

一种新的基于区域特征的快速步态识别方法

柴艳妹 赵荣椿

(西北工业大学计算机学院, 西安 710072)

摘要 步态是生物特征识别领域的一个新兴热点, 它有以下 3 大优势: 远距离识别、非侵犯性和难于隐藏。由于当前提出的大量步态特征提取算法要么过于复杂, 要么识别率不高, 难以满足自动步态识别的需要, 因此, 为了进行准确快速的步态识别, 提出了一种新的基于区域特征的快速步态识别方法。该方法首先将检测出的 2 维人体侧影分为头部、躯干和腿部 3 个区域, 然后分别提取每个区域的目标面积, 最后将这些面积特征和人体的宽高比特征一起构成步态特征矢量用于训练和识别。此外, 还改进了一种新的 N-best 分类器, 该分类器在一定程度上提高了算法的识别率。实验结果表明, 该新方法不仅简单快速, 而且在 UCSD 和 CMU 数据集上分别得到了 90% 和 98% 左右的高识别率。

关键词 步态识别 区域面积 N-best 分类器

中图法分类号: TP391.41 文献标识码: D3 文章编号: 1006-8961(2006)09-1260-06

A New Fast Gait Recognition Method Based on Region Feature

CHAI Yan-mei, ZHAO Rong-chun

(School of Computer, Northwestern Polytechnical University, Xi'an 710072)

Abstract Recognizing people by their gait is a recent research hotspot. Compared with other biometrics, gait has the following three advantages: distance recognition, uninvasive and difficult to conceal. A lot of research on the gait feature extraction has been done, but some of them are much expensive computation, while others are poor identification effect. A new fast gait recognition method based on the region object area is proposed in this paper. Firstly, we divide the two dimensional silhouette of the walker into three regions (head region, trunk region and legs region). Then their region object areas are computed respectively. Together with the ratio of the silhouette's height to width, the gait feature vectors are constructed to identify different subjects. Moreover, a kind of N-best classifier is improved in this paper. Experimental results show that the new method is not only simple and fast, but also very effective. It can also be used on different database. Recognition rate of approximately 90% on the UCSD database and over 98% on the CMU database are achieved.

Keywords gait recognition, region area, N-best classifier

1 引言

人们早在 20 世纪 60 年代就开始对步态识别进行了研究, 但当时多用于临床医学方面。由于序列图像的数据量大、计算复杂性高, 致使这一技术的发展相当缓慢。但是, 随着计算机硬件处理能力的飞速提高和自“9.11 事件”以来人们对监视和认定技术的需求, 步态识别在近几年得到了广泛关注和快速发展。和其他生物特征识别方法相比, 由于步态

识别具有非侵犯性、远距离识别性、清晰度要求低和难以隐藏等优点, 因此从安全监控的角度来看, 步态是在远距离情况下最具潜力的生物特征。

人们提出了大量的方法用于步态特征的提取, 例如速度矩^[1]、隐马尔科夫模型^[2]、Procrustes 形状分析^[3]、主成分分析 (principal component analysis, PCA)^[4,5]、正则分析 (canonical analysis, CA)^[5]、对称性分析^[6]以及各种基于运动学和动力学的模型^[7,8]等, 但这些方法要么过于复杂, 要么识别率不高, 难以满足自动步态识别的需求。

基金项目: 国家自然科学基金项目(60472072)

收稿日期: 2005-04-28 改回日期: 2005-09-12

第一作者简介: 柴艳妹(1978~), 女, 2004 年获西北工业大学硕士学位, 现为西北工业大学计算机学院博士生。主要研究方向为图像处理、数据融合、步态识别、数字视频等, 发表论文 6 篇。E-mail: chai-4@163.com

针对上述不足,本文提出了一种新的基于区域面积的快速步态识别方法。由于该方法是在检测出的二值人体侧影上进行,因此简单快速、识别效果颇佳。另外,本文还改进了一种 N-best 分类器,用于身份识别,以进一步提高算法识别率。

2 侧影提取及其模板化

2.1 侧影提取

由于实验中摄像机是静止的,即人从摄像机前走过,并且不被物体遮挡(见图 1(a)),因此侧影提取实质上就是从序列图像中消除背景,以便检测到运动目标。它的实现步骤如下:

(1) 背景估计 由于场景在整个视频序列中近似静止,即其背景对应低频信息,因此可以用序列图

像中像素的平均值来估计静止背景(见图 1(b))。令 I 代表一个包含 N 帧图像的序列, $I_t(x, y)$ 表示图像序 I 中第 t 帧图像在 (x, y) 位置上的像素值,则背景图像像素值 $B(x, y)$ 可表示为

$$B(x, y) = \frac{1}{N} \sum_{t=0}^{N-1} I_t(x, y) \quad (1)$$

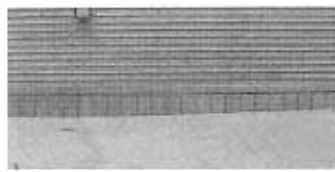
(2) 目标检测与分割 首先用背景差分法来检测序列图像中运动的目标,并在设定的阈值下对图像进行二值分割。令 T 为阈值,则序列图像与背景图像差分如下:

$$D_t(x, y) = \begin{cases} 1 & |I_t(x, y) - B(x, y)| \geq T \\ 0 & \text{else} \end{cases} \quad (2)$$

(3) 后处理 为了改善分割结果,分别使用了 3×3 的腐蚀和膨胀滤波算子来去除噪声点,并且用连通性分析算法进一步填充小的孔洞(见图 1(c))。



(a) 原始图像



(b) 背景图像



(c) 二值图像

图 1 人的侧影提取

Fig. 1 Silhouette extraction

2.2 模板化

从图 1 可以看到,目标只占了整幅图像的一小部分。为了去除冗余、提高计算效率,本文将目标侧影放置到一个图像模板中,这个过程称之为模板化。模板化有以下两种方法:一种是将目标侧影按比例

缩放到一个固定大小的模板中;另一种则是不做任何缩放,只将原始目标剪裁到大小合适的模板中去。图 2 分别表示了对 UCSD(University of California, San Diego)数据库进行两种模板化处理的结果,而接下来的实验过程也将分别在这两种模板上进行。



(a) 原始图像(320 × 160)



(b) 缩放的模板(64 × 64)



(c) 无缩放的模板(104 × 104)

图 2 两种模板化方式

Fig. 2 Two types of image templates

3 提取步态识别特征

自动步态识别的关键是如何从一段序列图像中

提取出具有较强区分能力的步态特征,而提取步态特征的现有方法可分为以下两大类:一类是基于运动和动力学模型的方法,例如将大腿建模为链接的钟摆^[7]、跟踪人体主要关节的角度变化^[8]等;另一

类则是基于整体的方法,其大多是建立在对侧影图像进行整体统计分析的基础上^[1,3,6,9]。由于对人体建模及跟踪相对比较困难,因此目前的自动步态识别算法中后者占多数。南安普敦大学的 James 等人通过提取出序列中侧影轮廓的对称性算子,并通过将其加权平均来得到步态识别特征,该识别算法虽在 UCSD 等数据集上得到了 90% 以上的高识别率^[6],但提取侧影轮廓和其对称性算子的方法繁琐复杂,且计算代价过大。南佛罗里达大学的 Liu Zongyi 等则提出了一种将二值侧影序列直接进行相加和平均来得到步态识别特征的方法^[9],该方法计算代价虽然小,但识别率却不理想。受到上述文献的启迪,本文针对其不足,提出了一种新的自动步态识别方法。该方法也是基于二值侧影图像,但不是对整体进行分析,而是把侧影图像分为头部区域、躯干区域和腿部区域 3 大块分别提取其面积特征,再和人体宽高比信息共同构成步态识别特征。这样既避免了建立模型和提取对称性算子方法的繁琐,也避免了整体分析法的笼统。

由观察可知,走路是个动态过程,且具有以下特征:①大多数人走路时,头部会有轻微运动;②除了手臂的摆动外(侧影序列中,手臂运动因受到自遮挡而影响很小),躯干基本保持不变;③腿部的运动呈现出一定的周期性。另外,对于不同的人,这几部分运动的幅度也不一样。由此就可将人体侧影按身体比例分为头部、躯干和腿部 3 个区域,并可分别求其目标面积作为区域特征。这里目标面积在二值图像中即目标像素的个数,它只和目标的大小有关。在行走过程中,由于身体各部分在某个固定区域内的面积会有所变化,且变化越大,说明该身体部分的运动幅度越大,因此可将其作为步态特征用于身份识别。本文所用的人体侧影比例分配可由大量实验和观察得出,其中头部占 20%,躯干和腿部各占 40%。求解区域目标面积的公式如下^[10]:

$$A_k = \sum_{(x,y) \in R_k} P(x,y) \quad (3)$$

其中 $P(x,y) = \begin{cases} 0 & (x,y) \in R_{\text{back}} \\ 1 & (x,y) \in R_{\text{target}} \end{cases}$, k 为子块号,分别代表头部、躯干和腿部, W_k 表示第 k 个子块区域, R_{back} 代表背景区域, R_{target} 代表目标区域。

人体的另外两个显著特征是人的高度和宽度,它们直接代表了人的高低和胖瘦,这也是人眼识别的重要依据之一。由于这两个特征会随着目标和摄

像机之间距离的远近而发生变化,因此它们单独不足以作为步态识别的特征,但是它们的比值却是相对稳定的,可以作为一个识别依据,其称之为目标的宽高比特征。人体侧影 3 个区域的面积和其宽高比特征共同构成了被测试者的动态步态特征矢量,可用于人体步态的训练和识别过程。

4 训练和识别

4.1 训练

给定 C 个训练序列,序列中的每一帧可由一个 4 维特征矢量 (A_h, A_r, A_l, r) 表示。其中 A_h, A_l, A_r 分别代表头部(head)、躯干(trunk)和腿部(legs)区域的面积,而 r 则代表该帧图像中目标的宽高比。 $X_{i,j} = (A_{i,j}^h, A_{i,j}^l, A_{i,j}^r, r_{i,j})$ 是第 i 个序列中的第 j 个特征矢量,且 N_i 是第 i 个序列的帧数,其中 $1 \leq i \leq C, 1 \leq j \leq N_i$ 。

4.2 识别

令训练序列为 $I_g = \{X_1^g, X_2^g, \dots, X_{N_g}^g\}$, 测试序列为 $I_p = \{X_1^p, X_2^p, \dots, X_{N_p}^p\}$, N_g 和 N_p 分别表示两序列的帧数(下角 g 代表训练,下角 p 代表测试)。由于步态具有周期性,因此本文的相似度比较是建立在步态周期的基础上的。假设一个步态循环约为 N 帧,则测试序列可被分成 $\lfloor N_p/N \rfloor$ 个子序列,其第 k 个子序列可表示为 $I_k^p = \{X_{k+1}^p, X_{k+2}^p, \dots, X_{k+N}^p\}$ 。这个测试子序列和训练序列之间的距离可表示为

$$dis(I_k^p, I_g) = \sum_{j=1}^N \|X_{k+j}^p - X_{l+j}^g\| \quad (4)$$

l 是训练序列中和测试子序列进行比较的起始帧,而整个测试和训练序列之间的距离为

$$Dis(I_p, I_g) = \frac{1}{K} \sum_{k=1}^K \min_l (dis(I_k^p, I_g)(l)) \quad (5)$$

大家知道,自动步态识别算法中大多选用简单、成熟的模式分类技术作为分类器,例如最近邻分类器(the nearest neighbor classifier, NN)和 K 近邻分类器(K -nearest neighbor classifier, KNN)。文献[11]中提出了一种 N -best 分类器,该分类器虽提高了算法的识别率,但由于该分类器是采用固定赋值系数的方法,从而使得算法缺乏灵活性。本文对其进行了改进,提出了一种自适应的 N -best 分类器,新分类器在实验中取得了良好的效果。

4.3 自适应的 N -best 分类器

在 NN 分类器中,每个序列被分类到离它最近

的邻居所属的类中,而KNN则是将前 k 个最近距离中出现次数最多的类作为所属类别。对于N-best分类器,则首先选取前 N 个最近距离;然后再按照接近的程度为其自适应地分配一个权重,距离越小,权重越大,最后将属于同一类别的权重进行相加,来得到其权重和。其中权重和最大的即为所属类别。具体的自适应权重分配方案如下:

(1) 将测试样本到所有训练样本的距离从小到大排序,然后选取前 N 个值;

(2) 对于第 i ($i=0 \dots N$)个值,其类别的权重 w 可定义为 $w = N - i / \sum_{j=1}^N j$;

(3) 对于前 N 个距离中出现 K 次的类别,其权重和 $W = \sum_{k=1}^K w_k$;

(4) 选取权重和最大的类为待识别序列所属的类。

5 实验

5.1 测试数据集

本文使用以下两个数据集来进行实验。它们分别代表了数据集的大/小、室内/室外、快走/慢走等因素对自动步态识别的影响。

(1) University of California, San Diego (UCSD) 数据集。该数据集在室外拍摄,由6人,每人7个序列组成,每个序列含2~3个步态循环。原始图像大小为 320×160 ,分别被模板化为 64×64 和 104×104 两种大小进行实验。

(2) Carnegie Mellon University (CMU MoBo) 数据集。该数据集在室内拍摄,由25个人组成,分快走、慢走、倾斜和抱球4种方式。本文分别实验了快

走和慢走两种方式的侧视序列,每个序列含7~8个步态循环。图像原始大小为 640×480 ,被模板化为 64×64 大小进行实验。

5.2 实验结果

本文用以下两个方法来评价识别结果的性能:一种是使用留一校验(leave-one-out cross validation)方法来获取识别率的无偏估计,即对于UCSD数据集的42个序列样本,每次留出一个样本作为测试样本,而对其余样本进行训练,然后将测试样本依据与训练样本集的相似性来进行分类,这样的过程共进行42次,而对于CMU的快走和慢走数据集,则将拆分后的每个步态循环作为一个序列用留一校验法来进行评价,这样的过程分别进行200次和175次(每个快走序列包含8个步态循环,每个慢走序列包含7个步态循环)。另外,为了了解速度对步态识别的影响,本文分别以快走和慢走数据集互为训练和识别样本进行了实验,并且上述所有实验均采用了NN、KNN和自适应N-best 3种分类器进行了分类。表1总结和对比了上述各实验的正确分类率(correct classification rate, CCR)。为了表述方便,本文为每个实验分别起了别名A~F。

另一种用来评价分类性能的度量方法是在FERET脸像识别算法中提出的ROS(rank order statistic)方法。它是累计匹配分值(cumulative match score, CMS)来报告性能统计特性(如图3所示)。图3横坐标代表阶次,纵坐标代表正确匹配百分比 $\rho(i)$ 。其中 $\rho(1)$ ($rank=1$)代表实际正确的分类率^[12]。由于在FERET识别算法中训练集的大小等于类别数,即训练集中每个类别只有一个样本^[12],因此对于同一数据集可以随机抽取每人的一个序列作为训练集,其余序列作为测试集,然后即可画出它们所对应的CMS图。其中,UCSD数据集上

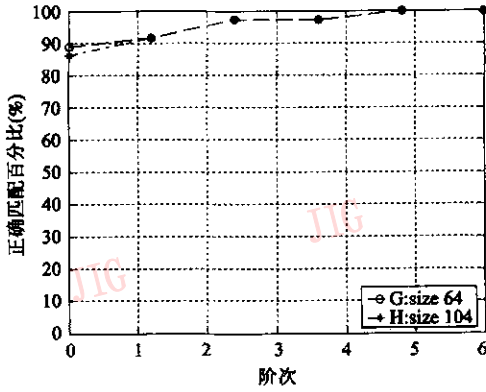
表1 算法的正确分类率

Tab.1 CCR of classifiers under our method

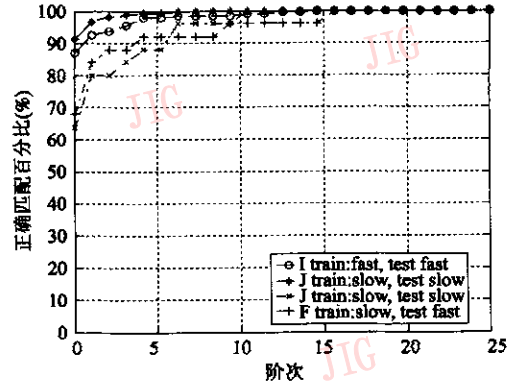
数据集	实验方法	正确分类率(%)		
		NN	KNN($k=3$)	自适应N-best
UCSD 数据集(6人)	A 64×64 留一校验	92.857	88.095	92.857
	B 104×104 留一校验	95.238	92.857	95.238
	C Fast Walk 留一校验	99	98	99
CMU 数据集(25人) 64×64	D Slow Walk 留一校验	98.857	97.142	98.857
	E Fast 训练, Slow 识别	64	64	64
	F Slow 训练, Fast 识别	68	68	68

的实验用 G 和 H 代表 ,CMU 数据集上的实验用 I 和 J 代表。对不同速度的测试实验仍沿用快走和慢走序列互为训练和识别样本集分别进行实验(实验 E

和实验 F)。它们的 CMS 如图 3 所示 ,图 3(a)是与 UCSD 数据集上的实验对应的 CMS ,图 3(b)是与 CMU 数据集上的实验对应的 CMS。



(a) UCSD 数据集上实验 G 和实验 H 的 CMS



(b) CMU 数据集上实验 I , J , E 和实验 F 的 CMS

图 3 基于 FERET 协议的识别结果

Fig. 3 Identification performance based on the FERET

5.3 性能分析

从上述的两种评价方式均可以看出 ,本文的方法在两个数据集上都得到了令人鼓舞的识别率 ,而且从整体上看 ,步行速度和拍摄距离(即图像质量)对算法识别率的影响很大。从 3 种分类器的性能比较来看 ,自适应 N-best 分类器要略胜一筹。实验 B 的结果明显好于实验 A ,这是因为实验 A 中使用了缩放模板 ,其在减少计算量的同时 ,也使原始图像中的部分信息丢失 ,从而导致了识别率的略微下降。实验 C 和实验 D 的识别率基本一致 ,都达到了 98% 左右。CMU 虽然比 UCSD 数据规模大 ,但是由于它的拍摄距离较近、图像质量高 ,因此取得了更好的识别效果。实验 E 和实验 F 由于步速的改变影响到了身体的动态变化和步幅、步频的变化 ,从而导致了识别率的较大幅度下降。但和文献 [2] 中的同一数据集上的相同实验(60% 以下)相比 ,本文方法所获得的 64% 以上的识别率是相当高的。图 3 中基于 FERET 协议的识别结果也从另外一个角度验证了上述分析。

6 结 论

本文最大的特点是在系统分析传统的模型法和整体分析法的基础上 ,把侧影图像分为头部、躯干和腿部 3 大区域来分别提取特征 ,并和人体宽高比这一重要特征一起来构成用于训练和分类的步态特征

矢量。这样做不仅避免了建立动力学模型的繁琐复杂 ,也避免了整体分析法的笼统。实验结果证明 ,本文方法不仅简单、易操作 ,而且取得了相当好的识别结果。但是由于它对步行速度和图像质量比较敏感 ,因此应进一步考察数据库规模对该算法性能的影响 ,以及如何改进和提取更有效的步态特征 ,以消除速度和图像质量对步态识别的影响 ,将是后续的研究重点。

参考文献(References)

- Shutler J , Nixon M. Zernike velocity moments for description and recognition of moving shapes [A]. In : Proceedings of British Machine Vision Conference(BMVC) [C] , Manchester , UK , 2001 : 705 ~ 714.
- Kale A , Rajagopalan A , Cuntoor N , et al. Gait based recognition of humans using continuous HMMs [A]. In : Proceedings of the International Conference on Face and Gesture Recognition [C] , Washinton D C , USA , 2002 321 ~ 326.
- Wang Liang , Ning Hua-zhong , Hu Wei-ming. Gait recognition based on procrustes shape analysis [A]. In : Proceedings of International Conference on Image Processing(ICIP) [C] , Rochester , New York , USA , 2002 : 22 ~ 25.
- Wang Liang , Hu Wei-ming , Tan Tie-niu. Gait-based human identification [J]. Chinese Journal of Computers , 2003 , 26(3) : 353 ~ 360. [王亮 胡卫明 谭铁牛. 基于步态的身份识别 [J]. 计算机学报 , 2003 26(3) 353 ~ 360.]
- Huang P , Harris C , Nixon M. Recognising humans by gait via parametric canonical space [J]. Journal of Artificial Intelligence in Engineering , 1999 , 13(4) : 359 ~ 366.

- 6 James B , Nixon M , Cater J. Automatic gait recognition by symmetry analysis [J]. Pattern Recognition Letters , 2003 , 24 (13) : 2175 ~ 2183.
- 7 Cunado D , Nixon M , Carter J. Using gait as a biometric via phase-weighted magnitude spectra [A]. In : Proceedings of International Conference on Audio-and Video-based Biometric Person Authentication [C] , Crans Montana , Switzerland , 1997 : 95 ~ 102.
- 8 Ning Hua-zhong , Tan Tie-niu , Wang Liang , *et al.* Kinematics-based tracking of human walking in monocular video sequences [J]. Image and Vision Computing , 2004 , 22 (5) : 429 ~ 441.
- 9 Liu Zong-yi , Sarkar S. Simplest representation yet for gait recognition : averaged silhouette [A]. In : Proceedings of the 17th International Conference on Pattern Recognition [C] , Cambridge , UK , 2004 : 211 ~ 214.
- 10 Zhao Rong-chun , *et al.* Digital image processing [M]. Xi 'an : The Publishing Company of Northwestern Poly-technical University. [赵荣椿等. 数字图像处理导论 [M]. 西安 :西北工业大学出版社 , 1995.]
- 11 Zhang Rong , Vogler C , Metaxas D. Human gait recognition [J]. Computer Vision and Pattern Recognition , 2004 , 27 (02) : 18 ~ 20.
- 12 Phillips J , Moon H , Rizvi S , *et al.* The FERET evaluation methodology for face recognition algorithms [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence , 2000 , 22 (10) : 1090 ~ 1104.