

## 基于分块字典学习理论的地震数据去噪

周俊捷, 吴相伶, 李文杰, 李静和

### Denoising of Seismic Data Based on Block Dictionary Learning Theory

ZHOU Junjie, WU Xiangling, LI Wenjie, and LI Jinghe

在线阅读 View online: <https://doi.org/10.15953/j.1004-4140.2022.31.05.03>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 地震资料宽频处理技术在天然气水合物矿体描述中的应用研究

Application of Broadband Seismic Data Processing Technology in the Description of Natural Gas Hydrate Ore Body  
CT理论与应用研究. 2021, 30(5): 547–554

#### 用于陆上宽频地震采集的设备及应用实例

Land Equipment for Broadband Seismic Data Acquisition and the Application  
CT理论与应用研究. 2019, 28(1): 99–110

#### 基于逆散射理论的介质参数与其地震响应的尺度对应关系研究

Research on the Scale Corresponding Relationship between the Medium Parameter and its Seismic Response Based on Inverse Scattering Theory  
CT理论与应用研究. 2018, 27(4): 423–432

#### 海洋拖缆地震资料初至波走时层析反演

First Arrive Wave Traveltime Tomographic Inversion of Marine Seismic Data  
CT理论与应用研究. 2018, 27(2): 197–204

#### 基于模糊粗糙集的机器学习储层参数预测

Prediction of Reservoir Parameters of Machine Learning Based on Fuzzy Rough Set  
CT理论与应用研究. 2018, 27(4): 455–464

#### 数据重构提高 $\tau-p$ 变换的精度

Data Reconstruction to Improve the Accuracy of  $\tau-p$  Transform  
CT理论与应用研究. 2018, 27(4): 465–476



关注微信公众号，获得更多资讯信息

周俊捷, 吴相伶, 李文杰, 等. 基于分块字典学习理论的地震数据去噪[J]. CT理论与应用研究, 2022, 31(5): 557-566.  
DOI:[10.15953/j.1004-4140.2022.31.05.03](https://doi.org/10.15953/j.1004-4140.2022.31.05.03).

ZHOU J J, WU X L, LI W J, et al. Denoising of seismic data based on block dictionary learning theory[J]. CT Theory and Applications, 2022, 31(5): 557-566. DOI:[10.15953/j.1004-4140.2022.31.05.03](https://doi.org/10.15953/j.1004-4140.2022.31.05.03). (in Chinese).

# 基于分块字典学习理论的地震数据去噪

周俊捷, 吴相伶<sup>✉</sup>, 李文杰, 李静和

桂林理工大学地球科学学院, 广西 桂林 541004

**摘要:** 随着油气勘探观测环境愈发复杂, 采集的地震数据常常掺杂各种噪声信号, 导致勘探目标引起的有效微弱信号被覆盖, 严重影响高精度的地震勘探数据解译, 因而有效的压制地震勘探数据噪声显得越发重要。本文采用字典学习策略, 将复杂地震数据进行分块, 通过分块数据的字典学习获取字典原子, 构建高精度的字典学习地震数据稀疏表示, 通过两次迭代更新字典原子, 进行数据去噪。将本文的字典学习算法应用于含随机噪声的模拟数据和实测地震勘探数据处理, 验证该算法的可行性和有效性。结果表明, 本文算法有效去除了随机噪声, 保留了有效信号同相轴, 提高了信噪比, 可为复杂含噪地震数据的去噪处理提供新的技术手段。

**关键词:** 分块字典学习; 地震数据去噪; 迭代更新; 实测数据处理

DOI:[10.15953/j.1004-4140.2022.31.05.03](https://doi.org/10.15953/j.1004-4140.2022.31.05.03) 中图分类号: O 242; P 315; P 631 文献标识码: A

地球物理勘探转战复杂环境区域, 使得勘探数据资料的质量要求也随之提高。由于地震勘探采集到的数据往往不仅包括目标体的信息, 同时含有各种各样的噪声干扰, 因而需要进行有效的数据去噪处理<sup>[1]</sup>。地震数据去噪处理主要包括原始采集数据加工、消除噪声影响, 提高目标资料的信噪比, 为后续处理和资料解释奠定基础。噪声的存在, 严重影响资料的准确度和保真性, 因而高效的去噪方法研究很有必要<sup>[2-3]</sup>。

传统的傅里叶变换、小波变换、曲波变换和F-X域反褶积滤波等方法广泛应用于地震勘探数据去噪领域<sup>[4]</sup>。但传统方法稀疏表示能力差, 不能有效反映地震数据的特征。字典学习(dictionary learning, DL)是跨学科研究领域的代表, 集合了稀疏表示、机器学习、图像应用和压缩感知等的理论思想, 主要用于解决稀疏表示模型的字典设计问题。贺月等<sup>[5]</sup>将其用于数据去噪、图像识别、数据重构等领域, 得到了出色的效果。字典学习包含固定基函数解析型字典和自适应学习型字典。后者可以不断学习训练更新字典, 提高字典的表达能力, 得到更稀疏的信号表示, 因此自适应学习型字典应用更为广泛。

基于非正交基理论, 分解信号数据的同时会得到稀疏系数矩阵, 提出了超完备字典方法作为字典学习理论基础。练秋生等<sup>[6]</sup>将匹配追踪方法作为稀疏表示方法延续技术, 取得了不错的进展; 随后Chen<sup>[7]</sup>提出了解决稀疏优化问题的基追踪方法, Olshausen等<sup>[8]</sup>提出了自适应学习型完备字典。自适应训练并更新字典, 使得大多系数等于零或者无限接近为零,

收稿日期: 2021-08-18。

基金项目: 广西自然科学基金(区域检测数据驱动下探地雷达隧道地质预报智能化解译技术研究与应用(2021GXNSFAA196056); 小波与曲波组合域探地雷达数据噪声与背景杂波压制技术(2018GXNSFAA281028))。

达到以更少的信息来反映样本特征的目的。为了更高效地训练得到自适应学习型字典, Aharon<sup>[9]</sup>提出了奇异值分解 (K-singular value decomposition, K-SVD) 算法。该方法基于奇异值分解方法对字典中的元素依次代替刷新, 最终获得所需的过完备字典, 且随着整体迭代的进行, 会同时更新系数矩阵与字典原子。压制噪声的原理为字典与稀疏表示系数重组结合, 因而自适应字典学习方法可以达到训练并不断更新字典和有效去噪同步进行的效果。

Tang 等<sup>[10]</sup>率先将学习型完备字典应用到地震勘探数据去噪中, 在对输入信号的不断学习训练后, 更新字典的同时得到稀疏表示系数, 并实现了优于传统方法的去噪效果, 但地震勘探数据的复杂性使得迭代次数要求更多, 以至于运算时间更长。Simon 等<sup>[11]</sup>提出一个基于变分稀疏表示模型的学习算法, 最优方向法 (method of optimal directions, MOD) 算法更新字典, 取得了一定去噪结果。周艳辉等<sup>[12]</sup>将基于 K-SVD 的范数稀疏字典学习方法与多震源混合地震记录分离方法结合, 证明了方法的可行性, 但相比固定基字典计算效率需要进一步加强。许德鑫<sup>[13]</sup>提出将 K-SVD 算法中的 MOD 算法换成 StOMP (stagewise orthogonal matching pursuit) 算法, 既克服正交匹配追踪 (orthogonal matching pursuit, OMP) 算法导致的过匹配现象, 又进一步使收敛速度明显提高。张成<sup>[14]</sup>提出了一种基于 K 均值序贯泛化算法字典的地震勘探数据去噪思路, 基于控制干扰的 Controlled-SGK 方法, 通过判定信噪比值是否大于设定的阈值来判断是否更新下一个原子, 以此来控制噪声干扰, 去噪质量得到大的提高。在压制随机噪声方面, 屈光中<sup>[15]</sup>提出了基于固定基字典的双稀疏学习型字典, 并首次成功应用于地震信号随机噪声的压制中, 该字典通过训练得到的稀疏表示系数结合固定基字典构成, 结构性与稀疏性比较高, 字典表达能力和去噪效果相比传统固定基字典学习均更有优势。张岩等<sup>[16]</sup>利用多道相似组稀疏表示方法压制随机噪声; 王惠迎等<sup>[17]</sup>提出自适应学习字典分类得压制方法; 王量等<sup>[18]</sup>提出在线字典学习算法进行地震数据去噪。

然而, 学习完整尺寸的数据字典原子是不容易计算的, 因为它需要大量的训练样本, 每个维度的原子都被学习。经过训练的字典集也太大, 无法在实践中使用, 因为超过的完整字典包含的原子比它们的维度还要多。分块理论的引入, 一定程度上可解决上述问题维度过大的问题。分块理论的应用较为成熟的领域为块匹配三维协同滤波, 该方法将数据域与频率域有效结合, 通过对观测数据进行分块, 采用块匹配和分组进行数据聚集, 可较好保留图像或数据的细节信息, 取得较好的去噪结果。然而, 块理论与字典学习的研究在地震数据处理领域应用较少, 因而开展分块字典学习算法地震勘探数据去噪研究具有重要的现实意义。

综上所述, 基于字典学习的跨学科去噪方法研究是当前研究热点, 新的稀疏编码方法和字典更新方法不断涌现, 适应范围也在扩大至三维地震数据去噪、重构和全波形反演、电法和电磁法数据去噪等方向。随着字典学习算法的研究, 将在地球物理勘探数据处理中发挥更大的作用。

## 1 方法理论

假设  $\mathbf{Y}$  为输入的时间域地震信号数据, 将时间域地震数据通过分块模块函数  $R$  进行分块, 因此, 基于分块理论, 字典学习地震勘探数据去噪问题可以分解为单个块数据的字典学习问题:

$$(\mathbf{X}, \mathbf{D}) \in \arg \min_{x, d} \sum_{i,j} \left( \frac{1}{2} \|R_{i,j}(Y) - d \cdot x_{i,j}\|^2 + \lambda \|x_{i,j}\|_1 \right), \quad (1)$$

其中,  $\mathbf{X}$  和  $\mathbf{D}$  为稀疏表示系数矩阵和计算的字典矩阵,  $x$  和  $d$  为各个块稀疏表示稀疏及字典原子,  $(i, j)$  的稀疏表示稀疏及字典的位置,  $\lambda$  为控制因子, 由试验确定取值。字典学习算法中更新稀疏表示的稀疏矩阵采用 OMP 方法, 更新字典则使用 MOD 方法。

### 1.1 正交匹配追踪 (OMP) 算法

OMP 算法的核心要求是找到输入信号与字典原子内积最大的原子, 定义其为匹配检索项, 从而除去与之相关的数据, 剩余部分再次与其他原子求内积值, 寻找结构上有一定相关性的原子, 直至剩余信号满足停止条件, 就结束迭代。内积表示其实就是反映输入信号和字典原子的相关性, 相关性越大, 则说明两者的内积越大, 结构越相似, 进一步表明两者一定程度上可以近似相等。以一维信号为例, 对于输入信号  $y \in R^{m \times 1}$ , 建立字典  $\mathbf{D} = (d_1, d_2, \dots, d_n)$ , 且满足  $\|d_i\|^2 = 1$ ,  $\mathbf{D}$  中的  $n$  个原子不需要正交。遍历字典中的每一个原子, 使其分别与信号做内积, 并选出最大内积对应原子  $d_1$ :

$$(y, d_1) > (y, d_i) \quad d_i \in \mathbf{D} (i \neq 1). \quad (2)$$

通过从信号  $y$  中减去其在  $d_1$  所张成空间上的正交投影得到残差  $e_1^{[14]}$ :

$$e_1 = y - d_1(y, d_1). \quad (3)$$

对残差重复继续以上两式, 当残差大小或者迭代次数满足条件时, 停止迭代。

其正交化所有当前选中原子, 继承了原子处理步骤的规则, 从而迭代的最优性得到保留。同样以一维信号为例, 信号经  $k$  步分解后, 当前选择出的原子用  $\hat{d}_l (l = 1, 2, \dots, k)$  表示, 信号  $y$  可以表示为:

$$y = \sum_{l=1}^k x_l \hat{d}_l + e_k, \quad (e_k, \hat{d}_l) = 0 (l = 1, 2, \dots, k-1). \quad (4)$$

在进行  $k+1$  步迭代时, 从剩余原子中选择出与  $e_k$  内积最大的原子  $d_{k+1}$ , 将  $d_{k+1}$  与之前选出的  $k$  个原子做正交化处理:

$$\hat{d}_{k+1} = d_{k+1} - \sum_{i=1}^k \frac{(d_{k+1}, \hat{d}_i)}{(\hat{d}_i, \hat{d}_i)} \hat{d}_i. \quad (5)$$

残差  $e_k$  投影到正交化后的  $\hat{d}_{k+1}$  上, 可得:

$$e_{k+1} = e_k - \frac{(e_k, \hat{d}_{k+1})}{(\hat{d}_{k+1}, \hat{d}_{k+1})} \hat{d}_{k+1}. \quad (6)$$

在  $k$  次迭代后, 信号表示可约为:

$$y = \sum_{i=1}^k \frac{(e_i, \hat{d}_i)}{(\hat{d}_i, \hat{d}_i)} \hat{d}_i. \quad (7)$$

### 1.2 最优方向 (MOD) 算法

最优方向 (MOD) 算法把字典学习问题作为嵌套的最小化问题: 内层看成给定字典的情

况下关于非零项稀疏表示系数  $\mathbf{X}_j$  个数的最小化问题，外层是定义在字典  $\mathbf{D}$  之上的最小化问题。这种交替最小化的思路将第  $k-1$  步得到的字典  $\mathbf{D}_{(k-1)}$  来求解第  $k$  步的稀疏系数，可以得到矩阵  $\mathbf{X}_{(k)}$ ，最后用最小二乘求解  $\mathbf{D}_{(k)}$ <sup>[19]</sup>，即

$$\mathbf{D}_{(k)} = \operatorname{argmin}_{\mathbf{D}} \left\| \mathbf{Y} - \mathbf{DX}_{(k)} \right\|_2^2 = \mathbf{YX}_{(k)}^T \left( \mathbf{X}_{(k)} \mathbf{X}_{(k)}^T \right)^{-1} = \mathbf{YX}_{(k)}^+。 \quad (8)$$

MOD 算法的详细描述如下所示：

**Input:** 给出用于训练的数据，稀疏度  $L$

**Output:** 重构数据字典  $\mathbf{D}$

1). 初始化：令  $k=0$ ，用  $M$  个随机选择的样本来构造字典  $\mathbf{A}_{(0)} \in \mathbb{R}^{n \times M}$ ，同时各列归一化处理；

2). 迭代： $k$  每次加1，并按照一下步骤执行；

3). 稀疏编码：用追踪算法近似求解

$$\hat{x}_j = \operatorname{argmin}_x \left\| \mathbf{y}_j - \mathbf{D}_{(k-1)} x \right\|_2^2 \text{ s.t. } \|x\|_0 \leq L, \text{ 得出稀疏系数 } \hat{x}_j, \text{ 组构成矩阵 } \mathbf{X}_{(k)};$$

4). 字典更新：用公式 (8) 更新字典；

5). 达标停止迭代：如果  $\mathbf{Y} - \mathbf{AX}_{(k)}^2$  的变化量足够的小则停止迭代，输出结果为  $\mathbf{A}_{(k)}$ ；否则继续转到步骤 2。

对于给定的信号集，MOD 算法既能得到冗余字典，又能得到稀疏表示矩阵，且表示误差最小，但其优化结果是一个高度非凸的组合问题，能得到局部最优解。一般来说，MOD 算法只需要少量的迭代就可以实现收敛，算法非常高效。结合 OMP 算法与 MOD 算法，本文求解字典学习问题流程如下：

(1) 归一化原始数据  $\mathbf{M}_0$ ，加高斯随机噪声后得含噪数据  $\mathbf{M}$ ；

(2) 用 OMP 思想对原始数据  $\mathbf{M}_0$  编码得稀疏系数  $\mathbf{X}$ ，用 MOD 思想更新字典  $\mathbf{D}_1$ ，重构计算信噪比；

(3) 用 OMP 思想对含噪数据  $\mathbf{M}$  编码得稀疏系数  $\mathbf{X}$ ，结合上一步的字典  $\mathbf{D}_1$  重构得到去噪后的数据  $\mathbf{M}'$ ；

(4) 重复第 (2) 和第 (3) 两步，即完成两次字典学习。

## 2 算例分析

### 2.1 数值模拟算例

数值模拟算例为合成地震勘探数据，主要包含水平及倾斜地震波同相轴，同时包含复杂的同相轴相交的情况，如图 1 (a) 所示 (Data 1)。图 1 (b) 所示为按照  $5 \times 5$  矩阵进行分块示意图，可见，分块对原始数据各个不同特征分布均有所涵盖。下文算例中均采用相同分块模式，更新迭代 100 次，不再图示及概述。对合成模拟数据加入 20% (噪声幅值与信号幅值的比值) 随机噪声，如图 1 (b) 所示，噪声遍布数据整个范围，地震波同相轴被噪声淹没。去噪结果表明 (图 1 (c))，字典学习数据去噪的效果是可行的，背景噪声滤除较为干净，有效成分被较好保留。图 2 为上述图 1 (a) 所示相同地震波同相轴成分，但不同分布情况 (Data 2)。对 Data 2 数据加入相同的 20% 随机噪声，去噪结果可见，不同类型的同相轴均有效保留，噪声消除效果较好，显示了本文算法的可行性和有效性。

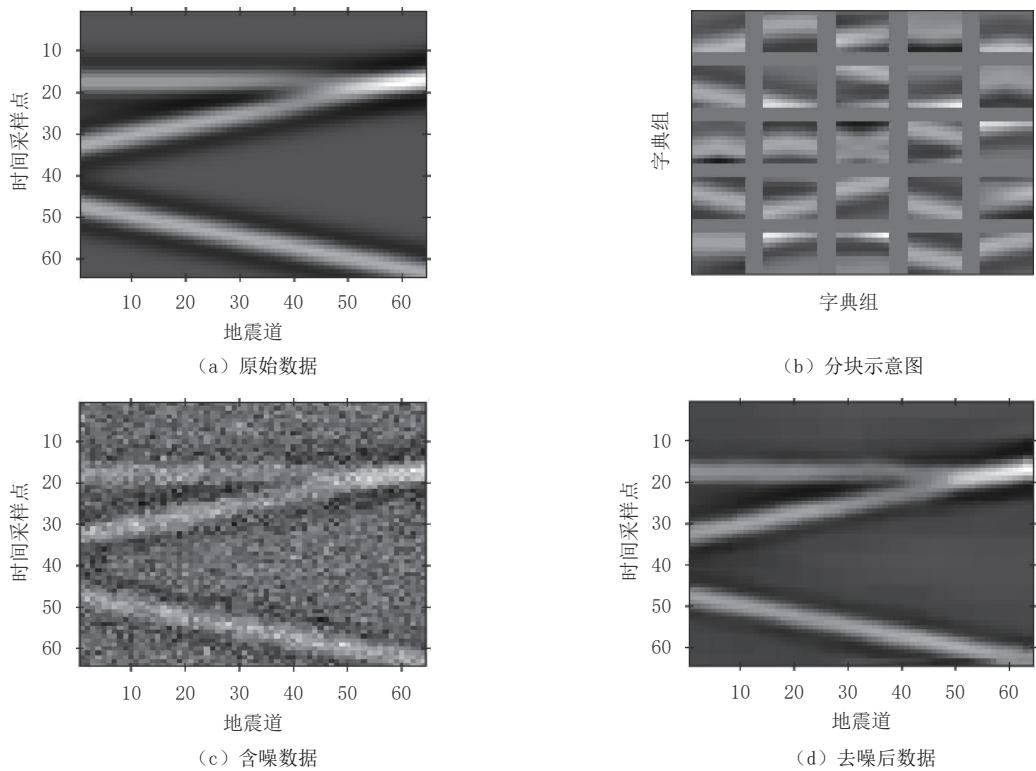


图 1 Data 1 合成数据字典学习去噪结果

Fig.1 Data 1 synthesis data dictionary learning denoising results

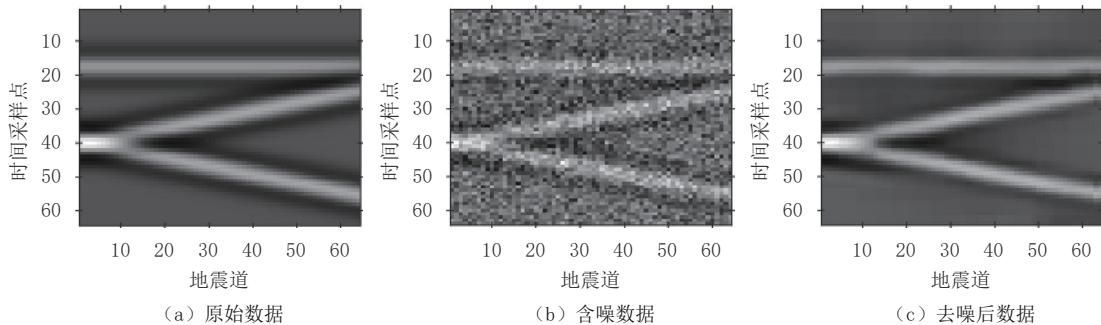


图 2 Data 2 合成数据字典学习去噪结果

Fig. 2 Data 2 synthetic data dictionary learning denoising results

图 3 是 Data 3 合成数据字典学习去噪结果，与前述两个合成算例不同，Data 3 数据为单独水平同相轴和相交同相轴地震波形。结合前述图 1 和图 2 所示结果分析，Data 3 数据分布较为规则（图 3 (a)），但交叉型地震波同相轴位于数据集中部，其两侧交叉位置同相轴较为复杂。含噪数据则使得上述交叉部分的同相轴更为复杂（图 3 (b)）。从图 3 (c) 去噪结果可见，去噪后同相轴分布与 Data 3 原始数据同相轴分布一致，复杂同相轴蕴含更多信息被有效保留。当交叉型地震波不对称时，字典学习过程分块对不对称交叉型同相轴的选取亦会不一样，其去噪的难度增加，如图 4 (a) 所示。当加入 20% 随机噪声，如图 4 (b) 所示，含噪数据在左侧不对称交叉型同相轴更加复杂化。图 4 (c) 去噪结果可见，

字典学习有效恢复了不对称交叉型同相轴分布，去噪效果良好，再次显示了该算法的可行性和有效性。

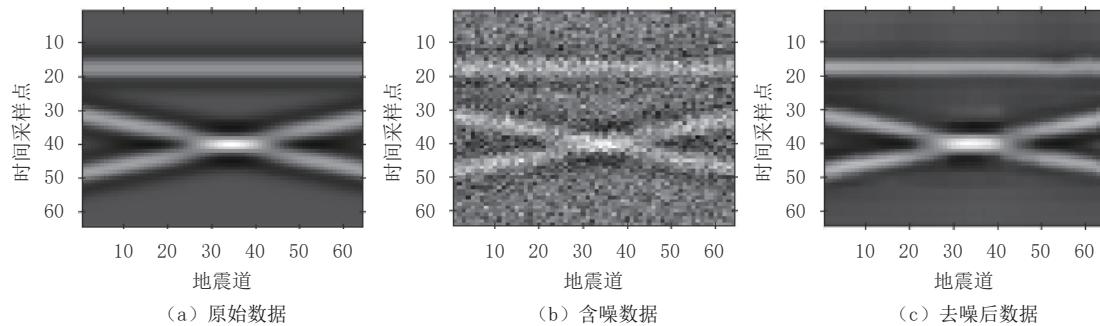


图 3 Data 3 合成数据字典学习去噪结果

Fig. 3 Data 3 synthetic data dictionary learning denoising results

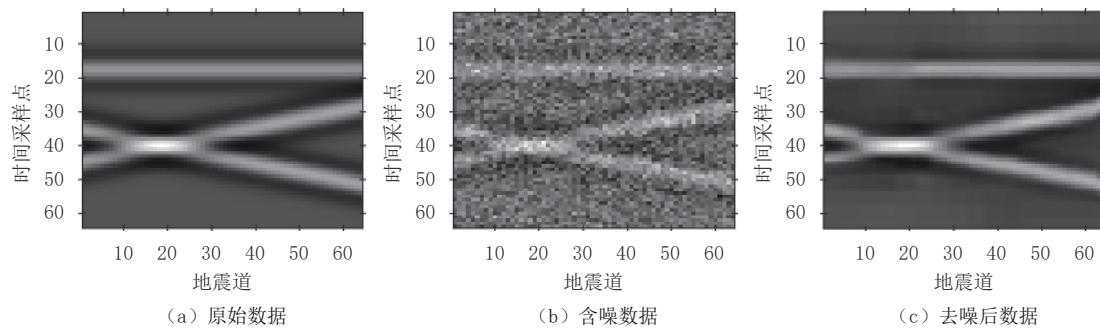


图 4 Data 4 合成数据字典学习去噪结果

Fig. 4 Data 4 synthetic data dictionary learning denoising results

表 1 所示为上述 Data 1–Data 4 合成地震数据及其对应去噪结果按照标准信噪比（单位 dB）与运行时间对比。由表格结果可见，加入 20% 随机噪声地震数据信噪比较低，均小于 4 dB；去噪后，4 个合成数据算例信噪比得到大大提高，最高算例提高约 15 dB。运行时间方面，针对上述 4 个算例按  $5 \times 5$  分块策略及  $64 \times 64$  数据集，在普通电脑（RAM 8 G, i7-CPU, 2.6 GHz）上需要耗时小于 35 s。可见，基于分块理论字典学习算法计算效率是可行的。

表 1 不同模拟数据去噪前后信噪比与运行时间

Table 1 SNR and running time of different analog data before and after denoising

类型	去噪前信噪比/dB	去噪后信噪比/dB	运行时间/s
Data 1	3.539	17.436	30.2
Data 2	2.909	18.553	26.1
Data 3	3.089	18.326	28.2
Data 4	2.883	16.618	32.9

## 2.2 实测数据算例

实测数据算例 1 如图 5 (a) 所示，地震波同相轴主要以水平状为主，从采集时间早期到晚期，地震波同相轴由强幅值变化为弱幅值。由于强幅值同相轴与弱幅值同相轴在字典

学习过程获得的原子各异,常规字典学习算法在处理此类地震数据过程中存在误处理问题。通过对原始数据进行 $5 \times 5$ 分块,选取了强弱不同幅值的同相轴均匀分布,保证了不同幅值同相轴在学习过程中出现的概率,因而可一定程度上改善上述误处理问题。如图5(b)所示,加入20%随机噪声作为干扰成分,可见,噪声的存在使得实测数据更为复杂。图5(c)所示为本文方法去噪结果,采集时间早期和晚期对应不同幅值同相轴得到有效恢复和凸显,噪声消除较为干净,验证了本文算法对不同幅值同相轴同时存在的地震数据去噪领域的可行性和有效性。

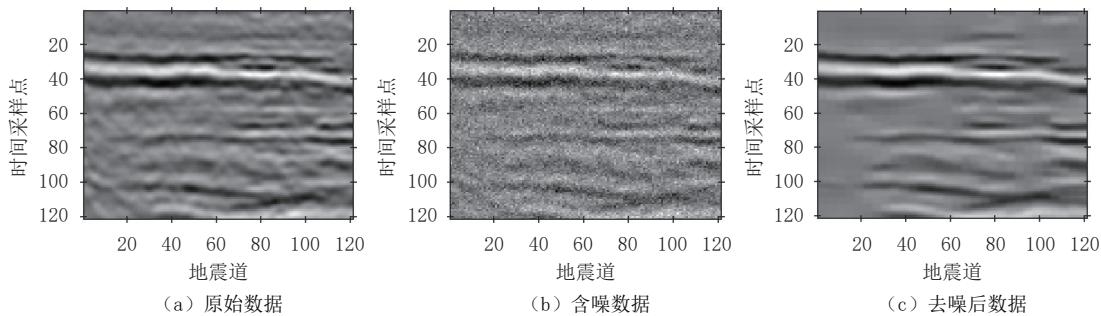


图5 实测数据算例1字典学习去噪结果

Fig. 5 Denoising results of dictionary learning for example 1 of measured data

图6所示为实测数据算例2字典学习去噪结果。图6(a)为原始数据,可见,算例2中地震波同相轴多以倾斜类型为主,部分区域掺杂了其他微弱杂波。加入20%随时噪声后,如图6(b)所示,仅可模糊分辨部分较强幅值同相轴,其他同相轴均被随机噪声覆盖,使得原始数据分布更为复杂。图6(c)为字典去噪结果,相对较原始数据分布,去噪结果较好的恢复了倾斜的地震波同相轴分布,其中,绝大部分噪声能得到压制,对于弱幅值地震波同相轴保护较好,波形同相轴质量得到有效保障。

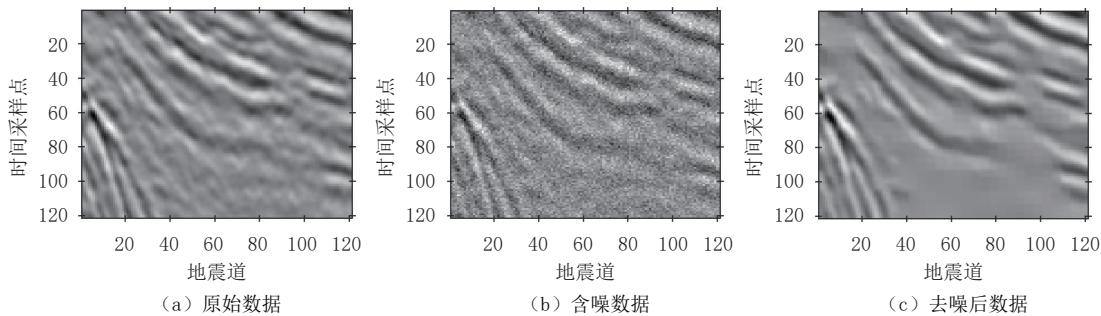


图6 实测数据算例2字典学习去噪结果

Fig. 6 Denoising results of dictionary learning for example 2 of measured data

图7所示为实测数据算例3字典学习与其他方法去噪结果对比图。图7(a)为原始实测地震数据,可见,原始数据包含了水平、倾斜及弯曲类型地震波同相轴,同时不同采集时间段内出现了同相轴断缺、强弱幅值同相轴现象,是较为复杂的地震数据。通过采用本文 $5 \times 5$ 分块策略,得到了分块字典学习示意图(图7(b)),分块图中囊括了原始数据大部分特征区域,为后续分块字典学习去噪奠定了基础。同样,对原始数据加入20%随机噪

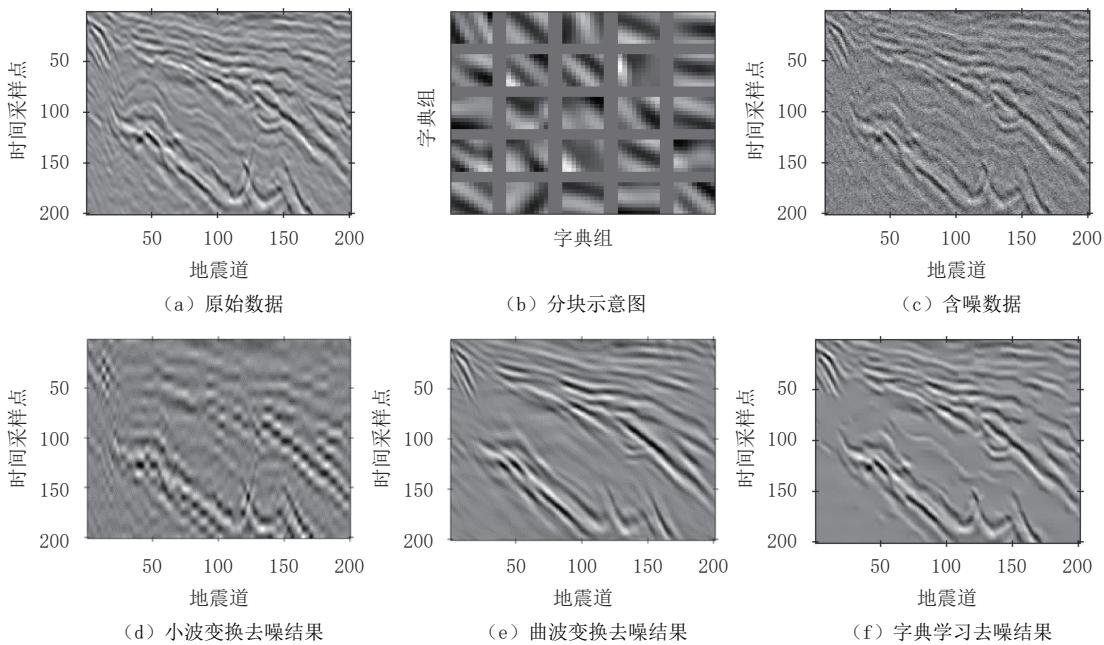


图 7 实测算例 3 字典学习去噪结果

Fig. 7 Denoising results of dictionary learning for example 3 of measured data

声, 如图 7 (c) 所示, 噪声的存在使得原始地震波形的分布更为复杂, 若不进行有效去噪, 无法对有效波形进行正确识别。为了进行有效性对比, 本文引入小波变换及曲波变换去噪算法进行对比分析, 其中, 小波变换采用 Sym8 小波基函数, 分解层数为 3 层; 曲波变换采用 L2 标准差作为阈值函数。图 7 (d) 和图 7 (e) 分别为小波变换和曲波变换算法去噪结果, 可见, 总体上随机噪声基本得到清除。但由于小波变换采用矩形单元窗口, 导致倾斜特征地震波同相轴出现断续波形现象; 而曲波变换由于计算的标准差阈值不能完全适用强弱幅值波形同相轴, 出现了曲波变换倾斜矩形单元窗口拉长同相轴现象。图 7 (f) 为本文字典学习算法去噪结果, 相比之下, 由于分块策略选择了各种不同特征波形同相轴, 因而去噪结果可有效保留不同强弱幅值、不同类型的地震波同相轴。整体上, 字典学习算法去噪效果由于小波变换或曲波变换算法的去噪效果, 验证了本文算法可用于复杂地震波数据的去噪领域。

整体上, 字典学习算法去噪效果由于小波变换或曲波变换算法的去噪效果, 验证了本文算法可用于复杂地震波数据的去噪领域。

### 3 结论

针对含随机噪声地震数据处理问题, 本文采用基于分块理论的字典学习算法开展了数值模拟算例及实测地震数据算例去噪研究。首先详细介绍了基于分块理论字典学习地震数据去噪算法的基本原理及实现流程。通过对水平及倾斜地震波同相轴、交叉型地震波同相轴合成数据算例加入随机噪声, 采用本文去噪方法进行数据恢复, 获得较好的去噪效果, 各算例的信噪比得到大幅度提升, 验证了本文算法的可行性。最后, 将本文方法用于加入随机噪声的实测数据算例去噪处理, 同时将去噪结果与小波变换、曲波变换算法去噪结果

对比,验证了本文算法用于实测数据的有效性。

## 参考文献

- [1] 王高成, 刘浩杰, 杨宏伟, 等. 一种有效的转换波去噪方法[J]. 地球物理学进展, 2021, 36(6): 2588-2598.  
WANG G C, LIU H J, YANG H W, et al. One good denoising method of converted wave[J]. *Progress in Geophysics*, 2021, 36(6): 2588-2598. (in Chinese).
- [2] 姜占东, 范彩伟, 黎孝璋, 等. 多道奇异谱分析压制强振幅窄带随机噪声研究[J]. 地球物理学进展, 2021, 36(2): 730-735.  
JIANG Z D, FAN C W, LI X Z, et al. Study on suppressing strong amplitude narrow band random noise based on multichannel singular spectrum analysis[J]. *Progress in Geophysics*, 2021, 36(2): 730-735. (in Chinese).
- [3] 高好天, 孙宁娜, 孙可奕, 等. DnCNN和U-Net对地震随机噪声压制的对比分析[J]. 地球物理学进展, 2021, 36(6): 2441-2453.  
GAO H T, SUN N N, SUN K Y, et al. Comparative analysis of DnCNN and U-Net on suppression of seismic random noise[J]. *Progress in Geophysics*, 2021, 36(6): 2441-2453. (in Chinese).
- [4] 李光辉, 张子豪, 李佳蓉. 基于自适应小波神经网络在压制地震勘探低频噪声中的应用[J]. 测试技术学报, 2021, 35(3): 185-189, 204. DOI:10.3969/j.issn.1671-7449.2021.03.001.  
LI G H, ZHANG Z H, LI J R. An applied on low-frequency noise suppression of seismic exploration based on adaptive WNN[J]. *Journal of Test and Measurement Technology*, 2021, 35(3): 185-189, 204. DOI:10.3969/j.issn.1671-7449.2021.03.001. (in Chinese).
- [5] 贺月, 柳建新, 王显莹, 等. 基于自适应字典学习的插值去噪的应用[J]. 地球物理学进展, 2021, 36(6): 2454-2461.  
HE Y, LIU J X, WANG X Y, et al. Application of interpolation denoising based on adaptive dictionary learning[J]. *Progress in Geophysics*, 2021, 36(6): 2454-2461. (in Chinese).
- [6] 练秋生, 石保顺, 陈书贞. 字典学习模型、算法及其应用研究进展[J]. 自动化学报, 2015, 41(2): 240-260.  
LIAN Q S, SHI B S, CHEN S Z. Research advances on dictionary learning models, algorithms and applications[J]. *Acta Automatica Sinica*, 2015, 41(2): 240-260. (in Chinese).
- [7] CHEN S S, DONOHO D L, SAUNDERS M A. Atomic decomposition by basis pursuit[J]. *SIAM Review*, 2001, 43(1): 129-159.
- [8] OLSHAUSEN B A, MILLMAN K J. Learning sparse overcomplete image representations[C]// Wavelet Applications in Signal and Image Processing VIII, International Society for Optics and Photonics. 2000, 4119: 445-452.
- [9] AHARON M, ELAD M, BRUCKSTEIN A. The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation[J]. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311-4322.
- [10] TANG G, MA J W, YANG H Z. Seismic data denoising based on learning-type overcomplete dictionaries[J]. *Applied Geophysics*, 2012, 9(1): 27-32.
- [11] SIMON B, MA J W. Simultaneous dictionary learning and denoising for seismic data[J]. *Geophysics*, 2014, 79(3): A27-A31.
- [12] 周艳辉, 高静怀, 陈文超. 基于稀疏表示的多震源混合地震记录分离方法研究[C]//中国石油学会物探技术研讨会论文集, 2015: 165-168.
- [13] 许德鑫. 基于稀疏表示和字典学习的地震数据去噪研究[D]. 长春: 吉林大学, 2016.  
XU D X. Research on seismic denoising based on the sparse representation and dictionary learning[D]. Changchun: Jilin University, 2016. (in Chinese).
- [14] 张成. 基于K均值序贯泛化(SGK)的二维地震信号重建与去噪研究[D]. 天津: 河北工业大学, 2016.  
ZHANG C. Study on reconstruction and denoising of two-demensional seismic signal based on sequential generalization of K-means[D]. Tianjin: Hebei University of Technology, 2016. (in Chinese).
- [15] 屈光中. 基于稀疏表示的地震信号随机噪声压制与面波分离[D]. 合肥: 合肥工业大学, 2016.

- QU G Z. The suppression of random noise and separation of ground roll in seismic signals based on sparse representation[D]. Hefei: Hefei University of Technology, 2016. (in Chinese).
- [16] 张岩, 任伟建, 唐国维. 利用多道相似组稀疏表示方法压制随机噪声[J]. 石油地球物理勘探, 2017, 52(3): 442–450. DOI:[10.13810/j.cnki.issn.1000-7210.2017.03.005](https://doi.org/10.13810/j.cnki.issn.1000-7210.2017.03.005).
- ZHANG Y, REN W J, TANG G W. Random noise suppression based on sparse representation of multi-trace similarity group[J]. Oil Geophysical Prospecting, 2017, 52(3): 442–450. DOI:[10.13810/j.cnki.issn.1000-7210.2017.03.005](https://doi.org/10.13810/j.cnki.issn.1000-7210.2017.03.005). (in Chinese).
- [17] 王惠迎, 陈昕, 陈建友, 等. 基于自适应学习字典分类的地震数据谐波噪声压制方法[C]//国际地球物理会议暨展览电子论文集, 2018: 438–441.
- [18] 王量, 买皓, 李勇. 基于在线字典学习算法的地震数据去噪应用[J]. 断块油气田, 2019, 26(2): 177–180. DOI:[10.6056/dkyqt201902009](https://doi.org/10.6056/dkyqt201902009).
- WANG L, MAI H, LI Y. Application of seismic data denoising based on online dictionary learning algorithm[J]. Fault-Block & Gas Field, 2019, 26(2): 177–180. DOI:[10.6056/dkyqt201902009](https://doi.org/10.6056/dkyqt201902009). (in Chinese).
- [19] 秦娜. 基于稀疏表示的航空磁测数据处理技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2016.
- QIN N. Research on aeromagnetic data processing based on sparse representation[D]. Harbin: Harbin Institute of Technology, 2016. (in Chinese).

## Denoising of Seismic Data Based on Block Dictionary Learning Theory

ZHOU Junjie, WU Xiangling<sup>✉</sup>, LI Wenjie, LI Jinghe

College of Earth Sciences, Guilin University of Technology, Guilin 541004, China

**Abstract:** With the increasingly complex observation environment of oil and gas exploration, the seismic data collected are often mixed with various noise signals, resulting in the effective weak signal caused by the exploration target is covered, which seriously affects the high-precision seismic data interpretation, so it is more and more important to effectively suppress the seismic data noise. In this paper, the dictionary learning strategy is used to block the complex seismic data. The dictionary atoms are obtained through the dictionary learning of the block data, and the sparse representation of the seismic data is constructed by high-precision dictionary learning. The dictionary atoms are updated through two iterations for data denoising. The dictionary learning algorithm is applied to the processing of simulated and measured seismic data with random noise. The analysis results show that the algorithm can effectively removes the random noise while retains the effective signal phase axis, improves the signal-to-noise ratio which verifies the feasibility and effectiveness of the algorithm. The research results provide a new technical means for complex noisy seismic data denoising.

**Keywords:** block dictionary learning; seismic data denoising; iterative updating; processing of measured data



**作者简介:** 周俊捷, 男, 桂林理工大学地球科学学院勘查技术与工程专业在读本科生, 主要从事地球物理数据处理与解释研究工作, E-mail: [843558604@qq.com](mailto:843558604@qq.com); 吴相伶<sup>✉</sup>, 男, 桂林理工大学硕士研究生, 主要从事地球物理正反演、数据处理与解释研究, E-mail: [wxl@glut.edu.cn](mailto:wxl@glut.edu.cn)。