

# 自适应压缩感知的语音压缩重构算法研究\*

毛 静<sup>1,2</sup>, 尹鹏亮<sup>1,2</sup>, 马 琦<sup>1,2</sup>, 张斌珍<sup>1,2</sup>

(1. 中北大学 电子测试国家重点实验室, 山西 太原 030051;

2. 中北大学 仪器科学与动态测试教育部重点实验室, 山西 太原 030051)

**摘 要:** 根据传统语音信号的处理过程和语音信号的特征, 提出了利用自适应冗余字典 KSVD 算法、自适应观测矩阵和 SAMP 重构算法的压缩重构方法, 通过仿真分析, 并与普通压缩感知对比平均帧重构信噪比、相对误差, 验证了压缩感知自适应算法的优越性。

**关键词:** 压缩感知; 语音信号; 自适应算法

中图分类号: TN912

文献标识码: A

文章编号: 0258-7998(2015)02-0072-03

DOI: 10.16157/j.issn.0258-7998.2015.02.016

## Research on algorithm of speech compression and recovery based on adaptive compressed sensing

Mao Jing<sup>1,2</sup>, Yin Pengliang<sup>1,2</sup>, Ma Qi<sup>1,2</sup>, Zhang Binzhen<sup>1,2</sup>

(1. National Key Laboratory for Electronic Measurement Technology, North University of China, Taiyuan 030051; China;

2. Key Laboratory of Instrumentation Science & Dynamic Measurement of Ministry of Education, North University of China, Taiyuan 030051, China)

**Abstract:** According to the speech characteristics and traditional signal processing, this paper proposes that the adaptive algorithm is joined in compressed sensing, combined with redundant dictionary KSVD adaptive algorithm, adaptive observation matrix and SAMP reconstruction algorithm. By simulation analysis and comparison with the general compressed sensing on the average frame reconstruction SNR and relative error, it is ended to verify the superiority of adaptive compressed sensing algorithm.

**Key words:** compressed sensing; speech signal; adaptive algorithm

### 0 引言

传统的奈奎斯特采样定理要求采样率高于信号最高频率的两倍, 来实现信号的采集、压缩和恢复。但随着采集数据和频率的急剧增加, 压缩过程中丢弃了绝大部分采集数据, 只保留了小部分来恢复信号, 导致增加大量不必要的存储和传输设备。Donoho、Candès 等人提出的压缩感知理论很好地解决了这个难题<sup>[1-2]</sup>。压缩感知将可稀疏的信号通过观测从高阶矩阵线性投影为低阶, 信号的采集和压缩在此过程同时进行, 最后高概率精确地重建原始信号。在语音信号的处理过程中, 模拟信号的高采样率必然会产生海量的冗余数据, 根据语音信号的可压缩性, DCT 变换域以及小波变换的稀疏特性, 满足压缩感知理论的先验条件<sup>[3-4]</sup>。压缩感知的理论包含三项内容: 选择信号稀疏基, 设计观测矩阵和重构原信

号。其中信号是否具有稀疏性是前提条件; 设计观测矩阵是实现的重要部分, 不仅关系到压缩和采样速率的快慢, 而且影响重构信号的质量; 重构原始信号是压缩感知理论的核心, 决定着恢复原信号质量的好坏。

### 1 语音自适应压缩感知的设计

信号在稀疏矩阵变换下越稀疏, 重构时所用的观测数目就越少, 重建信号的效率就越高。由于语音信号具有短时稳态性, 因此基于压缩感知的语音处理要先对信号进行编帧, 然后利用随机观测矩阵对每帧信号压缩采样, 最后逐帧恢复原信号。

实验仿真选用的一段女生声音内容为“hello, 酷狗”, 共有 44 100 个样点, 采样率为 22.05 kS/s<sup>[5-6]</sup>。

#### 1.1 自适应信号稀疏

DCT 是语音信号处理中应用较多的稀疏方法, 具有很强的能量集中功能。但它属于正交基变换, 信号的稀

\* 基金项目: 国家自然科学基金(61335008)

疏变换唯一,正交基的选择影响稀疏性,进而影响信号重构。冗余字典可以使信号呈现最佳稀疏,需要遵循各个基向量使输入信号达到最佳稀疏的原则。基于这种原则,冗余字典是一定非正交并且冗余的矩阵,可以通过增加稀疏基的梳理来提高变换系统的冗余性,进而增强信号逼近的灵活性,同时也提高了稀疏表示高阶信号的能力。研究采用基于K-均值的K-SVD算法作为自适应冗余字典对语音信号稀疏。K-SVD算法的方程表示为:

min\_{D,X} { ||x\_l||\_0 } s.t. ||Y-DX||\_F^2 \le \epsilon (1)

KSVD算法的相对误差和平均帧重构信噪比(AFSNR)示意如图1。KSVD算法的平均帧重构信噪比如表1所示。

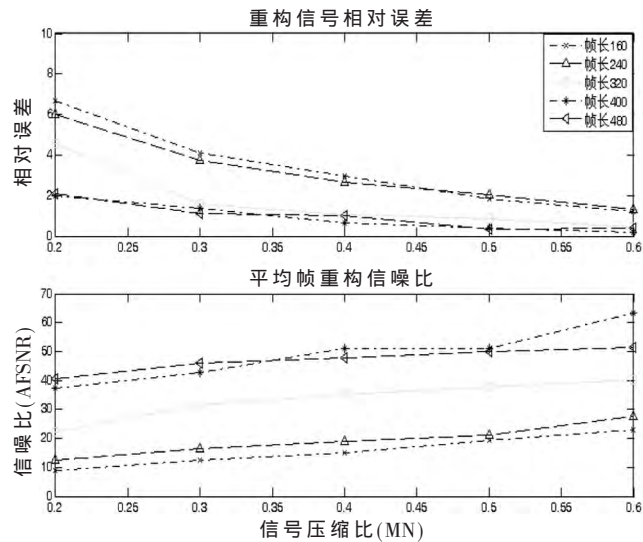


图1 KSVD算法的相对误差和AFSNR

表1 KSVD算法的平均帧重构信噪比

帧长	M/N				
	0.2	0.3	0.4	0.5	0.6
160	9.001 5	12.482 4	14.802 0	19.196 1	22.832 4
240	12.586 1	16.357 2	18.772 9	21.221 5	27.671 4
320	22.070 3	31.735 4	35.279 8	37.560 2	40.310 2
400	37.198 7	42.870 6	51.117 9	50.972 4	63.226 7
480	40.727 2	46.148 6	47.934 1	50.091 0	51.453 3

由表1的数据得到结论:经过KSVD算法稀疏得到重构语音的AFSNR随帧长和信号压缩比的增加而增大。

1.2 自适应观测矩阵

观测矩阵的设计原则是与稀疏矩阵尽可能不相干,且自身的列矩阵之间相互独立。随机观测矩阵具有上述特点,如高斯矩阵等。在语音信号观测投影时首先对信号进行编帧,然后选择与稀疏矩阵尽量不相干的观测矩阵相乘,得到观测值。

语音信号信号压缩比值越大,恢复信号的质量就越高;但观测值越多,其恢复时间也越长。根据语音信号每帧信息量的大小,选择相应的观测数,即自适应观测;能量大的帧分配较多的观测数,能量小的帧分配较小的观测数,噪声大部分存在于能量较小的帧,较少的观测数

能起到去噪的作用。

对已知的语音信号进行自适应观测,得到的相对误差和平均帧重构信噪比(AFSNR)如图2所示。

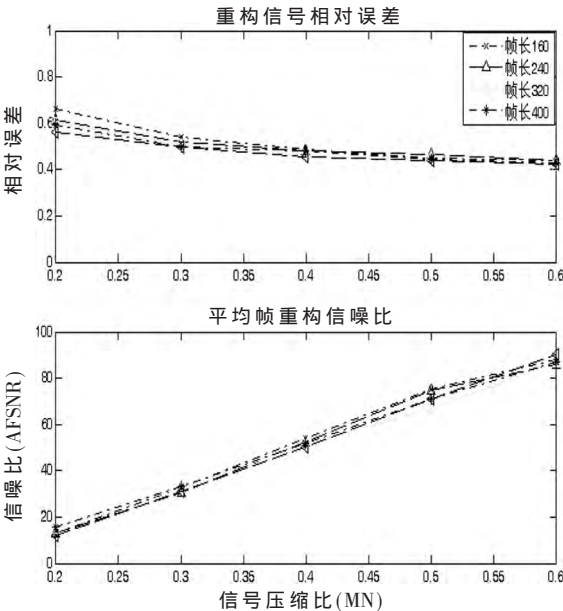


图2 自适应观测矩阵的相对误差和AFSNR

图2的相对误差表明,语音信号的帧长对自适应观测的相对误差值几乎没有影响,而压缩比的值对信号相对误差影响也较小,自适应观测的相对误差集中在0.45左右。

自适应观测矩阵的重构语音质量的好坏与压缩比有直接关系,但每帧信号的帧长大小对重构语音的影响微乎其微,如表2所示。

表2 自适应观测矩阵的平均帧重构信噪比

帧长	M/N				
	0.2	0.3	0.4	0.5	0.6
160	13.258 7	32.926 2	53.927 8	75.297 9	87.857 2
240	12.952 7	30.524 7	52.195 0	74.656 6	85.946 5
320	14.904 1	32.572 5	51.169 9	66.841 1	86.848 5
400	15.429 1	33.236 8	51.938 4	71.247 6	86.838 2
480	11.770 4	30.933 5	50.265 0	71.016 8	90.218 4

1.3 自适应重构算法

重构信号是利用优化求解的方法从观测值和稀疏矩阵中重构原信号。重建算法的设计应遵循如下原则:算法应使用较少的观测值,并且能精确和快速地恢复原信号。重构语音需要对每帧语音信号逐一进行恢复。常用的信号重构算法有BP(Basic pursuit)算法和OMP算法。

稀疏自适应匹配追踪算法(Sparse Adaptive Matching Pursuit,SAMP)是基于OMP算法基础上提出的,SAMP算法的主要思路是:在未知稀疏度的情况下进行信号重构,然后利用逐步加大步长的方法不断增加所用的原子规模,并引入回溯思想,在每次加大步长和选择原子后,都会与上步的原子合并,最后从中选择最佳的匹配原子。SAMP算法最大的优点是在稀疏度未知的前提下能

够自适应重构原信号。

对已知语音信号采用 SAMP 算法进行重构, 得到的相对误差和平均帧重构信噪比结果如图 3 所示。

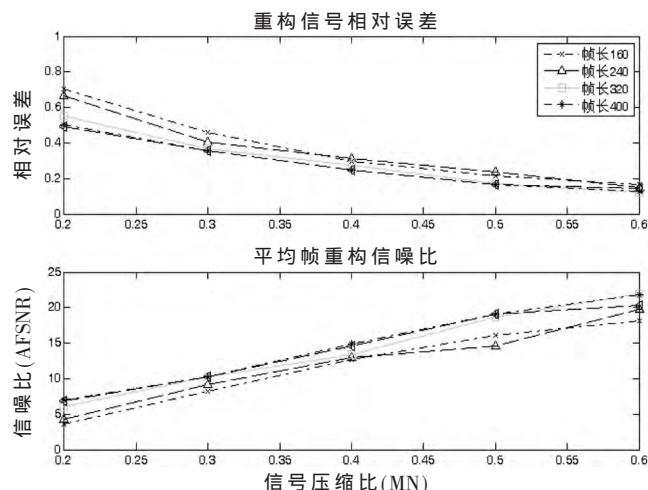


图 3 SAMP 算法的相对误差和 AFSNR

采用 BP 算法和 OMP 算法与 SAMP 算法进行比较, AFSNR 结果如图 4 所示。

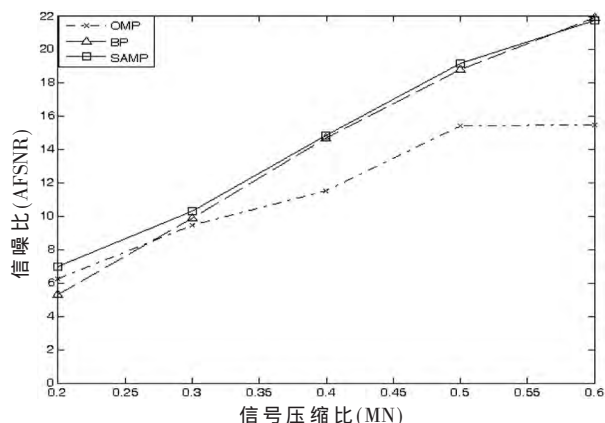


图 4 SAMP、BP 和 OMP 算法的 AFSNR 比较

以重构语音时间的角度评判, 运行观测矩阵约为 1 s, 剩余全部为重构时间, SAMP 算法用时在 10 s 内, OMP 算法最少时间是 10 s, 而随着压缩比的增大, 重构时间最高可达 400 s, BP 算法时间更长, 最小为 40 s, 最高为 2 500 s。因此使用 SAMP 算法会大大减小压缩感知重构语音的时间。

## 2 自适应压缩感知

语音自适应压缩感知的设计流程如图 5 所示。

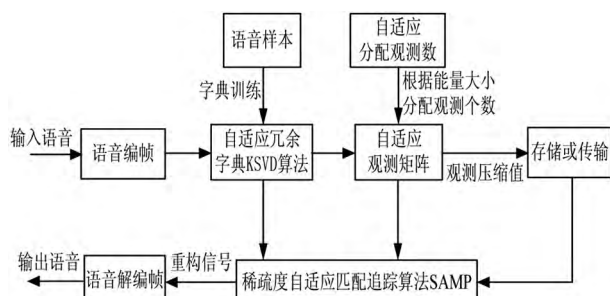


图 5 语音自适应压缩感知流程图

自适应压缩感知重构语音平均帧重构信噪比如表 3 所示, 表 3 的数据表明, 帧长与信号压缩比对重构语音的质量有积极影响, 而且信号压缩比值较小的条件下, 恢复的语音质量较高; 在帧长 400 时, 压缩比的大小对重构语音影响不大, 重构的语音信号达到优的级别, 同时要远远高于 KSVD 算法、自适应观测矩阵和 SAMP 算法单独重构语音的质量。3 种自适应算法组合一起能充分发挥各自的优点, 弥补了每个算法的缺陷: 减少了 KSVD 所产生的噪声, 重构语音的相对误差成倍的降低, 并且提高了重构的 AFSNR, 使得语音的质量更好, 同时提高了信号稀疏度, 降低了在重构时计算的复杂度, 从而进一步减少重构语音的时间。

表 3 自适应压缩感知重构语音平均帧重构信噪比

帧长	M/N				
	0.2	0.3	0.4	0.5	0.6
160	48.964 6	46.121 3	49.459 8	51.824 8	48.108 1
240	80.295 4	106.545 8	88.821 0	101.350 3	109.079 9
320	175.580 4	179.022 3	173.154 6	178.390 1	185.079 3
400	212.307 0	216.213 8	238.618 4	231.214 7	239.091 2
480	188.153 6	211.426 8	192.418 4	206.642 9	202.118 9

如图 6 所示, 相对于普通压缩感知, 自适应压缩感知具有更好的重构语音信号质量, 采样恢复语音时间更短, 而且在信号低压缩的条件下, 能恢复出良好的语音。

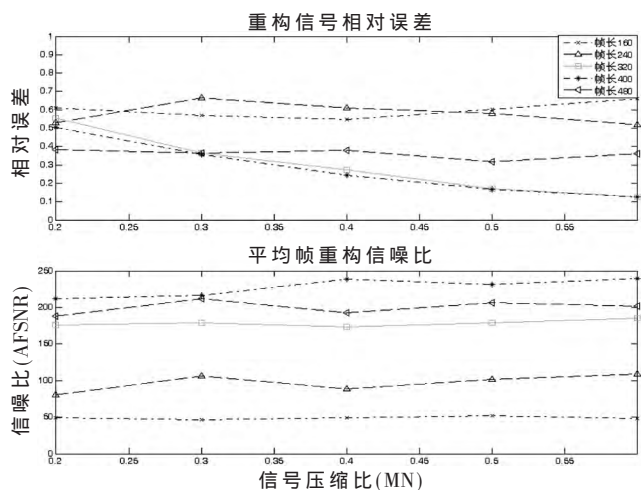


图 6 自适应压缩感知的相对误差和 AFSNR

## 3 结论

本文通过将自适应算法引入压缩感知, 结合自适应冗余字典 KSVD 算法、自适应观测矩阵和 SAMP 重构算法, 提出自适应压缩感知, 并分别进行仿真分析, 通过平均帧重构信噪比、相对误差, 与普通压缩感知进行对比, 验证了 KSVD 的稀疏的性能以及 SAMP 算法的优势。最后对自适应压缩感知进行仿真分析, 验证了将自适应算法引入压缩感知理论的可行性。

## 参考文献

- [1] CANDÈS E, WAKIN M. An introduction to compressive

(下转第 77 页)



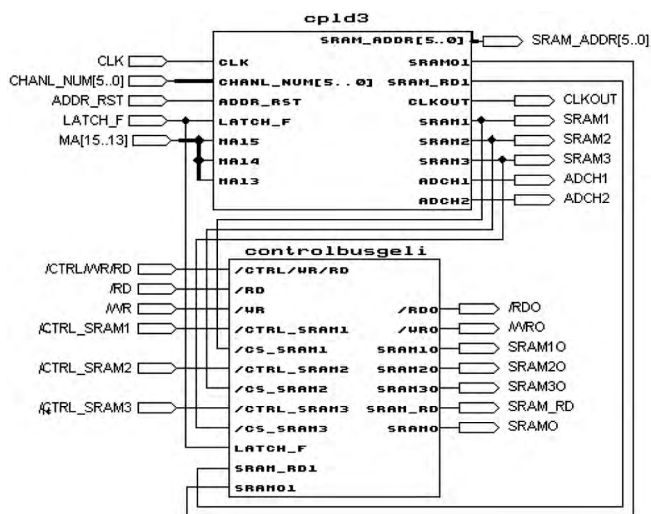


图 4 CPLD 框图

隔离。

正常工作时,系统的时序图如图 5 所示。

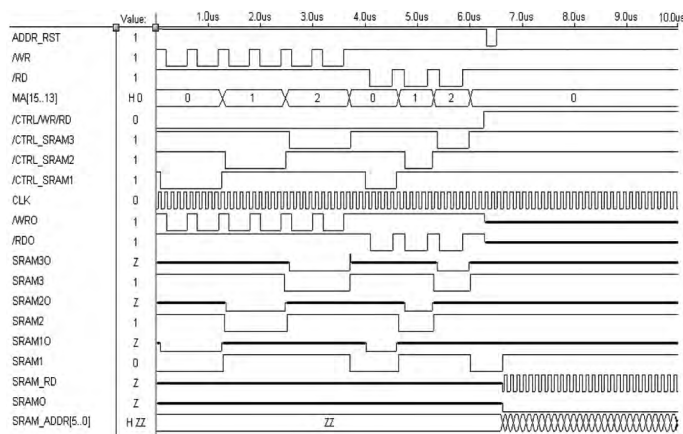


图 5 系统时序图

## 3 测试系统的软件设计

测试系统软件使用 VC++ 开发, 运行于 Windows 环境下, 人机界面友好, 包括系统测试软件和系统校准

软件。其中系统测试软件是在对被测系统进行测试时, 具体实现对电压、电流、相位、功率、功率因数等参数的稳态和瞬态进行数据测试、数据分析和数据存储的执行软件, 系统校准软件是对系统的精度进行计量的软件。

该软件包采用中文下拉菜单方式提示操作、显示、打印, 自动记录和存储所有测试数据, 以便于事后调出查询, 观察分析, 重新显示输出。测试数据套用相关标准限制曲线图形, 标准规定的极限曲线与试验数据填充曲线的对比图形可同时显示打印, 以便于判断测试结果是否满足设计要求。

## 4 结束语

本系统已经投入实际应用, 经过实践证明这一系统完全满足多通道同时采样并且速度快、精度高的要求, 稳定可靠, 取得了很好的效果, 说明了文章中所作的分析和讨论是合理和实用的。本系统可应用于高速、多通道且通道数可变的工程应用领域。

## 参考文献

- [1] 王育坚. Visual c++ 面向对象编程教程[M]. 北京: 清华大学出版社, 2003.
- [2] 夏宇闻. Verilog 数字系统设计教程[M]. 北京: 北京航空航天大学出版社, 2003.
- [3] 东方人华. MAX+PLUS 入门与提高[M]. 北京: 清华大学出版社, 2004.
- [4] 王金明, 杨吉斌. 数字系统设计与 Verilog HDL[M]. 北京: 电子工业出版社, 2002.
- [5] BHASKER J. Verilog HDL 硬件描述语言[M]. 徐振林, 译. 北京: 机械工业出版社, 2000.

(收稿日期: 2014-09-04)

## 作者简介:

马涛 (1981-), 男, 硕士, 工程师, 主要研究方向: 智能交通、智能仪表及工业控制。

郑鑫 (1981-), 女, 硕士, 工程师, 主要研究方向: 智能交通、微波传输技术。

(上接第 74 页)

- sampling[J]. IEEE Signal Processing Magazine, 2008, 25(2): 21-30.
- CANDÈS E, ROMBERG J, TAO T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. IEEE Transactions on Information Theory, 2006, 52(2): 489-509.
- 余恺, 李元实, 王智, 等. 基于压缩感知的新型声信号采集方法[J]. 仪器仪表学报, 2011, 33(1): 106-112.
- 周小星, 王安娜, 孙红英, 等. 基于压缩感知过程的语音增强[J]. 清华大学学报(自然科学版), 2011, 51(9): 1234-1238.

- [5] 罗武骏, 陶文凤, 左加阔, 等. 自适应语音压缩感知方法[J]. 东南大学学报(自然科学版), 2012, 42(6): 1027-1030.

- [6] 张雪英. 数字语音处理及 MATLAB 仿真[M]. 北京: 电子工业出版社, 2010.

(收稿日期: 2014-09-01)

## 作者简介:

毛静 (1987-), 女, 在读硕士, 主要研究方向: 测试计量技术与仪器。

尹鹏亮 (1987-), 男, 在读硕士, 主要研究方向: 测试计量技术与仪器。

马琦 (1982-), 男, 工程师, 主要研究方向: 测试计量技术与仪器。