Multidimensional Signal Processing:
Fast Transform, Sparse Representation, and Low Rank Analysis

多维信号处理：
快速变换、稀疏表示与低秩分析

戴琼海 著
Dai Qionghai

清华大学出版社

Multidimensional Signal Processing:
Fast Transform, Sparse Representation, and Low Rank Analysis

多维信号处理：
快速变换、稀疏表示与低秩分析

戴琼海 著

Dai Qionghai

清华大学出版社

北 京

内 容 简 介

多数信号处理论著主要针对理论与方法臻备的一维信号，而对于仍在发展完善中的多维信号处理少有涉及或涉之不深。本书凝聚著者在多媒体信号处理领域十余年的研究成果，以快速变换、稀疏表示、低秩分析为理论主线，内容涉及图像/视频的感采样、表示、编码、滤波、恢复、三维重建等应用。本书系统介绍了多维离散余弦变换与离散小波变换的快速分解方法、过完备双树小波变换包优选方法及其图像/视频编码与降噪应用、图像信号的自回归压缩感知方法、重加权矩阵低秩恢复模型以及对数和矩阵低秩填充模型、基于低秩分析的光照度立体重建与三维运动场估计等。

本书可以作为从事信号处理等领域科技工作者和工程技术人员的参考资料，同时也可作为高等院校相关专业高年级本科生和研究生的参考书。

图书在版编目(CIP)数据

多维信号处理：快速变换、稀疏表示与低秩分析/戴琼海著. —北京：清华大学出版社，2016
ISBN 978-7-302-43475-7

Ⅰ.①多…　Ⅱ.①戴…　Ⅲ.①信号处理－研究　Ⅳ.①TN911.7

中国版本图书馆 CIP 数据核字(2016)第 078270 号

责任编辑：王一玲
封面设计：常雪影
责任校对：梁　毅
责任印制：沈　露

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编**：100084
社 总 机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质 量 反 馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
课 件 下 载：http://www.tup.com.cn，010-62795954
印 装 者：三河市少明印务有限公司
经　　销：全国新华书店
开　　本：175mm×245mm　**印　张**：13.5　**彩　插**：3　**字　　数**：307 千字
版　　次：2016 年 7 月第 1 版　　**印　　次**：2016 年 7 月第1 次印刷
印　　数：1～1500
定　　价：49.00 元

产品编号：065715-01

前言

PREFACE

数字信号处理历久弥新，新的理论和应用层出不穷。20 世纪 40 年代，在傅里叶分析的理论基础上，以香农为代表的专家学者构建了信息论的基础理论。70 年代，专业数字信号处理器的问世与推广将数字信号处理带入到了实际应用阶段。90 年代，大众消费电子促进了数字信号处理的飞速发展。21 世纪以来，伴随着图像、视频等多维信号规模的不断增长，数据维数已经超越了数据处理能力的发展速度，"维数灾难"随之产生。更为严峻的是，芯片处理能力飞速增长所依赖的工艺尺度已经逼近物理极限，而数据仍以指数级别膨胀，多维信号处理成为新的研究热点和难点。

面对多维信号，学者们一方面致力于将多维信号处理转化为基本的单维信号单元来进行快速处理与应用，例如多维离散余弦变换（m-D DCT）与多维离散小波变换（m-D DWT）的快速变换、针对图像/视频的多尺度几何变换等。另一方面则探索新的信号表示与处理理论，例如压缩感知、稀疏表示、低秩分析等，为多维信号处理开启了新的篇章。作者在从事信号处理、图像处理和视频处理的长期科学研究中，深刻地感受到快速变换、稀疏表示与低秩分析在多维信号处理中的重要地位，并体现在作者和合作者在国际权威和著名杂志发表的一系列论文中。为了系统性地向读者介绍多维信号处理的相关知识和研究体会，特编写此书，以飨读者。

本书共 9 章，首先在第 1 章回顾与展望多维信号处理三大基本处理工具（快速变换、稀疏表示以及低秩分析）的发展历程，然后通过 8 章的内容系统性地阐述了三大工具的基础知识以及作者的研究体会与经验，主要内容包括：

第一部分　快速变换：针对应用最广泛的离散余弦变换和小波变换，介绍两者在多维变换情况下的快速算法。具体而言，第 2 章描述如何通过 m-D DCT 变换矩阵的张量积分解，使得 m-D DCT 与 m-D IDCT 转化为若干独立的 1-D DCT，从而便于并行与电路实现；第 3 章介绍 m-D DWT 的脉动阵列 VSLI 结构设计，以实现低硬件复杂度、规则数据流和低控制复杂度的架构。

第二部分　稀疏表示：在详细讨论压缩感知这一亚奈奎斯特稀疏采样的理论基础上，介绍了一种常用的稀疏表示方法——冗余小波变换。具体来讲，第 4 章介绍压缩感知的理论依据、核心优化问题与常见解法，并讨论其在医学成像、荧光显微、超分辨率重建、计算摄像等不同场景的应用；

第5章以双树离散小波为例，讨论小波基方向性对图像/视频表示的重要影响，给出基函数自适应分解与优选方法，展示冗余变换经过率失真优化的稀疏表示之后可以显著提升图像/视频编码性能，甚至超越现有采用正交变换的编码标准。

第三部分 低秩分析：在给出矩阵低秩分解理论后，通过视频修复、三维重建、点云融合等实例介绍了低秩分析的具体应用方法。其中，第6章介绍矩阵秩最小化与矩阵低秩分解理论，包括无噪矩阵填充、稀疏大噪声分离两种经典模型，给出其在视频恢复、超声波断层成像、光照传输特性重建、光度立体、图像对齐等方面的应用案例；第7章介绍如何通过两种非凸启发式模型的理论关联，将低秩分析引入矩阵低秩结构学习框架，并展示其在静态三维点云融合中的应用；第8章介绍矩阵低秩分析如何解决动态三维重建中的运动场恢复问题；第9章系统性地给出了信号自相关表示、自相关模型、堆积矩阵与低秩分析的理论联系，介绍如何通过堆积矩阵低秩分解实现信号最优稀疏基的逼近，展示其在压缩采样与超分辨率的成功应用。

在本书成稿及修订过程中，博士后颜成刚、博士生杨敬钰、李坤、彭义刚、付长军、邓岳等人参与了多次讨论，提出一些很好的想法，作者对他们表示深深的谢意，同时感谢业内人士对该书提出的宝贵意见。由于本书源自作者的教学和科学研究之体会，借此机会感谢国家九七三重点基础研究发展规划项目基金、教育部“长江学者奖励计划”、国家自然科学基金委、重大专项以及省部级产学研等课题的资助。

囿于作者的理解水平和能力，错误与不妥之处在所难免，敬请广大同行和读者不吝赐教，以便再版时补充和修改。

戴琼海

2015年10月于北京清华大学

目录

CONTENTS

第1章 多维信号处理的回顾与展望

1.1 引言

随着信息技术的日臻成熟以及现代社会的全方位信息化,多维信号的获取、处理、传输和呈现等已经在国防、经济、医疗、广播电视等领域获得广泛应用。典型应用包括对地观测、金融分析、影像诊断、影视媒体等。这些应用所关心的指标通常随多种因素变化,需要用多维信号进行表示与刻画。例如,高光谱图像是水平坐标、竖直坐标以及波长的三维信号;核磁共振三维造影是空间位置的三维信号;立体视频则是水平坐标、竖直坐标、视角、时间的四维信号。信号空间维度的增加使得信号的变化具有更多的自由度,高维信号的特征比一维信号更为丰富,也更为复杂。多维信号处理的挑战在于如何高效地刻画或者利用这些丰富而复杂的高维信号特征,这同时也是多维信号处理研究的中心主题。近年互联网技术进展迅猛,社交网络(social network)与云存储、与云计算的兴起更使得多维信号呈指数级增长,信息技术从而进入大数据(big data)时代。多维信号的高效处理因而比过去任何一个时代都显得更为重要,也更为迫切。

早期多维信号处理方法大都是一维信号处理方法的可分离扩展(separable extension),即将多维信号的各个维度分别按一维信号进行处理。多维离散傅里叶变换(m-dimensional discrete Fourier transform,m-D DFT)、多维离散余弦变换(m-dimensional discrete cosine transform,m-D DCT)、多维离散小波变换(m-dimensional discrete wavelet transform,m-D DWT)等多维变换都是相应一维变换以可分离的方式扩展得到。尽管这些多维变换计算复杂度通常低于非可分离的多维变换,近年仍有许多研究致力于设计更为精巧的快速变换,以进一步适应于大规模多维媒体数据的在线实时处理。

采用可分离方式扩展得到的多维变换具有复杂度低易于实现的优点。但是,其基函数是相应一维变换基函数在多个维度上的张量积。高维奇异特征(诸如图像与视频中的边缘与轮廓)不能简单地分解为两个一维信号的乘积,因而难以采用可分离变换高效表示。近几十年来,探索高维信号

特征的高效表达一直是多维信号处理方面的活跃领域。20 世纪 90 年代，信号稀疏表示研究兴起，给多维信号的高效表达带来了新的框架[1]。在这个框架下，设计高效表达多维信号特征的方法大致可分为两类：(1)通过观察信号抽象出共性特征设计相应的数学变换，例如，强调图像方向特征的 Curvelet[2]，Contourlet[3]，Directionlet[4]等；(2)构建合适的能量函数，从信号中学习一组可以获得高效表达的字典，例如稀疏编码(Spase Coding)[5]，K-SVD[6]，在线字典训练(online dictionary learning)[7]等。这两类方法通常突破正交性的约束，囊括更多的基函数/原子，在刻画信号特征方面拥有更大的灵活性；由于基函数/原子数目大于信号样本点数，因此稀疏表示中的过完备变换/字典，又称冗余变换/字典。

信号在过完备字典上具有稀疏表示本质上是因为它在信号外围空间中具有低维结构。矩阵低秩分析是另一种探求信号低维结构的数学框架。在稀疏表示中，高维信号被向量化为一维向量进行处理，其处理的基本对象是向量；与稀疏表示不同，低秩分析处理的基本对象是矩阵。近几年来，矩阵低秩与稀疏表示交相辉映，是近年信号处理学术界最具有影响力的两个领域。

1.2 多维信号快速变换

1.2.1 快速 m-D DCT

DCT 是信号处理的核心运算，是 JPEG、MPEG-2、MPEG-4 等国际编码标准的标准组件。因此，学术界与工业界有许多学者和工程师致力于 DCT 快速算法的研究与开发。

2-D DCT 的快速分解可分为行列分解法和直接分解法。行列分解法主要是将 $N\times N$ 的数据按行(列)进行 N 个 1-D DCT 计算，产生中间矩阵，然后对中间矩阵再按列(行)进行 N 个 1-D DCT 计算，最后得到 2-D DCT 结果。直接分解法直接在二维数据上设计变换结构，所需计算量通常少于行列分解法[8-15]。例如，Haque 将二维数据分解为更小的子块进行快速分解[9]；Ma 采用基 2×2 结构构造了一种与 2-D FFT 类似的递归快速算法[10]，节省约 25% 的乘法操作。将 2-D DCT 转换为 1-D DCT 或一维奇数 DFT 也是常用的方法，根据转换方法的不同可以分为多项式变换和切比雪夫多项式变换。Vetterli 等人[8]建立 2-D DCT 到实数 2-D DFT 的映射关系，通过多项式变换将 2-D DFT 转化为更小规模的 2-D DFT 和一维奇数 DFT，从而将乘法次数降至传统行列分解法的 50%。Duhamel 和 Guillemot[11]提出了二维 DCT 的间接多项式变换方法：先将 2-D DCT 输入行换序，使之等价于二维奇数 DFT，从而转换为 $2^{(n-1)}$ 个一维奇 DFT 求得，所需乘法运算量减至传统行列分解法的 50%。Morikawa 等人[16]采用切比雪夫多项式同余形式，把二维 DCT 转换为 N 个 N 点一维 DCT 和一个长度为 N 的切比雪夫多项式变换，所需计算量与经典行列

分解法较少约 50%。Cho and Lee[12] 采用三角函数法，将长度为 2^N 的二维 DCT，表示为两个新的二维变换之和，再利用换序移位和附加的实数加法运算，将两个新二维变换转换为 N 个 N 点一维 DCT，算法的思路和切比雪夫多项式变换相同，它的乘法复杂性和切比雪夫多项式变换相同。

对于 m-D DCT，国内外学者也提出了许多快速算法，典型思路包括：把尺寸较大的 m-D DCT 分解为若干尺寸较小的 m-D DCT 计算[17]；把 $N\times N\times\cdots\times N$（$N$ 为 2 的乘方）的 m-D DCT 分解为多个独立的长度为 N 的 1-D DCT，可大幅减少计算量，但是不能满足每维长度不同的应用[18]。Chen 等人[19] 通过运用变换矩阵的张量积分解将尺寸为 $N_1\times N_2\times\cdots\times N_m$（$N_i$ 是 2 的乘方且小于 256）的 m-D DCT 分解为 1-D DCT、加法、移位等基本操作。若所有维度的信号长度相同，乘法运算量将是传统行列分解法的 $1/m$。Dai 等人[20] 将 $N_1\times N_2\times\cdots\times N_m$（$N_i$ 是 2 的乘方）的 m 维 DCT 分解为 2^m 个子块来进行单独的快速运算，其中只涉及两步运算：r 维（$r=1,2,\cdots,m$）DCT 逆变换，以及一些乘法和加法运算。另外，DCT 可以通过多项式变换转化为 DFT，从而可以利用 FFT 进行快速计算。Zeng 等人[21] 提出一种新的多项式变换，将 m-D DCT 直接分解为一系列的 1-D DCT，所需乘法次数减少为行列分解方法的 $1/m$，同时其具有更好的计算结构。更进一步，在基于多项式变换的快速 m-D DCT 框架下，运用 Ramanujan 有序数（Ramanujan ordered numbers）来计算余弦角可以将乘法运算替换为换序移位和加法运算，从而获得无乘法的快速 m-D DCT[22]。

1.2.2 快速 m-D DWT

DCT 与 DFT 一脉相承，其基函数由一系列在信号区间做等幅振荡的周期三角函数构成，因而擅长表示与分析平稳信号。对于非平稳信号的表示与分析则存在一些缺陷。例如，基于 DCT 的低码率图像/视频编码在边缘与轮廓等瞬变信号周围常产生振铃效应（ringing artifacts）；与 DFT 类似，DCT 可以分析信号所含的频率分量，但是不能指出某一频率分量何时开始何时结束，也就是说它没有均衡可控的时频分辨率。随着时频分析的兴起，小波分析在 20 世纪 90 年代完成理论构建，并且在信号处理、数字通信、机器视觉、金融分析等领域均获得成功应用。因此，DWT 高效实现对多维信号实时处理具有重要意义。

DWT 等价于采用一组满足小波关系的低通滤波器和高通滤波器对信号进行迭代分解，其快速计算方法可以分为两类：(1) 卷积运算；(2) 提升方案（lifting scheme）。与卷积相比，提升方案在低通滤波与高通滤波之间共用一部分提升操作从而获得更低（约 50%）的计算量。与 m-D DCT 相似，最直接的算法是将 m-D DWT 以可分离方式转换为一系列各个维度上的 1-D DWT。由于需要储存中间数据，对内存需求量大，同时也存在输出延迟和系统延迟。非可分离方式实现不需要中间数据的传输，但是通常需要更多的乘法和累加运算。因此，许多研究者提出了许多针对大规模集成（very large scale integration，VLSI）电路的小波变换算法结构，

以对地观测，视频压缩等应用对数据规模、处理时间的要求。

虽然已有许多针对 1-D DWT 的 VLSI 架构，然而设计低硬件成本与高吞吐率的 m-D DWT VLSI 架构仍然充满挑战。Lewis 和 Knowles[23]针对四抽头 Daubechies 滤波器设计一种无乘法器的二维小波变换架构。Parhi 和 Nishitani[24]结合字并行（word-parallel）和数字串行（Digital-serial）方法提出两种 2-D DWT 架构。Vishwanath 等人[25]针对 2-D DWT 提出一种结合脉动滤波器和并行滤波器的脉动-并行（systolic-parallel）架构。Chakrabarti 和 Vishwanath[26]提出两种高效的不可分架构，即并行滤波器和 SIMD 二维阵列，从而对空间和时间进行优化。Chuang 和 Chen[27]提出一种针对 2-D DWT 的并行流水线式 VLSI 阵列架构。Chen 和 Bayoumi[28]提出一种可伸缩的脉动阵列架构。Chakrabarti 和 Mumford[29]提出一种 2-D DWT 折叠架构和调度算法。在各种 VLSI 架构中，有两种在硬件成本与计算速度均表现出色的 2-D DWT 架构。其中一种架构利用多相分解和系数折叠技术实现可分离 2-D DWT 计算[30]。另外一种架构基于修正递归塔型算法计算不可分分离 2-D DWT[31]。这两种架构均可在 $2N^2/3$ 个时钟周期内对 $N\times N$ 图像进行 DWT 分解。由于设计复杂，少有文献涉及更高维 m-D DWT 的 VLSI 架构。

1.3 多维信号稀疏表示

正交变换（如 DFT、DCT 与 DWT 等）为信号提供唯一的表示。变换的正交性是一种优良的数学性质，带来了许多处理与设计的方便，例如，易于进行信号处理失真的分析与优化，因而在信号处理中获得广泛应用。作为信号表示的新进展，冗余信号表示（包括冗余变换与冗余字典）松弛了正交性的要求，拥有比信号维度更多的基函数/原子，从而也为信号提供了不唯一（理论上无穷种）的表示。冗余系统的表达多样性给信号处理带来了更大的灵活性与自由度，同时也提出了新问题。例如，给定冗余变换/冗余字典，可以在广阔的表达空间中寻找满足某种需求的信号表示；另一方面，如何定义所需的信号表达，如何将其刻画为可计算的测度，并且给出求解该表达的算法？信号稀疏表示的根本出发点正是从无穷多组表示系数中找一组最紧致（稀疏）的表达。信号表示的紧致性（稀疏度）反映了基函数对信号特征的刻画能力，这就与信号逼近、特征选择、模式辨识等重要问题建立了直接联系——信号稀疏表示在图像/视频压缩、神经信号解码、人脸识别等诸多研究领域获得成功应用。信号稀疏表示有两个基本问题：(1)针对信号特征与应用需求设计合适的冗余变换/冗余字典；(2)给出该冗余系统下的稀疏表示高效求解算法。信号稀疏表示滥觞于稀疏表示求解算法的研究，下面简要回顾稀疏表示算法的发展。

早期的稀疏表示求解方法属于贪婪算法。1993 年，Mallat 与 Zhang 提出了匹配追踪（matching pursuit，MP）信号分解方法，并以基于时频字典（time-frequency dictionary）的自适应信号分解为例展示了这一信号表示新框架[32]。MP 算法非常简洁，即将当前残差信号投影到字典中与其内积最大的原子，选中该原子及其系数，减

去投影分量以更新残差信号,依次迭代直至收敛。此后,匹配追踪有许多扩展,包括最为人熟知的正交匹配追踪(orthogonal matching pursuit,OMP)[33]、阶梯正交匹配追踪(stagewise orthogonal matching pursuit,StOMP)[34]、基于树状结构的匹配追踪(tree-based matching pursuit,TBP)[35]等。MP 及其各种改进均为贪婪算法,即每一步都确保所选原子对残差信号表示的最优性,但是却无法保证最终所选字典子集对原始信号给予最稀疏的表示。需要注意的是,同时也存在其他求取线性系统稀疏表示的贪婪算法[36]。由于计算复杂度低,匹配追踪类算法至今仍不失为快速求解稀疏表示的有效方法。

在匹配追踪类算法崭露头角之时,许多研究者立足于对稀疏度的数学刻画,将稀疏系数的求解表述为现代优化理论擅长解决的模型,发展出许多高效的稀疏表示求解方法。1995 年,Chen 等人指出标架法(method of frame,MOF)、MP、最优正交基(best orthogonal basis,BOB)等方法在寻求信号稀疏表示的不足,以求解冗余系统稀疏表示为目标,提出了最小化系数 ℓ_1 范数的稀疏表示求解模型,称为基追踪方法(basis pursuit,BP)[37]。除 BP 之外,同时期各个领域也有一些研究工作采用 ℓ_1 范数最小化的信号处理研究,包括用于时间序列处理的 ℓ_1 范数解卷积[38]、针对线性回归的最小绝对收缩与选择算子(least absolute shrinkage and selection operator,LASSO)[39]、面向稀疏信号重建的聚焦欠定系统求解法(focal under-determined system solver,FOCUSS)[40]。ℓ_1 范数最小化尽管带来了比匹配追踪类算法更为稀疏的表示,但并不能保证得到最稀疏的表示。Donoho 与 Huo 指出信号稀疏度可以用 ℓ_0 范数(即非零元素数目)来表示,而稀疏表示问题可以表述为在重建约束下的 ℓ_0 范数最小化问题[41]。遗憾的是,该优化问题(在统计建模中也称"子集选择")是 NP 问题,不存在多项式时间复杂度的求解方法[36]。另一方面,BP 主要研究者 Chen 在博士论文中[42]给出了令人深思的试验结果:从冗余字典中选择少数几个原子合成信号,BP 算法能完全恢复信号合成中所选择的原子及其系数。这引起 Donoho 与 Huo 关于 ℓ_1 范数最小化与 ℓ_0 范数最小化等价性的思考[41]:(1)数学上,ℓ_1 范数是 ℓ_0 范数的凸松弛;(2)对于一些特定情况,ℓ_1 范数优化的解与 ℓ_0 范数优化的解相同。ℓ_1 最小化与 ℓ_0 最小化的等价性具有非常重要的理论价值,引起很多学者的兴趣,为稀疏表示的发展奠定了坚实的基础[43-45]。有了理论上的保障,许多研究力量集中在 ℓ_1 范数最小化的高效算法设计上。早期的算法,例如 BP,将 ℓ_1 范数最小化转化为线性规划,从而采用单纯形法(simplex method)或者内点法(interior point method)来求解。但是这些方法的计算与储存复杂度随问题规模增大而急剧增大。而图像处理、压缩感知、生物信息学等领域的问题规模通常都非常大,这些方法难以直接应用。研究人员将目光投向曾因收敛慢而被忽略的一阶方法(First-order Methods,即只需用梯度等一阶信息),提出了许多快速高效的稀疏表示求解算法[46],例如,迭代阈值-收缩法(iterative shinkage-thresholding methods)[47-49]、梯度投影法(gradient projection methods)[50]、邻近梯度法(proximal point methods)[51,52]、增强拉格朗日乘子法(augmented Lagrange multiplier methods)[53,54]。这些算法的核心步骤只需

要向量代数操作和矩阵向量相乘计算，而不像内点法中需要大量的矩阵相乘与矩阵分解运算。时至今日，稀疏表示算法仍然层出不穷。各种算法在收敛速度、空间与计算复杂度、求解精度、问题规模适应性等各方面均有所区别，要根据具体问题与运行环境择优选取。

早期稀疏表示研究的出发点是寻求参数化冗余字典下的紧致表达。这类冗余字典，包括 Gabor 字典、小波包（wavelet packets），余弦包（cosine packets）等，通常具有优良的数学性质，例如良好的时频局域化、时频均匀覆盖等[55]。例如，Mallat 与 Zhang 在提出 MP 的工作[32]中采用了以尺度、位置、频率为参数的 Gabor 字典。Neff 和 Zakhor 采用二维 Gabor 字典，提出了一种基于 MP 的低比特率时频编码[56]——这是在图像/视频编码中采用冗余表示的经典工作。从光滑性、时频局域化等数学性质为出发点设计的冗余字典具有较好的通用性，但是往往在信号特征的自适应刻画方面有所不足。另一类冗余字典由样本信号训练得到，通常比参数化冗余字典获得更稀疏的表达。作为该领域最早的代表性工作，Olshausen 与 Field 在 Nature 上发表了关于自然图像稀疏编码的著名研究结果[57]。他们发现，通过施加稀疏先验，训练得到的基函数同时具备局部性（localized）、方向性（oriented）、带通性（bandpass）三种性质。这与哺乳动物初级视觉皮层中由简单细胞组成的感受视野的基本性质相似——该研究揭示视神经响应与图像统计结构具有紧密联系，同时也指出了稀疏编码在图像处理中的巨大潜力。Vinje 和 Gallant[58]则用生理学实验验证了视觉初级皮层对自然场景的稀疏表示。Olshausen 与 Sallee 将稀疏编码基函数的训练方法拓展到多尺度的情况[59]。值得注意的是，独立成分分析（independent component analysis，ICA）从自然图像学习得到的独立成分分量也具有类似的性质[60]。Kreutz-Delgado 等人[61]系统地研究了面向稀疏表示的字典学习方法，指出了稀疏编码与矢量量化的内在联系，并探索字典学习在信号分离与图像编码中的应用。Aharon 等人提出了另一种字典学习算法，称为 K-SVD[6]。该方法可以看做是 K-means 分类算法的推广，又由于核心步骤采用 SVD 分解，因而得名。K-SVD 是近年应用较为广泛的字典学习方法，被拓展至视频稀疏字典训练、三维字典训练等场合，在去噪、压缩、超分辨率等经典问题中获得成功应用。Mairal 等人[7]基于随机近似技术提出了在线字典学习方法，解决了实际应用中百万量级样本的字典学习问题。目前，面向图像稀疏表示的字典学习方法已经趋向成熟。值得注意的是，目前大多数字典学习方法仅以稀疏表示为目标，而图像特征的高效表达除了稀疏之外还有许多重要的性质，例如平移不变性、尺度不变性、旋转不变性等。稀疏编码的早期贡献者 Olshausen 与合作者将实数域冗余字典学习推广到复数域学习，初步研究了稀疏表示变换不变性的学习机制。稀疏表示字典学习还有非常广阔的空间待进一步探索。

1.4 多维信号低秩分析

矩阵的低秩逼近在机器学习、控制、推荐系统及计算机视觉等领域具有重要的作用。近年来，基于低秩特性的矩阵重建（matrix reconstruction）继稀疏表示与压缩

感知之后成为热点研究领域。矩阵重建大致可分为：矩阵填充(matrix completion)和矩阵恢复(matrix recovery)两类问题。矩阵填充的任务是从矩阵的部分观测数据中预测出缺失元素的值；而矩阵恢复则从带噪数据中恢复原矩阵。两者可类比于图像处理中的图像补绘(inpainting)与图像去噪(denoising)。

正如稀疏表示中给定冗余系统存在无数种信号表示，矩阵填充也是一个欠定问题。如果不对矩阵进行某种约束，理论上也存在无数个满足观测的完整矩阵。在许多应用(例如推荐系统、运动检测、图像对齐)的数据(或直接或经过处理之后)所对应的矩阵具有低秩特性(low-rankness)——低秩成了矩阵重建领域最重要的先验。因而，矩阵重建被表述为观测约束下的秩最小化问题。与稀疏表示中 ℓ_0 范数最小化问题相似，矩阵的秩最小化是 NP 难问题。Candès 和 Recht[62]采用核范数(nuclear norm)对将矩阵秩进行凸松弛，给出了基于半正定规划的凸优化算法，推导出在该求解框架下进行精矩阵填充的矩阵元素采样界限，同时也展示了最小化秩与最小化核范数的等价性。注意到半正定规划不能求解大规模问题的缺点，Cai 等人[63]提出了一种近似求解核范数最小化的一阶算法，称为奇异值阈值(singular value thresholding,SVT)方法。随后，Toh 和 Yunti[64]采用 Nesterov 技术对 SVT 算法进行加速，提出加速邻近梯度法(accelerated proximal gradient,APG)，将收敛速度提升到 $O(K^{-2})$。

作为矩阵恢复的代表性模型，鲁棒主成分分析(robust principle component analysis,RPCA)[65,66]将观测矩阵分解为待恢复矩阵与误差矩阵之和，分别用核范数与 ℓ_1 范数进行约束，如何高效求解颇具挑战性。文献[65]将原问题转化为二阶近似问题，针对其拉格朗日函数提出了一种与矩阵填充 SVT 算法[63]类似的迭代阈值方法。Ganesh 等人[67]提出 RPCA 模型的 APG 算法。Toh 和 Yun[64]进一步提出了基于加速 APG 的核范数最小化算法。Lin[68]提出了一种基于增广拉格朗日乘子法的矩阵恢复方法，比迭代阈值与 APG 算法都具有更优的收敛性能。矩阵恢复模型也可以在 ALM 框架进行高效求解，其收敛性能常优于 APG 算法[68,69]。Tao 和 Yuan[70]进一步扩展 RPCA 模型，加入另外一误差分量以对稠密的高斯白噪声进行建模，给出了基于 ALM 的求解算法。虽然该算法通常可以得到很好的实验结果，但是收敛性证明还属于开放性问题。Peng 等人[71]提出了加权低秩矩阵恢复模型，将核范数与 ℓ_1 范数推广至更一般的加权核范数与加权 ℓ_1，给出增强低秩与稀疏性的权重计算方法以及基于 ALM 框架的加权低秩矩阵恢复算法。除了以迭代阈值为核心操作的一阶算法之外，还有一些研究采用其他函数来逼近矩阵的秩。Fornasier 等人[72]采用迭代重加权最小二乘(iteratively reweighted least squares)来近似求解核范数最小化问题。Log-det 函数也是另一种在优化中逼近矩阵秩的采用手段[73,74]。

随着模型与算法的日臻完善，矩阵重建在许多领域都获得了非常成功的应用。Ji 等人[75]将联合稀疏与低秩模型用于视频恢复；Peng 等人[76]提出了一种基于矩阵稀疏与低秩分解的图像对齐方法；Zhang 等人[77]提出了一种用低秩进行纹理描述的框架；Li 等人[78]将矩阵填充引入到三维运动估计领域中；Deng 等人[79]提出基于矩

阵恢复的三维重建方法。

1.5 本章小结

本章概览了多维信号的快速变换、稀疏表示以及低秩分析。在快速变换方面，分别介绍了 m-D DCT、m-D DWT、MCLT 等三类重要快速变换算法的发展；在稀疏表示方面，分别对冗余系统设计与稀疏表示算法的发展脉络进行梳理；对于低秩分析，则分别综述矩阵填充与矩阵恢复两类模型的求解方法。值得注意的是，稀疏表示（包括压缩感知[80]）与矩阵重建有着非常紧密的联系[62,63]。从模型层面，稀疏表示与矩阵重建的原始模型都属于 NP 问题，都需要通过适当的松弛以获得高效求解：矩阵重建理论采用核范数对矩阵秩进行凸松弛，这与稀疏表示理论采用 ℓ_1 范数对 ℓ_0 范数进行凸松具有非常工整的对偶关系。从算法层面，矩阵重建与稀疏表示在优化求解上可以找到很好的对应，例如两者均以迭代阈值/收缩作为核心操作，两者都有很多基于 APG 思想和 ALM 框架的求解算法，两者都可以采用 Nesterov 技巧进行加速。从某种程度上来说，两者具有物理意义上的等价关系，稀疏表示研究的是向量的稀疏，而矩阵重建探索的是矩阵的稀疏。虽然数学对象维度不同，但是它们追求紧致表达的精髓则高度一致。