**实验结果**

实验均采用15个人体关键点. FV编码选取高斯分布个数M=20.

**（1）JHMDB行为数据集**

在JHMDB数据集上实验结果的分类混淆矩阵如图所示.



图1 本文方法在JHMDB数据集上的分类混淆矩阵

为了评估不同类型特征的识别能力, 对姿态空间关系特征、姿态时间变化特征、时空特征串联融合、时空特征加权融合四种情形进行实验, 实验结果如表1所示.

表**1** 不同特征类型及不同特征融合方法的识别结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 特征类型 | Split1 | Split2 | Split3 | 平均识别率 |
| 姿态空间特征 | 69.4% | 61.1% | 65.6% | 63.8% |
| 姿态时间特征 | 70.9% | 72.6% | 70.9% | 71.5% |
| 时空特征串联 | 77.6% | 77.8% | 73.9% | 76.4% |
| 时空特征加权融合 | 79.5% | 78.2% | 75.5% | 77.7% |

为了评估不同特征编码表示的识别能力, 对四种聚合方式、BOF编码、FV编码三种情形进行实验. 实验结果如表2所示.

表**2** 不同特征编码表示的识别结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 编码方式 | Split1 | Split2 | Split3 | 平均识别率 |
| Min | 70.5% | 67.4% | 60.8% | 66.2% |
| Max | 67.5% | 68.9% | 62.6% | 66.4% |
| Mean | 60.1% | 55.9% | 54.7% | 56.9% |
| Mean-Max-Min | 75.0% | 72.2% | 64.9% | 70.7% |
| BOF编码 | 76.5% | 76.3% | 72.5% | 75.1% |
| FV编码 | 79.5% | 78.2% | 75.5% | 77.7% |

将人体划分为3种不同部位：上本身、下半身、全身, 这里对不同的肢体部位的识别准确率进行评估对比, 实验结果如表3所示.

表**3** 使用人体不同肢体部位的识别结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 肢体部位 | Split1 | Split2 | Split3 | 平均识别率 |
| 上半身 | 72.4% | 73.7% | 67.2% | 71.1% |
| 下半身 | 50.0% | 47.0% | 43.8% | 46.9% |
| 上本身+下半身 | 75.4% | 75.9% | 72.3% | 74.5% |
| 全身 | 79.5% | 78.2% | 75.5% | 77.7% |

本文方法与其他文献方法在JHMDB数据集上的实验结果对比表4所示.

表**4** 本文方法与其他方法识别结果比较

|  |  |  |  |
| --- | --- | --- | --- |
| 行为识别 | 特征提取 | 发表年份 | 平均识别率 |
| IDT-FV[6] | DT+HOG+HOF +MBH | 2013 | 65.9% |
| HLPF[8] | Pose | 2013 | 76.0% |
| 2Stream[9] | 2Stream CNN | 2014 | 59.4% |
| HR[20] | Pose+RNN | 2015 | 66.2% |
| P-CNN[12] | Pose+2Stream CNN | 2015 | 74.6% |
| JDDs[14] | Pose+C3D CNN | 2016 | 81.9% |
| Pose for Action[21] | Pose+IDT-FV | 2017 | 74.6% |
| 本文 | Pose |  | 77.7% |

**（2）Penn Action实验数据集**

在Penn Action数据集上实验结果的混淆矩阵如图2所示.



图2 本文方法在Penn Action数据集上的分类混淆矩阵

本文方法与其他文献方法的实验结果对比表5所示.

表**5** 本文方法与其他方法识别结果比较

|  |  |  |  |
| --- | --- | --- | --- |
| 行为识别 | 特征提取 | 发表年份 | 平均识别率 |
| HOG3D[23] | HOG3D | 2008 | 84.5% |
| STIP[24] | STIP | 2009 | 82.9% |
| Action Bank[25] | Bank of action detector | 2012 | 83.9% |
| IDT-FV[6] | DT+HOG+HOF +MBH | 2013 | 92.0% |
| Actemes[22] | a set of spatiotemporal patch classifiers | 2013 | 79.4% |
| HR[20] | Pose+RNN | 2015 | 92.8% |
| P-CNN[12] | Pose+2Stream CNN | 2015 | 97.7% |
| JDDs[14] | Pose+C3D CNN | 2016 | 95.7% |
| Pose for Action[21] | Pose+IDT-FV | 2017 | 92.9% |
| 本文 | Pose |  | 97.8% |