

基于联合过完备库的信号分离及重构

余付平¹, 冯有前², 杨荣^{1*1}

(1. 空军工程大学防空反导学院, 陕西西安, 710051; 2. 空军工程大学理学院, 陕西西安, 710051)

摘要 在信号稀疏分解理论的基础上, 提出了构建联合过完备库的思想。通过对包含不同特征成分的过完备子库进行联合构建联合过完备库, 它包含丰富的待分解信号的信息, 复合信号在其上具有更好的稀疏性, 同时由于每个分量信号在各自的过完备子库上均具有稀疏性, 利用基追踪算法实现各个分量信号的分离和重构。仿真结果表明: 联合过完备库对复合信号的重构以及分量信号的分离和重构具有较好地效果。

关键词 联合过完备库; 稀疏分解; 信号分离; 稀疏性; 重构误差

DOI 10.3969/j.issn.1009-3516.2012.04.012

中图分类号 TN957.51 **文献标识码** A **文章编号** 1009-3516(2012)04-0056-05

近年来, 由于稀疏分解在信号压缩、去噪、特征提取、信息分类、图像语音处理等信号处理问题中的广泛应用和良好前景, 受到广大学者的关注。然而关于信号稀疏分解的理论还不成熟, 还有许多问题亟待解决。其中过完备库的构建是稀疏分解理论和应用中都需要考虑的难点、重点, 在信号稀疏分解中起到至关重要的作用。

目前, 过完备库的构建一般有2种方法: 根据先验知识构建单一结构特征的过完备库和几个正交基级联构成过完备库。文献[1]提出反复调谐字典; 文献[2]利用多尺度线性调频基对机械振动信号进行稀疏分解; 文献[3]提出了一种多尺度脊波字典构造方法; 文献[4]建立了 Gabor 感知多成分字典; 文献[5-7]提出利用多组正交基构造过完备字典; 文献[8]通过构建混合基重构图像。但是, 当待分解信号由多个不同特征成分的分量信号组合而成时, 单一特征结构的过完备库由于不能包含待分解信号的丰富信息, 具有很大的局限性; 而正交基级联而成的过完备库虽然包含了丰富的被分解信号的特征信息, 但不能保证每个分量信号在各自相应的正交基上具有稀疏性。所以, 针对以上问题, 本文提出构建由不同特征成分的过完备库联合而成的联合过完备库, 并研究信号在联合过完备库下稀疏分解的稀疏性和信号重构效果。

1 稀疏分解原理

给定一个集合 $\psi = \{\varphi_r\}_{r \in \Gamma}$, 其元素是张成整个 N 维 Hilbert 空间 $H = \mathbf{R}^N$ 的单位矢量, 如果其元素个数 $K < N$, 则集合 ψ 构成一个过完备基, 也称之为字典, 其元素 φ_r 为由参数组 r 定义的原子, 且其范数为 1。由于过完备库中原子间的相关性, 则任意信号 $S \in \mathbf{R}^N$ 在过完备原子库 $\psi \in \mathbf{R}^{N \times K}$ 上具有无穷解。信号 $S \in \mathbf{R}^N$ 在过完备库上的稀疏分解就是在无穷多个解向量中寻找非零元素个数最少的解向量。则信号在过完备库上的稀疏分解可以描述为^[9-11]:

$$\min \|\alpha\|_0 \text{ 满足 } S = \psi\alpha \quad (1)$$

式中 $\|\alpha\|_0$ 是解向量 $\{\alpha_r\}$ 中非零项的个数。

信号在过完备库上稀疏分解的问题就是: 在一定误差范围内, 如何从各种可能的解集中选择出非零元素

* 收稿日期: 2012-02-15

基金项目: 国家自然科学基金资助项目(61003148)

作者简介: 余付平(1983-), 女, 河南商丘人, 博士生, 主要从事信号稀疏分解理论、雷达信号处理研究。

E-mail: juunjingji@163.com

最少的一个解向量,即最为稀疏的一个。然而,这个问题的求解是一个 NP 难题。为了解决上述问题,许多稀疏分解算法已经出现,如匹配追踪(Matching Pursuit, MP)算法^[11]、基追踪(Basis Pursuit, BP)算法^[11]、框架方法(Method of Frames, MOF)^[11-12]和最优正交基(Basis Orthogonal Best, BOB)算法^[11,13]。

2 联合过完备库下的信号分离和重构原理

假设信号 f 由 2 个不同结构特征信号 f_1 和 f_2 组成,即 $f = f_1 + f_2$ 。根据信号 f_1 的结构特征构造具有 f_1 信号特征成分的过完备库 $\theta = \{\phi_i, i \in \Gamma_1\}$ ($\Gamma_1 = 1, 2, \dots, L$)。根据稀疏分解理论,信号 f_1 在该过完备库上具有稀疏性,那么存在非零元素 i 个数最少的解向量 α_1 且满足 $f_1 = \theta\alpha_1$ 即:

$$\|\alpha_1\|_0 \leq \|\alpha\|_0 \text{ 且满足 } f_1 = \theta\alpha, \alpha_1 \in \alpha \quad (2)$$

同理,根据信号 f_2 的特征构造包含信号 f_2 特征成分的过完备库 $\varphi = \{\psi_j, j \in \Gamma_2\}$ ($\Gamma_2 = 1, 2, \dots, K$),那么存在:

$$\|\chi_1\|_0 \leq \|\chi\|_0 \text{ 且满足 } f_2 = \varphi\chi, \chi_1 \in \chi \quad (3)$$

同时,信号 f_1 在过完备库 φ 上不具有稀疏性,信号 f_2 在过完备库 θ 上不具有稀疏性,即过完备库 θ 和 φ 之间的相关性较小。

根据过完备库 θ 和 φ , 构建联合过完备库 $D = \theta \cup \varphi = \{\phi_1, \phi_2, \dots, \phi_L, \psi_1, \psi_2, \dots, \psi_K$ (其中 θ 和 φ 被称为联合过完备库 D 的子原子库)。联合过完备库需要下列条件:

1) 满足:

$$\alpha_1 = \arg \min_{\alpha} \|\alpha\|_0 \text{ 满足 } f_1 = \theta\alpha \quad (4)$$

$$\chi_1 = \arg \min_{\chi} \|\chi\|_0 \text{ 满足 } f_2 = \varphi\chi \quad (5)$$

条件的解存在。即信号 f_1 在过完备子字典 θ 上能够稀疏分解,信号 f_2 在过完备子字典 φ 上能够稀疏分解。

2) 满足:

$$f_2 = \theta x \quad (6)$$

$$f_1 = \varphi z \quad (7)$$

条件的解 x, z 的零范数均较大,即信号 f_2 在过完备子字典 θ 上不具有稀疏性,信号 f_1 在过完备子字典 φ 上不具有稀疏性。

$$3) u = \max_{1 \leq i \leq L, 1 \leq j \leq K} \frac{|\langle \phi_i, \psi_j \rangle|}{\|\phi_i\|_2 \cdot \|\psi_j\|_2} \text{ 较小, 即子字典 } \theta \text{ 和 } \varphi \text{ 之间的相关性很小。}$$

以上理论分析是针对 2 个过完备子库构成联合过完备字典问题进行的。将上述理论分析推广到由多个过完备子库组成过完备字典中同样适用。

由于 θ 和 φ 之间的弱相关性和分量信号在各自子原子库上的稀疏性,信号 f 在联合过完备库 D 上具有稀疏性,即存在:

$$\|\beta_1\|_0 \leq \|\beta\|_0 \text{ 且满足 } f = D\beta, \beta_1 \in \beta \quad (8)$$

由于 $f = f_1 + f_2, f_1 = \theta\alpha, f_2 = \varphi\chi, f = D\beta, D = \theta \cup \varphi$, 所以 $\{\alpha \cup \chi\} \subset \beta$, 且 $\alpha_1 \cup \chi_1$ 一定是方程 $f = D\beta$ 的一个解向量。那么由信号稀疏分解原理和联合过完备库的构建思想可知:

$$\|\beta_1\|_0 \leq \|\alpha_1\|_0 + \|\chi_1\|_0 \quad (9)$$

令 $p_1 = \|\alpha_1\|_0, p_2 = \|\chi_1\|_0, p = \|\beta_1\|_0$, 则:

$$p \leq \sum_{i=1}^2 p_i \quad (10)$$

所以,在联合过完备库下对信号进行分解,其分解系数的最小零范数一定小于或等于 2 个分量信号在各自过完备库上的分解系数最小零范数之和。

由于:

$$f = D\beta_1 = (\theta \quad \varphi) \begin{pmatrix} \beta_{1\theta} \\ \beta_{1\varphi} \end{pmatrix} = \theta\beta_{1\theta} + \varphi\beta_{1\varphi} \quad (11)$$

$$f_1 = \theta\beta_{1\theta}, \quad f_2 = \varphi\beta_{1\varphi} \quad (12)$$

同时联合过完备库 D 中的子原子库 θ 包含信号 f_1 的特征信息,子原子库 φ 包含信号 f_2 的特征信息,所以通过式(11) - (12)可以重构复合信号以及各分量信号,实现信号的分离和重构。同时,由于联合过完备库更能反映信号的特征,所以在联合过完备库下能够获得更好的信号重构效果。由于 BP 算法分解精确度较好,所以本文采用 BP 算法实现信号在过完备字典上的稀疏分解。信号重构效果的评价标准一般是重建信号与原始信号的均方误差或信噪比。本文采用均方误差(Mean Square Error, MSE)作为评价标准,公式为: $MSE = \sum_{i=1}^N [f(i) - \tilde{f}_r(i)]^2 / N$ 。

3 仿真分析

选取信号 $f_1 = \sin(30t)$, $f_2 = \exp(\sin(200t))$ 组成复合信号 f ,同时构造子原子库 $\theta = \{\sin(it), i = 1, 2, \dots, 2000\}$; $\varphi = \{\exp(\sin(jt)), j = 1, 2, \dots, 2000\}$ 。根据联合过完备库构造的理论构建过完备库 D ,则 D 中的原子为 $\left\{ \bigcup_{i=1}^{2000} \sin(it), \bigcup_{j=1}^{2000} \exp(\sin(jt)) \right\}$ 。根据待分解信号的先验信息,参量 i 和 j 以间隔 1 均匀取值。原子个数为 4000,采样点数 1000。

为了研究联合过完备库的有效性以及构建联合过完备库的必要性,选取子原子库 φ 作为单一特征成分的过完备库,在该过完备库下对复合信号进行稀疏分解,得到如图 1(a) 所示的分解系数。从图 1(a) 中可以得到:信号在单一特征过完备库下的分解系数虽然具有稀疏性,但是由于原子之间的相关性较强,属于同一类型的原子,即原子的可分性较弱,所以不能直接用于信号类型的检测和分离。所以单一特征过完备库仅仅能够重构信号,而不能实现信号的分离和识别。图 1(b) 为在联合过完备库下的分解系数,图 1(b) 中的分解系数明显具有稀疏性,且分解系数中远大于零的值分别位于不同的子原子库中。由于原子库包含不同类型的原子,原子的可分性较强,所以复合信号在联合过完备库下的分解不但具有稀疏性还可以实现分量信号的识别和分离。

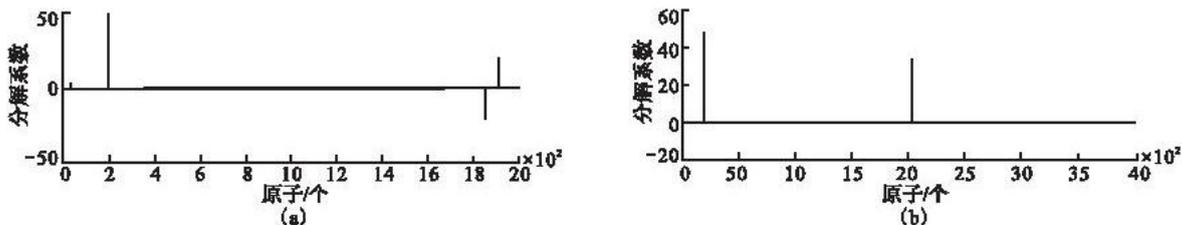


图 1 不同过完备库下复合信号的分解系数

Fig. 1 The decomposition coefficient on different overcomplete dictionary

表 1 为不同过完备库下的信号分解稀疏性比较。在单一特征过完备库下分量信号的稀疏性无法计算,复合信号的稀疏性(分解系数中绝对值大于零值的个数)近似为 5。在联合过完备库下,分量信号在各自的子原子库下的稀疏性均为 1,复合信号在由子原子库联合而成的过完备库下的稀疏性为 2。该结果与理论分析一致。

表 1 不同过完备库下信号的稀疏性

Tab. 1 The sparsity on different overcomplete dictionary

| 信号 | 单一特征过完备库下稀疏性 | 联合过完备库下稀疏性 |
|------------|--------------|------------|
| 分量信号 f_1 | 不具有稀疏性 | 1 |
| 分量信号 f_2 | 不具有稀疏性 | 1 |
| 复合信号 f | 5 | 2 |

图 2 为原始复合信号与不同过完备库下的重构信号对比图。从图中可以看出,2 种过完备库下重构的复合信号与原始复合信号相似,所以在 2 种过完备库下复合信号的重构效果均较好,但单一特征过完备库无法实现分量信号的分离。图 3 为在联合过完备库下重构的分量信号,从图中可以看出,重构的分量信号和原始分量信号几乎相同,所以联合过完备库能够非常精确地实现复合信号中分量信号的分离和重构。从图 2 和 3 中可以看出,联合过完备库不但能够很好地重构复合信号,对于各个分量信号也能实现较好地识别、分离和重构。

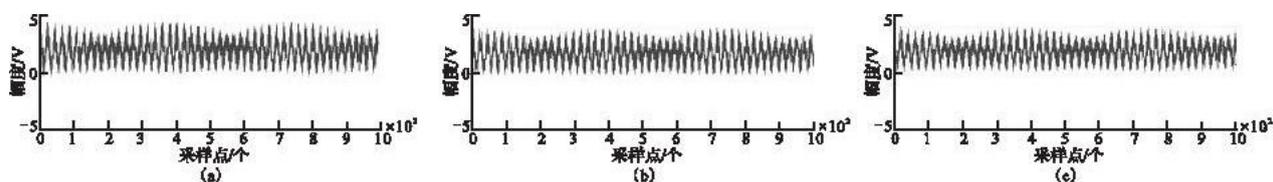


图2 不同过完备库下的重构信号

Fig. 2 The reconstructed signals on different overcomplete dictionary

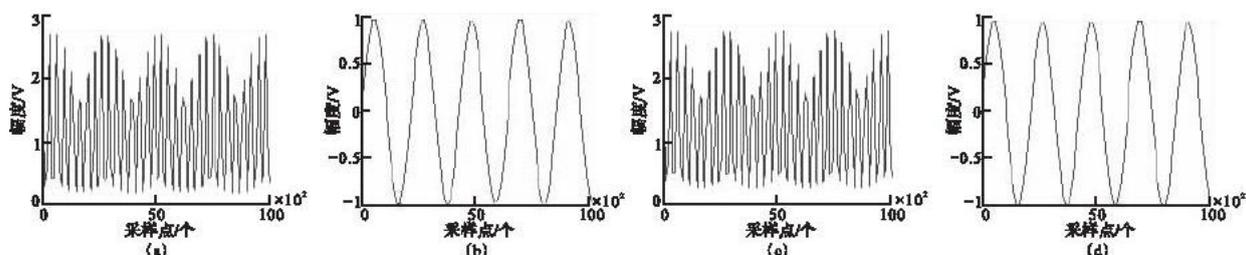


图3 分量信号的分离和重构

Fig. 3 The separation and reconstruction of single signals

表2 为不同过完备库下的重构误差。从表中可以看出,2种过完备库对复合信号的重构效果均较好,但联合过完备库下的重构误差小于单一特征过完备库下的重构误差;同时,分量信号在各子原子库上的重构误差也较小,重构效果也较好。

表2 不同过完备库下信号的重构误差

Tab. 2 The reconstruction errors on different overcomplete dictionary

| 信号 | 单一特征过完备库的重构误差 | 联合过完备库的重构误差 |
|------------|----------------------------|----------------------------|
| 分量信号 f_1 | 无法实现重构 | $7.761\ 9 \times 10^{-23}$ |
| 分量信号 f_2 | 无法实现重构 | $4.061\ 2 \times 10^{-28}$ |
| 复合信号 f | $8.785\ 9 \times 10^{-20}$ | $5.083\ 4 \times 10^{-28}$ |

4 结束语

本文提出了联合过完备库的思想,与单一特征成分的过完备库不同,联合过完备库包含更丰富的待分解信号的信息。通过理论分析和仿真实验可以得到:复合信号在联合过完备库下不但能够实现信号的重构,同时对于各分量信号也能够实现很好地识别、分离和重构,且重构效果较好。而复合信号在单一特征过完备库下虽然能够实现稀疏分解和重构,但是却无法实现分量信号的分离和重构。同时复合信号在联合过完备库上的稀疏性和重构效果均优于单一特征过完备库上的稀疏性和重构效果。因此,通过构建包含丰富待分解信号信息的联合过完备库不但可以实现复合信号的重构,对于构成复合信号的分量信号也可以实现分离和重构,且重构效果较好。

参考文献 (References):

[1] Joaquin Zepeda, Christine Guillemot, Ewa Kijak. The iteration - tuned dictionary for sparse representations [C]//MMSP'10. Saint - Malo, France: [s. n.], 2010: 93 - 98.

[2] 彭富强,于德介,刘坚. 一种基于多尺度线调频基的稀疏信号分解方法[J]. 振动工程学报, 2010, 23(3): 333 - 338. PENG Fuqiang, YU Dejie, LIU Jian. Sparse signal decomposition method based on multi - scale chirplet [J]. Journal of vibration engineering, 2010, 23(3): 333 - 338. (in Chinese)

[3] 邓承志,曹汉强. 多尺度脊波字典的构造及其在图像编码中的应用[J]. 图像图形学报, 2009, 14(7): 1273 - 1278. DENG Chengzhi, CAO Hanqiang. Construction of multiscale ridgelet dictionary and its application for image coding [J]. Journal of image and graphics, 2009, 14(7): 1273 - 1278. (in Chinese)

[4] 孙玉宝,肖亮,韦志辉,等. 基于 Gabor 感知多成份字典的图像稀疏表示算法研究[J]. 自动化学报, 2008, 34(11): 1379 - 1387. SUN Yubao, XIAO Liang, WEI Zhihui, et al. Sparse representations of images by a multi - component gabor perception dic-

- tionary[J]. *Acta automatica sinica*, 2008, 34(11): 1379 – 1387. (in Chinese)
- [5] Cédric Févotte, Simon J. Godsill. Sparse linear regression in unions of bases via bayesian variable selection[J]. *IEEE signal processing letters*, 2006, 13(7): 441 – 444.
- [6] Rémi Gribonval, Morten Nielsen. Sparse representations in unions of bases[J]. *IEEE transactions on information theory*, 2003, 49(12): 3320 – 3325.
- [7] 刘丹华, 石光明, 周佳社. 一种冗余字典下的信号稀疏分解新方法[J]. *西安电子科技大学学报*, 2008, 35(2): 228 – 232.
- LIU Danhua, SHI Guangming, ZHOU Jiashe. New method for signal sparse decomposition over a redundant dictionary[J]. *Journal of Xidian university*, 2008, 35(2): 228 – 232. (in Chinese)
- [8] 练秋生, 陈书贞. 基于混合基稀疏图像表示的压缩传感图像重构[J]. *自动化学报*, 2010, 36(3): 385 – 391.
- LIAN Qiusheng, CHEN Shuzhen. Image reconstruction for compressed sensing based on the combined sparse image representation[J]. *Acta automatica sinica*, 2010, 36(3): 385 – 391. (in Chinese)
- [9] Mohimani H, Babaie Zadeh M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed norm[J]. *IEEE trans on signal processing*, 2009, 52(1): 289 – 301.
- [10] Jokar S, Mehrmann V, Pfetsch M E, et al. Sparse approximate solution of partial differential equations[J]. *Applied numerical mathematics*, 2010, 60, (4): 452 – 472.
- [11] YANG Jingyu, PENG Yigang, XU Wenli, et al. Ways to sparse representation: an overview[J]. *Science in China series F: information sciences*, 2009, 52(4): 695 – 703.
- [12] Daubechies I. Time – frequency localization operators: a geometric phase space approach[J]. *IEEE trans on information theory*, 1998, 34(4): 605 – 612.
- [13] Coifman R R, Wickerhauser M V. Entropy – based algorithms for best – basis selection[J]. *IEEE trans on information theory*, 1992, 38(2): 713 – 718.

(编辑: 田新华)

Signal Separation and Reconstruction over Combined Over – complete Dictionary

YU Fu – ping¹, FENG You – qian², YANG Rong¹

(1. School of Air and Missile Defense, Air Force Engineering University, Xi'an 710051, China; 2. School of Science, Air Force Engineering University, Xi'an 710051, China)

Abstract: The combined over – complete dictionary is proposed based on the theory of sparse decomposition to achieve signal separation and reconstruction which is intensively studied and has been widely applied to many areas in signal processing. The combined over – complete dictionary is composed of many over – complete dictionaries with single feature, contains enough information of decomposed signal and makes complex signal more sparse, so separate signals can be separated and reconstructed much better by the basis pursuit algorithm. The simulation results show that the proposed idea of the combined over – complete dictionary is efficient to the complex signal reconstruction and separate signal separation and reconstruction.

Key words: combined over – complete dictionary; sparse decomposition; signal separation; sparsity; reconstruction error