

Information Fusion for
Distributed Networked Systems

分布式网络化系统 信息融合

刘莉 等 编著
Liu Li

清华大学出版社



Information Fusion for
Distributed Networked Systems

分布式网络化系统 信息融合

刘莉 周文举 张小峰 陶强 编著

清华大学出版社
北 京

内 容 简 介

分布式网络化系统由于受到通信带宽等因素的制约,降低了系统的性能。传统的信息融合方法主要考虑由单一属性传感器感知信息,难以满足被测目标状态精度的实际需要。本书主要内容如下:首先,提出基于活动轮廓模型的多相双曲线图像分割方法,用于提高图像分割的精度;其次,研究具有鲁棒性的数据集分类方法,降低分类过程中参数不确定性的影响;第三,提出加权融合重组新息和误差互协方差实现信息交互,进而提出基于估计的线性时滞补偿策略,重组带有时间戳的测量序列;最后,以仿真结果验证所提方案的有效性,进而提高系统的可靠性和定位精度。

本书可作控制科学与工程及其他相关专业和研究方向的研究生或高年级本科生的参考书使用,也可供最优估计理论及其应用的相关领域的工程人员和科技工作者自学和参考。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

分布式网络化系统信息融合/刘莉等编著. —北京:清华大学出版社,2018

ISBN 978-7-302-50781-9

I. ①分… II. ①刘… III. ①分布式算法 IV. ①TP301.6

中国版本图书馆 CIP 数据核字(2018)第 178639 号

责任编辑:郭 赛

封面设计:何凤霞

责任校对:焦丽丽

责任印制:沈 露

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者:北京鑫海金澳胶印有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:9.75

字 数:182 千字

版 次:2018 年 11 月第 1 版

印 次:2018 年 11 月第 1 次印刷

定 价:39.00 元

产品编号:079733-01

前

言

为了适应不断增长的信息计算和感知需求,网络化系统常采用分布式结构用于提高系统的适应性、效率性和自动化性能,成功地应用在信息物理系统、智能电网、智能交通以及通信网络等重要领域。但是由于受到通信带宽等因素的制约,网络化系统不可避免地存在网络诱导的通信约束现象,这些现象通常会降低系统的性能。因此,研究网络引起的现象对于复杂网络化系统具有重要的意义。本书依据网络化系统中的空间定位与目标跟踪问题,研究分布式结构中信息传输受通信约束的信息融合策略,目的在于提高系统的可靠性和定位精度。主要内容概括如下:

首先,研究由视觉传感器获取的图像信息对其进行精确识别,这是准确定位目标的基础。提出一种基于活动轮廓模型的多相双曲线图像分割方法,该方法采用最大似然估计和期望最大化算法建立 N 个水平集函数,并把灰度图像分割成 2^N 个子区域。在每个目标子区域,该方法主动将双曲线演化函数嵌入水平集方程,因此图像轮廓的演化曲线能够双边扩展。该方法对于抵抗噪声干扰具有鲁棒性,能够提高图像分割的精度。

其次,依据不同功能的传感器测量得到的复杂数据集,提出一种新的具有鲁棒性的数据集分类方法,该方法利用邻居搜索和内核模糊 C-均值进行数据遍历和聚类。一些优化的策略包括邻居搜索、控制聚类形状和自适应距离核函数,分别用于解决聚类数目、分类方法的稳定性和一致性的问题。经过理论分析,提出的方法具有抵抗噪声干扰的鲁棒性能,可以降低数据集分类过程中参数不确定性的影响,进而更好地将测量信息应用于分布式系统的信息融合。

第三,研究了受传输时延和交叉相关噪声影响的不确定网络化系统的状态估计问题。提出一种基于鲁棒卡尔曼滤波的分布式感知和集中式融合方法,该方法能够提高具有扰动的测量的估计精度。为了描述分布式系统

的信息交互,提出一种加权融合重组新息策略,用来降低计算负担并且抑制噪声的干扰。此外,为了获得最优的线性无偏估计,研究的融合估计方法使用加权误差互协方差实现信息交互,该方法是对每个子系统的再次优化,能够获得更精确的估计值。

第四,针对一类离散时间随机不确定系统研究了其建模和滤波问题。由于数据在传输过程中不可避免地受到随机传输延迟、数据包错序以及相关性噪声的影响,为了确定是否保持数据包错序现象,系统模型的建立分别采用了零阶保持器和逻辑零阶保持器两种信号选择方案。基于所建立的系统模型,应用扩展状态空间和最小化误差协方差矩阵方法,提出了鲁棒有限时域卡尔曼型滤波策略。该滤波方法的估计协方差矩阵可以通过其上界约束推导得到。为了提高滤波器的处理效率,提出一种基于估计的线性时滞补偿策略,用于处理传输时滞现象,并应用在带有时间戳的重组测量序列中。此外,为解决测量缺失以及减轻计算负担,采用一种假定的延迟补偿方法,实现一步预测估计。通过对动态跟踪系统的仿真,结果表明该滤波方法能够紧密跟随系统的实际状态。

最后,在上述理论研究并通过仿真结果验证了所提方案有效性的基础上,结合 3D 光电传感定位系统,分析了基于线性 CCD 的空间定位方法。针对测量得到的空间目标,对其运动轨迹进行了跟踪与估计,提出了一类离散时间不确定系统的建模与估计方法,考虑了数据在从设备到估计器的传输过程中含有网络诱导的随机时延、丢包以及数据包错序的情况。为了丢弃数据包错序进而提高系统性能,系统模型的建立依据逻辑零阶保持器方案。基于有时间戳的重组测量序列,设计鲁棒有限时域卡尔曼型滤波策略,用于估计误差协方差的约束上界。此外,为了获得最优的线性估计,提出了加权的融合估计方法,用于探测信息协作的误差互协方差矩阵。仿真结果表明所提出的方法能够降低网络通信负担,抑制通信约束对测量信息的影响,进而实现提高空间定位精度的目的,具有良好的应用前景。

编者

2018 年 10 月

第 1 章 绪论	1
1.1 研究背景	1
1.2 相关领域的国内外研究进展与现状	3
1.2.1 分布式系统测量方法	3
1.2.2 目标定位方法	4
1.2.3 水平集演化方法	6
1.2.4 数据集分类方法	7
1.2.5 信息融合理论	8
1.3 主要工作和特色	12
1.3.1 研究内容	12
1.3.2 本书特色	13
1.4 章节安排	14
第 2 章 基于多相双曲线水平集演化的图像分割方法	17
2.1 引言	17
2.2 水平集方法预备知识	17
2.2.1 水平集方法原理	18
2.2.2 多相水平集理论	20
2.2.3 基于区域的水平集方法	21
2.2.4 多项水平集方法的能量函数	23
2.3 多相双曲线方法	24
2.3.1 区域控制能量项	25
2.3.2 能量拟合函数	25

2.3.3	水平集曲线演化	29
2.3.4	算法的执行过程	32
2.4	方法实施和实验探讨	32
2.4.1	实验结果	32
2.4.2	结果分析	35
2.5	本章小结	37
第3章	基于邻居搜索和内核模糊 C-均值的鲁棒性数据集分类方法	38
3.1	引言	38
3.2	聚类算法的理论基础	38
3.2.1	聚类分析	39
3.2.2	内核模糊 C-均值算法	40
3.3	鲁棒的数据分类方法	42
3.3.1	邻居搜索遍历数据集	42
3.3.2	初始化隶属度矩阵和聚类原型	44
3.3.3	优化隶属度矩阵和聚类原型	45
3.3.4	鲁棒性能估计	47
3.3.5	算法执行步骤	49
3.4	实验验证	49
3.4.1	实验一: MEMS 加速度计数据集	50
3.4.2	实验二: 复杂图像数据集	53
3.5	本章小结	57
第4章	不确定网络化系统的分布式加权融合估计	58
4.1	引言	58
4.2	最优估计理论和问题描述	59
4.2.1	卡尔曼滤波理论	59
4.2.2	分布式网络化系统融合架构	60
4.2.3	传输时延和交叉相关噪声	61
4.3	分布式鲁棒卡尔曼类型的滤波	62
4.3.1	重组新息序列	62

4.3.2 分布式融合估计	65
4.4 数值验证	72
4.5 本章小节	77
第5章 随机不确定系统的建模与滤波	78
5.1 引言	78
5.2 问题阐述	79
5.2.1 系统描述	79
5.2.2 基于序列重排的建模	79
5.2.3 噪声的相关性	82
5.3 鲁棒有限时域滤波	82
5.3.1 基于 ZOH 的滤波	83
5.3.2 基于逻辑 ZOH 的滤波	88
5.4 仿真结果	92
5.5 本章小结	96
第6章 面向空间定位的不确定系统的建模与估计	97
6.1 引言	97
6.2 基于线性 CCD 的空间定位方法	97
6.2.1 空间定位原理	98
6.2.2 3D 光电传感定位系统工作流程	99
6.2.3 空间定位效果仿真	100
6.3 问题描述和分析	105
6.3.1 基于信号选择方案的系统模型	106
6.3.2 交叉相关性噪声	107
6.4 基于逻辑 ZOH 的估计器	107
6.4.1 子系统的增广状态向量	108
6.4.2 估计协方差的上界	110
6.4.3 随机延迟的线性补偿	111
6.4.4 分布式加权融合估计	112
6.5 数值仿真	114

6.6 本章小结	120
第7章 总结与展望	121
7.1 总结	121
7.2 未来展望	123
附录 A 命题 3.1~命题 3.3 的证明	124
附录 B 定理 4.2 的证明	129
附录 C 定理 5.2 的证明	131
参考文献	136

第 1 章

绪 论

1.1 研究背景

目前,构建于通信和计算基础上的复杂网络化大规模系统已经成为现代通信技术的核心。由于网络化系统具有扩展性好、灵活性强、维护便利等优点,广泛地应用在信息物理系统(cyber physical systems, CPSs)^[1]、通信网络(communication networks)^[2-4]、智能电网(smart grids)^[5]以及智能交通(intelligent transportation)^[6]、环境监测(environmental monitoring)^[7]、远程诊断和故障排除(remote diagnostics and troubleshooting)^[8]等重要研究领域。为了适应日益增长的信息计算量,并提高信息协作感知的需求,分布式结构策略被用于提高物理器件和计算单元之间的协同关系,增加系统的适应性、融合效率和自动化性能。研究基于复杂大规模系统的分布式信息感知与协同,能够有效地利用包含在多个数据集中的数据进行信息计算与估计。为了增加信息的可靠性,提高估算精度并且降低通信负担,分布式估算策略应用于网络化系统的信息融合。本书基于复杂网络化系统的研究与应用,从以下三个方面进行分析。

(1) 通信制约网络环境下的分布式测量方法。分布式网络化测量采用分布式感知信息技术,将测量信息用于目标定位的策略已在传感器网络^[9,10]、雷达^[11]、无线通信^[12]、探测^[13]等领域得到大量研究及广泛应用^[14]。以大型机械制造业领域内的测量与定位问题为例,其测量量程为几米到几十米,针对这个范围内的物体可以测量其三维空间坐标、形貌尺寸以及运动轨迹等,测量精度要求较高,并且面临着测量环境复杂、干扰因素众多、被测对象多样、测量效率低等一系列亟待解决的问题^[15]。通信网络的引入增加了系统的复杂性,同时由于通信网络受到带宽、容量和服务能力的限制,使得传输的数据不可避免地存在通信时延、网络拥塞、数据丢包、多包传输等问题,从

而导致系统性能的下降甚至不稳定^[16]。

(2) 网络化系统的分布式信息融合方法。通信信道参数包括量化、时延、数据丢包、带宽受限、比特率限制等约束,因此对分布式网络化系统的测量不能一味地只注重数据的获取而忽视了数据中所蕴含的信息。由此采用在节点处对其获得的局部数据进行融合和估计的方法以提取有效信息,然后再将融合信息传输给用户的分布式融合策略^[17]。网络化系统中的分布式信息融合方法的基本原理是充分利用多个传感器资源测量被测目标,把多个传感器在时间或空间上的冗余或互补信息依据某种准则进行组合,以获得被测对象的一致性描述。由于传感器之间的冗余信息增强了系统的可靠性,因而互补信息扩展了单个传感器的性能^[18]。网络化的多传感器融合系统(networked multi-sensor fusion systems, NMFSs)^[19,20]极大地增强了系统的可扩展性和可维护性,减少了系统布线,增强了系统的灵活性,并且可以方便地进行信息共享,使得分布式的融合系统结构更易于实现,提高了系统的可靠性和故障容错能力,且基于网络的信息融合系统允许融合中心具有处理远程传感器测量数据的能力。因此,对网络化的分布式信息融合估计已经成为一个重要的发展趋势,其应用范围和作用正在不断地扩大。

(3) 网络化系统分布式测量方法在定位中的应用。近几十年来,为解决高精度测量定位问题,相关的空间大尺寸测量定位系统与测量方法得到了迅速的发展^[21,22],同时网络化的大尺寸测量定位系统与测量方法所覆盖的空间范围越来越大,而且覆盖方式越来越灵活。随着网络规模的日益增大,其通信任务日渐复杂,由于每个节点的能量和计算能力十分有限,只有通过所有节点分担信号处理工作量,协作完成信号处理和通信任务,才能使得网络性能不随网络规模的增大而衰减^[23]。因此,采用分布式的网络化测量结构,在保证测量精度的前提下,通过将测量节点布置于被测物周围,以其较优的精度性能,以及可扩展性、灵活性、可靠性,乃至支持并行多任务测量、数据冗余测量和融合其他传感设备的潜在能力,达到覆盖任意测量空间的目的,且测量精度原理上没有损失^[24]。

然而,分布式网络化系统在通信过程中由于网络的引入使得信息传输模式发生了根本变化,也必然带来通信约束问题,一个典型的网络化多传感器融合系统的结构如图 1-1 所示。当有限的通信网络作为信息传输枢纽时,NMFSs 的重要特点是其用于信息传输的网络带宽是有限的,各节点要通过竞争以获得网络资源,有限的网络带宽使得信息传输出现延迟、丢失等不确定现象。因此,在 NMFSs 中不能通过无限制的信息采集与传输提高融合估计性能,融合中心端必须在通信受限的情况下(如有限的

通信量、信息延迟与丢包)设计融合估计算法。显然,通信受限将不可避免地影响融合估计性能,迫切需要提出适用于通信约束的网络化系统信息融合估计理论与算法。

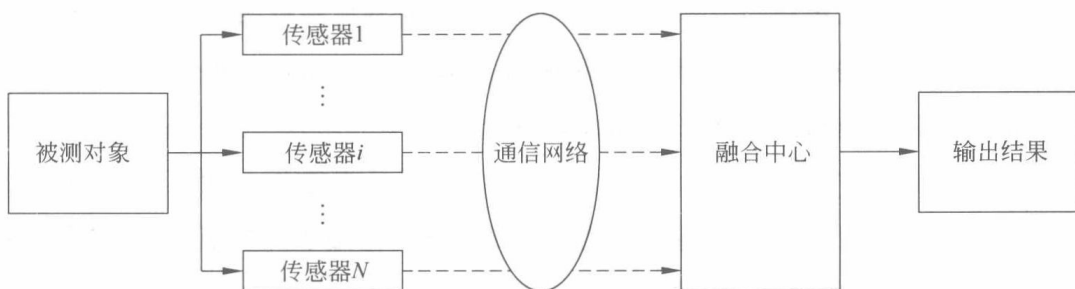


图 1-1 典型的网络化多传感器融合系统

在通信受限的分布式网络化系统中,首先,融合估计往往在受通信约束的不完整信息下进行,如何设计具有不完整信息的融合估计器并建立通信量、时延、丢包特性参数与估计性能之间的关系一直是个研究难点;其次,分布式信息融合估计的关键在于信息融合所处理的多传感器信息具有更复杂的形式,可出现在不同的信息层次上,而且估计方法与融合策略紧密联系,当进一步考虑通信受限条件时,相应的分析和设计显得更加困难;最后,网络化估计问题的研究成果^[25,26]大多数都是讨论单个传感器情形下的信号估计问题,因此研究受通信约束的分布式多信息融合方法是富有挑战性的研究课题。

1.2 相关领域的国内外研究进展与现状

目前,围绕网络化系统的分布式信息融合方法及其在空间目标定位中的应用,学术界开展了一系列研究工作,其中部分关键问题研究如下。

1.2.1 分布式系统测量方法

以分布式网络结构为核心的分布式测量定位系统可以具体分成以下形式。

(1) 分布式视觉测量系统^[27]。依靠相机成像以及相关图像处理技术^[28]实现位置测量。周婧等^[29]利用移动视觉测量系统解决有遮挡条件下被测点的坐标测量问题,用不同站位测量结果的合理估计求最佳坐标,依据特征向量的稳定理论分配融合权重,将全局归一化的空间坐标进行融合,得到融合估计的最终表达式,给出了最优的估计结果。考虑到摄像机运动产生的多聚焦图像的情况,基于梯度的多尺度加权的多聚焦图像融合方法^[30]被提出,该方法应用在多聚焦图像由于对象或摄像机运动而产生

的各向异性模糊问题,更好地解决了多尺度边缘和角落结构的清晰度。白相志等提出基于四叉树分解策略的加权焦点测量方法实现了多聚焦图像的融合^[31],该方法将源图像分解成通过四叉树结构中的最优尺寸的块,使用加权的焦点测量方法识别重组的区域,进而产生一个完全的焦点图像。另外,在一部分应用中,权重的分配是靠专家经验给出的,在实践中也得到了很好的应用效果。

(2) 分布式超声测距定位网络^[32,33]。依据超声测距原理^[34],并利用多点定位方法实现坐标测量,而对其进行优化分析需要从超声测距的原理和特点出发,以及与多点定位模型结合,寻找合适的处理方法。文献[35]以分析与优化分布式空间测量定位系统 wMPS 的性能为主题,指出分布式空间网络结构是影响系统性能优势的关键因素,从单发射站、网络布局以及加权测量等方面入手,对 wMPS 网络结构性能进行深入分析,并给出相关优化方法。天津大学的吴军博士^[34,36]等提出了一种基于构造三维温度场的方法,对超声传播路径上的声速作精确补偿,该方法通过温度传感器探测测量场的边界温度值,采用有限元算法精确拟合测量空间内的温度分布。相对于传统的通过计算温度平均值补偿声速的方法,该方法更适合补偿工业现场中非线性温度分布对声速造成的影响,优化超声收发器节点间的超声传播质量。关于超声定位中超声收发器节点的加权方法,Luca Mastrogiacomo 等在研究定位算法时,引入标准器件(其上任意两点间的距离被认为是准确的),这样就引入了基准,利用最小二乘算法迭代方差,可缓解不同精度测距带来的误差,其效果相当于为超声定位节点引入权重^[37]。

综上所述,被测信息经过传感器感知,用于测量的传感器采用分布式的网络结构,每个传感器相当于一个测量节点,用于测量的节点间的关系一般集中在融合算法方面。然而,在目前关于测量方法的研究成果中,多数侧重于融合算法中网络的优化以及加权融合,较少从降低数据的冗余和减少网络传输压力的角度进行研究。因此,需要结合感知节点的分布式结构、有限计算能力等特点,深入研究分布式系统被测信息的感知和传输方法,这也是后续进行信息融合设计的基础。

1.2.2 目标定位方法

当前主流的定位方法有基于测距(range-based)和非测距(range-free)两类。

基于测距的定位算法^[38,39]使用绝对值测量技术,由于可以得到准确的数值,例如节点之间的准确距离、邻居节点之间的相对角度值等,所以定位的精确度较高。基于测距的定位常用几何计算方法,包括三边测量定位方法(trilateration)^[40,41]、三角测量定位方法(triangulation)^[42]、极大似然估计测量定位方法(maximum likelihood

estimation)^[43]、双曲线方法^[44]以及混合定位方法等^[45]。由于上述方法都需要附加额外的昂贵硬件设备,使得基于测距的定位方法在以能量约束、低成本为特色的传感器网络中的应用受到限制。因此基于非测距的定位算法被提出。

基于非测距的定位算法由于没有使用附加的外部设备获得精确的用于节点坐标计算的测量数据,所以这类算法的定位精度相对较低,但是该算法的能量代价小,同时在性能提高方面存在巨大的潜在空间。例如质心(centroid)逼近定位算法^[46],每个未知位置的节点(unknown node)利用能够监听到的锚节点(anchor node)的坐标,计算这些锚节点坐标的中心,获得自己的位置估计。如果锚节点的分布是均匀的,并且未知节点能监听到的锚节点的数量越多,则定位精度就越高。基于节点的连通性和最大速率的定位算法(range-free localization algorithm based on connectivity and maximum velocity of nodes, LACMV)^[47]通过把移动节点的连通性以及它们的最大速率信息建立误差函数,并采用迭代求解最小优化参数,旨在跟踪人类行为所表征的轨迹。基于跳数的定位算法 DV-Hop^[48]使用了距离向量路由传播数据包,并且每个传感器节点都维护一个跳数计数器,用以记录此节点相对每个可收听到的锚节点的最小跳数。当某个节点接收到一个新的锚节点信息包时,如果其中包含的跳数值小于节点当前保存的相对于该锚节点的最小跳数值,那么这个接收节点就更新自己存储的跳数值,并且把锚节点信息包中的跳数值加1,同时把信息包转发出去。

针对分布式视觉测量系统,其定位方法是采用视觉传感器获取信息,确定位姿参数并建立数学模型。视觉技术指通过视觉传感器获取的先验知识完成分析处理工作,通过选择不同的图像特征,有感知地实现研究对象的选择。为了分析多重表达数据间的典型相关性,YUAN等提出了一种正规化的多集图像特征提取方法,用于处理多重表达数据的典型相关性、区别性和内部几何结构^[49]。由于数据收集阶段的不确定性,考虑所获取的图像信息受到光照不均和背景纹理的影响,导致图像噪声大和识别效率低的问题,文献[50, 51]提出了基于双边梯度演化曲线的多水平集图像分割方法,提升了图像分割速度,提高了图像识别精度。文献[52]针对环境不确定条件下的相机的移动,设计有条件的视觉同时定位与地图创建系统,该系统在预测阶段利用图像数据建立相机状态模型,以提高预测精度和降低计算负担。南京航空航天大学的冯春博士^[53]使用单目视觉技术提取图像轮廓,估计目标物的深度,提出了基于单目视觉的航天器间相对位姿参数确定算法,实现位姿参数估计的实时性和高精度。

在具体的实现过程中,由于视觉信息成像存在测量误差、特征提取及匹配误差、量化误差等不可避免的影响,因此针对复杂的环境设计出高性能的定位算法,有必要研

究更加稳定、可靠、实用的相对位姿参数估计算法以提高估计精度。对于定位而言,如果能处理好定位所需要的数据,则能够提高定位的精确度和优化定位算法的性能。

1.2.3 水平集演化方法

为了解决分布式视觉测量系统中图像信息的处理问题,水平集方法用于分割信息轮廓。基于水平集方法的曲线演化模型通常分为两类:参数模型和几何特征模型。参数活动轮廓模型^[54,55]由拉格朗日方程显式构成,由于经典的 Snake 模型^[54]容易陷入局部极值,并且该模型在初始位置对于拓扑结构的改变比较敏感,因此使用该模型处理复杂图像分割是困难的。几何主动轮廓模型^[56-58]由隐式欧拉方程组成,该模型对于处理闭合运动接口的几何与拓扑结构的变化是有效的。LI 等^[59]提出了一种基于边的方法,使用局部边缘信息描述对象边界的主动轮廓曲线,该方法在能量函数中引入惩罚项,使曲线在演化过程中能够保持符号距离函数(signed distance function, SDF)的值,但是需要预先设置曲线演化采用扩大或缩小的形式。基于边的图像分割方法对于抵抗噪声的干扰不具有鲁棒性,而且弱边缘容易被泄漏。CHAN 和 Vese^[60]利用变分水平集方法提出了 C-V 模型,具有更大的收敛范围,且对初始的设置更加灵活。Vemuri 和 CHEN^[61]提出了基于区域的分段平滑(piecewise smooth, PS)模型,用于组合形状的先验信息,以实现图像融合和分割。基于区域的主动轮廓线模型取决于每个分割区域的强度均匀性,然而在真实图像中,由不同方式获取的图像强度是不均匀的,因此基于边界和基于区域的方法对于被覆盖的目标和类似的灰度图所处理图像信息的结果不够理想。为了处理一类灰度不均匀图像,HOU 等^[62]介绍了图像预处理的步骤,Michailoich 等^[63]提出了一种基于 Bhattacharyya 梯度流的主动轮廓模型,用于区分演化曲线内部和外部之间的强度分布。多相区域分割方法被广泛地应用于处理强度不均匀的图像,如磁共振成像(magnetic resonance image, MRI)、车辆检测、复杂医学图像和药片包装图像等。为了实现多目标的图像分割,Do 等^[64]验证了多相水平集分割的框架,LI 等^[65]提出了一个加权的 K-均值变分水平集(weighted K-means variational level set, WKVLS)方法,用于在偏置校正场还原真实信号,其中偏置校正模型^[66]采用最大似然(maximum-likelihood, ML)和期望最大化(expectation maximization, EM)方法。然而,该方法的初始化变量需要复杂的计算过程。文献[67]研究了 N 个水平集函数分割 N 个区域的方法,该方法依赖于水平集函数的个数,增加了计算复杂度。为了解决这个问题,Mansouri 等^[68]提出了著名的多水平集分割模型,即多区域竞争模型。文献[69]提出了变分多相水平集方法,并且其偏置校

正对应于似然函数(likelihood function, LF)求解。此外,WANG等^[70]介绍了一阶变分水平集方法,用于分割多相的脑部核磁共振成像。值得注意的是,引入高斯分布是为了加强局部均值和方差统计。

为了分割另一种强度不均匀的图像,W. J. ZHOU等^[71,72]采用精确的图像捕捉控制,实现了快速瓶盖缺陷检测。H. Y. ZHOU等^[73]利用贝叶斯分析法解决了基于水平集的图像处理问题,进而,文献[74, 75]应用梯度向量流和均值漂移方法对皮肤病变图像进行分割。文献[76]提出了一种新的加权距离函数(weighted distance function, WDF),通过建立统一的张量表达多相水平集函数演化曲线。LI等^[77]研究了区域可扩展拟合(region-scalable fitting, RSF)能量最小化方法,用于分割图像的轮廓。文献[78]提出了一种隐式主动轮廓模型,采用反应-扩散水平集演化(reaction-diffusion level set evolution, R-D LSE)方式重新初始化自由水平集演化曲线。文献[79]中应用多相水平集方法分割强度不均匀的图像(3-Phase LSE)。

1.2.4 数据集分类方法

在网络化系统中,不同功能的传感器测量得到的信息表示不同的物理含义,需要对所获取的数据集进行分类。数据分类方法中,模糊聚类作为无监督的学习方法,通常可以分为硬 C-均值(hard C-means algorithm, HCA)^[80]和模糊 C-均值(fuzzy C-means algorithm, FCM)^[81]两种。相对于硬 C-均值聚类算法,模糊 C-均值算法对数据表示和聚类效果具有更好的性能。基于此,WU等^[82]提出了一种新的衡量标准,即替代硬 C-均值(alternative hard C-means, AHCM)和替代模糊 C-均值(alternative fuzzy C-means, AFCM)聚类算法。这两种算法修改了距离函数,但是聚类结果仍是模糊的,因为在迭代过程中,其距离函数值小于1。Ahmed等^[83]提出了一种偏差纠正模糊 C-均值(bias corrected FCM, BCFCM)算法,这种方法引入了一个在空间上的直接邻居项,用于给其直接邻居点贴上标签,其缺点是在迭代中会消耗大量的时间。此外,CHEN等^[84]研究了基于 BCFCM 算法的图像平滑预处理步骤,但它无法控制平滑以及聚类之间的权衡。文献[85]中,Szilagyi等提出了一种改进的模糊 C-均值聚类(enhanced FCM, EnFCM)算法,使用偏差解决分段均匀的点标记问题。虽然有关分类是有效的,但未能根据标签上的间隔进行分段划分。一个半监督模糊 C-均值聚类(semi-supervised FCM, SS FCM)算法由 LI等^[86]研究,提出了一种新的迭代方法计算有标记样本的隶属度,克服无标记样本的隶属度计算复杂的缺点。值得注意的是,上述 FCM 聚类方法应用了具有非鲁棒性能的欧几里得距离(Euclidean distance),它

不能抵抗噪声的干扰,因此在实际应用中会带来诸多不便。

虽然 FCM 方法已成功地应用于许多领域,但是它无法选择模糊加权指数,且聚类原型的确定依赖先验知识。文献[87]提出了分层多关系聚类算法,解决一对多关系的数据集分类。一种子空间聚类混合模型(subspace clustering mixed model, SCMM)^[88]应用在订单生产调度方法中。在此基础上,“内核法”是一种可行的数据集分类的方向^[89]。经典的内核模糊 C-均值(kernel fuzzy C-means, KFCM)^[90]方法受到自动可变权重系数的影响,在聚类过程中利用自适应的距离设置变量的权重系数。Krinidis 等基于窗口大小信息,提出了一种鲁棒的具有模糊局部信息的 C-均值(fuzzy local information C-means, FLICM)聚类算法^[91]。然而,窗口大小信息的引入使噪声对信息的影响被放大。根据 FLICM 算法,GONG 等^[92]研究了一种权衡加权模糊因子的自适应控制局部空间关系的聚类方法。为了进一步把 FLICM 算法扩展到更高的维度,Krinidis 等在文献[93]中介绍了一种通用的局部信息模糊 C-均值(Generalization clustering algorithm based on the fuzzy local information C-means, GFLIM)聚类算法,进而将模糊加权 C-均值聚类方法应用到高光谱数据^[94]。文献[95]提出了基于区间的多维不确定性数据流(uncertain interval data micro, UIDMicro)聚类算法。为了克服初始位置对分类结果的影响,文献[96]使用基于背景差分法的多通道内核模糊相关图检测运动物体,文献[97, 98]提出了基于主动轮廓模型的模糊聚类方法,并将其应用于医学图像数据集分割。快速空间内核模糊 C-均值(fast spatial kernel fuzzy C-means, FKFCM)方法使用核函数,同时分配更多的权重给相近像素的图像。LU 等^[99]研究了一种改进的模糊 C-均值算法,为了提高分类精度,基于相对密度等级,该方法引入了一种新型的密度性距离度量。QIU 等^[100]提出了增强的区间 2-类型 C-均值(enhanced interval type-2 fuzzy C-means, EIT2FCM)算法,用于提高初始中心的精确度。李永忠等^[101]提出了一种将云模型和半监督聚类相结合的入侵检测算法,应用于网络安全防御体系,强化数据分类。需要指出的是,上述方法的实现需要预先设置聚类数目,且初始聚类中心也不确定。为了克服这些问题,基于多尺度自回归(multiscale auto regressive, MAR)模型与模糊 C-均值(FCM)聚类的图像数据集分割方法被提出^[102]。在鲁棒性估计方面,文献[103-106]描述了鲁棒性理论和分析处理过程。然而,这些基于内核函数的 FCM 算法在减少时间消耗和优化数据中心资源方面是比较困难的。

1.2.5 信息融合理论

信息融合的功能可以概括为扩大时空搜索范围,提高目标可探测性,改进探测性