ISSN : 1226-7244 (Print)
ISSN : 2288-243X (Online)
논문번호 18-03-05

j.inst.Korean.electr.electron.eng.Vol.22,No.3,558~564,September 2018
http://dx.doi.org/10.7471/ikeee.2018.22.3.558
37

A novel approach to design of local quantizers for distributed estimation

Yoon Hak Kim*★

Abstract

In distributed estimation where each node can collect only partial information on the parameter of interest without communication between nodes and quantize it before transmission to a fusion node which conducts estimation of the parameter, we consider a novel quantization technique employed at local nodes. It should be noted that the performance can be greatly improved if each node can transmit its measurement to one designated node (namely, head node) which can quantize its estimate using the total rate available in the system. For this case, the best strategy at the head node would be simply to partition the parameter space using the generalized Lloyd algorithm, producing the global codewords, one of which is closest to the estimate is transmitted to a fusion node. In this paper, we propose an iterative design algorithm that seeks to efficiently assign the codewords into each of quantization partitions at nodes so as to achieve the performance close to that of the system with the head node. We show through extensive experiments that the proposed algorithm offers a performance improvement in rate-distortion perspective as compared with previous novel techniques.

Key words : Distributed compression, distributed source coding (DSC), distributed estimation, quantizer design, generalized Lloyd algorithm, source localization, sensor networks

Ⅰ. Introduction

In distributed systems where battery-operated sensor nodes randomly placed in a sensor field gather measurements from the parameter of interest, quantize the measurements and transmit to a fusion node which executes estimation using the received quantized data, various design techniques for local quantizers have been proposed in order to improve the estimation performance. In most of the designs, the generalized Lloyd algorithm has been employed to produce such local quantizers at nodes for distributed systems. Notably, designing local quantizers that optimize the global metric such as estimation error in this Lloyd framework without exchanging measurements between nodes would yield design difficulty since the global metric to be optimized should be a function of all local measurements. Specifically, encoding of local measurements into one of the quantization partitions constructed to minimize the global cost function cannot be properly performed without computation of the metric.

In order to make the algorithm manageable, the cost functions should be guaranteed to be non-increasing at iterations while maintaining

* Dept. of Electronic Engineering, Chosun University
★ Corresponding author
E-mail : yhk@chosun.ac.kr, Tel : +82-62-230-7129
※ Acknowledgment
This work was supported by a research fund from Chosun University, 2017.
Manuscript received Aug. 9, 2018; revised, Sep. 8, 2018; accepted Sep. 11, 2018

independent encoding at nodes. For instance, the probabilistic distance was presented under two hypotheses for distributed detection systems to ensure a controllable design process [11]. In acoustic sensor networks, the weighted cost function was suggested in the Lloyd design. The weight is iteratively searched so as to make the metric non-increasing and independent encoding possible [5], [8].

To further improve the design performance, quantizers for distributed systems could be designed in a *non-regular* structure where multiple disjoint partitions are assigned into the single codeword [13]. A novel algorithm that systematically transforms from regular quantizers to non-regular ones was proposed for source localization in acoustic sensor networks [7]. The non-regular design was incorporated into the Lloyd design framework where multiple quantization partitions are iteratively mapped to a single codeword to reduce the estimation error [2]. It was demonstrated that independent encoding can be ensured by allowing multiple codewords for each of quantization partitions, yielding a substantial performance gain [4]. In addition, an independent encoding method that relates global codewords to local ones was presented for distributed estimation [3].

Consider a distributed system with a head node which can collect the measurements from the nodes involved via local communication between nodes and transmit to a fusion node with limited transmission rate (say R). Then the best strategy would be not to quantize the received measurements but to simply cluster the parameter space into $K = 2^R$ egions where the global codeword is computed as a centroid of the region and send one of the codewords closest to the estimate of the parameter to the fusion node. However, in a situation we consider in this work, the parameter is not directly accessible at nodes and each node can measure only partial information on the parameter and transmit its quantized data to a fusion node without communicating with other nodes. Obviously, there would be a severe performance degradation due to increased uncertainty of the parameter at nodes which would be eliminated by allowing local communication between nodes. Furthermore, since each node can use the partial rate R_i where $\Sigma R_i = R$, to send its quantized measurement, the quantization algorithm that should conduct independently of the measurements at other nodes would lead to one of suboptimal design techniques as compared to the system with a head node.

In this work, we seek to find an iterative design algorithm that assigns K global codewords into $L_i = 2^{R_i}$ set of the codewords at each node while reducing the estimation error at iterations. In addition, we present the efficient independent encoding techniques that enable mapping of local measurements into one of the pre-computed codeword sets in real situations. Note that the fusion node can estimate the parameter by computing a centroid of the intersection of the received codeword sets, which are transmitted from the sensor nodes involved. We finally conduct experiments to demonstrate a noteworthy performance improvement of the proposed algorithm with respect to previous design techniques.

The main contributions of this work are: (a) we suggest quantization of the parameter space, not the measurements, which is motivated by the assertion that the global metric can be efficiently optimized by directly partitioning the parameter space even when only local measurements are available at nodes; (b) we propose a novel quantization technique that produces global codewords to be transmitted, which would be more likely to represent the quantization partitions of the parameter space; (c) we finally present the independent encoding rule which finds the corresponding quantization partitions constructed in the design process conducted prior to the real operation.

This paper is organized as follows. The problem formulation is presented in Section Ⅱ. The motivation for this work is given in Section Ⅲ and the partitioning of the parameter space and the proposed algorithm are elaborated in Section Ⅳ. The algorithm is summarized in Section Ⅴ and an application system for our algorithm is introduced in Section Ⅵ. Experimental results are provided in Section Ⅶ and conclusions are given in Section Ⅷ.

Ⅱ. Problem Formulation

In a distributed system where M sensor nodes are randomly placed at known locations, $\mathbf{x}_i \in \mathbf{R}^2$, $i = 1,\cdots,M$ in a sensor field $S \subset \mathbf{R}^N$, each node measures its sensor reading $z_i$on the unknown parameter Θ which is given by:

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

where $f_i(\Theta)$ is a sensing model employed at nodes and the measurement z_i is typically noise-corrupted by the noise ω_i assumed to be approximated by normal distribution $N(0, \sigma_i^2)$. The i-th node quantizes its measurement z_i by using a R_i- bit quantizer with quantization level $L_i = 2^{Ri}$:that is, the quantizer encodes its measurement z_i to one of the predetermined sets of the codewords $V_i^j(\hat{\theta}_k), j = 1,\ldots,L_i$ if the mapping of z_i to the j-th set minimizes the cost function such as the estimation error. In this work, we first generate the global codewords $\hat{\theta}_k, \ k = 1,\ldots,K = 2^R, \ R = \sum R_i$ using the well-known generalized Lloyd algorithm [12] and given those codewords, we seek to partition them into L_i sets at nodes by ensuring an efficient independent encoding technique.

Ⅲ. Motivation

Suppose that each node in distributed systems can collect only the measurement which contains partial information for estimation of the parameter. In other words, there are many applications where each node is not able to estimate the parameter based on only the local measurement, although the information obtained from the measurement would be related with the parameter. For example, in a source localization system in acoustic amplitude senor networks, each node senses the source energy and thus can measure the distance between nodes and source. Thus, it would be impossible to directly estimate the parameter (in this case, the source location) without the measurements from the other nodes. Quantization in most of previous work has been conducted in the measurement domain and quantized measurements would be transmitted to a fusion node which performs estimation using the quantized data. However, question arises that if each node can send direct information on the parameter, the estimation performance at fusion node would be greatly improved. The challenge here is how each node sends the direct information efficiently even when it has no direct access to the parameter. To solve the problem, we suggest partitioning of parameter space into a given number of regions and propose an efficient encoding scheme that allows us to map a local measurement into the corresponding subset of the regions that maximizes the global metric such as estimation performance at the fusion node.

Ⅳ. Partitioning of parameter space and encoding technique

We first make a partitioning of the parameter space into $K = 2^R$ regions using the generalized Lloyd algorithm, resulting in the global codewords $\{\hat{\theta}_k, k = 1,\ldots,K\}$, which will be iteratively clustered into $L_i = 2^{Ri}$ quantization partitions at nodes. Since we seek to encode the local measurement $z_i$into one of the partitions (i.e., sets of codewords), we should be able to find a set of the codeword $\hat{\theta}_k$ that is relatively most likely to occur given the local measurement z_i. In doing so, we regard each local measurement as a set of the corresponding parameter regions as follows:

$$C(z_i) = \{\hat{\theta}_k : \|\theta - \hat{\theta}_k\|^2 \le \|\theta - \hat{\theta}_l\|^2, \forall l \ne k, \in S(z_i)\} \quad (2)$$

where $S(z_i)$, the set of parameters that can most likely happen given z_i, is easily generated assuming noiseless condition in (1); that is, $S(z_i) = \{\theta : z_i = f_i(\theta),\ \theta \in S\}$, implying $P(\theta \in S(z_i) \mid z_i) \approx 1$. Clearly, $C(z_i)$ is the set of the global codewords closest to the parameters in the set $S(z_i)$. Once the construction of $S(z_i)$ and $C(z_i)$ for each sample measurement is completed, we can perform quantization process by assigning the codeword sets $C(z_i)$ into one of the quantization partitions $V_i^j, j = 1, \ldots, L_i$. Note that since quantization in this work deals with codeword sets $C(z_i)$ instead of measurements z_i, encoding into the j-th partition implies that the node would send its reconstructed codeword set $\hat{C}_i^j$ for estimation that represents the j-th quantization partition V_i^j, not the reconstructed value $\hat{z}_i^j$. Construction of such partitions is conducted as follows:

$$V_i^j = \{C(z_i) : \|\theta(z_i) - \hat{\theta}^{j*}\|^2 \le \|\theta - \hat{\theta}^{l*}\|^2, \forall l \ne k\} \quad (3)$$

where $\hat{\theta}^{j*}$ indicates the estimate of θ computed when the i-th node sends the j-th reconstructed codeword set $\hat{C}_i^j$. Specifically, the estimation based on the received M codeword sets during the design process is executed by simply taking the centroid as the estimate $\hat{\theta}^*$:

$$\hat{\theta}^* = E[\hat{\theta} \mid \hat{\theta} \in \hat{C}], \qquad \hat{C} = \cap_i^M \hat{C}_i \quad (4)$$

where $\hat{C}_i$ is the the reconstructed codeword set transmitted from node i and it is assumed that the intersection of M reconstructed codeword sets is not empty. However, when the intersection yields an empty set, the estimation can be also conducted by using M reconstructed values $\hat{z}_i, i = 1, \ldots, M$ and in this work, for this case, the maximum likelihood (ML) estimation is employed for fast computation.

Notice that the quantization algorithm is conducted using the training samples of local measurements prior to the real operation, whereas the resultant quantizers should operate only based on the single local measurement at each node, requiring an efficient independent encoding scheme; in other words, encoding of measurements into one of the pre-computed partitions without the measurement at the other nodes involved should be ensured. The proposed encoding that performs the mapping of z_i into the j-th partition V_i^j is provided below:

$$E_{\theta \in S(z_i)} \min_{\hat{\theta}_k \in \hat{C}_i^j} \|\theta - \hat{\theta}_k\|^2 \le E_{\theta \in S(z_i)} \min_{\hat{\theta}_k \in \hat{C}_i^l} \|\theta - \hat{\theta}_k\|^2, \forall l \ne j \quad (5)$$

It should be emphasized that the encoding is performed by computing the Euclidean distance $\|\theta - \hat{\theta}_k\|^2$, not the estimation error $\|\theta - \hat{\theta}^*\|^2$ where $\hat{\theta}^*$ is the estimate obtained from the M reconstructed codeword sets, $\hat{C}_1, \ldots, \hat{C}_M$ by using (4), yielding an low-weight encoding scheme. Note that the principle of the encoding is similar to that of the standard quantization where the measurements are encoded into one of the reconstructed codewords that is closest in Euclidean distance to them in the measurement domain.

Ⅴ. Proposed quantizer design algorithm

In this section, an iterative design algorithm that partitions the codewords $\{\hat{\theta}_k\}$ into $L_i$groups at node i is summarized as follows:

Algorithm: an iterative design algorithm at node *I*

Step 1: Initialize the local codewords $\hat{z}_i^j, j = 1, \ldots, L_i$ and construct $S(\hat{z}_i^j)$ and the reconstructed codeword sets $\hat{C}_i^j$ as follows:

$$\hat{C}_i^j = \{\hat{\theta}_k : \|\theta - \hat{\theta}_k\|^2 \le \|\theta - \hat{\theta}_l\|^2, \forall l \ne k, \theta \in S(\hat{z}_i^j)\} \quad (7)$$

Step 2: Group the codeword sets $C(z_i)$ constructed from (2) into $L_i$partitions $V_i^j, j=1,\dots,L_i$ so as to minimize the estimation error using (3)

Step 3: Update the reconstructed codeword sets $\hat{C}_i^j$ that minimize the estimation error over each partition with respect to $\hat{z}_i$:

$$\hat{C}_i^{j*}(\hat{z}_i^*) = \arg\min_{\hat{C}(\hat{z}_i)} E[\|\theta(z_i) - \hat{\theta}^*(\hat{C}_i(\hat{z}_i))\|^2 | C(z_i) \in V_i^j] \tag{8}$$

where $\hat{\theta}^*$ is computed by (4) with $\hat{C}_i = \hat{C}_i(\hat{z}_i)$.

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

Ⅵ. Application to source localization in acoustic sensor networks

For an application system, a source localization system in acoustic amplitude sensor networks is considered. It is assumed that M sensor nodes are randomly scattered in a sensor field S and each node equipped with acoustic amplitude sensor measures signal energy from a source and send its quantized signal energy to a fusion node for source localization. In this work, we employ an energy decay model for generating measurements at nodes, which was proposed and verified by the field experiment in [9]. Note that the sensor model was widely used for various applications [1], [10], [6]. The signal energy measurement at sensor i, denoted by $z_i$can be written as follows:

$$z_i(\theta) = g_i \frac{a}{\|\theta - x_i\|^{\alpha}} + \omega_i \tag{9}$$

where the sensing model in (1) is given by the energy decay model with source signal energy a, gain factor of the i-th sensor g_i, and energy decay factor α ($\approx$ 2). It is assumed that the energy measurement is noise corrupted and approximated by a normal distribution, $N(0, \sigma_i^2)$. In this work, it is also assumed that the signal energy is known to the fusion node for estimation but in real situations, the signal energy a is typically unknown and can be jointly estimated with its location (see [6] for the detail).

Ⅶ. Simulation results

We evaluate the performance of the proposed algorithm by considering a source localization system in acoustic sensor networks. We randomly deploy M (= 5) sensor nodes in $10 \times 10m^2$ sensor field for each of 100 configurations. For each configuration, we generate training samples from the model parameters a = 50, α = 2, g_i= 1 in (1), assuming the noiseless condition $\sigma_i^2 = \sigma^2 = 0$ and uniform distribution of source locations. Using the samples, we run the algorithm described in Section V to produce the proposed quantizer with a small partition level $K \leq 150$ for fast computation. Note that the proposed design provides performance improvement as K becomes large at the cost of increased complexity. We conduct typical designs such as uniform quantizers (Unif Q) and Lloyd 4 quantizers (Lloyd Q) for comparison and to further inspect our quantizer, we also design the previous novel quatizers. In investigating numerically the effectiveness of our algorithm, we generate the test sets of 1000 source locations with respect to noise level variation and calculate the average localization errors $E\|\theta - \hat{\theta}\|^2$ by using our simple estimation technique in (4) for our quantizer and MLE for the other ones, respectively.

To compare the proposed quantizers with typical designs (i.e., Unif Q and Lloyd Q), we compute the average localization error in meter over 100 node configurations by varying R_i.

In Figure 1, the rate-distortion (R-D) curves are depicted for performance evaluation. As expected, the proposed quantizer achieves a substantial localization performance improvement as compared with typical designs, indicating that

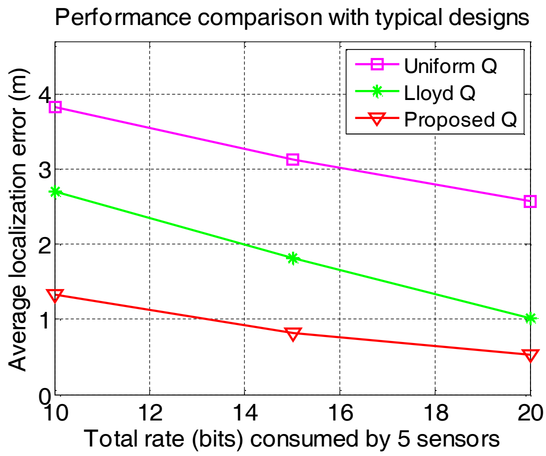

Fig. 1. The proposed quantizers are compared with uniform quantizer and Lloyd Q with $\sigma^2 = 0$.

transmitting the codeword sets at each node to reduce the global metric (i.e., estimation error) can be a useful design strategy for distributed estimation systems. We also run the proposed algorithm with $R_i = 3$, $\sigma^2 = 0$ and compare with previous novel designs such as the localization-specific quantizer (LSQ) [8] and the distributed optimized quantizer (DOQ) [2]. Note that those designs were proposed for design of local quantizers in distributed systems and evaluated for source localization in acoustic sensor networks. The R-D performance curves are illustrated in Figure 2. It should be observed that the proposed quantizer shows a comparable performance gain with a reduced design complexity as compared with the previous novel design techniques.

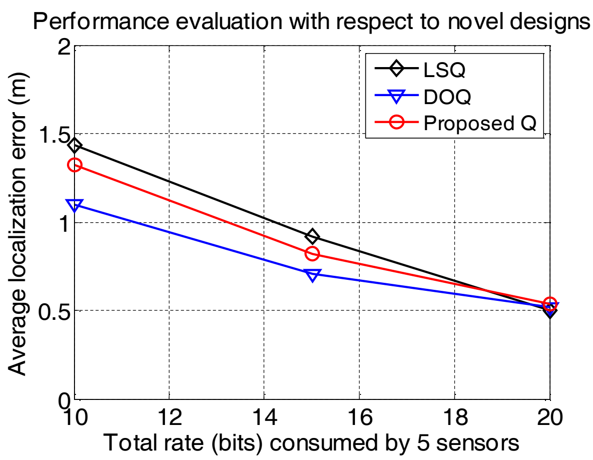

Fig. 2. The proposed quantizers are compared with novel design techniques with $R_i = 3$, $\sigma^2 = 0$.

To investigate how design algorithms can be affected with respect to noise level, we change the measurement noise σ_i in the range from SNR=40 to 100dB for each configuration to generate the test samples. Note that the SNR at 1 meter from the source given by $10\log_{10}a^2/\sigma^2$ is typically much higher than 40dB for noisy engine sound of practical vehicle targets [9], [10]. In Figure 3, it is demonstrated that our quantizers maintain good performance in noisy conditions as compared with the other algorithms.

Ⅷ. Conclusion

In this paper, we proposed a novel design algorithm for local quantizers in distributed systems. Instead of quantizing the measurements, we suggested partitioning the parameter space into sub-regions and sending the quantized version of sets of those regions. To make the algorithm manageable, we presented an iterative design procedure equipped with independent encoding scheme, which achieved a significant performance gain with respect to typical designs and offered a useful design strategy with the adjustable partition level as compared with the previous novel ones. In the future, we will focus on extension of the algorithm to accommodate various useful applications.

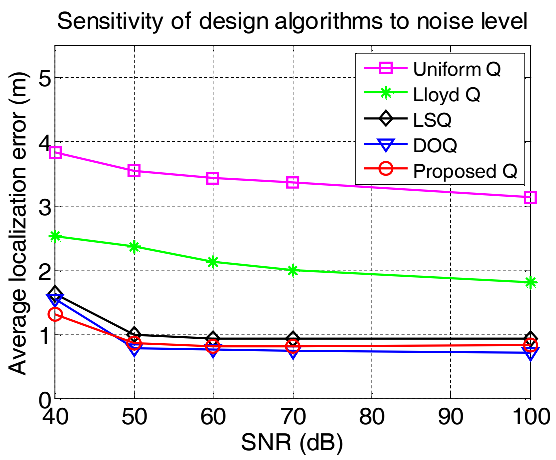

Fig. 3. Sensitivity to noise level with $R_i = 3$, a = 50.

References

[1] 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. DOI:10.1109/ICASSP.2005.1415803

[2] Y. H. Kim. "Quantizer design optimized for distributed estimation," *IEICE Trans. on information and systems*, vol.E97-D, no.6, pp.1639-1643, June 2014. DOI:10.1587/transinf.E97.D.1639

[3] 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, December 2016. DOI:10.6109/jicce.2016.14.4.215

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

[5] Y. H. Kim. and A. Ortega. "Quantizer design for source localization in sensor networks," *IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP)*, March 2005.

[6] 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. DOI:10.1109/ICASSP.2006.1661460

[7] Y. H. Kim and A. Ortega. "Distributed encoding algorithms for source localization in sensor networks," *EURASIP Journal on Advances in Signal Processing*, vol.2010, pp.1-13, Sept 2010. DOI:10.1155/2010/781720

[8] Y. H. Kim and A. Ortega. "Quantizer design for energy-based source localization in sensor networks," *IEEE Trans. on Signal Processing*, vol.59, no.11, pp.5577-5588, November 2011. DOI:10.1109/TSP.2011.2163401

[9] 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. DOI:10.1155/S1110865703212075

[10] 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. DOI:10.1155/S111086570321204X

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

[12] K. Sayood. "*Introduction to data compression-second edition*," Morgan Kaufmann Publishers, 2000.

[13] N. Wernersson, J. Karlsson, and M. Skoglund. "Distributed quantization over noisy channels," *IEEE Trans. on communications*, vol.57, pp.1693-1700, June 2009. DOI:10.1109/TCOMM.2009.06.070482

BIOGRAPHY

Yoon Hak Kim (Member)

1992 : BS degree in Electronic Engineering, Yonsei University.
1994 : MS degree in Electronic Engineering, Yonsei University.
2007 : PhD degree in Electrical Engineering, University of Southern California.

2012~Present : Associate professor, Chosun University.