

# 基于 MP 算法的语音信号稀疏分解

井爱雯<sup>1</sup>, 刘云<sup>2</sup>, 马轶丽<sup>3</sup>

JING Ai-wen<sup>1</sup>, LIU Yun<sup>2</sup>, MA Yi-li<sup>3</sup>

1.空军工程大学 理学院, 西安 710051

2.空军工程大学 导弹学院, 西安 710051

3.黄河科技学院, 郑州 450006

1.College of Science, Air Force Engineering University, Xi'an 710051, China

2.the Missile Institute, Air Force Engineering University, Xi'an 710051, China

3.Huanghe Science and Technology University, Zhengzhou 450006, China

E-mail: grgdh@163.com

**JING Ai-wen, LIU Yun, MA Yi-li. Speech signal sparse decomposition based on matching pursuit algorithm. Computer Engineering and Applications, 2009, 45(5): 144-146.**

**Abstract:** Speech signal sparse decomposition is a new method for decomposing audio data into a compact approximate representation. Speech signal sparse decomposition can be used in compression, denoising and recognition of speech signal. This paper researches speech signal sparse decomposition with matching pursuit algorithm. At last, the simulation experiments results show that speech signal sparse decomposition based on matching pursuit algorithm has better reconstructive precision and higher sparsity.

**Key words:** sparse decomposition; over-complete dictionary; matching pursuit algorithm

**摘要:** 语音信号稀疏分解是一种新的语音信号分解方法, 可以将语音信号分解为很简洁的近似表达形式。在语音信号稀疏分解的基础上, 可应用于语音处理的多个方面, 如语音压缩、语音去噪和语音识别等。研究利用 Matching Pursuit (MP) 算法实现语音信号的稀疏分解, 实验结果表明基于 MP 算法的语音信号稀疏分解具有较好的重建精度和较高的稀疏度。

**关键词:** 稀疏分解; 过完备库; 匹配追踪算法

**DOI:** 10.3778/j.issn.1002-8331.2009.05.042 **文章编号:** 1002-8331(2009)05-0144-03 **文献标识码:** A **中图分类号:** TP391.41

## 1 引言

传统的信号表示方法是基于“基”的展开, 如 Fourier 函数或者小波函数等。但这种建立在正交基上的信号分解有一定的局限性, 即当用这一类性质特定的基函数来表达任意信号时, 一旦基函数确定, 对于一个信号的展开也就随之确定, 往往不总能够达到好的稀疏表示效果, 尤其是对于时频变化范围很广的信号, 效果更差。一种更好的信号分解方式应该是根据信号的特点, 自适应地选择合适的基函数, 来完成信号的分解。因此近年来非正交分解引起人们越来越多的研究兴趣。为了实现信号更加灵活、简洁和自适应的表示, 在小波分析的基础上, Mallat 和 Zhang 于 1993 年提出基于过完备库 (over-complete dictionary) 的稀疏分解思想。

信号稀疏分解理论的基本思想是: 通过信号在过完备库 (又称冗余字典) 上的分解, 用来表示信号的基可以自适应地根据信号自身的结构特点灵活选取。分解的结果将可以得到信号非常简洁的表达 (即稀疏表示)。这个过程就称为稀疏分解 (Sparse decomposition)<sup>[1]</sup>。稀疏分解一方面可以实现数据压缩的高效性, 更重要的是可以利用字典的冗余特性捕捉原始信号

的自然特征。而语音信号稀疏分解也已经被成功应用于语音信号处理的多个方面, 如语音压缩、语音去噪和语音识别等。

至今已经发展了多种稀疏分解算法。常用的稀疏分解算法有匹配追踪 (Matching pursuit, MP) 算法<sup>[1-2]</sup>、基追踪 (Basis Pursuit, BP) 算法<sup>[2]</sup>、MOF (Method of Frames) 算法<sup>[3]</sup>、BOB (Basis Orthogonal Best) 算法<sup>[4]</sup>。BP 算法, 尤其 MP 算法是目前最常用的两种算法。

因此, 论文研究基于匹配追踪算法的语音信号稀疏分解方法, 这种算法根据语音信号的特点, 利用 Gabor 过完备库进行稀疏分解, 大大增强了信号的稀疏性, 从而提高了语音信号稀疏分解具有较好的重建精度和较高的稀疏度。

## 2 语音信号稀疏分解

基于过完备库的语音信号稀疏分解的数学描述是: 设集合  $D = \{g_i, i=1, 2, \dots, M\}$ , 且 Hilbert 空间  $R^N = \text{span}(D)$ ,  $M \gg N$ 。由于  $M \gg N$ , 则可称  $D$  为过完备库, 其元素称为原子。每一个原子都要能够充分代表语音信号的特点。对于语音信号  $f \in R^N$ , 在  $D$

**作者简介:** 井爱雯 (1962-), 女, 副教授, 主要研究领域为智能信息处理; 刘云 (1978-), 男, 助理工程师, 主要研究领域为后勤保障信息管理等; 马轶丽 (1978-), 黄河科技学院。

**收稿日期:** 2008-01-09 **修回日期:** 2008-05-08

中选取  $K(K \ll N)$  个原子对信号  $f$  作  $K$  项逼近:

$$f_k = \sum_{i \in I_k, |I_k|=K} \langle f, g_i \rangle g_i \quad (1)$$

其中  $I_k$  是  $g_i$  的下标集合。定义逼近误差:

$$\sigma = \inf_{f_k} \|f - f_k\| \quad (2)$$

从稀疏逼近的角度出发, 我们希望在满足条件(2)的前提下, 从各种可能的组合中, 挑选出分解系数最为稀疏的一组原子, 从而得到语音信号非常简洁的表达, 以利于后续的处理工作。

如果  $D$  是 Hilbert 空间  $R^N$  的一组正交基, 如何得到信号  $f$  的最佳  $K$  项逼近是一件显而易见的事, 即保留与  $f$  的内积  $\langle f, g_i \rangle$  最大的  $K$  个基, 然而对于过完备库  $D$  来说, 这是一个 NP(Non-deterministic polynomial time) 难问题<sup>[5]</sup>。

要找到最稀疏的信号表示, 等同于解决下述问题:

$$\min \|c\|_0, \text{ s.t. } f = \sum_{i=1}^K c_i g_i \quad (3)$$

其中  $\|c\|_0$  是序列  $c_i (i=1, 2, \dots, K)$  中非零项的个数。由于这是一个 NP 难问题, 为解决这一难点, Chen, Donoho 和 Saunders 提出解决下述稍有差别的问题:

$$\min \|c\|_1, \text{ s.t. } f = \sum_{i=1}^K c_i g_i \quad (4)$$

于是问题转化为在  $l_1$  范数的优化问题, 被称为 Basis Pursuit 或 BP。这是一个较为简单的问题, 可以通过线性规划的方法解决<sup>[2]</sup>。

### 3 基于 MP 的语音信号稀疏分解算法

尽管使用了线性规划方法, 基追踪法由于要在原子的不同组合中寻求满足算式(4)成立的解, 其计算量仍然是巨大的, 而采用能够求解局部最优的贪婪算法以取代求  $\|c\|_0$  的全局最优, 可以降低计算的复杂度。MP 就是这样一个迭代的贪婪算法, 它在每一次迭代过程中, 从过完备库中选择最能匹配信号结构的一个原子而构建的一种逼近过程<sup>[1, 5-6]</sup>。

基于 MP(Matching pursuit) 算法的语音信号分解过程如下<sup>[1]</sup>:

对于语音信号  $f$ , 首先从过完备库  $D$  中选出与它最匹配的原子  $g_{i_0}$ , 并且满足以下条件:

$$|\langle f, g_{i_0} \rangle| = \sup_{i \in I} |\langle f, g_i \rangle| \quad (5)$$

其中,  $I$  为参数组  $i$  的集合。于是, 语音信号分解为在最佳原子  $g_{i_0}$  上的分量和残余分量两部分, 即:

$$f = \langle f, g_{i_0} \rangle g_{i_0} + R^1 \quad (6)$$

其中,  $R^1$  是用最佳原子匹配后余留下的残留部分。对最佳匹配后的残余不断进行上述式(6)分解, 经过  $l+1$  次分解后, 即可得到如下等式:

$$R^l = \langle R^l, g_i \rangle g_i + R^{l+1} \quad (7)$$

其中,  $g_i$  满足下式:

$$|\langle R^l, g_i \rangle| = \sup_{i \in I} |\langle R^l, g_i \rangle| \quad (8)$$

经过上述  $K$  步分解后, 信号最终被分解为:

$$f = \sum_{l=0}^{K-1} \langle R^l, g_i \rangle g_i + R^K \quad (9)$$

式中  $R^K$  为  $K$  步分解后的残余(逼近误差)。由于每一步分解都满足式(8), 所以每一步分解的残余会随着分解步数  $l$  的增加而迅速衰减, 文献[1]已经证明, 在信号满足长度有限的条件下, 范数  $\|R^l\|$  随着  $l$  的增大而指数衰减至 0。

### 4 实验分析

在对语音信号进行稀疏分解时, 根据语音信号的特点, 利用 Gabor 原子对点奇异信号较强的表示能力, 选用 Gabor 过完备库<sup>[5]</sup>。因此过完备库  $D = \{g_\gamma\}_{\gamma \in \Gamma}$  选择由 Gabor 原子构成, 每个 Gabor 原子由一个高斯窗函数构成, 即

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) \cos(vt+w) \quad (10)$$

式中  $g(t) = e^{-\pi t^2}$  是高斯窗函数,  $\gamma = (s, u, v, w)$  是时频参数, 其中  $s$  为尺度因子,  $u$  为位移因子,  $v$  为原子频率,  $w$  为原子相位。将时频参数离散化可得  $\Gamma = \{(\alpha^j, p\alpha^j \Delta u, k\alpha^j \Delta v, i\Delta w)\}$ , 其中  $\alpha = 2, \Delta u = 1/2, \Delta v = \pi, \Delta w = \pi/6, 0 \leq j \leq \lfloor \log_2 N \rfloor, 0 \leq k \leq 2^{j+1}, 0 \leq i \leq 12$ 。该原子库中的原子形如图 1 所示。

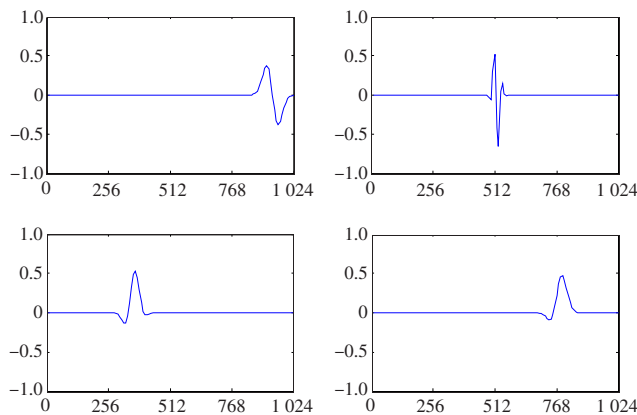


图 1 Gabor 过完备库中原子形状 ( $N=1024$ )

以一维语音信号为例, 取  $N=1024$ , 分解原子数  $K=128$ 。图 2(a) 为实验用的原始信号, 对其进行基于 MP 的语音信号稀疏分解算法, 然后利用分解得到的  $K$  个系数重构信号。图 2(b) 为 MP 算法分解后, 再利用 128 个原子进行重构后的结果。

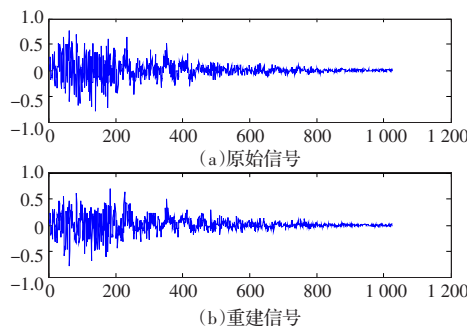


图 2 基于 MP 的语音信号稀疏分解

经过计算可得基于 MP 的语音信号稀疏分解算法的逼近误差  $\|f - f_k\|$  为 1.095 8。也就是说, 在利用 128 个稀疏系数对长度为 1024 的语音信号进行压缩的时候, 重建得到的逼近误差非常小, 为原始信号能量的 0.001 27 倍。因此, 在 Gabor 过完备库上, 基于 MP 算法的语音信号稀疏分解具有较好的重建精

度和较高的稀疏度。

## 5 总结与展望

在对语音信号研究和工程应用中,对语音信号分解和表达是一个根本性的问题,直接影响到后续的压缩、识别等处理,因而具有非常重要的意义。但正交分解有一定的局限性,因此近年来非正交分解引起人们越来越多的研究兴趣。本文研究利用 MP 算法实现语音信号的稀疏分解,实验结果表明基于 MP 算法的语音信号稀疏分解具有较好的重建精度和较高的稀疏度。

非线性逼近理论给出令人信服的实例,证明过完备库下的逼近优于已知的正交基<sup>[7]</sup>;稀疏分解目前所涉及的应用已扩展到很多方面<sup>[8]</sup>,如本文研究的语音信号。目前,需要研究的问题仍很多,除算法的改进外,如何构造更加适合语音信号特点的过完备库也是研究的重点。

### 参考文献:

[1] Mallat S,Zhang Z.Matching pursuit with time-frequency dictionary-

ies[J].IEEE Trans on Signal Processing,1993,41(12):3397-3415.

- [2] Chen S,Donoho D,Saunders M.Atomic decomposition by basis pursuit[J].SIAM J Sci Comput,1999,20:33-61.
- [3] Candes E J,Romberg J.Practical signal recovery from random projections[Z].2005-01.
- [4] Coifman R,Wickerhauser M.Wickerhauser,Entropy-based algorithms for best-basis selection[J].IEEE Transactions on Information Theory,1992,38:713-718.
- [5] 张春梅,尹忠科,肖明霞.基于冗余字典的信号过完备表示与稀疏分解[J].科学通报,2006(6):628-633.
- [6] Davis G,Mallat S,Avellaneda M.Adaptive greedy approximation[J].Constr Approx,1997,13(1):57-98.
- [7] Candes E,Donoho D.New tight frames of curvelets and optimal representations of objects with C2 singularities[R].Stanford University,2002.
- [8] Berg A,Mikhael W.A survey of mixed transform techniques for speech and image coding[C]//Proc IEEE Intern Symp Circ Syst,1999,4:106-109.

(上接 85 页)

- [8] Beasley J E,Chu P C.A Genetic Algorithm for the set covering problems[J].European Journal of Operational Research,1996:392-404.
- [9] Mitchell M.An introduction to genetic algorithms[M].Cambridge, MA:MIT Press,1996.

[10] AT&T IP Backbone[EB/OL].(2001).<http://www.ipservices.att.com/backbone/>.

- [11] Chen GL,Wang XF,Zhang ZQ.Genetic Algorithm and its application[M].Beijing:People's Post and Telecommunications Press,1996.
- [12] Cui J-H.Modelling the spatial properties of multicast group members,#030011,UCLA CSD Technical Report[R].2003.

(上接 131 页)

- [15] 张广兴,张大方,谢高岗,等.Internet 城域出口链路流量测量与特征分析[J].电子学报,2007.
- [16] 邵立松,窦文华.自相似网络通信量模型研究综述[J].电子与信息学报,2005,27(10):1671-1676.
- [17] Internet World Stats.Internet usage and population statistics[EB/

OL].(2005).<http://www.internetworldstats.com/stats.htm>.

- [18] Karagiannis T,Broido A,Brownlee N,et al.Is P2P dying or just hiding?"[C]//Proc of GLOBECOM 2004,Dalls,TX,2004.
- [19] Pacheco J C R,Roman D T.Accuracy of time-domain algorithms for self-similarity:an empirical study[C]//The 15th International Conference on Computing,2006:379-384.

(上接 143 页)

## 5 结论

提出一种新的机制来保证移动计算环境下已复制数据的一致性(单副本可串行化)并且所有副本最终收敛于一致性状态(最终一致性)。该协议对已复制数据库对象的读操作或写操作等临时信息进行探索,并且运用正确性标准的可串行化。除了使用联合正确性标准之外,该协议还释放隔离属性,采取随便读、随便写方式,允许更高的数据可用性。此外,事务操作仅仅执行一次,服务器不必为所有执行操作记录日志。并且提出的方法也减少了服务器间信息的交换量。

### 参考文献:

- [1] Bernstein P A,Hadzilacos V,Goodman N.Concurrency control and recovery in database systems[M].S.I.:Addison-Wesley,1987.
- [2] Saito Y,Shapiro M.Optimistic replication[J].ACM Computing Sur-

veys,2005.

- [3] Barreto J,Ferreira P.An efficient and fault-tolerant update commitment protocol for weakly connected replicas[C]//Proceedings of the EUROPAR,Lisboa,Portugal,2005.
- [4] Cetintemel U,Keleher P,Franklin M.Support for speculative update propagation and mobility in deno[C]//Proceedings of the The 21st IEEE International Conference on Distributed Computing Systems (ICDCS01),2001.
- [5] Maya R,Anthony L.Decentralized weighted voting for P2P data management[C]//Proceedings of the 3rd ACM International Workshop on Data Engineering for Wireless and Mobile Access,San Diego,CA,USA,(MobiDe 03),2003.
- [6] Casanova M A.The concurrency problem of database systems[M]//Lectures Notes in Computer Science,1981.
- [7] Terry D,Demers A,Petersen K,et al.Managing update conflicts in bayou[Z].a Weakly Connected Replicated Storage System,1995.