

センサネットワークと多端子情報理論

大濱 靖匡[†]

[†]九州大学システム情報科学研究所 〒812-8581 福岡市東区箱崎 6-10-1

E-mail: †oohama@csce.kyushu-u.ac.jp

あらまし センサネットワークの理論構築に多端子情報理論が深く関与することが指摘されている。本論文は、センサネットワークを利用した情報センシングのプロセスが多端子情報理論の問題として、どのように定式化されるのかを述べる。また、定式化により得られた符号化問題について、これまでどのような結果が得られているのかについて述べる。

キーワード センサネットワーク、多端子情報理論、分散符号化、多補助者問題、CEO 問題、

Sensor Network and Multiterminal Information Theory

Yasutada OOHAMA[†]

[†] Faculty of Information Science and Electrical Engineering,
Graduate Schools, Kyushu University

6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

E-mail: †oohama@csce.kyushu-u.ac.jp

Abstract The construction of theoretical framework to treat sensor networks has a close connection with multiterminal information theory. In this paper we explain how a process of signal sensing in sensor network to be formulated as a coding problem of multiterminal information theory. We further explain previous results on the formulated coding problem.

Key words Sensor Network, multiterminal information theory, distributed source coding, many-help-one problem, CEO problem

1. はじめに

多数のセンサデバイスをネットワークで結んだ通信システムを用いて情報センシングや遠隔制御を行うセンサネットワークとよばれる技術が注目を集めている。現在、環境測定、セキュリティ、知的空間の構築、大災害時の救助活動など多方面へのセンサネットワークの応用が検討されている。

従来の通信ネットワークと著しく異なり、センサネットワークの構成要素である信号源、センサ出力、ネットワークには、不安定性、不確実性などの種々の拘束条件が課せられている。この制約下で、センシングの高度化や情報通信、制御の最適化を図る必要がある。これは、センサネットワークの基本課題の1つである。

センサネットワークの要素技術の関与する学術研究分野は、センシング、信号処理、統計学、情報理論、通信理論、情報ネットワーク、人工知能、制御理論、システム理論など、極めて多方面にわたる。各分野の既存成果の組み合わせにより、センサネットワークを設計、構築して提供することが可能であり、このような研究開発が精力的に進められている。しかし、これだけでは、提供されるシステムの性能を客観的に評価できない。センサネットワークの基本

課題を解決し、また、その設計、構築の客観的指針を得るためには、分野横断的視点から、センサネットワークを取り扱う理論的枠組みを確立し、その枠組みにおいてシステムの限界と可能性を数理的に明らかにする必要がある。

情報通信の数理的基礎といえば、1948年にシャノンに築いた情報理論がある。これは、1970年代に入り、多端子情報理論とよばれる多数の送受信者による情報通信を扱う理論へ発展した。

ある情報信号が複数の地点に分散配置されたセンサによりセンシングされる場合を考える。分散センサの得た情報信号に関するデータは、ネットワークを介して一箇所に集められ、そこで情報信号に関する推論が行われる。このような分散センサによる信号推定システムは、センサネットワークにおける情報センシングのプロセスと捉えることが出来る。一方、このような通信システムの理論的モデルは、多端子情報理論における主要研究対象として、これまで研究されてきたものの一つである。

本稿では、センサネットワークによる情報センシングのプロセスが多端子情報源の分散符号化の問題として、自然な形で定式化されることを明らかにし、さらに定式化された問題に関するこれまでの研究結果について述べる。

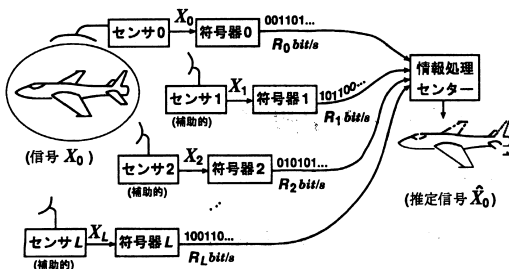


図1 センサネットワークによる情報センシング

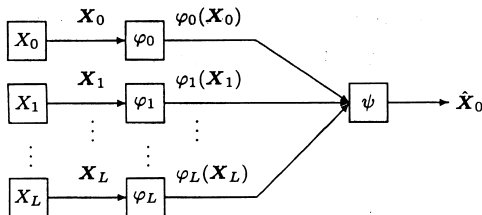


図2 センサネットワークの理論的モデル

2. センサネットワークの理論的モデル

集合 $\mathcal{X}_i$ に値をとる確率変数 $X_i, i = 0, 1, 2, \dots, L$ を考える。これら $(L+1)$ 個の確率変数は相関を持つものとする。本稿では、 X_i のとる値の集合 $\mathcal{X}_i, i = 0, 1, 2, \dots, L$ が有限集合、または実数の場合を考える。前者の場合、 $(X_0, X_1, \dots, X_L)$ の従同時確率分布 $p_{X_0 X_1 \dots X_L}$ を

$$p_{X_0 X_1 \dots X_L} \triangleq \{p_{X_0 X_1 \dots X_L}(x_0, x_1, \dots, x_L) \mid (x_0, x_1, \dots, x_L) \in \mathcal{X}_0 \times \mathcal{X}_1 \times \dots \times \mathcal{X}_L\}$$

と記す。後者の場合は確率分布を確率密度関数に置き換えた定義になる。確率変数が実数に値をとる場合に関しては、本稿では、 $X_0, X_1, \dots, X_L$ が $(L+1)$ 次元ガウス分布に従う場合を主に扱う。 $\{(X_{0,t}, X_{1,t}, \dots, X_{L,t})\}_{t=1}^{\infty}$ を定常無記憶情報源とする。各 $t = 1, 2, \dots$ に対し、 $(X_{0,t}, X_{1,t}, \dots, X_{L,t})$ は、 $(X_0, X_1, \dots, X_L)$ と同一の分布に従うものとする。情報源から発生する長さ n の確率変数の列を $\mathbf{X}_i = X_{i,1} X_{i,2} \dots X_{i,n}$ と記す。

さて、離れた $(L+1)$ 個の地点において、 $X_0, X_1, \dots, X_L$ が個別に観測される場合を考える。これらの $(L+1)$ 個のデータ系列は、それぞれの地点で独立に符号化され、復号器へ送られる。復号器は、送られてくる符号化データから X_0 の推定値 $\hat{X}_0$ を出力する。このようなセンサネットワークによる情報センシングの様子を図1に示す。 X_0 が情報センシングにより推定したい信号である。残りの L 個のデータ系列は復号器において符号化された形で X_0 の推定のための補助情報として働く。センサネットワークによる情報センシングのプロセスの理論的モデルを図2に示す。この図に示された通信システムは、形式的には次のように述べられる。各 $i = 0, 1, \dots, L$ に対し、情報源から発生する長さ n のデータ系列 $\mathbf{X}_i$ は、符号器 φ_i により各々独立に $\varphi_i(\mathbf{X}_i)$ と符号化され、これらは、復号器 ψ に送られる。復号器 ψ は、 $\varphi_i(\mathbf{X}_i), i = 0, 1, \dots, L$ を入力として X_0 の推定値 $\hat{X}_0$ を出力する。ここで、符号器 $\varphi_i, i = 0, 1, \dots, L$ は、 $\varphi_i: \mathcal{X}_i^n \rightarrow \mathcal{M}_i = \{1, 2, \dots, M_i\}$

なる写像で定義され、以下の伝送率制約を満たす。

$$\frac{1}{n} \log M_i \leq R_i + \delta, \quad i = 0, 1, \dots, L \quad (1)$$

ここで δ は予め任意に決められた正数である。なお、本稿を通じて対数の底は2であるとする。復号器 ψ は $\psi: \mathcal{M}_0 \times \mathcal{M}_1 \times \dots \times \mathcal{M}_L \rightarrow \mathcal{X}_0^n$ で定義される。伝送率制約(1)を満たす符号器と復号器の組 $(\varphi_0, \varphi_1, \dots, \varphi_L, \psi)$ 全体からなる集合を $\mathcal{F}_\delta^{(n)}(R_0, R_1, \dots, R_L)$ と書く。歪測度 $d(x, y), (x, y) \in \mathcal{X}_0^2$ を $x = y$ のとき0、それ以外のとき正の有限値をとるものと定義する。データ系列 $\mathbf{X}_0$ とその推定値

$$\hat{\mathbf{X}}_0 = \psi(\varphi_0(\mathbf{X}_0), \varphi_1(\mathbf{X}_1), \dots, \varphi_L(\mathbf{X}_L))$$

との一致の尺度として、以下の平均歪を考慮する。

$$\Delta(\mathbf{X}_0, \hat{\mathbf{X}}_0) = \frac{1}{n} \sum_{t=1}^n \text{Ed}(X_{0,t}, \hat{X}_{0,t})$$

与えられた $D > 0$ と任意の $\delta > 0$ に対し、整数 $n_0(\delta)$ が存在して、任意の $n \geq n_0(\delta)$ に対し、 $(\varphi_0, \varphi_1, \dots, \varphi_L, \psi) \in \mathcal{F}_\delta^{(n)}(R_0, R_1, \dots, R_L)$ なる符号器、復号器の組が存在して、 $\Delta(\mathbf{X}_0, \hat{\mathbf{X}}_0) \leq D + \delta$ となるとき、伝送率ベクトル $(R_0, R_1, \dots, R_L)$ は、許容であると定義する。許容であるような伝送率ベクトル全体からなる集合 $\mathcal{R}_L(D)$ を伝送率・歪領域とよぶ。 $\mathcal{R}_L(D)$ の具体形を明らかにすることが基本問題である。

条件(1)は通信回線に対する容量制約を表す。この制約は上記の問題設定において本質的である。実際、 $\mathcal{X}_0$ が有限集合の場合、その要素の個数を $|\mathcal{X}_0|$ として、 $R_0 \geq \log |\mathcal{X}_0|$ ならば、符号器 φ_0 は X_0 を誤りなしで送れる。したがって、容量制約がなければ、他の L 本の通信回線が不要となり、問題自体が意味をなさない。

次に、 L 個の符号器が、いずれも X_0 を観測できる場合を考える。この場合は、単一情報源に対する符号化の基本定理より $\mathcal{R}_L(D)$ は

$$\sum_{i=0}^L R_i \geq R_{X_0}(D) \quad (2)$$

を満たす伝送率ベクトルからなる領域となる。ここで、 $R_{X_0}(D)$ は伝送率・歪関数 (rate-distortion function) とよばれる量である。この関数の $\mathcal{X}_0$ が有限集合の場合の定義は次のようになる。 $V = \{V(\hat{x}_0|x_0)\}_{(x_0, \hat{x}_0) \in \mathcal{X}_0^2}$ を $\mathcal{X}_0$ の元が与えられた下での $\mathcal{X}_0$ 上の条件付確率分布とし、 $\mathcal{X}_0^2$ に値をとる確率変数の組 $(X_0, \hat{X}_0)$ の同時分布を $p_{X_0}(x_0)V(\hat{x}_0|x_0), (x_0, \hat{x}_0) \in \mathcal{X}_0^2$ と定めたとき

$$R_{X_0}(D) \triangleq \min_{V: \text{Ed}(X_0, \hat{X}_0) \leq D} I(X_0; \hat{X}_0)$$

で定義される。ここで $I(X_0; \hat{X}_0)$ は X_0 と $\hat{X}_0$ との間の相互情報量を表す。これは $\hat{X}_0$ のエントロピー

$$H(\hat{X}_0) \triangleq - \sum_{\hat{x}_0 \in \mathcal{X}_0} p_{\hat{X}_0}(\hat{x}_0) \log p_{\hat{X}_0}(\hat{x}_0)$$

$$p_{\hat{X}_0}(\hat{x}_0) = \sum_{x_0 \in \mathcal{X}_0} p_{X_0}(x_0)V(\hat{x}_0|x_0), \hat{x}_0 \in \mathcal{X}_0$$

と X_0 が与えられた下での $\hat{X}_0$ の条件付エントロピー

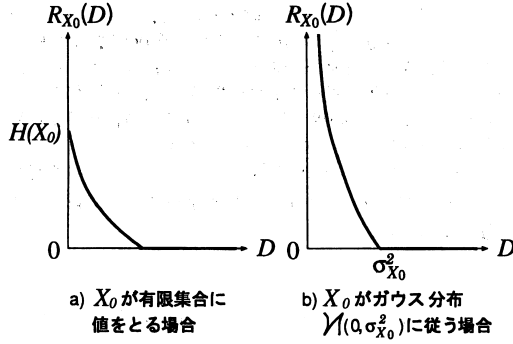


図3 関数 $R_{X_0}(D)$ の概形

$$H(\hat{X}_0|X_0) = - \sum_{(x_0, \hat{x}_0) \in \mathcal{X}_0^2} p_{X_0}(x_0) V(\hat{x}_0|x_0) \log V(\hat{x}_0|x_0)$$

との差 $H(\hat{X}_0) - H(\hat{X}_0|X_0)$ で表現できる。 X_0 が実数の場合の $R_{X_0}(D)$ の定義は、確率分布を確率密度関数に置き換え、和を積分に換えればよい。特に X_0 の従う分布が平均 0、分散 $\sigma_{X_0}^2$ のガウス分布で、歪み測度 d が 2 乗誤差 $d(x, y) = (x - y)^2$ の場合、 $R_{X_0}(D)$ は陽に求められ

$$R_{X_0}(D) = \frac{1}{2} \log^+ \left[\frac{\sigma_{X_0}^2}{D} \right] \quad (3)$$

となる。ここで、 $\log^+[a] = \max\{\log a, 0\}$ 。 X_0 が有限集合に値をとる場合およびガウス分布に従う場合の各場合に対する $R_{X_0}(D)$ の概形を図 3 に示す。以上のことから各符号器が X_0 を観測できる場合も、新しい問題は生じない。

これまでの考察からわかるように、定式化された問題において、観測データの不確実性、通信回線に課せられる厳しい容量制約といったセンサネットワークの有する拘束条件は、極めて本質的な意味を持つ。

3. 情報源出力が有限集合に値をとる場合

伝送率・歪領域 $\mathcal{R}_L(D)$ を求める問題は、 L 補助者問題とよばれる。この問題は現在も一般には未解決である。情報源からの出力系列が有限集合に値をとる場合を考える。この場合の L 補助者問題に対する主要結果が [1]-[8] によって得られている。これらの研究結果は、多端子情報源符号化に関する研究のこれまでの発展の歴史においても重要な位置づけにある。その結果の詳細は、本稿では省略する。詳細は [9], [10] を参照せよ。

これらの結果の中でも特にセンサネットワークとのつながりが非常に深いものとして、Gelfand and Pinsker [7] の結果、および CEO 問題 [8] がある。ここでは、それらについて述べる。

3.1 Gelfand and Pinsker [7] の結果

Gelfand and Pinsker [7] は情報源が次の相関条件を満たす場合に注目した。

条件: X_0 が与えられた条件の下で $X_1, X_2, \dots, X_L$ が条件付独立 (Conditionally Independent) である。この相関条件を CI 条件とよぶことにする。これは、同時

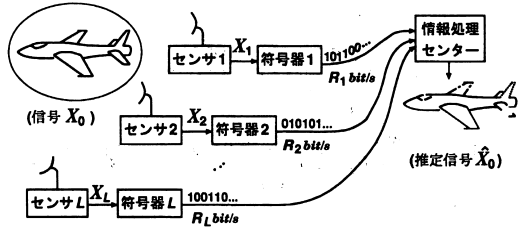


図4 補助情報のみからなる情報センシング

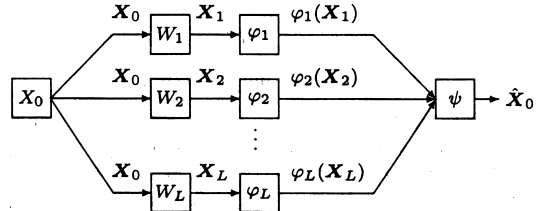


図5 CEO 問題の枠組み

確率分布 $p_{X_0 X_1 \dots X_L}$ が以下の形を持つことを意味する。
 $p_{X_0 X_1 \dots X_L}(x_1, x_2, \dots, x_L)$

$$= p_{X_0}(x_0) \prod_{i=1}^L p_{X_i|X_0}(x_i|x_0) \quad (4)$$

Gelfand and Pinsker [7] は、情報源が CI 条件を満たす場合に領域 $\mathcal{R}_L(0)$ が L 個の補助確率変数と情報理論的用量を用いた形で表現できることを示した。

3.2 CEO 問題

情報源 $(X_0, X_1, \dots, X_L)$ が (4) で与えられる CI 条件を満たす場合について、 $R_0 = 0$ の場合の符号化、即ち補助情報の符号化データのみから、情報源出力 X_0 を復号する問題は、CEO 問題とよばれている。これは、図 4 に示すように、補助情報のみから情報センシングを行うシステムとみなすこともできる。CEO は、最高経営責任者 (Chief Executive Officer) の略である。CEO 問題の枠組みを図 5 に示す。 X_0 が与えられた下での $X_i, i = 1, 2, \dots, L$ の条件付確率分布 $p_{X_i|X_0}$ の定義する雑音通信路を

$$W_i = p_{X_i|X_0} = \{p_{X_i|X_0}(x_i|x_0)\}_{(x_0, x_i) \in \mathcal{X}_0 \times \mathcal{X}_i}$$

とおく。CI 条件より、信号 X_0 は、各 $i = 1, 2, \dots, L$ について、統計的に独立な雑音を通信路 W_i から被る。各符号器 φ_i は、通信路 W_i から出力される信号 X_0 の雑音版 X_i を観測し、これを符号化して、CEO へ送る。CEO は、送られてくる L 個の符号化データから復号器 ψ を用いて信号 X_0 の推定値 $\hat{X}_0$ を出力する。

CEO 問題は、Berger, Zhang and Viswanathan [8] により、定式化され研究された。彼らは、雑音通信路 W_i が同一の確率特性をもつ場合、即ち $\mathcal{X}_i = \mathcal{X}_1, W_i = W_1, i = 2, \dots, L$ となる場合を扱い、伝送率・歪領域 $\mathcal{R}_L(D) \cap \{R_0 = 0\}$ について、伝送率ベクトルの和の下限と D との関係を表す以下の量を定義した。

$$R_L(D) \triangleq \min_{(0, R_1, \dots, R_L) \in \mathcal{R}_L(D)} \left\{ \sum_{i=1}^L R_i \right\}$$

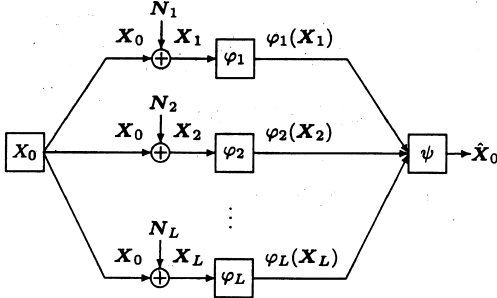


図6 ガウス CEO 問題の枠組み

さらに、 L が十分大の場合を考え、 $R_L(D)$ の極限

$$R(D) \triangleq \lim_{L \rightarrow \infty} R_L(D)$$

の関数形を論じた。 $X_0 = x \in \mathcal{X}_0$ が与えられた下での確率変数 X_i の条件付分布を確率分布にもつ確率変数を $X_i(x)$ と書く。Berger ら [8] は、 $X_1(x), X_2(x), \dots$ が有限個の未知パラメータ $x \in \mathcal{X}_0$ により指定される独立同分布 (i.i.d.) 確率変数列になることに着目し、統計的仮説検定問題と情報源符号化の手法を組み合わせた手法を用いて、CEO 問題を考察した。その結果、 D が十分小さいときの $R(D)$ の漸近式が

$$R(D) \approx \alpha(p_{X_0}, W_1) \log \frac{1}{D} \quad (5)$$

の形になることを示し、 $\alpha(p_{X_0}, W_1)$ の計算可能な表式を得た。有限の L に対する $R_L(D) \cap \{R_0 = 0\}$ の決定問題は依然として未解決である。Berger ら [8] は、 L が無限大の場合の漸近論を考察することにより、 L の有限性に起因する難しさがある意味で回避したといえる。また、 L が十分大きく、符号器の容量制約の総計が R 以下に押えらるという設定は、センサネットワークにおいて、実用上も意味のある問題設定といえる。以上の点で彼らの得た結果は大変興味深い。

4. 情報源出力が連続値をとる場合

情報源からの出力データが連続値をとる場合について、 $L = 1, R_1 = \infty$ の場合の R_0 の下限を求める問題は、Wyner [11] が考察し、また、 $L = 2$ で $R_0 = 0$ の場合は、Yamamoto and Ito [12], Flynn and Gray [13] が議論し、 $R_L(D) \cap \{R_0 = 0\}$ の内界を得ている。情報源出力が連続値をとる場合の一般論に関しては、これら以外の結果は殆んど得られていない。一方、情報源がガウス型の場合については、幾つかの具体的結果が得られている。ここではそれらの結果を述べる。

4.1 ガウス型情報源の場合

ここでは、情報源がガウス型情報源で、歪み測度 d が 2 乗誤差 $d(x, y) = (x - y)^2$ の場合を考える。 $L = 1, R_1 \rightarrow \infty$ の場合は、Wyner and Ziv [4], Wyner [11] が解決した。Oohama [14] は、彼らが得た結果の拡張を試み、 $L = 1$ の場合の領域 $\mathcal{R}_1(D)$ を決定した。さらに、 L 補助者問題について、情報源が有限集合に値をとる場合に Gelfand and Pinsker [7] が得た結果を情報源がガウス型で、2 乗誤差歪みの場合へ拡張した。ガウス型情報源の場合、CI 条件は、 $X_i = X_0 + N_i, i = 1, 2, \dots, L$ と書ける。ここで、 X_0 は、

平均 0、分散 $\sigma_{X_0}^2$ のガウス型確率変数、 $N_i, i = 1, 2, \dots, L$ は、平均 0、分散 $\sigma_{N_i}^2$ の独立なガウス型確率変数であり、これらは、 X_0 とも独立である。Oohama [15] は、情報源が上記の CI 条件を満たす場合に領域 $\mathcal{R}_L(D)$ を決定した。この結果は $L = 1$ の場合に Oohama [14] が得た結果を特別な場合として含む。結果を述べるために $\Lambda = \{1, 2, \dots, L\}$ とおき、また、 $S \subseteq \Lambda$ に対し、 $r_S = \{r_i\}_{i \in S}$ なる記法を用いる。 $D > 0, r_i \geq 0, i \in \Lambda$ および $S \subseteq \Lambda$ に対し、次の量を定義する。

$$J_S(r_S, D)$$

$$= \frac{1}{2} \log^+ \left[\left\{ \frac{1}{\sigma_{X_0}^2} + \sum_{i \in S} \frac{1}{\sigma_{N_i}^2} (1 - 2^{-2r_i}) \right\}^{-1} \cdot \frac{1}{D} \right]$$

ここで、 $S = \emptyset$ に対しては、 $J_\emptyset(r_\emptyset, D) = J(D)$ と記す。さらに、次の集合を定義する。

$$\mathcal{R}_L^*(D)$$

$$= \{(R_0, R_1, \dots, R_L) : \text{ある } r_i \geq 0, i \in \Lambda$$

が存在し、任意の $i \in \Lambda$ と任意の $S \subseteq \Lambda$ に対し

$$R_i \geq r_i$$

$$R_0 + \sum_{i \in S} R_i \geq J_{\Lambda-S}(r_{\Lambda-S}, D) + \sum_{i \in S} r_i \}.$$

このとき次の結果が成り立つ。

[定理 1] ガウス型情報源が CI 条件を満たすとき、 $\mathcal{R}_L(D) = \mathcal{R}_L^*(D)$ が成り立つ。

CI 条件を満たさない一般のガウス型情報源については、Pandya ら [16] が領域の外界を計算した。彼らは、Oohama [14] が $L = 1$ の場合に用いた手法を $L \geq 2$ の場合に適用することにより結果を得た。この結果は外界の評価としては弱く、完全解は得られていない。

4.2 ガウス CEO 問題

情報源が CI 条件を満たすガウス型で、平均 2 乗歪の場合において、 $R_0 = 0$ のときの伝送率・歪領域 $\mathcal{R}_L(D) \cap \{R_0 = 0\}$ を求める問題は、ガウス CEO 問題とよばれる。ガウス CEO 問題の枠組みを図 6 に示す。定理 1 より、この問題の解は、 $\mathcal{R}_L^*(D) \cap \{R_0 = 0\}$ で与えられる。ガウス CEO 問題は、Yamamoto and Ito [12] が初めて考察し、 $L = 2$ で $\sigma_{N_1}^2 = \sigma_{N_2}^2$ の場合に、 $\mathcal{R}_2(D) \cap \{R_0 = 0\}$ の内界を計算した。彼らの得た内界は最適である。即ち $\mathcal{R}_2^*(D) \cap \{R_0 = 0\}$ と一致する。

以下では、定理 1 より、Oohama [17] がガウス CEO 問題に対して得た結果が簡単に導出できることを述べる。ガウス CEO 問題において、 $N_i, i = 1, 2, \dots, L$ の雑音レベルが同一の値 $\sigma_{N_i}^2 = \sigma^2$ である場合に $\mathcal{R}_L(D), R(D)$ を CEO 問題と同様に定義する。雑音レベル σ が 0 の場合、各符号器は X_0 そのものを観測できる。よって、(2) で $R_0 = 0$ とすることにより、直ちに

$$R(D) = R_L(D) = R_{X_0}(D) = \frac{1}{2} \log^+ \left[\frac{\sigma_{X_0}^2}{D} \right] \quad (6)$$

を得る。雑音レベルが正の場合に、分散符号化による性能劣化を明らかにすることが問題である。

ガウス CEO 問題において、 $R(D)$ を求める問題は、Viswanathan and Berger [18] が初めて研究した。彼らは、 $X_1(x), X_2(x), \dots$ が $x \in \mathcal{X}_0$ を未知の実数パラメータとする i.i.d. 確率変数列とみなせることに注目し、統計学にお

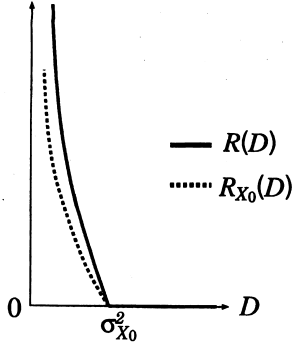


図7 $R(D)$ および $R_{X_0}(D)$ の概形

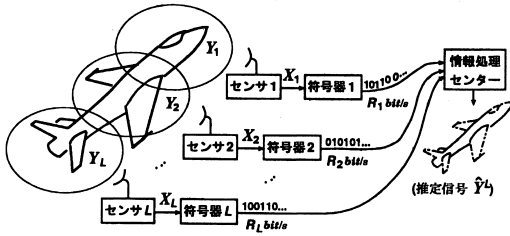


図8 観測信号が同一でない場合の情報センシング

けるパラメータ推定問題と情報源符号化を組み合わせた符号化の手法によりガウス CEO 問題を考察し、 $R(D)$ の $D \ll 1$ における漸近形の上界、下界を導出した。一方、Oohama [17] は、彼らとは異なる多端子情報源符号化の立場からこの問題を考察し、 $R(D)$ の形を完全に決定した。

さて、簡単な計算により、定理 1 から次の系を得る。

[系 1]

$$R_L(D) = -\frac{L}{2} \log \left(1 - \frac{2}{L} \hat{R}(D) \right) + R_{X_0}(D) \quad (7)$$

ここで、

$$\hat{R}(D) \triangleq \frac{\sigma_{X_0}^2}{2\sigma_{X_0}^2} \left[\frac{\sigma_{X_0}^2}{D} - 1 \right]^+, [a]^+ \triangleq \max\{a, 0\}.$$

系 1 の (7) において、 $L \rightarrow \infty$ とすると

$$R(D) = (\log e) \cdot \hat{R}(D) + R_{X_0}(D) \quad (8)$$

となり、Oohama [17] の得た結果が導かれる。(8) の右辺第一項である $(\log e) \cdot \hat{R}(D)$ が分散符号化による性能劣化を表す項になる。 $R(D)$ および $R_{X_0}(D)$ の概形を図 7 に示す。正数 D が十分小さいとき、(8) より

$$R(D) \approx (\log e) \cdot \frac{\sigma^2}{2\sigma_{X_0}^2} \cdot \frac{\sigma_{X_0}^2}{D} \quad (9)$$

なる漸近式が成立する。この式は Viswanathan and Berger [18] が予想した。Oohama [17] の得た $R(D)$ の完全形は、彼らの予想が正しいことを示している。

4.3 観測信号が同一でない場合

これまで、各センサが、同一の信号 X_0 の雑音信号を観測する場合を扱ってきた。一般には、図 8 に示すように、必ずしも同一でない信号が、雑音に汚された形で観

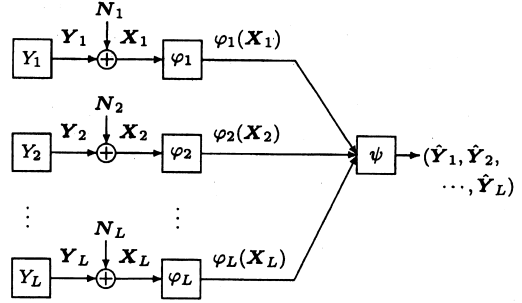


図9 同一でない情報信号の推定の理論的モデル

測される場合があり得る。このような場合の情報センシングの理論的モデルが図 9 に示されている。図において $Y^L = (Y_1, Y_2, \dots, Y_L)$ を相関を有する L 個のガウス型確率変数からなるベクトルとし、 $X_i = Y_i + N_i, i \in \Lambda$, とする。各 X_i は、 Y_i の雑音に汚された観測値である。このシステムにおいて、平均 2 乗歪

$$\Delta(Y^L, \hat{Y}^L) = \frac{1}{n} \sum_{i=1}^L \sum_{t=1}^n (Y_{i,t} - \hat{Y}_{i,t})^2 \quad (10)$$

を与えられた値 D 以下にするような伝送率ベクトル $R^L = (R_1, R_2, \dots, R_L)$ の占める領域 $\hat{\mathcal{R}}_L(D)$ を求めることが基本問題である。この問題は、Pandya et. al [19], Zhang and Wicker [20] が考察した。最近、Oohama [21], [22] は、伝送率・歪領域 $\hat{\mathcal{R}}_L(D)$ の内界と外界を計算し、さらに Oohama [22] は情報源の相関が、それほど強くない場合に、両者が一致することを示した。

5. 関連する話題と今後の課題

分散センサーからの符号化データを用いて発生信号の値を推定する問題をこれまで扱ってきた。集められた符号化データをもとに如何なる情報処理を行うかにより、様々な符号化問題が提起される。例えば分散センサーからの符号化データに基づき、発生信号の分布に関する統計的推論を行う問題が考えられる。これは、情報理論における符号化の理論と、統計学における仮説検定、未知母数推定の理論の双方に関わる新しい理論的枠組を提起する。この方面の研究については、Han and Amari [23] に詳しい解説がある。特に、Han and Kobayashi [24] は、実際的な設定の下で、この問題を考察し、結果を得ている。

本稿では、センサーの配置が分散型の場合を扱った。センサーの配置が縦列型の場合については、システムの定式化と理論解析が Draper and Wornell [25] によってなされている。またこの論文では、縦列配置型センサネットワークと多端子情報源符号化における縦列接続型符号化システムとの関係が述べられている。縦列接続型符号化システムは、Yamamoto [26] により初めて定式化され、研究された。その後、Gu and Effros [27], [28] が拡張を行っている。これらの研究成果がセンサネットワークの理論的モデルとしてどのような意味を持つかについては、今後さらに検討を進めていく必要があると思われる。

多端子情報源符号化の分野では、1974 年の Slepian and Wolf [1] の研究以来、様々な多端子通信システムの符号化問題が研究されてきた。最近、山本 [29] により、このような符号化問題に対するこれまでの研究成果が紹介されてい

る。これらの中には、本稿では、触れられなかったものも幾つかある。それらの符号化システムとセンサネットワークとの関連について、検討の余地があると思われる。

理論限界を達成する符号器、復号器の設計問題については、本稿では全く触れることができなかった。通信路符号化におけるシャノン限界に迫る符号手法として LDPC 符号、ターボ符号がある。これらの符号化技術を用い、多端子符号化システムの理論限界を実現する通信システムの構築の試みが、最近欧米において行われてきている。

本稿で述べてきた様に、センサネットワークの理論的定式化により提起される多端子情報源の分散符号化システムに関しては、現在也未解決の問題が多数ある。これらの問題の解決は容易ではないが、CEO 問題のように現実のシステムにおいて意味のある設定や状況を考えることは、問題解析の手法を拡張、新たな理論展開を生み出す可能性を持つ。センサネットワークによる情報センシングの理論の構築には、多端子情報理論の枠組みとともに、現実のシステムの本質を捉え、これを理論的モデルに反映させる視点が必要であると思われる。

文 献

- [1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471-480, July 1973.
- [2] A. D. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 294-300, May 1975.
- [3] R. F. Ahlswede and J. Körner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 629-637, Nov. 1975.
- [4] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 1-10, Jan. 1976.
- [5] T. Berger, K. B. Houswright, J. K. Omura, S. Tung, and J. Wolfowitz, "An upper bound on the rate distortion function for source coding with partial side information at the decoder," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 664-666, Nov. 1979.
- [6] J. Körner and K. Marton, "How to encode the module-two sum of binary sources," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 219-221, Mar. 1979.
- [7] S. I. Gelfand and M. S. Pinsker, "Source coding with incomplete side information," (in Russian) *Probl. Pered. Inform.*, vol. 15, no. 2, pp. 45-57, 1979.
- [8] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inform. Theory*, vol. 42, pp. 887-902, May 1996.
- [9] 大濱靖匡, "多端子情報理論とセンサネットワーク," 電子情報通信学会誌小特集「センサネットワーク」, pp.367-373, 2006 年 5 月.
- [10] 大濱 靖匡, "センサネットワークの理論と多端子情報理論," 電子情報通信学会技術研究報告, SN2006-19, pp.7-14, 2006 年 5 月.
- [11] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder-II: General sources," *Inform. Contr.*, vol. 38, pp. 60-80, July 1978.
- [12] H. Yamamoto and K. Itoh, "Source coding theory for multiterminal communication systems with a remote source," *Trans. of the IECE of Japan*, vol. E63, no.10, pp. 700-706, Oct. 1980.
- [13] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 773-787, Nov. 1987.
- [14] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1912-1923, Nov. 1997.
- [15] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2577-2593, July 2005.
- [16] A. Pandya, A. Kansal, G. Pottie and M. Srivastava, "Lossy source coding of multiple Gaussian sources: m-helper problem" *Proceedings of IEEE Information Theory Workshop, San Antonio, TX*, pp. 34-38, Oct. 2004.
- [17] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1057-1070, May 1998.
- [18] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1549-1559, Sept. 1997.
- [19] A. Pandya, A. Kansal, G. Pottie and M. Srivastava, "Fidelity and resource sensitive data gathering," *Proceedings of the 42nd Allerton Conference, Allerton, IL*, June 2004.
- [20] X. Zhang and S. Wicker, "On the rate region of the vector Gaussian CEO problem," *Proceedings of the 2005 Conference on Information Science and Computer Engineering*, the Johns Hopkins University, March 2005.
- [21] Y. Oohama, "Rate distortion region for separate coding of correlated Gaussian remote observations," *Proceedings of the 43rd Allerton Conference, Allerton, IL*, pp. 2237-2246, Sept. 2005.
- [22] Y. Oohama, "Separate source coding of correlated Gaussian remote sources," *Proceedings of Information Theory & Applications Inaugural Workshop*, UCSD, CA, Feb. 6-10, 2006.
- [23] T. S. Han and S. Amari, "Statistical inference under multiterminal data compression," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2300-2324, Oct. 1998.
- [24] T. S. Han and K. Kobayashi, "Multiterminal filtering for decentralized detection systems," *IEICE Trans. Commun.*, vol. E75-B, pp. 437-444, June 1992.
- [25] S. C. Draper and G. W. Wornell, "Side information aware coding strategies for sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 966 - 976, Aug. 2004.
- [26] H. Yamamoto, "Source coding theory for cascade and branching communication systems," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 299-308, May 1981.
- [27] W.-H. Gu and M. Effros, "Source coding for a multi-hop network," *Proceeding of the 2005 Data Compression Conference*, p. 461, Snowbird, Utah, March 2005.
- [28] W.-H. Gu and M. Effros, "Source coding for a multihop network," *Proceeding of the 2005 International Symposium on Information Theory*, pp. 2335-2339, Adelaide, Australia, Sept. 2005.
- [29] 山本博資, "[招待講演] 通信ネットワークへのレートひずみ理論の拡張," 電子情報通信学会技術研究報告, IT2005-74, pp. 57-62, 2006 年 3 月.