

JOINT OPTIMIZATION OF SOURCE CODING AND POWER ALLOCATION IN SENSOR NETWORKS

Jun Yuan and Wei Yu

Electrical and Computer Engineering Department, University of Toronto

ABSTRACT

In a sensor network, each sensor makes a local observation of some underlying physical phenomenon, and sends a quantized version of the observation to a central office via communication links. Since the sensors' observations are often partial and correlated, the network performance becomes a complicated and non-separable function of all individual data rates at each sensor. In this paper, we consider a joint optimization of source coding and power allocation in a sensor network. We model the sensor network from an information theoretical perspective, and propose a novel formulation for distributive source coding to characterize the tradeoff among source coding rates. The new formulation is capable of dealing with the case where the physical source is described by a vector of random variables. We further optimize the power allocation strategy among sensors. We show that the joint source coding and sensor power allocation problem can be solved optimally and efficiently via convex programming.

1. INTRODUCTION

Sensor networks have emerged as a promising application in military sensing, physical security, traffic control and environment monitoring. In a sensor network, a large number of sensors are deployed over a field. Each sensor makes a local observation of some underlying physical source, quantizes its observation, and transfers the quantized data back to a "central estimation office" (e.g. CEO) through communication links.

The design objective of the sensor network is to reconstruct the underlying physical source under resource limitation i.e., power. Thus, it is natural to define the overall network utility optimization problem as that of minimizing the overall distortion and power consumption. Here, the distortion is defined as the difference between the true underlying source and its estimation at CEO. Due to the partial observation at each sensor, overall estimation at CEO depends on all sensors' data rates in a complicated fashion. In particular, using distributed source coding techniques, it is possible to tradeoff transmission rate in one sensor with transmission rate in another sensor. Furthermore, since power is a scarce resource in a sensor network, it is important to allocate power

optimally among sensors when transmitting observation back to the CEO.

The joint quantizer design and power allocation problem has been considered before in the literature. In [1], Xiao, Cui, Luo and Goldsmith proposed a scheme of optimal power scheduling for decentralized estimation in sensor networks. The authors derived a distortion bound based on best linear unbiased estimation (BLUE). They analytically solved an optimization problem of minimizing total power consumption under distortion requirement.

In this paper, we address the joint optimization of source coding and power allocation from an information theory point of view. We focus on a complete *digital* approach, where each sensor compresses its local observation into a digital form and transmits the data digitally. Our main contribution is a source coding formulation that characterizes the coupled relation of source rates among nodes. We further show that the overall minimization of distortion and power can be solved jointly via convex programming. This paper differs from [1] in that (i) our source coding formulation is capable of dealing with physical source that is modelled as a vector of random variables, while the work of Xiao et al considered the scalar case only; (ii) we model the source coding problem using rate-distortion theory and model the power allocation problem using Shannon capacity formula from an information theoretical viewpoint, while [1] adopted a signal processing perspective. This paper also differs from our previous work [2]. Here, we explicitly characterize the rate-distortion region and the capacity region, while [2] assumes that both regions are given and convex.

2. SYSTEM MODEL

Consider an environment sensing application depicted in Fig. 1. The underlying physical source θ is a vector of random variables. N sensors are deployed in the field, each making a local (and possibly partial) observation of θ , while being corrupted by observation noise n_i . The observation channel is characterized by a matrix H . At each sensor i , the noisy observation y_i is quantized into a codeword u_i . The quantized information from all sensors is transmitted back directly to a remote CEO with source rates (s_1, \dots, s_N) . The source rates are bounded by the link capacities (c_1, \dots, c_N) , which de-

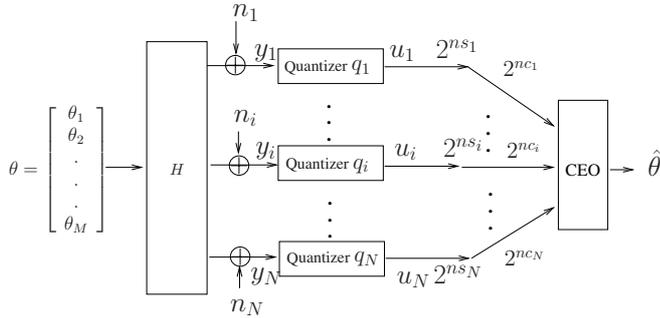


Fig. 1. Sensor Network

pend on the communication link gain, noise, and transmission powers. At the remote CEO, the decoder first jointly decodes the codewords \mathbf{u} , then estimates the physical source. The estimation is denoted as $\hat{\theta}$. The performance criterion is the mean-square error distortion, i.e., $D(\hat{\theta}, \theta) = \|\hat{\theta} - \theta\|^2$. In this paper, we assume the physical source is modelled as an independent Gaussian vector θ , and the observation noise is modelled as an independent Gaussian vector \mathbf{n} .

2.1. Source Coding

The task of source coding at each sensor is to design a good quantization scheme such that the distortion is minimized. The quantization process can be modelled as an independent quantization noise variable \mathbf{q} according to [3, 4]. In this case, the sensor quantization output can be expressed as

$$\mathbf{u} = \mathbf{y} + \mathbf{q} = H\theta + \mathbf{n} + \mathbf{q}$$

Now, suppose that the remote CEO can receive \mathbf{u} correctly, (this is justified in the next paragraph,) the minimum mean square error (MMSE) estimation, which provides an optimal estimation $\hat{\theta}$ with a corresponding error covariance matrix K_0 , can be expressed as follows:

$$\begin{aligned} \hat{\theta} &= \mathbb{E}(\theta|\mathbf{u}) = R_{\theta u} R_{uu}^{-1} \mathbf{u} \\ K_0 &= \mathbb{E} \left[\begin{bmatrix} \theta - \hat{\theta} \\ \theta - \hat{\theta} \end{bmatrix} \begin{bmatrix} \theta - \hat{\theta} \\ \theta - \hat{\theta} \end{bmatrix}^T \right] = R_\theta - R_{\theta u} R_{uu}^{-1} R_{u\theta} \\ &= R_\theta - R_\theta H^T (H R_\theta H^T + R_n + R_q)^{-1} H R_\theta^T \end{aligned}$$

where $[\cdot]^T$ denotes matrix transpose, $R_{\theta u} = \mathbb{E}[\theta \mathbf{u}^T] = R_\theta H^T$, R_θ is the covariance matrix of the underlying source, $R_{uu} = \mathbb{E}[\mathbf{u} \mathbf{u}^T] = H R_\theta H^T + R_n + R_q$, and R_n R_q are diagonal covariance matrices of observation noise and quantization variance respectively. The distortion is the trace of K_0

$$D(\hat{\theta}, \theta) = \text{tr}(K_0) \quad (1)$$

Note that the distortion actually depends on the quantization schemes at all sensors. This distortion expression naturally characterizes the coupled relation among sensor rates.

In order to correctly decode \mathbf{u} at the remote CEO, each sensor must communicate to the CEO at a sufficiently large rate. Since the sensors' observations are correlated, then quantized version \mathbf{u} are correlated as well. From an information theory point of view, it is possible to encode \mathbf{u} using distributive source coding methods (e.g., using the techniques of Slepian-Wolf [5] and Berger-Tung [6]).

However, Slepian-Wolf coding is not easy to implement in practice. In this paper, we adopt a suboptimal but computationally simple source coding expression, which preserves the coupled relation among sensor's quantization levels but does not take into account Slepian-Wolf coding:

$$s_i = I(Y_i; U_i) = \frac{1}{2} \log_2 \left(\frac{h_i^T R_\theta h_i + \sigma_{n_i}^2 + \sigma_{q_i}^2}{\sigma_{q_i}^2} \right) \quad (2)$$

where h_i^T is the i th row of channel matrix H , $\sigma_{n_i}^2$ and $\sigma_{q_i}^2$ are the noise variance and quantization variance respectively, which are the i th diagonal element of R_n and R_q . Note the source rate s_i in (2) is an upper bound to the true rate-distortion region.

For reasons that will be clear later, it is convenient to define a *quantization effort* variable $w_i = 1/(\sigma_{n_i}^2 + \sigma_{q_i}^2)$, which has a one-to-one map with the quantization variance $\sigma_{q_i}^2$. In this case, $\sigma_{q_i}^2 = 1/w_i - \sigma_{n_i}^2$, $R_w^{-1} = R_n + R_q$. Let $\sigma_{s_i}^2 = h_i^T R_\theta h_i$. The distortion and source rate can now be expressed as:

$$\begin{aligned} D &= \text{tr}(R_\theta) - \text{tr}(R_\theta H^T (H R_\theta H^T + R_w^{-1})^{-1} H R_\theta^T) \\ s_i &= \frac{1}{2} \log_2 \left(\frac{1 + \sigma_{s_i}^2 w_i}{1 - \sigma_{n_i}^2 w_i} \right) \\ \sigma_{s_i}^2 &= h_i^T R_\theta h_i, \quad 0 \leq w_i \leq \frac{1}{\sigma_{n_i}^2} \end{aligned}$$

where R_w is a diagonal matrix with w_i as the i th diagonal element. Note that larger the quantization effort, i.e., $\mathbf{w} = [w_1, \dots, w_N]^T$, smaller the distortion is, and higher the source rates are.

2.2. Power Allocation

In this paper, for the sake of simplicity, we assume that the communication links are independent, e.g., a time-division multiple access (TDMA) scheme or a frequency-division multiple access (FDMA) scheme is used. Therefore, the link capacity is determined by the power allocation according to Shannon formula:

$$c_i = \frac{1}{2} \log_2 \left(1 + \frac{G_i p_i}{\Gamma_i} \right), \quad p_i \in [0, p_{i,\max}], \quad \forall i \quad (3)$$

where G_i and Γ_i are the channel gain and noise respectively for the i th link. The transmission power p_i is constrained by its maximum $p_{i,\max}$.

3. JOINT OPTIMIZATION

Given the characterization of source coding and channel coding in the previous section, we are now ready to formulate the joint source coding and power allocation optimization problem in a sensor network.

$$\begin{aligned} & \text{minimize} && D + \beta^T \mathbf{p} && (4) \\ & \text{subject to} && s_i \leq c_i, \quad \forall i \end{aligned}$$

where the objective function illustrates a tradeoff between distortion minimization and power consumption, with β as the relative weighting vector, i.e., $\beta = [\beta_1, \dots, \beta_N]^T$. Here, D is the distortion and $\mathbf{p} = [p_1, \dots, p_N]^T$ is the vector of transmission power allocation at each node. The inequality constraint states that the source rate s_i at each node must be smaller than its link capacity c_i . Furthermore, (4) can be expressed in terms of quantization effort and power consumption as follows:

$$\begin{aligned} \min & \quad \text{tr}(R_\theta) - \text{tr}(R_\theta H^T (H R_\theta H^T + R_w^{-1})^{-1} H R_\theta^T) \\ & \quad + \beta^T \mathbf{p} && (5) \\ \text{s.t.} & \quad \frac{1}{2} \log_2 \left(\frac{1 + \sigma_{s_i}^2 w_i}{1 - \sigma_{n_i}^2 w_i} \right) \leq \frac{1}{2} \log_2 \left(1 + \frac{G_i p_i}{\Gamma_i} \right) \\ & \quad \sigma_{s_i}^2 = h_i^T R_\theta h_i, \quad w_i \in [0, 1/\sigma_{n_i}^2], \quad p_i \in [0, p_{i,\max}] \end{aligned}$$

Theorem 1 *The joint optimization problem of source coding and power allocation (5) is a convex optimization problem.*

Proof: We first prove that the objective function is a convex function over variables (\mathbf{w}, \mathbf{p}) by checking that its Hessian is positive semidefinite. Denote $Y = (H R_\theta H^T + R_w^{-1})$, $Z = H R_\theta^T$, and $f(\mathbf{w}, \mathbf{p}) = \beta^T \mathbf{p} - \text{tr}(Z^T Y^{-1} Z)$. Here, R_θ is a positive definite diagonal matrix; R_w is a diagonal matrix with the i th element w_i ; and $R_w^{-1} = R_n + R_q$, where R_n, R_q are positive definite diagonal matrices. Hence, Y is a positive definite, symmetric matrix. We use the following definition.

Definition 1 *If $A = [a_{ij}] \in M_{m,n}$ and $B = [b_{ij}] \in M_{m,n}$ are given, then the Hadamard (or Schur) product of A and B is the matrix $A \circ B \equiv [a_{ij} b_{ij}] \in M_{m,n}$.*

Let $\text{diag}(A)$ be the vector consisting of the diagonal elements of A . If a matrix A is positive semidefinite, we denote as $A \succeq 0$. By matrix chain rule, we have

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{w}} &= -\text{diag}(R_w^{-2} \circ [Y^{-1} Z Z^T (Y^{-1})^T]) \\ \frac{\partial^2 f}{\partial \mathbf{w}^2} &= 2 [R_w^{-3} - R_w^{-2} Y^{-1} R_w^{-2}] \circ [Y^{-1} Z Z^T Y^{-1}] \\ \frac{\partial f}{\partial \mathbf{p}} &= \beta, \quad \frac{\partial^2 f}{\partial \mathbf{p}^2} = 0, \quad \frac{\partial^2 f}{\partial \mathbf{w} \partial \mathbf{p}} = 0 \end{aligned}$$

Now, we prove that both matrices $(R_w^{-3} - R_w^{-2} Y^{-1} R_w^{-2})$ and $(Y^{-1} Z Z^T Y^{-1})$ are positive semidefinite. This is true, be-

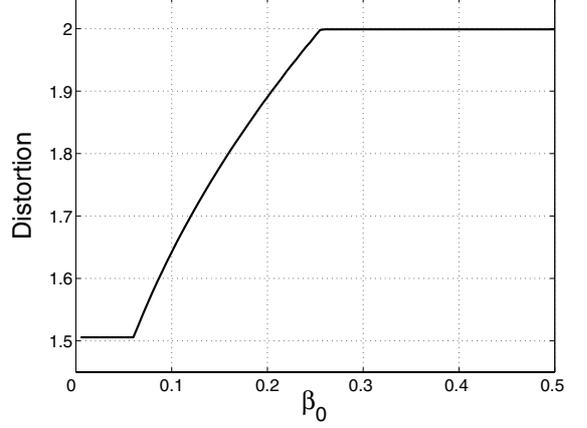


Fig. 2. Distortion vs. Beta

cause R_θ, R_w, R_w^{-1} are all positive definite diagonal matrices.

$$\begin{aligned} Y = H R_\theta H^T + R_w^{-1} \succeq R_w^{-1} &\Rightarrow R_w^{-3} - R_w^{-2} Y^{-1} R_w^{-2} \succeq 0 \\ Z Z^T \succeq 0 &\Rightarrow Y^{-1} Z Z^T Y^{-1} \succeq 0 \end{aligned}$$

According to Schur product theorem [7], the schur product of two positive semidefinite matrices is positive semidefinite. Hence, the Hessian is positive semidefinite, and $f(\mathbf{w}, \mathbf{p})$ is convex over (\mathbf{w}, \mathbf{p}) .

Furthermore, the first inequality in (5) is equivalent to

$$\left(\frac{1 + \sigma_{s_i}^2 w_i}{1 - \sigma_{n_i}^2 w_i} \right) - \left(1 + \frac{G_i p_i}{\Gamma_i} \right) \leq 0$$

It is not difficult to see that the first term above is convex over w_i under the domain of $w_i \in [0, 1/\sigma_{n_i}^2]$. And the second term is linear in p_i under the domain of $p_i \in [0, p_{i,\max}]$. Therefore, the constraint is convex.

In conclusion, both the objective function and constraint sets are convex, hence the joint optimization problem (5) is a convex optimization problem. Q.E.D.

Given the result of Theorem 1, the joint optimization problem of source coding and power allocation can be solved optimally and efficiently via convex programming [8].

4. SIMULATION EXAMPLE

We simulate an example of a sensor network in Fig. 1. The underlying physical source to be observed is modelled as a two dimension Gaussian vector with an identity covariance matrix. For the sake of simplicity, we consider the case of two sensors with the corresponding observation matrix $H = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix}$ and equal weighting elements, i.e., $\beta_1 = \beta_2 = \beta_0$. The communication links are independent. The optimization problem is solved using the interior-point method.

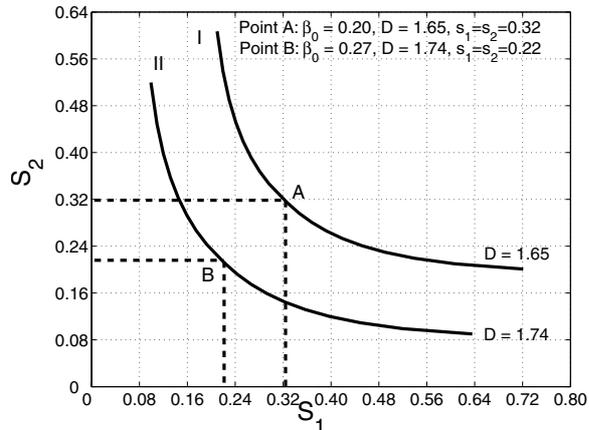


Fig. 3. Distortion-Power Tradeoff

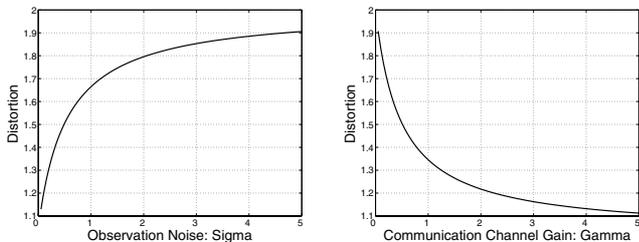


Fig. 4. (a) Distortion vs. Observation Noise Variance, (b) Distortion vs. Communication Channel Gain

Fig. 2 illustrates the tradeoff between distortion minimization and power consumption. The larger the cost weight β_0 is, the larger the optimal distortion becomes. As we can see when the cost is small i.e., $\beta_0 < 0.12$, the system operates at maximum quantization rate with maximum transmission power. At the other extreme i.e., $\beta_0 > 0.5$, it is too expensive to transmit, therefore the CEO just ‘guesses’, i.e., $D = \text{tr}(R_\theta) = 2$.

Fig. 3 plots different rate-distortion curves (solid line) and capacity regions (dash line). It is clear that multiple pairs of source coding rates (s_1, s_2) along the curve can achieve the same distortion. The question of which pair to operate on depends on how much link capacity is available. For example, given cost weight $\beta_0 = 0.20$, the optimal solution of (5) turns out to be point A, where the rate-distortion curve I intersects with the capacity region for a distortion level $D = 1.65$. As the power becomes more expensive $\beta_0 = 0.27$, the optimal solution turns out to be point B, where the distortion is larger $D = 1.74$ but the power consumption is also lower.

We further explore how the optimal distortion depends on other factors, such as the observation noise variance or the communication channel gain. Consider the configuration of zero cost i.e., $\beta_0 = 0$. As illustrated in Fig. 4(a), the optimal distortion monotonically increases as the variance of obser-

vation noise increases. Similarly when the communication channel gain becomes larger and larger, the distortion continuously decreases as shown in Fig. 4(b).

5. CONCLUSIONS

In this paper, we address the joint optimization of source coding and power allocation in a sensor network. An information theoretical approach is taken. Our main contribution is a novel source coding formulation for vector source. Under the assumption of independent communication links, we prove that the joint optimization problem is convex, therefore optimal solution can be obtained efficiently via convex programming.

Future extensions of this work will be conducted along two directions. First, we may consider a more complicated channel coding model such as the interference links. Second, the communication from each sensor to the CEO can be done through a multihop network. Therefore, routing becomes another factor to be optimized along with source and channel coding.

6. REFERENCES

- [1] J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, “Joint Estimation in Sensor Networks under Energy Constraint,” *IEEE Trans. on Signal Processing*, vol. 54, no. 2, pp. 413–422, February 2005.
- [2] W. Yu and J. Yuan, “Joint Source Coding, Routing and Resource Allocation for Wireless Sensor Networks,” *IEEE International Conference on Communications (ICC)*, May 2005.
- [3] W. R. Bennett, “Spectra of Quantized Signals,” *Bell Syst. Tech. J.*, vol. 27, pp. 446–472, 1948.
- [4] B. Widrow, “Statistical Analysis of Amplitude Quantized Sampled Data Systems,” *Trans. Amer. Inst. Elec. Eng. Pt. II: Applications and Industry*, vol. 79, no. 555-568, 1960.
- [5] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471–480, 1973.
- [6] T. Berger, “Multiterminal source coding,” *The Information Theory Approach to Communications (G. Longo ed.)*, vol. 229 of CISM courses and lectures, pp. 171–231, 1978.
- [7] R. Horn and C. Johnson, *Matrix Analysis.*, Cambridge Univ. Press, 1991.
- [8] S. Boyd and L. Vendenbergh, *Convex Optimization*, Cambridge University Press, 2003.