

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

叶 蕾¹, 杨 震², 王天荆², 孙林慧¹

(1. 南京邮电大学通信与信息工程学院, 江苏南京 210003; 2. 南京邮电大学信号处理与传输研究院, 江苏南京 210003)

摘 要: 基于语音信号在离散余弦域上的近似稀疏性, 针对采用随机高斯观测矩阵及线性规划方法进行语音压缩感知与重构时, 重构零(近似零)系数定位能力差而导致重构效果不好的缺点, 本文提出一种新的行阶梯矩阵做观测矩阵, 用对偶仿射尺度内点重构算法对语音进行压缩感知与重构, 并对该算法下的重构性能进行理论分析. 语音压缩感知仿真结果表明, 在离散余弦基下, 压缩比(观测序列与原始序列样值数之比)为 1:4 时, 行阶梯观测矩阵下的平均重构信噪比比随机高斯观测矩阵下提高 9.73dB, 平均 MOS 分比随机高斯观测矩阵下提高 1.22 分.

关键词: 压缩感知; 离散余弦基; 观测矩阵; 行阶梯矩阵; 对偶仿射尺度内点法

中图分类号: TN912.3 **文献标识码:** A **文章编号:** 0372-2112 (2012) 03-0429-06

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2012.03.003

Compressed Sensing of Speech Signal Based on Row Echelon Measurement Matrix and Dual Affine Scaling Interior Point Reconstruction Method

YE Lei¹, YANG Zhen², WANG Tian-jing², SUN Lin-hui¹

(1. College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210003, China;

2. Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210003, China)

Abstract: Based on the approximate sparsity of speech signal in discrete cosine basis, this paper proposes a new algorithm of compressed sensing of speech signal based on special row echelon measurement matrix and dual affine scaling interior point reconstruction method. This algorithm can resolve the problem of inaccuracy of location of reconstruction coefficient which is zero or nearly zero of compressed sensing based on Gaussian measurement matrix and linear programming to some extent. The reconstruction performance of this algorithm is analyzed theoretically. The simulation results of compressed sensing of speech signal show when the reduction ratio (the ratio of numbers of measurements and original samples) is 1:4 based on the discrete cosine basis, the average SNR of reconstruction signal based on the special row echelon measurement matrix is 9.73 dB higher than the Gaussian measurement matrix, and the average MOS score of reconstruction signal based on the special row echelon measurement matrix is 1.22 higher than the Gaussian measurement matrix.

Key words: compressed sensing; discrete cosine basis; measurement matrix; row echelon matrix; dual affine scaling interior point method

1 引言

压缩感知 (CS; Compressed Sensing) 理论^[1~4] 是近年来针对稀疏信号提出的一种新兴采样理论, 以远低于奈奎斯特频率的速度成功实现了对信号的同时采样与压缩. 该理论在日臻完善的同时已经在信源编码、传感器

网络、信号检测、医学图像处理、雷达遥感、生物传感、模式识别、盲源分离、频谱分析等领域获得了越来越多的应用^[5~14]. 本文研究适合语音信号的压缩感知, 基于语音信号在离散余弦域 (DCT; Discrete Cosine Transform) 上的近似稀疏性, 研究 l_1 范数优化重构算法下的语音信号压缩感知, 针对随机高斯观测矩阵下重构时对零系

数(或近似于零的系数)位置不能有效定位的缺点,提出一种新的行阶梯矩阵作为观测矩阵,并采用解决 l_1 范数问题的对偶仿射尺度内点法进行重构,并对该算法下的重构性能进行理论分析. 语音压缩感知的仿真结果表明,在压缩比(观测序列与原始序列样值数之比)为 1:4 时,行阶梯观测矩阵下的重构信噪比(SNR: Signal to Noise Ratio)比随机高斯观测矩阵下平均提高 9.73dB, MOS(Mean Opinion Score)分比随机高斯观测矩阵下平均提高 1.22 分.

2 行阶梯观测矩阵、对偶仿射尺度内点重构下的压缩感知算法

2.1 压缩感知基本理论

压缩感知理论主要涉及三个核心问题:信号稀疏表示即变换矩阵 Ψ , 观测矩阵 Φ , 以及重构算法的设计^[15].

假设原始信号 $\mathbf{x}_0 \in R^n$ 在正交基 $\Psi = \{\psi_i | \psi_i \in R^n, i = 1, 2, \dots, n\}$ 上是 K 稀疏的, 即

$$\mathbf{x}_0 = \Psi^{-1} \boldsymbol{\theta} \quad (1)$$

稀疏系数向量

$$\boldsymbol{\theta} = \Psi \mathbf{x}_0 = (\theta_1, \theta_2, \dots, \theta_n)^T, \|\boldsymbol{\theta}\|_0 = K.$$

观测值 $\mathbf{y} = (y_1, y_2, \dots, y_m)^T$.

$$\mathbf{y} = \Phi \mathbf{x}_0 = \Phi \Psi^{-1} \boldsymbol{\theta} = \mathbf{H} \boldsymbol{\theta} \quad (2)$$

其中 $\Phi = (\phi_1, \phi_2, \dots, \phi_n)$, $(\phi_i \in R^m, i = 1, 2, \dots, n)$. $\mathbf{H} = \Phi \Psi^{-1}$ 称为 CS 矩阵. 由 \mathbf{y} 重构 \mathbf{x}_0 或 $\boldsymbol{\theta}$ 时需要求解式(2), 当 $m < n$ 时, 它有无穷多个解, 所以, 由式(2)无法求得确切的重构信号. 但根据 BP (Basis Pursuit) 算法^[16~19], 可通过求解 l_1 最优化问题得到它的等价解:

$$\min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_1 \text{ subject to } \mathbf{y} = \Phi \Psi^{-1} \boldsymbol{\theta} = \mathbf{H} \boldsymbol{\theta} \quad (3)$$

2.2 行阶梯观测矩阵、对偶仿射尺度内点法下的压缩感知

根据文献^[16~20]将式(2)问题转为式(3)所示的线性规划问题, 通过线性规划解决 l_1 范数优化问题来重构信号. 线性规划的一系列算法, 都是基于杰出数学家 Dantzig 提出的单纯形法. 然而单纯形法计算时间过长, 对于较大型问题来说, 内点法比单纯形法占有明显的优势^[21, 22]. 本文通过对偶仿射尺度内点法(Dual Affine Scaling Interior Point Method^[21, 22])求解式(4)的对偶问题式(5)来求得最优解 $\boldsymbol{\theta}^*$.

$$\min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \text{ subject to } \mathbf{A} \mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq 0 \quad (4)$$

其中 $\mathbf{c} = (1, \dots, 1)^T$, $\mathbf{A} = (\mathbf{H}, -\mathbf{H})$, $\mathbf{b} = \mathbf{y}$, $\mathbf{x} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}$,

$$\boldsymbol{\theta}^* = \mathbf{u} - \mathbf{v}.$$

$$\max_{\mathbf{z}} \mathbf{b}^T \mathbf{z} \text{ subject to } \mathbf{A}^T \mathbf{z} = \mathbf{c}$$

$$\mathbf{z} \geq 0 \quad (5)$$

对偶仿射尺度内点法计算步骤如下:

步骤 1 给定初始内点 $\mathbf{z}^{(0)}$, 参数 $\gamma \in (0, 1)$, 容许限 $\epsilon > 0$, 置 $k = 1$.

步骤 2 计算松弛变量

$$\mathbf{v}^{(k)} = \mathbf{c} - \mathbf{A}^T \mathbf{z}^{(k)} \quad (6)$$

步骤 3 置对角矩阵

$$\mathbf{D}_k = \text{diag}\left(\frac{1}{v_1^{(k)}}, \frac{1}{v_2^{(k)}}, \dots, \frac{1}{v_m^{(k)}}\right) \quad (7)$$

步骤 4 计算可行方向

$$\mathbf{d}_z = (\mathbf{A} \mathbf{D}_k^2 \mathbf{A}^T)^{-1} \mathbf{b} \quad (8)$$

步骤 5 令

$$\mathbf{d}_v = -\mathbf{A}^T \mathbf{d}_z \quad (9)$$

步骤 6 令步长

$$\lambda = \gamma \cdot \min \left\{ \frac{v_i^{(k)}}{-(\mathbf{d}_v)_i} \mid (\mathbf{d}_v)_i < 0, i \in \{1, \dots, m\} \right\} \quad (10)$$

步骤 7 置

$$\mathbf{z}^{(k+1)} = \mathbf{z}^{(k)} + \lambda \mathbf{d}_z \quad (11)$$

步骤 8 若

$$\|\mathbf{b}^T \mathbf{z}^{(k+1)} - \mathbf{b}^T \mathbf{z}^{(k)}\| / \|\mathbf{b}^T \mathbf{z}^{(k)}\| < \epsilon \quad (12)$$

则停止计算, $\mathbf{z}^{(k)}$ 为式(5)最优解, 相应的式(4)最优解为

$$\mathbf{x}^{(k)} = \mathbf{D}_k^2 \mathbf{A}^T (\mathbf{A} \mathbf{D}_k^2 \mathbf{A}^T)^{-1} \mathbf{b} \quad (13)$$

否则置 $k = k + 1$, 返回步骤 2.

用对偶仿射尺度内点法得到重构系数 $\boldsymbol{\theta}^*$ 时, 初始内点 $\mathbf{z}_0 = (0.1, 0.1, \dots, 0.1)^T$, 参数 $\gamma = 0.5$, 容许限 $\epsilon = 0.001$. 由式(4)和式(13), 有

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}^{(k)} = \mathbf{x}^{(k)} = \mathbf{D}_k^2 \mathbf{A}^T (\mathbf{A} \mathbf{D}_k^2 \mathbf{A}^T)^{-1} \mathbf{b} \quad (14)$$

由式(3)、式(4)、式(14)得

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}^{(k)} = \mathbf{D}_k^2 \begin{pmatrix} \mathbf{H}^T \\ -\mathbf{H}^T \end{pmatrix} \left((\mathbf{H}, -\mathbf{H}) \mathbf{D}_k^2 \begin{pmatrix} \mathbf{H}^T \\ -\mathbf{H}^T \end{pmatrix} \right)^{-1} \mathbf{H} \boldsymbol{\theta} \quad (15)$$

由于式(7) \mathbf{D}_k 为对角矩阵, \mathbf{D}_k^2 可以用两个对角矩阵 \mathbf{D}_α 和 \mathbf{D}_β 来表示即

$$\mathbf{D}_k^2 = \begin{pmatrix} \mathbf{D}_\alpha & 0 \\ 0 & \mathbf{D}_\beta \end{pmatrix} \quad (16)$$

式(15)可表示为

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}^{(k)} = \begin{pmatrix} \mathbf{D}_\alpha & 0 \\ 0 & \mathbf{D}_\beta \end{pmatrix} \begin{pmatrix} \mathbf{H}^T \\ -\mathbf{H}^T \end{pmatrix} \left(\mathbf{H} \mathbf{D}_\alpha \mathbf{H}^T + \mathbf{H} \mathbf{D}_\beta \mathbf{H}^T \right)^{-1} \mathbf{H} \boldsymbol{\theta} \quad (17)$$

令对角矩阵

$$\mathbf{D}_\gamma = \mathbf{D}_\alpha + \mathbf{D}_\beta \quad (18)$$

式(17)简化为

经过式(32)的运算,可使 F 非对角线上的大部分元素为零. 迭代过程中, D_γ 不一定是单位阵,但由于 D_γ 为对角阵, F 矩阵非对角线上零元素的位置是不变的. 并且,随着迭代次数的增加,根据式(18)和式(22), $\theta^* = uu^{(k)} - vv^{(k)} = D_\gamma F \theta$ 应该逐步向 θ 逼近,在 F 非对角线大部分元素为零及 D_γ 为对角阵的条件下,逐次迭代最后得到的 $D_\gamma F$ 趋向于对角阵(其中前 K 行 K 列构成的子阵趋向单位阵, K 为稀疏度). 当采用式(30)形式的行阶梯矩阵做观测矩阵时,由于 D_γ 是对角阵, F 是趋向对角阵的,因此保证了重构零系数位置的正确性,而高斯观测矩阵做不到这一点.

下面用实验验证采用式(30)形式的行阶梯观测矩阵时的重构效果. 假设某信号 $x_0 \in R^n$ 对应的 n 个 DCT 系数只有前 K 个非零,这里取 $n = 48, K = 8$. 前 8 个非零系数随机产生,后 40 个系数设为 0,得到稀疏系数 θ . 然后由 θ 和 DCT 反变换求出原始信号 x_0 . 观测样值个数 $m = 12$,压缩比为 1:4. 图 1 和图 2 对行阶梯观测矩阵和随机高斯观测矩阵下的稀疏系数重构效果进行对比,分别给出两种矩阵下最后一次迭代的原始 DCT 系数与重构 DCT 系数及 $uu^{(k)}$ 和 $vv^{(k)}$. 可见,重构时式(30)形式的行阶梯观测矩阵比随机高斯观测矩阵对零系数定位能力强.

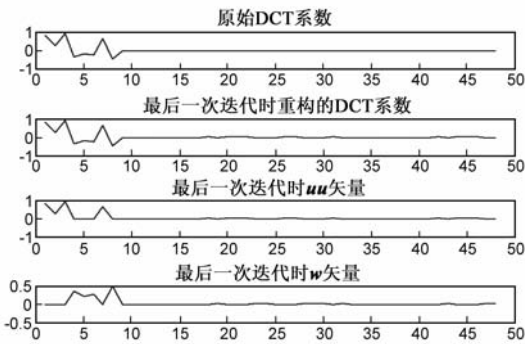


图1 行阶梯观测矩阵下最后一次迭代

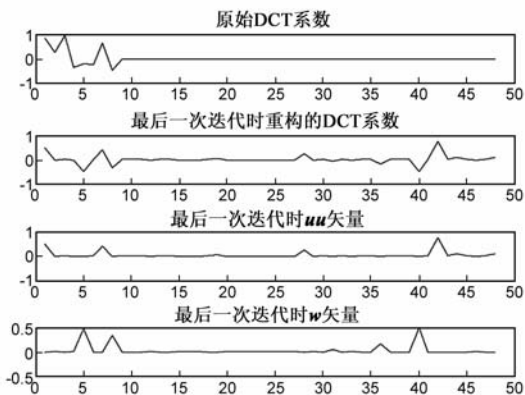


图2 随机高斯观测矩阵下最后一次迭代

表 1 给出不同稀疏度 K 时,两种观测矩阵下系数重构误差对比,误差采用 2 范数相对误差. 原始信号 48 个样值,先随机产生前 K 个非零 DCT 系数,其它系数设为 0,得到稀疏系数 θ . 然后由 θ 和 DCT 反变换求出原始信号 x_0 . 观测样值个数取 12,压缩比为 1:4. 各个 K 下取 100 次实验的平均结果. 可以看出,行阶梯矩阵在 $m \geq K$ (m 为观测样点数 12) 时均能较好的重构系数,而随机高斯矩阵在 $m \geq 4K$ 时才有较好的重构. 也就是说,与随机高斯观测矩阵相比,本文的行阶梯矩阵,可以采用更少的观测点数来获得同样好的重构效果,数据压缩程度更大.

表 1 不同稀疏度 K 两种观测矩阵下系数重构误差对比

K	行阶梯矩阵下系数重构误差	随机高斯矩阵下系数重构误差
1	1.9556×10^{-11}	3.2168×10^{-5}
2	2.6014×10^{-7}	8.1108×10^{-5}
3	2.1611×10^{-7}	6.6456×10^{-5}
4	1.5675×10^{-6}	0.2888
5	8.7123×10^{-7}	0.2656
6	2.9789×10^{-6}	0.7646
7	8.7452×10^{-5}	0.4243
8	2.2717×10^{-6}	0.9679
9	3.1092×10^{-6}	0.5604
10	2.2913×10^{-5}	0.7559
11	8.7011×10^{-6}	0.8951
12	2.4239×10^{-5}	0.9479

3 行阶梯矩阵、对偶仿射尺度内点重构下的语音压缩感知实验

当然,设计观测矩阵要满足 Φ 和 Ψ^{-1} 不相干特性条件,即保证 Φ 与 Ψ^{-1} 的相干度 (coherence) μ 尽可能的小. 相干度定义为^[20]

$$\mu(\Phi, \Psi^{-1}) = \sqrt{N} \max_{\substack{1 \leq k \leq m \\ 1 \leq j \leq n}} |\langle \phi_k, \psi_j^{-1} \rangle| \quad (33)$$

我们对式(30)形式的行阶梯矩阵和随机高斯矩阵与 DCT 基的相干度. 实验中的数据库采用中国科学院自动化所录制的语音库,共 600 个语句,原始语音的采样率为 16kHz. 当压缩比为 1:4,20ms 分帧时,实验结果显示随机高斯矩阵下的相干度为 84.5517,而行阶梯矩阵下的相干度为 5.6562,远远小于前者,满足 Φ 和 Ψ^{-1} 不相干条件,且观测矩阵容量小,计算复杂度低. 我们对行阶梯观测矩阵及随机高斯观测矩阵下的语音压缩感知重构算法性能进行了对比. 表 2 和表 3 分别给出压缩比取 1:4 时,男性和女性语音在行阶梯矩阵及随机高斯矩阵下,不同观测点数(对应不同帧长)时压缩感知的重构效果. 重构语音质量用分段信噪比及 MOS 分来度量, MOS 分均采用 P.862 标准算出.

表 2 男性行阶梯观测矩阵及随机高斯观测矩阵下语音压缩感知重构结果

观测样点数	行阶梯观测矩阵下重构 SNR(dB)	行阶梯观测矩阵下重构 MOS	随机高斯观测矩阵下重构 SNR(dB)	随机高斯观测矩阵下重构 MOS
60	25.5465	3.439	12.6702	1.851
80	25.5842	3.454	13.3145	2.126
100	25.9069	3.485	14.6491	2.414
120	25.9659	3.412	14.0882	2.323

表 3 女性行阶梯观测矩阵及随机高斯观测矩阵下语音压缩感知重构结果

观测样点数	行阶梯观测矩阵下重构 SNR(dB)	行阶梯观测矩阵下重构 MOS	随机高斯观测矩阵下重构 SNR(dB)	随机高斯观测矩阵下重构 MOS
60	16.5801	3.054	8.4841	1.649
80	16.6634	2.970	8.7053	1.742
100	16.4758	2.975	9.1016	1.847
120	16.3108	2.830	9.5479	1.774

从表中数据可以看出,当稀疏基采用 DCT 基时,基于行阶梯观测矩阵的语音压缩感知,信号重构效果大大优于传统的基于随机高斯观测矩阵的压缩感知。

当然,由于女性语音的 DCT 系数高频成份较多,即不如男性系数集中于低频,重构效果不如男性,但行阶梯观测矩阵下的重构效果仍明显优于随机高斯观测矩阵。

4 结束语

本文针对 DCT 稀疏基下,语音信号采用随机高斯观测矩阵及线性规划方法进行压缩感知与重构时,重构零(近似零)系数定位能力差而导致重构效果不好的缺点,提出采用一种行阶梯矩阵做观测矩阵,对偶仿射尺度内点法对语音进行压缩感知重构。文中首先分析当非零 DCT 系数集中于稀疏系数向量的前段,采用对偶仿射尺度内点法重构时,行阶梯观测矩阵对重构零(近似零)系数的定位能力优于随机高斯观测矩阵的特性,然后通过仿真实验,对行阶梯观测矩阵和随机高斯观测矩阵下语音信号的压缩感知重构效果进行对比,实验结果显示,行阶梯观测矩阵下的重构性能明显优于随机高斯观测矩阵。

参考文献

[1] D Donoho. Compressed sensing[J]. IEEE Trans on Information Theory, 2006, 52(4): 1289 – 1306.

[2] E Candès. Compressive sampling[A]. Proceedings of the International Congress of Mathematicians[C]. Madrid, Spain: European Mathematical Society Publishing House, 2006. 1433 – 1452.

[3] D L Donoho, Y Tsaig. Extensions of compressed sensing[J]. Signal Processing, 2006, 86(3): 533 – 548.

[4] M Andreucut, R A Este, S A Kauffman. Competitive optimization of compressed sensing[J]. Journal of Physics A: Mathematical and Theoretical, 2007, 40(16): 299 – 305.

[5] Giacobello D, Christensen M G, Murthi M N, Jensen S H, Moonen M. 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.

[6] J F Gemmeke, B Cranen. Using sparse representations for missing data imputation in noise robust speech recognition[A]. European Signal Processing Conf, (EUSIPCO) [C]. Lausanne, Switzerland: EUSIPCO, 2008. 987 – 991.

[7] H Luo, G Pottie. Routing explicit side information for data compression in wireless sensor networks[A]. Int Conf on Distributed Computing in Sensor Systems (DCOSS) [C]. Marina Del Rey, CA: V K Prasanna Kumar, 2005. 75 – 88.

[8] D Takhar, J Lsaka, M Wakin. A new compressive imaging camera architecture using optical domain compression[A]. Proceedings of SPIE [C]. Bellingham WA: International Society for Optical Engineering, 2006: 6065.

[9] Michael Lustig, David Donoho, John M Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging[J]. Magnetic Resonance in Medicine, 2007, 58(6): 1182 – 1195.

[10] P Borgnat, P Flandrin. Time-frequency localization from sparsity constraints[A]. IEEE Int Conf on Acoustics, Speech, and Signal Processing (ICASSP) [C]. Piscataway: Institute of Electrical and Electronics Engineers Inc, 2008. 3785 – 3788.

[11] Giacobello D, Christensen M G, Murthi M N, Jensen S H, Moonen M. 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.

[12] Tao Xu, Wenwu Wang. A compressed sensing approach for underdetermined blind audio source separation with sparse representation[A]. IEEE/SP 15th Workshop on Statistical Signal Processing[C]. Cardiff, United Kingdom: IEEE, 2009. 493 – 496.

[13] Willett R M, Raginsky M. Performance bounds on compressed sensing with Poisson noise[A]. IEEE International Symposium on Information Theory [C]. Seoul: IEEE, 2009. 174 – 178.

[14] S Ji, Y Xue, L Carin. Bayesian compressive sensing[J]. IEEE Transactions Signal Processing, 2008, 56(6): 2345 – 2356.

[15] 石光明, 刘丹华, 高大化, 刘哲, 林杰, 王良君. 压缩感知理论及其研究进展[J]. 电子学报, 2009 5(37): 1070 – 1081.

SHI Guang ming, LIU Dan hua, GAO Da hua, LIU Zhe, LIN Jie, WANG Liang jun. Advance s in theory and application of compressed sensing[J]. Acta Electronica Sinica, 2009, 5(37):

1070 – 1081. (in Chinese)

- [16] S Chen, David L Donoho, M A Saunders. Atomic decomposition by basis pursuit[J]. SIAM Review, 1999, 20(1): 33 – 61.
- [17] David L Donoho. For most large underdetermined systems of linear equations, the minimal L1 solution is also the sparsest solution[EB/OL]. <http://wwwstat.stanford.edu/~donoho/Reports,2007>.
- [18] David L Donoho. For most underdetermined systems of linear equations, the minimal l_1 norm near-solution approximates the sparsest near-solution [EB/OL]. <http://www-stat.stanford.edu/donoho/Reports,2007>.
- [19] E Candès, T Tao. Decoding by linear programming[J]. IEEE, Trans Inf Theory, 2005(51): 4203 – 4215.
- [20] 金坚, 谷源涛, 梅顺良. 压缩采样技术及其应用[J]. 电子与信息学报, 2010, 2(32): 470 – 475.
Jin Jian, Gu Yuan-tao, Mei Shun-liang. An introduction to compressive sampling and its applications[J]. Journal of Electronics & Information Technology, 2010, 2(32): 470 – 475.
- [21] 陈宝林. 最优化理论与算法[M]. 第 2 版. 北京: 清华大学出版社, 2005.
CHEN Bao ling. Optimal Theories & Algorithms[M]. Edition 2. Beijing: Tsinghua University publishing house, 2005.

- [22] 何坚勇. 最优化方法[M]. 第 1 版. 北京: 清华大学出版社, 2007. 1 – 74.
HE Jian yong. Optimal Method[M]. Edition 1. Beijing: Tsinghua University publishing house, 2007. 1 – 74.

作者简介



叶 蕾 女, 1978 年生于安徽合肥. 南京邮电大学通信与信息工程学院讲师, 信号与信息处理专业博士研究生. 研究方向为语音处理与现代语音通信.

E-mail: yel@njupt.edu.cn



杨 震 男, 1961 年生于江苏苏州. 南京邮电大学信号处理与传输研究院教授, 博士生导师. 研究方向为现代网络通信、语音处理与现代语音通信.