基于贝叶斯压缩感知的信号重构①

陆海东 [1]，顾美康 [1]，申晓磊 [2]

[1](上海师范大学 信息与机电工程学院，上海 200234)
[2](东华大学 信息科学与技术学院，上海 200050)

摘 要: 本文提出了基于贝叶斯压缩感知的信号重构算法，将压缩感知理论应用于信号的压缩传输以及重构，该算法将压缩感知问题转化为线性回归问题，逐步推演出结果向量之间的迭代关系，最后通过迭代以得到原始信号的精确重构. 仿真说明了贝叶斯压缩感知在信号处理中的应用，结果表明该算法对一维和二维信号的压缩重构有很好的效果.

关键词: 贝叶斯; 压缩感知; BCS 算法; 信号重构

Signal Reconstruction Based on Bayesian Compression Sensing

LU Hai-Dong[1], GU Mei-Kang[1], SHEN Xiao-Lei[2]

[1](College of Information, Mechanical&Electrical Engineer, Shanghai Normal University, Shanghai200234, China)
[2](School of Information Science& Technology, Donghua University, Shanghai200050, China)

Abstract: In this paper, a compressed sensing signal reconstruction algorithm that based on Bayesian compression perception theory is proposed. It can be applied to signal compression and transmission as well as reconstruction. The new algorithm inverted the compressed sensing problem into a linear regression problem. Firstly, then deduced an iterative relationship of the resulting vectors gradually, at last got the exact reconstruction of the original signal by iteration. The simulation experiment exploted that the Bayesian compressive sensing algorithm have a good reconstruction effect used in one-dimensional and two-dimensional signal processing and the reconstruction.

Key words: Bayes; compressed sensing; BCS algorithm; signal reconstruction

传统奈奎斯特采样定理提出：采样率至少是信号最高频率的两倍，才能保证信号避免失真，这样便产生问题：采样与传输时会面临巨大的数据量，直接导致处理效率下降，同时采样设备所面临的压力也急剧增加. 而压缩感知(CS)理论表明：如果信号具有稀疏性，或可压缩的，可以找到与变换基不相关的观测矩阵将高维原始信号投影到低维空间，通过最优化方法进行原始信号的精确重构.

目前 CS 理论得到不断的发展，主要的问题涉及到恢复重构算法的研究，一些学者提出了一些贪婪算法，如 OMP,STOMP 算法等. 但是这些得法重构的精度比较低，而且所耗时间多. 而另一部分学者提出了基于贝叶斯的压缩感知方法，如：拉普拉斯先验算法，但是它却受到一些噪声的干扰.

基于上述算法存在的问题，本文提出基于贝叶斯压缩感知的新型信号处理方法，将 CS 问题转换为贝叶斯线性回归问题，将观测值分解为有用信号与噪声信号，最后经过一系列推演，找到变量之间的迭代关系，进行相互迭代，无限趋近于原始信号，最后实现信号的精确重构. 仿真试验表明，该方法能够高概率重构原始信号.

1 CS模型

根据压缩感知理论，$N\times1$ 维的未知信号 f 在线性基 ψ [1] (例如 curvelet)下是可压缩的. 则信号 f 可以表示为 $f=\psi h$，其中 h 是 $N\times1$ 的稀疏信号，大部分

① 基金项目:上海师范大学创新性和前瞻性项目(DYL.201007)；国家自然科学基金(60971004)
收稿时间:2013-09-11;收到修改稿时间:2013-10-16

系数为零. 对原始信号观测, 得到观测值如下式所示:

$$G=\phi' f+n \tag{1}$$

(1)式中为原始未知信号 f 的 $M\times1$ 维观测值, 并具有 $M\times N$ 维观测矩阵 $\phi'=[\phi'_1,\ \phi'_2,...\phi'_n]$, 其中 n 是噪声向量, 由此可以得出下式

$$G=\phi h+n \tag{2}$$

(2)式中 $\phi=\phi'\psi$,由压缩感知理论可以得出: 当观测数目远小于信号的维数时($M<<N$), 可以使用适当的算法进行原始信号 f 的精确重构[2-3]. 原始信号的近似值求解就可以转化以下最优化问题.

$$\hat{h}=\arg\min\{\|g-\phi h\|_2^2+\tau\|h\|_0\} \tag{3}$$

这个最优化问题是 NP 难问题. 因此通常将其转化为范数问题, 如下所示:

$$\hat{h}=\arg\min\{\|g-\phi h\|_2^2+\tau\|h\|_1\} \tag{4}$$

其中 $\|\bullet\|_1$ 代表 l_1 范数.

2 基于贝叶斯的压缩感知(BCS)

对(2)式进行仔细分析, 并设置一个新的向量 h_s , h_s 由向量 h 中 M 个非零元素组成, 其余的 $N-M$ 个元素值近似为 0. 同理将 h 中剩余的 $N-M$ 个元素看做另一个向量 h_e . 所有 h_e 中剩余的元素设置为 0. 因此有 $h=h_s+h_e$, (2)式可以表示如下:

$$g=\phi h=\phi h_s+\phi h_e=\phi h_s+n_e \tag{5}$$

其中 $n_e=\varphi h_e$, ϕ 是通过随机采样形成的观测矩阵, n_e 由零均值的高斯矩阵组成, 以满足中心极限定理. 用 n_m 表示 CS 观测矩阵所包含的噪声, 满足零均值的高斯分布, 因此可得

$$g=\phi h_s+n_e+n_m=\phi h_s+n \tag{6}$$

其中 n 满足均值为零, 未知方差为 σ^2 的高斯噪声, 因此便可得出高斯模型

$$p(g\mid h_s,\sigma^2)=(2\pi\sigma^2)^{-K/2}\exp(-\frac{1}{2\sigma^2}\|g-\phi h_s\|^2) \tag{7}$$

上述分析将 CS 问题中的求解 h_s 转化为: 具有稀疏性 h_s 作为先验条件的线性回归问题. 假设已知观测矩阵 ϕ , 以及 CS 观测值 g, 稀疏权值 h_s 以及噪声方差 σ^2 . 在贝叶斯分析中, 就可以找到基于 h_s 和 σ^2 的后验密度函数.

2.1 观测模型

通过以上分析可知, 未知信号 h 的先验分布为 $p(g\mid h,r)$,观测值的条件分布为 $p(g\mid h,t)$, 而 t 是高斯白噪声方差的倒数. 此时的 r 与 t , 通常称作超参数.

观测噪声是独立的, 高斯噪声具有零均值, 方差为 σ^2 , 结合(2)式可得:

$$p=(g\mid h,\sigma^2)=N(y\mid\phi h,\sigma^2) \tag{8}$$

因为 $t=1/\sigma^2$, σ^2 的 Gamma 先验如下: 可以得出

$$p=(t\mid a^t,\ b^t)=\Gamma(t\mid a^t,b^t) \tag{9}$$

Gamma 分布可以定义为

$$\Gamma(s\mid a^t,b^t)=\frac{(b^s)^{a^s}}{\Gamma(a^s)}(s^{a^s-1}-1)\exp[-b^s s] \tag{10}$$

其中, $s>0$ 是一超参数, $b^s>0$,作为尺度参数, 同时 a^s 0,为形状参数, s 的均值 μ_s 和方差 σ^2 表示如下:

$$\mu_s=a^s/b^s \tag{11}$$

$$\sigma^2=a^s/(b^s)^2 \tag{12}$$

以上过程将用 Gamma 分布代替表示高斯分布的方差倒数的先验表示, 一定程度上简化了分析过程.

2.2 信号模型

通过贝叶斯方程可以得知: w_s 是通过稀疏提升正则化以达到稀疏状态. 拉普拉斯密度函数如[4]下:

$$p(h\mid\lambda)=(\frac{\lambda}{2})^N\exp(-\lambda\sum_{i=1}^{N}|h_i|) \tag{13}$$

而(4)式中的 l_1 正则化规划, 等同于使用基于 h 的拉普拉斯先验, 如下式所示:

$$p(h\mid\lambda)=(\frac{\lambda}{2})\exp(-\frac{\lambda}{2}\|h\|_1) \tag{14}$$

给定 CS 观测值 g , 假设存在(7)式中的函数, 我们可以将(4)式中最优化问题变换为寻求 h 的最大后验估计.

2.3 层次稀疏先验概率

上述分析将传统的 CS 求权值的 h 反演与 MAP 估计合成为贝叶斯线性回归分析, 具有拉普拉斯先验稀疏的 h . 再次可以定义 h 的每个元素具有一个零均值的高斯先验. 如下所示

$$p(h|\alpha)=\prod_{i=1}^{N}N(h_i|0,\alpha_i^{-1}) \tag{15}$$

$\alpha_i^{-1}=(\alpha_1^{-1},\alpha_2^{-1},...\alpha_3^{-1})$ 是高斯密度函数的逆方差. 同时由文献[5]层次分析第二阶段可知, $p(\alpha_i^{-1})\alpha\alpha_i$ 再者 gamma 超先验如下

$$p(\alpha|a,b)=\prod_{i=1}^{N}\Gamma(\alpha_i|a,b) \tag{16}$$

通过边缘化超参数, α,w 所有先验可以估计为

$$p(h|a,b)=\prod_{i=1}^{N}\int_0^\infty N(h_i|0,\alpha_i^{-1})\Gamma(\alpha_i|a,b)d\alpha_i \tag{17}$$

密度函数 $\Gamma(\alpha_i|a,b)$ 与 α_i 是共轭先验, 当 h_i 为观测的函数, $N(h_i|0,\alpha_i^{-1})$ 作为可能性函数. 就可以对积分 $\int_0^\infty N(h_i|0,\alpha_i^{-1})\Gamma(\alpha_i|a,b)d\alpha_i$ 进行预估分析, 以对应于学生分布[6](学生分布又称 t 分布, 主要应用在估计程正态分布的母群体之平均数). 只要选择合适的 a 和 b, 学生分布能够得出 $h_i=0$, 因此, (17)式中的大部分 h_i 可以置为 0(如果是稀疏先验的), 类似 gamma 先验 $\Gamma(\alpha_0|a,b)$ 可以通过噪声方差 $\alpha_0=1/\sigma^2$ 的逆得以实现.

假设超参 α 与 α_0,以及投影矩阵 ϕ,h 的后验可以表达为一个多元变量的高斯分布具有均值和方差

$$\mu=\alpha^0\sum\phi^T G \tag{18}$$

$$\sum=(\alpha^0\phi^T\phi+A)^{-1} \tag{19}$$

其中 A=diag($\alpha_1\alpha_{2,...}\alpha_N$).在这里问题就转化为寻求超参数 α 与 α_0, 在 RVM 中这些超参数估计来源于(type-II)ML(或者证据最大化)过程[5], 尤其是通过边缘化权值 h, 边缘估计 α 与 α_0. 该算法 ℓ (α, α_0) 能够表达为

$$\begin{aligned}\ell(\alpha,\alpha_0)&=\log p(g|\alpha,\alpha_0)\\&=\log\int p(g|h,\alpha_0)p(h|a)dh\\&=-\frac{1}{2}[K\log 2\pi+\log|C|+g^T C^{-1}g]\end{aligned} \tag{20}$$

其中 C= $\sigma^2 I+\phi A^{-1}\phi^{-T}$, 一个(type-II)ML[5]表明对于 α 与 α_0 的点估计以最大化, 可以通过 EM 算法实现(或其他算法)形成

$$\alpha_i^{new}=\frac{r_i}{\mu_i^2} \tag{21}$$

其中 μ_i 是(18)式中的第 i 个后验均值, 我们已经定义数目 $\gamma_i\overset{\Delta}{=}1-\alpha_i\sum ii$, $\sum ii$ 是第 i 个后验权值协方差的对角线矩阵元素. 对于噪声方差 $\sigma^2=1/\alpha_0$,微分之后进行预估计.

我们可以注意到 α^{new} 与 α_0^{new} 是 μ 与 $\sum$ 的函数, μ 与 $\sum$ 同时也是 α 与 α_0 的函数. 因此我们可以在(18),(19),(20),(21)之间进行迭代, 直到满足收敛性准则.

3 新的算法

上述第三部分的分析有一定的道理, 然而, 其中 a_i 是趋向于无穷大的, 在 $g=\phi h$ 中, h_i 并不能完全表现出稀疏性, 只对 h_i 小部分进行设置, 所以对应的 a_i 也只是一小部分, 其观测值与稀疏度的表现完全取决于 h_i, 同时我们可以在 Gamma 超先验设置四个超参数 a,b,c,d. 这等同于将 a,b,c,d 设置为 0, 这样 α 与 α_0, 可以被调用了.

当对权值 h 进行不确定的观测, 最需关注的是信号 f 的维数, 其中 $f=Bh$, 因为 h 是来源于多维高斯分布, 具有均值和方差, 在(18)(19)两式中已经定义, f 的后验密度函数具有均值与协方差如下式所示

$$\mathrm{E}(f)=B\mu \tag{22}$$

$$Cov(f)=B\sum B^T \tag{23}$$

(23)式中的协方差矩阵的对角元素提供了误差精度, 是基于(22)式中的方差, 以表明重构的精确度, 考虑 ℓ (α) 与 α_i 的相关性, $i\in\{1...N\}$, 我们可以将(20)式中的 C 分解为[6-7]

$$\mathrm{C}=\sigma^2\mathrm{I}+\sum\nolimits_{n\neq 1}\alpha_n^{-1}\phi_n\phi_n^T+\alpha_i^{-1}\phi_i\phi_i^T=C_{-i}+\alpha_i^{-1}\phi_i\phi_i^T \tag{24}$$

其中 C_{-i} 是减去基向量 i 后所得结果, 我们可以将 ℓ (α) 写成如下式子

$$\begin{aligned}|\mathrm{C}|&=|\mathrm{C}_{\text{-i}}||1+\alpha_{\mathrm{i}}^{1}\phi_i^T C_{-i}^{-1}\phi_i|\\C^{-1}&=C_{-i}^{-1}-\frac{C_{-i}^{-1}\phi_i C_{\text{-}i}^{-1}}{\alpha_i+\phi_i^T C_i^{-1}\phi_i}\end{aligned} \tag{25}$$

因此我们可以改写 ℓ (α) 如下所示

$$\begin{aligned}\ell(\alpha)&=-\frac{1}{2}[\mathrm{N}\log(2\pi)+\log|\mathrm{C}_{-i}|]+\mathrm{t}^T C_{-i}t\\&-\log\alpha_{-i}+\log(\alpha_i+\phi_i^T\mathrm{C}_{-i}^{-1}\phi_i)-\frac{(\phi_i^T\mathrm{C}_{-i}^{-1}\mathrm{t})^2}{\alpha_i+\phi_i^T C_{-i}^{-1}\phi_i}\\&=\ell(\alpha_{-i})+\frac{1}{2}[\log\alpha_i-\log(\alpha_i+\mu_i)+\frac{v_i^2}{\alpha_i+\mu_i}]\\&=\ell(\alpha_{-i})+\rho(\alpha_i)\end{aligned} \tag{26}$$

为了简化表达式我们定义:

$$\mu_i \overset{\Delta}{=} \phi_i^T C_{-i}^{-1} \phi_i \qquad (27)$$
$$v_i \overset{\Delta}{=} \phi_i^T C_{-i}^{-1} t$$

稀疏因子 u_i 可以看做基向量 ϕ 重叠部分的观测. 质量因子可以写作 $v_i = \sigma^{-2}\phi_i^T(t - g_{-i})$ 可以是具有误差模型的 ϕ 属性的观测. 目标函数被分解成 ℓ（α_{-i}）$,\phi$ 的边缘可能性被排除在外, 此时 $\rho(\alpha_i)$ 中, α_i 被孤立在外, 分析 $\rho(\alpha_i)$ 表明 ℓ（α）对应于 α_i 有一个唯一的最大值

$$\alpha_i = \frac{\mu_i^2}{v_i^2 - \mu_i}\text{，如果 } v_i^2 > \mu_i \qquad (28)$$

$$\alpha_i = \infty\text{，　如果 } v_i^2 < \mu_i \qquad (29)$$

用所有的基函数 ϕ_i 接计算 u_i 与 v_i，这样便可以很容易维持和更新下式的值

$$U_n = \phi_n^T C^{-1} \phi_n \qquad V_n = \phi_n^T C^{-1} t \qquad (30)$$

因此可以简化为如下

$$u_n = \frac{a_n u_n}{a_n - u_n} \qquad v_n = \frac{a_n v_n}{a_n - u_n} \qquad (31)$$

当且仅当上式满足条件 $a_n = \infty$，$u_n = U_n$，同时 $v_n = V_n$ 实际上易利用 woodbury[6-7]获得所需的数据量

$$\mathrm{U}_n = \phi_n^T X \phi_n - \phi_n^T X \phi \sum \phi^T X \phi_n \qquad (32)$$

$$V_n = \phi_n^T X \hat{t} - \phi_n^T X \phi \sum \phi^T X \hat{t} \qquad (33)$$

其中 $X = \sigma^{-2} I$，这里 $\hat{t} \equiv t$，因此在此可以重新估计 a_i，将式(20)应用于模型中的基向量，我们便可以实现一个相对算法

本文提出新的算法如下:

①对需处理的原始信号进行小波变换，增强其稀疏性，在这里选择 db1 作为小波基，通过下文的仿真实验表明在该小波基情况下，稀疏性比较好，同时原始信号的重构速度比较快,

②上一步得到小波分解向量即小波系数矩阵，选择观测矩阵，这里选择高斯白噪声，与观测矩阵不相关(满足受限等距原则(RIP))，得到观测值

③随后利用新型贝叶斯后验方法进行信号的重构，具体流程如下:

说明:　d：迭代次数($1 \le d \le 100$)

m：中间值

$curr$：当前重构值

$last$：上次重构值

acc：最小误差精度（$acc = 1e = 4$）

输入: 投影矩阵 ψ 和观测矩阵 ϕ

输出: 重构信号 Z

初始化: 根据与 ψ 与 ϕ 的值分别初始化 u_n 与 v, n, u, Σ.

设超参数 $a = a_0, b = b_0, m = 0$，

$curr = 3 * acc + 1, last = 1$;

将上述初始值代入进行计算,

$last = curr$;

$\mathrm{curr} = \rho(\alpha)$

如果(1)　d 达到最大迭代次数 100;

或者(2)误差精度 $\dfrac{|curr - last|}{last} < acc$

两个条件中满足一个，则停止迭代得出重构结果 Z. 否则,根据(31)(32)(33)以及(18)(19)(20)(21)继续循环计算 u_n, v_n, μ, Σ 直到满足以上二条件之一.

4　仿真实验及结果

这部分主要通过 Matlab 对一维信号和二维信号进行仿真，以验证本文所提出的新算法的性能表现，具体仿真分析如下.

4.1 一维信号的处理及结果

本文仿真中，原始信号设置为 N =512，选择间隔为 50 的位置上分布大小为±2 的随机信号，投影矩阵 ϕ 是 $M \times N$、观测矩阵采用均值为 0，标准差为 σ^2 的高斯白噪声

4.1.1 BCS 与新方法的对比

如图 1 所示为经 BCS 方法处理的信号重构，图 2 是通过新的方法进行处理后的结果，对比图 1 与图 2 可知，图 1 的相对误差为 0.16179，重构的时间为 2.2834s，图 2 第一张图的相对误差为 0.0030638，重构时间为 1.61s.在相同的前提下，图 2 无论是在信号重构所花的时间还是在重构相对误差精度上，效果都明显优越于 BCS 算法，与 BCS 方法进行处理对比发现效率得到很大提高，误差系数明显小很多.

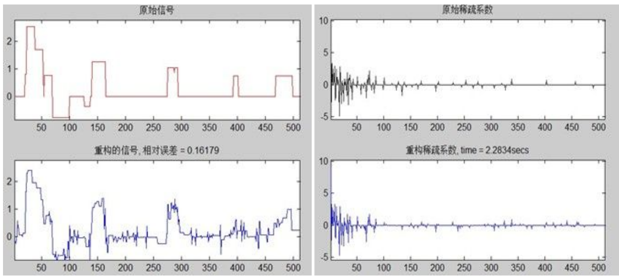

图 1 BCS 处理一维信号

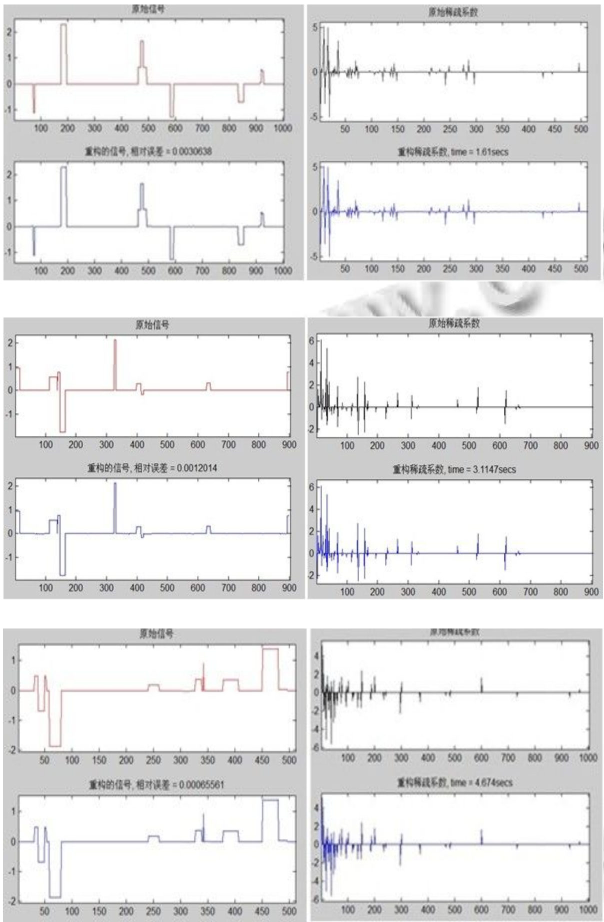

图 2 不同观测数目，新方法处理一维信号

4.1.2 新方法中不同观测量之间的对比

图 2 是在不断增加观测数目的前提下，对算法的效果进行仿真，通过对比发现，随着观测数目的增加，重构所花费的时间越来越长，但是误差系数显著降低，说明增加观测数目可以提高信号的重构精度.

4.2 对于二维信号的处理及结果

4.2.1 BCS 与新方法的对比

与一维信号类似，这里选择一组车牌图片，分别用 BCS 算法与新方法进行二维信号的重构处理，结果如图 3 与 4 所示，图 3 相对误差为 0.18087，重构时间为 50.2313s，图 4 第一张图相对误差为 0.13201，重构时间为 30.4627.图 4 明显比图 3，在时间复杂度和重构误差精度上有质的提高.

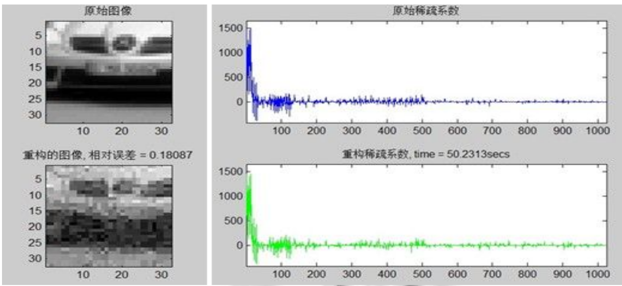

图 3 BCS 处理二维信号

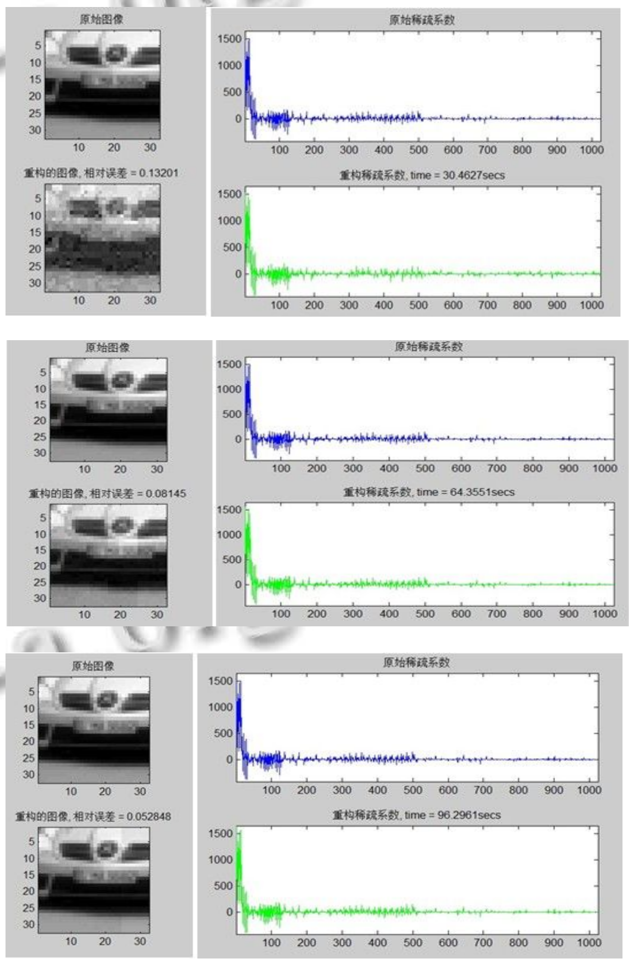

图 4 不同观测数目，新方法处理二维信号

4.2.2 新方法中不同观测量之间的对比

图 4 是在不同观测数目下对新的算法的验证，说明随着观测数目的增加，算法的时间效率逐步降低，重构效果逐渐明显.

4.3 不同二维图像数据量之间的对比

采用 lena 图像进行数据处理，对比车牌图像我们发现，如下图 5 发现随着二维图像数据的增加，重构的时间复杂度逐渐增加，但是二维信号的重构精度没有多大差别，这说明该算法还是可行的.

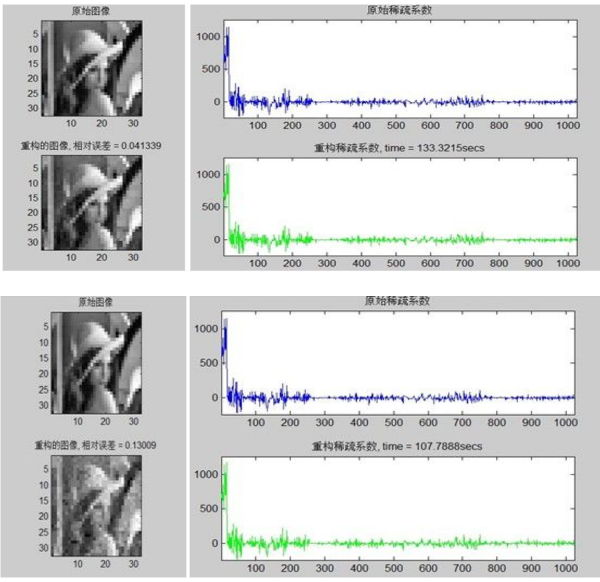

图 5 二维图像的处理

5 结论

本文对压缩感知理论进行简要叙述，提出 CS 中需要解决的信号重构问题，由此引出基于压缩感知理论的的贝叶斯方法，依据该方法提出了新的算法，将该算法应用于一维随机信号和二维图像信号的采样压缩与重构中，仿真实验得到很好的效果，原始信号与重构信号的误差度很小，尤其是二维图像信号，一定程度上能够满足一般图像处理研究所需的误差精度. 有待提高的是，当二维图像信号数据量比较大的时候，时间复杂度会逐渐增大，这是亟待改进的地方.

参考文献

1 沈明欣,刘文波.基于压缩感知的图像重构技术.电子科技, 2011,24(3):9–12.

2 Candes EJ, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory, 2006, 52(2): 489–509.

3 Donoho DL. Compressed sensing. IEEE Trans. Inf. Theory, 2006, 52(4):1289–1306.

4 Babacan SD, Molina R, Katsaggelos AK. Bayesian compressive sensing using Laplace priors. IEEE Trans. on Image Processing, 2010, 19(1): 53-63.

5 Tipping ME. Sparse Bayesian learning and the relevance vector machine. The Journal of Machine Learning Research, 2001, 1: 211-244.

6 Faul AC, Tipping ME. Analysis of sparse Bayesian learning. In: Dietterich TG, Becker S, Ghahramani Z, eds. Advances in Neural Information Processing Systems, 2002: 383–389.

7 Tipping ME, Faul AC. Fast marginal likelihood maximization for sparse Bayesian models. In: Bishop CM, Frey BJ, eds. Proc. 9th Int.Workshop Artificial Intelligence and Statistics. 2003.