

# 基于 Bayes 规则与 HMM 相结合的步态识别方法研究

余 涛 邹建华

(西安交通大学系统工程研究所 西安 710049)

(西安交通大学系统工程国家重点实验室 西安 710049)

**摘 要** 提出一种将 Bayes 规则与 HMM 相结合的步态识别方法. 检测环节通过采用重心变化作为特征序列来削弱时间差分算法中运动实体目标存在中空 of 负面影响, 通过对检测出的目标特征序列采用直线拟合提取对称轴, 并等效转化为具有方向性的点线距序列, 来简化运算, 降低失真度. 步态训练中, 通过初始化的修正使训练出的 HMM 接近全局最优, 并给出 Bayes 相关先验学习方法. 步态识别中, 应用 HMM 的前-后向算法, 并融合 Bayes 规则, 客观性增强, 最终结果经中国科学院 CASIA 的原始步态视频测试, 达到比较高的识别率, 且对衣着具有一定鲁棒性. 该文方法主要针对视角在  $0^{\circ} \sim 180^{\circ}$  间的室内监控直道行走场景.

**关键词** HMM; Bayes 规则; 步态; 识别; 序列

**中图法分类号** TP391 **DOI 号:** 10.3724/SP.J.1016.2012.00386

## A Method of Combining Bayes Rule with HMM in Gait Recognition

YU Tao ZOU Jian-Hua

(Systems Engineering Institute, Xi'an Jiaotong University, Xi'an 710049)

(State Key Laboratory for Systems Engineering, Xi'an Jiaotong University, Xi'an 710049)

**Abstract** This paper presents a framework of combining Bayes rule with Hidden Markov Model (HMM) to recognize human identification by gait, indoors. First, the monitored human motion is detected mainly by a three-frame differencing algorithm. Then, a curve of centroid on the object's motion can be acquired. The curve is transformed into the observation sequence of its corresponding HMM by adaptive filtering, median filtering, line fitting, rotating equivalently, normalizing, nearest neighbor clustering, and cycle extracting in turn. During the process of training the HMM with Baum-Welch algorithm, the original parameters of matrix  $\mathbf{B}$  is modified statistically by Viterbi algorithm, making the final trained model approximate global optimization further. And the prior knowledge in the Bayes rule is also acquired from relative learning. Lastly, the observation sequence is used to recognize the human identification by means of combining Bayes rule with the Forward-Backward algorithm in trained HMM. In the end, the performance of the method is illustrated by the videos of the CASIA Gait Database, the result acquires comparatively higher recognition rate and is robust for the objects' clothes to a certain extent. The framework of this paper is fitting for monitoring indoors. The type of the gallery monitored is straight and the sight angle for object is between  $0^{\circ}$  and  $180^{\circ}$ .

**Keywords** HMM; Bayes rule; gait; recognition; sequence

# 1 引言

步态作为一种生物特征中的行为特征,对其检测和识别不需要身体接触,可在低分辨率图像序列中进行,是唯一一个可在远距离识别的生物特征且难以伪装和隐藏,这些特点使得有关步态识别的研究近年来备受关注<sup>[1]</sup>.步态识别的基本任务是根据拍摄的人的行走视频,自动提取表征步行特征的视觉线索,通过人体走路的方式来鉴别人的身份.其主要环节包括:步态预处理、步态检测、步态特征的提取与表征和步态特征识别.

步态预处理是将前景运动目标从背景中分割出来.主要方法:背景减除法、时间差分法和光流法<sup>[2]</sup>.步态检测是用知识或统计的方法对步态建立模型,比较所有可能的待检测区域与步态模型的匹配度,从而得到可能存在步态的区域<sup>[3]</sup>.主要检测步态周期,而且步态轮廓的有效分割对于后期处理也非常重要.常用方法:差分图像、互相关分析、时空梯度<sup>[4]</sup>等.步态序列的特征提取及表征主要分为两类:结构化(模型)方法和非结构化方法.前者对人的身体结构或运动进行建模,通过提取图像特征并且映射它们到模型的结构分量或导出体元的运动时间轨迹来识别个体<sup>[3]</sup>.常对人体建立 2D 或 3D 模型<sup>[5-6]</sup>;后者不考虑潜在结构,而是特征化人体整个运动模式以得到简洁的运动特征表达<sup>[3]</sup>.通常输入的步态特征维数较高,一般要选取如主元分析(PCA)<sup>[7]</sup>、多元判别分析(MDA)<sup>[8]</sup>等适当方法来降维.步态特征的分类识别通常也分为两类:模板匹配方法(TM)和状态空间方法(SSA).前者将图像序列转换为的一组静态形状模式,然后在识别过程中和预先存储的行为标本比较,以某种相似度作为分类标准<sup>[9-11]</sup>;后者定义每个静态姿势为一个状态,状态间通过某种概率联系起来.任何运动序列可看作这些静态姿势的不同状态间的一次遍历过程,在遍历期间计算联合概率,其最大值被选择作为分类标准<sup>[9]</sup>.常用隐马尔可

夫模型(HMM)<sup>[12-13]</sup>、动态时间规划(DTW)<sup>[14]</sup>、神经网络(NN)<sup>[15]</sup>、支持向量机(SVM)<sup>[16]</sup>等.

由于目前计算机视觉领域中从运动目标的检测跟踪到目标识别,每个环节的问题都还没有完全解决,所以很多实际步态识别框架跳过了检测跟踪环节的实时性、精确性要求,用已经预处理过的,图像分割较好的运动目标且从多参数多维空间的角度进行研究.当前,时间序列问题成为运动行为识别分析研究的热点方向,以 HMM 为代表的状态空间法已被广泛应用于时间序列的检测、估计和预测识别中.而在识别时,无论是选择何种特征,随着样本数量增多,不同样本间的特征相似的可能性自然会增大,这时若仅依靠现有的特征去判别,很有可能识别错误,此时若能合理地融入一些有利于区别个体的先验客观因素,适当兼顾识别问题的实际客观背景情况,将更有利于准确识别.同时,对这种先验的获取,若能从实际历史的训练学习中获得,而不是主观设定,也将会有利于确保先验知识的客观性.

本文直接从原始步态视频入手,对实际目标先通过时间差分方法进行检测跟踪;通过检测目标的重心变化,得到相应特征序列;再对该特征序列通过自适应滤波、中值滤波、直线拟合、等效旋转、规则化、最近邻聚类 and 周期提取一系列处理后,作为特征观测序列;在进行身份识别时,根据事先通过 Baum-Welch 算法训练好的(训练前,初始的状态观测矩阵  $B$ ,通过 Viterbi 算法,结合相应观测序列  $O$ ,被统计修正)若干 HMM,通过前向-后向算法,并结合带有训练学习获得的先验知识的 Bayes 规则,对目标进行身份识别;最终通过用中国科学院的标准 CASIA 视频库的一些原始视频进行了相关实验,得到较高的识别率,且对衣着具有一定的鲁棒性.该方法适合于视角在  $0^\circ \sim 180^\circ$  范围之间的室内监控直道场景.

## 2 方法原理

本文的步态检测及识别的总体框架如图 1 所示,

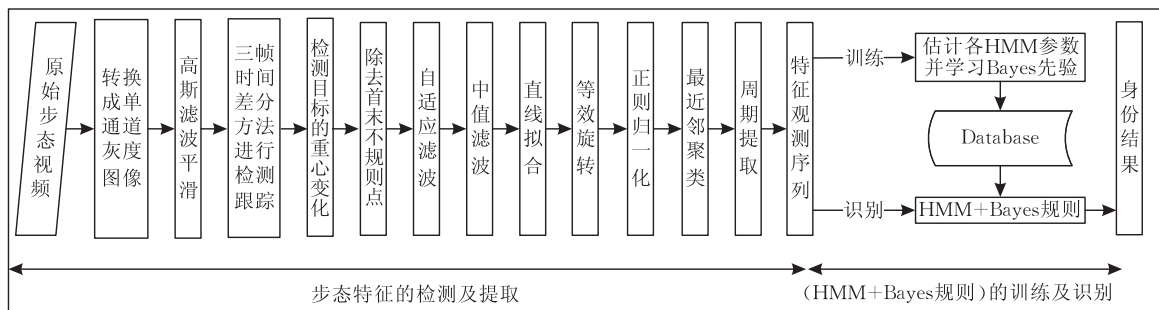


图 1 本文的步态检测及识别总体框架

主要包括两部分:步态特征的检测与提取;(HMM+Bayes)规则的训练和识别.以下将对框架中的两部分分别阐述.

## 2.1 步态特征的检测及提取

在对目标的检测跟踪环节中,先将原始视频图像转换为单通道灰度图像,再进行高斯滤波平滑图像,然后采用类似文献[17]中的三帧差分转化为二值图,具体计算如下:

$$\begin{cases} |I_n(x) - I_{n-1}(x)| > T_n(x) \\ |I_n(x) - I_{n-2}(x)| > T_n(x) \end{cases} \quad (1)$$

式中: $T_n(x)$ 表示在像素位置 $x$ 处的具有统计意义的描述灰度变化的域值.当某像素位置 $x$ 处的灰度值 $I_n(x)$ 相对上一帧 $I_{n-1}(x)$ 和前一帧 $I_{n-2}(x)$ 都有较大变化时,则认为该点像素属于运动目标,这样将运动目标从背景中分离出,重复操作即得到一系列前景目标图像.

由于该方法对检测出的运动目标区域内部通常产生中空,而目标重心的计算公式中的求和运算具有正负抵消的作用,这样一定程度上消除了中空对特征检测的影响,因此后续的目标特征选取为目标的重心变化,计算如下:

$$x_{\text{重}(t)} = \frac{1}{N} \sum_{x_i \in \text{目标区域}} x_i, \quad y_{\text{重}(t)} = \frac{1}{N} \sum_{y_i \in \text{目标区域}} y_i; \quad (2)$$

$N = \text{目标区域内像素点数}$

依次求出每一帧图像的目标重心,就可绘出目标的重心特征的运动轨迹,如图2所示.图(a)为原始视频截图,图(b)是将上述时间差分法得到的运动目标的灰度图转换成彩色图,并连续描绘出过去时间内的运动重心变化的特征曲线.

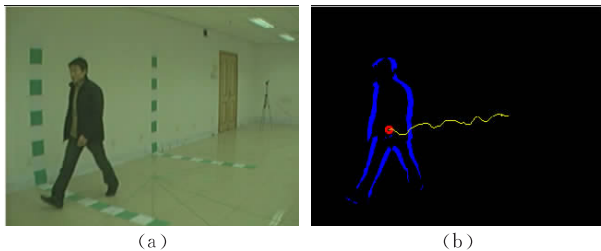


图2 目标步态的重心特征曲线检测图

将此特征序列处理如下:

由于目标在开始进入场景及最后即将离开场景时,只有部分轮廓可以检测到,所以那段时间得到的特征点为不规则点,因此首先通过设定域值去掉首尾两端部分的不规则点.然后滤波去噪.由于特征序列信号的周期变化形式多种多样,所以这里采用由文献[18]提出的一种自适应最小均方误差算法进行

滤波,即自适应滤波,滤波原理如图3所示.

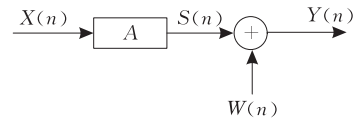


图3 自适应滤波原理图

结合本文应用,图3中: $X(n)$ 对应检测出的特征序列的横坐标(或者称为时间顺序号), $Y(n)$ 对应该序列的纵坐标, $A$ 为标量权系数, $W(n)$ 对应图像或视频中常见的高斯噪声信号,通常 $W(n)$ 为平稳随机干扰信号且与 $X(n)$ 互不相关,可看出:

$$\begin{cases} Y(n) = S(n) + W(n) \\ S(n) = AX(n) \end{cases} \quad (3)$$

利用 $Y(n)$ 和 $X(n)$ 可得到 $S(n)$ 的估计,相应的自适应估计算法如下:

$$\begin{cases} \hat{S}(n) = \hat{A}(n)X(n) \\ \hat{A}(n+1) = \hat{A}(n) + 2\mu(Y(n) - \hat{A}(n)X(n))X(n) \end{cases} \quad (4)$$

选择 $\hat{A}(n)$ 的初始值 $\hat{A}(0)$ ,根据文献[18],只要满足:

$$0 < \mu < 1/R, \text{ 其中: } R = E[x(n)x(n)^T] \quad (5)$$

$\hat{A}(n)$ 总可以收敛于最佳值 $A(n)$ , $\hat{S}(n)$ 也逐渐收敛于 $S(n)$ .

本文实验中发现:对检测出的特征序列按照上述自适应估计算法迭代两次(其中,第2次滤波各点的输入是在第1次滤波的末端点值的基础上继续累加)就得到比较平稳且原信号变化失真较小的特征序列,如图4中点划线所示.可看出,虽然由于滤波迭代计算中的初始累加,使得滤波后的信号幅度整体略大于滤波前,但相对变化趋势基本得以保持,本文研究的正是信号的这种相对变化特征.

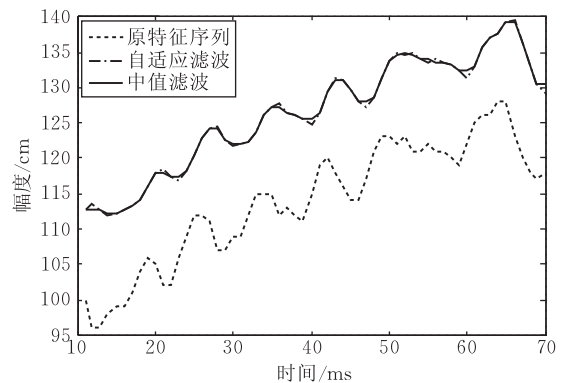


图4 对特征曲线进行滤波前后的效果比较图

自适应滤波后,再采用中值滤波进一步去除由噪声引起的序列局部抖动误差,或称为序列中的毛刺误差,具体将原序列中的每一点的值和其邻近两

侧值进行比较,三者中按大小排序,取中间值作为当前点滤波后的值,然后依次类推.滤波结果如图4中实线所示.

上述滤波完成后,再利用最小二乘法对序列进行直线拟合以求出曲线的拟合中心轴,拟合计算如下:

$$\begin{cases} \bar{x} = \frac{1}{M} \sum_{i=1}^M x_{重(i)}, \bar{y} = \frac{1}{M} \sum_{i=1}^M y_{重(i)}, \\ \overline{x^2} = \frac{1}{M} \sum_{i=1}^M x_{重(i)}^2, \overline{xy} = \frac{1}{M} \sum_{i=1}^M x_{重(i)} y_{重(i)} \\ k = \frac{(\overline{xy} - \bar{x} \cdot \bar{y})}{\overline{x^2} - \bar{x}^2}, b = \bar{y} - k \cdot \bar{x} \end{cases} \quad (6)$$

其中: $M$ 为特征序列点数目, $k$ 为拟合直线斜率, $b$ 为拟合直线截矩.

通常在得到曲线的拟合中心轴后,需要计算出该直线倾角,再按照该倾角大小根据对应象限,将序列各点反向旋转,即将曲线的中心轴旋转至水平轴上,这样的计算效果往往累积误差较大,失真度较严重.本文采用的方法是直接计算出特征序列的各点至该拟合直线的距离,以这些距离作为新的序列值的绝对值,以原特征序列的各点相对拟合直线上对应各点的高度正负作为新的序列值的正负号,这实际上相当于实现了原特征序列的中心轴向水平轴的旋转,相应原理如图5所示,计算如下:

$$l_i = \frac{|y_{重(i)} - b - k \cdot x_{重(i)}|}{\sqrt{k^2 + 1}}$$

其中, $l_i$ 为重心序列中某点 $(x_{重(i)}, y_{重(i)})$ 距拟合直线的距离,上式去掉绝对值后,正负号表示特征序列相对拟合直线的上下变化

$$l_i = \frac{y_{重(i)} - b - k \cdot x_{重(i)}}{\sqrt{k^2 + 1}} \quad (7)$$

式(7)即为新的特征序列的计算公式.

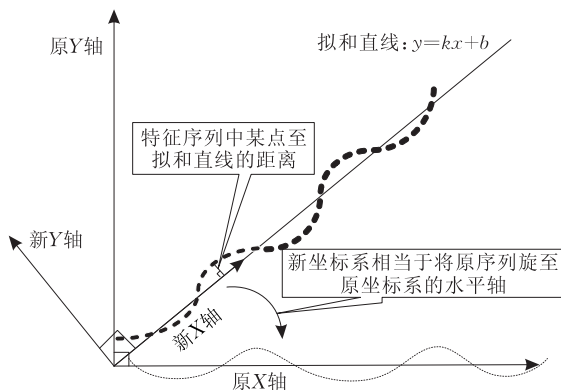


图5 特征序列等效旋转原理示意图

之所以要等效完成这样的旋转是因为实际监控

通道的方向往往相对监控终端界面的坐标轴有所倾斜,通过上面的变换等价于旋转监控通道方向轴,使其与监控终端界面的坐标轴平行.从而减小了由实际的通道倾斜使目标特征检测误差增大的影响.

然后将上述经过等效对称轴旋转后的序列,按下式所示,通过除以序列中最大幅度的绝对值实现归一化:

$$l_i = l_i / \max |l_i|, 0 \leq i \leq T \quad (8)$$

最后,根据步态的周期性,对得到的归一化序列,按图6所示,实现观测值的最近邻聚类,本文设定9种状态值,5种观测值.图6中可看出,观测值在对称轴附近为1,然后依次向对称轴两侧等间隔划分,观测值大小依次递增.这种划分与实际步态特征序列的变化幅度方向相一致,使其对步态信息的反映更直观.这样就得到一组观测值序列,其式如下:

$$\arg \min_{o_i \in O} (|l_i - o_i|) = o_i, 0 \leq i \leq T \quad (9)$$

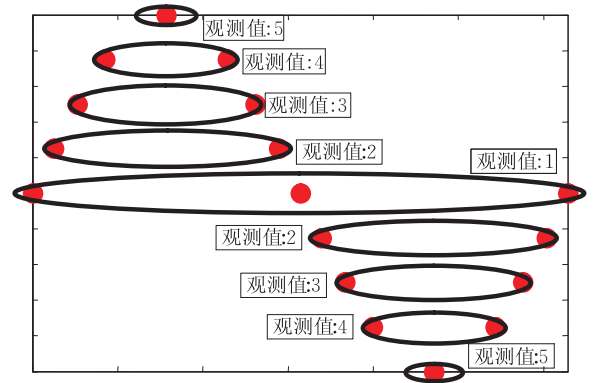


图6 特征观测值序列聚类示意图

本文在训练阶段,采用多个特征观测值序列训练,每个序列多个周期,这样使训练中的样本尽量充分;在识别阶段,采用单个特征观测值序列识别且对该序列只抽取一个周期.对于周期的确定,结合图6可以看出,需要求解出任一最高观测值与相邻的最低值的间距,考虑到实际曲线可能出现的非正规对称,这里是以某个最高观测值(或最低观测值)为基准,取其两侧当中相距较远的最低观测值(或最高观测值)对应的间距,将此间距乘以4倍即为一个周期.即式(10)所示:

$$T = 4 \times (n_{o_5} - n_{o_1} + 1) \quad (10)$$

式中: $n_{o_5}, n_{o_1}$ 为上图中相邻的最高、最低观测值在观测序列中的对应序号.

## 2.2 HMM 与 Bayes 规则的相结合的训练及识别

本文在对 HMM 训练中,参考了文献[19-20],主要综合应用了文献[20]中为克服可能因结果过小出现“下溢”现象而引入比例因子加以修正后的

HMM 相关的 3 种基本算法,即前向-后向算法,多序列训练的 Baum-Welch 算法和 Viterbi 算法,并设计出相关的修正初始化,实现了对 HMM:  $\lambda = (\pi, A, B)$  的训练. 这里设状态数目为  $N$ ; 每个状态有  $M$  个观测值;  $\pi = (\pi_1, \pi_2, \dots, \pi_N)$  为各状态初始向量;  $A = (a_{ij})_{N \times N}$  为状态转移概率矩阵;  $B = (b_{ij})_{N \times M}$  为状态观测概率矩阵;  $O = \{o_1, o_2, \dots, o_T\}$  为观测值序列;  $Q = \{q_1, q_2, \dots, q_T\}$  为状态序列. 对 HMM 的整个训练流程如图 7 所示.

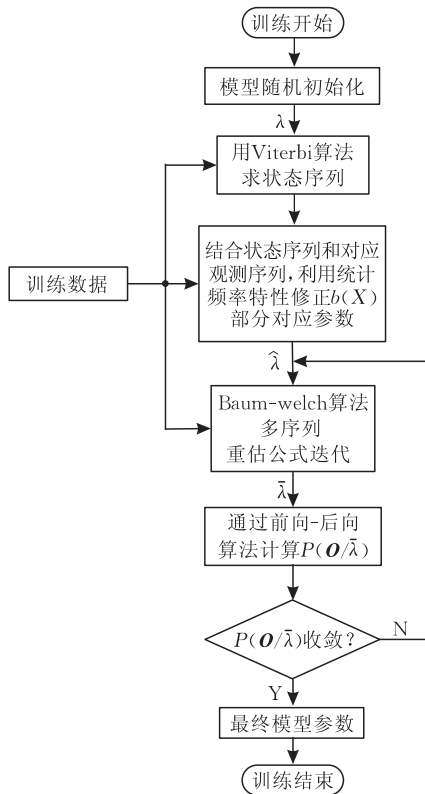


图 7 本文使用的 HMM 参数估计流程

由于实际对 HMM 训练时,不同初始模型将产生不同的训练结果,而训练算法是在使  $P(O/\lambda)$  局部极大时得到模型参数,因此应选取适当初始模型,使最后求出的局部极大尽量接近全局极大. 因为通常  $\pi_i, A$  的初值选取对 HMM 的训练影响不大,而  $B$  的初值影响较大,所以本文训练中的初始修正也主要针对状态观测概率矩阵  $B$ . 图 7 中,在最初始的 HMM 中,  $A, B$  和  $\pi$  都采用随机阵列表示,且满足下式条件:

$$\begin{cases} 0 \leq \pi_i \leq 1 & \text{且} \sum_{i=1}^N \pi_i = 1; \\ 0 \leq a_{ij} \leq 1 & \text{且} \sum_{j=1}^N a_{ij} = 1; \end{cases} \quad (11)$$

然后根据各样本的观测序列,通过 Viterbi 算法

得到相应状态序列,再结合状态序列和观测序列,得到部分状态观测的统计频率值,用此统计频率值代替最初  $B$  中相应的元素,从而修正  $B$  中部分元素,对于未修正的元素,仍采用最初的随机数表示. 最后用这种经过部分修正的  $B$ ,代替最初的  $B$ ,结合原先的  $A, \pi$ ,作为初始化的 HMM,通过 Baum-Welch 算法进行训练. 修正  $B$  的示例如下:

假设对某个 HMM,给定观测序列  $O$  及最初在满足式(11)的前提下,通过随机设置给定  $A, B$  和  $\pi$ ,如下所示(假设状态种类数  $N$  为 3,观测种类数  $M$  为 3):

$$O: 1 \quad 3 \quad 2 \quad 2 \quad 3 \quad 2 \quad 1 \quad 3 \quad 2$$

$$A: \begin{matrix} 0.20 & 0.60 & 0.20 \\ 0.25 & 0.50 & 0.25 \\ 0.35 & 0.15 & 0.50 \end{matrix} \quad B: \begin{matrix} 0.25 & 0.25 & 0.50 \\ 0.20 & 0.30 & 0.50 \\ 0.45 & 0.05 & 0.50 \end{matrix}$$

$$\pi: 0.50 \quad 0.25 \quad 0.25$$

根据上述  $O, A, B, \pi$ ,通过 Viterbi 算法得出对应的状态序列  $Q$  为

$$Q: 1 \quad 2 \quad 2 \quad 2 \quad 2 \quad 2 \quad 2 \quad 2 \quad 2$$

结合上述观测序列  $O$  和对应状态序列  $Q$  可以看出:需对状态 1,2 的观测概率分布(对应上述矩阵  $B$  的前两行)进行修正,而状态 3 的观测概率分布保持不变,则对上述矩阵  $B$  修正如下:

$$B: \begin{matrix} 1.000 & 0.000 & 0.000 \\ 0.125 & 0.500 & 0.375 \\ 0.450 & 0.050 & 0.500 \end{matrix}$$

将新的矩阵  $B$ ,结合上述的  $O, A, \pi$  作为初始值,通过 Baum-Welch 多序列训练算法并结合前向-后向算法判断  $P(O/\lambda)$  的收敛性,即可得出模型参数  $A, B, \pi$  的最终估计值. 从原理上可看出,该方法要求样本的观测序列要尽量充分. 后面的实验会发现:在较充分的样本下,该训练方法得到使得对应的  $P(O/\lambda)$  接近全局最大的比较精确的 HMM:  $\lambda$ .

由于 HMM 实质上是一种统计意义上的有限状态机,当前它已被广泛应用于对时间序列的识别过程中<sup>[21]</sup>. 人的步态可看作是从一个具有结构化概率本质关系的序列中采样的一个姿态集合,这些被个体采样的姿态可作为 HMM 的状态,它们对于个体来说具有典型代表性,因此可作为个体的判别手段<sup>[12]</sup>.

结合本文研究目标,在步态识别中可以对每一个人建立一个 HMM:  $\lambda = (\pi, A, B)$ ,即用 3 种参数

综合表示一个人的步态身份。 $\mathbf{O}$  表示步态序列在某一时刻的特征观测向量。随着步态运动的延续, 产生  $\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_T$  等一系列的观测向量, 这里的目标就是通过这一系列的观测向量序列, 寻找产生这一观测序列在概率意义上最大的 HMM:  $\lambda$ 。如果不考虑先验知识  $P(\lambda)$ , 通过前向-后向算法所求得的最大  $P(\mathbf{O}/\lambda)$  对应的 HMM:  $\lambda$ , 将作为相应身份识别的结果。如果考虑先验知识  $P(\lambda)$ , 根据 Bayes 规则, 进一步求取:

$$\begin{aligned} P(\lambda_i/\mathbf{O}) &= P(\mathbf{O}/\lambda_i)P(\lambda_i)/P(\mathbf{O}) \\ &= P(\mathbf{O}/\lambda_i)P(\lambda_i) / \sum_{i=1}^M P(\mathbf{O}/\lambda_i)P(\lambda_i) \quad (12) \end{aligned}$$

所求得最大的  $P(\lambda/\mathbf{O})$  对应的 HMM:  $\lambda$  将作为相应身份识别的结果。可以看出, 考虑先验知识时, 此时, 某 HMM:  $\lambda$  产生某观测序列  $\mathbf{O}$  的概率  $P(\mathbf{O}/\lambda)$  最大时, 也并不意味着该观测序列属于该模型的概率  $P(\lambda/\mathbf{O})$  最大, 它还与  $P(\lambda)$  有关。联系到步态识别中, 某人产生某种步态特征观测序列的概率  $P(\mathbf{O}/\lambda)$  最大时, 也并不意味着该观测序列属于该人的概率  $P(\lambda/\mathbf{O})$  最大, 还与该人以往出没于该场景的概率  $P(\lambda)$  即先验知识有关。由此看出, 先验知识的涉及将使得识别出的身份更加客观合理, 后面的实验将进一步验证该性质。

当然, 此处要求事先对检测场景的以往历史中, 出没的各种身份的人的概率进行统计, 统计方式为多种。计算某人的出没概率时, 通常有两种方式: (1) 可按照某段时间内, 目标出现的时间长度所占比率来表示。(2) 可按照某段人流量中, 目标出现的次数所占比率表示。上述统计还要求可在线更新。这里, 在步态序列训练的过程中, 同时对应用的 Bayes 规则中的相关先验  $P(\lambda_i)$  也进行了学习, 其方法原理如下:

假设总目标数为  $M$  个, 最初每个被识别目标的训练样本数目相同, 皆为  $m$  个, 即各目标出入场景的次数均等, 皆为  $m$  次, 根据统计学中的贝努里定理, 事件发生的频率以概率收敛于事件的概率, 即当试验次数很大时, 事件发生的频率与概率有较大偏差的可能性很小, 此时, 便可用事件发生的频率来代替事件的概率。而本文的训练样本, 可认为是各目标的全部历史事件, 即可认为试验次数充分大, 则各目标出入场景的概率  $P(\lambda) = m/(mM) = 1/M$ , 若对某人( $i$ )的目标的训练样本数目增加了  $\Delta m$  个, 则此时对应该目标( $i$ )的出入场景概率  $P(\lambda_i) = (m + \Delta m)/(mM + \Delta m)$ , 其余目标( $j$ )的出入场景概率

$P(\lambda_j) = m/(mM + \Delta m)$ ,  $j \neq i$ 。反之, 若某人( $i$ )的训练样本数目减少了  $\Delta m$  个, 则将上述式中加号换成减号即可。对于其它目标的样本数目变化情况, 可依次类推, 这样就实现了对所有目标的对应先验  $P(\lambda)$  的学习。

综上所述, 本文对步态特征序列训练及识别的理论框架如下:

(1) 步态序列训练时, 对随机初始化的观测矩阵  $\mathbf{B}$  通过 Viterbi 算法, 结合相应观测序列  $\mathbf{O}$ , 进行统计修正, 然后用 Baum-Welch 算法训练出与各类观测序列对应的 HMM:  $\lambda_i = (\pi_i, A_i, B_i)$ ,  $i = 1, \dots, M$  中的相关参数, 同时根据各类  $\lambda_i$  的样本数目, 对 Bayes 规则相关的先验知识  $P(\lambda_i)$  进行学习, 最后对训练出的各类  $\lambda_i$  的相关参数及相关先验  $P(\lambda_i)$  建立数据库 Database。

(2) 步态序列识别时, 对前面训练中建立的 Database 库中的每一类  $\lambda_i$ , 结合相应被测试的观测序列  $\mathbf{O}$ , 通过前向-后向算法得出  $P(\mathbf{O}/\lambda_i)$ 。

(3) 将上述结果结合 Database 中的相关先验知识  $P(\lambda_i)$ , 再通过 Bayes 规则(即式(12))得出  $P(\lambda_i/\mathbf{O})$ 。

(4) 比较各  $P(\lambda_i/\mathbf{O})$ , 其中最大  $P(\lambda_i/\mathbf{O})$  对应的  $\lambda_i$  其对应识别出的身份结果。

### 3 实验测试方案及结果分析

在对整个检测及识别框架的测试中, 本文以 Pentium1.73 GHz 单 CPU 的台式机为平台, 在 VC6.0 环境下编程实现其功能, 通过标准的原始视频步态数据库, 模拟测试在真实环境下对监控通道中实际目标的检测识别。最终, 将相应结果的统计数据进行了曲线描绘, 并进行了相关分析。

本文的测试视频库采用的是中国科学院的 CASIA 步态数据库中的 Dataset B<sup>[22]</sup>。该数据集是一个大规模多视角的步态库, 采集于 2005 年 1 月, 共有 124 人, 每人有 11 个视角( $0^\circ, 18^\circ, 36^\circ, \dots, 180^\circ$ ), 在 3 种条件下(穿较单薄的春夏季服装, 穿较厚重的冬季服装及携带包裹)行走。其中第 1 种每人 6 个样本, 第 2 和第 3 种每人各 2 个样本, 因为本文主要研究基本步态, 所以主要针对未携带任何物品的人的行走步态, 即对前两种行走条件进行识别。

试验中, 对步态数据库 B 中的原始视频, 选择样本数 124 人(即全部人数), 视角选择是在  $18^\circ \sim 162^\circ$  范围内以  $18^\circ$  为间隔划分成几个常见视角。这里

分为样本模型的先验知识相同和不同两种情况分别测试. 前者主要测试本文的训练 HMM 方案所对应的结果模型的精确性及最初的检测环节对识别效果的影响, 同时验证前向-后向算法的精度, 后者主要测试前者在融合了 Bayes 规则的客观合理性后对识别率的影响.

(1) 样本模型的先验知识相等的情况. 此时 Bayes 规则对各模型的影响均等, 因此从某种意义上, 可认为 Bayes 规则此时不起作用, 影响最终结果相对大小的, 除最初检测环节外, 识别环节只有前向-后向算法起作用. 这里主要进行两项测试, 并与文献[12-13, 23-24]的典型方法进行比较, 而且这些文献又各有两三种方法, 这里选择它们各自当中识别率最大的方法对应结果进行比较.

表 1 先验知识相等时第 1 种行走条件的不同视角下对应识别率

方法	识别率/%									
	18°	36°	54°	72°	90°	108°	126°	144°	162°	
本文方法	95.51	97.76	96.3	93.6	90.55	92.13	95.45	97.54	94.23	
文献[12]方法	91.12	92.03	92.8	92.96	93.75	92.89	93.58	93.37	93.27	
文献[13]方法	90.75	92.47	92.3	93.59	92.35	93.22	93.17	92.69	91.93	
文献[23]方法	80.03	77.86	76.9	79.12	83.04	80.21	78.78	77.93	79.78	
文献[24]方法	97.43	97.54	98.1	97.27	97.67	98.32	98.25	98.08	97.86	

表 2 先验知识相等时第 2 种行走条件的不同视角下对应识别率

方法	识别率/%									
	18°	36°	54°	72°	90°	108°	126°	144°	162°	
本文方法	80.35	86.24	81.20	78.65	72.46	75.76	83.23	85.88	77.69	
文献[12]方法	10.44	12.09	11.60	10.45	12.02	9.79	9.93	11.64	8.89	
文献[13]方法	7.08	8.57	6.89	8.35	11.04	9.53	7.79	6.56	6.35	
文献[23]方法	9.78	10.12	10.40	11.16	12.46	9.25	8.87	5.35	7.67	
文献[24]方法	22.34	30.56	25.10	30.02	36.23	25.68	26.95	29.73	27.46	

结合图 8 和图 9, 对比各视角可看出: 视角在 36° 和 144° 附近时, 如图 9(b)、(h) 所示, 被测试的人的正面轮廓和侧面轮廓, 几乎同时可以较完备地作为该人的整体轮廓观测出来. 这样正面的手臂和肩

第 1 项测试是在第 1 种行走条件下采用留一校验方法 (leave-one-out cross validation), 取每人 6 个样本的中的 5 个用于训练, 第 6 个用于识别, 求出识别率, 然后调换样本顺序, 重复实验, 最后求出识别率平均值; 第 2 项测试是将第 1 种行走条件下全部样本用于对 HMM 训练, 第 2 种行走条件下全部样本用于识别, 然后求出识别率, 以测试该方法对个体衣着的鲁棒性.

相应测试结果如表 1、表 2 所示, 图 8 为对应两表的相关变化曲线. 可以看出: 本文方法在多个视角下的识别率总体普遍较高. 这说明前面训练出的 HMM 具有较强的识别泛化能力, 使得对应的  $P(O/\lambda)$  逼近了全局极大. 也说明本文的检测识别框架具有一定合理性.

部的左右摆动及侧面两腿的前后摆动, 都会影响此时的整体轮廓大小、形状的变化, 从而影响轮廓的重心位置的变化. 也就是说, 此时目标的重心特征序列的变化, 相对其它视角, 包含了更多的人体步态运动

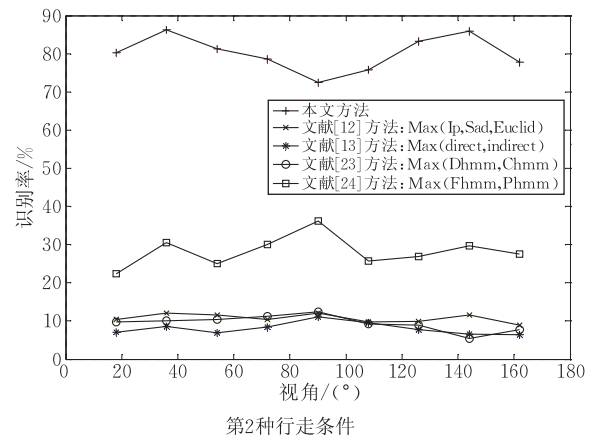
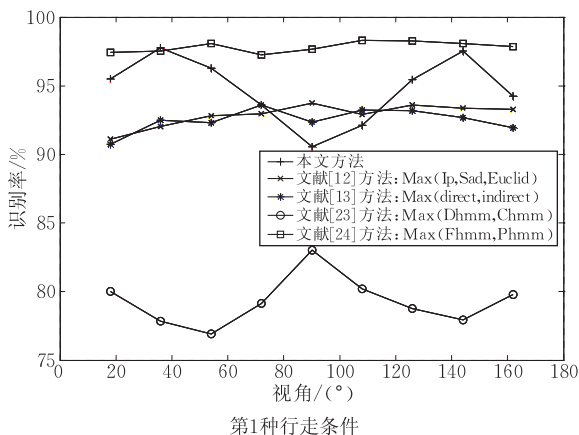


图 8 Bayes 规则先验知识相等时不同视角下的识别率比较

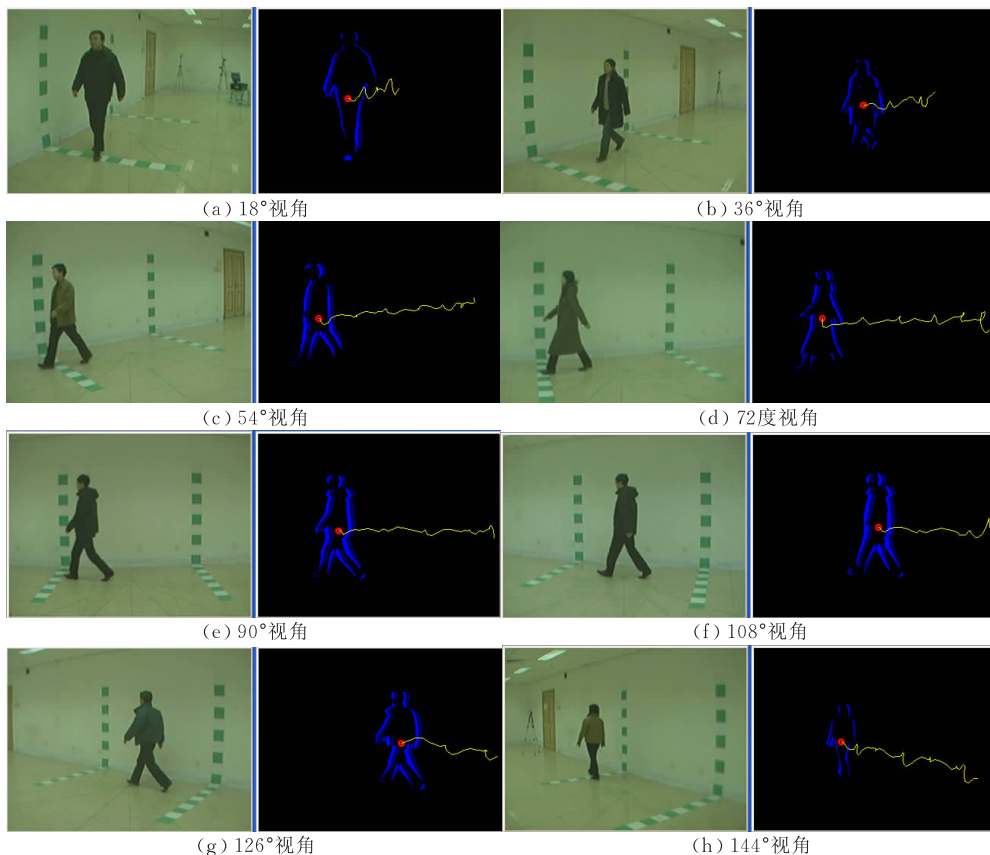


图 9 不同视角下目标步态的重心特征曲线检测图

正面和侧面两方面的个性化特征信息,个性化信息越强,从而使得特征序列间的区别增大,所以此时识别率相对其它视角较高。

视角在  $90^\circ$  附近时,如图 9(e)所示,由于主要是侧面轮廓影响重心特征变化,目标正面的步态特征如手臂,肩部在正面的摆动几乎未刻画出来,从而特征间的差异相对其它视角有所减小,识别率较低。

实验测试中还发现,人的步态多样性主要表现在每个人的步态对称性是不一样的。有些人的步态对称性在通常的视角范围基本可以观测到(具体又在幅度、周期及多种波的交替往复上存在差异),有些人的步态对称性在某一非常规方向才可以观察到,此时在常规方向的观测结果成杂乱无章的序列,从而使得这类人的步态在常规方向无法进行识别。当样本数量增多时,这类人出现的数量随着增大,从而使得识别率降低。因此可以设想:如果将步态特征曲线在三维空间中描绘时,可能会对各类人的步态得到更精确且更合理的特征描述,也将会使得试验对于大样本的识别率有所上升。

在和其它典型同类方法的比较中,除在特征检测环节中各有不同外,文献[12]用背景减除的二值化图像作为特征向量,用分别基于向量差分正则化

的 1 范数(SAD), 2 范数(Euclid)和向量正则化内积(IP)的 3 种不同距离矩阵,来度量特征向量间的相似度,从而估计相关的观测序列,用于 HMM 判别。文献[13]利用二值化侧面轮廓作为图像特征,分别采用直接(direct)训练 HMM 和通过生成低维的帧与样本的距离向量 FED 间接(indirect)训练 HMM 两种训练方法。两文献在对  $\mathbf{B}$  的初始修正都是在某种事先确定的函数下选取,最终识别率虽在不同视角下变化较小,但总体水平次于本文方法。文献[23]用自组织映射的码本向量及经过 PCA 降维后的向量两种特征向量分别来训练离散 HMM(Dhmm)和连续 HMM(Chmm),且未进行任何初始化,其识别结果总体上明显低于本文方法。文献[24]其特有的参数形 HMM(Fhmm)特征级融合与并行 HMM(Phmm)决策级融合的结构,很大程度上补偿了初始化  $\mathbf{B}$  的影响,这使得其在表 1 中总体效果优于本文方法,但在表 2 中的结果同其它 3 个被比较文献皆次于本文方法,即对目标衣着变化的鲁棒性方面性能均较弱。图 8 中的 Max 表示选取实验中识别率最大的文献对应处理方式进行比较。

(2) 样本模型的先验知识不等时,按前文所述,此时各 HMM 对应的目标出没于同一场景的概率

不同. 这里对前述相等时的两项试验中那些检测出的相似度较大的特征曲线对应的测试样本进行区分(因为相似度较小时, 测试样本的差异较大, 通过前向-后向算法区别已经很显著, 此时, Bayes 规则虽然有作用, 但无法直观看出, 即此时起作用的因素很多, 无法验证是哪个起作用), 以针对性地验证 Bayes 规则的合理性. 实验中将前面试验中部分识别错误的目标对应的训练样本数目改变, 在对相应的 HMM 重新训练的同时, 按第 2 节中先验的学习原理, 此时需同时将所有训练出的 HMM 的先验知识概率按前述计算方法进行相应改变, 最后全部测

试样本再次进行识别, 并与之前进行比较.

结果如表 3 和表 4 所示, 图 10 为相应比较图. 可看出, 在 Bayes 规则中先验不同情况下, 从总体样本的测试来看, 相比先验相同情况, 识别率普遍提高, 分析如下:

根据各 HMM 对应的先验知识, 当某人(对应  $\lambda$ )在过去的的时间里, 出没于某场景的概率较大时, 则  $P(\lambda)$  较大, 此时若该目标产生当前特征观测序列  $O$  的概率  $P(O/\lambda)$  也较大时, 式(12)中右式的值变大, 则对应的  $P(\lambda/O)$  值变大, 即该观测序列属于该目标的概率增大.

表 3 第 1 种行走条件的不同视角下 Bayes 规则先验变化的识别率比较

Bayes 影响	识别率/%								
	18°	36°	54°	72°	90°	108°	126°	144°	162°
Bayes 规则先验相同	95.51	97.76	96.3	93.6	90.55	92.13	95.45	97.54	94.23
Bayes 规则先验不同	96.79	99.01	97.45	96.69	93.56	95.44	96.37	98.82	97.58

表 4 第 2 种行走条件的不同视角下 Bayes 规则先验变化的识别率比较

Bayes 影响	识别率/%								
	18°	36°	54°	72°	90°	108°	126°	144°	162°
Bayes 规则先验相同	80.35	86.24	81.2	78.65	72.46	75.76	83.23	85.88	77.69
Bayes 规则先验不同	88.21	94.51	91.25	88.37	82.56	85.15	88.12	92.22	87.49

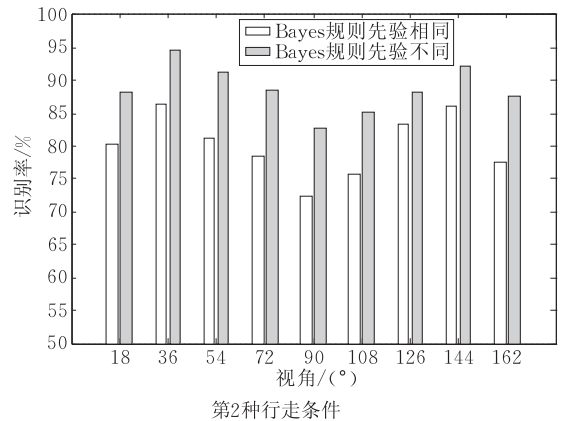
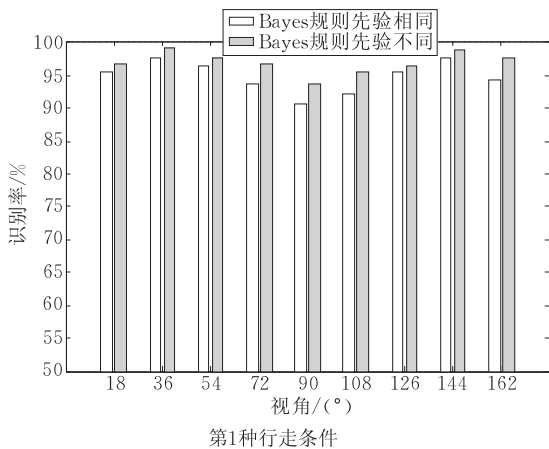


图 10 Bayes 规则先验变化时不同视角下的识别率比较

考虑到实际监控环境中, 被检测的各个目标在很多情况下其相关 Bayes 先验大都是不同的, 因此用实际信息与历史信息相结合的方法来推断当前, 使得识别结果更加客观而准确, 从而提高了目标识别率. 而此时个别的识别错误, 主要是由于在前面环节中, 对个别目标的检测部分及后续的前向-后向算法的误差相对偏大, 使得 Bayes 先验概率的补偿有所降低造成.

至于在某些情况下是否会使识别率降低, 从统计概率理论的意义上来讲, 若先验知识准确可靠, 则融入 Bayes 规则后, 识别率应该是可靠性增强, 即识别

率提高, 若先验知识不可靠, 则无法保证识别结果的正确性. 而本文是在假设具有可靠先验知识的前提下进行研究的.

## 4 结 论

本文通过对步态视频中运动目标的重心变化特征信息的获取及相关滤波、标准化、规则化、聚类等一系列处理后, 根据 Baum-Welch 算法训练过的 HMM, 通过前向-后向算法, 结合具有训练学习过的先验知识的 Bayes 规则实现了人体步态的身份识

别,由于融入 Bayes 规则后识别的客观性质增强,加上检测环节中重心特征曲线的检测减轻了时间差分算法目标中空的影响,检测出的曲线拟合的等效旋转简化了计算,降低了失真度。这些使得整个框架对直道类型的监控通道,具有较高的识别率,且对衣着具有一定的鲁棒性。由于从单个角度的获取步态信息仍具有局部性,本文的下一步工作将从多个角度同时获取步态信息,并对多种信息进行实时三维刻画及融合来对目标身份进行更准确的识别,并进行相关的行为理解。

### 参 考 文 献

- [1] Zhao Guo-Ying. Video-based Gait Recognition [Ph. D. dissertation]. Institute of Computing Technology, Chinese Academy of Sciences, Beijing, 2005(in Chinese)  
(赵国英. 基于视频的步态识别[博士学位论文]. 中国科学院计算技术研究所, 北京, 2005)
- [2] Sun Xiaoying, Zhang Qihong, Xu Yanqun. Application of segmentation based on optical flow for gait recognition//Proceedings of the 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE). Chengdu, 2010; 567-571
- [3] Yang Run-Hui, Wu Qing-Jiang. Research on gait-based identity recognition. Microcomputer Application, 2007, 23(12): 8-10(in Chinese)  
(杨润辉, 吴清江. 基于步态的身份识别研究. 微型电脑应用, 2007, 23(12): 8-10)
- [4] Srivastava S, Sural S. Human gait recognition using temporal slices. Pattern Recognition and Machine Intelligence, 2007, 4815: 592-599
- [5] Cunado D et al. Gait extraction and description by evidence-gathering//Proceedings of the International Conference on Audio and Video Based Biometric Person Authentication. Washington DC, USA, 1999; 43-48
- [6] Zhao G et al. 3D gait recognition using multiple cameras//Proceedings of the 7th IEEE International Conference on Automatic Face and Gesture Recognition (FGR'06). Southampton, 2006; 529-534
- [7] Wang L et al. Silhouette analysis-based gait recognition for human identification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(12): 1505-1518
- [8] Ma Qin-Yong, Nie Dong-Dong, Wang Shen-Kang. Gait expression and recognition based on primary motion contours. Acta Automatica Sinica, 2009, 35(5): 519-525(in Chinese)  
(马勤勇, 聂栋栋, 王申康. 基于主运动轮廓线的步态表示与识别. 自动化学报, 2009, 35(5): 519-525)
- [9] Wang Liang, Hu Wei-Ming, Tan Tie-Niu. A survey of visual analysis of human motion. Chinese Journal of Computers, 2002, 25(3): 225-237(in Chinese)  
(王亮, 胡卫明, 谭铁牛. 人运动的视觉分析综述. 计算机学报, 2002, 25(3): 225-237)
- [10] Lee L, Grimson W E L. Gait analysis for recognition and classification//Proceedings of 5th IEEE International Conference on Automatic Face and Gesture Recognition. Washington, DC, USA, 2002; 148-155
- [11] Lam T H W, Lee R S T, Zhang D. Human gait recognition by the fusion of motion and static spatio-temporal templates. Pattern Recognition, 2007, 40(9): 2563-2573
- [12] Sundaresan A, RoyChowdhury A, Chellappa R. A hidden Markov model based framework for recognition of humans from gait sequences//Proceedings of the International Conference on Image Processing. Barcelona, Catalonia, Spain, 2003; 3(II-93-6)
- [13] Kale A et al. Identification of humans using gait. IEEE Transactions on Image Processing, 2004, 13(9): 1163-1173
- [14] Boulgouris N V, Plataniotis K N, Hatzinakos D. Gait recognition using dynamic time warping//Proceedings of the IEEE International Symposium Multimedia Signal Processing. Siena, Italy, 2004; 263-266
- [15] Lee H, Hong S, Kim E. Neural network ensemble with probabilistic fusion and its application to gait recognition. Neurocomputing, 2009, 72(7/9): 1557-1564
- [16] Begg R K, Palaniswami M, Owen B. Support vector machines for automated gait classification. IEEE Transactions on Biomedical Engineering, 2005, 52(5): 828-838
- [17] Collins R et al. A system for video surveillance and monitoring: VSAM final report. Technical Report: CMU-RI-TR-00-12, 2000
- [18] Widrow B, Stearns S D. Adaptive Signal Processing, Upper Saddle River, NJ, USA; Prentice-Hall, Inc., 1985; 474
- [19] Rabiner L R. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 1989, 77(2): 257-286
- [20] Xie Jin-Hui. HMM and Its Application in Speech Processing. Wuhan: Huazhong University of Science and Technology Press, 1995(in Chinese)  
(谢锦辉. 隐 Markov 模型(HMM)及其在语音处理中的应用. 武汉: 华中理工大学出版社, 1995)
- [21] Wang Tian-Shu, Zheng Nan-Ning, Xu Ying-Qing, Shum Heung-Yong. Unsupervised clustering analysis of human motion. Journal of Software, 2003, 14(2): 209-214(in Chinese)  
(王天树, 郑南宁, 徐迎庆, 沈向洋. 人体运动非监督聚类分析. 软件学报, 2003, 14(2): 209-214)
- [22] Yu Shi-Qi, Tan Dao-Liang, Tan Tie-Niu. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition//Proceedings of the 18th International Conference on Pattern Recognition (ICPR 2006). Hong Kong, China, 2006; 441-444
- [23] Suk H I, Sin B K. HMM-based gait recognition with human profiles//Proceedings of the Structural, Syntactic, and Statistical Pattern Recognition. Lecture Notes in Computer Science; 4109/2006. Hong Kong, China, 2006; 596-603
- [24] Chen C et al. Factorial HMM and parallel HMM for gait recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2008, 39(1): 114-123



**YU Tao**, born in 1974, Ph.D. candidate. His current research interests include computer vision, pattern recognition, digital image processing and analysis, machine learning, multimedia, visual surveillance, etc.

**ZOU Jian-Hua**, born in 1964, professor, Ph.D. supervisor. His current research interest include image processing, machine and computer vision, pattern recognition, multimedia, complex systems analysis and networked control systems.

### Background

Recently, gait recognition, as a kind of identification recognition from far distance without contacting, has been received more and more attention at the safe-guard monitoring area in countries all over the world. Many universities and research institutes have carried out the associated studies in this direction. At present, in the area the computer vision, from object's detecting, tracking to recognition, all the problems are not solved completely at each tache. So, most frameworks for gait recognition, bypassing the real-time and accurate demands at phases of detecting and tracking, used the preprocessed segmented moving object and studied in multi-parameter and multi-dimension space. As the problems of time-sequence becomes a hot spot for the analysis and recognition of moving actions, the authors attempted to start from the original moving videos directly. At the phase of detecting and tracking, the important traits of relative change in

the time-sequences on motion and the important value of prior objective background are noticed and applied. Finally, the Bayes rules and HMM are combined to recognize identification. The associated testing results are encouraging. For the more sophisticated motional actions and associated analysis, the authors will carry out more deep exploring study.

This paper's study received support by National Natural Science Foundation of China (50177025). The authors are studying the direction with the associated working group on computer vision in the state key laboratory for systems engineering, systems engineering institute, Xi'an Jiaotong University. The work of this paper will be regarded as an important part with others' works, such as face detection and crowd detection, to be combined and applied in monitoring area.