

# 语音压缩感知研究进展与展望

孙林慧 杨震

(1. 南京邮电大学通信与信息工程学院, 南京, 210003, 2. 南京邮电大学“宽带无线通信与传感网技术”教育部重点实验室, 南京, 210003)

**摘要:** 压缩感知技术, 特别是语音压缩感知技术逐渐成为信号处理领域的研究热点。当前的语音压缩感知关键技术主要包括适合语音信号的稀疏分解矩阵构造, 观测矩阵的选择和重构算法的设计。稀疏分解矩阵的重要代表是正交基、基于语音特性的线性预测矩阵和过完备字典。观测矩阵方面主要采用随机观测矩阵分析语音压缩感知性能; 重构算法方面重点研究当观测序列或语音信号本身含有噪声时鲁棒的语音压缩感知重构算法。本文对上述语音压缩感知的3大关键技术进行了介绍和对比分析, 并对语音压缩感知的应用进行了总结, 最后对未来可能的研究热点进行了展望。

**关键词:** 压缩感知; 语音压缩感知; 稀疏表示; 观测矩阵; 自适应重构

**中图分类号:** TN912.3      **文献标志码:** A

## Compressed Speech Sensing for Research Progress and Prospect

Sun Linhui, Yang Zhen

(1. College of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China; 2. Key Lab of "Broadband Wireless Communication and Sensor Network Technology", Ministry of Education, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China)

**Abstract:** Compressed sensing technology, especially the compressed speech sensing technology has gradually become the research hotspot in signal processing. The currently key issues of compressed speech sensing include the construction of the sparse decomposition matrix, the selection of the measurement matrix and the design of the reconstruction algorithm for speech signal. The important representatives of sparse decomposition matrix are the orthogonal basis, linear prediction matrix based on speech characteristics and overcomplete dictionary. For measurement matrix, the performance of reconstructed speech signals based on random measurement matrix is analyzed. For reconstruction algorithm, the robust reconstruction algorithms with noisy measurement or noisy speech signal are researched. In the paper, the above three kinds of compressed speech sensing technologies are introduced and compared, and the main applications of compressed speech sensing are also provided. Finally, the possible future research points of compressed speech sensing are discussed.

**Key words:** compressed sensing; compressed speech sensing; sparse representation; measurement matrix; adaptive reconstruction

## 引 言

近年来,由 Donoho 与 Candes 等人提出的压缩感知<sup>[1-3]</sup>(Compressed sensing, CS)逐渐成为信号处理领域的研究热点。压缩感知理论的核心思想是边压缩边采样,将压缩与采样同时进行。首先对信号进行非自适应线性投影得到观测值;然后由观测值采用重构算法重构原始信号<sup>[4]</sup>。其优点在于采集的观测序列具有更小的冗余度,数据量远远小于基于传统 Nyquist 采样方法所获的数据量,突破了香农采样定理的瓶颈,使得高分辨率信号的采集成为可能。该理论的提出被美国科技评论评为 2007 年度十大科技进展。在 Candes 等提出 CS 理论之后, Duarte 等提出了分布 CS 理论(Distributed compressed sensing, DCS)<sup>[5]</sup>, Ji 等提出了 Bayesian CS 理论<sup>[6]</sup>, Laska 等提出了 1-BIT CS 理论<sup>[7]</sup>, Baraniuk 等提出了基于模型的 CS 理论<sup>[8]</sup>, Duarte 等又提出了谱 CS 理论<sup>[9]</sup>, Peyré 提出了最佳基 CS 理论<sup>[10]</sup>。这些理论都是在基本 CS 理论基础上进行了一些改进,如最佳基 CS 理论,基于信号在由正交基构造的树形字典上而不是基于信号在固定正交基上的稀疏先验,通过迭代阈值算法估计最佳基,实现对声音和图像等信号的特性的自适应捕获,从而提高了重构性能。

CS 理论的提出是建立在信号的稀疏表示基础上。信号的稀疏表示在信息压缩及分类、声音图像信号处理等信号处理领域、通信领域、欠定盲源分离、压缩感知中广泛应用。稀疏信号的定义<sup>[8]</sup>:若  $N$  维信号只有  $K$  个( $K < N$ )非零值,而其他值均为零,则称信号是  $K$  稀疏的。实际应用中,许多  $N$  维信号除了有少数大值外,其他的值都很小,其变换系数排序后以指数级衰减逐渐趋近于零,此时信号具有可压缩性,可以认为是近似稀疏。传统的信号稀疏表示都是基于正交变换的,但许多复杂的自然信号在单一的正交基下的表示不一定是稀疏的。Mallat 等提出了过完备稀疏表示理论<sup>[11]</sup>,即用过完备的冗余函数取代正交基函数,过完备冗余函数称为字典,字典中的原子可以尽可能地符合信号的结构,从而对信号更好的稀疏表示。在压缩感知中,为了确保信号非自适应线性压缩采样的观测值能够保持原信号的结构,需要选择稳定的观测矩阵。Tao 等提出压缩感知观测矩阵需要满足约束等距性(Restricted isometry property, RIP)条件<sup>[12-13]</sup>。Candes 等已证明随机高斯矩阵能以较大概率满足 RIP 条件<sup>[12,14]</sup>。Tsaig 采用满足 RIP 条件的随机高斯矩阵、随机贝努利矩阵、部分傅里叶矩阵、部分哈达玛矩阵作为观测矩阵,对稀疏信号进行观测并重构,重构信号的误差都比较小,并且随着观测数目的增加误差减小<sup>[3]</sup>。Bajwa 等将以很大概率满足 RIP 条件的托普利兹矩阵(Toeplitz)和循环矩阵作为观测矩阵<sup>[15-16]</sup>,这两种矩阵可以应用快速的重构算法,从而可以有效地减少高维信号问题的计算存储复杂度。Do 等提出了结构化随机矩阵<sup>[17]</sup>,该矩阵保持了随机高斯/伯努力矩阵和部分傅里叶变换矩阵的优点,可以看成是他们的混合模型。后来有学者提出采用自适应观测矩阵<sup>[18]</sup>,可以进一步减少重构误差一定时所需的观测个数。信号重构算法关系到压缩后信号能否精确重构和采样过程准确性的验证,是压缩感知理论中的关键部分。Cands 等已证明了可以通过求解最小  $l_0$  范数解决压缩感知信号重构问题,但由于求解最小  $l_0$  范数需要穷举信号非零值的所有排列可能,是一个 NP hard 问题,只能求其次最优解。常用的压缩感知的重构算法主要有基追踪(Basis pursuit, BP)算法<sup>[19]</sup>和匹配追踪(Matching pursuit, MP)系列的贪婪追踪算法。BP 算法解决的是在一定条件下与  $l_0$  优化问题等价的  $l_1$  优化问题。一般情况下 BP 算法可以找到全局最优解,重构信号质量较好,但是计算复杂度高。MP 算法比 BP 算法计算复杂度低,具有渐进收敛性,是最早的一种贪婪迭代算法,每次迭代的结果是次最优的。正交匹配追踪(Orthogonal matching pursuit, OMP)算法在 MP 算法的原子选择准则基础上,对已选择的原子集合进行了正交化,从而保证了每次迭代的最优性,减少了达到收敛的迭代次数。但是由于 OMP 算法在一次迭代中选取一个原子,完成重构需要较长的时间。后来出现了一系列改进的匹配追踪系列算法。Needdl 等在 OMP 的基础上提出

了正则化正交匹配追踪(Regularized orthogonal matching pursuit, ROMP)<sup>[20]</sup>算法,利用相关性原则和正则化原则一次筛选多个原子,该算法对于所有稀疏信号都可以准确重构。Donoho提出了分段正交匹配追踪(Stagewise orthogonal matching pursuit, StOMP)<sup>[21]</sup>算法,将迭代过程分为若干个阶段,在每次迭代时通过对残差信号进行匹配滤波得到相关矢量,再通过硬阈值方法选择重要的非零值位置,一次迭代可以选择一组原子,节省了重构时间,特别适于解决大数据量问题。ROMP和StOMP的优点是速度比OMP快,算法复杂度在 $O(KMN)$ 附近,远低于BP算法,缺点是需要较多数目的观测才能取得较好的重构效果。Needell等人还提出了压缩采样匹配追踪(Compressive sampling matching pursuit, CoSaMP)<sup>[22]</sup>算法,该算法引入回溯思想,重构过程具有一定的噪声鲁棒性。另外,子空间追踪(Subspace pursuit, SP)<sup>[23]</sup>算法也引入了回溯思想,首先建立一个候选集,然后再舍弃不需要的原子,最后形成支撑集,SP算法理论重构质量与BP相当,而重构复杂度与OMP相当。ROMP, StOMP, CoSaMP, SP算法都是建立在稀疏度已知的基础上。实际应用中稀疏度往往是未知的,稀疏自适应匹配追踪(Sparsity adaptive matching pursuit, SAMP)<sup>[24]</sup>算法就是在稀疏度未知的情况下通过固定步长逐步逼近实现重构,并获得了较好的重构效果。CS实现了直接信息采样,具有巨大的吸引力和应用前景,它的应用研究已经涉及到众多领域。

压缩感知在语音信号处理领域受到重视并得到应用,给语音信号处理领域带来新的发展变化。压缩感知是基于信号的稀疏性建立的边压缩边采样的新理论,从传统的信号采样转变成了信息采样,突破了传统的奈奎斯特采样定理的限制。国内外研究表明语音信号具有稀疏性或可压缩性,可以基于压缩感知实现低速率不失真地采样。用CS理论中的观测值来代替传统语音采样值,便于语音信号的采样、存储、传输和处理。基于CS理论研究语音信号处理方法,意味着对传统的基于Nyquist采样定理的语音分析方法的颠覆,由于压缩采样代替了Nyquist采样,使得采样信号特征发生了根本性的变化,进而影响到整个语音信号处理体系。因此,很多语音信号处理领域的专家和学者对语音压缩感知的关键技术及应用开展了积极的研究,并取得了一些可喜的进展。

## 1 语音信号的压缩感知采样

以压缩感知理论为背景,语音压缩感知采样技术主要包括适合语音信号的稀疏分解矩阵的构造、观测矩阵的构造、信号重构算法的设计、有噪声环境下语音信号稀疏性恶化相应的解决方案及基于观测和重构信号的应用等。基于CS理论的语音压缩感知采样系统框架如图1所示。

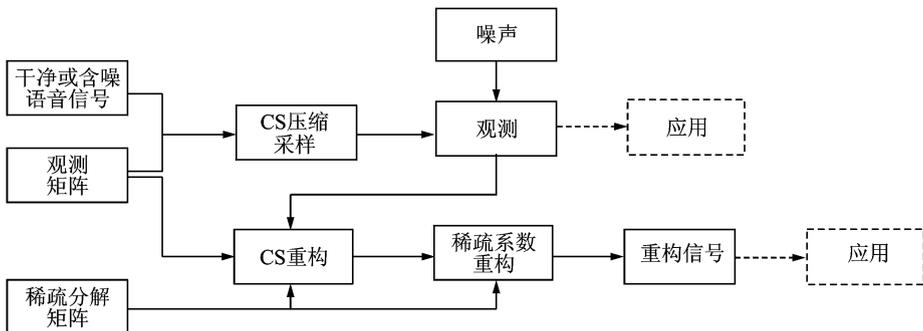


图1 基于CS的语音压缩感知系统框架

Fig. 1 Diagram of compressed speech sensing based on CS

$N \times 1$  维的语音帧信号  $\mathbf{x}$ , 可以用  $N \times 1$  维的基向量  $\{\boldsymbol{\psi}_i | i=1, 2, \dots, N\}$  的线性组合表示<sup>[2]</sup>, 因此  $\mathbf{x}$  可表示为

$$\mathbf{x} = \sum_{i=1}^N \alpha_i \boldsymbol{\psi}_i = \boldsymbol{\Psi} \boldsymbol{\alpha} \quad (1)$$

式中: 投影系数  $\alpha_i = \langle \mathbf{x}, \boldsymbol{\psi}_i \rangle$ , 正交基矩阵  $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \dots, \boldsymbol{\psi}_N]$ , 则投影系数矢量  $\boldsymbol{\alpha} = \boldsymbol{\Psi}^T \mathbf{x}$ ,  $\mathbf{x}, \boldsymbol{\alpha}$  分别为语音信号的时域和  $\boldsymbol{\Psi}$  域表示, 当  $\boldsymbol{\alpha}$  仅有  $K (\ll N)$  个非零系数(或远大于零的系数)时, 语音信号  $\mathbf{x}$  在正交基  $\boldsymbol{\Psi}$  上是  $K$  稀疏的, 则  $\boldsymbol{\Psi}$  为语音信号  $\mathbf{x}$  的稀疏基, 此时式(1)为语音信号  $\mathbf{x}$  的稀疏表示。如果  $\mathbf{x}$  是  $K$ -稀疏的, 根据 CS 理论, 可以对语音信号  $\mathbf{x}$  进行线性、非自适应的投影得到观测向量  $\mathbf{y} \in \mathbf{R}^{M \times 1}$

$$\mathbf{y} = \boldsymbol{\Phi} \mathbf{x} \quad (2)$$

式中: 观测矩阵  $\boldsymbol{\Phi} \in \mathbf{R}^{M \times N} (M \ll N)$  与正交基  $\boldsymbol{\Psi}$  不相关。由于  $\mathbf{y}$  的维数  $M$  远小于原始语音信号  $\mathbf{x}$  的维数  $N$ ,  $\mathbf{y}$  称为  $\mathbf{x}$  的压缩采样语音或观测值。将式(1)代入式(2), 有

$$\mathbf{y} = \boldsymbol{\Phi} \boldsymbol{\Psi} \boldsymbol{\alpha} = \boldsymbol{\Xi} \boldsymbol{\alpha} \quad (3)$$

式中: 语音感知矩阵  $\boldsymbol{\Xi} = \boldsymbol{\Phi} \boldsymbol{\Psi}$ 。求解式(3)的逆问题(根据  $\mathbf{y}$  求解  $\mathbf{x}$ )是一个病态问题, 所以无法直接从观测  $\mathbf{y}$  中解出信号  $\mathbf{x}$ 。然而当式(3)中  $\boldsymbol{\alpha}$  是  $K$  稀疏的, 且  $K < M \ll N$  时, 根据信号稀疏分解理论, 可以通过已有的稀疏分解算法由观测  $\mathbf{y}$  先求出稀疏系数  $\boldsymbol{\alpha}$ , 再由式(1)得到信号  $\mathbf{x}$ , 这称为语音信号的重构问题。

### 1.1 稀疏分解矩阵

在语音压缩感知系统中, 语音信号稀疏性的好坏直接影响重构语音的质量, 也就是说语音的稀疏分解矩阵对语音压缩重构性能影响很大。当采用同一观测矩阵时, 语音信号在稀疏分解矩阵下表示越稀疏, 准确或近似重构语音信号所需要的观测数越少; 当观测数固定时, 语音信号在稀疏分解矩阵下表示越稀疏, 重构语音的质量越高。

传统的稀疏分解矩阵主要采用正交基, 如 DCT(Discrete cosine transform)基<sup>[25]</sup>、小波基<sup>[26-27]</sup>、KLT(Karhunen-Loeve transform)基<sup>[28]</sup>等。文献[25]采用 DCT 基对语音压缩感知进行研究<sup>[25]</sup>。DCT 具有很强的能量集中作用, 能量主要集中在低频系数。语音信号在 DCT 基上分解, 大系数主要集中在低频。某帧语音帧信号  $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$  的 DCT 系数  $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_N\}$  可以表示为

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{N}} & \cdots & \frac{1}{\sqrt{N}} \\ \sqrt{\frac{2}{N}} \cos \frac{\pi}{2N} & \sqrt{\frac{2}{N}} \cos \frac{3\pi}{2N} & \cdots & \sqrt{\frac{2}{N}} \cos \frac{(2N-1)\pi}{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \sqrt{\frac{2}{N}} \cos \frac{(N-1)\pi}{2N} & \sqrt{\frac{2}{N}} \cos \frac{3(N-1)\pi}{2N} & \cdots & \sqrt{\frac{2}{N}} \cos \frac{(2N-1)(N-1)\pi}{2N} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \quad (4)$$

即

$$\boldsymbol{\alpha} = \boldsymbol{\Psi}^T \mathbf{x} \quad (5)$$

式中:  $\boldsymbol{\Psi}$  是 DCT 基。语音信号在 DCT 基上具有可压缩性, 采用 DCT 基的优势是 DCT 基是固定的正交基, 与输入信号特性无关, 具有普适性, 在压缩比大于 0.5 时, 基于 DCT 基对语音信号进行压缩感知时, 重构语音质量较好。

Sreenivas 等采用线性预测(Linear prediction, LP)矩阵对语音压缩感知进行研究<sup>[29]</sup>。由于语音信号帧内样值间存在强相关性, 通过线性预测可以使预测后的残差信号只有少数较大值, 出现较多的“零”值, 从而使语音信号在线性预测冗余域具有良好的稀疏性。LP 矩阵为

$$\Psi = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -a_1 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ -a_p & & & & \cdots & \vdots \\ 0 & & \cdots & & & 0 \\ 0 & \cdots & -a_p & \cdots & -a_1 & 1 \end{bmatrix}^{-1} \quad (6)$$

式中:  $a_i (i=1, 2, \dots, p)$  为第  $i$  个线性预测系数。通过  $\alpha = \Psi^{-1}x$  可以得到该帧信号的残差系数。语音信号的浊音在冗余域是近似稀疏的。清音的理想预测误差趋向白噪声, 样值间相关性较小, 清音在线性预测冗余域的稀疏性较差<sup>[30]</sup>。当式(6)中  $a_1=2, a_2=-1, p=2$  时, 原信号与预测信号的误差为

$$\alpha(n) = x(n) - 2x(n-1) + x(n-2) = [x(n) - x(n-1)] - [x(n-1) - x(n-2)] \quad (7)$$

由式(7)可知对语音信号做二阶前向差分就是对语音信号做二阶常系数的线性预测(Constant coefficient linear prediction, CCLP), 即通过前后相邻样值的二次差分得到冗余域系数矢量, 相应的矩阵  $\Psi^{-1}$  是二阶差分矩阵。

本文采用 LP, CCLP 和 DCT 矩阵作为稀疏分解矩阵, 分析语音压缩感知重构信号性能。图 2 (a~c) 为同一帧浊音信号, 图 2 (d~f) 为同一帧浊音信号分别为经过线性预测、二阶差分、DCT 变换得到的系数矢量, 系数绝对值大于 0.05 的个数分别为 6, 26, 59。可以看出, DCT 具有很强的能量集中的作用, 能量主要集中在低频系数, 而线性预测、二阶差分后的冗余系数分布较均匀, 不过都只有很少的大系数, 说明帧信号在 3 种变换下都具有近似稀疏性。就系数稀疏性而言, 线性预测优于二阶差分, 二阶差分优于 DCT。图 2 (g~i) 为压缩比为 0.3 的 CS 压缩采样后分别基于 LP, CLP, CT 矩阵并应用 BP 算法重构得到的系数矢量, 图 2 (j~l) 为基于 BP 算法重构的语音信号, 可以看出 LP 和 CCLP 作为压缩感知的稀疏分解矩阵, 其重建性能优于 DCT 变换矩阵, 能够既简单又更有效地恢复出原信号。低压缩比时, 采用 DCT 基重构语音信号性能欠佳, 主要是因为固定的 DCT 基对语音信号的稀疏特性的展示来说不是最佳的。基于 CCLP 矩阵的语音压缩感知可以获得较好的重构性能, 说明语音信号在常系数预测下的预测冗余域信号也具有一定的稀疏性。CS 压缩采样后基于 LP 矩阵的压缩感知性能好, 但是 LP 矩阵与信号相关, 压缩感知用于语音压缩重构时需要传输 LP 系数, 占用传输资源。

Rauhut 等对在冗余字典上稀疏的信号应用压缩感知<sup>[31]</sup>, 重构性能良好。由于语音信号在过完备字典下稀疏性进一步提高, 如何构造适合语音信号的冗余字典和快速有效的稀疏分解算法是当前语音压缩感知研究的主要任务之一。目前构造过完备字典的方法主要有 3 类:

(1) 利用传统基构造过完备字典。字典可以由单个正交基构造或者由若干个正交基的联合构造, 由单个正交基构造字典, 如通过更精细的采样 DCT 基的频率构造过完备 DCT 字典, 其优点是简单快速、复杂度较低。当字典为多个正交基的联合时, 对于由若干成分叠加的信号在此种结构的字典下的表示更加稀疏, 此时构造的过完备字典对复杂信号的稀疏分解较有效<sup>[32]</sup>。Ravelli 等采用不同尺度的修正 DCT (Modified discrete cosine transform, MDCT) 基函数构造了过完备字典, 音频信号在该字典下的稀疏性优于在单个 MDCT 基下的稀疏性<sup>[33]</sup>。

(2) 通过选取若干函数来设计字典。如 gabor 字典。

(3) 利用语音数据训练过完备字典。首先选取代表该类语音信号特性的大量训练信号, 通过学习得到该类信号的过完备字典, 由于该字典与该类信号具有较大的相关性, 所以能够高效精确地表示信号, 且具有一定的信号自适应性, 语音信号在训练得到的字典下呈现良好的稀疏性。由 Michal Aharon 等提出的基于 K 奇异值分解 (K-singular value decomposition, K-SVD) 算法<sup>[34-35]</sup> 就是利用训练数据通过迭代训练过完备字典的, 该算法在稀疏分解和字典的原子更新之间进行迭代, 主要优点: 灵活、简单、高效、

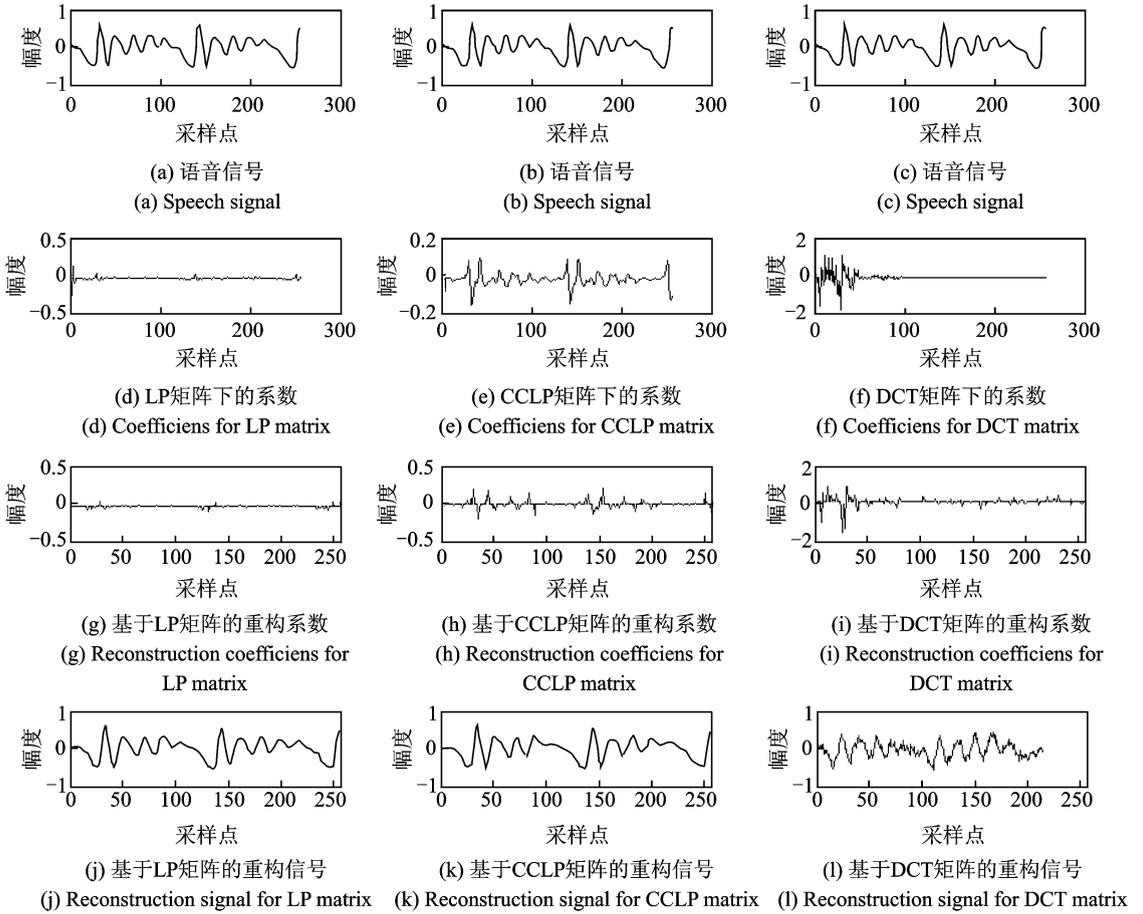


图2 基于不同稀疏分解矩阵的重构语音信号

Fig. 2 Reconstruction signal of speech for different sparse decomposition matrices

目标明确。Engan 等介绍了一系列基于迭代最小均方的字典学习算法(Iterative least squares based dictionary learning algorithms, ILS-DLA),这些算法都是在稀疏分解和字典的原子更新之间进行迭代<sup>[36]</sup>。由于信号基于字典稀疏分解需要较大的运算量,急需快速且有效的稀疏分解算法研究。Rubinstein 等提出了针对 K-SVD 算法的快速实现方法,利用 Batch-OMP 算法实现信号在字典下的稀疏表示,并用更快的近似算法代替奇异值分解算法实现字典中原子的更新<sup>[37]</sup>,因此提高了运算速度并且减少了所需的存储空间。Gowreesunker 等提出采用迭代子空间识别(Iterative subspace identification, ISI)学习算法构造过完备字典<sup>[38]</sup>,ISI 算法无需任何先验知识,可以通过学习推断出信号数据的子空间维数,生成数据驱动过的完备字典,信号在该字典下具有良好的稀疏性,该算法可以取得与 K-SVD 算法同样的稀疏表示但需要较少的训练时间,ISI 算法训练的过完备字典应用于音频盲源分离时需要较少的训练时间且可以取得好的分离效果。Skretting 等提出了基于递归最小均方的字典学习算法(Recursive least squares based dictionary learning algorithms, RLS-DLA)<sup>[39]</sup>,该算法每处理一个训练数据就对字典进行一次更新,而不是对全部训练数据进行处理来实现对字典的更新,其字典的更新是一个连续的过程,而且该算法通过引入遗忘因子,对初始字典的依赖性更小,算法的收敛性更好。

某帧语音信号波形如图 3(a),基于大小为  $256 \times 1024$  的 DCT 字典和过完备的 K-SVD 字典得到的

系数分别见图 3 (b,c)。可以看出:该帧语音信号在过完备 DCT 字典和过完备 K-SVD 字典下,只有少数系数数值较大,近似稀疏,语音信号可以基于这两种字典过完备稀疏表示。文献[40]采用 K-SVD 字典进行语音压缩感知,由于过完备的 K-SVD 字典由语音信号训练得到,更符合语音的结构,能够较好地表示语音信号,因此取得了较好的重构性能。

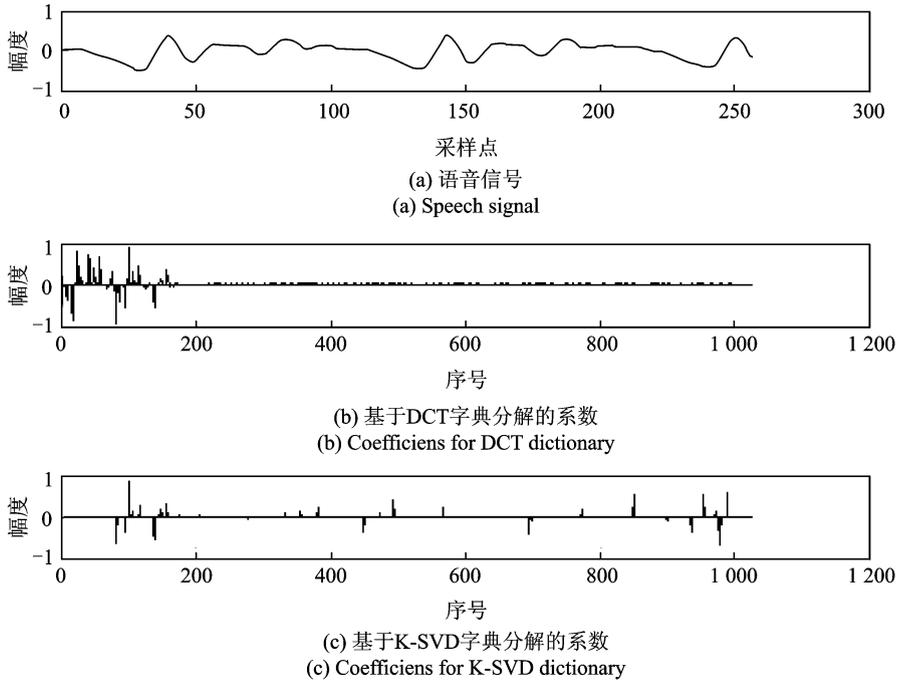


图 3 不同字典下浊音信号的稀疏表示

Fig. 3 Sparse representation of voiced speech for different dictionaries

文献[30]采用过完备线性预测(Overcomplete linear prediction, OLP)字典进行语音压缩感知。为了避免实测阶段对每帧语音信号求取 LP 系数,采用大量训练语音进行 LP 分析,再采用聚类算法构造 LP 系数矢量码本,进而由码本中的码矢构造 LP 矩阵,多个 LP 矩阵联合构造过完备的 LP 字典。基于此 OLP 字典,语音的浊音稀疏性进一步提高,清音的稀疏性也得到改善,语音压缩感知方法取得了较好的重构效果。该方法是基于训练的方法,需要较大的数据库用于训练,但是不占用实际测试阶段的时间。

文献[41]研究了一种新的基于自适应冗余字典的语音信号稀疏表示算法,该算法针对自相关函数为指数衰减的平稳信号,从 Karhunen-Loève (K-L)展开出发,建立了匹配信号结构的冗余字典,进而提出一种高效的基于非线性逼近的信号稀疏表示算法。该文献得出结论:冗余字典中原子的自适应性和代数结构可以使语音信号具有较高的稀疏度,应用于语音压缩感知有较好的重构精度。该方法基于信号的特性提取参数设计冗余字典,与信号紧密相关,因此提高了信号的稀疏性,可以应用于压缩编码中。曾理等基于此算法提出采用分段匹配追踪式 Karhunen-Loeve 非相干字典语音压缩感知方法<sup>[42-43]</sup>等。

相对基于正交基进行语音压缩采样而言,基于冗余字典进行语音压缩采样重构复杂度提高,因此实用中需要对字典的大小和重构精度进行折中考虑。

## 1.2 观测矩阵

观测矩阵需满足的约束等距性(Restricted isometry property, RIP)条件<sup>[12-13]</sup>

$$(1 - \delta_k) \alpha_2^2 \leq \Phi \Psi \alpha_2^2 \leq (1 + \delta_k) \alpha_2^2 \quad (8)$$

即

$$(1 - \delta_k) \alpha_2^2 \leq y_2^2 \leq (1 + \delta_k) \alpha_2^2 \quad (9)$$

式中:  $\delta_k \in (0, 1)$ , 可以通过极小化  $l_1$  范数恢复出稀疏信号。常见的满足约束等距性的观测矩阵有: 随机高斯矩阵、随机贝努利矩阵、部分傅里叶矩阵、部分哈达玛矩阵以及部分正交观测矩阵等<sup>[3]</sup>。

#### (1) 随机高斯矩阵

随机高斯(Gaussian)矩阵中每个元素独立的服从均值为 0, 方差为  $1/M$  的高斯分布, 即  $\Phi_{i,j} \sim N(0, 1/M)$ 。由于随机高斯矩阵与大多数固定正交基构成的矩阵不相关, 随机高斯矩阵是压缩感知中最常用的观测矩阵。可以采用 Matlab 中的函数 `randn(N, M)` 生成, 然后对列进行归一化。

#### (2) 随机贝努利矩阵

随机贝努利矩阵也是压缩感知常用的观测矩阵, 矩阵的每个元素独立地服从对称的贝努利分布, 即  $F_{i,j} = +\frac{1}{\sqrt{M}}$  和  $-\frac{1}{\sqrt{M}}$  的概率均为  $1/2$ 。

#### (3) 部分傅里叶矩阵

首先生成大小为  $N \times N$  的傅里叶方阵, 然后在生成矩阵中随机地选取  $M$  行向量, 构成一个  $M \times N$  的矩阵, 最后对列进行归一化。部分傅立叶观测矩阵是一个复数矩阵, 为了简单起见, 通常只选其实部作为观测矩阵。

#### (4) 部分哈达玛观测矩阵

首先生成大小为  $N \times N$  的哈达玛矩阵, 然后在生成矩阵中随机地选取  $M$  行向量, 构成一个  $M \times N$  的矩阵, 最后对列进行归一化。

#### (5) 部分正交观测矩阵

随机生成  $N \times N$  的正交矩阵, 然后在正交矩阵中随机地选取  $M$  行向量, 对  $M \times N$  矩阵的列向量进行单位化得到观测矩阵。

对一帧 512 点的语音数据, 采用不同随机观测矩阵, 在  $0.1 \sim 0.9$  的压缩比下压缩采样, 采用 BP 算法重构信噪比曲线如图 4 所示, 其中 USE, RSE, RST, Hadamard, URP 分别表示采用随机高斯观测矩阵、随机贝努利观测矩阵、实的部分傅里叶矩阵、部分哈达玛观测矩阵、正交观测矩阵作为观测矩阵时重构语音。可以看出: 采用 USE, RSE, URP 作为观测矩阵时重构语音性能相当, 且随着压缩比  $M/N$  的增大重构语音信噪比提高; 采用 Hadamard, RST 作为观测矩阵时, 虽然随着压缩比的增大重构语音信噪比也提高, 但整体重构语音性能较差。目前最常用的满足 RIP 条件的观测矩阵为随机高斯矩阵, 也有采用基于自相关特性的截断循环自相关矩阵<sup>[44]</sup>、固定矩阵<sup>[45]</sup>、随机滤波器组构造随机测量矩阵<sup>[46]</sup>等。采用随机矩阵投影得到的每一个观测值都包含了原来信号的所有信息, 传输观测值时丢失一小部分数据对重构语音质量影响相对较小。但是由于随机矩阵中每一个元素都需要存储, 所需存储空间大, 计算较复杂, 在硬件上难以实现。未来适合语音的结构化观测矩阵有待于进一步研究。

### 1.3 鲁棒的重构算法

无噪声环境下, 稀疏信号压缩采样后重构常用的求解方法有 BP 算法、MP 算法和正交 OMP 系列算法等。但是语音信号在常规正交基下是近似稀疏的, 即使在无噪声情况下采用这些重构方法恢复结果也存在一定误差, 导致大压缩下语音重构性能较差。因此, 要想在大压缩下取得好的重构性能, 一方面需要设计更适合语音信号的稀疏分解矩阵来提高语音信号的稀疏性, 另一方面需要设计对重构误差鲁棒的重构算法来增强重构性能。

针对重构误差鲁棒的重构方法有 LASSO(Least absolute shrinkage and selection operator)算法<sup>[47]</sup>、基追踪降噪算法(Basis pursuit de-noising, BPDN)、通过二次锥规划问题实现二次约束下极小化  $l_1$  范数

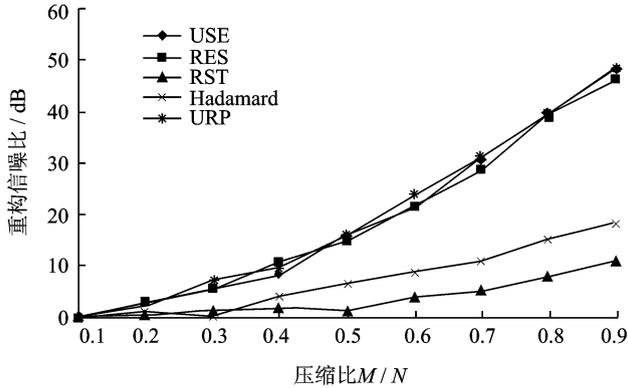


图4 基于不同观测矩阵重构语音信噪比曲线

Fig. 4 SNR curves of reconstructed speech signal for different measurement matrices

的方法<sup>[48]</sup>梯度投影稀疏重构法(Gradient projection for sparse reconstruction, GPSR)<sup>[49]</sup>等。LASSO方法和BPDN方法,把最优化问题等价于线性规划问题进行求解,利用这两种方法从包含观测噪声的感知数据中恢复的信号与原始信号的误差来源于两项。第1项是在无噪声情况下恢复结果的误差,而第2项正比于观测噪声大小。由此可以看出,这两种压缩感知重构方法对观测噪声有一定的鲁棒性。二次约束下极小化 $l_1$ 范数的方法是在满足观测约束的范围内找一个 $l_1$ 范数最小的向量作为恢复结果,该问题可以利用二次锥规划的求解方法进行求解。GPSR算法基于 $l_p$ 范数最小求解,解决一个受约束的二次规划问题。在已有的算法中,GPSR以高运算速度和良好的重构效果著称,主要策略是从可行点出发,沿着下降的可行方向搜索寻找使目标函数值下降的新的可行点。

实际应用中感知到的语音数据不可避免地会包含一些噪声和误差,针对语音信号本身含有噪声,由于语音信号稀疏性恶化,需要设计更具鲁棒性的重构算法。文献[50]前期研究了含噪语音信号的自适应基追踪去噪算法,自适应选择最佳重构参数,从而提高重构性能。文献[51]研究了含噪语音压缩与重构的自适应共轭梯度投影算法,在自适应选择最佳重构参数的同时进一步加快了收敛速度。

## 2 压缩感知在语音信号处理中的应用

国内外将CS用于语音信号处理领域的研究相对于图像处理领域较少,主要应用于语音识别、语音增强、语音编码、语音信息隐藏、语音丢包补偿和欠定盲源分离等领域。

(1)语音识别。Gemmeke等<sup>[52]</sup>提出预先由干净语音信号训练过完备基,单个字在该基上稀疏,利用CS原理对噪声环境下的语音进行识别,语音识别系统的抗噪性能大大提高。

(2)语音增强。在现实环境中的语音通常都混有各类噪声,因此,语音增强作为对声音信号的预处理技术,对于其后的各种语音应用的质量扮演着非常重要的角色。迄今已有许多方法运用于语音增强,如维纳滤波、卡尔曼滤波、最小均方误差的谱幅值估计方法及信号子空间分析等。Wu等基于压缩感知实现了语音增强<sup>[53-54]</sup>。

(3)语音编码。Giacobello等<sup>[55]</sup>提出把CS的理论框架应用到语音编码领域。利用语音信号冗余域的稀疏性,特别是浊音信号的稀疏性,通过CS方法计算稀疏的线性激励近似值,然后再进行编码,并与多脉冲激励等方法进行了对比,实验结果表明可以取得好的听觉效果。该文献也表明CS可以应用到非正交基上具有稀疏性的语音信号。另外,Griffin等采用CS理论对声音信号进行正弦模型编码<sup>[56]</sup>。

(4)语音信息隐藏。为了实现在语音信号中高容量的信息隐藏,考虑到人的听觉系统的敏感性,Xu

等<sup>[57]</sup>把 CS 理论引入信息隐藏。首先采用 CS 技术压缩秘密语音以减少信息比特率,然后通过采用 DCT 和 LWT(Lifting wavelet transform)两级转换技术提高系统性能。仿真结果显示该方法具有高的隐藏度和抗信道干扰能力。

(5) 语音丢包补偿。在 VoIP 和无线移动通信中提供有效可靠的语音传输是重要和具有挑战性的任务。在实际的通信网络中,恶劣条件下会发生严重的丢包现象。Ma<sup>[58]</sup>等基于 CS 理论提出丢包补偿策略,主要利用语音信号的稀疏性和 CS 的随机采样理论,基于收到的语音信号恢复丢失的部分,该部分独立于语音编码过程。该丢包补偿策略可以应用到语音通信的接收端。

(6) 欠定盲源分离。欠定盲源分离问题通常基于信号的稀疏表示。Xu 等<sup>[59]</sup>把 CS 引入到欠定盲源分离,提出的算法包括两个阶段。第 1 阶段在转换域利用聚类算法估计声音信号的未知的混合矩阵,第 2 阶段源信号通过 CS 方法恢复,取得了良好的实验性能。Bao 等采用压缩感知方法和两层稀疏字典讨论了欠定盲语音分离,认为语音信号低频部分在 K-SVD 字典下稀疏,语音信号高频部分在 DCT 基下稀疏,取得了良好的分离性能<sup>[60]</sup>。

另外,Defraene 等使用感知压缩感知消除音频信号的咔嚓声<sup>[61]</sup>,Mignot 等使用压缩传感来实现房间混响重建<sup>[62-63]</sup>。

### 3 结束语

相关研究进展表明,压缩感知理论的提出极大地丰富了信号采样理论,并为语音信号处理领域的研究提供了新技术和新思路,研究前景广阔。然而目前适用于近似稀疏的语音信号的压缩感知理论还不是很完善,相应的应用研究也刚刚起步,尚有较多问题需要在未来研究中得到突破。

#### (1) 自适应观测矩阵构造问题

从信息论的角度来说,最优的投影应当可以使得原信号与投影信号之间的互信息量最大,也即最优的投影信号应当最大程度去捕捉原信号的信息。在压缩感知中,观测矩阵需要满足约束等距性条件,目前观测矩阵大多采用随机观测矩阵,建立在非自适应线性观测基础上的压缩感知技术具有普适性,但不具有灵活性,考虑到语音信号具有很强的随机性和非平稳性,因而有必要针对语音信号研究自适应观测矩阵,即根据不同的信号类型采用不同的观测矩阵进行自适应压缩采样。进一步研究点可能在基于语音自相关特性的结构化自适应观测矩阵的设计方面,以便实现自适应观测。

#### (2) 语音感知矩阵的优化问题

当信号不能在正交基上稀疏表示时,可以将其扩展到过完备字典上进行稀疏表示。对于语音信号,在过完备字典上投影通常可以使其更加稀疏。然而在压缩感知中,利用过完备字典代替经典的正交基虽然可以更好地表示和重构信号,但由于相应的感知矩阵中会出现较多的相关列,这些相关列增加了算法的存储和计算的成本。因此过完备字典的冗余度和感知矩阵中相关列的数量需根据语音信号特性综合优化。未来可在语音感知矩阵迭代优化算法方面进一步研究,以便结合语音信号特性,在已知信号稀疏分解矩阵的条件下,逐步迭代以减少 CS 感知矩阵的列向量的相关性,从而进一步优化感知矩阵的近似正交特性以获得最优感知矩阵,最终使得语音压缩感知能够更准确地重构出原始信号。

#### (3) 语音信号分类稀疏分解矩阵的构造问题

压缩感知理论应用的前提是信号本身是稀疏的,或者信号在转换域是稀疏的。目前,常规的正交基和冗余字典下语音信号的浊音稀疏性较好,而清音不稀疏,这是一个难点,影响整体语音信号采用 CS 理论的效果。未来可在使语音清音和浊音分类稀疏表示的稀疏分解矩阵的构造方面进一步研究。可以结合语音信号样值间的相关特性,针对清浊音分别构造稀疏分解矩阵,提高浊音的稀疏性、改善清音的稀疏性,从而提高稀疏分解效率。

#### (4) 噪声环境下语音压缩感知的高质量重构问题

初步研究表明:在含噪环境下,语音信号稀疏性变差,需要更多的观测序列才能够保证高的重构概率以及好的重构质量,这样不仅降低了压缩率并增加了系统复杂度。如何以最少的观测序列精确重构原始信号是含噪语音压缩感知模型的重要课题。进一步研究点可能包括:鲁棒性抗噪观测矩阵的构建;鲁棒性压缩感知信号重构技术;语音压缩感知系统中线性和非线性消噪技术。鲁棒性抗噪观测矩阵的构建方面可以考虑利用反馈机制对噪声类型进行分析,结合序贯压缩感知(Sequential compressed sensing, SCS)原理,寻找重构信号、扩展观测空间与信号重构误差间的内在关系。通过估算信号的重构误差,确定扩展观测矩阵的最小维数,在确定最优观测序列长度的基础上,可以采用贝叶斯压缩感知(Bayesian compressed sensing, BCS)思想对新增的观测矩阵进行最优设计,最终构建出具有最低复杂度且对不同类型噪声均具有鲁棒性的观测矩阵。鲁棒性压缩感知信号重构技术方面,可以考虑根据噪声和语音统计特性的自适应重构算法和基于分布式压缩感知模型的协作降噪重构算法,侧重研究在语音信号或观测序列受到噪声影响的前提下,如何抑制压缩感知的重构算法对噪声的放大作用,进一步改善重构性能。语音压缩感知系统中线性和非线性消噪技术方面,可以考虑线性和非线性重构前置滤波器设计和后置消噪处理系统设计。

#### (5) 语音观测值的应用问题

语音信号处理的最终目的并不都是重构信号,而有时是为了得到有关的信息。由压缩感知理论可知,在一定条件下,通过少量的观测值可以准确重构原始信号,即少量的观测值能够保持原始语音信号的结构和信息。因此,进一步研究点包括:语音压缩感知观测值直接用于实现语音清浊音分类<sup>[64]</sup>、特征提取、目标检测、说话人识别等任务;对观测直接进行建模以便实现超低速率编码。

#### 参考文献:

- [1] Donoho D L. Compressed sensing [J]. IEEE Transactions on Information Theory, 2006,52(4):1289-1306.
- [2] Baraniuk R G. Compressive sensing [J]. IEEE Signal Processing Magazine, 2007,24(4):118-120.
- [3] Tsaig Y, Donoho D L. Extensions of compressed sensing [J]. Signal Processing, 2006,86(3):549-571.
- [4] Candes E J, 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.
- [5] Duarte M F, Sarvotham S, Baron D, et al. Distributed compressed sensing of jointly sparse signals[C]// Conference Record-Asilomar Conference on Signals, Systems and Computers. Pacific Grove, California:IEEE, 2005:1537-1541.
- [6] Ji S, Xue Y, Carin L. Bayesian compressive sensing [J]. IEEE Transactions on Signal Processing, 2008,56(6):2346-2356.
- [7] Laska J N, Wen Z W, Yin W T, et al. Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements [J]. IEEE Transactions on Signal Processing, 2011,59(11):5289-5301.
- [8] Baraniuk R G, Cevher, V, Duarte M F, et al. Model-based compressive sensing[J]. IEEE Transactions on Information Theory, 2010,56(4):1982-2001.
- [9] Duarte M F, Baraniuk R G. Spectral compressive sensing[J]. Applied and Computational Harmonic Analysis, 2013, 35(1): 111-129.
- [10] Peyré G. Best basis compressed sensing [J]. IEEE Transactions on Signal Processing, 2010,58(5):2613-2622.
- [11] Mallat S G, Zhang Z F. Matching pursuits with time-frequency dictionaries [J]. IEEE Transactions on Signal Processing, 1993,41(12):3397-3415.
- [12] Candes E J, Tao T. Decoding by linear programming [J]. IEEE Transactions on Information Theory, 2005,51(12):4203-4215.
- [13] Candes E, Romberg J. Sparsity and incoherence in compressive sampling [J]. Inverse Problems, 2007,23(3):969-985.
- [14] Candes E J, Tao T. Near-optimal signal recovery from random projections; Universal encoding strategies[J]. IEEE Transactions on Information Theory, 2006,52(12):5406-5425.
- [15] Bajwa W U, Haupt J D, Raz G M, et al. Toeplitz-structured compressed sensing matrices [C]//IEEE Workshop on Statistical Signal Processing Proceedings. Madison WI, United States:IEEE, 2007:294-298.
- [16] Yu L, Barbot J P, Zheng G, et al. Toeplitz-structured chaotic sensing matrix for compressive sensing [C]// International

Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP). Newcastle Upon Tyne, United Kingdom; IEEE, 2010:229-233.

- [17] Do T T, Tran T D, Gan L. Fast compressive sampling with structurally random matrices [C]//IEEE International Conference on Acoustics, Speech and Signal Processing-Proceedings (ICASSP). Las Vegas NV, United States; IEEE, 2008:3369-3372.
- [18] Babacan S D, Molina R, Katsaggelos A K. Bayesian compressive sensing using laplace priors [J]. *IEEE Transactions on Image Processing*, 2010, 19(1):53-63.
- [19] 石光明, 刘丹华, 高大化, 等. 压缩感知理论及其研究进展[J]. *电子学报*. 2009, 37(5):1070-1081.  
Shi Guangming, Liu Danhua, Gao Dahua, et al. Advances in theory and application of compressed sensing [J]. *Acta Electronica Sinica*, 2009, 37 (5):1070-1081.
- [20] Needell D, Vershynin R. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit [J]. *IEEE Journal on Selected Topics in Signal Processing*, 2010, 4(2):310-316.
- [21] Donoho D L, Tsaig Y, Drori I, et al. Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit [J]. *IEEE Transactions on Information Theory*, 2012, 58(2):1094-1121.
- [22] Needell D, Tropp J A. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples [J]. *Applied and Computation Harmonic Analysis*, 2009, 26:301-321.
- [23] Dai W, Milenkovic O. Subspace pursuit for compressive sensing signal reconstruction [J]. *IEEE Transactions on Information Theory*, 2009, 55(5):2230-2249.
- [24] Do T T, Gan L, Nguyen N, et al. Sparsity adaptive matching pursuit algorithm for practical compressed sensing [C]// Conference Record - Asilomar Conference on Signals, Systems and Computers. Pacific Grove, California; IEEE, 2008:581-587.
- [25] 郭海燕, 王天荆, 杨震. DCT 域的语音信号自适应压缩感知[J]. *仪器仪表学报*. 2010, 31(6):1262-1268.  
Guo Haiyan, Wang Tianjing, Yang Zhen. Adaptive speech compressed sensing in the DCT domain [J]. *Chinese Journal of Scientific Instrument*, 2010, 31(6):1262-1268.
- [26] 孙林慧, 杨震, 叶蕾. 基于自适应多尺度压缩感知的语音压缩与重构[J]. *电子学报*, 2011, 39(1):40-45.  
Sun Linhui, Yang Zhen, Ye Lei. Speech compression and reconstruction based on adaptive multiscale compressed sensing theory [J]. *Acta Electronica Sinica*, 2011, 39(1):40-45.
- [27] 梁瑞宇, 邹采荣, 赵力, 等. 语音压缩感知及其重构算法[J]. *东南大学学报:自然科学版*, 2011, 41(1):1-5.  
Liang Ruiyu, Zou Cairong, Zhao Li, et al. Compressed sensing in speech and its reconstruction algorithm [J]. *Journal of Southeast University: Natural Science Edition*, 2011, 41(1):1-5.
- [28] 郭海燕, 杨震. 基于近似 KLT 域的语音信号压缩感知[J]. *电子与信息学报*. 2009, 31(12):2948-2952.  
Guo Haiyan, Yang Zhen. Compressed speech signal sensing based on approximate KLT [J]. *Journal of Electronics and Information Technology*, 2009, 31(12):2948-2952.
- [29] Sreenivas T V, Kleijn W B. Compressive sensing for sparsely excited speech signals [C]// IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Taipei, China; IEEE, 2009:4125-4128.
- [30] 孙林慧, 杨震, 季云云, 等. 基于过完备线性预测字典的压缩感知语音重构. *仪器仪表学报*, 2012, 33(4):743-749.  
Sun Linhui, Yang Zhen, Ji Yunyun, et al. Reconstruction of compressed speech sensing based on overcomplete linear prediction dictionary [J]. *Chinese Journal of Scientific Instrument*, 2012, 33(4):743-749.
- [31] Rauhut H, Schnass K, Vandergheynst P. Compressed sensing and redundant dictionaries [J]. *IEEE Transactions on Information Theory*, 2008, 54(5):2210-2219.
- [32] Zhang Chunmei, Yin Zhongke, Chen Xiaodong, et al. Signal overcomplete representation and sparse decomposition based on redundant dictionaries [J]. *Chinese Science Bulletin*, 2005, 50(23):2672-2677.
- [33] Ravelli E, Richard G, Daudet L. Union of MDCT bases for audio coding [J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2008, 16(8):1361-1372.
- [34] 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.
- [35] Zhang Qiang, Li Baoxin. Discriminative K-SVD for dictionary learning in face recognition [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). San Francisco, CA, United States; IEEE, 2010:2691-

2698.

- [36] Engan K, Skretting K, Husøy J H. A family of iterative LS-based dictionary learning algorithms, ILS-DLA, for sparse signal representation [J]. *Digital Signal Process*, 2007,17(1):32-49.
- [37] Rubinstein R, Zibulevsky M, Elad M. Efficient implementation of the k-svd algorithm using batch orthogonal matching pursuit [R]. Technical Report CS-2008-08, Technion, 2008.
- [38] Gowreesunker B V, Tewfik A H. Learning sparse representation using iterative subspace identification [J]. *IEEE Transactions on Signal Processing*, 2010,58(6):3055-3065.
- [39] Skretting K, Engan K. Recursive least squares dictionary learning algorithm [J]. *IEEE Transactions on Signal Processing*, 2010,58(4):2121-2130.
- [40] Xu Tingting, Yang Zhen, Shao Xi. Adaptive compressed sensing of speech signal based on data-driven dictionary [C] // *Asia-Pacific Conference on Communications (APCC)*. Shanghai, China:IEEE, 2009:257-260.
- [41] 王天荆, 郑宝玉, 杨震. 基于自适应冗余字典的语音信号稀疏表示算法[J]. *电子与信息学报*, 2011,33(10):2372-2377.  
Wang Tianjing,Zheng Baoyu,Yang Zhen. A speech signal sparse representation algorithm based on adaptive overcomplete dictionary [J]. *Journal of Electronics and Information Technology*,2011,33(10):2372-2377.
- [42] 曾理,张雄伟,陈亮,等. 分段匹配追踪式 Karhunen-Loeve 非相干字典语音压缩感知[J]. *声学学报*,2013,38(4):493-500.  
Zeng Li,Zhang Xiongwei,Chen Liang,et al. Compressed speech signal sensing with K-L incoherent dictionary based on segment MP [J]. *Acta Acustica*,2013,38(4):493-500.
- [43] 曾理,张雄伟,陈亮,等. 基于压缩感知的 K-L 分解语音稀疏表示算法[J]. *数据采集与处理*,2013,28(3):267-273.  
Zeng Li, Zhang Xiongwei,Chen Liang, et al. Compressed sensing based speech sparse representation with K-L expansion [J]. *Journal of Data Acquisition and Processing*,2013,28(3):267-273.
- [44] 季云云,杨震. 基于自相关观测的语音信号压缩感知[J]. *信号处理*, 2011, 127(2):207-214.  
Ji Yunyun, Yang Zhen. Compressed speech signal sensing based on autocorrelative measurement [J]. *Signal Processing*, 2011,127(2):207-214.
- [45] 叶蕾, 杨震, 王天荆, 等. 行阶梯观测矩阵、对偶仿射尺度内点重构算法下的语音压缩感知[J]. *电子学报*. 2012,40(3):429-434.  
Ye Lei, Yang Zhen, Wang Tianjing, et al. Compressed sensing of speech signal based on row echelon measurement matrix and dual affine scaling interior point reconstruction method [J]. *Acta Electronica Sinica*, 2012,40(3):429-434.
- [46] 隋璐瑛, 黄建军, 张雄伟, 等. 压缩感知在语音处理中的应用研究[J]. *军事通信技术*,2012,02:100-104.  
Sui Luying, Huang Jianjun,Zhang Xiongwei,et al. Compressive sensing theory and its application in speech signal processing [J]. *Journal of Military Communications Technology*, 2012,02:100-104.
- [47] Candès E, Tao T. The Dantzig selector:Statistical estimation when  $p$  is much larger than  $n$ [J]. *Annals of Statistics*,2007,35(6):2313-2351.
- [48] Boyd S, Vandenberghe L. *Convex optimization* [M]. Cambridge:Cambridge University Press, 2004.
- [49] Figueiredo M A T, Nowak R D, Wright S J. Gradient projection for sparse reconstruction:Application to compressed sensing and other inverse problems [J]. *IEEE Journal on Selected Topics in Signal Processing*. 2007,1(4):586-597.
- [50] 孙林慧,杨震. 基于自适应基追踪去噪的含噪语音压缩感知[J]. *南京邮电大学学报*,2011,31(5):1-6.  
Sun Linhui, Yang Zhen. Compressed sensing of noisy speech signal based on adaptive basis pursuit de-noising [J]. *Journal of Nanjing University of Posts and Telecommunications*,2011,31(5):1-6.
- [51] 杨真真, 杨震. 含噪语音压缩与重构的自适应共轭梯度投影算法[J]. *仪器仪表学报*, 2012,33(10):2200-2207.  
Yang Zhenzhen,Yang Zhen. Adaptive conjugate gradient projection algorithm for noisy speech compression and reconstruction [J]. *Chinese Journal of Scientific Instrument*,2012, 33(10):2200-2207.
- [52] Gemmeke J F, Hamme H V, Cranen B, et al. Compressive sensing for missing data imputation in noise robust speech recognition [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2010,4(2):272-287.
- [53] Wu D, Zhu W P, Swamy M N S. Compressive sensing based speech enhancement in non-sparse noisy environments[J]. *Signal Processing, IET*, 2013,7(5):450-457.
- [54] Wu D, Zhu W P, Swamy M N S. The theory of compressive sensing matching pursuit considering time-domain noise with application to speech enhancement[J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2013,22(3):

682-696.

- [55] Giacobello D, Christensen M G, Murthi M N, et al. Retrieving sparse patterns using a compressed sensing framework: applications to speech coding based on sparse linear prediction [J]. *Signal Processing Letters, IEEE*, 2010, 17(1):103-106.
- [56] Griffin A, Hirvonen T, Tzagkarakis C, et al. Single-channel and multi-channel sinusoidal audio coding using compressed sensing [J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2011, 19(5):1382-1395.
- [57] Xu Tingting, Yang Zhen, Shao Xi. Novel speech secure communication system based on information hiding and compressed sensing [C] // *International Conference on Systems and Networks Communications (ICSNC)*. Porto, Portugal: IEEE, 2009: 201-206.
- [58] Ma Hongfei, Xiong Jing, Xu Yajun, et al. Packet loss concealment for speech transmission based on compressed sensing [C] // *IET International Communication Conference on Wireless Mobile and Computing (CCWMC)*. Shanghai, China: IEEE, 2009: 327-330.
- [59] Xu Tao, Wang Wenwu. A block-based compressed sensing method for underdetermined blind speech separation incorporating binary mask [C] // *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Dallas TX, United States: IEEE, 2010: 2022-2025.
- [60] Bao Guangzhao, Ye Zhongfu, Xu Xu, et al. A compressed sensing approach to blind separation of speech mixture based on a two-layer sparsity model [J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2013, 21(5): 899-906.
- [61] Defraene B, Mansour N, De Hertogh S, et al. Declipping of audio signals using perceptual compressed sensing [J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2013, 21(12): 2627-2637.
- [62] Mignot R, Chardon G, Daudet L. Low frequency interpolation of room impulse responses using compressed sensing [J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2014, 22(1): 205-216.
- [63] Mignot R, Daudet L, Ollivier F. Room reverberation reconstruction: Interpolation of the early part using compressed sensing [J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2013, 21(11): 2301-2312.
- [64] 王文娟, 杨震. 基于语音压缩感知观测序列非重构的清浊音判别法 [J]. *数据采集与处理*, 2013, 28(3): 274-279.  
Wang Wenjuan, Yang Zhen. Voicing-state identification based on speech observation sequence and non-reconstruction in compressing sensing [J]. *Journal of Data Acquisition and Processing*, 2013, 28(3): 274-279.

作者简介: 孙林慧(1979-), 女, 副教授, 博士, 研究方向: 语音处理与现代语音通信, E-mail: sunlh@njupt.edu.cn; 杨震(1961-), 男, 教授, 博士生导师, 研究方向: 语音处理与现代语音通信及网络通信技术, E-mail: yangz@njupt.edu.cn。

