

QR 분해에 기반한 저 복잡도 센서 선택 알고리즘

김윤학*

Low-complexity Sensor Selection Based on QR factorization

Yoon Hak Kim*

*Associate Professor, Department of Electronic Engineering, Chosun University, Gwangju, 61452 Korea

요 약

센서 네트워크에서 평균 추정 성능을 높이기 위한 저 복잡도를 갖는 센서 노드 선택 알고리즘에 대해 연구한다. 복잡도를 줄이기 위해 직접적인 비용함수인 평균 추정오차를 최소화 하는 대신, 평균 추정오차 공분산 역행렬의 로그 행렬식을 비용함수로 채택하고 이를 최대화하는 센서 노드 집합을 선택하기 위한 탐욕적 반복 알고리즘을 제안한다. 비용함수에 있는 관측행렬에 QR분해를 적용하여 단계마다 한 개의 노드를 선택하기 위한 저 복잡도를 갖는 수학적 관계식을 유도한다. 다양한 실험을 통해, 추정 성능 및 복잡도면에서 기존의 센서 노드 선택기술 대비 제안 알고리즘이 경쟁력있는 성능을 보임을 입증하고 실용적 센서 노드 선택기술로써 다양한 네트워크시스템에 적용할 수 있는 방안을 제시한다.

ABSTRACT

We study the problem of selecting a subset of sensor nodes in sensor networks in order to maximize the performance of parameter estimation. To achieve a low-complexity sensor selection algorithm, we propose a greedy iterative algorithm that allows us to select one sensor node at a time so as to maximize the log-determinant of the inverse of the estimation error covariance matrix without resort to direct minimization of the estimation error. We apply QR factorization to the observation matrix in the log-determinant to derive an analytic selection rule which enables a fast selection of the next node at each iteration. We conduct the extensive experiments to show that the proposed algorithm offers a competitive performance in terms of estimation performance and complexity as compared with previous sensor selection techniques and provides a practical solution to the selection problem for various network applications.

키워드 : 센서 선택, 탐욕 알고리즘, 로그 행렬식, 센서 네트워크, 파라미터 추정

Keywords : Sensor selection, Greedy algorithm, Log-determinant, Sensor networks, Parameter estimation

Received 22 October 2022, Revised 4 November 2022, Accepted 16 November 2022

*Corresponding Author Yoon Hak Kim (Email:yhk@chosun.ac.kr, Tel:+82-62-230-7129)

Associate Professor, Department of Electronic Engineering, Chosun University, Gwangju, 61452 Korea

Open Access <http://doi.org/10.6109/jkiice.2023.27.1.103>

print ISSN: 2234-4772 online ISSN: 2288-4165

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

최근 급부상하고 있는 응용 분야 중 하나인 센서네트워크는 다수의 센서노드를 공간상에 무작위로 배치하여, 시스템의 응용 목적(예를들면, 위치 추정, 주변 온도 및 이상현상 감시 등)을 달성하기 위해 각 노드에서 측정값을 수집하게 된다. 이때, 각 노드는 배치된 위치에서 센싱된 값을 획득하게 되므로, 적절한 위치에 있는 센서노드의 결정은 시스템 성능에 큰 영향을 미치게 된다. 즉, 센서 네트워크 시스템에서 특정 파라미터를 추정(estimation)하는 경우, 비용함수인 평균 추정오차를 최소화하는 방향으로 최적의 부분 노드집합을 결정하는 문제를 생각해 볼 수 있으며, 이는 역선형문제(linear inverse problem)를 해결하는 과정이고, 또한, 근본적으로 조합문제(combinatorial problem)로 상당량의 계산량을 초래한다 [1, 2].

이러한 문제를 해결하기 위해 한 단계에서 한 개의 노드를 선택하는 다양한 탐욕적(greedy) 반복 알고리즘이 개발되었다 [2-5]. 직접 비용함수인 평균 추정오차 대신에 하위 모듈(submodular) 함수를 비용함수로 제안하여 이를 최적화함으로써 평균 추정오차와 인접한 범위까지 최적화할 수 있음을 보였다 [2]. 또한 대행렬의 역행렬을 구하는 대신에 평균 추정오차 공분산 역행렬의 로그 행렬식(log-determinant)을 비용함수로 채택하여 이를 최대화하기 위한 반복적인 수학적 유도를 통해 빠른 선택 알고리즘을 제안하였으며 [3], 최적의 센서노드 수와 위치를 정하기 위해 관측행렬(observation matrix)의 최소 고유공간(eigenspace)에 최대 사영(projection)을 적용하는 방법이 제안되었다 [4]. 최근에, QR분해에 기반하여 평균 추정오차를 최소화하는 수학적 관계식을 정립하여 계산량을 합리적으로 유지하면서 최적의 센서노드집합을 선택하는 탐욕 알고리즘이 개발되었다 [5].

센서선택기술과 밀접한 분야로 그래프 샘플링이 있으며, 이는 다수의 그래프 노드에서 특정 노드들로 구성된 집합을 선택하여 그래프 신호의 복원오차를 최소화하는 기술로 다양한 방식이 개발되었다 [6-10]. 특히, 평균 복원오차의 상한값을 최소화하기 위해 가우시안 제거법(Gaussian elimination)을 적용하여 어느정도의 성능저하를 초래하지만 복잡도를 상당수준으로 낮춘 기술이 제안되었다 [7].

본 논문에서는 복잡도를 줄이기 위한 센서 노드 선택 탐욕 알고리즘을 개발하기 위해, 직접 비용함수인 평균 추정오차를 최소화하는 대신에, 평균 추정오차 공분산 역행렬(inverse covariance matrix)의 로그행렬식을 비용함수로 채택하고 이를 최대화 하게 된다. 이를 위해 기존의 방식 [3]과는 다르게 QR분해를 적용하여 반복적으로 한번에 한 개의 노드를 선택하는 수학적 관계식을 유도한다. 다양한 관측행렬에 대해 평균 추정오차 및 복잡도(실행시간)를 평가하여 본 논문에서 제안하는 알고리즘이 기존 선택기술 대비 우수한 성능을 보임을 입증한다.

본 논문은 2장에 비용함수인 평균 추정오차 공분산 역행렬의 로그행렬식을 수학적으로 전개하여 문제를 정립하며, 3장에서는 QR분해를 적용하여 반복적으로 노드를 선택하는 수학적 관계식을 유도한다. 그리고 평균 추정오차 및 실행시간 비교를 통한 성능평가가 4장에서 수행되며, 5장의 결론으로 마무리 한다.

II. 문제 정립

N 개의 센서노드가 배치되어 있는 센서네트워크 시스템에서 노드의 측정값을 통해 파라미터 $\theta \in R^p$ 를 추정하는 경우, 측정잡음 w 에 오염된 측정값 y 는 다음과 같이 표현할 수 있다.

$$y = H\theta + w \quad (1)$$

여기서 $y = [y_1, \dots, y_N]^T$ 는 $N \times 1$ 열벡터이고, H 는 N 개의 $1 \times p$ 행벡터 $h_j^T, j = 1, \dots, N$ 를 갖는 $N \times p$ 관측행렬이다. 그리고, 측정잡음 w 는 $\theta \in R^p$ 와 서로 독립이고 각각 정규 분포(normal distribution) $N(0, \sigma^2 I)$ 와 $N(0, \Sigma_\theta)$ 를 보인다고 가정한다. 전체집합 $V = \{1, \dots, N\}$ 에서 n ($\ll N$)개의 센서노드를 선택하여 집합 S 를 구성하고 선택된 노드의 측정값으로부터 파라미터 θ 의 추정치 $\hat{\theta}$ 을 구하게 된다. 이를 통해 평균 추정오차 $E\|\theta - \hat{\theta}\|^2$ 를 계산할 수 있으며, 또한 평균 추정오차는 추정오차 공분산 행렬 $\Sigma(S)$ 의 주대각합(trace)으로 표현할 수 있고, 여기서 $\Sigma(S)$ 은 다음과 같이 구할 수 있다 [5, 7].

$$\Sigma(S) = \left(\Sigma_\theta^{-1} + \frac{1}{\sigma^2} H_S^T H_S \right)^{-1} \quad (2)$$

여기서 H_S^T 는 선택된 집합 S 에 있는 노드의 인덱스에 해당하는 H 의 행벡터를 열벡터로 구성하여 만들어진 $p \times |S|$ 행렬 (예를들어, j 번째 노드가 선택된 경우 j 번째 행벡터 h_j^T 에 해당)이고, 파라미터 θ 는 베이지안 (Bayesian) 선형 추정방식을 사용하여 추정치 $\hat{\theta}$ 를 구할 수 있으며 이를 수식적으로 표현하면 다음과 같다 [5,7]:

$$\hat{\theta} = \left(\Sigma_\theta^{-1} + \frac{1}{\sigma^2} H_S^T H_S \right)^{-1} \frac{1}{\sigma^2} \sum_{i \in S} h_i y_i \quad (3)$$

본 논문에서는 직접적 비용함수인 평균 추정오차 $tr(\Sigma(S))$ 가 최소가 되는 S 를 구하는 대신 선택기술의 복잡도를 줄이기 위해 평균 추정오차 공분산 역행렬의 로그행렬식 (log-determinant)을 최대로 하는 집합 S 를 탐욕적 반복과정을 통해 구하게 된다. 이를 위해 $(i+1)$ 번째 단계를 수식적으로 나타내면 다음과 같다.

$$j^* = \arg \max_{S_{i+1} = S_i \cup \{j\}, j \in S^c} \log \det \left(\frac{1}{\sigma_\theta^2} I_p + \frac{1}{\sigma^2} H_{S_{i+1}}^T H_{S_{i+1}} \right) \quad (4)$$

$$S_{i+1}^* = S_i^* \cup \{j^*\} \quad (5)$$

여기서 $S_i^c = V - S_i^*$ 으로 $(i+1)$ 번째 단계에서 남아있는 노드의 집합을 의미하고, $\log \det$ 는 로그행렬식 함수를 나타낸다. 그리고, 파라미터의 공분산 (covariance) 행렬은 $\Sigma_\theta = \sigma_\theta^2 I_p$ 이라고 가정하였으며, 여기서 I_p 는 $p \times p$ 항등 (identity) 행렬을 나타낸다. 이러한 노드 선택과정 (식 (4), (5))은 $|S_i^*| = p$ 가 될 때까지 반복된다.

III. 저 복잡도 센서노드 선택 알고리즘

본 논문에서는 선택과정의 수학적 관계식을 유도하기 위해 QR 분해 $H_{S_{i+1}}^T = QR^{(i+1)}$ 를 적용한다. 여기서 Q 는 $p \times p$ 정규 직교 (orthonormal) 행렬이고, $R^{(i+1)}$ 은 $p \times (i+1)$ 상삼각 (upper triangular) 행렬이다. 이 경우 평균 추정오차 공분산 역행렬의 로그행렬식은 다음과 같이 표현될 수 있다.

$$\log \det \Sigma(S_{i+1})^{-1} \quad (6)$$

$$= \log \det \left(\frac{1}{\sigma_\theta^2} I_p + \frac{1}{\sigma^2} H_{S_{i+1}}^T H_{S_{i+1}} \right) \quad (7)$$

$$= \log \det \left(\frac{1}{\sigma_\theta^2} I_p + \frac{1}{\sigma^2} QR^{(i+1)}R^{(i+1)T}Q^T \right) \quad (8)$$

$$= \log \det Q \det \left(\frac{1}{\sigma_\theta^2} I_p + \frac{1}{\sigma^2} R^{(i+1)}R^{(i+1)T} \right) \det Q^T \quad (9)$$

$$= \log \det \left(\frac{1}{\sigma_\theta^2} I_p + \frac{1}{\sigma^2} R^{(i+1)}R^{(i+1)T} \right) \quad (10)$$

여기서 수식 (8)은 QR분해를 적용한 결과이고 수식 (9)는 행렬곱의 행렬식은 각 행렬식의 곱과 같음을 이용했으며 수식 (10)은 $\det Q \det Q^T = 1$ 이므로 성립한다.

또한, 상삼각행렬 $R^{(i+1)}$ 은 $(i+1) \times (i+1)$ 상삼각 정방행렬인 $\bar{R}^{(i+1)}$ 로 다음과 같이 표현할 수 있으며,

$$R^{(i+1)} = \begin{bmatrix} \bar{R}^{(i+1)} \\ 0 \end{bmatrix} \quad (11)$$

여기서, 0 는 $(p-i-1) \times (i+1)$ 영행렬 (zero matrix)을 나타낸다. 따라서, 수식 (10)은 다음과 같이 전개할 수 있다:

$$= \log \det \left(\frac{1}{\sigma_\theta^2} I_p + \frac{1}{\sigma^2} \begin{bmatrix} \bar{R}^{(i+1)}\bar{R}^{(i+1)T} & 0 \\ 0 & 0 \end{bmatrix} \right) \quad (12)$$

$$= \log \det \left(\frac{1}{\sigma^2} \begin{bmatrix} \bar{R}^{(i+1)}\bar{R}^{(i+1)T} + \frac{\sigma^2}{\sigma_\theta^2} I_{i+1} & 0 \\ 0 & \frac{\sigma^2}{\sigma_\theta^2} I_{p-i-1} \end{bmatrix} \right) \quad (13)$$

$$\propto \log \det \left(\bar{R}^{(i+1)}\bar{R}^{(i+1)T} + \frac{\sigma^2}{\sigma_\theta^2} I_{i+1} \right) \quad (14)$$

여기서, 수식 (14)는 $\det \left(\frac{\sigma^2}{\sigma_\theta^2} I_{p-i-1} \right)$ 값이 $(i+1)$ 번째 단계에서 최대의 값을 갖는 노드를 선택하는데 영향을 주지 않으므로 노드 선택 관계식에서 제외가 되어 성립하게 된다. 또한 행렬식은 행렬의 고유치 (eigenvalue)의 곱으로 표현할 수 있고, 상삼각 정방행렬인 경우 고유치는 주대각 (main diagonal)에 배치되므로 $\lambda_j, j = 1, \dots, i+1$ 가 $\bar{R}^{(i+1)}$ 의 고유치일 때 노드선택을 위한 수학적 관계

식은 다음과 같이 결정할 수 있다:

$$j^* = \operatorname{argmax}_{S_{i+1}} \log \left[\left(\lambda_i^2 + \frac{\sigma^2}{\sigma_\theta^2} \right) \cdots \left(\lambda_{i+1}^2 + \frac{\sigma^2}{\sigma_\theta^2} \right) \right] \quad (15)$$

$$= \operatorname{argmax}_{S_{i+1}} \left[\log \left(\lambda_i^2 + \frac{\sigma^2}{\sigma_\theta^2} \right) + \cdots + \log \left(\lambda_{i+1}^2 + \frac{\sigma^2}{\sigma_\theta^2} \right) \right] \quad (16)$$

$$= \operatorname{argmax}_{S_{i+1}} \log \left(\lambda_{i+1}^2 + \frac{\sigma^2}{\sigma_\theta^2} \right) \quad (17)$$

$$= \operatorname{argmax}_{S_{i+1}} \lambda_{i+1}^2 = \operatorname{argmax}_{j \in S^c} \left[q_{i+1}^T h_j \right]^2 \quad (18)$$

여기서, 수식 (17)은 (i+1)번째 단계에서 $\lambda_j, j = 1, \dots, i$ 값은 이미 결정되어 노드선택에 영향을 주지 않게 되므로 성립하며, 또한 (i+1)단계에서 관측행렬 H의 j번째 행벡터 h_j^T 가 선택된 경우, λ_{i+1} 는 QR분해에서 (i+1)번째 정규직교 벡터 q_{i+1} 와 h_j 와의 내적 (inner product)으로 간단하게 구할 수 있다. 따라서, 본 논문에서 유도한 수학적 관계식 (수식 (18))은 매 단계에서의 노드를 선택하는 과정의 복잡도를 상당 수준 감소시키게 된다. 제안 알고리즘의 복잡도 평가는 4장에서 기존기술과의 실행시간 비교를 통해 수행된다.

IV. 실험 및 분석

본 논문에서는 다양한 센서노드 선택 알고리즘의 성능을 평가하기 위해 다음 두 종류의 관측행렬 H에 대해 실험을 수행한다:

- 1) 동일하고 독립적인 (identical and independent) 정규 분포 $N(0, 1)$ 를 갖는 요소로 구성된 랜덤행렬
- 2) 동일하고 독립적인 베르누이 (Bernoulli) 분포를 갖는 요소 (0 또는 1)로 구성된 랜덤행렬

전체노드 수 $N=300$ 으로 하고 총 500번 랜덤으로 관측행렬을 발생하여 각각의 관측행렬에 대해 다양한 선택기술을 적용하여 집합 S를 구성하게 된다. 그리고, 집합 S에 있는 선택된 노드에서의 잡음에 오염된 측정값을 사용하여 수식 (3)을 통해 파라미터 θ 를 추정하고, 평균 추정오차 $E\|\theta - \hat{\theta}\|^2$ 를 계산하게 되는데, 이때 측정잡음 w는 신호 대 잡음비 $\text{SNR}(\text{dB})=10\log_{10}(\sigma_w^2/\sigma^2)$ 를 변화시키면서 잡음의 세기를 조절하게 된다. 본 논문에서 제안하는 노드선택 알고리즘 (수식 (18) 참

조)과 성능비교를 하기 위한 기존의 노드 선택 기술은 다음과 같다:

- 1) 가우시안 제거법에 기반 한 고속 그래프 샘플링 기술 (Efficient Sampling Method, ESM [6])
- 2) 제안 알고리즘과 같은 로그행렬식을 사용한 센서노드 선택기술 (Greedy Sensor Selection, GSS [3])
- 3) 직접 비용함수인 평균 추정오차를 최소화하는 센서노드 선택기술 (GSS based on QR factorization, GSS-QR [5])

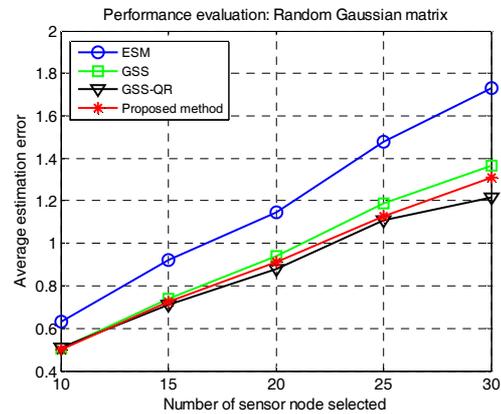


Fig. 1 Performance evaluation of different sensor selection methods for random Gaussian matrix by varying the number of selected nodes with SNR=2dB.

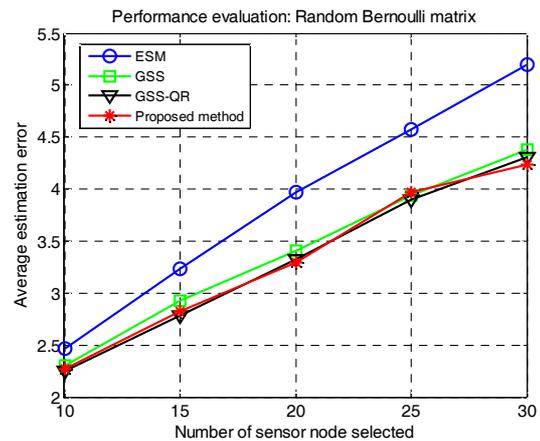


Fig. 2 Performance evaluation of different sensor selection methods for random Bernoulli matrix by varying the number of selected nodes with SNR=2dB.

첫 번째 실험으로 선택 노드수 ($|S|=p$)를 10,25,...,30로 변화해 가면서 SNR=2dB로 유지하여 측정값 y 를 발생하고 이를 사용하여 평균 추정오차의 성능을 비교하였다 (그림 1, 2 참조). 제안 알고리즘이 기존의 선택 기술 대비 경쟁력 있는 평균 추정성능을 보임을 확인할 수 있다. 두 번째 실험에서는 선택 노드수를 $|S|=p=20$ 로 고정하고 잡음세기를 SNR=0dB에서 10dB까지 변화하여 평균 추정오차를 비교하였다 (그림 3, 4 참조). 다양한 잡음환경에서도 제안 알고리즘이 우수한 추정성능을 보임을 알 수 있다.

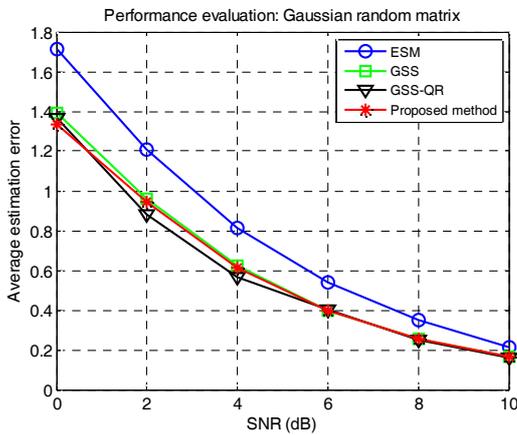


Fig. 3 Performance evaluation of different sensor selection methods for random Gaussian matrix by varying the SNR with 20 selected nodes.

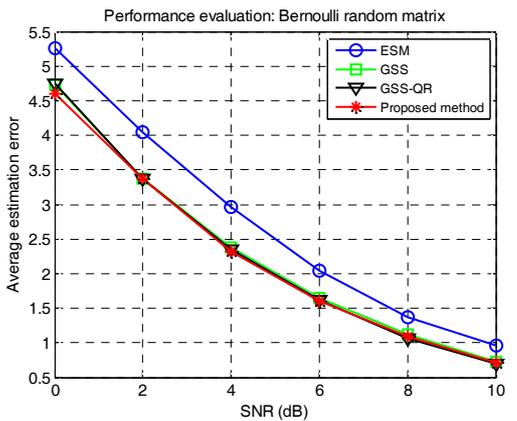


Fig. 4 Performance evaluation of different sensor selection methods for random Bernoulli matrix by varying the SNR with 20 selected nodes.

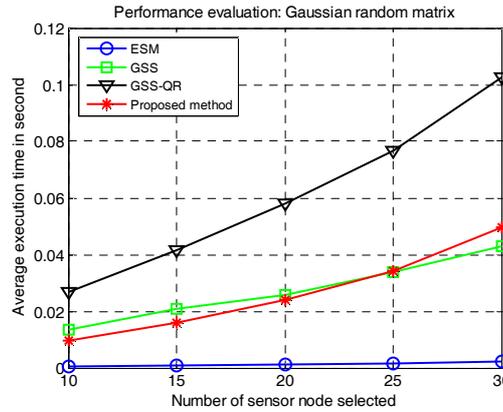


Fig. 5 Complexity evaluation of different sensor selection methods for random Gaussian matrix by varying the number of selected nodes with SNR=2dB.

제안 알고리즘의 복잡도 평가는 실행시간을 비교하여 수행되었으며, 이를 위해 정규 분포 랜덤행렬 H 에 대해 선택 노드수 $|S|=p$ 를 변화해 가면서 SNR=2dB인 경우 선택기술의 실행시간을 비교 하였다 (그림 5 참조). 예상했듯이 기존의 기술과 대비하여 평균 추정성능을 유지하면서 상대적으로 빠른 실행시간 (특히, $n \ll N$ 인 경우)을 보임을 확인할 수 있다.

V. 결론

QR분해에 기반하여 저 복잡도를 갖는 센서노드 선택 기술을 제안하였다. 이를 위해 직접비용함수인 평균 추정오차 대신 평균 추정오차 공분산 역행렬의 로그행렬 식을 최대화하기 위한 수학적 관계식을 유도하였다. 기존의 선택기술 대비 저 복잡도를 유지하면서 경쟁력 있는 평균복원성능을 보임을 실험적으로 입증하였다. 향후 다양한 네트워크 응용분야에 적용할 수 있는 실용적 센서선택 기술에 대한 연구를 지속적으로 수행할 계획이다.

ACKNOWLEDGEMENT

This study was supported by research fund from Chosun University, 2022.

References

- [1] S. Joshi and S. Boyd, "Sensor Selection via Convex Optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451-462, Feb. 2009. DOI: 10.1109/TSP.2008.2007095.
- [2] J. Ranieri, A. Chebira, and M. Vetterli, "Near-Optimal Sensor Placement for Linear Inverse Problems," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1135-1146, Mar. 2014. DOI: 10.1109/TSP.2014.2299518.
- [3] M. Shamaiah, S. Banerjee, H. Vikalo, "Greedy sensor selection: Leveraging submodularity," in *Proceedings of IEEE International Conference on Decision and Control (CDC)*, Atlanta: GA, USA, pp. 2572-2577, 2010. DOI: 10.1109/CDC.2010.5717225.
- [4] C. Jiang, Y. C. Soh, and H. Li, "Sensor Placement by Maximal Projection on Minimum Eigenspace for Linear Inverse Problems," *IEEE Transactions on Signal Processing*, vol. 64, no. 21, pp. 5595-5610, Nov. 2016. DOI: 10.1109/TSP.2016.2573767.
- [5] Y. H. Kim, "Greedy sensor selection based on QR factorization," *EURASIP journal on Advances in Signal Processing*, vol. 2021, no. 117, pp. 1-13, Dec. 2021. DOI: 10.1186/s13634-021-00824-5.
- [6] 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, pp. 3775-3789, Jul. 2016. DOI: 10.1109/TSP.2016.2546233.
- [7] L. F. O. Chamon and A. Ribeiro, "Greedy Sampling of Graph Signals," *IEEE Transactions on Signal Process.*, vol. 66, no. 1, pp. 34 - 47, Jan. 2018. DOI: 10.1109/TSP.2017.2755586.
- [8] Y. H. Kim, "Low-complexity Sampling Set Selection for Bandlimited Graph Signals," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 24, no. 12, pp. 1688-1693, Dec. 2020. DOI: 10.6109/jkiice.2020.24.12.1682.
- [9] Y.H. Kim, "QR factorization-based sampling set selection for bandlimited graph signals," *Signal Processing*, vol. 179, pp. 1-10, Feb. 2021. DOI: 10.1016/j.sigpro.2020.107847.
- [10] Y. H. Kim, "Efficient Sampling of Graph Signals with Reduced Complexity," *Journal of the Korea Institute of Electronic Communication Science*, vol. 17, no. 2, pp. 367-374, Apr. 2022. DOI: 10.13067/JKIECS.2022.17.2.367.



김윤학(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