

# 基于压缩感知的数据压缩与检测

李燕, 王博

(南京邮电大学 通信与信息工程学院, 江苏 南京 210003)

**摘要:**在无线传感器网络(WSN)中,以往都是采用奈奎斯特技术对信号进行采样并重构,而随着信号频率的增加,应用奈奎斯特技术会使成本大幅度的增加,这是人们所不乐见的。针对这一问题,近年来出现一种新的技术即压缩感知技术,它能利用更少的数据和合适的重构方法得到更精确的原始信号。将稀疏贝叶斯学习(SBL)和压缩感知联合起来,形成了一种在有噪声的情况下更好重建可压缩信号的方法,并进一步将这种方法应用在WSN中,可以在误差允许的范围内有效控制测量数据的维数,在保证一定误差的同时还减少了成本,提高了算法的效率。

**关键词:**无线传感网络;压缩感知;贝叶斯模型;信号重构

中图分类号:TN91

文献标识码:A

文章编号:1673-629X(2014)03-0198-04

doi:10.3969/j.issn.1673-629X.2014.03.049

## Data Compression and Detection Based on Compressive Sensing

LI Yan, WANG Bo

(College of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

**Abstract:**In wireless sensor networks, signal is sampled and reconstructed using the technology of Nyquist in the past. But it requires a substantial increase in the cost with the growth of the signal frequency, which is that people do not like to see. Recently a new technology is emerged, which is called compressive sensing technology. Compressive sensing can use less data and appropriate reconstruction method to get a more accurate original signal. Put Sparse Bayesian Learning (SBL) and compressive sensing together to form a better way of reconstructing compressible signal under the noise. This method can effectively control the dimension of measurement data within the range of allowed error in WSN, so you can ensure a certain degree of error while reducing the cost, improving the efficiency of the algorithm.

**Key words:**wireless sensor networks; compressive sensing; Bayesian model; signal reconstruction

## 0 引言

无线传感网络是在某个特定区域内由大量独立分布的传感器节点组成的。主要是由传感器节点感知数据,然后所感知的数据即监测数据经由其他传感器逐跳地进行传输,最终到达汇聚节点,汇聚节点把接收到的数据借助互联网传输到管理节点,用户可以通过管理器节点对传感器网络进行人工实时监控和管理。但是传感器节点一般是采用微型电池供电并且不可以频繁地进行更换,所以能量是有限的。尽可能减少能量的消耗,即降低功耗是无线感知网络研究的首要问题。应用近年来形成的压缩感知技术就是一种有效途径。压缩感知技术<sup>[1]</sup>在国内外许多领域都得到了广泛的应用并取得了很显著的效果。在信号的重构方面,相对于常用的贪婪追踪重构算法实现点的估计,文中

借助贝叶斯模型得到整体数据后验概率的分布。在有噪声条件下,BP和OMP<sup>[2]</sup>等方法都不能很好地重构原来信号,但贝叶斯模型<sup>[3]</sup>就可以。并且通过大量实验表明选择合适的反馈系数控制采集数据的维数可以在保证允许误差的前提下尽量减少测量次数,降低成本,提高算法效率。

## 1 压缩感知的总述

压缩感知理论采用非自适应线性投影保持信号的原始结构,以亚奈奎斯特频率对信号进行采样,通过解数值最优化问题准确重构出原始信号。压缩感知的一个基本观点就是:当信号在某一组已知基是稀疏的可以很大地减少需要储存数据的测量。所以感知信号必须可以用稀疏性或可压缩性形式表示。具体表述

收稿日期:2013-05-18

修回日期:2013-08-24

网络出版时间:2014-01-08

基金项目:国家自然科学基金资助项目(60972041,60972045)

作者简介:李燕(1989-),女,硕士研究生,研究方向为在无线传感网络中压缩感知的应用。

网络出版地址:<http://www.cnki.net/kcms/detail/61.1450.TP.20140108.0816.017.html>

为:假设信号  $x$  在基  $\Psi$  上是稀疏的,即

$$\mathbf{x}^{(k)} = \Psi \mathbf{s}^{(k)} \quad (1)$$

其中,  $\mathbf{x}^{(k)} \in R^N$  作为在离散的时间  $k=1,2,\dots,K$  上所感兴趣的信号;  $N$  维向量  $\mathbf{s}^{(k)} \in R^N$  是  $\mathbf{x}^{(k)}$  的稀疏表示。在 WSN 中观测数据可以表示为

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

其中,在  $k$  时刻观测向量  $\mathbf{y}^{(k)} \in R^L, M \leq L < N$  在无线传感网络中  $\Phi_{L \times N}$  被认为是路由选择矩阵,每一行最多有一个元素为 1,每一列也最多有一个元素为 1 且最多有  $L$  列非 0。最后可得

$$\mathbf{y}^{(k)} = \tilde{\Phi} \mathbf{x}^{(k)} = \tilde{\Phi} \Psi \mathbf{s}^{(k)} = \Phi \mathbf{s}^{(k)} \quad (3)$$

要想得到重构信号的稀疏表示  $\hat{\mathbf{s}}^{(k)}$ , 就要求下面公式的最优解:

$$\min_{\mathbf{s}} \|\mathbf{s}\|_q \quad \text{s. t. } \mathbf{y} = \tilde{\Phi} \Psi \mathbf{s} \quad (4)$$

这个系统是病态的,即方程的数量  $L$  小于变量的数量  $N$ 。所以对式(4)采用一些特殊的处理技术才可以求得  $\mathbf{s}^{(k)}$  的估计值  $\hat{\mathbf{s}}^{(k)}$ , 再代入(1)才可得到感兴趣的信号重构  $\hat{\mathbf{x}}^{(k)}$ 。但是这样处理数据有一个必要条件就是信号是稀疏的,并且  $\Phi \Psi$  满足 RIP 条件即要求稀疏基和测量基之间满足非相干性,也就是要求这两个基之间不能相互稀疏描述。

另外,有噪声的情况下,记为

$$\mathbf{y} = \tilde{\Phi} \mathbf{x} + \mathbf{n} \quad (5)$$

利用拉氏定理可以改写上式的重构公式为

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\operatorname{argmin}} \{ \|\mathbf{y} - \tilde{\Phi} \Psi \mathbf{s}\|_2^2 + \iota \|\mathbf{s}\|_1 \} \quad (6)$$

## 2 贝叶斯模型

应用贝叶斯模型<sup>[4-5]</sup>对信号重构过程被视为是一种线性回归的问题<sup>[6]</sup>。主要分为两个阶段:第一阶段是所有的未知数被认为是服从指定分布的随机量<sup>[7]</sup>。在有噪声的观测模型中通常假设噪声是服从独立的均值为 0,方差为  $\sigma^2$  的高斯分布,可以记为

$$p(\mathbf{n}) = \prod_{i=1}^N N(n_i | 0, \sigma^2) \quad (7)$$

超参数  $\sigma^2$  可以提前设置为常量,通常这个参数的估计值定为  $\sigma^{-2} = 0.01 \|\mathbf{y}\|_2^2$ , 也可以通过数据进行估计。设  $\beta = \sigma^{-2}$  可以进一步假设:

$$p(\beta/a, b) = \Gamma(\beta/a, b) \quad (8)$$

观测向量  $\mathbf{y}$  可以看作是多变量的高斯似然函数

$$p(\mathbf{y} | \Phi \mathbf{s}, \sigma^2) = (2\pi)^{-N/2} \sigma^{-N} \exp \left\{ -\frac{\|\mathbf{y} - \Phi \mathbf{s}\|_2^2}{2\sigma^2} \right\} \quad (9)$$

通常对信号模型<sup>[8-9]</sup>采用如下的形式

$$p(\mathbf{s} | \lambda) = \frac{\lambda}{2} \exp(-\lambda \|\mathbf{s}\|) \quad (10)$$

但是这个拉氏先验分布公式在贝叶斯分析中并不方便,因为不能和观测模型式(9)进行联合处理,所以这里采用多层先验概率进行分析。首先

$$p(\mathbf{s}/\alpha) = \prod_{i=1}^N N(s_i/0, \alpha_i) \quad (11)$$

其中超参数  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_M)$ , 超参数是指信息先验分布通常有其自身的参数,这些参数即被称作超参数。

进一步把参数  $\alpha$  看做变量并由超参数  $a, b$  决定

$$p(\alpha | a, b) = \Gamma(\alpha_i | a, b) = \Gamma(a)^{-1} b^a \alpha_i^{a-1} \exp(-b\alpha_i) \quad (12)$$

其中  $\Gamma(a) = \int_0^\infty t^{(a-1)} \exp(-t) dt$ , 为了简化模型,

初始化时可以假设这些参数是某些固定的值,例如文献[10-11]中假设  $a=b=c=d=0$ ;文中采用类似的处理手段固定  $a=1$ 。所以有

$$p(\mathbf{s}/a=1, b) = \int p(\mathbf{s}/\alpha) p(\alpha/a=1, b) d\alpha = \prod_i \int p(s_i/\alpha_i) p(\alpha_i/a=1, b) d\alpha_i = \frac{\lambda^{N/2}}{2^N} \exp(-\sqrt{\lambda} \sum_i s_i) \quad (13)$$

注意上式进行了  $\lambda = 2b$  时的改写,把概率论中著名的恒等式应用到文中可得  $\mathbf{s}$  的后验概率。

$$p(\mathbf{s}/\mathbf{y}, \lambda, \sigma^2) = p(\mathbf{y}/\mathbf{s}, \sigma^2) p(\mathbf{s}/\lambda) / p(\mathbf{y}/\lambda, \sigma^2) \quad (14)$$

把式(9)、式(13)代入,计算可知式(14)服从  $N(\mathbf{u}, \Sigma)$ 。

其中

$$\mathbf{u} = \sigma^{-2} \Sigma \Phi^T \mathbf{y} \quad \Sigma = (\Lambda + \sigma^{-2} \Phi^T \Phi)^{-1}$$

$$\Lambda = \operatorname{diag}(\alpha_i^{-1}) p(\mathbf{y}/\lambda, \sigma^2) = \int p(\mathbf{y}/\mathbf{s}, \sigma^2) p(\mathbf{s}/\lambda) d\mathbf{s}$$

通过统计分析可知上面的均值也就是要重构的信号稀疏表示形式,另外协方差  $\Sigma$  的对角线上的元素也就是重构过程中产生的误差。在无线传感网络中允许的误差最大值为  $\eta$ , 重构过程中产生的误差均值为  $\Delta$  通过大量的实验可得反馈系数  $p_{re}$

$$p_{re} = \begin{cases} \min \left\{ \frac{\Delta - \eta}{\eta} c_1, 1 \right\}, & \text{if } 1.2\Delta > \eta \\ \max \left\{ -\frac{\eta - \frac{3}{5}\Delta}{\eta} c_2, -\left(1 - \frac{K}{L} \log\left(\frac{N}{K}\right)\right) \right\}, & \text{if } 1.2\Delta < \eta \end{cases} \quad (15)$$

上面也就是第2阶段的目的,得到最匹配的超参数从而进一步获得权值参数的最优估计即待重建的信号<sup>[9]</sup>,采用的方法是最大化超参数的边缘对数似然函

数。对于边缘似然函数  $p(\lambda, \beta, \alpha/y)$  的对数形式的最大化,这里采用第二类型最大似然参数估计的方法<sup>[10-11]</sup>。边缘似然函数的对数形式通常表示为

由于  $p(\lambda, \beta, \alpha/y) = \frac{p(\lambda, \beta, \alpha, y)}{p(y)} \propto p(\lambda, \beta, \alpha, y)$

所以有

$$L = \lg(\int p(\lambda, \beta, \alpha, y, s) ds) =$$
$$\lg(\int p(y/s, \beta) p(s/\alpha) p(\alpha/a = 1, b) p(b) ds) =$$
$$-\frac{1}{2} \lg |C| - \frac{1}{2} y^T C^{-1} y + (N-1) \lg \lambda - \frac{\lambda}{2} \sum_i \alpha_i$$

(16)

其中  $C = \sigma^2 I + \Phi A^{-1} \Phi^T$ 。下面对上式进行处理。

$$L(\lambda) = -\frac{1}{2} [\lg |C_{-i}| + y^T C_{-i}^{-1} y + \frac{\lambda}{2} \sum_{j \neq i} \lambda_j] +$$
$$\frac{1}{2} [\log \frac{1}{1 + \lambda_i p_i} + \frac{q_i^2 \lambda_i}{1 + \lambda_i p_i} - \lambda_i b] =$$
$$L(\lambda_{-i}) + l(\lambda_i)$$

估计参数  $\lambda_i$ , 最大化  $L(\lambda)$  因为把  $\lambda_m (m \neq i)$  都视为常数,所以相当于最大化  $l(\lambda_i)$ 。

计算可知:

$$\alpha_i = \begin{cases} -\frac{p_i^2 + 2\lambda p_i + p_i \sqrt{p_i^2 + 4\lambda q_i^2}}{2\lambda p_i}, & q_i^2 - p_i > 0 \\ 0 & \text{其他} \end{cases}$$

(17)

另外设

$$P_m = \sigma^{-2} \varphi_m^T \varphi_m - \sigma^{-4} \varphi_m^T \varphi \Sigma \varphi^T \varphi_m, Q_m = \sigma^{-2} \varphi_m^T y -$$
$$\sigma^{-4} \varphi_m^T \varphi \Sigma \varphi^T y$$

(18)

更新公式

$$p_m = \frac{\lambda_m P_m}{\lambda_m - P_m}, q_m = \frac{\lambda_m Q_m}{\lambda_m - P_m}$$

(19)

3 实验仿真

在第一个实验中考虑的是稀疏  $M$  为 20 且长度  $N$  为 512 的稀疏信号,随机选择 20 个位置并随机赋值正 1 或负 1<sup>[12]</sup>,观测矩阵为随机矩阵并进行归一化处理。仿真结果如图 1 所示(其中横坐标代表信号的长度,纵坐标代表信号的幅值)。

在反馈系数更新公式中  $c_1 = 1.8, c_2 = 3$ 。下面的实验结果是在进行的 1 000 次测量进行重构过程中最后一次测量的实验结果。

在第 2 实验中,数据集的选取和实验一中的相同,主要是对观测向量  $y$  的维数对整体重构误差的影响进行实验。

实验结果如图 2。

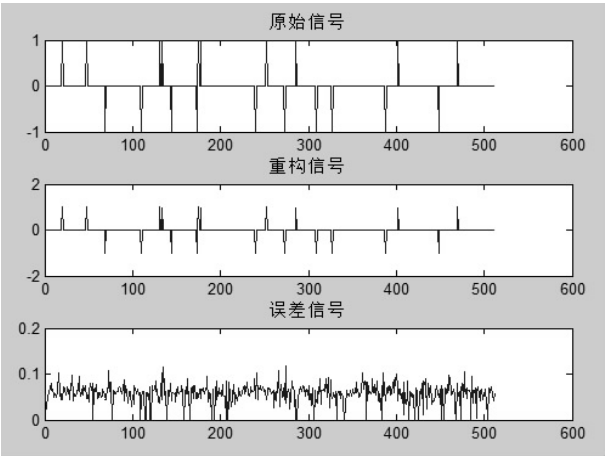


图 1 仿真结果

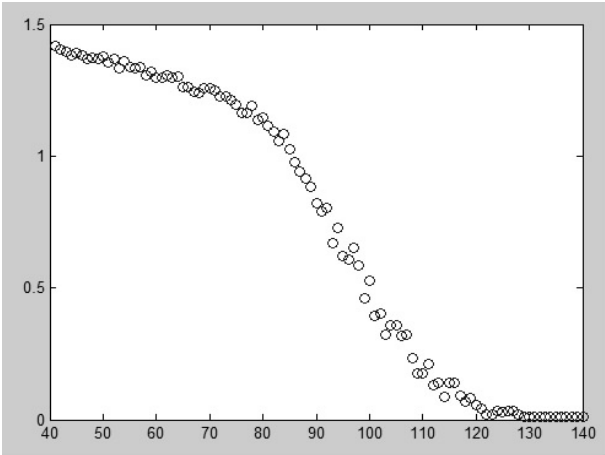


图 2 实验 2 结果

在自适应系统中如果误差太大可以通过增加测量向量的维数来减少误差,这是贝叶斯模型一个明显的优势。

在第 3 个实验中对 OMP、StOMP 和文中的重构算法,如图 3 所示(“\*”代表 OMP 算法的重构,“o”代表文中算法的重构,“x”代表改进的 OMP 算法即 StOMP 的重构,其中横坐标代表信号的长度,纵坐标代表重构信号的幅值)。

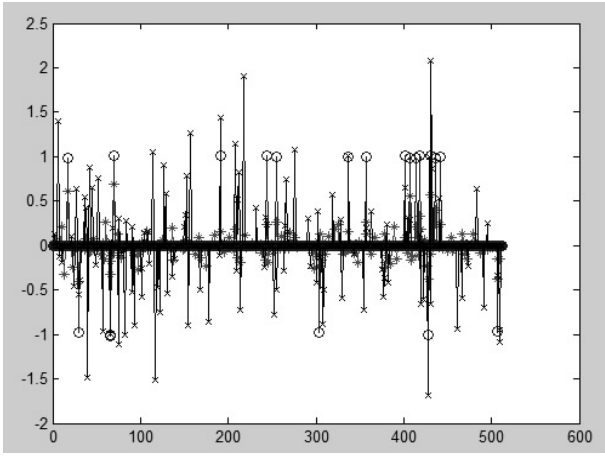


图 3 实验 3 结果

4 结束语

在无线传感网络中采用微型电池供电所以能量是有限的。这一限制在很大程度上缩短了传感器的使用寿命进而使得无线传感网络的稳定性受到了很大的影响。在实际的环境中,噪声是无处不在的,贝叶斯模型就是一种典型解决有噪情况下的框架模型。这就是文中把压缩感知技术和贝叶斯模型结合起来的最原始的出发点。通过大量的仿真求证反馈系数是为了自动地控制网络系统,以满足在允许误差范围内利用最少的测量数据获取精确的重构信号,降低数据的采集传输成本,延长传感器的使用寿命,保证无线传感网络的稳定性,为未来无线传感网络性能的优化提供一种有效的方法。

参考文献:

[1] Donoho D. Compressed sensing[J]. IEEE trans on information theory,2006,52(4):4036-4048.

[2] Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit[J]. IEEE trans on information theory,2007,53(12):4655-4666.

[3] 邓乃扬,田英杰. 数据挖掘中的新方法-支持向量机[M]. 北京:科学出版社,2004.

(上接第197页)

5 结束语

文中将LLE算法和OIF-Elman神经网络结合起来,提出了一种新的滤波电路故障诊断模型,并对该模型的具体实现给出了详细的说明与论证,最后利用所提出的故障诊断模型对二阶滤波电路进行分类。对仿真结果进行标记与分析,表明了文中所提出的故障诊断模型具有理想的电路故障诊断准确率,其收敛速度和稳定性对于传统的BP神经网络具有明显的优势。

参考文献:

[1] 刘新东. 基于LLE和SVM的模拟电路软故障诊断[C]//系统仿真技术及其应用学术会议论文集. 出版地不详:出版者不详,2010.

[2] 李云红,孟凡斌,谭阳红,等. 基于量子神经网络的容差模拟电路的软故障诊断[J]. 计算机测量与控制,2009,17(8):1463-1467.

[3] 孙永奎,陈光[祺],李辉. 模糊聚类与SVM诊断模拟电路单软故障的方法[J]. 计算机辅助设计与图形学学报,2008,20(5):612-617.

[4] 刘安,刘春生. 基于RBF神经网络的非线性系统故障诊断[J]. 计算机仿真,2007(2):141-144.

[4] Vapnik V N. 统计学习理论[M]. 许建华,张学工,译. 北京:电子工业出版社,2004.

[5] Vapink V N. 统计学习理论的本质[M]. 张学工,译. 北京:清华大学出版社,2000.

[6] 薛建儒,郑南宁,郑朝晖,等. 基于自适应高斯混合体模型的相控阵雷达TWS跟踪技术[J]. 电子学报,2003,31(3):433-436.

[7] Babacan S D, Molina R, Katsaggelos A K. Bayesian compressive sensing using Laplace priors[J]. IEEE trans on image processing,2010,19(1):53-63.

[8] Tipping M. Sparse Bayesian learning and the relevance vector machine[J]. Journal of machine learning research,2001,1(1):211-244.

[9] Figueiredo M, Nowak R, Wrih S J. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems[J]. IEEE transaction on selected topics in signal processing,2007,1(4):586-597.

[10] 王平波,蔡志明,刘旺锁. 混合高斯概率密度模型参数的期望最大化估计[J]. 声学技术,2007,26(3):498-502.

[11] 岳佳,王士同. 高斯混合模型聚类中EM算法及初始化的研究[J]. 微计算机信息,2006,22(11-3):244-246.

[12] Seeger M W, Nickisch H. Compressed sensing and Bayesian experimental design[C]//Proc of ICML. [s.l.]:[s.n.],2008:912-919.

[5] 张义忠,冯振声. 基于神经网络的复杂电子装备故障诊断系统的仿真研究[J]. 计算机仿真,2000,17(4):39-42.

[6] 祝世虎,董朝阳,张金鹏,等. 基于神经网络与专家系统的智能决策支持系统[J]. 光电与控制,2006,13(1):8-11.

[7] 吴进华,沈剑,段育红,等. 数模混合电路故障诊断的方法研究[J]. 海军航空工程学院学报,2008,23(3):297-301.

[8] 祝文姬,何怡刚. 容差模拟电路软故障诊断的神经网络方法[J]. 电工技术学报,2009,24(11):184-191.

[9] 高洪深,陶有德. BP神经网络模型的改进[J]. 系统工程理论与实践,1996(1):67-71.

[10] 陈龙,张可,罗配明. 改进的Elman神经网络在WSNs距离预测中的应用[J]. 传感器与微系统,2013,32(1):149-152.

[11] 郑杰. 一种改进的LLE方法[J]. 湖南理工学院学报(自然科学版),2007,20(3):30-32.

[12] Vin de Silva, Joshua B. Global versus local methods in nonlinear dimensionality reduction[C]//Proc of NIPS. [s.l.]:[s.n.],2002:705-712.

[13] Shi X H, Liang Y C, Lee H P, et al. Improved Elman networks and applications for controlling ultrasonic motors[J]. Applied artificial intelligence,2004,18:603-609.

基于压缩感知的数据压缩与检测

作者：[李燕](#)，[王博](#)，[LI Yan](#)，[WANG Bo](#)

作者单位：[南京邮电大学 通信与信息工程学院, 江苏 南京, 210003](#)

刊名：[计算机技术与发展](#)



英文刊名：[Computer Technology and Development](#)

年，卷(期)：2014(3)

本文链接：[http://d.g.wanfangdata.com.cn/Periodical\\_wjfz201403049.aspx](http://d.g.wanfangdata.com.cn/Periodical_wjfz201403049.aspx)