

# 基于动作串的人体行为识别

赵海勇 李俊青

(聊城大学计算机学院 聊城 252000)

**摘要** 提出了一种以运动人体侧影为特征的基于模板匹配的人体行为识别方法。首先,利用背景差分法和阴影消除技术提取运动人体侧影。利用 $\mathcal{R}$ 变换对人体侧影进行特征提取,将时变的2D区域形状转换为对应的1D距离向量。然后,利用谱系聚类方法提取动作序列的关键姿态,将关键姿态编码为称为动作串的模板。最后,利用动态时间规整算法度量测试序列与标准模板之间的相似性。实验结果表明,本方法对人的6种日常行为进行识别的正确识别率达到85%以上,具有简单实用的特点。

**关键词** 计算机视觉,人体侧影,关键姿态,聚类,行为识别

中图分类号 TP391.4 文献标识码 A

## Human Action Recognition Using Action String

ZHAO Hai-yong LI Jun-qing

(School of Computer Science and Technology, Liaocheng University, Liaocheng 252000, China)

**Abstract** A template match-based method was presented by using the silhouettes of image sequence as representative descriptors of human posture to achieve human actions recognition. First, human silhouette extraction was obtained by background subtraction and shadow elimination. A new silhouette descriptor based on  $\mathcal{R}$  transform was defined, which is employed to transform the silhouette into a 1D distance signal. Hierarchical clustering method was proposed to extract key postures of human action. Then, key postures were coded as a template called action string. At last, dynamic time warping was proposed to measure the similarity of the templates and the test sequences. This algorithm was evaluated using six human daily actions. The experimental results show that this method can achieve the correct recognition rate above 85%.

**Keywords** Computer vision, Silhouette, Key posture, Clustering, Action recognition

## 1 引言

近年来,人体行为识别已经成为计算机视觉与模式识别领域的研究热点<sup>[1-6]</sup>。它在智能视频监控、基于内容的视频检索、体育赛事分析、人机交互和虚拟现实等方面具有广泛的应用前景和潜在经济价值。然而,真实环境的多样性、行为序列中人体姿态的非刚性特点以及研究人员对行为类别的模糊定义,使得人体行为识别成为一个非常具有挑战性的研究领域。

人体行为识别最基本的两个问题是行为的描述与识别。行为描述方法按照特征的性质可分为两类<sup>[3]</sup>:一类是基于低层图像信息的方法,一类是基于高层人体结构的方法。由于低层图像信息特征获取简单,因此,该类算法成为行为描述的主要方法<sup>[7]</sup>。可利用的低层图像信息主要包括:前景目标、前景目标的运动速度、运动轨迹信息<sup>[8]</sup>、光流、前景目标的轮廓和侧影等等。人的高层结构信息是指人身体结构所呈现的姿态,与低层图像信息相比,它可以更精细地描述人的行为。根据提取特征过程中利用的人体模型不同,这类描述行

为的算法可分为3种:基于人体点模型的方法、基于2维人体模型的方法和基于3维人体模型的方法。基于高层人体结构特征的方法通常参数多,训练复杂,计算量较大。常用行为识别方法有两种:一是模板匹配法,一是状态空间法。在行为识别中,基于模板匹配的算法可以分为帧对帧匹配方法和融合匹配方法。帧对帧匹配方法是指直接用测试集的行为特征序列与参考集的行为特征序列进行逐帧匹配,而融合匹配方法是指先将整个行为过程融合为一个整体模板或者若干个固定数目的模板,然后再利用这有限个模板进行匹配。基于模板匹配的行为识别算法计算简单,但是对行为的执行时间比较敏感,从而造成相同的行为由于执行时间长短不同而识别结果不同。基于状态空间的行为识别算法可以避免对时间间隔建模的问题,但是需要的训练样本大,计算复杂度较高。

本文提出了一种以运动人体的侧影为特征、基于融合匹配的人体行为识别方法。两类基于形状的特征人体侧影和轮廓能够鲁棒性地获得,因此经常被用于运动人体行为特征提取。基于人体轮廓的方法只考虑人体形状边界所包含的像

到稿日期:2012-12-14 返修日期:2013-04-19 本文受国家自然科学基金(61104179),山东省高校智能信息处理与网络安全重点实验室资助。

赵海勇(1981—),男,博士,讲师,主要研究方向为计算机视觉, E-mail:zhaohaiyongym@163.com;李俊青(1976—),男,副教授,主要研究方向为智能信息处理。

素,因而计算复杂度较低,但是由于需要准确提取人体轮廓,限制了其应用范围。基于人体侧影的方法应用范围广泛,但是由于将整个体形状所包含的像素考虑在内,导致计算复杂度较高。 $\mathcal{R}$ 变换<sup>[9,10]</sup>是对 Radon 变换的一种改进,通过改进使其近似满足了几何不变性(平移不变性、旋转不变性和尺度不变性),因而更适合人体行为特征的提取与描述。这种描述方法计算简单,而且对人体侧影的全局特征能进行很好的描述。当提取的运动人体侧影不连续或者存在空洞时, $\mathcal{R}$ 变换仍然具有较好的鲁棒性。但是  $\mathcal{R}$ 变换所提取的形状向量维数较高,需要进行压缩,本文采用主成分分析(PCA)进行行为特征的压缩。借助形状特征,人体行为可以看作按时间顺序串连起来的静态姿态图像序列之间的形状时变过程<sup>[11]</sup>,通过观察其中的一些关键姿态就可很好地识别一种行为<sup>[12]</sup>。文献[13]提出了一种假设:“每个动作可以由3个关键的姿势来描述,这些关键姿势分布在一个完整运动周期的开始、中间和结尾阶段,其余姿势可以认为是这些关键姿势的过渡。与文献[13]不同,本文采用非监督学习方法<sup>[14]</sup>分析人体行为。首先利用谱系聚类方法,自动提取行为的关键姿态集,然后将行为的关键姿态建模为动作串,最后采用动态时间规整方法(DTW)实现动作串的匹配与识别。

## 2 人体姿态描述

人体姿态描述包括两个部分:(1)人体侧影提取;(2)人体侧影特征描述。运动人体的检测与分割是人体行为识别的第一步(算法流程图见图1),也是最重要的一步,因为识别结果很大程度取决于检测与分割的质量。为了尽量消除遮挡、阴影、光照变化以及噪声等因素给检测与分割所造成的影响,本文做了预处理与后期处理。

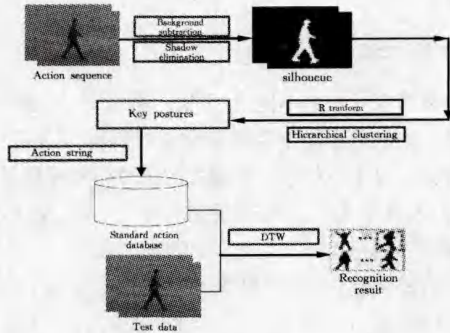


图1 算法流程图

### 2.1 人体侧影提取

假定包含人体行为的数据采集环境为:1)摄像机静止不动;2)摄像机方向与人体运动方向近似垂直;3)环境中只有一个人在运动。人体轮廓提取步骤如下:

(1)图像去噪:视频序列在获取和传输时常常会受到各种各样的噪声干扰,例如高斯白噪声、脉冲噪声和乘性噪声等。为了改善图像质量以及便于各种后续处理,本算法采用高斯滤波去噪和邻域去噪相结合的方法。

(2)目标检测与阴影去除:采用背景差分法检测视频序列中的运动人体。首先建立一个随时间和空间变化而自适应变化的背景模型<sup>[15]</sup>,然后将当前帧与背景模型进行差分,设定

阈值,二值化分割图像得到前景和背景。目标检测的目的是从背景图像中分割提取出运动变化区域,因此,在检测运动目标的同时,运动目标投射的阴影也会作为运动前景的一部分。本文采用文献[16]提出的算法对前景/阴影进行分割,该方法适用于室内户外等多种环境。

(3)轮廓提取:采用基于连通性的边界跟踪算法获取提取整个图像序列中行人的侧影。图2给出了其中一帧的人体侧影和轮廓提取结果。

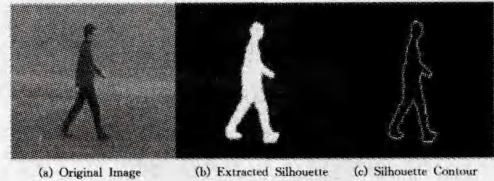


图2 人体检测及侧影提取图

### 2.2 人体侧影特征描述

$\mathcal{R}$ 变换是对 Radon 变换的一种改进,通过改进使其近似满足了几何不变性(平移不变性、旋转不变性和尺度不变性),因而更适合人体行为特征的提取与描述。Radon 变换是计算图像在某一特定角度射线方向上投影的变换方法,对一幅离散二值图像  $f(x, y)$ , Radon 变换定义(见图3)如下。

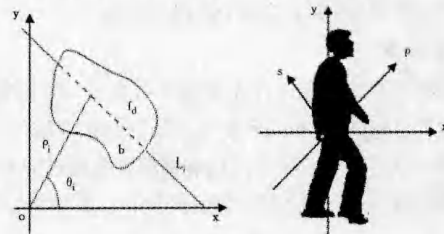


图3 Radon 变换示意图

$$T_{Rf}(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy = R\{f(x, y)\} \quad (1)$$

式中,  $\theta \in [0, \pi]$ ,  $\rho \in [-\infty, \infty]$ ,  $\delta(\cdot)$  为 dirac-delta 函数,

$$\delta(x) = \begin{cases} 1, & \text{if } x=0 \\ 0, & \text{else} \end{cases} \quad (2)$$

通过对式(1)计算分析可知, Radon 变换不满足几何不变性。为了解决上述问题,文献[9]提出了  $\mathcal{R}$ 变换,其定义如下:

$$\mathcal{R}_f(\theta) = \int_{-\infty}^{\infty} T_{Rf}^2(\rho, \theta) d\rho \quad (3)$$

$\mathcal{R}$ 变换本身所具有的一些属性使其适合于行为特征表示。

**属性1** 当图像  $f(x, y)$  平移向量  $\vec{\mu} = (x_0, y_0)$  时,

$$\int_{-\infty}^{\infty} T_{Rf}^2((\rho - x_0 \cos \theta - y_0 \sin \theta), \theta) d\rho = \int_{-\infty}^{\infty} T_{Rf}^2(\nu, \theta) d\rho = \mathcal{R}_f(\theta) \quad (4)$$

由式(4)可以看出,平移图像并不会改变  $\mathcal{R}$ 变换的结果。

**属性2** 改变图像  $f(x, y)$  的尺度,因子为  $\alpha$ ,

$$\frac{1}{\alpha^2} \int_{-\infty}^{\infty} T_{Rf}^2(\alpha \rho, \theta) d\rho = \frac{1}{\alpha^2} \int_{-\infty}^{\infty} T_{Rf}^2(\nu, \theta) d\rho = \frac{1}{\alpha^2} \mathcal{R}_f(\theta) \quad (5)$$

改变图像的尺度,会引起  $\mathcal{R}$ 变换结果振幅的改变,因此,

需要对图像进行标准化。

属性3 当图像旋转角度  $\pi$  时,

$$\begin{aligned} \int_{-\infty}^{\infty} T_{kf}(-\rho, \theta \pm \pi) d\rho &= - \int_{-\infty}^{\infty} T_{kf}(\nu, \theta \pm \pi) d\nu \\ &= \int_{-\infty}^{\infty} T_{kf}(\nu, \theta \pm \pi) d\nu = \mathcal{R}_f(\theta \pm \pi) \end{aligned} \quad (6)$$

由式(6)可以看出,  $\mathcal{R}$  变换的周期为  $\pi$ , 因此, 一个 180 度的向量就能充分描述人体侧影区域。图 4 给出了 6 种人体行为为姿态的  $\mathcal{R}$  变换结果。

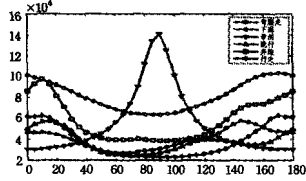


图 4 6 种行为  $\mathcal{R}$  变换结果

### 3 动作串构建

动作串的构建步骤主要分为两步: 步骤 1, 采用谱系聚类对图像帧集合进行聚类, 以获得连续运动的离散表示, 每一关键帧用一个字符(动作串中的字符)表示。步骤 2, 根据不同姿态的重要性, 对相应字符赋予不同的权重。

#### 3.1 谱系聚类

谱系聚类算法作为一种无监督学习算法, 其原理是首先视每个样本自成一类, 利用各样本之间的相似性进行合并, 以减少类别的数目。谱系聚类算法能够根据数据内部的特点获取最优聚类数目, 实现数据的快速离散化。具体算法描述如下:

步骤 1 假设  $A = \{X_1, X_2, \dots, X_N\}$  为待聚类的样本集, 构造  $N$  个类, 每个类包含一个样本。计算两两之间的距离, 构成一个对称距离矩阵  $D_{(0)}$ :

$$D_{(0)} = \begin{bmatrix} 0 & d_{12} & \dots & d_{1n} \\ d_{21} & 0 & \dots & d_{2n} \\ & & \ddots & \\ d_{n1} & d_{n2} & \dots & 0 \end{bmatrix} \quad (7)$$

步骤 2 选择  $D_{(0)}$  中主对角线以下(或者以上)的最小元素, 设该元素为  $d_{pq}$ , 如果  $d_{pq}$  大于阈值  $\tau$ , 则聚类过程停止, 转步骤 5, 否则继续步骤 3。

步骤 3 将类  $X_p$  和  $X_q$  合成为一个新类  $X_r$ ,  $X_r = \{X_p, X_q\}$ 。在  $D_{(0)}$  中消去类  $X_p$  和  $X_q$  所在的行和列, 并加入由新类  $X_r$  与剩下的其他未聚合的类间的距离所组成的一行和一列, 得到一个更新的距离矩阵  $D_{(1)}$ , 它是  $n-1$  阶方阵。

步骤 4 从  $D_{(1)}$  出发重复步骤 2。

步骤 5 最终得到的聚类集  $\{X_1, X_2, \dots, X_M\}$ , 取每个聚类的中心帧作为关键帧, 进而得到行为的关键姿态集。

#### 3.2 构建行为动作串

将关键姿态集中的每一个关键姿态用一个字符表示, 这样行为就可表示为一个动作串。如(见图 5)“拍手”行为可以表示为: “sabcdcbe”, 其中 ‘s’ 代表字符串的开始, ‘e’ 代表字符串的结束。

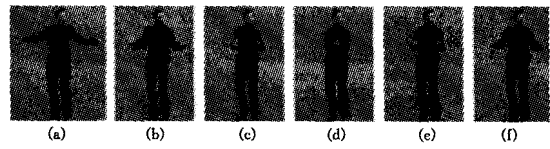


图 5 拍手行为字符串表示

在构建动作串时, 由于各关键姿态在识别过程中的重要性不相同, 需要赋予不同的权重。用二元组  $(p, \omega)$  表示一个赋予权重的字符, 其中  $p$  表示不同的姿态,  $\omega$  表示权重, 计算如下:

$$\omega_i = \frac{\text{sum}(X_i)}{\text{sum}(A)} \quad (8)$$

式中,  $\text{sum}(X_i)$  表示类  $X_i$  包含帧的帧数,  $\text{sum}(A)$  表示每种行为的总帧数。

赋权之后, 动作串可以表示为:  $s = \{P_1, P_2, \dots, P_n\}$ ,  $P_i = (p_i, \omega_i)$ ,  $i = 1, 2, \dots, n$ ,  $n$  表示关键帧数。假设有两个动作串  $s^1$  和  $s^2$ :  $s^1 = \{a_1^1, a_2^1, \dots, a_{n_1}^1\}$ ,  $a_i^1 = (p_i^1, \omega_i^1)$ ,  $i = 1, \dots, n_1$ ,  $s^2 = \{a_1^2, a_2^2, \dots, a_{n_2}^2\}$ ,  $a_i^2 = (p_i^2, \omega_i^2)$ ,  $i = 1, \dots, n_2$ , 它们之间的距离定义为将  $s^1$  转换成  $s^2$  所需要的操作成本之和:

$$D(s^1, s^2) = \sum_{k=1, \dots, K} \text{Cost}_k \quad (9)$$

式中,  $K$  是操作数量,  $\text{Cost}_k$  表示第  $k$  次操作所需要的成本。在转换过程中, 共有 3 类操作: 替换、插入和删除, 成本定义如下:

$$\text{Cost}^R(a^1, a^2) = |\omega^1 - \omega^2| + \min\{\omega^1, \omega^2\} \quad (10)$$

$$\text{Cost}^I(a^1, a^2) = \omega^2 \quad (11)$$

$$\text{Cost}^D(a^1, a^2) = \omega^1 \quad (12)$$

式中,  $\text{Cost}^R(a^1, a^2)$  表示将  $s^1$  中的  $a_1$  替换为动作串  $s^2$  中的  $a_2$  所需要的成本,  $\text{Cost}^I(a^1, a^2)$  表示将  $a_2$  插入  $s^1$  所需要的成本,  $\text{Cost}^D(a^1, a^2)$  表示当比较  $a_1$  和  $a_2$  时, 从  $s^1$  中删除  $a_1$  所需要的成本。

### 4 动作串识别

由于环境或习惯的不同, 对于同一种行为, 构建的运动串也会不同。如果采用最近距离算法进行匹配, 往往会出现运动抖动现象。由于动作串在构建过程中较好地保持了序列的时间尺度, 因此采用动态时间规整算法<sup>[17]</sup> (DTW) 可以较好地完成测试序列与参考序列之间的模式匹配。

假设有模板动作串  $s^1$  和待识别动作串  $s^2$ :  $s^1 = \{a_1^1, a_2^1, \dots, a_n^1\}$ ,  $a_i^1 = (p_i^1, \omega_i^1)$ ,  $i = 1, \dots, n$ ,  $s^2 = \{a_1^2, a_2^2, \dots, a_m^2\}$ ,  $a_j^2 = (p_j^2, \omega_j^2)$ ,  $j = 1, \dots, m$ , DTW 算法就是要寻找一个最佳的时间规整函数, 使待识别动作串的时间轴非线性地映射到模板动作串的时间轴上, 同时使总的累计失真量最小。将规整路径定义为:  $P = \{p_1, p_2, \dots, p_k\}$ , 其中  $p_k = (a_i^1, a_j^2)_k$ ;  $\max(m, n) \leq K \leq m+n$ 。DTW 可以表示为如下目标函数:

$$J = \frac{1}{\sum_{k=1}^K w_k} \min \left\{ \sum_{k=1}^K w_k d(a_{i_k}^1, a_{j_k}^2) \right\} \quad (13)$$

$$p^* = \text{argmin} \left\{ \sum_{k=1}^K w_k d(a_{i_k}^1, a_{j_k}^2) \right\} \quad (14)$$

式(13)和式(14)中,  $d(a_{i_k}^1, a_{j_k}^2)$  是两个动作串中  $a_{i_k}^1$  与  $a_{j_k}^2$  之间的局部距离,  $d(a_{i_k}^1, a_{j_k}^2) = |a_{i_k}^1 - a_{j_k}^2|$ ;  $w_k$  为加权因子, 对局部距离加权, 从而区别不同局部距离的重要性, 计算如下:

$$w_k = (a_{i_k}^1 - a_{i_{k-1}}^1) + (a_{j_k}^2 - a_{j_{k-1}}^2) \quad (15)$$

项  $\sum_{k=1}^K w_k$  对规整路径的不同长度加以补偿,从而使两个动作具有可比性。

## 5 实验结果与分析

采用 Blank 等人<sup>[18]</sup>的数据库对算法进行验证。该数据库共包含 90 个低分辨率的视频序列(180×144 像素,50 帧/秒),所有动作分别由 9 名实验人员完成,每人完成 10 个自然动作,本文对其中的“run”、“walk”、“skip”、“bend”、“jump”及“sideways”等进行试验。将 9 人所完成的动作序列分成 3 组,每组序列由 3 人完成,其中一组序列用于训练,建立标准行为数据库,其余两组用于测试。

### (1) 人体轮廓提取

利用轮廓提取方法,提取行为序列的所有人体轮廓,图 6 给出了示例图像及其相应的人体轮廓。

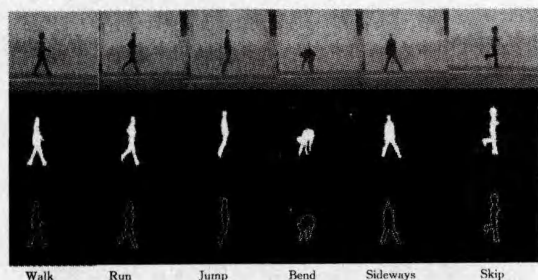


图 6 数据库的示例图像及对应的人体剪影和轮廓

### (2) $\mathcal{R}$ 变换属性分析

$\mathcal{R}$  变换本身所具有的一些属性使其适合于行为特征的表示,图 7 给出了行走在各种操作下的  $\mathcal{R}$  变换结果。

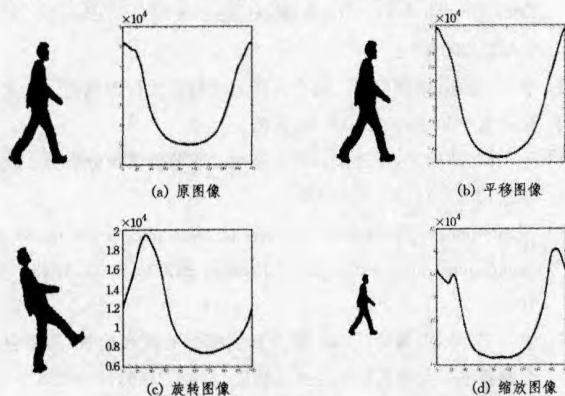


图 7 行走为姿态在各种数据情况下的  $\mathcal{R}$  变换结果

由图 7 可以看出,平移图像并不会改变  $\mathcal{R}$  变换的结果;旋转图像会使  $\mathcal{R}$  变换结果发生较大改变,但是现实中旋转图像的情况很少发生;改变图像的尺度,会引起  $\mathcal{R}$  变换结果振幅的改变,因此,需要对图像进行标准化。

### (3) $\mathcal{R}$ 变换鲁棒性分析

$\mathcal{R}$  变换是用于描述人体侧影区域的描述子,因此,侧影提取的质量直接影响着  $\mathcal{R}$  变换的结果。在理想情况下,将提取的侧影数据称为理想数据(Ground Truth Data)。如果侧影存在空洞或者不连续,则称之为残缺数据(Incomplete Data)。阴影和噪声会造成提取的侧影面积扩大,称这样的侧影为冗余数据(Redundant Data)。将视频序列有关键帧缺失时的数据称为关键帧缺失数据。为了测试  $\mathcal{R}$  变换的鲁棒性,图 8 给出了行走为姿态在各种情况下  $\mathcal{R}$  变换的结果。

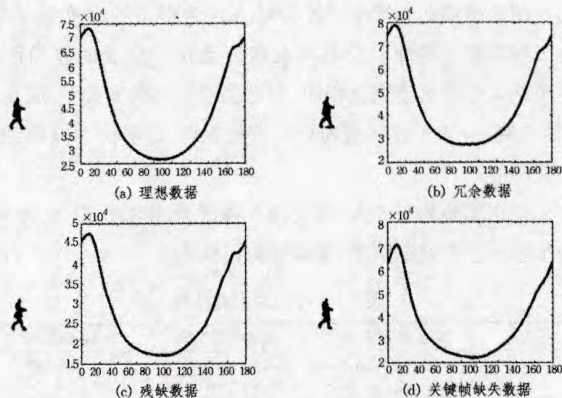


图 8 行走为姿态在各种数据情况下的  $\mathcal{R}$  变换结果

仿真结果显示在数据丢失或者不连续的情况下, $\mathcal{R}$  变换结果变化不大,但是在数据冗余的情况下,对变换结果影响较大。

### (4) 行为识别结果

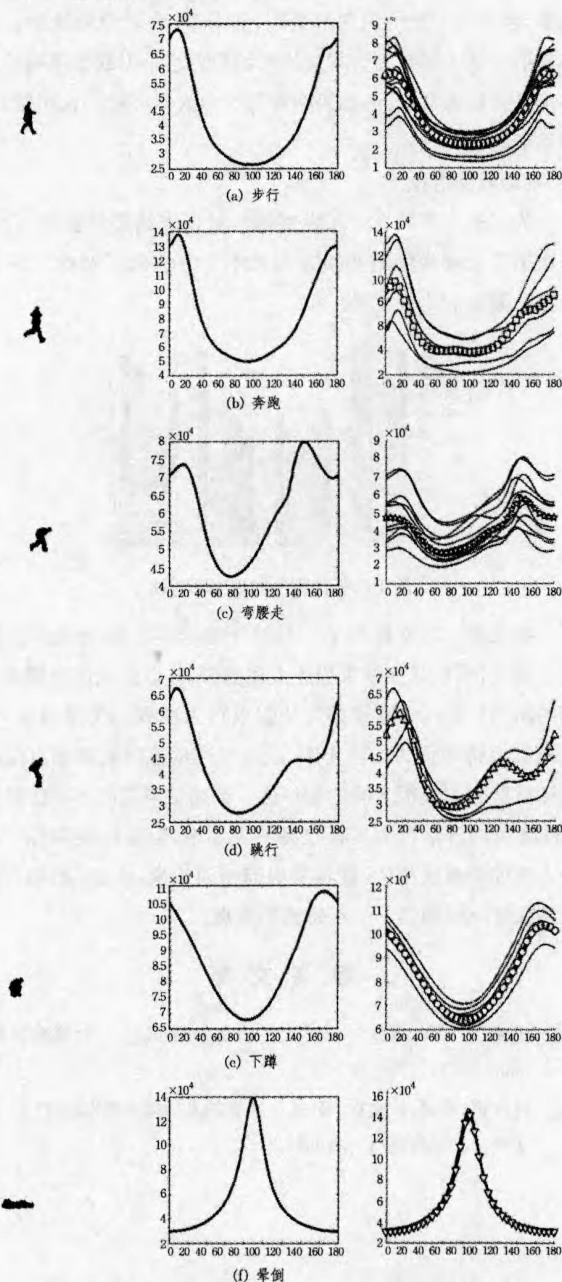


图 9 6 种行为姿态的  $\mathcal{R}$  变换结果

图9中第2列给出了6种单人行姿态的 $\mathcal{R}$ 变换结果。由于环境的多样性以及不同人执行动作所造成的差异性, $\mathcal{R}$ 变换结果可能不会完全相同,但是曲线的形状大致相同(见图9第3列)。为了使 $\mathcal{R}$ 变换更具有代表性,对多次 $\mathcal{R}$ 变换结果取平均值。

识别实验共有6人,每人做6种行为,实验次数为40次。表1给出了正确识别率、误认率和拒认率。

表1 6种行为识别结果

	正确识别率	误认率(FAR)	拒认率(FRR)
walk	34/40=85%	4/40=10%	6/40=15%
run	35/40=87.5%	5/40=12.5%	5/40=12.5%
jump	39/40=97.5%	3/40=12.4%	1/40=2.5%
bend	40/40=100%	0	0
sideways	38/40=95%	3/40=12.5%	2/40=5%
skip	36/40=90%	4/40=10%	4/40=10%

总体来说本算法对于6种简单的日常行为具有较高的识别率,表中对弯腰的识别率最高,而对走、跑和单腿跳的识别率较低。这主要是因为本文所用识别方法是以动作序列的人体形状特征为基础的,如果两种行为中人体姿势形状相似,就容易引起错误识别。

### (5)算法对比

为了测试不同算法和模型结构对识别结果的影响,在相同数据库上对常用矩变换方法和本文方法做了对比。图10给出了算法的比较结果。

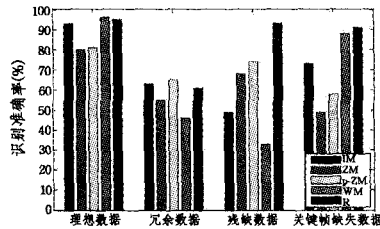


图10 5种算法的识别率对比

**结束语** 本文提出了一种基于模板匹配的行为识别方法。首先,采用 $\mathcal{R}$ 变换来描述人体侧影,并以此来度量侧影之间的距离。利用谱系聚类方法提取行为序列的关键姿态,在此基础上构建行为的动作串,最后采用动态时间规划算法度量测试序列与模板之间的相似性。试验结果表明本算法对于6种简单的日常行为具有较高的识别率,但是如果两种行为中人体姿势形状相似,就容易引起错误识别,因此,如何有效地解决这一问题将是下一步研究的重点。

### 参考文献

- [1] 王亮,胡卫明,谭铁牛. 人运动的视觉分析综述[J]. 计算机学报, 2002, 25(3): 225-237
- [2] 杜友田,陈峰,徐文立,等. 基于视觉的人的运动识别综述[J]. 电子学报, 2007, 35(1): 84-90
- [3] 徐光祐,曹媛媛. 动作识别与行为理解综述[J]. 中国图象图形学报, 2009, 14(2): 189-195
- [4] Aggarwal J K, Cai Q. Human motion analysis: A review[J]. Computer Vision and Image Understanding, 1999, 73(3): 428-440
- [5] Poppe R. A survey on vision-based human action recognition [J]. Image and Vision Computing, 2010, 28(6): 976-990
- [6] Candamo J, Shreve M, Goldof D B, et al. Understanding Transit Scenes: A Survey on Human Behavior-Recognition Algorithms [J]. IEEE Transactions on Intelligent Transportation Systems, 2010, 11(1): 206-224
- [7] 谷军霞,丁晓青,王生进. 行为分析算法综述[J]. 中国图象图形学报, 2009, 14(3): 377-387
- [8] Bashir F I, Khokhar A A, Schonfeld D. Object trajectory based activity classification and recognition using hidden Markov models[J]. IEEE Transaction on Image Processing, 2007, 16(7): 1912-1919
- [9] Abbone T S, Wendling L, Salmon J-P. A new shape descriptor defined on the Radon transform[J]. Computer Vision and Image Understanding, 2006, 102(1): 42-51
- [10] Wang Ying, Huang Kai-qi, Tan Tie-niu. Human Activity Recognition Based on R Transform [C]// Proc of the IEEE Conference on Computer Vision and Pattern Recognition. 2007: 1-7
- [11] 凌志刚,梁彦,潘泉,等. 基于张量子空间学习的人行为识别方法[J]. 中国图象图形学报, 2009, 14(3): 394-400
- [12] Schindler K, Van Gool L. Action Snippets: How Many Frames Does Human Action Recognition Require? [C]// Proc of IEEE CVPR. 2008: 1-8
- [13] 李宁,须德,傅晓英,等. 结合人体运动特征的行为识别[J]. 北京交通大学学报, 2009, 33(2): 6-16
- [14] 王天树,郑南宁,徐迎庆,等. 人体运动非监督聚类分析[J]. 软件学报, 2003, 14(2): 209-214
- [15] Elgammal A, Harwood D, Davis L. Non-parametric model for background subtraction[C]// Proc of ECCV. 2000, 1843: 751-767
- [16] 褚一平,陈勤,黄叶珏,等. 基于判别模型的视频前景/阴影自动分割算法[J]. 模式识别与人工智能, 2008, 21(6): 849-855
- [17] Sakoe H, Chiba S. Dynamic Programming Algorithm Optimization for Spoken Word Recognition [J]. IEEE Transaction on Acoustics, Speech, and Signal Processing, 1978, 26(1): 43-49
- [18] Gorelick L, Blank M, Shechtman E, et al. Action as space-time shapes [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(12): 2247-2253
- [19] Chen H S, Chen H T, Chen Y W, et al. Human action recognition using star skeleton[C]//Proc of the 4th ACM International Workshop on Video Surveillance and Sensor Networks. NY: ACM, 2006: 179-182