

引用格式：汪恺, 李风华, 杨习山. 利用字典学习的浅海被动声层析[J]. 声学技术, 2022, 41(5): 637-641. [WANG Kai, LI Fenghua, YANG Xishan. Shallow water passive acoustic tomography using dictionary learning[J]. Technical Acoustics, 2022, 41(5): 637-641.] DOI: 10.16300/j.cnki.1000-3630.2022.05.001

# 利用字典学习的浅海被动声层析

汪 恺<sup>1,2</sup>, 李风华<sup>1</sup>, 杨习山<sup>1</sup>

(1. 中国科学院声学研究所, 北京 100049; 2. 中国科学院大学, 北京 100190)

**摘要：**针对被动海洋声层析，提出了一种利用字典学习从海洋环境噪声反演浅海声速剖面的方法。首先，通过海洋噪声互相关函数提取出两个水平阵列间的经验格林函数；其次，通过字典学习从数据生成字典矩阵来稀疏表征声速剖面；最后通过搜索稀疏的系数来实现对浅海声速剖面的反演。通过南海实验数据对本方法进行了验证，相对于传统被动海洋声层析方法实际反演结果的均方根误差降低为 $0.53 \text{ m}\cdot\text{s}^{-1}$ 。而且搜索参数更少，同时具有较高的准确性。

**关键词：**格林函数；字典学习；声速剖面反演

中图分类号：O427.9

文献标志码：A

文章编号：1000-3630(2022)-05-0637-05

## Shallow water passive acoustic tomography using dictionary learning

WANG Kai<sup>1,2</sup>, LI Fenghua<sup>1</sup>, YANG Xishan<sup>1</sup>

(1. Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** A dictionary learning-based method for sound speed profile (SSP) inversion in shallow water is presented. The empirical Green's functions between two parallel horizontal arrays can be extracted from noise cross-correlation functions. The sound speed profiles are sparsely characterized by data-generated dictionary matrix, and they can be inverted by searching for sparse coefficients. This method is validated by experimental data in the South China Sea. Compared with the traditional empirical orthogonal function (EOF) methods, the inversion accuracy accuracy is reduced to  $0.53 \text{ m}\cdot\text{s}^{-1}$ , moreover this method has fewer search parameters and higher accuracy.

**Key words:** Green's function; dictionary learning; sound speed profile (SSP) inversion

## 0 引言

海洋环境特性对声传播的影响十分重要，其声场对水体声速变化极为敏感，这种声波和海洋环境特性之间的联系使得从声信号中提取海洋环境信息成为可能。自从 Godin 论证了噪声互相关函数的时延和两点之间的格林函数时间到达结构一致后，被动声层析方法就得到长足的发展<sup>[1]</sup>。通过噪声互相关函数，可以提取两个接收器之间的经验格林函数来反演声速剖面(Sound Speed Profile, SSP)<sup>[2]</sup>。Godin 等成功从 7 天平均的噪声互相关函数反演了海洋的平均声速<sup>[3]</sup>，在此基础上 Li 等通过波束形成将格林

函数的提取时间缩减到了 2 h<sup>[4]</sup>。

对于声速剖面的重建方法，Bianco 等<sup>[5]</sup>提出了字典学习表征声速剖面的方法，相对于传统的经验正交函数表征方法(Empirical Orthogonal Function, EOF)<sup>[6-8]</sup>，字典学习表征方法在声速剖面重建之后能恢复出更多声速剖面的起伏。由于不需要保证基函数正交，因此对训练数据的选择更加灵活<sup>[9]</sup>。

本文通过稀疏表达，用字典学习代替经验正交函数来表征声速剖面，在构建出字典矩阵后，从海洋环境噪声中准确反演得到了声速剖面，比传统的被动声层析方法精度更高，同时待反演参数更少。

## 1 基于字典学习的被动声层析方法

### 1.1 水平阵列噪声提取经验格林函数

图 1 为水平阵列被动声层析示意图，展示了水平阵列噪声提取格林函数的实验布放方式。海底两个平行的水平线阵连续不断地接收海洋环境噪声，

收稿日期：2021-06-16；修回日期：2021-07-28

基金项目：国家自然科学基金项目(11974017, 11874061)、中国科学院青年创新促进会人才项目(2017028)。

作者简介：汪恺(1995—)，男，安徽省合肥市人，博士，研究方向为水声物理。

通信作者：汪恺, E-mail:wangkai1995@mail.ioa.ac.cn

将其划分为  $K$  个快拍后, 通过波束形成加速能量累积, 两个水平阵列之间的噪声互相关函数可通过下式计算得到<sup>[4]</sup>:

$$\begin{aligned} \mathbf{F}^{(k)}(\omega) &= \mathbf{W}_1^H(\omega, \theta_0) \mathbf{P}_1^{(k)}(\omega) \\ &\quad \mathbf{P}_2^{(k)H}(\omega) \mathbf{W}_2(\omega, \theta_0) \end{aligned} \quad (1)$$

式中:  $\mathbf{P}_i^{(k)}(\omega)$  是第  $i$  个阵列第  $k$  个快拍在傅里叶变换后的频域信号  $\mathbf{P}_i^{(k)}(\omega) = [\mathbf{p}_i^{(1)}(\omega) \ \mathbf{p}_i^{(2)}(\omega) \ \dots \ \mathbf{p}_i^{(N)}(\omega)]$ , 其中  $i$  为阵列的编号,  $i=1, 2$ ,  $N$  代表了每个阵列的阵元总数;  $\mathbf{W}_i^{(k)}(\omega, \theta_0) = [w_i^{(1)}(\omega, \theta_0) \ w_i^{(2)}(\omega, \theta_0) \ \dots \ w_i^{(N)}(\omega, \theta_0)]$  是阵列  $i$  对应的权重向量, 其中  $w_i^{(j)}(\omega, \theta_0) = \exp[-j\omega\tau_j(\theta_0)]$ ,  $j$  表示第  $j$  个阵元,  $\theta_0$  为阵列 1 的第  $m$  个阵元水听器指向阵列 2 的第  $n$  个水听器的方向。

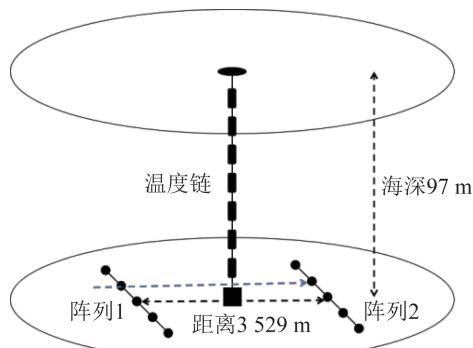


图1 水平阵列被动声层析示意图

Fig.1 Schematic diagram of passive acoustic tomography with two parallel horizontal arrays

在经过  $K$  个快拍的累积叠加后, 可以得到累积的经验格林函数(Empirical Green Function, EGF):

$$\mathbf{G}(\tau) \approx -\frac{\int_{-\infty}^{\infty} \left[ \sum_{k=1}^K \mathbf{F}^{(k)}(\omega) \right] \exp(j\omega\tau) d\omega}{d\tau} \quad (2)$$

在提取出的经验格林函数基础上可以对声速剖面进行进一步反演。

## 1.2 基于字典学习的被动声层析射线反演理论

由于海洋波导在实际传播过程中的多途效应,  $r_{1,m}$  和  $r_{2,n}$  两点间的格林函数可以表示为各条本征声线的叠加:

$$\mathbf{G}(r_{1,m}, r_{2,n}; \tau) \approx \sum_{i=1}^I a_i s(\tau - t_i) \quad (3)$$

式中:  $I$  是本征声线的数目;  $a_i$  是本征声线的幅度;  $t_i$  是本征声线的传播时间。在实际的声层析中往往只有  $J$  条可清晰分辨的声线路径, 其组成的时间到达向量为

$$\mathbf{t} = [t_1 \ t_2 \ \dots \ t_J]^T \quad (4)$$

通过给定不同的声速剖面进行声场仿真计算, 将模型计算得到各声线路径到达时间与提取格林函数的各个声线到达时间对比, 进而匹配出最接近实

际海洋环境的声速剖面, 也就是求解式(5):

$$\hat{\mathbf{c}}(z) = \arg \min \| \mathbf{t} - \mathbf{t}^{\text{model}}(\mathbf{c}(z)) \|^2 \quad (5)$$

其中:  $\mathbf{c}(z) \in \mathbb{R}^{Z \times 1}$  为深度上的声速剖面,  $Z$  为深度上的采样点数,  $\mathbb{R}$  表示实数集,  $\hat{\mathbf{c}}(z)$  为求解得到的代价函数最小时的声速剖面。

考虑到实际声速剖面往往深度采样点数较多, 为了减少反演参数, 与 EOF 方法类似, 用字典学习方法对声速剖面进行稀疏表征:

$$\mathbf{c}(z) = \mathbf{c}_0(z) + \mathbf{Q}\mathbf{x} \quad (6)$$

式中:  $\mathbf{c}(z) \in \mathbb{R}^{Z \times 1}$  是时间平均声速,  $\mathbf{Q} \in \mathbb{R}^{Z \times D}$  为字典矩阵,  $D$  为字典的维数,  $\mathbf{x} \in \mathbb{R}^{D \times 1}$  为稀疏的系数向量, 仅有  $d$  个非零值, 且  $d \ll D$ 。这里词典  $Q$  通过式(7)从训练数据中生成:

$$\begin{aligned} \mathbf{X}_{\text{opt}}, \mathbf{Q}_{\text{opt}} &= \arg \min \| \mathbf{Q}\mathbf{X} + \mathbf{c}_0 - \mathbf{C} \|^2 \\ \text{Subject to } \forall \|\mathbf{x}\|_0 &\leq d, \forall \|\mathbf{q}\|_2 \leq 1 \end{aligned} \quad (7)$$

其中:  $\mathbf{C}$  为  $\mathbf{c}(z)$  组成的声速剖面矩阵,  $\mathbf{X}$  为系数向量  $\mathbf{x}$  组成的系数矩阵,  $\mathbf{q}$  为  $\mathbf{Q}$  中的字典向量,  $\mathbf{Q}_{\text{opt}}$  为最后生成的字典矩阵,  $\mathbf{X}_{\text{opt}}$  为对应的系数矩阵。

式(5)的代价函数可进一步表示为

$$\hat{\mathbf{c}}(z) = \arg \min \| \mathbf{t} - \mathbf{t}^{\text{model}}(\mathbf{c}(\mathbf{x})) \|^2 \quad (8)$$

由于约束较少, 难以搜索到最优解, 因此借助卫星遥感测量的海洋表面温度进行联合同化反演<sup>[10]</sup>, 优化问题进一步表示为

$$\hat{\mathbf{c}}(z) = \arg \min \| \mathbf{t} - \mathbf{t}^{\text{model}}(\mathbf{c}(\mathbf{x})) \|^2 + \mu \| T_s - T(\mathbf{x}) \|^2 \quad (9)$$

其中:  $T_s$  为卫星遥感测量的海表温度,  $\mu$  为正则化参数,  $T(\mathbf{x})$  为依据构建的声速剖面计算得到的海表温度<sup>[11]</sup>。由于系数向量  $\mathbf{x}$  的稀疏性, 只需要搜索  $\mathbf{x}$  中  $d$  个参数即可对式(9)求解, 搜索出最优的声速剖面。

## 2 实验处理

实验布放示意图如图1所示, 在海深约为 97 m 的海底布放两个大致平行的水平阵列, 阵列水平距离为 3 529 m, 每个阵列有间隔 6 m 的 15 个等间距阵元, 实际阵列长度大约为 84 m, 满足远场假设。一条包含 38 个 TD 的温度链被布放在两个水平阵列的中央用以记录水体大约 2.5~95.0 m 的水文数据。接收到的噪声数据以 10 s 为一拍进行划分, 同时对  $\mathbf{P}_i^{(k)}(\omega)$  进行 20~400 Hz 的滤波。在满足远场近似并且两水平阵列型相似、阵元数目一致的条件时, 通过常规波束形成, 依据式(1)~(2), 通过叠加多个快拍可以提取出 2 h 时间平均的经验格林函数<sup>[4]</sup>。图2 展示了提取的一条典型的格林函数。

温度链间断测量了该海域从 2018 年 4 月 19 日~

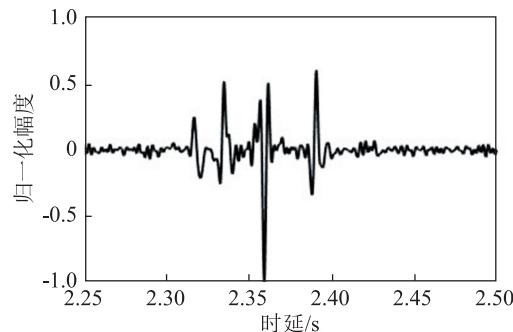


图2 2 h 平均的经验格林函数  
Fig.2 An empirical Green's function for 2 h average

7月30日的水文信息。声速剖面通过插值得到完整海深深度上的声速分布。图3展示了按1 h平均采样后的声速剖面，其中第146~240号、第740~793号样本的测量时间与接收信号时间相对应，因此这些水文与格林函数时间一致被降采样到2 h，用来测试反演结果。其余水文用来构建声速剖面的学习字典和EOF。字典维度 $D$ 设为200，稀疏参数 $d$ 设为1，使用K-SVD算法来从式(7)中生成更新字典<sup>[8-9]</sup>。K-SVD算法是一种经典的字典学习算法，在迭代过程中使用 $K$ 次奇异值分解(Singular Value Decomposition, SVD)<sup>[12]</sup>。

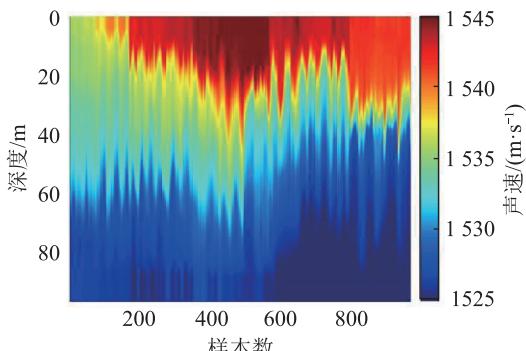


图3 温度链间断测量的声速剖面  
Fig.3 SSPs discontinuously measured by thermistor chain

图2中格林函数第2号包络的波峰和第3号包络的波谷以及第4号包络的波峰更加稳定，因此用这三个时延代入式(9)中反演声速剖面<sup>[13]</sup>。EOF方法采用3阶EOF系数，由于3阶组合的搜索参数较多，我们采用遗传算法对参数组合进行搜索。字典学习方法的参数，由于 $x$ 的稀疏性( $d=1$ )，直接遍历字典向量各维( $D=200$ )，进行一维的参数搜索即可反演声速剖面。

### 3 实验结果与讨论

图4(a)展示了前36阶字典向量(按方差排序)，图4(b)则是前36阶经验正交函数。可以看出图中看出，EOF的信息主要由前几阶EOF所表征，高阶

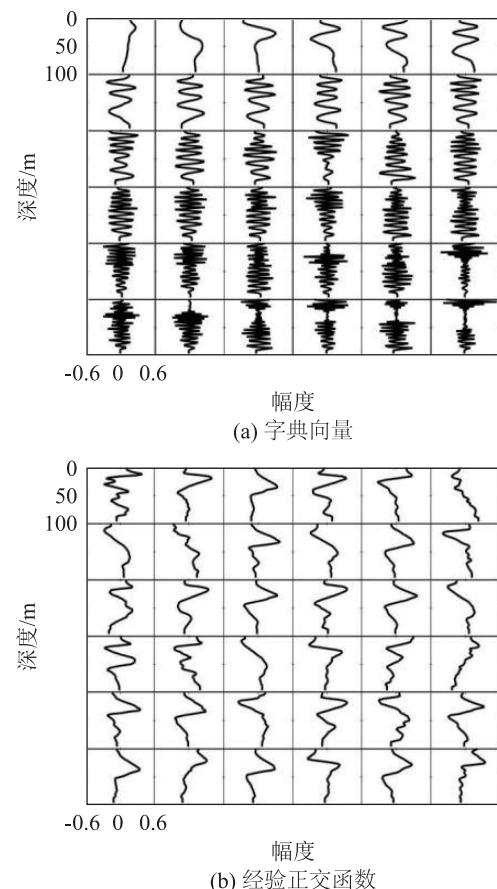


图4 前36阶字典和经验正交函数得到的声速剖面  
Fig.4 SSPs obtained from leading 36 order dictionaries and traditional empirical orthogonal function

EOF蕴含的信息很少，而每个字典向量都能表达声速起伏。

图5展示了用来构建EOF和字典的声速剖面在经过两种方法分解后，由1阶稀疏度的字典方法和前3阶EOF重建的结果。从图中5可以看到，3阶EOF和1阶稀疏度的字典学习方法都能较好地表征重建声速剖面。它们的平均均方误差分别约为 $0.28 \text{ m} \cdot \text{s}^{-1}$ 和 $0.61 \text{ m} \cdot \text{s}^{-1}$ ，字典学习方法重建的误差更小。

对式(9)进行参数搜索后，反演的声速剖面结果如图6所示。从图可以看到，字典学习方法更准确地反映了声速剖面的起伏变化，其均方误差约为 $0.53 \text{ m} \cdot \text{s}^{-1}$ ，而3阶EOF方法精度相对较低，均方误差约为 $0.78 \text{ m} \cdot \text{s}^{-1}$ 。图7为图6中第54号声速剖面的对比，从图中可以看到字典学习反演结果更加准确。

图8对比了EOF不同阶数和字典不同稀疏度对反演结果的影响。从图中可以看到，字典学习方法相对于EOF方法总体上反演结果精度更好，并且EOF在阶数增加到4，字典学习在稀疏度增加到3之后，均方误差没有继续下降，反而可能大大增

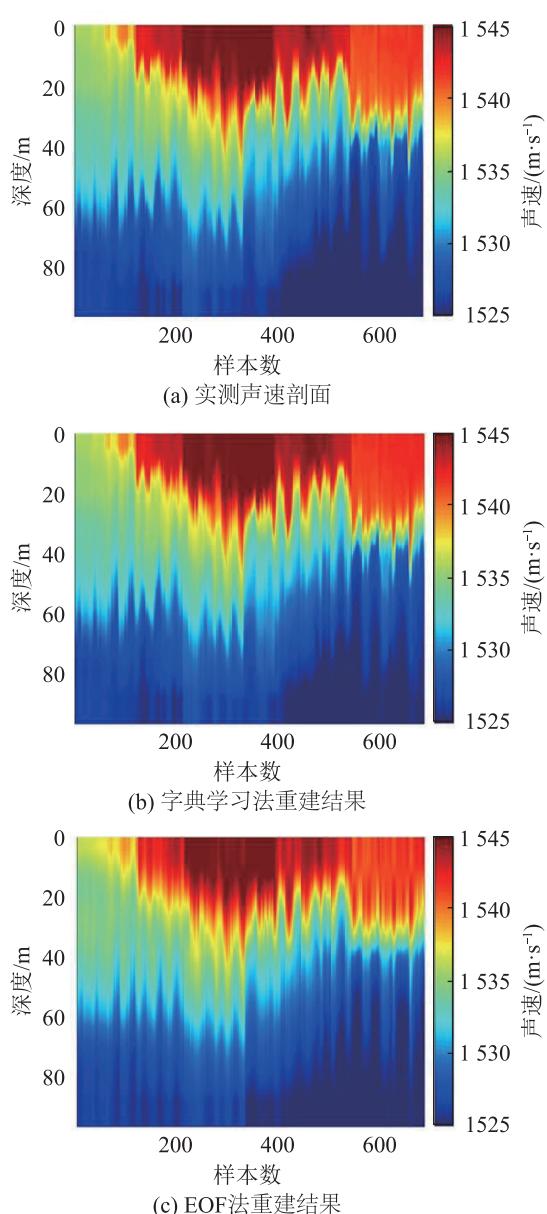


图5 实测声速剖面及字典学习法和EOF法的重建结果  
Fig.5 The measured SSPs and the SSPs reconstructed by Dictionary Learning and by EOF

加。这是由于在被动声层析中代价函数式(9)的约束项较少，随着阶数或稀疏度继续增加，可能难以搜索到更优的参数组合，反而由于多解性导致反演精度下降。

声速剖面的声学反演是一个非线性多解问题，往往使用遗传算法进行参数搜索，难以直接对比计算时间。通常对每个系数划分为 $H$ 个离散值，对各阶系数组合进行搜索，参数组合为 $O(H^n)$ 量级， $n$ 为阶数。从图8可以发现，在同等精度下，字典学习只需要1阶系数，而EOF方法需要3阶系数。假设系数划分的离散值个数 $H=1000$ ，字典维数为200，那么字典方法有 $2 \times 10^5$ 个可能的解，EOF方法有 $10^9$ 个可能的解。因此在同等精度下，

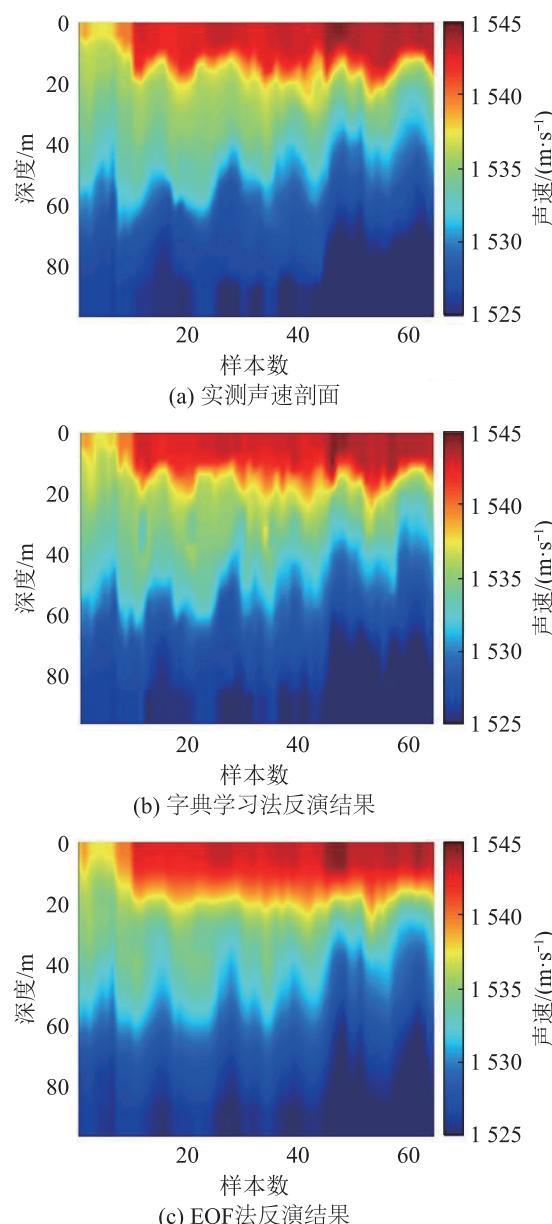


图6 实测声速剖面及字典学习法和EOF法的反演结果  
Fig.6 The measured SSPs and the SSPs obtained by inversion with Dictionary Learning and with EOF

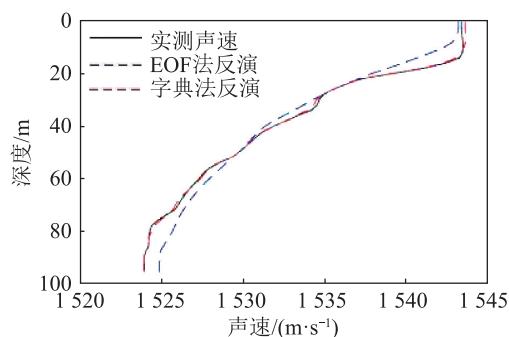


图7 不同反演方法得到的声速剖面对比  
Fig.7 Comparison of SSPs obtained by inversion with different methods

通常字典学习方法的待搜索参数的个数少于EOF方法。

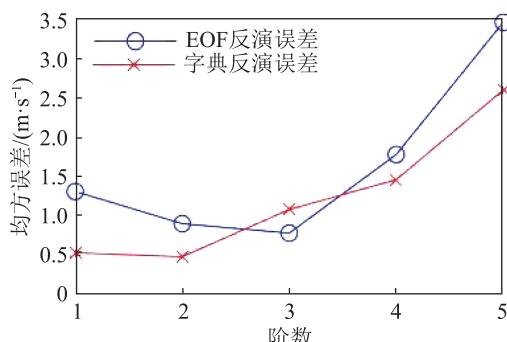


图8 不同方法的声速剖面反演结果的均方误差  
Fig.8 RMSEs of SSP inversion results of different methods

## 4 结论

从海洋噪声互相关函数提取出格林函数之后,通过字典学习方法代替EOF方法表征声速剖面,更准确地反演出了海洋声速剖面的变化趋势,反演精度为 $0.53 \text{ m}\cdot\text{s}^{-1}$ 。并且在同等精度下,字典学习方法的参数搜索量也远小于EOF方法。

但由于被动声层析代价函数约束项较少,难以通过增加稀疏度的方式来进一步提高精度。

## 参 考 文 献

- [1] GODIN O A. Recovering the acoustic Green's function from ambient noise cross correlation in an inhomogeneous moving medium[J]. Physical Review Letters, 2006, **97**(5): 054301.
- [2] GODIN O A, IRISOV V G, CHARNOTSKII M I. Passive acoustic measurements of wind velocity and sound speed in air [J]. The Journal of the Acoustical Society of America, 2014, **135**(2): EL68-EL74.
- [3] GODIN O A, ZABOTIN N A, GONCHAROV V V. Ocean tomography with acoustic daylight[J]. Geophysical Research Letters, 2010, **37**(13).
- [4] LI F H, YANG X S, ZHANG Y J, et al. Passive Ocean acoustic tomography in shallow water[J]. The Journal of the Acoustical Society of America, 2019, **145**(5): 2823.
- [5] BIANCO M, GERSTOFF P. Dictionary learning of sound speed profiles[J]. The Journal of the Acoustical Society of America, 2017, **141**(3): 1749.
- [6] 沈远海, 马远良, 屠庆平, 等. 浅水声速剖面用经验正交函数(EOF)表示的可行性研究[J]. 应用声学, 1999, **18**(2): 21-25.
- [7] SHEN Yuanhai, MA Yuanliang, TU Qingping, et al. Feasibility of description of the sound speed profile in shallow water via empirical orthogonal functions (EOF)[J]. Applied Acoustics, 1999, **18**(2): 21-25.
- [8] 周士弘, 张茂有, 周日鹏. 海洋声速场的经验正交函数描述及声速剖面预报[J]. 海洋通报, 1999, **18**(5): 27-34.
- [9] ZHOU Shihong, ZHANG Maoyou, Zhou Yuepen. Study on empirical orthogonal functions expression and prediction of the sound speed field[J]. Marine Science Bulletin, 1999, **18**(5): 27-34.
- [10] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006, **54**(11): 4311-4322.
- [11] RAJAN S D, FRISK G V. Seasonal variations of the sediment compressional wave-speed profile in the Gulf of Mexico[J]. The Journal of the Acoustical Society of America, 1992, **91**(1): 127-135.
- [12] LI Qianqian, SHI Juan. Sparse expression of sound speed profiles with dictionary learning[C]//Li Qianqian. Acoustical Society of China Underwater Acoustics Conference 2019.
- [13] ELISSEEFF P, SCHMIDT H, XU W. Ocean acoustic tomography as a data assimilation problem[J]. IEEE Journal of Oceanic Engineering, 2002, **27**(2): 275-282.
- [14] RAJAN S D, FRISK G V. Seasonal variations of the sediment compressional wave-speed profile in the Gulf of Mexico[J]. The Journal of the Acoustical Society of America, 1992, **91**(1): 127-135.
- [15] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006, **54**(11): 4311-4322.
- [16] LI F H, WANG K, YANG X S, et al. Passive Ocean acoustic thermometry with machine learning[J]. Applied Acoustics, 2021, **181**: 108167.