

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012 7
Vol.17 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年7月 第17卷 第7期(总第195期)

目次

综述

中国多媒体技术研究: 2011 史元春, 徐光祐, 高原(741)

图像处理和编码

图像引导滤波的局部多尺度 Retinex 算法 方帅, 杨静荣, 曹洋, 武鹏飞, 饶瑞中(748)

基于第一主成分方向稳定性的图像零水印 付剑晶, 王珂(756)

小波变换估计非线性扩散最优停止时间 蒋平, 张建州(770)

图像分析和识别

改进的 Beamlet 与 Canny 相结合提取复杂图像线特征 曾接贤, 周沥沥, 符祥(775)

动静态信息融合及动态贝叶斯网络的步态识别 杨旗, 薛定宇(783)

融合音频单词与视觉特征的成人视频检测 刘毅志, 唐胜, 王向东, 林守勋, 张勇东(791)

基于平面区域跟踪的目标位姿参数自动测量 回丙伟, 文贡坚, 赵竹新, 钟金荣(798)

3 维图像中边界曲面的分类追踪及抽取 丁德福, 程柳航, 王利生(806)

局部时空域模型的核密度估计目标检测方法 王兴宝, 刘纯平, 费兰英, 王朝晖, 季怡(813)

基于相位谱和调谐幅度谱的显著性检测方法 李崇飞, 高颖慧, 卢凯, 曲智国(821)

图像理解和计算机视觉

“目标-场景”语境关联的生成图模型分析 谢昭, 李姍琦, 高隼(828)

融合上下文信息的场景结构恢复 武晖, 于昕, 隋尧, 张利(839)

带有局部控制因子的图割光流估计 路子赟, 唐土生, 高隼, 沈琳, 刘伟(846)

改进的 Harris 亚像素角点快速定位 何海清, 黄声享(853)

计算机图形学

应用最小生成树实现点云分割 孙金虎, 周来水, 安鲁陵(858)

医学图像处理

病毒进化的离散差分进化超声图像特征选择 张巧荣, 朱长明, 倪军, 刘海波(866)

分水岭优化的 Snake 模型肝脏图像分割 兰红, 张璐(873)

遥感图像处理

Normalized Cut 与分水岭变换在高光谱影像混合像元端元提取中的应用 许菡, 李小娟(880)

改进非局部均值滤波的 SAR 图像降噪方法 郑永恒, 程建, 曹宗杰(886)

~~~~~  
第四届国际遥感考古会议 ..... (892)

**中国图象图形学报**

刊名题字: 宋 健

月刊(1996 年创刊)

第 17 卷 第 7 期

2012 年 7 月 16 日出版

**主管单位** 中国科学院  
**主 办** 中国科学院遥感应用研究所  
中国图象图形学学会  
北京应用物理与计算数学研究所  
**主 编** 李小文  
**编辑出版** 《中国图象图形学报》编辑出版委员会  
北京 9718 信箱 邮编 100101  
电子信箱:jig@irsa.ac.cn  
电话:010-64807995 010-82614429  
网 址:www.cjig.cn  
**印刷装订** 北京北林印刷厂  
**广告经营许可证** 京朝工商广字第 0346 号  
**总 发 行** 北京报刊发行局  
**订 购** 全国各地邮局  
**国外发行** 中国国际图书贸易总公司  
(中国国际书店)  
(北京 399 信箱 邮编 100044)

**Superintended by** Chinese Academy of Sciences  
**Sponsored by** Institute of Remote Sensing Application,  
CAS China Society of Image and Graphics  
Institute of Applied Physics and Computational  
Mathematics  
**Chief editor** LI Xiaowen  
**Editor, Publisher** Editorial and Publishing Board  
of Journal of Image and Graphics  
(P. O. Box 9718, Beijing 100101, China)  
E-mail:jig@irsa.ac.cn  
**Distributed by** Beijing Bureau for Distribution of Newspapers  
and Journals  
**Domestic** All Local Post Offices in China  
**Foreign** China International Book Trading Corporation  
(P. O. Box 399, Beijing 100044, China)  
**Printed by** Beijing Beilin Printing House

ISSN 1006-8961 CN11-3758/TB CODE ZTTFXZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00 元

# Journal of Image and Graphics

( Monthly , Started in 1996 )

Vol. 17 No. 7 July 2012

## Contents

### Review

Researches on multimedia technology in China, 2011 ..... Shi Yuanchun, Xu Guangyou, Gao Yuan (741)

### Image Processing and Coding

Local multi-scale Retinex algorithm based on guided image filtering  
..... Fang Shuai, Yang Jingrong, Cao Yang, Wu Pengfei, Rao Ruizhong (748)

Image zero-watermark based on direction stability of first principal component vector ..... Fu Jianjing, Wang Ke (756)

Stopping-time estimation for anisotropic diffusion using discrete wavelet transform ..... Jiang Ping, Zhang Jianzhou (770)

### Image Analysis and Recognition

Complex image line feature extraction based on improved Beamlet transform and the Canny operator  
..... Zeng Jiexian, Zhou Lili, Fu Xiang (775)

Gait recognition based on dynamic & static information fusion and dynamic bayesian network ..... Yang Qi, Xue Dingyu (783)

Fusing audio-words with visual features for adult video detection  
..... Liu Yizhi, Tang Sheng, Wang Xiangdong, Lin Shouxun, Zhang Yongdong (791)

Automatic measurement for an object's position and attitude via tracking planar regions  
..... Hui Bingwei, Wen Gongjian, Zhao Zhuxin, Zhong Jinrong (798)

Detection and extraction of boundary surface patches within 3D images ..... Ding Defu, Cheng Liuhang, Wang Lisheng (806)

Foreground object detection method using kernel density estimation of a local spatio-temporal model  
..... Wang Xingbao, Liu Chunping, Fei Lanying, Wang Zhaohui, Ji Yi (813)

Saliency detection method based on phase spectrum and amplitude spectrum tuning  
..... Li Chongfei, Gao Yinghui, Lu Kai, Qu Zhiguo (821)

### Image Understanding and Computer Vision

"Object-Scene" contextual associated generative graph model analysis ..... Xie Zhao, Li Shanqi, Gao Jun (828)

Structure recovery algorithm using contextual information ..... Wu Hui, Yu Xin, Sui Yao, Zhang Li (839)

Graph cut optical flow estimation with a local control factor ..... Lu Ziyun, Tang Tusheng, Gao Jun, Shen Lin, Liu Wei (846)

Improved algorithm for Harris rapid sub-pixel corners detection ..... He Haiqing, Huang Shengxiang (853)

### Computer Graphics

Research on point cloud segmentation using a minimum spanning tree ..... Sun Jinhu, Zhou Laishui, An Luling (858)

### Medical Image Processing

Virus-evolutionary discrete differential evolution algorithm for feature selection of cervical lymph nodes in ultrasound images  
..... Zhang Qiaorong, Zhu Changming, Ni Jun, Liu Haibo (866)

Liver image segmentation algorithm based on the Snake model And optimized by watershed transformation  
..... Lan Hong, Zhang Lu (873)

### Remote Sensing Image Processing

Endmember extraction for hyperspectral image based on normalized cut and watershed transformation  
..... Xu Han, Li Xiaojuan (880)

SAR image denoising via improved non-local means filter ..... Zheng Yongheng, Cheng Jian, Cao Zongjie (886)

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2012)07-0783-08

论文引用格式: 杨旗, 薛定宇. 动静态信息融合及动态贝叶斯网络的步态识别[J]. 中国图象图形学报, 2012, 17(7): 783-790.

## 动静态信息融合及动态贝叶斯网络的步态识别

杨旗<sup>1,2</sup>, 薛定宇<sup>1</sup>

1. 东北大学信息科学与工程学院, 沈阳 110819; 2. 沈阳理工大学机械工程学院, 沈阳 110159

**摘要:** 步态是远距离视频监控领域最具潜力的生物特征。目前对步态的识别研究大都是考虑单一条件下步态的识别率,但在穿外套、背包等混合条件下识别率较低,通过分析人体行走时步态的时序特征,提出一种基于动静态信息相结合的多信息融合的动态贝叶斯网络(DSIF-DBN)。模型含有3层状态,模型中每个时间片都为静态信息和动态信息的融合。此模型能很好地表达步态的时序特性,即步态行走时人体姿态,运动幅度等特征的节奏性变化。实验结果表明该方法有较高的识别率,能有机地融合步态的静态信息及动态信息,并且在有噪声及信息缺失的情况下有较好的鲁棒性,大大降低了外套及背包对步态识别的影响。

**关键词:** 步态识别;视频监控;动态贝叶斯网络;信息融合

### Gait recognition based on dynamic & static information fusion and dynamic bayesian network

Yang Qi<sup>1,2</sup>, Xue Dingyu<sup>1</sup>

1. School of Information Science & Engineering, Northeastern University, Liaoning province Shenyang 110819, China;
2. School of Mechanical Engineering, Shenyang Ligong University, Liaoning province Shenyang 110159, China

**Abstract:** Gait is an important biological characteristics in the long distance video surveillance field. Nowadays, almost all gait recognition researcher focus on gait recognition only under one single condition. However, the gait recognition rate rapidly decline in blended conditions, for example when somebody is wearing a coat or carrying a bag. Based on our analysis of the gait timing characteristics during the human movements, we propose a new gait recognition approach that expresses dynamic information and static information by using a dynamic Bayesian network (DSIF - DBN). The DSIF - DBN contains three levels of states and for every time slice of the DSIF - DBN model is expressed by the fusion of dynamic information and static information. This model can exactly express the timing characteristics of the gait, which are the body posture and the range of motion, as well as other gait rhythmic change characteristics. Experimental result show that the DSIF - DBN model recognizes gait with high rates and good robustness to noise and lost of information. The DSIF - DBN model can fuse the dynamic information as well as static information and can greatly reduce the impact of gait recognition rates when somebody is wearing a coat or carrying a bag.

**Key words:** gait recognition; video surveillance; dynamic Bayesian network; information fusion

收稿日期: 2011-09-26; 修回日期: 2012-01-04

基金项目: 国家自然科学基金项目(61174145)

第一作者简介: 杨旗(1976—), 男, 讲师, 东北大学模式识别与智能系统专业博士研究生, 主要研究方向为数字图像处理、模式识别与智能处理。E-mail: 9501133@163.com

## 0 引言

作为远程生物特征认证技术,步态识别越来越受到人们的重视。近些年来对步态识别的研究日益增多,大都是通过提取人行走的步态轮廓后进行识别,技术可分为两类;一是利用步态轮廓的静态信息,如 Kim 提出了基于主动轮廓模型及运动预测的步态识别<sup>[1]</sup>;二是采用提取轮廓的动态信息,进行这类研究的算法很多,如基于步态能量图像(GEI)和 2 维主成分分析的步态识别方法<sup>[2]</sup>,由于步态能量图像(GEI)受穿着影响较大,故识别率较低。为了改善 GEI 图像丢失太多的动态信息,Zhang 提出了基于主动能量图(AEI)及 2 维局部投影的方法进行识别<sup>[3]</sup>,其中 AEI 图像是由帧差图像叠加构成,可很好地反应人体运动时的动态步态特征。为了保留步态行走时的静态信息,Chen 提出了对提取的不完全步态轮廓采取建立帧差能量图(FDEI)的方式构建步态特征,同时建立隐马尔可夫模型(HMM)进行表述<sup>[4]</sup>,识别效果好。Chen 还提出了基于双层的动态贝叶斯网络的步态识别(DBN)<sup>[5]</sup>,首先把整个步态序列分若干段,第 1 层用动态纹理表达,第 2 层用隐马尔可夫模型表达,方法新颖,充分考虑了步态时序特性。Faezeh 提出了基于手臂和腿部运动的步态识别,分析腿部和手臂的运动来识别步态<sup>[6]</sup>。步态识别的识别步骤大都从建立步态的特征图像着手,利用步态的特征图像进行特征的提取识别。识别方法如基于特征匹配的方法<sup>[7-10]</sup>,基于动态贝叶斯网络的方法<sup>[5,11]</sup>,其中基于动态贝叶斯网络的方法是当前研究的热点,它充分考虑了步态的时序特性。而隐马尔可夫模型(HMM)及其改进模型作为动态贝叶斯网络的一种形式得到了广泛应用<sup>[12-13]</sup>,如耦合隐马尔可夫模型(CHMM)<sup>[14]</sup>,嵌入式马尔可夫模型(EHMM)。用于步态识别的基于嵌入式隐马尔可夫模型(EHMM)<sup>[15]</sup>张二虎提出的利用动态部位变化的步态识别<sup>[16]</sup>方法融入了时序的特性,用基于最大熵马尔可夫模型(MEMM)作为步态分类器,识别率高,模型的鲁棒性也较好。

综上,上述方法虽然表达了人体行走的步态动态信息,但是大都是把动态信息的总和叠加构成一个静态的图像作为识别的特征图像,只有少量方法考虑了步态行走时的时序特性,即随着时间的推移,步态行走姿势及运动幅度的变化,而这些特征恰恰

反映了一个人行走的独有特征,即步态行走时人体姿态、运动幅度等特性的节奏的变化。这些特征都表现为时序特性。本文提出了一种基于动静态信息相融合的动态贝叶斯网络(DSIF-DBN),在此网络中含用 3 层状态,模型中每个时间片都为静态信息和动态信息的融合,而每层状态分别反映了步态的静态信息或动态信息,不同层状态之间反映了动静态信息之间的关系。此模型能很好地表达步态的时序特性,即步态行走时人体姿态、运动幅度等特性的节奏的变化,能有机地融合步态的静态信息及动态信息,进行建模及识别。

## 1 动静态信息融合的动态贝叶斯模型

步态行走是个时序的过程,在 DSIF-DBN 模型中在单位时间片  $t$  内,不仅包含由步态序列帧表达的外形轮廓静态信息,还包含由帧差图像表达的运动幅度、速度等人体行走时的节奏性动态信息。在动态信息的表达中,人在步态行走时,通常表现为左右肢体交替性的有节奏的摆动。这种摆动从人体侧面  $90^\circ$  视角观察时,表现为人体的左右部分由前向后,再由后向前的运动过程。在此过程步态帧差中:帧差的前向部分和帧差的后向部分表达了人体运动左右部分的动态运动特性(如图 1),而通常的步态帧差图大都提取帧差的前向部分作为帧差图。在此定义:帧差的前向部分称为前向帧差图,帧差的后向部分称为后向帧差图,数学表达如下。

前向帧差图为

$$Fa = \begin{cases} 0 & B(x, y, t) \leq B(x, y, t-1) \\ B(x, y, t) - B(x, y, t-1) & \text{其他} \end{cases} \quad (1)$$

后向帧差图为

$$Fb = \begin{cases} 0 & B(x, y, t) \geq B(x, y, t-1) \\ B(x, y, t-1) - B(x, y, t) & \text{其他} \end{cases} \quad (2)$$

帧差图像为

$$F = Fa + Fb \quad (3)$$

式中,  $B(x, y, t)$  为步态帧在  $t$  时刻的图像  $x, y$  坐标点的像素值,如图 1(a);  $B(x, y, t-1)$  为步态帧在  $t-1$  时刻的图像  $x, y$  坐标点的像素值,如图 1(b);  $F$  为帧差图像,如图 1(c);  $x, y$  为 2 维图像的坐标值。

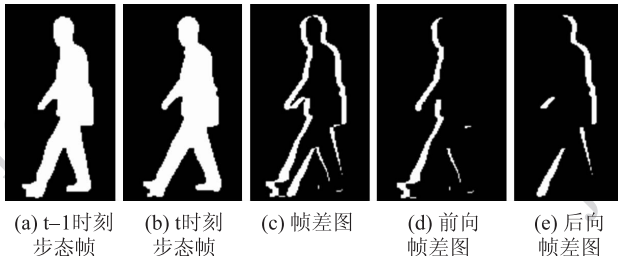


图 1 步态动态信息图像

Fig. 1 The dynamic information of gait

由如上动态信息和静态信息的融合构建动态贝叶斯网络 (DSIF-DBN) 模型如图 2 所示。

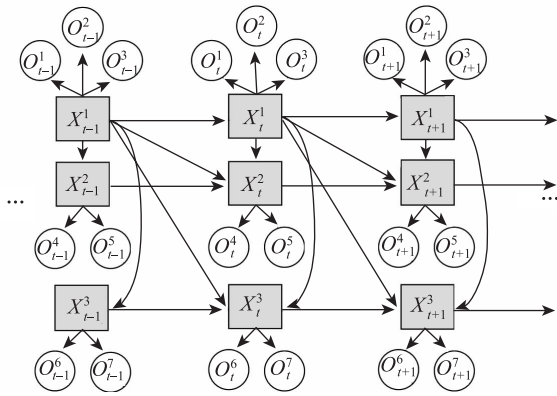


图 2 动态贝叶斯网络模型

Fig. 2 Dynamic Bayesian network model

DSIF-DBN 模型含有 3 个状态变量  $X^1, X^2, X^3$  及 7 个观测变量  $O^1 \sim O^7$ ; 其中  $X^1$  用来表达步态序列的静态帧信息, 即如图 1(a)(b) 的  $t-1$  时刻步态帧、 $t$  时刻步态帧, 此步态帧信息只包含了当前时刻的步态静态的信息, 如外形轮廓、姿态等, 不能表达步态行走时的单位时间内的运动速度、幅度等动态信息;  $X^2$  用来表达帧差图像的前向信息  $Fa$ , 如图 1(d) 的前向帧差图,  $X^3$  用来表达帧差图像的后向信息  $Fb$ , 如图 1(e) 的后向帧差图;  $X^2, X^3$  为相互独立的两个状态变量, 表达为人体的左右部分由前向后, 再由后向前的运动过程; 由于前向帧差图像和后向帧差图像表达的差别很大, 所以分开表达, 而没有用单一的帧差图像来表达。观测变量  $O^1, O^2, O^3$  为步态静态帧信息的关节角、高度、宽度信息; 观测变量  $O^4, O^5$  为步态运动的前向帧差的速度及幅度信息;  $O^6, O^7$  为步态运动的后向帧差的速度及幅度信息。

### 1.1 DSIF-DBN 模型分析

步态识别的 DSIF-DBN 模型采用 3 层的状态来描述人行走时步态的动态信息和静态信息, 而在每

层中动态概率过程假设是马氏的 (Markovian), 即未来时刻的概率只与当前时刻有关而与过去时刻无关:  $X_t$  依赖于  $X_{t-1}$ 。由于步态行走是个时序的过程, 反映的是随时刻  $t$  步态帧的姿态、幅度、轮廓以及节奏的变化; 随时刻  $t$  变化的动态信息能很好地反应出人行走时的步态幅度、节奏的变化, 而随时刻  $t$  变化的静态信息能很好地表达步态的姿态、轮廓外形的信息。在模型中动态信息受当前时刻及前一时刻静态信息的作用来表达, 模型能很好融合步态行走的动静态信息。

模型具有如下优点: 1) 能有效地融合人体行走时的动静态信息; 2) 可以捕捉人体运动时身体部位变化的细节, 即采用 2 层表达帧前差图像和 3 层表达帧后差图像; 3) 和传统的 HMM 相比, DSIF-DBN 模型大大降低参数空间。

DSIF-DBN 模型能很好融合步态行走的动静态信息, 充分表达步态的时序特性。与其他模型相比, DSIF-DBN 模型与 EHMM 模型不同, 后者采用嵌入式的方式在单马尔可夫链中内嵌多条马尔可夫链, 形式类似于分层的马尔可夫, 是从单一的观测序列中学习模型的状态和识别, 如文献 [15] 采用的方法; 而耦合隐马尔可夫模型 CHMM, 是由多条对称的马尔可夫链构成, 很难表达不同层次结构的关系, 即动态信息和静态信息的依属关系。

### 1.2 DSIF-DBN 模型推理与学习

#### 1.2.1 DSIF-DBN 模型推理

在模型的推理中, DSIF-DBN 是以概率网络为基础, 把原来的静态网络结构与时间信息结合, 形成具有处理时序数据的随机模型, 从静态贝叶斯 (BN) 出发, 导出 DBN 的概念及图形表达, 进而描述<sup>[17-22]</sup>。推理 DSIF-DBN 网络就是计算在给定观察序列  $O_{1:i} = O_1 O_2 \dots O_i$  计算隐状态变量  $X_i$  的边缘概率  $P(X_i | O_{1:i})$ , 可通过计算所有状态节点的联合概率分布, 再边缘化, 进而计算所有状态节点的概率分布, DSIF-DBN 推理的全局联合概率分布为

$$P(X_{1:T}^{1:3}, O_{1:T}^{1:7}) = P(X_{1:T}^{1:3})P(O_{1:T}^{1:7} | X_{1:T}^{1:3}) \quad (4)$$

式中, 任一状态节点的联合概率分布为

$$P(X_{1:T}^{1:3}) = \prod_{i=1}^3 P(X_i^i | Pa(X_i^i)) \times \prod_{t=2}^T \prod_{i=1}^3 P(X_t^i | Pa(X_t^i)) = P(X_1^1)P(X_1^2 | X_1^1)P(X_1^3 | X_1^1) \times$$

$$\prod_{i=2}^T [P(X_i^1 | X_{i-1}^1) P(X_i^2 | X_{i-1}^1, X_{i-1}^2, X_i^1) \times P(X_i^3 | X_{i-1}^3, X_{i-1}^1, X_i^1)] \quad (5)$$

条件概率分布为

$$P(O_{1:T}^{1:7} | X_{1:T}^{1:3}) = \prod_{i=1}^T P(O_i^1, O_i^2, O_i^3 | X_i^1) P(O_i^4, O_i^5 | X_i^2) P(O_i^6, O_i^7 | X_i^3) \quad (6)$$

数学表达中  $X_i^i$  表示第  $i$  个隐状态变量  $t$  时刻的取值, 上标 1:3 表示从第 1 到第 3 个状态变量,  $Pa(X_i^i)$  表示其父节点,

$$X_{1:T}^{1:3} = \begin{bmatrix} X_1^1 \\ X_2^1 \\ \vdots \\ X_T^1 \end{bmatrix} \cdots \begin{bmatrix} X_1^3 \\ X_2^3 \\ \vdots \\ X_T^3 \end{bmatrix} \quad (7)$$

$$O_{1:T}^{1:7} = \begin{bmatrix} O_1^1 \\ \vdots \\ O_1^7 \end{bmatrix} \cdots \begin{bmatrix} O_T^1 \\ \vdots \\ O_T^7 \end{bmatrix}$$

### 1.2.2 DSIF-DBN 模型参数表示与学习

DSIF-DBN 模型参数  $\Theta$  表示为  $\Theta = \{\pi^{1:3}, a^{1:3}, b\}$ , 其中  $\pi^{1:3}$  为初始概率,  $a^{1:3}$  为状态转移概率,  $b$  为观测概率分布。在模型中参数的更新公式及极大似然估计表示如下<sup>[23]</sup>: 在时刻  $t=1$ , 状态  $X^1$ 、 $X^2$ 、 $X^3$  的初始转换概率为

$$\hat{\pi}_i^1 = E[X_1^1 = i] = \frac{P(O_{1:T}^{1:7}, X_1^1 = i | \Theta)}{P(O_{1:T}^{1:7} | \Theta)} \quad (8)$$

$$\hat{\pi}_j^r = E[X_1^r = j | X_1^1 = i] = \frac{P(O_{1:T}^{1:7}, X_1^1 = i, X_1^r = j | \Theta)}{P(O_{1:T}^{1:7}, X_1^1 = i | \Theta)} \quad r \in \{2, 3\} \quad (9)$$

状态  $X^1$  从状态  $i$  到状态  $j$  的转换概率为

$$\hat{a}_{ij}^1 = \frac{E[X_i^1 = j | X_{i-1}^1 = i]}{E(X_{i-1}^1 = i)} = \frac{\sum_{i=2}^T \frac{P(O_{i-1:T}^{1:7}, X_{i-1}^1 = i, X_i^1 = j | \Theta)}{P(O_{i-1:T}^{1:7} | \Theta)}}{\sum_{i=2}^T \frac{P(O_{i-1:T}^{1:7}, X_{i-1}^1 = i | \Theta)}{P(O_{i-1:T}^{1:7} | \Theta)}} \quad (10)$$

状态  $X^2$  从状态  $k$  到状态  $h$  的转换概率为

$$\hat{a}_{ikh}^2 = \frac{E[X_i^2 = h | X_{i-1}^2 = k, X_{i-1}^1 = i, X_i^1 = j]}{E[X_{i-1}^2 = k, X_{i-1}^1 = i, X_i^1 = j]} = \frac{\sum_{i=2}^T \frac{P(O_{i-1:T}^{1:7}, X_{i-1}^1 = i, X_i^1 = j, X_{i-1}^2 = k, X_i^2 = h | \Theta)}{P(O_{i-1:T}^{1:7} | \Theta)}}{\sum_{i=2}^T \frac{P(O_{i-1:T}^{1:7}, X_{i-1}^1 = i, X_i^1 = j, X_{i-1}^2 = k | \Theta)}{P(O_{i-1:T}^{1:7} | \Theta)}} \quad (11)$$

状态  $X^3$  从状态  $m$  到状态  $n$  的转换概率为

$$\hat{a}_{ijmn}^3 = \frac{E[X_i^3 = n | X_{i-1}^3 = m, X_{i-1}^1 = i, X_i^1 = j]}{E[X_{i-1}^3 = m, X_{i-1}^1 = i, X_i^1 = j]} = \frac{\sum_{i=2}^T \frac{P(O_{i-1:T}^{1:7}, X_{i-1}^1 = i, X_i^1 = j, X_{i-1}^3 = m, X_i^3 = n | \Theta)}{P(O_{i-1:T}^{1:7} | \Theta)}}{\sum_{i=2}^T \frac{P(O_{i-1:T}^{1:7}, X_{i-1}^1 = i, X_i^1 = j, X_{i-1}^3 = m | \Theta)}{P(O_{i-1:T}^{1:7} | \Theta)}} \quad (12)$$

观测概率为

$$\hat{b}_{iq} = \frac{E[O_i = q | X_i = i]}{E[X_i = i]} = \frac{\sum_{i=1}^T \frac{P(O_i = q, X_i = i | \Theta)}{P(O_{1:T}^{1:7} | \Theta)}}{\sum_{i=1}^T \frac{P(X_i = i | \Theta)}{P(O_{1:T}^{1:7} | \Theta)}} \quad (13)$$

式中,  $(O_i, X_i) \in \{(O_i^1, X_i^1), (O_i^2, X_i^1), (O_i^3, X_i^1), (O_i^4, X_i^2), (O_i^5, X_i^2), (O_i^6, X_i^3), (O_i^7, X_i^3)\}$ 。

模型学习的任务是根据给定的训练数据来估计模型的参数。令  $\mathbf{X}_t = (X_t^1, X_t^2, X_t^3)$  表示  $t$  时刻的状态,  $\mathbf{X}_{1:T} = (X_1, X_2, \dots, X_T)$  表示状态序列,  $\mathbf{O}_t = (O_t^1, O_t^2, \dots, O_t^7)$  表示  $t$  时刻的观测数据,  $\mathbf{O}_{1:T} = (O_1, O_2, \dots, O_T)$  表示观测序列。模型学习的任务是根据给定的训练数据来估计模型的参数, 对于给定训练观测序列  $\mathbf{O}_{1:T}$ , 由 DSIF-DBN 模型参数  $\Theta$  最大似然方法估计, 即  $\hat{\Theta} = \operatorname{argmax}_{\Theta} P(\mathbf{O}_{1:T} | \Theta)$ ,  $\mathbf{X}_{1:T}$  观测是不完全的, 故通过 Expectation-maximization (EM) 算法来进行迭代求解<sup>[24-25]</sup>, 即

$$\Theta^{(n+1)} = \operatorname{argmax}_{\Theta} E\{\log P(\mathbf{O}_{1:T}, \mathbf{X}_{1:T} | \Theta) | \mathbf{O}_{1:T}, \Theta^{(n)}\} \quad (14)$$

式中,  $\Theta^{(n)}$  表示第  $n$  次迭代时的参数估计, 通过式 (14) 迭代收敛到一个局部极值, 至少可以达到局部最优, 联合概率密度表示如式 (4)。

### 1.3 DSIF-DBN 模型识别

从使用模型的步态识别上来看, 基于 DSIF-DBN 网络识别实际上是一个推理迭代求解过程, 如给定  $R$  个训练好的模型  $M_1 \cdots M_R$ , 其中每个模型对应一个人的步态, 通过测试, 观测序列为  $\mathbf{O}_{1:T}$ , 则由式 (15) 确定分类。

$$r = \operatorname{arg max}_i P(M_i | \mathbf{O}_{1:T}) = \frac{\operatorname{arg max}_i P(\mathbf{O}_{1:T} | M_i) P(M_i)}{P(\mathbf{O}_{1:T})} \quad (15)$$

式中,  $P(M_i)$  为模型  $M_i$  的先验概率, 取为平均值  $1/R$ , 设  $M_i$  模型参数为  $\Theta_i$ , 则  $P(O_{1:T}|M_i) = P(O_{1:T}|\Theta_i)$ , 观测序列给定, 则  $P(O_{1:T}) = 1$ , 则式(15)可推导为

$$r = \arg \max_i P(O_{1:T}|M_i)/R \quad (16)$$

由式(16)确定分类, 进而识别。

## 2 DSIF-DBN 模型特征提取

根据步态的动静态信息的特征, DSIF-DBN 模型在第 1 层中选取静态信息, 在第 2、3 层中选取的动态信息。

### 2.1 静态特征信息

在静态信息中, 步态轮廓中心用  $(x, y)$  表示。对于  $t$  时刻静态轮廓的信息为: 1) 步态轮廓信息的高宽比  $r$ ; 步态行走时是双臂、腿部、躯干交替的摆动过程, 从  $90^\circ$  视角来看, 人体轮廓宽度是个周期变化的过程, 从高度上看, 人体行走是从脚尖到脚跟的着地过程, 表现为人体身高的周期性变化, 这里对于静态帧特征采用高宽的比值作为静态特征之一。2) 腿部关节角度信息  $\theta$ , 即大腿与小腿的关节角度, 如图 3 所示。3) 人体步态轮廓中心高度值  $y$ , 如图 3 所示。

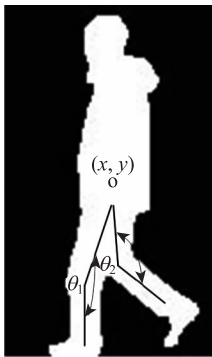


图 3 腿部关节角  
Fig. 3 Leg joint angles

### 2.2 动态特征信息

步态动态特征信息用前向及后向帧差图表达。

1) 对  $t$  时刻计算其水平运动速率  $v$ , 取值为: 对于前向帧差图像取其所有白色像素点个数作为水平运动的速度值, 前向及后向帧差图像黑色为背景, 白色为运动部位变化区域, 如图 4 所示。

2) 计算步态运动的幅度, 在  $t$  时间片内的身体各个部位动态的变化, 体现人体在行走时姿态、幅度



图 4 前向帧差图

Fig. 4 Fore-frame difference image

等特征; 计算中采用扇形区域法建立特征向量<sup>[16]</sup>。设步长为  $\Delta\theta$ , 在步态图像中心建立扇形区域, 如图 5 所示, 则步态图像被划分为  $N = \frac{360}{\Delta\theta}$ , 对每个分区计算动态区域像素点转动惯量作为此扇区的特征值, 即  $J = \sum m \times r^2$ , 其中  $m$  取为像素点的像素值, 帧差为二值图像, 白色区域应该为 255, 但为了便于计算试验中像素值设定常量值为 1。  $r$  为像素点到扇形中心的距离, 设定扇形中心点为坐标  $(0, 0)$  点, 即  $r = \sqrt{x^2 + y^2}$ , 扇形区域中动态部位的像素点坐标为  $(x, y)$ , 即图像中白色区域的坐标点, 以此特征值建立特征向量。

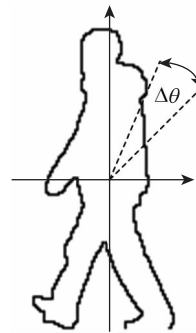


图 5 扇形区域建立特征向量

Fig. 5 Feature vectors establish based on fan-shaped area

## 3 实验及结果分析

实验采用 CASIA B 步态数据库, CASIA B 是由中国科学院提供的一个大规模、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 种行走条件下 (普通条件、穿大衣、携带包裹条件) 采集。本实验采用  $90^\circ$

视角,3种行走状态即普通、穿大衣、背包。

### 3.1 DSIF-DBN 模型下步态识别结果

**实验1** 在步态识别 DSIF-DBN 模型中,采用3种方式对模型进行训练和识别来研究环境对识别率的影响:1)普通条件下对模型进行训练识别;2)普通及穿大衣条件对模型进行训练和识别;3)普通、穿大衣和背包条件下对模型进行训练和识别。

表1中样本数为124人,其中普通条件下提供了每人6个视频序列,穿大衣条件下提供了2个视频序列,背包条件下提供了2个视频序列。在训练时取视频中的1个步态循环作为训练样本,通常一个视频中含有3个步态循环(1个步态循环为从当次左腿或右腿迈出到下次左腿或右腿迈出所经历的步态序列,称为一个步态循环),这样大大增加了训练样本数。从实验结果可以看出,方式1)的普通条件下识别率较高;方式2)的普通及穿大衣条件下识别率没有较大改变,说明在此模型下,穿大衣对识别的结果影响不大;方式3)识别结果有一定幅度降低,说明在步态识别中,普通、穿大衣和背包条件下,对模型的识别效果有较大影响。而在同样的数据库中采用帧差能量图及马尔可夫模型的方法<sup>[4]</sup>的识别率最高为91.1,而采用利用动态部位变化的步态识别方法<sup>[16]</sup>的识别率最高为94.4,但以上采用的方法,不是把3种状态的混合样本进行训练,而是分别对普通、穿大衣、背包的样本数据进行训练和识别。在文献[2]中,采用3种状态的混合样本进行训练的最高识别率为85.5。本文方法采用3种状态独立的训练识别率如下表2,实验结果说明本模型识别结果明显高于其他模型,而且受外套、背包的影响较低。

表1 不同方式下的识别率

Table 1 Recognition rate at different ways

| 训练识别方式 | 训练及识别视频数 | 识别率/% |
|--------|----------|-------|
| 方式1)   | 6 × 124  | 97.1  |
| 方式2)   | 8 × 124  | 95.6  |
| 方式3)   | 10 × 124 | 91.2  |

表2 相同方式下的识别率

Table 2 Recognition rate under same ways

| 训练识别方式 | 训练及识别视频数 | 识别率/% |
|--------|----------|-------|
| 普通     | 6 × 124  | 97.1  |
| 穿大衣    | 2 × 124  | 97.3  |
| 背包     | 2 × 124  | 96.1  |

### 3.2 模型间的步态识别结果比较

**实验2** 本文模型采用3层的动态贝叶斯网络,可以充分表达人体行走时的动态步态特征,即行走姿态、幅度等特性的变化。模型中用 $X^2$ 、 $X^3$ 分别表达人体运动的前向帧差和后向帧差图像,用以说明人体运动的身体摆动。如采用两层的动态贝叶斯网络,两层动态贝叶斯网络的第1层同3层动态贝叶斯网络一样来表达静态信息,而第2层表达帧差图像信息,实验结果如表3所示。实验结果表明3层模型识别效果更好。

表3 不同模型下的识别率

Table 3 Recognition rate under different models

| 训练识别方式 | 训练及识别视频数 | 3层 DSIF-DBN 模型识别率/% | 2层 DSIF-DBN 模型识别率/% |
|--------|----------|---------------------|---------------------|
| 普通     | 6 × 124  | 97.1                | 95.1                |
| 穿大衣    | 2 × 124  | 97.3                | 95.1                |
| 背包     | 2 × 124  | 96.1                | 94.8                |

### 3.3 模型鲁棒性分析

在数据获取中,观测数据样本会出现噪声或者数据缺失的情况,实验在噪声或者数据缺失情况下对 DSIF-DBN 模型定量分析了步态的识别结果。

1)数据中存在噪声 噪声表示如下:  $(x_t, y_t)$  为 $t$ 时刻人体的轮廓坐标,  $(x_t\%, y_t\%)$  为对应加噪声后的值,  $\xi_t, \eta_t$  为满足标准正态分布的高斯噪声,表示对人体步态轮廓的污染。

$$\begin{cases} x_t\% = x_t + w\xi_t & \xi_t: N(0,1) \\ y_t\% = y_t + w\eta_t & \eta_t: N(0,1) \end{cases} \quad (17)$$

从图6中可以看出,  $w \geq 3$  时,识别率下降迅速,但 DSIF-DBN 模型识别结果明显优于其他模型。

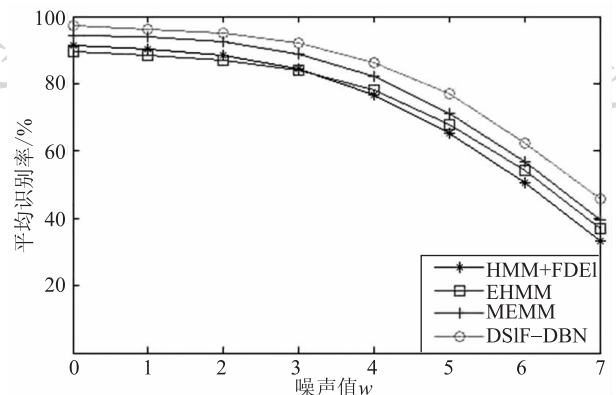


图6 识别率-污染程度曲线图

Fig. 6 Recognition rate-pollution level curve image

2)数据缺失的情况 在一个周期步态序列内,随机抽取  $n$  帧图像,抽取的方式采用随机方式,抽取比例占整个周期步态帧图像的 10%~90%。一个步态周期样本数选为 24 帧及 48 帧,识别率结果如图 7 所示。

从图 7 可以看出,在数据缺失条件下,随机抽取

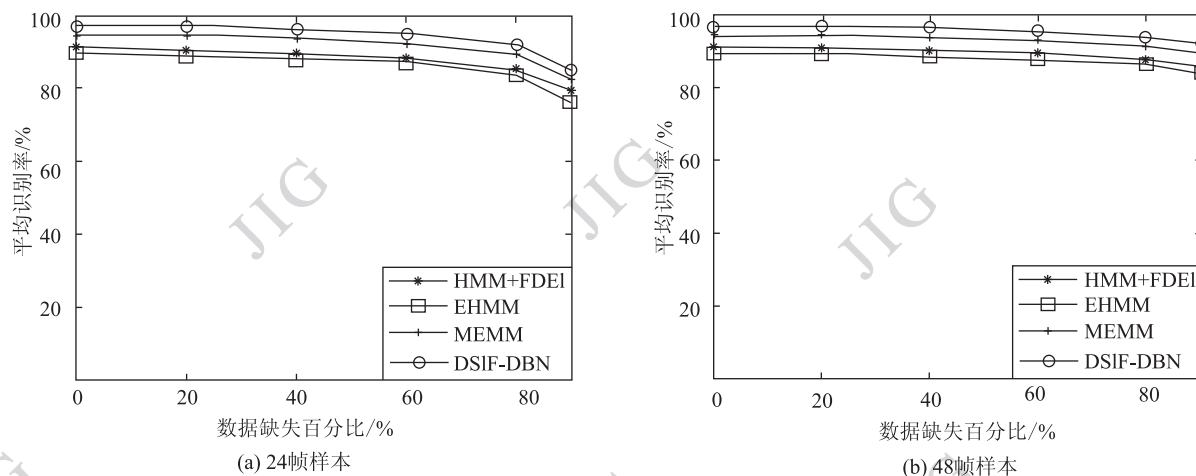


图 7 识别率-数据缺失曲线图

Fig. 7 Recognition rate-data missing curve image

## 4 结 论

提出一种基于动静态信息相结合的多信息融合的动态贝叶斯识别网络(DSIF-DBN),此模型中含有 3 层状态,每个时间片都为静态信息和动态信息的融合,而每层状态分别反映了步态的静态信息或动态信息,不同层状态之间反映了动静态信息之间的关系。此模型能很好地表达步态的时序特性,即步态行走时人体姿态,运动幅度等特性的节奏的变化,能有机地融合步态的静态信息及动态信息进行建模及识别。实验在 CASIA B 数据库进行,实验表明,DSIF-DBN 模型识别率高,并且可大大降低背包对步态识别的影响,模型在有噪声污染及数据缺失情况下,识别效果依然良好。

## 参考文献(References)

[1] Kim D, Paik J. Gait recognition using active shape model and motion prediction [J]. Computer Vision IET, 2010, 4(1): 25-36.  
 [2] Wang K J, Liu L L, Ben X Y. Gait recognition based on gait energy image and two dimensional principal component analysis

图像,单步态循环的样本帧数越多,受数据缺失的影响越小。同时在步态识别过程中,由于存在着大量的冗余帧信息,所以几乎所有模型都对数据缺失具有较好的鲁棒性,DSIF-DBN 模型的方法,可对帧信息动态与静态相结合,实现多信息融合,在数据缺失条件下鲁棒性更好。

[J]. Journal of Image and Graphics, 2009, 14(12): 2503-2509. [王科俊,刘丽丽, 贲晔焯. 基于步态能量图像和 2 维主成分分析的步态识别方法 [J]. 中国图象图形学报, 2009, 14(12): 2503-2509.]  
 [3] Zhang E H, Zhao Y W, Xiong W. Active energy image plus 2DLPP for gait recognition [J]. Signal Processing, 2010, 90(7): 2295-2302.  
 [4] Chen C H. Frame difference energy image for gait recognition with incomplete silhouettes [J]. Pattern Recognition Letters, 2009, 30(11): 977 - 984.  
 [5] Chen C H, Liang J M, Zhu X C. Gait recognition based on improved dynamic Bayesian networks [J]. Pattern Recognition, 2011, 44(4): 988 - 995.  
 [6] Tafazzoli F, Safabakhsh R. Model-based human gait recognition using leg and arm movements [J]. Engineering Applications of Artificial Intelligence, 2010, 23(8): 1237-1246.  
 [7] Zhao Y W, Zhang E H, Lu J W. Gait recognition via multiple features and views information fusion [J]. Journal of Image and Graphics, 2009, 14(3): 387-393. [赵永伟, 张二虎, 鲁继文. 多特征和多视角信息融合的步态识别 [J]. 中国图象图形学报, 2009, 14(3): 387-393.]  
 [8] Bashir K, Xiang T, Gong S G. Gait recognition without subject cooperation [J]. Pattern Recognition Letters, 2010, 31(13): 2052-2060.  
 [9] Xue Z J, Ming D, Song W, et al. Frame difference energy image for gait recognition with incomplete silhouettes [J]. Pattern Recognition, 2010, 43(8): 2904-2910.

- [10] Jeff P, Mark S, Adam P. Automatic gait recognition using area-based metrics [J]. *Pattern Recognition Letters*, 2003, 24(14): 2489-2497.
- [11] Suk H, Sin B K, Lee S W. Hand gesture recognition based on dynamic Bayesian network framework [J]. *Pattern Recognition*, 43(9): 3059-3072.
- [12] Bae J, Tomizuka M. Gait phase analysis based on a Hidden Markov Model [J]. *Mechatronics*, 2011, 21(6): 961-970.
- [13] Cheng M H, Ho M F, Huang C L. Gait analysis for human identification through manifold learning and HMM [J]. *Pattern Recognition*, 2008, 41(8): 2541-2553.
- [14] Waleed H, Kasabov N. Reduced feature-set based parallel CHMM speech recognition systems [J]. *Information Sciences*, 2003, 156(1-2): 21-38.
- [15] Zhang Q J, Xu S L. Gait recognition based on embedded hidden markov model [J]. *Information and Control*, 2010, 39(1): 25-29. [张前进, 徐素莉. 基于嵌入式隐马尔可夫模型的步态识别[J]. *信息与控制*, 2010, 39(1): 25-29.]
- [16] Zhang E H, Zhao Y W. Gait recognition using variance of dynamic region [J]. *Journal of Image and Graphics*, 2009, 14(9): 1756-1763. [张二虎, 赵永伟. 利用动态部位变化的步态识别[J]. *中国图象图形学报*, 2009, 14(9): 1756-1763.]
- [17] Yang G S, Lin Y Z, Bhattacharya P. A driver fatigue recognition model based on information fusion and dynamic Bayesian network [J]. *Information Sciences*, 2010, 180(10): 1942-1954.
- [18] Laurence L S, Sigelle M. Recognition of degraded characters using dynamic Bayesian networks [J]. *Pattern Recognition*, 2008, 41(10): 3092-3103.
- [19] Frankel J, Wester M, King S. Articulatory feature recognition using dynamic Bayesian networks [J]. *Computer Speech & Language*, 2007, 21(4): 620-640.
- [20] Jawad H, Pauplin O, Ren J C, et al. Performance of hidden Markov model and dynamic Bayesian network classifiers on handwritten Arabic word recognition [J]. *Knowledge-Based Systems*, 2011, 24(5): 680-688.
- [21] Hwang K A, Yang C H. Learner attending auto-monitor in distance learning using image recognition and Bayesian Networks [J]. *Expert Systems with Applications*, 2009, 36(9): 11461-11469.
- [22] Du Y T, Chen F, Xu W L. Approach to human activity multi-scale analysis and recognition based on multi-layer dynamic Bayesian network [J]. *Acta Automatica Sinica*, 2009, 35(3): 225-232. [杜友田, 陈峰, 徐文立. 基于多层动态贝叶斯网络的人的行为多尺度分析及识别方法[J]. *自动化学报*, 2009, 35(3): 225-232.]
- [23] Murphy K. *Dynamic Bayesian network: representation, inference and Learning* [D]. San Francisco, USA: Berkeley University of California, 2002.
- [24] North B, Blake A. Using expectation-maximisation to learn dynamical models from visual data [J]. *Image and Vision Computing*, 1999, 17(8): 611-616.
- [25] Steinborn A, Taut S, Brendler V, et al. TRLFS: Analysing spectra with an expectation-maximization (EM) algorithm [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2008, 71(4): 1425-1432.