論文 84-21-1-11

An ADPCM System with Improved Error Control

(개선된 전송오차 제어기능을 가진 ADPCM 시스템에 관한 연구)

金 熙 東*, 殷 鍾 官**

(Hee Dong Kim and Chong Kwan Un)

要 約

본 논문에서는 ADPCM 시스템의 noisy channel에서의 성능 개선을 위한 새로운 방법을 제시하였다. 이 방법은 robust quantizer를 사용하면서 주기적으로 maximum step size를 수신측에 보내준다. 또한 수신측 버퍼에서는 MSB 에러검출·수정을 행한다. Noisy channel 상태에서 실제의 음성에 대해 컴퓨터 시뮬레이션한 결과 제안된 시스템의 성능은 원래의 ADPCM의 성능보다 크게 향상되었다.

Abstract

In this paper a new method of improving the performance of ADPCM in noisy channel is proposed. The proposed method employs a robust quantizer, and transmits the information regarding the maximum step size periodically. Also, a scheme to correct most significant bit (MSB) errors is used in the receiver buffer. According to our computer simulation with real speech, the proposed ADPCM with error control yields an improvement of about 4 to 5 dB in noisy channel over the conventional ADPCM without error control.

I. Introduction

Adaptive differential pulse code modulation (ADPCM) is known to be an efficient method of speech coding at medium rates of 16 to 48 kbits/s[1]. Presently, CCITT is considering ADPCM as a candidate for a 32 kbits/s coder that can be used along with 64 kbits/s PCM systems.

It is well known that the performance of ADPCM is extremely sensitive to channel bit errors, since the transmission errors give degrading effects not only on the quantizer level but also on the step size. Moreover, such errors in differential coders propagate over many sample periods. The step size offset causes a damaging effect on the system performance more seriously than the quantizer level offset. Goodman and Wilkinson have developed a robust quantizer in which step size offset is decreased exponentially with sampling time[2]. The introduction of a leak factor in the robust quantizer, however, reduces the dynamic range. Jayant proposed a differential pulse code modulation with adaptive quantiza-

*非會員, **正會員, 韓國科學技術院 電氣 및 電子工學科

(Dept. of Electrical Eng., KAIST)

接受日字 : 1983年 10月 18日

tion feedforward (DPCM-AQF) for digital transmission of speech through noisy channel[3]. In this system, since the step size remains constant within a block, the effect of transmission error is less severe than in a conventional ADPCM system. A disadvantage of the DPCM-AQF is that it yields smaller signal-to-noise ratio (SNR) than ADPCM in ideal channel. To improve the performance in ideal channel, Evci et al. employed a sequential gradient estimation predictor in their DPCM-AQF system[4]. However, this predictor requires a large number of computations compared with other predictors.

In this work, we study another approach of improving the noisy channel performance of ADPCM. Basically, the proposed system is similar to DPCM-AQF. It employs a robust quantizer and transmits the maximum step size in a block. As a result, the step size mistracking due to channel errors is significantly reduced. Also, it utilizes a preprocessing unit at the receiver by which MSB errors are detected and corrected.

In what follows, we first describe details of the ADPCM system with improved error control including its hardware implementation. Next, we present computer simulation results for the performance of the proposed system at various conditions. Also, we compare its performance with those of other systems. Finally, we make a conclusion.

II. ADPCM System with Improved Error Control

The step size transmitting ADPCM system proposed in this work is shown in Fig. 1. The transmitter consists of a maximum step size searching unit, a conventional adaptive encoder and a multiplexer. And the receiver consists of a buffer to store the input code words, a demultiplexer and a preprocessing unit. Without the maximum stepsize searching unit and the preprocessing unit, the system becomes the same as the conventional ADPCM of Cummisky et al.[5].

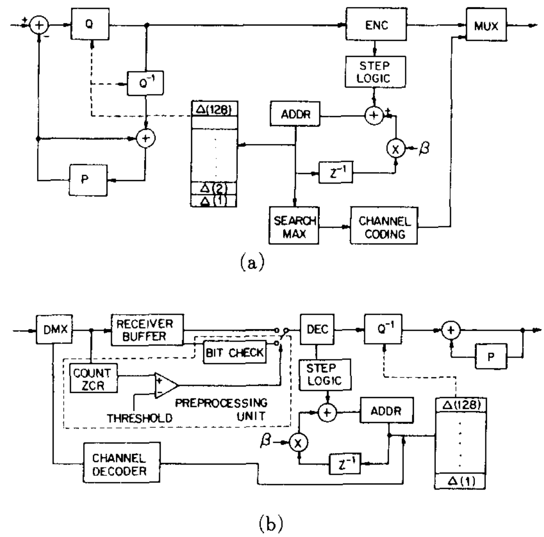

Fig. 1. Maximum step-size transmilting ADPCM system (a) Encoder (b) Decoter.

In our system we use the robust adaptive quantization scheme of Goodman and Wilkinson[2]. The quantizer step size is controlled by

$$\Delta_{k+1} = \Delta_k^{\beta} \cdot M(I_k). \tag{1}$$

That is, the new quantizer step size Δ_{k+1} is a function of the previous step size Δ_k multiplied by the pre-assigned multiplier M (I_k). The parameter β controls the rate at which the effect of transmission errors is dissipated. The mutiplier $M(I_k)$ is chosen so that $M(I_k) \geq 1$ when the code word is in the outer quantization region and $M(I_k) \leq 1$ when it is in the inner region.

To implement the adaptive quantizer digitally, we employ a table look-up method as implemented previously by Boddie et al.[6]. The possible 128 step sizes are stored in consecutive read only memory (ROM) locations in increasing order. Instead of calculating the step size given by (1) at each sampling instant, the address of ROM which indicates the memory location of corresponding step size is changed. Accordingly, a suitable mapping function is required to relate the ROM address with its contents. In this manner, the calculation required for determination of step size is simplified. The mapping function used in our system is derived in Appendix.

In the encoding process, the maximum step size searching unit finds the largest ROM address within a block, and transmits this address to the receiver as a side information after multiplexing with W-sample code words. In this case, some channel coding is required to prevent the side information from corrupting due to channel errors.

At the receiver, the buffer stores the code words. On receiving the side information, the decoding process starts, in which the quantizer step size is varied within the limits of maximum and minimum values.

The preprocessing unit detects and corrects MSB errors. It performs its operation in a time interval corresponding to the queueing delay time in the receiver buffer. It is noted that the MSB error does not change the step size multiplier because the code words are assigned symmetrically.

The algorithm used in the preprocessing unit has been made based on the following observation. There is a correlation between input speech and residual signal in voiced speech when a fixed predictor is used. It is well known that the voiced signal has low zero-crossing rate (ZCR). So does the residual signal of voiced speech. Since the MSB has the sign information of the residual signal, MSB alternation occurs infrequently in voiced speech.

The operation of the preprocessing unit is as follows. Before the code words enter the buffer, the bit alternation rate of MSB's is measured. With a suitable threshold value, the decision whether to preprocess or not is made. If the bit alternation rate of MSB's is lower than the given threshold, the code words in this block are assumed to represent the voiced signal and the preprocessing unit is activated. In this case, the threshold value THRS is, of course, a function of the block size W. To set the threshold value properly, we tested with speech samples and obtained an average value of bit alternation rate of MSB's for voiced speech. In this way, we have obtained a relation between THRS and W as

$$\mathrm{THRS} = 4. \times \frac{W}{32}. \tag{2}$$

This relation has been used in our simulation study. In the preprocessing unit, five consecutive MSB's are taken in a sliding manner. If the center MSB has different value from the other four, this MSB should be inverted. Its Boolean function is given by

$$M'_k = \left[(M_{k-2} \cdot M_{k-1} \cdot M_{k+1} \cdot M_{k+2} + \bar{M}_{k-2} \cdot \bar{M}_{k+1} \cdot \bar{M}_{k+1} \cdot \bar{M}_{k+2}) \cdot E \right] \oplus M_k, \quad k=3,4,\ldots,W-2, \tag{3}$$

where M_k is the MSB to be tested, M'_k is the corrected M_k, and E is the enable signal activated at voiced speech. Equation (3) can easily be implemented by standard logic gates.

With this simple procedure, most of MSB errors occurring in voiced speech can be corrected. Also, the transmitted maximum step size plays an important role in synchronizing the step size between encoder and decoder and in stabilizing the decoder output. With these schemes a remarkable performance improvement in noisy channel has been obtained.

III. Computer Simulation Results

The proposed ADPCM system with various parameter values was simulated on a computer, and the performances in ideal and noisy channels have been investigated. Also, we compared the performance of our ADPCM system with those of DPCM-AQF and conventional ADPCM without error control. The test utterances were five sentences spoken by a male and a female. They were low-pass filtered at 3.4 kHz, sampled at 8 kHz, and quantized with 12 bit resolution. The configuration used in this study is shown in Fig. 2. The low-pass filter (LPF) used here is an eight-pole Butterworth filter. To simulate the noisy channel with some specified bit error rate (BER), we changed a controlled number of

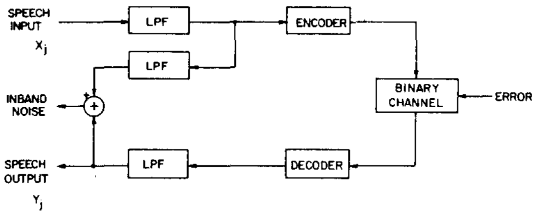

Fig. 2. Block diagram of simulation.

bits of the binary signal from 0 to 1 and vice versa.

To measure the performance of a coder, we have used the segmented signal-to-noise ratio (SNR_{SEG})[7]. In our simulation, we used 58,664 speech samples which are about 7.5 seconds long. In computation of SNR_{SEG}, we divided speech signal into segments of 128 samples or 16ms.

Prior to simulation, we optimized the ADPCM system parameters at 32 kbits/s. The optimal coefficient values used were a_1=0.91 for the first-order predictor, and a_1=1.2399 and a_2=−0.329 for the second-order predictor. The optimal step size was set to the minimum, and the initial predictor output was assumed to zero.

In order to make a fair comparison of performances of various coders, it is desirable to use the same input speech for each coder. In addition, it is important to have the input signal range, companding dynamic range and initial conditions identical among the systems being studied. In our simulation, the input signal was varied over the range of 80 dB. For quantizer step size variation, 60 dB companding was used.

1. *Performances in Ideal Channel*

The performance in SNR_{SEG} of conventional ADPCM system at 32 kbits/s was first obtained at different input signal level and with different step size leak factors in ideal channel. We used a first-order or a second-order fixed predictor in the system. The results are shown in Table 1. It is seen from this table that the performance of ADPCM with a second-order predictor is almost the same as that with a first-order predictor. Also, one can see that that the dynamic range of ADPCM tends to

Table 1. SNR_{SEC} conventional ADPCM with fixed predictor for different leak factors vs. input signal level (transmission rate: 32 Kbps).

ADPCM: No of Tap Coefficient	Leak Factor \ Input Signal Level (dB)	−60	−40	−20	0	20
1 Tap Predictor	$\beta = 1$	9.31	22.77	25;35	25.30	22.364
	$\beta = \frac{127}{128}$	9.3	22.43	25.39	25.35	22.53
	$\beta = \frac{63}{64}$	8.3	20.22	24.70	25.39	22.47
	$\beta = \frac{31}{32}$	3.0	6.03	15.53	24.25	22.48
2 Tap Predictor	$\beta = 1$	9.72	22.47	24.95	25.14	22.34
	$\beta = \frac{127}{128}$	9.73	22.34	24.87	24.94	22.27
	$\beta = \frac{64}{63}$	8.79	20.10	24.10	24.95	22.04
	$\beta = \frac{31}{32}$	−	4.42	15.73	23.91	22.38

be narrower as the leak factor decreases. But, at a low input signal level, the performance of ADPCM with the robust quantizer becomes slightly improved. The reason is believed to be due to the fact that the leak factor of the robust quantizer stabilizes the step size variation.

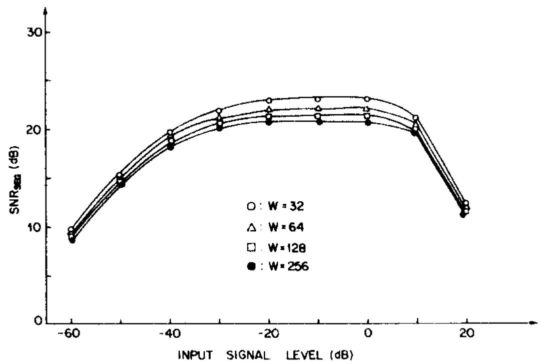

Fig. 3. SNR_{SEC} of DPCM-AQF vs. input signal level (transmission rate: 32 kbps).

Also, we obtained SNR_{SEG} of DPCM-AQF as a function of input signal level with various block sizes. The result is shown in Fig. 3. It is seen in the figure that as the block size decreases, the performance is improved due to the increased step size adaptation rate. Comparing the performance of DPCM-AQF with that of ADPCM, it has been found that ADPCM is superior to DPCM-AQF by 5 dB in SNR_{SEG}. Although the step size of DPCM-AQF is accurately determined from the input signal, it remains constant during the period of one block time (256 samples or 32 ms). Since speech signal is time-varying, this constant step size can degrade the system performance and make the dynamic range narrow.

Considering the performance of ADPCM with error control, we note that the step size adaptation sequence of the proposed ADPCM with error control both at the transmitter and at the receiver are identical in ideal channel. Thus, the maximum step size transmitted to the receiver does not give any effect on the step size adaptation logic. Accordingly, the performance of ADPCM with error control is the same as that of the conventional ADPCM in the error-free channel condition (See Table 1).

2. Performances in Noisy Channel

For simulation of a noisy channel, we introduced a controlled number of errors to the serial bit stream. We have made bit errors to occur randomly. Although we could consider two types of errors, burst errors and occasional single errors, we assumed in this study that the channel has only the later type of errors. The performance of the conventional ADPCM with various leak factors β at different error rates is shown in Fig. 4. As the bit error rate increases, the performance of ADPCM becomes extremely sensitive and degrades rapidly. We note that one bit error of ADPCM is equivalent to one word error. Consequently, ADPCM with $\beta=1$ produces significantly distorted speech at bit error rates beginning 10^{-4} and becomes unacceptable at the bit error rate of 10^{-3} and above. As the leak factor in the step size adaptation logic decreases, the vulnerability of ADPCM is significantly reduced. Considering the trade-off between the robustness and the dynamic range, we have chosen the leak factor as $\beta = 127/128$.

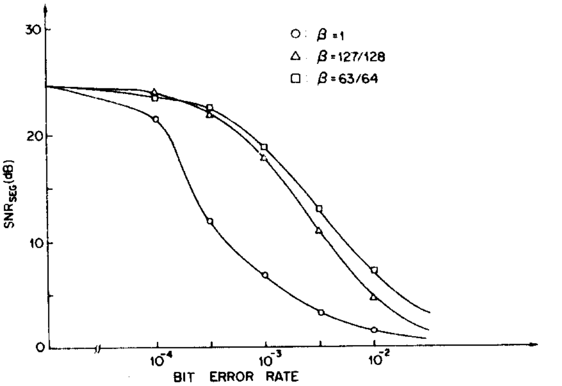

Fig. 4. SNR_{SEC} of ADPCM with one-tap fixed predictor vs. bit error rate (transmission rate: 32 kbps).

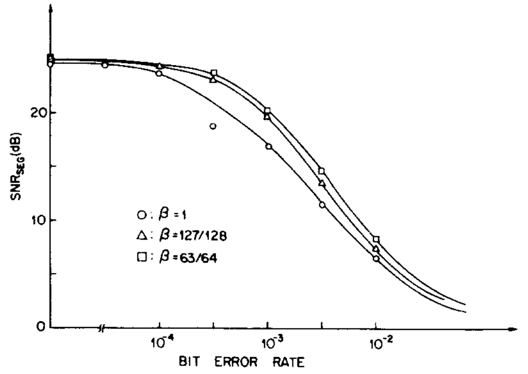

Fig. 5. SNR$_{SEG}$ of ADPCM with one tap fixed predictor and error control vs. bit error rate.

The performance of the proposed ADPCM system in the noisy channel is shown in Fig. 5. By transmitting the maximum step size within a block, thereby limiting the step size variation at the receiver, the performance is remarkably improved. In this case, we assume that the transmitted step size is not corrupted by channel errors. The performances of various ADPCM systems in noisy channel are shown in Fig. 6. Comparing the performance of the

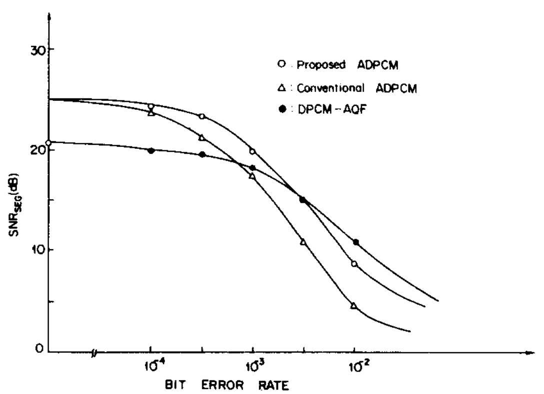

Fig. 6. SNR$_{SEG}$ of various coders vs. bit error rate.

proposed system with that of the conventional ADPCM, about 4 to 5 dB gain in SNR$_{SEG}$ can be obtained. Also, the performance of this system is superior to that of DPCM-AQF at the bit error rate lower than 3×10^{-3} but about the same at the rate above 3×10^{-3}.

To show the improvement resulting from

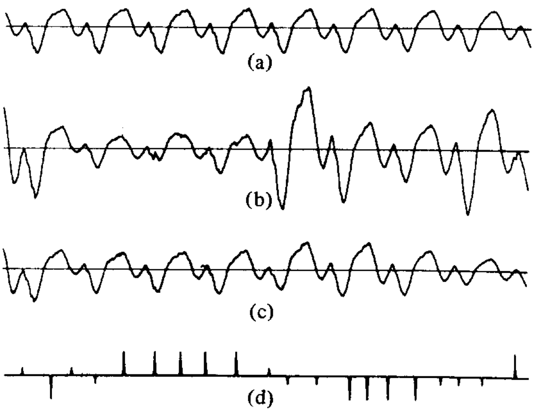

Fig. 7. Comparison of various waveforms of ADPCM in noisy channel (Bit error rate: 10^{-2}).
(a) orginal speech
(b) Error induced ADPCM output
(c) Error-controlled ADPCM output
(d) Error locations

error control, waveforms of ADPCM with and without error control are compared with original speech in Fig. 7. In this case, the bit error rate was 10^{-2} and the leak factor β was 127/128. Errors were generated at the locations where pulses occur as shown in (d) in the figure. The magnitude of each pulse indicates which bit is in error. That is, the largest value represents an MSB error, and the smallest value gives an LSB error. From this figure, one can see that the signal distortion due to channel errors can be reduced significantly by using the proposed error control scheme. The MSB correction improves the performance in such a way that the waveform of the error-controlled ADPCM is roughly the same as that of the original speech in voiced region. Perceptually, the MSB error gives the popping effect to the listener. By correcting MSB errors in voiced sound, the ADPCM system can yield fairly good quality of speech at 32 kbps.

IV. Conclusion

In this work, we have proposed a new

ADPCM that is robust in noisy channel condition. In this system, a robust quantizer is employed for instantaneously adaptive quantization. Also, a maximum step size in a given block is transmitted to the receiver. This side information controls the step size variation at the receiver. Further, a simple algorithm has been proposed to correct MSB errors in voiced speech. By using the proposed schemes, an improvement of about 4 to 5 dB in SNR_{SEG} could be obtained in noisy channel over the conventional ADPCM without error control.

Appendix

In this appendix, we derive a mapping function used in the quantizer adaptation logic. We have used ROM of 128 words, and its contents are arranged to have 60 dB companding. Let us denote the address of ROM and its content as ADR(i) and $\Delta(i)$, respectively. We set the maximum step size $\Delta_{max}=1$, and the minimum step size $\Delta_{min}=1/1024$. For convenience, let us consider the log transformed step size of (1),

$$d(i) \triangleq \log_Q \Delta(i), \tag{A-1}$$

and the log multiplier

$$m_k \triangleq \log_Q M(I_k). \tag{A-2}$$

Then, (1) is transformed to

$$d(i+1) = \beta \cdot d(i) + m_k. \tag{A-3}$$

By this equation, we can relate the address of ROM with a log transformed step size. Considering the dynamic range, we have

$$\frac{ADR(128)}{ADR(1)} = \frac{d(128)}{d(1)} = \log_Q \frac{\Delta_{max}}{\Delta_{min}} \tag{A-4}$$

and

$$Q=1.055645. \tag{A-5}$$

With this base Q, the step size can be obtained straightforwardly by

$$\Delta(i) = Q^{ADR(i)-128} \tag{A-6}$$

Thus,

$$\begin{aligned} &\Delta(128)=1 \\ &\Delta(127)=Q^{-1} \\ &\Delta(126)=Q^{-2} \\ &\quad\vdots \\ &\Delta(1) = Q^{-128} = \frac{1}{1024}. \end{aligned} \tag{A-7}$$

And the log multiplier is modified by (A-2). These values are listed in Table A-1. In this way, the address adaptation logic becomes

$$ARD(i+1) = [\beta \cdot ARD(i) + m \cdot (I_k)] \tag{A-8}$$

where [.] represents truncation.

Table A-1. Log transformed multiplier.

Symbol	Value	$\log_Q M_k$
M_1-M_4	0.9	−1.9
M_5	1.2	3.4
M_6	1.6	8.7
M_7	2.0	12.8
M_8	2.4	16.2

References

[1] N.S. Jayant, "Digital coding of speech waveforms; PCM, DPCM, and DM quantizers," *Proc. IEEE,* vol. 62, pp. 611-632, May. 1974.

[2] D.J. Goodman and R.M. Wilkinson, "A robust adaptive quantizer," *IEEE Trans. on Commun.,* vol. COM-23, Nov. 1975.

[3] N.S. Jayant, "Step-size transmitting differential coders for mobile telephony," *BSTJ,* vol. 54, pp.1557-1581, Nov. 1975.

[4] C.C. Evci, R. Steele and C.S. Xydeas, "DPCM-AQF using second order adaptive predictors for speech signals," *IEEE*

Trans. ASSP, Part 1. vol. ASSP-29, pp.337-341, June 1981.

[5] P. Cummiskey, N.S. Jayant, and J.L. Flanagan, "Adaptive quantization in differential PCM coding of speech," *BSTJ,* vol. 52, no. 7, pp.1105-1108, Sep. 1973.

[6] J.R. Boddie et al., "Digital signal processor: adaptive differential pulse code modulation coding," *BSTJ,* vol. 60, no. 7, pp. 1547-1561, Sep. 1981.

[7] D.J. Goodman et al. "Objective and subjective performance of tandem connections of waveform coders with a LPC vocoder," *BSTJ,* pp.601-630, Mar. 1979.

[8] N.S. Jayant, "Adaptive quantization with a one-word memory," *BSTJ,* vol. 52, no. 7, pp.1119-1144, Sep. 1973.