

DOI: 10.13382/j.jemi.2017.06.013

一种基于随机场多运动目标跟踪算法^{*}

汪济洲^{1,2} 鲁昌华¹ 蒋薇薇¹

(1. 合肥工业大学 计算机与信息学院 合肥 230031; 2. 合肥学院 合肥 230061)

摘 要:多运动目标跟踪是智能视频监控系统的重要内容。由于目标数量众多且目标之间相互遮挡,会造成跟踪轨迹不连续,这是多目标跟踪问题的难点,也是当前流行算法无法解决的研究的问题。利用目标对相互之间的位置信息受到外界干扰较少的特点,提出基于条件随机场多目标跟踪。与以前针对所有目标的外形特征不同,算法对跟踪子对进行标示。采用分层结构,由低阶状态作为输入序列到高阶,通过代价函数最小化,从而得到最终连续跟踪轨迹。能有效解决复杂环境下多目标不连续的问题。算法对于外形类似的被跟踪目标以及相互遮挡严重的情况有比较明显的优势,与当前优秀的算法相比,实验表明本文算法在定量分析和定性分析上均有较好的结果。

关键词:多目标跟踪;条件随机场;跟踪子对

中图分类号: TP274.3; TN99 **文献标识码:** A **国家标准学科分类代码:** 520.604

New multi-target tracking algorithm based on conditional random field

Wang Jizhou^{1,2} Lu Changhua¹ Jiang Weiwei¹

(1. School of Computer and Information, Hefei University of Technology, Hefei 230031, China;

2. Hefei University, Hefei 230061, China)

Abstract: Multi-target tracking is very important in intelligent video surveillance system. Occlusions among many moving objects and objects with similar appearance that could result in incontinuous trajectory are challenging problems in this field. Based on the fact that the relative positions remain stable between two close pairs, a novel multi-target tracking algorithm based on conditional random field is presented in this paper. Unlike previous approaches which focus only on appearance and motion models for all targets, we consider discriminative features for distinguishing difficult pairs of targets with sets of labels. Multi-level is adopted by this means, in which the set of tracklets produced by previous level is used as an input. And tracking problem is transformed into an energy minimization problem including a set of unary function for a continuous trajectory and a set of pairwise function for pairs of tracklets. This new method is more powerful in deal with objects with similar appearance and badly occlusions compared to state-of-art methods. Qualitative and quantitative experimental results show that this new method has a better performance.

Keywords: multi-target tracking; conditional random field; pair of tracklets

1 引 言

多目标跟踪是近年计算机视觉分析中重要的课题。实际应用环境中,由于外界光线、背景变化或者出现遮挡情况,往往导致采用固定模型的跟踪算法无法有效实施跟踪,而采用更新模型的算法在这种情况下,往往会出现

模型漂移现象,导致失效^[1-2]。

在目标众多的时候,例如热闹的商场和演出场所,由于目标之间间距极小,相互遮挡现象及其严重,而且当目标外形相近时,很难利用外形特征将目标识别出来。过去通常采用跟踪对的方法来实现对于这种类型的多目标进行跟踪。当镜头发生移动导致视角发生变化时,目标跟踪会受到破坏,从而形成不连续的轨迹^[3-8]。但是,由

收稿日期:2016-12 Received Date: 2016-12
^{*} 基金项目:大学生创客实验室建设计划(2015ckjh068)、基于 QoS 的 4G LTE 系统切换策略的研究(KJ2015B1105908)资助项目

于目标间的相对位置不会因为视角的变化而变化,其变化较小。因此,相对位置可以在多目标跟踪得到应用。

基于以上事实,通过对目标对进行标记,可以充分利用目标对之间相对位置信息。而条件随机场是一种判别式模型,具有训练样本少的特点。因此,本文提出一种基于随机场 Conditional Random Filed-CRF 的多目标跟踪算法^[9-12]。与文献[13-14]不同,在本文算法中,所有的目标对构成 CRF 的节点,并用标签标示该节点的目标对是否相关,若相关,则标签为 1,否则,标签为 0。本文对针对每个节点的能量函数进行优化计算获得全局最优解。而 CRF 的边的能量函数标示目标对是否有联系。采用目标的外形特征和运动信息做为建模特征。

2 多目标跟踪的随机场 CRF 模型简介

对于一个视频序列,首先采用预训练的检测模型对每一帧图像进行检测。与 Kuo 提出的算法类似,将相邻的视频帧转换得到跟踪子序列。令跟踪子序列为 $T_i = \{d_i^t, \dots, d_i^t\}$,其中, T_i 为连续视频得到的跟踪子。 t_i^s 和 t_i^e 表示第 i 个跟踪子的起始帧和终止帧。 $d_i^t = \{p_i^t, s_i^t, v_i^t\}$,是指第 t 帧图像,所检测到的目标。其中, p_i^t 为位置, s_i^t 为大小信息,而 v_i^t 为速度^[15]。算法采取多层结构,每层的输入为前一层的输出 $S = \{T_1, T_2, \dots, T_n\}$ 。对于每个节点表示单个跟踪子 ($T_{i1} \rightarrow T_{i2}$)。引入标签 l_i ,如果 l_i 为 1,表示 T_{i1} 与 T_{i2} 为一个平滑轨迹,否则 l_i 为 0。目标是找出标签的最大似然概率,在给定输入跟踪子序列 S 情况下,似然函数为 L 则可以表示如下:

$$L^* = \operatorname{argmax}_L P(L | S) = \operatorname{argmax}_L \frac{1}{Z} \exp(-\Psi(L | S)) \tag{1}$$

式中: Z 为归一化算子, ψ 为代价函数。假设只考虑 2 个轨迹之间的关系,公式如下:

$$L^* = \operatorname{argmax}_L \Psi(L | S) = \operatorname{argmax}_L \sum_i U(l_i | S) + \sum_{i,j} B(l_i, l_j | S) \tag{2}$$

式中: $U(l_i | S) = -\ln P(l_i | S)$ 为随机场的节点能量函数,而 $B(l_i, l_j | S) = -\ln P(l_i, l_j | S)$ 表示跟踪子对的能量函数。式(2)中的第 1 部分表示两个跟踪子的关系,而第 2 部分表示跟踪子对之间的关系,利用条件随机场对多目标跟踪问题建模。随机场图 $G = \{V, E\}$,其中, $V = \{v_1, \dots, v_p\}$ 为图的节点,其含义为具有关联性的跟踪子,即 $v_i = (T_{i1} \rightarrow T_{i2})$,而边 $e_j = \{(v_{j1}, v_{j2})\}$ 是指节点也就是跟踪子对之间关联性。标签序列 $L = \{l_1, \dots, l_p\}$ 表示当前层的跟踪子是否有关联。其中, L 满足条件:

$$\begin{aligned} \sum_{v_i \in \text{Head}_{i1}} l_i &\leq 1 \ \& \ \sum_{v_i \in \text{Tail}_{i2}} l_i &\leq 1 \\ \text{Head}_{i1} &= \{(T_{i1} \rightarrow T_j) \in V\} \quad \forall T_j \in S \\ \text{Tail}_{i2} &= \{(T_j \rightarrow T_{i2}) \in V\} \quad \forall T_j \in S \end{aligned} \tag{3}$$

为了提高计算效率,算法不对整帧进行计算,而是对一个滑动的窗口进行计算。

3 基于随机场多运动目标跟踪算法

如图 1 所示,算法为分层结构,视频序列输入后,经过底层目标检测算子获得初级跟踪子。分成两类:一类

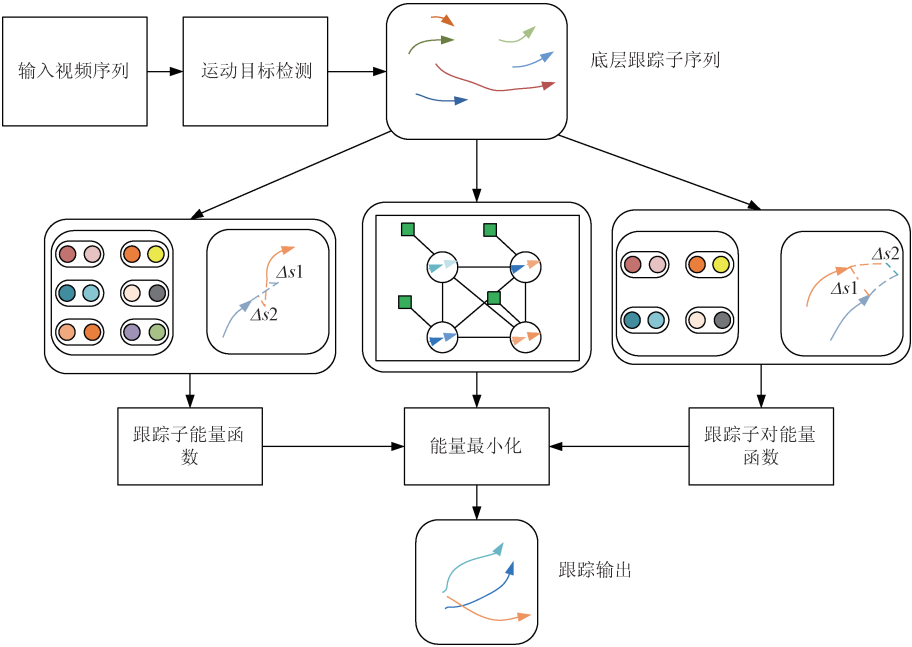


图 1 基于条件随机场多目标跟踪算法

Fig. 1 Block diagram of multi-track algorithm based on conditional random field

为全局跟踪子,目标为 CRF 节点,表示两个跟踪子是否为一个跟踪轨迹的概率;另一类是跟踪子对节点,表示一对跟踪是否为一对的概率。

3.1 CRF 图的创建

设输入为初级跟踪子序列 $S = \{T_1, \dots, T_N\}$ 利用条件随机场图对其建模。如果 T_i 的结束帧与 T_j 的起始帧的间距小于某个阈值,可以视为两个跟踪子互相关联。即 $0 < t_j^s - t_i^e < t_{\max}$ 其中, $t_{\max}$ 为阈值,而 t_i^e 是 T_i 的终止帧, t_j^s 为 T_j 的起始帧。因此,定义随机场的节点,记为 $V = \{v_i = (T_{i1} \rightarrow T_{i2})\}$ 满足 T_{i1} 与 T_{i2} 相互关联。使用跟踪子对定义随机场图边,这里先定义头相连和尾相连的概念,

如图 2 所示,两个跟踪子 T_i 和 T_j 如果满足以下条件:

$$t_i^s < t_j^e \ \& \ ||P_i^{t_i^e} - P_j^{t_j^e}|| < \gamma \min\{s_i^{t_i^e}, s_j^{t_j^e}\} \quad (4)$$

式中: γ 为阈值,本文取 3。式(4)说明 T_i 的起始轨迹与 T_j 起始轨迹头相连。同理,可以定义终止帧相连。因此,定义随机场的边:

$$E = \{v_i, v_j\} \quad \forall v_i, v_j \in V \quad (5)$$

式中: v_{i1} 和 v_{j1} 为轨迹头相连,或者 v_{i2} 和 v_{j2} 轨迹尾相连。因此,采用这样定义,将复杂难以识别多目标以目标对形式进行判断。

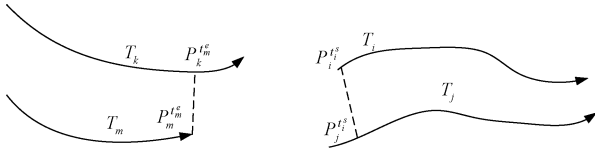


图 2 头相连和尾相连的定义

Fig. 2 Schematic diagram of definition of head linked together and tail linked together

3.2 单个跟踪子的学习算法

式(2)定义了跟踪子的能量函数。其中 $U(l_i | S) = -\ln P(l_i | S)$, 进一步分解成外形特征和运动信息特征。其定义如下:

$$U(l_i = 1 | S) = -\ln(P_m(T_{i1} \rightarrow T_{i2} | S) P_a(T_{i1} \rightarrow T_{i2} | S)) \quad (6)$$

式中: $P_m(\cdot)$ 是以运动信息特征的概率函数,而 $P_a(\cdot)$ 为外形特征的概率函数。如图 3 所示。

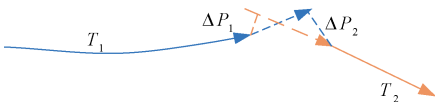


图 3 单个跟踪子为同一轨迹

Fig. 3 Schematic diagram of the same track with single tracklet

图 3 中跟踪子 T_1 和 T_2 之间为同一轨迹的概率以 $\Delta p_1, \Delta p_2$ 为基础,其定义如下:

$$\Delta p_1 = p^{\text{head}} - v^{\text{head}} \Delta t - p^{\text{tail}} \quad (7)$$

$$\Delta p_2 = p^{\text{tail}} + v^{\text{tail}} \Delta t - p^{\text{head}} \quad (8)$$

所以,概率函数定义如下:

$$P_m(T_{i1} \rightarrow T_{i2} | S) = G(\Delta p_1, \sum p) G(\Delta p_2, \sum p) \quad (9)$$

对于 $P_a(\cdot)$ 基于本文的 OLDAM。

3.3 双跟踪子对学习算法

根据以上观察结果,双跟踪子对可以用于识别相距很近的多目标。与单个跟踪子对类似,双跟踪子对也分成基于运动信息特征和外形特征。以双跟踪子对 (T_1, T_3) 和 (T_2, T_4) 为例,令:

$$t_x = \min\{t_1^e, t_2^e\}, t_y = \max\{t_3^s, t_4^s\} \quad (10)$$

T_1 和 T_2 为尾相连,预测在 t_y 时两者相对距离 Δp_1 , 计算 T_3 和 T_4 在 t_y 时刻的相对距离 Δp_2 , 其中, Δp_1 定义如下:

$$\Delta p_1 = (p_1^{t_y} + V_1^{\text{tail}}(t_y - t_1^e)) - (p_2^{t_y} + V_2^{\text{tail}}(t_y - t_2^e)) \quad (11)$$

如同单一跟踪对利用高斯函数一样,即定义概率如下:

$$P_m(T_{i1} \rightarrow T_{j1} | S) = G(\Delta p_1 - \Delta p_2, \sum p) \quad (12)$$

因此,如果 T_1, T_2 与 T_3, T_4 双跟踪子对,则, T_1 与 T_3 是同属一个跟踪子的可能性较高,同理, T_2 与 T_4 一样。这样,就实现了多目标的跟踪。

3.4 算法说明

单一跟踪子已经文献[1]被证明有效的,因此,在式(3)的约束下,可以使用匈牙利算法求得全局最优解。然后,将单一跟踪子能量函数按照从小到大的顺序排列。如下:

$$A = \{v_i = (T_{i1} \rightarrow T_{i2})\} \quad (13)$$

对于每一个被选中的节点,如果其能量函数更小,则保持变化。算法伪码如下:

Input: 当前 $S = \{T_1, T_2, \dots, T_n\}$, $G = \{V, E\}$

使用匈牙利算法在满足最小能量函数条件下,获得标签集合 L 。

按照单个跟踪子对的能量从小到大排序,

$\{v_1, v_2, \dots, v_k\}$ 计算能量 ϕ

for $i = 1, \dots, k$ do

$\phi' = \infty$

for $j = 1, \dots, t$ do

在满足式(3)条件下,切换 l_i 和 l_j 并计算新能量 Ω

if $\Omega < \phi'$ then $\phi' = \Omega$

if $\phi' < \phi$ then $\phi = \phi'$ 更新 L

output: L

4 实验结果

实验平台为 CPU 为 intel2. 4G, 内存 8 G, 采用 opencv2. 3。对于视频分析的效果, 一般很难使用单一指标来评判算法的效果。实验数据采用 TU Darmstadt。如表 1 所示, 定量分析指标采用以下指标: 1) 正确率: 正确

表 1 算法定量分析比较

Table 1 Quantitative analysis comparison of algorithm			
算法	正确率/%	每帧错误率	正确跟踪率/%
OLDAMS	75	0. 022	60
PRIMPT	78	0. 034	60
Offline CRF	80	0. 025	65
在线 CRF	81	0. 015	70

跟踪的帧数比例;2) 每帧错误率: 每帧平均错误率;3) 正确跟踪率: 正确跟踪超出 80% 的跟踪轨迹。

如图 4 所示, 定量分析结果表明, 在线 CRF 算法与 OLDAMS 和 PRIMPT^[2] 在正确率、每帧平均错误率和正确跟踪率上有大幅提高。

从定性上来说, 第一段视频背景不是很复杂, 但是, 行人相互靠近, 互相遮挡的现象比较严重, 所以在跟踪的时候, 跟踪轨迹算法一般比较容易丢失跟踪, 本文算法很好实施跟踪, 跟踪子对“1”和“2”, 以及“3”和“4”在算法结果中得到很好的检测与跟踪。第二段视频背景属于较为复杂街头, 而且多目标出现严重遮挡现象, 例如 578 帧中目标 9 被目标 2 完全遮挡, 本算法正确检测出来, 并且实施精确跟踪。基于 CRF 多目标跟踪算法对于 TUD 数据库的视频跟踪的计算速度为 10FPS, 而 ETH 为 6FPS。

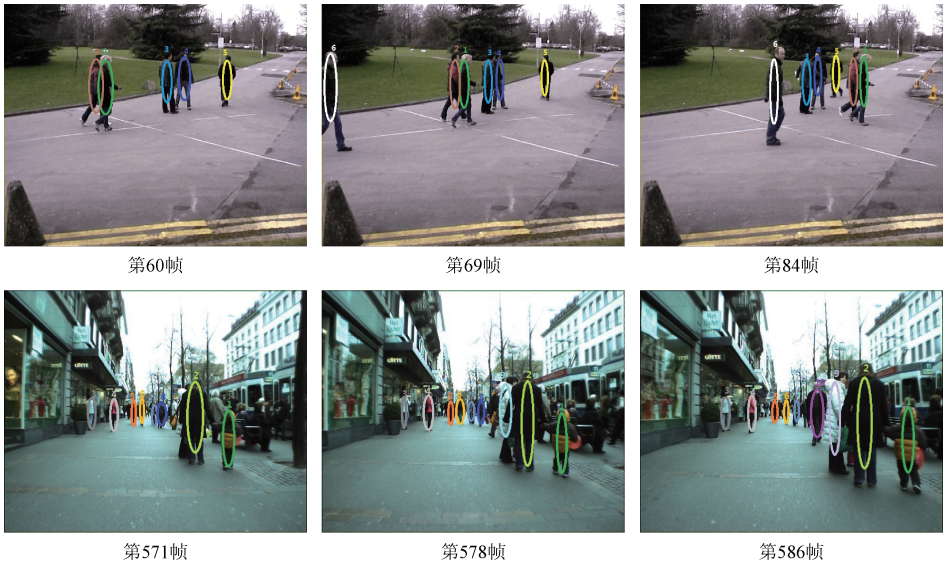


图 4 算法定性实验结果
Fig. 4 Qualitative experimental result of algorithm

5 结论

本文阐述了一种基于随机场的多目标跟踪算法, 这种算法针对复杂环境下众多运动目标相互粘连, 遮挡跟踪结果出现漏检、错检的问题, 采用单跟踪子用于追踪单一跟踪轨迹, 同时采用跟踪子对来实现正确跟踪多目标跟踪。与一般的全局描述算子不同, 基于外形特征模型基于单一跟踪子。并且, 对于复杂的被跟踪目标, 则使用跟踪子对来实现跟踪。算法采用一种新的算法来求解 CRF 的能量函数最小值。实验结果表明与当前比较流行的两种算法 OLDAMS 和 PRIMPT 相比, 本文算法对于复杂环境下多目标跟踪的效果更佳。

参考文献

[1] KUO C H, HUANG C, NEVATIA R. Multi-target tracking by online learned discriminative appearance models[C]. CVPR, 2010: 1-2, 4- 7.

[2] KUO C H, NEVATIA R. How does person identity recognition help multi-person tracking? [C]. CVPR, 2011: 1, 4, 6-8.

[3] ANDRIYENKO A, SCHINDLER K. Multi-target tracking by continuous energy minimization[C]. CVPR, 2011: 1-8.

[4] PERERA A G A, SRINIVAS C, HOOGS A, et al. Multi-object tracking through simultaneous long occlusions and split-merge conditions[C]. CVPR, 2006:

- 1-2, 5.
- [5] XING J, AI H, LAO S. Multi-object tracking through occlusions by local tracklets filtering and global tracklets association with detection responses [C]. CVPR, 2009: 1-2.
- [6] XING J, AI H, LIU L, et al. Multiple players tracking in sports video: a dual-mode two-way bayesian inference approach with progressive observation modeling [J]. IEEE Transactions on Image Processing, 2011, 20(6): 1652-1667.
- [7] ISARD M, BLAKE A. Condensation-conditional density propagation for visual tracking [J]. International Journal of Computer Vision, 1998, 29(1): 5-28.
- [8] ESS A, LEIBE B, SCHINDLER K, et al. Robust multiperson tracking from a mobile platform [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(10): 1831-1846.
- [9] 谢英红, 吴成东. 基于投影群和协方差流形双重建模的目标跟踪 [J]. 仪器仪表学报, 2014, 35(2): 375-379.
- XIE Y H, WU CH D. Object tracking with dual modeling based on projection group and covariance manifold [J]. Chinese Journal of Scientific Instrument, 2014, 35(2): 375-379.
- [10] 蒋鹏, 宋华华, 王兴民. 基于动态生成树和改进不敏卡尔曼滤波的传感器网络目标跟踪算法研究 [J]. 仪器仪表学报, 2015, 36(2): 415-421.
- JIANG P, SONG H H, WANG X M. Target tracking algorithm for sensor networks based on dynamic spanning tree and improved unscented Kalman filter [J]. Chinese Journal of Scientific Instrument, 2015, 36(2): 415-421.
- [11] 宋晓琳, 邬紫阳, 张伟伟. 基于阴影和类 Haar 特征的动态车辆检测 [J]. 电子测量与仪器学报, 2015, 29(9): 1341-1346.
- SONG X L, WU Z Y, ZHANG W W. Dynamic vehicle detection based on shadow and Haar-like feature [J]. Journal of Electronic Measurement and Instrumentation, 2015, 29(9): 1341-1346.
- [12] YANG B, HUANG C, NEVATIA R. Learning affinities and dependencies for multi-target tracking using a CRF model [C]. Proceedings of CVPR, 2011: 1-2, 4, 6.
- [13] SHITRIT H B, BERCLAZ J, FLEURET F, et al. Tracking multiple people under global appearance constraints [C]. Proceedings of ICCV, 2011: 1-2.
- [14] 高琳, 唐鹏, 盛鹏, 等. 复杂场景下基于条件随机场的视觉目标跟踪 [J]. 光学学报, 2010, 30(6): 1721-1727.
- GAO L, TANG P, SHENG P, et al. Visual object tracking based on conditional random field under complex scene [J]. Acta Optica Sinica, 2010, 30(6): 1721-1727.
- [15] PIRSIYAVASH h, RAMANAN D, FOWLKES C C. Globally-optimal greedy algorithms for tracking a variable number of objects [C]. Proceedings of CVPR, 2011: 1-2.

作者简介



image processing.



汪济洲, 合肥工业大学在读博士, 合肥学院讲师, 研究方向为数字图像处理。

Email: wangjizhou@hfu.edu.cn

Wang Jizhou, Ph. D. candidate in Hefei University of Technology, lecturer in Hefei University. His research interest is digital

鲁昌华, 现为合肥工业大学教授, 博导, 研究方向为信息处理。

Email: jsdzlch@hfu.edu.cn

Lu Changhua, professor and Ph. D. supervisor in Hefei University of Technology. His research interest is information processing.

蒋薇薇, 1978 年出生, 合肥工业大学在读博士, 合肥工业大学讲师, 研究方向为信号检测、智能信息处理。

Jiang Weiwei was born in 1978, Ph. D. candidate and lecturer in Hefei University of Technology. Her research interest is signal detection and intelligent information processing.