

指导教师： 杨涛

提交时间： 2015.3.29

The task of  
**D**igital Image Processing

数字图像处理

School of Computer Science

No: 1

姓名： 董营营

学号： 2012302505

班号： 10011205

# 社会意识的大规模人群预测

Alexandre Alahi Vignesh Ramanathany Li Fei-Fei  
Computer Science Department, Stanford University  
Department of Electrical Engineering, Stanford University  
falahi, vigneshr, feifeilig@cs.stanford.edu

## 摘要

在拥挤的空间里，如市中心、火车站等，人类的流动性看似复杂，但往往是由仅有的一些原因引起的。我们提出了通过引入一个在火车站收集的 42,000,000 条轨迹大小的数据集，定性研究拥挤环境。在这个数据集的前提下，我们提出了预测行人目的地，一个在大规模人群流动的核心问题。我们需要克服由有限观测次数带来的（如稀疏的摄像机），以及跨越不同相机，行人外在改变的挑战。此外，在一个场景中，我们通常有行人在道路上可移动方面的限制，编码为始发地和目的地信息（OD）的偏好。在此框架下利用 OD-prior，我们提出了一个新的描述符，称为社会亲和力地图（SAM），来连接破损或不可预测的个体在人群中的轨迹。实验表明，通过使用 SAM 和 OD-prior，性能有了很大的提升。据大家所知，我们针对在数以百万的行人调查范围内，更好地理解群体行为方面的工作，是第一个提出令人鼓舞的结果的研究。

## 1. 引言

最近的研究表明，我们在城市规模水平的移动性上有着相当高的预测能力[31]。在 3 平方公里的不确定性的假设前提下，一个人在任何给定时刻的位置可以平均 93% 的精确度预测。如何在更精细的分辨率，如在购物中心，在机场，或者在火车终端实现安全或资源更优化？最佳预测人类行为相关的线索是什么？Kitani 等人认为场景语义是一个预测行人轨迹的强烈信号。以前的工作【12、22】也表明，我们的流动性受来自于邻居、亲戚、朋友，有意识甚至是无意识的影响，例如跟随一个单独的个体方便导航。在公共场所，为了捕获和预测人类流动，观察低密



图一：在社会亲和力地图（SAM）的条件下，预测行人的行为。我们定义 SAM 作为径向二进制描述符来表示你的相邻空间配置。

度和高密度人群都会迎来以下挑战：(i) 人们较多地遮挡对方降低了外观线索的区别度，(ii) 在人群中，独立运动之前 [15, 6, 1] 成为一弱假设，因为社会干预可以影响社会动力学。(iii) 观察往往是有限的，因为稀疏和分散网络摄像头的普遍安装。在本文中，为解决以上挑战，我们提出一个预测算法，利用细和粗先知先觉预测群体行为。我们提出了一个新的描述符，称为社会亲和力地图（SAM），通过稀疏的网络摄像头优先连接 tracklets<sup>1</sup> 来解决外观信息匮乏和较弱的独立运动。

流动显示有三个层次的理解[13]：战略层面（预期目标），战术水平（路线选择）和执行层面（每一时刻的实际运动）。我们建议研究后面的操作层面来预测先前的战略一。也就是说，我们模拟行人的社会干预来预测他们的目的地（见图一）。目的地的预测通常被称为 OD 估算矩阵[20]。它代表了所有行人在同一时间段轨迹的开始和结束点。我们的论文的主要贡献如下：

1. 我们介绍了一个从现实世界的火车站提取的 42,000,000 轨迹的大型数据集。

---

<sup>1</sup>A tracklet 是由一台单相机精准捕获的一

个轨道碎片。

2. 我们提出了一个新的特征描述符来捕获相邻行人的行为信号，术语上称为社会亲和力地图 (SAM)。

3. 我们将这个问题表示为一个线性整数规划问题，利用提出的 SAM 特性和 OD-prior，以及用一个启发式优化的方法来解决它。

## 1. 1. 大规模的数据收集活动

社会干预是存在于个体之间的关系，它可能不会发生在所有的环境中。为了更好的模拟和理解它们，我们需要在现实世界中捕捉而不是大规模的模拟数据。因此，我们已经在火车站安装了超过一百台的密集网络摄像头，来捕获行人的完整轨迹（见图二）。在任意给定的时间，多达一千人可以在相同的区域出现（如图二中走廊）。这样的数据收集活动使得验证社会亲和力的发生以及它们对预测现实世界行人 42,000,000 行为轨迹的影响成为可能。我们分享捕获的 dataset2，使得各研究协会从心理学，从计算机视觉进行深入的大规模分析人类在拥挤环境中流动性。

收集来自于网络摄像头的行人行为意味着以下步骤：(i) 检测，(ii) tracklet 的产生，(iii) tracklet 的关联性。对于 (i) 和 (ii)，我们使用先进的检测【11】和 tracklet 产生算法。简单地说，为了实现高精度的检测性能，我们安装了前视图，光学和热成像，来稳健地应对光照变化和自我屏蔽问题。当顶视图不可能时（由于低天花板），我们安装深度摄像头 (rgb-d 传感器) 捕捉三维检测部分遮挡。我们评估了超过 10,000 多次手工标记行人的检测性能。这导致了在 95% 的召回中有着超过 99% 的精度，由于受控视角和遥控模式。给定一预计算理想轮廓的字典，我们使用一个稀疏度驱动的框架分割近景轮廓 [1]。一旦人在地面上被定位，我们就可以解决由 Leal-Taixe 等人提出的最小网络流问题 [19]。最后的 (iii) tracklet 关联性将在第四部分网络摄像头中作进一步讨论。

## 2. 相关研究

**大规模的行人跟踪。**过去的几十年见证了許多相机网络数据集的各种应用程序，举几个来说，例如，体育分析 [7]，安全（宠物车间） [10]，交通建模 [29] 以及最近的视频理解数据集 [8]。他们由几个小时的视频序列组成。然而，他们不解决大规模设置，在拥挤环境中长期跟踪的起讫才是他们的兴趣所在。群体行为通常以这样的方式解决：只有一个摄像头监控马拉松的一部分或一个政治集会，OD 分析是有限的 [2, 28]。

**轨迹片段关联。**大量的工作模型可视化外观通过跨越相机实现轨迹片段的连接 [17, 9, 34, 26]。Andriluka 等人 [3] 使用人为检测作为线索执行检测跟踪，反之亦然。Javed 等人 [14] 使用旅行时间和外观特征的相似性。Song 等人 [30] 使用一个随机图进化策略。从每台相机提取的轨迹片段与匈牙利算法 [25] 相关，多通道多选择 [33]，或者全局最优贪婪的方法 [26]。这些方法并没有解决在高度拥挤的场景中数十米开外的轨迹片段连接问题。此外，考虑到摄像机视角和人流量的因素，往往只有行人的部分身体部位可以捕获（例如头发）。因此，其他线索也需要被充分利用。然而这些方法并没有捕捉社会干预，这在拥挤长途的设置中是非常具有价值的线索。

**社会跟踪。**最近的社会行为被纳入现有的跟踪框架中，通过卡尔曼过滤器 [21]，扩展卡尔曼过滤器 [23] 或者线性规划 [19, 27] 模拟知名社会因素 [12]。Antonini 等人 [4] 使用离散选择模型来模拟人的步行行为。当少数帧丢失（例如当给定一个低帧率，或短暂闭塞情况下），这些方法改进了操作级别的跟踪。他们也经常模型分组提示解决数据关联问题 [24, 19, 27]。他们模拟它作为一类似的速度和空间接近性。同样，[18] 在分层框架中使用分组信号确定运动员角色。类似地，分组信号通常是作为一个二进制变量指示处理的。然而，核心挑战在于使用更好的表示捕获组关联，集成它成为轨迹片段的关联问题。Yang 等人 [32, 24] 使用条件随机场的框架联合估计组成员以及跟踪。Leal 等人 [19] 迭代计算不同速度的最低成本流程和分组任务，直到收敛或者迭代次数达到最大。Qin 等

监视区域	区域密度	行走时间	距离	速度	流量/天	总计
20.000 m <sup>2</sup>	多达 1 ped/m <sup>2</sup>	1 min	100 m	1.37 m/s	100.000–250.000/terminal	42 million trajectories

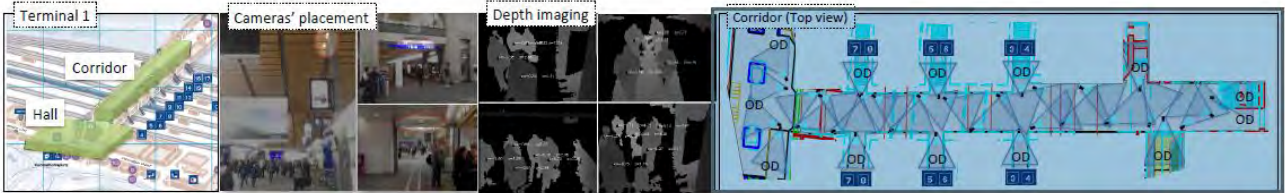


图 2: 实际设置。顶行提出了一些事实数据集(平均值)。底行说明了监视的通道之一。30 多个摄像机部署在如图所示的走廊, 而共有 132 个摄像头部署在 3 个走廊、一条小道和一个大厅。在任何给定的时间, 走廊的占用率可以达到一千多行人。标签“OD”表示入口/出口区。

人[27]使用匈牙利算法联合分组和连接轨迹片段。然而, 匈牙利算法并不能解决完整的长期跟踪的全局最小化问题, 然而, 最小网络流公式可以。在这项工作中, 我们提出了一个描述符, 将分组信号表示为高效匹配行人行为的特征。

### 3. 社会亲和力地图: SAM

我们收集的数据集使得我们可以进行在拥挤场景设置中的人类行为的研究。在本文中, 我们分析重点放在社会亲和力上, 它使得人们在大规模人群中维系在一起。我们对这样的行为线索非常感兴趣, 它在不同的时间以及基于不同的传感方式(例如光、热、深度), 依然保持稳定连接远距离的轨迹片段。

**定义 1** 我们定义“社会亲和力”为相邻个体之间的运动上的密切关系。

社会亲和力可以由朋友、亲戚或同事有意识地形成。然而, 在拥挤的环境中, 潜意识的亲和力依然存在。更正式地说, leaderfollower 模式捕获一个行人(即追随者)的行为, 为追随领导者, 他调整他的运动来完成流畅的运动。我们建议认识各种社会亲和力, 它将人们绑定到拥挤的场景中的这个新特性被称为社会亲和力地图(SAM)。

#### 3.1. The SAM 特性

我们观察到, 在公共场合中, 社会因素主

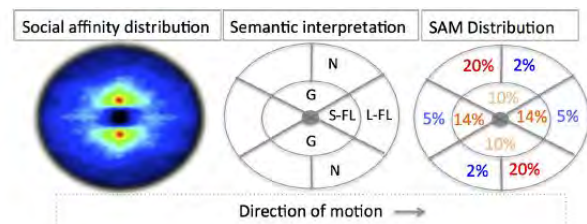


图 3: 左手边: 通过所有的轨迹片段绘制的关于所有相邻行人的相对位置的 Heatmap。中间一列: 按照我们的语义描述, 它表示了 SAM 中“G”组亲和力(例如夫妻、朋友)、“S-FL”是短距离 Follow-Leader 行为, “L-FL”是长途行为, 当走向同一个方向“N”可以被看作可保持的令人舒适的距离。右手边描述了提出的分配行为。

要是由彼此邻近的人决定的, 正如早先的工作提到的[12]。因为, 人们更容易受到其附近他人的影响, 我们提出了社会亲和力特性来捕捉相邻轨迹片段的空间位置。如图 3 所示, 我们基于放射性地装箱相邻轨迹片段的位置实现这一目标。我们进一步学习了空间装箱, 通过捕获得到的轨迹线, 归类合并周围个体的相对位置。我们认为相对位置应在 3m 的限制范围内, 来避免异常值(离群值)。在图 3 中, 相对位置的分布在百万轨迹线中是可视化的。如图所示, 我们获得了 10 箱作为分类归并的结果。汇集到这个箱子中相对位置的百分比也显示在图中。值得一提的是, 最常用的箱子竟然是最右边的(图 3 中的“N”

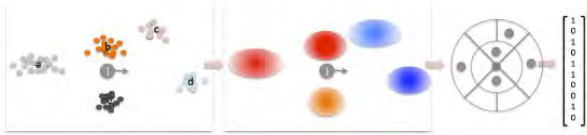


图 4: 社会关联映射提取(前视图)的说明。邻近的人的相对位置聚集到一个径向直方图。后者是一位量化的。

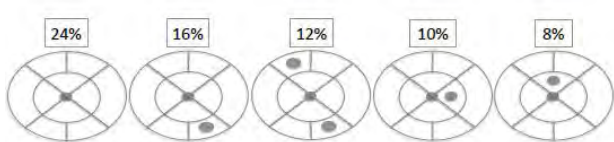


图 5: 所得数据中观察到的最多的 8 种社会亲和力。上述比例描述了相应 SAM 的出现频率。

的标签), 它可以解释为舒适的行走模式对于其他个体不是左手边。

考虑到一个新的轨迹片段, 我们进行了矢量量化 (VQ) 的编码来获得 SAM 特性。我们将高斯混合模型应用于周围轨迹片段的相对位置。在先前的推断 GMM 值学习空间离散获得二进制径向垃圾箱柱状图, 它代表了山姆的特征向量。完整的流程见图 4。将先前提到的空间箱子中的 GMM 推断值离散化来获得二进制径向柱状图, 它标志着 SAM 特征向量。完整的流程见图 4。汉明距离是用来比较轨迹片段间的 SAM。请注意, 二进制量化对特性的功效影响非常小, 它只是用来加快比较速度的方式。我们的 SAM 特性可以区分社会亲和力下的不同构造因素, 例如“夫妇散步”, 或者“Leaderfollower”行为。图 5 说明了在上百万轨迹中观察到的最多的 8 种。值得提出的是, 76% 的个体属于同一人群。因此, SAM 在拥挤场景设置中提供了重要信息, 推进了这些线索在预测行人流动性方面的应用。

#### 4. 预测流动: 问题公式化

我们有一个稀疏的摄像头网络, 来监视在公共场合中经过的行人, 如铁路枢纽。枢纽有一组入口点被称为原点, 和出口点被称为目标。我们工作的目标是确定每一个人

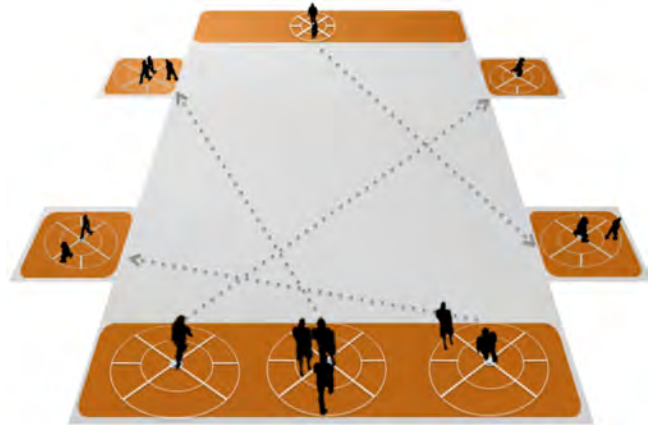


图 6: 考虑到社会亲和力地图 (SAM) 装有少数相机, 预测行人的行为。橙色部分代表了监控区域。我们举例说明了一极端的情况, 当相机只放置在入口或出口区, 称为 OD 相机。

在摄像头网络监控范围内的出发地与目的地 (OD)。我们通过识别相应的轨迹片段, 它连接了从位于起点的轨迹片段至位于终点的轨迹片段, 从而实现这一目标。被连接的中间轨迹片段的数量随着摄像头网络的稀疏度而减少。图 6 展示了一个只有起点轨迹片段和终点轨迹片段的极端情况。

我们有一组起始点 tracklets  $O$  和相同数目的一组目的地 tracklets  $D$ 。  $O$  中的每一个轨迹片段是在该区域诸多进口中的其中一个捕获的,  $D$  中的目标轨迹是在一个出口捕获的。我们也有一组有稀疏摄像头网络得到的中间轨迹片段  $X$ 。我们想要找到这样一组轨迹  $T$ , 其中的每条轨迹  $t \in T$ , 表示为一个轨迹片段的有序集合  $(o_t; X_t; d_t)$ , 其中  $o_t \in O$ ,  $d_t \in D$  是起点和终点轨迹中的轨迹片段。同样的,  $X_t = (x_t^{(1)}, \dots, x_t^{(n)})$  是一个与轨迹形成相关的轨迹片段的有序集合。这些轨迹片段是按照起始反应的时间排序的。这个问题可视为一个类似于 [34, 26] 最大后验估计问题:

$$T^* = \max_T P(X|T)P(T),$$

其中  $P(X|T)$  是  $X$  中的轨迹片段是真正的可能性。概率  $P(X|T)$  等于:

$$P(X|T) \propto \prod_{t \in T} \prod_{x \in X_t} \frac{P_{tp}(x)}{P_{fp}(x)},$$

其中  $P_{tp}(x)$  和  $P_{fp}(x)$  分别是轨迹片段真正、假正的概率。我们将 POD(o; d) 定义为 OD-prior 术语，它表明了一个人进入起始点对应于 o 和退出终止点对应 d。这样的 prior 经常被忽视和假定为统一的。然而，在许多应用中，它是一个 strong prior，例如避免在机场禁止路径。

接下来，类似于 [26]，我们假设一个马尔科夫链模型，用概率  $P(x_t^{(i+1)}|x_t^{(i)})$  来连接轨迹 T 中的每个中间轨迹  $x_t^{(i)}$  和随后的  $x_t^{(i+1)}$ 。轨迹概率 P(T) 是：

$$\begin{aligned} P(T) &= \prod_{t \in T} P(t), & (3) \\ P(t) &= P_{OD}(o_t, d_t) P(x_t^{(1)}|o_t) \\ &\quad \prod_{i=2}^n P(x_t^{(i)}|x_t^{(i-1)}) P(d_t|x_t^{(n)}), \end{aligned}$$

其中  $n = |X_t|$  是轨迹集合中中间轨迹片段的数量。

Eq. 1 中的 MAP 问题可以表述为一个类似于 [26] 的方式线性整数规划问题。

$$\begin{aligned} \min_f \quad & C(f) & (4) \\ C(f) = \quad & \sum_{x_i \in X} \alpha_i f_i + \sum_{x_i, x_j \in X} \beta_{ij} f_{ij} + \\ & \sum_{\substack{x_i \in X, \\ o \in O}} \beta_{oi} f_{oi} + \sum_{\substack{x_i \in X, \\ d \in D}} \beta_{id} f_{id} + \sum_{\substack{o \in O, \\ d \in D}} \gamma_{od} f_{od} \\ \text{s.t} \quad & f_i, f_{ij}, f_{od} \in \{0, 1\} \\ \text{和} \quad & f_i = \sum_j f_{ij} + \sum_d f_{id} = \sum_i f_{ji} + \sum_o f_{oi}, \\ & \sum_{od} f_{od} = |O| = |D|, \\ & \sum_d f_{od} = \sum_i f_{oi}, \\ & \sum_o f_{od} = \sum_i f_{id} \quad \forall x_i, x_j \in X, o \in O, d \in D, \end{aligned}$$

其中  $f_i$  是流量变数，表明相应的轨迹片段是

否是真正的， $f_{ij}$  表明相应的轨迹片段是否是连接在一起的。变量  $\beta_{ij}$  表示由  $\log P(x_i | x_j)$ ，其中  $x_i, x_j \in X$  给出的转型成本。对数似然函数  $\beta_{oi}$ ， $\beta_{od}$  也是类似的定义，从起始轨迹片段 o 到终止点轨迹片段 d。本地费用  $\alpha_i$  是中间真正轨迹的对数似然函数。最后，OD-prior 费用表示为  $\lambda$ 。

我们注意到 Eq. 4 中的最优化问题等价于在 [26, 34] 中不存在 OD-prior 术语时广泛讨论的流量最优化问题。这类问题可以通过 k 最短路径或 [26] 中更有效的贪心算法。然而，OD-prior 术语的添加导致了网络流问题中的循环，它已不再能通过最短路径算法被精确解决。因此，我们采用启发式近似法来解决 Eq. 4，正如在 Sec. 5 中提到的。

#### 4. 1. 本地成本

本地成本  $\alpha_i$  正比于轨迹片段的长度。它有利于帮助我们移除可能为假正的短轨迹片段。

#### 4. 2. 转型成本

转型成本  $\beta_{ij}$  是对于任意两个轨迹片段分成如下所示的双组分：

$$\beta_{ij} = \beta_{ij}^{SAM} + \beta_{ij}^M, \quad (5)$$

其中  $\beta^{SAM}$  是社会亲和力成本， $\beta^M$  是确保连接轨迹片段的平滑性的费用。

社会亲和力成本。在我们的模型中，我们希望确保移动在相似社会群体的轨迹片段有很大的可能性连接到彼此。这种亲和力在大范围的跟踪情境中正如我们的，个体的外观不是很具有区分性，形成了很重要的组成部分。在 Sec. 3 中介绍的 SAM 特性通常被用来测量移动在群体的轨迹片段之间的社会亲和力距离，如下所示：

$$\beta_{ij}^{SAM} = \mathbf{H}(sam_i, sam_j), \quad (6)$$

其中  $\mathbf{H}(\cdot)$  表示两二进制矢量之间的汉明距离,  $sam_i, sam_j$  表示两轨迹的 SAM 特征向量。

运动相似性。另外一个提示  $\beta^M$  通过测量两轨迹片段的运动模型之间的距离, 类似于

【33, 30】来获得, 它习惯于被用作确保轨道运动的平滑性。

### 4.3. OD-prior 成本

OD-prior 成本是从起始点到终止点运动中先验概率的对数似然函数。在大多数的监视场景中, 我们可以用终端地理学和粗略估计乘客运费的先验知识来获得 OD prior。另外, OD prior 可以被用做强制约束条件, 例如乘客进入到一个特定的入口点不能通过一个平行的入口返回相同的位置。在我们的后期实验中, OD prior 是通过定位的信息反馈来获得的。它将随着数据集被释放。

## 5. 优化

如上所述, Eq. 4 中的优化问题不能通过传统跟踪情况下现有的最短路径算法【26】琐细地解决。因此, 我们采用启发式近似法, 在后面有详细解释。

**OD prior 的贪心优化。** 首先我们运行一个贪心算法来确定图表中的低成本解决方案。

1. 找到在 Eq. 4 中连接起讫轨迹片段的最短路径。
2. 移除在以上的步骤中已获取和重复的轨迹片段。

贪心算法提供了这个问题的一个近似解, 而且计算比较高效。然而, 它不能解决全局最优化问题。我们用一个简单的启发式搜索来得到一个更好的解决方案。

**OD 再加权优化费用。** 贪心算法的解决方案帮助我们确定符合 OD-prior 的路径。因此, 这一算法设置的过渡流量参数提供了在 OD-prior 下的轨迹片段之间成对亲和力的粗略估计。我们利用这一直觉添加一额外费用,

它不利于在贪心算法中未被连接的轨迹片段的连接。当添加这一费用时, 我们删除了原本的 OD-prior 费用  $\gamma_{od}$ , 因而产生了一网络流量问题, 它可以通过 K-最短路径算法解决。

改进后的费用  $\tilde{C}$  如下所示:

$$\tilde{C}(f) = \sum_{x_i \in X} \alpha_i f_i + \sum_{x_i, x_j \in X} \tilde{\beta}_{ij} f_{ij} + \sum_{\substack{x_i \in X, \\ o \in O}} \tilde{\beta}_{oi} f_{oi} + \sum_{\substack{x_i \in X, \\ d \in D}} \tilde{\beta}_{id} f_{id}, \quad (7)$$

其中 OD 再加权费用  $\tilde{\beta}$  定义如下:

$$\tilde{\beta}_{ij} = \beta_{ij} + \lambda \mathbf{1}(f_{ij}^{greedy} = 1), \quad (8)$$

其中  $f_{ij}^{greedy}$  是由贪心算法获得的,  $\lambda$  参数表明 OD-prior 成本的优势。转型成本是对包括起讫在内的所有成对轨迹片段的再加权。

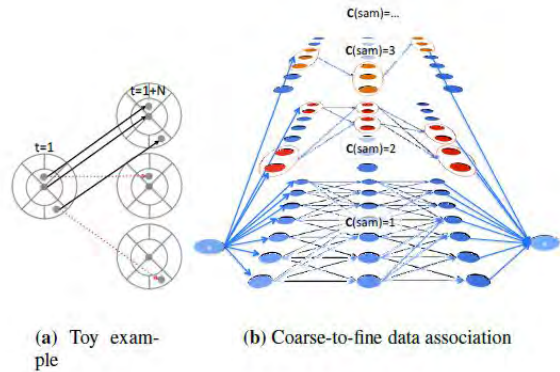


图 7: (一)玩具的例子中 3 个轨迹片段可能错误地联系在一起。红色虚线箭头说明在没有一个由粗到精的数据关联的情况下可能发生错误的分配。(b) 在 SAM 基数下的由粗到精数据关联。在特定基数的轨迹片段组中, 每个子图相当于轨迹片段的相关性问题, 由  $C(sam)$  表示 SAM 特征元素的总和。通过求解这些问题得到的流量变量是用来定义用于最终优化的额外转换成本。

## 6. 由粗到精的数据关联

在 Sec. 4 中出现的模型使用一社会亲和力费用来保证相似群体的轨迹片段是连接在一起的。然而，它并不能解释隶属于不同基数（群体中的人数）群体的人们可以共享 SAM 特性。在图 7 (a) 中的例子中，两个隶属于不同基数群体的轨迹片段因为相似的 SAM 错误地连接到一起（用红色标出）。但是，我们希望来自相似规模的群体的轨迹片段连接到一起（黑色箭头）。我们通过提出一个由粗到精的数据关联方式来解释它。

我们收集在相同时刻共同出现的轨迹片段，基于群体的分离分成不同的组。一轨迹片段的基数  $C(x_i)$  是对应于轨迹片段  $x_i$ ，该群体的人数。我们可以试想一下，如果聚类完美且人们按照相同的组态跨越整个摄像头网络，它将足以连接整个轨迹片段组而不是单独的轨迹片段。同时它也解决了跨越不同基数的组中轨迹片段的连接问题。然而，在实际的场景设置中，分组并不是完美的且人们常脱离分组。因此，我们连接相同基数的组，并且利用从这一组跟踪获得的连接来定义额外的转型成本。完整的流程见补充文件。这一方式简单形象地呈现在图 7 (b) 中。

## 7. 实验

### 7.1. 大规模的评估

数据收集活动帮助我们在现实场景中基于大量动态人群实施各种各样的实验。在本节内容中，我们介绍一组实验来提出有引进的数据集给出的预测问题。我们选定网络摄像头的一子集，测量我们算法的性能用仅有的摄像头来预测流动性。

**测试。** 先前的工作已研究了就详细数据（例如撤销/准确率，MOTA，MOTP 等到）而言，给定检测跟踪算法的影响[5]。在这项工作中，我们对拥挤的场景比较感兴趣，在这里只有一部分场景被摄像头网络覆盖。因此，我们更感兴趣的是，对进入摄像头网络的个体起讫评估的准确率。我们限定了 14 个起讫点的监视，产生了一轨迹中 196 种可能的 OD 路径。我们将摄像头分成两组：一组为属于 OD 位置（例如抓取远距离轨迹片段的起点和终点 0）另一组为这些位置之间的摄像头。我

们计算的 OD 出错率作为摄像头覆盖区域下总人数的错误百分比。

**地面真值。** 海量资料被收集后，标记数以百万的轨迹是不现实的。因此，我们安置了密集的摄像头网络来尽可能的减少盲点，并且连接距离彼此只有几公分的轨迹片段。由此密集网络计算得出的轨迹作为一基准线。然而由密集网络计算得到的轨迹（和 OD）并不是完美的地面真值，在实践中在我们的规模中，他们不太容易并以比手工注释轨迹更低的成本获得。我们的预测算法的目标是在使用稀疏摄像头网络时达到堪比使用秘籍摄像头网络的性能。

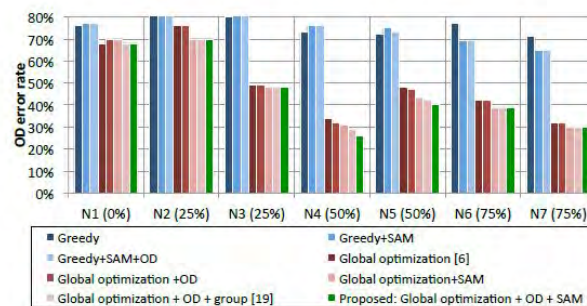


图 8: 基于摄像头间隔的不同数量，OD 预测的性能。摄像头间隔的百分比相机见括号内。七个网络配置的评估结果（简称为 N1 到 N7）。

### 7.2. OD 预测

图 8 显示了使用 7 个稀疏网络摄像头，OD 出错率的结果。这项评估是基于不同级别的网络稀疏度提出的，从摄像头之间的 0% 到 75%。例如，N4 和 N5 在走廊中使用可得的一半摄像头（见图 2）。这些摄像头被用来选作试探性地使得在任意给定稀疏度的情况下最小化它们之间的最小距离，同时我们也评估不同摄像头的配置，礼服 N4 和 N5 50% 的稀疏度。来自网络 N1 到 N3 的轨迹片段平均距离彼此几十米开外。为了验证我们的算法，我们评估了贪心优化算法和提出的全局优化算法的性能。我们测评了使用 SAM 作为一额外特性和由粗到精地跟踪模拟 OD prior 的影响。

正如预期的，在有或者没有 OD prior 的情况下全局优化方式总是优于贪心算法。使用



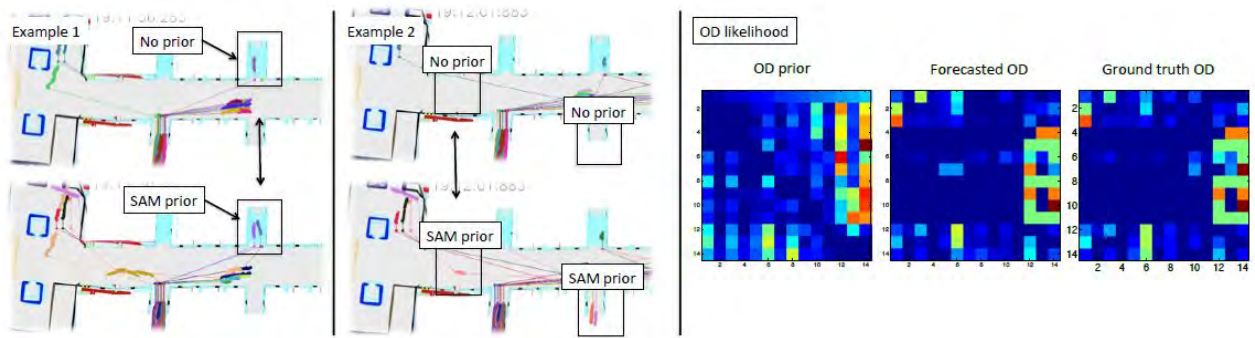


图 9: 稀疏网络 1 中连接轨迹片段的定性结果, 50% 的中间摄像头不适用。用这种方式选出的轨迹片段如图所示。连线表示连接的轨迹片段。在右边, 我们将 OD prior, ForecastedOD 和 Ground truthOD 绘制成热点图。我们可以看到尽管他们是不同的但是最终结果仍然类似于地面真值。

全局优化方法, 性能较之前改进了两倍多。SAM 特性和 OD 再加权成本的使用 (OD-prior 的使用) 在使用全局优化方法时都有很有效的作用。这证明了我们的决定的有效性, 在优化期间试探性地模拟 OD-prior 的效果。我们也比较了从【6】到【19】中的算法。我们的完整模型, “全局优化+OD+SAM”, 在观察范围限制在走廊是优于这些方式。请注意, 摄像头的放置对于预测是有影响的。虽然在 N2heN3, 或者 N4 和 N5 中使用相同数量的摄像头, 但是预测的精度因网络的不同而不同。如果一中间摄像头战略性地安放来捕获频繁的路线, 它就减少了耦合战略的不确定性。这导致了对于相同摄像头数量, 不同网络的不同性能, 如图 8 所示。

我们测试了极端设置, 例如在摄像头之间没有中间摄像头 (标记为 N1), 即我们仅在进出口区域安有摄像头。在此情况下, 轨迹片段距离彼此达到 100m。图 8 显示了性能的下降。贪心算法和全局优化方式之间的差距大大地减小了。另外, SAM 特性和 OD Prior 在此极端情况下并没有显著地影响。这些结果激励着我们未来的工作重点放在类似极端情况的处理上。

图 9 展示了一些定性的结果, 说明了 SAM 的力量及优势。同时我们也分析了 OD prior, 用半数的摄像头稀疏网络作为密集网络预测 OD (地面真值)。SAM 的影响。我们说明了有

完整模型实现的轨迹片段连接, 将它和在不使用 SAM 时利用全局优化方式实现作比较, 见图 9。正如预期的, 我们看到在使用 SAM 的情况下, 在相似群体结构移动的轨迹片段并没有连接在一起, 导致了错误的结果。另一方面, SAM 有利于消除相似轨迹片段的选择难度, 除了群体设置。OD prior 的影响, 在图 9 中, 我们展示了有我们的完整模型估算的 OD 矩阵, 并将它和在使用 OD prior 和地面真值 OD 的条件作比较 (基于密集网络摄像头)。很显然, 它仅仅提供了关于真正 OD 很弱的信号, 但是它在减少权重路径, 即非常不利的情况例如堵塞的走廊方面非常有用。我们的模型预测的 OD 矩阵非常接近于由于一密集网络摄像头获得的地面真值 OD 矩阵。

## 8. 总结

我们已经解决了在现实世界拥挤的火车站用有限数量的摄像头预测行人目的地的问题。我们定量地展示了社会亲和力的存在并帮助解决预测问题。提出的 SAM 描述符使得独立运动行为的轨迹片段连接问题得以全局优化。部署的摄像头网络使得在现实世界拥挤场景的运动得到大量分析。每天数十万轨迹被收集, 到目前为止产生了 42, 000, 000 个轨迹。另外为了提高 OD 矩阵的估算年未来的工作可以使用这些数据调整行人模拟器, 或者了解到一种在有限的摄像头条件下, 理想的摄像头安放方式。

## 致谢

我们在此感谢 P. Vanderghenst, M. Bierlaire, J. Paratte, 和 D. Chanel 提供的有用代码和点评。Alexandre Alahi 是由瑞士自然科学基金资助的, 为: PBELP2-141078。

## 参考文献

- [1] A. Alahi, L. Jacques, Y. Boursier, and P. Vanderghenst. Sparsity driven people localization with a heterogeneous network of cameras. *Journal of Mathematical Imaging and Vision*, 2011. 1, 2
- [2] S. Ali and M. Shah. Floor fields for tracking in high density crowd scenes. In *ECCV*. 2008. 2
- [3] M. Andriluka, S. Roth, and B. Schiele. People-tracking-by-detection and people-detection-by-tracking. In *CVPR*, 2008. 2
- [4] G. Antonini, M. Bierlaire, and M. Weber. Discrete choice models of pedestrian walking behavior. *Transportation Research Part B*, 2006. 2
- [5] F. Bashir and F. Porikli. Performance evaluation of object detection and tracking systems. In *PETS*, 2006. 6
- [6] J. Berclaz, F. Fleuret, E. Turetken, and P. Fua. Multiple Object Tracking using K-Shortest Paths Optimization. *TPAMI*, 2011. 1, 7
- [7] C. De Vleeschouwer and D. Delannay. Basket ball dataset from the european project apidis, 2009. 2
- [8] G. Denina, B. Bhanu, H. T. Nguyen, C. Ding, A. Kamal, C. Ravishankar, A. Roy-Chowdhury, A. Ivers, and B. Varda. Videoweb dataset for multi-camera activities and non-verbal communication. In *Distributed Video Sensor Networks*, pages 335 - 347. Springer, 2011. 2
- [9] A. Ess, B. Leibe, K. Schindler, and L. Van Gool. A mobile vision system for robust multi-person tracking. In *CVPR*. IEEE, 2008. 2
- [10] J. Ferryman and A. Ellis. Pets2010: Dataset and challenge. In *Advanced Video and Signal Based Surveillance (AVSS)*. IEEE, 2010. 2
- [11] M. Golbabaee, A. Alahi, and P. Vanderghenst. Scoop: A real-time sparsity driven people localization algorithm. *Journal of Mathematical Imaging and Vision*, 48(1):160 - 175, 2014. 2
- [12] D. Helbing and P. Molnar. Social force model for pedestrian dynamics. *Physical review E*, 1995. 1, 2, 3
- [13] S. P. Hoogendoorn and P. H. Bovy. Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, 38(2):169 - 190, 2004. 1
- [14] O. Javed, Z. Rasheed, K. Shafique, and M. Shah. Tracking across multiple cameras with disjoint views. In *Proc. IEEE International Conference on Computer Vision*, page 952, Washington, DC, USA, 2003. IEEE Computer Society. 2
- [15] S. M. Khan and M. Shah. Tracking multiple occluding people by localizing on multiple scene planes. *IEEE*

## 数字图像处理作业 The task of Digital Image Processing

- Transaction on Pattern Analysis and Machine Intelligence, 31(3):505 - 519, 2009. 1
- [16] K. Kitani, B. Ziebart, J. Bagnell, and M. Hebert. Activity forecasting. 1
- [17] C. Kuo, C. Huang, and R. Nevatia. Inter-camera association of multi-target tracks by on-line learned appearance affinity models. ECCV, 2010. 2
- [18] T. Lan, L. Sigal, and G. Mori. Social roles in hierarchical models for human activity recognition. In Computer Vision and Pattern Recognition (CVPR), 2012. 2
- [19] L. Leal-Taixe, G. Pons-Moll, and B. Rosenhahn. Everybody needs somebody: Modeling social and grouping behavior on a linear programming multiple people tracker. In ICCV Workshops, 2011. 2, 7
- [20] K. Lindveld. Dynamic OD matrix estimation: a behavioural approach. 2003. 1
- [21] M. Luber, J. Stork, G. Tipaldi, and K. Arras. People tracking with human motion predictions from social forces. In ICRA, pages 464 - 469, 2010. 2
- [22] M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz. The walking behaviour of pedestrian social groups and its impact on crowd dynamics. PloS one, 5(4):e10047, 2010. 1, 3
- [23] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool. You'll never walk alone: Modeling social behavior for multi-target tracking. In ICCV, 2009. 2
- [24] S. Pellegrini, A. Ess, and L. Van Gool. Improving data association by joint modeling of pedestrian trajectories and groupings. In ECCV. 2010. 2
- [25] A. A. Perera, C. Srinivas, A. Hoogs, G. Brooksby, and W. Hu. Multi-object tracking through simultaneous long occlusions and split-merge conditions. In CVPR, 2006. 2
- [26] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes. Globally optimal greedy algorithms for tracking a variable number of objects. In CVPR, 2011. 2, 4, 5
- [27] Z. Qin and C. R. Shelton. Improving multi-target tracking via social grouping. In CVPR. IEEE, 2012. 2
- [28] M. Rodriguez, I. Laptev, J. Sivic, and J.-Y. Audibert. Density-aware person detection and tracking in crowds. In ICCV, 2011. 2
- [29] Z. Shuai, S. Yoon, S. Oh, and M.-H. Yang. Traffic modeling and prediction using sensor networks: Who will go where and when? ACM Transactions on Sensor Networks (TOSN), 2012. 2
- [30] B. Song, T. Jeng, E. Staudt, and A. Roy-Chowdhury. A stochastic graph evolution framework for robust multi-target tracking. ECCV, 2010. 2, 5
- [31] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási. Limits of predictability in human mobility. Science, pages 1018 - 1021, 2010. 1
- [32] B. Yang, C. Huang, and R. Nevatia. Learning affinities and dependencies for multi-target tracking

using a crf model. In  
CVPR, 2011. 2

[33] Q. Yu, G. Medioni, and I. Cohen.  
Multiple target tracking  
using spatio-temporal markov chain monte  
carlo data association.  
In CVPR, pages 1 - 8, 2007. 2, 5

[34] L. Zhang, Y. Li, and R. Nevatia.  
Global data association for  
multi-object tracking using network  
flows. In CVPR, 2008.  
2, 4, 5