

저 복잡도를 갖는 효율적인 그래프 신호의 샘플링 알고리즘

김윤학*

Efficient Sampling of Graph Signals with Reduced Complexity

Yoon Hak Kim*

요약

그래프 노드상에서 발생하는 그래프 신호의 일부를 선택해서 만든 샘플링 신호로부터 원신호를 복원하기 위해, 복원오차를 최소화하기 위한 최적의 샘플링 집합을 선택하는 알고리즘에 관해 연구한다. 복잡도 개선을 위해 복원오차를 직접적으로 최소화하는 대신에, 복원오차의 상한값을 비용함수로 사용하고, QR분해 적용을 통해 발생하는 삼각행렬의 대각선상에 위치하는 값으로 샘플링을 결정할 수 있게 하는, 저 복잡도를 갖는 반복적 탐욕알고리즘을 제안한다. 기존의 샘플링 선택 방법과 비교하여, 제안 알고리즘이 복원 성능 저하를 평균 5%미만으로 유지하면서, 약 3.5배 빠른 실행시간을 보임을 다양한 그래프 상황에서 실험을 통해 확인한다.

ABSTRACT

A sampling set selection algorithm is proposed to reconstruct original graph signals from the sampled signals generated on the nodes in the sampling set. Instead of directly minimizing the reconstruction error, we focus on minimizing the upper bound on the reconstruction error to reduce the algorithm complexity. The metric is manipulated by using QR factorization to produce the upper triangular matrix and the analytic result is presented to enable a greedy selection of the next nodes at iterations by using the diagonal entries of the upper triangular matrix, leading to an efficient sampling process with reduced complexity. We run experiments for various graphs to demonstrate a competitive reconstruction performance of the proposed algorithm while offering the execution time about 3.5 times faster than one of the previous selection methods.

키워드

Graph Signal Processing, Sampling Set Selection, Greedy Algorithm, QR Factorization
그래프 신호 처리, 샘플링 집합 선택, 탐욕 알고리즘, QR 분해

1. 서론

신경망, 전력망, 교통망, 사회망 그리고 센서망과 같이 현재 급부상하고 있는 응용분야에서 발생하는 데이터는 비정형적이고 고차원적인 특성을 보이는 네

트워크 데이터로, 이는 기존의 데이터와는 다른 특성을 보이며, 이러한 네트워크 데이터를 처리하기 위해 그래프 신호처리 분야가 학계에 뜨거운 관심을 받고 있다[1,2].

그래프 신호처리의 중요 기술 중에서 샘플링집합

* 교신저자 : 조선대학교 전자공학부
• 접수일 : 2022. 03. 07
• 수정완료일 : 2022. 03. 27
• 게재확정일 : 2022. 04. 17

• Received : Mar. 07, 2022, Revised : Mar. 27, 2022, Accepted : Apr. 17, 2022
• Corresponding Author : Yoon-Hak Kim
Dept. Electronic Engineering, Chosun University,
Email : yhk@chosun.ac.kr

선택 기술은 많은 양의 네트워크 데이터를 처리하는 대신에, 최적의 부분 데이터를 선택하는 방식으로, 원 신호와의 복원오차를 최소화하면서 복원하기 위한 다양한 기술들이 제안되어 왔다[3-11]. 복원오차의 상한값을 최소화하기 위한 탐욕적(greedy) 반복 알고리즘이 제안되었고[3-5], 지역 불확실성원리(local uncertainty principle)에 기반하여 샘플링하는 방식[6], 그리고 계산량을 줄이기 위해, 샘플링하는 과정에서 필요한 고유값 분해(eigen-decomposition)를 사용하지 않는 샘플링 기술이 개발되었다[7,8]. 그래프 신호의 변화량에 기반한 비용함수를 도입하여 샘플링 속도를 높이는 기술이 또한 제안되었고[9], QR분해를 사용한 효율적이고 성능이 우수한 탐욕적 샘플링 알고리즘이 개발된 바 있다[10]. 최근에는 신호값이 가중치를 갖는 경우 기존의 샘플링방식을 어떻게 바꿔야 하는지에 대한 연구가 수행 되었다[11].

본 논문에서는 한번에 한 개의 노드를 샘플링하는 탐욕적(greedy) 반복 알고리즘을 제안하며, 샘플링 과정의 계산량을 줄이기 위해 복원오차를 직접적으로 최소화하는 대신에, 복원오차의 상한값을 비용함수로 취하고, 이를 최소화하기 위한 수학적 결과식을 유도한다. 기존의 상한값 최소화 방식과 다른 방법으로, 본 연구에서는, 비용함수에 QR분해를 적용하여 발생하는 상삼각행렬(upper triangular matrix)의 대각선에 위치하는 값을 사용하여 매 반복단계에서 노드를 선택함으로써 비용함수를 최소화할 수 있음을 보인다. 이는 상당한 계산량을 감소시킬 수 있으며, 기존의 샘플링 방식과 실행시간을 비교하여 복잡도면에서 성능이 우수함을 보인다. 또한, 복원오차에 대한 성능평가를 위해 다양한 그래프 환경에서 서로 다른 기존의 샘플링 기술을 비교 평가함으로써, 본 논문에서 제안하는 샘플링 기술의 우수성을 입증한다.

본 논문은 2장에서 비용함수를 제안하고, 문제를 정립하며, 3장에서 문제를 해결하기 위한 샘플링 알고리즘을 제안한다. 4장에서 알고리즘의 성능평가가 수행되고, 5장의 결론으로 마무리한다.

II. 비용함수 및 문제정립

본 논문에서는 N 개의 노드(node or vertex)로 구성

되고, 무방향(undirected)이면서 엣지로 표현되는 가중치를 갖는 그래프 $G(V,E)$ 를 다루며, 여기서 $V=\{1, \dots, N\}$ 와 $E=\{i,j,e_{ij}\}$ 는 각각 N 개의 노드로 구성된 전체집합과 i 번째 노드와 j 번째 노드를 연결하는 가중치 e_{ij} 의 집합을 의미한다. 그래프상에서 각 노드에서 발생하는 값으로 이루어진 그래프 신호 $\mathbf{f}=[f_1 \dots f_N]^T \in \mathbf{R}^N$ 를 정의할 수 있으며, f_i 는 i 번째 노드에서 발생한 신호값을 표시한다. 그래프의 비정형적 특성과 신호의 변화를 나타내기 위해 그래프 라플라시안 L (combinatorial graph Laplacian 또는 normalized Laplacian)등이 사용되는데[2,4], 여기서, 그래프 라플라시안 L 은 $N \times N$ 행렬이고 고유값(eigenvalue) $|\lambda_1| \leq \dots \leq |\lambda_N|$ 와 정규 직교 고유벡터(orthonormal eigenvector) $\mathbf{u}_1 \dots \mathbf{u}_N$ 를 갖는다고 가정한다. 이때, 그래프신호 \mathbf{f} 는 $\mathbf{f}=\mathbf{U}\mathbf{c}$ 로 표현될 수 있다. 여기서 $\mathbf{U}=[\mathbf{u}_1 \dots \mathbf{u}_N]$ 는 고유벡터행렬을 의미하고, \mathbf{c} 는 그래프 푸리에 변환(graph Fourier transform, GFT)을 나타내며, 수식적으로 $\mathbf{c}=\mathbf{U}^{-1}\mathbf{f}=\mathbf{U}^T\mathbf{f}$ 로 표현된다. 그래프 신호가 ω -대역폭 제한 신호라면, 즉, GFT의 구성요소가 $c_j=0, |\lambda_j|>\omega, \forall j>r$ 라면, \mathbf{f} 는 다음과 같이 표현가능하다.

$$\mathbf{f}=\sum_{i=1}^r c_i \mathbf{u}_i = \mathbf{U}_{VR} \mathbf{c}_R \quad (1)$$

여기서, \mathbf{c}_R 은 $r \times 1$ 열벡터(column vector) $[c_1 \dots c_r]^T$ 를 의미하며, 아랫첨자 $R=\{1, \dots, r\}$ 은 각 구성요소의 인덱스(index)를 표현한다. 마찬가지로, \mathbf{U}_{VR} 은 고유벡터행렬 \mathbf{U} 의 부분행렬로 각각의 행과열의 인덱스는 아랫첨자인 V 와 R 로 표현되며, 본 논문에서는 같은 방식으로 아랫첨자를 사용한다.

ω -대역폭 제한 그래프 신호 공간상에서 무 잡음 신호로부터 샘플링된 신호를 통해 원 신호를 오차없이 복원할 수 있는 샘플링 집합 S 를 고유 집합(uniqueness set)으로 정의할 수 있는데, 여기서 샘플링된 신호는 $\mathbf{f}_S=[f_1 \dots f_{|S|}]^T, f_j, j \in S$ 로 나타내고, 고유 집합내의 각각의 노드는 \mathbf{U}_{VR} 에서 독립적인 행 벡터를 선택함으로써 결정할 수 있으며, 이때, j 번째 행 벡터의 선택은 j 번째 노드를 선택했음을 의미한다

[4]. 따라서, U_{VR} 로부터 샘플링된 $|S|$ 개의 독립적인 행벡터로 구성된 $|S| \times r$ 행렬 U_{SR} 을 만들 수 있으며, 이를 통해 샘플링 신호를 얻을 수 있다. 즉, $f_S = U_{SR}c_R$ 이 된다. 이때, c_R 를 구하기 위해 최소 제곱추정 방법(least square estimate) 사용하면 $\hat{c}_R = (U_{SR}^T U_{SR})^{-1} U_{SR}^T f_S$ 로 나타낼 수 있으며, 이로부터 복원된 그래프신호 \hat{f} 를 다음과 같이 구할 수 있다:

$$\hat{f} = U_{VR} \hat{c}_R = U_{VR} U_{SR}^+ f_S \quad (2)$$

여기서 $U_{SR}^+ = (U_{SR}^T U_{SR})^{-1} U_{SR}^T$ 은 U_{SR} 의 유사역행렬(pseudo-inverse)를 의미하며, 샘플링집합이 고유집합일때, 같은 의미로 U_{SR} 의 랭크(rank)가 r 일때, 무잡음 ω -대역폭 제한신호를 가정하면 $\hat{f} = f$ 가 되어 오차없이 신호를 복원할 수 있게 된다.

본 논문에서는, 그래프 신호가 ω -대역폭 제한신호라고 가정하고, 샘플링 신호로부터 원신호를 복원하는에 있어, 복원오차를 최소화하는 대신에, 복원오차의 상한값(upper bound)을 최소화하는 방식을 택한다. 이는 직접적으로 복원오차를 최소화하는 방식에 비해 상당한 계산량을 줄일 수 있다. 수식적으로 표현하면, 그래프신호 $f_i, i \in V$ 가 가산성(additive)이고 서로 독립적인, 평균값이 0이고 분산이 1인 정규화된 잡음 n 에 의해 오염된다고 가정했을 때 오차 e 벡터는 다음과 같다:

$$e = \hat{f} - f = U_{VR} U_{SR}^+ (f_S + n_S) - f = U_{VR} U_{SR}^+ n_S \quad (3)$$

여기서 n_S 는 $|S| \times 1$ 열벡터를 나타낸다.

본 논문에서는 독립적인 r 개의 행벡터를 선택해서 샘플링집합을 구하게 되고, 즉, $|S|=r$ 이 되어 $U_{SR}^+ = U_{SR}^{-1}$ 이 성립하며, 이 경우에 복원오차(reconstruction mean squared error, MSE)는 다음과 같이 구할 수 있다.

$$MSE = tr[E e e^T] \quad (4)$$

$$= tr[U_{VR} U_{SR}^{-1} E n_S n_S^T (U_{SR}^{-1})^T U_{VR}^T] \quad (5)$$

$$= tr[(U_{SR}^T U_{SR})^{-1}] \quad (6)$$

여기서 (6)은 $E n_S n_S^T = I$ 와 U_{VR} 이 정규직교 고유벡터를 가지므로 성립하게 된다. (6)번식의 복원오차는 행렬노름(matrix norm)으로 표현하면 다음과 같다[12].

$$tr[(U_{SR}^T U_{SR})^{-1}] = \|U_{SR}^{-1}\|_F^2 \leq r \|U_{SR}^{-1}\|_2^2 \quad (7)$$

여기서, $\|\cdot\|_F^2$ 와 $\|\cdot\|_2^2$ 는 각각 Frobenius 행렬노름과 스펙트럴(spectral) 행렬노름을 의미한다. 그리고, 스펙트럴 행렬노름은 최대 특이값(singular value)의 제곱 σ_{\max}^2 으로 표현될 수 있으므로[12], 복원오차는 다음과 같이 전개할 수 있다.

$$MSE = tr[(U_{SR}^T U_{SR})^{-1}] \leq r \sigma_{\max}^2 (U_{SR}^{-1}) \quad (8)$$

결국, 본 논문에서는 최적의 샘플링 집합 S^* 를 구하기 위해, 복원오차의 상한값(upper bound)인 $\sigma_{\max}^2 (U_{SR}^{-1})$ 를 최소화하는 전략을 취함으로써, 알고리즘의 복잡도를 개선하게 된다. 이를 수식적으로 표현하면 아래와 같다.

$$S^* = \operatorname{argmin}_S \sigma_{\max}^2 (U_{SR}^{-1}) = \operatorname{argmax}_S \sigma_{\min}^2 (U_{SR}) \quad (9)$$

즉, 역행렬 U_{SR}^{-1} 을 구하기 위한 복잡한 연산 과정을 피하고 U_{SR} 의 최소 특이값의 제곱을 최대로 하는 r 개의 독립적인 행벡터를 U_{VR} 로부터 샘플링하는 과정을 통해 최적의 샘플링 집합 S^* 를 얻을 수 있게 된다.

또한, 전체 N 개의 노드에서 r 개의 독립적인 행벡터를 샘플링하는 과정도 N 이 큰 경우 막대한 계산량을 필요로 하므로, 본 논문에서는 실용적인 샘플링 방식인 탐욕적(greedy) 샘플링 방식을 제안한다. 즉, 한번에 한 개의 노드를 선택하며 r 개의 노드를 선택할때까지 반복적으로 수행하게 된다. 이를 수식적으로 표

현하면 다음과 같으며. 여기서 S_i 는 i 번째 단계까지 샘플링된 i 개의 노드들로 구성된 집합을 나타내고, $S_i^C \equiv (V - S_i)$ 는 전체 N 개에서 i 개의 노드를 선택하고 남은 노드들의 집합을 표현한다.

$$j^* = \arg \max_{S_{i+1} = S_i + \{j\}, j \in S_i^C} \sigma_{\min}^2(\mathbf{U}_{S_{i+1}, \mathbf{R}}) \quad (10)$$

$$S_{i+1}^* = S_i + \{j^*\}, i = 0, \dots, |S| - 1. \quad (11)$$

여기서 $\mathbf{U}_{S_{i+1}, \mathbf{R}}$ 은 $(i+1)$ 번째 단계에서 $(i+1)$ 개의 독립적인 행벡터로 구성된 행렬을 의미한다:

$$\mathbf{U}_{S_{i+1}, \mathbf{R}}^T = [\mathbf{u}^{(1)} \dots \mathbf{u}^{(i+1)}] \quad (12)$$

수식 (12)에서 $(\mathbf{u}^{(i)})^T$ 은 i 번째 단계에서 선택한 독립적인 행벡터를 표시한다. 결국, 제안하는 탐욕알고리즘은 (10)과 (11)을 r 번 반복적으로 수행하면서 r 개의 샘플링 집합 S^* 을 구하게 된다.

III. 샘플링 집합 선택 알고리즘

본 논문에서는 수식 (12)에 QR 분해를 적용하여 수식을 전개한다. 즉, $\mathbf{U}_{S_{i+1}, \mathbf{R}}^T = \mathbf{Q}^{i+1} \mathbf{R}^{i+1}$ 를 얻게 되고, 여기서 \mathbf{Q}^{i+1} 은 $r \times (i+1)$ 행렬이고 $i+1$ 개의 정규직교 열벡터 $\mathbf{q}_1, \dots, \mathbf{q}_{i+1}$ 로 구성되며, \mathbf{R}^{i+1} 은 $(i+1) \times (i+1)$ 상삼각 정방행렬(upper triangular square matrix)이다. 이를 이용하여 $(i+1)$ 번째 단계에서의 복원오차를 구하면 다음과 같다.

$$\begin{aligned} MSE_{i+1} &= \text{tr} \left[\left(\mathbf{Q}^{i+1} \mathbf{R}^{i+1} (\mathbf{R}^{i+1})^T (\mathbf{Q}^{i+1})^T \right)^+ \right] \\ &= \text{tr} \left[\mathbf{Q}^{i+1} (\mathbf{R}^{i+1} (\mathbf{R}^{i+1})^T)^{-1} (\mathbf{Q}^{i+1})^T \right] \end{aligned} \quad (13)$$

$$= \text{tr} \left[(\mathbf{R}^{i+1} (\mathbf{R}^{i+1})^T)^{-1} \right] \quad (14)$$

여기서 수식 (13)은 $((\mathbf{Q}^{i+1})^T)^+ = \mathbf{Q}^{i+1}$ 와 $(\mathbf{Q}^{i+1})^+ = (\mathbf{Q}^{i+1})^T$ 이 성립함으로 얻을 수 있고, 수식 (14)는 \mathbf{Q}^{i+1} 가 정규직교 열벡터를 가지므로 성립한다. 따라서, 본 논문에서 사용하는 비용함수는 복원오차의 상한값인 수식 (9)를 사용하여 $(i+1)$ 번째 단계에

서의 최적의 샘플링 집합을 아래와 같이 구할 수 있게 된다.

$$S_{i+1}^* = \arg \max_S \sigma_{\min}^2(\mathbf{R}^{i+1}) \quad (15)$$

그리고, QR분해 특성상 \mathbf{R}^{i+1} 은 이전단계의 \mathbf{R}^i 로부터 마지막에 열벡터를 첨가해서 구할 수 있다. 즉,

$$\mathbf{R}^{i+1} = \begin{bmatrix} \mathbf{R}^i & (\mathbf{Q}^i)^T \mathbf{u}^{(i+1)} \\ \mathbf{0}_{1 \times i} & \mathbf{q}_{i+1}^T \mathbf{u}^{(i+1)} \end{bmatrix} \quad (16)$$

여기서 $\mathbf{0}_{1 \times i}$ 은 모든 엔트리값이 0인 $1 \times i$ 행렬이다. 또한, 정방행렬 특이값의 제곱은 고유값의 절대치와 같으므로, $(i+1)$ 번째 단계에서 최적의 노드 즉 행벡터를 고르기 위해서는 최소 고유값의 절대치 중에 최대값을 갖는 행벡터를 선택하게 된다. 그리고, 삼각 정방행렬의 고유값은 행렬의 대각선에 위치하게 되고, 주의할 점은 고유집합을 구성하기 위해서 선택된 행벡터들이 독립적이어야 하므로, 고유값들이 0이 되지 않는 행벡터 중에서 선택하게 된다. 이를 수식적으로 나타내면 다음과 같다.

$$\begin{aligned} \mathbf{u}^{(i+1)*} &= \\ \arg \max_{\mathbf{u}^{(i+1)}, \mathbf{q}_{i+1}^T \mathbf{u}^{(i+1)} \neq 0} \min & \left[|\mathbf{q}_1^T \mathbf{u}^{(1)*}| \dots |\mathbf{q}_{i+1}^T \mathbf{u}^{(i+1)*}| \right] \end{aligned} \quad (17)$$

여기서 $\mathbf{u}^{(i+1)*}$ 는 $(i+1)$ 번째 단계에서 선택된 행벡터를 의미한다. 그리고, 매 단계에서 이전에 선택된 고유값 절대치보다 작은 고유값 절대치를 선택하게 함으로써, 현 단계에서의 고유값 절대치가 최소가 되어, 수식 (17)번식에서 $i+1$ 개의 고유값 절대치 중 최소를 고르는 과정 없이 최대 고유값 절대치를 취함으로써 계산량을 추가적으로 줄이게 된다. 즉,

$$\mathbf{u}^{(i+1)*} = \arg \max_{\mathbf{u}^{(i+1)}, \mathbf{q}_{i+1}^T \mathbf{u}^{(i+1)} \neq 0} |\mathbf{q}_{i+1}^T \mathbf{u}^{(i+1)}| \quad (18)$$

여기서, $|\mathbf{q}_i^T \mathbf{u}^{(i)*}| \geq |\mathbf{q}_{i+1}^T \mathbf{u}^{(i+1)*}|$ 을 만족한다.

본 논문에서 제안한 샘플링 집합 선택 알고리즘을 요약하면 다음과 같다.

단계 1: 샘플링 집합을 $S_0 = \emptyset$, $i = 0$ 으로 초기화

한다.

단계 2: 각각의 $\mathbf{u}^{(i+1)}$ 에 대해,

- 1) 수식 (18)를 통해 $\mathbf{u}^{(i+1)*}$ 를 구한다.
- 2) 이렇게 구한 $\mathbf{u}^{(i+1)*}$ 이 j^* 번째 행벡터 일때 $S_{i+1} = S_i + \{j^*\}$ 가 된다.
- 3) $\mathbf{u}^{(i+1)*}$ 를 구한 후, \mathbf{Q}^{i+1} , \mathbf{R}^{i+1} 를 QR분해를 통해 구한다.

단계 3: $i = i + 1$, $S = S_i$ 이 될 때까지 **단계 2**를 반복한다.

IV. 실험 및 분석

알고리즘의 성능평가를 위해 다음 3개의 그래프를 사용하여 기존의 샘플링 알고리즘과 비교한다.

- 1) 랜덤센서 그래프(Random sensor graph, RSG)
- 2) 랜덤정형 그래프(Random regular graph, RRG)
- 3) 미네소타 도로망 그래프(Minnesota road network graph, MRNG)

미네소타 그래프인 경우 총 노드수 $N=2642$ 이고, 다른 그래프인 경우 $N=1000$ 로 실험이 진행된다. 또한, 본 실험에서 RRG인 경우 각각의 노드가 6개의 주변 노드에 규칙적으로 연결되어 발생된다. 각각의 그래프를 총 30개를 발생하고 그래프 라플라시안 \mathbf{L} (combinatorial Laplacian)를 사용하여, 고유벡터 행렬 \mathbf{U} 및 GFT \mathbf{c} 를 얻게 된다. 본 실험에서는 고유집합을 구축하여 평가가 진행되며, 고유집합의 크기 $r=|S|$ 는 30에서 80까지 변화시켜 검증한다. 비교평가가 이루어질 샘플링 기술은 가우시안 제거기법에 기반한 효율적인 방식인 ESM(efficient sampling method[4], QR분해에 기반한 성능이 우수한 QRM(QR factorization-based method[10]), 그리고 본 논문에서 제안한 방식(Proposed method)이다. 대역폭 제한 그래프 신호를 가정하여 알고리즘이 제안되었지만, 실제 환경에서 발생하는 신호는 대역폭 제한신호가 아니므로 본 평가에서는 대역폭 제한이 아닌 임의의 그래프 신호를 사용하여 수행된다. 그래프 신호는 가우시안 분포를 사용하여 아래와 같이 발생할 수 있다:

$$p(\mathbf{f}) \propto \exp(-\mathbf{f}^T \mathbf{K}^{-1} \mathbf{f}) = \exp(-\mathbf{f}^T (\mathbf{L} + \delta \mathbf{I}) \mathbf{f}) \quad (19)$$

여기서, $\mathbf{L} + \delta \mathbf{I}$ 은 역공분산행렬(inverse covariance matrix)로 공분산행렬이 존재해야 하므로, δ 는 작은값 ($=0.01$)을 취하게 된다. 이렇게 발생된 그래프 신호는 가산성 정규분포 $\mathcal{N}(0, \sigma^2)$, $\sigma = 0.1$ 를 갖는 잡음에 의해 오염되며, 성능평가를 위해 100개의 오염된 그래프 신호를 발생하여 평균 복원오차를 구하게 된다. 그래프 발생 및 관련 속성값 (\mathbf{L} , \mathbf{U} , e_{ij})은 GSP 툴박스 [13]를 사용하여 평가가 진행되었다.

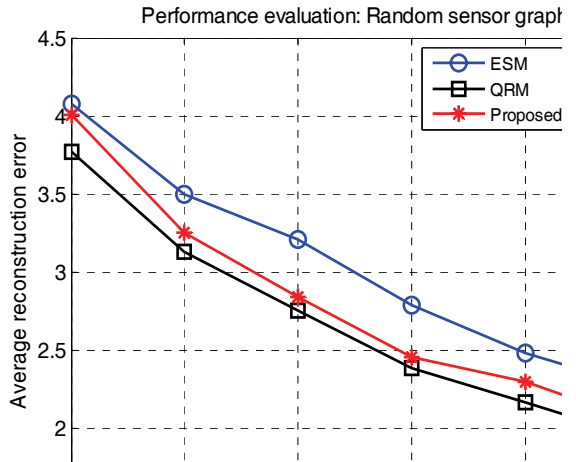


그림 1. RSG에서 샘플링 집합 크기 변화에 따른 샘플링기술의 성능평가

Fig. 1 Evaluation of reconstruction performance for RSG by varying sample size

3개의 그래프 RSG, RRG, MRNG 상에서 각각 발생된 그래프신호에 3개의 다른 샘플링 방법을 적용하여 고유 집합을 만들고 이로부터 샘플링된 신호를 얻은 후 수식 (2)를 통해 원신호를 복원하여 평균복원오차를 구했으며, 이를 그림 1, 2, 3에 각각 나타내었다. 예상한대로, ESM은 간단한 샘플링기법으로 빠른 샘플링을 수행하지만 저조한 복원성능을 보이고, QRM은 복잡도가 높은 대신, 우수한 복원성능을 보임을 알 수 있다. 본 논문에서 제안하는 알고리즘은 복잡도를 감소시키면서 만족할만한 복원성능을 보이며, QRM과 비교시, 평균 5%미만의 성능저하를 보임을 확인할 수 있다.

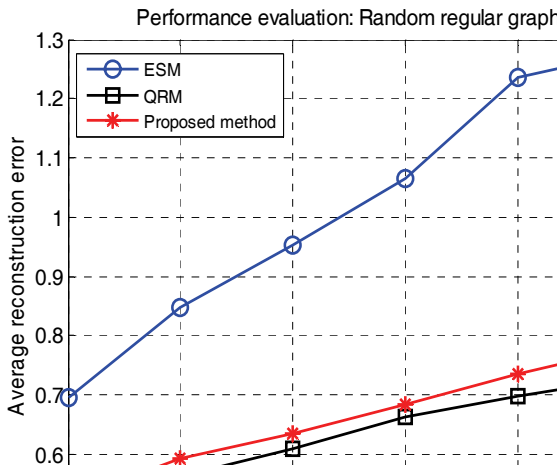


그림 2. RRG에서 샘플링 집합크기 변화에 따른 샘플링기술의 성능평가

Fig. 2 Evaluation of reconstruction performance for RRG by varying sample size

요구하는 실용적 응용 분야에 적합한 샘플링 기법이 라 할 수 있다.

V. 결론

그래프상에서 발생하는 신호를 샘플링하는 방법에 대해 연구가 수행되었으며, 복잡도 개선을 위해 복원 오차를 직접적으로 최소화하는 대신에, 복원오차의 상한값을 비용함수로 채택하고, 이를 최소화하기 위한 수학적 결과식을 유도하였다.

다양한 그래프 상황에서 복잡도 및 복원성능평가가 수행되었으며, 기존 기술대비 복잡도를 크게 감소하면서, 경쟁력있는 복원성능을 보임을 확인하였다.

향후 실용적 응용분야에 적용가능하리라 판단되며, 다양한 응용분야에 맞는 샘플링 기술 개발에 대한 연구를 지속적으로 수행할 계획이다.

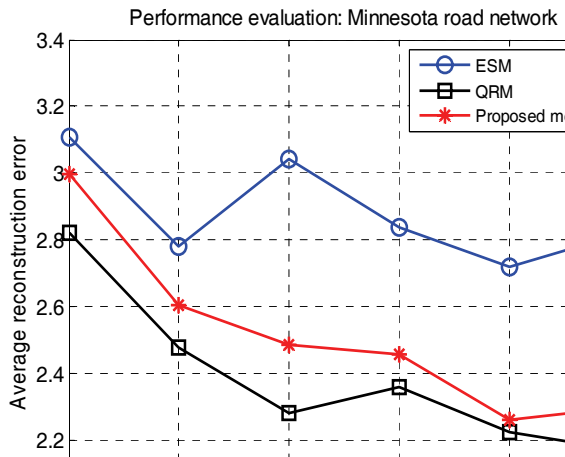


그림 3. MRNG에서 샘플링 집합크기 변화에 따른 샘플링기술의 성능평가

Fig.3 Evaluation of reconstruction performance for MRNG by varying sample size

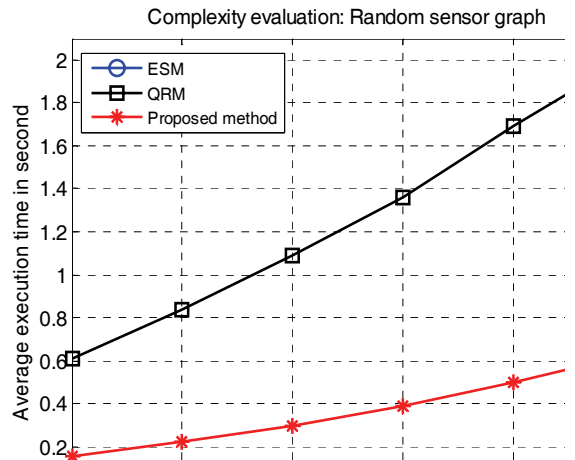


그림 4. RSG에서 샘플링 집합크기 변화에 따른 샘플링기술의 복잡도 평가

Fig. 4 Complexity evaluation for RSG by varying sample size

복잡도를 평가하기 위해 RSG상에서 서로 다른 샘플링 기술의 실행 시간을 비교 하였으며, 이를 그림 4에 나타내었다. 제안 알고리즘은 복원성능을 유지하면서 (그림 1, 2, 3 참조), QRM보다 평균 약 3.5배 빠른 샘플링속도를 보임을 확인할 수 있으며, 이는 고성능을

감사의 글

이 논문은 2021학년도 조선대학교 학술연구비의 지원을 받아 연구되었음.

References

- [1] D. Shuman, S. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, vol. 30, no. 3, 2013, pp. 83-98.
- [2] A. Ortega, P. Frossard, J. Kovacevic, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: overview, challenges and applications," *Proceedings of the IEEE*, vol. 106, no. 5, 2018, pp. 808-828.
- [3] A. Anis, A. Gadde, and A. Ortega, "Towards a sampling theorem for signals on arbitrary graphs," *IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP)*, Florence, Italy, 2014, pp. 3864-3858.
- [4] A. Anis, A. Gadde, and A. Ortega, "Efficient sampling set selection for bandlimited graph signals using graph spectral proxies," *IEEE Transactions on Signal Processing*, vol. 64, no. 14, 2016, pp. 3775-3789.
- [5] S. Chen, R. Varma, A. Sandryhaila, and J. Kovacevic, "Discrete signal processing on graphs: sampling theory," *IEEE Transactions on Signal Processing*, vol. 63, no. 24, 2015, pp. 6510-6523.
- [6] N. Perraudin, B. Ricaud, D. Shuman, and P. Vandergheynst, "Global and local uncertainty principles for signals on graphs," *APSIPA Transactions on Signal and Information Processing*, vol. 7, 2018, pp. 1-26.
- [7] A. Sakiyama, Y. Tanaka, T. Tanaka, and A. Ortega, "Eigendecomposition-free sampling set selection for graph signals," *IEEE Transactions on Signal Processing*, vol. 67, no. 10, 2019, pp. 2679-2692.
- [8] F. Wang, G. Cheung, and Y. Wang, "Low-complexity graph sampling with noise and signal reconstruction via Neumann series," *IEEE Transactions on Signal Processing*, vol. 67, no. 21, 2019, pp. 5511-5526.
- [9] Y. Kim, "Fast sampling set selection algorithm for arbitrary graph signals," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 15, no. 6, Dec. 2020, pp. 1023-1030.
- [10] Y. Kim, "QR factorization-based sampling set selection for bandlimited graph signals," *Signal Processing*, vol. 179, 2021, pp. 1-10.
- [11] Y. Kim, "Sampling set selection algorithm for weighted graph signals," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 17, no. 1, Feb. 2022, pp. 1023-1030.
- [12] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.
- [13] N. Perraudin, J. Paratte, D. Shuman, L. Martin, V. Kalofolias, P. Vandergheynst, and D. K. Hammond, "GSPBOX: A toolbox for signal processing on graphs," arXiv:1408.5781, 2014.

BIOGRAPHY

**Yoon-Hak Kim**

1992: BS degree in Electronic Engineering, Yonsei University
 1994: MS degree in Electronic Engineering, Yonsei University

2007: Ph.D. degree in Electrical Engineering, University of Southern California.

2012 - Present: Associate professor, Chosun University

