Journal of information and
communication convergence engineering
http://jicce.org

J. Inf. Commun. Converg. Eng. 16(3): 142-147, Sep. 2018 Regular paper

Low-Complexity Design of Quantizers for Distributed Systems

Yoon Hak Kim*, ***Member, KIICE***

Department of Electronic Engineering, College of IT Convergence Engineering, Chosun University, Gwangju 61452, Korea

Abstract

We present a practical design algorithm for quantizers at nodes in distributed systems in which each local measurement is quantized without communication between nodes and transmitted to a fusion node that conducts estimation of the parameter of interest. The benefits of vector quantization (VQ) motivate us to incorporate the VQ strategy into our design and we propose a low-complexity design technique that seeks to assign vector codewords into sets such that each codeword in the sets should be closest to its associated local codeword. In doing so, we introduce new distance metrics to measure the distance between vector codewords and local ones and construct the sets of vector codewords at each node to minimize the average distance, resulting in an efficient and independent encoding of the vector codewords. Through extensive experiments, we show that the proposed algorithm can maintain comparable performance with a substantially reduced design complexity.

Index Terms: Distributed compression, Generalized Lloyd algorithm, Quantizer design, Sensor networks, Vector quantization (VQ)

I. INTRODUCTION

Efficient design techniques for local quantizers have been presented to improve the rate-distortion (R-D) performance for distributed systems in which battery-operated sensor nodes randomly placed in a sensor field gather measurements from the parameter of interest, quantize the measurements and transmit them to a fusion node that executes estimation using the received quantized data. The generalized Lloyd algorithm has been employed to design quantizers that operate at local nodes. It should be noted that the Lloyd algorithm was devised to minimize the local metric (i.e., reconstruction error) but the design metric for distributed systems should be global ones (i.e., estimation error, a function of all local measurements). Consequently, design difficulty arises in optimizing the global metric at local nodes in distributed systems in which sensor nodes transmit their quantized measurements directly to a fusion node without exchanging data with the other nodes. In other words, encoding of local measurement into one of the quantization partitions or codewords to minimize the global metric cannot be properly conducted without computation of the global metric.

To avoid this difficulty, new related cost functions have been introduced to guarantee both the non-increasing global metric and the independent encoding at local nodes for distributed systems [1-3]. For distributed detection systems, the probabilistic distance was suggested under two hypotheses for distributed detection to enable a manageable design process [1]. For estimation in sensor networks, one weighted metric was presented to iteratively design quantizers by searching for the weight that ensures the non-increasing metric and makes the independent encoding possible [2, 3].

Novel encoding techniques have also been developed to handle the design difficulty. Noting that to achieve performance improvement, scalar quantizers for distributed sys-

Received 23 July 2018, **Revised** 21 August 2018, **Accepted** 21 August 2018
***Corresponding Author** Yoon Hak Kim (E-mail: yhk@chosun.ac.kr, Tel: +82-62-230-7129)
Department of Electronic Engineering, College of IT Convergence Engineering, Chosun University, 309 Pilmun-daero, Dong-gu, Gwangju 61452, Korea.

Open Access https://doi.org/10.6109/jicce.2018.16.3.142 print ISSN: 2234-8255 online ISSN: 2234-8883

tems should be designed in a non-regular framework, implying that the multiple disjoint partitions are mapped into the same codeword [4], a new encoding algorithm was presented to generate non-regular quantizers systematically in acoustic sensor networks [5]. To further improve the estimation performance a non-regular design was used to assign multiple partitions to a single codeword in an iterative manner [6]. A recently introduced novel approach to quantization allows mapping of multiple codewords into a single partition to ensure the independent encoding, yielding a substantial performance improvement [7]. Furthermore, an independent encoding method that links global codewords to local ones was presented for distributed estimation [8].

It is well known that vector quantization (VQ) yields a lower distortion than scalar quantizers for several reasons: more accurate representation of the samples by using correlation between samples and more flexibility owing to quantization in higher dimensions [9]. However, despite the benefit of VQ, it has been excluded for quantizer design for distributed systems considered in this work because the vector codewords can be properly encoded only if each node has access to all of the local measurements, which is not allowed in our scenario.

In this paper, we revisit VQ for quantizer design and focus on developing encoding methods that assign local measurements independently into a set of the vector codewords generated from the standard VQ design process. Specifically we propose an encoding technique that constructs a fixed number of sets of vector codewords, with each codeword in the sets closest in distance to its local codeword. To measure the distance between vector codewords and local ones, we introduce new metrics and test them under various conditions. It is noted that estimation at the fusion node can be conducted by simply computing a centroid of the intersection of the codeword sets transmitted from the nodes involved. Clearly, the proposed algorithm aims at minimizing simply the distance on average, not the estimation error which demands a high cost of computational complexity in previous designs. In addition, the proposed design operates on a relatively small number of vector codewords, not a huge number of training samples. These features of the proposed technique ensure a low-complexity design process. We finally conduct experiments to demonstrate a noteworthy performance improvement of the proposed algorithm at the cost of substantially reduced complexity with respect to previous design techniques.

In the rest of this paper, we formulate the problem in Section II and explain the VQ and new distance metrics in Section III. We present the proposed algorithm in Section IV and introduce an application system for our algorithm in Section V. We investigate the proposed algorithm through experiments in Section VI and provide conclusions in Section VII.

II. PROBLEM FORMULATION

In a sensor field $S \subset \mathbf{R}^N$, M sensor nodes are randomly deployed at known locations, denoted by $\boldsymbol{x}_i \in \mathbf{R}^2$, $i = 1, \ldots, M$. Each node collects its measurement denoted by z_i on the parameter θ to be estimated as follows:

$$z_i(\theta) = f_i(\theta) + \omega_i, \quad i = 1, \ldots, M, \tag{1}$$

where the measurement z_i at the i-th node is generated from the sensing model $f_i(\theta)$ and is typically noise-corrupted by the additive noise ω_i which is assumed to be drawn from the normal distribution $N(0, \sigma_i^2)$. The measurement z_i is quantized by an R_i-bit quantizer at node i (equivalently, the quantization level is $L_i = 2^{R_i}$) which assigns its measurement z_i to one of the pre-computed sets of vector codewords V_i^j, $j = 1, \ldots, L_i$ if the mapping or encoding of z_i to the j-th set minimizes the distance metric which will be introduced in the following section. In this work, we first generate the vector codewords using the well-known VQ algorithm [9] and given those codewords, we seek to partition those codewords into L_i sets by proposing a low-weight independent encoding method.

III. VECTOR QUANTIZATION AND NEW DISTANCE METRICS

In this section, we seek to integrate the VQ methodology into our design of local quantizers to retain the benefits of VQ. We first conduct the standard VQ to generate the vector codewords $\hat{z}^k = (\hat{z}_1, \ldots, \hat{z}_M)$ where K is the vector quantization level as follows:

$$V^k = \{\mathbf{z}: \|\mathbf{z} - \hat{\mathbf{z}}^k\|^2 \le \|\mathbf{z} - \hat{\mathbf{z}}^l\|^2, \forall k \ne l\}, \tag{2}$$

$$\hat{\mathbf{z}}^{k*} = arg\min_{\hat{\mathbf{z}}} E[\|\mathbf{z} - \hat{\mathbf{z}}\|^2 | \mathbf{z} \in V^k], \tag{3}$$

where $\mathbf{z}$ indicates M-tuple measurements $(z_1, \ldots, z_M)$ and V^k, $k = 1, \ldots, K$ is the k-th vector quantization partition constructed to minimize the standard VQ metric $E\|\mathbf{z} - \hat{\mathbf{z}}\|$. The two processes in (2) and (3) are iteratively repeated with $\hat{\mathbf{z}}^k$ replaced by $\hat{\mathbf{z}}^{k*}$ at the next iteration until the VQ metric is minimized.

We introduce the distortion metrics that define the distance between the scalar value (i.e., measurement z_i) and the M-dimensional vector (i.e., the k-th vector codeword $\hat{\mathbf{z}}^k$). In doing so, we construct the set of vector measurements $A_z(z_i)$ satisfying $P(\mathbf{z} \in A_z(z_i) | z_i) \approx 1$. Note that by assuming the noiseless condition ($\sigma_i^2 = 0$), the following set can be easily constructed:

$$A_{\mathbf{z}}(z_i) = \{\mathbf{z}: \ z_i = f_i(\theta), \ \forall i, \ \theta \in A_\theta(z_i)\}, \tag{4}$$

where $A_\theta(z_i) = \{\theta : z_i = f_i(\theta), \forall i, \theta \in S\}$. Using the set $A_z(z_i)$, we present the two distortion metrics denoted by D1 and D2, respectively:

$$\text{D1: } \|z_i - \hat{\mathbf{z}}^k\|^2 \equiv \min_{z \in A_z(z_i)} \|z_i - \hat{\mathbf{z}}^k\|^2, \quad (5)$$

$$\text{D2: } \|z_i - \hat{\mathbf{z}}^k\|^2 \equiv E_{\mathbf{z} \in A_z(z_i)} \|z_i - \hat{\mathbf{z}}^k\|^2. \quad (6)$$

Note that (5) is the metric similar to that in [8] where the distance was measured between the local measurement and the estimated value $\|z_i - \hat{\theta}^k\|^2$. It should be mentioned that the metric D1 allows us to find the vector codeword that maximizes the likelihood $p(\hat{\mathbf{z}}^k|z_i)$ and the metric D2 chooses the vector codeword that maximizes the average likelihood.

Now, we are in a position to partition the vector codewords in to L_i sets $V_i^j, j = 1, ..., L_i$ as follows:

$$V_i^j = \{\hat{\mathbf{z}}^k : \|\hat{z}_i^j - \hat{\mathbf{z}}^k\|^2 \le \|\hat{z}_i^l - \hat{\mathbf{z}}^k\|^2, \forall j \ne j\}, \forall k. \quad (7)$$

Obviously, the elements in V_i^j, $j = 1, ..., L_i$ are relatively more likely to occur given the corresponding local codeword $\hat{z}_i^j$ than those in V_i^l, $\forall l \ne j$. The metrics defined above can be modified by allowing multiple memberships for each vector codeword during the quantization process, providing possible design choices for various applications:

$$V_i^l = \{\hat{\mathbf{z}}^k : \|\hat{z}_i^l - \hat{\mathbf{z}}^k\|^2 - \min_j \|\hat{z}_i^j - \hat{\mathbf{z}}^k\|^2 \le Th_i\}, \forall l \quad (8)$$

where the design parameter Th_i determines how many sets each vector codeword can belong to. It should be noticed that computation of $\|\hat{z}_i^j - \hat{\mathbf{z}}^k\|^2$ in (7) and (8) can be conducted by using one of the two different metrics D1 and D2, resulting in four different quantizers. In addition, the sets constructed by (7) are disjoint to each other whereas the sets by (8) possibly generate non-empty intersections.

IV. PROPOSED QUANTIZER DESIGN ALGORITHM

We present an iterative procedure that enables the independent encoding that minimizes the distance cost function $E_{z_i}\|\hat{z}_i - \hat{\mathbf{z}}^k\|^2$, producing L_i sets of vector codewords at each node.

Algorithm: Low-complexity quantizer design procedure at node i

Step 1: Initialize the local codewords $\hat{z}_i^j, j = 1, ..., L_i$.

Step 2: Group the vector codewords $\hat{\mathbf{z}}^k$, $k = 1, ..., K$ into L_i sets V_i^j, $j = 1, ..., L_i$ to minimize the distance cost function $E_{z_i}\|z_i - \hat{\mathbf{z}}^k\|^2$ using (7) or (8).

Step 3: For each set V_i^j, the local codeword is updated as follows:

$$\hat{z}_i^{j^*} = \arg\min_{z_i} E[\|z_i - \hat{\mathbf{z}}^k\|^2 | \hat{\mathbf{z}}^k \in V_i^j] \quad (9)$$

Step 4: Repeat Step 2 to 3 with $\hat{z}_i^j$ replaced by $\hat{z}_i^{j^*}$ until there is no change in $\{\hat{z}_i^j\}, j = 1, ..., L_i$

It should be emphasized that the proposed design runs iteratively on the K vector codewords, (typically, 200 to 500 codewords generated in the experiment in this work) which is a much smaller set than that of the training measurement samples (≥1,000). Furthermore, it requires only simple computation of the distance metric D1 or D2, which takes much lower complexity than the estimation error $E\|\theta - \hat{\theta}\|^2$, the highly computational step in the previous designs in [3, 6] that should be computed for all of the measurement samples at each iteration.

The independent encoding of z_i at node i into one of the sets $V_i^j, j = 1, ..., L_i$ can be executed by computing the simple Euclidean distance between local codewords and measurements. Specifically, z_i is assigned to V_i^j at node i if $|z_i - \hat{z}_i^k|^2 \le |z_i - \hat{z}_i^l|^2$, $\forall l \ne j$. For estimation at the fusion node, simply taking the intersection of the M received sets can yield a good estimate. Formally,

$$V = \cap_{i=1}^M V_i, \ \hat{\theta}^* = E[\hat{\theta}(\hat{\mathbf{z}}) | \hat{\mathbf{z}} \in V], \quad (10)$$

where V_i indicates the set V_i^j if the j-th index at node i is transmitted to the fusion node. For performance evaluation, the estimate $\theta(\hat{\mathbf{z}})$ is computed by using the maximum likelihood estimation (MLE). Notice that the received M sets generated from noise-corrupted measurements may produce empty intersections. In the experiments in Section VI, the MLE method is also employed to deal with those cases for comparison with the other design algorithms that use the MLE for estimation.

V. APPLICATION OF PROPOSED ALGORITHM TO SENSOR NETWORKS

We test our proposed algorithm by applying it to a source localization system where M sensor nodes are randomly scattered in a sensor field S to gather signal energy from a source by using acoustic amplitude sensors. To collect the signal energy measurements at nodes, we adopt the energy decay model proposed in [10]. Note that the sensor model was widely used for various applications [10, 11]. Specifically, the signal energy measured at sensor i, denoted by z_i can be given by

$$z_i(\theta) = g_i \frac{a}{\|\theta - x_i\|^\alpha} + \omega_i, \quad (11)$$

where the energy model corresponds to the sensing model $f_i(\theta)$ in (1) which consists of the energy decay factor α (≈ 2), the i-th gain factor g_i, and the source signal energy a. It is assumed that the energy measurement is contaminated by the measurement noise w_i which is assumed to be generated from a normal distribution, $N(0, \sigma_i^2)$. In the experiments, it is assumed that the signal energy is known for estimation. However, in real applications, the signal energy can be jointly estimated with its location (see [12] for detail).

VI. SIMULATION RESULTS

In the experiments, we consider a source localization system in acoustic sensor networks where M (= 5) sensor nodes are scattered in a 10 m × 10 m sensor field and generate 100 different node configurations. For each configuration, we run the algorithm in Section IV that employs one of the four different clustering criteria: the set construction by (7) using one of the two distance metrics D1 and D2 in (5) and (6), denoted by low-complexity quantizers with distance metrics LCQ-D1 and LCQ-D2, respectively and the construction by (7) using the distance metrics, also denoted by LCQ-D3 and LCQ-D4, respectively. In designing them, we generate the training samples from the model parameters given by $\alpha = 2$, $g_i = 1$, $a = 50$, and the noiseless condition $\sigma_i^2 = \sigma^2 = 0$ in (1) by assuming a uniform distribution of source locations. We also design typical quantizers such as uniform quantizers (Unif Q) and Lloyd quantizers (Lloyd Q) for comparison. We generate the test sets of 1000 source locations to evaluate the design algorithms. Note that we can improve the system performance by choosing one of the four quantizers (LCQ-D1 to LCQ-D4) that shows the best performance for each node configuration, denoted by LCQ-select. For performance evaluation, we compare the average localization errors $E\|\theta - \hat{\theta}\|^2$ computed by the simple estimation method in (10) for the proposed quantizers. The MLE is used for the other quantizers. In this section, all of the experiments are conducted in MATLAB.

A. Performance Comparison with Various Different Designs

We first compare the proposed quantizers with typical quantizers (i.e., Lloyd Q and Unif Q). In Fig. 1, the localization error in meters is averaged for each rate R_i over 100 configurations. It can be seen that our quantizers (LCQ-D1, LCQ-D2, and LCQ-select) provide a significant performance gain over typical quantizers because the benefits of VQ are efficiently realized in our design framework for distributed systems by transmitting sets of the vector codewords whereas the typical designs focus on sending a local codeword without linking to ones of the other nodes involved. As expected, LCQ-select shows improved performance over LCQ-D1 and LCQ-D2, offering a useful design choice in real applications.

We also compare our quantizer (LCQ-select) with the previous designs—i.e., localization-specific quantizer (LSQ) [3] and distributed optimized quantizer (DOQ) [6]—which were presented for the design of quantizers in distributed systems. In Fig. 2, it should be noted that although the proposed algorithm seeks to minimize the indirect metric (i.e., distance cost function) whereas LSQ and DOQ focus on minimizing directly the localization error, LCQ-select achieves comparable performance with a substantially reduced complexity compared with the previous novel design techniques. In Table 1, the numerical data for the localization error are provided for various designs. In the experiment, the proposed algorithm operates by a factor of ~10 faster than in the previous work [3] for the case of $M = 5$, $R_i = 3$ and K=350.

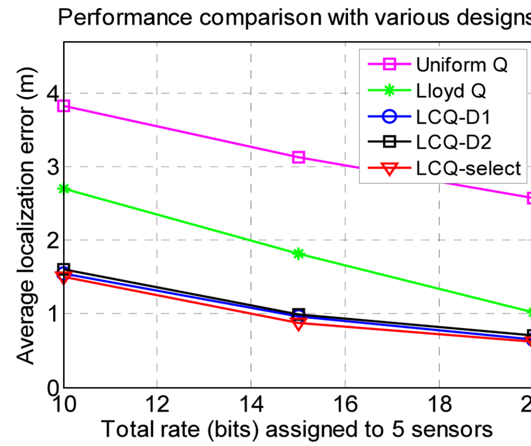

Fig. 1. LCQ-D1, LCQ-D2, and LCQ-select are designed with $\sigma^2 = 0$ and compared with typical designs (uniform Q and Lloyd Q).

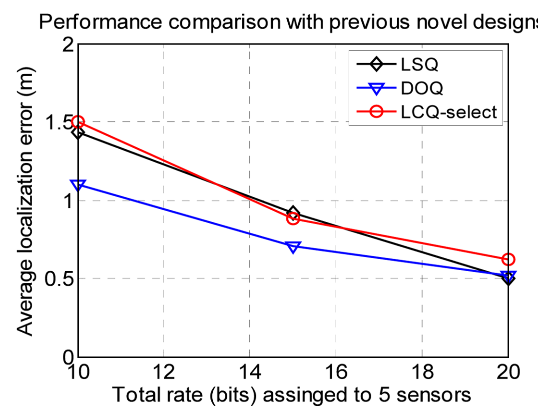

Fig. 2. LCQ-select designed with $R_i = 3$, $\sigma^2 = 0$ and $K = 350$ are compared with novel designs.

Table 1. Average localization error of various design algorithms with $M = 5$ and $R_i = 3$ and $K = 350$

Quantizer	Average localization error (m)
Unif Q	3.1246
Lloyd Q	1.8120
LSQ	0.9182
DOQ	0.7038
LCQ-select	0.8788

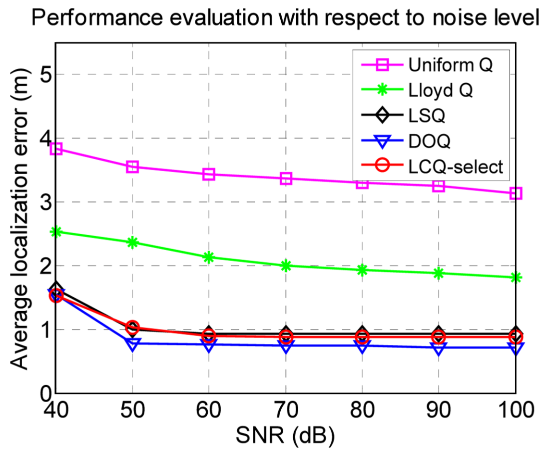

Fig. 3. Design algorithms are evaluated by varying noise level with M = 5, R_i = 3 and a = 50.

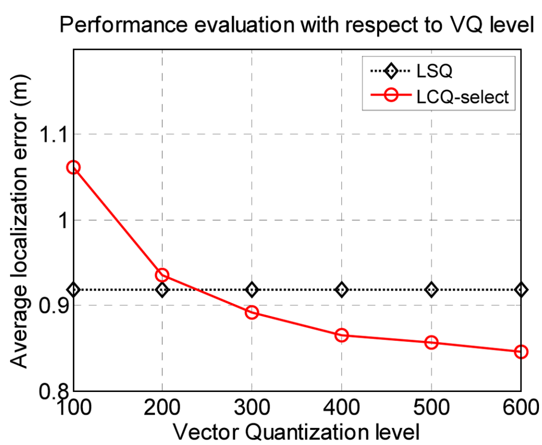

Fig. 4. Performance evaluation with respect to vector quantization level with M = 5, R_i = 3 and a = 50.

B. Performance Evaluation with Respect to Noise Level and Vector Quantization Level

We generate test samples by varying the measurement noise σ_i in the range from SNR = 40 to 100 dB for each configuration. The SNR is measured by 10 $\log_{10} a^2/\sigma^2$ and the noisy engine sound of practical vehicle targets is typically much higher than 40 dB [10, 11]. In Fig. 3, the R-D curves are provided; they exhibit a robustness to noise level compared with the other algorithms.

Furthermore, we investigate the characteristics of our quantizers with respect to the vector quantization level K. As expected in Fig. 4, as K becomes large, the proposed quantizer outperforms LSQ with an increased design complexity, providing a good design trade-off for K which clearly depends upon the application system.

VII. CONCLUSION

We presented a low-complexity design for local quantizers in distributed systems that incorporate the benefits of VQ design methodology into our design process. We introduced four distance metrics that measure the distance between vector codewords and local ones. We proposed an iterative design algorithm that constructs the sets of vector codewords such that the distance cost function is minimized. We conducted extensive experiments, demonstrating that the proposed technique offers significant performance improvement over typical designs and operates much faster than the previous novel ones. In the future, we will find various useful applications for our algorithm and study its theoretical aspects.

ACKNOWLEDGMENTS

This study was supported by a research fund from Chosun University 2017.

REFERENCES

[1] M. Longo, T. D. Lookabaugh, and R. M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Transactions on Information Theory*, vol. 36, no. 2, pp. 241-255, 1990. DOI: 10.1109/18.52470.

[2] Y. H. Kim and A. Ortega, "Quantizer design for source localization in sensor networks," in *Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP)*, Philadelphia, PA, 2005. DOI: 10.1109/ICASSP.2005.1416144.

[3] 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. DOI: 10.1109/TSP.2011.2163401.

[4] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed quantization over noisy channels," *IEEE Transactions on Communications*, vol. 57, no. 6, pp. 1693-1700, 2009. DOI: 10.1109/TCOMM.2009.06.070482.

[5] Y. H. Kim and A. Ortega, "Distributed encoding algorithms for source localization in sensor networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, article no. 781720, 2010. DOI: 10.1155/2010/781720.

[6] Y. H. Kim, "Quantizer design optimized for distributed estimation," *IEICE Transactions on Information and Systems*, vol. 97, no. 6, pp. 1639-1643, 2014. DOI: 10.1587/transinf.E97.D.1639.

[7] Y. H. Kim, "Encoding of quantisation partitions optimised for distributed estimation," *Electronics Letters*, vol. 52, no. 8, pp. 611-613, 2016. DOI: 10.1049/el.2015.3470.

[8] Y. H. Kim, "Cluster-based quantization and estimation for distributed systems," *Journal of Information and Communication Convergence Engineering*, vol. 14, no. 4, pp. 215-221, 2016. DOI: 10.6109/jicce.2016.14.4.215

[9] K. Sayood, *Introduction to Data Compression*, 2nd ed. San Francisco, CA: Morgan Kaufmann Publisher, 2000.

[10] D. Li and Y. H. Hu, "Energy-based collaborative source localization using acoustic microsensor array," *EURASIP Journal on Applied Signal Processing*, vol. 2003, article no. 985029, 2003. DOI: 10.1155

/S1110865703212075.
[11] J. Liu, J. Reich, and F. Zhao, "Collaborative in-network processing for target tracking," *EURASIP Journal on Applied Signal Processing*, vol. 2003, article no. 616720, 2003. DOI: 10.1155/S111086570321204X.
[12] Y. H. Kim and A. Ortega, "Maximum a posteriori (MAP)-based algorithm for distributed source localization using quantized acoustic sensor readings," in *Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP)*, Philadelphia, PA, 2006. DOI: 10.1109/ICASSP.2006.1661460.

Yoon Hak Kim
received the B.S. and M.S. degrees in Electronic Engineering from Yonsei University, Seoul, Korea in 1992 and 1994, respectively and the Ph.D. degree in Electrical Engineering from University of Southern California in 2007. He is currently with Department of Electronic Engineering, College of IT Convergence 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 graph signal processing.