

Dynamic Quantization for Control

Toshiharu Sugie, Shun-ichi Azuma, and Yuki Minami

Abstract This paper overviews a series of the authors' recent contributions to dynamic quantizer design for control. The problem considered here is to find a dynamic quantizer such that the resulting quantized system is an optimal approximation of an ideal unquantized system. We show here a fundamental solution to this problem and briefly review several results toward real applications.

1 Introduction

As a bridge between the infinite and finite worlds, the quantization has been a key issue in science and engineering fields. For example, it can be seen in

- *signal processing* analog-to-digital conversion,
- *information theory* source coding,
- *statistics* cluster analysis,
- *operations research* facility location.

By the quantization, one can convert infinitely large number of noise-corrupted data to compact data.

Also in the systems and control field, the quantization has attracted much attention in last decade, due to the increasing need for hybrid control and networked control (e.g., see [1, 2, 3, 4]). From various points of view, many results have been obtained so far; for example, the quantizer coarseness for

Toshiharu Sugie and Shun-ichi Azuma
Graduate School of Informatics, Kyoto University; Uji, Kyoto 611-0011, Japan; e-mail: {sugie, sazuma}@i.kyoto-u.ac.jp

Yuki Minami
Department of Control Engineering, Maizuru National College of Technology;
Maizuru, Kyoto 625-8511, Japan; e-mail: minami@maizuru-ct.ac.jp

stabilization has been characterized in e.g., [5, 6, 7, 8, 9], and good quantizers (or switching controllers) for control have been developed in e.g., [1, 2, 10, 11].

On the other hand, the authors have been interested in “dynamic quantizers”, which map continuous-valued signals into discrete-valued ones depending on the past history of both signals. Compared with static quantizers, dynamic quantizers have much better performance, which has motivated us to pursue their potential for control. So far, for a class of dynamic quantizers, called the $\Delta\Sigma$ modulators in the signal processing community [12], we have obtained several key results which clarify the optimal quantization structure and the performance limitation in control systems. In this paper, we briefly review a series of the authors’ results in [13, 14, 15, 16, 17].

The problem addressed here is as follows: when a plant and a controller are given for the quantized feedback system in Fig. 1 (a), find a dynamic quantizer such that the system in (a) optimally approximates the usual feedback system in Fig. 1 (b), in terms of the input-output relation. If the problem is solved with small approximation error, one can directly apply controllers designed for the usual system in (b) to the quantized system in (a), even if the plant input is restricted to belonging to a *fixed* discrete set. This gives a big advantage to construct quantized control systems subject to discrete-valued signal constraints.

In the following sections, we first derive an expression of the performance of dynamic quantizers. Based on this, an optimal dynamic quantizer in a closed form is presented. Finally, the authors’ recent studies toward real applications are introduced.

Notation: Let $\mathbf{R}$, $\mathbf{R}_+$, and $\mathbf{N}$ be the real number field, the set of positive real numbers, and the set of natural numbers (positive integers), respectively. We denote by 0 the zero matrix of appropriate dimensions. For the matrix $M := \{M_{ij}\}$, let $\text{abs}(M)$ denote the matrix composed of the absolute value of each element, i.e., $\text{abs}(M) = \{|M_{ij}|\}$, and let M^+ be the pseudo-inverse. For the vector sequences $X := (x_1, x_2, \dots)$ and $Y := (y_1, y_2, \dots)$, we use $X - Y$ to express the vector sequence $(x_1 - y_1, x_2 - y_2, \dots)$. For the vector x , the matrix M , and the vector sequence X , the symbols $\|x\|$, $\|M\|$, and $\|X\|$ express their ∞ -norms (i.e., $\|X\| := \sup_{i \in \mathbf{N}} \|x_i\|$).

2 Dynamic Quantizer Design Problem

Consider the feedback system Σ_Q in Fig. 1 (c), which is a generalized version of the quantized feedback system in (a).

The system G is given by

$$G : \begin{cases} x(k+1) = Ax(k) + B_1r(k) + B_2v(k), \\ z(k) = C_1x(k) + D_1r(k), \\ u(k) = C_2x(k) + D_2r(k) \end{cases} \quad (1)$$

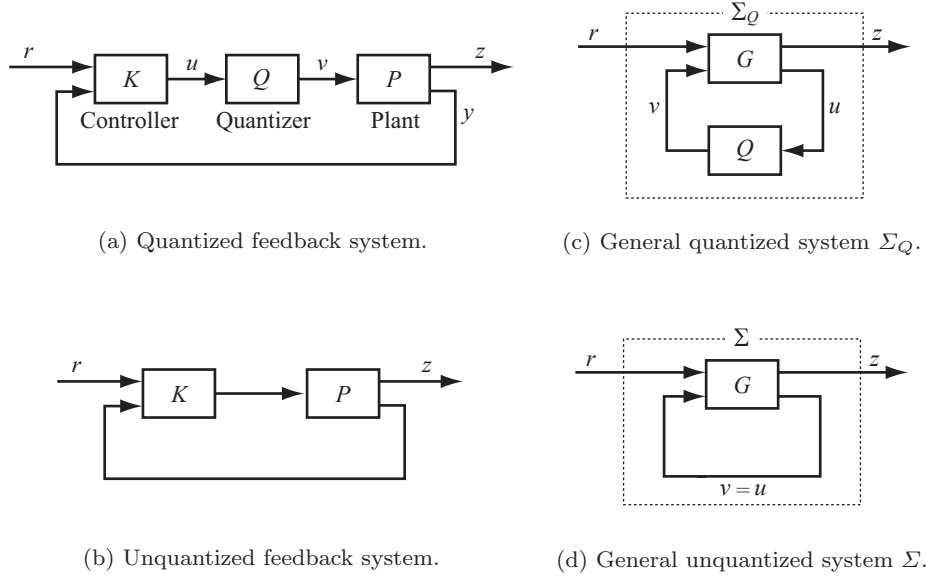


Fig. 1 Quantized and unquantized (usual) feedback systems.

where $x \in \mathbf{R}^n$ is the state, $r \in \mathbf{R}^p$ and $v \in \mathbf{R}^m$ are the inputs, $z \in \mathbf{R}^l$ and $u \in \mathbf{R}^m$ are the outputs, $k \in \{0\} \cup \mathbf{N}$ is the discrete time, and $A \in \mathbf{R}^{n \times n}$, $B_1 \in \mathbf{R}^{n \times p}$, $B_2 \in \mathbf{R}^{n \times m}$, $C_1 \in \mathbf{R}^{l \times n}$, $C_2 \in \mathbf{R}^{m \times n}$, $D_1 \in \mathbf{R}^{l \times p}$, $D_2 \in \mathbf{R}^{m \times p}$ are constant matrices. The initial state is given as $x(0) = x_0$ for $x_0 \in \mathbf{R}^n$.

On the other hand, Q is the dynamic quantizer in the following form:

$$Q : \begin{cases} \xi(k+1) = A\xi(k) + B_1u(k) + B_2v(k), \\ v(k) = q(\mathcal{C}\xi(k) + u(k)) \end{cases} \quad (2)$$

where $\xi \in \mathbf{R}^N$ is the state of dimension N , $u \in \mathbf{R}^m$ is the input, $v \in \mathbf{V}^m := \{0, \pm d, \pm 2d, \dots\}^m$ is the output, and $A \in \mathbf{R}^{N \times N}$, $B_1, B_2 \in \mathbf{R}^{N \times m}$, $\mathcal{C} \in \mathbf{R}^{m \times N}$ are constant matrices. The function $q : \mathbf{R}^m \rightarrow \mathbf{V}^m$ is the nearest-neighbor static quantizer, where $\mathbf{V}$ is the discrete set on which each output takes its value and $d \in \mathbf{R}_+$ is the quantization level. The initial state is given as $\xi(0) = 0$ for guaranteeing that Q is drift-free, i.e., $v(k) = 0$ for $u(k) = 0$ ($k = 0, 1, \dots$). This quantizer determines its output depending upon its current input and past input sequence.

Next, we prepare some symbols. For the system Σ_Q , let $Z_Q(x_0, R)$ denote the controlled output sequence $(z(1), z(2), \dots, z(\infty))$ for the initial state x_0 and the reference input $R := (r_0, r_1, \dots) \in \ell_\infty^p$ (i.e., $x(0) = x_0$ and $r(k) = r_k$), and let $z_Q(k, x_0, R)$ be the output at time k . In addition, we consider the

feedback system Σ in Fig. 1 (d), corresponding to a generalized version of (b), for which the symbols $Z(x_0, R)$ and $z(k, x_0, R)$ are similarly defined. Then as a performance index of Q , we define the *maximum output difference*:

$$E(Q) := \sup_{(x_0, R) \in \mathbf{R}^n \times \ell_\infty^p} \|Z_Q(x_0, R) - Z(x_0, R)\|. \quad (3)$$

This is the difference between the system Σ_Q in Fig. 1 (c) and the ideal system Σ in (d), in terms of the input-output relation. Then our problem is formulated as follows.

Problem 1. For the system Σ_Q , suppose that the quantization level $d \in \mathbf{R}_+$ is given and assume that Σ is stable (the matrix $A + B_2 C_2$ is Schur).

- (i) Determine the value of $E(Q)$ for given Q .
- (ii) Find a Q (i.e., parameters $(N, \mathcal{A}, \mathcal{B}_1, \mathcal{B}_2, \mathcal{C})$) minimizing $E(Q)$ under the stability condition:

$$\sup_{k \in \mathbf{N}} \|x(k)\| < \infty, \quad \sup_{k \in \mathbf{N}} \|\xi(k)\| < \infty \quad (\forall (x_0, R) \in \mathbf{R}^n \times \ell_\infty^p). \quad (4)$$

□

Problems (i) and (ii) respectively correspond to the analysis and design problems. If $E(Q)$ is small, the input-output relation of the ideal system Σ is almost preserved in Σ_Q . This provides us a practical method of control systems design with discrete-valued signal constraints. For example, consider the feedback system in Fig. 1 (a), and suppose that P has to be actuated by discrete-valued signal. Then the performance would be good with

- any controller K achieving desirable performance in the ideal system in Fig. 1 (b),
- any dynamic quantizer Q such that $E(Q)$ is small.

Note that Problem 1 is nonconvex. In fact, since Z_Q involves a discontinuous function q , the function $\|Z_Q(x_0, R) - Z(x_0, R)\|$ is nonconvex with respect to (x_0, R) . Furthermore, the problem (ii) is a minimax problem between (x_0, R) and $(N, \mathcal{A}, \mathcal{B}_1, \mathcal{B}_2, \mathcal{C})$.

3 Analytical Solutions

Under the following assumptions, we can obtain an analytical solution to Problem 1.

- (A1) $\text{rank } D_2 = m$ (D_2 is full row rank).
- (A2) For $\tilde{A} := A + B_2 C_2$, there exists a $k \in \{0\} \cup \mathbf{N}$ such that $C_1 \tilde{A}^0 B_2 = C_1 \tilde{A}^1 B_2 = \dots = C_1 \tilde{A}^{k-1} B_2 = 0$ (if $k \geq 1$) and $\text{rank } C_1 \tilde{A}^k B_2 = l$.
- (A3) The system Σ' in Fig. 2 has no unstable system zero.

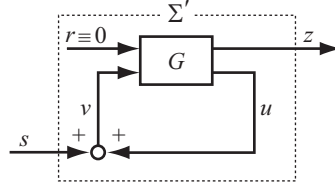


Fig. 2 A subsystem of quantized system Σ_Q .

Note in (A2) that $C_1 \tilde{A}^k B_2$ ($k = 0, 1, \dots$) express the impulse response matrices of the system Σ' in Fig. 2 (from s to z).

Even if these assumptions do not hold, a practical solution to Problem 1 is obtained, which will be explained in Remark 1 and Section 4.

3.1 Performance Expression

A solution to Problem 1 (i) is given as follows.

Theorem 1 ([14, 15]). *For the system Σ_Q , assume (A1) ((A2) and (A3) are not necessary). If*

$$[C_1 \ 0] \begin{bmatrix} \tilde{A} & B_2 \mathcal{C} \\ 0 & \mathcal{A} + B_2 \mathcal{C} \end{bmatrix}^k \begin{bmatrix} 0 \\ \mathcal{B}_1 + B_2 \end{bmatrix} = 0 \quad (\forall k \in \{0\} \cup \mathbf{N}), \quad (5)$$

then

$$E(Q) = \left\| \sum_{k=0}^{\infty} \text{abs} \left([C_1 \ 0] \begin{bmatrix} \tilde{A} & B_2 \mathcal{C} \\ 0 & \mathcal{A} + B_2 \mathcal{C} \end{bmatrix}^k \begin{bmatrix} B_2 \\ \mathcal{B}_2 \end{bmatrix} \right) \right\| \frac{d}{2}; \quad (6)$$

otherwise

$$E(Q) = \infty. \quad (7)$$

□

Theorem 1 gives an exact expression of $E(Q)$, which enables us to compute the value of $E(Q)$ for given dynamic quantizer Q .

The intuitive meaning of this result is as follows. Let us introduce the new variable $w \in [-d/2, d/2]^m$:

$$w(k) := q(\mathcal{C}\xi(k) + u(k)) - (\mathcal{C}\xi(k) + u(k)), \quad (8)$$

which expresses the quantization error of the static quantizer q in (2). This allows us to represent Q as

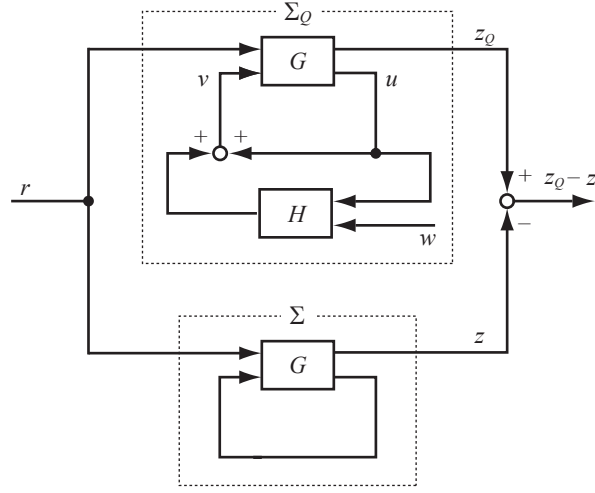


Fig. 3 Error system between quantized system Σ_Q and unquantized system Σ .

$$Q : \begin{cases} \xi(k+1) = (\mathcal{A} + \mathcal{B}_2 \mathcal{C})\xi(k) + (\mathcal{B}_1 + \mathcal{B}_2)u(k) + \mathcal{B}_2 w(k), \\ v(k) = \mathcal{C}\xi(k) + u(k) + w(k) \end{cases} \quad (9)$$

and to formally regard Q as a linear system with the external inputs u and w . With this expression, the error system for Σ_Q and Σ is illustrated as Fig. 3, where H is a subsystem (which is linear) of (9). Then (5) means that the impulse response matrices from r to $z_Q - z$ are zero. Thus if (5) does not hold, $\|z_Q - z\|$ can be arbitrarily large by some large r , which gives (7). On the other hand, the right hand side of (6) is composed of

- the impulse response matrices from w to $z_Q - z$,
- the upper bound of the static quantization error w , i.e., $d/2$.

So it follows that the right hand side represents the influence of the static quantization error on the output difference $z_Q - z$.

Remark 1. Even if (A1) is not satisfied, the weak version of Theorem 1, in which the right-hand side of (6) becomes an upper bound of $E(Q)$, holds. Therefore, although it is rather conservative, the value of $E(Q)$ can be estimated. $\square$

3.2 Optimal Dynamic Quantizers

We next show a solution to Problem 1 (ii).

Theorem 2 ([13, 15]). *For the system Σ_Q , assume (A1)–(A3). Then a solution to Problem 1 (ii) is given by*

$$Q^* : \begin{cases} \xi(k+1) = \tilde{A}\xi(k) - B_2u(k) + B_2v(k), \\ v(k) = q(-(C_1\tilde{A}^\tau B_2)^+ C_1\tilde{A}^{\tau+1} \xi(k) + u(k)) \end{cases} \quad (10)$$

and the minimum value of $E(Q)$ is given by

$$E(Q^*) = \|C_1\tilde{A}^\tau B_2\| \frac{d}{2} \quad (11)$$

where τ is the value of k satisfying the condition in (A2). $\square$

Theorem 2 provides an optimal quantizer, where Assumptions (A1) and (A2) relate to the minimality of $E(Q^*)$ and (A3) does to the stability of Σ_Q .

This result explains an optimal quantization structure as follows. Suppose that Q^* is applied to the error system in Fig. 3. Then the impulse response matrices from r to $z_Q - z$ are $(0, 0, \dots)$, and those from w to $z_Q - z$ are given by

$$(0, \dots, 0, \underset{\substack{\uparrow \\ \tau\text{-th}}}{C_1\tilde{A}^\tau B_2}, 0, 0, \dots) \quad (12)$$

which, actually, corresponds to the minimum. So Q^* plays a role to satisfy (5) and to reduce the signal transfer from w to $z_Q - z$ as small as possible.

Example 1. Consider the system Σ_Q for the feedback system in Fig. 1 (a). Here, P and K are the discrete-time plant and controller obtained from the continuous-time ones

$$\begin{aligned} P_c : \begin{cases} \dot{x}_P(t) = \begin{bmatrix} 0.1 & 3 \\ -0.8 & 2 \end{bmatrix} x_P(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} v(t), \\ z(t) = [1 \ 2]x_P(t), \quad y(t) = [1 \ 0]x_P(t), \end{cases} \\ K_c : \begin{cases} \dot{x}_K(t) = \begin{bmatrix} -10 & 3 \\ -14.7 & -6.1 \end{bmatrix} x_K(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} r(t) + \begin{bmatrix} 10.1 \\ 11.5 \end{bmatrix} y(t), \\ u(t) = -[2.4 \ 8.1]x_K(t) + r(t) \end{cases} \end{aligned}$$

and the zero-order hold with the sampling period $h := 0.1$. For Q^* , the quantization level is given by $d := 2$.

Fig. 4 shows the simulation result on the time responses of the system Σ_Q , where $x_0 := [0.5 \ -0.5 \ 0 \ 0]^\top$ ($x := [x_P^\top \ x_K^\top]^\top$) and $r(k) \equiv 0$. In addition, the output response of Σ in Fig. 1 (b) is also shown by the thin line in the third figure, where x_0 and r are set to the same values. Though v is a coarse discrete-valued signal in Σ_Q , we see that the output behavior of Σ_Q is almost the same as that of Σ . $\square$

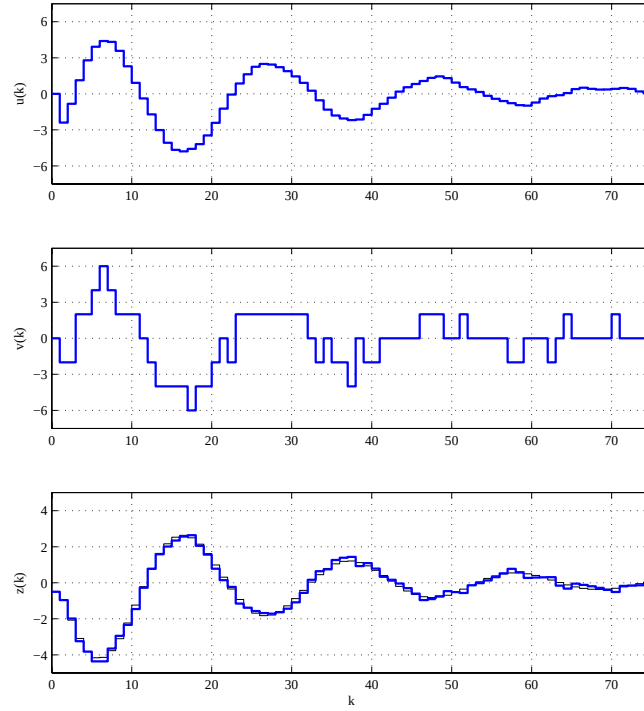


Fig. 4 Responses of optimally quantized system Σ_Q (thick lines) and output response of unquantized system Σ (thin line).

4 Advanced Topics

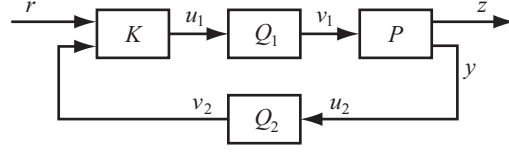
To apply to real systems, the above basic theory has to be generalized. In this section, we introduce recent results toward this direction.

Numerical optimization based design of dynamic quantizers

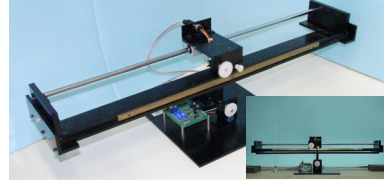
Since the above discussion holds under several assumptions, the results (especially, Theorem 2) can be applied to a limited class of systems. As an alternative, a design method based on numerical optimization has been developed in [14]. There, by exploiting special structure of Problem 1, the problem is reduced into a linear programming problem, which allows us to efficiently optimize the dynamic quantizer without any strong assumption. A MATLAB implementation entitled “ODQ toolbox” is available at the web site [18].

Decentralized dynamic quantizers

As shown in Fig. 5 (a), it is often necessary to have a decentralized structure in the quantizer. In [16], the above results have been extended to the decentralized case. Furthermore, an experimental evaluation has been per-

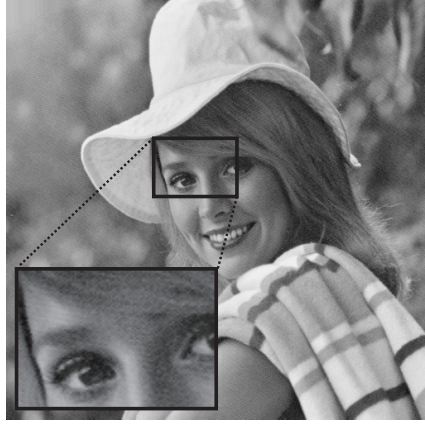


(a) Quantized feedback system with decentralized quantizer.

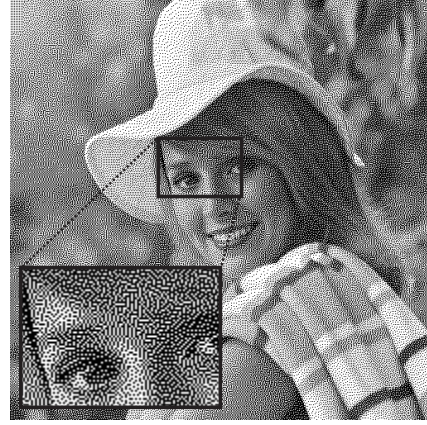


(b) Seesaw-cart system.

Fig. 5 Decentralized dynamic quantizers [16].



(a) Original image (8-bit colors).



(b) Halftone image (2-bit colors).

Fig. 6 Binary halftoning by a 2D optimal dynamic quantizer [17].

formed with the seesaw-cart system in Fig. 5 (b). There, it is successfully achieved to stabilize the unstable system under the severe condition that the plant input takes one of three values and the controller input does one of seven values.

n -dimensional dynamic quantizers

In [17], the authors have extended the optimal quantizers to an n -dimensional (n -D) version. This can be used not only for control of n -D systems but also for image processing. Fig. 6 shows an example of applying the result to the halftone image processing, which is to transform a grayscale image to a binary image keeping the quality to the eye. This shows the potential of our framework to other fields out of control.

5 Conclusion

The authors' recent results on the control-oriented dynamic quantizers have been reviewed. We hope that this will be utilized to real control applications.

Acknowledgment: The authors would like to thank Mr. Ryosuke Morita, Kyoto University, for supplying data.

References

1. A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, Vol. 35, No.3, pp. 407–427, 1999.
2. D. Liberzon, *Switching in systems and control*, Birkhauser, Boston, 2003.
3. G.N. Nair, F. Fagnani, S. Zampieri, and R.J. Evans, "Feedback control under data rate constraints: an overview," *Proc. of the IEEE*, Vol. 95, No. 1, pp. 108–137, 2007.
4. G.C. Goodwin, E.I. Silva, and D.E. Quevedo, "A brief introduction to the analysis and design of networked control systems," *Proc. of Chinese Control and Decision Conference*, pp. 1–13, 2008.
5. W.S. Wong and R.W. Brockett, "Systems with finite communication bandwidth constraints II: Stabilization with limited information feedback," *IEEE Trans. on Automatic Control*, Vol. 44, No. 5, pp. 1049–1053, 1999.
6. R.W. Brockett and D. Liberzon, "Quantizer feedback stabilization of linear systems," *IEEE Trans. on Automatic Control*, Vol. 45, No. 7, pp. 1279–1289, 2000.
7. N. Elia and S.K. Mitter, "Stabilization of linear systems with limited information," *IEEE Trans. on Automatic Control*, Vol. 46, No. 9, pp. 1384–1400, 2001.
8. M. Fu and L. Xie, "The sector bound approach to quantized feedback control," *IEEE Trans. on Automatic Control*, Vol. 50, No. 11, pp. 1698–1711, 2005.
9. D. Liberzon and D. Nesic, "Input-to-state stabilization of linear systems with quantized state measurements," *IEEE Trans. on Automatic Control*, Vol. 52, No. 5, pp. 767–781, 2007.
10. D.E. Quevedo, G.C. Goodwin, and J.A. De Dona, "Finite constraint set receding horizon quadratic control," *Int. J. of Robust and Nonlinear Control*, Vol. 14, No. 4, pp. 355–377, 2004.
11. C. Canudas-de-Wit, F.R. Rubio, J. Fornés, and F. Gómez-Estern, "Differential coding in networked controlled linear systems," *Proc. of 2006 American Control Conference*, pp. 4177–4182, 2006.
12. G.I. Bourdopoulos, *Delta-sigma modulators: modeling, design and applications*, Imperial College Press, 2003.
13. S. Azuma and T. Sugie, "Optimal dynamic quantizers for discrete-valued input control," *Automatica*, Vol. 44, No. 2, pp. 396–406, 2008.
14. S. Azuma and T. Sugie, "Synthesis of optimal dynamic quantizers for discrete-valued input control," *IEEE Trans. on Automatic Control*, Vol. 53, No. 9, pp. 2064–2075, 2008.
15. Y. Minami, S. Azuma, and T. Sugie, "Optimal dynamic quantizers for discrete-valued input feedback control," *Proc. of 46th IEEE Conference on Decision and Control*, pp. 2259–2264, 2007.
16. Y. Minami, S. Azuma, and T. Sugie, "Optimal decentralized dynamic quantizers for discrete-valued input control: a closed form solution and experimental evaluation," *Proc. of 2009 American Control Conference*, pp. 4367–4372, 2009.
17. Y. Minami, S. Azuma, and T. Sugie, "Optimal feedback quantizers for n -dimensional systems with discrete-valued input," to appear in *Nonlinear Analysis: Hybrid Systems*, 2009.
18. R. Morita, S. Azuma, Y. Minami, and T. Sugie, "ODQ Toolbox", www.robot.kuass.kyoto-u.ac.jp/~morita/odqtoolbox_1.0.0b.zip.