

DOI: 10.13382/j.jemi.2017.01.020

# 基于 EM 的多目标跟踪算法<sup>\*</sup>

周  娜  鲁昌华  徐婷佳  蒋薇薇  杜  雲  
(合肥工业大学  计算机与信息学院  合肥  230031)

**摘  要:**为了提高多目标跟踪的鲁棒性,增强目标之间的区别性,使用了一种基于能量最小化(energy minimization, EM)的多目标跟踪算法,不同于现有算法,本算法专注于将多目标跟踪中的复杂问题表示为能量函数的模型,模型中包括了更优的目标区分策略(相似度模型)。通过将每个能量函数成本值对应一个多目标的跟踪轨迹方案,算法将多目标跟踪问题转化为能量最小化的问题。在能量函数模型的优化方法上,算法采用共轭梯度算法和一系列的跳转运动来找到能量最小的值。公开数据集的实验结果证明了本算法的有效性,而且定量分析结果证明了本算法提高了目标与背景、目标之间的相互区别性从而与其他算法相比能获得更好的鲁棒性能。

**关键词:** 多目标跟踪;目标区分策略;能量最小化

**中图分类号:** TN911.73      **文献标识码:** A      **国家标准学科分类代码:** 510.4050

## Multi-target tracking algorithm based on EM method

Zhou Na  Lu Changhua  Xu Tingjia  Jiang Weiwei  Du Yun

(School of Computer and Information, Hefei University of Technology, Hefei 230031, China)

**Abstract:** In order to improve the multi-target tracking robustness and enhance the difference between the targets, this paper uses an energy minimization method for multi-target tracking. Different to the existing algorithm, the algorithm focuses on the representation of the complex problem in multi-target tracking as energy function model, which includes a better target segmentation strategy (similarity model). By assigns every possible solutions a cost (the “energy”), the algorithm transforms the multiple target tracking problem into an energy minimization problem. In the energy minimization optimization method, the algorithm uses the conjugate gradient algorithm and a series of jump moves to find the minimum energy value. The experimental results of open data demonstrate the effectiveness. And the quantitative analysis results show that this algorithm can improve the difference between targets or between target and background so as to obtain better robust performance compared with other algorithms.

**Keywords:** multi-target tracking; target discrimination strategy; energy minimization

## 1 引  言

目标跟踪一直以来都是计算机视觉应用中的重要内容,同时也是极具挑战性的内容课题。根据目标数量分类,目标跟踪可分为单目标跟踪与多目标跟踪。相对于单目标跟踪而言,要在错误检测和遮挡等复杂情况下准确的划分各检测目标,多目标跟踪无疑具有更大的难度。近期的跟踪算法大部分基于检测优先<sup>[1-4]</sup>,通过背景减法

或者分类器预先将目标检测出来,而后再估测其轨迹。文献[5]采用全局优化方法来实现对一个给定的时间窗口内的所有目标轨迹进行求解,为了降低模型的复杂度将目标的状态集约束在可数的状态空间中。Leibe 等人<sup>[6-7]</sup>通过二次布尔模型将跟踪器的估计值和检测器的检测值相关联匹配,然后优化解出轨迹,但是此模型要在一个有限的时间窗口进行优化,否则轨迹的搜索空间将会变得非常大。Jiang 等人<sup>[8]</sup>建立了一组子图,用图的形式来描述每个目标和对象,将多目标跟踪问题转化为路

收稿日期:2016-08      Received Date: 2016-08

<sup>\*</sup> 基金项目:公安网视频云计算平台应用支撑系统研制(1401b042002)、安徽省科技攻关计划资助项目

径搜索问题,通过线性规划算法来求解。为了解决目标间的遮挡问题,提高目标之间的相互区别性,Andriyenko等人<sup>[9]</sup>和Berclaz等人<sup>[10]</sup>在模型中使用0~1变量来表示局部相交区域的遮挡情况,然后用线性规划算法求得最优解,文献[11-12]中则使用表观模型来提高目标之间的相互区别性。

与当前为便于优化而简化问题的跟踪算法不同,本文旨在利用能量函数更全面真实的表达多目标跟踪中的轨迹问题,其中包括物理约束,如轨迹的平滑性,同帧中不同目标的互斥性等。为了使得不同的跟踪目标相互区别,算法建立了相似度模型,并且考虑遮挡情况结合遮挡模型提高相似度模型的适用性。鉴于能量函数的非凸性,所以不能用全局优化的方法来求解,本文使用共轭梯度算法和一系列的跳转来让能量函数摆脱弱极小而得到最优值。图1所示为本文的算法流程图,从已有的数据集中得到检测值,EKF跟踪算法得到初始跟踪状态,计算检测值和初始跟踪状态的能量函数值后利用共轭梯度算法得到极小值,然后通过使能量值减小的一系列跳转转换搜索空间从而找到能量最小值的解也就是最终的跟踪状态。

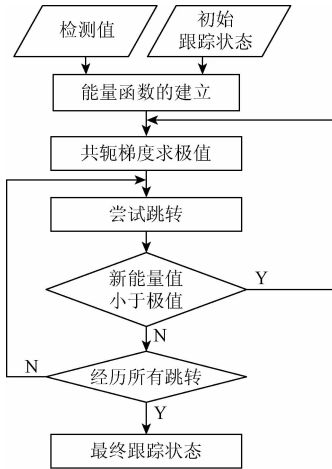


图1 算法流程

Fig. 1 Flow chart of algorithm

## 2 能量函数的建立

EM法主要是要建立目标函数<sup>[13]</sup>,使得每个可能的解对应一个跟踪状态,然后找到一个使得目标函数的代价最小也就是能量函数最小的值,则它的解即为最终的跟踪状态。本文中的能量函数为:

$$E = E_{\text{obs}} + \alpha E_{\text{cor}} + \beta E_{\text{mot}} + \chi E_{\text{exc}} + \delta E_{\text{per}} + \varepsilon E_{\text{reg}} \quad (1)$$

本文的目的则是寻找使得式(1)的值最小的解:

$$x^* = \arg \min E(x) \quad (2)$$

### 2.1 观测模型

观测模型的目的是为了让轨迹尽可能的和检测值拟合,当目标和检测值越接近  $E_{\text{obs}}$  越小:

$$E_{\text{obs}}(x) = \sum_{i=1}^n \sum_{t=s_i}^{e_i} \left[ \eta - \sum_{k=1}^{d(i)} \varpi_k^t \frac{s_k^2}{\|x_i^t - d_k^t\|^2 + s_k^2} \right] \quad (3)$$

式中:  $n$  表示跟踪的目标数目,  $s_i$  和  $e_i$  分别表示第  $i$  个目标轨迹的第一帧和最后一帧图像,  $d_k^t$  表示第  $t$  帧图像的第  $k$  个检测值的坐标,  $x_i^t$  则表示当前帧第  $i$  个目标的坐标值,  $\varpi_k^t$  表示检测值的可信度,  $s_k$  则表示检测目标的规模大小,  $\eta$  是一个常量。

### 2.2 运动模型

为了能更好的模拟平滑自然的轨迹曲线,算法使用简单的匀速运动模型表示目标的运动状态:

$$E_{\text{mot}}(x) = \sum_{i=1}^n \sum_{t=s_i}^{e_i-2} \|v_i^t - v_i^{t+1}\|^2 \quad (4)$$

式中:  $v_i^t = x_i^{t+1} - x_i^t$  表示目标  $i$  的当前速度向量。

### 2.3 互斥模型

互斥模型的目的是不允许在同一帧中的两个目标占据同一个空间位置,通过定义一个排斥项来实现这个约束:

$$E_{\text{exc}}(x) = \sum_{i=1}^f \sum_{j \neq i}^{n(t)} \frac{s}{\|x_i^t - x_j^t\|^2} \quad (5)$$

这里的  $f$  表示所有的图像数目,  $n(t)$  表示在第  $t$  帧中的目标数目,  $s$  表示目标大小。

### 2.4 轨迹持久性模型

为了防止目标在跟踪区域内(可以进入和离开跟踪区域)突然消失导致轨迹中断,  $E_{\text{per}}$  则要求目标轨迹沿着预定义的轨道长度,否则能量值变大:

$$E_{\text{per}}(x) = \sum_{i=1}^n \frac{1}{1 + \exp(-\frac{1}{s} \times b(x_i^t) + 1)} \quad t \in \{s_i, e_i\} \quad (6)$$

式中:  $b(x_i^{s_i})$  和  $b(x_i^{e_i})$  分别表示第  $i$  个目标的进入和退出跟踪区域的点的位置,  $s$  则表示目标的规模大小。

### 2.5 正则化模型

正则化模型则不允许目标数目过多,轨迹过短的情况:

$$E_{\text{reg}}(x) = n + \sum_{i=1}^n \frac{1}{f(i)} \quad (7)$$

式中:  $f(i)$  表示包含第  $i$  个目标的图像的数目。

### 2.6 相似度模型

为了提高目标与背景、目标之间的区别度,建立了相似度模型用于计算连续帧之间目标的相似度,相似度越高则能量函数值越小。目标间的相似度有多种度量方法,如颜色直方图的相似性,轮廓特征矩的相似性等方

法。由于物体运动的时空连续性,所以同一目标在连续帧之间的前景区域会有所重叠,因此本文使用重叠的前景区域面积与总面积的比值来计算两两目标间的相似性。每个目标可以表示为二元指示函数,第  $t$  帧的第  $i$  个目标表示为:

$$\Phi_i^t(\chi) = \begin{cases} 1, \chi \in B(x_i^t) \\ 0, \text{其他} \end{cases} \tag{8}$$

在包围盒里  $B(x_i^t)$  的前景区域表示为 1,其他区域则为 0。所以求解两个目标之间的重叠区域则是将两个目标的二元指示函数的积求积分,即:

$$\int \Phi_i^t(\chi) \cdot \Phi_i^{t+1}(\chi) d\chi \tag{9}$$

但是式(8)是不可微的,不能使用共轭梯度优化,于是本文将二元指示函数转化为高斯函数,则目标间相似性度量表示为:

$$CX(x_i^t) = \frac{\int N_i^t(\chi) \cdot N_i^{t+1}(\chi) d\chi}{o_i^t + o_i^{t+1} - \int N_i^t(\chi) \cdot N_i^{t+1}(\chi) d\chi} \tag{10}$$

式中:  $N_i^t(\chi)$  是  $N(\chi; u_i^t; \sum_i^t)$  的简写,  $u_i^t = x_i^t$ ,  $\sum_i^t = \begin{pmatrix} \frac{1}{2}(\frac{s_i^t}{2})^2 & 0 \\ 0 & (\frac{s_i^t}{2})^2 \end{pmatrix}$ ,而  $s_i^t$  是目标在图像区域的高度<sup>[14]</sup>。  
 $o_i^t$  为目标所占的区域大小(根据包围盒来计算区域大小),  $\int N_i^t(\chi) \cdot N_i^{t+1}(\chi) d\chi$  表示两目标的重叠面积,这样的积分可以转化成  $N(u_i^t; u_i^{t+1}; \sum_{t:t+1})$ ,而  $\sum_{t:t+1} = \sum_i^t + \sum_i^{t+1}$ 。

显然当物体发生遮挡时候,相似性度量的参考值就应该降低,所以这里结合文献[14]中的遮挡模型来提高相似度模型的适用性,文献[14]中通过计算目标被遮挡后的能见度来表示目标的遮挡状态。本文使用目标的遮挡状态的几何平均来加权到惩罚函数中,即:

$$CV(x_i^t) = \overline{v_i^t}(x) \cdot (1 - CX(x_i^t)) \tag{11}$$

式中:  $\overline{v_i^t}(x)$  表示  $\sqrt{(v_i^t(x) \cdot v_i^{t+1}(x))}$ ,  $v_i^t(x)$  表示第  $i$  个目标在  $t$  帧的能见度。然后在式(11)的基础上使用了软阈值让相似度模型的效果更好,则最终的表达式为:

$$E_{\text{cor}}(x) = \sum_{i=1}^n \sum_{t=s_i}^{e_i-1} \frac{1}{1 + \exp(\sigma - \tau \cdot CV(x_i^t))} \tag{12}$$

式中:  $\sigma$  和  $\tau$  都是常量。

3 优 化

本文的优化方法使用了共轭梯度法,使用 Polack-Ribiere 法来计算搜索方向,Wolfe-Powell 收敛准则来得到

步长。但是这样的共轭梯度优化法得到的结果很容易是弱极小,为了得到最优解,本文使用一系列的跳转来摆脱弱极小,当经过跳转能量值减少时则通过跳转到不同区域的搜索空间来寻找最小能量值<sup>[13]</sup>。这些跳转包括延长轨迹、缩短轨迹、合并轨迹、拆分轨迹、添加轨迹、移除轨迹。每个轨迹都要经历这 6 项跳转,而且所有的轨迹可以在同一时间内执行某一项跳转,当跳转到一个新的搜索空间时,继续使用共轭梯度法来寻找这一搜索空间的极小值,直到执行完全部跳转则得到最终的最优解即多目标的跟踪轨迹。

4 实验结果分析

由于本文建立的能量函数是非凸的,所以算法优化结果依赖于初始的跟踪态,但是本文中一系列跳转降低了这种依赖性能。而且在实验中划分了跟踪的区域,如图 2 和 3 的跟踪效果图中的黑色矩形区域即为标定的跟踪区域,这样能有效的计算进入点和退出点的距离,有利于执行轨迹的持久性。在式(1)中的各参数,在实验中设置为 {0.1, 0.02, 0.5, 0.5, 1},式(3)中的  $\eta$  设置为 0.125。

为了验证算法的可靠性,在 MATLAB 实验平台对两个数据集,对比了 3 种算法来进行比较实验,比较的方法使用的是 CLEAR MOT<sup>[15]</sup>度量方法,比较结果如表 1、2 所示。其中 MOTA 表示的是多目标跟踪的准确度,用于统计在跟踪中的误差累积情况,包括了误判数 FP、缺失数 FN 和误配数 ID,而 MOTP 表示的是多目标跟踪的精确度,通过计算跟踪到的目标位置和真实的目标位置的差距来衡量目标位置的精确程度。表 1 中数据集 PETS-S3-MF1 有 107 帧,包含 7 个目标;表 2 中数据集 PETS-S2L1 测试了 138 帧,包含 8 个目标。

表 1 PETES-S3-MF1 定量分析

Table 1 Quantitative analysis of PETES-S3-MF1

评估标准	MOTA/%	FP	FN	ID	MOTP/%
不包含相似度模型的 EM 法	97.1	3	12	0	84.4
KSP	83.7	22	62	0	77.8
本文算法	97.9	1	10	0	86.2

表 2 PETES-S2L1 定量分析

Table 2 Quantitative analysis of PETES-S2L1

评估标准	MOTA/%	FP	FN	ID	MOTP/%
不包含相似度模型的 EM 法	91.6	21	38	4	79.3
KSP	80.3	42	140	6	72.0
本文算法	91.8	19	37	3	80.2

算法首先在简单的测试集 PETS-S3-MF1 上进行了测试,部分跟踪效果图如图 2 所示。在这个测试集中一行人之间彼此没有遮挡,这种情况下能够充分的利用相似度模型来增加跟踪的鲁棒性能,从表 1 中可以看出本文算法较不包含相似度模型的算法虽然在 MOTA 进步不大,但是却能使跟踪器更加密切的跟踪目标从而在 MOTP 上有着明显的提升。为了检测算法的相似度模型在遮挡情况下的应用,又在较为复杂的数据集 PETS-S2L1 上进行实验,部分跟踪效果图如图 3 所示。这种情况下虽然相似度模型的参考值被约束,但是从表 2 中还是可以看出鲁棒性能的提升。而且对比 KSP 算法,本文中的算法无论在 MOTA 还是在 MOTP 指标上都有很大的提高。

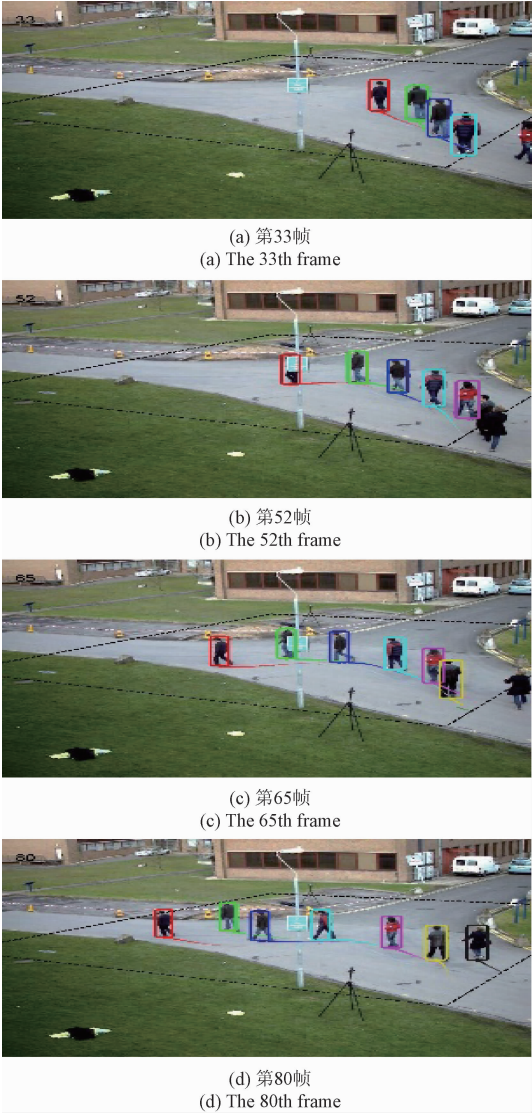


图 2 PETS-S3-MF1 跟踪效果  
Fig. 2 PETS-S3-MF1 tracking result

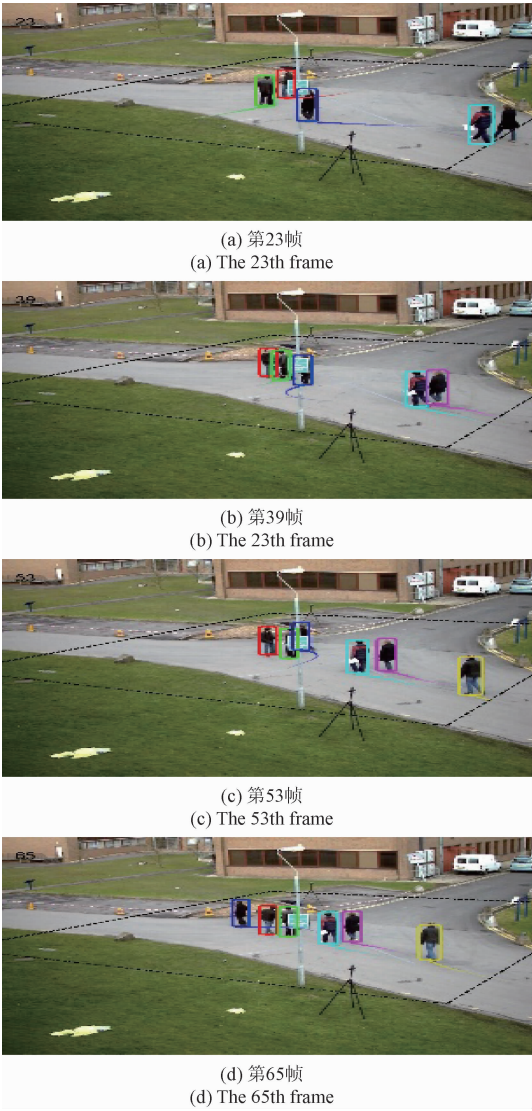


图 3 PETES-S2L1 跟踪效果  
Fig. 3 PETES-S2L1 tracking results

5 结 论

本文将多目标跟踪的复杂问题转化为能量最小化的问题,为了能更加真实的表达多个目标的轨迹情况,使用了非凸的能量函数模型,在此基础上建立的相似度模型和遮挡模型的结合能够充分将目标与背景、目标之间进行区分从而提高轨迹位置的精确度。通过共轭梯度法和一系列跳转运动而最终求得能量函数模型的最优解,得到多个目标的轨迹。从实验结果分析中能看出本算法的优越性,但是仍有需要改进的地方,本算法中的参数设置都是手动经验设置,未来的改进算法中可以根据机器学习来自动找到一套合适的参数,当然这就需要更多数据集来训练来达到效果。

## 参考文献

- [1] 李静宇,姚志军,田睿. 一致性特征点匹配在目标跟踪中的应用[J]. 电子测量技术, 2015, 38(10): 28-31.  
LI J Y, YAO ZH J, TIAN R. Target tracking based on consistency feature points matching [J]. Electronic Measurement Technology, 2015, 38(10): 28-31.
- [2] 蔡波,周大可,杨欣. 基于显著特征区域和条件随机场的目标跟踪[J]. 电子测量技术, 2016, 39(8): 64-69.  
CAI B, ZHOU D K, YANG X. Salient feature regions based conditional random field for object tracking [J]. Electronic Measurement Technology, 2016, 39(8): 64-69.
- [3] 高文,汤洋,朱明. 复杂背景下目标检测的级联分类器算法研究[J]. 物理学报, 2014, 63(9): 156-164.  
GAO W, TANG Y, ZHU M. Study on the cascade classifier in target detection under complex background[J]. Acta Physica Sinica, 2014, 63(9): 156-164.
- [4] 徐超,高梦珠,查宇锋,等. 基于 HOG 和 SVM 的公交乘客人流量统计算法[J]. 仪器仪表学报, 2015, 36(2): 446-452.  
XU CH, GAO M ZH, ZHA Y F, et al. Bus passenger traffic statistics algorithm based on HOG and SVM [J]. Chinese Journal of Scientific Instrument, 2015, 36(2): 446-452.
- [5] HENRIQUES J, CASEIRO R, BATISTA J. Globally optimal solution to multi-object tracking with merged measurements [C]. International Conference on Computer Vision, 2011.
- [6] LEIBE B, CORNELIS N, CORNELIS K. Dynamic 3d scene analysis from a moving vehicle [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2007.
- [7] LEIBE B, SCHINDLER K, CORNELIS N, et al. Coupled object detection and tracking from static cameras and moving vehicles [C]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008.
- [8] JIANG H, FELS S, LITTLE J. A linear programming approach for multiple object tracking [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2007.
- [9] ANDRIYENKO A, SCHINDLER K. Globally optimal multi-target tracking on a hexagonal lattice [C]. European Conference on Computer Vision, 2010.
- [10] BERCLAZ J, FLEURET F, FUA P. Multiple object tracking using flow linear programming [C]. 12th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS-Winter) IEEE, 2009.
- [11] SHITRIT H, BERCLAZ J, FLEURET F, and FUA P. Tracking multiple people under global appearance constraints [C]. International Conference on Computer Vision, 2011.
- [12] KUO C H, HUANG C, NEVATIA R. Multi-target tracking by online learned discriminative appearance models [C]. IEEE Conference on Computer Vision, 2010.
- [13] ANDRIYENKO A, SCHINDLER K. Multi-target tracking by continuous energy minimization [C]. IEEE Conference on Computer Vision, 2011.
- [14] ANDRIYENKO A, ROTH S, SCHINDLER K. An analytical formulation of global occlusion reasoning for multi-target tracking [C]. IEEE International Conference on Computer Vision, 2011.
- [15] BERNARDIN K, STIEFELHAGEN R. Evaluating multiple object tracking performance: The CLEAR MOT metrics [C]. Eurasip Journal on Image and Video Processing, 2008.
- [16] FERRYMAN J, SHAHROKNI A. PETS2009: Data set and challenge [C]. 12th IEEE International Workshop Performance Evaluation of Tracking and Surveillance, 2009.

## 作者简介



周娜, 1992 年出生, 现为合肥工业大学研究生, 目前主要研究方向为计算机视觉、目标跟踪。

E-mail: chouchou@mail.hfut.edu.cn

**Zhou Na** was born in 1992. And she is a M. Sc. candidate in Hefei University of Technology now. Her present research interest includes computer vision and target tracking.



鲁昌华, 1962 年出生, 现任合肥工业大学教授, 目前主要研究方向为智能信息处理等。

E-mail: jsdzlch@hfut.edu.cn

**Lu Changhua** was born in 1962. And he is a professor in Hefei University of Technology now. His present research interests include information theory, digital signal processing and so on.