

加权 DTW 距离的自动步态识别

张浩 刘志镜

(西安电子科技大学计算机学院计算机信息应用研究中心, 西安 710071)

摘要 为了满足智能监控系统自动、准确、实时识别行人的要求,提出了一种表示方法简单、计算复杂度低的步态识别算法。首先建立环境的高斯背景模型,从步态视频序列中提取轮廓图像,计算质心以及轮廓线上的点到质心的欧氏距离。再将轮廓线以最高点为起点顺时针展开,将2维轮廓线特征转换为1维距离特征并标准化,建立标准步态模型。然后计算训练序列与标准步态模型之间的动态时间规整距离,确定阈值。最后,输入测试序列,计算动态时间规整距离并与阈值比较,识别人体的步态。与常用步态识别方法相比,该方法兼顾了计算复杂度和识别率,符合智能监控系统的性能要求。

关键词 步态识别 特征提取 动态时间规整 智能监控

中图法分类号: TP391.41 **文献标志码:** A **文章编号:** 1006-8961(2010)05-830-07

Automated Gait Recognition Using Weighted DTW Distance

ZHANG Hao, LIU Zhijing

(Computer Information Research Center, School of Computer Science and Technology, Xidian University, Xi'an 710071)

Abstract This paper presents a new approach to gait identification and authentication with simple representation and lower computational complexity, which can meet intelligent surveillance's need in precision and response. It creates Gaussian Mixture Model for each scenario, and contour of gait is extracted from binary silhouette for Euclidean distance between the centroid and any pixel on it. Contour is unfolded clockwise by the distance from the uppermost pixel, and then 2D features are transformed into 1D and normalized according to a standard model of gait. Thresholds are determined by dynamic time warping (DTW) distance between training sequences and standard model. Finally, gait recognition is performed by comparing DTW distance of testing sequences with predetermined threshold. Compared with other methods, it balances both computational cost and recognition rate, and achieves performance of intelligent surveillance.

Keywords gait recognition, feature extraction, dynamic time warping(DTW), intelligent surveillance

0 引言

鉴于计算机自动识别人体的需要,生物特征识别已经成为一个研究热点^[1]。在许多领域,虹膜、指纹、掌纹和鞋印等生物信息已经得到了深入的研究和广泛的应用。这些生物信息主要存在两个缺点:1)在低分辨率的图像中,这些信息不易采集;2)被观测者必须主动有效地配合,才能得到预期的

效果。

人体步态或者行走方式也是一种生物特征,此特征不受以上两个缺点的影响。因为它可以从远距离图像中获得,而且不需要被观测者的主动配合。步态包含步行者物理形状和心理状况等方面的信息。在一定情况下,利用步态信息完全可以对人进行识别。

人体步态的观测最先起源于医学目的,如 Murray 等人^[2]按步态把病人分组,来研究不同类型

基金项目:广东省教育部产学研结合项目(2006D90704017)

收稿日期:2009-05-11;**改回日期:**2009-08-20

第一作者简介:张浩(1980—),男。现为西安电子科技大学计算机学院计算机应用技术专业博士研究生。主要研究方向为计算机视觉、模式识别和图像处理。E-mail: zhanghao@mail.xidian.edu.cn

的治疗方法。在计算机视觉领域中,已经对人体运动进行了大量研究^[3],并且已经在人体步态分析和识别方面做了大量的工作。步态识别的性能受到许多因素的影响^[4-5],例如,轮廓质量、步行速度、动静态因素、耗费时间、鞋袜和环境条件^[6]、物理条件、伪装、室内外等,不同因素的影响具有相关性。虽然步态受到许多因素的影响,但是对于识别^[4]依然是非常有用的。

步态表示是识别过程中的一个关键问题,有效的表示方法可以提取更加显著的特征,有利于步态的识别和分析。主要的模型有 4 种:基于外形的模型^[4,7-11]、随机统计模型^[8]、关节生物力学模型^[12-14]和其他参数模型^[15]。不同种类的模型之间可以交叉组合,利用不同模型的特点,通过特征融合更好地表示步态特征。

基于外形的步态模型主要有两种,完整人体轮廓模型^[7-8,16-17]和部分人体轮廓模型^[18]。其中,完整轮廓主要使用人体侧面的轮廓图像,而部分轮廓主要采用变化显著的运动部位图像,如下肢。Ju 和 Bir^[7]提出了步态能量图像方法,利用合成步态模型表示步态特征,然后利用统计特征进行识别。在识别的过程中,分别需要对训练序列和测试序列的原始特征和合成特征进行主元分析和多分辨分析,达

到较好的识别效果。Amit 等人^[8]提出了两种步态识别方法:一种是利用帧间距离和隐马尔科夫模型(HMMs)进行识别,另一种是直接利用特征向量和 HMMs 进行识别。与参考算法相比,两种算法的识别率都有所提高。陈实等人^[16]提出用一组同心矩形分割两个单频运动历史图像,提取出局部性的矩统计量作为步态特征向量,实现步态识别。叶波和文玉梅^[17]提出了基于运动人体轮廓宽度特征的步态识别方法,用正交小波变换提取低维特征,然后用支持向量机训练和识别。以上方法虽然都获得了较好的实验结果,但是由于智能监控系统对计算复杂度和实时性要求较高,所以这些方法在应用方面略显欠缺。

智能监控系统通常并发处理多路视频信号,依据步态实时分析视频中的行人。因此,设计一种具有较高识别率、较低计算复杂度的算法就成为一个亟待解决的问题。

该算法先建立高斯背景模型,从中提取二值轮廓图像,再计算质心以及质心到轮廓线上各点的距离,将 2 维特征转换为 1 维特征并标准化,然后计算训练序列的动态时间规整(DTW)距离并建立模型,确定不同插值时的阈值,最后计算测试序列的 DTW 距离,通过阈值比较判别步态的相似程度。算法流程如图 1 所示。

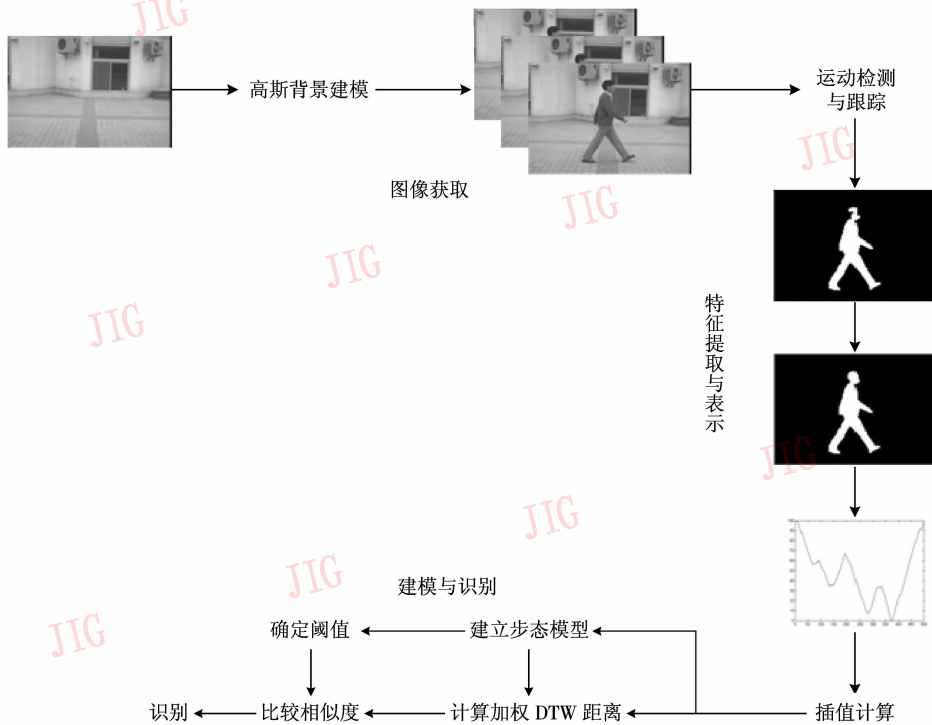


图 1 加权 DTW 距离算法的流程图

Fig. 1 The flowchart of gait recognition using weighted DTW distance

1 原理概述

1.1 轮廓线展开

人体轮廓形状是人运动的重要特征之一,它随时间呈周期性变化,在人体运动状态描述方面,具有一定的代表性。为了更加显著地表示人体特征、降低计算复杂度,将人体 2 维轮廓特征转化为 1 维距离特征,近似表示步态运动的时空变化,轮廓线展开过程如图 2 所示。

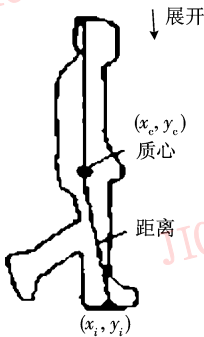


图 2 轮廓线展开示意图

Fig. 2 Contour unfolding

1) 质心计算 提取人运动轮廓后,可以使用边界提取算法(如 Canny 算法)提取轮廓线上的坐标点,利用下式计算轮廓的质心坐标:

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

式中, (x_c, y_c) 是质心坐标, N_b 是轮廓线上点的个数, (x_i, y_i) 是轮廓线上点的坐标。

2) 轮廓线展开 选定轮廓线上纵坐标最大的点为起点(取左下方为图像坐标原点),沿顺时针方向,利用下式将轮廓线展开为质心到轮廓点距离的 1 维特征:

$$d_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (2)$$

3) 数据标准化 为了消除图像尺度和特征长度不同造成的影响,在最大限度地保持特征完整性的前提下,将特征的幅度和长度值标准化,以达到利于分析数据规律和特征的目的。

1.2 动态时间规整

动态时间规整是一种柔性模式匹配算法,能够对存在全局或局部扩展、压缩或变形的模式进行匹配,解决动态模式的相似度量 and 分类问题。DTW 最早用于语音信号识别领域,现在已经在动态多变量

工业过程领域得到了广泛的应用。

若两个时间序列 A 和 B 分别表示为

$$A: \{a_1, a_2, \dots, a_i, \dots, a_n\}$$

$$B: \{b_1, b_2, \dots, b_j, \dots, b_m\}$$

规整路径 P 定义为

$$P: \{p_1, p_2, \dots, p_k, \dots, p_K\}$$

其中, $p_k = (i, j)_k; \max(m, n) \leq K \leq m + n$ 。

DTW 可以表示为如下的目标函数:

$$J = \frac{1}{\sum_{k=1}^K w_k} \min \left\{ \sum_{k=1}^K w_k d(i_k, j_k) \right\} \quad (3)$$

$$P^* = \arg \min \left\{ \sum_{k=1}^K w_k d(i_k, j_k) \right\} \quad (4)$$

式中, $d(i_k, j_k) = |a(i) - b(j)|$, $d(i_k, j_k)$ 是两个序列中 a_i 与 b_j 之间的局部距离; 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;分母项 $\sum_{k=1}^K w_k$ 对规整路径的不同长度加以补偿,从而使两个序列之间具有可比性。

DTW 对规整路径作了一定的限制,即规整路径应满足局部约束和全局约束的条件。局部约束包括端点约束、连续约束和单调约束 3 种类型。

1) 端点约束

$$p_1 = (1, 1), p_K = (n, m)$$

端点约束表明,两个时间序列的起始点和终止点是一致的。实际上,可以放宽对端点约束条件的限制,允许起始点和终止点分别对应地落在某个窗口内。

2) 连续性约束

对于局部路径中 $p_{k-1} = (i, j)$ 和 $p_k = (i', j')$ 两点,连续性约束可以表示为

$$i' - i \leq c, j' - j \leq c$$

c 为正整数(一般取常数 1 或 2)。

连续性约束用来限制 DTW 在规整过程中的过度扩张。图 3 为两种典型的连续性约束条件。

3) 单调性约束

对于局部路径 $p_{k-1} = (i, j)$, $p_k = (i', j')$, 要求

$$i' - i \geq 0, j' - j \geq 0$$

全局约束定义了 $n \times m$ 网格中的一个子集作为路径的搜索空间。全局约束将规整路径限制在一定的区域内,如带状区域或者平行四边形区域。

DTW 问题可通过最优化问题的方法来求解。式(5)可以由如下的递推关系,通过迭代的方法得到解。

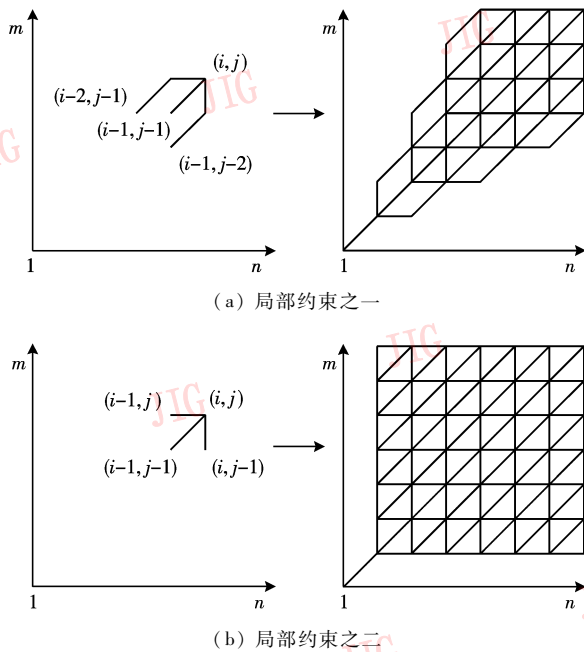


图 3 两种局部约束及其可能路径

Fig. 3 Two local restrictions and possible paths

$$D(i, j) = d(i, j) + \min \{ D(i-2, j-1) + d(i-1, j), D(i-1, j-1), D(i-1, j-2) + d(i, j-1) \} \quad (5)$$

根据不同的局部约束,可以得到不同的递推关系式。式(5)为图 3(a)局部约束的情况,其初始条件为

$$D(2, 2) = d(1, 1) + d(2, 2)$$

考虑图 3(b)局部约束的情况,递推关系为

$$D(i, j) = d(i, j) + \min \{ D(i-1, j), D(i-1, j-1) \} \quad (6)$$

2 步态建模

在步态识别过程中,设计并建立步态模型是核心问题。建模过程主要分为 3 个阶段:图像获取、特征提取与表示和模型建立,如图 1 所示。

2.1 图像获取

输入步态视频序列,利用混合高斯模型建立背景模型,检测和跟踪运动目标。

2.2 特征提取与表示

利用背景剪除法,提取人体轮廓的二值图像。然后对二值图像进行膨胀、腐蚀、连通等预处理,构成单连通人体轮廓。由于步态运动具有周期性(如图 4 所示),可以通过一个步行周期表示完整的步行过程。设视频序列帧的编号为 n ,人体轮廓的高

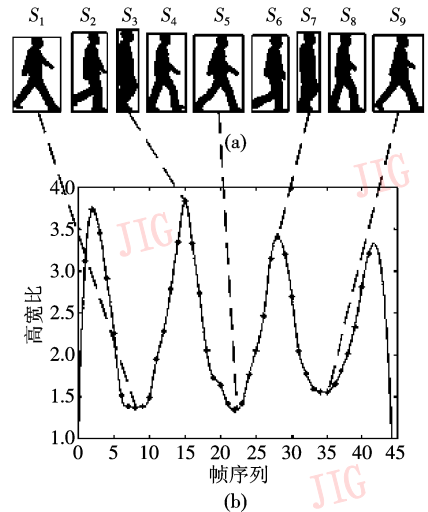


图 4 步态周期轮廓图像与轮廓外接矩形的高宽比
Fig. 4 Silhouette images of gait cycle in sequences and the ratios of height to width

宽比为 r ,因此两者之间的关系可表示为

$$r = f(n) \quad (7)$$

式中, $1 \leq n \leq N$, N 为视频帧数。

用轮廓线展开法对步态进行特征表示,由式(1)计算人体轮廓的质心坐标 (x_c, y_c) ,设 d_i 为人体质心到轮廓线上第 i 个点的距离,由式(2)得到 1 维特征 d_i 。统计视频序列中人体轮廓线的像素点个数,95%以上的视频序列满足 $N_b \in [400, 600]$,且呈正态分布。因此,标准化后取 $N_b = 500$ 。设 D 为标准化的距离范围,取 $D = 100$ 。轮廓特征降维并标准化的结果如图 5 所示。

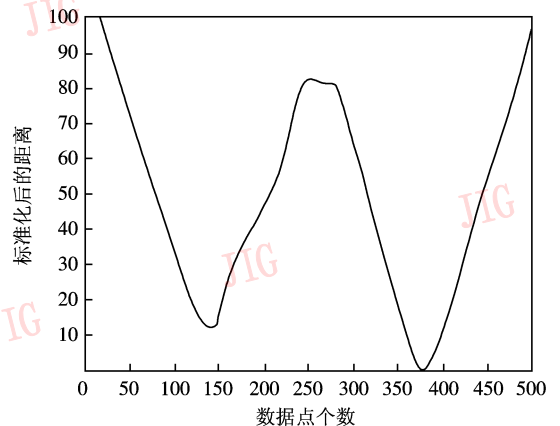


图 5 距离特征归一化

Fig. 5 Normalized distance

2.3 模型建立

为准确计算 DTW 距离,对标准化后的数据进

行插值,插值点为 $x_i (x_i \in \mathbf{Z}, 0 \leq x_i \leq 500)$, 间隔为 $\sigma = x_{i+1} - x_i$ 。根据已有的数据点, 计算 $500/\sigma + 1$ 个点的值。提取步态的一个完整周期(两步)作为模型, 按照双脚开立的距离, 分别标记为 $S_1 \sim S_8$ 8 个状态。其中, 状态 S_3, S_7 表示模型高宽比的极大值, 状

态 S_1, S_5 为高宽比的极小值, 状态 S_2, S_4, S_6, S_8 为过程值(如图 4 所示)。由式(7)可得到 $S_1 \sim S_8$ 所对应的帧编号 $n_1 \sim n_8$ 。为了减小误差, 分别取 $n_i (1 \leq i \leq 8)$ 的前后 3 帧, 即 $n_i - 1, n_i, n_i + 1$, 计算平均值, 得到图 4 状态所对应的 8 个状态参数, 如图 6 所示。

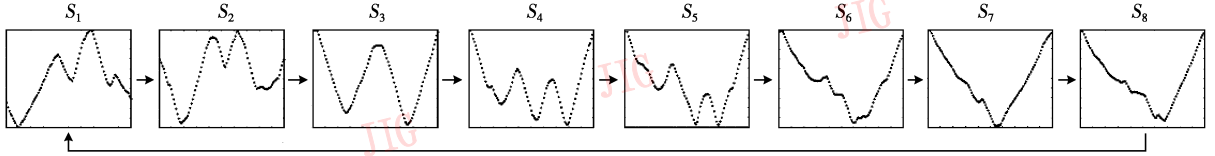


图 6 模型状态转换图

Fig. 6 Transitions of model with 8 states

输入多组视频序列进行训练。按照图 3 所示的步骤得到状态参数, 由式(6)计算训练数据与模型参数之间的 DTW 距离。并对每组序列进行 3 组插值, 插值间隔 σ 分别为 20、10 和 5, 比较对应状态之间 DTW 距离的变化程度, 统计结果如表 1 所示。

值为 0.2, 过程状态的权值为 0.05, 可得

$$D_s = (S_1 + S_3 + S_5 + S_7) \times 0.2 + (S_2 + S_4 + S_6 + S_8) \times 0.05 \quad (8)$$

由式(8)计算不同插值时, 步行周期的加权 DTW 距离值 D_s 的分布情况如表 2 所示。

表 1 σ 不同时 DTW 距离的比值关系

Tab. 1 Percentage of ratios variation with σ

状态	$\sigma_{20} : \sigma_{10}$	$\sigma_{10} : \sigma_5$	$\sigma_{20} : \sigma_5$
S_1	83	83	54
S_2	77	69	53
S_3	73	89	65
S_4	96	81	68
S_5	79	67	53
S_6	58	55	32
S_7	81	64	52
S_8	67	67	45
平均值	76.8	71.9	52.8

表 2 DTW 距离的分布统计

Tab. 2 Statistical distribution of DTW distance

D_s	$\sigma = 20$	$\sigma = 10$	$\sigma = 5$
$[0, 50)$	15	5	0
$[50, 100)$	35	20	5
$[100, 150)$	35	35	5
$[150, 200)$	5	30	25
$[200, 250)$	5	5	25
$[250, 300)$	5	5	30
$[300, 350)$	0	0	5
$[350, \infty)$	0	0	5

由图 4 可知, 由于状态 S_1, S_3, S_5 和 S_7 的变化幅度较小, 具有明显的步行特征, 根据这 4 个状态就可以较为准确地识别人的步态, 故称为特征状态。而状态 S_2, S_4, S_6 和 S_8 变化幅度较大, 步行特征欠显著, 故称为过程状态。选取数据库中步态序列进行训练, 计算一个完整周期(8 个状态)所对应的距离值, 在状态 $S_1 \sim S_8$ 的转换过程中, 不同状态的 DTW 距离值是相差较大的, DTW 距离值在特征状态相对较小, 而在过程状态相对较大。因此, 为了更精确地描述步态, 依据各个状态对识别结果的影响程度, 对各个状态的 DTW 距离加权求和, 取特征状态的权

由表 2 可知, 阈值不同时得到的识别率如表 3 所示。因此, 当 $\sigma = 20, 10$ 和 5 时, 阈值 α 可分别设定为 150、200 和 300。

表 3 不同阈值的识别率

Tab. 3 Recognition rate in different thresholds

	分布概率/%
$\sigma = 20$ 且 $D_s \in [0, 150]$	85
$\sigma = 10$ 且 $D_s \in [0, 200]$	90
$\sigma = 5$ 且 $D_s \in [0, 300]$	90
平均值	88.3

3 实验结果

3.1 实验数据库

1) CASIA(中国科学院自动化研究所)数据库

目前该步态数据库有 3 个数据集:Dataset A(即 NLPR 步态数据库)创建于 2001 年 12 月 10 日,包含 20 个人的数据,每人有 3 个行走方向(与图像平面分别成 $0^\circ, 45^\circ, 90^\circ$),每个方向有 4 个图像序列,每个序列的帧数介于 37 ~ 127 之间。Dataset B 是一个大规模的、多视角的步态库,采集于 2005 年 1 月。共有 124 个人,每个人有 11 个视角 ($0^\circ, 18^\circ, 36^\circ, \dots, 180^\circ$),在 3 种行走条件下(普通条件、穿大衣、携带包裹)采集。Dataset C 是一个用红外(热感)摄像机在夜间拍摄的大规模数据库,采集于 2005 年 7 ~ 8 月,包含 153 人,每个人在 4 种条件下

行走:正常走、快走、慢走、带包走。

2) SOTON(南安普顿大学)数据库

该数据库创建于 2002 年 5 月 15 日,后经过近两年的补充,成为一个完善的测试集。数据库由两部分组成,大数据集有 115 人,小数据集有 12 人,每人有 3 个行走方向(与图像平面分别成 $0^\circ, 45^\circ, 90^\circ$),每个方向有 4 个图像序列,包含室内、室外场景和不同速度的视频,在多种行走条件下(穿着不同鞋袜、外衣、携带包裹)采集。

3.2 实验结果

分别用 CASIA 数据库的 Dataset A 和 Dataset B,以及 SOTON 数据库进行建模和测试。对同一个人的多组步态视频序列,选择一组作为建模序列,从剩余序列中随机选择一组作为测试序列。其中,所有视频序列均与被观测者的运动方向成 90° ,测试结果如表 4 所示。

表 4 CASIA 和 SOTON 数据库的识别结果

Tab.4 Statistics of gait recognition using CASIA and SOTON database

数据库	测试集	测试人数	识别率/%			平均值	总平均值
			$\sigma = 20$ 且 $D_s \in [0, 150]$	$\sigma = 10$ 且 $D_s \in [0, 200]$	$\sigma = 5$ 且 $D_s \in [0, 300]$		
CASIA	Dataset A	20	90	95	95	93.3	91.1
	Dataset B	100	91	92	94	92.3	
SOTON	小数据集	12	83.3	91.7	91.7	88.9	90
	大数据集	100	88	91	91	90	

由于插值间隔 σ 不同,加权计算得到的 DTW 距离不同,故阈值也就不同。因此,确定符合实际情况的阈值,是识别算法的关键。从表 4 的数据可知,与建模时确定的阈值相比,对应阈值的识别率提高了 1% ~ 5%,总识别率提高了 2.8%,由此验证了算法的有效性,也间接证明了阈值的合理性。

由表 5 可知,通过与其他算法相比较,可以发现本文算法在识别率方面有所提高,在计算复杂度方面有较大的降低,同时兼顾了两者,使该算法的整体性能有了较大的提高,满足了智能监控系统的性能要求,在实际应用方面具有重要的意义。

表 5 几种算法的结果比较

Tab.5 The comparison of algorithms

算法	平均识别率/%	计算复杂度
文献[16]	85.57	中
文献[17]	91	高
本文算法	91.1	中

4 结 论

结合智能监控系统性能的要求,本文提出了一种加权 DTW 距离的步态识别算法。该算法将 2 维特征转换为 1 维特征,然后建立外形的加权 DTW 距离模型,利用训练序列确定阈值并进行识别。与其他几种算法相比,在识别率方面有所提高,但是对步态特征表示简单,计算复杂度较低,又无需学习,符合智能监控系统的应用要求。与基于统计特征的方法相比,此算法在特征分析方面略显欠缺。未来的工作主要集中在将模型的全局特征与局部特征相结合,进一步提高步态识别率。

参考文献 (References)

[1] Turaga P, Chellappa R, Subrahmanian V S, et al. Machine recognition of human activities: A survey [J]. IEEE Transactions

- on Circuits and Systems for Video Technology, 2008, 18(11): 1473-1488.
- [2] Murray M, Drought A, Kory R. Walking pattern of normal men[J]. Journal of Bone and Joint Surgery, 1964, 46-A(2): 335-360.
- [3] Moeslund T B, Granum E. A survey of computer vision-based human motion capture [J]. Computer Vision and Image Understanding, 2001, 81(3): 231-268.
- [4] Sarkar S, Phillips P, Liu Z, et al. The human ID gait challenge problem: Data sets, performance and analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(2): 162-177.
- [5] Boyd J E. Synchronization of oscillations for machine perception of gaits [J]. Computer Vision and Image Understanding, 2004, 96(1): 35-59.
- [6] Haritaoglu I, Cutler R, Harwood D, et al. Backpack: Detection of people carrying objects using silhouettes [J]. Computer Vision and Image Understanding, 2001, 6(3): 385-397.
- [7] Han J, Bhanu B. Statistical feature fusion for gait-based human recognition [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: Institute of Electrical and Electronics Engineers Inc., 2004: 842-847.
- [8] Kale A, Sundaresan A, Rajagopalan A N, et al. Identification of humans using gait [J]. IEEE Transactions on Image Processing, 2004, 13(9): 1163-1173.
- [9] Liu Z, Sarkar S. Simplest representation yet for gait recognition: Averaged silhouette [C]//Proceedings of International Conference on Pattern Recognition. Washington, DC, USA: Institute of Electrical and Electronics Engineers Inc., 2004: 211-214.
- [10] Tanawongsuwan R, Bobick A. Modelling the effects of walking speed on appearance-based gait recognition [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: Institute of Electrical and Electronics Engineers Inc., 2004: 783-790.
- [11] Wang L, Tan T, Ning H, et al. Silhouette analysis-based gait recognition for human identification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(12): 1505-1518.
- [12] Bobick A, Johnson A. Gait recognition using static activity-specific parameters [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: Institute of Electrical and Electronics Engineers Inc., 2001: 423-430.
- [13] Tanawongsuwan R, Bobick A. Gait recognition from time-normalized joint-angle trajectories in the walking plane [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: Institute of Electrical and Electronics Engineers Inc., 2001: 726-731.
- [14] Cunado D, Nixon M, Carter J. Automatic extraction and description of human gait models for recognition purposes [J]. Computer Vision and Image Understanding, 2003, 90(1): 1-41.
- [15] Cutler R, Davis L. Robust periodic motion and motion symmetry detection [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: Institute of Electrical and Electronics Engineers Inc., 2000: 615-622.
- [16] Chen Shi, Tian Yumin, Huang Wanhong, et al. Automatic human gait recognition using temporal template [J]. Journal of Xidian University, 2007, 34(4): 605-610. [陈实, 田玉敏, 黄万红, 等. 用时变模板自动识别行人的步态 [J]. 西安电子科技大学学报:自然科学版, 2007, 34(4): 605-610.]
- [17] Ye Bo, Wen Yumei. Gait recognition based on DWT and SVM [J]. Journal of Image and Graphics, 2007, 12(6): 1055-1063. [叶波, 文玉梅. 基于小波变换和支持向量机的步态识别算法 [J]. 中国图象图形学报, 2007, 12(6): 1055-1063.]
- [18] Lee L, Dalley G, Tieu K. Learning pedestrian models for silhouette refinement [C]//Proceedings of the IEEE International Conference on Computer Vision. Washington, DC, USA: Institute of Electrical and Electronics Engineers Inc., 2003: 663-670.