A Study of Adaptive Quantization

扈堯盛 / 人力開發部

〈ABSTRACT〉

In this paper, the basic philosophy and adaptation rule of adaptive quantization are described. By computer simulations, some properties of adaptive quantization for a first-order Gauss-Markov process are considered.

I. INTRODUCTION

A logarithmic companding technique is known to enhance the performance of a quantizer with an input that has a large dynamic range [1]. A log-quantizer is still time-invariant, however; and in that sense, it is not the ideal device for quantizing nonstationary inputs.

Serveral schemes have been proposed for using adaptive quantizers which track input signals with variable dynamic range [2]-[8]. These techniques have been applied primarily to speech coding systems for lower bit rates. In Ref. [2], an adjustable uniform quantizer dependent on observation of blocks of quantized samples is investigated. In Ref. [3], a differential RCM coder in which the adaptive quantizer is used together with a fixed first-order predictor in the feedback loop is considered. In Ref. [4], Jayant presents the results of extensive computer simulations undertaken to determine the multiplier coefficients which maximize various performance functionals. In Ref. [5], theoretical aspects of the adaptive quantizer are deeply discussed. In Ref. [6], a quantizer scheme is described for situations in which the signal power is unknown a prior but remains constant for the duration of a communication. In Ref. [7], a robust adaptive quantizer which dissipates the effects of transmission errors is described. It implies imperfect, or sluggish adaptations that may slightly diminish coder performance over error free channel. In Ref. [8], an adaptive quantizer is viewed as one that estimates the variance of its input and normalizes the input by the square root of the estimate. In Ref. [9], the performance limits are described for different speech-encoding schemes including adaptive quantization and adaptive prediction schemes. In Ref. [10], a mathematical analysis of an adaptive quantizer is described. Adaptive quantization has provided a better quantization error performance than companded quantization if the input is nonstationary.

The basic principle of the adaptation is that in each time interval the quantizer characteristic is controlled by the input or output signal with the aim of maintaining a constant loading factor [7]. Thus the quantizer range tends to track the dynamic range of the input. An adaptive quantizer can be viewed either as one with a variable range of one that normalizes the input with a variable and uses a fixed range [8]. The idea of the former model is to adapt the quantizer step size to an input value. The goal of the latter adaptation strategy is to offer a unit level input to a fixed quantizer.

A symmetric uniform quantizer and adaptation rule are described in Section II. In Section III, the bias function is defined and design rules are considered. In Section IV, by computer simulations of a first-order Gauss-Markov input, some properties of adaptive quantizer are discussed.

II. THE ADAPTIVE QUANTIZER

The symmetric uniform quantizer that we consider is shown in Fig. 1.

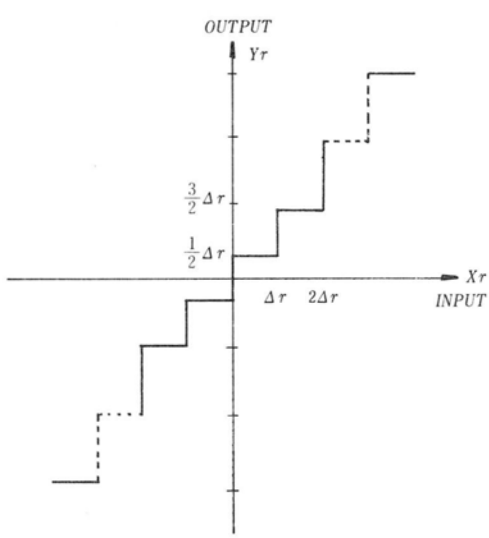

Fig. 1. *Symmertric Uniform Quantizer.*

For a nonnegative input, the output of the uniform quantizer are of the form

$$Y_r = (2I_r - 1)\frac{\Delta_r}{2} \quad \cdots\cdots (1)$$

where $I_r = k$ for $(k-1)\,\Delta_r \leq X_r < k\,\Delta_r$

By the symmetry, the response to negative inputs is

$$Y_r(X_r) = -Y_r(-X_r), \text{ for } X_r < 0 \quad \cdots\cdots (2)$$

The step size Δ_r is scaled by a multiplicative factor determined by the magnitude of previous output level:

$$\Delta_{r+1} = \Delta_r M(I_r) \quad \cdots\cdots (3)$$

In a quantizer with 2N output levesl, an output of $I_r=1$ suggests that the quantizer is probably underloaded, and that Δ_r should be decreased by a factor M(1)< 1. Conversely, an output sample of I_r =N suggests that the quantizer may be overloaded, and that Δ_r should be increased by a factor M(N)< 1. It follows that the N multipliers should satisfy [5]-[7].

$$M(1) < 1,\ M(N) > 1 \quad \cdots\cdots (4)$$

$$M(1) \leq M(2) \leq \ldots \leq M(N) \quad \cdots\cdots (5)$$

III. THE BIAS FUNCTION AND QUANTIZER DESIGN

As a measure of the stability, we define the adaptation bias function [6].

$$\beta(y) = \sum_{i=1}^{N} p_i(y) \log M(i) \quad \cdots\cdots (6)$$

where $p_i(y)$ represents the probability of output magnitude level i and hence the probability of using multiplier $M(i)$ in the steady-state. It is assured that all sets of properly ordered $M(i)$ for which $\sum p_i(y) \log M(i) = 0$ provide the desired static performance [5].

Because $\beta(y)$ is monotonically decreasing from $\beta(0) = \log M(N) > 0$ to $\beta(\infty) = \log M(1) < 0$ and $\beta(y)$ is continuous, $\beta(y)$ has a unique zero crossing. The adaptation algorithm may be viewed as an iterative procedure for finding this zero crossing [6].

Clearly there is an infinite number of such sets. Hence we can say that the static range depends on the N-1 ratio, log M(i)/log M(N). The multiplier magnitudes determine adaptation speed [5].

IV. COMPUTER SIMULATIONS

Using the same method as Jayant's [4], some properties of adaptive quantizer are investigated.

1. Inputs and Performance Criterion

Our simulation have employed a first-order Gauss-Markov sequence as quantizer input. It is generated by the recursive rule

$$X_r = C X_{r-1} + \sqrt{1-C^2}\, N_r \; ; \; X_0 = 0, \quad \cdots\cdots (7)$$

where the samples N_r are drawn from a zero-mean, unit variance, white Gaussian sequence.

The quantizer output is the output level nearest to the input X_r

$$Y_r = \left((2\left[\frac{X_r}{\Delta}\right]+1)\frac{\Delta}{2} \right) sigX_r, \quad \frac{X_r}{\Delta} < 2^{B-1}$$

$$= \left((2^B-1)\frac{\Delta}{2}) \right) sig X_r, \quad \frac{X_r}{\Delta} \geq 2^{B-1}, \quad \cdots\cdots (8)$$

where [.] stands for "greatest integer in."

The quantization error $\quad E_r = Y_r - X_r. \quad \cdots\cdots (9)$

A conventional performance measure is the signal-to-quantization error ratio (SNR)

$$SNR(N, \Delta_{START}) = \frac{\sum_1^N X_r^2}{\sum_1^N E_r^2} \quad \cdots\cdots (10)$$

In adaptive quantization, a suitable multiplier function for a given signal should provide a compromise between quickness of response and steady-state performance. We define an average performance index

$$SNR_{AVE} = \frac{1}{20} \sum_N \sum_{\Delta_{START}} SNR(N, \Delta_{START}) \cdot\cdot (11)$$

for values of N=10, 100, 500, 1000, and

$$\Delta_{START} = \left[\frac{1}{10}, \frac{1}{\sqrt{10}}, 1, \sqrt{10}, 10 \right] \Delta_{OPT}.$$

where Δ_{OPT}=0.9957 for B=2, Δ_{OPT}=0.5860 for B=3 (1) [1].

2. Multiplier Functions for B=2, C=0

Table 1 illustrates the nature of the SNR function (10) for two multiplier functions in B=2.

<Table 1> SNR (dB) for B=2, C=0

20log $\left(\frac{\Delta_{START}}{\Delta_{OPT}}\right)$	Value of N			
	10	100	500	1000
	M(1)=0.90		M(2)=1.20	
-20	1.66	6.39	7.92	8.04
-10	3.23	7.46	7.55	8.00
0	13.03	9.31	8.23	8.36
10	4.55	7.87	8.36	8.45
20	-12.76	-0.72	4.90	6.16
	M(1)=0.98		M(2)=1.04	
-20	1.31	4.20	7.48	8.15
-10	3.99	6.84	8.61	8.91
0	8.45	7.64	8.67	8.87
10	2.34	6.32	8.21	8.39
20	-9.84	-4.42	1.06	3.68

The first multiplier function shows faster response (better SNR values for N=10 or 100), while the second function achieves a better asymptotic value of SNR (at N=1000). Obviously, the poor asymptotic performance of the first one is due to overly abrupt step-size oscillations in the steady-state, while the inferior performance of the second one for small N is due to sluggish adaptations of Δ when Δ_{START} is suboptimal [4].

<Table 2> Comparison of Multiplier Function (B=2, C=0)

M(1)	M(2)	SNR_{AVE}(dB)	$\prod_{i=1}^{2} M(i)^{P_i}$
0.71	2.00	5.36	0.999
0.80	1.60	5.77	1.006
0.90	1.20	5.80	0.990
0.95	1.10	5.51	0.997
0.98	1.04	4.94	0.999
0.95	1.20	4.61	1.026
0.50	2.00	4.47	0.790
0.90	1.10	5.28	0.962

Table 2 compares several multipliers for a 2-bit quantizer on the basis of (11). The first five functions satisfy $\sum_{i=1}^{2} p_i \log M(i)=0$, where $P_1=0.67$ and $P_2=0.33$ [1].

3. Multiplier Functions for B=3, C=0

Table 3 shows the nature of the SNR function (10) for B=3 and a specific multiplier function.

<Table 3> SNR(dB) for M(1)=0.90, M(2)=0.90, M(3)=1.25,(M)4=1.75 (B=3, C=0)

20log $\left(\frac{\Delta_{START}}{\Delta_{OPT}}\right)$	Value of N			
	10	100	500	1000
-20	11.2	12.2	12.9	12.4
-10	9.91	12.9	12.1	12.7
0	14.4	13.3	13.8	12.7
10	9.20	12.4	12.8	12.8
20	-4.93	6.12	10.2	11.4

Table 4 uses the performance criterion (11). The first three functions satisfy a stability constraint, where $p_1=0.47$, $p_2=0.30$, $p_3=0.14$, $p_4=0.09$ [1]. Notice that the reduction of the number of distinct step-size multipliers leads to a marginal decrease of SNR_{AVE} [4].

<Table 4> Comparison of Multiplier Function (B=3, C=0)

M(1)	M(2)	M(3)	M(4)	SNR_{AVE}(dB)	$\prod_{i=1}^{4} M(i)^{p_i}$
0.90	0.90	1.25	1.75	11.02	1.0005
0.90	1.00	1.00	1.75	11.01	1.0008
0.50	1.00	1.00	2.00	8.27	0.7684
0.30	0.90	1.50	2.10	5.98	0.6225

4. Comparison of Adaptive and Nonadaptive Quantizers

Table 5 summarizes the nature of optimal multiplier functions for B=2 and B=3. These functions are obtained from [11].

Although the quantizer problem for C=0.5 is qualitatively similar to that for C=0.99, we note that results for C=0.5 are nearly identical with those for C=0. The SNR gain resulting from adaptation is seen to be negative. The reason for using an adaptive quantizer in these situations is only to facilitate quantizations with much less knowledge of the input. In other words, step-size adaptations increase the dynamic range of the quantizer and enable it to handle to inputs with large amplitude variations, such as nonstationary signals.

<Table 5>Quantization of Gauss-Markov Inputs

B	C	0.00	0.50	0.99
2	SNR_{NA}	9.30	9.30	9.30
	SNR_A	5.80	6.23	6.87
	M(1)	0.90	0.90	0.50
	M(2)	1.20	1.20	2.00
3	SNR_{NA}	14.62	14.62	14.62
	SNR_A	11.02	10.52	12.95
	M(1)	0.90	0.90	0.30
	M(2)	0.90	0.90	0.90
	M(3)	1.25	1.25	1.50
	M(4)	1.75	1.75	2.10

* SNR_{NA} stands for SNR of the nonadaptive optimun quantizer [11], [12].

** SNR_A stands for SNR of an adaptive quantizer.

Notice that step-size increases are always faster than step-size decrease. The need for fast increases of step-size and slow decrease may be physically explained as follows [4]. Quantization errors during overload tend to be more harmful than those during granulity, since the magnitude of granular error is restricted to half step-size, while no such constraint exist for an overload error. It is therefore reasonable to decrease step-size slowly to avoid unduly small step-sized leading to the harmful overload errors.

V. CONCLUDING REMARKS

We considered a quantizer which adapts its step-size by a factor depending only on the knowledge of the previous quantizer output level. Adaptation rule and design function are described. By computer simulations, we know that M_{OPT} has the property of calling for fast increases and slow decreases of step-size.

REFERENCES

[1] Jayant, N.S. and P. Noll, Digital Coding of Waveform: Practices and Applications to Speech and Video, Englewood Cliffs, NJ: Prentice-Hall, 1984.

[2] Golding, L.S. and P.M. Schulthesis, "Study of an Adaptive Quantizer," Proc. IEEE, Vol. 55, pp. 293-297, Mar. 1967.

[3] Cummiskey, P. N.S. Jayant, and J.L. Flanagan, "Adaptive Quantization in Differential PCM Coding of Speech," Bell Syst. Tech. J., Vol. 52, pp. 1105-1118, Sept. 1973.

[4] Jayant, N.S., "Adaptive Quantization with a One-Word Memory," Bell Syst. Tech. J., Vol. 52, pp. 1119-1144, Sept. 1973.

[5] Goodman, D.J. and A. Gersho, "Theory of Adaptive quantizer," IEEE Trans. on Commun., Vol. COM-22, pp. 1037-1045, Aug. 1974.

[6] Gersho, A. and D.J. Goodman, "A Training Mode Adaptive Quantizer," IEEE Trans. on

[7] Goodman, D.J. and R.M. Wilkinson, "A Robust Adaptive Quantizer," IEEE Trans. on Commun., pp. 1362-1365, Nov. 1975.

[8] Cohn, D.L. and J.L. Melsa, "The Relationship Between an Adaptive Quantizer and a Variance Estimator," IEEE Trans. on Inform. Theory, pp. 669-671, Nov. 1975.

[9] Noll, P. "A Comparative Study of Various Quantization Schemes for Speech Encoding," Bell Syst. Tech. J., pp. 1597-1614, Nov. 1975.

[10] Mitra, D. "Mathematical Analysis of Adaptive Quantizer," Bell Syst. Tech. J., pp. 867-898, May-June 1974.

[11] Max, J. "Quantizing for Minimum Distortion," IRE Trans., Vol. PGIT-6, pp. 7-12, Mar. 1960.

[12] Lloyd, S.P. "Least Squares Quantization in PCM," IEEE Trans. on Inform. Theory, pp. 129-136, Mar. 1982.