

视觉传感网络分布式在线数据关联

刘 莉¹ 万九卿¹

摘 要 数据关联是视觉传感网络监控系统的基本问题之一. 本文针对无重叠视域视觉监控网络的多目标跟踪问题提出一种基于多外观模型的视觉传感网络在线分布式数据关联方法, 将同一目标在不同摄像机节点上的外观用不同的高斯模型描述, 由分布式推理算法综合利用外观与时空观测计算关联变量的后验概率, 同时通过近似最大似然估计算法对各传感节点上的外观模型参数进行在线估计. 实验结果表明了所提方法的有效性.

关键词 视觉传感网络, 数据关联, 分布式在线推理, 极大似然估计, 多模型

引用格式 刘莉, 万九卿. 视觉传感网络分布式在线数据关联. 自动化学报, 2014, 40(1): 117–125

DOI 10.3724/SP.J.1004.2014.00117

Distributed Online Data Association in Visual Sensor Networks

LIU Li¹ WAN Jiu-Qing¹

Abstract One of the fundamental requirements for visual surveillance with smart camera networks is the correct association of camera's observations. In this paper, we present a distributed online approach based on multiple appearance models for multi-object tracking with distributed non-overlapping cameras. Firstly, we use multiple Gaussian models to describe each object's appearances under different camera nodes. Secondly, we develop a novel distributed online framework, in which the posterior margins of association variables are calculated using appearance and spatio-temporal information by a distributed inference algorithm, and the model parameters are updated online on each camera by approximate maximum likelihood estimation. Experimental results show the validity of the proposed method.

Key words Smart camera networks, data association, online distributed inference, maximum likelihood estimation, multiple models

Citation Liu Li, Wan Jiu-Qing. Distributed online data association in visual sensor networks. *Acta Automatica Sinica*, 2014, 40(1): 117–125

随着网络技术和数字视频技术的飞速发展, 视觉监控技术越来越多地应用在民用跟踪、交通流量监测等领域中. 视觉传感网络监控系统由多个摄像机组成, 每个摄像机可对其视场范围内的局部区域进行观测. 在大规模视觉监控应用中, 不同摄像机的视场往往是非重叠的. 很多复杂的视觉监控任务, 比如在一定范围内对感兴趣的人或车辆的联合跟踪、智能交通中的违规检测、行人行为的分析任务等, 其核心问题均为对不同视觉传感器的观测数据进行数据关联, 即如何确定每个观测所对应的目标标号. 由于视频监控系统的成像质量往往较低, 且同一目标在不同观测条件下外观会发生很大变化, 使得单纯

根据外观特征进行数据关联无法满足使用要求^[1–3]. 因此需要结合监控区域的拓扑结构和视觉传感节点的时空观测信息进行综合推理以提高关联准确率.

近年来, 有关视觉传感网络的数据关联问题引起了国内外学者的高度关注, 研究人员在此方面开展了大量的研究工作. 目前有关视觉传感网络的数据关联方法可大致分为两类: 基于优化的方法和基于推理的方法. 通过建立适当的模型, 数据关联问题可转化为最优化问题. 文献 [4] 提出了一种多目标优化框架, 将数据关联问题建模为最优路径的搜索问题, 通过动态规划算法对上述问题求解. 文献 [5] 提出一种基于最小费用流模型的数据关联算法, 引入关联数的限制, 通过最小费用流优化算法对问题求解. 文献 [6] 考虑观测数据的分划集合, 每个分划代表一个目标的运动轨迹, 根据目标观测数据的生成模型计算各个分划的后验概率, 关联结果由具有最大后验概率的分划确定. 另一种研究思路是引入表征目标标号的随机变量, 即关联变量, 建立观测数据和关联变量之间的概率模型, 将数据关联转化为关联变量边缘分布的推理问题. 文献 [7] 利用动态贝叶斯网络, 描述观测数据和关联变量之间的依赖关系,

收稿日期 2012-08-15 录用日期 2013-01-11

Manuscript received August 15, 2012; accepted January 11, 2013

国家自然科学基金 (61174020), 北京市自然科学基金 (4113072) 资助
Supported by National Natural Science Foundation of China (61174020), Natural Science Foundation of Beijing (4113072)

本文责任编辑 赵千川

Recommended by Associate Editor ZHAO Qian-Chuan

1. 北京航空航天大学自动化科学与电气工程学院 北京 100191

1. School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191

用假设密度滤波器 (Assumed density filter, ADF) 计算关联变量的边缘分布. 文献 [8] 基于条件独立性假设对信念状态进行分解, 进一步提高了假设密度滤波器的估计精度.

然而, 上述数据关联算法均是集中式的, 即需要将网络中所有观测数据汇集到中心处理节点上进行处理. 这种集中式结构的计算和通信复杂度随着网络规模迅速膨胀, 难以应用于大规模视觉传感网络. 同时, 中心处理节点的故障会使整个系统陷入瘫痪, 系统鲁棒性较低. 目前, 智能摄像机^[9] 的迅速发展使得每个视觉传感节点不仅可以采集视频数据, 同时具有信息处理和通信功能, 这为构建分布式视觉监控系统奠定了基础. 因此, 近年来分布式数据关联方法的研究得到了高度关注. 文献 [10] 将分布式最大期望 (Expectation maximization, EM) 算法应用于视觉传感网络数据关联中, 实现了基于纯外观观测的分布式关联推理以及目标外观模型的估计. 事实上, 目前受到广泛关注的分布式 EM 算法^[11-15] 均可以看作是基于纯外观的分布式数据关联算法. 在此基础上, 文献 [16] 引入时空观测进行分布式关联推理, 进一步提高了关联准确率.

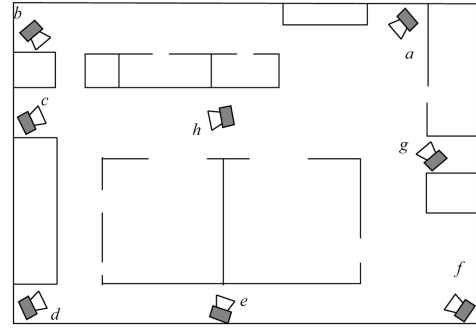
在分布式数据关联的研究中, 仍存在如下两个问题需要解决: 1) 以往的研究中通常假设同一目标在各传感节点的外观观测服从单一的高斯分布, 但由于视觉传感网络各节点的观测条件往往存在较大差异, 同一目标在不同节点处的外观观测难以用单峰高斯分布来描述. 因此, 在这种情况下, 传统方法难以取得理想的效果; 2) 前述算法均属于离线推理方法, 无法应用于在线跟踪等应用场合. 值得注意的是, 文献 [17] 提出一种在线分布式 EM 算法, 然而该算法只是利用外观信息进行关联推理, 同时要基于单外观模型的假设.

针对上述问题, 本文提出一种基于多外观模型的视觉传感网络分布式在线数据关联方法, 其主要特点在于: 1) 同一目标在不同传感节点处的外观观测由不同的高斯模型描述, 这种建模上的灵活性使得本文方法在各节点观测条件变化较大时表现出明显优势; 2) 建立了一种分布式关联推理与模型估计算法框架, 其中每个摄像机节点同时利用外观和时空观测以并行的方式进行关联变量后验概率的计算以及各目标外观模型的更新; 3) 所提方法中关联推理与模型估计均由网络节点上的目标检测事件驱动, 根据以往的观测数据进行计算, 不需要离线的训练过程, 可用于视觉传感网络在线多目标联合跟踪等应用场合.

1 问题描述

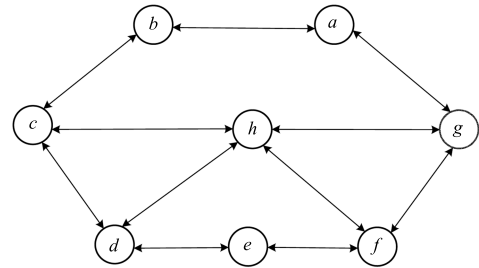
假设有 H 个目标在具有 M 个摄像头的监控网

络中运动, 如图 1(a) 所示. 图中每个节点代表一个摄像头, 假定网络节点时间已经同步且各节点具有一定的储存、通信和计算能力. 相应的网络拓扑结构如图 1(b) 所示, 其中每个节点对应于一个摄像机, 节点间的连线代表观测节点间存在目标运动的连通路程. 我们用 $\mathcal{N}_u$ 来表示摄像头节点 u 的邻居节点集合, 且 $\bar{\mathcal{N}}_u = u \cup \mathcal{N}_u$.



(a) 监控区域不重叠的智能摄像机网络

(a) Non-overlapping smart cameras networks



(b) 网络拓扑结构

(b) Topology of visual sensor networks

图 1 网络建模

Fig. 1 Networks modeling

当运动目标经过某个摄像机节点 u 时, 运行在该节点上的目标检测与特征提取算法根据采集到的视频数据产生一个观测 $y_{u,i}^k$, 索引 i 代表该观测是产生在节点 u 上的第 i 个观测, 索引 k 代表该观测是整个网络中的第 k 个观测. 我们使用 $Y^{1:k}$ 表示到时间 k 为止产生的所有观测集合. 对于每个观测 $y_{u,i}^k$, 我们都分配一个关联变量 $x_{u,i}^k$, 用于表征 $y_{u,i}^k$ 由哪一个目标产生, $x_{u,i}^k$ 的取值范围为 $\{1, \dots, H\}$. 本文的基本目标是要通过计算每个观测对应的关联变量的后验概率 $p(x_{u,i}^k | Y^{1:k})$, 选取概率最大的关联变量取值与观测关联.

每个观测 $y_{u,i}^k$ 由外观观测和时空观测两部分组成. 我们用 $o_{u,i}^k$ 表示目标的外观信息, 比如目标的颜色或形状特征等; $s_{u,i}^k$ 表示观测的时空信息, 包括观测发生的时间、位置以及目标运动方向等信息. 在给定关联变量的条件下, 我们假设外观观测和时空观

测相互独立. 下面将分别对外观观测模型与时空观测模型进行阐述.

1.1 外观观测模型

为了提取目标的外观观测, 我们采用文献 [6] 中的方法, 将图像中人体区域分为上、中、下三个部分, 分别计算每个区域 RGB 分量的均值, 由此组成 9 维向量 $o_{u,i}^k$ 作为外观观测. 由于不同摄像机节点观测条件可能存在很大差异, 我们用不同的高斯模型描述同一目标在不同节点上的外观分布. 目标 h 在节点 u 的外观观测模型为

$$o_{u,i}^k | x_{u,i}^k = h \sim N(o_{u,i}^k; \mu_u^h, \sigma_u^h) \quad (1)$$

其中, μ_u^h 和 σ_u^h 分别代表目标 h 在节点 u 下的外观均值和方差. 上述模型参数需要在数据关联的同时进行在线估计.

1.2 时空观测模型

时空观测包括两部分: 目标被摄像机捕获的时间和目标在摄像机视域内的运动方向. 我们取目标自进入至离开摄像机视域的中间时刻作为捕获时间, 用 $d_{u,i}$ 表示第 i 个观测在摄像机 u 上的捕获时间. 假设产生当前观测 $d_{u,i}$ 的目标的上一次观测 $d_{v,j}$ 来自于摄像机 v , 则 $d_{u,i}$ 和 $d_{v,j}$ 之间存在如下约束关系:

$$d_{u,i} | d_{v,j} \sim N(d_{u,i} - d_{v,j}; \mu_{uv}, \sigma_{uv}^2) \quad (2)$$

其中, μ_{uv} , σ_{uv}^2 分别代表两节点运动时间的均值和方差, 假设二者均与两节点之间连同路径的长度成正比. 当网络拓扑结构确定时, 上述参数可作为已知的先验信息. 当目标进入摄像头视域范围时, 摄像头采集到一段视频, 目标跟踪算法可以从视频中提取目标在此节点下的移动轨迹. 根据目标的移动轨迹, 可以得到目标从视频图像的哪一边界进入 (或离开) 摄像头视域范围, 将此边界信息作为目标进入 (或离开) 摄像机的运动方向信息 $in_{u,i}$ (或 $out_{u,i}$). 图 2 截取了一段视频的的开始和结束画面, 由图 2 可以看出, 目标自视频图像的左侧进入摄像头视域范围后从右侧离开, 则我们采集方向信息 $in_{u,i} = \text{“左”}$, $out_{u,i} = \text{“右”}$. 由于目标离开某一节点后必将进入其相邻节点的视域范围, 因此考虑相邻节点 u 和 v 上的两观测 $y_{u,i}^k$ 和 $y_{v,j}^{k'}$ ($k > k'$), 结合网络节点的相对方位我们可得到离散概率分布:

$$in_{u,i} | out_{v,j} \sim p(in_{u,i} | out_{v,j}) \quad (3)$$

方向变量 $in_{u,i}$ ($out_{u,i}$) 的取值空间为摄像机视域的各个边界, 因此上述分布为离散概率分布.



图 2 目标运动方向

Fig. 2 The moving directions of one person

2 分布式在线数据关联与模型估计

如前所述, 视觉传感网络数据关联需要同时进行关联变量后验分布的计算和目标外观模型的更新, 本节中我们将给出一种分布式在线数据关联与模型估计算法. 该算法首先对每个摄像机 u 上的模型参数 $\Theta_{u,0} = \{\pi_{u,0}^h, \mu_{u,0}^h, \sigma_{u,0}^h\}$ 赋初值, 其中 $\pi_{u,0}^h$ 是目标 h 出现在节点 u 的先验概率, $\mu_{u,0}^h$ 和 $\sigma_{u,0}^h$ 是目标 h 在节点 u 的外观观测的均值和方差. 当摄像机 u 产生第 k 个观测 $y_{u,i}^k$ 后, 首先利用第 2.1 节中的分布式数据关联算法计算相应的关联变量 $x_{u,i}^k$ 的后验分布, 接着利用第 2.2 节中的模型更新算法对该摄像机上的模型参数 $\Theta_{u,i}$ 进行在线更新. 在实际应用中, 算法 1 将部署到每个智能摄像机节点上独立并行运行, 每个摄像机上的计算过程由目标检测事件驱动. 后面我们将看到, 每个摄像机上的关联推理与模型更新过程仅包括进行摄像机节点上的局部信息处理以及相邻节点之间的信息交互, 全部计算过程只需利用以往的观测信息, 因而该算法是一种在线分布式算法.

算法 1. 分布式在线数据关联与模型估计算法

- 1: 初始化
- 2: For 摄像机 $u = 1$ to M
- 3: 初始化每个摄像机上的模型参数 $\Theta_{u,0}$;
- 4: End for
- 5: 迭代过程:
- 6: For 迭代步骤 $k = 1$ to ∞
- 7: For 摄像机 $u = 1$ to M 并行计算
- 8: 等待观测: $y_{u,i}^k$;
- 9: 关联推理: 根据当前模型参数 $\Theta_{u,i-1}$
 计算后验概率 $p(x_{u,i}^k | Y^{1:k})$;
- 10: 模型更新: 根据 $p(x_{u,i}^k | Y^{1:k})$
 更新模型参数, 计算 $\Theta_{u,i}$;
- 11: End 并行计算
- 12: End for

2.1 分布式关联推理

数据关联推理的目标是要计算当前观测 $y_{u,i}^k$ 对应的关联变量 $x_{u,i}^k$ 的后验分布 $p(x_{u,i}^k | Y^{1:k})$, 值得注意的是, $Y^{1:k}$ 表示到目前为止网络产生的全部观测集合, 因此在传统的数据关联算法中, 需要将整个网络中产生的观测汇聚到统一的中心处理节点上处理. 实现分布式数据关联的关键在于如下的基本假设: 当摄像机 u 捕获到一个观测时, 如果该观测的目标不是在网络中首次出现的话, 那么该目标在到达 u 之前必然经过了 u 的相邻节点之一 $v \in \mathcal{N}_u$. 我们用 $Y_{\mathcal{N}_u}^{1:k-1}$ 表示观测 $y_{u,i}^k$ 产生之前, 在节点集合 $\mathcal{N}_u$ (节点 u 及其相邻节点 $\mathcal{N}_u$) 上产生的所有观测, 相应的关联变量集合记为 $X_{\mathcal{N}_u}^{1:k-1}$.

根据上述基本假设可知, 在给定 $X_{\mathcal{N}_u}^{1:k-1}$ 和 $Y_{\mathcal{N}_u}^{1:k-1}$ 的条件下, 当前的关联变量 $x_{u,i}^k$ 独立于网络中除 $\mathcal{N}_u$ 之外的摄像机节点上产生的观测, 即

$$p(x_{u,i}^k | X_{\mathcal{N}_u}^{1:k-1}, Y^{1:k-1}) = p(x_{u,i}^k | X_{\mathcal{N}_u}^{1:k-1}, Y_{\mathcal{N}_u}^{1:k-1}) \quad (4)$$

为了描述时空约束, 对每个 $x_{u,i}^k$, 引入一组辅助变量 $z_{u,i}^{1:H} \in \{0, 1, \dots, L\}^H$, 其中 L 为集合 $Y_{\mathcal{N}_u}^{1:k-1}$ 中元素的个数. $z_{u,i}^{(h)} = l$ 表示当前观测 $y_{u,i}^k$ 的上一观测是观测集合 $Y_{\mathcal{N}_u}^{1:k-1}$ 中的第 l 个元素, $z_{u,i}^{(h)} = 0$ 意味着 $y_{u,i}^k$ 是目标 h 产生的第一个观测, H 表示目标的总个数. 由贝叶斯公式, $(x_{u,i}^k, z_{u,i}^{1:H})$ 的联合信念状态为

$$\begin{aligned} b_{u,i}^k(x_{u,i}^k, z_{u,i}^{1:H}) &= p(x_{u,i}^k, z_{u,i}^{1:H} | Y^{1:k}) \propto \\ & p(y_{u,i}^k | x_{u,i}^k, z_{u,i}^{1:H}, Y^{1:k-1}) p(x_{u,i}^k, z_{u,i}^{1:H} | Y^{1:k-1}) = \\ & p(o_{u,i}^k | x_{u,i}^k = h) p(d_{u,i}^k | z_{u,i}^{(h)} = l, Y^{1:k-1}) \times \\ & p(x_{u,i}^k | Y^{1:k-1}) p(z_{u,i}^{1:H} | Y^{1:k-1}) \approx \\ & \pi_{u,i-1}^h \lambda(x_{u,i}^k = h) \eta(z_{u,i}^{(h)} = l) \prod_h p(z_{u,i}^{(h)} | Y^{1:k-1}) \end{aligned} \quad (5)$$

关联变量的信念状态 $b_{u,i}^k(x_{u,i}^k) = p(x_{u,i}^k | Y^{1:k})$ 可由上式对 $z_{u,i}^{1:H}$ 积分而得. 式 (5) 中 $\pi_{u,i-1}^h$ 是目标 h 出现在节点 u 的先验概率, $\lambda(x_{u,i}^k = h)$ 是外观似然函数, 可由外观观测模型计算:

$$\lambda(x_{u,i}^k = h) = N(o_{u,i}^k; \mu_{u,i-1}^h, \sigma_{u,i-1}^h) \quad (6)$$

$\eta(z_{u,i}^{(h)} = l)$ 为时空似然函数, 可由时空观测模型计算:

$$\eta(z_{u,i}^{(h)} = l) = N(d_{u,i} - d_{v,j}; \mu_{uv}, \sigma_{uv}^2) p(in_{u,i} | out_{v,j}) \quad (7)$$

$p(z_{u,i}^{(h)} | Y^{1:k-1})$ 是辅助变量 $z_{u,i}^{(h)}$ 的预测分布, 由上述条件独立性假设, 该预测分布可写为

$$p(z_{u,i}^{(h)} = l | Y^{1:k-1}) = p(X_{\mathcal{N}_u,l}^{1:k-1} = h, X_{\mathcal{N}_u,1:l-1}^{1:k-1} \neq h | Y^{1:k-1}) \quad (8)$$

进一步, 将关联变量的联合概率近似为各个关联变量边缘概率的乘积, 即

$$p(X_{\mathcal{N}_u}^{1:k-1} | Y^{1:k-1}) \approx \prod_l p(X_{\mathcal{N}_u,l}^{1:k-1} | Y^{1:k-1}) = \prod_l b_{\mathcal{N}_u,l}^\tau \quad (9)$$

综合式 (5)~(9) 可以看到, 在计算当前关联变量 $x_{u,i}^k$ 的信念状态 (即后验概率) 时, 摄像机 u 首先从其相邻节点收集观测信息 $Y_{\mathcal{N}_u}^{1:k-1}$ 和关联变量 $X_{\mathcal{N}_u}^{1:k-1}$ 的信念状态, 然后由式 (6) 和式 (7) 计算外观似然和时空似然, 由式 (8) 和式 (9) 计算辅助变量的预测分布, 最后由式 (5) 经积分计算 $x_{u,i}^k$ 的信念状态.

2.2 在线模型更新

本节讨论模型参数 Θ_u 的在线更新算法. 前面假设每个目标在各个摄像机下的外观观测服从高斯分布, 因此目标外观模型的更新可以看作是一个高斯混合模型的估计问题. 由于目标外观在每个摄像机节点上都用单独的高斯模型描述, 因此模型的更新可在各个摄像机节点上独立进行. 在下面的讨论中, 用 i 来对摄像机 u 上的观测进行索引. 我们采用与文献 [18] 类似的推导过程, 通过最大化经衰减函数调制的似然函数而得到在线模型更新算法. 假定历史观测数据对当前估计的影响函数 $S_i(t) = \alpha \exp((i-t)/\tau)$ 呈指数衰减, 其中时间常数 τ 控制着衰减速度, $\alpha = 1 - e^{-1/\tau}$ 为归一化常数. 假设当前模型参数为 $\Theta_{u,i-1}$, 则在给定观测 $y_{u,1:i}$ 的条件下, 模型参数由下式给出:

$$\Theta_{u,i} = \arg \max_{\Theta} \sum_{t=1}^i S_i(t) \log(L(y_{u,t}; \Theta_{u,i-1})) \quad (10)$$

其中, $L(\cdot)$ 为观测数据的似然函数. 对上式进行求解就可得到模型更新算法, 具体过程如下: 首先, 根据第 2.1 节给出的关联变量后验分布以及当前观测计算如下几个中间变量:

$$\theta_{u,i}^h \approx \sum_{t=1}^i S_i(t) p(x_{u,t} = h; \Theta_{u,t-1}) = (1 - \alpha) \theta_{u,i-1}^h + \alpha p(x_{u,i} = h; \Theta_{u,i-1}) \quad (11)$$

$$\beta_{u,i}^h \approx \sum_{t=1}^i S_i(t) o_{u,i} p(x_{u,t} = h; \Theta_{u,t-1}) = (1 - \alpha) \beta_{u,i-1}^h + \alpha o_{u,i} p(x_{u,i} = h; \Theta_{u,i-1}) \quad (12)$$

$$\gamma_{u,i}^h \approx \sum_{t=1}^i S_i(t) o_{u,i}^2 p(x_{u,t} = h; \Theta_{u,t-1}) = (1 - \alpha) \gamma_{u,i-1}^h + \alpha o_{u,i}^2 p(x_{u,i} = h; \Theta_{u,i-1}) \quad (13)$$

接着, 根据上述结果对模型参数进行如下更新:

$$\pi_{u,i}^h = \frac{\theta_{u,i}^h}{\sum_j \theta_{u,i}^j} \quad (14)$$

$$\mu_{u,i}^h = \frac{\beta_{u,i}^h}{\theta_{u,i}^h} \quad (15)$$

$$\sigma_{u,i}^h = \frac{\gamma_{u,i}^h}{\theta_{u,i}^h} \quad (16)$$

值得注意的是, 在目前的分布式 EM 算法^[10-15, 17]中, 关联变量的后验概率 (即隶属度) 是根据当前摄像机节点上的外观观测计算的, 即

$$p(x_{u,i} = h; \Theta_{u,i-1}) = \frac{\pi_{u,i}^h N(o_{u,i}; \mu_{u,i-1}^h, \sigma_{u,i-1}^h)}{\sum_j \pi_{u,i}^j N(o_{u,i}; \mu_{u,i-1}^j, \sigma_{u,i-1}^j)} \quad (17)$$

而在本文中, 隶属度是根据整个网络中的外观与时空观测由分布式推理算法进行计算, 因而本文的方法可以更有效地利用网络中的时空观测信息. 此外, 在目前的分布式 EM 算法中, 通常假定目标在不同节点下的外观观测服从同一高斯分布. 因此, 分布式 EM 中的 E 步可以在各个摄像机节点上独立进行, 而 M 步则需要利用某种共识算法^[10-15, 17]通过摄像机节点间的通信完成. 就分布式数据关联问题而言, 分布式 EM 算法可以看作是一种基于纯外观观测的单模型分布式关联算法.

3 实验结果

3.1 实验数据

我们在教学楼一层的 8 个不同地点安装摄像头节点, 采用分辨率为 752 像素 × 582 像素, 快门 1/50 ~ 1/100 000 s 的摄像头组建监控网络. 摄像头视场没有重叠, 网络结构如图 1(a) 所示. 我们采用 PAL 制式来采集视频, 共采集了 5 个行人在 8 个摄像头上产生的 200 段视频信息, 人工提取每段视频对应的观测数据 $y_{u,i}^k = \{o_{u,i}^k, d_{u,i}^k, in_{u,i}^k, out_{u,i}^k\}$, 其中, $o_{u,i}^k$ 为一段视频中中间帧图像目标所在区域外观 9 维向量的平均值, $d_{u,i}^k$ 为该段视频中某一帧的采集时间, $in_{u,i}^k$ 和 $out_{u,i}^k$ 为该段视频中目标进入和离开的方位.

值得注意的是, 由于光照等观测条件的影响, 同一行人在不同摄像机下的外观具有较大差异. 图 3 是某个行人在 8 个摄像机下的外观观测, 可以看到在摄像机 D、H 处, 该行人的外观与其他摄像机下有较大差别, 这主要是由于摄像机 D、H 处的光照较暗造成的.

3.2 算法比较和评价指标

为验证本文数据关联算法效果, 我们进行了如下三种在线算法的比较: 1) 算法 1, 文献 [17] 提出的分布式在线 EM 算法, 该算法假设同一目标在不同传感节点上的观测数据服从同一高斯模型, 可以看作是一种基于纯外观观测的单模型数据关联算法; 2) 算法 2, 在文献 [17] 的基础上, 通过第 2.1 节的算法, 利用外观和时空观测计算观测数据的隶属度, 属于一种单模型分布式在线数据关联算法, 与算法 1 相比, 可以更为有效地利用时空信息; 3) 算法 3, 即本文提出的多模型分布式在线数据关联算法.

数据关联算法的目标是将观测数据划分到对应于不同目标的集合中, 从而形成每个目标的重构轨迹. 我们用以下准则来评价算法优劣^[7]:

查准率:

$$P = \frac{1}{K} \sum_{s=1}^K \frac{\max_i |\hat{C}_s \cap C_i|}{|\hat{C}_s|} \quad (18)$$



图 3 行人外观观测

Fig. 3 The appearance observations under different cameras

查全率:

$$R = \frac{1}{K} \sum_{i=1}^K \frac{\max_s |\hat{C}_s \cap C_i|}{|C_i|} \quad (19)$$

F1 测度:

$$F1 = \frac{2 \times P \times R}{P + R} \quad (20)$$

其中, K 代表目标数量, C_i 代表第 i 个目标的真实轨迹划分, $\hat{C}_s$ 代表跟踪算法产生的轨迹.

3.3 结果

在算法实现过程中, 按如下方式对模型初始值 $\Theta_{u,0} = \{\pi_{u,0}^h, \mu_{u,0}^h, \sigma_{u,0}^h\}$ 进行设定: 目标 h 出现在摄像机 u 的先验概率 $\pi_{u,0}^h$ 按均匀分布给定; 每个目标的外观均值初值 $\mu_{u,0}^h$ 在外观观测值域内随机设定; 外观方差初值 $\sigma_{u,0}^h$ 设定为对角元素为 30 的对角阵. 对于中间变量 $\theta_{u,0}^h, \beta_{u,0}^h, \gamma_{u,0}^h$, 设 $\theta_{u,0}^h = \pi_{u,0}^h$, 依据式 (14), 可得到 $\beta_{u,0}^h, \gamma_{u,0}^h$ 的值. 在模型在线更新过程中, 遗忘因子 α 取为 0.2. 我们用上述三种算法对实验观测数据进行处理, 关联结果如表 1 所示.

可以看到, 由于目标在不同节点外观变化较大, 使得基于纯外观的关联推理准确率较低. 算法 2 通过引入时空观测使得关联效果有了一定的改善, 但由于使用单一的高斯模型描述目标外观, 因而关联

效果仍不够理想. 本文算法采用多外观模型, 具有最高的关联准确率.

表 1 三种数据关联算法实验结果 (%)

Table 1 Results of different inference algorithms under real data (%)

	查准率	查全率	F1 测度
算法 1	50.08	52.50	51.66
算法 2	75.79	74.00	74.88
算法 3	88.22	87.50	87.86

图 4 给出了摄像机节点 A 、 D 上产生的观测对应的关联变量边缘分布的推理结果, 其中, A 节点处光照较强, 而 D 节点处光照较暗. 图中每一列表示一个观测对应的关联变量的边缘分布, 灰度代表概率值, 纯黑色代表概率为 1, 纯白色代表概率为 0. 该观测对应目标的真实标号用星号表示.

由图 4 可以看到, 在两个节点处算法 1 均具有较大的推理误差. 算法 2 和算法 3 可以明显地提高推理准确率, 将较大的概率值赋予正确的标号. 对比算法 2 和算法 3 的推理结果可以看到, 由于大部分摄像机节点处光照较强, 因此在 A 节点处二者推理准确率相近, 但在 D 节点处, 算法 3 准确率较高. 这是由于 D 点的光照较弱, 运用算法 2 进行外观模型

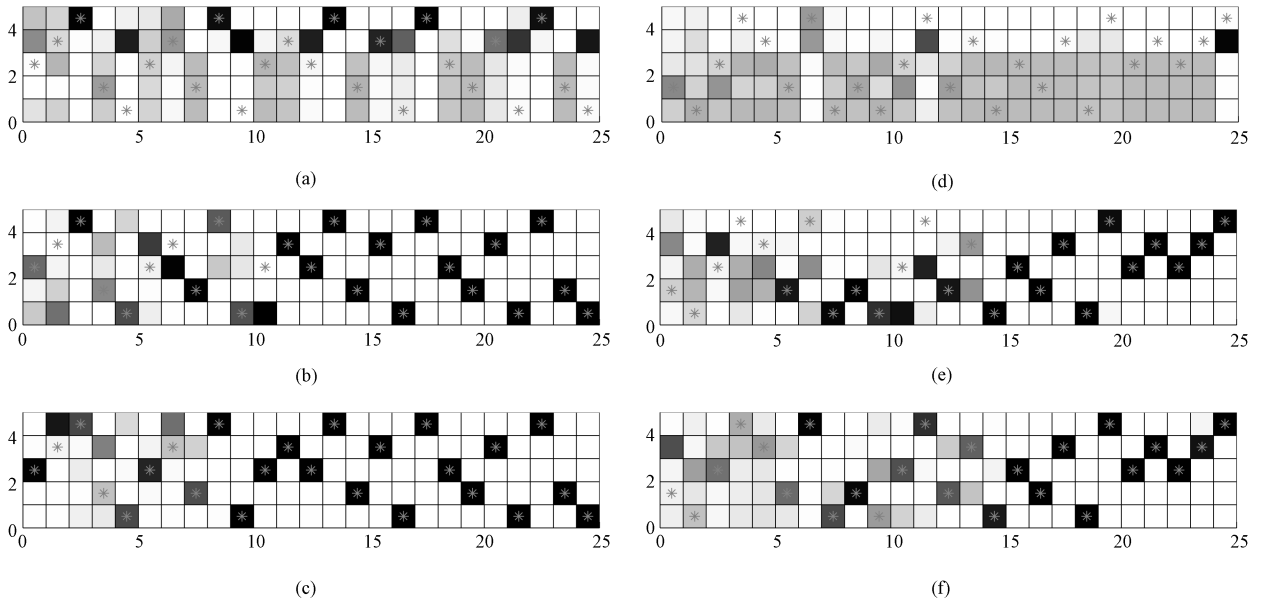


图 4 边缘概率分布 ((a)~(c) 为算法 1~3 在摄像机 A 上的推理结果; (d)~(f) 为算法 1~3 在摄像机 D 上的推理结果)

Fig. 4 Marginal distribution of labeling variable in exact inference ((a)~(c) are the results under camera A based on the above algorithms; (d)~(f) are the results under camera D based on the above algorithms.)

更新时, 每当观测出现在 D 点, 由于外观真实值与前一步单外观模型的更新值相差较大, 使单外观模型难以描述 D 节点处的外观. 而在算法 3 中, 假设目标外观值为多峰分布, 在 D 节点只利用了本节点的观测信息来更新模型状态, 因此准确率较高. 值得注意的是, 在算法 3 中, 当观测进入一个新的摄像头视域时, 由于外观高斯模型需要从初始状态进行更新, 此时的匹配度取决于初始状态与目标外观真实值的差异, 因而会出现部分的误判 (如图 4(f) 所示), 但随着本节点观测的增多, 外观模型参数会迅速收敛到真实值附近 (图 4(e) 与 4(f) 比较), 极大提高了之后的推理准确度.

图 5 给出了算法 3 的重构轨迹, 其中每一栏代表一个行人的轨迹, 虚框代表关联错误的行人. 可以看到, 在观测数较少时误判较多, 随着观测数目的增加, 外观模型的估计越来越准确, 关联准确率也随之增高.

在实际应用中, 不同行人的运动速度会有所区别, 同时行人在摄像头观测节点之间可能出现停顿, 这些因素增大了节点间运动时间的不确定性. 为了进一步考察算法的实用效果, 我们在捕获时间数据中人为地加入零均值高斯噪声, 进行蒙特卡洛仿真并统计结果.

图 6 是 100 次仿真后各性能指标的统计结果. 横坐标表示加入的高斯噪声的方差与真实捕获时间之比, 纵坐标分别为查准率、查全率与 $F1$ 测度值, 黑色细线表示统计方差.

由图 6 可以看出, 随着时空噪声的增加, 基于纯外观观测的算法 1 的准确率几乎没有变化, 而算法 2 和 3 的准确率逐渐下降, 但始终高于算法 1. 值得注意的是, 算法 2 的准确率随时空噪声的增加下降较快. 这主要是因为当时空噪声较大时, 外观观测在关联推理中的作用更加明显, 因而基于多外观模型的关联推理算法表现出较为明显的优势.

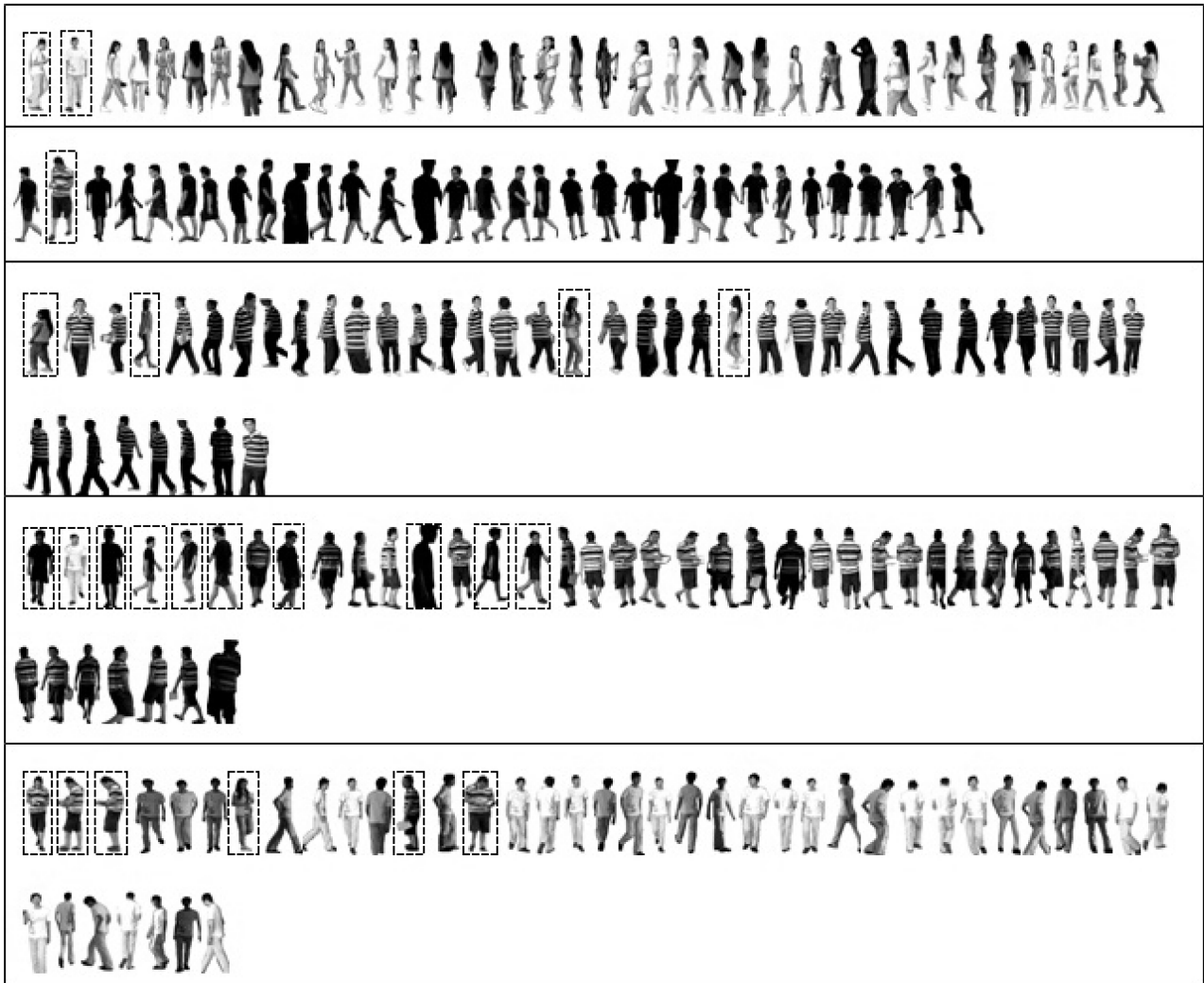


图 5 重构轨迹

Fig. 5 Trajectories recovery

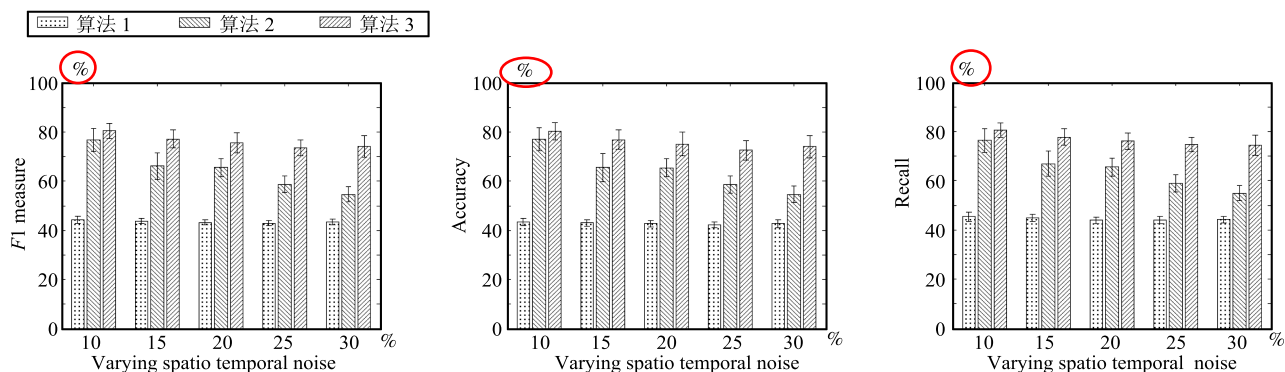


图 6 性能指标统计

Fig. 6 Statistical analysis of performance index

此外, 我们亦将本文算法与文献 [8] 中的集中式离线算法从关联精度和运算效率两方面进行了比较. 由于存储容量的限制, 文献 [8] 中的算法只能一次处理 100 个观测, 因此我们只比较了实验数据中前 100 个观测上的结果. 文献 [8] 中的算法的准确率为 88.46%, 查全率为 88.46%, $F1$ 测度为 88.46%. 本文算法的准确率为 86.31%, 查全率为 86.50%, $F1$ 测度为 86.40%. 可见, 本文算法的关联精度略低于集中式离线推理算法. 这主要是因为, 在进行 E 步关联推理时, 本文算法基于关联变量 x 和辅助变量 z 的边缘独立性假设, 见式 (5) 和式 (9), 而文献 [8] 算法基于条件独立性假设, 相对本文算法而言保留了变量间更多的相关性. 另外, 在 M 步更新过程中, 本文算法只利用了历史观测数据, 而文献 [8] 算法利用了历史和将来的全部观测数据, 并且进行了多次迭代. 虽然本文算法在关联精度上有所降低, 但在运算效率方面却有明显的提高. 在 E 步关联推理中, 本文算法推进一步需要计算 $H + Hl$ 个状态的概率, 而文献 [8] 中的算法进行一步关联推理需计算 $H + H^2k$ 个状态的概率, 其中 H 表示目标数, l 表示到当前时刻为止邻居节点上的观测数, k 表示到当前时刻为止全部网络中的观测数. 在普通 PC 机上用 Matlab 实现本文算法和文献 [8] 中的集中式推理算法, 处理 5 个目标产生的 100 个观测数据, 前者仅需要 0.43s, 而后者需要进行多次 EM 迭代, 每次 EM 迭代处理 100 个观测需要 149.49s.

4 结论

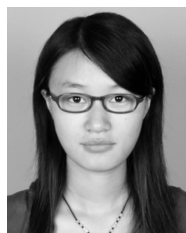
本文研究了视觉传感网络中多目标数据关联问题, 其中每个目标的外观由分布在各个摄像机节点的多个高斯模型描述, 给出了分布式关联推理算法和在线模型更新算法. 该算法中每个摄像机只需根据自身及相邻节点的历史观测数据进行局部信息处理, 适用于分布式在线应用场合, 实验结果表明了所

提算法的有效性. 本文方法的主要局限性在于需要事先给定目标的数目. 在进一步的研究工作中将考虑如何根据观测数据对目标数目进行在线估计.

References

- 1 Gilbert A, Bowden R. Tracking objects across cameras by incrementally learning inter-camera colour calibration and patterns of activity. In: Proceedings of the 9th European Conference on Computer Vision. Berlin, Heidelberg: Springer-Verlag, 2006. 125–136
- 2 Javed O, Shafique K, Rasheed Z, Shah M. Modeling inter-camera space-time and appearance relationships for tracking across non-overlapping views. *Computer Vision and Image Understanding*, 2008, **109**(2): 146–162
- 3 Kuo C H, Huang C, Nevatia R. Inter-camera association of multi-target tracks by on-line learned appearance affinity models. In: Proceedings of the 11th European Conference on Computer Vision. Berlin, Heidelberg: Springer-Verlag, 2010. 383–396
- 4 Song B, Roy-Chowdhury A K. Robust tracking in a camera network: a multi-objective optimization framework. *IEEE Journal of Selected Topics in Signal Processing*, 2008, **2**(4): 582–596
- 5 Liu Shao-Hua, Lai Shi-Ming, Zhang Mao-Jun. A min-cost flow based algorithm for objects association of multiple non-overlapping cameras. *Acta Automatica Sinica*, 2010, **36**(10): 1484–1489
(刘少华, 赖世铭, 张茂军. 基于最小费用流模型的无重叠视域多摄像机目标关联算法. 自动化学报, 2010, **36**(10): 1484–1489)
- 6 Zajdel W, Klöse B J A. A sequential Bayesian algorithm for surveillance with nonoverlapping cameras. *International Journal of Pattern Recognition and Artificial Intelligence*, 2005, **19**(8): 977–996
- 7 Zajdel W. Bayesian Visual Surveillance: from Object Detection to Distributed Cameras [Ph. D. dissertation], University of Amsterdam, Amsterdam, 2006
- 8 Wan Jiu-Qing, Liu Qing-Yun. Data association in visual sensor networks based on high-order spatio-temporal model. *Acta Automatica Sinica*, 2012, **38**(2): 236–247
(万九卿, 刘青云. 基于高阶时空模型的视觉传感网络数据关联方法. 自动化学报, 2012, **38**(2): 236–247)

- 9 Rinner B, Wolf W. A bright future for distributed smart cameras. *Proceedings of the IEEE*, 2008, **96**(10): 1562–1564
- 10 Mensink T, Zajdel W, Krose B. Distributed EM learning for appearance based multi-camera tracking. In: *Proceedings of the 1st ACM/IEEE International Conference on Distributed Smart Cameras*. Vienna, Austria: IEEE, 2007. 178–185
- 11 Nowak R D. Distributed EM algorithms for density estimation and clustering in sensor networks. *IEEE Transactions on Signal Processing*, 2003, **51**(8): 2245–2253
- 12 Dimakis A G, Kar S, Mouraand J F. Gossip algorithms for distributed signal processing. *Proceedings of the IEEE*, 2010, **98**(11): 1847–1864
- 13 Gu D B. Distributed EM algorithm for Gaussian mixtures in sensor networks. *IEEE Transactions on Neural Networks*, 2008, **19**(7): 1154–1166
- 14 Saul L K, Weiss Y, Bottou L. *Advances in Neural Information Processing Systems Vol. 17*. Cambridge: The MIT Press, 2005. 713–720
- 15 Olfati-Saber R, Shamma J S. Consensus filters for sensor networks and distributed sensor fusion. In: *Proceedings of the 44th IEEE Conference on Decision Control*. Seville, Spain: IEEE, 2005. 6698–6703
- 16 Wan J Q, Liu Q Y. Distributed data association in smart camera networks. In: *Proceedings of the 5th ACM/IEEE International Conference on Distributed Smart Cameras*. Ghent, Belgium: IEEE, 2011. 1–8
- 17 Morral G, Bianchi P, Jakubowicz J. On-line Gossip-based distributed expectation maximization algorithm. In: *Proceedings of the 20th IEEE Statistical Signal Processing Workshop*. Wisconsin, USA: IEEE, 2012. 305–308
- 18 Jepson A D, Fleet D J, El-Maraghi T F. Robust online appearance models for visual tracking. In: *Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition*. Kauai, Hawaii, USA: IEEE, 2001. 415–422



刘莉 北京航空航天大学自动化科学与电气工程学院硕士研究生. 主要研究方向为数字图像处理和目标跟踪.

E-mail: liulizi123@sina.com

(LIU Li Master student at the School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics. Her research interest covers digital signal processing and target tracking.)



万九卿 北京航空航天大学自动化科学与电气工程学院副教授. 主要研究方向为信号/图像处理, 统计推理与机器学习, 目标检测、跟踪与识别, 复杂系统故障诊断与健康管理. 本文通信作者.

E-mail: wanjiuqing@gmail.com

(WAN Jiu-Qing Associate professor at the School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics. His research interest covers signal/image/video processing, statistical inference, target detection, tracking and recognition, and fault diagnosis and prognosis of complex system. Corresponding author of the paper.)