

多媒体信号处理的数学理论前沿进展

熊红凯¹, 戴文睿¹, 林宙晨², 吴飞³, 于俊清⁴, 申扬眉¹, 徐明星¹

1. 电子信息与电气工程学院, 上海交通大学, 上海 200240;

2. 信息科学技术学院, 北京大学, 北京 100080;

3. 计算机科学与技术学院, 浙江大学, 杭州 310027;

4. 计算机科学与技术学院, 华中科技大学, 武汉 430074

摘要: **目的** 深度学习模型广泛应用于多媒体信号处理领域, 通过引入非线性能够获得极大地性能提升, 但是其黑箱结构无法解析地给出最优点和优化条件。因此如何基于传统信号处理理论, 利用基于变换/基映射模型逼近深度学习模型, 解析优化问题, 成为当前研究的前沿问题。**方法** 本文从信号处理的基础理论出发, 分析了当前针对高维非线性非规则结构方法的数学模型和理论边界, 主要包括: 结构化稀疏表示模型、基于框架理论的深度网络模型、多层卷积稀疏编码模型以及图信号处理理论。具体地, 详细描述了基于组稀疏性和层次化稀疏性的表示模型和优化方法, 分析基于半离散框架和卷积稀疏编码构建深度/多层网络模型, 进一步在非欧空间上扩展形成图信号处理模型, 并对国内外关于记忆网络的研究进展进行了比较。**结果** 结构化稀疏方法通过混合的联合范式正则项, 利用近端算法或网络流优化求解组 Lasso 目标问题, 获取包含结构信息的更优的稀疏表示。鉴于结构化稀疏表示仍然是立足于基集的线性映射, 基于框架理论的深度网络成为焦点。它通过延拓多尺度几何分析中的半离散框架来生成深度卷积网络, 引入非线性形成对于形变稳定的特征表示, 并继承框架理论良好的尺度和方向分解性质。受散射网络启发, 多层卷积稀疏编码在稀疏编码基础上, 引入组合正则约束, 拟合最大化池化操作, 并递归逐层学习和分解稀疏表示的完备字典, 形成表示多尺度非规则结构的稀疏表示。图信号处理理论基于非欧空间上的拓扑结构扩展传统信号处理理论, 通过结合卷积神经网络学习复杂关系网络, 适用于数据驱动的高维不规则信号的分析 and 处理。**结论** 最后, 本文展望了多媒体信号处理的理论模型发展, 认为图信号处理通过解析谱图模型的数学性质, 解释其中的关联性, 为建立广义的大规模非规则多媒体信号处理模型提供理论基础, 是未来研究的重要领域之一。

关键词: 结构化稀疏表示模型; 基于框架理论的深度网络模型; 多层卷积稀疏编码模型; 图信号处理理论; 多媒体信号处理

Advances in mathematical theory for multimedia signal processing

Xiong Hongkai¹, Dai Wenrui¹, Lin Zhouchen², Wu Fei³, Yu Junqing⁴, Shen Yangmei¹, Xu Mingxing¹

1.School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240

2.School of Electronic Engineering and Computer Science, Peking University, Beijing 100080

3.College of Computer Science and Technology, Zhejiang University, Hangzhou, 310027

4.School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, 430074

基金项目: (要求有项目编号) 不同基金之间用分号隔开, 国家级的基金需英译

收稿日期: ; **修回日期:**

Supported by: 国家级基金项目英译 (基金项目英译见学报网站下载中心)

Abstract: Objective Deep learning models have been widely used in multimedia signals processing. They significantly improve the performance of signal processing tasks by introducing non-linearities, but lacks analytical formulation of optimum and optimality conditions due to their black-box architectures. In recent years, it is popular to analyze the optimal formulation and approximate the deep learning models based on classical signal processing theory for multimedia, i.e., transform/basis projection based models. **Method** This report presents and analyzes the mathematical models and their theoretical bounds for high-dimensional nonlinear and irregular structured methods based on the fundamental theories of signal processing. The main content of this paper includes structured sparse representation, frame-based deep networks, multi-layer convolutional sparse coding and graph signal processing. We begin with sparse representation models based on group sparsity and hierarchical sparsity with their optimization methods, and subsequently, analyze the deep/multi-layer networks developed using semi-discrete frames and convolutional sparse coding. We also present graph signal processing models by extending classical signal processing to the non-Euclidean geometry. Recent advances in these topics achieved by domestic and abroad researchers are compared and discussed. **Result** Structured sparse representation introduces the mixed norms to formulate a group Lasso problem for structural information, which can be solved using proximal method or network flow optimization. Considering that structured sparse representation is still based on the linear projection onto dictionary atoms, frame-based deep networks are developed to extend the semi-discrete frames in multiscale geometric analysis. They inherit the scale and directional decomposition led by frame theory and introduce nonlinearities to guarantee deformation stability. Inspired by scattering networks, multi-layer convolutional sparse coding introduces combined regularization into sparse representation to fit max pooling operation. Sparse representation of irregular multiscale structures can be achieved with the trained overcomplete dictionary in a recursive manner. Graph signal processing extends conventional signal processing into the non-Euclidean spaces. When integrated with convolutional neural networks, graph neural networks learn complex relational networks and are desirable for data-driven large-scale high-dimensional irregular signal processing. **Conclusion** In the end, this paper forecasts the future work of mathematical theories and models for multimedia signal processing. It is significant to develop generalized graph signal processing model for large-scale irregular multimedia signals by analyzing the mathematical properties and linkages of conventional signal processing and graph spectral model.

Key words: Structured sparse representation; frame-based deep convolutional network; multi-layer convolutional sparse coding; graph signal processing; multimedia signal processing

0 引言

传统多媒体信号处理的基础理论主要建立在各类解析表示(变换)上。傅里叶变换(Fourier Transform)作为周期平稳信号描述,可扩展到多维空间来描述图像和视频;小波变换则通过有效逼近非平稳信号的点奇异性,适用于表示分段光滑函数。二维小波由一维小波的张量积构建得到,不同分辨率下,它的基具有不同尺寸大小的正方形支撑区间。因此,二维小波以点奇异性来逼近曲线奇异性,缺乏方向性。多尺度几何分析理论构建一种新的最优多维信号表示方法,将重要结构特征映射于低维子集中,如二维影像的主要特征可由其中光滑方向性边缘所刻画。自然影像统计模型和对人类视觉系统的研究显示,基于多尺度几何分析的最优图像表示方法,具有多分辨率、局部特性、方向性和各向异性等性质。

多尺度几何分析分为自适应和非自适应两类。自适应方法一般先进行边缘检测,再利用边缘信息对原函数进行最优表示,主要代表为楔形波(Wedgelet)变换(Donoho, 1999),条带波(Bandelet)变换(Le Pennec等, 2005)和方向波(Directionlet)变换(VELISAVLJEVIĆ等, 2006)。非自适应方法则不需要图像几何特性的先验知识,而是通过可解析变换逼近信号奇异性,以脊波(Ridgelet)变换(Candès, 1998)、曲线波(Curvelet)变换(Candès等, 2000)和轮廓波(Contourlet)变换(Do等, 2005)为代表,并衍生出一系列扩展和改进工作。如剪切波(Shearlet)变换(Labate等, 2005)以剪切操作代替旋转来改进曲线波变换;轮廓波可通过引入树状结构的多维方向滤波器组表示多维信号中曲面奇异的表面波(Surfacelet)变换(Lu等, 2007)和实现非均匀方向频率分解(Xiong等, 2011)等进行扩展;基于图像内容规律性在任意支集延拓哈尔(Haar)小波的群波(Grouplet)变换(Mallat, 2009)。

针对图像多尺度分析,以过完备原子集合(字典)来替代全局固定的基函数,是稀疏表示的另一种重要途径。鉴于冗余系统的欠定性,信号的最优分解方式并不唯一,因此需要基于误差最小化原则自适应地寻找最稀疏的系数分布。1993年, Mallat和Zhang首先提出匹配追踪算法(Mallat等, 1993)实现这种想法,通过信号在过完备字典上的分解,根据信号统计特性自适应地选取基函数,得到稀疏

表示。随后,发展出了一系列寻找基函数的算法,诸如基追踪算法(Basis Pursuit, BP)(Chen等, 1998)、正交匹配追踪算法(Orthogonal Matching Pursuit, OMP)(Pati等, 1993)和迭代阈值算法(Iterative Thresholding)(Blumensath等, 2008)等,能够在给定约束下,将图像/视频信号的能量集中到分布更为稀疏的系数上去,从而取得比基于离散余弦变换(DCT)和离散小波变换(DWT)更好的分析和逼近结果。随着学习算法的引入,基于采样自适应训练字典受到广泛关注,如最优方向方法(Method of Optimal Directions, MOD)(Engan等, 1999)、正交基联合(Union of Orthobases)(Lesage等, 2005)和广义主成分分析(Generalized PCA, GPCA)(Vidal等, 2005)等,通过迭代学习寻找信号能量最为集中的基函数或者最优的方向,自适应构建字典。2006年, Elad等提出了一种基于迭代调整和更新的K-SVD算法(Aharon等, 2006),相较于传统冗余字典,能有效减少稀疏表示所需的字典原子数,并且仍然可以训练后原子的线性组合表示初始字典,提高重构视觉信号的效率。K-SVD可以通过组合基于分析和基于学习字典(Rubinstein等, 2010),引入基于图像分析的稀疏表示(Rubinstein等, 2013)以及建立多任务字典学习模型(Ruvolo等, 2013)来进一步优化。2010年和2011年, Krause等和Mairal等分别提出亚模字典选择(Krause等, 2010)和在线字典学习(Mairal等, 2010)等学习算法,也能在保证表示和逼近性能的同时,提升其运行效率。

虽然基于过完备字典学习能够有效提升多媒体信号的稀疏表示性能,然而基于字典的线性映射限制了对于非规则结构的表达能力。随着卷积神经网络在多媒体信号表示、逼近和分析中取得重大突破,却由于其“黑箱”性质,缺少数学可描述的理论模型,解析优化条件和收敛性质,限制了重构稳定性和适用大规模信号的泛化性。然而,受到卷积神经网络的启发,其中关键方法如非线性池化操作和局部平移等变卷积操作逐渐受到关注,被引入到传统信号处理理论框架中,在数学可解释性的保证下,有效提升了多媒体信号处理和分析能力,并被引入到图信号处理中,通过与表示相关性的图结构结合,利用多媒体信号中的高维非规则拓扑奇异性。

本文从信号处理的基础理论出发,重点介绍目前受到广泛关注的针对高维非规则多媒体信号处理的数学模型和理论方法,主要包括:结构化稀疏表示模型、基于框架理论的深度网络模型、多层卷积

稀疏编码模型以及图信号处理理论。鉴于传统低维信号的稀疏表示采用 ℓ_0 或 ℓ_1 范式约束，孤立地优化字典原子，无法表征信号潜在的结构关联，结构化稀疏表示方法通过混合联合范式正则项，利用近端算法或网络流优化求解组 Lasso 目标问题，获取包含结构信息的更优稀疏表示。扩展互不相交的组稀疏为相互重叠的层次化稀疏，衍生符合多维网格结构的层次化分组 Lasso，拓展至有向非循环图的分层结构，形成非规则稀疏性的图 Lasso 问题。鉴于结构化稀疏表示仍然立足于基集的线性映射，基于框架理论的深度网络通过延拓多尺度几何分析中的半离散框架来生成深度卷积网络，引入非线性形成对于形变稳定的特征表示，并继承框架理论的尺度和方向分解性质。散射网络通过多尺度小波变换系数的期望，等效平均池化操作，延拓形成卷积网络，实现对于多媒体信号的任意尺度和方向分解，并扩展至一般化的半离散框架。受散射网络启发，多层卷积稀疏编码在稀疏编码基础上，引入 ℓ_0 和 ℓ_∞ 的组合正则约束，拟合最大化池化操作，并递归逐层分解学习过完备字典，形成多尺度非规则结构的稀疏表示。考虑到卷积神经网络多基于传统网格结构，而忽略非规则结构的内在拓扑特征，近年来，图信号处理理论，旨在非欧空间上进行基于拓扑结构的信号处理，将传统的信号处理理论扩展到图表示的非规则信号领域，考虑图谱结构与图结构上信号的关联性。利用谱图理论可以图傅立叶变换扩展包括时频变换及其算子、滤波器设计等的传统信号处理理论，然而其泛化性能和可扩展性受制于对特定图结构的依赖性。因此，基于传统信号处理的理论框架，解析谱图模型的数学性质，解释其中的关联性，为广义的大规模非规则多媒体信号处理模型提供理论基础，成为当前研究中的重要领域。

本文的章节安排如下：第 1 章和第 2 章分别介绍了在结构化稀疏表示模型、基于框架理论的深度网络模型、多层卷积稀疏编码模型以及图信号处理理论方面的国际和国内研究现状；第 3 章对于国际和国内研究进展做了比较；第 4 章分析目前的研究趋势，并对未来的研究方向进行展望。

1 国际研究现状

1.1 结构化稀疏表示模型

在统计、信号处理与机器学习等研究领域，稀疏性是一个核心的科学概念。稀疏表示通常采取变量或特征选择的问题形式，旨在通过对少量特征基

原子的线性组合，为目标数据寻求一种最为简洁、紧致的表示方法。从目标函数角度，稀疏估计本质上为 ℓ_0 范式约束的组合优化问题，而引入 ℓ_1 范式正则项可以将该非凸的组合优化进行有效的凸松弛（Bach 等，2011；Beck 等，2009；Efron 等，2004；Yuan，2010）。针对这类经典的稀疏表示方法，一系列的理论框架与应用实践迅速发展（Mairal 等，2014）。

以 ℓ_1 范式正则化主导的经典稀疏表示问题，对每一个变量单独进行选择，而不考虑变量之间的结构特征与潜在关系。然而，在许多实际应用场景中，表示估计可以受益于与具体问题相关的结构性先验知识，进而提升预测与近似性能。例如，在基于功能磁共振（fMRI）或脑磁图（MEG）的神经影像中，体素集合往往组织形成小尺寸的局部连通区域，通过观察这种结构特征，有助于对不同脑状态的区分与医学诊断（Gramfort 等，2009）。类似的，在人脸识别任务中，以像素集合形成面部的小尺寸凸区域作为特征，可以有效提升识别对遮挡问题的鲁棒性（Jenatton 等，2010）。朴素的 ℓ_1 范式无法编码这些高阶结构性约束信息，因此，结构化稀疏表示方法（Structured Sparsity）应运而生，通过 ℓ_1/ℓ_q （ $q \in \{2, \infty\}$ ）联合范式诱导稀疏性的正则项，能够编码信号的稀疏性模式所隐含的高阶的结构化信息，从而得到更优的稀疏逼近（Bach 等，2012；Jenatton 等，2009）。

1.1.1 组稀疏性

组稀疏性（Group Sparsity）为最基本的结构化稀疏形式，所适配的结构性先验为，在稀疏分解过程中，变量分布于彼此互不相交的分组（Disjoint Groups）中，而同一分组中的变量会被同时选中或丢弃。这样的结构约束可通过联合 ℓ_1/ℓ_q 范式表达，在最小二乘回归的背景中，该正则项被称之为组 Lasso（Group Lasso）问题（Turlach 等，2005；Yuan 等，2006），如公式（1）所示。

$$\min_{\beta \in \mathbb{R}^p} \left\| y - \beta_0 - \sum_{i=1}^g X_i^T \beta_i \right\|_F^2 + \lambda \sum_{i=1}^g \sqrt{p_i} \|\beta_i\|_2 \quad (1)$$

其中， $X \in \mathbb{R}^{d \times n}$ 为字典， $y \in \mathbb{R}^d$ 为输入信号， $\beta \in \mathbb{R}^p$ 为稀疏向量， λ 为正则化参数平衡稀疏性与保真度， $p = \sum_{i=1}^g p_i$ ，稀疏向量的索引 $\{1, \dots, p\}$ 分为 g 个不相交的组。大量的理论与实践表明，当实际问题与组结构特性相关，引入这种正则形式能够有效提升模型的预测性能与可解释性（Huang 等，2010）。例如，

组稀疏字典对图像的视觉描述子袋 (Bag of Visual Descriptors) 构造分组结构, 通过组稀疏正则项来自适应地择取对表达原始信号最相关的组, 获取图像级的稀疏性, 进而提升图像分类的预测精度 (Bengio 等, 2009)。同时, 组稀疏约束在多任务学习 (Multi-task Learning) 中可以利用不同任务之间共享的特征 (Obozinski 等, 2010), 在多核学习 (Multiple Kernel Learning) 中对不同核进行有效选取 (Bach, 2007)。

1.1.2 层次化稀疏性

互不相交的组稀疏可进一步扩展为相互重叠的分组模式 (Overlapping Groups), 即不同分组中可以包含相互重叠的变量, 而同一分组中的变量在稀疏分解过程中被同时选中或落选。由这种结构性先验出发, 可以设计一系列的范式形式, 匹配不同的问题结构。例如, 相互毗连的稀疏性模式适用于一维的时序或基因序列, 而二维网格中的矩形稀疏性模式则适用于图像特征提取。通过构建二维网格的分块结构以矩形空间邻近的 $p \times p$ 个变量进行分组, 与之对应的结构性稀疏约束在视频序列的背景抽取问题上取得了很好的性能 (Huang 等, 2011)。受到地形独立成分分析 (Topographic Independent Component Analysis) 启发, 法国 INRIA 的 Mairal 等将这种结构应用于字典学习问题上, 提出地形字典 (Topographic Dictionary) 的概念, 利用网络流优化算法 (Network Flow Optimization) 求解非平滑的目标问题, 使得学习得到的字典具有显著的空间平滑性 (Mairal 等, 2011)。另一种层次化稀疏结构 (Hierarchical Sparsity) 受到了广泛关注与研究, 通过将变量组织为分层的树状结构, 该结构约束项使顶层变量优先于底层变量选择。该稀疏结构最早以回归问题中的稀疏分组 Lasso (Sparse Group Lasso) 形式出现 (Friedman 等, 2010), 如公式 (2) 所示。

$$\min_{\beta \in \mathbb{R}^p} \left\| y - \beta_0 - \sum_{i=1}^g X_i^T \beta_i \right\|_F^2 + \lambda \sum_{g \in \mathcal{G}} \omega_g \|\beta_g\| \quad (2)$$

其中, 稀疏向量的索引 $\{1, \dots, p\}$ 分为 $|\mathcal{G}|$ 个相互重叠的组, ω_g 为第 g 个组的非负权重, β_g 为 β 中以 g 索引的子向量。层次化稀疏先后在基于小波的图像降噪问题中结合多尺度分解的天然层次化结构 (Zhao 等, 2009), 在生物信息学中利用基因网络的树形结构进行多任务回归 (Multi-task Regression) 与 fMRI 数据的多尺度挖掘 (Multi-scale Mining), 预测简单的认知任务 (Jenatton 等, 2012; Kim 等, 2010)。Jenatton

提出层次化稀疏字典学习 (Hierarchical Dictionary Learning) 问题, 以字典原子为节点构造树状分层结构, 并利用层次化分组 Lasso (Hierarchical Group Lasso) 惩罚项对信号分解的稀疏性模式进行约束, 通过近端梯度下降算法 (Proximal Gradient Descent) 求解非平滑线性不可分离的优化目标问题, 使之形成与字典树同根的连通的子树结构。该字典可以应用于自然图像复原问题, 学习得到的字典原子自顶向下具有信号频率逐层递增的关系, 也可应用于文本文档的主题建模 (Topic Modeling), 形成自根至叶的由粗到精的主题层级 (Topic Hierarchy), 提升文本的分类精度 (Jenatton 等, 2010)。

此外, 考虑更为复杂的拓扑结构, 可以进一步推演出更多的扩展, 如三维空间的立方体离散化 (3D Spaces Discretized in Cubes), 或球体的切片离散化 (Spherical Volumes Discretized in Slices) (Varoquaux 等, 2010), 以及基于有向非循环图 (Directed Acyclic Graphs) 的分层结构 (Bach, 2008), 和以图的边 (Edges) 进行分组来选择连通分量的图 Lasso (Graph Lasso) 问题等 (Jacob 等, 2009)。假定 G 是无向高斯图模型 (Undirected Gaussian Graphical Models), 则顶点的条件独立性对应于精度矩阵 (Precision Matrix) 中的非零项, 模型选择等价于精度矩阵中非零元素的选择。据此, 图 Lasso 如公式 (3) 所示。

$$\min_{\Theta} \log \det \Theta - \text{trace}(S\Theta) - \lambda \|\Theta\|_1 \quad (3)$$

其中, $S = \sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x})^T$ 为经验协方差矩阵, \bar{x} 为样本均值, $\ell(\Theta) = \log \det \Theta - \text{trace}(S\Theta)$ 为样本的对数最大似然。在深度学习方面, Wen 等提出结构化稀疏性学习 (Structured Sparsity Learning) 方法来正则化深度神经网络 (Deep Neural Networks, DNNs) 结构 (滤波器通道数、滤波器形状、网络层数等), 在保证预测精度前提下, 学习紧实的、硬件友好的网络架构, 降低计算复杂度, 减少所需运算资源 (Wen 等, 2016)。

1.2 基于框架理论的深度网络模型

自 2010 年 AlexNet (Krizhevsky 等, 2012) 在 ImageNet 大规模视觉识别竞赛 (Large Scale Visual Recognition Challenge, LSVRC) 拔得头筹, 卷积神经网络由于对形变稳定的不变性受到广泛关注。一类手工设计的卷积网络, 通过设计局部化的滤波器, 能够分析网络最优化配置和性质, 成为了新的研究热点。其中, 最具代表性的工作是由 Mallat 等提出

表 1 基于框架的深度网络模型概述

Table 1 An overview of frame-based deep networks

深度网络模型	深度网络特点
散射网络 (Bruna 等, 2013; Mallat, 2012)	群平移不变、结构稳定的 Morlet 小波非线性特征提取子
旋转不变散射网络 (Oyallon 等, 2015; Sifre 等, 2012)	叠加散射网络, 得到具有旋转不变性的特征提取子
混合散射网络 (Oyallon 等, 2017)	结合无监督散射网络与神经网络, 减少训练和预测开销
核学习的散射网络 (Xiong 等, 2016)	利用支持向量机 (SVM) 最大边际估计特点, 学习散射网络中的卷积核与路径
基于半离散框架的通用特征提取子 (Wiatowski 等, 2018)	扩展散射网络理论概念, 拓展散射网络为基于多种核函数和非线性单元的半离散框架
基于滤波器组的卷积网络 (Xu 等, 2017, 2016)	利用滤波器组构建深度卷积网络, 具有完美重构的能力, 克服了散射网络可逆性很弱的问题
分数小波散射网络 (Liu 等, 2019)	广义平移不变性的分数阶小波的散射网络理论模型
深度卷积框架波 (Ye 等, 2017)	分析网络不具备完美重构能力的原因, 提出基于框架理论的框架波

的散射网络 (Scattering Network) (Mallat, 2012), 能够实现平移不变和形变稳定的非线性特征提取。散射网络参考卷积神经网络的架构, 利用尺度和方向上变化的 Morlet 小波基作为卷积核, 以取模 (Module) 运算引入非线性操作, 递归地将取模运算结果作为卷积核的输入, 堆叠形成网络, 在不同尺度和方向上对图像进行特征提取。同卷积神经网络比较, 散射网络在逐层递归迭代中, 能够保持特征的尺度不变性和方向协变性。并且由于取模运算保持层间能量不变, 因此网络的稳定性能够得到保证。Mallat 等将散射该网络用于特征提取, 并以支持向量机 (Support Vector Machine, SVM) 基于所提取特征进行分类, 能够获得优于传统手工设计特征的方法, 并接近深度学习方法 (Bruna 等, 2013)。在散射网络基础上, 可以进一步引入新的性质, 或者与卷积神经网络进一步结合, 满足广泛的图像处理任务。例如, 在纹理分类任务中, 基于空间和角度变量的一维小波滤波器, 构建深度散射网络, 将方向协变性扩展为方向不变性, 使得产生的特征对于旋转后的纹理图像保持稳定, 减小分类器的学习难度 (Oyallon 等, 2015; Sifre 等, 2012)。散射网络可以进一步结合卷积神经网络, 形成混合深度网络。将无监督散射网络作为混合网络的初始特征提取层, 为后续卷积神经网络提供高维特征, 在保证性能的情况下, 减少网络层数及训练时间 (Oyallon 等, 2017)。

鉴于散射网络基于方向小波的半离散紧框架和取模运算构成, Wiatowski 等考虑更具一般性的半离散框架 (Semi-discrete Frame) 及 Lipschitz 连续的非线性函数和池化操作, 构造通用特征提取网络 (Wiatowski 等, 2018)。针对包括小波、脊波、曲

线波、剪切波和 Gabor 框架等在内的方法, 基于框架理论构造深度网络, 进行理论分析, 并组合修正线性单元 (Rectified Linear Units, ReLU)、平移 Sigmoid 函数 (Shifted Logistic Sigmoids), 双曲正切函数 (Hyperbolic Tangents) 和取模运算等 Lipschitz 连续的非线性操作, 以及 Lipschitz 连续池化操作模拟亚采样和平均等离散池化操作。半离散框架的深度网络放宽了构成散射网络的允许条件 (Admissibility Condition), 特征提取的垂直平移不变性随网络层数增加而渐进增强。在不引入池化操作时, 具有与散射网络相同的水平平移不变性。半离散框架的深度网络对带限函数、分段平滑的结构函数 (Cartoon Function) 和 Lipschitz 函数描述的信号, 能够保持形变稳定性, 保证形变敏感性上界。

然而, 上述网络主要针对特征提取, 基于 Morlet 小波构造分解算子, 不能完成从所提取特征到原始图像的逆变换, 保证完美重构。考虑框架波泛化了用于求解逆问题的低秩汉克尔 (Hankel) 矩阵方法 (Yin 等, 2017), 深度卷积框架波 (Deep Convolutional Framelets) (Ye 等, 2017) 利用多层卷积框架波和修正线性单元建立完美重构的深度神经网络。类似于卷积神经网络的构造方式, 深度卷积框架波基于框架波引入冗余滤波器通道和建立残差块, 拼接修正线性单元引入非线性, 利用高通分支的信息增强池化/反池化层, 保证完美重构。在分解过程中, 将线性修正单元及其原点对称的函数同时作用于特征图上, 获得分别对应隐含层表示 (Hidden Representation) 正负两部分系数的特征图。合成时, 拼接这两张特征图获得原隐含层表示, 进行重构。上述过程可逐层进行, 并最终保证整个网络的重构性能。基于框架波的性质, 深度卷积框架

波揭示了深度学习的信号表示，能够基于非局部基和数据驱动的局部基的组合来描述，延伸了经典信号处理理论。同时，深度卷积框架波的证明过程为设计新的卷积神经网络提供了思路。基于拼接非线性单元，能够提升网络的抗噪性能，优化图像重建。

基于深度学习的完美重构网络的研究在近来也在逐渐兴起。为了减少深度神经网络训练时所需的内存开销，Gomez 等（Gomez 等，2017）受到传统信号处理中提升算法的启发，提出了可逆残差网络（Reversible Residual Network）。由于该网络中每一层的特征图可以由反向传播过程的上一层计算而来，解决了反向传播算法需要记录每一层特征图的问题，从而减少了内存开销。而这种方法引起了 Jacobsen 等的注意而被引入了神经网络的解释体系中（Jacobsen 等，2018）。具体地，由信息瓶颈理论出发的神经网络解释方法说明深度网络在训练的过程中逐渐丢弃了与标签信息无关的信息，而可逆网络在分类任务上的优异表现则与之相异。因此，基于信息瓶颈理论的深度神经网络解释方法仍然需要进一步地补充完善。

1.3 卷积稀疏编码模型

传统基于稀疏编码假设信号表示可以基于一个全局不变的字典，对于所有局部采样进行稀疏表示。然而，与当前广泛应用的深度卷积神经网络相比，该假设无法充分利用信号在不同尺度上的局部性特征。随着信号局部化特征提取问题逐渐受到关注，多层卷积稀疏编码被提出来融合传统稀疏表示理论和深度学习以及深度卷积神经网络。

卷积稀疏编码（Convolutional Sparse Coding, CSC）利用字典所衍生的一系列滤波器，通过卷积操作获得稀疏表示，在 2010 年首先由美国纽约大学 Zeiler 等提出（Zeiler 等，2010）。卷积稀疏编码通过卷积字典操作捕捉信号的可变局部特征，生成移不变（shift invariant）字典，表示局部关联性，捕捉局部化特征，并由 ℓ_0/ℓ_∞ 稀疏项引入非线性最大化操作（Pappyan 等，2017）。然而，不同于卷积神经网络的不可解释性（“黑盒”结构）和数学模型的不可分析性（无法分析最优解），卷积稀疏编码继承了稀疏编码的特点，能够通过凸优化理论进行求解和分析，并适用于无监督学习。为了适应多尺度信号表示，Sulam 等拓展了卷积稀疏编码，通过递归迭代建立多层字典学习和编码结构，并通过基于追踪的算法求解（Sulam 等，2018）。随后，Pappyan 等为基于 ℓ_0/ℓ_∞ 稀疏性的非凸优化建立理论误差上下界，保证了卷

表 2 卷积稀疏编码模型和优化方法

Table 2 Convolutional sparse coding models and optimization methods

相关工作	稀疏性模型	优化方法	
Papayan 等人 (2017)	ℓ_1 约束问题	Toeplitz 展开	匹配追踪 正交匹配追踪
Papayan 等人 (2017)	ℓ_0 约束问题		迭代阈值
Sulam 等人 (2018)			
Plaut 等人 (2018)			
Chalasani 等人 (2013)	ℓ_1 约束问题	傅里叶变换	
Silva 等人 (2018)			
Sreter 等人 (2018)			
Lin 等人 (2017)			
Bristow 等人 (2013)	ℓ_0 约束问题		
Wohlberg (2014)			
Heide 等人 (2015)			
Peng (2019)	ℓ_0 约束问题		
Wang 等人 (2018)	ℓ_1 约束问题		
Wang 等人 (2018)			
Chun 等人 (2018)			

积稀疏编码在普适性信号表示应用的性能边界（Pappyan 等，2018）。

1.3.1 快速傅里叶变换（FFT）

与稀疏编码有所不同，卷积稀疏编码将矩阵乘法变为矩阵卷积，估计的信号由多个卷积生成的子信号求和得到，由于直接卷积大大提升了空间复杂度和时间复杂度，对卷积稀疏编码的实现多是利用快速傅里叶变换将原问题转化到频域，再利用交替方向乘法（Alternating Direction Method of Multipliers, ADMM）（Boyd, 2010）迭代求解。Bristow 等对于 ℓ_1 约束的卷积稀疏编码问题，采用 ADMM 进行四重分解，进行快速求解（Bristow 等，2013）。Wohlberg 在 ADMM 优化框架中，利用快速傅里叶变换（FFT），从而达成与滤波器数量成线性关系的计算复杂度（Wohlberg, 2014）。Heide 等将优化目标函数高效分解为凸函数的组合，并能够利用合适的边界条件获得更优解（Heide 等，2015）。Chun 和 Fessler 基于控制不等式提出块近端梯度下降算法，来取代 ADMM 达成快速收敛（Chun 等，2018）。Choudhury 等考虑降低采用 ADMM 求解卷积稀疏编码的内存消耗，分解优化问题，提出基于协商一致的优化方法（Consensus Optimization），用于大规模

高维信号 (Choudhury 等, 2017)。

1.3.2 Toeplitz 展开

另一类方法则仍然在空域上, 拓展经典稀疏表示的求解算法。主要考虑到虽然利用快速傅里叶变换显著降低了卷积稀疏编码问题的复杂度, 但这类算法的复杂度提升上限被快速傅里叶变换及其逆变换所决定。利用 Toeplitz 矩阵变换, 将矩阵卷积转变为矩阵乘法, 成为理解和实现卷积稀疏编码提供了一种新的思路 (Papayan 等, 2017)。基于 Toeplitz 变换生成的全局字典存在极大的结构冗余, 因此该方向的研究者采用基于局部字典的优化方法实现卷积稀疏编码, 并在计算复杂度上有望突破快速傅里叶变换的限制 (Papayan 等, 2017)。最近一系列工作探讨了利用匹配追踪 (Plaut 等, 2018; Sulam 等, 2018) 和迭代阈值算法 (Chalasani 等, 2013; Silva 等, 2018; Sreter 等, 2018) 求解的可能性。并且可通过在线学习, 在频域采用迭代阈值算法大幅提高优化问题求解效率 (Liu 等, 2017)。由于 ℓ_0 范数稀疏约束的卷积稀疏编码问题是 NP 问题, 且在迭代过程中使用贪婪法不能保证收敛的性质, 大多数先前的研究都是引入 ℓ_1 范数保证迭代算法最终收敛。但 ℓ_1 范数的惩罚函数在字典学习过程中会在较大系数上产生偏移和发生过度惩罚 (Bao 等, 2016)。

1.3.3 应用

多层卷积稀疏编码有望成为理解和探究卷积神经网络“黑盒子”的一种手段。研究多层卷积稀疏编码各个层上的尺度缩放和误差传递过程, 对于人们理解卷积神经网络具有非常重要的意义。目前已应用于多个信号处理实例中。如 Zhu 和 Lucey 利用卷积稀疏编码, 从自然发生的点轨迹集中训练过完备字典, 使其满足稀疏重构的有限等距性, 从非刚性物体的二维映射轨迹重构各个点的三维轨迹 (Zhu 等, 2015)。在钢琴曲谱转录誊写中, 基于短时傅里叶变换, 利用卷积稀疏编码构建时域转录誊写算法, 并在时域基于音符波形和起止时间脉冲的卷积字典学习, 提升转录誊写准确率和时间精度 (Cogliati 等, 2015)。Zhang 和 Patel 利用卷积稀疏编码, 直接基于整幅图片进行分解 (轮廓和纹理), 不需要根据重叠块学习局部字典 (Zhang 等, 2016), 该方法也可应用于图像前景和背景分解, 如雨痕清除 (Zhang 等, 2017)。

近年来, 鉴于生物医学影像的对于假象和噪声的稳定性要求, 卷积稀疏编码模型逐渐受到关注。Zhou 等基于卷积稀疏编码进行多光谱特征学习, 训

练一系列卷积滤波器组表示不同生物对象的光谱, 探索组织形态学特征 (Zhou 等, 2014)。Chang 等将卷积稀疏编码与迁移学习结合, 应用于无监督的细胞组织分类 (Chang 等, 2018)。Jas 等基于重尾分布提出概率卷积稀疏编码, 利用蒙特卡罗期望最大化 (EM) 算法, 从神经时间序列数据中推断移不变原型波形 (原子), 并且对于人为假象和噪声干扰稳定 (Jas 等, 2017)。在动态对比增强核磁共振成像的重构中, Quan 和 Jeong 基于卷积稀疏编码建立基于能量的重构优化问题, 利用傅里叶卷积定理加速滤波器组学习, 并通过 ADMM 求解 (Quan 等, 2016)。Serrano 等在高动态范围成像中应用卷积稀疏编码 (Serrano 等, 2016)。对核磁共振成像问题, Huang 等提出弱监督联合卷积稀疏编码方法同时解决超分辨率和跨模态图像合成 (Huang 等, 2017)。

1.4 图卷积神经网络

随着数据采集设备的发展, 信号的形式不再局限于低维规则网格结构 (如语音、图片和视频等), 表现出复杂不规则的拓扑结构 (社交网络、蛋白质结构以及点云)。传统的基于傅立叶变换和小波变换的信号处理理论针对的是定义在欧式空间中低维规则网格结构上的信号, 因此无法直接适用于高维不规则信号。如何表示、分析和处理这些高维不规则信号成为当下亟待解决的难题。图作为一种灵活的数据表示形式, 能够很好的刻画高维信号复杂的拓扑关系, 成为理想的高维非规则信号的表示形式。图信号处理旨在建立在非欧空间上进行基于信号内在拓扑结构的信号处理理论。目前的图信号处理理论可以分为两个理论体系: 一是基于谐波分析和谱图理论 (Chung, 1997), 将传统信号处理理论扩展到图域形成的谱图信号处理理论 (Hammond 等, 2011; Shuman 等, 2013)。其以对称半正定的图拉普拉斯矩阵的完备正交基作为傅立叶基定义图傅立叶变换, 并在此基础上形成谱图小波变换理论 (Hammond 等, 2011)、图信号采样理论 (Chen 等, 2015; Tsitsvero 等, 2016) 以及图滤波器组理论 (Narang 等, 2012; Tanaka 等, 2014; Teke 等, 2017)。为了获得对称正定的拉普拉斯矩阵, 信号必须定义在具有非负权重的无向图上; 二是基于代数图论的建立的代数图信号处理理论 (Sandryhaila 等, 2013)。代数图信号处理理论将图邻接矩阵定义为图平移算子, 并在此基础上设计平移不变图滤波器构建信号处理理论体系。相较于谱图信号处理, 代数信号处理理论可以扩展到有向图。以上两者构成了图信号

表 3 谱图卷积神经网络

Table 3 Spectral graph convolutional neural networks

网络模型	模型特性
Niepert 等人 (2016)	参数化图滤波器构造可学习图卷积操作。对图拉普拉斯矩阵行特征分解, 计算和存储复杂度高。同时滤波器无权值共享机制且不具备局部性质。
Bruna 等人 (2013)	利用空域频域对偶特性, 学习平滑乘子构造具备一定局部定位特性的图滤波器。参数大幅减少, 降低过拟合风险, 计算过程不需要特征分解。只适用于小规模图。
Defferrard 等人 (2016)	切比雪夫多项式逼近构造参数化图卷积滤波器。图滤波器具备严格局部特性, 计算和存储复杂度大幅降低。多项式函数拟合, 滤波器逼近性能受限且过于平滑。
Levie 等人 (2019)	使用 Cayley 多项式为基构造可学习图卷积滤波器, 可以很好的对窄带信号进行定位和滤波。滤波器不具备严格局部定位特性, 计算复杂度高。
Bianchi 等 (2019)	设计自回归滑动平均图滤波器, 利用有理函数更精确地逼近任意图滤波器, 使用递归方式对图信号进行分布式处理。具备更广泛的滤波器表示空间, 能够建模任意转换函数。
Xu 等人 (2019)	使用图小波基代替图傅立叶基设计图卷积操作子形成深度图小波网络。图小波变换具备高稀疏性、严格局部定位特性、适用快速算法。
Zou 等人 (2018)	基于散射变换构造图卷积神经网络, 具有能量守恒性、排列不变性和图调制稳定性

处理的基本理论基础。

鉴于谱图信号处理是传统信号处理理论在图域上的扩展, 本节主要介绍谱图卷积神经网络的最新进展。谱图信号处理适用于所有定义在具有非负权重的无向图上的信号。图傅立叶变换以对称正定的图拉普拉斯矩阵的完备正交基作为傅立叶基, 以特征值作为图频谱 (Shuman 等, 2013)。其中具有较小特征值的特征向量表征在图上缓慢变化的低频图信号, 而大特征值对应的特征向量表征高频图信号。基于图傅立叶变换, 可将传统信号处理理论中的平移、卷积和调制等概念被扩展到图信号域, 形成图滤波器设计的基本理论 (Hammond 等, 2011; Shuman 等, 2013)。考虑图傅立叶变换是一种全局变换, 其对应的基包含了所有节点, 可以考虑扩展传统信号处理中的短时傅里叶变换和小波变化, 提升局部特性 (Hammond 等, 2011; Shuman 等, 2016)。并可进一步考虑图信号的多尺度奇异性, Coifman 和 Maggioni 根据扩散模型利用不同尺度的扩散算子的压缩表示来捕捉图信号在不同分辨率下的奇异性 (Coifman 等, 2006)。Shuman 等进一步提出了自适应于图频谱分布的紧框架的小波变换 (Shuman 等, 2015)。由此, 发展出基于图下采样、图缩减、图滤波和图上采样框架的多尺度图拉普拉斯变换 (Shuman 等, 2016)。为了使得信号变换适应于特定的图信号, 获得对于信号最优的变换, 扩散小波包 (Bremer 等, 2006)、基于深度学习的图小波变换 (Rustamov 等, 2013) 以及基于树结构的小波变换

(Ram 等, 2011) 被相继提出。

随着深度学习的发展, 卷积神经网络强大的建模能力以及数据处理能力得到了检验和发展, 将卷积神经网络向非欧式空间扩展成为新的研究热点。图卷积神经网络由此被提出用于数据驱动的高维不规则信号的分析 and 处理 (Wu 等, 2019)。2013 年 Brunu 等首先在谱图信号处理理论上通过将图滤波器参数化构造图卷积操作形成深度图卷积网络, 从而使得卷积神经网络可以应用于处理非规则高维信号 (Brunu 等, 2013)。但是由于滤波器学习过程中需要对拉普拉斯算子进行的特征分解, 具有很高的计算和存储复杂度, 限制了该网络的可扩展性仅适用于处理小规模的数据并且该滤波器不具备局部特性, 其感受野覆盖全图, 另外其滤波器的学习依赖于先验的图结构信息。在此基础上, Henaff 等尝试将深度学习应用于不具备图结构先验知识的高维数据上, 将图结构的学习过程融入卷积神经网络中, 利用空域和频域的对偶性质, 通过在图频域内学习以平滑乘子为基的平滑函数获得具备一定局部特性的图卷积算子同时减少卷积滤波器参数量降低模型过拟合风险 (Henaff 等, 2015)。Defferrard 等以切比雪夫多项式逼近图滤波器获得具备严格局部特性的可以快速计算的图卷积操作, 并且利用最大割集算法构造多分辨率图结构进行多分辨率图卷积神经网络学习 (Defferrard 等, 2016)。多项式参数化的图滤波为有限脉冲响应滤波器, 其对于图信号和图结构的变换非常敏感, 并且由于平滑性不能在图

频域内对窄带信号进行建模。为了克服上述问题, CayleyNet 以 Cayley 多项式为基础构造可学习图卷积滤波器, 可以很好的对窄带信号进行定位和滤波 (Levie 等, 2019)。Bianchi 等提出可学习的自回归滑动平均图滤波器, 利用有理函数对任意图滤波器进行逼近, 使用递归方式对图信号进行分布式处理 (Bianchi 等, 2019)。相较于多项式图滤波器, 其具备更加广泛的滤波器表示空间, 能够建模任意转换函数, 可以更加精确地逼近任意图滤波器。由于其滤波器构造不依赖于拉普拉斯矩阵的特征向量和特征值, 对于图信号和图结构的变化更加鲁棒, 可以扩展到空时图信号处理中。最近, Zou 和 Lerman 基于冗余的多尺度 Hammond 图小波 (Hammond 等, 2011), 将散射变换扩展到图卷积神经网络中, 并证明其能量守恒性、排列不变性和图调制稳定性 (Zou 等, 2018)。

2 国内研究进展

2.1 结构化稀疏表示模型

基于结构化稀疏表示方法, 国内的研究机构取得了一系列相关的科研成果。上海交通大学的 Li 等将组稀疏应用于视频压缩感知, 自适应地将信号分解为数据驱动的联合子空间 (Union of Data-driven Subspace, UoDS), 利用线性与多线性子空间学习法 (Linear and Multilinear Subspace Learning) 推演出最优的结构化稀疏基, 显著提升了压缩感知的鲁棒性与重构质量 (Li 等, 2017)。北京大学的研究人员提出基于低秩表示 (Low-Rank Representation, LRR) 的子空间结构的鲁棒重建, 在给定的字典空间中, 寻求一组基于数据样本的最低秩表示, 利用凸规划, 精确恢复出真实的子空间结构 (Liu 等, 2013); 在此基础上, 进一步提出利用分块对角表示 (Block Diagonal Representation, BDR) 诱导的正则项来通用地解决子空间聚类问题 (Lu 等, 2019)。西安电子科技大学的 Dong 等利用基于组稀疏的同步稀疏编码 (Simultaneous sparse coding, SSC), 提出高斯尺度混合 (Gaussian scale mixture, GSM) 模型的非局部 (Non-local) 扩展, 应用于自然图像复原任务上 (Dong 等, 2015)。同时, 他们还利用非负结构化稀疏编码进行高光谱图像的超分辨率 (Hyperspectral Image Super-Resolution), 基于空域-谱域稀疏性先验知识, 联合估计高光谱字典与稀疏编码 (Dong 等, 2016)。北京大学的 Lu 等设计 ℓ_1 范式的超图正则项 (Hypergraph Regularization) 引入结构化稀疏性,

使得所构建的图表示与谱嵌入 (Spectral Embedding) 能够利用流形结构 (Manifold Structure), 抽取中级特征的鲁棒相关性, 对于基于隐式语义学习 (Latent Semantic Learning) 的人体动作识别具有明显的性能改进 (Lu 等, 2013)。同时, Zhou 等将结构化稀疏应用于 ℓ_1 -图构建 (ℓ_1 -Graph Construction), 提出了 k 近邻融合 Lasso 图 (K-Nearest Neighbor Fused Lasso Graph), 应用与复杂的大尺寸图像数据集的谱聚类 (Spectral Clustering) 问题 (Zhou 等, 2013)。中科院自动化所的 Peng 等针对显著性目标检测问题 (Salient Object Detection) 提出结构化矩阵分解模型 (Structured Matrix Decomposition), 分别利用两种结构化稀疏正则项, 树状稀疏性诱导的正则项用于捕捉图像结构并约束来自相同目标的图像块具有相似的显著性, 拉普拉斯正则项用于扩大显著性目标与特征空间背景的分离程度 (Peng 等, 2017)。华南理工大学的 Wen 等将有遮挡的人脸识别问题 (Face Recognition with Occlusion) 建模为在附加遮挡的字典 (Occlusion-appended Dictionary) 上求解结构化稀疏表示的问题, 通过替换 ℓ_1 -范式稀疏性为结构化稀疏性, 增强了人脸识别的鲁棒性 (Wen 等, 2016)。

2.2 基于框架理论的深度网络模型

在基于框架理论的深度网络模型方面, 国内研究基于散射网络的想法做了模型扩展和应用。在应用问题中, 主要将其作为特征提取子引入图像处理应用。例如, 复旦大学研究人员在合成孔径雷达的目标识别中, 对于不同空间和角度变量, 引入平移和旋转不变的复小波滤波器的深度散射网络, 进行特征提取和表示, 并利用散射网络的无监督特性避免训练, 提升效率 (Wang 等, 2018)。上海交通大学研究人员将散射网络与卷积神经网络相结合, 构造端到端的混合小波卷积网络, 用于稀疏编码和图像超分辨率任务 (Gao 等, 2016)。通过堆叠卷积, 混合小波卷积网络能够同时利用散射网络根据频域性质选取稀疏编码候选原子, 并利用卷积神经网络获得稀疏编码。

而在理论模型方法的研究, 清华大学研究人员受到散射网络的启发, 将网络的尺度不变特性引入卷积神经网络中, 结合多尺度特征提取和识别 (Xu 等, 2014)。按照多列架构, 使得每一列对应于某个特定尺度, 并通过尺度变换使得各列共用相同的滤波器参数集, 从而在不增加模型参数量的情况下处理尺度变化。东南大学研究人员则从卷积核本身的

特性出发,提出了分数阶小波的散射网络理论模型及基础应用,提供广义的平移不变性(Liu等,2019)。通过迭代计算分数小波变换和取模操作,构建散射网络,并利用不同分数阶数对信号进行分析,提升医学图像识别和分割的性能。而上海交通大学研究人员则从信号分解重构的角度出发,通过设计可逆变换,添加完美重构条件,利用完美重构滤波器组(Perfect Reconstruction Filter Bank)对应的滤波器作为卷积网络的卷积核,构造具有完美重构特性的卷积网络(Xu等,2017,2016)。利用伪极坐标傅里叶变换(Pseudo Polar Fourier Transform)对图像的傅里叶域进行变换,实现分离的方向和尺度分解。在方向和尺度上,引入一维完美重构滤波器组作为卷积核,利用下采样矩阵实现临界采样,递归叠加构造深度卷积网络。深度网络中,分解尺度和方向的数目可由卷积和下采样层数来控制,从而获得原始信号在任意方向和尺度上的分解子带。合成网络可直接基于与分解滤波器组(Analysis Filter Bank)对称的合成滤波器组(Synthesis Filter Bank)构建,实现完美重构。

2.3 卷积稀疏编码模型

对于基于变换的卷积稀疏编码求解问题,尤其是基于非可微的 ℓ_0 稀疏性约束问题,国内研究者在2018和2019年分别提出算法,旨在改进ADMM来提升收敛速度。Wang等利用多层卷积字典学习建立可分级信号表示,并通过在线学习更新多层字典,提高收敛速度(Wang等,2018,2018)。Peng则在基于快速傅里叶变换的求解方法中,自适应地更新ADMM算法的正则系数,提高ADMM算法的收敛速度(Peng,2019)。

国内研究机构将卷积稀疏编码应用于图像超分辨率和图像识别等应用,取得了一系列成果。在合成孔径雷达目标配置识别中,陕西师范大学研究人员将训练采样嵌入稀疏表示模型,按照基于高斯混合分布的统计模型,逐一构建字典去除方位角敏感性带来的干扰,提升识别准确率(Liu等,2017)。华南理工大学研究人员在图像识别问题中应用卷积稀疏编码,通过训练卷积滤波器组决定分类标签,并结合稀疏表示分类方法,获得更多表示信息,提升分类性能(Chen等,2016)。西安交通大学研究人员在视频雨痕去除问题中应用多尺度卷积稀疏编码,通过训练卷积滤波器捕捉不同位置散射雨痕的局部重复稀疏模式和相距摄像机不同距离位置的多尺度特性,进行雨痕去除(Li等,2018)。

在基于图像重建的问题中,卷积稀疏编码被用于解决传统稀疏表示问题中基于重叠块重构的像素一致性问题。图像超分辨率应用中,香港科技大学、哈尔滨工业大学等机构研究人员,考虑基于重叠块的图像重构时像素一致性的问题,通过卷积稀疏编码基于高/低分辨率图像的特征图映射进行图像重构,从而避免将图像划分为重叠块产生像素不一致的问题(Gu等,2015)。四川大学等研究机构人员,在计算机断层扫描(CT)的稀疏视角重建问题中,引入卷积稀疏编码作为重构优化问题的目标函数正则项,实现对于整个图像的重构,从而避免基于重叠块方法所导致的像素一致性问题(Bao等,2019)。

2.4 图卷积神经网络

国内对于图信号处理理论的研究较少,随着图卷积神经网络成为深度学习领域的热点,国内的高校及研究机构也开展了相应的研究工作,主要方向集中于图卷积神经网络的设计和相关应用。在网络模型结构方面,清华大学的研究人员在2018年提出基于控制变量的图卷积网络,有效减少感受野的大小(Chen等,2017);腾讯AI Lab提出自适应图卷积神经网络(Li等,2018),可接受任意结构和规模的图作为输入;最近,中国科学院大学的相关研究人员提出图小波网络,使用图小波基代替图傅立叶基设计图卷积运算形成深度图小波网络(Xu等,2019)。图小波变换相对于图傅立叶变换具有稀疏、局部以及快速计算的优势,在半监督学习任务中取得了很好的表现,同时也使得图信号处理和图网络的结合成为一个具有前景的研究方向。

在应用方面,清华大学Dai等关注图神经网络的鲁棒性(Dai等,2018),通过对抗训练的方法来增强图卷积神经网络滤波器的分类稳定性。香港中文大学研究人员提出了一种时空图卷积网络,并利用它们进行人类行为识别(Yan等,2018)。这种算法基于人类关节位置的时间序列表示而对动态骨骼建模,并将图卷积扩展为时空图卷积网络而捕捉这种时空的变化关系。此外,清华大学的两组研究人员分别针对此领域发表了综述文章(Wu等,2019),介绍了图卷积神经网络发展和相关理论。

3 国内外研究进展比较

结构化稀疏(组稀疏)表示模型最早由美国佐治亚理工学院和威斯康辛大学学者2006年提出(Yuan等,2006),至今,共发表大约7000多篇学术文章。以发现实际问题中的结构信息,建立稀疏

性诱导的正则项的数学模型,设计优化算法求解正组化问题,以及扩展表示模型至更广泛的应用问题为主要研究脉络,结构化稀疏表示模型在国内外的众多研究领域取得成功。国外的研究人员在观点、模型与算法这三个维度上做出了许多贡献,如组稀疏(Yuan等,2006)、分层稀疏(Zhao等,2009)与图稀疏(Meinshausen等,2006)模型的提出、交替方向乘子法(Boyd等,2010)、近端算法(Jenatton等,2010)、网络流优化(Mairal等,2011)等高效算法的设计。2011年,机器学习顶级期刊JMLR(Journal of Machine Learning Research)上,法国INRIA研究院的F. Bach等共发表了三篇结构化稀疏表示的经典工作(Jenatton等,2009, 2010; Mairal等,2011),并在2012年从数学优化的角度出发,发表结构化稀疏的系统专题论文,涵盖不同的优化算法的对比与充分的理论基础分析(Bach等,2011)。国内对结构化稀疏表示的研究约在2012年兴起,由于起步较晚,此时理论发展已较为完善,且随着大数据时代对计算速度的需求上升,国内学者的研究内容则更加侧重于更为有效的优化算法设计,以及在不同场景下的应用拓展。比如,上海交通大学设计的分块坐标下降法(Wen等,2012),北京大学设计的解决低秩优化问题的加速梯度法(Li等,2019)。在应用方面,上海交通大学研究人员对视频压缩感知应用的拓展(Li等,2017),西安电子科技大学研究人员基于自然图像复原与高光谱图像的超分辨率的工作等(Dong等,2016, 2015),均取得了显著的性能提升。

基于框架理论的深度网络模型和卷积稀疏编码模型的研究方兴未艾,其理论均由国外学者最先于2012年和2010年提出,并逐渐受到关注。其中,基于框架理论的深度网络模型方面工作主要基于法国巴黎高师的Mallat和美国纽约大学的Bruna所提出的散射网络(Bruna等,2013),并且在此基础上就其旋转不变性、谱分析等问题进行了系统性研究。随后,苏黎世高工Wiatowski等和韩国科学技术院Ye等将其扩展到半离散框架的深度网络。国内包括清华大学、上海交通大学及东南大学的研究人员主要就散射网络与卷积神经网络的尺度联系、完美重构和广义平移不变性等性质,做了进一步研究,并将散射网络作为一种无监督高效特征提取子应用到特征提取、图像识别和分割等应用中。卷积稀疏编码最早由纽约大学Zeiler等提出,受到散射网络的启发,随后被用于卷积神经网络的解释和拟合。以

色列理工学院的研究人员就卷积稀疏编码的空域优化方法,尤其是基于追踪算法的优化求解及其逼近误差理论上界做了系统性的研究。与此同时,美国洛斯阿拉莫斯国家实验室Wohlberg和斯坦福大学Heide等研究人员开始从事基于交替方向乘子法的变换域优化方法,并受到国内研究人员关注,提出了提高收敛速率的集合在线学习的优化方法,以及基于 ℓ_0 稀疏约束的卷积稀疏编码方法。鉴于卷积稀疏编码的平移不变性和全局字典学习,在图像和视频的应用受到国内外研究人员的一致关注,并取得一系列研究成果。

图信号处理理论最早由斯坦福和瑞士联邦理工相关研究人员提出并发展,在此基础上与卷积网络结合,发展出不同的图卷积神经网络理论模型。目前的图卷积网络的相关模型设计以及理论均由国外学者提出并发展,国内学者主要致力于图卷积神经网络在非规则数据上的应用,如点云、交通数据以及社交网络等,对理论方面的研究较为缺乏。随着图小波网络的提出和发展,将图信号处理的相关理论应用于图卷积神经网络,形成具备严谨理论支撑的图卷积神经网络体系引起了国内外研究人员的广泛的研究兴趣。

4 发展趋势与展望

现有的结构化稀疏表示方法只能对传统的在欧式空间中均匀分布的信号(如语音、图像和视频)进行优化表示,而随着大规模3D传感设备与采集系统的迅猛发展,虚拟/增强现实(VR/AR)、自动驾驶、智慧城市等新型技术的兴起,对点云与全景视频等非规则立体大数据的需求激增。这类大规模非规则数据的结构往往天然具备高度稀疏性,非欧式信号空间的拓扑结构奇异性,以及多维数据间存在的抽象依赖关系。结构化稀疏表示正是一种有力工具,能够通过混合范式约束优化目标,利用交替优化的思想,自适应地表示信号尺度间的频率演变关系,有效编码信号之间潜在的结构先验信息,从而进行渐进最优的稀疏逼近。因此,如何将结构化稀疏方法扩展至更加复杂的、在空间上完全不均匀分布的非规则信号,同时动态捕捉非规则信号内在的结构化信息与多尺度特征,实现这类信号的高效表示、编码与传输,是一个潜在的机遇与挑战。

图作为一种灵活的数据表示形式,能够很好地表示数据之间的关系,表征拓扑信息,在社交网络分析、互联网、生物医学、分子物理和化学以及金

融、高维信号处理领域具有重要的价值和意义,也会成为推动社会和产业发 展的下一个助力点。同时,图信号处理将传统信号处理理论体系向高维非欧空间信号的进行扩展,建立一整套完善的图信号处理理论,为图信号在各个领域的应用和发展奠定理论基础。图信号处理以及图卷积神经网络的未来发展方向将重点在以下几个方面:

(1) 动态图分析和处理。由于图信号处理体系建立在图傅立叶基的基础上,因此不同的图所建立的傅立叶基是不同的,因此对于动态图的分析 和处理成为图信号处理体系建立的必须解决的问题。

(2) 传统的信号处理工具在图上的扩展如图滤波器组和相关快速算法的理论建立和实现。

(3) 基于数据驱动的图信号处理方法的建立。随着计算力和数据的增长,基于数据驱动的信号处理如图字典和卷积稀疏码等表现了很好的性能和泛化性能,成为研究的热点。

(4) 基于图信号处理理论的图卷积神经网络的发展和理论解释。包括图卷积、下采样、上采样和模型稳定性和泛化性的理论保证。

(5) 基于图卷积神经网络在相关领域的应用。如点云、蛋白质和化合物分析、社交网络分析。

(6) 基于图卷积神经网络的图表示学习方法。

参考文献(References)

Aharon M, Elad M and Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006. [DOI:10.1109/TSP.2006.881199]

Bach F. Consistency of the group Lasso and multiple kernel learning[J]. Journal of Machine Learning Research, 2007, 9(6): 1179–1225. [http://arxiv.org/abs/0707.3390]

Bach F. Exploring large feature spaces with hierarchical multiple kernel learning[C]//Advances in Neural Information Processing Systems 21. Vancouver, BC, Canada: Curran Associates, Inc., 2008: 105–112. [https://papers.nips.cc/paper/3418-exploring-large-feature-spaces-with-hierarchical-multiple-kernel-learning.pdf]

Bach F, Jenatton R, Mairal J and Obozinski G. Optimization with sparsity-inducing penalties[J]. Foundations and Trends in Machine Learning in Machine Learning, 2011, 4(1): 1–106. [DOI:10.1561/22000000015]

Bach F, Jenatton R, Mairal J and Obozinski G. Structured Sparsity through Convex Optimization[J]. Statistical Science, 2012, 27(4): 450–468. [DOI:10.1214/12-STS394]

Bao C, Ji H, Quan Y and Shen Z. Dictionary learning for sparse coding: Algorithms and convergence analysis[J]. IEEE

Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(7): 1356–1369. [DOI:10.1109/TPAMI.2015.2487966]

Bao P, Xia W, Yang K, Chen W, Chen M, Xi Y, Niu S, Zhou J, Zhang H, Sun H, Wang Z and Zhang Y. Convolutional Sparse Coding for Compressed Sensing CT Reconstruction[J]. IEEE Transactions on Medical Imaging, 2019: 1–1. [DOI:10.1109/tmi.2019.2906853]

Beck A and Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems[J]. SIAM Journal on Imaging Sciences, 2009, 2(1): 183–202. [DOI:10.1137/080716542]

Bengio S, Pereira F and Strelow D. Group sparse coding[C]//Advances in Neural Information Processing Systems 22. Vancouver, BC, Canada: Curran Associates, Inc., 2009: 82–89. [https://papers.nips.cc/paper/3691-group-sparse-coding.pdf]

Bianchi F M, Grattarola L, Aiuppi C and Livi L. Graph neural networks with convolutional ARMA filters[J]. arXiv preprint, 2019. [http://arxiv.org/abs/1901.01343]

Blumensath T and Davies M E. Iterative thresholding for sparse approximations[J]. Journal of Fourier Analysis and Applications, 2008, 14(5–6): 629–654. [DOI:10.1007/s00041-008-9035-z]

Boyd S. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. Foundations and Trends in Machine Learning, 2010, 3(1): 1–122. [DOI:10.1561/22000000016]

Bremer J C, Coifman R R, Maggioni M and Szlam A D. Diffusion wavelet packets[J]. Applied and Computational Harmonic Analysis, 2006, 21(1): 95–112. [DOI:10.1016/j.acha.2006.04.005]

Bristow H, Eriksson A and Lucey S. Fast convolutional sparse coding[C]//2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA: IEEE, 2013: 391–398. [DOI:10.1109/CVPR.2013.57]

Bruna J and Mallat S. Invariant scattering convolution networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1872–1886. [DOI:10.1109/TPAMI.2012.230]

Bruna J, Zaremba W, Szlam A and LeCun Y. Spectral networks and locally connected networks on graphs[J]. arXiv preprint, 2013. [http://arxiv.org/abs/1312.6203]

Candès E J. Ridgelets: Theory and Applications[D]. Stanford University, 1998. [http://statweb.stanford.edu/~candes/papers/thesis.ps]

Candès E J and Donoho D L. Curvelets: A surprisingly effective nonadaptive representation of objects with edges[J]. Curves and Surface Fitting, 2000, C(2): 1–10. [DOI:10.1016/j.biopsycho.2009.12.002]

Chalasanani R, Principe J C and Ramakrishnan N. A fast proximal method for convolutional sparse coding[C]//The 2013 International Joint Conference on Neural Networks (IJCNN). Dallas, TX, USA: IEEE, 2013.

- [DOI:10.1109/IJCNN.2013.6706854]
- Chang H, Han J, Zhong C, Snijders A M and Mao J H. Unsupervised transfer learning via multi-scale convolutional sparse coding for biomedical applications[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(5): 1182–1194. [DOI:10.1109/TPAMI.2017.2656884]
- Chen B, Li J, Ma B and Wei G. Convolutional sparse coding classification model for image classification[C]//2016 International Conference on Image Processing (ICIP). Phoenix, AZ, USA: IEEE, 2016: 1918–1922. [DOI:10.1109/ICIP.2016.7532692]
- Chen J, Zhu J and Song L. Stochastic training of graph convolutional networks with variance reduction[J]. arXiv preprint, 2017
- Chen S S, Donoho D L and Saunders M A. Atomic decomposition by basis pursuit[J]. SIAM Journal on Scientific Computing, 1998, 20(1): 33–61. [DOI:10.1137/S1064827596304010]
- Chen S, Varma R, Sandryhaila A and Kovačević J. Discrete signal processing on graphs: Sampling theory[J]. IEEE Transactions on Signal Processing, 2015, 63(24): 6510–6523. [DOI:10.1109/TSP.2015.2469645]
- Choudhury B, Heide F, Wetzstein G and Heidrich W. Consensus convolutional sparse coding[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 4280–4288. [DOI:10.1109/ICCV.2017.459]
- Chun I Y and Fessler J A. Convolutional dictionary learning: Acceleration and convergence[J]. IEEE Transactions on Image Processing, 2018, 27(4): 1697–1712. [DOI:10.1109/TIP.2017.2761545]
- Chung F. Spectral Graph Theory[M]. American Mathematical Society, 1997. [DOI:10.1145/568547.568553]
- Cogliati A, Duan Z and Wohlberg B. Piano music transcription with fast convolutional sparse coding[C]//2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP). Boston, MA, USA: IEEE, 2015. [DOI:10.1109/MLSP.2015.7324332]
- Coifman R R and Maggioni M. Diffusion wavelets[J]. Applied and Computational Harmonic Analysis, 2006, 21(1): 53–94. [DOI:10.1016/j.acha.2006.04.004]
- Dai H, Li H, Tian T, Huang X, Wang L, Zhu J and Song L. Adversarial attack on graph structured data[J]. arXiv preprint, 2018
- Defferrard M, Bresson X and Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering[C]//Advances in Neural Information Processing Systems 29. Barcelona, Spain: Curran Associates, Inc., 2016: 3844–3852. [https://papers.nips.cc/paper/6081-convolutional-neural-networks-on-graphs-with-fast-localized-spectral-filtering.pdf]
- Do M N and Vetterli M. The contourlet transform: An efficient directional multiresolution image representation[J]. IEEE Transactions on Image Processing, 2005, 14(12): 2091–2106. [DOI:10.1109/TIP.2005.859376]
- Dong W, Fu F, Shi G, Cao X, Wu J, Li G and Li X. Hyperspectral image super-resolution via non-negative structured sparse representation[J]. IEEE Transactions on Image Processing, 2016, 25(5): 2337–2352. [DOI:10.1109/TIP.2016.2542350]
- Dong W, Shi G, Ma Y and Li X. Image restoration via simultaneous sparse coding: Where structured sparsity meets gaussian scale mixture[J]. International Journal of Computer Vision, 2015, 114(2–3): 217–232. [DOI:10.1007/s11263-015-0808-y]
- Donoho D L. Wedgelets: Nearly minimax estimation of edges[J]. Annals of Statistics, 1999, 27(3): 859–897. [DOI:10.1214/aos/1018031261]
- Efron B, Hastie T, Johnstone I and Tibshirani R. Least angle regression[J]. Annals of Statistics, 2004, 32(2): 494–499. [DOI:10.1214/009052604000000067]
- Engan K, Aase S O and Husoy J H. Method of optimal directions for frame design[C]//Proceedings. 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Phoenix, AZ, USA: IEEE, 1999: 2443–2446. [DOI:10.1109/ICASSP.1999.760624]
- Friedman J, Hastie T and Tibshirani R. A note on the group Lasso and a sparse group Lasso[J]. arXiv preprint, 2010. [http://arxiv.org/abs/1001.0736]
- Gao X and Xiong H. A hybrid wavelet convolution network with sparse-coding for image super-resolution[C]//2016 International Conference on Image Processing (ICIP). Phoenix, AZ, USA: IEEE, 2016: 1439–1443. [DOI:10.1109/ICIP.2016.7532596]
- Gomez A N, Ren M, Urtasun R and Grosse R B. The reversible residual network: Backpropagation without storing activations[C]//Advances in Neural Information Processing Systems 30. Long Beach, CA, USA: Curran Associates, Inc., 2017: 2214–2224. [https://papers.nips.cc/paper/6816-the-reversible-residual-network-backpropagation-without-storing-activations.pdf]
- Gramfort A and Kowalski M. Improving M/EEG source localization with an inter-condition sparse prior[C]//Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009. 2009. [DOI:10.1109/ISBI.2009.5193003]
- Gu S, Zuo W, Xie Q, Meng D, Feng X and Zhang L. Convolutional sparse coding for image super-resolution[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015: 1823–1831. [DOI:10.1109/ICCV.2015.212]
- Hammond D K and Vandergheynst P. Wavelets on graphs via spectral graph theory[J]. Applied and Computational Harmonic Analysis, 2011, 30(2): 129–150. [DOI:10.1016/j.acha.2010.04.005]
- Heide F, Heidrich W and Wetzstein G. Fast and flexible convolutional sparse coding[C]//2015 IEEE Conference on

- Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA: IEEE, 2015: 5135–5143. [DOI:10.1109/CVPR.2015.7299149]
- Henaff M, Bruna J and LeCun Y. Deep convolutional networks on graph-structured data[J]. arXiv preprint, 2015
- Huang J and Zhang T. The benefit of group sparsity[J]. Annals of Statistics, 2010, 38(4): 1978–2004. [DOI:10.1214/09-AOS778]
- Huang J, Zhang T and Metaxas D. Learning with structured sparsity[J]. Journal of Machine Learning Research, 2011, 12: 3371–3412. [DOI:10.1145/1553374.1553429]
- Huang Y, Shao L and Frangi A F. Simultaneous super-resolution and cross-modality synthesis of 3D medical images using weakly-supervised joint convolutional sparse coding[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 5787–5796. [DOI:10.1109/CVPR.2017.613]
- Jacob L, Obozinski G and Vert M-P. Group Lasso with overlap and graph Lasso[C]//Proceedings of the 26th International Conference on Machine Learning. Montreal, QC, Canada: Omnipress, 2009: 433–440. [DOI:10.1145/1553374.1553431]
- Jacobsen J-H, Smeulders A and Oyallon E. i-RevNet: Deep invertible networks[C]//6th International Conference on Learning Representation. Vancouver, BC, Canada: 2018: 1–11
- Jas M, La Tour T D, Simsekli U and Gramfort A. Learning the morphology of brain signals using alpha-stable convolutional sparse coding[C]//Advances in Neural Information Processing Systems 30. Long Beach, CA, USA: Curran Associates, Inc., 2017: 1099–1108. [https://papers.nips.cc/paper/6710-learning-the-morphology-of-brain-signals-using-alpha-stable-convolutional-sparse-coding.pdf]
- Jenatton R, Audibert J-Y and Bach F. Structured variable selection with sparsity-inducing norms[J]. Journal of Machine Learning Research, 2009, 12: 2777–2824. [http://jmlr.org/papers/volume12/jenatton11b/jenatton11b.pdf]
- Jenatton R, Gramfort A, Michel V, Obozinski G, Eger E, Bach F and Thirion B. Multiscale Mining of fMRI Data with Hierarchical Structured Sparsity[J]. SIAM Journal on Imaging Sciences, 2012, 5(3): 835–856. [DOI:10.1137/110832380]
- Jenatton R, Mairal J, Obozinski G and Bach F. Proximal methods for hierarchical sparse coding[J]. Journal of Machine Learning Research, 2010, 12: 2297–2334. [http://jmlr.csail.mit.edu/papers/volume12/jenatton11a/jenatton11a.pdf]
- Jenatton R, Obozinski G and Bach F. Structured sparse principal component analysis[C]//13th International Conference on Artificial Intelligence and Statistics. Sardinia, Italy: PMLR, 2010: 366–373. [http://arxiv.org/abs/0909.1440]
- Kim S and Xing E P. Tree-guided group Lasso for multi-task regression with structured sparsity[C]//Proceedings of the 27th International Conference on Machine Learning. Haifa, Israel: Omnipress, 2010: 543–550. [DOI:10.1214/12-AOAS549]
- Krause A and Cevher V. Submodular dictionary selection for sparse representation[C]//Proceedings of the 27th International Conference on Machine Learning. Haifa, Israel: Omnipress, 2010: 567–574. [http://www.icml2010.org/papers/366.pdf]
- Krizhevsky A, Sutskever I and Hinton G E. ImageNet classification with deep convolutional neural networks[C]//Advances in Neural Information Processing Systems 25. Lake Tahoe, NV, USA: Curran Associates, Inc., 2012: 1097–1105. [DOI:10.1145/3065386]
- Labate D, Lim W-Q, Kutyniok G and Weiss G. Sparse multidimensional representation using shearlets[C]//Wavelets XI. San Diego, CA, USA: SPIE, 2005: 59140U. [DOI:10.1117/12.613494]
- Le Pennec E and Mallat S. Sparse geometric image representations with bandelets[J]. IEEE Transactions on Image Processing, 2005, 14(4): 423–428. [DOI:10.1109/TIP.2005.843753]
- Lesage S, Gribonval R, Bimbot F and Benaroya L. Learning unions of orthonormal bases with thresholded singular value decomposition[C]//Proceedings. IEEE International Conference on Acoustics, Speech and Signal Processing, 2005. Philadelphia, PA, USA: 2005: V-293–V–296. [DOI:10.1109/ICASSP.2005.1416298]
- Levie R, Monti F, Bresson X and Bronstein M M. CayleyNets: Graph convolutional neural networks with complex rational spectral filters[J]. IEEE Transactions on Signal Processing, 2019, 67(1): 97–109. [DOI:10.1109/TSP.2018.2879624]
- Li H and Lin Z. Provable accelerated gradient method for nonconvex low rank optimization[J]. Machine Learning, 2019. [DOI:10.1007/s10994-019-05819-w]
- Li M, Xie Q, Zhao Q, Wei W, Gu S, Tao J and Meng D. Video Rain Streak Removal by Multiscale Convolutional Sparse Coding[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2018: 6644–6653. [DOI:10.1109/CVPR.2018.00695]
- Li R, Wang S, Zhu F, Artificial J H-T-S A C on and 2018 undefined. Adaptive graph convolutional neural networks[C]//Thirty-Second AAAI Conference on Artificial Intelligence. New Orleans, LA, USA: AAAI, 2018: 3546–3553. [https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/viewPaper/16642]
- Li Y, Dai W, Zou J, Xiong H and Zheng Y F. Structured sparse representation with union of data-driven linear and multilinear subspaces model for compressive video sampling[J]. IEEE Transactions on Signal Processing, 2017, 65(19): 5062–5077. [DOI:10.1109/TSP.2017.2721905]
- Liu G, Lin Z, Yan S, Sun J, Yu Y and Ma Y. Robust recovery of subspace structures by low-rank representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 171–184. [DOI:10.1109/TPAMI.2012.88]
- Liu J, Garcia-Cardona C, Wohlberg B and Yin W. Online convolutional dictionary learning[C]//2017 International Conference on Image Processing (ICIP). Beijing, China: IEEE, 2017: 1707–1711. [DOI:10.1109/ICIP.2017.8296573]

- Liu L, Wu J, Li D, Senhadji L and Shu H. Fractional wavelet scattering network and applications[J]. IEEE Transactions on Biomedical Engineering, 2019, 66(2): 553–563. [DOI:10.1109/TBME.2018.2850356]
- Liu M and Chen S. Label-dependent sparse representation for synthetic aperture radar target configuration recognition[J]. International Journal of Remote Sensing, 2017, 38(17): 4868–4887. [DOI:10.1080/01431161.2017.1328144]
- Lu C, Feng J, Lin Z, Mei T and Yan S. Subspace clustering by block diagonal representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(2): 487–501. [DOI:10.1109/TPAMI.2018.2794348]
- Lu Y M and Do M N. Multidimensional directional filter banks and surfacelets[J]. IEEE Transactions on Image Processing, 2007, 16(4): 918–931. [DOI:10.1109/TIP.2007.891785]
- Lu Z and Peng Y. Latent semantic learning with structured sparse representation for human action recognition[J]. Pattern Recognition, 2013, 46(7): 1799–1809. [DOI:10.1016/j.patcog.2012.09.027]
- Mairal J, Bach F and Ponce J. Sparse modeling for image and vision processing[J]. Foundations and Trends in Computer Graphics and Vision, 2014, 8(2–3): 85–283. [DOI:10.1561/06000000058]
- Mairal J, Bach F, Ponce J and Sapiro G. Online Learning for matrix factorization and sparse coding[J]. Journal of Machine Learning Research, 2010, 11(1): 19–50. [<http://www.jmlr.org/papers/volume11/mairal10a/mairal10a.pdf>]
- Mairal J, Jenatton R, Obozinski G and Bach F. Convex and network flow optimization for structured sparsity[J]. Journal of Machine Learning Research, 2011, 12(9): 2681–2720. [<http://jmlr.org/papers/volume12/mairal11a/mairal11a.pdf>]
- Mallat S. Geometrical grouplets[J]. Applied and Computational Harmonic Analysis, 2009, 26(2): 161–180. [DOI:10.1016/j.acha.2008.03.004]
- Mallat S. Group invariant scattering[J]. Communications on Pure and Applied Mathematics, 2012, 65(10): 1331–1398. [DOI:10.1002/cpa.21413]
- Mallat S G and Zhang Z. Matching pursuits with time-frequency dictionaries[J]. IEEE Transactions on Signal Processing, 1993, 41(12): 3397–3415. [DOI:10.1109/78.258082]
- Meinshausen N and Bühlmann P. High-dimensional graphs and variable selection with the Lasso[J]. Annals of Statistics, 2006, 34(3): 1436–1462. [DOI:10.1214/0090536060000000281]
- Narang S K and Ortega A. Perfect reconstruction two-channel wavelet filter banks for graph structured data[J]. IEEE Transactions on Signal Processing, 2012, 60(6): 2786–2799. [DOI:10.1109/TSP.2012.2188718]
- Obozinski G, Taskar B and Jordan M I. Joint covariate selection and joint subspace selection for multiple classification problems[J]. Statistics and Computing, 2010. [DOI:10.1007/s11222-008-9111-x]
- Oyallon E, Belilovsky E and Zagoruyko S. Scaling the scattering transform: Deep hybrid networks[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 5619–5628. [DOI:10.1109/ICCV.2017.599]
- Oyallon E and Mallat S. Deep roto-translation scattering for object classification[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA: IEEE, 2015: 2865–2873. [DOI:10.1109/CVPR.2015.7258904]
- Papayan V, Romano Y and Elad M. Convolutional Neural Networks Analyzed via Convolutional Sparse Coding[R]. 2017, 18
- Papayan V, Romano Y, Elad M and Sulam J. Convolutional dictionary learning via local processing[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 5306–5314. [DOI:10.1109/ICCV.2017.5661]
- Papayan V, Romano Y, Sulam J and Elad M. Theoretical foundations of deep learning via sparse representations: A multilayer sparse model and its connection to convolutional neural networks[J]. IEEE Signal Processing Magazine, 2018, 35(4): 72–89. [DOI:10.1109/MSP.2018.2820224]
- Papayan V, Sulam J and Elad M. Working locally thinking globally: Theoretical guarantees for convolutional sparse coding[J]. IEEE Transactions on Signal Processing, 2017, 65(21): 5687–5701. [DOI:10.1109/TSP.2017.2733447]
- Pati Y C, Rezaiifar R and Krishnaprasad P S. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition[C]//Proceeding of 27th Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA, USA: IEEE, 1993: 40–44. [DOI:10.1109/ACSSC.1993.342465]
- Peng G-J. Adaptive ADMM for Dictionary Learning in Convolutional Sparse Representation[J]. IEEE Transactions on Image Processing, 2019. [DOI:10.1109/tip.2019.2896541]
- Peng H, Li B, Ling H, Hu W, Xiong W and Maybank S J. Salient object detection via structured matrix decomposition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 818–832. [DOI:10.1109/TPAMI.2016.2562626]
- Plaut E and Giryes R. Matching pursuit based convolutional sparse coding[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary, AB, Canada: IEEE, 2018: 6847–6851. [DOI:10.1109/ICASSP.2018.8461543]
- Quan T M and Jeong W K. Compressed sensing reconstruction of dynamic contrast enhanced MRI using GPU-accelerated convolutional sparse coding[C]//2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). Prague, Czech Republic: IEEE, 2016: 518–521. [DOI:10.1109/ISBI.2016.7493321]
- Ram I, Elad M and Cohen I. Generalized tree-based wavelet transform[J]. IEEE Transactions on Signal Processing, 2011, 59(9): 4199–4209. [DOI:10.1109/TSP.2011.2158428]
- Rubinstein R, Peleg T and Elad M. Analysis K-SVD: A

- dictionary-learning algorithm for the analysis sparse model[J]. IEEE Transactions on Signal Processing, 2013, 61(3): 661–677. [DOI:10.1109/TSP.2012.2226445]
- Rubinstein R, Zibulevsky M and Elad M. Double sparsity: Learning sparse dictionaries for sparse signal approximation[J]. IEEE Transactions on Signal Processing, 2010, 58(3): 1553–1564. [DOI:10.1109/TSP.2009.2036477]
- Kustamov R and Guibas L J. Wavelets on graphs via deep learning[C]//Advances in Neural Information Processing Systems 26. Lake Tahoe, NV, USA: Curran Associates, Inc., 2013: 998–1006. [DOI:10.1007/978-3-030-03574-7_5]
- Ruvolo P and Eaton E. Online multi-task learning based on K-SVD[C]//The ICML2013 Workshop on Theoretically Grounded Transfer Learning. Atlanta, GA, USA: PMLR, 2013. [https://www.seas.upenn.edu/~eeaton/papers/Ruvolo2013Online.pdf]
- Sandryhaila A and Moura J M F. Discrete signal processing on graphs[J]. IEEE Transactions on Signal Processing, 2013, 61(7): 1644–1656. [DOI:10.1109/TSP.2013.2238935]
- Serrano A, Heide F, Gutierrez D, Wetzstein G and Masia B. Convolutional sparse coding for high dynamic range imaging[J]. Computer Graphics Forum, 2016, 35(2): 153–163. [DOI:10.1111/cgf.12819]
- Shuman D I, Faraji M J and Vandergheynst P. A multiscale pyramid transform for graph signals[J]. IEEE Transactions on Signal Processing, 2016, 64(8): 2119–2124. [DOI:10.1109/TSP.2015.2512529]
- Shuman D I, Wismeyr C, Holighaus N and Vandergheynst P. Spectrum-adapted tight graph wavelet and vertex-frequency frames[J]. IEEE Transactions on Signal Processing, 2015, 63(16): 4223–4235. [DOI:10.1109/TSP.2015.2424203]
- Shuman D I, Narang S K, Frossard P, Ortega A and Vandergheynst P. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains[J]. IEEE Signal Processing Magazine, 2013, 30(3): 83–98. [DOI:10.1109/MSP.2012.2235192]
- Shuman D I, Ricaud B and Vandergheynst P. Vertex-frequency analysis on graphs[J]. Applied and Computational Harmonic Analysis, 2016, 40(2): 260–291. [DOI:10.1016/j.acha.2015.02.005]
- Sifre L and Mallat S. Combined scattering for rotation invariant texture analysis[C]//Proceedings of 2012 European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges, Belgium: i6doc, 2012: 127–132. [http://www.i6doc.com/en/livre/?GCOI=28001100967420]
- Silva G and Rodriguez P. Efficient convolutional dictionary learning using partial update fast iterative shrinkage-thresholding algorithm[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary, AB, Canada: IEEE, 2018: 4674–4678. [DOI:10.1109/ICASSP.2018.8462315]
- Sreter H and Giryes R. Learned convolutional sparse coding[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary, AB, Canada: IEEE, 2018: 2191–2195. [DOI:10.1109/ICASSP.2018.8462313]
- Sulam J, Pappayan V, Romano Y and Elad M. Multilayer convolutional sparse modeling: Pursuit and dictionary learning[J]. IEEE Transactions on Signal Processing, 2018, 66(15): 4090–4104. [DOI:10.1109/TSP.2018.2846226]
- Tanaka Y and Sakiyama A. M-channel oversampled graph filter banks[J]. IEEE Transactions on Signal Processing, 2014, 62(14): 3578–3590. [DOI:10.1109/tsp.2014.2328983]
- Teke O and Vaidyanathan P P. Extending classical multirate signal processing theory to graphs-Part II: M-channel filter banks[J]. IEEE Transactions on Signal Processing, 2017, 65(2): 423–437. [DOI:10.1109/TSP.2016.2620111]
- Titsvero M, Barbarossa S and Di Lorenzo P. Signals on graphs: Uncertainty principle and sampling[J]. IEEE Transactions on Signal Processing, 2016, 64(18): 4845–4860. [DOI:10.1109/TSP.2016.2573748]
- Turlach B A, Venables W N and Wright S J. Simultaneous variable selection[J]. Technometrics, 2005, 47(3): 349–363. [DOI:10.1198/004017005000000139]
- Varoquaux G, Jenatton R, Gramfort A, ... G O-N W on P and 2010 U. Sparse structured dictionary learning for brain resting-state activity modeling[C]//NIPS Workshop on Practical Applications of Sparse Modeling: Open Issues and New Directions. 2010: 1–7
- Velisavljević V, Beferull-Lozano B, Vetterli M and Dragotti P L. Directionlets: Anisotropic multidirectional representation with separable filtering[J]. IEEE Transactions on Image Processing, 2006, 15(7): 1916–1933. [DOI:10.1109/TIP.2006.877076]
- Vidal R, Ma Y and Sastry S. Generalized principal component analysis (GPCA)[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(12): 1945–1959. [DOI:10.1109/IJCNN.2005.1555814]
- Wang H, Li S, Zhou Y and Chen S. SAR automatic target recognition using a roto-translational invariant wavelet-scattering convolution network[J]. Remote Sensing, 2018, 10(4): 501. [DOI:10.3390/rs10040501]
- Wang Y, Yao Q, Kwok J T and Ni L M. Scalable online convolutional sparse coding[J]. IEEE Transactions on Image Processing, 2018, 27(10): 4850–4859. [DOI:10.1109/TIP.2018.2842152]
- Wang Y, Yao Q, Kwok J T and Ni L M. Online convolutional sparse coding with sample-dependent dictionary[C]//Proceedings of the 35th International Conference on Machine Learning. Stockholm, Sweden: PMLR, 2018: 5209–5218. [http://proceedings.mlr.press/v80/wang18k.html]
- Wen W, Wu C, Wang Y, Chen Y and Li H. Learning structured sparsity in deep neural networks[C]//Advances in Neural Information Processing Systems 29. Barcelona, Spain: Curran Associates, Inc., 2016: 2074–2082. [https://papers.nips.cc/paper/6504-learning-structured-sparsity-]

- in-deep-neural-networks.pdf]
- Wen Y, Liu W, Yang M, Fu Y, Xiang Y and Hu R. Structured occlusion coding for robust face recognition[J]. *Neurocomputing*, 2016, 178: 11–24. [DOI:10.1016/j.neucom.2015.05.132]
- Wen Z, Goldfarb D and Scheinberg K. Block coordinate descent methods for semidefinite programming[G]//International Series in Operations Research and Management Science. 2012. [DOI:10.1007/978-1-4614-0769-0_19]
- Wiatowski T and Bolcskei H. A mathematical theory of deep convolutional neural networks for feature extraction[J]. *IEEE Transactions on Information Theory*, 2018, 64(3): 1845–1866. [DOI:10.1109/TIT.2017.2776228]
- Wohlberg B. Efficient convolutional sparse coding[C]//2014 IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings (ICASSP). Florence, Italy: IEEE, 2014: 7173–7177. [DOI:10.1109/ICASSP.2014.6854992]
- Wu Z, Pan S, Chen F, Long G, Zhang C and Yu P S. A comprehensive survey on graph neural networks[J]. *arXiv preprint*, 2019
- Xiong H, Zhu L, Ma N and Zheng Y F. Scalable video compression framework with adaptive orientational multiresolution transform and nonuniform directional filterbank design[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2011, 21(8): 1085–1099. [DOI:10.1109/TCSVT.2011.2133310]
- Xiong Y, Xu C and Xiong H. Kernelized learning in deep scattering convolution networks[C]//2016 IEEE International Conference on Multimedia and Expo (ICME). Seattle, WA, USA: IEEE, 2016. [DOI:10.1109/ICME.2016.7552973]
- Xu B, Shen H, Cao Q, Qiu Y and Cheng X. Graph wavelet neural network[C]//7th International Conference on Learning Representation. New Orleans, LA, USA: 2019: 1–13 [http://arxiv.org/abs/1904.07735]
- Xu C, Dai W and Xiong H. Extended conjugate polar Fourier transform in convolution network[C]//2017 IEEE International Conference on Image Processing (ICIP). Beijing, China: IEEE, 2017: 2453–2457. [DOI:10.1109/ICIP.2017.8296723]
- Xu C and Xiong H. Reversible deep convolutional networks with iterated directional filter bank[C]//2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Shanghai, China: IEEE, 2016: 1586–1590. [DOI:10.1109/ICASSP.2016.7471944]
- Xu Y, Xiao T, Zhang J, Yang K and Zhang Z. Scale-invariant convolutional neural networks[J]. *arXiv preprint*, 2014. [http://arxiv.org/abs/1412.5104]
- Yan S, Xiong Y and Lin D. Spatial temporal graph convolutional networks for skeleton-based action recognition[C]//Thirty-Second AAAI Conference on Artificial Intelligence. New Orleans, LA, USA: AAAI, 2018: 7444–7452. [https://aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17135]
- Ye J C, Han Y and Cha E. Deep convolutional framelets: A general deep learning framework for inverse problems[J]. *SIAM Journal on Imaging Sciences*, 2017, 11(2): 991–1048. [DOI:10.1137/17M1141771]
- Yin R, Gao T, Lu Y M and Daubechies I. A Tale of Two Bases: Local-Nonlocal Regularization on Image Patches with Convolution Framelets[J]. *SIAM Journal on Imaging Sciences*, Society for Industrial & Applied Mathematics (SIAM), 2017, 10(2): 711–750. [DOI:10.1137/16m1091447]
- Yuan M. High dimensional inverse covariance matrix estimation via linear programming[J]. *Journal of Machine Learning Research*, 2010, 11: 2261–2286. [http://jmlr.csail.mit.edu/papers/v11/yuan10b.html]
- Yuan M and Lin Y. Model selection and estimation in regression with grouped variables[R]. 2006, 68
- Zeiler M D, Krishnan D, Taylor G W and Fergus R. Deconvolutional networks[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, CA, USA: IEEE, 2010: 2528–2535. [DOI:10.1109/CVPR.2010.5539957]
- Zhang H and Patel V. Convolutional sparse coding-based image decomposition[C]//Proceedings of the British Machine Vision Conference 2016. York, UK: BMVA Press, 2016: 125.1–125.11. [DOI:10.5244/C.30.125]
- Zhang H and Patel V M. Convolutional sparse and low-rank coding-based rain streak removal[C]//2017 IEEE Winter Conference on Applications of Computer Vision (WACV). Santa Rosa, CA, USA: IEEE, 2017: 1259–1267. [DOI:10.1109/WACV.2017.145]
- Zhao P, Rocha G and Yu B. The composite absolute penalties family for grouped and hierarchical variable selection[J]. *Annals of Statistics*, 2009, 37(6 A): 3468–3497. [DOI:10.1214/07-AOS584]
- Zhou G, Lu Z and Peng Y. L1-graph construction using structured sparsity[J]. *Neurocomputing*, 2013, 120: 441–452. [DOI:10.1016/j.neucom.2013.03.045]
- Zhou Y, Chang H, Barner K, Spellman P and Parvin B. Classification of histology sections via multispectral convolutional sparse coding[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014: 3081–3088. [DOI:10.1109/CVPR.2014.394]
- Zhu Y and Lucey S. Convolutional sparse coding for trajectory reconstruction[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(3): 529–540. [DOI:10.1109/TPAMI.2013.2295311]
- Zou D and Lerman G. Graph convolutional neural networks via scattering[J]. *arXiv preprint*, 2018

作者简介



熊红凯, 1973 年生, 男, 特聘教授, 主要研究方向为信号处理与编码、多媒体网络通信、数据表示与学习。Email: xionghongkai@sjtu.edu.cn.



于俊清, 通信作者, 男, 教授, 主要研究方向为基于内容的视频分析、索引与检索、多核计算与流编译、视频情感计算、网络安全与大数据处理。Email: yjqhust@qq.com.

戴文睿, 男, 副教授, 主要研究方向为图像/视频编码, 数据压缩和生物医学信息处理。Email: daiwenrui@sjtu.edu.cn.

林宙晨, 男, 教授, 主要研究方向为机器学习、模式识别、计算机视觉、图像处理和数值优化。Email: zlin@pku.edu.cn.

吴飞, 男, 教授, 主要研究方向为人工智能、跨媒体计算、多媒体分析与检索和统计学习理论。Email: wufei@zju.edu.cn.

申扬眉, 女, 博士研究生, 主要研究方向为结构化稀疏表示与字典学习, 非规则数据压缩。Email: shenyangmei0214@sjtu.edu.cn.

徐明星, 男, 博士研究生, 主要研究方向为图信号处理, 图卷积神经网络理论及其应用。Email: xumingxing@sjtu.edu.cn.