射为二维图像,将行为识别问题转化成为图像分类问题。

## 3 行为分类

### 3.1 CRC

CRC 是一种非常有效的分类器,它认为测试样本可以由所有训练样本近似线性表示。给定一个具有  $K$  类的训练样本集合  $\mathbf{D} = \{\mathbf{D}_1, \dots, \mathbf{D}_k, \dots, \mathbf{D}_K\}$ ,其中  $\mathbf{D}_k$  ( $k=1, 2, \dots, K$ ) 是类别  $k$  所对应的样本向量集合。用  $\mathbf{y}$  来表示一个测试样本,则可以由  $\mathbf{D}$  协作表征为  $\mathbf{y} \approx \mathbf{D}\boldsymbol{\alpha}$ ,其中  $\boldsymbol{\alpha}$  是测试样本的协作表征系数向量。协作表征分类器算法过程如下。

输入: 训练样本集合, 测试样本  $\mathbf{y}$ , 正则化参数  $\lambda$ 。

1) 基于目标函数, 优化协作表征系数向量:

$$\hat{\boldsymbol{\alpha}} = \operatorname{argmin}_{\boldsymbol{\alpha}} \{ \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \lambda \|f(\boldsymbol{\alpha})\|_2^2 \}, \quad (2)$$

式中  $f(\boldsymbol{\alpha})$  为关于  $\boldsymbol{\alpha}$  的函数。

2) 计算重构样本和测试样本间的残差:

$$e_i = \|\mathbf{y} - \mathbf{D}_i \hat{\boldsymbol{\alpha}}_i\|_2^2, \quad (3)$$

式中  $\hat{\boldsymbol{\alpha}}_i (i=1, 2, \dots, K)$  为类别  $i$  的训练样本所对应的系数向量。

3) 计算测试样本的标签

$$\operatorname{label}(\mathbf{y}) = \operatorname{argmin}_i (e_i). \quad (4)$$

输出:  $\operatorname{label}(\mathbf{y})$ 。

L2-CRC 使用 Tikhonov 矩阵作为正则项对协作表征系数进行约束, 目标函数为

$$\hat{\boldsymbol{\alpha}} = \operatorname{argmin}_{\boldsymbol{\alpha}} \{ \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\Gamma}\boldsymbol{\alpha}\|_2^2 \}, \quad (5)$$

式中  $\boldsymbol{\Gamma}$  为 Tikhonov 正则矩阵, 形式如下:

$$\boldsymbol{\Gamma} = \begin{bmatrix} \|\mathbf{y} - \mathbf{D}_1\|_2^2 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \|\mathbf{y} - \mathbf{D}_K\|_2^2 \end{bmatrix}. \quad (6)$$

在对(5)式进行优化求解时,  $\boldsymbol{\Gamma}$  可以对系数向量  $\boldsymbol{\alpha}$  提供先验知识, 从而更精确地逼近最优解, 得到的表征系数向量  $\boldsymbol{\alpha}$  相较于 CRC 更加稳健。但对于具有相似动作的训练样本, 利用表征系数向量  $\boldsymbol{\alpha}$  重构出来的相似样本之间判别性不强。本文通过二次约束项来提升表征系数  $\boldsymbol{\alpha}$  的判别性和分类器的性能。

### 3.2 DCRC

用  $S$  来表示由所有训练样本协作张成的线性空间, 用  $S_i$  表示由一个相同类别的样本  $\mathbf{D}_i (i=1, 2, \dots, K)$  张成的线性子空间, 用  $L = \{1, 2, \dots, K\}$  来表示所有类别的集合。由 3.1 节可知, 对于不属于空间  $S$  内的测试样本可以表示为  $\mathbf{y} \approx \mathbf{D}\boldsymbol{\alpha}$ , 这仅仅能够表明测试样本的类别是属于  $L$  中的一类。接着利用测试样本和每一个类别重构样本之间的残差来近似判定出样本的类别。然而, 当训练样本中存在 2 类样本非常相似时 (如  $\mathbf{D}_i$  和  $\mathbf{D}_j$ ), CRC 得到的表征系数向量中相对应的系数  $\boldsymbol{\alpha}_i$  和  $\boldsymbol{\alpha}_j$  重构出的样本相似度很高, 那么基于残差的分类规则产生误分类的概率就会变大。

为了提高 CRC 对相似动作的判别性, 提高分类器的性能, 本文通过对系数进行二次约束获得判别性强的协作表征系数, 据此提出 DCRC。首先, 确定空间  $S$  中一个共享样本点  $\hat{\mathbf{y}} = \mathbf{D}\boldsymbol{\alpha}^* = \mathbf{D}(\boldsymbol{\alpha}_1^*, \dots, \boldsymbol{\alpha}_K^*)$ , 其中,  $\boldsymbol{\alpha}^*$  为样本点  $\hat{\mathbf{y}}$  相对应的表征系数向量, 该共享

样本点要满足两个条件: 1) 与测试样本之间的相似度很高; 2) 与每一个子空间  $S_i$  中重构的样本点  $\hat{\mathbf{y}}_i = \mathbf{D}_i \boldsymbol{\alpha}_i^*$  距离之和最小。然后, 经过对目标函数的不断优化, 可以得到最佳的协作表征系数  $\boldsymbol{\alpha}^*$ , 同时可以得到最佳的共享样本点  $\hat{\mathbf{y}}$ 。最后, 判定共享样本点  $\hat{\mathbf{y}}$  与某一子空间中的重构样本点  $\hat{\mathbf{y}}_i (i=1, \dots, K)$  的残差最小, 得到测试样本的类别 (图 1)。

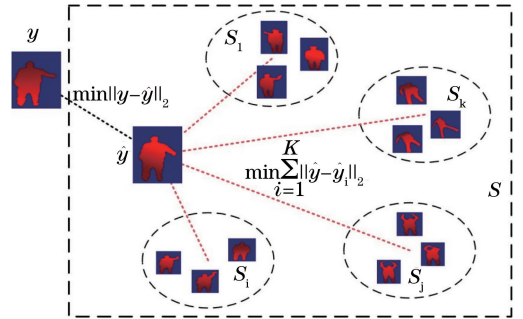


图 1 DCRC 的图示说明

Fig. 1 Illustration of DCRC

由经验可知, 当 2 个样本距离越近时, 这 2 个样本属于同一类的概率就越大。假设所有样本都是独立分布的, 用  $l(\mathbf{y})$  表示样本  $\mathbf{y}$  的标签, 则测试样本  $\mathbf{y}$  属于类别  $i$  的概率为

$$\begin{aligned} P[l(\mathbf{y}) = i] &= \\ P[l(\mathbf{y}) = l(\hat{\mathbf{y}}) | l(\hat{\mathbf{y}}) = i] P[l(\hat{\mathbf{y}}) = i] &= \\ P[(l(\mathbf{y}) = l(\hat{\mathbf{y}}) | l(\hat{\mathbf{y}}) = i)] & \\ P[l(\hat{\mathbf{y}}) = i | l(\hat{\mathbf{y}}) \in L] P[l(\hat{\mathbf{y}}) \in L] & \quad (7) \end{aligned}$$

由于样本之间相互独立, 那么  $P[l(\mathbf{y}) = l(\hat{\mathbf{y}}) | l(\hat{\mathbf{y}}) = i] = P[l(\mathbf{y}) = l(\hat{\mathbf{y}}) | l(\hat{\mathbf{y}}) \in L]$ , 则(7)式等价于:

$$\begin{aligned} P[l(\mathbf{y}) = i] &= \\ P[l(\mathbf{y}) = l(\hat{\mathbf{y}}) | l(\hat{\mathbf{y}}) = i] P[l(\hat{\mathbf{y}}) = i] &= \\ P[l(\mathbf{y}) = l(\hat{\mathbf{y}})] P[l(\hat{\mathbf{y}}) = i | l(\hat{\mathbf{y}}) \in L] & \quad (8) \end{aligned}$$

式中  $P[l(\mathbf{y}) = l(\hat{\mathbf{y}})]$  可以度量测试样本  $\mathbf{y}$  与定点  $\hat{\mathbf{y}}$  之间的距离, 即等价于  $\|\mathbf{y} - \hat{\mathbf{y}}\|_2^2$ 。因为  $\hat{\mathbf{y}}_i$  是落在子空间  $S_i$  内部的, 因此  $P[l(\hat{\mathbf{y}}) = i | l(\hat{\mathbf{y}}) \in L]$  可以认为度量了定点  $\hat{\mathbf{y}}$  和  $\hat{\mathbf{y}}_i$  之间的距离, 即  $\sum_{i=1}^K \|\hat{\mathbf{y}} - \mathbf{D}_i \boldsymbol{\alpha}_i\|_2^2$ 。为了得到测试样本的标签, 则有:

$$\max P[l(\mathbf{y}) = i] = \min \left( \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \mu \sum_{i=1}^K \|\hat{\mathbf{y}} - \mathbf{D}_i \boldsymbol{\alpha}_i\|_2^2 \right). \quad (9)$$

为了降低过拟合的风险与计算复杂度, 利用文献[6]中的 Tikhonov 矩阵正则项进行约束, 得到最终的目标函数:

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \left\{ \|\mathbf{y} - \mathbf{D}\alpha\|_2^2 + \lambda \|\Gamma\alpha\|_2^2 + \mu \sum_{i=1}^K \|\hat{\mathbf{y}} - \mathbf{D}_i\alpha_i\|_2^2 \right\}, \quad (10)$$

式中  $\lambda$  和  $\mu$  为正则化约束参数, 都是标量;  $\sum_{i=1}^K \|\hat{\mathbf{y}} - \mathbf{D}_i\alpha_i\|_2^2$  基于  $\mathbf{D}_i$  对  $\alpha_i$  进行二次约束, 可以使最终的系数向量  $\alpha$  判别性增强。当  $\mu = 0$  时, (10)

式的模型就变为 CRC 模型, 因此要保证  $\mu > 0$ 。(7)~(10)式从概率角度证明了 DCRC 的可行性, 能够通过取测试样本  $\mathbf{y}$  属于单一类别的概率的最大值来判定其所属类别。

约束项  $\sum_{i=1}^K \|\hat{\mathbf{y}} - \mathbf{D}_i\alpha_i\|_2^2$  对系数向量  $\alpha$  的偏导求解过程如下:

$$\begin{aligned} \frac{\partial}{\partial \alpha} \left( \sum_{i=1}^K \|\hat{\mathbf{y}} - \mathbf{D}_i\alpha_i\|_2^2 \right) &= \frac{\partial}{\partial \alpha} \sum_{i=1}^K \operatorname{tr} [(\hat{\mathbf{y}} - \mathbf{D}_i\alpha_i)^T (\hat{\mathbf{y}} - \mathbf{D}_i\alpha_i)] = \\ &= \sum_{i=1}^K \frac{\partial}{\partial \alpha} \operatorname{tr} (\alpha^T \mathbf{D}^T \mathbf{D} \alpha - \alpha^T \mathbf{D}^T \mathbf{D}_i \alpha_i - \alpha_i^T \mathbf{D}_i^T \mathbf{D} \alpha + \alpha_i^T \mathbf{D}_i^T \mathbf{D}_i \alpha_i) = \\ &= \sum_{i=1}^K \{ 2\mathbf{D}^T \mathbf{D} \alpha - 2(\mathbf{D}^T [0, \dots, \mathbf{D}_i, \dots, 0] \alpha + [0, \dots, \mathbf{D}_i^T, \dots, 0] \mathbf{D} \alpha) + 2[0, \dots, \mathbf{D}_i^T, \dots, 0] [0, \dots, \mathbf{D}_i, \dots, 0] \alpha \}. \quad (11) \end{aligned}$$

令  $\bar{\mathbf{D}}_i = [0, \dots, \mathbf{D}_i, \dots, 0]$ , 则(11)式可化简为

$$\sum_{i=1}^K 2\mathbf{D}^T \mathbf{D} \alpha - 2(\mathbf{D}^T \bar{\mathbf{D}}_i \alpha + \bar{\mathbf{D}}_i^T \mathbf{D} \alpha) + 2\bar{\mathbf{D}}_i^T \bar{\mathbf{D}}_i \alpha. \quad (12)$$

与 CRC 模型的最优解结合起来, 可以得到判别协作表征分类器的最优解:

$$\hat{\alpha} = [\mathbf{D}^T \mathbf{D} + \lambda \Gamma^T \Gamma + \mu \sum_{i=1}^K (\mathbf{D} - \bar{\mathbf{D}}_i)^T (\mathbf{D} - \bar{\mathbf{D}}_i)]^{-1} \mathbf{D}^T \mathbf{y}. \quad (13)$$

最后, 采用一个新的判别规则来判定测试样本的类别:

$$e_i = \|\mathbf{D}\hat{\alpha} - \mathbf{D}_i\hat{\alpha}_i\|_2^2, \quad (14)$$

$$\operatorname{label}(\mathbf{y}) = \underset{i}{\operatorname{argmin}} [e(i)]. \quad (15)$$

## 4 实验结果

在 MSR Action 3D Dataset 和 UTD-MHAD 数据集上对所提出的行为识别方法进行验证。

### 4.1 参数分析

(10)式中包含 2 个正则化参数  $\lambda$  和  $\mu$ , 有利于表征系数更好地拟合训练数据, 且保持表征系数的值较小, 防止出现参数过大、过拟合的情况。正则化参数值不能过大或者过小: 过大的正则化参数会使最终的表征系数趋近于 0, 产生欠拟合的情况, 模型的泛化能力会很差; 过小的表征系数容易造成过拟合的情况。根据经验, 正则化参数的取值一般在 0.1~1 之间。

### 4.2 MSR Action 3D Dataset

MSR Action 3D Dataset 是深度摄像头采集到的深度序列动作集, 包含 20 个动作, 每个动作由 10

个人进行演示, 每次演示 2~3 次, 采集到的深度图分辨率为 320 pixel×240 pixel。为了保证试验的公正性, 采用文献[7]中的设置。

从表 1 可以看到, 在所有的实验中本文方法的识别率都高于文献[6]和文献[13], 并且大部分实验结果较文献[7]均有所提升。在实验 2 中, 本文方法在所有子集上的识别率都达到 100%。在实验 3 中, 本文方法的识别率较文献[9]均有所提升。

为了验证 DCRC 的稳定性, 随机选择实验对象的演示动作作为训练数据, 其余作为测试数据, 进行 100 次实验, 计算平均识别率和平均标准差, 结果如表 2 所示。与前 3 项实验的结果相比, Test 1 和 Test 2 的识别率相差不大, 并且标准差都比较小。Test 3 虽然识别率有所下降, 但是平均识别率达到了 80%, 并且标准差相差不大, 说明 DCRC 具有一定的稳定性。

将编号为 1、3、5、7、9 的对象所演示的所有动作作为训练数据, 其余的作为测试数据, 使用相同的特征描述子, 不同分类器的识别率如表 3 所示, 其中 SRC 使用的是 1 范数约束, 不同方法的识别率如表 5 所示。从表 3 可以看到, DCRC 的识别率优于 SRC 和 L2-CRC, 相比 L2-CRC, DCRC 的识别率提升了 3.78%。从表 4 可以看出, 本文方法的识别精度高于其他方法, 识别精度最高提高了 4.67%。

图 2 为在 Action 3D 数据集上使用 DMMs+L2-CRC 和 DMMs+DCRC 方法的混淆图。使用 L2-CRC 时, 相似动作 Tennis serve 和 Tennis swing, 以及 Draw x、Draw tick 和 Draw circle 非常容易产生混淆。当使用 DCRC 进行识别时, Tennis swing 的识别率提



高到 93%，Draw x 的识别率提高到 85.7%，Draw circle 的识别率提高到 60%，同时，其他动作，如

bend 的识别率达到 100%。这表明 DCRC 可以提高对相似动作的识别率。

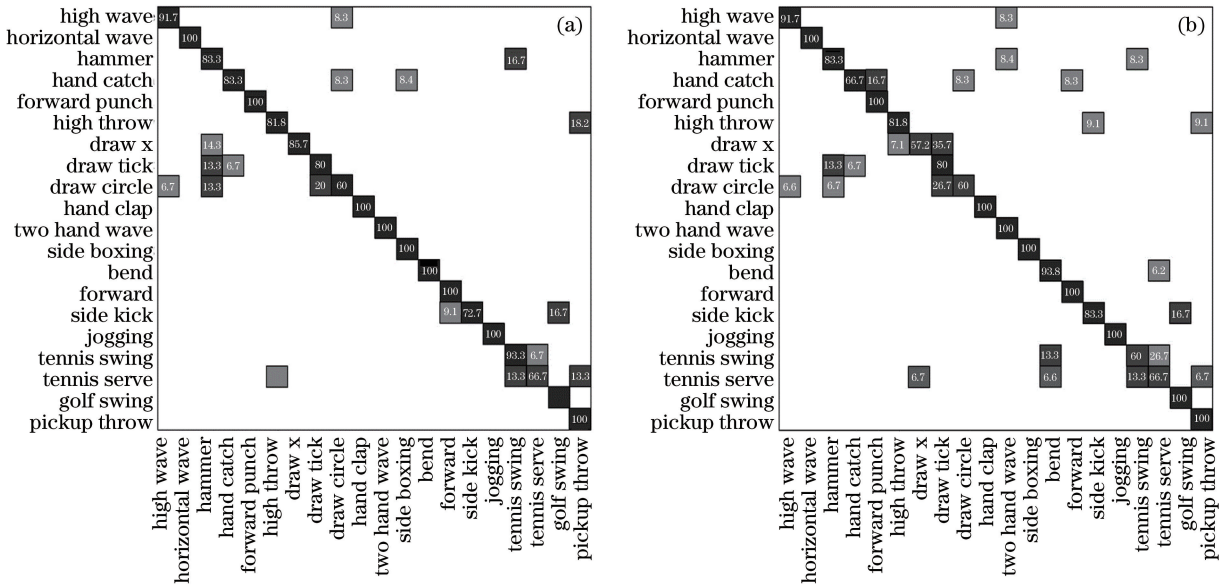


图 2 MSR Action 3D 数据集上 DMMs+CRC(a)与 DMMs+DCRC(b)的混淆图

Fig. 2 Confusion matrix of DMMs+CRC (a) and DMMs+DCRC (b) on MSR Action 3D dataset

表 1 MSR Action 3D 数据集上不同方法的识别率

Table 1 Recognition rates of different methods on MSR Action 3D dataset

Test	Subset	Ref.[6]	Ref.[7]	Ref.[13]	Ref.[14]	Ref.[15]	Proposed
		method	method	method	method	method	
Test 1	AS1	97.3	97.3	89.5	98.2	94.7	98.0
	AS2	92.2	96.1	89.0	94.8	95.4	96.7
	AS3	98.0	98.7	96.3	97.4	97.3	96.7
	Average	95.8	97.4	91.6	96.8	95.8	97.2
Test 2	AS1	98.7	98.6	93.4	99.1	97.3	100
	AS2	94.7	98.7	92.9	97.0	98.7	100
	AS3	98.7	100	96.3	98.7	97.3	100
	Average	97.4	99.1	94.2	98.3	97.8	100
Test 3	AS1	96.2	96.2	72.9	84.7	74.5	97.1
	AS2	84.1	83.2	71.9	81.3	76.1	83.2
	AS3	94.6	92.0	79.2	88.4	96.4	92.9
	Average	91.6	90.5	74.7	84.8	82.3	91.1

表 2 MSR Action 3D 数据集上实验 4 的平均识别率和平均标准差

Table 2 Average recognition rates and its standard deviation of MSR Action 3D dataset on test 4

Subset	Test 1		Test 2		Test 3	
	Recognition rate / %	Standard deviation	Recognition rate / %	Standard deviation	Recognition rate / %	Standard deviation
AS1	97.6	0.8	98.8	1.1	80.0	3.2
AS2	96.8	1.0	98.4	1.2	70.0	4.0
AS3	97.1	1.0	98.9	0.9	90.0	3.4
Average	97.2	0.9	98.7	1.1	80.0	3.5

表3 MSR Action 3D数据集 DMMs  
特征下不同分类器的识别率

Table 3 Recognition rates of different classifier on MSR  
Action 3D dataset with DMMs features

Classifier	Recognition rate / %
PRF-SVM	85.50
SRC	86.01
L2-CRC	86.23
DCRC	90.01

表4 MSR Action 3D数据集上不同方法的识别率  
Table 4 Recognition rates of different methods  
on MSR Action 3D dataset

Method	Recognition rate / %
Ref.[6] method	85.52
Ref.[16] method	88.89
Ref.[17] method	88.80
Ref.[18] method	89.45
Ref.[19] method	89.86
Proposed method	90.19

### 4.3 UTD-MHAD

Chen 等<sup>[20]</sup>提出了一个新的行为识别数据集,命名为 UTD Multimodal Human Action Dataset (UTD-MHAD)。该数据集包含多种模式:RGB 视

频序列,深度序列,由 Kinect 采集的骨骼点位置信息,可穿戴式传感器采集的加速和旋转信号。共包含 27 个动作,每个动作由 8 个人来演示(4 男 4 女),每个动作演示 4 遍。除去 3 个损坏的序列,该数据集一共包含 861 个动作序列。

在 UTD-MHAD 数据集上,选择编号为 1、3、5、7 的对象所演示的动作作为训练数据,剩余的作为测试数据,最终的识别率如表 5 所示。本文方法的识别率比文献[6]方法的识别率提高了 4.27%。图 3 为这两种方法在该数据集上结果的混淆图。比较发现,本文方法对相似动作 baseball\_swing 和 tennis\_swing 的识别率有所提升,对 sit2stand 和 stand2sit 的识别率显著提升,表明 DCRC 的二次约束对相似动作具有一定的判别性。

表5 UTD-MHAD数据集上不同方法的识别率

Table 5 Recognition rate of different  
methods on UTD-MHAD dataset

Method	Recognition rate / %
Ref.[20] method	66.10
Proposed method	70.37

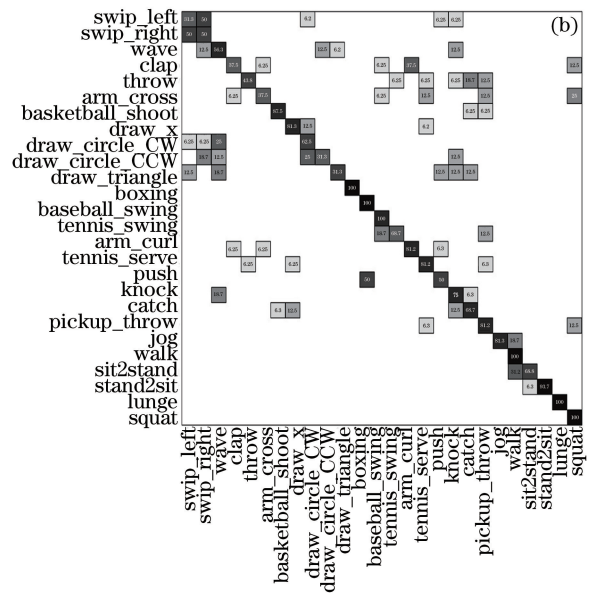
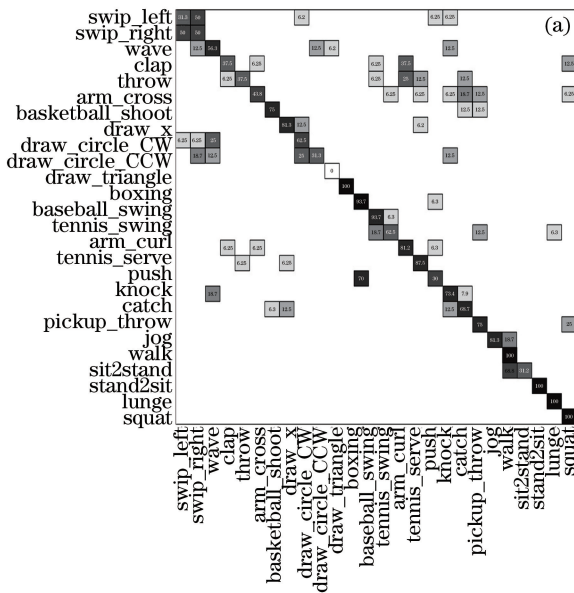


图3 UTD-MHAD数据集上 DMMs+CRC(a)与 DMMs+DCRC(b)的混淆图

Fig. 3 Confusion matrix of DMMs+CRC (a) and DMMs+DCRC (b) on UTD-MHAD dataset

## 5 结 论

提出了 DCRC,通过对协作表征系数的二次约束,增强了对相似动作的判别性,提升了分类器的性能。基于该分类器进行人体行为识别。首先,采用复杂度低、实时性好的 DMMs 作为动作的特征描述子,然后,利用 DCRC 对提取的 DMMs 特征描述子

进行分类识别,从而实现人体行为识别。在两个数据集上的实验结果表明,DCRC 对相似动作具有一定的判别性,且较现有的一些方法具有更高的识别率。虽然 DCRC 可以提升对相似动作的识别率,但提升的效果有限,因此下一步工作将使用多尺度的深度运动映射图特征来进一步提高对相似动作的识别能力。

## 参 考 文 献

- [1] Chen C, Jafari R, Kehtarnavaz N. Improving human action recognition using fusion of depth camera and inertial sensors [J]. *IEEE Transactions on Human-Machine Systems*, 2015, 45(1): 51-61.
- [2] Chen C, Kehtarnavaz N, Jafari R. A medication adherence monitoring system for pill bottles based on a wearable inertial sensor [C]// *Proceedings of the 36th International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014: 4135-4138.
- [3] Zhang X G, Liu C X, Zuo J Q. Small scale crowd behavior recognition based on causality network analysis [J]. *Acta Optica Sinica*, 2015, 38 (8): 0815001.  
张旭光, 刘春霞, 左佳倩. 基于因果网络分析的小规模人群行为识别 [J]. *光学学报*, 2015, 35 (8): 0815001.
- [4] Cai J X, Feng G C, Tang X, *et al.* Human action recognition based on local image contour and random forest [J]. *Acta Optica Sinica*, 2014, 34 (10): 1015006.  
蔡加欣, 冯国灿, 汤鑫, 等. 基于局部轮廓和随机森林的人体行为识别 [J]. *光学学报*, 2014, 34(10): 1015006.
- [5] Cai J X, Feng G C, Tang X, *et al.* Human action recognition by learning pose dictionary [J]. *Acta Optica Sinica*, 2014, 34(12): 1215002.  
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.
- [6] Yang X, Zhang C, Tian Y. Recognizing actions using depth motion maps-based histograms of oriented gradients [C]// *Proceedings of the 20th ACM International Conference on Multimedia*, 2012: 1057-1060.
- [7] Chen C, Liu K, Kehtarnavaz N. Real-time human action recognition based on depth motion maps [J]. *Journal of Real-Time Image Processing*, 2013, 12 (1): 155-163.
- [8] Wright J, Yang A Y, Sastry S S, *et al.* Robust face recognition via sparse representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(2): 210-227.
- [9] Zhang L, Yang M, Feng X. Sparse representation or collaborative representation: which helps face recognition [C]. *IEEE International Conference on Computer Vision, IEEE*, 2012: 471-478.
- [10] Akhtar N, Shafait F, Mian A. Sparseness helps: sparsity augmented collaborative representation for classification [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 1-10.
- [11] Zhang H, Wang F, Chen Y, *et al.* Sample pair based sparse representation classification for face recognition [J]. *Expert Systems with Applications*, 2016, 45(1): 352-358.
- [12] Cui M, Prasad S. Class-dependent sparse representation classifier for robust hyperspectral image classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(5): 2683-2695.
- [13] Li W, Zhang Z, Liu Z. Action recognition based on a bag of 3D points [C]. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2010: 9-14.
- [14] Vieira I A W, Nascimento E R, Oliveira I G L, *et al.* Stop: space-time occupancy patterns for 3D action recognition from depth map sequences [M]. Heidelberg: Springer, 2012: 252-259.
- [15] Yang X, Tian Y. Eigen joints-based action recognition using naive-bayes-nearest-neighbor [C]. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2012: 14-19.
- [16] Oreifej O, Liu Z. HON4D: histogram of oriented 4D normals for activity recognition from depth sequences [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2013: 716-723.
- [17] Rahmani H, Mahmood A, Huynh D Q, *et al.* Real time action recognition using histograms of depth gradients and random decision forests [C]. *Applications of Computer Vision, IEEE*, 2014: 626-633.
- [18] Zhang C, Tian Y. Histogram of 3D facets: a depth descriptor for human action and hand gesture recognition [J]. *Computer Vision and Image Understanding*, 2015, 139: 29-39.
- [19] Evangelidis G, Singh G, Horaud R. Skeletal quads: human action recognition using joint quadruples [C]. *International Conference on Pattern Recognition*, 2014: 4513-4518.
- [20] Chen C, Jafari R, Kehtarnavaz N. UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor [C]. *IEEE International Conference on Image Processing*, 2015: 168-172.