ISSN : 1226-7244 (Print)
ISSN : 2288-243X (Online)
논문번호 19-02-53
j.inst.Korean.electr.electron.eng.Vol.23,No.2,710～715,June 2019
http://dx.doi.org/10.7471/ikeee.2019.23.2.710
361

Efficient distributed estimation based on non-regular quantized data

Yoon Hak Kim*★

Abstract

We consider parameter estimation in distributed systems in which measurements at local nodes are quantized in a non-regular manner, where multiple codewords are mapped into a single local measurement. For the system with non-regular quantization, to ensure a perfect independent encoding at local nodes, a local measurement can be encoded into a set of a great number of codewords which are transmitted to a fusion node where estimation is conducted with enormous computational cost due to the large cardinality of the sets. In this paper, we propose an efficient estimation technique that can handle the non-regular quantized data by efficiently finding the feasible combination of codewords without searching all of the possible combinations. We conduct experiments to show that the proposed estimation performs well with respect to previous novel techniques with a reasonable complexity.

Key words : Distributed estimation,, non-regular quantizer design, generalized Lloyd algorithm, source localization, sensor networks

Ⅰ. Introduction

Parameter estimation is an inevitable task in distributed systems where sensor nodes randomly deployed in a sensor field collect measurements related with the parameter, quantize and transmit them to a fusion node which then conducts estimation of the parameter based on the received quantized data. In order to improve the estimation performance, much effort has been made in two folds: *first*, various design algorithms for quantizers at local nodes have been proposed. For most of the designs, the generalized Lloyd framework has been adopted to optimize the global metric such as estimation error, leading to efficient design techniques that achieve a significant performance gain over typical designs [1-3]. Furthermore, *non-regular* quantizers which enable encoding of a local measurement into multiple partitions or codewords can be designed to show performance gain over regular ones [4-7]. It should be noted that if quantization partitions are constructed to minimize the estimation error, independent encoding at local nodes without communicating with other nodes would be impossible without computation of estimation error. However, it was shown that independent encoding can be perfectly ensured in a non-regular framework where multiple codewords are allowed for each local measurement while maintaining optimized

* Dept. of Electronic Engineering, Chosun University
★ Corresponding author
E-mail : yhk@chosun.ac.kr, Tel : +82-62-230-7129
※ Acknowledgment
Manuscript received, Jun. 1, 2019; revised Jun. 26, 2019; accepted Jun. 26, 2019.

performance gain [6].

Second, the estimation accuracy can be further improved by developing efficient estimation methods based on quantized data [8-10]. In acoustic sensor networks, source localization algorithm was presented to outperform typical techniques [8, 9]. However, most of estimation techniques operate on regular quantized data and thus direct application to non-regular data would fail to make a full use of information in those non-regular data. To tackle this problem, estimation based on non-regular quantized data was recently proposed in [10] and shown to produce an additional performance improvement at the cost of increased estimation complexity which is proportional to the size of codeword sets, each with a great number of codewords.

In this work, we propose an estimation algorithm that allows us to efficiently find the most feasible combination of codewords by searching the space of combinations of codewords with high probability. This would dramatically decrease the estimation complexity by avoiding an exhaustive search of the spaces of all of combinations. We evaluate the proposed estimation algorithm through experiments and show performance improvement over previous design techniques at the cost of a reasonable complexity.

The main contributions of this work are: (a) we propose a novel estimation technique that operates on non-regular quantized data, even with codeword sets of large cardinality, which ensures a perfect encoding at local nodes, indicating no performance loss in encoding process; (b) we suggest a simple technique to limit the search space to find the combination that is the most likely to occur given received non-regular quantized data and to be able to adjust the complexity level by finding a proper pre-determined parameter that depends upon applications.

This paper is organized as follows. The problem formulation is presented in Section II. Construction of optimized quantization partitions and codeword sets are elaborated in Section III. The estimation algorithm is proposed in Section IV and applied to a source localization system. Experimental results and conclusions are given in Section V and VI, respectively.

II. Problem Formulation

We consider a distributed system where M nodes equipped with sensors modeled by f_i are randomly deployed at known locations, $\mathbf{x}_i \in \mathbf{R}^2$, $i = 1, \cdots, M$ in a sensor field $S \subset \mathbf{R}^N$. In this system, the local measurement at node i denoted by z_i is given by:

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

where θ is the parameter to be estimated and the measurement noise ω_i is assumed to be normal distributed, i.e., $\omega_i \sim N(0, \sigma_i^2)$. In generating quantized measurements, each node uses a R_i bit quantizer with quantization level $L_i = 2^{Ri}$.

For design of non-regular quantizers at nodes, a novel encoding algorithm proposed in [6] is assumed to be employed at each node and thus each node transmits a quantization index interpreted as one of the codeword sets, each with a great number of codewords constructed for a perfect independent encoding. In this work, we present an efficient estimation technique which handles multiple codewords sent from nodes at cost of a reasonable computational complexity. Notice that most of previous estimation algorithms use a single codeword per node for estimation to avoid a huge complexity due to the large cardinality of codeword sets.

III. Construction of optimized quantization partitions and codeword sets

Designing local quantizers that operate independently and minimize the estimation error is explained in what follows. First, quantization partitions are constructed to optimally cluster local measurements into $L_i = 2^{Ri}$ partitions, V_i^j and for each partition, a codeword $\hat{z}_i^{j*}$ is obtained to minimize the estimation error computed over the corresponding partition:

$$\begin{aligned} V_i^j = \{ & z_i : E_{\theta|z_i} \| \theta(z_i) - \hat{\theta}(\hat{z}_i^j) \|^2 \\ & \le E_{\theta|z_i} \| \theta - \hat{\theta}(\hat{z}_i^k) \|^2, \forall k \neq j \} \\ \hat{z}_i^{j*} = & \arg\min_{\hat{z}_i} E[\| \theta(z_i) - \hat{\theta}(\hat{z}_i) \|^2 | z_i \in V_i^j] \end{aligned} \tag{2}$$

where $\hat{\theta}(\hat{z}_i^j)$ is the abbreviated expression for $\hat{\theta}(\hat{z}_{1,\ldots,}\hat{z}_i^j,\ldots,\hat{z}_M)$ which is the estimate of θ computed with $\hat{z}_i = \hat{z}_i^j$. These two steps are iteratively executed until no change in the partitions is attained.

Secondly, the codeword sets C_i^j that enable a perfect independent encoding are shown to be constructed by using the optimized partitions and their corresponding codewords, i.e., V_i^j and $\hat{z}_i^{j*}$, respectively (see [6] for details):

$$C_i^j = \{ \hat{z}_i^j(0), \ldots, \hat{z}_i^j(n_j) \}, \ j = 1, \ldots, L_i$$

where $\hat{z}_i^j(0)$ is the codeword $\hat{z}_i^{j*}$ obtained from the two iterative steps in (2) and n_j is the index for the codewords in the set. In addition, the perfect encoding at local nodes is performed as follows:

$$j^* = \arg\min_{j,n_j} |z_i - \hat{z}_i^j(n_j)|^2, \ \forall j, n_j = 0, \ldots, |C_i^j| \tag{3}$$

For example, if local measurement z_i is closest in one of the codewords in the set C_i^k, then the codeword $\hat{z}_i^k(0)$ is transmitted at node i to a fusion node where the estimation of θ is carried out based on a combination of M codewords, $(\hat{z}_1(0), \ldots, \hat{z}_M(0))$ where $\hat{z}_i(0) = \hat{z}_i^k(0)$.

IV. Proposed estimation algorithm

In this section, we first find the admissible space $S_\theta(\triangle)$ consisting of feasible combinations of codewords given the received quantized data, $\hat{z}_1(0), \ldots, \hat{z}_M(0)$:

$$S_\theta(\triangle) = \{ \theta : \| \theta - \hat{\theta}(\hat{z}_1(0), \ldots, \hat{z}_M(0)) \| \le \triangle \} \tag{4}$$

where $\triangle$ is the pre-set parameter that can be adjusted to achieve a reasonable search complexity and $\hat{\theta}(\hat{z}_1(0), \ldots, \hat{z}_M(0))$ is the estimate of the parameter θ based on the combination of M received codewords.

Second, we generate M measurements $z_i(\theta)$, $i = 1, \ldots, M$ from each of the samples in $S_\theta(\triangle)$ by (1) and encode those measurements to construct the set $\hat{C}$ of the combinations of codewords that are likely to occur given the received quantized data $\hat{z}_1(0), \ldots, \hat{z}_M(0)$:

$$\hat{C}(\triangle) = \{ \hat{z} : \hat{z} = (\hat{z}_1(z_1(\theta)), \ldots, \hat{z}_M(z_M(\theta))), \theta \in S_\theta(\triangle) \} \tag{5}$$

where $\hat{z}_i(\bullet)$ is the codeword encoded from $z_i(\theta)$ at node i by using the encoding rule (3). Clearly, the cardinality of the set $\hat{C}(\triangle)$ is substantially reduced as compared with the whole space of all of the combinations of codewords, each from the cordword set C_i and the size of the set $\hat{C}(\triangle)$ can be simply changed by adjusting the pre-determined parameter $\triangle$, depending upon applications.

The estimation based on the set $\hat{C}(\triangle)$can be performed as follows:

$$\begin{aligned} \hat{z}^* &= \arg\max_{\hat{z}} \Pr[\hat{z} \in \hat{C} \mid \hat{z}_1(0), \ldots, \hat{z}_M(0)] \\ \hat{C}^* &= \hat{C}(\hat{z}^*) \end{aligned} \tag{6}$$

where the estimate $\hat{\theta}^*$ is the maximum likelihood estimate (MLE) obtained from the most feasible combination $\hat{z}^*$.

Note that the construction of the set $\hat{C}(\triangle)$ is conducted at the fusion node for estimation prior to real operation and a perfect encoding at nodes is ensured by using the codeword set $C_i^j, j = 1, \ldots, L_i$ while $\hat{C}(\triangle)$ is generated by using the codeword sets for estimation.

In summary, the proposed algorithm to estimate the parameter θ based on the set $\hat{C}(\triangle)$ consisting of the feasible combinations of codewords is briefly described as follows:

Step 1: Construct the admissible space $S_\theta(\triangle)$ by (4).

Step 2: Generate M measurements from each of samples in $z_i(\theta), i=1,\ldots,M$ by (1).

Step 3: Apply the encoding rule (3) to construct the set $\hat{C}(\triangle)$.

Step 4: Find the most feasible combination $\hat{z}^*$ by searching the reduced set $\hat{C}(\triangle)$ and perform the estimation based on $\hat{z}^*$ by (6).

Ⅴ. Simulation results

We apply our estimation technique to a source localization system with acoustic amplitude sensors where an energy decay model for gathering energy measurements at nodes is employed. Note that the sensor model was widely used for various applications [11, 12]. The measurement at sensor i, denoted by z_i can be expressed as follows:

$$z_i(\theta) = g_i \frac{a}{\|\theta - x_i\|^\alpha} + \omega_i \tag{7}$$

where the sensor model consists of gain factor of the i-th sensor g_i, and energy decay factor α (≈ 2) and source signal energy a which is assumed to be known to the fusion node for estimation.

In the experiments, we randomly scatter M (= 5) nodes in $10 \times 10m^2$ sensor field for 50 different node configurations. We first construct optimal quantization partitions and codewords given in (2) by using training samples generated from the model parameters a = 50, α = 2, g_i = 1 and = 0 in (7) and 1000 uniform distributed source locations for each configuration. From the partitions and codewords, we execute the estimation algorithm in Section IV by varying the pre-set parameter $\triangle$ to create the reduced codeword sets $\hat{C}(\triangle)$ which are employed at a fusion node for estimation. We investigate the performance of our proposed estimation technique (Prop Est) by comparing the previous methods such as the Lloyd quantizer (Lloyd Q) equipped with the maximum likelihood estimator (MLE), the localization specific quantizer (LSQ) in [3] with MLE and the optimized encoding quantizer (OEQ) in [6] with MLE. Notice that Lloyd Q is the standard quantizer designed by simply applying the well known Lloyd algorithm to local training samples and Lloyd Q and LSQ generates regular quantized data. In addition, MLE operates on regular quantized data and the proposed estimation technique runs on non-regular quantized data generated from OEQs at nodes. We generate noise-corrupted measurements from 1000 source locations for each configuration and compute the average localization errors $E\|\theta - \theta\|^2$ by using the four different techniques for comparison.

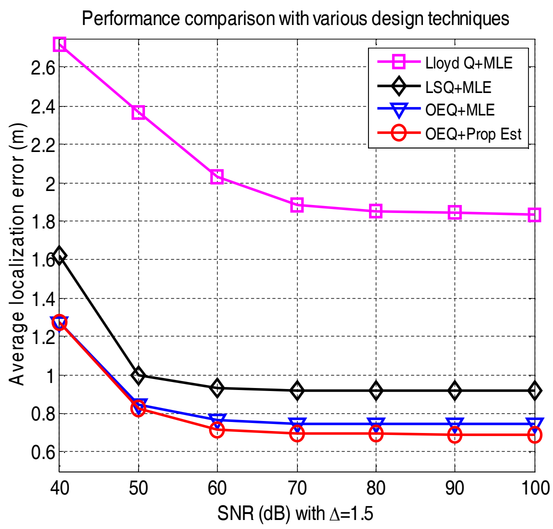

Fig. 1. Performance comparison in the presence of measurement noise: R_i = 3 and, $\triangle = 1.5$.

In Figure 1, the performance curves are plotted for the four cases. In generating noisy test samples, we vary the measurement noise σ_i in the range of SNR=40 to 100dB where SNR is given by $10 \log_{10} a^2/\sigma^2$ (typically, much higher than 40dB for practical vehicle targets [11, 12]). As expected, our estimator (OEQ+Prop Est) outperforms the previous novel designs in noisy conditions since our method makes a good use of non-regular quantized data for estimation with added complexity as compared with OEQ+MLE. Notably, LSQ with MLE shows the worst of all of the optimized quantizers except the typical design, i.e., Lloyd Q, implying an obvious advantage of non-regular quantizers over regular ones.

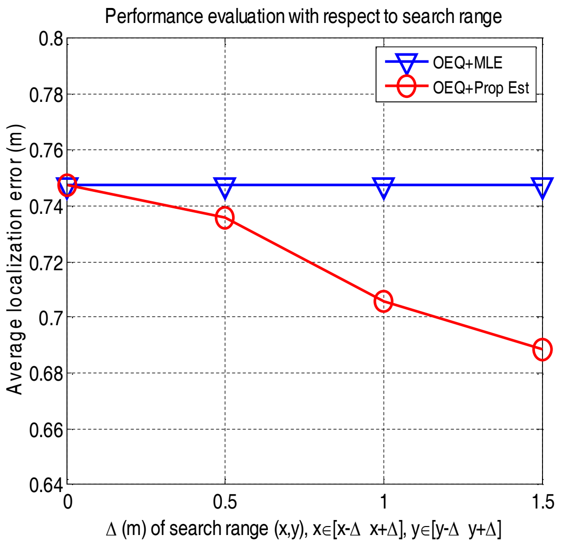

Fig. 2. Performance evaluation with respect to : $R_i = 3$, and $\sigma^2 = 0$.

We also examine how our estimation technique works as the parameter △ is changed. In Figure 2, the performance of our algorithm is compared with OEQ+MLE to emphasize effectiveness that our estimator can offer by using non-regular data for estimation with respect to estimation complexity level. As seen, the proposed estimator achieves a significant localization performance gain as compared with the novel design (non-regular quantizer equipped with a typical estimator).

Ⅵ. Conclusion

In this paper, we proposed an efficient estimation algorithm in distributed systems that operates on non-regular quantized data. In order to utilize a great number of codewords transmitted by each node for estimation, we presented a simple technique that reduces the size of the space of the combinations of codewords to be searched and showed that our estimator performs better than previous novel estimators with a reasonable complexity. We also demonstrated that the performance of our technique can be greatly improved by increasing the pre-determined parameter. In the future, we will work on compression of the codeword sets while maintaining a perfect independent encoding at nodes.

References

[1] 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, 1990. DOI: 10.1109/18.52470

[2] Y. H. Kim. and A. Ortega. “Quantizer design for source localization in sensor networks,” *IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP)*, 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 Trans. on Signal Processing*, vol.59, no.11, pp. 577-5588, 2011. DOI: 10.1109/TSP.2011.2163401

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

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

[6] 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

[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, 2010. DOI: 10.1155/2010/781720

[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] 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)*, 2006. DOI: 10.1109/ICASSP.2006.1661460

[10] Y. H. Kim. "Distributed estimation using non-regular quantized data," *Journal of information and communication convergence engineering*, vol.15, no.1, pp.7-13, 2017. DOI: 10.6109/jicce.2017.15.1.7

[11] 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

[12] 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

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