

- [16] H. Murakami, Y. Hatori, and H. Yamamoto, "Comparison between DPCM and Hadamard transform coding efficiencies for the variation of statistical characteristics of NTSC signals" (in Japanese), Paper Tech. Group, CS 78-53, IECE Japan, June 1978.
- [17] N. Ahmed and K. R. Rao, *Orthogonal Transforms for Digital Signal Processing*. New York: Springer-Verlag, 1975.
- [18] W. Mauersberger, "Experimental results on the performance of mismatched quantizers," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 381-386, July 1979.
- [19] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inform. Theory*, vol. IT-6, pp. 7-12, Mar. 1960.
- [20] P. A. Wintz and A. J. Kuternback, "Waveform error control in PCM telemetry," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 650-661, Sept. 1968.
- [21] R. C. Reininger and J. D. Gibson, "Distribution of the two-dimensional DCT coefficients for image," *IEEE Trans. Commun.*, vol. COM-31, June 1983.

## Hybrid Adaptive Quantization for Speech Coding

M. E. M. NASR AND C. V. CHAKRAVARTHY

**Abstract**—This correspondence describes a new quantization technique called hybrid adaptive quantization (HAQ) that uses instantaneous [1], [2] as well as syllabic [3] adaptation of the step size. Two types of instantaneous adaptive algorithms have been used—Jayant's adaptive quantizer (JAQ) and the incremental adaptive quantizer (IAQ). Computer simulations have been performed for a sine-wave, correlated Gaussian signal and digitized speech. Signal-to-noise ratio (SNR) computation for PCM and DPCM coders indicates that the hybrid technique is superior to the normal adaptive quantizer, when both have the same ratio of maximum to minimum step size.

### I. INTRODUCTION

Waveform encoding methods based on pulse code modulation (PCM) or differential PCM (DPCM) may employ linear, nonlinear, or adaptive quantizers [1]–[3]. Adaptive delta modulation (ADM) is a special case of a DPCM coder. Adaptive DPCM coders, which use only half the bandwidth needed by conventional PCM coders, are being standardized for voice communication. While the existing  $\mu$ -law and A-law coders have fixed nonuniform quantizers, ADPCM coders are based upon adaptive quantizers.

Adaptive quantization is of two types—instantaneous and syllabic. The former can change its step size for every sample, while in the latter it is much slower (following the envelope variation in speech). Although instantaneous quantizers have good attack and decay times, they are prone to a little more

quantizing noise, unlike the slower syllabic adaptive systems. Un *et al.* [4] have successfully employed a combination of both methods for ADM while we have investigated the hybrid technique for PCM and DPCM coders. The main idea is as follows.

An instantaneous adaptive quantizer changes its step size at each sampling instant, based upon a certain rule, subject to a minimum step size  $\Delta_0$  and a maximum step size  $\Delta_{\max}$ , resulting in a range factor  $R = \Delta_{\max}/\Delta_0$ . In an HAQ these limits are not fixed values, but vary proportionally to the envelope of the input. If the envelope is considered to be constant over short periods of time, the values of  $\Delta_0$  and  $\Delta_{\max}$  within these intervals will be constant (but different from interval to interval).

Another method that is somewhat similar to the HAQ (called the "pseudosyllabic" adaptation) has been reported by Raulin<sup>1</sup> *et al.* [6], [7]. In the simplest case the quantizer can be a linear one with a minimum step size of  $\Delta_0$ . Whenever the magnitude  $|B_n|$  of the binary output exceeds a predetermined value, a constant value of 256 is fed to a low-pass filter with transfer function  $H(z) = (1 - (511/512)z^{-1})^{-1}$ , and a zero is fed otherwise. The effect is to increase the output (denoted by  $S(n)$ , say) by 256 when the threshold is exceeded, or to reduce it by a factor of 511/512 otherwise. The step size  $\Delta(n)$  is given by  $\Delta(n) = \Delta_0 2^{i(n)}$  where  $i(n)$  is defined as the largest integer in  $[(S(n)/2048) - 4]$ . This allows the step size to vary over the range of  $\Delta_0$  to  $2048 \Delta_0$ . It is easy to see that the step size can double in eight samples (1 ms for a sampling rate of 8 kHz), while the step size reduces by half in a time that varies between 4 and 16 ms, depending upon the initial step size.

As an example, a 4 bit quantizer can have 15 values of  $B_n$  between  $-7$  and  $+7$  (since the implementation reported excludes the all-zero code). If the threshold is selected as 5, the filter output is increased by 256 every time  $|B_n| > 5$ . However, when the step size is doubled, even though the threshold is 5, the absolute value represented by this threshold is doubled. To make the transition smoother, four nonlinear laws have been used, each with different thresholds. It must be noted, however, that during the time it takes  $i(n)$  to change by one unit (up or down), the quantizer will be the same.

While it seems to be similar to the hybrid scheme, the pseudosyllabic system is different for the following reasons. During the time of eight samples required for step increase or during the time required for step size decrease, the quantizer is fixed. On the other hand, in the hybrid scheme, even during this time the step size can change based on the instantaneous algorithm. In fact, if JAQ is used, the step size can double in just one or two samples. This is particularly efficient in avoiding clipping in PCM or slope overload in DPCM. Likewise, the time taken to halve the step size would be longer in the pseudosyllabic system, and during this time we have the fixed step size. Of the different instantaneous adaptive algorithms available, we have chosen two for our study—Jayant's adaptive quantizer (JAQ) [1] and the incremental adaptive quantizer (IAQ) [2]. PCM and DPCM (single tap predictor) coders have been simulated on a DEC-10 computer using sinusoidal, correlated Gaussian random signal, and digitized speech inputs.

In the case of the IAQ, the hybrid quantizer offers a substantial increase in dynamic range, while in the case of JAQ, there is not any significant increase. The reasons for and the implications of this are presented in the concluding section.

Paper approved by the Editor for Signal Processing and Communication Electronics of the IEEE Communications Society for publication without oral presentation. Manuscript received April 20, 1983; revised May 4, 1983. This work was supported in part by a fellowship from the Government of India and the Government of Egypt.

M. E. M. Nasr is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India, on leave from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt.

C. V. Chakravarty is with the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803, on leave from the Department of Electrical and Computer Engineering, Indian Institute of Science, Bangalore 560012, India.

<sup>1</sup>The authors thank an anonymous referee for bringing these references to their notice and for providing a copy of the CCITT Study Group Report.

## II. QUANTIZER DESCRIPTION

One of the forms of JAQ is shown in Fig. 1. The step size  $\Delta_n$  of the quantizer (which need not necessarily be uniform) at the  $n$ th sampling time is

$$\Delta_n = \Delta_{n-1} \cdot M(|B_{n-1}|), \quad (1)$$

$M$  being a multiplier dependent on  $|B_{n-1}|$ . The HAQ is realized using the system shown in Fig. 2. The quantizer output is low-pass filtered to extract the envelope, which is used to multiply the minimum step size  $\Delta_0$ . In the simulation of JAQ [1], the value of  $\Delta_0$  used is  $0.1 \Delta_{opt}$  where  $\Delta_{opt}$  is the optimum step size for a nonadaptive quantizer given by the rule of thumb  $\Delta_{opt} = 8\sigma/2^N$  ( $\sigma^2$  is the variance and  $N$  the number of bits).

Starting with a step size of  $\Delta_0$ , if the envelope is constant for  $t_1$  seconds, then at  $t = t_1$ , the step size will be

$$\Delta(t_1) = \Delta_0 \left\{ \sum_{i=1}^l M(|B_i|) \right\} \quad (2)$$

such that  $lT_s = t_1$  ( $T_s$  being the sampling period). During the next sampling time, if the envelope changes by a factor  $P$ , then the step size for the next sample is  $P\Delta(t_1)$  and not  $\Delta(t_1)$ . This assumes an instantaneous response. In practice, however, the envelope change will take place in a finite time (albeit small) and during the transitions the step size would not have stabilized to the optimum value.

The syllabic adaptation is achieved as follows. The output of the quantizer (analog samples) is rectified and filtered to produce the envelope which then multiplies the minimum step size. The filter is a two-pole LP filter with a 20 ms time constant.

The adaptation rule for the incremental adaptive quantizer [3] is

$$\Delta_n = \begin{cases} \Delta_{n-1} - \Delta_0(n) & \text{if } |B_{n-1}| < (2^N - 1)/4 \\ \Delta_{n-1} + \Delta_0(n) & \text{if } |B_{n-1}| > (2^N - 1)/4 \\ \Delta_{n-1} & \text{otherwise.} \end{cases} \quad (3)$$

$\Delta_0(n)$  is the minimum step size at the  $n$ th sampling instant ( $=\Delta_0$  if it is not a hybrid quantizer, and  $=\Delta_0 \times \text{envelope of the input at the } (n-1)\text{th instant}$  if it is a hybrid quantizer).

The quantizer schematic was shown in Fig. 2, and Fig. 4 shows the hardware realization. Both types of the HAQ have been used for PCM and the familiar single tap predictor ADPCM (Fig. 3). The multipliers used for the JAQ in PCM and DPCM coders are different, as discussed by Jayant [1]. The IAQ is shown in Fig. 5.

## III. SIMULATION AND RESULTS

Figs. 6-10 highlight the performance of the hybrid quantizer. Figs. 6 and 7 provide the results for PCM coders with sinusoidal and random inputs, respectively, while Figs. 8-10 are for DPCM coders, the inputs being, respectively, a 1020 Hz sinewave, random noise band limited to within 300 and 3400 Hz, and speech. The digitized speech has a sampling rate of 10 kHz. For speech signals, the segment SNR [8] is computed. All the values of SNR are computed over 10 000 samples. In the normal adaptive quantizer, the ratio of the maximum to minimum step size is limited to 60 dB. In the hybrid quantizer the ratio remains the same but with a difference. The ratio of the largest and smallest steps at any instant is fixed at 100 for the *instantaneous adaptation* only. Therefore, whenever the envelope is constant, this ratio is 100 but the absolute values of the largest and smallest steps will vary.

There is not much difference between the JAQ and its

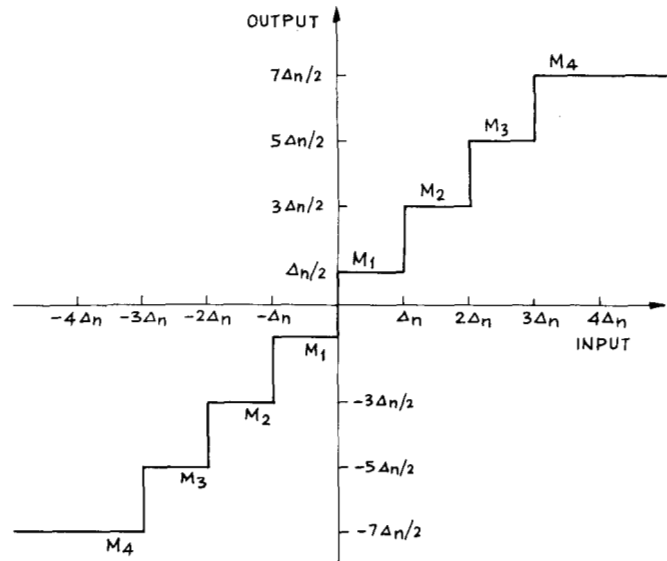


Fig. 1. Jayant's adaptive quantizer.

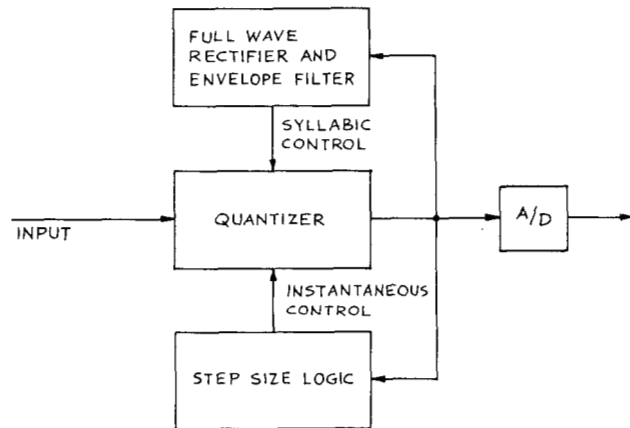


Fig. 2. Schematic of a hybrid adaptive quantizer.

hybrid version. The reason for this is that the adaptation algorithm is ideally suited for very quick changes in the step size, even when the starting value is grossly suboptimal. This is not the case with the IAQ. If the step increment is too small, the step size change is too slow. So the SNR degrades. On the other hand, if the increment is made proportional to the envelope (hybrid version) the performance is excellent. While an ordinary version of JAQ seems enough, the hybrid version has certain advantages. While the computer simulation provides a continuum of step sizes, a hardware implementation has to restrict the values to a finite set. This disadvantage can be offset to some extent with a hybrid coder.

## IV. CONCLUSIONS

A new method of realizing an adaptive quantizer, called the hybrid adaptive quantizer, has been described. The step size adaptation is both instantaneous as well as syllabic—the first part controlled by the digital output and the second part by the envelope of the quantizer output. Simulation results for deterministic, random, and speech signals have been obtained in the case of 3 bit and 4 bit PCM and DPCM quantizers. The instantaneous adaptation algorithms used are Jayant's adaptive quantizer (JAQ) and the incremental adaptive quantizer (IAQ). Although among them there is not much difference in performance, the hybrid versions can offer a large dynamic range and should prove to be especially useful for speech coding.

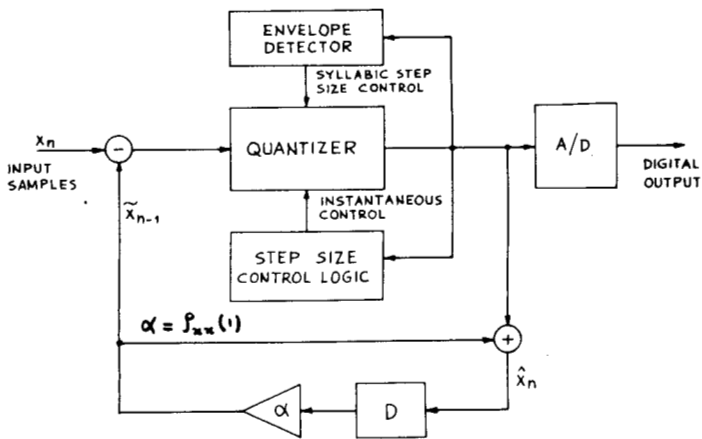


Fig. 3. DPCM coder with a single tap predictor.

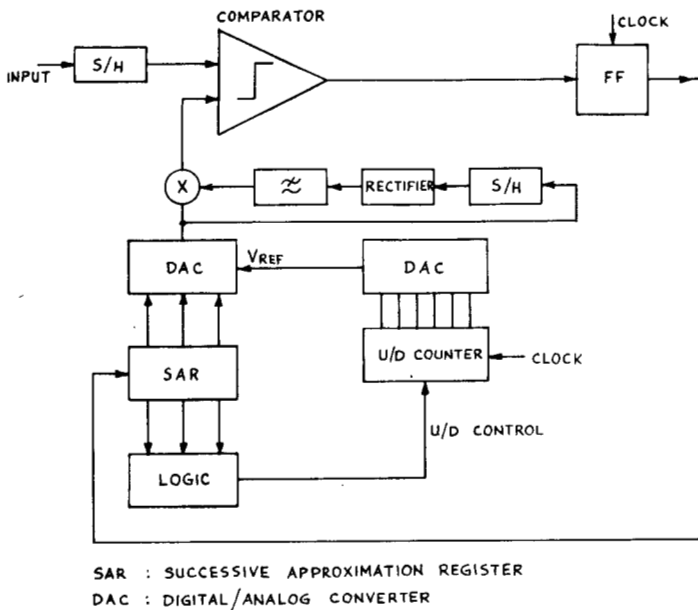


Fig. 4. Suggested hardware realization scheme for the HAQ (incremental version).

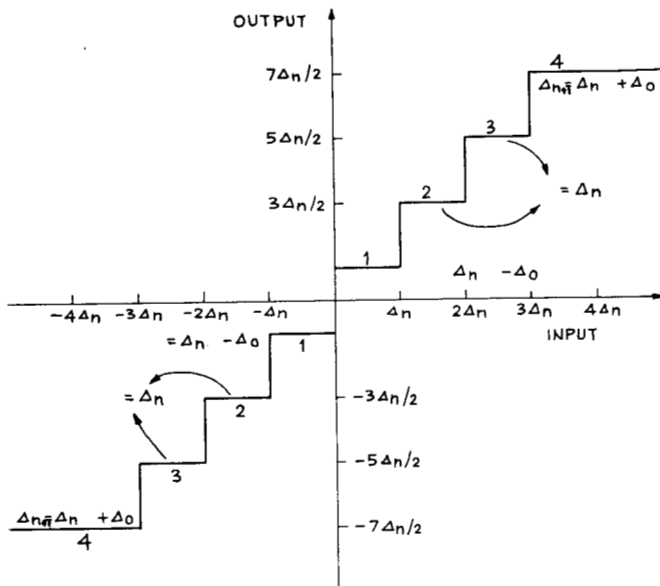


Fig. 5. A 3 bit instantaneous adaptive quantizer.

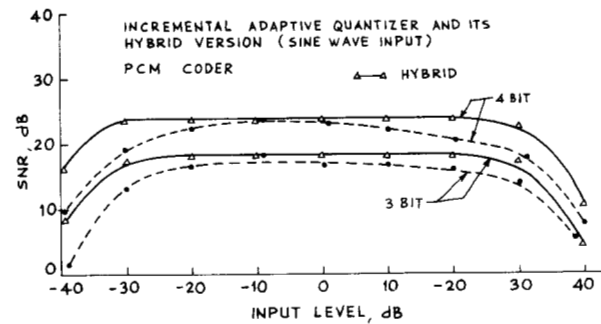
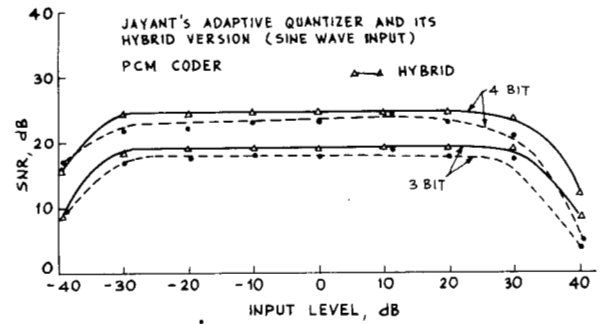


Fig. 6. Comparison of SNR versus input for sinusoidal signal (PCM quantizer).

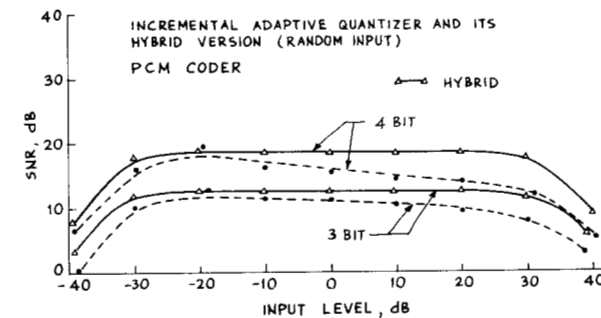
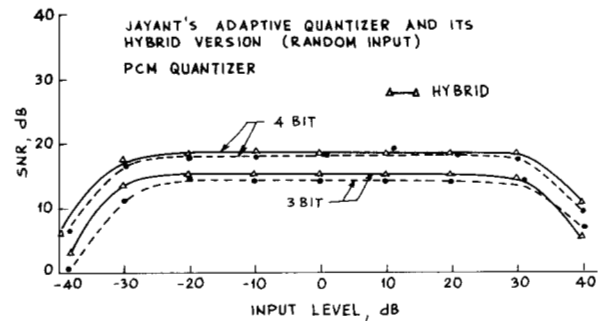


Fig. 7. Comparison of SNR versus input for a Gaussian input (PCM quantizer).

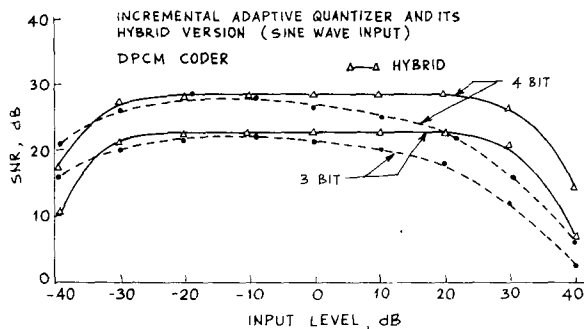
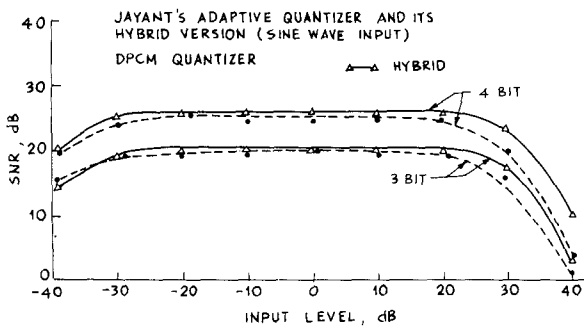


Fig. 8. Comparison of SNR variation with input level (sine-wave, DPCM quantizer).

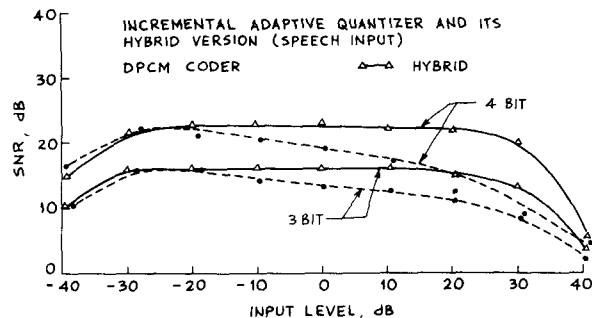
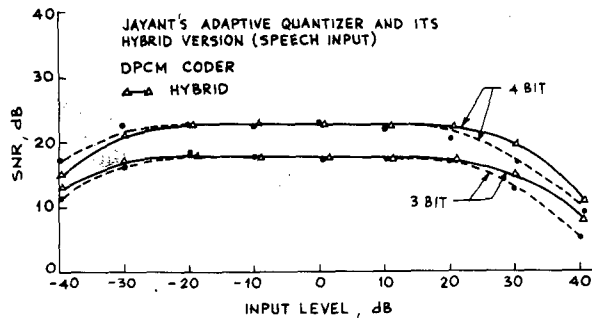


Fig. 10. SNR variation for speech signals (DPCM quantizer).

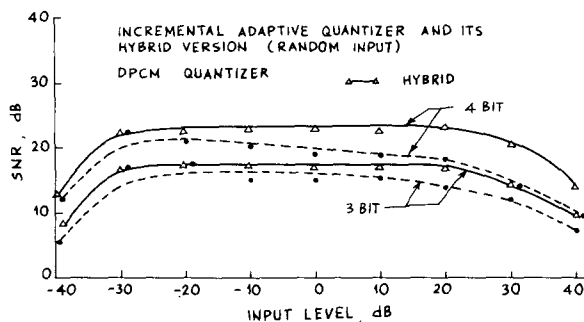
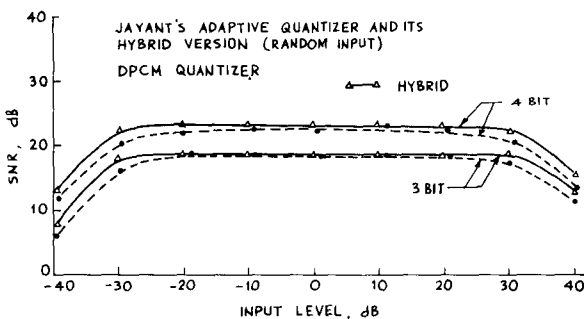


Fig. 9. Comparison of SNR variation with input level (random input, DPCM quantizer).

#### ACKNOWLEDGMENT

The authors are grateful to the Computer Centre, I.I.Sc., for the facilities provided and to Ms. Bharati Devi of the School of Automation for providing the digitized speech samples. The first author thanks the Government of India and Government of Egypt for providing a fellowship. The authors thank the reviewers for their suggestions.

#### REFERENCES

- [1] N. S. Jayant, "Adaptive quantization with one work memory," *Bell Syst. Tech. J.*, vol. 52, pp. 1119-1144, Sept. 1973.
- [2] C. V. Chakravorthy, N. D. Georganas, and S. G. S. Shiva, "An incremental adaptive quantizer," *IEEE Trans. Commun.*, vol. COM-29, pp. 1056-1061, July 1981.
- [3] M. S. Sodhi, J. Das, and C. V. Chakravorthy, "An adaptive PCM with syllabic step size adaptation," in *Conf. Rec. Nat. Telecommun. Conf.*, Washington, DC, 1979, pp. 4.3.1-4.3.6.
- [4] C. K. Un, H. S. Lee, and J. S. Song, "Hybrid companded delta-modulation," *IEEE Trans. Commun.*, vol. COM-29, pp. 1337-1343, Sept. 1981.
- [5] J. L. Flanagan, M. R. Schroeder, B. S. Atal, N. S. Jayant, R. Crochiere, and J. M. Tribolet, "Speech coding," *IEEE Trans. Commun.*, vol. COM-27, pp. 711-742, Apr. 1979.
- [6] J. M. Raulin, G. Bonnerot, J. Jeandot, and R. Lacroix, "A 60 channel PCM-ADPCM converter," *IEEE Trans. Commun.*, vol. COM-30, pp. 567-573, Apr. 1982.
- [7] CCITT Contrib. Study Group XVIII, no. 35, "32 kb/s ADPCM codecs," Mar. 1981.
- [8] P. Noll, "Adaptive quantization in speech coding systems," in *Proc. Int. Zurich Sem. Digital Commun.*, Zurich, Switzerland, 1974, pp. B3(1)-B3(6).