

# 基于流密度的多重交互集体行为识别算法

吴云鹏 赵晨阳 时增林 叶阳东

(郑州大学信息工程学院 郑州 450052)

**摘要** 作为当前模式识别领域的研究热点之一,识别视频场景中的集体行为有着重要的科研意义和应用价值。已有的识别方法都忽略了对集体行为中多重交互关系的描述和分析,使得复杂视频场景中的集体行为识别仍存在局限性。该文提出了一种基于流密度的多重交互集体行为识别算法 FDA,能够识别集体行为的局部和全局模式。为了准确地衡量个体间的行为一致性程度,受流体力学启发定义了流密度,并通过基于流密度的聚类算法识别局部子群组;为了精确描述集体行为客观存在中的多重交互关系,提出了多重邻接关系模型。该模型能够分析子群组间的全局一致性,进而通过合并具有较高一致性的子群组得到全局集体行为模式。在多个真实视频数据集上的实验结果表明, FDA 算法相比于已有方法具有更高的识别精度。

**关键词** 集体行为;多重交互;流密度;密度峰值

中图法分类号 TP18 DOI 号 10.11897/SP.J.1016.2017.02519

## A Flow Density Based Algorithm for Detecting Coherent Motion with Multiple Interaction

WU Yun-Peng ZHAO Chen-Yang SHI Zeng-Lin YE Yang-Dong

(School of Information Engineering, Zhengzhou University, Zhengzhou 450052)

**Abstract** As one of the most common phenomenon in nature, coherent motion widely exists in different crowd systems such as pedestrians, cells, bacterial colonies and interactive physical particles. Detecting coherent motion in crowd from video sequences is an active area of research in pattern recognition community, and has received increasing attentions in recent years. Generally, the clusters of coherent motion provide a mid-level representation of crowd dynamics, and could be used for high-level semantic analysis such as scene understanding, abnormal detection and activity recognition. Even more important, detecting coherent motion has many practical applications. For example, people always gather at some places of security interests, such as shopping malls, train stations, sports events and public square. Automatically detecting the crowd behavior in video surveillance of such crowded environment can be used to develop crowd management strategies and assist the movement of the crowd or individuals, to avoid the crowd disasters and ensure the public safety. In addition, traffic pattern recognition and traffic flow estimation, which provide valuable information for the transportation department, play important roles in traffic management. Basically, existing methods for coherent motion detection can be divided into two categories: the LK optical flow tracking-based methods and KLT tracker-based methods. These two methods achieve detection by segmenting optical flow field and analyzing individuals' topological structure, respectively. Despite the extensive research efforts spent on

this problem, it remains a challenging problem due to the complex interactions between individuals in crowded scenes. To this end we propose a Flow Density based Algorithm (FDA) for detecting coherent motion with multiple interactions. In FDA, the multiple interactions inherent in coherent motion are described and analyzed, which are widely ignored in the literature. FDA detects coherent motion from local and global level. At the local level, to detect the consistency among local individuals, we define a flow density by regarding coherent motion as fluid motion. Then a flow density peaks clustering is introduced to detect the subgroups with strong local consistency with the merits of single step of clustering process, the automatic decision of clustering centers and the accurate identification of outliers. At the global one, the multiple interactions among subgroups are described by a multiple adjacency model. Considering the connected components in graph theory, we define a global consistency to describe the complex interaction between collective subgroups. Based on this, global consistency of subgroups can be quantified. Then we achieve coherent motion detection by merging the collective subgroups according to the degree of global consistency. Experimental results on several benchmark datasets (Collective Motion Dataset and Bacterial Colony Dataset) including the scenes of pedestrians, traffic and bacterial colony demonstrate that FDA has better detection performance with comparative efficiency, compared with the existing methods. In order to further show the remarkable performance of the proposed method, we compare the number of the coherent motions obtained respectively by our methods and other state-of-the-art methods in several common seen scenes. As expected, the proposed method achieves better results than the existing methods in agreement with other experimental results.

**Keywords** coherent motion; multiple interaction; flow density; density peaks

## 1 引言

集体行为是指运动场景中多个运动个体所呈现的一致性行为模式,在自然界和人类生活之中广泛存在(如图 1 所示)。视频场景中集体行为的识别有着重要的科研意义和应用价值。例如,生物群的研究使人类更加科学地认知生物群体智能并加以运用<sup>[1-2]</sup>;对视频监控中行人流进行分析对于公共安全有着重要的意义<sup>[3]</sup>;道路监控中车流的识别有利于交通资源配置的优化<sup>[4]</sup>等。因此,在计算机视觉领域,视频中集体行为的识别受到越来越多的关注,成为重要的研究内容之一。



图 1 自然界中的集体行为

基于视频中提取的运动个体跟踪点,集体行为识别是根本任务,是求解跟踪点的模式划分,使得每个划分内的跟踪点具有较高的行为一致性。以此为目标,计算机视觉领域的研究者相继提出了很多集体行为识别方法。这些方法中,一类以 Lucas-Kanade 光流法(LK)<sup>[17]</sup>作为跟踪策略,以行为一致性为约束,将光流粒子划分至不同区域<sup>[5-10]</sup>;另一类方法通过跟踪运动区域的 Kanade-Lucas-Tomasi 粒子(KLT)<sup>[18]</sup>间的一致性交互,识别集体行为模式<sup>[11-16]</sup>。

已有工作均忽略了对视频场景集体行为中多重交互关系的描述和分析,使得识别出的集体行为模式仍存在局限性。一般地,集体行为中跟踪点的位置分布散乱、密度分布不均,使得从跟踪点的层面描述和分析集体行为中存在的复杂交互关系十分困难。进一步研究发现,局部某些行为一致性较强的跟踪点会以子群组的形式存在于集体行为模式中。本文从子群组层面出发,对复杂交互关系进行深入的描述和分析。子群组间的交互中,有一致性交互,例如,分流交互、合并交互,还有不一致性交互,例如,穿插交互、对流交互(如图 2 所示)。这些交互关系呈现出多重性,具体体现在两个方面:(1)多个,形状各异

的子群组间通常存在多个交互区域; (2) 多样, 交互区域内跟踪点的行为一致程度具有多样性. 这种存在多重交互关系的集体行为识别主要存在以下难点: (1) 具有行为一致性的集体行为局部子群组形状各异, 且其内部跟踪点分布散乱. 如何根据跟踪点间的行为一致性程度, 将一致性程度较高的跟踪点划分至局部子群组是亟待解决的问题; (2) 具有较高全局一致性的子群组呈现出集体行为的全局模式. 如何通过描述和分析多重交互关系, 呈现出子群组间全局一致性是难点.

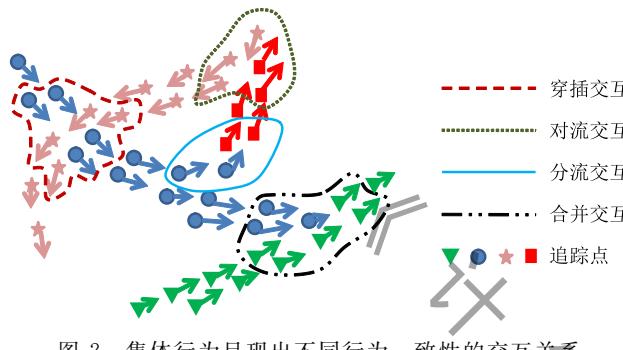


图 2 集体行为呈现出不同行为一致性的交互关系

本文提出基于流密度的多重交互集体行为识别算法 FDA(A Flow Density based Algorithm for Detecting Coherent Motion with Multiple Interaction), 能够实现局部和全局集体行为的识别. 在局部, 为了准确地衡量个体间的行为一致性程度, 基于流体力学提出流密度的定义. 进而提出基于流密度的聚类算法将跟踪点划分为局部子群组; 全局上, 为了精确描述集体行为客观存在的多重交互关系, 提出多重邻接关系模型. 该模型能够分析子群组间的全局一致性, 进而通过合并具有较高一致性的子群组得到最终的集体行为模式. 在真实视频数据集的实验结果表明, 相比于已有方法, 本文中提出的方法在保证识别效率的前提下, 进一步提升了集体行为的识别精度.

与国内外同类研究工作相比, 本文贡献点总结如下: (1) 基于流体力学定义的流密度, 揭示了跟踪个体间的行为一致性程度, 能够有效地识别局部一致性较强的子群组; (2) 首次对集体行为中客观存在的多重交互关系进行了描述和分析, 提出了多重邻接关系模型, 该模型能够有效识别具有全局一致性的集体行为模式; (3) 提出了多重交互集体行为识别算法. 在多个真实视频数据集上的实验结果表明该算法具有较高的识别精度.

## 2 相关工作

不同科研领域研究者都广泛地关注存在于各种

群体系统中的集体行为. 已有工作分别从产生机理、演化过程、运动建模等方面对集体行为展开相关研究, 取得了阶段性成果. 在生物学中, 研究人员通过实证方法揭示不同群体中集体行为的产生机理<sup>[21-22]</sup>; 物理学家将群体看作运动粒子, 通过相关理论为集体行为建立粒子运动模型<sup>[19-20]</sup>; 社会学研究表明, 群体中的个体受到周围具有行为一致性其它个体的影响, 会呈现出跟随行为<sup>[23-24]</sup>.

近年来, 视频中集体行为识别的研究已成为计算机视觉领域的重要分支. 已有识别方法根据跟踪点提取算法的不同可以分为两类: 基于 LK 光流法<sup>[17]</sup>和基于 KLT 跟踪粒子<sup>[18]</sup>. 基于 LK 光流法的识别方法提取光流跟踪点, 通过分割场景区域实现集体行为的识别. 文献[8]通过对连续视频帧跟踪点的分析识别行为模式, 但仅能识别简单场景中显著的个体行为; 文献[6]通过分割行为光流场实现集体行为的识别, 但需要人工指定行为模式的数目; 文献[5]提出拉格朗日粒子模型, 分割场景的光流粒子, 识别具有高度流动稳定性的集体行为; 文献[7]基于拉格朗日粒子模型, 提出面向群体流的脉线表示法. 然而, 场景中的群体行为模式较为复杂时, 光流跟踪噪声会导致识别出的行为模式精度较低. 文献[9]通过行为光流场中局部转化域的分割识别集体行为模式; 文献[10]基于热力扩散方法建立集体行为光流场, 通过分层聚类算法识别行为模式. 虽然构建更加精细的光流场可以一定程度上提升识别精度, 但这些方法在识别多重交互集体行为方面仍存在局限性.

另一类方法使用 KLT 跟踪粒子表示行为跟踪点. 位置分布杂乱无章, 密度各异是这种跟踪点的显著特征, 但跟踪点规模比光流法有所减少. 文献[11]通过行人效能的方法对局部群体行为呈现出的统计特性进行分析; 文献[12]提出一种自底向上的分层聚类方法实现识别局部行为相似的小群组; 文献[13]提出近邻不变性的概念, 通过基于图论的聚类算法划分集体模式. 文献[14-15]基于集体行为中行为一致性程度的度量实现集体行为识别. 然而如果场景中存在极低密度的情况, 上述方法对于全局集体行为的识别仍存在不足. 文献[16]提出的一致性密度聚类算法虽然在全局层面集体行为的识别方法有所改善, 但仍无法有效地识别场景中的多重交互关系集体行为.

## 3 多重交互集体行为识别

本文针对集体行为在局部和全局层面客观存在

多重交互关系这一问题,提出基于流密度的多重交互集体行为识别算法 FDA. 一方面,受流体力学启发,定义的流密度能够准确地衡量个体间的行为一致性程度. 进而局部具有行为一致性子群组的挖掘是通过基于流密度的聚类算法实现的; 另一方面,为了精确描述集体行为客观存在中的多重交互关系,提出多重邻接关系模型. 该模型能够分析子群组间的全局一致性,进而通过合并具有较高一致性的子群组得到最终的集体行为模式. 图 3 展示了 FDA 的算法流程示意图.

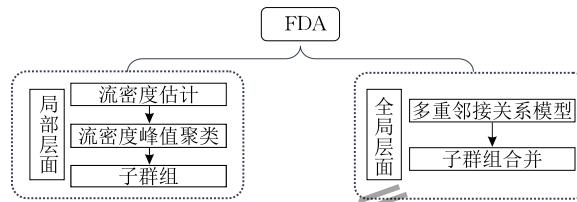


图 3 提出算法的流程图

### 3.1 流密度峰值聚类

视频中的集体行为呈现出典型的流体运动<sup>[25]</sup>, 内部个体由于受到视域的限制, 其运动会主要参考邻近个体的行为<sup>[15]</sup>. 这些具有近邻关系的个体将会形成子群组, 且组内个体间往往具有较高的行为一致性. 因此, 本文从子群组的层面出发理解集体行为模式中蕴含的流式结构. 然而, 形状各异和内部跟踪点分布杂乱无章为这种子群组的识别带来了困难. 一般地, 在待识别簇的形状各异的情况下, 基于密度的聚类算法具有优势<sup>[26]</sup>. 在视频场景中, 跟踪点的位置分布随机性很强, 没有统一的规律. 现阶段, 如何量化个体间的局部一致性仍然是开放性问题. 针对该问题, 本文从流式结构的角度出发, 提出一种流密度.

运动流体的基本元素为流体质点<sup>[27-28]</sup>. 这些质点在运动过程中, 会在相反方向上产生一种相互作用力——黏性应力. 该作用力对流体的滑动或变形起到一种阻碍的效果. 流体力学的概念指出, 黏度系数和剪切速率是影响黏性应力的主要因素. 黏度系数是指流体受外力作用运动时, 内部质点对于流动的阻尼. 例如, 由于在黏度系数方面油大于水, 因此受到外力作用时油的流动会明显慢于水. 剪切速率反映的是指流体质点在不同流层间的流速差异. 不同流层间的质点流速越一致, 剪切速率就越小, 则当前的流动较为稳定; 反之, 若流速差异较大, 剪切速率会增大, 就会出现乱流、漩涡的现象. 令  $\gamma$  表示黏性应力,  $\mu$  表示黏度系数,  $\lambda$  表示剪切速率, 则有

$$\gamma = \mu \lambda \quad (1)$$

从上述黏性应力的定义可以看出, 较小的黏性应力表示流体质点间的相互作用较弱, 也就意味着质点的运动较为一致. 因此, 黏性应力与质点间运动的一致性成反比关系.

集体行为中跟踪点可以等价为流体中的质点, 而集体行为就可以视为具有较强稳定性的流体. 因此, 跟踪点间的行为一致性程度可以通过求解“黏性应力”获得, 其中“黏度系数”表示跟踪点在位置上的相对关系, 而“剪切速率”则表示跟踪点在运动方向上的差异. 在视频场景所在的二维平面  $R^2$  上, 令  $C$  表示跟踪点的集合, 任意跟踪点  $c_i (i=1..N)$  的位置和运动方向分别记作  $P_i(p_i^x, p_i^y)$  和  $O_i(o_i^x, o_i^y)$ . 下面分别给出“黏度系数”、“剪切速率”和“黏性应力”在集体行为中的定义.

**定义 1.** 黏度系数. 跟踪点对  $(c_i, c_j)$  间“黏度系数”定义为

$$\mu(i, j) = d_{pos}(i, j) / Thr_{dis} \quad (2)$$

其中,  $d_{pos}(i, j) = \sqrt{(p_i^x - p_j^x)^2 + (p_i^y - p_j^y)^2}$  表示任意点对  $(c_i, c_j)$  间的距离,  $Thr_{dis}$  是距离阈值. 特别地, 若  $d_{pos}(i, j) < Thr_{dis}$ , 则跟踪点在位置方面的相对关系更为紧密.

**定义 2.** 剪切速率. 跟踪点对  $(c_i, c_j)$  间“剪切速率”定义为

$$\lambda(i, j) = e^{\frac{d_{ori}(i, j)}{Thr_{ori}} - 1} \quad (3)$$

其中  $d_{ori}(i, j) = |\arccos(o_i^x \cdot o_i^y / |o_i^x| \cdot |o_i^y|) - \arccos(o_j^x \cdot o_j^y / |o_j^x| \cdot |o_j^y|)|$  表示任意点对  $(c_i, c_j)$  间的运动方向差异,  $Thr_{ori}$  是方向阈值. 尤其是  $d_{ori}(i, j)$  小于  $Thr_{ori}$  时, 点对  $(c_i, c_j)$  之间运动方向的差异更为显著.

**定义 3.** 黏性应力. 任意点对  $(c_i, c_j)$  之间的“黏性应力”定义为

$$\gamma(i, j) = \mu(i, j) \lambda(i, j) \quad (4)$$

得到的“黏性应力”是阻止流体相对运动的力, 其与跟踪点的行为一致性成反比关系, 而流密度是行为一致性的综合体现, 因此流密度定义如下.

**定义 4.** 流密度. 任意跟踪点  $c_i$  的流密度定义为

$$\rho_i = \sum_{j=1..N, j \neq i} 2\pi e^{-\frac{1}{2}(\gamma(i, j))^2} \quad (5)$$

由式(5)可知, 跟踪点  $c_i$  的流密度综合考虑了  $c_i$  与其它所有点在“行为一致性”方面关联程度. 因此, 面对集体行为中跟踪点位置分布散乱的情况, 流密

度更加有利于后续聚类算法有效地识别集体行为子群组.

集体行为子群组通常由局部行为一致性较强的跟踪点构成, 形状各异是子群组的典型特征. 在定义流密度的基础上, 基于密度的聚类算法能够适用于识别形状不规则的集体行为子群组. 文献[26]提出了密度峰值聚类算法, 其核心思想是通过辨别质心与非质心点在密度分布上呈现出的不同特征, 进而能够确定质心, 实现聚类分配. 然而, 该算法需要人工确定质心且仅能识别边缘噪声点, 无法解决实时视频中的集体行为识别问题. 针对该问题, 文献[16]进一步提出的聚类算法通过分析不同集体行为子群组的行为特征差异自动确定质心, 并实现了集体行为跟踪点中类内和边缘噪声点的过滤. 但是, 提出的一致性密度忽略了集体行为的流体特性, 而密度估计是影响后续质心确定和非质心点分配的重要因素. 因此, 该算法在集体行为子群组识别方面仍有局限性. 与一致性密度相比, 定义的流密度具有以下优点: (1) 从流体的角度理解集体行为, 定义的流密度更能体现跟踪点在流式结构中的行为一致性; (2) 一致性密度仅粗略地统计了近邻中行为一致的跟踪点个数, 而流密度是建立在流体力学中“黏性应力”的基础上, 通过行为特征的精细度量反应跟踪点的局部行为一致性. 上述优点决定了流密度能够在跟踪点分布散乱的情况下有效地识别子群组. 本文进一步提出了流密度峰值聚类算法, 图 4 展示了算法的核心流程.

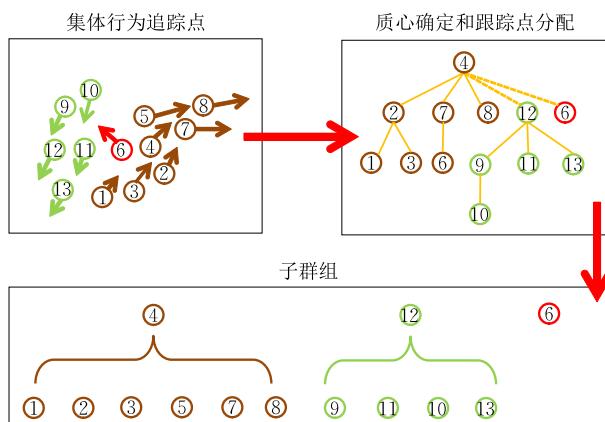


图 4 流密度峰值聚类算法中数据点的指派过程示意图  
(棕色和绿色是由正常的跟踪点构成的子群组,  
红色为噪声点)

流密度峰值聚类算法需解决的首要问题是确定子群组质心. 通过对集体行为深入分析, 本文认为集体行为子群组的质心有以下两方面特点. 一方面, 质心点往往具有较高的流密度, 且周围点密度相对较

低; 另一方面, 子群组质心点与具有更高流密度跟踪点相比, 两者的位置和运动方向通常具有较大差异. 通过上述分析, 本文给出质心自动确定策略. 对于任意跟踪点  $c_i$ , 分配向量  $father(i)$  表示其质心. 特别地, 若  $c_i$  本身就是质心点, 则  $father(i)=i$ . 在开始阶段, 将具有最大流密度值的所有跟踪点都初始化为质心点, 令跟踪点  $c_j$  表示任意具有密度最大值的点, 则  $father(j)=j$ , 而其他跟踪点  $father(i)=-1$ ,  $i \neq j$ . 同时, 令  $neigh(i)$  表示比  $c_i$  流密度高且距离  $c_i$  最近点的索引下标, 则

$$neigh(i) = \arg \min_{j, \rho_j > \rho_i} (d_{pos}(i, j)) \quad (6)$$

$neigh(i)$  表示的是由密度较低点指向密度较高点的连接, 为每个跟踪点提供了潜在的质心选择.

根据密度由高到低的顺序, 依次判定每个跟踪点的潜在质心选择是否合适, 即任意跟踪点  $c_i$  的质心为

$$father(i) =$$

$$\begin{cases} father(neigh(i)), & d_{ori}(i, neigh(i)) \leq Thr_{ori} \\ i, & \text{其他} \end{cases} \quad (7)$$

式(7)中, 若跟踪点  $c_i$  的运动方向与  $c_{neigh(i)}$  行为较一致, 则  $c_i$  的质心选择与密度较高的点  $c_{neigh(i)}$  保持一致; 否则, 若  $c_i$  与  $c_{neigh(i)}$  在运动方向上差异较大, 即两者不在同一个局部子群组中, 因此  $c_i$  会成为新的质心. 通过该质心分配策略, 初始时的非质心点会依次分配给现有质心或者成为新质心, 从而实现所有跟踪点的聚类划分. 令  $T$  表示所有  $M$  个子群组集合, 即  $T = \{t_1, t_2, \dots, t_M\}$ , 其中  $t_\alpha = \{c_i | father(i) = \text{子群组 } \alpha \text{ 的质心}\}$ .

噪声点过滤也是集体行为子群组识别的关键问题. 集体行为中的噪声点主要来源于两方面, 跟踪误差造成的类内噪声点和密度极低区域的边缘噪声点. 在提出方法的聚类分配过程中, 无论类内还是边缘噪声点, 与潜在分配质心点  $neigh(\cdot)$  间存在较大的行为特征差异, 因此这些噪声点都会成为新的质心. 此外, 噪声点与其他跟踪点间的行为一致性较弱, 即噪声点通常具有极低的流密度, 因此其他点也不会选择噪声点作为其质心. 这时, 噪声点所在的子群组仅包含其自身. 通过剔除仅包含一个跟踪点的簇, 即可实现噪声点的过滤.

最终, 流密度峰值聚类通过确定质心、分配非质心点以及过滤噪声点, 准确地识别集体行为子群组.

### 3.2 多重邻接关系模型

流密度峰值聚类算法根据跟踪点的行为一致性

程度,将这些点划分至不同的子群组.然而,最终的集体行为识别需要进一步考虑跟踪点间在全局层面呈现出的行为一致性,其中子群组间的多重交互关系是衡量全局一致性的重要因素.对多重交互关系的描述和分析存在以下两方面难点.首先,子群组间交互关系是多样的,例如穿插、对流、分流、合并,其中的关键问题是如何有效地对多重交互关系进行描述和分析;其次,子群组间交互不是单一的,而是可能会同时存在多种交互关系,这对于分析多重交互关系提出了更高的要求.针对上述问题,本文通过构建多重邻接关系模型用于描述子群组间的多重交互关系,进而分析子群组间的全局一致性,将具有较高一致性的子群组合并,实现最终的集体行为识别.

处于不同子群组内的跟踪点存在邻接性是多重交互关系形成的必要条件,因此,首先对跟踪点邻接性进行定义.

**定义 5.** 邻接性. 令  $\mathbf{W}$  表示邻接性矩阵,则任意两个跟踪点  $c_i, c_j$  的邻接性定义为

$$W(i, j) = \begin{cases} 1, & d_{pos}(i, j) \leq Thr_{dis} \\ 0, & d_{pos}(i, j) > Thr_{dis} \end{cases} \quad (8)$$

式(8)中,跟踪点  $c_i, c_j$  间的距离小于或等于距离阈值  $Thr_{dis}$  即表示它们是邻接的.

处于不同子群组  $t_\alpha$  和  $t_\beta$  的两个跟踪点  $c_i$  和  $c_j$ ,  $c_i \in t_\alpha$  且  $c_j \in t_\beta$ , 如果  $c_i$  和  $c_j$  是邻接的  $W(i, j) = 1$ , 则认为子群组  $t_\alpha$  和  $t_\beta$  存在交互关系. 本文使用邻接关系图  $G_{\alpha\beta}(V_{\alpha\beta}, E_{\alpha\beta})$  表示这种交互关系,  $V_{\alpha\beta} = \{c_i\} \cup \{c_j\}$ , 表示  $t_\alpha$  和  $t_\beta$  中所有具有邻接性的跟踪点集合;  $E_{\alpha\beta} = \{\langle c_i, c_j \rangle\}$ , 是邻接点间的边集合. 在图  $G_{\alpha\beta}(V_{\alpha\beta}, E_{\alpha\beta})$  中, 子群组间往往存在着多个邻接区域, 每个邻接区域都体现了某种交互关系, 且每个区域内的点是连通的, 区域间的点不连通. 这样的邻接区域可以表示为连通分量  $g_k(v_k, e_k)$ ,  $v_k$  和  $e_k$  分别是邻接区域内的点和边. 令  $S_{\alpha\beta}$  表示所有连通分量的集合, 则有  $S_{\alpha\beta} = \{g_k(v_k, e_k) | \bigcup_k g_k = G_{\alpha\beta}, g_k \cap g_l = \emptyset\}$ , 其中  $g_k$  为  $G_{\alpha\beta}$  的第  $k$  个连通分量. 因此, 求解图  $G_{\alpha\beta}$  中的所有连通分量可以对多重交互关系进行有效地描述, 如图 5 所示. 基于上述分析, 本文给出多重邻接关系模型的定义.

**定义 6.** 多重邻接关系模型. 使用五元组  $\{T, \mathbf{W}, G_{\alpha\beta}, S_{\alpha\beta}, f\}$  表示多重邻接关系模型, 其中每个元素分别为

(1) 子群组集合  $T = \{t_1, t_2, \dots, t_M\}$ ;

(2) 邻接性矩阵  $\mathbf{W}$ ;

(3) 邻接关系图  $G_{\alpha\beta}$ ;

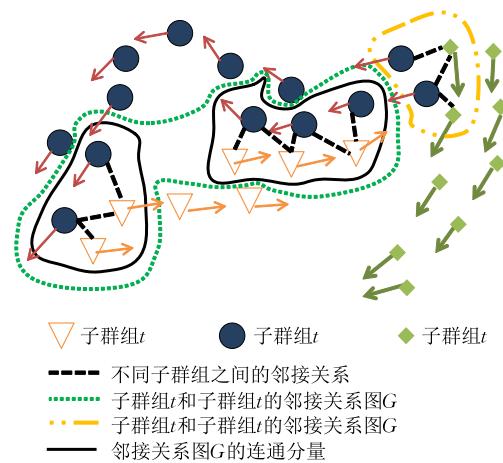


图 5 连通分量与多重交互关系间关联性的示意图

(4) 连通分量的集合  $S_{\alpha\beta}$ ;

(5) 求解函数  $f, S_{\alpha\beta} = f(\{t_\alpha, t_\beta\}, \mathbf{W}, G_{\alpha\beta})$ .

提出模型根据邻接性矩阵  $\mathbf{W}$  和邻接关系图  $G_{\alpha\beta}$ , 通过函数  $f$  求解子群组  $t_\alpha$  和  $t_\beta$  对应邻接关系图  $G_{\alpha\beta}$  的连通分量集合  $S_{\alpha\beta}$ . 得到的连通分量集合  $S_{\alpha\beta}$  就是子群组  $t_\alpha$  和  $t_\beta$  间多重交互关系的有效描述, 为进一步分析子群组间的全局一致性奠定了基础.

不同交互关系内跟踪点的行为一致性是分析子群组间全局一致性的重要因素. 多重交互关系按照行为一致性的不同可以分为两类. 一类反映的是子群组间具有全局一致性, 如分流、合并; 另一类则呈现出子群组间的行为差异, 如穿插、对流. 提出的多重邻接关系模型从连通分量的角度给出了多重交互关系的有效描述. 后续的关键步骤是基于这种描述, 在全局层面度量子群组间的行为一致性.

给定任意两个子群组  $t_\alpha$  和  $t_\beta$ , 对于连通分量集合  $S_{\alpha\beta} = \{g_1(v_1, e_1), g_2(v_2, e_2), \dots, g_k(v_k, e_k)\}$  中任意的连通分量  $g_i(v_i, e_i)$ , 令  $v_i^\alpha, v_i^\beta$  分别表示  $g_i(v_i, e_i)$  中跟踪点的子集且  $v_i^\alpha \subset t_\alpha, v_i^\beta \subset t_\beta$ , 函数  $AVR(v_i)$  表示  $g_i(v_i, e_i)$  中所有跟踪点的平均运动方向, 则子群组间的全局一致性度量函数为

$$GC(\alpha, \beta) = \min_i (|AVR(v_i^\alpha) - AVR(v_i^\beta)|) \quad (9)$$

从函数  $GC$  中可以看出, 在子群组  $t_\alpha$  和  $t_\beta$  间的所有连通分量中, 选取行为差异最小的连通分量作为子群组间的全局一致性度量. 而子群组的合并取决于是否存在连通分量的全局一致性度量满足  $GC \leq Thr_{ori}$ . 通过合并具有全局一致性的子群组, 提出的多重邻接关系模型能够得到最终的集体行为模式.

### 3.3 多重交互集体行为识别算法 FDA

#### 3.3.1 FDA 算法描述

FDA 算法分为两个阶段. 第一阶段是流密度峰

值聚类识别具有局部一致性的子群组,第二阶段基于多重邻接关系模型合并具有全局一致性的子群组. 具体算法描述如下所示:

### 算法 1. FDA 算法.

输入: 跟踪点集合  $C$  及其位置和运动方向  $P, O$

输出: 集体行为模式  $T^{final}$

/\* 局部层面 \*/

1. 通过式(1)~(7),计算流密度  $\rho$

2. 初始化向量  $father, neigh$

3. FOR ALL  $c_i \in C$  //按  $\rho$  的降序处理

4. 根据式(9)更新分配向量  $father$

5. END FOR

6. IF  $\exists \sum (father == i) = 1$  THEN  $father(i) = -1$

7. 由向量  $father$  得到子群组划分  $T = \{t_1, t_2, \dots, t_M\}$ ,  
其中子群组  $t_\alpha = \{c_i | father(i) = \text{第 } \alpha \text{ 个子群组}\}$

/\* 全局层面 \*/

8. 根据式(10)计算邻接矩阵  $W$ ,将  $T^{final}$  初始化为  $T$

9. FOR  $\alpha = 1$  TO  $M-1$

10. FOR  $\beta = \alpha+1$  TO  $M$

11. 对于子群组  $t_\alpha$  和  $t_\beta$ ,构建邻接关系图  $G_{\alpha\beta}$

12. 基于多重邻接关系模式求解连通分量集合  $S_{\alpha\beta}$

13. IF  $GC(\alpha, \beta) \leq Thr_{ori}$  THEN //合并子群组

14.  $T^{final} = T^{final} - \{t_\alpha, t_\beta\}$

15.  $t_{new} = t_\alpha \cup t_\beta$

16.  $T^{final} = T^{final} \cup \{t_{new}\}$

17. END IF

18. END FOR

19. END FOR

20. RETURN  $T^{final}$

算法仅涉及两个参数: 距离阈值  $Thr_{dis}$  和方向阈值  $Thr_{ori}$ . 距离阈值  $Thr_{dis}$  主要用来判定跟踪点间的邻接性, 从而影响多重交互关系的描述和分析. 方向阈值  $Thr_{ori}$  是局部行为的一致性在运动方向上的约束, 能够影响局部子群组的形成. 通过大量实验, 本文给出了参数的经验性取值.  $Thr_{dis}$  采用所有跟踪点间的距离  $d_{pos}$  中第  $k$  小元素的值,  $k = N_{pair} \times 2.0\%$ , 其中  $N_{pair} = N \times (N-1)/2$ .  $Thr_{ori}$  的取值为  $\pi/3$ . 第 4 节会给出具体的参数对比实验及相关分析.

### 3.3.2 时间复杂度分析

FDA 算法涉及局部和全局两个层面. 一方面, 流密度估计和局部子群组识别的时间复杂度为  $O(N^2)$ ,  $N$  为点的数量; 另一方面, 构建多重邻接关系模型, 进而根据全局行为一致性合并子群组得到最终的集体行为模式. 由于连通分量求解的时间复杂度为  $O(V+E)$  ( $V$  和  $E$  分别点和对应连接边的数量), 因此分析任意两对子群组(共  $M^2$  对)间全局一

致性的时间复杂度为  $O(M^2(V+E))$ , 其中  $M$  为子群组个数. 又因为所有邻接点和对应边数量存在近似关系  $M(V+E) \approx N$ , 且满足  $M \ll N$ , 因此全局层面的时间复杂度  $O(M^2(V+E)) \ll O(N^2)$ . 综上所述, 提出算法的整体时间复杂度为  $O(N^2)$ .

## 4 实验结果与分析

本文通过视频监控数据集<sup>[14-15]</sup>和细菌群落数据集<sup>[1]</sup>验证 FDA 算法的有效性, 并对实验结果进行了深入地分析.

### 4.1 实验数据集

视频监控: 该数据集包含 413 视频序列, 其中每个片段有 100 帧图片. 视频内容涉及室内外人群、车流等真实监控场景.

细菌菌落: 该数据集中视频的内容是芽孢杆菌在培养基上生长过程, 其中细菌游动也是典型的集体行为场景.

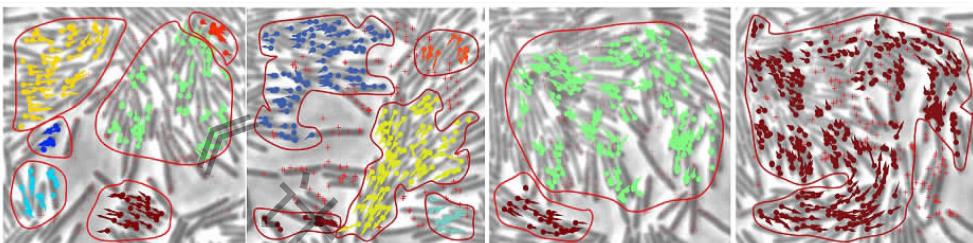
选取的实验数据集具有以下特点: (1) 包含人群、车流以及细菌等不同群体行为场景; (2) 不同的场景中存在群体间交互的一致性程度不同, 例如商场内自由走动人群、十字路口穿插而过的行人、马拉松比赛中按照预设赛道跑动的参赛者、道路上有序的车流、培养基中游动的细菌等.

### 4.2 实验结果

本文分别在视频监控和细菌群落两个数据集上进行了集体行为的识别实验. FDA 算法通过 gKLT (generalized KLT)<sup>[15]</sup> 跟踪策略提取视频场景中个体行为特征. 算法的参数  $Thr_{dis}$  采用所有跟踪点间的距离  $d_{pos}$  中第  $k$  小元素的值,  $k = N_{pair} \times 2.0\%$ , 参数  $Thr_{ori}$  的取值为  $\pi/3$ . 图 6(a) 是 FDA 算法在视频监控数据集上的识别效果, 图 6(b) 是细菌菌落数据集中具有代表性的四帧识别结果. 从图中可以看出, FDA 算法在不同场景、不同个体、不同交互的情况下都能有效地识别出正确的集体行为. 为了进一步验证 FDA 算法对于复杂交互场景的有效性, 本文给出了两个典型场景的实验: (1) 异常个体穿越人群; (2) 行人流在路口交汇, 如图 7 所示. 图 7(a) 给出了异常行为个体穿越人群的三个阶段. 在场景中, 异常个体与正常群体的规模存在较大的不平衡性, 且其与正常群体中的多个个体存在交互关系, 因此进一步增加了集体行为的识别难度. 实验结果说明了提出算法的有效性: (1) 异常个体在穿越正常人群的过程中, 始终与周围正常个体存在着复杂的交

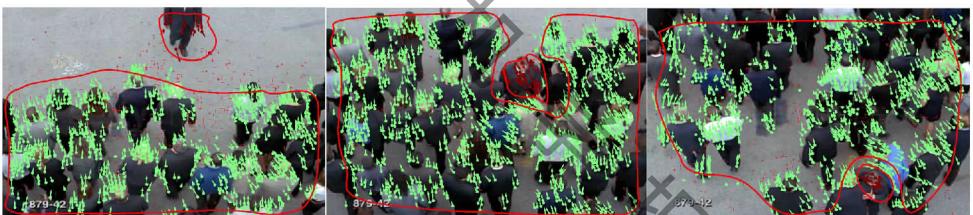


(a) 视频监控数据集



(b) 细菌菌落数据集

图 6 FDA 算法在不同数据集上的识别结果(不同集体行为模式使用红色实线圈出, 噪声点用红色十字表示)



(a) 异常个体穿越正常人群(深红色跟踪点为异常个体对应的跟踪点, 绿色为正常群体跟踪点, 红色十字为噪声点)



(b) 行人穿越马路的三个阶段(相遇、交汇和分离)

图 7 FDA 算法在两个特殊视频场景的识别效果(红线圈出的跟踪点代表不同的集体行为模式, 红色十字表示噪声点)

互, 提出算法通过多重交互关系模型保证了其识别的准确性; (2) 识别效果中, 异常个体所对应的“小簇”与正常群体对应的“大簇”体现了本文算法对于簇规模不平衡问题的鲁棒性. 在图 7(b)中, 两个群组在人行道交汇, 包括三个阶段: 相遇、交汇和分离. 尤其是行人的交汇阶段, 群体行为呈现出更为复杂的交互关系. 实验结果同样验证了 FDA 算法的有效性. 此外, FDA 算法除了有效地识别出行人的集体行为模式之外, 同时准确地识别出后两个阶段中

出租车由行驶到静止的状态变化.

#### 4.3 实验分析

本节将提出的 FDA 算法与已有集体行为识别算法从识别效果和量化指标两个方面进行了对比分析, 这些算法包括一致性密度聚类 (Collective Density Clustering, CDC)<sup>[16]</sup>、一致性度量 (Collectiveness Measurement, CM)<sup>[14-15]</sup>、一致性过滤 (Coherent Filter, CF)<sup>[13]</sup>、拉格朗日粒子 (Lagrangian Particle, LP)<sup>[5]</sup> 和脉线表示法 (Streakline Represen-

tation, SR)<sup>[7]</sup>. 同时, 对提出算法不同参数设置对实验结果的影响进行了分析.

#### 4.3.1 FDA 与 CDC 的对比实验

为了说明 FDA 算法在识别精度方面的提升, 首先引入 FDA 算法和 Collective Density Clustering (CDC)<sup>[16]</sup> 算法的对比实验. 图 8 分别给出了 CDC、CDC+流密度、CDC+多重邻接模型和 FDA 在细菌菌落数据集中某一帧识别效果的对比. 在图 8(a) 中, CDC 的识别效果仍存在以下问题: (1) 左上角和右侧中部稀疏区域的子群组被识别为噪声点; (2) 红色和蓝色跟踪点、浅黄色和橘黄色跟踪点构成的全局集体行为识别仍存在局限性. 针对第一个问题, 在图 8(b) 中的 CDC+流密度有了进一步改善 (黑色圈出区域). 因此, 本文定义的流密度从流体的

角度出发, 基于“黏性应力”的概念理解集体行为中的运动个体, 能够更加准确地衡量个体间的局部行为一致性. 然而由于 CDC 中没有精细地考虑多重交互关系, 在图 8(b) 的全局一致性判定方面仍存在局限性. 图 8(c) 中给出了 CDC+多重邻接关系的实验结果, 其在全局集体行为的识别方面有了进一步改善 (黑色圈出区域), 然而由于 CDC 将某些极低密度区域的子群组识别为噪声点, 导致最终集体行为的识别效果仍存在局限性. 在图 8(d) 中, 提出的 FDA 算法分别从局部、全局两个层面都有所改善, 这说明了: (1) 流密度能够更好地反应局部一致性, 进而使算法能够得到更准确的局部子群组; (2) 多重邻接关系模型的提出能够给出更加精确的全局一致性判定, 从而准确地识别出集体行为.

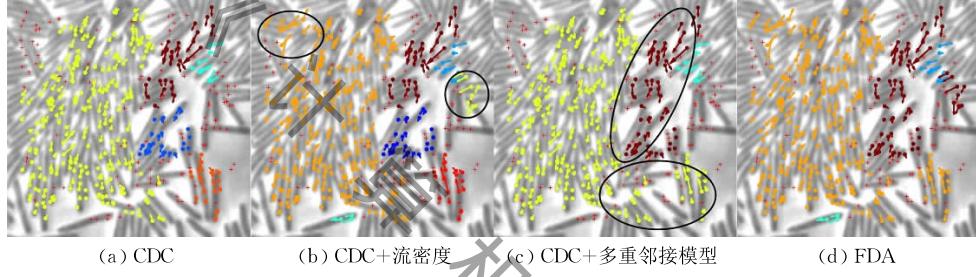


图 8 FDA 和 CDC 在细菌数据集上的对比实验

图 9 展示了 FDA 和 CDC 在细菌数据集上连续四帧的识别效果. 图中显著的流式结构以及多重交互关系使得 FDA 的识别准确性更高. 因此,

相比于 CDC, FDA 中流密度的定义以及多重邻接关系模型的构建都进一步提升了集体行为的识别精度.

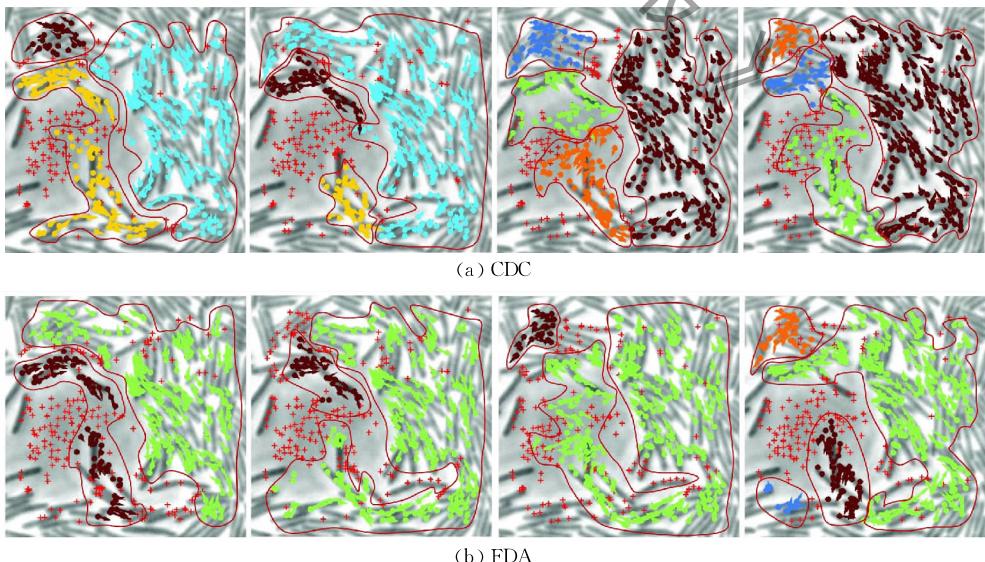


图 9 FDA 和 CDC 在连续四帧细菌运动场景上的识别结果对比

图 10 进一步给出了 FDA 和 CDC 在视频监控数据集上的对比实验. 场景一中白色圈出区域以及场景二中黑色圈出区域都进一步说明了 FDA 在行

人交互较为复杂的区域有着更为精确的识别效果.

#### 4.3.2 FDA 与其它集体行为识别方法的对比实验

为了充分说明 FDA 算法的有效性, 在视频监控

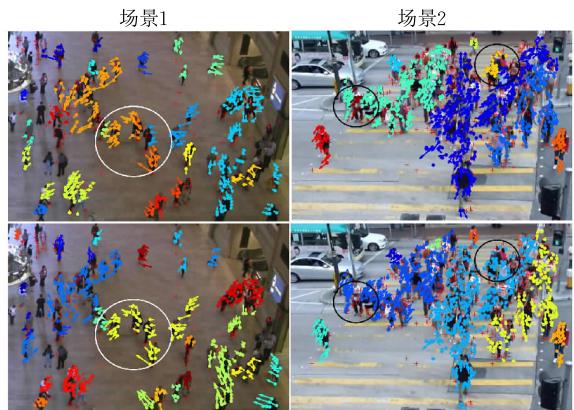


图 10 FDA 和 CDC 在视频监控数据集上的对比实验  
((a) FDA; (b) CDC)

数据集设计了 FDA 与一致性度量 (Collectiveness Measurement, CM)<sup>[14-15]</sup>、一致性过滤 (Coherent Filter, CF)<sup>[13]</sup>、拉格朗日粒子 (Lagrangian Particle, LP)<sup>[5]</sup> 和脉线表示法 (Streakline Representation, SR)<sup>[7]</sup> 的对比实验, 如图 11 所示. CM 和 CF 使用  $K$

邻近结构描述局部一致性. 在 CF 中, 由于参数  $K$  没有固定, 对比实验选取 10、15 和 25 三个值; CM 给出了参数  $K$  固定值 20. LP 和 SR 均按照论文给出的参数设置. 图 11 中, (a) 是场景 1 到场景 6 的集体行为真实标注, (b)~(h) 分别是 FDA 以及其它算法的识别效果. (c)~(e) 给出的是 CF 分别取  $K=10, 15, 25$  时的识别结果. 从图中可以看出, CF 仅能识别具有局部一致性的群组, 即使是  $K=25$  时. 在 (f) 中, CM 同样在识别具有全局一致性的集体行为方面存在局限性. (g) 和 (h) 是基于光流场的识别效果, 可以看出, 其对于较为复杂群体行为场景的识别结果与真实标注间差异较大. 因此, (b) 给出的 FDA 识别效果相比于上述几种方法是最优的.

图 12 进一步给出了 FDA 与相关工作在细菌数据集上的对比实验. 中间和右下角区域均有两股细菌群落汇集的现象, FDA 有效地识别出了具有这种交互关系的集体行为, 而相关工作均存在局限性.

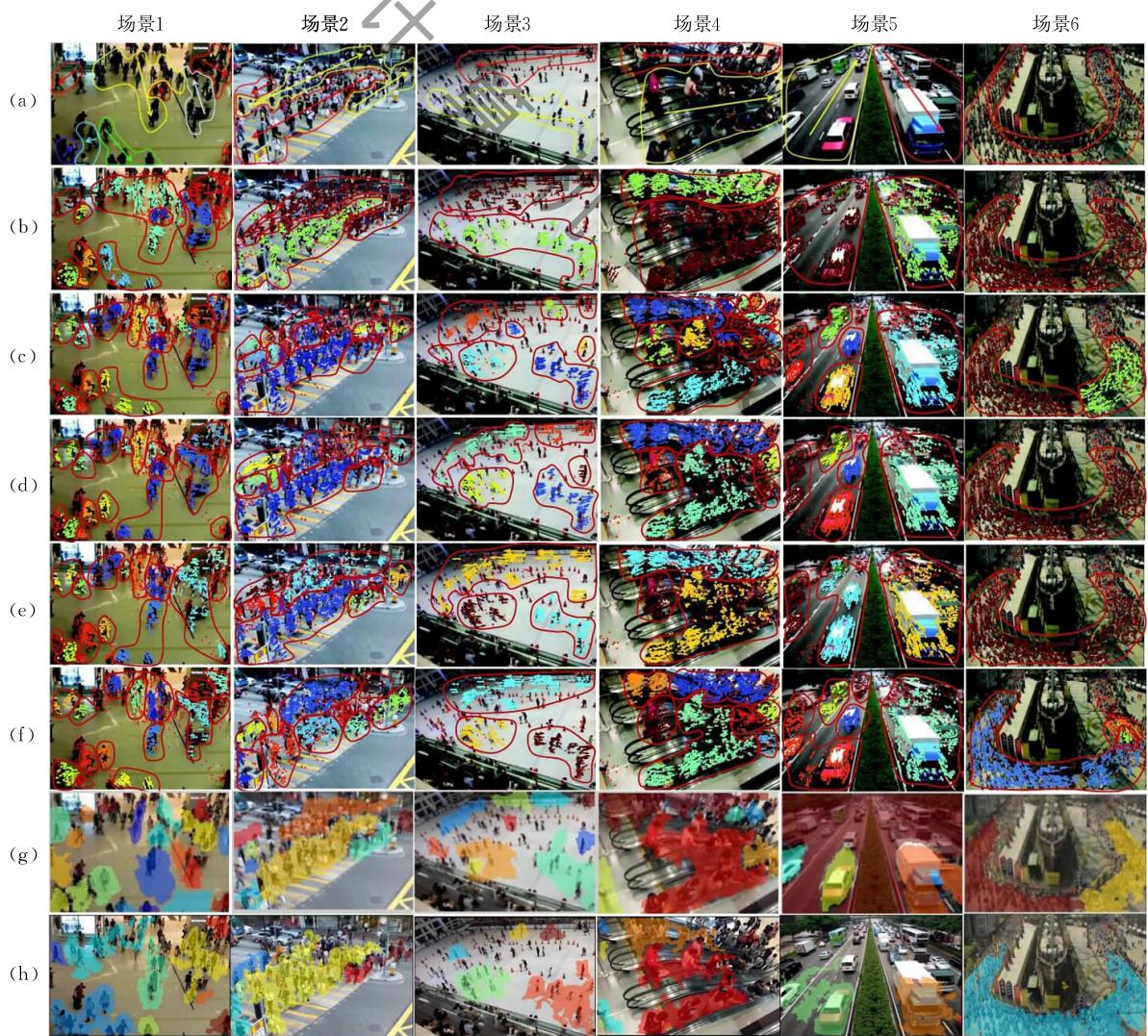


图 11 FDA 与其它识别方法在视频监控数据集上的对比实验

((a) 真实标注; (b) FDA; (c) CM; (d) CF,  $K=10$ ; (e) CF,  $K=15$ ; (f) CF,  $K=25$ ; (g) LP; (h) SR)

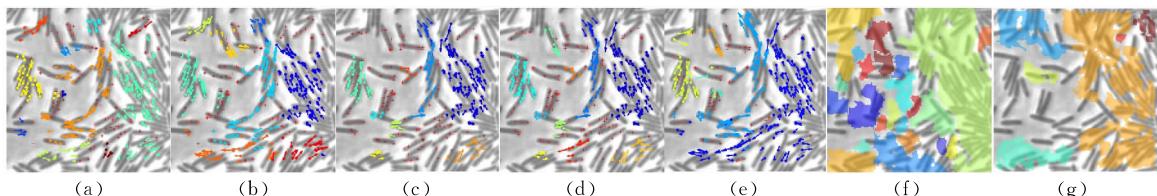


图 12 FDA 与其它识别方法在细菌菌落数据集上的对比实验  
((a) FDA; (b) CM; (c) CF,  $K=10$ ; (d) CF,  $K=15$ ; (e) CF,  $K=25$ ; (f) LP; (g) SR)

#### 4.3.3 量化指标分析

本文同样给出了量化指标的对比实验. 实验中, 集体行为簇的个数作为分析依据. 为此, 我们从视频监控数据集的 40 000 多帧中均匀地选取了 2031 帧场景(隔 20 帧选取 1 帧), 并为每帧人工标注了集体行为簇的真实个数. 如图 13 所示, 该柱状统计图展示了不同算法识别出的集体行为簇个数与真实标注的差异. 由图可见, FDA 在更多数场景的识别中簇数误差为 0, 其识别精度优于其它方法. 表 1 进一步给出了识别的簇个数与真实个数间差异值的期望  $E_{\text{cluster}}$ 、方差  $D_{\text{cluster}}$  以及每帧的平均识别时间这些指标的量化结果. FDA 在簇个数差值的期望、方差两项指标上都取得最小值. 从集体行为簇的个数方面表明 FDA 的识别精度更高, 且与人工标注真实值偏差的波动性更小. 在每帧的平均识别时间方面, FDA 虽然引入了多重邻接关系模型, 但在单帧识别方面仅比效率最高的CDC方法慢了 0.0194s. 因

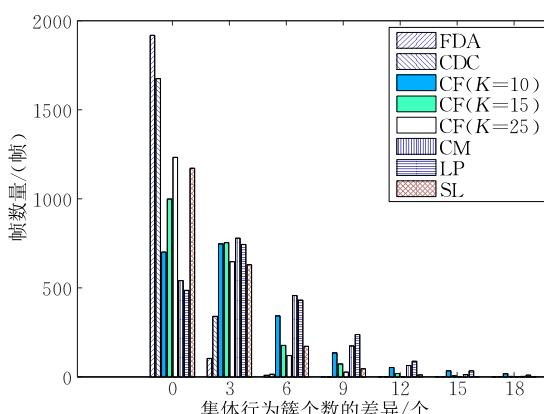


图 13 识别的簇个数与真实个数差异的柱状统计图(横坐标为各个算法识别出的簇个数与标注簇个数间的差异值, 纵坐标为在当前差异值范围内帧的个数)

表 1 集体行为簇个数差异(个)的期望与方差以及平均识别时间的对比

方法	$E_{\text{cluster}}$	$D_{\text{cluster}}$	识别时间/s
提出方法	<b>0.759</b>	<b>0.993</b>	0.1107
一致性密度聚类	0.807	1.043	<b>0.0913</b>
一致性过滤( $K=10$ )	3.543	12.583	0.3287
一致性过滤( $K=15$ )	2.287	6.271	0.4028
一致性过滤( $K=25$ )	1.604	3.423	0.4775
一致性度量	3.839	9.485	0.4202

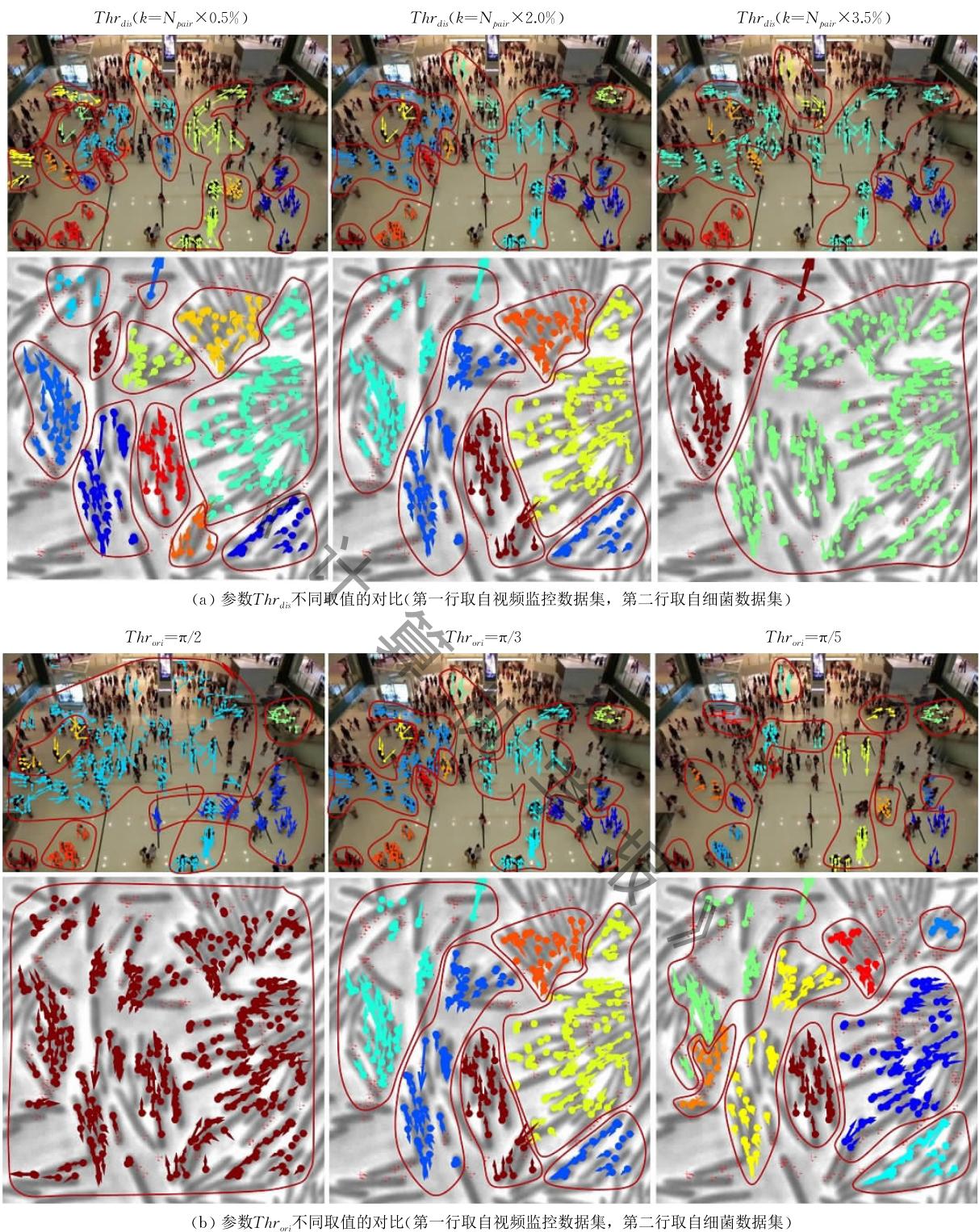
此, FDA 在保证效率前提下, 具有更优的集体行为识别精度.

#### 4.3.4 参数设置分析

FDA 算法涉及两个参数: 距离阈值  $Thr_{\text{dis}}$  和方向阈值  $Thr_{\text{ori}}$ . 本文给出了参数的经验性取值,  $Thr_{\text{dis}}$  设定为所有跟踪点对的距离中第  $k$  小的距离值且  $k = N_{\text{pair}} \times 2.0\%$  ( $N_{\text{pair}}$  为所有跟踪点对间距离的数量,  $N_{\text{pair}} = N \times (N-1)/2$ ),  $Thr_{\text{ori}} = \pi/3$ . 图 14 给出了不同参数取值的对比实验. 图 14(a) 中,  $Thr_{\text{ori}}$  的值固定为  $\pi/3$ , 计算  $Thr_{\text{dis}}$  时  $k$  的取值分别为  $N_{\text{pair}} \times 0.5\%$ 、 $N_{\text{pair}} \times 2.0\%$  和  $N_{\text{pair}} \times 3.5\%$ . 当  $Thr_{\text{dis}}$  取值过小时, 因为跟踪点邻接性的要求过于严格, 弱化了子群组间邻接关系, 最终难以识别全局一致性较强的集体行为. 当  $Thr_{\text{dis}}$  取值过大时, 跟踪点间邻接性的约束过于宽松, 导致子群组间出现过度合并的现象. 图 14(b) 中,  $Thr_{\text{dis}}$  计算中  $k$  值固定为  $N_{\text{pair}} \times 2.0\%$ ,  $Thr_{\text{ori}}$  的取值分别为  $\pi/2$ 、 $\pi/3$  和  $\pi/5$ . 当  $Thr_{\text{ori}}$  取值过大时, 局部一致性的约束过于宽松, 形成的子群组中会包含很多噪声点. 当  $Thr_{\text{ori}}$  取值过小时, 局部一致性的约束过于严格, 有些在合理的角度差异范围内的跟踪点被错误划分为噪声点. 因此, 通过大量实验,  $Thr_{\text{dis}}$  设定为所有跟踪点对距离中第  $N_{\text{pair}} \times 2.0\%$  小的距离值,  $Thr_{\text{ori}}$  的取值为  $\pi/3$ .

## 5 结 论

本文针对已有方法均忽略了集体行为中多重交互关系的描述和分析这一问题, 提出基于流密度的多重交互集体行为识别算法. 该算法能够有效地识别出视频中的局部和全局集体行为. 在局部集体行为识别方面, 通过引入流体力学的概念定义流密度, 其能够反映个体间局部行为一致性; 通过流密度峰值聚类进而识别出行为一致性较强的局部子群组. 在全局集体行为识别方面, 提出多重邻接模型对子群组间的多重交互关系进行有效描述, 进而分析交



(b) 参数  $Thr_{ori}$  不同取值的对比(第一行取自视频监控数据集, 第二行取自细菌数据集)

图 14 在视频监控和细菌数据集上参数  $Thr_{dis}$  和  $Thr_{ori}$  的对比分析

互关系反映出的全局一致性,通过合并具有全局一致性的子群组实现了集体行为的识别.在视频监控和细菌菌落两个数据集上的实验结果同时表明了提出算法的有效性.与已有一致性密度聚类、一致性过滤、一致性度量、拉格朗日粒子、脉线表示等方法

相比,提出算法在不失效率的前提下,进一步提升了集体行为的识别精度.

提出算法能够有效地识别出集体行为这种视频中群体的中层模式,为带有高层语义分析的领域具有重要的应用价值,例如异常行为检测、事件识别等.

## 参 考 文 献

- [1] Zhang H P, Avraham B E, Florin E L, Swinney H L. Collective motion and density fluctuations in bacterial colonies. *Proceedings of the National Academy of Sciences*, 2010, 107(31): 13626-13630
- [2] Couzin I. Collective minds. *Nature*, 2007, 445(7129): 715-715
- [3] Huang Kai-Qi, Chen Xiao-Tang, Kang Yun-Feng, Tan Tie-Niu. Intelligent visual surveillance: A review. *Chinese Journal of Computers*, 2015, 38(6): 1093-1118(in Chinese) (黄凯奇, 陈晓棠, 康运锋, 谭铁牛. 智能视频监控技术综述. *计算机学报*, 2015, 38(6): 1093-1118)
- [4] Menze M, Geiger A. Object scene flow for autonomous vehicles//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA, 2015: 3061-3070
- [5] Ali S, Shah M. A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA, 2007: 1-6
- [6] Min H, Ali S, Shah M. Learning motion patterns in crowded scenes using motion flow field//Proceedings of the International Conference on Pattern Recognition, Tampa, USA, 2008: 1-5
- [7] Mehran R, Moore B E, Shah M. A streakline representation of flow in crowded scenes//Proceedings of the European Conference on Computer Vision. Crete, Greece, 2010: 439-452
- [8] Brox T, Malik J. Object segmentation by long term analysis of point trajectories//Proceedings of the European Conference on Computer Vision. Crete, Greece, 2010: 282-295
- [9] Wu S, Wong H. Crowd motion partitioning in a scattered motion field. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2012, 42(5): 1443-1454
- [10] Wang W, Lin W, Chen Y, et al. Finding coherent motions and semantic regions in crowd scenes: A diffusion and clustering approach//Proceedings of the European Conference on Computer Vision. Zurich, Switzerland, 2014: 756-771
- [11] Kratz L, Nishino K. Going with the flow: Pedestrian efficiency in crowded scenes//Proceedings of the European Conference on Computer Vision. Florence, Italy, 2012: 558-572
- [12] Ge W, Collins R T, Ruback R B. Vision-based analysis of small groups in pedestrian crowds. *IEEE Transactions on Software Engineering*, 2012, 34(5): 1003-1016
- [13] Zhou B, Tang X, Wang X. Coherent filtering: Detecting coherent motions from crowd clutters//Proceedings of the European Conference on Computer Vision. Florence, Italy, 2012: 857-871
- [14] Zhou B, Tang X, Wang X. Measuring crowd collectiveness //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA, 2013: 3049-3056
- [15] Zhou B, Tang X, Zhang H, Wang X. Measuring crowd collectiveness. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(8): 1586-1599
- [16] Wu Y, Ye Y, Zhao C. Coherent motion detection with collective density clustering//Proceedings of the ACM Conference on Multimedia. Brisbane, Australia, 2015: 361-370
- [17] Bruhn A, Weickert J, Schnorr C. Lucas/Kanade meets horn/schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, 2005, 61(3): 211-231
- [18] Tomasi C, Kanade T. Detection and tracking of point features. *Technical Report*, 1991, 9(21): 9795-9802
- [19] Vicsek T, Czirók A, Ben-Jacob E, et al. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 1995, 75(6): 1226-1229
- [20] Vicsek T, Zafeiris A. Collective motion. *Physics Reports*, 2012, 517(3): 71-140
- [21] Couzin I D. Collective cognition in animal groups. *Trends in Cognitive Sciences*, 2009, 13(1): 36-43
- [22] Moussaid M, Garnier S, Theraulaz G, Helbing D. Collective information processing and pattern formation in swarms, flocks, and crowds. *Topics in Cognitive Science*, 2010, 1(3): 469-497
- [23] Le Bon G. *The crowd: A study of the popular mind*. London, England: T Fisher Unwin, 1908
- [24] Raafat R M, Nick C, Chris F. Herding in humans. *Trends in Cognitive Sciences*, 2009, 13(10): 504-504
- [25] Hughes R L. The flow of human crowds. *Annual Review of Fluid Mechanics*, 2003, 35(1): 169-182
- [26] Alex R, Alessandro L. Clustering by fast search and find of density peaks. *Science*, 2014, 344(6191): 1492-1496
- [27] Panton R L. *Incompressible Flow*. New York City, United States: John Wiley & Sons, 2013
- [28] Kirby B J. *Micro and Nanoscale Fluid Mechanics: Transport in Microfluidic Devices*. Cambridge, England: Cambridge University Press, 2010



**WU Yun-Peng**, born in 1987, Ph. D. His research interests include machine learning and computer vision.

**ZHAO Chen-Yang**, born in 1990, M. S. candidate. His research interests focus on video behavior analysis.

**SHI Zeng-Lin**, born in 1992, M. S. candidate. His research interests focus on machine learning.

**YE Yang-Dong**, born in 1962, Ph. D., professor, Ph. D. supervisor. His research interests include machine learning, knowledge engineering and intelligent systems.

## Background

Coherent motion detection is an active area of research in computer vision with a wide range of applications such as event recognition, traffic flow estimation and behavior prediction, and has received increasing attention in recent years. It remains a challenging problem despite the extensive research effort spent on this problem. These challenges are typically caused by the complex interactions between individuals in crowded scenes. In general, the existing methods to coherent motion detection in videos can be summarized as the optical flow tracking-based methods, which achieved the detection through the segmentation of optical flow field, and KLT tracker-based methods, which accomplished the recognition through analyzing the topological structure of the individuals. However, the multiple interaction inherent in coherent motion is widely ignored in these works. Actually it is useful for improving detection performance.

In this work, the Flow Density based multiple interaction

coherent motion detection Algorithm (FDA) is proposed, which achieves the detection from local and global level. At the local level, the coherent motion is regarded as the fluid motion, and flow density is defined. Then the flow density peaks clustering is introduced to detect the subgroups with strong local consistency. At the global level, the multiple interaction in subgroups is described by the proposed multiple adjacency model. Based on this, we can quantify the global consistency of subgroups. The coherent motion can be obtained by merging the subgroups according to the global consistency. Experimental results on several real videos demonstrate that FDA has better detection performance compared to the existing methods.

This work is supported by the National Natural Science Foundation of China (Grant Nos. 61170223, 61502434, 61672469, 61772475) and the Henan Provincial Key Technology Research and Development Program (Grant No. 172102210011).

