

一种基于光流的多区域分割在步态识别中的应用

徐艳群^{1,2} 张 斌¹

(南阳理工学院计算机科学与技术系 南阳 473004)¹ (西安交通大学经济与金融学院 西安 710000)²

摘 要 人体目标分割的质量对步态识别的性能有直接的影响。提出了一种鲁棒性的步态表示方法,即利用光流特征提取视频中的运动信息,并将目标人体区域部分按人体结构特点划分为多个子区域,每个子区域通过基于光流特征的椭圆模型进行拟合,建立多区域椭圆模型的人体结构模型。识别过程中将模型参数作为步态特征,结合动态时间规整技术解决了动态模式的相似度量及匹配问题。实验表明,该算法可以有效地提高识别算法的鲁棒性,并且具有较好的识别性能。

关键词 步态识别,人体轮廓,光流,椭圆模型,动态时间规整

中图法分类号 TP391.41 **文献标识码** A

Application of Segmentation Based on Optical Flow in Gait Recognition

XU Yan-qun^{1,2} ZHANG Bin¹

(Department of Computer Science and Technology, Nanyang Institute of Technology, Nanyang 473004, China)¹

(School of Finance and Economics, Xi'an Jiaotong University, Xi'an 710000, China)²

Abstract The quality of human silhouettes has a direct effect on gait recognition performance. This paper proposed a robust gait representation scheme to suppress the influence of silhouette incompleteness. By means of dividing human body area in a video sequence into several sub-areas and representing each sub-area by an ellipse whose parameters can be calculated from the corresponding motion information extracted from optical flow field, a new body structure model called multi-linked ellipse model was established. In the recognition stage, the parameters of model are finally used to achieve gait recognition based on dynamic time warping technology. Experimental results prove the higher performance of the method.

Keywords Gait recognition, Silhouettes, Optical flow, Ellipse model, DTW(dynamic time warping)

步态识别是近年来计算机视觉和生物特征识别领域的一个备受关注的研究方向,它旨在根据人们走路的姿势进行身份识别,通常包括视频捕捉、侧影分割、特征提取和识别4个阶段。步态识别针对的仅仅是视频中的前景目标(行人),因此为了消除背景、衣服颜色或纹理等因素的影响,需要从序列图像中将变化区域从背景图像中提取出来。然而许多因素都会导致分割的不准确,比如光照的改变,目标和背景颜色相似,行人与摄像机的距离,遮挡等,最终得到不完整的人体侧影。为了克服分割质量对识别性能的影响,研究人员提出了诸多的策略,可分为如下3类:

第一类策略专注于对侧影的重建。Liu等^[1]提出了一种基于隐马尔可夫(HMM)模型对侧影进行重建,该算法能够恢复由背景和阴影造成的不完整侧影,并且对视角改变具有较好的鲁棒性。但是目标所特有的特征可能在重建中失去,因此导致识别性能不高。

第二类策略专注于对轮廓的校正。Yu^[2]等利用改进动态时间规整(IDTW)降低在人体轮廓上噪音点的影响,该算法利用空间域上的DTW和轮廓相似度量来增强基于轮廓的

步态识别算法对噪音的鲁棒性。该算法能够有效地提高识别性能,但是时间复杂度较高。

第三类策略专注于提取鲁棒性的特征向量,或者使用鲁棒性的识别算法。Han和Bhanu提出了步态能量图^[3](Gait Energy Image),这种方法将步态周期的平均图像作为行人的步态特征。步态能量图具有良好的识别率,对于噪音点不敏感,但是每个周期仅仅只获得一个能量图,在实际中往往导致训练样本不足。

Little等人^[4]讨论了行为识别中的一个重要的问题:该提取什么样的特征。序列图像中可选取的特征是多种多样的,如目标形状特征:大小、颜色、轮廓等;动力学特征:速度、方向、加速度等。由于动力学特征在行为识别中比目标形状特征具有更好的鲁棒性, Little等选择提取序列图像中的光流特征。光流不需要预先获取图像背景,而且计算结果仅仅依靠连续帧的相对运动,不受复杂环境的影响,具有较强的鲁棒性。Lee等^[5]讨论了实际应用中随时都可能出现的遮挡问题,认为对步态识别研究而言,成功解决遮挡问题是至关重要的。而基于模型的步态分析就有这方面的优势,原因在于模

到稿日期:2011-06-26 返修日期:2011-09-07 本文受国家大学生创新实验计划(081063510),2009年度河南省国际合作项目(094300510 027),2011年河南省软科学计划项目(112400450137)资助。

徐艳群(1978-),女,博士,讲师,主要研究方向为软件工程及计算机应用,E-mail: xuyanqun@yahoo.com.cn;张 斌(1979-),男,博士,讲师,主要研究方向为软件工程及计算机应用。

型是依赖于序列图像中人的移动模式而建立的,能够反映当前的变化,还能对过去和将来的变化进行估算。文中构建了椭圆模型:首先对人体侧面投影所构成轮廓图像进行二值化处理,依据质心比例关系将人体分成7个部分,用椭圆对每一部分建模,以椭圆的质心、离心率等参数作为反映人体步态的特征参数。

本文提出了一种基于光流的多区域椭圆模型,利用光流特征提取视频中的运动信息,并将目标人体区域部分按人体结构特点划分为多个子区域,每个子区域通过结合光流特征的椭圆模型进行拟合,建立多区域椭圆模型的人体结构模型。识别过程中将模型参数作为步态特征,结合动态时间规整技术解决了动态模式的相似度和匹配问题。

1 改进的光流场基本方程

所谓光流是指亮度模式引起的表现运动,是景物中可见点的三维速度矢量在成像平面上的投影,它表示了景物表面上的点在图像中位置的瞬时变化;同时光流场携带了有关运动和结构的丰富信息^[6]。

假设在时间 t , 图像中坐标为 (x, y) 的像元, 其灰度为 $f(x, y, t)$, 经过时间 Δt , 这一点运动到 $(x+\Delta x, y+\Delta y, t+\Delta t)$, 其灰度为 $f(x+\Delta x, y+\Delta y, t+\Delta t)$ 。根据灰度守恒假设有:

$$f(x, y, t) = f(x+\Delta x, y+\Delta y, t+\Delta t) \quad (1)$$

如果认为灰度随 (x, y, t) 的变化是平滑、连续的, 则式(1)的右边可以用泰勒级数展开得:

$$f(x+\Delta x, y+\Delta y, t+\Delta t) = f(x, y, t) + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial t}dt + e \quad (2)$$

式中, e 包含 dx, dy, dt 的二次和高次项。根据式(1), 消去 $f(x, y, t)$, 且当 $dt \rightarrow 0$ 时, 得到:

$$\frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial f}{\partial t} = 0 \quad (3)$$

$$\text{令 } \frac{\partial f}{\partial x} = E_x, \frac{\partial f}{\partial y} = E_y, \frac{\partial f}{\partial t} = E_t, \frac{dx}{dt} = u, \frac{dy}{dt} = v, \text{ 得到光流场}$$

的基本方程:

$$E_x u + E_y v + E_t = 0 \quad (4)$$

但是在实际应用情况中, 光流场基本方程的灰度守恒假设往往不能满足, 如光源的运动和不均匀性, 对象的遮挡性和透明等原因, 这样计算得到的光流场精度较低。引入金字塔式多分辨率结构由粗糙到精确来计算光流场可以解决这个问题。

首先将前一帧和当前帧的图像作金字塔分解。Gaussian 金字塔是一组图像序列, 序列中的每一级:

$$G_i(x, y) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) G_{i-1}(2x+m, 2y+n) \quad (5)$$

$$\text{式中, } w(m, n) = (1/400) \begin{bmatrix} 1 & 5 & 8 & 5 & 1 \\ 5 & 25 & 40 & 25 & 5 \\ 8 & 40 & 64 & 40 & 85 \\ 5 & 25 & 40 & 25 & 5 \\ 1 & 5 & 8 & 5 & 1 \end{bmatrix}.$$

图像均是其前级图像低通滤波所得到的。 G_0 表示输入原图像, 作为 Gaussian 金字塔的底层; G_i 表示第 i 层图像, 其每一像素值均可用高斯窗口函数 w 对第 $i-1$ 级矩阵进行加权平均而得到, 依此类推, 图像大小以 $1/4$ 倍率减小。设图像元素的横纵坐标分别用 x, y 表示, 则层间的运算可以表示为: 在计算得出金字塔图像后, 就可以用多分辨率结构的光流来

计算光流场。其计算流程如下^[6]:

(1) 在最高层 n 采用 Lucas-Kanade 算法得到光流^[7]。

(2) 在第 i 层时:

① 获取第 $i+1$ 层的光流 $u_{i+1}(x, y), v_{i+1}(x, y)$;

② 用双线性插值计算 $u_i^*(x, y), v_i^*(x, y)$;

③ 将 $u_i^*(x, y), v_i^*(x, y)$ 都乘以 2;

④ 用 $u_i^*(x, y), v_i^*(x, y)$ 补偿第 i 层当前帧的金字塔图像;

⑤ 在补偿后的金字塔图像间用标准 Lucas-Kanade 光流计算 $u_i'(x, y), v_i'(x, y)$;

⑥ 最终计算得到第 i 层光流场:

$$u_i(x, y) = u_i^*(x, y) + u_i'(x, y)$$

$$v_i(x, y) = v_i^*(x, y) + v_i'(x, y)$$

(3) 迭代步骤(2), 直至计算出最终第 0 层的光流场。

对处于复杂背景中的目标图像, 外界的干扰因素一般是比较多的, 所以引入金字塔式多分辨率结构, 更显得合适与可靠。根据实际需要, 本文选用了三层金字塔图像来计算相邻两帧间的光流场。当相邻两帧图像间的运动不超过 5 个像素时, 算法能获得较好的精度。

2 运动目标检测与分割

基于光流法的运动检测采用了运动目标随时间变化的光流特性, Meyer^[8]等通过计算位移向量光流场来初始化基于轮廓的跟踪算法, 从而有效地提取和跟踪运动目标。该方法的优点是在摄像机运动存在的前提下也能检测出独立的运动目标, 即不需要图像间的特征对应。

当然, 在运动变化检测中还有一些其它的方法, 如 Friedman 与 Russell 利用扩展的 EM(Expectation Maximization)算法^[9], 为每个像素建立了混合高斯分类模型, 该模型可以自动更新, 并能自适应地将每个像素分类为背景、影子或者运动前景, 在目标运动检测缓慢的情况下亦能较好地完成运动区域的分割, 并可以有效地消除影子的影响; 另外, Friedman N 和 Russell S 也提出了一种新颖的基于数学形态学的场景变化检测算法, 在变化的环境条件下获得了相当稳定的分割效果。

本文采用了文献[4]提出的基于像素位移进行目标分割的方法。在理论上, 视频帧中每像素的最小位移为 1.0 像素。因此经过预处理后, 非零值位移的点组成了运动点集合, 即

$$T(u, v) = \begin{cases} 1, & \text{if } |(u, v)| \geq 1.0 \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

式中, (u, v) 为通过光流场得到的每一像素的位移, $T(u, v)$ 将运动像素从静止的背景中分割出来。图 1 显示了测试原图及其运动目标分割后的结果。

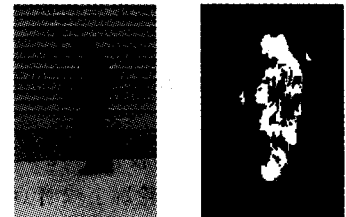


图 1 测试原图及其运动点集合

3 步态表征及特征分析

3.1 人体区域划分

传统的步态识别方法大多以身体质心作为中心点, 通过

计算侧影轮廓上关键点到质心的距离和两腿与质心形成的夹角作为步态识别的特征。但人在行走过程中,双手的摆动及双脚的交替运动导致身体轮廓质心在垂直和水平方向上产生平移。同样,当手上抱着物体时,侧影轮廓受到物体的影响,质心位置则更容易被改变。而质心位置的改变直接影响特征值的计算,因此需要采用新的坐标点作为中心点,根据 Lee^[5]提出的方法将人体轮廓划分成 7 个不同的子区域。如图 2 所示,头部区域为 1,左上半身区域为 2,右上半身区域为 3,左大腿区域为 4,右大腿区域为 5,左小腿区域为 6,右小腿区域为 7。

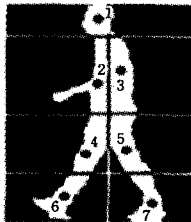


图 2 身体区域分割

3.2 人体子区域的椭圆拟合

考虑对于侧影的 7 个区域,利用椭圆对各个子区域进行拟合,拟合的重点是如何确定椭圆的参数。视频光流场提供了非静止点集的运动信息,利用点光流 $|u, v|$ 对每一像素进行加权,如图 3(a)所示。



(a) 光流 $|u, v|$ 加权后的图像 (b) 椭圆拟合后人体侧影图像

图 3

椭圆拟合的过程涉及到前景像素均值和协方差矩阵的计算,假设 $I(x, y)$ 表示某一区域像素值,前景像素值为 1.0,背景像素的值为 0,则加权后图像在区域 i 的质心 $(\bar{x}_i, \bar{y}_i)$ 为:

$$\bar{x}_i = \frac{1}{N} \sum_{x,y} |u, v| I(x, y) x, \bar{y}_i = \frac{1}{N} \sum_{x,y} |u, v| I(x, y) y$$

式中,该区域内所有像素总数为 $N = \sum_{x,y} I(x, y)$ 。

该区域的协方差矩阵为:

$$\begin{bmatrix} a & d \\ c & b \end{bmatrix} = \frac{1}{N} \sum_{x,y} |u, v| I(x, y) \begin{bmatrix} (x - \bar{x}_i)^2 & (x - \bar{x}_i)(y - \bar{y}_i) \\ (x - \bar{x}_i)(y - \bar{y}_i) & (y - \bar{y}_i)^2 \end{bmatrix}$$

该协方差矩阵的特征值 λ_1, λ_2 与对应的特征向量 v_1, v_2 表示椭圆长短轴的长度和方向。特征值之比为 $l_i = \frac{\lambda_1}{\lambda_2}$; 长轴的方向 $a_i = \text{angle}(v_1) = \arccos\left(\frac{v_1 \cdot x}{|v_1|}\right)$, 其中, x 表示单位矢量 $[1, 0]$, 图 3(b) 表示人体子区域的椭圆拟合后的结果。

3.3 步态特征的提取与识别

在视频序列中,人在行走过程中随着姿态的不断变化,序列中各帧图像子区域中的步态侧影也是变化的。假设点 $(\bar{x}_{iw}, \bar{y}_{iw})$ 表示未经光流加权的区域 i 质心坐标,点 $(\bar{x}_i, \bar{y}_i)$ 表

示光流加权后区域 i 质心坐标,图 4 给出了一个步态序列中人体侧影下肢部分右区域中 $\bar{y}_{iw}$ 与 $\bar{y}_i$ 值随帧序变化的情况,很明显地可以观察到两坐标之间的周期性规律。取一步态序列,起始点为步态周期的起始,长度为一个步态周期,设为 L 。提取 v_i 为第 i 区域的特征向量:

$$v_i = [\bar{x}_i - \bar{x}_{iw}, \bar{y}_i - \bar{y}_{iw}, l_i, a_i], i = 1, \dots, 7 \quad (7)$$

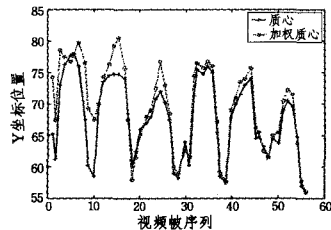


图 4 $\bar{y}_{iw}$ 和 $\bar{y}_i$ 值随帧序变化的情况

人体模型共由 7 个椭圆组成,因此 t 时刻的人体姿态可由一个向量表示:

$$V_t = (v_1, \dots, v_7) \quad (8)$$

取一个周期长度为 L 的步态序列,并从中得到各帧中的人体姿势表达 V_t ,将其按时间顺序排列,组成一个 $L \times 28$ 的矩阵作为此步态序列的特征向量。由于两个待比较步态周期子序列的长度通常不相等,如果对于测试序列中每个姿态都在参考序列中根据最近距离匹配,往往会出现动作匹配误对应。采用动态时间规整(dynamic time warping, DTW)技术来解决该问题^[2]。DTW 概念简单、算法鲁棒,即使测试序列模式与参考序列模式的时间尺度不一致,只要时间次序的约束存在,DTW 仍能较好地完成测试序列和参考序列之间的模式匹配。

最后,本文将求解所得的 L 个步态序列的特征向量送入最近邻即 NN(Nearest Neighbor)分类器进行分类,并以欧式距离来度量特征向量间的相似性。

4 实验与分析

本文采用了中国科学院自动化研究所提供的 NLPR 数据库^[10]进行实验。NLPR 中的步态图像序列是使用固定在三角架上的数字摄像机(Panasonic NV-DX100EN)在户外环境中捕捉的,行人相对于摄像机而言以侧面、正面和倾斜 3 个视角行走(相对于图像平面分别 $0^\circ, 90^\circ, 45^\circ$)。该数据库包含 20 个人,每个人每个视角 4 个序列,总计包括 240 个序列。图像序列以 25 帧每秒的速度拍摄,原始尺寸为 352×240 ,且平均长度约为 100 帧。在检测并提取步态侧影之后,为了避免人行方向对识别效果的影响,本文通过图像的镜像翻转,把各不同视角下的行走方向统一为从右往左。

实验中使用 NN 分类器,并采用留一校验法(leave-one-out cross validation)来获取步态识别率的无偏估计。在步态库中每次留出一个作为测试样本,余下的作为训练样本,最后依据与余下样本的相似性来分类被留出的样本。

表 1 给出了本文算法的正确分类率 CCR(Correct Classification Rate),对比了其他文献中的算法在使用相同数据库条件下的分类结果。表 2 给出了本文的方法和文献^[11]所列出的 4 种方法的识别结果。从表 1 和表 2 看出,我们提出的方法相比之下可以取得较为满意的结果。此外,图 5 使用国

(下转第 292 页)

personal communicate devices [C]//Proceedings of International Symposium on Consumer Electronics, 1997

- [3] Chang L P, Kuo T W. An Adaptive Striping Architecture for Flash Memory Storage Systems of Embedded Systems [C]//8th IEEE Real-time and Embedded Technology and Applications Symposium(RTAS). September 2002;187-196
- [4] Inoue A, Wong D. NAND Flash Applications Design Guide [R]. Toshiba America Electronic Components, Inc, 2003
- [5] Kawaguchi A, Nishioka S, Motoda H. A Flash-memory based File System [C]//Proceedings of the 1995 USENIX Technical Conference, January 1995;155-164
- [6] Intel Corporation. Ftl Logger Exchanging Data With Ftl Systems [R]. 1995
- [7] Intel Corporation. Understanding the Flash Translation Layer Specification [EB/OL]. <http://www.intel.com/design/fl-comp/applnots/297816.htm>, 1998
- [8] M-systems. Flash-Memory Translation Layer for NAND Flash (NFTL) [R]. 1998
- [9] WoodHouse D. Jffs: The Journaling Flash File System [EB/OL]. <http://sources.redhat.com/jffs2/jffs2-html/>

- [10] Bityutskiy A B. JFFS3 Design Issues. Version 0. 32(draft) [EB/OL]. <http://www.linux-mtd.infradead.org/tech/JFFS3design.pdf>
- [11] Alpha One Limited. Yet Another Flash Filing System [EB/OL]. <http://www.yaffs.net/yaffs-overview>, 2006
- [12] Wu M, Zwaenepoel W. eNvy: A Non-volatile Main Memory Storage System [C]//Proceedings of the Sixth International Conference on Architectural Support for Programming Languages and Operating Systems, 1994;86-97
- [13] Lee S W, Moon B. Design of flash-based DBMS: An in-page logging approach[C]//Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data, Beijing, China, 2007;55-66
- [14] 周大, 梁智超, 孟小峰. HF_Tree: 一种闪存数据库的高更新性能索引结构[J]. 计算机研究与发展, 2010, 47(5): 832-840
- [15] Wu C H, Kuo T W, Chang L P. An Efficient B-Tree Layer Implementation for Flash-memory Storage Systems [J]. ACM Transactions on Embedded Computing Systems, 2007, 6(3)
- [16] 卢泽萍, 孟小峰, 周大. HV_Recovery: 一种闪存数据库的高效恢复方法[J]. 计算机学报, 2010, 33(12): 2258-2266

(上接第 277 页)

际上在人脸识别算法中通用的分类性能度量方法 ROS(rank order statistic)^[12], 给出了本文提出方法的 CMS(cumulative match scores)即累积匹配分值图。CMS 定义为一个测试度量的实际类别在它的最前 k 个匹配值之间的累积概率, CCR 等价于 $k=1$ 时的 CMS。

表 1 不同算法正确分类率的比较

分类器	算法	CCR(%)
NN	Little et al. [4]	77.5
	Lee et al. [5]	78.25
	本文方法	84.25

表 2 COT 方法的压缩性能比较

算法	Top1(%)	Top5(%)	Top10(%)
STC, NN, No validation	65.00	---	---
Wang 等的方法	68.75	---	---
NED, NN, No validation	65.00	---	---
NED, NN, With validation	70.00	---	---
BenAbdelkader 等的方法	72.50	88.75	96.25
Collins 等的方法	71.25	78.75	87.5
Philips 等的方法	78.75	91.25	98.75
本文的方法	84.25	90.55	96.5

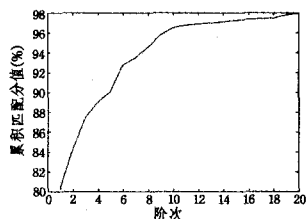


图 5 累积匹配分值图

结束语 本文提出了一种新的鲁棒性的步态识别算法, 这种算法基于鲁棒性较强的动力学特征光流, 并建立多区域椭圆模型的人体结构模型, 将目标人体区域部分按人体结构特点划分为多个子区域, 每个子区域通过结合光流特征的椭圆模型进行拟合。使用中国科学院自动化研究所提供的 NL-

PR 数据库进行实验, 实验通过最近邻分类器 NN 进行分类, 取得了较理想的识别效果。

参 考 文 献

- [1] Liu Z, Sarkar S. Effect of silhouette quality and hard problems in gait recognition[J]. IEEE Trans. Systems Man and Cybernetics, part B: Cybernetics, 2005, 35(2): 170-183
- [2] Yu S, Tan D, Huang K, et al. Reducing the Effect of Noise on Human Contour in Gait Recognition[C]//Internat. Conf. on Biometrics, 2007;338-346
- [3] 方政. OpenCV 技术在数字图像处理中的应用[J]. 北京教育学院学报: 自然科学版, 2011(01): 7-11
- [4] Varadhan K. The NS Manual[EB/OL]. <http://www.isi.edu/nsnam/ns/doc/>, 2010-02-08
- [5] 任亚峰, 郑金华. 多目标进化算法鲁棒性实验研究[J]. 计算机应用研究, 2011(02): 451-454
- [6] Chen Chang-hong, Liang Ji-min, Zhao Heng, et al. Frame Difference Energy Image for Gait Recognition with Incomplete Silhouettes[J]. Pattern Recognition Letters, 2009, 30(11): 977-984
- [7] Lucas B, Kanade T. An iterative image registration technique with an application to stereo vision[C]//Proc of DARPA U Workshop, 1981;121-130
- [8] 赵永伟, 张二虎, 鲁继文, 等. 多特征和多视角信息融合的步态识别[J]. 中国图像图形学报, 2009, 14(3): 388-392
- [9] McLachlan G, Krishnan T. The EM Algorithm and Extensions [Z]. Wiley Interscience, 1997
- [10] 郑帅. CASIA 步态数据库申请说明[EB/OL]. <http://www.cb-sc.ia.ac.cn/china/Gait%20Database%20CH.asp>, 2011-03-03
- [11] Chen Chang-hong, Liang Ji-min, Zhao Heng, et al. Frame Difference Energy Image for Gait Recognition with Incomplete Silhouettes[J]. Pattern Recognition Letters, 2009, 30(11): 977-984
- [12] Dwass M. Simple random walk and rank order statistics[J]. Ann. Math. Statist., 1967, 38: 1042-1053