PAPER

Region-Based Distributed Estimation Using Quantized Data*

Yoon Hak KIM^{†a)}, Member

SUMMARY In this paper, we consider distributed estimation where the measurement at each of the distributed sensor nodes is quantized before being transmitted to a fusion node which produces an estimate of the parameter of interest. Since each quantized measurement can be linked to a region where the parameter is found, aggregating the information obtained from multiple nodes corresponds to generating intersections between the regions. Thus, we develop estimation algorithms that seek to find the intersection region with the maximum likelihood rather than the parameter itself. Specifically, we propose two practical techniques that facilitate fast search with significantly reduced complexity and apply the proposed techniques to a system where an acoustic amplitude sensor model is employed at each node for source localization. Our simulation results show that our proposed algorithms achieve good performance with reasonable complexity as compared with the minimum mean squared error (MMSE) and the maximum likelihood (ML) estimators.

key words: distributed estimation, maximum a posteriori (MAP), quantization, source localization, sensor networks

1. Introduction

Estimation algorithms have been proposed for distributed systems where each sensor node equipped with various sensor types, e.g., acoustic, seismic or thermal, collects its measurement and sends it with limited communication bandwidth to a fusion node physically separate from the nodes. One important distributed estimation with broad application is source localization in sensor networks. Localization based on acoustic signal energy measured at individual acoustic amplitude sensors was proposed [10], where each sensor transmits unquantized acoustic energy readings to a fusion node, which then computes an estimate of the location of the source. The localization problem has been solved mostly through nonlinear least squares estimation, which is sensitive to local optima and saddle points. To overcome this drawback, alternative approaches that cast the problem as a convex feasibility problem have also been proposed [4].

Note that quantization of measurements arises in practical estimation systems, especially power-constrained systems such as sensor networks where estimation is conducted on quantized data. In addition, most standard statistical methods for estimation were developed on the mod-

els that interpret quantized data as continuous variables, which can render those methods inappropriate and misleading when quantization process is involved in estimation [18]. Thus, estimation algorithms should be assessed in terms of both the accuracy and the amount of communication data the nodes require to transmit to a fusion node, while efficient quantization schemes will also be needed in order to achieve improvements in rate-distortion performance. Lam and Reibman [9] discussed construction of optimal quantization partitions for quantizer design in distributed estimation systems. In the high resolution regime, asymptotic quantizer designs for distributed estimation were derived in terms of the point density function which is the limiting density of quantizer partitions [13], [14]. The necessary conditions for optimal quantization rules and linear estimation fusion rules were derived and shown to be searched simultaneously by using an iterative algorithm [17]. For acoustic sensor networks, quantizer design algorithms optimized for source localization were proposed and shown to achieve significant performance gain as compared with typical standard designs [6], [8]. A survey for quantization incorporated to distributed estimation and classification is provided [2].

In a simple distributed framework where a parameter of interest is directly estimated at each sensor, distributed estimators based on quantized data were derived [19]; these results rely on the availability of measurement noise statistics. A source localization system where each sensor measures the signal energy, and sends the quantized sensor reading to a fusion node is considered [15]. In this framework, the maximum likelihood (ML) estimation problem based on quantized data was addressed and the Cramer-Rao bound (CRB) was derived for comparison. The ML estimation over noisy channel was addressed for bandwidth-constrained sensor networks [1]. In the absence of noise distribution, (quasi-) optimal distributed estimation algorithms were proposed so as to minimize the mean squared error (MSE) for sensor networks by assigning optimal rate to sensors [11].

In order to make distributed estimation of the parameter possible, each different parameter has to produce a different set of readings at the nodes so that the vector of readings received by the fusion node uniquely determines the parameter itself. Quantization of the readings at each sensor node reduces the accuracy of the estimation by linking each quantized value to a region with a shape that depends on the characteristics of the sensor. Thus, a vector of quantized readings at a set of nodes should uniquely determine

Manuscript received February 24, 2014.

Manuscript revised June 30, 2014.

[†]The author is with the Department of Electronic Engineering, College of Electronic & Information Engineering, Chosun University, Korea.

*This work was presented in part in International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2006.

a) E-mail: yhk@chosun.ac.kr

DOI: 10.1587/transinf.2014EDP7063

an intersection region where the parameter can be found. However, estimation becomes more difficult when measurement noise is not negligible since the situation makes it more likely that a vector of readings will lead to an empty intersection region.

To solve this problem, we considered distributed estimation for source localization in acoustic sensor networks where the source is expected to be located in the intersection of all the regions, one constructed from each quantized measurement and proposed a novel estimation technique that enables a search for the intersection region with the maximum likelihood [7]. We demonstrated that good performance can be achieved with reduced complexity when compared with the minimum mean squared error (MMSE) estimator which is optimal in a sense of minimizing the mean squared estimation error. We also showed that the intersection regions can be expressed as average likelihood functions, which can be efficiently used for obtaining two optimal estimators in the maximum a posteriori (MAP) sense and Kullback-Leibler (KL) divergence sense, respectively [5].

In this paper, we extend our previous work [7] presenting novel region-based estimation techniques in a general distributed estimation framework with more extensive experiments. We believe that our approach is applicable to general cases where quantized measurements at sensor nodes can be associated with regions of the parameter. To address the approach we use a probabilistic formulation, where we model the likelihood that a candidate parameter would produce one of the given intersection regions or vector readings which could be obtained from noiseless measurements (e.g., a high signal-to-noise ratio (SNR) case); we first formulate the estimation problem as an MMSE estimation problem and this is usually avoided for energy-constrained systems such as sensor networks because of its high computational complexity.

To circumvent the difficulty, we construct a region-based estimation framework where the possible intersection region or the vector reading with maximum posterior is searched for estimation rather than the parameter itself. To curtail the search complexity, we then propose two practical search techniques that yield a significant reduction in computational complexity while maintaining good performance. The benefit of the proposed techniques is illustrated by the complexity analysis and the comprehensive experiments in which our proposed algorithms are applied for source localization in acoustic sensor networks for evaluation. Simulation results show that the proposed techniques achieve good localization accuracy with reasonable complexity as compared with the typical estimation methods such as the ML estimators (ML-E) and MMSE estimators (MMSE-E). It is also demonstrated that the proposed algorithms show performance close to the MMSE estimation, as the rate or the SNR becomes increased.

This paper is organized as follows. We consider distributed estimation when the sensor measurements are quantized and formulate this as the MMSE problem in Sect. 2. We then construct a region-based distributed estimation

framework based on the MAP criterion in Sect. 3 which will be efficiently employed to propose the two practical estimation techniques in Sect. 4. We present a brief analysis of the computational complexity for our methods in Sect. 5. As an application system, we consider source localization in acoustic sensor networks in Sect. 6 and then evaluate the proposed techniques for this system in Sect. 7. Finally, we draw conclusions and give future directions for research on distributed systems in Sect. 8.

2. Problem Formulation

Within the sensor field $S \subset \mathbf{R}^N$ where N is the number of dimensions, assume that there are M sensor nodes located at known spatial locations denoted by $\mathbf{x}_i \in \mathbf{R}^2, i = 1, \dots, M$. The sensor nodes collect measurements on the unknown parameter $\theta \in S$ to be estimated and the measurement z_i at node i can be expressed as follows:

$$z_i(\theta) = f_i(\theta) + \omega_i, \quad i = 1, \dots, M \quad (1)$$

where $f_i(\theta)$ is the sensing model employed at sensor node i and it is assumed that the measurement noise ω_i is independent of θ and can be approximated by a normal distribution $N(0, \sigma_i^2)$. It is also assumed that each node equipped with a quantizer of R_i bits collects its noise-corrupted measurement, quantizes it and sends it to a fusion node which produces the estimate of the parameter $\hat{\theta}$ based on the received noisy M -tuple $\mathbf{Q}_r = (Q_1, \dots, Q_M)$ or $\hat{\mathbf{z}}_r = (\hat{z}_1, \dots, \hat{z}_M)$, where Q_i and $\hat{z}_i$ are the quantization index and the corresponding reconstruction value sent from the i -th node, respectively. Denote $\alpha_i(\cdot)$ the encoder at sensor i , which generates a quantization index $Q_i \in I_i = \{1, \dots, 2^{R_i}\}$. In what follows Q_i will be used to denote both a quantization index and the quantization interval $[Q_{i,l}, Q_{i,h}]$ where the noisy measurement z_i at node i belongs.

Throughout the paper we assume that there is only one-way communication from sensor nodes to fusion node, i.e., there is no feedback channel and the sensors do not communicate with each other; we also assume that the fusion node receives $\mathbf{Q}_r$ without transmission error.

2.1 Distributed Estimation Based on Quantized Data

We first consider the case of no measurement noise ($\omega_i = 0$ in (1)). Since each quantized sensor reading corresponds to a region where a parameter can be found, all quantized sensor readings lead to a partition of the parameter space S obtained by intersecting the regions corresponding to each Q_i . Thus, estimation based on the quantized M -tuple $(Q_1, \dots, Q_M)$ is conducted by constructing an intersection region A as follows:

$$A = \bigcap_{i=1}^M A_i, \quad A_i = \{\theta : f_i(\theta) \in Q_i\}, \quad (2)$$

$$\hat{\theta} = \hat{\theta}(Q_1, \dots, Q_M) = E(\theta | \theta \in A), \quad (3)$$

where A_i is the region corresponding to Q_i where the parameter θ is expected to be found and $\hat{\theta} = \hat{g}(\cdot)$ is the estimator that operates on quantized data. Then (3) would be the MMSE estimate of the parameter in the noiseless case. If the probability density function $p(\theta)$ associated with our a priori knowledge about the parameter is uniform, $\hat{\theta}$ would simply be the sample mean in A .

Now, consider the noisy case where there is no guarantee that the construction in (2) always leads to the non-empty intersection region A . Assuming that the statistics of the measurement noise w_i are known, we can formulate the problem as an MMSE estimation problem:

$$\begin{aligned}\hat{\theta} &= \int_{\theta \in S} \theta p(\theta | \mathbf{Q}_r) d\theta \\ &= \frac{\int_{\theta \in S} \theta p(\mathbf{Q}_r | \theta) p(\theta) d\theta}{p(\mathbf{Q}_r)}\end{aligned}\quad (4)$$

$$= \frac{\int_{\theta \in S} \theta \left[\prod_{i=1}^M p(Q_i | \theta) \right] p(\theta) d\theta}{p(\mathbf{Q}_r)}\quad (5)$$

$$= \frac{\int_{\theta \in S} \theta \left[\prod_{i=1}^M \int_{z_i \in Q_i} p(z_i | \theta) dz_i \right] p(\theta) d\theta}{p(\mathbf{Q}_r)},\quad (6)$$

where (5) follows from the measurement noise w_i independent of (θ) and $p(\mathbf{Q}_r) = \int_{\theta \in S} \left[\prod_{i=1}^M \int_{z_i \in Q_i} p(z_i | \theta) dz_i \right] p(\theta) d\theta$. The conditional probability $p(z_i | \theta)$, $i = 1, \dots, M$ in (6) can be obtained using the sensor model in (1): that is, $p(z_i | \theta)$ is normal distributed $N(f_i(\theta), \sigma_i^2)$. When $p(z_i | \theta)$ and $p(\theta)$ are available, (6) provides the MMSE estimator. However, computation of $\hat{\theta}$ in (6) is highly complicated, yielding drawbacks in real-time applications.

In this work, it is assumed that a prior probability of the parameter is available for estimation and if there is ignorance about the parameter, the best strategy would be to use an uninformative or uniform prior distribution, since this allows us to obtain a posteriori distribution which will be approximately proportional to the likelihood. In addition, the uninformative prior has another advantage of keeping computations relatively simple.

3. Construction of Region-Based Estimation Framework

It is noted that typical estimators such as ML-E and MMSE-E are supposed to make a full search of the parameter space S to find the estimate of the parameter, leading to a huge computational complexity. However, if we can find the region where the parameter is very likely to be located and perform the estimation within the region, then we would achieve a substantial reduction in computation. In this section, in order to make an efficient search for the region, we suggest partitioning the parameter space prior to the region search by constructing a region-based estimation framework. Furthermore, since the region search will be the most computational process in this framework, we propose the two practical techniques to reduce the search complexity without causing much loss in performance in Sect. 4.

We first consider the case of no measurement noise for which only certain M -tuples can be produced by the sensor nodes and those M -tuples can be linked to non-empty intersection regions in S . We construct the set of M -tuples S_Q^f that can be generated in a noise-free environment as follows:

$$S_Q^f = \{(Q_1, \dots, Q_M) | f_i(\theta) \in Q_i, \forall i, \theta \in S\}.\quad (7)$$

For this case, the region-based estimation framework can be simply constructed by obtaining the intersection region from the received M -tuple $\mathbf{Q}_r \in S_Q^f$ from (2) and (3). Denote A^j the intersection region constructed from the j -th M -tuple $\mathbf{Q}_j^f \in S_Q^f$. Then, $A^j \cap A^i = \emptyset$, if $j \neq i$, so that we can partition the parameter space S into $|S_Q^f|$ regions. Note that our framework can be employed for many sensor types; the only assumption we make is that quantization of measurements at the sensor nodes leads to a partition of the parameter space $S \subset \mathbf{R}^N$.

We now consider the noisy case for which there is no guarantee that $\mathbf{Q}_r$ will belong to S_Q^f . Our approach is to first identify the most likely intersection region corresponding to an M -tuple in S_Q^f , given the noisy observation $\mathbf{Q}_r$ and then compute the estimate of the parameter within the chosen region. To construct the region-based estimation framework for this case, we first denote H_j the hypothesis associated with the j -th region A^j and its vector reading $\mathbf{Q}_j^f$ and employ the maximum a posteriori (MAP) criterion to find H^* that is the most likely to occur, given the received noisy M -tuple $\mathbf{Q}_r$ as follows:

$$\begin{aligned}H^* &= \arg \max_j p(H_j | \mathbf{Q}_r), \quad j = 1, \dots, |S_Q^f| \\ &= \arg \max_j p(\mathbf{Q}_r | H_j) p_j \\ &= \arg \max_j p(\mathbf{Q}_r | \theta \in A^j) p_j \\ &= \arg \max_j \prod_{i=1}^M p(Q_i | \theta \in A^j) p_j,\end{aligned}\quad (8)$$

where $p_j = p(\theta \in A^j)$ and $p(Q_i | \theta \in A^j)$ is computed as

$$\begin{aligned}p(Q_i | \theta \in A^j) &= p(f_i(\theta) + w_i \in Q_i | \theta \in A^j) \\ &= \int_{\theta \in A^j} \left[\Phi\left(\frac{Q_{i,h} - f_i(\theta)}{\sigma_i}\right) - \Phi\left(\frac{Q_{i,l} - f_i(\theta)}{\sigma_i}\right) \right] p_j(\theta) d\theta\end{aligned}\quad (9)$$

where $p_j(\theta) = p(\theta | \theta \in A^j)^\dagger$ and $\Phi(\cdot)$ is the cumulative distribution function for the normal distribution $N(0, 1)$. Once H^* is obtained, the parameter estimate $\hat{\theta}$ is computed by $E(\theta | H^*) = E(\theta | \theta \in A^*)$, which clearly becomes the MMSE estimator in a noiseless case.

4. Two Proposed Estimation Techniques with Reduced Complexity

The most computationally intensive step in our estimation

[†]Note that $p_j(\theta)$ and p_j should be differentiated without confusion since p_j is not a function of θ .

framework is the integration in (9), which should be conducted for each hypothesis. To simplify the process, we first propose to reduce the number of hypotheses to evaluate, given $\mathbf{Q}_r$. Note that even in noisy cases, if the parameter is in A^j corresponding to $\mathbf{Q}_j^f \in S_Q^f$, the received M -tuple $\mathbf{Q}_r$ is likely to be $\mathbf{Q}_j^f$. From this consideration, we construct the set $A_s(\subset S)$ such that $p(\theta \in A_s | \mathbf{Q}_r) \approx 1$, i.e., the set containing the hypotheses that are likely to happen. By considering only the hypotheses in A_s , we can reduce the computational complexity for estimation. Construction of A_s can be accomplished by noting that given the parameter θ and the received noisy M -tuple $\mathbf{Q}_r$, it would be very likely that $p(\theta \in A^i | \mathbf{Q}_r) > p(\theta \in A^j | \mathbf{Q}_r)$ as long as $p(\hat{\theta}_i | \mathbf{Q}_r) \gg p(\hat{\theta}_j | \mathbf{Q}_r)$, where $\hat{\theta}_j$ is the centroid of the set A^j . Thus, we first find a coarse estimate of the source location $\hat{\theta}_c$ by determining which of the centroids of all A^j is the most likely, i.e., we compute

$$\begin{aligned} \hat{\theta}_c &= \arg \max_{\hat{\theta}_j} p(\hat{\theta}_j | \mathbf{Q}_r), \quad j = 1, \dots, |S_Q^f| \\ &= \arg \max_{\hat{\theta}_j} p(\mathbf{Q}_r | \hat{\theta}_j) p_j \\ &= \arg \max_{\hat{\theta}_j} \prod_{i=1}^M \left[\Phi \left(\frac{Q_{i,h} - f_i(\hat{\theta}_j)}{\sigma_i} \right) - \Phi \left(\frac{Q_{i,l} - f_i(\hat{\theta}_j)}{\sigma_i} \right) \right] p_j \end{aligned}$$

and then construct $A_s(\delta) = \bigcup_{k=1}^K A^k$ such that $|\hat{\theta}_c - \hat{\theta}_k| < \delta$, $\forall k = 1, \dots, K$ where $\hat{\theta}_k = E(\theta | \theta \in A^k)$, i.e., we include all of the hypotheses that are within a certain distance of our coarse estimate:

$$H^* = \arg \max_k p(H_k | \mathbf{Q}_r), \quad H_k \in A_s, \quad k = 1, \dots, K \quad (10)$$

We select the most likely hypothesis following (8) but among only those candidates in $A_s(\delta)$. Thus, by choosing δ we can control the trade-off between search complexity and estimation accuracy (a smaller δ will lead to a faster search, since fewer hypotheses will be tested). Notice that K would take a small value ($\ll |S_Q^f|$) with a good choice of δ , depending upon the applications.

Second, we propose a simple minimum distance technique that allows us to find the most likely hypothesis H^* without computing the integration in (9). This technique is based on the additive noise model for quantization error which has been shown to be a valid model in cases of high resolution quantization with smooth probability density functions [3]. Specifically, denoting $Q(\cdot)$ a quantization function at each sensor node, we have

$$\begin{aligned} \hat{z}_i &= Q(z_i) = Q(f_i(\theta) + \omega_i) = f_i(\theta) + \omega_i + e_i \\ &= f_i(\theta) + e_i + \omega_i \approx Q(f_i(\theta)) + \hat{\omega}_i \\ &= \hat{z}_i^f + \hat{\omega}_i \end{aligned} \quad (11)$$

where $\hat{z}_i^f$ is the reconstruction value generated in a noiseless case at the i -th sensor node, e_i is the quantization error induced from the additive noise model and $\hat{\omega}_i$ is the equivalent measurement noise that can be approximated by a normal distribution. Note that each combination of M -tuple

$\hat{z}_i^f, i = 1, \dots, M$ will have one-to-one mapping to the corresponding element in S_Q^f . Denoting $\hat{\mathbf{z}}_j^f = (\hat{z}_1^f, \dots, \hat{z}_M^f)$ the j -th M -tuple of the reconstruction values (corresponding to $\mathbf{Q}_j^f$ and thus H_j) that could be generated in a noise-free case, we calculate the distance $\|\hat{\mathbf{z}}_r - \hat{\mathbf{z}}_j^f\|^2$ where $\hat{\mathbf{z}}_r$ is the received noisy M -tuple of reconstruction values and find $\hat{\mathbf{z}}_j^f$ with the minimum distance which corresponds to the most likely hypothesis H^* . Formally,

$$\begin{aligned} H^* &= \arg \max_j p(H_j | \hat{\mathbf{z}}_r), \quad j = 1, \dots, |S_Q^f| \\ &= \arg \max_j p(\hat{\mathbf{z}}_r | \theta \in A^j) p_j \\ &= \arg \max_j p(\hat{\mathbf{z}}_r | \hat{\mathbf{z}}_j^f) p_j \\ &= \arg \max_j \left[\prod_{i=1}^M p(\hat{z}_i | \hat{z}_i^f) \right] p_j \\ &= \arg \min_j \|\hat{\mathbf{z}}_r - \hat{\mathbf{z}}_j^f\|^2 - \ln p_j, \end{aligned} \quad (12)$$

where (12) follows from (11) and the assumption that $\hat{\omega}_i$ is approximated as a normal distribution $N(0, \hat{\sigma}_i^2 = \hat{\sigma}^2)$. Notice that this technique could allow us to detect H^* that maximizes the posterior distribution, provided that (11) is validated. It will obviously produce a substantial reduction in the search complexity but would be justified at a high rate ($R_i \gg 1$) or in the high SNR regime. Thus it may generate erroneous detections of H^* in an operating rate or SNR, especially for the parameters θ where $\|\hat{\mathbf{z}}_r(\theta) - \hat{\mathbf{z}}_j^f\|^2 > \delta_d > 0$. In this work, it is suggested that estimation for such parameter samples is executed by using the first proposed technique in (10) to avoid possible wrong detections.

5. Analysis of Computational Complexity

In this section, we discuss the computational complexity of the proposed algorithms by comparing with ML-E and MMSE-E denoted by $\hat{\theta}_{ML-E}$ and $\hat{\theta}_{MMSE-E}$, respectively. First, the MMSE estimation in (4) can be further manipulated as follows:

$$\begin{aligned} \hat{\theta}_{MMSE-E} &= \frac{\int_{\theta \in S} \theta p(\mathbf{Q}_r | \theta) p(\theta) d\theta}{p(\mathbf{Q}_r)} \\ &= \frac{\sum_j^{|S_Q^f|} \int_{\theta \in A^j} \theta p(\mathbf{Q}_r | \theta) p(\theta) d\theta}{p(\mathbf{Q}_r)} \\ &= \frac{\sum_j^{|S_Q^f|} \int_{\theta \in A^j} \theta p(\mathbf{Q}_r | \theta) p_j(\theta) p_j d\theta}{p(\mathbf{Q}_r)} \end{aligned}$$

Note that the computational complexity required to test all of the hypotheses is given by $\sum_j^{|S_Q^f|} \int_{\theta \in A^j} p(\mathbf{Q}_r | \theta) p_j(\theta) p_j d\theta$, implying that our region-based framework would facilitate a much faster estimation than MMSE-E. Furthermore, since the first proposed estimator in (10) (namely, the MAP-based estimator (MAP-E) or $\hat{\theta}_{MAP-E}$) is devised to test only the hypotheses within a distance δ , it will permit an estimation

process with a significantly reduced complexity. More precisely, the reduction can be approximately expressed by using $C_{\hat{\theta}}$ which indicates the complexity of the estimator $\hat{\theta}$:

$$C_{\hat{\theta}_{MAP-E}} \approx \frac{2\pi\delta^2}{|S|}C \quad (13)$$

where C is the complexity for test of all of the hypotheses. For example, in our experiments with $\delta = 1m$ in a $10 \times 10m^2$ sensor field, almost 94% ($C_{\hat{\theta}_{MAP-E}} = \frac{2\pi\delta^2}{|S|}C = \frac{2\pi \cdot 1^2}{10 \times 10}C$) reduction in computation can be achieved with respect to C .

In particular, the second proposed technique in (12) (namely, the MAP-based minimum distance estimator (MAP-MDE) or $\hat{\theta}_{MAP-MDE}$) selects the most likely hypothesis by simply finding the minimum distance between the received data $\hat{\mathbf{z}}_r$ and $\hat{\mathbf{z}}_j^f$, $j = 1, \dots, |S_Q^f|$. Thus, its computational complexity is comparable to ML-E which is equivalent to finding the parameter with the minimum distance between the received data and their noiseless measurements; formally,

$$\begin{aligned} \hat{\theta}_{ML-E} &= \arg \max_{\theta \in S} p(\mathbf{Q}_r|\theta) \\ &= \arg \min_{\theta \in S} \sum_i^M [\hat{z}_i - f_i(\theta)]^2 \end{aligned}$$

Clearly, since there are no integration and probability calculation involved for ML-E and MAP-MDE, their complexity will be greatly lowered. It should be observed that at a high rate or in the high SNR, our second estimator $\hat{\theta}_{MAP-MDE}$ can perform an accurate detection of hypotheses with high probability, enabling a faster operation than ML-E. However, as the rate or SNR decreases, more samples will show $\|\hat{\mathbf{z}}_r(\theta) - \hat{\mathbf{z}}_s^f\|^2 > \delta_d$ and thus should be estimated by MAP-E to maintain a good performance. For example, in our experiments with $\delta_d = 1m$, $R_i = 3$ and SNR=50dB, almost 10% samples are processed by MAP-E, indicating that MAP-MDE operates over 10 times faster on the average than MAP-E.

6. Application to Acoustic Amplitude Sensor Case

As an example of the application, we consider source localization for acoustic amplitude sensor networks where an energy decay model of sensor signal readings proposed in [10] is used for localization. Note that the energy decay model was verified by the field experiment in [10] and was also used in [4], [7], [12]. Each sensor node measures signal energy generated by a source located at an unknown location $\theta \in \mathbf{R}^2$. This model is based on the fact that the acoustic energy emitted omnidirectionally from a sound source will attenuate at a rate that is inversely proportional to the square of the distance in free space [16]. When an acoustic sensor is employed at each node, the signal energy measured at node i and denoted by z_i , can be expressed as follows:

$$z_i(\theta) = g_i \frac{a}{\|\theta - \mathbf{x}_i\|^\alpha} + w_i, \quad (14)$$

Fig. 1 Source Localization based on quantized energy readings.

where the acoustic sensor model consists of the gain factor of the i -th node g_i , an energy decay factor α which is approximately equal to 2 in free space, and the source signal energy a . The measurement noise term w_i can be approximated using a normal distribution $N(0, \sigma_i^2)$. It is assumed that the signal energy a is uniformly distributed over the range $[a_{min}, a_{max}]$.

Assuming that the source signal energy a is known[†], localization based on quantized sensor readings can be illustrated by Fig. 1, where the ring-shaped area A_i , $i = 1, \dots, M$ corresponds to one quantized measurement at the i -th sensor and can be constructed as follows:

$$A_i = \{\theta : g_i \frac{a}{\|\theta - \mathbf{x}_i\|^\alpha} \in Q_i\} \quad (15)$$

By computing the intersection of all the ring areas (one per sensor), it is possible to define the area where the source is expected to be located (see (2) and (3)). Note that for this application, at least three quantized measurements are required to achieve a connected intersection region.

7. Simulation Results

In our experiments, we consider an acoustic sensor network with $M(= 3, 4, 5)$ nodes deployed randomly in a $10 \times 10 m^2$ sensor field. Each sensor measures an acoustic source energy based on the energy decay model in (14), quantizes it using a quantizer of $R_i(= 2, 3, 4)$ bits designed by the algorithm in [8]. The measurement noise is assumed to be normal distributed, $N(0, \sigma_i^2 = \sigma^2)$ and the SNR is computed by $10 \log_{10} \frac{a^2}{\sigma^2}$. Note that the SNR is measured at 1 meter from the source. For example, SNR=40dB corresponds to SNR $\approx$ 5.5dB measured at each sensor location on the average.^{††} Extensive simulation is conducted to compare the performance of various estimation algorithms and to inves-

[†]In practice, the signal energy is unknown and should be jointly estimated along with the source location as described in [7].

^{††}For practical vehicle target, it is often much higher than 40dB. A typical value of the variance of measurement noise σ^2 is 0.05^2 (= 60dB) [10], [12].

tigate the sensitivity of the proposed algorithms to perturbation of the model parameters. Furthermore, a larger sensor network is also considered for testing our estimators since typical sensor networks involve many sensor nodes in large sensor fields. In the experiments, the simulation parameters δ and δ_d are set as 1 meter for our proposed estimators.

7.1 Comparison with Typical Estimators

In this experiment, 100 different sensor configurations are generated for $M (= 3, 4, 5)$ and each sensor node employs a quantizer of $R_i (= 3, 4, 5)$ bits and for each configuration, the proposed estimators in Sect. 4 are evaluated by generating a test set of 1,000 source locations from a uniform distribution of $p(\theta)$, the model parameters given by $a = 50$, $\alpha = 2$, $g_i = 1$ and the measurement noise $\sigma_i^2 = \sigma^2 = 0.15$ ($= 50$ dB). In Fig. 2, the localization error is averaged over 100 configurations and compared with ML-E and MMSE-E. Note that MMSE-E gives us a good lower bound at the cost of high computational complexity.

We also investigate the sensitivity to noise level by generating a test set of 1,000 source locations for each 5-sensor configuration with $R_i = 3$ and SNR in the range from 40 dB to 100 dB by varying σ . The localization results averaged over 100 configurations are plotted in Fig. 3. These experimental comparisons demonstrate the significant advantage of the proposed estimators over the typical estimators in terms of both computational complexity and localization accuracy. Moreover, as mentioned in Sect. 4, MAP-MDE shows much faster execution time than MAP-E while maintaining comparable performance. Note that the parameter δ specifies the size of the region to be searched for hypotheses testing and δ_d determines the number of parameter samples to be estimated by MAP-E, both affecting compu-

tational complexity. With larger values of the parameters, the localization performance would be improved at the cost of increased complexity.

7.2 Sensitivity to Parameter Perturbation

In this experiment, we examine how the performance of the proposed algorithms can be affected by various types of the model parameter mismatches. Specifically, we evaluate the estimation techniques by generating test sets with respect to the presence of the perturbation of α , g_i and $\mathbf{x}_i = (x, y)$. In applying the proposed estimators, the prior distribution $p(\theta)$ is assumed to be uniform and we give random perturbation to one of the sensor model parameters for each test set of 1,000 source locations generated with $R_i = 3$ bits and SNR=60 dB: that is, the actual value of α is varied from $2 - \Delta\alpha$ to $2 + \Delta\alpha$ for $\Delta\alpha = 0.2$ and the actual gain is also varied from $1 - \Delta g$ to $1 + \Delta g$ for $\Delta g = 0.4$. Similarly, each sensor location (x, y) is generated from $x - \Delta x$ to $x + \Delta x$ and from $y - \Delta y$ to $y + \Delta y$ for $\Delta x = \Delta y = 0.4$, respectively.

The localization results for MAP-E and MAP-MDE are tabulated in Table 1. It is observed that the localization error shows worse performance in the cases of $\Delta g < 0$ (e.g., $g_i = 1 - 0.4 = 0.6$) than the cases of $\Delta g > 0$ (e.g., $g_i = 1 + 0.4 = 1.4$) because the sensor measurements are less amplified before being added to the measurement noise, resulting in relatively lower SNR than the positive Δg s. For the cases of $\pm\Delta\alpha$, there are also slightly different results which depend mainly upon quantizer design algorithms rather than localization techniques. When it comes to the impact of $(\pm\Delta x, \pm\Delta y)$, there is no noticeable difference in the simulation results (minor variations in the localization error can be typically seen from experiments using random measurement data). Nonetheless, it can be said that small perturbation can be allowed to maintain good localization accuracy.

Fig. 2 Comparison of proposed estimators in Sect. 4 with MMSE-E and ML-E. The average localization error is plotted vs. the number of sensors M with $R_i = 3$ bits and SNR=50 dB (left) and vs. the number of bits R_i , assigned to each sensor with $M = 5$ and SNR=50 dB (right).

Fig. 3 Comparison of proposed estimators in Sect. 4 with MMSE-E and ML-E. The average localization error is plotted vs SNR =40,...,100 dB with $M = 5$ and $R_i = 3$ bits.

Table 1 Localization Error (LE) (m) of proposed estimators under various parameter mismatches. $LE = \frac{1}{100} \sum_{l=1}^{100} E_l(\|\theta - \hat{\theta}\|^2)$, where E_l is the average localization error for the l -th 5-sensor configuration with $R_i = 3$ and SNR=60 dB.

$\Delta\alpha$	-0.2	-0.15	-0.1	-0.05	0	0.05	0.1	0.15	0.2
LE(MAP-E)	1.4512	1.1246	0.8540	0.6391	0.5208	0.5831	0.7480	0.9739	1.2394
LE(MAP-MDE)	1.2227	1.0043	0.8341	0.6932	0.6121	0.6547	0.7579	0.9090	1.0582
Δg_i	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4
LE(MAP-D)	1.9459	1.3413	0.8807	0.6040	0.5100	0.6362	0.8443	1.0751	1.3382
LE(MAP-MDE)	1.5205	1.1132	0.8275	0.6635	0.6089	0.6836	0.8099	0.9550	1.1207
$(\Delta x, \Delta y)$	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4
LE(MAP-E)	0.7677	0.6582	0.5738	0.5233	0.5022	0.5308	0.5968	0.6785	0.7958
LE(MAP-MDE)	0.7466	0.6876	0.6440	0.6113	0.5987	0.6169	0.6583	0.7134	0.7812

Fig. 4 Average localization error (m) vs. number of sensors ($M = 12, 16, 20$) in a larger sensor field $20 \times 20 m^2$. The proposed estimators are compared with traditional estimators with $R_i = 3$ bits and SNR=50 dB.

7.3 Performance Analysis in a Larger Sensor Network

In this experiment, 20 different sensor configurations in a larger sensor field, $20 \times 20 m^2$, are generated for $M = 12, 16, 20$ and for each sensor configuration, our estimators are tested with a given rate of $R_i = 3$ bits and SNR=50 dB in Fig.4. As expected, the proposed estimators show good performance with respect to the traditional estimators. It is noted that the system with a larger number of sensors outperforms the system with a smaller number of sensors ($M = 3, 4, 5$) although the sensor density is kept the same. In the experiments, the sensor density for $M = 20$ in $20 \times 20 m^2$ given by $\frac{20}{20 \times 20} = 0.05$ is equal to that for $M = 5$ in $10 \times 10 m^2$. This additional performance gain is achieved because localization performance degrades around the edges of the sensor field and thus in a larger sensor field there is a relatively smaller number of source locations near the edge as compared to a smaller field with the same sensor density.

8. Conclusions

In this paper, we have constructed a region-based estima-

tion framework for general distributed systems which is efficiently used to search the region with maximum likelihood, instead of attempting to estimate the parameter directly. In this framework, we have proposed two practical search techniques so as to yield a significant reduction in computational complexity and showed that the proposed algorithms achieve good performance with reasonable complexity as compared with typical standard estimators such as the ML and MMSE estimators. In the future, we will work on joint designs of compression and estimation algorithms optimized for distributed systems.

Acknowledgments

This study was supported by a research fund from Chosun University, 2014. The author would like to thank the Associate Editor and the anonymous reviewers for their careful reading of the manuscript and useful suggestions which led to significant improvements in the manuscript.

References

- [1] T. Aysal and K. Barner, "Constrained decentralized estimation over noisy channels for sensor networks," IEEE Trans. Signal Process., vol.56, no.4, pp.1398–1410, April 2008.
- [2] R.M. Gray, "Quantization in task-driven sensing and distributed processing," IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP), Toulouse, France, May 2006.
- [3] R.M. Gray and D.L. Neuhoff, "Quantization," IEEE Trans. Inf. Theory, vol.44, no.6, pp.2325–2383, Oct. 1998.
- [4] A.O. Hero and D. Blatt, "Sensor network source localization via projection onto convex sets (POCS)," IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP), March 2005.
- [5] Y.H. Kim, "Distributed estimation based on quantized data," IEICE Electronics Express, vol.8, pp.699–704, May 2011.
- [6] Y.H. Kim, "Functional quantizer design for source localization in sensor networks," EURASIP J. Advances in Signal Processing, vol.2013, no.151, 2013.
- [7] Y.H. Kim and A. Ortega, "Maximun a posteriori (MAP)-based algorithm for distributed source localization using quantized acoustic sensor readings," IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP), May 2006.
- [8] Y.H. Kim and A. Ortega, "Quantizer design for source localization in sensor networks," IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP), pp.857–860, March 2005.
- [9] W. Lam and A. Reibman, "Design of quantizers for decentralized estimation systems," IEEE Trans. Commun., vol.41, no.11, pp.1602–1605, Nov. 1993.
- [10] D. Li and Y.H. Hu, "Energy-based collaborative source localization

- using acoustic microsensor array,” *EURASIP J. Applied Signal Processing*, pp.321–337, 2003.
- [11] J. Li and G. AlRegib, “Rate-constrained distributed estimation in wireless sensor networks,” *IEEE Trans. Signal Process.*, vol.55, no.5, pp.1634–1643, May 2007.
 - [12] J. Liu, J. Reich, and F. Zhao, “Collaborative in-network processing for target tracking,” *EURASIP J. Applied Signal Processing*, pp.378–391, 2003.
 - [13] S. Marano, V. Matta, and P. Willett, “Asymptotic design of quantizers for decentralized MMSE estimation,” *IEEE Trans. Signal Process.*, vol.55, no.11, pp.5485–5496, Nov. 2007.
 - [14] V. Misra, V.K. Goyal, and L.R. Varshney, “High-resolution functional quantization,” *Data Compression Conference*, pp.113–122, Snowbird, UT, March 2008.
 - [15] R. Niu and P.K. Varshney, “Target location estimation in wireless sensor networks using binary data,” *38th Annual Conference on Information Sciences and Systems*, Princeton, NJ, March 2004.
 - [16] T.S. Rappaport, *Wireless communications: principles and practice*. Prentice-Hall, New Jersey, 1996.
 - [17] X. Shen, Y. Zhu, and Z. You, “An efficient sensor quantization algorithm for decentralized estimation fusion,” *Automatica*, vol.47, no.5, pp.1053–1059, May 2011.
 - [18] S. Vardeman and C.-S. Lee, “Likelihood based statistical estimation from quantized data,” *IEEE Trans. Instrum. Meas.*, vol.54, pp.409–414, Feb. 2005.
 - [19] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G.B. Giannakis, “Distributed compression-estimation using wireless sensor networks,” *IEEE Signal Process. Mag.*, vol.23, no.4, pp.27–41, July 2006.

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, Chosun University, Korea. 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.