

多层时空切片联合的快速人体跟踪算法

杨景景* 苏小红 马培军

(哈尔滨工业大学计算机科学与技术学院 哈尔滨 150001)

摘要: 针对传统人体跟踪方法中目标模型复杂、计算量大等问题, 该文提出一种无目标模型的多层时空切片联合的人体跟踪算法。用多层时空切片中的多个动态区域表示人体, 区域的选择无需使用任何预定义的目标区域模型。使用时空切片方法在图像序列空间中提取多层次水平时空切片图像, 在每层时空切片图像中, 检测和跟踪潜在的运动区域, 并根据区域运动一致性和空间一致性关系, 将多个区域关联成不同的人体目标, 实现多个人体目标的跟踪, 从而将 XYT 3 维空间中的人体跟踪问题转化为多个 XT 2 维空间的区域联合跟踪问题。实验表明, 该算法降低了跟踪的轨迹误差, 满足实时性跟踪要求, 同时通过多区域的联合增强了跟踪算法的抗干扰能力, 即在人体部分区域丢失的情况下仍能有效跟踪。

关键词: 目标跟踪; 人体跟踪; 时空切片; 目标映射; 层次聚类

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2012)10-2382-07

DOI: 10.3724/SP.J.1146.2011.01313

Combination of Multiple Spatial-temporal Slices for Fast Human Tracking

Yang Jing-jing Su Xiao-hong Ma Pei-jun

(School of Computer Science and Technology Harbin Institute of Technology, Harbin 150001, China)

Abstract: Considering the high computational cost and complex object representation problems in human tracking, this paper presents a model-free tracking approach using a combination of multiple spatial-temporal slices. The human is represented with a variable number of components in different spatial-temporal slice images. The component initialization requires no pre-defined object part model. By introducing the spatial-temporal slice method, the original image sequence volume is divided into multiple horizontal spatial-temporal slice images. In each slice image, candidate components are detected and tracked across frames. A combination scheme is proposed to assemble these components into different human objects based on their motion and position consistence. Thus, the traditional human tracking issue in the XYT 3D space is transformed into a combined component tracking issue in the XT 2D space. Experiments show that the proposed method reduces the trajectory errors, is real-time computational efficient and robust to human component missing.

Key words: Object tracking; Human tracking; Spatial-temporal slice; Object correspondence; Hierarchical clustering

1 引言

人体跟踪一直是目标跟踪领域中一个具有挑战性的研究课题。许多计算机应用技术的实现都依赖于准确的人体跟踪结果, 例如行为理解、安全监控和智能交通系统等。目前主要的人体跟踪方法有基于特征点的跟踪方法、基于剪影的跟踪方法和基于区域的跟踪方法^[1]。

基于特征点的跟踪算法^[2]使用多个特征点表示人体目标, 通过找到最优的多特征点的关联关系实

现人体跟踪。这种方法无需建立目标模型, 计算相对简单, 但是不能提供准确的目标区域信息, 因此不能适于实际人体跟踪的要求。与之相反, 基于剪影的跟踪算法^[3]通过轮廓进化或形状匹配的方法能够跟踪出精确的人体轮廓, 但是相应地该类算法需要建立复杂的目标轮廓模型, 模型需要大量的初始化和训练, 不能适用于实时的人体跟踪要求。

基于区域的跟踪算法是以上两类算法在区域准确率和计算复杂度上的一个平衡, 是目前应用比较广泛的一类人体跟踪算法, 该类算法一般使用模板和密度外观模型来表示人体区域。其中比较经典的有均值漂移算法和粒子滤波算法等。均值漂移算法^[4]采用取自圆形区域内的加权颜色直方图表示目标,

2011-12-12 收到, 2012-07-02 改回

国家自然科学基金(61175027)资助课题

*通信作者: 杨景景 yangjjhit@hotmail.com

并通过计算均值漂移偏移量来估计目标下一时刻的位置来实现跟踪。粒子滤波器算法^[5]通过多个加权粒子的形式建立目标概率密度方程，采用非参数化的方法实现目标状态空间估计，可以解决非线性、非高斯分布问题，具有更广泛的应用性和鲁棒性。此外，多模型粒子滤波算法^[6]、粒子概率密度假设滤波器^[7]、高斯粒子联合概率数据关联滤波算法^[8]和准蒙特卡罗滤波算法^[9]等也相继提出以解决复杂环境下的机动多目标跟踪问题。

基于区域的跟踪算法仍然在一定程度上依赖于准确的目标外观模型，并且需要对模型进行初始化和训练，增加了算法的计算量。针对以上问题，本文提出无目标模型的多区域联合的人体跟踪算法，在保证区域跟踪准确率的前提下降低跟踪所需的计算量。本文利用时空切片^[10]的方法，将人体表示为在不同时空切片图像中的多个区域(本文称为切片)的组合，切片无需使用任何预定义的目标模型。首先，在各时空切片图像中通过最小化一个改进的目标映射代价方程，确立人体切片间的映射关系，实现人体切片的跟踪。然后，根据人体的不同切片运动和空间一致性，使用聚类算法组合人体的多个切片的跟踪结果，重构人体轮廓，进而实现整个人体的跟踪。

2 时空切片方法的引入

假设视频是一个 $\mathcal{R}(x, y, t)$ 的 3 维图像序列，其中 (x, y) 为图像维， t 为时间维。那么一个水平时空切片图像就是固定 y 坐标后连续地从一个视频中提取相同位置的一行像素点组合得到的一个 $\mathcal{R}(x, y = y_k, t)$ 2 维图像。设切片图像 \mathcal{R}_0 源自原始视频中 $y = y_0$ 的区域，则图像 \mathcal{R}_0 中 $t = t_i$ 行对应的是原始视频中 $t = t_i$ 帧(实际对应 $t = t_i$ 帧图像中 $y = y_0$ 的那一行像素)。

如果将人体近似地看成是一个垂直的“条形”刚体目标，则当人体在视频中运动时，人体在不同时空切片图像中生成的人体切片不仅具有相同的水平坐标，而且也具有一致的运动模式^[11]。图 1 显示的是人体在 4 个不同水平时空切片图像中的轨迹对比。

如果将视频 $\mathcal{R}(x, y, t)$ 中第 i 帧的第 j 个人体表示成一个 2 维函数 $\gamma_j^i(x, y)$ 的形式， γ 既可代表一片区域也可代表一段轮廓。传统的人体跟踪问题实际上就是在图像序列中连续寻找一个 γ 人体序列的问题，可表示为

$$\langle \gamma_j^i(x, y), \gamma_j^{i+1}(x, y), \gamma_j^{i+2}(x, y), \dots \rangle \quad (1)$$

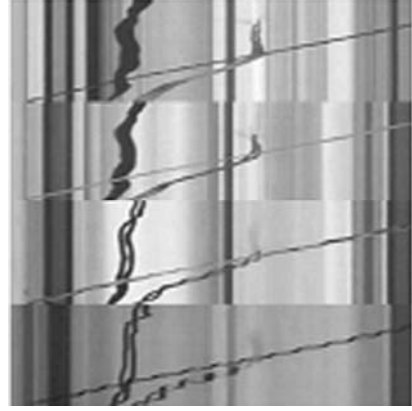


图 1 在 $y = 160, 150, 140, 130$ 时的水平时空切片图像

本文引入时空切片方法，将原始的 $\mathcal{R}(x, y, t)$ 图像序列分割成水平时空切片图像的集合，即

$$\{\mathcal{R}(x, y = y_k, t) | 1 \leq k \leq n\} \quad (2)$$

其中 n 表示时空切片的切分数目，且 $y_{k+1} > y_k$ 。相应地，视频中的每个 2 维 $\gamma_j^i(x, y)$ 人体也被分割成人体切片的集合，形如

$$\{\gamma_j^i(x, y = y_k) | 1 \leq k \leq n\} \quad (3)$$

其中 $\gamma_j^i(x, y = y_k)$ 表示位于时空切片图像 $\mathcal{R}(x, y = y_k, t)$ 中的人体切片； x 的取值范围可以为空，表示该时空切片图像中没有对应 j 人体的人体切片。

这样，人体跟踪就由寻找 γ 人体序列变成寻找人体切片集合的序列，即

$$\begin{aligned} & \langle \{\gamma_j^i(x, y = y_k) | 1 \leq k \leq n\}, \\ & \{\gamma_j^{i+1}(x, y = y_k) | 1 \leq k \leq n\}, \\ & \{\gamma_j^{i+2}(x, y = y_k) | 1 \leq k \leq n\}, \dots \rangle \quad (4) \end{aligned}$$

对于每个切片集合中的任意两个人体切片 $\gamma_j^i(x, y = y_k)$ 和 $\gamma_j^l(x, y = y_l)$ 。如果 $k \neq l$ ，则它们始终位于不同的时空切片图像中，彼此互不影响，对这两个人体切片的跟踪完全可以并行地在不同的时空切片图像中进行。因此，式(4)中的人体切片集合的序列就可以转化为人体切片序列的集合，形如

$$\langle \{\gamma_j^i(x, y = y_k), \gamma_j^{i+1}(x, y = y_k), \gamma_j^{i+2}(x, y = y_k), \dots\} | 1 \leq k \leq n \rangle \quad (5)$$

其中人体切片序列

$$\langle \gamma_j^i(x, y = y_k), \gamma_j^{i+1}(x, y = y_k), \gamma_j^{i+2}(x, y = y_k), \dots \rangle \quad (6)$$

在对应的时空切片图像 $\mathcal{R}(x, y = y_k, t)$ 中使用本文提出的目标映射算法实现。

3 人体切片的检测与跟踪

在时空切片图像中，人体切片是一个区别于背

景的,分布在某一 x 区间内的多个像素点的集合。人体切片在图像中所占的区域一般较小,但是又与纯粹的点目标不同,它还包含一定的形状、颜色和纹理等信息。因此针对这一特点,对时空切片图像 $\mathcal{R}(x, y = y_k, t)$ 中的第 i 行的第 j 个人体切片 γ_j^i , 本文使用下式来描述人体切片:

$$\gamma_j^i = [H_j^i \ W_j^i \ L_j^i \ T_j^i \ R_j^i \ Q_j^i(u)]^T \quad (7)$$

其中 H_j^i 表示当前切片在原始视频中的 y 坐标,即 y_k ; W_j^i 表示当前切片的宽度; L_j^i 表示当前切片的中心 x 坐标; T_j^i 表示当前切片的平移运动参数; R_j^i 表示当前切片旋转运动参数; $Q_j^i(u)$ 表示当前切片的颜色直方图。实验中,为了提高精确度,计算当前切片 5 行邻域内的颜色信息,即在 $5 \times W_j^i$ 邻域内计算颜色直方图。

3.1 人体切片检测

如果将时空切片图像看成是一个在时间域上的“行”序列,那么参照视频帧图像序列中的前景和背景定义,“行”就相当于序列中的“帧图像”,而人体切片就相当于“帧图像”中的前景目标。因此,本文使用如下所示的混合高斯背景模型对时空切片图像中的“行”进行建模,进而检测人体切片。

$$p(x_N) = \sum_{j=1}^k w_j \eta(x_N; \theta_j) \quad (8)$$

其中 w_j 为第 j 个高斯成分的权重系数; $\eta(x_N; \theta_j)$ 表示第 j 个成分的正态分布。对不同 y 取值的时空切片图像,算法都需要训练一个高斯背景模型。

值得注意的是,由于时空切片图像中的“行”仅包含有限的颜色信息,为了提高检测准确率和获得更丰富的颜色信息,本文在高斯背景建模及颜色直方图提取时,在当前行的相邻 m 行邻域内进行特征提取(实验中 $m=5$)。

3.2 人体切片跟踪

人体切片的跟踪实际上就是建立多个切片目标之间的一对一映射的问题。本文通过最小化目标映射代价方程来确立多个切片目标之间的映射关系。为了提高映射准确率,在现有的几何约束代价方程的基础上,本文提出新的颜色直方图相似性约束,并使用以下两种约束的组合来计算目标映射代价。

(1)最近一致的几何约束^[2]。本约束假设目标始终沿着一条最近一致的路径运动。也就是说,如果帧间的时间间隔足够小,那么可以认为人体切片在下一时刻将始终停留在离前一时刻最近的位置上,而且它们的运动方向和速度都不会发生大的变化。

设 γ_j^i 表示第 i 行的第 j 个切片, Φ^k 表示从第 k 行到第 $(k+1)$ 行的切片的映射关系。则最近一致约

束的代价方程 δ 定义为

$$\begin{aligned} \delta(\gamma_p^{k-1}, \gamma_{\Phi^{k-1}(p)}^k, \gamma_r^{k+1}) &= \frac{\left\| \gamma_p^{k-1} \gamma_{\Phi^{k-1}(p)}^k - \gamma_{\Phi^{k-1}(p)}^k \gamma_r^{k+1} \right\|}{\sum_{x=1}^m \sum_{z=1}^l \left\| \gamma_x^{k-1} \gamma_{\Phi^{k-1}(x)}^k - \gamma_{\Phi^{k-1}(x)}^k \gamma_z^{k+1} \right\|} \\ &+ \frac{\left\| \gamma_{\Phi^{k-1}(p)}^k \gamma_r^{k+1} \right\|}{\sum_{x=1}^m \sum_{z=1}^l \left\| \gamma_{\Phi^{k-1}(x)}^k \gamma_z^{k+1} \right\|} \end{aligned} \quad (9)$$

其中, m 和 l 分别是第 $(k-1)$ 行和第 $(k+1)$ 行中的人体切片个数; $1 \leq p \leq m$, $1 \leq r \leq l$ 。

(2)颜色直方图相似性约束。不同于纯粹的点目标,人体切片还包含一定的颜色信息。因此,为了进一步提高映射准确率,本文在前一约束基础上增加一种新的颜色直方图相似性约束,对目标进行颜色直方图相似性判断。其中直方图相似性使用巴氏距离定义。

令 $Q_j^i(u)$ 表示切片 γ_j^i 的颜色直方图。那么直方图相似性约束代价方程 φ 可以定义为

$$\varphi(\gamma_p^{k-1}, \gamma_{\Phi^{k-1}(p)}^k, \gamma_r^{k+1}) = \frac{\sum_{u=1}^h (Q_{\Phi^{k-1}(p)}^k(u) Q_r^{k+1}(u))}{\sum_{x=1}^m \sum_{z=1}^l \sum_{u=1}^h (Q_{\Phi^{k-1}(x)}^k(u) Q_z^{k+1}(u))} \quad (10)$$

其中 m , l , p 和 r 的定义同式(9); u 表示颜色直方图的灰度级分量, h 表示直方图的灰度级。

由以上两个约束方程组合计算得到最终的目标映射代价函数 ϕ 如下:

$$\phi(\gamma_p^{k-1}, \gamma_{\Phi^{k-1}(p)}^k, \gamma_r^{k+1}) = \lambda \delta + \beta \varphi \quad (11)$$

其中 λ 和 β 分别是以上两种约束的权重系数,且 $\lambda + \beta = 1$ 。实验发现,几何约束对映射准确率的影响要大于颜色直方图相似性约束,因此 $\lambda > \beta$; 并且 β 取值过大或过小时,映射准确率都较低。实际上, λ 取 0.6~0.7, β 取 0.3~0.4 时,准确率比较理想。本文中, $\lambda = 0.6$ 。使用非递归的贪心搜索算法最小化 $\sum \phi$ 求得其中可能的一个最优解。

4 人体切片的聚类与重构

基于人体切片运动与空间一致性,建立人体切片到候选人体目标的隶属关系,将人体切片集合划分为多个人体类,然后在此基础上重构相应的人体区域信息,实现人体切片到完整人体的组合跟踪。

4.1 人体切片聚类

算法使用式(7)中的 3 个参数 (L_j^i T_j^i R_j^i) 来表示第 i 行第 j 个人体切片的当前运动状态,进行人体切

片聚类。 T_j^i 和 R_j^i 的具体定义如下：

$$T_j^i = \overline{\gamma_j^{i-1} \gamma_j^i} \quad (12)$$

$$R_j^i = \overline{\gamma_j^{i-1} \gamma_j^i - \overline{\gamma_j^{i-2} \gamma_j^{i-1}}} \quad (13)$$

其中 γ_j^{i-2} , γ_j^{i-1} 和 γ_j^i 可以由切片映射关系确定。

图 2 显示的是包含两个运动人体的人体切片 LTR 分布图。其中不同的坐标点代表不同的人体切片。图中的坐标点自然聚集为两个区域，每个区域代表一个人体。其中同一人体内的坐标点，分布集中，具有较小的类内距离；而不同人体间的坐标点，距离较远，具有较大的类间距离。由于聚类使用的是与尺度无关的切片中心 x 坐标及运动参数等信息，算法具有处理一定尺度变化范围内的不同人体聚类的能力。而在实际视频监控环境中，由场景远近带来的人体尺度变化是有限的，因此本文算法能够满足实际监控环境的跟踪需要。本文使用层次聚类方法在 LTR 3 维空间中进行人体切片聚类，并使用欧氏距离表示切片间的距离度量。由于未知聚类数目，根据层次聚类的类间距离来判定聚类终止条件。

4.2 人体轮廓重构

本文使用式(7)中的 3 个参数 (H_j^i W_j^i L_j^i)，计算聚类后的人体切片的边缘点，来重构人体的完整轮廓。理论上，时空切片的采样频率越高，采样间隔越小，算法生成的轮廓越接近于目标的真实轮廓。具体重构算法如下：

步骤 1 对集合内的所有人体切片，按各切片 y 坐标 H_j^i 的升序对 γ_j^i 进行排序；

步骤 2 在最大 H_{\max}^i 和最小 H_{\min}^i ，连接 $(L_{\max}^i - W_{\max}^i / 2, H_{\max}^i)$ 和 $(L_{\max}^i + W_{\max}^i / 2, H_{\max}^i)$ 得到最高边缘，连接 $(L_{\min}^i - W_{\min}^i / 2, H_{\min}^i)$ 和 $(L_{\min}^i + W_{\min}^i / 2, H_{\min}^i)$ 得到最低边缘；

步骤 3 对每个 $H_{\min}^i < H_j^i < H_{\max}^i$ ，寻找当前高度处的最左边缘点坐标 $LP_j^i = (L_j^i - W_j^i / 2, H_j^i)$ 和最右边缘点坐标 $RP_j^i = (L_j^i + W_j^i / 2, H_j^i)$ ；

步骤 4 对每个 $H_{\min}^i < H_j^i < H_{\max}^i$ ，连接 LP_j^i

和 RP_{j+1}^i 得到人体左边缘，连接 RP_j^i 和 RP_{j+1}^i 得到人体右边缘。

5 实验结果和分析

本文采用 CAVIAR 视频对算法进行测试，其中，CAVIAR 视频分辨率 384×288 ，帧率 25 fps，分别采自架设在监控区域正面和拐角的两个摄像头。实验中，视频取 64 个时空切片图像^[3]。

5.1 轨迹误差以及鲁棒性分析

图 3(a)和图 3(b)是两个视频的人体跟踪轨迹与真实轨迹的对比，其中加号轨迹代表真实轨迹，点状轨迹代表算法估计轨迹。图 3(a)对应的视频 Video1 仅包含一个人体目标，而图 3(b)对应的视频 Video2 则包含多个人体目标。图 3(c), 3(d)分别显示的是图 3(a)和图 3(b)两段视频的轨迹误差曲线(以像素为单位)。从图中可以看出，在某些帧，由于复杂背景干扰造成的切片丢失及检测误差，导致估计区域大于或小于真实区域，进而区域中心点坐标出现偏差，轨迹误差增大。但总体来说，其轨迹误差依然保持在较低的水平上，平均轨迹误差分别为 2.389 像素和 5.27 像素。值得注意的是，图 3(c)中的轨迹误差递增现象是由 Video1 人体目标面向镜头行走造成尺寸递增而引起的，算法本身没有跟踪迭代或者累加的过程。

由于复杂背景干扰而引起人体的部分切片丢失，使算法在某些帧内跟踪到的只是人体的部分区域，在结果上表现为跟踪的人体的中心点发生微小变化，导致跟踪的轨迹是一条围绕着真实轨迹上下扰动的曲线，平均每帧的轨迹误差变化并不明显，从而使其相对于其他跟踪算法而言对环境干扰具有更强的鲁棒性。

图 4 显示的是不同切片丢失情况下的人体区域跟踪结果。从图 4 结果可以看出，本文的跟踪算法对环境干扰引起的人体切片丢失具有一定的鲁棒性。具体地讲，如图 4(b)，位于人体内部的躯干切片的丢失不会对人体区域的大小和位置产生任何影

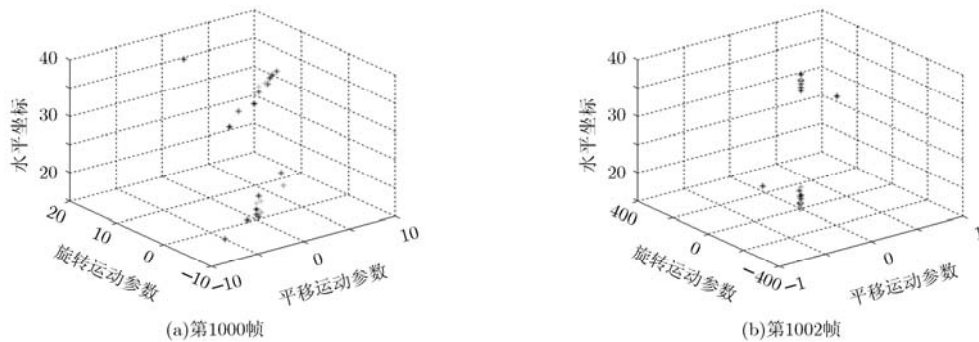


图 2 包含两个行人的视频中人体切片 LTR 分布示意图

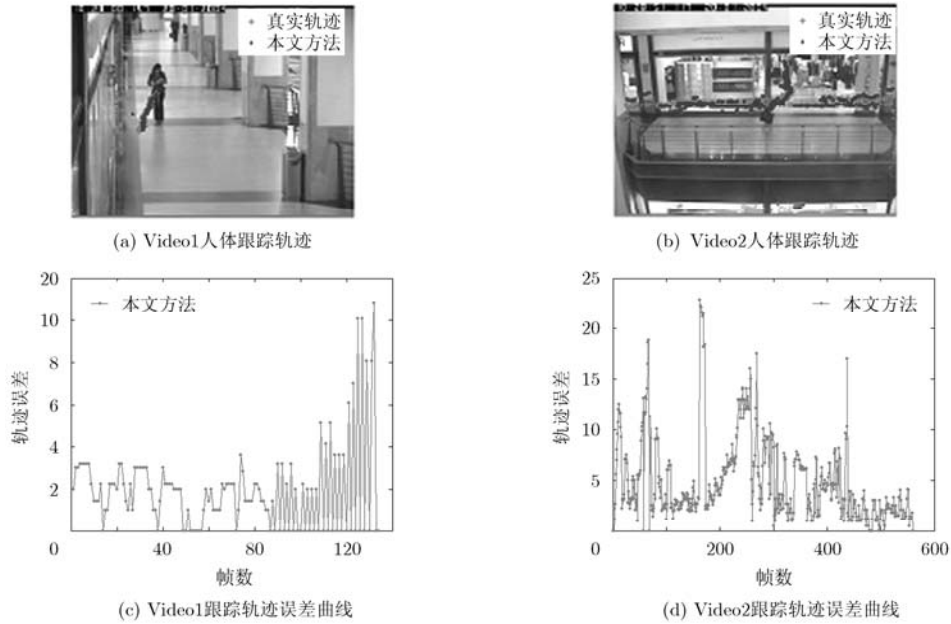


图 3 Video1 和 Video2 人体跟踪轨迹及轨迹误差

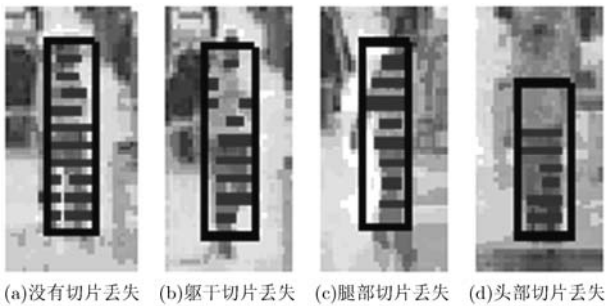


图 4 不同切片丢失情况下的人体区域跟踪结果

响,因而不会影响轨迹误差。而腿部切片和头部切片对生成最终的人体区域的影响相对较大。其中,当某帧中的切片丢失数量较少(如图 4(c)所示)时对轨迹误差的影响在误差范围内可以忽略不计,当某帧中的切片数量丢失较多(如图 4(d)所示)时,会对当前帧的轨迹误差产生一定的影响,但由于每一帧检测的切片数量是变化的,切片丢失的情况仅在个别帧中出现,因此个别帧中切片的丢失对平均每帧轨迹误差影响不大。

5.2 与其他方法的对比

为了验证本文提出的跟踪算法的有效性,与现有的其他目标跟踪方法如 MS^[4], MSPF^[14], CCMPF^[15]和 MSFG^[16]进行了对比。

表 1 显示的是 CAVIAR 数据库中 4 个视频的平均轨迹误差、运行时间、查准率和查全率^[1]的对比结果,其中查准率和查全率用于验证估计区域的准确性。从表 1 都可以看出,本文跟踪算法的平均轨迹误差均低于 CCMPF, MS, MSFG 和 MSPF 算法;

在平均运行时间上,本文跟踪算法与 CCMPF, MS 和 MSFG 算法接近,远优于 MSPF 算法;查全率明显优于其他算法,查准率与其他算法接近,其中平均查准率为 93.33%,而平均查全率为 60.62%。平均查全率较低的原因是:本文算法提取的是人体轮廓,而真实值给定的是人体边界盒矩形区域。如图 5 所示,轮廓剪影区域为本文跟踪结果,而矩形框区域为真实结果。本文的人体轮廓能更精确地描述目标形状,这也是论文的一个优势,但是该轮廓始终位于边界盒的内部,自然造成查全率较低,而查准率则会较高。而其他算法由于采用位于目标内部的椭圆模型来表示目标区域,造成查全率也普遍较低。

6 结束语

本文采用多层时空切片方法,将传统的 XYT 3 维空间上的人体跟踪问题转化为多个 XT 2 维空间中的人体切片联合跟踪问题。使用多个人体切片的联合来表示人体目标,无需初始化和训练人体目标模型。实验结果表明,本文算法在轨迹误差上优于传统跟踪算法,满足实时性视频处理要求,并且对于由于环境干扰引起的人体区域丢失具有一定的鲁棒性。现阶段本文仅使用运动和颜色特征进行跟踪,并且特征提取也是在当前帧的切片较小邻域内进行,信息量不够丰富,不能处理较大尺度变化场景下的人体跟踪。为了进一步提高跟踪准确率和鲁棒性、解决尺度变化问题,在下一步的研究工作中考虑结合连续多帧的切片区域,在其基础上提取更丰富的颜色、梯度或纹理特征用于跟踪。



图 5 部分人体跟踪结果(轮廓剪影为本文估计区域, 矩形框为真实区域)

表 1 Video1-Video4 实验对比结果

		轨迹误差(像素)	查全率(%)	查准率(%)	运行时间(ms)
Video1	本文方法	2.389	89.284	97.413	34.489
	CCMSPF	3.690	67.119	98.357	32.161
	MS	19.038	13.079	98.475	45.146
	MSFG	4.938	68.620	95.734	45.520
	MSPF	15.695	16.949	100.000	268.090
Video2	本文方法	5.270	62.945	92.334	31.777
	CCMSPF	9.184	31.546	90.867	48.133
	MS	13.060	30.915	83.907	52.216
	MSFG	9.321	35.075	91.734	53.071
	MSPF	9.930	34.446	82.053	334.160
Video3	本文方法	6.325	51.143	94.559	30.352
	CCMSPF	8.927	29.141	97.403	25.137
	MS	11.195	31.643	89.798	49.735
	MSFG	7.434	34.934	94.351	50.192
	MSPF	9.768	33.192	82.406	192.740
Video4	本文方法	8.712	53.028	83.556	33.445
	CCMSPF	11.300	21.648	89.890	48.270
	MS	13.391	23.696	84.866	54.551
	MSFG	11.538	29.891	85.964	54.713
	MSPF	10.120	22.972	83.227	391.270

参 考 文 献

[1] Yilmaz A, Javed O, and Shah M. Object tracking: a survey[J]. *ACM Computing Surveys*, 2006, 38(4): 13-58.

[2] Vidal R and Ma Y. A unified algebraic approach to 2-d and 3-d motion segmentation[C]. *European Conference on*

Computer Vision, Prague, Czech, 2004: 1-15.

[3] Yilmaz A, Li X, and Shah M. Contour based object tracking with occlusion handling in video acquired using mobile cameras[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(11): 1531-1536.

[4] Comaniciu D, Ramesh V, and Meer P. Kernel-based object tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(5): 564-575.

[5] Isard M and MacCormick J. BraMBLE: a Bayesian multiple-blob tracker[C]. *IEEE International Conference on Computer Vision*, Vancouver, Canada, 2001, 2: 34-41.

[6] 鉴福升, 徐跃民, 阴泽杰. 多模型粒子滤波跟踪算法研究[J]. *电子与信息学报*, 2010, 32(6): 1271-1276.

Jian Fu-sheng, Xu Yue-min, and Yin Ze-jie. Research of multiple model particle filter tracking algorithms[J]. *Journal of Electronics & Information Technology*, 2010, 32(6): 1271-1276.

[7] 连峰, 韩崇昭, 刘伟峰, 等. 基于 SMC-PHDF 的部分可分辨的群目标跟踪算法[J]. *自动化学报*, 2010, 36(5): 731-741.

Lian Feng, Han Chong-zhao, Liu Wei-feng, et al. Tracking partly resolvable group targets using SMC-PHDF[J]. *Acta Automatica Sinica*, 2010, 36(5): 731-741.

[8] 张俊根, 姬红兵, 蔡绍晓. 基于高斯粒子 JPDA 滤波的多目标跟踪算法[J]. *电子与信息学报*, 2010, 32(11): 2686-2690.

Zhang Jun-gen, Ji Hong-bing, and Cai Shao-xiao. Gaussian particle JPDA filter based multi-target tracking[J]. *Journal of Electronics & Information Technology*, 2010, 32(11): 2686-2690.

[9] 赵玲玲, 马培军, 苏小红. 一种快速准蒙特卡罗粒子滤波算法[J]. *自动化学报*, 2010, 36(9): 1351-1356.

Zhao Ling-ling, Ma Pei-jun, and Su Xiao-hong. A fast quasi-Monte Carlo-based particle filter algorithm[J]. *Acta*

- Automatica Sinica*, 2010, 36(9): 1351-1356.
- [10] Adelson E H and Bergen J R. Spatiotemporal energy models for the perception of motion[J]. *Journal of the Optical Society of America A*, 1985, 2(2): 284-299.
- [11] 杨景景, 苏小红, 马培军. 基于时空切片轨迹分析的复杂人体运动跟踪[J]. *信号处理*, 2012, 28(2): 246-256.
Yang Jing-jing, Su Xiao-hong, and Ma Pei-jun. Complex human motion tracking based on spatio-temporal slice trajectory analysis[J]. *Signal Processing*, 2012, 28(2): 246-256.
- [12] Shafique K and Shah M. A non-iterative greedy algorithm for multi-frame point correspondence[C]. *IEEE International Conference on Computer Vision*, Nice, France, 2003: 110-115.
- [13] Yang J, Su X, and Ma P. Fast pedestrian detection using slice-based motion analysis[C]. *International Conference on Pervasive Computing, Signal Processing and Applications*, Harbin, China, 2010: 74-77.
- [14] Zhang B, Tian W, and Jin Z. Joint tracking algorithm using particle filter and mean shift with target model updating[J]. *Chinese Optics Letters*, 2006, 4(10): 569-572.
- [15] Silva D D, Cesar R M, Vieira M B, et al. Tracking and matching connected components from 3D video[C]. *Brazilian Symposium on Computer Graphics and Image Processing*, Natal, Brazil, 2005: 257-264.
- [16] Li L, Huang W, Gu I Y H, et al. Statistical modeling of complex backgrounds for foreground object detection[J]. *IEEE Transactions on Image Process*, 2004, 13(11): 1459-1472.
- 杨景景: 男, 1982年生, 博士生, 研究方向为计算机视觉、行人检测、目标跟踪.
- 苏小红: 女, 1966年生, 教授, 博士生导师, 研究方向为图形与图像处理、软件缺陷检测、信息融合、计算智能等.
- 马培军: 男, 1963年生, 教授, 博士生导师, 研究方向为空间计算、信息融合、色彩匹配、图像处理及智能控制等.