

# 基于改进K-SVD字典学习的地震信号去噪

侯晓敏

成都理工大学数学科学学院, 四川 成都

收稿日期: 2025年4月29日; 录用日期: 2025年5月23日; 发布日期: 2025年5月30日

## 摘要

字典学习作为近年来的研究热点, 因其在稀疏表示方面的优势而受到广泛关注。该方法通过构建合适的稀疏字典, 将地震数据表示为一组原子的线性组合, 从而实现地震信号与噪声的有效分离。然而, 传统的K-奇异值分解(K-SVD)字典学习算法在应用于地震数据去噪的过程中, 尽管展现出了良好的适应性和去噪性能, 但仍存在一些不足之处。本文在传统的K-SVD字典学习算法的基础上创新性地结合统计理论, 在多个重要步骤对其进行改进, 提出一种新的优化字典学习去噪算法。并通过对合成地震数据和实际地震数据的去噪实验验证: 本文提出的优化字典学习算法与传统的K-SVD字典学习算法相比, 在衡量去噪效果的多个维度都具有明显优势。

## 关键词

地震数据去噪, 稀疏表示, 字典学习, 贝叶斯分析

# Seismic Signal Denoising Based on Improved K-SVD Dictionary Learning

Xiaomin Hou

School of Mathematical Sciences, Chengdu University of Technology, Chengdu Sichuan

Received: Apr. 29<sup>th</sup>, 2025; accepted: May 23<sup>rd</sup>, 2025; published: May 30<sup>th</sup>, 2025

## Abstract

Dictionary learning, as a research hotspot in recent years, has received widespread attention due to its advantages in sparse representation. This method constructs a suitable sparse dictionary to represent seismic data as a linear combination of atoms, thereby achieving effective separation of seismic signals and noise. However, the traditional K-Singular Value Decomposition (K-SVD) dictionary learning algorithm, although exhibiting good adaptability and denoising performance when applied to seismic data denoising, still has some shortcomings. This article innovatively combines statistical

theory with the traditional K-SVD dictionary learning algorithm, improves it in multiple important steps, and proposes a new optimized dictionary learning denoising algorithm. And through denoising experiments on synthetic seismic data and actual seismic data, it was verified that the optimized dictionary learning algorithm proposed in this paper has significant advantages compared to the traditional K-SVD dictionary learning algorithm in multiple dimensions of measuring denoising effectiveness.

## Keywords

Earthquake Data Denoising, Sparse Representation, Dictionary Learning, Bayesian Analysis

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

地震数据是通过地震波在地壳内传播时,地震传感器(如地震仪)所接收的信号。在数据采集过程中,受限于复杂的观测环境和野外地形条件,所获取的信号往往不仅包含地下地质结构的有效信息,还常常受到多种噪声源(如环境干扰、设备误差、信号多次反射等)的污染,导致数据质量下降,进而影响后续的地震数据处理与解释[1]-REF\_Ref19978\r\h[5]。地震数据去噪的核心目标,就是从含噪的原始数据中提取出真实的地震信号,为更高精度的分析和解释提供可靠的数据基础。

字典学习(Dictionary Learning)作为一种高效的信号表示技术,近年来已被广泛应用于图像处理、语音识别以及地震数据分析等多个领域,展现出良好的适应性与表现力。在字典学习正式被提出之前,一些基础性的研究已经为其发展奠定了理论框架[6]-[10]。1999年,Engan等人首次提出了MOD(Method of Optimal Directions)算法,这是最早的数据驱动字典学习方法[11]-[19]。其核心思想是:先固定字典,利用稀疏编码(如OMP、BP)求解信号的稀疏表示;再固定稀疏系数,更新字典以最小化重构误差。MOD方法开启了从数据中学习字典的研究方向,但它的优化过程较为简单,收敛性较差。2006年,Aharon、Elad和Bruckstein提出了K-SVD(K-means Singular Value Decomposition)算法,这标志着字典学习方法的正式成熟[19]-[25]。K-SVD采用迭代更新策略:在稀疏编码阶段:使用OMP(正交匹配追踪)等方法求解稀疏系数;在字典更新阶段:利用奇异值分解(SVD)更新字典原子,以最小化重构误差。

传统的K-奇异值分解(K-SVD)字典学习算法在应用于地震数据去噪的过程中,尽管展现出了良好的适应性和去噪性能,但仍存在一些不足之处。如部分原始信号在去噪过程中被削弱甚至丢失,以及算法计算效率较低。为此,需对传统K-SVD方法进行优化或引入更高效的稀疏表示与字典更新策略,以提升去噪性能并保留更多的原始地震信息。针对当前字典学习算法在地震去噪应用中的不足与挑战,本文提出一种优化K-SVD字典学习算法,并通过使用几种去噪方法对合成地震记录和实际地震记录的数值进行实验,对比验证该方法能够实现更有效的噪声抑制,提升去噪性能。

## 2. 理论方法

### 2.1. K-SVD字典学习算法

在地震数据采集过程中,噪声的来源多种多样,且由于各种因素的干扰,噪声的存在是不可避免的。这些噪声会降低地震信号的质量,影响地下结构的准确解释。传统去噪方法如小波变换、经验模态分解

(EMD)、低秩矩阵分解等存在一定的局限性,如信号损失、对复杂噪声适应能力弱等。K-SVD (K-Singular Value Decomposition)字典学习算法能够自适应学习数据特征,提高信号稀疏表示的能力,在地震数据去噪方面展现出良好的性能。

给定地震数据矩阵  $X \in R^{m \times n}$ , 其中  $m$  为信号长度,  $n$  为道数, K-SVD 通过最小化以下目标函数来学习字典  $D$  和稀疏表示矩阵  $A$ :

$$\min_{D,A} \|X - DA\|_F^2 \text{ s.t. } \forall i, \|a_i\|_0 \leq T_0 \quad (1)$$

其中,  $D \in R^{m \times k}$  为学习得到的字典,  $k$  为字典原子的个数;  $A \in R^{k \times n}$  为稀疏表示矩阵;  $T_0$  是稀疏度约束(即每个数据点最多由  $T_0$  个字典原子表示)。

传统 K-SVD 字典学习的主要步骤:

1) 初始化字典: ① 随机选取数据样本(从地震数据矩阵  $X$  中随机抽取  $k$  列作为初始字典) ② 给定一个字典作为初始字典(如 DCT)。

2) 稀疏编码: 在固定字典  $D$  的情况下, 使用正交匹配追踪(OMP)算法求解稀疏表示系数  $A$ 。

3) 字典更新: 在固定系数表示  $A$  的情况下, 使用奇异值分解(SVD)更新字典  $D$ ; 逐列优化字典原子, 提取主导特征, 提高表示能力。

4) 迭代更新, 直至收敛到最大迭代次数。

K-SVD 字典学习算法在地震数据去噪中的应用: 地震数据去噪的核心目标是分离有效信号与噪声, K-SVD 通过学习地震信号的自适应字典, 使得地震信号能在该字典上稀疏表示, 而噪声部分难以被稀疏表达, 从而实现去噪。

具体处理流程如下:

1) 数据预处理: 选取原始地震数据  $X$ , 进行标准化(零均值、单位方差); 设定字典大小  $k$  和稀疏度  $T_0$ 。

2) K-SVD 字典学习: 使用 K-SVD 训练自适应学习字典  $D$ , 并计算稀疏表示矩阵  $A$ 。

3) 信号重构: 用优化后的  $D$  和  $A$  进行重构。

## 2.2. 贝叶斯分析

贝叶斯算法基于贝叶斯定理(Bayes' Theorem), 结合先验信息和数据观测, 构建概率模型, 从而进行推理、估计和预测。

贝叶斯定理描述了在已知先验信息的情况下, 如何更新对未知量的概率分布:

$$p(\theta|X) = \frac{p(X|\theta)p(\theta)}{p(X)} \quad (2)$$

其中:

$p(\theta|X)$  为后验概率, 表示在观测数据  $X$  之后, 参数  $\theta$  取值的概率。  $p(X|\theta)$  为似然函数, 表示在参数为  $\theta$  时, 数据  $X$  发生的概率。

$p(\theta)$  为先验概率, 表示在未观察数据前, 对  $\theta$  的信念。

$p(X)$  为证据, 表示数据  $X$  发生的总概率:

$$p(X) = \int p(X|\theta)p(\theta)d\theta \quad (3)$$

贝叶斯方法的优势在于: 能够结合先验知识, 在数据稀缺时仍能提供合理的推断, 并且避免过拟合, 具有更高的鲁棒性。

### 2.3. 基于贝叶斯分析的优化字典学习算法

在传统的 K-SVD 字典学习算法中, 稀疏编码步骤常使用 OMP、LASSO 等贪心或凸优化方法进行求解。但这些方法通常只是最小化重构误差, 并未利用数据本身的统计特性。贝叶斯 OMP (Bayesian OMP, BOMP) 是对传统 OMP 算法的改进版本, 将统计理论引入稀疏编码阶段, 通过结合先验分布和噪声模型, 对稀疏系数进行后验估计, 从而提高了对噪声的鲁棒性和自适应能力。相对于传统 OMP, 贝叶斯 OMP 不仅能够更好地平衡重构误差和稀疏性, 还能提高算法的鲁棒性和稳定性, 使其更准确地分辨有效信号与噪声, 在地震数据等高噪声环境下, 发挥明显的优势。

贝叶斯 OMP 的数学模型:

假设观测信号  $x$  由稀疏表示  $\alpha$  乘以字典  $D$  生成, 并加上高斯噪声:

$$x = D\alpha + \varepsilon, \varepsilon \sim N(0, \sigma^2 I) \quad (4)$$

1) 先验假设

① 稀疏系数  $\alpha$  服从稀疏先验, 可以使用拉普拉斯分布或高斯-拉普拉斯混合分布:

$$P(\alpha) = \prod_i \frac{\lambda}{2} e^{-\lambda|\alpha_i|} \quad (5)$$

其中,  $\lambda$  控制稀疏度。

② 噪声服从高斯分布:

$$P(x|D, \alpha) = N(D\alpha, \sigma^2 I) \quad (6)$$

2) 后验概率计算

根据贝叶斯公式, 目标是计算:

$$P(\alpha|x, D) \propto P(x|D, \alpha)P(\alpha) \quad (7)$$

但是由于后验分布不可解析, BOMP 采用最大后验估计(MAP)或变分推断来估计稀疏系数。

贝叶斯 OMP 的算法步骤:

BOMP 采用与传统 OMP 类似的逐步选择原子策略, 但引入了贝叶斯估计进行优化:

1) 初始化: 设定残差  $r = x$ , 原子索引集合  $S = \emptyset$ , 初始稀疏解  $\alpha = 0$ 。

2) 贝叶斯原子选择:

① 计算字典原子的后验概率, 选择最可能的字典原子。

$$j^* = \arg \max_j P(\alpha_j | x, D) \quad (8)$$

其中,  $P(\alpha_j | x, D)$  可通过变分贝叶斯或最大后验估计(MAP)近似求解。

② 考虑噪声水平计算马氏距离(Mahalanobis Distance)来提高原子选择的鲁棒性:

$$j^* = \arg \max_j \frac{|D_j^T r|}{\sigma_j} \quad (9)$$

其中,  $\sigma_j$  由后验估计得到, 衡量噪声对原子选择的影响。

3) 更新稀疏解: 在选择了原子  $j^*$  后, 使用贝叶斯估计优化稀疏系数:

$$\alpha_S = \arg \min_{\alpha} \left\{ \|x - D_S \alpha\|^2 + \lambda \|\alpha\|_1 \right\} \quad (10)$$

这通常可以使用贝叶斯线性回归或期望最大化(EM)算法迭代求解。

4) 更新残差:

$$r = x - D_S \alpha_S \quad (11)$$

5) 稀疏度自动判断: 传统 OMP 需要预设稀疏度  $T_0$ , 而 BOMP 通过计算后验概率收敛条件自适应停止:

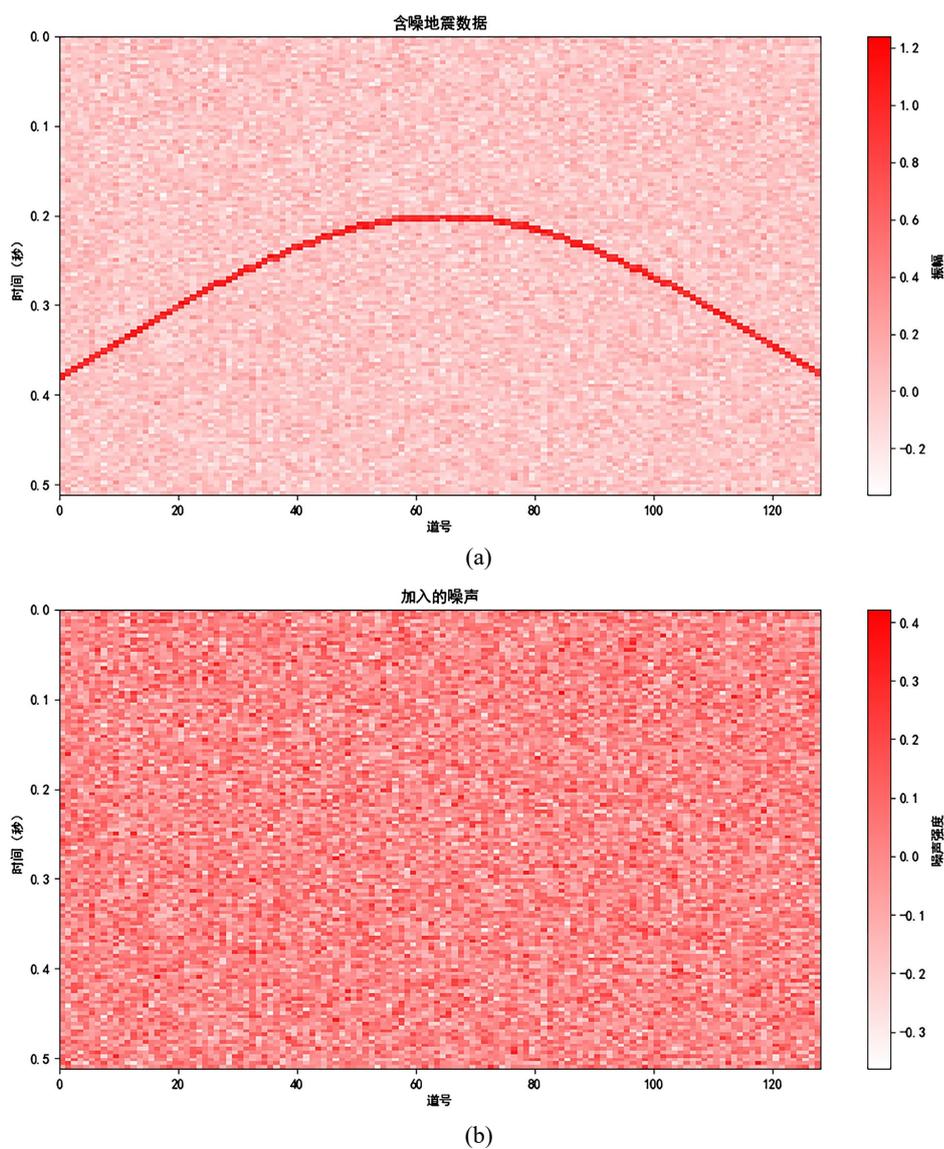
$$P(\alpha_j | x, D) < \text{阈值} \Rightarrow \text{停止迭代} \quad (12)$$

当新增的原子贡献很小, 则终止。

### 3. 数值实验

为验证本文提出的基于统计理论的优化字典学习算法在地震数据去噪的优越性, 分别采用中值滤波、PCA、K-SVD 以及本文方法对合成地震数据和实际地震数据进行对比实验。

#### 3.1. 合成地震数据



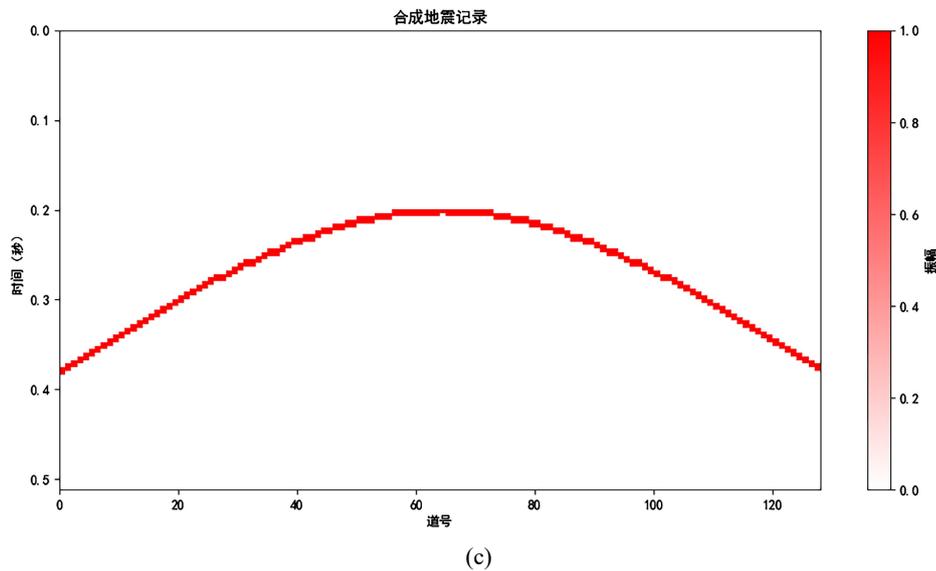


Figure 1. Synthetic seismic data  
图 1. 合成地震数据

1) 首先, 通过双曲同相轴合成方法生成原始地震数据, 并在此基础上添加随机噪声, 从而构造出含噪地震数据, 作为实验对象进行去噪测试。图 1(a)是原始无噪地震数据, 其尺寸为  $128 \times 128$ , 即包含 128 道, 每道包含 128 个采样点, 采样时间间隔为 4 ms。在此基础上加入标准差为 0.1 的随机噪声(图 1(c)), 生成图 1(b)所示含噪地震数据。在本文的模拟地震数据研究中, 所添加的随机噪声均通过调用 `randn` 函数产生。由于 `randn` 函数产生的是服从均值为 0、标准差为 1 的标准正态分布(高斯分布)的随机数, 因此加入的随机噪声属于高斯噪声。

对于去噪效果, 这里主要是比较地震数据去噪前后的信噪比提升幅度, 即采用信噪比来衡量两组实验的去噪效果。将信噪比定义为原始地震数据的功率与噪声数据的功率之比值取对数, 再乘以 10:

$$SNR(dB) = 10 \lg \frac{P_s}{P_n} \quad (13)$$

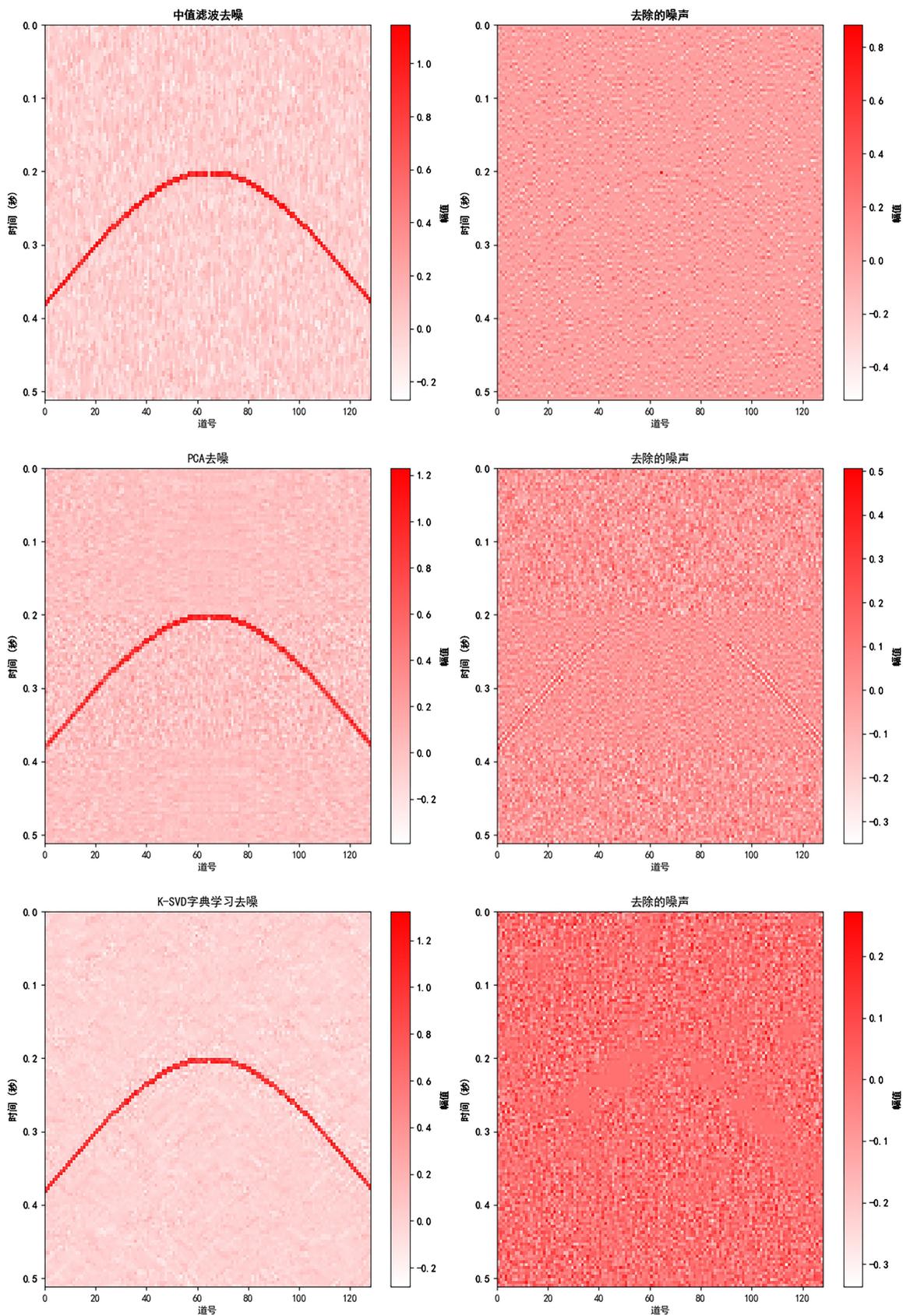
其中,  $P_s$  是信号功率(power of signal): 计算原始信号的平方和的均值;  $P_n$  是噪声功率(power of noise): 计算噪声(即含噪信号减去原始信号)的平方和的均值; SNR: 使用上述公式计算信噪比, 单位是分贝(dB)。

实验结果如表 1 所示, 从表中可以看出, 对于具有相同信噪比的同一含噪地震数据进行去噪处理, 几种去噪方法都可以提高地震数据的信噪比。相较于中值滤波和 PCA 这两种传统去噪方法, 字典学习方法去噪能够更进一步提高信噪比, 且本文方法去噪后的信噪比又明显高于 K-SVD 方法。图 2 的四张图依次为中值滤波、PCA、传统 K-SVD 和本文方法对含噪地震数据去噪后的结果以及去除的噪声。从中可以直观地看出, 本文提出的基于统计理论的优化字典学习方法相较于其他三种方法, 在去除更多随机噪声的同时, 还能进一步保护有效信号, 因此本文提出的方法在地震数据处理应用中具有更佳的去噪性能。

Table 1. Denoising signal-to-noise ratio for different methods ( $\sigma_n = 0.1$ )

表 1. 不同方法去噪信噪比( $\sigma_n = 0.1$ )

	含噪地震数据	中值滤波	PCA	K-SVD	本文方法
SNR (dB)	1.92	4.24	5.06	7.70	10.54



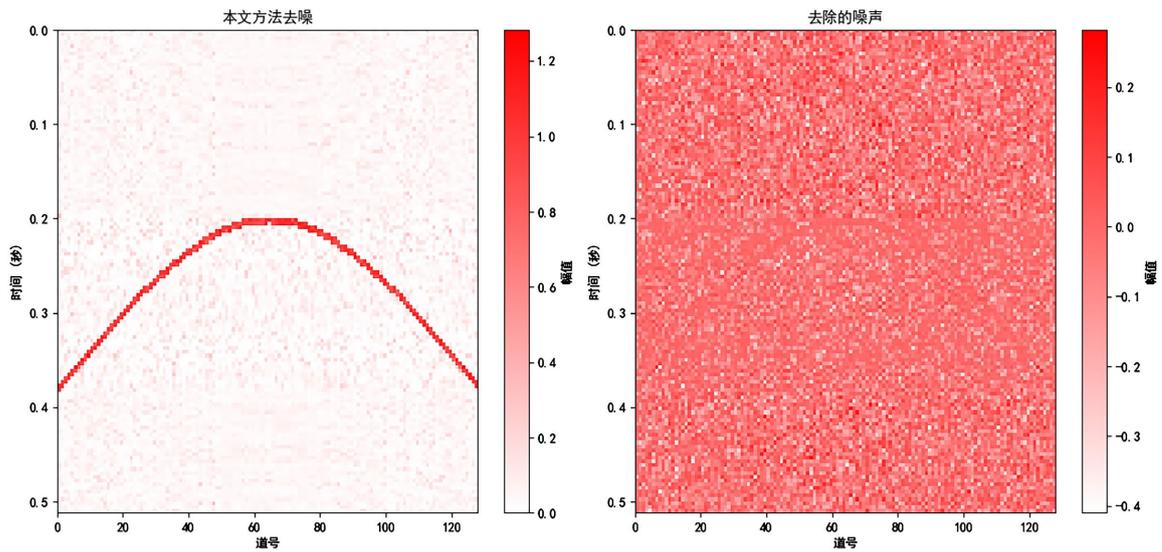


Figure 2. Denoising effect and noise removed by different methods

图 2. 不同方法去噪效果及去除的噪声

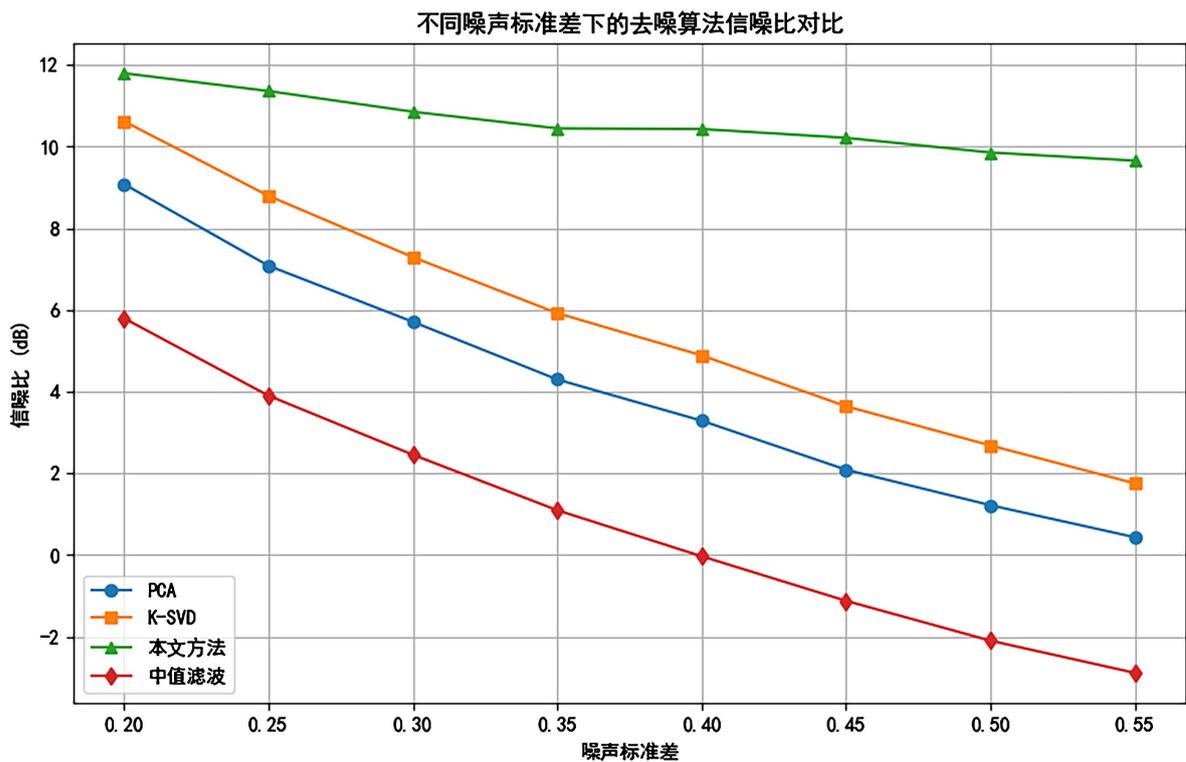


Figure 3. Signal-to-noise ratio curve under different noise intensities

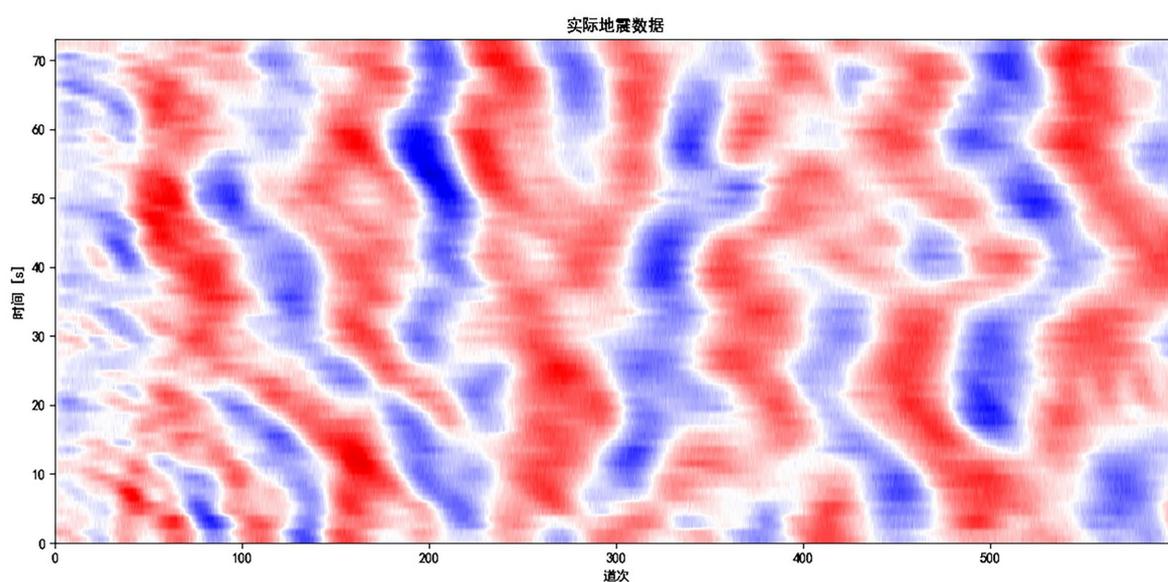
图 3. 不同噪声强度下的信噪比曲线

2) 稳健性分析: 稳健性分析指的是评估去噪方法在面对不同类型、强度或分布的噪声时, 依然能够有效去除噪声并保留信号特征的能力。稳健性分析不仅能检验去噪方法的效果, 也能帮助改进方法以适应更复杂的地震数据环境。这里使用四种地震数据去噪方法, 在不同噪声强度(改变噪声的标准差)下进行去噪操作。每个噪声强度下都需要运行去噪模型, 并记录去噪后的信噪比。图 3 即为在不同噪声强度下

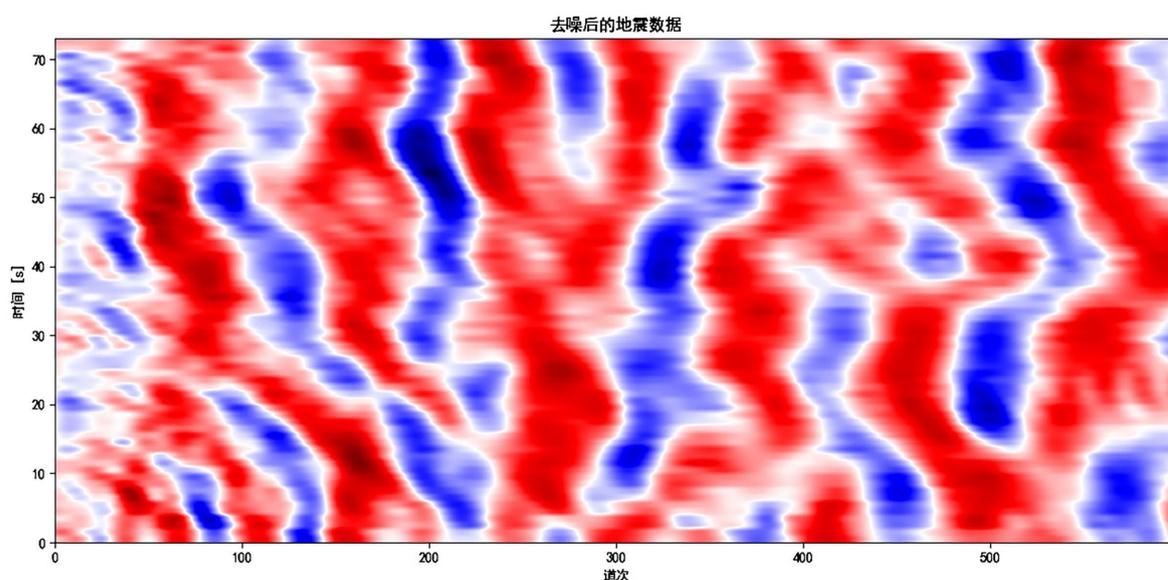
绘制的信噪比曲线，从中可以直观地看出不同的去噪方法在各个噪声水平下的表现：首先在相同的噪声水平下，本文方法去噪后的信噪比始终是最高的。其次，随着噪声强度的增加，其他三种方法去噪效果均明显下降，而本文提出的去噪方法依旧能够保持较高的稳定性。

### 3.2. 实际地震数据

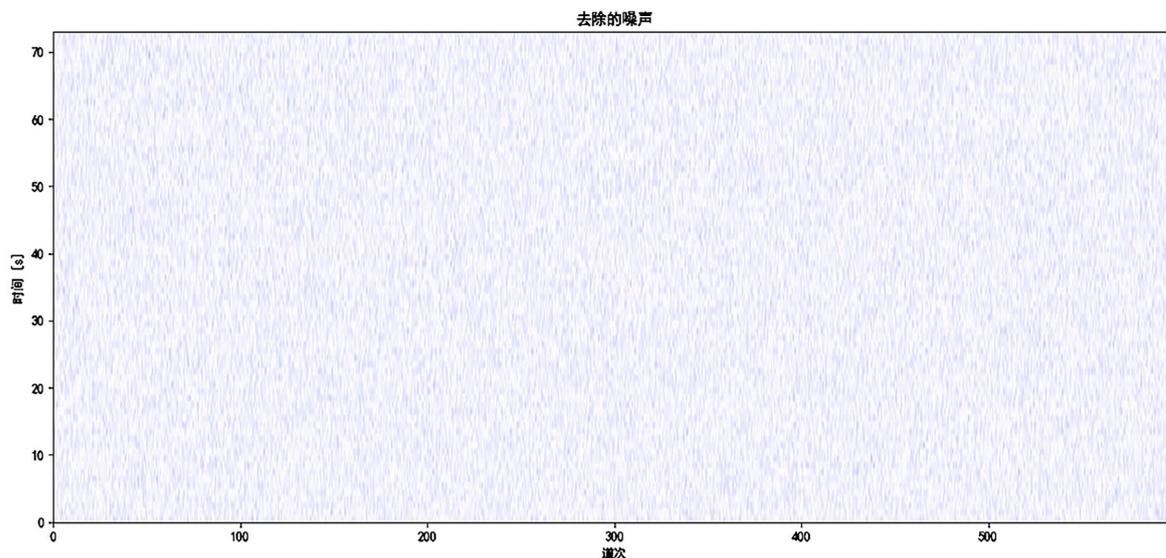
为验证本文所提的地震数据去噪算法在真实地震数据处理中的效果，将其应用于一条实际采集的地震剖面。图 4(a)展示了该地震剖面，共 73 道，每道包含 5999 个采样点，原始地震记录中充斥着大量的无规则随机噪声，严重干扰了有效信号的识别与解释，使用本文方法对该地震数据进行了去噪处理，结果如图 4(b)所示：可以看出随机噪声得到明显去除，有效信号进一步突显。图 4(c)为去除的噪声。



(a) 实际地震数据剖面图



(b) 本文方法去噪后的地震剖面图



(c) 去除的噪声

**Figure 4.** Denoising effect of the method proposed in this article on actual seismic data  
**图 4.** 本文方法对实际地震数据的去噪效果

### 3.3. 总结

贝叶斯估计的核心在于利用先验分布对信号和噪声进行建模。地震数据的噪声通常是复杂且具有多种来源的,而贝叶斯方法通过合理的先验信息(例如信号的平滑性、噪声的分布等)可以显著提升估计的精度。本文主要是利用统计理论中贝叶斯方法的独特性质,在传统 K-SVD 字典学习算法的基础上做出进一步改进优化,提出一种新的优化字典学习算法,并通过实验验证该方法在地震数据去噪中的性能进一步提高。

### 致 谢

在我完成这篇论文的过程中,得到了许多人的帮助和支持。在此,我谨向他们表示衷心的感谢。首先,我要感谢我的导师,感谢您在我整个研究过程中给予的悉心指导和不懈支持。您的严谨治学态度和渊博的知识让我受益匪浅,您的耐心和鼓励让我在遇到困难时始终保持信心。没有您的帮助,这篇论文无法顺利完成。

感谢我的家人,尤其是我的父母,感谢你们一直以来的理解、支持和鼓励。在我追求学术道路的过程中,你们无私的爱和鼓励让我始终保持动力,克服了一次次的挑战。

感谢我的同学和朋友们,感谢你们在学术和生活上的帮助与陪伴,尤其是在研究过程中一起讨论、分享思想,让我受益良多。你们的建议和反馈对于我的研究有着不可或缺的推动作用。同时,感谢实验室的所有老师和同学们,感谢你们在学术和生活上的支持,尤其是在实验设备和数据分析方面的帮助,让我的研究工作能够顺利进行。

最后,感谢所有曾经帮助过我的人,感谢你们为我的学术成长和论文完成所作出的贡献。

### 参考文献

- [1] 周绪文. 反射波地震勘探方法[M]. 北京: 石油工业出版社, 1989.
- [2] 牟永光. 地震数据处理方法[M]. 北京: 石油工业出版社, 2007.

- [3] 王有新. 应用地震数据处理方法[M]. 北京: 石油工业出版社, 2009.
- [4] 熊翥. 地震数据处理应用技术(修订版)(BZ) [M]. 北京: 石油工业出版社, 2008.
- [5] 吴招才, 刘天佑. 地震数据去噪中的小波方法[J]. 地球物理学进展, 2008, 23(2): 493-499.
- [6] 彭才, 朱仕军, 孙建库, 等. 基于独立成分分析的地震数据去噪[J]. 勘探地球物理进展, 2007, 30(1): 30-32+68.
- [7] 张博. 基于稀疏变换的地震数据去噪方法研究[D]: [博士学位论文]. 杭州: 浙江大学, 2013.
- [8] 宋西林. 基于自监督网络学习与自适应中值滤波的地震数据去噪方法研究[J]. 应用数学进展, 2024, 13(10): 4478-4485.
- [9] 代丽艳, 董宏丽, 李学贵. 微地震数据去噪方法综述[J]. 吉林大学学报(地球科学版), 2019, 49(4): 1145-1159.
- [10] 张孝珍, 董汉强, 侯国文, 等. 地震勘探中的去噪技术新进展[J]. 勘探地球物理进展, 2009, 32(3): 172-178+155.
- [11] Engan, K., Aase, S.O. and Husoy, J.H. (1999) Method of Optimal Directions for Frame Design. *International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, 15-19 March 1999, 2443-2446. <https://doi.org/10.1109/icassp.1999.760624>
- [12] 王威, 闫军部. 一种数字图像低通滤波方法研究[J]. 软件导刊, 2012, 11(2): 145-146.
- [13] 李新辉. 高通滤波方法, 高通滤波器和主动降噪系统[P]. 中国专利, CN202010754075.4. 2025-04-09.
- [14] 党人庆, 万志强. 一种带通滤波方法及其初步应用[J]. 气象科学研究院院刊, 1989(1): 29-34.
- [15] 李建国, 吴俊军, 余刚, 等. 一种地震数据的时变带通滤波方法及装置[P]. 中国专利, CN202010502784.3. 2025-04-09.
- [16] 杨柳, 郑国莘, 倪平哲, 等. 一种用于窄带信号的数字带通滤波器及滤波方法[P]. 中国专利, CN201210089898. 2025-04-09.
- [17] 刘艳霞. 基于小波变换的图像去噪算法研究[D]: [博士学位论文]. 天津: 天津大学, 2011.
- [18] 徐宏斌, 李庶林, 陈际经. 基于小波变换的大尺度岩体结构微震监测信号去噪方法研究[J]. 地震学报, 2012, 34(1): 85-96.
- [19] Aharon, M., Elad, M. and Bruckstein, A. (2006) K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE Transactions on Signal Processing*, **54**, 4311-4322. <https://doi.org/10.1109/tsp.2006.881199>
- [20] 林克正, 李殿璞. 基于小波变换的去噪方法[J]. 哈尔滨工程大学学报, 2000, 21(4): 21-23+25.
- [21] 王爱民, 袁昌旺, 霍耀璞, 等. 一种基于交叉验证的傅里叶变换去噪方法[P]. 中国专利, CN201711096062.7. 2025-04-09.
- [22] 司祯祯. 傅里叶变换与小波变换在信号去噪中的应用[J]. 电子设计工程, 2011, 19(4): 155-157.
- [23] 刘炜. 基于小波变换的遥感图像去噪与融合方法研究[D]: [博士学位论文]. 咸阳: 西北农林科技大学, 2008.
- [24] 杜鑫, 邓小凡, 杨晨莹. 基于匹配傅里叶变换插值规则化的叠前去噪[C]//国际勘探地球物理学家学会. CPS/SEG 北京 2018 国际地球物理会议暨展览电子论文集. 2018: 311-313.
- [25] 刘伟, 曹思远, 崔震. 稀疏准则约束的曲波变换去噪[C]//中国地球物理学会, 全国岩石学与地球动力学研讨会组委会, 等. 2014 年中国地球科学联合学术年会论文集. 2014: 36.