

文章编号:1004-9037(2013)03-0267-07

# 基于压缩感知的 K-L 分解语音稀疏表示算法

曾 理<sup>1</sup> 张雄伟<sup>1</sup> 陈 亮<sup>2</sup> 杨吉斌<sup>1</sup> 贾 冲<sup>1</sup>

(1. 解放军理工大学指挥信息系统学院,南京,210007;2. 解放军理工大学通信工程学院,南京,210007)

**摘要:**为克服稀疏基在 KLT 域不便传输的不足,提出一种基于 Karhunen-Loeve(K-L)正交分解的语音稀疏表示算法。结合压缩感知理论,建立语音自相关模型并求解 Fredholm 积分方程,采用二分法估计出可实时传输的模型参数,构造非相干字典;然后用随机矩阵对语音在字典上的稀疏投影系数进行观测获得低维观测值。重构结果表明:相比已有的稀疏表示算法,本文算法的字典匹配性更好,且具有较好的语音质量。

**关键词:**压缩感知;K-L 分解;稀疏表示;匹配追踪

**中图分类号:**TN912.3 **文献标志码:**A

## Compressed-Sensing-Based Speech Sparse Representation with K-L Expansion

Zeng Li<sup>1</sup>, Zhang Xiongwei<sup>1</sup>, Chen Liang<sup>2</sup>, Yang Jibin<sup>1</sup>, Jia Chong<sup>1</sup>

(1. College of Command Information Systems, PLA University of Science and Technology, Nanjing, 210007, China;  
2. College of Communications Engineering, PLA University of Science and Technology, Nanjing, 210007, China)

**Abstract:** To overcome the infeasibility of real-time transmission of KLT sparsifying basis, a speech sparse representation algorithm is presented based on Karhunen-Loeve(K-L) expansion. With compressed sensing(CS) theory a speech autocorrelation model is built. Fredholm integral equation is solved and the model parameter is estimated. This parameter is used to construct incoherent dictionary and can be easily transferred. Thereafter, low-dimensional measurements are obtained by sensing the sparse vector with a stochastic matrix. Reconstruction experiments show that the proposed algorithm outperforms the existing methods in dictionary adaptability and reconstruction quality.

**Key words:** compressed sensing; K-L expansion; sparse representation; matching pursuit

## 引 言

在“信息爆炸”的时代,如何获取和处理海量信号,成为人们重点关注的问题。信号稀疏表示思想为解决这一问题提供了新的途径;若对信号进行正交变换、投影或用不同向量线性组合逼近,则可以用很少的数据重新进行刻画。2006年,Candès等撰文详细论述了压缩感知(Compressed sensing,CS)理论<sup>[1,2]</sup>。该理论将信号稀疏作为前提,以远低于 Nyquist 定理的速率对信号进行观测和处理,进一步提高了稀疏化研究的意义。信号稀

疏基的寻找和构造,也成为近年研究的热点问题。

实际信号由于成份复杂,多数情况下无法由单一基函数得到其稀疏表示。Mallat等在总结前人研究成果的基础上,提出了信号在冗余字典上分解的思想<sup>[3]</sup>,指出了冗余字典对于信号稀疏表示的必要性和重要性。但如何从冗余字典中选取原子对信号进行最优描述属于非确定多项式(Nondeterministic polynomial complete, NP)难题<sup>[4]</sup>。为解决该问题,Donoho等提出了利用最小化  $l_1$  范数寻找最优稀疏基的方法,被称为基追踪(Basis pursuit, BP)<sup>[5]</sup>,随后 Michael Elad等从理论上分析给出了联合冗余字典信号稀疏表示惟一性条件<sup>[6]</sup>。

**基金项目:**国家自然科学基金(61072042)资助项目;江苏省自然科学基金(BK2012510)资助项目;解放军理工大学预研基金(20110211)资助项目。

**收稿日期:**2012-04-18; **修订日期:**2013-02-28

另一方面,由于相干度能为字典冗余性提供新的测度,学者们开始对非相干冗余字典进行研究<sup>[7]</sup>。Peyré 提出了把单个正交基扩展到多个正交基构成联合冗余字典的方法<sup>[8]</sup>。

目前,CS 研究更多地关注其在图像、雷达信号处理中的应用<sup>[2,9]</sup>,对语音信号的压缩感知应用关注较少。其原因在于语音产生的生理机制导致信号具有短时平稳特性,在不同的时段特征区别比较明显,这使得语音稀疏化更加困难。文献<sup>[10-14]</sup>对语音信号稀疏性进行了讨论,常见的正交变换域包括离散傅里叶变换(Discrete Fourier transform, DFT),离散小波变换(Discrete wavelet transform, DWT),离散余弦变换(Discrete cosine transform, DCT)等。其中 DFT 是最常见的正交变换,适用于以正弦函数为内核而构成的信号;DWT 变换适用于非平稳信号,可以多尺度自适应地对信号进行分解,得到稀疏表达。DWT 对于信号的稀疏描述结果优于 DFT,但变换后的系数中包含较多的非零小分量,影响信号重构效果;DCT 变换的基函数缺乏时间/空间分辨率,不能有效地提取对应语音具有时频局部化性质的特征,并且与 DWT 类似,其变换后的非稀疏部分仍有大量小系数非零。

KLT 变换是建立在统计特性基础上的一种变换,其相关性好且为均方误差(Mean square error, MSE)意义下的最佳变换,可以对语音信号进行较好的稀疏表达。然而,构造 KLT 变换矩阵的每帧语音具有不同形式,导致在降低观测维度的同时又增加了需要传输或存储的数据量。文献<sup>[12]</sup>提出了模板匹配方法对矩阵进行逼近,取得了一定效果,但清音部分的稀疏化效果仍不尽理想;文献<sup>[14]</sup>将基于自相关求取的 KLT 矩阵进行了模型逼近,但参数确定方法不够准确。

针对上述不足,本文对基于 Karhunen-Loeve (K-L)展开的语音稀疏表示算法进行了研究。首先对语音信号自相关函数进行建模,然后结合模型函数求解 Fredholm 积分方程,构造完备非相干基,并求取模型参数,最终得到语音信号离散非相干冗余字典。在此基础上实现了语音信号压缩感知和重构。实验结果表明,基于 K-L 正交展开的语音自适应冗余字典可对信号进行良好的稀疏化;精确的语音自相关模型参数计算能够进一步提高重构精度。

## 1 基于 K-L 分解的正交字典

对于随机信号,K-L 分解把随机过程描述为一组由互不相关的随机系数(主成分)所调制的确定性函数(正交模态)的线性组合形式,仅需少量的展开项就可以表征随机过程的主要能量和相干结构<sup>[15]</sup>,为信号稀疏化提供了新的途径。

### 1.1 K-L 展开

假设二阶矩连续实随机过程  $\{x(t), t \in [0, 1]\}$ ,其 K-L 展开记为

$$x(t) = \sum_{n=1}^{\infty} a_n \varphi_n(t) \quad (1)$$

式中:系数  $a_n = \int_0^1 x(t) \varphi_n(t) dt$ ,正交 K-L 基  $\{\varphi_n(t)\}$  是信号自相关函数  $R_x(t, u)$  的特征函数组,  $\varphi_n(t)$  和与之对应的特征值满足 Fredholm 积分方程

$$\lambda \varphi_n(t) = \int_0^1 R_x(t, u) \varphi_n(u) du \quad (2)$$

正交基  $\varphi_n(t)$  即可作为信号稀疏字典。其获得的关键在于对 Fredholm 积分方程的求解。在实际问题中,由于自相关函数形式复杂,积分方程求解较为困难,需将函数进行简化。对于语音信号,利用其自相关函数在时延较小时衰减较快的特性,用指数函数来逼近:  $R_x(t, u) = r_x(0) e^{-\mu|t-u|}$ ,  $\mu$  为衰减系数。

为将信号稀疏表示,从字典中选择部分原子线性组合对原信号进行逼近。由于字典原子正交,只需选出与  $x$  的内积  $|\langle x, \varphi_i(t) \rangle|$  最大的  $K$  个基向量即可。即选出的原子下标集合为  $U = \{i: |\varphi_i|_{i \in U} > |\varphi_i|_{i \notin U}\}$ ,则信号可记为

$$\hat{x}(t) = \sum_{i=1}^K a_i \varphi_i(t) \quad (3)$$

### 1.2 Fredholm 积分方程求解

将语音自相关逼近函数代入式(2),求解下列微分方程<sup>[14]</sup>

$$\lambda \varphi(t) = \int_0^1 R_x(t, u) \varphi(u) du = r_x(0) \left( e^{-\mu t} \int_0^t e^{\mu u} \varphi(u) du + e^{\mu t} \int_t^1 e^{-\mu u} \varphi(u) du \right) \quad (4)$$

对式(4)两边分别求 1, 2 阶导

$$\lambda \varphi'(t) = r_x(0) \left( -\mu e^{-\mu t} \int_0^t e^{\mu u} \varphi(u) du + \mu e^{\mu t} \int_t^1 e^{-\mu u} \varphi(u) du \right) \quad (5)$$

$$\lambda\varphi''(t) = r_x(0) \left( \mu^2 \int_0^t e^{-\mu(t-u)} \varphi(u) du + \mu^2 \int_t^1 e^{-\mu(u-t)} \varphi(u) du - 2\mu\varphi(t) \right) \quad (6)$$

当  $t=0$  和  $t=1$  时,由式(5,6)可得边界条件

$$\varphi'(0) = \mu\varphi(0), \varphi'(1) = -\mu\varphi(1) \quad (7)$$

代入边界条件,求解式(6),可得方程通解

$$\varphi(t) = c_1 \cos(\omega t) + c_2 \sin(\omega t) \quad c_1, c_2 \in \mathbf{R} \quad (8)$$

其中  $\omega = \sqrt{[2r_x(0)\mu - \mu^2\lambda]/\lambda}$ 。进一步地,可得方程特解

$$\varphi_n(t) = (n\pi/\mu) \cos(n\pi t) + \sin(n\pi t) \quad n \in \mathbf{Z} \quad (9)$$

### 1.3 字典完备化

特解中,  $\varphi_0(t) = 0$  不能作为特征函数,故取正交特征函数组为

$$\varphi_n(t) = \frac{n\pi}{\mu} \cos(n\pi t) + \sin(n\pi t) \quad n \in \mathbf{Z} \setminus \{0\} \quad (10)$$

再加入  $\varphi_0(t) = 1$ ,则可构建完备的正交 K-L 基函数组:  $\mathbf{E} = \{\varphi_0(t)\} \cup \{\varphi_n(t), n \in \mathbf{Z} \setminus \{0\}\}$ ,其中任意不同的基均正交,即

$$\langle \varphi_i(t), \varphi_j(t) \rangle = \int_{-\infty}^{\infty} \varphi_i(t) \varphi_j(t) dt = 0 \quad i \neq j, \forall i, j \in \mathbf{Z} \quad (11)$$

## 2 语音信号自适应冗余字典

由语音自相关模型所获得的 K-L 正交基已具备结构化特征,仅需传输信号自相关函数模型参数  $\mu$ ,在解码端即可根据该结构直接获得对应的码本,解决了 KLT 语音稀疏化方法中稀疏基矩阵不便实时传输的问题。

### 2.1 参数 $\mu$ 的确定

结构化的正交字典具有统一模式,解码端仅需参数  $\mu$  即可构造出相应的字典,因此精确的参数计算可以保证信号更好地进行稀疏表示。文献[14]在这里采用了曲线拟合的方式获得该值,但拟合的方式存在一定的误差,导致获得的字典不能最大化地匹配当前语音帧,对信号逼近残差较大。本文采用定量数值分析的方法,来获得更为精确的模型参数。

语音信号  $\mathbf{x} \in \mathbf{R}^M$  在时延  $\tau$  的自相关函数无偏估计为

$$R_x(\tau) = \frac{1}{(M-\tau)} \sum_{i=1}^{M-\tau} x(i)x(i+\tau), \text{ 通过求}$$

$\min G(\mu) = \|R_x - \hat{R}_x\|_2^2$  来获得最优值  $\mu^*$ ,其中  $\hat{R}_x(\tau) = r_x(0)e^{-\mu|\tau|}, \tau = 0, \dots, M-1$ 。

$$\min_{\mu} G(\mu) = \|R_x - \hat{R}_x\|_2^2 = \sum_{\tau=1}^{M-1} (R_x(\tau) - r_x(0)e^{-\mu|\tau|})^2 \quad (12)$$

求导并令

$$G'(\mu) = 2 \sum_{\tau=1}^{M-1} (R_x(\tau) - r_x(0)e^{-|\tau|\mu}) \cdot [-r_x(0)(-|\tau|)] = 0 \quad (13)$$

得到

$$\sum_{\tau=1}^{M-1} \tau (R_x(\tau) - r_x(0)e^{-|\tau|\mu}) = \sum_{\tau=1}^{M-1} \tau R_x(\tau) - \tau \sum_{\tau=1}^{M-1} r_x(0)e^{-\tau\mu} = 0 \quad (14)$$

式(14)中第1个求和项和  $r_x(0)$  均为常量,可将式子简记为

$$\sum_{\tau} \tau e^{-\tau\mu} = \frac{\sum_{\tau} \tau R_x(\tau)}{\sum_{\tau} r_x(0)} = C \quad (15)$$

记  $e^{-\mu} = q$ ,则式(15)成为一个高次方程

$$\sum_{\tau=1}^{M-1} \tau \cdot q^{\tau} = C \quad (16)$$

式中  $C$  是常数。注意到  $q \in (0, 1)$ ,且方程左边为单调函数,采用二分法在  $(0, 1)$  区间搜索,可获得参数  $\mu$  的精确值,具体步骤为

**步骤 1:** 令  $f(t) = \sum_{n=1}^{M-1} n \cdot t^n - C$ ,初始化  $i=1$ ,

$a_i=0, b_i=1, c_i=(a_i+b_i)/2$ ;

**步骤 2:** 计算  $f(a_i), f(b_i), f(c_i), i=i+1$ ;

**步骤 3:** 若  $f(c_i)$  与  $f(a_i)$  异号,则  $a_i = a_{i-1}, b_i = c_{i-1}, c_i = (a_i + b_i)/2$ ,否则  $a_i = c_{i-1}, b_i = b_{i-1}, c_i = (a_i + b_i)/2$ ;

**步骤 4:** 若  $|f(c_i)| < \epsilon$  (误差阈值),则  $\mu^* = c_i$ ,算法停止;否则返回步骤 2。

### 2.2 字典离散化

针对数字信号处理,需要离散的字典原子,故对  $\varphi_n(t)$  在  $0 \leq t \leq 1$  进行均匀采样获得原子  $\mathbf{e}_n = [e_n(1), \dots, e_n(i), \dots, e_n(M)]^T$ ,其中

$$e_n(i) = \frac{n\pi}{\mu} \cos\left(\frac{n\pi(i-1)}{M-1}\right) + \sin\left(\frac{n\pi(i-1)}{M-1}\right) \quad i=1, \dots, M \quad (17)$$

再联合  $\mathbf{e}_0 = [1, \dots, 1]^T$ ,构建离散语音信号稀疏字典

$$\mathbf{D} = \{\mathbf{e}_0\} \cup \{\mathbf{e}_n, n \in \mathbf{Z} \setminus \{0\}\} \quad (18)$$

由于离散区间采样,此处的字典原子已不再是数学意义上严格正交;但在信号维度  $M$  较大的情况下,各原子间相关性极小,字典成为非相干冗余字典。同时,由于每帧语音自相关模型参数  $\mu$  均由当前帧信号获得,因此字典具有自适应性。

### 3 信号稀疏表示及压缩感知

#### 3.1 信号稀疏观测

鉴于字典的冗余性(原子数  $n > \text{Card}(\mathbf{D})$ ),从字典中选取原子对信号进行逼近,可能有多种组合方式。这对于随机的冗余字典,属于 NP 问题。而对于正交字典,如何得到信号最佳逼近就较为容易,即保留与信号  $\mathbf{x}$  的内积  $|\langle \mathbf{x}, \mathbf{d}_i \rangle|$  最大的一部分原子即可。由于非相干字典中原子相干系数充分小,也可采用正交字典的原子选择方式。K-L 分解得到的字典原子相关度较小,可以认为是非相干冗余字典。

获得信号稀疏向量  $\mathbf{s}$  后,用随机矩阵  $\Phi \in \mathbf{R}^{N \times M}$  对稀疏向量进行观测,获得低维观测值  $\mathbf{y} = \Phi \mathbf{s} \in \mathbf{R}^N$  ( $N \ll M$ )。联合参数  $\mu^*$  经信道传输,在解码端对原语音进行重构。

#### 3.2 重构方法

根据 CS 理论,稀疏信号的重构即通过该信号在已知测量矩阵  $\Phi$  上的  $M$  个非相干的线性投影  $\mathbf{y} = \Phi \mathbf{s} \in \mathbf{R}^N$  来获得重建。由于待重构信号  $\mathbf{s}$  仅有  $K$  个非零元,因此测量向量  $\mathbf{y}$  可以看作  $\Phi$  中  $K$  个列向量的线性组合。利用  $\mathbf{y}$  和  $\Phi$  重构  $\mathbf{s}$ ,需在测量向量集合  $\Phi$  中准确找到参与测量的  $K$  个列向量(原子)。该问题与信号稀疏分解时搜索字典原子问题有着相似的数学模型,因此 CS 理论中的可压缩信号的重构问题可以类似地看成信号稀疏分解问题。

基于贪婪迭代思想的匹配追踪算法的原理是:通过迭代的方法依次找出待重建信号的支撑,基于某种贪婪准则一次确定一个或多个待重构信号的构成原子。本文采用匹配追踪(Matching pursuit, MP)和正交匹配追踪(Orthogonal matching pursuit, OMP)算法进行信号重构。MP 算法流程如下:

**步骤 1:** 初始化迭代残差向量  $\mathbf{r}_0 = \mathbf{y}$ , 初始重构向量  $\hat{\mathbf{s}} = \mathbf{0}$ , 初始化迭代次数  $i = 1$ ;

**步骤 2:** 搜索和残差  $\mathbf{r}_{i-1}$  相关度最大的原子索引  $\lambda_i: \lambda_i = \underset{j=1,2,\dots,n}{\operatorname{argmax}} |\langle \mathbf{r}_{i-1}, \phi_j \rangle|$ ;

**步骤 3:** 更新迭代残差,并计算稀疏向量重构值:  $\hat{\mathbf{s}}_i = \langle \mathbf{r}_{i-1}, \phi_{\lambda_i} \rangle \phi_{\lambda_i}, \mathbf{r}_i = \mathbf{r}_{i-1} - \hat{\mathbf{s}}_i$ ;

**步骤 4:**  $i = i + 1$ , 若  $i = K$ , 迭代结束, 否则返回步骤 2。

MP 算法需要其迭代残差  $\mathbf{r}_i$  与当次所选原子  $\phi_{\lambda_i}$  的正交性来保证算法收敛,但无法保证  $\mathbf{r}_i$  与之前所选择的原子集合  $\{\phi_{\lambda_1}, \phi_{\lambda_2}, \dots, \phi_{\lambda_{i-1}}\}$  的正交性,导致原子搜索常常陷入局部最优。OMP 算法通过递归对已选择原子集合进行正交化,有效克服了 MP 的不足。

### 4 实验结果及分析

实验分别采用约 20 000 帧实验室环境下男、女声语音,以 8 kHz 速率采样,每帧长度为 37.5 ms,即 300 样点。分别针对清音、浊音、过渡音三类信号的稀疏性,利用文中方案进行了测试,比较了自相关函数模型参数曲线拟合和数值计算方法,并对重构算法性能进行了主、客观评价。实验中,取字典参数估计中的误差阈值为  $e = 10^{-4}$ 。同时,对信号及对应稀疏向量均做归一化处理。

#### 4.1 信号稀疏性及重构效果

语音信号清音、浊音、过渡帧在 K-L 非相干冗余字典上的稀疏性见图 1。由字典原子结构可知,随着  $n$  增大,原子能量和频率也随之增大,使得语音信号在字典上的分解能量主要集中在稀疏向量的前段,且大系数数量很少,保证了分解系数的稀疏性。

3 类语音帧中,浊音稀疏性最好,清音和过渡音次之。对于结构性很强的浊音帧,分解系数能量全部集中在前段系数中;清音帧和过渡帧信号具有一定随机信号特征,需要少量高频结构对其进行描述,因此稀疏向量的中、后段也存在少量非零系数。由于决定听觉效果的音素成份主要为浊音,且清音、过渡音在语音信号中所占比例远不及浊音,因此在稀疏化的过程中,认为这部分小系数对信号的观测和重构质量影响较小,将其作为零向量进行观测。

字典原子决定了信号稀疏表示结果,同时影响重构精度。基于 K-L 分解的冗余字典需要信号自相关模型参数对其原子进行刻画。本文对已有的最优参数计算方法进行了改进。定义稀疏度  $S = K/M$ ,图 2 比较了  $S = 5\%$  时两种字典对应的信号重构结果。

可以看到,采用曲线拟合方法得到的字典对信号稀疏化后重构结果有包络上的衰减,其原因用负指数函数曲线拟合信号自相关时,存在一定误差,所得原子过分集中于低频部分,导致描述信号

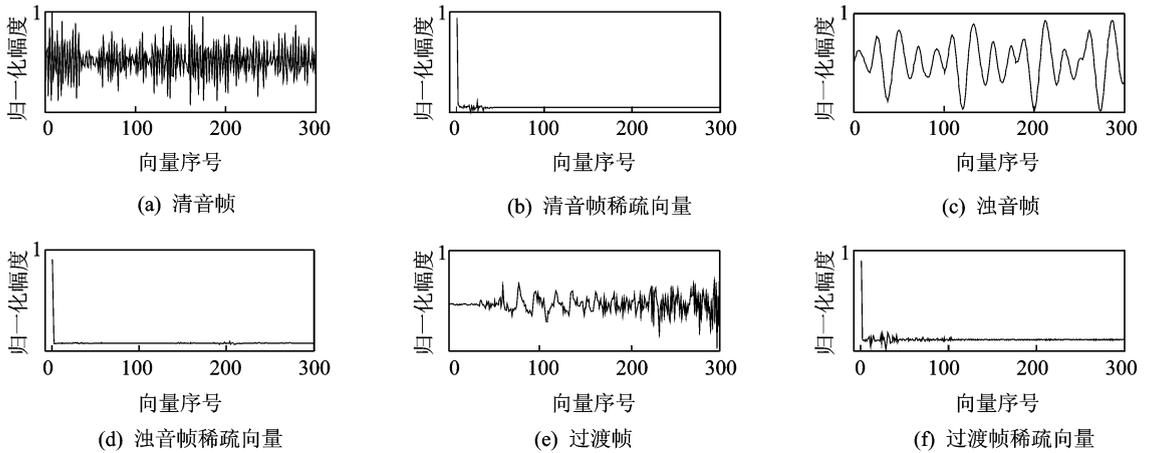


图 1 三类语音信号稀疏性

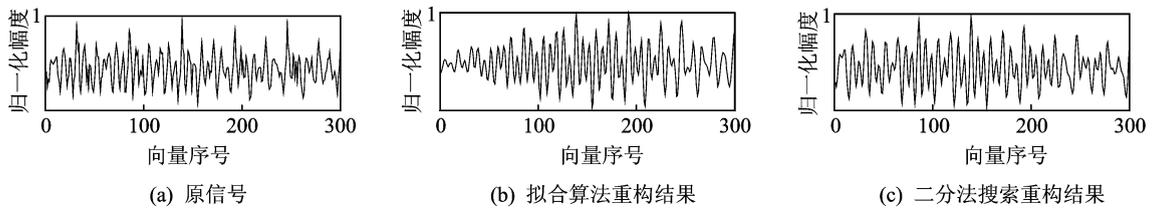


图 2 两种参数估计方法重构结果对比

低频的原子过多,从而部分中高频分量在观测时未能保留。采用二分法搜索,能更好地匹配当前语音帧自相关函数,减少函数逼近误差,从而得到更优的字典,减少了信号稀疏化时高频细节分量的丢失。

### 4.2 主观评价结果

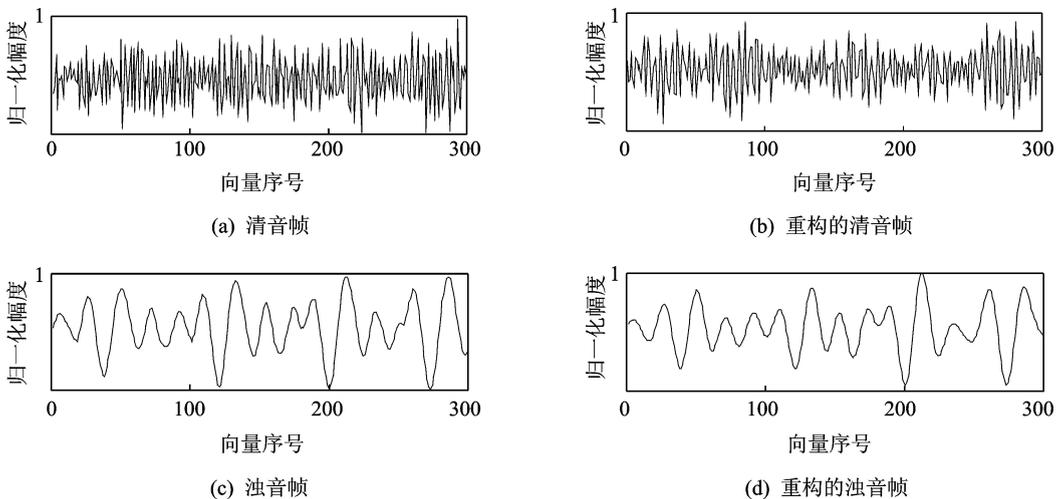
结合 MP,OMP 算法,用 ITU-T P. 862 推荐的平均意见得分(Mean opinion score, MOS)对两种方法的重构语音进行了评估。测试了稀疏度 5%~20%时两者的 MOS 分。图 3(a)~(f)给出了 3 类典型语音帧采用拟合法和数值分析方法的

重构波形,图 3(g)为两种字典 MOS 分比较结果。结果表明:采用数值分析方法的重构语音质量高于拟合方法;在不同稀疏度下,采用数值分析方法得到的字典,MP 和 OMP 算法重构信号 MOS 分平均提高为 0.091 3,0.093 7。

### 4.3 客观评价结果

实验采用平均分信噪比(Average-subsection signal-to-noise-ratio, SegSNR)<sup>[16]</sup>作为语音质量客观评价标准,对重构前后男女声进行评价

$$\text{SegSNR} = 10 \left( \frac{1}{\text{Frames}} \sum_{i=1}^{\text{Frames}} \text{SNR}_i \right) \quad (19)$$



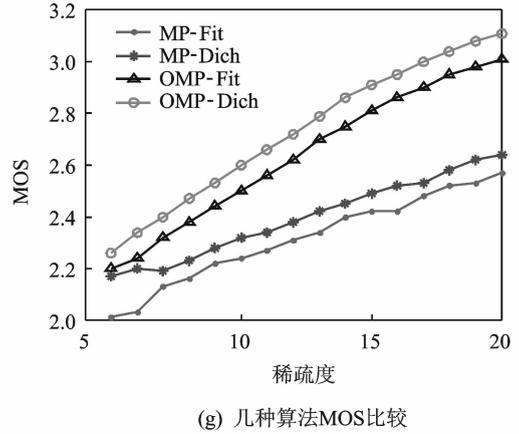
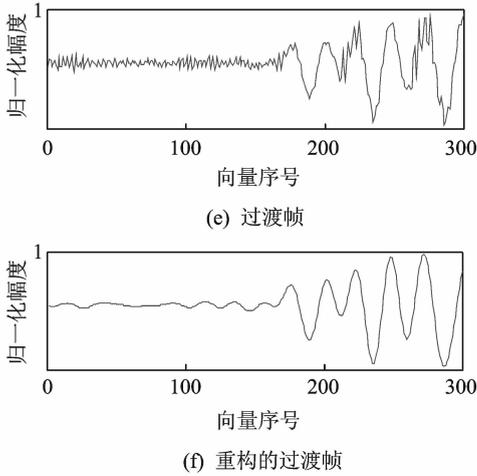


图3 信号重构结果及主观评价结果比较

$$\text{其中 } \text{SNR}_i = \sum_{j=1}^N \frac{x_i^2(j)}{[x_i(j) - \hat{x}_i(j)]^2}。$$

将重构信号按信噪比不大于 0, 0~10, 10~20, 20~50 dB, 大于 50 dB 分别划分为质量很差、差、中等、好、很好。表 1 给出了不同稀疏度对应的信号重构质量。

表 1 男女声不同稀疏度下原始/重构信号

S/%	5	6	7	8	9	10	11	12	13	14	15
男声	8.70	9.82	10.49	11.72	12.93	13.51	14.61	15.31	15.88	16.43	16.86
女声	8.12	8.73	9.95	10.67	11.37	12.55	13.84	14.66	15.54	16.05	16.52

## 5 结束语

本文基于 K-L 分解设计了一种改进型语音信号非相干冗余字典。通过对语音自相关函数进行建模逼近, 采用数值分析的方法计算更为精确的模型参数, 解决了 KLT 变换算法中变换基矩阵不便实时传输的不足。在此基础上, 设计并实现了语音信号压缩感知及高精度重构。实验结果表明, 本文算法可以获得适应性更好的语音稀疏字典, 在稀疏度较低条件下重构语音质量较好, 从而说明文中算法对语音信号稀疏表示具有指导意义。

### 参考文献:

- [1] Donoho D. Compressed sensing[J]. IEEE Trans Inf Theory, 2006, 52(4): 1289-1306.
- [2] Candes E, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. IEEE Trans Inf Theory, 2006, 52(2): 489-509.
- [3] Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries[J]. IEEE Trans Signal Process, 1993, 41: 3397-3415.

语音重构误差随着稀疏度的提高而减少; 在相同稀疏度条件下, 男声重构误差小于女声重构误差, 这是由于女声基音高于男声, 因此在高频原子部分小系数分量多于男声, 在观测时将小系数作为零向量, 这导致女声信号重构误差略大于男声。

- [4] Davis G, Mallat S, Avellaneda M. Adaptive greedy approximation[J]. Constr Approx, 1997, 13(1): 57-98.
- [5] Donoho D, Huo X. Uncertainty principles and ideal atomic decompositions [J]. IEEE Trans Inf Theory, 2001, 47: 2845-2862.
- [6] Elad M, Bruckstein A M. A generalized uncertainty principle and sparse representation in pairs of bases [J]. IEEE Trans Inf Theory, 2003, 49: 1579-1581.
- [7] 张春梅, 尹忠科, 肖明霞. 基于冗余字典的信号超完备表示与稀疏分解[J]. 科学通报, 2006, 51(6): 628-633.  
Zhang Chunmei, Yin Zhongke, Xiao Mingxia. Signal overcomplete representation and sparse decomposition based on redundant dictionary[J]. Chinese Science Bulletin, 2006, 51(6): 628-633.
- [8] Peyre G C. Best basis compressed sensing[J]. IEEE Trans Signal Process, 2010, 58(5): 2613-2622.
- [9] 张弓, 杨萌, 张劲东, 等. 压缩感知在雷达目标探测与识别中的研究进展[J]. 数据采集与处理, 2012, 27(1): 1-12.

- Zhang Gong, Yang Meng, Zhang Jingdong, et al. Advances in theory and application of compressed sensing in radar target detection and recognition[J]. Journal of Data Acquisition and Processing, 2012, 27(1): 1-12.
- [10] 马蜂, 张宁, 戴礼荣. 基于语音信号稀疏性的 FDICA 初始化和后处理方法 [J]. 数据采集与处理, 2012, 27(2): 210-217.  
Ma Feng, Zhang Ning, Dai Lirong. FDICA Initialization and Post-Processing Method Based on Sparseness of Speech[J]. Journal of Data Acquisition and Processing, 2012, 27(2): 210-217.
- [11] Griffin A, Tsakalides P. Compressed sensing of audio signal using multiple sensors[C]//16 th European Signal Processing Conference. Lausanne, Switzerland:[s. n.], 2008: 16:1-4.
- [12] 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 & Information Technology, 2009, 31(12), 2948-2952.
- [13] Glacobello 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]. IEEE Signal Processing Letters, 2010, 17(1): 103-106.
- [14] 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 & Information Technology, 2011, 33(10): 2372-2377.
- [15] 李杰, 刘章军. 基于标准正交基的随机过程展开法 [J]. 同济大学学报: 自然科学版, 2006, 34(10): 1279-1283.  
Li Jie, Liu Zhangjun. Expansion method of stochastic processes based on normalized orthogonal bases [J]. Journal of Tongji University ; Natural Science, 2006, 34(10): 1279-1283.
- [16] Tingting Xu, Zhen Yang, Xi Shao. Novel speech secure communication system based on information hiding and compressed sensing[C]// 4th International Conference on Systems and Networks Communications. Porto, Portugal: IEEE, 2009: 201-206.

**作者简介:**曾理(1985-),男,博士研究生,研究方向:语音信号处理、压缩感知理论及应用,E-mail:38392177@qq.com;张雄伟(1965-),男,教授,博士生导师,研究方向:多媒体信号处理、数字通信;陈亮(1974-),男,教授,研究方向:语音编码、隐蔽通信;杨吉斌(1978-),男,副教授,研究方向:语音编码、语音增强;贾冲(1979-),男,博士,研究方向:语音编码、多媒体信号处理。