曹静杰，王彦飞，杨长春. 地震数据压缩重构的正则化与零范数稀疏最优化方法. 地球物理学报，2012，55(2)：596-607，doi：10.6038/j.issn.0001-5733.2012.02.022.

Cao J J, Wang Y F, Yang C C. Seismic data restoration based on compressive sensing using the regularization and zero-norm sparse optimization. *Chinese J. Geophys.* (in Chinese), 2012, 55(2): 596-607, doi: 10.6038/j.issn.0001-5733.2012.02.022.

地震数据压缩重构的正则化与零范数稀疏最优化方法

曹静杰[1,2,3]，王彦飞[2*]，杨长春[2]

1 石家庄经济学院，石家庄 050031

2 油气资源研究重点实验室，中国科学院地质与地球物理研究所，北京 100029

3 中国科学院研究生院，北京 100049

摘 要 地震数据重构问题是一个病态的反演问题. 本文基于地震数据在 curvelet 域的稀疏性，将地震数据重构变为一个稀疏优化问题，构造 0 范数的逼近函数作为目标函数，提出了一种投影梯度求解算法. 本文还运用最近提出的分段随机采样方式进行采样，该采样方式能够有效地控制采样间隔并且保持采样的随机性. 地震数值模拟表明，基于 0 范数逼近的投影梯度法计算效率有明显的提高；分段随机采样方式比随机欠采样有更加稳定的重构结果.

关键词 波场重构，curvelet 变换，压缩传感，0 范数逼近，反问题，不适定性，稀疏优化

doi:10.6038/j.issn.0001-5733.2012.02.022 **中图分类号** P631 **收稿日期** 2011-06-15，2011-10-14 收修定稿

Seismic data restoration based on compressive sensing using the regularization and zero-norm sparse optimization

CAO Jing-Jie[1,2,3], WANG Yan-Fei[2*], YANG Chang-Chun[2]

1 *Shijiazhuang University of Economics*, *Shijiazhuang* 050031, *China*

2 *Key Laboratory of Petroleum Resources Research*, *Institute of Geology and Geophysics*, *Chinese Academy of Sciences*, *Beijing* 100029, *China*

3 *Graduate University*, *Chinese Academy of Sciences*, *Beijing* 100049, *China*

Abstract Seismic data restoration is an ill-posed inverse problem. Based on the sparseness of seismic data in the curvelet domain, this problem can be transformed into a sparse optimization problem. This paper proposes to use the approximation of zero-norm as the objective function and develop a projected gradient method to solve the corresponding minimization problem. We also employ a recently proposed piecewise random sampling method which can both control the sampling gap and keep the randomness of sampling. Numerical results show that the projected

基金项目 国家自然科学基金项目(10871191,40974075)，中国科学院知识创新工程重要方向性项目(KZCX2-YW-QN107)和石家庄经济学院博士科研启动基金联合资助.

作者简介 曹静杰，男，博士，讲师，1982 年出生，2011 年于中国科学院地质与地球物理研究所获博士学位，现在石家庄经济学院工作，主要从事勘探地球物理研究. E-mail：cao18601861@163.com

* **通讯作者** 王彦飞，男，研究员，1997 年毕业于河北师范大学数学系，2002 年于中国科学院数学与系统科学研究院获博士学位. 主要从事计算及勘探地球物理反演问题理论及计算研究. E-mails：yfwang@mail.iggcas.ac.cn；yfwang_ucf@yahoo.com

gradient method can reduce the amount of computation greatly, and the restoration based on the piecewise random sampling are better than that of random sub-sampling.

Keywords Wavefield recovery, Curvelet transform, Compressed sensing, Zero-norm approximation, Inverse problems, Ill-posedness, Sparse optimization

1 引 言

在地震勘探中由于采集成本、坏道、噪声、地形等诸多因素的影响，采集的数据常常不满足采样定理，会影响多次波消除、偏移等处理的效果[1-2]，因此在数据处理过程中常常需要对地震数据进行重构. 地震数据重构主要有三大类方法. 最常用的方法是基于变换的方法，这类方法基于信号在变换域的特征进行数据重构，计算结果比较稳健，因此得到广泛的应用. 常用的变换有傅里叶变换[3-5]，Radon 变换，局部 Radon 变换[6]等. 第二类方法是预测滤波方法，基于分频预测的思想，用低频的信息预测高频信息. 这类方法主要有 F-X 域预测滤波[7]，F-K 域预测滤波[8-9]和 T-X 域预测滤波[10]. 第三类方法是基于波动方程的方法，利用波传播的物理性质重构地震波场[11-12]，这类方法由于需要速度信息且计算量大，因此没有得到广泛的应用.

地震数据重构问题可以变为一个稀疏优化问题. 文献[13]中采用 curvelet 变换作为稀疏变换，利用压缩传感的理论对随机欠采样的地震数据进行重构. 采用的迭代阈值法需要多次迭代，计算效率比较低；而随机欠采样方式不能够有效控制采样间隔，因此需要进一步改进. 本文将地震数据重建问题变为一个稀疏优化问题，并构造 0 范数的逼近函数作为目标函数，利用 curvelet 变换的紧框架性，采用投影梯度法解，同时本文运用作者近期提出的一种能够控制最大采样间隔并且保持采样随机性的分段随机采样方式. 数值试验表明基于 0 范数逼近的投影梯度法的计算效率比现有的稀疏优化解法有很大的提高；分段随机采样能够有效地控制采样间隔并且保持采样的随机性，比随机欠采样有更好的重构效果.

本文结构如下：第二节阐明了地震数据重构问题的数学模型；第三节简单介绍所采用的稀疏变换；第四节回顾了稀疏优化算法并详细介绍了 0 范数逼近方法和梯度迭代技巧；第五节探讨采样方式并提出新的采样方式；第六节的数值模拟表明分段随机采样比随机欠采样能够更好地改进重构效果；0 范数逼近方法可以显著地提高计算效率，其 CPU 计算时间大约是通常采用的迭代阈值法(IST)和谱投影法(SPGL1)的三分之一.

2 地震数据重构的数学模型

地震数据采集可以表示为如下数学模型：

$$\boldsymbol{y} = \boldsymbol{R}\boldsymbol{f}, \tag{1}$$

其中 $\boldsymbol{f} \in \boldsymbol{R}^N$ 为原始波场数据，$\boldsymbol{R}$ 为采样矩阵，$\boldsymbol{y} \in \boldsymbol{R}^M (M < N)$ 为采集的波场数据. 由于采集数据的不完整性，因此 $\boldsymbol{R}$ 是一个欠定矩阵. 地震数据重构由采集的数据 $\boldsymbol{y}$ 和采样算子 $\boldsymbol{R}$ 重构完整的数据 $\boldsymbol{f}$. 由于 $M<N$，因此存在无穷多的 $\boldsymbol{f}$ 满足公式(1). 因此地震数据重构是一个不适定的反问题[1].

方程(1)可以通过优化问题

$$\min \| \boldsymbol{y} - \boldsymbol{R}\boldsymbol{f} \|_2^2, \tag{2}$$

来求解，其中 $\|\cdot\|_2$ 是向量的 2 范数. 由于问题(1)的不适定性，必须采用正则化方法求解，其中最经典的方法是 Tikhonov 正则化[14-15]，即解

$$\min \| \boldsymbol{y} - \boldsymbol{R}\boldsymbol{f} \|_2^2 + \alpha \| \boldsymbol{f} \|_2^2. \tag{3}$$

该优化问题的最优解为 $\boldsymbol{f} = (\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} + \alpha\boldsymbol{I})^{-1}\boldsymbol{y}$，其中 $\boldsymbol{I}$ 为单位矩阵，α 是正则参数. 此外，正则化方法还包括迭代正则化、自适应正则化、截断的奇异值分解、共轭梯度法等方法[14-16].

若存在某个变换 $\boldsymbol{C}$ 使得 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{f}$ 是稀疏的，则可以先解方程 $\boldsymbol{R}\boldsymbol{C}^{-1}\boldsymbol{x} = \boldsymbol{y}$，求出稀疏的解 $\boldsymbol{x}$，然后再求出 $\boldsymbol{f} = \boldsymbol{C}^{-1}\boldsymbol{x}$. 此时问题(1)变为约束极小化问题

$$\min \| \boldsymbol{x} \|_0, \quad \text{s. t.}\ \boldsymbol{R}\boldsymbol{C}^{-1}\boldsymbol{x} = \boldsymbol{y}, \tag{4}$$

其中零范数 $\|\cdot\|_0$ 表示向量的非零元素个数，记 $\boldsymbol{R}\boldsymbol{C}^{-1} = \boldsymbol{A}$. 问题(4)需要穷举所有的组合才能找到最优解[1,17]. 在一定的条件下，可以利用 1 范数优化问题求解欠定矩阵的稀疏解[17]. 这种基于数据的稀疏性，由欠采样数据重构完整数据的问题就是压缩传感问题. 压缩传感主要由三部分组成：稀疏变换、采样方法和求解方法. 这三部分共同影响重构的结果. 本文基于压缩传感的理论进行地震数据重构，下面简要地介绍采用的稀疏变换.

3　稀疏变换

稀疏变换是压缩传感的重要组成部分．变换域中的数据越稀疏，需要的采样个数就越少,而且重构效果越好．地震数据处理中最常用的变换有傅里叶变换、Radon变换、小波变换以及Gabor变换．但是这些变换不能直接用于压缩曲线形状的地震记录，因此Candes提出了curvelet变换[18]．这种多尺度、多方向、各向异性的变换在局部窗内用线段逼近曲线，因此能够很好地压缩曲线形状的记录．连续的curvelet变换可以表示为

$$c(s,\theta,t_0,x_0)=\int_t\int_x f(t,x)\varphi(s,\theta,t_0-t,x_0-x)\mathrm{d}x\mathrm{d}t,\quad(5)$$

其中$f(t,x)$为T-X域的数据，函数$\varphi(s,\theta,t,x)$是变换的核函数，其中s为尺度，θ为角度或倾角，t_0，x_0为T-X域的位置参数．该变换将数据分解成不同的尺度和角度，其离散形式记为$c=\boldsymbol{\Psi}f$，其中c为curvelet系数，$\boldsymbol{\Psi}$为算子的离散形式，f为T-X域数据的离散形式．由于离散curvelet变换是一个紧的框架，因此$\boldsymbol{\Psi}^{\mathrm{H}}$为$\boldsymbol{\Psi}$的广义逆．

4　稀疏优化方法

基于稀疏变换的地震数据重构问题需要解一个稀疏优化问题，对于线性系统的稀疏解法研究已经有了很多成果．下面首先总结常见的稀疏解法，然后介绍在地震数据重构中得到应用的稀疏解法，在4.2节提出了地震数据重构的0范数逼近算法．

4.1　稀疏解法介绍

最常见的稀疏解法是基于1范数的方法．主要有内点法、迭代重赋权1范数方法、投影梯度类法等等．除此之外，匹配追踪类方法是著名的贪婪类方法，包括匹配追踪、正交匹配追踪和匹配追踪的其他变形,这类方法不适合大规模问题和解非精确稀疏的情况．第三类方法是阈值法[19]，该方法简单易行，能够解大规模问题，但是需要多次迭代．此外还有同伦法和最小角度回归方法，该类方法适合维数较小且解为精确稀疏的情形,迭代重赋权最小二乘法通过在迭代过程中调节解的权重得到稀疏解.

若信号在某个变换域是稀疏的，则在一定条件下可以利用1范数优化问题求出问题的稀疏解[17]，即解以下问题：

$$\min\|\boldsymbol{x}\|_1,\ \text{s.t.}\ \boldsymbol{Ax}=\boldsymbol{y}.\quad(6)$$

问题(6)可以变为线性规划问题，并用内点法等方法[1,20]求解．若采集数据含有加法噪声n，即$\boldsymbol{y}=\boldsymbol{Ax}+n$，则问题(4)变为

$$\min\|\boldsymbol{x}\|_1,\ \text{s.t.}\ \|\boldsymbol{Ax}-\boldsymbol{y}\|_2<\varepsilon.\quad(7)$$

问题(7)可以变为二阶锥优化问题求解．其无约束优化形式为

$$\min\|\boldsymbol{Ax}-\boldsymbol{y}\|_2^2+\lambda\|\boldsymbol{x}\|_1,\quad(8)$$

其中λ为拉格朗日乘子．问题(8)可以用内点法、投影梯度法及同伦算法等方法求解[21-22]．

文献[23]提出一种解不适定问题的l_p-l_q模型：

$$\min\|\boldsymbol{Ax}-\boldsymbol{y}\|_p^p+\lambda\|\boldsymbol{x}\|_q^q,\ p,q>0.\quad(9)$$

该模型给出了解优化问题的一个统一框架，可以求解一般的反问题．基于变换的地震数据重构问题的先验条件是解的稀疏性．当$p=2,q\to1$时，问题(9)可以求出稀疏解．文献[24]进一步提出了1范数信赖域解法．在$q=0$的情形，文献[25]提出基于0范数的稀疏分解法.

对于线性问题求稀疏解有许多方法，但是由于地震数据量大，目前只有基于梯度的方法被用到地震数据重构中．文献[26]利用傅里叶变换作为稀疏变换，并采用迭代重赋权最小二乘法(IRLS)求解.

Hennenfent等对问题(8)采用迭代软阈值法(IST)求解[13]．该方法的迭代方式为：若$\boldsymbol{x}_k$为当前迭代点，则下一个迭代点为$x_{k+1}=T_\lambda(x_k+\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{Ax}_k))$，其中$T_\lambda(x)=\mathrm{sgn}(x)\cdot\max(0,|x|-|\lambda|)$．该迭代方式简单，参数设定容易，但是不能保证解的稀疏性且可能收敛到局部解.

问题(7)也可以采用投影梯度法来求解，Ewout等提出了一种谱投影梯度法(SPGL1)[27]．该方法利用Pareto曲线$\phi(\tau)=\|\boldsymbol{r}_\tau\|_2$(其中$r_\tau=\boldsymbol{y}-\boldsymbol{Ax}_\tau$)，将以下两个优化问题联系起来

$$(BP_\sigma)\quad\min\|\boldsymbol{x}\|_1,\ \text{s.t.}\ \|\boldsymbol{Ax}-\boldsymbol{y}\|_2\leqslant\sigma,\quad(10)$$

$$(LS_\tau)\quad\min\|\boldsymbol{Ax}-\boldsymbol{y}\|_2,\ \text{s.t.}\ \|\boldsymbol{x}\|_1\leqslant\tau.\quad(11)$$

令x_τ为(LS_τ)的最优解，Pareto曲线是关于τ的连续可微凸函数．当τ取合适的值时，(BP_σ)和(LS_τ)有相同的解，因此可以通过解(LS_τ)得到(BP_σ)的最优解．对于(LS_τ)采用投影梯度法求出解x_τ，然后通过Pareto曲线来更新τ，通过解新的(LS_τ)问题逐渐逼近(BP_σ)的准确解．在求解(LS_τ)问题时，利用对偶间隙作为停机准则，每次迭代将负梯度方向投影到1范数球上，然后采用非单调梯度法求步长，该方法能够保证解的快速收敛性[28]．

不难发现谱投影梯度法中的参数τ的求解方法

实质上是利用了反问题研究领域中的后验选取正则参数的偏差准则，并且该参数 τ 可以用高阶收敛算法求得[15,29].

以上三种方法具有相同的收敛性[30]，但是地震数据处理需要更加快速稳定的解法. 因此，研究大规模问题的稳定高效的快速计算方法对于地震数据处理有重要意义.

下面提出一种地震数据重构的快速方法. 该算法构造简单，参数设置比较容易，而且在相同重构效果情况下，其计算时间大约是谱投影梯度法和迭代阈值法的三分之一. 数值模拟证明了算法的高效性.

4.2　0 范数逼近算法

按照 $l_p - l_q$ 的正则化思想，当 $q \to 0$ 时可以根据问题(9)求出稀疏解. 因此我们研究 0 范数优化问题：

$$\min \| \boldsymbol{x} \|_0, \mathrm{s.t.}\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}. \tag{12}$$

大多数方法采用 1 范数作为 0 范数的逼近，然后解 1 范数优化问题. 我们知道 1 范数是 0 范数的最好的凸逼近形式，这是一种近似转化的方法. 下面我们构造非凸函数作为 0 范数的逼近. 为此我们构造满足以下性质的函数 $f_\sigma(x)$：(1) 对于固定的 σ，$f_\sigma(x)$ 在 x 的一个合适的邻域内是连续的凸函数；(2) 当 σ 趋向于 0 时，函数逼近于 0 范数：

$$\lim_{\sigma \to 0} f_\sigma(x) = \begin{cases} 0, & x = 0, \\ 1, & x \neq 0. \end{cases} \tag{13}$$

通过研究发现函数

$f_\sigma(x) = 1 - \exp(-x^2/2\sigma^2)$，第 1 种函数 (14)

和

$f_\sigma(x) = 1 - \sigma^2/(x^2 + \sigma^2)$，第 2 种函数 (15)

满足上述两个性质. 图 1 给出了一维情形下这两个函数的图像.

这两种函数是非凸的，而且随着参数 σ 变小，逐渐逼近 0 范数. 因此对于问题(12)，我们可以构造连续函数来作为目标函数，然后求其极小值. 这样问题(12)就变为

$$\min \sum_{i=1}^{N} f_\sigma(x_i), \mathrm{s.t.}\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}. \tag{16}$$

由于目标函数是 σ 的函数，σ 越小，函数越逼近 0 范数，由于函数的非凸性，对一个小的 σ 值，直接求解问题(16)很难获得全局最优解. 但是对于较大的 σ 值，函数 $\sum_{i=1}^{N} f_\sigma(x_i)$ 是光滑的，容易求出全局最优解，因此可以通过针对逐渐下降的 σ 序列来解(16)，用上次的近似解作为下一次迭代的初始解，并逐渐降低 σ. 这样就可以防止迭代得到局部最优解. 该方法本质上是一种逼近的投影梯度法. 对于上述两种函数，该方法具有相同的效果.

(1) 给出内部循环次数 L，外部循环次数 J，步长 ω，初始解 x_0(x_0 为 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 的 2 范数解)，初始的 σ 值，并令 $l = 0, j = 0, x = x_0$；

(2) 求出当前迭代点 $\boldsymbol{x}$ 的梯度 $g = \nabla\left(\sum_{i=1}^{M}(1 - f_\sigma(x_i))\right)$，更新迭代点：$\boldsymbol{x} := \boldsymbol{x} - \omega g$($\omega$ 为迭代步长，需要线搜索计算)；

(3) 将 $\boldsymbol{x}$ 投影到可行集 $\boldsymbol{Z} = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}\}$ 上，即 $\boldsymbol{x} := \boldsymbol{x} - \boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y})$；

(4) 令 $l := l + 1$，如果 $l < L$，转第(2)步；否则，令 $\sigma = \sigma/2, j = j + 1, l = 0$，转步 2；

(5) 判断停机准则：如果 $j > J$，则停机，输出解 $\hat{\boldsymbol{x}}$；否则，转步(2)，继续迭代.

由于本文的 curvelet 变换是正交的，因此矩阵 $\boldsymbol{A}$ 的广义逆的求取比较方便，$\boldsymbol{x} := \boldsymbol{x} - \boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y})$ 可以很容易求出，因此大大地提高了计算效率. 该算法结合了正则化和最优化技巧，因此具有稳定的收敛性. 关于梯度算法的正则性和收敛性的理论性结果见文献[14,25].

5　分段随机采样

在地震勘探中，观测系统的设计对地震数据处理有一定的影响. 在相同采样个数下，如何设计好的采样方式是值得研究的方向. 本节给出一种分段随机采样方法，该采样方法能够保持随机性和控制采样间隔，因此能够改进数据重构的效果.

采样方式是压缩传感的重要组成部分，采样方法关系到重构的效果. 虽然在其他领域已经有各种采样方法[31]，但是由于地震采集只能在地表放置检波器接收地震信号，因此采样方法受到很大的限制. 目前的欠采样方法主要分为规则欠采样和随机欠采样[24]. 规则欠采样在测线上等间隔地放置检波器，但是采样间隔大于完整采样所要求的间隔. 规则欠采样如图 2 所示.

其中的黄点代表放置了检波器，白点表示没有放置检波器. 这种规则欠采样在 F-K 域会产生严重的相干假频.

随机欠采样在测线上随机的放置检波器，其采样间隔不能得到控制. 随机欠采样如图 3 所示.

随机欠采样在F-K域产生的噪声分布在整个

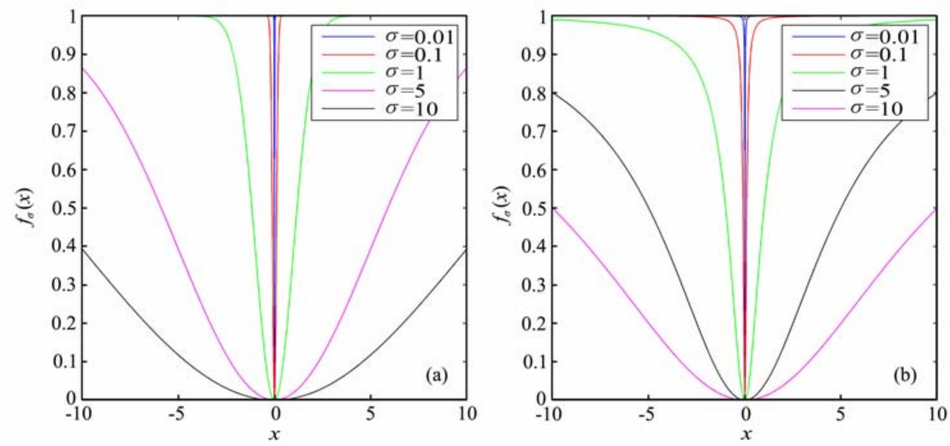

图 1 第一种函数 $f_\sigma(\boldsymbol{x}) = 1 - \exp(-x^2/2\sigma^2)$(a)和第二种函数 $f_\sigma(\boldsymbol{x}) = 1 - \sigma^2/(x^2+\sigma^2)$(b)

Fig. 1 (a) Case 1 funcation: $f_\sigma(\boldsymbol{x}) = 1 - \exp(-x^2/2\sigma^2)$; (b) Case2 function: $f_\sigma(\boldsymbol{x}) = 1 - \sigma^2/(x^2+\sigma^2)$

图 2 规则欠采样方式

Fig. 2 Schematic of regular sub-sampling

图 3 随机欠采样方式

Fig. 3 Schematic of random sub-sampling

频率域，可以看做随机噪声，因此可将波场重构看作去噪问题[13]．由于随机欠采样会产生大的采样间隔，当这些间隔大于 curvelet 的某个尺度时，重构结果就会丢失原始信息，因此需要提出新的采样方法来克服随机欠采样的缺点.

Hennenfent 引进一种 jittered 采样方法[32]．这种采样方法先等间隔的选取点，然后以这些点为中心随机的扰动．jittered 采样能够控制采样间隔而且具有随机性，但是该采样方法缺乏灵活性，只能取完整采样个数的整数分之一.

针对上述三种采样方法的缺点，本文详细描述文献[24]给出的一种分段采样方法．对于随机欠采样来说，每个点被采样到的概率是相等的，假设 Nyquist 采样个数为 N，实际采样个数为 K，则采样比例为 K/N．该采样方法不能够控制最大间隔，其最大间隔为 $N-K$，当采样间隔大于 curvelet 的某个尺度时，则在这个尺度内的采样为零，因此重构的数据必然会丢失原始信息．分段随机采样首先将这 N 点均匀的分成 M 段，每段的长度为 N/M，在每段内按照一定的比例随机的采样，这种采样方式下，其最大采样间隔为 $2(N/M)(1-K/N)$，但是这种概率出现的极小．因此这种分段随机采样方法能够有效地减小采样间隔，同时保持了采样的随机性.

分段随机采样的具体的操作如下：(1) 将所有的采样个数 N 划分为 M 段($M<K$)，每段长度 N/M 小于 curvelet 的尺度 Q. (2) 在每段上随机的取 K/M 个点，每段的采样点的总和等于所有的采样点数 K．当划分的段数 M 足够多时，这种方法既能够有效地控制最大采样间隔又能够保持采样的随机性，因此能够得到好的重构效果.

6 数值试验

6.1 分段随机采样的数值模拟

为了证明新的分段随机采样的效果，本文对模拟的炮记录进行数值试验．图 4 是一个水平层状介质的地震炮记录及其频谱，空间采样间隔为 15 m，时间采样间隔为 2 ms，该数据含有直线形状的直达波和双曲线形状的反射波．图 5a 是规则欠采样数据，其采样个数为图 4a 采样个数的 1/3，图 5b 产生严重的相干假频．图 6a 是随机欠采样数据，和图 5a 有相同的采样个数，图 6b 是图 6a 的频谱．图 7a 是图 6a 的重构数据，其信噪比为 7.1606，图 7b 是图 7a 的频谱．信噪比公式为

$$\mathrm{SNR} = 10 * \lg\left(\frac{\|\mathrm{orig}\|_2}{\|\mathrm{orig}-\mathrm{rest}\|_2}\right)^2, \quad (17)$$

其中 orig 为完整的记录，rest 为重构的记录．图 8 是 jittered 采样数据及其频谱，图 8a 和图 5a 的采样个数相同．图 9 表示图 8a 的重构数据及其频谱，重构结果的信噪比为 9.3008．图 10 表示分段采样数据及其频谱，图10a和图5a具有同样的采样个数．图 11 是分段采样的重构数据及其频谱，重构结果的信噪比为 9.8417．经过多次试验，随机欠采样

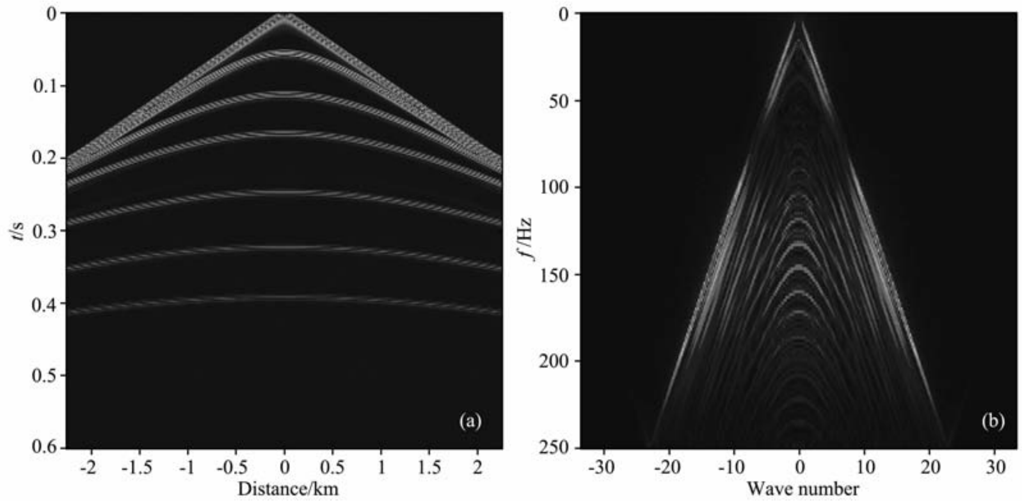

图 4　原始数据 (a)及其频谱 (b)

Fig. 4　(a) The original data; (b) Frequency spectrum of (a)

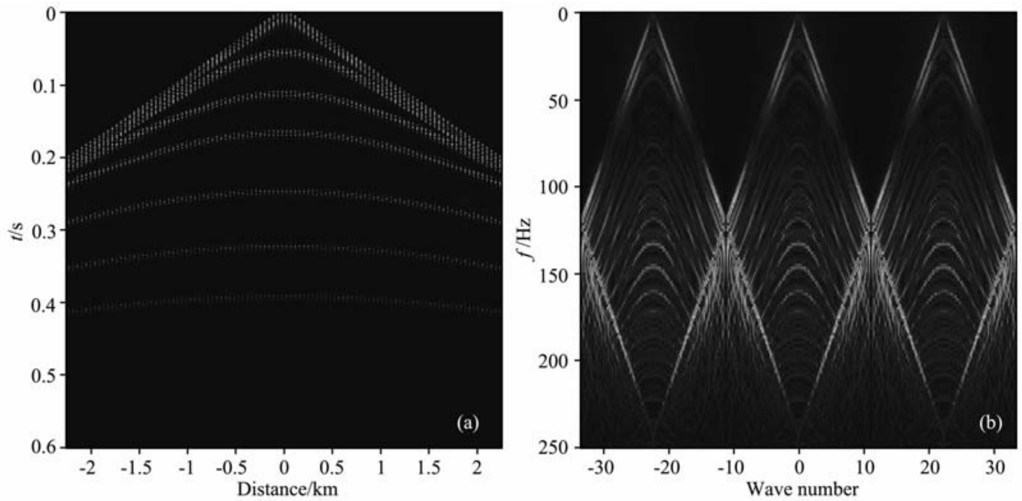

图 5　规则欠采样的数据 (a)及其频谱 (b)

Fig. 5　(a) Data of regular sub-sampling; (b) Frequency spectrum of (a)

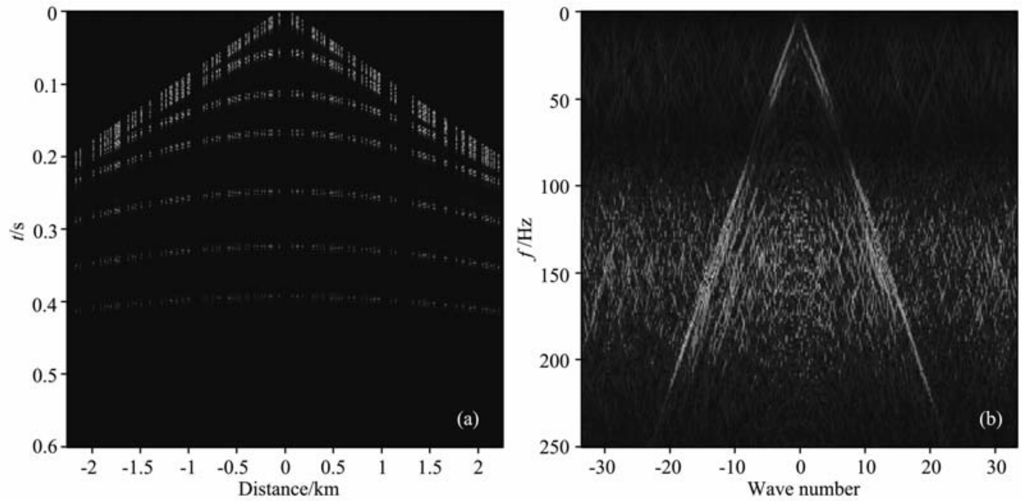

图 6　随机欠采样数据(a)及其频谱(b)

Fig. 6　(a) Data of random sub-sampling; (b) Frequency spectrum of (a)

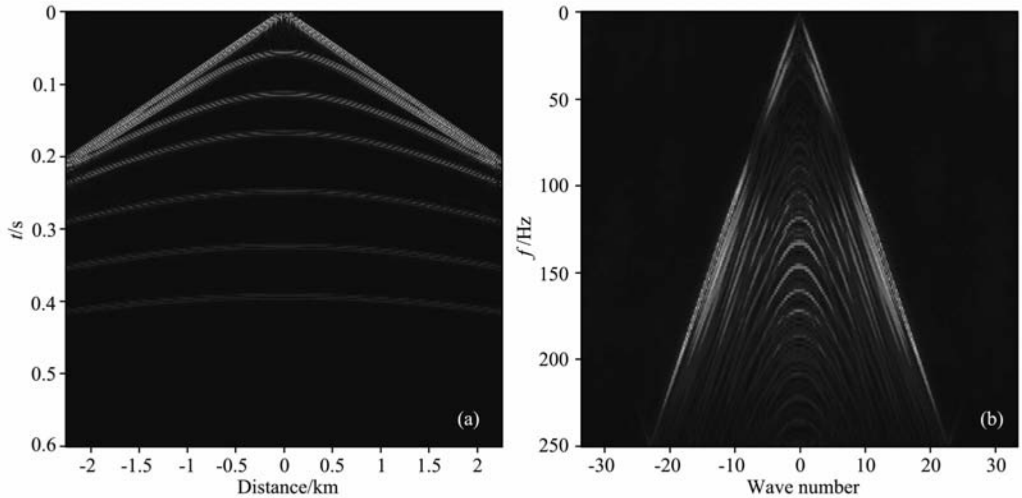

图 7 随机欠采样的重构数据(a)及其频谱(b)

Fig. 7 (a) Restoration of random sub-sampling; (b) Frequency spectrum of (a)

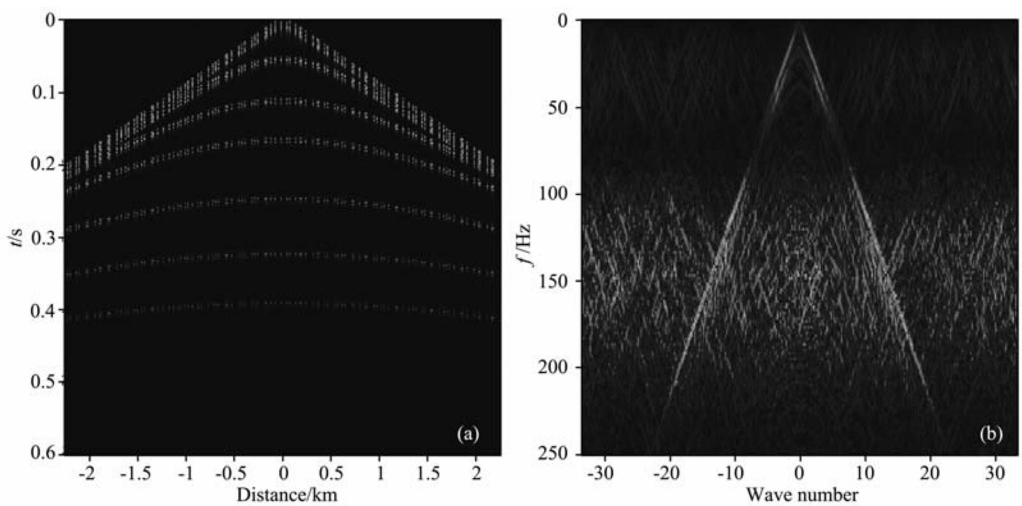

图 8 jittered 采样数据(a)及其频谱 (b)

Fig. 8 (a) Data of jittered sampling; (b) Frequency spectrum of (a)

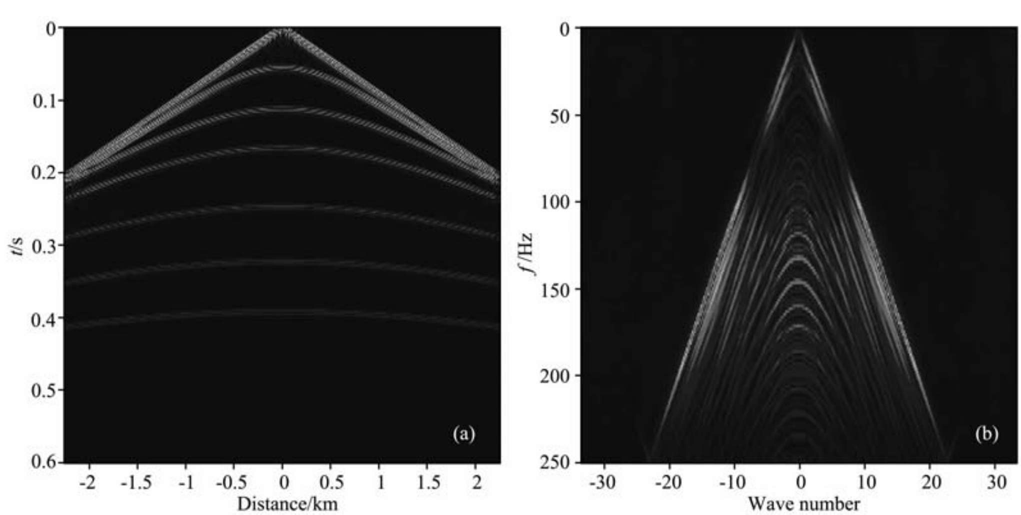

图 9 jittered 采样的重构数据(a)及其频谱(b)

Fig. 9 (a) Restoration of jittered sampling; (b) Frequency spectrum of (a)

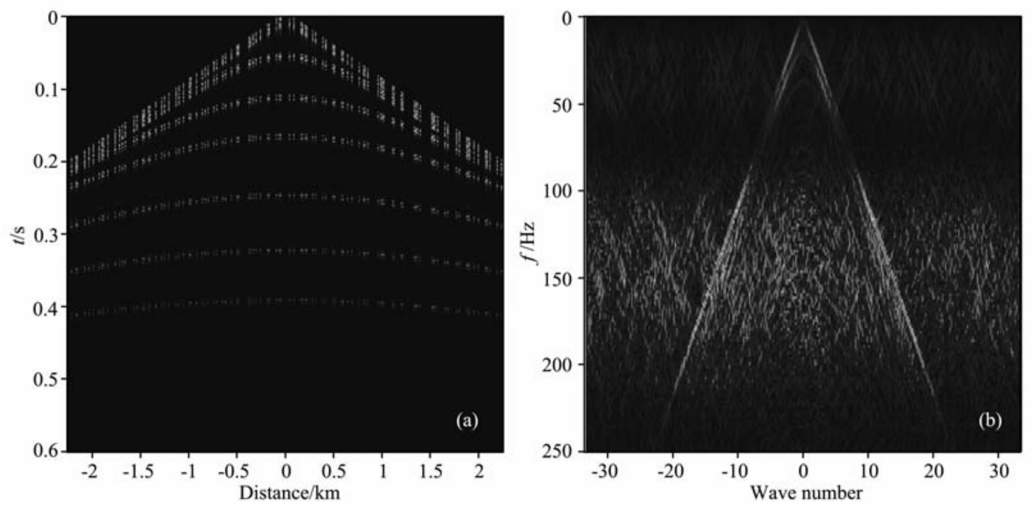

图 10　分段随机采样数据(a)及其频谱(b)

Fig. 10　(a) Data of piecewise random sampling;(b) Frequency spectrum of (a)

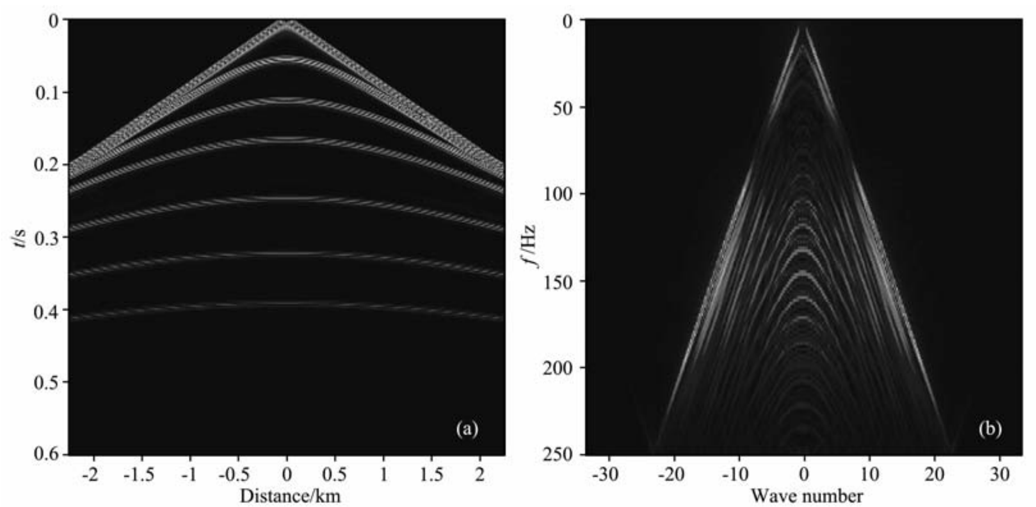

图 11　分段随机采样的重构数据(a)及其频谱(b)

Fig. 11　(a) Restoration of piecewise random sampling;(b) Frequency spectrum of (a)

重构的信噪比在 7.5 左右；jittered 随机欠采样的信噪比在9.2左右；分段采样重构的信噪比在 9.8 左右. 因此分段随机采样的重构效果有明显的改进.

6.2　0 范数逼近算法的数值模拟

下面给出 0 范数逼近方法的数值模拟. 图 12a 中的地震数据的空间采样间隔为 15m，共有 300 道，时间采样间隔为 2ms，采样时间为 0.6s，图 12b 是已知的采样数据，采样个数为图 12a 的一半. 图 13a 是第 1 种函数 $f_\sigma(x)=1-\exp(-x^2/2\sigma^2)$ 时 0 范数逼近算法的重构数据，信噪比为 25.5282，计算时间为 306 s，图 13b 是图 13a 与原始数据的误差. 图 14a 是第 2 种函数 $f_\sigma(x)=1-\sigma^2/(x^2+\sigma^2)$ 时 0 范数逼近算法的重构结果，信噪比为 25.6524，计算时间为 318 s，图 14b 是图 14a 与原始数据的误差. 图 15a 是迭代阈值法(IST)的重构数据，信噪比为 24.8597，计算时间为 1326 s，图 15b 是图 15a 与原始数据的误差. 图 16a 是谱投影梯度法(SPGL1)的重构数据，信噪比为25.1210，计算时间为 1069 s，图 16b 是图 16a 与原始数据的误差. 表 1 列出了以上三种方法重构的 CPU 计算时间、相对误差以及信噪比. 综上可见，0 范数逼近算法的计算时间明显减少，约为迭代阈值法和谱投影梯度法的三分之一；因此 0 范数逼近方法更适合地震数据重构.

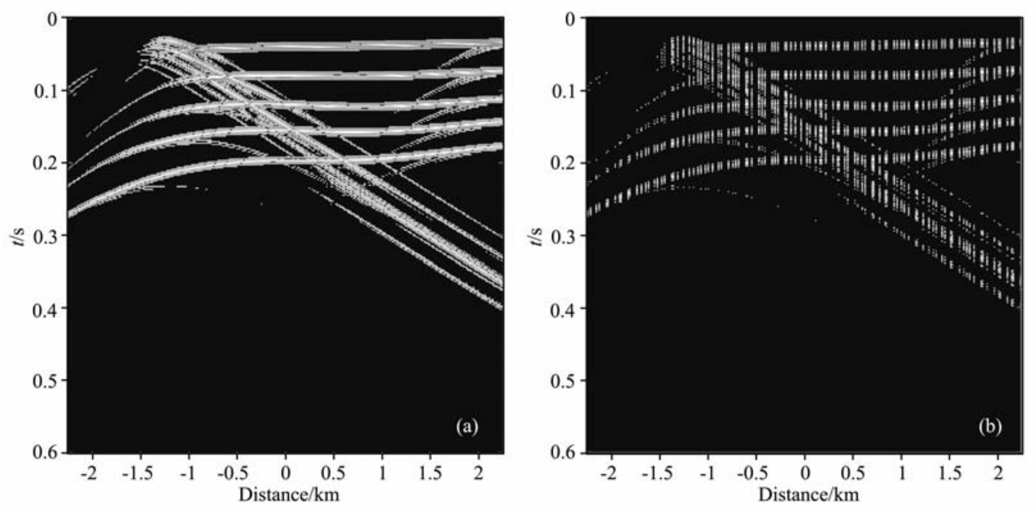

图 12 原始数据 (a)和采样数据 (b)

Fig. 12 (a) The original data;(b) The sampled data

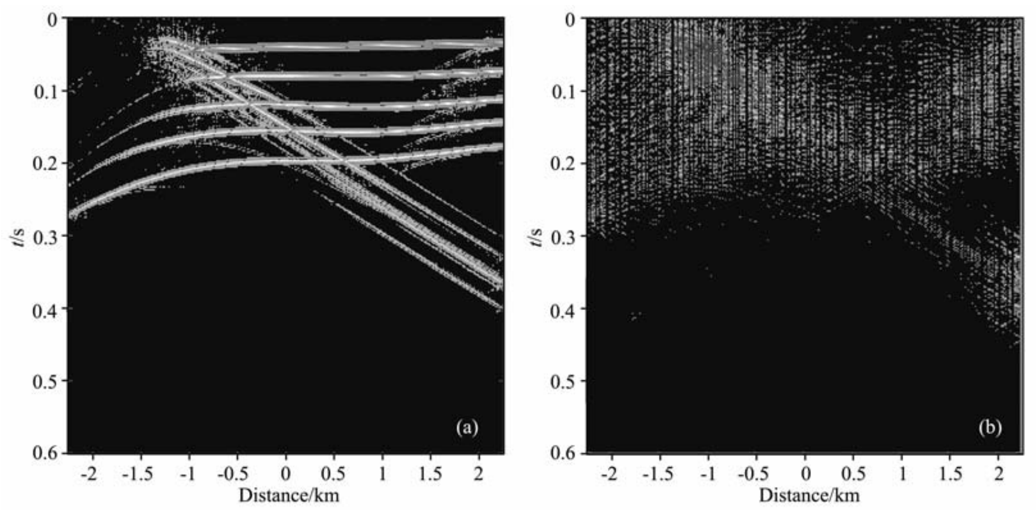

图 13 基于函数 1 的 0 范数逼近算法的重构数据(a)及其与原始数据的误差(b)

Fig. 13 (a) Restoration of zero-norm approximate method based on case 1 function;(b) Difference between the original data and (a)

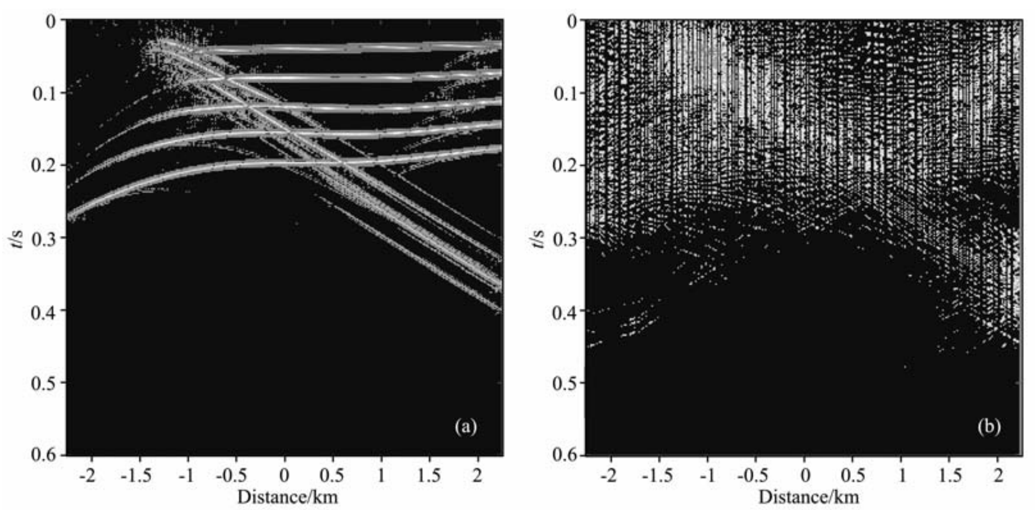

图 14 基于函数 2 的 0 范数逼近算法的重构数据 (a)及其与原始数据的误差 (b)

Fig. 14 (a) Restoration of zero-norm approximate method based on case 2 function;(b) Difference between the original data and (a)

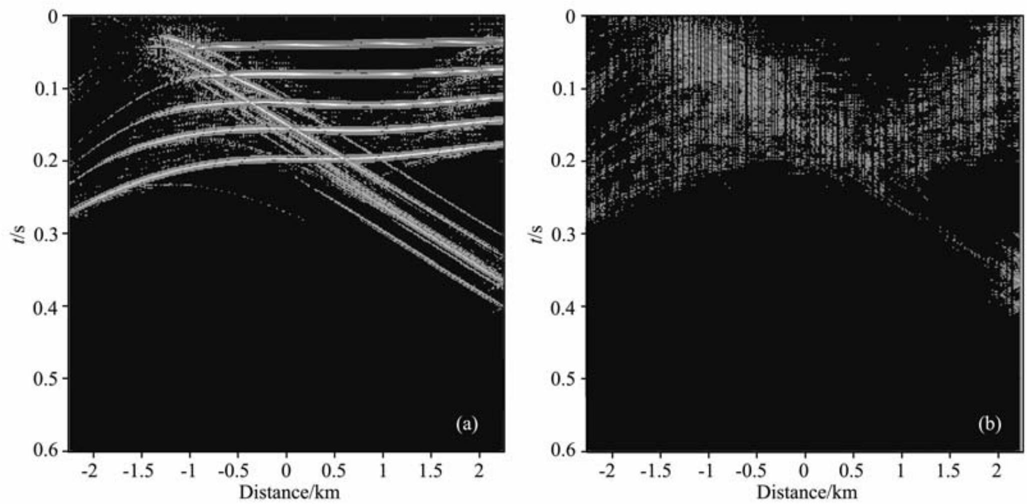

图 15 迭代阈值法(IST)的重构数据(a)及其与原始数据的误差(b)

Fig. 15 (a) Restoration of IST method;(b) Difference between the original data and (a)

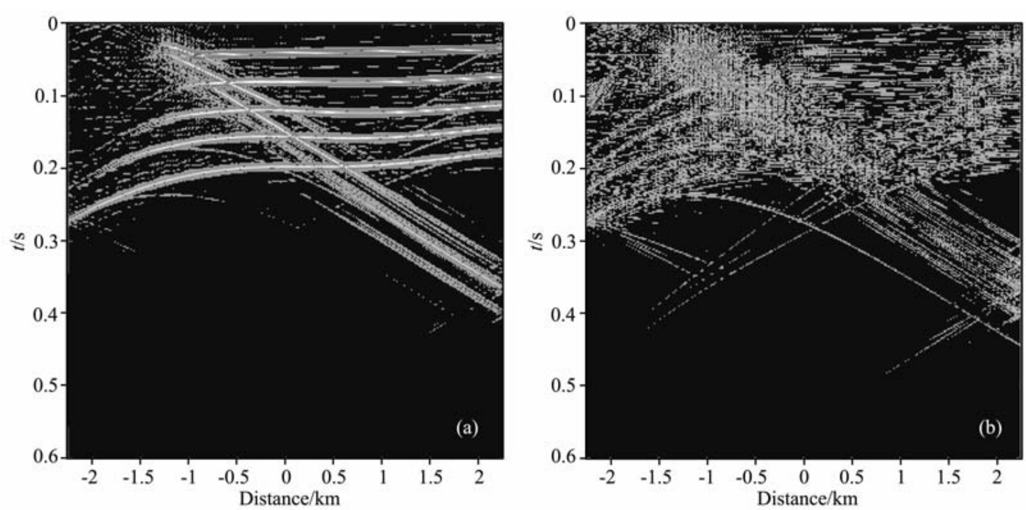

图 16 谱投影梯度法(SPGL1)的重构结果 (a)及其与原始数据的误差 (b)

Fig. 16 (a) Restoration of the SPGL1 method;(b) Difference between the original data and (a)

表 1 三种算法重构结果的比较

Table 1 The restoration results comparison of the above three methods

方法	相对误差	信噪比	计算时间/s
0 范数逼近方法:第 1 种函数	0.0529	25.5282	306
0 范数逼近方法:第 2 种函数	0.0522	25.6524	318
迭代阈值法(IST)	0.0571	24.8597	1326
谱投影梯度法(SPGL1)	0.0555	25.1210	1069

7 结论和讨论

本文采用 curvelet 变换作为稀疏变换,将地震数据重构问题变为稀疏优化问题,提出了一种基于 0 范数逼近的投影梯度法求解地震数据重构问题;并且利用最近提出的分段随机采样方式来改进重构的效果. 数值模拟显示,在相同重构结果情况下,基于 0 范数逼近的投影梯度算法的计算时间是谱投影梯度法和迭代阈值法的三分之一,因此具有更快的计算效率. 分段随机采样能够控制采样间隔和保持采样的随机性,因此比随机欠采样方式更适合地震数据重构.

稀疏优化技术在地震数据处理中具有广阔的应用前景,比如时频分析、去噪、多次波消除和偏移等等. 高精度反褶积问题[33]是时间域的插值,因此可以利用压缩传感的理论解高精度反褶积问题,进行时间域的信息重构.

稀疏变换直接影响重构的结果,变换后的数据越稀疏,越有利于数据恢复;快速高效的求解方法

是稀疏优化的核心；好的采样方式在保证质量的前提下，能够节约采样时间和成本；三者都是值得研究的重要方向.

基于稀疏重构的正则化方法和最优化方法是当前的研究热点. 稀疏优化解法目前主要集中在基于1范数的解法，真正的基于0范数的快速稳定方法会有更好的数值结果.

致 谢 感谢评审人提出的宝贵意见，使得本文的研究内容有了很大的充实.

参考文献(References)

[1] Wang Y F, Yang C C, Cao J J. On Tikhonov regularization and compressive sensing for seismic signal processing. *Mathematical Models and Methods in Applied Sciences*, 2012, 22(2), doi : 10.1142/S0218202511500084.

[2] Cao J J, Wang Y F, Zhao J T, et al. A review on restoration of seismic wavefields based on regularization and compressive sensing. *Inverse Problems in Science and Engineering*, 2011, 19(5): 679-704.

[3] Sacchi M D, Ulrych T J. Estimation of the discrete Fourier transform, a linear inversion approach. *Geophysics*, 1996, 61(4): 1128-1136.

[4] Duijndam A J W, Schonewille M A, Hindriks C O H. Reconstruction of band-limited signals, irregularly sampled along one spatial direction. *Geophysics*, 1999, 64(2): 524-538.

[5] Liu B. Multi-dimensional reconstruction of seismic data[D]. Alberta: University of Alberta, 2004.

[6] Sacchi M D, Verschuur D J, Zwartjes P M. Data reconstruction by generalized deconvolution. 74th Annual International Meeting, SEG, Expanded Abstracts, 2004: 1989-1992.

[7] Spitz S. Seismic trace interpolation in the F-X domain. *Geophysics*, 1991, 56(6): 785-794.

[8] Gülünary N. Seismic trace interpolation in the Fourier transform domain. *Geophysics*, 2003, 68(1): 355-369.

[9] Claerbout J F. Earth Soundings Analysis: Processing Versus Inversion. Blackwell Science, 1992.

[10] Crawley S. Seismic trace interpolation with nonstationary prediction-error filters[D]. Stanford: Stanford University, 2000.

[11] Ronen J. Wave-equation trace interpolation. *Geophysics*, 1987, 52(7): 973-984.

[12] Stolt R H. Seismic data mapping and reconstruction. *Geophysics*, 2002, 67(3): 890-908.

[13] Hennenfent G, Herrmann F J. Random sampling: New insights into the reconstruction of coarsely-sampled wavefields. *SEG, Expanded Abstracts*, 2007, 26: 2575-2579.

[14] 王彦飞. 反问题的计算方法及其应用. 北京：高等教育出版社，2007.
Wang Y F. Computational Methods for Inverse Problems and Their Applications (in Chinese). Beijing: Higher Education Press, 2007.

[15] 肖庭延，于慎根，王彦飞. 反问题的数值解法. 北京：科学出版社，2003.
Xiao T Y, Yu S G, Wang Y F. Numerical Methods for the Solution of Inverse Problems (in Chinese). Beijing: Science Press, 2003.

[16] 袁亚湘，孙文瑜. 最优化理论与方法. 北京：科学出版社，1997.
Yuan Y X, Sun W Y. Theory and Methods for Optimization (in Chinese). Beijing: Science Press, 1997.

[17] Candes E J. Compressive sampling. // Proceedings of International Congress of Mathematicians. Madrid: European Mathematical Society Publishing House, 2006: 33-52.

[18] Candes E J, Donoho D L. Curvelets: A surprisingly effective nonadaptive representation of objects with edges. // Cohen A, Rabut C, Schumaker L L, eds. Curves and Surfaces Fitting: Saint-Malo 1999. Nashville: Vanderbilt University Press, 2000: 105-120.

[19] Herrity K K, Gilbert A C, Troop J A. Sparse approximation via iterative thresholding. // Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Proceeding, 2006: 624-627.

[20] Wang Y F, Fan S F, Feng X. Retrieval of the aerosol particle size distribution function by incorporating a priori information. *Journal of Aerosol Science*, 2007, 38(8): 885-901.

[21] Wang Y F. Seismic impedance inversion using l_1-norm regularization and gradient descent methods. *J. Inv. Ill-Posed Problems*, 2011, 18(7): 823-838.

[22] Wang Y F, Ma S Q, Yang H, et al. On the effective inversion by imposing a priori information for retrieval of land surface parameters. *Science in China D*, 2009, 52(4): 540-549.

[23] Wang Y F, Cao J J, Yuan Y X, et al. Regularizing active set method for nonnegatively constrained ill-posed multichannel image restoration problem. *Applied Optics*, 2009, 48(7): 1389-1401.

[24] Wang Y F, Cao J J, Yang C C. Recovery of seismic wavefields based on compressive sensing by an l_1-norm constrained trust region method and the piecewise random subsampling. *Geophysical Journal International*, 2011, 187(1): 199-213.

[25] Mohimani H, Babaie-Zadeh M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed l_0 norm. *IEEE Transactions on Signal Processing*, 2009, 57(1): 289-301.

[26] Zwartjes P M, Sacchi M D. Fourier reconstruction of nonuniformly sampled, aliased seismic data. *Geophysics*, 2007, 72(1): V21-V32.

[27] van den Berg E, Friedlander M P. Probing the pareto frontier

for basis pursuit solutions. *SIAM J. Sci. Comput.*, 2008, 31(2): 890-912.

[28] Wang Y F, Yang C C. Accelerating migration deconvolution using a nonmonotone gradient method. *Geophysics*, 2010, 75(4): S131-S137.

[29] Wang Y F, Xiao T Y. Fast realization algorithms for determining regularization parameters in linear inverse problems. *Inverse Problems*, 2001, 17(2): 281-291.

[30] Hennenfent G, van den Berg E, Friedlander M P, et al. New insights into one-norm solvers from the Pareto curve. *Geophysics*, 2008, 73(4): A23-A26.

[31] Lustig M, Donoho D L, Pauly J M. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 2007, 58(6): 1182-1195.

[32] Hennenfent G, Herrmann F J. Simply denoise: wavefield reconstruction via jittered undersampling. *Geophysics*, 2008, 73(3): V19-V28.

[33] Taylor H L, Banks S C, McCoy J F. Deconvolution with the l_1 norm. *Geophysics*, 1979, 44(1): 39-52.

（本文编辑　胡素芳）