

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： 刘子宁

学号： 2012302514

班号： 10011206

# 基于场景无关的拥挤人群群体分析

邵京 香港中文大学 电子工程系

Chen Change Loy 香港中文大学 信息工程系

王晓刚 香港中文大学 电子工程系

## 摘要

研究组是由拥挤人群组成的实体。在群体的层次上理解运动以及特性在科学上是非常重要的,在大规模的应用领域是非常有用的,尤其是对拥挤人群的理解。在这个研究中,我们会展示基本的群体水平上的属性,比如群体内部的稳定性和群体之间的冲突,能够通过视觉描述系统地量化。通过学习一种新的之前的集体过渡成为可能,这会产生在公共领域进行群体分割的一种有效的方法。一直以来,我们进一步设计一组丰富的群体视觉描述属性。这些描述符是场景无关的,能够有效地适应灰度以及描述符多样的公共场景。在群体分析以及拥挤场景理解方面,数以百计的带有属性描述公共场景视频演示的广泛实验不仅仅是有用的,而且是必须的。



图一.群众行为能更好地被理解通过内部和外部固有的特性。在这个研究中,我们展示了量化诸如场景无关视觉描述的特性。最好通过颜色观察。

## 1.简介

群体运动已经在社会心理学以及生物学调查上被广泛地研究,作为能够影响拥挤人群行为的基本处理。在这些研究中,群体分析在群体内部以及外部属性都有特征的。群体内部特征,比如群体性,稳定性以及一致性,表明在相同群体内部成员的一致性。同时,群体外部特性,比如冲突,反映在不同群体成员间的外部作用。这样的属性在动物已及昆虫群体系统中广泛存在(比如细菌寄生以及鸟类成群),另外,在社会心理学上也被经常研究。以实例来说,细菌寄生被研究发现展示群体行为是为了实现一个共同的目标,扩展疾病。从社会学观点来看,冲突来自资源的竞争或者目标的不一致。

在视觉监视环境方面来看,群体性活动主要由人类群体组成。的确,丰富的

文献表明人们倾向于和他们的朋友和家庭成员组成一个群体。和其他行人形成一个群体的倾向在拥挤的人群中显得更加突出，人们不得不和其他人形成统一的行为来代替个人行为。

当人们来自群体，他们展示了一些令人感兴趣的行为，分享在心理以及生理上的共同点。举例来说，群体行为被观察到，当在一个集体里的人们有共同的目标。在一个群体灾难的时候，混乱行为能被稳定性所代替。集体倾向于不均匀的分布当它的成员有不同的社会关系以及在限制区域活动较少的话。有着不同目标的两人组展示冲突当从不同方向穿过公路的时候。明显的，理解这些特性提供了对集体行为分析决定性的表述，以及能促进其它更高水平的语义分析，比如群众场景分析，集体场景分割，以及集体事件恢复。

我们的目标是描述和量化这些集体的视觉特征，研究他们关于集体行为分析和集体场景理解的潜能。我们考虑仅仅由空间近似个人的一个集合体，但是也能展示多样的内部和外部集体行为特征的动态单元，能被用来通过不同的群众系统比较群体活动。就我们所知，这个研究是第一次人们想准确系统地研究群众集体的视觉属性。我们做了如下的贡献：

- 1) 一次功能强大的群体-探测器-我们介绍了一个有关捕获潜在的群体动能的群体过渡器。基于以前的知识，我们规划了一个功能强大的群体探测器来超过目前最先进的方法。
- 2) 场景无关的群体描述符-基于以前的 CT，我们设计了一组视觉描述符来适应四条最基本的群体内外的属性，分别是集合性，稳定性，均匀性和冲突性。这些描述符传达了更加丰富的有关比较常见的群体大小和变动群体水平的信息。特别重要地，这些描述符是场景不变，有效的描述变化人群的公共场景。
- 3) 迅速变化群体群众场景理解-我们表明被提到的描述符在描述固有的群体属性（气态，液态和固态）随着群众类比。我们也表明了他们的优越性对于场景无关群体情况分析和拥挤人群视频分类，对于已经存在的活动描述符。

实验是经过数以百计的从超过 200 个群体场景收集来的录像带的指导。这些资料和地面事实在群体人群分析方面更适应未来的调查。

## 2. 相关研究

大多数存在的基于影像的群众分析方法不是倾向于将群众作为个人的集合就是作为一个整体。对比这些研究，我们在群体水平上分析群众。以目标为中心的方法要求明确的检测和从群体中的个人分割。这些技术在拥挤场景是不可行的当群体内部目标被观察时。这些活动表明整体的方法，例如光线变化，人群流动，元素的分布，对于学习时空方面的场景水平模式是非常有用的，但是没有直接地从学习群体水平属性方面分析，这些要求有效的群体分割。

最先进的方法实现通过轨迹关联实现群体检测。周等人代表着连续滤波方法在分割集体行为方面。同时 Ge 等人发现小群众基于成对地通过分层聚类对象的速度和距离。在我们的实验中，上面的方法对跟踪噪声不是太敏感就是不能覆盖整个群众场景。特别重要的，他们研究的群众属性没有一个在群众水平上分析躯体行为。

大量的方法被提出来认识群众活动，例如开会和斗争。这些研究倾向于分析

小规模的社会集体，将注意力特别放在特定场景学习，交互建模，以及社会信号分析。然而，很多集合模型方法是特定场景的，例如，特定场景的活动模型不能适应其它场景。我们的研究明显地不同于之前提到的研究：我们将注意力放在了识别的不同方面能适应基本的群体属性，在原则上不同于其它群众系统，我们将注意力放在有大量成员的集体场景上面。

### 3. 分析组属性

我们将群体当作许多有共同目标和统一行为的成员的集合。给定一个较短的帧的录像，一组群体  $\{G\}_{i=1}^m$  被检测到。每一组  $G_i$  包含了一组轨迹  $\{z\}$  被 KLT 特征点追踪检测到。对于每一个被检测到的集合，我们希望提取出一系列的视觉描述符来描述它的属性。

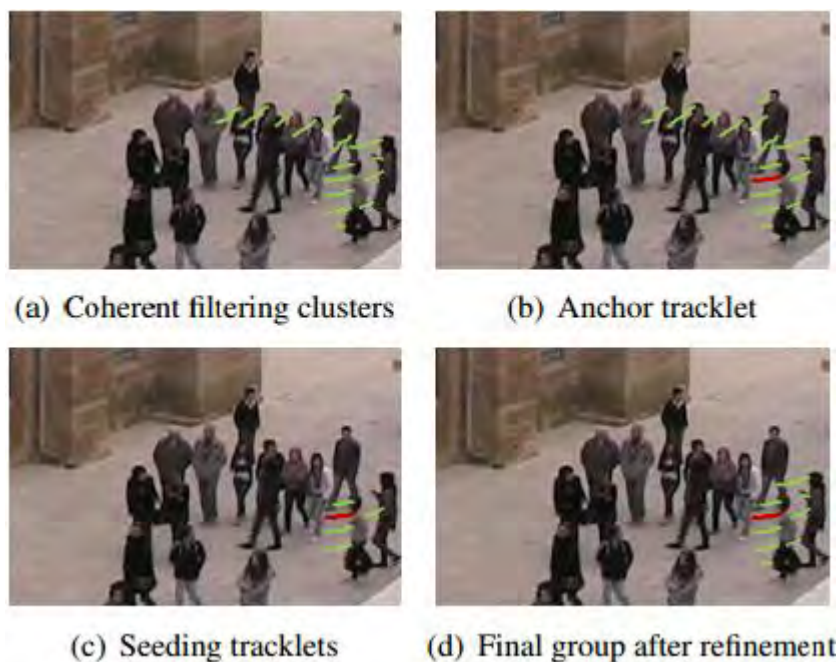
#### 3.1 之前的集体过度

在人群中精确的集合检测因为行人之间复杂的相互作用而具有挑战性。我们假定在一个场景里的行人运动熟悉的被有限数量的集合过度器所控制。这些方法同时发现群体检测过程。我们展示群体检测能预先知道群体检测更好地考虑时间上的连续和一致。

每个行人集合有一个特定的 CT 检测，是从录像中被发现的。更准确的说，对于  $n$  个轨迹， $\{z\}_{k=1}^n$ ，我们假定  $m$  个马尔可夫链，此时  $m < n$  和  $m$  是被自主推测的。每个马尔可夫链是时间序列模型，有如下形式

$$z_k^t = A z_k^{t-1} + v^t,$$

连续的观测量  $z_k^t$  通过变换矩阵  $A \in \mathbb{R}^{3 \times 3}$



图二 (a) 相关滤波，不能区分两个细微差别的集合。我们找到解决它的方法通过代表性的轨迹（红色）和一组种子轨迹来推断一组特定群体的 CT 值。在更细分的基础上，把两幅图片分开。

高斯噪声  $v_t \sim N(0, Q)$  在变换间求和。让  $z_k = [x_t, y_t, 1]^T$  代表行人的均匀位置坐标，初始观测点  $z_{1k}$  服从高斯分布  $N(u, E)$ ，我们表明  $\theta = \{A, Q, u, E\}$  作为链的参数。一个代表 CT 预测，能够揭示集合中所有成员的统一行为， $\{u, E\}$  保证集合成员时间上连续，在初始帧。接下来，我们讨论怎么学习这种方法，以及展示功能强大的实时检测。

### 3.2 通过集体变换进行组检测

最关键的想法是通过录像调查行人团体被检测的程度。这个方法允许分散轨迹一直不能跟踪整个录像。丢失的数据  $z_k$  被预测通过 EM. 所以不能适应密度大的人群。另外，它依赖于时空关系和没有考虑行人集合的整体形状。因此它能够适应不同变换和视点的场景。

学习 CT priors 关键的步骤对于组水平来说，是学会总结。

步骤一：产生相关滤波集合：我们首先用 CF（图 2（a））发现一组初始轨迹集群。这些集合没有完美的感知力，但是能够为发现最后的轨迹组  $\{G\}_{i=1}^m$  提供基础。

步骤二：确定锚轨迹：进行迭代，首先随机选择一个集群  $C_i$ ，找到他的锚轨迹  $z_i^*$  通过长时间和低方差。（图 2（b））

步骤三：通过 CT 学习发现种子轨迹：正如图 2(c)所展示的那样，一组种子轨迹  $S_i$ ，被选中了，用如下的标准：（1）它们也来自  $C_i$ ；（2）和  $z_i^*$  高度相关，

$$\frac{\langle v_{z \in C_i}, v_{z_i^*} \rangle}{\|v_{z \in C_i}\| \cdot \|v_{z_i^*}\|} > \eta, \quad (2)$$

#### 算法 1：组检测通过集体变换

输入：轨迹  $\{Z\}_{k=1}$  到  $n$  在一个录像带

输出： $m$  个轨迹组， $\{G\}_i$  从 1 到  $m$

步骤 1:  $i$ ，获得连续滤波集合  $\{C\}_j$  从 1 到  $r$

If  $\{C\}_i \neq \emptyset$ ，然后

    步骤 2: 统一一个锚轨迹， $z_i^*$ ；

    步骤 3: 从  $C_i$  发现种子轨迹集合  $S_i$ ；

    步骤 4: 细化  $G_i$

$\{G\} = \{G\} \cup G_i, \{C\} = \{C\} \setminus C_i, i = i + 1;$

end

$N$  是一个阈值。Si 包括可靠的轨迹，被用来代表 CT，这个将被用来第四布中的群体自身。

步骤四：组细化：我们改善初始集合  $C_i$  中的每一个轨迹  $z$ ，使用  $A_i$  中的第  $i$  个马尔可夫链。这个轨迹的改善误差  $e$  被定义为

$$\epsilon = \frac{1}{\tau - 1} \sum_{t=1}^{\tau-1} \|A z^t - z^{t+1}\|_2^2. \quad (3)$$

所有的误差小于  $\$$  的轨迹来构想  $G_i$ 。不适合的轨迹将会重复这个迭代过程来产生不同的群体。

### 3.3 人群场景的群体描述符

我们规划了一组描述符来适应群体的属性（表 1）。前三个属性适应群体之间结构的时空变换，同时第四个描述群体内部的相互作用。第四步展示了他们相互补充来时场景无光的组分析表现得更好，以及更加已处理

Table 1. List of group descriptors.

Property	Descriptor	Equation
Collectiveness	$\phi^{\text{coll}}(\mathcal{G})$	4
Stability	$\Phi^{\text{stab}}(\mathcal{G})$	10
Uniformity	$\Phi^{\text{unif}}(\mathcal{G})$	13
Conflict	$\Phi^{\text{conf}}(\mathcal{G})$	14

为了进一步说明，我们做了一个分析在点和成员之间。一个被观测的组在一个帧里有  $n$  个成员，形成了  $K$ -NN 图片， $G(V, E)$ ，它们的定点  $V$  代表了它们的成员，成员对之间通过边来联系， $E$ 。这些边被一个同类的矩阵  $W$  进行加权，通过元素  $w_{ij} = \exp(-d_{ij}^2 / \sigma^2)$ ， $d_{ij}$  是两个成员之间的空间距离。我们表明集合中成员  $z$  的邻域作为  $Nz_1, \dots, Nz_t$ ，在每一个给定录像的视频帧。接下来我们详细讨论描述符。

**集体性：**集体属性表明个体的活动作为一个联合统一的行为。它是基本一般的处理方法对于拥挤系统来说。一个统一量度对于一整个录像来说被提及用来多方面学习。对比来说，我们适应一致行为在群体水平上通过被提及的几何变换  $A$ ，因为它捕获了所有成员的连续行为。特别地，我们计算出群体的一致性为：

$$\phi^{\text{coll}}(\mathcal{G}) = \frac{1}{|\mathcal{G}|} \sum_{z \in \mathcal{G}} \epsilon(z, A),$$

在这里  $|\cdot|$  表明输入集合的基， $e(z, A)$  在 Eqn 中被定义。

在  $\phi^{\text{coll}}(\mathcal{G})$  中的低数值表明一个集体里的成员向一个共同的目的地移动时非常一致。这个描述符对有区别的低统一性的群体非常有用，例如，在一个火车站或一个雨天市场，从一个高统一性的群体，例如，在一个马拉松赛跑或一个自动扶梯上。

**稳定性：**稳定特性是一个集体中能够全程保持内部拓扑结构。它可以类比于

在一个化学系统中的稳定的结构。特别地，稳定的成员倾向于保持和邻域保持一个相同的集合；和它的邻域保持一致的拓扑距离；一个成员很少喜欢离开它目前最近的集合。按照这个思想，我们规划三个稳定描述符。

我们计算第一个稳定描述符通过计算以及平均每一个成员不变邻域的数目在一副 K—NN 图片中

$$\phi_a^{\text{stab}}(\mathcal{G}) = \frac{1}{|\mathcal{G}|} \sum_{z \in \mathcal{G}} (K - |\mathcal{N}_z^1 \setminus \mathcal{N}_z^T|), \quad (5)$$

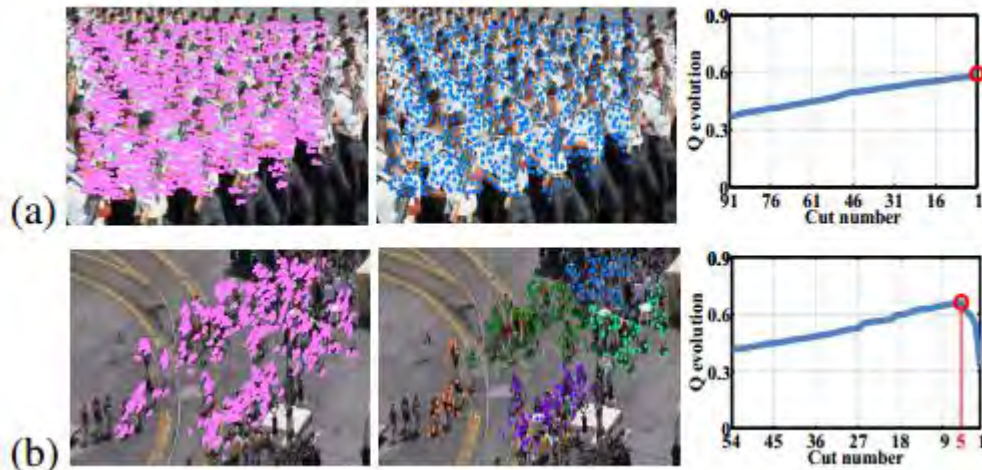
这时  $|\mathcal{N}_z^1 \setminus \mathcal{N}_z^T| = |\{z : z \in \mathcal{N}_z^1 \text{ and } z \notin \mathcal{N}_z^T\}|$ .

第二个稳定性描述符是规划检测如果成员保持一定的拓扑结构和他的最近的邻居。它的实现是通过第一排序在一个成员最近的邻居中，来将他们两两结合，然后应用距离度量字符串来比较每两个连续的帧的排名。如果两个排名相同，

$d_z^t = 0$ ，如果所有成员的排名改变了  $d_z^t = K$ 。通过聚集 dtz 在整个帧，我们构建它的直方图， $h(z)$ ，对于每一个成员第二个稳定性描述符通过均匀直方图来获得

$$\Phi_b^{\text{stab}}(\mathcal{G}) = \frac{1}{|\mathcal{G}|} \sum_{z \in \mathcal{G}} h(z). \quad (6)$$

它揭示了在一个群体中成员之间的拓扑结构距离改变的信息。



图三：(a) 和 (b) 展示了一副均匀分布和非均匀分布的群体，特别地。从左到右，我们展示了初始连续群体检测，子集合通过进一步归类得到，最佳分割通过模型函数得到。

第三个稳定性描述符度量成员离他最近的邻域集合。我们假定一个随机的行走行为在所有群体成员之间，例如，我们允许成员自由组合在组内，还可以加入其它成员形成新的邻域。我们然后度量这些成员的稳定性作为不同点在它的初始

情况和最后变换的情况。我们初始化变换概率矩阵  $\mathbf{P} \in \mathbb{R}^{n \times r}$

$$\mathbf{P} = \mathbf{D}^{-1}\mathbf{W}, \quad (7)$$

$\mathbf{D}$  是对角矩阵，它的元素  $D_{ii} = \sum_j w_{ij}$ 。第  $i$  个成员的密度分布“走动”和“加入”其它成员被定义为

$$\mathbf{q}_i = \mathbf{e}_i^T [(\mathbf{I} - \alpha\mathbf{P})^{-1} - \mathbf{I}], \quad (8)$$

其中  $\mathbf{q} \in \mathbb{R}^{1 \times n}$ ， $\mathbf{I}$  是单位矩阵， $\mathbf{e}_i = (\mathbf{e}_1, \dots, \mathbf{e}_n)$  的转置是指示向量当  $\mathbf{e}_i = 1$ ， $\mathbf{e}_j = 0$ 。参数  $\alpha$  的范围是  $0 < \alpha < 1/\rho(\mathbf{P})$ ， $\rho(\mathbf{P})$  表明  $\mathbf{P}$  的频谱半径。我们设置  $\alpha = 0.9/K$ 。第  $i$  个成员的稳定性通过测量  $\mathbf{q}_i$  的 KL 散度来得到， $\mathbf{q}_i$  在第一帧和最后一帧之间。更低水平的 KL 度量得分， $s_{kl}$  表明更高的稳定性。我们计算第三个描述符，通过平均所有成员的得分

$$\phi_c^{\text{stab}}(\mathcal{G}) = \frac{1}{|\mathcal{G}|} \sum_{z \in \mathcal{G}} s^{\text{kl}}(z). \quad (9)$$

最后稳定性描述符

$$\Phi^{\text{stab}}(\mathcal{G}) = [\phi_a^{\text{stab}}(\mathcal{G}), \Phi_b^{\text{stab}}(\mathcal{G}), \phi_c^{\text{stab}}(\mathcal{G})]. \quad (10)$$

**均匀性:** 均匀性是非常重要的属性对于标记一个群体的同质性，在时域描述方面。这个特征对比于前两个描述暂时方面的特性。一个群体是均匀分布的如果他们的成员相互保持亲近甚至在空间中分散式分布。一个非均匀分布的群体能够分成更多的亚群体。一个均匀分布和非均匀分布集合的对比例子如图 3 (a) 和 3 (b) 所示。

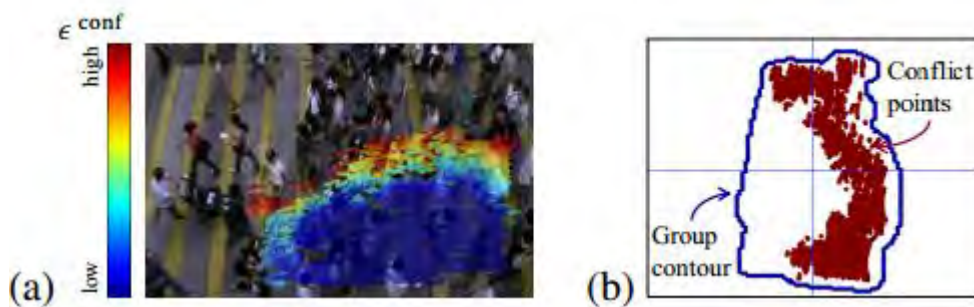


图 4: (a) 图中的冲突位置。深颜色表明高度冲突。(b) 集合的冲突分布。这个地图在集合中心被标准化，运动检测，最大的距离以集合为中心的等高线

我们量化均匀分布通过推测图像分割的最佳数字。一个高  $C^*$  表明图像是高度非均匀的。一个集合的层次是通过凝聚的集群得到的，以及模型函数  $Q$  被用来找到  $c^*$ 。特别地，给定一个集合数字  $c$ ，一个图像分割  $\{V_1, \dots, V_2\}$  从  $H$  获得。对  $C\{1, \dots, C\}$  计算  $Q_c$  它的最大值表明最佳数字分割：



$$c^* = \underset{c \in \{1, \dots, C\}}{\operatorname{argmax}} Q_c \quad (11)$$

$$\text{given } Q_c = \sum_{i=1}^c \left[ \frac{\mathcal{A}(V_i, V_i)}{\mathcal{A}(V, V)} - \left( \frac{\mathcal{A}(V_i, V)}{\mathcal{A}(V, V)} \right)^2 \right], \quad (12)$$

$$\mathcal{A}(V', V'') = \sum_{i \in V', j \in V''} w(i, j).$$

。在图 3 最后一列中展示的例子。他们表明非均匀分布的集合有一个相当高的数字分割。随着群体进化，均匀的群体可能改变，我们度量它的均匀性通过最优数字的均值  $\mu_{c^*}$  和方差  $\sigma_{c^*}$

$$\Phi^{\text{unif}}(\mathcal{G}_i) = \{\mu_{c^*}, \sigma_{c^*}\}. \quad (13)$$

**冲突性：** 冲突属性描述群体之间的摩擦或斗争，当他们相互接近时。空间分布和冲突能被图 4 展示的 2 维标准地图所展示。这样的一副地图对于群体识别是很有信息量的，它包含了在不同场景之间群体内部相互作用的不同点的丰富信息。在这幅地图里，集合等高线被获得，作为内部成员间的外部边界，一个冲突点被定义为一个成员 对比于外部群体成员的 K-NN 集合， $\mathbf{N}$ 。注意 K-NN 集合被定义为不同于我们先前提到的那样，目前的集合被包含在外部集体的成员。

为了简洁地代表冲突地图的不变性，我们规划一个冲突形状概念描述符。第一步是捕获每一个空间分布上的冲突点通过计算相关集合等高线的点的直方图。这正好被以每一个冲突点为中心的极坐标实现，计算在图像中每一个等高线点的频谱。8 个相等的空间角度和 5 个相等的空间半径。第二步时表明 K-均值集群通过建立直方图的词汇表，产生 Bow 代表。使用局部的线性编码，第 i 个冲突点通过词汇表有一个分布  $u_i$ 。我们更深一步了解冲突的水平基于在 3.1 节介绍的 CT

$$\epsilon_i^{\text{conf}} = \frac{1}{|\mathcal{N}_i|} \sum_{\mathbf{z} \in \mathcal{N}_i} \epsilon(\mathbf{z}, \mathbf{A}),$$

在第三节定义的误差  $\epsilon(\mathbf{z}, \mathbf{A})$ ， $\mathbf{A}$  是群体中冲突点 CT 。直观地，如果一个冲突点的最近邻域大多数外部成员不适合  $\mathbf{A}$ ，一个高数值在  $\epsilon^{\text{conf}}$  中被获得。最后冲突的群体属性被计算通过被  $\{\epsilon^{\text{conf}}\}$  加权的  $u_i$ 。

## 4. 操作及实验结果

我们评估群体检测，证明我们关于两个操作的实验符的影响：群体水平分析和人群录像分类。都是场景无关的。

### 4.1 人群数据

评估是由一个新的香港中文大学人群资料组做的。它包括不同灰度和远景标定的人群录像，从不同的环境搜集而来，比如，街道，商店，飞机场，停车场。它包含了从 215 个场景搜集来的 474 个录像带，419 个录像从 Pond5 和 Getty Image

网站搜集来，55 个是我们获得的。它比在人群数据和任何存在的人群数据都大（他们确实包括我们的资料组）。虽然录像有不同的长度，我们仅仅从每个录像中的头三十帧中实现我们的方法。在数据方面，整盘录像合适的。组检测的基础事实，组水平的分析，组录像带的检测时自动的，它们被倍数注释器检测。

## 4.2 群体检测

从 300 个录像检测的轨迹是手动获得的，在同一个群体的成员有共同的目标，形成了统一的行为。不属于任何群体的轨迹被注释为异常值。我们比较用统一过度方法和三种场景有关方法：动态混合，分层归类，继承滤波。基础事实的例子和检测结果的比较在图 5 中展示。

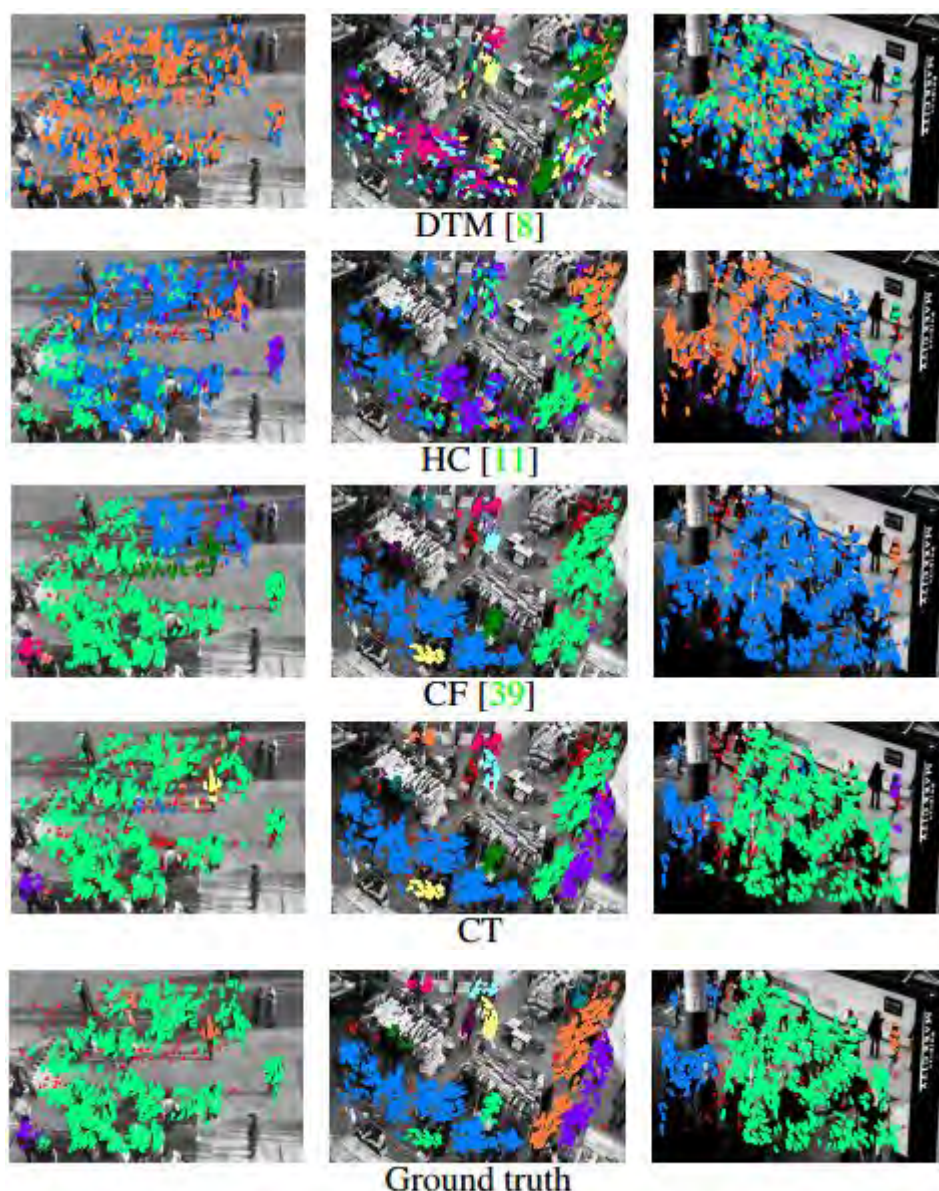
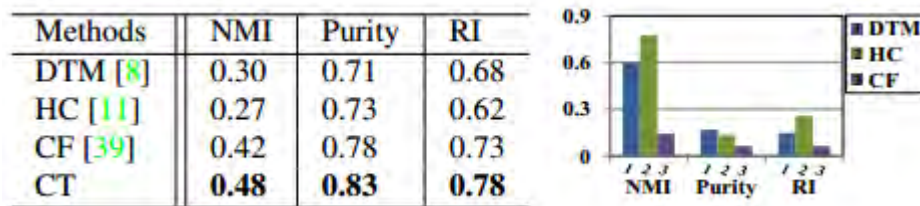


图 5. 群体检测和四种方法的比较结果。颜色区别是明显的。红色表示异常值。箭头表示移动方向。在颜色上的最好展示。

DTM 很好的将背景和简单群体运动分离，但是复杂和混合群体运动效果不明显。另外，它要求手动群体规范（我们提供输入的基础事实），对每一个录像它比我们的方法花费了更多的时间。HC 等级簇轨迹速率和空间约束，没有考虑

群体动态。它因此比我们导致了更多的错误。CF 检测连贯动作，用邻域测量，没有模型分析，被整个组分享。它因此对轨迹失败检测很敏感。这是能在图 5 的第一列观测到，CF 在亚组同一方向检测了分裂的轨迹。另外，CF 第一次检测组用继承运动在连贯的帧和群体搜索之间，通过整个录像带。它因此是误差累积的。在第二和第三列，CF 联合了两个不同方向的群体运动，产生于各个帧的误差。

对于定量的评价，我们考虑群体检测作为一个族群我呢提，广泛地适应族群评估度量，例如正常信息交互，纯度，随机指数。图 6 展示了这种比较。在右边的表展示了我们方法的相关提升，和 DTM,HC,和 CF 相比。



图六：左边：定量比较的群体检测方法:和 DTM,HC 和 CF 相比我们的方法有很大提高。

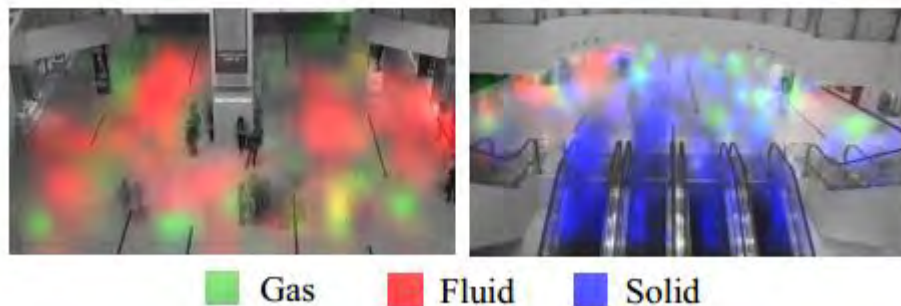


图 7：不同类型的群体分布，在一个商城和一个自动扶梯场景。颜色表明群体自动检测用我们的描述符。颜色的最好描述。

### 4.3 应用 I:群体水平分析

立体和分析人群研究通常将人群分类为下面的组别用类比的分类不同的短语的物质平衡统计力学。它假定潜在的物理模型是不同的。

气体：没有和其它成员形成统一的行为，在不同的方向有离散的运动。

固态：在同一方向有统一的行动。他们相关的位置保持不变，被内部力量绑定。

纯液态：趋向同一方向；然而，他们相关的位置经常改变，因为缺少内部力量。

不纯液态：它和纯液态很像，但是受到其他群体的入侵。

这些状态被社会心理学和生理因素决定，包括群体灰度，目标，反应和群体成员的关系和场景结构。在图 7 展示的例子，在一个大的开放区域，行人大多表现得像气态和液态，但是在一个自动扶梯或队列中移动的像飞的固态。在左图中，流动的组在路径联系和存在区域展现的非常频繁，同时气体的组位置随机，他们就像走动的服装个体。在右图中，固态和液态之间的组变换的不这么频繁。

在一个竞争资源的群体中，他们表现得像液态。群体位置很好地反映了这种因素，对不同的应用有兴趣。我们使用提议的特征描述符来展示群体组的位置，在群体场景理解中非常有用。

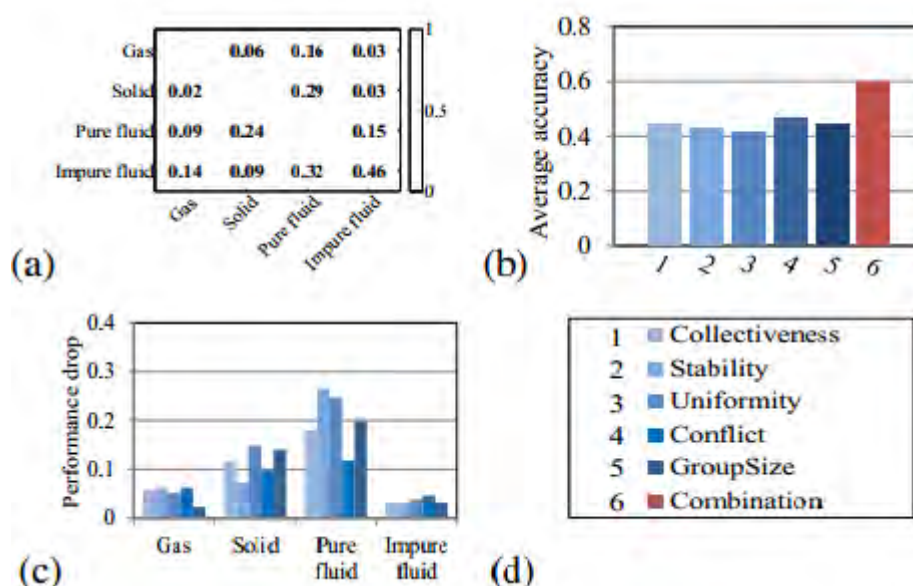


图 8,(a)混合所有组描述符的混合矩阵表明群体位置。(b)混合每一个描述符的平均精度。(c)仅仅用一个描述符性能下降。更低的线表明描述符对一个特殊的群体位置更有效果。(d) (b)和(c)的说明。1~5 是单个描述符，6 是五个描述符的混合。

作为基础事实，有 927 组手动标签：128 个气态组，291 个固态组，249 个纯液态组合 159 个不纯液态组。半数的数据对训练和保持文本时随机选择的（训练和文本设置不包含一样的场景）。所有我们提议的组水平的描述符被作为 SVM 输入的特征混合。这个混合矩阵在图 8(a)有超过 10 个考验。平均准确度是 60%，同时随机猜中的机会是 25%。这个结果展示了我们群体描述符的有效性，以及他们普遍在多个场景中更有效果。主要的但是微妙的不同在两个班级之间是冲突点的时间分布。轨迹错误也是困难增加的，在两个分离的场景。图 8 (b) 和图 8(c)展示了每一个群体描述符的有效性在不同的群体场景之间。观测到的是，稳定性和冲突时最有效果的在分析固体组方面。统一性和冲突时最有效果的在分析纯液态组方面。群体大小对气体组是有影响的。

#### 4.4 应用 II:群体录像分析

我们也证明目标群体描述符的强大功能和有效性在分析群体录像带方面，以代替个体。有关使用整体描述符来分析群体录像调查研究是存在的。以实例来说，比如 Kratz 将录像分成空间和提取运动特征。我们展示我们的描述符通过群体特征设计，这些比一般特征更有效果。

在资料库中的所有 474 组录像时手动设计成 8 类，在标签 2 中展示。这 8 类在群体录像经常看到，还有一些在群体管理和交通控制特殊的兴趣。比如，群体合并和群体交叉可能导致交通拥挤和群体灾难，比如逃散。在入口保持自动交通平稳以及在出口避免阻塞，碰撞和潜在危险是重要的。在类 1，在行走场景的行人，有高度混合行为的不同方向。在类 2 和 3，大多数的行人跟随主要行为。在类 2，行人相关的位置是平稳的，很少改变，同时，在类 3，行人没有很好组

织。大多数的群体录像能够一般的被这三个情形描述。然而，我们很少分析类（4~8），这些在群体中有特殊的特征，以及区别他们从现有的群体录像中。因此类3~4将类4~8排除在外。所有8个类别被一起分析。分析被使用。每一个时间场景（可能包含很多录像）被用来测试，未来测试保持的场景。因此他测试了交叉场景的普遍的情形。如果一个录像有多样的组，我们取群体描述符的平均值作为录像描述符。SVM 用来分析。混合矩阵在图9中展示。我们方法的平均精准度是70%，远高于随机猜测的12.5%，使用整体群体场景的描述符的结果是(44%)。

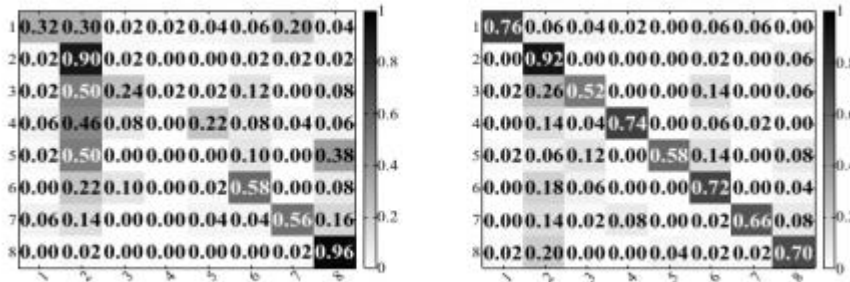


图9 群体录像分析的混合矩阵。左图:使用整体特征。平均精度是44%。右图:使用我们的描述符。平均精度是70%。

Table 2. List of crowd video classes.

	Class name
1	Highly mixed pedestrian walking
2	Crowd walking following a mainstream and well organized
3	Crowd walking following a mainstream but poorly organized
4	Crowd merge
5	Crowd split
6	Crowd crossing in opposite directions
7	Intervened escalator traffic
8	Smooth escalator traffic

## 5. 总结

在这篇论文中，我们从不同的视觉点系统的研究了存在于多样的群体系统的基本一般的群体属性。它们被社会心理学的研究所提醒，在群体场景分析中非常重要。以个功能强大的群体检测算法和丰富的群体属性描述符被提议通过学习群体过渡。它们很好地适应了场景无关群体分析，以及群体录像分析。这次研究也在未来的工作中激发了新的应用，比如交叉群体事件检测和模拟行人混合群体动态。

**确认：** 这次研究由香港委员会支持的一般研究基金会赞助研究经费。

## 参考文献：

- (1) c . c . Aggarwal. 人机互动投影的方法集群. 16(4):448-460, 2004.

- (2) Aggarwal and M. S. Ryoo. 人类活动分析: 审查, ACM 计算调查, 43(3): 2011。
- (3) S. Ali and M. Shah. 拉格朗日质点动力学的方法, 人群流分割和稳定性分析。
- (4) M. R. Amer, D. Xie, M. Zhao, S. Todorovic, and S.-C. Zhu. 多尺度活动的厂商自顶向下或自底向上的推理认可。
- (5) A. F. Aveni. 友谊团体的集体行为 40:96 - 99,99
- (6) L. Bazzani, M. Cristani, G. Paggetti, D. Tosato, G. Menegaz, and V. Murino. 分析组: 一个社会信号视角商业智能视频分析,
- (7) L. Bazzani, M. Cristani, G. Paggetti, D. Tosato, G. Menegaz, and V. Murino. 与形状上下文匹配。
- 在统计和分析的形状
- (8) A. B. Chan and N. Vasconcelos. 建模、聚类和分割视频动态纹理的混合物
- (9) M.-C. Chang, N. Krahnstoeber, and W. Ge. 概率组级别运动分析和场景识别
- (10) W. Choi and S. Savarese. 一个统一的框架, 用于多目标跟踪识别和集体活动
- (11) W. Ge, R. T. Collins, and R. B. Ruback. 建立分析小组的行人人群
- (12) D. Helbing and A. Johansson. 行人, 人群疏散动力学
- (13) L. F. Henderson. 人群中液体的统计数据
- (14) I. Karamouzas and M. Overmars. 模拟和评估当地小行人组织的行为
- (15) S. Kim, S. J. Guy, and D. Manocha. velocity 的建模在多主体模拟系统物理交互
- (16) L. Kratz and K. Nishino. 异常检测的非常拥挤场景使用时空运动模式模型
- (17) D. Kuettel, M. D. Breitenstein, L. Van Gool, and V. Ferrari. 发生了什么? 发现在动态时空依赖性场景
- (18) S. Kullback and R. A. Leibler. 信息和充分性数理统计年鉴
- (19) T. Lan, Y. Wang, W. Yang, S. N. Robinovitch, and G. Mori. 有识别力的潜在模型识别上下文团体活动
- (20) V. I. Levenshtein. 二进制代码能够纠正删除, 插入和逆转
- (21) R. Li, R. Chellappa, and S. K. Zhou. 认识互动集团使用时间交互矩阵和黎曼的活动统计数据
- (22) C. C. Loy, T. Xiang, and S. Gong. 通过高斯过程造型多目标活动
- (23) C. C. Loy, T. Xiang, and S. Gong. 检测和识别行为异常
- (24) R. Mehran, A. Oyama, and M. Shah. 人群异常行为检测利用社会力模型
- (25) M. Moussaid, S. Garnier, G. Theraulaz, and D. Helbing. 集体信息处理和模式形成成群的羊群, 人群
- (26) M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz. 行人的步行行为社会群体及其影响人群动力学
- (27) S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool. 社会行为建模的多目标跟踪
- (28) W. M. Rand. 聚类方法的评价的客观标准
- (29) M. Rodriguez, J. Sivic, I. Laptev, and J.-Y. Audibert. 数据驱动的人群分析视频
- (30) S. Smyth and S. White. 谱聚类方法来发现社区的图表
- (31) J. Toner, Y. Tu, and S. Ramaswamy. 流体动力学 abcd 的阶段
- (32) T. Vicsek and A. Zafeiris. 集体运动
- (33) J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong. Localityconstrained 线性编码的图像分类

- (34) X. Wang, X. Ma, and W. E. L. Grimson. 无人监督的活动使用分层感知在拥挤和复杂的场景贝叶斯模型
- (35) S. A. Wheelan. 组织研究和实践的手册
- (36) M. Wu and B. Schölkopf 本地集群学习的方法
- (37) H.-P. Zhang, A. Beer, E.-L. Florin, and H. L. Swinney. 细菌菌落集体运动和密度波动
- (38) W. Zhang, X. Wang, D. Zhao, and X. Tang. 图联系程度:有向图集群
- (39) B. Zhou, X. Tang, and X. Wang. 人群杂波相干过滤:检测连贯的动作
- (40) B. Zhou, X. Tang, and X. Wang. 测量人群集体主义
- (41) B. Zhou, X. Wang, and X. Tang. 理解集体人群行为:学习动态行人-事件的混合模型。