



# Geometry-Based Sensor Selection for Large Wireless Sensor Networks

Yoon Hak Kim\*, *Member, KIICE*

Department of Electronic Engineering, Chosun University, Gwangju 501-759, Korea

## Abstract

We consider the sensor selection problem in large sensor networks where the goal is to find the best set of sensors that maximizes application objectives. Since sensor selection typically involves a large number of sensors, a low complexity should be maintained for practical applications. We propose a geometry-based sensor selection algorithm that utilizes only the information of sensor locations. In particular, by observing that sensors clustered together tend to have redundant information, we theorize that the redundancy is inversely proportional to the distance between sensors and seek to minimize this redundancy by searching for a set of sensors with the maximum average distance. To further reduce the computational complexity, we perform an iterative sequential search without losing optimality. We apply the proposed algorithm to an acoustic sensor network for source localization, and demonstrate using simulations that the proposed algorithm yields significant improvements in the localization performance with respect to the randomly generated sets of sensors.

**Index Terms:** Iterative sequential search, Localization, Sensor networks, Sensor selection

## I. INTRODUCTION

In sensor networks, a large number of inexpensive, low-power communication sensor devices with one or more sensing capabilities are deployed in a sensor field, providing physical sensing phenomena, processing them, and communicating this information to other sensors. These sensors are battery-powered and the network lifetime is a crucial concern for sensor networks. Collecting measurements from many sensors and transmitting them for various tasks (e.g., localization, tracking targets, and temperature monitoring) will reduce the lifetime of a network. Thus, the sensing tasks should be able to minimize the number of sensors involved in order to prolong the network lifetime. This will motivate the selective use of the most informative sensors, resulting in a reduction of the number of sensors in

operation.

In the scenario of localization using wireless sensor networks, the localization accuracy can be significantly improved by selecting the most informative sensors [1-3]. For example, the entropy-based heuristic approach proposed in [1] greedily selects the next sensor in order to reduce the overall uncertainty. The authors of [2] focus on finding the best set of sensors that maximizes the localization accuracy, under the assumption of a known uncertainty region for the source location; further, they present analytical bounds on the performance of their sensor selection algorithm. In [3], the author addresses the sensor selection problem when the quantization of measurements is taken into account and shows that this sensor selection problem can be treated as a rate allocation problem where the goal is to allocate the rate to each sensor so as to minimize the localization error. To

Received 02 June 2013, Revised 01 July 2013, Accepted 22 July 2013

\*Corresponding Author Yoon Hak Kim (E-mail: [yhk@chosun.ac.kr](mailto:yhk@chosun.ac.kr), Tel: +82-62-230-7129)

Department of Electronic Engineering, Chosun University, 309, Pilmun-daero, Dong-gu, Gwangju 501-759, Korea.

**Open Access** <http://dx.doi.org/10.6109/jicce.2014.12.1.008>

print ISSN: 2234-8255 online ISSN: 2234-8883

© 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

solve this problem, the generalized Breiman, Friedman, Olshen, and Stone (BFOS) algorithm (see [4]) is applied with the rate-distortion (R-D) calculation for each candidate rate allocation.

However, most of the existing algorithms suffer from a high computational complexity, since the sensor networks consist of a large number of sensors and the process of selecting sensors must be executed online whenever every event of interest occurs. Thus, fast and powerful sensor selection algorithms are required to facilitate the practical applications of large sensor networks. In this paper, we propose a simple geometry-based sensor selection algorithm by utilizing only the information of sensor locations. Motivated by the observation that sensors clustered together tend to generate redundant information [3], we seek to find a set of sensors that are located at the maximum distance from each other. More specifically, we search the best set of sensors that maximizes the average distance between the sensors within the set. Furthermore, in order to accelerate the selection process, we propose an iterative sequential search algorithm without losing optimality, instead of directly conducting an exhaustive search of sensors over the sensor field. We show that the proposed algorithm reduces the average distance in each iteration, ensuring the convergence of the algorithm. Our experiments demonstrate that the proposed sensor selection achieves a significant performance improvement as compared to random sensor selections and exhibits a performance close to that of the exhaustive search among sensors over the sensor field.

Throughout this paper, it is assumed that each sensor can obtain a noise-corrupted measurement, such as signal energy or temperature using actual measurements (e.g., time-series measurements or spatial measurements) and there are only one-way noiseless communication channels from the sensors to the fusion node; i.e., there is no feedback channel, and the sensors do not communicate with each other (no relay between sensors).

The rest of this paper is organized as follows: The problem formulation is given in Section II. The geometry-based sensor selection algorithm is explained in Section III. The iterative sequential algorithm is summarized in Section IV and the proposed algorithm is applied to the acoustic amplitude sensor system for source localization, which is introduced in Section V. Simulation results are presented in Section VI and the conclusion is stated in Section VII.

## II. PROBLEM FORMULATION

Within the sensor field  $S$  of interest, suppose that there are  $M$  sensors located at the known spatial locations, denoted by  $\mathbf{x}_i$ ,  $i = 1, \dots, M$ , where  $\mathbf{x}_i \in S \subset \mathbf{R}^2$ . It is assumed that  $K \ll M$  sensors are selected prior to the

sensing operation and the selected sensors will measure signals related to the unknown parameter  $\mathbf{x}$  and send them to a fusion node that estimates the parameter by using  $K$  measurements. The measurement at sensor  $i$  over a given time interval  $k$ , denoted by  $z_i$  can be expressed as follows:

$$z_i(x, k) = f(x, x_i, P_i) + w_i(k), \forall i = 1, \dots, K, \quad (1)$$

where  $f(\mathbf{x}, \mathbf{x}_i, \mathbf{P}_i)$  indicates the sensor model employed at sensor  $i$  and  $w_i$  denotes a combined noise term that includes both measurement noise and modeling error and can be approximated using a normal distribution,  $N(0, \sigma_i^2)$ .  $\mathbf{P}_i$  is the parameter vector for the sensor model (an example of  $\mathbf{P}_i$  for an acoustic amplitude sensor case is given in Section V). It is assumed that each sensor measures its sensor reading  $z_i(x, k)$  at time interval  $k$  and sends it to a fusion node, where all sensor measurements are used to obtain an estimate  $\hat{x}$ . For the case of quantized measurements, we use an  $R_i$ -bit quantizer with a dynamic range at sensor  $i$ . We assume that the quantization range can be selected for each sensor on the basis of the desirable properties of their respective sensing ranges (see [5] for the details). Each sensor generates a quantization index  $Q_i \in I_i = \{1, \dots, 2^{R_i}\}$  to which the measurement  $z_i$  belongs to.

We believe that this formulation is general and captures many scenarios of practical interest. Each scenario will obviously lead to a different sensor model  $f(\mathbf{x}, \mathbf{x}_i, \mathbf{P}_i)$ . For example, for the source localization in acoustic sensor networks, each sensor captures acoustic energy from a source located at  $\mathbf{x} \in S$ . In the following paragraphs, we explain the motivation for a geometry-based sensor selection algorithm and propose an iterative sequential search algorithm.

## III. GEOMETRY-BASED SENSOR SELECTION ALGORITHM

It is observed that sensors clustered together tend to send redundant information to a fusion node [3] and thus, a good strategy would be to select sensors that are at the maximum distance from each other. To be specific, we seek to search the best set of  $K$  sensors from  $M$  sensors in a sensor field that maximizes the average distance over the set:

$$\{\hat{x}_k^*\}_{k=1}^K = \arg \max_{\{x_i\}_{i=1}^K} \sum_{x_i, x_j \in \{x_i\}_{i=1}^K} \|x_i - x_j\|^2, \quad (2)$$

where the distance  $\|x_i - x_j\|$  should be in the range  $[d_{\min}, d_{\max}]$  and  $\{x_i\}_{i=1}^K$  is one of the possible sets that can be constructed by selecting  $K$  from  $M$  sensors. Note that sensors located very far from each other should be excluded

from our selection process because sensors typically assume their sensing ranges for reliable measurements. The parameters  $(d_{\min}, d_{\max})$  that determine the range of the acceptable distance between sensors will depend upon the applications. Clearly, the best set  $\{\hat{x}_k^*\}_{k=1}^K$  will consist of the sensors that are at the maximum distance from each other on average under the constraint of the distance range.

In search for the best set, we take an iterative sequential approach rather than an exhaustive search to make the algorithm practically applicable to large sensor networks ( $M$  large). Suppose that we are initially given the set of  $K$  sensors  $V = \{\hat{x}_1, \dots, \hat{x}_K\}$ . Then, we find the next set that reduces the metric in Eq. (2) in each iteration. This search is repeated until there is no change in the set  $V$ . The procedure for finding the next set can be elaborated as follows: First, we construct the  $K$  regions such that the  $k$ -th region includes the candidate sensors for  $\hat{x}_k$  in the next set. Formally, we obtain from the set  $V = \{\hat{x}_1, \dots, \hat{x}_K\}$

$$S_k = \{x_i : \|x_i - \hat{x}_k\|^2 \leq \|x_i - \hat{x}_j\|^2, j \neq k, j = 1, \dots, K, i = 1, \dots, M\}, \quad (3)$$

$k = 1, \dots, K$

where  $S_k$  consists of the sensors closer to  $\hat{x}_k$  than other sensors,  $\hat{x}_j, j \neq k; j = 1, \dots, K$ . Clearly, each sensor in  $S_k$  can be selected as the next candidate for  $\hat{x}_k$  because most of the assignments of the sensors in  $S_k$  to the other elements,  $\hat{x}_j, j \neq k; j = 1, \dots, K$ , will likely cause the average distance over the next set to decrease. Notice that this step will reduce the search range for the next element  $\hat{x}_k$ . The next step is to choose one from the sensors in each region that maximizes the metric:

$$\hat{x}_k^* = \arg \max_{x_i \in S_k} \sum_{j=1, j \neq k} \|x_i - \hat{x}_j\|^2. \quad (4)$$

Note that this search should be conducted under the constraint of the distance range. It is obvious that the construction of  $K$  regions,  $S_k, k = 1, \dots, K$  and the subsequent selection of the next sensor over each region, the two main tasks in the proposed algorithm will guarantee that the metric remains non-decreasing over the iterations, implying the convergence of the algorithm. Note that each sensor in the next set is searched for over a small region and in a sequential manner, enabling fast operation.

#### IV. PROPOSED ALGORITHM

The proposed algorithm is iteratively executed over all sensors  $i = 1, \dots, M$  until no change in the set  $\{\hat{x}_1, \dots, \hat{x}_K\}$  is achieved; it can be summarized as follows:

Algorithm 1: Geometry-based iterative design algorithm.

**Step 1** : Initially, set iteration index  $n = 1$  and choose randomly  $K$  out of  $M$  sensors to construct a set,

$$V_n = \{\hat{x}_1, \dots, \hat{x}_K\}$$

**Step 2** : Construct  $K$  regions,  $S_k, k = 1, \dots, K$  by using Eq. (3).

**Step 3** : Find the  $k$ -th sensor  $\hat{x}_k^*$  from its corresponding region  $S_k$  that maximizes the average distance by using Eq. (4).

**Step 4** : Set  $\hat{x}_k = \hat{x}_k^*$  and repeat Step 3 for the other  $K - 1$  sensors.

**Step 5** : Set  $n = n + 1$  and construct the new set  $V_n = \{\hat{x}_1, \dots, \hat{x}_K\}$

**Step 6** : If  $V_{n-1} = V_n$ , then stop; otherwise, go to Step 2.

Note that the proposed iterative algorithm suffers from the presence of numerous poor local optima depending on the initialization of set  $V$ . Hence, it is recommended that the sensors sufficiently distant to each other be selected for the initial set.

#### V. APPLICATION TO ACOUSTIC AMPLITUDE SENSOR CASE

As an example of an application of the proposed algorithm, we consider the acoustic amplitude sensor system for source localization where an energy decay model of sensor signal readings proposed in [6] is used for localization. This model is based on the fact that the acoustic energy emitted omni-directionally from a sound source will attenuate at a rate that is inversely proportional to the square of the distance in free space [7] and was verified by the field experiment in [6] and used in [8-10]. When an acoustic sensor is employed at each sensor, the signal energy measured at sensor  $i$  over a given time interval  $k$ , and denoted by  $z_i$ , can be expressed as follows:

$$z_i(x, k) = \gamma_i \frac{a}{\|x - x_i\|^\alpha} + w_i(k), \forall i = 1, \dots, M, \quad (5)$$

where the parameter vector  $\mathbf{P}_i$  in Eq. (1) consists of the gain factor of the  $i$ -th sensor  $\gamma_i$ , an energy decay factor  $\alpha$ , which is approximately equal to 2 in free space, and the source signal energy  $a$ . The measurement noise term  $w_i(k)$  can be approximated using a normal distribution  $N(0, \sigma_i^2)$ . In (5), it is assumed that the signal energy  $a$ , is uniformly distributed over the range  $[a_{\min}, a_{\max}]$ .

The  $K$  sensors are first searched by applying the algorithm proposed in Section IV, and the selected sensors will capture the signal energies generated from a source by using Eq. (5) and send them to a fusion node that will produce the estimate  $\hat{x}$  of the source location from the measurements  $z_i, i = 1, \dots, K$ . For the case of quantized

measurements, it is assumed that each sensor quantizes its measurement with a predesigned quantizer of  $R_i$ -bits before sending them to a fusion node.

### A. Location Estimation Technique Based on Quantized Data

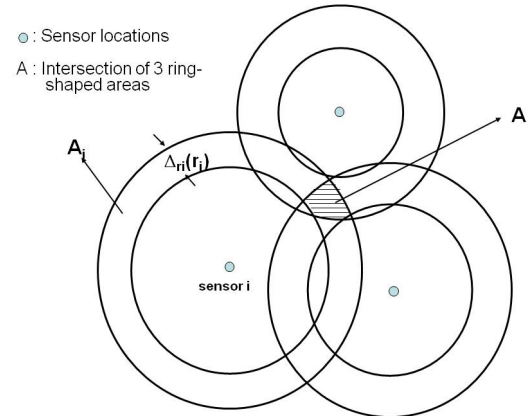
Clearly, for  $z_i(x, k)$  to be useful for localization, it must be a function of the relative positions of the source and the sensor. Thus, there exists some function that can provide an estimate of the source location  $\hat{x}$  on the basis of the original, unquantized measurements; these estimators have been the focus of most of the literature to date on both sensor networks and other source localization scenarios. Instead, we consider the corresponding estimators  $g(\cdot)$  that operate on quantized data to estimate the source location  $\hat{x}$ .

While specific  $g(\cdot)$  choices depend on the sensor model  $f(\mathbf{x}, \mathbf{x}_i, \mathbf{P}_i)$ , we can sketch some of the general properties of this estimator. First, since  $z_i(x, k)$  is used for localization, it must provide distance information about the relative position of the sensor and the source. Thus, after quantization, each transmitted symbol will represent a range of positions (e.g., a range of distances from the sensor). Second, once information is obtained from all sensors, it will be optimal to exploit the range information corresponding to each quantized measurement without obtaining its reconstructed value that has been used in a standard estimator (i.e., by replacing a range of distances by a single distance). That is, an optimal estimator  $g(\cdot)$  should be a function of the range information rather than the reconstructed values.

More specifically, here, we consider the estimator  $g(\cdot)$  that performs energy-based source localization by using quantized data. Note that the localization algorithm to be explained in this section is designed for the high signal-to-noise ratio (SNR) regime ( $\sigma_i = 0$ ) but will also provide a foundation for the localization based on noisy quantized data [9]. Since each quantized sensor reading corresponds to a region where a source is located, all quantized sensor readings lead to a partition of a sensor field obtained by intersecting the regions corresponding to each sensor reading. The localization based on the quantized sensor readings is illustrated in Fig. 1, where the ring-shaped area,  $A_i$ ,  $i = 1, \dots, K$ , corresponds to a quantized measurement at the  $i$ -th sensor. By computing the intersection of all the ring areas (one per sensor), it is possible to define an area where the source is expected to be located. Note that at least three quantized readings are required to achieve a connected intersection. Formally,

$$A_i = \{x : \gamma_i \frac{a}{\|x - x_i\|^\alpha} \in Q_i\}, \forall i = 1, \dots, K, \quad (6)$$

$$\hat{x} = E[x | x \in \bigcap_{i=1}^K A_i]$$

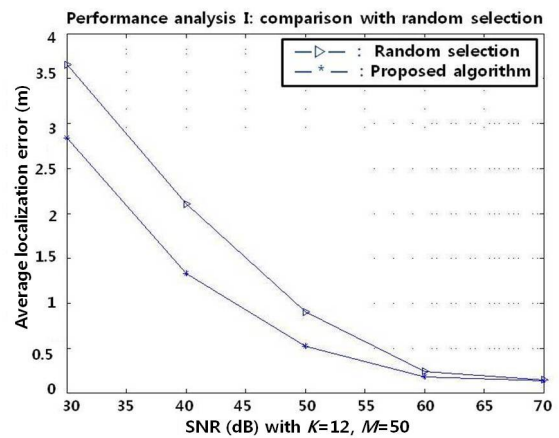


**Fig. 1.** Localization of the source on the basis of the quantized sensor readings.

where  $A_i$  denotes the ring-shaped region obtained from the quantization index  $Q_i$  that  $z_i$  falls into (Fig. 1). If the source is uniformly distributed in the sensor field, the estimate  $\hat{x}$  will be the sample mean in intersection  $A$ . Clearly,  $\hat{x}$  is the minimum mean squared error (MMSE) estimator under the assumption of no measurement noise.

## VI. SIMULATION RESULTS

In the experiments, we consider an acoustic sensor network where  $M (=50)$  sensors are randomly deployed in a sensor field measuring  $20 \text{ m} \times 20 \text{ m}$ . Prior to performing the localization process, the best set of  $K$  sensors is selected by using the proposed algorithm and denoted by  $V^*$ . It is assumed that the source locations are uniformly distributed and the measurements are generated from the model parameters in Eq. (5) by setting  $a = 50$ ,  $\alpha = 2$ , and  $\gamma_i = 1$ .



**Fig. 2.** Comparison of the proposed algorithm with random selection. The average localization error is plotted vs. SNR (dB) with  $K = 12$  and  $M = 50$ . SNR: signal-to-noise ratio.

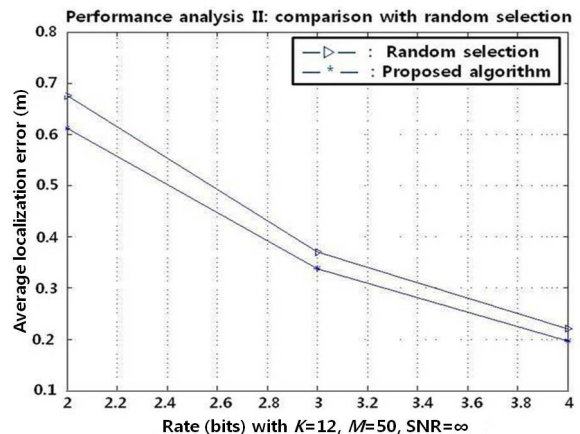
The sensor selection of the proposed algorithm is compared with many sensor selections  $V^R$  that can be constructed by randomly selecting  $K$  sensors from  $M$  sensors. We also investigate the performance of the proposed algorithm when the sensor measurements are quantized before being transmitted to the fusion node. Finally, our search algorithm is evaluated by comparing its solution with the optimal solution obtained by an exhaustive search that requires a huge computational complexity. For the evaluation, the localization error  $E\|x - \hat{x}\|^2$  is computed using the maximum likelihood (ML) estimation except where stated otherwise.

### A. Performance Analysis I: Comparison with Random Sensor Selections

In this experiment, 100 different sensor configurations are generated for  $M = 50$ , and for each configuration, the proposed algorithm is executed to find the best set of  $K = 12$  sensors and is compared with the sensor selection executed by randomly choosing  $K$  from  $M$  sensors. For testing the sensor selections, 1,000 source locations are generated with the SNR ranging from 30 to 70 dB obtained by varying  $\sigma_1 = \sigma$ . Note that the SNR computed by  $10 \log_{10} a^2/\sigma^2$  is measured at a distance of 1 m from the source, and for a practical vehicle target, it is often considerably higher than 40 dB. A typical value of the variance of the measurement noise is  $\sigma^2 = 0.05^2$  (=60 dB) [6, 8]. In Fig. 2, the localization error averaged over 100 configurations is plotted for comparison. As expected, our sensor selection performs considerably well with respect to random selections because the sensors that are remotely distant to each other are likely to be selected and they send considerably less redundant information to a fusion node that can estimate the source location with a relatively high accuracy. In addition, our sensor selection produces a better localization result for noisy cases.

### B. Performance Analysis II: Effect of Quantization

One hundred different sensor configurations are generated for  $M = 50$ , and for each configuration, we execute our algorithm for  $K = 8$  in order to obtain the best set and examine the effect of quantization on the performance of the proposed algorithm with respect to that of random selections. In the experiment, it is assumed that each sensor employs the equally distance-divided quantizer (EDQ), which is designed with the rate  $R_i = 2, 3, 4$  bits. The EDQ is introduced in [11] and shown to achieve good localization performance in an acoustic amplitude sensor system. Fig. 3 demonstrates that the proposed selection shows improved performance over the random selections when the quantization



**Fig. 3.** Comparison of the proposed algorithm with random selection: The average localization error is plotted vs. the rate  $R_i = 2, 3, 4$  bits with  $K = 8, M = 50$ , and signal-to-noise ratio (SNR) =  $\infty$ .

**Table 1.** Comparison of the selection techniques,  $V^R, V^*$  with optimal solutions, and  $V_{opt}$  for  $K = 8, M = 20$ , and SNR = 40 dB

$K$	$V^R$	$V^*$	$V_{opt}$
8	2.8201 m	1.7903 m	1.7354 m

The localization results in meters are obtained by averaging over 100 different sensor configurations. SNR: signal-to-noise ratio.

process is involved. Obviously, a good sensor selection will have a dominant impact on the performance as sensor measurements are coarsely quantized because the importance of the information associated with sensor locations becomes more significant. While computing the localization error based on quantized measurements, the localization algorithm in Eq. (6) is applied for fast computation.

### C. Performance Evaluation: Comparison with Optimal Selection

The proposed algorithm is evaluated by comparing its selection with the optimal selection  $V_{opt}$ , which is generated by an exhaustive search. In this experiment, because the computational complexity required for the search is extremely high, we consider the case of  $K = 8$  and  $M = 20$  in a sensor field measuring  $20 \text{ m} \times 20 \text{ m}$  in order to curtail the search size. It is noted that the proposed sensor selection is suboptimal and is affected by the initial sets.

Thus, for each configuration, the localization errors are computed for  $V^R, V^*$ , and  $V_{opt}$ , respectively, and are averaged over 100 different sensor configurations for the evaluation. In Table 1, the localization results for three different selection techniques (i.e., random selection, the proposed algorithm, and the exhaustive search) are tabulated for the sake of comparison. It can be seen that the proposed technique achieves a performance close to the optimal one

while maintaining a considerably low complexity, implying that the geometric metric in Eq. (2) is effective for the acoustic source localization system.

## VII. CONCLUSION

In this paper, we have proposed a geometry-based sensor selection algorithm optimized for large sensor networks. To achieve low complexity, we suggest the use of the average distance between sensors within a set as a metric to be maximized and propose an iterative sequential search algorithm to find the best set of sensors that maximizes the average distance. This algorithm was applied to an acoustic sensor network for source localization and was shown to perform quite well in comparison to random sensor selections. In the future, we will continue to improve the proposed algorithm in terms of performance and complexity.

## ACKNOWLEDGMENTS

This study was supported by a research fund from Chosun University, 2013.

## REFERENCES

- [ 1 ] H. Wang, K. Yao, G. Pottie, and D. Estrin, "Entropy-based sensor selection heuristic for target localization," in *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks*, Berkeley, CA, pp. 36-45, 2004.
- [ 2 ] V. Isler and R. Bajcsy, "The sensor selection problem for bounded uncertainty sensing models," in *Proceedings of the 4th International Symposium on Information Processing in Sensor Networks*, Los Angeles, CA, pp. 151-158, 2005.
- [ 3 ] Y. H. Kim, "Quantization-aware sensor selection for source localization in sensor networks," *International Journal of KIMICS*, vol. 9, no. 2, pp. 155-160, 2011.
- [ 4 ] E. A. Riskin, "Optimal bit allocation via the generalized BFOS algorithm," *IEEE Transactions on Information Theory*, vol. 37, no. 2, pp. 400-402, 1991.
- [ 5 ] H. Yang and B. Sikdar, "A protocol for tracking mobile targets using sensor networks," in *Proceedings of the 1st IEEE International Workshop on Sensor Network Protocol and Applications*, Anchorage, AK, pp. 71-81, 2003.
- [ 6 ] D. Li and Y. H. Hu, "Energy-based collaborative source localization using acoustic microsensor array," *EURASIP Journal on Applied Signal Processing*, vol. 2003, pp. 321-337, 2003.
- [ 7 ] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ: Prentice-Hall, 1996.
- [ 8 ] J. Liu, J. Reich, and F. Zhao, "Collaborative in-network processing for target tracking," *EURASIP Journal on Applied Signal Processing*, vol. 2003, pp. 378-391, 2003.
- [ 9 ] Y. H. Kim and A. Ortega, "Maximum a posteriori (MAP)-based algorithm for distributed source localization using quantized acoustic sensor readings," in *Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing*, Toulouse, France, pp. 1-5, 2006.
- [10] A. O. Hero and D. Blatt, "Sensor network source localization via projection onto convex sets (POCS)," in *Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing*, Philadelphia, PA, pp. iii/689-iii/692, 2005.
- [11] Y. H. Kim and A. Ortega, "Quantizer design for energy-based source localization in sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5577-5588, 2011.



**Yoon Hak Kim**

received his B.S. and M.S. in Electronic Engineering from Yonsei University, Seoul, Korea, in 1992 and 1994, respectively, and his Ph.D. in Electrical Engineering from University of Southern California in 2007. He is currently with the Department of Electronic Engineering, College of Electronics & Information Engineering, Chosun University. His research interests include distributed compression/estimation in sensor networks with a focus on application-specific compression techniques, distributed source coding, and image compression/enhancement.