

AI

INTELLIGENT INFORMATION FUSION AND
TARGET RECOGNITION METHOD

智能信息融合 与目标识别方法

胡玉兰 郝博 王东明◎等著



机械工业出版社
CHINA MACHINE PRESS

智能信息融合与 目标识别方法

胡玉兰 郝 博 王东明 郝伟光 著
秦丽娟 胡南南 李 芳



机械工业出版社

本书研究了信息融合目标识别技术,首先分析了特征级融合目标识别的基本理论,然后研究了多源图像的预处理、结合阈值分割的分水岭算法、结合聚类分割的分水岭算法、目标特征提取方法。对于特征融合方法,研究了基于协方差矩阵多特征信息融合、基于主成分分析的特征融合方法、基于改进免疫遗传的特征融合方法、基于独立分量的特征融合、对典型相关分析特征融合方法的改进。最后介绍了基于优化改进的反向传播神经网络目标识别、模糊支持向量机理论与编程实现、基于模糊支持向量机的识别系统实现。

本书适合从事多源信息融合理论、目标识别技术研究和工程应用的技术人员阅读,也可作为高等院校相关专业高年级本科生和研究生的参考教材。

图书在版编目 (CIP) 数据

智能信息融合与目标识别方法/胡玉兰等著. —北京:
机械工业出版社, 2018. 4
ISBN 978-7-111-59317-1

I. ①智… II. ①胡… III. ①信息融合—图象识别
IV. ①G202②TP391. 41

中国版本图书馆 CIP 数据核字 (2018) 第 041505 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑: 任 鑫 责任编辑: 阎洪庆

责任校对: 肖 琳 封面设计: 马精明

责任印制: 张 博

三河市国英印务有限公司印刷

2018 年 5 月第 1 版第 1 次印刷

169mm × 239mm · 13 印张 · 246 千字

0001—2600 册

标准书号: ISBN 978-7-111-59317-1

定价: 59.00 元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

服务咨询热线: 010-88361066

机工官网: www.cmpbook.com

读者购书热线: 010-68326294

机工官博: weibo.com/cmp1952

010-88379203

金书网: www.golden-book.com

封面无防伪标均为盗版

教育服务网: www.cmpedu.com

前 言 >>>

随着科学技术的发展,信息化已经在各个领域得到了重视和发展,特别是在军事方面的应用,如何获得有效信息是现代化战争的一个关键点。各个国家都投入了很大的人力和物力,用以实现对目标的全天候、全方位、多角度的有效识别与追踪。单一传感器的工作环境正变得日益复杂,如遇到移动、遮挡、缩放和旋转的目标时,需要通过利用多传感器来同时获取目标多方位特征,对其进行有效的融合分析,从而能够实现对目标的识别。目前,多传感器信息融合大多被应用在模式识别、目标跟踪等方面,同时在交通、网络安全、工业、军事和生物医学等领域也有着非常好的应用前景。

在传统的目标识别系统中,由于技术局限、环境复杂度不高和工作要求低,一般都以单一传感器来获取数据,所以获得目标或场景的信息一般都是片面的,并且受一定的光线和环境影响导致了精确度低。为此人们提出了基于多传感器多特征的目标识别方法。

本书研究了信息融合目标识别技术,首先分析了特征级融合目标识别的基本理论,然后研究了多源图像的预处理、结合阈值分割的分水岭算法、结合聚类分割的分水岭算法、目标特征提取方法。对于特征融合方法。研究了基于协方差矩阵多特征信息融合、基于主成分分析的特征融合方法、基于改进免疫遗传的特征融合方法、基于独立分量的特征融合、对典型相关分析特征融合方法的改进。最后介绍了基于优化改进的反向传播神经网络目标识别、模糊支持向量机理论与编程实现、基于模糊支持向量机的识别系统实现。

本书由胡玉兰、郝博、王东明、郝伟光、秦丽娟、胡南南、李芳著。本书的研究工作得到了国家自然科学基金面上项目(项目号:61373089)的资助,作者在此表示诚挚的感谢。同时感谢机械工业出版社参加本书出版工作的同志们的大力支持与帮助。

信息融合目标识别技术目前仍处于不断发展的阶段,许多理论和算法还不够完善,加之作者水平所限,书中难免有疏漏之处,恳请广大读者批评指正。

作者

目 录 >>>

前言

第1章 绪论	1
1.1 概述	1
1.2 研究现状	3
第2章 特征级融合目标识别的基本理论	5
2.1 信息融合目标识别结构	5
2.2 特征级融合目标识别系统基本结构	6
2.3 特征级融合目标识别的关键问题分析	8
2.3.1 特征数据库	8
2.3.2 特征提取	8
2.3.3 特征融合	8
2.3.4 目标识别	10
2.4 本章小结	11
第3章 多源图像的预处理	13
3.1 多源图像的去噪	13
3.1.1 噪声分析	13
3.1.2 常用的图像去噪方法	14
3.1.3 几种较新的去噪方法	15
3.2 多源图像的平滑	20
3.2.1 均值平滑法	20
3.2.2 邻域平均法	20
3.2.3 加权平均法	21
3.2.4 选择式掩模平滑法	21
3.3 多源图像的滤波	23
3.3.1 低通滤波	23
3.3.2 高通滤波	24
3.3.3 中值滤波	26

3.3.4	维纳滤波	27
3.4	本章小结	29
第4章	结合阈值分割的分水岭算法	30
4.1	引言	30
4.2	图像阈值分割概述	30
4.3	常用的阈值分割方法	31
4.3.1	全局阈值法	31
4.3.2	局部阈值法	36
4.3.3	动态阈值法	36
4.4	改进的最大类间方差法	36
4.5	本章小结	43
第5章	结合聚类分割的分水岭算法	44
5.1	图像聚类分割概述	44
5.2	常用的图像聚类分割算法	46
5.2.1	划分聚类算法	46
5.2.2	层次聚类算法	47
5.2.3	基于密度的聚类算法	47
5.2.4	基于模型的聚类算法	48
5.2.5	基于网格的聚类算法	48
5.3	改进的FCM聚类分割算法	49
5.4	本章小结	53
第6章	目标特征提取方法	54
6.1	传统的特征提取方法	57
6.1.1	经典的特征提取方法	57
6.1.2	多项式不变矩	59
6.1.3	共生矩阵	67
6.2	红外和可见光图像特征提取和融合	70
6.2.1	多传感器特征提取	70
6.2.2	目标区域分割和检测	71
6.2.3	特征提取与融合	73
6.3	本章小结	75
第7章	基于协方差矩阵多特征信息融合	77
7.1	图像特征提取	77
7.1.1	多传感器特征提取	77

7.1.2	灰度特征	78
7.1.3	空间信息特征	78
7.1.4	梯度信息特征	79
7.2	协方差矩阵的构造	80
7.2.1	协方差矩阵与协方差	80
7.2.2	协方差矩阵距离的度量	81
7.2.3	特征协方差	82
7.3	多特征信息融合	83
7.4	融合实验结果与分析	83
7.5	本章小结	85
第8章	基于主成分分析的特征融合方法	87
8.1	特征提取	87
8.1.1	灰度共生矩阵	87
8.1.2	Hu 不变矩	88
8.2	几何特征	90
8.3	主成分分析的定义	90
8.3.1	主成分分析的基本原理	90
8.3.2	数学模型	91
8.3.3	主成分的推导	93
8.4	基于主成分分析的图像特征级融合实现	96
8.4.1	目标特征的提取	96
8.4.2	特征融合	96
8.4.3	实验结果与分析	96
8.5	本章小结	99
第9章	基于改进免疫遗传的特征融合方法	100
9.1	遗传算法基础理论	100
9.1.1	遗传算法概述	100
9.1.2	遗传算法流程	101
9.2	一般的免疫算法基础理论	103
9.2.1	免疫算法概述	104
9.2.2	免疫算法流程	105
9.2.3	一般的免疫遗传算法	105
9.3	基于改进免疫遗传的图像特征级融合实现	107
9.3.1	改进算法的关键技术	107
9.3.2	基于改进免疫遗传的特征融合原理	108

9.3.3	特征级融合实现	108
9.3.4	实验结果与分析	111
9.4	本章小结	112
第 10 章	基于独立分量的特征融合	113
10.1	ICA 的定义	113
10.2	随机变量的独立性概念	114
10.3	ICA 独立性的度量	115
10.3.1	非高斯性	115
10.3.2	互信息	117
10.4	快速固定点 ICA 算法	117
10.5	基于 ICA 的图像特征级融合实现	119
10.5.1	ICA 的预处理	119
10.5.2	特征融合	120
10.6	实验结果与分析	121
10.7	本章小结	125
第 11 章	对典型相关分析特征融合方法的改进	126
11.1	CCA 的基本思想	126
11.2	CCA 的基本原理	127
11.2.1	CCA 的数学描述	127
11.2.2	典型相关的定义及导出	128
11.2.3	CCA 的基本性质	130
11.3	典型相关变量和相关系数的求解步骤	132
11.4	基于 CCA 改进算法的图像特征级融合实现	133
11.4.1	改进算法的关键技术	133
11.4.2	特征融合过程	133
11.5	实验结果与分析	134
11.6	本章小结	137
第 12 章	基于优化改进的反向传播神经网络目标识别	138
12.1	BP 神经网络	139
12.1.1	BP 网络模型	139
12.1.2	BP 神经网络结构	139
12.1.3	BP 神经网络训练学习	142
12.1.4	BP 神经网络主要特点	144
12.2	改进的 BP 神经网络	145

12.3	PSO 算法的基本原理和理论基础	145
12.3.1	PSO 算法的基本原理	145
12.3.2	PSO 算法的基本流程	147
12.4	PSO 优化改进的 BP 神经网络	147
12.4.1	PSO 优化问题分析	147
12.4.2	PSO 优化算法流程	149
12.5	仿真结果分析和识别系统实现	149
12.5.1	仿真结果分析	149
12.5.2	识别系统实现	152
12.6	本章小结	155
第 13 章	模糊支持向量机理论与编程实现	157
13.1	模糊数学理论	157
13.1.1	模糊数学的发展历程	157
13.1.2	模糊集	158
13.1.3	几种常见的模糊隶属度函数	158
13.2	支持向量机理论	160
13.2.1	经验风险	160
13.2.2	结构风险最小化	162
13.2.3	最优超平面	164
13.2.4	VC 维数	165
13.2.5	核函数	166
13.3	模糊支持向量机理论	167
13.3.1	第一种模糊支持向量机	167
13.3.2	第二种模糊支持向量机	168
13.4	模糊隶属度核函数的选择	170
13.4.1	基于类中心距离的隶属度函数	170
13.4.2	基于 S 形函数的隶属度函数	171
13.4.3	基于 π 形函数的隶属度函数	171
13.5	模糊支持向量机算法的编程实现	172
13.6	本章小结	174
第 14 章	基于模糊支持向量机的识别系统实现	175
14.1	模糊支持向量机识别系统的架构	175
14.2	图像预处理部分的实现	176
14.2.1	文件的打开与显示	177
14.2.2	预处理模块图像增强的执行	178

14. 2. 3	预处理模块图像分割的执行	178
14. 3	特征提取部分的实现	179
14. 3. 1	文件的打开与显示	179
14. 3. 2	传感器类型的选择	179
14. 3. 3	计算相应的特征值	180
14. 3. 4	特征值的存储	181
14. 4	目标特征数据库系统的技术实现	181
14. 4. 1	特征数据库管理系统的构建	181
14. 4. 2	目标动态特征库管理技术	182
14. 4. 3	目标特征数据库建立	184
14. 4. 4	数据库管理系统编程实现	185
14. 5	模型训练部分	188
14. 6	目标识别部分	189
14. 6. 1	目标识别模块工作流程与结果显示	189
14. 6. 2	目标识别结果分析与对比	189
14. 7	本章小结	192
参考文献		193

1.1 概述

随着科学技术的发展，信息化已经在各个领域里得到了重视和发展，特别是在军事方面的应用，如何获得有效信息是现代化战争的一个关键点，它决定了战争的胜负，从而使得各个国家都对实时获取信息投入了很大的人力和物力，以达到对目标实时、全天候、全方位、多角度有效而准确的识别与追踪。而目标识别的主要问题是获得我们想要识别的目标的特征，使用单一传感器一般在特定环境下可对目标进行识别，但是随着现在的工作环境变得日益复杂，在遇到可移动、遮挡、缩放和可旋转姿态变化的目标时，它已经不能满足需求了。此时我们就需要通过利用多传感器来同时获取目标多方位特征，对其进行有效的融合分析，才能对目标进行识别。目前，多传感器信息融合大多被应用在模式识别、目标跟踪、态势评估等方面，同时在交通、网络安全、工业、军事和生物医学等领域也有着非常好的应用前景。

在传统的目标识别系统中，由于技术局限、环境复杂度不高和工作要求低，一般都以单一传感器来获取数据，所以获得目标或场景的信息一般都是片面的，并且在受一定的光线和环境影响导致了精确度低，此时人们为此提出了基于多传感器多特征的目标识别方法。在 20 世纪 70 年代后期人们提出了基于多传感器的信息融合概念，通过将传感器、图像信号处理、计算机仿真和人工智能有效地结合起来形成现代处理技术。它是通过对人类和其他受感知的器官（四肢、眼、鼻、耳）所感知的信息（触觉、图像、味觉、声音）来进行有效的模拟，再利用先验知识对其进行综合分析，从而达到了对周围陌生环境做出有效可靠的分析判断的目的。当然这里的处理运行过程相当的复杂，同时也是根据自适应去有效完成的，通过将各种有效信息进行转化，来达到对有价值的具体环境的解释，这又涉及许多不同的智能结合应用，以及通过利用先验知识库来对信息含义进行有效解释。通过将不同的或者相同的多个传感器进行有效结合的方法称为信息融合处理方法，同时人们也称它为多传感器融合，而信息融合则是我们应用中最广泛的称呼。

在目标识别系统中，可见光传感器和红外传感器是常用的两种传感器。主要负责在白天与黑夜对目标图像数据进行有效的采集，充分结合了它们在图像信息的互补性，从而使得它们能够协同进行并得到更高的识别精确度、环境抗干扰性和适应能力。所以相对于单一的传感器目标识别系统而言，多传感器协同的目标识别系统具有以下优点：

- 1) 分辨率高，可信度高，生存能力强，有利于识别目标。
- 2) 范围广，扩展了目标识别有效空间、时间的覆盖范围，对隐身目标能有效地进行识别。
- 3) 可见光和红外传感器能够采集图像的不同特征，从而降低了目标识别的模糊性，对目标能够进行更有效的识别。
- 4) 可实时、全天候、全方位、多角度地对目标进行有效识别，改善了系统的整体性能。

本书研究的目的是针对单传感器在目标识别系统中的不足，研究、设计以及实现采用多传感器协同处理技术的目标识别系统，充分利用红外和可见光采集目标数据信息进行互补，从而获得比任何单传感器对目标更为全面和精确的表述，为多传感器协同的目标识别系统打下了良好的基础。因此，基于多传感器协同的目标识别系统有很高的研究价值。

随着科学技术的发展，虽然基于特征信息融合的目标识别技术被广泛地应用在国防、医学、安防、信息安全、遥感探测等领域，但是在模式识别领域中对研究的难点和一些薄弱环节仍需要我们去完善和解决，所以选择多特征信息融合来作为我们的研究方向，主要有以下几点原因：

- 1) 目前有文献对图像的特征信息融合方法进行了描述，并说明了其在目标识别中的表现明显好于像素级图像融合和决策级图像融合。
- 2) 通过图像的特征信息融合不仅能够帮助我们提取更多有效的特征信息，来对图像进行综合分析和融合处理，而且还可获得图像中附带可利用的其他特征，在一定程度上帮助系统降低了训练学习的复杂度，从而增强了算法的鲁棒性。
- 3) 特征信息图像融合通过外部传感器获得的图像数据既能对目标的多特征信息进行融合和保留，也能帮助我们在一定程度上消除原始图像的一些冗余信息。这样就能达到信息压缩的目的，有利于对信息进行有效的实时处理。
- 4) 目标识别主要就是依靠目标特征数据库，利用多传感器图像特征信息融合就是为了从不同角度、不同时间、不同空间去对目标信息进行有效采集，通过有效的特征提取，以及特征信息融合，最终的结果就是为了更好地去识别目标，提高目标识别系统的可靠性。目前对特征融合的研究最为滞后和困难，由于特征的千差万别，特征提取对算法的要求也是千差万别，所以缺乏一套成熟理论框架和结构。目前所使用的融合思路仍处于沿用模糊方法、聚类方法、人工智能等方法阶段。这里我们面对的问题是怎样才能有效地突破传统的技术并

将其应用于特征融合，来帮助我们获得更好的实质性效果。由于这一领域具有非常大的挑战性和应用前景，专家学者都在努力地深入探索。

1.2 研究现状

在信息融合目标识别方面，国外起步相对国内要早。早在20世纪70年代末期，美国科研机构利用多个悬浮水下声呐，再通过使用当时先进的计算机处理技术对信号进行有效的融合处理后发现，其根据处理后的数据能够对敌方潜艇的位置进行有效的检测。事实证明这一发现可以说是现代战争一次变革的开始，从而使得信息融合技术进入了军事研究部门的视野。

早在1973年，在美国国防部的全额资助下，美国科研机构利用基于多传感器对信息融合技术展开了深入的研究。至此，信息融合技术这一新的名词诞生了，并且在学术界很快得到了广泛的重视并发展起来。在著名的军事C³I系统中，它不仅得到了应用——同时通过采用多种传感器来有效地收集信息，而且在工业控制、智能机器人、海洋纹理监视、空中交通管制和综合导航等领域中，基于多传感器开发也被赋予很大的应用前景。到了20世纪80年代，随着科学技术的快速发展，美国海陆空部队也对军事战略监视系统展开了研发，通过利用信息融合技术来实现对目标进行有效的跟踪、识别、态势评估和威胁估计，并得到了很好的效果。同时基于此项研究成果，成功地开发出了世界上第一代信息融合系统，即集战略系统、海洋监视系统和小型战术系统为一体的系统。美国国防部于1998年将信息融合保密作为一项重点研究开发的关键技术之一。如今英、美、德、法等发达国家中都有许多顶尖的技术人才在政府的资助下对数据融合技术进行深入研究，他们将大量的研究成果公开发表在各种学术期刊和学术会议上。进入21世纪后，在实际应用中信息融合识别技术得到了快速的发展和开发。比如在2000年，加拿大Safe Software公司成功地将自己的产品空间数据操作引擎（FME）打入到中国市场，它是通过基于多年来对数据融合技术研究成果的积累，才成功开发出来的空间数据格式转换的实际应用软件。目前根据用户反馈的数据统计来看，其软件得到了非常好的认可度。美国国防部曾委托雷声公司，利用多传感器信息融合技术为美国陆军研制可有效提供远程先进侦察监视的目标识别系统。这套传感器主要包含了视频传感器、红外传输系统、基于人眼安全的激光测距仪和GPS干涉仪。充分利用了各个传感器的特点，既可以成像，也可通过电视传感器传输实时融合成像，至少帮助图像提高了50%的清晰度，而且关键的是系统的目标探测有效距离相对原来的系统提高了3倍。

国内虽然在多传感器数据融合目标识别这方面相对国外起步比较晚，技术不够成熟，但是在军事应用方面也时刻关注国外的发展，并努力积极进行研究和开发。在非军事领域，我国也投入了充足的科研经费，也得到了很大的进展。如基

于人类生物特征指纹和人脸身份识别系统，它们在考勤、安防等方面取得了很好的开发成果。主要开发成果如下：基于指纹和人脸的融合身份识别、掌纹和视网膜的融合身份识别、可见光和红外融合的人脸身份识别等；基于遥感图像特征融合目标识别，通过使用遥感技术来对目标进行有效探测，并通过利用基于多传感器和多光谱的遥感图像来进行特征融合，从而能够获得目标的位置、性质和身份属性等信息；基于特征级融合的交通智能系统，主要应用在对驾驶员疲劳驾驶警示系统中；基于多传感器特征融合的工业系统，主要应用在工业系统故障诊断中。

特征融合在目标识别中处于一个非常重要的处理环节，但是目前人们对特征融合的研究仍然停留在对样本能提取什么样的特征、如何选择合适的融合规则。目前对于融合算法开发，主要基于应用对象需求来对其展开设计。理论上还没有一套统一体系与评价体系。对于在特征融合目标识别的研究中，广大学者的注意力主要在处理算法、融合识别的结构与框架和对特征融合识别系统如何进行建模与实现等问题上，这几方面中研究热度最高的是对融合识别中使用的算法及信息融合系统在各领域中具体的应用。

目前信息融合技术主要可分为特征级融合、数据级融合和决策级融合这三部分。数据信息融合主要是解决实际中遇到的问题，如传感器中解决时域协调性的问题，在全局坐标中解决相关的一致性检验、坐标的变换、信息的同化、数据的关联等问题，其中最棘手的问题是解决一致性检验和数据的关联。而特征级与决策级融合的识别算法与人工智能紧密结合进行应用是比较常见的。当前比较主流算法主要有卡尔曼滤波法、加权平均法、N-P法、错误率最小化方法、贝叶斯法、证据组合(D-S)法、神经网络法、模糊理论法、聚类分析法等。目前在基于特征级融合识别方法的应用中，大家都广泛采用的算法主要是N-P法、神经网络法、模糊理论法、聚类分析法及各类混合算法。

目前都是基于同一传感器通过对图像进行不同方向、时刻和角度的特征采集来进行有效的融合^{[7][8]}，而由于技术原因，对基于多传感器图像特征信息融合的应用还相对不多。由于信息融合处于目标识别系统的前一级，属于“心脏”地位，其处理过程是非常复杂且具有一定的不确定性，以目前我们所知的理论、技术、设备来对信息融合方法进行研究还是达不到很高的要求。虽然如此，我们的一些成果还是成功地广泛应用在了生活和军事诸多领域中，如医疗领域中的CRT图像检测；生物领域中的指纹识别、人脸识别等；军事战场上的目标识别（如坦克、飞机等）。随着计算机的不断快速发展，处理速度越来越快，融合的计算处理速度效果也越来越快，但是在面对数量非常庞大的目标和特征信息时，由于缺少成熟的特征关联融合理论及其理论工具，在模式识别方法中，模糊方法、聚类方法、神经网络方法、人工智能等方法依旧是研究特征级融合目标识别的主要方法。目前国内外专家学者都在积极探索如何将这些理论和技术应用于特征融合中，从而获得令人满意的效果。

特征级融合目标识别的基本理论

2.1 信息融合目标识别结构

多传感器信息融合目标识别系统一般处理流程有：传感器采集信息预处理、信息配准、信息融合、特征提取以及目标识别。按信息抽象程度，多传感器信息融合一般分为三个层次，即数据级信息融合、特征级信息融合、决策级信息融合。多传感器信息融合目标识别系统结构如图 2-1 所示。

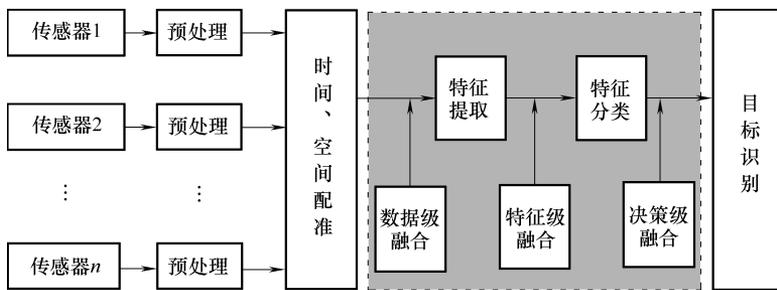


图 2-1 多传感器信息融合目标识别系统结构图

数据级融合是较低层次的融合，它是对来自同等量级的传感器原始数据直接进行综合和分析，在基于传感器的融合数据上进行特征提取和身份估计。它要求传感器必须是同类型或是相同量级，例如若干个可见光传感器或红外传感器和可见光传感器的组合。数据级融合通过对传感器原始数据进行关联和配准，来确定已融合的数据是否与同一目标或实体有关，通过传感器融合数据完成与单传感器相同的识别处理过程。

决策级融合目标识别是指每个传感器对目标的位置、属性、身份等做出初步识别后，再对这些多个识别结果进行融合判断决策，给出最终的目标识别结果。该级融合对通信带宽要求最低，但识别的结果相对来说最不准确。

特征级融合目标识别就是利用各个传感器获取的目标信息，利用特征提取方法提取目标特征数据，运用特征融合算法融合特征数据使原始目标特征信息

变为融合特征数据，再利用目标识别分类算法，进行目标识别分类。通过特征级融合，目标有效信息得到足够保留，且有效信息损失比决策级低很多，但是目标识别能力比具有庞大数据量处理的数据级融合目标识别稍低，但比决策级融合目标识别精确很多，提高了目标识别的实时性、准确性。

2.2 特征级融合目标识别系统基本结构

特征级融合的目标识别是指从多传感器获取的目标原始信息中提取目标特征信息，应用融合算法将其进行融合成复合特征信息量，利用分类识别方法识别目标的过程，如图 2-2 所示。经过预处理（增强、滤波、消噪、分割等）的多传感器系统采集的目标原始信息，经特征提取系统获取目标信息的充分表示量或充分统计量的特征信息，典型的特征信息有边缘、形状、轮廓、角、纹理、相似亮度区域、相似景深区域等，在进行融合目标识别处理时特征信息与多传感器图像融合的应用目的和场合相关，通过特征融合目标识别算法对目标特征信息进行分类识别。通过特征级融合目标识别可以在原始图像中挖掘相关特征信息、增强特征信息的有效性、排除干扰识别的虚假特征、建立融合特征等，从而提高识别率和识别效率。特征级融合的目标识别对传感器配准要求不是十分严格，因此传感器可以分布在不同的平台上。特征级融合的目标识别优点在于可观的信息压缩，可实现实时处理。

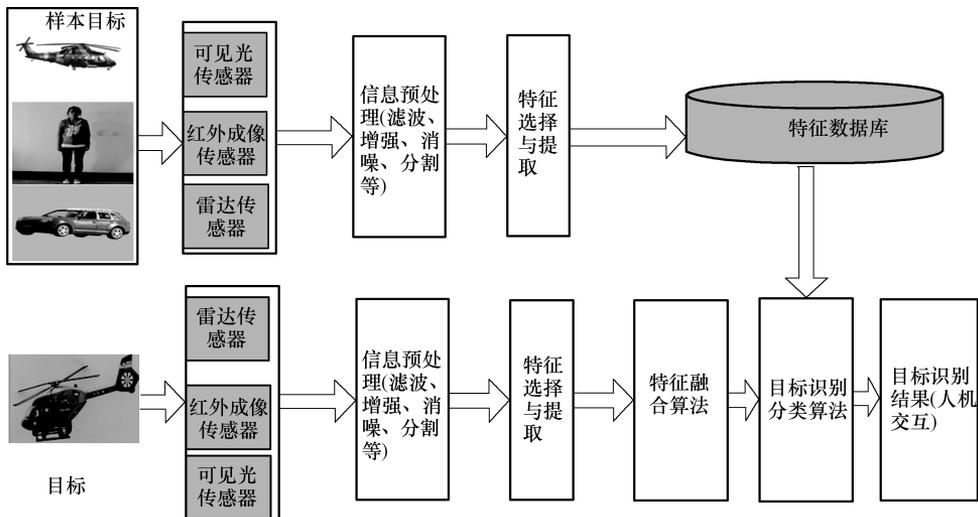


图 2-2 特征级融合的目标识别系统结构图

特征级融合的目标识别原理如图 2-3 所示。

目标特征获取：从物理传感器原始数据中利用特征提取算法提取和选择有

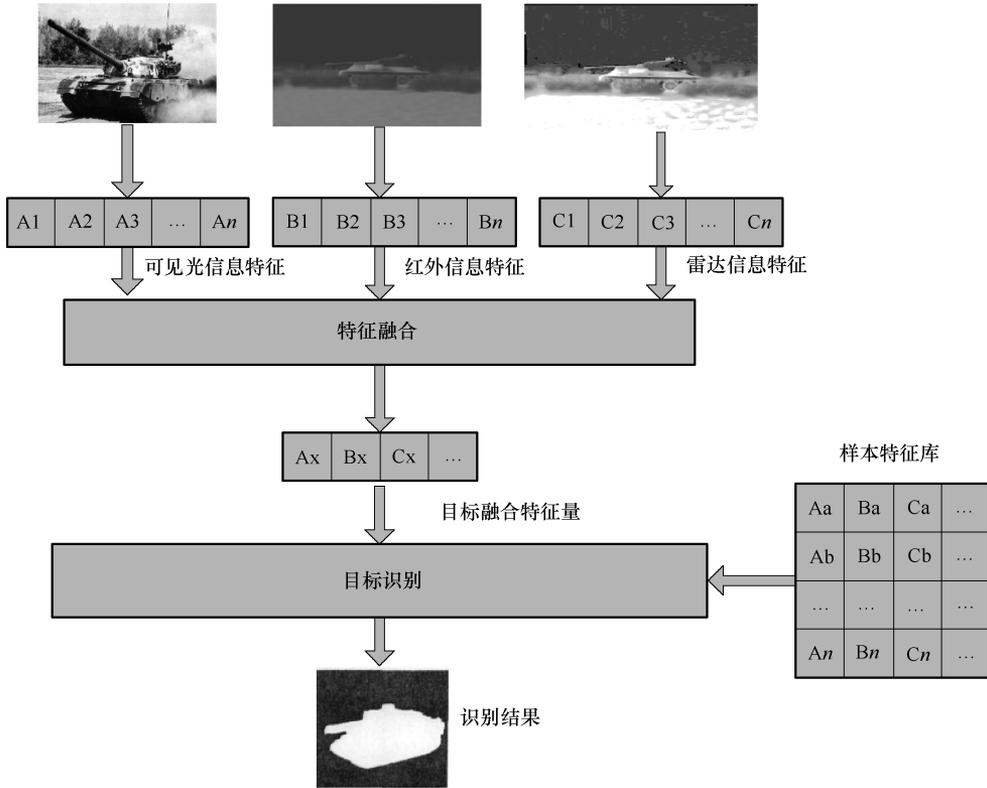


图 2-3 特征级融合的目标识别原理图

效的包含目标信息的目标特征数据 ($A_1, A_2, A_3, \dots; B_1, B_2, B_3, \dots; C_1, C_2, C_3, \dots; \dots$)。

特征融合：对以上目标特征数据通过融合算法，降低特征空间维数，保证维数压缩后特征数据的熵、能量和相关性不改变，消除提取的特征空间中特征表示向量之间的冗余信息，获得融合特征量 ($A_x, B_x, C_x, \dots; \dots$)。

分类识别：通过目标识别算法结合目标样本特征数据库中存放的各种潜在目标特征融合数据 ($A_1, A_2, A_3, \dots; B_1, B_2, B_3, \dots; C_1, C_2, C_3, \dots; \dots$) 进行分类识别待识别目标。

传感器获取的目标原始信息一般目标信息都较弱且包含噪声，必须对其进行预处理（如消除噪声、增强目标、图像分割）。然后从经过预处理的原始目标信息中通过特征提取和选择，获取目标的融合识别特征信息。在目标识别之前，应建立特征数据库（包含了可能的目标特征信息），采用适当的分类识别算法比较传感器提取的目标特征与特征数据库中的特征信息，识别出目标类别。多传感器系统由于存在异类传感器，从传感器提取的特征信息内容与形式或许存在

不同，通过融合算法将这些信息转换成统一形式，使目标特征信息之间可以相互比较和综合。对于单一目标，特征提取后即可进行目标的分类识别；对于多目标，需要通过各传感器提供的信息，判断目标属于系统航迹数据库中的哪一条航迹，判定各传感器提取的目标特征属于哪一个目标，进行信息相关确定。根据信息相关的结论，可由位置融合获得的目标属性信息（如经纬度、速度、加速度等）辅助完成目标识别。最后由分类识别算法根据特征数据库中的目标特征信息对目标进行分类识别。

目前，信息融合目标识别系统受软、硬件等技术能力制约，还与传感器物理性质和质量、通信等有关，实际应用环境并不普遍适用。

2.3 特征级融合目标识别的关键问题分析

2.3.1 特征数据库

多传感器信息融合目标识别的识别率和实时性在很大程度上取决于特征数据库建立的完备程度。现有的目标识别算法模式，都需要与先验模式进行比较。将各种可能存在的目标特征信息存放在特征数据库中。因为目标特征数据库构成了目标识别分类器的先验模式集。它包含了各种目标的特征信息，也可以包含目标的身份信息和辐射源参数等详细信息。特征信息是可被各种传感器直接或间接测量的实变量提取的特征数据，如幅度的直观特征、统计特征、变换系数特征、代数特征等。身份信息属于非实值变量，是对目标的详细分类描述，例如目标的类型、种类、级别和发射机型号等。辐射源参数包括可被电子支援措施（ESM）检测到的辐射源名字和类别，例如载频、重频、脉宽和脉幅等。

2.3.2 特征提取

面对各种异质传感器组合的多传感器系统，传感器的物理探测原理不同，获取的目标信息有图像和非图像类等较大的性质差异，并且信息量巨大。如何选择特征提取方法，能提取出目标独有的、可用于区别其他目标的最本质的信息属性是特征级融合目标识别的难点，也是关键点。多传感器获取的目标特征有代数特征、图像视觉特征、图像统计特征、变换特征等。

2.3.3 特征融合

特征融合的过程就是将特征提取后特征信息通过融合处理，获得多个特征数据的特征综合——一种新的联合特征量，为目标的识别处理、决策提供服务。

特征融合算法是对提取的原始目标特征空间降低了特征空间维数，消除提取的特征空间中特征表示向量之间的冗余信息，并且保证维数压缩后特征数据

的熵、能量、相关性不改变，利于后继的目标识别。总之，特征融合算法就是获得有利于目标识别的有效的、低维的特征表示向量。

特征融合在部分文献和研究者中也称为特征选择，特征提取和特征选择有些系统是合在一起，有些是分开，算法中它们没有明确的界限，即特征提取和特征选择、特征融合是融合在一起的算法。特征融合分为目标状态信息融合和目标特性信息融合。目标状态信息融合用于目标跟踪，主要是实现状态矢量的估计和参数的相关。而特征级目标特性融合主要实现目标属性特性的联合，实现目标属性种类的判定。

特征融合算法按特征向量的产生方式可划分为特征选择和特征组合两类实现方法。特征选择方法是待融合的所有特征量放在一起，用某种方法从原向量选择产生一个新的特征向量，例如，遗传算法、Tabu 搜索算法等。特征组合方法将所有特征向量按某种规则组合成新的特征向量，例如，串行和并行融合策略。遗传算法详见第 9 章，串行和并行融合策略是简单而有效的特征融合方法，其缺点就是所有大量特征信息全部保留，存在大量的冗余特征量，数据维数高，计算复杂，影响识别的精确性和实时性。

1) 串行特征融合方法：设样本模式空间 Ω 的两个特征空间为 A 和 B ，对 $\Gamma \in \Omega$ 的任意样本，特征量可以表示为 $\alpha \in A$ 和 $\beta \in B$ 。串行特征融合方法就是将 α 、 β 串成了一个特征量 γ ，公式如下：

$$\gamma = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (2-1)$$

由式 (2-1) 可知，若 α 是 n 维的， β 是 m 维，那么串行组合的特征量 γ 是 $(m+n)$ 维。因此，由串行融合而成的特征量构成 $(m+n)$ 维的融合特征空间，后续的目标分类识别就是在这个融合特征空间中进行。

2) 并行特征融合方法：设样本模式空间 Ω 的两个特征空间为 A 和 B 。对于 $\Gamma \in \Omega$ 的任意样本，特征量可以表示为 $\alpha \in A$ 和 $\beta \in B$ 。并行特征融合方法将 α 、 β 两个特征量并成了一个复合量 γ ，公式如下：

$$\gamma = \alpha + i\beta \quad (2-2)$$

式中， i 为虚数单位。 α 和 β 特征量维数不一致时，低维的特征量需要补 0，两个特征量才能并行融合。例如， $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ ， $\beta = (\beta_1, \beta_2)^T$ ，则首先将 β 变为 $\beta = (\beta_1, \beta_2, 0)^T$ ，然后融合成融合特征量 $\gamma = (\alpha_1 + i\beta_1, \alpha_2 + i\beta_2, \alpha_3 + i0)$ 。

定义一个在 Ω 上的并行融合特征空间 $C = \{\alpha + i\beta \mid \alpha \in A, \beta \in B\}$ 。这是一个 n 维的复向量空间，其中 $n = \max\{\dim A, \dim B\}$ 。在这个空间里，内积可定义为

$$(X, Y) = X^H Y \quad (2-3)$$

式中， $X, Y \in C$ ， H 表示共轭转置。

定义以上内积的复向量空间，称为酉空间。在酉空间可引入以下范数：

$$\|Z\| = \sqrt{Z^H Z} = \sqrt{\sum_{j=1}^n (\alpha_j^2 + \beta_j^2)} \quad (2-4)$$

式中, $Z = (\alpha_1 + i\beta_1, \alpha_2 + i\beta_2, \dots, \alpha_n + i\beta_n)^T$ 。相应地, 复向量 Z_1 和 Z_2 之间的距离可定义为

$$\|Z_1 - Z_2\| = \sqrt{(Z_1 - Z_2)^H (Z_1 - Z_2)} \quad (2-5)$$

遗传算法等优化算法的特征融合方法与串行、并行特征融合方法相比, 可以获得更有效的、低维的融合特征表示向量, 但是融合过程较串行、并行特征融合方法耗费的时间长, 但是降低了后续的目标识别分类时间。

2.3.4 目标识别

特征级融合的目标识别算法目前都是基于模式识别的方法, 可分为统计模式识别、句法模式识别、模糊模式识别、人工智能方法。统计模式识别、句法模式识别这两种方法比较成熟, 发展得较早, 但是自身缺点很多, 现在用得较少。后两种方法目前都是研究的热点, 应用也较多, 由于模糊模式识别和人工智能方法都允许样本有较大的缺损、畸变, 目前这两种方法都在积极的探索中。

(1) 统计模式识别

统计模式识别又称统计决策法, 它的理论基础是概率论和数理统计。其主要方法包括线性分类、非线性分类、贝叶斯决策、聚类分析等分类方法, 是发展较早、也比较成熟的一种方法。为此而引入鉴别函数, 由特征矢量计算出对应于各类别的鉴别函数值, 通过鉴别函数值的比较实行分类。其主要优点是成熟、能考虑干扰噪声等影响、识别模式基元能力强; 缺点是对结构复杂的模式抽取特征困难、不能反映模式的结构特征, 难以描述模式的性质, 难以从整体角度考虑识别问题。

(2) 句法模式识别

句法模式识别也称为语言学方法、结构方法。句法模式识别的方法是用简单的子模式组合描述一个模式, 更简单的子模式的组合又可描述子模式, 结构描述最终得到一个树形, 模式基元就是在底层的最简单的子模式。在句法模式中选取基元就是目标识别中选取特征。对模式提供紧凑并反映其结构关系的描述就是基元的要求, 并且抽取可以用非句法方法。重要的结构信息并不包含在基元本身。模式描述语句就是一组基元与基元的组合关系形成模式的描述, 类似于语言中词组合成为短语和句子的作用, 而词是用字符组合而成的。语法来指定基元组合成模式的规则。一旦基元被鉴别, 识别过程可通过句法分析进行, 即分析给定的模式语句是否符合指定的语法, 满足某类语法的即被分入该类。

(3) 模糊模式识别

面对甲、乙两类归属问题, 传统的二值逻辑模式, 对样本 A 或者属于甲, 或者属于乙。Zadeh 在 1965 年提出模糊集理论, 将传统的二值 0 和 1 的逻辑模

式转化为逻辑 $(0, 1)$ 区间, 此种对事物刻画的模式改变了片面、单纯的通过事物内涵描述其特征的方式, 并且提出了综合事物内涵与外延性态的数学模型, 构建了隶属度函数。模糊逻辑面对甲、乙两类归属问题, 认为 A 既属于甲, 又属于乙, 判断两者归属的结果是根据 A 在甲、乙两类中的隶属度即归属程度不同。模糊模式识别与普通的模式识别方法相比, 具有客体信息表达更加合理, 信息利用充分, 各种算法简单灵巧, 识别稳定性好, 推理能力强的特点。模糊模式识别思想方法对解决模式识别问题起到了很大的推动作用。

如何建立隶属度函数是模糊模式识别的关键, 但是如何建立比较合理的隶属度函数目前还是难点, 是需要进一步解决的问题。模糊统计法、模糊分布法、二元对比排序法、相对比较法和专家评分法等是目前的主要方法。这些方法虽然具有一定的客观规律性与科学性, 也已经开始应用并取得巨大成果, 但建立时也需要人为的主观因素, 因此怎样建立准确合理的隶属度函数还需要进一步摸索出合理的、普适的理论和方法。

(4) 人工神经网络识别

人工神经网络是模拟动物神经系统的某些功能, 采用软件或硬件的方法, 建立了许多以大量处理单元为节点, 处理单元间实现(加权值的)互联的拓朴网络, 进行模拟, 称为人工神经网络。这种方法可以看作是对原始特征空间进行非线性变换, 产生一个新的样本空间, 使得变换后的特征线性可分。同传统统计方法相比, 其分类器是与概率分布无关的。人工神经网络的主要特点在于其具有信息处理的并行性、自组织性和自适应性, 具有很强的学习能力和联想功能以及容错性能等, 在解决一些复杂的模式识别问题中显示出了其独特的优势。

人工神经网络是一种复杂的非线性映射方法, 其物理意义比较难解释, 在理论上还存在一系列亟待解决的问题。例如在设计上, 网络层数的确定和节点个数的选取带有很大的经验性和盲目性, 缺乏理论指导, 网络结构的设计仍是一个尚未解决的问题。在算法复杂度方面, 神经网络计算复杂度大, 在特征维数比较高时, 样本训练时间比较长; 在算法稳定性方面, 学习过程中容易陷入局部极小, 并且存在欠学习与过学习的现象, 泛化能力不容易控制。这些也是制约人工神经网络进一步发展的关键问题。

2.4 本章小结

在各种环境下对目标进行有效的目标特征提取并识别是非常困难的, 仅仅依靠一种或少数几种特征很难准确地进行特征级融合的目标识别, 必须尽可能利用异类多传感器收集到的多种目标特征信息综合出完善的目标特征信息及有效的识别算法进行目标特征级融合识别。

在努力提高特征提取的有效性和可分性的同时要兼顾识别效率，使目标特征信息数据少而精，从而使识别算法能实时地在工程应用中发挥效用。基于异类传感器特征级融合的目标识别技术正在迅猛发展，已经逐步渗透到现代生活的各个领域。如何恰当和充分地利用这一技术和概念进行信息处理，解决关键问题，提高目标识别的准确率，从而增强识别的效率，还有诸多课题、算法需要探索和研究。

多源图像的预处理

本书研究的多源图像的分割方法主要采用分水岭算法，但是采用分水岭算法分割的缺陷之一是对图像中的噪声很敏感，从而使得分割后的图像不能得到正确的轮廓信息。因此，对于改进分水岭算法的方法中，图像预处理是必不可少的步骤。图像预处理是整个图像处理系统的关键环节，而且也是图像识别、图像理解、图像分割的前期处理的程序。图像预处理的目的是改善处理质量下降的图像，从而提高图像在图像处理系统中和视觉上的质量。图像预处理主要包括：图像去噪、图像平滑、图像滤波、图像特征增强、图像恢复等。

3.1 多源图像的去噪

多源图像成像传感器在拍摄图像的过程中，由于使用的器件的局限性，在一定程度上会增加噪声，严重影响了多源图像的视觉效果，甚至影响了正常识别。因此，去除多源图像中的噪声是多源图像预处理的一项重要重要的步骤。

3.1.1 噪声分析

目前，在多源图像系统中，扫描多维图像转换为一维信号，再进行存储、传输、转换和处理，最后通过信号的重组，变成多维图像信号输出。与此同时，多源图像中携带的图像噪声也同样受到分解和合成，在这个过程中，由于受外界和电气系统本身的影响使得多源图像的噪声变得相当复杂。多源图像的噪声按其起源不同，常见的有电子噪声、抖动噪声、电噪声、胶片颗粒噪声。

电子噪声是由于阻性器件中的电子不规则热运动而形成，经常使用均值为零的高斯白噪声表示电子噪声的模型，并用零均值高斯白噪声的标准差来完全代表电子噪声的强弱；抖动噪声是由承载物体或采集装置等机械运动引起的一种噪声；光电子噪声的本质是光的统计性质和传感器中光诱导的电转换引发的一种噪声，经常使用泊松分布作为其模型，特别的，当在光线较强时，泊松分布趋于高斯分布；胶片颗粒噪声是因为在曝光过程中只有个别被曝光，而引起颗粒密度发生转变而产生的一种噪声。通常情况下，胶片颗粒噪声采样用高斯白噪声表示其有效模型，其分布呈现一种随机性。

多源图像中可见光图像的噪声除了有电子噪声、抖动噪声、光电子噪声、胶片颗粒噪声之外，还有转移噪声、复位噪声、散粒噪声、暗流噪声等。多源图像中的红外图像，由于受环境的干扰和本身物理量转变的影响，使得红外图像的噪声情况极为复杂。红外图像中的噪声有热噪声、 $1/f$ 噪声、散粒噪声、辐射噪声、温度起伏噪声、有色噪声。热噪声是由于分子热运动引发载体分子的热运动而产生的噪声； $1/f$ 噪声是电流运动而产生的噪声，其功率谱随频率成反比变换，因而在低频段时， $1/f$ 噪声较显著；散粒噪声是由半导体器件中产生的重组载体的变化率引发的；辐射噪声通常称为光子噪声，是接收器件的极限噪声且属于白噪声，该噪声功率谱是均匀的，且该噪声与频率无关；有色噪声的光谱响应是非均匀的。

总之，多源图像的噪声是内、外部噪声于一体的综合噪声源，为了减少多源图像噪声对分水岭分割结果的影响，提高多源图像的品质，多源图像进行有效的去噪处理是特别重要的。

多源图像理想消噪效果体现在消除多源图像噪声的同时保护或加强图像的特点信息。于是，多源图像消噪的研究工作也是围绕这一效应开展的。只有兼顾去噪和保留多源图像正确的轮廓信息的前提下，才能保证通过去噪处理后的多源图像最大程度地接近不含噪声的原始图像。因而，多源图像去噪方法的研究具有重要意义，主要体现在以下几个方面：

1) 多源图像的去噪处理使得多源图像的信息能够被正确识别。由于成像机理的不同，当多源图像中含有的噪声比较严重时，多源图像就变得极为模糊，不仅使多源图像失去了存储信息的实际意义，而且也影响了对多源图像的观察，同时也干扰了对多源图像信息的正确理解。

2) 多源图像的去噪处理使视觉识别信息的精度得到了提高。

3) 多源图像的去噪处理为多源图像做进一步图像预处理提供了有力的保证。

4) 多源图像的去噪处理方法依据有差异的数学理论知识。这些方法的内部机理及之间的相应关系，对探求出较好消噪效果起着重大作用。

5) 多源图像去噪方法的研究具有广阔的应用前景。

3.1.2 常用的图像去噪方法

多源图像经常使用的消噪方式有空间域去噪和变换域去噪。空间域去噪方法是原图像上的像素对灰度值进行数据运算操作。这种数据运算又包含点运算和局部运算。点运算是像素点逐点运算；局部运算是图像中像素点域的相关空间域的运算。变换域去噪方法是对多源图像进行某种变换处理，即将多源图像从空间域转到变换域后，对其变换后的系数采取相应的处理，最终将处理后的系数通过反变换再转换到空间域，来完成多源图像消噪的目的。

空间域去噪方法拥有较完整的理论基础并且是典型的去噪方法。局部平均

法是一种空间域消噪方法中的经典方法。局部平均法的基本思想是像素灰度值用像素邻域各像素的均值灰度值代替，从而实现图像去噪的目的。在此方法中，图像被认为是由多个灰度值不变的小块构成，邻像素间空间相关性极强，而图像中的噪声则是统计独立的，因此由于噪声的统计独立性，该方法能简单快速地去除噪声。局部平均法中普通的方法是加权邻域平均法，该方法的优点是邻域中的每个像素被均等地对待，缺点是在去噪的同时会在边缘和细节处发生模糊现象。

变换域去噪方法又称频域去噪法，该方法利用噪声信号和有用信号在变换域表现出的不同特征来完成有效地去除噪声的目的。常见的变换域去噪方法有基于独立分量分析的去噪算法、基于傅里叶变换的去噪算法、基于沃尔什-哈达玛变换的去噪算法、基于小波变换的去噪算法、基于离散余弦变换的去噪算法、基于多尺度几何分析的去噪算法。变换域去噪方法对那些空域中无法进行有效分析的信号进行去噪有良好的效果。下面介绍几种具有代表性的变换域去噪方法。

(1) 基于独立分量分析的去噪算法

基于独立分量分析的去噪算法起源于 20 世纪 90 年代的盲信号分离，能够对图像实施特定的线性分解，并使图像的像素分解成为独立统计分量，从而达到去除噪声的目的。该方法的缺陷是运算复杂度高并且运算时间较长，其突出的优点是，由于该方法假设图像中独立成分为非高斯分布，因而能够有效地区分有用信号和噪声。

(2) 基于傅里叶变换的去噪算法

基于傅里叶变换的去噪算法的基本原理是图像中的有用信号频谱主要集中在某个限定的低频带，而高频带主要是噪声的频谱，利用低通滤波方法在频域内完成图像去噪的目的。该算法的性能优劣关键依靠于低通滤波器的性能优劣和截止频率的合适选择。此算法的优点是能够有效地去除高频噪声，缺点是去噪后的图像具有模糊边缘和细节。

(3) 基于小波变换的去噪算法

基于小波变换的去噪算法的原理是利用一个母函数在空间尺度上的伸缩与时间上的平移获得了一个函数族，然后将这族函数去逼近或表示信号，从而获得一种能自适应各种频变成分的有效信号分析手段。基于小波变换的去噪算法不仅能够解决非整体信息的图像的去噪处理问题，而且还能获得图像的局部频谱信息，具有不同频率特性，弥补了基于傅里叶变换的去噪算法中不能描述随时间变化的频率特性的不足。

3.1.3 几种较新的去噪方法

多源图像的去噪方法有许多种，但是随着新问题的不断出现而使学者们研

究出了许多新的去噪方法。例如，基于阈值的小波变换的去噪方法、基于模糊加权均值的去噪方法、基于形态学开闭运算的去噪方法等。

1. 基于阈值的小波变换的去噪方法

多源图像经过基于小波变换的去噪算法处理后，其图像中幅值较大的小波系数一般为有用信号，而幅值较小的小波系数一般情况下是噪声信号，因此选择一个最佳的阈值进行阈值处理小波系数，就可以把有用信号的系数保留下来，同时去除大部分的噪声系数，甚至将噪声系数减小到零，从而完成去噪处理。

基于阈值的小波变换的去噪方法，就是对有噪声的图像先进行一次小波变换，得到几个分量，分别用 $W^1 y_{j,(k,l)}$ 、 $W^2 y_{j,(k,l)}$ 、 $W^3 y_{j,(k,l)}$ 来表示，用 $HL^j_{(k,l)}$ 、 $LH^j_{(k,l)}$ 、 $HH^j_{(k,l)}$ 来表示其高频分量，将这些分量作为小波系数，其中 (k, l) 为二维小波变换系数， $j = 1, 2, \dots, J_m$ 表示分解层次。得到的低频分量 $LL^j_{(k,l)}$ 用 $W^0 y_{m,(k,l)}$ 表示，且认为 $LL^j_{(k,l)} = W^0 y_{m,(k,l)}$ ，设计算得到的阈值为 λ 。通常采用的阈值法有假设检验法、最大限度地减少风险阈值的方法、贝叶斯收缩阈值法。一般采用通用的阈值法： $\lambda = \sigma_n \sqrt{2 \log(N)}$ ，其中 N 是图像中小波系数的总数， σ_n 是噪声的标准差。一般采用噪声标准差鲁棒中的估计值 σ_n 为 $MAD/0.6745$ ，这里的 MAD 代表对原图像的第一次小波分解获得的小波系数的中值。一般情况下采用直接阈值法对小波变换系数 $W^i y_{j,(k,l)}$ （其中 $i=0, 1, 2, 3; j=1, 2, 3, \dots, J_m$ ）进行阈值处理。主要有以下两种方式。

(1) 硬阈值法

先对噪声图像实施小波变换，忽略不计粗尺度信号，只对细节信号做阈值处理，当某个位置小波变换值小于阈值时将其置零，大于阈值时将其保留原值。其表达式为

$$W^i y_{j,(k,l)} = \begin{cases} W^i y_{j,(k,l)}, & |W^i y_{j,(k,l)}| \geq \lambda \\ 0, & |W^i y_{j,(k,l)}| < \lambda \end{cases} \quad (3-1)$$

最后，利用小波变换重构，求出去噪后的图像。

(2) 软阈值法

先对噪声图像实施小波变换，并求得带有噪声的小波系数，再将细节信号阈值化处理。当小波变换在图像中的位置值小于阈值时将被设置为零，并且当小波变换位置值大于阈值时，进行以下运算，其中 $\text{sgn}(x)$ 为符号函数。

$$W^i y_{j,(k,l)} = \begin{cases} \text{sgn}(W^i y_{j,(k,l)}) [|W^i y_{j,(k,l)}| - \lambda], & |W^i y_{j,(k,l)}| \geq \lambda \\ 0, & |W^i y_{j,(k,l)}| < \lambda \end{cases} \quad (3-2)$$

最后对小波系数 $W^i y_{j,(k,l)}$ 的估计值 $W^i y_{j,(k,l)}$ 做小波变换的反变换以重建图像，得到去噪后的图像。

2. 基于模糊加权均值的去噪方法

基于模糊加权均值的去噪方法依赖于模糊隶属度函数的概念，对加权均值

优化处理，使其不仅能有效地降低高斯噪声，同时也对混合噪声及脉冲噪声有较好的抑制能力。该方法的优点是既能保护图像的边缘信息，也能有效地减弱边缘模糊现象。

3. 基于形态学开闭运算的去噪方法

基于形态学开闭运算的去噪方法的基本思想是借助数学形态学，用特定形态的元素去衡量并提取与之相应的原图像中的形状，当结构元件移动到原始图像，我们可以研究图像中各部分的相互关系，以了解图中各部分的结构特点，区分有用的图像信息和图像信息的噪声，为了达到目的，从而完成去噪处理。因此，该方法去噪效果主要依赖于其结构元素的选取。常用的结构元素包括：圆形、方形、菱形、六边形、线段等。目前没有特定的方法选取结构元素，只能根据图像要求的处理结果及图像本身特征来决定。

数学形态学是由一组形态学基本的代数算子构成，它的基本运算包含：开、闭运算，腐蚀，膨胀。

集合 B 腐蚀集合 A 表示为 $A \ominus B$ ，其定义为

$$A \ominus B = \bigcap_{b \in B} \{A - b\} = \bigcap_{b \in B} A_{-b} \quad (3-3)$$

通过腐蚀运算对图像进行腐蚀操作后，若所有结构元素均为正值，则输出图像比输入图像偏暗；若输入图像的细节小于结构元素的区域，则细节亮的效果会被削弱。因此，腐蚀运算能够使图像边界向内部紧缩及消除边界点的特征。

腐蚀的对偶操作是膨胀，集合 B 膨胀集合 A 表示为 $A \oplus B$ ，其定义为

$$A \oplus B = \bigcup_{b \in B} \{A + b\} = \bigcup_{b \in B} A_b \quad (3-4)$$

膨胀运算对图像进行膨胀操作后，若所有结构元素均为正值，则输出图像比输入图像偏亮，与此同时暗细节的效果会被减弱。因此，膨胀运算能够使图像边界向外部扩张。

基于腐蚀和膨胀运算可以推导出各种数学形态学的基本运算。开运算和闭运算就是这两种运算的组合运算。集合 B 对集合 A 的开运算就是先做腐蚀运算后做膨胀运算，表示为 $A \circ B$ ，其定义为

$$A \circ B = (A \ominus B) \oplus B \quad (3-5)$$

开运算的优势体现在能够去除比结构元素要小的突出的亮点噪声，并且能够维持图像整体区域的特征以及大部分的灰度层级。开运算的对偶运算是闭运算，集合 B 对集合 A 作闭运算就是先做膨胀运算再做腐蚀运算，表示为 $A \cdot B$ ，定义为

$$A \cdot B = (A \oplus B) \ominus B \quad (3-6)$$

闭运算的优点是不仅能够在不明显改变物体的形状和面积的前提下对图像做去噪处理，而且能够在连接邻近物体的同时填充图像内部的细小空洞。因此，对同一幅图像使用相同的结构元素做形态学开闭运算的去噪处理，闭运算能够

起到削弱膨胀运算对边界扩张的作用；而开运算则能够起到补充被腐蚀消除的部分图像的作用。因而形态学开闭运算的去噪方法具有良好的去噪效果和保持清晰的目标轮廓的优点，与其他去噪算法相比，它的这个优点更为显著。

图像质量评价中的主观评价也是人类视觉特性，但评估质量受人为因素的影响，不能准确地描述图像的定量分析，因此本章采用图像的两种主要的客观评价来比较基于阈值的小波变换的去噪方法、基于模糊加权均值的去噪方法、基于形态学开闭运算的去噪方法这几种新的去噪方法的优劣。采用的两种主要的客观评价方法是峰值信噪比（PSNR）和信噪比（SNR），其中峰值信噪比可定量描述对图像的去噪质量；信噪比用来比较含噪声的原图像与去噪后图像的质量，信噪比越大，图像品质就越好。表 3-1 列出了几种较新的去噪方法的峰值信噪比和信噪比。

表 3-1 几种较新的去噪方法的峰值信噪比和信噪比 （单位：dB）

方法种类	峰值信噪比	信噪比
基于硬阈值的小波变换的去噪方法	18.75	6.32
基于软阈值的小波变换的去噪方法	19.62	7.18
基于模糊加权均值的去噪方法	18.69	6.29
基于形态学开闭运算的去噪方法	19.85	7.37

因此，由表 3-1 中的数据可知，基于模糊加权均值的去噪方法的峰值信噪比和信噪比是最低的，而基于形态学开闭运算的去噪方法的峰值信噪比和信噪比是最高的。因而，相对于其他去噪方法，采用基于形态学开闭运算的去噪方法得到的图像去噪质量较高、效果较好。

本章采用了基于形态学开闭运算的去噪方法对多源图像做去噪处理，使得多源图像在被分水岭算法分割之前不仅消除了无规则的明暗噪声和细节扰动，而且也消除了易导致过分割的区域噪声与细节，更重要的是使多源图像的重要轮廓得以保留。因此，此形态学开闭运算去噪方法对改进分水岭算法起到了关键作用，它使分水岭分割算法对图像中噪声敏感的缺陷有所改善，更为重要的是它使分水岭算法分割前输入的多源图像具有正确的轮廓信息，即对修正多源图像起到了关键作用，为分水岭算法正确、有效的分割奠定了一定的基础。

采用了基于形态学开闭运算的去噪方法对分水岭算法的改进，不仅是利用它的去噪处理改善多源图像的输入质量，更重要的是利用它对多源图像去噪的同时也将多源图像进行梯度化处理，将多源图像变成它的梯度图输出。因为分水岭算法的基本思想来自地理学，将被分割图像视作地形图并求取梯度图像，因此采用基于形态学开闭运算的预处理恰好是分水岭分割前的第一步，即图像梯度化处理。传统的梯度化处理方法是通过微分算子，例如 Sobel、Prewitt、

Canny 等求得梯度图。而本章选用形态学开闭运算的图像即为梯度图像，是因为形态学梯度运算有非线性的特点，能够加强图像中极值对比度，并且能够保护图像中目标轮廓的区域。

本章采用 VC ++ 6.0 开发环境编程调试完成形态学开闭运算的去噪算法，实验结果如图 3-1 和图 3-2 所示。其中图 3-1 为可见光图像的形态学开闭运算去噪后的图像，图 3-2 为红外图像的形态学开闭运算去噪后的图像。

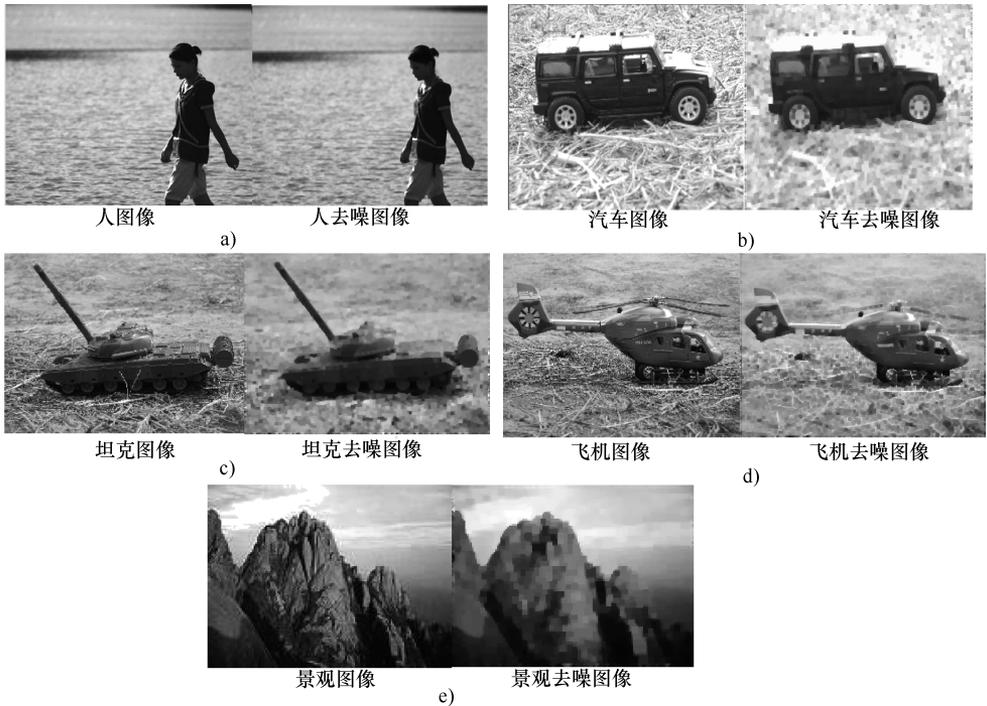


图 3-1 可见光图像去噪后的图像

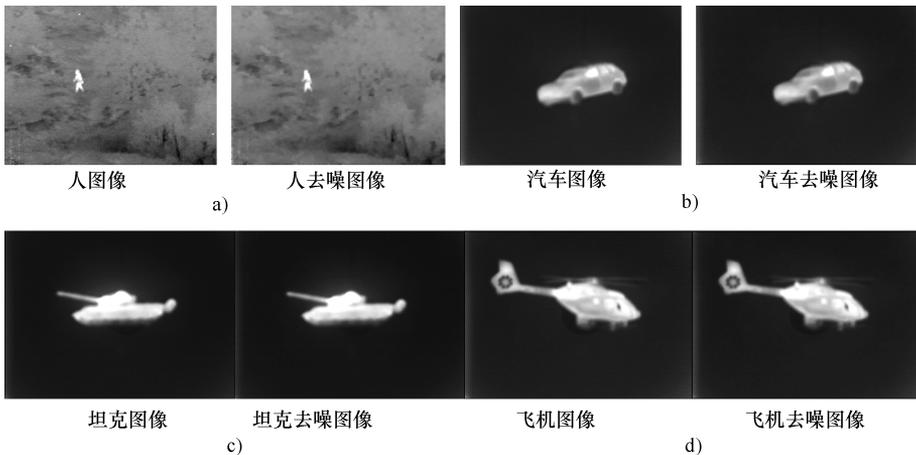


图 3-2 红外图像去噪后的图像

3.2 多源图像的平滑

由于本书采用的分水岭算法在分割过程中会出现很微小琐碎、无语义学含义的分割的小区域，因此，为了克服分水岭初始分割的区域数众多的弊端，本章采用图像平滑算法对去噪后的多源图像进行处理，目的是补充邻域信息同时填补空穴，并且连接细小零碎的区域断口以减少初始分割的区域数。

常见的图像平滑方法分为线性方法和非线性方法。线性方法一般使用一个固定的平滑模板作用于图像，或称其移动平均法；非线性方法一般采用不同方向系数的平滑模板平滑图像，而这种模板是由某些统计特性的图像中相邻像素的分类排序所决定的。线性方法包括：均值平滑法、邻域平均法、算术均值平滑法、几何均值平滑法、谐波均值平滑法、逆谐波均值平滑法、阿尔法均值平滑法等算法；非线性方法包括：加权平均法、梯度倒数加权法、算数加权法等算法。综合上述讨论，本节将重点研究均值平滑法、邻域平均法、加权平均法以及图像平滑中一种典型的算法，即选择式掩模平滑法。

3.2.1 均值平滑法

均值平滑法是求图像中像素的平均，设图像 $f(x, y)$ 是 $N \times N$ 的阵列，图像 $g(x, y)$ 是平滑后的图像，并且包括 (x, y) 限定邻域的一些像素灰度级的均值确定平滑操作后的图像 $g(x, y)$ 中的每一个像素的灰度值，因而平滑操作后的图像 $g(x, y)$ 的表达式表示如下：

$$g(x, y) = \frac{1}{M} \sum_{(i, j) \in S} f(i, j) \quad (3-7)$$

式中， S 为 (x, y) 点邻接中心点的坐标的集合且不包含点 (x, y) ； M 是 S 内坐标点的数目， $x, y = 0, 1, 2, 3, \dots, N-1$ 。这种均值平滑法用一定范畴的灰度平均值取代像素值，减弱了图像灰度的突变现象，但是因为图像的边界存在这种突变，因而采用均值平滑法会使图像的边缘产生模糊效应，不能达到好的平滑效果。

3.2.2 邻域平均法

均值平滑法是最常用的线性平滑方法，它的 3×3 的模板系数均为 1，当对此模板的输出响应除以 9 再平滑图像时，这种方法被称为邻域平均法。邻域平均法通常采用全部系数都取一样值的 3×3 和 5×5 的模板，表示如下：

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \frac{1}{25} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (3-8)$$

邻域平均法的基本思想：设图像 $f(x, y)$ 为输入图像，图像 $g(x, y)$ 是平滑后的图像， S 是点 (x, y) 的邻域，那么它的四邻域和八邻域的坐标集合分别表示如下：

$$S = \{(x, y + 1), (x, y - 1), (x + 1, y), (x - 1, y)\} \quad (3-9)$$

$$S = \left\{ \begin{array}{l} (x - 1, y - 1), (x, y - 1), (x + 1, y - 1), (x - 1, y) \\ (x + 1, y), (x - 1, y + 1), (x, y + 1), (x + 1, y + 1) \end{array} \right\} \quad (3-10)$$

四邻域和八邻域的邻域平均法平滑图像，其图像的边缘也会有一定程度的模糊效应，为了减弱这种模糊效应，对四邻域和八邻域的邻域平均法进行了改进，演变为阈值邻域平均方法。阈值邻域平均法的基本思想是：首先要处理的像素的平均灰度值不包含待处理像元像素值，若此平均值与待处理像元像素差的绝对值大于阈值，则用灰度平均值代替像素值；相反，如果该差值的绝对值小于阈值，此时该像素的灰度值保持不变。阈值邻域平均法的数学表达式如下：

$$g(x, y) = \begin{cases} f(x, y) & \left| \frac{1}{M_{(m,n)} \in S} \sum f(x - m, y - n) - f(x, y) \right| \leq T \\ \frac{1}{M_{(m,n)} \in S} \sum f(x - m, y - n) & \left| \frac{1}{M_{(m,n)} \in S} \sum f(x - m, y - n) - f(x, y) \right| > T \end{cases} \quad (3-11)$$

阈值邻域平均法的关键是阈值 T 的预先选定，若 T 太小，会使图像边缘的模糊效应显著，若 T 太大，则会使去噪效果变差，因此依靠先验知识选取恰当的阈值尤为重要。可见，上述均值平滑法、邻域平均法对图像平滑后，图像均有一定的模糊程度，给后续图像处理带来一定的困难。为了克服均值平滑法、邻域平均法中简单局部平均的缺陷，可采用加权平均法平滑图像。

3.2.3 加权平均法

加权平均法的基本原理是将参与平均的像素赋予不同的权重。一般情况下依据距离待处理像素的远近确定权值或依据与待处理像素的灰度值的接近程度确定权值。当图像中的像素距离待处理像素较远时，此像素赋予较小的权值；当图像中像素的灰度值与待处理像素灰度值偏差大时，此像素被赋予较小的权值。因此，从权值上可以区分出一些重要的像素，从而达到平滑图像的目的。但是，该方法受牵制的因素很多，例如，怎么选取邻域点的权重系数，怎么选取参与平均的点数以及怎么选用邻域的形状、大小以及方向。因此这种方法也不是最佳的图像平滑方法。

3.2.4 选择式掩模平滑法

由于均值平滑法、邻域平均法在平滑时，均存在由平均化引起的尖锐边缘和轮廓线条模糊的缺陷；而加权平均法又受很多牵制因素的影响，因此研究一

种自适应的局部平滑算法显得尤为必要。选择式掩模平滑法就是依据图像的背景与目标通常情况下具备不近相同的统计特征，因而具有不同的方差和均值，它以模板运算为基础，分别计算每个模板的平均值及方差，然后选取方差最小的模板进行平均化而完成自适应局部平滑操作。选择式掩模平滑算法的优势是能够较好地保留图像细节；补充邻域信息同时填补空穴，并且连接细小零碎的区域断口；能够获得比较清晰的边缘轮廓。

选择式掩模平滑算法选用 5×5 的模板，在模板中选用中心像素 (x, y) 作为基准点，构造 4 个六边形、4 个五边形、1 个边长为 3 的正方形的模板。

使用上述类型的模板依次计算出各个模板下的均值和方差，其中式 (3-12) 是均值的计算公式，式 (3-13) 是方差的计算公式。

$$M_x = \frac{\sum_{k=1}^{k=N} f(x, y)}{N} \quad (3-12)$$

$$\sigma_x = \sum_{k=1}^{k=N} [f^2(x, y) - M_x^2] \quad (3-13)$$

在式 (3-12) 与式 (3-13) 中， N 对应于各掩模的像素个数， $k=1, 2, \dots, N$ 。通过式 (3-12) 计算得到的 M_x 进行排序，再计算相应的方差，最后找到方差最小的掩模的相应灰度值的均值 M_x ，并将这个均值作平滑结果输出。

本章采用 VC++6.0 开发环境编程调试完成选择式掩模平滑算法，并将此平滑算法对形态学开闭运算去噪后的结果图像做平滑操作，其实验结果如图 3-3 和图 3-4 所示，其中图 3-3 为可见光图像的选择式掩模算法的平滑结果图像，图 3-4 为红外图像的选择式掩模算法的平滑结果图像。



图 3-3 可见光图像平滑结果图像

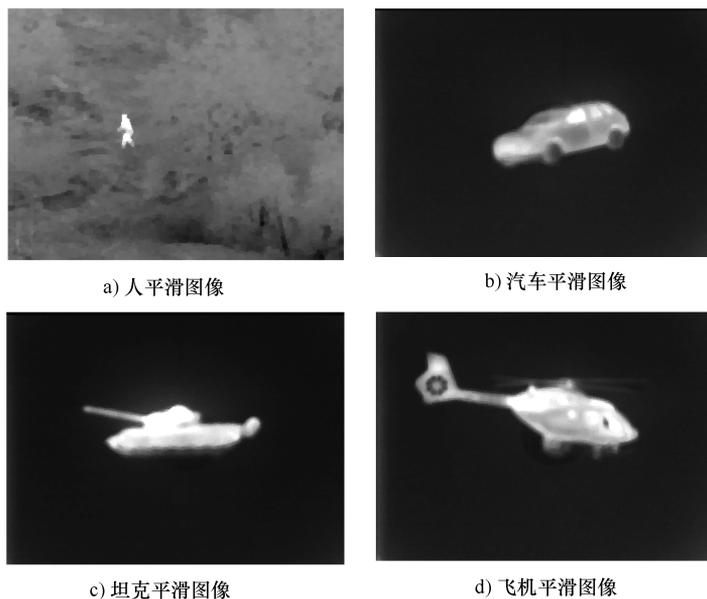


图 3-4 红外图像平滑结果图像

3.3 多源图像的滤波

本书选用分水岭算法对多源图像实施分割操作，而分水岭算法要求梯度图像输入，因此本章已经采用基于形态学开闭运算的预处理将多源图像转变成梯度图像，但是这种载入的梯度图像极易受噪声的影响，可能会使图像中的局部分割线发生偏移，严重时可能会丢失局部分割线，而造成多源图像中的目标和背景粘连在一起，从而导致不能完成分割的目的。因此，对平滑处理后的多源图像的梯度图像进行滤波处理，去除梯度图像中的噪声和杂波，并且提高输出的梯度图像的质量及信噪比，对改善分水岭的分割效果有极其重要的作用。

目前，主要的滤波方法有低通滤波、高通滤波、中值滤波、维纳滤波等。其中低通滤波又包括：理想低通滤波、指数低通滤波和巴特沃兹（Butterworth）低通滤波；高通滤波又包含：理想高通滤波、指数高通滤波和巴特沃兹高通滤波。

3.3.1 低通滤波

低通滤波的原理是去除傅里叶变换中所有高频成分，即通低频、阻高频，保留有用的低频信息。其原理框图如图 3-5 所示。图中， $f(x, y)$ 为原始图像，

$F(u, v)$ 为经过傅里叶变换的图像, $G(u, v)$ 为低通滤波后的图像, 且滤波后的图像 $G(u, v)$ 要经过低通滤波的传递函数 $H(u, v)$ 获得, 最后经过傅里叶逆变换而获得去噪后的图像 $g(u, v)$ 。滤波后的图像 $G(u, v)$ 的表达式如下:



图 3-5 低通滤波原理框图

$$G(u, v) = F(u, v) H(u, v) \tag{3-14}$$

低通滤波包括: 理想低通滤波、指数低通滤波和巴特沃兹低通滤波。其中理想低通滤波的描述见式 (3-15), 指数低通滤波的描述见式 (3-16), 巴特沃兹低通滤波的描述见式 (3-17)。其中, D_0 代表被指定距离且为非负数值, $D(u, v)$ 代表点 (u, v) 与频率矩形原点的距离。

$$H(u, v) = \begin{cases} 1 & D(u, v) \leq D_0 \\ 0 & D(u, v) > D_0 \end{cases} \tag{3-15}$$

$$H(u, v) = e^{-[\ln(1/\sqrt{2})][D(u, v)/D_0]^n} \tag{3-16}$$

$$H(u, v) = \frac{1}{1 + (\sqrt{2} - 1)[D(u, v)/D_0]^{2n}} \tag{3-17}$$

由于图像中高频段多集中噪声频谱, 而低、中频段常集中图像的有用信息, 所以这几种低通滤波可以去掉突变的高频成分与尖峰噪声。低通滤波滤掉了跳变性的高频分量, 这就意味着滤除了图像边缘的跳变信息, 从而造成了图像边缘模糊现象, 因此, 这种滤波算法是以牺牲图像的清晰度为代价而换取的。

3.3.2 高通滤波

高通滤波的原理是衰减甚至截断大部分的低频成分, 保留高频成分, 即通高频、阻低频。高通滤波的实质是某一邻域内像素灰度值的平均值被每一个像素灰度值减去, 若这个像素是目标像素, 则它的值与邻域内像素卷积后的值相差很远, 且此差值也极大; 反之, 若这个像素是非目标像素, 因此它的值与邻域中的像素卷积的值很相似, 且此差值也极小, 因而可以滤掉每个像素中背景噪声的平均值。高通滤波的原理框图如图 3-6 所示。其中, $f(x, y)$ 是原始图像, $F(u, v)$ 是经过傅里叶转换后的图像, $G(u, v)$ 是高通滤波后的图像, 且滤波后的图像 $G(u, v)$ 要经过高通滤波的传递函数 $H(u, v)$ 获得, 最后经过傅里叶逆变换而获得去噪后的图像 $g(u, v)$ 。

若高通滤波的脉冲响应函数为 $h(x, y)$, 输入的图像信号为 $f(x, y)$, 则输出的滤波后的图像信号为 $g(x, y)$, 且输出的滤波后的图像信号 $g(x, y)$ 的表



图 3-6 高通滤波原理框图

达式见式 (3-18)；其频域表式见式 (3-19)，其中 $F(u, v)$ 是输入的图像信号 $f(x, y)$ 的傅里叶变换， $H(u, v)$ 为高通滤波的脉冲响应函数 $h(x, y)$ 的傅里叶变换， $G(u, v)$ 为输出的滤波后的图像信号 $g(x, y)$ 的傅里叶变换。因此，整个高通滤波过程的表达式见式 (3-20)，其中 F 表示傅里叶变换， F^{-1} 表示傅里叶逆变换。

$$g(x, y) = h(x, y) * f(x, y) \quad (3-18)$$

$$G(u, v) = H(u, v) F(u, v) \quad (3-19)$$

$$g(x, y) = F^{-1} \{ F[f(x, y)] H(u, v) \} \quad (3-20)$$

对于高通滤波的整个过程，最关键的就是选取适当的脉冲响应函数 $h(x, y)$ ，利用 $h(x, y)$ 可以使 $f(x, y)$ 中的低频分量衰减，保留高频分量，起到高通滤波的作用。对于离散的多源图像的数字图像，高通滤波的脉冲响应函数通常情况下选用卷积模板，表示为矩阵 H 。关于模板 H ，它的中心像素权值越小，信号就越不容易通过；反之，其中心像素权值越大，信号越容易通过，且权值周围部分全为 1，因此，高通滤波保留了图像中信号强度很高的微小目标和独立噪声点。常用的一些高通模板有 h_1 、 h_2 、 h_3 、 h_4 等，如图 3-7 所示。由这些模板演变出许多不同形式的高通滤波模板，它们的尺寸也不尽相同，因此滤波效果也不相同。

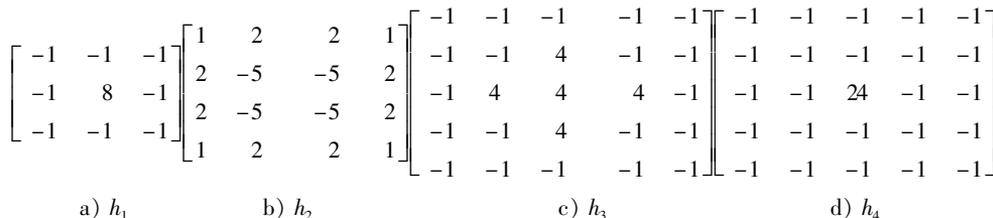


图 3-7 常用的几种高通滤波模板

高通滤波方法一般情况下由输入图像与经过低通滤波后的图像的差值取得，它的一维表达式见式 (3-21)，其中 $f(n)$ 是一维输入信号， $l(n)$ 是低通滤波的输出信号， $g(n)$ 是滤波输出图像的一维输出信号， M 是低通滤波所使用的点数。 $l(n)$ 的描述见式 (3-22)。高通滤波的二维表达式见式 (3-23)，其中， $f(i, j)$ 是输入图像的信号， $l(i, j)$ 是低通滤波的输出信号， $g(i, j)$ 是滤波输出图像的二维输出信号， M 是低通滤波半径， $s(i, j, m)$ 是输入图像的信号求和运算。 $l(i, j)$ 的描述见式 (3-24)。 $s(i, j, m)$ 的描述见式 (3-25)。

$$g(n) = f(n) - l(n) \quad (3-21)$$

$$l(n) = \frac{1}{M}f(n) + \left(1 - \frac{1}{M}\right)l(n-1) \quad (3-22)$$

$$g(i, j) = f(i, j) - l(i, j) \quad (3-23)$$

$$l(i, j) = \frac{1}{M} \left[\sum_{m=0}^{M-1} \left(1 - \frac{1}{M}\right)^m s(i, j, m) + \left(1 - \frac{1}{M}\right)^M s(i, j, M) \right] \quad (3-24)$$

$$s(i, j, m) = \frac{1}{2^{2m+1}} \sum_{\substack{|i-k|=m \\ |j-n|=m}} f(i-k, j-n) \quad (3-25)$$

常用的高通滤波包含：理想高通滤波、指数高通滤波和巴特沃兹高通滤波。其中理想高通滤波的描述见式 (3-26)，指数高通滤波的描述见式 (3-27)，巴特沃兹高通滤波的描述见式 (3-28)。其中， D_0 代表原点与截止频率之间的距离， $D(u, v)$ 代表点 (u, v) 与原点间的距离，且 $D(u, v)$ 的描述见式 (3-29)， n 代表控制传递函数 $H(u, v)$ 从原点开始计算的增长率。

$$H(u, v) = \begin{cases} 0 & D(u, v) \leq D_0 \\ 1 & D(u, v) > D_0 \end{cases} \quad (3-26)$$

$$H(u, v) = e^{-0.347[D_0/D(u, v)]^n} \quad (3-27)$$

$$H(u, v) = \frac{1}{1 + (\sqrt{2} - 1)[D_0/D(u, v)]^{2n}} \quad (3-28)$$

$$D(u, v) = \sqrt{u^2 + v^2} \quad (3-29)$$

以上几种高通滤波方法均可以使图像的部分边缘变得更加清晰，而使物体在背景中更为突出。其中指数高通滤波由于具有较快的增长率，因而相对于其他方法来说更能保留图像的高频分量而起到对图像中的物体勾边的作用。但是，高通滤波虽然能去除背景，但对系统自身噪声（例如点噪声等）却无法滤除，而且在处理图像时，运算复杂，运算量极大，也不利于硬件实现，因此，高通滤波对多源图像的梯度图像的滤波效果并不是最佳的。

3.3.3 中值滤波

中值滤波是一种邻域运算，是通过邻域像素进行排序并取它们的中值来确定中心像素灰度的一种处理方法。它首先选择某个像素为中心点的邻域，再对此邻域像素的灰度值实施排序操作，找到它们的中间值，并将此中间值替换中心点像素的灰度值。中值滤波的描述见式 (3-30)，其中 $f(x, y)$ 是输入图像， $g(x, y)$ 是输出的中值滤波图像， $A(x, y)$ 是像素点 (x, y) 的邻域。

$$g(x, y) = f(x, y) - \text{med}[f(x, y)] \quad (3-30)$$

虽然中值滤波能消除一定的噪声，但会使图像边缘模糊，这主要与所选用的滤波窗口的大小相关，若窗口太大，会使图像中线段和边缘细节受到破坏；若窗口太小，就不能起到滤波的效果。因此，窗口尺寸直接关系到中值滤波的效果，必须考

虑通过窗口时，要有足够的参考样本，并且还要求窗口尺寸大于目标尺寸的两倍；与此同时还要求窗口中所选取的参考样本的形式须与中值滤波模板匹配。可见，中值滤波是受许多因素限制的。在中值滤波中，对于每次移动窗口计算中值时，均要求对窗口内的所有像素值进行排序，显然这是个大开销的计算量，因此，中值滤波的计算效率较低，如果将其应用于实时目标检测中，它将会占用大量的系统时间，从而导致系统的实时性降低。中值滤波对于一些细节较多的图像，尤其是那些点、线、尖顶细节众多的图像，往往中值滤波会抑制有用的信号，严重时甚至导致图像中的有用信息丢失。因此，中值滤波对多源图像的梯度图像的滤波效果也不是最佳的。

3.3.4 维纳滤波

维纳滤波的基本思想是通过估计每个像素的局部邻域的均值和方差来完成滤波的目的，具体地，它将图像中的噪声和有用信息均看作是随机信号，并在分析统计的基础上对随机信号采用最优准则进行滤波处理。假设图像 $g(i, j)$ 的大小为 $M \times N$ ，且图像 $g(i, j)$ 是由噪声 $n(i, j)$ 和真实图像 $f(i, j)$ 组成的，维纳滤波的最终目的是尽可能地降低输出图像 $I(i, j)$ 的噪声信号 $n(i, j)$ ，同时恢复真实图像 $f(i, j)$ 。 $e(i, j)$ 被定义为误差信号，且其描述见式 (3-31)。 μ 是图像的均值，其描述见式 (3-32)； σ^2 是图像的方差，其描述见式 (3-33)；MSE 代表平均误差的度量，其描述见式 (3-34)；输出图像 $I(i, j)$ 的描述见式 (3-35)。

$$e(i, j) = f(i, j) - I(i, j) \quad (3-31)$$

$$\mu = \frac{1}{ab} \sum_{(i, j) \in S} f(i, j) \quad (3-32)$$

$$\sigma^2 = \frac{1}{ab} \sum_{(i, j) \in S} f(i, j) - \mu^2 \quad (3-33)$$

$$\text{MSE} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} e^2(i, j) \quad (3-34)$$

$$I(i, j) = \mu + \frac{\sigma^2 - \mu^2}{\sigma^2} [f(i, j) - \mu] \quad (3-35)$$

维纳滤波的关键是最优准则的选取，且它的性能好坏是通过均分误差值评价的。均分误差值是输出图像与均值之差的二次方的结果，当均分误差值越小时，维纳滤波的效果就越好。由于维纳滤波是一种有效的自适应滤波方法并且具有良好的选择特性，因此其方差可以自适应地调整并不断地修正维纳滤波的输出信号，从而降低了维纳滤波的均分误差值，并更好地保护了图像的高频特征细节信息和边沿信息，获得令人满意的滤波结果。因此，采用维纳滤波对多源图像的梯度图像进行滤波处理，来提高输出的梯度图像的质量及信噪比，为改善分水岭的分割效果奠定一定的基础。

本章采用 VC ++6.0 开发环境编程调试完成维纳滤波算法，并将此滤波算法对选择式掩模平滑后的梯度图像做滤波处理，其实验结果如图 3-8 和图 3-9 所示。



图 3-8 可见光图像的维纳滤波图像

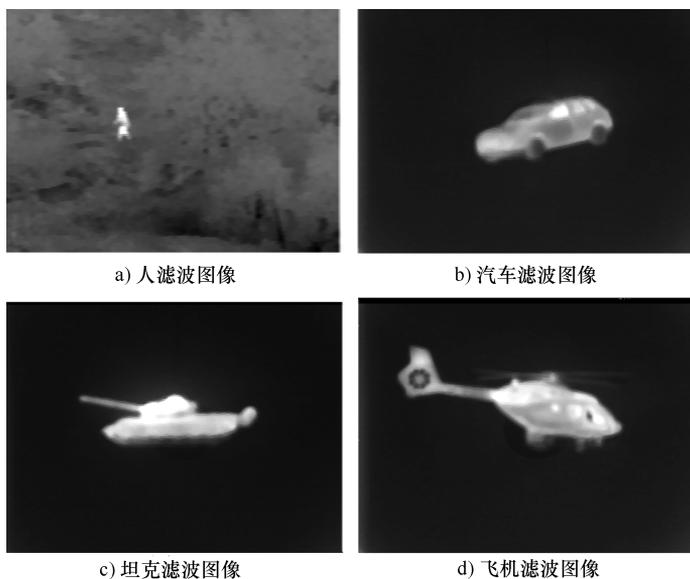


图 3-9 红外图像的维纳滤波图像

综上所述，本章对多源图像做了以下图像预处理：形态学开闭运算的去噪处理、选择式掩模算法的平滑处理以及维纳滤波处理，从而改善多源图像在图像处理系统中视觉上的质量，从而对改善分水岭算法的分割效果起到积极的作用。

3.4 本章小结

本章主要介绍了对多源图像的图像预处理过程，主要包括：多源图像的去噪、平滑以及滤波。首先，介绍了多源图像的去噪处理，例如，对多源图像的噪声分析、多源图像的几种常用的去噪方法以及几种较新的去噪方法，对这些去噪方法的去噪效果进行了分析，并最终选择了形态学开闭运算作为去噪算法；其次，阐述了多源图像重要的平滑方法以及对这些平滑方法的平滑效果做了比较分析，且最终选择了选择式掩模平滑算法作为平滑算法；再次，还讲解了多源图像常用的滤波方法并对这些滤波方法的滤波效果进行了分析，且最终选择了维纳滤波作为滤波算法。综上所述，本章介绍的多源图像的预处理将对改善分水岭算法的分割效果起到重要的作用。

结合阈值分割的分水岭算法

4.1 引言

由于多源图像的不规则性，其局部区域中会隐含存在一些虚假的局部极小值，并且这些局部极小值的数量远远大于目标提取的个数，因而在采用分水岭算法对多源图像进行分割操作时，就会产生过分割现象，从而造成图像的目标轮廓难以被提取，目标不能被正确分割的后果。因此，为了解决传统分水岭算法中虚假的局部极小值过多与过分割问题，本章采用结合阈值分割的分水岭算法对多源图像进行阈值优化操作，不仅消除了易造成过多区域的极小值，而且可以使分水岭算法中的分水点较准确地定位，同时也得到了具有一定意义的边界特征的分割图像，从而改善了传统分水岭算法中易产生过分割的缺陷。

4.2 图像阈值分割概述

阈值分割是一种有效且简单的基于区域的图像分割方法，其目的是将图像中的像素集合依据灰度级划分，从而得到拥有同等性质的每个区域以及具有相同性质的邻接区域，且每个区域都与实际景物对应一致。阈值分割方法的基本思想是依据图像中要分割的目标、背景的不同灰度特征，通过取一个最佳阈值，将图像中像素点隶属于目标还是背景划分出来，从而将图像划分为两类区域的组合而完成图像分割的目的。

阈值分割法的基本原理是，设置有差别的特征阈值，对要分割图像的所有像素点实现分类，即利用确定的阈值和待分割图像的像素点的灰度值进行比较，并根据对比结果将图像像素分为目标和背景这两类像素点的集合。假设原始图像为 $g(x, y)$ ，接着遵循一定的准则确定出待分割图像的阈值 T ；根据此阈值将小于阈值 T 的像素点的值设为 I_1 ，将大于或等于阈值 T 的像素点设为 I_2 ，此时初始图像 $g(x, y)$ 就被分割为两部分，其分割结果为 $f(x, y)$ ，且图像 $f(x, y)$ 的表达式如下：

$$f(x,y) = \begin{cases} I_1 & g(x,y) < T \\ I_2 & g(x,y) \geq T \end{cases} \quad (4-1)$$

从阈值分割的表达式中可以看出，如何确定出最佳阈值是整个阈值法的关键，也是该方法的难点所在。实际中，图像中的像素往往要被分为几类，因此一个单一的阈值是不够的，需要多个阈值对图像中的像素点进行分类操作。因而单阈值算法是多阈值算法的一种特例，多阈值算法又是单阈值算法的进一步发展。图像阈值分割的框图如图 4-1 所示。

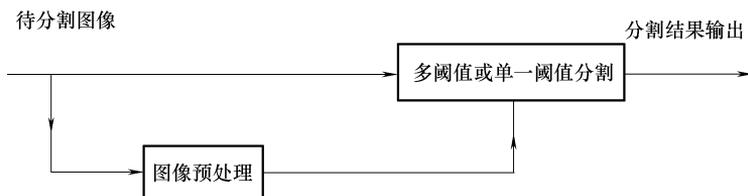


图 4-1 图像阈值分割框图

图像阈值分割的步骤分为两步：首先，确定待分割图像的阈值 T ；其次，将待分割图像的阈值 T 分别与待分割图像的所有像素的灰度级进行比较，从而最终得到划分了区域的图像，即分割后的图像。

4.3 常用的阈值分割方法

常用的阈值分割方法分为全局阈值法、局部阈值法以及动态阈值法。全局阈值法是求出图像最佳分割的阈值，它包含基于点的全局阈值法和基于区域的全局阈值法；局部阈值法是指将要分割的图像划分为一些子图，然后对这些子图采用全局阈值法依次求出最佳分割阈值进行图像分割。换句话说，局部阈值法是图像中多个子图的全局阈值法的组合；动态阈值法是指依据待分割图像每一个像素与其邻接像素灰度值的状况，动态地做运算求出图像的最佳阈值。

4.3.1 全局阈值法

全局阈值法包含基于点的全局阈值法和基于区域的全局阈值法。其中基于点的全局阈值法是依据图像中每个点的灰度值确定阈值的；基于区域的全局阈值法是依据图像中邻域像素点的局部灰度值的分布来确定阈值的。基于点的全局阈值法包含直方图方法、最小误差法、矩量保持法以及最大熵方法等方法；基于区域的全局阈值法包含直方图变换法、概率松弛方法以及灰度分布统计方法等方法。

1. 直方图方法

直方图方法的基本思想是依据待分割图像的灰度分布直方图来确定图像的阈值，包括 p -分位数法 (p -tile 法)、双峰法以及直方图凹面分析法。

p -tile 法是最早的灰度直方图的自动阈值分割方法，它依靠顺次积累灰度直方图的值来获取图像阈值。假设待分割图像中灰度级高的背景中存在一个灰度级低的目标，且已知目标在待分割图像中面积的比例为 p ，接着依次累计灰度直方图的值直到大于或等于 p 时，求得的灰度级就是图像的待求阈值。 p -tile 法计算相当简单，但是在使用该方法时必须预先知道目标在图像中的面积比 p ，如果 p 是未知的或随图像的不同发生变化，该方法就不能发挥分割作用了。

双峰法只适用于理想状态下目标和背景较清晰的图像，且此图像的灰度直方图呈双峰状分布，此时波谷与待分割图像的边缘相对应，因此当选取的阈值恰好是波谷的位置时，就可以获得较好的分割效果。但是，由于实际图像的情况一般来说较复杂，其图像的直方图参差不齐，分辨不出波峰、波谷，这个方法就不适用图像分割了。

直方图凹面分析法适用于直方图中没有波谷的图像。该方法中的凹面与直方图中的谷底和“肩”相对应，且在直方图的“肩”所处的位置上可以确定一个最佳阈值。但是由于原始灰度直方图往往会存在许多毛刺，使得该方法也不能准确定位直方图的“肩”所处的位置，从而导致产生错误的分割结果。

2. 最小误差法

最小误差法的基本思想是将灰度直方图视为对其先验概率 p_i 的正态分布的概率密度函数 $p(f)$ 的估计，并假定图像中的目标和背景的灰度分布的均值是 μ_i ，方差是 σ_i 。 $p(f)$ 的表达式见式 (4-2)，且最佳阈值可通过式 (4-3) 求得。

$$p(f) = \sum_{i=1}^2 p_i p(f|i) \tag{4-2}$$

$$\frac{(f-\mu_1)^2}{2\sigma_1^2} + \ln\sigma_1^2 - 2\ln p_1 = \frac{(f-\mu_2)^2}{2\sigma_2^2} + \ln\sigma_2^2 - 2\ln p_2 \tag{4-3}$$

用此方法计算图像的阈值，需要知道统计参量 σ_i 、 μ_i 以及 p_i 的值，但是这些值均是未知的，因此引入准则函数 $J(m)$ ，通过最小化准则函数 $J(m)$ 求得最佳阈值 T 。其中 $J(m)$ 的表达式见式 (4-4)， $p_i(m)$ 的表达式见式 (4-5)， $\mu_i(m)$ 的表达式见式 (4-6)， $\sigma_i(m)$ 的表达式见式 (4-7) 以及 T 的表达式见式 (4-8)。

$$J(m) = 1 + 2 \{ p_1(m) \ln \sigma_1^{(m)} + p_2(m) \ln \sigma_2^{(m)} \} - 2 \{ p_1(m) \ln p_1^{(m)} + p_2(m) \ln p_2^{(m)} \} \tag{4-4}$$

$$p_i(m) = \begin{cases} p_1(m) = \sum_{f=0}^m h(f), & i = 1 \\ p_2(m) = \sum_{f=m+1}^{L-1} h(f), & i = 2 \end{cases} \quad (4-5)$$

$$\mu_i(m) = \begin{cases} \mu_1(m) = \frac{\left\{ \sum_{f=0}^m h(f)f \right\}}{p_1(m)}, & i = 1 \\ \mu_2(m) = \frac{\left\{ \sum_{f=m+1}^{L-1} h(f)f \right\}}{p_2(m)}, & i = 2 \end{cases} \quad (4-6)$$

$$\sigma_i(m) = \begin{cases} \sigma_1(m) = \frac{\left\{ \sum_{f=0}^m h(f) [f - \mu_1(m)]^2 \right\}}{p_1(m)}, & i = 1 \\ \sigma_2(m) = \frac{\left\{ \sum_{f=m+1}^{L-1} h(f) [f - \mu_2(m)]^2 \right\}}{p_2(m)}, & i = 2 \end{cases} \quad (4-7)$$

$$T = \text{Arg Min}_{m \in G} J(m) \quad (4-8)$$

但是当图像中有多个目标和背景时， σ_i 、 μ_i 以及 p_i 的个数增多，计算量及计算难度都有所增加，因此该方法具有一定的局限性。

3. 矩量保持法

矩量保持法又称矩守恒阈值法，即最优阈值在待分割图像和分割后的图像中的矩量保持恒定不变。若待分割图像的像素点的灰度值 $r \in \{0, 1, 2, 3, \dots, L-1\}$ ，且 L 代表灰度级总数， $M_i (i \in \{0, 1, 2, 3, \dots, L-1\})$ 代表灰度级 i 在待分割图像中出现的次数； p_i 代表灰度级 i 出现的频率，其表达式见式(4-9)， n_k 代表第 k 阶矩，其表达式见式 (4-10)。在此方法中，阈值分割应当满足前三阶矩恒定不变，其矩量保持方程见式 (4-11)，其中 W_0 、 W_1 分别代表阈值化后的每一个种别的灰度值， p_0 、 p_1 表示阈值化后的灰度值的分布概率，通过求解方程组式 (4-11)，可以求出 p_0 以及参数 c_0 、 c_1 、 d ，且其结果见式 (4-12)，此时求出的 p_0 即为 p_0 的分位数，也就是待分割图像的最优阈值 T 。

$$p_i = \frac{M_i}{M} \quad (4-9)$$

$$n_k = \begin{cases} n_0 = 1 \\ n_k = \sum_{i=0}^{L-1} p_i i^k, & k = 2, 3, \dots \end{cases} \quad (4-10)$$

$$\begin{cases} p_0 W_0^0 + p_1 W_1^0 = n_0 \\ p_0 W_0^1 + p_1 W_1^1 = n_1 \\ p_0 W_0^2 + p_1 W_1^2 = n_2 \\ p_0 W_0^3 + p_1 W_1^3 = n_3 \end{cases} \quad (4-11)$$

$$\begin{cases} c_0 = \frac{n_1 n_3 - n_2^2}{n_2 - n_1^2} \\ c_1 = \frac{n_1 n_2 - n_3^2}{n_2 - n_1^2} \\ d = \left((c_1^2 - 4c_0)^{\frac{1}{2}} - c_1 \right) / 2 \\ p_0 = \frac{d - n_1}{(c_1^2 - 4c_0)^{\frac{1}{2}}} \end{cases} \quad (4-12)$$

此方法是以理想状态下假设分割前与分割后的图像在矩量上保持恒定不变为前提的，但一般情况下由于图像受外界或自身复杂因素的影响，图像在分割后矩量均会有所变化，因此该方法不具有普遍性。

4. 最大熵方法

最大熵方法是由 Pun 最早使用香农关于熵的概念提出的。该方法以图像完全由其灰度分布直方图表示为前提，用图像的熵作为度量实现对图像中的目标进行分割的目的。熵代表一个系统的信息量，记作 H ，其定义式见式 (4-13)，其中 p_r 代表符号 r 出现的概率。因此对应定义图像的熵， p_i 代表图像中灰度级 i 出现的概率，其图像熵的表达式见式 (4-14)。

$$H = - \sum p_r \log p_r \quad (4-13)$$

$$H = - \sum_{g=0}^{255} p_i \log p_i \quad (4-14)$$

图像阈值分割中应用熵的概念，其实质是通过研究灰度直方图熵的测量，从而寻找到图像的最优阈值，若熵的定义角度不同，那么就会有不同的最大熵方法。例如，如果考虑图像中灰度的空间分布，并通过像素的灰度值与其邻接像素的灰度均值来计算直方图，那么就会出现二维熵的阈值分割方法。较常用的最大熵方法有一维最大熵阈值法和二维最大熵阈值法。

一维最大熵阈值法的基本原理是假设灰度级 T 分割图像，将大于 T 的像素点的集合定义为背景 N ，且将小于 T 的像素点的集合定义为目标 M ；接着利用图像中目标和背景的灰度级的熵来计算图像的最大熵 H ，其表达式见式 (4-15)，其中， H_N 代表图像中背景的熵，其表达式见式 (4-16)； H_M 代表图像中目标的熵，其表达式见式 (4-17)； p_N 代表图像中背景所占面积的比例，其表达式见式

(4-18); p_M 代表图像中目标所占面积的比例, 其表达式见式 (4-19); p_i 代表图像中灰度级 i 出现的概率。

$$H = H_M + H_N \quad (4-15)$$

$$H_N = - \sum_{i=r+1} p_N \log p_N \quad i = r+1, r+2, r+3, \dots, l-1 \quad (4-16)$$

$$H_M = - \sum_1 p_M \log p_M \quad i = 1, 2, 3, \dots, r \quad (4-17)$$

$$p_N = \frac{p_i}{1 - \sum_{i=0} p_i} \quad i = r+1, r+2, r+3, \dots, l-1 \quad (4-18)$$

$$p_M = \frac{p_i}{\sum_{i=0} p_i} \quad i = 1, 2, 3, \dots, r \quad (4-19)$$

二维最大熵阈值法的基本原理与一维最大熵阈值法不同的是计算图像中灰度级的概率是用二维坐标表示的。例如, p_{ij} 代表图像中像素点的灰度均值对 (i, j) 的概率, 即图像中点的灰度为 i , 其区域灰度均值通过计算灰度为 i 的各个像素点及其八邻域像素点所组成区域的均值为 j 的概率。综上所述, 当图像中的像素点变得相当多时, 最大熵方法的计算量及计算难度都会有所增加, 因此该方法也具有一定的局限性。

5. 直方图变换法

直方图变换法的基本思想是依照图像中像素点的部分特征, 将像素点加权后取得新的直方图而间接求取原图像的最佳阈值。因此, 该方法取得最佳阈值的关键是要选取合适的加权方法以及最佳的权值。

6. 概率松弛方法

概率松弛法的基本原理是假设图像是由较暗的目标和较亮的背景这两部分组成, 并且用概率表示该图像中像素点及像素点之间的空间相关性, 然后选择一个合适的阈值调整每个像素的概率, 使得这个像素点只属于暗的目标区域或亮的背景区域, 从而通过多次迭代使得图像中的目标与背景被明显地分割出来。该方法原理简单, 但仅仅适用于待分割图像中的目标与背景灰度级分别处于直方图前后两半的情况。

7. 灰度分布统计方法

灰度分布统计方法中一种典型的方法是灰度共生矩阵法。灰度共生矩阵中的元素是两个不同的灰度级在其相邻位置处出现的概率。该方法对应两个直方图, 并且这两个直方图分别与灰度共生矩阵中的近对角线的元素、远对角线的元素相对应, 那么该方法的最佳阈值就是在这两个直方图对应的波谷与波峰重叠的区域中获得。该方法在两个直方图对应的波谷与波峰不能重叠时失效, 因而不能寻找到最佳阈值而无法完成图像分割的目的。

4.3.2 局部阈值法

全局阈值法在待分割图像中的目标和背景比较清晰时，可以取得较好的分割效果，但当图像中的目标和背景比较模糊、背景分布不匀称、目标残缺或目标灰度转变率很大的情况下，该方法就不适用了。这种情况下就需要采用局部阈值法将待分割图像分为几个子图，然后对这些子图分别采用全局阈值法依次求出最佳分割阈值进行图像分割，这样就可以完成对不具有统一阈值的图像进行分割操作。

4.3.3 动态阈值法

对于那些目标和背景比较模糊、背景不均匀、目标残缺或目标灰度变化率较大的图像，要获得一个适用于整幅图像的阈值，须采用动态阈值法。该方法根据待分割图像每个像素及其邻域像素灰度值的状况，动态地计算出图像的最佳阈值。常用的动态阈值法包括：Kaneko 和 Chow 提出的方法、Brunckstein 和 Yanowitz 提出的方法以及 Pietikainen 和 Sauvola 提出的方法。

综上所述，常用的阈值分割方法总结如下：全局阈值记作 $T = T[p(x, y)]$ ，只与图像中所有像素点的灰度值有关；局部阈值记作 $T = T[p(x, y), q(x, y)]$ ，不仅与图像区域内的各像素点的值有关，而且还与各个像素相邻像素点的值有关；动态阈值记作 $T = T[x, y, p(x, y), q(x, y)]$ ，不仅与图像中像素的空间坐标有关，而且还与该像素点的局部区域和本身性质有关。

4.4 改进的最大类间方差法

最大类间方差法（OTSU 法）是一种典型的无参数、无监督的自动确定阈值的阈值分割方法。早在 1979 年，该算法由日本学者大津展之提出，是在最小二乘法的基础上被推导出求出的。该方法无须预先知道其他先验知识，具有算法简便、运算速度快、目标和背景分离性较好以及拥有自适应阈值等优点，从而该方法被广泛地应用到图像分割中。

最大类间方差法的基本原理是依据待分割图像中像素的灰度特点，将图像分成目标和背景两种类型；若图像中的目标与背景类间方差大，则表示目标和背景之间的差别大，因此如果当图像中的背景信息被误判为目标信息，或者图像中的目标信息被错判为背景信息时，该图像中的目标与背景类间方差就会取很小的值。由于这一特性，该方法只要满足图像中像素类间方差最大或类内方差最小，就意味着最小的误差，进而可以取得最优阈值，从而对图像进行精确分割。

最大类间方差法的思想如下：

1) 一维最大类间方差法的基本原理：假设待分割图像为 h ，该图像的灰度

级共 L , n_j 代表图像中灰度值是 j 的像素点的个数, N 代表图像像素的总个数, 其表达式见式 (4-20); 图像中各像素点的灰度值的概率是 p_j , 其表达式见式 (4-21)。接着, 该方法自动确定阈值为 T , 将图像中的像素按照灰度级分为以下两组像素点的集合 D_0 和 D_1 , 其表达式分别见式 (4-22) 和式 (4-23); 像素点集合 D_0 的概率为 M_0 (或表示为 $M(T)$), 其表达式见式 (4-24), 像素点集合 D_1 的概率为 M_1 , 其表达式见式 (4-25); D_0 平均值为 μ_0 , 其表达式见式 (4-26), D_1 平均值为 μ_1 , 其表达式见式 (4-27); 整个图像的灰度平均值为 μ , 其表达式见式 (4-28); $\mu(T)$ 代表当图像的阈值为 T 时, 整个图像灰度值的平均值, 其表达式见式 (4-29); 所有采样的灰度值的平均值为 μ' , 其表达式见式 (4-30); $\delta^2(T)$ 代表两组像素点集合 D_0 与 D_1 间的方差, 称其为阈值选取函数, 其表达式见式 (4-31), 并且求解式 (4-31) 中 $\delta^2(T)$ 的最大值, 记作 $\max \delta^2(T)$, 这个值就是该方法的最佳阈值。

$$N = \sum_{j=1}^L n_j \quad (4-20)$$

$$p_j = \frac{n_j}{N} \quad (4-21)$$

$$D_0 = \{1, 2, \dots, T\} \quad (4-22)$$

$$D_1 = \{T+1, T+2, \dots, L\} \quad (4-23)$$

$$M_0 = \sum_{j=1}^T p_j = M(T) \quad (4-24)$$

$$M_1 = \sum_{j=T+1}^L p_j = 1 - M_0 \quad (4-25)$$

$$\mu_0 = \sum_{j=1}^T \frac{j p_j}{M_0} = \frac{\mu(T)}{M(T)} \quad (4-26)$$

$$\mu_1 = \sum_{j=T+1}^L \frac{j p_j}{M_1} = \frac{\mu - \mu(T)}{1 - M(T)} \quad (4-27)$$

$$\mu = \sum_{j=1}^L j p_j \quad (4-28)$$

$$\mu(T) = \sum_{j=1}^T j p_j \quad (4-29)$$

$$\mu' = M_0 \mu_0 + M_1 \mu_1 \quad (4-30)$$

$$\delta^2(T) = M_0(\mu_0 - \mu)^2 + M_1(\mu_1 - \mu)^2 = M_0 M_1 (\mu_1 - \mu_0)^2 = \frac{[\mu M(T) - \mu(T)]^2}{M(T)[1 - M(T)]} \quad (4-31)$$

2) 二维最大类间方差法的基本原理: 假设待分割图像为 $h(x, y)$, 图像中像素点的坐标是 (x, y) , 图像的灰度级共 L , $f(x, y)$ 代表 $k \times k$ 邻域的平均灰度值, 其表达式见式 (4-32); 该方法用图像的灰度值 $h(x, y)$ 和 $k \times k$ 邻域的

平均灰度值 $f(x, y)$ 来表示图像，记作 $h_{i,j}$ ， $h_{i,j} = (h(x, y), f(x, y))$ ，且代表像素点坐标为 (i, j) ，该像素点的灰度值为 $h(x, y)$ ，其邻域的平均灰度值为 $f(x, y)$ ； $p_{i,j}$ 代表点 (i, j) 的频率，其表达式见式 (4-33)。

$$f(x, y) = \frac{1}{k^2} \sum_{m=-\frac{k}{2}}^{\frac{k}{2}} \sum_{n=-\frac{k}{2}}^{\frac{k}{2}} h(s+m, y+n) \quad (4-32)$$

$$p_{i,j} = \frac{h_{i,j}}{M \times N} \quad (4-33)$$

二维最大类间方差法的二维灰度直方图如图 4-2 所示。图中横、纵坐标代表二维最大类间方差法的分割阈值 (s, t) ，直方图被分为 A、B、C、D 四个区域；图像中的目标和背景主要集中在 B、C 这两个区域，而对于 A、D 这两个区域几乎没有图像的目标和背景，因此在这两个区域目标和背景的分布概率几乎接近于 0。像素点集合 D_0 和 D_1 分别代表直方图中的目标和背景，即在直方图中的 B、C 两个区域，且它们的概率分别为 $p(D_0)$ 、 $p(D_1)$ ，其表达式分别见式 (4-34) 和式 (4-35)；像素点集合 D_0 和 D_1 的灰度均值矢量分别为 μ_0 、 μ_1 ，其表达式分别见式 (4-36) 和式 (4-37)；待分割图像总的灰度均值矢量是 μ_T ，其表达式见式 (4-38)；该直方图的离散矩阵为 S_E ，其表达式见式 (4-39)；离散矩阵 S_E 的迹是 $\text{tr}(S_E)$ ，且当离散矩阵 S_E 的迹 $\text{tr}(S_E)$ 取最大值时，此时就可以取得该图像的最佳阈值 (S, T) ，这个最佳阈值记作 $\text{tr}(S_E(S, T))$ ， $\text{tr}(S_E(S, T))$ 的表达式见式 (4-40)。

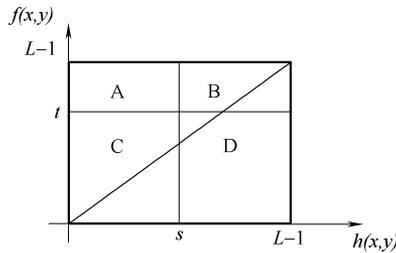


图 4-2 二维灰度直方图

$$p(D_0) = \sum_{i=0}^s \sum_{j=0}^t p_{i,j} = p_0(s, t) \quad (4-34)$$

$$p(D_1) = \sum_{i=s+1}^{L-1} \sum_{j=t+1}^{L-1} p_{i,j} = p_1(s, t) \quad (4-35)$$

$$\mu_0 = (\mu_{0,i}, \mu_{0,j})^T = \begin{bmatrix} \sum_{i=0}^s ip(i|D_0) \\ \sum_{j=0}^t jp(j|D_0) \end{bmatrix} \quad (4-36)$$

$$\mu_1 = (\mu_{1,i}, \mu_{1,j})^T = \begin{bmatrix} \sum_{i=s+1}^{L-1} ip(i | D_1) \\ \sum_{j=t+1}^{L-1} jp(i | D_1) \end{bmatrix} \quad (4-37)$$

$$\mu_T = (\mu_{T,i}, \mu_{T,j})^T = \begin{bmatrix} \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} ip_{i,j} \\ \sum_{j=0}^{L-1} \sum_{i=0}^{L-1} jp_{i,j} \end{bmatrix} \quad (4-38)$$

$$S_E = \sum_{k=0}^1 p(D_k) [(\mu_k - \mu_T)(\mu_k - \mu_T)^T] \quad (4-39)$$

$$\text{tr}(S_E) = p_0 [(\mu_{0,i} - \mu_{T,i})^2 + (\mu_{0,j} - \mu_{T,j})^2] + p_1 [(\mu_{1,i} - \mu_{T,i})^2 + (\mu_{1,j} - \mu_{T,j})^2] \quad (4-40)$$

由于待分割的多源图像中的目标和背景有时差别不是很大，背景呈不均匀分布，此时采用最大类间方差法分割多源图像，不仅使多源图像的分割阈值难以把握，而且还会造成图像的边缘信息残缺，因此，针对此缺陷本章对最大类间方差法进行了改进。由于大量的实验已证明，图像经过梯度运算转为梯度图像后可以保存图像的边缘信息，同时去除图像中的局部极小值及干扰噪声，从而使得图像的目标和背景更为突出，使得图像的直方图具有更加明显的双峰，因而获得最优阈值，并且结合了梯度特征的阈值分割能够使多源图像取得较好的分割效果，因此本章使用梯度算法与最大类间方差法结合的方法来对多源图像进行分割。

结合梯度特征的最大类间方差法的步骤：首先将多源图像梯度化获取梯度信息，保存图像完整的边缘信息，从而为准确寻找最优阈值奠定基础；然后再使用最大类间方差法对多源图像进行分割。图像梯度化通常使用一阶差分表示梯度，且常用的计算一阶差分的方法包含：Prewitt 边缘检测算子、Roberts 边缘检测算子以及 Sobel 边缘检测算子等；计算梯度值通常使用图像的卷积和图像小区域模板，Kirsch 边缘检测算子是利用卷积和图像小区域模板计算梯度值的典型方法之一。

Prewitt 边缘检测算子计算图像 $h(x, y)$ 对 x, y 的一阶差分，其中 h_0, h_1, \dots, h_9 是以 h_5 为中心的各个像素点，则 x, y 的一阶差分分别表示为 G_x, G_y ，表达式如下：

$$G_x = (h_7 + h_8 + h_9) - (h_1 + h_2 + h_3) \quad (4-41)$$

$$G_y = (h_3 + h_6 + h_9) - (h_1 + h_4 + h_7) \quad (4-42)$$

Roberts 边缘检测算子计算图像 $h(x, y)$ 对 x, y 的一阶差分，其中 h_0, h_1, \dots, h_9 是以 h_5 为中心的各个像素点，则 x, y 的一阶差分分别表示为 $G_x,$

G_y , 表达式如下:

$$G_x = h_9 - h_5 \quad (4-43)$$

$$G_y = h_8 - h_6 \quad (4-44)$$

Sobel 边缘检测算子计算图像 $h(x, y)$ 对 x, y 的一阶差分, 其中 h_0, h_1, \dots, h_9 是以 h_5 为中心的各个像素点, 则 x, y 的一阶差分分别表示为 G_x, G_y , 表达式如下:

$$G_x = (h_7 + 2h_8 + h_9) - (h_1 + 2h_2 + h_3) \quad (4-45)$$

$$G_y = (h_3 + 2h_6 + h_9) - (h_1 + 2h_4 + h_7) \quad (4-46)$$

Kirsch 边缘检测算子由八个代表八个方向的模板组成算子, 利用该算子进行卷积, 即可得到八个方向中的最大值及其该方向所对应的方向数值, 从而得到由边缘方向及强度表示的具有梯度特征的图像。

若图像为 $h(x, y)$, 且该图像的灰度取值范围是 $G \in (0, L-1)$, 则图像 $h(x, y)$ 的梯度为 $\nabla h(x, y)$, 其表达式见式 (4-47); $\nabla h(x, y)$ 的幅度值为 $\text{mag}(h(x, y))$, 其表达式见式 (4-48); $\nabla h(x, y)$ 的相位为 $\theta(x, y)$, 其表达式见式 (4-49)。

$$\nabla h(x, y) = [G_x, G_y]^T = \left[\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y} \right]^T \quad (4-47)$$

$$\text{mag}(h(x, y)) = (G_x^2 + G_y^2)^{\frac{1}{2}} \quad (4-48)$$

$$\theta(x, y) = \arctan\left(\frac{\partial h / \partial y}{\partial h / \partial x}\right) \quad (4-49)$$

本章采用形态学开闭运算对多源图像梯度化, 是因为形态学梯度运算具有非线性, 与 Prewitt 边缘检测算子、Roberts 边缘检测算子以及 Sobel 边缘检测算子等常用的梯度化方法相比, 更能增强图形中极值的对比度, 使图像中的目标与背景的对比度增加, 从而使目标和背景的分度变得明显; 其次, 该方法也具有有良好的保持图像中重要轮廓区域的效果, 从而保存了图像完整的边缘信息, 为准确寻找最佳阈值打下良好的基础。图 4-3 为原测试的多源图像, 图 4-4 为最大类间方差法的分割图像, 图 4-5 为改进的最大类间方差法即结合梯度特征的最大类间方差法的分割图像。



图 4-3 测试图



图 4-4 最大类间方差法分割图



图 4-5 改进的最大类间方差法分割图

通过对比图 4-4 与图 4-5，可知直接采用最大类间方差法，由于多源图像中的背景有时比较复杂，背景呈不均匀分布，使得多源图像的分割阈值难以寻找到，甚至出现图 4-4 中大量的与目标无关的极小值区域，分割效果远差于图 4-5 的结合梯度特征的最大类间方差法的分割效果。

因此，本章采用结合梯度特征的最大类间方差法对分水岭算法进行改进，形成结合梯度特征的最大类间方差法的分水岭算法，从而得到了具有一定意义的边界特征的分割图像，解决了传统分水岭算法中虚假的局部极小值过多与过分割问题，而且使分水岭算法中的分水点较准确地定位，因此改善了传统分水岭算法中易产生过分割的缺陷。选取图像去噪、平滑、滤波预处理后的图像作为输入图像，对其采用结合梯度特征的最大类间方差法进行分割，其分割结果分别如图 4-6 和图 4-7 所示，其中图 4-6 为可见光图像阈值分割图像，图 4-7 为红外图像阈值分割图像；之后再对阈值分割后的图像进行分水岭分割，其分割结果分别如图 4-8 和图 4-9 所示，其中图 4-8 为可见光图像结合阈值算法的分水岭分割图像，图 4-9 为红外图像结合阈值算法的分水岭分割图像。



图 4-6 可见光图像阈值分割图像

上述实验结果可以看出，结合阈值算法的分水岭算法比只使用分水岭算法的分割效果要好，该算法消除了易造成过多区域的极小值，使分水岭算法的分水点较准确地定位，同时去除了大部分的复杂背景，并且也得到了具有一

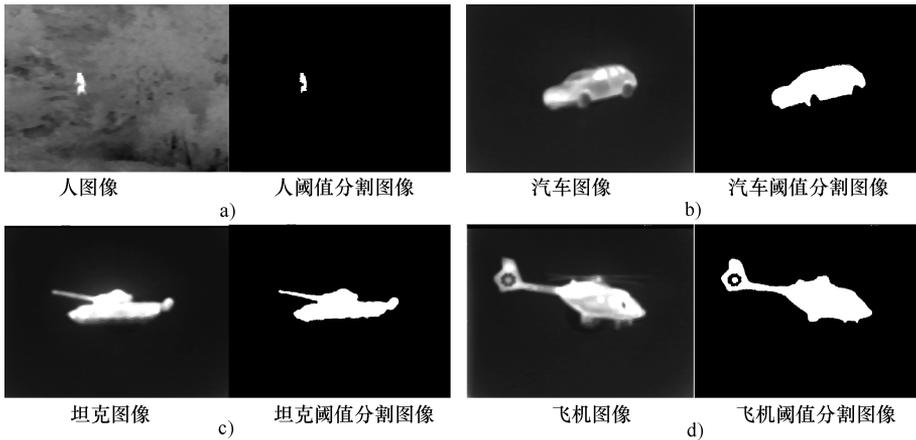


图 4-7 红外图像阈值分割图像

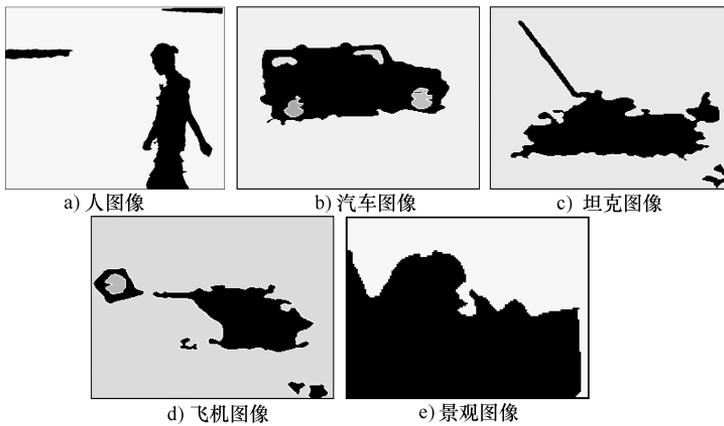


图 4-8 可见光图像结合阈值分割算法的分水岭分割图像

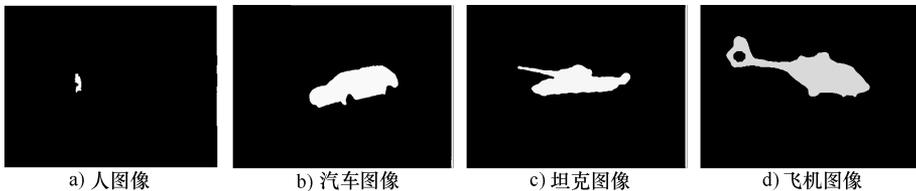


图 4-9 红外图像结合阈值分割算法的分水岭分割图像

定意义的边界特征的分割图像，从而改善了传统分水岭算法中易产生过分割的缺陷。

4.5 本章小结

本章主要研究了多源图像的阈值分割的方法，首先研究了阈值分割的基本思想、基本原理以及实现阈值分割算法的基本步骤，从而为寻找合适的阈值分割方法奠定了理论基础；其次，通过分析比较常用的阈值分割方法，从而为寻找最佳阈值分割方法提供了一定的思路；再次，主要研究了最大类间方差法，并对该方法进行改进，提出了结合梯度特征的最大类间方差法，并使用改进的最大类间方差法对分水岭算法进行改进，因而又提出了结合阈值算法的分水岭分割方法，并且将这两种改进的算法分别在 VC++6.0 开发环境下进行了实验，并对实验结果进行了分析。

结合聚类分割的分水岭算法

本书已采用结合阈值分割的分水岭算法对多源图像进行阈值优化操作，消除了易造成过多区域的极小值，虽然已经消除了大多数的过分割区域，但由于分水岭算法易受图像中的量化误差的影响，在水分岭算法结束后仍有一些过分割区域，因此为了得到良好的分割效果，本章在结合阈值分割的分水岭算法后又采用聚类分割算法合并那些无语义学意义的过分割的小区域，从而进一步改善了传统分水岭算法中易产生过分割的缺陷，并能获得更有意义的分割效果。

5.1 图像聚类分割概述

聚类是指依照某一个准则将数据集划分为某几个类或簇，使得属于同一类内的数据集合具有较高的相似度，而属于不同类的数据集合具有较低的相似度，因而聚类过程的关键就是尽可能地将同类事物聚集在一起，将不同类别的数据集合尽可能地分离。聚类分析属于多元统计方法中的一种，在样本进行聚类分析的时候，在样本所属的类别和类别数目未知的情况下，该方法依据样本数据，采用数学方法来处理数据集的分类问题。聚类分析在图像处理领域，尤其在图像分割方面发挥着相当重要的作用，因而产生了许多基于聚类算法的图像分割方法。

聚类的定义：由许多数据点组成的数据集 X ，此数据集是一个 $N \times d$ 的矩阵，且 X 中包含的数据点 x_j 在数据空间 B 中，并且表示为 $x_j = (x_{j1}, x_{j2}, \dots, x_{jd}) \in B$ ， x_j 的属性定义为数值型或枚举型的 x_{ji} ；若数据集 X 包括 N 个 $x_j (j=1, 2, \dots, N)$ 数据点作为其对象，然后把数据集 X 分裂为 k 个分割 $D_m (m=1, 2, \dots, k)$ 和噪声 D_n ，并且当数据集 X 是噪声 D_n 与没有交集的多个分割 $D_m (m=1, 2, \dots, k)$ 并集的时候，完成聚类算法，其表达式如下：

$$\begin{cases} X = D_1 \cup D_2 \cup \dots \cup D_k \cup D_n \\ D_i \cap D_j = \emptyset (i \neq j) \end{cases} \quad (5-1)$$

数据类型包括：数据聚类矩阵和相异矩阵。数据矩阵是目标与变量组织的数据表达方式，若聚类中有 N 个对象 $x_j (j=1, 2, \dots, N)$ ，且每个对象选择 P

个变量，则第 j 个对象的第 i 个变量的观测值表示为 x_{ji} ，因此，聚类中有 N 个对象的所有观测值组成的 $N \times P$ 的矩阵，即为此聚类过程的数据矩阵，其表达式见式 (5-2)，且第 j 个对象的 P 个变量的观测值表示为向量 x_j ，其表达式见式 (5-3)；相异度矩阵是目标与其结构的数据表达方式，若聚类数据集中的对象 i 与 j 之间的相异性用一个非负数值 $e(i, j)$ 量化表示，则聚类数据集存储的 N 个对象的相异度矩阵，表示为一个 $N \times N$ 的矩阵，见式 (5-4)。从式 (5-4) 可以看出， $e(i, j) = e(j, i)$ 且 $e(i, i) = 0$ ；当对象 i 与 j 之间越相近或相似时， $e(i, j)$ 就趋近于零，反之当对象 i 与 j 之间相距越远或不相同时， $e(i, j)$ 的取值就越大，因此相异度矩阵可以表示出聚类数据集存储的 N 个对象间的相似或相异程度。

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1P} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2P} \\ x_{31} & x_{32} & x_{33} & \cdots & x_{3P} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ x_{N1} & x_{N2} & x_{N3} & \cdots & x_{NP} \end{bmatrix} \quad (5-2)$$

$$x_j = (x_{j1}, x_{j2}, x_{j3}, \cdots, x_{jP})^T \quad (5-3)$$

$$\begin{bmatrix} 0 & & & & \\ e(2,1) & 0 & & & \\ e(3,1) & e(3,2) & 0 & & \\ \vdots & \vdots & \vdots & \ddots & \\ e(N,1) & e(N,2) & \cdots & \cdots & 0 \end{bmatrix} \quad (5-4)$$

完整的图像聚类过程不仅包含聚类算法本身，还包括图像的特征选择与提取以及数据集的相似度度量的计算，其图像聚类过程可用图 5-1 表示，从图 5-1 中可以反映出图像聚类过程的前两步：图像特征的选择与提取以及数据集相似度的计算会受到聚类输出反馈的影响。



图 5-1 图像聚类过程

到目前为止，评价聚类方法的优劣还没有量化的客观标准，因而聚类方法效果的好与差主要采用以下几个标准来衡量：是否具有处理大量数据集的能力；是否具有处理数据抗噪声的能力；是否具有处理携带间隔或嵌套的任意类型数据的能力；是否具有处理后的输出结果与数据输入的顺序无关的能力；是否具有处理多维数据的能力；在聚类过程中是否需要先验知识。

5.2 常用的图像聚类分割算法

在图像分割中，根据图像中要处理的数据、分割的目的以及用途，可选取不同的聚类算法以实现图像分割的目的，目前常用于图像分割的聚类算法大体上可分为划分聚类算法、层次聚类算法、基于密度的聚类算法、基于模型的聚类算法以及基于网格的聚类算法。

5.2.1 划分聚类算法

划分聚类算法采用目标函数最小化策略把一个确定的 N 个数据对象的数据集分成 k 个组，并且该算法使得每一组中的对象相似度相当高，而不同组的对象相似度比较低，由此可知相似度的定义是划分聚类算法的关键环节。该算法的目标函数一般定义为式 (5-5)，其中 x_i 表示对象空间中一个数据对象，且 \bar{x}_i 是第 i 类的均值， J 为集合 A 中全部对象与对应的聚类中心的均方差之和。最常用的划分聚类算法包含 k -means 算法和 k -medoids 算法。

$$J = \frac{1}{N} \sum_{i=1}^k \sum_{n=1}^N A_n d^2(\bar{x}_i - x_n) \quad (5-5)$$

1. k -means 算法

k -means 算法的原理是将恒定的 N 个数据目标的数据集分成事先给定的数目为 k 的簇，首先随机选择 k 个对象作为初始的 k 个聚类中心 $C = (C_1, C_2, \dots, C_k)$ ，接着通过算出其余的所有样本到各自聚类中心的距离，把该样本划分到距离它很近的类中，之后再使用平均值的方法计算调整后的新类的聚类中心，重复上述步骤直到计算出的两次类中心保持不变时，标志着数据集中的样本分类结束且聚类平均误差准则函数 F 处于收敛状态，聚类平均误差准则函数 F 的表达式见式 (5-6)，其中 q 为数据集的数据对象， m_i 是第 i 类聚类中心 C_i 的平均值， F 代表数据集中全部数据对象的平方误差的总和。

$$F = \sum_{i=1}^k \sum_{q \in C_i} |q - m_i|^2 \quad (5-6)$$

k -means 算法虽然容易实现，但是该算法也具有一些缺陷： k -means 算法当选用不一样的初始值时，能够得到不相同的聚类结果，因此该算法对初始聚类中心的依赖性较大； k -means 算法对独立点及噪声点反应较为敏锐，严重时会导致聚类中心的偏离； k -means 算法需要预先掌握的知识来求出待生成的簇的数目。

2. k -medoids 算法

k -medoids 算法的处理过程如下：先随意选出 k 个对象作为初始的 k 个聚类的代表点，接着算出其余的样本到其最靠近聚类中心的对象的距离，把该样本

归类到离它最近的聚类中，并依据某代价函数估算目标与代表点间的相异度平均值，若对象与代表点相似则替换代表点，反复进行上述过程直到不再有对象替换代表点为止。 k -medoids 算法包括 PAM 算法、CLARANS 算法以及 CLARA 算法。当数据集和簇的数目较大时，PAM 算法的性能就会变得很差；CLARANS 算法不能辨认套嵌或其余繁杂形状的聚类形状，而且该算法具有运算效率低、没有处理高维数据的能力以及不能准确找到局部极小点产生错误的聚类结果的缺陷；CLARA 算法的聚类结果与抽样的样本大小有关，当抽样的样本发生偏差时，CLARA 算法的性能就会变差而不能得到良好的聚类结果。

5.2.2 层次聚类算法

层次聚类算法分为凝聚层次聚类算法和分裂层次聚类算法。凝聚层次聚类认为每一个对象是一个簇，遵循自下而上的原则逐步归并簇，继而构成很大的簇，反复这一过程直到图像中所有的对象都在同一个簇中或满足某一约束条件时，该算法结束。分裂层次聚类的过程与凝聚层次聚类的过程相反，分裂层次聚类认为图像中所有的对象已经在一个簇中，遵循自上而下的原则逐渐将图像中的对象从一个簇中划分为越来越小的簇，反复这一过程直到图像中每一个对象被分为单独的一簇或满足某一约束条件时，该算法结束。经常使用的层次聚类算法有 BIRCH 算法、ROCK 算法、CURE 算法、Chameleon 算法等。BIRCH 算法通过扫描数据来建立一个有关聚类结构的 CF 树，并对此 CF 树的叶节点进行聚类；ROCK 算法根据相似度阈值与共同邻域的基本概念计算出图像的相似度矩阵，接着从图像的相似度矩阵中构建一个稀疏图，并对该稀疏图进行聚类；CURE 算法依据收缩因子的值调整每个簇的大小和形状，从而形成不同类型的簇而完成聚类；Chameleon 算法依据若图像中存在两个簇之间相似性以及互联性高度相关的对象，则动态地合并这两个簇，重复上述过程直到不能合并为止，该算法结束。

层次聚类算法虽然易处理不同粒度水平上的数据，但是该算法的结束条件模糊；其扩展性不良，因而要求预先算出图像中大部分的簇才可完成合并或分裂操作；并且该算法的归并或割裂簇的处理是不可修正的，因此该算法聚类质量较低。

5.2.3 基于密度的聚类算法

基于密度的聚类算法弥补了划分聚类算法和层次聚类算法的不足，不仅可以处理凸形簇的聚类，而且可以处理任意形状簇的聚类，该算法依据图像中数据集密度的相似度，把密度相近的数据集划分为一个簇，反之，把密度不接近的数据集划分为不同的簇，从而完成聚类的目的。常用的基于密度的聚类算法有 DBSCAN 算法、DENCLUE 算法以及 OPTICS 算法等。DBSCAN 算法通过图像

中对对象集合的每个对象的特定邻域来确定簇的区域，该算法的聚类结果不受数据输入顺序的影响，但此算法在执行的时候，需要事先知道图像中确定输入的聚类参数，但由于现实中的高维数据集不容易确定出聚类参数，因此该方法具有一定的局限性；DENCLUE 算法依据图像中数据集的影响函数来计算数据空间的整体密度，接着确定出密度吸引点并寻找到确定的各个簇的区域而完成聚类的目的，但是该算法受聚类参数的影响较大，往往参数值的轻微变化会引发差别较大的聚类结果；OPTICS 算法可以自动、交互地算出图像中簇的次序，并且此次序表示数据集的聚类结构，但是由于该算法所确定的聚类结构是从一个宽泛的参数所设置的范围中所获得，因此该算法不能产生一个数据集的合簇，因而聚类结果不太理想。

5.2.4 基于模型的聚类算法

基于模型的聚类算法依据图像中的数据集符合某一概率分布这一假设，把数据集表示为某一数学模型来实现聚类的目的，因而该方法划分的每一个簇的形式均是通过概率描述来表示的。常用的基于模型的聚类算法有统计方法和神经网络方法，此外还有一些新的模型聚类算法，例如，支持矢量方法的聚类算法、SPC 算法以及 SyMP 算法等。

统计聚类方法有 COBWEB 算法、CLASSIT 算法、AutoClass 算法以及高斯混合模型算法等。COBWEB 算法是最著名的基于统计聚类的方法，该算法用一个启发式估算度量将数据集中的对象加入到能够产生最高分类效果分类树的位置，于是会不断地创建出新的类，从而完成聚类的目的。COBWEB 算法不需要事先提供数据集的聚类参数就可以自动地修正并划分出数据集的簇的数目，但是由于该方法进行的前提是假设每个簇的概率分布是相互独立的，因而该方法具有局限性；此外该方法在存储和更新数据集的每个簇的概率分布的时候，均会付出较高的代价而效率变低。CLASSIT 算法可以处理连续性数据集的增量的聚类，并且该算法是 COBWEB 算法的一个衍生算法，因而该算法存在与 COBWEB 算法相同的缺陷，因此该算法也不适用于解决大型数据集的聚类问题。

神经网络方法将数据集的每一个簇看作是一个例证，并将该例证视为聚类的初始点，接着该算法依据某种相似度，将新的对象分配到与其最相似的簇中而完成聚类的目的。主要的神经网络方法包含竞争学习神经网络方法和自组织特征映射神经网络方法。基于神经网络的聚类算法处理数据需要的时间较长，并且不适合将其用于大型数据集的处理。

5.2.5 基于网格的聚类算法

基于网格的聚类算法是首先将图像空间数据量化成某些单元，然后该算法对这些量化单元进行聚类。经典的基于网格的聚类算法有 STING 算法、CLIQUE

算法以及 WaveCluster 算法。STING 算法是一种针对不同级别的分辨率将图像空间分为多个级别的长方形单元的多分辨率的聚类方法；CLIQUE 算法是一种综合密度与网格的针对处理高维数据集的聚类算法；WaveCluster 算法采用小波变换把图像的数据集的空间域转变为其频率域，并在这个频率域中找到密集的数据区域而实现数据聚类。基于网格的聚类算法虽然能快速聚类，但是该算法只能对垂直和水平边界执行聚类，不能对斜边界执行聚类，因此该算法具有一定的局限性；其时间复杂程度通常与数据集的规模无关而与网格数目相关，若网格单元数太大，则其时间复杂度就会变大，反之若网格单元数太小，该算法的聚类精确度就会受影响，因此该算法选取恰当的网格数是取得良好的聚类效果的关键环节。

5.3 改进的 FCM 聚类分割算法

依据图像数据集中的簇的重叠程度可以将聚类分割算法分为硬性聚类算法和模糊聚类算法。硬性聚类算法体现了对象与簇之间的非此即彼的关系，即严格地把每个待聚类的对象分类到某个簇中，设图像的数据集为 $X = \{x_1, x_2, \dots, x_n\}$ ，其分类的数目为 c ，且满足条件 ($2 \leq c \leq n$)；假设将数据集 X 划分为互不重叠的 c 个簇，使得任何一个 X 中的样本点均属于某一个簇，并且每一个簇中最少有一个样本点，那么这个硬性聚类算法的聚类结果就可以表示为一个 $c \times n$ 阶的矩阵 U ，则数据集中的第 j 类可以表示为 u_{jk} ，其表达式见式 (5-7)。通常在实践中事物间的边界并不是很分明，就会出现模糊划分的聚类算法，而模糊集合理论又为模糊聚类算法提供了优良的数学工具。模糊聚类算法矩阵 U 中元素的取值并不只局限于 0 与 1 这两个值，而是在 0 与 1 之间的区间上取值，因而此时的硬性聚类算法就转变为模糊聚类算法。

$$u_{jk} = \begin{cases} 1 & x_k \in A_j \quad A_j(j=1, 2, \dots, c) \\ 0 & x_k \notin A_j \quad A_j(j=1, 2, \dots, c) \end{cases} \quad (5-7)$$

硬性聚类算法一般采用最小平方误差和作为其聚类准则，设 $v_j(j=1, 2, \dots, c)$ 是该聚类的聚类中心， d_{jk} 代表图像数据集中的第 j 类样本 x_k 到聚类中心 v_j 的距离，各簇中样本到聚类中心的距离平方和用 $J(U, V)$ 表示，则此硬性聚类算法的目标函数表示为式 (5-8) 或式 (5-9)，并且当用迭代算法来求取最小的 $J(U, V)$ 时，就可以找到最佳的图像矩阵与聚类中心的配对，即 (U, V) ；模糊聚类算法一般采用拉格朗日乘数法求目标函数，设 x_k, v_j 都是 q 维向量， $\forall x_k, v_j \in R^q$ ， A 是 $q \times q$ 的矩阵， $A \in M_{qq}$ ，数据集中的样本 x_k 与聚类中心 v_j 之间距离的平方表示为 $(d_{jk})^2$ ，其表达式见式 (5-10)，其中 T 表示矩阵的转置，且 A 为对称矩阵，当 $A=I$ 时，式 (5-10) 转化为欧式距离，该模糊聚类算法的目标函数的

表达式见式 (5-11)，利用约束条件 $\sum_{j=1}^c u_{jk} = 1, \forall k$ 可求得 $J(U, V)$ 的最小

值，从而找到最佳的 u_{jk} 与 v_j 。

$$J(U, V) = \sum_{j=1}^c \left(\sum_{x_k \in A_j} (d_{jk})^2 \right) \quad (5-8)$$

$$J(U, V) = \sum_{j=1}^c \left(\sum_{x_k \in A_j} u_{jk} (d_{jk})^2 \right) \quad u_{jk} \in \{0, 1\} \quad (5-9)$$

$$(d_{jk})^2 = \|x_k - v_j\|^2 = (x_k - v_j)^T A (x_k - v_j) \quad (5-10)$$

$$J(U, V) = \sum_{j=1}^c \sum_{x_k \in S_j} (u_{jk})^m (d_{jk})^2 \quad (5-11)$$

HCM 聚类算法是典型的硬性聚类算法，该算法收敛速度较快，但由于此算法只能将图像中的样本点以概率 0 和 1 划分到各个簇中，因此该方法不能够充分表现出客观事物的模糊特性，因而在实际的图像分割中往往会产生错误分割。相反，模糊聚类算法中的一种典型算法 FCM 算法就可以解决 HCM 算法中不能充分表达事物模糊性的这一缺陷，该算法在图像分割中利用隶属度函数解决了像素同时属于多个不同类别的可能性问题；该算法能够避免阈值的设定问题，并能解决阈值分割中多个分支的分割问题；该算法可以形成较细致的特征空间，其分割结果不会像硬聚类分割那样产生某种偏差；该算法的聚类过程是自动进行的，是无监督的聚类方法，因而该算法在图像分割领域被广泛应用。

FCM 算法的图像分割的步骤如下：设待分割的图像为 M ，其分割门限为 ∂ ，然后对该图像采用迭代优化方案求其 FCM 算法中目标函数 $J(U, V)$ 的最小值。首先，确定图像的聚类数目 c ($2 \leq c \leq n$) 以及加权指数 m ($m \in [2, \infty]$)；接着，初始化模糊聚类矩阵 U ，其初始值 $U^{(l)} = [\mu_i(x, y)^l]$ ，且 $l=0$ ；然后，依据式 (5-12) 计算各个簇的聚类中心 v_i 并且计算新的模糊聚类矩阵 U^{l+1} ，再依据式 (5-13) 和式 (5-14) 分别计算 $M(x, y)$ 、 $\bar{M}(x, y)$ ；若 $M(x, y) = \emptyset$ ，则满足式 (5-15)，反之，若 $M(x, y) \neq \emptyset$ ，则满足式 (5-16)；最后，检查 $\|U^{(l+1)} - U^{(l)}\|$ 的值是否小于 FCM 算法预先设的阈值，若该值小于此阈值则标志着 FCM 算法结束，图像分割已完成，反之，若该值大于或等于 FCM 算法预先设的阈值，则算法就会继续计算新的聚类中心，当达到停止条件时停止，此时算法收敛后，其分割图像可表示为式 (5-17)。

$$v_i = \sum_{x,y} (\mu_i(x,y))^m f(x,y) / \sum_{x,y} (\mu_i(x,y))^m \quad (5-12)$$

$$M(x,y) = \{i | 0 \leq i < c; d_i(x,y) = \|f(x,y) - v_i\| = 0\} \quad (5-13)$$

$$\bar{M}(x,y) = \{0, 1, 2, \dots, c-1\} - M(x,y) \quad (5-14)$$

$$u_i(x,y) = 1 / \sum_{q=0}^{c-1} (d_i(x,y)/d_q(x,y))^{2/(m-1)} \quad (5-15)$$

$$\mu_i(x,y) = 0, \sum_{i \in \bar{M}(x,y)} \mu_i(x,y) = 1, i \in \bar{M}(x,y) \quad (5-16)$$

$$\mu_i(x, y) = \max_i \{ \mu_0(x, y), \mu_1(x, y), \mu_2(x, y), \dots, \mu_{c-1}(x, y) \} \geq \partial \quad (5-17)$$

影响 FCM 聚类算法分割质量的因素如下：

1) FCM 聚类算法的初始隶属度矩阵，该矩阵直接影响 FCM 聚类算法是否能找到最佳的聚类中心以及影响 FCM 聚类算法的运算时间和迭代次数。

2) FCM 聚类算法的对称矩阵，该矩阵的矩阵形式直接影响 FCM 聚类算法的聚类分布，是球状分布、条状分布、带形分布、矩形分布还是菱形分布等。

3) FCM 聚类算法的聚类数目，这个参数直接影响 FCM 聚类算法的运算时间，假如此数目很大，其运算时间就会成倍增加，因而此时该算法在一般的实验设备就难以完成。

4) FCM 聚类算法的加权指数，该参数控制着模糊聚类的类间模糊程度，若该加权指数变大，则分类矩阵模糊程度就会变大；当该加权指数趋于无穷的时候，隶属度矩阵中的元素均接近 $1/c$ (c 为聚类的总数)，此时隶属度矩阵就失去了意义。

5) FCM 聚类算法的阈值设定，如果阈值设置得太大，则每次运算的聚类结果就会有较大的差异，且聚类结果也不稳定；反之，如果阈值设置得太小，则该算法的运算量就会增加，从而导致运算时间变长以及产生该算法不能收敛的结果。

因此，改进 FCM 聚类算法的分割效果实质就是对以上几个因素的改进，本章对传统的 FCM 聚类算法进行了如下改进：

1) 隶属度矩阵的改进：利用图像中像素的空间特征加强原有的隶属度函数，使同类区域的邻域像素携带少量的噪声点的权重，最终获得较准确的聚类中心；改进的隶属度迭代公式见式 (5-18)，其中 p 、 q 分别为控制原有隶属度和控制空间函数间关系的参数；改进隶属度后得到的聚类中心的迭代函数的公式见式 (5-19)。

$$u_{ij} = (u_{ij}^p \cdot f_{ij}^q) / \sum_{r=1}^c u_{ij}^p \cdot f_{ij}^q \quad (5-18)$$

$$v_i = \left(\sum_{j=1}^N (u_{ij})^m x_j \right) / \sum_{j=1}^N (u_{ij})^m \quad (5-19)$$

2) 对称矩阵的选择：根据图像的聚类分布对应地选择出对称矩阵的形式。

3) 聚类数目的确定：本章采用结合阈值分割的分水岭算法中的分水岭盆地的个数作为初始的聚类数目，从而解决了传统的 FCM 聚类算法不能自动确定总的聚类数目的问题。

4) 加权指数的确定：对于不同的图像，加权指数也具有不同的取值范围，但加权指数有一个经验范围是 $[1, 1.5]$ ，之后又从物理上求出，当加权指数取 2 的时候，其表示的模糊聚类的类间模糊程度最有意义。

5) 阈值设定：本章采用结合梯度特征的最大类间方差法无监督、自适应地确定了多源图像的分割阈值，从而为 FCM 聚类算法收敛到较好的聚类结果奠定了一定的基础。

因此，通过上述改进的 FCM 聚类分割算法可以求出新的目标函数的表达式为式 (5-20)，并且用迭代算法可求出目标函数 $J(U, V)$ 最小值，从而得到最佳的图像矩阵与聚类中心的配对，即 (U, V) ，最终完成聚类分割的目的。

$$J(U, V) = \sum_{i=1}^c \sum_{j=1}^N (u_{ij})^2 (d_{ij})^2 \quad (5-20)$$

因此，本章采用结合空间特征的 FCM 聚类算法对分水岭算法进行改进操作，从而形成了结合改进的 FCM 聚类分割的分水岭算法，该算法减弱了分水岭算法易受图像中的量化误差的影响；与此同时该算法也归并了图像中那些细微的无语义学意义的过分割的小区域，使得被分割的目标轮廓更为清晰；并且该算法进一步改善了传统分水岭算法中易产生过分割的缺陷，最终得到具有较高分割质量的图像。本章采用 VC++ 6.0 开发环境编程调试完成的结合空间特征的 FCM 聚类分割算法对结合阈值算法的分水岭分割图像进行图像分割，其分割结果分别如图 5-2 和图 5-3 所示，其中图 5-2 为可见光图像的结合聚类算法的分水岭分割图像，图 5-3 为红外图像的结合聚类算法的分水岭分割图像。

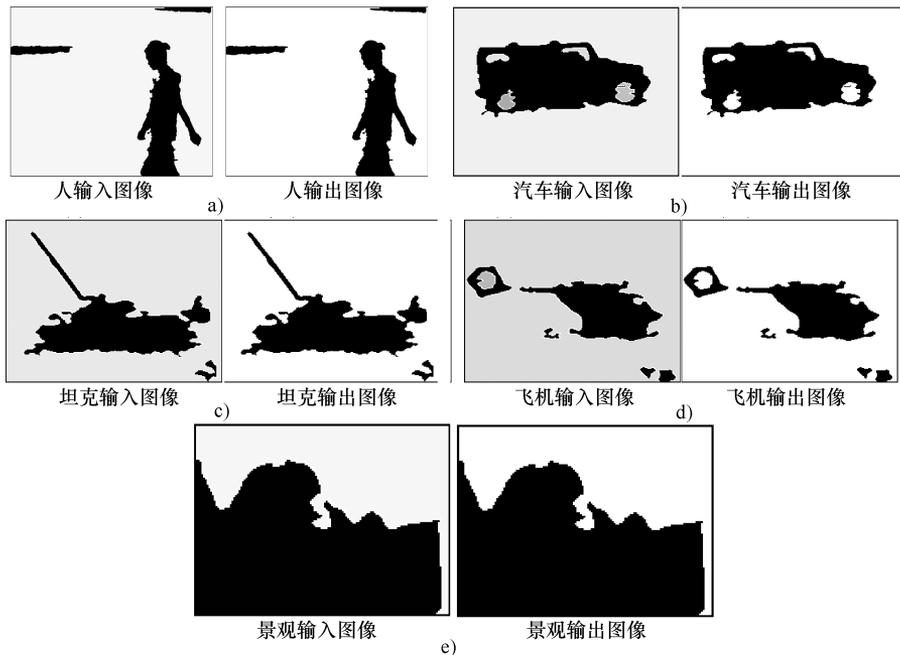


图 5-2 可见光图像的结合聚类算法的分水岭分割图像

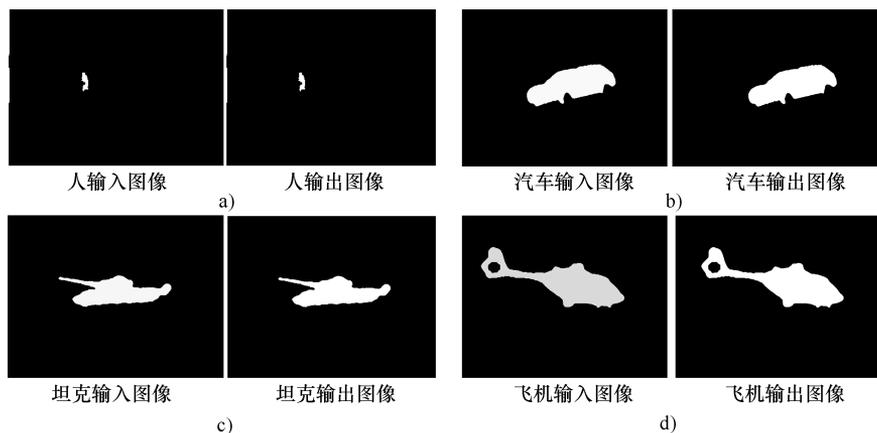


图 5-3 红外图像的结合聚类算法的分水岭分割图像

5.4 本章小结

本章主要研究了多源图像的聚类分割的方法，首先研究了聚类算法的基本思想、基本原理、定义、所需的数据类型、评价聚类算法的标准以及实现图像聚类分割算法的基本步骤，从而为寻找合适的聚类分割方法奠定了扎实的理论基础；其次，本章通过分析比较常用的聚类分割方法的优缺点，从而为寻找最佳的聚类分割方法提供了一定的思路；再次，本章主要研究了模糊聚类算法中的一种典型算法 FCM 算法，并对影响该算法分割效果的因素进行了深入分析，从而为改进 FCM 算法提供了一定的思路，并且因此本章提出了一些改进 FCM 算法的措施，使 FCM 算法能够自动地确定聚类数目、具有较好的最佳收敛阈值，并且该算法最终被改进成结合空间特征的 FCM 聚类分割算法，最后再使用改进的 FCM 聚类分割算法对本章的分水岭分割方法进行改进，从而又提出了结合改进的 FCM 聚类分割的分水岭算法，并且本章将改进的 FCM 聚类分割算法在 VC++6.0 开发环境下进行了实验，从而完成了对分水岭算法的更大改进。

目标特征提取方法

随着微电子、计算机、信息技术等学科的迅速发展和广泛应用，传感器技术也得到了快速发展，越来越多的应用系统配备了多个传感器以满足实际需要。因此，如何利用多个传感器所蕴含的丰富信息来确保系统更可靠、性能更卓越是摆在研究人员面前的一大任务。多传感器所蕴含的信息具有多样性、复杂性和冗余性，并且大多数应用环境需要对信息进行实时处理，单凭人是无法实现的，因此需要利用计算机进行计算，实现特征融合。所谓多传感器信息融合是指对来自多个传感器的信息进行多级别、多方面、多层次的处理与综合，从而获得更丰富、更精确、更可靠的有用信息。

多传感器图像特征融合是多传感器信息融合的一个分支。多传感器图像特征融合的主要思想是利用不同输入信道图像特征信息的冗余性和互补性，采用一定的特征提取方法，把两个或多个不同传感器图像进行特征提取和融合，从而使融合的目标特征向量能更全面地描述目标特征，进而提高目标识别系统的可靠性。例如，可见光传感器对图像的亮度变化敏感，能较好地显示对比度和纹理细节等，可以提取灰度图像的边缘纹理特征。红外传感器反映目标和场景的红外辐射特性，可以全天候监测，可以提取图像的形状特征。利用可见光和红外图像的互补性，分别提取同一场景可见光和红外图像的独有特征，得到一组目标融合特征量，进一步进行目标分类或识别。

随着科技发展，多传感器应用越来越多，是未来发展的趋势。而目标识别是多传感器应用的重要领域。特征提取是目标识别的关键技术，对目标能否有效识别起决定性作用。近年来对单一传感器的特征进行提取，已有很多学者在研究，并取得了很大的进展，但是对多种传感器进行特征提取和融合的研究并不多，有一些学者和院校刚刚开始。

无论采用什么融合算法进行自动目标识别，特征提取是关键。目标的特征是目标所具有的最基本的内容，是该目标特有的、用于区别于其他类型目标的最本质的属性。在实际目标识别应用系统中，如何使目标特征化、提取能够全面描述目标的特征是实现实时、准确目标识别的关键所在。作为目标识别的关键步骤，特征提取的目的是获取一组能准确描述目标的“少而精”的分类特征向量，通过对图像目标的特征提取，可以有效地减少冗余信息，减少系统的计

算量，从而增强识别系统的可靠性。

在现有的目标识别系统中，常用的特征有角点特征、矩特征、纹理特征、变换特征、统计性特征等。

1. 角点特征

在形状分析中，目标轮廓上的角点是形状常用的特征。相对于其他特征量而言，不受目标遮挡、缺损的影响，因此角点特征在目标识别中非常重要。然而，角点的定义一直很模糊，近年来学者提出了许多角点检测算法，如 Kitchen 发现以局部梯度乘以梯度方向的变化可以很好地提取角点。Harris 和 Stephens (1988) 采用了相同的思想对 Moravec 算子进行改进，提出了著名的 Plessey 角点提取算子。Smith 和 Brady (1997) 提出了一种完全不同的角点提取方法，即“SUSAN (Smallest Univalued Segment Assimilating Nucleus)”提取算子。SUSAN 提取算子的基本原理是，与每一图像点相关的局部区域具有相同的亮度。

在目标识别中，角点特征应用广泛。卢汉清等人把形心到相邻两角点的直线所成的夹角作为特征量用于目标识别。Shutal Li 等人把角点与线矩融合用于缺损目标识别，有较高的识别率。由于检测角点时容易出现漏检，通常利用角点和其他特征相结合，可以取得较好地识别效果。曹健等人提出了一种不变性的角点构造方法，用于目标识别中。

2. 矩特征

不变矩是具有平移、比例和旋转不变性的几何矩。它是利用各阶矩的组合矩作为特征量，进行目标识别。Hu 于 1962 年定义了连续函数矩，给出了七个不变矩，证明了矩具有平移、比例以及旋转不变性，并用于字母自动识别中。R. Y. Hong 在 Hu 不变矩的基础上，推导出了离散函数不变矩的多项式。近年来，几何矩应用越来越多，许多学者和院校在理论方面做了许多研究，并提出了一些新的矩定义和性质。针对多项式不变矩含有冗余信息，引入了正交矩，人们提出了 Zernike 矩的基本定义和性质，及其计算方法和应用领域。进而，为了满足径向和角度旋转不变性，把几何矩推广到径向矩。随着小波理论的发展，Shen 等提出了基于小波变换的小波不变矩，不仅能描述目标全局特征，还能描述目标局部特征。

3. 纹理特征

图像的纹理是图像像素值在灰度空间上的重复和变化，或是反复出现的局部纹理模式及其排列规则。纹理特征是图像的最基本特征，并在视觉系统起着关键作用，为图像理解和分析提供了重要信息。Sutton 和 Hall 研究了利用纹理的特征来分类识别肺病。Herlidou 等人利用纹理的特征对磁共振脑组织进行识别。

4. 变换特征

图像变换特征就是首先把图像变换成频域，利用频域中变换系数中的相关性来识别目标。在图像有随机噪声时，不影响变换特征的分类效果，较为常用

的傅里叶变换就是用的图像频谱特征。

5. 统计性特征

基于统计参数特征的目标识别是将一幅图像看成是一个二维随机过程的一次实现，这样便可以使用各种统计参数来描述图像的特征，这些统计参数有均值、方差、能量、熵等特征量。Haralick 等用灰度共生矩阵纹理特征对遥感图像进行分类研究，并获得了大约 80% 的分类精度。

目前国内外计算机视觉、模式识别与人工智能等领域都对图像的目标特征提取及其应用进行了深入研究，并取得了快速发展，一些成果已具备初步的实用价值。随着遗传算法、神经网络、形态学、统计学、小波理论等深入研究广泛应用，图像目标特征提取发展趋势如下：

多种特征融合：除了利用图像的原始灰度特征外，还可利用图像的高层次特征，如视觉特征、统计特征、变换系数特征等，通过多种特征的融合，能够更全面地描述图像目标，提取的特征更准确。多特征融合已得到广泛应用。例如，张建军等人把小波能量信息特征与图像矩特征结合起来，用于制导武器红外图像的目标识别，结果表明，有较高使用价值。将基于显著性特征提取的目标识别方法与序贯融合方法相结合，用于飞机目标识别。

多种提取方法结合：由于目标的多样性和复杂性，单一的特征提取方法难以对含复杂目标的图像进行提取。在这种情况下，除需要利用多种特征的融合外，还需要将多种提取方法结合使用，使提取方法充分发挥各自的优势，避免各自的劣势。比如于吉红等人把部分 Hu 矩、仿射矩和小波矩组合在一起，用于舰船图像目标的分类识别，提高了识别率。张劲锋等人把 Hu 不变矩的部分分量和仿射不变矩结合成新的特征向量，用于空间目标的识别。

多种传感器融合：由于不同传感器描述目标的多样性，采用单一传感器不能全面、准确描述目标，需要利用多种传感器的互补特性，提取目标的不同特征，进行多传感器特征融合，全面描述目标特征，提高识别系统的鲁棒性和识别率。例如，熊大容等人利用红外和可见光的互补优势，对远距离的目标进行检测，增强了系统的可靠性。凌虎等人分别提取不同传感器的轮廓特征，融合在一起用于目标检测。

与图像分割方法相适应：由于图像目标的多样性和各种应用需求的复杂性，图像的特征提取应与图像的分割方法相结合，特定的提取方法选择特定的分割方法，来获得最好的图像识别结果。

由于图像目标特征的复杂性和多样性，与图像分割相对应，现有的方法不能满足实际要求，一些根本问题有待进一步研究，还没有统一的应用所有模式识别的特征提取方法。虽然各种特征提取方法在提取能力和处理速度方面各有优势，但是在通用性、性能、准确率、自动化程度方面还有很多不足。因此，对于图像目标特征提取方法的研究需要付出更多努力和关注。

6.1 传统的特征提取方法

由于目标的多样性及其复杂性，使得寻找具有准确、全面描述目标特征信息和具有良好分类性能的图像目标特征以及提取这些特征就成为解决图像目标识别问题的关键。一般来说，不同的传感器使用不同的特征描述，特征提取方法也不相同。特征提取与选择就是对于预处理后的图像目标数据进行降维处理、去粗取精的过程。由于原始图像数据量相当大，为了快速计算出目标识别结果，减少计算量，必须把这些数据转换为若干个特征量，称为特征提取。为了提高识别的速度和精度，对提取的特征还必须进行降维，选择信息冗余度较小的特征量，并且具有比例、旋转和位移不变性等特性，增强提取方法的鲁棒性。

在提取目标特征时，提取的目标特征要尽可能地反映目标重要的、本原的特性。重要特征是指以它们作为特征分量能实现同类目标聚集、异类目标分散。本原特征是指特征绝对性强，尽可能地不依赖于提取目标特征时的条件和环境。这两个特征也就是要求所提取的特征量在同一类型目标上具有唯一性和稳定性，不同类型目标之间具有可区分性。特征提取和选择是目标识别的核心，也是识别分类器能准确识别目标的前提。保证所提取的目标的特征量稳定、可靠和实用是整个识别算法取得成功的关键。

6.1.1 经典的特征提取方法

图像的二维特征有形状、区域和纹理等特征。不同特征描述具有不同的提取算法。

1. 形状特征

形状特征指图像中目标的几何特征，根据几何形状分为线形特征和块状形特征。

(1) 周长、长宽比、复杂度、面积

目标长宽比是目标最小外接矩形的长度和宽度之比。它可以把不同几何形状的物体区分开来，比如长方形和正方形。假设目标的最小外界矩形长为 L ，宽为 W ，长宽比为

$$\psi_1 = L/W \quad (6-1)$$

目标周长是指目标边界的长度，而面积是目标区域的像素总数。利用图像目标的面积和周长可简单和有效地把复杂与简单形状的物体区别开来。设目标周长为 C ，面积为 S 。

目标形状的复杂度是目标周长的 2 次方和面积的比值，目标形状越复杂，则相同面积目标的边缘长度就越长。公式为

$$\psi_2 = C^2/S \quad (6-2)$$

可以证明，这种目标复杂度特征定义与场景到镜头的距离无关，不失一般性。

假定目标在横坐标和纵坐标方向上进行了缩放因子为 k 的平移，那么目标的面积变为原来的 $1/k^2$ ，周长变为原先的 $1/k$ ，代入上式，形状复杂度不变。

(2) 矩特征

图像矩特征是由下式所定义的 $M(p, q)$ 决定的

$$M(p, q) = \sum_{(i,j)} i^p j^q f_{ij} \quad (6-3)$$

式中， $\{f_{ij}\}$ 是在目标区域内设为 1，外设为 0 的二值化图像。由 (p, q) 值决定各种特征向量，常用的特征有主轴、重心等。

(3) 傅里叶描述

在分析图像目标形状时，首先跟踪目标边界线，并把对应的封闭曲线，展开为傅里叶级数，则用展开的系数表示目标形状特征。

除以上特征外，一些具有连通性的特征参数，如欧拉数、孔数、连接成分数等，也可用于表示图形特征。

2. 区域分割

区域分割是把图像对象物区域从背景中分割出来。最简单的方法就是把图像进行二值化处理，关键在于阈值的选取。较复杂的分割方法有：

(1) 区域扩张法

区域扩张法就是把图像进行小区域分类，根据小区域的特征相似性，把图像分割成特征相似的连续区域群。根据区域形成过程的不同，又可分为合并、像素结合等方法。合并法的基本原理是把图像分割为 $N_x \times N_y$ 的子区域，通过研究邻接小区域的相似性，合并相似性较高的区域，反复进行，直到不能合并为止。

(2) 聚类算法

该方法就是把像素作为一种模式，然后利用模式识别的理论进行区域分类。首先把图像像素变换到特征空间，然后利用分类法进行像素分类。在实际中，常用小区域来代替像素，有时用一维投影代替多维特征来进行目标区域分割。

3. 纹理特征

纹理是区域具有的典型特征之一。常用的特征有：

(1) 直方图特征

直方图特征是纹理区域的灰度直方图、方差和平均值等。但直方图并不能得到纹理的二维灰度的变化情况。通常需要进一步处理，可以与其他特征相结合，作为识别的特征。例如可以用二维统计量、能量、惯性矩等特征量，识别

目标。

(2) 傅里叶特征

纹理特征不仅可以在空域中描述，还可以在频域求得。首先计算出图像 $f(x, y)$ 傅里叶变换 $F(u, v)$ 的功率谱 $P(u, v)$ ，转换成极坐标为 $p(r, \theta)$ ，再求出：

$$p(r) = 2 \sum_{\theta=0}^{\pi} p(r, \theta) \quad q(\theta) = \sum_{r=0}^{\omega/2} p(r, \theta) \quad (6-4)$$

式中， ω 表示 $P(u, v)$ 的大小，于是可用 $p(r)$ 和 $q(\theta)$ 的波峰的大小和位置， $p(r)$ 和 $q(\theta)$ 的方差和平均值作为纹理特征。

通常图像有多个纹理区域，要想获得纹理特征，必须进行纹理分割。可以使用直方图分割法，也可以进行边缘检测。边缘检测法就是图像目标相邻像素急剧变化，可以用边缘检测算子检测，算子可分为两大类：一次微分得到的，如 Roberts 算子、Sobel 算子、Prewitt 算子等；二次微分得到的，如拉普拉斯算子、高斯-拉普拉斯算子。

6.1.2 多项式不变矩

6.1.2.1 几何矩

在直角坐标系中，连续函数 $f(x, y)$ 的 $p+q$ 阶几何矩定义为

$$M_{p,q} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad p, q = 0, 1, 2, \dots \quad (6-5)$$

这种矩是将函数 $f(x, y)$ 投影到 $x^p y^q$ 上，其中基本集 $\{x^p y^q\}$ 具有完备性、不正交性。

唯一性定理：如果函数 $f(x, y)$ 分段连续，且只在 (x, y) 的有限区域有非零值，则所有阶的矩都存在。由 $f(x, y)$ 可唯一确定矩序列 $\{M_{pq}\}$ ；反之，由矩序列 $\{M_{pq}\}$ 也可唯一确定 $f(x, y)$ 。

一幅图的面积必然有限和分段连续，因此，图像的所有阶矩必然存在，用这些矩描述图像具有唯一性。理论上，表征图像的所有信息，需要计算无限多矩值。而实际应用中，只需计算包含足够多图像信息的矩值子集就行。

下面分别介绍各低阶矩及其物理意义：

(1) 零阶矩

由式 (6-5) 的定义可知，图像 $f(x, y)$ 的零阶矩为

$$M_{00} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy \quad (6-6)$$

由上式可知，对于一幅灰度图像， M_{00} 表示图像所有灰度值的和；对于一幅二值图像， M_{00} 表示图像面积。

(2) 一阶矩

图像 $f(x, y)$ 的一阶矩 M_{01} 和 M_{10} 可用来表示图像重心。设图像重心坐标为 (\bar{x}, \bar{y}) ，计算公式为

$$\bar{x} = \frac{M_{10}}{M_{00}} \quad \bar{y} = \frac{M_{01}}{M_{00}} \tag{6-7}$$

对于二值图像 (\bar{x}, \bar{y}) 来说它表示几何中心。若把图像重心与坐标系原点重合，令 $\bar{x}=0, \bar{y}=0$ ，那么式 (6-7) 就为图像中心矩。这样计算出的矩就与图像坐标系无关。中心矩计算公式为

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})(y - \bar{y})f(x, y) dx dy \tag{6-8}$$

由上式计算可得，一阶中心矩 μ_{01}, μ_{10} 满足 $\mu_{01} = \mu_{10} = 0$ 。

(3) 二阶矩

二阶中心矩 $\mu_{02}, \mu_{20}, \mu_{11}$ ，又称惯性矩。可用来描述物体的主轴和椭圆等重要特性。物体目标的主轴由长轴和短轴组成，它们分别表示最大和最小二阶矩的方向。主轴的方向角 θ 的公式为

$$\theta = -\frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{6-9}$$

式中， θ 为主轴与坐标轴的夹角，它的取值范围为 $\left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$ ，见表 6-1。

表 6-1 主轴方向的确定

μ_{11}	$\mu_{20} - \mu_{02}$	θ
0	-	0
+	-	$-\pi/4 < \theta < 0$
+	0	0
+	+	$0 < \theta < \pi/4$
0	0	0
-	+	$-\pi/4 < \theta < 0$
-	0	0
-	-	$0 < \theta < -\pi/4$

需要说明的是，仅使用主轴角并不能确定物体目标的方向。比如物体旋转 180° 后，主轴的方向不变，这就需要利用三阶中心矩，只有这样才能唯一确定物体方向。由一、二阶中心矩还可确定一个图像椭圆。图像椭圆就是一个与原图像的二阶矩和灰度总和都相等的椭圆。椭圆有长半轴 a 和短半轴 b ，定义为

$$a = \left(\frac{\mu_{20} + \mu_{02} + [(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2]^{1/2}}{\mu_{00}^{1/2}} \right) \tag{6-10}$$

$$b = \left(\frac{\mu_{20} + \mu_{02} + [(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2]^{1/2}}{\mu_{00}^{1/2}} \right) \quad (6-11)$$

椭圆内的灰度表示为

$$I = \frac{\mu_{00}}{\pi ab} \quad (6-12)$$

为了便于分析图像性质，一般将图像重心与椭圆中心和主轴方向重合，如图 6-1 所示。

(4) 三阶矩

三阶中心矩 μ_{30} 和 μ_{03} 描述图像投影后的扭曲程度。扭曲可用来衡量均值对称分布中的偏差程度。式 (6-13)、式 (6-14) 为图像在 x 、 y 轴投影的扭曲系数计算公式：

$$sk_x = \frac{\mu_{30}}{\mu_{20}^{3/2}} \quad (6-13)$$

$$sk_y = \frac{\mu_{03}}{\mu_{20}^{3/2}} \quad (6-14)$$

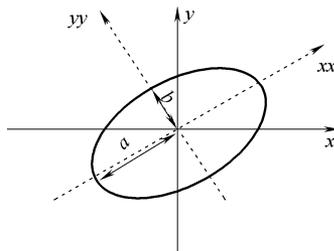


图 6-1 图像椭圆

扭曲系数的正负与图像扭曲的关系见表 6-2。对于前面提到的主轴方向的问题，是根据扭曲系数的正负来确定的。

表 6-2 扭曲系数的正负与图像扭曲的关系

图像扭曲	扭曲系数符号
x 轴投影扭曲方向	sk_x
y 轴左侧	+
对称于 y 轴	0
y 轴右侧	-
y 轴投影扭曲方向	sk_y
x 轴下方	+
对称于 x 轴	0
x 轴上方	-

(5) 四阶矩

四阶中心矩 μ_{40} 和 μ_{04} 又称投影峰度。式 (6-15)、式 (6-16) 为图像在 x 、 y 轴投影的投影峰度系数计算公式：

$$k_x = \frac{\mu_{40}}{\mu_{20}^2} \cdot 3 \quad (6-15)$$

$$k_y = \frac{\mu_{04}}{\mu_{20}^2} \cdot 3 \quad (6-16)$$

式中, 令峰度系数等于 0, 为高斯分布; 小于 0, 为平坦而少峰; 大于 0, 为狭窄而多峰。

几何矩经过简单变化, 就可以满足平移、缩放和旋转不变性。

平移变换: 将图像 $f(x, y)$ 在 x 轴和 y 轴方向分别平移 a 和 b , 平移后的图像 $f'(x, y)$ 表示为

$$f'(x, y) = f(x - a, y - b) \quad (6-17)$$

图像 $f'(x, y)$ 的 $p+q$ 阶矩为

$$M'_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x+a)^p (y+b)^q f(x, y) dx dy \quad (6-18)$$

设数字图像为 $g(m, n)$, 则式 (6-18) 变换为

$$M'_{pq} = \sum_m \sum_n (m+a)^p (n+b)^q g(m, n) \quad (6-19)$$

平移前后矩 M_{pq} 和 M'_{pq} 的关系为

$$M'_{pq} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \binom{q}{s} a^{(p-r)} b^{(q-s)} M_{rs} \quad (6-20)$$

由式 (6-20) 可得图像 $f(x, y)$ 的 $p+q$ 阶中心矩为

$$\mu_{pq} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \binom{q}{s} (x-\bar{x})^{(p-r)} (y-\bar{y})^{(q-s)} M_{rs} \quad (6-21)$$

则平移变换前后有

$$\mu'_{pq} = \mu_{pq} \quad (6-22)$$

比例变换: 将图像 $f(x, y)$ 在 x 轴和 y 轴方向分别缩放 α, β 倍, 则比例变换后的新图像 $f'(x, y)$ 为

$$f'(x, y) = f(x/\alpha, y/\beta) \quad (6-23)$$

变换前后的矩 M'_{pq} 和 M_{pq} 关系如式 (6-24)、式 (6-25) 所示:

$$M'_{pq} = \alpha^{1+p} \beta^{1+q} M_{pq} \quad \alpha \neq \beta \quad (6-24)$$

$$M'_{pq} = \alpha^{2+p+q} M_{pq} \quad \alpha = \beta \quad (6-25)$$

令 $M_{00} = 1$ 就可实现图像比例归一化。对于二值图像, 可以把图像面积的值设为 1。实现归一化的尺度因子表示为

$$p = \frac{1}{\sqrt{m_{00}}} \quad (6-26)$$

归一化后原点矩为

$$M'_{pq} = \left(\frac{1}{\sqrt{m_{00}}} \right)^{(p+q+2)} M_{pq} \quad (6-27)$$

通过上述的处理方法, 可得出具有比例 (缩放) 不变性的中心矩为

$$\eta_{pq} = \frac{1}{(M_{00})^y} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-\bar{x})^p (y-\bar{y})^q f(x, y) dx dy \quad (6-28)$$

式中,
$$y = \frac{p+q+2}{2} \quad (6-29)$$

旋转变换: 将图像 $f(x, y)$ 绕坐标原点作逆时针 θ 角度旋转后, 得到的图像 $f'(x, y)$ 为

$$f'(x, y) = f[(x\cos\theta + y\sin\theta), (-x\sin\theta + y\cos\theta)] \quad (6-30)$$

旋转变换后图像 $f'(x, y)$ 的 $p+q$ 阶矩为

$$M'_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x\cos\theta - y\sin\theta)^p (x\sin\theta + y\cos\theta)^q f(x, y) dx dy \quad (6-31)$$

对于数字图像 $g(m, n)$, 式 (6-31) 变为

$$M'_{pq} = \sum_m \sum_n (m\cos\theta - n\sin\theta)^p (m\sin\theta + n\cos\theta)^q g(m, n) \quad (6-32)$$

旋转变换前后的矩 M_{pq} 和 M'_{pq} 的关系如下:

$$M'_{pq} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \binom{q}{s} (-1)^{q-s} (\cos\theta)^{q+r+s} M_{p+q-r-s, r+s} \quad (6-33)$$

采用主轴法可以实现矩的旋转不变性。对于有唯一分布的主轴图像, 计算它的各阶矩, 这些矩在分布方向上是不变的。令 $\mu_{11} = 0$, 可得到主轴。

归一化后标准矩定义如下:

$$(RMI)_{pq} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \binom{q}{s} (-1)^{q-s} (\cos\theta)^{q+r+s} (\sin\theta)^{q+r-s} \eta_{p+q-r-s, r+s} \quad (6-34)$$

6.1.2.2 Hu 矩的构造

事实上, 几何矩 M_{pq} 并不具有不变性; 图像的中心矩 μ_{pq} 仅具有平移的不变性; 归一化中心矩有多种构造方式。

方式一为

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{20} + \mu_{02})^{(p+q+2)/4}} \quad p+q=2, 3, \dots \quad (6-35)$$

另一种方式为

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{(p+q+2)/2}} \quad p+q=2, 3, \dots \quad (6-36)$$

归一化的中心矩 η_{pq} 具有平移和比例不变性, 但不具有旋转不变性。针对这一问题, Hu 在 1961 年给出了七个不变矩特征量, 并证明了它们具有平移、比例不变性和旋转不变性的性质。七个不变矩多项式表达式为

$$M_1 = \eta_{20} + \eta_{02} \quad (6-37)$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (6-38)$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (6-39)$$

$$M_4 = (\eta_{30} + 3\eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (6-40)$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (6-41)$$

$$(3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$M_6 = (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{21} + \eta_{03})(\eta_{30} + \eta_{12}) \quad (6-42)$$

$$M_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (6-43)$$

$$(\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

由于七个不变矩值的变化范围很大，为了便于比较，可利用取对数的方法进行数据压缩。因此，实际采用的不变矩为

$$M_k = |\log |M_k|| \quad k = 1, 2, \dots, 7 \quad (6-44)$$

Hu 不变矩的计算流程如图 6-2 所示。

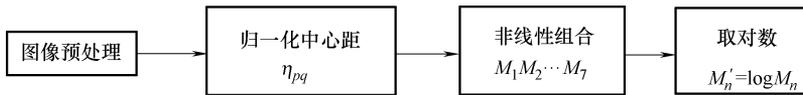


图 6-2 Hu 不变矩的计算流程

表 6-3 为对坦克图片 Hu 矩的特征提取结果。

表 6-3 坦克 Hu 矩特征提取结果

矩 图片	M_1	M_2	M_3	M_4	M_5	M_6	M_7
原始图像	2.8912	5.9812	9.9958	11.2356	21.8595	14.3698	22.5653
旋转 90°	2.8912	5.9812	9.9958	11.2356	21.8595	14.3698	22.5653
旋转 180°	2.8912	5.9812	9.9958	11.2356	21.8595	14.3698	22.5653
缩小一倍	2.8911	5.9808	9.9942	11.2340	21.8559	14.3661	22.5750
放大一倍	2.8910	5.9807	9.9945	11.2350	21.8579	14.3687	22.5684
左移 20%	2.8912	5.9812	9.99584	11.2356	21.8595	14.3698	22.5653
右移 20%	2.8912	5.9812	9.99584	11.2356	21.8595	14.3698	22.5653

由表 6-3 中数据可知，通过对坦克图片进行旋转、尺度和平移变换，Hu 矩提取的特征参数不发生变化，满足矩不变性。

6.1.2.3 仿射不变矩的构造

Hu 提出的不变矩具有平移、尺度和旋转不变性，但当由于拍摄角度不同，图像出现扭曲等变形时，这三个不变性并不能满足要求，因此需要构造一种目标发生扭曲、拉伸等仿射变换条件下的不变矩特征。此特征能用于不同拍摄角度下目标的识别。

仿射变换是一种典型的线性变换，它是通过不同的线性变换进行构造的。设

图像平面上一点 P 的坐标为 (x, y) 和对应仿射变换坐标系中点 P' 的坐标为 (x', y') ，则其二维的仿射变换公式如下：

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} + B \quad (6-45)$$

式中， $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 是仿射变换矩阵。仿射变换矩阵 A 必须满足非奇异性条件，

即 $\det(A) \neq 0$ 。 $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ 是平移量，在二维矩阵空间中，矩阵 A 可以按扭曲、伸缩、尺度、旋转等进行构造。

尺度变换模型为

$$A_m = \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}, m \geq 0 \quad (6-46)$$

伸缩变换模型为

$$A_n = \begin{bmatrix} 1 & 0 \\ 0 & n \end{bmatrix} \quad (6-47)$$

扭曲变换模型为

$$A_i = \begin{bmatrix} 1 & i \\ 0 & 1 \end{bmatrix} \quad (6-48)$$

旋转变换模型为

$$A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, 0 \leq \theta \leq 2\pi \quad (6-49)$$

因此，矩阵 A 可表示为

$$A = A_m A_n A_i A_\theta = \begin{bmatrix} \theta \cos m & ni\theta \cos m - i\theta \sin m \\ \theta \sin m & ni\theta \sin m + i\theta \cos m \end{bmatrix} \quad (6-50)$$

尺度、平移、伸缩、旋转和扭曲五个变换是仿射变换中的特例。

由仿射变换模型可知，如果一个特征量，在尺度、平移、伸缩、旋转和扭曲变换条件下仍然不发生变换，那么我们就认为该特征量是仿射不变量。

设图像 $f(x, y)$ 仿射变换后的图像为 $f_T(x, y)$ ，则仿射变换后图像的重心变为

$$\begin{bmatrix} \bar{x}' \\ \bar{y}' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (6-51)$$

把式 (6-45) 与式 (6-51) 相减，得

$$\begin{bmatrix} x' - \bar{x}' \\ y' - \bar{y}' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} \quad (6-52)$$

平移参数 b_1 、 b_2 被抵消掉。

由前面的内容可知，归一化中心具有比例、平移和旋转不变性。若利用归一化中心矩来构造仿射不变矩，只需要满足扭曲和拉伸不变性，就可达到仿射变换不变性。我们可以构造中心矩多项式，来抵消仿射变换矩阵 A ，就可满足常用的仿射不变性。构造多项式的方法有配极多项式、Hankel 行列式、多项式判别式等方法。本章采用 Jan Flusser 等人构造的 6 个仿射不变矩，作为目标图像的特征不变量，进行目标识别。

$$I_1 = (\mu_{20}\mu_{02} - \mu_{11}^2) / \mu_{00}^4 \tag{6-53}$$

$$I_2 = (\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{03}\mu_{21}\mu_{12} + 4\mu_{03}\mu_{30}^3 + 4\mu_{30}\mu_{12}^3) / \mu_{00}^{10} \tag{6-54}$$

$$I_3 = (\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{12}\mu_{30} - \mu_{21}^2)) / \mu_{00}^7 \tag{6-55}$$

$$I_4 = (\mu_{20}^3\mu_{03}^2 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{03}\mu_{12}^2 + 12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} + 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} - 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} - 8\mu_{11}^3\mu_{03}\mu_{30} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} + 9\mu_{20}\mu_{02}^2\mu_{21}^2 + 12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21} + \mu_{02}^3\mu_{30}^2) / \mu_{00}^{11} \tag{6-56}$$

$$I_5 = (\mu_{40}\mu_{04} - 4\mu_{13}\mu_{31} + 3\mu_{22}^2) / \mu_{00}^6 \tag{6-57}$$

$$I_6 = (\mu_{04}\mu_{22}\mu_{40} + 2\mu_{13}\mu_{22}\mu_{31} - \mu_{04}\mu_{31}^2 - \mu_{40}\mu_{13}^2 - \mu_{22}^3) / \mu_{00}^9 \tag{6-58}$$

仿射不变矩的计算流程如图 6-3 所示。

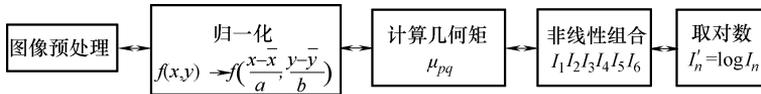


图 6-3 仿射不变矩的计算流程

表 6-4 为坦克图片仿射不变矩提取结果。

表 6-4 坦克图片仿射不变矩的特征提取结果

矩 图片	I_1	I_2	I_3	I_4	I_5	I_6
原始图像	6.7827	21.6081	14.8189	20.1820	12.3744	19.8410
旋转 90°	6.7081	20.3604	14.4650	19.9798	12.2705	19.5925
旋转 180°	6.7827	21.6081	14.8189	20.1820	12.3744	19.8410
缩小一倍	6.7826	21.6152	14.8079	20.1798	12.3737	19.8403
放大一倍	6.7824	21.6130	14.8063	20.1792	12.3731	19.8393
左移 20%	5.7491	15.8991	12.7326	16.6104	10.3175	15.7908
右移 20%	5.6145	15.0651	11.9479	16.4614	10.2095	15.3156
拉伸 20%	5.6359	15.9815	11.9941	16.5255	10.2595	15.7902

由表 6-4 中数据可知，当对坦克图片进行旋转、放缩变换时，仿射不变矩

不发生变化，对图片进行平移和拉伸变换时，提取的特征参数基本不发生变化，满足矩不变性。

6.1.3 共生矩阵

纹理反映的是图像的空间分布、灰度统计和结构信息。它是由一定大小和形状的像素集合组成的，是所有图像都具有的特性。纹理特征提取是指通过检测算法，检测出纹理基元并建立纹理模型，最终用特征量来描述。灰度共生矩阵是典型的纹理特征提取方法。它由两个位置像素的联合概率密度来定义，反映像素亮度特性及其像素之间位置关系。

共生矩阵表示像素空间的关系和依赖程度。设灰度共生矩阵中某一元素 (i, j) 的灰度值为 i ，另一个元素的灰度值为 j ，它们之间的距离为 d ，方向为 θ ，那么共生矩阵的值就是满足以上条件像素的个数。实际中， θ 一般选为 0° 、 45° 、 90° 、 135° 。图像的灰度级一般为 256，实际计算中远小于它。因为若矩阵维数过大，窗口较小，那么共生矩阵表示纹理的效果不好。若窗口较大，维数小，这样大大增加计算量，降低了实时性。所以在计算时，首先需要降低维数或减小灰度级。

设数字图像 $f(x, y)$ 的大小为 $M \times N$ ，灰度级为 N_g ，则灰度共生矩阵为

$$P(i, j) = \#\{(x_1, y_1), (x_2, y_2) \in M \times N \mid f(x_1, y_1) = i, f(x_2, y_2) = j\} \quad (6-59)$$

式中， $\#(x)$ 表示集合 x 中的元素个数，显然 P 为 $N_g \times N_g$ 的矩阵，若 (x_1, y_1) 与 (x_2, y_2) 间距离为 d ，并与横轴的夹角为 θ ，则可得任何间距和角度的共生矩阵 $P(i, j, d, \theta)$ 。图 6-4 为四个不同方向上的共生矩阵示意图。

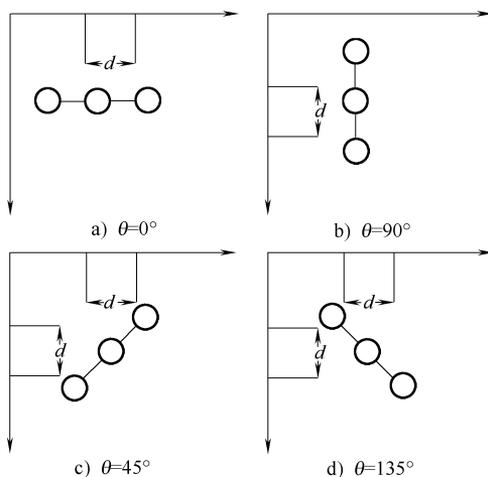


图 6-4 四个不同方向上的共生矩阵示意图

纹理特征提取是以灰度级的空间相关矩阵为基础，这是因为灰度共生矩阵表示了图像中相距 $(\Delta x, \Delta y)$ 的两个像素同时出现的联合频率分布。若图像灰度级为 N ，那么共生矩阵为 $N \times N$ 矩阵，可表示为 $M(h, k)$ ，其中位于 (h, k) 的元素 M 的值表示两个灰度分别为 h 和 k 、相距为 $(\Delta x, \Delta y)$ 的像素对的出现次数。

为了使共生矩阵能更直观描述纹理，用一些统计参数来表述纹理状况，常用的有：

(1) 角二阶矩

$$f_1 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p^2(i, j) \quad (6-60)$$

角二阶矩是图像灰度分布均匀性的度量。当共生矩阵中的元素分布比较集中于主对角线时，说明从局部区域观察图像的灰度分布是均匀的，从式 (6-60) 可知，它是共生矩阵中所有像素值的平方和，也可称为能量。从整个图像来看，若角二阶矩 f_1 越大，则纹理较粗，能量越多；反之亦然。

(2) 惯性矩

$$f_2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} |i - j| p(i, j) \quad (6-61)$$

若纹理较粗， $p(i, j)$ 的值主要集中在主对角线的附近，这时 $\|i - j\|$ 的值比较小，因此惯性矩也小；而纹理较细， $p(i, j)$ 数值的分布较均匀，惯性矩比较大。

(3) 对比度

$$f_3 = \sum_{i=0}^{L-1} n^2 \left[\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i, j) \right] \quad (6-62)$$

式中， $|i - j| = n$ 。

图像的对比度可以理解为图像的清晰度，即纹理的清晰程度。图像中纹理的沟纹越深，则其对比度 f_3 越大，图像越清晰。反之，对比度小，则沟纹浅，效果也较模糊。

(4) 相关

$$f_4 = \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} ij p(i, j) - u_1 u_2}{\sigma_1^2 \sigma_2^2} \quad (6-63)$$

式中，

$$u_1 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} ip(i, j), u_2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} jp(i, j);$$

$$\sigma_1^2 = \sum_{i=0}^{L-1} (i - u_1)^2 \sum_{j=0}^{L-1} p(i, j), \sigma_2^2 = \sum_{i=0}^{L-1} (j - u_2)^2 \sum_{j=0}^{L-1} p(i, j)$$

相关是用来衡量灰度共生矩阵的元素在行的方向或列的方向的相似程度。例如，某图像具有水平方向的纹理，则图像在 $\theta=0^\circ$ 的灰度共生矩阵的相关值往往大于 $\theta=45^\circ$ 、 $\theta=135^\circ$ 或 $\theta=90^\circ$ 的灰度共生矩阵的相关值。相关值的大小反映的是局部灰度的相关性。若矩阵中元素的值均匀相等，相关值就大；若矩阵中元素值相差很大，则相关值较小。

(5) 熵

$$f_5 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i,j) \log p(i,j) \quad (6-64)$$

熵值是对图像中信息量的度量，纹理属于图像信息。若图像没有纹理，则熵值几乎为零。如果图像有很多细纹理，并且 $p(i, j)$ 的数值比较接近，那么该图像熵值将达到最大。反之亦然。

(6) 方差

$$f_6 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - u)^2 p(i,j) \quad (6-65)$$

式中， u 为 $p(i, j)$ 的均值。

(7) 和方差

$$f_7 = \sum_{k=0}^{2L-2} \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (k - f_9)^2 p(i,j) \quad (6-66)$$

方差、和方差反映了纹理的周期大小，它们的值越大，表明纹理的周期越大。

(8) 逆差矩

$$f_8 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{p(i,j)}{1 + (i - j)^2} \quad (6-67)$$

逆差矩又称为局部平稳，它是图像纹理局部变化的度量，反映了纹理的规则程度。纹理越规则，值就越大。反之亦然。

(9) 和平均

$$f_9 = \sum_{k=0}^{2L-2} \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} kp(i,j) \quad (6-68)$$

式中， $k=i+j$ 。和平均是图像区域内像素点平均灰度值的度量，适用于灰度图像。

(10) 差平均

$$f_{10} = \sum_{k=0}^{L-1} \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i,j) \quad (6-69)$$

(11) 差方差

$$f_{11} = \sum_{k=0}^{L-1} (k - f_9)^2 \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i,j) \quad (6-70)$$

差方差是邻近像素灰度值差异的方差，若像素对比得越强烈，值越大。

(12) 差熵

$$f_{12} = \sum_{k=0}^{L-1} \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i,j) \sum_{j=0}^{L-1} p(i,j) \quad (6-71)$$

式中, $k = |i - j|$ 。

(13) 和熵

$$f_{13} = \sum_{k=0}^{2L-2} \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i,j) \log \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i,j) \quad (6-72)$$

式中, $k = i + j$ 。

和熵、差熵代表的意义参考熵。

综上所述, 取不同的 θ 、 d 和灰度级组合就会生成不同的共生矩阵, 得到的特征参数值也就不同, 这样描述纹理的能力也会有差异。因此, 通过分析特征参数, 可以得到纹理的不同描述。

6.2 红外和可见光图像特征提取和融合

6.2.1 多传感器特征提取

多传感器图像融合是多传感器数据融合中的一个重要内容。图像融合是指对多个传感器同一时间获取的关于某一具体场景的图像进行特征提取和融合, 融合后的特征更能全面、准确地描述目标特征, 这是单一的传感器不可比拟的。图像融合的目的是获取同一场景目标的更加准确、全面和可靠的特征。

多传感器能够提供互补的信息。利用这些互补信息, 可提高多传感器系统的可靠性, 使得对某一传感器的依赖性降低。由于实际应用中, 环境和气候经常发生变化, 这时我们可以发挥各个传感器的优势, 结合多种传感器用于进行目标检测和识别。比如, 当光照受雨、云、雾、烟等条件限制或光照强度较低时, 可见光传感器很难探测到目标, 这时可以利用具有较强穿透力的毫米波雷达进行探测, 尽管图像信号的衰减较严重, 但勉强可以看到目标。当环境和目标的温差差别较大时, 可以利用红外传感器的热辐射特性来辨识目标, 而可见光传感器能获取包含丰富的形状和彩色等的图像细节信息, 可以清楚地看到目标的精细特征。

特征级的图像融合指中间层次的融合方法。它首先利用各个传感器的特点, 分别提取出同一场景的不同的或相同的特征, 如目标形状、轮廓、边缘纹理等典型的目标特征, 然后利用某些融合算法, 把这些特征进行组合, 组合特征能更全面、准确地描述目标特征, 最后对融合后的特征进行目标分类识别。特征级融合包括图像分割、特征提取和特征层信息融合, 用于后续的目标分类识别中。目前常用的特征级的图像融合方法主要有: 神经网络方法、Dempster-Shafer

推理方法、聚类分析方法、信息熵方法、贝叶斯估计方法、表决方法及加权平均法。

可见光传感器和红外传感器是两种常用的图像传感器，它们的性能和工作机理有很大差异。红外传感器获取的图像为目标和场景的红外辐射特性，记录的是背景与目标的红外辐射强度信息。由于红外传感器反映的是背景与目标间的热辐射能量大小，一般目标热辐射能量强，在图像中显示较亮，因此在检测目标方面有很大优势，但其对场景的亮度变化并不敏感，成像的清晰度比较低，不能用于提取目标的细节特征；可见光传感器对目标和场景的反射较敏感，能够获取目标和场景的细节信息，可以用于提取目标的细节纹理特征，但对目标和场景的热对比度不敏感，很难获取目标的空间位置和结构信息。

本章主要研究了可见光传感器和红外传感器采集图像的特征提取方法。综合利用红外图像较好的空间探测能力和可见光图像的清晰场景信息，对同一场景的目标进行有效的特征提取。

6.2.2 目标区域分割和检测

由红外和可见光传感器的成像图像特点可知，红外图像具有较强的目标位置的检测能力，而可见光图像对目标的细节信息把握较强，因此，可以结合两种传感器的优势，对地面目标进行检测和特征提取，用于分类识别系统中。

首先，利用红外传感器检测出目标区域，这里计算目标的最小外接矩形，然后把目标区域映射到可见光图像中，确定可见光图像的目标区域，再对目标区域分别进行特征提取，由于红外图像的结构特征较明显，先对红外图像进行二值化处理，对处理后的图像进行形状矩特征提取；对于可见光图像，能够很好地反映图像的细节信息，先对目标区域进行边缘检测，而后对检测后的图像进行边缘特征提取。最后对不同传感器提取的特征进行特征融合，利用融合后的特征进行目标分类识别。具体的框图如图 6-5 所示。

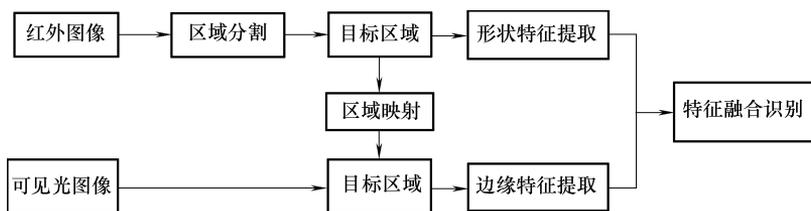


图 6-5 红外和可见光图像特征融合示意图

6.2.2.1 红外图像的区域分割

从红外图像的灰度图中可看出，感兴趣目标区域与背景中的路面的灰度相

差不很明显，若采用简单的阈值分割算法，若选取的阈值过低，则路面会被看成感兴趣目标而被分割出来，不仅提高了目标的虚警率，使后续特征提取的计算量也大大增加；若选取的阈值过高，则分割后目标区域过小，这样映射到可见光图像中，目标的轮廓特征丢失，这样提取的目标特征不准确。

因此，本章采用区域生长的图像分割方法对图像中感兴趣区域进行分割。区域生长法是一种基于某个或某些区域的串行图像分割技术，基本思想是把具有相似灰度级像素集合在一起构成目标区域。首先要在要分割的目标区域找一个种子像素点，把该点作为区域生长的起点，然后按照一定的规则，把种子像素周围邻域中像素点与种子像素进行比较，若相同或相近则把该点合并到种子像素所在的区域；依次进行比较，直到不能找到满足条件的像素点为止。这样就得到目标区域。

区域生长法的 3 个重要因素包括：

- 1) 确定能正确代表所需区域的种子像素。
- 2) 制定在生长过程中能将相邻像素包含进来的准则。
- 3) 指定让生长条件停止的条件或规则。

种子像素点选取可根据具体问题进行分析而定。由于红外图像的目标辐射较大，可以选择图像中最亮（灰度级最大）的像素点为种子像素。生长准则的选取与具体问题本身无关，是由所用的图像数据决定的。

本章采用的区域生长法的选取准则如下：

1) 种子像素点的选取：选取图像中灰度最亮点作为种子像素点。为了避免将强背景噪声选为种子像素点，在选点之前，对原红外图像进行平滑处理。平滑处理选取的模板为 $1/4 \begin{Bmatrix} 0, 1, 0 \\ 1, 0, 1 \\ 0, 1, 0 \end{Bmatrix}$ 。

2) 生长准则的确定：根据红外图像灰度分布的特点，采用基于区域灰度差的生长准则，即在种子像素点的八邻域内的像素，按照下式准则进行处理：

$$g(x, y) = \begin{cases} \max(v) & d(x, y) \geq v_0 \\ 0 & d(x, y) < v_0 \end{cases} \quad (6-73)$$

式中， $d(x, y)$ 是种子像素点八邻域内的点； $g(x, y)$ 是分割后的二值化图像，若 $d(x, y)$ 满足条件，即作为新的种子像素点进行区域生长； $\max(v)$ 为最大值； v_0 为满足区域生长条件的门限，这里选取：

$$v_0 = \frac{\max(v)}{m} \quad (6-74)$$

实验发现， m 可以在 5~8 之间任意选择，对最终的分割结果影响不大。

3) 生长方式的确定：以像素作为基本单位逐行扫描进行操作。扫描流程如图 6-6 所示。

采用区域生长的图像分割算法，不仅避免了采用直方图统计阈值分割方法需要复杂的处理步骤，而且还避免了采用全局阈值分割图像时引入背景噪声的

问题。这样提取的目标感兴趣区域更准确。

6.2.2.2 可见光图像区域分割

针对可见光图像难以获得目标区域的问题，本章利用红外图像能检测目标的优势，利用红外图像确定目标区域，然后利用图像配准，将红外图像的目标区域映射到可见光图像中的目标区域。

本章提取的目标感兴趣区域为红外分割后目标轮廓的最小外接矩形。最小外接矩形的确定方法是对二值化后的红外目标图像进行扫描，找到最左边和最右边，最上边和最下边的坐标，这样构成一个矩形。为了避免目标的边缘模糊，在宽中分别增加了 5 个像素点。经过区域分割后的可见光和红外图像的感兴趣区域如图 6-7 所示。

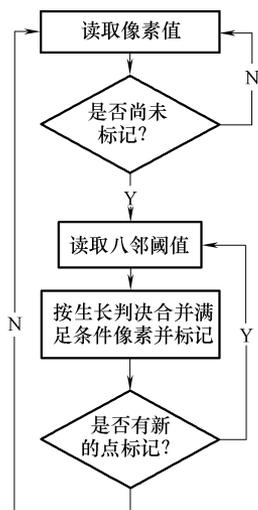


图 6-6 区域生长法流程

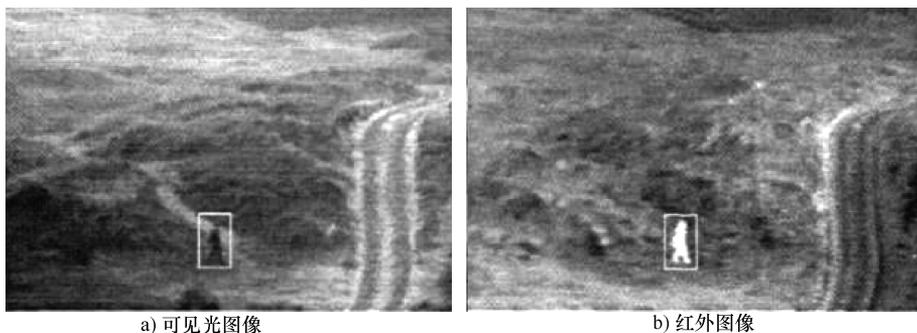


图 6-7 可见光和红外图像

6.2.3 特征提取与融合

特征提取是目标分类识别的前提，特征提取的好坏，直接影响到识别结果。选取时应遵循的原则如下：

- 1) 在保证目标正确识别的前提下，选取尽可能少的特征参数。
- 2) 尽量选取计算量小，准确率高的特征参数。
- 3) 选取的特征量鲁棒性要高，即具有平移、比例和旋转等不变性。
- 4) 特征量之间的相关性要尽量小。

在特征提取的过程中，提取的特征量是具有独立性、可靠性、数据量少和可区别性的特点，这样才能增强目标识别系统的可靠性和准确性。

根据以上原则，可以提取红外图像的形状特征和可见光图像的边缘特征。

特征融合，即把从不同传感器提取的特征，通过某种算法，重新组合成一个新的特征向量，新的特征向量用作后续目标分类和识别的判断依据。

特征级融合算法可分为两大类：特征选择和特征组合。把所有的特征量放在一起，用某种方法产生一个新的特征向量，新向量中的元素都是从原向量选择得到的，称为特征选择，例如，遗传算法。将所有向量直接组合成新向量，称为特征组合，例如，串行和并行融合策略。

(1) 串行融合策略

设 A 和 B 是在样本模式空间 Ω 的两个特征空间。对于任意样本 $\Gamma \in \Omega$ ，相应的特征表示向量为 $A \in \alpha$ 和 $B \in \beta$ 。串行融合策略将这两个特征表示向量串成了一个复向量 γ ，公式如下：

$$\gamma = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{6-75}$$

由式 (6-75) 可知，若 α 是 n 维的， β 是 m 维的，那么合成的向量 γ 为 $(m+n)$ 维的。因此，所有由串行融合而成的向量，是一个 $(m+n)$ 维的新特征空间，后续的分类识别就是在这个新特征空间中进行的。

(2) 并行融合策略

设 A 和 B 是在样本模式空间 Ω 的两个特征空间。对于任意样本 $\Gamma \in \Omega$ ，相应的特征表示向量为 $A \in \alpha$ 和 $B \in \beta$ 。并行融合策略将这两个特征表示向量合成了一个复向量 γ ，公式如下：

$$\gamma = \alpha + i\beta \tag{6-76}$$

式中， i 为虚数单位。需要注意，如果 α 和 β 维数不一致，那么需要对低维的向量补 0。例如， $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ ， $\beta = (\beta_1, \beta_2)^T$ ，则首先将 β 变为 $\beta = (\beta_1, \beta_2, 0)^T$ ，然后合成向量 $\gamma = (\alpha_1 + i\beta_1, \alpha_2 + i\beta_2, \alpha_3 + i0)$ 。

在 Ω 上定义一个并行融合的特征空间 $C = \{\alpha + i\beta \mid \alpha \in A, \beta \in B\}$ 。显然，这是一个 n 维的复向量空间，其中 $n = \max(\dim A, \dim B)$ 。在这个空间里，内积可定义为

$$(X, Y) = X^H Y \tag{6-77}$$

式中， $X, Y \in C$ ， H 表示共轭转置。

定义了以上内积的复向量空间，被称为酉空间。在酉空间可引入以下范数：

$$\|Z\| = \sqrt{Z^H Z} = \sqrt{\sum_{j=1}^n (\alpha_j^2 + \beta_j^2)} \tag{6-78}$$

式中， $Z = (\alpha_1 + i\beta_1, \alpha_2 + i\beta_2, \dots, \alpha_n + i\beta_n)^T$ 。相应地，复向量 Z_1 和 Z_2 之间的距离可定义为

$$\|Z_1 - Z_2\| = \sqrt{(Z_1 - Z_2)^H (Z_1 - Z_2)} \tag{6-79}$$

并行融合策略与串行策略相比，降低了融合后向量的维数。更重要的是，

它引入了酉空间的概念，从而把两个实向量空间的融合问题转化为一个复向量空间的数学问题。

(3) 遗传算法

遗传算法模仿了生物的进化过程。该算法将问题的可能解编成 0、1 代码串，称为染色体。若给定一组初始的染色体，遗传算法就会利用遗传算子对其进行操作，产生新一代染色体。新一代染色体可能包含了较前代更好的解。每一条染色体都要通过适应度函数去评价其适宜程度，遗传算法的目标是找到最适宜的染色体。

遗传算法主要由四部分组成：遗传算子、编码机制、控制参数、适应度函数。

编码机制是遗传算法的基础。遗传算法不是对研究对象直接进行讨论，而是通过某种编码机制把对象统一赋予由特定符号按一定顺序排成的串。在常用的遗传算法中，染色体由 0 与 1 组成，码为二进制串。对遗传算法的码可以有十分广泛的理解。在优化问题中，一条染色体对应于一个可能解。

在遗传算法中，用适应度函数来描述染色体的适宜程度，即根据其适应度来评估染色体优劣。

遗传算法最重要的算子有：选择、交叉、变异。选择的作用是根据染色体的优劣程度决定它在下一代是被淘汰还是被复制。交叉算子是让不同的染色体可以进行信息交换。变异算子就是改变染色体的某个位置上的值。

在实际操作过程中，为提高选优的效果，需先适当地确定某些参数的取值。例如，每一代的群体大小、交叉率和变异率，此外还有遗传的代数，或其他可供确定中止繁殖的指标。

例如，假设 α 和 β 分别表示某一目标的两类不同特征。通过遗传算法，可以得到融合的特征向量 γ ，表达式如下：

$$\gamma = f(\alpha, \beta, x) \quad (6-80)$$

式中， x 为最优染色体。 x 的每一位与特定位置的特征成分相关，该位的取值决定了这个位置的特征成分是从 α 选择（值为 1）还是从 β 选择（值为 0）。

6.3 本章小结

随着科技发展，多传感器应用越来越多，是未来发展的趋势。而目标识别是多传感器应用的重要领域。特征提取是目标识别的关键技术，对目标能否有效识别起决定性作用。近年来对单一传感器的特征进行提取，已有很多学者在研究，并取得了很大的进展，但是对多种传感器进行特征提取和融合的研究并不多，有一些学者和院校刚开始研究。本章针对红外和可见光两种常用的传感器，研究了图像目标的特征提取方法。本章的主要工作及创新点如下。

1) 首先介绍了常用的特征提取方法, Hu 矩和仿射不变矩。针对组合多项式包含冗余信息, 并且不能有效地提取目标的细节信息, 研究了 Zernike 矩和小波矩。介绍了利用正交性去除冗余信息的 Zernike 矩, 这些方法都只能提取图像的全局信息, 进而研究了小波矩的特征提取方法, 利用小波变换的特征不仅能提取图像的全局信息, 还能提取图像的局部信息。对于这些只能满足平移、尺度和旋转不变性, 不能满足由于拍摄视角不同而使采集的图像有仿射变换的问题, 提出了把小波矩和放射不变矩结合, 组成新的特征向量, 进行目标特征提取, 提高了目标识别系统的鲁棒性。

2) 针对红外和可见光的特点, 研究了红外图像形状特征提取和可见光图像的边缘特征提取。红外传感器获取的图像目标和背景的差异明显, 讨论了红外图像的阈值分割算法, 主要研究了二值化处理后图像的各种不同类型目标的不同矩特征提取方法, 进行特征提取实验, 对实验结果进行分析; 可见光传感器能获取图像目标的细节信息, 讨论了可见光图像的边缘检测算子。重点研究了不同类型目标的边缘特征提取方法, 分析了共生矩阵的四个特征量(能量、熵、逆差矩和惯性矩), 并进行特征提取仿真实验。还对可见光图像的边缘矩特征进行提取, 分析不同矩特征提取的结果。

3) 针对红外和可见光图像传感器的互补优势, 研究了红外和可见光图像的特征提取和融合方法。主要针对单一传感器对地面目标特征提取的不稳定性问题, 提出了一种基于感兴趣区域进行特征提取的方法。首先分析红外和可见光图像的特点, 利用红外图像的目标高对比度特性, 提取出图像中的感兴趣区域, 然后通过空间投影映射在可见光图像中找出同样的区域。其次分别对感兴趣区域可见光图像进行边缘特征提取和红外图像进行形状特征提取, 把提取的特征进行融合, 用于目标识别。本章提出的算法得到的融合特征, 不仅包含目标的外部形状特征, 还包括目标的内部边缘特征, 得到的目标特征更稳健, 对后续的目标分类识别提供更好的判决依据。

基于协方差矩阵多特征信息融合

特征信息融合在整个目标识别过程中处于一个中间核心位置，在获取了有效的目标特征数据之后，还需要对目标特征信息之间的相关性进行综合分析和进行有效的特征融合。在查阅了众多融合算法之后，可以通过构造有效的图像特征的协方差矩阵来对目标的特征信息进行融合，所以本章采用基于协方差矩阵的特征信息融合方法。与传统的串联特征信息融合相比，通过利用构造的协方差矩阵可以对不同的特征信息数据进行融合可以对获取的目标特征数据进行降维，而且获得的特征之间的相关性更小。

7.1 图像特征提取

7.1.1 多传感器特征提取

多传感器特征提取是指通过目标识别系统外部的各种类型的传感器获取的图像数据进行特征提取，这样获取的特征数据通过融合对目标特征的信息描述更准确、更可靠、更全面、更有效，从而使得最后提取的目标特征受周边环境、时间和空间角度影响降到最低，传统的单一传感器对多传感相比一点优势都不具备，通过多传感器目标特征提取就是为了在相同的环境能够对目标进行更精确，更可靠的特征提取。

多传感器还有另外一个好处就是对采集到的目标图像信息能够进行互补。通过这些互补的信息，不仅降低了识别系统对其中某种传感器过度的依赖性，而且由外到内提高了目标识别系统的性能。在实际应用中会针对不同的环境配置不同的组合传感器，这样能够充分发挥各个传感器在系统中的作用。在图像目标采集的过程中，受环境的影响是最为显著的，例如大雨、大雾、浓烟等情况下因为能见度低，如果利用可见光来采集图像数据会比较困难，但搭配雷达对目标进行探测，效果相对于可见光传感器会好很多，虽然在采集过程会有信号衰减，但是在一定程度上与可见光传感器进行了信息互补。在夜晚或者亮度比较低的环境下，此时最好是利用红外传感器对目标进行图像信息采集。因为红外传感器主要是根据物体自身向外辐射的能量来获取目标的形状特征信息，

此时可以获取目标的体型特征。

在实际应用中，最常用的就是可见光传感器和红外传感器。它们对环境的适应是互补的，利用它们彼此的成像原理不同，可见光传感器主要是在能见度正常或者能见度高的情况下效果比较好，获取的目标图像细节纹理比较清晰，能够帮助系统对其进行准确有效的特征数据提取，红外传感器一般根据目标自身向外辐射的能量不同对其进行成像，在亮度低的环境下对目标信息采集具有很大的优势，对目标的外形具备很好的成像效果，但是所获取的图像不够清晰，根据上述可知它们在现实应用中是最理想的组合搭配。此时如果想要获取目标相应地理信息，可以通过 GPS 对其进行定位，来获取其精确的经纬度信息，这就是利用多传感器特征提取的巨大优势，即全面，精确，不受时空限制。

7.1.2 灰度特征

利用系统外部传感器采集到的图像都为彩色图像，一般情况下图像的色彩都是由红 (R)、绿 (G)、蓝 (B) 三种颜色组合而成，它们是我们人眼所能感知到的颜色，也是目前应用比较广泛的原因。而我们所需要的灰度特征就是将图像中的像素用其中某一单色形成灰度图，然后再对其进行灰度特征提取。因为彩图和灰度图都能客观真实地反映图像中的目标信息，而且通过对灰度图的特征提取可以减少相应的特征数，从而帮助我们降低计算量来提高效率。在转换的过程中，并不会因为灰度化了就降低了图像的亮度，亮度只是人眼视觉对颜色的一种敏感度，所以对彩图进行灰度化不会影响对目标特征提取的有效性。

在灰度转换的过程中，一般通过颜色空间转换将源图像从空间 RGB 转化到空间 YCbCr，在空间 YCbCr 中用 Y 分量来表示图像的灰度，其实它也是图像的亮度信息，其中的 Cb 分量和 Cr 分量表示转换后灰度图像的色度信息。下面是 RGB 转 YCbCr 的转换公式：

$$\begin{cases} Y = 0.299R + 0.587G + 0.114B \\ Cb = -0.172R - 0.339G + 0.511B + 128 \\ Cr = 0.511R - 0.428G - 0.083B + 128 \end{cases} \quad (7-1)$$

根据上述转换公式可知最后的灰度转换公式如下：

$$\text{Gray} = 0.229R + 0.587G + 0.114B \quad (7-2)$$

7.1.3 空间信息特征

在理论运用过程中，简单的利用数学中的统计特性来对目标特征进行描述是不够全面的，这样处理导致的结果就是没有足够高的准确率。根据统计特性对目标特性描述的不足，可以在融合的过程中添加相应的空间位置信息，这样对目标的特征信息描述会更准确，由于上述要对传感器采集的图像进行灰度化处理，对灰度后图像其像素是单一的，所以在本章中，将使用距离 $d(x, y)$ ，即

图像中的像素点与在整个图像的中心区域的距离，利用这样的距离来作为我们所需要的空间位置信息，加入到特征信息融合中去，其可以表示为

$$d(x, y) = \sqrt{x'^2 + y'^2}, (x', y') = (x - x_0, y - y_0) \quad (7-3)$$

在式 (7-3) 中，我们用 (x', y') 来表示图像中任意一个像素点，而 (x_0, y_0) 则表示图像中的中心区域坐标点。在本章中不考虑像素点与图像中目标中的方向角，通过对目标的长和宽进行归一化处理，从而获取的目标特征是与目标的尺寸不相关的。

7.1.4 梯度信息特征

梯度特征是图像的基本特征之一，它主要是体现目标图像中的边缘信息，并且它具有对光不敏感的特性。在对图像的边缘进行检测时，首先会检测像素点八邻域灰度变化。通过数学公式推导可知，一般情况下梯度算子就是通过对图像进行一阶求导即可，但是当我们想对一个二维图像求梯度算子时，这时我们对图像求导时会存在方向性。一般都是从简单到复杂，所以方向可定为 x 方向和 y 方向这两个方向梯度特征。一般我们通常使用的梯度算子有 Sobel 算子、Canny 算子和拉普拉斯算子等。

假设图像中的像素可以用 $p_{i,j}$ 来表示，如果使用上述常用的 Sobel 算子来作为梯度算子，则其提取的梯度信息向量可以表示为

$$V_{i,j} = \{dx_{i,j}, dy_{i,j}\} \quad (7-4)$$

梯度信息向量中的 $dx_{i,j}$ 和 $dy_{i,j}$ 分两个梯度方向来表示，即如图 7-1 中的水平梯度方向 G_x 和垂直梯度方向 G_y 。

-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1

a) 垂直 G_y 方向
b) 水平 G_x 方向

图 7-1 Sobel 算子

然后通过上述的算子进行相应的卷积运算，这样就能够获取 Sobel 算子在水平方向和垂直方向上对应的梯度值。

x 方向梯度幅度值：

$$\Delta f_x(x, y) = \{f(x+1, y-1) + f(x+1, y+1) + 2f(x+1, y)\} - \{f(x-1, y-1) + f(x-1, y+1) + 2f(x-1, y)\} \quad (7-5)$$

y 方向梯度幅度值：

$$\Delta f_y(x, y) = \{f(x-1, y+1) + f(x+1, y+1) + 2f(x, y+1)\} - \{f(x-1, y-1) + f(x+1, y-1) + 2f(x, y-1)\} \quad (7-6)$$

$$f(x, y) = \sqrt{(\Delta f_x(x, y))^2 + (\Delta f_y(x, y))^2} \quad (7-7)$$

7.2 协方差矩阵的构造

7.2.1 协方差矩阵与协方差

存在一组随机变量为 X_1, X_2, \dots, X_N , 假设能够通过将上述的随机变量来随机地组成一组随机的向量 $X = [X_1, X_2, \dots, X_N]^T$, 在这组由随机变量构成的向量中, 假设其中每个变量都有 m 个有效样本, 这样就可以得到相应的样本矩阵:

$$M = [\vec{\alpha}_1, \vec{\alpha}_2, \dots, \vec{\alpha}_n]^T = [\vec{\beta}_1, \vec{\beta}_2, \dots, \vec{\beta}_m] \quad (7-8)$$

在式 (7-8) 中可以用矩阵参数 $\vec{\beta}_j$ ($j=1, 2, \dots, m$) 来表示每个随机向量 X 中的样本向量, 随机变量中样本所构成向量用 $\vec{\alpha}_i$ ($i=1, 2, \dots, n$) 来表示。

根据上述的随机变量 X_i, X_j , 可以将它们两者之间的协方差表示为

$$c_{ij} = E\{[X_i - E(X_i)][X_j - E(X_j)]\} \quad (7-9)$$

通过已经给出的样本值, 这样就能对它们的协方差进行有效的估计, 得到的估计值如下:

$$c_{ij} = \frac{1}{m} \sum_{k=1}^m \left[\left(M_{ik} - \frac{1}{m} \sum_{a=1}^m M_{ia} \right) \left(M_{jk} - \frac{1}{m} \sum_{b=1}^m M_{jb} \right) \right] \quad (7-10)$$

对式 (7-10) 进一步进行化简为

$$\begin{aligned} c_{ij} &= \frac{1}{m} \sum_{k=1}^m M_{ik} M_{jk} - \frac{1}{m^2} \sum_{a=1}^m M_{ia} \sum_{b=1}^m M_{jb} \\ &= \frac{1}{m} \vec{\alpha}_i^T \cdot \vec{\alpha}_j - \frac{1}{m^2} \sum_{a=1}^m M_{ia} \sum_{b=1}^m M_{jb} \end{aligned} \quad (7-11)$$

此时, 我们所需要的协方差矩阵就能够表示如下:

$$\begin{aligned} \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix} &= \begin{bmatrix} \frac{1}{m} \vec{\alpha}_1^T \cdot \vec{\alpha}_1 - \frac{1}{m^2} \sum_{a=1}^m M_{1a} \sum_{b=1}^m M_{1b} & \dots & \frac{1}{m} \vec{\alpha}_1^T \cdot \vec{\alpha}_n - \frac{1}{m^2} \sum_{a=1}^m M_{1a} \sum_{b=1}^m M_{nb} \\ \frac{1}{m} \vec{\alpha}_2^T \cdot \vec{\alpha}_1 - \frac{1}{m^2} \sum_{a=1}^m M_{2a} \sum_{b=1}^m M_{1b} & \dots & \frac{1}{m} \vec{\alpha}_2^T \cdot \vec{\alpha}_n - \frac{1}{m^2} \sum_{a=1}^m M_{2a} \sum_{b=1}^m M_{nb} \\ \vdots & \ddots & \vdots \\ \frac{1}{m} \vec{\alpha}_n^T \cdot \vec{\alpha}_1 - \frac{1}{m^2} \sum_{a=1}^m M_{na} \sum_{b=1}^m M_{1b} & \dots & \frac{1}{m} \vec{\alpha}_n^T \cdot \vec{\alpha}_n - \frac{1}{m^2} \sum_{a=1}^m M_{na} \sum_{b=1}^m M_{nb} \end{bmatrix} \\ &= \frac{1}{m} M M^T - \frac{1}{m^2} \begin{bmatrix} \sum_{a=1}^m M_{1a} \\ \sum_{a=1}^m M_{2a} \\ \vdots \\ \sum_{a=1}^m M_{na} \end{bmatrix} \begin{bmatrix} \sum_{a=1}^m M_{1a} & \sum_{a=1}^m M_{2a} & \dots & \sum_{a=1}^m M_{na} \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{m} [\vec{\beta}_1, \vec{\beta}_2, \dots, \vec{\beta}_m] [\vec{\beta}_1, \vec{\beta}_2, \dots, \vec{\beta}_m]^T - \frac{1}{m^2} [\vec{\beta}_1 + \vec{\beta}_2 + \dots + \vec{\beta}_m] \\
 &\quad [\vec{\beta}_1 + \vec{\beta}_2 + \dots + \vec{\beta}_m]^T \\
 &= \frac{1}{m} \sum_{i=1}^m [\vec{\beta}_i - \vec{\beta}_0] [\vec{\beta}_i - \vec{\beta}_0]^T \tag{7-12}
 \end{aligned}$$

式中, $\vec{\beta}_0 = \frac{1}{m} [\vec{\beta}_1 + \vec{\beta}_2 + \dots + \vec{\beta}_m]$, 由式 (7-12) 可对其进行进一步化简, 最后我们可以得到相应的最简式, 即我们所需要的协方差矩阵的表达式。

如果其所有样本存在平均值为零向量, 那么式 (7-12) 就能够进一步化简为

$$C = \frac{1}{m} \sum_{i=1}^m \vec{\beta}_i \vec{\beta}_i^T \tag{7-13}$$

根据上述对协方差和协方差矩阵的推导和对其值进行有效的估计, 我们必须清楚地知道根据公式推导得到的协方差的真实值只是一个估计值, 它是依靠大量的样本数据得到的估计值, 如果我们的样本数据越多, 样本在整体中的占比就会越大, 最后获得的协方差矩阵就会越可靠。

在协方差矩阵中, 每个变量都是随机向量 X 分量间不同的协方差, 这里的协方差不是样本之间的协方差, 在协方差矩阵如果用 c_{ij} 来表示矩阵中的元素, 则它代表的就是随机变量 X_i 与 X_j 之间的协方差。我们利用协方差主要是为了对变量间的二阶统计特性进行表示, 在随机变量中, 如果它们的不同变量之间的相关性比较小, 那么最终所得到的协方差矩阵类似于一个对角矩阵。

7.2.2 协方差矩阵距离的度量

在矩阵理论发展的历史中, 如何比较两个矩阵是一个非常经典的问题, 在我们传统的欧式空间运算中, 因为协方差在其运算中封闭性比较差, 所以不能在传统的欧式空间中解决协方差矩阵之间的度量问题, 所以通过查阅了一些文献后, 这里主要借鉴一种解决协方差直接距离度量的方法, 如下:

$$d(C_1, C_2) = \sqrt{\sum_{i=1}^n \ln^2 \lambda_i(C_1, C_2)} \tag{7-14}$$

在式 (7-14) 中, $\{\lambda_i(C_1, C_2)\}_{i=1,2,\dots,n}$ 是一个特征方程, 它是由 C_1 和 C_2 特征值组成的, 其中这些特征都是正定矩阵, 它具有下面几个特性:

- 1) 正定性 $d(C_1, C_2) \geq 0$, 其中等号当且仅当 $C_1 = C_2$ 时才会成立。
- 2) 对称性 $d(C_1, C_2) = d(C_2, C_1)$ 。
- 3) 三角不等式 $d(C_1, C_2) + d(C_1, C_3) \geq d(C_2, C_3)$ 。

根据式 (7-14), 这里存在式 (7-15):

$$\det(\lambda C_1 - C_2) = 0 \tag{7-15}$$

通过式 (7-15) 再对其进行有效的简化推导可得下式:

$$\lambda C_1 - C_2 = \sqrt{C_1}(\lambda E - \sqrt{C_1^{-1}C_2}\sqrt{C_1^{-1}})\sqrt{C_1} \quad (7-16)$$

因为 C_1 和 C_2 都属于正定矩阵, 所以在式 (7-16) 中 $\sqrt{C_1^{-1}C_2}\sqrt{C_1^{-1}}$ 的特征值都是正定数。所以可以通过利用雅克比过关法来对式 (7-16) 解得特征值, 其中最关键的是需要将 $\lambda_1(C_1, C_2), \dots, \lambda_d(C_1, C_2)$ 通过转换为正定实矩阵来进行计算, 这样最终才能正确且更少计算量地求得特征值。

图 7-2 是协方差矩阵的一个构造流程图。

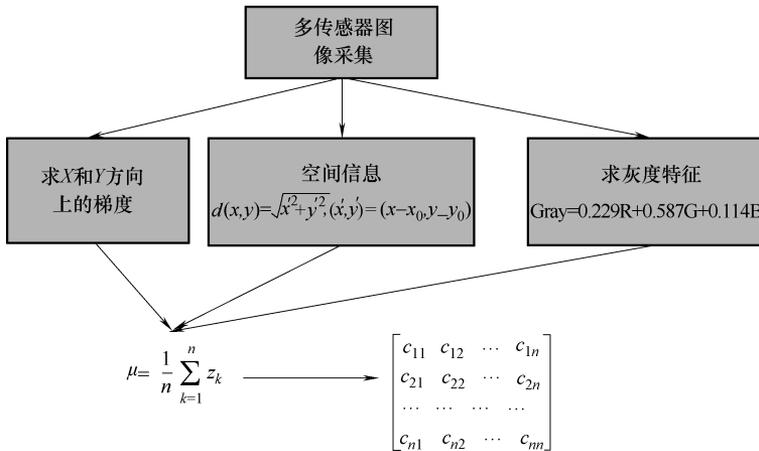


图 7-2 协方差矩阵的构造流程

7.2.3 特征协方差

假如通过外部传感器获得的图像 I 为二维数字图像, 此时可以假设目标在该图像区域 R 中的大小为 $M \times N$, 那么接下来在区域 R 内特征点就能够用特征向量 $\{z_k\}_{k=1,2,\dots,n}$ 来表示, 其中 n 代表在区域 R 中的特征点数, 在特征向量中, 它的变量种类信息很多, 不仅能够用图像中像素的亮度、纹理、梯度, 而且还用一些滤波器对图像中像素的响应值来表示。根据上述假设的目标区域 R , 它的协方差矩阵 Q_R 能够通过使用一个大小为 $d \times d$ 矩阵对其定义:

$$Q_R = \frac{1}{n-1} \sum_{k=1}^n (z_k - \mu)(z_k - \mu)^T \quad (7-17)$$

式中, $\mu = \frac{1}{n} \sum_{k=1}^n z_k$ 。

利用协方差矩阵的好处就是它不仅可以将图像区域中的有效目标特征信息进行融合, 其他的还能将一些图像中的表象特征信息包含在其中, 比如图像直方图的特性。在上述过程中我们能够知道协方差矩阵 Q_R 其实是一个大小

为 $d \times d$ 的实对称矩阵，在几何原理中也称为正定矩阵，我们知道正定矩阵的特点是在它的两条对角线上的分量都是互相独立的图像特征量，在矩阵其他区域的值就是代表与其具有相关性关系的特征量，协方差矩阵 Q_R 之前设它的大小为 $d \times d$ ，其中它的大小只跟我们所获取的特征向量的维数 d 有关系，其他的如目标区域都跟它没有关系，这点是非常重要的。通过对协方差矩阵中多特征值进行平均值运算，它们中的一些特征噪声就会被滤除，从而使得协方差矩阵具有很好的鲁棒性。比如在传感器对目标进行图像信息采集时，由于受到外界光照强度的干扰，会使得获取的目标图像颜色信息发生变化，但是这不会影响到协方差特征值，所以目标采集数据不会受很大的影响。

7.3 多特征信息融合

由于在这个识别系统中，主要利用的是多传感器进行图像数据采集，而且其中负责图像信息采集的传感器主要为可见光和红外传感器，在前面比较详细地介绍了它们采集到的图像的特点后，所以可分两块对图像目标特征进行特征提取，通过上述介绍的特征提取算法进行相应的特征提取，然后对图像中目标所在区域中的特征向量的均值进行有效计算获取，最后通过式 (7-17)，来对所获取的目标特征向量进行融合，得到我们所需要的多特征信息融合协方差矩阵。

7.4 融合实验结果与分析

上述详细地介绍了利用协方差矩阵多特征融合算法的每个详细步骤公式，下面就验证此算法的可靠性和有效性，通过对它进行基于 Visual C++ 6.0 编程和具体图像数据仿真来进行试验。在本次仿真实验过程中，因为涉及的图像数据比较多，考虑到复杂度和易操作性，我们选取了坦克、越野车和直升机来作为图像目标，以下的图片都是通过识别系统外部可见光和红外传感器真实采集到的图像信息进行特征融合处理，对采集到的源图片进行预处理去噪增强，通过训练和建模建立图像目标的特征数据库，在实验中，考虑到数据非常大，暂时每组 200 张图片，其中可见光和红外各 100 张，尺寸均为 128 像素 \times 128 像素。在图 7-3 中显示的就是数据库中的一些图片，其中上面是可见光图片，下面为红外图片。

根据上述对可见光和红外图像特征提取可分为八个维度对目标特征进行提取。通过在 Visual C++ 6.0 上进行编程处理得到了相应的特征值，见表 7-1，它是图 7-3a 的部分特征值数据。

通过对图 7-4a、b、c、d 进行不同方向的旋转、侧翻和缩放的图像协方差特

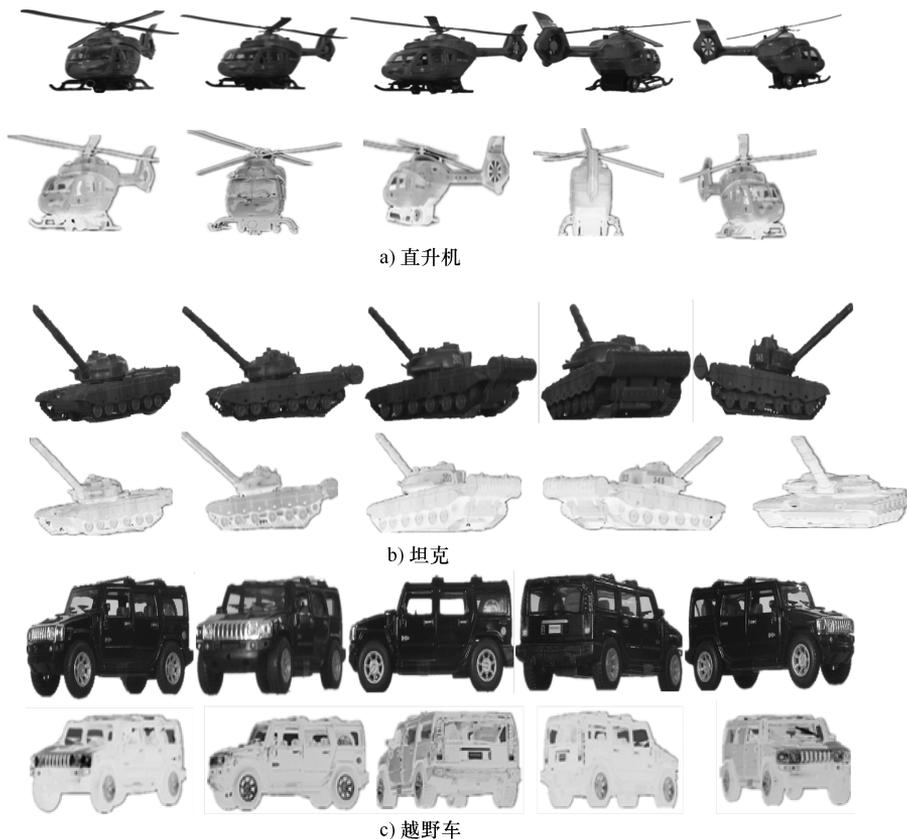


图 7-3 融合的部分类别图像

征值测试。然后求得它们的特征协方差曲线图如图 7-5 所示，根据图中曲线的变化可以看出，对于图像中目标的旋转、侧翻和适度缩放对其特征值的影响非常小，所以能得出特征协方差矩阵具备非常强的旋转不变性、尺度不变性和适应鲁棒性的结论。

表 7-1 图 7-3a 中部分图像的特征值

	可见光图像特征					红外图像特征		
	Gray	G_x	G_y	$\frac{\partial^2}{\partial x}$	$\frac{\partial^2}{\partial y}$	x	y	s
1	76	57	90	16	3	120	29	0.17832
2	79	40	19	7	6	119	38	0.20463
3	157	27	31	2	17	57	79	0.31031
4	49	6	5	5	4	151	83	0.40307

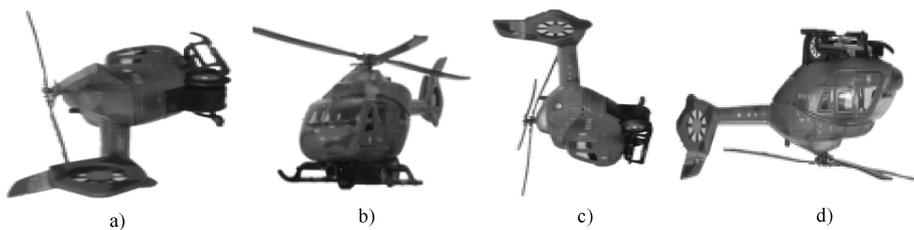


图 7-4 测试图像

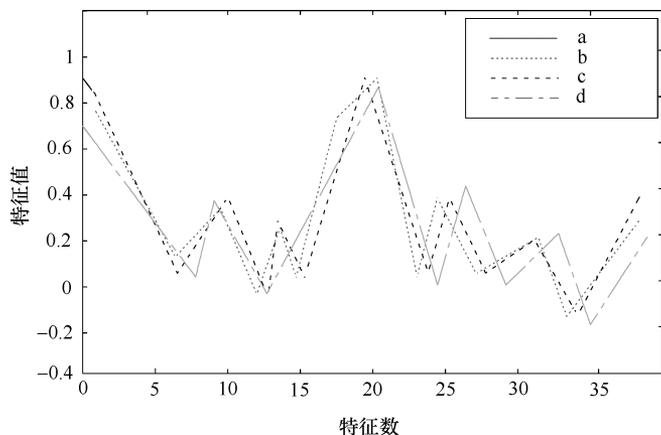


图 7-5 协方差矩阵特征值曲线

根据上面实验仿真数据可以总结出利用协方差矩阵对目标特征信息融合的优点如下：在公式计算和推导中我们能够清楚地知道，协方差矩阵的大小只和我们提取目标特征向量的维数有关系，如果我们想要再对目标特征进行更多的特征添加也只需要加上相应的维数就行，从而有效地降低了相应的计算量；在后面目标图像进行旋转、侧翻和缩放测试中，协方差矩阵获得特征值基本是不变或者相差非常小，从而从实验到理论验证了协方差矩阵具备非常好的旋转不变性、尺度不变性和鲁棒性。

7.5 本章小结

本章主要研究了基于协方差矩阵多特征信息融合算法。在本章主要介绍了该算法中对灰度化特征、图像目标区域的坐标、梯度算子、图像二阶导和空间位置信息等目标特征提取方法，然后详细描述了对协方差进行可靠有效的构造，协方差矩阵的推导和距离度量等理论，然后通过协方差矩阵对目标多特征进行

有效融合，最后我们通过理论加实践的方法对其进行可行性和有效性验证，实验结果表明利用协方差矩阵能够有效可靠地对目标的多特征进行融合处理，而且还显示了利用协方差矩阵，使得提取的特征值具有非常优异的旋转不变性、尺度不变性和鲁棒性。

基于主成分分析的特征融合方法

多元统计分析是统计学中一个非常重要的分支，它主要研究客观事物中多个变量（或多个因素）之间相互依赖的统计规律性。如果每个个体都有多个观测数据，或者从数学角度上说，如果观测数据能表示为 p 维欧氏空间的点，那么这样的数据叫作多元数据，而分析多元数据的统计方法就叫作多元统计分析。多元统计分析可以同时处理异度量的多个变量间的各种关系，能够较好地实现数据降维且损失很少的原有信息量。所以某些多元统计分析方法已经成为数据融合分析的一种有效工具。主成分分析是一种从数据的相关性入手，将多个变量通过线性变换以选出少数重要变量的多元统计分析方法。在本章中，我们将介绍主成分分析原理和方法，论述如何运用主成分分析方法实现红外与可见光图像的特征融合。

8.1 特征提取

8.1.1 灰度共生矩阵

灰度共生矩阵是建立在估计图像的二阶组合条件概率密度函数基础上的统计方法，主要描述纹理基元或局部模式随机和空间统计特征，以表示区域的一致性及其区域间的相对性。它能反映出图像灰度关于方向、相邻间隔、变化幅度的综合信息，是当前人们公认的一种重要的纹理分析方法。

灰度共生矩阵描述的是在某方向上间隔一定距离的一对像素点对灰度出现的统计规律，即从图像灰度值为 i 的像元 (x, y) 出发，统计与其距离为 d 、方向为 θ 、灰度值为 j 的像元 $(x+a, y+b)$ 同时出现的概率 $p(i, j, d, \theta)$ ，其数学表达式为

$$p(i, j, d, \theta) = \{ [(x, y), (x+a, y+b)] | f(x, y) = i, f(x+a, y+b) = j \} \quad (8-1)$$

式中， $x=0, 1, 2, \dots, N_x$ ， $y=0, 1, 2, \dots, N_y$ ， $i \in [0, L-1]$ ， $j \in [0, L-1]$ 。 θ 为灰度共生矩阵的生成方向，通常取 0° 、 45° 、 90° 、 135° 四个方向。 N_x 、 N_y 分别表示水平和垂直方向的像素总值， L 为图像的灰度级。

以下是一个灰度共生矩阵的生成例子。图 8-1a 所示是一个灰度级为 4 的图

像。由于图像的灰度级决定了共生矩阵的大小，所以共生矩阵为 4×4 的矩阵。行表示 i 的灰度变换，列表示 j 的灰度变换。由式 (8-1) 可以计算出 $d=1$ 时 0° 、 90° 、 135° 和 45° 的灰度共生矩阵 P_H 、 P_V 、 P_{LD} 、 P_{RD} ，如图 8-1b 所示。

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

a) 图像灰度值

$$P_H = \begin{bmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}, P_V = \begin{bmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}, P_{LD} = \begin{bmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}, P_{RD} = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

b) $d=1$ 时 0° 、 90° 、 135° 和 45° 的灰度共生矩阵

图 8-1 图像灰度值和灰度共生矩阵

Haralick 等人定义了 14 个用于纹理分析的灰度共生矩阵特征参数，虽然这 14 种特征参数都能表达纹理的某些特定信息，但存在信息冗杂、重复表述的问题。根据各个参数的意义，选用以下四个参数：

能量

$$W_1 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p^2(i, j, d, \theta) \quad (8-2)$$

熵

$$W_2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i, j, d, \theta) \log[p(i, j, d, \theta)] \quad (8-3)$$

惯性矩

$$W_3 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i-j)^2 p(i, j, d, \theta) \quad (8-4)$$

局部平稳性

$$W_4 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{1}{1 + (i-j)^2} p(i, j, d, \theta) \quad (8-5)$$

式中， L 为图像灰度级，为保证参数的旋转不变性，取 0° 、 45° 、 90° 和 135° 四个方向参数的均值作为灰度共生矩阵的值。

8.1.2 Hu 不变矩

概率密度分布函数为 $f(x, y)$ 的二维连续随机函数的 $(p+q)$ 阶几何矩定义为

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad (8-6)$$

由式 (8-6) 可以看出, m_{pq} 的值是由 $f(x, y)$ 唯一确定的。由于 m_{pq} 不具备平移不变性, 于是定义其 $(p+q)$ 阶中心矩为

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (8-7)$$

式中, $\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}, p, q = 0, 1, 2, \dots$ 。

对于一幅大小为 $M \times N$ 的数字图像 $f(x, y)$, 积分用求和代替, 二维 $(p+q)$ 阶几何矩和中心矩定义为

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y) \quad (8-8)$$

$$u_{pq} = \sum_{x=1}^M \sum_{y=1}^N (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (8-9)$$

式中, $p, q = 0, 1, 2, \dots, (\bar{x}, \bar{y})$ 为灰度图像 $f(x, y)$ 的灰度质心, 定义为 $\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$ 。

归一化的中心矩, 用 η_{pq} 表示, 定义为

$$\eta_{pq} = \frac{u_{pq}}{u_{00}^\gamma} \quad (8-10)$$

式中,

$$\gamma = \frac{p+q}{2} + 1 \quad p+q=2,3,\dots \quad (8-11)$$

1962 年, M. K. Hu 利用归一化的二阶和三阶中心矩构造了七个不变矩, 它们在连续图像条件下可保持平移、缩放和旋转不变性, 具体定义如下:

$$\phi_1 = \eta_{20} + \eta_{02} \quad (8-12)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (8-13)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (8-14)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (8-15)$$

$$\begin{aligned} \phi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03}) \\ & (\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (8-16)$$

$$\phi_6 = (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (8-17)$$

$$\begin{aligned} \phi_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30}) \\ & (\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (8-18)$$

8.2 几何特征

本章使用的几何特征是根据图像中目标的几何特性、目标与背景之间的关系等来设计的，依据本章需求选择以下几个稳定性较好的特征量。

复杂度：图像中目标边界像素点数与整个目标像素点数的比值。

长宽比：图像中目标最小外接矩形的长度与宽度之比，长宽比可以将较纤细的物体与方形或者圆形区分开来。

均值对比度：图像中目标最小外接矩形中的灰度平均值与图像其余部分灰度平均值的比值。

紧凑度：图像中整个目标像素点数与最小外接矩形内的总像素点数之间的比值。

8.3 主成分分析的定义

主成分分析也称主分量分析，其研究可以追溯到1901年，由 Pearson 首次提出，不过当时只是对非随机变量进行讨论。到1933年，Hotelling 将主成分分析的概念推广到随机向量。Hotelling 对主成分分析的定义如下：对于一个 d 维的观察向量序列 $\{t_n\}$, $n \in \{1, 2, \dots, n\}$ ，主成分分子就是要找到 q 个正交的主方向 w_j , $j \in \{1, 2, \dots, q\}$ ，使得观察向量序列 $\{t_n\}$ 在这 q 个主方向张成的子空间上的投影保留的方差最大。1972年，K. Fukunaga 在其出版的专著《Introduction to Statistical Recognition》中系统地阐述主成分分析的理论与方法。主成分分析可以在损失很少信息的前提下，将原有的多个指标转化为少数的几个综合指标，通常把转化生成的综合指标称为主成分。其中每个主成分都是原始指标的线性组合，并且彼此之间互不相关，这使得主成分比原始指标具有某些更优越的性能。因而，在多指标问题的研究中，可以只考虑少数的几个主成分，这样做不仅不会造成信息的过多损失，而且更易于把握问题的本质，并且揭示事物内部变量之间的规律性，同时使问题得到简化，提高分析效率。

8.3.1 主成分分析的基本原理

主成分分析的思想主要是运用统计学的知识从一个高维空间中的大量数据中提取比较主要的特征，形成新的子空间，实现一个从高维空间到低维空间的转换，从而达到去除数据间的冗余性和相关性，同时又保留了原始数据中最大信息的目的。它借助了一个正交变换，使其分量相关的原始随机向量转化为其分量不相关的新随机向量。这在数学上表现为将原始随机向量的协方差矩阵变换成对角形阵，在几何上表现为将原坐标系变换成新的正交坐标系，使得指向

样本点散布最开的 p 个正交方向，然后对多维变量系统进行降维处理，使之能以一个较高的精度转换成低维变量系统。下面依次对主成分分析的数学原理和几何原理进行阐述。

8.3.2 数学模型

根据方差最大化原理，用一组新的、线性无关且相互正交的向量来表征原数据矩阵的行或者列，这组原数据向量的线性组合即为主成分。

设有 n 个样品，每个样品观测 p 个指标，这 p 个指标构成的 p 维向量分别为 X_1, X_2, \dots, X_p ，得到原始数据资料阵：

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = (X_1, X_2, \dots, X_p) \quad (8-19)$$

式中， $X_i = \begin{bmatrix} X_{1i} \\ X_{2i} \\ \vdots \\ X_{ni} \end{bmatrix} \quad i = 1, 2, \dots, p。$

设数据矩阵 X 的均值为 μ ，协方差矩阵为 Σ 。将 X 的 p 个指标 X_1, X_2, \dots, X_p 作线性组合，形成新的综合指标，即

$$\begin{cases} F_1 = a_{11}X_1 + a_{21}X_2 + \cdots + a_{p1}X_p \\ F_2 = a_{12}X_1 + a_{22}X_2 + \cdots + a_{p2}X_p \\ \cdots \\ F_p = a_{1p}X_1 + a_{2p}X_2 + \cdots + a_{pp}X_p \end{cases} \quad (8-20)$$

简写成：

$$F_i = a_{1i}X_1 + a_{2i}X_2 + \cdots + a_{pi}X_p, i = 1, 2, \dots, p \quad (8-21)$$

若线性变换满足以下约束：

- 1) $a'_i a_i = 1$ ，即 $a_{1i}^2 + a_{2i}^2 + \cdots + a_{pi}^2 = 1$ ，其中 $i = 1, 2, \dots, p。$
- 2) F_i 与 F_j ($i \neq j, j = 1, 2, \dots, p$) 互不相关。
- 3) F_1 是 X_1, X_2, \dots, X_p 的一切线性组合（系数满足上述方程组）中方差最大的， F_2 是与 F_1 不相关的 X_1, X_2, \dots, X_p 一切线性组合中方差最大的， \dots ， F_p 是与 F_1, F_2, \dots, F_{p-1} 都不相关的 X_1, X_2, \dots, X_p 一切线性组合中方差最大的。

那么，称式 (8-20) 为指标 X_1, X_2, \dots, X_p 的主成分分析模型，并将式 (8-20) 确定的综合指标 F_1, F_2, \dots, F_p 分别称为原始指标的第一，第二， \dots ，第 p 个主成分。

主成分，从代数学的观点看就是 p 个随机变量 X_1, X_2, \dots, X_p 的一些特殊的线性组合，而在几何上这些线性组合代表以 X_1, X_2, \dots, X_p 构成的坐标系经过平移、尺度伸缩和坐标旋转产生的一个新坐标系，新坐标轴为样品偏差最大的方向（或者说具有最大的样品方差）。为了能够更直观地理解主成分分析的基本思想，下面以最简单的二元正态变量来讨论主成分分析的几何意义，所得结论可以很容易地扩展到多维的情况。

设有 n 个样品，每个样品有 p 个观察变量，记为 X_1, X_2, \dots, X_p ，它们的综合变量记为 F_1, F_2, \dots, F_p 。当 $p=2$ 时，由变量是 X_1, X_2 组成的坐标空间中， n 个样品散布的情况如图 8-2 所示。

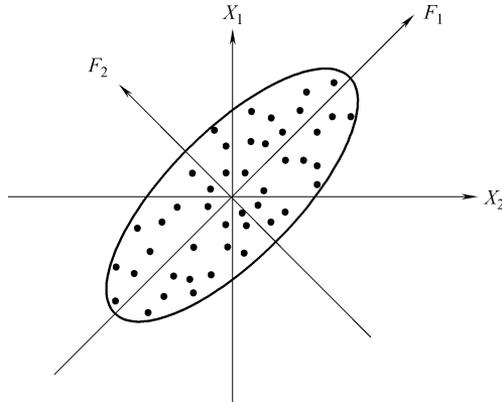


图 8-2 主成分的意义

由图可以看出， n 个分散的点大致形成一个椭圆。若在椭圆长轴方向取坐标轴 F_1 ，在短轴方向取 F_2 ，这相当于将坐标轴按逆时针方向旋转 θ 角度，根据旋转轴变换公式，新老坐标之间有如下关系：

$$\begin{cases} F_1 = X_1 \cos\theta + X_2 \sin\theta \\ F_2 = -X_1 \sin\theta + X_2 \cos\theta \end{cases} \quad (8-22)$$

其矩阵形式为

$$\begin{bmatrix} F_1 \\ F_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = UX \quad (8-23)$$

式中， U 为旋转变换矩阵，由式 (8-23) 可知 U 为正交阵，即满足 $U^T = U^{-1}$ ， $U^T U = I$ 。

从图 8-2 可看出，经过旋转后的 n 个点的波动（可用方差表示）大部分可以归结为在 F_1 轴上的波动，而在 F_2 轴上的波动较小。当图 8-2 椭圆相当扁平的时候，那么我们可以只考虑 F_1 方向上的波动，而忽略 F_2 方向的波动。这样，二维就可以降为一维，只取第一个综合变量 F_1 即可，即椭圆的长轴。

一般情况下， p 个变量组成 p 维空间， n 个样品就是 p 维空间的 n 个点，对 p 元正态分布变量来说，找主成分的问题就相当于找 p 维空间中椭球体的主轴问题。

8.3.3 主成分的推导

在下面推导主成分的过程中，需要用到线性代数中的两个定理：

定理一：若 A 是 $p \times p$ 阶是对称阵，则一定可以找到正交阵 U ，使

$$U^{-1}AU = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_p \end{bmatrix}$$

式中， $\lambda_1, \lambda_2, \dots, \lambda_p$ 是 A 的特征根。

定理二：若上述矩阵 A 的特征根所对应的单位特征向量为 u_1, u_2, \dots, u_p ，令

$$U = (u_1, u_2, \dots, u_p) = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1p} \\ u_{21} & u_{22} & \cdots & u_{2p} \\ \vdots & \vdots & & \vdots \\ u_{p1} & u_{p2} & \cdots & u_{pp} \end{bmatrix}$$

则实对称 A 属于不同特征根所对应的特征向量是正交的，即

$$u_i u_j = 0 \Rightarrow UU' = U'U = I$$

主成分的推导过程如下：

设 $F = a_1 X_1 + a_2 X_2 + \dots + a_p X_p = a'X$ ，其中 $a = (a_1, a_2, \dots, a_p)'$ ， $X = (X_1, X_2, \dots, X_p)'$ ，由主成分的定义可知，求取主成分也就是寻找 X 的线性函数 $a'X$ 使相应的方差尽可能的大，即使

$$\begin{aligned} \text{Var}(a'X) &= E(a'X - E(a'X))(a'X - E(a'X))' \\ &= a'E(X - EX)(X - EX)'a \\ &= a'\Sigma a \end{aligned}$$

达到最大值，且 $a'a = 1$ 。

设协方差矩阵 Σ 的特征根为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$ ，其对应的单位特征向量为 u_1, u_2, \dots, u_p 。令

$$U = (u_1, u_2, \dots, u_p) = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1p} \\ u_{21} & u_{22} & \cdots & u_{2p} \\ \vdots & \vdots & & \vdots \\ u_{p1} & u_{p2} & \cdots & u_{pp} \end{bmatrix}$$

由前面的线性代数定理可知， $UU' = U'U = I$ ，且

$$\Sigma = U \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_p \end{bmatrix} U' = \sum_{i=1}^p \lambda_i u_i u_i'$$

因此

$$a' \Sigma a = \sum_{i=1}^p \lambda_i a' u_i u_i' a = \sum_{i=1}^p \lambda_i (a' u_i) (a' u_i)' = \sum_{i=1}^p \lambda_i (a' u_i)^2 \quad (8-24)$$

所以

$$a' \Sigma a \leq \lambda_1 \sum_{i=1}^p (a' u_i)^2 = \lambda_1 (a' u_1) (a' u_1)' = \lambda_1 a' U U' a = \lambda_1 a' a = \lambda_1 \quad (8-25)$$

而且, 当 $a = u_1$ 时, 有

$$u_1' \Sigma u_1 = u_1' \left(\sum_{i=1}^p \lambda_i u_i u_i' \right) u_1 = \sum_{i=1}^p \lambda_i u_1' u_i u_i' u_1 = \lambda_1 (u_1' u_1)^2 = \lambda_1 \quad (8-26)$$

因此, $a = u_1$ 使 $\text{Var}(a' X) = a' \Sigma a$ 达到最大值, 且

$$\text{Var}(u_1' X) = u_1' \Sigma u_1 = \lambda_1 \quad (8-27)$$

同理

$$\text{Var}(u_i' X) = \lambda_i \quad (8-28)$$

而且

$$\text{cov}(u_i' X, u_j' X) = u_i' \Sigma u_j = u_i' \left[\sum_{a=1}^p \lambda_a (u_i' u_a u_a') \right] u_j = \sum_{a=1}^p \lambda_a (u_i' u_a) (u_a' u_j) = 0, i \neq j \quad (8-29)$$

上述推导过程表明, 以 Σ 的特征向量为系数的线性组合就是 X_1, X_2, \dots, X_p 的主成分, 它们彼此互不相关, 其方差为 Σ 的特征根。

由于 Σ 的特征根 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$, 所以有 $\text{Var}(F_1) \geq \text{Var}(F_2) \geq \dots \geq \text{Var}(F_p) > 0$ 。这就是主成分的次序按照特征根取值大小顺序排列的原因。

在解决实际问题时, 一般不是取全部的 p 个主成分, 确定新变量的个数 k 是一个关键问题。 k 越小, 越能降低数据维数, 便于分析, 同时也能降低噪声。但是如果 k 过小, 会导致一些有用的信息丢失, 累计贡献率的大小反映了前 k 个主成分代替原始变量时的可靠性。贡献率越大, 可靠性越大; 反之, 则可靠性越小。因此, 根据累计贡献率的大小取前 k 个主成分。称第一主成分的贡献率为 $\lambda_1 / \sum_{i=1}^p \lambda_i$, 由于 $\text{Var}(F_1) = \lambda_1$, 所以 $\lambda_1 / \sum_{i=1}^p \lambda_i = \text{Var}(F_1) / \sum_{i=1}^p \text{Var}(F_i)$ 。因此第一主成分的贡献率就是第一主成分的方差占全部方差 $\sum_{i=1}^p \lambda_i$ 的比例。这个值越大, 表明第一主成分综合 X_1, X_2, \dots, X_p 信息的能力越强。

前两个主成分的累计贡献率定义为 $(\lambda_1 + \lambda_2) / \sum_{i=1}^p \lambda_i$ ，以此类推，前 k 个主成分的累计贡献率为 $\sum_{i=1}^k \lambda_i / \sum_{i=1}^p \lambda_i$ 。如果前 k 个主成分的贡献率达到 85%，就表明前 k 个主成分中所包含的信息和那 p 个主成分所包含的几乎一样多。取这 k 个主成分就可以取代原始的那 p 个主成分，于是对 p 个主成分的 n 次测量值所组成的原始数据集，就压缩为对 k 个主成分的 n 次测量值所组成的数据集，这样既减少了变量的个数，又便于对实际问题的分析和研究。

值得指出的是，当协方差矩阵 Σ 未知时，可用其估计值 S ，即样本协方差矩阵来代替。

设原始数据资料阵为

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}$$

则 $S = (s_{ij})$ ，其中 $s_{ij} = \frac{1}{n} \sum_{a=1}^n (x_{ai} - \bar{x}_i)(x_{aj} - \bar{x}_j)$ 。

而相关系数矩阵为

$$R = (\gamma_{ij})$$

式中， $\gamma_{ij} = s_{ij} / (\sqrt{s_{ii}}\sqrt{s_{jj}})$ 。

当原始变量 X_1, X_2, \dots, X_p 标准化后，则

$$S = R = \frac{1}{n} X'X \quad (8-30)$$

在实际应用时，指标的量纲往往不同，因此在计算之前应先消除量纲的影响，即将原始的数据标准化。这样一来 S 和 R 相同。所以一般求 R 的特征值和特征向量，就不妨取 $R = X'X$ ，因为这时的 R 与 $\frac{1}{n}X'X$ 只相差一个系数。虽然 $X'X$ 与 $\frac{1}{n}X'X$ 的特征根相差 n 倍，但是它们的特征向量不变，并不影响主成分的求取。

由主成分分析的定义和基本原理的讨论大体已经可以看出进行主成分分析的步骤，在此概括如下：

1) 将原始观察数据组成样本矩阵 X ，每一行代表一维数据，每一列为一个观察样本。

2) 将样本矩阵 X 进行标准化处理，计算其协方差矩阵 Σ ：

$$\Sigma = E(XX') = E(X - EX)(X - EX)'$$

- 3) 计算协方差矩阵 Σ 的特征值 λ_i 及相应的特征向量 $u_i, i = 1, 2, \dots, p$ 。
- 4) 计算各主成分的贡献率 a_k 和累计贡献率 $a(k)$ 。
- 5) 确定主成分的个数, 构成特征空间。

通常情况下, $a(k)$ 取 0.9 ~ 1 之间的数值。将特征值按贡献率由大到小的顺序排列, 选取前 k 个较大特征值对应的特征向量, 构成变换矩阵 $U = (u_1, u_2, \dots, u_k)$, 即为特征空间。

将样本矩阵投影到这个特征空间, 求出其系数向量, 这就是样本矩阵 X 的主成分, 求出融合后的特征, 即 $F_i = U^T X_i, i = 1, 2, \dots, p$ 。

8.4 基于主成分分析的图像特征级融合实现

8.4.1 目标特征的提取

对于数据库中每一组图像, 提取红外图像目标的 Hu 不变矩 (7 维) 和几何特征 (4 维), 可见光图像目标的灰度共生矩阵 (4 维), 共计 15 维特征。然后将这些特征组成特征向量 $x = (\varphi_1, \varphi_2, \dots, \varphi_{15})^T$ 。写出样本矩阵: $X = (x_1, x_2, \dots, x_N)^T$, 其中向量 x_i 为第 i 个图像的特征向量组成的列向量, N 为训练集样本个数。

8.4.2 特征融合

求取标准化后的特征向量 $X' = (x'_1, x'_2, \dots, x'_N)^T$ 的协方差矩阵 C_x , 根据主成分分析的原理, 对 C_x 进行特征分解, 得到特征值和特征向量。根据特征值的贡献率选取前 k 个较大特征值及其对应的特征向量。

设原协方差矩阵的特征向量为 u_i , 则特征空间表示为 $U = (u_1, u_2, \dots, u_k)$ 。

将每一幅样本图像标准化后的特征向量投影到特征空间, 这就是融合特征, 即 $F_i = U^T X_i, i = 1, 2, \dots, N$ 。

8.4.3 实验结果与分析

为衡量本章融合算法的效果, 使用上述方法分别提取红外图像和可见光图像的特征组成特征向量, 根据主成分分析的原理融合特征, 并在 Visual C++ 6.0 环境下进行仿真实验。对图 7-3 预处理后的图像提取特征, 标准化并进行主成分分析特征融合的数据见表 8-1。

计算前 10 个主成分的贡献率以及特征值的累计贡献率, 计算结果见表 8-2。其中 F_1 是“信息最多”的指标, 称为第一主成分; F_2 为信息第二多的指标, 称为第二主成分, 以此类推。

表 8-1 图 7-3 特征提取、标准化并进行主成分分析特征融合的数据

	可见光图像特征				红外图像特征										
	能量	熵	惯性矩	局部平 稳性	复杂度	长宽比	均值对 比度	紧凑度	Hu 矩 特征						
特征向量	0.0929	0.4722	0.0929	0.4722	0.0929	0.4722	0.0929	0.9049	3.0768	8.6392	13.3505	13.3457	26.7719	18.1633	26.953
标准化	-0.7427	-1.4728	-0.7427	-1.4728	-0.7427	-1.4728	-0.7427	0.2043	0.0567	-0.3313	-0.0604	1.6954	1.0844	1.0136	1.288
融合特征	1.6083		1.8367		-1.5077		0.6800	0.6840	0.9009			0.0823		-0.7187	

表 8-2 前 10 个主成分的贡献率以及特征值的累计贡献率

	主成分				
	F_1	F_2	F_3	F_4	F_5
贡献率	0.4978	0.2069	0.1428	0.0515	0.0498
累计贡献率	0.4978	0.7047	0.8475	0.8990	0.9488
	主成分				
	F_6	F_7	F_8	F_9	F_{10}
贡献率	0.0273	0.0123	0.0071	0.0034	0.0007
累计贡献率	0.9761	0.9884	0.9955	0.9989	0.9996

通过对主成分分析特征数据融合得到的各主成分的贡献率（见表 8-2）及累计贡献率的分析（见图 8-3 和图 8-4）表明，经过主成分分析特征融合后的特征向量大于或等于 8 维时，特征向量的累计贡献率不再发生变化，充分说明主成分分析融合后 8 维向量可以 99.55% 地表示融合前 15 维特征向量。当贡献率设为 99% 时，特征的主成分个数为 8 个，实现了特征量的有效降维。

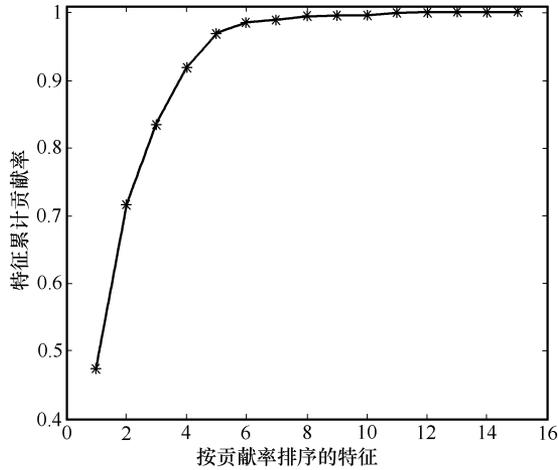


图 8-3 各主成分累计贡献率

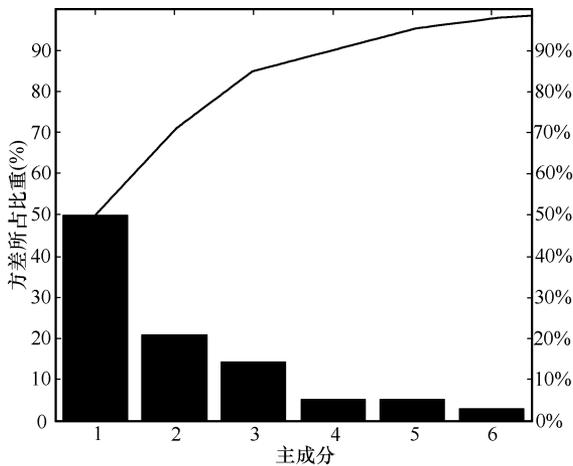


图 8-4 各主成分方差的 Pareto 图

最后，在 Visual C ++ 6.0 环境中分别使用传统的串联法和主成分分析方法进行融合识别实验。其中，主成分分析方法选择贡献率大于 99% 的特征值为主特征，两种方法使用支持向量机识别目标的识别率见表 8-3。比较分析表 8-3 中

的实验数据可以看出，传统串联法虽然特征维数多，但识别率相对较低，这是因为特征之间存在冗余，影响了识别的准确率；而基于主成分分析的红外与可见光图像特征融合方法，在未降低目标识别率的前提下，减少了系统运行时间，改善了整个识别系统的效率。

表 8-3 算法识别率比较

	传统串联法 (15 维特征)	主成分分析融合法 (8 维特征)
	识别率 (%)	识别率 (%)
人员目标	96.66	95.55
汽车目标	94.44	95.55
坦克目标	92.22	92.22
飞机目标	93.33	94.44

8.5 本章小结

本章主要研究了基于主成分分析的特征融合方法。首先根据红外与可见光图像的特点和算法的需求，介绍了图像的灰度共生矩阵、Hu 不变矩和几何特征。接着从主成分分析的概念入手，介绍了算法的数学模型和几何解释，从协方差矩阵出发给出了主成分推导过程，总结了主成分分析执行的步骤。最后，结合主成分分析思想，根据红外与可见光图像特征级融合的实际需求，将主成分分析引入到特征融合中，通过实验对主成分分析方法做出了有的放矢的比较和评价。

基于改进免疫遗传的特征融合方法

在模式识别中，由特征提取形成的原始特征空间往往维数很高，而且特征在子空间的分布是非线性的，采用基于线性的特征融合方法可能导致不理想的识别结果。这样自然想到对特征进行编码，将特征融合转化为组合优化问题，可以通过解决优化问题的方法获得最优融合方案，生成融合特征。本章正是从这个角度出发，将免疫遗传算法应用在红外与可见光图像特征级融合。

9.1 遗传算法基础理论

遗传算法起源于 20 世纪 60 年代末期到 70 年代初期，是由美国密歇根大学的 Holland 教授及其学生提出的。在 1975 年出版的著名专著《Adaptation in Natural and Artificial Systems》中，Holland 详细地阐述了遗传算法的基本理论和方法，提出了对遗传算法的发展极为重要的模板理论。同年，De Jong 在计算机上基于遗传算法的思想进行了大量的纯数值优化计算实验，首次将遗传算法运用到函数优化。80 年代，Goldberg 在一系列研究工作的基础上系统地总结了遗传算法的研究成果，完整地论述了遗传算法的基本原理及应用，形成了遗传算法的基本框架。近年来，由于遗传算法在求解复杂优化问题的巨大潜力及其在人工智能、自动控制、生物工程等各领域的成功应用，使遗传算法受到了广泛的关注。

9.1.1 遗传算法概述

遗传算法借鉴了达尔文的进化论和孟德尔的遗传学说，模仿自然界生物进化机制发展起来的一种随机全局搜索和优化方法，因而在这个算法中常涉及各种生物进化和遗传学的一些基本概念，如：

染色体：个体的表现性质。在遗传学中，染色体是生物细胞中含有的一种微小丝状化合物。它是遗传物质的主要载体，由多个遗传因子——基因组成。

基因：控制生物性状的遗传物质的功能单元和结构单元，又称为遗传因子。

基因型：它是性状染色体的内部表现，或者说，由遗传因子组合的模型。

表现型：由染色体决定性状的外部表现，或者说，根据遗传因子形成的个体，称为表现型。

个体：指染色体带有特征的实体。

群体：由一定数量的个体组成的集合。

适应度：反映个体性能的数量值，表示某一个体对生存环境的适应程度。

编码：从表现型到基因型的映射。

解码：从基因型到表现型的映射。

对上述术语的理解，有助于更好地理解遗传算法的基本思想。遗传算法是从优化问题可能潜在解集的一个种群开始，而一个种群是由经过基因编码的一定数目的个体组成。产生初始种群后，按照适者生存和优胜劣汰的原理，经过多代繁衍，产生出一个越来越好的种群。在每一代，由问题的目标函数构造一个适应度函数，根据个体适应度的优劣选择一部分优良个体，并对其进行交叉和变异操作，产生出代表新解集合的种群。这个过程将使种群像自然进化一样，子代种群比父代更加适应于环境，整个进化过程中的最优个体经过解码就可以作为问题的近似最优解。

9.1.2 遗传算法流程

遗传算法从任一初始种群出发，通过随机选择、交叉和变异等操作，产生一群更适应环境的个体，使种群进化到搜索空间中更好的区域。经过一代代的不断繁衍进化，最后收敛到最适应环境的个体，即问题的最优解。遗传算法的一般流程如图 9-1 所示，由图可以看出遗传算法是一个典型的迭代寻优的过程，它的实现包括编码方法；初始化种群的产生；适应度函数以及由选择、交叉和变异组成的遗传算子；终止条件。

1. 编码方法

由于遗传算法不能直接处理解空间的数据，因此必须通过编码将解空间的解数据转化为遗传算法所能处理的搜索空间的基因型串结构数据。常用的编码方法包括二进制编码、顺序编码、实数编码、整数编码等。编码方法对算法搜索能力、种群多样性等性能具有很大影响，需要根据所解决问题的实际需要，选择合适的编码方法。

2. 初始群体的产生

遗传算法是一种基于种群寻优的方法，算法运行时是以一个种群在搜索空间进行搜索。我们必须为遗传算法操作准备一个由若干初始解组成的初始群体。初始种群是随机产生的，具体的产生方式依赖于编码方法。种群中个体的数量

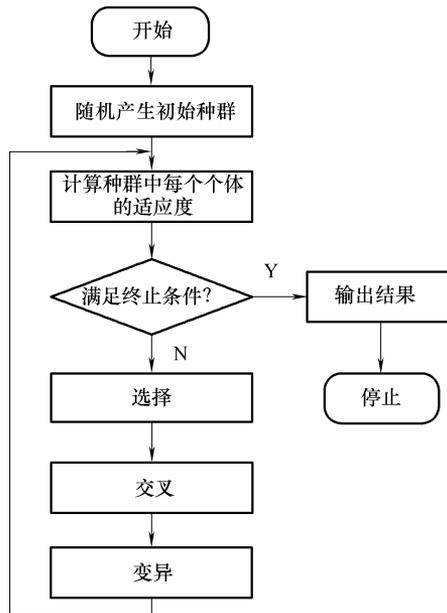


图 9-1 遗传算法流程图

称为种群规模，记为 N 。种群规模影响着遗传优化的最终结果和遗传算法的执行效率。一般来说，遗传种群规模越大越好，但规模越大，其适应度评估次数增加越多，将导致运算时间的增大，通常设为 100 ~ 1000。

3. 适应度函数

在遗传算法中使用适应度函数来表征种群中每个个体对其生存环境的适应能力，每个个体具有一个适应度，适应度是群体中个体生存机会的唯一确定性指标。因此适应度函数的选取直接影响遗传算法的收敛速度以及能否找到问题的最优解。适应度函数基本是依据优化的目标函数来确定，目标函数一般表示为 $f(x)$ ，适应度函数一般为 $F(x)$ 。

对于目标函数最小值的优化问题，将其转化为求目标函数最大值的优化问题，即

$$F(x) = -\min f(x) \tag{9-1}$$

对于目标函数最大值的优化问题，可以直接设定目标函数为其适应度函数，即

$$F(x) = \max f(x) \tag{9-2}$$

4. 遗传算子

遗传算子模拟了每一代中创造后代的繁殖过程，是遗传算法的精髓，包括选择、交叉和变异。

选择是以一定的概率从当前种群中选择一部分个体的操作。在遗传算法中自然选择规律的体现就是以适应度大小决定的概率分布来进行选择。个体的适应度越大，该个体被遗传到下一代的概率越大；反之，个体的适应度越小，该个体被遗传到下一代的概率也就越小。选择操作的目的是保留优秀的基因，改善算法的全局收敛性以提高收敛速度。常用的选择策略是正比选择策略，即每个个体被选中进行遗传运算的概率为该个体的适应度与群体中所有个体适应度总和的比。对于个体 i ，设其适应度为 F_i ，种群规模为 N ，则该个体被选中的概率为

$$P_i = \frac{F_i}{\sum_{i=1}^N F_i} \quad (9-3)$$

交叉是指同时对两个相互配对的染色体操作，组合两者的特性形成两个新的个体。通过交叉操作能从种群中寻找父代双亲已有但未被合理利用的优良基因信息，扩大了遗传算法在解空间中的搜索范围，实现其全局搜索的目的。交叉率定义为各代中交叉产生的后代数与种群规模的比，记为 P_c ，一般取为 0.4 ~ 0.99。

变异是在染色体上自发产生随机的变化，一种简单的变异方式是替换一个或者多个基因，从而形成一个新的个体。变异的目的是：一是使算法具有局部搜索的能力，当遗传算法通过交叉算子的作用已接近最优领域时，利用变异算子的局部搜索性能可以加速向最优解收敛；二是维持群体多样性，防止出现未成熟的收敛现象。变异率定义为种群中变异基因数在总基因数中的百分比，记为 P_m ，一般取为 0.0001 ~ 0.1。

5. 终止条件

遗传算法的终止条件一般采用设定最大进化代数的方法，最大进化代数常表示为 T ，一般取为 100 ~ 500。

9.2 一般的免疫算法基础理论

免疫是指生物体接触到抗原性异物（如各种细胞、微生物、病毒）后，能产生一种特异的生理反应，其作用是排除异物以保护机体。它是生物机体的一种生理反应，每当抗原性异物进入生物机体后，生物机体能识别“自己”和“非己”，并且发生特异性的免疫应答，排除抗原性的非己物质。免疫算法是基于免疫系统机制和免疫学理论开发的工程应用算法。根据构成系统中各元素的关系，一般分为基于群体的免疫算法和基于网络的免疫算法。第一类包括所有不考虑免疫网络的免疫算法，如阴性选择、克隆选择算法等。基于网络的算法是所有受免疫系统网络理论启发的算法，一般免疫算法本质上是基于网络的算法。

9.2.1 免疫算法概述

免疫算法是对生物免疫系统机理抽象而得的，算法中的许多概念和算子与免疫系统中的概念和免疫机理存在着对应关系。为了便于更好地说明免疫算法，下面阐述几个常用的免疫系统概念：

抗原：是指能够刺激和诱导机体的免疫系统使其产生免疫应答，并且能与相应的免疫应答产物在体内或体外发生特异性反应的物质。

抗体：是指免疫系统受抗原刺激后，免疫细胞转化为浆细胞并且产生能与抗原发生特异性结合的免疫球蛋白，该免疫球蛋白即为抗体。

抗原识别：通过表达在抗原表面的表位和抗体分子表面的对位化学基进行相互匹配选择完成识别，这种匹配过程也是一个不断对抗原学习的过程，最终能选择产生最适当的抗体与抗原结合。

疫苗：根据进化环境或者待求解问题的先验知识，所得到的相对最佳个体基因的估计。

抗体抗原的亲和力：抗体与抗原之间的结合能力。亲和度越高，免疫细胞与抗原的结合力越强，免疫细胞质量越好。

记忆细胞：指免疫系统将能与抗原发生反应的抗体作为记忆细胞保存记忆下来，当同类抗原再次侵入时，相应的记忆细胞被激活而产生大量的抗体，缩短免疫反应时间。

克隆：一般是指生物的增长过程。

免疫算法是将待求解的问题看作抗原，问题的解对应为抗体。在使用免疫算法解决具体问题时，首先要把需要解决的问题抽象成抗原形式，再产生初始抗体。接着计算抗体与抗原、抗体与抗体之间的亲和性，也就是对问题解的评估，根据评估结果，对记忆细胞进行更新。再通过免疫算子产生新的抗体，对新的群体进行评估，若终止条件满足，则其为该问题的最佳解，否则重新计算亲和性，进行下一轮的克隆选择，直至满足终止条件为止。免疫系统和免疫算法之间的比较见表 9-1。

表 9-1 免疫系统和免疫算法的比较

免疫系统	免疫算法
抗原	要解决的问题
抗体	最佳解向量
抗原识别	问题识别
从记忆细胞产生抗体	联想过去的成功的解
淋巴细胞分化	优良解（记忆）的保持
细胞抑制	剩余候选解的消除
抗体增加（细胞克隆）	利用遗传算子产生新抗体

9.2.2 免疫算法流程

一般免疫算法的主要步骤如下：

抗原识别：抗原识别就是用来判断侵入的抗原是否已经产生了相应的抗体。

初始抗体产生：如果侵入的抗原已经具有相应的抗体，则模拟免疫系统的记忆机制，从记忆细胞中取出相应的抗体作为初始群体；如果侵入的抗原没有相应的抗体，就随机产生抗体群。这样做的好处是，可以从记忆细胞中迅速找到适应度高的抗体，提高抗体多样性，无须进化，从而加快搜索速度。

计算适应度和抗体浓度：在当前产生的群体中，计算所有抗体的适应度。在免疫算法中，适应度的计算是通过适应度函数来完成。抗体的浓度是指与某一个体相同或者相近的个体在种群中所占的比例。

产生记忆细胞：若抗原是新的，对于当前群体取出适应度很高的抗体，计算它与记忆细胞中抗体的相似度，从而用它来代替记忆细胞中适应度最高的抗体，这样做可以提高记忆细胞中抗体的多样性。

抗体的促进和抑制：结合前面的抗体适应度和群体中抗体的浓度，来确定抗体的死亡和分裂的概率。抗体受到促进是指该个体的选择概率增加，抗体受到抑制是指该个体的选择概率减少。高适应度和低浓度的个体选择概率较高。

群体更新：把新产生的抗体群作为当前群体，按照一定比例淘汰掉适应度差的群体，再随机生成同样数量的抗体群加入到下一代中。

终止条件判断：如果满足终止条件（找到最优解或达到最大迭代次数），则算法结束，否则继续寻优。

免疫算法的基本流程如图 9-2 所示。

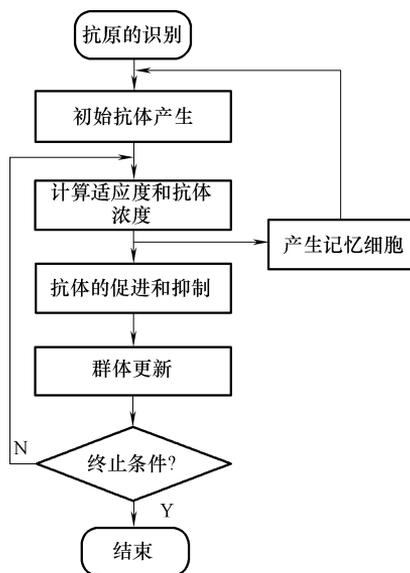


图 9-2 免疫算法流程图

9.2.3 一般的免疫遗传算法

为了克服遗传算法在解决全局最优化问题时出现早熟收敛问题，提高遗传算法的全局搜索能力，将免疫算子与传统遗传算法相结合，构造出一种改进的遗传算法——免疫遗传算法。形式上比较免疫算法和免疫遗传算法，就是在免疫算法中添加了遗传算子，与遗传算法相比，增加了抗原识别、记忆功能和调节功

能，并没有附加复杂的操作，也没有降低遗传算法的鲁棒性，免疫遗传算法兼顾了搜索速度、全局搜索能力和局部搜索能力。

在解决实际问题时，将待求解的优化设计问题作为抗原，将问题的解作为抗体，通过抗原和抗体的亲和度描述可行解与最优解的逼近程度。免疫遗传算法是从随机生成的初始解抗体种群出发，对父代（当前代）抗体群进行交叉、变异等遗传操作后，计算抗体的浓度，根据抗体的期望繁殖率进行选择操作，期望繁殖率高的个体有较高的概率被选中并复制到下一代，如此产生的子代通常优于父代。上述过程循环执行直到满足停止条件，最终使优化过程以大概率趋于全局最优解。免疫遗传算法的一般算法流程图如图 9-3 所示。

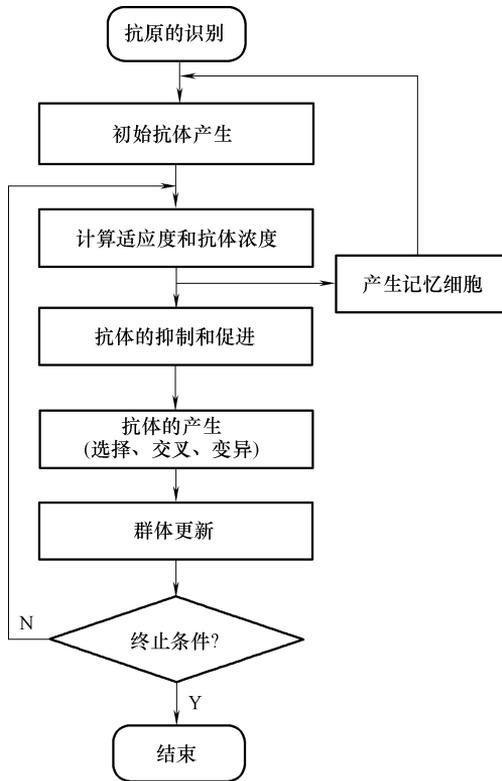


图 9-3 免疫遗传算法的一般流程图

免疫遗传算法可以克服遗传算法收敛方向无法控制的缺陷，把目标函数和约束条件作为抗原，这样做能保证所产生的抗体直接与问题相关联，收敛方向能得到控制。生成的抗体可以有效地排除抗原，也就相当于求得问题的最优解。对抗原亲和力高的抗体进行记忆，能够促进快速求解，即当遇到同类的抗原时可以快速产生相应的抗体。在许多方面表现出超越遗传算法和免疫算法的优点。

9.3

基于改进免疫遗传的图像特征级融合实现

9.3.1 改进算法的关键技术

针对传统免疫遗传算法存在的早熟收敛，固定交叉率、变异率带来的搜索过程缓慢，以至于种群进化停滞不前等，研究提出一种加入“精英选择”策略和自适应策略的改进免疫遗传算法。

1. “精英选择”策略

遗传算法的最大问题就是容易陷入局部最优，为了避免当前种群的最优个体在下一代发生丢失，导致遗传算法陷入局部最优而不能收敛到全局最优解，De Jong 在其博士论文中提出了“精英选择”策略，也称为“精英保留”策略。精英选择策略是把群体在进化过程中迄今出现的最好个体（称为精英个体）不进行配对交叉而直接复制到下一代中。

要使得遗传算法能够快速地收敛到全局最优点，应该尽可能保证适应度高的优秀个体的存在，充分发挥遗传算法“优胜劣汰”的特点。这种精英选择策略中，在保留最优个体的基础上，添加精英个体取代新一代群体中适应度最小的个体，该精英个体具有与最优个体的相异因子较大，而适应度不过小的特点。这样做不但保持了每一代最佳适应度单调递增，而且通过新添加的个体保持了种群个体染色体的多样性，提高算法的整体搜索能力，使算法具有全局收敛性。

2. 自适应策略

免疫遗传算法控制参数中的交叉概率和变异概率的选择是影响遗传算法行为和性能的关键所在，直接影响算法的收敛性。交叉率 P_c 越大，新个体产生的速度就越快，然而取值过大会破坏群体的优良特性，不利于进化；但是如果 P_c 取值过小，会使个体的搜索速度慢，以至于停滞不前。对于变异率 P_m ， P_m 越大，种群的多样性就越好，发生早熟的可能性就越小，然而较大的 P_m 将使个体改变方向，扩大搜索范围，使免疫遗传算法退化为随机搜索，进化速度变慢；如果 P_m 过小，使产生新个体和抑制早熟的能力较差，不能达到变异操作的效果。要为某个特定的优化问题设置好交叉概率和变异概率，算法需要经过反复地试验且难以丰富种群中优良解的多样性。

自适应调节能根据个体的适应度与当前群体进化情况，自动改变 P_c 和 P_m 。当种群中个体适应度趋于一致或趋于局部最优时，它使 P_c 、 P_m 增加，而当适应度比较分散时，则令 P_c 、 P_m 减少。同时对于适应度高于群体平均适应度的个体，给予较低的 P_c 、 P_m ，使它得以保护进入下一代；而低于平均适应度的个体，就给予较高的 P_c 、 P_m ，使之被淘汰。自适应变化的交叉率和变异率，能使算法具有更高的鲁棒性、全局最优性和效率。

本章 P_C 按下式进行自适应调整：

$$P_C = \begin{cases} P_{C1} - \frac{(P_{C1} - P_{C2})(F - F_{avg})}{F_{max} - F_{avg}} & F \geq F_{avg} \\ P_{C1} & F < F_{avg} \end{cases} \quad (9-4)$$

式中， F_{max} 是种群中最大的适应度函数； F_{avg} 是每代种群的平均适应度函数； F 是要交叉的两个个体中较大的适应度函数值； P_{C1} 、 P_{C2} 是初始化时第一、第二代的交叉概率。

P_M 按下式进行自适应调整：

$$P_M = \begin{cases} P_{M1} - \frac{(P_{M1} - P_{M2})(F_{max} - F')}{F_{max} - F_{avg}} & F' \geq F_{avg} \\ P_{M1} & F' < F_{avg} \end{cases} \quad (9-5)$$

式中， F' 是要变异个体的适应度函数值； P_{M1} 、 P_{M2} 是初始化时第一、第二代的变异概率。

9.3.2 基于改进免疫遗传的特征融合原理

基于改进免疫遗传的特征融合原理图如图 9-4 所示。分别提取红外和可见光图像中目标的特征，将得到的特征组成特征向量。利用改进免疫遗传算法对特征的组合进行优化，经过评估找到识别能力最好的特征融合组合。

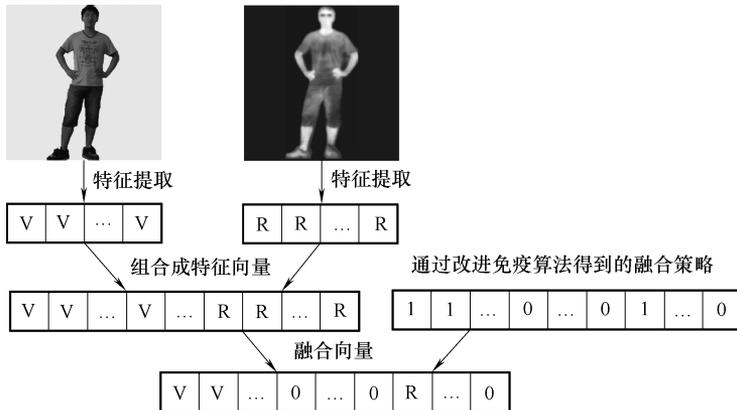


图 9-4 基于改进免疫遗传的特征融合原理图

9.3.3 特征级融合实现

以上述介绍的抗体相似度、抗体浓度等定义以及精英选择策略、自适应策略为基础，下面给出基于改进遗传免疫算法的图像特征级融合的具体步骤：

1. 确定编码方式和参数

基于免疫遗传的特征融合目的是从原始特征向量中选出 d 个特征的最优组

合，组成融合特征。这里采用二进制编码，选中的特征取为 1，否则为 0，则抗体编码长度为原始特征向量个数。根据实际需求确定种群规模、选择率、变异率和终止进化代数。

2. 生成初始抗体群

对于初次应答，随机产生 N 个初始抗体，而对于再次应答，则借助免疫机制的记忆功能，部分初始抗体由记忆细胞获取，剩余的抗体随机地产生。

3. 计算抗体的亲和度、浓度和激励度

抗体的亲和度表征免疫细胞与抗原的结合强度，定义为

$$\text{aff}(x_i) = 100p_{\text{sum}} - 10d/L \quad (9-6)$$

式中， p_{sum} 是抗体 x_i 十折交叉验证训练模型的识别率， d 为选中的特征个数， L 为原始特征向量的特征个数。

十折交叉验证是指将每类目标的训练样本分为十份，轮流将其中九份作为训练数据，一份作为测试数据，进行试验。一般需循环十次，直到所有十份数据全部被选择一遍为止，十次验证的正确率平均值即为 p_{sum} 。

抗体的浓度表征了抗体种群多样性的好坏，过高的抗体浓度意味着种群中存在大量相似的个体，寻优搜索会集中于可行解区间的一个区域，不利于全局优化。因此应抑制种群中浓度过高的个体，以保证个体的多样性。抗体 x_i 的浓度定义为

$$\text{den}(x_i) = \frac{1}{N} \sum_{j=1}^N \text{aff}(x_i, x_j) \quad (9-7)$$

式中， $\text{aff}(x_i, x_j) = \sum_{k=1}^L \sigma_k$ ， $\sigma_k = \begin{cases} 1 & x_{ik} = x_{jk} \\ 0 & x_{ik} \neq x_{jk} \end{cases}$ ， x_i 为种群中的第 i 个抗体， $\text{aff}(x_i, x_j)$

为抗体 x_i 与抗体 x_j 的亲度， x_{ik} 为抗体 x_i 的第 k 位。

抗体的激励度是对抗体质量的最终评价结果，与抗体浓度和抗体亲密度密不可分，通常浓度低而亲密度大的抗体具有较高的激励度。抗体 x_i 的激励度可以利用抗体亲和度和抗体浓度进行简单的数学运算得到，如：

$$\text{sim}(x_i) = \text{aff}(x_i) - 10\text{den}(x_i) \quad (9-8)$$

计算所有抗体的激励度，把抗体按激励度从小到大的顺序排列，将适应度最大的抗体保留为精英抗体，然后将其复制和保存。

4. 抗体记忆库初始

设定记忆库的大小为十个，从初始化种群中选择抗体激励度最好的两个个体进行存储。

5. 选择、交叉和变异操作

本章采用正比选择策略对抗体执行选择操作，以 P_c 为概率交叉抗体，对于经过交叉操作后的抗体，以 P_M 为概率随机对种群进行变异，形成新一代种群。

根据式 (9-4) 和式 (9-5), 自适应改变 P_C 、 P_M 。

6. 群体更新

把抗体按激励度从小到大的顺序排列, 用记忆抗体代替遗传操作后产生的新种群中前几个个体。如果这一代抗体群中没有与精英抗体激励度相同的抗体, 则将保存的精英抗体复制一个到该抗体群中, 并将该抗体群中激励度最小的抗体删除; 如果这一代抗体群中激励度最大的抗体其激励度的值大于精英抗体的激励度的值, 则将这个激励度最大的抗体复制一个, 并以它作为新的精英抗体替代保存的精英抗体。

7. 更新抗体记忆库

如果抗体记忆库未滿, 选择最好的两个新个体更新补充记忆库; 如果抗体记忆库已滿, 则选择最好的两个新个体替换记忆库中最差的两个个体。

8. 终止检验

若达到终止进化代数, 则结束整个算法的运行, 输出融合特征; 否则, 返回步骤 3。

改进免疫遗传算法的图像特征级融合流程图如图 9-5 所示。

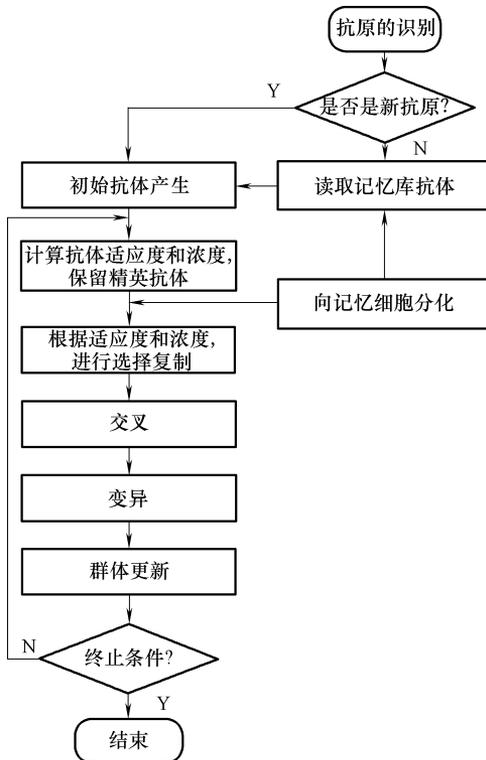


图 9-5 改进的免疫遗传算法流程图

9.3.4 实验结果与分析

为验证本章算法改进后的优越性，采用本章改进算法对 8.1 节提取方法生成的特征向量进行融合实验。设定算法的初始种群规模为 50，初始交叉概率为 0.9，第二代交叉概率为 0.6，初始变异率为 0.1，第二代变异概率为 0.001，停止准则设为迭代次数达到 200 代。运行 10 次之后，激励度最高为 95.33，此时融合后的特征选为复杂度、长宽比、局部平稳性、Hu 不变矩中的 ϕ_3 、 ϕ_4 、 ϕ_6 、 ϕ_7 。将本章算法与基于遗传算法和基于免疫遗传算法的特征融合方法进行比较，为了便于对比分析，参数选择同上。绘制适应度与迭代次数的变化曲线如图 9-6 所示。

从图 9-6 可以看出，遗传算法在迭代过程中多次出现适应度较长时间停留在一个数值上的情况，与其他两种算法相比，遗传算法适应度的变化速度较慢，这说明种群向一些相同或相似的串收敛。另外，遗传算法收敛到局部最优解，而不是全局最优解，即早熟现象。也就是说在遗传算法进化过程中，出现了一个适应度大大超过当前种群平均适应度的个体，在选择、交叉和变异算子的作用下，使得这个在种群中占据绝对的优势，搜索解的范围迅速变窄，从而使算法较早的收敛于局部最优解。

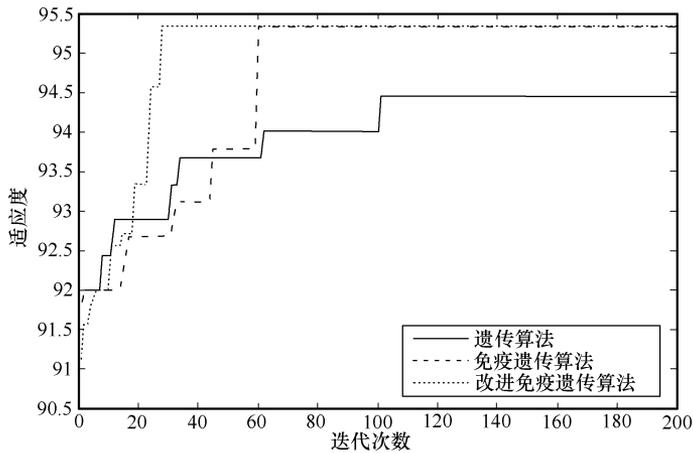


图 9-6 收敛曲线

免疫遗传算法在第 60 代左右搜索到了全局最优解，而遗传算法是在第 100 代左右才搜索到局部最优解，说明与遗传算法相比，免疫遗传算法不仅有较快的收敛速度，而且能有效克服遗传算法中出现的早熟收敛现象，提高了全局搜索能力。这是由于记忆细胞的引入使得每次交叉和变异的父代抗体都可以从性能良好、适应生存的抗体中挑选，免疫遗传算法能以很快的速度收敛于全局最优。从曲线的收敛程度可以看出，使用免疫遗传算法得到的抗体的亲和度是不

断上升的，这说明抗体不断朝着优良方向进行进化。

本章研究提出的改进免疫遗传算法，在第 30 代左右搜索到了全局最优解。“精英选择”策略把最好的个体作为精英直接带入下一代个体中，而经过任何改变，保持了种群中个体染色体的多样性，提高了全局搜索能力。“精英选择”策略也能避免算法对冗余空间的过多搜索，加快了搜索速度。在交叉率和变异率中加入自适应策略，使得交叉操作时，组合两个个体中的优良基因产生新的后代，可以在种群进化期间加快搜索速度；变异操作时保持种群中基因的多样性，提高了算法的计算速度和效率，使算法在保持种群多样性的同时保证了算法的收敛性，确保快速收敛于全局最优解。

将各算法融合得到的主要实验数据汇总在一起，见表 9-2。prob 为使用融合特征和支持向量机方法识别目标的平均成功率，IMEAN 为寻找最优特征组合所需要的平均迭代次数。

表 9-2 各算法融合效率评价

	遗传算法	免疫遗传算法	改进算法
prob	94.44	97.32	97.32
IMEAN	104	75	30

表 9-2 的数据可以更直观地证明上述结论。算法在第 30 代左右就能够收敛到最优解，这说明本章研究提出的改进免疫遗传算法有较快的收敛速度和不易出现早熟现象的优点。

9.4 本章小结

本章在介绍了一般的遗传算法、免疫算子与遗传算法相结合的免疫遗传算法的基础上，针对免疫遗传算法的不足，将“精英选择”策略和自适应策略加入到免疫遗传算法中，提出并研究了一种改进免疫遗传算法。随后研究了基于这种改进免疫遗传算法的图像特征融合，先对特征进行编码，随后初始化种群，再由改进的免疫遗传算法融合特征，最后通过实验对比分析了改进算法与遗传算法和免疫遗传算法的性能。实验结果证明改进的免疫遗传算法有较快的收敛速度和较好的搜索能力，使用基于改进的免疫遗传算法的图像融合特征具有较好的识别率。

基于独立分量的特征融合

独立分量分析 (ICA) 算法是一种新兴的信号处理技术, 用于从多维统计数据中找出隐含因子或独立分量, 从线性变换的角度解释, ICA 就是要在混合的观测信号中估计出源信号的基本结构。ICA 的基本理论是在 20 世纪 80 年代初期由法国学者 J. Herault 和 C. Jutter 首先提出的, 到 90 年代中期, ICA 的研究已经得到了国际信号处理界的广泛关注。目前 ICA 的研究工作都集中在算法理论研究和实际应用方面, 各国学者相继提出了一系列的估计算法, 包括 FastICA 算法、Infomax 算法、极大似然估计算法和高阶累积量算法等。随着 ICA 算法理论研究的不断改进和日益成熟, 该算法也被广泛应用到了通信数据处理、金融数据分析、生物医学信号处理、语音信号处理、图像处理等各个领域, 并取得了很好的成绩。ICA 算法起初是在信号处理领域发展起来并得到了广泛关注, 但随着 ICA 算法本身的研究以及在信号处理方面的成功应用, 很多学者开始尝试把 ICA 算法引入到图像处理领域, 并且也取得令人惊奇和满意的效果。

10.1 ICA 的定义

ICA 的基本思想主要是在假设随机变量的各分量之间是统计独立或近似独立的情况下用一组基函数来表示这一随机变量, 也就是能从训练样本中找到一组相互独立的分量, 并以此来描述原来的样本数据。

ICA 的目的是将观测数据通过某种线性变换分解成统计独立的分量。假设输入的随机向量 x 确实是通过统计独立的线性组合来合成的, 并且从严格意义上讲, 成分是固定的, 见式 (10-1)。

$$x = As \quad (10-1)$$

式中, $x = [x_1, x_2, \dots, x_n]^T$ 是 n 维观测信号向量; A 是 $n \times m$ 大小的混合矩阵; $s = [s_1, s_2, \dots, s_m]^T$ 是 m 维零均值的源信号向量, 其分量是相互独立的。

ICA 的任务是给定一个输入样本 x , 通过确定一个可逆矩阵 W 得到变换后的向量 y , 即式 (10-2) 成立。

$$y = Wx \quad (10-2)$$

式中, $y_i (i=1, 2, \dots, m)$ 是源信号 s 的估计值。

假设所有独立分量都有相同的分布, 且 $m = n$, 这样就可以认为 W^{-1} 是 A 的估计, 在理想情况下, $W = A^{-1}$ 。由式 (10-2) 可知, 为了估计其中的一个独立分量 y_i , 可考虑 $y_i = w_i^T x$ 的线性组合, 这里 w 是一个待定的向量, 如果 w_i^T 是矩阵 $W(A$ 的逆) 的一个行向量, 则这个线性组合实际上就是一个独立分量。但是由于矩阵 A 是未知的, 所以不能准确地确定出 w , 但是我们可以找到一个很接近的估计。若将变量进行如下变换, 即定义 $z = A^T w$, 则有

$$y = w^T x = w^T A s = z^T s \quad (10-3)$$

式中, y 是 s 的一个线性组合, 其权重由 z 给出。

由式 (10-3) 可知最大化 $w^T x$ 的非高斯性, 即可得到一个独立分量。

10.2 随机变量的独立性概念

统计独立性是构成 ICA 基础的一个关键概念。考虑两个不同随机变量 x 和 y 的情形, 如果知道随机变量 y 的值并不能给出随机变量 x 取值的任何信息, 那么我们说 x 独立于 y 。数学上, 统计独立性是通过联合概率密度函数来定义的。若随机变量 x 和 y 为独立的, 需要满足式 (10-4) 的条件:

$$p_{x,y}(x,y) = p_x(x)p_y(y) \quad (10-4)$$

换句话说, x 和 y 的联合概率密度 $p_{x,y}(x,y)$ 必须能分解成它们的边缘密度 $p_x(x)$ 与 $p_y(y)$ 的乘积。可以通过累积分布函数等价地来定义: 在式 (10-4) 中将概率密度函数换成相应的累积分布函数, 那么联合累积分布函数也必须是可分解的。

满足独立性的随机变量具有如下基本性质:

$$E\{g(x)h(y)\} = E\{g(x)\}E\{h(y)\} \quad (10-5)$$

式中, $g(x)$ 和 $h(y)$ 分别是关于 x 和 y 的任意绝对可积函数。这是因为:

$$\begin{aligned} E\{g(x)h(y)\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)p_{x,y}(x,y) dy dx \\ &= \int_{-\infty}^{\infty} g(x)p_x(x) dx \int_{-\infty}^{\infty} h(y)p_y(y) dy \\ &= E\{g(x)\}E\{h(y)\} \end{aligned} \quad (10-6)$$

由式 (10-4) 可以推广到多个随机变量的情形。令 $x, y, z \dots$ 是随机向量, 则 $x, y, z \dots$ 的独立性条件见式 (10-7):

$$p_{x,y,z \dots}(x,y,z \dots) = p_x(x)p_y(y)p_z(z) \dots \quad (10-7)$$

而基本性质式 (10-5) 则推广为式 (10-8) 所示:

$$E\{g_x(x)g_y(y)g_z(z) \dots\} = E\{g_x(x)\}E\{g_y(y)\}E\{g_z(z)\} \dots \quad (10-8)$$

式中, $g_x(x)$ 、 $g_y(y)$ 和 $g_z(z)$ 分别是随机变量 x 、 y 、 z 的任意函数, 要求这些

函数使得式 (10-8) 中定义的期望存在。

式 (10-8) 给出了统计独立性标准观念的一种推广。随机向量 x 的分量本身就是标量值的随机变量, y 和 z 也是一样。显而易见, x 的各分量之间是可以相互关联的, 但需要与其他随机向量的分量之间相互独立, 对随机向量 y 和 z 也有类似的论断, 这样才能使得式 (10-8) 成立。

10.3 ICA 独立性的度量

ICA 的目标是通过寻找合适的线性变换 [见式 (10-9)], 来得到一组相互独立的随机变量, 也就意味着如果源信号 $S(t)$ 之间是相互独立的, 那么 $Y(t)$ 是 $S(t)$ 的估计。下面我们讨论几个信号之间独立性的度量标准。

$$Y(t) = BX(t) \tag{10-9}$$

10.3.1 非高斯性

非高斯性在 ICA 模型的估计中扮演着极为重要的角色, 是实现 ICA 估计的一个基础, 独立成分可以通过寻找数据具有最大非高斯性的方向而得到, 而非高斯性可以利用熵或者累积量 (如峭度) 来度量。实际在自然界中真正满足高斯分布的信号很少, 所以 ICA 方法的研究有非常重要的意义, 并且应用前景广阔。ICA 以统计独立性为基本原则, 非高斯性极大是度量 ICA 独立性的重要方法, 它的思想来源于中心极限定理。

定理 10.1 (中心极限定理): 设随机变量 x_1, x_2, \dots, x_n 相互独立, 它们的期望和方差见式 (10-10) 所示:

$$E\{x_k\} = \mu_k \quad D\{x_k\} = \sigma^2 \neq 0 \quad k=1, 2, \dots, n \tag{10-10}$$

记 $B_n^2 = \sum_{k=1}^n \sigma_k^2$, 若存在正数 δ , 使得当 $n \rightarrow \infty$ 时, 令 $x_k = \sum_{i=1}^k z_i$ 为某独立同分布随机变量 $\{z_i\}$ 的部分和序列。因为当 $k \rightarrow \infty$ 时, x_k 可能无界地增长, 进而考虑标准化的变量, 其公式见式 (10-11) 所示:

$$y_k = \frac{x_k - E\{x_k\}}{D\{x_k\}} \tag{10-11}$$

可以说明, 当 $k \rightarrow \infty$ 时, y_k 的分布收敛于具有零均值和单位方差的某个高斯分布。

中心极限定理可以推广到具有共同均值 m_z 和协方差矩阵 C_z 的随机向量 z_i 的情形。随机向量序列见式 (10-12) 所示:

$$y_k = \frac{1}{\sqrt{k}} \sum_{i=1}^k (z_i - m_z) \tag{10-12}$$

它的极限分布是一个具有零均值和协方差为 C_z 的多元高斯分布。

由中心极限定理可知，在一定条件下，相互独立的各个随机变量之和的分布趋向于高斯分布，两个独立随机变量之和形成的分布比两个原始的随机变量中的任意一个更接近于高斯分布。也就是说，最大化信号的非高斯性和独立性是一致的，这也就是所谓的“非高斯性就是独立的”。

为了在 ICA 估计中使用非高斯性，我们必须对一个随机变量（如 y ）的非高斯性定义一个度量指标。下面将介绍两个度量信号非高斯性的重要指标：峭度和负熵。

(1) 峭度

随机变量的四阶累计量又叫作峭度，是经典的非高斯性度量指标。

在零均值的情况下， y 的峭度 $\text{kurt}(y)$ 可定义为

$$\text{kurt}(y) = E\{y^4\} - 3(E\{y^2\})^2 \quad (10-13)$$

也可以用规范化的峭度，其定义为

$$\tilde{k}(y) = \frac{E\{y^4\}}{(E\{y^2\})^2} - 3 \quad (10-14)$$

对于白化的数据， $E\{y^2\} = 1$ ，因此峭度的两个定义都归结为

$$\text{kurt}(y) = \tilde{k}(y) = E\{y^4\} - 3 \quad (10-15)$$

这意味着对于白化数据，四阶矩 $E\{y^4\}$ 可以替代峭度来刻画 y 的分布，这也说明，峭度实际上就是四阶矩的一种规范形式。高斯变量的峭度为零，对于高斯分布的变量 y ，其四阶矩等于 $3(E\{y^2\})^2$ 。

在 ICA 及其相关领域，峭度的绝对值或 2 次方已被广泛地用作非高斯性的度量，这主要是因为它无论从理论研究还是计算上都非常简单。从计算角度，在样本方差保持不变的情况下，峭度可以简单地用样本数据的四阶矩来估计。由于在两个随机变量相互独立的情况下，峭度具有以下两种线性特点，因此理论分析也相应得到简化。设 y_1 和 y_2 是两个相互独立的随机变量，则式 (10-16) 和式 (10-17) 两式恒成立：

$$\text{kurt}(y_1 + y_2) = \text{kurt}(y_1) + \text{kurt}(y_2) \quad (10-16)$$

$$\text{kurt}(\alpha y_1) = \alpha^4 \text{kurt}(y_1) \quad (10-17)$$

式中， α 为常数。

(2) 负熵

熵是信息论中的基本概念，用来衡量随机变量出现的期望值。对于一个离散取值的随机变量 X ，它的熵 H 定义为

$$H(X) = - \sum_i P(X = a_i) \log P(X = a_i) \quad (10-18)$$

式中， a_i 为 X 的可能取值。

取不同的对数基底，将得到熵的不同单位。通常情况下取 2 作为基底，这时的单位是 bit。

负熵也是度量非高斯性的一个重要指标，负熵 J 定义为

$$J(x) = H(\mathbf{x}_{\text{gauss}}) - H(x) \quad (10-19)$$

式中， $\mathbf{x}_{\text{gauss}}$ 为高斯随机向量。

它与 \mathbf{x} 具有相同的协方差矩阵。它的熵可以由式 (10-20) 计算得到，即

$$H(\mathbf{x}_{\text{gauss}}) = \frac{1}{2} \log |\det \Sigma| + \frac{n}{2} [1 + \log 2\pi] \quad (10-20)$$

式中， n 为 x 的维数。

负熵总是非负的，它在可逆线性变换下是不变的，这是因为，对 $\mathbf{y} = \mathbf{M}\mathbf{x}$ ，有 $E\{\mathbf{y}\mathbf{y}^T\} = \mathbf{M}\Sigma\mathbf{M}^T$ ，负熵可以用式 (10-21) 计算：

$$\begin{aligned} J(\mathbf{M}\mathbf{x}) &= \frac{1}{2} \log |\det(\mathbf{M}\Sigma\mathbf{M}^T)| + \frac{n}{2} [1 + \log 2\pi] - (H(\mathbf{x}) + \log |\det \mathbf{M}|) \\ &= \frac{1}{2} \log |\det(\Sigma)| + 2 \times \frac{1}{2} \log |\det(\mathbf{M})| + \frac{n}{2} [1 + \log 2\pi] - H(\mathbf{x}) - \log |\det \mathbf{M}| \\ &= \frac{1}{2} \log |\det(\Sigma)| + \frac{n}{2} [1 + \log 2\pi] - H(\mathbf{x}) \\ &= H(\mathbf{x}_{\text{gauss}}) - H(\mathbf{x}) = J(\mathbf{x}) \end{aligned} \quad (10-21)$$

负熵是尺度不变的，也就是说，把一个随机变量乘以某个常数，不会改变它的负熵。

10.3.2 互信息

互信息是指一个集合中的一组随机变量和其他随机变量之间同时具有某种信息的度量，互信息最小是度量 ICA 独立性的另一个重要指标。利用熵，可以将 n 个（标量）随机变量 $x_i, i=1, \dots, n$ 之间的互信息 I 定义为

$$I(x_1, x_2, \dots, x_n) = \sum_{i=1}^n H(x_i) - H(X) \quad (10-22)$$

式中， X 为包含所有 x_i 的向量。

互信息可以像熵一样解释成代码长度。 $H(x_i)$ 这些项给出了当 x_i 这些变量各自单独编码时的码长，而 $H(X)$ 给出了当 X 作为一个随机向量，即所有分量都编码在同一个代码中的码长。这样，互信息表明了对整个向量进行编码时的码长与对单个分量分别编码时的码长的长度差值。一般而言，对整个向量编码会得到更好的代码。然而，如果 x_i 之间相互独立，它们都不给出对方的信息，这时我们只需对各个变量分别单独编码，而不会增加码长。互信息总是非负的，当且仅当变量之间统计独立时为零，所以互信息最小也是衡量 ICA 独立性的重要方法。

10.4 快速固定点 ICA 算法

快速固定点 ICA 算法又称为 FastICA 算法，是一种快速寻优的迭代算法。我

们通常用负熵来度量非高斯性，基于负熵最大的 FastICA 算法是寻找一个使负熵最大的方向，即一个单位长度向量 w ，使得对应的投影 $y = w^T z$ 具有最大的非高斯性。非高斯性在这里是利用负熵的近似 $J(w^T z)$ 来度量的。基于负熵最大的 FastICA 算法结合了固定点迭代的优良算法特性与负熵良好的统计特性，因此是一种有效并且快速的 ICA 算法。

FastICA 算法的推导如下： $w^T z$ 的近似负熵的极大值通常在 $E\{G(w^T z)\}$ 的极值点处取得，根据拉格朗日条件， $E\{G(w^T z)\}$ 在约束 $E\{(w^T z)^2\} = \|w\|^2 = 1$ 条件下的极值是在使下式梯度为零的点处取得：

$$E\{zg(w^T z)\} + \beta w = 0 \tag{10-23}$$

式中， $\beta = E\{w_0^T Xg(w_0^T X)\}$ ，是一个恒定值，其中 w_0 是优化后的 w 值。下面利用牛顿法来求解式 (10-23)，用 F 表示式 (10-23) 左侧的部分，求得其梯度为

$$\frac{\partial F}{\partial w} = E\{zz^T g'(w^T z)\} + \beta I \tag{10-24}$$

为了简化矩阵求逆，需要对式 (10-24) 中右侧的第一项进行近似。因为数据已经经过白化处理， $E\{zz^T\} = I$ ，所以 $E\{zz^T g'(w^T z)\} \approx E\{zz^T\} E\{g'(w^T z)\} = E\{g'(w^T z)\} I$ ，这时梯度变成了对角化的矩阵，可以简单地求逆。因而可以得到近似的牛顿迭代算法如下所示：

$$w^* = w - [E\{zg(w^T z)\} - \beta w] / [E\{g'(w^T z)\} - \beta] \tag{10-25}$$

$$w = w^* / \|w^*\| \tag{10-26}$$

式中， w^* 为 w 的新值； $\beta = E\{w^T zg(w^T z)\}$ ； $g(\cdot)$ 为非线性函数。 $g(\cdot)$ 常选用下列函数：

$$g_1(u) = \tanh(a_1 u) \tag{10-27}$$

$$g_2(u) = u \exp(-u^2/2) \tag{10-28}$$

$$g_3(u) = u^3 \tag{10-29}$$

经过简化可以得到 FastICA 算法的迭代公式为

$$w^* = E\{zg(w^T z)\} - E\{g'(w^T z)\} w \tag{10-30}$$

经过标准化得到 $w = w^* / \|w^*\|$ 。

FastICA 算法的基本步骤如下：

- 1) 将观测数据进行中心化使其均值为 0。
- 2) 然后进行白化，得到 z 。
- 3) 选择一个具有单位范数的初始化（可随机选取）向量 w 。
- 4) 更新 $w^* = E\{zg(w^T z)\} - E\{g'(w^T z)\} w$ ，函数 g 的定义见式 (10-27) ~ 式 (10-29)。
- 5) 标准化 w ， $w = w^* / \|w^*\|$ 。
- 6) 如果尚未收敛，返回步骤 4。

同理，可以推广到估计多个分量的情况，计算步骤如下：

- 1) 对观测数据 X 进行中心化, 使它的均值为 0。
- 2) 对数据进行白化, $X \rightarrow Z$ 。
- 3) 选择需要估计的分量的个数 m , 设迭代次数 $p \leftarrow 1$ 。
- 4) 选择一个初始权矢量 (随机的) W_p 。
- 5) 令 $W_p = E\{Zg(W_p^T Z)\} - E\{g'(W_p^T Z)\}W$, 非线性函数 g 的选取见式(10-27) ~ 式(10-29)。

$$6) W_p = W_p - \sum_{j=1}^{p-1} (W_p^T W_j) W_j。$$

$$7) \text{ 令 } W_p = W_p / \|W_p\|。$$

8) 假如 W_p 不收敛, 返回步骤 5。

9) 令 $p = p + 1$, 如果 $p \leq m$, 返回步骤 4。

FastICA 算法的收敛速度快, 和梯度算法不同, 无须选择步长参数, 易于使用。它的性能能够通过选择适当的非线性函数 g 来优化, 并且能够利用非线性函数 g 直接分离出任何非高斯分布的独立分量, 不必像其他算法一样需要进行概率密度函数的估计。它与投影追踪相类似, 独立分量可以被逐个估计出来, 因此在只需要估计出一小部分独立分量的情况下, 可以减少计算量。FastICA 算法与神经网络算法一样有很多优点, 它是并行的、分布式的且计算简单, 内存要求很少。

10.5 基于 ICA 的图像特征级融合实现

10.5.1 ICA 的预处理

1. 中心化

对观测数据进行中心化处理是 ICA 算法最基本的预处理步骤, 其处理过程是把每个观测数据 (也可看作随机变量) X 减去其均值向量 $m = E\{X\}$, 使 X 成为零均值变量。

这意味着源信号 S 的估计信号 Y 也同时变为零均值的。因为:

$$E\{S\} = A^{-1}E\{X\} \quad (10-31)$$

不过, 通过这样的中心化预处理, 混合矩阵仍可保持不变, 因此不用担心会对混合矩阵的估计有影响。当利用零均值矩阵对混合矩阵与独立成分进行估计以后, 被减掉的均值还可以通过在独立成分上加 $A^{-1}E\{X'\}$ 而简单地恢复。

2. 白化

一般情况下, 我们所获取的观测数据都具有相关性, 而白化过程能够很好地去除数据之间的相关, 所以在预处理时通常都会对数据做白化处理, 从而简化后续独立成分的提取过程。而且与没有经过白化处理的数据相比, 白化后的

数据收敛性更好。

一个零均值的随机向量 z 的各分量具有相同的单位方差且相互不相关，我们称 z 是白化的，换句话说， z 的协方差矩阵（包括相关矩阵）是单位矩阵，如下所示：

$$E\{zz^T\} = I \quad (10-32)$$

进一步来说，白化过程意味着我们将观测数据向量 x 与某个矩阵 V 线性相乘后得

$$z = Vx \quad (10-33)$$

得到一个白化的新向量 z 。白化过程有时也称为球面化。

令 $E = (e_1, \dots, e_n)$ 是以协方差矩阵 $C_x = E\{xx^T\}$ 的单位范数特征向量为列的矩阵， $D = \text{diag}(d_1, \dots, d_n)$ 是以 C_x 的特征值为对角元素的对角矩阵，则线性白化变换可以由式 (10-34) 给出：

$$V = D^{-1/2}E^T \quad (10-34)$$

C_x 可以用特征向量和特征值矩阵 E 和 D 写成 $C_x = EDE^T$ ， E 为正交矩阵，满足 $EE^T = E^TE = I$ ，则 $E\{zz^T\} = VE\{xx^T\}V^T = D^{-1/2}E^TEDE^TDE^{-1/2} = I$ 成立， z 的协方差为单位矩阵，所以 z 是白化的。

式 (10-34) 中的线性算子 V 不是唯一的白化矩阵，容易看到，任何矩阵 UV (U 为正交矩阵) 也是白化矩阵。最常用的白化方法是利用协方差矩阵的特征值分解：

$$E\{xx^T\} = EDE^T \quad (10-35)$$

式中， E 为 $E\{xx^T\}$ 的特征向量的正交矩阵； D 为相应的特征向量的对角矩阵。

这样，白化过程可以利用下面的白化矩阵来实现：

$$V = ED^{-1/2}E^T \quad (10-36)$$

矩阵 $D^{-1/2}$ 只需通过简单的逐元素开方计算得到 $D^{-1/2} = \text{diag}(d_1^{-1/2}, \dots, d_n^{-1/2})$ 。这样得到的白化矩阵记为 $E\{xx^T\}^{-1/2}$ 或 $C^{-1/2}$ 。

白化矢量 z 可以由式 (10-37) 获得：

$$z = ED^{-1/2}E^Tx \quad (10-37)$$

对信号的白化可以得到一个新的混合矩阵 \tilde{A} ，而且 \tilde{A} 是正交矩阵，对于混合矩阵 A ，使用 ICA 算法需要估计 n^2 个参数，而对于新的混合矩阵 \tilde{A} ，由于它是正交的，只需要估计 $n(n-1)/2$ 个参数，因此大大减少了 ICA 的工作量。

10.5.2 特征融合

$x = (x_1, x_2, \dots, x_n)$ 是观测数据，经过中心化处理得到矩阵 $x' = (x'_1, x'_2, \dots, x'_n)$ ，经过白化处理后得到矩阵 z ，然后根据 10.4 节叙述的 FastICA 原理及步骤求出分离矩阵 w ，通过式 (10-3) 计算出各独立分量 y 。

10.6 实验结果与分析

实验中选取人、汽车、卡车和飞机四类目标，而我们还没有找到此方面公开的标准融合识别数据库，于是通过自行采集红外摄像机和 CCD 照相机的源图像，并且对源图像进行预处理，建立了上述四类目标的红外与可见光融合识别数据库，其中每类目标有 90 组红外和可见光图像，尺寸均为 256 像素 × 256 像素。每类目标的第 1, 3, 5, …, 89 幅图片组成训练集，第 2, 4, 6, …, 90 幅图片组成测试集。数据库的部分图像如图 10-1 所示，上面一行为可见光，下面一行为红外图像。

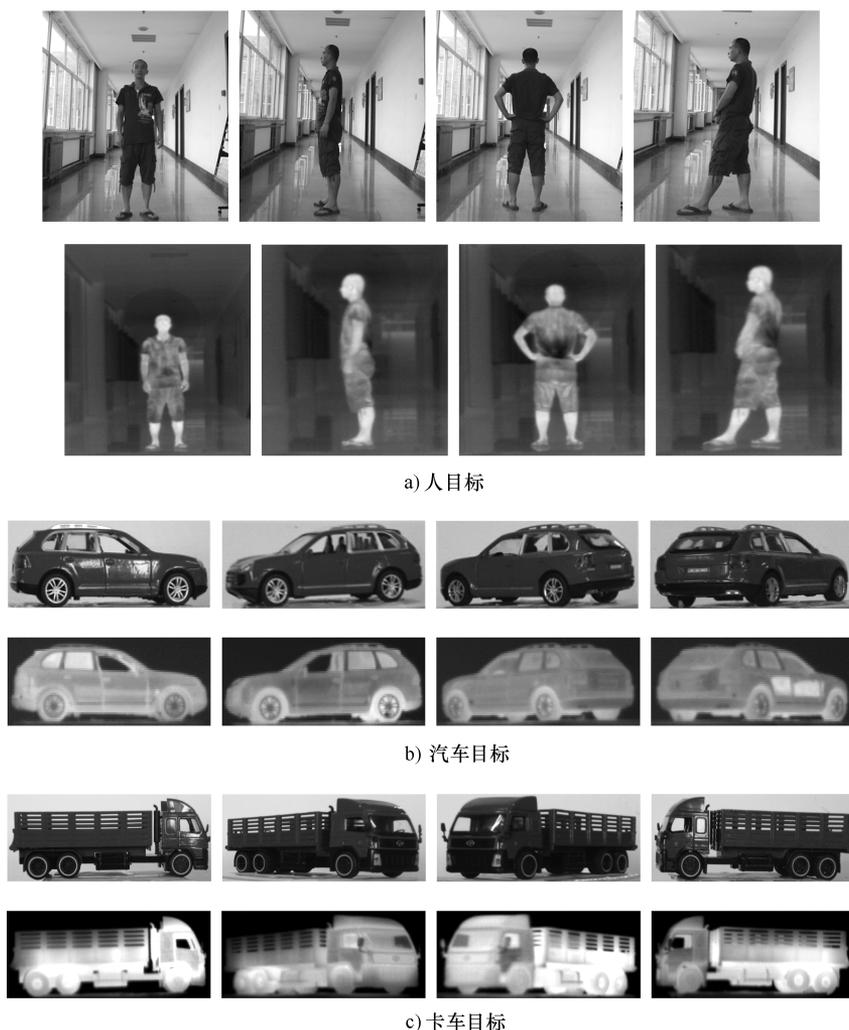
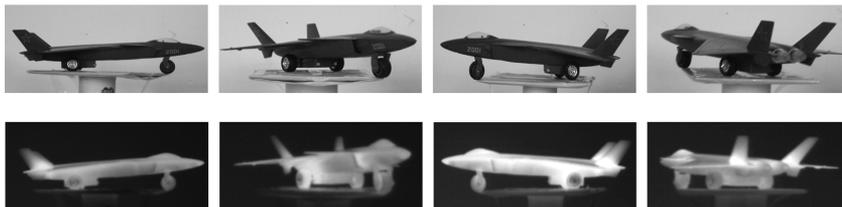


图 10-1 可见与红外融合识别数据库部分图像



d) 飞机目标

图 10-1 可见与红外融合识别数据库部分图像 (续)

首先,对可见光和红外图像进行增强、滤波、分割等预处理,然后对处理后的图像进行特征提取,提取的图像特征包括灰度共生矩阵、Hu 不变矩、仿射不变矩、小波矩和 Zernike 矩。

提取图 10-1 中部分可见光图像的灰度共生矩阵特征值见表 10-1,其特征值曲线如图 10-2 所示。

表 10-1 灰度共生矩阵特征值

	能量 (φ_1)	熵 (φ_2)	局部平稳性 (φ_3)	惯性矩 (φ_4)
人 (可见光)	0.39606	1.38461	0.77592	2.89037
汽车 (可见光)	0.25205	2.25716	0.71144	4.49152
卡车 (可见光)	0.44609	1.53458	0.79103	3.77675
飞机 (可见光)	0.47866	1.37811	0.82633	3.21457

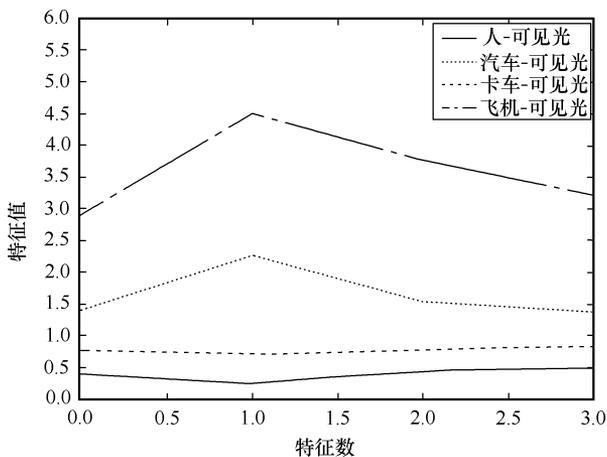


图 10-2 灰度共生矩阵的特征值曲线图

提取图 10-1 中部分可见光与红外图像的 Hu 不变矩特征值见表 10-2。

表 10-2 Hu 不变矩特征值

	φ_5	φ_6	φ_7	φ_8	φ_9	φ_{10}	φ_{11}
人 (可见光)	2.63209	5.70381	10.3790	10.2179	20.5167	13.0699	21.9237
汽车 (可见光)	2.61453	5.50907	11.9081	10.8593	22.2562	14.0645	22.8579
卡车 (可见光)	2.50155	5.26340	10.4782	10.0683	21.0262	13.4318	20.3510
飞机 (可见光)	2.73477	5.84439	10.2342	10.6508	22.4924	13.8035	21.0937
人 (红外)	2.54830	5.68260	9.61936	10.9072	21.3638	14.2065	21.2853
汽车 (红外)	2.68970	5.56961	9.56586	9.68588	19.3145	12.4722	20.2585
卡车 (红外)	2.69772	5.51222	9.97093	10.2700	20.3916	13.0262	21.5395
飞机 (红外)	2.80530	5.93502	9.40829	9.70917	19.2880	12.7031	19.7948

提取图 10-1 中部分可见光与红外图像的仿射不变矩特征值见表 10-3。

表 10-3 仿射不变矩特征值

	φ_{12}	φ_{13}	φ_{14}	φ_{15}	φ_{16}	φ_{17}
人 (可见光)	5.29765	14.6616	11.1345	15.5509	9.5191	14.3592
汽车 (可见光)	5.22568	15.0976	11.1388	15.2136	9.37162	14.5834
卡车 (可见光)	5.06785	14.6033	10.787	14.7399	9.03077	14.0719
飞机 (可见光)	5.74418	17.2043	11.8656	17.0144	10.353	16.2898
人 (红外)	5.30922	15.1088	10.979	15.8743	9.50448	14.6908
汽车 (红外)	5.30278	14.4494	11.5489	11.5489	9.52838	14.4513
卡车 (红外)	5.04291	13.4845	11.2943	14.577	9.10079	13.6846
飞机 (红外)	5.45966	15.4499	11.8045	15.8372	9.80844	15.2156

提取图 10-1 中部分可见光与红外图像的小波不变矩特征值见表 10-4。

表 10-4 小波不变矩特征值

	φ_{18}	φ_{19}	φ_{20}	φ_{21}	φ_{22}	φ_{23}	φ_{24}
人 (可见光)	11.0745	1.34946	1.84384	0.22468	0.15887	7.83086	0.95422
汽车 (可见光)	12.585	2.68857	2.09533	0.44763	0.31652	8.89897	1.90111
卡车 (可见光)	10.1098	4.31766	1.68321	0.71886	0.50831	7.14868	3.05305
飞机 (可见光)	10.6718	3.06946	1.7768	0.51105	0.36137	7.54612	2.17044
人 (红外)	9.23674	1.24159	1.53786	0.20672	0.14617	6.53136	0.87794
汽车 (红外)	18.7286	1.12235	3.1182	0.18686	0.13213	0.13213	0.79362
卡车 (红外)	22.0282	0.83549	3.66756	0.1391	0.09836	15.5763	0.59078
飞机 (红外)	26.2281	1.17233	4.36681	0.19519	0.13802	18.546	0.82897

提取图 10-1 中部分可见光与红外图像的 Zernike 不变矩特征值见表 10-5。

表 10-5 Zernike 不变矩特征值

	φ_{26}	φ_{27}	φ_{28}	φ_{29}	φ_{30}	φ_{31}
人 (可见光)	58.0556	143.715	0.85359	2.45495	0.12537	0.64544
汽车 (可见光)	80.3201	230.19	0.95885	3.54656	0.05969	0.02158
卡车 (可见光)	61.9834	178.208	0.79413	3.03872	0.08932	0.00893
飞机 (可见光)	53.1681	147.787	0.21789	0.822242	0.16815	0.04079
人 (红外)	49.7788	126.408	2.85157	9.59831	0.30807	0.16238
汽车 (红外)	79.5074	225.56	0.36124	1.20259	0.19240	0.02304
卡车 (红外)	117.976	344.095	0.63488	2.41283	0.14489	0.01845
飞机 (红外)	127.602	366.791	0.73832	2.74678	0.12007	0.02838

由表 10-1 ~ 表 10-5 的特征数据可以看出，四种特征提取方法提取的特征数据都在不同程度上与其他目标有区别，能够把两个不同的目标区分开，但是不同目标的同一特征之间存在着相似性和冗余性，这样在识别的时候就会受到干扰，容易影响识别结果。由表 10-3 和表 10-4 可以看出，提取的仿射不变矩和小波矩特征都有很好的聚类性，不同目标之间距离较大，区分性较强。灰度共生矩阵也有很好的聚类性和区分能力。由表 10-3 可以看出，Hu 不变矩随着矩阶数的增加，它的特征值有增加的趋势，计算量也随之增大。Zernike 矩随着阶数的增加相似性也随之增大，冗余信息增多。

由以上分析可知，如果只是单纯地把目标特征串联在一起直接用于识别，那么就会受到冗余特征的干扰，不能对目标有很好的区分效果。所以有必要在识别之前对特征进行融合处理，也可以称作特征选择，选出能用于区分目标的有效特征，降低特征维数，去除冗余信息，提高识别效率，同时也减少计算量。

选取可见光图像的灰度共生矩阵、Hu 矩和红外图像的小波矩特征作为实验数据，实验在 MATLAB 7.1 环境中进行，表 10-6 为 FastICA 方法提取出的独立成分。

表 10-6 对特征数据提取的独立成分

序 号	第一独立成分	第二独立成分	第三独立成分	第四独立成分	第五独立成分
1	1.003771	0.763324	-0.817523	-1.008725	0.684751
2	-0.876953	1.005321	0.756894	-0.947521	0.894562
3	1.003881	-0.992227	1.038942	0.563624	-0.774568
4	-1.002177	0.699915	-0.654983	0.892145	0.645829

(续)

序 号	第一独立成分	第二独立成分	第三独立成分	第四独立成分	第五独立成分
5	1.011362	-0.995892	-0.875246	-0.110547	1.000895
6	-1.002459	-1.005471	-1.015894	0.985246	1.002479
7	0.562757	1.011563	0.998514	0.772549	-0.789624
8	-1.005646	-1.005409	0.847569	-0.578921	0.994816

最后，在 MATLAB 7.1 环境中对传统的串联法和快速固定点 ICA 方法进行融合识别实验。使用支持向量机方法对两种方法的融合特征进行分类识别，识别率见表 10-7。

表 10-7 算法识别率比较

	传统串联法识别率 (%)	ICA 融合法识别率 (%)
人员目标	86.66	91.55
汽车目标	84.44	89.55
卡车目标	82.22	87.38
飞机目标	85.36	88.45

分析比较表 10-7 中的数据可以看出，快速固定点 ICA 方法比传统的串联算法识别率要高，有更好的性能。

10.7 本章小结

本章主要研究了基于 ICA 的特征融合方法。首先阐述了 ICA 的定义和随机变量独立性的概念，以及 ICA 独立性的度量标准，包括非高斯性、峭度、负熵和互信息。然后重点介绍了快速固定点 ICA 算法。最后给出 ICA 数据预处理的方法，即中心化和白化，将 ICA 算法应用到红外与可见光图像特征级融合中，在图像库中选取了四种图像，即人、汽车、卡车、飞机，提取图像的特征包括灰度共生矩阵、Hu 不变矩、仿射不变矩、小波矩和 Zernike。

对典型相关分析特征融合方法的改进

典型相关分析 (Canonical Correlation Analysis, CCA) 是研究两组变量之间相关关系的多元统计分析方法, 它能揭示两组变量之间的内在联系。CCA 的目的是识别并量化两组变量之间的联系, 将两组变量之间的相关关系分析, 转化为一组变量的线性组合和另一组变量的线性组合之间的相关关系分析。CCA 的应用不论在宏观经济分析还是市场分析, 不论在农业研究还是医学应用上, 还有气象学、管理学、教育学等领域都取得了很好的效果, 它在理论研究上的深刻意义也受到许多统计学学者的重视。从某种意义上讲多元回归分析、判别分析或者对应分析等许多重要的数据分析方法都可以归结为 CCA 的一种特例, 同时它还是偏最小二乘回归分析的理论基石。近些年来, CCA 方法被一些学者引入到图像处理、信号处理、计算机视觉、语音识别及人脸识别等领域, 并取得了一定进展, 同时也成为一种新的特征融合技术。

11.1 CCA 的基本思想

CCA 的研究焦点是一组变量的线性组合和另一组变量的线性组合之间的相关关系, 借助具有最大相关性的线性组合来描述两组随机变量之间的关系。它的基本思想与主成分分析非常类似, 首先分别为两组变量寻找一个线性组合, 并且要求提取出的一对线性组合之间具有最大相关性, 然后接着选取与这对线性组合最不相关的另外一对线性组合, 并且要求选取的两组线性组合之间的相关系数最大, 如此继续下去, 每一次都提取一对具有最大相关性且与前面的组合互不相关的线性组合, 直到把原来两组变量之间的相关性提取完为止。被选出的每一对线性组合都被称为典型相关变量, 它们的相关系数被称为典型相关系数。典型相关系数度量了这两组变量之间联系的强弱, 而选用最大相关系数则是通过努力将两组变量间的高维关系浓缩到少数几对典型相关变量来体现。Sun 等人提出的特征融合方法, 就是以 CCA 方法中求最大相关系数的函数为准则函数, 通过求解最大特征值所对应的特征向量来确定投影轴, 两组向量通过投影轴的变换得到具有最大相关性的典型相关变量, 即为要求解的融合特征。

11.2 CCA 的基本原理

11.2.1 CCA 的数学描述

设有两随机变量组 $\mathbf{X} = [X_1, X_2, \dots, X_p]'$ 和 $\mathbf{Y} = [Y_1, Y_2, \dots, Y_q]'$, 不妨设 $p \leq q$ 。

对于 \mathbf{X} 、 \mathbf{Y} , 设第一组变量的均值和协方差矩阵为 $E(\mathbf{X}) = \boldsymbol{\mu}_1$, $\text{cov}(\mathbf{X}) = \boldsymbol{\Sigma}_{11}$ 。

第二组变量的均值和协方差矩阵为 $E(\mathbf{Y}) = \boldsymbol{\mu}_2$, $\text{cov}(\mathbf{Y}) = \boldsymbol{\Sigma}_{22}$ 。第一组与第二组变量的协方差矩阵为 $\text{cov}(\mathbf{X}, \mathbf{Y}) = \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}'_{21}$ 。于是, 对于矩阵 $\mathbf{Z} = \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}$ 有均值向量

$$\boldsymbol{\mu} = E(\mathbf{Z}) = E \begin{bmatrix} E(\mathbf{X}) \\ E(\mathbf{Y}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \quad (11-1)$$

则协方差矩阵为

$$\begin{aligned} \sum_{(p+q)(p+q)} &= E(\mathbf{Z} - \boldsymbol{\mu})(\mathbf{Z} - \boldsymbol{\mu})' \\ &= \begin{bmatrix} E(\mathbf{X} - \boldsymbol{\mu}_1)(\mathbf{X} - \boldsymbol{\mu}_1)'E(\mathbf{X} - \boldsymbol{\mu}_1)(\mathbf{Y} - \boldsymbol{\mu}_2)' \\ E(\mathbf{Y} - \boldsymbol{\mu}_2)(\mathbf{X} - \boldsymbol{\mu}_1)'E(\mathbf{Y} - \boldsymbol{\mu}_2)(\mathbf{Y} - \boldsymbol{\mu}_2)' \end{bmatrix} \\ &= \begin{bmatrix} \sum_{(p \times p)} & \sum_{(p \times q)} \\ \sum_{(q \times p)} & \sum_{(q \times q)} \end{bmatrix} \end{aligned}$$

要研究两组变量 X_1, X_2, \dots, X_p 和 Y_1, Y_2, \dots, Y_q 之间的相关关系, 首先分别作两组变量的线性组合, 即

$$\begin{aligned} \mathbf{U} &= a_1 X_1 + a_2 X_2 + \dots + a_p X_p = \mathbf{a}'\mathbf{X} \\ \mathbf{V} &= b_1 Y_1 + b_2 Y_2 + \dots + b_q Y_q = \mathbf{b}'\mathbf{Y} \end{aligned} \quad (11-2)$$

式中, $\mathbf{a} = (a_1, a_2, \dots, a_p)'$, $\mathbf{b} = (b_1, b_2, \dots, b_q)'$ 为任意非零常系数向量, 则可得

$$\begin{aligned} \text{var}(\mathbf{U}) &= \mathbf{a}'\text{cov}(\mathbf{X})\mathbf{a} = \mathbf{a}'\boldsymbol{\Sigma}_{11}\mathbf{a} \\ \text{var}(\mathbf{V}) &= \mathbf{b}'\text{cov}(\mathbf{Y})\mathbf{b} = \mathbf{b}'\boldsymbol{\Sigma}_{22}\mathbf{b} \\ \text{cov}(\mathbf{U}, \mathbf{V}) &= \mathbf{a}'\text{cov}(\mathbf{X}, \mathbf{Y})\mathbf{b} = \mathbf{a}'\boldsymbol{\Sigma}_{12}\mathbf{b} \end{aligned}$$

式中, \mathbf{U} 与 \mathbf{V} 称为典型相关变量, 它们之间的相关系数 ρ 称为典型相关系数, 可以由下式求出 ρ :

$$\rho = \text{cov}(\mathbf{U}, \mathbf{V}) = \frac{\mathbf{a}'\boldsymbol{\Sigma}_{12}\mathbf{b}}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_{11}\mathbf{a}}\sqrt{\mathbf{b}'\boldsymbol{\Sigma}_{22}\mathbf{b}}} \quad (11-3)$$

式中, $\boldsymbol{\Sigma}_{11}$ 为变量 \mathbf{X} 的协方差矩阵; $\boldsymbol{\Sigma}_{22}$ 为变量 \mathbf{Y} 的协方差矩阵; $\boldsymbol{\Sigma}_{12}$ 为变量 \mathbf{X} 与

Y 的互协方差矩阵。

11.2.2 典型相关的定义及导出

设有两组相互关联的随机向量 $X' = [X_1, X_2, \dots, X_p]$ 和 $Y' = [Y_1, Y_2, \dots, Y_q]$, U_i 、 V_i 分别为在两组变量内选取的几个具有代表性的综合变量, 每个综合变量都是原变量的一个线性组合, 如下所示:

$$U_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p = a'X \quad (11-4)$$

$$V_i = b_{i1}Y_1 + b_{i2}Y_2 + \dots + b_{iq}Y_q = b'Y \quad (11-5)$$

CCA 要做的工作就是要找到一对投影方向 a 、 b , 使得投影 $U_i = a'X$ 和 $V_i = b'Y$ 之间具有最大的相关性, 这种相关被称为典型相关。 $U_1 = a'_1X$, $V_1 = b'_1Y$ 为 X 、 Y 的第一对典型相关变量, 类似地可以求出第 2 对 (U_2, V_2) , 第 3 对 (U_3, V_3) , \dots , 第 i 对 (U_i, V_i) , 而且每对典型相关变量之间互不相关。利用这些典型相关变量就可以反映出 X 与 Y 之间的线性相关情况。

一般地, 投影方向 a 与 b 通过最大化 U 与 V 之间的相关系数 ρ 获得。

由于随机变量 U 、 V 与任意常数相乘并不改变它们之间的相关系数, 所以为防止结果重复, 令

$$\begin{cases} \text{var}(U) = \text{var}(a'X) = a'\Sigma_{11}a = 1 \\ \text{var}(V) = \text{var}(b'Y) = b'\Sigma_{22}b = 1 \end{cases} \quad (11-6)$$

所以

$$\rho = \text{cov}(a'X, b'Y) = a'\text{cov}(X, Y)b = a'\Sigma_{12}b \quad (11-7)$$

于是求解问题即在式 (11-6) 约束下, 求出 $a \in R^p$, $b \in R^q$ 使得式 (11-7) 达到最大。

根据拉格朗日定理可将问题转化为求

$$\varphi(a, b) = a'\Sigma_{12}b - \frac{\lambda}{2}(a'\Sigma_{11}a - 1) - \frac{v}{2}(b'\Sigma_{22}b - 1) \quad (11-8)$$

的极大值, 其中 λ 、 v 是拉格朗日乘数。

根据求极值的必要条件, 得

$$\begin{cases} \frac{\partial \varphi}{\partial a} = \Sigma_{12}b - \lambda \Sigma_{11}a = 0 \\ \frac{\partial \varphi}{\partial b} = \Sigma_{21}a - v \Sigma_{22}b = 0 \end{cases} \quad (11-9)$$

将式 (11-9) 的两式分别左乘 a' 与 b' , 则得

$$\begin{cases} a'\Sigma_{12}b - \lambda a'\Sigma_{11}a = 0 \\ b'\Sigma_{21}a - v b'\Sigma_{22}b = 0 \end{cases} \quad (11-10)$$

即有

$$\begin{cases} a'\Sigma_{12}b = \lambda a'\Sigma_{11}a = \lambda \\ b'\Sigma_{21}a = v b'\Sigma_{22}b = v \end{cases} \quad (11-11)$$

因为 $(\mathbf{b}'\Sigma_{21}\mathbf{a}) = \mathbf{a}'\Sigma_{12}\mathbf{b}$, 所以 $\lambda = v = \mathbf{a}'\Sigma_{12}\mathbf{b}$, 可知 λ 为线性组合 U 、 V 的相关系数。用 λ 代替式中的 v , 则式 (11-11) 写为

$$\begin{cases} \Sigma_{12}\mathbf{b} - \lambda\Sigma_{11}\mathbf{a} = 0 \\ \Sigma_{21}\mathbf{a} - \lambda\Sigma_{22}\mathbf{b} = 0 \end{cases} \quad (11-12)$$

假定各随机变量协方差矩阵都有逆矩阵, 则由式 (11-12) 中的第二式可得

$$\mathbf{b} = \frac{1}{\lambda}\Sigma_{22}^{-1}\Sigma_{21}\mathbf{a} \quad (11-13)$$

将式 (11-13) 代入式 (11-12) 的第一式, 得

$$\frac{1}{\lambda}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\mathbf{a} - \lambda\Sigma_{11}\mathbf{a} = 0 \quad (11-14)$$

即有

$$\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\mathbf{a} - \lambda^2\Sigma_{11}\mathbf{a} = 0 \quad (11-15)$$

同理, 由式 (11-12) 可得

$$\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\mathbf{b} - \lambda^2\Sigma_{22}\mathbf{b} = 0 \quad (11-16)$$

用 Σ_{11}^{-1} 和 Σ_{22}^{-1} 分别左乘式 (11-15) 和式 (11-16), 得

$$\begin{cases} \Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\mathbf{a} - \lambda^2\mathbf{a} = 0 \\ \Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\mathbf{b} - \lambda^2\mathbf{b} = 0 \end{cases} \quad (11-17)$$

即

$$\begin{cases} (\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \lambda^2\mathbf{I}_p)\mathbf{a} = 0 \\ (\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \lambda^2\mathbf{I}_q)\mathbf{b} = 0 \end{cases} \quad (11-18)$$

由上述分析可知, $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ 和 $\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ 具有相同的特征根 λ^2 , \mathbf{a} 、 \mathbf{b} 是其特征根对应的特征向量。特征根 λ 即为我们要求的典型相关系数, \mathbf{a} 、 \mathbf{b} 为典型相关向量。

由于我们所求的是最大特征根及其对应的特征向量, 因此, 最大特征根 λ_1^2 对应的特征向量 $\mathbf{a}_1 = (a_{11}, a_{12}, \dots, a_{1p})'$ 和 $\mathbf{b}_1 = (b_{11}, b_{12}, \dots, b_{1q})'$ 就是所求的典型变量的系数向量, 即可得

$$\mathbf{U}_1 = \mathbf{a}'_1\mathbf{X} = a_{11}\mathbf{X}_1 + a_{12}\mathbf{X}_2 + \dots + a_{1p}\mathbf{X}_p \quad (11-19)$$

$$\mathbf{V}_1 = \mathbf{b}'_1\mathbf{Y} = b_{11}\mathbf{Y}_1 + b_{12}\mathbf{Y}_2 + \dots + b_{1q}\mathbf{Y}_q \quad (11-20)$$

式中, \mathbf{U}_1 、 \mathbf{V}_1 被称为第一对典型变量, 最大特征根的平方根 λ_1 为两典型变量的相关系数, 称为第一典型相关系数。

第二对典型变量 $\mathbf{U}_2 = \mathbf{a}'_2\mathbf{X}$ 和 $\mathbf{V}_2 = \mathbf{b}'_2\mathbf{Y}$ 需要满足如下约束条件:

$$D(\mathbf{U}_2) = \mathbf{a}'_2\Sigma_{11}\mathbf{a}_2 = 1 \quad (11-21)$$

$$D(\mathbf{V}_2) = \mathbf{b}'_2\Sigma_{22}\mathbf{b}_2 = 1$$

另外, 为了有效测量两组变量之间的相关信息, 第二对典型变量应不再包含第一对典型变量中已有的信息, 因此, 需要增加约束条件:

$$\begin{aligned} \text{cov}(U_1, U_2) &= \text{cov}(a'_1 X, a'_2 X) = a'_1 \Sigma_{11} a_2 = 0 \\ \text{cov}(V_1, V_2) &= \text{cov}(b'_1 Y, b'_2 Y) = b'_1 \Sigma_{22} b_2 = 0 \end{aligned} \quad (11-22)$$

λ_2 为第二典型相关系数。

如此下去，依次可求出第 r 对典型变量： $U_r = a'_r X$ 和 $V_r = b'_r Y$ ，其系数向量 a_r 和 b_r 分别为矩阵 $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$ 和 $\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$ 的第 r 特征根 λ_r^2 对应的特征向量， λ_r 即为第 r 典型相关系数。

综上所述，计算典型相关变量和典型相关系数就是要求解矩阵 A 和 B 的特征根及相应的特征向量。若矩阵 $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$ 和 $\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$ 的秩为 r ，则共有 r 对典型相关变量， r 个典型相关系数。

11.2.3 CCA 的基本性质

设 $A = \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$ ， $B = \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$ ，可以证明， A 和 B 的特征向量和特征根具有如下性质：

- 1) A 和 B 具有相同的非零特征根，且所有特征根非负。
- 2) A 和 B 的特征根均在 $0 \sim 1$ 之间。
- 3) 设 A 和 B 的非零特征根为 $\lambda_1^2 \geq \lambda_2^2 \geq \dots \geq \lambda_r^2$ ， $r = \text{rank}(A) = \text{rank}(B)$ ， $a^{(1)}$ ， $a^{(2)}$ ， \dots ， $a^{(r)}$ 为 A 对应于 λ_1^2 ， λ_2^2 ， \dots ， λ_r^2 的特征向量， $b^{(1)}$ ， $b^{(2)}$ ， \dots ， $b^{(r)}$ 为 B 对应于 λ_1^2 ， λ_2^2 ， \dots ， λ_r^2 的特征向量，即典型变量。

典型变量具有如下性质：

- (1) 同一组的典型变量互不相关

设 x 、 y 的第 i 对典型变量为

$$u_i = a'_i x, v_i = b'_i y, i = 1, 2, \dots, m \quad (11-23)$$

则有

$$\begin{aligned} V(u_i) &= a'_i \Sigma_{11} a_i = 1, V(v_i) = b'_i \Sigma_{11} b_i = 1, i = 1, 2, \dots, m \\ \rho(u_i, u_j) &= \text{cov}(u_i, u_j) = a'_j \Sigma_{11} a_j = 0, 1 \leq i \neq j \leq m \\ \rho(v_i, v_j) &= \text{cov}(v_i, v_j) = b'_i \Sigma_{22} b_j = 0, 1 \leq i \neq j \leq m \end{aligned} \quad (11-24)$$

式 (11-24) 表明 x 组成的第一组典型变量 u_1, u_2, \dots, u_m 互不相关，且均有相同的方差 1；由 y 组成的第二组典型变量 v_1, v_2, \dots, v_m 也互不相关，且也均有相同的方差 1。

- (2) 不同组的典型变量之间的相关性

$$\begin{aligned} \rho(u_i, v_j) &= \rho_{ij}, i = 1, 2, \dots, m \\ \rho(u_i, v_j) &= \text{cov}(u_i, v_j) = \text{cov}(a'_i x, b'_j y) = a'_i \text{cov}(x, y) b_j = a'_i \Sigma_{12} b_j \\ &= (\Sigma_{11}^{-\frac{1}{2}} \alpha_i)' \Sigma_{12} (\Sigma_{22}^{-\frac{1}{2}} \beta_j)' = \alpha'_i \Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-\frac{1}{2}} \beta_j \\ &= \alpha'_i (\Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-\frac{1}{2}} \beta_j) = \rho_{ij} \alpha'_i \left(\frac{1}{\rho_j} \Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-\frac{1}{2}} \beta_j \right) \end{aligned}$$

$$V \begin{pmatrix} x \\ y \end{pmatrix} = \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} & \sigma_{1,p+1} & \sigma_{1,p+2} & \cdots & \sigma_{1,p+q} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} & \sigma_{2,p+1} & \sigma_{1,p+2} & \cdots & \sigma_{2,p+q} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} & \sigma_{p,p+1} & \sigma_{p,p+q} & \cdots & \sigma_{p,p+q} \\ \hline \sigma_{p+1,1} & \sigma_{p+1,2} & \cdots & \sigma_{p+1,p} & \sigma_{p+1,p+1} & \sigma_{p+1,p+2} & \cdots & \sigma_{p+1,p+q} \\ \sigma_{p+2,1} & \sigma_{p+2,2} & \cdots & \sigma_{p+2,p} & \sigma_{p+2,p+1} & \sigma_{p+q,p+2} & \cdots & \sigma_{p+2,p+q} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{p+q,1} & \sigma_{p+q,2} & \cdots & \sigma_{p+q,p} & \sigma_{p+q,p+1} & \sigma_{p+q,p+2} & \cdots & \sigma_{p+q,p+q} \end{pmatrix}$$

$$\begin{aligned} \text{cov}(x, u) &= \text{cov}(x, A'x) = \Sigma_{11}A \\ \text{cov}(x, v) &= \text{cov}(x, B'y) = \Sigma_{12}B \\ \text{cov}(y, u) &= \text{cov}(y, A'x) = \Sigma_{21}A \\ \text{cov}(y, v) &= \text{cov}(y, B'y) = \Sigma_{22}B \end{aligned}$$

11.3 典型相关变量和相关系数的求解步骤

1. 计算原始数据的协方差矩阵

设有两组变量， X 代表第一组的 p 个变量， Y 代表第二组的 q 个变量，不妨假设 $p \leq q$ ，令

$$X_{(p+q) \times 1} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \\ \hline Y_1 \\ Y_2 \\ \vdots \\ Y_q \end{bmatrix}, \quad \text{cov}(X, Y) = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \quad (11-27)$$

$\begin{matrix} \Sigma_{11} & \Sigma_{12} \\ (p \times p) & (p \times q) \\ \hline \Sigma_{21} & \Sigma_{22} \\ (q \times p) & (q \times q) \end{matrix}$

式中， $\Sigma_{11} = \text{cov}(X)$ ； $\Sigma_{22} = \text{cov}(Y)$ ； $\Sigma_{12} = \text{cov}(X, Y) = \Sigma_{21}'$ 。 Σ_{11} 为第一组变量的协方差矩阵， Σ_{22} 为第二组变量的协方差矩阵， Σ_{12} 和 Σ_{21} 为两组变量之间的协方差矩阵，并且有 $\Sigma_{21} = \Sigma_{12}'$ 。

2. 计算两个矩阵 A 和 B

$$A = \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \quad (11-28)$$

$$B = \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \quad (11-29)$$

式中, A 为 $p \times p$ 阶矩阵; B 为 $q \times q$ 阶矩阵。

3. 计算 A 和 B 的特征向量和特征根

$\lambda_1^2 \geq \lambda_2^2 \geq \dots \geq \lambda_r^2$ 为矩阵 A 和 B 的非零特征根, a_1, a_2, \dots, a_r 为 A 对应于 $\lambda_1^2, \lambda_2^2, \dots, \lambda_r^2$ 的特征向量, b_1, b_2, \dots, b_r 为 B 对应于 $\lambda_1^2, \lambda_2^2, \dots, \lambda_r^2$ 的特征向量, 则最大特征根 λ_1^2 对应的特征向量 $a_1 = (a_{11}, a_{12}, \dots, a_{1p})'$ 和 $b_1 = (b_{11}, b_{12}, \dots, b_{1q})'$ 就是所求的第一对典型变量的系数向量, 第一对典型相关变量如下所示:

$$U_1 = a_1' X = a_{11} X_1 + a_{12} X_2 + \dots + a_{1p} X_p \quad (11-30)$$

$$V_1 = b_1' Y = b_{11} Y_1 + b_{12} Y_2 + \dots + b_{1q} Y_q \quad (11-31)$$

依次还可以求出 r 个典型相关系数和 r 对典型相关变量。

11.4 基于 CCA 改进算法的图像特征级融合实现

11.4.1 改进算法的关键技术

孙权森等提出将 CCA 方法用于特征融合, 获得了优于串行融合和并行融合的结果, 但是该方法是基于向量形式的特征融合方法, 应用时必须将二维图像矩阵转换为一维向量形式, 通常转换后的数据量非常大, 这样在应用 CCA 进行特征融合时的运算量会很大, 运行时间长。同时, 如果没有足够的训练样本还会导致高维小样本问题, 使得 CCA 算法无法正常求解。本章研究红外图像与可见光图像的特征级融合, 提出基于 PCA-CCA 的图像特征融合方法, 首先利用 PCA 方法对样本特征进行降维, 然后在低维空间中利用 CCA 方法进行特征融合。

11.4.2 特征融合过程

本章基于 PCA-CCA 方法进行可见光图像和红外图像特征级融合的具体算法流程如下:

- 1) 提取目标图像的特征。对待识别的目标图像进行预处理, 然后对可见光图像提取灰度共生矩阵、Hu 不变矩和小波矩特征, 对红外图像提取仿射不变矩、小波矩和 Zernike 不变矩特征。
- 2) 利用 PCA 方法对可见光图像的特征向量进行降维处理, 提取出向量 X 。
- 3) 利用 PCA 方法对红外图像的特征向量进行降维处理, 提取出向量 Y 。
- 4) 将样本空间内的样本特征向量进行归一化处理。
- 5) 基于 CCA 方法的特征融合。首先计算出 X 与 Y 中样本的总协方差矩阵 Σ_{11} 、 Σ_{22} 及互协方差矩阵 Σ_{12} ; 然后计算 $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$ 和 $\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$ 矩阵的非零

特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d \geq \dots \geq 0$ ，以及对应的特征向量 \mathbf{a} 和 \mathbf{b} ；最后由 $\mathbf{U}_i = \mathbf{a}'\mathbf{X}$ 和 $\mathbf{V}_i = \mathbf{b}'\mathbf{Y}$ 分别求出典型变量，如下所示：

$$\mathbf{U}_i = \mathbf{a}_{i1}\mathbf{X}_1 + \mathbf{a}_{i2}\mathbf{X}_2 + \dots + \mathbf{a}_{ip}\mathbf{X}_p = \mathbf{a}'\mathbf{X} \quad (11-32)$$

$$\mathbf{V}_i = \mathbf{b}_{i1}\mathbf{Y}_1 + \mathbf{b}_{i2}\mathbf{Y}_2 + \dots + \mathbf{b}_{iq}\mathbf{Y}_q = \mathbf{b}'\mathbf{Y} \quad (11-33)$$

如果第一组典型变量 $(\mathbf{U}_1, \mathbf{V}_1)$ 不足以代表两组原始变量的信息，那么继续求出第二组 $(\mathbf{U}_2, \mathbf{V}_2)$ ，第三组 $(\mathbf{U}_3, \mathbf{V}_3)$ ， \dots ，第 i 组 $(\mathbf{U}_i, \mathbf{V}_i)$ ，然后将线性变换

$$\begin{pmatrix} \mathbf{U} \\ \mathbf{V} \end{pmatrix} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}^T \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \quad (11-34)$$

作为投影后的组合特征用于分类。

完整的基于 PCA-CCA 的特征融合识别过程如图 11-1 所示。

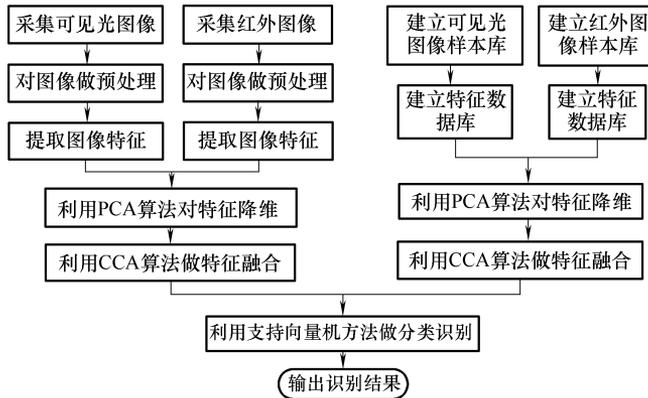


图 11-1 基于 PCA-CCA 特征融合识别过程的流程图

11.5 实验结果与分析

为了衡量本章提出的 PCA-CCA 方法的有效性，将提取的可见光与红外图像的特征组成特征向量，首先利用 PCA 方法对标准化后的特征向量进行降维，然后再利用 CCA 方法求出融合特征，最后利用支持向量机方法对融合特征进行识别分类。

图 11-2 和图 11-3 分别为可见光和红外图像的主成分 Pareto 图和累计贡献率图。

利用 CCA 方法对由 PCA 提取的特征向量 \mathbf{X} 和 \mathbf{Y} 进行融合处理，首先计算出 \mathbf{X} 与 \mathbf{Y} 中样本的总协方差矩阵 Σ_{11} 、 Σ_{22} 及互协方差矩阵 Σ_{12} ，再计算出矩阵 \mathbf{A} （即 $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ ）和矩阵 \mathbf{B} （即 $\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ ），见表 11-1 和表 11-2。

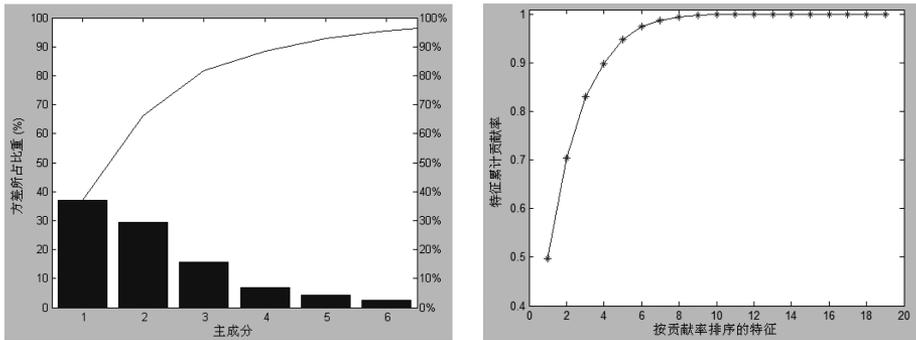


图 11-2 可见光图像各主成分及其累计贡献率图

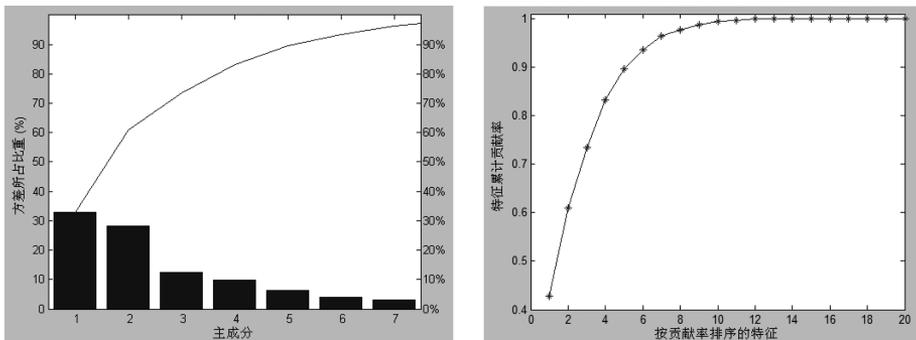


图 11-3 红外图像各主成分及其累计贡献率图

表 11-1 CCA 方法计算出的矩阵 A

1	2	3	4	5	6	7	8	9
0.6528	-0.1068	-0.0450	-0.0019	0.0352	-0.0241	-0.0219	0.0007	-0.0024
-0.1343	0.7605	-0.0310	-0.0561	-0.0621	0.0046	-0.0002	0.0206	0.0038
-0.1066	-0.0583	0.7990	0.0304	-0.0478	-0.0467	-0.0406	-0.0480	0.0392
-0.0101	-0.2441	0.0702	0.6189	-0.0563	-0.0444	0.0541	0.1404	0.0280
0.3016	-0.4237	-0.1731	-0.0882	0.6456	0.0575	0.0626	-0.0413	-0.0365
-0.3458	0.0528	-0.2836	-0.1166	0.0963	0.8255	-0.0994	0.0141	0.1353
-0.4994	-0.0037	-0.3908	0.2252	0.1663	-0.1577	0.7741	-0.0582	0.2047
0.0212	0.5057	-0.6252	0.7906	-0.1483	0.0303	-0.0787	0.5138	-0.0777
-0.1002	0.1228	0.6789	0.2096	-0.1745	0.3860	0.3683	-0.1034	0.3410

表 11-2 CCA 方法计算出的矩阵 B

1	2	3	4	5	6	7	8	9
0.8688	-0.0686	-0.0935	-0.0356	0.0726	-0.0006	-0.0079	0.0135	-0.0001
-0.0800	0.7202	0.0084	0.0895	0.1325	0.0368	0.0191	0.0246	0.0217
-0.2455	0.0189	0.6485	-0.2582	0.0866	-0.0845	-0.0646	0.0218	0.0065
-0.1186	0.2560	-0.3278	0.6132	-0.0406	-0.1167	-0.0355	-0.0356	-0.0147
0.3777	0.5911	0.1714	-0.0633	0.4306	0.0598	-0.0527	-0.0437	0.0398
-0.0047	0.2634	-0.2685	-0.2924	0.0961	0.3353	-0.0807	0.0183	-0.0344
-0.0887	0.1834	-0.2759	-0.1194	-0.1136	-0.1084	0.8215	0.0857	0.0480
0.3638	0.5688	0.2242	-0.2878	-0.2268	0.0593	0.2060	0.7742	-0.1217
-0.0039	0.6104	0.0813	-0.1452	0.2513	-0.1352	0.1406	-0.1481	0.7188

分别求出上述矩阵 A 和矩阵 B 的特征值和特征向量 a 和向量 b ，见表 11-3 和表 11-4。

表 11-3 典型相关向量矩阵 a

1	2	3	4	5	6	7	8	9
0.0286	0.0079	0.3214	0.0084	-0.0128	0.0530	-0.1738	0.0162	-0.0729
0.0099	-0.0550	0.1432	-0.1858	0.1591	0.0226	0.2646	-0.0565	0.0545
0.0644	0.0887	0.0475	-0.2088	-0.1409	-0.0833	0.0665	0.3767	0.2607
0.0469	-0.3247	0.0037	-0.1435	-0.3520	0.0972	0.1376	-0.0402	0.2870
-0.0969	-0.0359	-0.1823	-0.8092	0.0337	-0.1117	-0.5846	-0.1919	-0.4184
0.2418	0.0133	0.1330	0.0074	-0.5883	-0.2398	0.2510	-0.6398	-0.6584
0.3356	0.2624	0.2079	0.0779	0.3789	-0.5388	0.0676	-0.3645	0.3558
-0.1318	0.9018	-0.0405	-0.4504	-0.4162	0.5281	0.5175	-0.5167	0.1694
-0.8916	0.0028	0.8821	0.1935	-0.4084	-0.5840	0.4464	-0.0842	0.2785

表 11-4 典型相关向量矩阵 b

1	2	3	4	5	6	7	8	9
-0.0596	-0.1054	-0.1054	0.4688	0.0338	0.0797	-0.0225	-0.0097	0.0049
0.0708	-0.2289	0.0010	-0.0270	-0.1062	-0.0223	-0.0242	0.2495	0.1793
-0.3207	-0.0942	0.2160	-0.0619	0.0186	0.2122	-0.1038	0.1054	0.1963
-0.3722	0.1530	-0.0545	-0.1511	-0.1335	0.2926	-0.0166	-0.0467	-0.2061
0.1333	0.7212	0.1300	0.4125	-0.0335	-0.0699	-0.1624	0.3145	0.2339
-0.7786	0.0432	-0.0075	0.2258	-0.1249	-0.6911	0.1421	0.0567	0.0737
-0.2632	0.0774	-0.5813	-0.6953	0.3596	-0.1354	-0.3099	0.2349	0.2930
0.0635	0.6014	-0.7053	-0.2182	-0.4721	0.5018	0.4944	0.2357	0.6965
-0.2295	0.1202	0.2949	0.0255	0.7749	0.3313	0.7751	0.8435	0.5057

利用 CCA 方法求出的前 5 对典型相关变量见表 11-5。

表 11-5 典型相关变量

0.0144	-0.0502	0.3351	-0.0358	-0.1707	-0.5692	-0.3243	-0.4597	-0.0749
-0.0060	0.1349	0.2558	-0.2690	0.2014	0.0742	0.2026	0.7320	0.3657
-0.0649	0.0485	0.2320	0.2554	-0.3724	-0.5860	0.3166	0.1508	0.2479
-0.0832	0.1591	0.2474	-0.0974	0.2981	0.0356	-0.2665	-0.2957	0.7403
-0.1330	0.2026	0.0509	0.1053	-0.4476	0.1922	0.0517	0.3962	0.3417
-0.0263	0.1856	-0.1155	0.2429	0.2215	-0.0107	-0.0251	-0.6730	0.7898
0.0752	0.0321	-0.1182	0.1380	-0.1585	-0.3402	-0.7644	0.7493	-0.8287
0.2630	-0.0151	-0.0347	-0.0847	0.2314	0.1267	-0.3900	-0.1224	0.0143
-0.0125	0.1555	-0.1378	0.3441	0.0330	-0.5751	0.3703	-0.4068	-0.3992
0.0446	-0.1401	0.0245	-0.1762	-0.0442	-0.1649	0.4745	-0.6230	1.0226

最后，在 MATLAB 7.1 环境中使用改进的 CCA 方法进行融合识别实验，并使用支持向量机方法对融合特征进行分类识别，识别率见表 11-6。

表 11-6 三种方法识别率及运行时间比较

融合方法	平均识别率 (%)	平均运行时间/s
PCA	89	1.844
CCA	90	0.438
PCA-CCA	95	0.631

分析比较表 11-6 中数据可以看出，PCA-CCA 方法比单独的 PCA 方法和 CCA 方法识别率要高，且运行时间较短。

11.6 本章小结

本章研究了基于 CCA 的特征融合算法并对算法做出改进。首先介绍了 CCA 的原理和求解过程，然后对 CCA 进行改进，引入 PCA 算法先对原始特征向量降维，解决了 CCA 算法运算量大，运行时间长，并且有可能出现高维小样本的情况，提高了算法的效率，有效地提高了目标识别率。最后对三种图像特征级融合算法做了比较分析。

基于优化改进的反向传播神经网络目标识别

目前关于目标识别分类的方法有许多，比如支持向量机以及在它基础上发展出来的其他分类算法，如模糊支持向量机、先进支持向量机等，还有一些基于图像中目标结构进行分类的方法，除此以外，现在比较热门的人工智能，在分类算法中，神经网络也是图像处理的重要发展领域，因为神经网络模仿的是人类的大脑神经，所以在自主学习、泛化能力和高度并行性等方面更出色，能够很好地去解决非线性复杂目标识别问题，所以具有很高的研究价值和应用前景，所以在本章中将主要利用其作为识别算法来研究和验证。

目前在神经网络领域经常被使用的网络主要有径向基函数（RBF）神经网络、反向传播（BP）神经网络和 Elman 网络等多种网络。其中在理论研究中，我们通常使用得最多的是前两种神经网络。BP 神经网络属于监督学习，一般都被应用于多层感知进行训练，而在实际工业应用中，BP 神经网络虽然本身存在一些不足，但是它是被应用最多的网络。在非连续函数中，BP 神经网络使得多层前向神经网络中的权重有非常好的调整效果。在 BP 神经网络中，它能够容纳大量的输入与输出之间的映射关系，最重要的是它具有强大的逼近能力。所以本章选择 BP 神经网络来作为系统最终的识别算法进行研究和验证。

通过理论分析和查阅文献，我们知道 BP 神经网络对任何复杂的神经网络具有非常强大的非线性映射能力，如果内部机制很复杂，它是不二之选，因为本章中目标识别是一个非常复杂的机制系统，所以选择它来做识别算法。另外它的映射条件没有那么苛刻，具备非常良好的学习能力，非常强的鲁棒性，且易用计算机操作处理实现。除上面非常好的优点外，它也存在许多不足的地方，最重要的就是 BP 神经网络的学习速度慢，这也是目前研究的重点；其次它本身存在容易陷入局部极值这样的问题，这样学习训练的效果会非常差；而且受参数设置的影响比较大，如果设置不同的参数，那么它得到结果也都不会相同。所以本章根据上述 BP 神经网络存在的不足之处对其进行优化，目前因为其参数可优化性比较强，所以利用目前一些比较好的优化算法对其进行融合，来提高 BP 神经网络的学习效率和收敛速度，这样通过局部参数优化来到全局优化的目的。本章选择粒子群来对其参数进行优化处理，然后再应用于多特征信息融合目标识别中去。

因为本书前面已经详细介绍了利用多传感图像信息采集，再对采集到的图像进行有效的预处理，主要包括对图像的去噪和图像增强，然后通过各种特征提取算法将图像中的目标特征提取，然后利用特征融合算法将获得的目标特征数据进行有效可靠的融合处理，本章将进行目标识别，所以下面主要对基于粒子群去优化神经网络进行论述和实验验证其有效性。

12.1 BP 神经网络

BP 神经网络是由 Rumelhart 和 McClelland 等科学家于 1986 年正式提出来的，它属于遵循误差逆传播算法进行训练的多层前馈网络，根据目前我们所知道的人工神经网络的许多模型，在这么多模型中被使用最多的一种模型就是 BP 神经网络模型，BP 神经网络对输入和输出的映射关系学习能力比较强大，而且它能存储这样的映射关系，并且不需要任何事前对映射描述的数学公式。BP 神经网络主要运用的学习方法是速下降法，通过利用网络中的方向传播来达到对网络中的权值和阈值不停调整的目的，这样最终能够使得该网络中的误差平方和值达到最小。神经网络分为前向型和后向型，而 BP 神经网络属于前向型，并且它本身具备监督学习方式，在一些分类处理中的表现是非常优秀的，现在已经被广泛地应用于模式识别中。

12.1.1 BP 网络模型

BP 神经网络主要将复杂的分类进行非线性映射到高维空间，这样使得比直接映射到低维空间更有可能是线性可分，简而言之就是在低维空间不一定是线性可分，但是映射到高维空间上就可能存在线性可分。通过对 BP 神经网络结构分析，我们主要将其内部结构分为三层：输入层，隐层，输出层。如图 12-1 所示，输入层主要由一些感知单元组成，好比我们的嗅觉，就是通过嗅觉感知细胞将外界的信息传达到我们的神经；第二层是隐层，它是主要包含一些激励函数，如阈值函数、S 形函数、阶跃函数等，主要负责对输入信号响应，然后进行转换输出，所以它具有局部逼近的能力；第三层输出层是线性的，主要是为输入层的输入向量提供相应的响应输出。其中如何能够设计一个好的 BP 神经网络，关键在于如何设计隐层中节点的个数。

12.1.2 BP 神经网络结构

在 BP 神经网络的结构中，它主要是由三部分组成，即输入层、隐层和输出层，每层中都包含了大量的神经元，该神经元模型可以用图 12-1 来描述。

根据图 12-1 可知，其中 x_1, x_2, \dots, x_n 表示神经元的输入量， w_1, w_2, \dots ,

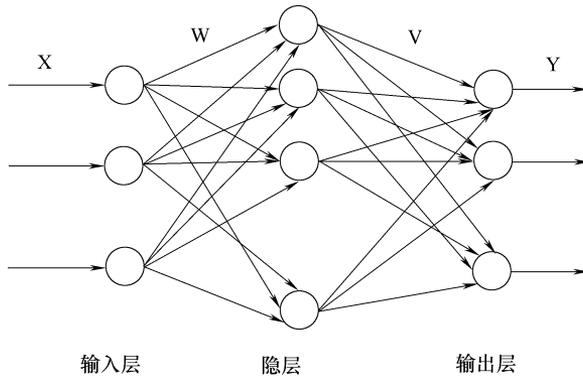


图 12-1 BP 神经网络结构图

w_n 表示网络结构层之间的权值，简单地讲就是层之间连接的强弱；输入量对应权值的和为 $\sum_{i=1}^n w_i x_i$ ，它代表神经元的激活值，当这个值大于神经元自身的阈值 θ 时，此时神经网络就可以学习训练了； O 表示神经元最终的输出量。那么根据上述介绍，此时输出量就可以用下面的公式来表示，即

$$O = f \left(\sum_{i=1}^n w_i x_i - \theta \right) \tag{12-1}$$

在式 (12-1) 中，函数 f 代表神经元内输入与输出之间的关系，一般称这样的函数为激活函数或者是输出函数。在神经元内，我们可以用 W 来作为神经元中相应的权值量，用 X 表示神经元的输入量，它们可以分别用下式来表示：

$$W = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \tag{12-2}$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{12-3}$$

我们可以通过对激活值进行假设，即 $\text{net} = W^T X$ ，类似于我们对神经元进行外部刺激来影响神经元内部反应。此时式 (12-1) 可简化为 $f(\text{net} - \theta)$ 。在实际使用中，存在很多类型的激活函数，但是一般常用的有这么几种，即 S(Sigmoid) 形函数、阈值函数和双面正切函数，它们的函数图如图 12-2 所示。

阈值函数，该函数公式如下：

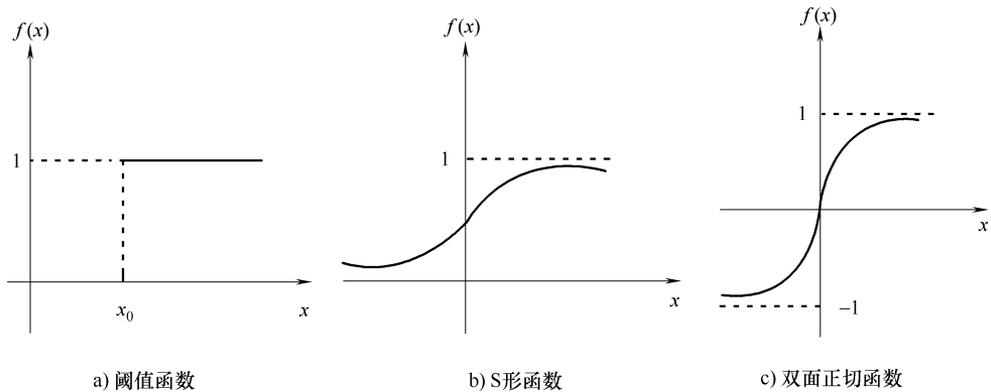


图 12-2 常用激活函数图

$$f(x) = \begin{cases} 1, & x \geq x_0 \\ 0, & x < x_0 \end{cases} \quad (12-4)$$

S 形函数，该函数公式如下：

$$f(x) = \frac{1}{1 + e^{-x}}, \quad 0 < f(x) < 1 \quad (12-5)$$

双面正切函数，该函数公式如下：

$$f(x) = \frac{1 - e^{-ax}}{1 + e^{-ax}}, \quad -1 < f(x) < 1 \quad (12-6)$$

在对 BP 神经网络的隐层函数设计和进行选取时，一般都是选用 S 形函数，因为其可导，容错性比较强，而且它还能对信号的增益有非常好的控制能力。在式 (12-5) 中，如果参数 x 比较小，这时获取的函数值就比较大，相应的增益也会很大；反之就是增益比较小，此时神经网络不容易进入饱和状态。

BP 神经网络主要是通过有监督的学习方式来对样本进行训练，神经网络通过训练好的模型，那么网络中的神经元可以通过相应的激活值，然后从输入层通过中间的隐层，最后通过输出层进行输出，此时通过输出层中的神经元来进行相应的网络响应来实现输出。然后，通过对比实际网络中的误差与期望误差进行对比，如果达不到期望误差值，此时会通过输出层到隐层，然后回到原始输入层进行修改相应的权值，依次循环上述的过程直到达到期望的误差值或者循环的最大次数才会停止。我们称这样的修改权值过程为“误差逆传播”。BP 神经网络正是通过这样的误差逆传播来提高网络的学习效率和正确输入响应的能力。在整个神经网络运行的过程中主要存在两种信号：一种是误差信号，即网络中的实际输出和期望输出之间的误差值；另一种是工作信号，即在输入端进行相应的输入时，最后输出端也会有相应的输出响应信号。

另外 BP 神经网络还具有很强的映射能力，假设存在一组输入向量 x ，并有

$x \in R^n$ ，那么 $x = (x_0, x_1, \dots, x_{n-1})^T$ ，此时如果网络中隐层的神经元一共有 n_1 ，这里可以用 x' 来表示输入向量对应隐层中的输出，那么 $x' = (x'_0, x'_1, \dots, x'_{n_1-1})^T$ ，假设在输出层中神经元总数为 m ，且 $y \in R^m$ ， $y = (y_0, y_1, \dots, y_{m-1})^T$ 。在网络中的输入层到隐层之间的权值可以用 w_{ij} 来表示，相应的阈值可记为 θ_j ，隐层到输出层之间的权值可用 w'_{jl} 来表示，相应的阈值可用 θ'_l 来表示。此时网络中各层相应的输出可表示如下：

$$\begin{cases} x'_j = f \left(\sum_{i=0}^{n_1-1} w_{ij} x_i - \theta_j \right), & j = 0, 1, 2, \dots, n_1 - 1 \\ y_l = f \left(\sum_{k=0}^{n_2} w'_{lk} x'_k - \theta'_l \right), & l = 0, 1, 2, \dots, m - 1 \end{cases} \quad (12-7)$$

式 (12-7) 可以看成是 n 维向 m 维空间进行映射。

12.1.3 BP 神经网络训练学习

在 BP 层神经元训练的过程中，主要通过对 w_{ji} 不停地进行有效的调整，来帮助神经元达到权值训练的目的，最后使得 w_{ji} 与 u_i 无限接近。当调整到一个极限值时，此时 BP 神经元的输出是 1，但是条件是在 BP 神经网络隐层中的神经元必须处于 $w_{ji} = u_i$ 时，神经元的输出 1 才会有效。当神经网络内部的训练已经全部结束时，输入层有向量信号输入到此神经网络中，此时在 BP 神经网络中，它里面的所有神经元，不管是输入层、输出层还是隐层都会根据输入向量与其相应的权值向量对比，如果互相比接近就对其进行有效的结果输出。如果在相应的对比过程中出现与权值向量的差距比较大，这时神经网络的输出是个非常小的数，反之就是上述提到的输出为 1。因为 BP 神经网络中的输出层可以为线性，若此时输出层的神经元是线性函数，则此时的神经网络输出是经过加权并求和作为输出的结果。通过理论分析和推导得出的结论是，在 BP 内部只要存在所需要的神经元个数，此时 BP 神经网络能够精确地无限逼近存在的任何有效函数。

BP 神经网络中还必须选取合适的参数，其中隐层中神经元的个数确定是主要参数问题，其中主要对神经网络中的层数、神经元总数、设置初始值、学习速率和期望误差等进行分析。在 3 层 BP 神经网络中，根据式 (12-7) 可知，它是可以将 n 维向 m 维空间进行映射，这样就能够对任何函数进行逼近。在隐层中的神经元，因为神经元的数量直接影响了训练的效果，进而影响最终的识别结果。但是当在隐层的神经元数量过大的时候，这时候网络是非常复杂的，这样会导致训练时间过长。一般都是通过在一定范围内进行随机取数的方式来解决，当然也可以通过以下公式来作为理论参考，即

$$r = \sqrt{m+n} + j \quad (12-8)$$

式中， r 代表隐层中神经元的总数， m 代表输入层中神经元的个数， n 代表输出层中神经元的个数， j 一般在 $1 \sim 10$ 中取整。

除上述以外，我们还得对权值进行相应的初始化，它的大小直接关系到学习的时间长短，一般取值是在 $(-1, 1)$ 或者 $[-2.4/n, 2.4/n]$ 中的随机数，其中 n 表示输入特征量的数目。为了保证系统能够稳定对输入的样本进行学习和训练，学习率一般在 $0.01 \sim 0.8$ 之间进行取值。最后就是设置期望误差，这是决定神经网络学习迭代的一个重要条件。

图 12-3 是 BP 神经网络学训练习训练的具体流程图。

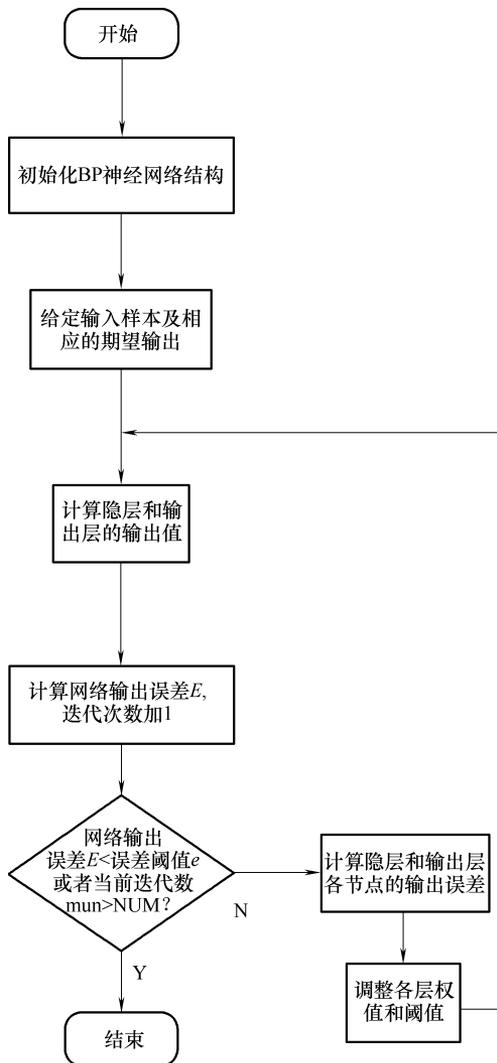


图 12-3 BP 神经网络训练流程图

12.1.4 BP 神经网络主要特点

BP 神经网络主要特点如下：

1) BP 神经网络属于前向型网络，也称为反馈神经网络，所谓前向型网络就是在该神经网络中，各神经元在接受前一层的输入后，然后将其在输出给自己的下一层，但是就是没有反馈。其中它们的节点分为两大类，即输入节点和计算节点，其中每个计算节点能够拥有许多不同的输入，但是只能存在一个输出，其中它们的输入和输出节点都是与外界相连而成的，中间层就是隐层。

2) BP 神经网络拥有强大的推广能力，因为我们获得的目标样本数据中，它们之间肯定会存在一定的相似性，所以它可以根据这种相似性来处理数据间相似性问题。正常情况下我们获取的数据一般多少会包含一定的噪声，也有可能对目标某些特征描述存在缺陷，此时通过利用该网络的推广能力就能够对失真的样本数据进行相应的恢复处理，因为现在所用的识别模式都是实时传输的，所以利用该方法能够有效地解决实时传输中图像的失真问题，这也是它被广泛使用的一个原因。

3) BP 神经网络拥有自组织和自学能力，它的自学能力主要体现在对样本数据处理的过程中。通过这样的自学能力能够主动地利用输入样本数据中的特性和相应的规律，来加强学习效果。这种学习方式不需要相应的先验知识，不需要我们在处理过程中假设具有先验的函数，而且不需要通过对样本中的空间分布进行状态假设，只是非常简单智能通过样本的数据对其进行相关关系的学习，这样大大增加了智能性。所以如果在不了解具体样本数据的情况下，利用这种学习能够帮助解决识别中未知目标的问题。

4) BP 神经网络拥有对数据并行处理的能力，这种并行处理能力类似于计算机中的多线程处理方式，这样的好处就是提高了神经网络的学习效率。

5) BP 神经网络中隐层存在许多种不同的激励函数。一般用得比较多的是非线性激励函数，有中心限幅函数、阶跃函数和 S 型函数，其中 S 型函数是用得最多的激励函数，也是目前对其研究改进和优化的热点。

BP 神经网络的缺点是，算法优化程度不是很深，它作为一种局部搜索的算法，它的权值是通过局部改善来进行相应的调整，这个时候就会出现该网络自动地陷进局部极小化这样的问题，那么它的权值就进行相应的收敛，直至达到极小的对应点，这样的后果就是训练不能成功；其次 BP 神经网络还存在收敛速度过慢这样的问题，因为在它的算法里，主要使用的是梯度下降法，由于其主要针对复杂的目标进行优化，所以花费的时间会比较长，这样就导致了其算法效率不高；另外还存在对样本过度依赖的问题，这点在 RBF 神经网络中同样存在，因为神经网络的函数的逼近主要与它本身的推广能力和学习样本的能力有

关，这样导致的结果就是训练样本选取的问题。

12.2 改进的 BP 神经网络

在本章中对 BP 神经网络的改进，主要是通过假设动量因子和自适应学习速率改善 BP 网络自身收敛速度慢和易陷入局部极小问题，然后再利用粒子群对其改进的 BP 神经网络进行优化处理。采用附加动量主要是来调整其权值，利用该方法是因为它将误差关联到了梯度上，它是基于反向传播将调整后的所有权值与调整前的权值进行正比，同时利用反向传播来产生新的权值。改进后权值调整公式如下：

$$\Delta w_{ij}(n+1) = (1+nb)\eta E_{i,x_j} + nb\Delta w_{ij}(n) \quad (12-9)$$

$$\Delta c_j(n+1) = (1-nb)\eta E_i + nb\Delta c_j(k) \quad (12-10)$$

式中， i 表示隐层数； j 表示隐层中包含的节点数； n 表示网络中训练的次数； Δw 代表权值的变化增量； η 代表网络的学习速率； E 表示网络的误差； nb 为动量因子。

由上述对 BP 神经网络的分析 and 介绍可知，如何选择学习效率对训练的速度有很大的影响，所以在网络中通过添加自适应学习算法来自动调节学习效率，这样就能相应准确地选择学习效率，它主要是通过对修改后的权值进行检查，如果该权值降低了，网络中的误差就会相应地对其添加一个量，相反就会降低学习效率，该调整公式如下所示：

$$\eta(k+1) = \begin{cases} 1.05\eta(k) & \text{SSE}(k+1) < \text{SSE}(k) \\ 0.7\eta(k) & \text{SSE}(k+1) > 1.04\text{SSE}(k) \\ (k) & \text{其他} \end{cases} \quad (12-11)$$

$$\text{SSE} = \sum_{i=1}^n (y_i - y'_i)^2 \quad (i = 1, 2, \dots, n) \quad (12-12)$$

上述对 BP 神经网络本身进行改进，因为改进本身还存在一定的局限性，所以接下来还会利用粒子群优化算法对改进后的 BP 神经网络进一步进行优化，这样会使得该网络更不容易出现陷入局部极小或者收敛过慢等不稳定的情况。

12.3 PSO 算法的基本原理和理论基础

12.3.1 PSO 算法的基本原理

粒子群优化 (Particle Swarm Optimization, PSO) 是由 Kennedy 和 Eberhart 等人在 1995 年首次提出来的方法，此算法被提出后就立即吸引了广大国内外学者对其进行研究，到现在已经有了多种改进增强版本，下面将对其原理进行详细

概述。

粒子群优化算法来源于动物群体的生活行为，通过对鸟群飞行活动的分析，发现鸟在飞行时只要跟着它附近的邻居鸟就行了，但是在空中我们能够注意到整个鸟群就像绕着某个中心在进行有方向、有序的飞行活动。简单地说就是一些全局行为都是由一些简单而不复杂的准则通过它们之间相互作用产生的，这就是群智能的主要意义。本节中 PSO 算法主要是源于鸟群在群体觅食中的行为，一群鸟在空中寻找食物的时候，为了能够最快地找到食物，它们会对含有食物区域内的鸟群进行搜寻，这样就能找到食物。PSO 算法就是通过上述对鸟类习性的研究产生的，并被用于解决一些优化问题。PSO 的基本思想就是先在相应的空间找到一颗粒子的位置，此时这颗粒子不是静止而是运动的，它们会不断改变自己的状态，它们有自己的速度，此外还需要一个目标函数对相应的适应值进行决策。当找到这些粒子后，它们自身是有记忆的，而且它们会在相应的解空间中自动地去找当前它们觉得最优的粒子，如果此时它们找到了一个更好解，接下来会基于这个解再寻找下一个更好解。在迭代过程中，粒子会利用空间中两个所谓的“极值”来对自己当前位置进行更新处理，这两个“极值”可以分为个体极值点 p_{best} 和全局极值点 g_{best} ，个体极值点是粒子本身搜寻到的最好解，全局极值点是该粒子所在群的极值点，粒子就是根据这两个极值点来调整自身的搜寻方向和速度，使之达到最佳状态。

假设存在一个大小为 D 维的搜寻空间，而且在此空间中存在一个大小为 N 的粒子群，那么这个 D 维空间向量与空间中粒子群的关系就可表示如下（下式中的 i 都表示群中对应的粒子的位置）：

$$X_i = (x_{i1}, x_{i2}, \dots, x_{iD}), i = 1, 2, \dots, N \quad (12-13)$$

在群空间中，其中粒子对应的速度如下：

$$V_i = (v_{i1}, v_{i2}, \dots, v_{iD}), i = 1, 2, \dots, N \quad (12-14)$$

粒子本身通过搜寻得到的最好解即为上述提到的个体极值，其可定义为

$$p_{best} = (p_{i1}, p_{i2}, \dots, p_{iD}), i = 1, 2, \dots, N \quad (12-15)$$

该粒子所在群得到的极值点即是上述提到的全局极值，其可定义为

$$g_{best} = (p_{g1}, p_{g2}, \dots, p_{gD}) \quad (12-16)$$

当粒子在得到上述相应的两个极值时，此时粒子会对自身相应的速度与位置做出有效的更新，它的更新表达式如下：

$$v_{id} = w^* v_{id} + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 (p_{gd} - x_{id}) \quad (12-17)$$

$$x_{id} = x_{id} + v_{id} \quad (12-18)$$

式中， c_1 和 c_2 是更新过程中的学习因子； r_1 和 r_2 表示在 $[0, 1]$ 区间内的随机数； $w^* v_{id}$ 在物理学中可定义为“动量 (momentum)”，即粒子对自身速度控制的能； $c_1 r_1 (p_{id} - x_{id})$ 是粒子群中粒子的局部“认知 (cognition)”，即粒子对历史

的记载，通过这样的记载来达到向最佳位置进行有效的逼近； $c_2 r_2 (p_{gd} - x_{id})$ 描述在粒子群中粒子间相互作用的记载，体现了粒子自身能够向粒子群进行有效逼近的能力。

12.3.2 PSO 算法的基本流程

PSO 算法的流程步骤如下：

1) 初始化：在大小为 D 维的空间中进行随机处理，然后产生相应粒子的位置和速度。

2) 粒子的评价：对在 D 维空间中的粒子进行相应的适应值 $F_{ij}[i]$ 有效计算。

3) 位置最优更新：通过上述获取的适应值，用其来和粒子本身个体极值进行相应大小的比较，如果比 $p_{best}(i)$ 好，则进行当前粒子的位置调换，然后再通过比较适应值和全局极值，如果 $F_{ij}[i] > g_{best}(i)$ ，就用全局极值点的位置更换为粒子的位置。

4) 粒子自身更新：通过式 (12-17) 和式 (12-18) 对粒子自身的速度和位置进行有效的更新处理。

5) 停止依据：通过上述进行不断循环，一般只要满足相应的最佳适应值和最大的循环迭代数就能达到停止循环的条件。

12.4 PSO 优化改进的 BP 神经网络

在 BP 神经网络中主要通过网络参数来决定它的性能的优劣，即权值和节点阈值这两个参数值。如果是利用传统的 BP 神经网络，如何确定这两个参数对神经网络学习的影响是非常巨大的，而且关键是只能通过局部空间中寻找有效的最优解这样的方式来确定整个神经网络中的参数。所以这完全不是我们想要的最佳解，最直接的一个后果就是这些参数设置不合理将引发函数逼近性能下降，更或者引起神经网络发散，所以在本章我们分析了 PSO 算法的原理和优点后，选择其来对传统的 BP 神经网络进行学习优化，主要优化神经网络中权值和阈值这些参数，这些参数在 PSO 中主要是用来表示粒子的空间具体位置。

12.4.1 PSO 优化问题分析

在对 BP 神经网络基函数参数优化之前，我们需要对下面几个问题进行分析和解决：

(1) 确定对最优激活函数中参数的评价标准

在神经网络中对其内部进行数据训练主要是想要通过获得的相应网络参数

可以使该神经网络中的平均误差 2 次方值的和达到最小，其计算公式为

$$D = \frac{1}{P} \sum_{j=1}^P \sum_{i=1}^m (y_{ji}'' - \hat{y}_{ij})^2 \quad (12-19)$$

式中，用参数 P 来表示训练样本的总数； $j=1, 2, \dots, p$ ； $i=1, 2, \dots, m$ ，参数 m 代表的是输出层在神经网络中的节点数； y_{ji}'' 表示节点输出值，即在总样本中第 j 个样本通过输出层中的第 i 个节点输出； \hat{y}_{ij} 表示节点预计输出值，即在总样本中第 j 个样本通过输出层中的第 i 个节点预计输出。

下面给出激活函数中参数评价的标准，即先激活函数中参数进行假设，即参数 $p=(c_1, \sigma_1^2, c_2, \sigma_2^2, \dots, c_k, \sigma_k^2)$ ，在参数集中，利用 $c_i=(c_{i1}, c_{i2}, \dots, c_{im})$ ， $i=1, 2, \dots, k$ ，来表示节点中 S 型激活函数对应的中心，其中用 k 表示在隐层中的节点数，用 m 表示在输入层输入特征向量对应的维数，在上述参数中用 σ_i 来表示第 i 个节点的 S 型激活函数对应的宽度，然后对输入样本的特征向量进行训练，再对其进行归一化处理，其中 $0 \leq c_{ij} \leq 1$ ， $0 \leq \sigma_i^2 \leq 1$ ， $j \leq 1, i=1, 2, \dots, n$ 。根据上述参数假设，通过假设的参数 P 来对 BP 神经网络进行有效可靠的组建，该网络平均误差和用 D 来表示，在激活函数参数的有效区域内如果对于其每一个相关的参数 P 而言，都会存在 $D(P^*) \leq D(P)$ ，则此时 P^* 就是我们所要寻找的最优激活函数参数。

(2) 粒子群中粒子初始化

假设在目标空间中粒子群的大小为 N ，那么可以通过这个空间来随机产生一个相应的粒子位置矩阵 X 和对应的速度矩阵 V ，它们两者可表示如下：

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1M} \\ \vdots & & \vdots \\ x_{N1} & \cdots & x_{NM} \end{bmatrix}, V = \begin{bmatrix} v_{11} & \cdots & v_{1M} \\ \vdots & & \vdots \\ v_{N1} & \cdots & v_{NM} \end{bmatrix}$$

$$x_{ij} = x_{\min} + (x_{\max} - x_{\min}) \times \text{rand}() \quad (12-20)$$

$$v_{ij} = -v_{\max} + 2v_{\max} \times \text{rand}() \quad (12-21)$$

式中， $M=k(n+1)$ ，其中用 k 来表示在神经网络隐层的节点数，输入层中相应输入特征向量的维数用 n 来表示； $\{x_{ij} \mid i=1, 2, \dots, N; j=1, 2, \dots, M\}$ 指的是在空间粒子的相应位置向量所对应的分量， i 表示粒子在空间的位置， j 表示相应维数分量； $(x_{i1}, x_{i2}, \dots, x_{iM})$ 与 BP 神经网络核函数中的参数进行对应； $x_{\max}=1, x_{\min}=0, 0 \leq x_{ij} \leq 10$ ； $\{v_{ij} \mid i=1, 2, \dots, N; j=1, 2, \dots, M\}$ 指的是在空间粒子的相应速度向量所对应的分量， i 表示粒子在空间的速度， j 表示相应维数分量。

(3) 确定粒子的适应度函数

适应度函数决定了粒子在空间中相应位置的优劣度量，如果适应度函数的值越小，那么粒子的位置就越不好，具体定义为

$$\text{FitnessFunc}_i = \frac{1}{1 + D(x_{i1}, x_{i2}, \dots, x_{iM})} \quad (12-22)$$

式 (12-22) 代表的是在粒子群中第 i 个粒子对应的适应度函数值, 其中当核函数的参数是 $(x_{i1}, x_{i2}, \dots, x_{iM})$ 的时候, 此时神经网络的平均误差平方和就是 $D(x_{i1}, x_{i2}, \dots, x_{iM})$ 。

12.4.2 PSO 优化算法流程

下面主要介绍如何利用 PSO 去优化 BP 神经网络核函数参数的具体步骤流程:

1) 通过上述式 (12-20) 和式 (12-21) 来对粒子群进行相应的初始化;

2) 再根据上述式 (12-22) 对粒子群中粒子的适应度函数值进行有效的计算;

3) 通过粒子自身的适应度的函数与历史最佳位置 P_i 的适应度的函数比较, 判断优劣, 然后更新 P_g ;

4) 根据上述式 (12-17) 和式 (12-18) 来对粒子群中的粒子相应的速度位置进行有效的计算, 通过式 (12-21) 来检验获取粒子的速度和位置是否有效; 如果此时粒子群的平均适应度变化率比较小或者循环的次数已经到了最大值, 则就能够输出相应有效的核函数参数 P_g , 如果不满足上述条件就转到步骤 2 继续循环处理。

图 12-4 所示为 PSO 优化 BP 神经网络算法的基本流程图。

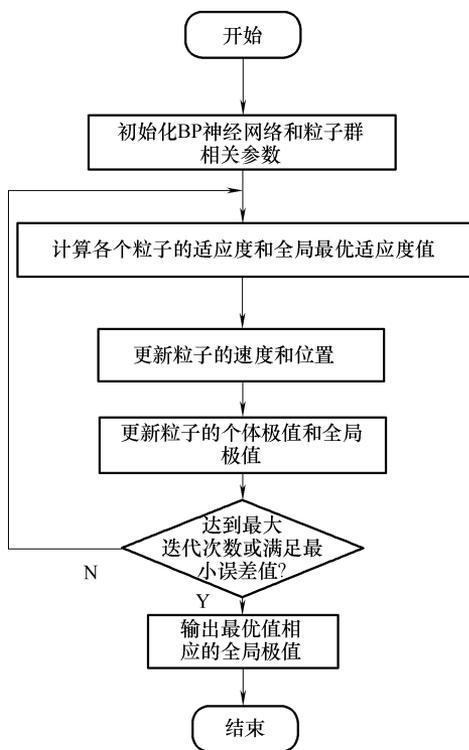


图 12-4 PSO 优化 BP 神经网络算法的主要流程图

12.5 仿真结果分析和识别系统实现

12.5.1 仿真结果分析

本实验主要是基于 MATLAB7.0 编程来实现, 在该仿真实验中主要利用了现有的目标识别系统的目标特征库, 该特征库主要包括直升机、坦克、越野车等

目标图像数据库，里面包含了目标的训练集和测试集数据，在训练集中主要包含红外和可见光图像数据，通过对它们不同角度的图像数据采集，训练中每类有 144 张样本图片，本次仿真实验主要针对上述三类目标进行仿真，一共有 432 张样本图片，每个类别目标有 50 个，图 12-5 所示是通过识别系统外部传感器获得直升机、坦克和越野车的照片信息。测试目标图片主要来自测试集中的图片，这里每组用 20 张来做测试。



图 12-5 目标图像

通过实验仿真图（见图 12-6、图 12-7）可以看出，没有经过优化的 BP 神经网络其耗时比较长，主要是网络的学习时间太长，而且容易陷入局部最优，导致最终的识别结果不是很理想，识别效果差。根据表 12-1、表 12-2 数据的比较可以发现利用 PSO 优化后的 BP 神经网络的识别率明显地提高了，所以本实验

仿真能够有效地验证该优化算法的有效性，同时能够有效地提高基于多传感器多特征融合目标识别的识别率。

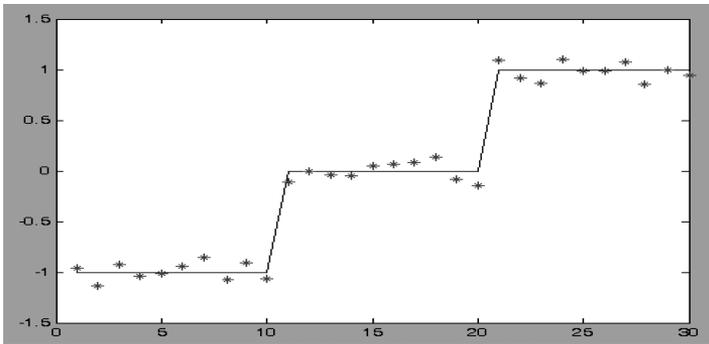


图 12-6 基于 PSO 优化 BP 神经网络特征信息融合的目标识别结果

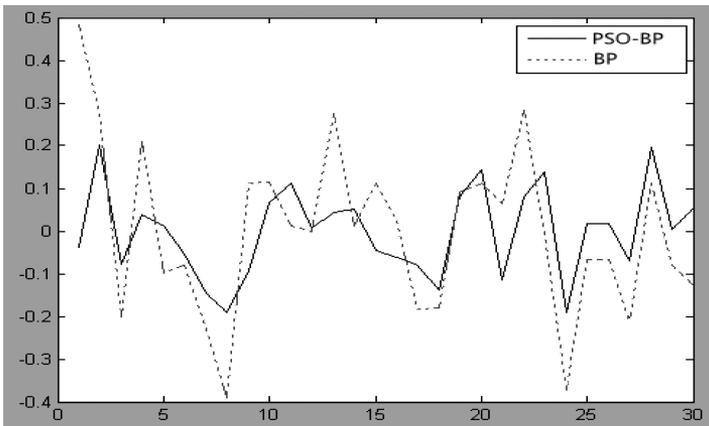


图 12-7 非优化 BP 与优化 BP 神经网络多特征目标识别的误差

表 12-1 基于 BP 神经网络特征信息融合的目标识别结果

目标项目	样本数	目标数	正确数	识别率	耗时
轿车	144	20	14	70%	45.5s
坦克	144	20	12	60%	37.8s
直升机	144	20	16	80%	43.4s

表 12-2 基于 PSO 优化 BP 神经网络多特征目标识别的结果

项目目标	样本数	目标数	正确数	识别率	总耗时
轿车	144	20	18	90%	35.5s
坦克	144	20	16	80%	33.6s
直升机	144	20	18	90%	37.1s

12.5.2 识别系统实现

在该识别系统中，后期数据处理到识别实现主要依靠特征数据库、目标训练和目标识别。

1. 特征数据库

特征数据库主要用来存储传感器采集到的目标特征数据，为训练模型做数据准备。在本系统中主要运用目前比较成熟的数据库 SQL Server 2005 来建立相应的特征数据库，它的优点就是能够将各种复杂的数据用二维表格来直观地表示出来，而且它的存储结构最为突出的优点就是各种复杂的数据结构之间的关系能够以二维表格的形式直观表示。图 12-8 为系统中样本特征数据。

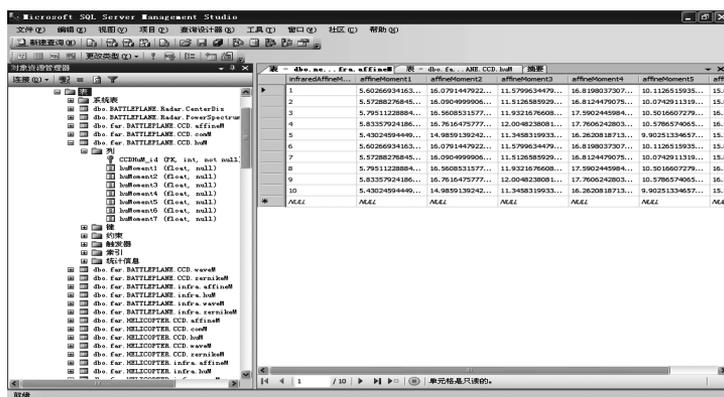


图 12-8 基于 SQL 的样本特征数据

2. 目标特征数据模型训练

在图 12-9 和图 12-10 中，先对样本图像进行选择。



图 12-9 样本文件选择



图 12-10 样图显示

选择好样本后，根据图 12-11 和图 12-12 来对目标进行特征提取。提取后的目标特征数据会自动存入到特征数据库中，存储完成如图 12-13 所示。

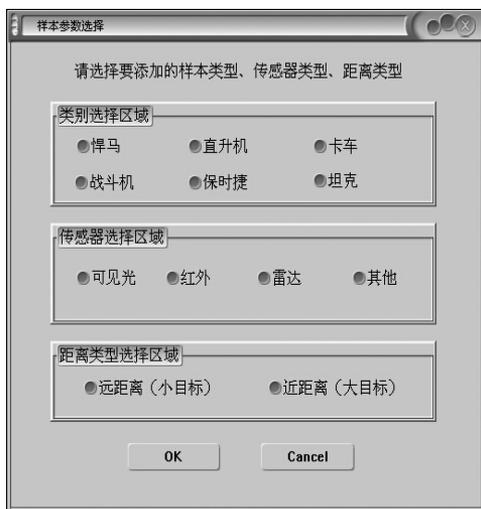


图 12-11 目标参数选择

通过上面对样本特征的提取和存储，下面就是系统对特征数据进行训练，建立相应的目标特征模型。图 12-14 是识别模型训练的界面。

3. 目标识别

目标识别方法主要是采用基于 PSO 优化改进的 BP 神经网络目标识别方法，通过仿真实验我们证明该方法的有效性，而现在我们将其应用到具体的目标识别系统中，来证明它的可用性和可行性。在图 12-15 中，主要介绍了识别系统



图 12-12 多文件选择界面



图 12-13 存储完成

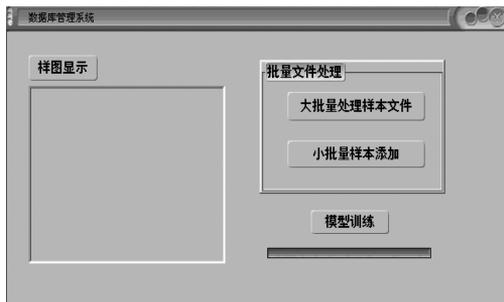


图 12-14 模型训练

的操作流程，通过外界人为对系统进行操作，该系统中利用外部传感器对目标进行图像采集，然后通过对图像的预处理，利用特征提取算法对目标特征进行特征提取，然后再利用特征融合算法对提取到的目标特征进行多特征信息融合，最后再利用识别算法对目标进行识别处理。在图 12-16 中，针对战斗机进行了目标识别试验，通过外部传感器捕获到的战斗机图像信息，单击界面中的目标识别按钮，在很短的时间内就能在“识别结果”中显示出“战斗机”这样的识别结果文字，所以在理论到实践，再到具体应用，都证明了该方法的可行和可靠。

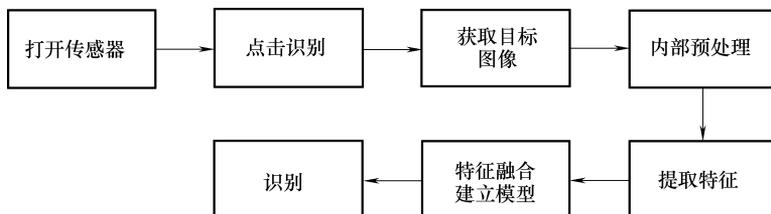


图 12-15 识别系统中内部的运行流程



图 12-16 识别系统结果

12.6 本章小结

本章先详细介绍了理论知识，然后在此基础上对 PSO 和 BP 神经网络优缺点

进行了详细分析，然后利用互补优势将其充分结合起来对目标进行识别。本章先对 PSO 算法和 BP 神经网络的算法原理进行分析，并得到相应的流程图，然后提出一种改进的 BP 神经网络优化方法，再利用 PSO 对改进的 BP 神经网络进行优化，通过利用现有的直升机、坦克和越野车等目标图像数据对它们进行分类识别。最后通过仿真结果得出了基于 PSO 优化改进的 BP 神经网络相应的误差图和识别率，通过对比它们的误差和识别率来进行分析。

模糊支持向量机理论与编程实现

13.1 模糊数学理论

13.1.1 模糊数学的发展历程

模糊性可以定义为现实中的不确定的没有明确标准的现象。比如“很好”与“不太好”、“健康”与“不健康”之间没有明确的界定标准，无法知道满足这些概念需要具备什么资格，从而便有了模糊性。模糊数学是一种定量处理方法，但是不同于其他学科，它是专门为了解决模糊现象而产生的数学方法。

L. A. Zadeh 教授是美国著名的控制论专家，1965 年他发表了一篇论文《Fuzzy Sets》，至此模糊数学诞生了。

人类具有认知的能力，可以通过五官的感知从外界获取信息，通过大脑对这些信息进行整合与分析。在对这些信息进行处理的过程中，信息中包含了大量的模糊信息。但是人类却能够很好地通过这些信息进行推理、抽象，最终做出必要的正确的决策。这是人类特有的能力，看似简单的能力在计算机中实现起来却不是一件容易的事。计算机具有人类无法比拟的运算能力，对信息的记忆能力也要远超常人。虽然计算机具备这么多让人望尘莫及的超级能力，但是当计算机面对具有模糊性质的问题时却依然是束手无策。对于计算机不能如同人类的大脑一样思考问题，基于对计算机与大规模系统之间矛盾的研究，L. A. Zadeh 教授对这一问题进行了详尽的分析。要解释计算机不能处理具有模糊性质的问题，首先要对康托尔集合论（Cantor's Set）进行了解。康托尔集合论指出，论域中的任何一个元素与某个集合 A 之间的关系只有两种，属于 A 或者不属于 A ，不可能出现第三种情况。所以，康托尔集合只能描述“是”与“否”的概念，对于具有不确定性的问题就无法进行描述了。类似的，正是由于计算机只有处理康托尔集合的逻辑造成了它无法处理模糊问题。了解了原因之后，要解决这一问题就只剩下寻找解决的办法了，而模糊集合论正是为解决此类问题而诞生的。经过多年的发展，模糊集合论现在已经形成了一套完备的模糊数

学体系。

虽然发展的时间很短，但是模糊数学却在很多领域得到了非常广泛的应用，它所涉及的领域几乎包含了迄今为止人类所有的活动。因此模糊数学理论正逐渐成为与人们息息相关的科学技术。很多具有模糊性的问题都需要使用模糊数学进行处理，尤其是面对庞大的系统决策时更是如此。

13.1.2 模糊集

模糊数学是描述模糊现象的数学。而模糊集是模糊数学的理论基础。对于我们熟悉的普通集论要求：论域 U 中每个元 u ，对于子集 $A \subseteq U$ 来说， u 与 A 只存在 $u \in A$ 或者 $u \notin A$ 两种关系，不会出现模棱两可的现象。因而，子集 A 可用 0 和 1 两个数字来刻画。设论域 U 为自然数集，那么上述关系可以用特征函数法表示为

$$\begin{aligned} C_A: U \rightarrow \{0,1\} \\ u \rightarrow C_A(u) \end{aligned} \quad (13-1)$$

$$\text{式中, } C_A(u) = \begin{cases} 0 & u \in A \\ 1 & u \notin A \end{cases}$$

可见，给定了论域 U 上的一个子集，就等于给定了特征函数，反之亦然。因此，一个特征函数对应着集合里的一个值。

传统的数学理论中特征函数的取值范围是 $\{0, 1\}$ ，而 L. A. Zadeh 教授则将这个范围进行了拓展，将取值范围由原来确定的两个数值拓展到了一个闭区间 $[0, 1]$ 。这也是模糊集概念的来源。

设在论域 U 上给定了一个映射：

$$\begin{aligned} A: U \rightarrow [0,1] \\ u \mapsto A(u) \end{aligned} \quad (13-2)$$

则称 A 为论域 U 上的模糊集 (Fuzzy Set)， $A(u)$ 称为 A 的隶属度函数，用来求得 u 对 A 的隶属度。

隶属度可以理解为是用来表示一个样本与模糊集合之间相关性程度的值。当 $A(u) = 1$ 时，表示样本 u 完全属于模糊集合 A ；反之，当 $A(u) = 0$ 时，表示样本 u 完全不属于模糊集合 A 。 $A(u)$ 的值与样本 u 属于模糊集合 A 的程度成正比。

13.1.3 几种常见的模糊隶属度函数

1. 对某个值靠近程度的表示函数

实际应用中常用一些特殊的隶属度函数用来表示当前 u 值对某个值的靠近程度，比如高斯函数：

$$A(u) = \exp\left[-\frac{(x-a)^2}{2\sigma^2}\right] \quad (13-3)$$

式中， a 和 σ 为参数。函数图形如图 13-1 所示。

从 Gaussian 型函数的图形中我们可以看出， u 越是靠近 a ，其隶属度越接近 1，当 $u = a$ 时隶属度为 1，因此我们可以用 Gaussian 型隶属度函数来表示“靠近 a 的程度”。

另外还经常用三角形函数定义如下：

$$A(u) = \begin{cases} 0 & u \leq a \\ (u-a)/(b-a) & a < u \leq b \\ (c-u)/(c-b) & b < u \leq c \\ 0 & u > c \end{cases} \quad (13-4)$$

式中， a 、 b 、 c 为参数，三角形函数图形如图 13-2 所示。

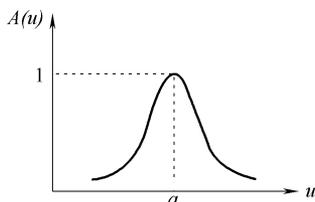


图 13-1 高斯型函数图形

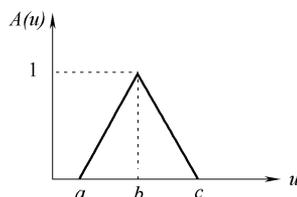


图 13-2 三角形隶属度函数图形

由图 13-2 可以看出，三角形函数与 Gaussian 型函数所表示的含义基本一致，都是对某个值的靠近程度，三角形靠近程度为线性递增和递减，Gaussian 型靠近程度为非线性递增和递减。

除此之外还有梯形函数等都可以用来表示同样的含义。

2. 对某个值远离程度的表示函数

用于表示 $u > a$ 的程度的隶属度函数，常用函数如下：

$$A(u) = \begin{cases} 0 & 0 \leq u \leq a \\ \left[1 + \left(\frac{u-a}{b}\right)^{-2}\right]^{-1} & a < u \leq 2a \end{cases} \quad (13-5)$$

函数图形如图 13-3 所示。

另外的一个常用的函数可以表示为

$$A(u) = \begin{cases} 1 & 0 \leq u \leq a \\ \left[1 + \left(\frac{u-a}{b}\right)^2\right]^{-1} & a < u \leq 2a \end{cases} \quad (13-6)$$

函数图形如图 13-4 所示。

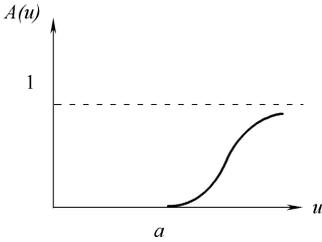


图 13-3 表示 $u > a$ 程度的隶属度函数

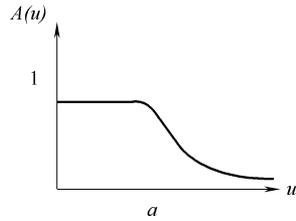


图 13-4 表示远离 a 的程度的隶属度函数

13.2 支持向量机理论

机器学习领域凸现出了许多优秀的理论与算法，但是支持向量机（SVM）理论绝对是其中的集大成者。该理论由 Vapnik 和他的合作伙伴共同提出，一经提出就受到了世界范围的关注，随后便有了突飞猛进的发展势头，不但成为机器学习领域的标准工具之一，也成为众多其他领域经常采用的算法。支持向量机理论之所以如此优秀，是因为它建立在了众多优秀的理论的基础之上，比如凸二次规划、统计学理论、核函数理论等一批完备的数学理论。

近几年涌现出的大量理论研究成果，更为其应用研究奠定了坚实的基础。然而对于支持向量机来说，最大的问题在于解二次规划问题时的运行速度较慢。为了对此问题进行改善，众多国内外科学家和研究人员对其算法做出了研究与改进，从而形成了多种基于传统支持向量机的变形算法，而且这些经过变形的算法往往在某一领域具有较好的运行速度。下面具体列出了几种经常用到的变形算法，它们都是比较典型的变形算法： ν -SVM、Fuzzy SVM、GSVM（Generalized SVM）、LS-SVM（Least-Square SVM）、One-Class SVM、RSVM（Reduced SVM）等算法。另一方面，许多国内外科学家和研究学者从学习策略方面入手，针对实际情况中不同的应用问题结合各种支持向量机方法各自的优点，提出了多种新型的基于组合形式的支持向量机算法。下面我们对支持向量机的理论知识简单做一些介绍。

13.2.1 经验风险

机器是不具有思考能力的，因此机器的学习过程是需要人为地制定学习规则的，也就是对学习内容、学习方式以及学习的停止准则进行规定。其中学习停止准则最为重要，因为如果机器在还没有完全获得必要的知识之前就停止学习，那么它的识别能力也就可想而知了；相反，如果出现了过学习现象，那么机器的泛化能力也会下降，这也是我们不希望看到的。所以对学习准则的制定需要进行多方面的考虑，而经验风险最小化的原则是较为实用的停止准则之一，

下面我们对这一停止原则进行解释。

假设有 l 个观测样本 $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$ ，它们之间是相互独立的关系，但是遵循一样的分布规律， y 和 x 之间的关系是满足一个未知的联合概率密度函数 $F(x, y)$ 。预测函数集 $\{f(x, w)\}$ 中存在可以使得期望风险 $R(w)$ 最小的一个预测函数 $f(x, w_0)$ ，称之为最优函数。对期望风险 $R(w)$ 的定义为

$$R(w) = \int L(y, f(x, w)) dF(x, y) \quad (13-7)$$

式中，我们将 $\{f(x, w)\}$ 称之为预测函数集合， w 为函数的广义定义参数。 $\{f(x, w)\}$ 可以表示任何函数集。 $L(y, f(x, w))$ 为由于用 $f(x, w)$ 对 y 进行预测而造成的损失，针对学习的情况不同，损失函数的形式也不相同。

机器学习最常应用于以下三类：分类问题、回归估计问题和概率密度估计问题。对于分类问题，输出 y 的形式为 $y = \{0, 1\}$ 或 $y = \{1, 1\}$ ，该类问题的损失函数可以定义为以下形式：

$$L(y, f(x, w)) = \begin{cases} 0 & y = f(x, w) \\ 1 & y \neq f(x, w) \end{cases} \quad (13-8)$$

针对目标识别问题，我们期望通过对现有样本数据学习的过程，找到一个使得错分概率最小的分类函数 $f(x, w_0)$ ，并且学习过程是在 $F(x, y)$ 未知的情形下进行的。如果分类器的输出与期望值不符，那么我们称该次分类为错分。

在函数回归估计问题中， y 的值是一个连续变量，损失函数定义为

$$L(y, f(x, w)) = (y - f(x, w))^2 \quad (13-9)$$

回归函数就是在损失函数下能够使得风险泛函式 (13-7) 取值最小的函数。

概率密度估计问题不同于模式识别，模式识别具有期望输出，而概率密度估计没有这样一个期望值来做对比，它的目的是通过对现有样本的学习来大概估计 x 的概率密度，其中被估计的密度函数为 $p(x, w)$ 。此类问题的损失函数也不同于模式识别，具体形式定义如下：

$$L(y, f(x, w)) = -\log p(x, w) \quad (13-10)$$

虽然各问题的应用方向不一样，但通过对以上问题的描述可以看出，它们都是以期望风险最小化为学习目的的。但是由于受到各种条件限制，比如样本数量较少、概率密度 $F(x, y)$ 未知，所以式 (13-7) 表示的期望风险无法估计。针对此类问题，传统的学习方法是通过计算样本的误差来定义经验风险的：

$$R_{\text{emp}}(w) = \frac{1}{l} \sum_{i=1}^l L(y_i, f(x_i, w)) \quad (13-11)$$

我们将这种利用样本误差来定义经验风险的方法称之为经验风险最小化 (Empirical Risk Minimization, ERM) 原则。

统计学习的任务就是要找到合适的方法最小化 $R_{emp}(w)$ 来替代式 (13-7)。针对不同的问题, 具有不同的损失函数, 其对应的经验风险的含义就会不同。比如, 模式识别问题中的经验风险是错分率, 而回归问题中的经验风险却是平方训练误差。许多经典算法都采用了基于经验风险最小化的原则, 比如最小二乘法、最大似然估计法这些算法都是采用了经验风险最小化的原则, 另外还有神经网络学习方法也是采用了同样的学习原则。

13.2.2 结构风险最小化

经验风险最小化原则是经过实践验证的有效的学习停止规则, 并被应用到了众多的识别算法中, 但是它所应用的情况也是具有局限性的。比如观测样本必须最够多, 在大量的数据中寻找某个规律并总结经验, 使得在以后的预测过程中犯错的可能性达到最小, 也就是先验知识和经验足够丰富才行。对于小样本数据问题, 往往不能通过现有的数据从中得到较为有效的先验知识, 从而造成置信范围过大。神经网络等方法就是采用经验风险最小的原则, 然而神经网络的网络结构是具有不确定性的, 针对不同问题采用的网络结构不同, 最后得出的效果差异也较大。因此在网络的设计过程中对操作人员的技巧具有很强的依赖性。

针对经验风险最小化的问题, 统计学习理论提出了一种新的学习停止规则, 被称为结构风险最小化准则。结构风险最小化准则的基本思想可以通过图 13-5 进行说明。通过图 13-5 我们可以看出, 所谓结构风险最小是指首先通过函数集中不同子集 VC 维 (Vapnik-Chervonenkis dimension) 的大小对这些子集进行排序, 之后在子集的内部搜寻最小经验风险, 然后在子集之间添加对置信范围的考虑, 最后取一个折中值。

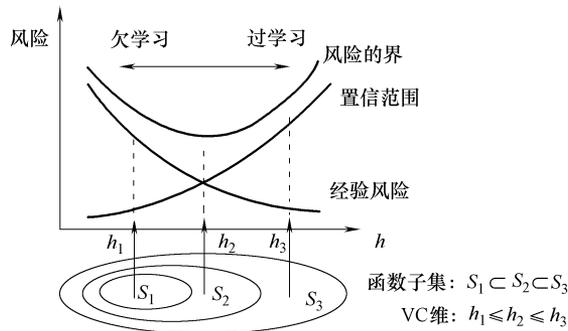


图 13-5 结构风险最小化准则

通过上述描述可知，经验风险和置信范围分别取决于不同的对象，置信范围是取决于函数集的，而经验风险往往是由某个具体的函数来决定的。统计学习理论对结构风险最小化原则进行了详细的论证。首先对结构风险最小化准则下实际风险能够收敛的性质进行了证明，还对模型的选择给出了严密的论证过程，不仅如此，它还对函数子集结构的合理性提出了相应的标准。

基于对函数集重要性的考虑，统计学习理论对其各种类型进行了细致的研究。希望能够通过对它们详细的研究来更好地了解泛化界限，也可以简单地理解为是寻找它与经验风险和实际风险之间的内在关联。

不失一般性，我们这里对两分类问题进行讨论，可以得到如下的结论：经验风险 $R_{\text{emp}}(w)$ 和实际风险 $R(w)$ 之间有如下关系：

$$R(w) \leq R_{\text{emp}}(w) + \sqrt{\frac{h(\ln(2l/h+1) - \ln(\eta/4))}{l}} \quad (13-12)$$

式中， h 为 VC 维， l 是样本个数的具体数值， η 取值范围为 $0 \leq \eta \leq 1$ 。根据公式可以看出，小于等于号右边由两项构成，也就是实际风险 $R(w)$ 可以分为两部分：一部分是我们所熟知的经验风险 $R_{\text{emp}}(w)$ ，另一部分我们可以称其为置信界限。并且公式中的置信界限是由分类器的 VC 维 h 和训练样本数 l 共同确定的，它是对复杂结构所带来的风险的一个度量。所以式 (13-12) 也可以简单地表示为

$$R(w) \leq R_{\text{emp}}(w) + \Phi(h/l) \quad (13-13)$$

由式 (13-13) 右侧第二项可以看出， $\Phi(h/l)$ 与 h 成正比。过学习的现象就是由于在训练样本有限的情况下，学习机器的复杂性过高、VC 维较大，此时正比于 VC 维的置信界限自然也就较大。通过式 (13-13)，置信界限大意味着真实风险与经验风险之间的差别大，从而出现了过学习的现象。所以，我们可以通过调整分类器的泛化能力来解决这个问题，这样就可以构造适于小样本的分类器。对于有界非负函数， $0 \leq L(z, w) \leq B$ ， $w \in \Lambda$ ，以大于或等于 $1 - \eta$ 的概率满足下面的公式：

$$R(w_l) \leq R_{\text{emp}}(w) + \frac{B\varepsilon}{2} \left[1 + \sqrt{1 + \frac{4R_{\text{emp}}(w_l)}{B\varepsilon}} \right] \quad (13-14)$$

同样的考虑无界限集合，我们可以得到分类器的泛化能力：

$$R(w_l) \leq \frac{R_{\text{emp}}(w_l)}{1 - \alpha(p)\tau\sqrt{\varepsilon}} \quad (13-15)$$

式中， $\alpha(p) = p \sqrt{\frac{1}{2} \left(\frac{p-1}{p-2} \right)^{p-1}}$ ， $\varepsilon = 2 \frac{\ln N - \ln \eta}{l}$

通过以上描述可知，经验风险越小，风险的上界也在减少，最小化经验风险或者最小化式 (13-14) 小于等于号右边的第二项，都可以调控最小化过程。

显然，后者适合小样本情况。

13.2.3 最优超平面

假设存在一个线性可分的训练集 $S = \{(x_j, y_j)\}_{j=1}^l$ ，线性可分是指训练集中不同类别的样本点能够被一个线性超平面正确分开。我们可以将这个分类超平面用公式描述为

$$w \cdot x + b = 0 \quad (13-16)$$

式中， w 是超平面的法线方向， $\frac{w}{\|w\|}$ 为单位法线方向， $\|w\|$ 为 w 的模， $w \cdot x$ 为 w 与 x 的内积，这里 $b \in \mathbf{R}$ ， $w \in \mathbf{R}^n$ ， $x \in \mathbf{R}^n$ 。

分类结果可以表示为

$$\begin{cases} w \cdot x_i + b \leq 0, & y_i = -1 \\ w \cdot x_i + b \geq 0, & y_i = +1 \end{cases} \quad (13-17)$$

我们可以假设空间：

$$f(w, b) = \text{sgn}(w \cdot x + b) \quad (13-18)$$

对 (w, b) 进行如下约束：

$$\min |w \cdot x_i + b| = 1 \quad (13-19)$$

我们的目的是为了完全正确地进行分类，显然能够满足分类目的的分类超平面未必只有一个。如图 13-6 所示，该分类超平面也可以完全正确地对两类样本进行正确的分类。尽管能够正确分类的超平面不唯一，但是最优分类超平面是唯一的。所谓最优分类超平面是指能够满足距离该超平面最近的两类区域边界处的训练点之间的距离最大（见图 13-6），我们称间隔为 $\frac{2}{\|w\|}$ 的分类超平面为最优分类超平面。支持向量就是指这些距离最优超平面最近的处于类别区域边界上的样本点，正是利用这些支持向量才唯一地确定了最优分类超平面。

图 13-6 和图 13-7 对比描述了最优分类超平面与普通分类超平面的结构比较，通过两幅图形的对比可以很直观地看出两类分类超平面的差异。图 13-6 中 H 为最优分类超平面， H_1 和 H_2 分别为两个类别区域的边界超平面，两者经过各自类别区域中位于边界上的样本点并与 H 平行，称 H_1 和 H_2 之间的距离为分类间隔。这里的最优是指分类间隔能够达到最大值 $\frac{2}{\|w\|}$ 。由约束条件式（13-19）可知：

$$\begin{aligned} H: w \cdot x + b &= 0 \\ H_1: w \cdot x_i + b &\geq 1, \quad y_i = +1 \\ H_2: w \cdot x_i + b &\leq -1, \quad y_i = -1 \end{aligned} \quad (13-20)$$

两种情况可以简写成：

$$y_i(w \cdot x_i + b) \geq 1, i = 1, 2, \dots, l \quad (13-21)$$

式中， H_1 和 H_2 到 H 的距离为 $\frac{1}{\|w\|}$ ，分类间隔为 $\frac{2}{\|w\|}$ 。

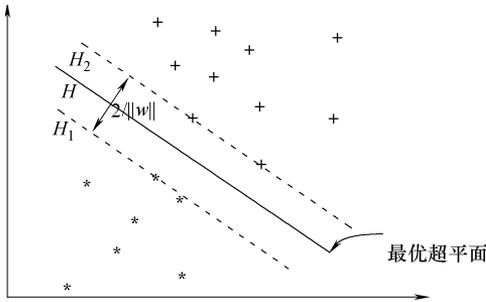


图 13-6 最优分类超平面

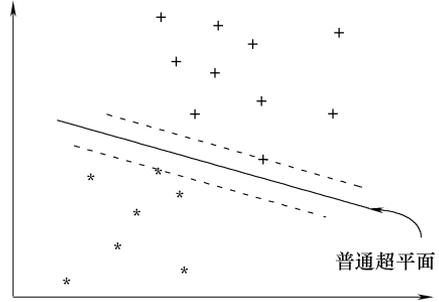


图 13-7 普通分类超平面

13.2.4 VC 维数

VC 维 (Vapnik-Chervonenkis Dimension) 是统计学习理论中一个重要的概念。它是一个与几何概念的维没有关系的纯粹的组合概念。它在统计学习理论中扮演着一个中心的角色。VC 维是一个用来度量具有分类作用的函数集的具体分类能力高低的数值。粗略地说，为了可靠的学习，一个类所需要的样本的数量正比于那个类的 VC 维。因此对 VC 维的估计需要首先关注。

由数学知识中的排列组合理论可以知道，如果存在 h 个一样的相互独立的小球，现在要将这 h 个小球随机地放入两个不同的箱子 A 和 B 中，那么可以求出放入的方式有 2^h 种。可以理解为每个小球有两种选择，选择放入 A 箱子或者 B 箱子，那么 h 个小球共有 $2 \times 2 \times \dots \times 2 = 2^h$ 种情况。现在将 A 和 B 两个箱子看作两个不同的类别，将 h 个小球看作是观测样本，也就是说样本的划分情况共有 2^h 种。那么我们规定如果一个函数集，它最多能够实现将 h 个样本以任意的组合进行分类，那么称此函数集的 VC 维为 h 。如果此函数集能够实现将无穷多个样本以任意的组合进行分类，那么该函数集的 VC 维为 ∞ 。

虽然 VC 维的概念已经清晰了，但是对 VC 维的计算却至今没有通用的理论，所以怎样得到一个通用的计算 VC 维的公式是当前有待解决的问题。虽然计算公式尚未存在，但是对于某些特殊的函数集我们是知道它们的 VC 维的，如一个线性分类器，它的 VC 维是样本的维度加 1； $f(x, a) = \sin(ax)$ 的 VC 维则为 ∞ 。

13.2.5 核函数

对于非线性问题，支持向量机通过一个映射过程将原始空间中的问题映射到更高维度的空间中，使之在高维度空间中成为一个线性问题，并在高维度空间中进行求解。但是，由于维数急剧增长，通常情况下在特征空间计算最优分类面往往很难实现。但是，在求解支持向量机的对偶问题过程中，仅在高维空间中的点积运算，如果能找到一个合适的函数 $K(x_i, x_j)$ ，使得

$$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j) \quad (13-22)$$

这样，在高维特征空间中的内积运算就可以在原空间中利用函数来实现，避免了维数转换。函数 $K(x_j, x_i)$ 就称为核函数 (Kernel Function)。

支持向量机的核函数种类繁多，不同的核函数可以满足不同算法的需要，但是基本所有的核函数都会满足 Mercer 条件：

$K(\mathbf{x}, \mathbf{x}')$ 为连续对称核，其中 \mathbf{x} 和 \mathbf{x}' 的取值范围为区间 $[a, b]$ 。那么该对称核 $K(\mathbf{x}, \mathbf{x}')$ 可以展开为下式中的级数形式：

$$K(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}') \quad (13-23)$$

式中， $\varphi_i(\mathbf{x})$ 为特征函数展开式， λ_i 称为特征值。如果特征值全都是正数，那么称核 $K(\mathbf{x}, \mathbf{x}')$ 是正定核。

式 (13-23) 满足展开式的合理性和绝对一致收敛性的充要条件为

$$\int_b^a \int_b^a K(\mathbf{x}, \mathbf{x}') \Psi(\mathbf{x}) \Psi(\mathbf{x}') dx dx' \geq 0 \quad (13-24)$$

对于所有满足 $\int_b^a \Psi^2(\mathbf{x}) dx < \infty$ 的 $\Psi(\mathbf{x})$ 成立。

通过 Mercer 定理我们只能得知该候选核函数是否可以表示高维空间的内积以及是否可以应用于支持向量机。但是它并没有告诉我们如何去构造函数 $\varphi_i(\mathbf{x})$ ，因此我们需要自己构造，目前常用的核函数有：

(1) 多项式核函数 (Polynomial Kernel)：

$$K(x, x_i) = [(x \cdot x_i) + 1]^q \quad (13-25)$$

式中， q 为阶数。多项式核函数的优点较多，主要的优点有：全局性能较好、外推能力较强，其中阶数 q 与外推能力成反比。

(2) 径向基核函数 (RBF Kernel)

$$K(x, x_i) = \exp\left(-\frac{|x - x_i|^2}{2\sigma^2}\right) \quad (13-26)$$

径向基核函数具有很好的局部性，也是运用最为广泛的核函数之一。其中公式中的参数 σ 对分类器影响较大，其值将直接对分类器的分类能力起到关键性的影响。因此，选取合适的参数在实际应用中是非常重要的环节。

(3) S 形核函数

$$K(x, x_i) = \tanh(v(x \cdot x_i) + c) \tag{13-27}$$

基于 S 形核函数的支持向量机可以看作一个含有隐含层的多层感知机。而且可以由算法自动确定隐含层内部的节点数，不需要人为设置，所以这是 S 形核函数的一个优点。另外，具有全局极小值并可以避免局部极小值也是 S 形核函数的一个重要优点。

13.3 模糊支持向量机理论

模糊支持向量机 (FSVM) 与传统支持向量机不同之处在于引入了模糊数学的理论。主要思想是通过对样本数据引入隶属度来实现。其中，各样本点隶属度的数值由隶属度函数来进行确定，而且不同的隶属度函数对于模糊支持向量机的分类结果的影响也是不同的。由于传统支持向量机是平等地对待所有的样本点的，对于位于分类边界处的点相当敏感，所以当训练数据中存在野值或者是噪声的时候，将对分类超平面产生很大的影响。所以我们引入隶属度函数来克服这个问题。

模糊支持向量机的概念主要有两种，下面将对这两种模糊支持向量机进行介绍。

13.3.1 第一种模糊支持向量机

第一种模糊支持向量机的概念是由日本学者 Takuga 和 Shigeo Abe 于 2001 年提出的。这种方法主要针对在多分类问题中“一对一”和“一对多”支持向量机存在分类盲区而提出的。

图 13-8 对第一类模糊支持向量机的概念进行了直观的说明。训练好的分类函数可能会出现对某个待分类样本无法进行分类的问题。为了避免出现图中不可分区域的情况引入了模糊隶属度函数来构造模糊支持向量机进行分类。具体过程是，首先对样本 x 通过采用某种计算隶属度的方法来计算该样本 x 与第 i 类

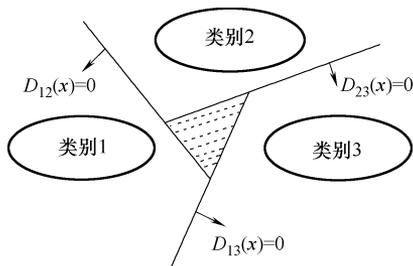


图 13-8 “一对一”模糊支持向量机多类分类器中的不可分区域

之间的隶属度值，然后将其归类为隶属度值最大的类别中去。

13.3.2 第二种模糊支持向量机

第二种模糊支持向量机的概念是由 Chun-Fu Lin、Sheng-De Wang、Yi-Hung Liu 和 Han-Pang Huang 等人于 2002 年提出。在传统的支持向量机的优化问题中，为了减少对样本点错分的情况，提出了惩罚参数的概念。惩罚参数 C 的具体作用就是当发现分类器出现错分的情况时给予一个惩罚，惩罚力度由惩罚参数 C 来控制，而且惩罚力度正比于惩罚参数的值。但是不管 C 的值是多少，一旦 C 的值确定了，那么在支持向量机的整个训练过程中这个值都不再改变，所以，传统的支持向量机对所有的训练样本点其实是同等对待的，没有任何区别。这种情况导致了支持向量机对一些特别情形的过分敏感，而且这种敏感是我们所不希望的。例如我们前面讲到的距离类别中心位置很远的孤立的样本点或者是噪声，这样就会出现过学习现象，使得分类器泛化能力下降。

既然已经找到了问题出现的原因，那么要解决该问题就只需要找到正确的解决方法就可以了。既然噪声点和距离类别中心较远的孤立样本点对于分类造成了不利的影 响，那么我们只要对这些特殊的样本点采取一些措施来减少它们对分类的贡献度，应该就能得到更为理想的分类超平面了。因此我们针对样本点的性质分别赋予不同的隶属度，原则就是，贡献越大，隶属度越高，反之隶属度就低。采取这种措施就可以尽量减小噪声点和孤立样本点的隶属度数值，从而有效地减少噪声和孤立点对训练过程中分类性能造成的不利影响。

针对两分类问题，首先对训练样本集定义一个隶属度函数。之后根据隶属度函数，将所有的训练样本进行模糊化处理，得到每个训练样本 x_i 的隶属度 s_i 。最后将训练样本集改写为模糊训练样本集，将隶属度信息添加到样本点数据中。

假设模糊训练集合为

$$S = \{(x_1, y_1, s_1), (x_2, y_2, s_2), \dots, (x_l, y_l, s_l)\} \quad (13-28)$$

式中， $x_i \in \mathbf{R}^n$ ， $y_i \in \{-1, +1\}$ ， $s_i \in (0, 1]$ 。

s_i 表示训练样本 x_i 与其对应的输出 y_i 之间的隶属度数值。隶属度的值表示训练样本 x_i 对分类结果的重要程度，而参数 ξ_i 为支持向量机的松弛变量，表示容错程度。因此， $s_i \xi_i$ 可以用来衡量对于分类重要性不同的变量的容错程度。

对于前面所描述的训练集，通过解如下的优化问题可以得到我们所期望的最优分类超平面：

$$\left. \begin{aligned} & \min_{w \in H, b \in R, \xi \in R^l} \frac{1}{2} \|w\|^2 + C \sum_{j=1}^l s_j \xi_j \\ \text{s. t. } & y_j(W \cdot \Phi(x_j) + b) + \xi, j = 1, \dots, l \\ & \xi \geq 0, j = 1, \dots, l \end{aligned} \right\} \quad (13-29)$$

式中, $C > 0$ 是惩罚参数, s_j 是模糊隶属度, $\xi = (\xi_1, \dots, \xi_l)^T$ 。

求解式 (13-29) 二次规划的对偶规划, 构造拉格朗日函数:

$$L(w, b, \xi, \alpha, \beta) = \frac{1}{2} \|w\|^2 + c \sum_{j=1}^l s_j \xi_j - \sum_{j=1}^l \alpha_j (y_j (w \cdot x_j + b) + \xi_j - 1) - \sum_{j=1}^l \beta_j \xi_j \quad (13-30)$$

式中, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)^T$, $\beta = (\beta_1, \beta_2, \dots, \beta_l)^T$, $\alpha_j \geq 0$, $\beta_j \geq 0$ ($j = 1, 2, \dots, l$)。

根据对偶的定义, 对式 (13-30) 关于 w 、 b 和 ξ_j 求极小值, 求偏导如下:

$$\left. \begin{aligned} \frac{\partial L}{\partial w} &= w - \sum_{j=1}^l \alpha_j y_j x_j = 0 \\ \frac{\partial L}{\partial b} &= - \sum_{j=1}^l \alpha_j y_j x_j = 0 \\ \frac{\partial L}{\partial \xi_j} &= s_j c - \alpha_j - \beta_j = 0 \end{aligned} \right\} \quad (13-31)$$

将式 (13-31) 代入式 (13-30) 中, 得二次规划的对偶规划为

$$\left. \begin{aligned} & \min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j < x_i \cdot x_j > - \sum_{j=1}^l \alpha_j \\ \text{s. t. } & \sum_{j=1}^l y_j \alpha_j = 0 \\ & 0 \leq \alpha_j \leq s_j C, j = 1, \dots, l \end{aligned} \right\} \quad (13-32)$$

因此, 通过前面的推导我们得到了最优超平面问题的对偶形式, 根据对偶的性质, 通过求得对偶形式的最优解也就得到了我们所期望的最优分类超平面。可以看出式 (13-32) 为一个凸二次规划问题, 所以可以解得最优解 $\alpha^* = (\alpha_1^*, \dots, \alpha_l^*)^T$, 据此我们得到模糊最优分类函数公式如下所示:

$$f(x) = \text{sgn} \left(\sum_{j=1}^l \alpha_j^* y_j < x \cdot x_j > + b^* \right), x \in R^n \quad (13-33)$$

若存在 α^* 的正分量 α_j^* 使得 $\alpha_j^* \in (0, s_j C)$, 则 $b^* = y_i - \sum_{j=1}^l y_i \alpha_j^* (x_i \cdot x_j)$, 其中下标 $i \in \{i | 0 < \alpha_i^* < s_j C\}$ 。若 $\alpha_j^* > s_j C$, 为错分的样本点。这种模糊支持向量机与传统的支持向量机最大的不同在于, 由于 s_j 的出现, 使得模糊支持向量机中 α_j^* 所对应的支持向量与传统的支持向量机中对应的支持向量有可能发生变化, 也就是前后两者所对应的不再是同一类支持向量。

对于非线性问题，我们引入核函数 $K(x_i, x_j)$ ，那么分类问题可用如下的二次规划表示：

$$\left. \begin{aligned} \min_{\alpha} & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{j=1}^l \alpha_j \\ \text{s. t.} & \sum_{j=1}^l y_j \alpha_j = 0 \\ & 0 \leq \alpha_j \leq s_j C, j = 1, \dots, l \end{aligned} \right\} \quad (13-34)$$

通过式 (13-34) 可以看出，这是一个凸二次规划问题，对其求解可以得到 $\alpha^* = (\alpha_1^*, \dots, \alpha_l^*)^T$ 为最优解，据此我们所求的模糊最优分类函数可以表达为如下形式：

$$f(x) = \text{sgn} \left(\sum_{j=1}^l \alpha_j^* y_j K(x, x_j) + b^* \right), x \in R^n \quad (13-35)$$

式中， $b^* = y_i - \sum_{j=1}^l y_i \alpha_j^* (x_i \cdot x_j)$ ， $j \in \{j \mid 0 < \alpha_i^* < s_j C\}$ 。

对于式 (13-35) 中 α_j 值，只有当样本点为支持向量时才会有非零值的可能，其他情况均为零值。因此，只有支持向量才会使用式 (13-35) 进行求和。

综上所述，可以得出模糊支持向量机的基本思想是，通过引入模糊数学的理论构造隶属度函数，通过计算各样本点的隶属度来控制惩罚系数的大小，从而减少噪声和孤立点对分类的贡献度。采取这种措施就可以尽量减小噪声点和孤立样本点的隶属度数值，从而有效地减少噪声和孤立点对训练过程中分类性能造成的不利影响。

13.4 模糊隶属度核函数的选择

通过前面的讨论我们可以得知，模糊支持向量机分类性能的好坏很大程度上取决于模糊隶属度函数。所以，所构造的隶属度函数的好坏很大程度上决定了分类器的分类能力，对隶属度函数的构造也就自然而然地成为模糊支持向量机的关键问题之一。Lin 等学者提出选取一个合适的隶属度函数需要满足以下两点：

- 1) 定义隶属度的下界。
- 2) 隶属度函数的选取必须和样本的特性存在着一些联系，即必须根据样本特性选取合适的隶属度函数。

下面我们介绍几种常用的隶属度函数。

13.4.1 基于类中心距离的隶属度函数

基于类中心距离的隶属度函数具有很多优点，比如它的结构非常简单，所

以函数构造起来非常方便，而且隶属度就是样本点到类中心点的距离，隶属度所表达的概念直观易于理解，所以该方法是目前最常用到的构造隶属度函数的方法。基于类中心距离的隶属度函数，顾名思义，隶属度的值就是距离值，是样本点到它所属类的中心点位置的距离，因此它是一个线性函数。所以，当某个样本点的隶属度较大时意味着该点距离所属类的中心位置很近，此样本点在类的中心位置附近，相反，如果隶属度值较小，意味着它远离类的中心位置。

不妨假设对于 R^n 上某类的样本点 $\{x_1, x_2, \dots, x_k\}$ ，设 x_c 为该类中心点，

$$x_c = \frac{1}{k} \sum_{i=1}^k x_i \quad (13-36)$$

类半径 r 为样本到类中心的最远距离，即

$$r = \max \|x_c - x_i\| \quad (13-37)$$

式中， $i \in [1, k]$ 。

我们按照距离确定隶属度，那么该类中每个样本的隶属度可以表示为

$$\mu(x_i) = 1 - \frac{\|x_c - x_i\|}{r + \delta} \quad (13-38)$$

式中， $\delta > 0$ 为一个非常小的常数，是为了避免出现隶属度为零值的情况。

13.4.2 基于 S 形函数的隶属度函数

基于类中心距离的隶属度函数所呈现的隶属度简单来说就是类中心与样本点之间的一种线性关系，但是在 S 形隶属度函数中就不是简单的线性关系了，它是一种在 $[0, 1]$ 上单调递增的非线性函数关系。Zadeg 设计的 S 形函数公式为

$$\mu_i(x, a, b, c) = \begin{cases} 0 & x_i < a \\ 2[(x_i - a)/(c - a)]^2 & a \leq x_i < b \\ 1 - 2[(x_i - a)/(c - a)]^2 & b \leq x_i < c \\ 1 & x_i \geq c \end{cases} \quad (13-39)$$

式中， $b = \frac{a+c}{2}$ ，且 $x_i = b$ 时，样本的隶属度为 0.5。

13.4.3 基于 π 形函数的隶属度函数

π 形函数，顾名思义，该函数的图形与字符 π 的结构很相似，中间平坦，两边衰减较快。 π 形函数可以用公式表示为如下形式：

$$\mu_\pi(x) = \begin{cases} \mu_s(x, a, b, c) & x_i \leq c \\ 1 - \mu_s(x, a, b, c) & x_i > c \end{cases} \quad (13-40)$$

不难看出， π 形函数可以用 S 形函数来定义。

13.5 模糊支持向量机算法的编程实现

本章考虑到实时性与可移植性的要求，采用 C++ 语言进行开发设计。结合 Libsvm 工具包，加入了基于类中心距离的模糊隶属度函数，构建了模糊支持向量机识别算法，并对序列最小优化算法进行了编程实现。使用 Visual C++ 6.0 环境进行编译运行，开发界面如图 13-9 所示。

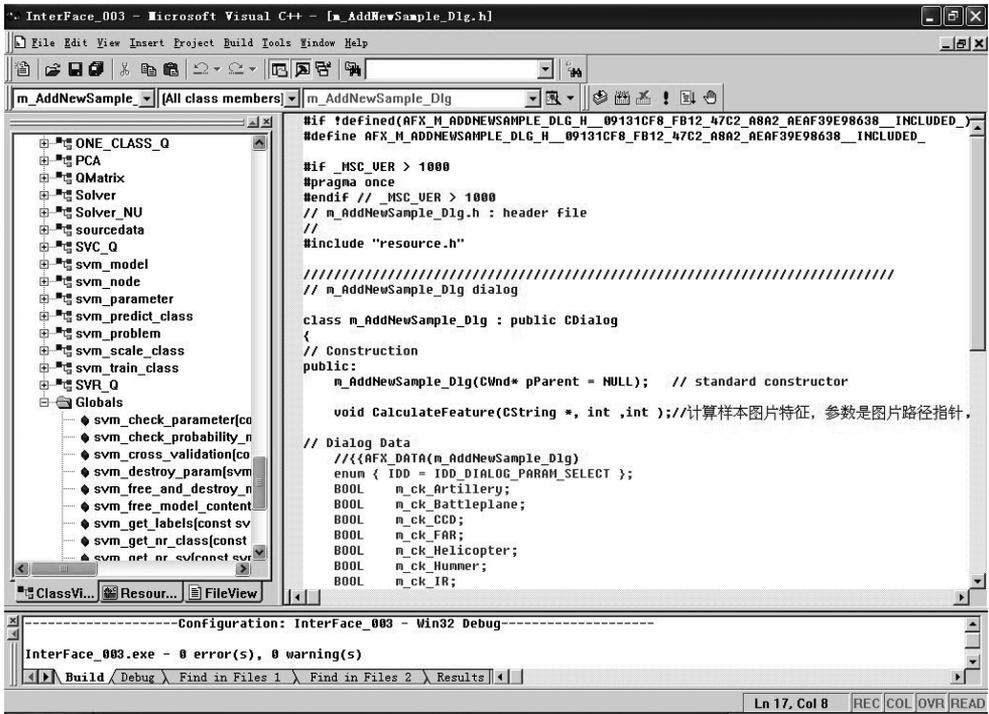


图 13-9 模糊支持向量机编程开发界面

在 LIBSVM 软件的基础上进行编程，其中核函数定义编码如下：

```
static double dot(const svm_node * px,const svm_node * py);
double kernel_linear(int i,int j) const
{ return dot(x[i],x[j]); }
double kernel_poly(int i,int j) const
{ return powi(gamma * dot(x[i],x[j]) + coef0,degree); }
double kernel_rbf(int i,int j) const
```

```

{ return exp(-gamma * (x_square[i] + x_square[j] - 2 * dot(x[i], x[j]))) ); }
double kernel_sigmoid( int i, int j) const
{ return tanh( gamma * dot( x[i], x[j] ) + coef0 ); }
double kernel_precomputed( int i, int j) const
{ return x[i][ (int)(x[j][0].value) ].value; }

```

基于类中心距离的模糊隶属度值获取函数定义如下：

```

void calculate_S()
{ for( int i=0; i < N; ++i )
  { xc + = x[i]; }
xc / = N;
for( i=0; i < N; ++i )
  { s[i] = 1 - mod( xc, x[i] ) / ( r + delt ); } }

```

具体实现模糊支持向量机的函数调用流程如图 13-10 所示。

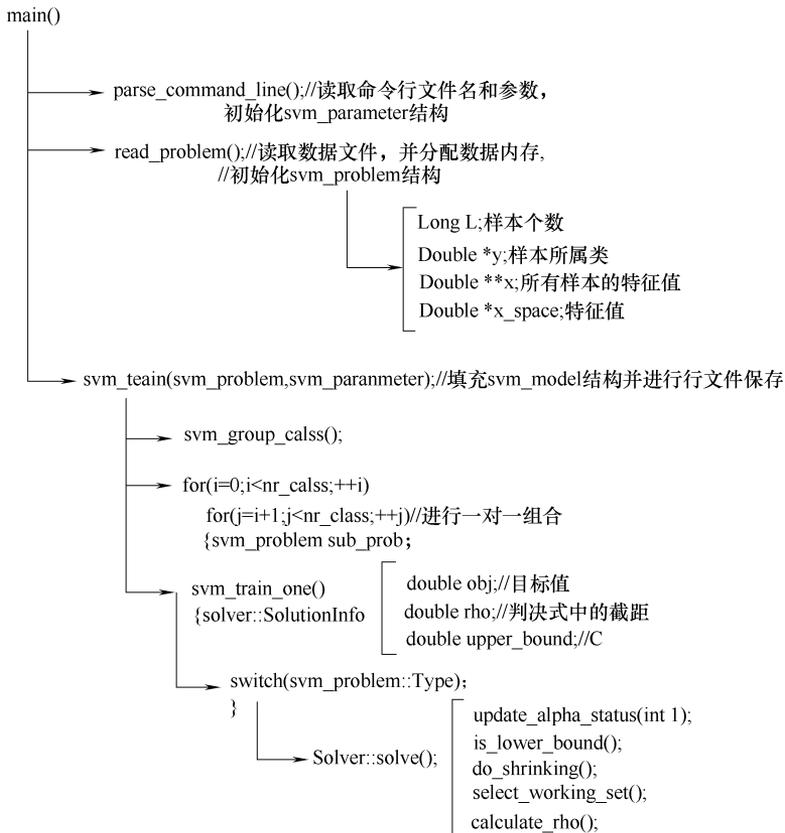


图 13-10 模糊支持向量机函数调用过程的流程图

13.6 本章小结

本章主要讲述了模糊数学的基础理论知识、支持向量机的理论知识和模糊支持向量机的基础理论知识。首先通过对模糊数学中的模糊集概念、隶属度函数的概念的介绍，以及对支持向量机中的统计学习理论、最优分类超平面、核函数等知识点的介绍，让我们对模糊支持向量机有了初步的认识。之后对模糊支持向量机理论进行了详细的介绍。主要介绍了两种不同类型的模糊支持向量机的理论知识、公式的推导过程以及模糊核函数的选择问题等内容，使我们对模糊支持向量机有了更深刻的认识和理解。最后结合 LIBSVM 软件工具包，通过编程对模糊支持向量机算法进行了实现，为下一步工作打下了良好的基础。

基于模糊支持向量机的识别系统实现

14.1 模糊支持向量机识别系统的架构

基于模糊支持向量机 (FSVM) 的目标识别系统由图像采集模块、图像预处理模块、目标特征提取模块、目标特征数据库管理模块、识别模型训练模块、目标识别模块等多个模块构成。其中系统的各模块主要负责完成以下功能：

图像采集模块：负责可见光、红外图像的采集工作，并将采集到的图像正确显示在系统显示端。

图像预处理模块：负责数字图像的预处理，包括去噪、增强、分割等过程，是目标识别系统中的首要环节，是非常重要的模块之一，其性能的好坏直接决定了识别率的高低。

目标特征数据库管理模块：目标特征数据库存储目标特征数据，支持基于专家系统的特征识别的识别规则库和目标特征描述知识库，通过数据库管理模块可以对数据库中的数据进行添加、删除等操作。

识别模型训练模块：使用特征数据库中已有的训练样本的特征数据进行识别模型的训练工作，通过训练得到模型参数，从而获得可以使用的识别模型。

目标识别模块，使用训练好的识别模型对计算得到的图像中目标的特征值进行计算，根据计算结果得到最终的目标类别。

设计框图如图 14-1 所示。

为了提高系统的实时性，系统整体采用 C++ 语言设计，为保证系统的稳定运行大量借鉴了微软 MFC 类库进行开发，开发平台为 Windows XP 操作系统，开发软件为 Visual C++ 6.0，硬件配置为 Intel (R) Core i3 处理器，RAM 容量为 2GB (1.74GB 可用)，CPU 主频为 2.53GHz。系统运行时主界面如图 14-2 所示。

可以看到识别系统界面美观简洁，没有复杂的数据项目。系统实行一键式操作，最大程度地简化了操作人员的操作过程，易于使用。

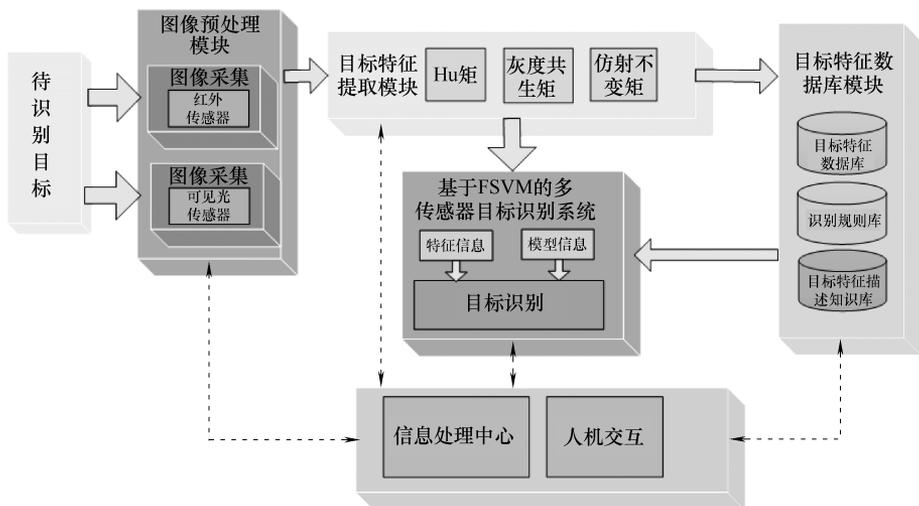


图 14-1 基于 FSVM 的目标识别系统设计框图



图 14-2 基于 FSVM 的目标识别系统主界面

14.2 图像预处理部分的实现

图像预处理模块采用多文档视图进行开发，界面简洁美观，运行时界面图如图 14-3 所示。可以看到图像预处理模块中包含了数字图像预处理的一些基本处理方法，包括几何变换、灰度变换、图像空间域去噪增强、图像频域去噪增强、图像形态学处理以及图像分割等工具。

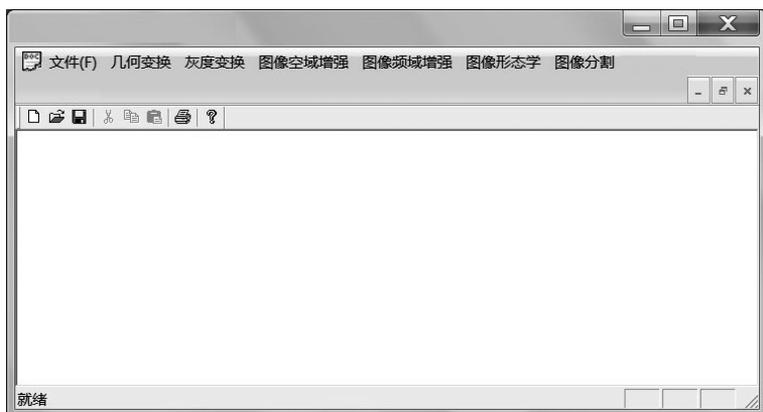


图 14-3 图像预处理模块界面

14.2.1 文件的打开与显示

为方便使用人员进行操作，在设计图像预处理模块时考虑了多种常用情况，开发实现了多种便捷操作。使用人员可以通过菜单栏中的文件下拉菜单对文件进行新建文件、文件打开、关闭文件、文件保存、文件另存为、退出系统等操作，并支持快捷键操作。这里首先选择需要进行特征提取的图片并显示于工作区域，操作结果如图 14-4 所示。



图 14-4 图像预处理模块文件显示界面

通过图 14-4 我们可以看到，选中的文件被正确加载到图像预处理模块中，并能够在界面中正确显示，图像显示清晰正常，没有出现信息缺失的问题。图片默认显示在工作区域左上角位置。

14.2.2 预处理模块图像增强的执行

通过上一步操作，我们已经成功将需要处理的图片加载到图像预处理模块中并显示出来，下面我们就开始对图像进行图像增强的操作。如图 14-5a 中标注所示，单击工具栏中的“图像空域增强”，选择中值滤波，滤波之后的图像如图 14-5b 所示。



图 14-5 图像预处理模块空域增强前后对比图

通过图 14-5 中滤波前后的对比图我们可以看出，中值滤波后的图像边缘明显比滤波之前变得平滑了。

“图像空域增强”下拉菜单栏中还包含了梯度锐化算法、邻域平均算法、拉普拉斯锐化算法等常用的数字图像增强方法，操作人员可以根据需要选择适当的增强方法即可实现图像的去噪增强等操作。

14.2.3 预处理模块图像分割的执行

对于图像分割的操作，它与图像增强的操作类似，操作人员只需单击“图像分割”选项，在弹出的下拉菜单栏中选择需要的分割方法即可。图像分割模块中主要实现了基本的 Otsu 算法、K-means 算法、区域生长、边缘检测等算法，另外根据需要添加了改进后的分割算法。

图 14-6 中给出了使用改进后的分割算法对图像进行分割的效果图。

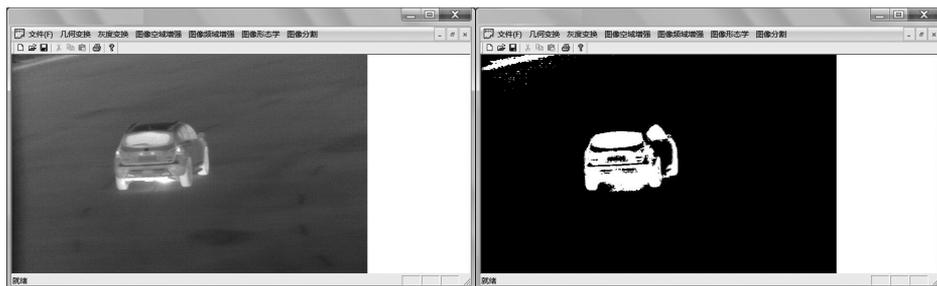


图 14-6 预处理模块实现改进后的分割效果图

图像预处理中其他操作与图像增强和图像分割的操作过程基本一致。这里不再赘述。

14.3 特征提取部分的实现

该模块使用 Visual C++ 6.0 平台进行搭建,采用 C++ 语言编程。该模块具有较好的实时性,模块化编程程度较高,可移植性好。模块主要实现了对图像数据以及非图像数据的各种特征的提取,主要有 Hu 矩、Zernike 矩、仿射不变矩、灰度共生矩。系统采用多文档界面进行操作。

目标特征获取模块初始界面如图 14-7 所示。

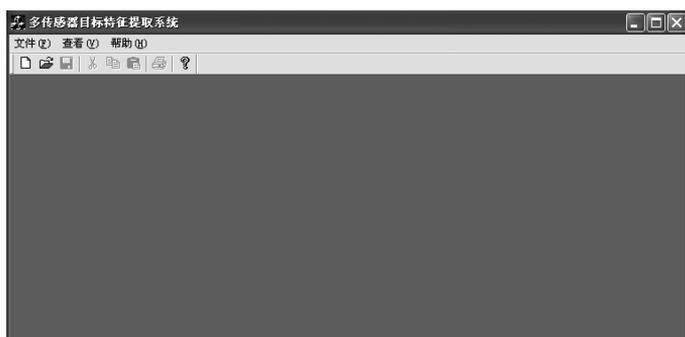


图 14-7 目标特征获取系统界面

下面将详细介绍具体操作步骤。

14.3.1 文件的打开与显示

与图像预处理模块类似,目标特征提取模块也最大化简化了使用人员的操作过程,比如最近打开过的文件名也在文件菜单中进行显示等。使用人员可以通过菜单栏中的文件下拉菜单对文件进行新建文件、文件打开、关闭文件、文件保存、文件另存为、退出系统等各种操作,并支持快捷键操作。这里首先选择需要进行特征提取的图片并显示于工作区域,操作结果如图 14-8 所示。

可以看到选中的文件被正确加载到特征提取模块中,并在界面中正确显示。图片默认显示在工作区域左上角位置。

14.3.2 传感器类型的选择

通过上一步操作,使用人员选择好需要进行特征提取的文件之后便可开始进行特征提取。这里对应不同的传感器类型,所获得文件格式和类型可能会有所差异,比如传感器数据包含红外图像数据和可见光图像数据,因此需要选择



图 14-8 文件的显示

对应传感器类型菜单下的特征提取方法进行特征提取。操作如图 14-9 所示。



图 14-9 传感器类型选择

14.3.3 计算相应的特征值

传感器类型菜单中有多种不同的特征类型，单击需要计算的特征类型之后系统开始计算图像特征，之后特征数据就会以对话框的方式显示出来，如图 14-10 所示。



图 14-10 特征值计算

14.3.4 特征值的存储

特征数据计算完毕之后操作人员可以选择对其进行存储。单击“确定”按钮后，系统自动调用数据库 API 函数与数据库连接，并对特征数据进行存储。之后我们便可以通过数据库模块访问和使用已保存过的特征数据。

14.4 目标特征数据库系统的技术实现

14.4.1 特征数据库管理系统的构建

目标识别系统数据库管理系统采用关系数据模型构建，包含多个二维数据表以及模型数据和必要的密钥数据。通过数据库管理系统操作人员可以获得数据库中的特征数据以及识别模型数据，当然操作人员首先要通过密钥验证。

图 14-11 所示为本数据库管理系统开发的技术流程。

数据库管理系统主要是由以下几个模块构成，包括数据添加模块、数据删除模块、数据更新模块、数据获取模块和身份验证模块。通过知识库管理系统，技术人员可对知识库中的数据进行各种操作。常用的操作有新的数据的添加完善、数据内容的修改编辑、无效或过期数据的删除、数据的更新等操作。

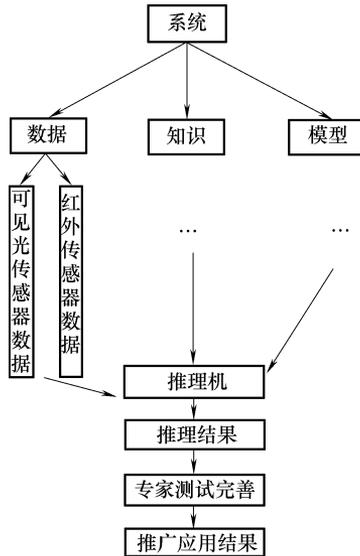


图 14-11 数据库管理系统开发流程图

14.4.2 目标动态特征库管理技术

本章系统平台上构建的目标识别系统面对的是多种目标类别的识别分类知识的组织。新类别的知识、规则和数据的组织从结构上很难完全与原有类别的知识、规则和数据的结构一致。即使采用相同的识别算法也难于做到结构上的一致。因此，本章中所有类别均采用相同的 FSVM 识别算法。在系统的 FSVM 算法中需要存放经学习后的权阵表，用以在系统运行时进行分类识别使用。

由于不同目标类别的特征数据和类别型号不一样，在对新目标进行模型训练后，获得的模型数据与原有数据格式不一致，因此我们需要动态地改变识别模型数据在数据库中的存储格式。对于同一类别的权阵表结构而言，随着系统的运行，子类别数有可能增加或改变，提取的特征数据的维数也可能增加或改变。

针对上述两种原因，本章所建数据库不能使用数据库结构模板，也不能使用 SQL 程序简单构建其数据库结构。由于数据库管理系统的局限性，要实现数据库的动态模式创建必须要通过借助静态模式的创建工具。可以简单理解为使用已存在的静态表创建可变的动态表。据此，我们给出动态数据库的构建流程如下：

- 1) 首先将创建动态表所需要的数值作为数据存入静态表中，并将此静态表命名为结构参数表。

2) 动态表中的所有数据顺序存于另一个静态表中, 保持数据类型一致, 并将存放数据的表命名为流式数据表。

3) 完成以上两步, 动态数据表就基本创建完毕了, 当我们需要使用动态表时, 只需要通过给定的结构参数去流式数据表中搜寻对应的数据, 然后恢复出来即可。其系统结构如图 14-12 所示。

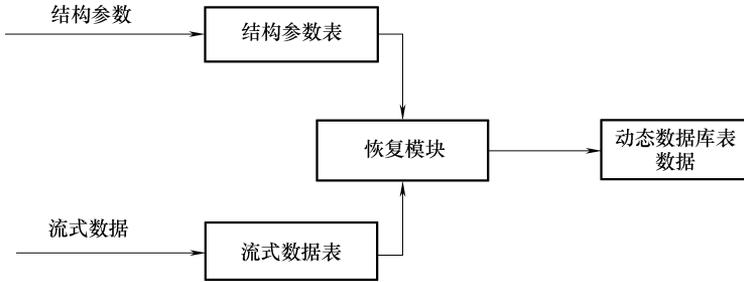


图 14-12 系统结构图

本章中, 系统采用 FSVM 算法进行目标识别, 因此无论是新增一种目标类别, 或是重新训练识别模型, 都要求数据库能够针对这些改变进行动态的调整。

因此, 在对数据库进行开发时, 设计了两种表来解决上述问题。其一是目标的 FSVM 算法参数表, 它存放各类目标的权阵表的结构及其他可用信息。另一类是权阵表, 用以存放每种目标的权阵数据。

图 14-13 所示为基于 SQL 的样本特征对象资源管理。

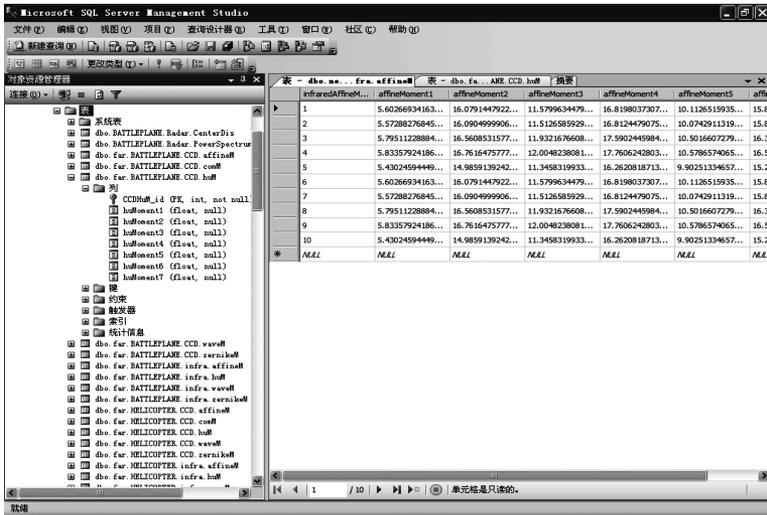


图 14-13 基于 SQL 的样本特征对象资源管理

14.4.3 目标特征数据库建立

本章目标特征知识数据库设计采用关系结构模型。关系结构模型最为突出的优点就是能够以简单的二维表格的形式直观地表示出各种复杂的数据结构之间的关系。结构简单，清晰直观。图 14-14 为本章中样本特征数据的存储结构。



图 14-14 基于 SQL 的样本特征数据

图 14-15 和图 14-16 所示分别本章中样本特征数据表的事例。

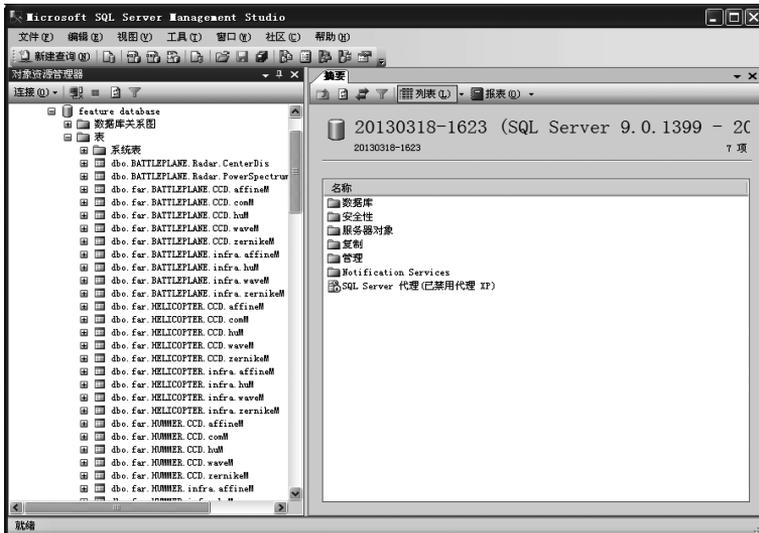


图 14-15 基于 SQL 的样本特征数据表 1

为了保证项目的识别准确率，保证软件识别的精准性，本章建立了独立的特征数据库，数据库一共对汽车、卡车、战斗机、直升机四大类图像中 20 余种型号的模型，进行了可见光和红外图像的提取，每个模型按照垂直方向（90°每

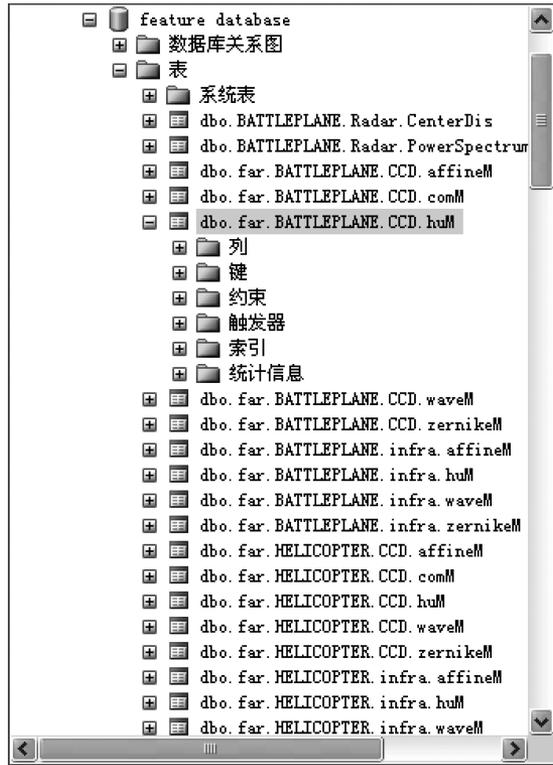


图 14-16 基于 SQL 的样本特征数据表 2

隔 5°) 和水平方向 (360°每隔 5°) 进行样本拍摄, 每个模型共提取 1250 张图片, 数据库中共计包含 5000 张图片。

14.4.4 数据库管理系统编程实现

数据库管理子系统使用开放数据库互连将 Visual C++ 6.0 与数据库相连接, 使用 MFC 的标准类库函数对数据库进行各种必要的操作以及数据的读取与写入。数据库系统可分为四个部分: 样图显示部分、文件处理部分、模型训练部分, 以及数据显示部分。

图 14-17 为数据库登录密钥界面。

数据库管理系统界面如图 14-18 所示。

(1) 样图的显示

样图显示功能用于操作人员直观判断图像传感器拍摄的样本图片质量、图片类型的好坏。操作人员单击“样图显示”按钮, 选择需要观察的图像数据, 单击“确定”按钮之后样图便会显示在图像显示区域。操作过程如图 14-19 和图 14-20 所示。



图 14-17 数据库登录密钥界面

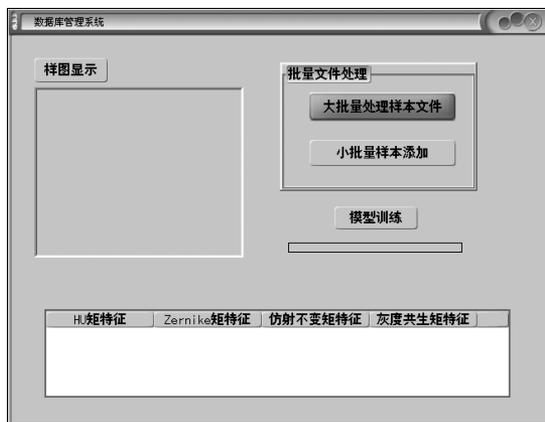


图 14-18 数据库管理系统界面



图 14-19 文件选择



图 14-20 样图显示

(2) 文件的处理

批量文件处理通过针对不同的情况采用不同的操作方式进行特征数据提取，最大化简化操作人员的操作复杂度。样本数量较少时可以采用小批量处理模式进行文件处理。单击“小批量样本添加”按钮，在弹出的对话框中首先按照样本类型进行参数的设定，如图 14-21 所示。

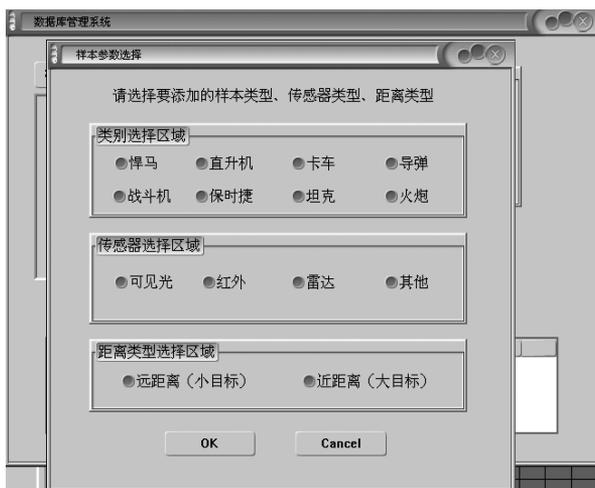


图 14-21 参数选择

设置好参数之后单击“确定”按钮，这时系统会弹出文件选择对话框，如图 14-22 所示。

系统允许一次性选择多个文件同时进行处理。选择好文件之后单击“打开”按钮进行特征提取。特征提取完毕之后会将数据显示在数据显示区域内，并弹出“存储完成”对话框，如图 14-23 所示。



图 14-22 多文件选择界面



图 14-23 存储完成提示

14.5 模型训练部分

上一步提取的特征数据存入数据库后，操作人员就可以在数据库管理界面进行模型的训练操作。单击“模型训练”按钮后系统开始执行模型训练功能，对样本库中所有数据进行训练，通过网络寻优、交叉验证等训练过程后得到最优模型参数。操作界面如图 14-24 所示。

模型训练过程结束后会弹出“训练结束”提示对话框表示训练成功结束，如图 14-25 所示。训练结束后，训练得到的识别模型参数将自动保存到数据库



图 14-24 模型的训练过程

中。当识别模块执行目标识别过程时便可以直接从数据库中取得该模型进行目标识别。

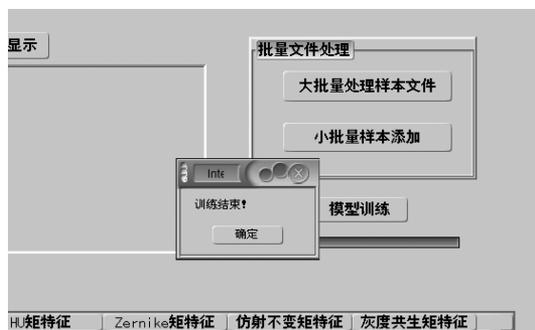


图 14-25 模型训练结束

14.6 目标识别部分

14.6.1 目标识别模块工作流程与结果显示

“目标识别”按钮位置如图 14-2 所示，位于主界面右上角部分。单击“目标识别”按钮后系统将会自动执行目标识别过程，包括图像预处理、特征提取、识别模型获取以及目标值计算，最终得到目标类别，并在识别结果显示区域进行显示。具体流程图如图 14-26 所示。

目标识别过程为一键式操作，操作人员只需要单击“目标识别”按钮即可，实用性高。最终识别结果将显示在文本控件中，如图 14-27 所示。

14.6.2 目标识别结果分析与对比

识别模型的识别能力是目标识别系统最为重要的部分。我们将通过实验对

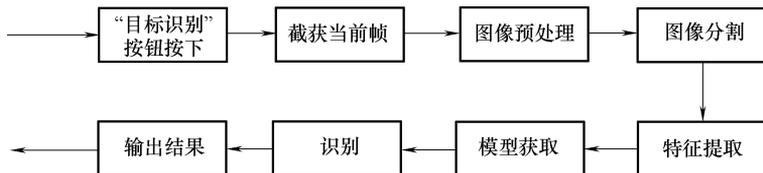


图 14-26 “目标识别”按钮响应流程

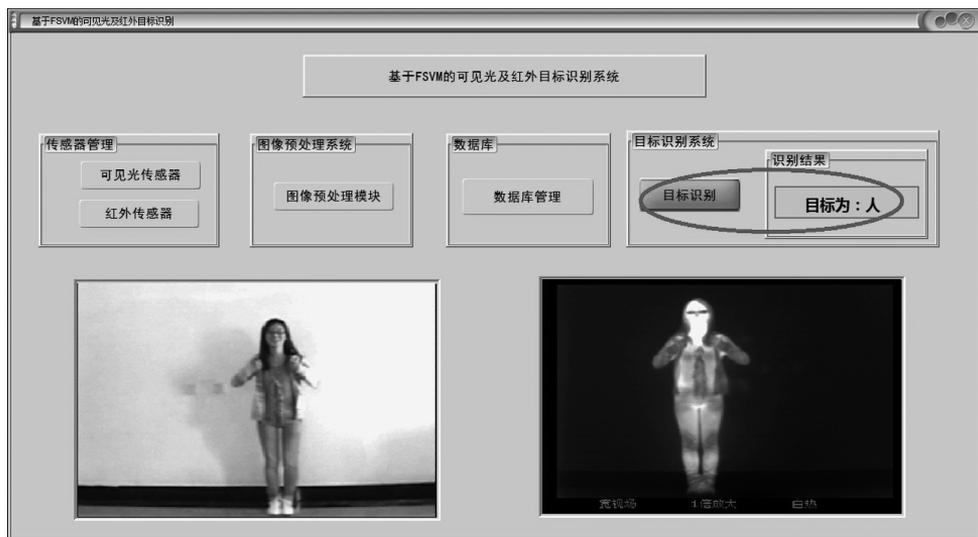


图 14-27 识别结果显示图

识别模型进行验证。

本章使用反向传播 (Back Propagation, BP) 神经网络算法与 FSVM 算法进行识别精度比对。本次实验共有四类训练目标, 包括汽车、坦克、战斗机以及人物。对每类目标分别进行了可见光和红外图像的提取, 每类目标按照垂直方向 (90°每隔 10°) 和水平方向 (360°每隔 10°) 进行样本拍摄, 每个模型共提取 324 张图片, 每张图片计算出目标特征值存于数据库中。每类目标取数据集约 80% 数据用作训练数据, 剩余约 20% 数据用作测试数据。实验对比结果见表 14-1。

表 14-1 BP 神经网络算法与 FSVM 算法识别结果比较

算 法	识别样本总数	分类正确数	分类错误数	识别正确率
BP 神经网络	270	203	67	75.3%
FSVM 算法	270	241	29	89.2%

通过对比可以看出 FSVM 算法相较于 BP 神经网络算法具有更高的识别率,

识别精度良好。其中 FSVM 识别模型经过训练样本训练之后的识别能力测试图如图 14-28 和图 14-29 所示。

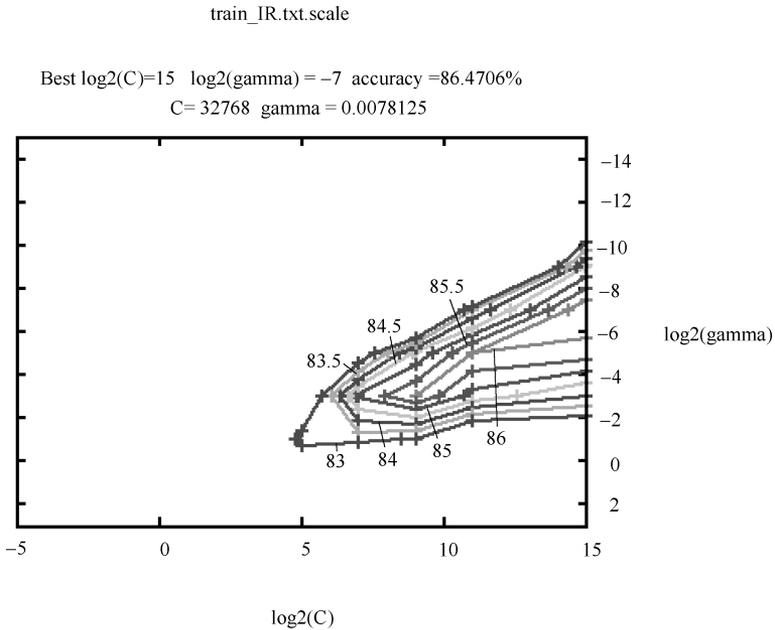


图 14-28 识别模型对红外图像识别测试结果图

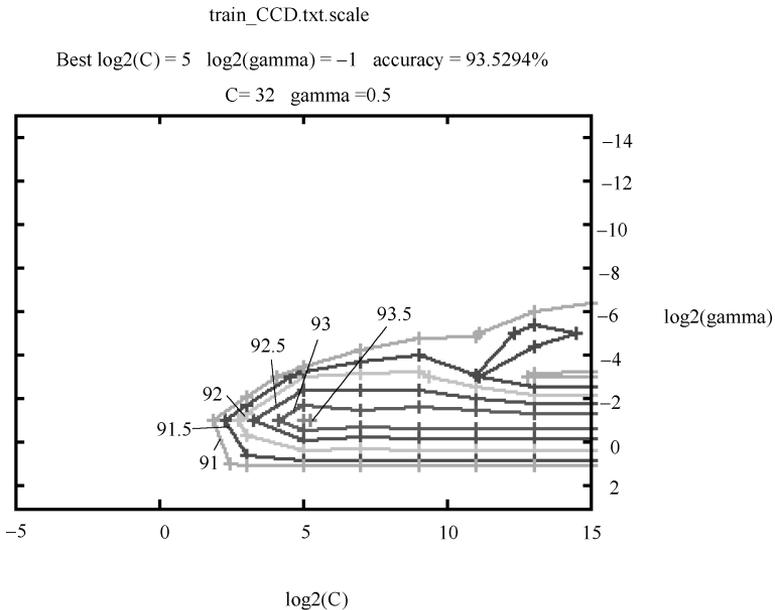


图 14-29 识别模型对可见光图像识别测试结果图

图中主要记录了识别模型的重要参数数值，以及正确识别率等高线的绘制。由图 14-29 可以看出，训练后的识别模型对于红外图像中目标的正确识别率最高可以达到 86%，对于可见光图像的最大正确识别率更是高达 93.5%。本次试验结果表明，基于 FSVM 的目标识别系统识别速度与识别精度令人满意。系统运行稳定，操作简单，可用性较高。

14.7 本章小结

本章从整个识别系统的实现情况出发，首先讲述了基于 FSVM 的红外及可见光目标识别系统的设计框架，然后对各模块设计流程、运行情况以及基本的操作流程进行了详尽的讲解。最后，对系统的目标识别模型的正确识别率进行了实验验证。对采集到的大量目标图像进行识别，实验结果表明系统整体运行稳定，识别速度以及识别精度令人满意。该系统具有一定实用价值。

参 考 文 献

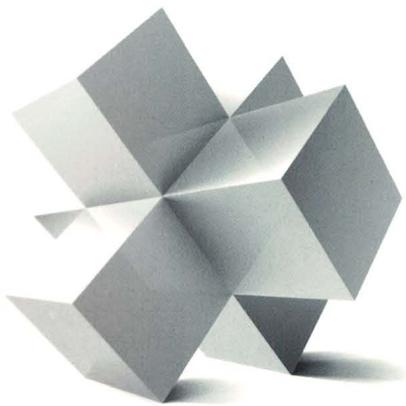
- [1] 孙红岩, 毛士艺. 多传感器目标识别的数据融合 [J]. 电子学报, 1995, 10: 188-193.
- [2] Roger Schmidt. Benefits of IR-visible fusion [J]. Proc. of SPIE 2007. 2007; 6541-6545.
- [3] M. D. Shahbe and S. Hati. Decision Fusion Based on Voting Scheme for IR and Visible Face Recognition [J]. Computer Graphics, Imaging and Visualization, 2007. CGIV' 07 14-17. 2007, 8: 358-364.
- [4] 李新德, 杨伟东, DEZERT Jean. 一种飞机图像目标多特征信息融合识别方法 [J]. 自动化学报, 2012, 08: 1298-1307.
- [5] Gunatilaka, A. H, et. al. Feature-Level Fusion and Decision-Level Fusion of Noncoincidentally Sampled Sensors for Land Mine Detection [J]. IEEE Transactions on pattern analysis and machine intelligence, Vol 23, No. 6 2001: 577-589.
- [6] 徐南荣, 翟荣贞. 飞行器的红外特性 [J]. 红外与激光工程, 1999, V01. 28, No. 1.
- [7] 彭真明, 张启衡. 基于多特征融合的图像匹配模式 [J]. 强激光与粒子束. 2004, V01. 16, No. 3.
- [8] 王彪, 姜志国, 赵丹培. 基于多特征提取的遥感图像机场目标自动检测 [J]. 中国体视学与图像分析, 2009, 14 (2): 120-124.
- [9] 贾平, 徐宁, 张叶. 基于局部特征提取的目标自动识别 [J]. 光学精密工程, 2013 (7): 1898-1905.
- [10] PENG H, LONG F, DING C. Feature selection based on mutual information criteria of max-dependency. maxrelevance and min-redundancy [J]. IEEE Trans. Patt. Analy. And Mach Intell, 2005, 27 (8): 1226-1238.
- [11] Graf A B A, Smola A J, Borer S. Classification in a normalized feature space using support vector machines [J]. Neural Networks, IEEE Transactions on, 2003, 14 (3): 597-605.
- [12] Weston J, Watkins C. Multi-class support vector machines [J]. Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, May, 1998.
- [13] Tan Y, Wang J. A support vector machine with a hybrid kernel and minimal Vapnik-Chervonenkis dimension [J]. Knowledge and Data Engineering, IEEE Transactions on, 2004, 16 (4): 385-395.
- [14] 王元斌, 夏学知. 多传感器综合目标识别技术研究 [J]. 舰船电子工程, 2004 (4): 8-11, 16.
- [15] Vapnik V. The nature of statistical learning theory [M]. springer, 2000.
- [16] X. J. Wu, D. X. Su, X. Q. Luo, et. A new similarity function for region based image fusion incorporating gabor filters and fuzzy C-means clustering. Proc. SPIE6625, 66250Z (2007); doi: 10. 1117/12. 791022.
- [17] 蒋晓瑜, 梁浩聪, 王加, 等. 目标识别中多传感器信息融合算法比较 [J]. 计算机系统应用, 2013 (4): 10-13, 5.
- [18] 牛丽红, 倪国强. 多传感器目标识别系统的特征优化方法 [J]. 光学技术, 2005 (3): 420-423, 426.
- [19] Lin C F, Wang S D. Fuzzy support vector machines. Neural Networks [J]. IEEE Transactions on, 2002, 13 (2): 464-471.
- [20] Inoue T, Abe S. Fuzzy support vector machines for pattern classification. Neural Networks, 2001. Proceedings. IJCNN01. International Joint Conference on. IEEE, 2001, 2: 1449-1454.
- [21] 周静, 黄心汉, 彭刚. 基于多特征融合的飞机目标识别 [J]. 华中科技大学学报, 2009, 37 (1):

- 38-41.
- [22] 王广宇. 车辆牌照识别系统综述 [J]. 郑州轻工业学院学报: 自然科学版, 2001 (2).
- [23] 高文琦, 张复春, 李雁, 等. 基于多特征融合和极坐标变换的目标识别 [J]. 计算机工程, 2011, 37 (11): 197-199.
- [24] 金敏, 徐守时, 汪行. 不变矩在模式识别中的应用研究 [J]. 计算机工程与应用, 2004, 25: 65-67.
- [25] 余静, 游志胜. 自动目标识别与跟踪技术研究综述 [J]. 计算机应用研究, 2005 (1): 12-15.
- [26] 陈绵书, 陈贺新, 桑爱军. 计算机人脸识别技术综述 [J]. 吉林大学学报: 信息科学版, 2003, 21 (5): 101-109.
- [27] 郑坤鹏, 王普凯, 周国印, 等. 基于 MATLAB 的红外图像中坦克特征量的提取方法 [J]. 红外技术, 2010. 32 (11): 625-628.
- [28] 翟俊海, 赵文秀, 王熙照. 图像特征提取研究 [J]. 河北大学学报: 自然科学版, 2009, 29 (1): 106-112.
- [29] 余丽萍, 黎明, 杨小芹, 等. 基于灰度共生矩阵的断口图像识别 [J]. 计算机仿真, 2010, 4 (27): 224-227.
- [30] M. K. Hu. Visual pattern recognition by moment invariants [J]. IRETrans. Information theory, 1962, IT (8): 179-187.
- [31] 郭军, 倪晋平, 王会峰, 等. 一种使用小波矩的飞行器目标识别方法 [J]. 西安工业大学学报, 2008, 28 (2): 108-110.
- [32] 杨阳, 胡玉兰. 主成分分析的红外与可见光图像特征融合 [J]. 沈阳理工大学学报, 2012, 4: 7.
- [33] 马炎. 小样本人脸图像识别研究 [D]. 南京: 南京信息工程大学, 2011.
- [34] 杨阳. 基于红外与可见光图像的特征融合方法研究 [D]. 沈阳: 沈阳理工大学, 2013.
- [35] Deng Yong, Jiang Wen, Xu Xiaobin, Li Qi and Wang Dong. Determining BPA under uncertainty environments and its application in data fusion [J]. Journal of Electronics (China) 2009, 26 (1): 13-17.
- [36] A. L. Cunha, J. Zhou, M. N. Do. The nonsubsampled contourlet transform: Theory, design and application [J]. IEEE Transactions on Image Processing, 2006, 15 (10): 3089-3101.
- [37] 杨露菁, 郝威. 多传感器目标识别的神经网络与证据理论结合方法 [J]. 探测与控制学报, 2006 (1): 40-43.
- [38] 史颜玲, 王忠义. 基于形状特征的图像特征提取方法及其在医学图像分析中的应用 [J]. 许昌学院学报, 2011, 30 (6): 69-71.
- [39] 程显毅, 李小燕, 任越美. 图像空间关系特征描述 [J]. 江南大学学报: 自然科学版, 2007, 6 (6): 637-641.
- [40] 宋玮. 基于双源多特征的步态信息融合技术研究 [D]. 天津: 天津大学, 2009.
- [41] 潘春雨, 卢志刚, 杜静. Zernike 矩的不变性分析及其改进 [J]. 火力与指挥控制, 2010, 35 (增): 10-11.
- [42] O. Tuzel, F. Porikli, P. Meer. Region covariance: a fast descriptor for detection and classification [C]. In proceedings of the 9th European Conf. on Computer Vision, Graz, Austria, 2006. 2, 589-600.
- [43] PEREZ M, RUBIN D M, SCOTT L E, et al. A hybrid fuzzy-SVM classifier, applied to gene expression profiling for automated leukaemia diagnosis [C]. Proceedings of the IEEE Conference Israel. Eilat, Israel: IEEE, 2008.
- [44] 陆其明. DirectShow 实务精选 [M]. 北京: 北京科海电子出版社, 2004.

- [45] 陆其明. DirectShow 开发指南 [M]. 北京: 清华大学出版社, 2003.
- [46] 熊超, 陆起涌, 田小芳. 基于 DirectShow 的多线程视频采集系统 [J]. 计算机工程, 2006, 10: 221-223.
- [47] 黄颖. 基于 DirectShow 的视频播放系统研究与实现 [D]. 武汉: 华中科技大学, 2006.
- [48] 姚军, 蒋晓瑜, 黄应清. 目标识别中 Hu 矩、Zernike 矩和小波矩的比较 [J]. 装甲兵工程学院学报, 2006, 20 (3): 34-36.
- [49] 王大伟. 基于特征级图像融合的目标识别技术研究 [D]. 北京: 中国科学院研究生院, 2010.
- [50] 刘波. 粒子群优化算法及其工程应用 [M]. 北京: 电子工业出版社, 2010.
- [51] 高鹏毅. BP 神经网络优化分类器研究 [D]. 武汉: 华中科技大学, 2012.
- [52] 李松, 刘力军, 解永乐. 遗传算法优化 BP 神经网络的短时交通混沌预测 [J]. 河北大学管理学院, 2011 (6): 1-3.
- [53] 刘彩红. BP 神经网络学习算法的研究 [D]. 重庆: 重庆师范大学, 2008.
- [54] 徐鹤鸣. 多目标粒子群优化算法的研究 [D]. 上海: 上海交通大学, 2013.
- [55] 徐小琴. 多传感器数据融合目标识别算法综述 [J]. 红外与激光工程, 2006, S4: 321-328.
- [56] Philip J. A noniterative algorithm for determining all essential matrices corresponding to five point pairs [J]. Photogrammetric Record, 1996, 15 (88): 589-599.
- [57] 吴朝福, 胡占义. 关于 P5P 问题的研究 [J]. 软件学报, 2001, 12 (5): 768-774.
- [58] Quan L, Lan Z D. Linear N-point camera pose determination [J]. IEEE Transactions on PatternAnalysis and Machine Intelligence, 1999, 21 (8): 774-780.

本书内容包括：

- ◎ 特征级融合目标识别的基本理论。
- ◎ 多源图像的预处理、结合阈值分割的分水岭算法、结合聚类分割的分水岭算法、目标特征提取方法。
- ◎ 基于协方差矩阵的多特征信息融合、基于主成分分析的特征融合方法、基于改进免疫遗传的特征融合方法、基于独立分量的特征融合，以及对典型相关分析特征融合方法的改进。
- ◎ 基于优化改进的反向传播神经网络目标识别、模糊支持向量机理论与编程，以及基于模糊支持向量机的识别系统。



INTELLIGENT INFORMATION FUSION AND TARGET RECOGNITION METHOD

地址：北京市百万庄大街22号
邮政编码：100037

电话服务

服务咨询热线：010-88361066

读者购书热线：010-68326294

010-88379203

网络服务

机工官网：www.cmpbook.com

机工官博：weibo.com/cmp1952

金书网：www.golden-book.com

教育服务网：www.cmpedu.com

封面无防伪标均为盗版



机械工业出版社微信公众号



E视界

传播电类内容提升专业知识



科技电眼

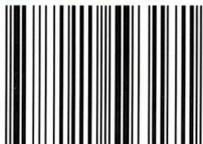
关注电类行业动态 聚焦前沿科技

上架指导 信息处理

ISBN 978-7-111-59317-1

策划编辑◎任鑫 / 封面设计◎马精明

ISBN 978-7-111-59317-1



9 787111 593171 >

定价：59.00元