

## 基于稀疏性理论与子问题耦合的多目标跟踪方法

叶 润, 邓 煜

(上海交通大学 电子信息与电气工程学院, 上海 200240)

**摘 要:** 针对多数目标跟踪方法在非受控环境中稳定性不高以及检测-跟踪模块分离的缺点, 提出一种稀疏性检测器与网络数据关联技术相结合的多目标跟踪方法。离散化目标的移动空间, 对于 3D 的每个可能位置, 将目标投影到图像平面, 形成码字并构建字典。扩展模型至多类别跟踪情况, 并根据耦合公式分配给子问题和协调局部解以实现解的最优化。使用网络单纯形算法解决最小成本流数据关联问题。在 BU-Marathon, PETS2009 等公开数据集上的实验结果表明, 与能处理遮挡的多目标跟踪方法相比, 该方法具有较高的跟踪精度, 误检率和漏检率更低。

**关键词:** 多目标跟踪; 检测-跟踪方法; 稀疏性; 耦合公式; 网络单纯形算法

**中文引用格式:** 叶 润, 邓 煜. 基于稀疏性理论与子问题耦合的多目标跟踪方法[J]. 计算机工程, 2017, 43(6): 219-224, 229.

**英文引用格式:** Ye Run, Deng Yu. Multi-target Tracking Method Based on Sparse Theory and Sub-problem Coupling[J]. Computer Engineering, 2017, 43(6): 219-224, 229.

### Multi-target Tracking Method Based on Sparse Theory and Sub-problem Coupling

YE Run, DENG Yu

(School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai 200240, China)

**[Abstract]** Aiming at the inherent shortcomings that many existing target tracking methods lack of stability in non controlled environment and the detecting-tracking module are separate, a multi-target tracking method based on sparseness detector and network data association technique is proposed. The moving space of the target is discretized, and the target is projected onto the image plane for each possible 3D position. The dictionary is built after formation of code. The model is extended to multiple classes. The optimal solution is achieved by assigning coupling equation to the sub problems and coordinating local solutions. The minimum cost flow problem of data association is solved by network simplex algorithm. BU-Marathon, PETS2009 and other public data sets are used in the experiments. Experimental results show that, compared with the multiple target tracking algorithms with occlusion handling, the proposed method has higher tracking accuracy and less false detection rate and missing detection rate.

**[Key words]** multi-target tracking; detecting-tracking method; sparseness; coupling equation; network simplex algorithm  
**DOI:** 10.3969/j.issn.1000-3428.2017.06.035

### 0 概述

对大群体中的物体运动进行参数分析是计算机视觉领域的一个重要研究课题<sup>[1]</sup>。一个完整的多目标跟踪<sup>[2]</sup>系统通常由 3 个阶段组成: 目标检测, 时态数据关联<sup>[3]</sup> (即当前目标轨迹的观测) 和每个目标的状态估计 (可选), 这些模块通常在系统中被独立优化。然而, 由于非受控环境以及检测的可扩展性差等问题, 几乎所有的跟踪系统都使用检测-跟踪和 2 个分离的模块来处理检测和数据关联任务。本文的主要研究目的是将 2 个模块合并, 尝试同时解决

2 个任务。在多目标跟踪问题中, 需要重点解决的问题是遮挡<sup>[4]</sup>。在跟踪过程中, 进行遮挡推理有 2 个主要的难题: 1) 当一个目标开始被遮挡, 并且在图像平面上侦测不到该目标, 或仅能检测到局部, 这时基于模型的探测器必须处理难以预测的不确定性; 2) 必须得到近似假设, 同时也需要解释真实检测结果。第 1) 个难题可以通过基于局部探测器来解决, 在可见部分有足够高的分辨率情况下<sup>[5]</sup>, 探测器可能会检测到局部遮挡目标, 但是当目标完全遮挡或目标的分辨率太低时, 基于局部探测器会失效, 在局部遮挡和低分辨率的情况下表现大幅下降或退化。

**作者简介:** 叶 润 (1991—), 男, 硕士研究生, 主研方向为智能算法、目标跟踪; 邓 煜, 硕士研究生。

**收稿日期:** 2016-06-06 **修回日期:** 2016-07-07 **E-mail:** shyerui@126.com

此外,总体上对遮挡直接建模很难,因为目标模型需要明确表示出局部遮挡的程度。然而,一个详细的目标模型或探测器对很多监控应用来说并非必需。有时因为有限的分辨率或困难的成像条件,它甚至是无用的。常见的思路是通过相机标定或多视图几何<sup>[6-7]</sup>,使二进制形状模板适应观察值。但这些方法依赖于背景差分法的预处理步骤。第2)个难题是可采用非极大值抑制技术,但是当目标在图像平面上有很大重叠时,滤出真实检测结果的副作用尤其不理想,参数调整也是问题。文献[8]对遮挡进行建模和优化,将遮挡模型整合到全局目标函数中。但是,随着目标函数变得越来越复杂,该函数的非凸性变强,函数优化依赖于初始化。

另外,也可以采用多个独立跟踪器,如文献[9]使用多个独立跟踪器(Multiple Independent Trackers, MIT)跟踪多个目标,方法简单高效,然而MIT方法容易受到身份切换问题的限制。文献[10]使用预先训练的检测器来扫描视频的所有帧,然后用“Tracklets”连接检测的跟踪实例,该类问题为线性分配问题,即连接一个Tracklets到另一个函数,并表示为外观和运动特征函数。

本文提出一种结合稀疏性检测器与网络数据关联技术的多目标跟踪方法,其特点是整体目标函数为一个简单的形式,且可以根据拉格朗日对偶分解法解决,通过将耦合公式分配给子问题并协调局部解来实现全局解最优化,并允许分布式计算。给出一种耦合子问题,在单一的目标函数中对多目标跟踪的检测和数据关联,使用一个简单的全局优化技术,采用稀疏理论进行目标检测,以二进制的前景/背景分割作为输出,推断相互遮挡关系。

## 1 本文方法

### 1.1 稀疏性驱动探测器

受稀疏性的人体定位法<sup>[11]</sup>启发,本文以 $L_1$ -norm最小化公式制作目标探测器。首先将目标的移动空间离散化。如果相机信息可用,那么对于3D的每个可能位置,将目标投影到图像平面。被投影的图像前景可被视作模板或“码字”。该码字可以仅是单视图场景中的一个画面,也可以是多视图场景中多个画面的拼接。每个码字的长度即观察图像的大小,而字典中词的数量由离散化程度决定。通常创建码字字典的步骤可以离线完成。

已知背景差分后的二进制前景图像 $Y$ ,本文试图找到最佳方法实例化码字,使生成的图像尽可能接近观察值 $Y$ 。在数学上,试图最小化 $L_0$ -norm,其中 $X$ 为二进制向量,指示字典中的码字选择, $N$ 为码字数量。

$$\min_X \|Y - DX\|_0, X \in \{0,1\}^N \quad (1)$$

如果 $X$ 极为稀疏,使用 $L_1$ -norm对 $X$ 还原“近似精确”。由于遮挡,本文对真实成像过程的建模实际上应为码字的线性组合。如果目标间没有相互遮挡,即每个像素属于一个唯一的目标,那么在不需要考虑量化影响的情况下,式(1)是一个精确的模型,在式(1)中,原公式可以转换为以下线性规划问题:

$$\begin{aligned} \min_{X,U} & U \\ \text{s. t.} & -DX - U + Y \leq 0 \\ & DX - U - Y \leq 0 \\ & 0 \leq X \leq 1 \end{aligned} \quad (2)$$

其中, $U$ 是一个辅助变量,式(2)中的 $X$ 是连续的。分枝界定法可应用于进一步产生精确的整数解。 $L_1$ -norm在解中引入稀疏性,对使用最小数量的假设来解释图像观察,这是一个理想的属性。本文把式(2)中的解称作线性去噪解码器(Linear Denoising Decoder, LDND)。

若要考虑目标的形状变化,则通过提供多个模板丰富字典。特定类别目标的形状模板可以从无监督聚类的训练样本中获得<sup>[12]</sup>。本文对选择变量 $X$ 进行唯一性约束,即该系统只能选择多个模板中的一个作为有效解。改良的最小化公式为:

$$\begin{aligned} \min_{x^{(1)}, x^{(2)}, \dots, x^{(k)}} & \|Y - \sum_{k=1}^k D_k X^{(k)}\|_1 \\ \text{s. t.} & \sum_{k=1}^k X^{(k)} \leq 1, X^{(k)} \in \{0,1\}^N \end{aligned} \quad (3)$$

1-bit解量化一般是不适定问题,即使在无噪声环境中,原始信号也会完全失去信号幅度。这里用遮挡层变量 $Y$ 来表示对每个像素信号解量的价值,解释在像素中被遮挡的目标数量。这个遮挡层变量同时对1-bit解量化遮挡层和稀疏信号恢复进行估算,即:

$$\begin{aligned} \min_{x, \hat{Y}} & [\|\hat{Y} - DX\| + \beta \|\hat{Y} - Y\|_1] \\ \text{s. t.} & \hat{y}_i \geq y_i, \forall i: y_i > 0 \\ & \hat{y}_j = 0, \forall j: y_j = 0 \\ & \hat{Y} \in \mathbb{Z}_+^M, X \in \{0,1\}^N \end{aligned} \quad (4)$$

其中, $Y$ 是待估算的遮挡层,其必须保持量化正确性: $Q(\hat{Y}) = Y$ 。目标函数中新的附加项为 $L_1$ 正则化,其对解量化和量化量度之间的差异进行惩罚。参数 $\beta$ 在目标函数中具有双项性, $\mathbb{Z}_+$ 是非负整数的集合。通过线性松弛,上述问题能够被转化为与式(3)相似的线性规划问题。正则化是必要的,以确保两组变量的估算不会出现不适定性,实验确定加权参数 $\beta$ 控制解量化的质量。本文把式(4)的解称作线性量化译码器(Linear De-quantization Decoder,

LDQD)。

### 1.2 数据关联

经典的网络流数据关联方法将探测器回送的每个检测表述为网络中每帧的一个节点, 将跨时间检测中每个潜在匹配表述为弧, 并伴随关联成本。本文加大了网络规模, 将场景中目标的所有可能位置设为节点。图 1 中的黑色圆圈代表每一时刻所有可能位置, 总数为 1 的流沿源  $S$  (跟踪启动) 至汇点  $T$  (跟踪终止) 的路径表示一个单一目标。3 个候选检测 (1, 2, 3), (4, 5, 6), (7, 8, 9) 分别在 3 个帧中。重复节点和容量为 1 的弧确保互不相交路径的计算。这里最多可以表达 27 个路径。图标中每个边缘都与成本关联。增加了 2 个特殊的顶点数“源”和“汇集”, 以表示跟踪的启动和终止。为确保多个轨迹不共享同一检测, 节点在每个时间步被复制, 并在节点间加入一个单一的、单位容量的零成本弧。通过执行每个边缘的流上限为 1, 确保通过图标的路径或流互斥。使得流的总成本最小, 即最小成本流问题。

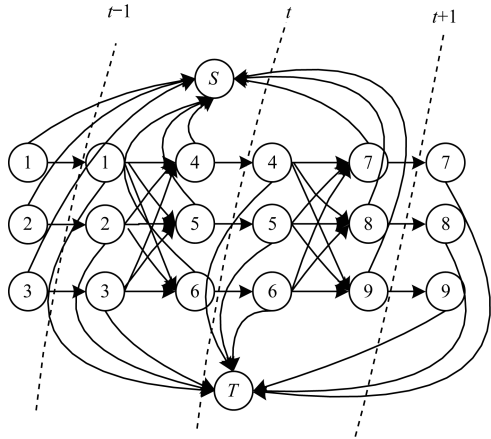


图 1 基于数据关联的网络流量问题

由于存在的目标数目未知, 因此该方法需要搜索产生的最小成本流量。然后, 可以连接汇集节点和源节点, 形成一个等价运输问题, 而不必明确规定推入流量。值得注意的是, 网络流数据关联假定单一轨迹的成本函数是可加的, 即路径中边缘成本的总和。最小化问题如下:

$$\begin{aligned} \min \sum_{i \geq 0} \sum_j c_{i,j} f_{i,j} \\ \text{s. t. } \sum_i f_{i,n} = \sum_j f_{n,j}, \forall n \end{aligned} \quad (5)$$

其中,  $c_{i,j}$  是与每个边缘关联的成本, 连接节点  $i$  和  $j$ ;  $f_{i,j}$  是与每个边缘关联的流量变量, 该网络中最佳值总是整数。在每个候选检测节点中, 传入流量与输出流量相等。如果边缘定义的所有成本有正值, 那么网络中将不会有流推入。该网络只有当强检测证据出现时才开始发挥作用。在这种情况下, 每个节点将与一个负检测分关联, 可能制造网络的完全负成本通道。

网络流模型的其他版本见跟踪类文献<sup>[13]</sup>等, 这些文献均设计为单一类别目标跟踪。本文将经典的网络流模型扩展为一个多层网络模型, 以处理多目标、多类别跟踪问题。多目标、多类别跟踪问题是最常见的跟踪问题, 但在文献中少有讨论。在许多应用中经常出现, 如(行人、汽车、自行车等)。

除了处理传统约束的多目标单类别问题, 本文还需要确保轨迹与目标类别的唯一关联性, 创建了网络中复制结构的多个图层, 将图 1 中的网络复制为多层, 每个图层描述一个特定类别下的多目标跟踪问题。不允许任何转换边缘跨越图层, 防止流向其他层。然而, 不同类别的相互排斥约束同样需要得到满足, 即不同目标类别的轨迹不能在相同时间占据相同位置。本文将这种约束表述为一组线性不等式约束:  $\sum_l f_{l,n^+,n^-} \leq 1$ , 其中,  $L$  层对应模型  $L$  类别下的目标;  $f_{l,n^+,n^-}$  表示流通过帧  $t$  中第  $l$  层的第  $n$  个节点。这些约束通常被称为“边约束”。改良后的优化问题可以放宽到一个线性优化问题框架予以解决。

## 2 耦合过程

为耦合本文的稀疏性驱动探测器和经典网络流数据关联方法, 本文提出一个共同目标函数,  $\sum_t \|\hat{Y}_t - DX_t\|_1 + \beta \|\hat{Y}_t - Y_t\|_1$ , 逼近图像似然负对数函数  $p(\|Y|X)$ , 流成本总和  $\sum_j c_{i,j} f_{i,j}$  逼近 Markov 运动先验负对数函数  $p(X)$ 。通过对式 (5) 中给出的网络流问题变量重新排列, 使用式 (4), 得到以下的耦合最小化问题:

$$\begin{aligned} \min_{x_t, f_t, \hat{Y}_t, t=1,2,\dots,T} [ & \sum_t \|\hat{Y}_t - DX_t\|_1 \\ & + \beta \|\hat{Y}_t - Y_t\|_1 + \sum_{t,n} c_{in,n^+}^t f_{in,n^+}^t \\ & + \sum_{t,n} c_{n^-,out}^t f_{n^-,out}^t + \sum_{t,n} c_{n^+,n^-}^t f_{n^+,n^-}^t \\ & + \sum_{t,m,n} c_{m^-,n^+}^t f_{m^-,n^+}^t ] \\ \text{s. t. } & f_{in,n^+}^t + \sum_m f_{m^-,n^+}^t = \sum_k f_{n^-,k^+}^t + f_{n^-,out}^t \\ & \sum_{t,n} f_{in,n^+}^t = \sum_{t,n} f_{n^-,out}^t \\ & x_{t,n} = f_{n^+,n^-}^t, \forall t, \forall n \\ & \hat{y}_{t,i} = y_{t,i}, \forall t, \forall i: y_{t,i} > 0 \\ & \hat{y}_{t,i} = 0, \forall t, \forall j: y_{t,j} = 0 \\ & f \geq 0, \hat{y}_t \in \mathbb{Z}_+^M \text{ and } X_t \in \{0, 1\}^N \end{aligned} \quad (6)$$

选择变量  $X_t$  表示在时间  $t$  时, 目标在离散空间的一个特定位置。在式 (3) 给出的模板  $K$  中,  $X_t$  可以被  $\sum_{k=1}^k X_t^{(k)}$  结合线性约束取代:  $\sum_{k=1}^k X_t^{(k)} \leq 1$ 。流量变量

$f$ 用于最小成本流问题,其中, $f_{i,j} = 1$ 表示在属于同一个轨迹的位置*i*和*j*的检测中有一个匹配,成本函数式(6)是2个局部项总和最小化,第1个包含流变量的约束集确保流量平衡。第2个包含流变量及选择变量的约束集确保2个局部变量*X*和*f*之间的一致性。即假定时刻*t*时在位置*n*处有一个真检测,即 $x_{t,n} = 1$ ,那么在同一时间必定有一个流经过同一位置,即 $f_{n^+,n^-}^{(t)} = 1$ 。

因为这是一个线性整数规划问题,可以通过应用线性规划求解器得到最优解。当需要计算数百帧或考虑目标函数的另一个高阶形式时,其限制了可扩展性和泛化性。相反,由于目标函数的特殊结构,可以将问题分解为两类子问题,每个子问题都可以通过有效算法解决。此方法可以通过将拉格朗日对偶问题公式化为最小值问题,即:

$$L(\lambda) = \min_{x,f,\hat{Y}} \left[ \sum_t \|\hat{Y}_t - DX_t\|_1 + \beta \|\hat{Y}_t - Y_t\|_1 + \lambda_t^T X_t + \sum_{t,n} c_{in,n^+}^{(t)} f_{in,n^+}^{(t)} + \sum_{t,n} c_{n^-,out}^{(t)} f_{n^-,out}^{(t)} + \sum_{t,n} (c_{n^+,n^-}^{(t)} - \lambda_{t,n}) f_{n^+,n^-}^{(t)} + \sum_{t,m,n} c_{m^-,n^+}^{(t)} f_{m^-,n^+}^{(t)} \right] \quad (7)$$

其中, $\lambda_t$ 为拉格朗日乘数向量,与式(6)中定义的耦合约束 $x_{t,n} = f_{n^+,n^-}^{(t)}$ 关联。所谓的主问题可被分为如下独立子问题(忽略约束),其中*T*表示帧数目。

$$g_t(\lambda) = \min_{x_t,\hat{Y}_t} \left[ \|\hat{Y}_t - DX_t\|_1 + \beta \|\hat{Y}_t - Y_t\|_1 + \lambda_t^T X_t \right]$$

$$h(\lambda) = \min_f \left[ \sum_{t,n} c_{in,n^+}^{(t)} f_{in,n^+}^{(t)} + \sum_{t,n} c_{n^-,out}^{(t)} f_{n^-,out}^{(t)} + \sum_{t,n} (c_{n^+,n^-}^{(t)} - \lambda_{t,n}) f_{n^+,n^-}^{(t)} + \sum_{t,m,n} c_{m^-,n^+}^{(t)} f_{m^-,n^+}^{(t)} \right] \quad (8)$$

其中,对偶问题是使用变量 $\lambda$ 最大化 $\sum_t g_t(\lambda) + h(\lambda)$ 。本文在这里使用梯度法解决主问题。拉格朗日乘数 $\lambda$ 将在每个迭代中被修正,这可以视作对不含拉格朗日项的原有子问题的扰动。因此,重优化技术应纳入考量,以便不再需要从头开始解决每个迭代的子问题。且扰动仅在目标函数中出现,约束集不变。这种情况下适用原始法,因为从上次迭代中得到的解依然可用。作为结果,原始法可以复用先前的解作为出发点,而无需寻找一个起始可行的解。特别地,本文采用原始单纯形算法以解决第1个子问题的并行计算,采用网络单纯形算法解决最小成本流子问题。耦合算法如算法1所示。

#### 算法1 跟踪耦合算法

$k = 1, 2, \dots$ , 最大迭代次数, 运行:

1) 使用原始单纯形算法解决稀疏性检测问题:

$X_t \leftarrow \text{argmin}_t(\lambda)$

2) 使用网络单纯形算法解决最小成本流数据关联问题:

$f \leftarrow \text{argmin}_h(\lambda)$

3) 如果 $x_{t,n} = f_{n^+,n^-}^{(t)}$ 为所有*n*和*t*,则返回 $X_t, f$ 。

4) 修正对偶变量 $\lambda_{t,n} = \lambda_{t,n} + \alpha_k (x_{t,n} - f_{n^+,n^-}^{(t)})$ ,  $\alpha_k = \frac{1}{k}$ (步长)。

5) 返回 $X_t, f$ 。

对于检测子问题,较高阶的 $\lambda$ 表示在特定位置的较低。对于数据关联子问题,较高阶的 $\lambda$ 导致较低边缘成本,因此其吸引流通过边缘。当协议达成,得到原始目标函数的全局最优解。数据关联的影响保证了检测输出的平稳性。通过对偶分解动态改变 $\lambda$ 值,可以抑制误警,恢复因遮挡产生的漏报。

## 3 实验结果与分析

### 3.1 细节描述

为了获得二进制图像,本文使用自适应高斯混合方法进行背景减除<sup>[14]</sup>。地平面采用0.3 m × 0.3 m大小的网格离散,该网格与一个行人可占据的空间一半大小相当。为了提高计算速度,将二进制图像的像素分辨率调整为320像素 × 240像素。使用2个形状模板,如图2所示,该模板通过K-均值聚类得到。

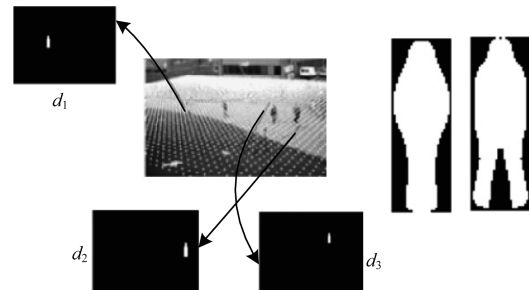


图2 形状模板

在运行线性规划求解器前,减少字典的大小。如果码字没有从图像中得到足够支持,即网格中50%的前景像素没有被删除,那么字典中相应的列将被移除。码字的原始长度为320像素 × 240像素。然而,图像中很大一部分像素将不会被字典中的任何码字覆盖,因为该像素在监控区域外或本来可能覆盖该像素的码字在第1个步骤中被移除了。这些无法覆盖的像素对应地从字典行中移除。为此,数据集构建的字典最后大小约为500码字,每个码字表示约20 000像素。正则化参数 $\beta$ 被选择为0.1。

为了建立在数据关联中使用的网络,需要定义边缘成本。有4种类型的边缘:时间帧中重复节点间的边缘;不同时间节点间的边缘;从源节点开始的边缘以及从汇集节点结束的边缘。本文将这些边缘的成本称作检测成本、转换成本、起始成本、终止成本。检测成本被计算为 $-\ln \frac{\rho}{1-\rho}$ ,这里 $\rho$ 是可被码

字解释的前景像素数量和码字中的前景像素数量的比率。对于红外数据集, 本文使用节点间的欧式距离作为转换成本。对于可见光相机捕获的数据集, 本文首先计算每个码字颜色直方图, 已知连续帧中 2 个实例化的码字, 通过计算直方图的交点距离表达为转换成本。

本文实验配置为英特尔 i5 双核 CPU, 2.29 GHz 主频, 4.0 GB 内存, 在 Windows 8 操作系统的笔记本上实现, 编程环境为 Matlab2011b。

### 3.2 重要参数设置与评估标准

有极少的用户定义参数需要通过实验来确定。式(4)中管理解量化质量的加权参数  $\beta$  对所有包含刚性目标的序列设为 0.1, 对红外数据集设为 10。通常对支持多视图数据集的解量化效果进行估算, 为确保  $L_1$ -norm 项与式(6)中的网络流项数值平衡, 本文通过加权参数  $\gamma$  重调  $L_1$ -norm, 对所有数据集设为 0.01。此类加权参数一般不敏感, 在本文的实验中, 并没有频繁地对其调整。一旦发现此类加权参数的合理值, 附近各种不同的值也同样适用。

本文使用标准 CLEAR MOT 度量<sup>[15]</sup>进行评估。

多目标跟踪准确性 (MOTA) 度量将误警率、漏报率和标识切换率结合为一个理想值为 100% 的单一数值。多目标跟踪精度 (MOTP) 测量地面实况和定位器输出之间的平均距离。为评估 2D 跟踪性能, 精度由区域重叠标准定义, 理想值为 1。3D 的精度由 3D 空间中基于相机标定信息的欧式距离定义, 理想值为 0。另外, 将主要跟踪 (MT) 大于 80% 的轨迹数量、主要丢失 (ML) 小于 20% 的轨迹、轨道碎片 (FM) 和标识切换 (IDS) 作为评估标准, 其中, FM 为轨迹被干扰的时间量; IDS 为跟踪轨迹改变与之匹配的时间量。上述度量定义具体见文献[15]。

### 3.3 性能比较

本文的视频序列包括不同的目标类别、密度、动作方式、成像模式和相机视图 (如表 1 所示): 1) BU<sup>[16]</sup> 的 3 个视频中飞行蝙蝠的 2 组视频序列, 选择 0.3 m 为 3D 空间中击中-丢失阈值, 其近似目标物理尺寸。2) 红外多类别多目标跟踪序列 BU-Marathon 捕捉行人、自行车、车辆和摩托车的街景, 选择 0.5 作为区域重叠阈值。3) PETS2009 中的多个序列使用文献[8]给出的注释。

表 1 数据集说明

数据	相机	帧数	类别	目标	密度	视图	跟踪空间	击中或丢失/m
BU Bat-1	红外	1 200	蝙蝠	205	稀疏	3	3D 空间	$\leq 0.3$
BU Bat-2	红外	300	蝙蝠	201	密集	3	3D 空间	$\leq 0.3$
BU Marathon	红外	2 500	跑步者, 车, 摩托, 自行车	178	中等	1	图像平面	$\leq 0.5$
PETS S2L1	可见光	795	行人	22	稀疏	1	地平面	$\leq 1$
PETS S1L1-1357	可见光	221	行人	44	密集	1	地平面	$\leq 1$
PETS S1L1-1359	可见光	241	行人	31	中等	1	地平面	$\leq 1$
PETS S1L2-1406	可见光	201	行人	42	密集	1	地平面	$\leq 1$

各方法的量化结果如表 2 ~ 表 4 所示, 在有明显区别的数据集实验中, 基于 MOTA, MT 和 ML 的量化结果证明了本文方法比其他方法更加可靠, 基于

MOTP 的分数则证明了本文方法的准确性。本文方法在包含密集人群目标且目标局部被频繁遮挡的序列中有极佳表现, 如表 4 的 PETS 数据集结果。

表 2 BU Bat 数据集各算法的量化结果

方法	MT/%	ML/%	FM	IDS	MOTA/%	MOTP/%
文献[8]方法	89.1	0	104	97	84.1	87.4
文献[9]方法	85.2	0	142	128	78.1	81.3
文献[16]方法	91.1	0	117	128	88.7	86.7
LDQD 方法	84.2	3.5	178	156	80.9	80.1
本文方法	<b>94.9</b>	<b>0</b>	<b>94</b>	<b>91</b>	<b>92.1</b>	<b>94.5</b>

表 3 BU Marathon 数据集各算法的量化结果

方法	MT/%	ML/%	FM	IDS	MOTA/%	MOTP/%
文献[8]方法	96.6	0	121	90	91.1	92.7
文献[9]方法	95.7	0	162	117	92.9	90.3
文献[16]方法	97.0	0	174	170	95.3	92.1
LDQD 方法	81.2	1.5	185	149	89.5	87.4
本文方法	<b>98.9</b>	<b>0</b>	<b>109</b>	<b>84</b>	<b>96.1</b>	<b>97.1</b>

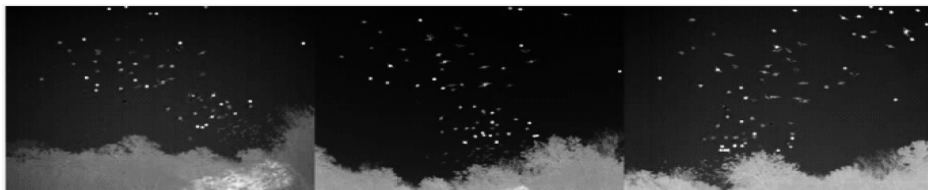
表4 PETS数据集各算法的量化结果

方法	MT/%	ML/%	FM	IDS	MOTA/%	MOTP/%
文献[8]方法	<b>91.5</b>	<b>0</b>	125	87	90.5	90.1
文献[9]方法	90.4	<b>0</b>	171	106	91.7	93.2
文献[16]方法	92.5	<b>0</b>	183	129	94.1	95.2
LDQD方法	86.7	0.5	180	147	89.4	90.0
本文方法	90.9	<b>0</b>	<b>115</b>	<b>85</b>	<b>95.7</b>	<b>96.5</b>

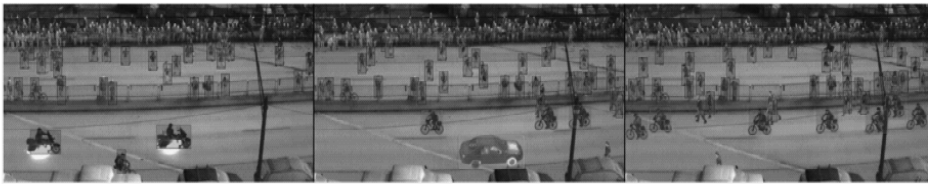
表2的蝙蝠红外数据结果突出表明本文检测跟踪方法能够成功抑制3D空间中的重影重建,其他方法由于重建步骤中持续的重影效应,其误警率高(数字加粗的为最优结果)。一旦将大群体目标组纳入考量,对消除重影的需求开始出现,这证明了本文耦合理念的重要影响。

文献[8]对遮挡进行建模和优化,从表2~表4可以看出,其优于其他方法(除了本文方法),但该方法

的函数优化依赖于初始化。文献[9]使用多个独立跟踪器,但容易受到身份切换问题的限制。文献[16]强烈依靠行人探测器,适用于由可见光相机记录的相对较高分辨率的图像中,表现出较大的波动性。而本文方法比较稳定,漏检和误检更少。本文方法的跟踪结果部分视频帧如图3所示。从图3中可以看出,即使有很多遮挡,跟踪结果依然比较清晰,拥有比较好的稳定性。



(a) BU Bat数据集跟踪的样本帧



(b) BU Marathon数据集跟踪的样本帧



(c) PETS数据集跟踪的样本帧

图3 本文算法的跟踪结果

## 4 结束语

本文提出一种多目标跟踪方法,该方法耦合目标检测和数据关联,并结合了稀疏性驱动探测方法与网络流数据关联方法。其中,稀疏性驱动探测器可以为整个图像似然建模,可避开一般的启发性算法,如非极大值抑制等。通过对耦合目标函数逐步优化,使每个子问题都可采用现有的有效算法。实验结果表明,该方法能明显提高跟踪性能。本文方法的复杂性很大程度上取决于字典的大小,这与形状模板数量和图像大小有关,下一步研究是将更多的先验知识纳入本文框架,如社会群体模型,从而更有效地提高本文算法的稳定性。

## 参考文献

- [1] 李介谷. 计算机视觉的理论和实践[M]. 上海: 上海交通大学出版社, 1991.
- [2] 李方敏, 姜娜, 熊迹, 等. 融合热释电红外传感器与视频监控器的多目标跟踪算法[J]. 电子学报, 2014, 42(4): 672-678.
- [3] 冯洋. 多目标跟踪的数据关联算法研究[D]. 西安: 西安电子科技大学, 2008.
- [4] 张彦超, 许宏丽. 遮挡目标的分片跟踪处理[J]. 中国图象图形学报, 2014, 19(1): 92-100.
- [5] Girshick R B, Felzenszwalb P F, McAllester D. Object Detection with Grammar Models[C]//Proceedings of IEEE NIPS' 11. Washington D. C., USA: IEEE Press, 2011: 442-450.

(下转第229页)

- [ 4 ] Hemami S, Estrada F, Susstrunk S. Frequency-tuned Salient Region Detection [ C ] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA; IEEE Computer Society, 2009; 1597-1604.
- [ 5 ] Cheng Mingming, Mitra N J, Huang Xiaolei, et al. Global Contrast Based Salient Region Detection [ J ]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(3): 409-416.
- [ 6 ] Li Yin, Hou Xiaodi, Koch C, et al. The Secrets of Salient Object Segmentation [ C ] // Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA; IEEE Computer Society, 2014; 280-287.
- [ 7 ] Herreras E B. Cognitive Neuroscience, The Biology of the Mind [ J ]. Cuadernos De Neuropsicología, 2010, 4(1): 87-90.
- [ 8 ] 刘建伟, 刘媛, 罗雄麟. 深度学习研究进展 [ J ]. 计算机应用研究, 2014, 31(7): 1921-1930.
- [ 9 ] Le Cun Y, Bengio Y. Convolutional Networks for Images, Speech, and Time Series [ J ]. The Handbook of Brain Theory and Neural Networks, 1995, 361(10): 1-14.
- [ 10 ] Harel J, Koch C, Perona P. Graph-based Visual Saliency [ C ] // Proceedings of NIPS '06. Vancouver, Canada; [ s. n. ], 2006; 545-552.
- [ 11 ] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [ C ] // Proceedings of CVPR '14. Columbus, USA; IEEE Computer Society, 2014; 580-587.
- [ 12 ] Deng Jia, Dong Wei, Socher R, et al. Imagenet: A Large-scale Hierarchical Image Database [ C ] // Proceedings of CVPR '09. Miami, USA; IEEE Computer Society, 2009; 248-255.
- [ 13 ] Uijlings J R R, van de Sande K E A, Gevers T, et al. Selective Search for Object Recognition [ J ]. International Journal of Computer Vision, 2013, 104(2): 154-171.
- [ 14 ] Kanizsa G, Gerbino W. Convexity and Symmetry in Figure-ground Organization [ J ]. Vision and Artifact, 1976, 1(1): 25-32.
- [ 15 ] 陆遥. 感知组织在图像显著性及视觉注意力预测上的应用 [ D ]. 上海: 复旦大学, 2013.
- [ 16 ] Felzenszwalb P F, Huttenlocher D P. Efficient Graph-based Image Segmentation [ J ]. International Journal of Computer Vision, 2004, 59(2): 167-181.
- [ 17 ] Shi Jianping, Yan Qiong, Xu Li, et al. Hierarchical Image Saliency Detection on Extended CSSD [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(4): 717-729.

编辑 索书志

(上接第 224 页)

- [ 6 ] Thangali A. Coupling Detection and Data Association for Multiple Object Tracking [ C ] // Proceedings of IEEE Conference on Computer Vision & Pattern Recognition. Washington D. C., USA; IEEE Press, 2012; 1948-1955.
- [ 7 ] 夏侯凯顺, 陈善星, 邬依林. 基于双目云台相机的目标跟踪系统建模与仿真 [ J ]. 系统仿真学报, 2015, 27(2): 362-368.
- [ 8 ] Andriyenko A, Roth S, Schindler K. An Analytical Formulation of Global Occlusion Reasoning for Multi-target Tracking [ C ] // Proceedings of IEEE International Conference on Computer Vision Workshops. Barcelona, Spain; IEEE Press, 2011; 1839-1846.
- [ 9 ] Kyriazis N, Argyros A. Scalable 3D Tracking of Multiple Interacting Objects [ C ] // Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. [ S. l. ]; IEEE Computer Society, 2014; 3430-3437.
- [ 10 ] Pirsiavash H, Ramanan D, Fowlkes C C. Globally-optimal Greedy Algorithms for Tracking a Variable Number of Objects [ C ] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. [ S. l. ]; IEEE Computer Society, 2011; 1201-1208.
- [ 11 ] Alahi A, Jacques L, Boursier Y, et al. Sparsity Driven People Localization with a Heterogeneous Network of Cameras [ J ]. Journal of Mathematical Imaging & Vision, 2011, 41(12): 39-58.
- [ 12 ] 沈红斌, 杨杰, 王士同, 等. 采样定理、视觉原理及无监督聚类分析理论 [ J ]. 上海交通大学学报, 2005, 39(4): 544-548.
- [ 13 ] Henriques J F, Caseiro R, Batista J. Globally Optimal Solution to Multi-object Tracking with Merged Measurements [ C ] // Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA; IEEE Press, 2011; 2470-2477.
- [ 14 ] 邵奇可, 周宇, 李路, 等. 复杂场景下自适应背景减除算法 [ J ]. 中国图象图形学报, 2015, 20(6): 756-763.
- [ 15 ] Bernardin K, Stiefelhagen R. Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics [ J ]. Eurasip Journal on Image & Video Processing, 2008, 48(1): 1-10.
- [ 16 ] Wu Z, Fuller N, Theriault D, et al. A Thermal Infrared Video Benchmark for Visual Analysis [ C ] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA; IEEE Press, 2014; 201-208.

编辑 索书志