

单位代码: 10293 密 级: \_\_\_\_\_

南京邮电大学

专业学位硕士学位论文



论文题目: 基于监控视频动态信息分析的人流密度检测  
算法的研究

学 号 121202201

姓 名 乐 勇

导 师 龚建荣

专业学位类别 工程硕士

类 型 全 日 制

专业（领域） 电子与通信工程

论文提交日期 二〇一五年四月

# **The study of people flow density detection algorithm based on video surveillance of dynamic information**

Thesis Submitted to Nanjing University of Posts and  
Telecommunications for the Degree of  
Master of Engineering



By

YueYong

Supervisor: Prof. Gong Jianrong

April 2015

## 南京邮电大学学位论文原创性声明

本人声明所呈交的学位论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得南京邮电大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

本人学位论文及涉及相关资料若有不实，愿意承担一切相关的法律责任。

研究生签名：\_\_\_\_\_ 日期：\_\_\_\_\_

## 南京邮电大学学位论文使用授权声明

本人授权南京邮电大学可以保留并向国家有关部门或机构送交论文的复印件和电子文档；允许论文被查阅和借阅；可以将学位论文的全部或部分内容编入有关数据库进行检索；可以采用影印、缩印或扫描等复制手段保存、汇编本学位论文。本文电子文档的内容和纸质论文的内容相一致。论文的公布（包括刊登）授权南京邮电大学研究生院办理。

涉密学位论文在解密后适用本授权书。

研究生签名：\_\_\_\_\_ 导师签名：\_\_\_\_\_ 日期：\_\_\_\_\_

# 摘要

随着经济的快速发展，人员往来的日益频繁，人口流动的速度在不断地加快。这也导致城市里、景区中高密度人流的公共场所越来越多。而这些场所的人员安全问题也日益成为一个十分迫切的问题。视频监控技术的高速发展，特别是人流密度检测技术的兴起，为保障这些场所的安全提供了一种手段。通过实时的人流密度检测系统，管理者或者组织者能够及时方便的采取相对应的措施，避免险情的发生，保证人民群众的生命财产安全。

本文首先介绍了人流密度检测的意义以及国内外人流密度检测研究的现状。然后研究了图像预处理的知识，这也是人流密度检测的理论基础。在特征提取与跟踪算法的研究中，重点探讨了基于 KLT(Kanade-Lucas-Tomasi)算法的特征提取与跟踪，并将其与传统特征提取方法进行了对比分析，最后利用 KLT 算法完成了视频图像特征点的提取。这一章的研究是为聚集度反应人流密度奠定了基础。在人流密度检测算法的研究中，分析了目前主流的基于像素统计和基于纹理分析的人流密度检测算法。通过实验完成了两种算法的对比，其中基于纹理分析的算法采用了灰度共生矩阵。在深入分析了 SDP(Self-Driven Particles)模型后阐述了聚集度的概念后，通过仿真检验了 SDP 模型。聚集度描述了个体参与群体运动的程度。通过聚集度，能够帮助管理者迅速锁定视频运动中的起主导作用的个体。同时聚集度也能够定性地反应人流密度的大小，人流密度高时，聚集度也高，反之亦然。最后通过将人流密度与聚集度相结合，实现了聚集度实用性与准确性的验证，并给出五种密度等级下聚集度的取值范围。聚集度与人流密度的结合应用是本文的创新之处，这也为日后研发完整的系统提供了参考。

**关键词：**人流密度检测、数字图像处理、图像特征、SDP 模型、聚集度

# Abstract

With the rapid development of economy, personnel exchanges have become increasingly frequent, the population migration velocity is speeding up too. More and more public places are in a high density flow. The safety of these places is becoming a very urgent problem. With the high speed development of video surveillance technology, especially the rise of crowd density detection technology, which provides the guarantee for the safety of these places. With the help of the crowd density detection system in real time, manager or organizer can take corresponding measures timely and conveniently to avoid the dangers, guarantee the safety of life and property.

Firstly, we introduced the current situation of research of people flow density detection at home and abroad. Secondly, this thesis studied some theory and knowledge of flow detection in the field of image process. most of them is the knowledge of digital image processing, including image enhancement, image denoising, morphological processing and edge detection etc. Then we discussed the feature extraction and tracking based on KLT algorithm. It laid foundation of collectiveness that reflected crowd density. There are two traditional methods in crowd density detection based on the statistics of pixel and based on texture analysis. We compared them with experiment. In the description of the SDP model we introduced the concept of Collectiveness, which indicates the degree of individuals acting as a union in collective motion, is a fundamental and universal measurement for various crowd systems. While people density is high, Collectiveness is high, and vice versa. Finally we combined flow density and Collectiveness and applied it to the practical application, show its practicality and accuracy. We also give the range of collectiveness in five different crowd density. This combination was just the creation of this thesis and it provided a reference for designing a complete system in the future.

**Key words: Crowd density detection, digital image processing, image feature, SDP model, Collectiveness**

# 目录

第一章 绪论	1
1.1 背景描述与选题意义	1
1.1.1 论文的研究背景	1
1.1.2 论文的研究意义与应用	2
1.2 人流密度检测研究现状	3
1.3 研究内容及章节安排	3
第二章 图像预处理技术	5
2.1 数字图像处理概念	5
2.1.1 数字图像分类	6
2.1.2 像素间的基本关系	6
2.2 图像预处理	7
2.2.1 图像的灰度化	8
2.2.2 图像去噪	8
2.2.3 图像增强	10
2.2.4 形态学处理	10
2.3 边缘检测	12
2.3.1 Sobel 算子	13
2.3.2 Roberts 算子	14
2.3.3 Canny 边缘检测	15
2.3.4 边缘检测总结	17
2.4 本章小结	17
第三章 特征提取与跟踪算法研究	18
3.1 图像特征分类	18
3.1.1 纹理特征	18
3.1.2 形状特征	19
3.1.3 颜色特征	20
3.1.4 空间关系特征	21
3.2 KLT 算法	22
3.2.1 特征点概念	22
3.2.2 KLT 算法发展	24
3.2.3 KLT 算法原理	24
3.3 结合 KLT 算法的特征跟踪概述	27
3.3.1 图像特征点的选取	27
3.3.2 图像特征点的跟踪	28
3.3.3 特征点检测和跟踪算法分析	28
3.4 本章小结	30
第四章 人流密度检测算法的研究	31
4.1 人流密度等级分类与系统流程	31
4.1.1 人流密度等级分类	31
4.1.2 人流密度监控系统工作流程	32
4.2 前景提取	32
4.2.1 光流法	33
4.2.2 帧间差分法	33
4.2.3 概率方法	34

4.3 人流密度检测的常用方法.....	35
4.3.1 基于像素统计的人流密度检测方法.....	35
4.3.2 基于纹理分析的人流密度检测方法.....	36
4.4 实验结果对比.....	40
4.4.1 基于像素统计的实验结果.....	40
4.4.2 基于纹理分析的实验结果.....	42
4.5 本章小结.....	44
第五章 SDP 模型的研究与应用.....	45
5.1 SDP 模型简介.....	45
5.2 聚集度的测量.....	46
5.2.1 邻域范围内的行为相似性.....	46
5.2.2 群体范围的行为相似性.....	47
5.2.3 个体聚集度与群体聚集度.....	48
5.2.4 聚集度的性质.....	49
5.2.5 群体聚集算法步骤.....	49
5.3 基于 SDP 模型的算法验证.....	50
5.3.1 SDP 模型的群体聚集度.....	50
5.3.2 混合群体的聚集.....	51
5.3.3 聚集度的收敛性.....	52
5.4 SDP 模型的进一步研究与应用.....	53
5.4.1 聚集度与人类感知.....	53
5.4.2 视频信息中的人群运动监测.....	55
5.4.3 细菌菌落中的群体运动.....	56
5.5 聚集度与人流密度的结合应用.....	57
5.6 本章小结.....	59
第六章 总结与展望.....	60
6.1 总结.....	60
6.2 展望.....	60
参考文献 .....	61
致谢.....	65

# 第一章 绪论

人流密度检测是运用摄像机等设备采集人流的视频信息，使用现代图像处理技术分析并处理视频画面中的活体运动目标，并计算视频中目标区域内行人的数量和密度，当人群的数量值超过预先设置的阈值时系统自动给出报警信息。工作人员结合报警信息能够及时地采取相对应的解决措施，可应用于诸如庙会，机场，车站等公共场所。

## 1.1 背景描述与选题意义

自然界是一个充满神奇的地方，而人类作为当今社会活动的主体，研究并解释人类的行为，不仅对于人类认识自身，而且对于了解自然具有十分重要的意义。在科技尚未发达以前，人们更多的通过目测观察来加深对这个世界的了解，而信息科技革命以来，尤其是计算机技术的兴起，人们可以借助计算机及辅助工具探知未知领域。而在计算机视觉中，视觉分析完成了对人的图像序列的识别、提取以及跟踪等一系列理解和分析的过程，因而成为当前的一个热门研究领域。计算机视觉学科交叉性强，涉及到了图像处理，模式识别，人工智能等学科知识；同时动态视频信息中，人体与人体，人体与周围建筑之间的相互遮挡也给我们的研究带来一定挑战。在监控视频的动态信息中，人是运动场景中最主要的目标，因此对视频中人体信息进行正确地特征提取，跟踪和识别有利于我们准确迅速的完成对人群活动的行为分析与理解。

### 1.1.1 论文的研究背景

伴随着经济社会的高速发展，城镇化进程的加速，人们之间交流活动的也不断密集，车站、商场、展览会场等大型现代活动场所发生人群聚集进而堵塞的现象时有发生。这已经成为安全隐患，尤其是节假日或者体育赛事，庆祝活动期间，一旦疏散不及时，造成的后果难以想象。元旦发生的上海外滩踩踏事件的惨剧依然历历在目。下面是几个在不同场所由于人员过度拥挤而导致的踩踏惨剧。

娱乐场所：2014年12月31日23时35分许，正值跨年夜活动，因很多游客市民聚集在上海外滩迎接新年，黄浦区外滩陈毅广场进入和退出的人流对冲，致使有人摔倒，发生踩踏事件。截至2015年1月23日11点，事件造成36人死亡49人受伤，已有47人经诊治后出



院，2 名重伤员继续在院治疗，其中 1 名生命体征还不平稳。

集会场所：2010 年 3 月 4 日，印度北部一处庙宇发生踩踏事件，造成 63 人死亡，上百人受伤。2011 年 1 月 14 日，该国南部喀拉拉邦发生了导致约 100 人死亡，90 人受伤的踩踏事件。

车站场所：2012 年 5 月 10 日，南非城市比勒陀利亚一个火车站发生踩踏事件，事故共造成 17 人受伤。2013 年 2 月 10 日，印度北部城市阿拉哈巴德因举办大壶节，导致火车站人员骤增，最终酿成踩踏的惨剧，事故导致了至少 40 人的伤亡。

仔细分析这些灾祸发生的原因，都是因为人流密度过大，致使人群过度拥挤，一旦无法流通疏散或者发生某些突发情况，人群行为就会失控，从而发生惨剧。管理者对群众性活动预防准备不足、现场管理不力、应对处置不当。近些年来，无论国内还是国外，这样的事故时有发生，因此如果能够对这些场所进行有效的监控管理，事实动态地观测人群流动情况，当异常或危险情况发生时，系统能够自动检测出来并立刻发出报警，相关部门就可以果断采取措施，避免意外发生。

### 1.1.2 论文的研究意义与应用

伴随着我国经济社会的飞速发展，人们之间的交往不断频繁，人员往来不断聚集，而由于人群密度过高导致的踩踏事件时有发生。在日常生活中，对大家频繁出入的地铁，超市，广场，车站等场所进行人流密度的估计是十分重要的。因此，基于动态视频信息的人流密度的检测具有深远的现实意义和研究价值。

人流密度检测的意义可以表现在以下几方面：

一、人群集中的高密度场所。如大型娱乐场，会议展览中心，体育馆等。此类场景下，可以通过算法得到人流的实时密度，获知他们动态的行为和状态，并且能够分析人群的下一步运动信息。通过对人群的实时动态监控，能够达到避免因密度过大引起灾祸的目的。

二、人群流动繁忙的公共交通场所。比如可以对地铁站、火车站以及重点公交站台等区域的乘客数量进行人流密度估计，就能够方便车站工作人员合理安排车辆调度，调整策略，最大限度的方便老百姓出行。

三、市场调查方面：对于超市、商场等新进的货物，通过检测货物柜台前的人流密度，可以得知一定时间以来对该产品顾客的关注程度，并将此信息反馈给商家，方便其及时调整营销策略，提高市场运营速度，提高收益。

四、智能化建筑设计：对现有的建筑中某些场所例如电梯，楼道口进行人流密度统计，

能够分析出哪些地方容易发生拥堵，从而帮助设计师对以后的设计方案进行优化改进，合理的调整设计面积布局，这在城市土地资源越来越稀缺的当下，具有举足轻重的意义。

## 1.2 人流密度检测研究现状

国外对人流密度检测的研究主要兴起于上世纪八九十年代。人流检测的手段和其他许多学科一样，经历了人工检测和自动化分析检测的阶段，目前，仍在进一步前进发展中。早期人工检测的方法主要由人工完成，在重要场所安装视频监控系统，操作者根据目测所得结果并结合行人身高、性别、年龄等信息判断行人的运动速度和方向等信息。但是人工检测的方法结果准确性直接取决于操作者经验，并且很受周围条件干扰，若当操作者疲劳时，容易忽视监视器上的突发情况，漏掉重要信息，导致不可挽回的后果。

随着图像处理技术的不断发展，视频分析理论的不完善，众多的基于视频信息的处理算法不断涌现出来。也正是因为这些算法的出现才使许多视频监控系统得以投入使用。在人流密度估计上主要有 Davies, Chow<sup>[1,2]</sup>等人提出基于监控区域背景图像的方法。香港地铁采用了基于该算法的人流密度估计的监控系统，不过这套系统监控范围小，分类器结构比较复杂。而 Rourke<sup>[3]</sup>等人的人流运动估计方法需要提取行人的个体信息，主要应用在低密度人群的情形，如果人流密度较大，则容易出现个体间的相互遮挡，使得个体信息提取不完整，这又影响了方法的准确性。为了解决高密度情况下的问题，Marana 提出了基于纹理分析的人流密度估计的算法。该算法依据的是不同密度的人群图像对应着不同纹理模式。低密度人群纹理上表现为粗模式，高密度人群表现为细模式。基于纹理分析的算法解决了高密度人群情况下区分度差的缺点，目前仍在进一步发展中。现今，人流密度检测技术大体上可以分为基于像素统计的人流密度检测方法、基于纹理分析的人流密度检测方法，除此以外，还有诸如结合了支持向量机（Support Vector Machine, SVM）等新工具新手段的其他的方法。

## 1.3 研究内容及章节安排

本文研究的内容是基于监控视频动态信息的人流密度检测算法。通过分析和研究已有的算法，了解并掌握了近年来在人流密度领域的最新研究成果。深入探讨了 SDP 模型，并将聚集度的概念与人流密度结合应用，这也是本文的创新点。文章内容可以分为以下几个部分：

第一章绪论，探讨了人流密度检测的研究背景以及选题意义，研究了国内外人流密度检测的发展历程与现状。最后给出了本文的研究内容与章节安排。

第二章图像预处理技术，为下文人流密度研究中做一系列的技术铺垫。这些知识也是我

们研究数字图像的基础。

第三章特征提取与跟踪算法的研究，主要探讨了几种图像特征的概念，并分析了每种特征常用的特征提取与跟踪算法，之后重点研究了 KLT(Kanade-Lucas-Tomasi)跟踪算法。

第四章人流密度检测算法研究，完成了人流密度定义等级的划分，研究了前景提取的知识，主要分析并总结了两种主流的人流密度检测算法：基于像素统计的算法、基于纹理分析的算法，在基于纹理分析的算法中，又展开了讨论了灰度共生矩阵。

第五章 SDP 模型的研究与应用(Self-Driven Particles)。深入分析了聚集运动中常用的研究模型——SDP 模型，并探讨其相关性质，以及群体聚集算法的步骤。通过实验验证了聚集度概念的准确性，发现了其与人流密度之间的关系。最后将人流密度与聚集度的概念在实际应用中相结合，展示了广阔的应用前景。

第六章是对文章内容的总结，并就人流密度检测的研究提出了自己的展望。

## 第二章 图像预处理技术

在人流密度检测系统中，摄像头采集完视频信息后传输给视频处理模块进行分析处理。但是由于实际环境中，获取的视频信息容易受光照、大气等因素的影响，因此所得的图像与理想的结果有不少差距。若不做处理加工，直接用来提取人流特征，则会影响到后续对人流密度估计结果的准确性。对此，在采集完视频信息并数字化后，我们先要对图像进行预处理操作。包括图像灰度化，图像噪声消除，图像增强以及形态学处理等。经过预处理操作的图像，我们可以去除其中无关紧要的信息，提高人流密度检测结果的准确性。

### 2.1 数字图像处理概念

数字图像处理作为一门学科大约形成约上世纪 60 年代初期，是指将图像信号转换成数字信号并利用计算机对其进行处理的过程。包括了图像去噪、图像增强、图像复原以及图像分割等处理的手段和方法。数字图像处理应用领域广泛，在航空航天，生物医学，军事公安等方面都有着许多应用。

数字图像处理的目的主要分成三类：

- (1) 提升图像的视感质量，如对图像进行亮度或者彩色的变换，抑制或者增强某些分量，对图像实施几何变换等，以改善图像的质量。
- (2) 提取图像中某些对我们有帮助的信息。提取出来这些信息通常都能够提高计算机分析图像的速度。而提取这些信息的过程实际上是模式识别或者计算机视觉分析。提取的特征可以包括频域特征、灰度特征以及颜色特征等诸多信息。
- (3) 图像数据的转换、编码和压缩，方便图像的存储和传输。

不管图像处理的目的是什么，都需要借由计算机以及其他专用设备构成的图像处理系统完成对图像数据的读取、处理和输出。

我们通常用二维函数  $f(x, y)$  来表示一副图像。其中我们把坐标  $(x, y)$  称作数字图像的像素点。其中  $x$  表示图像上某一点的横坐标， $y$  表示纵坐标。位于该点的图像强度或者图像灰度用坐标点  $(x, y)$  上的函数值来表示。坐标点和函数幅值  $f$  均为离散有限的数值。

### 2.1.1 数字图像分类

根据采用何种表示方式，数字图像可以分为位图和矢量图两种。常见的位图有 BMP、JPG、GIF 等，它们都是用数字阵列表示的；而矢量图是使用矢量数据库来表示的，例如 PNG 等格式。

若以每个像素所代表的的不同信息来分类，我们可以将数字图像分成二值图像、灰度图像、RGB 图像、索引图像：

(1) 二值图像：每个像素的取值只有 0 和 1 两个选择的数字图像。一般而言，黑色用 0 表示，白色用 1 表示。二值图像的用处多是在描述图形或者文字的场所。其优点是占用空间少；缺点在于当用在人物或者风景的图像时，二值图像只能简单地描绘它们的轮廓，不能描述具体的细节。这时候要用更高的灰度级。

(2) 灰度图像：每个像素采样只有一种颜色的图像。在计算机领域中，这类图像的颜色通常表现为从最暗的黑色过渡到最亮的白色之间的任意一种颜色。理论上这个采样可以具有任何不同深浅的颜色，甚至还有可能不同亮度上具有不同的颜色。与只有黑白两色的二值图像不同，灰度图像在黑色与白色两者之间还存在许多不同深度级别的颜色。

(3) RGB 图像：一幅 RGB 图像本质上就是一个彩色像素的数组，其大小为  $M \times N \times 3$ 。其中每个位置的彩色像素点都是对应位置上彩色图像所对应的红、绿、蓝三个分量。RGB 图像也可以看做由三个灰度图像形成的“堆栈”，当这些“堆栈”输入到彩色显示器的红、绿、蓝输入端时，就会在屏幕上输出彩色图像。

(4) 索引图像：RGB 调色板中的下标直接设置为像素值的图像。在索引图像中像素值的大小与调色板数值之间是“直接映射”的关系。通常索引图像与调色板是一起存储的。装载图像时，索引图像也会和调色板一起装载。当提取完像素值后，根据像素值在调色板中的位置搜索对应的索引值。与 RGB 图像相比，索引图像的优点在于更加节省存储空间<sup>[15]</sup>。

### 2.1.2 像素间的基本关系

图像中的基本单元是像素，像素有一定的空间位置，在图像空间中按某种规律排列，相互之间存在着一定的联系。

(1) 相邻像素：位于坐标  $(x, y)$  的像素点 P 有水平方向和垂直方向共计 4 个相邻像素点。其坐标由公式 (2.1) 给出：

$$(x-1, y), (x+1, y), (x, y-1), (x, y+1) \quad (2.1)$$

这四个像素点集成为  $P$  的领域，用  $N_4(P)$  表示。 $P$  的 4 个对角线方向上的相邻像素坐标：

$$(x+1, y+1), (x+1, y-1), (x-1, y-1), (x-1, y+1) \quad (2.2)$$

可用  $N_D(P)$  表示此像素集。水平垂直方向以及对角线方向的 8 个点一起构成  $P$  点的 8 邻域，用  $N_D(P)$  表示。

(2) 联通性：用来描述图像中边界和区域的重要概念。两个像素是否联通可以用下述两个条件来判断：

- (a) 两个像素点处于相邻的位置；
- (b) 两个像素点的灰度值满足的一定的相似性法则。

对于灰度值为  $V$  的像素点  $p$  和  $q$ ，如果  $q$  在集合  $N_D(P)$  中，则说  $p$  和  $q$  是四邻接的。倘若  $q$  在集合  $N_D(P)$  中，则我们可以称  $p$  和  $q$  是八邻接的。

从坐标为  $(x, y)$  的像素  $p$  到坐标为  $(s, t)$  的像素  $q$  的路径是符合如下定义的像素序列，其坐标序列为： $(x_0, y_0), (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 。

其中满足  $(x_0, y_0) = (x, y), (s, t) = (x_n, y_n)$ 。并且像素  $(x_i, y_i)$  和  $(x_{i+1}, y_{i+1})$  是邻接的， $i$  满足条件  $(x_0, y_0) = (x, y), (s, t) = (x_n, y_n)$ 。假设  $S$  是一副图像中像素的子集，若  $S$  中任意的像素点两两之间都存在着一一条路径，那我们认为  $p$  和  $q$  是联通的。

(3) 区域和边界：令  $R$  是图像中某一个像素的子集，如果子集  $R$  是联通的，则称作  $R$  是一个区域。区域的边界是指区域中全部的包含单个或多个相邻像素所构成的集合，这些像素不在区域  $R$  中的。区域的边缘则是由部分特定像素点的一阶或者二阶导数组成的。是一个像素直接邻域的局部性质的反应。它是一个矢量，有着方向和大小的特性<sup>[16]</sup>。

## 2.2 图像预处理

在现实应用中，视频采集设备收集到的图像信息容易受到外界因素的影响，若不进行预处理，得到的图像将与我们理想中的结果相差甚远。图像预处理的主要目的是为了剔除图像中我们不需要的无关信息，尽量还原我们需要的有用信息，从而达到有效提取图像特征的目的。图像预处理一般包括图像的灰度化、图像去噪、图像增强和形态学处理等。

### 2.2.1 图像的灰度化

出于人流密度检测系统对实时性要求高的特点,我们需要将视频中的 RGB 图像转换成灰度图像。这不仅能够节约存储空间,还能够提高运算效率。虽然灰度图像一定程度上舍弃掉了图像中的某些信息,但是整幅图像色度和特征等重要信息依旧可以通过灰度图像来反映。在 RGB 模型中,  $R=G=B$  表示为仅有一种颜色的灰度图像,其中  $R(G,B)$  值的大小称为灰度值。所以在灰度图像中,一个像素的灰度值存储空间仅需要一个字节。灰度图像的取样范围从黑到白,对应的灰度值范围介于 0 到 255 之间。灰度值能够反映出颜色的深浅程度。故灰度图像相比于只有黑白两色的二值图像,具有更大的颜色变化范围。常用的图像灰度化处理有四种方法:

(1) 分量法: 灰度图像的灰度值是从 RGB 图像中三个分量的亮度值得到的。能够根据实际需求灵活地选取某种灰度图像。

$$f_1(i, j) = R(i, j) \quad f_2(i, j) = G(i, j) \quad f_3(i, j) = B(i, j) \quad (2.3)$$

其中  $f_k(i, j)$  ( $k=1,2,3$ ) 代表的是转换完成后原图像在坐标  $(i, j)$  处的灰度值。

(2) 最大值法: 其原理是灰度图像的灰度值是由 RGB 图像中三个分量里面最高亮度的分量得来的。

$$f(i, j) = \max(R(i, j), G(i, j), B(i, j)) \quad (2.4)$$

(3) 平均值法: 灰度图像的灰度值是由 RGB 图像中三个分量的亮度均值求得的。

$$f_1(i, j) = (R(i, j) + G(i, j) + B(i, j)) / 3 \quad (2.5)$$

(4) 加权平均法: 根据不同分量重要性的大小以及一些其他的因素,赋予三个分量不同的权值(如红色取 0.30, 绿色取 0.59, 蓝色取 0.11)以计算每个分量的加权平均值。由于人眼对绿色来说更加敏感,而对蓝色敏感度较低,因此通过计算 RGB 这三个分量的加权平均就能得到相对准确的灰度图像。加权公式如下:

$$f(i, j) = 0.30 * R(i, j) + 0.59 * G(i, j) + 0.11 * B(i, j) \quad (2.6)$$

### 2.2.2 图像去噪

噪声是产生图像干扰的重要原因<sup>[17]</sup>。图像的获取和传输是产生噪音的主要过程。为了方

便描述，我们假设原始图像用  $f(x,y)$  表示，图像信号用  $g(x,y)$  表示，噪声用  $n(x,y)$  表示。

通常所见的噪声主要有下面四种形式：

(1) 加性噪声：是指与输入图像信号没有关系的噪声。含有加性噪声的图像可以这样表示： $f(x,y) = g(x,y) + n(x,y)$ 。比如信道噪声就是加性噪声的一种。

(2) 乘性噪声：与加性噪声和输入图像没有关系不同，乘性噪声是和输入的图像信号紧密相关的。可以用  $f(x,y) = g(x,y) + n(x,y) * g(x,y)$  表示含有乘性噪声的图像。比较常见的该类噪声有：产生于胶片的颗粒噪声和电视机产生的噪声。

(3) 量化噪声：此类噪声与输入图像信号也没有关系。不过不同的是量化噪声的大小反应的是数字图像和原始图像的差异。

(4) 椒盐噪声：该类噪声经数字化处理成最大值（非黑即白）。它是唯一一种外观上能够引起退化的并且视觉上能够区分的噪声。

常见的图像去噪方法有三种：

(1) 均值滤波器：主要包括了算术均值滤波器、几何均值滤波器、谐波均值滤波器以及逆谐波均值滤波器。其中最简单的均值滤波器便是算术均值滤波器。它用某一个像素点邻域内灰度的算术平均值代替该点的像素值。不管该滤波器有模糊图像的缺点。而几何均值滤波器在平滑图像时比算术均值滤波器处理的图像丢失的细节更少。谐波均值滤波器比较擅于处理椒盐噪声以及高斯噪声，但不适合处理胡椒噪声。而脉冲噪声带来的影响更适合用逆谐波均值滤波器来消除。但它缺点在于事先一定要知道处理的是暗噪声还是亮噪声，否则无法选择合适的滤波器阶数符号。如果选择不当，偏大或者偏小，都可能会导致意想不到的后果。总体来说，高斯或均匀随机噪声的情形适合用算术均值滤波器或者几何均值滤波器处理（后者更佳）。若是需要处理脉冲噪声，逆谐波均值滤波器更宜。

(2) 统计排序滤波器：统计排序滤波器主要包括了最大值和最小值滤波器、中点滤波器、修正的阿尔法滤波器、中值滤波器。最大值滤波器比较适宜处理胡椒噪声，原因是此类滤波器能够迅速地发现图像中最明亮的像素点。最小值滤波器适宜处理盐粒噪声，原因是它能够准确的发现图像中的最暗的像素点。顾名思义，中点滤波器简略地计算被滤波器所包围区域中介于最大值和最小值之间的中点，这类滤波器适合处理类似高斯噪声这样的随机分布噪声。修正的阿尔法滤波器在多种噪声类型混合的情形适合能够获得很好的效果。中值滤波器应用范围十分广泛，它是用一个像素点邻域中的灰度级大小的中值替换该点像素的值。相比于其他同尺寸的滤波器，中值滤波器产生的模糊更少，在存在单级或双脉冲噪声的情形时特别适



用。

(3) 自适应滤波器：自适应滤波器的综合性能是优于其他滤波器的，不过该滤波器相对地也比较复杂。自适应滤波器可以进一步分为自适应中值滤波器和自适应局部降低噪声滤波器。这里面自适应中值滤波器不仅能够处理具有大概率的脉冲噪声而且它还可以在平滑非脉冲噪声时候保留细节。和一般的滤波器一样，它的工作范围也在矩形区域内，不过它在滤波工作的时候会根据实际条件灵活地调整矩形区域的大小。

### 2.2.3 图像增强

改善并提高图像的视觉效果是图像增强最重要的目标。对于事先设定了应用场景的图像，为了突出图像的局部或者整体特性，需要强化图像中不同对象之间特征的差异性，从而明显地提高图像质量，改善图像的识别能力，最终实现分析图像的目标。

目前主流的图像增强方法可以分成两种：频率域法和空间域法。其中频率域法是指把图像看成简单的二维信号，对它进行二维傅里叶变换。其中，为了通低频信号，采用低通滤波器，可以有效地去除图像中的噪声；而为了通高频信号，采用高通滤波器，能够使模糊的图像变的清晰。空间域算法主要有两类：中值滤波法和局部求平均值法。它们可以有效的去除或抑制噪声。

### 2.2.4 形态学处理

图像处理中的形态学处理<sup>[18]</sup>是指通过一系列数字形态学操作帮助我们从图像中提取出能够描绘和表达图像区域形状的有用分量。比较典型的有用分量有边界、凹壳、骨架等。数字形态学主要使用结构元素探测二值图像形成的集合，其实质是通过结构元素和图像集合之间的相互作用。二值图像是形态学处理中我们感兴趣的研究对象，针对二值图像常用的形态学处理方法有膨胀运算、腐蚀运算、开闭运算四种。

(1) 膨胀运算：膨胀运算的运算符为 $\oplus$ ，结构元素  $B$  膨胀  $A$  可以如下定义：

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\} \quad (2.7)$$

该公式理解为以  $B$  关于原点的映射，并且对  $z$  的映像进行平移为基础的。其中  $\hat{B}$  为  $B$  的映射，定义是  $\hat{B} = \{w \mid w = -b, b \in B\}$ 。膨胀的操作步骤如下：首先对  $B$  做关于原点的映射，再将得到的映射平移  $x$ 。当平移  $A$  和  $B$  达到  $x$  个单位后的两者映射集合为空集时，膨胀集合

的像素可以认定为是 B 的原点。如果把 B 当作一个卷积模板，膨胀就是对 B 做关于原点的映射，接着将映射得到的结果在 A 上连续地平移实现的。膨胀运算如图所示 2.1 所示，图 (d) 中黑色的方框表示膨胀过后扩张的部分。

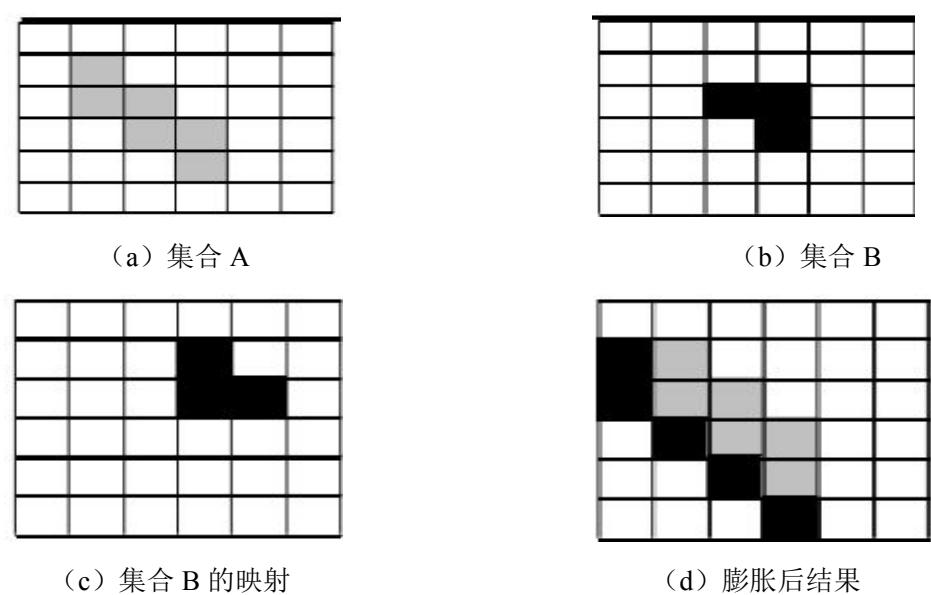


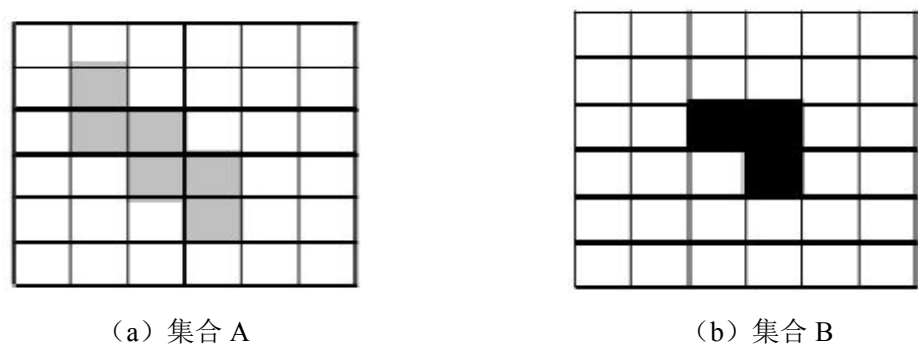
图 2.1 膨胀运算过程

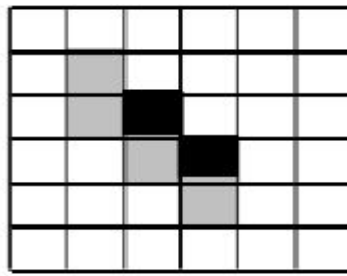
图 (d) 中黑色的方框表示膨胀过后扩张的部分。

(2) 腐蚀运算：腐蚀运算的运算符为  $\odot$ ，结构元素 B 膨胀 A 可以定义为：

$$A \odot B = \{ z | (B)_z \subseteq A \} \tag{2.8}$$

B 对 A 的腐蚀过程看做是把所有满足条件的 B 平移 x 单位后，B 仍然全部包含在 A 中的 x 的集合。腐蚀运算的过程如图 2.2 所示：





(c) 腐蚀后结果（黑色部分）

图 2.2 腐蚀运算过程

腐蚀运算和膨胀运算其实是对偶的，即：

$$(A \odot B)^c = A^c \oplus \hat{B} \quad (2.9)$$

$$(A \oplus B)^c = A^c \odot \hat{B} \quad (2.10)$$

(3) 开运算：开运算一般用于平滑物体轮廓、断开窄的瓶颈以及消除细的突出物等。B 对 A 的开运算实际上就是先 B 对 A 腐蚀操作，再用 B 对腐蚀的结果进行膨胀，也即：

$$A \circ B = (A \odot B) \oplus B \quad (2.11)$$

(4) 闭运算：闭运算虽然也会平滑轮廓的一部分，但与开操作相反的是，它通常用来弥合细长的沟壑和较窄的间断，消除小的洞孔，填补轮廓线中的裂痕。

闭运算实际上是对图像先膨胀后腐蚀的过程。结构元素 B 对集合 A 的闭运算的计算公式如下：

$$A \bullet B = (A \oplus B) \odot B \quad (2.12)$$

## 2.3 边缘检测

边缘是图像的最基本的特征，图像的轮廓、细节等基础信息都存在于图像的边缘之中。在图像分析处理中借助边缘不仅能够大大的减少需要处理的信息，减轻工作量，而且图像中物体的形状信息也不会丢失。因此边缘检测在模式识别、机器视觉和图像处理中有着十分广泛的应用。本质上讲，边缘是图像的局部特性间断性的反应，包括灰度、纹理、颜色以及结构突变等。区域在边缘结束，同时也在哪里开始了一个新的区域。

若两个相邻区域的灰度值不相等，则区域交界处存在边缘。灰度值的不连续性导致了边缘的产生，可以用导数的方法检测出这种不连续性。数学中边缘的检测常用一阶导数和二阶导数。边缘检测的基本步骤是：首先用边缘增强算子，增强图像的局部边缘，随后定义像素的边缘“强度”，利用设置好的阈值提取边缘点集。主流的边缘检测算法的思想是构造关于

原始图像中某个像素点小邻域的边缘检测算子。常用的有 Sobel 算子、Robert 算子和 Canny 算子等，下面将分别介绍这几种算子。

### 2.3.1 Sobel 算子

Sobel 算子是一种梯度算子，图像  $f(x,y)$  在位置  $(x,y)$  的梯度定义式如下：

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (2.13)$$

该算法是在图像空间中利用水平和竖直两个方向的模板分别与图像进行邻域卷积实现的。这两个方向的模板分别用来检测水平边缘和垂直边缘。模板系数为模板中的数字，边缘方向与梯度方向是垂直正交的关系。

$$M_A = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (2.14)$$

$$M_B = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (2.15)$$

式 (2.14) 与 (2.15) 为水平梯度方向和垂直梯度方向，可以分别用来检测垂直边缘与水平边缘。

以  $3 \times 3$  窗口为例，窗口像素和模板元素之间的对应关系定义如式 (2.16)：

$$[M]^i = \begin{bmatrix} M_{-1,-1}^i & M_{-1,0}^i & M_{-1,1}^i \\ M_{0,-1}^i & M_{0,0}^i & M_{0,1}^i \\ M_{1,-1}^i & M_{1,0}^i & M_{1,1}^i \end{bmatrix} \quad (2.16)$$

上式中  $i=1,2$  分别代表垂直方向和水平方向的模板。设窗口灰度为：

$$[f] = \begin{bmatrix} f(j-1,k-1) & f(j-1,k) & f(j-1,k+1) \\ f(j,k-1) & f(j,k) & f(j,k+1) \\ f(j+1,k-1) & f(j+1,k) & f(j+1,k+1) \end{bmatrix} \quad (2.17)$$

模板卷积过程实际上就是求乘积的过程：

$$f_i(j, k) = \sum_{m=-1}^i \sum_{n=-1}^i F(j+m, k+n) M_{m,n}^i \quad (2.18)$$

上式中等式左边为模板卷积法边缘检测的输出， $i = \frac{l}{2}$ （ $l$ 为窗口宽度）。对于  $3 \times 3$  窗口（ $l=1$ ），图像中对应模板中心新像素值即两个卷积的最大值，为：

$$f_{\max} = \max[f_i(j, k)] \quad i=1, 2 \quad (2.19)$$

判断一个点是不是边缘点可根据图像边缘附近邻域内灰度变化较大的特点，设置一阈值，当邻域内的像素新灰度值大于阈值时该点就作为边缘点。

Sobel 算子的边缘检测效果较好，而且对噪声有平滑作用，能够降低噪声对检测的影响。但是 Sobel 算子得到的结果边缘较粗，这是因为产生了伪边缘的缘故，这也降低了边缘检测的准确度。故当在对边缘检测结果要求不高的情况下，采用 Sobel 算子还是很有效的。

### 2.3.2 Roberts 算子

图像灰度具有一阶导数快速变化到极值的特性<sup>[40]</sup>。而 Roberts 算子正是利用这一性质来进行奇异点的检测。它在某一点上的数值正好表示的是这一点的梯度值。任意两个方向上互相垂直的差分算子都可以近似的用来求取梯度。Roberts 利用了这一原理，提出了 Roberts 交叉算子。

一幅连续的图像假设用  $f(x, y)$  表示，为了计算某一点梯度值的大小以和方向，需要求得该点的偏导数。而又由于数字图像是离散的，于是能够用一阶差分代替一阶偏导。图像的二维离散函数在水平方向的一阶差分函数定义为：

$$f(x+1, y) - f(x, y) \quad (2.20)$$

相应地，竖直方向上的一阶差分定义为：

$$f(x, y+1) - f(x, y) \quad (2.21)$$

Roberts 算子利用局部差分法来寻找图像的边缘，因而比较简单。它检测边缘的原理是采用像素的梯度幅值近似于对角线方向上相邻两像素差的绝对值。即：

$$\text{grad}(x, y) = |f(x+1, y+1) - f(x, y)| + |f(x, y+1) - f(x+1, y)| \quad (2.22)$$

其边缘检测算子是一个  $2 \times 2$  的模板，水平方向和垂直方向的两个卷积模板如表 2.1 所示：

表 2.1 Roberts 算子卷积模板

1	0	0	1
0	-1	-1	0

Roberts 算子定位边缘精度较高，但是也易对噪声敏感。适用于有明显边缘且只有少量噪声的图像分割场景。Roberts 算子使用局部差分法来求得图像的边缘，不过使用 Roberts 算子处理后的图像结果中边缘存在不平滑的缺点。原因是 Robert 算子通常会在邻近图像边缘的区域内产生较宽的脉冲响应，所以在使用 Roberts 算子检测前常需做细化处理，这反过来也影响了边缘定位的精度。

2.3.3 Canny 边缘检测

上述的两种 sobel 和 Roberts 边缘检测算子都是使用模板对图像中每个像素点的邻域依次做卷积运算，但是主要存在着两个缺点：一是容易对噪声敏感，二是定位的边缘精度不高。Canny 针对上述缺点，提出了边缘检测准则，该准则主要包括三个方面的指标<sup>[19]</sup>：

- （1）低误判率：即将正确的边缘错误地判断成非边缘点的概率低；
- （2）高定位性能：检测出的边缘点要最大可能地处在实际边缘点的中心位置；
- （3）虚假响应低：尽量避免虚假响应边缘的产生。

Canny 边缘检测算法首先采用高斯滤波器过滤平滑阈值化后产生的二值图像，然后用“非极大值抑制”平滑过滤后的图像，之后再使用双阈值法提取图像边缘。该算法流程大致分成四个步骤：

- （1）高斯滤波器过滤图像噪声。二维高斯分布函数为：

$$G(x,y)=\frac{1}{2\pi\sigma^2}e^{\frac{-(x^2+y^2)}{2\sigma^2}}$$

(2.23)

上式中方差  $\sigma$  控制图像的平均滑度。如果要平滑离散的数字图像，则需要对高斯分布函数离散化。而从理论上来说，若要实现这一目标，需要采用一个无穷大的卷积模板。为了简化运算，利用卷积特性可以得出当到中心的距离大于方差的 3 倍时，模板系数可近似为 0。表 3.2 给出的是当方差  $\sigma=1.4$  时近似高斯模板：

表 2.2 高斯模板

2	4	5	4	2
4	9	12	9	4
5	12	15	12	5
4	9	12	9	4
2	4	5	4	2

(2) 利用一阶微分模板求像素点的梯度大小和方向。使用如下的模板分别计算水平方向和垂直方向的一阶差分值。

表 2.3 一阶梯度模板

1	1
-1	-1

-1	1
-1	1

(3) 梯度的非极大值抑制：以梯度幅值大小和方向为依据确定图像边缘。为了提高边缘像素点提取的准确率，需要保留局部邻域内像素梯度最大的点。又为了去除同一区域内的非极大值，中心像素的梯度方向被映射成四个区域，这样方便使用 3\*3 窗口进行抑制。举例来说，如果中心点像素的梯度方向在第二区，则将其梯度大小与它左上与右下相邻的像素点的梯度值比较。如果是局部最大值，则可以认为该像素点是梯度像素点，否则，将此像素点的梯度值设为 0。

(4) 双阈值法检测并连接边缘：双阈值法的原理是设定一个高阈值  $a$  和一个低阈值  $b(b=0.4a)$ 。假设以  $a$  为阈值进行二值化后的图像为  $A$ ， $b$  为阈值进行二值化后的图像为  $B$ 。由于  $a$  值较大，所以所得的图像  $A$  虽然除去了大部分噪声，但是也无法避免的丢失了一些边缘信息；而图像  $B$  由于  $b$  值较小，虽然也保留了一些噪声信息，但是边缘信息保存的较好。把两者结合起来，以图像  $A$  为基础，图像  $B$  为补充，连接边缘。

实际应用中，Canny 算子检测的边缘获得了比 sobel、Roberts 算子更好的优点。并且还能够通过改变方差  $\sigma$  来调整边缘检测系统的抗噪能力和定位的准确性。因此获得了更好的应用。但是 Canny 算子在 2\*2 邻域内利用一阶差分均值来求求梯度的幅值，对噪声敏感，容易丢失真实的边缘信息；并且利用固定的高低阈值提取图像边缘，阈值的设置完全取决于人工获得，缺乏不同图像时的自适应性，自动化程度不高。

### 2.3.4 边缘检测总结

除了以上介绍的三种边缘检测方法，还有 LOG，prewitt 等边缘检测算子，这里不再详细介绍，只简单比较下文中提到的三种算子。

(1) sobel 边缘检测算子是对邻域范围内的像素加权，模板中心位置的数值较大。使用该算子检测的图像不仅能够产生较好的边缘，而且还起到对噪声的平滑作用。不过也存在着伪边缘，而且比较粗，并且定位的精度也不高。

(2) Roberts 算子中信号的突变是利用对角线方向上相邻两像素的差值来表示的。其水平方向和垂直方向的边缘检测性能优于斜线方向，获得的定位精度也比较高，不过其缺点是容易对噪声敏感，结果图像的边缘较细。

(3) Canny 算子为了获得较高的噪声抑制能力，采用了高斯函数对图像进行平滑处理；但这也平滑掉了部分高频边缘，造成边缘丢失；但是由于采用了双阈值法检测和边缘连接，获得的图像结果有较好的连续边缘。

## 2.4 本章小结

本章主要研究了图像预处理的一些基本知识。首先分析了数字图像处理的基本概念，包括数字图像的分类、像素间的位置关系；然后讨论了图像的灰度化、图像去噪、图像增强和形态学处理等几个基本的图像处理操作；最后探讨了边缘检测的知识，其中又重点分析了 sobel、Roberts、Canny 三种边缘检测算子并比较了它们的优劣性。



## 第三章 特征提取与跟踪算法研究

在视频监控的实际应用中，图像特征的提取十分重要。诸如阴影消除、前景检测等众多人流密度检测的研究步骤中都广泛的应用到了图像特征的提取。不论我们在研究人流密度的时候采用哪种算法，我们都要首先提取出视频中行人的特征信息，然后再根据行人特征并结合所用算法分析得出当前场景的人流密度信息。

### 3.1 图像特征分类

在数字图像分析中常用的图像特征包括纹理特征、形状特征、颜色特征、空间关系特征等。下面我们将简单介绍这四种常见的图像特征。

#### 3.1.1 纹理特征

##### （一）纹理特征的概念

纹理特征是一种全局特征，它能够描述场景里对应图像的表面特征，但是由于纹理只是物体表面的一种特性，它并不能完全反应出对象的所有固有特征。所有我们不能冀希望于仅仅通过图像的纹理特征就能得到更深层次的图像内容。纹理特征其实是统计各像素点邻域内像素。这种区域性的特征在模式匹配中具有较强的优越性，不会出现由于局部的偏差导致无法匹配的情况。纹理特征的通常是用于纹理信息差异明显的图像之间的检索。作为一种统计特征，其通常具有旋转不变性，并且有较强的抗噪能力。但是当图像分辨率较低的时候，纹理计算会有较大的偏差，或者纹理相差不大时，人的视觉感受很难区分纹理之间的细小差异，此时不再适合使用纹理特征区分图像。

监控场景中的人流情况十分复杂，不同密度的人流在图像中呈现出不同的纹理信息。特别是在高密度人流情况下，由于遮挡比较严重，若依旧使用传统的基于像素统计的特征方法估计人流密度，误差会很大。但如果使用基于纹理特征的人流密度检测算法会显著提高高密度人流情况下的准确率。

##### （二）特征提取与匹配方法

目前纹理分析方法主要包括四种：基于几何的方法、基于统计的方法、基于信号处理的方法和基于模型的方法。

（1）基于几何的方法：是一种建立在纹理基元基础上的纹理特征分析方法。纹理基元是最基

本的纹理元素。该理论认为若干简单的纹理基元可以以一定的方式有规律的重复构成复杂的纹理。Voronio 棋盘格特征法和结构法是两种比较有影响的几何方法。

(2) 基于统计的方法：灰度共生矩阵的纹理特征分析是该方法的典型代表。建立在研究共生矩阵大量统计特征基础上后，Kreyszig 和 Gotlieb 等人通过实验，得到了关于灰度共生矩阵的四大关键特征：熵、能量、惯量以及相关性。而利用自相关函数提取图像的纹理特征是另一种常见的统计方法。其中自相关函数实际上是图像的能量谱密度函数。通过对能量谱密度函数的计算我们能够提取出纹理的粗细程度和方向等诸多参数。

(3) 基于信号处理的方法：该方法中提取与匹配特征的手段主要有 Tamura 纹理特征、灰度共生矩阵、小波变换、自回归纹理模型等。其中 Tamura 纹理特征提出了对比度、粗糙度、线像度、方向度、规整度和粗略度共 6 中属性，这些属性的提出其实是从人类自身对纹理特征的视觉感知切入的；而灰度共生矩阵主要依赖于熵、能量、惯量和相关性这四个参数；而自回归纹理模型我们可以把它当作是马尔可夫随机场模型的实际应用。

### 3.1.2 形状特征

#### (一) 形状特征的特点

当从图象中分割出物体以后，将描述形状特征与测量尺寸有机结合能够区分出不同的物体，这样的方法在机器视觉中扮演着十分重要的作用。不仅如此，基于形状特征的检索方法能够高效地使用图像中感兴趣的目标来完成检索任务，不过它们也有一些共性的问题：①目前基于形状特征的检索方法缺少完整的数学理论模型②当目标形状变化时检索结果容易产生偏差③大量的形状特征只是简单描述了目标的局部特性，若想要完整的表达出目标的整体信息，对计算时间已经存储空间要求较高④很多形状特征反应出的目标信息与人眼的视觉感受不完全一致，甚至相差很大。⑤从二维图像中表现的三维物体实际上只是物体在空间上某一平面的投影，反应出的形状并不是物体的真实形状。由于观察角度的不同，会产生各种失真。

#### (二) 特征提取与匹配方法

一般而言，我们通常使用两类方法表示形状特征，一是轮廓特征，二是区域特征。轮廓特征主要位于图像边界，区域特征则关系到整个形状区域。边界特征法通过对图像边界特征的描述来获得图像自身的形状参数。

(1) Hough 变换：Hough 变换是一种用来检测平行直线以及边界方向直方图的经典算法。Hough 变换的大致过程是利用图像的全局特性，连接边缘像素用以构成封闭的区域边界。该方法的本质是利用了点与线的对偶性；而为了获得图像边缘，边界方向直方图首先对图像做

微分运算，然后，绘出关于边缘的直方图（包括大小和方向），计算图像灰度的梯度方向矩阵是其常用的方法。

（2）傅利叶形状描述法：该算法的形状描述是以图像边缘的傅利叶变换作为基础的，利用区域边界存在的周期性和封闭性这两个特点，将二维问题降维处理成一维问题。由边界点导出曲率函数、复坐标函数、质心距离三种形状表达。

（3）几何参数法：在几何参数法中，形状的匹配与表达采用更直观的简单的区域特征方法。如使用矩、面积、周长等参数。例如 IBM 的 QBIC 系统中，基于形状特征的图像检索即是采用偏心率、圆度、主轴方向以及代数不变矩等几个几何参数。

需要注意的是，提取形状参数，必须以图像处理及图像分割为前提，分割效果必然会影响到参数的准确性。如果图像的分割效果很差，形状参数提取误差将会很大甚至无法提取。

（4）形状不变矩法：该方法的描述参数是目标所占区域的矩的形状。

（5）其它方法：近几年来，在形状描述和匹配领域还取得了包括有限元法、旋转函数和小波描述符等新的进展。基于小波描述的方法首先用小波变换获得多尺度的图像边缘，由于每个尺度有 7 个不变矩，计算完这个 7 个不变矩之后再将它们转化为 10 个相对矩，而图像的特征向量即是所有尺度上的相对矩，从而将区域中封闭和不封闭的结构统一了起来。

### 3.1.3 颜色特征

#### （一）颜色特征的概念

同纹理特征一样，颜色特征也是一种局部特征，它同样能够描述场景里对应图像的表面特征。因为颜色对图像的大小、方向的变化不敏感，所以颜色特征虽然能够描述图像目标的局部特性，但是效果比较差。另外，查询时仅使用颜色特征，倘若数据量很大，常会将检索出许多无关的图像。

#### （二）特征提取与匹配方法

颜色特征提取和匹配的常用方法有颜色集、颜色矩、颜色相关图、颜色聚合向量和颜色直方图等集中。

（1）颜色直方图：通常来说我们常用直方图来间接表示各种颜色信息。因此颜色特征也叫做直方图特征。颜色直方图实际上是统计图像中不同颜色区间的像素数。它实现起来方法简单，获取容易，不受图像平移和旋转变化的影响。它能够简单描述图像中颜色的全局分布，即每种色彩在图像中的比列，适用于那些无需考虑物体空间位置关系以及难以分割自动分割的图像。不过它无法描述颜色的局部分布和不同颜色所处的位置空间，即不能描述具体的物体和

对象。难以处理运动较快的目标，此外在摄像机本身运动的情况下也不适宜使用。

常用的颜色空间有 RGB 颜色空间、HSV 颜色空间。颜色直方图的匹配方法有直方图相交法、中心距法、距离法、累加颜色直方图法、参考颜色表法。

(2) 颜色集：因为颜色直方图特征提取和匹配颜色特征的过程是全局的，无法区分局部的信息。为了解决这一问题，引入了颜色集。不过也可以近似认为颜色集就是颜色直方图。首先将图像从 RGB 颜色空间转换到诸如 HSV 空间的视觉均衡的颜色空间，并且把颜色空间量化成若干柄。然后用色彩分割技术把图像分割成若干区域，使用颜色空间量化后的某个色彩分量为每个区域建立索引。这样就可以用一个二进制的图像索引集表示图像。图像匹配可以使用对比不同颜色索引集之间的距离并且参考色彩区域的空间关系的方法进行。

(3) 颜色矩：颜色矩的数据基础在于可以用颜色的矩阵来表示该颜色的任意分布。另外，颜色的信息分布很多都聚集在低阶矩中，因此图像的颜色分布我们只需要使用一、二、三阶颜色就可以表示。

(4) 颜色聚合向量：主要思想是将直方图柄中每个柄的像素分成两个部分，如果该柄内某些像素所占的区域面积大于阈值，则这些像素当作聚合像素，否则是非聚合像素。

### 3.1.4 空间关系特征

#### (一) 空间特征关系概念

空间关系特征的概念是指在图像分割后多个目标之间在位置或者方向上的关系。这些关系有邻接、重叠、包含等。空间位置信息通常可以分为绝对空间位置信息以及相对空间位置信息这两类。绝对着重的是目标间距离长度以及方向等信息。而相对则是目标之间诸如如上下左右的相对关系。很明显的可以看出，两者之间相对空间位置可以由绝对空间位置引出，但是相对空间位置在表达起来通常比较简单。

空间关系特征的使用可以帮助我们强化对图像内容的描述区分，但它容易对图像或目标的旋转或者尺度变化等位置信息的变动敏感。另外，真实的应用环境中，单独使用空间信息是不可取的，无法完全地描绘出表达场景信息。因此为了实现检索图像的目的，除了空间关系特征外，还需要结合其它特征。

#### (二) 特征提取与匹配方法

空间关系特征中提取图像特征主要有两种方法，一种是首先分割图像，划分出图像中包含对象或颜色的区域，之后根据这些划分好的区域提取图像特征，并建立索引；另一种方法则比较简单。均匀地将图像分成若干有规则的子块，之后提取出每个子块的特征，并建立相

应的索引。

## 3.2 KLT 算法

学术界有许多研究特征跟踪的方法，例如粒子滤波(pf)、meanshift 跟踪，以及 KLT 跟踪（或叫 Lucas 光流法）。这些方法各有各的特点，举例来说粒子滤波虽然它能够比较好的在全局搜索到最优解，但求解速度相对较慢，并且由于其计算是基于颜色直方图的，所以对相同颜色的东西区别的效果不是很明显；meanshift 虽然运行速度快，但是容易陷入局部最优的情况。因而 KLT 算法实时性的特点使其应用广泛。

### 3.2.1 特征点概念

除了 3.1 节介绍的图像特征以外，近些年以来，研究者也开始利用特征点进行人流密度检测的研究，并取得了良好的效果。特征点跟踪在视频图像处理中有着十分重要的作用。比如视觉导航、电视电影的后期制作、目标识别等领域都有着广泛的应用。因此如何实现对特征点的准确跟踪成为一个很重要的课题。现今，研究特征点跟踪的方法主要有两种：第一种方法建立在长序列的基础上，假设运动的目标不会发生剧烈的变化，通过记录特征点的运动得到目标的最佳运动轨迹。它主要用在目标缓慢平稳变化的场合，在摄像头位置无法固定或者目标的运动姿态飘忽不定的情形不再适合；二是通过处理相邻帧间的连续性信息，得到特征点的最佳匹配位置。而 KLT 算法就是一种典型的利用相邻帧的连续信息进行处理算法。

特征点数量与行人数量大致成线性关系。特征点主要用在运动物体的跟踪等。文献<sup>[22]</sup>直接利用统计特征点的数量估计场景中人数。文献<sup>[23]</sup>首先对场景中运动前景的特征点分类，然后对每一类特征点数量进行等参数的 SVM 训练，从而最终估计出场景中的人数。本文的研究算法采用的正是基于 KLT 算法提取的特征点的人流密度检测。

目前，学术界中还没有给特征点一个准确的定义，不同的文献叫法不一，比如兴趣点，显著点等。最简单的图像特征就是用点的位置来表示的。事实上，特征点既反应了一个点的位置关系，也同时体现了它在局部邻域范围内呈现出一定的模式特征。我们可以将特征点可以分为狭义特征点和广义特征点两类。这其中狭义特征点所处的空间位置本身只是如角点、交叉点等常规的属性意义。而相比之下，广义特征点的定义是基于区域的，它自身的位置不代表特征属性，只表示符合一定特征条件的区域的位置。某个特征区域的任一相对位置也可以是广义特征点。这种特征不一定必须是物理意义上的特征，当符合特定的数学描述时就可以认定其为特征，因而有时是抽象的。综上，从实质上讲，可以将广义特征点认为是一个抽

象的特征区域。特征区域具有的属性就是广义特征点的属性。之所以把它称作点，是为了将其简化为一个位置的概念，易于理解和研究。

不过在已发表的国内外文献著作中，狭义特征点多用于计算机视觉分析、图像匹配和模式识别中。而广义特征点则多在图像配准的场合下使用。无论是基于狭义特征点还是广义特征点，两者的方法和步骤基本是一样的，都首先要检测出特征点，之后再匹配特征点（包括特征描述），接着利用特征点匹配的结果得出图像之间变换关系，最后的操作是对图像进行变换和融合。

虽然狭义特征点自身所处的空间位置仅仅具有常规的属性意义，但无论是狭义还是广义特征点，在特征匹配的过程中不一定必须用到具有特征点本身属性的特征。前面说过，特征点既反应了一个点的位置关系，也同时体现了它在局部邻域范围内呈现出一定的模式特征。事实上，特征点标示了一个具有某个或者某些特定特征的局部区域的位置，把它称作点，是将其简化为一个位置甚至是坐标的概念，以便于确定同一个坐标点在两幅不同图像之间的对应关系而为图像匹配提供准备。所以，特征匹配的过程概括起来说是以该特征点为中心，匹配该特征点邻域内的局部特征。换言之，是在特征匹配之前首先要为这些特征点（包括了狭义特征点和广义特征点）建立特征描述，通常把这些特征描述叫作描述符。在基于特征点的图像匹配和拼接技术中，描述符拥有和特征点相同的地位，倘若无法得到一个好的描述符支持，即使是最精确的特征点也很难匹配到正确的结果。

由于传统的匹配算法通常直接提取边缘或者角点，易受光照、位置等外在环境因素的影响，所以随着对特征点研究的不断深入，出现了一些新的研究方法手段，比较有名的是 SIFT 和 SURF。

SIFT(尺度不变特征变换)算法是由 British Columbia 大学 David G.Lowe 教授 1999 年首次提出并于 2004 年加以完善。这种算法的主要优点有：SIFT 特征属于图像的局部特征，具有旋转不变的特性，并且对噪声的影响具有一定的稳定性；独特性好，适合在大容量的数据库中进行匹配操作和检索；多量性，即使只有少数对象的图像也能够产生许多 SIFT 特征向量；优化之后的 SIFT 算法在速度要求上能够达到一定的标准；可扩展性，能够和其他类似的特征向量整合；SURF 算法由 Bay 于 2006 年提出，其可以看成 SIFT 算法的增强版；它计算量小，运行速度快，提取出的特征点几乎与 SIFT 的结果相同。两个算法的比较如下表所示：

表 3.1 SIFT 与 SURF 算法比较

	SIFT	SURF
特征点检测	用尺度不一的图片与高斯函数做卷积运算	用不同大小的 box filter 与原始图像做卷积，易于并行
描述符生成	20*20 的像素区域划分成 4*4 或 2*2 的子区域，每个子域计算直方图	20*20sigma 单位的区域划分 4*4 子域，各个子域计算 5*5 个样本点 Haar 小波响应
方向	在邻接特征点的矩形区域内，利用梯度直方图计算	邻接特征点的圆域内，计算 x、y 方向上的 Haar 小波响应

3.2.2 KLT 算法发展

KLT 算法全称是叫做 Kanade-Lucas-Tomasi 算法。时间上要追溯到 Kanade-Lucas 两人在 20 世纪 80 年代发表的论文：An Iterative Image Registration Technique with an Application to Stereo Vision，文章里介绍了一种图像点定位的方法。该方法将图像匹配的传统问题，从通常的搜索滑动窗口的方法创新地转变为一个求解偏移量  $d$  的过程。后来 Jianbo Shi 和 Carlo Tomasi 两人发表了一篇名为 Good Features To Track 的文章。随着越来越多的研究者加入，KLT 算法也在不断发展完善并被国内外研究者应用到许多场景中。如 Ajmal S.Mian 等利用 KLT 算法成功完成了对结构简单的飞机的跟踪<sup>[24]</sup>；文献<sup>[43]</sup>中戢玲玲、陈黎、张晓龙利用 KLT 算法应用在了驾驶员的疲劳检测中。本文主要描述的是在求解  $d$  的过程中，哪些情况下可以保证一定能够得到  $d$  的解，该算法不仅在跟踪方面表现不俗，而且实时计算速度也比较快。KLT 算法得到的特征点，是很多点的轨迹“trajectory”，并且还伴随着一些漂移的点，所以，得到跟踪点的结果后要再进一步处理。

3.2.3 KLT 算法原理

假设待跟踪的窗口为  $W$ ，KLT 算法是一种计算该窗口在视频帧间的灰度差平方和 (Sum of Squared intensity Differences) 作为标准的跟踪算法。

对于灰度图像，在 KLT 算法中我们假定一个特征窗口  $W$ ，它包含了特征纹理信息。在  $t$  时刻图像帧  $I(x,y,t)$  和在  $t+\tau$  时刻  $I(x,y,t+\tau)$  中的位置满足

$$I(x, y, t + \tau) = I(x - \Delta x, y - \Delta y, t) \quad (3.1)$$

即属于下一时刻帧  $I(x, y, t + \tau)$  的每一个像素点都可以由上一帧  $I(x, y, t)$  中响应窗口的像素点平移  $d = (\Delta x, \Delta y)$  个单位得到。这也正如上一小节所述的, KLT 算法就是为了寻找求解  $d$  的过程。

假设在  $t + \tau$  时刻的特征窗口  $B(x) = B(x, y, t + \tau)$ , 其中  $X = (x, y)$  为视口坐标。 $t$  时刻的特征窗口为:  $A(X - d) = A(x - \Delta x, y - \Delta y, t)$ 。

在一般情况下有:

$$B(X) = A(x - d) + n(X) \quad (3.2)$$

$n(X)$  为  $\tau$  时间内由光照变化产生的噪声。在整个窗口上对  $n(X)$  的平方积分, 于是就可以求得窗口图像的灰度差平方和(SSD)。

$$\begin{aligned} \varepsilon &= \iint_V n(X)^2 w(X) dX \\ &= \iint_V [A(X - d) - B(X)]^2 w(X) dX \end{aligned} \quad (3.3)$$

其中, 通常  $X = [x, y]^T$ ,  $d = [dx, dy]^T$ ,  $w(X)$  的取值为 1。如果想要强调中心部分纹理的作用结果,  $w(X)$  应该采用高斯分布函数。

当  $d$  远小于  $X$  时候, 将  $A(X - d)$  泰勒展开, 舍去高次项, 得到:

$$A(X - d) = A(X) - g \bullet d \quad (3.4)$$

将 (3.4) 式代入到 (3.3) 并同时 (3.3) 式等号左右两边对  $d$  求导并令结果为 0, 得到:

$$\frac{\partial \varepsilon}{\partial d} = \iint_V [A(X) - B(X) - g^T d] \times g w(X) dX = 0 \quad (3.5)$$

这时,  $\varepsilon$  取极小值。式 (3.5) 又可以变换为:

$$\begin{aligned} &(\iint_V g g^T w dX) \\ &= \iint_V [A(X) - B(X)] g w dX \end{aligned} \quad (3.6)$$

假设有



$$\begin{aligned}
Z &= \iint_V g(X)g^T(X)w(X)dX \\
&= \iint_V \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix} w(X)dX \\
&= \begin{bmatrix} G_{xx} & G_{xy} \\ G_{xy} & G_{yy} \end{bmatrix}
\end{aligned} \tag{3.7}$$

$$\text{并且有 } e = \iint_V [A(X) - B(X)]g(X)w(X)d(X) \tag{3.8}$$

则 (3.6) 式可以改写为:

$$Zd = e \tag{3.9}$$

对于前后两帧图像, 通过求解式 (3.8) 能够得到特征窗口的位移  $d = (\Delta x, \Delta y)$ 。

在 KLT 跟踪算法中, 即使使用了包含纹理信息的窗口跟踪并不一定能取得很好的效果。对于某一个特征窗口, 当其矩阵值  $Z$  的两个特征值  $\lambda_1, \lambda_2$  如果满足条件  $\lambda_2 > \lambda_1 > \lambda_{\max}$ , 则此特征窗口跟踪起来具有良好的效果。阈值  $\lambda_{\max}$  和拍摄时候所处的环境有关, 可以通过实验确定。

展开 (3.8) 式得到:

$$\begin{aligned}
e &= \iint_V \begin{bmatrix} (A(X) - B(X))_x g_x w_x(X) \\ (A(X) - B(X))_y g_y w_y(X) \end{bmatrix} d(X) \\
&\begin{bmatrix} \iint_V (A(X) - B(X))_x g_x w_x(X) d(X) \\ \iint_V (A(X) - B(X))_y g_y w_y(X) d(X) \end{bmatrix} \\
&\begin{bmatrix} E_x \\ E_y \end{bmatrix}
\end{aligned} \tag{3.10}$$

将其代入 (3.9) 得到

$$\begin{bmatrix} G_{xx} & G_{xy} \\ G_{yx} & G_{yy} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} E_x \\ E_y \end{bmatrix} \tag{3.11}$$

解方程 (3.11) 得到我们需要求解的位移量  $d$  的坐标

$$\Delta x = \frac{\left( E_x - \frac{G_{xx}}{G_{yy}} E_y \right)}{\left( G_{xx} - \frac{G_{xy}^2}{G_{yy}} \right)}$$

$$\Delta y = \frac{\left( E_x - \frac{G_{xx}}{G_{yy}} E_y \right)}{\left( G_{xy} - \frac{G_{xx} G_{yy}}{G_{xy}} \right)} \quad (3.12)$$

### 3.3 结合 KLT 算法的特征跟踪概述

为了减少运动对图像处理质量的影响，我们在图像序列中建立了一个仿射变换的模型。而仿射模型的建立需要在图像上选择特征点并跟踪它们，通过拟合所有特征点偏移量得到整幅图像的仿射变换参数，从而建立起图像间的仿射关系。因此特征的跟踪与提取成为解决问题的关键。

图像处理中的所指特征提取概念实际上是利用计算机视觉的理论知识分析并处理外部世界的过程。即利用图像的边缘、纹理、颜色等特征来描述、处理图像。我们可以通过图像边缘的灰度变化率、纹理基元的排列顺序、图像边缘的几何特征来选择特征点。总体来说，特征点跟踪与提取的两个基本问题是：

- (1) 如何选取适合跟踪的特征点；
- (2) 如何实现帧与帧之间特征点的连续跟踪

国内外的研究者提出了很多关于特征提取与跟踪的方法，如基于图像的 Laplace 变换方法；基于灰度图像边缘的一阶导数二阶导数的方法。而本文里我们采用了基于 KLT 算法的特征提取与跟踪算法。

#### 3.3.1 图像特征点的选取

由于不易实现单个像素的特征点跟踪，因此需要跟踪的匹配窗口应该具有足够的纹理特征信息。系统根据图像的拐角、光线的明显变化等纹理特征，采用拉普拉色变换，得到图像的高频分量，而这些高频分量又是各个特征匹配窗口的中心。假设  $\lambda$  是设定的阈值， $\lambda_1$  和  $\lambda_2$  是匹配窗口的特征取值范围，并且满足  $(\lambda_1, \lambda_2) > \lambda$ 。此时，对于每一个特征匹配窗口而言，应该选择高频分量值最大的点，选取作为匹配窗口内的特征点。同时为了降低噪声的影响，通过设置一个独立的参数来控制特征点的分布，这个参数即是特征点的最小距离。

### 3.3.2 图像特征点的跟踪

假设  $I(x, y, t)$  表示  $t$  时刻图像序列中  $(x, y)$  点处的灰度值，则图像的偏移可以表示为：

$$I(x, y, t + \tau) = I(x - \xi, y - \eta, t) \quad (3.13)$$

而偏移量可以用  $d(\xi, \eta)$  表示。为了保证跟踪的窗口具有足够的纹理特征信息，并且假设该窗口内的灰度变化是线性的，为此采用如下的两种方法解决特征点跟踪：

(1) 偏移量检测：用来判断前后两帧是否为同一帧，两帧如果变化太大，则抛弃该窗口；

(2) 仿射变换：由于特征匹配窗口中不同点呈现不同的运动特性，所以一般在特征窗口的仿射变换中，假设窗口中每个像素的运动特性是相似的，用偏移量  $d$  来表示。

用公式表示这个跟踪偏移量的方法，定义：

$$J(X) = I(x, y, t + \tau) \quad (3.14)$$

$$I(x - d) = I(x - \xi, y - \eta, t) \quad (3.15)$$

其中  $X = (x, y)$ ，则式 (3.14) 可以改写为  $J(X) = I(x - d) + n(X)$ ， $n(X)$  表示噪声。留数  $\varepsilon$  的定义为：

$$\varepsilon = \iint_W [J(Ax + d) - I(x)]^2 w(x) dx \quad (3.16)$$

参数说明： $W$  为需要求解的特征匹配窗口， $w$  为加权函数，缺省时值为 1 或者类高斯函数，目的是在于突出特征匹配窗口  $W$  的中心区域。此时，为了保证能够有效的跟踪每个窗口的特征点，只需要最小化留数  $\varepsilon$  即可。

### 3.3.3 特征点检测和跟踪算法分析

在 3.3 节的开头，我们叙述过特征提取与跟踪主要分为两个问题，即如何选择合适的特征点以及如何实现帧与帧之间这些点的连续跟踪。通过 KLT 算法，在相邻两帧之前定义特征匹配窗口，建立两帧之间大量相对应的特征点，可以得到对应的特征点的集合。利用这些特征点集合并结合最小二乘法拟合得到图像间的仿射变换模型。这样在前一帧图像里就得到了后一帧图像的映射。本文为了表面 KLT 算法能降低因物体的运动而给检测带来的干扰，摄像头录制的场景中只有待检测的物体，其他物体均静止不动。否则如果对前后两帧直接进行时间差分检测物体的运动，物体的运动带来的噪声会导给检测结果致不利的影响。

检测到的场景中的物体运动，实际上是由含待测物体的运动和场景中环境的运动这两个

相互独立的子运动组成的。采用 KLT 算法检测，目的之一就是从运动的背景环境中补偿出需要检测的物体。同时利用 KLT 算法检测到的特征点在运动区域和不运动的区域都有分布。为了分析跟踪的特征点的偏移量，我们选取图像序列两帧之间跟踪成功的特征点，这些特征点必须是偏移量最大的点。通过分析发现，这些偏移量最大的点大多分布在背景中带有“拐角”的纹理中，相对地，运动物体上的特征点分布较少。原因是运动的物体的面积占整个图像的比例比较小，并且它们自身的纹理较为平滑，部分背景相对拐角点更为明显。比较计算背景与物体的特征点的平均偏移量发现（用每个特征点的偏移量相加得到的总合除以特征点个数），背景上的特征点运动不及物体上的特征点。

由于运动物体的特征点与背景的特征点在一定程度上呈现差异，所以可以利用物体上分布的特征点的运动信息，来确定物体在下一帧图像中的运动区域。这可以降低噪声以及待测物体的运动对检测效果的影响，进而提高测得结果的准确率。但是，特征点在运动物体上的分布比较少，因此，若仅凭跟踪分布在运动物体上的特征点对运动物体的区域进行检测，无法包含完整的区域。而且容易受到图像上其他特征点的影响，导致出现较大误差。

本文出于这样的考虑，利用了物体特征点区域上的信息，跟踪上一帧物体上的特征点在该当前的位置，得到一个特征点区域。再通过全局仿射变换得到特征运动区域。检验该区域，降低噪声以及运动对结果的影响。

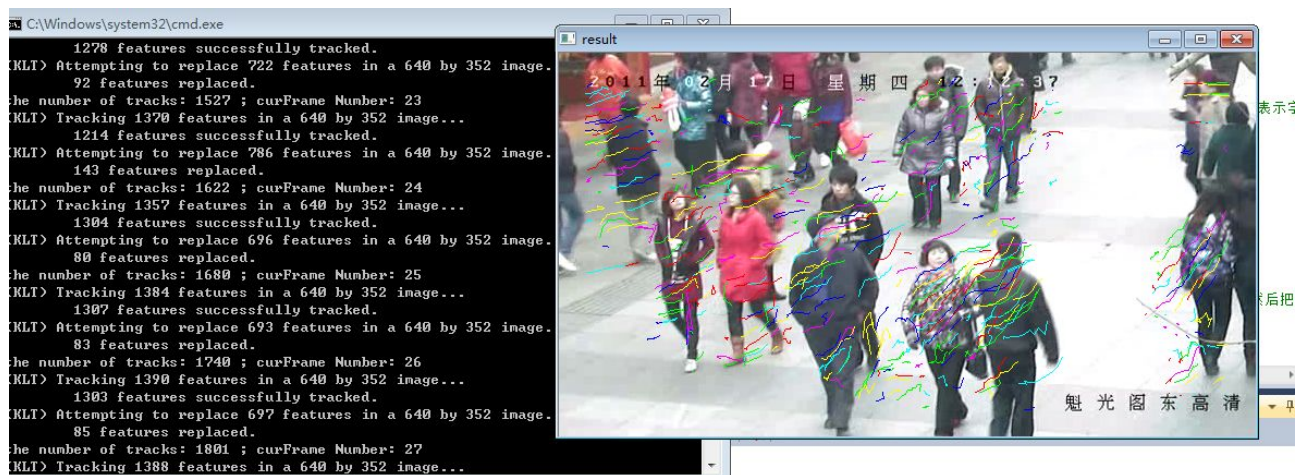


图 3.1 实验结果

图 3.1 实验结果中左侧为程序运行时的提示信息，提示读取到多少帧，这一帧里有多少特征点。右侧为实际运动视频。视频中不同颜色的线代表了特征点的运动轨迹。从实验结果中可以看出，这些轨迹基本都在行人上。得到的特征点保存在一个 txt 文本中，保存的有效内容是特征点坐标以及帧数信息。其中文本的每一行第一个数据表示某一个特征点总共出现的帧数。

### 3.4 本章小结

本章节主要介绍了图像特征的概念，包括颜色特征、纹理特征、空间关系特征、形状特征。之后重点介绍了 KLT 算法。我们该算法的求解步骤分解成两个子问题：1、如何选择特征；2、如何在帧与帧之间跟踪图像。在第二个问题中，我们计算待跟踪窗口  $W$  在视频帧间的灰度差平方和(Sum of Squared intensity Differences)作为度量的标准。由于帧与帧之间的微小运动，当前帧可以近似看成上一帧的一种细小的转变。实际操作中，这些近似会带来误差，但是我可以多次迭代的方法来减小误差。至于如何选择合适的窗口，我们的定义是如果能够很好的被跟踪，则认为该窗口是一个好窗口。本章的研究为第五章通过聚集度反应人流密度奠定了基础。

## 第四章 人流密度检测算法的研究

人流密度检测通常用在商场、车站、大型娱乐设施等公共场所。所谓人流密度是指场景区域内行人的平均数量。对于商家或者活动管理者来说，这个数值不能太大，也不能太小。太大，人员不易疏散，容易导致事故的发生。太小，则达不到预期的效果。可以将人流密度分成五个等级，组织者或者管理者可用根据实时的等级及时作出调整。

### 4.1 人流密度等级分类与系统流程

#### 4.1.1 人流密度等级分类

现实生活中，人流密度等级的定义根据应用场景的不同可以多样化。本文参考采用了Polus 提出的定义<sup>[14]</sup>，即把人流密度从低到高分成 5 个等级，如表 4.1 所示：

表 4.1 人流密度等级分类

密度等级	极低密度	低密度	中等密度	高密度	极高密度
场景人数	0—10	11—30	31—60	61—100	>100



(a) 极低密度



(b) 低密度



(c) 中等密度



(d) 高密度





(d) 极高密度

图 4.1 人流密度等级分类实例图

图 4.1 通过对真实场景的获取，实例展示了这五种人流密度对应的不同场景。

4.1.2 人流密度监控系统工作流程

人流密度监控系统是借助计算机视觉技术对包含人群场景的图像信号实时分析、实时处理的系统。人流密度检测主要包括运动检测跟踪和密度估计两块。一般的处理步骤如下：首先提取出人流的运动行为与运动特征作为运动前景。利用基于像素统计、纹理分析或者个体特征的方法对运动前景进行人流分类。再将分类结果传输给控制系统由控制系统做处理。流程如图所示：



图 4.2 人流密度检测系统工作流程

- （1）视频采集：摄像头连续摄取目标区域场景内的人流图像并反馈给视频处理模块。
- （2）视频处理：将采集得到的视频图像转变为视频序列，采用视频图像处理算法对这些序列分析，得到人流密度信息。
- （3）控制单元：根据视频处理模块处理得出的密度信息并结合相应的场景作出相应判断，自动或者由工作人员采取措施。

在整个框图中最重要的是视频处理模块，而该模块中最核心的又是视频图像处理算法。处理算法的选取直接影响了观测结果的准确性。

4.2 前景提取

前景提取是人流密度检测研究中常用的手段。目的是从图像序列的背景中提取出我们感兴趣的运动对象，忽略对我们帮助不大的其他信息。对视频中的图像可以使用运动检测技

术得到运动对象或者目标。这对于跟踪运动目标和分析图像异常事件等后期处理都具有十分重要的帮助。通常，视频监控中常用的前景提取方法有光流法、帧间差分法、概率法、前景建模法等。

### 4.2.1 光流法

光流法是由于观察者和场景的相对运动导致的视觉场景中对象、边界、表面呈现明显运动特征的一种模式。光流概念的提出是上世纪 40 年代美国心理学家 James J. Gibson 首次提出的，通过会移动的动物来描述视觉刺激。James J. Gibson 强调了光流法在启示感知中的重要性，具有辨识环境行为的能力。光流分为稀疏光流和稠密光流<sup>[20]</sup>。在计算稀疏光流前，需要在被跟踪前选定一组点。如果指定的这组点特性比较明显，那么跟踪效果也比较稳定可靠。借助 OpenCV 我们能够找到最适合跟踪的特征点。稠密光流是将每一个像素都与速度相关联，OpenCV 中实现了一种块匹配方法，即将相邻两帧同一个像素点的邻域进行匹配。相比于稀疏光流，稠密光流的开销比较大，所以更多只用在理论研究中。光流的计算方法按照数学角度来划分可以分为基于匹配的方法、基于梯度的方法、基于相位的方法和基于能量的方法。

利用光流法进行前景提取的步骤包括：①利用相邻的前后两帧图像估计运动场并得到运动向量，②使用运动向量进行前景提取，③对上一步得到的分割结果后续处理，降低噪声，合并细小区域，④最后统计区域个数。如果同样的区域超过一个，则表示场景已经发生了变化。光流法的优点是运动目标的检测不受摄像机运动状态的影响，缺点是计算复杂度高，实时性差，易受噪声干扰。

### 4.2.2 帧间差分法

帧间差分法是对前后两帧图像做差分运算来获得运动目标的图像轮廓。差分运算得到的帧差图像中，运动图像与背景图像之间存在比较大的差异。对两帧图像做差分运算可以得到这两帧图像亮度差的绝对值。最后通过判断该值与设定好的临界阈值的大小判断当前场景是否存在运动对象（大于阈值，则存在运动对象，否则不存在）。

相邻两帧的帧间差分函数为：

$$D_k = |I_k - I_{k-1}| \quad (4.1)$$

其中  $I_k$  代表第  $k$  帧图像的灰度值大小。

而通过阈值(T)判断可以得到前景目标函数，可由以下公式给出：



$$F_k = \begin{cases} I_k(i, j), D_k \geq T \\ 0, D_k \leq T \end{cases} \quad (4.2)$$

帧间差分法的优点是算法的复杂度低，实现简单；对于光线经常变化或者存在多个运动对象的场景都能很好的检测出前景，检测能力强，稳定性高。但是，该法也存在着如下缺点：只能提取出目标对象的边缘轮廓，不能提取全部轮廓；

容易受到两帧之间时间间隔的影响。如果是快速移动的目标，若选取的时间间隔不当，系统会认为相邻两帧间没有发生重叠，从而检测出示两个不同的对象，因此需要选择较小的时间间隔；而对于速度较慢的目标，应该选择较长的时间间隔，否则会导致目标对象在相邻两帧间完全重叠，从而检测不出对象。因此帧间差分法在不同的场合下选择合适的时间间隔十分重要。

### 4.2.3 概率方法

概率方法检测运动对象的基本思想是：静止状态下的运动对象会在一定帧数内持续固定的颜色信息，所以在特定的像素点出现该像素值的概率也越大。通过这个方法可以判断视频图像中的某一点是属于前景还是背景。概率方法可以分为两类：单高斯背景模型和高斯混合背景模型。

(1) 单高斯背景模型：单高斯背景模型最为简单，适用于背景单一不变的场合，采取参数迭代方法，不需要每次都建模处理。其认为对于某个背景图像  $B$ ，特定点像素点亮度的分布符合高斯分布的特性，即坐标点  $(x, y)$  处的亮度满足：

$$IB(x, y) \sim N(u, d) \quad (4.3)$$

其中  $u$  和  $d$  代表背景模型中每个像素的平均值和方差属性。对于一幅给定的图像  $G$ ，若满足：

$$\exp\left(-\frac{(IG(x, y) - u(x, y))^2}{2d^2}\right) > T \quad (4.4)$$

则认为点  $(x, y)$  是图像的背景点，否则是前景点。其中  $T$  是阈值。又因为图像背景随着时间的推移慢慢地变化，我们需要不断更新每个

像素点的参数值：

$$u(t+1, x, y) = a * u(t, x, y) + (1-a)I(x, y) \quad (4.5)$$

其中， $a$  为反应背景变化的更新参数。

(2) 混合高斯背景模型：混合高斯模式实际上是单高斯模型的扩充。使用  $K$  ( $K$  值一般取 3~5 之间整数) 个单高斯模型<sup>[21]</sup>表示图像中像素点的特征。获得新的图像帧后更新混合高斯噪声模型，并将混合高斯模型与当前图像中的每个像素点匹配。如果匹配成功，则认为该点是背景点，否则属于前景点。

混合高斯模型的原理可以表述为：为每个像素建立  $K$  个单高斯模型。如果某一个像素点的采样值为  $\{X_1, X_2, \dots, X_t\}$  ( $X_t$  表示当前像素值)，则观察到  $X_t$  的概率为

$$p(X_t) = \sum_{i=1}^k w_i \frac{1}{2\pi^{n/2} |p_i|^{1/2}} e^{-1/2(X_t - u_i) p_i^{-1} (X_t - u_i)} \quad (4.6)$$

中  $w_i$ 、 $u_i$  和  $p_i$  分别表示第  $i$  个模型的权值、均值以及颜色值的协方差矩阵， $n$  为颜色通道数。该模型其

中假设了 RGB 三种颜色通道相互独立并且方差相同

$$p_i = \sigma_k^2 I \quad (4.7)$$

### 4.3 人流密度检测的常用方法

在绪论中介绍过，目前人流密度检测的研究方法主要分成两类，一是基于像素统计的方法，另一个是基于纹理分析的方法。

#### 4.3.1 基于像素统计的人流密度检测方法

基于像素统计的人流密度检测的方法是由 Davies 等人首先提出的，该方法的依据是图像上人数所占的像素比例是与人数在整个人群中的比例成正比的<sup>[4]</sup>。首先，使用边缘检测算子与背景差分法，对运动前景进行边缘检测，统计图像的边缘像素数量以及前景像素数数量。然后，通过采用最小二乘拟合法或者卡尔曼滤波器，求得其线性模型的参数。选择合适的顶值  $T_1$ ，可将人群图像粗略的分成高密度人群和低密度人群两大类。由表 4.1 可知，低密度人群可划分成极低密度和低密度，高密度人群又可细分成中等密度、高密度、极高密度。其中极低密度和低密度的划分根据底值  $T_2$  来判断。该方法虽然实现简单，但是底值只用于分类低密度，因此仅能在低密度人群的场景时获得较高的准确度。

在 1999 年，Chow 提出了基于混合全局学习算法的神经网络来估计人流密度检测的方法。该方法的提出的依据是人流密度与图像中运动区域有关<sup>[5,6]</sup>。首先通过背景差分法和边缘检测，

获得背景、前景以及前景边缘。接着将背景和前景像素各自的占比以及前景边缘像素数作为参数，输入到已经训练好的神经网络，估计人群的密度。2002 年，Chow 等人进一步改进了该算法，在其中添加了行人模板的匹配技术。首先通过行人模板对人流对象区域进行搜索匹配，然后将匹配得到的结果作为第四个特征值，再组成一个新的四维特征矢量，然后把该矢量送入神经网络估计人流密度。

美国的 Paragios 等人于 2001 年提出了一种基于马尔科夫随机场(MRF)的实时动态的人流密度估计的方法。首先运用了检测基于马尔科夫随机场变化的方法来区分背景和前景。然后对所得的结果图像结合几何模块进行透视校正，从而得到监控场景内的人流密度<sup>[7]</sup>。

上述的方法都是建立在像素统计基础上的，例如对视频中人群边缘的像素统计，人群所占面积的像素统计等。由于人群的人数与像素数有着线性关系，我们可以用这一性质来估计人流密度。由于上述几种方法都要首先进行背景差分操作，因此需要提供一个没有行人走动的干净的背景。他们的优点是实现简单，提取的特征计算量小，在人流不是很拥挤的情况下，能得到比较精确的结果。但是在高密度情况下，人与人之间比较拥挤，人群相互重叠，从而导致较大的误差。

#### 4.3.2 基于纹理分析的人流密度检测方法

图像的纹理虽然从局部的角度来看没有规律可循，但整体上来说却呈现一定的规则。它不仅反映出图像的灰度信息，还能够帮助研究人员获得图像的空间分布以及具体结构等信息。1998 年，巴西的 A.N.Marana 研究了图像的纹理特征，并提出了一种基于纹理分析技术的人流密度估计方法<sup>[8]</sup>，它依据的是不同密度的人流在图像上对应着不同的纹理模式的原理。此类方法的理论依据是高密度人群具有较细的纹理模式，而低密度人群具有较粗的纹理模式。把人流图像当作纹理来处理。通过纹理分析提取出人流图像的纹理特征。该方法克服了基于像素统计方法在高密度人流情况下因人群遮挡导致的误差大的缺点。该算法流程如图 4.1 所示：



图 4.3 基于纹理分析的人流密度检测算法流程图

纹理分析，顾名思义，就是运用图像分析技术提取出图像的纹理，并分析走向、密集程度以及粗细等信息的过程。为的就是找出隐藏在其中的内容，发现规律，以达到研究的目的。近年来，随着研究的不断深入，研究者们提出了一系列不同维度，不同方法的纹理分析探究手段。较为常见的有统计分析法，结构分析法，频谱分析法和模型分析法。统计纹理分析是

一种常用的纹理分析方法，也是应用时间最早最长的研究方法。例如 Laws 纹理能量法、自相关法等在众多领域有着广泛的应用。结构分析法主要分析纹理基元的单位的密度、面积、周长等几何特征，可以把纹理分析转变成数学运算，直观易懂，但在自然纹理分析时受限。频谱分析法，利用傅里叶变换、Garbor 变换、小波变换以及各种滤波器进行滤波变换。不过频谱分析法运算量较大，而且稳定性不高。模型分析法（例如马尔科夫随机场模型），创建模型的过程十分复杂，这也导致了繁琐的计算，对我们分析图像的纹理特征帮助甚微。

统计分析方法，虽然简单，但无法解决人群重叠的问题。为了解决这一问题，可以采用灰度共生矩阵来分析图像的纹理特征。灰度共生矩阵是对纹理的一种度量方法。每一副图像都存在与之对应的灰度共生矩阵，它描绘的是图像灰度在空间上的联系。灰度是图像的固有特征，而纹理则看成是灰度在图像空间上互相联系、互相作用形成的具有一定结构的隐形特征。而分析这种隐形特征前提是需要确定一定范围内灰度之间的关系，这也是纹理分析的关键步骤。而灰度共生矩阵正是为了寻找这种关系而被提出的。

设  $f(x, y)$  表示一副二维数字图像，其大小为  $M \times N$ ，灰度级为  $L$ ，两个像素点间距为  $d$ ，像素点直接连线与水平轴方向夹角为  $\theta$ ，则我们可以定义图像的灰度共生矩阵为：

$$p(i, j, d, \theta) = \#\{(x_1, y_1), (x_2, y_2) \in M \times N \mid f(x_1, y_1) = i, f(x_2, y_2) = j\} \quad (4.8)$$

其中  $\#(x)$  表示集合  $x$  中元素的个数。运算时通常  $d$  值唯一，夹角一般取  $0^\circ, 45^\circ, 90^\circ, 135^\circ$ ，即水平、垂直、对角线三个方向。

当计算出灰度共生矩阵后，就可以提取用来描述图像纹理特征的统计量了。常用的能量、熵、对比度、均匀度以及相关性五种。

#### (1) 能量

$$E = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} f(i, j, d, \theta)^2 \quad (4.9)$$

能量也称为一致性，是灰度共生矩阵中所有元素值的平方和。当灰度共生矩阵的一些值大而其他值小时，能量值较大；反之，如果所有值均相等，则能量值小。对于粗纹理的图像区域，其灰度共生矩阵中的元素值主要集中在主对角线附件。这是因为粗纹理中许多像素对具有相同的灰度。细纹理的区域，其灰度共生矩阵中的元素值则较为分散。总之，从图像整体看，纹理粗对应的能量值较大；反之，能量值较小。

#### (2) 熵

$$H = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} f(i, j, d, \theta) \log f(i, j, d, \theta) \quad (4.10)$$

熵是图像中所有元素随机性的度量。当灰度共生矩阵中所有元素值近似相等、元素具有较大的随机性时，灰度共生矩阵中元素的分部较为离散时，熵值较大。反之，值较小。

### (3) 对比度

$$I = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i-j)^2 f(i, j, d, \theta) \quad (4.11)$$

对比度是指一个像素在整个图像上与其邻居间的灰度对比的度量。可以体现纹理的清晰程度和纹理纹沟的深浅程度。当灰度共生矩阵的值比较集中于对角线时，对比度比较小，纹理变化较小，视觉效果较为模糊；反之，对比度较大，纹理变化也比较大，视觉效果较为清晰。

### (4) 均匀度

$$S = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{f(i, j, d, \theta)}{1 + (i-j)^2} \quad (4.12)$$

均匀度也称为同质性，是灰度共生矩阵中元素对角线分布的空间紧密程度的度量。粗纹理的灰度共生矩阵的数值较集中于主对角线附件，相应的 S 值也较大，说明图像纹理的不同区域之间缺少变化。相反，细纹理的均匀度值较小。

### (5) 相关性

$$C = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{(i-u_x)(j-u_y)f(i, j, d, \theta)}{\sigma_x \sigma_y} \quad (4.13)$$

其中  $u_x = \sum_{i=0}^{L-1} i \sum_{j=0}^{L-1} f(i, j, d, \theta)$ ,  $u_y = \sum_{j=0}^{L-1} j \sum_{i=0}^{L-1} f(i, j, d, \theta)$ ,  $\sigma_x = \sum_{i=0}^{L-1} (i-u_x)^2 \sum_{j=0}^{L-1} f(i, j, d, \theta)$ ,

$$\sigma_y = \sum_{j=0}^{L-1} (j-u_y)^2 \sum_{i=0}^{L-1} f(i, j, d, \theta)。$$

相关性指一个像素在整个图像上与其邻居间相关程度的度量<sup>[41]</sup>，相关性值的大小反应了图像中局部灰度的相关性。如果灰度共生矩阵中所有元素的值均相等，则相关性的值就大，说明区域图像灰度分布比较均匀；如果灰度共生矩阵中元素值相差很大，则相关性的值就小。在以上五种特征值中，能量、熵、对比度和均匀度因为计算方便并且又能给出较高的分类精度，因此在实际应用中使用较多。

灰度共生矩阵可以用来估计高密度人流的情形。但是在低密度情况下，它的有效性却比较低。随着信号处理技术在计算机视觉和图像处理中的不断发展，2001 年 Marana<sup>[9,10]</sup>等人提出了利用小波分析检测人流密度的方法。因为摄像机摆放的位置不同，人流图像也会呈现出不同的透视效果，所以具有不同的尺度，针对此问题，Marana 等人在基于人流图像纹理分析

多年理论经验积累基础上，提出了一种多分辨率分析的人流密度检测算法。该算法 利用基于小波包分解的方法提取图像的纹理特征。首先，进行图像的二级小波包分解，得到图像的小波包分解系数矩阵。然后计算系数矩阵的能量并把这些能量其作为特征矢量送入自组织神经网络(SOM)进行分类，实现人流密度检测。但是此算法数据量庞大，运算繁琐，因此在实际应用中，一般不被使用。

近年以来，人们对于此类方法的研究一直在深入进行中。例如，H.Rahmalan 提出了基于Chebyshen 矩描述纹理特征的方法<sup>[11]</sup>。

下表给出了基于像素统计和基于纹理分析算法的对比，从中可以看出，基于像素统计的方法使用于低密度人流的场景，而基于纹理分析的方法在高密度人流时能取得更好的效果。

表 4.2 人流密度检测基于像素统计与基于纹理分析方法比较

	优点	缺点
基于像素统计	计算简单，复杂度低，适合低密度人流	高密度场景准确率低
基于纹理分析	能准确的检测高密度人流场景	计算复杂，不适合低密度

除了基于像素统计、基于纹理分析的算法外，还有一种基于个体特征的人流密度检测算法。该类算法着重于从人体的个体特征着手来进行人流密度的研究。香港的 Lin Shengfun，提出了一种利用小波变换算法提取个体头部特征的方法来定位个体，以此确定拥挤场景下的人群密度即是个体人数<sup>[12]</sup>。但此算法计算复杂，通用性低，而且无法解决严重遮挡时对分割目标所带来的难题，因此应用很少。

此外，一些新的分析方法，分析工具随着研究的不断深入、技术的日趋进步而涌现出来。比如，随着支持向量机理论的日臻完善，SVM 已经广泛的应用到了图像处理分析领域。支持向量机方法是建立在统计学习理论的 VC 维（Vapnik Chervonenkis Dimension）理论<sup>[42]</sup>和结构风险最小原理的基础上，根据有限的样本信息在模型的复杂性（即对特定训练样本的学习精度）和学习能力（即无错误地识别任意样本的能力）之间寻求最佳折衷，以期获得最好的推广能力。SVM 有着严谨的数学基础和科学的理论依据，虽然算法主要适用于小样本，但依旧具有全局性最优性。不仅如此，相比于神经网络，它的泛化能力也更加突出。出于以上的优点，并结合上述的三种研究方法，SVM 也开始逐渐地被研究人员应用到人流密度分析检测中去。

2007 年,Damian 利用基于像素统计的方法并结合支持向量机实现了人流密度检测。首先，提取网格化的像素特征用以实现对射影畸形的矫正<sup>[13]</sup>。如图 4.2 所示，其中(a)图是需要网格化的原图，(b)图是采用网格化处理后得到的图像结果。(b)图中图像被分成一个个小的网格，

上层的网格所占面积小，下层的网格逐步扩大。其次，由于每一个网格内的像素计算时都有不同的权值。通过计算提取出畸形矫正后像素的统计特征，然后采用支持向量机分析。其得到的人流密度结果在人群不是很密集的情况下准确率接近 90%。



图 4.4 网格化的像素统计方法

## 4.4 实验结果对比

### 4.4.1 基于像素统计的实验结果

对于基于像素统计的人群密度分类，选择视频作为实验对象，利用 OPENCV 对视频进行处理，大致将其分为低密度，极低密度和纹理分析人群三大类。基于像素统计的人群密度分类方法进行人群密度分类算法简单，但是只是适用于低密度人群。特别注意的是这里的阈值选取采用的是分析多幅从视频中提取的低密度，极低密度，以及纹理分析人群中相应图片获取的，这样就减少了人为主观性，更具说服力。



图 4.5 实验原图

先对视频进行前景提取，利用背景差分提取前景。再将前景提取的结果使用 Canny 算子边缘检测，如图 4.6 和 4.7 所示。



图 4.6 前景提取结果



图 4.7 边缘检测效果图



(a)极低密度

(b)中密度

图 4.8 基于像素统计的结果图



4.4.2 基于纹理分析的实验结果

高密度人群会出现人群中人与人的重叠，如果利用基于像素统计的人群密度分类方法虽然算法简单，但是所得出的结果误差比较大，因此利用基于纹理分析的人群密度分类方法，这种方法算法复杂，计算量比较大。一般的先提取出基于灰度共生矩阵，并在此基础上得到特征值，把特征值利用支持向量机的分类方法进行分类。

对于纹理分析的视频，截取不同密度种类的图片进行实验分析，构成图片训练库。图片的灰度在 0-255 之间均匀取值，根据灰度共生矩阵的含义和生成过程，如果图片的灰度级越大，那么灰度共生矩阵的维数就越大，对灰度共生矩阵做相应处理时，计算量就越大，计算速度就越慢。因此为了方便计算需要多图像进行压缩，这样既保持图像的原型，又不影响图像的纹理特征。下面是对相同密度级的图像进行不同级别灰度级压缩后的各特征比较图，这里取  $d$  的范围为 0~40， $\theta=0^\circ$ 。

(1)相同密度级不同灰度级下能量变化，如图 4-9 所示。

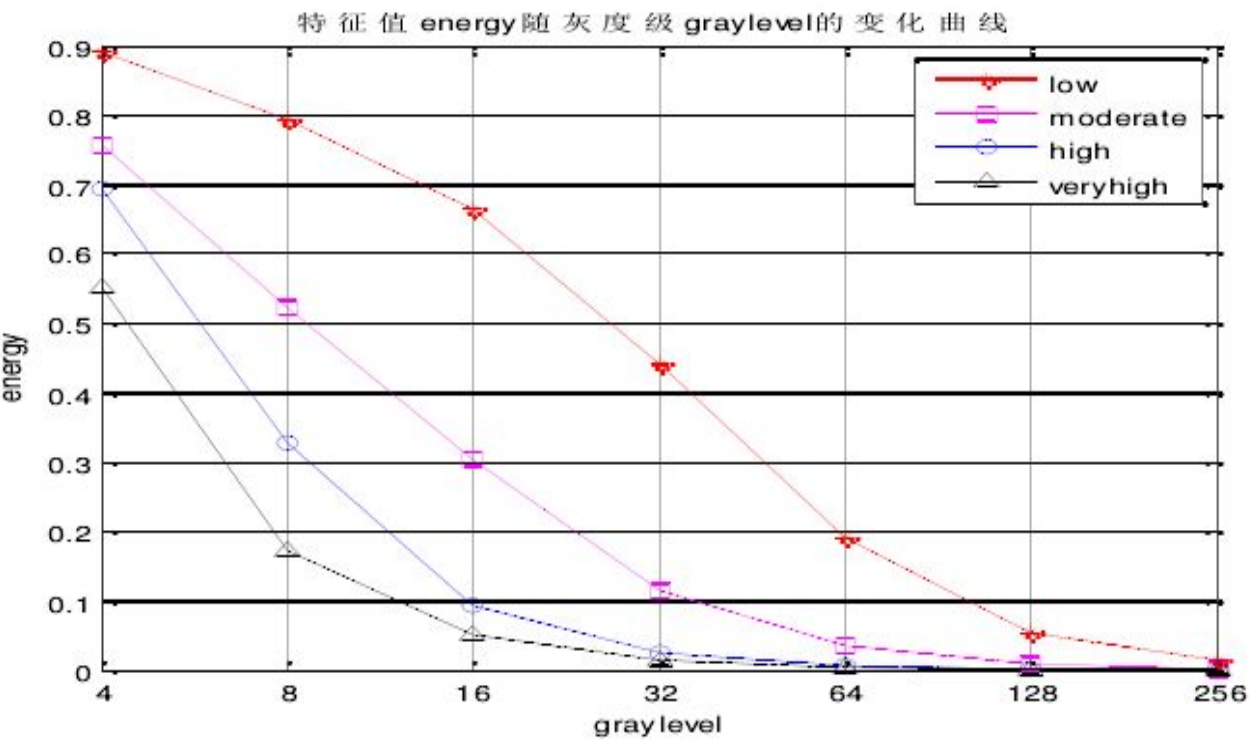


图 4.9 相同密度级不同灰度级下能量变化示意图

(2) 相同密度级不同灰度级下熵变化，如图 4-10 所示。

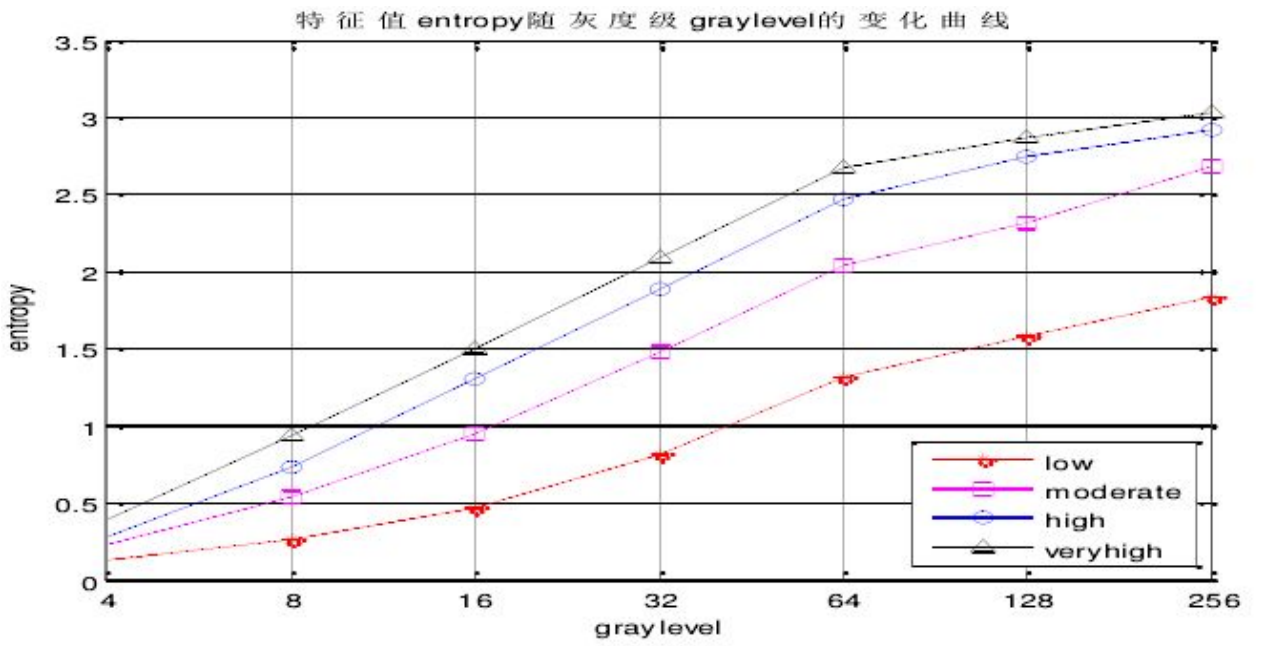


图 4.10 相同密度级不同灰度级下熵变化

(3) 相同密度级不同灰度级下对比度变化，如图 4-11 所示：

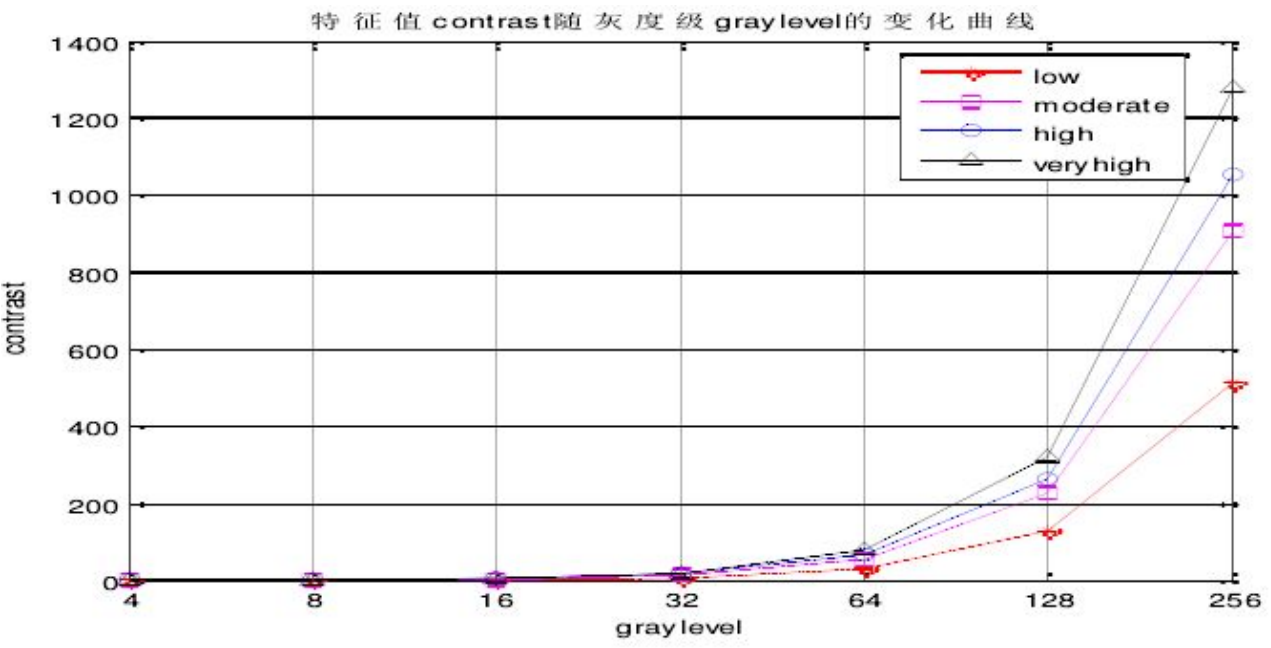


图 4.11 相同密度级不同灰度级下对比度变化

由上图可知，将图像压缩到 8 灰度级，16 灰度级，32 灰度级，64 级灰度级，128 级灰度级时，虽然图片有变化，相同人群密度的五种特征在不同灰度级下相互的对比关系并没有变化，所以在对人群进行分类时可以压缩图像到一定的灰度级，另外本文选择 16 灰度级对图像进行处理。在选定灰度级后，将研究不同密度级别的人群进行各特征值变化比。

## 4.5 本章小结

本章先是探讨了人流密度的五个等级分类以及人流密度检测系统的工作流程，然后分析了开展人流密度研究之前重要的前景提取的知识。紧接着分别研究了基于像素统计的人流密度检测算法、基于纹理分析的人流密度检测算法，并对它们做了简单的对比。其中在基于纹理分析的算法中，还进一步讨论了灰度共生矩阵的概念。最后通过实验对比了基于纹理分析的算法和利用灰度共生矩阵的纹理分析方法。

## 第五章 SDP 模型的研究与应用

聚集运动是自然界中十分普通但又十分特别的一种运动，小到微生物，大到人类，聚集运动无处不在。研究中发现，聚集运动都会由一开始的杂乱无章逐渐到转变到最后呈现出贵规则运动。而在聚集运动的研究中，SDP(Self-Driven Particles)<sup>[26]</sup>模型由于表现出了和自然界中的群体聚集运动高度的相似性，所以它被广泛的用来研究人流运动。利用 SDP 研究中的理论，可以帮助我们更好的研究人流运动。

### 5.1 SDP 模型简介

SDP 模型是为了研究系统中粒子的聚集运动而提出的。在粒子的状态变化过程中，它们呈现出一种十分复杂的彼此之间相互影响的运动方式。这也是粒子运动中一个十分有趣的值得研究方面。而原先研究这些粒子运动的理论都是平衡理论。这些内部的复杂运动导致了平衡理论难以深入的研究下去，于是在 1995 年 T. Vicsek、A. Czirók、E. Ben-Jacob、I. Cohen, 和 O. Shochet 等人提出了 SDP 模型<sup>[31]</sup>。

SDP 模型中的粒子运动速度由一个简单的规则决定并伴有随机波动：半径为  $r$  的领域范围的粒子以一个恒定的速度运动，粒子群体的运动方向被更新为上一时刻的平均方向。邻域内排列方向上的随机波动  $\eta$  会导致整个系统由杂乱无章的运动逐渐的向有规则的运动转变。

假设在 SDP 模型里面，有个体  $i$ ，其运动方向为  $\theta$ ，每一个时刻更新完速度方向后，方向可以表达为：

$$\theta(t+1) = \langle \theta_j(t) \rangle_{j \in N(i)} + \Delta\theta \quad (5.1)$$

$\langle \theta_j(t) \rangle_{j \in N(i)}$  代表个体  $i$  所属邻域上一刻的平均方向， $\Delta\theta \in [-\eta/2, \eta/2]$ ，它是一个随机的角度。

运动角度表示为：

$$\arctg[\langle \sin(\theta(t)) \rangle_r / \langle \cos(\theta(t)) \rangle_r] \quad (5.2)$$

在 SDP 模型中，我们提出一个十分重要的概念——聚集度（collectiveness），这个值反应了个体参与群体运动的程度。它是衡量不同运动群体的最基本也是最普遍的值。

## 5.2 聚集度的测量

在群体运动中，局部邻域范围内的个体行为呈现较高的一致性，当出了邻域范围，行为的一致性不再明显。这是因为群体运动中的地理感知能力只在很小的范围内起作用，并且通常它们的运动是建立在诸如位置或者它“邻居”的运动之类的局部信息上。一些实验研究表明了地理拓扑结构关系等信息传递在群体运动中的重要性。但是很多现有的文献资料仅仅计算了个体的平均速度，来表征整个群体的聚集度，这既不准确也不可靠。

传统的计算机视觉领域，很多方法更多的注重于运动模式分割或者分析行人的行为。例如 Rabaud<sup>[27]</sup>Zhou<sup>[28]</sup>等人计数和聚类的独立/群体运动对象检测。Zhou 等人使用了混合动力系统（一个应用于人群仿真的动力学应用）分析行人行为<sup>[29]</sup>。遗憾的是，他们没有提及到如何度量群体行为的聚集度，也没有探索相应的未来可能的应用。

聚集度描述了个体参与群体运动的程度。它和很多因素相关，密度不同或者主导的个体不同都会导致聚集度结果的不同。通过计算聚集度，能够帮助我们了解不同种类的群体之间的共性。另外聚集度在监控群体从无序到有序的运动变化、研究聚集度与人流密度之间的关系、描绘群体运动动态变化的过程等领域存在广阔的应用前景。

一个群体不单单是众多个体的集合。在某些条件下，不同运动状态的群体汇聚到一起形成一个新的群体。因此整个群体的聚集度是由组成它的个体的聚集度决定的。同时整体的聚集度反应了在同一个群体中个体之间行为的相似性。

通过量化群体的地理位置特性，我们提出了一种既能够描述整个群体也能够描述单个个体聚集度的算法——群体汇聚算法（Collective Merging）。采用从上到下的描述方法：即先从邻域内部个体的行为相似性到所有成对的个体间，再从个体聚集度到群体聚集度层层推进。

### 5.2.1 邻域范围内的行为相似性

我们首先度量邻域范围内个体行为的相似性。假设个体  $i$  和个体  $j$ ，其中个体  $j$  在  $i$  的邻域内，设时间为  $t$ ，则邻域的数学表达式为  $j \in N(i)$ ，此刻的相似性定义为：

$$w_t(i, j) = \max(C_t(i, j), 0) \quad (5.3)$$

$C_t(i, j)$  是  $t$  时刻个体  $i$  和个体  $j$  之间的速度关联矩阵。 $N$  是由 K-NN(K-nearest-neighbour) 分类算法决定的，这是因为大量的已有实验数据表明不仅是人类，其他动物也都有着局部对

自己周围一定数量的邻居的地理感知能力<sup>[30]</sup>。 $w_t(i, j) \in [0, 1]$ 衡量了个体在邻域中相比于整个群体的行为一致性。当两个个体不属于同一个邻域时，这个值不再准确。下图可以说明这一点。

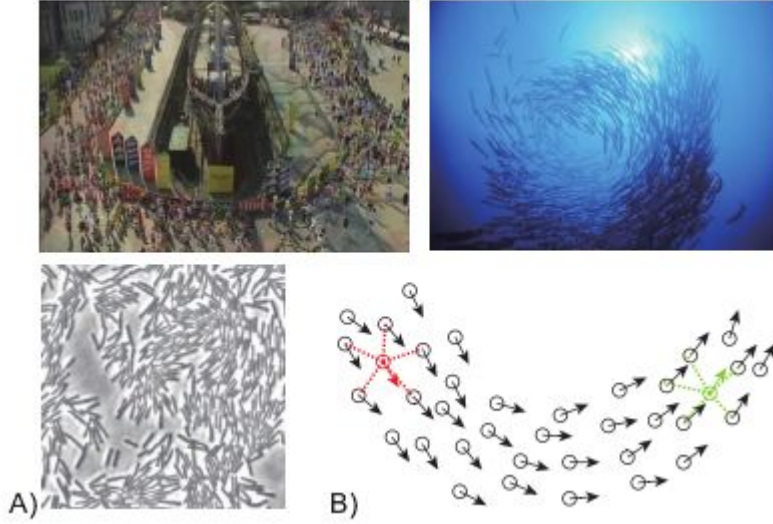


图 5.1 聚集运动示例

图 5.1 中 A 图表示的是人群、鱼群、细菌的聚集运动。B 图说明的是空间地理关系对运动的影响。即使是同一个群体中，不同个体当超过一定距离时候，它们的速度关联不再紧密。正如图中红色个体和绿色个体。

### 5.2.2 群体范围的行为相似性

由于当两个个体相隔一段距离时，相似性不能很好的计算。我们提出了基于地理拓扑结构的个体两两之间成对的相似性——路径（path）<sup>[32]</sup>。在群体系统中，路径在个体之间的信息传递与交流中扮演着十分重要的角色。因此基于路径的相似性能够更好的刻画群体中个体的行为一致性。

令  $W$  为群体  $C$  的加权邻接矩阵，邻接矩阵的边缘值  $w_t(i, j)$  是个体  $i$  和个体  $j$  的行为相似性，再令  $\gamma_l = \{p_0 \rightarrow p_0 \rightarrow \dots p_l\}$  表示在邻接矩阵中从个体  $i$  到个体  $j$  经过节点  $p_0, p_1 \dots p_l$  ( $p_0 = i, p_l = j$ ) 长度为  $L$  的路径。则路径  $\gamma_l$  上的路径相似性为：

$$v_{\gamma_l} = \prod_{k=0}^l w_t(p_k, p_{k+1}) \quad (5.4)$$

由于个体  $i, j$  之间长度为  $L$  的路径不止一条，假定  $P_l$  包含了  $i$  和  $j$  之间所有长度为  $L$

的路径，则 L 长路径上的路径相似性定义为：

$$v_l(i, j) = \sum_{\gamma_l \in p_l} v_{\gamma_l}(i, j) \quad (5.5)$$

其中  $v_l(i, j)$  是邻接矩阵  $W^l$  在  $(i, j)$  处的输入。

### 5.2.3 个体聚集度与群体聚集度

接下来由路径相似性导出个体聚集度。由于 L 长路径上的相似性  $v_l(i, j)$  表示的是 i 和 j 这两个个体在所有长度为 L 的路径上的相似性，那我们可以定义个体 i 在所有长度为 L 的路径上的个体聚集度为：

$$\phi(i) = \sum_{l=1}^{\infty} z^l \phi_l(i) = [Ze]_i \quad (5.6)$$

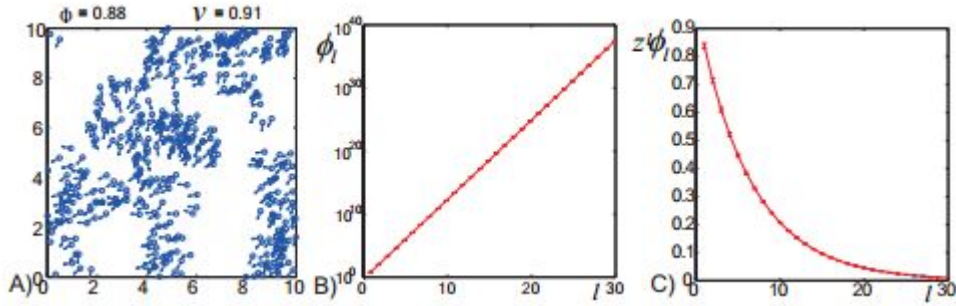


图 5.2 聚集度与速度关系

图 5.2 中 A 图是有 SDP 模型合成的群体，B 图表征的是个体聚集度  $\phi_l$  与路径长度 L 的指数变化关系，图 C 是归一化路径长度下的个体聚集度与路径之间的关系。B 图之所以是指数型增长关系是因为可能的路径个数随着路径长度的增加而指数变化。下面来分析 C 图。

为了研究群体聚集度，我们需要整合所有路径上的个体聚集度，即  $\{\phi_1, \dots, \phi_l, \dots, \phi_\infty\}$ ，但是由于指数增长的关系，不同路径下的个体聚集度无法精确地计算，所以我们定义一个生成函数来整合所有的路径相似性。定义 L 长路径相似性的生成函数为：

$$\overline{w}_{i,j} = \sum_{l=1}^{\infty} z^l v_l(i, j) \quad (5.7)$$

$z$  为实系数的归一化因子， $z^l$  理解成路径相似性的加权系数，因为  $z$  小于 1，这样可以抵消  $\phi_l$  指数增长过快而带来的影响。 $\overline{w}_{i,j}$  是矩阵  $Z$  在  $(i, j)$  处的输入，其中  $Z = (I - zW)^{-1} - I$ ，

$1 < z < 1/\rho(w)$ ， $\rho(w)$  表示邻接矩阵  $W$  的谱半径<sup>[40]</sup>。

由所有路径相似性的生成函数得到的个体聚集度表示为:

$$\varphi(i) = \sum_{l=1}^{\infty} z^l \varphi_l(i) = [Ze]_i \quad (5.8)$$

此时再看图 5.2(C), 令  $z = 0.9 / \rho(W)$ , 计算  $z^l \varphi_l(i)$ , 当长度  $L$  逐渐增加时候,  $z^l \varphi_l(i)$  值逐渐趋于 0。即所有路径长度上个体聚集度归一化后值是收敛的。实系数归一化因子  $z$  控制着收敛的速度, 并且相比于其他因素,  $z$  对相似性的贡献更大。

群体  $C$  的群体聚集度可以看成所有个体聚集度的求和之后平均, 即:

$$\Phi = 1/|C| \sum_{i=1}^{|C|} \varphi(i) = 1/|C| e^T ((I - zW)^{-1} - I)e \quad (5.9)$$

聚集度  $\Phi$  描述了整个群体的结构特性。

### 5.2.4 聚集度的性质

由于群体运动本身具有很多的特性, 描绘其的聚集度也有着三大性质:

(1) 矩阵  $Z$  的收敛性: 矩阵  $Z$  在  $z < 1/K$  收敛。如果我们通过比较  $Z$  的谱半径  $\rho(W)$  来选择系数  $z$ , 则计算起来是十分复杂的(计算  $W$  的特征值得到  $\rho(W)$  的时间复杂度达到  $O(n^3)$ ), 尤其是在复杂的群体时候更是如此。而通过性质(1)我们无需计算谱半径就能得到  $z$  的大小。

(2) 群体聚集度  $\Phi$  的有界性:  $0 \leq \Phi \leq zK / (1 - zK)$ ,  $z < 1/K$ 。假设  $A$  是由 K-NN 算法得到的单位矩阵, 则当  $W=A$  时, 上式取等号。这意味着邻域之间有完美的速度关系, 也就是对于任意的个体  $i$ , 满足  $j \in N(i)$ , 则有  $w_i(i, j) = 1$ 。这其实表明组成这一邻域的所有个体有着相同的运动方向。为了简单起见, 我们的实验中都令  $K=20$ ,  $z=0.025$ , 那么  $\Phi$  的上限值就为 1。

(3) 矩阵  $Z$  输入的上限性:  $\overline{w_{i,j}} < z / (1 - zK)$ 。这一特性可以帮助我们从杂波中检测群体运动。

### 5.2.5 群体聚集算法步骤

至此, 通过上述的论述, 我们可以总结出群体汇聚算法的步骤:

- ①由速度与坐标信息, 利用公式 (4.3) 计算个体  $i$  个个体  $j$  的相似性  $w_i(i, j)$ ;
- ②利用 K-NN 算法计算邻接矩阵  $W$ ;



- ③当  $Z(i,j) \geq K$  时，令其为 1，否则为 0；
- ④设置阈值，剔除矩阵  $Z$  中相似度差的个体。

首先，根据已经给定的坐标  $x$  和速度  $v$ ，计算出  $w_i(i,j)$ ，然后得到邻接矩阵  $W$ 。通过设置矩阵  $Z$  的矩阵，我们很方便的剔除低聚集度的个体并且得到运动群体的集合。根据性质 3，我们令阈值  $k = az / 1 - zK$ ，其中  $0.4 < a < 0.8$ 。建立在对聚集度准确测量的基础上，我们的群体聚集算法能够实时的运行，并且有着潜在而又广阔的前景。下面我们将通过数据验证我们的算法。

### 5.3 基于 SDP 模型的算法验证

借助 matlab 工具，我们通过仿真验证 SDP 模型的群体聚集度、混合群体的聚集度以及收敛条件下的聚集度。在仿真中，我们取 SDP 模型中粒子数为  $N=400$ ，速度大小  $\|v\|=0.03$ ，粒子相互影响的作用范围半径  $r$  取 1。实际上，这三个参数值的选取对 SDP 模型中整体行为的影响并不大。

#### 5.3.1 SDP 模型的群体聚集度

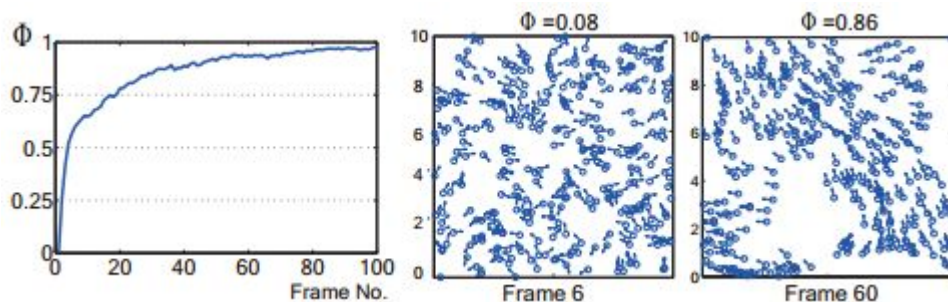


图 5.3  $\Phi$  的变化图

如上图所示，我们计算不同时刻下的群体聚集度  $\Phi$ ， $\Phi$  监控着群体运动的变化过程。开始的时候， $\Phi$  值很低，这是因为所有粒子的空间位置与速度方向是随机分布的，随着时间的推移，群体逐渐向有序的运动集体运动发展。在第 6 帧中，粒子杂乱无章的运动中，而第 60 帧中，粒子的运动以及出现很大的规律性。 $\Phi$  值大小准确的记录了这个变化过程。

而随机波动值  $\eta$  的大小增大的时候，SDP 模型中的粒子运动又变得无序了。而图 5.4 记录的正是随机波动对聚集度的影响。

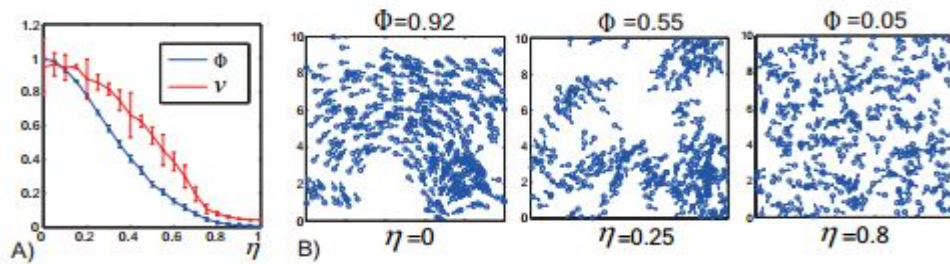


图 5.4 动态变化图

A 图中，红蓝两色的柱状线分别表示的是速度标准差与聚集度的标准差。其中，为了方便，速度  $v$  采用统计概念上的均值，这在已有的研究中很常用：

$$v = \left\| \frac{1}{N} \sum_{i=1}^N \frac{v_i}{\|v_i\|} \right\| \quad (5.10)$$

速度的标准差的大小明显大于聚集度的标准差，说明在 SDP 模型刚开始阶段，速度是不稳定并且敏感的。而聚集度  $\Phi$  证明了其的稳定性。

B 图中， $\eta$  为 0 的时候，所有粒子呈现有规则的集体运动， $\Phi$  值接近它的上限 1；当  $\eta$  增大的时候，粒子的运动开始分散；而当  $\eta$  值更高时，粒子个体随机运动，聚集度也很低。

### 5.3.2 混合群体的聚集

SDP 模型中假设所有的个体都是属于同一群体的。但是对自然界中大多数的复杂群体研究表明它们并不是同一群体的<sup>[31]</sup>。为了证明我们提出的聚集度这一概念在混合群体中的适用性，我们在 SDP 模型中加入部分自有移动的粒子，然后计算混合群体中的个体聚集度。

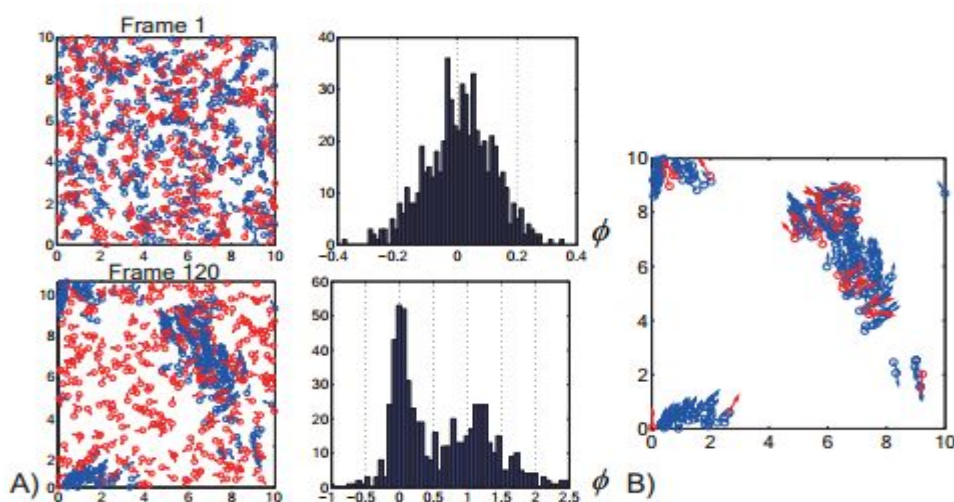


图 5.5 直方图与运动变化图

如上图所示，蓝色的粒子代表 SDP 模型中的粒子，红色的加入的杂波粒子，其中红色杂波粒子个数与蓝色粒子个数相当，并且随机波动  $\eta$  值为 0。刚开始的时候，红蓝两个群体的

粒子随机的运动，对应的个体聚集度的直方图只有一个峰值，当不同颜色的粒子形成各自的群体运动时，直方图中峰谷中间形成明显的分割。剔除个体聚集度小于 0.5 的粒子，我们能够得到群体运动中的个体粒子的集合，如 B 图所示。

### 5.3.3 聚集度的收敛性

前文论述的聚集度三大性质之一便是收敛性。也就是当  $z < 1/k$  时，矩阵  $Z$  收敛。那么如果  $z \geq 1/k$  时，又会怎样呢？令  $z = 1/k$ ，并且计算  $L$  长路径下的归一化聚集度  $[z\phi_1, z^2\phi_2, \dots, z^L\phi_L]$ 。

实验发现，当 SDP 模型逐渐形成聚集运动时候， $z^L\phi_L$  也在逐渐趋向于 1，这也使得  $\sum_{l=1}^{\infty} z^l\phi_l$  趋向于无穷，发散。群体聚集度  $\Phi$  也因此不稳定。如图 5.6 所示。

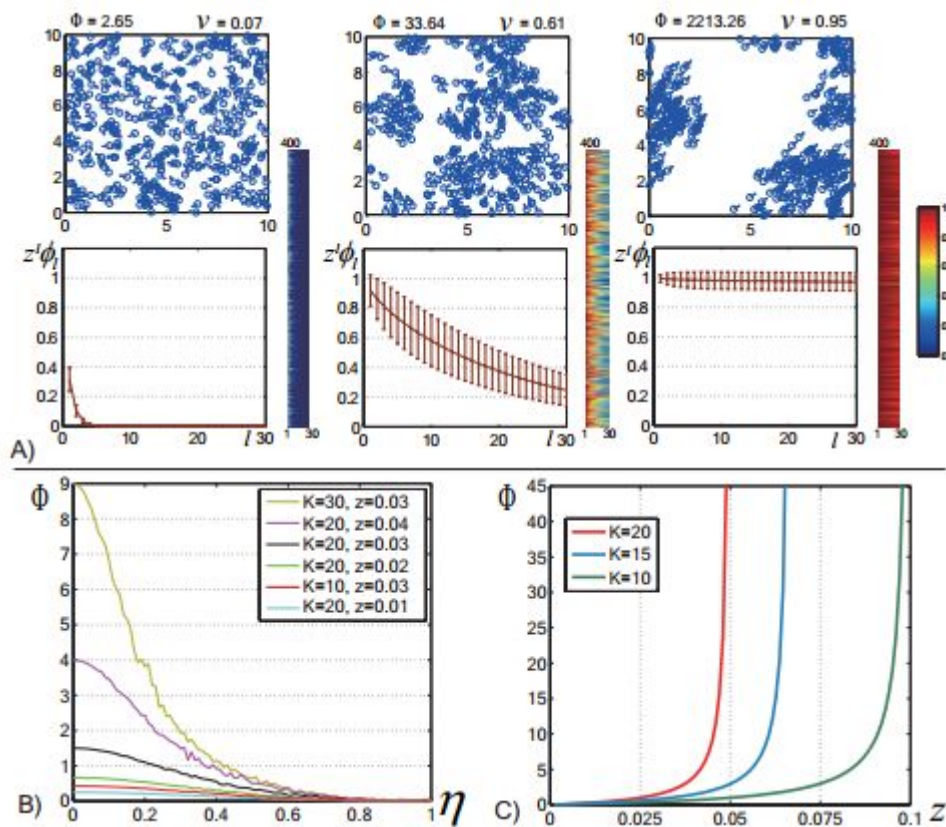


图 5.6 多参数变化图

上图中，图 A 对应的是当  $\eta=0$  并且  $z=1/k$  时三种不同运动水平下的群体运动结果。其中第一行图是运动的直观展示，第二行图中，左边对应的是  $L$  取值 1~30 时的  $z^L\phi_L$  平均值的变化，右边显示的是  $L$  取值 1~30，个体从 1~400 时间  $z^L\phi_L$  所有的值。由于收敛条件不再满足， $\Phi$  值变化的很不稳定。最右边运动的聚集度值甚至超过了 2200。而图 B 反应了不同的  $z$  值与  $K$

值下当随机波动  $\eta$  增大时  $\Phi$  的变化。C 图描绘了在给定的三种不同的  $K$  值条件下, 当  $z$  趋近  $1/k$  时,  $\Phi$  的上限值迅速的变大, 这也导致了其值的不稳定。

实际应用中在计算聚集度时有两个参数  $z$  和  $K$ ,  $K$  决定了邻域的范围大小( $K$ -NN 分类算法),  $z$  决定了矩阵或者说集合  $Z$  是否收敛。当  $K$  取值比较大时, 会使测量结果不再准确, 而  $K$  特别小时, 结果又对噪声敏感, 实际操作中,  $K$  的理想取值是群体  $C$  范围大小的 5%~10%。在我们的仿真中, 取  $K=20$ , 由性质 1 中  $z < 1/K$ , 那么  $z$  的取值也是有范围的, 不同的  $K$  和

$z$  使得  $\Phi$  值的上限也发生变化, 正如图 5.6(B)中那样。上限值的增大, 也使得  $\frac{d\Phi}{d\eta}$  变大, 我们

的测量更容易受群体运动变化的影响。若除以上限值,  $\Phi$  可以限定在 0~1 之间。因此在实际应用中, 我们选择合适的  $K$  和  $z$  值, 我们可以控制聚集度的敏感性。当  $z$  接近  $1/k$  时, 聚集度  $\Phi$  的上界迅速的增大, 这使得  $\Phi$  也变的不稳定, 如图 5.6(C)所示。理想中  $z$  的取值范围是:

$$\frac{0.4}{K} < z < \frac{0.8}{K}。$$

## 5.4 SDP 模型的进一步研究与应用

这一小节中我们进一步的验证聚集度与人类行为的一致性, 并且在各种人群运动的视频和细菌菌落中应用验证我们的算法。

### 5.4.1 聚集度与人类感知

为了量化验证我们提出的聚集度的概念, 我们将其与人类运动作比较。实验中采用的人群运动的数据库包含 62 个运动场景下的 413 幅视频剪辑。每个剪辑的视频包含 100 帧图片。为了得到接近于实况的结果, 我们将视频中人群运动分成高中低三个等级。之后我们提出了两个标准来评价论文中我们阐述的聚集度与划分好的实际人群运动之间的一致性。

第一个标准是群得分与聚集度之间的联系。所谓人群得分 (score), 我们低密度人群设置权值为 0, 中密度为 1, 高密度为 2。我们将这些分数相加作为这个视频的人群得分。这个分数的范围是[0,20]。人群得分的直方图如图 5.7 所示。





图 5.7 实际场景分析

图 5.7 整个视频数据库中部分视频的分析结果。从该图我们可以看出聚集度，人群得分以及速度的大小。利用第三章介绍的 KLT 算法，我们提取出每一个视频剪辑的特征点，可以计算出这段视频剪辑中的聚集度以及人群的运动速度。而图 5.8(A)中描绘的是所有剪辑视频中人群得分与  $\Phi$  和速度之间的关系。可以看出，聚集度  $\Phi$  与人群得分之间的关联更紧密，并且聚集度的结果与我们人类自己感知一致。

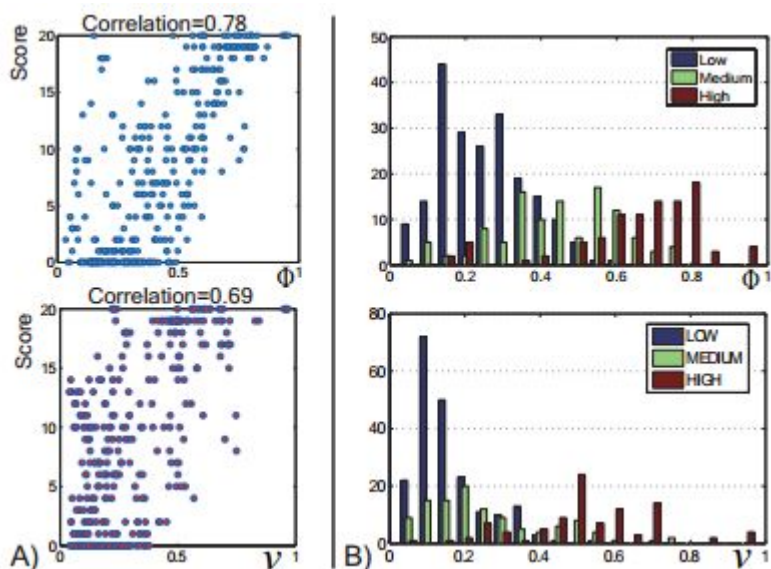


图 5.8 得分与聚集度直方图

第二个评判标准就是基于聚集度的分类准确性。我们将所有的视频分成三类，然后评估聚集度对它们的分类效果。图 5.8(B)展示了三种分类方式下聚集度  $\Phi$  与速度  $v$  的直方图。从

图中可以看出  $\Phi$  相比于  $v$  对视频有着更好的区分效果。图 5.9(A)中刻画了 ROC 曲线以及在高-低、高-中、低-中三种二分分类方式下  $\Phi$  雨  $v$  的最佳准确率和边界值。我们能够看出  $\Phi$  能够比  $v$  更好的区分不同运动状态下的人群运动，尤其是在高-中，中-低分类当中。这表明我们用来描述聚集度的  $\Phi$  能够很好的反应人群运动的动态变化。

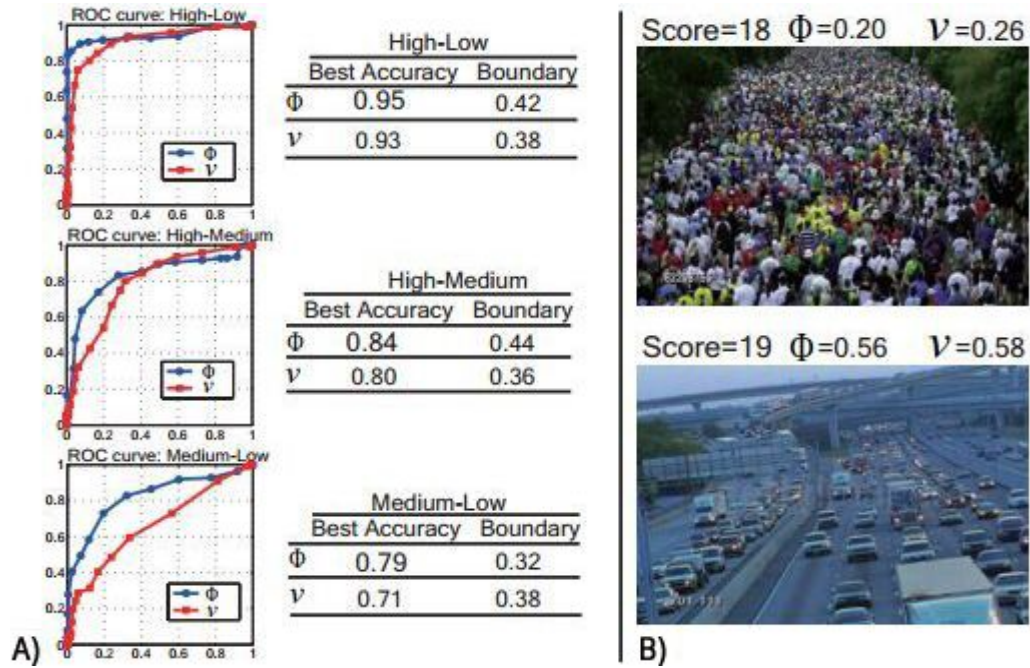


图 5.9 聚集度分类场景

不过图 5.9(B)反应了使用聚集度分类失败的两种场景。由于在高-中，中-低的运动中存在灰度区域，并且这些例子人类自己也无法很好的识别。此外，聚集度  $\Phi$  在跟踪失败、投影失真和特殊场景结构中也不能很好的工作。(B)图中两处场景计算得出的聚集度值比较低，这是由于透视畸变使得 KLT 算法跟踪时没有很好的捕捉运动特征。而虽然人群得分很高，但仅仅是因为场景中有规则的行人走动以及汽车运动。

5.4.2 视频信息中的人群运动监测

我们依旧利用上一小节中的人群运动的视频库监测人群运动。实际上由于提取的特征运动轨迹短而且分散以及外部特征的轨迹干扰，为研究视频中的人群运动带来一定的困扰。



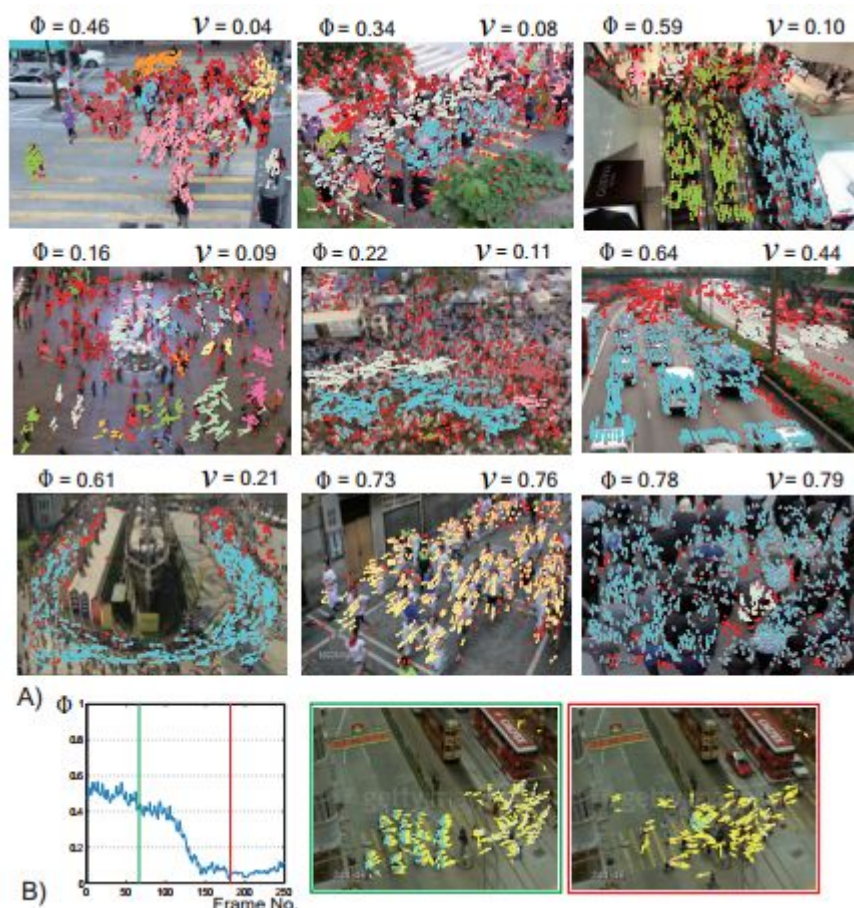


图 5.10 大规模应用场景

在图 5.10 中，分别计算显示了九种视频场景下的聚集度  $\Phi$  与运动速度大小。这些视频包含了我们研究人流密度的很多实用场景，比如道路、购物中心、广场等。其中每个视频上显示的带有箭头的颜色点表示运动的特征点，箭头方向表示特征点的方向。相同颜色的特征点表示相同的特征聚类，红色表示外部无关的特征。运动场景与运动状态的不同，人群运动的聚集度  $\Phi$  也随之变化。另外， $\Phi$  还能够监控人群运动中的动态变化。(B)图中， $\Phi$  值出现急剧的下降，这是因为两组人群相遇时候彼此擦身而过。 $\Phi$  值类似于这样的变化表明人群运动的状态变化或者到达了一个新的关键。这对于人群运动的研究十分有帮助意义。

### 5.4.3 细菌菌落中的群体运动

在这一小节，我们利用聚集度  $\Phi$  的概念研究细菌的集体运动。野生型枯草芽孢杆菌菌落生长在琼脂培养基，细菌在其中自有的游动。细菌个体的真实运动数据来源于<sup>[32]</sup>。每一帧中大约有 200~400 的细菌。实验结果如下图所示：

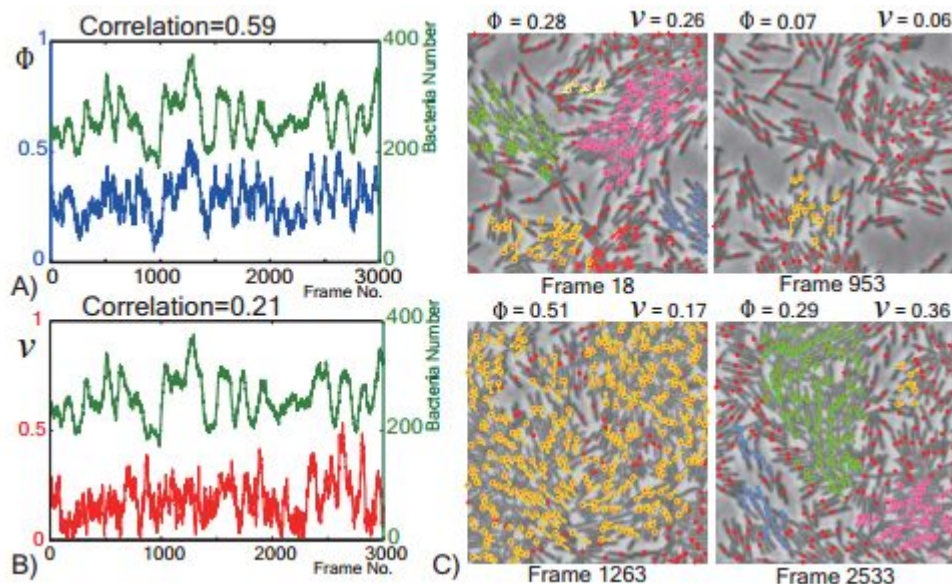


图 5.11 细菌菌落运动信息图

图(A)与图(B)反应的是  $\Phi$  与  $v$  随细菌个数的变化而变化的结果图。其中相比于 B 图, A 图中聚集度  $\Phi$  与个数变化的联系更为紧密。C 图是选取第 18、953、1263、2533 这四帧下实验结果展示。不同的颜色代表不同的特征分类,红色的十字表示的是随机运动的与特征群无关的外部细菌。群体密度被证明是形成集体运动的十分重要的原因<sup>[32,26]</sup>。为了研究它们之间的关系,学者们已经进行了大量的研究。在相同的环境中,同一种类的细菌聚集度会随着密度的增加而单调增加,图 5.11(A)恰好反映了这一点。

## 5.5 聚集度与人流密度的结合应用

聚集度不仅能够反应了人群的运动状态。通过聚集度我们能够从数值上了解到个体参与群体运动的程度。同时聚集度的概念也能够定性的反应出人流密度。在实际应用中,如果将聚集度的计算与人流密度的检测合二为一,能够更好地帮助我们掌握人群运动的实际状况,包括聚集度,人流密度、实时运动速度以及运动方向等信息,避免意外的发生。突发特殊情况时,也有助于我们及时的做出处理,尽可能将损失降低到最小。以下三张图片结果均 Microsoft Visual Studio 环境下实现。

### (一) 低密度人流

如图 5.12 所示,此视频画面中只出现了一个行人。人流密度值较小,相应的聚集度值也比较低。程序共提取出了 6 个特征点,其中黄色带有箭头的实线表示特征点的运动方向,实线的长度相对表明了特征点的速度大小。Order 值表示速度的大小。没有黄色实线的特征点表示该点此刻处于静止状态,瞬时速度为 0。不过我们可以发现视频中有三个特征点不属于行人,占到了所有特征点的一半,故而误差率较大,需要日后的进一步研究探索,降低低密度



人流情况下的误差率。



图 5.12 低密度人流

(二) 中密度人流

图 5.13 中不同颜色的特征点表示属于不同的群体。随着人数的增多，视频中的特征点的数目也有明显增长。有的行人身体上特征密集分布，有的则分布地比较稀疏。可以看出相比于低密度人流情况，误差得到了改善，并且聚集度值也增大。箭头指示的方向基本与行人的运动方向吻合。



图 5.13 中密度人流



图 5.14 高密度人流

高密度人流场景下，人数与特征点数目进一步增加。特征点误判的概率进一步的下降，并且聚集度的值达到了 0.75，也反应了此刻较高的人流密度。不过同中等密度情景一样，高密度人流特征点分布依旧不均匀。并且由于人流较为聚集，人群运动速度较为缓慢，代表速度大小的箭头长度也比较短。这些特点符合客观事实。

表 5.1 人流密度与聚集度对应表

密度等级	极低密度	低密度	中等密度	高密度	极高密度
场景人数	0-10	11-30	31-60	61-100	>100
聚集度	0-0.2	0.2-0.4	0.4-0.7	0.7-0.85	0.85-1

5.6 本章小结

本章主要研究了群体聚集运动中十分常用的研究模型——SDP 模型，首先阐述了该模型的特性：速度基本恒定，方向是所有个体的平均方向。然后引出并着重分析了 SDP 模型中聚集度的概念以及测量方法。之后简单探讨了聚集度三大性质。随后列举了群体聚集算法的步骤。接着又通过仿真验证了聚集度在单个群体和混合群体运动中的准确性。在论证 SDP 模型的进一步研究中，我们通过人类自己对实际人流状况的感知与聚集度作出比较，得出聚集度反应的情况与我们的实际感知相吻合。紧接着又通过在细菌菌落中的运动再次验证了 SDP 模型广泛的适用性。通过将聚集度与人流密度相结合，得到运动人群的速度大小、方向、密度、聚集度等信息，更好地分析人群的运动状态。最后给出了不同密度等级下聚集度的范围。

## 第六章 总结与展望

### 6.1 总结

在诸如车站、重点景区等公共场所，人流密度检测系统的应用越来越普及，对此，本文主要对监控视频信息中人流密度这一热点方向进行了研究。

本文完成的工作主要有以下几点：

(1) 在图像预处理技术中，在探讨了一些基础知识后着重研究了边缘检测，并对比了三种边缘检测算子。因为边缘检测在分析提取特征中有着较多的应用。

(2) 研究了特征的概念，在简单分析几种特征提取的方法之后重点研究了 KLT 算法。其中 KLT 算法可以分成两个步骤：一是如何选取特征；二是如何跟踪特征。我们计算待跟踪窗口在视频帧间的灰度差平方和作为度量的标准。

(3) 通过实验总结对比两种常见的人流密度检测算法后发现基于像素统计的算法计算的复杂度低，适合低密度人流，在高密度场景准确率低；基于纹理分析的算法虽然能在高密度场景下获得不错的效果，但是实现复杂。

(4) 介绍了 SDP 模型的概念，探讨了 SDP 模型中聚集度测量的步骤，并介绍了聚集度的三大特征。结合仿真验证了聚集度的准确性。在深入的研究 SDP 模型后，总结出聚集度与人流密度的关系，并因此将人流密度与聚集度有机结合在实际应用中，展示了其良好的应用前景。

### 6.2 展望

本文对人流密度检测的算法进行了初步的探索，在实验验证中取得了一定的成功，不过在实际应用中还会不断涌现新的问题，需要进一步的研究。

(1) 本文的验证场景仅仅包含单纯有人流的理想场合，但如果是车站等场景，当行人携带行李时候，难以区分行李与行人，会对实验结果造成影响。

(2) 伴随着嵌入式技术的发展，在嵌入式系统中已经移植了很多人流密度监控系统的应用。如何将本文中论述的方法与嵌入式应用相结合，是今后一个研究方向。

## 参考文献

- [1] Tommy Chow, Siu Yeung Cho and Chi Tat Leung, A Neural — Based Crowd Estimation by Hybrid Global Learning Algorithm [J]. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics. VOL 29, No. 4, AUGUST 2001:535—554.
- [2] Nikos Paragios, Visvanathan Ramesh. AMRF-based Approach for Real-time Subway Monitoring [C]. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics. 2001 (1):1034-1040.
- [3] Daies A.C., Yin J.H., Velastin S. A. et al. Crowd monitoring using image processing [J]. IEEE Electronics and Communication Engineering Journal, 2003, 7 (1):37-47.
- [4] Chow T.W.S., Yam J.Y.F., Cho S.Y. Fast training algorithm for feed forward neural networks: application to crowd estimation at underground stations [J]. Artificial Intelligence in Engineering, 2004, 13(3):301~307
- [5] Paragios N., Ramesh V. A MRF-based approach for real-time subway monitoring [J]. IEEE Computer Vision and Pattern Recognition, 2001, 14(1):1034~1040
- [6] Marana A.N., Costa L.F., Lotufo R.A., et al. On the efficacy of texture analysis for crowd monitoring [C]. International Symposium on Computer Graphics, Image Processing and Vision, 2008:354~361
- [7] Siu Yeung Cho, Tommy W.S. Chow, and Chi—Tat Leung. A Neural—Based Crowd Estimation by Hybrid Global Learning Algorithm [J]. IEEE Transactions on Systems, Man, and Cybernetics-- Part B: Cybernetics, Vol. 29, No. 4, August 2009:535-541.
- [8] Tommy W.S. Chow, Siu Yeung Cho. Industrial neural vision system for underground railway station for surveillance [J]. Advanced Engineering Informatics, 16 (2002):73—83.
- [9] H. Rahmalan, M. Nixon, and J. Carter. On crowd density estimation for surveillance [J]. Institution of Engineering and Technology, 2006, 8(12):315~325
- [10] Lin S.F., Chen J.Y., Chao H.X. Estimation of number of people in crowded scenes using perspective transformation [J]. IEEE Transactions on Systems, Man and Cybernetics—Part A: Systems and Humans, 2011, 41(6):645~653
- [11] Siu Yeung Cho, Tommy W.S. Chow, and Chi Tat Leung. A Neural—Based Crowd Estimation by Hybrid Global Learning Algorithm [J]. IEEE Transactions on Systems, Man, and Cybernetics-- Part B: Cybernetics, Vol. 29, No. 4, August 1999: 535-541.
- [12] Schofer J., Ushpiz A., Polus, A. Pedestrian Flow and Level of Service. Journal of Transportation Engineering [J]. 1983, 109(1):46-56.

- [13] 夏令.铁路客运枢纽客流人数自动识别系统[D].北京交通大学硕士学位论文, 2011, 6
- [14] Haralick R M. Statistical and structural approaches to texture[J]. Proc. IEEE, 1979, 67(5):786-804.
- [15] 刘祝华. 图像去噪方法的研究[D].江西师范大学.2005.
- [16] Rafael C. Gonzalez, Richard E. Woods 著, 阮秋琦, 阮宇智等译.数字图像处理[M].第三版.电子工业出版社.2011.
- [17] 罗培全.图像处理技术在客流统计系统中的应用及研究[D].上海: 上海交通大学, 2006.
- [18] Gray Bradski, Adrian Kaebler 著,于仕琪,刘瑞祯译.学习 OpenCV (中文版) [M].第一版.清华大学出版社.2011.
- [19] 王选贺.跨摄像机多人目标的跟踪研究[D].浙江大学.2011.
- [20] Antonio Albiol, Maria Julia Silla, Alberto Albiol and Jos"e Manuel Mossi: Video Analysis using Corner Motion Statistics. Proceedings 11th IEEE International Workshop on PETS[C]. 2009:31-37.
- [21] Donatello Conte, Pasquale Foggia, Gennaro Percannella et al. A Method for Counting Moving People in Video Surveillance Videos. EURASIP Journal on Advances in Signal Processing[J]. 2010.
- [22] Mian A S. Realtime visual tracking of aircrafts[C].Proceedings of the 2008 Digital Image Computing: "Techniques and Applications, Washington, DC, USA, December 1 — 3, 2008.
- [23] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. Physical Review Letters, 2005.
- [24] V. Rabaud and S. Belongie. Counting crowded moving objects. In Proc. CVPR, 2006.
- [25] B. Zhou, X. Tang, and X. Wang. Detecting coherent motions from crowd clutters. In Proc. ECCV, 2012.
- [26] B. Zhou, X. Wang, and X. Tang. Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents. In Proc. CVPR, 2012.
- [27] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, et al. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. Proceedings of the National Academy of Sciences, 2008.
- [28] J. Miller and S. Page. Complex adaptive systems: An introduction to computational models of social life. Princeton Univ Pr, 2007.
- [29] H. Zhang, A. Ber, E. Florin, and H. Swinney. Collective motion and density fluctuations in bacterial colonies. Proceedings of the National Academy of Sciences, 2010.
- [30] 岳晓娟. 智能监控系统的人群密度检测算法研究[D].哈尔滨工业大学, 2013.
- [31] 王雅琳. 基于灰度共生矩阵的人群密度估计算法研究[D].西安科技大学, 2013.
- [32] 蒋绪团, 李庆敏. 基于实时背景提取的人群密度检测[J]. 电脑知识与技术, 2010, 15: 3996-3998.

- [33] 陈鸿. 重点视频监控区域的人群密度估计改进算法的研究与实现[D].南京邮电大学,2013.
- [34] 黄璐,林燕. 一种基于视频的车站人群密度估计算法[J]. 计算机时代,2012,07:23-25.
- [35] 张文倩. 视频监控中人群密度估计研究[D].中北大学,2013.
- [36] 柴进. 视频监控中的人数统计和人群密度分析[D].西安电子科技大学,2011.
- [37] 刘世清. 实时背景提取到人群密度检测技术[J]. 数字技术与应用,2014,05:127.
- [38] 黄敬堃. 视频监控中的人群密度估计和人数统计技术研究[D].杭州电子科技大学,2014.
- [39] 张燕. 机场航站楼视频监控系统的人群密度估计研究[D].南京航空航天大学,2012.
- [40] 刘福美,黎宁,张燕,张可. 一种基于图像处理的人群密度估计方法[J]. 计算机与数字工程,2011,05:118-122+170.
- [41] 李亚飞. 人群密度估计的算法研究[D].河北师范大学,2014.
- [42] 沈娜. 智能视频监控系统中人群密度估计与实现[D].南京航空航天大学,2012.
- [43] 戢玲玲,陈黎,张晓龙. 基于 KLT 算法的驾驶员疲劳检测方法[J]. 计算机工程与应用,2010,31(2):436-442



## 致谢

三年的研究生生活即将结束，也即将离开热爱了七年的南邮校园。在论文即将结束之际，首先要感谢我的导师龚建荣老师。龚老师两年多以来对我的敦敦教导，言传身教，使我受益匪浅。不但能够在具体问题上给我帮助，更能在更高的学科方向上为我指路。龚老师渊博的学识，严谨的治学态度，平易近人的风范都给我留下了深刻的印象，是我终身学习的楷模。在此，谨向龚老师致以我崇高的敬意与衷心的感谢。

另外也要感谢李晓飞老师，为我们提供了学习与生活上的莫大帮助，使我们能够在研二一年赴昆山市工业技术研究院学习锻炼的机会。感谢在昆山期间与我们朝夕相处的刘浏老师对我们学习生活的解疑答惑，使我们不断进步。

还要感谢我的同门们张聪、赵晓伟、张群还有同处一个实验室的顾婷婷、张文义、南玮建、朱文、车少帅、赵光明、刘梦。感谢你们这两年多时间里的陪伴，给我们的学习与生活带来了无限的欢歌笑语。

最后向百忙之中参加答辩和审阅论文的每一位老师表示衷心的感谢。