

文章编号: 2096-1618(2021)01-0007-08

# MCA 框架下 K-SVD 构建字典对地震信号去噪与重建

陈思琪<sup>1</sup>, 薛雅娟<sup>1</sup>, 杨清蜜<sup>2</sup>, 方立鑫<sup>2</sup>, 郑书琳<sup>2</sup>

(1. 成都信息工程大学通信工程学院, 四川 成都 610225; 2. 中国石化地球物理重点实验室, 江苏 南京, 210000)

**摘要:**针对由于地震信号采集环境的复杂性带来的采样信号不完整,存在大量噪声等情况,对采集的叠前地震信号进行去噪和重建。在传统的 K-奇异值分解(K-singular value decomposition, K-SVD)建造冗余字典的压缩感知重建的基础上,提出了基于形态分量分析(morphometric principal components analysis, MCA)的 K-SVD 地震信号的去噪与重建。即使用 MCA 对地震信号的结构和平滑部分进行分类,并针对上述两种类别分别构建由 K-SVD 算法计算的冗余字典,将两种类别分别置于不同字典中进行去噪与重建。与传统的方法相比,该方法在减少了地震信号采集的成本和难度的基础上,精确辨别地震信号细节,并取得良好的去噪效果。

**关键词:**地震信号;压缩感知;K-奇异值分解;形态分量分析;冗余字典

**中图分类号:** TN911.4

**文献标志码:** A

**doi:** 10.16836/j.cnki.jcuit.2021.01.002

## 0 引言

地震信号采集时,由于周围环境的复杂多变性,采集的地震信号有不规则和不完整等情况。降低地震信号的采样质量,不利于地震信号的分析研究。在实际工作中,对不完整的地震信号进行插值采样重建<sup>[1]</sup>,在一定程度上解决了该项问题。然而由于对地震信号的采集精度要求提升,该地震信号重建存在大量不足。寻求更好的地震信号重建方式,对地震信号的分析有重大意义。

现有的地震信号重建技术,可归纳为三大类。第一类为滤波器重建技术,即需要低通滤波器把周期拖延的高频部分过滤,达到恢复原有信号的目的,如 Spitz-S 提出的 F-X 预测误差滤波算法<sup>[2]</sup>。由于此类算法是通过高斯窗进行插值计算,不确定因素较大。第二类方法是算子迭代的重建方法,即通过共倾角时差矫正(dip move out, DMO)等正演、反演算子进行迭代求解。DMO 与逆 DMO 叠加法,此类方法能处理不均匀采样的数据。需先采集地下结构的先验信息,对计算要求极大,处理较费时<sup>[3]</sup>。第三类是目前比较广泛使用的变换法,即通过变换将地震数据放在不同的变换域处理。如傅里叶变换重建<sup>[4]</sup>, Curvelet 曲波变换重建, Radon 变换重建<sup>[5]</sup>。虽然该类方法可以处理不均匀采样,但受空间带宽影响对采样的要求很高。

近年来,发展起来的压缩感知技术(compressed

sensing, CS)<sup>[6]</sup>,与传统的奈奎斯特采样要求不同,降低了采样要求进而减少了采样的难度和时间。该技术可以通过较少的数据恢复满足一定程度的原始数据。由于该技术需要处理的数据必须满足稀疏条件,而地震信号往往不满足。研究者们采用某种变换或完备字典的形式满足信号的稀疏要求。由于学习型完备字典<sup>[7-8]</sup>能够自适应的提取样本信号数据,更优化于某种正交变换,常常采用两者相结合的方式,即通过某种正交变换建立学习型字典的初始字典,再通过 K-奇异值分解(K-singular value decomposition, K-SVD)等方法的训练建立超完备冗余字典<sup>[9-10]</sup>。但由于在进行字典训练时,缺少对地震信号特征的考虑,对地震信号的纹理部分和结构部分没有进行分块处理,将使得其不能更好地稀疏地表达采样信号。

针对上述问题,提出基于形态成分分析(morphometric principal components analysis, MCA)的 K-SVD 地震信号的去噪与重建。首先将地震信号分为纹理部分和结构部分。由于地震信号的纹理和结构分别对应地震信号的高频和低频部分,采用聚合经验模态分解(empirical mode decomposition, EMA),将信号的低频和低频信号分开<sup>[11]</sup>。将大量高频地震信号样本和低频地震信号样本分别单独训练,形成属于该部分的完备字典。通过 K-SVD 完善训练字典,构成优化完备冗余字典。最后将每部分分别通过对应的优化字典达到信号稀疏的目的。该方法可以更稀疏的表示原始地震信号,更好的恢复和重建地震信号。最后通过正则化正交匹配追踪算法进行地震信号的重建。

# 1 MCA 框架下基于 K-SVD 构建字典

## 1.1 信号的稀疏表示

信号稀疏即信号的绝大部分元素为 0。传统的奈奎斯特采样由于对信号采集的频率有一定的要求,往往在实际自然信号的采集中如地震信号存在太多局限性。该采集方法不利于应对自然信号采集面临的复杂多变环境。面对该问题,压缩感知理论成为地震信号采集以及重建整理的新方向。由于该方法要求信号是稀疏的,而自然信号往往不符合条件。信号的稀疏表示尤为重要。

信号往往可以看成是几个线性原子的叠加,将大量原子聚集在同一处的集合称之为字典。信号的稀疏表示即通过在字典中寻找能组成的自身的线性原子,进而达到稀疏的目的。当字典的列数大于行数时,称其为冗余字典。该字典能更好地捕捉信号的原始特征,更有利于进行信号的重建算法<sup>[12]</sup>。

## 1.2 K-SVD 算法构造冗余字典

K-SVD 算法是由 k-means 算法的一种泛化形式。与 k-means 算法中信号量只能由一种原子表示不同,K-SVD 算法是用多种原子的组合表示。其原理是通过某种变换构建初始字典,在其基础上使用 SVD 算法进行字典的更新以及稀疏表示。由该算法构建的超完备学习字典能够适应多种不同类型的信号并能更好地稀疏信号<sup>[13]</sup>。信号往往可以在合适的变化域中稀疏化。

给定训练数据由于很难一次构建最优字典,需要通过逐步逼近的方法寻找最优解。K-SVD 可通过以下 3 步实现:

(1)初始字典构造。传统方式一般采用随机挑选  $k$  个样本数据构成,文中将针对不同地震信号类型来寻求不同的初始字典。

(2)稀疏表示。用第一步得到的初始字典  $D$  对样本数据  $Y$  进行稀疏表达,公式为

$$D, X = \arg \min_{D, X} \{ \|x\|_0 \} \quad \text{st.} \quad \|Y - DX\|^2 \leq \varepsilon \quad (1)$$

$X$  为稀疏后的数据; $\varepsilon$  为重构误差的最大范围。

(3)更新字典。对更新字典的方法 K-SVD 采取逐步更新。即每次只更新一列原子,在更新该列原子时其他原子保持固定不动。其目标函数为

$$\|Y - DX\|^2 = \|y - \sum_{k=1}^K a_k \cdot x_k\|^2 = \|E_k - a_k \cdot x_k\|^2 \quad (2)$$

其中: $Y$  为列原子, $a$  为字典数据, $x$  为更新数据通过上述目标函数的迭代更新,达到 K-SVD 字典更新的过程。

## 1.3 形态分量分析

MCA 是 Starck 等于 2004 年提出的结合变分方法和稀疏表示将信号分离成几种不同形态特征的方法,广泛用于二维信号的去噪、重建、高分辨等,为二维信号的处理提供了更加良好的处理机制<sup>[14]</sup>。该方法认为一种信号是由不同的线性分量组成,通过信号之间不同成分的差异可将信号视为一系列不同形态特征组成的。即给定二维信号  $M$ ,该信号可由  $K$  个形态分量所组成。

$$M = \sum_{k=1}^K M_k \quad (3)$$

式中,每一个  $M_k$  代表着该二维信号的一个形态分量。每一个二维分量分别有一个专属的字典表示其稀疏分量。选取不同字典的结合,能更好的将二维信号的结构和纹理细分。达到算法更加精细,计算更加简单的目的。

## 1.4 MCA 框架下字典的设计

MCA 框架下的字典设计可以通过图 1 显示,即将大量样本信号分解成多个形态分量,并针对不同的形态分量进行 K-SVD 算法下的冗余字典构建。初始字典的不同会造成冗余字典的效果不同,初始字典包括 DCT 字典、曲波字典、局部脊波字典等。MCA 框架设计主要内容依赖于初始字典的选择,不同初始字典的选择造成的重建效果不同,找出字典选择的最优比是关键。针对地震信号特点,主要采样 DCT 字典和 Curvelet 字典作为初始字典进行字典设计。

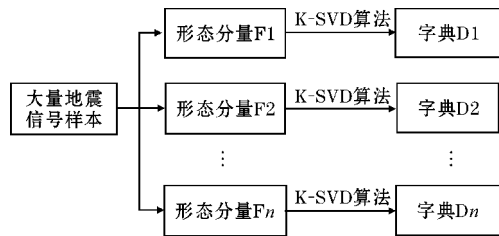


图1 MCA 框架下的字典构建

# 2 基于 MCA 的 K-SVD 地震信号去噪与重建

传统的 K-SVD 算法能较好地适用在地震信号的去噪与重建中,但依旧存在一些不足。由于传统 K-SVD 算法未能将地震信号的结构和细节划分开,缺少对细节的考虑。再加上由于噪声信号影响,进入训练字典中会干扰地震信号的重建。

为解决上述问题,提出基于 MCA 的 K-SVD 字典的训练。通过对地震信号进行分类处理,将地震信号划分为纹理部分和结构部分,针对不同的部分,通过不同的初始字典用 K-SVD 算法进行字典训练,从而达到细节部分的优化,并能更好地恢复地震信号并起到更强的去噪作用。

## 2.1 形态成分分析的地震信号处理

地震信号能通过高低频划分为不同部分。通过对地震信号的高低频读取,将地震信号划分为结构和纹理部分。采用聚合经验模态分解化为地震信号的高低频。由于该方法可以将地震信号转变为由高到低的一系列频率分步,可以很好地将地震信号进行细节分类。通过将高低频率分别分类,形成地震信号的结构部分(高频部分)以及纹理部分(低频部分)。将不同的部分通过不同的字典培训<sup>[15-16]</sup>。

## 2.2 字典选择

根据地震信号的实际情况,地震信号通过细节可分两类。即高频部分组成的奇异值信号及低频部分组成的平滑部分。分别将 DCT 字典和 curvelet 字典作为培训的初始字典,通过对大量地震信号奇异值的样本和大量平滑地震信号样本训练出不同的字典。将需重建地震样本不同部分通过所训练的不同字典。

## 2.3 地震信号重建

对地震数据重建算法采用的是正则化正交匹配追踪算法(regularized orthogonal matching pursuit, ROMP)和正交化匹配算法(orthogonal matching pursuit, OMP)一次只能选出与残差内积绝对值相差最大的一个原子的不同量,该方法能选出满足上述条件的多列原子。具体重建方法如下:

(1)初始化。将残差分量  $r_0$  置为 0,迭代次数  $\tau$  设为 1;

(2)选定相关原子。通过传感矩阵  $A = \varphi \cdot \varphi$ ,即稀疏信号与测量矩阵的乘积矩阵,计算残差内积绝对值  $u = \text{abs}[A^T r_{i-1}]$ ,并将  $u$  中  $k$  个最大值(或者所有非零值)与对应  $A$  的列序号构成集合  $J_0$ 。

(3)正则化找寻索引。在集合  $J$  中寻找满足  $|u_i| \leq 2|u_j|, j \in J_0$  的子集  $J_0$ ,选取具有最大能量  $\sum_j |u_j|^2, j \in J_0$  的子集  $J_0$ 。

(4)放置集合。将选择出的  $J_0$  放入集合  $A_\tau, \tau$  为迭代次数,选去除  $A$  中相对应的列向量  $a_j, j \in J_0$  放入集合  $A_\tau$ 。

(5)更新残差。残差分量  $r_\tau = y - A_\tau \hat{\theta}_\tau = y - A_\tau (A_\tau^T A)^{-1} A_\tau^T y$ 。

(6)判断是否完成迭代。更新迭代次数  $\tau = \tau + 1$ ,如果未到达指定迭代次数返回步骤若满足指定迭代次数或残差分量计算为 0,则结束迭代。

由于 ROMP 一次选择的原子更多,有利于简化计算步骤,提高重建精度,相比于 OMP 能得到更加精确的结果。

## 2.4 方法步骤

第一步,构建冗余字典;使用 K-SVD 方法,建立 DCT 初始字典训练大量的地震信号中的奇异部分构建冗余字典。同时建立 Curvelet 字典作为初始字典训练大量地震信号中的平滑部分构建另一个冗余字典。

第二步,数据预处理;对重建地震信号进行预处理,通过聚合经验模态分解出该地震信号的各频率分量。进而将该地震信号分别分成结构部分和纹理部分。

第三步,稀疏表示;将地震信号的不同部分分别通过不同的培训字典,得到地震信号不同部分的稀疏分量。

第四步,数据重建;采用贪婪算法中的 ROMP 算法进行数据重建,在减少计算量的同时增加重建的精确度。

# 3 地震数据的去噪与重建实验与分析

为验证该方法在地震信号的去噪和重建中是否优于传统的 K-SVD 地震信号去噪与重建方法,对合成地震信号的去噪实验以及实际带噪地震信号在规律及随机缺失 10%、30%、50%、70% 信号道分别进行实验,验证文中提出方法。

## 3.1 合成地震信号去噪

实验合成数地震为 100 道,90 个采样点。采样频率为 1 kHz 的合成地震数据(图 2a),并将该合成地震信号数据中添加信噪比为 10 dB 的高斯白噪声(图 2b),通过传统 K-SVD 算法合成信号进行去噪(图 2c),再将该合成信号通过形态分量分析进行分类,将不同部分的地震信号通过不同培训的冗余字典,以此进行去噪与重建(图 2d)。

通过对比可以看出,两类方法都能很好的对合成地震信号起到去噪作用。文中采取的方法对合成信号做了分类对比,可以看出去噪效果得到明显的改善,与此同时,细节部分得到突出。由上述实验可以看出,改进的 K-SVD 算法更利于信号去噪。



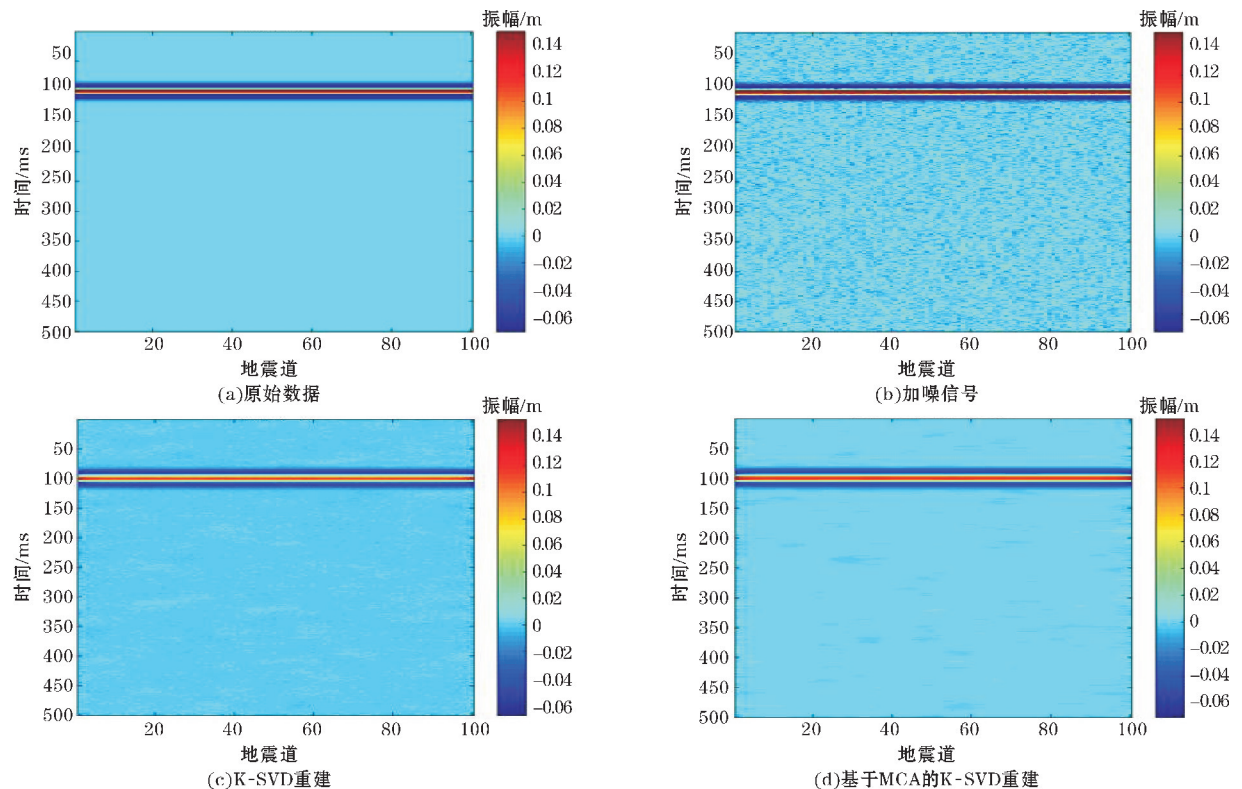


图 2 K-SVD 和改进 K-SVD 去噪对比图

3.2 对合成信号的去噪重建

实验中依然采用上一节的合成信号,合成数地震为 100 道,90 个采样点。采样频率为 1 kHz 的合成地震数据。

在信号实验中,假定该实验信号规律缺失 30%,并在去噪的基础上对其进行重建(图 3)。从信噪比为 20 dB、10 dB 依次为该合成信号添加噪声(图 4~5)。并对比实验结果(见表 1)。

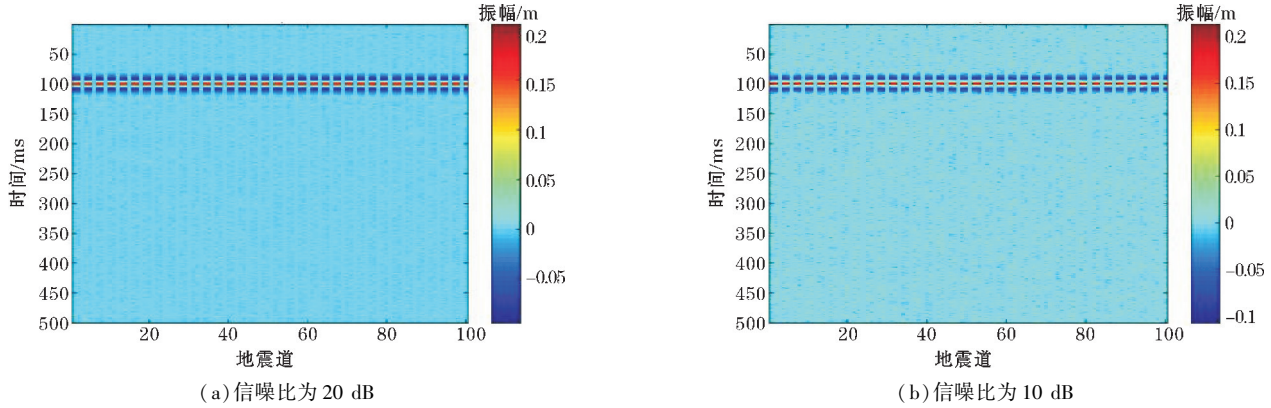


图 3 规律缺失 30% 且在 20 dB、10 dB 信噪比下的合成信号

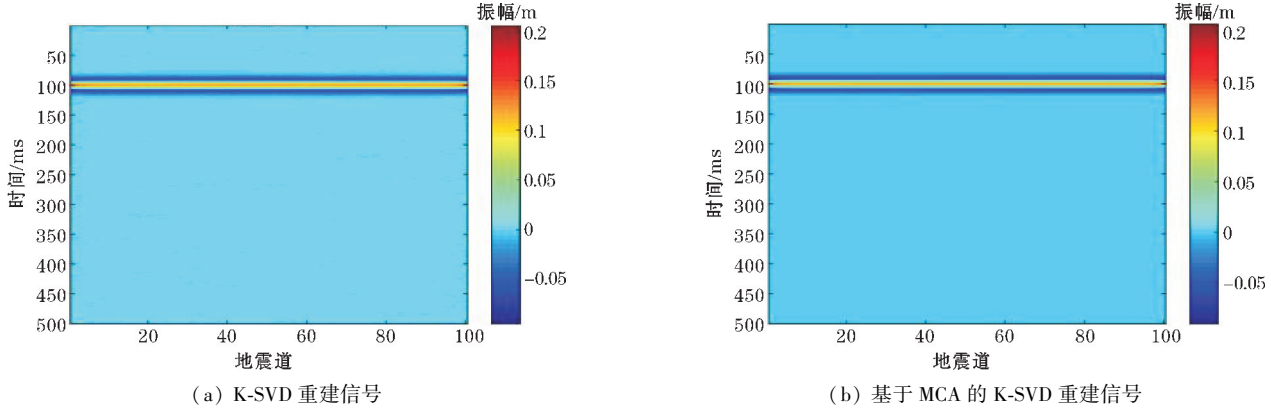


图 4 信噪比为 20 dB,规律缺失 30% 的去噪重建图

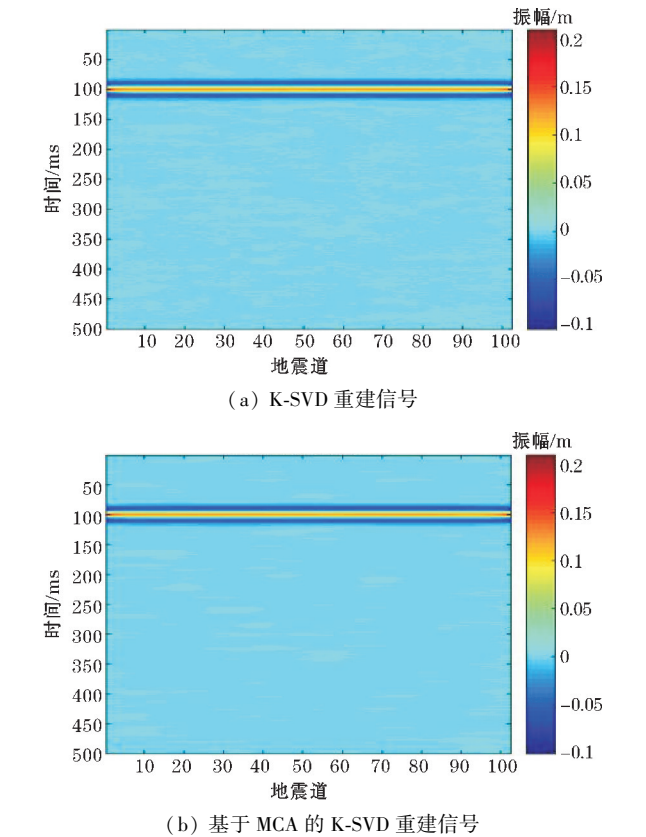


图 5 信噪比为 10 dB,规律缺失 30% 的去噪重建图

表 1 不同噪声下重建信号信噪比对比表 单位: dB

合成信号的 信噪比	K-SVD 重建 信号信噪比	基于 MCA 的 K-SVD 重建信号信噪比
20	55.203 3	55.304 3
10	40.108 0	46.143 2

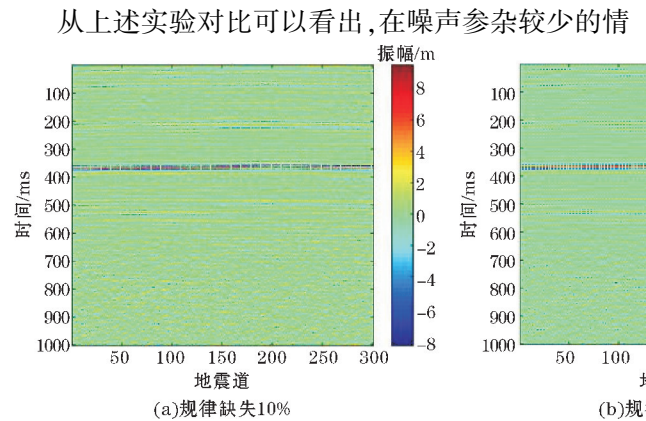


图 7 规律缺失 10%、30%、50% 的实际信号

通过 K-SVD 和基于 MCA 的 K-SVD 算法对上述 3 种规律缺失的地震信号进行去噪与重建,得到图 8~10。

通过对比可得出,对地震信号进行分块处理后,细节更突出,重建效果更好,并能很好地起到去掉原信号所带的高斯白噪声。

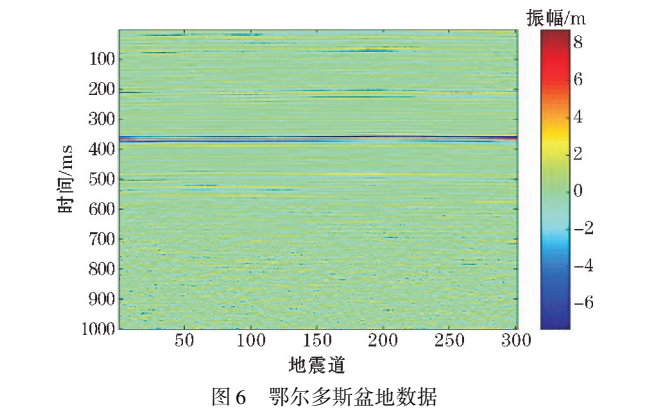
3.3.2 对实际信号进行随机信号道缺失重建实验

为节约采样成本,采样信号需要进行重建处理。由于地震信号采集环境复杂,采集过程中会出现地震

况下,两种方法的去噪效果相似,但文中方法于细节上更加突出,重建效果更好,更接近于合成信号。在噪声较大的情况下,两种方法都起到一定的去噪效果。相比于传统方法,文中方法在重建效果良好的情况下去噪效果更优。综上,改进的 K-SVD 算法对于带噪信号的重建结果远优于传统方法。

3.3 对实际带噪地震信号的去噪与重建

实验数据为鄂尔多斯盆地数据中 CG-H096224 数据,该数据为叠后地震信号。为方便观察及实验,对该信号的观测截取在 300 道左右进行去噪与重建,如图 6 所示。



3.3.1 对实际信号进行规律地震道缺失实验

在实际地震信号的测量中,为减少数据处理的计算量和采样成本,地震样本通过离散点组成。即地震样本的采样间隔较大。为进行数据资料分析需要对这

些采样信号进行重建和去噪处理。

所以,对实际地震信号进行一定程度的规律地震道缺失实验,分别为规律缺失 10%、30%、50% (图 7)。

道缺失或者炮道为废道的现象。由此造成采样信号的缺失,也需要通过信号处理,对采样信号进行重建,以此还原高精度的原始地震信号。

采取上述实际信号,分别对该信号进行随机信号道模拟实验,信号缺失率分别为 30%、50%、70%。并用 K-SVD 算法和文中算法进行去噪重建实验,得出结果如图 11~14 所示。



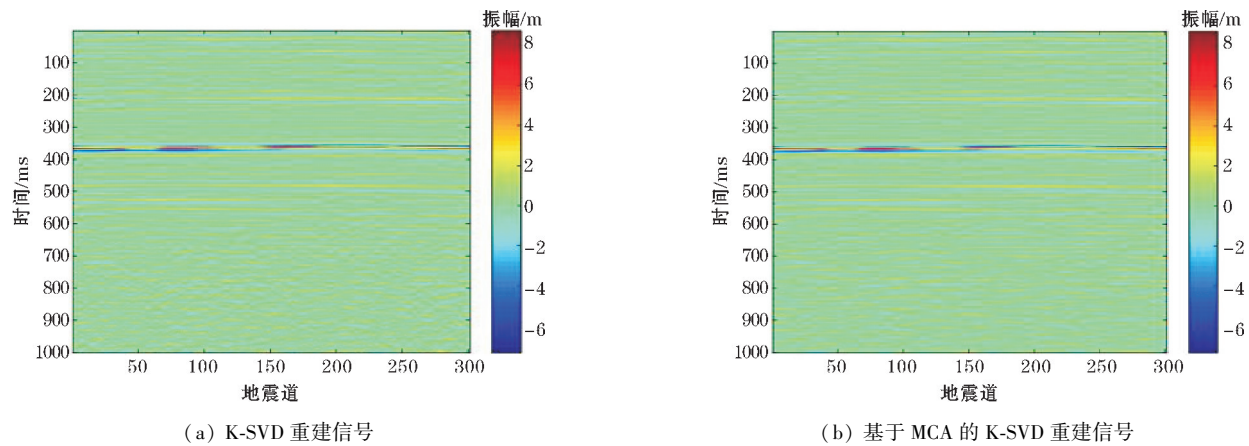


图8 实际地震信号规律缺失 10% 的去噪重建图

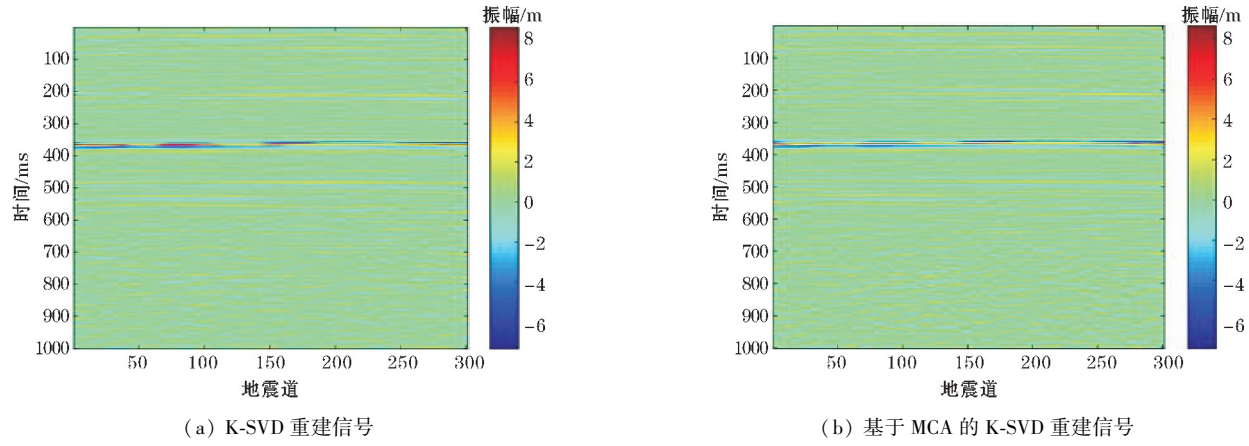


图9 实际地震信号规律缺失 30% 的去噪重建图

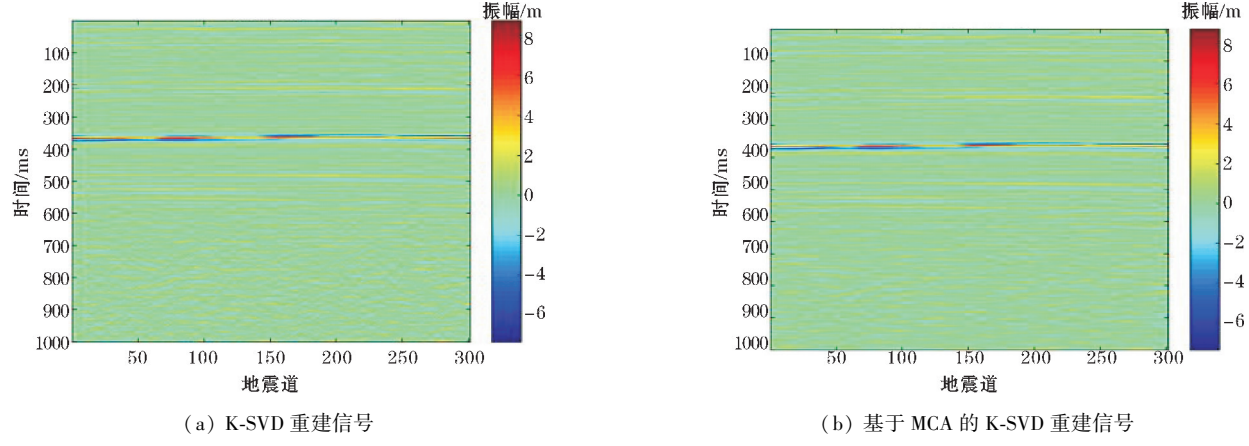


图10 实际地震信号规律缺失 50% 的去噪重建图

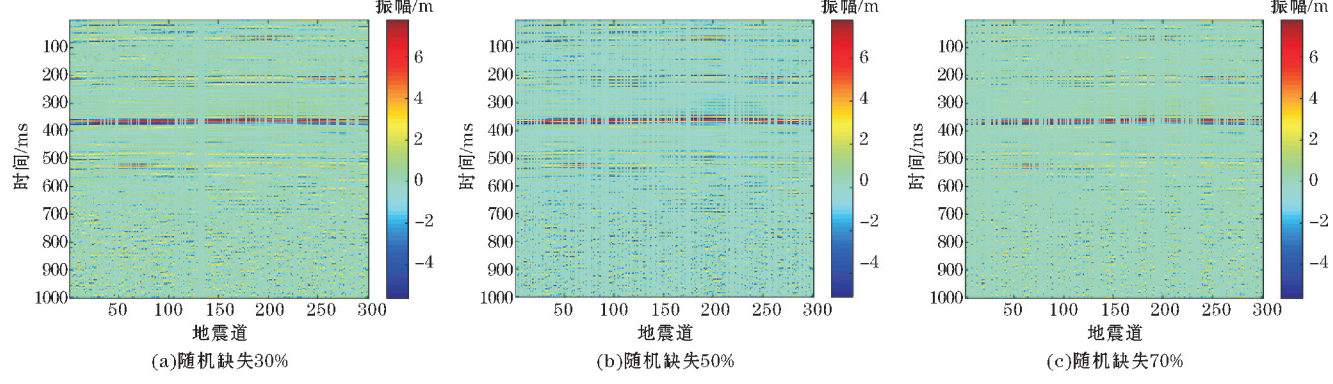


图11 随机缺失 30%、50%、70% 的实际信号

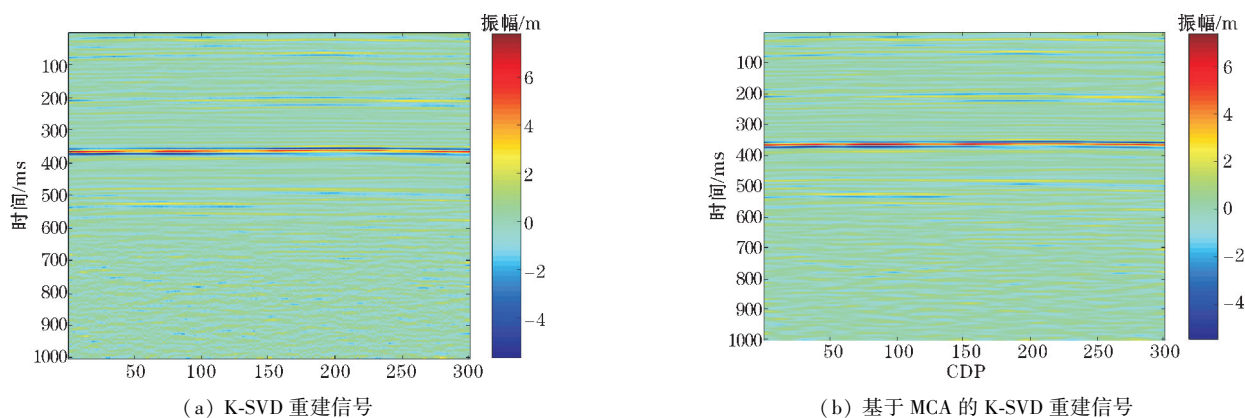


图 12 实际地震信号随机缺失 30% 的去噪重建图

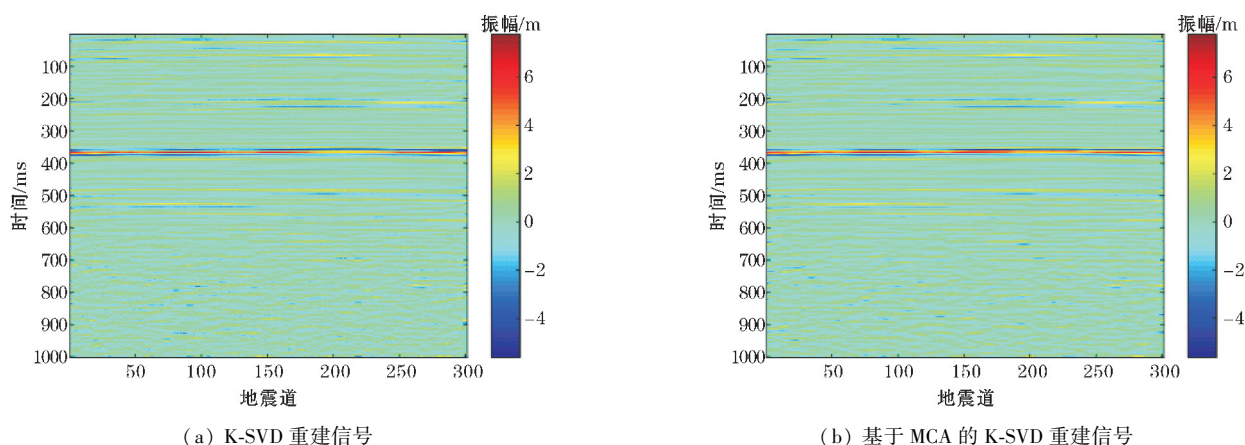


图 13 实际地震信号随机缺失 50% 的去噪重建图

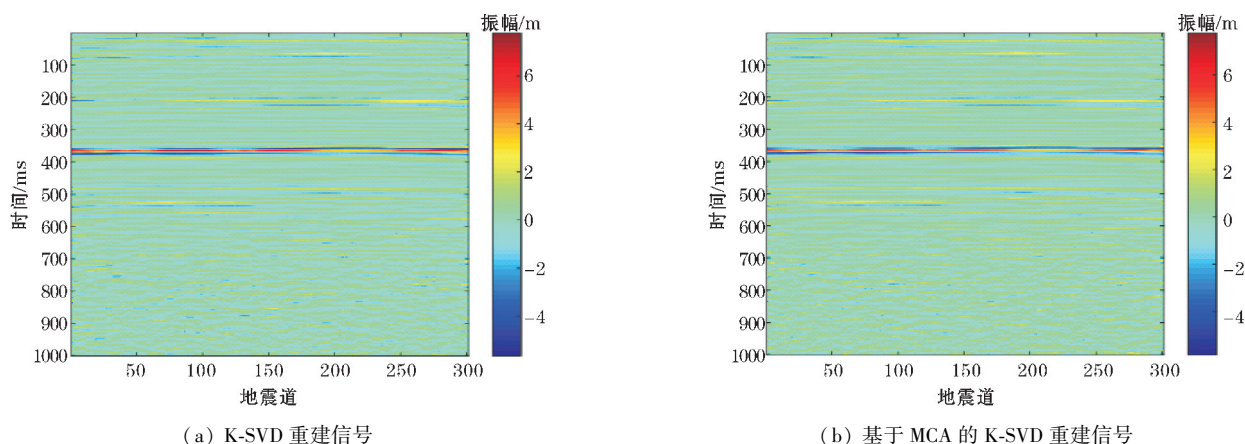


图 14 实际地震信号随机缺失 70% 的去噪重建图

实验结论可以看出,随机信号道缺失细节方面能得到恰好的去噪,文中提出的方法有更好的去噪效果,并能更加突出细节。

通过上述两个实验可以看出,改进的 K-SVD 算法不论是在人为减少炮道的情况还是自然采集缺失炮道的情况都能对地震信号进行很好的去噪重建。相比于传统的 K-SVD 算法,重建的信号精度更高,更贴近实际信号。

## 4 结束语

在地震信号的重建中,压缩感知算法可以很好的

进行运用,并在实际信号分析中起到节省成本以及应对复杂环境采样良好的作用。通过对合成信号、缺失含噪合成信号、实际信号规律缺失、实际信号随机缺失一系列实验得出基于 MCA 对 K-SVD 地震信号的去噪与重建是优于传统算法的。即通过 MCA 算法对地震信号进行平滑和结构的分类,并对不同分类进行针对性的去噪重建。但对于过高噪声的信号,该方法也存在一定的局限性。在接下来的研究中,会对如何应对高过噪声地震信号的去噪与重建进行进一步扩展与讨论。



## 参考文献:

- [1] 石战战,夏艳晴,王元君,等.基于 $L_1-L_1$ 范数稀疏表示的共偏移距道集地震数据重建方法[J].地球物理学进展,2019,34(5):1893-1899.
- [2] Serdyukov A, Yablokov A, Duchkov, et al. Slant f-k transform of multichannel seismic surface wave data[J]. Geophysics, 2019, 84(1): A19-A24.
- [3] 杨海涛,贺振华,陈丹,等.转换波静校正技术在四川地区的应用研究[J].地球物理学进展, 2017, 32(4): 1701-1704.
- [4] 池越,丁木,周亚同,等.地震信号的 Landweber 迭代傅里叶快速重建[J].煤炭学报, 2018, 43(9): 2562-2569.
- [5] Tonegawa T, Nishida K, Watanabe T, et al. Seismic interferometry of teleseismic S-wave coda for retrieval of body waves: An application to the Philippine Sea slab underneath the Japanese Islands[J]. Geophysical Journal International, 2018, 178(3): 1574-1586.
- [6] 舒国旭,吕公河,吕尧,等.基于压缩感知的地震数据重建[J].石油物探, 2018, 57(4): 66-71.
- [7] 曹慧慧.压缩感知方法在地震数据中的应用探讨[D].成都:西南石油大学, 2018.
- [8] 弓震,宋维琪.基于 K-SVD 字典学习方法的微地震去噪[Z].中国石油学会 2019 年物探技术研讨会, 2019.
- [9] 张良,韩立国,许德鑫,等.基于压缩感知技术的 Shearlet 变换重建地震数据[J].石油地球物理勘探, 2017(2): 8-42.
- [10] 程时俊,韩立国,于江龙,等.基于改进 K-SVD 字典学习方法的地震数据去噪[J].世界地质, 2018, 37(2): 627-634.
- [11] 黄翔.基于 EMD 重构地震信号的去噪方法[J].油气地球物理, 2017, 15(2): 18-23.
- [12] Jawad S, IM Qi. Reconstruction of Sparse Signals and Compressively Sampled Images Based on Smooth  $l_1$ -Norm Approximation[J]. Journal of Signal Processing Systems, 2017, 88(3).
- [13] FU Jun, YUAN Haikou, ZHAO Rongqiang, et al. Clustering K-SVD for sparse representation of images[J]. EURASIP Journal on Advances in Signal Processing, 2019(1).
- [14] 胡平平.基于形态分量分析的图像融合算法研究[D].北京:北京交通大学, 2019.
- [15] 江萍,张凯,张医奎,等.基于形态分量分析的含噪地震数据重建方法[J].地球物理学进展, 2019, 34(2): 573-580.
- [16] 张凯,张医奎,李振,等. MCA 框架下 Shearlet 和 DCT 字典组合地震数据重建[J].石油地球物理勘探, 2019. 54(5): 1005-1013.

## Denoising and Reconstruction of Seismic Signals based on K-SVD under the MCA Framework

CHEN Siqi<sup>1</sup>, XUE Yajuan<sup>1</sup>, YANG Qingmi<sup>2</sup>, FANG Lixin<sup>2</sup>, ZHENG Shuling<sup>2</sup>

(1. College of Communication Engineering, Chengdu University of Information Technology, Chengdu 610225, China; 2. Sinopec Key Laboratory of Geophysics, Nanjing 21000, China)

**Abstract:** For the incompleteness of sample signal due to the complexity of seismic signal collection environment, and quantities of noise. Therefore, before analyzing it, we need to denoise and reconstruct the pre-stack seismic data. On the basis of building the compress sensing reconstruction of redundant dictionaries by traditional K-singular value decomposition(K-SVD), we propose the K-SVD seismic signal denoise and reconstruction combined with morphometric principal components analysis(MCA). It realizes that we use MCA to classify the structure and smooth part of the seismic signal. And build redundant dictionaries calculated by the K-singular value decomposition for the above two parts, respectively. Finally, the two parts are placed in different dictionaries for denoising and reconstruction. Compared with the traditional method, this method reduces the cost and difficulty of seismic data acquisition, accurately discriminates signal details, and achieves good denoising effect.

**Keywords:** seismic data; compressed sensing; K-singular value decomposition(K-SVD); morphometric principal components analysis(MCA); redundant dictionary