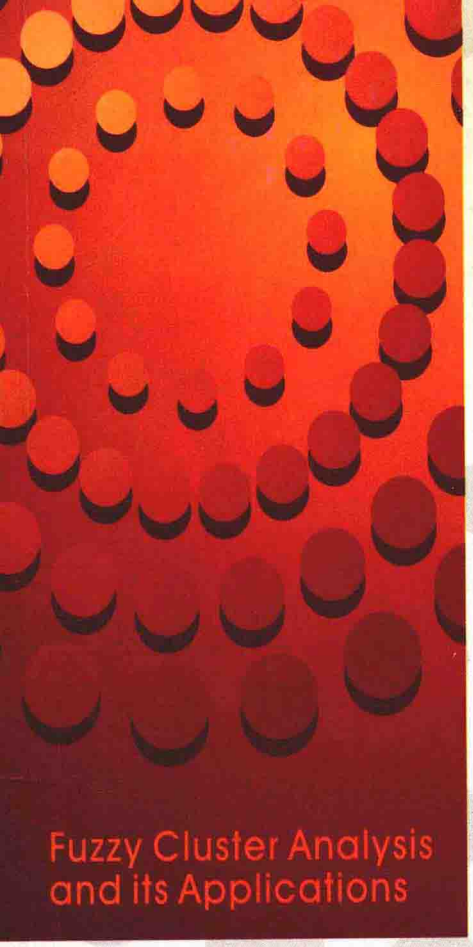
现代电子信息工程理论与技术丛书



模糊聚类分析及其应用

Fuzzy Cluster Analysis and its Applications

高新波 著

西安电子科技大学出版社
http:// www.xduph. com

现代电子信息工程理论与技术丛书

模糊聚类分析及其应用

高新波　著

西安电子科技大学出版社
2004

内 容 简 介

模糊聚类分析是非监督模式识别的重要分支，在模式识别、数据挖掘、计算机视觉以及模糊控制等领域具有广泛的应用，也是近年来得到迅速发展的一个研究热点。本书系统地论述了基于目标函数模糊聚类的基本理论、方法，以及现存的许多开放性的问题与初步的研究成果，主要内容有：模糊数学与可能性理论基础，谱系聚类、基于等价关系的聚类和图论聚类方法，c 均值类型的基于目标函数的模糊聚类方法及存在的问题，模糊聚类神经网络，模糊聚类遗传算法和进化策略，模糊聚类的原型初始化方法，模糊聚类的有效性分析，模糊聚类的聚类趋势分析，区间值数据的模糊聚类分析及其推广，以及模糊聚类在图像分割和模式识别中的应用。

本书可以作为理工科大学计算机、自动控制、信号与信息处理、电路与系统、系统工程等专业的博士生、硕士生及高年级本科生的教材，同时对有关领域的研究人员和工程技术人员也有重要的参考价值。

图书在版编目(CIP)数据

模糊聚类分析及其应用/高新波著.

—西安：西安电子科技大学出版社，2004.1

(现代电子信息工程理论与技术丛书)

ISBN 7-5606-1301-2

Ⅰ.模… Ⅱ.高… Ⅲ.模糊集理论 Ⅳ.O159

中国版本图书馆 CIP 数据核字(2003)第 085819 号

策　　划　臧延新　陈宇光

责任编辑　龙　晖

出版发行　西安电子科技大学出版社(西安市太白南路 2 号)

电　　话　(029)8242885　8201467　　邮　编　710071

http://www.xduph.com　　E-mail：xdupfxb@pub.xaonline.com

经　　销　新华书店

印刷单位　西安兰翔印刷厂

版　　次　2004 年 1 月第 1 版　2004 年 1 月第 1 次印刷

开　　本　787 毫米×1092 毫米　1/16　印张 13.875

字　　数　323 千字

印　　数　1～4 000 册

定　　价　21.00 元

ISBN 7-5606-1301-2/TN·0242

XDUP　1572001-1

序

西安电子科技大学出版社一直把视角的焦点放在电子信息领域的最新发展和对于生产的应用方面。针对当前新经济时代，信息化水平已成为衡量我国现代化程度和综合国力的主要标志，现在出版“现代电子信息工程理论与技术丛书”，显然是一个十分恰当的时机。这套丛书的主要对象是从事电子信息领域研究和开发的科技工作者、工程师、在读的研究生，以及希望了解该领域发展的各类相关人员。因此本套丛书的重点不在于艰深的理论探讨，而是力求理论联系实际，揭示新应用，发展新领域；总之，我们希望通过这套丛书能帮助读者对电子信息领域的总体、全貌和发展趋势有所了解。

西安电子科技大学出版社一直以电子信息领域的热心读者作为自己的服务对象。这套丛书的好与坏，起的作用大与小都要靠每一位读者来检验。因此在成立编委会和着手编辑这套丛书的时候，我们对读者的对象、读者的需求和读者的兴趣做了多方面的设想。为了使多方面的读者都有所收获，我们力求把每本书每个章节都做到简单明了、深入浅出；每本书都是读者了解电子信息领域的忠实“导游”；每本书都是作者与读者交换思想和促膝谈心的最佳机会。

西安电子科技大学出版社一直有着广泛且相对联系紧密的作者群，他们大多是熟悉电子信息领域发展的一线专家，其中不乏是该领域的知名学者、教授，正是由于这么一个群体，使我们有信心把这套丛书的学术水平和实用价值提到一个新的水平。

尽管如此，这套丛书的编撰还是新的尝试，作者和编辑们缺乏经验，加之本领域发展十分迅速，使我们难于全面把握。衷心希望每一位读者都作为这套丛书的实践检验者，你们的每一条意见，将是丛书提高的重要依据。

丛书编委会

现代电子信息工程理论与技术丛书编委会

前　言

聚类分析是多元统计分析的一种，也是非监督模式识别的一个重要分支。它把一个没有类别标记的样本集按某种准则划分成若干个子集(类)，使相似的样本尽可能归为一类，而不相似的样本尽量划分到不同的类中。作为一种无监督分类方法，聚类分析已经被广泛地应用于模式识别、数据挖掘、计算机视觉和模糊控制等许多领域。“人以群分，物以类聚”。聚类是一个古老的问题，它伴随着人类社会的产生和发展而不断深化，人类要认识世界就必须区别不同的事物并认识事物间的相似性，而每个概念的最初形成无不借助于事物的聚类分析。因此，聚类分析的研究不仅具有重要的理论意义，也具有重要的工程应用价值和人文价值。

传统的聚类分析是一种硬划分，它把每个待辨识的对象严格地划分到某类中，具有“非此即彼”的性质，因此这种类别划分的界限是分明的。而实际上大多数对象并没有严格的属性，它们在性态和类属方面存在着中介性，具有“亦此亦彼”的性质，因此适合进行软划分。模糊集理论的提出为这种软划分提供了有力的分析工具，人们开始用模糊的方法来处理聚类问题，并称之为模糊聚类分析。由于模糊聚类得到了样本属于各个类别的不确定性程度，表达了样本类属的中介性，即建立起了样本对于类别的不确定性描述，更能客观地反映现实世界，从而成为聚类分析研究的主流。

我对模糊理论与模糊聚类的研究始于 1994 年。1994 年 9 月，我被西安电子科技大学免试推荐攻读信号与信息处理专业的硕士学位，师从模糊信息处理专家谢维信教授，作为第二完成人完成了国家自然科学基金项目“模糊聚类新方法研究”(批准号：69472046)，并于 1996 年上半年提前完成硕士学位论文《基于进化计算和神经网络的模糊聚类新算法研究》，研究了模糊逻辑、神经网络和进化计算的集成及其在聚类分析问题中的应用。通过硕士阶段的研究工作，我对模糊聚类分析有了更深刻的认识，产生了深厚的感情。因此，我决定继续留在谢维信教授的课题组研究模糊聚类分析，并攻读博士学位。1997 年，我受学校派遣赴日本静冈大学情报科学部计算机科学系计算机博弈研究所交流学习，在饭田弘之(Hiroyuki Iida)教授的指导下从事模糊聚类与人工智能相结合的研究。1998 年，我回国后完成了国家自然科学基金项目的技术报告《模糊聚类的新方法研究》，同时在导师谢维信教授的悉心指导下，于 1999 年 5 月完成博士学位论文《模糊聚类算法的优化及应用研究》，并顺利通过论文答辩。此论文经著名的多值逻辑专家南京航空航天大学的朱梧贾教授，模糊逻辑专家陕西师范大学的王国俊教授，模糊逻辑应用方面的专家国防科技大学的庄钊文教授、郁文贤教授，华南理工大学的郑启伦教授以及神经网络专家西安电子科技大学的焦李成教授等审阅，并获得了高度评价。2000 年，我到香港中文大学讯息工程系多媒体研究室做副研究员，在汤晓鸥教授的指导下从事基于内容的视频信息检索方面的研究，研究发现模糊聚类在视频信息检索方面具有重要的应用而且能够取得相当好的效果，从而更加坚定了我对模糊聚类分析更深层次的理论研究和应用研究。2001 年，我回到学校后，积极查阅资料，准备申报模糊聚类在视频检索方面的课题，并于 2002 年非常幸运地得到国家自然科

学基金的资助，使我继续从事我喜爱的研究课题。在查阅资料的过程中，我发现国内缺乏一本系统的专门研究有关模糊聚类分析的书籍，而实际应用中又往往需要这样一本教材或专著。作为该领域的科研工作者我感到自己责无旁贷，尽管自己对模糊聚类的研究还很肤浅，但是我愿意抛砖引玉，让更多的学者出版这方面的专著，以飨读者。

从 1996 年到 1998 年我协助谢维信教授主办了两届“多值逻辑与模糊逻辑”国内学术会议，认识了很多模糊逻辑界的专家学者，他们都勉励我把模糊聚类分析方面的研究成果整理出版，我受到莫大的鼓舞。2003 年，在北京清华大学举办的“模糊信息处理理论与应用”国际会议上，我遇到模糊数学的创始人 Zadeh 教授，当谈及模糊聚类分析时，他认为模糊聚类分析是模糊数学的成功应用的范例之一，具有非常广阔的研究前景，同时他也勉励我继续开展该课题的研究。在众人的鼓励下，我开始着手整理近年来的初步研究成果，其中包括我本人的 50 余篇有关模糊聚类的论文和课题的部分成果，也参考和吸收了我的师兄范九伦教授、钱云涛教授和裴继红教授的部分成果，在此向他们表示深深的谢意。没有谢维信教授的指导，我不可能接触模糊聚类这一具有挑战性的课题，也就不会有本书的出版，在此向谢维信教授表示崇高的敬意。

我特别感谢西安电子科技大学出版社的臧延新同志，本书能够得以顺利出版发行，与他的耐心指导和辛勤劳动是分不开的，同时也非常感谢西安电子科技大学姬红兵教授、杨兵副教授，我们在同一课题组工作，他们给我提供了大量的方便和帮助，焦李成院长、刘应南书记对本书的出版给予了大力关照，在此一并表示感谢。在写作过程中，参考了大量的文献，作者尽可能一一注明，但由于文献较多，疏漏在所难免，在此向被遗漏的作者表示歉意，并向所有的参考文献作者表示衷心的感谢！

感谢自然科学基金的资助，正是自然科学基金的资助才使我努力钻研，深入探索，在模糊聚类分析研究方面取得了一定进展，先后在重要的期刊上发表学术论文 50 余篇。本书充分地吸收了这些论文的精华，并在此基础上进行了增删和补改。现在呈献给广大读者的是我从事模糊聚类分析研究的一点体会和心得，但愿对您有所帮助。说句实在话，我非数学专业毕业，对模糊数学和模糊聚类的认识还很肤浅，真心地希望您特别是有关专家提出宝贵意见，以便再版时改正。

谨以此书献给我的父母、岳母和我挚爱的妻子、儿子！

高新波

2003 年 11 月

目　录

第1章 绪 论

1.1 模糊数学的产生和发展

我们知道，数学是从量的侧面研究客观世界的一门科学，因此，一提起数学，人们自然会想到它是精确的。然而精确的数学有时不能有效地描述现实世界中存在的大量模糊现象，例如，“好与坏”、“长与短”、“一大堆”、“一小撮”、“太热”、“有点冷”、“比较甜”、“不太苦”、“物美价廉”、“地大物博”等等。但是这些“量”在人们的头脑里的确有个“标准”，而且为人们所普遍接受。利用这些模糊量非但不会影响人们之间的信息交流，反倒更便于理解与记忆。

精确数学是建立在经典集合论基础之上的。根据集合论的要求，一个对象对于一个给定的集合，要么属于($\in$)，要么不属于($\notin$)，两者必居其一，绝不允许模棱两可。由此而产生了我们熟知的二值逻辑，即对于一个“命题”，或者是真(真值为 1)，或者是假(真值为 0)，两者必居其一。19 世纪由于英国数学家布尔(Bool：1815－1864)等人的研究，这种基于二值逻辑的绝对思维方法经过抽象后成为布尔代数，也叫逻辑代数，它用代数方法研究推理、证明等逻辑问题，它的出现促使数理逻辑成为一门很有实用价值的特殊学科，同时也成为计算机科学的基础。尽管如此，二值逻辑却无法解决一些逻辑悖论或诡辩问题，例如，著名的罗素(Russell)“理发师悖论”问题，“秃头悖论”问题和“克利特岛人(Cretan)说谎悖论”问题等等。

日常生活中的“模糊性”现象的存在、逻辑悖论的发现以及海森堡(Heisenberg)测不准原理的提出导致了多值逻辑或“模糊逻辑”在 20 世纪二三十年代的诞生。量子理论学家在二值逻辑的框架中引入第三值或中间真值来表示不确定性，并进一步引入了不确定性程度，把真假看作不确定性的两个极限情况。

20 世纪 30 年代早期，波兰逻辑学家卢卡塞维克兹(Lukasiewicz)首次正式提出了三值逻辑体系，把逻辑真值的值域由{0,1}二值扩展到{0,1/2,1}三值，其中 1/2 表示不确定。后来，他又把真值范围从{0,1/2,1}进一步扩展到[0,1]之间的有理数，并最终扩展为[0,1]区间。逻辑学家们利用常用的真值函数 t:{命题}→[0,1]来定义连续或“模糊”逻辑，并将该体系命名为 L_1[174]。

量子哲学家马克思·布莱克(Max Black)利用连续逻辑为集合中的成员赋值，可以说，他在历史上第一个构造了模糊集的隶属度函数。布莱克称结构的不确定性为“模糊性(Vagueness)”。

1965 年，美国自动控制专家、数学家扎德(L. A. Zadeh)发表了他的论文《模糊集(Fuzzy Sets)》[297]，正式提出了多值集合理论，并把“Fuzzy(模糊)”一词引入技术文献中，从而掀起了多值数学结构研究的第二次浪潮，研究兴趣遍及模糊系统到模糊拓扑的各个方面。此后的二三十年，随着模糊商业产品和新理论、新应用的不断涌现，形成了多值系统研究的第三次浪潮。

扎德的主要贡献在于把模糊性和数学统一在了一起[217]。模糊数学决不是把已经很精确的数学变得模模糊糊，而是用精确的数学方法来处理过去无法用数学描述的模糊事物，因为在现实世界中要想绝对精确是不可能的，实际上也就只能将所谓的不准确程度降低到无关紧要的水平罢了。扎德充分注意到这一点，他的观点不是让数学放弃严格性去迁就模糊性，而是要把数学方法打入具有模糊现象的“禁区”里去，也就是让数学回过头来吸取人脑对于模糊现象识别和判决等处理中的优点，这样就为电子计算机开辟了进一步模拟人脑思维特点的道路，使它变得更加“聪明”。

模糊数学从它诞生的那天起，便和计算机的发展息息相关，相辅相成。没有电子计算机，就没有模糊数学；没有模糊数学，计算机的应用也会大大受到限制。因为利用模糊数学构造数学模型，来编制计算机程序，可以更广泛、更深入地模拟人的思维。而且，模糊数学既认识到事物“非此即彼”的明晰性状态，又认识到事物的“亦此亦彼”的模糊性状态，因此它的适应面也就比传统数学广泛得多。迄今为止，模糊数学已在模式识别、自动控制、信息处理、天气预报、地震研究、人工智能、医疗诊断、农作物选种以及心理学、生态学、语言学等许多领域内得到应用。

当前，模糊数学的研究领域可大体分为三个方面[278]：模糊数学理论及其与经典数学、统计数学的关系，模糊语言和模糊逻辑，模糊数学的应用等。尽管模糊数学诞生很晚，但其发展十分迅速。1978 年，Zadeh 教授提出了可能性理论，阐述了随机性和可能性的区别。这被认为是模糊数学发展的第二个里程碑。可能性理论的出现为模糊数学更广泛地应用于模式识别和其他领域提供了强有力的理论基础和有效的工具。

目前，尽管模糊数学已在自然科学及社会科学领域内获得了较为广泛的应用，但它的理论体系和实际推广应用仍处于发展之中。这就需要我们从理论和实践两个方面进一步深入地研究它、发展它和完善它。

1.2 信息科学与模式识别

随着现代科学特别是计算机科学的发展，社会科学与自然科学之间正在相互渗透，形成许多新的边缘学科。其中最具生命力的，莫如信息科学，因为它用量化公式把人的思维过程表现了出来。这样，配合以现代电子计算机的巨大信息存储能力，便可以解决许多人的才智所不能解决的复杂问题。有人说，当今是一个脑力延伸的时代，确实一点也不过分。

信息科学的最新研究发展表明，建立在概率论基础上的香农(Shannon)信息论，只着重表达了信息的传递，但难以表达信息本身的含义。而信息科学不仅要研究信息“量”的问题，更重要的还在于信息的结构，即信息的定性描述问题。这就涉及到信息的提取、描述、

分析、推理、判断和决策等富有挑战性的处理工作。在信息处理这一领域中，模式识别起着举足轻重的作用，具有信息感知与理解等处理功能，是研究信息结构与含义的重要分析工具[74]。

模式识别(Pattern Recognition)是 20 世纪 60 年代初迅速发展起来的、与高技术的研究开发有着密切联系的一门新兴学科。它所研究的理论和方法在很多科学和技术领域中得到了广泛的应用，推动了人工智能系统的发展，扩大了计算机应用的领域，在向人类智能逼近这一永恒的前沿课题中占有一席之地。可以说，在高度自动化的今天，模式识别几乎已经进入人类生活的各个领域[28]。

“模式”(Pattern)这个词与保护神(Patron)来自同一个词根，本意是指供模仿用的理想标本[60]。因此，形象地讲，模式识别是指从待识别的对象中分辨出哪个对象与标本相同或相似。人脑就是一个可靠的识别系统，人们在感受外界现象的时候，总要把它们进行分类，即把相似而又不完全相同的现象分成一组。这时，在同一组中不同的物体和现象之间总有某些方面是相似的。人们只要熟悉现象中为数不多的代表，就能从现象形成组的概念，正是人脑的这种能力才构成模式的概念。显然，分类(Classification)是建立和识别模型的重要基础和手段，因此模式识别与分类是密切相关的。此外，任何一门学科都要通过分类来建立自己的概念，也要通过分类来发现和总结规律。这样，作为一种强有力的工具，分类的研究具有十分重要的意义。

1.3　模式识别与模糊聚类

模式识别又常称作模式分类。从处理问题的性质和解决问题的方法等角度，模式识别或者模式分类可分为有监督的分类(Supervised Classification)和无监督的分类(Unsupervised Classification)两种类型。

所谓有监督的分类，又称为有教师的分类或有指导的分类。在这类问题中，已知模式的类别和某些样本的类别属性，首先用具有类别标记的样本对分类系统进行学习或训练，使该分类系统能够对这些已知样本进行正确分类，然后用学习好的分类系统对未知的样本进行分类。这就要求我们对分类的问题要有足够的先验知识，而要做到这一点，往往要付出相当大的代价。

在没有先验知识的情况下，则需要借助无监督的分类技术。无监督的分类又称为聚类分析(Cluster Analysis)，这是本书将要研究的主要内容。从学科的谱系图上看，聚类分析属于信息科学这棵大树上模式识别分支中的一片树叶。希望本书的研究能为丰富和发展信息科学这一前沿学科起到一定的推动作用，能为完善和提高模式识别这一实用性极强的自动化技术起到积极的促进作用。在展开讨论之前，首先让我们对聚类问题作一简要介绍。

聚类就是按照一定的要求和规律对事物进行区分和分类的过程，在这一过程中没有任何关于分类的先验知识，没有教师指导，仅靠事物间的相似性作为类属划分的准则，因此属于无监督分类的范畴。聚类分析则是指用数学的方法研究和处理给定对象的分类。

“人以群分，物以类聚”。聚类是一个古老的问题，它伴随着人类社会的产生和发展而不断深化，人类要认识世界就必须区别不同的事物并认识事物间的相似性[142]。聚类分析

是多元统计分析的一种，也是非监督模式识别的一个重要分支。它把一个没有类别标记的样本集按某种准则划分成若干个子集(类)，使相似的样本尽可能归为一类，而不相似的样本尽量划分到不同的类中。

传统的聚类分析是一种硬划分(Crisp Partition)，它把每个待辨识的对象严格地划分到某类中，具有"非此即彼"的性质，因此这种类别划分的界限是分明的。而实际上大多数对象并没有严格的属性，它们在性态和类属方面存在着中介性，具有"亦此亦彼"的性质，因此适合进行软划分。模糊集理论的提出为这种软划分提供了有力的分析工具，人们开始用模糊的方法来处理聚类问题，并称之为模糊聚类分析。由于模糊聚类得到了样本属于各个类别的不确定性程度，表达了样本类属的中介性，即建立起了样本对于类别的不确定性描述，更能客观地反映现实世界，从而成为聚类分析研究的主流[313]。有关模糊聚类的研究现状将在后续章节中详细介绍。

1.4 模糊聚类研究的意义

虽说聚类分析应用于模式识别的时间不长，但它并非一个新领域，早已被应用在其他学科中。Dubes 和 Jain 关于聚类分析的综述包括了从 77 份杂志和 40 本书中摘取出来的 250 条引文[48]，如此巨大的文献量说明了聚类分析的重要性和交叉学科性，也足以说明它的发展及应用前景的广阔性。

同时，国际和国内的学者都对聚类分析的研究非常重视，IEEE 的汇刊中《模式分析与机器智能》(PAMI，Pattern Analysis and Machine Intelligence)、《系统、人和控制》(SMC，Systems，Man，and Cybernetics)、《模糊系统》(FS，Fuzzy Systems)、《神经网络》(NN，Neural Networks)、《信号处理》(SP，Signal Processing)等杂志中几乎每期都有讨论聚类分析问题的文章。从 1992 年开始的由 IEEE 和神经网络理事会共同主办的 FUZZ-IEEE 会议，每两年召开一次，每次至少有 3 到 4 个专题讨论聚类和模糊聚类分析的最新研究进展和发展现状。另外，我国作为模糊数学研究的大国，不仅在基础理论研究上取得了丰硕的成果，而且在模糊聚类等的应用研究上亦令世人瞩目，比如基于模糊聚类的天气预报、矿藏识别和医学诊断等等。为了积极引导(模糊)聚类分析的理论和应用的研究热潮，国家自然科学基金委员会还专门对"模糊聚类的新方法研究"(批准号：69472046)和"无监督新闻视频语义分割和自动标注算法研究"(批准号：60202004)立了项，重点资助我们的研究。在这样的背景下，研究(模糊)聚类分析的意义也就不言而喻了。

1.5 模糊聚类的应用

模糊聚类理论的发展推动了它在生产实践中的应用，反过来实际应用的需求又促进了模糊聚类理论的不断丰富和完善。随着理论的发展，模糊聚类已经在诸多领域获得了广泛的应用，并取得了满意的效果和可观的效益。其应用范围涉及到通信系统中的信道均衡[252]、矢量量化编码中的码书设计[163,226,290]、时间序列的预测[215,243]、神经网络的训

练[165,288]、非线性系统辨识[241]、参数估计[123]、医学诊断[19]、天气预报[220]、食品分类[277]、水质分析[218]等众多领域。在此，我们只简要介绍模糊聚类在模式识别和图像处理中的应用情况。

1.5.1 模糊聚类在模式识别中的应用

模式识别中两大主要的分支为有监督的分类和无监督的分类，而其中无监督的分类与聚类分析相对应。正是由于模糊聚类与模式识别的天然联系，使得它首先在模式识别领域获得了成功的应用。

模式识别中一个最重要的问题是特征提取，模糊聚类不但能从原始数据中直接提取特征[237]，还能对已经得到的特征进行优选和降维操作[18]，以免造成“维数灾难”；在提取完特征后就需要设计分类器，模糊聚类算法既可以提供最近邻原型分类器[29,166]，还可以用来进行特征空间划分和模糊规则提取[6]，以构造基于模糊 IF - THEN 规则的分类器[58,107,265]；在物体识别或线条检测中，模糊聚类既可以直接作用于原始数据上[51,53,54,182,183]，也可以用于变换域中，比如 Hough 变换一直受峰值检测的困扰，Jolion[154]提出基于模糊聚类的检测方法，解决了这一难题，使得 Hough 变换可以自动执行，使用起来方便快捷。

在一些具体的识别应用中，模糊聚类也取得了较好的效果，比如汉字字符识别中的字符预分类[169,280]，语音识别中的分类和匹配[144,311]，雷达目标识别中目标库的建立和新到目标的归类[257,284]等等，在此不再一一列举。

1.5.2 模糊聚类在图像处理中的应用

图像处理是计算机视觉的重要组成部分。由于人眼视觉的主观性使图像适合用模糊手段处理，训练样本图像的匮乏又需要无监督分析，而模糊聚类正好满足这两方面的要求，因此成为图像处理中一个强大的分析工具。

模糊聚类在图像处理中最为广泛的应用为图像分割，由于分割问题可以等效为像素的无监督分类，因此早在 1979 年 Coleman 和 Andrews 就提出用聚类算法做图像分割[49]，此后又涌现出如基于二维直方图[197,200]、塔型结构[193,231]、小波分析[234]、分形分维[38]、空间约束[266]、可能性理论[185]和有效性指导[15]等一系列的灰度图像的聚类分割方法。在纹理图像[234]、彩色图像[198,209]、序列图像[233]以及航空遥感图像[42,269]等分割方面也获得了很大的进展。

另外，基于模糊聚类的方法在边缘检测[54,73,177]、图像增强[253]、图像压缩[189]、图像平滑[60]、图像匹配[258]等众多方面也同样取得了丰硕的成果。

随着在图像处理中的应用的发展，对模糊聚类理论又提出了许多新的要求，因此必须进一步丰富和完善聚类理论，指导实际应用，使(模糊)聚类更好地服务于人类。

第 2 章 模糊理论基础

模糊理论(Fuzzy Theory)是建立在模糊集合基础之上的，是描述和处理人类语言中所特有的模糊信息的理论。它的主要概念包括模糊集合(Fuzzy Sets)、隶属度函数(Membership Function)、模糊算子(Fuzzy Operator)、模糊运算(Fuzzy Operation)和模糊关系(Fuzzy Relation)等。本章将分别介绍这些概念。

2.1 普通集合

我们把被讨论的全体对象或范围叫做论域(Domain)，常用 $U,V,E,\cdots,X,Y,\cdots$ 大写字母表示。把论域中的每个对象称为元素(Element)，用相应的小写字母 $u,v,e,\cdots,x,y,\cdots$ 表示。

定义 2.1-1 给定论域 X 和某一性质或属性 P，X 中满足性质 P 的所有元素所组成的全体叫做集合(Set)，简称集。

其实，这里集合不是定义的概念，而是一种用“数学语言”进行的数学刻画。通常，我们习惯用大写字母 $A,B,\cdots$ 来表示集合。从 X 中任意取出一个元素 x，对于给定的集合 A，要么有 x 属于 A，记做 $x\in A$，要么有 x 不属于 A，记做 $x\notin A$，二者必居其一且仅居其一，这就是普通集合论中最基本的要求。

2.1.1 集合的表示方法

如果一个集合所包含的元素为有限个，则称之为有限集，否则就叫做无限集。常用的集合表示方法有如下三种形式：

1. 列举法(枚举法)

对于有限集，可以将所有的元素一一列出，并用大括号括起来表示，如

$$A=\{a_1,a_2,\cdots,a_n\}$$

2. 描述法(定义法)

对于无限集，由于元素数目无限，可通过元素的定义来描述集合，如

$$A=\{x|P(x)\}$$

其中，$P(x)$ 是指“x 具有性质 P”。