

利用 FFT 实现语音信号稀疏分解的改进算法

刘 强, 尹忠科, 王建英

LIU Qiang, YIN Zhong-ke, WANG Jian-ying

西南交通大学 信息科学与技术工程学院, 成都 610031

School of Science & Technology of Information, Southwest Jiaotong University, Chengdu 610031, China

E-mail: liuqiang110810@mars.swjtu.edu.cn

LIU Qiang, YIN Zhong-ke, WANG Jian-ying. Improvement of MP based voice signal sparse decomposition by using FFT. Computer Engineering and Applications, 2007, 43(26): 74-75.

Abstract: After studying of Matching Pursuit (MP) based voice signal sparse decomposition, a new improvement of sparse decomposition algorithm is presented based on analysis of structure property of the over-complete atom dictionary used in voice signal sparse decomposition. According to the character of the voice signal, the new algorithm based on the FFT algorithm that reduces the range of the searching atoms, can not only heighten the speed of the decompositions, but also reconstruct the speech signal with less atoms. Finally the experimental results show that the performance of the proposed algorithm is good.

Key words: signal processing; sparse representation; sparse decomposition; Matching Pursuit (MP); FFT

摘 要: 研究基于 Matching Pursuit (MP) 方法实现的语音信号稀疏分解问题, 通过对语音信号稀疏分解中使用的过完备原子库结构特性的分析, 提出了一种改进的信号稀疏分解算法。该算法针对语音信号的特点, 以 FFT 算法实现的稀疏分解为基础缩小了原子的搜索范围, 从而不仅进一步提高分解速度, 还能以更稀疏的形式表示语音信号。算法的有效性为实验结果所证实。

关键词: 信号处理; 稀疏表示; 稀疏分解; Matching Pursuit (MP); FFT

文章编号: 1002-8331(2007)26-0074-02 文献标识码: A 中图分类号: TN911.72

Mallat 和 Zhang 首先于 1993 年提出了信号在过完备库 (over-complete dictionary) 上分解的思想^[1]。通过在过完备库上的分解, 用来表示信号的基可以自适应地根据信号本身的特点灵活选取。分解的结果, 将得到信号的一个非常简洁的表达 (即: 稀疏表示 sparse representation)。而得到信号稀疏表示的过程称为信号的稀疏分解 (sparse decomposition)。由于信号的稀疏表示的优良特性, 信号稀疏表示已经被应用到信号处理的许多方面, 如信号去噪^[2]、信号编码^[3]和识别^[4]等。其中, 在信号时频分布研究方面的应用特别值得关注^[5]。但是目前信号的稀疏分解在信号处理中的实际应用很难被推广而产业化。阻碍信号稀疏分解研究及应用发展的关键因素是信号稀疏分解的计算量十分巨大, 计算时间在现有计算条件下令人无法忍受。国内有研究人员指出, 信号长度为 1024 采样点时, 信号的稀疏分解的难度将十分巨大^[6]。针对此问题, 本文研究基于 MP 的信号稀疏分解, 通过分析过完备原子库的形成过程, 在 FFT 算法^[7]基础上通过在由已选取的原子所构建的更小的原子库中搜索局部最佳原子来解决耗时高的问题。同时以更少的原子表示信号, 有利于信号的压缩编码。

1 基于 MP 的信号稀疏分解

基于 MP 的稀疏分解, 是目前信号稀疏分解的最常用方法。假设研究的信号为 f , 信号长度为 N 。若将信号分解在一组完备正交的基上, 则这组基的数目应为 N 。由于基的正交性, 因而基在由信号所组成的空间中的分布是稀疏的, 从而, 信号的能量在分解以后将分散分布在不同的基上。这种能量分布的分散最后将导致用基的组合表示信号时表达的不简洁性, 即信号表示不是稀疏的。非稀疏的表示, 不利于信号的处理, 如识别和压缩等。为了得到信号的稀疏表示, 基的构造必须使得基在信号组成的空间中足够的密。由此, 基的正交性将不再被保证, 所以此时的基也不再是真正意义上的基了, 而改称为原子。由这些原子组成的集合, 是过完备的, 被称为过完备库 (over-complete dictionary of atoms)。信号在过完备库上的分解结果一定是稀疏的^[1]。

设 $D = \{g_r\}_r$ 为信号稀疏分解用到的过完备库, g_r 为由参数组 γ 定义的原子。用不同的方法构造原子, 参数组 γ 所含有的参数及参数个数也不一样。原子 g_r 的长度与信号长度相同, 但原子应作归一化处理, 即 $\|g_r\| = 1$ 。 I 为参数组 γ 的集合。由库的过完备性可知, 参数组 γ 的个数应远远大于信号的长度,

基金项目: 四川省科技攻关计划 the Key Technologies R&D Program of Sichuan Province, China under Grant No.04GG021-020-5; 教育部留学回国人员科研启动基金 The Project-sponsored by SRF for ROCS, SEM No.教外司[2004]517号; 四川省应用基础研究项目 (No. 04JY029-059-2); 西南交通大学校基金。

作者简介: 刘强 (1982-), 男, 硕士研究生, 研究方向: 信号与信息处理。

即若用 P 表示过完备库 $D=\{g_r\}_r$ 中原子的个数, 则 P 应远远大于信号长度 N 。MP 方法分解信号过程如下^[1]:

首先从过完备库中选出与待分解信号最为匹配的原子 g_{γ_0} , 其满足以下条件:

$$|f, g_{\gamma_0}| = \sup_r |f, g_r| \quad (1)$$

信号因此可以分解为投影到最佳原子 g_{γ_0} 上的分量和残余两部分, 即为:

$$f = f, g_{\gamma_0} g_{\gamma_0} + R^1 f \quad (2)$$

其中 $R^1 f$ 是用最佳原子对原信号进行最佳匹配后的残余。对最佳匹配后的残余可以不断进行上面同样的分解过程, 可得:

$$R^k f = R^k f, g_{\gamma_k} g_{\gamma_k} + R^{k+1} f \quad (3)$$

其中 g_{γ_k} 满足:

$$|R^k f, g_{\gamma_k}| = \sup_r |R^k f, g_r| \quad (4)$$

由公式 2) 和 3) 可知, 经过 n 步分解后, 信号被分解为:

$$f = \sum_{k=0}^{n-1} R^k f, g_{\gamma_k} g_{\gamma_k} + R^n f \quad (5)$$

其中 $R^n f$ 为原信号分解为 n 个原子的线性组合后, 用这样的线性组合表示信号所产生的误差。由于每一步分解中, 所选取的最佳原子满足公式 4), 所以分解的残余 $R^n f$ 随着分解的进行, 迅速地减小。已经证明^[1], 在信号满足长度有限的条件下 (对数字信号而言, 这是完全可以而且一定满足的), $R^n f$ 随 n 的增大而指数衰减为 0。从而信号可以分解为:

$$f = \sum_{k=0}^{\infty} R^k f, g_{\gamma_k} g_{\gamma_k} \quad (6)$$

事实上, 由于 $R^n f$ 的衰减特性, 一般而论, 用少数的原子与信号长度相比较而言) 就可以表示信号的主要成分, 即:

$$f \approx \sum_{k=0}^n R^k f, g_{\gamma_k} g_{\gamma_k} \quad (7)$$

其中, $n \ll N$ 。公式 7) 和条件 $n \ll N$ 集中体现了稀疏表示的思想。

虽然基于 MP 的稀疏分解是目前信号稀疏分解的最常用方法, 也是几乎所有算法中速度最快的, 但和别的信号稀疏分解方法一样, 存在的关键问题仍是计算量十分巨大。在基于 MP 的信号稀疏分解中, 每一步都要完成信号或信号分解的残余在过完备库中的每一个原子上的投影计算。按式 4) 所要求, 选取每个最佳原子要进行的内积计算 $R^k f, g_{\gamma_k}$ 是在一个很高维 (N 维) 空间的内积计算, 而且要进行很多次选取, 这是基于 MP 的信号稀疏分解计算量巨大的根本原因所在。因此减小计算量的一个途径就是缩小搜索最佳原子的范围, 下面的改进算法对语音信号稀疏分解计算速度的提高有明显效果。

2 改进算法

通过深入研究过完备原子库^[6]的结构和语音信号的特点, 针对语音信号具有的类似余弦信号的周期特性 (如图 1), 本文提出以下改进算法: 利用 FFT 在过完备原子库中选取了最佳原子后, 以这个原子为基原子, 通过形成过完备原子库的离散化参数 p 和 w^{Δ} 改变基原子的位置和相位。这样就可以由具有相同尺度, 不同位置和相位的原子组成一个子库 D , 这个新原子库比过完备原子库小得多。正是由于语音信号具有的类周期特性, 因此可以按照 FFT 算法的流程, 在得到一个全局最佳原子后, 以这个最佳原子为基原子构造子库 D , 再将上一次匹配所得到的信号残差在子库中进行 MP 分解。由图 1 可知, 若由

FFT 算法选取的最佳原子与语音信号中具有周期特性的部分相匹配, 那么残差信号将在子库中找到与其很好匹配的原子。具体做法为: (1) 利用 FFT 算法选取最佳原子 g_{γ} , 并计算信号或信号残差在原子 g_{γ} 上投影后的残差; (2) 以原子 g_{γ} 为基原子生成子库 D , 让残差信号在子库 D 上做基于 MP 的稀疏分解, 这一步也是逐次逼近信号的迭代过程, 可设定总迭代次数或以残差信号能量为结束迭代的条件; (3) 重复以上两步即可完成信号在过完备原子库上的初步稀疏分解。如果初步分解的效果令人满意, 即可结束语音信号的稀疏分解; 如果初步分解的效果不令人满意, 以原来基于 MP 的稀疏分解方法继续进行分解, 直到得到满意的稀疏分解结果。

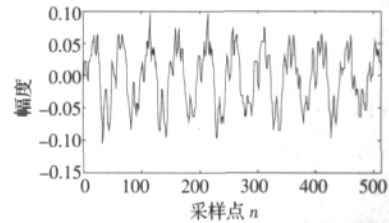


图 1 具有周期特性的语音信号

分析改进算法可知, 原来在过完备原子库上的投影工作现在大部分将在一个很小的子库中完成。因为搜索与信号或残差信号相匹配的原子的范围缩小, 所以要达相同的重建语音信号质量所需的时间更少。另外由改进算法中构建子库的方法不难得出这个结论: 因为子库中的所有原子都具有相同的尺度, 只是所在位置和相位不同, 所以若对信号进行压缩编码, 改进算法需要处理的数据量将更少。

3 实验结果与分析

实验中采用长度为 512 经归一化处理的实际语音信号。过完备原子库的构造方法按文献[3]。若将原始的基于 FFT 算法的稀疏分解记为算法 1, 改进后的算法记为算法 2。综合多次实验的结果, 改进的 FFT 算法在计算速度上比原始的 FFT 算法提高 20%到 50%不等, 见表 1。

表 1 语音信号稀疏分解计算速度比较

	算法 1 所需时间/s	算法 2 所需时间/s
语音信号 1	1.407 8	0.668 9
语音信号 2	1.070 3	0.767 7
语音信号 3	1.123 5	0.786 2
语音信号 4	1.384 1	0.903 6
语音信号 5	1.222 2	0.692 5
语音信号 6	1.726 8	0.940 9

由于子库的出现, 使得分解过程中所选取的原子中有很多都具有相同尺寸, 从原子库的构造看, 这些原子具有相同的尺度因子、频率因子, 不同的只是平移因子和相位因子。

从压缩编码角度看, 这有利于简化编码过程和减少编码的数据量。对 512 长的语音信号, 根据文献[3]中原子参数的量化方案可知, 若频率、相位和位移因子可分别用 4 位、6 位、9 位编码, 语音信号在原子库的投影值可用 10 位编码。表 2 比较了算法 1 和算法 2 分解相同语音信号达到相同信噪比时编码所需二进制位数。

由表 2 看出, 改进算法分解后的数据量比原始 FFT 算法数据量有明显的减少, 最少为原始 FFT 算法的 63%。

(下转 78 页)

多运行 $n-1$ 其中 $(n-2)$ 步就能得到 T 即可。由定理 2 的充分性证明知： $\zeta \in \mathcal{S}$ 中词 T 可以从初始语言 I 中的词 n_i 开始，运用粘贴规则序列 r_1, r_2, \dots, r_{j-1} ，其中 $r_s \in \{n_i, n_{i_2} \dots n_{i_s}, n_{i_{s+1}}\}$ ， $s=1, 2, \dots, j-1$ ，得到词 $n_{i_1} n_{i_2} \dots n_{i_j}$ 。

由定理 1 知： $|T| \leq n$ ，所以 $|T| \leq j+1 \leq n$ ，即 $j \leq n-1$ 。因此粘贴系统 S 至多进行 $(n-2)$ 步就可得到 T 。

由上述定理可知，对于粘贴系统 S 来说，它模拟的有向图 G 与其产生的语言 $\zeta \in \mathcal{S}$ 是一一对应的。

下面给出一个实例（图 2）来说明粘贴系统模拟哈密顿路问题的整个过程。模拟图 2 的粘贴系统的字母表 $V = \{v_1, v_{12}, v_{21}, v_{22}, \dots, v_{n1}, v_{n2}\}$ ，初始语言 $I = \{n_1, n_2, \dots, n_n\}$ ，规则 R 同上。第 1 步粘贴运算：

$$\begin{aligned} & (n_1, n_2) \triangleright (n_1, n_2), (n_1, n_3) \triangleright (n_1, n_3), (n_2, n_1) \triangleright (n_2, n_1), (n_2, n_3) \\ & \triangleright (n_2, n_3), (n_3, n_4) \triangleright (n_3, n_4), (n_3, n_5) \triangleright (n_3, n_5), (n_4, n_5) \triangleright (n_4, n_5), \\ & (n_5, n_2) \triangleright (n_5, n_2), (n_5, n_3) \triangleright (n_5, n_3), (n_5, n_4) \triangleright (n_5, n_4) \end{aligned}$$

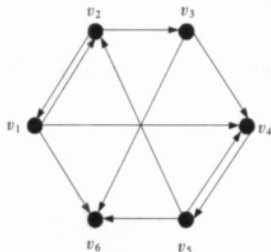


图 2 6 个顶点有向哈密顿路

令 $I_1 = \{n_i | (n_i, n_j) \triangleright (n_i, n_j), i, j=1, 2, \dots, 6 \text{ 且 } i \neq j\}$ 。这样，经过第 1 步粘贴运算后得到 $\alpha(1) = I_1$ 。根据粘贴规则 $R \in T_1, T_2$ ，第 2 步粘贴运算可分为 4 类：

- (1) $|T_1|=|T_2|=1$ 。这类运算与第 1 步粘贴中的运算相同，它不会给 $\alpha(1)$ 中增加新词。
- (2) $|T_1|=1, |T_2|=2$ 。例：词 n_1 与第 1 步粘贴得到的词 $n_2 n_3$ 在规则下发生运算 $n_1, n_2 n_3 \triangleright (n_1 n_2 n_3)$ 。
- (3) $|T_1|=2, |T_2|=1$ 。例：第 1 步粘贴得到的词 $n_2 n_3$ 与词 n_4 在规则 $r \in R$ 发生运算 $n_2 n_3, n_4 \triangleright (n_2 n_3 n_4)$ 。

(上接 75 页)

表 2 语音信号稀疏分解编码所需二进制位数比较

	算法 1 所需二进制位数	算法 2 所需二进制位数
语音信号 1	1 450	915
语音信号 2	1 450	926
语音信号 3	1 450	1 005
语音信号 4	1 450	1 035
语音信号 5	1 450	1 117
语音信号 6	1 450	1 131

图 2 为一段实际语音信号和稀疏分解的重构信号，图 2 (b) 为基于原始 FFT 算法用 50 个原子的重构信号；图 2 (c) 为与原始基于 FFT 的稀疏分解算法重建语音信号相同信噪比情况下改进算法的重构信号。从图 2 中可看出，两图基本一样，只是某些细节有出入。由于子库的出现，可以将信号在平移原子库中选取的原子视为一个原子，这样改进算法在相同信噪比下所需原子数更少，因此改进算法更为稀疏。

(收稿日期：2007 年 1 月)

参考文献：

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

(4) $|T_1|=2, |T_2|=2$ 。例：第 1 步粘贴得到的词 $n_1 n_2$ 与词 $n_3 n_4$ 在规则 $r \in R$ 下发生运算 $n_1 n_2, n_3 n_4 \triangleright (n_1 n_2 n_3 n_4)$ 。

第 3 步，第 4 步粘贴运算与第 2 步粘贴类似。

下面看一下第 5 步也就是最后 1 步粘贴得到的结果。长度为 6 的词有： $n_1 n_2 n_3 n_4 n_5 n_6, n_1 n_4 n_5 n_2 n_3 n_6$ 和 $n_3 n_4 n_5 n_2 n_1 n_6$ 。由推论 2：图 2 中有三条不同的哈密顿路，再由定理 2 或推论 1 知：这三条哈密顿路分别为 $n_1 n_2 n_3 n_4 n_5 n_6, n_1 n_4 n_5 n_2 n_3 n_6$ 和 $n_3 n_4 n_5 n_2 n_1 n_6$ 。

5 结论

本文将 DNA 计算中的形式模型应用于图论中的有向哈密顿路完全问题。利用粘贴系统的高度并行性，对有向哈密顿问题进行了模拟。通过构造的粘贴系统所产生语言的性质对有向哈密顿路问题进行了分析，给出了有向图存在哈密顿路的充要条件。(收稿日期：2007 年 5 月)

参考文献：

[1] Adleman L. Molecular computation of solutions to combinational problems[J]. Science, 1994, 266: 1021- 1024.
 [2] Roweis S, Winfree E, Burgoyne R, et al. A sticker based architecture for DNA computation[C]//Baum E B. Proc 2nd Annual Meeting on DNA Based Computers, Princeton, 1999: 1- 27.
 [3] Kari L, Paun G H, Rozenberg G, et al. DNA computing, sticker systems and universality[J]. Acta Inform, 1998, 33: 401- 420.
 [4] Paun G, Rozenberg G. Sticker system[J]. Theoretical Computer Science, 1998, 204: 183- 203.
 [5] Paun G, Rozenberg G, Salomaa A. DNA computing—new computing paradigms[M]. Berlin: Springer, 1998.
 [6] Lin Wen-bin, Xu Jin. A DNA solution to weighted Hamilton path problem[J]. System Engineering and Electronics, 2002, 6: 99- 102.
 [7] Gao Lin, Ma Rui-nian, Xu Jin. DNA algorithm to the directed shortest Hamilton problem[J]. Systems Engineering and Electronics, 2002, 8: 102- 105.
 [8] 王淑栋, 许进, 刘文. 基于剪裁系统的有向哈密顿路问题分析[J]. 电子学报, 2005, 3: 774- 777.

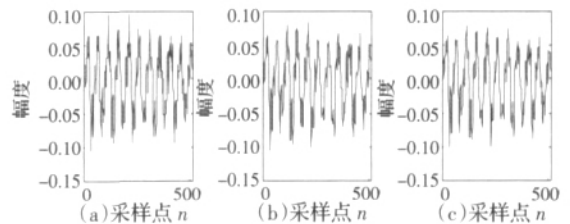


图 2 信号和稀疏分解后重建的信号

ies[J]. IEEE Trans on Signal Processing, 1993, 41(12): 3397- 3415.
 [2] 张文耀. 基于匹配跟踪的低速率语音编码研究[D]. 中国科学院研究生院, 软件研究所, 2002.
 [3] Arthur P L, Philipos C L. Voiced/unvoiced speech discrimination in noise using gabor atomic decomposition[C]//Proc of IEEE ICASSP, Hong Kong, April 2003, 1: 820- 828.
 [4] 邹红星, 周小波, 李衍达. 时频分析: 回溯与前瞻[J]. 电子学报, 2000, 28(9): 78- 84.
 [5] 邵君, 尹忠科, 王建英. 利用 FFT 实现基于 MP 信号稀疏分解的改进算法[J]. 西南交通大学学报, 2006, 41(4): 466- 470.
 [6] 邵君, 尹忠科, 王建英. 信号稀疏分解中过完备原子库的集合划分[J]. 铁道学报, 2006, 28(1): 68- 71.