

基于扩散算子的带限图信号加权重建策略

杨立山 游康勇 郭文彬*

(北京邮电大学信息与通信工程学院 北京 100876)

摘要: 图信号处理技术将经典信号处理的概念和算法延展到图结构信号的处理领域。对于带限图信号,可以通过分析信号之间的关联性,重建出未采样的信号。该文分析了未采样信号的构成架构,提出一个基于扩散算子的未采样信号迭代重建算法。在每次迭代过程中,将已采样信号的重建残差扩散至所有未采样的信号节点,并进一步通过初步估计结果与重建残差的加权处理,提升算法的收敛速度。采用合成数据和真实数据进行仿真验证,实验结果显示所提出的算法具有低重建误差和快速收敛的特点。

关键词: 带限图信号处理; 图信号扩散算子; 下采样与重建

中图分类号: TN911.7

文献标识码: A

文章编号: 1009-5896(2017)12-2937-08

DOI: 10.11999/JEIT170106

Graph Diffusion Operator Based Weighted Reconstruction Strategy for Band-limited Graph Signals

YANG Lishan YOU Kangyong GUO Wenbin

(School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: Signal processing on graphs extends signal processing concepts and methodologies from the classical signal processing theory to data indexed by general graphs. For a band-limited graph signal, the unsampled data can be reconstructed from the sampled data by exploiting the relationship of the graph signals. This paper proposes a concept of graph diffusion operator for signal processing on graphs, and uses the operator to reconstruct band-limited graph signals from the sampled data. In each iteration, the residuals of the sampled vertices are propagated to all the unsampled vertices, and the known information and initial estimated results are further exploited via weighted process, aiming at accelerating the convergence. An analysis framework is proposed for the unsampled graph signals. The simulation results of synthetic data and real-world data demonstrate the wonderful effectiveness of the proposed reconstruction strategy.

Key words: Band-limited graph signals processing; Graph diffusion operator; Downsampling and reconstruction

1 引言

近些年来,随着传感器技术的发展与工业化革命,在气象预测、智能交通等监控网络内部署了大量传感器设备。由于传感器设备空间分布的不规则性,其所采集的大规模数据也呈现空间分布的不规则性,从而限制了 PCA, SVD 等经典规则域算法的应用。因此,面对不规则域下的数据处理问题,需要新的算法、思维和处理模式,促进图信号处理技术的快速发展^[1,2]。

图信号处理技术中大规模数据之间的关系被映射为 2 维平面的图结构,图中的节点代表单个数据

元素,节点之间的连接边表示数据元素之间的关联关系,所有节点的信号总称为图信号。图信号可以理解为一个具有空间关联关系的矢量信号,由于信号间的关系复杂并且量级庞大,利用小规模信号集表征原始信号是非常必要的。图信号采样与重建的实质从图信号节点集合中选取合适的部分节点,利用部分节点的数据,根据图拓扑结构以及图信号的特点,实现对原始数据的表征与恢复。目前关于图信号采样与重建问题的研究有以下几个方面。文献[3]分析和利用图结构的邻接矩阵,建立了基于优化理论的采样与非迭代重建理论,给出了能够实现无失真重建的条件。文献[4]采用数据汇聚的思想,将图结构中的节点通过图邻接矩阵汇聚至单个节点,从而实现对图结构的采样。与上述两个方面不同,本文重点研究基于图结构的拉普拉斯矩阵以及凸集映射的采样信号迭代重建策略。Persenson^[5]建立了

收稿日期: 2017-02-08; 改回日期: 2017-09-11; 网络出版: 2017-10-27

*通信作者: 郭文彬 gwb@bupt.edu.cn

基金项目: 国家自然科学基金(61271181)

Foundation Item: The National Natural Science Foundation of China (61271181)

基于 Paley-Wiener 函数的图信号采样理论基础, 提出了采样信号唯一集的概念, 给出了信号无失真重建的充分条件。关于采样图信号的重建问题, 现有重建算法的核心思想是将经典信号处理中的 Papoulis-Gerchberg 算法^[6,7]引入到不规则域, 采用凸集映射的思想迭代重建信号。文献^[8,9]提出了一个最小均方迭代重建算法 (Iterative Least Square Reconstruction, ILSR), 将已采样的图信号交替投影至顶点域下采样子空间和图谱域低通滤波器子空间, 寻找能够同时满足两个空间限制条件的信号。Wang 等人^[10,11]提出了局部集的概念及其划分方法, 利用局部集的划分选取合适节点构成下采样集, 并在此概念的基础上提出了两个迭代重建算法 (Iterative Propagating Reconstruction, IPR 和 Iterative Weighting Reconstruction, IWR)。这两个算法设定的前提是将图拓扑结构划分成若干个互不相连的子集, 单个子集称之为局部集。与 ILSR 相比, IPR 最主要的特点是建立了采样信号重建残差的传播机制。在每次迭代重建过程中, 重建残差从已采样的节点传播到位于同一局部集中的未采样节点, 然后将组合后的图信号投影至图谱域低通滤波器子空间。得益于采样信号重建残差的传播, IPR 能够得到比 ILSR 更快的重建收敛速度。但是, IPR 将采样节点的重建残差直接赋予同一局部集合内的非采样节点, 不仅忽略了局部集合外的节点对局部集合内节点影响, 而且无法反映真实图信号重建残差之间的差异性, 从而影响重建收敛速度。另外, 现阶段的研究缺少未采样信号构成要素的分析。

本文针对从采样信号中重建原始带限图信号的问题, 充分利用带限图信号的平滑特性, 构建了恢复未采样信号的分析框架, 从均值与偏置的角度建立了采样信号重建残差的扩散模型, 定义了局部均值和全局偏置扩散算子的概念, 并在此基础上提出了基于扩散算子的迭代重建算法。建立了采样信号与未采样信号之间的关联关系, 利用局部均值与全局偏置的协同处理机制消除了未采样信号的重建残差在局部区域内的过度一致化, 并进一步通过初步估计结果与重建误差的加权处理, 加速信号重建的收敛速度。仿真结果显示, 所提算法具有较低的迭代重建误差和快速的重建收敛速度。

2 图信号处理相关工作

用 $G = (\mathbf{V}, \mathbf{E})$ 表示一个简单无向图的图结构, $\mathbf{V} = \{1, 2, \dots, N\}$ 表示图结构中节点的集合, \mathbf{E} 表示图结构中边的集合, 其中 $\mathbf{E} = \{w_{ij}, \forall i, j \in \mathbf{V}\}$, w_{ij} 表示节点 i 和节点 j 相连接边的权重关系。 \mathbf{A} 表示图结

构的邻接矩阵。节点 i 的度 d_i 定义为所有与节点 i 相连的边的权重之和, 图结构的度矩阵是一个对角阵, 定义为 $\mathbf{D} = \text{diag}\{d_1, d_2, \dots, d_N\}$ 。图结构的拉普拉斯矩阵定义为 $\mathbf{L} = \mathbf{D} - \mathbf{A}$, 归一化的拉普拉斯矩阵 \mathcal{L} 是对称半正定矩阵, 可以分解为

$$\mathcal{L} = \mathbf{D}^{-1/2} \mathbf{L} \mathbf{D}^{-1/2} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T = \sum_{i=1}^N \lambda_i \mathbf{u}_i \mathbf{u}_i^T \quad (1)$$

其中, $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}$ 为图拉普拉斯特征值构成的对角矩阵, $\mathbf{U} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N\}$ 为图拉普拉斯特征值相对应的特征向量矩阵, \mathbf{u}_i 表示图拉普拉斯特征向量矩阵的第 i 列元素。

图 1 为明尼苏达州路网图^[12], 含有 2642 个节点和 6606 条边, 图信号定义为信号与节点之间的矢量映射 $\mathbf{f}: \mathbf{V} \rightarrow \mathbf{R}^N$, 其中 $f(i)$ 表示在节点 i 上的信号值。图傅里叶变换是以图拉普拉斯特征向量矩阵作为基元素, 将图信号从顶点域变换到图谱域, 实现对图信号的图谱域特征分析与处理。 $\hat{\mathbf{f}}(i) = \langle \mathbf{f}, \mathbf{u}_i \rangle$ 为图信号 \mathbf{f} 的图傅里叶变换^[13,14]。

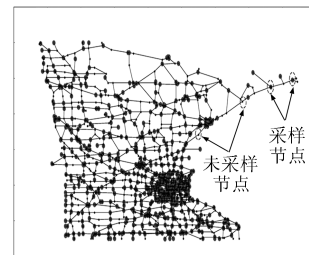


图 1 明尼苏达州路网图

2.1 带限图信号的下采样方法

Pesenson^[5]定义了带限图信号的概念。当与某个图信号相关联的图拉普拉斯特征值大于 w 时, 其相应的图谱域系数全为零, 这个图信号则被称之为带限图信号, w 被称之为图谱域截止频率。此外, 图谱域截止频率可以通过观察历史数据或者抓取快照等方式获取, 因此将其作为已知的先验信息。带限的图信号所组成空间称之为 Paley-Wiener 空间 (Paley-Wiener Space, PWS), 表示为

$$\text{PW}_w(G) = \{\mathbf{f}: \hat{\mathbf{f}}(\lambda) = 0, \lambda \geq w\} \quad (2)$$

\mathbf{S} 表示顶点域的下采样集合, 用 $\mathbf{S}^c = \mathbf{V} / \mathbf{S}$ 表示下采样集合的补集, 图信号采样过程的目标是从所有节点集合中选取代表性的节点构成下采样集合 \mathbf{S} 。关于采样集合的构成问题, 现阶段的研究主要聚焦于两种解决方法: 贪婪获取^[8,9]和局部集分割^[10,11]。贪婪获取方法采用贪婪逼近截止频率的方式选取最优的节点构成下采样集合。局部集分割方法首先将图拓扑结构划分成若干个互不相连的独立

子集, 每个子集选取一个关键节点进入下采样集合。对于局部集的划分, 在文献[10]中作者提出了基于节点度的分割方法, 主要适用于对采样集中的元素没有数量或者位置限制的情形。如图 1 所示, 采用基于节点度的分割方法构成下采样集, 采样节点用大实心圆形表示, 未采样节点用小实心圆形状表示。

2.2 带限图信号的重建方法

带限图信号的重建算法是通过分析图结构建立采样节点与未采样节点之间的关联关系, 从而利用采样节点估计出未采样节点。如图 1 所示, 重建算法的目的就是希望利用大实心圆形节点的信号得到小实心圆形状节点的信号。借鉴 Papoulis-Gerchberg 算法的思想, 在文献[8]中作者提出了一个最小均方迭代重建算法(ILSR), 交替将采样图信号投影至顶点域下采样子空间和图谱域低通滤波器子空间, 利用已知的采样图信号估计未采样图信号。在每次迭代过程中, ILSR 首先将重建信号中的下采样部分替换为已知的采样部分, 然后将组合后的图信号映射至图谱域低通滤波器子空间。令 P_T 为顶点域下采样算子, f_d 为采样信号, 则采样过程为

$$f_d = P_T f \Rightarrow f_d(\mathcal{S}) = f(\mathcal{S}) \text{ and } f_d(\mathcal{S}^c) = 0 \quad (3)$$

ILSR 的迭代重建过程可表示为

$$f_0 = \mathcal{F}^{-1} P_W \mathcal{F} f_d \quad (4)$$

$$f_{k+1}^L = \mathcal{F}^{-1} P_W \mathcal{F} \left(f_k^L + (f_d - P_T f_k^L) \right) \quad (5)$$

其中, f_k^L 表示 ILSR 第 k 次迭代重建的图信号, P_W 表示图谱域截取算子, \mathcal{F} 和 \mathcal{F}^{-1} 表示图傅里叶变换和反变换。

文献[10]提出了局部传播算子的概念以及局部集的划分方法, 其给出的局部传播重建算法(IPR)也是基于迭代凸集投影来重建图信号。与 ILSR 的重建残差处理机制不同, 在每次迭代重建过程中, IPR 将局部集内采样节点的重建残差传播至局部集内的未采样节点, 然后再将合成的信号映射到图谱域低通滤波器子空间, 其迭代重建过程可表示为

$$f_0 = P_w \left(\sum_{v \in \mathcal{S}} f(v) \delta_{N(v)} \right) \quad (6)$$

$$f_{k+1}^P = f_k^P + P_w \left(\sum_{v \in \mathcal{S}} (f(v) - f_k^P(v)) \delta_{N(v)} \right) \quad (7)$$

其中, f_k^P 表示 IPR 第 k 次迭代重建信号, $P_w(\cdot)$ 为图信号的图谱域投影算子, 局部集滤波算子用 $\delta_{N(v)} = (\delta_{N(v)}(1), \delta_{N(v)}(2), \dots, \delta_{N(v)}(N))^T$ 表示。由于在局部集内进行重建残差传播, IPR 比 ILSR 具有更快的重建收敛速度。

3 基于扩散算子的重建策略

本节建立了未采样信号的分析架构, 定义了局部均值扩散算子和全局偏置算子的概念, 构建了采样信号的重建残差局部扩散与全局扩散策略, 提出了基于扩散算子的加权重建算法。

3.1 重建残差的扩散机制框架

采样信号重建问题的实质是依据图拓扑结构建立已知的采样图信号与待定的未采样图信号之间的关联关系。鉴于原始信号与重建信号都具有图谱域带限的特性, 可知重建残差同样呈现带限的特点, 也即, 与重建残差相关联的节点之间的差异性较小。那么, 可以利用这个关键特性, 依据已知的采样节点的重建残差去估计未采样节点的重建残差, 这也是本文研究的最主要动机。首先, 依据未采样节点图信号各分量构成方式的不同, 将其分解为

$$f_{\text{acu}} = f_{\text{rec}} + \alpha f_{\text{res}} \quad (8)$$

其中, f_{acu} 表示未采样节点的图信号, f_{rec} 表示将信号投影至图谱域低通滤波器子空间后估计的未采样节点的信号分量, f_{res} 表示扩散至未采样信号的重建残差分量, α 表示未采样信号重建残差分量的控制因子, 并且 $\alpha \in \{0, 1\}$ 。基于上一节分析可知, ILSR 和 IPR 具有相同的图谱域低通滤波器子空间估计分量 f_{rec} ; 不同的是, ILSR 中的重建残差分量 f_{res} 为零, 即控制因子 $\alpha = 0$, 而 IPR 则将采样节点的重建残差直接赋予同一局部集内的未采样节点, 可知其控制因子 $\alpha = 1$ 。但是, 如果仅仅考虑局部赋值, 不仅忽略了采样节点对其局部区域外未采样节点的影响, 而且局部区域内未采样节点重建残差的完全等值也与原始信号不符, 从而使得估计信号向原始信号逼近的步伐变慢, 增加重建过程中的迭代次数。因此, 需要引入重建残差的差异化获取机制。本文从均值和偏置的角度进一步精细化分析重建残差的构成机制, 将其分解为

$$f_{\text{acu}} = f_{\text{rec}} + \alpha_m f_{\text{res}}^{\text{mean}} + \alpha_b f_{\text{res}}^{\text{bias}} \quad (9)$$

其中, $f_{\text{res}}^{\text{mean}}$ 为重建残差分量的局部均值部分, α_m 表示局部均值部分的控制因子, $\alpha_m \in \{0, 1\}$, $f_{\text{res}}^{\text{bias}}$ 为重建残差分量的全局偏置部分, α_b 表示全局偏置部分的控制因子, $\alpha_b \in \{0, 1\}$ 。局部均值部分的目标是利用带限图信号平滑特性, 将已知的采样节点重建残差作为局部区域的均值, 并扩散至与其连接的未采样节点。全局偏置部分的目标是进一步挖掘和利用图拓扑结构的先验信息, 建立单个采样节点与所有未采样节点之间的关联关系, 为各个局部区域内的未采样节点提供全局的偏置信息, 消除局部区域的

过度一致化。分量控制因子 α_m 和 α_b 的引入是为了广义化算法架构, 模块化信号分析, 强化分量部分的管理, 有助于分析所提算法与现有算法之间的关系(具体阐释在第5节)。在下文中, 将定义局部均值和全局偏置扩散算子的概念, 实现采样节点重建残差的局部与全局扩散。

3.2 局部均值扩散算子

首先, 定义 $\mathbf{f}^s = \mathbf{f}_{\text{du}} - P_T \mathbf{f}_k$ 为采样信号的重建残差。

定义1 局部均值扩散算子 对于一个给定的重建残差分量 $\mathbf{f}^s = [f_1^s \ f_2^s \ \dots \ f_N^s]^T$, 定义局部均值扩散算子 P_m 为

$$\mathbf{f}_{\text{res}}^{\text{mean}} = P_m \mathbf{f}^s = \mathbf{A}^* \mathbf{f}^s \quad (10)$$

其中, $\mathbf{f}_{\text{res}}^{\text{mean}}$ 为重建残差分量的局部均值部分, \mathbf{A}^* 中的元素为邻接矩阵 \mathbf{A} 中未采样节点元素与该行对应的度的比值。

引入局部均值扩散算子的目的是实现采样节点重建残差的局部扩散, 将采样节点的重建残差视为局部区域内的均值, 其实质是依托带限图信号的平滑性, 充分利用节点之间差异化较小的特性, 为未采样节点重建残差的估计提供有效的基础。

3.3 全局偏置扩散算子

全局偏置算子的核心目的是建立单个采样节点与所有未采样节点之间的关联关系, 实现采样节点重建残差的全局偏置扩散, 为局部区域内的未采样节点引入重建残差的差异化信息。为了有效阐述全局偏置算子的特性, 首先以单个节点元素视角分析采样节点重建残差的扩散特性, 进而以所有采样节点的全局视角总结概括。首先分析重建残差中单个元素的扩散特性, 若重建残差 \mathbf{f}^s 中除了第1个节点之外的信号值为零, 即 $\mathbf{f}^s = [f^s(1) \ 0 \ \dots \ 0]^T$, 利用图傅里叶变换将其变换至图谱域进行分析, 用 $\hat{\mathbf{f}}^s$ 表示其图谱域形式。根据上一部分的分析可知, 重建残差也具有带限特性, 则假定其在图谱域截止频率 w 内只有 m 个频率分量, 将 $\hat{\mathbf{f}}^s$ 中高于图截止频率 w 的分量设置为零。而后对 $\hat{\mathbf{f}}^s$ 进行图傅里叶反变换, 并定义 $\tilde{\mathbf{f}}^s$ 为新的顶点域表达形式, 其第 i 个元素 $\tilde{f}^s(i)$ 可以表述为

$$\begin{aligned} \tilde{f}^s(i) &= u_{i1} \hat{f}^s(1) + \dots + u_{im} \hat{f}^s(m) + 0 + \dots + 0 \\ &= u_{i1} u_{11} f^s(1) + \dots + u_{im} u_{1m} f^s(1) + 0 + \dots + 0 \\ &= (u_{i1} u_{11} + \dots + u_{im} u_{1m}) f^s(1) \\ &= \sum_{j=1}^m u_{ij} u_{1j} f^s(1) \end{aligned} \quad (11)$$

基于上述表述可知, 等式(10)建立了第1个采样节

点的重建残差 $f^s(1)$ 与所有节点信号之间的关联关系。依此类推, 依次设置采样集 \mathbf{S} 中的采样节点, 那么, 单个未采样节点的重建残差估计值应为所有采样节点扩散至该未采样节点的重建残差之和。依据单个节点元素的分析结果, 可以看出, 连接采样节点与未采样节点之间关联关系的桥梁是图谱域带限滤波, 其本质是充分挖掘所研究图信号的带限特性, 依托图拓扑结构扩散重建残差。因此, 以所有采样节点的全局视角分析其扩散过程, 可将其概括为如下形式

$$\tilde{\mathbf{f}}^s = (\mathbf{I} - P_T) \mathcal{F}^{-1} P_W \mathcal{F} \mathbf{f}^s \quad (12)$$

可以从两个角度理解全局偏置扩散过程, 从采样节点的角度来看, 单个采样节点的重建残差分量以某种方式扩散至所有未采样节点; 从未采样节点的角度来看, 单个未采样节点的全局偏置部分则是所有采样节点的重建残差分量扩散到该节点上的重建残差分量之和。

定义2 全局偏置扩散算子 对于给定的重建残差 $\mathbf{f}^s = [f_1^s \ f_2^s \ \dots \ f_N^s]^T$, 定义全局偏置扩散算子 P_b 为

$$\mathbf{f}_{\text{res}}^{\text{bias}} = P_b \mathbf{f}^s = (\mathbf{I} - P_T) \mathcal{F}^{-1} P_W \mathcal{F} \mathbf{f}^s = \mathbf{U}' \mathbf{U}'' \mathbf{f}^s \quad (13)$$

其中, $\mathbf{f}_{\text{res}}^{\text{bias}}$ 表示重建残差的全局偏置部分, m 表示图谱域截止频率 w 内频率分量的数目, \mathbf{I} 表示单位矩阵, \mathbf{U}' 表示为变形的拉普拉斯特征向量矩阵, 其采样节点所对应的行为零序列, 且其第 $N-m$ 列至第 N 列元素为零序列, \mathbf{U}'' 表示为变形的拉普拉斯特征向量矩阵, 其未采样节点所对应的行为零序列, 且其第 $N-m$ 列至第 N 列元素为零序列。

3.4 扩散算子以及加权重建算法

定义3 扩散算子 对于给定信号的重建残差 $\mathbf{f}^s = [f_1^s \ f_2^s \ \dots \ f_N^s]^T$, 定义扩散算子 P_d 为

$$\begin{aligned} \mathbf{f}_{\text{res}} &= P_d \mathbf{f}^s = (\alpha_m P_m + \alpha_b P_b) \mathbf{f}^s \\ &= (\alpha_m \mathbf{A}^* + \alpha_b \mathbf{U}' \mathbf{U}'') \mathbf{f}^s \end{aligned} \quad (14)$$

其中, P_m 表示局部均值扩散算子, P_b 表示全局偏置扩散算子, α_m 和 α_b 分别为局部均值分量和全局偏置分量的控制因子。

上述定义中, 扩散算子由局部均值扩散算子和全局偏置扩散算子构成, 扩散算子将采样信号的重建残差在图拓扑结构内扩散, 其中局部均值扩散算子将采样信号的重建残差扩散至与该节点局部区域内的未采样节点, 全局偏置扩散算子则将采样信号的重建残差扩散至所有未采样节点。基于上述扩散过程分析, 根据凸集映射的思想, 本文提出基于扩散算子的图信号迭代重建算法 (Iterative Graph diffusion operator based Weighted Signal Reconstruction, IGSR-W)。首先将初始的采样信号

投影至图谱域低通滤波子空间，获取已知采样节点的重建残差，完成重建算法的初始过程。在每次迭代过程中，利用扩散算子将采样节点的重建残差扩散至所有未采样节点，然后将扩散后的信号投影至图谱域低通滤波子空间，赋予重建信号图谱域带限特性，使其与原始信号的特性相匹配，完成重建信号的第 1 步估计；而后进一步利用局部均值与全局偏置扩散部分，并与第 1 步估计的重建信号进行加权处理，最后将加权后的信号投影至图谱域低通滤波子空间，从而完成迭代过程中重建信号的估计。上述重建过程可归纳为

$$\mathbf{f}_0 = \mathcal{F}^{-1} P_W \mathcal{F} \mathbf{f}_d \quad (15)$$

$$\mathbf{f}_{k+1}^G = \mathcal{F}^{-1} P_W \mathcal{F} (\mathbf{f}_k^G + (\mathbf{I} + P_d)(\mathbf{f}_d - P_T \mathbf{f}_k^G)) \quad (16)$$

$$\tilde{\mathbf{f}}_{k+1}^G = \mathcal{F}^{-1} P_W \mathcal{F} (\mathbf{f}_{k+1}^G + (\beta \mathbf{A}^* + (1 - \beta) \mathbf{U} \mathbf{U}^n) \cdot (\mathbf{f}_d - P_T \mathbf{f}_{k+1}^G)) \quad (17)$$

其中， \mathbf{f}_k^G 表示第 k 次迭代的重建信号， $\tilde{\mathbf{f}}_{k+1}^G$ 表示经过加权处理的重建信号， β 表示为局部均值与全局偏置部分的加权因子，并且 $\beta \in [0, 1]$ 。可以看出，由于式(16)采用了局部均值和全局偏置扩散过程，在投影至低通滤波子空间之后， \mathbf{f}_{k+1}^G 比 \mathbf{f}_k^G 拥有更接近于原始信号的 \mathbf{f}_{rec} 分量。式(17)在此基础上，进一步利用从已知的采样信号中获取的局部均值和全局偏置分量，从而更深层次挖掘利用采样信号，将有效提升重建收敛速度，快速完成原始信号的重建。

4 仿真结果

本次仿真采用美国明尼苏达州路网图^[1]作为图结构，其节点表示道路交叉口，边表示交叉口之间的道路，如图 1 所示，其节点数目为 2642。该路网图主要用于验证本文提出的 IGSR-W 算法与现有算法的重建性能，对比分析所有算法在不同图谱域截止频率上的重建误差性能，并采用 3 种典型的图结构对比分析所有算法的重建效果，验证所提出算法的鲁棒性和适用性。另外，为了对比验证加权处理的性能，在以下仿真实验中，采用 IGSR(Iterative Graph diffusion operator based Signal Reconstruction)表示未进行加权处理的重建方式，即仅利用式(16)中的过程。由于 IPR 需要以局部分割作为基础，本次仿真采用基于节点度的分割方式构成下采样集合，依次选取度最大的节点以及其邻居节点构成局部集合，每个局部集合中选取一个节点作为采样节点。下采样集合中含有 873 个节点。首先需要模拟生成仿真实验中与明尼苏达州路网图相适应的带限信号，本文主要利用图傅里叶变换、图谱域截取与图傅里叶反变换进行生成，其具体过程

如下所示：

(1)首先产生一个 2642 行 1 列的随机信号，其每一行元素对应明尼苏达州路网图上每个节点；

(2)通过图傅里叶变换将图信号变换至图谱域，设定图谱域截止频率，将高于图截止频率的分量设置为零；

(3)通过图傅里叶反变换将图谱域信号变换至顶点域。

上述步骤中，第 1 步产生仿真所需的初始信号，并使其单个元素对应于明尼苏达州路网图上的单个节点；第 2 步与第 3 步则利用图傅里叶变换，使得初始信号中元素之间的关联关系与明尼苏达州路网图上对应节点之间的边相符合，并且使该信号满足图谱域带限的特性。需要说明的是，基于目前的研究工作可知，所研究的图信号仅需要其满足图谱域带限特性，并无其他限制条件。此外，采用山东电网作为图结构，并将山东电力消耗数据作为真实世界的数据对比分析了所有算法的实际应用性能。

4.1 重建性能

本节仿真实验的目的是验证 IGSR-W 算法中加权因子对算法性能的影响，并且对比分析所提算法与现有算法的性能表现。采用 0.45 作为本次仿真的图谱域截止频率(归一化图谱域频率的变化范围为 0 至 2)。首先，鉴于加权因子的变化范围为 0 至 1，设置加权因子的参数值为 0, 0.2, 0.4, 0.6, 0.8 和 1.0，分析加权因子对 IGSR-W 算法的鲁棒性。如图 2(a)所示，纵坐标为相对误差，横坐标为迭代次数，IGSR-W 算法在不同的加权因子下都具有稳定的表现，且其加权因子为 0.6 时获得最快收敛速度。因此，将设置加权因子的参数值为 0.6，以此分析所提算法与现有算法之间的重建性能表现。重建性能的对比如图 2(b)所示，纵坐标为相对误差，横坐标为迭代次数。从图中可以看出，IGSR-W 和 IGSR 比 ILSR 和 IPR 拥有较快的信号重建收敛速度，与前文阐述的理论分析相同；得益于加权处理，IGSR-W 在 IGSR 的基础上进一步提升了算法的收敛速度。

4.2 图谱域截止频率

由于现有算法都是基于迭代收敛，交替投影至下采样子空间和低通滤波子空间。因此，图谱域截止频率也是一个非常关键的影响因素。本节是为了验证重建算法对于图谱域截止频率的鲁棒性，观察其在不同的图谱域截止频率下是否能够保持稳定。设置图谱域的截止频率的变化范围为 0.40 至 0.50。基于上一节仿真实验的结果，设置加权因子为 0.60。如图 3 所示，横坐标为图谱域截止频率，

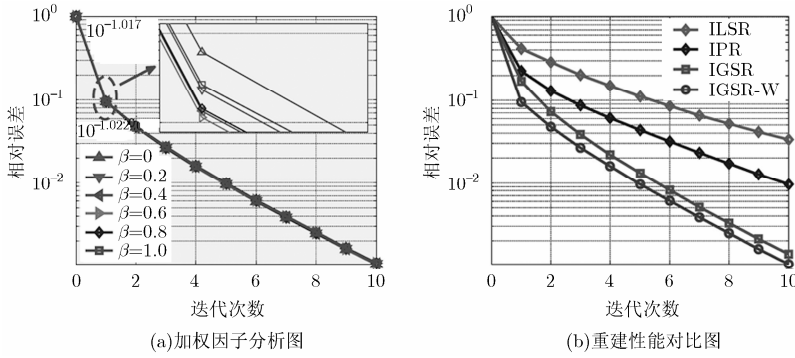


图 2 参数分析与重建收敛曲线对比图

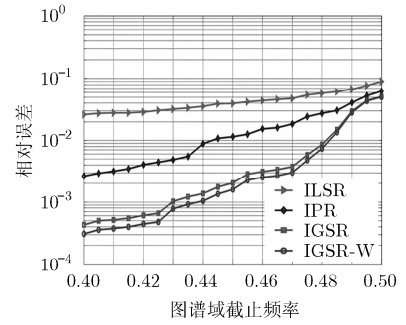


图 3 图频域截止频率变化对比图

纵坐标为相对误差，图中曲线为 ILSR, IPR, IGSR 和 IGSR-W 在不同的图谱域截止频率设定下第 10 次迭代重建的重建误差。从图中可以看出，随着图谱域截止频率的变化，所有重建算法具有相类似的增长变化，其相对误差都随着图谱域截止频率的增大而增大。由于采用了重建残差的局部均值和全局偏置扩散，IGSR 和 IGSR-W 比 ILSR 和 IPR 在第 10 次迭代重建的误差拥有较低的重建误差。由于采用了加权处理，IGSR-W 在 IGSR 的基础上提升了算法性能，并且在各个图谱域截止频率上都保持稳定。

4.3 典型图结构

现实世界中的大部分网络都可以划归为 3 种典型的图结构：Erdos-Renyi 随机图(Erdos-Renyi Graph)、小世界图(Small-world Graph)和无标度图(Scale-free Graph)。Erdos-Renyi 随机图是一种典型的随机图结构，节点之间的连接关系由某个概率确定。本次仿真生成了 40 个节点 Erdos-Renyi 随机图，节点之间的连接概率为 0.10。小世界图是介于随机图与规则图之间的典型图结构，图中大部分的节点

彼此不邻接，但是大部分节点可以从任一其他节点经过有限步数跳转到达。采用 Watts-Strogatz 模型^[15]产生 100 个节点的小世界图结构。第 3 个典型的图结构是无标度图，其最主要的特点是其节点的度分布服从幂律分布。采用 Barabasi-Albert 模型^[16]产生 200 个节点的无标度图结构。3 种典型结构中的图谱域截止频率值均为 0.50。为了有效验证所有算法重建性能，消除随机因素的影响，每种图结构的实验结果为 100 次随机图结构的平均值。首先，设置加权因子的参数值为 0, 0.2, 0.4, 0.6, 0.8 和 1.0，分析 IGSR-W 算法的不同加权因子在 3 种典型图结构下的表现。如图 4(a)所示(鉴于线条过多，仅选取部分结果展示)，纵坐标为相对误差，横坐标为迭代次数，IGSR-W 的不同加权因子在 3 种典型图结构下具有稳定的性能，且在 Erdos-Renyi 随机图、小世界图和无标度图上具有最快收敛速度的加权因子分别为 0.2, 0.2 和 0.8。因此，依次将相关的加权因子设置为 0.2, 0.2 和 0.8，分析所提算法与现有算法之间的性能。如图 4(b)所示，对比 3 种图结构，算法在小世界图结构的重建性能为最佳。对比其他重

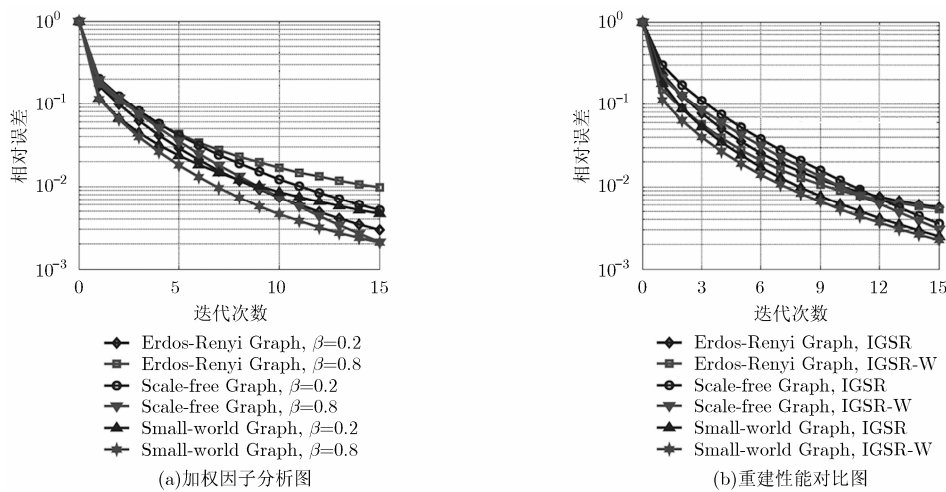


图 4 基于典型图结构的性能对比图

重建算法，IGSR-W在3种典型图结构上具有较低的迭代重建误差。由此可见，IGSR-W重建算法不拘泥于某种特定的拓扑结构，具有普适性的特点，其局部均值扩散算子与全局偏置扩散算子相辅相成，实现对采样图信号的快速与稳定的重建。

4.4 真实世界数据

采用山东电网 2015 年的电力消耗数据^[7]作为真实世界的的数据，验证所提出的重建算法与现有的重建算法。图结构如图 5(a)所示，图结构的节点为山东省 17 个地级市，节点之间的边为地级市之间的电力连接线。采用局部集合分割方式构成采样集合，其含有 7 个节点。关于真实系统中的图谱域截止频率和加权因子的选定，可以将存储的历史数据或者抓取的快照数据(snapshot data)投影至图谱域进行观测，从而获取合适的参数值。因此，本次仿真设定的图谱域截止频率为 0.55，设置加权因子为 0.2。重建性能如图 5(b)所示，横坐标为迭代次数，纵坐标为相对误差。从仿真结果中可以看出，与 ILSR 和 IPR 相比，IGSR 和 IGSR-W 具有更快的迭代收敛速度；由于引入加权处理方式，IGSR-W 比 IGSR 具有更快的收敛速度。

5 讨论与分析

本文所提算法的研究对象为带限图信号，其最主要的特点为平滑性，也即，图结构中相关节点之间的图信号变化平缓，差异性较小。所提算法的主要目的是建立采样信号与未采样信号的 3 层关联体系：第 1 层是将采样信号投影至图谱域低通滤波器空间获取的关联分量；第 2 层是将采样信号的重建残差扩散至邻近区域内未采样节点的局部均值关联分量；第 3 层是将采样信号的重建残差扩散至所有未采样节点的全局偏置关联分量。由于节点之间在邻近区域的紧密程度要大于图拓扑结构中的其它区域，关联体系的第 2 层着重突出局部性与定位性；而第 3 层则考量对原始信号的拟合与逼近，进一步

充分挖掘和利用图拓扑结构，为局部区域提供差异化信息，消除局部区域的过度一致化。所提算法的应用场景可为：重建传感器网络中受损节点的数据、估计调查问卷中未被问询的信息、以及分析推荐系统中的未知用户信息等，具体内容参见文献[1-3,10]。

当 IGSR-W 算法不进行加权处理，IGSR-W 算法退化为 IGSR 算法；当 IGSR-W 算法中不进行加权处理、局部均值扩散控制因子和全局偏置扩散控制因子为零时，也就是在重建过程中不进行采样节点重建残差的扩散，IGSR-W 算法退化为 ILSR 算法；当 IGSR-W 算法中不进行加权处理且全局偏置扩散控制因子为零时，也就是在重建过程中采样节点的重建残差值仅扩散至与其有连接关系的邻居节点，IGSR-W 算法退化为 IPR 算法。因此，IGSR-W 具有统一性与模块化的优越特点。

此外，由于引入局部均值扩散过程和全局偏置扩散过程，IGSR-W 要比 ILSR 和 IPR 消耗更多的计算资源。但是，采用集中式的数据采样方式，原始信号的重建过程通常是在一个被认为计算资源与能量资源无限的数据处理中心完成，从而弥补了 IGSR-W 所引起的不足。

6 结束语

本文针对采样的带限图信号重建问题，提出了未采样信号的构成框架，定义了局部均值扩散算子和全局偏置扩散算子的概念，提出了基于扩散算子的加权重建算法。最后采用实验方法，验证了所提出算法的有效性。之后的研究者，可以根据实际系统的需要，依托未采样信号的 3 层关联体系，设计不同的局部均值扩散算子和全局偏置扩散算子，灵活实现图信号的采样与重建过程。此外，本文采用从历史数据或抓取的数据中获取加权因子，其最优的获取方式将在后续研究中继续展开。

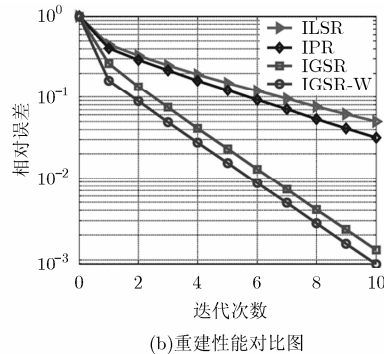
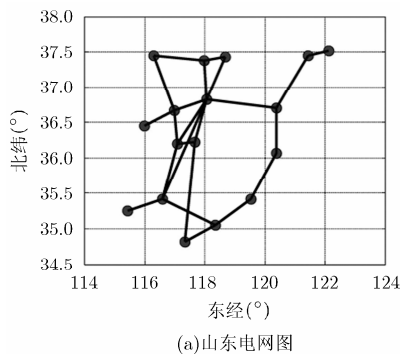


图 5 真实世界数据的性能对比图

参考文献

- [1] SHUMAN D I, NARANG S K, and FROSSARD P. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains[J]. *IEEE Signal Processing Magazine*, 2013, 30(3): 83–98. doi: 10.1109/MSP.2012.2235192.
- [2] SANDRYHAILA A and MOURA J M F. Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure[J]. *IEEE Signal Processing Magazine*, 2014, 31(5): 80–90. doi: 10.1109/MSP.2014.2329213.
- [3] CHEN Siheng, VARMA R, SANDRYHAILA A, *et al.* Discrete signal processing on graphs: Sampling theory[J]. *IEEE Transactions on Signal Processing*, 2015, 63(24): 6510–6523. doi: 10.1109/TSP.2015.2469645.
- [4] MARQUES A G, SEGARRA S, LEUS G, *et al.* Sampling of graph signals with successive local aggregations[J]. *IEEE Transactions on Signal Processing*, 2016, 64(7): 1832–1843. doi: 10.1109/TSP.2015.2507546.
- [5] PESENSON I. Sampling in Paley-Wiener spaces on combinatorial graphs[J]. *Transactions of the American Mathematical Society*, 2008, 360(10): 5603–5627. doi: 10.1090/S0002-9947-08-04511-X.
- [6] PAPOULIS A. A new algorithm in spectral analysis and band-limited extrapolation[J]. *IEEE Transactions on Circuits and Systems*, 1975, 22(9): 735–742. doi: 10.1109/TCS.1975.1084118.
- [7] GERCHBERG R W. Super-resolution through error energy reduction[J]. *Journal of Modern Optics*, 1974, 21(9): 709–720. doi: 10.1080/713818946.
- [8] NARANG S K, GADDE A, and SANOU E. Localized iterative methods for interpolation in graph structured data[C]. *IEEE Global Conference on Signal and Information Processing*, Texas, USA, 2013: 491–494. doi: 10.1109/GlobalSIP.2013.6736922.
- [9] ANIS A, GADDE A, and ORTEGA A. Towards a sampling theorem for signals on arbitrary graphs[C]. *IEEE International Conference on Acoustics, Speech and Signal Processing*, Florence, Italy, 2014: 3864–3868. doi: 10.1109/ICASSP.2014.6854325.
- [10] WANG X, LIU P, and GU Y. Local-set-based graph signal reconstruction[J]. *IEEE Transactions on Signal Processing*, 2015, 63(9): 2432–2444. doi: 10.1109/TSP.2015.2411217.
- [11] WANG X, CHEN J, and GU Y. Generalized graph signal sampling and reconstruction[C]. *IEEE Global Conference on Signal and Information Processing*, Orlando, USA, 2015: 567–571. doi: 10.1109/GlobalSIP.2015.7418259.
- [12] GLEICH D. The MatlabBGL Matlab Library[OL]. <http://www.cs.purdue.edu/homes/dgleich/packages/matlab-bgl/index.html>.2016.9.
- [13] SANDRYHAILA A and MOURA J. Discrete signal processing on graphs: Frequency analysis[J]. *IEEE Transactions on Signal Processing*, 2014, 62(12): 3042–3054. doi: 10.1109/TSP.2014.2321121.
- [14] SANDRYHAILA A and MOURA J. Discrete signal processing on graphs: graph Fourier transform[C]. *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, Canada, 2013: 6167–6170. doi: 10.1109/ICASSP.2013.6638850.
- [15] WATTS D J and STROGATZ S H. Collective dynamics of ‘small-world’ networks[J]. *Nature*, 1998, 393(6684): 440–442. doi: 10.1038/30918.
- [16] BARABASI A L and ALBERT R. Emergence of scaling in random networks[J]. *Science*, 1999, 286(5439): 509–512. doi: 10.1126/science.286.5439.509.
- [17] SHANDONG BUREAU OF STATISTICS. Shandong Statistical Yearbook (2015)[OL]. <http://www.stats-sd.gov.cn.2016.8>.
- 杨立山: 男, 1986年生, 博士生, 研究方向为图信号处理与压缩感知、大数据处理。
- 游康勇: 男, 1993年生, 博士生, 研究方向为压缩感知技术。
- 郭文彬: 男, 1971年生, 副教授, 研究方向为信号处理、认知无线电及其关键技术。