

Action Recognition Based on Key Poses Sequences with Searching-Based K-Means Algorithm

Yin Xin¹, Shengrong Gong², Chunping Liu²

¹School of Mathematics, Soochow University, Suzhou Jiangsu

²School of Computer Science and Technology, Soochow University, Suzhou Jiangsu

Email: 1207402095@suda.edu.cn, shrgong@suda.edu.cn, cpliu@suda.edu.cn

Received: Feb. 2nd, 2015; accepted: Feb. 13th, 2015; published: Feb. 16th, 2015

Copyright © 2015 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Vision-based human action recognition is currently one of the most active research fields. Action recognition is a cross-disciplinary field which contains theories of image processing, computer vision and artificial intelligence. Firstly, we get contours and pose presentation through background subtraction algorithm based on algebra theory and then we get the key poses of action through improved searching-based K-means algorithm. Finally actions are recognized through dynamic time warping algorithm. Experimental results on the main datasets show suitability for online recognition and real-time scenarios.

Keywords

Action Recognition, Set Theory, K-Means, Dynamic Time Warping

基于搜索算法K-Means动作关键特征序列的行为识别方法

殷鑫¹, 龚声蓉², 刘纯平²

¹苏州大学数学科学学院, 江苏 苏州

²苏州大学计算机科学与技术学院, 江苏 苏州

Email: 1207402095@suda.edu.cn, shrgong@suda.edu.cn, cpliu@suda.edu.cn

收稿日期: 2015年2月2日; 录用日期: 2015年2月13日; 发布日期: 2015年2月16日

摘要

行为识别是近年来计算机视觉领域的一个研究热点。本文在当今已有的行为识别算法的基础之上进行优化改进。通过基于代数理论的背景减法提取轮廓并进行姿势表达、通过聚类算法提取动作关键特征,并基于DTW动态时间规整算法完成动作识别。由于原始K-means算法中聚类结果对于初值的依赖性,我们引入基于搜索算法的K-means聚类算法,避免了初值对聚类结果的影响。通过在国际主要数据库上的实验,达到了较高的准确率和稳定度,并能够实现在线实时识别。

关键词

行为识别, 集合论, 平均聚类算法, 动态时间规整算法

1. 引言

动作识别是当今视觉领域的研究热点之一,被广泛应用于视频监控与人工智能等领域。基于计算机视觉的人体运动分析研究涵盖了图像处理、计算机视觉、模式识别和人工智能的理论,是一个学科交叉研究方向。由于人体轮廓的差异性以及光照变化、个体动作差异等因素的影响,模式识别领域亟需在理论上产生新的处理方法,以提高应用的广度和精度[1]。由于人体动作的特性,识别过程中难点来自空间复杂性和时间差异性。

在不同光照、视角和背景等条件下,相同行为的动作特征在不同的动作场景中会产生较为明显的差异。衣物遮挡、前景遮挡、个体动作差异等问题都会造成在空间中动作识别的复杂性[2]。空间复杂性最终会决定识别的高效性与准确率。时间差异性[2]是指人体动作发生时间点的不可预测性,以及动作持续间隔不尽相同性。这要求在识别过程中能够较为准确地辨别动作的起止时间,同时有效计算动作持续时间,对动作的时序性和连贯性做出具体分析。时间差异性不仅对识别精确性产生影响,也带来计算实时性和效率等影响识别高效性的问题[2]。近年来国内外人体动作行为识别的研究已经取得较为重大的进展,但动作的高复杂性和多变化性使得识别的精确性和高效性没有得以实现,因此识别复杂动作依旧是当前最具挑战的研究领域[2]。

现在有很多方法来表示动作和提取动作特征,例如基于运动信息的动态特征方法、基于光流信息的动态特征算法等,由于算法的复杂性和鲁棒性不足,只能较好地识别简单动作,对于复杂动作的识别就存在一定局限性,如表1所示。鉴于此,论文提出一种基于关键特征的描述性方法的行为识别,其基本流程如图1所示。由于基于关键特征描述能够较好的识别连续动作和交互动作,具有较高的鲁棒性。

2. 基于背景减除法的运动目标检测

运动目标检测是动作识别的基础,其关键在于如何快速精确地在动作视频序列的每帧图像中找到运动目标所在的位置。背景减除法是目前运动检测中最为常用的一种方法,主要原理是利用当前图像与背景的差分以检测出运动区域。背景减除法的基本思路是通过先建立一个与实际场景相似的背景模型,然后将待检测帧与背景模型比较、并进行差分运算,从而找出其中有变化的区域作为运动目标[3]。背景减除法的基本流程如图2所示,主要通过预处理、背景建模、目标检测、后处理四个步骤实现。

Table 1. Comparison among features for action recognition
表 1. 人体动作识别特征比较

特征	基本技术	鲁棒性	计算复杂度	适合动作类别
静态特征	姿态估计方法	低	低	简单动作
基于运动信息的动态特征	底层跟踪方法	低	中	简单动作
基于光流信息的动态特征	光流计算方法	中	高	简单动作
时空特征	图像处理方法	高	高	连续动作/交互动作
描述性特征	机器学习方法	高	高	连续动作/交互动作

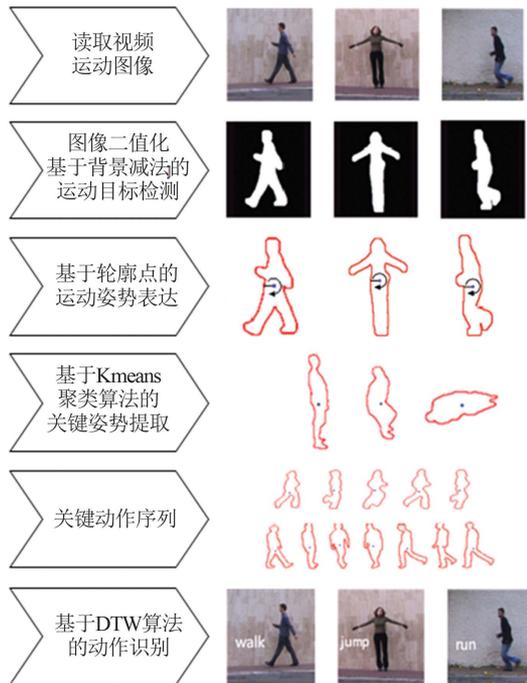


Figure 1. Basic framework of action recognition
图 1. 行为识别基本流程

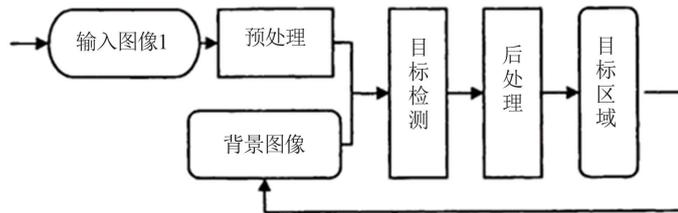


Figure 2. Flow chart of background subtraction algorithm
图 2. 背景减除算法基本流程图

图像二值形态预处理

利用数学原理处理图像是本文实验方法的创新点，将建立在集合代数理论基础之上[4]，用集合论的方法定量描述几何结构特征。集合全体代表图像中物体的形状，在二值图像中所有的黑色像素点构成的集合就是对该图像的完整表达。在二值图像中，集合全体为二维空间中的像素点的全体，集合中的每个

元素就是一个二维变量，用 (x, y) 表示，并按一定的对应法则代表图像中的一个黑色像素点。而在灰度数字图像中可以用三维集合来进行表示，集合中每个元素的前两个维度的值表示像素点的位置坐标，第三个维度表示离散的灰度值。

假定 S 表示进行运算后的二值图像集合， B 表示用来进行运算的结构元素，结构元素内的每一个元素取值为 0 或 1，由此可以组合成任何一种形状的图形，在图形中有一个中心点： X 表示原图像经过二值化后的像素集合。本文引入腐蚀、膨胀和闭运算的概念。腐蚀的作用是消除物体的边界点，使边界向内部收缩的过程，将小于结构元素的物体进行剔除[4]。膨胀运算的作用是对二值化物体边界点的扩充，将与物体有接触的背景点合并到物体中，使得物体边界向外部扩张[4]。闭运算是进行先膨胀后腐蚀的过程，其作用是填充物体内部的细小空洞、连接邻近物体、平滑边界，但不会明显改变其面积[5]。结构元素是形态学的基本算子，合理选取结构元素将直接影响图像处理的效果和质量。实验中选取圆盘形结构元素，同时设大结构元素尺寸为 R ，小结构元素尺寸为 r ，膨胀或腐蚀的次数为 k 。结构元素尺寸与运算次数的关系为：

$$R = (r - 1) \times k + 1 \tag{1}$$

对数据库中的动作序列进行运动目标检测，选取 $r = 2$ ， $k = 5$ ，即小结构元素尺寸为 2，膨胀 5 次，腐蚀 5 次，实验结果如图 3 所示，运动目标提取效果较好，同时除去了光照等因素对实验结果的影响。

3. 姿势表达

姿势表达方法建立在人体整体轮廓的基础之上。第 2 节中采用背景减除法可以提取出运动视频中的运动目标(如图 3 所示)，再利用边界跟踪算法可以提取出人体轮廓。利用轮廓点表示人体姿势的优势在于，去除人的轮廓内部的冗余点的同时，能够减少因小的视点或光照变化在前期形态学预处理过程产生的影响。本文借鉴 Dedeoglu et al.的思路[6]，以轮廓点为特征对姿势进行表达。其具体步骤如下：

Step 1 通过 8 邻域边界跟踪算法获得轮廓点集 $P = \{p_1, p_2, \dots, p_n\}$ 。如图 4 所示设当前点 $P(x, y)$ 在上边界点 C 的 8 邻域内的位置编码为 n ，则从当前点 (x, y) 的 8 邻域内的编码为 n 的位置，顺时针方向移动 2 个像素的位置就是下一边界点的起始搜索位置。若不是边界点，则从搜索的起始点开始按照逆时针方向顺次搜索，共搜索 5 次便可以找到下一个边界点。8 邻域边界跟踪算法执行步骤如下：设图像中背景点值为 0，目标点为 1， P_k 为第 k 个边界点， k 的初值为 0， t 表示边界终结点的个数。

- 1) 定目标区域标签值 $label$;

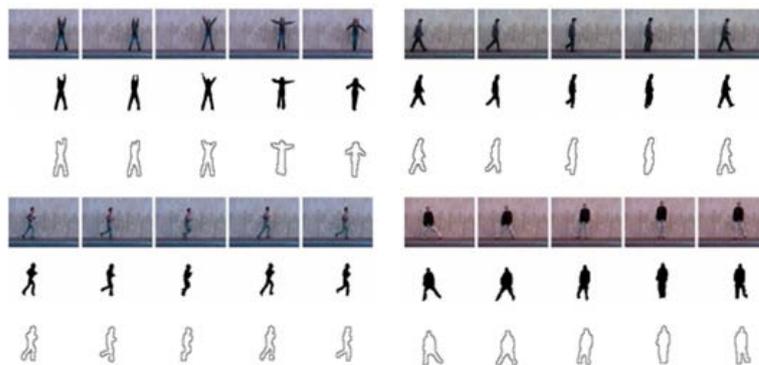


Figure 3. Results of motion detection and contour extraction (row 1 and 4 original image, row 2 and 5 motion object, and row 3 and 6 contour of object)

图 3. 运动目标检测及轮廓提取实验结果(第一、四行原始图像，第二、五行运动目标，第三、六行轮廓提取)

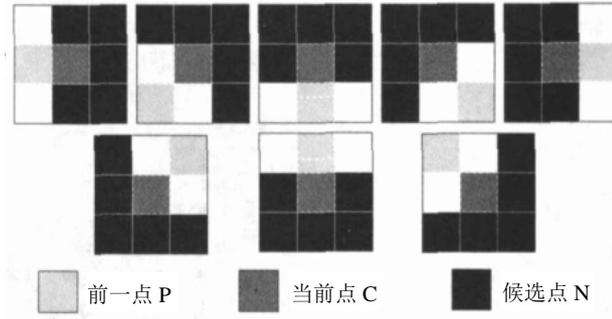


Figure 4. Position relationship between the formal point and the candidate points
图 4. 搜索的候选点与前一点的位置关系

- 2) 从左到右、自上而下按行扫描图像，找到第一个像素为 1 的点且标签值为 $label$ 的点即为边界的起始点 P_0 ，并将其坐标 (x, y) 存入边界点序列，预置 t 为 0，位置编码为 0；
- 3) 确定下一目标点搜索的起始位置，然后从该位置开始按顺时针方向依次检查当前边界点的 8 邻近像素，当第一次出现其像素点的标签值等于预定标签值时，这一像素点就是新的边界点 $P_k (k = k + 1)$ ，并记下它在 8 邻域中位置编码值；
- 4) 若新边界点 $P_k = P_0$ ，即回到了起始点，边界跟踪结束，此时边界点序列中存放的就是该目标的外边界点坐标，转(7)；
- 5) 若新边界点 $P_k \neq P_0$ ，则以 P_k 作为当前点，记下其位置编码，转③；
- 6) 若没有找到目标点，说明当前点是轮廓的终结点，终结点个数 t 加 1。如果 $t=1$ ，则令 P_0 为当前点，设位置编码为 4，转③；(即对不封闭轮廓从起始点开始反向搜索，直至找到另一个终结点为止)；如果 $t=2$ ，转⑦；
- 7) 若还需跟踪其他目标的轮廓，转②；否则算法结束。

Step 2 计算出轮廓点集的重心 $C_m = (x_c, y_c)$ ，其中 x_c, y_c 的计算公式如下：

$$x_c = \frac{\sum_{i=1}^n x_i}{n}, y_c = \frac{\sum_{i=1}^n y_i}{n} \quad (2)$$

其中 n 为选定轮廓点的个数。

Step 3 定义距离算子 DS 。距离算子是每个轮廓点到轮廓重心的欧几里得距离，为方便标记，规定以最左侧的点为初始点顺时针标记排序。

$$DS = \{d_1, d_2, \dots, d_n\}$$

$$d_i = \|C_m - p_i\|, \forall i \in \mathbb{N} \quad (3)$$

Step 4 通过固定距离算子的大小，对特征尺度进行二次采样成为连续的长度 ℓ ，同时对总体进行标准单位化处理，以达到实验过程中的尺度不变性要求。

$$D\hat{S}[j] = DS \left[j * \frac{n}{L} \right], \forall j \in [1, \dots, \ell]$$

$$D\bar{S}[j] = \frac{D\hat{S}[j]}{\sum_{j=1}^{\ell} D\hat{S}[j]}, \forall j \in [1, \dots, \ell] \quad (4)$$

虽然相较于其他相近的方法，利用该方法生成的信息更为广泛详实、数据占用空间较大，但该方法提取出的依旧是一个低维特征，处理过程中的计算量较小。

4. 关键特征提取

通过前期对行为视频的处理，我们得到视频序列中的人体动作的表达方式，要准确识别出人体动作，就需要提取出每种动作中的关键特征。不同于 Chaaaraoui 等人提出的采用 *K-means* 提取特征[11]，本文借鉴学者 Cheema et al. [7]和 Baysal et al. [8]提出的基于欧氏距离的 *K*-平均聚类算法对关键特征进行提取并给出相应的改进。

基于搜索算法的 *K*-平均聚类算法

由于开始选取任意 *K* 个点作为初始的聚类点再进行迭代操作，因此初始点的不同可能会影响到聚类结果，为聚类结果对初值的依赖性，同时提高算法稳定性，我们对 *K* 平均聚类算法进行优化改进，通过搜索算法得到一组较好的初始聚类中心[9]。算法步骤描述如下：

```

Input: n 个数据构成的集合  $X_i$  及  $K'(K' > K)$ ;
Output: K 个聚类中心  $Z_j$  以及 K 个聚类对象集合  $C_j$ ;
Begin
 $m = 1; S = \bigcup_{i=1}^j S_i$ ; //对 S 进行划分，划分为互不相交的 j 个子集//
K-means( $S_m, K'$ ); //执行 K-means 算法，产生  $K'$  个聚类中心//
Calculate  $J_c(m) = \sum_{j=1}^{K'} \sum_{X_i \in C_j} \|X_i - Z_j\|^2$ ;  $Z_j = \min \{J_c(m)\}$ ,  $j \in [1, K']$ ;
K-means( $S, K'$ ); //再次执行 K-means 算法//
Calculate  $D(Z_i, Z_j) = \min \{\|Z_i - Z_j\|\}$ ;
if  $D(Z_{i_k}, Z_{j_k}) = \min D(Z_i, Z_j)$  then  $Z_{i_k} = Z_{i_k} \cup Z_{j_k}$ ;
//合并距离最小的相邻测试组，每次合并后聚类中心减少 1 个//
Repeat
Until 聚类中心数减小为 K
End
    
```

通过基于搜索算法的 *K*-平均聚类算法，可以对数据库中的每个训练集进行单独的聚类分析，从而一族动作关键特征就被提取出来。如图 5(a)、(b)所示，从 Weizmann 数据库中我们提取出跑步动作在原始 *K-means* 聚类算法(图 5(a))和改进的基于搜索的 *K-means* 聚类算法(图 5(b))的动作关键特征。从图 5(a)、(b)可以看出，两种聚类算法提取的 *run* 动作关键特征有所不同，由改进 *K-means* 算法提取的关键特征序列较为明晰丰富，能够较好地代表跑步动作的关键特征序列，一定程度上避免了相似动作的关键特征序列之间的相近问题，对后期实验结果的准确性提供了技术支持。



Figure 5. Key features of “run” action: (a) Results of original *K-means* algorithm; (b) Results of improved *K-means* algorithm

图 5. *run* 动作关键特征: (a) 原始 *K-means*; (b) 改进 *K-means*

5. 人体动作识别

5.1. 基本思路

通过对一族运动动作视频的反复训练，可以得到一族运动动作的较为精准的关键特征动作。待测运动的动作序列为：

$$S_p = \{p_1, p_2, \dots, p_n\} \quad (5)$$

而关键特征序列为 $S_{key} = \{kp_1, kp_2, \dots, kp_n\}$ ，通过比较可以得到相邻关键特征动作，从而可以对动作进行识别。

5.2. 基于 DTW 动态时间规整算法的动作识别

为了减小组间方差效应的影响，需要寻找一个适合的距离矩阵来比较两个训练视频提取出的动作关键特征。由于人体轮廓和个体动作的差异性，对动作识别的精准度产生了较大的影响。而动态时间规整算法[10] (Dynamic Time Warping)在不调整现有时间序列的前提下，较好解决了序列比较中的时间尺度不一致性，而且能够将时间长短不一的动作序列排列整齐。在相同条件下，DTW 算法与相对常用的 HMM 算法，识别效果相差不大，但 HMM 算法要复杂得多，需要在训练阶段提供大量的训练集，通过反复计算才能得到参数模型，而 DTW 算法本身既简单又有效，DTW 算法的训练中几乎不需要额外的计算[10]。

已知训练序列动作关键特征集合和测试序列动作关键特征集合，可以计算得到 DTW 距离[11]为：

$$S_{train} - S_{test} = dtw(n, m) \quad (6)$$

$$dtw(i, j) = \min \begin{cases} dtw(i-1, j) \\ dtw(i, j-1) \\ dtw(i-1, j-1) \end{cases} + d(kp_i, kp'_j) \quad (7)$$

其中 $d(kp_i, kp'_j)$ 是欧氏距离函数，用以比较两关键动作特征的差异性。

6. 实验结果与分析

6.1. 实验中参数设定

在本文所提出的行为识别的方法中，实验的参数设定如下：

MIN_FRAMES = 5; //检测出动作区域的最小帧数

KMeansIterations = 6; //K 平均聚类算法的最佳迭代次数

Delta = 5; //累计达到后开始重新检测的最小帧数

Gamma = 10; //最优步长迭代

HistoryCount = 22; //高斯平滑算法用到的视频帧数

EPSILON = 1e-6; //聚类中心误差界

ITERATIONS = 1000; //K 平均聚类算法最大循环次数

INIT_ITER = 100; //K 平均聚类算法最大初始循环次数

Lameda = 3; //聚类数

6.2. 实验结果

为了论证本文提出方法的行为识别性能，在 Weizmann 数据集、KTH 数据集和 MuHAVi 数据集进行了实验和分析。

6.2.1. Weizmann 数据集上结果

在 Weizmann 数据集通过采取留一交叉校验法进行数据集的训练。Weizmann 是由 9 个人分别执行 10 个不同的动作(*bend, jack, jump, pjump, run, side, skip, walk, wave1, wave2*)构成的数据集[12], 视频背景、视角都是静止的。多数学者在实验过程中都会舍弃 *skip* 这一动作的识别, 因为 *skip* 动作的训练效果不佳, 同时会影响其他动作的判别。图 6 是在剔除 *skip* 动作后的训练集判别分析矩阵, 在实验中的平均正确率为 91.36%。从图 6 可以发现, *run* 与 *walk*、*jack* 和 *wave2* 存在互相干扰。这是因为 *run* 与 *walk* 的动作关键特征较为相近, 导致判别过程中关键动作的差异性较小。而 *jack* 与 *wave2* 动作中手部动作相同, 从而对判别分析产生一定影响。从图 7 可知在添加 *skip* 后再次进行训练, 实验平均正确率减少为 85.87%, 此时第 3 节 Step 4 中 $\ell = 200$, $K' = 120$ 。可以发现使用本文方法后, 虽然 *skip* 的判别正确率较高, 但仍对其他有相似关键动作特征的动作判别产生一定干扰, 如 *jump*、*pjump*、*walk*、*run*、*side*。

6.2.3. KTH 数据集上结果

KTH 数据集包括由 25 个不同的人分别在四个场景下执行的 6 类行为(*walking, jogging, running, boxing, hand waving, hand clapping*), 一共有 599 段视频。视频背景相对静止, 除了镜头的拉近拉远, 摄像机的运动比较轻微。图 8 给出了实验中得到的在 KTH 数据库上的识别判别分析矩阵。此时 $\ell = 200$, $K' = 120$, 平均正确率达到 94.47%。

6.2.3. MuHAVi 数据集上结果

MuHAVi 数据库是行为识别领域较新的多视角数据集之一。包括由 7 个不同的人分别在室内执行的 14 类行为(*CollapseLeft, CollapseRight, GuardToKick, GuardToPunch, KickRight, PunchRight, RunLeftToRight,*

	bend	jack	jump	pjump	run	side	walk	wave1	wave2
bend	9/9								
jack		9/9							
jump			9/9						
pjump				9/9					
run					8/10	2/10			
side			2/9			7/9			
walk							10/10		
wave1		1/9						8/9	
wave2		2/9							7/9

Figure 6. Confusion matrix of the Weizmann dataset without the skip action. $\ell = 200$, $K' = 120$, average accuracy is 91.36%

图 6. 动作判别分析矩阵(Weizmann 数据库不含 skip 动作)
 $\ell = 200$, $K' = 120$, 平均正确率 = 91.36%

	bend	jack	jump	pjump	run	side	skip	walk	wave1	wave2
bend	9/9									
jack		9/9								
jump			8/9					1/9		
pjump				8/9						
run					8/10	2/10				
side			1/9			8/9				
skip			1/9	1/9			6/9	1/9		
walk							2/10	8/10		
wave1		1/9							8/9	
wave2		2/9								7/9

Figure 7. Confusion matrix of the Weizmann dataset with the skip action. $\ell = 200$, $K' = 120$, average accuracy is 85.87%

图 7. 动作判别分析矩阵(Weizmann 数据库含 skip 动作)
 $\ell = 200$, $K' = 120$, 平均正确率 = 85.87%

Run Right To Left, Standup Left, Standup Right, Turn Back Left, Turn Back Right, Walk Left To Right, Walk Right To Left)。由本文方法检测多视角的 *MuHAVi* 数据库, 得到如图 9 所示的判别分析矩阵。此时 $\ell = 340$, $K' = 136$, 平均正确率为 92.10%。

6.3. 现有方法精度比较

为了与已有算法进行比较, 对 Singh et al. (2010)、Martinez-Contreras et al. (2009)、Cheema et al. (2011)、Chaaroui et al. (2013) 的算法进行了测试, 得到了四种算法在 *Weizmann* 数据集、*KTH* 数据集、*HuHAVi* 数据集上的平均正确率, 如表 2 所示。从表 2 可以发现, 在 Chaaroui et al. (2013) 方法的基础之上进行的基于搜索 *K-means* 聚类算法优化, 平均正确率有较为明显的提高。

6.4. 后期改进

实验中发现, *Weizmann* 数据集[12]中个体运动的速度不同, 也有人体自遮挡问题等, 这些对识别算

	walking	jogging	running	boxing	handwaving	handclapping
walking	0.97	0.03				
jogging	0.03	0.94	0.03			
running	0.02	0.11	0.87			
boxing	0.01			0.99		
handwaving				0.07	0.93	
handclapping				0.07		0.93

Figure 8. Confusion matrix of the *KTH* dataset. $\ell = 200$, $K' = 120$, average accuracy is 94.47%

图 8. 动作判别分析矩阵(*KTH*). $\ell = 200$, $K' = 120$, 平均正确率 = 94.47%

	CollapseLeft	CollapseRight	GuardToKick	GuardToPunch	KickRight	PunchRight	RunLeftToRight	RunRightToLeft	StandupLeft	StandupRight	TurnBackLeft	TurnBackRight	WalkLeftToRight	WalkRightToLeft
CollapseLeft	4/4													
CollapseRight		4/4												
GuardToKick			7/8	1/8										
GuardToPunch				8/8										
KickRight					8/8									
PunchRight						8/8								
RunLeftToRight							3/4	1/4						
RunRightToLeft								4/4						
StandupLeft									1/2	1/2				
StandupRight										4/4				
TurnBackLeft			1/2								1/2			
TurnBackRight			1/4									3/4		
WalkLeftToRight													4/4	
WalkRightToLeft														4/4

Figure 9. Confusion matrix of the *MuHAVi* dataset. $\ell = 340$, $K' = 136$, average accuracy is 92.10%

图 9. 动作判别分析矩阵(*MuHAVi*) $\ell = 340$, $K' = 136$, 平均正确率 = 92.10%

Table 2. Comparison of average success rate between some algorithms
表 2. 各算法平均正确率对比表

	<i>Weizmann</i>	<i>KTH</i>	<i>HuHAVi</i>
Singh et al. (2010)	90.4%	90.4%	82.4%
Martinez-Contreras et al. (2009)	89.3%	87.3%	81.8%
Cheema et al. (2011)	91.6%	89.6%	86.0%
Chaarouai et al. (2013)	92.8%	93.3%	91.4%
Our Method	91.4%	94.5%	92.1%

法在实际生活中的应用提供了改进和优化的思路。由于 *Weizmann* 数据库使用已经较为广泛，识别效果较好，现实生活中的多维摄像头问题也为识别精度的提高提出新挑战。实验后期将对更加复杂的 *IXMAX* 多视角数据库问题进行研究，同时改进优化算法，以提高算法的稳定性和准确率，增强算法实用性。

致 谢

本课题受“秦惠着和李政道中国大学生见习进修基金”资助。感谢秦惠着女士、李政道先生、亚洲基督教高等教育联合管理委员会及苏州大学相关机构对本课题的支持。感谢导师龚声蓉教授、刘纯平教授的悉心指导。

基金项目

本课题受“秦惠着和李政道中国大学生见习进修基金”资助。

参考文献 (References)

- [1] 杜友田, 陈峰, 徐文立, 李永彬 (2007) 基于视觉的人的运动识别综述. *电子学报*, **1**, 84-90.
- [2] 李瑞峰, 王亮亮, 王珂, 等 (2014) 人体动作行为识别研究综述. *模式识别与人工智能*, **1**, 35-48.
- [3] 李莉莉 (2010) 基于视频的指尖检测与跟踪算法及实现. 内蒙古大学, 呼和浩特.
- [4] 王忠礼 (2008) 智能交通系统车辆检测算法研究. 哈尔滨工业大学, 哈尔滨.
- [5] 金慧 (2005) 基于多视图图像的三维建模方法研究. 南京理工大学, 南京.
- [6] Dedeoglu, Y., Toreyin, B., Gudukbay, U. and Cetin, A. (2006) Silhouette-based method for object classification and human action recognition in video. In: Huang, T., Sebe, N., Lew, M., Pavlovic, V., Kolsch, M., Galata, A. and Kisacanin, B., Eds., *Computer Vision in Human-Computer Interaction*, Springer, Berlin, Heidelberg, Volume 3979 of Lecture Notes in Computer Science, 64-77.
- [7] Cheema, S., Eweiwi, A., Thureau, C. and Bauckhage, C. (2011) Action recognition by learning discriminative key poses. *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, Barcelona, 6-13 November 2011, 1302-1309.
- [8] Baysal, S., Kurt, M. and Duygulu, P. (2010) Recognizing human actions using key poses. *20th International Conference on Pattern Recognition (ICPR)*, Istanbul, 23-26 August 2010, 1727-1730.
- [9] 严勇 (2007) 数据挖掘中聚类分析算法研究与应用. 电子科技大学, 成都.
- [10] 李邵梅, 刘力雄, 陈鸿昶, 等 (2008) 实时说话人辨识系统中改进的 DTW 算法. *计算机工程*, **4**, 218-219.
- [11] Chaarouai, A.A., Climent-Pérez, P. and Flórez-Revuelta, F. (2013) Silhouette-based human action recognition using sequences of key poses. *Pattern Recognition Letters*, **34**, 1799-1807.
- [12] Blank, M., Gorelick, L., Shechtman, E., Irani, M. and Basri, R. (2005) Actions as space-time shapes. *Tenth IEEE International Conference on Computer Vision*, Beijing, 17-21 October 2005, Vol. 2, 1395-1402.

汉斯出版社为全球科研工作者搭建开放的网络学术中文交流平台。自2011年创办以来，汉斯一直保持着稳健快速发展。随着国内外知名高校学者的陆续加入，汉斯电子期刊已被450多所大中华地区高校图书馆的电子资源采用，并被中国知网全文收录，被学术界广为认同。

汉斯出版社是国内开源（Open Access）电子期刊模式的先行者，其创办的所有期刊全部开放阅读，即读者可以通过互联网免费获取期刊内容，在非商业性使用的前提下，读者不支付任何费用就可引用、复制、传播期刊的部分或全部内容。

