

分类号 TP957.52

学号 13049052

U D C

密级 公开

工学博士学位论文

基于单演信号的 SAR 图像目标识别技术研究

博士生姓名 董刚刚

学科专业 通信与信息工程

研究方向 图形与图像处理技术

指导教师 匡纲要 教授

国防科学技术大学研究生院

二〇一六年九月

Study on Target Recognition in SAR Image via the Monogenic Signal

**Candidate: Ganggang Dong
Supervisor: Professor Gangyao Kuang**

A dissertation

**Submitted in partial fulfillment of the requirements
for the degree of Doctor of Engineering
in Information and Communication Engineering**

Graduate School of National University of Defense Technology

Changsha, Hunan, P. R. China

September 10, 2016

独 创 性 声 明

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

学位论文题目： 基于单演信号的 SAR 图像目标识别技术研究

学位论文作者签名： 董闪闪 日期： 2016 年 9 月 10 日

学 位 论 文 版 权 使用 授 权 书

本人完全了解国防科学技术大学有关保留、使用学位论文的规定。本人授权国防科学技术大学可以保留并向国家有关部门或机构送交论文的复印件和电子文档，允许论文被查阅和借阅；可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密学位论文在解密后适用本授权书。)

学位论文题目： 基于单演信号的 SAR 图像目标识别技术研究

学位论文作者签名： 董闪闪 日期： 2016 年 9 月 10 日

作者指导教师签名： 董闪闪 日期： 2016 年 9 月 10 日

目 录

摘要	i
ABSTRACT	iii
第一章 绪论	1
1.1 研究背景及意义	1
1.2 研究现状及面临的挑战	2
1.2.1 当前研究现状	5
1.2.2 面临的困难与挑战	15
1.2.3 主要关键技术	20
1.3 论文的主要工作及创新	31
1.3.1 论文的主要工作	31
1.3.2 论文的创新点	33
第二章 基于稀疏表征的 SAR 图像目标分类可行性研究	35
2.1 稀疏表示	36
2.1.1 线性空间以及线性空间的基	36
2.1.2 过完备冗余字典	38
2.2 压缩感知	44
2.2.1 工作原理	45
2.2.2 测量矩阵	47
2.2.3 重构算法	48
2.2.4 常用的优化工具箱	49
2.3 基于稀疏表示的分类学习	51
2.4 稀疏表示在 SAR 图像目标识别中的应用	54
2.4.1 可行性分析	54
2.4.2 改进方案	57
2.4.3 实验验证	66
2.5 本章小结	79
第三章 基于单演信号的 SAR 图像目标特性描述	81
3.1 Riesz 变换与单演信号	81
3.1.1 解析信号	81
3.1.2 Riesz 变换	82
3.1.3 单演信号	85

3.2	单演信号对 SAR 图像散射特性的变化分析	91
3.2.1	物理配置变化分析	92
3.2.2	姿态角变化分析	92
3.2.3	俯仰角变化分析	93
3.2.4	连接体变体分析	94
3.2.5	随机噪声污染分析	94
3.3	本章小结	96
第四章	基于单演信号信息融合的 SAR 图像目标分类学习	97
4.1	空域信息融合方案	97
4.1.1	特征级融合 (Feature-level Fusion)	98
4.1.2	决策级融合 (Decision-level Fusion)	99
4.1.3	多特征多任务联合稀疏表示	107
4.2	实验验证	112
4.2.1	标准工作条件	112
4.2.2	扩展工作条件	115
4.2.3	杂波样本拒识	117
4.3	本章小结	120
第五章	基于单演信号多核合成学习的 SAR 图像目标分类	121
5.1	再生核 Hilbert 空间	121
5.1.1	Mercer 核	122
5.1.2	再生核 Hilbert 空间	123
5.1.3	正定核函数的特性	126
5.2	Hilbert 核空间线性表示	126
5.3	单演信号 Hilbert 空间多核合成学习	129
5.3.1	“先组合再映射” —— 单演信号特征组合	132
5.3.2	“先映射再组合” —— 格莱姆矩阵组合	132
5.4	实验验证	133
5.4.1	标准工作条件	134
5.4.2	扩展工作条件	136
5.4.3	随机噪声污染	139
5.5	本章小结	140
第六章	基于单演信号流形学习的 SAR 图像目标分类	141
6.1	微分几何基础理论	141
6.1.1	黎曼流形	141
6.1.2	Stiefel 和 Grassmann 流形	147

6.2 单演信号尺度空间的流形学习	149
6.2.1 单演信号协方差矩阵构建的黎曼流形	149
6.2.2 单演信号系数矩阵构建的 Grassmann 流形	154
6.2.3 方向可控 Riesz 小波构建的 Grassmann 流形	159
6.3 实验验证	166
6.3.1 基础实验验证	167
6.3.2 扩展工作条件实验比较	173
6.4 本章小结	177
第七章 结束语	179
致谢	183
参考文献	185
作者在学期间取得的学术成果	207
附录 A 向量空间	209
附录 B Dempster-Shafer 证据理论	211
B.1 DS 数据融合	211
B.2 Bayes 数据融合	212
附录 C 线性回归与核回归分析	214

表 目 录

表 1.1 鉴别特征汇总表	5
表 1.2 SAIP 使用的鉴别特征	7
表 1.3 MSTAR 计划使用的特征参量	10
表 1.4 MSTAR 计划扩展工作条件	11
表 1.5 经典的相关滤波器方法	13
表 2.1 基于 DS 证据推理的“软”决策证据合并框架	66
表 2.2 多类目标成像样本总数	68
表 2.3 俯仰角变化的成像样本数	68
表 2.4 掩模尺寸对频域稀疏表示的影响	70
表 2.5 目标配置变化实验数据	70
表 2.6 标准实验条件频域分类学习	71
表 2.7 空间偏移对频域分类学习的影响	72
表 2.8 噪声污染对频域分类学习的影响	74
表 2.9 俯仰角变化对频域分类学习的影响	75
表 2.10 基于证据推理的“软”决策方法示例一	76
表 2.11 基于证据推理的“软”决策方法示例二	77
表 2.12 四类目标在不同俯仰角的视数	78
表 2.13 目标配置变化实现两种决策方法比较	78
表 2.14 配置变化实验两种决策方法比较	78
表 4.1 决策级融合的证据合并框架	106
表 4.2 贝叶斯估计与 DS 证据理论的决策融合例证	106
表 4.3 空域信息融合实验的基准比较算法	113
表 4.4 标准工作条件的空域信息融合比较	113
表 4.5 目标配置变化实验数据	115
表 4.6 目标配置变化对信息融合的影响	116
表 4.7 俯仰角变化对空域信息融合的影响	118
表 5.1 再生核 Hilbert 空间分类学习实验基准算法	134
表 5.2 噪声污染对多核合成学习的影响	139
表 6.1 常用的黎曼流形距离函数	144
表 6.2 流形学习实验的基准比较算法	167
表 6.3 三种协方差描述子计算模式	168
表 6.4 协方差矩阵描述子计算模式	169

表 6.5 Grassmann 流形距离函数比较	170
表 6.6 笛卡尔坐标表示和极坐标表示对分类学习的影响	170
表 6.7 Riesz 小波尺度对流形学习的影响	171
表 6.8 Riesz 变换阶数对流形学习的影响	172
表 6.9 流形学习受目标配置变化的影响	173
表 6.10 噪声污染对流形学习的影响	176
表 B.1 不同假设的证据质量分配	212
表 C.1 线性回归分析与核回归分析优劣对比	214

图 目 录

图 1.1 经典的 SAR ATR 三级处理流程示意图	2
图 1.2 双参数 CFAR 检测器示意图	3
图 1.3 目标鉴别环节剔除的自然杂波虚警	4
图 1.4 单类鉴别器算法原理	4
图 1.5 特征空间分类示意图	4
图 1.6 SAIP 半自动图像智能处理系统流程	7
图 1.7 模板匹配方法的模板库生成示例	8
图 1.8 MSTAR 计划基于模型视觉的目标识别方法简明流程	9
图 1.9 基于相关滤波的目标识别方法流程图	13
图 1.10 基于统计分布模型的目标识别方法简明流程	14
图 1.11 构成目标特性的典型散射现象	16
图 1.12 SAR 图像平面投影原理	17
图 1.13 SAR 图像像素相干叠加受成像几何变化的影响	17
图 1.14 俯仰角变化引起的图像目标特性变化	18
图 1.15 坦克部件配置引起的变体示意图	18
图 1.16 树木遮挡对目标成像的影响	19
图 1.17 目标与植被的重影对成像结果的影响	20
图 1.18 SAR 目标切片特征提取流程	24
图 1.19 最近邻分类示意图	25
图 1.20 线性不可分数据	26
图 1.21 支持向量机示意图	27
图 1.22 前向反馈神经网络示意图	29
图 1.23 多元线性回归模型的稀疏求解原理	30
图 1.24 特征的有效性示意图	31
图 1.25 论文的组织结构图	32
图 2.1 Brodatz 库的 Grass 纹理	42
图 2.2 MOD 算法与 KSVD 算法的平均表示误差	42
图 2.3 MOD 与 KSVD 字典学习	43
图 2.4 稀疏表示模型	45
图 2.5 广义压缩感知原理	46
图 2.6 不同范数约束的稀疏表示	49
图 2.7 理想情况 ℓ_1 范优化问题与噪声扰动 ℓ_1 范数优化问题的示意图	50

图 2.8 稀疏表示分类流程图	55
图 2.9 从流形的角度解释稀疏表示分类算法	56
图 2.10 频域稀疏表示分类流程图	58
图 2.11 SAR 图像及其傅立叶频谱	59
图 2.12 SAR 图像空间平移前后的傅立叶频谱对比	59
图 2.13 SAR 图像平移前后对角线区域幅度谱对比	60
图 2.14 SAR 图像平移前后对角线区域相位谱对比	60
图 2.15 频谱筛选前后的 SAR 图像对比	60
图 2.16 DS 证据理论的置信函数与似然函数	63
图 2.17 MSTAR 机动目标切片 SAR 图像	67
图 2.18 MSTAR 场景 SAR 图像	67
图 2.19 姿态角变化成像比较	68
图 2.20 目标配置变化成像比较	69
图 2.21 频谱筛选掩模示意图	70
图 2.22 图像的空间平移示意图	72
图 2.23 频域稀疏表示随图像空间偏移正确率曲线	73
图 2.24 图像随机噪声示意图	73
图 2.25 频域稀疏表示随噪声污染等级的正确率曲线	73
图 2.26 三类车辆目标在 4 种雷达俯仰角 SAR 图像	74
图 2.27 连接体变体图示	75
图 2.28 频域稀疏表示随俯仰角变化的正确率曲线	76
图 2.29 十类目标识别实验 DS “软” 决策与经典 “硬” 决策的比较	79
图 3.1 六种典型特征及其对应的相位	82
图 3.2 二维信号的两个方向 Riesz 变换核	84
图 3.3 单演信号极坐标变换	87
图 3.4 信号执行 φ 角度旋转后的相位表示	88
图 3.5 Log-Gabor 滤波器频率响应	89
图 3.6 三尺度 Log-Gabor 滤波器组	89
图 3.7 单演信号多尺度表示	91
图 3.8 单演信号多尺度表示实现目标物理配置不变描述	93
图 3.9 单演信号多尺度表示实现目标姿态角不变描述	93
图 3.10 单演信号多尺度表示实现雷达俯仰角的不变描述	94
图 3.11 单演信号多尺度表示实现连接体变体的不变描述 (BRDM2)	95
图 3.12 单演信号多尺度表示实现连接体变体的不变描述 (ZSU23/4)	95

图 3.13 单演信号多尺度表示实现随机噪声污染的稳健描述	95
图 4.1 特征级融合简明流程	99
图 4.2 决策级融合简明流程	101
图 4.3 贝叶斯决策融合示例	103
图 4.4 基本概率赋值函数曲线	105
图 4.5 多任务联合稀疏表示的目标识别方法示例	108
图 4.6 多任务学习示意图	109
图 4.7 多任务多特征联合稀疏表示的框架流程	111
图 4.8 多任务联合稀疏表示实例引证	111
图 4.9 单任务稀疏表示与多任务联合稀疏表示的比较	112
图 4.10 空域信息融合的十类目标识别实验结果	114
图 4.11 杂波样本示例	119
图 4.12 自然杂波拒识的 ROC 曲线	119
图 4.13 混淆目标拒识的 ROC 曲线	119
图 5.1 非线性映射引导的 Hilbert 空间原理	122
图 5.2 基于 Hilbert 空间核稀疏表示的分类流程	129
图 5.3 基于多核合成学习的分类学习	131
图 5.4 多核合成学习的分类示例	134
图 5.5 多核合成学习十类目标标准工作条件的识别实验	135
图 5.6 目标配置变化对多核合成学习的影响	137
图 5.7 俯仰角变化对多核合成学习的影响	138
图 5.8 多核合成学习受噪声污染的识别率变化曲线	140
图 6.1 黎曼流形任意点的切向空间	143
图 6.2 区域协方差描述子图例	146
图 6.3 主角与测地距离图示	148
图 6.4 单演信号协方差描述子生成示意	151
图 6.5 对数协方差矩阵稀疏表示流程图	152
图 6.6 对称正定矩阵嵌入到 Hilbert 空间的线性表示流程	154
图 6.7 单演信号分量矩阵的 Grassmann 流形测地距离	156
图 6.8 Grassmann 流形嵌入到 Hilbert 空间的分类示意图	159
图 6.9 第 5 阶 Riesz 变换的频率响应	161
图 6.10 基于 Riesz 小波框架的流形学习流程图	162
图 6.11 五阶三尺度 Riesz 小波变换示例	163
图 6.12 SAR 图像 Riesz 表示模型距离度量	164

图 6.13 基于 Riesz 小波框架的核稀疏表示流程图	167
图 6.14 流形学习受 Riesz 小波尺度的影响	172
图 6.15 流形学习受 Riesz 变换阶数的影响	172
图 6.16 俯仰角变化对流形学习的影响	175
图 6.17 流形学习受噪声污染的识别率变化曲线	177
图 C.1 线性岭回归与核岭回归示意图	214

摘要

合成孔径雷达是一种工作于微波波段的主动式传感器，它不受光照、气候等条件限制，可全天时、全天候对地观测；通过超宽带成像技术还可透过地表或植被捕获浅埋目标的有用信息。自动目标识别是合成孔径雷达图像解译的基础性课题，它在战场监视、防空反导、战略预警等方面发挥着十分重要的作用，因而能为战场态势感知和情报生成提供重要保障。自 20 世纪 80 年代以来，各国学者围绕着雷达自动目标识别开展了大量富有成效的研究工作。然而，已有的研究成果距离实际战场应用还有一定差距。合成孔径雷达是一种相干成像雷达，其图像对于目标姿态角、雷达俯仰角、装备物理配置、遮挡与连接体变体、噪声污染等因素十分敏感。这些因素的细微变化，也可能导致迥然各异的目标散射现象。为了应对这些敏感因素带来的不利影响，本文深入系统地分析前人开展的工作，通过总结已有研究成果、分析存在的问题和不足，提炼出自动目标识别系统工程的两大关键技术：**SAR 图像特征描述和分类学习**。然后分别从这两个方面入手，寻找改进方案，提高识别的稳健性和可靠性。在深入分析经典的特征描述方法基础上，论文引入了近年来提出的一种高维空间解析信号——单演信号，借助带通滤波器组构建原始信号的单演信号尺度空间，利用不同尺度空间分解生成的分量对原始 SAR 图像目标散射现象进行刻画描述，捕捉信号内部蕴含的稳健信息；在分类学习环节引入稀疏表示理论，与经典方法推崇的正交编码不同，我们采取基于过完备冗余字典的稀疏表示方案，首先利用给定样本构造过完备冗余字典，根据字典原子的线性组合表示未知信号，生成多元线性回归模型，借助最稀疏约束求取最优回归系数，根据回归系数计算重构误差，实现决策分类。论文的主要贡献包括：

1. 系统回顾了 SAR 图像自动目标识别的发展历程，对现有的识别方法进行总结归类，分析梳理各种方法的优点与不足，进而概括自动目标识别技术应用于实际战场需要解决的关键问题；**围绕着这些关键问题提出本文的研究主线：借助单演信号多尺度表示刻画 SAR 图像目标散射现象，捕捉不变信息；引入稀疏表示理论进行分类学习，实现目标类别的判决。**
2. 回顾了稀疏表示的发展历程，介绍相关背景知识，包括工作原理、前提条件及重构算法，重点阐述稀疏表示在模式识别中的应用现状；详细解释了稀疏表示在遥感图像解译中的应用可行性；结合 SAR 图像目标识别技术需要解决的关键问题，分别从**样本表示和决策规则**两个方面进行了尝试性改

进，提出了基于频域稀疏表示的分类学习和基于 DS 证据推理的“软”决策两种算法。

3. 针对扩展工作条件下的目标识别问题，引入了高维空间解析信号—单演信号，借助带通滤波器组构造单演信号尺度空间，利用不同尺度空间的系数分量实现对 SAR 图像的表示，刻画目标散射现象，设计特征描述向量，构造回归模型，实现稀疏表示分类。主要内容包括：借助数据融合、再生核 Hilbert 空间合成核学习和黎曼流形学习等方法，分别提出基于单演信号图像域信息融合的分类学习、基于单演信号 Hilbert 空间多核合成的分类学习、基于单演信号尺度空间流形学习的分类三种识别框架和思路，其中图像域信息融合包括单演信号尺度空间分量的特征级融合、决策级融合以及多特征多任务联合稀疏表示三种实现方案，再生核 Hilbert 空间合成核学习包括数据空间合成学习与 Hilbert 空间多核合成学习两种具体算法，黎曼流形学习包括对称正定矩阵空间的流形学习、单演信号尺度空间 Grassmann 流形学习和方向可控 Riesz 小波框架 Grassmann 流形学习三种具体方案。
4. 借助 MSTAR 目标切片实测数据设计了大量目标识别实验，按照先基础验证再实验比较的步骤，对所提出的方法的合理性做基础验证，并通过与经典算法的比较来检验算法的有效性和扩展能力。

关键词: 合成孔径雷达，目标识别，分类，单演信号，稀疏表示，压缩感知，过完备字典，Hilbert 空间，合成核训练，黎曼流形

ABSTRACT

Synthetic aperture radar (SAR) is an active sensor worked in the microwave band. It could work in 24-hour a day and all-weather, due to the ability not to be limited by the lighting and environment conditions. With the development of ultra-wideband imaging technique, SAR could penetrate the earth surface, vegetation and hence get the valuable information. Automatic target recognition (ATR) is a basic research topic for SAR image interpretation. It plays an important role in battlefield surveillance, anti-ballistic missile, air defense, strategic warning, *etc*, and hence achieves the perceive of battlefield situation and the generation of important intelligence. Though a great many works have been done previously, SAR ATR is still an open problem, especially for various extended operating conditions. As widely reported, target scattering phenomenology is susceptible to the variation of target pose, radar depression angle, configuration, articulation and occlusion, noise corruption, and hence results in the mutable characteristics of SAR image. It is therefore necessary to make full consideration on these issues. By deeply analyzing the preceding works on SAR target recognition, this dissertation comes to the conclusion that the crux of target recognition lies in two procedures, the representation of target scattering characteristics and the implementation of classification learning. To improve recognition performance, this dissertation introduces two recently developed techniques, the monogenic signal and sparse signal modeling. Specifically, the monogenic signal is developed to achieve the representation of target scattering phenomenology, while sparse signal modeling is modified to perform efficient classification. It is expected to promote target recognition performance by advanced learning skill and effective feature descriptor. To jointly consider the components of multi-resolution monogenic signal into the framework of sparse signal modeling, this dissertation presents three different frameworks, including information fusion in the spatial domain, composite kernel learning in Hilbert space, and Riemannian manifold learning. The contribution and novelty of this dissertation include:

- **First**, this dissertation comprehensively reviews the development of SAR ATR, especially for the famous three-stage processing framework. The current approaches to SAR ATR are categorized as template matching scheme, model-based vision, and the improved strategy. Both advantages and disadvantages of each algorithms have been summarized, followed by the generalization of the crucial issues. To

solve these issues, the dissertation introduces two recently developed techniques, sparse signal modeling and the monogenic signal. The former is used to improve the accuracy of classification learning, while the latter is employed to capture the characteristics of target scattering phenomenology.

- **Second**, this dissertation provides the development of sparse signal modeling, as well as compressed sensing. The related background knowledges include the work mechanism, the prerequisites, and the reconstruction algorithms. Afterward, this dissertation gives the applicability of sparse signal modeling in SAR image interpretation, including image super-resolution, image recovery, speckle reduction, image compression, terrain classification. To adopt sparse signal modeling for target recognition in SAR image, this dissertation proposes two modified schemes, *i.e.*, performing sparse representation via Fourier spectrum descriptor and reaching the inference via Dempster-Shafer's theory of evidence.

Improvement of signal representation via Fourier transform. To deal with signal alignment and noise corruption, this dissertation proposes to implement sparse representation by Fourier spectrum bands of signal. Since signal energy in the frequency domain concentrates in a small portion of frequency bands, they are discriminative and hence used to represent the original signal. The proposed scheme is then used to feed into the framework of sparse signal modeling. The decision is made according to the characteristics of representation on reconstruction.

Improvement of decision rule via Dempster-Shafer's theory of evidence. To improve the reliability of decision rule, this dissertation recommends a soft decision via Dempster-Shafer's theory of evidence. The imprecision on uncertainty measurement is modeled by sample-wise ambiguity and the class-wise ambiguity during the quantification of probability mass. Various pieces of evidence derived from the candidate samples are pooled by means of Dempster's rule of combination, from which an inference can be reached.

- **Third**, to cope with target recognition under the extended operating conditions, the monogenic signal is proposed. Log-Gabor filter bank is employed to implement

multi-resolution monogenic signal. To improve recognition accuracy, this dissertation jointly considers the components derived from multi-resolution monogenic signal into the framework of sparse signal modeling, by which three different strategies are developed.

The 1st Strategy—Spatial Information Aggregation To aggregate various information derived from the monogenic signal, this dissertation designs a novel framework by performing information fusion in the spatial domain. According to the presented thought, three implementation schemes, *feature-level fusion*, *score-level fusion*, and *multi-feature & multi-task joint sparse representation* are recommended.

The 2nd Strategy—Composite Kernel Learning To deal with the dataset whose target classes are not linearly separable, this dissertation proposes to achieve classification by composite kernel learning. According to this thought, two implementation algorithms, *stacked features* and *summation kernels* are developed. These algorithms reflect the thoughts of mapping followed by feature combination, and kernel combination followed by mapping.

The 3rd Strategy—Riemannian Manifold Learning Recent studies show that great advantages can be achieved by considering the batch of signals with non-Euclidean geometry. Therefore, this dissertation proposes to perform target recognition with manifold learning skill. By applying the monogenic signal into the framework of manifold learning, three strategies, *Riemannian manifold learning on symmetry positive definite matrix space*, *Grassmann manifold learning via the components of multi-resolution monogenic signal*, and *Grassmann manifold learning via steerable Riesz wavelet frames* are developed.

- **Fourth**, multiple comparative experiments have been performed on MSTAR SAR database to validate the proposed algorithms.

Key Words: Synthetic Aperture Radar, target recognition, Classification, the monogenic signal, sparse representation, compressed sensing, over-complete dictionary, Reproducing Kernel Hilbert Space, Riemannian manifold

缩略语

SAR	\Leftrightarrow	Synthetic Aperture Radar	\Leftrightarrow	合成孔径雷达
ATR	\Leftrightarrow	Automatic Target Recognition	\Leftrightarrow	自动目标识别
SOC	\Leftrightarrow	Standard Operating Conditions	\Leftrightarrow	标准工作条件
EOC	\Leftrightarrow	Extended Operating Conditions	\Leftrightarrow	扩展工作条件
CFAR	\Leftrightarrow	Constant False Alarm Rate	\Leftrightarrow	恒虚警检测器
ROI	\Leftrightarrow	Region of Interest	\Leftrightarrow	感兴趣区域
MSTAR	\Leftrightarrow	Moving and Stationary Target Acquisition and Recognition		
PDF	\Leftrightarrow	Probability Density Function	\Leftrightarrow	概率密度函数
MAP	\Leftrightarrow	Maximum a Posterior Probability	\Leftrightarrow	最大后验概率
ROC	\Leftrightarrow	Receiver Operating Characteristics	\Leftrightarrow	接收机工作特性
RKHS	\Leftrightarrow	Reproducing Kernel Hilbert Space	\Leftrightarrow	再生核希尔伯特空间
PCA	\Leftrightarrow	Principal Component Analysis	\Leftrightarrow	主成分分析
ICA	\Leftrightarrow	Independent Component Analysis	\Leftrightarrow	独立成分分析
NMF	\Leftrightarrow	Nonnegative Matrix Factorization	\Leftrightarrow	非负矩阵分解
LPP	\Leftrightarrow	Locality Preserving Projections	\Leftrightarrow	局部保持投影
SPP	\Leftrightarrow	Sparse Preserving Projections	\Leftrightarrow	稀疏保持投影
LDA	\Leftrightarrow	Linear Discriminant Analysis	\Leftrightarrow	线性判决分析
SVD	\Leftrightarrow	Singular Value Decomposition	\Leftrightarrow	矩阵奇异值分解
SCR	\Leftrightarrow	Signal-to-Clutter Ratio	\Leftrightarrow	信杂比
SNR	\Leftrightarrow	Signal-to-Noise Ratio	\Leftrightarrow	信噪比
SVM	\Leftrightarrow	Support Vector Machine	\Leftrightarrow	支持向量机
SRC	\Leftrightarrow	Sparse Representation-based Classification		
BP	\Leftrightarrow	Basis Pursuit	\Leftrightarrow	基追踪
OMP	\Leftrightarrow	Orthogonal Matching Pursuit	\Leftrightarrow	正交匹配追踪
LASSO	\Leftrightarrow	Least Absolute Shrinkage and Selection Operator		
BPA	\Leftrightarrow	Basic Probability Assignment	\Leftrightarrow	基本概率赋值

第一章 绪论

1.1 研究背景及意义

二十一世纪以来，以美军为主导的西方军事强国先后发起了多场高技术条件下的局部战争，南联盟、伊拉克、叙利亚等国先后成为他们新武器的试验场。纵观这几场不对称的局部战争，共同之处都是围绕着制信息权展开。可见，现今的战争比拼的焦点已经悄悄转变为对信息权的争夺，其关键是如何实现从侦察单元（传感器）到火力打击单元的点对点实时传递与无缝链接，也就是军事专家所提出的发现即摧毁、查打一体化。可能大家对于基地组织多个头目在阿富汗、也门等地被美军捕食者无人机击毙的画面仍记忆犹新。整个事件的概略过程为：首先通过太空中的星载传感器获取有价值的目标信息，通过数据链将其传输给地面情报人员，由情报人员将其分发至最近的火力打击单元，瞬间即可完成对目标的攻击。类似的新闻报道还有很多，这些战例对军事专家及普通民众都有很强的启发意义。

近年来，我军的信息化建设已取得了长足进步，但与欧美发达国家相比，仍相形见绌，特别是信息获取与实时处理方面，差距仍然比较明显。究其原因，一方面是我军信息获取手段单一，已建成的传感器平台信息捕捉能力有限；二是信息辨别、传递、分发等环节的技术储备仍然比较滞后。因此，信息化水平已经成为制约我军战斗力提升的“瓶颈”，是未来若干年亟待解决的难题。

合成孔径雷达（Synthetic Aperture Radar, SAR）是一种工作于微波波段的主动式传感器，它可搭载于空中移动平台对较大范围内的区域进行侦察和监视。其主要优点可归纳为[1]：

- 通过综合孔径和脉冲压缩技术获取类似于光学摄影的高分辨率雷达图像；
- 成像不受气候、时间的影响，可全天候成像；
- 雷达图像分辨率与波长、载机飞行高度、雷达作用距离无关；
- 测绘带较少受载机航迹远近的影响；
- 通过超宽带成像技术，可对浅埋于地表之下以及隐藏于丛林叶簇之内的目标进行探测与挖掘。

因其本身的诸多优良特性，合成孔径雷达已经成为美国、以色列、德国等科技大国的一种重要军事侦察手段，在实际战场发挥着不可替代的作用。而国内研究目前尚处于起步阶段，中科院遥感与数字地球研究所、中国电科十四所和三十

八所、国防科技大学、西安电子科技大学等科研院所开展了大量的研究工作，但无论是传感器平台的信息采集技术或是图像的实时解译能力都无法满足我军信息化建设的需求。作为其中的一项基础性研究课题，合成孔径雷达自动目标识别（Automatic Target Recognition, ATR）就是在这种背景下应运而生。自动目标识别是指从获取的战场 SAR 图像快速提取和定位有重大军事意义的战术目标（如坦克、装甲车等），根据目标和周围环境的回波信号，设计能有效刻画目标散射特性的特征描述方法，区分不同类型的战术目标，判定目标属性、类别甚至型号，捕捉有用的情报信息 [2, 3]。然后根据目标属性及重要程度将有关信息分发到相应的处理单元。目前的 SAR 图像目标识别技术所涵盖的范围较广，如目标属性的识别、目标类别的识别、目标型号的识别、机载武器挂载情况的判断，地面机动目标的识别、海上舰船目标的识别、空中飞行物的识别等，这些课题对于战场监视、防空反导和战略预警都具有十分重要的作用。因而，自动目标识别技术对于我军的信息化建设不仅具有重大的理论研究意义，更具有广阔的应用前景和实用价值 [2–5]。

1.2 研究现状及面临的挑战

合成孔径雷达自动目标识别（SAR ATR）目的是从战场 SAR 图像中找出感兴趣目标，迅速判定其类别属性 [2, 3]，因而是一项系统性工程。在早期的研究中，西方学者将 SAR ATR 划分为三个步骤（如图 1.1 所示）：检测（Detection）、鉴别（Discrimination）和分类（Classification） [6–9]。

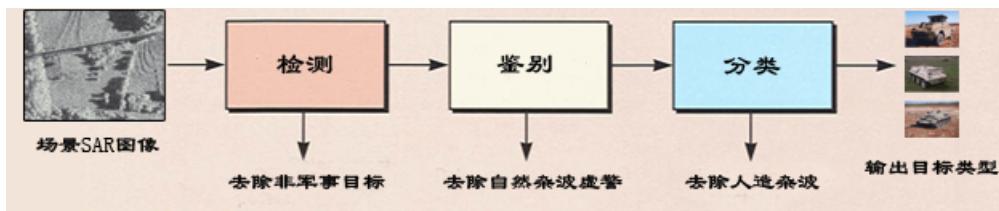


图 1.1 经典的 SAR ATR 三级处理流程示意图。

检测阶段用于挖掘与概略定位感兴趣区域（Region of interest, ROI），即潜在的重要军事目标。由于自然植被与人造物体的雷达回波其统计特性存在着明显差异，研究人员设计了一种基于高斯统计分布的双参数恒虚警率检测器（Constant False Alarm Rate, CFAR） [10–13]。在给定虚警率的前提下，通过比对待测区域与周边区域的广义似然比搜索 SAR 图像的感兴趣区域，将感兴趣区域的位置与相关信息传递给后续处理步骤即鉴别，等待执行下一个处理环节。也有学者通过引入其它统计模型改进双参数 CFAR 检测，如 Weibull 分布、K 分布、Rayleigh 分布等。恒虚警率检测器工作示意图如图 1.2 所示，首先对背景区域和待测区域分别进行统计建模，根据背景杂波区的样本数据估计均值和方差等统计量，为了消除背景

区域统计模型参数估计精度对于待测的目标区域的影响，在背景区之内又分别设定警戒区和目标区，其中目标区由被检测的样本数据组成，然后将估算的检验统计量与预先设定的阈值（根据虚警率计算得到）进行比较，推断被检测样本是目标或杂波，其最终输出 ROI 的概略位置。

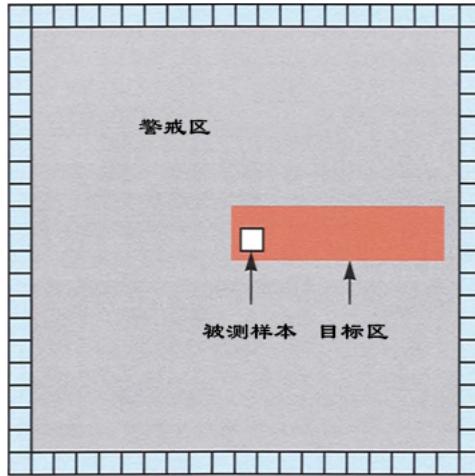


图 1.2 双参数 CFAR 检测器示意图。

鉴别阶段主要实现自然杂波（如树木、水塘、灌木丛等，如图 1.3 所示）的筛选与剔除。林肯实验室设计的鉴别算法实质是单类二次距离鉴别器（One-class Quadratic Discriminator）[14]。假设自然杂波虚警数据集在特征空间存在许多未知的属性，故仅使用目标集进行学习。如图 1.4 所示，对于检测出的每个未知 ROI 样本，鉴别器首先计算 ROI 样本与目标类别在特征空间的二次距离，其中目标类别的特征向量根据训练样本计算得到，如果目标类别的特征向量选择合理，则疑似目标类的 ROI 样本将会得出较小的二次距离，同样，虚警类 ROI 样本将会生成较大的二次距离，从而剔除自然杂波虚警 [15-17]。显然，鉴别算法的关键是所用于计算分离度的鉴别特征，早期林肯实验室设计的鉴别特征包括标准差（Standard-Deviation Feature）、分形维（Fractal-Dimension Feature）、权秩填充比（Weighted-Fill Feature），后来的学者又陆续提出了其它一些鉴别特征，表 1.2 对常用的鉴别特征进行了归纳和总结 [15, 18]。鉴别阶段最终输出潜在目标的精确位置。

鉴别后保留下来的未知 ROI 样本将传递给分类器，筛除人造虚警，辨识剩余样本类别。MSTAR 计划涵盖了目前为止最全的样本类别，包括装甲车、主战坦克、防空导弹、军用卡车等，分类的主要任务就是为保留下来的每个 ROI 切片合理地分配类别标记。在执行分类前通常先设计特征描述子来表示未知 ROI 与给定的训练样本，与鉴别阶段所设计的特征不同，分类阶段设计的特征主要用于区分不同类型的人造目标，如坦克与装甲车，火炮与卡车、自行火炮与反坦克发射装

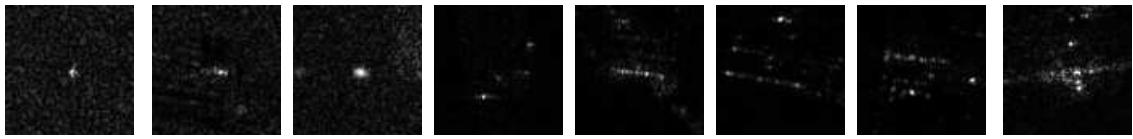


图 1.3 目标鉴别环节剔除的自然杂波虚警，图中列出一组 MSTAR 数据自然杂波虚警，从左到右，虚警与目标的相似程度逐渐提高。

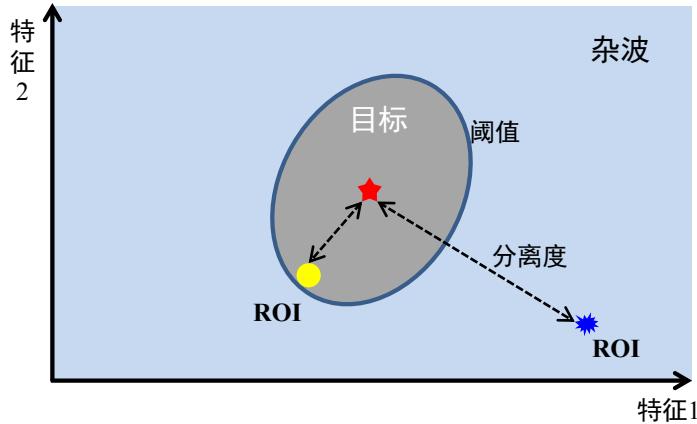


图 1.4 单类鉴别器算法原理，图示的特征空间，灰色椭圆表示目标区域，淡蓝色背景为杂波区域，红星五角星表示目标的训练样本生成的特征向量，黄色圆圈和蓝色不规则形状是未知 ROI 的特征向量，ROI 与目标向量间的距离是二者的分离度（二次距离）。图中黄色圆圈与目标向量的分离度小于边界阈值，判定它是目标，予以保留；蓝色不规则形状与目标向量的分离度大于边界阈值，判定它为虚警，予以剔除。

置等，其实质是设计一个抽象的特征空间，使得未知 ROI 样本与其真实属性类别中心的空间距离较小，而与其它类别中心的空间距离较大，然后在特征空间中进行分类学习，去除人造杂波虚警（如道路、路灯、建筑物等），判定目标类别，如图 1.5 所示，分类学习的主要任务就是在特征空间中寻找不同类别之间的决策边界。这也是本文所关注和研究的重点。

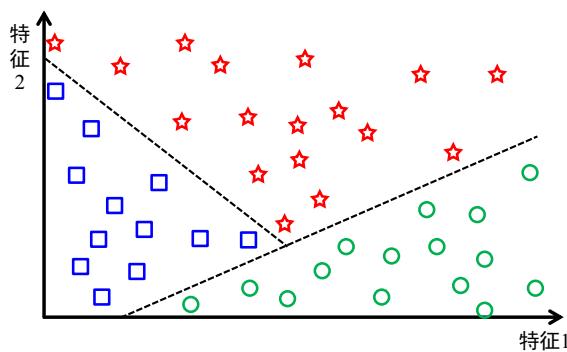


图 1.5 特征空间分类示意图，图中方块、圆圈、五角星表示特征空间的不同类别，分类的是寻找黑色虚线表示的类别间的决策边界。

表 1.1 鉴别特征汇总表。

开发者	鉴别特征
林肯实验室	标准差、分形维、权秩填充比 [9]
ERIM ^a	尺寸特征：质量、直径、归一化转动惯量 对比度特征：MaxCFAR, MeanCFAR, Percent Bright CFAR [15, 19] 极化特征：Percent Pure, Percent Pure Even, Percent Bright Even
Rockwell	尺寸、形状、方向角、极化相位比率、特定熵 [15]
Loral ^b	间隙度 [15]
RSIP ^c	主轴转动惯量、面积直径比、均值和峰值信噪比等 [18, 20, 21] 投影特征：水平投影、垂直投影、主轴投影、辅轴投影
Bir Bhanue	统计特征：98 分位像素数 [22]
Yingqiang Lin	距离特征：最短距离、最长距离、平均距离 矩特征：水平 μ_{20} 、垂直 μ_{02} 和对角线中心矩 μ_{22}
Jong-II Park 等	投影长度、投影误差、频域投影能量、目标像素数 [23]
其它	空间离散变化 [24]

^aEnvironmental Research Institute of Michigan, 密歇根环境研究所。

^bLoral Defense Corporation, 主要为美国海军开发雷达、红外、声呐检测算法的承包商。

^c国防科技大学遥感信息处理实验室, Lab of Remote Sensing Information Processing。

以美国为代表的西方国家对自动目标识别的研究起步较早, 技术比较成熟, 目前已经研制出多个 SAR ATR 系统, 比较有代表性的包括: 林肯实验室的半自动图像智能处理系统 (Semi-Automated Image Intelligence Processing, 即 SAIP 系统)、美国空军实验室的 ADTS (Advanced Detection Technology System) 系统、美国陆军实验室的 SAR ATR 系统、美国 Sandia 国家实验室的 MSTAR 计划、美国怀特 - 彼特森空军基地的 SHARP (System-Oriented High Range Resolution Automatic Recognition Program) 计划、美国 Sandia 国家实验室的 SAR 目标识别与定位系统 (SAR Target Recognition and Location System, STARLOS)、北约联合研制的 RG20 项目等 [25–35]。

1.2.1 当前研究现状

围绕着 SAR ATR 应用中面临实际困难, 研究人员开展了大量有针对性的研究工作, 推出了许多面向不同应用背景自动目标识别系统, 追本溯源, 这些识别系统的核心框架可归结为模板匹配 (Template-matching) 和模型视觉 (Model-based Vision) 两大类。之后, 通过借鉴模式识别和机器学习等领域的研究成果, 学者们也提出了许多改进方法, 下文将对已有的自动目标识别方法进行梳理和回顾。

1.2.1.1 基础方法

- 基于模板匹配的目标识别

基于模板匹配的目标识别方法由离线生成模板和在线匹配识别两部分组成 [36–39]。首先根据特定准则去除非潜在目标即自然杂波虚警，根据战场 SAR 图像生成 ROI，并粗略估算 ROI 的相关状态参量，将状态参量输入至分类器，依据某种概率准则寻找未知 ROI 与模板库参考样本间的最佳匹配，实现目标识别。常用概率准则有最小均方误差、最小欧氏距离、最大相关系数等。影响模板匹配识别精度的因素较多，其中比较重要的如模板库的构成、图像分辨率、未知样本与模板库样本间的距离或相似度定义等。

模板匹配的识别方法其经典应用是林肯实验室研制开发的 SAIP 系统 [40]，他们首先提出了一种 HDVI (High Definition Vector Imaging) [41] 超分辨率成像方法，对原始图像进行超分辨操作，然后再将超分辨的结果输入到均方误差模式匹配分类器进行比对识别，从而扩宽了模板匹配的适用范围，使得中低分辨率 SAR 图像也可执行目标识别。SAIP 系统的核心是自动目标识别框架，如图 1.6 所示，首先使用 CFAR 检测器执行目标检测，生成未知 ROI 样本，然后对生成的 ROI 进行鉴别，剔除自然杂波虚警，SAIP 计划中使用的鉴别特征包括林肯实验室早期设计的鉴别特征和密歇根环境研究所设计的鉴别特征 (Environmental Research Institute of Michigan, ERIM)。其中 ERIM 设计的特征主要包括三大类：尺寸特征、极化特征（仅针对全极化成像数据）和对比度特征 [19]，详见表 1.2。鉴别流程可简述为：先从四类十二种特征中选出一部分特征构成一个鉴别统计量，该统计量实质是检测出的 ROI 样本与目标类别之间的二次距离度量 (Quadratic distance metric)，

$$d(\mathbf{X}) = \frac{1}{n} (\mathbf{X} - \mathbf{M})^T \Sigma^{-1} (\mathbf{X} - \mathbf{M})$$

其中 \mathbf{X} 是未知 ROI 切片的特征向量， n 是选出的特征个数， \mathbf{M}, Σ 分别是根据训练样本所估算出目标的特征向量的均值和协方差矩阵；实际的人造军事目标其二次统计距离很小，大部分的自然杂波虚警其二次距离相对较大，因而可根据设计的统计距离滤除自然杂波虚警，剩余的那些与军事目标非常相似的人造虚警将会在分类阶段进行处理。

经鉴别后保留下来的 ROI 样本将会最终输送到一个均方误差模式匹配分类器 (Mean-Squared-Error Pattern-Matching Classifier)，其原理为：首先根据切片中的背景区域数据计算均值杂波水平 (Mean clutter level)，使用均值杂波水平对输入的 ROI 样本进行归一化处理，然后逐一计算未知 ROI 样本与模板库中的所有参考样本的均方误差，误差最小的视为最佳匹配。在具体的计算过程中，为了补偿目标质心的微小估算误差，通常会将参考样本沿距离向和方位向在一定范围内进行滑

动，在所有滑动情况下均方误差的最小值被视为未知样本与当前参考样本的均方误差。如果未知 ROI 与所有的参考样本的均方误差都大于某个容许值，则认为该样本为杂波虚警，即实现对人造虚警的剔除；相反，就将最佳匹配的模板的类别赋予给未知 ROI。若均方误差数值越小，则分类结果越可靠，因而，最小均方误差也可以作为分类置信水平的一个评估标准。

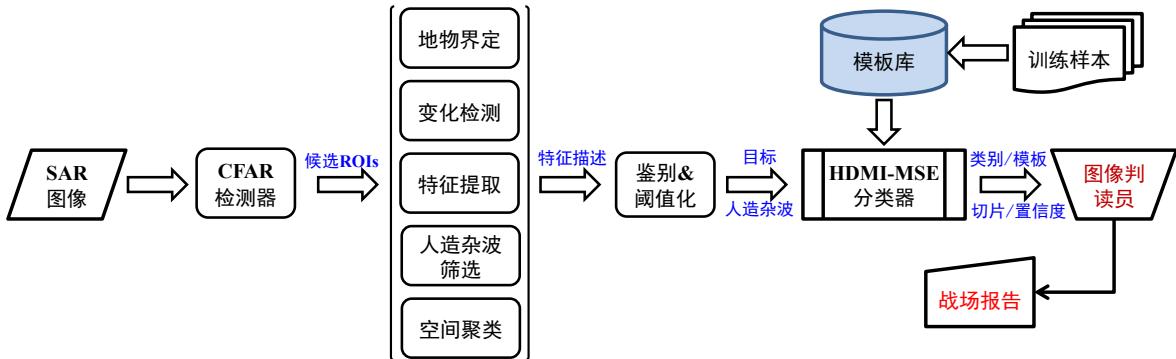


图 1.6 SAIP 半自动图像智能处理系统流程。

表 1.2 SAIP 使用的鉴别特征。

纹理特征	尺寸特征	对比度特征	极化特征
标准差	质量	峰值 CFAR	Percent pure odd
分形维	直径	均值 CFAR	Percent pure even
权秩填充比	归一化转动惯量	Percent bright CFAR	Percent bright even

由于 SAR 图像目标散射特性对姿态角变化的敏感性，兼顾计算效率方面的考虑，SAIP 计划以相邻 5 幅图像（以 1° 方位角为步长递增）数据的均值来为每类目标生成模板库，即每类 360 度全方位成像数据会生成 72 组均值模板（如图 1.7 所示），这些均值模板将会作为该类目标的代表，逐一与未知 ROI 样本进行比对，最后根据所有类别的均值模板与未知样本的比对结果预判目标类别。早期的研究主要以 3 类标准目标为主进行实验，后来逐步扩展到 10 类、20 类 [42]。研究表明：即使是同一类型目标，若物理配置、姿态角、俯仰角等因素发生变化，其 SAR 图像的散射特性将会发生剧烈的变化，而目标特性的动态变化，降低了目标类别的类内聚合性和类间差异性，从而加剧了 SAR 图像目标识别的复杂度。

影响 SAR 图像目标散射特性的因素比较复杂，加上实际战场目标本身物理配置的变化多样性、周围环境的偶然性等，都会形成 SAR 成像结果的千差万别。若要保证目标识别的可靠性，通常需要预想大量不同情况的工作条件，然后再根据

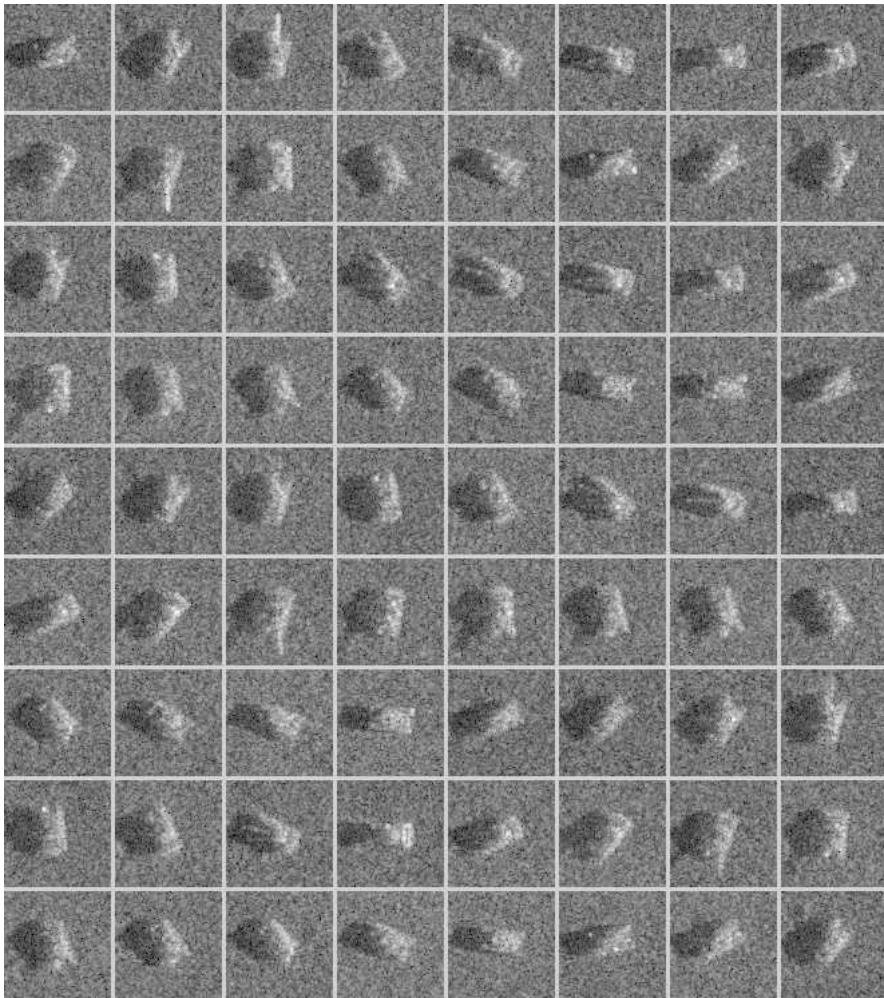


图 1.7 模板匹配方法的模板库生成示例，以 T72 坦克为例，使用俯仰角为 17° 每间隔 5° 方位角的序列号为 SN_132 的成像数据构建模板库，共 72 组参考样本（对原始成像数据执行对数变换增强对比度）。

这些工作条件下的成像数据来生成模板库的参考样本，这无论是在实现难度或是计算效率方面都是个很大的挑战。

- 基于模型视觉的目标识别

为了解决模板匹配中所遇到的实际问题，扩展其适用范围，研究人员提出了另外一种更加复杂的方案——基于模型视觉的目标识别方法 [27, 29, 43–47]。该方法实质是一种特征级的匹配识别。其核心思想概括为：首先从战场 SAR 图像中提取 ROI，方法与模板匹配的 CFAR 检测类似，然后估算未知 ROI 样本的特征属性及状态参量，凭借特征参量预估出许多与之相关的候选目标，然后根据雷达感知的过程原理和计算机辅助设计的 CAD 候选目标模型的几何特性预测并生成与未知目标相同的特征集，将候选目标与未知目标的特征进行比对，判决目标类型。

其实质上就是不断重复“假设—检验”这个过程，直至寻找到未知 ROI 样本的最佳匹配类型。候选假设成立与否是通过不断积累起来的匹配得分进行判断，即获得最大匹配得分的假设将会视为成立，从而得出未知 ROI 的类型。模型视觉采用在线类型预测的方法，它仅受限于 CAD 模型库的模型生成能力，因而，模型视觉能够处理因俯仰角和入射角改变所引发的目标几何变形或局部遮挡隐藏，与之不同，模板匹配则受限于模板库中所存储的参考样本。其次，模型视觉的分步推理策略（Coarse-to-fine）能够应付实际战场环境中许多无法预想的组合变化，如一类目标的多个不同配置变体。再者，模型视觉在识别决策的核心环节中采用了渐进式的证据累积策略，因而可从一定程度上容许侦察图像中关于目标的不完整或是缺陷知识。

基于模型视觉的识别方法最为典型的应用是由美国 Sandia 国家实验室等多家单联合推出的著名的 MSTAR 计划 [27, 29]，其核心的自动目标识别算法概要流程如图 1.8 所示。识别系统的主要输入是覆盖大幅地域的 SAR 图像以及相应的传感器、图像形式和几何参数。首先，利用对焦（Focus of Attention, FOA）操作消除场景 SAR 图像中不太可能包含目标的区域，类似于前文中的检测，其输出未知 ROIs；然后将每个 ROI 传递给索引模块进行处理，开辟候选目标的假设空间。候选目标假设包括目标类别（如 T72, BMP2）、姿态（位置和方位角）、连接状态（如炮塔方位、火炮位置等）以及可能的局部隐藏（遮挡、重叠）等内容。每个 ROI 都会有多个候选目标假设条目与其对应。聚焦和索引模块生成的候选目标假设连同 ROI 将会一起传递给 PEMS 循环（Predict, Extract, Match and Search），执行“假设—检验”处理，压缩假设空间，最终得出结果。

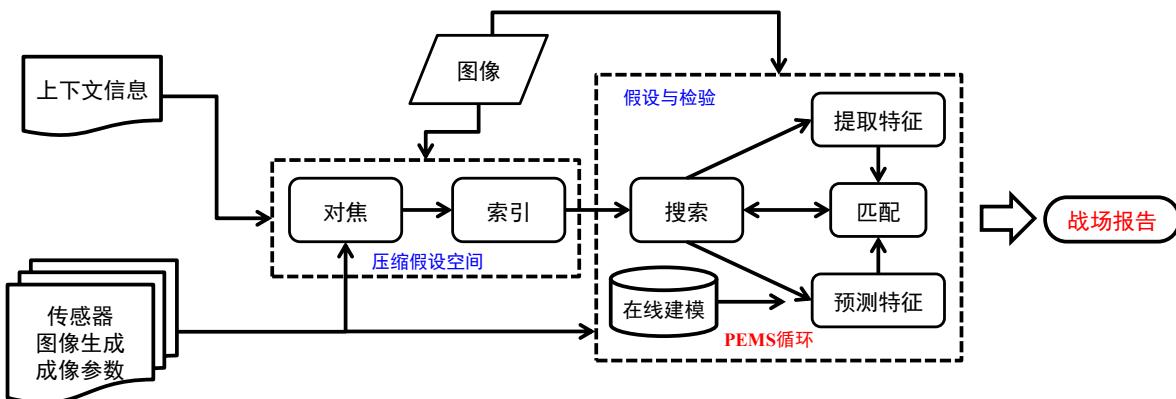


图 1.8 MSTAR 计划基于模型视觉的目标识别方法的简明流程。

PEMS 循环通过分析给定 ROI 及其候选目标假设推断出未知样本最可能的类别 [44, 48]。搜索模块（Search module）通过执行证据累积和假设推理过程来控制整个 PEMS 循环。针对每一条候选假设，搜索模块都会触发 MSTAR 的特征预测

模块 (MSTAR feature Prediction Module, MPM) 的特征提取模块 (Extract Module), 分别生成一组预测的目标特征和实时提取的目标特征 [43, 47]。再将预测和提取的目标特征传递给匹配模块, 评估它们之间的一致性, 将返回的匹配度分值再传递给搜索环节 [45, 46], 进行下一次循环处理。

需要注意的是模型视觉的识别方法中模拟信号特征完全可以不与灰度级的特征一致, 它与预测模块生成的特征有很大地区别。模拟特征其目的是为了实现模拟器输出与测量数据样本之间的像素级一致性, 而预测模块生成的特征和不确定性度量则是用于 PEMS 循环中的渐进式推理环节 [43]。表 1.3 列出了不同阶段执行搜索模块所需要的一些特征参量。

表 1.3 MSTAR 计划使用的特征参量。

特征类型	描述	不确定性测度
提取特征	维度 / 比率, 矩	统计量
显著点	位置, 幅度 / 相位	协方差, 概率密度函数
区域	掩模 / 像素列	边界位置协方差, 元概率密度函数
阴影	掩模 / 像素列, 噪声点	边界位置协方差, 元概率密度函数
纹理	分形维	元统计量 / 概率密度函数
全灰度	掩模 / 像素列	元幅度概率密度函数

MSTAR 的 PEMS 循环推理流程是通过嵌在搜索模块中的子程序来控制, 它能针对特定的 ROI 样本调整特定的假设空间, 所以无法预先进行设计, 这就给 MPM 模块的设计提出了两个相互矛盾的要求: 一是 MPM 的设计应该尽量容纳可能的假设条目或特征集, 即最大在线弹性; 二是为了在 PEMS 循环的每次迭代中能够生成大量的潜在候选假设, MPM 必须具备很高的在线执行效率。此外, 为了保证最终决策的可靠性, 维持电磁散射模型的高保真度也是 MPM 模块设计的一个核心理念之一。

通过对比模板匹配与模型视觉两类方法, 可以看出它们的前端聚焦模块功能大致相同, 不同之处在于后端 PEMS 回路。SAR ATR 是逆向问题, 为了能准确测量未知目标的真实状态, 构建有针对性的假设空间, 需要构建高精度的 3 维 CAD (Computed Aided Design) 模型, 并使用高频电磁仿真软件 (如 CST) 预测目标的姿态、装备配置、以及相应的特征集 [44, 48]。然而, 在实际应用中, 逼真地描述目标的测量特性是很困难的, 而忽略或是简化物体与周围环境之间的相互作用本身就会造成很大估算误差, 因而基于模型视觉的识别方法实现难度远远高于基于模板匹配的识别方法。为了充分利用目标相关的信息, 减少图像分析员的工作量,

提高目标识别系统的工作效率，也有学者提出将两种方法结合起来，尽可能地减少人工干预 [49]。

实际战场遇到的目标识别问题，远比理论建模复杂很多。为了与理想实验条件相区别，着眼于实际战场需要，E. Keydel 等学者提出扩展工作条件下 (Extended Operating Conditions, EOC) 的目标识别 [27]，内容包括入射角变化、目标连接体和物理配置变化、遮挡和重叠引起的局部隐藏、类内目标的易变性等，将其归纳起来主要有目标本身变化、周围遮挡伪装、图像获取参数等，表 1.4 列出了 MSTAR 中目标状态和获取参数方面的变化内容，显然，这些现实问题的提出无疑为自动目标识别从理论研究走向实际应用指明了方向，也为后来的学者们提供了很好地研究思路。

表 1.4 MSTAR 计划扩展工作条件

扩展工作条件	具体化实例
目标类别数目	3 类、10 类 \mapsto 20 类目标
图像获取参数	俯仰角 $\Rightarrow 40^\circ$ 、入射角 $\Rightarrow \pm 35^\circ$ 、方位角 $\Rightarrow 360^\circ$
目标状态变化	连接体变体、选择性配置、类内易变性
局部散射变化	遮挡重叠引起的模糊、伪装、欺骗

1.2.1.2 改进方法

在 SAIP 和 MSTAR 等科研计划的引领下，中外各国的研究人员纷纷开展了大量相关的研究。许多学者通过借鉴模式识别和机器学习等领域的研究成果，对模板匹配和模型视觉两类处理框架的子环节进行了优化与改进，其中检测环节的改进集中在背景杂波的统计建模方面的研究，例如使用 K 分布、 G 分布、 G^0 分布、Rician 分布等 [3, 50–57] 统计模型对背景区域进行建模，在保持虚警率恒定的前提下提高目标检测率；目标鉴别的改进则集中在设计新的鉴别特征，强化自然杂波与人造目标的区分能力，从而提高鉴别准确率，提出的特征有扩展分形维、均值信噪比、峰值信噪比、矩特征、形状特征、目标投影长度、目标统计特征等 [23, 24, 58]；分类环节的改进则集中在图像预处理（如噪声抑制、目标分割）、构造新特征（如变换域特征 [59, 60]、统计特征、散射中心等 [61, 62]）和设计分类器（如支持向量机、神经网络、贝叶斯分类）[63–67]。归纳起来，常见的改进方法有以下几类：

- **基于相关滤波的识别方法 (Correlation Pattern Recognition)**。基于模板匹配的识别方法主要由离线生成模板与在线匹配识别两部分组成，模板库是由特定方位角区间内固定步长的训练样本（如 SAIP 计划是以每 5° 方位角）所生成

的均值模板构成，即模板库中的每幅参考样本就是某个方位角区间的样本均值，然后将未知目标 ROI 切片与模板库中的参考样本逐一进行比对，以统计的方法衡量它们之间的差异性（如均方误差），实现目标分类识别。由于未知 ROI 与参考样本在空域（或者某个特征域）进行比对，因而 ROI 切片必须与参考样本在空间上进行精确配准，否则，空域的像素偏移会显著地影响分类精度，这无疑制约了模板匹配的适用范围。为了解决 ROI 与模板库样本的空间配准问题，构造空域平移不变识别算法，Abhijit Mahalanobis 和 Bhagavatula Vijaya Kumar 等人提出了相关滤波器的目标识别思路 [68–72]。相关滤波器的核心思想是利用二维傅立叶变换系数的优良性质实现分类识别的旋转不变性、平移不变性、扭曲不变性等。如图 1.9 所示，首先对给定的所有类别训练样本执行二维傅立叶变换，分别取出每类样本在特定方位角度区间内以固定步长为间隔的傅立叶变换系数，依据某种统计准则构造滤波器组（类似于模板匹配的参考样本）；对于每个输入的未知 ROI 样本，先估计它的方位角（如主导边界法，网络盒法等）[17, 73]，然后对 ROI 切片做二维傅立叶变换，将变换后的系数与所有目标类别在 ROI 切片方位角所在区间生成的滤波器组进行二维卷积，根据响应结果来进行决策，若响应小于某个阈值，则有理由认为该 ROI 是人造虚警，反之则赋予其类别。显然，相关滤波的方法其滤波器组的生成过程与模板匹配的类似，不同之处在于相关滤波是在频域执行，而模板匹配是在图像域进行。表 1.5 对经典的滤波器构建准则进行了梳理和总结。

虽然相关滤波器的识别方法能够解决粗略配准的问题，实现了平移不变性，但它也存在着一些明显的缺陷：一是为了提高运算效率，该方法需要先估计未知 ROI 切片的方位角，然后根据估计的方位角去搜索相应的滤波器组，若方位角估计超过一定误差，则后续的相关卷积就失去意义；二是算法涉及到大量的二维傅立叶变换以及滤波器组与未知 ROI 样本之间的相关卷积计算，计算开销较大；三是该方法识别精度有限，通常只对当前数据有效，若更改实验数据，则效果不能让人满意，即对于扩展工作条件的适用性不好。

- **基于统计模型的识别方法 (Statistical Distribution Model)**。SAR 图像的生成与传统的光学成像不同，它是镜面反射和漫反射共同作用的结果，为了量化地表述 SAR 图像这种特殊的形成机理，Joseph A. O’Sullivan 和 Michael D. DeVore 等学者提出了一种基于统计分布模型的目标识别思路 [84, 85]。该方法的核心框架如图 1.10 所示，将特定方位角下的 SAR 图像所有像素（将二维图像延伸为一个列向量）建模为服从某种参数化统计模型的样本，根据每

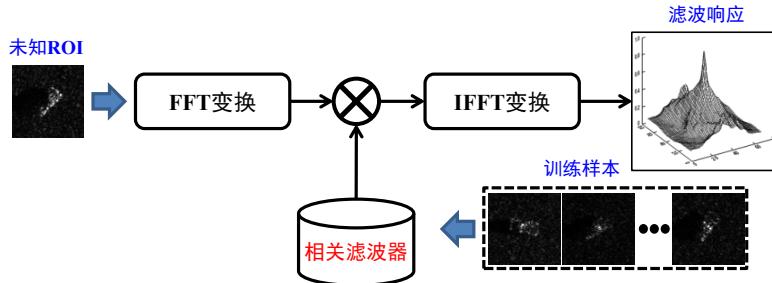


图 1.9 基于相关滤波的目标识别方法流程图。

表 1.5 经典的相关滤波器方法。

提出者	相关滤波器
David Casasent	Spatial Matched Filter [9, 74]
B. Vijaya Kumar	Optimal Trade-off Synthetic Discriminant Function Filter [75, 76]
Abhijit Mahalanobis	Maximum Average Correlation Height Filter [77]
Abhijit Mahalanobis	Distance Classifier Correlation Filters [78]
Mohamed Alkanhal	Extended Maximum Average Correlation Height Filter [79]
Mohamed Alkanhal	Polynomial Distance Classifier Correlation Filter [80]
David Casasent	Minimum Noise and Correlation Energy Filter [81–83]

个目标类别的训练样本估计其相应的统计分布的模型参数；对于未知 ROI 切片，首先利用希尔伯特－施密特估计器（Hilbert—Schmidt Estimator）估算其方位角，然后分别计算它在每类目标（每个角度区间）的条件分布概率函数（也可以看作是似然函数），最后根据最大后验概率准则判定目标类型。Harry L. Van Trees 最早在他的著作中使用瑞利分布模型 (Rayleigh model) 描述目标雷达回波和一组散射中心 [86]。Michael D. DeVore 等人使用广义条件高斯分布－条件 Rician 模型描述 SAR 图像复数据的幅度值，从而解释合成孔径雷达的镜面散射中心和漫反射中心 [84]，他的科研同事 Joseph A. O’ Sullivan 则认为条件高斯分布族更适合 SAR 图像 [87]。类似地，K. Copsey 和 A. Webb 将贝叶斯伽玛混合模型用于雷达距离向的信号分类问题 [88]。Michael D. DeVore 和 Joseph A. O’ Sullivan 利用 Kolmogorov-Smirnov 检验、D’Agostino-Pearson 检验以及拟合优度对不同条件下的统计模型分布的识别性能和计算复杂度进行了量化地比较和评估，并得出了许多有价值的结论 [89]。基于统计分布模型的目标识别方法需要根据大量训练样本数据来拟合统计模型分布参数，数据越丰富则拟合优度越高，因此，如果待测数据与训练数据之间不存在较强的统计关系（如仅有少量样本则估计的模型参数就有失偏颇），该方法的准确率就比较低。

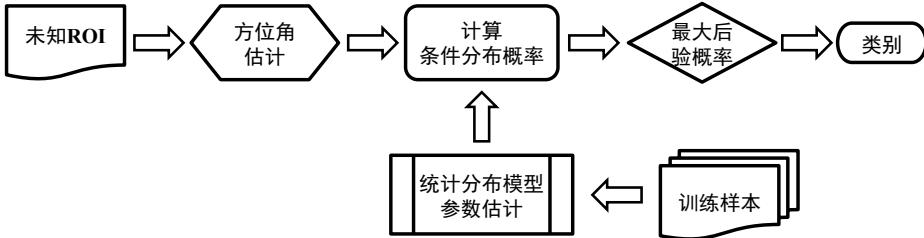


图 1.10 基于统计分布模型的目标识别方法简明流程。

- **基于机器学习的识别方法 (Machine Learning)**。目标识别是项系统性工程，每个处理环节实质就是一种机器学习方法，比如目标鉴别就是利用给定的目标样本进行学习，得到目标类别在特定的特征空间的表示向量，然后通过计算未知 ROI 与目标类别在特征空间的分离度来判断它属于虚警或是目标。因此，直观的想法就是借鉴模式识别和机器学习等相关领域的研究成果，改进目标识别的子环节，提高处理精度和运行效率。Steven K. Rogers 等人提出了一种生理驱动和人工智能神经网络的 SAR 图像自动目标识别方法，他们分别针对图像预处理（分割）、特征提取和选择以及分类三个进行改进 [63]。Li B. 等人将多个学习方法：卷积神经网络 (Convolutional Neural Network)、主成分分析、线性鉴别分析、向量量化学习、模块神经网络 (Modular Neural Network) 等引入自动目标识别，分别对这些学习方法进行了横向量化评估，结果表明算法未能取得好的结果其原因在于缺少理想的训练集 [90]。Zhao Qun 和 Jose C. Principe 使用了将能够产生较大分类边界的多个学习方法，如支持向量机 (Support Vector Machine)、AdaBoost 等用于 SAR 自动目标识别，提高分类正确率以及虚警样本的拒识率 [65]，在此基础上 David Casasent 等人提出了基于支持向量机的分层分类器方法 [91, 92]。Chen Yin 等人对于多个特征提取方法如主成分分析 (Principal Component Analysis)、线性判决策分析 (Linear Discriminant Analysis)、核主成分分析 (Kernel Principal Component Analysis) 以及多个分类方法如最近邻分类器、贝叶斯分类器、支持向量机从实验的角度进行了量化比较和评估 [93]。Sun Yijun 等人提出了基于 AdaBoost 学习的 SAR 图像目标识别算法，他们将 SAR 切片图像分别表示为原始特征和精炼特征，其中原始特征目的是为补偿目标姿态角估计误差对精炼特征造成的影响，然后使用高斯径向基函数的自适应 Boosting 算法完成分类识别 [66]。胡利平等人认为 SAR 图像数据服从多模分布，通过 k 均值聚类生成每类目标的子类划分，根据子类划分寻找最优投影矢量，使得不同类别的子类样本差异度更大，而子类内部样本的聚合度更强 [94]。Jayaraman J. Thiagarajan 等人将稀疏表示理论引入 SAR 自动目标识别，他们使用不同视

角下的训练样本矢量集生成字典，根据字典的基原子对未知 ROI 进行线性表示，通过施加稀疏约束项获取唯一的表示系数，根据重构误差判决目标类型 [95]；王燕霞和张弓提出使用广义二维主分量分析来构造过完备冗余字典，实现稀疏表示分类 [96]；Yang Shuyuan 等人使用脉冲轮廓波来刻画目标的方向选择特征，将其投影到核空间进行稀疏表示，实现目标识别 [97]；Zhang Haichao 等人提出了多视联合稀疏表示的目标识别方法，他们将稀疏信号表示用于同一目标在多个视角下的 SAR 图像目标识别 [98]；Cheng Jian 等人还提出了改进的联合稀疏表示学习方法，他们将分类划分为凸松弛和低秩矩阵复原，从而实现目标识别 [99]。Liu Ming 等人提出了一种将局部保留属性 (Locality Preserving Projections) 嵌入到统计模型的识别策略，以获取图像数据中需要的信息。他们使用高斯混合分布模型 (Gaussian Mixture Distribution) 描述 SAR 图像的斑点噪声，将局部保留属性嵌入到该统计模型中，借此来研究 SAR 图像的配置变体问题 [60]。Umamahesh Srinivas 等人将概率图模型引入 SAR 图像目标识别，以解决多特征集的融合和小样本分类学习两个突出问题。他们提出了两级处理框架，首先将 SAR 切片图像投影至富含信息的特征空间，从而得到多个互补的 SAR 图像表示，接下来将这些特征表示生成的图组合成为概率图模型，实现分类识别 [67]。Douglas R. Morgan 和 Timothy D. Ross 设计了一种贝叶斯框架，用于将多个分类器结果进行决策级融合，提取分类正确率 [100]，类似的思想也在文献 [101, 102] 中被采用，Liu Haicang 和 Li Shutao 则对 SVM 与稀疏表示两种分类器的识别结果进行决策级融合 [103]。Timothy D. Ross 等学者依托美国空军实验室的 FITE (Fusion for Identifying Targets Experiment program) 计划，对当前的决策级融合方法进行了回顾与总结 [104]。

1.2.2 面临的困难与挑战

为了研究 SAR 图像目标识别在实际战场应用中遇到的困难和挑战，论文对合成孔径雷达成像机理的要点进行简要回顾，其中涉及到目标特性的主要有：

1. 后向散射机制——SAR 图像中区分不同类别目标（如 T72 坦克 Vs. BMP2 装甲车）的关键信息取决于雷达入射波形与目标结构材料之间的交互结果，而从探测目标所返回的后向散射电磁波能量就携带了这些关键信息。如图 1.11 所示，目标的后向散射信号通常是由多种散射现象相互交织而生成，包括：平面反射、二面角反射、三面角反射、电磁波与地面的交互以及其它一些更为复杂的高阶多次反射（如类似于洞穴结构的多次反射）。目标结构材料所独有的散射机制反映出一定的类别信息。

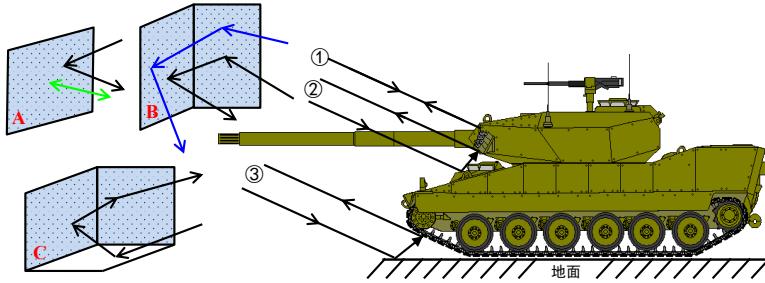


图 1.11 构成目标特性的典型散射现象：A. 平面反射示意、B. 二面角反射示意、C. 三面角反射示意，其中平面反射最不稳定，三面角反射最稳定，二面角反射居中；目标 SAR 图像生成过程的①平面反射、②二面角反射、③目标与地面的相互反射实例。

2. 像素物理意义——SAR 图像实质是源自于目标的三维后向散射回波所生成的二维平面图像。在平台工作过程中，传感器和目标之间存在着相对运动，通常使用距离 / 多普勒成像来进行解释。雷达回波所生成的平面图像中，距离向对应着传感器发射的电磁脉冲与接收到的反射回波之间的时延，而与某个时延 (Δt_i) 相关的所有回波位于一个直径为 $c\Delta t_i$ (c 表示光速) 的特定球体上；类似，切距离向 (cross-range) 对应于平台运动所导致的回波多普勒频移（增加或衰减），与某个多普勒频率 (f_{dj}) 相关的所有回波位于一个顶角半角满足 $\cos \theta_j = \lambda \frac{f_{dj}}{2V}$ (λ 是发射电磁波波长， V 是平台的瞬时速度) 的特定锥体上。因此，SAR 图像第 (i, j) 个索引位置的像素值就是第 i 个距离向球体与第 j 个多普勒锥体的横断面所定义的三维空间中的一个圆上的目标回波强度值。
3. 三维投影原理——SAR 图像的形成实质是接收到的后向散射回波在某个斜面的投影，这个斜面是由雷达到目标中心的视线和平台速度矢量所决定，如图 1.12 所示。这与传统的光学传感器成像截然不同，光学成像是将目标投影到与视线垂直的平面上，SAR 成像的投影平面本身就包含传感器视线，这种特殊的投影方式（投影于雷达视线所在的斜面）几何特性决定了图像中目标的几何特征。
4. 相干累积叠加——作用于观测目标表面的二次散射机制随着成像几何的变化而变化。目标成像过程每个像素位置的回波都是由投影于其表面的多个独立的后向散射回波的相干叠加作用而成。最终的响应幅度取决于这些单个独立回波是产生建设性交互或是破坏性干涉的相干叠加。单个回波间的相对相位改变或是任意回波幅度的变化都会使像素的整体幅度发生改变。如图 1.13 所示，雷达俯仰角变化将会改变投影于给定目标像素位置的分量的相对相位 [27]。

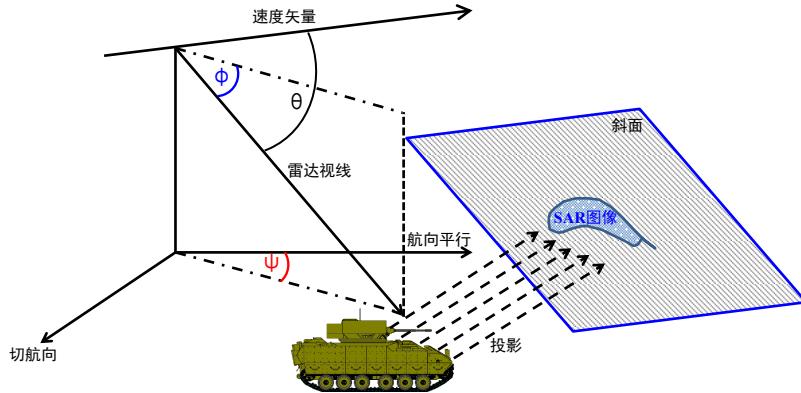


图 1.12 SAR 图像平面投影原理, ψ : 斜视角 (Squint Angle); ϕ : 俯仰角 (Depression Angle); θ : 多普勒频率构成的锥体顶角 (Cone Angle)。图中蓝色方框表示 SAR 图像投影斜面, 该平面包含雷达视线。

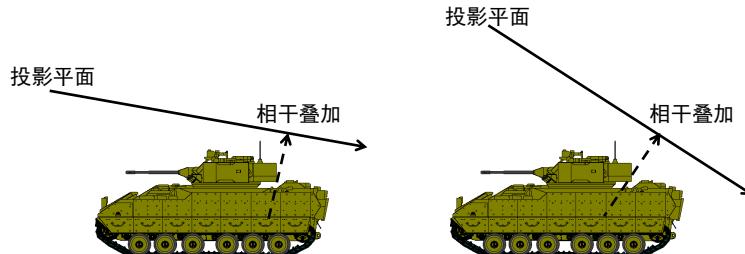


图 1.13 SAR 图像像素相干叠加受成像几何变化的影响。

与普通光学成像不同, SAR 图像的生成过程有其本身的特殊性, 因而, 合成孔径雷达自动目标识别在实际战场应用时将会面临许多特殊的困难和挑战, 其中最关键的就是目标特性的易变性, 根据 SAR 图像的生成原理可以将影响目标特性的主要因素概括为成像参数、目标变化和周围环境三类。

1. 成像参数: 即传感器平台的工作参数。

俯仰角与入射角: 在 SAR 图像的投影生成过程中, 投影斜面的构成主要取决于雷达视线的俯仰角和斜视角两个参数。对于水平直线飞行的传感器平台, 其俯仰角 (ϕ_{mc})、斜视角 (ψ_{mc}) 和多普勒锥体半角 (θ_{mc}) 之间满足关系: $\cos \theta_{mc} = \cos \phi_{mc} \cos \psi_{mc}$ 。从图 1.14 可以看出, 若俯仰角或是入射角发生改变, 即便是目标的相同位置的后向散射信号会发生显著改变。

投影平面方向角: SAR 图像是目标空间几何形状的平面展示。平台接收到的信号是由多种不同类型的回波构成, 如平面反射, 多次反射回波等 (如图 1.11 所示), 其中平面反射信号是源自于目标表面轮廓, 多次反射回波生成于目标的凹凸不平等内部区域, 而且大部分回波信号都会投影于斜面上, 因而, 目标特性将会随着 SAR 平面方向角的改变而改变。

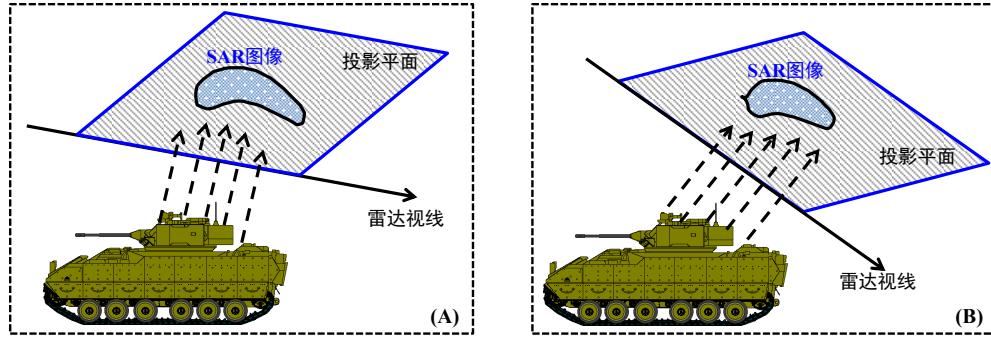


图 1.14 同一目标 (A) 缓俯仰角、(B) 陡俯仰角的投影成像。

其它成像参数: 如雷达分辨率、极化方式 (HH、VV、HV、全极化)、工作模式 (条带式、聚束式、滑动聚束式)、噪声水平、运动补偿等，这些参数的细微改变会对生成的目标特性产生显著的影响。

2. 目标结构：即目标本身结构、部件升级、姿态等。

结构部件变化: 如连接部件变化 (炮塔旋转、炮管直立与否等)、局部损坏、配置变异 (如增加副油箱)、战斗部署等，如图 1.15 所示。同一目标不同姿态的目标特性将会呈现出明显差异，同一目标战斗状态和行军状态其目标特性也会有显著变化；同类目标不同变体其目标特性呈现明显变化。

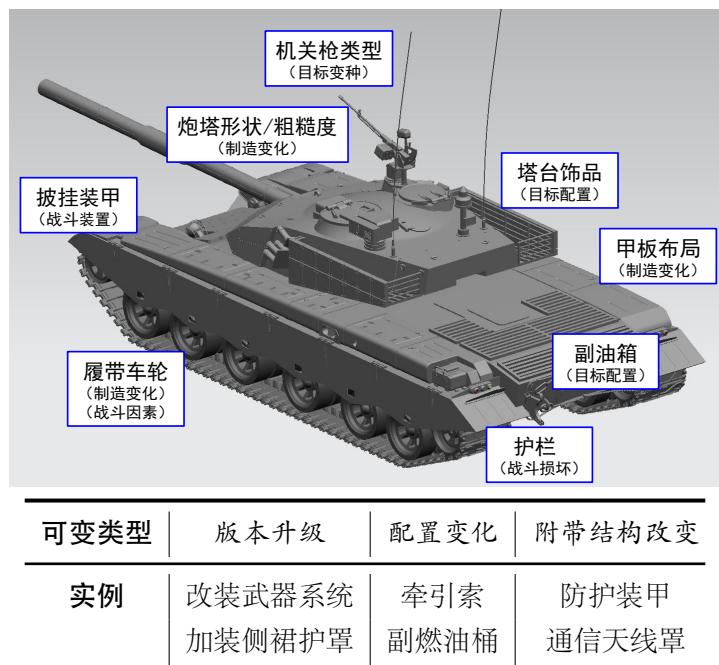


图 1.15 坦克部件配置引起的变体示意图。

目标姿态视角: SAR 图像是入射电磁波信号与目标物质结构之间的电磁交互作用的结果, 是被探测目标不同位置部件的回波相干叠加并投影到斜面而成, 因而, 即使是入射波相对视角的微小变化也会显著地改变 SAR 图像目标特性。这与普通光学成像又有显著差异, 光学成像过程中, 表面散射是典型的 Lambertian 反射。因而, SAR 图像的目标特性对视角的敏感程度要远比光学图像严重很多。

备注: *Lambertian* 表面是指在一个固定的照明分布下从所有的视场方向观测都具有相同亮度的表面, *Lambertian* 表面不吸收任何入射光。*Lambertian* 反射也叫散光反射或理想散射, 不管照明分布如何, *Lambertian* 表面在所有的表面方向接收并发散所有入射照明, 且每个方向都能看到相同数量的能量。

3. 周围环境: 即目标周围的遮挡、伪装、欺骗等环境设置。

环境遮挡伪装: 为了躲避雷达监测, 敌方目标通常会部署于建筑物、高大树木等类似环境中, 这些周边环境与目标回波相互作用之后就会使目标特性更加模糊, 产生遮挡、重影等。如图 1.16 所示, 遮挡通常是因目标与传感器之间存在着物体(或障碍)而阻滞或是衰减从目标返回的雷达电磁波能量。遮挡将会使目标特性的某个子元素在回波信号中消失或是严重削弱, 它的影响程度取决于雷达工作的俯仰角以及遮挡物的高度与位置。重影是指从目标附近的物体返回的电磁波与目标的回波相叠加, 从而模糊目标的结构特性。重影发生的根本原因在于三维空间电磁波在距离向和方位向二维 SAR 图像上平面投影的内在模糊性, 如图 1.17 所示, 若物体本身在空间上位于目标之后, 但却高于目标, 此时它的回波就会与目标回波相互叠加, 重影的程度通常会随着俯仰角的提升而增加。

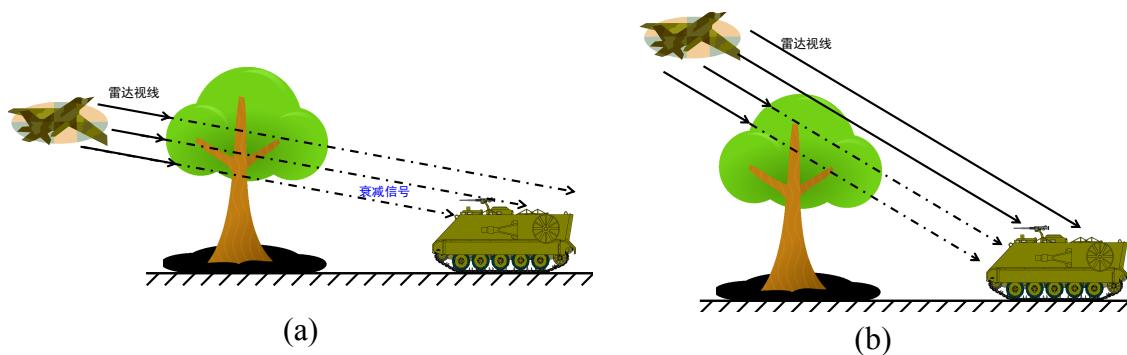


图 1.16 树木遮挡对目标成像的影响, (a) 缓俯仰角时, 所有探测到目标的入射波都因穿透树木而发生衰减; (b) 陡俯仰角时, 雷达发射的部分电磁波产生衰减信号, 而另一部分不受影响。

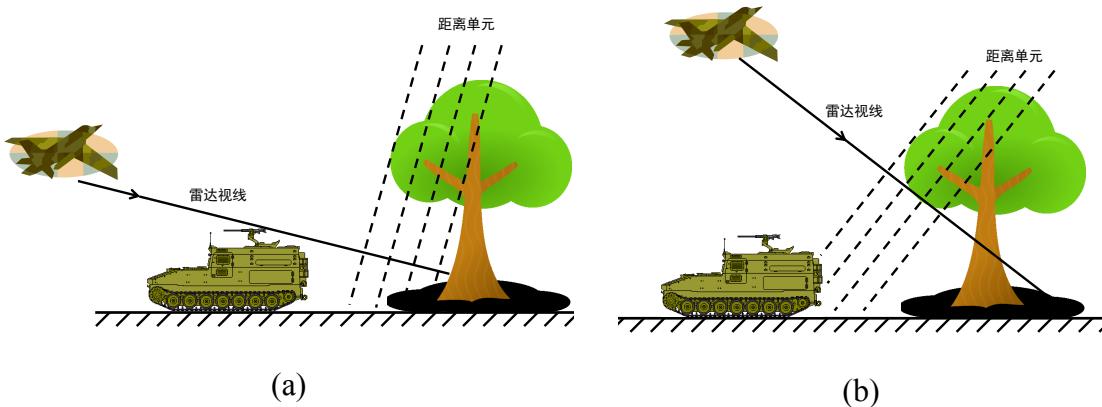


图 1.17 目标与植被的重影对成像结果的影响, (a) 小俯仰角、(b) 大俯仰角时树木枝叶回波对目标特性距离单元的污染。

目标欺骗伪装: 给重要目标铺设伪装网是战场中常用的一种方法。伪装网通常是由嵌入了碳或金属的纤维布构成。当雷达发射的电磁波触碰伪装网时,一部分能量将会被吸收和均匀向周围重新扩散,从而衰减与目标交互的发射电磁波能量,削弱期望的回波响应。此外,电磁波触碰伪装网时所引发的二次扩散辐射,二次扩散将产生目标次生回波响应,这是目标识别系统所并不期望的,因为这种组合响应将会模糊并减少目标的结构对比特性。另一种常用的欺骗方法就是地面诱饵,其目的是为了降低真实目标被检测的概率。诱饵通常由廉价材料构成,通过合理配置会生成与真实目标相似的雷达特性,如常见的角反射器。

1.2.3 主要关键技术

根据前文第 1.2.1 节对现有研究方法的总结以及第 1.2.2 节梳理的困难与挑战,可以得出 SAR 图像目标识别所涉及到的主要关键技术包括特征描述和分类学习,其中特征描述主要任务是寻找设计一类有效的、普适的、计算高效的 SAR 图像不变特征描述方法,分类学习就是改进现有的学习方法、完成分类识别。围绕着这两个关键环节,论文将分别进行阐述。

1.2.3.1 特征描述

根据前文的描述可知, SAR 是一种斜距投影的相干成像雷达,其成像机理与其它遥感系统特别是光学传感器有着显著的差异。目标 ROI (车辆、坦克、火炮等军事目标) 通常是由金属制成,它的 SAR 图像中表现为距离向和切距离向的二维散射强度分布,是若干个平面反射、二面角反射、棱锥反射、圆柱体反射等现象共同作用的结果 (如第 1.2.2 节图 1.11 所示),而目标的散射特性反映了雷达电磁波与被探测目标相互作用时表现出的物理结构 (如大小尺寸、结构构造、几何形

状、材料属性等)与雷达工作参数之间的某特殊关系 [2]。这种成像模式使得 SAR 图像目标识别可以使用传统的特征描述方法，也可以借助类光学图像特征，还可以使用 SAR 图像特有的目标电磁散射特征。归纳起来，SAR 图像目标识别中所使用的特征描述方法主要有：

强度散布值 即像素级特征，指使用原始的强度值或是经过简单的预处理（如归一化、超分辨、图像增强等）作为识别的特征。

早期的 SAR 图像目标识别方法主要依赖模板匹配的识别策略。林肯实验室的 Leslie M. Novak 等人提出用目标的原始强度分布数值作为特征，使用均匀划分的方位角区间内的均值样本作为模板库参考样本，通过搜索未知 ROI 与参考样本之间的最小均方误差来实现目标识别 [40, 42]，Gregory J. Owirka 等人提出对原始强度分布图进行增强预处理 [36]（如 HDVI 算法 [41]）；L. Novak 等也提出一种超分辨率图像处理方法对原始散布矩阵做图像增强 [105]。使用原始散布矩阵简单直观，易于计算，但它对目标方位角以及噪声扰动十分敏感。若未知图像与参考图像方位角接近，或噪声污染等级较小，则识别率较好；反之识别结果就会受到很大影响。此外，林肯实验室设计的匹配方法仅能应付微小范围的图像配准误差，若未知 ROI 与参考样本存在着明显的空间偏移，则识别效果并不理想。

变换域特征 就是利用数学工具、数理统计方法等将原始散布矩阵投影到某个特定子空间，寻找不变特征。

- 参数化特征。参数化特征就是将 SAR 图像像素值（或是变换后系数）建模为参数化统计分布模型，利用不同类别的样本估计该类的模型参数。Joseph A. O’ Sullivan 和 Michael D. DeVore 将 SAR 强度分布矩阵建模为服从某种参数化统计分布模型的样本，然后根据每类目标的训练样本估计该类型统计分布的模型参数，对于未知 ROI，分别计算它在每类目标（每个方位角区间）的条件分布概率，根据最大似然函数准则判定目标类型 [84, 85, 87]，因而，统计分布的模型参数就是他们执行目标识别所采用的特征。K. Copsey 和 A. Webb 提出使用混合 Gamma 模型对目标的高分辨率距离向统计特性进行建模，因此，区分不同类别目标的特征是混合 Gamma 模型参数 [88]；Igal Bilik 等使用高斯混合模型（Gaussian Mixture Model）对低分辨率地面雷达的目标识别问题进行建模 [106]；Duan Rong 等使用高斯混合模型对区域划分后的 SAR 切片图像进行拟合 [107]。

- 投影特征。投影特征是指根据某种准则，利用给定的训练样本生成一组特定基，然后将训练样本和未知 ROI 同时投影到基空间，进行相似度量。韩萍等使用非线性主成分分析 (Kernel Principal Component Analysis) 对 SAR 图像强度矩阵进行投影，提取目标特征 [108]；胡利平等针对二维 PCA 在计算过程中消除了散布矩阵行像素间的相关性，提出了一种两级二维 PCA 的投影策略 [109]；韩征等人为改进识别模板的可分性，强化不同类别在投影空间的差异性，提出了一种样本正交子空间的投影方法 [110]；王世希和贺志国为了避免传统 PCA 应用于 SAR 切片图像时的维数灾难，引入了基于 Hebb 学习规则的自组织神经网络来实现主成分分析 [111]；龙泓琳等利用水平集图像分割方法生成初始值，将其做非负矩阵分解，使用得到的权向量作为图像特征，最后借助 Fisher 线性判决进行分类识别 [112]；苏志刚等针对空间投影方法对于具有相似形态结构的目标效果不理想，提出了一种投影空间构造方法，改进不同目标间的可分性 [113]；李映等将核奇异值分解 (Singular Value Decomposition) 得到的图像非线性变换系数输入到 PCA 提取特征 [114]；其它的投影变换特征还包括线性判决分析 LDA、独立成分分析 ICA、非负矩阵分解 NMF、局部保留投影 LPP、稀疏保留投影 SPP 等 [59, 60, 115]。
- 时频特征。时频特征是指使用傅立叶变换、离散余弦变换、小波变换等时频分析工具的变换系数对 SAR 图像进行描述的特征。相关滤波器利用每类目标在每个固定方位角度区间内的傅立叶变换系数生成滤波器组，根据未知 ROI 对每组滤波器的响应判定目标类型 [9, 71, 83, 116]，这类识别策略是利用傅立叶变换系数进行特征描述，实现平移不变和扭曲不变特性；宦若虹和杨汝良对预处理后的 SAR 切片图像做二维离散小波变换，将分解后的低频子带系数作为图像特征 [117]，Nicholas Sandirasegaram 和 Ryan English 对二维傅立叶变换和二维小波变换提取的时频信息的目标鉴别能力进行了量化分析 [118]；Hu Fengming 等人根据 Gabor 滤波器组响应的子区域系数设计统计量，将统计量作为目标识别的特征 [119]；王爽等人对 SAR 图像目标区域时行曲线小波 Curvelet 变换，将得到的低频子带变换系数作为目标的轮廓特征，再借助数学形态学工具处理低频系数，生成区域特征，实现对目标的特征表述 [120]；学者 Lance M. Kaplan 和 Romain Murenzi 使用与目标尺寸相适应的二维连续小波变换增强图像的信杂比 (Signal-to-Clutter Ratio,

SCR), 提高对焦检测性能 [121]; 其它的时频分析工具还包括轮廓小波变换 Contourlet、Shearlet、Bandlet、Ridgelet 等。

变换域特征在某些数据集和某些实验中取得了较好的识别性能，但这些投影变换通常并不具有明确的物理意义；雷达目标识别是一种带有军事背景的研究课题，研究人员大都是在封闭的环境中开展研究，用于算法验证的公测数据集很少，因而很难去验证这些特征的通用性，识别的稳健性和可靠性也不得而知；投影基矩阵是变换域特征是否有效的关键，而投影基矩阵通常是建立很多的训练样本的基础上，运算开销较大，若不借助一些数值计算技巧，普通的电脑动态存储容量很难容纳。

几何形状特征 指点、线、面、纹理、尺寸、形状、区域等描述方法。典型的几何特征提取方法是：先通过对原始 SAR 图像进行预处理，将目标区域和阴影区域从背景杂波中分离出来，然后再计算提取目标或阴影的有关几何特征，如大小尺寸、面积、对比度、椭圆主轴、轮廓、不变矩等。Clark F. Olson 和 Daniel P. Huttenlocher 通过比较 SAR 图像不同类型目标的边缘图（边缘像素所组成）之间的距离实现目标识别 [122]；Chanin Nilubol 等使用 Radon 变换将 SAR 切片图像映射成向量，然后使用隐马尔可夫模型实现目标识别 [123]；计科峰等构造峰值位置、峰值幅度等特征来估计目标方位角，实现目标识别 [17]；John A. Saghri 和 Andrew DeKelaita 在 SAR 图像目标识别中引入阴影信息，将阴影轮廓与传统的特征如峰值、边缘、角点等联合起来进行目标识别，此外，阴影信息还可做目标方位角估计，从而选择相应的参考样本进行匹配，减少计算量 [124]；Scott Papson 和 Ram M. Narayanan 认为阴影是目标外形的背向投影，故某些方位角成像将会包含更多的鉴别信息，他们使用阴影轮廓作为特征进行目标识别 [125, 126]；徐牧等使用 Hough 变换进行 SAR 图像线段检测，估计目标方位角 [73]；

电磁散射特征 指雷达电磁波与物体交互时产生的散射回波。与普通的光学图像不同，SAR 图像目标通常表现为一系列不连续的二维峰值分布，实际上这些峰值点就是散射中心在投影斜面的表现形式，它与目标的材料结构有着密切关系。Lee Potter 和 Randolph Moses 认为人造目标的高频测量值取决于那些孤立的散射中心回波，如角点、平坦的块状区域等，这些散射中心的特征描述能够有效区分不同类别目标，因而他们提出通过雷达回波的参数预测实现特征提取，对于每个散射中心，模型参数的稳健估计能够提供包括位置、几何形状、极化响应在内的高分辨率属性，从而实现目标识别 [28]；Grinnell Jones 和 Bir Bhanu 根据目标散射中心的不变性和目标散射幅度不变性设计

不变特征，识别目标的物理变体（如装备配置的变化）[127, 128]；Chiang Hung-Chih 等提出了一种点位置与物理属性向量的特征描述方法，属性特征采用高频雷达散射参数模型，然后利散射中心特征计算目标候选假设的后验概率，实现目标识别[64]；Jon Gudnason 等使用自适应目标掩模和图像超分辨率算法估计目标散射中心的切距离向位置，从而生成目标距离向伪图，将伪图的低阶离散余弦变换系数作为目标识别的特征[129]；为了避免传统的点对点匹配方法，周剑雄等提出了一种区域到点的比对策略，首先对 SAR 切片图像做阈值化和形态学滤波处理，然后提取的目标区域用于估计散射中心模型，解决模板库参考样本不足的问题[130]；为了从高分辨率 SAR 图像中估计目标位置，判断目标的散射体类型，提取雷达回波承载的有用信息，娄军等提出了一种基于 Separate Ensemble Empirical Mode Decomposition (SEEMD) 的目标散射特征提取方法[131]。

几何形状特征和电磁散射特征两类特征描述方法需要将目标（或阴影）从背景杂波中分离出来，这个过程涉及到多个图像预处理步骤和技巧，如相干斑抑制、图像分割、数学形态学处理（消除孔洞、区域连通）、目标聚类等，如图 1.18 所示，每个中间处理步骤的精确程度都会对特征的鉴别能力产生重要的影响，且涉及到许多繁琐的数值计算。

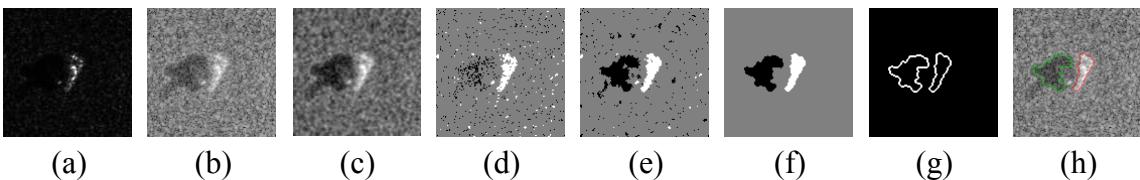


图 1.18 SAR 目标切片特征提取流程。图中给出一组 T72 坦克切片图像特征提取流程。(a) 姿态角为 240° 的 T72_132 在 15° 俯仰角的成像样本；(b) 对数变换；(c) 抑制噪声；(d) 粗分割结果；(e) 做数学形态学；(f) 做几何聚类；(g) 目标边缘轮廓和阴影边缘轮廓；(h) 目标与阴影边缘轮廓与原始图像叠加。从可以看出，从原始的强度分布图到最终的目标散射中心和阴影区域需要执行许多中间环节，每个环节的处理精度直接影响最终结果。图中所示的处理结果并没能提取出炮管阴影。

1.2.3.2 分类学习

分类是机器学习的范畴，指根据给定已标记的样本数据推测未知样本的标记。SAR 图像目标识别中所使用的分类方法都是从机器学习等相关领域借鉴而来，然后根据 SAR 图像目标识别的需要重新定义相似准则（或距离度量），典型的分类学习方法有：

- 最近邻分类器 (k-Nearest-Neighbor Classification) 是最基础、最直观、最常用的分类方案 [132]。首先定义特征空间中未知样本特征向量与训练样本特征向量之间的相似性度量，如欧氏距离、Mahalanobis 距离、Manhattan 距离、Minkowski 距离、Chebyshev 距离等，然后搜索寻找与未知样本特征向量距离最小的训练样本，将找到的训练样本的类别标记赋予该未知样本。一种常用的扩展形式就是 k 近邻，即寻找未知样本距离最近的 k 个训练样本，然后对 k 个训练样本标记进行投票，票数最多的类别即是未知样本的标记，如图 1.19 所示。 k 近邻的更广义扩展就是最近子空间分类 [133]，该方法认为每个类别的所有样本共同构成一个线性子空间，通过定义单个样本与子空间之间的距离度量，寻找未知样本的最近子空间，从而完成分类标记。对于最近邻、 k 近邻或是最近子空间分类，关键都是如何合理地定义距离，即样本与样本之间的距离（点对点距离）、样本与子空间的距离（点对面距离）。

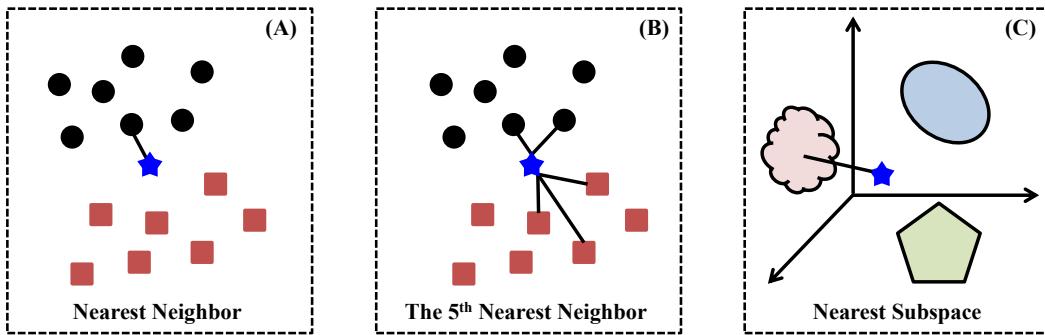


图 1.19 最近邻分类示意图，蓝色五角星表示待分类样本，图 (A)、图 (B) 中黑色圆圈和淡红色方框表示两个不同类别，黑色线段代表样本与样本之间的距离，图 (C) 中三个不同颜色的形状分别代表三类不同样本构成的线性子空间，黑色线段表示样本与线性子空间之间的距离。

经典的模板匹配策略通常采用 MSE 分类器，该方法的思路是先定义参考模板与未知 ROI 的均方误差

$$\varepsilon = \frac{\sum_{i=1}^N (R_i - T_i)^2}{N}$$

其中 N 是当前参考模板窗像素个数， R_i 和 T_i 分别表示参考样本和测试样本的像素。均方误差最小的参考模板即认为是最佳匹配 [19, 36, 105, 134]。因而，MSE 分类器也可以看作是最近邻分类，它的距离定义为均方误差。此外，变换域特征大都采用最近邻分类进行决策，决策依据是未知样本的投影向量与训练样本投影向量之间的欧氏距离 [109–111, 113, 114]。Jong-IL Park 和 Kyung-Tea Kim 提出了一种改进的极化映射分类器，首先寻找目标质心，

以质心为中心做 360° 极化映射，将映射后得到直方图作为图像特征，之后的分类学习划分为相关匹配和最近邻分类两步，实现目标识别，它的距离度量就是两列直方图之间的相似性 [135]。

- **线性分类 (Line Classification)** 通常先定义决策函数为特征向量各元素的线性组合，然后对决策函数进行阈值化，完成分类。若以 f 表示决策函数，以 $\mathbf{x} \in \mathbb{R}^N$ 表示未知样本的 N 维特征向量，以两个类别的分类任务为例，则线性分类可表述为

$$f(\mathbf{x}) = \sum_{n=1}^N w_n \cdot x_n + w_0 \Rightarrow \begin{cases} \geq 0 & \text{assign } \mathbf{x} \in \text{class 1} \\ < 0 & \text{assign } \mathbf{x} \in \text{class 2} \end{cases}$$

其中 w_0 称为偏置项， $w_n, n = 1, \dots, N$ 是特征向量中每个元素的权重。多类别的分类以此类推。

线性分类的决策边界为 $f = 0$ ，它是个 $N - 1$ 维的超平面，二维特征空间的决策边界退化为一条直线。线性分类的关键是如何确定决策边界，而影响决策边界的主要因素就是偏置项和权重系数。线性分类学习对于数据集的要求较高，必要线性可分，因而对于非线性的数据集（如图 1.20¹）效果不佳，此时就需要引入 kernel 映射，将非线性数据投影到某个抽象空间，使其线性可分，这就是非线性分类的处理思路。

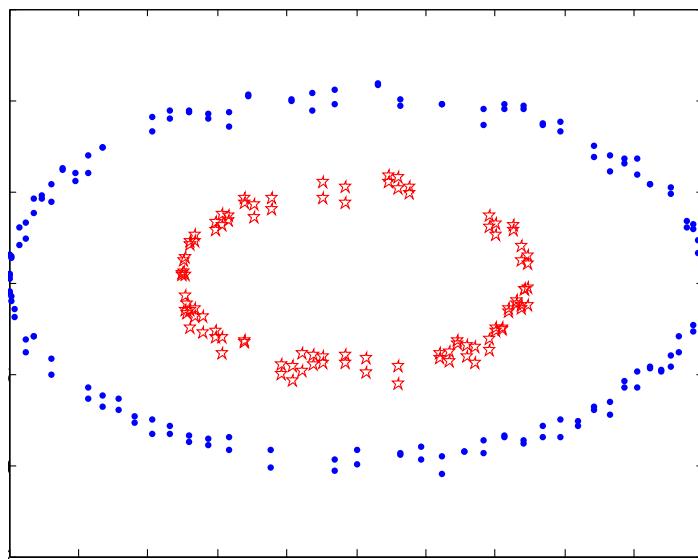


图 1.20 线性不可分数据，图中的蓝色实心圈和红色五角星分别代表两种不同的类别，无法找到一条直线（超平面）将两类数据分开，理想的分界应该是个“椭圆圈”。

¹参考网友 pluskid 的博客 <http://blog.pluskid.org/>。

- **非线性分类** (Nonlinear Classification 或 Kernel-based Classification) 借助非线性映射将原始数据映射到另一个虚拟的抽象空间，然后使用一个线性判决函数对待标记的未知数据进行分类，从而解决原始数据空间中类别线性不可分的分类问题。非线性分类最典型的代表就是支持向量机 (Support Vector Machine) [136, 137]，若以 \mathbf{x} 表示未知数据，以 $\phi(\cdot)$ 表示非线性映射，SVM 的决策函数可表示为

$$f_{SVM}(\mathbf{x}) = \mathbf{w}^T \cdot \phi(\mathbf{x}) + b$$

其中 b 是偏置项， \mathbf{w} 是权重系数，它们可通过训练样本学习得到。该方法利用非线性映射将原始数据投影至一个高维抽象空间，使得抽象空间中不同类别的数据的像的可分性得到改善，然后通过最大决策边界超平面和支持向量来解决因非线性映射带来的计算复杂度和过度拟合问题，支持向量是指那些用于估计权重系数和偏置项的一小部分训练样本，如图 1.21 所示，支持向量的挑选原则就是使样本点与决策边界之间的间距最大。

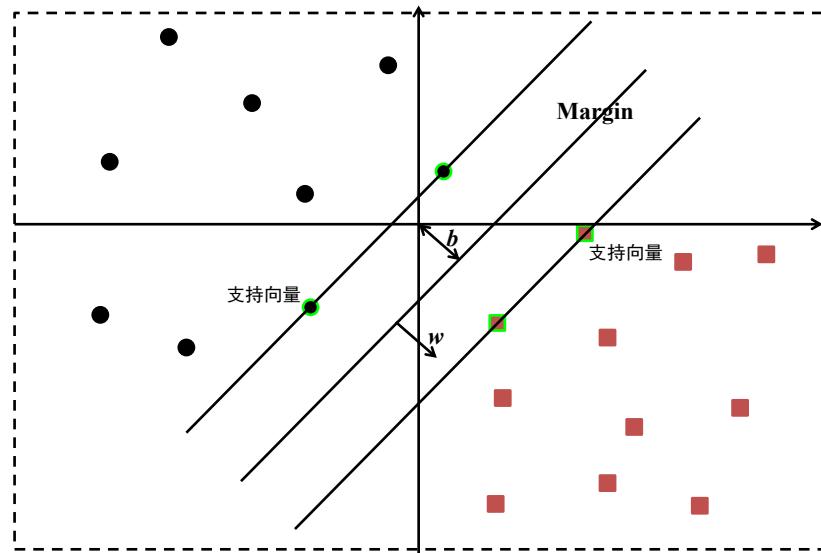


图 1.21 支持向量机示意图，图中黑色圆圈和淡红色方框分别表示两个不同的类别，绿色边框的 4 个样本就是挑选出的支持向量用于估计偏置项 b 和权重系数 \mathbf{w} 的样本。

Zhao Qun 和 Jose C. Principe 使用支持向量机解决公开条件的目标分类识别，并通过实验验证了高斯核支持向量机的分类正确率和拒识率均优于模板匹配；David Casasent 和 Yu-Chiang Wang 提出一种层级分类学习方案，层级中的每个节点都使用鉴别机器支持向量机，提高拒识性能 [91]；此外，文献 [92, 108, 138] 等采用支持向量机学习进行分类。

- **贝叶斯分类** (Bayesian Classification) 通常使用概率密度函数 (Probability Density Function) 将样本原始数据 (或是特征向量) 同时生成的联合概率进行建模，然后利用贝叶斯准则根据其统计特性进行分类。首先根据已标记的训练样本计算未知样本 \mathbf{x} 属于第 k 类的条件似然函数 $p(\mathbf{x}|k)$ ，其中 $k \in [1, 2, \dots, K]$ 是类别标记，然后再根据似然函数和先验概率 $p(k)$ 计算未知样本属于第 k 类的后验概率 $p(k|\mathbf{x}) = \frac{p(\mathbf{x}|k)p(k)}{p(\mathbf{x})}$ 。最后根据最大后验概率准则进行判决。隐马尔可夫模型 (Hidden Markov Model) [139] 是一种典型的贝叶斯分类。

J. A. O’Sullivan 等将未知 ROI 属于某类目标的似然概率定义为条件高斯分布模型，然后根据最大似然估计判决未知目标类型 [87]；Myron Z. Brown 将 SAR 图像目标分类问题建模为在给定方位角和统计分布模型参数的条件下的贝叶斯估计，根据训练样本学习得到每类目标的统计分布参数，然后对未知 ROI 进行方位角估计，分别计算未知样本位于特定方位角属于每类目标的后验概率，实现目标识别 [140]；为了应对高分辨率距离向目标识别中的视角变化、时变以及幅度尺度敏感性问题，Du Lan 等使用 Gamma 分布和高斯混合分布将不同类型距离单元的回波建模为相应的统计分布，利用训练样本估计统计分布的参数，从而得出未知样本属于每个类别的后验概率，实现目标识别 [141]；类似的学习方法也被 [64, 66, 67, 100, 106, 142–144] 等采纳。

- **神经网络** (Neural Network) 分类源自于人工智能，用于模拟人类大脑皮层的方向性选择活动，它是由许多相互连结的神经元组成。每个神经元的输入源自于另外的神经元，其输出是其它神经元所有输入的加权组合，

$$f_{NN}(\mathbf{x}, \mathbf{w}) = g(\mathbf{w}^T \cdot \phi(\mathbf{x}))$$

其中 \mathbf{w} 代表权重矩阵，它通过训练样本学习得到， $g(\cdot)$ 函数执行非线性变换， \mathbf{x} 是输入向量。

一类特殊且使用广泛的方法就是前向反馈神经网络 (Feed-Foward Neural Network)，或称为多层感知器 (Multilayer Perceptrons)，如图 1.22 所示，该方法将神经元以层来排列，即每层神经元从上一层的输出来获取它们的输入，同时它们的输出又会为下一层提供输入，实质就是深度学习 (Deep Learning) 的雏形 [145]；若将输入数据 \mathbf{x} 同时看作是输入层和输出层，则输出模式就可以用作分类，训练过程中权重矩阵 \mathbf{w} 和函数 ϕ 将会针对每个神经元分别进行调整。

Jose C. Principe 等通过对双参数 CFAR 检测器的准则进行扩展来构造一个基于图像局部强度的非参数分类器，称之为二次 Gamma 鉴别器（Quadratic Gamma Discriminator），然后通过对线性处理元素做非线性扩展从而得到一个多层感知器 [146]；Zhao Qun 等在估计目标方位角的前提下，分别对支持向量机、基于二次互信息代价函数的神经网络、多分辨率主成分分析三种学习和表示方法进行比较，结果表示方位角估计能够有效降低计算复杂度，同时提取目标识别准确率 [147]；Nicholas Sandirasegaram 和 Ryan English 将图像的 2D 傅立叶变换和 2D 小波变换作为特征，分别输入到 p 范数距离、多层次感知器和另外一种神经网络的商业算法三种不同的分类学习方案，评估它们的分类正确率与拒识别率 [118]；杨露菁等使用 Independent Component Analysis 为每个视角的 SAR 图像生成一组独立的基原子构成特征空间，每个特征空间分别训练一个模糊极小极大神经网络（Fuzzy Min-Max）用于分类，最后利用 DS 证据理论进行融合决策 [148]；此外，近年来也有一些学者尝试使用深度学习解决 SAR 图像目标识别问题 [149–153]。

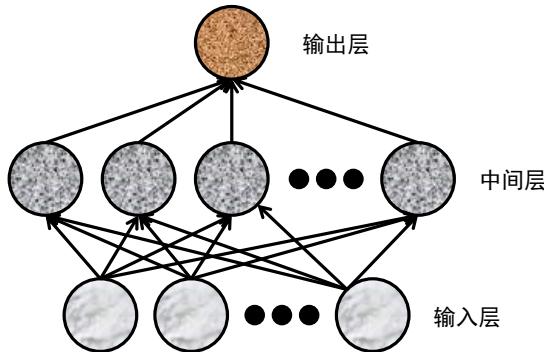


图 1.22 前向反馈神经网络，也称多层次感知器，图中不同颜色圆圈代表神经网络中不同层级的神经元，每个神经元的输入都是源自于网络中上一层级神经元的输出，同时它又是下一层级神经元的输入。近年来提出的深度学习（Deep Learning）实质是对多层次感知器的扩展。

- **稀疏表示分类**（Sparse Representation-based Classification）是近年来的一个热门课题 [154–156]。该方法假设同一类别的训练样本位于一个线性子空间中，将每个样本看作是线性子空间的一个元素。对于任意的未知样本，将其表示为所有训练样本的线性组合，从而构造多元线性回归模型，其中训练样本扮演回归参量的角色，未知样本是回归响应，待求解的表示系数是回归系数。由于未知量个数多于已知的方程式，这个病态的欠定问题有无穷多组解。若约束该线性系统的表示系数中仅有很少数的项为非零值，则该系数向量中就包含有关于未知样本类别的重要信息，从而会得到较好地分类效果。

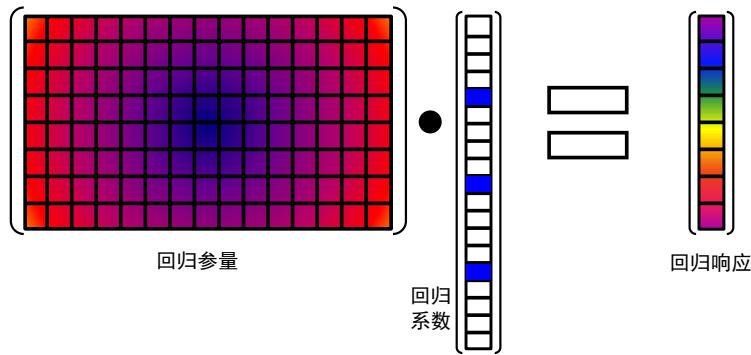


图 1.23 多元线性回归模型的稀疏求解原理，其中训练样本是回归参量或是预测量，待分类样本作为回归响应，最后利用回归系数进行决策。

E. Candès 和 T. Tao 首先证明了若线性回归模型满足稀疏约束条件时最小化 ℓ_1 范数问题和最小化 ℓ_0 范数是等价的，从而为稀疏表示最优化问题的高效求解奠定了基础 [157, 158]；Honglak Lee 等研究比较了 ℓ_1 和 ℓ_2 范数约束下的分类性能，并提出了一种 ℓ_1 范数的高效开源算法，特征符号搜索 [159]；John Wright 等系统地论述和推导了 ℓ_1 范数约束的稀疏表示分类方法，通过实验将该方法与最近邻分类、最近子空间分类、线性支持向量机等经典方法进行比较 [156]。为了避免 SAR 图像目标识别遇到的姿态角估计、图像预处理等问题，J. Thiagarajan 等将 John Wright 等提出的分类方法 [156] 用于 SAR 图像目标识别，使用不同方位角的成像数据生成过完备字典，根据该字典构建未知数据的线性表示模型，通过求解最小化 ℓ_1 范数获取最优表示系数，根据重构误差推断目标类型 [95]。之后，稀疏表示引起了研究人员的广泛关注与研究 [97–99, 160–164]。

此外，也有学者提出多分类器融合的思想，即使用数学工具如贝叶斯理论、DS 证据理论、模糊数学等将多个不同的分类学习结果进行合并，规避各自学习方法的缺点，最终得出联合决策。Lun Zhengjun 等将支持向量机与后向传播神经网络（Back Propagation Neural Networks）的学习结果进行融合 [165]；Huan R. 和 Pan Y. 对多视融合、多特征融合以及多分类器决策融合分别进行归类和讨论 [166]；Liu Haicang 和 Li Shutao 将稀疏表示分类与支持向量机的学习结果进行合并 [103]；Dong Ganggang 和 Kuang Gangyao 将不同输入的稀疏表示分类结果利用 DS 证据理论进行融合 [167]。

1.2.3.3 研究思路

针对特征提取和分类学习两大关键技术，论文期望通过引入相关领域最新的研究成果，分别进行解决。特征提取目的是寻找一种有效的、普适的、高效的方法

描述 SAR 图像目标散射现象，**有效性**是指特征应具备较强的鉴别能力，能够区分不同类别的目标，如图 1.24 所示，变换后的特征空间应具有较强的类内聚合性和较大的类间差别性，这是最基本的要求；**普适性**是指设计的特征能够适应不同的数据集、不同的实验设置等；**高效性**是指特征的数值实现过程应该计算简便、易于操作。

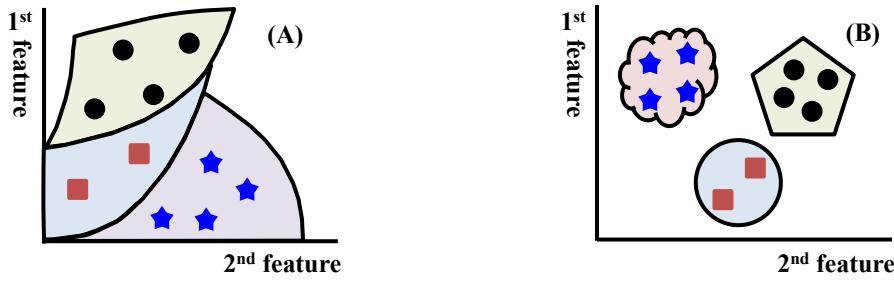


图 1.24 特征的有效性示意图，以二维特征空间为例，(A) 图设计的特征空间类内聚合性和类间差异性都逊于 (B) 图。

依据上述准则，本文把 SAR 图像方位向与距离向强度分布矩阵视为一类特殊的二维有限长信号，通过引入一种新的信号处理方法——单演信号，描述 SAR 图像二维散布矩阵，借助带通滤波器对原始信号做无限延拓，构成尺度空间，生成代表不同类型信息的局部分量，根据得到的局部分量设计特征描述子。论文的研究核心是构造 SAR 图像强度矩阵的单演信号尺度空间，只涉及到信号的调和分析，不需要执行繁琐的中间处理环节，不需要执行方位角估计。特征的有效性、普适性和高效性准则是通过单演信号尺度空间不同类型信息的相互融合来实现。

对于分类学习，论文拟采用稀疏信号表示进行改进，使得单演信号特征的鉴别性能够最大程度地被发挥。需要面临或是解决的主要问题包括：一是如何构造多元线性回归模型，即利用构造的过完备冗余字典对未知样本进行线性表示；二是如何将单演信号生成的众多局部分量与稀疏表示分类相结合；三是如何根据最优表示系数进行分类判决。

1.3 论文的主要工作及创新

论文的研究工作围绕着“高分辨率对地观测系统重大专项”“XXXX 目标检测与识别技术”子项目和装备预先研究演示验证项目“XXXX 高分辨率 SAR 图像 XXXX 判读系统”等课题展开。针对 SAR 图像目标识别面临的问题，以设计普适、简便、有效的特征描述方法和高效的分类学习方案为目的，重点研究借助

单演信号描述 SAR 图像目标散射现象，以及如何围绕着提出的特征描述方法设计分类学习方案，判决类别归属，并借助 MSTAR 实测数据进行验证。

1.3.1 论文的主要工作

论文围绕 SAR 图像自动目标识别的特征描述和分类学习展开。针对经典的特征计算复杂、中间环节繁多、扩展性不强等问题，引入单演信号；针对传统的分类学习效果不好，分类精度过度依赖于特征描述，引入稀疏表示；将原始 SAR 图像强度矩阵视为二维有限长信号，利用 Riesz 变换生成单演信号表示，通过信号分解生成不同尺度空间的局部分量，借助机器学习技巧实现信息融合，解决扩展工作条件的 SAR 图像目标识别问题。论文的主要框架如图 1.25 所示，主要内容包括：

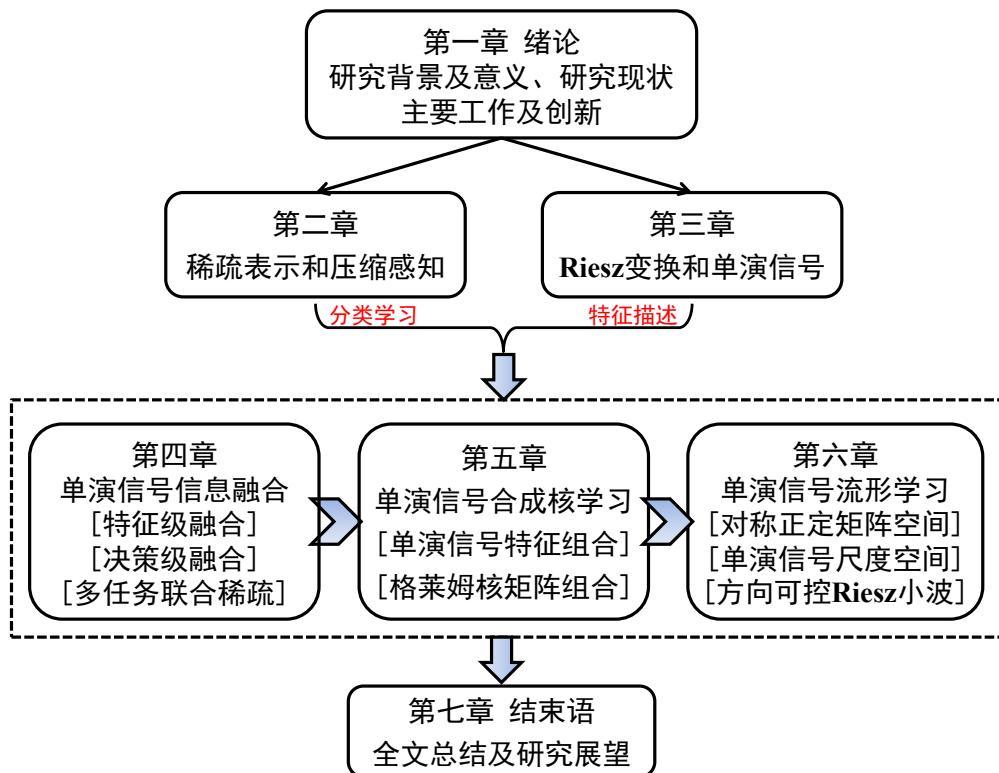


图 1.25 论文的组织结构图

- 介绍自动目标识别的研究背景及意义。首先通过近些年的局部战争引出对国防建设的思考，指出自动目标识别对军队信息化的重要意义。继而阐述自动目标识别的处理流程，总结现有的 SAR 图像目标识别方法，提炼自动目标识别涉及的关键环节，据此引出本文的研究主线和主要贡献。
- 介绍稀疏表示的相关背景知识。1) 从传统的正交基信号表示引出过完备冗余字典的信号表示，总结冗余字典的生成方式；2) 围绕着稀疏表示遇到的

“瓶颈”引出压缩感知，介绍它的应用前提、约束条件及典型算法；3) 总结稀疏表示分类的简要流程和数值实现；3) 阐述稀疏表示在 SAR 图像目标识别中的应用可行性；4) 针对 SAR 图像目标识别的实际，提出改进方案，围绕着表示方法提出频域稀疏表示，围绕着决策规则提出 DS 证据“软”决策。

- 介绍单演信号基础知识。1) 从传统的 Hilbert 变换引出 Riesz 变换，归纳它的诸多优良特性；2) 利用 Riesz 变换构造高维解析信号即单演信号，利用带通滤波器组实现单演信号多尺度表示；3) 分析单演信号多尺度表示描述 SAR 图像目标散射特性具有的优势。
- 借鉴经典数据融合技巧，设计基于图像域信息融合的分类识别框架，并在该框架下提出三种具体实现算法：特征级融合，将单演信号多尺度表示生成的分量进行降采样、归一化等处理生成增强的特征向量，利用生成的特征构造稀疏表示模型；决策级融合，分别设计面向单演信号各分量的稀疏线性表示模型，借助统计工具将各模型的输出进行融合，得出联合决策；联合稀疏表示，利用单演信号各分量构造多任务学习系统，通过施加特定约束项使每个子系统的表示系数具有类似的稀疏项，从而提取各任务之间的内在关联性。
- 借助多核合成学习实现 Hilbert 空间信息融合。通过构造非线性映射将单演信号分量投影到 Hilbert 空间，借助多核合成学习将面向不同分量的 Hilbert 子空间合并成全局核空间，合并不同类别的信息。包括：1) 再生核 Hilbert 空间背景知识，正定核函数及其特性；2) 提出两种具体方案，即单演信号特征组合和格莱姆核矩阵组合，实现“先合并再映射”和“先映射再合并”两种融合思路。
- 借助流形学习实现目标识别。将单演信号生成的多尺度序列整体视为黎曼流形的点，通过黎曼测度量化元素之间的相似性，实现分类识别；将流形嵌入到再生核 Hilbert 空间，构造核线性表示模型，利用黎曼核函数实现核空间像与像的相似比较，包括：1) 提出单演信号区域协方差描述子，并根据其非欧氏特性设计学习方案；2) 将黎曼流形嵌入到再生核 Hilbert 空间，根据核空间的像构造稀疏表示模型，实现分类学习；3) 将单演信号多尺度序列构造为 Grassmann 流形的点，通过 Grassmann 距离量化元素之间的相似性；4) 利用方向可控 Riesz 小波框架构造 Grassmann 流形，通过尺度间距离函数的累加定义全局度量；5) 将 Grassmann 流形嵌入到再生核 Hilbert 空间，根据核空间的像构造稀疏表示模型，利用 Grassmann 核函数实现问题求解。

1.3.2 论文的创新点

为了实现稳健的 SAR 图像目标识别，论文引入稀疏表示改进分类学习环节，引入单演信号实现目标散射特性的不变描述，其中稀疏表示是研究基础，它为后续分类框架的提出奠定基础，单演信号表示是论文的研究重点，它是论文的主要创新所在。围绕着如何依据单演信号设计特征描述算子以及如何利用设计的特征实现分类学习，论文分阶段、分层次逐步展开，按照“梳理总结成果 \mapsto 寻找存在不足 \mapsto 提出解决方法 \mapsto 实测数据验证”的流程完成既定目标。归纳起来，创新贡献包括：

1. 将稀疏表示引入 SAR 图像目标识别，对其可行性进行详细分析和充分论证，根据目标识别的实际，分别从表示方式和决策规则进行改进，提出频域稀疏表示和 DS 证据“软”决策；
2. 为了增强扩展工作条件的耐受性，论文引入单演信号用于 SAR 图像目标特性的描述；借助带通滤波器生成单演信号尺度空间，得到多个代表不同类型信息的局部分量；借鉴数据融合技巧提出基于空域信息融合的分类思路，依据该思路形成三种方案：特征级融合、决策级融合和多特征多任务联合稀疏表示。
3. 针对类间线性不可分的情况，构造非线性映射将单演信号投影至 Hilbert 空间，利用核空间的像构造线性表示模型，利用正定核函数的特性提出基于多核合成学习的分类框架，围绕着该框架形成两种具体解决方案，即单演信号特征组合和格莱姆核矩阵组合，分别反映“先合后映射”和“先映射后合成”两种思路。
4. 为了探索数据内部的非线性几何特性，将单演信号多尺度序列形成的线性子空间提升到流形，借助流形学习实现目标分类识别，在该思路指导下提出了三种实施方案：一是根据单演信号多尺度序列构造区域协方差矩阵描述子，使用生成的协方差矩阵构造黎曼流形；二是根据单演信号多尺度表示系数矩阵构造 Grassmann 流形，利用 Grassmann 度量实现分类学习；三是借助方向可控 Riesz 小波构建 Grassmann 流形，根据尺度间的距离累加定义全局度量，根据尺度内正定核函数的累加构成全局核函数。实现分类识别。

第二章 基于稀疏表征的 SAR 图像目标分类可行性研究

传统的信号获取与处理框架是以奈奎斯特采样定理 (Nyquist) 为基础，要精确重构带宽有限信号，采样速率必须达到信号带宽的两倍以上，即奈奎斯特采样率。随着现代社会对于巨量信息的需求，携带信息的信号带宽飞速增加，通信和感知系统的容量以前所未有的速度在增长，以奈奎斯特采样为基础的信号处理框架对宽带信号的处理其困难日益显现，比如遥感图像解译遇到的突出困难就是海量的高分辨率对地观测数据，如何执行传输与存储。为了处理成指数倍增长的大量数据，一种能够提供信号的稀疏或压缩表示的新技术亟待开发。稀疏表示理论就是在这种背景下应运而生。所谓的稀疏表示是指将一个 N 维空间的信号表示成 M 个基原子的线性组合，其中 M 远小于 N 。目前所接触的信号或时域稀疏，或频域稀疏，或是在某个基空间上可以稀疏展开，这为信号的稀疏表示提供了理论可能。之后，Bruno A. Olshausen 和 David J. Field 将稀疏表示引入计算机视觉，他们通过实验证明哺乳动物的细胞接收单元可被描述为局部的、有选择的、带通的过滤行为，类似于小波变换中的基原子，那种接收属性可以解释为动物大脑对于自然图像的输出活动的稀疏响应，转化为数学表述就是在最小二乘的基础上增加正则化，得到具有方向特性的图像块，从而解释了哺乳动物初级视觉皮层 (V1) 的工作机理，即对于线段的方向选择特性 [154, 168]。

然而稀疏表示在实际应用中遇到许多“瓶颈”问题，为了解决这些实际问题，近年来形成了另一种崭新理论即压缩感知 (Compressed Sensing 或 Compressive Sampling)，它是针对信号采样与重构所提出的，打破传统的奈奎斯特采样的限制，建立全新的信号描述和处理框架，能有效缓解巨量数据存储和传输的压力，是由 David L. Donoho、Emmanuel J. Candès、Terrence Tao 等数学家于 2005 年到 2006 年间共同提出 [158, 169, 170]。与奈奎斯特采样不同，压缩感知指出任意信号如果其本身可压缩，或是在某个变换域空间满足稀疏条件，就可以用一个与变换基不相关的观测矩阵将变换后的高维信号投影到一个低维子空间中，然后通过求解最优化问题从远小于奈奎斯特采样的少量测量值中以高概率重构原始信号，概括起来就是“通过对信号的高度不完备线性测量的高精确重建”¹。压缩感知理论框架中采样速率取决于信息在信号中的结构，而非信号带宽。尽管时域信号大都不是稀疏，但却可以寻找某个稀疏基使得信号在该基空间上稀疏，常见的稀疏基有小波基、离散余弦变换基、沃尔什变换、傅立叶变换等，这极大地扩展了压缩感知的

¹参考网友 Rachel-Zhang 的 CSDN 博客：<http://blog.csdn.net/abcjennifer/article/details/7721834>。

应用范围。压缩感知为信号处理提供了一种全新的思路，提出以来，迅速得到了研究人员的青睐，席卷了信号（图像）处理的多个应用领域。

本章论文首先按照“小波变换→多尺度分析→压缩感知”的顺序对信号紧凑表示的发展历程作以简要回顾，重点介绍了稀疏表示的最新研究成果—压缩感知，阐述其对于信号稀疏表征的推动作用，继而对于稀疏表征在 SAR 图像目标分类中的可行性进行了重点研究，并结合实际提出改进方案，借助实测 SAR 数据验证方案的合理性和有效性。本章内容主要包括：1) 稀疏表示基础理论；2) 压缩感知以及它所解决的主要问题；3) 基于稀疏表示的分类学习；4) 稀疏表示对于 SAR 图像目标分类的可行性分析。

2.1 稀疏表示

稀疏表示是指未知的观测信号可用特定空间中的极少数基原子的线性组合来进行重构，因此，需要对线性空间及相关的背景知识作以简要回顾。

2.1.1 线性空间以及线性空间的基

线性代数通过 n 维向量空间 \mathbb{R}^n （或 \mathbb{C}^n ）定义线性空间，线性空间包含两大基本要素，集合本身以及赋予集合的线性运算结构。不同的集合与不同的线性运算结构构成了不同的线性空间。

定义 2.1(线性空间): 设 X 为一非空集合， \mathbb{F} 为数域，对集合 X 中的任意两个元素定义加法“+”；对数域 \mathbb{F} 中的元素和集合 X 中的元素定义乘法“.”（符号可省略），且满足：对于任意 $x, y \in X, \alpha, \beta \in \mathbb{F}$ ，有 1) 交换律： $x + y = y + x$ ；2) 结合律： $(x + y) + z = x + (y + z)$ ；3) 存在零元 $0 \in X$ ，使得对一切 $x \in X$ 有 $x + 0 = x$ ；4) 存在负元，即对任意 $x \in X$ ，存在 $y \in X$ ，使得 $x + y = 0$ ；5) 存在单位元 $1x = x$ ；6) $\alpha(\beta x) = (\alpha\beta)x$ ；7) $\alpha(x + y) = \alpha x + \alpha y$ ；8) $(\alpha + \beta)x = \alpha x + \beta x$ ；称 $(X, \mathbb{F}, +, \cdot)$ 为线性空间（简记为 X ）。

若数域为 \mathbb{R} 或 \mathbb{C} ，称 X 为实线性空间或复线性空间。线性空间中的元素称作向量，定义在线性空间的两种运算结构称为线性运算。

定义 2.2(线性子空间): 设 Y 是线性空间 X 的一个非空集合，满足对任意 $x, y \in Y$ 及 $\alpha \in \mathbb{F}$ 有 $x + y \in Y, \alpha Y \in Y$ ，称 Y 是线性空间 X 的线性子空间。

从定义 (2.2) 可以看出，线性子空间本身也是线性空间，因此，关于线性空间的运算结构（加法和数乘）同样也满足于线性子空间 [171]。

推论 2.1 (最小线性子空间): 设 M 是线性空间 X 的非空子集,

$$\text{span}\{M\} = \{\alpha_1x_1 + \alpha_2x_2 + \cdots + \alpha_nx_n | n \in \mathbb{Z}_+, x_i \in M, \alpha_i \in \mathbb{F}\}, \quad (2.1)$$

是包含 M 的最小线性子空间, 亦即是一切包含 M 的线性子空间的交集, 故又称 $\text{span}\{M\}$ 是 M 的线性包或由 M 生成的子空间。

由推论 2.1 可知, 任意的线性空间总可以由相对较少的若干个元素线性生成, 如 n 维向量空间 \mathbb{R}^n 中有无穷多个向量, 但是它却可以由 n 个特殊的向量张成, 这 n 个向量需要做一定的约束, 即线性无关 [171]。

定义 2.3 (线性无关): 设 X 为线性空间, $M \subset X$ 是非空子集, 称 M 是线性无关,

- 若 $M = \{x_1, x_2, \dots, x_n\}$ 为有限集, $\sum_{i=1}^n \alpha_i x_i = 0$ 当且仅当 α_i 全为 0, $\alpha_i \in \mathbb{F}$;
- 若 M 为无穷集, M 的任意有限子集都是线性无关。

定义 2.4 (维数): 设 X 为线性空间, 若存在 $n \in \mathbb{Z}_+$, 使

- X 包含了一个由 n 个元素组成的线性无关集;
- 任何多于 n 个元素的集合都是线性相关的;

称 n 为 X 的维数, 记为 $\dim\{X\} = n$, 若 X 不存在这样的 n , 称 X 是无穷维。

与维数有密切联系的是线性空间的基。

定义 2.5 (基): 设 X 为线性空间, \mathbb{F} 是数域, $B \subset X$ 是线性无关集, 若 $\text{span}\{B\} = X$, 即对任意 $x \in X$, 存在 $n \in \mathbb{Z}_+$ 及 $x_1, x_2, \dots, x_n \in B$ 和 $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{F}$, 使得 $x = \sum_{i=1}^n \alpha_i x_i$, 则称 B 是 X 的一个基 (或 Hamel 基)。

本文只讨论有限维线性空间, 因而, 线性无关元素的个数即为线性空间的维数, 线性空间的基可由施密特正交化定理得到。由定义 2.1 和推论 2.1 可知, n 维线性空间可由 n 个线性无关向量 $\Psi = \{\psi_i\}_{i=1}^n$ 所张成, 这一组向量即称之为线性空间的基, 线性空间中的任意元素 (信号) $\mathbf{x} \in \mathbb{R}^n$ 可由这一组基原子的线性组合来表示, $\mathbf{x} = \sum_{i=1}^n \alpha_i \psi_i$, 其中表示系数 α_i 对于特定信号 \mathbf{x} 是唯一确定的, 将其改写成矩阵形式 $\mathbf{x} = \Psi \alpha$, Ψ 的所有原子两两正交,

$$\begin{aligned} \langle \psi_i, \psi_j \rangle &= 1, \text{ for } i = j \\ \langle \psi_i, \psi_j \rangle &= 0, \text{ for } i \neq j \end{aligned} \quad (2.2)$$

线性表示系数 α 可通过简单的矩阵求逆解析得到

$$\alpha = \Psi^{-1} \mathbf{x} = \Psi^T \mathbf{x}, \quad (2.3)$$

Ψ 是正定矩阵，所以有 $\Psi^{-1} = \Psi^T$ 。然而，这仅仅是从计算角度的一种理想情况，正定矩阵的要求过于苛刻，实际应用中很难满足，因而应用范围有限。

20世纪90年代，法国科学家 Mallat 提出了一种基于冗余字典的稀疏分解思想 [172]。该思想是一种新的信号表示理论，它使用超完备的冗余函数族取代传统的正交基函数族，从而极大地丰富了信号的自适应稀疏表示形式。一方面这可以实现高效的数据压缩，更为重要的是可利用字典的冗余特性捕捉原始信号的自然特征。因而，论文转而讨论这种更普遍的情况，即包含许多冗余原子的线性组合，也即使用矩阵 $\Psi \in \mathbb{R}^{m \times n}$ 表示向量（或信号） \mathbf{x} 。矩阵原子 $\{\psi_i \in \mathbb{R}^m\}_{i=1}^n$ 所张成的 m 维线性空间 $\mathbb{R}^{m \times n}$ ，任意元素 $\mathbf{x} \in \mathbb{R}^m$ 满足结论 [173]

$$A\|\mathbf{x}\|_2^2 \leq \|\Psi^T \mathbf{x}\|_2^2 \leq B\|\mathbf{x}\|_2^2, \quad (2.4)$$

其中 A 和 B 分别表示正交矩阵 $\Psi\Psi^T$ 的最小特征值和最大特征值，且有 $0 < A < B < \infty$ 。

调和分析理论将张成线性空间的基向量所构成的矩阵称作为字典，广义地讲， n 个线性无关的 m 维元素或者向量 ($m < n$) 所构成的矩阵通常被称作是过完备冗余字典，然后问题就转化成为根据已知字典寻找观测信号的线性表示

$$\mathbf{x} = \Psi\alpha, \quad (2.5)$$

其中 Ψ 是冗余字典，其维数小于原子的个数， $\alpha \in \mathbb{R}^n$ 是表示系数。

2.1.2 过完备冗余字典

所谓稀疏（或冗余）表示指使用特定字典中的极少数原子的线性组合对观测信号进行表示。因此，字典的选择对于稀疏表示的成功与否有着非常关键的作用。目前，对于过完备冗余字典的构成通过有两大类方法，一是预定义，根据某种数学模型所生成的序列解析地构成，如小波字典、离散余弦字典、Curvelet 字典等；二是学习，针对给定的数据集按照一定的最优化准则在线学习生成。预定义的解析字典通用性较好且计算简单，而学习字典弹性更强，能很好地适应特定的数据集或特定的任务。Ron Rubinstein 等对于稀疏表示模型中的字典生成方法进行详细地归纳与总结 [174]。

2.1.2.1 解析字典

解析字典是指通过采样特定函数的某些参数所形成的数值序列构成预定义矩阵。其优点是计算简便、易于执行，可离线生成。常见的解析字典生成方法包括：

- 离散余弦变换（Discrete Cosine Transform）的基函数（ $\mathbb{R}^{m \times n}$ 维）定义为

$$\phi_{l,k}(x, y) = \cos\left(\frac{(2x+1)k\pi}{2m}\right) \cos\left(\frac{(2y+1)l\pi}{2n}\right)$$

其中 $0 \leq x \leq m - 1, 0 \leq y \leq n - 1$ 表示信号空间位置索引, $0 \leq k \leq m - 1, 0 \leq l \leq n - 1$ 是频率位置索引。通过对不同频率的采样所生成的原子所构成的矩阵作为过完备冗余字典, 其中每个采样得到的原子重新排列为 mn 维列向量。

- 小波变换 (Wavelet Transform) 的基函数为 $\psi_{k,j}(x) = 2^{\frac{j}{2}}\psi(2^j x - k)$, 其中 $\psi(x)$ 是母函数, 常见的母函数有 haar 小波、Daubechies 小波、Meyer 小波等。以 Morlet 小波为例, $\psi_{k,j}(x) = \sin(5x) \exp\left\{-\frac{x^2}{2}\right\}$, 二维小波变换为

$$\psi_{k_1,j_1,k_2,j_2}(x,y) = 2^{\frac{j_1}{2}} \sin\left(5(2^{j_1}x - k_1)\right) \times 2^{\frac{j_2}{2}} \sin\left(5(2^{j_2}y - k_2)\right)$$

通过采样频率索引 k, j 得到函数序列, 并将其重新排成列向量组合起来构成的矩阵即为过完备冗余字典。

- 脊波变换 (Ridgelet Transform) 是由小波变换衍生而来, 其定义为

$$\psi_{a,b,\theta}(x,y) = a^{-\frac{1}{2}}\psi\left(\frac{x \cos \theta + y \sin \theta - b}{a}\right)$$

其中 $\psi(\cdot)$ 是小波母函数。通过对参数 a, b, θ 不同频率点进行采样得到一组脊波基序列, 即可构成过完备冗余字典。

- 曲线波变换 (Curvelet Transform) 是 Candes 和 Donoho 提出的一种连续变换, 它能以最优变化率描述具有光滑曲线跳跃性的 2 维点光滑函数。其离散化变换为

$$\psi_{j,k,l}(x) = \psi(R_{\theta_l}(x - x_k^{j,l}))$$

其中 R_θ 是以弧度 θ 进行旋转, 旋转角度的平均量化区间为 $\theta_l = 2\pi 2^{\lfloor \frac{l}{2} \rfloor} l, l = 0, 1, \dots$, $\lfloor \cdot \rfloor$ 表示取整运算。通过采样 j, k, l 可得到一族曲线波基, 将其重排并组合即可得到过完备字典。

- 轮廓波变换 (Contourlet Transform) 是 Minh N. Do 和 Martin Vetterli 提出 [175], 它利用拉普拉斯塔形分解 (Laplace Pyramids) 和方向滤波器线 (Directional Filter Bank) 实现一种多分辨率、局部的、多方向图像表示方法。它继承了曲线波变换的各向异性的多尺度关系, 可认为是曲线波变换的一种快速有效数字实现。通过在不同频点的采样就能生成轮廓波字典。

尽管由某种数学模型生成的函数序列构成的过完备冗余字典计算简单, 但是它也存在着明显的不足, 不能很好地表述复杂的自然信号 [174], 因而, 科学家提出另一种思路, 即利用给定的标记信号, 依据一定的最优化准则进行学习, 从而生成字典。

2.1.2.2 学习字典

学习字典是近年来才提出的一种字典设计方法，它对稀疏表示的进一步扩展产生了重要作用。近年来的字典学习方法大多都聚焦于 ℓ_0 范数和 ℓ_1 范数的稀疏度量，它借鉴了许多稀疏编码方面的最新研究成果。通过学习生成的字典其最主要的优点在于处理许多实际应用问题，然而，为之而付出的代价就是高额的计算复杂度。目前，一种借助参数化模型所生成结构化字典开始受到关注，这或许是未来字典学习的一个重要研究方向 [160, 176, 177]。

- **Method of Optimal Direction** MOD 是由 K. Engan 等于 1999 年提出的稀疏化执行算法 [178, 179]，它为后续的学习方法提供了很好的研究思路。任意给定一组标记样本 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ ，MOD 算法的宗旨是对给定样本表示误差最小的前提之下生成一个字典和一个系数矩阵²，

$$\arg \min \|\mathbf{X} - \mathbf{D}\Gamma\|_F^2 \quad s.t. \quad \|\gamma_i\|_0 \leq L \quad \forall i \quad (2.6)$$

其中 $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots]$ 是学习生成的字典， Γ 是由系统矩阵，它的每个列向量都是稀疏的， $\|\cdot\|_F$ 是矩阵的 F 范数， $\|\cdot\|_0$ 是指向量 0 范数（或称为伪范数），计数向量的非零元素个数。MOD 算法需要求解组合优化问题，且目标函数完全非凸，因此无法求取全局最优解，只能通过近似得到其局部最优解。与此前的组合优化问题的求解类似，MOD 算法通过交替执行稀疏编码和字典更新两个步骤来得到局部最优解。稀疏编码环节，首先预设一个初始字典，使用正交匹配追踪算法单独求取每个给定信号的稀疏表示，之后利用求解得到的表示系数更新字典，则通过最小化表示误差得到新的字典， $\mathbf{D} = \mathbf{X}\Gamma^{-1}$ （矩阵的 Moor-Penrose 伪逆）。MOD 算法只需要少数几次迭代即可收敛，因而整体计算效率较高。然而，该方法涉及高维矩阵求逆，计算复杂度较高，且算法的数值实现容易陷入异常。

- **Union of Orthobases** 是 S. Lesage 等于 2005 年提出 [180]，其目的是为了降低 MOD 算法的计算复杂度。该算法尝试学习生成一种结构化的过完备字典，它的思路是将多个正交基矩阵的组合 $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_k]$ 进行组合，其中 \mathbf{D}_i 是酉矩阵，从而生成冗余字典。该算法采用块坐标松弛优化算法，因而生成的稀疏编码模型执行效率较高。

然而，该方法也存在着明显的缺点，即它的约束条件（酉矩阵）太过于苛刻，该方法所生成的冗余字典是通过约束子字典 $\{\mathbf{D}_i\}$ 的正交性来实现，实际应用中很难展现出较强的扩展性，因而对于复杂信号的描述缺乏灵活性。

²s.t. 表示使得满足，*subject to* 的缩写。

- **Generalized Principal Component Analysis** 是 R. Vidal 等于 2005 年提出来的 [181]，是一种字典学习的新思路，它从一个低维线性子空间中近似得出一组原子样本，将样本建模为数个低维子空间的组合，子空间个数未知，且它的维数可以发生变化。

与经典的表示模型不同，广义主成分分析认为每个样本仅仅是通过众多子空间中的某一个来进行线性表示，它不允许不同子空间中的原子联合起来表示某个特定信号。该思路生成的字典能有效避免过度描述问题，然而这种独特的逻辑架构使它存在着致命的缺陷，若使用该方法描述复杂的自然信号，生成的冗余字典的结构将会非常复杂。另一方面，广义主成分分析能够判定特定场合字典中的原子数目，但是它的计算开销不容忽视，特别是当子空间的数目和维数增加时，其计算复杂度将难以承受。

将广义主成分分析所生成的稀疏表示与经典的稀疏表示相结合是很有前景的应用思路，因为在广义主成分分析中允许多个截然不同类型的字典同时共存，对于任意信号，都假定它可以用其中的某个字典进行稀疏表示。

- **K-SVD** 是迄今为止应用最广泛、技术最成熟的字典学习方法。它由以色列理工学院的 Michal Aharon 等学者于 2005 年提出 [182]，其初衷是为了寻找一种针对稀疏信号表示的通用字典学习方法，是对经典的 k 均值聚类算法的广义扩展。与 MOD 算法相同，K-SVD 也使用类似的块松弛方法求解非凸优化问题，不同之处是 K-SVD 提出一种更简单、更高效的逐原子字典更新方法，取代 MOD 算法的矩阵伪逆运算。为了提高运算效率，K-SVD 同时更新当前字典原子和与之相关的稀疏表示系数，因而计算速度和效率都优于传统的 MOD 算法。

K-SVD 算法中字典原子更新的核心是连续 K 次的矩阵奇异值分解，这也是该方法的命名由来。对于给定的第 k 个原子，通过借鉴 MOD 算法思路，将优化问题的二次项改写为

$$\|\mathbf{X} - \sum_{j \neq k} \mathbf{d}_j \gamma_j^T\|_F^2 = \|\mathbf{E}_k - \mathbf{d}_k \gamma_k^T\|_F^2 \quad (2.7)$$

其中 γ_j^T 表示系数矩阵的第 j 行， \mathbf{E}_k 表示第 k 个原子更新时的重构误差矩阵。原子的更新是通过调节 \mathbf{d}_k 和 γ_j^T 从而使得式 (2.7) 最小化来实现。为了避免在系数矩阵中引入新的非零项，原子的更新过程仅仅只涉及到参与当前原子稀疏表示的样本（即非零系数项对应的原子）。

实际应用中 K-SVD 只是对低维信号的表示效果较好 [174]，与 MOD 算法类似，它也存在诸多不足，其目标函数的非凸性决定其只能得到局部最优甚至是鞍部极值，很难得到全局最优解。针对其缺点，研究人员也提出了许多改进方法，如 multilevel K-SVD, discriminative K-SVD, Label Consistent K-SVD 等 [183, 184]。

图 2.1、图 2.2 和图 2.3 通过一组字典学习的实例将 MOD 和 K-SVD 算法进行比较。首先从图 2.1 Bradatz 植被纹理库³随机选出一组大小为 8×8 像素的子块纹理，使用其中 121 组子块的离散余弦变换作为初始字典原子，然后分别使用 MOD 算法和 KSVD 算法进行字典学习，最终生成的字典如图 2.3 所示。分析学习结果，MOD 算法和 KSVD 算法最终生成的子矩阵都是分段光滑的纹理化原子，体现图像局部角点、边缘等显著性特征，且部分原子十分相似。图 2.2 是两种算法的平均表示误差，通过比较初始误差与迭代 50 次之后的误差可知：离散余弦变换生成字典的表示误差（即初始迭代误差）远远大于 MOD 算法和 KSVD 算法学习生成的字典，而 MOD 算法与 KSVD 算法的平均表示误差几乎相等，然而 MOD 算法涉及矩阵求逆（时间复杂度为 $O(n^3)$ ），KSVD 算法需要执行矩阵奇异值分解（时间复杂度为 $O(n^3)$ ），若增大字典原子的维数，则计算开销将难以负担。前面的例子仅仅使用了 Grass 纹理图中很少部分子块（大约 10%）进行学习，接下来将再次进行实验，使用全部原子进行学习，最终经过 50 次迭代 MOD 算法的平均表示误差为 22.2771，KSVD 算法的平均表示误差为 22.6174，可以看出两种学习方法没有明显的优劣差别。

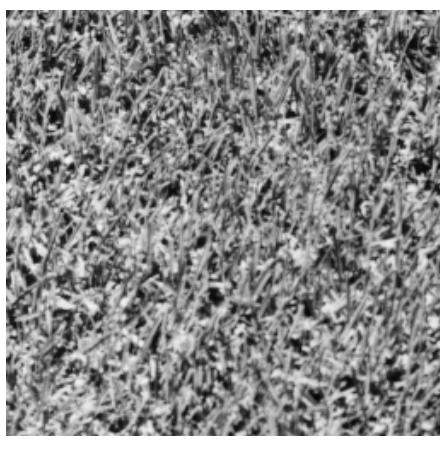


图 2.1 Brodatz 库的 Grass 纹理

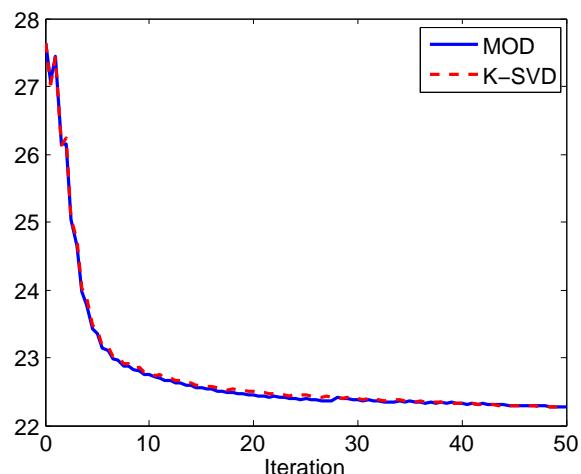


图 2.2 MOD、KSVD 算法的平均表示误差

³Brodatz 纹理库：<http://sipi.usc.edu/database/database.php>，图片的原始大小为 512×512 。

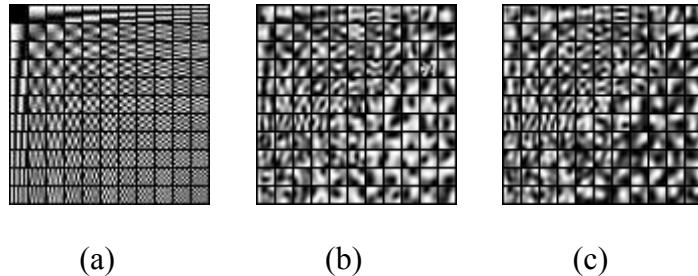


图 2.3 图 (a) 离散余弦变换生成的初始字典, 图 (b) MOD 学习算法学习的字典, 图 (c) 是 KSVD 学习算法学习的字典。

- **Parametric Training Methods**—为了加速收敛、降低搜索到局部最优的概率, 研究人员通过在字典上施加各种约束条件、减少自由参数数目, 期望得到理想的结果, 这就是参数化的字典学习。即使是少数的参数也能够改进学习过程的普适性, 并减少所需的样本数目。参数化的另外一个优点就是能生成更紧凑的表示, 因而执行效率更高。此外, 参数化的字典学习还可设计用于无限长或是任意长信号。近年来提出的参数化的字典学习方法有: 平移不变性字典, 即通过收集所有训练样本的平移变换结果构成的任意尺寸字典 [185]; 多尺度字典, 就是借助类似小波变换的金字塔信号分解生成字典 [186, 187]; 稀疏字典, 如由解析信号所构成的字典与稀疏矩阵的组合 [188]。
- **Online Disctionary Learning**—法国学者 Julien Mairal 和 Francis Bach 提出一种在线字典学习的思路 [189–191], 他们采用随机近似的优化算法求解组合优化问题, 实现从海量的训练样本中快速学习生成能够适应稀疏表示的过完备冗余字典。同时, Yannis Kopsinis 等学者也开展了相关的研究 [192]。

根据预定义或学习的方式生成冗余字典之后, 稀疏信号表示就可描述为已知观测信号 \mathbf{x} 和过完备冗余字典 \mathbf{D} 求解未知的表示系数 α , 即

$$\min_{\alpha} O(\alpha) \quad s.t. \quad \mathbf{x} = \mathbf{D}\alpha \quad (2.8)$$

其中 $O(\alpha)$ 是实现稀疏的目标函数, $\mathbf{x} = \mathbf{D}\alpha$ 是约束条件, 即满足精确重构原始信号。故问题的关键就是如何具体化目标函数, 通常期望系数向量 α 非零项的数目相比较向量长度能足够小, 参与信号 \mathbf{x} 线性表示的原子尽可能少, 即“稀疏”, 机器学习领域也称为稀疏编码 (Sparse Coding)。基于该思想, 一种合理的目标函数就是容忍大的表示系数, 而惩罚小的表示系数, 也就是 ℓ_p 范数约束的目标函数。围绕着这些问题的研究就引出另一种理论成果, 压缩感知。

2.2 压缩感知

研究表明一维信号 $\mathbf{x} \in \mathbb{R}^n$ 通常都含有冗余信息，为了提高存储和传输效率，需要执行信号压缩，传统思路是借助正交变换，即构造矩阵 $\Psi \in \mathbb{R}^{n \times n}$ ，信号压缩的正变换表示为 $\mathbf{y} = \Psi\mathbf{x}$ ，信号重构的逆变换表示为 $\mathbf{x} = \mathbf{D}^H\mathbf{y}$ ，其中变换矩阵 Ψ 满足条件： $\Psi\Psi^H = \Psi^H\Psi = \mathbf{I}$ (\mathbf{I} 是单位矩阵)。变换生成的信号 \mathbf{y} 能去除原始信号的相关性，使得能量更加集中，因此只需要保留 k 个较大分量，直接将剩余 $n - k$ 项置零，通过逆变换就能近乎完整地重构原始信号，因为舍弃的 $n - k$ 项系数对于原始信号的贡献微乎其微。满足这种重构框架的信号被称为 k 稀疏 (k-Sparse)。正交变换的编码策略包括⁴：

Encode 编码：构造正交矩阵 Ψ ，执行正交变换 $\mathbf{y} = \Psi\mathbf{x}$ ，变换后的信号 \mathbf{y} 仅保留最重要的 k 个分量及其对应的位置。

Decode 解码：将信号 \mathbf{y} 中最重要的 k 个分量位置不变，其它位置的系数全部置零，构造共轭转置矩阵 Ψ^H ，执行逆变换 $\hat{\mathbf{x}} = \mathbf{D}^H\hat{\mathbf{y}}$ 恢复 \mathbf{x} ，恢复原始信号。

信号压缩的优劣通常使用解码重构误差 $\|\mathbf{x} - \hat{\mathbf{x}}\|_2 = \|\mathbf{y} - \hat{\mathbf{y}}\|_2$ 来衡量，通常期望该误差尽可能小，更合理的评判指标是相对误差 $\frac{\|\mathbf{y} - \hat{\mathbf{y}}\|_2}{\|\mathbf{y}\|_2} \leq \delta$ (δ 是误差容限)。传统的信号压缩方法通过构造正交矩阵进行编码，需要将所有 n 维信号全部存储下来，其弊端包括：

- 受奈奎斯特采样定理限制，只能通过较小的采样间隔换取较高的信号分辨率，若减少采样数据量必然以牺牲分辨率代价，所以若要高精度重现原始信号就会消耗大量时间和空间；
- k 个保留的重要分量的位置随着信号的不同而不同，这种“自适应”策略需要额外分配存储空间；
- k 个保留下来的分量如果在传输过程中丢失，很难精确恢复原始信号；

为了解决传统的压缩方法遇到的问题，Emmanuel Candes，David Donoho 和 Terrence Tao 等数学家共同提出了一种新颖的信号处理框架即压缩感知，其核心是通过对信号的高度不完备线性测量的高精确的重建 [169, 170]。

⁴ 参照了香港大学电机电子工程系 Dr. Wei E.I. Sha 的在线公开课程。

2.2.1 工作原理

与传统的采样方式不同，压缩感知旨在解决过完备冗余字典的信号压缩问题。对于给定的 n 维的 k 稀疏信号 $\mathbf{x} \in \mathbb{R}^n$ ，若使用算子 $\Psi \in \mathbb{R}^{m \times n}$ 进行采样，得到的 m ($m < n$) 个采样值排列组成向量 $\mathbf{y} = \Psi \mathbf{x}$ ，其中 Ψ 称为采样矩阵。压缩感知的理论框架认为：只需要 m 个测量值而并非原来的 n 个样本且 $m \ll n$ 即可表示原始信号 \mathbf{x} 。⁵ 采样后的解码就是根据 m 个测量值 \mathbf{y} 精确重构原始信号 \mathbf{x} 。然而根据少数测量值 \mathbf{y} 重构原始信号 \mathbf{x} 是欠定的“病态”问题，不存在唯一解。通常希望向量 \mathbf{x} 中只有与原始信号相关的项为非零值，其余全部为零，根据这个思路最直接的方法就是约束 \mathbf{x} 的非零元个数⁵

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad s.t. \quad \mathbf{y} = \Psi \mathbf{x} \quad (2.9)$$

实质就是使用 ℓ_0 范数 $O(\mathbf{x}) = \|\mathbf{x}\|_0$ ，约束 \mathbf{x} 的可行集。以图 2.4 为例，原始信号的 3 个彩色方格表示它的非零元，对应位置的字典原子用黑框强调显示，只有这 3 个原子参与了观测信号的线性表示，需要解决的问题是如何根据给定的一给 m 维空间过完备基，寻找最少个数的基向量重构观测信号。

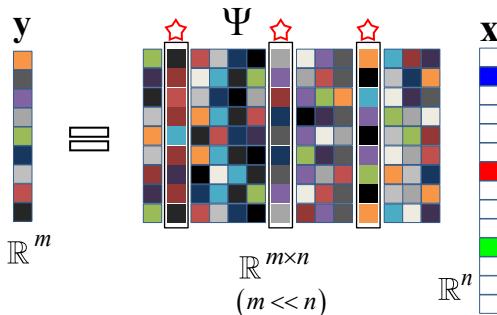


图 2.4 利用过完备字典 Ψ 对观测信号 \mathbf{y} 进行重构，生成原始稀疏信号 \mathbf{x} ，其中 \mathbf{x} 中仅有 3 项值非零，即字典中只有 3 个原子参与 \mathbf{y} 的表示。

然而目标函数 $O(\mathbf{x}) = \|\mathbf{x}\|_0$ 不连续且非凸，是个 NP 难⁶问题，直接求解比较困难，虽然使用穷举法、贪婪算法等策略能近似求解，但精度有限且计算复杂。数学家 David Donoho 和 Michael Elad 通过大量研究证明

定理 2.1：如果不等式 $\sigma(\Psi) \geq 2\|\mathbf{x}\|_0$ 成立，则最小化 ℓ_0 范数问题有唯一解，其中 $\sigma(\Psi)$ 表示字典中最小的线性相关的列向量集所含的向量个数。

⁵ $\|\cdot\|_0$ 是向量的零范数或是伪范数，用于计数向量的非零项的数目。

⁶NP-hard，NP 是是指非确定性多项式 (Non-deterministic Polynomial)。所谓的确定性多项式是指可用一定数量的运算去解决多项式时间内可解决的问题。

定理 2.1 的提出将问题 (2.9) 的求解又向前推进了一步，然而 $\sigma(\Psi)$ 的计算比较困难，实际应用中不具可操作性。2006 年，数学家 Terrence Tao 与 Emmanuel Candes 又证明了若字典 Ψ 满足约束等距性 (Restricted Isometric Property, RIP) 条件，最小化 ℓ_0 范数问题与最小化 ℓ_1 范数问题等价或是同解，

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad s.t. \quad \mathbf{y} = \Psi \mathbf{x} \quad (2.10)$$

(2.10) 式的目标函数 $O(\mathbf{x}) = \|\mathbf{x}\|_1$ 是连续可微且是凸函数。尽管最小化 ℓ_1 范数问题计算效率高于最小化 ℓ_0 范数问题，它需要的测量值的数目多于后者。D. Baron 等学者证明了求解 (2.10) 所需的最少的测量值数目满足 $m \approx \log_2(1 + \frac{m}{K})K$ [193]。John Wright 等还证明了非零值与测量值数目之间满足如下关系 $cm < K < \lfloor \frac{m+1}{3} \rfloor$ [156]，即只要 \mathbf{x} 中的非零数目远小于观测数目 m ，通过最小化 ℓ_1 范数完全可以重构原始信号。对图 2.4 所示的模型做进一步扩展，若原始 \mathbf{x} 并非稀疏，但它可压缩，即在某些变换域 $\Theta = [\theta_1, \theta_2, \dots, \theta_n]$ 稀疏，其中 θ_i 是变换域的基。假设 α 是 \mathbf{x} 在 Θ 域的表示系数，

$$\mathbf{x} = \Theta^{-1}\alpha = \Theta^T\alpha, \quad (2.11)$$

Θ 是正交矩阵，它满足 $\Theta^{-1} = \Theta^T$ 。然后，可压缩信号 \mathbf{x} 的采样过程就改写为

$$\mathbf{y} = \Psi \Theta^T \alpha = \mathbf{A} \alpha, \quad (2.12)$$

其中 \mathbf{A} 被称为测量矩阵。广义的压缩感知理论框架如图 2.5 所示。

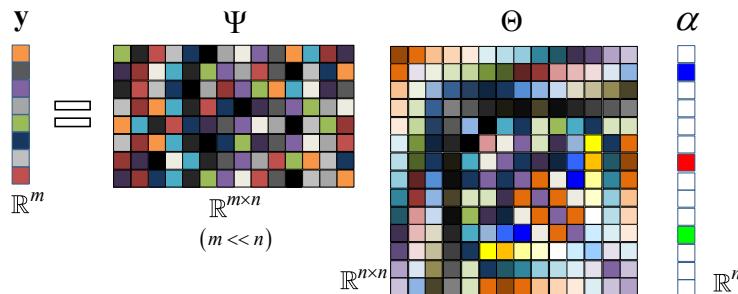


图 2.5 广义压缩感知原理。

类似地，问题的求解即改写为

$$\min \|\alpha\|_1 \quad s.t. \quad \mathbf{y} = \mathbf{A} \alpha, \quad (2.13)$$

根据计算得出 α 的值很容易重构原始信号 \mathbf{x} 。然而， ℓ_0 最小化和 ℓ_1 最小化同解是有前提的，即测量矩阵 \mathbf{A} 必须满足约束等距性，且采样矩阵 Ψ 和基矩阵 Θ 不相关，论文将在下文中进行讨论。

2.2.2 测量矩阵

在压缩感知的理论框架中，原始信号的测量过程不是自适应完成，测量矩阵的原子是预定义的，它不随前一次采样结果而发生变化。因而面临的问题就是如何设计测量矩阵使得它能尽可能地保留原始信号的信息，并确保原始信号得以高概率恢复。为了回答这个问题，数学家 Emmanuel Candes 和 Terrence Tao 证明了要从少数测量值中精确重构原始信号，测量矩阵必须满足约束等距性 [169]

定义 2.6 (约束等距性): 对于任意的 K 稀疏信号 \mathbf{x} 和常数 $\delta_k \in (0, 1)$ ，若不等式

$$1 - \delta_k \leq \frac{\|\mathbf{Ax}\|_2^2}{\|\mathbf{x}\|_2^2} \leq 1 + \delta_k \quad (2.14)$$

成立，称矩阵 \mathbf{A} 满足约束等距性。

从约束等距性的定义可以看出，要完全重构原始信号，必须保证观测矩阵不会把两个不同的 k 稀疏信号映射到同一个采样集合⁷，即从观测矩阵中抽取的任意 m 个列向量构成的矩阵非奇异。与约束等距性相关联的条件是不相干性，即感知矩阵的每一行不能对基矩阵的列向量做稀疏表示。

定理 2.2 (信号重构): 对于 K 稀疏的信号 \mathbf{x} ，假设测量矩阵 $\mathbf{A} \in \mathbb{R}^{m \times n}$ 中任意 $2K$ 列都是线性无关的，则可以根据测量值 $\mathbf{y} = \mathbf{Ax}$ 唯一地恢复原始信号 \mathbf{x} 。

定理 2.2 指出，完全可以仅仅使用 $2K$ 个测量值精确地感知 K 稀疏信号。数学家 David L. Donoho 给出测量矩阵的要求归纳为：

- 测量矩阵的列向量构成的子矩阵最小奇异值必须大于一定常数；
- 测量矩阵的列向量体现某种类似噪声的独立随机性；
- 满足稀疏度的解是满足 1 范数最小的向量。

若稀疏基和观测基不相关就能在很大程度上保证约束等距性。E. CandeS 和 T. Tao 证明，独立同分布的高斯随机测量矩阵可以成为普适的压缩感知测量矩阵。据此，莱斯大学 R. Baraniuk 教授等研制出单像素相机，麻省理工学院 L.L.Wald 教授等研制出 MRI RF 脉冲设备，伊利诺伊州立大学 O. Milenkovic 等研制出 DNA 微阵列传感器等。这些硬件的推出促进了压缩感知的实用化进程。目前最为常用的是高斯随机矩阵，除此之外还有其它一些生成方法，归纳起来有：

⁷可近似理解为避免出现多对一映射。

- 高斯随机矩阵：独立地服从均值为 0，方差为 $\frac{1}{\sqrt{m}}$ 的高斯分布。高斯随机测量矩阵几乎与任意稀疏信号都不相关，无论如何选择正交变换基 Θ ， $\Psi\Theta^T$ 仍以高概率满足约束等距性，因而使用较为广泛。
- 随机贝努利矩阵：独立服从等概率为 $\frac{1}{\sqrt{m}}$ 或是 $-\frac{1}{\sqrt{m}}$ 的对称贝努利分布。
- 正交变换矩阵：首先生成 $n \times n$ 维正交矩阵 \mathbf{U} ，再从矩阵 \mathbf{U} 中随机地选取 m 行向量，对 $m \times n$ 维矩阵的列向量进行单位化。
- 哈达玛矩阵：首先生成 $n \times n$ 维哈达玛矩阵，然后从生成的矩阵中随机地选取 m 行向量构成一个 $m \times n$ 维矩阵。
- 托普利兹循环矩阵：首先生成一个向量 \mathbf{u} ，再由向量 \mathbf{u} 生成相应的轮换矩阵或托普利兹矩阵 \mathbf{U} ，从矩阵 \mathbf{U} 中随机选取其中的 m 行构成矩阵。
- 稀疏随机矩阵：首先生成一个全零元素的矩阵，然后在该矩阵的每个列向量中，随机地选取若干个位置，将其值赋为 1。

2.2.3 重构算法

压缩感知最终要实现从不完全测量值中重构原始信号，或是使用过完备冗余字典对观测信号进行线性表示。确立约束等距性前提后，E. Candes 和 T. Tao 指出信号的重构可以通过最小化 ℓ_0 范数最优化来实现，

$$\min_{\alpha} \|\alpha\|_0 \quad s.t. \quad \mathbf{y} = \mathbf{A}\alpha \quad (2.15)$$

然而，最小化 ℓ_0 范数问题是个 NP 难的组合优化问题，若使用穷举法则有 C_K^n 种组合方法，显然在实际应用中不可行。因而，常见的解决方法主要分为两大类：

(1) 利用追踪算法获取次优解。通过选择一系列局部最优解来逐渐逼近原始信号，常见的方法包括 Matching Pursuit、Orthogonal Matching Pursuit、以及改进方法（如 Stage-wise OMP、Regularizer OMP）、Iterative Hard Thresholding 等。

(2) 凸松弛优化算法。根据 2.2.1 节的论述可知，若测量矩阵满足约束等距性，最小化 ℓ_0 范数与最小化 ℓ_1 范数有着相同的唯一解，

$$\min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \mathbf{y} = \mathbf{A}\alpha \quad (2.16)$$

目标函数 ℓ_1 范数最小化连续可微，是典型的凸优化问题，可实现多项式时间求解。常见的求解思路有：

- 直接优化：利用优化工具包直接求解，即 LASSO 模型 [194]；

- 给定拉格朗日乘子的无约束凸优化: $\min_{\alpha} \|\mathbf{y} - \mathbf{A}\alpha\|_2^2 + \lambda\|\alpha\|_1$;
- 未知拉格朗日乘子的无约束凸优化问题: $\min_{\alpha, \lambda} \|\mathbf{y} - \mathbf{A}\alpha\|_2^2 + \lambda\|\alpha\|_1$ 。

最小化 ℓ_1 范数的凸松弛求解有多种不同称法, 信号处理称为基追踪 Basis Pursuit, 应用统计称为 ℓ_1 正则化, 其实质都是同一种求解思路。除了 ℓ_1 凸松弛策略之外, 近年来也提出利用 $\ell_p, p \in (0, 1)$ 逐次逼近 ℓ_0 范数的方法等。图 2.6 对不同范数约束的最优化问题进行了详细比较。

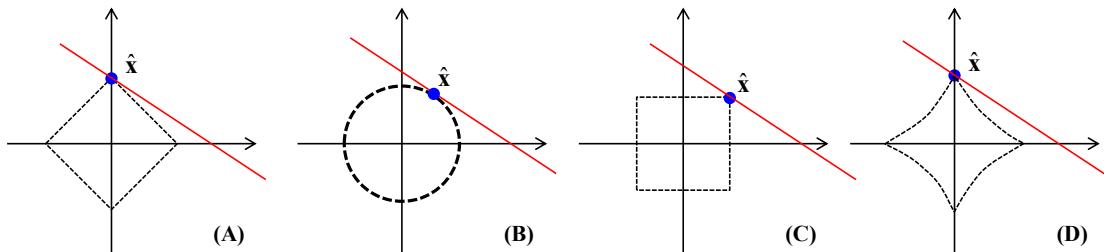


图 2.6 不同范数约束的稀疏表示, 图中直线表示 $\mathbf{y} = \mathbf{A}\mathbf{x}$ 约束条件, 虚线框分别表示 (A) ℓ_1 、(B) ℓ_2 、(C) ℓ_∞ 、(D) $\ell_{\frac{1}{2}}$ 范数平面。逐渐增大各范数平面的半径直到与直线相切为止, 切点即为最小化 ℓ_p 范数问题的最优解。

考虑到噪声扰动, 问题 (2.16) 可扩展为含噪信号的稀疏重构问题。若采样过程为 $\mathbf{y} = \mathbf{A}\alpha + \mathbf{n}$, 其中 \mathbf{n} 表示随机噪声向量, 信号重构可改写为

$$\min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{A}\alpha\|_2^2 < \epsilon \quad (2.17)$$

其中 ϵ 表示允许的误差范围。图 2.7 对于理想情况的优化问题与噪声扰动的最优化问题进行了详细解释。

至此, 稀疏表示以及与之有着密切联系的压缩感知理论框架进行了完整地介绍, 其推导过程可以总结为三个理论层次:

- 如果测量矩阵 (即字典) 满足 $\sigma(\mathbf{A}) \geq \|\alpha\|_0$, 最小化 ℓ_0 范数问题有唯一解;
- 若测量矩阵满足约束等距性, 最小化 ℓ_0 范数与最小化 ℓ_1 范数同解;
- 最小化 ℓ_1 范数是凸优化, 其唯一解就是最小化 ℓ_0 范数最优化的唯一解。

2.2.4 常用的优化工具箱

针对稀疏表示的数值实现, 数学家编写了许多开源工具包, 包括:

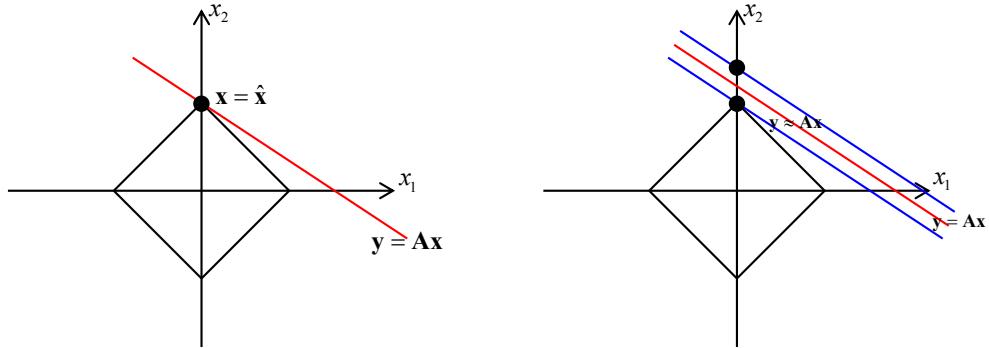


图 2.7 理想情况 ℓ_1 范数优化问题与含噪情况 ℓ_1 范数优化问题的示意图，左图表示理想情况的 ℓ_1 范数优化问题示意图，若以 $\mathbf{x} = (x_1, x_2)$ 表示任意向量，则 ℓ_1 范数为 $\|\mathbf{x}\|_1 = \|x_1\| + \|x_2\|$ ，为了寻找问题的最优解，就是从原点开始逐渐增大菱形的边长，直到约束条件 $\mathbf{y} = \mathbf{A}\mathbf{x}$ 与菱形相切为止，即 $\mathbf{y} = \mathbf{A}\mathbf{x}$ ，切点即为最优解；(b) 考虑测量噪声时 ℓ_1 范数优化问题示意图，图中两条红线所夹的区域即为噪声容限，即约束条件松弛为 $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon$ 。

- Spectral Projected Gradient for L1 minimization⁸ SPGL1 开源工具包是由 E. van den Berg 和 M. P. Friedlander 两位学者于 2008 年 6 月公开发布，主要用于求解大规模复杂数据范数正则化最小二乘问题，包括基追踪、含噪基追踪、LASSO 模型等 [195]。
- A MatlabToolbox for Sparse Statistical Modeling⁹ SpaSM 是丹麦科技大学的 Karl Sjøstrand 等学者编写，主要解决 LAR 模型 (Least Angle Regression)、LASSO 模型、Elastic Net 模型、稀疏主成分分析、稀疏判决分析等问题。
- SparseLab¹⁰ 工具包是斯坦福大学的 David Donoho, Victoria Stodden, Yaakov Tsaig 学者编写并开源发布，主要实现基追踪、匹配追踪、正交匹配追踪、分段匹配追踪等常用算法。
- a SPArse Modeling Software¹¹ SPAMS 是法国 INRIA 研究所的 Julien Mairal 教授开发，其底层使用 C++ 语言进行编写，并配有 Matlab、R、Python 等多种语言的用户接口，由于底层使用了 Intel 的 MKL 技术，是目前运算效率最高的工具包。主要包括字典学习和矩阵分解工具包、稀疏分解工具包、逼近工具包三部分组成，能够实现与稀疏表示相关问题的数值和工程实现。

⁸<http://www.cs.ubc.ca/labs/scl/spgl1/>

⁹<http://www.imm.dtu.dk/projects/spasm/>

¹⁰<http://sparselab.stanford.edu/>

¹¹<http://lear.inrialpes.fr/people/mairal/>

- Sparse Learning with Efficient Projections¹² SLEP 工具包是亚历桑那州立大学的 Jun Liu、Shuiwang Ji 和 Jieping Ye 学者开发，主要解决 ℓ_1 范数约束问题、 ℓ_1/ℓ_q 正则化优化问题、 ℓ_1/ℓ_2 约束问题、多任务稀疏表示以及群组稀疏表示问题等。
- A Matlab Solver for l1-Regularized Least Squares Problems¹³ ℓ_1_ls 工具包是斯坦福大学的 Kwangmoo Koh、Seungjean Kim 和 Stephen Boyd 于 2008 年 5 月公布的开源程度，它使用截断牛顿内点优化算法求解 ℓ_1 正则化最小二乘问题 [196]。
- Matlab Software for Disciplined Convex Programming¹⁴ CVX 是由 Michael C. Grant, Stephen P. Boyd 两位工程师开发并开源，主要用于求解严格凸规划问题，如最小二乘及衍生问题、最小化 ℓ_p 范数等。

2.3 基于稀疏表示的分类学习

稀疏表示自提出后就引起了学者的广泛关注，从早期的小波变换，到后来的多尺度分析（曲线波变换、轮廓波变换、脊波变换等）以及后来的压缩感知，都是寻求一种信号的紧致的、稀疏的表示方法。由于它本身的优良特性，稀疏表示也被广泛应用于计算机视觉。Bruno A. Olshausen 等使用稀疏表示描述哺乳动物初级视觉皮层对线段方向的选择性 [154]；Davi Geiger 等使用一组过完备基对给定的图像进行稀疏表示，通过对表示系数施加一种权值约束 $\ell_p, p \in (0, 1)$ 实现稀疏表示 [155]；Honglak Lee 等提出了一种高效的稀疏表示求解算法，特征符号搜索算法（Feature-Sign Search Algorithm），该算法通过交替求解 ℓ_1 范数正则化最小二乘和 ℓ_2 范数约束最小二乘得出最优系数向量，提高了稀疏编码的计算效率，为后续的改进方法提供了很好地解决思路 [159]；Ke Huang 和 Selin Aviyente 将稀疏表示应用于一维信号分类，提出了一套完整的分类框架，它将信号的重构保真度与线性表示模型稀疏度相结合，应付污染信号的分类，如噪声、数据丢失、奇异样本等，通过实验证明了分类方法的稳健性能；Alfred M. Bruckstein 等回顾了线性系统的稀疏解的理论求解、图像与信号稀疏模型的经验求解以及图像处理的求逆和压缩问题 [197]；John Wright 等提出了基于稀疏信号表示的分类方法，该方法使用 ℓ_2 归一化的训练样本构成过完备冗余字典，将未知样本表示成为字典原子的线性组合，通过约束表示系数的可行集获取问题的唯一稀疏解，并通过实验证明方法

¹²<http://www.public.asu.edu/~jye02/Software/SLEP>

¹³<http://www.stanford.edu/~boyd/>

¹⁴<http://cvxr.com/cvx/>

无论是对噪声污染或是遮挡信号都有优异的分类性能 [156]; Jian Yang 等对 ℓ_1 最优化问题在模式分类中所扮演的角色从实验的角度重新进行解释 [198]。此外，业界的顶级期刊之一，Proceedings of the IEEE 于 2010 年第 98 卷第 6 期分别邀请多位著名学者撰文回顾稀疏表示的研究现状，总结了稀疏表示的典型应用，并对下一步的研究前景进行展望。Michael Elad 等对于图像处理中稀疏（或冗余）表示的作用进行了深刻剖析，通过分析该理论比较成熟的应用领域，如图像超分辨、图像去模糊、图像去噪声等，指出目前所面临的诸多困难，如基于稀疏表示的通用图像压缩方法、多尺度训练字典及其应用、基于稀疏表示的图像合成技术等 [199]。John Wright 等总结了稀疏表示在计算机视觉和模式识别中的应用。文中指出稀疏表示在计算机视觉特别是非传统应用中展现出重大的影响力和应用潜力，其目的并不仅仅是寻找一种对于观测信号的紧致的、高保真的表示方法，更重要的是从中提取语义信息。文章重点给出了关于稀疏信号表示和计算机视觉交互的典型示例，提出了一些公开的问题供后续研究 [200]。Ron Rubinstein 等学者回顾总结了稀疏表示模型中字典的生成方法，将现有的方法系统地梳理为解析字典和学习字典两大类，并分别介绍了每类方法中比较有代表性的算法，就字典学习的未来研究方向进行了展望 [174]。这些著作作为稀疏表示的进一步挖掘与发展产生了深远的影响。

与此同时，稀疏信号表示也被用于应用于遥感图像解译的众多领域。K. Varshney 等学者将过完备冗余字典下稀疏表示方法用于大角度 SAR 成像，他们提出了一种使用结构化字典进行稀疏表示的 SAR 成像方法，该方法将过完备字典分解成多个子字典，且每个子字典的基原子能够按层级进行排列，然后通过梯度下降法来同时优化参数化字典并使用该字典进行稀疏表示，其优点在于描述大角度合成孔径雷达所引发的牵引散射中心现象 [160]；与之类似，S. Samadi 等学者提出了一种基于稀疏表示的合成孔径雷达成像方法，考虑到合成孔径雷达对物体的反射率是复数的特性，文中使用感兴趣的特征将复数据的幅度进行稀疏表示，并推导出了一种能够同时优化潜在反射域的幅度和相位的算法 [161]，Alessandra Budillon 等学者借助压缩感知理论提出了一种源自于多程信号的三维 SAR 聚焦方法 [201]，Christian D. Austin 等学者将稀疏表示用于雷达三维成像问题 [163]；Jayaraman J. Thiagarajan 等学者将稀疏信号表示用于 SAR 图像自动目标识别，其中不同目标在不同姿态角下的成像数据（做 ℓ_2 归一化）用于生成过完备字典，使用字典的基原子对未知样本做稀疏表示，这样就避免了繁琐的姿态角估计和诸多预处理环节，而求解未知样本基于过完备字典的稀疏表示可看作是寻找针对不同类流形的一种局部线性近似（Locally Linear Approximation），从而解释了该方法的合理性 [95]；Nuo Zhou 和 Wei Chen 使用稀疏信号表示进行高分辨率距离向自动

目标识别，其中稀疏信号表示用于分析高分辨率距离向诸分量，稀疏系数看作是描述其特性的特征 [202]；Haichao Zhang 等提出一种基于联合稀疏表示的多视角 SAR 图像目标识别方法，对多个视角下的 SAR 图像分别进行相似约束的稀疏表示，探索同一目标不同视角之间的内在相关性，并借助这种内部关联性实现目标识别 [98]；Chu He 等学者提出了一种基于稀疏表示的 SAR 图像超分辨方法 [203]；Yuntao Qian 等学者通过结构化稀疏表示实现超光谱图像的复原 [177]。

在稀疏表示的众多应用领域中，模式分类 (Pattern Classification) 研究得比较成熟，同时它也是本文的重点，下面首先介绍 John Wright 等学者提出的稀疏表示分类方法 [156] 及其在 SAR 图像目标识别中的应用 [95, 98, 204]。

分类 (或识别) 是根据给定的标记训练样本推测未知样本的类别属性。假设现有 K 个不同的样本类别，其中第 i 类所有 n_i 个训练样本，表示为 $\mathbf{X}_i = [\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \dots, \mathbf{x}_{i,n_i}]$ ，每个样本 $\mathbf{x}_{i,j}$ 都是 m 维的列向量 (如将图像沿列依次排列)。通常认为源自于同一类别的所有样本位于一个线性子空间。对于来自于第 i 类的任意未知样本 \mathbf{y} ，若第 i 类的训练样本足够多，就可以认为未知样本位于由第 i 类训练样本所张成的线性子空间，因而可将其表示为样本的线性组合，

$$\mathbf{y} = \mathbf{x}_{i,1}\alpha_{i,1} + \mathbf{x}_{i,2}\alpha_{i,2} + \cdots + \mathbf{x}_{i,n_i}\alpha_{i,n_i} \quad (2.18)$$

其中 $\alpha_{i,j} \in \mathbb{R}$, $j = 1, 2, \dots, n_i$ 是表示系数。然而，未知样本的类别属性通常是未知的。为了寻找未知样本的类别，将所有 K 类的所有样本沿列拼接起来组成字典， $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_K] = [\mathbf{x}_{1,1}, \mathbf{x}_{1,2}, \dots, \mathbf{x}_{K,n_K}]$ 字典共有 $n = \sum_{i=1}^K n_i$ 个原子。然后，使用字典的原子对未知样本的进行表示，

$$\mathbf{y} = \mathbf{x}_{1,1}\alpha_{1,1} + \mathbf{x}_{1,2}\alpha_{1,2} + \cdots + \mathbf{x}_{K,n_K}\alpha_{K,n_K} = \mathbf{X}\alpha \quad (2.19)$$

其中 $\alpha = [\alpha_{1,1}, \alpha_{1,2}, \dots, \alpha_{K,n_K}] \in \mathbb{R}^n$ 是系数向量。理论上除了第 i 类训练样本对应位置的系数为非零值，其余位置的系数均为零，即 $\alpha = [0, \dots, 0, \alpha_{i,1}, \dots, \alpha_{i,n_i}, 0, \dots, 0]$ ，从而包含未知样本的类别信息。因而通过求解线性系统 (2.19) 可推断未知样本的类别。考虑欠定系统 $m < n$ ，待求的“病态”问题其解并不唯一，最直接的方法是通过最小化 ℓ_2 范数来限制系数向量，如最小二乘 $\min_{\alpha} \|\alpha\|_2 \quad s.t. \quad \mathbf{y} = \mathbf{X}\alpha$ 。尽管 ℓ_2 最小化通过矩阵求逆很容易实现，表示系数并不能提供关于未知样本类别的有用信息。最小化 ℓ_2 范数生成的系数向量几乎所有项均为非零值，而期望的结果是仅有真实类别对应的系数项为非零值，其余系数项为零，即求解的系数向量稀疏，表示系数非零项越少，越容易推测未知样本类别，故最小化 ℓ_2 范数策略无法实现预期目的，转而寻找线性系统的稀疏解系，通过最小化 ℓ_0 范数来约束可行集，

$$\min_{\alpha} \|\alpha\|_0 \quad s.t. \quad \mathbf{y} = \mathbf{X}\alpha \quad (2.20)$$

根据前文介绍可知，若字典 \mathbf{X} 满足某种排列要求，只要系数向量的非零项少于 $\frac{m}{2}$ ，它即是线性系数唯一的最稀疏解。然而，求解最稀疏解是 NP 难问题，甚至很难得到近似解，唯一的方法是穷举其所有的情况逐一进行判断，计算过于复杂。根据压缩感知理论可知，若测量矩阵满足约束等距性最小化 ℓ_0 范数问题与最小化 ℓ_1 范数问题同解，

$$\min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \mathbf{y} = \mathbf{X}\alpha \quad (2.21)$$

线性规则算法即可实现多项式时间求解。此外，其它更为高效的算法较多，如同伦算法能够以复杂度 $O(n + L^3)$ (L 表示稀疏度，即非零数目) 求解问题。考虑到轻微地噪声扰动情况， $\mathbf{y} = \mathbf{X}\alpha + \mathbf{n}$ ，其中 \mathbf{n} 是能量有限的噪声， $\|\mathbf{n}\|_2 < \infty$ ，以类似地方法进行求解

$$\min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{X}\alpha\|_2^2 < \epsilon \quad (2.22)$$

针对噪声扰动的最小化 ℓ_1 范数问题与最小化 ℓ_0 范数问题，D. Donoho 证明 [205]

定理 2.3： 存在常数 ρ 和 ζ ，若最小化 ℓ_0 范数的解 $\hat{\alpha}_0$ 满足 $\|\hat{\alpha}_0\|_0 < \rho m$ 且噪声能量有限 $\|\mathbf{n}\|_2 < \epsilon$ ，则噪声扰动的最小化 ℓ_1 范数问题的解 $\hat{\alpha}_1$ 满足 $\|\hat{\alpha}_1 - \hat{\alpha}_0\|_2 \leq \zeta \epsilon$

通过最优化算法求解得出表示系数，接下来的问题是如何进行判决。理想情况下最优表示系数 $\hat{\alpha}$ 的非零项对应着未知样本的真实类型，然而实际应用中噪声和模型误差通常会导致系数向量中含有多个不同类别的样本的权值系数为非零值。因此，为了得到稳健而精确的估计结果，充分挖掘每个类别子空间的内部结构信息，John Wright 等提出根据每个类别的重构误差确定未知样本类别 [156]。具体方法是：遍历所有类别 $\{1, 2, \dots, K\}$ ，分别取出与之相关的表示系数项，然后计算该类别的重构误差

$$r_i(\mathbf{y}) = \|\mathbf{y} - \mathbf{X}_i \hat{\alpha}_i\|_2$$

，其中 \mathbf{X}_i 和 $\hat{\alpha}_i$ 分别是第 i 类的训练样本和表示系数，最终将未知样本判定为重构误差最小的类别， $\min_{i=1, \dots, K} \{r_i(\mathbf{y})\}$ 。算法流程如图 2.8 所示。

2.4 稀疏表示在 SAR 图像目标识别中的应用

2.4.1 可行性分析

根据论文第一章第 1.2.1 节所述，传统的 SAR 图像目标识别算法，特征提取通常需要估计姿态角以及许多繁琐的预处理环节（如相干斑抑制、图像分割、数学形态学处理等），为了避开这些预处理环节，Jayaraman J. Thiagarajan 等提出基

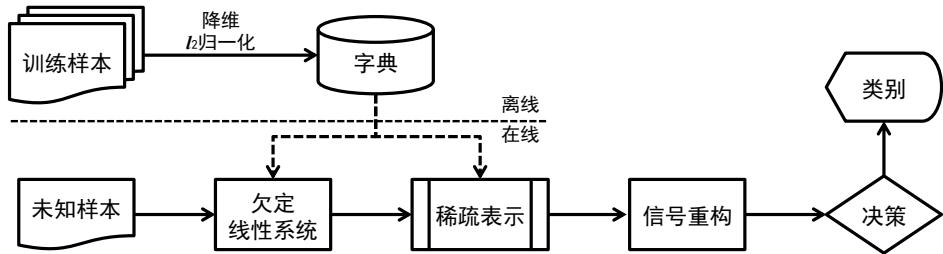


图 2.8 稀疏表示分类：首先使用训练样本构造过完备字典，根据字典对未知样本进行线性表示，通过最小化 ℓ_1 （或者 ℓ_0 ）范数寻找线性系统的最稀疏解，根据重构误差判决目标类型。

于稀疏表示的 SAR 图像目标分类方法 [95]，它实质是 John Wright 等人提出的稀疏表示分类 [156] 在 SAR 图像目标识别中的应用。首先使用不同目标在 $0^\circ \sim 359^\circ$ 范围内姿态角的成像结果组成过完备冗余字典，利用生成的字典对未知样本做线性表示，依据最小化 ℓ_1 （或 ℓ_0 ）范数的优化策略求解最优表示系数，最后根据重构误差来判定未知样本的类别属性。

首先将不同姿态角的 SAR 图像重新排列成列向量并做 ℓ_2 范数归一化，每个类别训练数据组成一个矩阵 $\mathbf{X}_i = [\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \dots, \mathbf{x}_{i,n_i}]$ ，其中 n_i 表示第 i 类的训练样本数，然后将所有 K 类的样本矩阵组合起来构成过完备字典 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_K]$ 。对于任意的未知样本 \mathbf{y} ，将其表示成字典原子的线性组合 $\mathbf{y} = \mathbf{X}\boldsymbol{\alpha}$ ，利用最小化 ℓ_1 的优化策略约束系数可行集

$$\min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{X}\alpha\|_2 < \epsilon \quad (2.23)$$

根据每个类别的重构误差推断未知样本类别。Thiagarajan 算法与 2.3 节 Wright 方法类似，然而，该方法成立是有其自身前提，同一目标在不同姿态解下的成像结果近似张成一个线性子空间。下面将从流形学习地角度解释方法的合理性和可行性。

V. Berisha 等学者的研究指出，对于任意给定的目标，它在不同方位角的 SAR 图像位于一个特定的流形，且该流形的维数远远低于图像的实际维数 [204]。由于目标的 SAR 图像取决于目标本身的姿态角和雷达发射电磁波时的俯仰角，生成的图像位于一个特殊的流形—Image Appearance Manifold (IAM) [206]。因而，SAR 图像的分类就等价于寻找与未知图像距离最近的流形。若以 \mathcal{M}_i 表示第 i 类目标的 SAR 图像所构成的流形，假设该类的所有训练图像都是流形 \mathcal{M}_i 上的样本，由于 SAR 图像存在边缘，该流形不可微。

根据 S. T. Roweis 和 L. K. Saul 提出的局部线性嵌入算法 (Locally Linear Embedding) [207] 可知，任意的非线性流形，如果只考虑它的局部区域，必然存在样本的线性表示。假设在给定姿态角的条件下，雷达俯仰角的变化并不会明显

地改变 SAR 图像的类内聚合性。对于未知样本 \mathbf{y} , 假设它相距第 j 类流形 M_j 较近, 则它可以表示为第 j 类训练样本的局部线性组合,

$$\mathbf{y} = \hat{\mathbf{y}}_j + r_j \quad (2.24)$$

其中 $\hat{\mathbf{y}}_j$ 表示样本 \mathbf{y} 在流形 M_j 上的投影, r_j 是其对应的误差, 投影 $\hat{\mathbf{y}}_j$ 由第 j 类训练样本的局部线性组合来表示

$$\hat{\mathbf{y}}_j = \sum_{l \in \Delta_y^j} \mathbf{x}_{l,j} \alpha_{l,j} \quad (2.25)$$

其中 Δ_y^j 表示参与未知样本 \mathbf{y} 线性表示的第 j 类训练样本的索引集合, $\alpha_{l,j} \in \mathbb{R}$ 是表示相应的表示系数。定义 $\Delta_y = \bigcup_{j=1}^k \Delta_y^j$ 表示训练集中参与未知样本表示的所有样本集合。对于未知样本 \mathbf{y} , 集合 Δ_y 和对应的表示系数 α 是待求的, 参与未知样本线性表示的训练样本数目(稀疏度)远小于整体样本总数。因而, 问题转化成为已知观测值 \mathbf{y} 和字典 \mathbf{X} , 计算 K 稀疏的系数向量 α , 是典型的稀疏表示模型, 通过匹配追踪可近似求解, 也可进行凸松弛求解。得到系数向量后计算残差 $r_j(\mathbf{y}) = \|r_j\|_2$, 根据残差判定样本的类型。

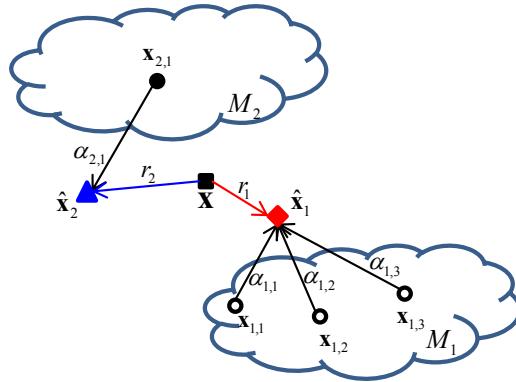


图 2.9 从流形的角度解释稀疏表示分类算法。根据稀疏表示计算未知样本的局部线性投影以及它在流形上的投影残差, 实现分类。图中 $\alpha_{1,1}, \alpha_{1,2}, \alpha_{1,3}$ 是源自于流形 M_1 的样本的表示系数, $\alpha_{2,1}$ 是源自于流形 M_2 的样本的表示系数, r_1 是未知样本在流形 M_1 上的投影误差, r_2 是未知样本在流形 M_2 上的投影误差。由于 $r_1 < r_2$, 可以认为未知样本是源自于流形 M_1 。

为了解释稀疏表示分类方法, 图 2.9 给出了一个简单的例子, 未知样本 \mathbf{y} 分别在 M_1 和 M_2 两个流形进行投影, 其中 M_1 流形上有三个训练样本参与了 \mathbf{y} 的表示 (用空心圆圈标记), M_2 流形上仅有 1 个样本参与了 \mathbf{y} 的表示 (用实心圆标记), 未知样本在流形 M_1 上的投影是 $\hat{\mathbf{x}}_1$ (用菱形块标记), 对应的重构误差 (也就是未知样本距离流形 M_1 的距离) 为 r_1 ; 与之类似, 未知样本在流形 M_2 上的投影是 $\hat{\mathbf{x}}_2$ (用三角形标记), 对应的重构误差 (也就是未知样本距离流形 M_1 的距

离) 为 r_2 。显然, \mathcal{M}_1 流形的表示误差 r_1 小于 \mathcal{M}_2 流形的表示误差 r_2 , 因此就认为未知样本属于 \mathcal{M}_1 类。

基于稀疏表示的分类算法只涉及流形的线性投影, 根据投影误差进行决策, 不需要对姿态角进行估计, 也不需要执行预处理。通过求解稀疏表示的局部线性近似方法挑选所需的训练样本, 判定目标类别。

2.4.2 改进方案

2.4.2.1 表示方法改进—频域稀疏表示

为了构造欠定系统, 稀疏表示的字典原子维度小于字典中原子个数。以 SAR 图像为例, 图像大小通常为 128×128 像素, 重新排列成向量, 维度为 \mathbb{R}^{16384} , 很难直接构建欠定线性系统, 因而需要对原始图像做降维处理, John Wright 等学者分别对比了一些常见的降维方法 (如降采样、主成分分析) 和一些不常见的降维方法 (如随机投影) [156]。Jayaraman J. Thiagarajan 等采用随机投影策略压缩数据维度, 构造线性回归模型 [95]。经典方法前提是理想数据集, 即测试样本与训练样本在图像域执行严格配准。尽管这些方法取得理想的实验结果。其前提条件无法直接应用于实际战场。经检测、鉴别处理之后, 传输到分类阶段的未知 ROI 虽位于切片的中心, 但不可能与训练图像进行严格配准, 这就需要考虑图像的空间偏移问题; 此外, 实际应用遇到的目标结构变化、传感器测量误差、噪声污染等因素也值得关注。

为了解决图像的空间偏移问题, 国外学者曾提出一种基于相关滤波的目标识别方法 (见论文第 1.2.1 节)。该方法将图像由空域转换至频域, 使用每类目标在某个姿态角区间内的傅立叶变换系数生成相关滤波器组, 根据未知样本的傅立叶变换系数与滤波器的相关卷积结果进行决策。相关滤波的目标识别方法能够解决图像的空间偏移和变形问题, 实现平移不变和失真不变。受相关滤波器的启发, 论文摒弃传统的图像空域数据, 使用图像的傅立叶变换系数构造多元回归分析, 通过稀疏约束求解回归系数。首先利用训练样本的傅立叶变换系数 (列向量) 作为原子构成过完备字典, 使用频域字典的原子对未知样本的傅立叶变换进行线性表示, 通过最小化 ℓ_1 范数获取最优表示系数, 计算重构误差, 判定未知样本的类别属性。

1. 算法整体框架

对于任意给定的二维有限长信号 $f(x, y) \in \mathbb{R}^{M \times N}$, 其中 M 和 N 表示信号的行数和列数, 它的二维离散傅立叶变换可以表示为

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp\left\{-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right\} \quad (2.26)$$

其中 $u = 0, \dots, M - 1, v = 0, \dots, N - 1$ 是频域索引。实信号 $f(x, y)$ 的傅立叶变换结果通常是复信号，可表示为笛卡尔坐标的实部与虚部

$$F(u, v) = R(u, v) + jI(u, v) \quad (2.27)$$

同塑钢，也可转换为极坐标的幅度谱和相位谱，

$$F(u, v) = |F(u, v)| \exp \{ -j\phi(u, v)\} \quad (2.28)$$

傅立叶变换频谱的能量集中在极少数波段，大部分变换系数都是冗余的，不含有鉴别信息。为了构造欠定表示系统，过完备字典的原子数目须远大于每个原子的特征维数，这就要求每个字典原子只包含能量集中的极少数波段，直接舍弃大多数冗余频谱，即设计频域特征描述子，筛选有鉴别信息的频谱构成特征向量，使用频域特征向量构建欠定线性系统。若以映射 $\mathcal{T} : \mathbb{R}^m \mapsto \mathcal{F}$ (\mathcal{F} 表示频域) 表示频域特征描述，使用训练样本的频域特征对未知样本 \mathbf{y} 的频域特征进行线性表示，

$$\mathcal{T}\{\mathbf{y}\} = \mathcal{T}\{\mathbf{X}\}\alpha = \mathcal{T}\mathbf{x}_{1,1}\alpha_{1,1} + \mathcal{T}\mathbf{x}_{1,2}\alpha_{1,2} + \dots + \mathcal{T}\mathbf{x}_{K,n_K}\alpha_{K,n_K} \quad (2.29)$$

其中 $\mathcal{T}\{\mathbf{X}\} = [\mathcal{T}\{\mathbf{X}_1\}, \mathcal{T}\{\mathbf{X}_2\}, \dots, \mathcal{T}\{\mathbf{X}_K\}]$ 是训练样本的频域特征构成的冗余字典。考虑欠定问题 $\text{card}(\mathcal{F}) < n$ ，通常期望系数向量 α 尽可能稀疏，从而提供关于未知样本类别归属的有用信息，该问题可通过最小化 ℓ_1 范数来实现

$$\min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \|\mathcal{T}\{\mathbf{y}\} - \mathcal{T}\{\mathbf{X}\}\alpha\|_2 < \epsilon \quad (2.30)$$

根据表示系数 $\hat{\alpha}$ 计算重构误差，推断未知样本类别， $\min_{i=1, \dots, K} \|\mathcal{T}\{\mathbf{y}\} - \mathcal{T}\{\mathbf{X}_i\}\hat{\alpha}_i\|_2$ 。算法的整个流程如图 2.10 所示。

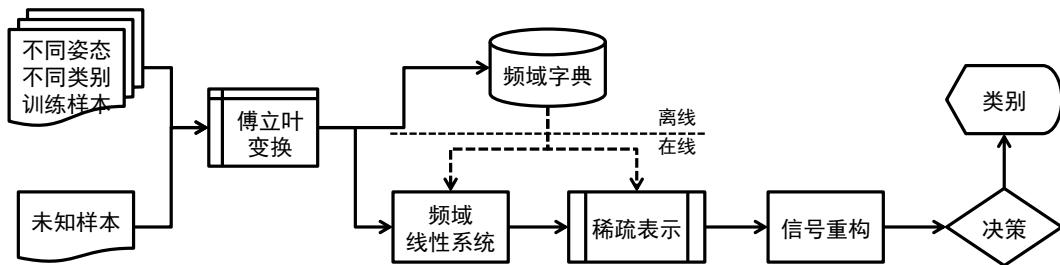


图 2.10 频域稀疏表示分类流程图。首先使用训练样本的频域特征构造过完备字典，根据字典原子对未知样本频域特征进行线性表示，求解最优表示系数，根据重构误差进行判决。

2. 频域特征描述

信号的频域能量主要集中于很少数频率波段 [208]，因此那些能量集中的频率波段比其余的冗余频率波段具有更强的鉴别能力，它们包含信号的主要鉴别信息。

受这个结论启发，能否使用能量集中的极少数频率波段构造频域特征描述子，摒弃冗余频谱。

图 2.11 给出了一幅 SAR 图像及其傅立叶变换频谱，包括实部、虚部、幅度谱和相位谱。可以看出，相位谱的能量分布则是杂乱无章，幅度谱能量集中于四个对角线区域的频谱波段。因此，仅取出对角线区域频谱，重新排列生成向量即为本文的频域特征，这样即降低特征向量的维数，又能提高鉴别能力。



图 2.11 SAR 图像及其傅立叶频谱。图中依次给出 SAR 图像的幅度谱、实部、虚部和幅度谱。

为了观察傅立叶频谱随空间平移的变化情况，图 2.12 给出了一组实例，将原始 SAR 切片图像在图像域手动平移 10 个像素，然后观察平移前后的傅立叶频谱变化情况，可以看出，平移前后的幅度谱基本没有变化，能量都集中在四个角落的频率波段，相位谱是杂乱无章，不存在特定规律。为了更仔细地观察对角线区域幅度谱与相位谱的差别，取出四个对角区域的频谱数据，重新组成一个子矩阵，如图 2.13 和 2.14 所示。从图中可知，两组幅度谱系数十分近似，相位谱系数没有任何相似性，因此，可以认为傅立叶变换幅度谱具备平移不变特性，而相位谱不具备。

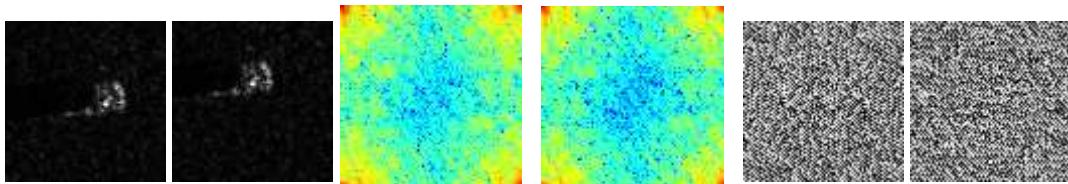


图 2.12 SAR 图像空间平移前后的傅立叶频谱对比，三组六幅图片分别是 SAR 图像与空间平移 10 像素之后图像的傅立叶幅度谱、相位谱对比，其中幅度谱是进行对数变换压缩动态范围之后的结果。

接下来的实验用于验证剔除大部分冗余频谱后图像重要信息是否能得以保留，如图 2.15 所示，原始图像为 80×80 像素，其二维傅立叶变换结果为 80×80 ，共 6400 组傅立叶频谱，实验中只保留对角线区域 20×20 共 400 组傅立叶频谱，即原始频谱的 $\frac{1}{16}$ ，其余频谱全部置零，然后对频谱筛选结果再做逆傅立叶变换，从复原结果可以看出，目标轮廓清楚，细节比较清晰，重要信息得以保留。虽然仅使用极少部分傅立叶频谱而摒弃绝大部分频谱，但原始图像的重要信息能够得以保留，实验结果很好地印证了前文的相关论述。

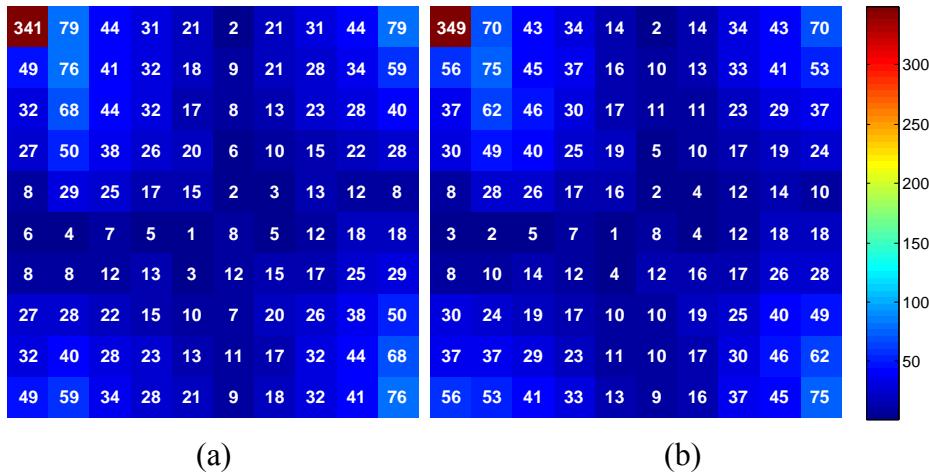


图 2.13 SAR 图像空间平移前后对角线 10×10 像素区域幅度谱对比, (a) 原始 SAR 图像对角线区域幅度谱系数, (b) 平移 10 像素后对角线区域幅度谱系数。

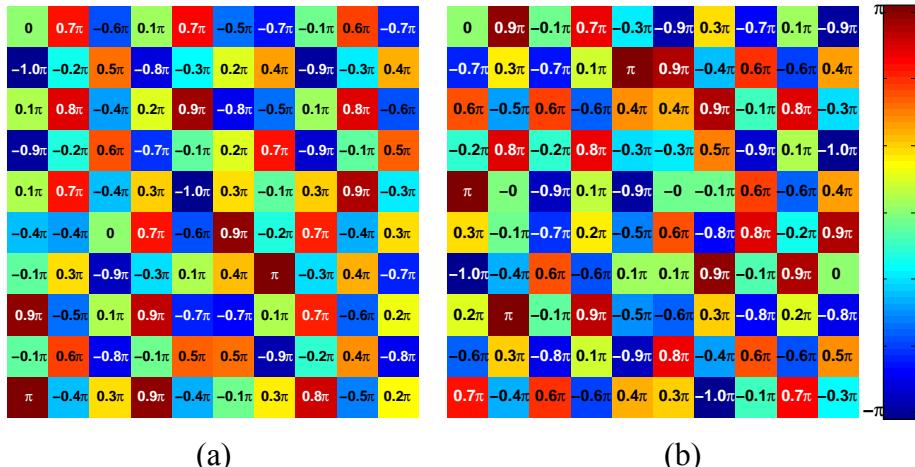


图 2.14 SAR 图像空间平移前后对角线 10×10 像素区域相位对比, (a) 原始 SAR 图像对角线区域相位谱系数, (b) 平移 10 像素后对角线区域相位谱系数。

2.4.2.2 决策规则改进—基于 Dempster-Shafer 证据融合的“软”决策

稀疏表示分类根据每个类别的重构误差估计待测样本的类别属性, 实质就是寻找距离未知样本最近的流形, 根据投影残差进行决策。换个角度看, 这种决策

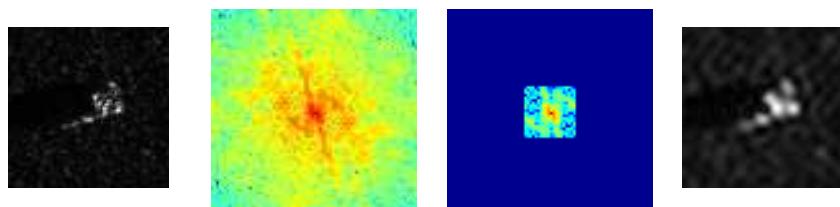


图 2.15 频谱筛选前后的 SAR 图像对比, (a) 原始 SAR 图像, (b) 原始 SAR 图像幅度谱 (利用 fftshift 将 DC 分量搬至频谱中心), (c) 保留中心区域频谱, 其余置零, (d) 筛选后的频谱做逆傅立叶变换, 生成的复原图像 (幅度谱做对数变换压缩动态范围)。

方法是一种贝叶斯估计，若将似然概率定义为样本的重构误差的函数，经典的决策方法就是最大似然估计（或最大后验概率）。由于贝叶斯理论的限制，如不能对未知测量状态的不确定性进行建模（详见 [209, 210] 等文献描述），经典的“硬”决策通常会生成错误判决。为了减少贝叶斯“硬”决策的误判，提高决策可靠性，本文提出一种基于数据融合的“软”决策。

为了充分利用信号蕴含的丰富信息，实现可靠而稳健的联合决策，常用的方法就是数据融合。现有的数据融合方法包括模糊理论、贝叶斯估计、纽曼—皮尔逊准则、Dempster-Shafer 证据理论等，其中 DS 证据理论因其本身的诸多优点而受到广泛关注 [211]。Thierry Denaux 使用 DS 证据理论对传统的最近邻分类进行改进 [212]。他将待测样本的每个近邻样本都看作是能够提供关于未知样本类别属性的一组证据，并定义概率置信函数与两个样本之间的欧氏距离满足某种特定函数关系，然后将所有 k 个近邻样本所提供的证据条目进行合并得到全局证据，并根据全局证据得出最终的判决。Sylvie Le Hegarat-Mascle 等将 DS 证据理论用于多源遥感图像的无监督分类问题 [209]。她们将不同信源图像分类结果进行对比，确定辨识框架和焦元，定义证据质量函数，通过 Shafer 融合规则摒弃无效类别，实现遥感图像的无监督分类。Florence Tupin 等学者将 DS 证据理论用于 SAR 图像自动解译的多个检测器融合检测问题 [213]。他们将 SAR 图像的自动解译工作划分成三步，第一步使用多个低层次的检测器从场景 SAR 图像中提取丰富的有用信息；在第二步，他们将各种检测器的检测结果使用 DS 证据理论进行融合，最后使用上下文知识给出最终的解译结果。Zhun-ga Liu 等学者将 DS 证据理论用于遥感图像的变化检测问题 [214]，Jiaojiao Tian 等学者使用 DS 证据理论将两种不同的检测指标进行融合，实现建筑物的变化检测 [215]。

受这些文献启发，本文提出一种基于 DS 证据推理的稀疏表示“软”决策。传统的决策规则只考虑未知样本属于单个类别的概率可能，根据最大似然估计（或最大后验概率）进行决策，是基于贝叶斯估计的“硬”决策。本文对其做进一步扩展，考虑未知样本属于不确定类别（如多个类别组合）的可能性，并对这些命题的不精确性进行量化。对于每一组参与未知样本线性表示的样本，根据它对未知样本的重构误差来定义一个基本概率赋值条目（即局部证据），然后，使用合并规则将局部证据合并为全局证据。合并过程分为两步，第一步是类内合并，用于合并由那些参与未知样本线性表示的源自于同一个类别的训练样本所提供的证据；第二步是类间合并，用于对每个类别内部的局部证据条目进行合并，从而得到全局证据。最后根据某种特定的规则，如最大置信度函数，最大似然度函数等，判决未知样本的类别。

1. Dempster-Shafer 证据理论

分类是将未知样本分配给 K 个互斥的目标类别之一，寻找最可能类别。论文用 $\Theta = [\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_K]$ 表示 K 个类别，用 $2^\Theta = \{\emptyset, \mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_K, \mathcal{L}_1 \cup \mathcal{L}_2, \dots, \mathcal{L}_1 \cup \mathcal{L}_2 \cup \dots \cup \mathcal{L}_K\}$ 表示它的幂集，即所有可能的类别组合。基本概率赋值 (Basic Probability Assignment) 是定义在从幂集 2^Θ 到区间 $(0, 1)$ 上的一个映射， $m(\cdot) : 2^\Theta \mapsto (0, 1)$ ，它满足 $A \in 2^\Theta$ ，

$$m : \begin{cases} m(\emptyset) = 0 \\ m(A) \geq 0 \\ \sum_{A \subseteq 2^\Theta} m(A) = 1 \end{cases} \quad (2.31)$$

其中 \emptyset 是空集。基本概率赋值也称为证据质量函数 (Mass function)。如果复合命题 $A \cup B$ 的基本概率赋值 $m(A \cup B)$ 非零， $m(A \cup B) > 0$ ，认为假设命题 A 或是 B 成立存在一定的可能性，不需要讨论它具体属于某个类别。类似的，如果 $\mathcal{L}_1 \cup \mathcal{L}_2 \cup \dots \cup \mathcal{L}_K$ 非零， $m(\Theta) > 0$ ，则认为未知样本属于所有类别的可能性。基本概率赋值函数提供某个命题为真的可信度，所有基本概率赋值非零的假设被称之为焦元。所有可能的假设命题组成的集合称为辨识框架。Dempster 和 Shafer 在基本概率赋值的基础上设计似然函数 (Plausibility) 和置信度函数 (Belief or Support) 表示不确定性和不精确性，它们的定义为

$$\begin{aligned} Bel(A) &= \sum_{B \subseteq A} m(B) \\ Pls(A) &= \sum_{A \cap B \neq \emptyset} m(B) \end{aligned} \quad (2.32)$$

从定义 (2.32) 可以看出，置信度函数和似然函数是命题的最低和最高概率函数，

$$\begin{aligned} Bel(A) &\leq Pls(A) \\ Pls(A) &= 1 - Bel(\bar{A}) \end{aligned} \quad (2.33)$$

其中 \bar{A} 表示 A 的补集命题，且有结论 $A \cup \bar{A} = \Theta$ 和 $A \cup \bar{A} = \emptyset$ 。类似地也容易得出不等式关系 $Bel(A) \leq m(A) \leq Pls(A)$ 。显然，不等式关系取等号即贝叶斯估计，因而，DS 证据理论也可以看作是贝叶斯理论的广义扩展。图 2.16 解释了 DS 证据理论置信度函数与似然函数的关系，置信函数也被视为命题的悲观估计 (下限)，似然函数也被视为命题的乐观估计 (上限)。

DS 证据理论体系中，源自于两个独立信息源且建立在同一辨识框架上的两组置信函数 Bel_1 和 Bel_2 ，可以被合并成一组置信函数。Dempster 组合规则就是针对这一需求而提出的。首先要求这两组置信函数是可组合的，即两种假设之间无交集 $\sum_{A \cap B = \emptyset} m_1(A)m_2(B) < 1$ 。对于两组可组合置信函数，它们的证据合并规则为

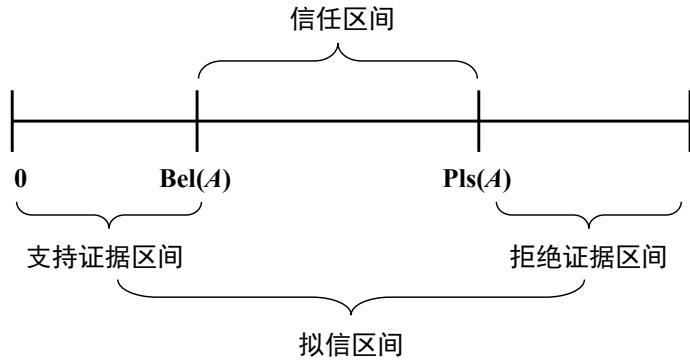


图 2.16 DS 证据理论的置信函数与似然函数。

$$m(\theta) = \frac{\sum_{A \cap B=\theta} m_1(A)m_2(B)}{1 - \sum_{A \cap B=\emptyset} m_1(A)m_2(B)}, \theta \neq \emptyset \quad (2.34)$$

且它们的组合结果满足基本概率赋值定义 (2.31)，合并后空集的证据质量仍为 0， $m(\emptyset) = 0$ ，这种组合规则也被称作直和 $Bel = Bel_1 \oplus Bel_2$ ，直和 (\oplus) 原指两个线性子空间的加法运算 “+”，详见附录 A。

根据直和规则的定义 (2.34) 可知它满足交换律和结合律，因而，很容易将其扩展到多个不同信息源证据质量的合并 $Bel = Bel_1 \oplus Bel_2 \oplus \dots \oplus Bel_n$ 。在完成证据合并之后，需要某个准则（类似于贝叶斯统计中的最大似然估计、最大后验概率等）来判断哪种假设更符合实际，目前比较常用的决策准则有：最大似然度函数、最大置信度函数和最大无重叠置信区间置信度函数等 [209]。

从 DS 证据理论的框架结构来看，实际应用中最为关键的是如何定义基本概率赋值函数，通过该函数引入不确定性命题以及它的不精确性度量，构建的模型是否合理直接决定着后续的合并及推理能否成功。然而，关于基本概率赋值函数如何定义这个难题，目前仍没有一种普适的、通用方法可供借鉴，现有的策略都是面向各自任务的函数构造方法，包括图像像素差异 [212]、模糊理论 [216] 以及借助第三方学习方法 [217] 等。

2. 基于证据融合的决策新规则

为了改进传统的稀疏表示分类算法决策规则的不足，本文提出一种基于 DS 证据推理的决策新方法 [218]。与传统的判决准则不同，本文将参与未知样本线性表示的所有训练样本看作是提供未知样本类别归属的信息源，每组训练样本都会提供一组关于未知样本类别属性的证据条目，每个条目证据质量是通过它对未知样本的重构误差进行量化，即根据它们对未知样本的重构误差（残差）定义基本概率赋值函数，实现数值量化，然后利用 Dempster 直和规则对所有训练样本生成的证据条目进行合并，得到全局证据，根据全局证据判决未知样本的类别属性。

将 K 个类别的训练样本组合成的字典 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_K] \in \mathbb{R}^{m \times n}$, 将字典原子的索引序号按照 $1 \sim n$ 的顺序重新排列, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, 每个原子对应的类别标记为 C_1, C_2, \dots, C_n , 其中 $C_i \in \{\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_K\}$ 。对于未知样本 \mathbf{y} , 使用字典原子的线性组合对其进行表示, 构成多元线性回归模型

$$\mathbf{y} = \mathbf{x}_1\alpha_1 + \mathbf{x}_2\alpha_2 + \cdots + \mathbf{x}_n\alpha_n = \mathbf{X}\alpha$$

其中 $\alpha \in \mathbb{R}^n$ 是回归系数, 回归模型的构建方法与前文相同, 通常期望回归系数尽可能稀疏 $\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_1$, 即参与未知样本线性表示的训练样本尽可能少, 从而包含期望的关于未知样本类别属性的足够信息。借助经典的最优化算法即可得到最稀疏表示系数, 然而如何根据表示系数进行证据质量描述是本文的关键?

首先根据最优回归系数 $\hat{\alpha}$ 中非零元素的位置来寻找参与未知信号线性表示的所有训练样本, 将这些样本组成集合 Φ 。对于集合中的任意元素 $\mathbf{x}_j \in \Phi$, 找出它对应的类别标记 $C_j = q$, 根据它的标记可以推测出关于未知样本 \mathbf{y} 属于第 q 类目标的证据质量 $m_{\mathbf{x}_j}$, 即参与未知样本线性表示的每个系数不为零的训练样本都会提供一条关于未知样本归属的证据。然而, 由样本 \mathbf{x}_j 所提供的信息也不可能百分之百确定。因而在 DS 证据理论框架中, 对于样本 \mathbf{y} 属于第 q 类的假设仅仅是部分可信。由于在这个假设过程中不涉及其它假设, 剩余的置信度就分配给不确定状态 Θ 。因此, 在这个推理过程中仅有假设 $\{C_j\}$ 和不确定类 Θ 非零, 其余假设均为零。上述的证据框架可使用如下公式进行表示

$$\begin{aligned} m^j(\{C_j\}) &= \eta \\ m^j(\Theta) &= 1 - \eta \end{aligned} \tag{2.35}$$

其中 $0 \leq \eta \leq 1$, 剩余命题的证据质量均为零 $\forall A \in 2^\Theta, A \notin \{C_j\}, \Theta\}, m^j(A) = 0$ 。根据前文的分析可知, 若重构误差 $e^j = \|\mathbf{y} - \mathbf{x}_j\alpha_j\|$ 越小, 则表示未知样本 \mathbf{y} 属于第 C_j 类的可能性越大, 样本 \mathbf{x}_j 所提供的信息越丰富, 因此证据质量越大即 η 的值越大; 相反地, 若重构误差 e^j 越大, 则样本 \mathbf{x}_j 提供的信息越少, 因此对应的 η 值就越小。根据这个结论可知证据质量 η 应该是重构误差 e^j 的减函数, 据此本文构造以重构误差 e^j 为自变量的函数

$$\eta = \eta_0 \cdot g(e^j),$$

从而量化地描述重构误差与基本概率赋值之间的关系, 式中 $0 \leq \eta_0 \leq 1$ 是权值系数, $g(\cdot)$ 是辅助函数, 它满足

$$g(0) = 1, \quad \lim_{x \rightarrow \infty} g(x) = 0 \tag{2.36}$$

式 (2.36) 中引入权值系数 η_0 是因为：即使样本 \mathbf{x}_j 对未知样本的重构误差为 0 (理想情况)，也不会完全相信由 \mathbf{x}_j 一个样本所提供的信息。因此，参数 η_0 的引入是完全有必要的。

辅助函数 $g(\cdot)$ 如何设计？ $g(\cdot)$ 应该是自变量的减函数，且必须满足约束条件 (2.35)，但满足该条件的函数有成千上万之多，且不太可能给出某个函数优于其它的函数，因而只能凭经验来设计。受文献 [212] 的启发，本文使用负指数函数来描述重构误差与基本概率赋值之间的关系，

$$g(x) = \exp\{-\lambda\|x\|_1\} \quad (2.37)$$

其中 $\lambda > 0$ 是常数。借助于函数 (2.37) 可对参与未知样本线性表示的字典原子 $\mathbf{x}_j \in \Phi$ 提供的信息进行量化，其基本概率赋值不仅取决于它对未知样本的重构误差，同时也与它的类别有关。

根据设计的基本概率赋值函数可对参与未知样本线性表示的每个字典原子提供的证据条目进行量化，接下来所要解决的问题就是如何将量化之后的碎片化证据条目进行合并，得出全局证据，从而判决未知样本类型。结合实际，论文将碎片化证据的合并过程划分成两个连续步骤，类内合并和类间合并，下文将分别进行介绍。

类内合并 *Intra-Class Combination*

对于源自于同一个类别 \mathcal{L}_q 的任意两组样本 $\mathbf{x}_i, \mathbf{x}_j \in \Phi$ ，根据 (2.35) 式计算它们的基本概率赋值， $m^i(\{\mathcal{L}_q\})$ 和 $m^j(\{\mathcal{L}_q\})$ 。然后利用 Dempster 合并规则（直和）将类内样本提供的证据进行合并，得到简单假设和复合假设的基本概率赋值为

$$\begin{aligned} m^{(i,j)}(\{\mathcal{L}_q\}) &= 1 - (1 - \eta_0 \cdot g(e^i)) \cdot (1 - \eta_0 \cdot g(e^j)) \\ m^{(i,j)}(\Theta) &= (1 - \eta_0 \cdot g(e^i)) \cdot (1 - \eta_0 \cdot g(e^j)) \end{aligned} \quad (2.38)$$

如果用 Φ_q 表示参与未知样本线性表示的源自于第 q 类原子的集合，将集合 Φ_q 中所有元素提供的证据进行合并，结果表示为 $m_q = \bigoplus_{\mathbf{x}_j \in \Phi_q} m^j$ ，其中类内简单假设和复合假设的基本概率赋值为

$$\begin{aligned} m_q(\{\mathcal{L}_q\}) &= 1 - \prod_{\mathbf{x}_j \in \Phi_q} (1 - \eta_0 \cdot g(e^j)) \\ m_q(\Theta) &= \prod_{\mathbf{x}_j \in \Phi_q} (1 - \eta_0 \cdot g(e^j)) \end{aligned} \quad (2.39)$$

类间合并 *Inter-Class Combination*

类内合并目的是将参与未知样本线性表示的源自同一个类别的原子提供的碎片化证据进行聚合，在此基础上需要将面向每个类别的证据进行聚合，生成全局

证据。使用 Dempster 规则将每个类别提供的证据 $m_q(\cdot), q = 1, \dots, K$ 组合起来，结果为 $m = \bigoplus_{q=1}^K m_q$ ，其中类间简单假设和复合假设的基本概率赋值为

$$\begin{aligned} m(\{\mathcal{L}_q\}) &= \frac{m_q(\{\mathcal{L}_q\}) \prod_{r \neq q} m_r(\Theta)}{1 - \mathcal{E}_c} \\ m(\Theta) &= \frac{\prod_{q=1}^K m_q(\Theta)}{1 - \mathcal{E}_c} \end{aligned} \quad (2.40)$$

$\mathcal{E}_c = \sum_{j=1}^K \prod_{q \neq j} m_q(\{\mathcal{L}_q\})$ 是冲突证据，作用类似于归一化因子。

论文设计的焦元只包含全局不确定命题，具体量化过程中又分为每个原子的不确定命题和每个类别的不确定命题，即样本复合假设和类间复合假设，根据合并生成的全局基本概率赋值可计算似然函数和置信度函数，

$$\begin{aligned} Bel(\{\mathcal{L}_q\}) &= m(\{\mathcal{L}_q\}) \\ Pls(\{\mathcal{L}_q\}) &= m(\{\mathcal{L}_q\}) + m(\Theta). \end{aligned} \quad (2.41)$$

显然此前提到的几种决策方法（如最大似然度函数、最大置信度函数等 [209]）给出的结论是一致的。整个证据合并框架可概括为表 2.1。

表 2.1 基于 DS 证据推理的“软”决策证据合并框架 ($\Theta = \mathcal{L}_1 \cup \mathcal{L}_2 \cup \dots \cup \mathcal{L}_K$)

	\mathcal{L}_1	\mathcal{L}_2	\dots	\mathcal{L}_K	Θ
类内合并	$m^{j_1}(\mathcal{L}_1) \dots$ $m^{j_1}(\Theta) \dots$	$m^{j_2}(\mathcal{L}_2) \dots$ $m^{j_2}(\Theta) \dots$	\dots	$m^{j_p}(\mathcal{L}_K) \dots$ $m^{j_p}(\Theta) \dots$	
	$m_1(\mathcal{L}_1)$	$m_2(\mathcal{L}_2)$	\dots	$m_K(\mathcal{L}_K)$	$m_1(\Theta)$ $m_2(\Theta)$ \dots $m_K(\Theta)$
全局证据	$m(\mathcal{L}_1)$	$m(\mathcal{L}_2)$	\dots	$m(\mathcal{L}_K)$	$m(\Theta)$

2.4.3 实验验证

2.4.3.1 实测数据集

目前，用于 SAR 图像解译算法验证的标准数据十分难得。由于数据采集的成本较高，很少有免费公开的数据集。因而，论文使用目前使用最为广泛的 MSTAR 数据，即美国预先研究计划署和空军实验室联合资助的面向 SAR 自动目标识别的公测数据集，于 1996 年和 1997 年利用 X 波段、HH 极化方式、 $1ft \times 1ft$ 分辨率聚束式合成孔径雷达采集，包括机动目标切片图像（如图 2.17）和场景图像（如图 2.18）两部分，其中机动目标成像数据分别是在 15° 、 17° 、 30° 、 45° 等雷达俯

仰角采集得到，官方推荐使用俯仰角为 17° 的成像数据进行训练，使用俯仰角为 15° 的成像数据进行测试，详见表 2.2，2S1、BRDM2 和 ZSU23/4 还采集了多个雷达俯仰角的成像数据，见表 2.3。切片图像大小有 128×128 像素、 129×128 像素、 158×158 像素等，场景图像大小为 1478×1784 像素。为了方便实验设置，论文将切片图像的大小统一裁剪成 80×80 像素或是 64×64 （只取中心部分）。



图 2.17 姿态角为 40° 的 MSTAR 机动目标切片图像。

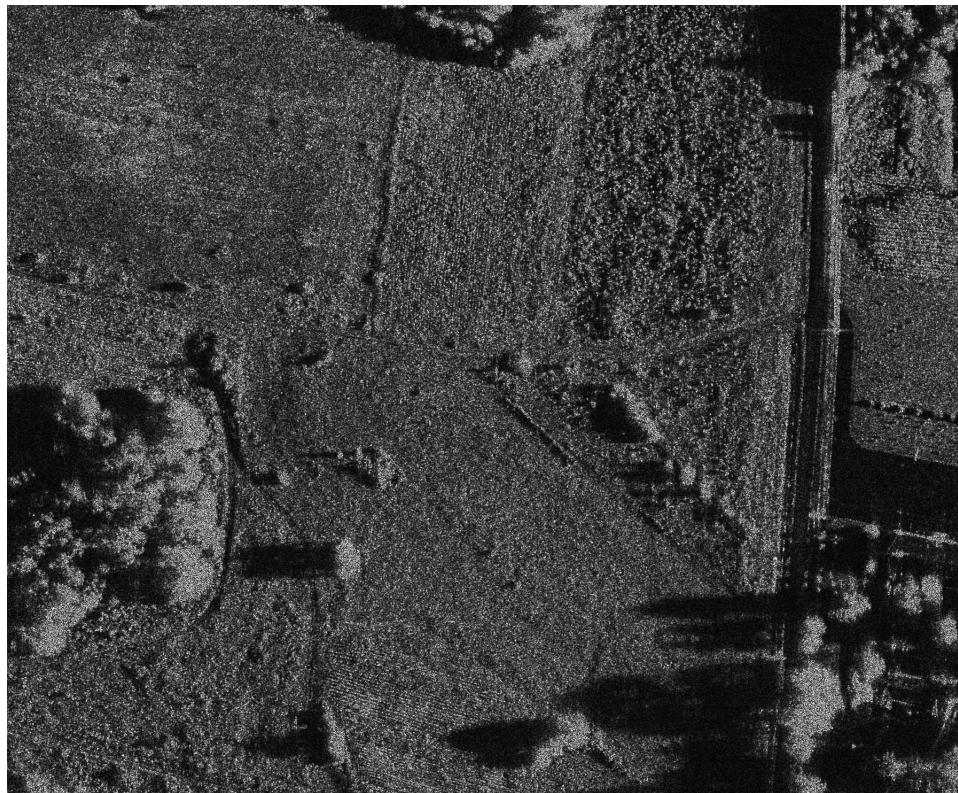


图 2.18 MSTAR 场景 SAR 图像

首先利用一组实测数据观察 SAR 图像目标散射现象的姿态敏感性、配置敏感性和成像参数敏感性，更详细的论述参见论文第一章第 1.2.2 节的相关阐述。

- **姿态敏感性和俯仰角敏感性：**图 2.19 (a) 是姿态角为 13° 、 14° 、 15° 的 BMP2 装甲输送车 (SN_9563) 在 15° 雷达俯仰角的成像，即使是 1° 的姿态角差异

表 2.2 十类苏制军事目标在 17° 和 15° 俯仰角下的成像样本总数，每组样本是该类目标在某个姿态角下的成像， $TG_1 \rightarrow BMP2$ 是步兵战车， $TG_2 \rightarrow BTR70$ 是装甲输送车， $TG_3 \rightarrow T72$ 是主战坦克， $TG_4 \rightarrow BTR60$ 是轮式装甲车， $TG_5 \rightarrow 2S1$ 是自行榴弹炮， $TG_6 \rightarrow BRDM_2$ 是两栖装甲侦察车， $TG_7 \rightarrow D7$ 是民用推土机， $TG_8 \rightarrow T62$ 是主战坦克， $TG_9 \rightarrow ZIL131$ 是军用卡车， $TG_{10} \rightarrow ZSU23/4$ 是自行高炮，BMP2 和 T72 括号内是不同配置序列号。

<i>Depr.</i>	TG_1	TG_2	TG_3	TG_4	TG_5	TG_6	TG_7	TG_8	TG_9	TG_{10}
17°	233 (SN_9563)		232 (SN_132)							
	232 (SN_9566)	233	231 (SN_812)	256	299	298	299	299	299	299
	233 (SN_c21)		228 (SN_s7)							
15°	195 (SN_9563)		196 (SN_132)							
	196 (SN_9566)	196	195 (SN_812)	195	274	274	274	273	274	274
	196 (SN_c21)		191 (SN_s7)							

表 2.3 2S1、BRDM_2 和 ZSU23/4 在 $15^\circ, 17^\circ, 30^\circ, 45^\circ$ 四种雷达俯仰角的成像样本数，其中括号中的数目是连接体等变体情况下的成像样本数。

<i>Depr.</i>	2S1	BRDM_2	ZSU23/4	Total
Training (17°)	299	298	299	896
(15°)	274	274	274	822
<i>Testing</i> (30°)	288	287 (133)	288 (118)	1114
(45°)	303	303 (120)	303 (119)	1148

其成像结果也存在着明显变化，这验证了论文第 1.2.2 节关于扩展工作条件的姿态敏感性；图 2.19 (b) 是姿态角为 $13^\circ, 14^\circ, 15^\circ$ 的 BMP2 装甲输送车 (SN_9563) 在 17° 雷达俯仰角的成像，图 2.19 (a) 和 (b) 两组图像仅仅存在 2° 的俯仰角变化，但它们的成像结果却呈现出显著变化，该现象也验证了论文第一章 1.2.2 节关于扩展工作条件的成像参数敏感性。

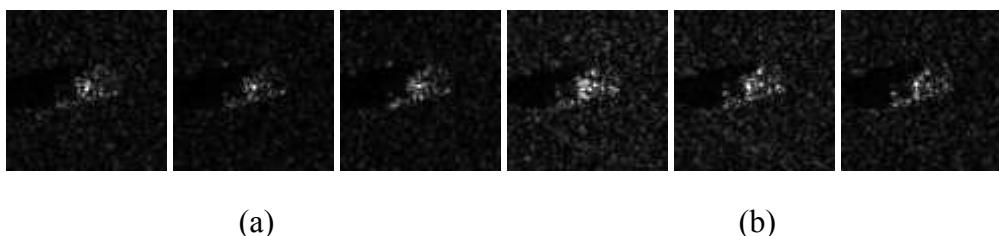


图 2.19 姿态角变化成像比较。图中给出的是姿态角分别为 $13^\circ, 14^\circ, 15^\circ$ 的 BMP2 (SN_9563) 在俯仰角为 (a) 15° 和 (b) 17° 时的 SAR 图像。

- **结构配置敏感性：**图 2.20 是姿态角为 45° 的 BMP2 三种不同配置 (SN_9563、SN_9566、SN_c21) 在 17° 雷达俯仰角的成像结果。可以看出，目标前身炮管部分、主炮塔部分的后向散射系数都有着显著变化，该现象也验证了论文第一章 1.2.2 节关于扩展工作条件的结构部件敏感性。

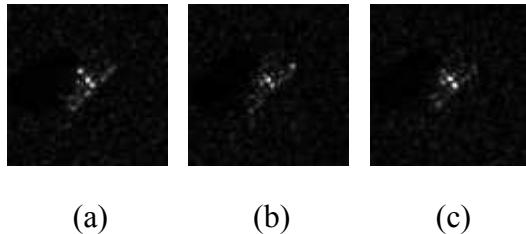


图 2.20 目标配置变化成像比较。图中给出的是姿态角为 45° 的 BMP2 (a) SN_9563、(b) SN_9566、(c) SN_c21 在俯仰角为 17° 的 SAR 图像比较。

2.4.3.2 频域稀疏表示的实验论证

为了弱化稀疏表示对于空间配准的苛刻条件，论文第 2.4.2 节中提出一种频域稀疏表示的分类方法，目的在于应对图像的空间平移以及噪声污染等问题。本节首先讨论不同频域特征的有效性，根据离散傅立叶变换即式 (2.27) 和式 (2.28) 有特征表示方法：

- 对角线区域幅度谱系数组成的特征向量；
- 对角线区域相位谱系数组成的特征向量；
- 对角线区域实部和虚部系数组合而成的特征向量。

为了进行实验比对，论文将三种特征描述的稀疏表示分类分别记为 SRC^1 , SRC^2 , SRC^3 。首先观察选取的频谱区域大小对分类性能的影响，尝试使用四种不同的掩模尺寸，如图 2.21 所示，即选取原始频谱中心部分的波段（对原始步谱进行搬移 fftshift），对应 100D、64D、25D、16D 傅立叶系数，其维度大小等同于对原始图像执行以 $\frac{1}{8}, \frac{1}{10}, \frac{1}{16}, \frac{1}{20}$ 为因子的下采样。对 BMP2, T72, BTR70, BRDM2 四种军事目标进行分类，每类样本的数据说明见表 2.5，实验结果如表 2.4 所示，可以看出：1) 幅度谱特征性能远远优于相位谱特征，也优于实部与虚部的组合特征；2) 实部与虚部的组合特征维度是其余两种特征的两倍，但它的实验结果并没有表现出很优异的性能；3) 100D 和 64D 特征优于 25D 和 16D 特征，说明 100D 和 64D 特征包含更多的鉴别信息；4) 特征尺寸从 64D 增加到 100D，实验准确率并没能显著提高，说明二者包含的鉴别信息相仿。通过这组验证实验可以看出，相位谱特

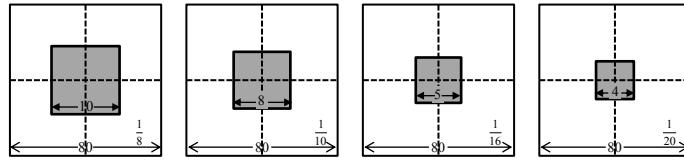
图 2.21 图中分别是大小为 $\frac{1}{8}, \frac{1}{10}, \frac{1}{16}, \frac{1}{20}$ 倍原始尺寸的掩模。

表 2.4 掩模尺寸对频域稀疏表示的影响，表中列出 3 种频域特征的 4 种掩模尺寸识别率。

特征	16D	25D	64D	100D
SRC ¹	0.9038	0.9367	0.9591	0.9559
SRC ²	0.7468	0.7516	0.7829	0.7756
SRC ³	0.8966	0.9143	0.9463	0.9423

征、实部与虚部的组合特征可靠性逊于幅度谱特征，因而后续实验将摒弃这两种特征方法，以幅度谱特征构成的分类算法记为 SRC^{FT}。

为了进一步验证频域稀疏表示算法的合理性和有效性，论文组织两类扩展条工作件的目标识别实验，包括目标配置变化与雷达俯仰角变化。

I. 目标配置变化实验：首先讨论不同结构配置的目标识别问题，选取了 BMP2、T72、BTR70、BRDM2 四类车辆目标，其中 BMP2 和 T72 分别有三种配置变体，型号分别以序列号 SN_9563、SN_9566、SN_c21 和 SN_132、SN_812、SN_s7 标记，BTR70 和 BRDM2 都只有单种型号，SN_C71 和 SN_E71，如表 2.5 所示。只选取 BMP2 和 T72 的单个型号 SN_9563 和 SN_132 进行训练，使用 17° 雷达俯仰角的数据训练分类算法，使用剩余四个型号的 15° 俯仰角数据进行测试，用于训练的型号不参与测试，提高实验难度。

表 2.5 目标配置变化实验数据，表中括号内加粗的数字表示训练用的 SAR 图像数据视角数目，下划线数字表示测试所用 SAR 图像数据视角数目。

类别	BMP2			BTR70	T72			BRDM2
型号	SN_9563	SN_9566	SN_c21	SN_c71	SN_132	SN_812	SN_s7	SN_E71
17°	[233]	232	233	[233]	[232]	231	228	[298]
15°	195	<u>196</u>	<u>196</u>	<u>196</u>	196	<u>195</u>	<u>191</u>	274

I. A 标准实验设置

标准实验条件是指训练数据和测试数据不做任何变动，进行实验。实验考察的方法主要有稀疏表示分类与支持向量机。输入的特征有空域特征与频域特征，其中空域特征是指原始强度矩阵做均匀降采样后的结果，频域特征是傅立叶幅

度谱特征，空域特征与频域特征等维，构成的分类算法分别记为 SRC、 SRC^{FT} 、SVM、 SVM^{FT} 。实验中的特征维数设为 16D、25D、64D、100D，分别是 $\frac{1}{20}, \frac{1}{16}, \frac{1}{10}, \frac{1}{8}$ 均匀降采样之后重排成的向量维数。实验结果如表 2.6 所示。

表 2.6 标准实验条件频域分类学习，表中列出空域表示和频域表示的 SVM 与 SRC 识别率。

维数	16D	25D	64D	100D
SVM	0.6931	0.7980	0.8798	0.8838
SVM^{FT}	0.8108	0.8934	0.9175	0.9190
SRC	0.7901	0.8942	0.9311	0.9503
SRC^{FT}	0.9038	0.9367	0.9591	0.9559

可以看出，若输入相同的特征向量，稀疏表示分类的正确率明显优于支持向量机的分类结果，这与已有的研究成果相吻合。若采用相同的分类方法，频域特征的分类结果远远优于图像域特征，特别是低维特征，如 16D 的 SVM^{FT} 比 SVM 的正确率高出 11.77%， SRC^{FT} 比 SRC 的正确率高出 11.37%。此外，若将特征设为 100D 维，SRC 与 SRC^{FT} 的分类结果十分接近，0.9503 Vs. 0.9559。这表明频域稀疏在标准实验条件并没有展示出明显的优越性。下面将重点关注扩展实验条件的性能比较。

I. B 扩展实验条件

从标准实验条件的实验结果可以看出整体表现都比较好，这是由于实验设置过于理想化，下面将会对实验设置进行调整，提高实验难度，观察各种学习算法对于苛刻实验条件的实际表现。

a. 空间平移。为了弱化空间配准的约束条件，本文提出频域稀疏表示分类，实现平移不变和扭曲、失真不变。为了验证算法的有效性，对测试图像的成像数据进行调整，使其在空域产生不同程度地偏移，同时训练图像保持不变，迫使测试图像不与训练图像严格配准，模拟实际的战场环境。图 2.22 给出了一组图像空间手动平移的实例。为了观察图像偏移对分类算法的影响，重新组织四类目标识别实验，与前面不同，所有的测试图像（即俯仰角为 15° 的成像数据）都做不同程度的空间平移，对平移之后的图像进行分类，结果如表 2.7 所示。由于频域稀疏表示受相关滤波器算法的启发，因此基于相关滤波的识别方法将作为基准方法进行比较，其中涉及的相关滤波器包括 MINACE 滤波器 [116] 和 OTSDF 滤波器 [71]。

可以看出，空域的识别方法，SRC 和 SVM，随着空间平移程度的加剧，其识别率下降得十分明显，正确率几乎都降到 40% 以下。与之相反，频域的分类方法表现较为理想，分类性能基本不受图像平移影响。与相关滤波器相比，频域稀疏表示的性能远远优于 MINACE 滤波器和 OTSDF 滤波器的表现，分类正确率高出

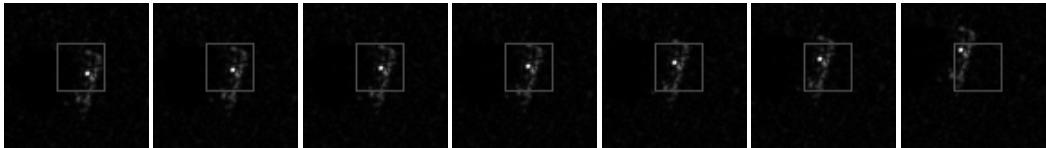


图 2.22 图像的空间平移示意图，图中从左至右分别是 -3 像素、-1 像素、0 像素、1 像素、3 像素、5 像素、10 像素平移后的图像，其中 0 像素偏移就是原始的配准图像。这里定义沿第四象限方向的偏移是正向偏移，沿第二象限方向的偏移是负向偏。图中的白色方框是原始图像的目标中心。

表 2.7 空间偏移对频域分类学习的影响，表中列出了 7 种不同程度空间偏移的识别正确率。

偏移程度	-3-pixel	-1-pixel	0-pixel	1-pixel	3-pixel	5-pixel	10-pixel
MINACE	0.8718	0.8582	0.8694	0.8702	0.8582	0.8654	0.8486
OTSDF	0.8365	0.8454	0.8462	0.8470	0.8486	0.8373	0.8470
SRC	0.6106	0.8926	0.9503	0.9271	0.6955	0.5385	0.3838
SVM	0.6995	0.8613	0.8838	0.8661	0.7115	0.5024	0.3060
SVM ^{FT}	0.8782	0.8950	0.8918	0.8926	0.8910	0.8741	0.8092
SRC ^{FT}	0.9543	0.9543	0.9559	0.9543	0.9575	0.9583	0.9479

约 10 个百分点左右。这是因为频域稀疏表示兼具了傅立叶频谱平移不变性和稀疏信号表示的优良特性。

b. 桑声污染。相干成像传感器 (SAR) 获取的图像中遍布着乘性噪声 (相干斑)，因此，良好地抗噪性能应该是 SAR ATR 所必备的。为了模拟图像中的噪声，从测试图像中随机挑选出部分像素，将其替换成为服从独立同分布的均匀分布的样本点 [156]。由于噪声的不可预知性，被污染的像素在图像中的所在位置是随机确定的。图 2.24 给出了一幅 SAR 切片图像及其噪声污染后的结果。

为了观察随机噪声污染对分类算法的影响，论文给所有的测试图像都施加不同程度的随机噪声污染，然后进行分类识别实验。图 2.25 是频域稀疏表示与基准算法在不同等级的噪声污染的识别率变化曲线，对应的识别率如表 2.8 所列。可以看出，随着噪声强度的逐渐提升，分类算法的识别性能都有不同程度地下降。当噪声等级从 1% 增加到 30%，频域稀疏表示 (SRC^{FT}) 的识别率分别为 0.9663，0.9159，0.9030，0.8918，0.8798 和 0.8766。基准算法中与其最接近的至少存在着 1.92%，3.14%，9.45%，11.54%，12.34% 和 18.87% 的性能差距。当噪声等级为 30% 时，频域稀疏依然取得了 0.8766 的正确率，相应地 MINACE 滤波器正确率为 0.6002，OTSDF 滤波器正确率为 0.6811，SRC 正确率是 0.5048，SVM 正确率是 0.3718，SVM^{FT} 正确率是 0.6161。在噪声等级的提升区间内，频域稀疏的正确

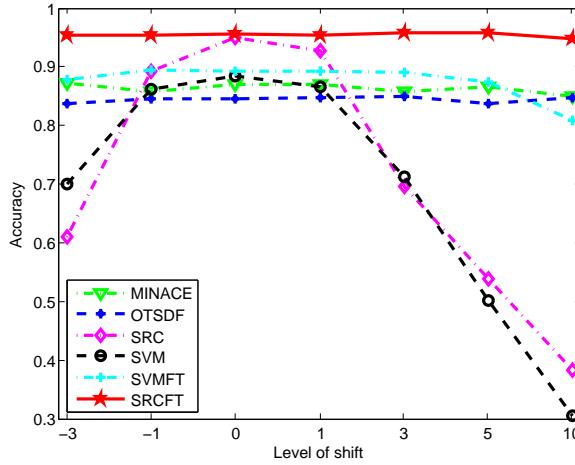


图 2.23 频域稀疏表示随图像空间偏移正确率曲线。

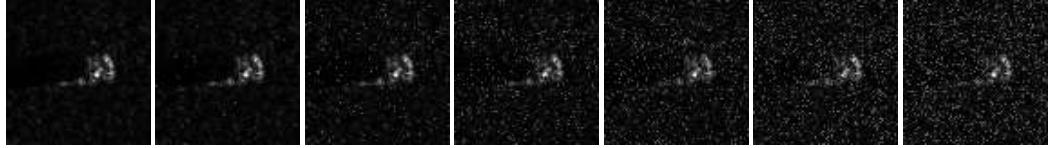


图 2.24 图像随机噪声示意图。图中给出了 SAR 切片图像中 0%、1%、5%、10%、15%、20%、30% 的像素受到随机噪声污染。

率下降了 8.49%，而 MINACE 滤波器、OTSDF 滤波器、SRC、SVM 和 SVM^{FT} 分别降低了 25.90%，15.86%，44.23%，52.24% 和 27.89%。

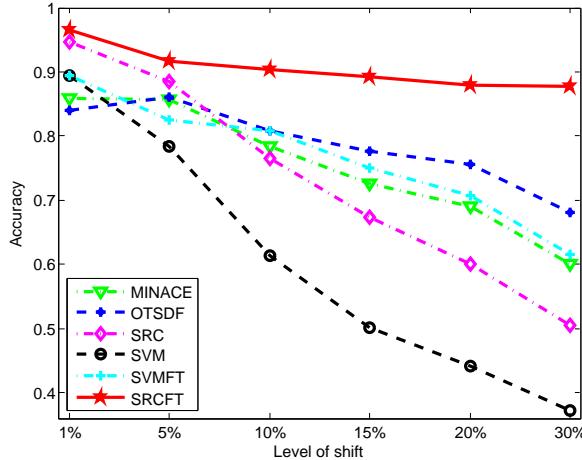


图 2.25 频域稀疏表示随噪声污染等级的正确率曲线。

II. 倾仰角变化实验：接下来的实验用于评估雷达倾仰角变化对目标识别性能的影响。MSTAR 数据集有三类目标包含多个雷达倾仰角的成像数据（如图 2.26 所示）：2S1、BRDM2 和 ZSU23/4，其中 BRDM2 和 ZSU23/4 还存在连接体变体。所谓的连接体变体是指车辆本身的物理变化，如打开防护盖、旋转炮塔、升降炮管等。图 2.27 给出了一组目标连接体变体的图示。为了观察雷达倾仰角变化以分类

表 2.8 噪声污染对频域分类学习的影响，表中列出算法受 6 种不同等级噪声污染的正确率。

算法	噪声等级						
	1%	5%	10%	15%	20%	30%	下降幅度
MINACE	0.8592	0.8574	0.7845	0.7260	0.6899	0.6002	25.90%
OTSDF	0.8397	0.8598	0.8085	0.7764	0.7564	0.6811	15.86%
SRC	0.9471	0.8846	0.7660	0.6739	0.6002	0.5048	44.23%
SVM	0.8942	0.7844	0.6129	0.5016	0.4415	0.3718	52.24%
SVM ^{FT}	0.8950	0.8253	0.8084	0.7508	0.7075	0.6161	27.89%
SRC ^{FT}	0.9663	0.9159	0.9030	0.8918	0.8798	0.8766	8.97%

算法的影响，使用 17° 俯仰角的成像数据进行训练，利用 30° 和 45° 俯仰角的成像数据进行测试。不同雷达俯仰角的目标的不同视角成像样本数见 2.4.3.1 节表 2.3。分类算法随俯仰角变化的正确率如表 2.9 所列，相应正确率变化曲线如图 2.28。

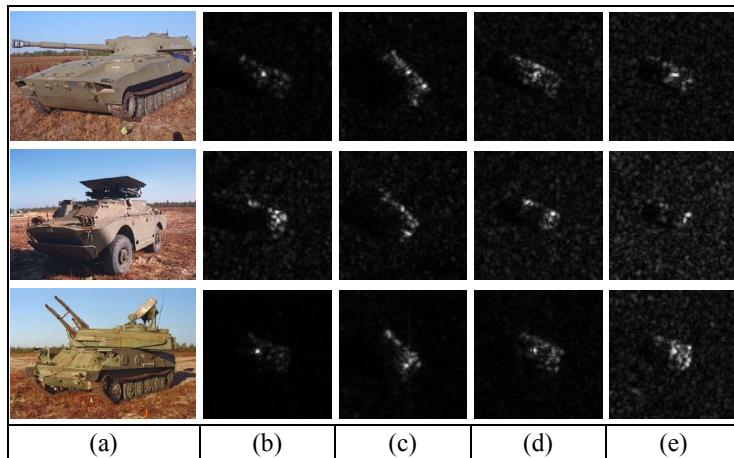


图 2.26 三类车辆目标在 4 种雷达俯仰角 SAR 图像。第一行、第二行和第三行分别是 2S1、BRDM2 和 ZSU23/4 在 (a) 15° , (b) 17° , (c) 30° , (d) 45° 俯仰角的 SAR 图像。

从实验结果来看，若以 15° 俯仰角的成像数据进行测试，所有算法的正确率都很完美，OTSDF 滤波器甚至达到了 0.9987。然而，该方法需要进行未知样本的姿态角估计，且需要部分测试样本训练折衷参数，实际意义不大。若增加俯仰角至 30° ，除了 SRC 和 SRC^{FT} 之外的算法性能都有显著下降，SRC 和 SRC^{FT} 分别取得了 0.9641 和 0.9632 的正确率，其余方法的正确率均低于 0.92。若进一步增加俯仰角至 45° ，所有方法的识别率都急剧下降，相关滤波器的方法识别率甚至低于 0.5。频域稀疏表示 SRC^{FT} 仍取得了 0.7030 的正确率，比最接近的三种基准方法高出 8.3%，10.2% 和 25.4%。

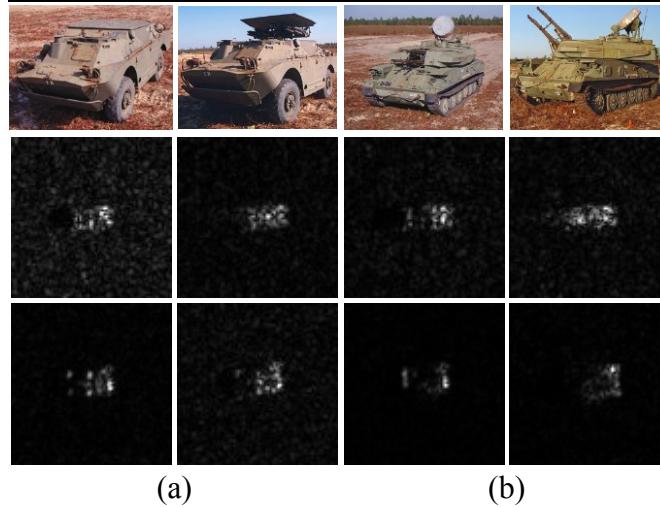


图 2.27 连接体变体图示。图中，第一行是变体的光学图像，第二行和第三行分别是 (a) BRDM2，(b) ZSU23/4 在俯仰角为 30° 和 45° 的 SAR 图像，其中目标的方位角为 4° 。

表 2.9 俯仰角变化对频域分类学习的影响，表中列出算法的 15° 、 30° 、 45° 测试样本识别率。

俯仰角	OTSDF	MINACE	SVM	SVM ^{FT}	SRC	SRC ^{FT}
15°	0.9987	0.9007	0.9647	0.9708	0.9927	0.9964
30°	0.7621	0.6679	0.8934	0.8950	0.9641	0.9623
45°	0.4486	0.3998	0.4207	0.6001	0.6193	0.7030

小结：根据目标配置变化与雷达俯仰角变化两类实验结果有如下结论：1) 稀疏表示分类的表现优于支持向量机、相关滤波器等方法；2) 相比较图像域的分类方法（如 SRC 和 SVM），频域分类方法（SRC^{FT}、SVM^{FT}、MINACE 和 OTSDF）对于空间平移和噪声污染有着更强的鲁棒性；3) 频域稀疏表示在扩展实验条件，如空间平移、随机噪声污染、俯仰角变化等能展现出更显著的优越性，更适应于实际应用。

2.4.3.3 DS 证据推理决策的实验验证

I. 数字实例：为了直观地说明基于 DS 证据推理的“软”决策方法，论文首先给出两组实例，其中基于贝叶斯估计的“硬”决策方法作为基准算法。选择 BMP2、T72、BTR60 和 T62 共 4 类军事目标进行分类实验，记为 $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \mathcal{L}_4$ 。实验结果如表 2.10 和表 2.11 所示。

第一组测试样本源自于 BMP2，其真实标记为 \mathcal{L}_1 。从表 2.10 列出的数据可以看出，共有 11 组基原子对应的表示系数不为零，它们共同参与了未知样本的线性表示，其中 8 组基原子属于 \mathcal{L}_1 ，残差分别为 0.7419、0.7056、0.7629、0.7503、0.7334、0.7091、0.8030、0.7416，整体重构误差为 0.1707，剩余的三组原子分别属

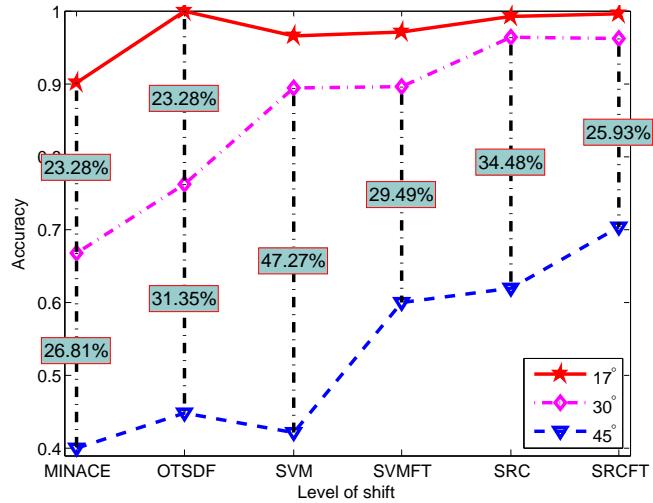


图 2.28 频域稀疏表示随俯仰角变化的正确率曲线

表 2.10 基于证据推理的“软”决策方法示例一。

Label	Bayesian		Dempster-Shafer			
	e^j	e_{SRC}	$m^j(C_j)$	$m^j(\Theta_j)$	$m(L_q)$	$m(\Theta)$
\mathcal{L}_1	0.7419		0.4524	0.5476		0.0081
\mathcal{L}_1	0.7056		0.4691	0.5309		
\mathcal{L}_1	0.7629		0.4430	0.5570		
\mathcal{L}_1	0.7503	0.1707	0.4486	0.5514	0.9713	
\mathcal{L}_1	0.7334		0.4563	0.5437		
\mathcal{L}_1	0.7091		0.4675	0.5325		
\mathcal{L}_1	0.8030		0.4256	0.5744		
\mathcal{L}_1	0.7416		0.4525	0.5475		
\mathcal{L}_2	0.6973	0.7416	0.4730	0.5270	0.0069	
\mathcal{L}_3	0.6929	0.7566	0.4751	0.5249	0.0075	
\mathcal{L}_4	0.7873	0.9527	0.4323	0.5677	0.0062	

于其它三类 \mathcal{L}_2 , \mathcal{L}_3 和 \mathcal{L}_4 , 对应的重构误差分别为 0.7416、0.7566 和 0.9527¹⁵。两种决策方法都给出了正确的决策, \mathcal{L}_1 对于未知样本的重构误差是 0.1707, DS 证据理论得出未知样本属于 L_1 的置信度为 0.9713。尽管两种决策方法都能正确地推测未知样本类别, 但是 DS 证据理论决策的置信水平优于传统的贝叶斯决策¹⁶。

每二组测试样本来自于 BMP2, 其真实标记为 \mathcal{L}_1 。从表 2.11 列出的数据可以看出, 共有 11 组基原子对应的表示系数不为零, 它们共同参与了未知样本的线性表示, 其中 4 组原子来自于 \mathcal{L}_1 , 2 组原子来自于 \mathcal{L}_2 , 4 组原子来自于 \mathcal{L}_3 , 一

¹⁵每个类别的整体重构误差没有做归一化处理, 因此, 四个类别的重构误差之和不等于单位值。

¹⁶决策的置信水平定义为不同类别之间的置信度(或似然概率)的差。

表 2.11 基于证据推理的“软”决策方法示例二。

	<i>Bayesian</i>		<i>Dempster-Shafer</i>			
<i>Label</i>	\mathbf{e}^j	\mathbf{e}_{SRC}	$m^j(C_j)$	$m^j(\Theta_j)$	$m(L_q)$	$m(\Theta)$
\mathcal{L}_1	0.6465	0.4899	0.4977	0.5023		
\mathcal{L}_1	0.6763		0.4831	0.5169		
\mathcal{L}_1	0.6908		0.4761	0.5239	0.4333	0.0330
\mathcal{L}_1	0.6835		0.4796	0.5204		
\mathcal{L}_2	0.6217	0.3745	0.5102	0.4898		0.0330
\mathcal{L}_2	0.5654		0.5397	0.4603	0.1134	
\mathcal{L}_3	0.7236		0.4608	0.5392		
\mathcal{L}_3	0.6853	0.6729	0.4787	0.5213		
\mathcal{L}_3	0.6980		0.4727	0.5273	0.3922	
\mathcal{L}_3	0.6907		0.4762	0.5238		
\mathcal{L}_4	0.7254	0.9732	0.4599	0.5401	0.0281	

组原子来自于 \mathcal{L}_4 ，分布比较均匀，因而比较有代表性。从结果可知，两种决策方法得出了截然不同的结论，四个类别的重构误差是 [0.4899, 0.3745, 0.6729, 0.9732]，贝叶斯估计认为该样本属于 \mathcal{L}_1 ，误差为 0.3745；与之不同，DS 证据理论推导出全局证据质量为 [0.4333, 0.1134, 0.3922, 0.0281, 0.0330]，对应的假设分别是 $\mathbf{y} \in \mathcal{L}_1, \mathbf{y} \in \mathcal{L}_2, \mathbf{y} \in \mathcal{L}_3, \mathbf{y} \in \mathcal{L}_4, \Theta$ ，得出的推论是 $\mathbf{y} \in \mathcal{L}_1$ 。贝叶斯估计给出的是错误结论，DS 证据理论正确识别未知样本，这是因为 DS 证据理论对不确定类别及不精确性进行假设建模。

上面通过两组实例说明 DS 证据“软”决策的工作原理，下面将使用 MSTAR 数据集对方法的有效性进行验证，其中图像的尺寸与前文介绍的相同，80×80。而定义基本概率赋值时所涉及到的参数，即公式 (2.35) 和公式 (2.37) 中的 η_0, λ ，它们分别根据经验设置为 $\eta_0 = 0.95$ 和 $\lambda = 1$ 。

II. 实验验证：前面使用两组实例说明 DS 证据“软”决策的推断原理，接下来将借助 MSTAR 实测数据组织识别实验，通过与经典的贝叶斯估计的性能比较，说明 DS 证据“软”决策的优良特性。

II. A. 目标配置变化：选择 BMP2、BTR60、T72 和 T62 四类目标在不同配置与雷达俯仰角的成像数据进行实验，其中 17° 雷达俯仰角的成像数据用于训练算法，15° 雷达俯仰角的成像数据进行性能测试，BMP2 和 T72 包含有三种物理变体，与前面的实验设置相同，仅使用 SN_9563 和 SN_132 进行训练，剩余四种型号进行测试，即测试数据和训练数据的俯仰角和目标配置都有变化，详见表 2.12。

表 2.12 四类目标在不同俯仰角的视数。

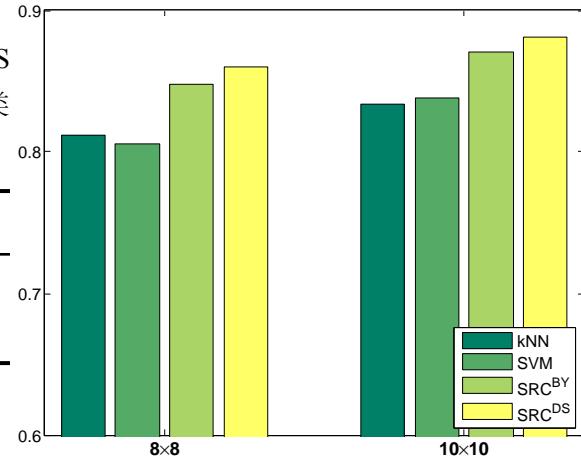
<i>Depr.</i>	BMP2	T72	BTR60	T62
17°	233 (SN_9563)	232 (SN_132)	256	299
15°	196 (SN_9566) 196 (SN_c2I)	195 (SN_812) 191 (SN_s7)	195	273

考虑到最小化 ℓ_1 范数优化问题的计算效率，将原始 SAR 图像进行以 ρ 为因子的均匀降采样，降低数据维数并去除冗余信息，采样因子设为 $\rho = \{\frac{1}{10}, \frac{1}{8}\}$ ，对应的采样结果为 8×8 和 10×10 像素。实验结果如表 2.13 和表 2.14 所示，其中 SRC^{BY} 表示传统的贝叶斯估计“硬”决策，SRC^{DS} 记为 DS 证据推理“软”决策。

表 2.14 配置变化实验两种决策方法比较。

表 2.13 目标配置变化实验贝叶斯估计与 DS “软”决策比较，表中给出两种采样因子的分类正确率。

维度	kNN	SVM	SRC ^{BY}	SRC ^{DS}
8×8	0.8122	0.8058	0.8475	0.8604
10×10	0.8339	0.8379	0.8708	0.8812



从实验结果可以看出，若采用 8×8 的采样因子对原始图像进行均匀降采样，传统的贝叶斯决策 SRC^{BY} 和基于 DS 证据融合的“软”决策 SRC^{DS} 分类正确率分别为 0.8475 和 0.8604，“软”决策比“硬”决策高 1.29%，另外两种基准算法线性核支持向量机和最近邻分类器的识别率分别为 0.8058 和 0.8122。相似的结论也可从原始图像的 10×10 下采样得出。这组实验结果展示了稀疏表示分类相比传统的训练方法的优势。若采用 $\frac{1}{10}$ 的采样因子进行均匀降采样，DS 证据推理的“软”决策分类正确率为 0.8604，比传统的贝叶斯“硬”决策高出 1.29%，若采用 $\frac{1}{8}$ 的采样因子进行均匀降采样，DS 证据推理的“软”决策分类正确率为 0.8812，比传统的“硬”决策方法高出 1.04%。

3. 十类目标识别

最后将实验扩展至 10 类目标，表 2.2 中的所有目标同时都用于实验。对于多变体目标，BMP2 和 T72，使用标准型号 (SN_9563 和 SN_132) 的 17° 俯仰角成像数据进行训练，所有型号的 15° 俯仰角成像数据进行测试，详见表 2.2。采样因子设置为 $\frac{1}{8}$ ，对应的特征维数为 100D。实验结果如图 2.29 所示。

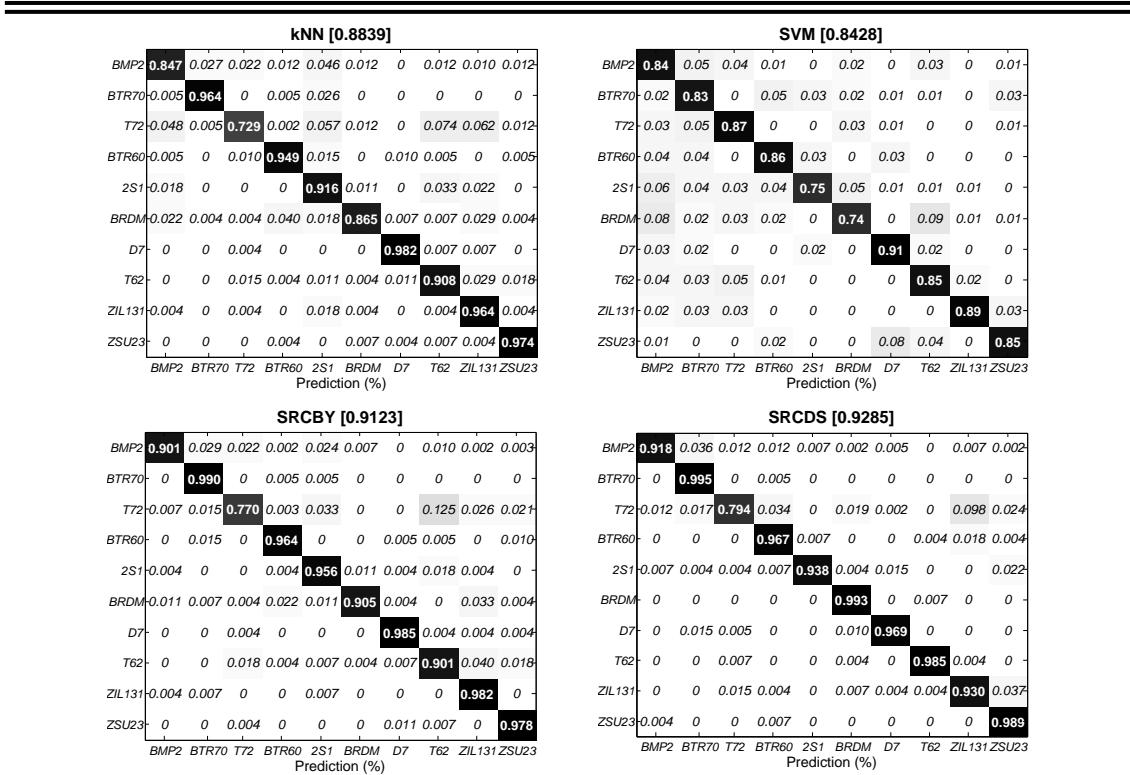


图 2.29 十类目标识别实验 DS “软” 决策与经典“硬”决策的比较，每幅子图的行表示未知样本的真实类别，列表示分类算法估计出的类别；对角线方格内的数字是每类目标的正确率，非对角线方格中的数字是误分率；顶端的括号中给出的是整体识别。

从四种方法的混淆矩阵来看，十类目标识别实验结论与前面类似，稀疏表示分类方法优于其余的两种基准比较算法。基于贝叶斯“硬”决策的稀疏表示分类 SRC^{BY} 的整体识别率为 0.9123，与之对应的基准算法 SVM 和 kNN 的分类正确率分别为 0.8428 和 0.8839，稀疏表示分类方法的表现比线性核支持向量机和最近邻分类分别高 6.95% 和 2.84%。基于 DS 证据融合“软”决策的稀疏表示分类方法， SRC^{DS} 整体识别率为 0.9285，比传统的“硬”决策方法高出了 1.62%。实验结果进一步验证了 DS 证据理论比传统的贝叶斯估计的优点，即能够对不确定状态的不准确性进行度量。

2.5 本章小结

本章主要介绍稀疏表示与压缩感知基础理论。首先简要回顾稀疏表示从小波变换到多尺度分析以及压缩感知的发展历程，梳理稀疏表示的典型应用以及遇到的“瓶颈”问题。围绕着稀疏表示遇到的实际问题，引出压缩感知这个新的调和分析理论，归纳了压缩感知的工作原理、测量矩阵、稀疏求解等关键技术。之后，借助传统的正交编码引出论文研究的重点内容之一：基于过完备冗余字典的稀疏

表示，概括稀疏表示在遥感图像解译及相关领域的主要应用。针对 SAR 图像目标识别，分别从表示方法和决策规则两方面入手，提出改进方法：

1) 表示方法改进——频域稀疏表示。受相关滤波器的启发，本文使用信号的傅立叶变换频谱构造过完备冗余字典，利用字典原子的线性组合表示未知样本的频谱，生成线性回归模型，通过求解回归系数进行判决；

2) 决策规则改进——DS 证据推理“软”决策。首先找出过完备冗余字典中参与未知样本线性表示的所有原子，本文认为每个原子都能够提供一条关于未知样本类别归属的证据，通过原子的表示误差量化证据质量，根据 Dempster 规则进行证据合并，进行判决。

第三章 基于单演信号的 SAR 图像目标特性描述

单演信号是一维解析信号在高维信号空间的广义扩展，是由 Michael Felsberg 和 Gerald Sommer 于 2001 年在文献 [219] 中首次提出，在文献 [220] 和 [221] 中做进一步完善形成单演信号尺度空间理论。早在单演信号问世之前，许多学者也曾对解析信号在高维信号空间的扩展作了相关尝试，其中比较有代表性的是以下两类方法：第一类方法倾向于将 Hilbert 变换分别沿两个方向轴单独执行，如学者 S. Hahn 将解析信号表示为两个独立的复分量，将这两个复分量的傅立叶频谱分别限制在第一象限和第四象限 [222]，从而形成具有 4 个冗余度的双树复小波变换，其优势在于平移不变性和方向选择性；第二类方法是分别使用相同的独立 Hilbert 变换元素，然后借助一种更加复杂的四元素代数理论将其组合成为四重积分，如学者 T. Bülow 和 G. Sommer 提出首先分别构造 Hilbert 变换分量，然后将其组合成四重积分 [223]。

与这些方法相比，单演信号有其特有的优点，即它的潜在的特征提取过程（转换成极坐标表示为局部幅度、局部相位、局部方位）具有旋转不变和尺度不变等特性（见 3.1.2 所述）。这也是它在信号分析和图像处理中受到广泛研究的重要原因。单演信号在信号分析中的应用包括单演小波变换、Riesz-Laplace 小波变换、单演曲线波变换、随机单演信号、高阶 Riesz 变换小波、可变 Riesz 小波、向量化单演小波等 [224–229]，它在图像分析中的应用包括：边缘检测、图像局部结构分析、运动估计、光流场估计、图像配准、图像分割、模式识别等 [230–239]。

本章主要介绍 Riesz 变换和单演信号两种信号分析工具，根据定义总结归纳其主要特性，据此提出论文的研究动机。本章与第二章是论文的基础部分，目的是为后续第四章、第五章和第六章的核心部分做铺垫，详见绪论章节的图 1.25 所示。本章内容包括：1) Riesz 变换及其主要性质；2) 单演信号及其主要性质；3) 单演信号对于目标散射特性的变化分析；4) 图例论证。

3.1 Riesz 变换与单演信号

3.1.1 解析信号

解析信号是一种信号描述与分析方法，自提出来在许多领域（如雷达信号处理）都受到广泛应用。任意给定的一维实信号 $f(x), x \in \mathbb{R}$ ，解析信号 $f_A(x)$ 定义为信号的复数化函数，其中实部为信号本身，虚部（共轭函数）是它的 Hilbert 变换

$$f_A(x) = f(x) - i \cdot f_H(x) \quad (3.1)$$

式中 $f_H(\cdot) = f(x) * h(x)$ 表示原始信号的 Hilbert 变换，变换核为 $h(x) = \frac{1}{\pi x}$ ，相应的频率响应为 $H(w) = -j \cdot \text{sgn}(w)$ 。将解析信号转换成极坐标，可表示为局部幅度 $A(x)$ 和局部相位 $\varphi(x)$ ，

$$\begin{aligned} A(x) &= \|f_A(x)\| = \sqrt{f^2(x) + f_H^2(x)} \\ \varphi(x) &= \text{atan}2(f_H(x), f(x)), \varphi(x) \in [0, 2\pi] \end{aligned} \quad (3.2)$$

局部相位随着信号的局部结构变化而变化，关于信号的局部能量不变；同样地，局部幅度表示信号的局部能量，它关于信号的局部结构不变。图 3.1 列举了几种典型的特征类型与相位之间的关系 [240]。

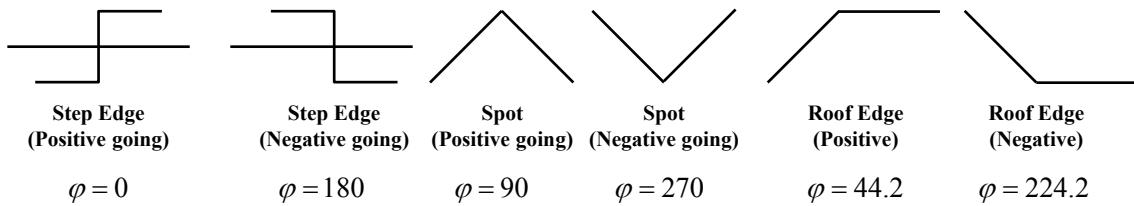


图 3.1 六种典型特征及其对应的相位。

解析信号广泛应用于信号的分析与处理。若要将解析信号扩展到高维信号空间，首先需要对 Hilbert 变换作各向同性的广义延伸，即寻找出具有等向能量分布的奇函数滤波器，然而该问题无论是在实数域或是复数域均没有通用解。目前为止，一个有效的解决方案就是向量化 Hilbert 变换即 Riesz 变换（就是多维等向广义 Hilbert 变换）。

3.1.2 Riesz 变换

对于任意 n 维信号 $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n$ ，它的 Riesz 变换 $f_{\mathcal{R}}(\cdot)$ 定义为

$$f_{\mathcal{R}_j}(\mathbf{x}) \equiv \mathcal{R}_j(f)(\mathbf{x}) = \lim_{\epsilon \rightarrow \infty} c_n \int_{\|\mathbf{y}\| > \epsilon} \frac{y_j}{\|\mathbf{y}\|^{n+1}} f(\mathbf{x} - \mathbf{y}) d\mathbf{y} \quad (3.3)$$

其中 $c_n = \frac{\Gamma[(n+1)/2]}{\pi^{(n+1)/2}}$ ， c_n 是取决于维数 n 的常数， $\Gamma(\cdot)$ 是 Gamma 函数， $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 是位置索引向量， $\mathbf{y} = (y_1, y_2, \dots, y_n)$ 是虚变量， $\|\cdot\|$ 是向量范数。以卷积核的形式可将 Riesz 变换的定义改写为

$$K_j(\mathbf{x}) = c_n \frac{x}{\|\mathbf{x}\|^{n+1}}, j = 1, \dots, n \quad (3.4)$$

任意的 n 维信号 $f : \mathbb{R}^n \mapsto \mathbb{R}$ ，总存在 n 个 Riesz 变换与之对应。数学家 E. M. Stein 指出并证明了 Riesz 变换就是 Hilbert 变换在多维空间中的推广 [241]。对于给定的 n 维实信号 $f(\mathbf{x})$ ，它的共轭函数恰巧是向量化的函数，

$$f_{\mathcal{R}}(\mathbf{x}) = \begin{bmatrix} f_{\mathcal{R}_1}(\mathbf{x}) \\ f_{\mathcal{R}_2}(\mathbf{x}) \\ \vdots \\ f_{\mathcal{R}_n}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} (K_1 * f)(\mathbf{x}) \\ (K_2 * f)(\mathbf{x}) \\ \vdots \\ (K_n * f)(\mathbf{x}) \end{bmatrix}$$

即任意 n 维信号的共轭函数在每个索引位置都包含 n 个元素。若将高维信号 $f(\mathbf{x})$ 的傅立叶变换 \mathcal{F} 定义为

$$\mathcal{F}(f)(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{x}) \exp(-i2\pi\langle \mathbf{u}, \mathbf{x} \rangle) d\mathbf{x}$$

其中 $\mathbf{u} = (u_1, u_2, \dots, u_n)$ 是频域向量， $\langle \cdot, \cdot \rangle$ 是内积运算（矢量或标量），则第 j 个 Riesz 变换核的傅立叶变换为

$$\mathbb{K}_{\mathcal{R}_j}(\mathbf{u}) = -\frac{u_j}{\|\mathbf{u}\|}, j = 1, \dots, n$$

由卷积定理 [242] 可知信号 f 的第 j 个 Riesz 变换的传递函数为

$$\mathcal{R}_j(r)(\mathbf{u}) \equiv F_{\mathcal{R}_j}(\mathbf{u}) = -\frac{u_j}{\|\mathbf{u}\|} F(\mathbf{u})$$

以二维图像信号为例，Riesz 变换的数值实现如列表所示，生成的两个方向变换核见图 3.2，相应的数值实现如下列表。

```

1 function [Hx, Hy] = Riesz_Kernel(nRow, nCol)
2
3 [u1, u2] = meshgrid(((1:nCol)-(fix(nCol/2)+1))/(nCol-mod(nCol
4 , 2)), ((1:nRow)-(fix(nRow/2)+1))/(nRow-mod(nRow, 2)));
5 u1 = ifftshift(u1);
6 u2 = ifftshift(u2);
7
8 radius = sqrt(u1.^2 + u2.^2);
9 radius(1,1) = 1;
10
11 Hx = 1i*u1./radius; % 虚部x轴的卷积核的频率响应
12 Hy = 1i*u2./radius; % 虚部y轴的卷积核的频率响应
13
14 end

```

显然，频域 Riesz 变换核（用 $\mathbb{K}(\cdot)$ 表示）所刻画的向量化线性滤波器 $\mathbb{K}(\mathbf{u}) = (-\frac{u_1}{\|\mathbf{u}\|}, -\frac{u_2}{\|\mathbf{u}\|}, \dots, -\frac{u_n}{\|\mathbf{u}\|})$ 完全满足构建多维解析信号的需要：即它具备等向能量分布，它在任意方向上的都是单位能量，且是关于坐标中心的奇函数。根据 Riesz 变换的定义可知它满足如下特性：

反对称性（关于坐标中心奇对称） $\mathbb{K}(-\mathbf{u}) = -\mathbb{K}(\mathbf{u})$ 。作为 Hilbert 变换的广义扩展，对称性是必须满足的基础条件。二维空间的对称通常是针对一个点或是一条

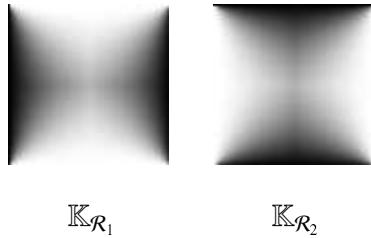


图 3.2 二维信号的两个方向 Riesz 变换核 (幅度矩阵)。

线。Riesz 变换是关于点对称，另一种扩展算法 — 局部 Hilbert 变换 (Partial Hilbert Transform) [243] 是关于某个偏爱的方向线对称。

平移不变性 (Translation-Invariant) 根据 Riesz 变换的时域定义有：

$$\forall \mathbf{x}_0 \in \mathbb{R}^2, \mathcal{R}_j\{f(\cdot - \mathbf{x}_0)\}(\mathbf{x}) = \mathcal{R}_j\{f(\cdot)\}(\mathbf{x} - \mathbf{x}_0)$$

尺度不变性 (Scale-Invariant) 根据 Riesz 的频域变换核可知：

$$\forall a \in \mathbb{R}^+, \mathcal{R}_j\{f(\cdot/a)\}(\mathbf{x}) = \mathcal{R}_j\{f(\cdot)\}(\mathbf{x}/a)$$

方向可控 (Steerability)¹

$$\mathbb{K}_{\mathcal{R}_j}(\mathbf{R}_\theta \mathbf{x}) = \exp\{-j\theta\} \mathbb{K}_{\mathcal{R}_j}(\mathbf{x})$$

其中 \mathbf{R}_θ 为 $\begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$

酉变换

$$\mathcal{R}^* f(\mathbf{x}) \longleftrightarrow \frac{(+j \cdot u_1 + u_2)}{\|\mathbf{u}\|} F(\mathbf{u})$$

其中 $\mathcal{R}^* = \mathcal{R}^{-1}$ 。

内积守恒 (Inner-Product Preservation) :

$$\forall f, \phi \in L_2(\mathbb{R}^n), \langle \mathcal{R}(f), \mathcal{R}(\phi) \rangle_{L_2^n} = \sum_{k=1}^n \langle \mathcal{R}_k(f), \mathcal{R}_k(\phi) \rangle_{L_2} = \langle f, \phi \rangle_{L_2}$$

¹Steerability 是指能使用一组基函数的线性组合来合成滤波器经过任意角度旋转之后的响应。

证明：内积守恒特性可根据傅立叶变换的 Parseval 关系得出，具体地推导过程如下：

$$\begin{aligned}
 \sum_{k=1}^n \langle \mathcal{R}_k(f), \mathcal{R}_k(\phi) \rangle_{L_2} &= \frac{1}{(2\pi)^n} \sum_{k=1}^n \int_{\mathbb{R}^n} -j \frac{u_k}{\|\mathbf{u}\|} \hat{f}(\mathbf{u}) j \frac{u_k}{\|\mathbf{u}\|} \hat{\phi}^*(\mathbf{u}) du_1 du_2 \cdots du_n \\
 &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \frac{\sum_{k=1}^n u_k^2}{\|\mathbf{u}\|_2^2} \hat{f}(\mathbf{u}) \hat{\phi}^*(\mathbf{u}) du_1 du_2 \cdots du_n \quad (3.5) \blacksquare \\
 &= \frac{1}{(2\pi)^n} \langle \hat{f}, \hat{\phi} \rangle_{L_2} = \langle f, \phi \rangle_{L_2}
 \end{aligned}$$

正交特性 (Orthogonality) 若给定一个实函数 $\psi(\mathbf{x}) \in L_2(\mathbb{R}^n)$ ，并且它的傅立叶变换 $\Psi(\mathbf{u})$ 满足点对称关系，

$$\|\Psi(u_1, u_2, \dots, u_n)\|^2 = \|\Psi(\pm u_1, \pm u_2, \dots, \pm u_n)\|^2$$

则必有

$$\begin{aligned}
 \langle \mathcal{R}_k(\psi), \psi \rangle_{L_2} &= 0, \quad \text{for } k = 1, 2, \dots, n \\
 \langle \mathcal{R}_{k_1}(\psi), \mathcal{R}_{k_2}(\psi) \rangle_{L_2} &= 0, \quad \text{when } k_1 \neq k_2 \quad (3.6)
 \end{aligned}$$

证明：该性质同样也可根据 Parseval 定理来证明，推导过程如下：

$$\langle \mathcal{R}_{k_1}(\psi), \mathcal{R}_{k_2}(\psi) \rangle_{L_2} = \left[\begin{array}{l} = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} -j \frac{u_{k_1}}{\|\mathbf{u}\|} \hat{f}(\mathbf{u}) j \frac{u_{k_2}}{\|\mathbf{u}\|} \hat{\phi}^*(\mathbf{u}) du_1 du_2 \cdots du_n \\ = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} u_{k_1} u_{k_2} \frac{\|\Psi(\mathbf{u})\|^2}{\|\mathbf{u}\|^2} du_1 du_2 \cdots du_n \\ = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} u_{k_1} \underbrace{\int_{\mathbb{R}^{n-1}} u_{k_2} \frac{\|\Psi(\mathbf{u})\|^2}{\|\mathbf{u}\|^2} du_1 du_2 \cdots du_n du_{k_1}}_{\psi(u_{k_1})} \end{array} \right] \blacksquare$$

3.1.3 单演信号

Hilbert 变换可以构造一维信号的解析信号表示，类似地，Riesz 变换可以生成多维信号的解析信号表示。本篇论文重点讨论二维有限长信号（即图像）的解析信号表示，也就是 Riesz 变换中的维数 $n = 2$ 。二维空间中的任意点记为 $\mathbf{x} = (x_1, x_2)$ ，对应的频域点记为 $\mathbf{u} = (u_1, u_2)$ 。为了标记一致，下文将交替使用空域卷积和频域相乘来表示 Riesz 变换。由 Riesz 变换的定义可知，对于给定的任意二维信号 $f(\mathbf{x})$ ，它两个方向的 Riesz 变换 ($n = 2$) 为

$$\begin{aligned}
 f_{R_1}(\mathbf{x}) &= p.v. \int_{\mathbb{R}^2} -\frac{x'_1}{2\pi\|\mathbf{x}'\|^3} f(\mathbf{x} - \mathbf{x}') d\mathbf{x}' \\
 f_{R_2}(\mathbf{x}) &= p.v. \int_{\mathbb{R}^2} -\frac{x'_2}{2\pi\|\mathbf{x}'\|^3} f(\mathbf{x} - \mathbf{x}') d\mathbf{x}' \quad (3.7)
 \end{aligned}$$

其中 $\mathbf{x}' = (x'_1, x'_2)$, $p.v.$ 表示柯西主值积分, 由于 $n = 2$, 故 $c_n = \frac{1}{2\pi}$ 。根据 Riesz 变换的定义 (3.7) 可知, 二维信号空间的 Riesz 变换第一和第二变换核函数分别为

$$\begin{aligned} K_1(\mathbf{x}) &= -\frac{x_1}{2\pi\|\mathbf{x}\|^3} \\ K_2(\mathbf{x}) &= -\frac{x_2}{2\pi\|\mathbf{x}\|^3} \end{aligned} \quad (3.8)$$

频域空间对应的第一传递函数和第二传递函数分别为

$$\begin{aligned} \mathbb{K}_1(\mathbf{u}) &= \frac{u_1}{\|\mathbf{u}\|} \\ \mathbb{K}_2(\mathbf{u}) &= \frac{u_2}{\|\mathbf{u}\|} \end{aligned} \quad (3.9)$$

若取 $n = 1$, 则 Riesz 变换就蜕化为传统的 Hilbert 变换 (见式 (3.1)), 其空域和频域的变换核分别为

$$h(x) = -\frac{1}{\pi x} \Rightarrow H(u) = -j \cdot \frac{u}{\|u\|}.$$

对于给定的任意二维信号 $f(\mathbf{x})$, 它的单演信号 $f_M(\mathbf{x})$ 定义为原始信号与它的 Riesz 变换的线性组合,

$$f_M(\mathbf{x}) = f(\mathbf{x}) - (i, j)f_{\mathcal{R}}(\mathbf{x}) = f(\mathbf{x}) - i \cdot f_{\mathcal{R}_1}(\mathbf{x}) - j \cdot f_{\mathcal{R}_2}(\mathbf{x}) \quad (3.10)$$

其中 i, j 表示两个虚部单元, $\{i, j, 1\}$ 构成 \mathbb{R}^3 空间的一组正交基。就像 Riesz 变换之于 Hilbert 变换, 单演信号是个多维等向广义解析信号, 因而, 它具备等向广义解析信号的一切特性, 如旋转不变性、非负谱、两倍信号能量等。

单演信号的定义涉及复杂的 Clifford 几何, 实际应用中会造成诸多不便。因而, 图像的单演信号通常借助向量化函数或向量域 $\mathbf{f}_M : \mathbb{R}^2 \mapsto \mathbb{R}^3$ 来定义,

$$\mathbf{f}_M(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ -\mathbf{f}_{\mathcal{R}}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} f(\mathbf{x}) \\ -f_{\mathcal{R}_1}(\mathbf{x}) \\ -f_{\mathcal{R}_2}(\mathbf{x}) \end{bmatrix} \quad (3.11)$$

式 (3.11) 是单演信号的笛卡尔坐标表示, 另外一种常用的表示方法就是极坐标正交表示。与解析信号的极坐标分解类似, 单演信号的极坐标半径 (或局部幅度) 为

$$\|\mathbf{f}_M(\mathbf{x})\| = \sqrt{f^2(\mathbf{x}) + \|\mathbf{f}_{\mathcal{R}}(\mathbf{x})\|^2} = \sqrt{f^2(\mathbf{x}) + f_{\mathcal{R}_1}^2(\mathbf{x}) + f_{\mathcal{R}_2}^2(\mathbf{x})} \quad (3.12)$$

其中 $\|\mathbf{f}_{\mathcal{R}}(\mathbf{x})\|$ 表示共轭函数 $\mathbf{f}_{\mathcal{R}}(\mathbf{x})$ 的逐点 ℓ_2 范数。同样, 可推导出它的方位角 (Azimuth or Orientation) 为

$$\theta_M(\mathbf{x}) = \arctan \frac{f_{\mathcal{R}_2}(\mathbf{x})}{f_{\mathcal{R}_1}(\mathbf{x})}, \theta_M \in (-\frac{\pi}{2}, \frac{\pi}{2}] \quad (3.13)$$

顶角（或称俯仰角，该分量目前的文献中名称较多，如 Elevation Angle, Zenith Angle 等）为

$$\varphi_M(\mathbf{x}) = \arg[f(\mathbf{x}) + i\|\mathbf{f}_R(\mathbf{x})\|], \varphi_M \in [0, \pi] \quad (3.14)$$

从单演信号的向量化定义可知，局部幅度 $\|f_M(\mathbf{x})\|$ 是原始图像的空间包络，对它取平方即得到图像的局部能量，而方位角 $\theta_M(\mathbf{x})$ 则表示局部的方向性，是图像的局部几何信息描述，顶角 $\varphi_M(\mathbf{x})$ 则表示局部的相位，是图像的结构信息描述。局部方位角和局部相位也被称作是局部相位向量或是单演相位， \mathbf{r} ，

$$\mathbf{r}(\mathbf{x}) = \frac{\mathbf{f}_R(\mathbf{x})}{\|\mathbf{f}_R(\mathbf{x})\|} \arctan \frac{\|\mathbf{f}_R(\mathbf{x})\|}{f(\mathbf{x})} = \|\varphi_M(\mathbf{x})\| [sgn(\varphi_M(\mathbf{x})) \cos(\theta_M(\mathbf{x})), sgn(\varphi_M(\mathbf{x})) \sin(\theta_M(\mathbf{x}))]$$

其中 $sgn(\cdot)$ 表示符号函数。根据局部分量的定义可将单演信号极坐标变换表示为图 3.3，其中 $|f_M|, \theta, \varphi$ 代表局部能量、局部相位、局部方位。

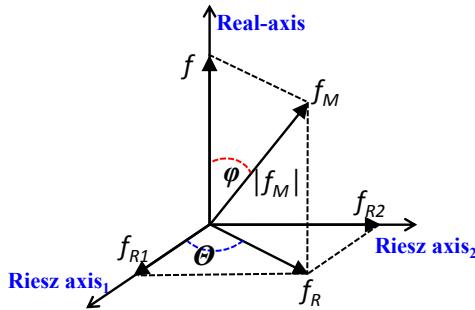


图 3.3 单演信号极坐标变换，图中 f, f_{R_1}, f_{R_2} 是原始笛卡尔坐标分量，分别表示实部和两个方向虚部， $|f_M|, \theta, \varphi$ 是极坐标分量。

通过与 Riesz 变换的性质类比，我们可以归纳单演信号具备如下特性：

- 两倍能量：如果忽略直流分量，则单演信号的能量是原始信号的两倍；

$$\int \|\mathbf{f}_M(\mathbf{x})\|^2 d\mathbf{x} = \int f^2(\mathbf{x}) + \|\mathbf{f}_R(\mathbf{x})\|^2 d\mathbf{x} = 2 \int f^2(\mathbf{x}) d\mathbf{x}$$

- 分解同一性：单演信号的极坐标分解满足分解同一性。单演信号的局部幅度（包络，能量）是等向分布的，即它不受信号方位角变化的影响，这类似于结构张量的各向同性，即关于能量和方位满足不变等价性。

- 冗余表示：单演信号的谱并非单边，它包含冗余，但完全可以使用非冗余的表示方式。

- 相位表示不唯一：若将信号执行 π 角度旋转后，极坐标变换得到的方位角相同而相位角相反，如图 3.4 例所示。

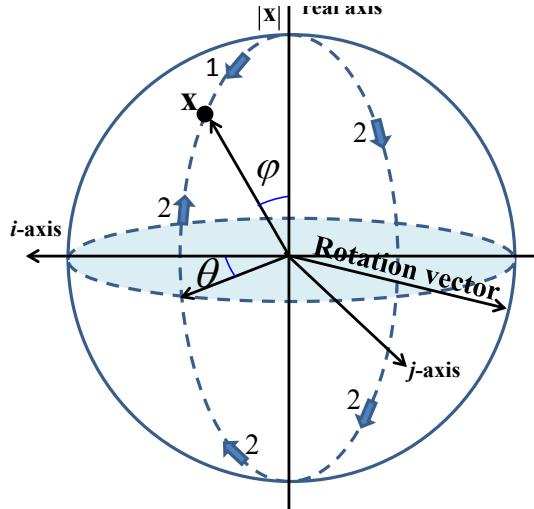


图 3.4 将实值 $|x|$ 先沿第二虚轴 (j -axis) 旋转 φ ，然后沿实轴旋转 θ_0 。旋转向量与实轴和向量 x 生成的平面正交，其长度取决于实轴和 x 的夹角。尽管旋转向量是唯一的，极坐标变换的相位表示却不唯一，有两种表示方式：(1) (θ_0, φ) ，(2) $(\theta_0 + \pi, -\varphi)$ 。[219]

实际应用中所遇到的信号通常都是有限长（如图像），它们的傅立叶频谱是周期无限长，因而需要借助带通滤波器对原始信号做空域的无限长延拓，然后再执行 Riesz 变换。对于任意给定的二维有限长信号，其经过空域延拓后的单演信号实质就是各分量通过滤波器之后的响应，

$$\mathbf{f}_M(\mathbf{x}) = (h_{bp} * f)(\mathbf{x}) - (i, j)(h_{bp} * \mathbf{f}_R)(\mathbf{x}) \quad (3.15)$$

其中 h_{bp} 表示带通滤波器的变换核。

为了保持 Riesz 变换核的奇对称性及零直流分量的特性，带通滤波器必需满足对称性条件。考虑到数学计算等原因，本文中使用 log-Gabor 滤波器来实现带通滤波。其它方法如柯西概率密度函数、泊松函数等也可以取得类似的结果。Log-Gabor 滤波器的频率响应为 [244]

$$G(\omega) = \exp \left\{ - \frac{[\log(\omega/\omega_0)]^2}{[\sigma/\omega_0]^2} \right\} \quad (3.16)$$

其中 ω_0 是中心频率， σ 是调节带宽的尺度因子。Peter Kovesi 指出，为了保持固定形状比率的滤波器组，通常将 $\sigma_r = \sigma/\omega_0$ 保持为常数 [245]。若以对数频率尺度观察，Log-Gabor 滤波器是类高斯的频率响应，如图 3.5 所示。因而，它可以截取信号高频部分的有用信息，实质就是我们期望的高通特性。以三尺度为例，Log-Gabor 波波器组的频率响应如图 3.6 所示，可以看出，它具有显著高通特性。

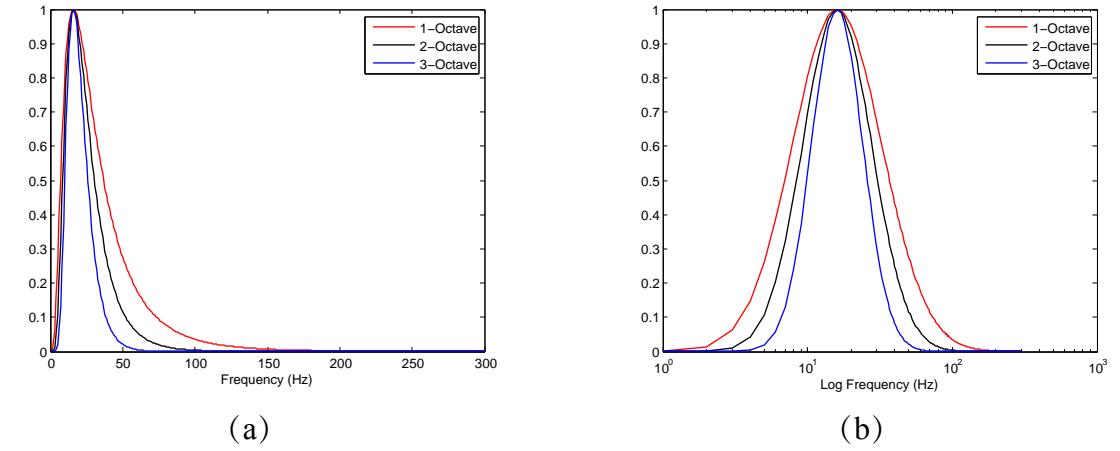


图 3.5 Log-Gabor 滤波器频率响应, (a) 线性频率坐标, (b) 对数频率坐标。

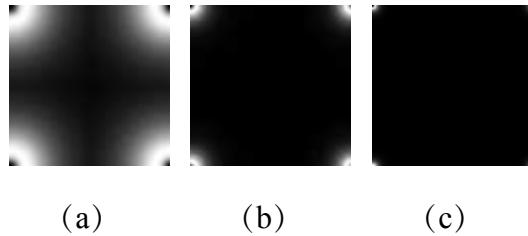


图 3.6 三尺度 Log-Gabor 滤波器组, (a) 第一尺度、(b) 第二尺度、(c) 第三尺度滤波器系数矩阵, 图中黑色区域表示零或接近零, 白色区域表示 1 或接近 1。

Log-Gabor 滤波器在零频点 $\omega = 0$ 是间断点, 因而它没有解析的空域表达式。其空域表示可通过逆傅立叶变换获取。Log-Gabor 滤波器与其姊妹函数 – Gabor 滤波器有着相似的低频特性, 但它的高频部分更陡峭。3 倍频带宽的 Log-Gabor 滤波器与单倍频 Gabor 滤波器有相同的空间宽度 [246]。这表明它能以一个紧致的空域滤波器获取更宽的谱信息。

为了充分挖掘信号特性, 通常使用信号的多尺度分析。Log-Gabor 滤波器的多尺度表示可通过调节尺度因子和中心频率来实现信号,

$$\omega_0 = \frac{1}{\lambda_{\min} \mu^{s-1}} \quad \sigma = \sigma_r \omega_0 \quad (3.17)$$

式中 λ_{\min} 表示最小波长, μ 表示波长的倍乘因子, s 是尺度索引, $\sigma_r = \sigma/\omega_0$ 是比值常数。将原始信号和 Riesz 变换结果分别与带通滤波器组进行卷积, 实现 Log-Gabor 滤波器引导的单演信号尺度空间。同样也可利用其它滤波器组构成单演信号尺度空间, 如 Cauchy 尺度空间 [247]、Poisson 尺度空间 [220]、一阶高斯函数尺度空间 [247]、Difference of Poisson 尺度空间等 [248]。由于 Log-Gabor 滤波器只有频域表达式, 不存在显式的时域形式, 时域卷积只能借助频域相乘然后将滤波器响应结果做逆傅立叶变换来实现, 单演信号尺度空间生成方法如下:

```

1 function [Component] = Multiscale_Monogenic_Signal(image,
2     min_wavel, sigma_f, nScale, mult)
3 % image: 输入的原始图像;
4 % min_wavel: Log-Gabor滤波器的最小波长;
5 % sigma_f: Log-Gabor滤波器sigma与中心频率的比值(保持为常数);
6 % nScale: Log-Gabor滤波器的尺度;
7 % mult: Log-Gabor滤波器的倍乘因子;
8 % Component: 输出的单演信号多尺度表示。
9
10 [nRow, nCol] = size(image);
11
12 [u1, u2] = meshgrid((1:nCol)-(fix(nCol/2)+1))/(nCol-mod(nCol
13 , 2)), ((1:nRow)-(fix(nRow/2)+1))/(nRow-mod(nRow, 2)));
14 u1 = ifftshift(u1);
15 u2 = ifftshift(u2);
16 radius = sqrt(u1.^2 + u2.^2);
17 radius(1,1) = 1;
18
19 H1 = 1i*u1./radius;
20 H2 = 1i*u2./radius;
21 IM = fft2(double(image));
22
23 Amp = zeros(nRow, nCol, nScale);
24 Ori = Amp; Pha = Amp;
25 Rel = Amp; Igi = Amp; Igj = Amp;
26
27 for s = 1:nScale
28 wavelength = min_wavel*mult^(s-1);
29 wo = 1.0/wavelength;
30 logGabor = exp((-log(radius/wo)).^2)/(2*log(sigmaOnf)^2));
31 logGabor(1,1) = 0;
32 H1s = H1.*logGabor;
33 H2s = H2.*logGabor;
34
35 f = real(ifft2(IM.*logGabor));
36 h1f = real(ifft2(IM.*H1s));
37 h2f = real(ifft2(IM.*H2s));
38
39 A = sqrt(f.^2 + h1f.^2 + h2f.^2);
40 q = sqrt(h1f.^2 + h2f.^2);
41
42 Amp(:,:,s) = A;
43 Ori(:,:,s) = atan(h2f./h1f);
44 Pha(:,:,s) = atan2(q,f);
45
46 Rel(:,:,s) = f;
47 Igi(:,:,s) = h1f;
48 Igj(:,:,s) = h2f;
49 end
50
51 Component.local_amplitude = Amp;
52 Component.local_orientation = Ori;
53 Component.local_phase = Pha;
54 Component.real_part = Rel;
55 Component.imaginary_x = Igi;
56 Component.imaginary_y = Igj;
57 end

```

为了直观地展示 Log-Gabor 滤波器的特性及单演信号表示, 图 3.7 给出了一组 SAR 切片图像及其单演信号多尺度表示, 图中使用三尺度 Log-Gabor 带通滤波对原始信号做无限延拓, 滤波器组如图 3.6 所示, 可以看出, Log-Gabor 滤波器允许四个对角线区域的频率分量(即高频部分)通过, 从而剔除掉中心部分的频率分量(低频部分)。图 3.7 (a)、(b) 分解展示了单演信号的笛卡尔坐标表示(实部、两个虚部即 Riesz 变换结果)和极坐标表示(局部幅度、局部相位、局部方位)。

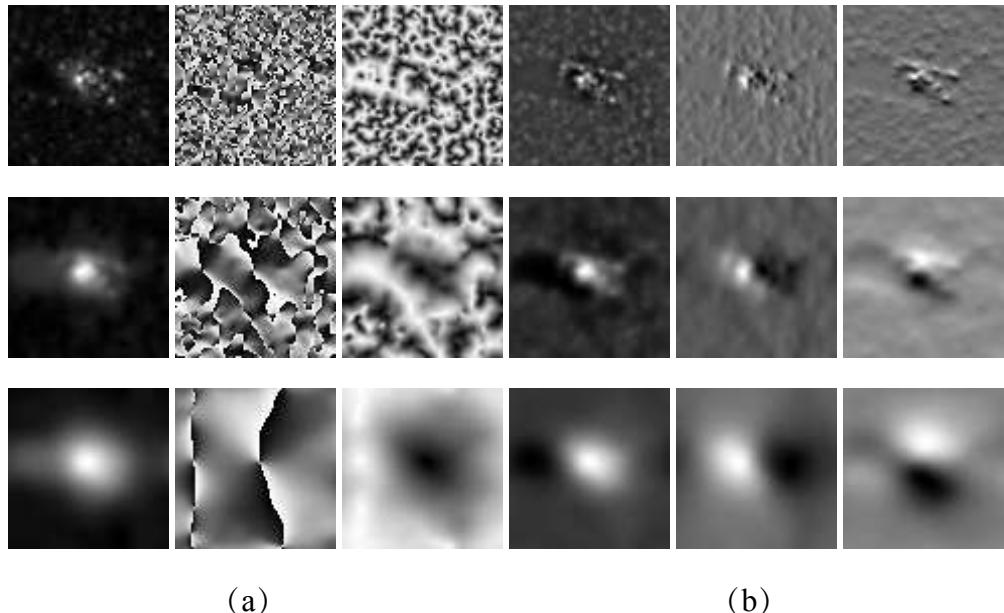


图 3.7 单演信号多尺度表示。图中是姿态角为 346° 的 BMP2_9563 在 17° 倾仰角的 SAR 切片图像的单演信号多尺度表示, 其中 (a) 是极坐标表示的系数矩阵, 即局部幅度、局部相位和局部方位, (b) 是笛卡尔坐标表示的系数矩阵, 即实部、 i 虚部和 j 虚部。

3.2 单演信号对 SAR 图像散射特性的变化分析

论文通篇的研究工作围绕着解决扩展工作条件的 SAR 图像自动目标识别问题展开, 第一章第 1.2.2 节对扩展工作条件涉及的诸多因素进行系统地梳理与建模, 概括为雷达成像参数、目标本身变化和周围环境变化三大类, 每类都列举了实例与图示。为了刻画 SAR 图像目标散射特性, 寻找有效的不变特征, 论文引入一种最新提出的信号分析与处理方法即单演信号, 并借助带通滤波器组构成单演信号尺度空间, 挖掘 SAR 图像本身蕴含的不同类型局部信息。为了观察单演信号多尺度表示对 SAR 图像目标散射现象刻画的效果, 我们借助 MSTAR 实测数据进行实例分析, 包括:

- 观察物理配置改变的 SAR 图像使用单演信号多尺度表示，能否实现配置不变描述；
- 观察姿态角改变的 SAR 图像使用单演信号多尺度表示，能否实现姿态角不变描述；
- 观察俯仰角改变的 SAR 图像使用单演信号多尺度表示，能否实现俯仰角不变描述；
- 观察变体后的 SAR 图像使用单演信号多尺度表示，能否实现不变特征描述；
- 观察噪声污染前后的 SAR 图像使用单演信号多尺度表示，能否找出噪声不变特征²。

3.2.1 物理配置变化分析

实际战场通常会存在同一目标的多种不同型号，所有不同型号的目标在军事层面理解应该归为同一类。为了观察单演信号表示对于目标配置型号变化的稳健性，我们选取一组 45° 姿态角的 BMP2 装甲输送车在 17° 俯仰角关于三种不同型号的成像数据，对它们做单演信号多尺度分析，生成的极坐标表示如图 3.8 所示。从图 3.8 (a) 可以看出，三种不同型号的 BMP2 其散射系数呈现出一定的差异性，特别是炮管头部的散射中心有明显的差别；图 3.8 (b)、(c)、(d) 展示的单演信号分量则呈现出较高的相似性，特别是局部相位和局部方位较为相似，这表明单演信号多尺度表示能有效应付目标配置的细小变化，也启发我们借助单演信号多尺度表示实现 SAR 图像物理配置不变描述。

3.2.2 姿态角变化分析

SAR 图像的相干成像原理决定了即使是微小的姿态角变化也会导致成像结果的显著差异。为了观察单演信号表示受目标姿态角变化的影响，论文选出姿态角为 13° 、 14° 、 15° 的 BMP2（型号为 SN_9563）在 17° 俯仰角的成像数据，对其分别进行单演信号多尺度表示，生成的极坐标分量如图 3.9 所示。从图 3.9 (a) 可以看出，即使是 1° 的姿态角变化，其成像结果还是有明显的变化，较为明显的是车身前侧的散射系数；然而它们的单演信号多尺度表示则呈现出较高的相似性，特别是局部相位和局部方位完全吻合，这也表明单演信号多尺度表示能够有效应对姿态角变化，因而，启发我们借助单演信号多尺度表示实现姿态角的不变描述。

²用于模拟传感器误差、成像算法对后向散射系数的估计误差等。

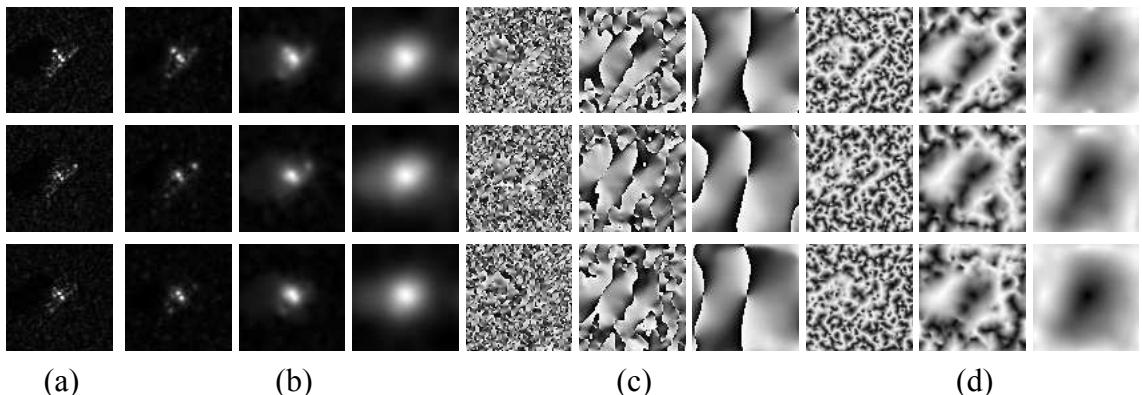


图 3.8 单演信号多尺度表示实现目标物理配置不变描述。图中分别是姿态角为 45° 的三种不同型号 BMP2 (SN_9563、SN_9566 和 SN_c21) 在俯仰角为 17° 的 SAR 图像单演信号表示, (a) 不同型号的 SAR 图像, (b) 局部幅度, (c) 局部相位, (d) 局部方位。

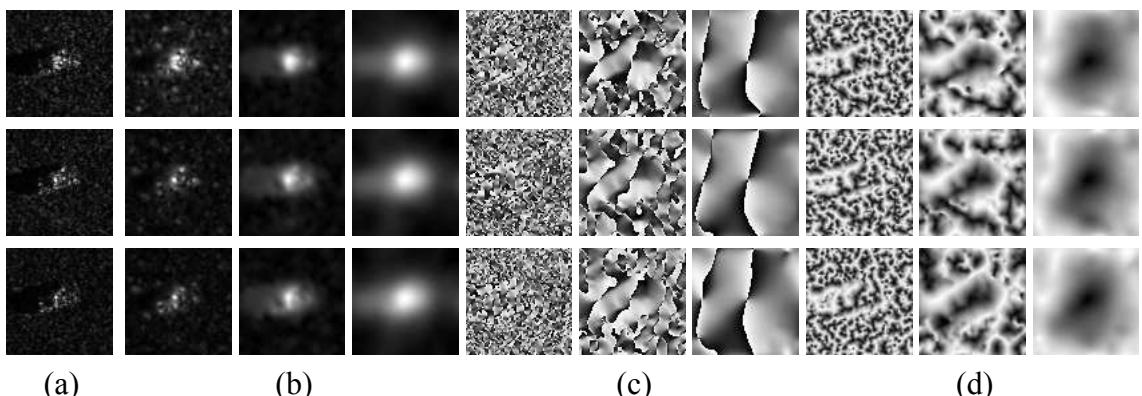


图 3.9 单演信号多尺度表示实现目标姿态角不变描述。图中分别是姿态角为 13° 、 14° 、 15° 的 BMP2 在俯仰角为 17° 的 SAR 图像单演信号表示, (a) 不同姿态角的 SAR 图像, (b) 局部幅度, (c) 局部相位, (d) 局部方位。

3.2.3 俯仰角变化分析

根据第一章分析可知, 同一目标在不同俯仰角下的雷达散射特性呈现出显著特性。为了观察单演信号受 SAR 图像俯仰角变化的影响, 论文选出 2S1 防空火炮在四种不同俯仰角的成像数据, 对其分别进行单演信号多尺度表示, 结果如图 3.10 所示。从图 3.10 (a) 可以看出, 同一目标在四种不同俯仰角下的 SAR 图像散射现象呈现出一定的差别, 特别是 45° 俯仰角的 SAR 图像, 其炮管方向的后向散射系数与其余三个角度的成像结果有明显的差异; 从图 3.10 (b)、(c)、(d) 可以看出, 它们的局部幅度、局部相位和局部方位则呈现出较高的相似性。这说明可以借助单演信号多尺度表示应对 SAR 图像的俯仰角变化, 实现俯仰角的不变描述。

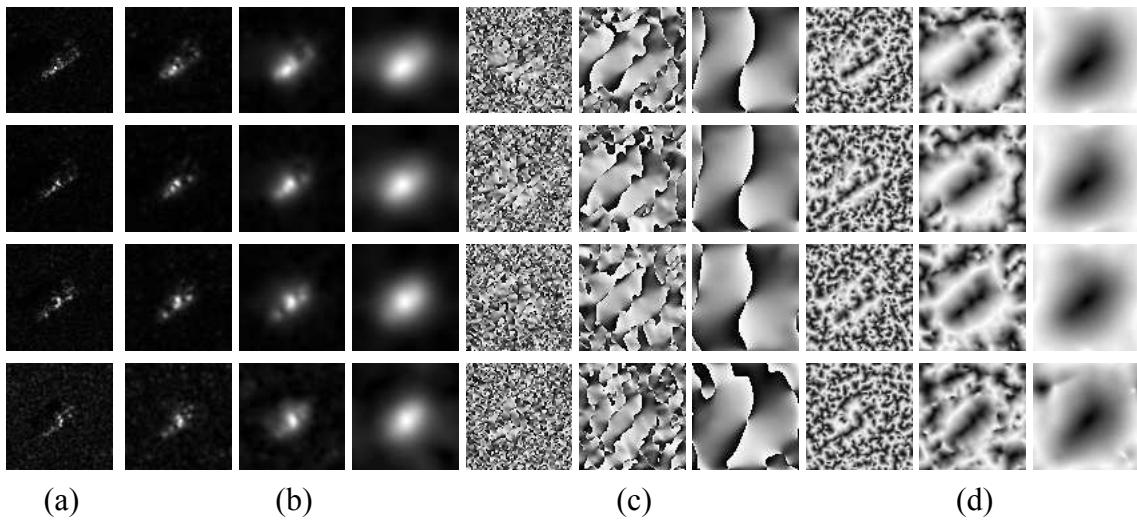


图 3.10 单演信号多尺度表示实现雷达俯仰角的不变描述。图中分别是姿态角为 30° 的 2S1 自行防空火炮在俯仰角为 15° 、 17° 、 30° 和 45° 时的 SAR 图像的单演信号表示，(a) 不同俯仰角获取的 SAR 图像，(b) 局部幅度，(c) 局部相位，(d) 局部方位。

3.2.4 连接体变体分析

根据前文分析可知，同一个目标在不同物理状态下其散射特性会呈现出明显的差异，比如行军途中的自行火炮和战斗状态的自行火炮其 SAR 图像会有显著地变化。为了观察单演信号对于同一目标不同状态的 SAR 的表示的稳健性，我们分别选取了一组 BRDM2 和 ZSU23/4 的标准状态 SAR 图像与连接体变体 SAR 图像，其中图像采集时的俯仰角和姿态角相同，仅仅是连接体的物理变化，如炮管的水平放置与竖起放置。对选取的 SAR 图像进行单演信号多尺度表示，结果分别如图 3.11 和图 3.12 所示。

通过比对结果可以看出，尽管两组目标存在连接体变化，但是它们的局部相位和局部方位呈现出较高的相似性，局部相位和局部方位对于连接体变体的稳健性优于局部幅度，可以想象，使用单演信号多尺度表示对 SAR 图像进行描述，能够从一定程度应对连接体变体所带来的不利影响，实现连接体变体的不变描述。

3.2.5 随机噪声污染分析

为了模拟 SAR 图像生成过程中的随机噪声，论文选取一组姿态角为 15° 的 T72 主战坦克 SAR 图像，对其施加等级为 10% 和 20% 的随机噪声污染，具体方法是：使用满足独立同分布的均匀分布的统计样本来替代原始图像中随机选取出的强度值，由于噪声是不可预知的，被污染的像素的位置是随机确定。对于原始图像以及污染后的图像使用 3 尺度单演信号表示，其极坐标分解如图 3.13 所示。

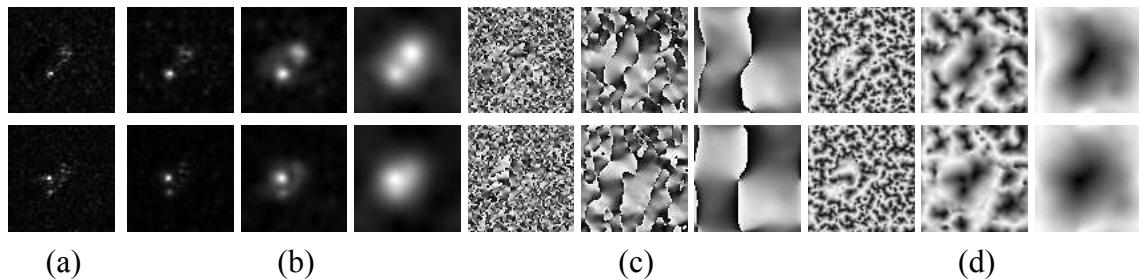


图 3.11 单演信号多尺度表示实现连接体变体的不变描述 (BRDM2)。图中是姿态角为 40° 的 BRDM2 装甲输送车在俯仰角为 30° SAR 图像及连接体变体图像的单演信号多尺度表示, (a) 原始 SAR 图像和连接体变体, (b) 局部幅度, (c) 局部相位, (d) 局部方位。

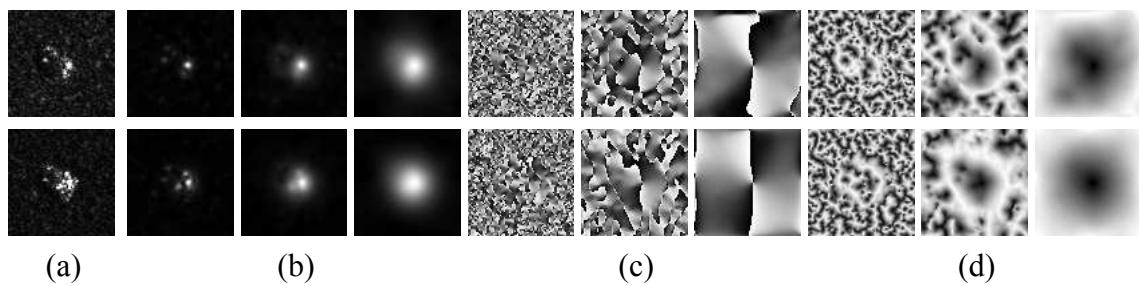


图 3.12 单演信号多尺度表示实现连接体变体的不变描述 (ZSU23/4)。图中是姿态角为 150° 的 ZSU23/4 在俯仰角为 45° 的 SAR 图像及连接体变体图像的单演信号多尺度表示, (a) 原始 SAR 图像和连接体变体, (b) 局部幅度, (c) 局部相位, (d) 局部方位。

从图 3.13 可以看出, 单演信号多尺度表示的局部幅度对噪声较为敏感, 局部相位和局部方位对噪声较为稳健; 小尺度分量受噪声影响较为明显, 大尺度分量

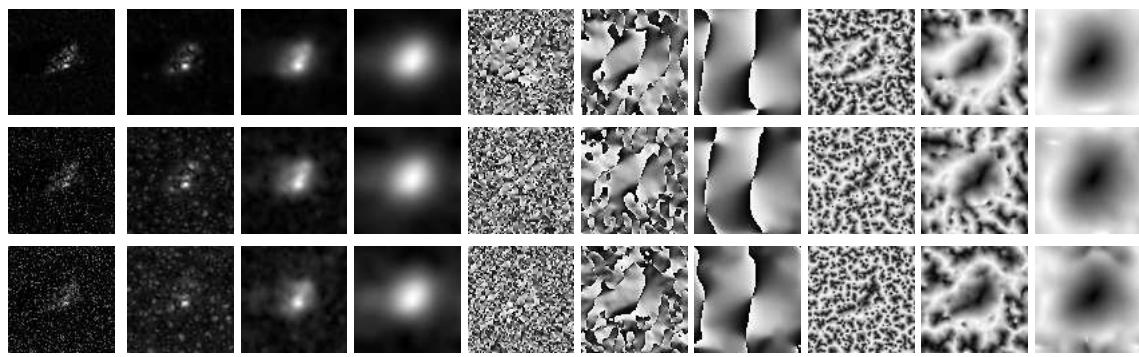


图 3.13 单演信号多尺度表示实现随机噪声污染的稳健描述。图中是方位角为 15° 的 T72 主战坦克 SAR 图像及噪声污染后的图像的单演信号多尺度表示, (a) 原始 SAR 图像和噪声污染图像, (b) 局部幅度, (c) 局部相位, (d) 局部方位。

能有效应对噪声的影响。因此，单演信号多尺度表示能够从一定程度提高噪声容限，实现噪声污染的不变描述，有利于目标识别的鲁棒性和准确性。

3.3 本章小结

本章主要介绍了一种新的信号分析与处理方法——单演信号。首先借助 Hilbert 变换的广义扩展引出 Riesz 变换，归纳总结 Riesz 变换的主要特性，包括反对称性、平移不变性、尺度不变性、可控性、内积守恒、正交特性等；通过 Riesz 变换定义高维信号空间的解析信号即单演信号。介绍了单演信号的构成原理及尺度空间构成，列举单演信号在信号分析与处理中的典型应用。之后，分别从目标配置变化、姿态角变化、俯仰角变化、连接体变体、随机噪声污染等方面以例证地方式介绍了单演信号多尺度表示用于 SAR 图像目标散射现象描述的诸多优点。

第四章 基于单演信号信息融合的 SAR 图像目标分类学习

根据前文的分析可知，实际战场的 SAR 图像目标识别需要考虑许多实际因素的影响，总结起来大致有 [27]：

- 目标本身的结构改变，如姿态角变化、结构和连接变体变化、外形轮廓的损坏、配置型号变化等；
- 周围环境产生的影响，如叶簇树木遮挡、气候条件变化、伪装隐蔽、欺骗等；
- 雷达工作参数的变化，如雷达俯仰角变化、斜视角变化、分辨率、极化方式、运动补偿等；

这些影响因素统称为扩展工作条件 (EOC)。列举出来的条目仅仅是理论因素，实际战场中的自动目标识别远比理论研究复杂许多。为了应对扩展工作条件，降低这些变化因素带来的不利影响，实现目标的不变特征描述从而提高目标识别可靠性，论文引入一种新的信号分析方法 — 单演信号 [219–221]，构造单演信号尺度空间，应对 SAR 图像俯仰角、姿态角、连接体变体和噪声污染等因素，第三章第 3.2 节对单演信号多尺度表示进行详细解释。

单演信号是解析信号在二维信号空间的广义扩展，若转换为极坐标表示，可分解成相互正交的三个局部分量：局部幅度、局部相位和局部方位，分别代表着原始信号的局部能量信息、局部几何信息和局部结构信息。能否将不同类型的信息整合起来用于目标识别，受此启发本章提出基于图像域信息融合的目标分类框架，将单演信号多尺度表示生成的不同类型分量，借助机器学习技巧进行图像域融合，整合不同类型信息用于目标识别，从而提高决策的稳健性与可靠性。本章的主要内容包括：1) 特征组融合，2) 决策级融合；3) 多特征多任务联合稀疏表示。

4.1 空域信息融合方案

为了应对扩展工作条件，论文期望能借助单演信号多尺度表示描述 SAR 图像，克服实际应用中遇到的困难。然而，单演信号仅仅是一种建立在 Riesz 变换基础上的调和分析理论成果，如何将数学工具与 SAR 图像目标识别的具体问题相结合，提出具体的实现方案是我们需要考虑的问题。一种直观的思想就是将单演信号生成的诸分量执行图像域融合。对此论文通过借鉴机器学习相关理论，提出基于图像域信息融合的分类思路，首先借助带通滤波器组构造单演信号尺度空间，然后将生成的不同尺度空间分量进行合并。根据这种思路框架的指导，设计三种具体实现方案，包括特征级融合、决策级融合和多任务多特征联合稀疏表示。

4.1.1 特征级融合 (Feature-level Fusion)

对于任意给定的二维有限长信号 (图像)，借助 Riesz 变换可构造解析信号表示，即单演信号。生成的单演信号可在笛卡尔坐标下表示为实部分量和两个虚部分量 (即 Riesz 变换结果)，也可以在极坐标下表示为局部幅度、局部相位和局部方位三个相互正交的分量。不同的分量代表着图像的不同类型信息，若将这些信息进行融合，将会提高 SAR 图像目标识别的可靠性和稳健性。对于如何合并使用单演信号诸分量，进行目标识别，一种直观的方法就是将不同分量矩阵拼接起来构造特征向量，然而直接进行组合生成的特征向量维数巨大，计算开销难以负担。单演信号分解得到的分量与原始信号维数相等，经过多尺度表示，其维数是原始信号的数倍，生成的向量将会是数万维，很难被直接使用。要解决维数灾难，传统的方法有很多，如数学变换 (PCA、LDA、NMF、LPP、SPP 等)、投影矩阵等，考虑到计算效率，去除冗余信息，论文采取均匀降采样的方案。对分量矩阵进行降维之后，将不同信息合并构造一种增强的单演信号特征向量。具体方法是：首先将单演信号多尺度分量作双线性内插的均匀降采样，采样后的结果分别重新排列成向量 (沿列排列或是沿行排列)，生成的向量做 ℓ_2 归一化，然后将所有向量组形成一个增强的单演信号特征向量。为了进一步降低增强的特征向量维数，提高运算效率，且不损失重要信息，我们对其做主成分分析，保留其 95% 的能量，最终生成的向量就是用于分类学习的单演信号特征。

假设使用 S 尺度 Log-Gabor 滤波对原始信号 $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^2$ 做无限延拓，所得到的单演信号尺度空间为 $\{\mathbf{f}_{M,1}, \mathbf{f}_{M,2}, \dots, \mathbf{f}_{M,S}\}$ ¹。对尺度空间中的所有信号沿极坐标进行分解 (笛卡尔坐标分解方法类似)，得到的局部幅度、局部相位、局部方位记为 $\{A_s, \varphi_s, \theta_s\}_{s=1}^S$ 。为了压缩数据维数并去除冗余信息，将所有分量矩阵作以 ρ 为因子的均匀降采样，采样结果得到的系数矩阵逐行 (或列) 扫描重新排成向量，对向量做 ℓ_2 范数能量归一化，然后将所有尺度空间的向量组合成一个增强的向量，即单演信号特征向量 [249]，

$$\chi_\rho = [vec_\rho(\mathbf{f}_{M,1})^T, vec_\rho(\mathbf{f}_{M,2})^T, \dots, vec_\rho(\mathbf{f}_{M,S})^T] \quad (4.1)$$

其中 $vec_\rho(\cdot)$ 是指将不同尺度的信号所分解的分量组合而成特征向量²。

以 $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,n_k}] \in \mathbb{R}^{m \times n_k}$ 表示第 k 类的训练样本，假设该类训练样本数目足够多。包含所有 K 个类别的训练集 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_K] \in \mathbb{R}^{m \times n}$ ($n = \sum_{k=1}^K n_k$)，其中每个类别的样本近似张成一个线性子空间，源自于该类的样本都位于该子空间中。为了刻画 SAR 图像的独有特性，充分利用其宽广的谱信息与

¹论文通篇将 Log-Gabor 滤波器的尺度设为 $S = 3$ 。

²为了避免相位缠绕，通常使用相位角和方向角的三角函数替代原始的角度，详见 [249] 所述。

空间局部特性，本文使用单演信号来对原始 SAR 图像进行描述，具体地推导过程见 3.1.3 节。然后根据单演信号特征向量构建稀疏表示系统，进行分类学习。首先根据训练样本的单演信号特征构造过完备字典 $\mathbf{D}_\chi = [\chi(\mathbf{X}_1), \chi(\mathbf{X}_2), \dots, \chi(\mathbf{X}_K)]$ ，其中 $\chi(\mathbf{X}_k) = [\chi(\mathbf{x}_{k,1}), \chi(\mathbf{x}_{k,2}), \dots, \chi(\mathbf{x}_{k,n_k})]$ ，然后使用字典原子对未知样本的单演信号特征向量 $\chi(\mathbf{y})$ 进行线性表示，

$$\chi(\mathbf{y}) = \chi(\mathbf{x}_{1,1})\alpha_{1,1} + \chi(\mathbf{x}_{1,2})\alpha_{1,2} + \dots + \chi(\mathbf{x}_{K,n_K})\alpha_{K,n_K} = \mathbf{D}_\chi\boldsymbol{\alpha} \quad (4.2)$$

其中 $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_K] \in \mathbb{R}^n$ 是未知的表示系数。理论上仅有那些与未知样本源于相同类别的训练样本参与线性表示，即表示系数中仅有很少数非零项，从而包含关于未知样本类别的重要信息，借助 ℓ_1 范数最小化的最优化策略即可实现，

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \quad s.t. \quad \|\chi(\mathbf{y}) - \mathbf{D}_\chi\boldsymbol{\alpha}\| \leq \epsilon \quad (4.3)$$

使用最优化算法即可求得最优表示系数，根据最优稀疏解 $\hat{\boldsymbol{\alpha}}$ 计算每个类别的重构误差，未知样本通常是源自于重构误差最小的类别，即

$$identity(\mathbf{y}) = \min_{k=1,\dots,K} \|\chi(\mathbf{y}) - \chi(\mathbf{X}_k)\hat{\alpha}_k\|_2 \quad (4.4)$$

若我们认为第 k 类样本的特征向量 $\chi(\mathbf{x}_{k,1}), \chi(\mathbf{x}_{k,2}), \dots, \chi(\mathbf{x}_{k,n_k})$ 构成一个微分流形，则 $\chi(\mathbf{X}_k)\hat{\alpha}_k$ 可以看作是未知样本 \mathbf{y} 在该流形上的投影， $\hat{\alpha}_k$ 是投影系数或权值，与未知样本距离最近的流形所在的类别即是未知样本的类别，该判决方法也就是最小小信号—流形距离判决准则。图 4.1 给出了特征级融合方法的简明流程图。

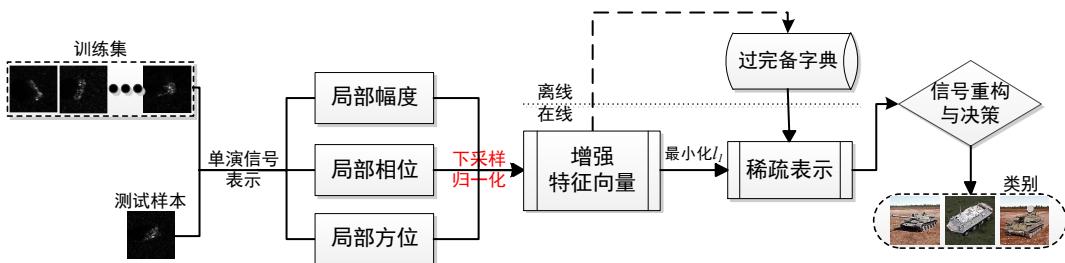


图 4.1 特征级融合简明流程。首先对训练样本做单演信号多尺度表示，将得到的分量组合起来生成增强的特征向量，再把特征向量组合构成过完备字典，并根据字典对未知样本的增强特征向量进行线性表示，借助稀疏约束求解线性模型，根据重构误差判定未知样本的类别。

4.1.2 决策级融合 (Decision-level Fusion)

SAR 图像反映的是目标表面的点散射分布，当雷达从不同的角度对目标进行探测时，部分结构通过会被遮挡。因而，即使是细微的姿态角变化也可能会造成图像与图像之间剧烈的特性变化。为了刻画 SAR 图像特性，前文使用单演信号对

SAR 图像进行描述，为了将不同类型信息进行融合，第 4.1.1 节中将单演信号诸分量组合起来构成一个增强的特征向量，这在数据融合中称之为特征级融合。尽管特征级的融合方法降低了数据维数以及冗余信息，从一定程度上提高了目标识别的准确性和稳健性，但它同时也会导致信息丢失。为了避免空域的简单组合，论文提出决策级的信息融合框架。首先，根据单演信号多尺度表示生成的分量构造特征向量，每一种分量分别构造各自的特征向量，方法与前文相同，根据这些特征向量分别建立线性表示模型，通过最稀疏约束求解，最后从统计地角度综合吸纳每个子模型的输出（即重构误差），得出联合决策。

以 S 尺度单演信号空间为例，对于第 k 类训练样本 $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,n_k}]$ 使用单演信号进行表示，对于单演信号所分解的各分量采取与 4.1.1 节相类似的处理方法，不同之处在于每种分量进行均匀降采样、 ℓ_2 范数归一化后重新排列生成各自的特征向量，通过合并同一分量在不同尺度空间下的表示来生成面向分量的特征向量，取代原来的增强特征向量，生成的分量特征为

$$\begin{cases} \chi_A = [A_1^\rho(:)^T, A_2^\rho(:)^T, \dots, A_S^\rho(:)^T] \\ \chi_\varphi = [\varphi_1^\rho(:)^T, \varphi_2^\rho(:)^T, \dots, \varphi_S^\rho(:)^T] \\ \chi_\theta = [\theta_1^\rho(:)^T, \theta_2^\rho(:)^T, \dots, \theta_S^\rho(:)^T] \end{cases} \quad (4.5)$$

根据特征向量 (4.5) 构造面向单演信号分量的线性表示模型。首先使用训练样本的特征构造过完备字典，

$$\begin{cases} \mathbf{D}_{\chi_A} = [\chi_A(\mathbf{x}_{1,1}), \chi_A(\mathbf{x}_{1,2}), \dots, \chi_A(\mathbf{x}_{K,n_K})] & \text{Local Amplitude} \\ \mathbf{D}_{\chi_\varphi} = [\chi_\varphi(\mathbf{x}_{1,1}), \chi_\varphi(\mathbf{x}_{1,2}), \dots, \chi_\varphi(\mathbf{x}_{K,n_K})] & \text{Local Phase} \\ \mathbf{D}_{\chi_\theta} = [\chi_\theta(\mathbf{x}_{1,1}), \chi_\theta(\mathbf{x}_{1,2}), \dots, \chi_\theta(\mathbf{x}_{K,n_K})] & \text{Local Orientation} \end{cases}$$

根据生成的字典对未知样本的特征向量 $\chi_A(\mathbf{y}), \chi_\varphi(\mathbf{y}), \chi_\theta(\mathbf{y})$ 进行线性表示，

$$\begin{aligned} \chi_A(\mathbf{y}) &= \chi_A(\mathbf{x}_{1,1})\alpha_{1,1}^A + \chi_A(\mathbf{x}_{1,2})\alpha_{1,2}^A + \dots + \chi_A(\mathbf{x}_{K,n_K})\alpha_{K,n_K}^A = \mathbf{D}_{\chi_A}\alpha^A \\ \chi_\varphi(\mathbf{y}) &= \chi_\varphi(\mathbf{x}_{1,1})\alpha_{1,1}^\varphi + \chi_\varphi(\mathbf{x}_{1,2})\alpha_{1,2}^\varphi + \dots + \chi_\varphi(\mathbf{x}_{K,n_K})\alpha_{K,n_K}^\varphi = \mathbf{D}_{\chi_\varphi}\alpha^\varphi \\ \chi_\theta(\mathbf{y}) &= \chi_\theta(\mathbf{x}_{1,1})\alpha_{1,1}^\theta + \chi_\theta(\mathbf{x}_{1,2})\alpha_{1,2}^\theta + \dots + \chi_\theta(\mathbf{x}_{K,n_K})\alpha_{K,n_K}^\theta = \mathbf{D}_{\chi_\theta}\alpha^\theta \end{aligned} \quad (4.6)$$

其中 $\alpha^A, \alpha^\varphi, \alpha^\theta$ 分别是三个线性表示模型的表示系数。同样期望 (4.6) 式回归模型的回归系数尽可能稀疏，即含有的非零项尽可能少

$$\begin{aligned} \min_{\alpha^A} \|\alpha^A\|_1 &\quad s.t. \quad \|\chi_A(\mathbf{y}) - \mathbf{D}_{\chi_A}\alpha^A\| \leq \epsilon \\ \min_{\alpha^\varphi} \|\alpha^\varphi\|_1 &\quad s.t. \quad \|\chi_\varphi(\mathbf{y}) - \mathbf{D}_{\chi_\varphi}\alpha^\varphi\| \leq \epsilon \\ \min_{\alpha^\theta} \|\alpha^\theta\|_1 &\quad s.t. \quad \|\chi_\theta(\mathbf{y}) - \mathbf{D}_{\chi_\theta}\alpha^\theta\| \leq \epsilon \end{aligned} \quad (4.7)$$

根据最优化算法求解回归系数，根据回归系统计算每类目标的重构误差 $\mathbf{e}^A, \mathbf{e}^\varphi, \mathbf{e}^\theta \in \mathbb{R}^K$ ，其中

$$\left\{ \begin{array}{l} e_k^A = \|\chi_A(\mathbf{y}) - \chi_A(\mathbf{X}_k)\hat{\alpha}_k^A\|_2 \\ e_k^\varphi = \|\chi_\varphi(\mathbf{y}) - \chi_\varphi(\mathbf{X}_k)\hat{\alpha}_k^\varphi\|_2 \\ e_k^\theta = \|\chi_\theta(\mathbf{y}) - \chi_\theta(\mathbf{X}_k)\hat{\alpha}_k^\theta\|_2 \end{array} \right\} \text{for } k = 1, \dots, K$$

稀疏表示分类的残差能量（即重构误差）反映的是未知信号到某个流形之间的距离，流形是由给定类别的所有样本构成，残差越小，决策越可靠，模型输出的置信度越高，反之亦然。因此，一种可行的方案是使用残差对分类系统的置信度进行量化评估，再根据统计理论对量化结果进行决策融合，如图 4.2 所示。

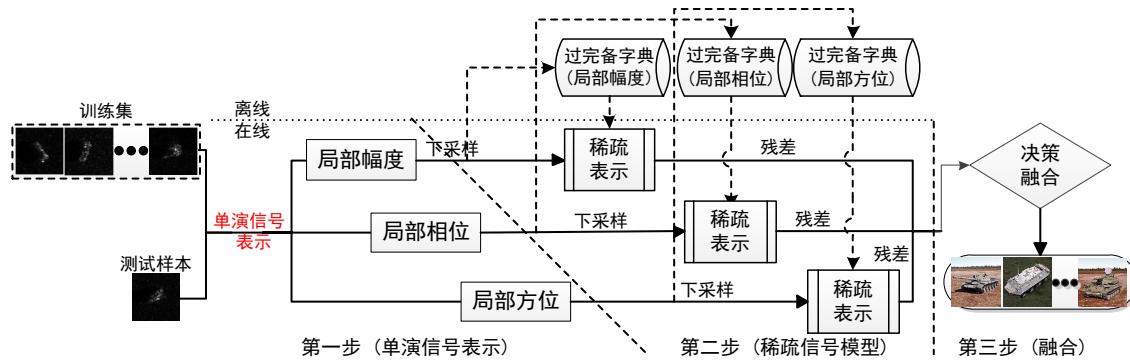


图 4.2 决策级融合简明流程。首先对训练样本做单演信号多尺度表示，将得到的每种分量在不同尺度的系数矩阵单独组合起来生成各自的特征向量，再把特征向量组合构成过完备字典，并根据字典构造未知样本的特征向量的稀疏线性模型，计算每个模型的重构误差，借助数学工具（如贝叶斯估计、DS 证据推理等）进行融合得出最终的决策。

围绕着图 4.2 所示的框架，本文提出两种实现方法：基于贝叶斯估计的决策融合和基于 DS 证据理论的决策融合。

4.1.2.1 贝叶斯估计

使用贝叶斯理论融合不同的决策输出，需要先定义每种假设命题的条件似然函数和先验概率，根据似然概率和先验概率计算后验概率，然后进行最终决策。首先对单演信号特征构建的稀疏表示模型的残差进行归一化处理 $e_k = \frac{e_k}{\sum_j e_j}, k = 1, \dots, K$ 。稀疏表示的重构误差是一种信号—流形距离，距离越小，未知样本源自于这个流形的可能性越大，反之亦然。因此，我们认为重构误差的大小与分类器的决策可靠性成反比，即重构误差越小，分类器的决策越可靠；反之，重构误差越大，分类器决策越不可信。

假设有 K 类目标分别记为 TG_1, TG_2, \dots, TG_K , 定义假设命题 $H_k : \mathbf{y} \in TG_k$ 的似然函数 (或是条件概率函数) 与该类样本的残差之间有如下的反比关系,

$$P(\mathbf{y}|H_k) = \frac{1/e_k}{\sum_j e_j} \quad (4.8)$$

根据贝叶斯理论可知, 假设 H_k 的后验概率为

$$P(H_k|\mathbf{y}) = \frac{P(\mathbf{y}|H_k)P(H_k)}{P(\mathbf{y})} \quad (4.9)$$

其中 $P(H_k)$ 表示未知样本属于第 k 类的先验概率。在没有任何先验知识的情况下, 我们通常认为未知样本属于每个类别是等可能概率, 即 $P(H_k) = \frac{1}{K}$ 。由于 $P(\mathbf{y})$ 是常数, 假设 H_k 的最大后验概率与它的最大似然函数等价 $\max P(H_k|\mathbf{y}) \propto \max P(\mathbf{y}|H_k)$ 。若给定相关的先验知识, 也可以量化地计算先验概率。

假设每个单演信号特征构成的稀疏表示模型的判决相对于其它模型的判决是条件统计独立的 [250], 即任何一个系统的决策并不影响其它系统的结果, 则最大联合后验概率就可以近似为

$$\begin{aligned} \max\{P(H_k|\chi_A(\mathbf{y}), \chi_\varphi(\mathbf{y}), \chi_\theta(\mathbf{y}))\} &\propto \max\{P(\chi_A(\mathbf{y}), \chi_\varphi(\mathbf{y}), \chi_\theta(\mathbf{y})|H_k)\} \\ &\propto \max\{P(\chi_A(\mathbf{y})|H_k)P(\chi_\varphi(\mathbf{y})|H_k)P(\chi_\theta(\mathbf{y})|H_k)\} \end{aligned} \quad (4.10)$$

根据最大后验概率准则, 我们就可以得到最终的判决,

$$\max_k (P(\chi_A(\mathbf{y})|H_k)P(\chi_\varphi(\mathbf{y})|H_k)P(\chi_\theta(\mathbf{y})|H_k)). \quad (4.11)$$

此外, Josef Kittler 等学者系统地归纳了一套基于贝叶斯估计的决策融合框架 [250], 即分类器合并方法, 包括: 加法准则、乘法准则、最大化准则、最小化准则、中值准则、多数投票准则等, 这些方法的原理与本文方法类似, 不同之处在于具体地合并过程各异。

为了展示决策级融合的优点, 图 4.3 给出了四类目标 (BMP2、T72、BTR60、T62) 识别的实例, 首列是三种单演信号分量构成的稀疏表示模型以及它们的子判决, 之后是传统的加法准则和基于贝叶斯估计的决策融合。未知样本源自于 BMP2, 局部幅度和局部方位特征所构成的稀疏表示模型将未知样本错误地识别为 BTR60 和 T72, 决策级融合方法能充分不同类别的信息, 规避各分量特征缺陷, 得出正确判决。

4.1.2.2 DS 证据推理

尽管贝叶斯融合从一定程度上提高了识别精度, 但受限于本身的一些理论缺陷, 如无法对不确定状态的不精确性进行描述 [209–211], 决策的可靠性还有一定的提升空间。关于贝叶斯估计与 DS 证据推理的优劣前文中已经进行了详细地分

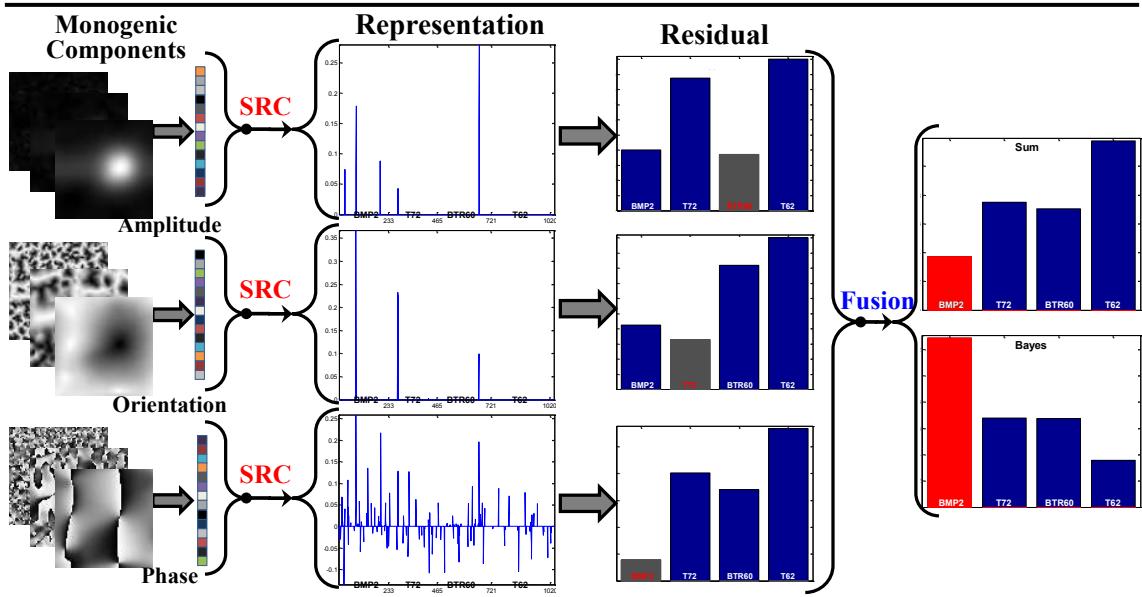


图 4.3 贝叶斯决策融合示例，图中依次是单演信号多尺度表示、稀疏表示模型的系数向量与子判决、贝叶斯估计决策融合，SUM 指将子判决的输出误差直接相加，寻找累加误差最小；Bayes 是指使用式 (4.11) 的决策结果，图中红柱表示融合决策。

析，此处不再赘述。为了改进基于贝叶斯估计的融合决策的不足，论文提出基于 DS 证据推理的决策级融合，提高判决的可靠性与合理性。

稀疏表示分类实质是多元线性回归模型的稀疏求解，其中训练样本扮演着预测量 (Regressor or Predictor) 的角色，而未知样本则是回归响应 (Response)，表示向量是回归系数。线性系统的重构误差反映未知信号到给定类别样本构成的流形之间的距离，所以一个合理的设想是根据重构误差来定义基本概率赋值函数³。

若将 K 个目标类别记为 $\{\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_K\}$ ，则辨识框架为

$$\{\mathcal{L}_1\}, \dots, \{\mathcal{L}_K\}, \{\mathcal{L}_1 \cup \mathcal{L}_2\}, \dots, \{\mathcal{L}_1 \cup \mathcal{L}_2 \cup \dots \cup \mathcal{L}_K\}$$

，其中 $\{\mathcal{L}_1 \cup \mathcal{L}_2\}$ 表示两个类别 \mathcal{L}_1 和 \mathcal{L}_2 的组合类别，即不确定状态。下面讨论如何使用残差定义基本概率赋值。稀疏表示分类使用最小残差判定未知样本的目标类别。残差越小，判决越可靠。对于给定的稀疏表示模型及其残差向量 $\mathbf{e} = [e_1, e_2, \dots, e_K] \in \mathbb{R}^K$ ，若某个类别的残差大于某预设阈值 $e_i > \tau_0$ 则我们就认为假设 $\mathbf{y} \in \{\mathcal{L}_i\}$ 不可能成立，其中 τ_0 是个预设的阈值，它是判断一个候选命题是否成立的临界值。此时命题 $\mathbf{y} \in \{\mathcal{L}_i\}$ 的基本概率赋值就认为是 0， $m(\{\mathcal{L}_i\}) = 0$ 。反之，如果某个类别的残差小于该阈值 $e_i < \tau_0$ ，则我们有理由认为假设 $\mathbf{y} \in \{\mathcal{L}_i\}$ 的成立存在着一定的可能性，因此它的基本概率赋值就不为 0，该命题就是一个焦元命

³关于 Dempster-Shafer 证据理论及基本概率赋值的相关基础知识参见论文第二章第 2.4.2.2 节。

题。归纳起来，我们仅仅将非零的基本概率赋值分配给残差小于预设阈值 τ_0 的那些假设命题。接下来讨论 $e_i < \tau_0$ 的情形。

对于任意特征构成的分类系统，第 i 类目标的残差 e_i 值越小，未知样本属于第 i 类目标的可能性越大，我们认为它的基本概率赋值就越大，反之 e_i 值越大（这里的前提条件是 $e_i < \tau_0$ ），未知样本属于第 i 类目标的可能性就越小，我们认为该假设的基本概率赋值越小。依据这个思路，我们通过构造数学模型描述残差与基本概率赋值之间的关系 [167]，

$$m(\{\mathcal{L}_i\}) = \begin{cases} W_i \cdot (1 - e_i) & e_i \leq \tau_1 \\ W_i \cdot g(e_i) & \tau_1 < e_i < \tau_0 \\ 0 & e_i \geq \tau_0 \end{cases} \quad (4.12)$$

其中 W_i 是个权值系数； τ_1 和 τ_0 是两个预设阈值，分别代表两种临界情况，绝对发生与绝对不发生； $g(\cdot)$ 是引入的辅助函数。函数 (4.12) 中设计权值 W_i 是为了引入不确定状态 (uncertainty)，即组合类别，避免出现某个命题的基本概率赋值为 1 的情形。

论文设计的框架模型可解释为：若某类的残差大于预设的阈值 τ_1 ，我们认为假设 $y \in \{\mathcal{L}_i\}$ 不可能成立，若该类的残差小于预设的阈值 τ_0 ，我们认为该假设成立的可能性很大，它的基本概率赋值与对应的残差成反比例关系，当残差介于两个阈值之间时，它们之间满足某种以 $g(\cdot)$ 描述的函数关系。根据模型 (4.12) 我们可以对每组简单假设进行量化，剩余的问题就是如何对复合假设（组合类别，即不确定状态）进行量化，这也是 DS 证据理论应用的关键。结合实际，为了兼顾计算效率和识别性能，本文将非零基本概率赋值仅仅赋予残差小于阈值 τ_1 的类别（即焦元），其余的全部归入不确定状态，即只考虑复合假设 $\Theta = \mathcal{L}_1 \cup \mathcal{L}_2 \cup \dots \cup \mathcal{L}_K$ ，将其视为不确定状态，基本概率赋值为 $m(\{\mathcal{L}_1 \cup \mathcal{L}_2 \cup \dots \cup \mathcal{L}_K\}) = 1 - \sum_{i=1}^K m(\{\mathcal{L}_i\})$ 。最终的证据描述框架中，焦元只包括 $\{\mathcal{L}_1\}, \{\mathcal{L}_2\}, \dots, \{\mathcal{L}_K\}, \{\mathcal{L}_1 \cup \mathcal{L}_2 \cup \dots \cup \mathcal{L}_K\}$ ⁴，总共只计算 $K + 1$ 个假设命题的证据质量，其余假设的证据均为零。为了满足基本概率赋值的定义，即所有焦元的证据之和是单位 1，本文将权值系数 W_i 设定为

$$W_i = \frac{1/e_i}{\sum 1/\mathbf{e}'}$$

其中 \mathbf{e}' 表示残差向量中小于阈值 τ_1 的项，从而整个证据描述的框架设计完毕。剩余的任务就是如何确定辅助函数 $g(\cdot)$ 。

⁴对于任意 K 个类别，总会存在 2^K 种假设（含一个空集 \emptyset ），其中 K 个类别对应的简单假设与贝叶斯估计相同，剩余的复合假设就是 DS 证据理论框架所独创，然而如何描述这些复合假设的证据质量，目前仍是个公开的难题。另一文方面，如何考虑所有的复合假设，则计算复杂度将以指数倍增长。因此，本文只讨论 Θ 这一种复合假设，更多的探索将会在后续研究中展开。

在本文提出的证据描述框架中，辅助函数 $g(\cdot)$ 是残差 e 是单调减函数，它的值域区间在 0 到 1 之间， $g(\cdot) : e \mapsto (0, 1)$ 。显然，满足该条件的函数有成千上万之多，且很难给出孰优孰劣的结论。受到 Franz Rottensteiner 等学者的研究工作 [251] 的启发，本文设计如下辅助函数，

$$g(e_i) = P_1 + (P_2 - P_1)[3 \times \left(\frac{\tau_0 - e_i}{\tau_0 - \tau_1}\right)^2 - 2 \times \left(\frac{\tau_0 - e_i}{\tau_0 - \tau_1}\right)^3] \quad (4.13)$$

其中 P_1, P_2 是一组预设的临界基本概率赋值， $P_1 < P_2$ 。函数 (4.13) 的曲线见图 4.4。可以看出，当残差大于阈值 τ_0 时，基本概率赋值固定为 P_1 ，当残差小于阈值 τ_1 时，基本概率赋值与残差成反比例关系，当残差位于 τ_1 和 τ_0 之间时， $\tau_1 \leq e \leq \tau_0$ ，它们之间的关系满足 $g(\cdot)$ 的函数关系。

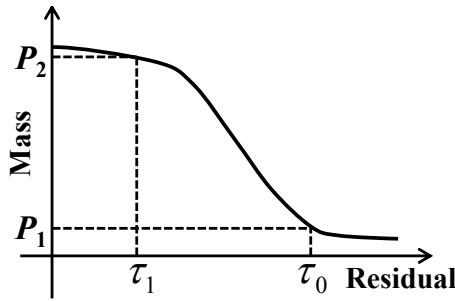


图 4.4 基本概率赋值函数曲线。若残差大于阈值 τ_0 ，基本概率赋值为 P_1 ；若残差小于阈值 τ_1 ，基本概率赋值与残差近似成反比例关系；若残差 $\tau_1 \leq e \leq \tau_0$ ，它们之间满足式 (4.13)。

至此，整个 DS 证据理论的应用框架搭建完毕。若单演信号分量构成的线性表示模型求解后的残差向量为 $\mathbf{e}^A, \mathbf{e}^\varphi, \mathbf{e}^\theta$ ，对其做归一化处理，然后根据 (4.12) 所定义的数学模型分别对 $\mathbf{e}^A, \mathbf{e}^\varphi, \mathbf{e}^\theta$ 进行量化，计算每个类别的基本概率赋值，然后使用直和规则将线性表示模型的证据描述进行合并，根据合并证据进行判决。证据合并流程如表 4.1 所示，根据 Dempster 合并规则可以算出最终的证据为

$$\left\{ \begin{array}{l} m(\mathcal{L}_1) = \frac{m^A(\mathcal{L}_1)m^\varphi(\mathcal{L}_1)m^\theta(\mathcal{L}_1) + m^A(\mathcal{L}_1)m^\varphi(\mathcal{L}_1)m^\theta(\Theta) + m^A(\mathcal{L}_1)m^\varphi(\Theta)m^\theta(\mathcal{L}_1) + m^A(\Theta)m^\varphi(\mathcal{L}_1)m^\theta(\mathcal{L}_1)}{1-K} \\ m(\mathcal{L}_2) = \frac{m^A(\mathcal{L}_2)m^\varphi(\mathcal{L}_2)m^\theta(\mathcal{L}_2) + m^A(\mathcal{L}_2)m^\varphi(\mathcal{L}_2)m^\theta(\Theta) + m^A(\mathcal{L}_2)m^\varphi(\Theta)m^\theta(\mathcal{L}_2) + m^A(\Theta)m^\varphi(\mathcal{L}_2)m^\theta(\mathcal{L}_2)}{1-K} \\ \vdots \\ m(\mathcal{L}_K) = \frac{m^A(\mathcal{L}_K)m^\varphi(\mathcal{L}_K)m^\theta(\mathcal{L}_K) + m^A(\mathcal{L}_K)m^\varphi(\mathcal{L}_K)m^\theta(\Theta) + m^A(\mathcal{L}_K)m^\varphi(\Theta)m^\theta(\mathcal{L}_K) + m^A(\Theta)m^\varphi(\mathcal{L}_K)m^\theta(\mathcal{L}_K)}{1-K} \end{array} \right.$$

其中 $K = \sum_{\mathcal{L}_i \cap \mathcal{L}_j = \emptyset} m(\mathcal{L}_i)m(\mathcal{L}_j)$ 是冲突证据。论文设计的证据描述框架中，简单假设的不确定区间为 $Bel(\{\mathcal{L}_i\}), Pls(\{\mathcal{L}_i\})$ ，都是相等的，因而前文提到的几种判决准则将会取得相同的排序结果，即相同的判决结论。

为了直观地说明本文算法，我们给出四组四类目标识别 (BMP2、T72、BTR60 和 T62，分别记为 $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \mathcal{L}_4$) 的例子，关于数据集的详细描述参见论文第二

表 4.1 决策级融合的证据合并框架, $\Theta = \mathcal{L}_1 \cup \mathcal{L}_2 \cup \dots \cup \mathcal{L}_K$ 。

假设命题	\mathcal{L}_1	\mathcal{L}_2	\dots	\mathcal{L}_K	Θ
幅度特征	$m^A(\mathcal{L}_1)$	$m^A(\mathcal{L}_2)$	\dots	$m^A(\mathcal{L}_K)$	$m^A(\Theta)$
相位特征	$m^\varphi(\mathcal{L}_1)$	$m^\varphi(\mathcal{L}_2)$	\dots	$m^\varphi(\mathcal{L}_K)$	$m^\varphi(\Theta)$
方位特征	$m^\theta(\mathcal{L}_1)$	$m^\theta(\mathcal{L}_2)$	\dots	$m^\theta(\mathcal{L}_K)$	$m^\theta(\Theta)$
融合结果	$m(\mathcal{L}_1)$	$m(\mathcal{L}_2)$	\dots	$m(\mathcal{L}_K)$	—

章。实验如表 4.2 所示, 列出了贝叶斯估计与 DS 证据推理两种融合方法的处理流程, 其中贝叶斯估计的条件似然概率定义为类别残差的反比例函数,

$$P(\mathcal{L}_i) = \frac{1/e_i}{\sum_j 1/e_j}$$

DS 证据理论的证据质量描述参考公式 (4.12)。

表 4.2 贝叶斯估计与 DS 证据理论决策融合例证。

规则	<i>Bayesian</i>				<i>Dempster-Shafer</i>				
	e_1	e_2	e_3	e_4	$m(\mathcal{L}_1)$	$m(\mathcal{L}_2)$	$m(\mathcal{L}_3)$	$m(\mathcal{L}_4)$	$m(\Theta)$
命题假设	0.0920	0.3716	0.1659	0.3705	0.5820	0	0.1458	0	0.2722
幅度特征	0.0404	0.3553	0.2489	0.3553	0.9596	0	0	0	0.0404
相位特征	0.1022	0.3618	0.1188	0.4171	0.4679	0	0.3607	0	0.1714
方位特征	0.5375	0.1158	0.2363	0.1103	0.5461	0.0019	0.0090	0.0019	—
合并结果	0.2638	0.2501	0.2024	0.2837	0.3898	0.0126	0.0102	0.0226	—
幅度特征	0.1877	0.3475	0.3777	0.0872	0.0771	0	0	0.6244	0.2985
相位特征	0.3383	0.1226	0.3289	0.2101	0	0.4768	0	0.0443	0.4789
方位特征	0.2626	0.1172	0.4011	0.2190	0	0.5149	0	0.0324	0.4528
合并结果	0.2183	0.2931	0.1554	0.3333	0.0814	0.2761	0.0647	0.2342	—
幅度特征	0.2998	0.2340	0.2830	0.1832	0	0	0	0.2741	0.7259
相位特征	0.2921	0.1373	0.3170	0.2537	0	0.6461	0	0	0.3539
方位特征	0.2008	0.2285	0.4137	0.1570	0.0517	0.0223	0	0.1958	0.7302
合并结果	0.2253	0.2978	0.1762	0.3007	0.2009	0.5463	0.1876	0.3277	—

第一组实验的未知样本源自于 \mathcal{L}_1 。贝叶斯估计赋予命题 $y \in \mathcal{L}_1$ 的似然概率是 0.5375, DS 证据理论则赋予 $y \in \mathcal{L}_1$ 的置信度函数是 0.5461。两种方法都得出了

正确的决策，但它们的置信度水平存在着明显的差异⁵，贝叶斯估计的概率差异度为 0.5604，DS 证据理论的证据差异度为 0.9835，贝叶斯估计的置信水平逊于 DS 证据理论。第二组实验的未知样本也来自于 \mathcal{L}_1 。两种融合方法分别给出截然不同的结论，贝叶斯估计将未知样本识别为 \mathcal{L}_4 ，而 DS 证据理论则给出了正确的决策。第三组和第四组实验的未知样本都源自于 \mathcal{L}_2 。贝叶斯估计将其识别为 \mathcal{L}_4 ，而 DS 证据理论则正确地将其识别为 \mathcal{L}_2 。这是因为：相比较于贝叶斯估计，DS 证据推理能够同时对不确定性状态以及它的不精确性度量进行建模，从而提高决策的可靠性。

4.1.3 多特征多任务联合稀疏表示

多任务学习是近年来提出一种学习方法 [252]，若子任务之间相关，联合任务学习会比单独学习效果好，特别是当每个子任务的数据相当少导致单独学习不成功时，多任务联合学习能够将每个子任务学习的好的学习结果传递到另一个新的任务上，因而展现出显著的优越性。受多任务学习启发，论文提出一种多特征多任务联合稀疏表示的目标识别方法。首先对原始信号进行单演信号多尺度表示，并沿极坐标对不同尺度空间的信号进行分解，将获取的分量按照均匀降采样、归一化、组合等步骤构造特征向量，然后根据特征向量分别构建线性表示系统，与前文方法不同的是，我们并不是为每个线性系统单独进行稀疏约束，而是通过施加一个特殊的联合稀疏约束，迫使每个线性系统能够生成具有相似稀疏项的表示系数，从而在不同的线性系统之间实现信息共享，提高目标识别率。多任务联合稀疏表示的工作机理如图 4.5 所示。

4.1.3.1 算法描述

首先构造单演信号尺度空间，将不同尺度空间转换成极坐标表示（笛卡尔坐标表示类似），得到的局部幅度、局部相位、局部方位按照第 4.1.2 节中的公式 (4.5) 构造特征向量 χ_A , χ_φ 和 χ_θ 。单演信号各分量分别代表着原始信号的不同类型信息，局部幅度代表信号的局部能量信息，单演相位（局部相位和局部方位）则蕴含着信号的局部结构信息和局部几何信息。为了避免特征组合过程的信息丢失问题，我们为每种分量设计各自的特征向量，然后构造面向特征向量的欠定线性系统。将未知样本的识别问题转化成为三个面向不同特征向量的多任务学习。对于每个学习任务，训练样本的特征向量扮演着预测者的角色，未知样本的特征向量则是系统对应的响应，未知的系数向量就是回归系数。该问题就是多任务学习的典型应用。研究表明若每个待学习的任务都能与其它任务共享某些信息，通过

⁵论文的置信度水平定义为不同命题之间的概率差异，本文定义差异度为 $\frac{\|p_1-p_2\|}{p_1}$ ，其中 p_1, p_2 分别是置信度最大的前两个命题的证据或概率。

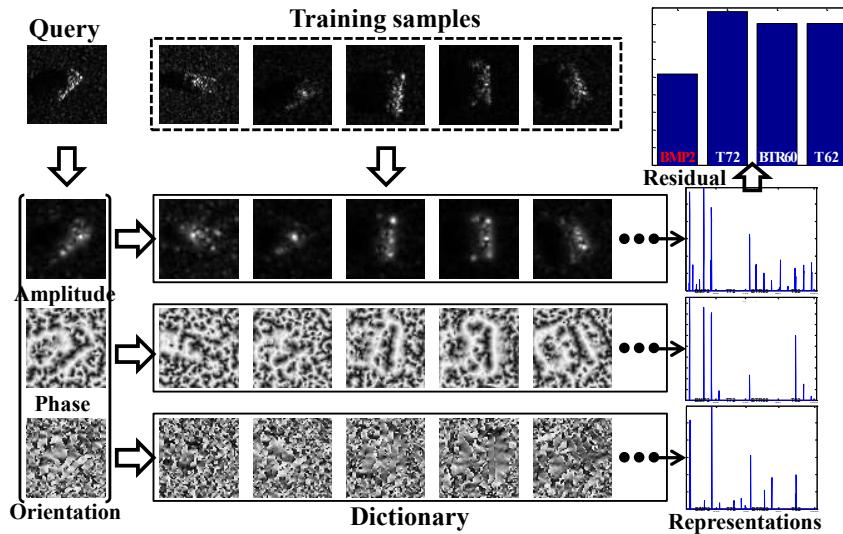


图 4.5 多任务联合稀疏表示的目标识别方法示例。对于任意的未知样本，对其做单演信号多尺度表示并将不同尺度空间的信号沿极坐标分解，从而得到三组面向各分量的特征向量。然后将得到的特征向量分别表示成为对应的训练样本的特征向量的线性组合，通过混合范数约束系数矩阵的可行集，使得各线性系统具有类似的稀疏项，从而实现不同任务之间的信息共享。最后根据整体重构误差判决目标类型。

考虑各任务之间的相互关系就可以极大地提高最终的决策可靠性 [253]。就目标识别问题而言，三个分量特征源自于一个样本，它们构建的线性表示模型的判决结果理论上应该是一致的，这就是多任务联合稀疏表示的根本出发点。

假设有足够的训练样本， $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,1}, \dots, \mathbf{x}_{k,n_k}]$, $k = 1, \dots, K$ ，我们根据它们的单深信号特征向量分别构造过完备冗余字典 [254, 255]，

$$\begin{cases} \mathbf{D}_{\chi_A} = [\chi_A(\mathbf{x}_{1,1}), \chi_A(\mathbf{x}_{1,2}), \dots, \chi_A(\mathbf{x}_{K,n_K})] \\ \mathbf{D}_{\chi_\varphi} = [\chi_\varphi(\mathbf{x}_{1,1}), \chi_\varphi(\mathbf{x}_{1,2}), \dots, \chi_\varphi(\mathbf{x}_{K,n_K})] \\ \mathbf{D}_{\chi_\theta} = [\chi_\theta(\mathbf{x}_{1,1}), \chi_\theta(\mathbf{x}_{1,2}), \dots, \chi_\theta(\mathbf{x}_{K,n_K})] \end{cases} \quad (4.14)$$

对于任意给定的未知样本 \mathbf{y} ，它的特征向量可线性表示为

$$\begin{cases} \chi_A(\mathbf{y}) = \mathbf{D}_A \alpha_{\chi_A} + \epsilon_A \\ \chi_\varphi(\mathbf{y}) = \mathbf{D}_\varphi \alpha_{\chi_\varphi} + \epsilon_\varphi \\ \chi_\theta(\mathbf{y}) = \mathbf{D}_\theta \alpha_{\chi_\theta} + \epsilon_\theta \end{cases} \quad (4.15)$$

其中 $\epsilon_A, \epsilon_\varphi, \epsilon_\theta$ 是每个系统的误差容限。式 (4.15) 是待求解的欠定线性系统，即三个独立地学习任务，传统的思想是分别求解单个线性系统，然后将每个系统的输出进行融合，如第 4.1.2 节的思路。这种思路易于理解和执行，但它不能有效利用各个学习任务的内在相关性，实现任务间的信息共享。三个线性表示模型是由同一个样本的单演信号分量特征构成，它们的判决输出理论上应该相一致。为了使

三个学习任务共享未知样本的类别属性信息，论文提出多特征多任务联合稀疏表示。首先，将线性表示系统的表示系数组合起来，整体进行考虑，得出所求问题的矩阵描述，

$$\min_{\alpha} \left\{ g(\mathbf{A}) = \sum_{l \in \{A, \varphi, \theta\}} \|\chi_l(\mathbf{y}) - \sum_{k=1}^K \chi_l(\mathbf{X}_k) \alpha_{\chi_l, k}\|_2 \right\} \quad (4.16)$$

其中 $\mathbf{A} = [\alpha_{\chi_A}, \alpha_{\chi_\varphi}, \alpha_{\chi_\theta}] \in \mathbb{R}^{n \times 3}$ 是由每个任务的系数向量组成的矩阵， $g(\mathbf{A})$ 是整体的重构误差。理论上 $g(\mathbf{A})$ 的值越小，则表示重构越精确。

为了挖掘不同学习任务间的内在相关性，本文将三个学习任务在同一个识别框架下同时进行考虑。我们认为三个线性系统有着相似的稀疏项，它们的系数向量有着类似的非零项索引位置，仅仅是权值系数不同。基于这一思路，我们使用最小化 ℓ_1/ℓ_q 范数的策略约束系数矩阵 \mathbf{A} 的可行集，

$$\min_{\mathbf{A}} g(\mathbf{A}) + \lambda \|\mathbf{A}\|_{q,1} \quad (4.17)$$

其中 λ 是正则化参数，用于保真度（重构误差）和稀疏度（表示系数）之间折衷， $\|\cdot\|_{q,1}$ 是混合范数，指先对矩阵每行求 ℓ_1 范数，再对结果求 ℓ_q 范数，该策略并不单独考虑每个向量，而是对系数矩阵进行整体约束，如图 4.6 所示。

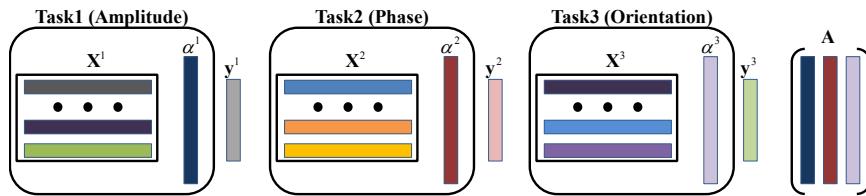


图 4.6 多任务学习示意图 [256]，每个方框代表一项训练任务，通过系数矩阵的联合稀疏约束探索每个学习任务间的内部关联性。

借助优化算法得出最优系数矩阵 $\hat{\mathbf{A}}$ 之后，根据所有任务的重构误差的累加判决未知样本的类别，

$$\min_{k=1, \dots, K} \sum_{l \in \{A, \varphi, \theta\}} \|\chi_l(\mathbf{y}) - \chi_l(\mathbf{X}_k) \hat{\alpha}_{\chi_l, k}\|_2 \quad (4.18)$$

4.1.3.2 数值实现

对于多任务学习问题，学者 Andreas Argyriou 和学者 Guillaume Obozinski 等开展了大量系统研究 [253, 256, 257]，通过借鉴他们研究成果，我们提炼出多特征多任务联合稀疏表示的数值实现。若以 $q = 2$ 为例，问题 (4.17) 演化为

$$\min_{\mathbf{A}} \sum_{l \in \{A, \varphi, \theta\}} \chi_l(\mathbf{y}) - \sum_{k=1}^K \|\chi_l(\mathbf{X}_k) \alpha_{\chi_l, k}\|_2^2 + \lambda \|\mathbf{A}\|_{1,2} \quad (4.19)$$

为了标记方便，我们将标识分量的下标 $l = \{A, \varphi, \theta\}$ 按顺序记为数字 $\{1, 2, 3\}$ ，用 α^i 表示矩阵 \mathbf{A} 的第 i 行，而 α_j 表示矩阵 \mathbf{A} 的第 j 列。公式 (4.19) 的目标函数在整个定义域区间可微，且它是凸函数，最优解 $\hat{\mathbf{A}}$ 满足

$$\begin{aligned}\hat{\alpha}_j &= 0, \|\nabla_{\alpha_j} g(\hat{\mathbf{A}})\|_2 \leq \lambda, \text{ or} \\ \hat{\alpha}_j &\propto -\nabla_{\alpha_j} g(\hat{\mathbf{A}}), \|\nabla_{\alpha_j} g(\hat{\mathbf{A}})\|_2 = \lambda\end{aligned}\quad (4.20)$$

其中 $\nabla_{\alpha_j} g(\hat{\mathbf{A}})$ 是线性系统的系数向量的偏微分（梯度）。H. Lee 等学者针对 ℓ_1 范数最小化最优化问题的求解提出了特征符号搜索算法 (Feature-Sign Search Algorithm) [159]。根据特征符号搜索算法可知只有那些活动集 (Active Set) 元素，即梯度向量的范数大于等于常数 λ' 参与稀疏系数向量的求解， $\lambda \frac{\hat{\alpha}_j}{\|\hat{\alpha}_j\|} = \nabla_{\alpha_j} g(\hat{\mathbf{A}})$ 。然而，公式 (4.20) 的条件太过苛刻，实际中很难满足，通常是转而求解它的次梯度，相应的不等式关系转化为

$$\begin{aligned}\hat{\alpha}_j &= 0, \|\nabla_{\alpha_j} g(\hat{\mathbf{A}})\|_2 < \lambda + \xi_0, \text{ or} \\ \|\nabla_{\alpha_j} g(\hat{\mathbf{A}}) + (\lambda - \xi) \frac{\alpha_j}{\|\alpha_j\|}\| &\leq \xi\end{aligned}\quad (4.21)$$

其中 ξ 和 ξ_0 是松弛参数。根据文献 [256] 可知，(4.17) 可通过交替迭代更新系数向量 α 和参数 λ' 求解，

$$\begin{aligned}\alpha^{(t+1)} &= \alpha^{(t)} - \varepsilon \frac{\nabla_{\alpha} g}{\|\nabla_{\alpha} g\|} \text{ and} \\ \lambda'^{t+1} &= \min(\lambda', \frac{g(\mathbf{A}^t) - g(\mathbf{A}^{t+1})}{\varepsilon})\end{aligned}\quad (4.22)$$

其中 t 是迭代索引， ε 是步长索引。得到问题 (4.17) 的最优稀疏系数矩阵之后，根据该矩阵计算每个类别在不同训练任务的重构误差的和即可判决未知样本的类别。多特征多任务联合稀疏表示的整个框架可归纳为图 4.7。

4.1.3.3 实例引证

为了说明多任务稀疏表示的优点，我们给出若干组四类目标 (BMP2、T72、BTR60 和 T62) 的识别实验，关于实验数据的具体描述参见论文的第二章。图 4.8 列出了其中的 6 组实验结果，包括每个子任务单独求解的结果以及多任务学习的结果。可以看出，若将各单信号分量构成的稀疏表示模型单独进行考虑，都存在着不同程度的错误判决。如基于幅度特征的稀疏表示仅能正确识别前三组测试样本，基于相位特征的稀疏表示仅能正确识别后两组测试样本，基于方位特征的稀疏表示只正确识别了第四组测试样本，其余的判决都是错误的。与之相比，多特征多任务联合稀疏表示能够正确识别所有 6 组测试样本。实验结果表明若单独考虑某个子任务通常都会偏爱于某种类别的表示系数，从而导致错误判决。多任务

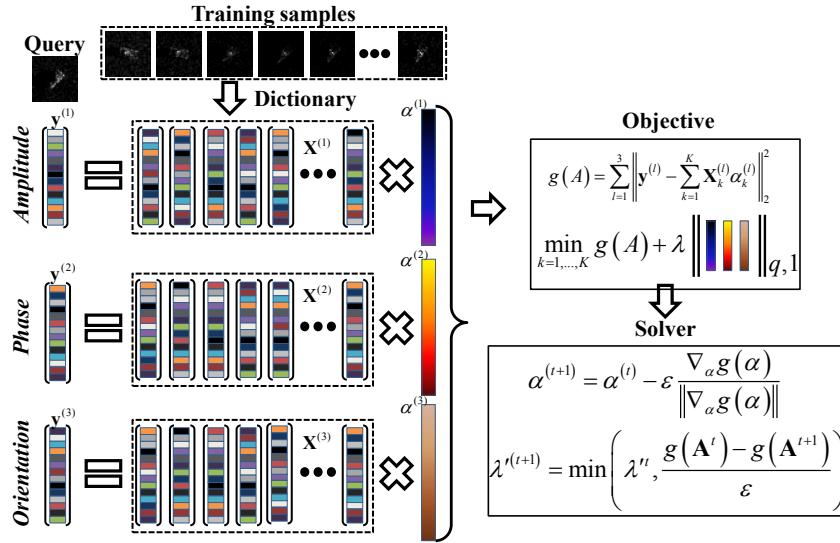


图 4.7 多任务多特征联合稀疏表示的框架流程。首先将 SAR 图像做单演信号多尺度表示，从而生成三组面向不同分量的特征，根据得到的特征构建线性表示模型。然后使用多任务训练方法求解线性模型，使其表示向量具有相似的稀疏项，从而在不同训练任务之间实现信息共享。根据不同任务所聚集的整体重构误差判决目标类型。

学习能将面向不同特征的子任务在同一个识别框架进行考虑，挖掘它们之间的内部关联性，在各子任务之间传递有价值的信息，因而得出的决策更稳健、更可靠。

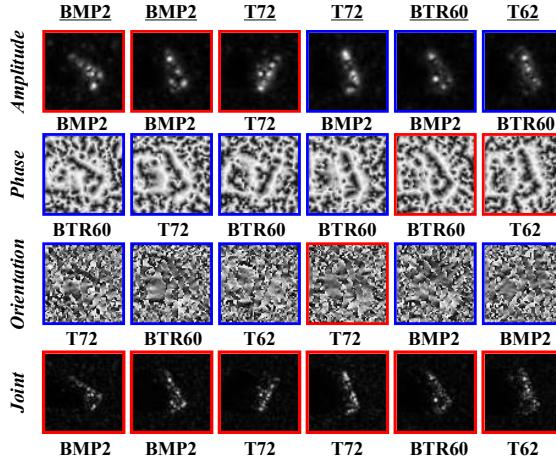


图 4.8 多任务联合稀疏表示实例引证。每一列表示的是一组分类实验，最顶端划线字典标识的是真实类别，前三行分别给出的是单独考虑每种分量（即不考虑任务之间的内部关联性）的决策结果，最后一行给出的是本文方法的联合决策，其中红色边框的图像表示决策正确，而蓝色边框的图像表示错误决策。

为直观地说明多特征多任务联合稀疏表示的原理，探寻多个学习任务间的内部相关性，图 4.9 给出一组实例，图示每个学习任务的系数向量与重构误差，比较单一任务学习和多任务学习的差别。未知样本源自于 T72，基于幅度特征的稀疏

表示决策为 T62，基于相位特征的稀疏表示将未知样本识别为 BMP2，基于方位特征的稀疏表示得出正确的判决，若单独执行三个学习任务，将得出完全不同的判决。多任务学习通过探寻不同任务间的内在相关性，将有价值的信息在不同任务间进行传递，实现信息共享，因而决策结果更可靠。

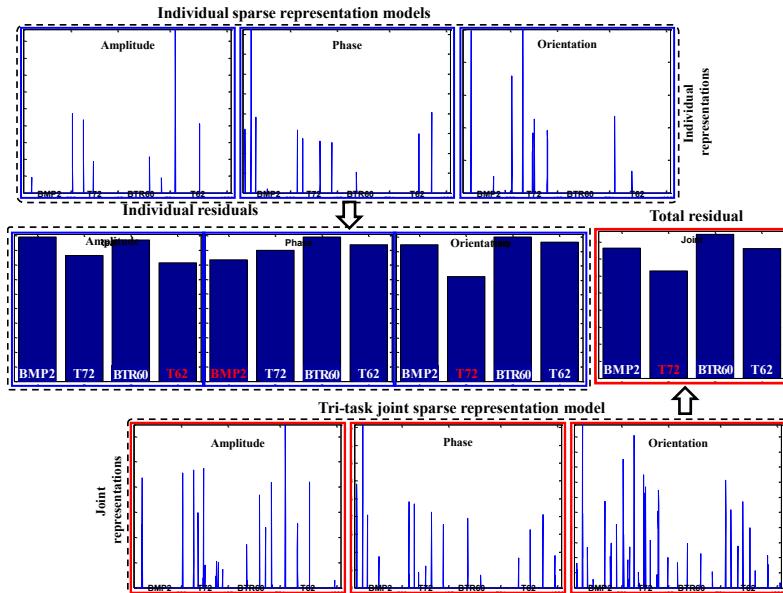


图 4.9 单任务稀疏表示与多特征多任务联合稀疏表示的比较。图中第一行是单独考虑每个学习任务得出的系数向量，相应的多任务联合稀疏表示的系数向量展示在第三行，第二行是单任务学习与多任务学习的重构误差，其中蓝色边框标识的是单训练任务的重构误差，红色边框标识的是多任务联合稀疏表示的累积重构误差。

4.2 实验验证

本章论文提出一种基于单演信号图像域信息融合的分类框架，并在该框架指导下形成三种具体算法：特征级融合、决策级融合和多特征多任务联合稀疏表示。下面将借助 MSTAR SAR 实测数据对算法的有效性从不同方面进行验证。实验设置从标准工作条件到扩展工作条件，其中扩展工作条件分为目标配置变化、俯仰角变化等。单演信号分量矩阵的下采样因子设为 $\rho = \frac{1}{8}$ ，对应的降采样后的图像尺寸为 10×10 像素。参与比较的基准算法即包括经典方法，如支持向量机、最近邻分类、稀疏表示分类、相关滤波器等，以及自适应 Boosting 分类、迭代图模式等最新识别方法 [66, 67]，详见表 4.3。

4.2.1 标准工作条件

首先考虑标准工作条件的目标识别问题。MSTAR 数据库中所有 10 类目标数据全部用于实验， 17° 雷达俯仰角的标准型号成像数据用作训练， 15° 雷达俯仰角

表 4.3 空域信息融合实验的基准比较算法。

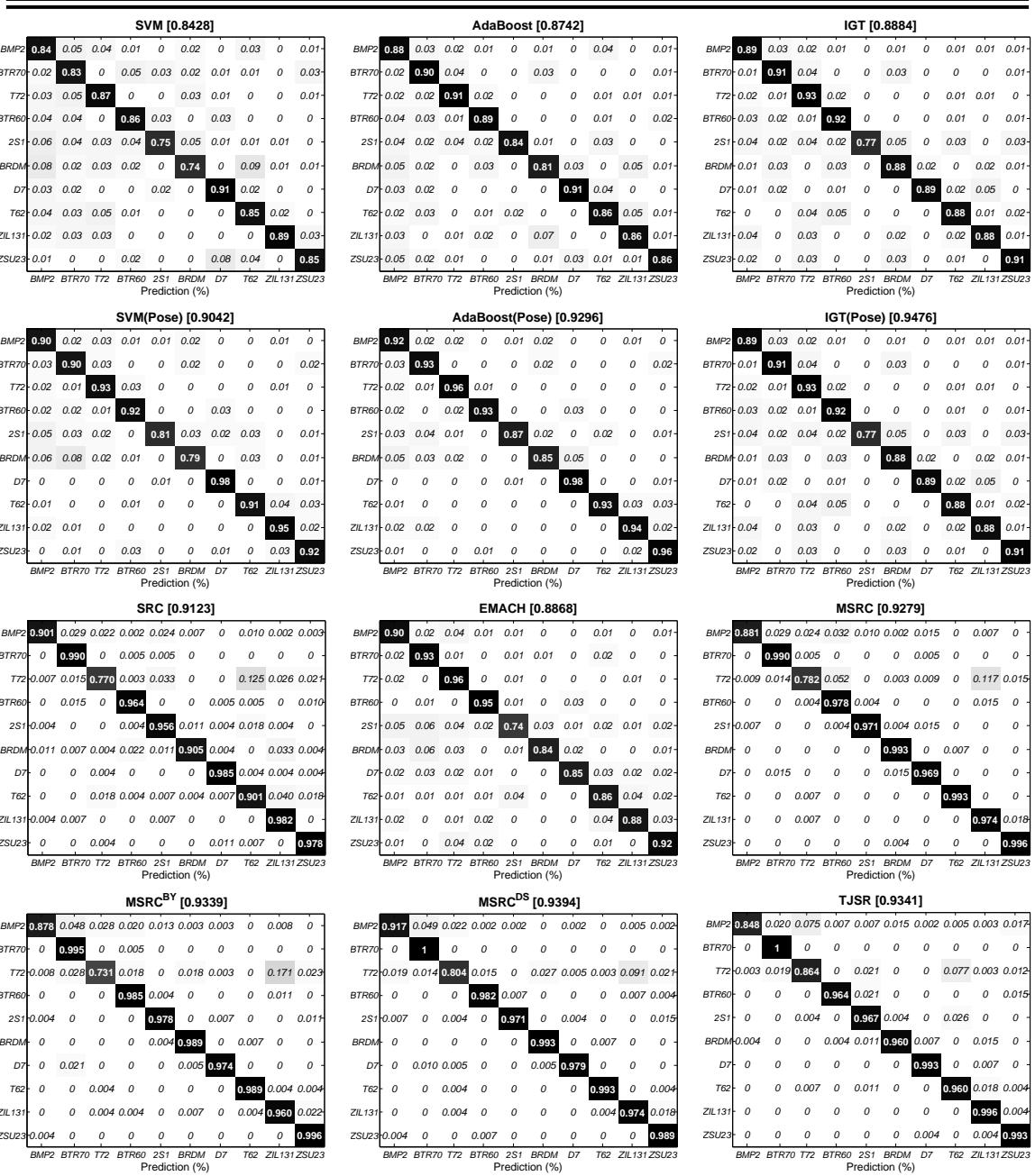
缩写	算法名称	输入特征
kNN	Nearest Neighbor Classifier [258]	原始强度值
SVM	Support Vector Machine Learning [65]	原始强度值
EMACH	Extended Maximum Average Correlation Height [71]	傅立叶变换系数
AdaBoost	Adaptive Boosting with RBF Network [66]	原始强度值及傅立叶变换
IGT	Iterative Graph Thickening [67]	小波变换（除高频外的频段）
SRC	Sparse Representation-based Classifier [95, 156]	原始强度值
MSRC	Feature-level Fusion (第 4.1.1 节)	增强的特征向量
MSRC ^{BY}	Decision-level Fusion via Bayes (第 4.1.2 节)	单演信号特征向量
MSRC ^{DS}	Decision-level Fusion via DS Theory (第 4.1.2 节)	单演信号特征向量
TJSR	Multi-task Joint Sparse Representation (第 4.1.3 节)	单演信号特征向量

的成像数据用于测试。三种数据融合方案与单任务学习的正确率如表 4.4 所示。可以看出，通过融合单演信号蕴含的不同类型信息，能有效弥补各分量单独执行学习的不足，提高决策的可靠性。同样的实验条件，三种信息融合方案的分类正确率分别 0.9292、0.9339、0.9346 和 0.9341，对应的幅度特征、相位特征、方位特征单任务学习的识别率为 0.8919、0.8501 和 0.8289，图像域信息融合方法将识别率提高了大约 4%~11%。单演信号诸分量蕴含着原始信号的不同类型信息（能量信息、结构信息和几何信息），将这些信息在图像域进行融合，能有效改进识别性能。

表 4.4 标准工作条件的空域信息融合，表中列出了单演信号分量信息融合的分类正确率。

Amplitude	Phase	Orientation	MSRC	决策级融合		TJSR
				MSRC ^{BY}	MSRC ^{DS}	
0.8919	0.8501	0.8289	0.9292	0.9339	0.9394	0.9341

三种图像域信息融合方法与基准算法在标准实验条件下的比较如图 4.10 所示，其中 AdaBoost、IGT、EMACH 和 SVM 四种算法的分类结果源自于文献 [65–67, 71]，SVM、AdaBoost 和 IGT 三种分类算法分两种情况进行讨论，即执行姿态角估计和盲适应（不执行姿态角估计，由算法自适应调整）。从实验结果可以看出，SVM、AdaBoost 和 IGT 受姿态角估计的影响较为明显，是否执行姿态角估计其分类正确率有显著差异，执行姿态角估计的正确率为 0.9042、0.9296 和 0.9475，盲适应的正确率为 0.8428、0.8742 和 0.8884，正确率下降 6.14%、5.54% 和 5.91%。然而，根据 SAR 切片图像估计目标的姿态角，需要执行相干斑抑制、图像分割、边



本章提出的分类算法无论是从稳健性还是有效性都优于传统的方法。在相同的实验环境下，特征级融合 MSRC、决策级融合 MSRC^{BY}、MSRC^{DS}、多特征多任务联合稀疏表示 TJSR 的整体正确率分别为 0.9292、0.9339、0.9394 和 0.9341，与参考算法执行姿态角估计的性能相仿，优于盲适应时的基准算法。基于 DS 证据理论的决策级融合方法表现最好，0.9394。

4.2.2 扩展工作条件

图像域信息融合算法对标准实验条件并没有比基准算法展现出显著优越性。为了挖掘图像域信息融合算法的优点，探索它对实际战场环境的可行性，需要提升实验难度，组织扩展工作条件的目标识别。扩展工作条件是指使用某种特定载机参数的成像数据训练分类算法，使用另一种不同载机工作条件的成像数据进行测试，论文的第二章对扩展工作条件进行了大量地举例说明，此处不再赘述。接下来的实验主要围绕扩展工作条件的目标识别展开，内容包括目标配置变化、俯仰角变化、杂波样本拒识别。

4.2.2.1 目标配置变化

实际战场通常存在许多军事意义上同一类别的多个物理变体目标，如型号升级，这就是目标配置变化。为了验证本章的分类算法在目标配置发生变化时的表现，论文组织四类目标识别实验，包括 BMP2、T72、BTR60 和 T62，其中 BMP2 是步兵战车，BTR60 是装甲输送车，T72 和 T62 是主战坦克，两两都具有十分相似的散射现象。 17° 雷达俯仰角标准型号 (BMP2_SN_9563 和 T72_SN_132) 成像数据用于训练算法， 15° 俯仰角其余型号的成像数据用于测试，测试型号与训练所用型号截然不同，实验样本总数如表 4.5 所示。论文的实验设置条件远比传统的方法（如文献 [65–67] 等）苛刻。

表 4.5 目标配置变化实验数据，粗体数字是训练样本数，下划线标示测试样本数。

类别	BMP2			T72			BTR60	T62
型号	SN_9563	SN_9566	SN_c21	SN_132	SN_812	SN_s7	k10yt7532	SN_A51
17°	[233]	232	233	[232]	231	228	[256]	[299]
15°	195	<u>196</u>	<u>196</u>	196	<u>195</u>	<u>191</u>	<u>195</u>	<u>273</u>

表 4.6 给出了不同算法的实验结果，包括混淆矩阵和整体识别率。从混淆矩阵可以看出，分类算法的识别性能与标准实验条件有不同程度的下降，性能差异主要集中在 BMP2 和 T72 两类多型号目标：BMP2 容易被误分为其余三类目标，以多特征多任务稀疏表示为例，BMP2 的分类正确率为 0.9107，将其误分为 T72、

表 4.6 目标配置变化对信息融合的影响，每个子表格括号中的数字是算法的整体识别率，表格内的数字是分类的混淆矩阵，其中对角线方格的粗体数字表示每类目标的分类正确率。

SVM [0.8459]					SRC [0.8692]				
Target	BMP2	T72	BTR60	T62	Target	BMP2	T72	BTR60	T62
BMP2	0.7985	0.1352	0.0434	0.0230	BMP2	0.9490	0.0306	0.0077	0.0128
T72	0.0466	0.7694	0.0	0.1839	T72	0.0596	0.6658	0.0052	0.2694
BTR60	0.0564	0.0154	0.8821	0.0462	BTR60	0.0051	0.0	0.9846	0.0103
T62	0.0037	0.0	0.0	0.9963	T62	0.0073	0.0183	0.0147	0.9597

MSRC ^{BY} [0.8804]					MSRC ^{DS} [0.8925]				
Target	BMP2	T72	BTR60	T62	Target	BMP2	T72	BTR60	T62
BMP2	0.9235	0.0357	0.0179	0.0230	BMP2	0.9490	0.0255	0.0102	0.0153
T72	0.0570	0.7021	0.0026	0.2383	T72	0.0492	0.7098	0.0026	0.2383
BTR60	0.0	0.0051	0.9949	0.0	BTR60	0.0	0.0051	0.9949	0.0
T62	0.0037	0.0073	0.0	0.9890	T62	0.0037	0.0073	0.0	0.9890

MSRC [0.8748]					TJSR [0.9117]				
Target	BMP2	T72	BTR60	T62	Target	BMP2	T72	BTR60	T62
BMP2	0.9286	0.0230	0.0255	0.0230	BMP2	0.9107	0.0510	0.0230	0.0153
T72	0.0415	0.6736	0.0259	0.2591	T72	0.0181	0.8394	0.0	0.1425
BTR60	0.0	0.0	0.9897	0.0103	BTR60	0.0256	0.0103	0.9385	0.0256
T62	0.0	0.0	0.0	1.0	T62	0.0	0.0037	0.0	0.9963

BTR60、T62 的比例分别为 0.0510、0.0230、0.0153；T72 很容易被误分类 T62，以稀疏表示 SRC 为例，T72 的分类正确率为 0.6658，将 T72 误分为 T62 和比例高达 0.2694。基于图像域信息融合的分类方法表现优于传统方法，基于 DS 证据推理的决策级融合 MSRC^{DS} BMP2 的正确率最高，0.9490；联合稀疏表示 TJSR 的 T72 正确率最高，0.8394。

从整体识别率看，算法的从优到劣的排序为：

$$\text{TJSR} > \text{MSRC}^{\text{DS}} > \text{MSRC}^{\text{BY}} > \text{MSRC} > \text{SRC} > \text{SVM},$$

多任务联合稀疏表示识别率为 0.9117，优于决策级融合方法，MSRC^{DS} 和 MSRC^{BY}，正确识别率分别为 0.8804 和 0.8925，优于特征级融合方法，正确识别率为 0.9279。这与第 4.1.1 节和第 4.1.2 节分析的结论基本一致，即特征级融合在信息合并过程中容易导致信息损失，决策级融合单独考虑每个学习任务，不能有效挖掘各学习任务间的内在关联性，也就不能实现信息传递与共享。与基准参考算法相比，基于单演信号空域信息融合的分类框架能够有效应对目标配置变化带来的不利影响，

这是因为我们引入了单演信号多尺度表示对 SAR 图像散射现象进行描述，并将单演信号多尺度分量所蕴含的不同类型的信息进行图像域融合，用于目标识别。

4.2.2.2 雷达俯仰角变化的连接体变体

接下来讨论不同俯仰角的目标识别问题，通过组织 2S1、BRDM2 和 ZSU23/4 三类目标识别实验观察分类算法对俯仰角变化的耐受性，其中 BRDM2 和 ZSU23/4 还分别包含连接体变体。连接体变体是指目标的空间结构改变，如炮管方向旋转和高度变化等，图 2.27 给出了一组连接体变体的实例。17° 雷达俯仰角标准型号目标的成像数据用于训练分类算法，30° 和 45° 雷达俯仰角变体型号目标的成像样本用于测试，即训练样本和测试样本的雷达俯仰角和目标型号同时发生改变。实验结果如表 4.7 所示，子表格分别给出了 30° 和 45° 俯仰角的混淆矩阵和整体识别率。

第一个实验场景使用 30° 雷达俯仰角工作条件下的成像样本进行测试，即训练样本（17° 俯仰角）和测试样本之间存在着 13° 的俯仰角差异。整体上各分类算法的识别率都很好，特征级融合和决策级融合分类甚至取得了 0.9829、0.9829 和 0.9838 的整体正确率，多特征多任务联合稀疏表示的整体正确率为 0.9524。结果表明微小的俯仰角变化无法区别算法之间的优劣，因而有必要组织条件更加苛刻的实验。第二个实验场景加剧俯仰角变化程度，使用将 45° 雷达俯仰角工作条件的成像数据进行测试，即训练样本和测试样本之间存在着 28° 的俯仰角差异。所有算法的识别率都有明显地下降，传统的稀疏表示分类的正确率甚至由 0.9488 下降到 0.5366，降幅达到 41.22%，波动比较剧烈。决策级融合和多任务联合稀疏表示则表现较好，分别取得了 0.7221、0.7613 和 0.7073 的识别率，基于 DS 证据推理的决策级融合 MSRC^{DS} 比其余方法 SRC、MSRC、MSRC^{BY}、TJSR 的正确率分别高 22.47%、12.19%、3.92%、5.39%。从实验结果的对比来看，基于图像域信息融合的目标识别方法受俯仰角变化的影响相对较小，这是因为它能够将不同类型信息组合起来用于目标识别，决策更加可靠和稳健。

4.2.3 杂波样本拒识

下面的实验用于讨论杂波样本的拒识问题，其中杂波包括各种自然虚警（如树木、灌木丛等）和人造虚警（如建筑物、桥梁、路灯等）。根据第二章第 2.4.3.1 节的介绍可知，MSTAR 公测数据集包含机动目标切片图像和场景图像两种。论文实验所用的杂波切片图像就是源自于 MSTAR 的场景图像截取而来，具体方法见 Angela R. Wise 等学者的相关研究内容 [35]。图 4.11 给出了一组杂波样本的示例，其中包括树木、路灯、建筑物的 SAR 图像。

表 4.7 俯仰角变化对空域信息融合的影响，图中子表格分别给出了 30° 和 45° 雷达俯仰角成像数据作为测试图像的分类结果，括号内粗体是算法的整体识别率，混淆矩阵对角线方格是每个类别的正确率。

SRC						
Depression	30° (0.9488)			45° (0.5366)		
	Target	2S1	BRDM2	ZSU23/4	2S1	BRDM2
2S1	0.9757	0.0243	0.0	0.7789	0.2013	0.0198
BRDM2	0.0595	0.9190	0.0214	0.4232	0.5650	0.0118
ZSU23/4	0.0172	0.0222	0.9606	0.6706	0.0237	0.3057
MSRC						
Depression	30° (0.9829)			45° (0.6394)		
	Target	2S1	BRDM2	ZSU23/4	2S1	BRDM2
2S1	0.9792	0.0714	0.0035	0.8647	0.0825	0.0528
BRDM2	0.0024	0.9833	0.0143	0.2648	0.4610	0.2742
ZSU23/4	0.0025	0.0123	0.9852	0.1848	0.1588	0.6564
MSRC ^{BY}						
Depression	30° (0.9829)			45° (0.7221)		
	Target	2S1	BRDM2	ZSU23/4	2S1	BRDM2
2S1	0.9965	0.0035	0.0	0.6601	0.3135	0.0264
BRDM2	0.0	0.9714	0.0286	0.0260	0.8038	0.1702
ZSU23/4	0.0049	0.0099	0.9852	0.0877	0.2370	0.6754
MSRC ^{DS}						
Depression	30° (0.9839)			45° (0.7613)		
	Target	2S1	BRDM2	ZSU23/4	2S1	BRDM2
2S1	0.9965	0.0035	0.0	0.7393	0.2277	0.0330
BRDM2	0.0	0.9738	0.0262	0.0449	0.7518	0.2033
ZSU23/4	0.0049	0.0099	0.9852	0.0569	0.1777	0.7654
TJSR						
Depression	30° (0.9524)			45° (0.7073)		
	Target	2S1	BRDM2	ZSU23/4	2S1	BRDM2
2S1	0.9861	0.0069	0.0069	0.5644	0.3366	0.0990
BRDM2	0.0214	0.9000	0.0786	0.0757	0.6738	0.2506
ZSU23/4	0.0025	0.0148	0.9828	0.0095	0.1469	0.8436

对于稀疏表示分类以及改进算法，MSRC、MSRC^{DS}、TJSR 等，都是根据最小重构误差来判决目标类型。重构误差越小，判决越可靠；反之，重构误差越大，判

决策结果越值得怀疑。根据这种工作原理，我们使用最小重构误差这一指标来实现非法样本的拒识。异常样本的拒识问题最重要的两个指标是检测率与虚警率。对于一个合法的目标样本，若它的最小重构误差小于某个特定的阈值 τ ，则它就对检测率 P_d 产生贡献；与之相反，对于一个非法的目标样本，若它的最小重构误差也小于该特定的阈值，则它就会对虚警率 P_{fa} 产生贡献。检测率就定义为所有被判定为合法的合法样本占合法样本总数的百分比，虚警率则定义为被误判为合法的杂波样本数目占杂波样本总数的百分比。

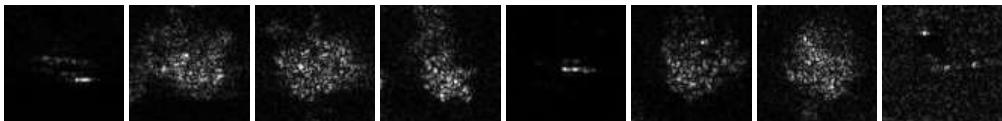


图 4.11 杂波样本示例。

在异常目标拒识实验中，通过调节特定阈值 τ 在某个范围内波动，就能生成检测率与虚警率变化曲线，即接收机工作特性曲线 (Receiver Operating Characteristics Curve)，这是一种最常用的检验识别算法拒识能力的指标。本文的未知样本拒识实验共分为自然杂波拒识和混淆目标拒识两组。

第一组：自然杂波拒识。拒识实验是从 MSTAR 自然场景图像中选出的 1616 组自然杂波作为非法样本，因而虚警率就是被误判为合法的杂波样本数占杂波样本总数 (1616) 的比值；而合法的目标设置为第三章第 2.4.3.1 节的表 2.2 中的前三类目标，BMP2、T72 和 BTR70，其总数为 1365，则检测率就是被识别为合法目标的总数占目标总数 (1365) 的比值。实验结果如图 4.12 所示。

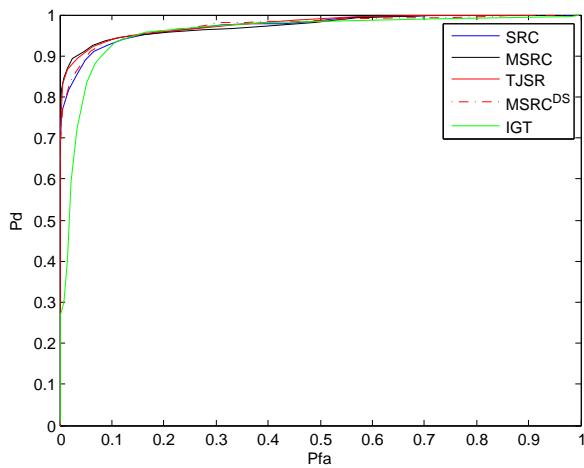


图 4.12 自然杂波拒识的 ROC 曲线。

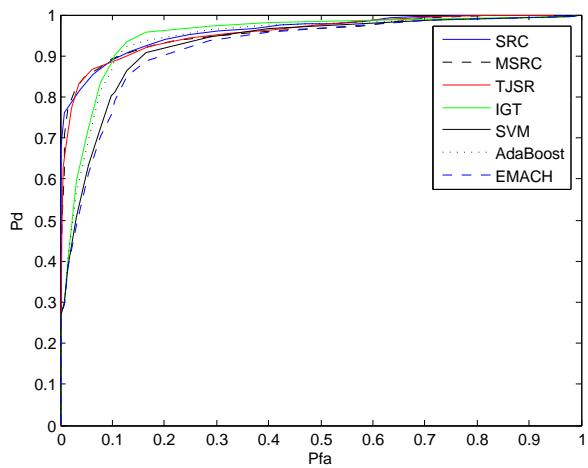


图 4.13 混淆目标拒识的 ROC 曲线。

从图 4.12 中可以看出，当虚警率大于 0.12，所有五种分类方法的检测几乎相等，优劣很难辨别；当虚警率小于 0.12，本章推荐的三种信息融合分类方法，

MSRC、MSRC^{DS} 和 TJSR 的检测率明显优于其余两种基准算法，SRC [95, 156] 和 IGT [67]。在虚警率保持不变的前提下，特征级融合、决策级融合和多任务联合稀疏表示三种分类算法的检测率明显优于其余的参考算法。实验结果表明，基于信息融合的分类框架能够有效地拒绝自然杂波为代表的非法样本。

第二组：混淆目标拒识：拒识实验选取 2S1 和 D7 两类目标作为非法样本，因而虚警率就被误判为真实目标的混淆目标占混淆目标总数 548 的比值；第二拒识实验的真实样本与第一组实验相同，BMP2、T72 和 BTR70 三类目标共 1365 组样本，则检测率就是被识别为合法目标的总数占目标总数（1365）的比值。实验结果如图 4.13 所示。

从实验结果来看，七种算法的性能差异主要集中在虚警率小于 0.2 的区域，迭代图鉴别模式 IGT 先抑后扬，在虚警率小于 0.1 时，SVM、IGT、EMACH 三种算法的检测率较低，SRC、MSRC、TJSR 三种算法的检测率较高，当虚警率大于 0.1 时，IGT 的检测率最高，AdaBoost 的检测率其次，SRC、TJSR 和 MSRC 居中，EMACH 和 SVM 最低。实验表明，虽然空域信息融合的分类框架表现逊于 IGT 和 AdaBoost，但也能较好地拒绝混淆目标。

4.3 本章小结

围绕着如何将单演信号应用于 SAR 图像目标识别问题，解决扩展工作条件的目标识别，论文提出基于单演信号空域信息融合的分类框架，整合单演信号蕴藏的不同类型信息，并在该框架指导下提出三种具体实现方案：

- 特征级融合：将单演信号多尺度表示生成的分量矩阵在空域直接进行合并，生成增强的单演信号特征向量，利用生成的特征向量构造稀疏线性表示模型，实现分类判决；
- 决策级融合：利用单演信号多尺度表示生成的系数矩阵分别设计面向各分量的独立的稀疏表示分类系统，然后借助贝叶斯理论或 DS 证据推理执行决策级融合，提高分类精度；
- 联合稀疏表示：利用单演信号多尺度表示生成的系数矩阵分别设计各自的特征向量，构建多任务学习模型，借助系数矩阵的稀疏约束实现多特征多任务联合稀疏表示，使得不同学习任务间能共享某些有用信息。

最后，借助 MSTAR 实测数据组织标准工作条件和扩展工作条件的目标识别实验，验证空域信息融合对目标配置变化、雷达俯仰角变化、连接体变体等因素的耐受性。

第五章 基于单演信号多核合成学习的 SAR 图像目标分类

为了刻画各种扩展工作条件下的目标散射现象，论文引入新的信号分析工具单演信号，于第四章提出基于图像域信息融合的分类学习，在该思路指导下形成特征级融合、决策级融合和多特征多任务联合稀疏表示三种具体方案。通过融合原始信号中的不同类型信息，如局部能量信息、局部结构信息、局部几何信息，实现分类学习。然而，这些学习方法建立在数据集线性可分的前提假设，对于线性不可分或是线性可分性较弱的数据集，需要重新进行考虑，寻找有效可行的办法。

根据第一章第 1.2.3.2 节关于分类学习的描述可知，借助非线性映射将原始数据空间投影到另一个抽象的特定空间即再生核 Hilbert 空间，投影空间数据可分性会明显改善，利用普通的学习方法就可实现分类识别 [137, 259, 260]。受其启发，本章提出基于再生核 Hilbert 空间合成核学习的 SAR 图像目标识别框架。首先使用非线性映射将单演信号特征投影至抽象空间中，投影空间的维数可以足够高甚至是无穷维，从而使得空间中的样本¹线性可分，然后利用合成核学习将投影空间中不同类型的信息进行融合，在投影空间中构建多元回归模型（即岭回归），通过求解回归系数判决样本类别。本章内容主要包括：1) 再生核 Hilbert 空间；2) 再生核 Hilbert 空间的分类学习；3) 基于多核合成学习的分类。

5.1 再生核 Hilbert 空间

线性不可分（或是线性可分性较弱）的数据空间将会对传统的线性分类学习造成不利影响，显著降低分类的可靠性。若存在一个非线性映射将原始数据空间投影到另一个抽象空间中，使得数据的线性可分性得到改善，遇到的问题就迎刃而解。机器学习将这个投影空间称为特征空间，其维数可以很高甚至是无穷维，导出特征空间的非线性映射称为是特征映射。如图 5.1 所示 [137]，假设我们需要解决两类分类问题，我们期望的理想决策边界是椭圆形²，然而传统的线性分类学习生成的决策边界却是一条直线，无法实现精确分类。若定义非线性映射 $\phi : \mathbb{R}^2 \mapsto \mathbb{R}^3$ 将原始数据空间投影至特征空间，

$$\phi : \mathbb{R}^2 \mapsto \mathbb{R}^3, (x_1, x_2)^T \mapsto (x_1^2, \sqrt{2}x_1x_2, x_2^2) \quad (5.1)$$

使得特征空间的像线性可分，类别之间的理想决策边界衍变成为一个超平面，故线性分类学习即可实现精确分类。这就是 Hilbert 空间投影的基本原理。

¹再生核 Hilbert 空间中的样本就是原始单演信号特征在投影空间的像， $\phi(\cdot)$ 。

²图 5.1 与论文第一章第 1.2.3.2 节的图 1.20 所表达的寓意类似。

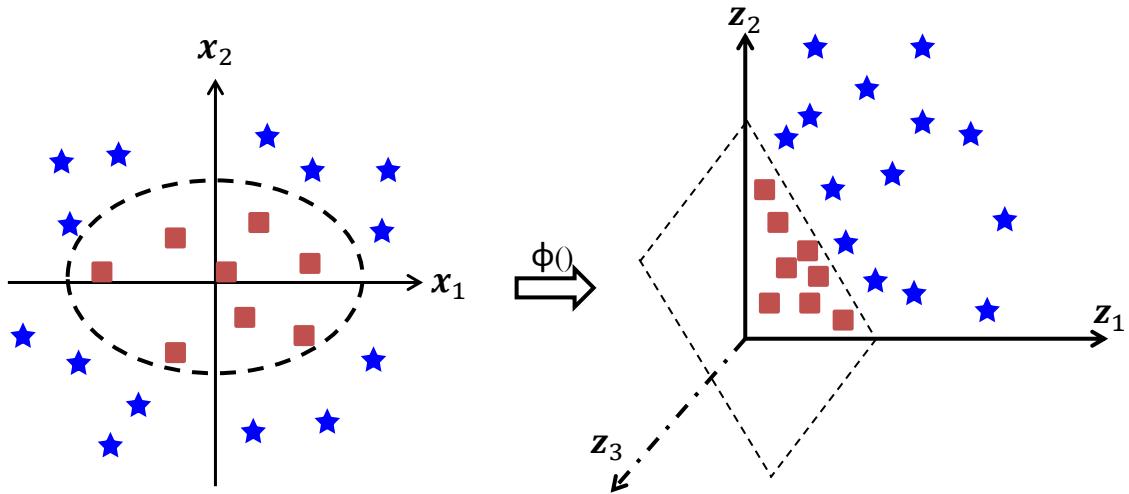


图 5.1 非线性映射引导的 Hilbert 空间原理，原始空间样本数据线性不可分，两类样本的决策边界是椭圆，因而无法找到一条直线（超平面）区分它们；若借助非线性映射将原始数据空间投影到高维 Hilbert 空间，使得投影生成的像线性可分，因而借助一个决策平面可实现分类。

为了将 Hilbert 空间中的像与像的相似计算与数据空间中的样本距离联系起来，数学家提出定义在超平面上的向量内积 (Inner Product)，通过二元正定核函数将其与原始空间的数据运算联系起来。以图 5.1 为例，对于二维多项式空间，投影空间像的内积与原始空间的样本有，

$$\kappa(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle = x_1^2 y_1^2 + 2x_1 x_2 y_1 y_2 + x_2^2 y_2^2 \quad (5.2)$$

定义 5.1： 欧氏空间 \mathcal{E} 的任意元素 $\mathbf{x}, \mathbf{y} \in \mathcal{E}$ 的核函数为 $\kappa(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle_{\mathcal{H}}$ ，其中 $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ 表示定义在投影空间 \mathcal{H} 的内积， $\phi(\cdot) : \mathcal{E} \mapsto \mathcal{H}$ 是从欧氏空间 \mathcal{E} 到投影空间 \mathcal{H} 的非线性映射。

构造核函数的目的是将投影空间中像的度量与原始空间的样本的运算联系起来，它可以看作是投影空间中像与像的关系在原始数据空间的具体数值实现，从而将特征空间与数据空间联系起来。

5.1.1 Mercer 核

任意核函数 $\kappa(\cdot, \cdot)$ ，若要诱导特征映射和特征空间必须满足一定的约束条件，即 Mercer 理论。下面分别讨论有限数据集和紧集上的核映射 [261]。

5.1.1.1 有限维特征空间

首先从最简单的有限维特征空间的特例开始进行讨论。

定义 5.2： 任意的实数对称矩阵 $\mathbf{K} \in \mathbb{R}^{N \times N}$ 正定当且仅当它满足 $\sum_{i,j} c_i c_j \mathbf{K}_{ij} \geq 0$ ， $c_1, c_2, \dots, c_N \in \mathbb{R}$ 。

给定实数集 $\Omega = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ 和实函数 $\kappa(\mathbf{x}, \mathbf{y})$, 用 \mathbf{K} 来表示 $N \times N$ 维矩阵, 其中 $\mathbf{K}_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$ 。由于矩阵 \mathbf{K} 是对称的, 对其进行对角化分解 $\mathbf{K} = \mathbf{U}\Lambda\mathbf{U}^T$, 其中 $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}$ 是所有特征值从小到大的顺序排列的对角矩阵, 而 $\mathbf{U} = [\mathbf{u}_1^T, \mathbf{u}_2^T, \dots, \mathbf{u}_N^T]^T$ 是由特征向量组成的酉矩阵, \mathbf{u}_i 是第 i 个特征值 λ_i 对应的特征向量。若矩阵 \mathbf{K} 正定, 它的所有特征值非负, 因而可以定义特征映射 $\phi : \Omega \mapsto \mathcal{H}(\mathbb{R}^N), x_i \mapsto \mathbf{u}_i \mathbf{D}^{1/2}$, 其中 $\mathbf{D}^{1/2} = \text{diag}\{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_N}\}$ 表示由所有特征值取平方根后组成的对角阵, 显然特征空间的内积 $\langle \cdot, \cdot \rangle$ 就与数据空间的核矩阵相联系起来,

$$\langle \phi(x_i), \phi(x_j) \rangle_{\mathcal{H}} = \mathbf{u}_i \mathbf{D} \mathbf{u}_j^T = (\mathbf{U} \mathbf{D} \mathbf{U}^T)_{ij} = \mathbf{K}_{ij} \quad (5.3)$$

5.1.1.2 紧子集上的 Mercer 核

我们以类似的方式讨论 N 维空间紧子集的 Mercer 条件。

定义 5.3 (Mercer 条件): 若用 Ω 表示 \mathbb{R}^N 空间上的紧子集, 假设 $\kappa : \Omega \times \Omega \mapsto \mathbb{R}$ 是连续对称函数, 且积分算子 $(\mathcal{T}_k f)(\mathbf{x}) = \int_{\Omega} \kappa(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$ 对于任意的函数 $f \in \mathcal{L}_2(\Omega)$ 都满足 $\int_{\Omega^2} \kappa(\mathbf{x}, \mathbf{y}) f(\mathbf{x}) f(\mathbf{y}) d\mathbf{x} d\mathbf{y} \geq 0$, 则根据算子 \mathcal{T}_k 的归一化特征函数 $\psi_i \in \mathcal{L}_2(\Omega)$ 可生成一组均匀的收敛序列 $\kappa(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{\infty} \lambda_i \psi_i(\mathbf{x}) \psi_i(\mathbf{y})$, 其中的归一化是指 $\|\psi_i\|_{\mathcal{L}_2} = 1$ 。

根据定义 5.3 可知, 算子 \mathcal{T}_k 的约束条件就是矩阵的正定约束条件的扩展。若根据算子 \mathcal{T}_k 的特征函数 ψ_i 定义特征映射序列, $\phi_d : \Omega \mapsto \mathcal{H} \mathbf{x} \mapsto (\sqrt{\lambda_1} \psi_1(\mathbf{x}), \dots, \sqrt{\lambda_d} \psi_d(\mathbf{x}))$ 映射序列 ψ_1, ψ_2, \dots 最终的收敛结果是 $\phi : \Omega \mapsto \mathcal{H}$, 且该映射满足 $\langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle_{\mathcal{H}} = \kappa(\mathbf{x}, \mathbf{y})$ 。据此, 我们给出 Mercer 核的定义 [262]:

定义 5.4 (Mercer 核): 若 Ω 表示 \mathbb{R}^N 空间的紧集, κ 是满足定义 5.3 的函数, 则存在从数据空间到特征空间的映射, $\phi : \Omega \mapsto \mathcal{H}$, 使得对任意元素 $\mathbf{x}, \mathbf{y} \in \Omega$ 都有

$$\kappa(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle_{\mathcal{H}}$$

对于任意正数 $\epsilon > 0$, 总存在一个数据空间到 Hilbert 空间的映射 ϕ_n , 使得对任意 $\mathbf{x}, \mathbf{y} \in \Omega$ 都有 $\|\kappa(\mathbf{x}, \mathbf{y}) - \langle \phi_n(\mathbf{x}), \phi_n(\mathbf{y}) \rangle_{\mathcal{H}}\| < \epsilon$ 。

推论 5.1: 数据空间 Ω 的对称函数 $\kappa : \Omega \times \Omega \mapsto \mathbb{R}$ 是 Mercer 核当且仅当由 Ω 上的任意子集所构成的核矩阵是半正定矩阵, 即它的特征值非负。

5.1.2 再生核 Hilbert 空间

图 5.1 是 Hilbert 空间的简单示意, 原始数据空间为 \mathbb{R}^2 , 投影空间为 \mathbb{R}^3 。若投影后的空间设为无穷维, 问题就比较复杂, 需要重新进行讨论。将投影空间从有限维推广到无限维, 处理的基本思路不变, 但是需要进行严格的数学推导。

若非线性映射 $\phi(\cdot)$ 将原始数据空间的样本投影到另一个抽象的高维空间，这个高维空间就是“再生核希尔伯特空间”(Reproducing Kernel Hilbert Space)。所谓“再生核”是指用于内积计算的核函数，内积是内积空间赋予的一种运算结构，是标量点积到线性空间的推广，用于度量两个向量间的“夹角”。内积空间是线性空间的一个推广，线性空间中的元素对加法和数乘两种运算结构封闭。若在线性空间中定义一个内积操作，且空间满足正定性、共轭对称性和对第一变元线性，就会构成内积空间。内积空间可以成为赋范线性空间，它满足非负性、齐次性和三角不等式关系，其范数可由 $\|\mathbf{x}\|_2^2 = \langle \mathbf{x}, \mathbf{x} \rangle$ 导出，该范数可诱导线性空间的度量， $d(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{x}_1 - \mathbf{x}_2\|$ 。因而，内积空间是赋予了内积的赋范线性空间，根据赋范线性空间可导出度量空间，度量空间是更抽象的一种线性空间。内积空间是一类特殊的赋范线性空间，完备的内积空间称为希尔伯特空间(Hilbert Space)。完备的概念源自于度量空间，它是指度量空间中的任意柯西序列都收敛于该空间之内的某个点³。

定义 5.5 (完备的度量空间): 设 (\mathbf{X}, d) 为度量空间， $\{x_n\} \subset \mathbf{X}$.

- 1) $\forall \varepsilon > 0, \exists N \in \mathbb{Z}$, 使得对一切 $n, m > N$ 都有 $d(x_n, x_m) < \varepsilon$, 称 $\{x_n\}$ 是 \mathbf{X} 中的基本列(或称 Cauchy);
- 2) 如果 \mathbf{X} 中的每个基本列都收敛于 \mathbf{X} 中的点, 称 \mathbf{X} 是完备的度量空间。

赋范线性空间是度量空间的特例，它的完备性与度量空间有一些差别。赋范线性空间的 Cauchy 列 $\{x_n\}$ 表述为: $\forall \varepsilon > 0, \exists N \in \mathbb{Z}$, 使得对一切 $n, m > N$ 有 $\|x_n - x_m\| < \varepsilon$, 或是 $\|x_n - x_m\| \rightarrow 0(n, m \rightarrow \infty)$ 。实数域的收敛数列当且仅当它是 Cauchy 列，对于赋范线性空间 \mathbf{X} , 若 $\{x_n\}$ 收敛到 $x \in \mathbf{X}$, 则它必是 Cauchy 列 $\|x_n - x_m\| \leq \|x_n - x\| + \|x_m - x\|$ 。然而逆命题并不成立，若 $\{x_n\}$ 是 \mathbf{X} 中的 Cauchy 列，它未必在 \mathbf{X} 中收敛。赋范线性空间中的每个 Cauchy 列收敛当且仅当空间是完备的，同一线性空间在不同的范数之下，其完备性可能不同。再生核 Hilbert 空间是定义在希尔伯特空间的函数空间。

定义 5.6 (再生核 Hilbert 空间): 若以 Ω 表示非空集合, 用 \mathcal{H} 表示函数族 $f : \Omega \mapsto \mathbb{R}$ 的 Hilbert 空间, 如果存在函数 $\kappa : \Omega \times \Omega \mapsto \mathbb{R}$ 满足条件:

- \mathcal{H} 空间的任意元素 $f \in \mathcal{H}$ 都有, $\langle f, \kappa(\mathbf{x}, \cdot) \rangle_{\mathcal{H}} = f(\mathbf{x})$, 且 $\langle \kappa(\mathbf{x}, \cdot), \kappa(\mathbf{y}, \cdot) \rangle_{\mathcal{H}} = f(\mathbf{x})$;
- 由函数 κ 张成的空间 $\mathcal{H} = \overline{\text{span}\{\kappa(\mathbf{x}, \cdot) | \mathbf{x} \in \Omega\}}$, 其中 \cdot 表示补集。

则称 \mathcal{H} 为赋予内积运算 $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ 的再生核 Hilbert 空间。

³ 度量空间、赋范线性空间、内积空间的定义及相互关系详见附录 A。

任意的再生核 Hilbert 空间必然存在由该空间内积定义的唯一的正定核函数。

定义 5.7 (正定核): 对于任意集合 Ω 中的元素 $\mathbf{x}_i, \mathbf{x}_j$, 对称实函数 $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \kappa(\mathbf{x}_j, \mathbf{x}_i)$ 是正定核函数当且仅当对于集合 Ω 的所有元素 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ 总存在实数 c_1, c_2, \dots, c_N 使得 $\sum_{i,j} c_i c_j \kappa(\mathbf{x}_i, \mathbf{x}_j) \geq 0$ 都成立, 其中 N 是任意的自然数。

维数相同的再生核 Hilbert 空间相互同构, 即空间的各种运算结构 (如内积、范数、度量和向量运算等) 在不同空间转换时能够得以保持。这条重要性质为后续计算奠定非常重要的基础, 我们不需要关注再生核 Hilbert 空间的具体形式, 借助正定核函数即将不同类型空间不同性质的运算联系起来。

对于非线性映射 $\phi(\cdot) : \mathcal{E} \mapsto \mathcal{H}$, 我们并不要求原始数据空间必须是欧氏空间或是向量空间。投影空间中的像的具体形态并不影响数值计算, 我们可以认为它是定义在原始数据空间 \mathcal{E} 的函数, 在一定条件下, 通常可以找到对应于希尔伯特空间的一个唯一再生核函数, $\kappa : \mathcal{S} \times \mathcal{E} \rightarrow \mathbb{R}$, 它满足

1. 对于任意 $\mathbf{x}_0 \in \mathcal{E}$, 都有 $\kappa(\mathbf{x}, \mathbf{x}_0) \in \mathcal{H}$;
2. 对于任意 $\mathbf{x} \in \mathcal{E}$ 和 $\mathbf{f}(\cdot) \in \mathcal{H}$, 都有 $f(\mathbf{x}) = \langle f(\cdot), \kappa(\cdot, \mathbf{x}) \rangle$ 。

第二条性质就是再生性。若我们把映射 $\phi(\cdot)$ 定义为 $\phi(\mathbf{x}) = \kappa(\cdot, \mathbf{x})$, 根据再生性有:

$$\langle \phi(\mathbf{x}_1), \phi(\mathbf{x}_2) \rangle = \langle \kappa(\cdot, \mathbf{x}_1), \kappa(\cdot, \mathbf{x}_2) \rangle = \kappa(\mathbf{x}_1, \mathbf{x}_2)$$

从而实现原始数据空间与投影空间联系。

欧氏空间常用的核函数有 ($\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$):

- Linear Kernel— $\kappa(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$
- Polynomial Kernel— $\kappa(\mathbf{x}, \mathbf{y}) = (\alpha \mathbf{x}^T \mathbf{y} + b)^d, \alpha, b, d \in \mathbb{R}$
- Sigmoid Kernel— $\kappa(\mathbf{x}, \mathbf{y}) = \tanh(\alpha \mathbf{x}^T \mathbf{y} + c), \alpha, c \in \mathbb{R}$
- Gaussian Kernel— $\kappa(\mathbf{x}, \mathbf{y}) = \exp\{-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\sigma^2}\}, \sigma > 0$
- ANONA Kernel— $\kappa(\mathbf{x}, \mathbf{y}) = \exp(-\sigma(\mathbf{x}^k - \mathbf{y}^k)^2)^d, k, d, \sigma \in \mathbb{R}$ ⁴。
- Cauchy Kernel— $\kappa(\mathbf{x}, \mathbf{y}) = \frac{1}{\frac{\|\mathbf{x}-\mathbf{y}\|^2}{\sigma^2} + 1}$

此外, 还有一些不太常用的核函数, 如 Circular Kernel、Wavelet Kernel、Spherical Kernel、Power Kernel、Log Kernel 等⁵。

⁴ANONA 是方差分析 Analysis of Variance 的简称, 又称变异数分析, 是 Ronald Aylmer Fisher 提出, 用于两个及两个以上样本均数差别的显著性检验。

⁵关于核函数的更多论述请参照博客 <http://www.cnblogs.com/yingying0907/p/3651223.html>。

5.1.3 正定核函数的特性

定理 5.1: 正定核函数对加法、数乘、积分、与单秩核函数相乘、取极限封闭。

若有正定核函数 $\kappa_1(\mathbf{x}, \mathbf{y}), \kappa_2(\mathbf{x}, \mathbf{y})$, 它们的线性组合 $\alpha_1\kappa_1(\mathbf{x}, \mathbf{y}) + \alpha_2\kappa_2(\mathbf{x}, \mathbf{y})$ 、点乘运算 $\kappa_1(\mathbf{x}, \mathbf{y})\kappa_2(\mathbf{x}, \mathbf{y})$ 、积分运算 $\int \kappa_1(\mathbf{x}, \mathbf{z})\kappa_2(\mathbf{y}, \mathbf{z})$ 、与单秩核函数相乘 $\kappa(\mathbf{x}, \mathbf{z})f(\mathbf{x})f(\mathbf{y})$; 是正定核函数, 若正定核函数序列 $\kappa_1(\mathbf{x}, \mathbf{y}), \kappa_2(\mathbf{x}, \mathbf{y}), \dots$ 收敛, 它们的极限 $\lim_{i \rightarrow \infty} \kappa_i(\mathbf{x}, \mathbf{y})$ 也是正定的。

定理 5.2: 对于任意的正定核函数 $\kappa(\mathbf{x}, \mathbf{y})$, 则 $\exp\{\kappa(\mathbf{x}, \mathbf{y})\}$ 和 $f(\kappa(\mathbf{x}, \mathbf{y}))$ 也是正定的, 其中 $f : \mathbb{R} \mapsto \mathbb{R}$ 是任意的多项式函数。

定理 5.1 和定理 5.2 为正定核函数的应用提供了很大便利, 根据这些特性就能方便地用一些简单的核函数来构建复合核函数。

5.2 Hilbert 核空间线性表示

对于线性不可分的数据集, 通常借助一个非线性映射将原始数据空间投影到一个抽象的特征空间, 挖掘数据内部的非线性几何特性, 改善数据的可分性。受该思路的启发, 学者们提出了一种核稀疏表示的分类方法 [97, 263–265], 首先使用与正定核函数相关联的一组非线性映射将原始不可分数据 (或可分性较差) 投影到高维特征空间, 在抽象的特征空间中构造多元线性回归模型, 即将未知样本的像表示成为所有训练样本的像的线性组合, 通过求解最优回归系数判定未知样本的类别。

假设训练集中 K 个类别都有足够的训练样本 $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,n_k}]$ 。对于任意未知样本的分类问题, 传统的稀疏信号表示首先使用训练样本构造过完备冗余字典, 再使用冗余字典对未知样本进行线性表示, 通过求解线性系统的最优稀疏解进行判决, 详细描述参见第二章第 2.3 节的相关内容。借助核映射的思路可对稀疏表示分类做进一步扩展, 应对数据线性不可分或是线性可分性较差等情形。假设存在非线性映射 $\phi(\cdot) : \mathbb{R}^m \mapsto \mathcal{H}$, 将数据空间 (\mathbb{R}^m) 元素投影至另一个抽象的特征空间 (用 \mathcal{H} 表示), 即 $\forall \mathbf{x} \in \mathbb{R}^m$, 总存在一个位于 Hilbert 空间的像, $\exists \phi(\mathbf{x}) \in \mathcal{H}$ 。投影的像构成的特征空间数据的线性可分性有很大程度的改善, 因而使用简单的线性分类器就能完成分类任务。在投影空间构建线性表示模型, 根据非线性映射将不同类别的训练样本投影到特征空间, 使用特征空间的像构造过完备字典 $\mathbf{D}_{\mathcal{H}} = [\phi(\mathbf{X}_1), \phi(\mathbf{X}_2), \dots, \phi(\mathbf{X}_K)]$ 其中子字典 $\phi(\mathbf{X}_k) = [\phi(\mathbf{x}_{k,1}), \phi(\mathbf{x}_{k,2}), \dots, \phi(\mathbf{x}_{k,n_k})]$ 的每个原子是训练样本在投影空间的像。然后根据投影空间的字典对未知样本的像进行线性表示,

$$\phi(\mathbf{y}) = \phi(\mathbf{x}_{1,1})\alpha_{1,1} + \phi(\mathbf{x}_{1,2})\alpha_{1,2} + \cdots + \phi(\mathbf{x}_{K,n_K})\alpha_{K,n_K} = \mathbf{D}_{\mathcal{H}}\boldsymbol{\alpha} \quad (5.4)$$

式中 $\alpha = [\alpha_{1,1}, \alpha_{1,2}, \dots, \alpha_{K,n_K}] \in \mathbb{R}^n$ 是系数向量。为了从系数向量中寻找未知样本的类别信息，通常希望 α 仅有少量的非零值，其余全部为 0，即尽可能稀疏。该思想可通过最小化 ℓ_1 范数的最优化策略实现（详见第二章第 2.3 所述）[266]，

$$\min_{\alpha} \|\alpha\|_1 \quad \text{subject to} \quad \|\phi(\mathbf{y}) - \mathbf{D}_{\mathcal{H}}\alpha\|_2 \leq \epsilon \quad (5.5)$$

但非线性映射 $\phi(\cdot)$ 的具体形式未知，直接求解问题 (5.5) 似乎不太可能，借助核函数可实现问题求解，即使映射 $\phi(\cdot)$ 的具体形式未知也能计算最优回归系数。对公式 (5.5) 做变形化简，将它的约束项（记为 $g(\alpha)$ ）取平方做向量展开 [267]，

$$\begin{aligned} g(\alpha) &= \|\phi(\mathbf{y}) - \mathbf{D}_{\mathcal{H}}\alpha\|_2^2 = (\phi(\mathbf{y}) - \mathbf{D}_{\mathcal{H}}\alpha)^T (\phi(\mathbf{y}) - \mathbf{D}_{\mathcal{H}}\alpha) \\ &= \alpha^T \mathbf{D}_{\mathcal{H}}^T \mathbf{D}_{\mathcal{H}} \alpha - 2\alpha^T \mathbf{D}_{\mathcal{H}}^T \phi(\mathbf{y}) + \phi(\mathbf{y})^T \phi(\mathbf{y}) \end{aligned} \quad (5.6)$$

式中 $\phi(\mathbf{y})^T \phi(\mathbf{y})$ 是与 α 无关的常数项，涉及到 α 最优求解的只有 $\mathbf{D}_{\mathcal{H}}^T \mathbf{D}_{\mathcal{H}}$ 和 $\mathbf{D}_{\mathcal{H}}^T \phi(\mathbf{y})$ 两项，分别对这两项进行展开

$$\mathbf{D}_{\mathcal{H}}^T \mathbf{D}_{\mathcal{H}} = \begin{bmatrix} \kappa(\mathbf{x}_{1,1}, \mathbf{x}_{1,1}) & \kappa(\mathbf{x}_{1,1}, \mathbf{x}_{1,2}) & \dots & \kappa(\mathbf{x}_{1,1}, \mathbf{x}_{K,n_K}) \\ \kappa(\mathbf{x}_{1,2}, \mathbf{x}_{1,1}) & \kappa(\mathbf{x}_{1,2}, \mathbf{x}_{1,2}) & \dots & \kappa(\mathbf{x}_{1,2}, \mathbf{x}_{K,n_K}) \\ \vdots & \vdots & \ddots & \vdots \\ \kappa(\mathbf{x}_{K,n_K}, \mathbf{x}_{1,1}) & \kappa(\mathbf{x}_{K,n_K}, \mathbf{x}_{1,2}) & \dots & \kappa(\mathbf{x}_{K,n_K}, \mathbf{x}_{K,n_K}) \end{bmatrix} = \kappa(\mathbf{X}, \mathbf{X})$$

其中 $\kappa(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle$ 是正定核函数（详见定义 5.7）， $\kappa(\mathbf{X}, \mathbf{X}) = \mathbf{D}_{\mathcal{H}}^T \mathbf{D}_{\mathcal{H}} \in \mathbb{R}^{n \times n}$ ⁶ 也被称作是格莱姆矩阵⁷，格莱姆矩阵中的每个元素是再生核 Hilbert 空间中两个像的内积运算（类似于线性空间中两个向量的夹角）。同样的方法将 $\mathbf{D}_{\mathcal{H}}^T \phi(\mathbf{y})$ 展开，

$$\mathbf{D}_{\mathcal{H}}^T \phi(\mathbf{y}) = \begin{bmatrix} \phi(\mathbf{x}_{1,1}) \\ \phi(\mathbf{x}_{1,2}) \\ \vdots \\ \phi(\mathbf{x}_{K,n_K}) \end{bmatrix} \quad \phi(\mathbf{y}) = \begin{bmatrix} \kappa(\mathbf{x}_{1,1}, \mathbf{y}) \\ \kappa(\mathbf{x}_{1,2}, \mathbf{y}) \\ \vdots \\ \kappa(\mathbf{x}_{K,n_K}, \mathbf{y}) \end{bmatrix} = \kappa(\mathbf{X}, \mathbf{y})$$

展开后的结果是列向量，向量的每一项数值是未知样本的像 \mathbf{y} 与相应的训练样本的像的内积，因而可将其看作是未知样本与训练样本在 Hilbert 空间的相似向量。将展开后的格莱矩阵和相似向量代回公式 (5.5) 可得⁸

$$\min_{\alpha} \|\alpha\|_1 \quad \text{subject to} \quad \alpha^T \kappa(\mathbf{X}, \mathbf{X}) \alpha - 2\alpha^T \kappa(\mathbf{X}, \mathbf{y}) + \kappa(\mathbf{y}, \mathbf{y}) \leq \epsilon' \quad (5.7)$$

⁶ $n = \sum_{i=1}^K n_i$ 是所有训练样本的总数。

⁷Gram Matrix 是由 n 维欧氏空间中任意 $k \leq n$ 个向量内积构成的半正定矩阵，详见附录 A。

⁸代入的误差项是原来的误差项取平方即误差能量，使用误差能量不改变原来问题的最优解。

显然不需要关注非线性映射 $\phi(\cdot)$ 的具体形式，借助核函数实现内积计算即可完成最优化求解。

若选定某个核函数，问题 (5.7) 就可以使用传统的优化算法如线性规则进行求解。算出最优稀疏解 $\hat{\alpha}$ 之后，可根据每个类别的残差判决未知样本，

$$\min_{k=1,\dots,K} \|\phi(\mathbf{y}) - \phi(\mathbf{X}_k)\hat{\alpha}_k\|_2 \quad (5.8)$$

对于投影后的特征空间，为了使样本之间的相似度量更加精确，它的空间维数可以设想成很高甚至是无穷维，因而数据的线性可分性会有极大地改善，故对于线性系统 (5.4) 我们也可以进一步松弛它的约束项，使用最小化 ℓ_2 范数约束系数向量的可行集 [267]，

$$\min_{\alpha} \|\alpha\|_2 \quad s.t. \quad \|\phi(\mathbf{y}) - \mathbf{D}_{\mathcal{H}}\alpha\|_2 \leq \epsilon \quad (5.9)$$

根据第二章的分析可知，问题 (5.9) 等价于无约束优化问题，

$$\min_{\alpha} \|\phi(\mathbf{y}) - \mathbf{D}_{\mathcal{H}}\alpha\|_2^2 + \lambda \|\alpha\|_2^2 \quad (5.10)$$

其中 λ 是常数，用于折衷重构误差与稀疏程度之间的矛盾。对于问题 (5.10) 的目标函数我们可以采样与上面类似的方法将其展开成 α 的函数（记为 $g'(\alpha)$ ），

$$g'(\alpha) = \alpha^T \kappa(\mathbf{X}, \mathbf{X})\alpha - 2\alpha^T \langle \kappa(\mathbf{X}, \mathbf{y}) + \kappa(\mathbf{y}, \mathbf{y}) + \lambda \alpha^T \alpha \rangle \quad (5.11)$$

函数 $g'(\alpha)$ 在定义域内连续可微，对其求微分即可直接得到函数的极小值，

$$\frac{dg'(\alpha)}{d\alpha} = 2\kappa(\mathbf{X}, \mathbf{X})\alpha - 2\kappa(\mathbf{X}, \mathbf{y}) + 2\lambda\alpha \quad (5.12)$$

令 $\frac{dg'(\alpha)}{d\alpha} = 0$ 可得函数的极小值为

$$\hat{\alpha} = (\kappa(\mathbf{X}, \mathbf{X}) + \lambda \mathbf{I})^{-1} \kappa(\mathbf{X}, \mathbf{y}) \quad (5.13)$$

其中 \mathbf{I} 表示单位矩阵。可以看出，核线性表示 (5.9) 的解有解析表达式，因而计算效率优于稀疏表示 (5.5)。算出系数向量之后，可根据 (5.8) 判决未知样本的类别（算法流程如图 5.2 所示）。再生核 Hilbert 空间线性表示分类的数值实现可以概括为算法 5.1。

对于无约束 ℓ_2 范数最优化问题 (5.10)，统计学也称之为 Kernel Ridge Regression（核岭回归），其实质是利用设计和映射将原来的线性回归模型的回归参量和回归响应分别投影到抽象的再生核 Hilbert 空间，根据回归参量和回归响应投影得到的像求解再生核 Hilbert 空间岭回归模型的回归系数。可以看出，论文第 4 章借助稀疏表示求解线性岭回归模型，而本章是利用稀疏表示求解核岭回归模型，然

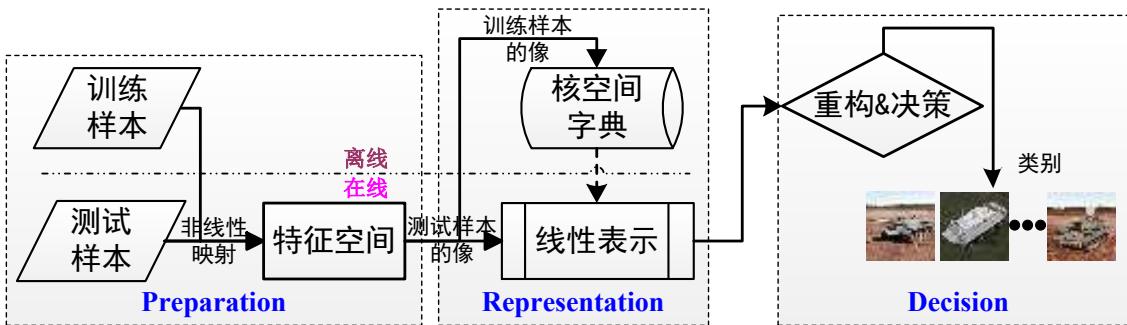


图 5.2 基于 Hilbert 空间核稀疏表示的分类流程，该流程可划分为三个阶段：(1) 准备，使用非线性映射将测试样本和训练样本投影到 Hilbert 空间；(2) 重构：使用训练样本的 Hilbert 空间的像构成字典，根据该字典原子对测试样本的像进行线性表示；(3) 决策：求解最优表示系数，重构原始信号，根据重构误差进行判决。

算法 5.1 再生核 Hilbert 空间稀疏表示分类算法

已知: K 类的训练样本 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_K]$, 其中 $\mathbf{X}_i = [\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \dots, \mathbf{x}_{i,n_i}]$;

未知样本 $\mathbf{y} \in \mathbb{R}^m$;

允许的噪声误差 ϵ 或正则化参数 λ ;

求：未知样本的类别；

- 1: 选择 Mercer 核函数, 合理设置参数;
 - 2: 根据训练样本计算格莱姆核矩阵 $\kappa(\mathbf{X}, \mathbf{X})$ 以及相似向量 $\kappa(\mathbf{X}, \mathbf{y})$;
 - 3: 求解最小化 ℓ_1 范数优化问题 (5.7 或者 5.10);
 - 4: 根据最优稀疏解 $\hat{\alpha}$ 求解每个类别样本的重构误差;
 - 5: **for** $i = 1$ to K **do**
 - 6: 计算重构误差 $r_i(\mathbf{y}) = \|\phi(\mathbf{y}) - \phi(\mathbf{X}_i)\hat{\alpha}_i\|_2$;
 - 7: **end for**
 - 8: 推断未知样本的类别 $\min_{i=1, \dots, K} r_i(\mathbf{y})$

后根据得到的回归系数判决未知样本的类别。具体求解方法前文已经给出了详细步骤，有关 Linear Ridge Regression 和 Kernel Ridge Regression 的详细介绍参见论文附录 C 部分。

5.3 单演信号 Hilbert 空间多核合成学习

为了应对线性不可分的数据集，前文中使用非线性映射将原始数据空间投影到抽象的特征空间中，改善数据的线性可分性。为了在特征空间充分利用不同类型的信息（如谱信息，空间上下文信息），西班牙学者 Gustavo Camps-Valls 等提出了一种用于超光谱地物分类的合成核学习方法（Composite Kernels Learning），根据第 5.1.3 节的论述可知，若干个正定核函数的线性组合是正定核函数，因此把代表空间信息和谱信息的一簇核函数组合起来仍是合法的核函数，从而融合不同

类型信息用于超光谱分类 [268, 269]，他们借助多核合成将空间上下文附加信息带入到支持向量机分类器，最终生成的核函数是源自于像素本身特征的谱信息核与周边区域特征（如均值、标准差等）的上下文信息核的组合。之后，Jón Atli Benediktsson 等提出一种替代方法，他们使用数学形态学表示像素间的空间关系。Li Jun 等学者进一步将多核合成扩展为多项逻辑回归分析，从不同的角度提出了多种改进方案 [270]，其核心思想都是将不同类型信息所构造的多个 Mercer 核进行合成实现信息融合。这些方法建立的理论基础就是定理 5.1 和定理 5.2，即任意正定核函数的线性组合、点乘、积分、求极限，其结果仍是正定核函数。受这些文献的启发，我们考虑是否可以将单演信号诸分量（局部幅度、局部相位、局部方位）通过非线性映射投影到抽象的特征空间，然后借助多核合成学习将各分量所构成的 Mercer 核组合起来进行 SAR 图像目标识别 [271]。

对于给定的 K 个类别的训练样本 $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,n_k}], k = 1, \dots, K$ ，为了方便标示格莱姆矩阵的计算，我们对训练样本 $\mathbf{X} \in \mathbb{R}^{m \times n}, n = \sum_{k=1}^K n_k$ 的标号重新按单下标的顺序进行排列，即 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ ，样本本身及其顺序保持不变，只是变换一种标记方式。假设存在一组非线性映射，可将单演信号的诸分量分别投影到 Hilbert 特征空间，然后将每个分量所构成的 Hilbert 空间通过合成核学习组合起来构成全局 Hilbert 空间，使用全局 Hilbert 空间的像构造过完备字典，使用字典原子对未知样本（像）进行线性表示，形成核回归模型，通过求解最优回归系数判决未知样本的类别。算法的具体实现可以划分成四个步骤：

1. 单演信号表示：对每个样本的单演信号多尺度表示分量进行均匀下采样、沿列排列、能量归一化等处理，生成特征向量，详见论文第四章。
2. Mercer 核计算：构造非线性映射将单演信号特征向量投影到再生核 Hilbert 空间，计算各分量的格莱姆核矩阵和未知样本的相似度向量；
3. 多核合成学习：利用分量 Mercer 核定义整体核函数，将单演信号分量构成的子 Hilbert 核空间组合成全局 Hilbert 空间，构建线性表示模型，求解回归系数；
4. 最终判决输出：根据最优化求解得到的表示系数重构未知样本的像，利用每个类别重构误差进行判决。

对于单演信号多尺度表示论文第四章已经进行了详细描述，Mercer 核函数以及 Hilbert 空间中的线性表示在第 5.1 节和第 5.2 节分别进行了介绍。剩余的问题就是如何使用合成核学习实现 Hilbert 空间的信息融合。为了将单演信号中所蕴含的原始信号的信息（如能量信息、结构信息、几何信息等）在 Hilbert 空间整合起

来，共同用于 SAR 图像目标识别，我们提出一种基于多核合成学习的分类框架，将提出的框架应用于单演信号尺度空间表示，简明流程如图 5.3 所示，并在该框架思路的指导下形成单演信号特征组合和格莱姆矩阵组合两种具体方案，分别反映“先组合后映射”和“先映射后组合”两种不同的信息融合思路。

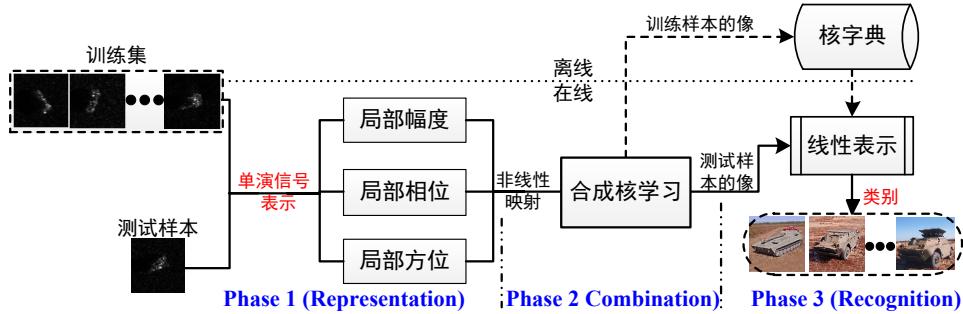


图 5.3 基于多核合成学习的分类学习，算法流程划分为三个连续步骤：(1) 表示，分别对训练样本和测试样本进行单演信号多尺度表示；(2) 组合，使用非线性映射将单演信号诸分量投影到 Hilbert 空间，将定义在每个子 Hilbert 空间的正定核函数进行组合；(3) 线性表示，使用训练样本的像组成核空间字典，根据该字典对测试样本的像进行稀疏线性表示，根据重构误差判决目标类型。

首先，对训练样本做单演信号多尺度表示，以样本 \mathbf{x}_k 为例，将其单演信号尺度空间上的诸分量做均匀下采样、能量归一化、逐列排列生成三个特征向量 $\chi_A(\mathbf{x}_k)$ 、 $\chi_\varphi(\mathbf{x}_k)$ 和 $\chi_\theta(\mathbf{x}_k)$ ，具体方法参见第三章 4.1.2 节的公式 (4.5)。然后，使用一组非线性映射 $\phi_A(\cdot)$ 、 $\phi_\varphi(\cdot)$ 和 $\phi_\theta(\cdot)$ 分别将单演信号特征向量映射到 Hilbert 空间中，对应的子 Hilbert 特征空间分别记为 \mathcal{H}_A 、 \mathcal{H}_φ 和 \mathcal{H}_θ 。最后使用合成核学习将子 Hilbert 特征空间合并起来构成全局 Hilbert 空间。在全局 Hilbert 空间中构造线性系统，使用训练样本的像对未知样本的像 ($\phi(\mathbf{y})$) 进行线性表示，

$$\phi(\chi(\mathbf{y})) = \phi(\chi(\mathbf{x}_1))\alpha_1 + \cdots + \phi(\chi(\mathbf{x}_n))\alpha_n = \chi(\mathbf{X})\alpha \quad (5.14)$$

其中 $\phi(\cdot)$ 是投影于全局再生核 Hilbert 空间的非线性映射，它和子再生核 Hilbert 空间映射 ϕ_A 、 ϕ_φ 和 ϕ_θ 之间的关系将分别在 5.3.1 节和 5.3.2 节中进行介绍。对线性系统 (5.14) 施加稀疏约束，

$$\min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \|\phi(\chi(\mathbf{y})) - \chi(\mathbf{X})\alpha\|_2 \leq \epsilon \quad (5.15)$$

或者

$$\min_{\alpha} \|\alpha\|_2 \quad s.t. \quad \|\phi(\chi(\mathbf{y})) - \chi(\mathbf{X})\alpha\|_2 \leq \epsilon \quad (5.16)$$

公式 (5.15) 和公式 (5.16) 是核回归模型回归系数的两种求解方式，前者是采用稀疏约束的目标函数，后者是 ℓ_2 范数约束，利用最优化算法即可得到唯一的回归系数⁹。然而由于映射 $\phi(\cdot)$ 是构造出来的辅助函数，具体形式未知，因而需要定义正定核函数实现 Hilbert 空间的像的内积，主要涉及字典原子的像生成的格莱姆矩阵 $\langle \phi(\chi(\mathbf{X})), \phi(\chi(\mathbf{X})) \rangle \in \mathbb{R}^{n \times n}$ ，以及字典原子的像和未知样本的像生成的核相似向量 $\langle \phi(\chi(\mathbf{X})), \phi(\chi(\mathbf{y})) \rangle \in \mathbb{R}^n$ 。围绕着如何由子再生核 Hilbert 空间中像与像间的内积导出全局 Hilbert 空间中像与像间的内积，我们给出两个具体实现方案：

5.3.1 “先组合再映射” —— 单演信号特征组合

要根据子 Hilbert 空间中像与像间的内积推导全局 Hilbert 空间中像与像间的内积，最直接、最易于操作的方法就是将单演信号多尺度表示生成的三种特征向量组合成一个增强的特征向量，类似于论文第四章第 4.1.1 节的特征级融合，然后构造一个非线性映射 $\phi(\cdot)$ 将增强的单演信号特征向量直接投影到再生核 Hilbert 空间 (\mathcal{H}) 中，继而定义再生核 Hilbert 空间的正定核函数为：

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\chi(\mathbf{x}_i)), \phi(\chi(\mathbf{x}_j)) \rangle_{\mathcal{H}} \quad (5.17)$$

其中 $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ 表示是定义在全局 Hilbert 空间 \mathcal{H} 上的内积运算。然后将定义的核函数带入前文介绍的核回归模型框架，求解最优回归系数，计算每个类别的重构误差，据此进行判决。这个方案反映的是“先组合再映射”的思路，将不同类型的信息首先在原始数据空间进行合并，将合并的结果投影到再生核 Hilbert 空间。

5.3.2 “先映射再组合” —— 格莱姆矩阵组合

单演信号特征组合是先在数据空间中对不同信息进行组合，然后借助非线性映射将其投影到再生核 Hilbert 空间。与这种先组合后投影的思路不同，本方案提出另一种先映射后组合的方法，即格莱姆矩阵组合。格莱姆矩阵组合是指借助不同的非线性映射将不同的单演信号特征向量从原始的数据空间分别投影到不同的子再生核 Hilbert 空间中，然后将子再生核 Hilbert 空间组合起来构成全局再生核 Hilbert 空间。若以 $\phi_A(\cdot)$ 、 $\phi_\varphi(\cdot)$ 和 $\phi_\theta(\cdot)$ 表示单演信号数据空间（记为 $\mathcal{E}_A, \mathcal{E}_\varphi, \mathcal{E}_\theta$ ）到再生核 Hilbert 空间（记为 $\mathcal{H}_A, \mathcal{H}_\varphi, \mathcal{H}_\theta$ ）的非线性映射，

$$\phi_j : \mathcal{E}_j \mapsto \mathcal{H}_j, j = \{A, \varphi, \theta\} \quad (5.18)$$

⁹ ℓ_2 范数约束的策略实质就是核岭回归分析，详细描述参见附录 C 所述。

它们将幅度、相位、方位特征向量投影到子 Hilbert 空间 \mathcal{H}_A 、 \mathcal{H}_φ 和 \mathcal{H}_θ ，然后利用这三个非线性映射的组合定义个局映射，

$$\phi(\mathbf{x}) = \{\phi_A(\chi_A(\mathbf{x})), \phi_\varphi(\chi_\varphi(\mathbf{x})), \phi_\theta(\mathbf{x})\} \quad (5.19)$$

(5.19) 式将三种非线性映射等概率组合，若给定先验知识，还可以施加相应地权值系数，如

$$\phi(\mathbf{x}) = \{\beta_A \phi_A(\chi_A(\mathbf{x})), \beta_\varphi \phi_\varphi(\chi_\varphi(\mathbf{x})), \beta_\theta \phi_\theta(\mathbf{x})\}$$

其中 β_A 、 β_φ 和 β_θ 是权重系数。对应的全局 Hilbert 空间的核函数为

$$\begin{aligned} \kappa(\mathbf{x}_i, \mathbf{x}_j) &= \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle \\ &= \langle \{\phi_A(\chi_A(\mathbf{x}_i)), \phi_\varphi(\chi_\varphi(\mathbf{x}_i)), \phi_\theta(\chi_\theta(\mathbf{x}_i))\}, \{\phi_A(\chi_A(\mathbf{x}_j)), \phi_\varphi(\chi_\varphi(\mathbf{x}_j)), \phi_\theta(\chi_\theta(\mathbf{x}_j))\} \rangle_{\mathcal{H}} \\ &= \langle \phi_A(\chi_A(\mathbf{x}_i)), \phi_A(\chi_A(\mathbf{x}_j)) \rangle_{\mathcal{H}_A} + \langle \phi_\varphi(\chi_\varphi(\mathbf{x}_i)), \phi_\varphi(\chi_\varphi(\mathbf{x}_j)) \rangle_{\mathcal{H}_\varphi} + \langle \phi_\theta(\chi_\theta(\mathbf{x}_i)), \phi_\theta(\chi_\theta(\mathbf{x}_j)) \rangle_{\mathcal{H}_\theta} \\ &= \kappa_A(\chi_A(\mathbf{x}_i), \chi_A(\mathbf{x}_j)) + \kappa_\varphi(\chi_\varphi(\mathbf{x}_i), \chi_\varphi(\mathbf{x}_j)) + \kappa_\theta(\chi_\theta(\mathbf{x}_i), \chi_\theta(\mathbf{x}_j)) \end{aligned} \quad (5.20)$$

式中 $\langle \cdot, \cdot \rangle_{\mathcal{H}_A}$ 、 $\langle \cdot, \cdot \rangle_{\mathcal{H}_\varphi}$ 和 $\langle \cdot, \cdot \rangle_{\mathcal{H}_\theta}$ 分别是定义在子 Hilbert 空间的内积， $\kappa_A(\cdot, \cdot)$ 、 $\kappa_\varphi(\cdot, \cdot)$ 和 $\kappa_\theta(\cdot, \cdot)$ 是与之对应的核函数。

将子 Hilbert 空间中的核函数进行组合，论文推导出全局 Hilbert 空间的核函数 $\mathcal{H}_{\mathcal{G}'}$ ，将全局核函数代入 $\langle \phi(\chi(\mathbf{X})), \phi(\chi(\mathbf{X})) \rangle_{\mathcal{G}'}$ 和 $\langle \phi(\chi(\mathbf{X})), \phi(\chi(\mathbf{y})) \rangle_{\mathcal{G}'}$ 计算格莱姆核矩阵和未知样本的相似度向量，然后通过求解 (5.5) 或 (5.9) 最优回归系数，即可判决未知样本的类别。为了直观地展示多核合成学习，图 5.4 给出一组 BMP2 分类识别实例，分别给出两种 Mercer 核函数的构造示意图，及其分类流程图，其中 *Stacked Features* 是指特征组合的合并方法，*Summation Kernel* 是指格莱姆矩阵组合的合并方法。

5.4 实验验证

本章论文提出多核合成学习，实现单演信号分量在 Hilbert 空间的信息融合，实现目标识别。接下来将借助 MSTAR 实测数据对算法的有效性进行验证。与前文的实验方案类似，先标准实验条件，后扩展工作条件，包括目标物理配置变化、雷达俯仰角度变化以及均匀随机噪声污染，逐渐增加实验难度，考察多核合成学习对于扩展工作条件的耐受性。单演信号尺度空间的构成方法与前面章节相同，采用 log-Gabor 引导的带通滤波器组生成单演信号尺度空间，生成的分量系数矩阵执行采样因子为 $\frac{1}{8}$ 的降采样，对应的采样后的图像尺寸为 10×10 ，然后利用非线性矩阵将生成的表示系数投影到 Hilbert 空间。实验中的主要基准比较算法包括

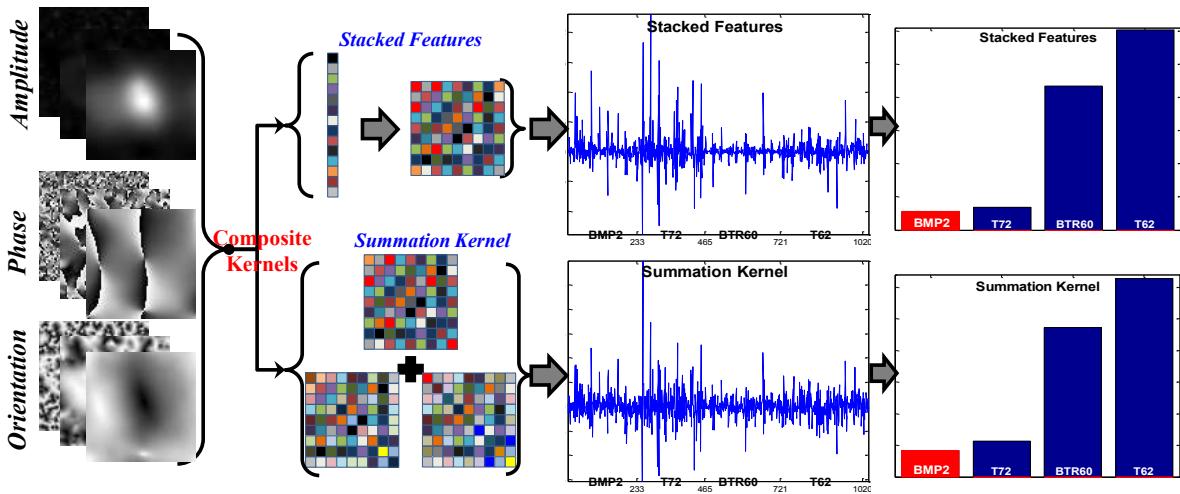


图 5.4 多核合成学习的分类示例，图中第一列给出的是未知样本的单演信号多尺度表示，第二列是两种合成核学习 Gram 核矩阵的构成示意图，第三列是最优表示系数，是根据 ℓ_2 范数最小化策略生成的最优解（详见 (5.9)），第四列是每个类别的重构误差。

第 4 章提出的空域信息融合算法，详见表 5.1，其中使用高斯核函数求解核线性回归。

表 5.1 再生核 Hilbert 空间分类学习实验基准算法。

缩写	算法名称	输入特征
KSVM	Support Vector Machine with nonlinear Kernel [65]	原始强度值
KSRC	Kernel Sparse Representation-based Classification [97, 264]	原始强度值
KLRC	Kernel Linear Representation-based Classification [267]	原始强度值
MSRC	Feature-level Fusion (第 4.1.1 节)	增强的单演特征
MSRC ^{BY}	Decision-level Fusion via Bayes (第 4.1.2 节)	单演信号特征
MSRC ^{DS}	Decision-level Fusion via DS Theory (第 4.1.2 节)	单演信号特征
CKLR1	Composite Kernel with Feature Combination (第 5.3.1 节)	单演信号特征
CKLR2	Composite Kernel with Gram Matrix Combination (第 5.3.2 节)	单演信号特征

5.4.1 标准工作条件

首先考虑的是标准工作条件的目标识别问题。仍然采用的第二章第 2.4.3.1 节的表 2.2 中的所有 10 类目标。标准型号（针对多变体目标 BMP2 和 T72 而言，标准型号为 SN_9563 和 SN_132）在俯仰角为 17° 的工作条件的成像样本用作训练算法，而所有型号在俯仰角为 15° 的工作条件的成像样本用于测试，实验结果如图 5.5。根据实验结果可以得出如下结论，

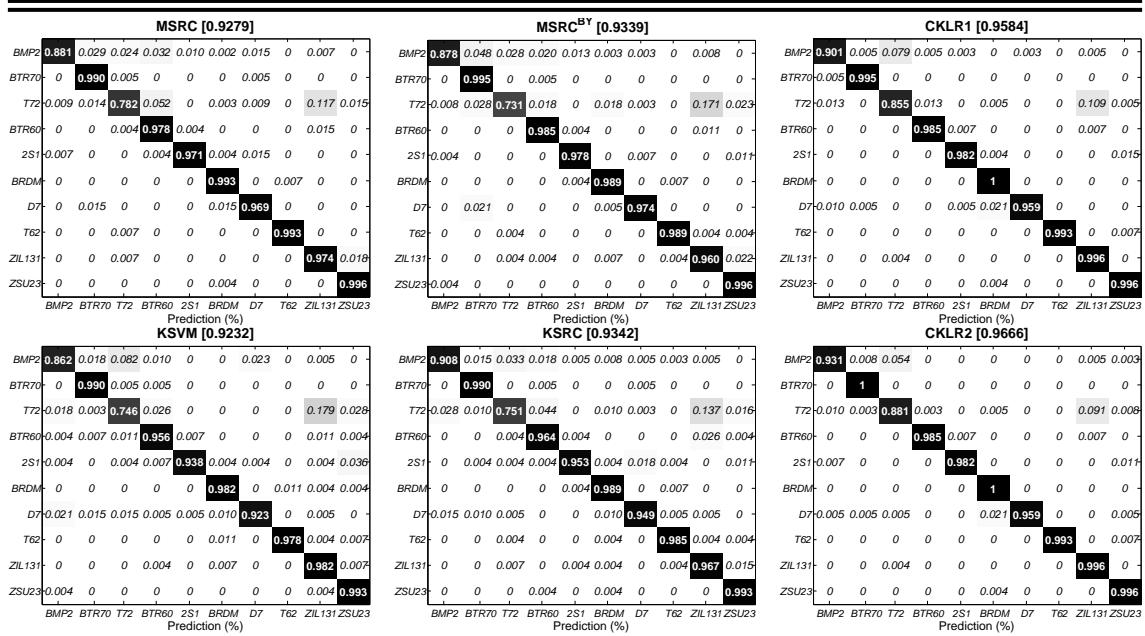


图 5.5 多核合成学习十类目标标准工作条件的识别实验，每幅子图表示分类算法得出的混淆矩阵，其列座标标识的是测试样本的真实类别，行座标标识的是各算法估计的类别，对角线元素表示每类目标的正确识别率，非对角线元素是误分率，方括号中给出的是整体识别率。

从整体识别率看：多核合成学习比空域信息融合性能有显著的改进。实验中，单演信号特征组合和格莱姆矩阵组合分别取得了 0.9584 和 0.9666 的分类识别率，比特征级融合方法 (0.9292) 高出 2.92% 和 3.74%，比决策级融合方法高出 2.45% 和 3.27%。这是因为本章的方法使用非线性映射将数据空间投影到 Hilbert 空间，改进数据集类别间的线性可分性。因此，Hilbert 空间的信息融合效率要明显优于图像域的信息融合。

从核映射算法比较来看：多核合成学习的表现优于传统的核回归学习。两种合成核学习 CKLR1 和 CKLR2 的分类正确率为 0.9584 和 0.9666，与之相比，非线性核支持向量机 KSVM 的正确率为 0.9232，核稀疏表示分类 KSRC 的识别率为 0.9342，多核合成学习比非线性核支持向量机高出 3.52% 和 4.34%，比核稀疏表示模型高出 2.42% 和 3.24%。这是因为多核合成学习能将单演信号所包含的不同类型信息在再生核 Hilbert 空间进行融合。

从目标的混淆矩阵来看：BMP2 和 T72 两个多型号目标的正确识别率明显低于其余类别，它们的最低识别率仅有 0.862 和 0.731，这是因为这两类目标存在多个物理变体，详见第二章表 2.2 所示，17° 俯仰角的标准型号目标用于训练，15° 俯仰角的所有型号目标进行测试，即测试样本和训练样本的物理配置与俯仰角同时发生变化。本章提出的多核合成学习对于 BMP2 和 T72 的分类正确率分别为 0.901、0.931 和 0.855、0.881，优于其它算法。

5.4.2 扩展工作条件

为了验证多核合成学习受实际因素影响的表现，接下来组织扩展工作条件的目标识别实验。仍然沿用第四章第 4.2.2 节的实验方案，从目标配置变化、雷达俯仰角度变化和随机噪声污染三个角度进行考察。

5.4.2.1 目标配置变化

为了验证本文算法对目标配置变化的稳健性，组织 BMP2、T72、BTR60 和 T62 四类目标识别实验。标准型号在 17° 雷达俯仰角的成像样本用于训练，其余的型号在 15° 雷达俯仰角的成像样本用于测试，用于测试的型号不包括在训练集中，具体设置见第四章第 4.2.2 节的表 4.5。实验结果如图 5.6 所示。下面分别从整体识别率、每个类别识别率以及核映射比较进行分析：

1. 整体分类正确率：基于多核合成的分类学习框架 CKLR1 和 CKLR2 的整体正确率分别为 0.9325、0.9349，基于图像域信息融合的分类框架 MSRC、 $MSRC^{BY}$ 、 $MSRC^{DS}$ 的整体正确率分别为 0.8748、0.8804 和 0.8925，多核合成学习比空域信息框架的整体正确率高 5.77%、5.69%、4.48% 和 6.01%、5.93%、4.72%。
2. 类别分类正确率：每个类别的分类正确率差异主要集中于 BMP2 和 T72 两个多型号目标，CKLR1 的 BMP2 类别正确率最优，0.9719，对应 MSRC、 $MSRC^{BY}$ 、 $MSRC^{DS}$ 的正确率分别为 0.9286、0.9235、0.9490，CKLR2 的 T72 类别正确率最高，0.8523，对应 MSRC、 $MSRC^{BY}$ 、 $MSRC^{DS}$ 的正确率分别为 0.6736、0.7021、0.7098。
3. 核映射相互比较：与多核分别学习比较的核映射方法包括非线性核支持向量机 KSVM 和核稀疏表示分类，它们的整体正确率为 0.8796 和 0.9045，比 CKLR1 和 CKLR2 分别低 5.29%、5.21%、4.0% 和 5.53%、5.45%、4.24%。每个类别的分类正确率也低于多核合成学习。

结果表明多核合成学习的分类框架能够应对目标配置变化带来的不利影响。

5.4.2.2 俯仰角变化

下面的实验用于考察雷达俯仰角变化对于分类算法的影响，包括 2S1、BRDM2 和 ZSU23/4 在不同俯仰角的成像样本，其中 BRDM2 和 ZSU23/4 包含连接体变体，如炮管的高低起落、左右旋转等。实验使用 17° 雷达俯仰角标准目标的成像样本训练分类算法，使用 30° 和 45° 雷达俯仰角连接体目标变体的成像数据进行测试，结果如图 5.7 所示。

SRC [0.8692]					MSRC [0.8748]				
Target	BMP2	T72	BTR60	T62	Target	BMP2	T72	BTR60	T62
BMP2	0.9490	0.0306	0.0077	0.0128	BMP2	0.9286	0.0230	0.0255	0.0230
T72	0.0596	0.6658	0.0052	0.2694	T72	0.0415	0.6736	0.0259	0.2591
BTR60	0.0051	0.0	0.9846	0.0103	BTR60	0.0	0.0	0.9897	0.0103
T62	0.0073	0.0183	0.0147	0.9597	T62	0.0	0.0	0.0	1.0

MSRC ^{BY} [0.8804]					MSRC ^{DS} [0.8925]				
Target	BMP2	T72	BTR60	T62	Target	BMP2	T72	BTR60	T62
BMP2	0.9235	0.0357	0.0179	0.0230	BMP2	0.9490	0.0255	0.0102	0.0153
T72	0.0570	0.7021	0.0026	0.2383	T72	0.0492	0.7098	0.0026	0.2383
BTR60	0.0	0.0051	0.9949	0.0	BTR60	0.0	0.0051	0.9949	0.0
T62	0.0037	0.0073	0.0	0.9890	T62	0.0037	0.0073	0.0	0.9890

KSVM [0.8796]					KSRC [0.9045]				
Target	BMP2	T72	BTR60	T62	Target	BMP2	T72	BTR60	T62
BMP2	0.8903	0.0816	0.0230	0.0051	BMP2	0.9464	0.0383	0.0102	0.0051
T72	0.0337	0.7539	0.0	0.2124	T72	0.0363	0.7746	0.0052	0.1839
BTR60	0.0256	0.0205	0.9436	0.0103	BTR60	0.0256	0.0	0.9641	0.0103
T62	0.0037	0.0	0.0	0.9963	T62	0.0	0.0073	0.0073	0.9853

CKLR1 [0.9325]					CKLR2 [0.9349]				
Target	BMP2	T72	BTR60	T62	Target	BMP2	T72	BTR60	T62
BMP2	0.9719	0.0281	0.0	0.0	BMP2	0.9541	0.0408	0.0051	0.0
T72	0.0259	0.8394	0.0026	0.1321	T72	0.0518	0.8523	0.0026	0.0933
BTR60	0.0256	0.0	0.9641	0.0103	BTR60	0.0205	0.0	0.9692	0.0103
T62	0.0	0.0037	0.0	0.9963	T62	0.0	0.0037	0.0	0.9963

图 5.6 目标配置变化对多核合成学习的影响，图中每个子表格是分类算法得出的混淆矩阵，括号中的数字是算法的整体识别率，表格内的数字是分类的混淆矩阵，其中对角线方格的粗体数字表示每类目标的分类正确率。

第一个实验场景指定 30° 俯仰角的成像数据为测试样本，训练样本和测试样本间存在着 13° 的俯仰角度差异。各算法的分类正确率都在 0.95 以上，基于多核合成学习的分类框架的正确率为 0.9829 和 0.9712，特征级融合的决策融合分类的正确率为 0.9829、0.9829 和 0.9839，整体表现都很好¹⁰。

¹⁰MSRC、MSRC^{BY}、MSRC^{DS} 的分类结果见第四章第 4.2.2 节表 4.7。

KSRC

<i>Depression</i>	30° (0.9722)			45° (0.5444)		
<i>Target</i>	2S1	BRDM2	ZSU23/4	2S1	BRDM2	ZSU23/4
2S1	0.9688	0.0313	0.0	0.3795	0.6205	0.0
BRDM2	0.0095	0.9833	0.0071	0.0757	0.9125	0.0118
ZSU23/4	0.0148	0.0222	0.9631	0.1209	0.5853	0.2938

KLRC

<i>Depression</i>	30° (0.9686)			45° (0.5453)		
<i>Target</i>	2S1	BRDM2	ZSU23/4	2S1	BRDM2	ZSU23/4
2S1	0.9826	0.0174	0.0	0.3762	0.6238	0.0
BRDM2	0.0119	0.9810	0.0071	0.0709	0.9149	0.0142
ZSU23/4	0.0222	0.0320	0.9458	0.1209	0.5829	0.2962

CKLR1

<i>Depression</i>	30° (0.9712)			45° (0.7238)		
<i>Target</i>	2S1	BRDM2	ZSU23/4	2S1	BRDM2	ZSU23/4
2S1	0.9896	0.0104	0.0	0.7030	0.2937	0.0033
BRDM2	0.0667	0.9310	0.0024	0.0700	0.7991	0.1300
ZSU23/4	0.0	0.0	1.0	0.0900	0.2464	0.6635

CKLR2

<i>Depression</i>	30° (0.9829)			45° (0.7456)		
<i>Target</i>	2S1	BRDM2	ZSU23/4	2S1	BRDM2	ZSU23/4
2S1	0.9965	0.0035	0.0	0.7393	0.1650	0.0957
BRDM2	0.0238	0.9571	0.0190	0.1182	0.5816	0.3002
ZSU23/4	0.0	0.0	1.0	0.0213	0.0640	0.9147

图 5.7 俯仰角变化对多核合成学习的影响，表格中分别给出了雷达俯仰角为 30° 和 45° 的成像数据作为测试图像的分类结果，括号内粗体数字是算法的整体正解率，混淆矩阵中对角线方格数字是每个类别的正确率。

第二个实验场景使用 45° 俯仰角的成像作为测试样本，训练样本和测试样本之间存在着 28° 的俯仰角度差异，实验难度骤增。各算法的识别率都有明显地下降，KSRC 和 KLRC 甚至降到了 0.5444 和 0.5453，比前一个实验场景的正确率下降了 42.78 和 42.33。与之相比，本章提出的两种方法的表现依然稳健，取得了 0.7238 和 0.7456 的识别率，优于参考算法。从类别的正确率来看，尽管基于多核合成学习的分类框架 CKLR1、CKLR2 的 BRDM_2 的正确率为 0.7991、0.5816，低于 KSRC 和 KLRC，但 2S1 和 ZSU23/4 两类的正确率远高于基准算法。KSRC

和 KLRC 很容易将 2S1 误分为 BRDM_2，将 ZSU23/4 误分为 BRDM_2，这是因为 KSRC 和 KLRC 容易造成过训练（或是过拟合，Over-fitting）。

5.4.3 随机噪声污染

接下来的实验讨论噪声染污环境的目标识别问题，实验设置与第二章第 2.4.3 节类似，仍以 BMP2、T72、BTR60 和 T62 四类目标的 17° 的 15° 雷达俯仰角的成数数据进行目标识别实验。为了模拟噪声，我们从测试图像中随机选取部分像素（即强度值），将其替换为满足独立同分布的均匀分布的样本，训练样本保持不变。随机噪声污染的示例见第 2 章图 2.24。

表 5.2 噪声污染对多核合成学习的影响，表中列出了受 5 种不同等级噪声污染时的分类正确率，下降程度表示噪声从 0 增加到 20% 算法的正确率降幅。

算法	噪声等级					
	0	5%	10%	15%	20%	下降程度
SRC	0.8708	0.8186	0.7464	0.6091	0.5040	36.68%
SVM	0.8331	0.7841	0.5979	0.4598	0.4550	37.81%
MSRC ^{BY}	0.8804	0.8620	0.8339	0.8082	0.7994	8.10%
KSVM	0.8796	0.8162	0.6276	0.4775	0.4536	42.60%
KSRC	0.9045	0.8451	0.7255	0.5971	0.5185	38.60%
KLRC	0.9061	0.8499	0.7793	0.6734	0.5738	33.23%
CKLR1	0.9325	0.8603	0.8547	0.8330	0.8234	10.91%
CKLR2	0.9349	0.8876	0.8667	0.8451	0.8435	9.14%

表 5.2 和图 5.8 给出了各分类算法受不同程度噪声染污的识别率变化情况。可以看出，当随机噪声污染等级从 0% 逐渐提升到 20%，所有算法的识别率都有不同程度的下降，其中贝叶斯决策级融合方法和基于多核合成学习的分类框架下降幅度较小，分别为 8.10%、10.91% 和 9.14%，与之相比 SRC、SVM、KSRC、KSVM、KLRC 下降幅度分别为 36.68%、37.81%、42.60%、38.60%、33.23%，说明随机噪声污染极大地影响了基准算法的分类精度。当噪声等提升到 30% 时，基于多核合成学习的分类 CKLR1 和 CKLR2 仍然取得了 0.8234 和 0.8435 的正确率，比 SRC、SVM、KSRC、KSVM、KLRC 分别高出 31.94%、36.84%、30.49%、36.90%、24.96% 和 33.95%、38.85%、32.50%、38.91%、26.97%。从实验结果可以得出多核合成学习的分类能够在一定程度上增强噪声污染的耐受性。

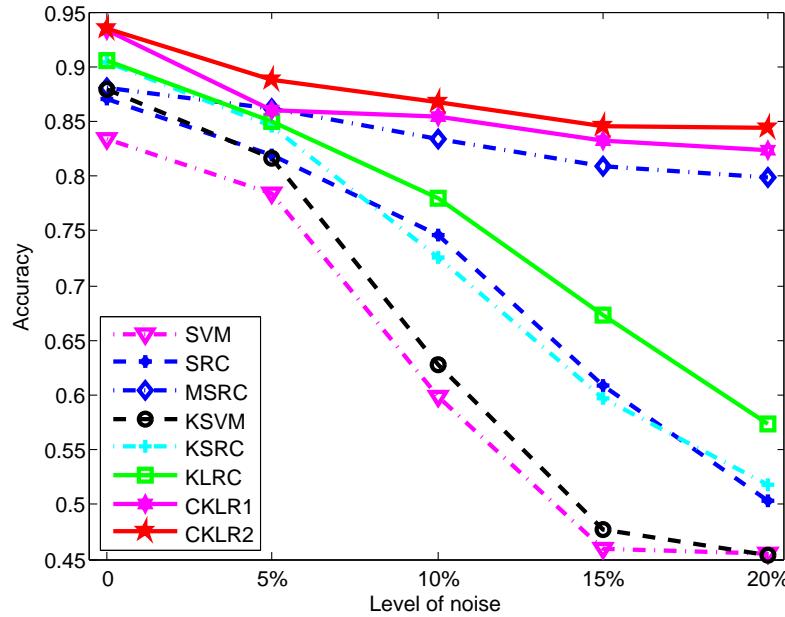


图 5.8 多核合成学习受噪声污染的识别率变化曲线。

5.5 本章小结

本章论文主要提出一种利用再生核 Hilbert 空间多核合成学习实现单演信号尺度空间信息融合的分类框架。首先，以一个简单直观的示例阐述了核空间映射理论的定义，在此基础上引出 Mercer 核理论和再生核 Hilbert 空间的概念，梳理列举出一些常用的核函数，重点分析了正定核函数的主要特性，为下文中的多核合成学习进行铺垫；然后，介绍了 Hilbert 空间中的线性表示模型以及分类算法，并通过正定核函数将再生核 Hilbert 空间像与像的内积与原始数据空间元素与元素的内积联系起来，实现再生核 Hilbert 空间核回归模型的求解；最后，借助多核合成学习，将单演信号尺度空间蕴含的不同类型的信息在再生核 Hilbert 空间进行融合，形成两种具体实现算法：

- 先组合后映射：将单演信号多尺度表示生成的分量在数据空间进行合并，生成特征向量，借助一个非线性映射将其投影到再生核 Hilbert 空间，构成核回归模型；
- 先映射后组合：借助多个非线性映射将单演信号多尺度表示生成的多种分量分别映射到再生核 Hilbert 空间，根据核函数的累加是合法的核函数构造全局再生核 Hilbert 空间，生成核回归模型；

通过求解全局再生核 Hilbert 空间核回归模型，实现未知样本的分类识别。通过大量的实验比较得出：多核合成学习的分类框架能够有效应对目标配置变化、雷达俯仰角变化、随机噪声污染等因素带来的不利影响。

第六章 基于单演信号流形学习的 SAR 图像目标分类

论文前面章节讨论了基于线性子空间的分类问题，内容包括图像域多元回归分析与再生核 Hilbert 空间核回归分析，这些方法使用线性子空间样本构造过完备冗余字典，利用字典原子对未知样本进行线性表示，生成回归模型，借助 ℓ_1 范数最小化的最优化策略获取最优回归系数，使用回归系数重构未知样本，根据重构误差进行归属决策。

机器学习和计算机视觉的最新研究进展表明，基于非欧几里德的子空间学习更有利于不同领域的多类分类问题求解。单个子空间通常具有线性空间的一般特性，而一族线性子空间的集合则与它本身完全不同，它即被称作是流形。流形是一种特殊的李群（Lie Groups）均匀空间，具有十分奇妙的数学特性，尽管该理论提出已经有很长时间，它在信息科学领域的应用研究也是近些年才开展。比较典型的应用包括：信号处理与控制、数字最优化、机器视觉等 [272–279]。

与前文不同，本章论文将单演信号分解生成的系数矩阵整体看作是非欧氏几何空间元素，借助数学技巧对单演信号系数矩阵进行处理，使其满足一定的约束条件，然后将其整体看作是特定流形的元素，通过定义在流形的距离度量来衡量元素的相似性，实现对未知样本的分类识别；另一方面，借助黎曼正定核函数将单演信号生成的线性子空间集嵌入到再生核 Hilbert 空间，处理数据内部的非线性几何结构，在 Hilbert 空间构造核回归模型，通过求解最优回归系数计算每个类别的残差，推断未知样本的类别属性。本章内容主要包括：1) 黎曼流形的基础理论；2) 区域协方差描述子；3) Grassmann 流形及其距离测度；4) 方向可控 Riesz 小波框架；5) 黎曼流形的 SAR 图像目标分类识别方法。

6.1 微分几何基础理论

6.1.1 黎曼流形

6.1.1.1 黎曼流形背景知识

流形是一类重要的拓扑空间，它是欧氏空间中曲线、曲面的推广，其显著特征就是局部为欧氏空间。流形的概念是著名数学家高斯（Gauss）首先提出，用于描述地球表面的绘图过程并予以推广、再进一步抽象而来。因而，流形的定义反映出它的显著特征：它在局部是欧氏的，对于流形的任意点，必有一个邻域与欧氏空间的某个开集同胚，因而可以使用局部坐标系对其进行讨论。流形也可直观地想象成为是由一块块“欧氏空间”粘贴而成 [280]。

定义 6.1(可微函数): n 维空间 \mathbb{R}^n 的任意开子集 U , 若映射 $\alpha : U \mapsto \mathbb{R}$ 直到 k (非负整数) 阶的所有偏导数在 U 的每一点都存在且连续, 称其为 C^k 类可微函数。

若 α 是 C^0 类, 它是连续的; 若 α 在 U 中每一点都有任意阶的连续偏导数, 称 α 是 C^∞ 类可微。与可微函数类似, 若映射 $f : U \mapsto \mathbb{R}^p$ 的每个分量函数 $f_i, i = 1, \dots, p : U \mapsto \mathbb{R}$ 是 C^k , 称 f 是 C^k 类可微映射。

定义 6.2(同胚流形): 设 M 是任意的 Hausdorff 拓扑空间, 若 M 中的任意 p 都存在它的开邻域和映射 $\phi : U \mapsto \mathbb{R}^m$, 使得 ϕ 是 U 到 \mathbb{R}^m 的开子集 $\phi(U)$ 的同胚映射 (一对一, 连续双射), 称 M 是个 m 维的拓扑流形。 (U, ϕ) 称为流形 M 的一个局部坐标系或局部坐标卡, U 称为局部坐标域, ϕ 称为坐标映射。

Hausdorff 拓扑空间是指空间中的每个点都有与之不相交的若干个邻域。该性质为微分流形的定义奠定了基础, 它保证任何收敛序列会收敛至唯一的极限点。

定义 6.3: 设 M 是拓扑流形, 若 M 的局部坐标卡集 $D = \{(U_\alpha, \phi_\alpha) | \alpha \in \Lambda\}$ 满足:

- D 中各坐标卡的定义域覆盖流形 M , 即 $\bigcup_{\alpha \in \Lambda} U_\alpha = M$;
- D 中任意两个局部坐标卡是 C^r 相容;
- 与 D 中各局部坐标卡 C^r 相容的任何局部坐标卡必属于 D ;

称 D 是 M 的一个 C^r 微分结构, M 连同给定的 C^r 微分结构称为 C^r 微分流形。

论文通篇讨论 C^∞ 流形, 微分流形特指的是这类光滑流形, 简称为流形。微分流形的任意曲线可以定义导数。流形的任意点 $\mathbf{X} \in M$ 的导数位于向量空间 $T_x M$ 内, $T_x M$ 是由所有经过 \mathbf{X} 的曲线的切向量所组成的向量空间。

定义 6.4: 黎曼流形是在其每个切向空间均赋予内积运算, 且该运算在点与点之间变化平滑的微分流形。流形内积可诱导切向空间范数 $\|\mathbf{y}\|_x^2 = \langle \mathbf{y}, \mathbf{y} \rangle$ 。

连接流形任意两点间的所有弧长中长度最短的曲线称测地线, 两点间的距离 $d(\mathbf{X}, \mathbf{Y})$ 定义为测地线的长度, 即测地距离。直观的理解, 测地线类似于欧氏空间 \mathbb{R}^d 的直线。对于流形的任意点 $\mathbf{X} \in M$, 它的切向空间 $T_x M$ 是由经过 \mathbf{X} 的所有曲线的切向量组成, 是标准的欧氏空间。切向空间必存在唯一的以某个 $\Delta \in T_x M$ 为起点的测地线。定义在微分流形的指数算子和对数算子将流形与经过流形的点 $\mathbf{X} \in M$ 的切向空间联系起来 (见图 6.1 所示)。

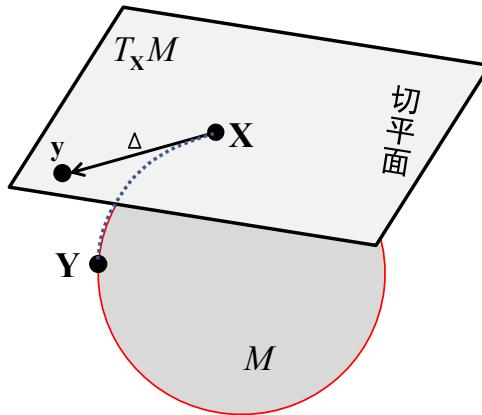


图 6.1 黎曼流形任意点 $\mathbf{X} \in \mathcal{M}$ 的切向空间 $T_{\mathbf{X}}\mathcal{M}$ 。通过矩阵对数算子 $\log(\cdot)$ 将流形的点转换为它的切矢量 $\Delta = \log_{\mathbf{X}}(\mathbf{Y})$, 同样切向空间 $T_{\mathbf{X}}\mathcal{M}$ 的切矢量可通过矩阵指数算子映射到流形 \mathcal{M} , $\exp_{\mathbf{X}}(\Delta) = \mathbf{Y}$ 。图中的虚线就是以 \mathbf{X} 为起点到 \mathbf{Y} 为终点的测地线。

指数算子 $\exp_{\mathbf{X}}(\cdot) : T_{\mathbf{X}}\mathcal{M} \mapsto \mathcal{M}$ 将切向量 $\Delta \in T_{\mathbf{X}}\mathcal{M}$ 映射成流形中的点 $\mathbf{Y} \in \mathcal{M}$ 。指数映射的性质确保切向量 Δ 的长度与流形元素 \mathbf{X} 和 \mathbf{Y} 之间的测地距离相等, 故测地线的距离为 $d(\mathbf{X}, \exp_{\mathbf{X}}(\Delta)) = \|\Delta\|_{\mathbf{X}}$ 。

对数算子 $\log_{\mathbf{X}}(\cdot) : \mathcal{M} \mapsto T_{\mathbf{X}}\mathcal{M}$ 是指数映射的逆过程 (并非是完全意义上的逆映射), 它将流形中的点映射到切向空间中去。从其定义可知指数算子 $\exp_{\mathbf{X}}(\cdot)$ 仅仅是在 \mathbf{X} 的邻域内是一一映射, 因此, $\log_{\mathbf{X}}(\cdot)$ 也只是唯一地定义在 \mathbf{X} 的邻域内。

对于任意的 $\mathbf{Y} \in \mathcal{M}$, 存在多个 $\Delta \in T_{\mathbf{X}}\mathcal{M}$, 使得 $\mathbf{Y} = \exp_{\mathbf{X}}(\Delta)$, 而 $\log_{\mathbf{X}}(\mathbf{Y})$ 也只是范数最小的某个切向量。因而, 这两种算子依赖于流形与其切向空间相切的流形的点 \mathbf{X} , 随着 \mathbf{X} 变化而变化。

6.1.1.2 对称正定矩阵空间

对于 $d \times d$ 维的对称正定矩阵所构成的空间 Sym_d^+ , 任意两个元素之和可能是非正定矩阵, 任意元素与实数相乘其结果也并不一定是对称正定矩阵, 因而, 对称正定矩阵空间对于加法和数乘两种运算结构并不封闭, 它无法构成欧氏空间 (或向量空间)。若给对称正定矩阵空间赋予黎曼度量, 就构成黎曼流形。由黎曼度量所诱导的测地距离就是对称正定矩阵空间中的任意两个矩阵之间的不相似性的衡量, 显然, 这种衡量方式要比传统的欧氏距离更为合理。

公理 6.1: 任意的实函数 $d : \Omega \times \Omega \mapsto \mathbb{R}$ 被称作度量当且仅当它满足:

非负性: $d(x_1, x_2) \geq 0$, 取等号当且仅当 $x_1 = x_2$;

对称性: $d(x_1, x_2) = d(x_2, x_1)$;

三角形不等式: $d(x_1, x_2) + d(x_2, x_3) \geq d(x_1, x_3)$;

其中 $x_1, x_2, x_3 \in \Omega$ 是数据空间 Ω 中的任意元素。

所谓的三角形不等式是指“三角形两边之和大于第三边”。度量就是满足公理 6.1 的距离函数，黎曼度量是指一族能够导出任意点切向空间向量范数的内积。近年来的研究提出了多个用于对称正定矩阵空间的距离函数，用于捕获流形的非线性特性，但并不是所有的距离函数都能由切向空间平滑变化的内积所引出，不一定是测地距离，详见表 6.1 所归纳。

表 6.1 常用的黎曼流形距离函数 ($\|\cdot\|_F$ 是矩阵的 F 范数)。

名称	计算公式	测地距离
对数欧氏距离	$\ \log(\mathbf{X}_1) - \log(\mathbf{X}_2)\ _F$	是
仿射不变距离	$\ \log(\mathbf{X}_1^{-1/2} \mathbf{X}_2 \mathbf{X}_1^{-1/2})\ _F$	是
Cholesky 距离	$\ chol(\mathbf{X}_1) - chol(\mathbf{X}_2)\ _F$	否
冥次欧氏距离	$\frac{1}{\alpha} \ \mathbf{X}_1^\alpha - \mathbf{X}_2^\alpha\ _F$	否
Stein 散度距离	$\sqrt{\log \det(\frac{\mathbf{X}_1 + \mathbf{X}_2}{2}) - \frac{1}{2} \log \det(\mathbf{X}_1 \mathbf{X}_2)}$	否

仿射不变度量 (Affine Invariant Riemannian Metric) 是法国学者 Xavier Pennec 首先提出 [274]，其思想是将黎曼流形的曲率连接起来，将沿着流形曲率对应的测地线长度定义为距离。黎曼流形切向空间的度量为 $\langle \Delta_1, \Delta_2 \rangle_{\mathbf{X}} = tr(\mathbf{X}^{-1/2} \Delta_1 \mathbf{X}^{-1} \Delta_2 \mathbf{X}^{-1/2})$ ，其中 $tr(\cdot)$ 表示矩阵的迹，与黎曼度量关联的指数算子与对数算子分别为

$$\begin{aligned}\exp_{\mathbf{X}}(\Delta) &= \mathbf{X}^{1/2} \exp(\mathbf{X}^{-1/2} \Delta \mathbf{X}^{-1/2}) \mathbf{X}^{1/2} \\ \log_{\mathbf{X}}(\mathbf{Y}) &= \mathbf{X}^{1/2} \log(\mathbf{X}^{-1/2} \mathbf{Y} \mathbf{X}^{-1/2}) \mathbf{X}^{1/2}\end{aligned}\tag{6.1}$$

为了与传统的标量运算的标记区别， $\exp(\cdot)$ 和 $\log(\cdot)$ 表示普通的矩阵运算， $\exp_{\mathbf{X}}(\cdot)$ 和 $\log_{\mathbf{X}}(\cdot)$ 表示面向流形的指数算子和对数算子， $\mathbf{X} \in \mathbf{Sym}_d^+$ 。若对称正定矩阵 Σ 的做特征值分解 $\Sigma = \mathbf{U} \Lambda \mathbf{U}^T$ ，其中 \mathbf{U} 是特征向量构成的酉矩阵， Λ 是由特征值组成的对角矩阵，式 (6.1) 可分别简化为

$$\begin{aligned}\exp(\Sigma) &= \sum_{k=0}^{\infty} \frac{\Sigma^k}{k!} = \mathbf{U} \exp(\Lambda) \mathbf{U}^T \\ \log(\Sigma) &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} (\Sigma - \mathbf{I})^k = \mathbf{U} \log(\Lambda) \mathbf{U}^T\end{aligned}\tag{6.2}$$

显然，只有当 Σ 的所有特征值均为正数时，对数算子才有意义。对于任意的对称正定矩阵 $\mathbf{X}, \mathbf{Y} \in \mathbf{Sym}_d^+$ ，它们的仿射不变黎曼度量为

$$\begin{aligned}d_{\mathcal{R}}^2(\mathbf{X}, \mathbf{Y}) &= \langle \log_{\mathbf{X}}(\mathbf{Y}), \log_{\mathbf{Y}}(\mathbf{X}) \rangle_{\mathbf{X}} \\ &= tr(\log^2(\mathbf{X}^{-1/2} \mathbf{Y} \mathbf{X}^{-1/2})) = \|\log(\mathbf{X}^{-1/2} \mathbf{Y} \mathbf{X}^{-1/2})\|_F^2\end{aligned}\tag{6.3}$$

对数欧氏黎曼度量 (Log-Euclidean Riemannian Metric) 是法国 INRIA 的著名学者 Vincent Arsigny 提出的与仿射不变度量有密切关系的黎曼度量 [273]，其目的是在保持仿射不变度量仿射不变 $d(\mathbf{AXB}, \mathbf{AYB}) = d(\mathbf{X}, \mathbf{Y})$ 的前提下减小计算复杂度。通过引入对数欧氏映射 $\log(\cdot) : \mathbf{Sym}_d^+ \mapsto \mathbf{Sym}_d$ 弱化对称下定矩阵空间的正定约束，然后定义对数欧氏黎曼度量为

$$d_{le}^2(\mathbf{X}, \mathbf{Y}) = \|\log(\mathbf{X}) - \log(\mathbf{Y})\|_F^2 \quad (6.4)$$

对数欧氏度量通过将对称正定矩阵映射到扁平的黎曼空间 (零曲率)，然后再使用传统的欧氏距离进行衡量，极大地降低了仿射不变度量的计算复杂度，且保留了它的部分特性 (如对相似变形的不变性)。

Cholesky 距离是 Wang Zhizhou 等提出的一种基于矩阵 Cholesky 分解¹的距离函数 [272]。若对称正定矩阵 \mathbf{X}, \mathbf{Y} 的 Cholesky 分解为 $\mathbf{X} = L_x L_x^T, \mathbf{Y} = L_y L_y^T$ ，其中 L_x 和 L_y 是下三角矩阵， \mathbf{X}, \mathbf{Y} 的 Cholesky 距离的定义为

$$d_{chol}^2(\mathbf{X}, \mathbf{Y}) = \|L_x - L_y\|_F^2$$

KL 散度距离 (Jeffrey's Kullback-Leibler Divergence) 是根据信息论的 f 散度 (f -divergence) 的概率函数定义的。 f 散度不满足对称性，需要对其作对称化处理。M. Moakher 和 P. Batcher 将对称正定矩阵看作是零均值高斯分布的方差，并使用对称化 KL 散度作为不同分布之间的距离度量 [281]

$$d_{kl}^2(\mathbf{X}, \mathbf{Y}) = \frac{1}{2} \text{tr}(\mathbf{X}^{-1} \mathbf{Y} + \mathbf{Y}^{-1} \mathbf{X} - 2\mathbf{I})$$

Jensen-Bregman 对数行列式散度 (Jensen-Bregman LogDet Divergence) 是印度学者 Anoop Cherian 等借鉴 Bregman 散度提出的， $d_\phi(\mathbf{X}, \mathbf{Y}) = \phi(\mathbf{X}) - \phi(\mathbf{Y}) - \langle \mathbf{X} - \mathbf{Y}, \nabla_\phi(\mathbf{Y}) \rangle$ ，其中 $\phi(\cdot)$ 是 Legendre 类型的严格凸函数。Bregman 散度是严格凸函数，不对称、非负、正定的。Bregman 散度不满足对称性和三角形不等式，因而不是合法度量。与 KL 散度类似，这些问题可以通过对称化处理解决，即生成 Jensen-Bregman 散度， $J_\phi(\mathbf{X}, \mathbf{Y}) = \frac{1}{2}(d_\phi(\mathbf{X}, S) + d_\phi(S, \mathbf{Y}))$ ，其中 $S = \frac{\mathbf{X} + \mathbf{Y}}{2}$ 。将 Jensen-Bregman 散度扩展成埃米特矩阵散度即得到 JBLD 度量 [276, 278]，

$$J_{ld}(\mathbf{X}, \mathbf{Y}) = \log \det \left\{ \frac{\mathbf{X} + \mathbf{Y}}{2} \right\} - \frac{1}{2} \log \det \{ \mathbf{XY} \} \quad (6.5)$$

根据 JBLD 的定义可知，它满足非负性、对称性和三角形不等式，是个合法的度量。此外，JBLD 还具备：仿射不变性 $J_{ld}(\mathbf{AXB}, \mathbf{AYB}) = J_{ld}(\mathbf{X}, \mathbf{Y})$ 和 求逆不变性 $J_{ld}(\mathbf{X}^{-1}, \mathbf{Y}^{-1}) = J_{ld}(\mathbf{X}, \mathbf{Y})$ 。

¹Cholesky 矩阵分解是指将对称正定矩阵分解成一个下三角矩阵以及它的共轭转置矩阵的乘积。

6.1.1.3 协方差矩阵算子

对称正定矩阵空间对加法和数乘不封闭，若赋予其黎曼度量就构成黎曼流形。据此土耳其学者 Oncel Tuzel 等提出了一种新的特征描述方法，即区域协方差矩阵，并将其应用于行人步态检测 [275]。区域协方差描述子生成的结果是对称正定矩阵，因而可使用黎曼距离量化它们之间的相似性，实现分类、检测、跟踪等任务。若以 \mathbf{I} 表示原始图像，以 $\mathbf{F} \in \mathbb{R}^{W \times H \times d}$ 表示从原始图像中提取的 d 维特征数组， $\mathbf{F}(x, y) = \Phi(\mathbf{I}(x, y))$ ，其中函数 Φ 表示特征映射，如强度、彩色、梯度、滤波响应等。对于图像中的任意给定区域 $Rg \subset \mathbf{F}$ ，以 $\{\mathbf{z}_i\}_{i=1}^S$ 表示区域 Rg 内的特征点集，区域协方差矩阵定义为

$$\mathbf{C}_{Rg} = \frac{1}{S-1} \sum_{i=1}^S (\mathbf{z}_i - \mu)(\mathbf{z}_i - \mu)^T \quad (6.6)$$

其中 μ 是特征均值。图 6.2 以 SAR 图像为例简要说明了协方差描述子的计算流程。

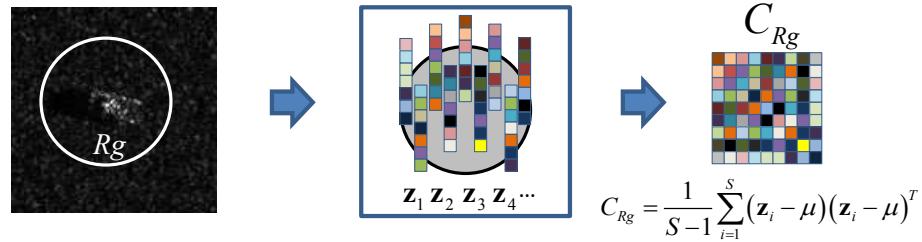


图 6.2 区域协方差描述子图例，根据输入图像提取区域 Rg 的 d 组特征矢量， $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_d$ ，然后根据 d 组特征矢量计算其区域协方差矩阵 C_{Rg} ，使用该矩阵作为区域描述子。

由 d 个特征所构成的区域协方差矩阵 $\mathbf{C}_{Rg} \in \mathbb{R}^{d \times d}$ ，它的对角线元素反映每类特征数组的自身方差，非对角线元素反映不同特征之间的相关性。区域协方差矩阵是一种有效的区域描述子，它广泛应用于目标检测、目标跟踪、视频识别等 [277-279, 282]。它提供了一种融合多个不同类型特征的直观方法，量化地描述这些特征之间的相关性。联想到 SAR 图像的目标识别，目标区域的协方差描述子可以将多个不同特征进行重新组合，根据它们的统计量（如方差）生成特征描述子，因而能有效应对各种非标准实验设置。

区域协方差描述子是 d 维对称矩阵，它只有 $d(d + 1)/2$ 个独立元素，其维数远低于原始的特征数组。对于任意给定区域 Rg ，协方差 \mathbf{C}_{Rg} 的计算与其每类特征点的顺序与个数无关，这就暗含了在不同图像同一区域实现旋转不变与尺度不变，因而能改进目标识别性能。然而，协方差矩阵描述子的表现优劣直接取决于所选择的特征，如何选取特征、以及特征数量是需要解决的关键问题。

6.1.2 Stiefel 和 Grassmann 流形

6.1.2.1 Grassmann 流形的定义

任意的 $d \times m$ 维矩阵空间 $\mathbb{R}^{d \times m}$ 是一种典型的黎曼流形 [283, 284]。若矩阵的所有列两两正交，称其为 Stiefel 流形 $S(m, d)$ ，正交是指满足条件

$$S(m, d) \triangleq \mathbf{Y} \in \mathbb{R}^{d \times m} : \mathbf{Y}^T \mathbf{Y} = \mathbf{I}_m \quad (6.7)$$

其中 \mathbf{I}_m 是 $m \times m$ 单位矩阵。Stiefel 流形可看作是正交矩阵右乘的商空间 $S(m, d) = \frac{O(d)}{O(d-m)}$ ，其中 $O(d)$ 表示 $d \times d$ 维正交群 (Orthogonal Group)。Grassmann 流形 $G(m, d)$ 是与 Stiefel 流形有着部分相似性的数学对象，它定义为满足等价关系

$$\mathbf{Y}_1 \sim \mathbf{Y}_2 \quad \text{if and only if} \quad \text{Span}(\mathbf{Y}_1) = \text{Span}(\mathbf{Y}_2)$$

的 Stiefel 流形 [285]，其中 $\text{Span}(\mathbf{Y})$ 表示矩阵 \mathbf{Y} 的列向量张成的线性子空间²。

定义 6.5: Grassmann 流形 $G(m, d)$ 是一组定义在 \mathbb{R}^d 上的 m 维线性子空间。

流形的欧几里德表示可以看作是由所有 $d \times m$ 维满秩矩阵 $\mathbf{Y} \in \mathbb{R}^{d \times m}$ 构成的空间 $\mathbb{R}_{d,m}^{(0)}$ ，矩阵空间 $\mathbb{R}_{d,m}^{(0)}$ 满足等价关系： $\mathbb{R}_{d,m}^{(0)}$ 空间中的任意元素， $\mathbf{Y}_1, \mathbf{Y}_2 \in \mathbb{R}_{d,m}^{(0)}$ ，若满足 $\text{Span}(\mathbf{Y}_1) = \text{Span}(\mathbf{Y}_2)$ ，则它们相等。因此， $\mathbb{R}_{d,m}^{(0)}$ 中的元素与 Grassmann 流形 $G(m, d)$ 上的点是一一对应的。Grassmann 流形 $G(m, d)$ 是解析流形，它的维数是 $md - mm = m(d - m)$ 。由于 Grassmann 流形 $G(m, d)$ 是个 $m \times (d - m)$ 维的紧支撑黎曼流形，其中的任意元素可表示为 $d \times m$ 维正交矩阵 \mathbf{Y} ，该矩阵满足 $\mathbf{Y}^T \mathbf{Y} = \mathbf{I}_m$ ，因此，通常更关注由 \mathbf{Y} 张成的线性子空间 $\text{Span}(\mathbf{Y})$ 而非矩阵本身。Grassmann 流形任意点的元素矩阵表示并不唯一，判断 Grassmann 流形的两个元素是否相等只需要判断它们张成的子空间是否相等。

定义 6.6: 任意的正交矩阵 $\mathbf{Y}_1, \mathbf{Y}_2 \in \mathbb{R}^{d \times m}$ ， $\mathbf{Y}_1^T \mathbf{Y}_1 = \mathbf{I}_m, \mathbf{Y}_2^T \mathbf{Y}_2 = \mathbf{I}_m$ ，若它们满足： $\mathbf{Y}_1 \cong \mathbf{Y}_2$ 当且仅当 $\text{Span}(\mathbf{Y}_1) = \text{Span}(\mathbf{Y}_2)$ ，可将其看作 Grassmann 流形的元素。

推论 6.1: $\mathbf{Y}_1 \cong \mathbf{Y}_2$ 当且仅当 $\mathbf{Y}_1 = \mathbf{Y}_2 \mathbf{R}$ ，其中 $\mathbf{R}_m \in O(m)$ 是任意的正交矩阵。

6.1.2.2 Grassmann 流形的距离度量

黎曼度量是衡量两组线性子空间之间距离的一种规范化度量，即测地线长度。一种直观地、易于计算的 Grassmann 流形测地距离是源自于两个子空间之间的主角 (Principal Angle)，或称为规范化相关系数 [286]，如图 6.3 所示。

² \mathbf{Y} 的列向量张成的包含 \mathbf{Y} 的最小线性子空间，详见第二章推论 2.1。

定义 6.7 (主角): 若以 $\mathbf{Y}_1, \mathbf{Y}_2$ 表示两个 $d \times m$ 维的正交矩阵，则子空间 $Span(\mathbf{Y}_1)$ 和 $Span(\mathbf{Y}_2)$ 之间的主角 $0 \leq \theta_1 \leq \dots \leq \theta_m \leq \frac{\pi}{2}$ 可递归地定义为：

$$\cos \theta_k = \max_{\mathbf{u}_k \in Span(\mathbf{Y}_1)} \max_{\mathbf{v}_k \in Span(\mathbf{Y}_2)} \mathbf{u}_k^T \mathbf{v} \quad subject \ to \left\{ \begin{array}{l} \mathbf{u}_k^T \mathbf{u}_k = 1, \mathbf{v}_k^T \mathbf{v}_k = 1 \\ \mathbf{u}_k^T \mathbf{u}_j = 0, \mathbf{v}_k^T \mathbf{v}_j = 0, j = 1, \dots, m, j \neq k \end{array} \right\}$$

第一主角 θ_1 定义为两组线性子空间的单位向量的最小夹角，即第一规范相关系数，以此类推，可得第 k 主角或规范相关系数的定义。

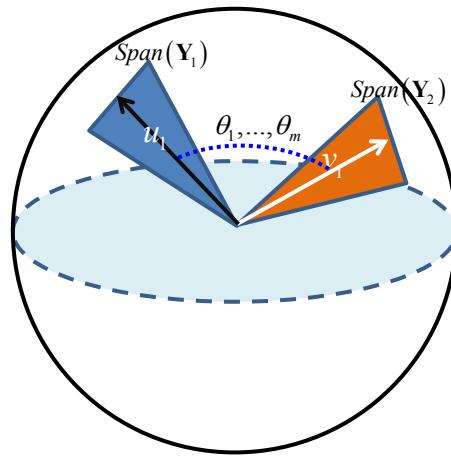


图 6.3 主角与测地距离图示。图中 $Span(\mathbf{f}_1)$ 和 $Span(\mathbf{f}_2)$ 表示两个信号的单演信号表示构成的子空间，它们之间的距离根据主角 $\theta_1, \dots, \theta_m$ 计算。Grassmann 流形两点的主角所计算的距离就是测地距离。

黎曼流形是一类赋予光滑内积的特殊的微分流形 [283]，Grassmann 流形是建立在正交约束条件的一种特殊黎曼流形，因而可将测地距离与主角联系起来。对于流形的点 \mathbf{Y} 处的任意两个切向量 Δ_1, Δ_2 ，它们之间的黎曼度量定义为

$$\langle \Delta_1, \Delta_2 \rangle_{\mathbf{Y}} = tr(\Delta_1^T (\mathbf{I}_d - \mathbf{Y} \mathbf{Y}^T / 2) \Delta_2^T) = tr(\Delta_1^T \Delta_2) \quad (6.8)$$

根据度量 (6.8) 可以导出测地距离，即连接流形任意两点间的最短曲线的长度，称为是弧长。根据主角计算弧长距离为，

$$d_{arc}^2(\mathbf{Y}_1, \mathbf{Y}_2) = \sum_{j=1}^k \theta_j^2 = \|\Theta\|_2^2 \quad (6.9)$$

其中 $\Theta = [\theta_1, \theta_2, \dots, \theta_m] \in \mathbb{R}^m$ 是主角。根据定义可知，要计算两组线性子空间之间的主角首先要求解定义 6.7 中的最优化问题，显然计算效率不高。实际上并不直

接求解该最优化问题，而是根据两个矩阵乘积 $\mathbf{Y}_1^T \mathbf{Y}_2$ 的奇异值分解（Singular Value Decomposition, SVD）来计算主角。若矩阵 $\mathbf{Y}_1^T \mathbf{Y}_2$ 的奇异值分解为

$$\mathbf{Y}_1^T \mathbf{Y}_2 = \mathbf{U} \Lambda \mathbf{V}^T \quad (6.10)$$

其中 $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m]$ 和 $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_m]$ 是酉矩阵， $\Lambda = \text{diag}(\cos \theta_1, \dots, \cos \theta_m)$ 是主角的三角函数组成的对角矩阵。根据定义可知主角构成一个单调不减序列 $0 \leq \theta_1 \leq \dots \leq \theta_m \leq \frac{\pi}{2}$ ，相应地规范相关系数构成一个单调不增序列， $1 \geq \cos \theta_1 \geq \dots \geq \cos \theta_m \geq 0$ 。

根据定义可以计算两个不同维数的矩阵所构成的子空间之间的主角³，但本文主要讨论相同维数的两个矩阵之间的距离。根据主角的定义，学者们提出了多种 Grassmann 流形距离，包括 [287, 288]：

- Projection Metric⁴ $d_{Pj}^2(\mathbf{Y}_1, \mathbf{Y}_2) = \sum_{j=1}^m \sin^2(\theta_m) = m - \|\mathbf{Y}_1^T \mathbf{Y}_2\|_F^2$
- Binet-Cauchy Metric $d_{BC}^2(\mathbf{Y}_1, \mathbf{Y}_2) = 1 - \prod_{j=1}^m \cos^2(\theta_j) = 1 - \|\mathbf{Y}_1^T \mathbf{Y}_2\|^2$
- Max Correlation $d_{max}^2(\mathbf{Y}_1, \mathbf{Y}_2) = 1 - \cos^2(\theta_1) = \sin^2(\theta_1)$
- Min Correlation $d_{min}^2(\mathbf{Y}_1, \mathbf{Y}_2) = 1 - \cos^2(\theta_m) = \sin^2(\theta_m)$
- Procrustes Metric $d_{Pc}^2(\mathbf{Y}_1, \mathbf{Y}_2) = 4 \sum_{j=1}^m \sin^2(\theta_1/2)$
- Fubini-Study Metric $d_{FS}^2(\mathbf{Y}_1, \mathbf{Y}_2) = \arccos \|\mathbf{Y}_1^T \mathbf{Y}_2\| = \arccos(\prod_{j=1}^m \cos(\theta_j))$

投影距离和 Finet-Cauchy 距离满足度量公理 6.1，是合法的 Grassmann 度量。

6.2 单演信号尺度空间的流形学习

论文引入单演信号对 SAR 图像特性进行描述，并在前文分别提出了欧氏空间与 Hilbert 空间的分类识别框架。与前面两种框架不同，本章将使用单演信号的多尺度表示生成的分量来构造黎曼流形和 Grassmann 流形，充分挖掘数据内部的非线性几何特性，并将其用于目标的分类识别问题。

6.2.1 单演信号协方差矩阵构建的黎曼流形

根据第 6.1.1.3 节所述可知，区域协方差描述子可以融合不同类型的区域特征。受该思路的启发，联想到将单演信号多尺度表示生成的多个分量使用区域协

³因为两个维数不相同的矩阵它们的奇异值分解可以实现。

⁴投影度量通常有两种，投影 2 范数和投影 F 范数，本文列出的是投影 F 范数，投影 2 范数定义为 $d_{p2}(\mathbf{Y}_1, \mathbf{Y}_2) = \|\mathbf{Y}_1 \mathbf{Y}_1^T - \mathbf{Y}_2 \mathbf{Y}_2^T\|_2$ 。

方差特征描述子进行融合，生成的区域协方差描述子是对称正定矩阵，所有该矩阵组合起来形成对称正定矩阵空间中 Sym_d^+ ，若赋予其某种黎曼度量就构成黎曼流形。基于这一想法，本节分别提出两种黎曼流形学习方案：一是借助面向流形的对数算子将流形的元素映射成切向空间的矢量，由于流形的切向空间是个典型的欧氏空间，可以构造线性表示系统，通过求解最稀疏表示判决目标类别；二是通过非线性映射将流形嵌入到再生核 Hilbert 空间，使用再生核 Hilbert 空间的像设计线性表示系统，求解最优表示系数，根据最优表示系数对未知样本的像进行重构，根据重构误差判决目标类别。

6.2.1.1 单演信号协方差描述子

对于二维信号 f ，首先对其做单演信号多尺度表示。假设构造 S 尺度 Log-Gabor 滤波器对原始信号进行无限延拓， $\{\mathbf{f}_{M,1}, \mathbf{f}_{M,2}, \dots, \mathbf{f}_{M,S}\}$ ，将生成的不同尺度空间的信号进行极坐标分解生成多个局部分量

$$\mathbf{f} = [\underbrace{A_1, \varphi_1, \theta_1}_{\mathbf{f}_{M,1}}, \underbrace{A_2, \varphi_2, \theta_2}_{\mathbf{f}_{M,2}}, \dots, \underbrace{A_S, \varphi_S, \theta_S}_{\mathbf{f}_{M,S}}]$$

其中 $\{A_s, \varphi_s, \theta_s\}_{s=1}^S$ 是第 s 尺度单演信号所分解的局部幅度、局部相位、局部方位。根据区域协方差描述子的定义，计算单演信号特征簇 \mathbf{f} 的协方差描述子，假设只计算信号中 $w \times h$ 区域的协方差，

$$\mathbf{C}_M(p, q) = \frac{1}{wh - 1} \sum_{k=1}^{wh} (\mathbf{f}_p(k) - \bar{\mathbf{f}}_p)(\mathbf{f}_q(k) - \bar{\mathbf{f}}_q) \quad (6.11)$$

其中 $\bar{\mathbf{f}}_p$ 表示第 p 类分量数组的均值。利用 S 尺度 Log-Gabor 滤波进行周期延拓，共有 $d = 3S$ 组单演信号分量，生成 $d \times d$ 维区域协方差矩阵 $\mathbf{C}_M \in \mathbb{R}^{d \times d}$ 。协方差矩阵的对角线元素表示每组单演信号分量的方差，非对角线元素表示不同分量之间的相关性。协方差矩阵是对称的，所以它只有 $d(d + 1)/2$ 个独立元素，即它的维数为 $d(d + 1)/2$ 。图 6.4 给出了单演信号协方差描述子的计算过程。

6.2.1.2 基于协方差矩阵的分类学习

单演信号多尺度分量所生成的协方差描述子 \mathbf{C}_M 是 $d \times d$ 维对称正定矩阵 $\mathbf{C}_M \in \text{Sym}_d^+$ ，它对于数乘并不封闭，不能形成传统的向量空间，若赋予某种黎曼度量构成黎曼流形。因此，如果将单演信号协方差矩阵直接应用于传统的学习方法显然是不合理的，该结论将会在实验部分做进一步验证分析。针对 SAR 图像目标识别问题，本节提出了两种解决方案：一是使用矩阵对数算子将协方差矩阵映射到它的切向空间，由于切向空间是典型的欧氏向量空间，可以使用稀疏信号表示来对识别问题进行建模；二是将黎曼流形嵌入到再生核 Hilbert 空间，使用非线

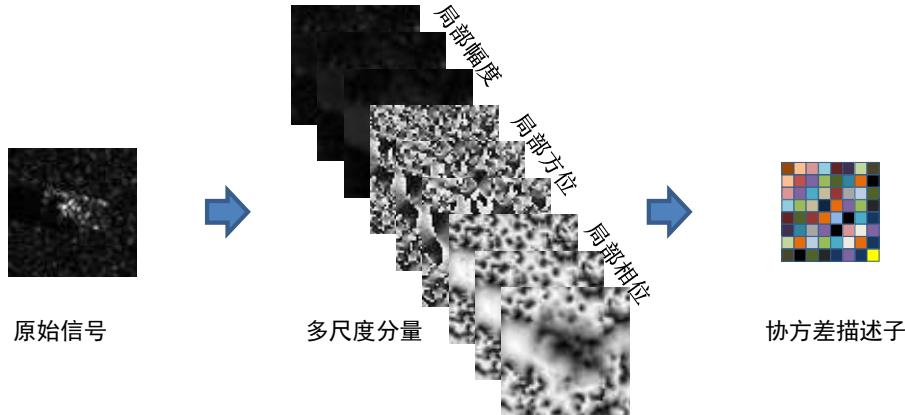


图 6.4 单演信号协方差描述子生成示意。

性映射将协方差矩阵描述子投影到再生核 Hilbert 空间，根据协方差矩阵在 Hilbert 空间中的像构建线性表示系统，通过求解最优表示系数对未知样本进行重构，根据重构误差判决目标类型 [289]。

方案一：对数协方差矩阵的稀疏表示：根据第 6.1.1 节介绍的基础知识可知，利用矩阵对数算子（面向流形）和矩阵指数算子能够实现对称正定矩阵（流形的元素）与其切向空间向量的相互转换。黎曼流形的切向空间是典型的欧氏向量空间，因此可以使用传统的学习方法，从而解决特征表示向量与分类学习方法之间的不同质矛盾。

任意对称正定矩阵 $\mathbf{C}_M \in \mathbf{Sym}_d^+$ ，若它的特征值分解 $\mathbf{C}_M = \mathbf{U}\Lambda\mathbf{U}^T$ ，其中 \mathbf{U} 是特征向量组成的酉矩阵， $\Lambda = diag\{\lambda_1, \lambda_2, \dots, \lambda_d\}$ 是特征值组成的对角矩阵，则 $\log(\mathbf{C}_M) = \mathbf{U}\log(\Lambda)\mathbf{U}^T$ 。对数运算得到的矩阵仍然对称，但不满足正定或半正定的要求，它对数乘和加法封闭，所以对数矩阵构成的空间是欧氏空间 [290]。由于 \mathbf{C}_M 只有 $d(d+1)/2$ 个独立元素，为了得到它的最紧致表示，只保留其上三角矩阵位置的元素从而生成表示向量

$$vec(\mathbf{C}_M) = [\mathbf{C}_{1,1}, \sqrt{2}\mathbf{C}_{1,2}, \sqrt{2}\mathbf{C}_{1,3}, \dots, \mathbf{C}_{2,2}, \sqrt{2}\mathbf{C}_{2,3}, \dots, \mathbf{C}_{d,d}] \quad (6.12)$$

其中非对角元素乘以 $\sqrt{2}$ 是为了保持其能量不变。

与所有同阶对称矩阵族相一致，对数协方差矩阵族对加法和数乘（线性组合）两种运算封闭，可构成标准的欧氏向量空间 [291]，因而可应用经典的学习方法如稀疏表示。首先使用训练样本的对数协方差矩阵构造未知样本协方差矩阵的线性表示。假若 K 个类别都有足够多的训练样本 $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,n_k}]$ ，首先计算所有类别训练样本的单演信号协方差描述子， $[\mathbf{C}_{k,1}, \mathbf{C}_{k,2}, \dots, \mathbf{C}_{k,n_k}]$ ，然后对所有协方差矩阵取对数运算得到对数协方差矩阵， $[\log(\mathbf{C}_{k,1}), \log(\mathbf{C}_{k,2}), \dots, \log(\mathbf{C}_{k,n_k})]$ ，使用 (6.12) 生成它们的紧致向量表示 $[vec(\log(\mathbf{C}_{k,1})), vec(\log(\mathbf{C}_{k,2})), \dots, vec(\log(\mathbf{C}_{k,n_k}))]$ ，将所有的训练样本的紧致向量组合构成过完备字典 $\mathbf{D}_{cov} = [\mathbf{D}_{cov,1}, \mathbf{D}_{cov,2}, \dots, \mathbf{D}_{cov,K}] \in \mathbb{R}^{m \times n}$ ，

其中 $\mathbf{D}_{cov,k} = [vec(\log(\mathbf{C}_{k,1})), vec(\log(\mathbf{C}_{k,2})), \dots, vec(\log(\mathbf{C}_{k,n_k}))]$ 表示由第 k 类训练样本的协方差描述子最紧致表示向量所构成的子字典， $m = d(d+1)/2$ 是描述向量的维数， $n = \sum_{k=1}^K n_k$ 是训练样本总数。对于未知样本 \mathbf{y} ，类似的流程可得到它的描述向量 $\mathbf{r}_y \in \mathbb{R}^m$ ，使用字典原子对未知样本的描述向量进行线性表示

$$\mathbf{r}_y = \mathbf{D}_{cov,1}\alpha_1 + \mathbf{D}_{cov,2}\alpha_2 + \cdots + \mathbf{D}_{cov,K}\alpha_K = \mathbf{D}_{cov}\alpha \quad (6.13)$$

其中 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_K] \in \mathbb{R}^n$ 是表示向量。通常期望系数向量 α 仅有很少数的非零项，从而包含未知样本类别信息，通过最小化 ℓ_1 范数即可实现，

$$\min_{\alpha} \|\alpha\|_1 \quad subject to \quad \|\mathbf{r}_y - \mathbf{D}_{cov}\alpha\|_2 \leq \epsilon \quad (6.14)$$

借助传统的优化算法如线性规划可得到问题的最优稀疏解 $\hat{\alpha}$ ，根据最优解推断未知样本的类别 $\min_{k=1,\dots,K} \|\mathbf{r}_y - \mathbf{D}_{cov,k}\hat{\alpha}_k\|_2$ 。算法的整个流程如图 6.5 所示。

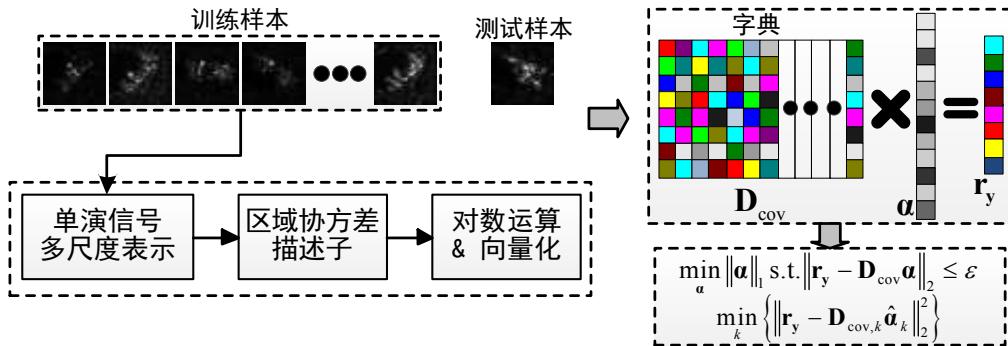


图 6.5 对数协方差矩阵稀疏表示流程图，首先使用训练样本的向量化对数协方差矩阵构成过完备字典，根据字典原子对未知样本的向量化对数协方差矩阵进行线性表示，使用最小化 ℓ_1 范数求取最优解，利用最优表示系数重构未知样本，根据重构误差进行判决。

方案二：再生核 Hilbert 空间核回归模型：借助矩阵对数算子可将对称正定矩阵映射到切向空间，从而解决特征表示向量与分类学习方法之间不同质的矛盾。然而，面向流形的对数算子仅仅是在某个很小的邻域内有唯一的定义，若在整个空间都使用对数映射通常会造成不同程度的信息损失。为了尽量避免信息损失，本节将黎曼流形嵌入到抽象的再生核 Hilbert 特征空间，然后使用再生核 Hilbert 空间的像对未知样本进行线性表示。

根据第五章第 5.1 节的论述可知，核函数能够将再生核 Hilbert 空间像的内积运算与原始数据空间样本的内积运算连接起来。因而，首先要讨论黎曼流形的正定核函数。尽管研究人员提出了许多流形核函数，其中大多都不满足 Mercer 条件（见论文第五章定义 5.4），因而没有坚实的理论支撑。印度学者 Anoop Cherian 等在吸取了对称化 KL 散度的研究成果基础上，提出了一种对称正定矩阵空间的仿

射不变、求逆不变的距离度量，即 Jensen-Bregman 对数行列式散度。之后，德国学者 Suvrit Sra、澳大利亚学者 Mehrtash T. Harandi 等根据 JBLD 度量构造合法的黎曼核函数，从而将对称正定矩阵空间（黎曼流形）嵌入到 Hilbert 空间中 [277, 292]。

定义 6.8：若以 $\Omega = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$ 表示对称正定矩阵空间 Sym_d^+ 的非空集合，若函数 $\kappa(\cdot, \cdot) : \Omega \times \Omega \mapsto \mathbb{R}^+$ 满足对称性 $\kappa(\mathbf{X}_1, \mathbf{X}_2) = \kappa(\mathbf{X}_2, \mathbf{X}_1)$ ，且对于任意的实数 $\alpha_i, \alpha_j \in \mathbb{R}$ 都有不等式 $\sum_{i,j} \alpha_i \alpha_j \kappa(\mathbf{X}_i, \mathbf{X}_j) \geq 0$ ，称其为黎曼核函数。

根据黎曼核函数的定义，德国学者 Suvrit Sra 证明了核函数 [292]

$$\kappa(\mathbf{X}, \mathbf{Y}) = \exp\{-\beta J_{ld}^2(\mathbf{X}, \mathbf{Y})\} = \frac{\det^\beta(\mathbf{XY})}{(\det(\mathbf{X} + \mathbf{Y})/2)^\beta} \quad (6.15)$$

存在，且它是正定核函数对于任意的 $\beta > 0$ 。文献 [292] 对正定核函数进行了详细地推导证明，此处不再赘述。

定理 6.1：对于可相互交换的对称正定矩阵集合 $\Omega \subset \text{Sym}_d^+$ ，若赋予 Jensen-Bregman 对数行列式散度，则子空间 (Ω, J_{ld}) 可以等距地嵌入到 Hilbert 空间。

根据定理 6.1 有，若给定一个抽象函数 $\phi(\cdot) : \text{Sym}_d^+ \mapsto \mathcal{H}$ ，可将对称正定矩阵空间嵌入到再生核 Hilbert 空间，且再生核 Hilbert 空间的像对数乘和加法运算封闭。对给定的训练样本 $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,n_k}]$ ，首先计算它们的单演信号协方差描述子，得到的是一组对称正定矩阵 $[\mathbf{C}_{k,1}, \mathbf{C}_{k,2}, \dots, \mathbf{C}_{k,n_k}]$ ，然后使用映射 $\phi(\cdot)$ 将协方差描述子投影到 Hilbert 空间中，得到的像分别为 $\phi(\mathbf{C}_{k,1}), \phi(\mathbf{C}_{k,2}), \dots, \phi(\mathbf{C}_{k,n_k})$ 。将所有训练样本协方差描述子的像组合起来构成过完备字典 $\mathbf{D}_\phi = \phi(\mathbf{C}_1), \phi(\mathbf{C}_2), \dots, \phi(\mathbf{C}_K)$ ，其中 \mathbf{C}_k 表示第 k 类训练样本的单演信号协方差描述子。对于未知样本 \mathbf{y} 可以类似的方法计算它在 Hilbert 空间中的像。由于 Hilbert 空间中对数乘法是封闭的，可以使用训练样本协方差描述子的像对未知样本协方差描述子 \mathbf{C}_y 的像进行线性表示，即核回归模型

$$\phi(\mathbf{C}_y) = \phi(\mathbf{C}_1)\alpha_1 + \phi(\mathbf{C}_2)\alpha_2 + \dots + \phi(\mathbf{C}_K)\alpha_K = \mathbf{D}_\phi\alpha \quad (6.16)$$

其中 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_K] \in \mathbb{R}^n$ 是表示系数。根据第五章第 5.2 节介绍的方法，利用最小化 ℓ_2 范数的策略求解核回归模型 (6.16) 的最优回归系数，

$$\min_{\alpha} \|\alpha\|_2 \quad subject to \quad \|\phi(\mathbf{C}_y) - \mathbf{D}_\phi\alpha\|_2 \leq \epsilon \quad (6.17)$$

然而嵌入函数 $\phi(\cdot)$ 的具体形式未知，协方差描述子在 Hilbert 核空间的像只是个抽象概念，无法直接计算。借鉴第五章第 5.2 节的处理方法，首先将约束最优化问题转化成为无约束优化问题，

$$\min_{\alpha} \{\|\phi(\mathbf{C}_y) - \mathbf{D}_\phi\alpha\|_2^2 + \lambda \|\alpha\|_2^2\} \quad (6.18)$$

并将目标函数（记为 $g(\alpha)$ ）做向量展开，

$$\begin{aligned} g(\alpha) &= (\phi(\mathbf{C}_y) - \mathbf{D}_\phi \alpha)^T (\phi(\mathbf{C}_y) - \mathbf{D}_\phi \alpha) + \lambda \alpha^T \alpha \\ &= \alpha^T \langle \mathbf{D}_\phi, \mathbf{D}_\phi \rangle \alpha - 2 \alpha^T \langle \mathbf{D}_\phi, \phi(\mathbf{C}_y) \rangle + \langle \phi(\mathbf{C}_y), \phi(\mathbf{C}_y) \rangle + \lambda \alpha^T \alpha \\ &= \alpha^T \kappa(\mathbf{C}_X, \mathbf{C}_X) \alpha - 2 \alpha^T \kappa(\mathbf{C}, \mathbf{C}_y) + \kappa(\mathbf{C}_y, \mathbf{C}_y) + \lambda \alpha^T \alpha \end{aligned} \quad (6.19)$$

其中 \mathbf{C}_X 表示所有训练样本的协方差矩阵集合，

$$\begin{aligned} \langle \mathbf{D}_\phi, \mathbf{D}_\phi \rangle &= \begin{bmatrix} \kappa(\mathbf{C}_{1,1}, \mathbf{C}_{1,1}) & \kappa(\mathbf{C}_{1,1}, \mathbf{C}_{1,2}) & \dots & \kappa(\mathbf{C}_{1,1}, \mathbf{C}_{K,n_K}) \\ \kappa(\mathbf{C}_{1,2}, \mathbf{C}_{1,1}) & \kappa(\mathbf{C}_{1,2}, \mathbf{C}_{1,2}) & \dots & \kappa(\mathbf{C}_{1,2}, \mathbf{C}_{K,n_K}) \\ \vdots & \vdots & \ddots & \vdots \\ \kappa(\mathbf{C}_{K,n_K}, \mathbf{C}_{1,1}) & \kappa(\mathbf{C}_{K,n_K}, \mathbf{C}_{1,2}) & \dots & \kappa(\mathbf{C}_{K,n_K}, \mathbf{C}_{K,n_K}) \end{bmatrix} = \kappa(\mathbf{C}_X, \mathbf{C}_X) \\ \langle \mathbf{D}_\phi, \phi(\mathbf{C}_y) \rangle &= \begin{bmatrix} \langle \phi(\mathbf{C}_{1,1}), \phi(\mathbf{C}_y) \rangle \\ \langle \phi(\mathbf{C}_{1,2}), \phi(\mathbf{C}_y) \rangle \\ \vdots \\ \langle \phi(\mathbf{C}_{K,n_K}), \phi(\mathbf{C}_y) \rangle \end{bmatrix} = \begin{bmatrix} \kappa(\mathbf{C}_{1,1}, \mathbf{C}_y) \\ \kappa(\mathbf{C}_{1,2}, \mathbf{C}_y) \\ \vdots \\ \kappa(\mathbf{C}_{K,n_K}, \mathbf{C}_y) \end{bmatrix} = \kappa(\mathbf{C}, \mathbf{C}_y) \end{aligned}$$

对函数 (6.19) 求自变量 α 的偏导数，令偏导数为零 $\frac{\partial g(\alpha)}{\partial \alpha} = 0$ 即可得到它的极小值，最优系数向量为 $\hat{\alpha} = (\kappa(\mathbf{C}_X, \mathbf{C}_X) + \lambda \mathbf{I})^{-1} \kappa(\mathbf{C}, \mathbf{C}_y)$ ，根据最优系数向量计算每个类别的重构误差，判决目标类型 $\min_k \|\phi(\mathbf{C}_y) - \phi(\mathbf{C}_k) \hat{\alpha}_k\|_2$ 。图 6.6 给出了算法的流程示意图。

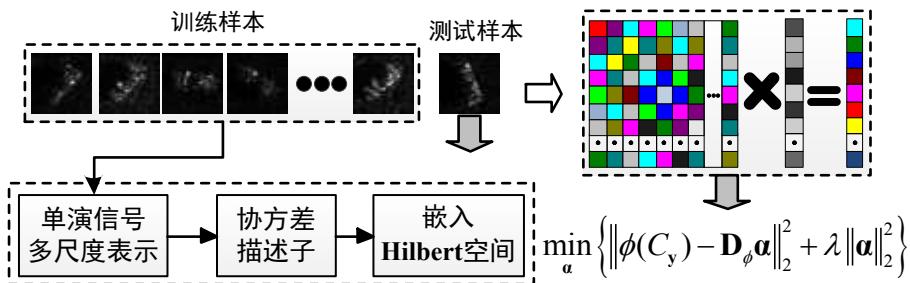


图 6.6 对称正定矩阵嵌入到 Hilbert 空间的线性表示流程。首先使用训练样本单演信号协方差描述子的 Hilbert 空间的像构造字典，根据字典对未知样本协方差描述子的像进行线性表示，使用用最小化 ℓ_2 范数求解最优表示向量，根据重构误差判决目标类型。

6.2.2 单演信号系数矩阵构建的 Grassmann 流形

前文提出一种单演信号协方差矩阵描述子，并将其看作是黎曼流形的元素进行处理。与其不同，本节提出一种基于原始矩阵的 Grassmann 流形学习方法，根据单演信号多尺度表示生成的系数矩阵构造一种特殊的黎曼流形——Grassmann

流形，将单演信号系数矩阵整体看作是 Grassmann 流形的元素，挖掘它们的非线性几何特性 [293, 294]。

对于任意的二维有限长信号 $f \in \mathbb{R}^{W \times H}$ ，使用 S 尺度 Log-Gabor 滤波器对信号进行无限延拓，得到的 S 尺度单演信号为 $\{\mathbf{f}_{M,1}, \mathbf{f}_{M,2}, \dots, \mathbf{f}_{M,S}\}$ 。将不同尺度空间的信号执行极坐标分解

$$\mathbf{f} = [\underbrace{A_1, \varphi_1, \theta_1}_{{\mathbf{f}}_{M,1}}, \dots, \underbrace{A_S, \varphi_S, \theta_S}_{{\mathbf{f}}_{M,S}}] \in \mathbb{R}^{W \times H \times 3S}$$

其中 $\{A_k, \varphi_k, \theta_k\}_{k=1}^S$ 是第 k 尺度单演信号分解的局部幅度、局部相位、局部方位。将所有的单演信号分量重新排列成列向量得到 $WH \times 3S$ 维的矩阵，

$$\mathbf{f} = [A_1(:), \varphi_1(:), \theta_1(:), \dots, A_S(:), \varphi_S(:), \theta_S(:)] \in \mathbb{R}^{WH \times 3S}$$

尽管矩阵 \mathbf{f} 的列向量之间并不正交，借助数处理技巧如格莱姆－施密特正交化等很容易实现正交化约束，使其满足正交性 $\mathbf{f}^T \mathbf{f} = \mathbf{I}$ 和等价性 $\mathbf{f}_1 \cong \mathbf{f}_2$ 当且仅当 $\text{Span}(\mathbf{f}_1) = \text{Span}(\mathbf{f}_2)$ 。从而将单演信号系数矩阵 \mathbf{f} 整体看作 Grassmann 流形 $\mathcal{G}(m, d)$ 的点，其中 $m = 3S, d = WH$ 。然后设计 Grassmann 流形的分类方案。

6.2.2.1 方案一：最近邻分类

对于任意两组单演信号分量系数矩阵 \mathbf{f}_1 和 \mathbf{f}_2 ，它们所张成的线性子空间 $\text{Span}(\mathbf{f}_1)$ 和 $\text{Span}(\mathbf{f}_2)$ 之间的距离可以用子空间之间的主角来计算，如图 6.3 所示。根据流形的度量可计算两组单演信号系数矩阵之间的距离，然后借助简单的最近邻分类器即可实现对未知样本的类别的判定，即逐个比较未知样本单演信号分量矩阵所构成的子空间与训练样本单演信号分量矩阵所构成的子空间之间的距离。

给定训练样本 $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,n_k}]$ ，计算它们的单演信号多尺度表示 $\mathbf{f}_{k,1}, \mathbf{f}_{k,2}, \dots, \mathbf{f}_{k,n_k}$ ，由这些系数矩阵张成的子空间为 $\text{Span}(\mathbf{f}_{k,1}), \text{Span}(\mathbf{f}_{k,2}), \dots, \text{Span}(\mathbf{f}_{k,n_k})$ 。对于未知样本 \mathbf{y} ，使用类似的方法得到子空间 $\text{Span}(\mathbf{f}_y)$ ，根据 Grassmann 流形的距离度量（详见 6.1.2.2）分别比较 $\text{Span}(\mathbf{f}_y)$ 与每个训练样本的子空间之间的距离，然后将其判定为距离最小的训练样本所在的类别。图 6.7 给出了由单演信号分量矩阵所构成的 Grassmann 流形两点间的距离度量示意图。

6.2.2.2 方案二：核稀疏表示模型

论文 6.2.1.2 节提出将单演信号协方差描述子构成的黎曼流形嵌入到 Hilbert 特征空间中，借助黎曼核函数将 Hilbert 特征空间像与像的相似度量与原始的协方差矩阵的运算联系起来。受该思路的启发，本节考虑将单演信号分量矩阵构成的 Grassmann 流形以相似的方式嵌入到再生核 Hilbert 空间，然后设计核回归模型。

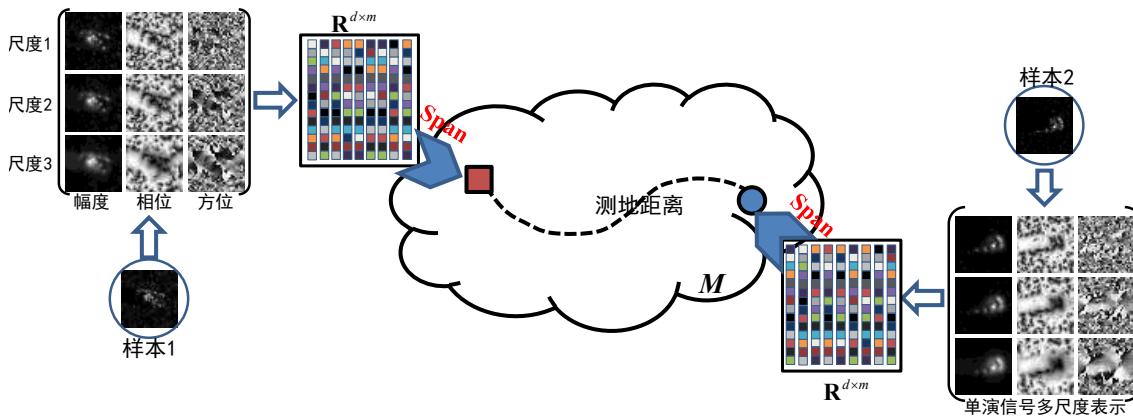


图 6.7 单演信号分量矩阵的 Grassmann 流形测地距离，首先定义两组单演信号分量矩阵的距离度量，然后寻找与未知样本距离最近的训练样本，根据匹配的训练样本的类别进行判决。

该思路实现的关键在于 Grassmann 流形核函数的设计，从而将再生核 Hilbert 空间像与像之间的相似度量与单演信号系数矩阵之间的距离联系起来。

1. Grassmann 核函数的设计

论文的 6.1.2.2 节给出了多个 Grassmann 流形的距离度量，包括投影度量、比奈特－柯西度量等。受到黎曼核函数的启发，论文考虑能否定义与 Grassmann 流形距离度量相容的正定核函数。

定义 6.9：Grassmann 核函数就是定义在 Grassmann 流形的正定核函数。

学者 Jihun Hamm 和 Daniel D. Lee 提出并证明了与投影距离和比奈特－柯西距离相容的投影核函数和比奈特－柯西核函数是 Grassmann 核函数 [288, 295]。

投影核函数：投影距离度量可以理解是将 Grassmann 流形的点 $\text{Span}(\mathbf{Y}) \in \mathcal{G}(m, d)$ 与它的投影矩阵 $\mathbf{Y}\mathbf{Y}^T$ 通过一个嵌入函数联系起来，

$$\phi_P : \mathcal{G}(m, d) \rightarrow \mathbb{R}^{d \times d}, \text{Span}(\mathbf{Y}) \mapsto \mathbf{Y}\mathbf{Y}^T \quad (6.20)$$

其中像 $\phi_P(\mathcal{G}(m, d))$ 是一组秩为 m 的正交投影矩阵。映射 (6.20) 是个等距映射，与之相关联的投影度量是 $\mathbb{R}^{d \times d}$ 空间中的欧氏距离，该空间中相应的内积是

$$\text{tr}[(\mathbf{Y}_1\mathbf{Y}_1^T)(\mathbf{Y}_2\mathbf{Y}_2^T)] = \|\mathbf{Y}_1^T\mathbf{Y}_2\|_F^2$$

从而有推论 (6.2)。

推论 6.2：与投影距离相容的核函数 $\kappa_P(\mathbf{Y}_1, \mathbf{Y}_2) = \|\mathbf{Y}_1^T\mathbf{Y}_2\|_F^2$ 是合法的 Grassmann 核函数。

证明：首先由 Grassmann 流形的定义可以推出，对任意的正交群 $\mathbf{R}_1, \mathbf{R}_2 \in O(m)$ 都有 $\kappa(\mathbf{Y}_1, \mathbf{Y}_2) = \kappa_P(\mathbf{Y}_1\mathbf{R}_1, \mathbf{Y}_2\mathbf{R}_2)$ ，因此投影核是良定义的。投影核函数的正

定性可通过矩阵的 Frobenius 范数的性质来证明。对于 Grassmann 流形的点 $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n \in \mathcal{G}(m, d)$ 及任意实数 $\alpha_1, \alpha_2, \dots, \alpha_n$, 有

$$\sum_{i,j} \alpha_i \alpha_j \|\mathbf{Y}_i^T \mathbf{Y}_j\|_F^2 = \sum_{i,j} \alpha_i \alpha_j \text{tr}(\mathbf{Y}_i \mathbf{Y}_i^T \mathbf{Y}_j \mathbf{Y}_j^T) = \text{tr}\left(\sum_i \alpha_i \mathbf{Y}_i \mathbf{Y}_i^T\right)^2 = \left\|\sum_i \alpha_i \mathbf{Y}_i \mathbf{Y}_i^T\right\|_F^2 \geq 0$$

■

故与投影距离相容的核函数是正定的。此外, 由投影核函数还可以衍生出一族 Grassmann 核函数, 如它的平方根 $\|\mathbf{Y}_1^T \mathbf{Y}_1\|_F$ 也是正定核函数。

比奈特—柯西核函数: 比奈特—柯西距离可理解是一种嵌入映射。若以 s 表示由 $\{1, 2, \dots, d\}$ 中任意 m 个元素所组成的子集 $s = \{r_1, r_2, \dots, r_m\}$, 以 $\mathbf{Y}^{(s)}$ 表示由矩阵 \mathbf{Y} 的第 r_1, r_2, \dots, r_m 行组成的 $m \times m$ 维矩阵, 以 s_1, s_2, \dots, s_n 表示所有以这种方式选取并按顺序排列的子集, 则比奈特—柯西嵌入函数就可以定义为

$$\phi_P : \mathcal{G}(m, d) \rightarrow \mathbb{R}^n, \mathbf{Y} \mapsto (\det(\mathbf{Y}^{(s_1)}), \det(\mathbf{Y}^{(s_2)}), \dots, \det(\mathbf{Y}^{(s_n)})) \quad (6.21)$$

其中 n 表示从 d 行中选出 m 行的数目, 这种情况的自然内积为 $\sum_{r=1}^n \det(\mathbf{Y}_1^{(s_r)}) \det(\mathbf{Y}_2^{(s_r)})$ 。

推论 6.3: 与比奈特—柯西距离相容的核函数 $\kappa_{BC}(\mathbf{Y}_1, \mathbf{Y}_2) = \det(\mathbf{Y}_1^T \mathbf{Y}_2)^2 = \det(\mathbf{Y}_1^T \mathbf{Y}_2 \mathbf{Y}_2^T \mathbf{Y}_1)$ 是合法的 Grassmann 核函数。

证明: 对于正交群 $\mathbf{R}_1, \mathbf{R}_2 \in O(m)$, 有 $\kappa_{BC}(\mathbf{Y}_1, \mathbf{Y}_2) = \kappa_{BC}(\mathbf{Y}_1 \mathbf{R}_1, \mathbf{Y}_2 \mathbf{R}_2)$, 因而比奈特—柯西核函数是良定义的。要证明比奈特—柯西核函数是正定的, 只需要证明 $\kappa(\mathbf{Y}_1, \mathbf{Y}_2) = \det(\mathbf{Y}_1^T \mathbf{Y}_2)$ 是正定的。

根据比奈特—柯西一致性有 $\det(\mathbf{Y}_1^T \mathbf{Y}_2) = \sum_s \det(\mathbf{Y}_1^{(s)}) \det(\mathbf{Y}_2^{(s)})$ 。因此, 对于 Grassmann 流形的任意点 $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n \in \mathcal{G}(m, d)$ 及任意实数 $\alpha_1, \alpha_2, \dots, \alpha_n$, 有

$$\sum_{i,j} \alpha_i \alpha_j \det(\mathbf{Y}_i^T \mathbf{Y}_j) = \sum_{i,j} \alpha_i \alpha_j \sum_s \det(\mathbf{Y}_i^{(s)}) \det(\mathbf{Y}_j^{(s)}) = \sum_s \left(\sum_i \alpha_i \det(\mathbf{Y}_i^{(s)}) \right)^2 \geq 0$$

■

因此, 比奈特—柯西核是正定核函数, 从而是 Grassmann 核函数。

2. 核稀疏表示

借助嵌入函数 $\phi(\cdot) : \mathcal{G}(m, d) \rightarrow \mathcal{H}$ 将单演信号系数矩阵映射到再生核 Hilbert 空间。对于第 k 类训练样本 $\mathbf{X}_k = [\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,n_k}]$, 将其单演信号多尺度表示生成的系数矩阵重新排列成向量 $\mathbf{f}_{k,1}, \mathbf{f}_{k,2}, \dots, \mathbf{f}_{k,n_k}$, 然后使用嵌入函数 $\phi(\cdot)$ 将系数矩阵投影到再生核 Hilbert 空间, $\phi(\mathbf{f}_{k,1}), \phi(\mathbf{f}_{k,2}), \dots, \phi(\mathbf{f}_{k,n_k})$ 。将所有类别的训练样本组合起来构成字典 $\mathbf{D}_\phi = [\phi(\mathbf{f}_1), \phi(\mathbf{f}_2), \dots, \phi(\mathbf{f}_K)]$, 其中 $\phi(\mathbf{f}_k) = [\phi(\mathbf{f}_{k,1}), \phi(\mathbf{f}_{k,2}), \dots, \phi(\mathbf{f}_{k,n_k})]$ 表示第 k 类训练样本在 Hilbert 空间的像的集合。对于未知样本 \mathbf{y} , 使用类似的

方法得到它的单演信号分量矩阵在 Hilbert 空间的像 $\phi(\mathbf{f}_y)$ 。使用字典原子对未知样本的像进行线性表示 $\phi(\mathbf{f}_y) = \phi(\mathbf{f}_1)\alpha_1 + \phi(\mathbf{f}_2)\alpha_2 + \cdots + \phi(\mathbf{f}_K)\alpha_K = \mathbf{D}_\phi\alpha$, 其中 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_K] \in \mathbb{R}^n$ 是系数向量。通常期望系数向量 α 仅有极少数非零值从而包含未知样本的类别信息, 通过最小化 ℓ_1 范数即可实现,

$$\min_{\alpha} \|\alpha\|_1 \quad \text{subject to} \quad \|\phi(\mathbf{f}_y) - \mathbf{D}_\phi\alpha\|_2 \leq \epsilon \quad (6.22)$$

将问题 (6.22) 转化为无约束优化问题,

$$\min_{\alpha} \left\{ \|\phi(\mathbf{f}_y) - \mathbf{D}_\phi\alpha\|_2^2 + \lambda \|\alpha\|_1 \right\} \quad (6.23)$$

然后将误差项 ($g(\alpha)$) 做向量展开,

$$\begin{aligned} g(\alpha) &= (\phi(\mathbf{f}_y) - \mathbf{D}_\phi\alpha)^T (\phi(\mathbf{f}_y) - \mathbf{D}_\phi\alpha) \\ &= \alpha^T \langle \mathbf{D}_\phi, \mathbf{D}_\phi \rangle \alpha - 2\alpha^T \langle \mathbf{D}_\phi, \phi(\mathbf{f}_y) \rangle + \langle \phi(\mathbf{f}_y), \phi(\mathbf{f}_y) \rangle \\ &= \alpha^T \kappa(\mathbf{f}_X, \mathbf{f}_X) \alpha - 2\alpha^T \kappa(\mathbf{f}_X, \mathbf{f}_y) + \kappa(\mathbf{f}_y, \mathbf{f}_y) \end{aligned} \quad (6.24)$$

其中 \mathbf{f}_X 表示所有训练样本的协方差矩阵集合,

$$\langle \mathbf{D}_\phi, \mathbf{D}_\phi \rangle = \begin{bmatrix} \kappa(\mathbf{f}_{1,1}, \mathbf{f}_{1,1}) & \kappa(\mathbf{f}_{1,1}, \mathbf{f}_{1,2}) & \dots & \kappa(\mathbf{f}_{1,1}, \mathbf{f}_{K,n_K}) \\ \kappa(\mathbf{f}_{1,2}, \mathbf{f}_{1,1}) & \kappa(\mathbf{f}_{1,2}, \mathbf{f}_{1,2}) & \dots & \kappa(\mathbf{f}_{1,2}, \mathbf{f}_{K,n_K}) \\ \vdots & \vdots & \ddots & \vdots \\ \kappa(\mathbf{f}_{K,n_K}, \mathbf{f}_{1,1}) & \kappa(\mathbf{f}_{K,n_K}, \mathbf{f}_{1,2}) & \dots & \kappa(\mathbf{f}_{K,n_K}, \mathbf{f}_{K,n_K}) \end{bmatrix} = \kappa(\mathbf{f}_X, \mathbf{f}_X)$$

式中的 $\langle \mathbf{D}_\phi, \mathbf{D}_\phi \rangle \in \mathbb{R}^{n \times n}$ 是格莱姆矩阵, 它的每个元素代表两个训练样本在 Hilbert 核空间的像的内积。类似的方法可以得出 $\kappa(\mathbf{f}_X, \mathbf{f}_y)$,

$$\langle \mathbf{D}_\phi, \phi(\mathbf{f}_y) \rangle = \begin{bmatrix} \langle \phi(\mathbf{f}_{1,1}), \phi(\mathbf{f}_y) \rangle \\ \langle \phi(\mathbf{f}_{1,2}), \phi(\mathbf{f}_y) \rangle \\ \vdots \\ \langle \phi(\mathbf{f}_{K,n_K}), \phi(\mathbf{f}_y) \rangle \end{bmatrix} = \begin{bmatrix} \kappa(\mathbf{f}_{1,1}, \mathbf{f}_y) \\ \kappa(\mathbf{f}_{1,2}, \mathbf{f}_y) \\ \vdots \\ \kappa(\mathbf{f}_{K,n_K}, \mathbf{f}_y) \end{bmatrix} = \kappa(\mathbf{f}_X, \mathbf{f}_y)$$

$\langle \mathbf{D}_\phi, \phi(\mathbf{f}_y) \rangle \in \mathbb{R}^n$ 是未知样本与训练样本在再生核 Hilbert 空间的相似向量, 向量的每个元素代表未知样本与某个训练样本在再生核 Hilbert 空间的像的内积。将 Grassmann 核函数代入 $g(\alpha)$ 中就可以求得问题 (6.22) 的最优稀疏向量。根据最优表示系数计算每个类别的重构误差, 从而判决未知样本的目标类别,

$$\min_{k=1,\dots,K} \|\phi(\mathbf{f}_y) - \phi(\mathbf{f}_k)\hat{\alpha}_k\|_2 \quad (6.25)$$

算法的整个流程如图 6.8 所示。

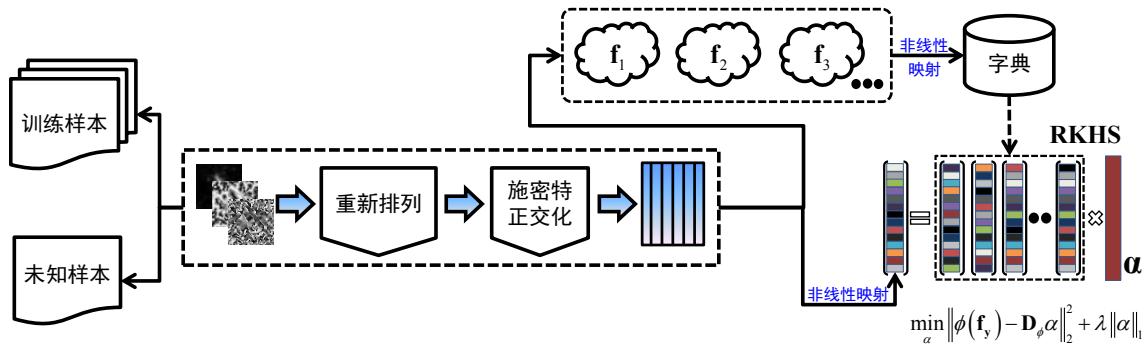


图 6.8 Grassmann 流形嵌入到 Hilbert 空间的分类示意图。使用嵌入函数将训练样本单演信号分量矩阵投影到 Hilbert 空间中，将 Hilbert 空间中的像组合起来构成字典，根据字典对未知样本的像进行线性表示，使用最小化 ℓ_1 范数计算最优表示向量，根据重构误差判决目标类型。

6.2.3 方向可控 Riesz 小波构建的 Grassmann 流形

Riesz 变换是传统的一维解析信号在高维信号空间的广义扩展，具有平移不变、尺度不变、方向可控等多个优良性质，单演信号就是根据二维信号的 Riesz 变换构成。为了充分考虑它的尺度与方向特性，许多学者利用 Riesz 变换构造小波框架。为了探索单演信号的局部性，Sofia C. Olhede 和 Georgios Metikas 通过定义一种广义的向量化四元母小波函数来构造一种四元单演小波族 [226]；Michael Unser 等借助 Riesz-Laplace 小波变换引出单演信号多分辨率分析，其中的复数化 Riesz 变换能够将能量有限 \mathbb{R}^2 空间的实数小波基映射到相应的复数空间 [225]；Stefan Held 等根据 Riesz 变换和各向同性小波框架定义一种超复平面单演信号，通过构造旋转兼容的特性将一维解析小波扩展到 n 维，该方法可以实现小波框架的方向可控 [296]；Michael Unser 和 Dimitri Van De Ville 提出一种用于设计方向可控、可逆的信号变换的广义连续域小波框架，该框架能够实现平移不变、尺度不变、方向可控、保持内积等优良特性 [227]；Raphaël Soulard 等提出一种向量化的单演信号框架，以实现多通道彩色信号的分析与处理 [239]。Adrien Depeursinge 和 Antonio Foncubierta-Rodriguez 提出一种基于局部尺度与方向的纹理学习方法，他们将 Riesz 小波的线性组合用于支持向量机学习。由于高阶 Riesz 变换所具备的诸多优良特性，本节考虑使用高阶 Riesz 变换生成的方向可控小波框架构建 Grassmann 流形，实现分类学习。

6.2.3.1 方向可控高阶 Riesz 小波

1. Riesz 变换：是 Hilbert 变换的广义扩展⁵。对于任意 ℓ_p 范数能量有限空间 $L^p(\mathbb{R}^n)$, $1 \leq p < \infty$ 的 n 维信号 $f(\mathbf{x}), \mathbf{x} = [x_1, x_2, \dots, x_n]^T$ ，它的 Riesz 变换 (\mathcal{R}) 定义为

⁵Hilbert 变换的传递函数是 $H(\omega) = -j\text{sign}(\omega)$ ，相应的脉冲响应是 $h(x) = \frac{1}{\pi x}$

$f_{\mathcal{R}_r}(\mathbf{x}) = \mathcal{R}_r\{f\}(\mathbf{x}) = \lim_{\epsilon \rightarrow \infty} c_n \int_{|\mathbf{y}| < \epsilon} \frac{y_r}{|\mathbf{y}|^{n+1}} f(\mathbf{x} - \mathbf{y}) d\mathbf{y}$, 其中 $r = 1, 2, \dots, n$, $c_n = \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}}$ 。
Riesz 变换的第 r 维子变换的卷积核定义为 $h_r(\mathbf{x}) = c_n \frac{x_r}{|\mathbf{x}|^{n+1}}$ 所以 Riesz 变换的 n 维滤波器组可写为

$$\mathcal{R}\{f\}(\mathbf{x}) = \begin{pmatrix} \mathcal{R}_1\{f\}(\mathbf{x}) \\ \mathcal{R}_2\{f\}(\mathbf{x}) \\ \vdots \\ \mathcal{R}_n\{f\}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} (h_1 * f)(\mathbf{x}) \\ (h_2 * f)(\mathbf{x}) \\ \vdots \\ (h_n * f)(\mathbf{x}) \end{pmatrix}. \quad (6.26)$$

若将 n 维信号 $f(\mathbf{x})$ 的傅立叶变换表示为 $F(\mathbf{u}) = \int_{\mathbb{R}^n} f(\mathbf{x}) \exp(-j2\pi\langle \mathbf{u}, \mathbf{x} \rangle) d\mathbf{x}$, 其中 $\mathbf{u} = [u_1, u_2, \dots, u_n]^T$ 。则第 r 维子变换的传递函数 (或是傅立叶乘子) 为 $H_r(\mathbf{u}) = -i \frac{u_r}{|\mathbf{u}|}$, 变换后的分量为 $F_{\mathcal{R}_r}(\mathbf{u}) = -i \frac{u_r}{|\mathbf{u}|} F(\mathbf{u})$ 。频域的向量化线性滤波器就表示为

$$H(\mathbf{u}) = [-j \frac{u_1}{|\mathbf{u}|}, -j \frac{u_2}{|\mathbf{u}|}, \dots, -j \frac{u_n}{|\mathbf{u}|}]^T \quad (6.27)$$

显然, Riesz 变换在所有方向都能保持单位长度, 因而具备各向同性的能量分布, 论文第三章对 Riesz 变换的特性进行了详细阐述。

2. 高阶 Riesz 变换: Michael Felsberg 和 Gerald Sommer 提出单演信号时只讨论了二维信号的一阶 Riesz 变换 f, f_x, f_y ⁷, 之后为了边缘检测、光流场估计等应用的需要, 逐渐形成了二维信号的二阶 Riesz 变换 $f, f_{xx}, f_{xy}, f_{yy}$ 。为了进一步挖掘 Riesz 变换的方向选择特性, Michael Unser 和 Dimitri Van De Ville 两人提出了高阶 Riesz 变换 [227]。高阶 Riesz 变换是根据一组多索引序列 $N = (N_1, N_2, \dots, N_n)$ 来定义 $\mathcal{R}^N = \sqrt{\frac{|N|!}{N!}} \mathcal{R}_1^{N_1} \mathcal{R}_2^{N_2} \dots \mathcal{R}_n^{N_n}$, 其中 $\frac{|N|!}{N!} = \binom{N}{N_1, \dots, N_n} = \frac{N!}{N_1! N_2! \dots N_n!}$ 。高阶变换 \mathcal{R}^N 的结果通常进行归一化以保持全局变换的能量。对于任意信号 $f(\mathbf{x})$, 它的 N 阶 Riesz 变换具体表示为

$$\mathcal{R}^{(N)}\{f\}(\mathbf{x}) = \begin{pmatrix} \mathcal{R}^{(N,0,\dots,0)}\{f\}(\mathbf{x}) \\ \vdots \\ \mathcal{R}^{(N_1,N_2,\dots,N_n)}\{f\}(\mathbf{x}) \\ \vdots \\ \mathcal{R}^{(0,0,\dots,N)}\{f\}(\mathbf{x}) \end{pmatrix} \quad (6.28)$$

它涉及到 $M = \binom{N+n-1}{n-1}$ 个不同的高阶 Riesz 变换分量。高阶 Riesz 变换的傅立叶算子为 $\mathcal{R}^N\{F\}(\mathbf{u}) = \sqrt{\frac{|N|!}{N!}} \frac{(-j\mathbf{u})^N}{\|\mathbf{u}\|^{|N|}} F(\mathbf{u})$, 信号的 N 阶 Riesz 变换将其分解成为 $\binom{N+n-1}{n-1}$ 个

⁶ $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ 是 Gamma 函数, 它的数值计算依赖于其递归性质, $\Gamma(x+1) = x\Gamma(x)$, 它通常是作为阶乘在实数集的延拓, $\Gamma(n) = (n-1)!$ 。

⁷以 f 表示原始信号, f_x, f_y 表示两个虚部分量, 是由原始信号的 Riesz 变换生成。

代表不同方向的分量，生成的分量概略以原始信号维数的 $(d - 1)$ 指数增加。对于 N 维信号，其二阶 Riesz 变换生成 $N + 1$ 个分量，三阶 Riesz 变换生成 $\frac{(N+2)(N+1)}{2}$ 组分量。对于低维信号 $d = 2, 3$ ，这种运算量的增长是完全可以应付。高阶 Riesz 变换同样也具备 Riesz 变换的诸多特性，如平移不变、尺度不变、能量保持等。图 6.9 给出了一组 5 阶 Riesz 变换的频率响应（幅度），可以看出高阶 Riesz 变换的方向选择性。

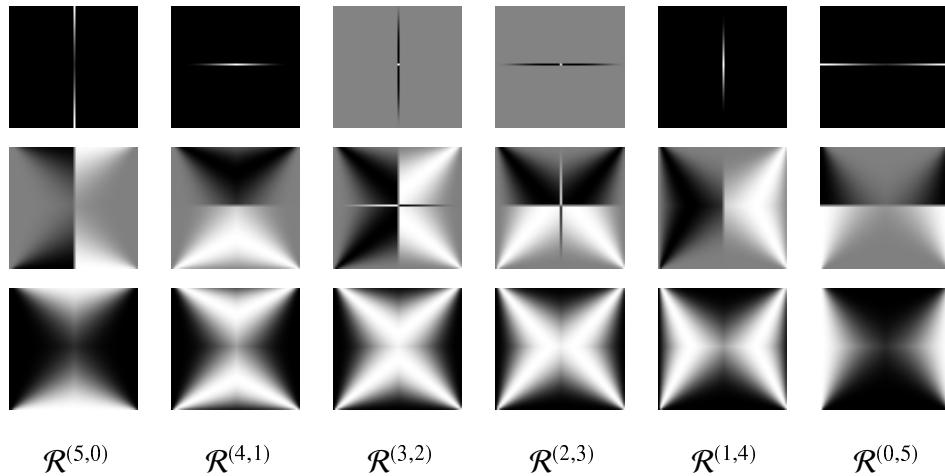


图 6.9 第 5 阶 Riesz 变换的频率响应，从第一行到第三行分分别是实部、虚部和幅度。

3. 方向可控小波：高阶 Riesz 变换能够将任意 ℓ_2 范数能量有限空间 $L_2(\mathbb{R}^n)$ 中的任意框架映射成另外一个 [297]。通常更关注该框架，因为它能够实现信号的小波分解与完美重建，即实现信号的自我可逆变换。

定义 6.10：一族小波 $\{\psi_{i,\mathbf{k}}\}_{\mathbf{k} \in \mathbb{Z}^n, i \in \mathbb{Z}}$ 若满足 $\|f\|_{L_2(\mathbb{R}^n)}^2 = \sum_{i \in \mathbb{Z}} \sum_{\mathbf{k} \in \mathbb{Z}^n} \|\langle \psi_{i,\mathbf{k}}, f \rangle_{\ell_2}\|$ ，即能量保持特性，称其为 $L_2(\mathbb{R}^n)$ 空间的帕塞瓦尔小波框架。

帕塞瓦尔小波框架最显著的特点就是它满足分解与重建规则

$$\forall f \in L_2(\mathbb{R}^n), f = \sum_{i \in \mathbb{Z}} \sum_{\mathbf{k} \in \mathbb{Z}^n} \langle \psi_{i,\mathbf{k}}, f \rangle_{\ell_2} \psi_{i,\mathbf{k}} \quad (6.29)$$

这与第二章介绍的正交基有着类似的性质。根据 (6.29) 可以生成一组小波框架，这些小波框架就是由位置索引和尺度（或称为分辨率）索引 (\mathbf{k}, i) 生成的一族函数 $\{\psi_{i,\mathbf{k}}\}_{\mathbf{k} \in \mathbb{Z}^n, i \in \mathbb{Z}}$ 。第 i 分辨率的基函数是原始分辨率（即 0 分辨率）的函数的尺度调整和膨胀， $\psi_{\mathbf{k},i}(\mathbf{x}) = \det(\mathbf{D})^{\frac{i}{2}} \psi_{\mathbf{k},0}(\mathbf{D}^{-i} \mathbf{x})$ ，式中 \mathbf{D} 是允许的 $n \times n$ 维膨胀矩阵，它的行列式必须为正整数。因此，就可以借助高阶 Riesz 变换

$$\psi_{\mathbf{k},i}^N = \mathcal{R}^N \psi_{\mathbf{k},i} = \sqrt{\frac{N!}{N_1! \dots N_n!}} \mathcal{R}_1^{N_1} \dots \mathcal{R}_n^{N_n} \psi_{\mathbf{k},i} \quad (6.30)$$

$N = (N_1, N_2, \dots, N_n)$, 生成方向可控高阶 Riesz 小波框架。方向可控高阶 Riesz 小波框架继承了许多 Riesz 变换和小波变换的优良特性，能够实现平移不变、尺度不变、方向可控，为信号的分析与处理提供了很大的便利。

6.2.3.2 方向可控小波构建 Grassmann 流形

方向可控 Riesz 小波具备的诸多优良性质启发我们，能否利用生成的特征描述方法应对扩展工作条件的目标识别，包括雷达俯仰角变化、随机噪声污染、装备型号升级等。因此，首先利用方向可控高阶 Riesz 小波框架生成 SAR 图像特征表示模型，描述目标的散射现象。然后将 Riesz 小波框架整体看作是 Grassmann 流形的点，通过构造非线性映射将其嵌入到再生核 Hilbert 空间，利用它们的像构造线性表示模型，通过求解表示系数对未知样本进行重构，根据重构误差输出判决，整个学习的流程如图 6.10 所示。

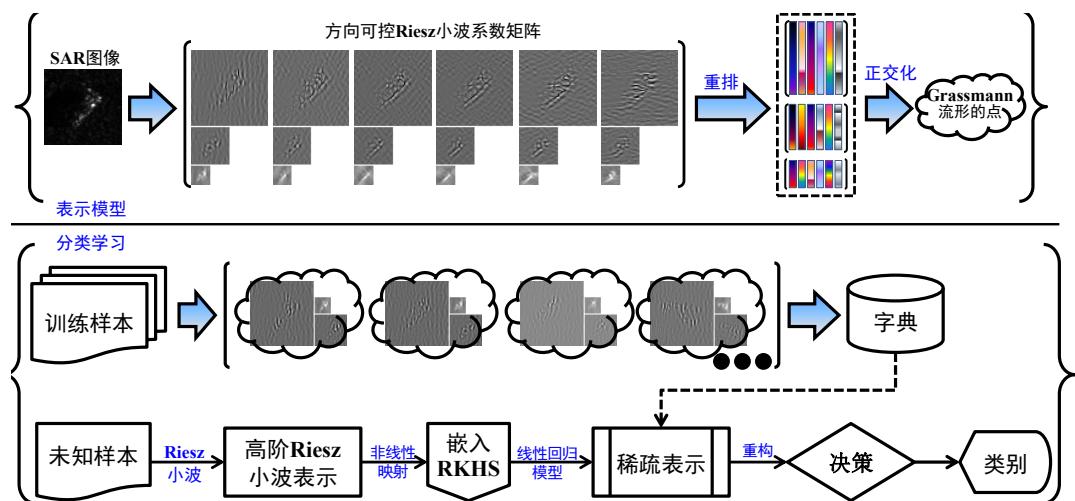


图 6.10 基于 Riesz 小波框架的流形学习流程图，算法由表示模型与分类学习两部分构成，表示模型是指根据 N 阶 J 尺度方向可控 Riesz 小波框架的系数矩阵进行重新排列、能量归一化和施密特正交化等步骤生成表示模型；然后将生成的表示模型整体看作中 Grassmann 流形的点，将其嵌入到再生核 Hilbert 空间，使用训练样本的像对未知样本的像进行线性表示，构建多元线性回归模型，通过求解回归系数（稀疏的）重构未知样本，根据重构误差进行判决。

1. SAR 图像表示模型

根据前文第 6.2.3.1 的论述可知，对于任意二维信号 f ，通过调节位置索引和尺度（或分辨率）索引能够生成一族自我可逆的紧框架，即方向可控小波。图 6.11 给出了一幅 SAR 目标图像的五阶三尺度方向可控 Riesz 小波变换。利用高阶多尺度方向可控 Riesz 小波变换对原始信号进行表示，生成的是一组类似于金字塔的系数矩阵，因此，对于未知样本的分类就转化成为根据训练样本的系数矩阵推测未知样本系数矩阵的归属类别，任意两个样本的相似性度量就成为两簇系数矩阵集

的距离。对于给定的 $H \times W$ 维 SAR 图像， $f \in \mathbb{R}^{H \times W}$ ，它的 N 阶 J 尺度方向可控 Riesz 小波框架表示为 $\{\psi_j^{(N,0)}, \psi_j^{(N-1,1)}, \dots, \psi_j^{(0,N)}\}$ ，其中 $\psi_j^{(\cdot,\cdot)} \in \mathbb{R}^{(2^{-j} \cdot H) \times (2^{-j} \cdot W)}$ 。然后将每个尺度的所有阶数系数矩阵重新排列为 $d^j = (\frac{H}{2^j}) \cdot (\frac{W}{2^j})$ 维向量。将同一尺度内重新排列生成的系数向量组合起来构成 $d \times m$ 维矩阵， $\Psi_j \in \mathbb{R}^{d^j \times m}$ ，其中 $m = N + 1$ ，对该矩阵进行施密特正交化，使其满足 $\Psi_j^T \Psi_j = \mathbf{I}$ 。根据第 6.1.2.1 节可知，最终生成的矩阵满足：任意两组图像 f_1, f_2 有 $\Psi_j(f_1) \sim \Psi_j(f_2)$ 当且仅当 (*if and only if*) $\text{Span}(\Psi_j(f_1)) = \text{Span}(\Psi_j(f_2))$ ，因而可将其整体看作 Grassmann 流形的点。

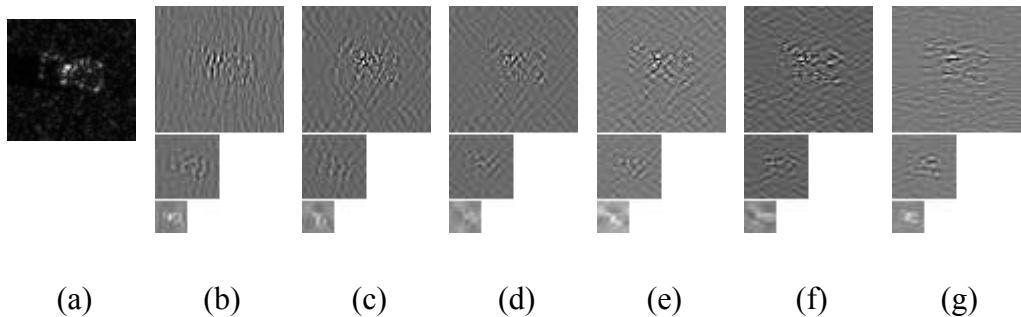


图 6.11 五阶三尺度 Riesz 小波变换示例，(a) 原始 SAR 图像，(b) $\mathcal{R}^{(5,0)}$ ，(c) $\mathcal{R}^{(4,1)}$ ，(d) $\mathcal{R}^{(3,2)}$ ，(e) $\mathcal{R}^{(2,3)}$ ，(f) $\mathcal{R}^{(1,4)}$ ，(g) $\mathcal{R}^{(0,5)}$ 系数矩阵，第一行是 0 尺度系数，后面分别是第一尺度和第二尺度系数。

若给定 Grassmann 距离函数，可衡量每个尺度系数矩阵的距离，将所有 J 尺度的距离值进行累加就生成了两组小波系数矩阵的整体距离度量。为了推测未知样本的类别，一种直观的想法就是逐个计算未知样本与训练样本在 N 阶 J 尺度 Riesz 小波框架下的距离，与未知样本距离最小的训练样本应该与未知样本属于同一类别。对于给定的 n 个训练样本， $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ ，它们分别来自于 L 个不同的样本类别。首先使用 N 阶 J 尺度方向可控小波框架 $\Psi_{N,J}(\cdot)$ 对训练样本进行表示，生成它们的表示模型 $\{\Psi_{N,J}(\mathbf{x}_1), \Psi_{N,J}(\mathbf{x}_2), \dots, \Psi_{N,J}(\mathbf{x}_n)\}$ 。将每个样本每个尺度的系数矩阵重新排列、做施密特征正交化，每个类别的训练样本的表示模型构成一个 Grassmann 流形，其中的样本的系数矩阵整体看作是 Grassmann 流形的点。对于未知样本 \mathbf{y} ，类似的方式生成它的表示模型 $\Psi_{N,J}(\mathbf{y})$ 。为衡量两组表示模型间的相似性，将它们每个尺度的距离值进行累加，生成整体相似度量，

$$d_G(\Psi_{N,J}(\mathbf{y}), \Psi_{N,J}(\mathbf{x}_k)) = \sum_{j=0}^{J-1} d_j(\Psi_j(\mathbf{y}), \Psi_j(\mathbf{x}_k)) \quad (6.31)$$

其中 $d_j(\cdot, \cdot)$ 表示两组表示模型的第 j 尺度系数矩阵之间的 Grassmann 距离。根据度量公理 6.1 可知，多个距离函数的累加仍然满足度量公理，因而是个合法的度量。图 6.12 以图例的形式描述了 Grassmann 流形两点的距离形成过程。根据 (6.31) 定

义的累加距离规则，就可逐个比对未知样本与训练样本的表示模型的距离，判定其类别属性。

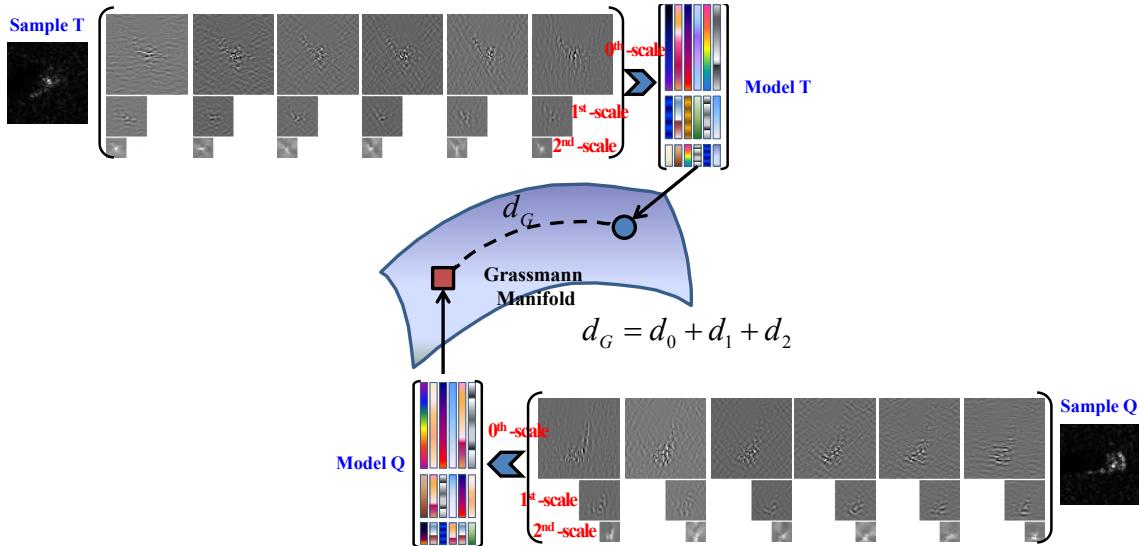


图 6.12 SAR 图像 Riesz 表示模型距离度量，五阶 3 尺度方向可控 Riesz 小波表示模型的距离定义为三个尺度内的 Grassmann 流形距离的累加， $d_G = d_0 + d_1 + d_2$ 。

2. Grassmann 流形的核稀疏表示

借助方向可控 Riesz 小波框架，前文提出一种 SAR 图像表示模式，该表示模式包含一簇不同维度的矩阵，由于该模式整体看作是流形的点，直接进行分类学习不太现实。为此，构造非线性映射将表示模式投影到再生核 Hilbert 空间 [295]，关于再生核 Hilbert 空间，请参照论文第五章第 5.1.2 节的论述。然后在 Hilbert 核空间中应用核分析方法进行学习。

若给定非线性映射 $\phi(\cdot) : \mathcal{M} \mapsto \mathcal{H}$ ，将 Grassmann 流形嵌入到再生核 Hilbert 空间。利用该映射把表示模型第 j 尺度的系数矩阵 Ψ_j 投影到 Hilbert 核空间。对于一组 SAR 图像 $\mathbf{x}_r, \mathbf{x}_l$ 的表示模型， $\Psi_{N,J}(\mathbf{x}_r), \Psi_{N,J}(\mathbf{x}_l)$ ，定义第 j 尺度的核函数 $\kappa_j(\Psi_j(\mathbf{x}_r), \Psi_j(\mathbf{x}_l))$ 实现内积运算。根据第五章第 5.1.3 节正定核函数的性质可知，任意个正定核函数之各仍为正定核函数。因此，将所有 J 个尺度定义的核函数相加，其结果是个合法的核函数

$$\kappa(\Psi_{N,J}(\mathbf{x}_r), \Psi_{N,J}(\mathbf{x}_l)) = \sum_{j=0}^{J-1} \beta_j \kappa_j(\Psi_j(\mathbf{x}_r), \Psi_j(\mathbf{x}_l)) \quad (6.32)$$

其中 $\beta_j \geq 0, j = 0, \dots, J - 1$ 是权重系数⁸。将表示模型整体看作是 Grassmann 流形的点，每个尺度内的核函数是合法的 Grassmann 核，所以公式 (6.32) 的全局核函

⁸设置权重系数是为了量化表示模式中不同尺度的系数矩阵的重要程度，在没有先验知识的情况下通常认为所有尺度的系数矩阵重要程度相同。

数也是合法的 Grassmann 核函数，详见定义 6.9。定义核函数之后，就可以通过构造核回归模型对未知样本进行线性表示，通过求解最优回归系数进行分类，具体方法参见第五章第 5.2 节的相关描述。

对于给定的一组训练样本 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ ，首先计算它们的 N 阶 J 尺度方向可控 Riesz 小波框架，生成表示模型。然后使用定义在流形 \mathcal{M} 的非线性映射 $\phi(\cdot)$ 将表示模型投影到再生核 Hilbert 空间，生成的 Hilbert 核空间的像组合构成一个字典 $\mathcal{D} = [\phi(\Psi_{N,J}(\mathbf{x}_1)), \phi(\Psi_{N,J}(\mathbf{x}_2)), \dots, \phi(\Psi_{N,J}(\mathbf{x}_n))]$ ，然后以类似的方式将未知样本的表示模型 $\Psi_{N,J}(\mathbf{y})$ 投影到再生核 Hilbert 空间， $\phi(\Psi_{N,J}(\mathbf{y}))$ 。将未知样本的像表示成为字典基原子的线性组合，构成核回归模型

$$\phi(\Psi_{N,J}(\mathbf{y})) = \phi(\Psi_{N,J}(\mathbf{x}_1))\alpha_1 + \phi(\Psi_{N,J}(\mathbf{x}_2))\alpha_2 + \dots + \phi(\Psi_{N,J}(\mathbf{x}_n))\alpha_n = \mathcal{D}\alpha \quad (6.33)$$

其中 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n] \in \mathbb{R}^n$ 是待求解的回归系数。从 (6.33) 可以看出，训练样本的像 $\{\phi(\Psi_{N,J}(\mathbf{x}_r))\}_{r=1}^n$ 是核回归模型的回归参量或预测量，未知样本的像 $\phi(\Psi_{N,J}(\mathbf{y}))$ 是对应的回归响应，它们之间是通过一组标量数值即回归系数 α 联系起来。为了方便求解，将未知样本和训练样本在 Hilbert 空间的像分别简写为 Φ_y 和 ϕ_{x_r} 。

根据第五章第 5.2 节的论述可知，通常期望回归系数 α 足够稀疏，只有很少一部分训练样本的像参与未知样本的像的表示，参与未知样本线性表示的训练样本的回归系数不为零，其余位置的系数全部为零。参照前文，采用 ℓ_1 范数最小化的最优化策略寻找最稀疏的回归系数

$$\min_{\alpha} \|\alpha\|_1 \text{ s.t. } \|\phi_y - \mathcal{D}\alpha\|_2^2 < \varepsilon \quad (6.34)$$

其中 ε 是允许的误差容限。将 (6.34) 式转化为等价的无约束优化

$$\min_{\alpha} \{\|\phi_y - \mathcal{D}\alpha\|_2^2 + \lambda \|\alpha\|_1\} \quad (6.35)$$

其中 λ 是正则化参数，用于稀疏度和保真度之间的折衷。为了求解 (6.35)，先将误差项展开 $\|\phi_y - \mathcal{D}\alpha\|_2^2 = (\phi_y - \mathcal{D}\alpha)^T (\phi_y - \mathcal{D}\alpha) = \phi_y^T \phi_y - 2\phi_y^T \mathcal{D}\alpha + \alpha^T \mathcal{D}^T \mathcal{D}\alpha$ ，然后使用 (6.32) 式定义的 Grassmann 核函数对误差项进行化简，其中 $\phi_y^T \phi_y = \langle \phi_y, \phi_y \rangle = \kappa(\Psi_{N,J}(\mathbf{y}), \Psi_{N,J}(\mathbf{y}))$ 是常数项，不参与回归系数的最优求解，剩余两项即 Hilbert 空间相似向量和格莱姆矩阵分别为

$$\phi_y^T \mathcal{D} = \begin{bmatrix} \langle \phi_y, \phi_{x_1} \rangle \\ \langle \phi_y, \phi_{x_2} \rangle \\ \vdots \\ \langle \phi_y, \phi_{x_n} \rangle \end{bmatrix} = \begin{bmatrix} \kappa(\Psi_{N,J}(\mathbf{y}), \Psi_{N,J}(\mathbf{x}_1)) \\ \kappa(\Psi_{N,J}(\mathbf{y}), \Psi_{N,J}(\mathbf{x}_2)) \\ \vdots \\ \kappa(\Psi_{N,J}(\mathbf{y}), \Psi_{N,J}(\mathbf{x}_n)) \end{bmatrix} \quad (6.36)$$

和

$$\mathcal{D}^T \mathcal{D} = \begin{bmatrix} \langle \phi_{\mathbf{x}_1}, \phi_{\mathbf{x}_1} \rangle & \langle \phi_{\mathbf{x}_1}, \phi_{\mathbf{x}_2} \rangle & \cdots & \langle \phi_{\mathbf{x}_1}, \phi_{\mathbf{x}_n} \rangle \\ \langle \phi_{\mathbf{x}_2}, \phi_{\mathbf{x}_1} \rangle & \langle \phi_{\mathbf{x}_2}, \phi_{\mathbf{x}_2} \rangle & \cdots & \langle \phi_{\mathbf{x}_2}, \phi_{\mathbf{x}_n} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \phi_{\mathbf{x}_n}, \phi_{\mathbf{x}_1} \rangle & \langle \phi_{\mathbf{x}_n}, \phi_{\mathbf{x}_2} \rangle & \cdots & \langle \phi_{\mathbf{x}_n}, \phi_{\mathbf{x}_n} \rangle \end{bmatrix} \quad (6.37)$$

从结果来看，非线性映射 $\phi(\cdot)$ 的具体形式未知，但并不影响回归系数的求解，只需要计算再生核 Hilbert 空间像与像的内积。若选定合适的 Grassmann 核函数，就可以求出最优回归系数 $\hat{\alpha}$ ，然后根据重构误差判决未知样本的类别属性，

$$\min_{k=1,\dots,L} \{ \|\phi_{\mathbf{y}} - \mathcal{D}\delta_k(\hat{\alpha})\|_2^2 \} \quad (6.38)$$

其中 $\delta_k(\cdot)$ 用于挑出第 k 类的回归系数，将剩余系数全部置零。整个分类学习的流程见图 6.13 和算法 6.1 所示。

算法 6.1 基于方向可控 Riesz 小波的核稀疏表示

已知: $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$: 训练集和未知样本 \mathbf{y} ;

κ : Grassmann 核函数;

λ : 稀疏度与保真度的正则化参数;

求: 未知样本 \mathbf{y} 的类别。

- 1: 未知样本 \mathbf{y} 的 N 阶 J 尺度方向可控 Riesz 小波框架表示 $\Psi_{N,J}(\mathbf{y})$;
 - 2: 训练样本的 N 阶 J 尺度方向可控 Riesz 小波框架表示 $\{\Psi_{N,J}(\mathbf{x}_k)\}_{k=1}^n$;
 - 3: **for** $r \leftarrow 1, n$ **do**
 - 4: 计算未知样本与第 r 个训练样本的 Riesz 小波表示核函数，构成相似向量 $\phi_{\mathbf{y}} \mathcal{D}$ 的第 r 个元素， $\kappa(\Psi_{N,J}(\mathbf{y}), \Psi_{N,J}(\mathbf{x}_r)) \rightarrow [\phi_{\mathbf{y}} \mathcal{D}]_r$;
 - 5: **for** $l \leftarrow 1, n$ **do**
 - 6: 计算第 r 个和第 l 个训练样本的 Riesz 小波表示核函数，构成格莱姆矩阵的第 (r, l) 个元素， $\kappa(\Psi_{N,J}(\mathbf{x}_r), \Psi_{N,J}(\mathbf{x}_l)) \rightarrow [\mathcal{D}^T \mathcal{D}]_{r,l}$;
 - 7: **end for**
 - 8: **end for**
 - 9: 通过 ℓ_1 范数最小化 (6.35) 计算最优回归系数 $\hat{\alpha}$;
 - 10: 利用判决规则 (6.38) 进行判决: $\min_k \{ \|\phi_{\mathbf{y}} - \mathcal{D}\delta_k(\hat{\alpha})\|_2^2 \}$ 。
-

6.3 实验验证

本节借助 MSTAR SAR 实测数据验证基于流形学习的分类框架。实验设置与第四章和第五章相似，先进行基础验证实验，再进行扩展工作条件的实验比较，

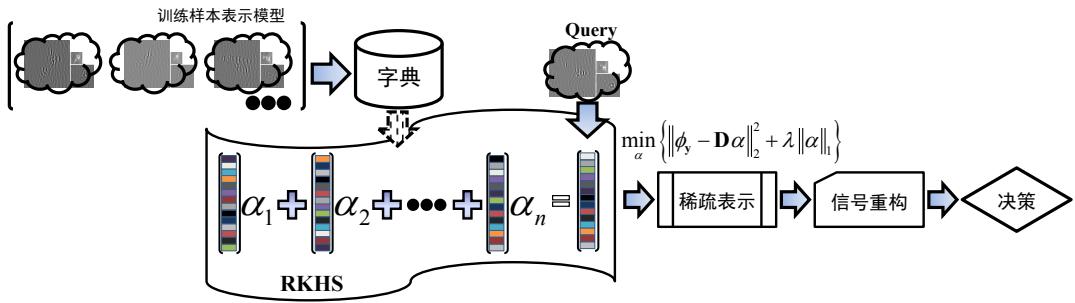


图 6.13 基于方向可控 Riesz 小波框架的稀疏表示，将未知样本和训练样本的 Riesz 小波表示投影到再生核 Hilbert 空间，其中训练样本的像用于生成字典，未知样本的像表示为字典基原子的线性组合，通过构造的 Grassmann 核函数实现问题的求解，根据求解得到的回归系数重构未知样本的像，使用重构误差进行判决。

逐渐增加实验难度。实验中的基准比较算法包括经典算法和论文第四章、第五章提出的分类学习方案，详见表 6.2 所列。

表 6.2 流形学习实验的基准比较算法。

缩写	算法名称	输入特征
SVM	Support Vector Machine with Linear Kernel [65]	原始强度值
SRC	Sparse Representation-based Classification [95, 156]	原始强度值
KSVM	Support Vector Machine with nonlinear Kernel [65]	原始强度值
KSRC	Kernel Sparse Representation-based Classification [97, 264]	原始强度值
MSRC	Sparse Representation on Feature-level Fusion (第4.1.1节)	增强特征向量
TJSR	Multi-task Joint Sparse Representation (第 4.1.3 节)	单演信号特征
CKLR	Linear Representation of Composite Kernel (第 5.3.1 节)	单演信号特征
SRC ^R	Sparse Representation of Log-Covariance (第6.2.1节)	协方差矩阵
KLSF	Embedding Riemannian Manifolds into RKHS 第6.2.1节)	协方差矩阵
kNN ^G	kNN on Grassmann Manifolds (第6.2.2节)	单演信号系数
KSR ^{G1}	Kernel Sparse Representation of Monogenic Signal (第6.2.2 节)	单演信号系数
KSR ^{G2}	Kernel Sparse Representation of Monogenic Signal (第6.2.2 节)	单演信号系数
kNN ^R	kNN of Riesz Wavelet Frames (第6.2.3节)	Riesz 小波框架
KSR ^R	Kernel Sparse Representation of Riesz Wavelet Frames (第6.2.3节)	Riesz 小波框架

6.3.1 基础实验证

首先是结合实测数据对提出的黎曼流形和 Grassmann 流形距离度量的合理性和有效性进行验证，之后与经典方法进行比较。为了提高计算效率，区域协方差描述子的区域 Rg 指定为目标中心的 64×64 像素的区域，周围的背景杂波直接舍弃；

在执行 Grassmann 流形学习时，考虑到计算负担论文仅使用目标中心的 64×64 像素，对中心区域图像做降采样处理，若以采样因子 $\frac{1}{4}$ 对原始图像做降采样，然后执行 3 尺度单演信号表示，形成 Grassmann 流形 $\mathcal{G}(m, d)$ ，其中 $m = 256, d = 9$ 。

6.3.1.1 协方差矩阵描述子的数值实现

论文第 6.2.1.2 节提出协方差矩阵描述子，用于将单演信号多尺度表示生成的不同类型分量进行融合，生成特征向量。围绕协方差描述子的数值计算，通过借鉴现有研究成果，论文提出三种实现方案，记为 Mode 1, Mode 2 和 Mode 3，其中 Mode 1 是仅使用单演信号多尺度表示生成的分量矩阵，Mode 2 加入了原始图像的强度信息，实现“原始信号 + 单演信号”组合，Mode 3 进一步加入了位置索引信息，实现“位置索引 + 原始信号 + 单演信号”组合，详见表 6.3 所示。

表 6.3 三种协方差描述子计算模式。

模式	特征组合	协方差矩阵维数
Mode 1	$[A_1, \varphi_1, \theta_1, A_2, \varphi_2, \theta_2, \dots, A_S, \varphi_S, \theta_S]$	$d = 3S$
Mode 2	$[\mathbf{x}, A_1, \varphi_1, \theta_1, A_2, \varphi_2, \theta_2, \dots, A_S, \varphi_S, \theta_S]^a$	$d = 3S + 1$
Mode 3	$[\mathbf{r}, \mathbf{c}, \mathbf{x}, A_1, \varphi_1, \theta_1, A_2, \varphi_2, \theta_2, \dots, A_S, \varphi_S, \theta_S]^b$	$d = 3S + 3$

^a \mathbf{x} 表示原始图像的强度值。

^b \mathbf{r} 和 \mathbf{c} 分别表示原始图像的行和列位置索引。

为了观察不同模式协方差矩阵描述子的实际表现，论文进行了一组四类目标的识别实验，包括 BMP2、T72、BTR60 和 T62，实验设置与前文相同， 17° 雷达俯仰角标准型号目标成像样本用于训练， 15° 雷达俯仰角变体型号目标成像样本用于测试，实验结果如表 6.4 所示，其中 Euclidean 表示将传统的欧氏空间学习方法应用于单演信号协方差矩阵描述子，生成的协方差矩阵描述子取出它的 $\frac{d(d+1)}{2}$ 个独立元素组成特征向量，构成稀疏表示模型，执行分类学习；Riemannian 表示将单演信号协方差矩阵看作是流形的元素，使用矩阵对数算子（面向流形）将其映射到切向空间，利用对数协方差矩阵的 $\frac{d(d+1)}{2}$ 个独立元素构成特征向量，执行分类学习；Hilbert 表示将对称正定矩阵所构成的黎曼流形嵌入到 Hilbert 空间，再执行分类学习，具体方法详见论文第 6.2.1.2 节所述。

从实验结果可以看出，三种协方差矩阵的计算模式表现从劣到优依次排列为 Mode1 < Mode2 < Mode3，Mode 3 的识别率最高，Mode 1 的识别率最低，这是因为 Mode 3 即使用了单演信号多尺度表示生成的分量，也使用了原始图像的强度信息与坐标索引信息，信息利用更充分。从三种不同类型空间的分类学习方案比较来看，欧氏空间中的分类学习方法远远逊色于其余两种方法，三种数值实现模式

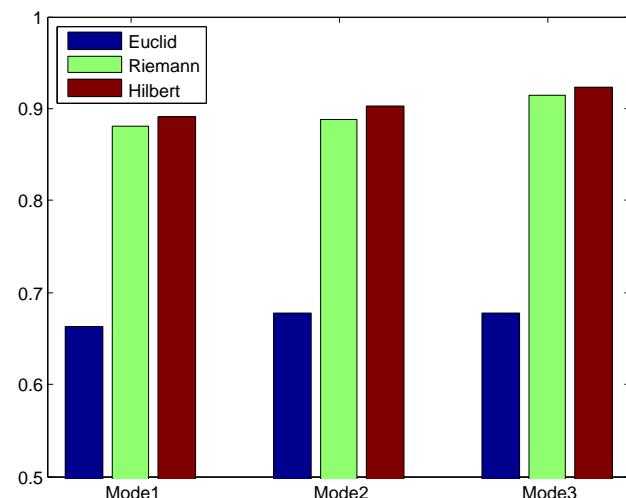
表 6.4 协方差矩阵描述子计算模式，表中列出了协方差矩阵描述子三种计算模式在欧氏空间、非欧流形及 Hilbert 空间的分类学习。

方法	Mode 1	Mode 2	Mode 3
Euclidean ^a	0.6637	0.6782	0.6782
Riemann ^b	0.8804	0.8884	0.9141
Hilbert ^c	0.8908	0.9029	0.9238

^a协方差矩阵特征执行稀疏表示分类。

^b对数协方差矩阵特征稀疏表示分类。

^c协方差矩阵的像进行稀疏表示分类。



在欧氏空间的分类学习正确率为 0.6637、0.6782、0.6782，相同模式在非欧流形的分类学习正确率为 0.8804、0.8884、0.9141，再生核 Hilbert 空间的分类学习的正确率为 0.8908、0.9029、0.9238，欧氏空间的分类学习方法比流形分类学习正确率平均大约低 22.09%，比再生核 Hilbert 空间的分类学习正确率平均大约低 23.25%。对实验结果做进一步分析：单演信号协方差矩阵描述子是对称正定矩阵，对数乘和加法两种运算结构并不封闭，若仍沿用欧氏空间的学习方法进行处理显然并不合理，这就不难解释为什么实验中 Euclidean 的正解率远低于其余两类；若赋予协方差矩阵空间黎曼度量就构成流形，借助矩阵对数算子能将流形的元素转换到对应的切向空间，切空间是标准的向量空间，对加法和数乘封闭，因而可以使用传统的学习方法进行分类识别。通过实验结果分析，论文后续的比较实验将采用 Mode 3 生成协方差矩阵特征。

6.3.1.2 Grassmann 流形距离函数比较

论文第 6.1.2.2 节介绍了多种 Grassmann 流形距离度量，本节实验将对这些距离函数进行直观地比较。仍然以 BMP2、T72、BTR60、T62 四类目标的识别实验为例，借助 Grassmann 流形的最近邻分类进行识别，实验目的并不是为了简单地比较分类正确率，而是观察不同 Grassmann 度量的特点。实验设置与前文相同，表 6.5 列出了采样因子为 $\frac{1}{8}$ 和 $\frac{1}{4}$ 构成的两种尺寸 Grassmann 流形实验结果。可以看出，Fubini-Study 距离与比奈特—柯西距离的识别正确率相等，因为它们两者之间是单调相关的；最大相关系数与最小相关系数两种距离的表现比其余几种的距离差，以 $\frac{1}{4}$ 采样因子为例，它们的识别率仅为 0.6100 和 0.7921，而其余方法的识别率都在 0.9 左右，这是因为这两种距离仅仅涉及单个主角 (θ_1 或 θ_m) 的计算，很容易受到噪声的影响，因而表现不稳健；尽管最大相关系数与最小相关系数都

只使用一种主角进行度量，它们的正确率也有明显的差别，最小相关系数大约优于最大相关系数 20% 左右，这是因为若子空间集中在某个点附近并且有非零交集时，大部分点的最大相关系数都走近于零；弧长距离、投影距离始终表现比较稳健，识别率也比较令人满意。根据得出的结论论文后续的比较实验将选取投影距离和 Binet-Cauchy 距离度量。

表 6.5 Grassmann 流形距离函数比较。

(a) 使用采样因子 $\rho = \frac{1}{8}$ 生成的 Grassmann 流形 $\mathcal{G}(64, 9)$ 。

距离函数	d_{Arc}	d_{Pj}	d_{BC}	d_{Max}	d_{Min}	d_{Pc}	d_{FS}
最近邻分类	0.8668	0.8844	0.8547	0.5947	0.7255	0.8708	0.8547
k 近邻分类	0.8941	0.8925	0.8700	0.6132	0.7456	0.8973	0.8700

(b) 使用采样因子 $\rho = \frac{1}{4}$ 生成的 Grassmann 流形 $\mathcal{G}(256, 9)$

距离函数	d_{Arc}	d_{Pj}	d_{BC}	d_{Max}	d_{Min}	d_{Pc}	d_{FS}
最近邻分类	0.9021	0.8941	0.8957	0.6100	0.7921	0.9053	0.8957
k 近邻分类	0.9133	0.9077	0.9101	0.6172	0.8226	0.9109	0.9101

6.3.1.3 不同单演信号分解构成的 Grassmann 流形

根据论文第三章介绍可知，单演信号定义为原始信号（二维 \mathbb{R}^2 ）与它的 Riesz 变换结果的线性组合，其中原始信号是实部分量，Riesz 变换生成的信号是虚部分量，虚部又包含两个不同方向的分量，因而可将其表示为实部分量与两个方向的虚部分量，同样可转换为极坐标的局部幅度、局部相位和局部方位，详见第三章图 3.3 所示。为了观察不同的信号表示方式对于分类学习结果的影响，论文开展一组四类目标的分类识别实验，仍然选取 BMP2、T72、BTR60、T62 四类军事目标，训练样本与测试样本的选取以前文相同，分别使用笛卡尔坐标与极坐标表示的单演信号分量矩阵，即实部与两个虚部组合、局部幅度、局部相位与局部方位组合，构成维数相同的 Grassmann 流形 $\mathcal{G}(d, m)$ ，其中 d 是降采样后的矩阵维数， $m = 3S$ 是系数矩阵总数。实验结果如表 6.6 所示。

表 6.6 笛卡尔坐标表示和极坐标表示对分类学习的影响 ($\rho = \frac{1}{4}$)。

距离函数	d_{Arc}	d_{Pj}	d_{BC}	d_{Max}	d_{Min}	d_{Pc}	d_{FS}
笛卡尔坐标	0.8949	0.8981	0.8828	0.7014	0.8475	0.8989	0.8828
极坐标	0.9133	0.9077	0.9101	0.6172	0.8226	0.9109	0.9101

可以看出：最大相关系数和最小相关系数受信号分解方式的影响比较大，其中笛卡尔坐标分解的识别正确率比极坐标分解的识别正确率大约高出近 10%，其余几种距离函数没有明显变化；横向比较，最大相关系数和最小相关系数的识别率远低于其余的距离函数，这与上一组实验的结论类似；弧长距离（测地距离）取得最高的识别率，0.9133。论文重点讨论的投影度量和 BC 度量两个距离函数，极坐标表示优于笛卡尔坐标表示，特别是 BC 距离，前者高于后者 2.73%。原因在于极坐标表示的三个局部分量相互正交，不存在冗余信息。根据实验结果，后续实验将选择极坐标表示构建 Grassmann 流形。

6.3.1.4 高阶 Riesz 小波构成 Grassmann 流形的影响因素

为了分析高阶 Riesz 小波构成 Grassmann 流形对分类学习的影响，论文组织 BMP2、T72、BTR60 和 T62 四类目标识别实验，实验设置与前文相同，重点分析不同阶数 Riesz 变换、不同尺度 Riesz 小波的流形学习结果。

1. 不同尺度 Riesz 小波：对于不同尺度的方向可控 Riesz 小波框架构成的 Grassmann 流形，论文定义全局距离为尺度内 Grassmann 距离的累加，故不同尺度小波框架必然导致全局距离度量的不同。为了观察尺度变化对于分类学习结果的影响，固定 Riesz 变换为五阶 $N = 5$ ，调整尺度从 $J = 0$ 到 $J = 2$ ，分别对应 $\{64 \times 64\}$, $\{64 \times 64, 32 \times 32\}$, $\{64 \times 64, 32 \times 32, 16 \times 16\}$ 三种小波框架。实验结果如表 6.7 和图 6.14 所示。随着尺度的增加，分类正确率逐渐提高，三种距离函数都是在 $J = 2$ 尺度 Riesz 小波框架下取得最好的效果。 $J = 2$ 尺度的 Riesz 小波框架 Binet-Cauchy 距离度量的分类正确率为 0.8957，稍稍优于其余两种距离度量。同样尺度的核稀疏表示分类正确率优于最近邻分类学习，两种 Grassmann 核函数构成的核稀疏表示分类正确率分别比相应的距离度量最近邻分类平均高出 3.34%、3.29%。从中可以看出，每个尺度的 Riesz 小波系数都含有一定的鉴别信息，将不同尺度包含的鉴别信息进行累积能提高识别正确率⁹。

表 6.7 Riesz 小波尺度对流形学习的影响，表中列出三种尺度 Riesz 小波的学习正确率。

(a) 最近邻分类			(b) 核稀疏表示				
度量函数	$J = 0$	$J = 1$	$J = 2$	核函数	$J = 0$	$J = 1$	$J = 2$
投影度量	0.8307	0.8604	0.8836	投影核	0.8684	0.8788	0.9278
BC 度量	0.8202	0.8644	0.8957	BC 核	0.8620	0.8941	0.9230
弧长度量	0.8283	0.8555	0.8844				

⁹ 鉴别信息的累积主要有两种方式，1) 通过尺度间局部距离的累加生成全局距离，见公式 (6.31)；2) 通过尺度间 Grassmann 核函数的累加生成全局 Grassmann 核函数，见公式 (6.32)。

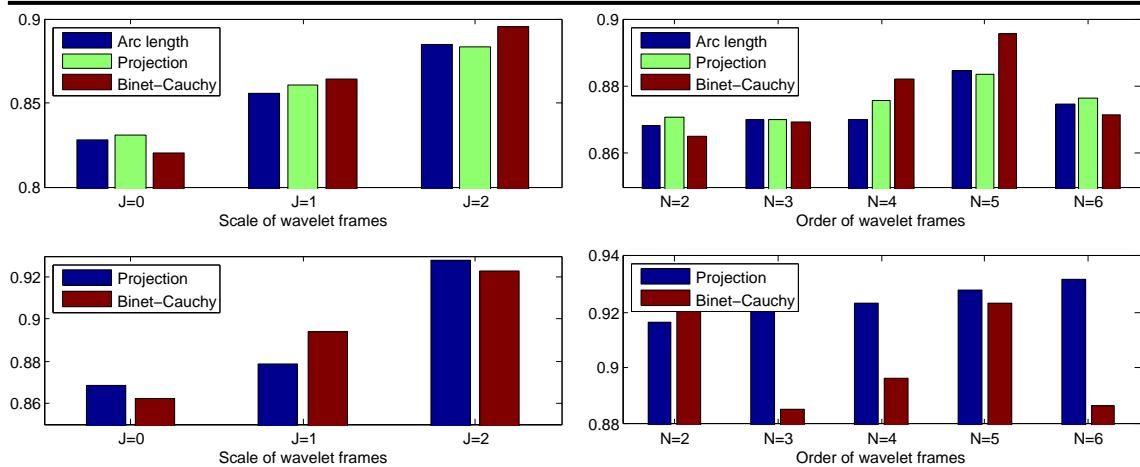


图 6.14 流形学习受 Riesz 小波尺度的影响。图 6.15 流形学习受 Riesz 变换阶数的影响。

2. 不同阶数 Riesz 变换：前文定义的 Grassmann 距离是与 Riesz 变换的阶数有关，不同阶数 Riesz 变换构成的 Grassmann 流形 $\mathcal{G}(d, m)$ 也发生变化¹⁰。为了观察广义 Riesz 变换的阶数对于分类学习结果的影响，分别生成 $N = 2$ 到 $N = 6$ 阶 Riesz 变换，据此构造 $J = 2$ 尺度 Riesz 小波框架，进行四类目标的分类实验，结果如表 6.8 和图 6.15 所示。

表 6.8 Riesz 变换阶数对流形学习的影响，表中列出 5 种阶数 Riesz 变换的学习正确率。

(a) 最近邻分类					
度量函数	$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$
投影度量	0.8708	0.8700	0.8756	0.8836	0.8764
BC 度量	0.8652	0.8692	0.8820	0.8957	0.8716
弧长度量	0.8684	0.8700	0.8700	0.8844	0.8748

(b) 核稀疏表示					
核函数	$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$
投影核	0.9165	0.9213	0.9230	0.9278	0.9318
BC 核	0.9238	0.8852	0.8965	0.9230	0.8868

对于最近邻分类学习，三种 Grassmann 距离函数的分类正确率随着阶数的提升先升后降， $N = 5$ 阶 Riesz 变换取得了最好的效果，正确率分别为 0.8844、0.8836、0.8957， $N = 6$ 阶 Riesz 变换的正确率低于 $N = 5$ 阶。三种距离的表现优劣并未呈现出一致的趋势，BC 度量取得了最高的正确率。对于核稀疏表示分类学习，两种 Grassmann 核函数的表现也有明显的差别，投影核随着 Riesz 变换阶数的

¹⁰二维信号的高阶 Riesz 变换的系数矩阵为 $N + 1$ 项，即生成的 Grassmann 流形维数 $m = N + 1$ 。

提升，其正确率也逐渐提高，但是提高的幅度较小；BC 核函数在 $N = 2$ 和 $N = 5$ 两个阶数的 Riesz 变换正确率优于其余三种情况，没有呈现出一致的变化趋势。

根据两小节的验证实验，论文将选取 $J = 2$ 尺度 $N = 5$ 阶方向可控 Riesz 小波框架来构建 Grassmann 流形，使用该设置与经典的算法进行比较。

6.3.2 扩展工作条件实验比较

前面的验证实验用于观察 SAR 图像特征表示受不同变化因素的影响，通过实验合理选择参数，提高实验效果。本节将会展开多组扩展工作条件的目标分类识别实验，用于将本章提出的三个分类学习方案与经典的算法进行比较，考察算法的有效性与稳健性。实验设置主要包括三部分：目标配置变化、雷达俯仰角变化和随机噪声污染，分别从不同侧面考察流形学习的可靠性。

6.3.2.1 目标配置变化

目标配置变化的具体内容包括装备的版本型号升级、武器系统配置、加装防护装置等，详细说明见论文第一章第 1.2.2 节。为了考察流形学习受目标配置变化的影响，选取 BMP2、T72、BTR60 和 T62 四类目标进行分类识别实验，其中 BMP2 与 BTR60 是装甲输送车，T72 与 T62 是苏制主战坦克，它们具有十分相似的散射现象，因而分类识别较为困难。训练样本和测试样本的选取与前文相同， 17° 雷达俯仰角的标准型号成像数据用于训练， 15° 雷达俯仰角的变体型号成像数据用于测试，即训练样本与测试样本的目标配置、雷达俯仰角均发生变化。参与实验的流形学习方法包括 SRC R 、kNN G 、KLMC、KSR G_1 、KSR G_2 和 KSR R ，实验结果如表 6.9 所示，其中表 6.9 (a) 是向量空间的分类学习，借助矩阵对数算子将流形的元素投影到切向空间，表 6.9 (b) 是再生核 Hilbert 空间的分类学习，借用非线性映射将黎曼流形嵌入到再生核 Hilbert 空间。从实验结果可以看出：

表 6.9 流形学习受目标配置变化的影响，表中给出欧氏空间与 Hilbert 空间目标配置变化的分类精度。

(a) 向量空间的分类学习

SVM	SRC	MSRC	SRC R	TJSR	kNN G	kNN R
0.8443	0.8764	0.8868	0.9117	0.9117	0.9133	0.8957

(b) 再生核 Hilbert 空间的分类学习

KLRC	KSVM	KSRC	CKLR	KLSF	KSR G_1	KSR G_2	KSR R
0.9165	0.8876	0.9181	0.9325	0.9238	0.9270	0.9390	0.9318

- 空域信息融合与流形学习的对比：黎曼流形学习的分类方案优于特征级融合方法。本章提出三种黎曼流形的表示方法 SRC^R 、 kNN^G 和 kNN^R 的识别正确率分别为 0.9141、0.9133 和 0.8957，MSRC 的识别率为 0.8868，黎曼流形的表示方法高出 MSRC 纸头 2.73%、2.65%、0.89%，它们与多特征多任务联合稀疏表示取得相似的正确率。特征级融合方法在信息合并过程中通常会导致信息损失，流形学习通过挖掘掩藏在数据内部的非线性几何特性能够弥补这种不足。
- 对称正定矩阵空间与 Grassmann 流形的对比：Grassmann 流形学习的分类方法与对称正定矩阵空间的分类方法正确率相仿， kNN^G 、 kNN^R 的正确率为 0.9133 和 0.8957，相应的对数协方差矩阵稀疏表示 SRC^R 的正确率为 0.9141，需要声明的是对称正定矩阵空间采用了稀疏表示的学习方法，而 Grassmann 流形仅仅采用了简单的最近邻分类学习，因而二者的比对结果并不客观，不能得出有说服力的结论。
- 再生核 Hilbert 空间与黎曼流形的对比：为了实现黎曼流形的分类学习，论文通过构造非线性映射将其嵌入到再生核 Hilbert 空间，实现线性可分，结果表明再生核 Hilbert 空间的分类学习方法明显优于其它方法。KLSF、 KSR_2^G 、 KSR_2^G 、 KSR^R 的分类正确率分别 0.9238、0.9270、0.9390 和 0.9318，高于其它的分类算法，合成核学习 CKLR、Binet-Cauchy 核函数 KSR^{G_2} 、Riesz 小波框架的正确率高于 0.93，Binet-Cauchy 核函数的表现最好。

6.3.2.2 俯仰角度变化

雷达俯仰角、入射角、投影平面方向角等参数的变化通常会显著改变 SAR 图像的电磁散射现象，作为一种典型的影响因素，雷达俯仰角对于 SAR 图像目标散射现象产生重要的作用。为了考察俯仰角变化对于分类学习结果的影响，论文组织三类目标分类识别实验，主要选取 2S1、BRDM2、ZSU23/4 三类军事目标于不同雷达俯仰角的成像数据，包括 17° 、 30° 和 45° 三种雷达俯仰的成像样本，其中 BRDM2 和 ZSU23/4 还包括 30° 和 45° 俯仰角的连接体变体成像数据。论文使用 17° 俯仰角标准目标成像样本进行学习，使用 30° 和 45° 俯仰角连接体和变体目标成像样本进行测试，数据的详细说明请参见第二章第 2.4.3.1 节。流形学习的具体设置见前文分析，学习方法的分类正确率随雷达俯仰角变化的降幅曲线如图 6.16 所示。

- 第一个实验场景测试雷达俯仰角度的微小变化：使用 30° 雷达俯仰角的成像样本进行测试，即训练样本与测试样本之间存在着 13° 俯仰角差异。从实验

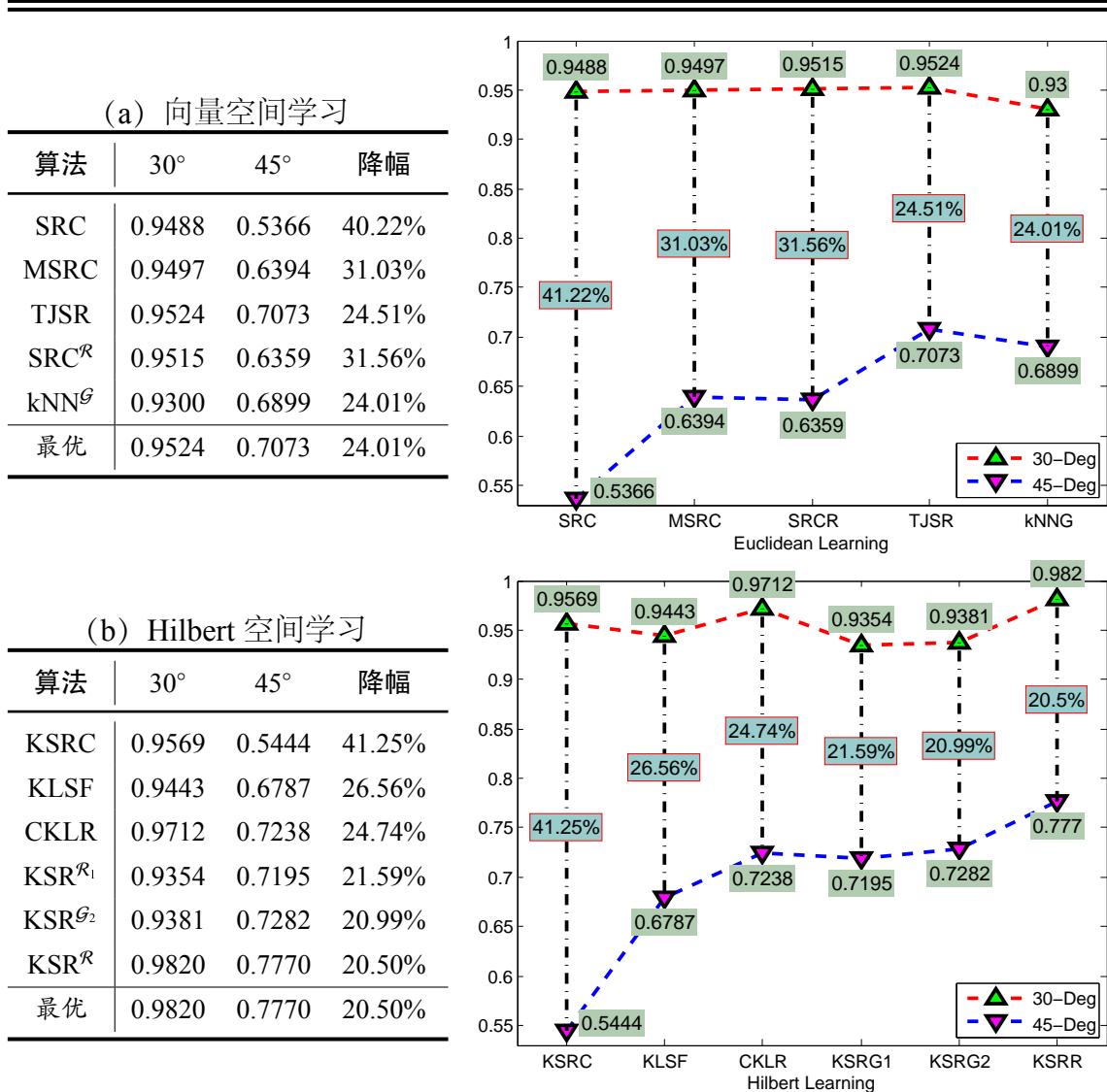


图 6.16 倾仰角变化对流形学习的影响，图中给出了 30° 和 45° 倾仰角测试样本的分类正确率，红色曲线与蓝色曲线间的差距是分类正确率的降幅。

结果来看，所有方法表现都比较好，识别率都能达到 0.9 以上。多特征多任务联合稀疏表示 TJSR 和方向可控 Riesz 小波核稀疏表示 KSR^R 分别取得最优识别率，0.9524 和 0.9820，其余流形学习方法的正确率稍稍低于它们两种，分别为 0.9300、0.9354、0.9381。

- 第二个实验场景测试雷达倾仰角度的剧烈变化：选取倾仰角为 45° 的成像样本进行测试，即训练样本与测试样本之间存在着 28° 的倾仰角差异。从实验结果来看，所有方法的识别率下降都比较明显，核稀疏表示 KSRC 的识别率甚至下降到 0.5444，下降了 41.25%，同样图像域稀疏表示 SRC 和特征组融合 MSRC 也下降了 41.22% 和 31.03%。多特征多任务联合稀疏表示

TJSR 和方向可控 Riesz 小波核稀疏表示 KSR^R 分别取得最优识别率, 0.7073 和 0.7770, 单演信号流形表示和方向可控 Riesz 小波在雷达俯仰角变化时的分类正确率降幅最小, 24.01% 和 20.50%。此外, 投影核函数的核稀疏表示 KSR^{G₁}、Binet-Cauchy 核函数的核稀疏表示 KSR^{G₂}、合成核学习 CKLR、联合稀疏表示 TJSR 表现也比较稳健, 它们的正确率降幅分别为 21.59%、20.99%、24.74%、24.51%, 远远低于其余的比较算法。综合来看, 方向可控 Riesz 小波框架核稀疏表示 KSR^R 表现最佳。

6.3.2.3 随机噪声污染

噪声在 SAR 图像形成过程比较普遍, 即有相干成像时固有的相干斑, 还有各种不同形式的测量误差, 以及迁移校正、运动补偿等成像算法的误差。为了观察噪声对于流形学习的影响, 论文开展一组随机噪声污染的分类识别实验, 选取 BMP2、T72、BTR60、T62 四类典型的军事目标进行实验。与前面实验设置不同, 对 15° 俯仰角的成像数据进行调整。为了模拟 SAR 图像成像过程中遇到的诸多噪声污染和测量误差, 论文随机挑选每幅测试图像的一部分强度值, 将其用满足独立同分步的均匀分布 $\mathcal{U}[0, u_{max}]$ 的样本代替, 其中 u_{max} 是可以设置的最大强度值。被污染的像素是从图像中随机选取。实验结果如表 6.10 和图 6.17 所示, 需要声明的是 SRC 算法采取针对噪声和遮挡的扩展 ℓ_1 范数最小化优化策略¹¹。

表 6.10 噪声污染对流形学习的影响, 表中列出了 7 种不同程度噪声污染的识别率。

算法	噪声等级							下降幅度
	0	5%	10%	15%	20%	25%	30%	
SVM	0.8443	0.7833	0.6404	0.4935	0.4606	0.4510	0.4454	39.89%
SRC	0.8764	0.8587	0.7945	0.6653	0.5618	0.5016	0.4655	41.09%
MSRC	0.8868	0.8515	0.8090	0.7504	0.6661	0.5939	0.5393	34.75%
SRC ^R	0.9117	0.7873	0.7408	0.7022	0.6846	0.6100	0.5923	31.94%
kNN ^{G₁}	0.8949	0.8796	0.8989	0.8788	0.8836	0.8740	0.8732	2.17%
kNN ^{G₂}	0.8981	0.8852	0.8989	0.8820	0.8812	0.8796	0.8724	2.25%
kNN ^{G₃}	0.8989	0.8836	0.8997	0.8812	0.8860	0.8780	0.8756	2.33%
kNN ^R	0.8957	0.8900	0.8796	0.8604	0.8347	0.8194	0.7945	12.12%
最优	0.9117	0.8900	0.8997	0.8820	0.8860	0.8796	0.8756	2.17%

¹¹扩展 ℓ_1 范数最小化是由 John Wright 在文献 [156] 提出, 用于应对噪声污染和遮挡, 该方法是指给原有的字典增加一个原有行维的方阵构成新字典, 该方阵可以是单位矩阵或是傅立叶变换、Haar 小波等生成, 然后通过求解回归系数进行判决。

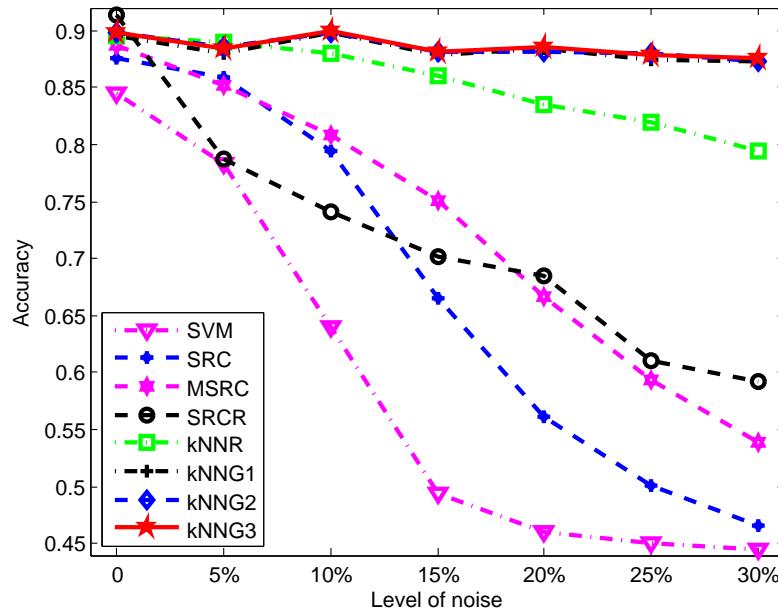


图 6.17 流形学习受噪声污染的识别率变化曲线。

为了观察随机噪声污染对流形学习的影响，分别采用了对称正定矩阵构成的黎曼流形、单演信号分量构成的 Grassmann 流形、方向可控 Riesz 小波构成的 Grassmann 流形，其中单演信号分量构成的 Grassmann 流形验证了弧长、投影度量、Binet-Cauchy 度量三种距离函数，分别记为 $\text{kNN}^{\mathcal{G}_1}$ 、 $\text{kNN}^{\mathcal{G}_2}$ 、 $\text{kNN}^{\mathcal{G}_3}$ 。从实验结果来看，随着噪声污染的不断加剧，所有算法的分类正确率都有不同程度地下降，然而不同算法的下降幅度的差别很明显。当噪声从 0% 增加到 30%，SVM、SRC、MSRC 的正确率分别下降了 39.89%、41.09%、34.75%，对称正定矩阵空间的流形学习下降了 31.94%，与之形成鲜明对比的是 Grassmann 流形学习表现比较稳健，三种 Grassmann 距离的分类算法 $\text{kNN}^{\mathcal{G}_1}$ 、 $\text{kNN}^{\mathcal{G}_2}$ 、 $\text{kNN}^{\mathcal{G}_3}$ 仅仅下降了 2.17%、2.25%、2.33%，方向可控 Riesz 小波框架的流形学习下降了 12.12%。若将噪声污染提升到 30%，Grassmann 流形学习方案仍然取得了 0.8740、0.8796、0.8780 和 0.7954 的正确率，远远高于经典的算法。实验结果充分说明了以原始矩阵构成的 Grassmann 流形的距离函数对于随机噪声污染有着很强的鲁棒性。

6.4 本章小结

本章主要介绍流形学习及其在单演信号尺度空间中的应用。首先介绍黎曼流形及相关的背景知识，给出黎曼流形距离度量，在此基础上提出一种具有非欧几何特性的区域特征描述方法即区域协方差矩阵，并针对对称正定矩阵空间设计分类学习方案；之后介绍了一种特殊的黎曼流形，Stiefel 和 Grassmann 流形，给出了 Grassmann 流形的约束条件及距离度量，设计基于黎曼（Grassmann）流形学习的分类框架，并在该框架下形成多种具体算法。主要内容包括：

1. 协方差矩阵空间形成的黎曼流形学习：协方差矩阵即对称正定矩阵空间对于数乘和加法不封闭，它不是欧氏空间的元素，若赋予它合适的度量可形成黎曼流形，从而可构造基于黎曼流形学习的分类，具体实现方案包括：
 - 对数协方差矩阵稀疏表示：借助面向流形的对数算子将协方差矩阵转换为切向空间的向量，根据该向量构成线性回归模型，通过求解回归系数进行分类；
 - 再生核 Hilbert 空间核回归：借助非线性映射将黎曼流形嵌入到再生核 Hilbert 空间，利用再生核 Hilbert 空间的像构造核回归模型，通过定义黎曼核函数求解核回归模型，实现分类；
2. 单演信号多尺度分量形成的 Grassmann 流形学习：将单演信号多尺度表示生成的分量排列为向量，对不同尺度空间的向量组合后的矩阵进行正文化处理，将其整体看作是 Grassmann 流形的点，通过非线性映射将 Grassmann 流形嵌入到再生核 Hilbert 空间，利用再生核 Hilbert 空间的像构造核回归模型，通过定义 Grassmann 核函数求解核回归模型实现分类；
3. 方向可控 Riesz 小波形成的 Grassmann 流形学习：将多方向多尺度 Riesz 小波框架的系数矩阵重新排列，进行正文化处理，将其整体看作 Grassmann 流形的点，借助非线性映射将 Grassmann 流形投影到再生核 Hilbert 空间，利用再生核 Hilbert 空间的像构造核回归模型，定义全局 Grassmann 核函数为所有尺度内核函数的累加，实现回归系数求解，完成样本分类。

最后，借助 MSTAR SAR 实测数据组织验证实验，分析生成流形的诸多因素对分类学习的影响，寻找最佳组合，实现与经典算法的性能比较。

第七章 结束语

论文紧紧围绕“基于单演信号的 SAR 图像目标识别研究”课题，着眼于扩展工作条件的 SAR 图像目标识别问题，通过引入一种新的信号分析与处理方法即单演信号来刻画 SAR 图像易变的目标散射现象，应对因姿态角变化、俯仰角变化、目标配置变化、噪声污染等因素对分类学习的不利影响；以简单、高效、易执行的目标识别技术为目标，采取先基础理论分析、后实测数据验证的步骤，对单演信号多尺度表示在不同学习框架内的应用进行深入而广泛地研究，通过大量的实验论证所提出的方法的实际性能表现。论文开展的工作以及取得的创新性成果可总结为：

1. **梳理现有成果：**论文对现有的 SAR 图像目标识别方法进行了全面而系统的归纳梳理，分析现有自动目标识别方法的优点与存在的不足，提炼出自动目标识别涉及到的关键技术以及面临的挑战；
2. **改进稀疏表示：**从信号的正交编码与解码引出基于过完备冗余字典的信号稀疏表示，解释稀疏表示相比正交表示所独有的优点，总结了稀疏表示在信号处理与机器学习中的典型应用，以例证的方式详细分析了稀疏表示在 SAR 图像目标识别中的应用可行性，结合实际分别对表示方法和决策规则进行改进，提出：
 - 基于傅立叶频谱的稀疏表示分类：利用原始信号傅立叶变换系数设计频域特征描述子，实现目标散射特性的不变刻画，根据频域特征描述子构造线性回归模型，通过求解回归系数实现分类；
 - 基于 DS 证据推理的“软”决策：根据线性回归模型中与回归响应相对应的回归参量的表示误差定义证据质量，借助 Dempster 证据合并规则对回归参量提供的证据进行合并，根据全局证据进行决策。
3. **引入单演信号：**首次将单演信号表示这种新提出的调和分析工具引入 SAR 图像解译，借助带通滤波器组构成单演信号尺度空间，实现单演信号多尺度表示，采用信号分解技巧将原始信号表示为多个代表不同类型信息的分量，借助机器学习算法将不同类型信息进行融合，提高 SAR 图像目标识别算法对扩展工作条件的耐受性；
4. **提出学习框架：**围绕着如何有效地融合单演信号多尺度表示所挖掘的不同类型信息，论文提出了三种学习框架，包括

框架一：空域信息融合，借助数据融合方法将单演信号多尺度表示生成的诸多分量进行空域融合，利用融合生成的特征描述构造线性回归模型，通过寻找最稀疏回归系数完成分类。在该框架的指导下提出三种具体实现算法：

算法一：特征级融合，将单演信号多尺度分量进行均匀下采样、能量归一化、逐行扫描组成增强的单演信号特征向量，根据增强的特征向量构造线性回归模型，通过求解回归系数实现分类；

算法二：决策级融合，将单演信号多尺度表示的多种分量分别进行均匀下采样和能量归一化，利用生成的特征向量分别构造线性回归模型，然后利用统计工具如贝叶斯理论、DS 证据理论把回归模型的输出进行融合实现分类；

算法三：联合稀疏表示，将单演信号多尺度表示的多种分量分别进行均匀下采样和能量归一化，利用生成的特征向量分别构造线性回归模型，通过对回归系数矩阵施加联合稀疏约束迫使所有回归模型的回归系数都有相似的稀疏项，实现分类。

框架二：多核合成学习，借助非线性映射将单演信号多尺度表示投影到再生核 Hilbert 空间，借助多核合成学习将不同类型的信息在 Hilbert 空间进行合并，构造核回归模型实现分类。在该框架的指导下提出两种算法：

算法一：“先合并后映射”，将单演信号多尺度表示生成的分量在数据空间进行合并，借助一个非线性映射将合并的结果投影到再生核 Hilbert 空间，利用投影得到的像构造核回归模型；

算法二：“先映射后合并”，借助多个非线性映射将单演信号多尺度表示分量分别投影到再生核 Hilbert 空间，利用投影得到的像构造核回归模型，引入多核合成学习将不同分量的核函数进行合并；

通过求解核回归模型的回归系数重构原始信号，根据重构误差进行判决。

框架三：黎曼流形学习，将单演信号表示生成的分量整体看作是黎曼流形的点，借助流形学习实现分类，具体分类算法包括：

算法一：协方差矩阵空间的流形学习，利用单演信号多尺度表示生成的诸多分量构造协方差矩阵，通过给协方差矩阵空间赋予黎曼度量构成黎曼流形，根据黎曼流形的特点设计两种算法：方案一是通过

面向流形的对数算子将协方差矩阵转换成为它切向空间的向量，然后构建线性回归模型；方案二是借助非线性映射将协方差矩阵投影到再生核 Hilbert 空间，利用投影得到的像构造核回归模型，借助黎曼核函数求解回归系数，根据回归系数实现分类；

算法二：单演信号系数矩阵生成的 Grassmann 流形学习，将单演信号多尺度表示生成的诸分量重新改造为列向量两两正交的矩阵，生成的矩阵整体看作是 Grassmann 流形的点，借助非线性映射将 Grassmann 流形嵌入到再生核 Hilbert 空间，利用再生核 Hilbert 空间的像构建核回归模型，通过定义在 Grassmann 流形的核函数求解回归系数，实现分类；

算法三：Riesz 小波系数矩阵生成的 Grassmann 流形学习，构造多尺度多方向 Riesz 小波框架，将生成的系数矩阵改造为 Grassmann 流形的点，借助非线性映射将 Grassmann 流形嵌入到再生核 Hilbert 空间，利用再生核 Hilbert 空间的像构建核回归模型，通过尺度内的 Grassmann 核函数的累加定义全局核函数，借此求解回归系数，实现分类。

利用 MSTAR SAR 实测数据针对每种框架的每类算法设计多组对比验证实验，内容包括目标本身的变化（如目标配置、版本型号、姿态角等）、周围环境的变化、成像参数的变化（如俯仰角等）、连接体变体、随机噪声污染等，通过实验结果展示提出的分类算法的优异性能。

论文详细回顾了近些年 SAR 图像自动目标识别技术的发展历程，归纳梳理已有的研究成果，分析它们的优点与不足，围绕着经典方法存在的不足引入单演信号和稀疏表示，提出多个分类框架，借助实测数据对提出的框架算法进行验证。展望未来，除了论文所做的工作以外，还有多个方面的研究有待进一步探索：

- 目标散射现象的不变描述：论文中将 SAR 图像的原始强度分布矩阵看作是二维有限长信号，通过引入一种新的信号分析与处理方法即单演信号对目标散射现象进行刻画，提取原始信号蕴含的多种不同类型信息，利用多种学习方法将这些信息进行融合，实现目标分类识别。未来有待持续研究的永恒课题就是探索简单有效、易于计算的目标散射特性描述方法；
- 分类学习子环节的改进：贯穿论文的两条主线是稀疏表示和单演信号，其中稀疏表示用于实现分类学习，通过将数据空间映射到再生核 Hilbert 空间应对线性不可分或线性可分性较差的情形。最近短短几年，稀疏表示的发展日

新月异，能否借鉴稀疏表示的最新研究进展，结合 SAR 图像目标识别的实际设计更稳健、更有效的学习方法是未来的又一个研究重点；

- 研究工作的相关领域扩展：论文通篇围绕着合成孔径雷达自动目标识别的分类开展了大量的研究工作，将稀疏表示与单演信号相结合提出了多个各具特色的分类框架，未来一个很有潜力的研究方向就是将相关研究扩展到目标检测与鉴别。目前，部分学者已经开展一些启发性的尝试，如基于稀疏表示的目标检测，但系统性的研究并不多；此外，目标与杂波之间的散射特性存在着显著差异，如何利用单演信号的诸多优良性质实现杂波的鉴别与剔除也有待进一步探索。

致 谢

光阴似箭，转眼间六年的硕博生涯就将结束，回首过去的日日夜夜，仿佛像电影荧幕般从我眼前滑过，镜头里看到了喜悦、也看到了苦涩，看到了温暖，也看到了沮丧。然而，一切的一切都随着学位论文的终稿而划上了句号，尘封在回忆里。在此谨向所有关心和帮助我的老师、同学和亲友表示感谢。没有你们的支持，我很难坚持下来。

衷心地感谢我的导师匡纲要教授。六年是个漫长而又短暂的过程，期间每篇小论文、每次学术报告匡老师都会与我反复讨论，提出中肯的意见，指明下一步的努力方向。我是参加基层工作六年之后才重回校园，开始读书，学习和工作中遇到许多困难和阻力，每每遇到困惑产生放弃的思想，匡老师都会耐心地开导我，以他渊博的学识和丰富的人生阅历消除我心中的疑虑。六年间，无论是学习、工作或日常生活，匡老师给我创造了很大的便利，让我少走很多弯路，能够遇到这样的良师益友，是我来到科大最幸运的事情。这些年从匡老师身上学到的不仅仅是一些书本知识，更多的是他那严谨求实的治学理念、精益求精的科研态度以及永不放弃的人生信条。在这即将离别的时刻，我衷心地祝福匡老师在以后的日子工作顺利、身体健康、家庭幸福。感谢粟毅教授，粟老师发散的思维、渊博的知识给我留下了深刻地印象，虽然与他的学术交流不多，但每次讨论之后我都会有很大的收获。感谢课题组的蒋咏梅老师、陆军老师、黄春琳老师等，我是后知后觉型学生，刚开始读研时，很多简单的知识点需要他（她）们多次讲解才能做到半知半懂，谢谢他（她）们将我引入学术之门，为我后来的科研道路做了丰富的知识储备和坚实铺垫。现在蒋老师已经离开课题组，祝她以后诸事顺利。感谢已经毕业的师兄贺治国、陈琪、朱宇涛、李禹、唐涛、熊博蕊与师姐赵凌君、刘丽等，谢谢他（她）们教我如何撰写科技论文、如何调试程序代码、如何进行中外文互译、如何检索文献资料等等；感谢课题组同学胡粲彬、黄嘉欣、王玮、涂松、张思乾、范海燕、项德良等，与他（她）们的多次讨论很好地拓宽了我的视野，谢谢他（她）们学习上给予的诸多支持与帮助；感谢实验楼 431 和 436 房间的刘涛、张小强、管冬冬、全斯农、赵逸超、于美婷等师弟和师妹们，祝他（她）们学业有成，日后走上心仪的工作岗位；感谢四院五队 10 级、七队 13 级的诸位同学，从研究生到博士的六年期间，大家相互学习、共同进步，度过了一段愉快而充实的美好时光，祝愿他（她）们在科研战线上再创佳绩，为母校争光。感谢西安电子科技大学雷达信号处理国家重点实验室的周峰教授为我提供的难得的访学机会。感谢

香港理工大学 *Yang Meng* 博士、西澳大学 *Peter Kovesi* 学者无私分享他们的研究成果与思路。还要感谢 IEEE SPL, GRSL, TIP, JSTARS 等期刊的编辑 Prof. *Rodrigo C. de Lamare*, Prof. *Alejandro Frery*, Prof. *Edwin A. Marengo*, Prof. *Jocelyn Chanussot* 等, 读博期间取得的学术成果也得益于他们的热心指导与宝贵建议。感谢校医院的赵晓冬、何晓霞等医生耐心地为我所做的多次检查, 感谢图书馆工作人员提供丰富的馆藏书籍和数字文献, 感谢四院博士队的历届领导给予的支持, 感谢机关和学位办工作人员在学位申请时给予的支持, 同时也感谢大学为我们提供的先进的科研设施和优美的校园环境, 我会怀念在这里奋斗的日子。

感谢同窗好友王毅在我读博期间给予的支持, 使我能专心完成学业, 还要感谢他们全家在我生病期间的细心照顾, 为我提供的诸多便利, 人生苦短、岁月无常, 愿我们的友谊长存, 祝他们全家身体健康、生活幸福。感谢老同学汪昶江、熊选华对于自己以后去向的关心与帮助, 祝他们工作顺利, 家庭美满。特别感谢我的好友尹华锋在诸多事情给予的巨大帮助, 感谢 68210 部队、68206 部队、93856 部队历任领导以及所有同事、战友在我求学路上的支持。

最后的笔墨留给我的家人。感谢父母的养育之恩, 自从大学毕业后我就很少有时间陪伴他们, 每次匆匆回家总感觉父亲的脚步又迟缓了一些, 母亲的白发又多了一些, 每想到此我的内心就有一种难以言表的伤感, 祈祷他(她)们能够陪伴我在人生路上多走一程。谢谢我的妻子王娜女士一直以来的鼓励和支持, 虽然这些年的许多事情都不顺利, 但我们能齐心协力、共同面对, 或许这也是生活对于家庭和成员的考验。就此搁笔, 再拜!

董刚刚, 于科大四院 431 房间

参考文献

- [1] 张澄波. 综合孔径雷达—原理、系统分析与应用 [M]. 北京: 科学出版社, 1989.
- [2] 黄培康, 殷红成, 许小剑. 雷达目标特性分析 [M]. 北京: 电子工业出版社, 2005.
- [3] 匡纲要, 高贵, 蒋咏梅等. 合成孔径雷达目标检测理论、算法及应用 [M]. 长沙: 国防科技大学出版社, 2007.
- [4] 保铮, 邢孟道, 王彤. 雷达成像技术 [M]. 北京: 电子工业出版社, 2005.
- [5] 张直中. 机载和星载合成孔径雷达导论 [M]. 北京: 电子工业出版社, 2003.
- [6] Bhanu B. Automatic Target Recognition: state-of-the-art survey [J]. IEEE Trans. Aerosp. Electron. Syst. 1986, 22: 364–379.
- [7] Bhanu B, Dudgeon D, Zelnio E, et al. Introduction to the Special Issue on Automatic Target Detection and Recognition [J]. IEEE Trans. Image Process. 1997, 6 (1): 1–6.
- [8] Dudgeon D, Lacoss R. An Overview of Automatic Target Recognition [J]. Lincoln Lab J. 1993, 6 (1): 3–10.
- [9] Novak L, Owirka G, Netishen C. Performance of a High-Resolution Polarimetric SAR Automatic Target Recognition System [J]. Lincoln Lab J. 1993, 6 (1): 11–23.
- [10] Novak L, Burl M, Chaney R, et al. Optimal Polarimetric Processing for Enhanced Target Detection [J]. IEEE Trans. Aerosp. Electron. Syst. 1993, 29 (1): 234–244.
- [11] Kuttikkad S, Chellappa R. Non-Gaussian CFAR Techniques for Target Detection in High Resolution SAR Images [C]. In Proc. IEEE Int'l Conf. Image Process. Nov. 1994: 910–914.
- [12] Subotic N, Thelen B, Gorman J, et al. Multiresolution Detection of Coherent Radar Targets [J]. IEEE Trans. Image Process. 1997, 6 (1): 21–35.
- [13] Anastassopoulos V, Lampropoulos G. Optimal CFAR Detection in Weibull Clutter [J]. IEEE Trans. Aerosp. Electron. Syst. 1995, 31 (1): 52–64.
- [14] Fukunaga K, Hayes R, Novak L. The Acquisition Probability for a Minimum Distance One-Class Classifier [J]. IEEE Trans. Aerosp. Electron. Syst. 1987, 23 (4): 493–499.
- [15] Kreithen D, Halversen S, Owirka G. Discriminating Targets from Clutter [J]. Lincoln Lab J. 1993, 6 (1): 25–52.

-
- [16] 张翠, 邹涛, 王正志. 一种基于目标主导边界的方位角估计算法 [J]. 模式识别与人工智能. 2004, 17 (4): 462–466.
 - [17] 计科峰, 匡纲要, 粟毅, 等. SAR 图像目标峰值特征提取与方位角估计方法研究 [J]. 宇航学报. 2004, 25 (1): 102–108.
 - [18] 王世希, 李禹, 刘军, 等. 一种基于间隙度特征的 SAR 图像车辆目标鉴别算法 [J]. 电子与信息学报. 2008, 30 (8): 1944–1948.
 - [19] Novak L M, Halversen S D, Owirka G J, et al. Effects of Polarization and Resolution on the Performance of a SAR Automatic Target Recognition System [J]. Lincoln Lab J. 1995, 8 (1): 49–68.
 - [20] 陈琪, 王娜, 陆军, 等. 一种基于 SAR 图像鉴别的港口区域舰船目标新方法 [J]. 宇航学报. 2011, 32 (12): 2582–2588.
 - [21] Gao G, Kuang G, Zhang Q, et al. Fast Detecting and Locating Groups of Targets in High-Resolution SAR Images [J]. Pattern Recognition. 2007, 40: 1378–1384.
 - [22] Bhanu B, Lin Y. Genetic Algorithm based Feature Selection for Target Detection in SAR Images [J]. Image and Vision Computing. 2003, 21: 591–608.
 - [23] Park J, Park S, Kim K. New Discrimination Features for SAR Automatic Target Recognition [J]. IEEE Geosci. Remote Sens. Lett. 2013, 10 (3): 476–480.
 - [24] Verbout S M, Weaver A L, Novak L M. New Image Features for Discriminating Targets from Clutter [C]. In SPIE Radar Sensor Technology III. 1998: 120–137.
 - [25] Owirka G, Novak L. A New SAR ATR Suite [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery I. 1994: 336–343.
 - [26] Williams A, Clark B. Evaluation of SAR ATR [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery III. 1996: 36–45.
 - [27] Keydel E, Lee S W, Moore J. MSTAR Extended Operating Conditions-A tutorial [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery III. Jun. 1996: 228–242.
 - [28] Potter L, Moses R. Attributed Scattering Centering for SAR ATR [J]. IEEE Trans. Image Process. 1997, 6 (1): 79–91.
 - [29] Mossing J, Ross T. An Evaluation of SAT ATR Algorithm Performance Sensitivity to MSTAR Extended Operating Conditions. [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery V. Orlando, Florida, 1998: 554–565.
 - [30] Pham Q, Brosnan T, Smith M, et al. A Efficient End-to-End Feature Based System for SAR ATR [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery VI. Orlando, Florida, Apr. 1998: 292–301.
-

-
- [31] Pham Q, Ezekiel A, Campbell M, et al. A New End-to-End SAR ATR System [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery VI. Orlando, Florida, Apr. 1999: 292–301.
 - [32] Novak L. State-of-the-art of SAR Automatic Target Recognition [C]. In Proc. IEEE Int'l Radar Conf. Alexandria, VA, 2000: 836–843.
 - [33] Blacknell D. Adaptive Design and Characterization of Features for SAR ATR [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery VIII. 2001: 252–263.
 - [34] Blacknell D. A Comparison of SAR ATR Performance with Information Theoretic Predictions [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery X. 2003: 385–395.
 - [35] Wise A, Fitzgerald D, Ross T. The Adaptive SAR ATR Problem Set [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery XI. 2004: 366–375.
 - [36] Owirka G, Weaver A, Novak L. Performance of a Multiresolution Classifier using Enhanced Resolution SAR Data [C]. In SPIE Radar Technology. Orlando, Florida, 1997: 90–100.
 - [37] Ross T, Worrell S, Welten V, et al. Standard SAR ATR Evaluation Experiments using the MSTAR Public Release Data Set [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery V. Apr. 1998: 566–573.
 - [38] Novak L, Owirka G, Weaver A. Automatic Target Recognition using Enhanced Resolution SAR Data [J]. IEEE Trans. Aerosp. Electron. Syst. 1999, 35 (1): 157–175.
 - [39] Owirka G, Verbout S, Novak L. Template-based SAR ATR Performance using Different Image Enhancement Techniques [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery VI. Orlando, Florida, Apr. 1999: 302–319.
 - [40] Novak L, Owirka G, Brower W. The Automatic Target Recognition System in SAIP [J]. Lincoln Lab J. 1997, 10 (2): 187–202.
 - [41] Benitz G. High-Definition Vector Imaging [J]. Lincoln Lab J. 1997, 10 (2): 147–170.
 - [42] Novak L, Owirka G, Brower W. Performance of 10- and 20-Target MSE Classifiers [J]. IEEE Trans. Aerosp. Electron. Syst. 2000, 36 (4): 1279–1289.
 - [43] Diemunsch J, Washbum R, Morgan D. Search Algorithms for Model-based SAR ATR [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery III. Orlando, Florida, 1996: 279–293.
-

-
- [44] Keydel E, Lee S W, Moore J. Signature Prediction for Model-based Automatic Target Recognition [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery III. 1996: 306–317.
 - [45] Keydel E, Williams W, Sieron R. Reasoning Support and Uncertainty Prediction in Model-based Vision SAR ATR [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery VI. 1999: 620–631.
 - [46] Williams W, Wojdacki W, Keydel E. Development of Class Models for Model-based Automatic Target Recognition [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery VI. 1999: 650–660.
 - [47] Diemunsch J, Wissinger J. Moving and Stationary Target Acquisition and Recognition (MSTAR) Model-Based Automatic Target Recognition: Search Technology for a Robust AIR [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery V. Orlando, Florida, 1998: 481–492.
 - [48] Andersh D, Hazlett M, Lee S. XPatch: a High Frequency Electromagnetic-Scattering Prediction Code and Environment for Complex Three-Dimensional Objects [J]. IEEE Antennas Propag. Mag. 1994, 36 (1): 65–69.
 - [49] Sharma R, Subotic N S. Construction of Hybrid Templates from Collected and Simulated Data for SAR ATR Algorithms [C]. In SPIE Automatic Target Recognition VII. Orlando, FL, 1998: 480–486.
 - [50] Xie H, Pierce L, Ulaby F. Statistical Properties of Logarithmically Transformed Speckle [J]. IEEE Trans. Geosci. Remote Sens. 2002, 40 (3): 721–727.
 - [51] Frery A. A Model for Extremely Heterogeneous Clutter [J]. IEEE Trans. Geosci. Remote Sens. 1997, 35 (3): 648–659.
 - [52] Blake A. High Resolution SAR Clutter Textural Analysis and Simulation [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery II. 1995: 101–108.
 - [53] Salazar J. Detection Schemes for Synthetic Aperture Radar Imagery Based On a Beta Prime Statistical Model [D]. USA: Florida University, 1999.
 - [54] Anastassopoulos V, Lampropoulos G, Drosopoulos A, et al. High Resolution Radar Clutter Statistics [J]. IEEE Trans. Aerosp. Electron. Syst. 1999, 35 (1): 43–59.
 - [55] Shnidman D. Generalized Radar Clutter Model [J]. IEEE Trans. Aerosp. Electron. Syst. 1999, 35 (3): 857–865.
 - [56] Kuruoglu E, Zerubia J. Modeling SAR Images With a Generalization of the Rayleigh Distribution [J]. IEEE Trans. Image Process. 2004, 13 (4): 527–533.
-

-
- [57] Delignon Y, Pieczynski W. Modeling non-Rayleigh Speckle Distribution in SAR Images [J]. IEEE Trans. Geosci. Remote Sens. 2002, 40 (6): 1430–1435.
- [58] Gao G. An Improved Scheme for Target Discrimination in High-Resolution SAR Images [J]. IEEE Trans. Geosci. Remote Sens. 2011, 49 (1): 277–294.
- [59] Suvorova S, Schroeder J. Automated Target Recognition using the Karhunen-Loéve Transform with Invariance [J]. Digital Signal Processing. 2002, 12: 295–306.
- [60] Liu M, Wu Y, Zhang P, et al. SAR Target Configuration Recognition Using Locality Preserving Property and Gaussian Mixture Distribution [J]. IEEE Geosci. Remote Sens. Lett. 2013, 10 (2): 268–272.
- [61] Dungan K, Potter L. Classifying Sets of Attributed Scattering Centers using a Hash Coded Database [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery XVII. Orlando, Florida, Apr. 2010: 5–9.
- [62] Dungan K, Potter L. Classifying Transformation-Variant Attributed Point Patterns [J]. Pattern Recognition. 2010, 43 (11): 3805–3816.
- [63] Rogers S, Colombi J, Martin C, et al. Neural Networks for Automatic Target Recognition [J]. Neural Networks. 1995, 8 (7): 1153–1184.
- [64] Chiang H-C, Moses R L, Potter L C. Model-Based Classification of Radar Images [J]. IEEE Trans. Inf. Theory. 2000, 46 (5): 1842–1854.
- [65] Zhao Q, Principe J. Support Vector Machines for SAR Automatic Target Recognition [J]. IEEE Trans. Aerosp. Electron. Syst. 2001, 37 (2): 643–654.
- [66] Sun Y, Liu Z, Todorovic S, et al. Adaptive boosting for SAR automatic target recognition [J]. IEEE Trans. Aerosp. Electron. Syst. 2007, 43 (1): 112–125.
- [67] Srinivas U, Monga V, Raj R. SAR Automatic Target Recognition Using Discriminative Graphical Models [J]. IEEE Trans. Aerosp. Electron. Syst. 2014, 50 (1): 591–606.
- [68] Kumar B V K V. Tutorial Survey of Composite Filter Designs for Optical Correlators [J]. Applied Optics. 1992, 31 (23): 4773–4801.
- [69] Mahalanobis A, Carlson D W, Kumar B V. Evaluation of MACH and DCCF Correlation Filters for SAR ATR using the MSTAR Public Database [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery V. Orlando, FL, 1998: 460–468.
- [70] Walls B, Mahalanobis A. Performance of the MACH filter and DCCF algorithms in the Presence of Data Compression [C]. In SPIE Automatic Target Recognition IX. Orlando, FL, 1999: 376–387.

-
- [71] Singh R, Kumar B. Performance of the extended maximum average correlation height filter and the polynomial distance classifier correlation filter for multiclass SAR detection and classification [C]. In SPIE Algorithms for SAR Imagery IX. Orlando, FL, 2002: 265–279.
- [72] Kumar B V K V, Mahalanobis A, Juday R. Correlation Pattern Recognition [M]. Cambridge, UK: Cambridge University Press, 2005.
- [73] 徐牧, 王雪松, 肖顺平. 基于 Hough 变换与目标主轴提取的 SAR 图像目标方位角估计方法 [J]. 电子与信息学报. 2007, 29 (2): 370–374.
- [74] Casasent D, Chang W-T. Correlation Synthetic Discriminant Functions [J]. Applied Optics. 1986, 25 (14): 2343–2350.
- [75] Kumar B V, Mahalanobis A, Carlson D W. Optimal Trade-off Synthetic Discriminant Function Filters for Arbitrary Devices [J]. Optics Letters. 1994, 19 (19): 1556–1558.
- [76] Carlson D W, Kumar B V, Mitchell R R, et al. Composite Correlation Filters for SAR Image Recognition [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery III. Orlando, FL, 1996: 338–349.
- [77] Mahalanobis A, Kumar B V K V, Song S, et al. Unconstrained Correlation Filters [J]. Applied Optics. 1994, 33 (17): 3751–3759.
- [78] Mahalanobis A, Carlson D W, Kumar B V, et al. Distance Classifier Correlation Filters [C]. In SPIE Hybrid Image and Signal Processing IV. Orlando, FL, 1994: 2–13.
- [79] Alkanhal M, Kumar B V, Mahalanobis A. Improving the False Alarm Capabilities of the Maximum Average Correlation Height Correlation Filter [J]. SPIE Optical Engineering. 2000, 39 (5): 1133–1141.
- [80] Alkanhal M, Kumar B V. Polynomial Distance Classifier Correlation Filter for Pattern Recognition [J]. Applied Optics. 2003, 42 (23): 4688–4708.
- [81] Ravichandran G, Casasent D P. Minimum Noise and Correlation Energy Optical Correlation Filter [J]. Applied Optics. 1992, 31 (11): 1823–1833.
- [82] Patnaik R, Casasent D. MINACE Filter Classification Algorithms for ATR using MSTAR Data [C]. In SPIE Automatic Target Recognition XV. Orlando, FL, 2005: 100–111.
- [83] Casasent D, Nehemiah A. Confuser Rejection Performance of EMACH Filters for MSTAR ATR [C]. In SPIE Optical Pattern Recognition XVII. 2006.

-
- [84] DeVore M D, Lanterman A D, OSullivan J A. ATR Performance of a Rician Model for SAR Images [C]. In SPIE Automatic Target Recognition X. 2000: 34–45.
- [85] DeVore M D, OSullivan J A, Montagnino L J. Conditionally Gamma- and K-distribution models for SAR ATR [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery IX. Orlando, FL, 2000: 27–37.
- [86] Trees H L V. Detection, Estimation, and Modulation Theory, Part III: Radar-Sonar Signal Processing and Gaussian Signals in Noise [M]. Canada: John Wiley Sons, 2001.
- [87] O'Sullivan J A, DeVore M D, Kedia V, et al. SAR ATR Performance using a Conditionally Gaussian Model [J]. IEEE Trans. Aerosp. Electron. Syst. 2001, 37 (1): 91–108.
- [88] Copsey K, Webb A. Bayesian gamma mixture model approach to radar target recognition [J]. IEEE Trans. Aerosp. Electron. Syst. 2003, 39 (4): 1201–1217.
- [89] DeVore M, O'Sullivan J. Quantitative Statistical Assessment of Conditional Models for Synthetic Aperture Radar [J]. IEEE Trans. Image Process. 2004, 13 (2): 113–125.
- [90] Li B. Experimental Evaluation of FLIR ATR Approach-A Comparative Study [J]. Computer Vision and Image Understanding. 2001, 84: 5–24.
- [91] Casasent D, Wang Y-C. A Hierarchical Classifier using New Support Vector Machines for Automatic Target Recognition [J]. Neural Networks. 2005, 18: 541–548.
- [92] Yuan C, Casasent D. MSTAR 10-Class Classification and Confuser and Clutter Rejection using SVRDM [C]. In Optical Pattern Recognition XVII. 2006.
- [93] Chen Y, Blasch E, Chen H, et al. Experimental Feature-Based SAR ATR Performance Evaluation under Different Operational Conditions [C]. In SPIE Signal Processing, Sensor Fusion, and Target Recognition XVII. 2008: 1–12.
- [94] 胡利平, 刘宏伟, 吴顺君. 基于改进的子类判决分析的 SAR 目标特征提取与识别 [J]. 电子与信息学报. 2009, 31 (9): 2264–2268.
- [95] Thiagarajan J, Ramamurthy K, Knee P, et al. Sparse Representation for Automatic Target Classification in SAR Images [C]. In Int'l Sym. Communcitaion, Control and Signal Processing. Mar. 2010.
- [96] 王燕霞, 张弓. 基于特征参数稀疏表示的 SAR 图像目标识别 [J]. 重庆邮电大学学报 (自然科学版). 2012, 24 (3): 308–313.

-
- [97] Yang S, Ma Y, Wang M, et al. Compressive Feature and Kernel Sparse Coding-based Radar Target Recognition [J]. IET Radar Sonar Navig. 2013, 7 (7): 755–763.
- [98] Zhang H, Nasrabadi N, Zhang Y, et al. Multi-View Automatic Target Recognition using Joint Sparse Representation [J]. IEEE Trans. Aerosp. Electron. Syst. 2011, 48 (3): 2481–2495.
- [99] Cheng J, Li L, Li H, et al. SAR Target Recognition based on Improved Joint Sparse Representation [J]. EURASIP Signal Process. 2014, 2014 (1): 1–12.
- [100] Morgan D R, Ross T D. A Bayesian Framework for ATR Decision-Level Fusion Experiments [C]. In Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications. 2007.
- [101] Rizvi S A, Nasrabadi N M. Fusion of Automatic Target Recognition Algorithms [C]. In SPIE Automatic Target Recognition XII. 2002: 122–132.
- [102] Liu M, Wu Y, Zhao W, et al. Dempster-Shafer Fusion of Multiple Sparse Representation and Statistical Property for SAR Target Configuration Recognition [J]. IEEE Geosci. Remote Sens. Lett. 2014, 11 (6): 1106–1110.
- [103] Liu H, Li S. Decision Fusion of Sparse Representation and Support Vector Machine for SAR Image Target Recognition [J]. Neurocomputing. 2013, 113 (7): 97–104.
- [104] Ross T D, Morgan D R, Blasch E P, et al. Survey of Approaches and Experiments in Decision-Level Fusion of Automatic Target Recognition (ATR) Products [C]. In SPIE Signal Processing, Sensor Fusion, and Target Recognition XVI. 2007.
- [105] Novak L, Owirka G, Weaver A. Automatic Target Recognition Using Enhanced Resolution SAR Data [J]. IEEE Trans. Aerosp. Electron. Syst. 1999, 35 (1): 267–278.
- [106] Bilik I, Tabrikian J, Cohen A. GMM-Based Target Classification for Ground Surveillance Doppler Radar [J]. IEEE Trans. Aerosp. Electron. Syst. 2006, 42 (1): 267–278.
- [107] Duan R, Chen L, Man H. Adaptive SAR ATR in likelihood space [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery XII. Orlando, FL, 2005: 294–305.
- [108] 韩萍, 吴仁彪, 王兆华, 等. 基于 KPCA 准则的 SAR 图像目标特征提取与识别 [J]. 电子与信息学报. 2003, 25 (10): 1297–1301.
- [109] 胡利平, 刘宏伟, 吴顺君. 基于两级 2DPCA 的 SAR 目标特征提取与识别 [J]. 电子与信息学报. 2008, 30 (7): 1722–1726.
- [110] 韩征, 苏志刚, 韩萍, 等. 基于样本正交子空间的 SAR 目标识别方法 [J]. 电子与信息学报. 2009, 31 (11): 2581–2584.

- [111] 王世希, 贺志国. 基于 PCA 特征的快速 SAR 图像目标识别方法 [J]. 国防科技大学学报. 2008, 30 (3): 136–140.
- [112] 龙泓琳, 皮亦明, 曹宗杰. 基于非负矩阵分解的 SAR 图像目标识别 [J]. 电子学报. 2010, 38 (6): 1–5.
- [113] 苏志刚, 常明, 吴仁彪, 等. 相似目标识别的投影子空间构造方法 [J]. 系统工程与电子技术. 2011, 33 (3): 473–475.
- [114] 李映, 龚红丽, 梁佳熙, 等. 基于 KSVD 和 PCA 的 SAR 图像目标特征提取 [J]. 吉林大学学报 (工学版). 2010, 40 (5): 1336–1339.
- [115] 张静, 杨智勇, 王国宏, 等. 基于局部约束编码的稀疏保持投影降维识别方法研究 [J]. 电子学报. 2016, 44 (3): 658–664.
- [116] Patnaik R, Casasent D. SAR Classification and Confuser and Clutter Rejection Tests on MSTAR Ten-class Data using MINACE Filters [C]. In SPIE Optical Pattern Recognition XVIII, Proc. SPIE. Orlando (Kissimmee), FL, Aug. 2007.
- [117] 宣若虹, 杨汝良. 基于小波域 NMF 特征提取的 SAR 图像目标识别方法 [J]. 电子与信息学报. 2009, 31 (3): 588–591.
- [118] Sandirasegaram N, English R. Comparative Analysis of Feature Extraction (2D FFT and wavelet) and Classification (L_p metric distances, MLP NN, and HNeT) Algorithms for SAR Imagery [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery XII. Orlando, FL, 2005.
- [119] Hu F, Zhang P, Yang R, et al. SAR Target Recognition based on Gabor Filter and Sub-block Statistical Feature [C]. In Proc. IET Int'l Radar Conf. 2009: 1–4.
- [120] Wang S, Liu Z, Jiao L, et al. Synthetic Aperture Radar Automatic Target Recognition Based On Curvelet Transform [C]. In SPIE MIPPR 2009, Automatic Target Recognition and Image Analysis. 2009.
- [121] Kaplan L M, Murenzi R. SAR Target Feature Extraction using the 2-D Continuous Wavelet Transform [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery IV. 1997: 101–112.
- [122] Olson C F, Huttenlocher D P. Automatic Target Recognition by Matching Oriented Edge Pixels [J]. IEEE Trans. Image Process. 1997, 6 (1): 103–113.
- [123] Nilubol C, Mersereau R M, Smith M J. A SAR Target Classifier Using Radon Transforms and Hidden Markov Models [J]. Digital Signal Processing. 2002, 12: 274–283.

-
- [124] Saghi J A, DeKelaita A. Exploitation of Target Shadows in Synthetic Aperture Radar Imagery for Automatic Target Recognition [C]. In SPIE Applications of Digital Image Processing XXIX. 2006.
 - [125] Papson S, Narayanan R M. Applying Target Shadow Models for SAR ATR [C]. In SPIE Signal Processing, Sensor Fusion, and Target Recognition XVI. 2007.
 - [126] Papson S, Narayanan R M. Classification via the Shadow Region in SAR Imagery [J]. IEEE Trans. Aerosp. Electron. Syst. 2012, 48 (2): 969–980.
 - [127] JonesIII G, Bhanu B. Quasi-Invariants for Recognition of Articulated and Non-Standard Objects in SAR Images [C]. In IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications. Fort Collins, CO, 1999: 88–97.
 - [128] JonesIII G, Bhanu B. Recognition of Articulated and Occluded Objects [J]. IEEE Trans. Pattern Anal. Mach. Intell. 1999, 21 (7): 603–613.
 - [129] Gudnason J, Cui J, Brookes M. HRR Automatic Target Recognition from Superresolution Scattering Center Features [J]. IEEE Trans. Aerosp. Electron. Syst. 2009, 45 (4): 1512–1526.
 - [130] Zhou J, Shi Z, Cheng X, et al. Automatic Target Recognition of SAR Images Based on Global Scattering Center Model [J]. IEEE Trans. Geosci. Remote Sens. 2011, 49 (10): 3713–3729.
 - [131] 娄军, 金添, 周智敏. 基于SEEMD的高分辨率SAR图像目标特征提取方法 [J]. 电子学报. 2012, 40 (4): 793–798.
 - [132] Cover T, Hart P. Nearest Neighbor Pattern Classification [J]. IEEE Trans. Inf. Theory. 1967, 13 (1): 21–27.
 - [133] Basri R, Hassner T, Zelnikmanor L. Approximate Nearest Subspace Search with Applications to Pattern Recognition [C]. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Minneapolis, MN, 2007: 1–8.
 - [134] Ikeuchi K, Shakunaga T, Wheeler M, et al. Invariant Histograms and Deformable Template Matching for SAR Target Recognition [C]. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 1996: 100–105.
 - [135] Park J-I, Kim K-T. Modified Polar Mapping Classifier for SAR Automatic Target Recognition [J]. IEEE Trans. Aerosp. Electron. Syst. 2014, 50 (2): 969–980.
 - [136] Vapnik V. The Nature of Statistical Learning Theory [M]. New York, USA: Springer, 1995.
-

- [137] Schölkopf B, Smola A. Learning With Kernels: Support Vector Machines, Regularization, Optimization and Beyond [M]. Cambridge, UK: MIT Press, 1995.
- [138] Bryant M, Garber F. SVM Classifier Applied to the MSTAR Public Data Set [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery VI. 1999: 355–360.
- [139] Eddy S R. Hidden Markov Models [J]. Current Opinion in Structural Biology. 1996, 6 (6): 361–365.
- [140] Brown M Z. Analysis of Multiple-View Bayesian Classification for SAR ATR [C]. In SPIE Algorithms for Synthetic Aperture Radar Imagery X. 2003: 265–274.
- [141] Du L, Liu H, Bao Z, et al. A Two-Distribution Compounded Statistical Model for Radar HRRP Target Recognition [J]. IEEE Trans. Signal Process. 2006, 54 (6): 2226–2239.
- [142] Lin Y, Bhanu B. Evolutionary Feature Synthesis for Object Recognition [J]. IEEE Trans. Syst., Man, Cybern. C. 2005, 35 (2): 156–171.
- [143] Erickson K J, Ross T D. A Testbed for Architecture and Fidelity Trade Studies in the Bayesian Decision-Level Fusion of ATR Products [C]. In SPIE Signal Processing, Sensor Fusion, and Target Recognition XVI. 2007.
- [144] Shi L, Wang P, Liu H, et al. Radar HRRP Statistical Recognition with Local Factor Analysis by Automatic Bayesian Ying-Yang Harmony Learning [J]. IEEE Trans. Signal Process. 2011, 59 (2): 610–617.
- [145] Hinton G E, Osindero S, Teh Y W. A Fast Learning Algorithm for Deep Belief Nets [J]. Neural Computation. 2006, 18 (7): 1527–54.
- [146] Principe J C, Kim M, Fisher J W. Target Discrimination in Synthetic Aperture Radar Using Artificial Neural Networks [J]. IEEE Trans. Image Process. 1998, 7 (8): 1136–1149.
- [147] Zhao Q, Principe J C, Brennan V L, et al. Synthetic aperture radar automatic target recognition with three strategies of learning and representation [J]. SPIE Optical Engineering. 2000, 39 (5): 1230–1244.
- [148] 杨露菁, 郝威, 刘忠, 等. 基于多特征空间与神经网络的 SAR 图像识别方法 [J]. 系统工程与电子技术. 2009, 31 (12): 2859–2862.
- [149] Neagoe V-E, Carausu D-C, Strugaru G-E. A concurrent neural module classifier for automated target recognition in SAR imagery [C]. In Proceedings of the 14th WSEAS international conference on Computers. 2010: 208–213.

-
- [150] Briones J C, Flores B, Cruz-Cano R. Multi-mode radar target detection and recognition using neural networks [J]. International Journal of Advanced Robotic Systems. 2012, 9.
 - [151] 李帅, 许悦雷, 马时平, 等. 一种深度神经网络 SAR 遮挡目标识别方法 [J]. 西安电子科技大学学报 (自然科学版). 2015, 42 (3): 154–160.
 - [152] Wang H, Chen S, Xu F, et al. Application of deep-learning algorithms to MSTAR data [C]. In Proc. IEEE Int'l Geosci. Remote Sens. Symp. 2015: 3743–3745.
 - [153] Ermushev S, Balashov A. A Complex Machine Learning Technique For Ground Target Detection and Classification [J]. International Journal of Applied Engineering Research. 2016, 11 (1): 158–161.
 - [154] Olshausen B A, Field D J. Emergence of Simple-cell Receptive Field Properties by Learning a Sparse Code for Natural Images [J]. Nature. 1996, 381 (6583): 607–609.
 - [155] Geiger D, Liu T-L, Donahue M J. Sparse Representations for Image Decompositions [J]. International Journal of Computer Vision. 1999, 33 (2): 139–156.
 - [156] Wright J, Yang A, Ganesh A, et al. Robust Face Recognition via Sparse Representation [J]. IEEE Trans. Pattern Anal. Mach. Intell. 2009, 31 (2): 210–227.
 - [157] Tropp J. Greed is good: Algorithmic results for sparse approximation [J]. IEEE Trans. Inf. Theory. 2004, 50 (10): 2231–2242.
 - [158] Candès E J, Tao T. Near-Optimal Signal Recovery from Random Projections: Universal Encoding Strategies [J]. IEEE Trans. Inf. Theory. 2006, 52 (12): 5406–5425.
 - [159] Lee H, Battle A, Raina R, et al. Efficient Sparse Coding Algorithms [C]. In Proc. Adv. Neural Inf. Process. Syst. Vancouver, Canada, 2007: 801–808.
 - [160] Varshney K, Çetin M, Fisher J, et al. Sparse Representation in Structured Dictionaries With Application to Synthetic Aperture Radar [J]. IEEE Trans. Signal Process. 2008, 56 (8): 3548–3561.
 - [161] Samadi S, Cetin M, Masnadi-Shirazi M. Sparse Representation-Based Synthetic Aperture Radar Imaging [J]. IET Radar Sonar Navig. 2009, 5 (2): 182–193.
 - [162] Zhou N, Chen W. High Range Resolution Profile Automatic Target Recognition Using Sparse Representation [J]. Chinese Journal of Aeronautics. 2010, 23: 556–562.
 - [163] Austin C D, Ertin E, Moses R L. Sparse Signal Methods for 3-D Radar Imaging [J]. IEEE J. Sel. Topics Signal Process. 2014, 5 (3): 408–423.
-

-
- [164] Fang J, Xu Z, Zhang B, et al. Fast Compressed Sensing SAR Imaging Based on Approximated Observation [J]. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. 2014, 7 (1): 352–363.
 - [165] Liu Z-J, Li Q, Xia Z-W, et al. Target Recognition for Small Samples of Ladar Range Image using Classifier Ensembles [J]. SPIE Optical Engineering. 2012, 51 (8): 1–12.
 - [166] Huan R, Pan Y. Decision Fusion Strategies for SAR Image Target Recognition [J]. IET Radar Sonar Navig. 2011, 5 (7): 747–755.
 - [167] Dong G, Kuang G. Target Recognition via Information Aggregation Through Dempster-Shafer's Evidence Theory [J]. IEEE Geosci. Remote Sens. Lett. 2015, 12 (6): 1247–1251.
 - [168] Olshausen B A, Field D J. Sparse Coding with an Overcomplete Basis Set: A Strategy Employed by V1? [J]. Vision Research. 1997, 37 (23): 3311–3325.
 - [169] Candès E, Tao T. Decoding by Linear Programming [J]. IEEE Trans. Inf. Theory. 2006, 51 (12): 4203–4215.
 - [170] Donoho D L. Compressed Sensing [J]. IEEE Trans. Inf. Theory. 2006, 52 (4): 1289–1306.
 - [171] 吴翊, 李超, 罗建书, 等. 应用数学基础 [M]. 北京: 高等教育出版社, 2006.
 - [172] Mallat S, Zhang Z. Matching Pursuits with Time Frequency Dictionaries [J]. IEEE Trans. Signal Process. 1993, 41 (12): 3397–3415.
 - [173] Eldar Y, Kutyniok G. Compressed Sensing: Theory and Applications [M]. Cambridge, UK: Cambridge University Press, 2012.
 - [174] Rubinstein R, Bruckstein A, Elad M. Dictionaries for Sparse Representation Modeling [J]. Proc. IEEE. 2010, 98 (6): 1045–1057.
 - [175] Do M N, Vetterli M. The Contourlet Transform: an Efficient Directional Multiresolution Image Representation [J]. IEEE Trans. Image Process. 2005, 14 (12): 2091–2106.
 - [176] Peyre G. Best Basis Compressed Sensing [J]. IEEE Trans. Signal Process. 2010, 58 (5): 2613–2622.
 - [177] Qian Y, Ye M. Hyperspectral Imagery Restoration Using Nonlocal Spectral-Spatial Structured Sparse Representation With Noise Estimation [J]. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. 2013, 6 (2): 499–515.
 - [178] Engan K, Aase S, Husoy J H. Method of Optimal Directions for Frame Design [C]. In Proc. IEEE Conf. Acoustic, Sonar and Signal Processing. 1999: 2443–2446.
-

- [179] Engan K, Aase S, Husoy J H. Multi-frame Compression: Theory and Design [J]. EURASIP Signal Process. 2000, 80 (10): 2121–2140.
- [180] Lesage S, Gribonval R, Bimbot F, et al. Learning Unions of Orthonormal Bases with Thresholded Singular Value Decomposition [C]. In Proc. IEEE Conf. Acoustic, Sonar and Signal Processing. 2005: 293–296.
- [181] Vidal R, Ma Y, Sastry S. Generalized Principal Component Analysis (GPCA) [J]. IEEE Trans. Pattern Anal. Mach. Intell. 2005, 27 (12): 1945–1959.
- [182] Aharon M, Elad M, Bruckstein A. K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation [J]. IEEE Trans. Signal Process. 2006, 54 (11): 4311–4321.
- [183] Zhang Q, Li B. Discriminative K-SVD for Dictionary Learning in Face Recognition [C]. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. San Francisco, CA, Jun. 2010: 2691–2698.
- [184] Jiang Z, Zhang G, Davis L. Submodular Dictionary Learning for Sparse Coding [C]. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Providence, RI, Jun. 2012: 3814–3825.
- [185] Blumensath T, Davies M. Sparse and Shift-Invariant Representations of Music [J]. 2006, 14 (1): 50–57.
- [186] Sallee P, Olshausen B A. Learning Sparse Multiscale Image Representations [C]. In Proc. Adv. Neural Inf. Process. Syst. 2003: 1327–1334.
- [187] Mairal J, Sapiro G, Elad M. Learning Multiscale Sparse Representations for Image and Video Restoration [J]. 2008, 7 (1): 214–241.
- [188] Rubinstein R, Zibulevsky M, Elad M. Double Sparsity: Learning Sparse Dictionaries for Sparse Signal Approximation [J]. 2010, 58 (3): 1553–1564.
- [189] Mairal J, Bach F. Online Dictionary Learning for Sparse Coding [C]. In Proc. Int'l Conf. Machine Learning. Montreal, Canada, 2009: 689–696.
- [190] Mairal J, Bach F. Online Learning for Matrix Factorization and Sparse Coding [J]. Journal of Machine Learning Research. 2010, 11: 19–60.
- [191] Mairal J, Bach F, Ponce J. Task-Driven Dictionary Learning [J]. IEEE Trans. Pattern Anal. Mach. Intell. 2012, 34 (4): 791–804.
- [192] Kopsinis Y, Slavakis K, Theodoridis S. Online Sparse System Identification and Signal Reconstruction Using Projections onto Weighted ℓ_1 Balls [J]. IEEE Trans. Signal Process. 2011, 59 (3): 936–952.

-
- [193] Baron D, Duarte M, Wakin M, et al. Distributed Compressed Sensing [J]. 2009.
 - [194] Tibshirani R. Regression Shrinkage and Selection via the LASSO [J]. *J. Roy. Stat. Soc., Ser. B*. 1996, 58 (1): 267–288.
 - [195] van den Berg E, Friedlander M P. Probing the Pareto Frontier for Basis Pursuit Solutions [J]. *SIAM J. Scientific Computing*. 2008, 31 (2): 890–912.
 - [196] Kim S-J, Koh K, Lustig M, et al. An Interior-Point Method for Large-Scale ℓ_1 -Regularized Least Squares [J]. *IEEE J. Sel. Topics Signal Process*. 2007, 1 (4): 606–617.
 - [197] Bruckstein A M, Donoho D L, Elad M. From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images [J]. *SIAM Review*. 2009, 51 (1): 34–81.
 - [198] Yang J, Zhang L, Xu Y, et al. Beyond sparsity: The role of ℓ_1 -Optimizer in Pattern Classification [J]. *Pattern Recognition*. 2012, 45: 1104–1118.
 - [199] Elad M, Figueiredo M, Ma Y. On the Role of Sparse and Redundant Representations in Image Processing [J]. *Proc. IEEE*. 2010, 98 (6): 972–982.
 - [200] Wright J, Ma Y, Mairal J, et al. Sparse Representation for Computer Vision and Pattern Recognition [J]. *Proc. IEEE*. 2010, 98 (6): 1031–1043.
 - [201] Alessandra Budillon A E, Schirinzi G. Three-Dimensional SAR Focusing From Multipass Signals Using Compressive Sampling [J]. *IEEE Trans. Geosci. Remote Sens.* 2011, 49 (1): 488–499.
 - [202] Zhou Y, Jin R, CHHoi S. Exclusive Lasso for Multi-task Feature Selection [C]. In *Int'l Conf. Artificial Intelligence and Statistics*. Sardinia, Italy, 2010: 988–995.
 - [203] He C, Liu L, Xu L, et al. Learning Based Compressed Sensing for SAR Image Super-Resolution [J]. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 2012, 5 (4): 1271–1280.
 - [204] Berisha V, Shah N, Waagen D, et al. Sparse Manifold Learning with Applications to SAR Image Classification [C]. In *Proc. IEEE Conf. Acoustic, Sonar and Signal Processing*. 2007: 1089–1092.
 - [205] Donoho D L. For Most Large Underdetermined Systems of Linear Equations the Minimal ℓ^1 -Norm Near Solution Approximates the Sparsest Solution [J]. *Comm. Pure and Applied Math.* 2006, 59 (10): 907–934.
 - [206] Wakin M B, Donoho D L, Choi H, et al. The Multiscale Structure of Non-differentiable Image Manifolds [C]. In *SPIE Wavelets XI*. San Diego, California, 2005.
-

- [207] Roweis S T, Saul L K. Nonlinear dimensionality analysis by locally linear embedding [J]. *Science*. 2000, 290: 2323–2326.
- [208] Hwang W, Wang H, Kim H, et al. Face Recognition System Using Multiple Face Model of Hybrid Fourier Feature Under Uncontrolled Illumination Variation [J]. *IEEE Trans. Image Process.* 2011, 20 (4): 1152–1165.
- [209] Hegarat-Mascle S, Bloch I, Vidal-Madjar D. Application of Dempster-Shafer Evidence Theory to Unsupervised Classification in Multisource Remote Sensing [J]. *IEEE Trans. Geosci. Remote Sens.* 1997, 35 (4): 1018–1031.
- [210] Koks D, Challa S. An Introduction to Bayesian and Dempster-Shafer Data Fusion [R/OL]. 2005. <http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR-1436.pdf>.
- [211] Dempster A. Upper and lower probabilities induced by a multivalued mapping [J]. *The Annals of Mathematical Statistics*. 1967, 38 (2): 325–339.
- [212] Denceux T. A k-Nearest Neighbor Classification Rule Based on Dempster-Shafer Theory [J]. *IEEE Trans. Syst., Man, Cybern.* 1995, 25 (5): 804–813.
- [213] Tupin F, Bloch I, Maitre H. A First Step Toward Automatic Interpretation of SAR Images using Evidential Fusion of Several Structure Detectors [J]. *IEEE Trans. Geosci. Remote Sens.* 1999, 37 (3): 1327–1343.
- [214] Liu Z, Dezert J, Mercier G, et al. Dynamic Evidential Reasoning for Change Detection in Remote Sensing Images [J]. *IEEE Trans. Geosci. Remote Sens.* 2012, 50 (5): 1955–1967.
- [215] Tian J, Cui S, Reinartz P. Building Change Detection based on Satellite Stereo Imagery and Digital Surface Models [J]. *IEEE Trans. Geosci. Remote Sens.* 2014, 52 (1): 406–417.
- [216] Laha A, Pal N, Das J. Land Cover Classification using Fuzzy Rules and Aggregation of Contextual Information Through Evidence Theory [J]. *IEEE Trans. Geosci. Remote Sens.* 2006, 44 (6): 1633–1641.
- [217] Sarkar A, Banerjee A, Banerjee N, et al. Landcover Classification in MRF Context using Dempster-Shafer Fusion for Multisensor Imagery [J]. *IEEE Trans. Image Process.* 2005, 14 (5): 634–644.
- [218] Dong G, Kuang G. A Soft Decision Rule for Sparse Signal Modeling via Dempster-Shafer Evidential Reasoning [J]. *IEEE Geosci. Remote Sens. Lett.* 2016, 13 (10): 1–5.

- [219] Felsberg M, Sommer G. The Monogenic Signal [J]. IEEE Trans. Signal Process. 2001, 49 (12): 3136–3144.
- [220] Felsberg M, Sommer G. The monogenic scale-space: A unifying approach to phase-based image processing in scale space [J]. Journal of Mathematical Imaging and Vision. 2004, 21 (1): 5–26.
- [221] Felsberg M, Duits R, Florack L. The Monogenic Scale Space on a Rectangular Domain and its Features [J]. International Journal of Computer Vision. 2005, 64 (2): 187–201.
- [222] Hahn S. Multidimensional Complex Signals with Single-Orthant Spectra [J]. Proc. IEEE. 1992, 80 (8): 1287–1300.
- [223] Bülow T, Sommer G. Hypercomplex Signals-A Novel Extension of the Analytic Signal to the Multidimensional Case [J]. IEEE Trans. Signal Process. 2001, 49 (11): 2844–2852.
- [224] Wietzke L, Sommer G. The Conformal Monogenic Signal [C]. In LNCS. 2008: 527–536.
- [225] Unser M, Sage D, Ville D. Multiresolution Monogenic Signal Analysis Using the Riesz-Laplace Wavelet Transform [J]. IEEE Trans. Image Process. 2009, 18 (11): 2042–2058.
- [226] Olhede S C, Metikas G. The Monogenic Wavelet Transform [J]. IEEE Trans. Signal Process. 2009, 57 (9): 3426–3441.
- [227] Unser M, Ville D V D. Wavelet Steerability and the Higher-Order Riesz Transform [J]. IEEE Trans. Image Process. 2010, 19 (3): 3426–3441.
- [228] Wietzke L, Sommer G. The Signal Multi-Vector [J]. Journal of Mathematical Imaging and Vision. 2010, 37: 132–150.
- [229] Depeursinge A, Foncubierta-Rodriguez A, de Ville D V, et al. Rotation-Covariant Texture Learning Using Steerable Riesz Wavelets [J]. IEEE Trans. Image Process. 2014, 23 (2): 898–908.
- [230] Mellor M, Brady M. Phase Mutual Information as a Similarity Measurement for Registration [J]. Med. Image Anal. 2005, 9 (4): 330–343.
- [231] Zang D, Sommer G. Signal Modeling for Two-Dimensional Image Structures [J]. J. Vis. Commun. Image Represent. 2007, 18 (1): 81–99.
- [232] Demarcq G, Mascarilla L, Berthier M, et al. The Color Monogenic Signal: Application to Color Edge Detection and Color Optical Flow [C]. 2011: 269–284.

- [233] Sierra-Vázquez V, Serrano-Pedraza I. Application of Riesz Transforms to the Isotropic AM-PM Decomposition of Geometrical-Optical Illusion Images [J]. *J. Opt. Soc. Am. A.* 2010, 27 (4): 781–796.
- [234] Belaid A, Boukerroui D, Maingourd Y, et al. Phase-Based Level Set Segmentation of Ultrasound Images [J]. *IEEE Trans. Inf. Technol. Biomed.* 2011, 15 (1): 3426–3441.
- [235] Zhang L, Li H. Encoding Local Image Patterns using Riesz Transforms: With Applications to Palmprint and Finger-knuckle-print Recognition [J]. *Image and Vision Computing.* 2012, 30: 1043–1051.
- [236] Huang X, Zhao G, Zheng W, et al. Spatiotemporal Local Monogenic Binary Patterns for Facial Expression Recognition [J]. *IEEE Signal Process. Lett.* 2012, 19 (5): 476–480.
- [237] Yang M, Zhang L, Shiu S, et al. Monogenic Binary Coding: An Efficient Local Feature Extraction Approach to Face Recognition [J]. *IEEE Trans. Inf. Forensics Security.* 2012, 7 (6): 1738–1751.
- [238] Alessandrini M, Liebgott H, Friboulet D, et al. Myocardial Motion Estimation from Medical Images Using the Monogenic Signal [J]. *IEEE Trans. Image Process.* 2013, 22 (3): 1084–1095.
- [239] Soulard R, Carré P, Fernandez-Maloigne C. Vector Extension of Monogenic Wavelets for Geometric Representation of Color Images [J]. *IEEE Trans. Image Process.* 2013, 22 (3): 1070–1083.
- [240] Venkatesh S, Owens R. On the Classification of Image Features [J]. *Pattern Recognition Letters.* 1990, 11: 339–349.
- [241] Stein E M. *Singular Integrals and Differentiability Properties of Functions* [M]. Princeton, New Jersey: Princeton University Press, 1970.
- [242] Bracewell R N. *The Fourier Transform and Its Applications* [M]. McGraw-Hill, 1978.
- [243] Peyrin F, Zhu Y, Goutte R. Extension of the Notion of Analytic Signal for Multidimensional Signals. Application to Images [J]. *Signal Processing III: Theories and Applications.* 1986: 677–680.
- [244] Fields D. Relations between the Statistics of Natural Images and the Response Properties of Cortical Cells [J]. *J. Opt. Soc. Am. A.* 1987, 4 (12): 2379–2394.
- [245] Kovesi P. *Invariant Measurement of Image Features from Phase Information* [D]. Australia: The University of Western Australia, 1996.

- [246] Cook J, Chandran V, Fookes C. 3D Face Recognition using Log-Gabor Templates [C]. In Proc. British Machine Vision Conf. Edinburgh, 2006: 769–778.
- [247] Boukerroui D, Noble J A, Brady M. On the Choice of Band-pass Quadrature Filters [J]. Journal of Mathematical Imaging and Vision. 2004, 21 (1-2): 53–80.
- [248] Wietzke L, Sommer G, Fleischmann O. The Geometry of 2D Image Signals [C]. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 2009: 1690–1697.
- [249] Dong G, Wang N, Kuang G. Sparse Representation of Monogenic Signal: with Application to Target Recognition in SAR Images [J]. IEEE Signal Process. Lett. 2014, 21 (8): 952–956.
- [250] Kittler J, Hatef M, Duin R, et al. On Combining Classifiers [J]. IEEE Trans. Pattern Anal. Mach. Intell. 1998, 20 (3): 226–239.
- [251] Rottensteiner F, Trinder J, Clode S, et al. Using the Dempster-Shafer Method for the Fusion of LIDAR Data and Multi-spectral Images for Building Detection [J]. Information Fusion. 2005, 6: 283–300.
- [252] Caruana R. Multi-Task Learning [J]. Machine Learning. 1997, 28: 41–75.
- [253] Yuan X, Liu X, Yan S. Visual Classification With Multitask Joint Sparse Representation [J]. IEEE Trans. Image Process. 2012, 21 (10): 4349–4360.
- [254] Dong G, Kuang G, Zhao L, et al. Joint Sparse Representation of Monogenic Components: with Application to Automatic Target Recognition in SAR imagery [C]. In Proc. IEEE Int'l Geosci. Remote Sens. Symp. Jul. 2014: 549–552.
- [255] Dong G, Kuang G. SAR Target Recognition via Joint Sparse Representation of Monogenic Signal [J]. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. 2015, 8 (7): 3316–3328.
- [256] Obozinski G, Taskar B, Jordan M. Joint Covariate Selection and Joint Subspace Selection for Multiple Classification Problems [J]. Statistics and Computing. 2009, 20 (2): 231–252.
- [257] Argyriou A, Evgeniou T, Pontil M. Convex Multi-Task Feature Learning [J]. Machine Learning. 2008, 73 (3): 243–272.
- [258] Bishop C. Pattern Recognition and Machine Learning [M]. New York: Springer, 2006.
- [259] Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machine: and Other Kernel-Based Learning Methods [M]. New York, USA: Cambridge University Press, 2000.

- [260] Shawe-Taylor J, Cristianini N. Kernel Methods for Pattern Analysis [M]. New York, USA: Cambridge University Press, 2000.
- [261] Hamm J. Subspace-based Learning with Grassmann Kernel [D]. USA: Computer Science Department, University of Pennsylvania, 2008.
- [262] Lyu S. Mercer Kernels for Object Recognition with Local Features [C]. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 2005: 223–229.
- [263] Zhang L, Zhou W, Chang P, et al. Kernel Sparse Representation-Based Classifier [J]. IEEE Trans. Signal Process. 2012, 60 (4): 1684–1695.
- [264] Liu J, Wu Z, Wei Z, et al. Spatial-Spectral Kernel Sparse Representation for Hyperspectral Image Classification [J]. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. 2013, 6 (6): 2462–2471.
- [265] Gao S, Tsang I W, Chia L. Sparse Representation With Kernels [J]. IEEE Trans. Image Process. 2013, 22 (2): 423–434.
- [266] Chen Y, Nasrabadi N, Tran T. Hyperspectral Image Classification via Kernel Sparse Representation [J]. IEEE Trans. Geosci. Remote Sens. 2013, 51 (1): 217–231.
- [267] Dong G, Wang N, Kuang G, et al. Kernel linear representation: application to target recognition in synthetic aperture radar images [J]. SPIE J. App. Remote Sens. 2014, 8 (1): 083613.1–13.
- [268] Camps-Valls G, Gomez-Chova L, Muñoz-Mari J, et al. Composite Kernels for Hyperspectral Image Classification [J]. IEEE Geosci. Remote Sens. Lett. 2006, 3 (1): 93–97.
- [269] Tuia D, Ratle F, Pozdnoukhov A, et al. Composite Kernels for Hyperspectral Image Classification [J]. IEEE Geosci. Remote Sens. Lett. 2010, 7 (1): 88–92.
- [270] Li J, Marpu P R, Plaza A, et al. Generalized Composite Kernel Framework for Hyperspectral Image Classification [J]. IEEE Trans. Geosci. Remote Sens. 2013, 51 (9): 4816–4829.
- [271] Dong G, Kuang G. Classification on the Monogenic Scale Space: Application to Target Recognition in SAR Image [J]. IEEE Trans. Image Process. 2015, 24 (8): 2527–2539.
- [272] Wang Z, Vemuri B, Chen Y, et al. A Constrained Variational Principle for Direct Estimation and Smoothing of the Diffusion Tensor Field from Complex DWI [J]. IEEE Trans. Med. Imag. 2004, 23 (8): 930–939.

- [273] Arsigny V, Pennec P, Ayache X. Log-Euclidean Metrics for Fast and Simple Calculus on Diffusion Tensors [J]. Magn. Reson. Med. 2006, 56 (2): 411–421.
- [274] Pennec X. Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements [J]. Journal of Mathematical Imaging and Vision. 2006, 25: 127–154.
- [275] Tuzel O, Porikli F, Meer P. Pedestrian Detection via Classification on Riemannian Manifolds [J]. IEEE Trans. Pattern Anal. Mach. Intell. 2008, 30 (10): 1713–1727.
- [276] Cherian A, Sra S, Banerjee A, et al. Efficient Similarity Search for Covariance Matrices via the Jensen-Bregman LogDet Divergence [C]. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Barcelona, Nov. 2011: 2399–2406.
- [277] Harandi M, Sanderson C, Hartley R, et al. Sparse Coding and Dictionary Learning for Symmetric Positive Definite Matrices: A Kernel Approach [C]. In Proc. Eur. Conf. Comput. Vis. 2012: 216–229.
- [278] Cherian A, Sra S, Banerjee A, et al. Jensen-Bregman LogDet Divergence with Application to Efficient Similarity Search for Covariance Matrices [J]. IEEE Trans. Pattern Anal. Mach. Intell. 2013, 35 (9): 2162–2175.
- [279] Guo K, Ishwar P, Konrad J. Action Recognition from Video Using Feature Covariance Matrices [J]. IEEE Trans. Image Process. 2013, 22 (6): 2479–2494.
- [280] 梅加强. 微分与几何初步 [M]. 北京: 科学出版社, 2011.
- [281] Moakher M, Batchelor P. Visualization and Processing of Tensor Fields [M]. NY: Springer, 2006.
- [282] Harandi M, Sanderson C, Wiliem A, et al. Kernel Analysis over Riemannian Manifolds for Visual Recognition of Actions, Pedestrians and Textures [C]. In IEEE Workshop on Applications of Computer Vision. 2012: 433–439.
- [283] Boothby W. An Introduction to Differentiable Manifolds and Riemannian Geometry [M]. NY: Academic Press, 2002.
- [284] Absil P-A, Mahony R, Sepulchre R. Riemannian Geometry of Grassmann Manifolds with a View on Algorithmic Computation [J]. Acta Applicandae Mathematica. 2004, 80: 199–220.
- [285] Turaga P, Veeraraghavan A, Srivastava A, et al. Statistical Computations on Grassmann and Stiefel Manifolds for Image and Video-Based Recognition [J]. IEEE Trans. Pattern Anal. Mach. Intell. 2011, 33 (11): 2273–2286.

-
- [286] Kim T-K, Kittler J, Cipolla R. Discriminative Learning and Recognition of Image Set Classes Using Canonical Correlations [J]. IEEE Trans. Pattern Anal. Mach. Intell. 2007, 29 (6): 1005–1018.
- [287] Edelman A, Arias T, Smith S. The Geometry of Algorithm with Orthogonality Constraints [J]. SIAM J. Matrix Analysis and Applications. 1999, 20 (2): 303–353.
- [288] Hamm J, Lee D. Grassmann Discriminant Analysis: a Unifying View on Subspace-based Learning [C]. In Proc. Int'l Conf. Machine Learning. 2008: 376–383.
- [289] Dong G, Kuang G. Target Recognition in SAR images via classification on Riemannian Manifolds [J]. IEEE Geosci. Remote Sens. Lett. 2015, 12 (1): 199–203.
- [290] Sivalingam R, Boley D, Morellas V, et al. Tensor Sparse Coding for Region Covariances [C]. In Proc. Eur. Conf. Comput. Vis. Heraklion, Greece, Sep. 2010: 722–735.
- [291] Zhang Y, Jiang Z, Davis L. Discriminative Tensor Sparse Coding for Image Classification [C]. In Proc. British Machine Vision Conf. 2013: 1–11.
- [292] Sra S. A New Metric on the Manifold of Kernel Matrices with Application to Matrix Geometric Means [C]. In Proc. Neural Inf. Process. Syst. 2012: 1–9.
- [293] Hu Y, Mian A, Owens R. Face Recognition Using Sparse Approximated Nearest Points between Image Sets [J]. IEEE Trans. Pattern Anal. Mach. Intell. 2012, 34 (10): 1992–2004.
- [294] Dong G, Kuang G. SAR Target Recognition Via Sparse Representation of Monogenic Signal on Grassmann Manifolds [J]. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. 2016, 9 (3): 1308–1319.
- [295] Harandi M, Sanderson C, Shen C, et al. Dictionary Learning and Sparse Coding on Grassmann Manifolds: An Extrinsic Solution [C]. In Proc. IEEE Int'l Conf. Comput. Vis. Sydney, Austrian, Dec. 2013: 3120–3127.
- [296] Held S, Storath M, Massopust P, et al. Steerable Wavelet Frames Based on the Riesz Transform [J]. IEEE Trans. Image Process. 2010, 19 (3): 653–667.
- [297] Unser M, Chenouard N, Ville D V D. Steerable Pyramids and Tight Wavelet Frames in $L_2(\mathbb{R}^d)$ [J]. IEEE Trans. Image Process. 2011, 20 (10): 2705–2721.
- [298] Dempster A P. A Generalization of Bayesian Inference [J]. Journal of the Royal Statistical Society. Series B (Methodological). 1968: 205–247.

作者在学期间取得的学术成果

发表的学术论文

- [1] Ganggang Dong, Gangyao Kuang. Classification on the Monogenic Scale-Space: Application to Target Recognition in SAR Image. *IEEE Transaction on Image Processing*, 2015, 24(8):2527-2539. (SCI 源刊, WOS:000354459700001, 被引次数: 4 次¹)
- [2] Ganggang Dong, Gangyao Kuang, Linjun Zhao, Jun Lu. SAR Target Recognition via Joint Sparse Representation of Monogenic Signal. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(6): 3316-3328. (SCI 源刊, WOS:000359538700007, 被引次数: 10 次)
- [3] Ganggang Dong, Gangyao Kuang. SAR Target Recognition Via Sparse Representation of Monogenic Signal on Grassmann Manifolds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 9(3): 1308-1319. (SCI 源刊, WOS:000373054100028, 被引次数: 2 次)
- [4] Ganggang Dong, Gangyao Kuang. Sparse Representation of Monogenic Signal: With Application to Target Recognition in SAR Images, *IEEE Signal Processing Letters*, 2014, 21(8):952-956. (SCI 源刊, WOS:000336792600004, 被引次数: 22 次)
- [5] Ganggang Dong, Gangyao Kuang. Target Recognition in SAR Images via Classification on Riemannian Manifolds. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(1):199-203. (SCI 源刊, WOS:000340952500041, 被引次数: 14 次)
- [6] Ganggang Dong, Gangyao Kuang. Target Recognition via Information Aggregation Through Dempster-Shafer's Evidence Theory. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(6):1247-1251. (SCI 源刊, WOS:000352571800018, 被引次数: 12 次)
- [7] Ganggang Dong, Gangyao Kuang. A Soft Decision Rule for Sparse Signal Modeling via Dempster-Shafer Evidential Reasoning. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(10): 1567-1571. (SCI 源刊)

¹被引频次数据来源于 Google Scholar <http://scholar.google.com.cn/>, 截止时间: September 10, 2016。

-
- [8] Ganggang Dong, Na Wang, Gangyao Kuang, Yinfang Zhang. Kernel Linear Representation: Application to Target Recognition in Synthetic Aperture Radar Images. *SPIE Journal of Applied Remote Sensing*, 2014, 8:083613-1-14. (SCI 源刊, WOS:000338499400001, 被引次数: 5 次)
 - [9] Ganggang Dong, Gangyao Kuang, Linjun Zhao, Jun Lu, Joint Sparse Representation of Monogenic Components: With application to Automatic Target Recognition in SAR Imagery, *IEEE International Geoscience and Remote Sensing Symposium 2014*, 549-552. (EI: 20144800241501)
 - [10] Ganggang Dong, Gangyao Kuang, Linjun Zhao, Jun Lu, Nonnegative and Local Linear Regression for Classification in SAR Imagery, *IEEE International Geoscience and Remote Sensing Symposium 2014*, 1702-1705. (EI: 20144800241333)
 - [11] Ganggang Dong, Gangyao Kuang. Target Recognition via Sparse Representation of Frequency Spectrum. *IEEE Transaction on Geoscience and Remote Sensing*, —审.
 - [12] Ganggang Dong, Gangyao Kuang. Classification via Sparse Representation of Steerable Wavelet Frames on Grassmann Manifold: Application to Target Recognition in SAR Image. *IEEE Transaction on Image Processing*, 二审.
 - [13] Siqian Zhang, Yutao Zhu, **Ganggang Dong**, Gangyao Kuang. Truncated SVD-Based Compressive Sensing for Downward-Looking Three-Dimensional SAR Imaging With Uniform/Nonuniform Linear Array. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(9):1853-1857. (SCI 源刊, WOS:000359579000011)
 - [14] Siqian Zhang, **Ganggang Dong**, Gangyao Kuang. Super-Resolution Downward-Looking Linear Array Three-Dimensional SAR Imaging Based on Two-Dimensional Compressive Sensing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 9(6): 2184-2196. (SCI 源刊, WOS:000379935100008)
 - [15] Siqian Zhang, Gangyao Kuang, Yutao Zhu, **Ganggang Dong**. Compressive Sensing Algorithm for Downward-looking Sparse Array 3-D SAR Imaging. *IET International Radar Conference*, 2015. (EI: 20162402497692)

研究成果

- [1] 匡纲要, 赵凌君, 熊博莅, 项德良, 董刚刚. SAR 图像目标识别方法: 中国, CN103268496A. (中国专利公开号)

附录 A 向量空间

度量空间是满足度量公理的非空集合，若度量空间的基本列都收敛于其中的点称其为完备的度量空间；赋范线性空间是一类特殊的度量空间，它是在线性空间中赋予范数以后得到，它不仅满足加法和数乘两种运算结构，还考虑了它们之间的逼近关系，完备的赋范线性空间称为 Banach 空间，它与范数有关；内积空间是一种重要的赋范线性空间，它是欧氏空间向无穷维的推广，完备的内积空间称为 Hilbert 空间。

Metric Space 度量空间

定义 A.1 (度量空间): 设 \mathbf{X} 是非空集合，映射 $d : \mathbf{X} \times \mathbf{X} \rightarrow \mathbb{R}$ 满足：对任意 $x, y, z \in \mathbf{X}$, 有

非负性 $d(x, y) \geq 0$, 取等号当且仅当 $x = y$;

对称性 $d(x, y) = d(y, x)$;

三角形不等式 $d(x, z) \leq d(x, y) + d(y, z)$;

称 d 是 \mathbf{X} 上的度量（或距离函数），集合连同其上的度量即二元对 (\mathbf{X}, d) （简记为 \mathbf{X} 称为度量空间）。

Normed Linear Space 赋范线性空间

定义 A.2 (赋范线性空间): 对于数域 Ω 上的任意线性空间 \mathbf{X} , 若映射 $\|\cdot\| : \mathbf{X} \mapsto \mathbb{R}$ 满足：对一切 $\mathbf{x}, \mathbf{y} \in \mathbf{X}, \alpha \in \Omega$ 都有：

非负性 $\|\mathbf{x}\| \geq 0$, $\|\mathbf{x}\| = 0$ 当且仅当 $x = 0$;

齐次性 $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$;

三角形不等式 $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$;

则称映射 $\|\cdot\|$ 是 \mathbf{X} 上的范数， $(\mathbf{X}, \|\cdot\|)$ 是赋范线性空间，简记 \mathbf{X} 。

Inner Product Space 内积空间

定义 A.3 (内积空间): 设 \mathbf{X} 是数域对于数域 Ω （实数域或复数域）上的线性空间，若映射 $\langle \cdot, \cdot \rangle : \mathbf{X} \times \mathbf{X} \mapsto \Omega$ 满足：对于任意 $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{X}, \alpha, \beta \in \Omega$,

正定性 $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, 取等号当且仅当 $\mathbf{x} = 0$;

共轭对称性 $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$;

对第一变元线性 $\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle$, 其中 $\alpha, \beta \in \Omega$;

则称 $\langle \cdot, \cdot \rangle$ 是内积, $(\mathbf{X}, \langle \cdot, \cdot \rangle)$ 称为内积空间。

内积空间是一类赋范线性空间, 由内积可以导出范数, 完备的内积空间称为 Hilbert 空间。

引理 A.1 (Schwarz 不等式): 设 \mathbf{X} 是内积空间, $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, 有 $\|\langle \mathbf{x}, \mathbf{y} \rangle\| \leq \sqrt{\mathbf{x} \cdot \mathbf{x}} \sqrt{\mathbf{y} \cdot \mathbf{y}}$.

内积作为映射 $\langle \cdot, \cdot \rangle \rightarrow \Omega$ 是连续的。

定义 A.4 (直和): 设 $\mathbf{X}_1, \mathbf{X}_2$ 是线性空间 \mathbf{X} 的两个子空间, 若它们满足

- $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$
- $\mathbf{X}_1 \cap \mathbf{X}_2 = \{0\}$

称 \mathbf{X} 是 \mathbf{X}_1 和 \mathbf{X}_2 的直和, 记为 $\mathbf{X}_1 \oplus \mathbf{X}_2$, $\mathbf{X}_1, \mathbf{X}_2$ 是各自的余子空间。

定义 A.5 (格莱姆矩阵): n 维欧氏空间中任意 $k, k \leq n$ 个向量 $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}^n$ 的内积所组成的矩阵

$$\Delta(\alpha_1, \alpha_2, \dots, \alpha_k) = \begin{pmatrix} \langle \alpha_1, \alpha_1 \rangle & \langle \alpha_1, \alpha_2 \rangle & \dots & \langle \alpha_1, \alpha_k \rangle \\ \langle \alpha_2, \alpha_1 \rangle & \langle \alpha_2, \alpha_2 \rangle & \dots & \langle \alpha_2, \alpha_k \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \alpha_k, \alpha_1 \rangle & \langle \alpha_k, \alpha_2 \rangle & \dots & \langle \alpha_k, \alpha_k \rangle \end{pmatrix}$$

称为 k 个向量 $\alpha_1, \alpha_2, \dots, \alpha_k$ 的格莱姆矩阵。格莱姆矩阵是半正定矩阵, 它成为正定矩阵的前提是 $\alpha_1, \alpha_2, \dots, \alpha_k$ 线性无关。

附录 B Dempster-Shafer 证据理论

DS 证据理论是由哈佛大学和伦敦经济学院数学家 Arthur P. Dempster 在利用上、下限概率解决多值映射问题时形成雏形 [211]，并在“*A Generalization of Bayes Inference*”一文正式提出 [298]，后经他的学生 Glenn Shafer 进一步扩展，引入置信函数，形成一整套基于“证据”和“合并”来处理不确定性推理问题的数学方法。

B.1 DS 数据融合

贝叶斯理论和 DS 证据理论目的都是赋予事件多种可能状态的权重，并最终对事件的状态进行推测，其中贝叶斯给所有互斥状态指定它们的概率，而 DS 证据理论进一步加入“不确定状态”，例如我们可以将两个或两个以上的状态组合起来生成一个新的状态。贝叶斯理论将每种状态的权值称为“概率”，而 DS 证据理论则引入另一种性质类似的概率——“证据质量”(Masses)，并据此引出“置信函数”(Belief Function) 和“似然函数”(Plausibility Function)，其中前者是对假设信任程度的下限估计，即悲观估计；后者是对假设信任程度的上限估计，即乐观估计。不同信息源的证据通过 Dempster 组合规则进行合并，对于冥集内的假设 A, B, C ，两个信源证据的合并规则为

$$m(C) = \frac{\sum_{A \cap B = C} m_1(A)m_2(B)}{1 - \sum_{A \cap B = \emptyset} m_1(A)m_2(B)}. \quad (\text{B.1})$$

DS 证据理论应用于实际问题时最大的难题是如何定义证据质量，目前该问题并没有统一的结论，已有的解决方法包括贝叶斯概率值、模糊数学理论、误差项等。

假定有两部传感器同时对未知的空中目标进行监测，已经掌握的情报信息表明，该目标属于 F-111 “土豚”战斗机、F/A-18 “大黄蜂”舰载机和 P-3C 海上巡逻机。在现有三种假设的基础上，我们通常还会增加两种额外的组合状态：

Fast 我们认为未知目标速度很快，但是无法辨别它是 F-111 和 F/A-18 哪种超音速战斗机，用 DS 证据理论的术语概括为 $\{F_{111} \cup F/A_{18}\}$ 。

Unknown 我们没有关于未知目标的更多信息，只能认为它是个飞行器，无法根据现有信息辨别它具体属于哪一类，即 $\{F_{111} \cup F/A_{18} \cup P_{3C}\}$ 。

证据分配如表 B.1 所示，第二列和第三列分别是每个传感器分配的证据质量，最后一列是使用 Dempster 合并规则计算生成的全局证据。根据三种飞行器类别总共有 8 种可能的假设（即冥集包括 2^3 种假设），其中焦元只有五种假设，DS 证据理论要求冥集包含的所有假设的证据质量之和为 1，需要声明的是此处的证据质

量并非概率，以 Sensor1 为例，Fast 和 Unknown 两种假设的证据质量分别为 42% 和 10%，若用证概率根本无法解释。

表 B.1 不同假设的证据质量分配。

Hypothesis	Sensor 1 (Mass m ¹)	Sensor 2 (Mass m ²)	Fused Evidence (Mass m ^{1,2})
F-111	30%	40%	55%
F/A-18	15%	10%	16%
P-3C	3%	2%	0.4%
Fast	42%	45%	29%
Unknown	10%	3%	0.3%
Total mass	100%	100%	100%

使用 Dempster 合并规则对 Sensor 1 和 Sensor 2 的证据质量进行融合，为了便于计算，我们不再计算用于归一化的冲突证据，以比例的形式分别给出前三种简单假设的证据质量：

$$\begin{aligned}
 m^{1,2}(F_{111}) &\propto \sum_{A \cap B = F_{111}} m^1(A)m^2(B) \\
 &= m^1(Unknown) \cdot m^2(F_{111}) + m^1(Fast) \cdot m^2(F_{111}) \\
 &\quad + m^1(F_{111}) \cdot m^2(F_{111}) + m^1(F_{111}) \cdot m^2(Fast) + m^1(F_{111}) \cdot m^2(Unknown)
 \end{aligned} \tag{B.2}$$

$$\begin{aligned}
 m^{1,2}(F/A_{18}) &\propto \sum_{A \cap B = F/A_{18}} m^1(A)m^2(B) \\
 &= m^1(Unknown) \cdot m^2(F/A_{18}) + m^1(Fast) \cdot m^2(F/A_{18}) \\
 &\quad + m^1(F/A_{18}) \cdot m^2(F/A_{18}) + m^1(F/A_{18}) \cdot m^2(Fast) + m^1(F/A_{18}) \cdot m^2(Unknown)
 \end{aligned} \tag{B.3}$$

$$\begin{aligned}
 m^{1,2}(P_{3C}) &\propto \sum_{A \cap B = P_{3C}} m^1(A)m^2(B) \\
 &= m^1(Unknown) \cdot m^2(P_{3C}) + m^1(P_{3C}) \cdot m^2(P_{3C}) + m^1(P_{3C}) \cdot m^2(Unknown)
 \end{aligned} \tag{B.4}$$

B.2 Bayes 数据融合

贝叶斯理论将两个事件 A, B 同时发生的联合概率记为 $p(A, B)$ ，它与条件概率 $p(A|B)$ 满足

$$p(A, B) = p(A|B)p(B) \tag{B.5}$$

联合概率 $p(A, B) = p(B, A)$, 因此我们可推导条件概率

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} \quad (\text{B.6})$$

如果事件 A 含有多子与事件 B 相区别的事件, 我们可继续推导

$$p(A|B) = \frac{p(B|A)p(A)}{\sum_i p(B|A_i)p(A_i)} \quad (\text{B.7})$$

公式 (B.7) 就是著名的贝叶斯规则, 广泛用于数据融合。贝叶斯规则将整个统计事件划分成为多个独立事件, 期望能够从一组数据中最准确地估计未知参数。贝叶斯理论需要寻找出某个最优值 x 进行决策, 两种最常用的最优估计是:

Maximum Likelihood Estimate 最大似然估计, $\hat{x} = \arg \max p(\text{data}|x)$

Maximum A Posterior Estimate 最大后验估计, $\hat{x} = \arg \max p(x|\text{data})$

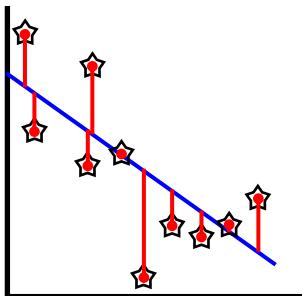
贝叶斯理论应用时遇到的通常困难就是所谓的先验概率, 即公式 (B.7) 中的 $p(A)$ 。仍然讨论第 B.1 节目标追踪的例子, 我们通常期望能够从有用参数的一组测量值确认飞行器的类别, 根据贝叶斯规则可知未知目标属于三种机型的概率分别为

$$\begin{aligned} p(F_{117}|\text{data}) &= \frac{p(\text{data}|F_{117})p(F_{117})}{p(\text{data}|F_{117})p(F_{117}) + p(\text{data}|F/A_{18})p(F/A_{18}) + p(\text{data}|P_{3C})p(P_{3C})} \\ p(F_{117}|\text{data}) &= \frac{p(\text{data}|F/A_{18})p(F/A_{18})}{p(\text{data}|F_{117})p(F_{117}) + p(\text{data}|F/A_{18})p(F/A_{18}) + p(\text{data}|P_{3C})p(P_{3C})} \\ p(F_{117}|\text{data}) &= \frac{p(\text{data}|P_{3C})p(P_{3C})}{p(\text{data}|F_{117})p(F_{117}) + p(\text{data}|F/A_{18})p(F/A_{18}) + p(\text{data}|P_{3C})p(P_{3C})} \end{aligned} \quad (\text{B.8})$$

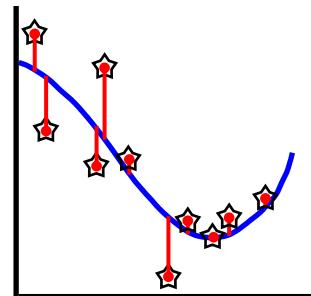
对于似然概率 $p(\text{data}|F_{117}), p(\text{data}|F/A_{18}), p(\text{data}|P_{3C})$, 通常可以根据测量值进行计算, 但是先验概率 $p(F_{117}), p(F/A_{18}), p(P_{3C})$ 如何计算止前尚没有合理的解决方法。

附录 C 线性回归与核回归分析

线性回归是利用线性回归方程的最小平方函数对一个或多个自变量和因变量之间关系进行建模的一种回归分析。如果回归分析中只包括一个自变量和一个因变量，且二者的关系可用一条直线近似表示，称其为一元线性回归分析；如果回归分析中包括两个或两个以上的自变量，且因变量和自变量之间满足线性关系，称其为多元线性回归分析。多元线性回归模型的求解方法中最为常见的是岭回归，它是一种用于共线性数据分析的有偏估计回归方法，实质是一种改良的最小二乘估计，通过放弃最小二乘法的无偏性，以损失部分信息、降低精度为代价，获得对病态数据的耐受性。为了应对线性不可分或是线性可分性较弱等情形，通常借助非线性映射将回归参量和回归响应投影到 Hilbert 空间，使用它们的像构成核回归模型。论文的第 4 章和第 5 章实质是用稀疏表示求解岭回归和核回归的回归系数，然后利用回归系数进行判决。图 C.1 以图例的方式对线性回归和核回归进行介绍，它们的优劣对比见表 C.1，其中 \mathbf{X}, \mathbf{y} 表示回归参量和响应， \mathbf{K} 表示回归参量的像构成的格莱姆矩阵， \mathbf{K}_y 表示回归参量的像与回归响应的像生成的相似向量， d, n 表示回归参量维度和数目。



(a) Linear Ridge Regression



(b) Kernel Ridge Regression

图 C.1 线性岭回归与核岭回归示意图，图中给出了线性回归与核回归模型的拟合示意曲线。

表 C.1 线性回归分析与核回归分析优劣对比。

	Linear Ridge Regression	Kernel Ridge Regression
回归系数	$\alpha = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y}$	$\beta = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{K}_y$
约束程度	苛刻	富有弹性
训练复杂度	$O(d^3 + d^2 n)$	$O(n^3)$
预测复杂度	$O(d)$	$O(n)$
优劣说明	$n \gg d$ 计算更高效	样本数较多时效率不高