

文章编号: 1001 - 2486(2011)02 - 0044 - 05

## 双向时空连续性轨迹片段关联的目标跟踪方法\*

王江峰, 张茂军, 包卫东, 熊志辉

(国防科技大学 信息系统与管理学院, 湖南 长沙 410073)

**摘要:**提出一种使用双向时空连续性关联轨迹片段的目标跟踪方法。首先对检测结果进行简单的帧间匹配关联,生成可靠的轨迹片段;然后对每个轨迹片段通过卡尔曼滤波以及有权重的均值法分别计算修正轨迹片段的速度、位置与颜色特征;最后通过计算轨迹片段之间的双向时空连续性迭代关联,找到最符合时空连续性的轨迹片段关联。实验证明本文方法可以有效解决目标间以及目标被背景遮挡问题,实现对目标的稳定跟踪。

**关键词:**目标跟踪;遮挡;时空连续性;轨迹片段关联

**中图分类号:**TP391.4 **文献标识码:**A

## Object Tracking Using Bidirectional Spatio-temporal Continuity Tracklet Association

WANG Jiang-feng, ZHANG Mao-jun, BAO Wei-dong, XIONG Zhi-hui

(College of Information Systems and Management, National Univ. of Defense Technology, Changsha 410073, China)

**Abstract:** An object tracking algorithm by associating tracklets with the best bidirectional spatio-temporal continuity was proposed. First, reliable tracklets were generated by a primitive frame-by-frame association; then tracklet's motion, position and color features were computed and refined by applying Kalman filter and weighted mean method respectively; finally, the best spatio-temporal association of tracklets was achieved through an iterative association by computing spatio-temporal continuity between tracklets. Experimental results prove that multiple objects can be successfully tracked under occlusion both by other object and scene object.

**Key words:** object tracking; occlusion; spatio-temporal continuity; tracklet association

目标跟踪是视频分析与智能视频监控系统中一个基础且关键的任务<sup>[1]</sup>。目标检测和跟踪方法都存在不稳定性,因此基于统计的目标跟踪方法成为研究的热点。这类方法通过模型预测来解决检测值的不确定性,一般是利用状态矢量空间来表示目标的特征,包括位置、速度、外形等。比较经典的有两种:联合概率数据关联滤波(JPDAF, Joint Probabilistic Data Association Filters)技术以及多假设跟踪(MHT, Multi-Hypothesis Tracking)。JPDAF方法对于不同的假设计算一个概率,错误假设的概率会随着时间的推进慢慢降低,最终被排除掉<sup>[2]</sup>。在JPDAF的基础上,加上时空约束实现了对多个目标的稳定跟踪。但JPDAF方法假设被检测的目标个数是固定的,一旦场景中目标个数发生变化,很容易产生跟踪错误。MHT同JPDAF的原理类似,不同的是MHT可以适应场景目标变化的情景,即场景中可以有目标的离开或

者新目标的进入。这两种方法共同的问题是随着目标个数的增多,假设空间呈指数增长,计算代价也随之增长<sup>[3]</sup>。通过采用最优的 $k$ 个假设来缩小搜索空间<sup>[4]</sup>,使用马尔科夫链蒙特卡洛(MCMC)方法对搜索空间进行抽样降低计算维度。近来,轨迹片段关联(tracklet associating/stitching)技术成为基于统计的关联方法研究中一个新的热点。这类方法一般分为两个阶段,首先在目标能够实现稳定跟踪时生成轨迹片段,然后再通过高层的轨迹片段关联形成目标的长轨迹,实现对目标的稳定跟踪<sup>[5]</sup>。使用局部粒子滤波与全局轨迹片段关联实现多个目标跟踪<sup>[6]</sup>,使用轨迹片段关联解决目标长时间重叠<sup>[7]</sup>,用轨迹片段关联实现多个摄像机协同跟踪<sup>[9]</sup>,使用费用流来解决数据关联问题实现对行人的跟踪,文献[10]使用轨迹片段关联方法实现目标跟踪与简单事件检测。

本文提出一种基于计算修正的双向时空连续

\* 收稿日期:2010-09-07

基金项目:国家自然科学基金资助项目(60705013);中国博士后科学基金资助项目(20070410977)

作者简介:王江峰(1980—),男,博士生。

性的轨迹片段关联方法。轨迹片段的速度和位置特征使用卡尔曼滤波进行计算和修正,颜色特征使用带有权重的均值统计模型计算和修正。通过逐步增加被关联轨迹片段之间时间间隔实现迭代关联,每次关联通过计算轨迹片段之间的时空连续性将短轨迹关联形成更长的轨迹,最终得到目标轨迹。实验证明本文方法能够解决目标之间重叠遮挡以及目标被背景遮挡等问题,实现对目标的稳定跟踪。

## 1 使用双向时空连续性轨迹片段关联的目标跟踪

### 1.1 方法概述

我们的方法是通过初步关联生成轨迹片段和更高层的轨迹片段关联实现对目标轨迹的跟踪。首先,对于输入的前景区域做简单的帧间匹配,生成置信度高的轨迹片段,以牺牲轨迹片段长度为代价保证每个轨迹片段只属于一个目标。初步匹配关联时采用前景区域的位置,面积以及颜色特征。然后,对初步关联生成的每个轨迹片段,使用卡尔曼滤波对轨迹片段的位置和速度进行计算修正,使用带有权重的均值统计模型计算修正轨迹片段的颜色特征,计算生成轨迹片段的属性集。接着通过迭代计算片段之间正反两个方向的时空连续性来寻找轨迹片段的可能关联对象,得到属于同一目标的轨迹片段集合,通过关联每个集合内的轨迹片段,最终得到目标的长轨迹。

### 1.2 初步关联生成轨迹片段

为了生成可信度高的轨迹片段以保证每个轨迹片段只属于一个目标,我们只对连续两帧图像的检测值进行匹配关联。鉴于连续两帧图像之间目标的位置与面积特征不会发生强烈的变化,初步关联匹配时采用位置、面积再加上颜色特征的相似性实现。获得的前景检测值可表示为  $D = \{d_i\}$ ,则初步关联的相似性可以表示为

$$\begin{aligned} P_{\text{associate}}(d_1, d_2) \\ = A_{\text{position}}(d_1, d_2) A_{\text{area}}(d_1, d_2) A_{\text{color}}(d_1, d_2) \end{aligned} \quad (1)$$

位置与面积的相似性计算可分别表示为

$$\begin{aligned} A_{\text{position}}(d_1, d_2) \\ = \exp\left[-\frac{(x_1 - x_2)^2}{\sigma_x^2}\right] \exp\left[-\frac{(y_1 - y_2)^2}{\sigma_y^2}\right] \end{aligned} \quad (2)$$

$$A_{\text{area}}(d_1, d_2) = \exp\left[-\frac{(s_1 - s_2)^2}{\sigma_s^2}\right] \quad (3)$$

颜色的相似性采用巴氏距离

$$A_{\text{color}}(d_1, d_2) = \sqrt{1 - BC(d_1, d_2)} \quad (4)$$

其中  $BC(d_1, d_2)$  是巴氏系数 (Bhattacharyya Coefficient):

$$BC(d_1, d_2) = \sum_{n=1}^N \sqrt{q_{(d_1)_n} p_{(d_2)_n}} \quad (5)$$

其中  $N$  是直方图的级数,  $d_1$  和  $d_2$  有相同的灰度级数(如 256)。  $q_{(d_1)_n}$  是第  $n$  级灰度所占  $d_1$  直方图的权重,  $p_{(d_2)_n}$  是第  $n$  级灰度所占  $d_2$  直方图的权重。

初步关联时阈值  $\theta$  的选取决定了生成轨迹片段的数量。对于视频 OneLeaveShop1cor(如图 2 所示)初步关联阈值与所产生轨迹片段数量的关系如图 1 所示。一方面为了获得置信度较高的轨迹片段,初步关联的阈值要尽量大;另一方面,阈值过大,会生成过多的轨迹片段。阈值越接近 1,轨迹片段越接近前景检测值,即未做任何关联。从图 1 可以看出,当关联阈值达到 0.8 左右时轨迹片段数目的增加速度出现拐点,实验中我们一般选择拐点处的  $\theta$  值作为初步关联阈值。

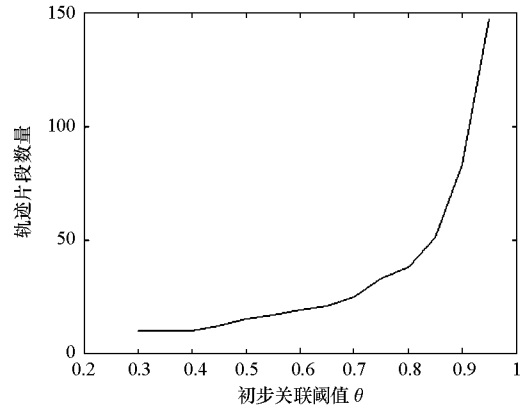


图 1 初步关联阈值与轨迹片段数量的关系曲线

Fig.1 Relational curve of primitive association threshold and number of tracklet

通过初步关联可以得到轨迹片段集合  $S_{\text{Tracklet}} = \{T_1, T_2, \dots, T_{N_{\text{Tracklet}}}\}$ ,其中  $N_{\text{Tracklet}}$  是轨迹片段的个数。对每个轨迹片段都计算生成记录其位置、大小、颜色以及速度特征的向量集:

$$T_i^{\text{attr}} = \langle T_i \text{sf}, T_i \text{ef}, T_i \text{vx}_i, T_i \text{vy}_i, T_i \text{area}_i, T_i \text{color}_i \rangle \quad (6)$$

其中  $T_i \text{sf}$  是轨迹片段  $i$  的起始帧,  $T_i \text{ef}$  是轨迹片段  $i$  的终止帧;  $T_i \text{vx}_i$  与  $T_i \text{vy}_i$  分别是轨迹片段  $i$  在第  $t$  帧的速度  $X$ 、 $Y$  轴分量,  $T_i \text{area}_i$  是轨迹片段  $i$  在第  $t$  帧的大小,  $T_i \text{color}_i$  是轨迹片段  $i$  在第  $t$  帧的颜色向量。每段轨迹片段的速度以及位置通过卡尔曼滤波进行预测和校正,轨迹片段的颜色使用

带有权重的均值统计模型,通过统计轨迹片段各帧得到修正后的轨迹片段颜色。

我们对 CAVIAR<sup>[8]</sup> 的一个数据集 OneLeaveShop1cor 进行了轨迹生成实验(见图 2):



图2 轨迹片段生成结果(不同颜色深度表示不同的轨迹片段)

Fig.2 Tracklet generation result (different color represents different tracklet)

### 1.3 双向时空连续性的轨迹片段关联

经过初步关联得到了轨迹片段集合  $S_{Tracklet} = \{T_1, T_2, \dots, T_{N_{Tracklet}}\}$ ,在计算每一对轨迹片段之间的关联相似性时,我们在颜色约束的基础上,加上时空约束。通过计算选择最符合时空连续性的轨迹片段进行关联。时间连续性可以表示为

$$P_t(T_i, T_j) = \begin{cases} 1 & \text{if } f_{j_s} - f_{i_e} < \tau^{gap} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

其中  $P_t(T_i, T_j)$  是轨迹片段关联的时间相关性,  $f_{i_e}$  是轨迹片段  $i$  的结束帧,  $f_{j_s}$  是轨迹片段  $j$  的起始帧,  $\tau^{gap}$  是关联时间阈值(两个轨迹片段的最大间隔帧数)。空间的连续性可以表示为

$$P_s(T_i, T_j) = P_s^{forward}(T_i, T_j) P_s^{backward}(T_i, T_j) \quad (8)$$

$$\begin{cases} P_s^{forward}(T_i, T_j) \\ = \exp\left[-\frac{(\hat{T}_i px_{\Delta f} - T_j sx)^2}{\sigma_x^2}\right] \exp\left[-\frac{(\hat{T}_i py_{\Delta f} - T_j sy)^2}{\sigma_x^2}\right] \\ P_s^{backward}(T_i, T_j) \\ = \exp\left[-\frac{(\hat{T}_j px_{\Delta f} - T_i ex)^2}{\sigma_x^2}\right] \exp\left[-\frac{(\hat{T}_j py_{\Delta f} - T_i ey)^2}{\sigma_x^2}\right] \end{cases} \quad (9)$$

其中  $P_s^{forward}$  和  $P_s^{backward}$  分别是正、反向空间关联相似性预测位置,计算如下:

$$\hat{T}_i px_{\Delta f} = T_i ex + T_i vx \times \Delta f \quad (10)$$

$$\hat{T}_i py_{\Delta f} = T_i ey + T_i vy \times \Delta f \quad (11)$$

当计算正向时空连续性时,使用被关联两个轨迹片段中前者的速度和尾部位置数据以及后者的头部位置数据进行计算;计算反向时空连续性时,采用后者的速度和头部位置数据以及前者的尾部位置数据进行计算。

两个轨迹的颜色相似性计算与初步关联的计算类似,匹配计算时采用有权重的均值法修正后的颜色直方图。两个轨迹片段关联的相似性则可

视频包含 295 帧图像,6 个移动目标。当初步关联阈值  $\theta = 0.8$  时共生成 38 个轨迹片段,初步关联阈值与轨迹片段数目的关系见图 1。

以表示为

$$P_{associate}(T_i, T_j) = P_t(T_i, T_j) P_s(T_i, T_j) P_c(T_i, T_j) \quad (12)$$

轨迹片段的关联是一个迭代过程,通过逐步增加关联时间阈值  $\tau^{gap}$  将时间间隔更大的轨迹片段关联起来(图 3)。这样可以有效解决基于关联的跟踪方法一个最大的缺陷,即随着目标个数的增多以及关联时间的增加,匹配算法的搜索空间会呈指数增长,算法的效率急剧下降。由于限制了轨迹之间的距离,每一次迭代时只对整个轨迹片段集合中的部分进行关联,将搜索空间划分为几个子空间,分步进行搜索,提高了算法的搜索效率。

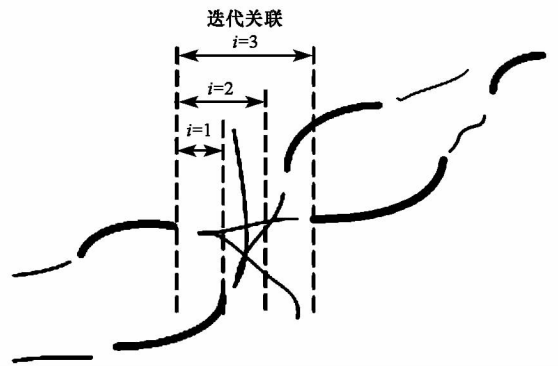


图3 轨迹片段迭代关联算法图示

(粗线条代表轨迹片段,细线条代表关联片段)

Fig.3 Graph model of iterative association process

(dark line = tracklet, thin line = association)

每次迭代的输入为上次关联形成的轨迹片段,第一次迭代的输入即为初步关联生成的轨迹片段集合  $\{T_i\}$ ,轨迹片段关联时间阈值的初始值  $\tau_0^{gap}$ ,距离增加步长  $\tau_{\Delta}^{gap}$  以及迭代次数  $N_{iter}$ 。每次迭代只计算间隔距离小于  $\tau_i^{gap}$  的轨迹片段相似性,输出关联后的轨迹片段集合作为下次迭代的输入,完成一次迭代。每迭代一次,间隔距离增加一个步长,逐渐将短的轨迹关联形成长的轨迹。

## 2 实验及讨论

采用公共视频数据集 CAVIAR<sup>[8]</sup> 对算法进行实验。CAVIAR 数据集由架设在购物中心走廊上方的单摄像机拍摄的视频序列组成,帧率为 25 帧/s,图像大小为 384 像素 × 288 像素。图像中行人之间以及行人被场景中固定物体的遮挡非常频繁。对数据集中两个视频序列进行了实验,视频一 OneLeaveShop1cor(图 4 第一行)包含 295 帧图像,视频二 WalkByShop1cor(图 4 第二、三行)包含 2360 帧图像。视频中最远端的行人由于太小(宽度小于 30 像素)或者只有局部出现在场景中,处理时未做考虑。

论文采用文献[9]中的评价标准:

- MT(Mostly Tracked):目标被成功跟踪帧数占目标实际存在帧数 80% 以上的轨迹个数;
- ML(Mostly Lost):目标被成功跟踪帧数占目标实际存在帧数 20% 以下的轨迹个数;
- PT(Partially Tracked):目标成功被跟踪帧数占目标实际存在帧数 20% ~ 80% 的轨迹个数;
- FRMT(Fragmentation):一条轨迹被分成了几段;
- IDS(ID Switches):轨迹标志交换的次数。

MT 的计算如下,其中  $O_i$  表示目标  $i$ :

$$MT = \left\lfloor \left\{ O_i \mid \frac{O_i \text{ 成功被跟踪的帧数}}{O_i \text{ 实际存在帧数}} > 80\% \right\} \right\rfloor \quad (13)$$

我们分别对 CAVIAR 数据集中的视频 OneLeaveShop1cor(295 帧)和 WalkByShop1cor(2360 帧)进行实验,行人的检测采用文献[12]中的方法。为了进行对比,将 MT、PT 以及 ML 三个指标转化成绝对数量与真实数据的百分比来评价,FRMT 与 IDS 由于采用绝对次数不具有可比性,未做对比,跟踪结果见表 1。

表 1 在 CAVIAR 数据集上的跟踪结果  
Tab.1 Tracking result on CAVIAR dataset

算法	MT	PT	ML
文献[11]方法	75.71%	17.86%	6.43%
文献[9]方法 1	74.29%	20.71%	5.00%
文献[9]方法 2	85.71%	10.71%	3.57%
初步关联(阈值 0.8)	66.67%	29.05%	4.29%
第一次迭代(阈值 30 帧)	80.00%	15.71%	4.29%
第二次迭代(阈值 75 帧)	84.29%	12.14%	3.57%



图 4 视频 OneLeaveShop1cor(第一行)与 WalkByShop1cor(第二、三行)的部分跟踪结果  
Fig.4 Tracking results, top row on OneLeaveShop1cor and bottom two rows on WalkByShop1cor

实验中初步关联阈值  $\theta$  取 0.8,迭代次数  $N_{iter}$  为 2,迭代关联中关联时间阈值  $\tau^{gap}$  分别为 30 和 75 帧,从表 1 可以看出从初步关联到迭代次数的增加,关联的效果越来越理想。以视频 OneLeaveShop1cor 为例,共有 6 个有效(远端小像

素目标未做处理)移动目标,实际轨迹为 6 条。初步关联时得到 38 条轨迹片段,第一次迭代后得到 12 条轨迹片段,第二次迭代后为 10 条轨迹片段。从结果看我们的方法优于文献[11]和文献[9]的方法 1,与文献[9]的方法 2 效果基本接近。部分

跟踪结果以及最终绘制轨迹如图4所示,第一行为视频 OneLeaveShop1cor 的跟踪结果,第二、三行为视频 WalkByShop1cor 的跟踪结果,相同颜色的框表示相同的目标。OneLeaveShop1cor 视频中行人被场景中的柱子遮挡,WalkByShop1cor 视频中行人之间的遮挡非常频繁,可以看到,通过轨迹片段关联方法对目标可以被稳定检测和跟踪到的部分进行关联,对丢失的部分根据轨迹双向的运动特征进行补充,可以实现对目标的稳定跟踪,有效处理目标之间遮挡以及目标被场景物体遮挡的情况,两个视频中所有目标的完整轨迹分别在最后一帧图像中绘出。

本文实验环境为 2.4GHz CPU、1GB 内存,检测部分使用文献[12]的 C++ 代码,对文中 384 像素 × 288 像素大小图像速度为 3.53 帧/s,关联部分使用 MATLAB7.1 实现,平均速度为 7.56 帧/s。

### 3 结束语

本文提出一种通过计算修正的双向时空连续性的轨迹片段关联方法,实现了对多个目标的稳定跟踪,有效解决了目标之间重叠遮挡,目标短暂消失以及被场景遮挡问题。轨迹片段的速度和位置特征通过卡尔曼滤波进行计算和修正,颜色特征通过有权重的均值法进行计算和修正。轨迹片段之间的关联使用迭代的方式,通过增加轨迹片段之间的间隔距离,将待匹配轨迹片段的搜索空间划分为几个子空间,提高了搜索效率。实验结果表明本文方法可以有效解决目标之间相互重叠以及目标被场景遮挡等问题。本文方法具有很强的扩展性,针对不同的应用,可以采用不同关联匹配属性以及不同的相似性度量来提高跟踪效果。

### 参考文献:

- [1] Alper Y, Omar J, Mubarak S. Object Tracking: A Survey [J]. ACM Computer Survey, 2006, 38(4): 13 - 57.
- [2] Bennett B, Magee D R, Cohn A G, et al. Enhanced Tracking and Recognition of Moving Objects by Reasoning about Spatio-temporal Continuity [J]. Image and Vision Computing, 2008, 26(1): 67 - 81.
- [3] Ingemar J C, Sunita L H. An Efficient Implementation of Reid's Multiple Hypothesis Tracking Algorithm and Its Evaluation for the Purpose of Visual Tracking [J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 1996, 18(2): 138 - 150.
- [4] Bardet F, Chateau T, Laprest'e J T. Illumination Aware MCMC Particle Filter for Long-term Outdoor Multi-object Simultaneous Tracking and Classification [C]//Proceeding of 12<sup>th</sup> IEEE International Conference on Computer Vision, 2009:1623 - 1630.
- [5] Xing J L, Ai H Z, Lao S H. Multi-object Tracking Through Occlusions by Local Tracklets Filtering and Global Tracklets Association with Detection Responses [C]//Proceeding of 2009 IEEE Computer Vision and Pattern Recognition, 2009:1200 - 1207.
- [6] Perera A, Srinivas C, Hoogs A, et al. Multi-object Tracking Through Simultaneous Long Occlusions and Split-merge Conditions [C]//Proceeding of 2006 IEEE Computer Vision and Pattern Recognition, 2006:666 - 673.
- [7] Liu S H, Lai S M. Schematic Visualization of Object Trajectories across Multiple Cameras for Indoor Surveillances [C]//Proceeding of 5<sup>th</sup> IEEE International Conference on Image and Graphics, 2009: 406 - 411.
- [8] CAVIAR[DB]. <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>.
- [9] Zhang L, Li Y, Nevatia R. Global Data Association for Multi-object Tracking using Network Flows [C]//Proceeding of 2008 IEEE Computer Vision and Pattern Recognition, 2008:666 - 673.
- [10] Wang J F, Zhang M J, Cohn A G. Object Tracking and Primitive Event Detection by Spatio-temporal Tracklet Association [C]// Proceeding of 5<sup>th</sup> IEEE International Conference on Image and Graphics, 2009: 457 - 462.
- [11] Wu B, Nevatia R. Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet Based Part Detectors [J]. International Journal of Computer Vision, 2007, 75(2), 247 - 266.
- [12] Lin Z, Davis L S. A Pose-invariant Descriptor for Human Detection and Segmentation [C]//Proceeding of 10<sup>th</sup> European Conference of Computer Vision, Marseilles, LNCS, 2008: 423 - 436.