EXPERIMENTAL COMPARISON OF FORWARD AND BACKWARD ADAPTIVE PREDICTION IN ADPCM SPEECH CODING

A Thesis

by

LOUIS CLYDE SAUTER

Submitted to the Graduate College of
Texas A&M University
in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE

August 1979

Major Subject: Electrical Engineering

EXPERIMENTAL COMPARISON OF FORWARD

AND BACKWARD ADAPTIVE PREDICTION

IN ADPCM SPEECH CODING

A Thesis

by

LOUIS CLYDE SAUTER

Approved as to style and content by:

(Chairman of Committee)

(Member)

(Member)

(Head of Department)

August 1979

ABSTRACT

Experimental Comparison of Forward
and Backward Adaptive Prediction in
ADPCM Speech Coding. (August 1979)

Louis Clyde Sauter

Chairman of Advisory Committee: Dr. Jerry D. Gibson

In this paper, two different adaptive differential pulse code modulation (ADPCM) speech coders are designed and compared. One is a fourth order Kalman predictor with a (3/5) pitch compensated quantizer (PCQ). The other is an eighth order forward adaptive predictor based on the autocorrelation of the input, with an optimum 3 level quantizer. The block length is 17.5 ms. Both systems achieve a transmission rate of 16 kilobits/sec. (Kb/s), with appropriate source codes.

Comparison of the signal-to-noise ratios (SNR) and subjective comparisons of processed speech lead to the conclusion that the backward predictive system performs best. The average SNR was 16.1 dB for the Kalman predictor with PCQ.

The speech processed consisted of four sentences in different languages and for a single male speaker, filtered and sampled at 8000 Hz.

Some remarks on unstable systems consisting of forward predictors with coarse backward quantizers are made.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Many adaptive differential pulse code modulation (ADPCM) systems have been proposed and studied since about 1970 for speech coding. These systems are characterized by an adaptive quantizer and a fixed or adaptive predictor. They achieve bit-rates between 4 kilobits/sec. (Kb/s) and 40 Kb/s. Adaptation can be based on data available to both the transmitter and the receiver (backward adaptation) or on data available to the transmitter only (forward adaptation). The latter scheme then requires transmission of the adaptation parameters to the receiver.

The purpose of this study was to determine whether backward adaptive prediction performs better than forward adaptive prediction in speech coding. Both schemes have been widely studied but their performances have never been compared. Of course, it is very important in systems design to compare performances of various techniques. The knowledge thus acquired will enable the systems designer to choose the best technique.

Pioneer work in differential pulse code modulation (DPCM) speech coding was done by McDonald [1] in 1966. He dealt with fixed predictors and fixed quantizers, and showed DPCM to perform better than pulse code modulation (PCM).

This thesis follows rules of style consistent with those found in IEEE Transactions on Communications.

Atal and Schroeder [2] and Noll [3] described and used forward adaptive prediction. In [2], the authors described a forward adaptive predictor used with a two-level adaptive quantizer. The performances of this system were subjectively compared with log-PCM speech coding, and found to be equivalent to those of 6 bits/sample log-PCM for a transmission rate several times lower. In [3], Noll compared various schemes using fixed and forward adaptive prediction with fixed and adaptive quantizers. The criterion used for the comparison was the signal-to-noise ratio (SNR) and the bit-rates considered ranged from 16 Kb/s to 40 Kb/s. A maximum SNR was reached for a scheme including both adaptive quantization and adaptive prediction.

Backward adaptive prediction was first studied by Stroh [4] and Cummiskey [5] in Ph.D. dissertations, and by Gibson Jones and Melsa [6] and Gibson [7]. Various adaptation schemes were used (gradient, steepest descent, stochastic approximation and Kalman type algorithms). In [7], Gibson compared the performances of a fixed predictor ADPCM, an adaptive gradient predictor ADPCM and a Kalman type predictor ADPCM and found the Kalman type predictor to perform best.

Little work has been done to compare performances of backward and forward adaptive prediction. Gibson [8] analytically compared the two schemes under certain assumptions, and found backward adaptive prediction to perform better than forward adaptive prediction.

In this study, an experimental comparison of the two schemes was made by computer simulation at a bit-rate of 16 Kb/s. For each technique, the best possible system was searched for, under the constraint of a maximum bit-rate of 16 Kb/s. Source codes were designed when needed to achieve this bit-rate. The speech processed consisted of four sentences spoken in different languages by a single male speaker (see Appendix A). The performances of the systems were judged both numerically by SNR comparisons and subjectively by listening tests.

This thesis is organized as follows: in Chapter II, a brief description is given of ADPCM systems. The notation introduced in this chapter will be used throughout the thesis. In Chapter III, adaptive prediction, both forward and backward, is studied, with descriptions of the systems used for this research. Chapter IV contains a description of various adaptive quantizers, stressing those used in the study. In Chapter V, the simulation itself is described, with details on the data used and the systems chosen for the comparison. Results are given in Chapter VI. Finally, conclusions and remarks drawn from the study are presented in Chapter VII.

CHAPTER II

ADPCM SPEECH CODING

Figures 1 and 2 show block diagrams of the ADPCM system transmitter and receiver respectively. The quantizer is indicated by a Q and the predictor by a P.

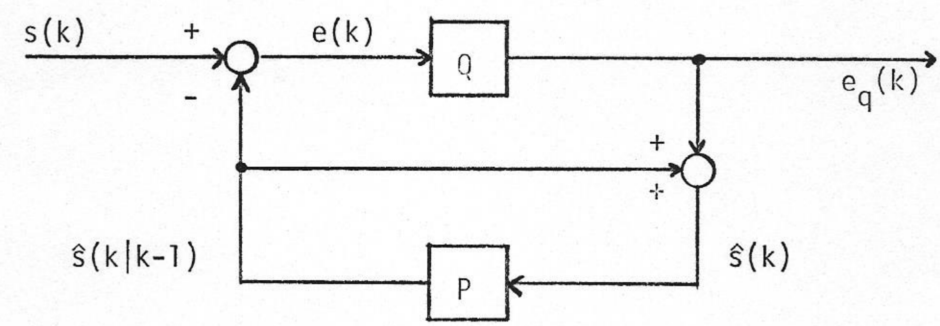

Fig. 1. ADPCM transmitter.

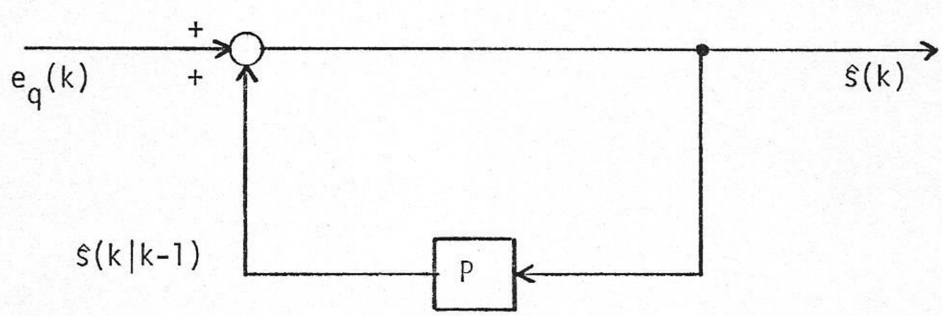

Fig. 2. ADPCM receiver.

The predicted value $\hat{s}(k|k-1)$ is usually a linear combination of the past input signal estimates:

$$\hat{s}(k|k-1) = \sum_{i=1}^{N} a_i \cdot \hat{s}(k-i) = \underline{a}_N^T \cdot \hat{S}_N(k-1) \tag{1}$$

where $\underline{a}_N^T = (a_1, a_2, \ldots a_N)$ is a vector of predictor coefficients, $\hat{S}_N(k-1) = (\hat{s}(k-1), \hat{s}(k-2), \ldots \hat{s}(k-N))$ and N is the order of the predictor.

The prediction error e(k) is given by

$$e(k) = s(k) - \hat{s}(k|k-1).$$

e(k) is quantized before transmission, yielding $e_q(k)$ and finally the estimate of the input sample (receiver output) is

$$\hat{s}(k) = \hat{s}(k|k-1) + e_q(k). \tag{2}$$

The difference

$$n_q(k) = e(k) - e_q(k) \tag{3}$$

is called quantization noise. Note that because of the coarse quantization used in this study, $n_q(k)$ is in general nonwhite and is highly correlated with e(k) and s(k).

Finally, Eqs. (1), (2) and (3) yield

$$\hat{s}(k) = s(k) + n_q(k).$$

The performances of the system can then be measured by the SNR:

$$SNR = \frac{\langle s^2(k) \rangle}{\langle n_q^2(k) \rangle} \quad \text{or} \quad SNR(dB) = 10 \log_{10} SNR$$

where $\langle \cdot \rangle$ denotes time averaging over the entire utterance.

System performances are also judged by subjective listening tests which may yield different results since the SNR is not necessarily a good indicator of how processed speech sounds to a listener.

CHAPTER III

ADAPTIVE PREDICTION

A. Forward Adaptive Prediction.

In this scheme, the predictor coefficients are calculated using a finite number of buffered input samples. They are then transmitted to the receiver. In [2], the predictor has a more complex structure than standard ADPCM since a long term predictor is included. This system is shown in Fig. 3.

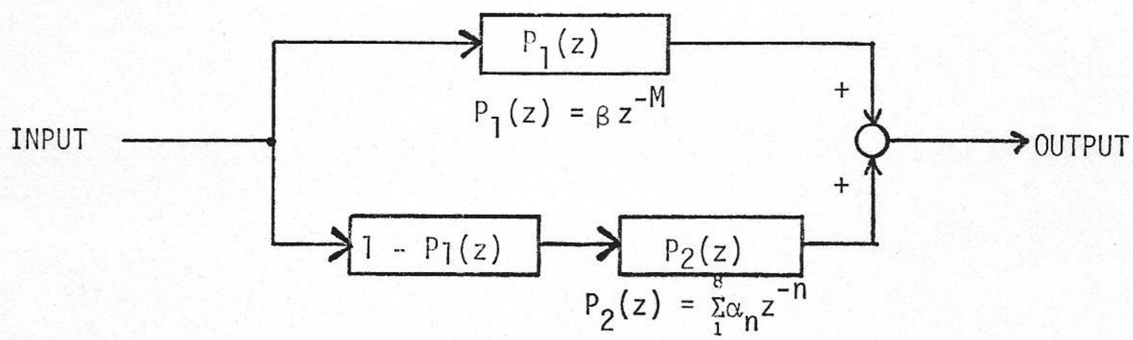

Fig. 3. Predictor in [2].

The coefficients β, M, α_1 through α_8 are determined so as to minimize the mean square error. This problem does not have a straight-forward solution and a sub-optimum solution was used. This solution was based on the short term autocorrelation of the input to calculate β and M, and on the short term autocorrelation of the output of $(1 - P_1(z))$ to calculate α_1 through α_8. The predictor was used with a two-level adaptive quantizer and the coefficients were updated every

5 msec. This system achieved performances equivalent to those of a 6 bits/sample log-PCM for a transmission rate several times lower (10 Kb/s instead of 40 Kb/s).

In [3], Noll uses a predictor based on the short term autocorrelation of the input. The optimum predictor coefficients are then given by

$$\underline{a}_N = R_N^{-1} \cdot \rho_N \tag{4}$$

where R_N and ρ_N are

$$R_N = \begin{bmatrix} R(0) & R(1) & R(2) & \cdots & R(N-1) \\ R(1) & R(0) & R(1) & \cdots & R(N-2) \\ R(2) & R(1) & R(0) & \cdots & R(N-3) \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ R(N-1) & R(N-2) & R(N-3) & \cdots & R(0) \end{bmatrix} \tag{5}$$

$$\rho_N = (R(1), R(2), \ldots R(N))^T. \tag{6}$$

The R(i) in Eqs. (5) and (6) are the short term autocorrelation coefficients given by

$$R(i) = \langle s(k) \cdot s(k+i) \rangle$$

where $\langle \cdot \rangle$ denotes the average over the buffered input. The length of the buffer is called the block length.

A proof of the optimality of $\underline{a}_N$ as given by Eq. (4) can be found in [9].

Note that R_N is a symmetric Toeplitz matrix (A Toeplitz matrix is a matrix in which all the entries along each diagonal are equal) and also that the entries of ρ_N repeat those of R_N. These properties enabled Durbin [10] to design a fast algorithm for solving Eq. (4). This algorithm is a recursive procedure and is specified as follows:

$$E(0) = R(0)$$

$$k_i = - \left[R(i) + \sum_{j=1}^{i-1} a_j^{(i-1)} R(i-j) \right] / E_{i-1}$$

$$a_i^{(i)} = k_i$$

$$a_j^{(i)} = a_j^{(i-1)} + k_i \, a_{i-j}^{(i-1)} \qquad j=1,2,\ldots i-1$$

$$E_i = (1 - k_i^2) \, E_{i-1}$$

The equations are solved recursively for i = 1, 2, ... N. Finally, $\underline{a}_N$ is given by

$$a_i = a_i^{(N)} \qquad i=1,2,\ldots N$$

As before, these coefficients must be transmitted to the receiver.

Noll used this predictor with various quantizers for different values of N. The bit-rates achieved ranged from 16 Kb/s to 40 Kb/s. The best system studied by Noll in [3] included both an adaptive quantizer and an adaptive predictor. It achieved a SNR of approximately 27 dB for 3 bits/sample encoding.

B. Backward Adaptive Prediction.

Stroh [4] and Cummiskey [5] wrote dissertations on this subject centered on gradient and stochastic approximation algorithms. These algorithms, however, did not produce a significant improvement [7, 11]. In [6], Gibson, Jones and Melsa used a Kalman type coefficient identification algorithm. The results were not very promising but in [7] Gibson compared fixed tap, adaptive gradient and Kalman type algorithms and found the latter to perform best. The SNR improvement attained 3 dB. The Kalman type algorithm is described as follows:

The predictor coefficients are updated by

$$\underline{a}_N(k+1) = \underline{a}_N(k) + K_{KF}(k+1)\ e_q(k+1) \tag{7a}$$

with zero initial conditions and where the Kalman gain matrix K_{KF} is given by

$$\begin{aligned} K_{KF}(k+1) = {} & V_{\tilde{a}}(k+1|k)\ \hat{S}_N(k) \\ & \cdot [\ \hat{S}_N^T(k)\ V_{\tilde{a}}(k+1|k)\ \hat{S}_N(k) + V_v]^{-1} \end{aligned} \tag{7b}$$

and

$$V_{\tilde{a}}(k+1|k) = [I - K_{KF}(k)\ \hat{S}_N^T(k-1)]\ V_{\tilde{a}}(k|k-1) + V_w \tag{7c}$$

with initial conditions $V_{\tilde{a}}(0|-1) = 0.01 \cdot I$, $V_v = 100$ and $V_w = 10^{-7} \cdot I$. Eq. (7c) is processed first, followed by Eq. (7b) and finally Eq. (7a). This is repeated recursively for each input sample. Note that the predictor coefficients depend only on the previous receiver output and the transmitted quantized error. The receiver can therefore compute the predictor without transmission of any extra parameters.

The results in [7] were obtained with a non-robust backward adaptive quantizer. Better performances were obtained by Gibson, Berglund and Sauter in [12] with the Kalman type predictor and a pitch compensated quantizer (PCQ). These results will be discussed further in Chapter V.

CHAPTER IV

ADAPTIVE QUANTIZATION

The adaptive quantizer used in [2] was a two-level adaptive quantizer. The step size (denoted X) was adjusted with the predictor coefficients so as to minimize the quantizer noise. X was then transmitted. In [3], Noll describes several quantizers, both forward and backward adaptive. For ADPCM shemes, he used a forward adaptive variance scheme, as well as various backward adaptive quantizers.

The forward adaptive variance quantizer readjusts the step size X according to an estimate of the error variance based on the short term autocorrelation of the input. X is given by

$$X^{-2} = \alpha[R(0) - \rho_N^T R_N^{-1} \rho_N] \qquad (8)$$

where α is a coefficient to be optimized and depends on the statistics of the error. $R(0)$, ρ_N and R_N were defined on p. 8. Justification of Eq. (8) can be found in [9].

Backward adaptive quantizers can be described by

$$X(k) = \phi_{k-1}[\ X(k-1)\ ]$$

where ϕ_{k-1} depends on the past quantizer output. Backward adaptive quantizers were studied by Jayant [13]. In a Jayant quantizer, $\phi_{k-1}(\cdot)$ is defined by

$$\phi_{k-1}[\ X(k-1)\ ] = f[e_q(k-1)]\ X_{k-1}$$

and f depends only on the last quantizer output. A robust (with respect to transmission errors) version of this quantizer consists in raising $X(k-1)$ to the power β ($\beta<1$) before multiplying by $f[e_q(k-1)]$. The quantizer then forgets the effect of transmission errors at a rate dependant of β.

The PCQ quantizer, formulated by Cohn and Melsa [14] and used by Qureshi and Forney [11] has a certain number of inner levels intended for normal usage and two rarely used outer levels with large expansion factors. The evolution of the step size is given by the following algorithm for a three inner levels (3/5) PCQ illustrated in Fig. 4,

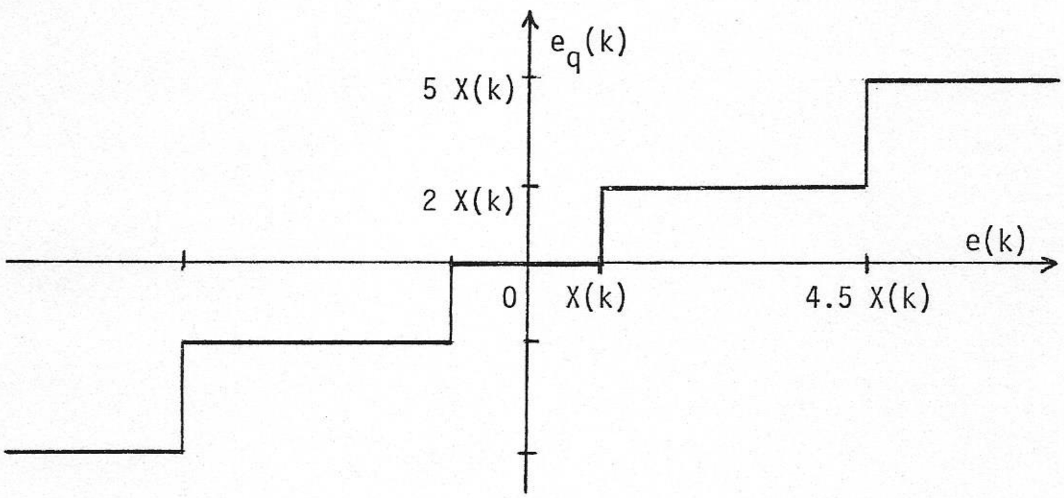

Fig. 4. The (3/5) PCQ.

$$X(k) = 2^{G(k)}$$

where
$$G(k) = GP(k) + C(k) + 1$$

and
$$GP(k+1) = \beta_{GP}\ GP(k) + f_1[e_q(k)]$$
$$C(k+1) = 3/4\ C(k) + f_2[e_q(k)].$$

f_1 and f_2 depend on the previous output:

$$f_1[e_q(k)] = \begin{cases} -7/128 & \text{if} \quad e_q(k) = 0 \\ 7/64 & \text{if} \quad e_q(k) = \pm\ 2\ X(k) \\ 15/64 & \text{if} \quad e_q(k) = \pm\ 5\ X(k) \end{cases}$$

$$f_2[e_q(k)] = \begin{cases} 3/4 & \text{if} \quad e_q(k) = \pm\ 5\ X(k) \\ 0 & \text{otherwise.} \end{cases}$$

β_{GP} is a coefficient to be optimized.

A modified version of the PCQ (MPCQ) was studied by Gibson, Berglund and Sauter in [12]. A minimum step size eliminated certain distortions during silence. This will be discussed further in the next chapter.

CHAPTER V

SIMULATION

A. The Data.

The data used for the simulation consisted of four sentences of about three seconds, in different languages, and spoken by a single male speaker. (See Appendix A for the text of each sentence.) The sentences are representative of the main characteristics of the language. For example, the Arabic sentence contains most of the harsh and guttural phonemes that are typical of that language. Such data should give information as to what type of speech fits the encoding process best. Further details can be found in Appendix A.

B. The Backward Predictive System.

The system used for the simulation was described in [12]. The predictor is a fourth order Kalman type algorithm as described above, with the following modification: the predictor is reset to its initial value and adaptation is inhibited when silence is detected. Silence is detected by the condition

$$\langle \hat{s} \rangle = 1/M \sum_{i=1}^{M} |\hat{s}(k+1-i)| < 10.$$

This modification was brought mainly to give robustness with respect to transmission errors: as long as the system remains stable, it forgets the past transmission errors when silence is detected (usually between words). Also, it seems intuitive to reset the predictor during silence since the predictor coefficients at the end of a word are not in general the optimum initial values for the beginning of the next word. (Note that during silence, the predictor coefficients do not vary since K_{KF} goes to zero with $\hat{S}_N(k)$.)

The quantizer is the robust (3/5) MPCQ used in [12]. It is a (3/5) PCQ with a minimum step size of 1. The optimized value for β_{GP} is 63/64.

Table I gives the frequencies of the different output levels for this system.

TABLE I

OUTPUT FREQUENCIES FOR (3/5) MPCQ WITH KALMAN PREDICTOR (%)

Sentence	-5 X(k)	-2 X(k)	0	2 X(k)	5 X(k)
A	2.25	29.38	40.61	26.09	1.67
E	2.76	28.20	41.50	25.89	1.65
F	2.60	28.11	45.57	22.07	1.65
G	2.00	27.79	42.91	25.75	1.54
Average	2.4	28.	43.	25.	1.6

The entropy, averaged over the four sentences, is 1.765 bits per sample. A source code was designed that achieves the required bit-rate of 16 Kb/s. (See Appendix B.)

C. <u>The Forward Predictive System.</u>

The predictor used is based on the one used in [3] and described above. An eighth order predictor was used. The predictor coefficients were calculated by Durbin's algorithm. A stability check was added to ensure the stability of the predictor. Several quantizers were used with this predictor.

In order to compare predictor performances, the (3/5) MPCQ was used. The predictor coefficients were updated every 25 ms. (Every 200 samples.) The output frequencies are given in Table II.

TABLE II

OUTPUT FREQUENCIES FOR (3/5) MPCQ WITH FORWARD PREDICTOR (%)

Sentence	-5 X(k)	-2 X(k)	0	2 X(k)	5 X(k)
A	2.00	26.86	43.54	26.00	1.60
E	2.18	27.89	44.34	23.84	1.75
F	2.02	30.61	46.41	19.37	1.58
G	1.58	26.60	46.37	23.88	1.57
Average	2.	28.	45.	23.	2.

Although the entropy is slightly less than above (1.75 bits per sample), the transmission rate for this system cannot be brought to 16 Kb/s. The transmission of the predictor coefficients requires approximately 5 bits per coefficient, yielding 1.6 Kb/s for this system. The needed rate is therefore roughly 17.6 Kb/s. A different quantizer was used in order to achieve a lower bit-rate.

First, a robust four level Jayant quantizer was used. One problem was that the average entropy remained too high. (Actually higher than for the MPCQ.)

Also, the system became unstable. This problem was also encountered with a three level Jayant quantizer. It can be intuitively explained as follows: the predictor coefficients are based on the actual input; if the quantizer performs poorly, the predictor does not fit the output (feedback) process and predicts with large errors. This increases the quantizer step size since the quantizer adapts using the actual error process. This continues until saturation. When the predictor coefficients are changed, the system may either remain unstable or recover.

It was finally decided to use a forward adaptive quantizer. A three level quantizer as described in [3] was used. Note that the values needed to calculate the step size X are available from the predictor computations. Eqs. (4) and (8) (p. 8 and p. 12 respectively) yield

$$X^{-2} = \alpha \, [R(0) - \rho_N^T \, \underline{a}_N]$$

and X is obtained easily after solving Eq. (4). The coefficient α must be optimized for the actual error process statistics. Analytically, assuming that e(k) is Laplacian yields the optimal values of $\alpha = 1$ (optimal in the mean square error sense) and $\alpha = .41$ (optimal in the sense of maximizing information). Experimentally, the optimum SNR was obtained at $\alpha = .75$. (Table III) The block length was kept equal to 200 samples. This resulted in the quantizer shown in Fig. 5.

The block length was also optimized within the values allowed by the transmission rate of 16 Kb/s. It was found to be optimum at 140 samples (i.e. 17.5 ms.). (See Table IV)

TABLE III

SNR FOR SENT. E FOR VARIOUS α

α	.41	.5	.6	.7	.75	.8	1
SNR:	9.41	10.66	11.98	11.88	12.23	12.06	11.87

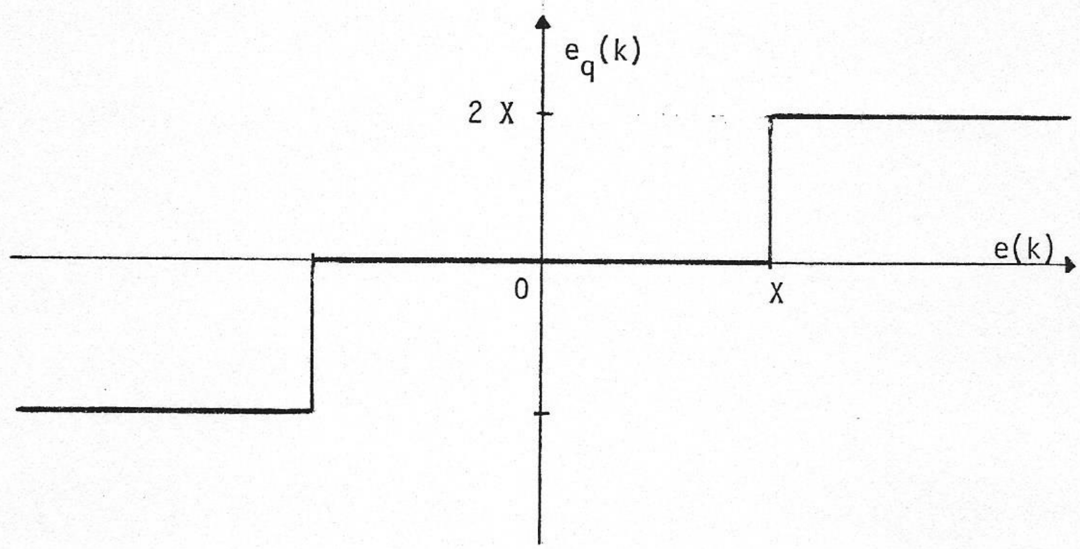

Fig. 5. 3 level optimized quantizer.

TABLE IV

SNR FOR SENT. E FOR DIFFERENT BLOCK LENGTHS

Block length:	128	135	140	145	200
SNR:	12.51	12.60	12.66	12.30	12.23

The output probabilities are given in Table V.

TABLE V

OUTPUT FREQUENCIES FOR THE OPTIMIZED 3 LEVEL QUANTIZER

WITH THE FORWARD ADAPTIVE PREDICTOR (%)

Sentence	-2 X	0	2 X
A	23.32	55.68	21.00
E	23.55	56.33	20.11
F	23.72	59.74	16.53
G	19.69	62.44	17.87
Average	22.57	58.56	18.87

The system then achieves a bit-rate of 16 Kb/s. (See Appendix C.)

CHAPTER VI

RESULTS OF THE SIMULATION

A. Forward Prediction with MPCQ versus Backward Prediction with MPCQ.

Although this forward adaptive system does not achieve the desired transmission rate, the two systems were compared with respect to the predictor performance. The resulting SNRs are given in Table VI.

TABLE VI

COMPARED SNR FOR MPCQ WITH FORWARD AND BACKWARD PREDICTION (DATA RATES ARE 17.6 Kb/s AND 16 Kb/s RESPECTIVELY)

Sentence	Forward Prediction	Backward Prediction
A	17.93	16.29
E	16.30	14.51
F	16.23	14.90
G	19.90	17.85
Average	17.87	16.10

It appears that the eighth order forward predictor performs better than the fourth order Kalman predictor, with an average gain of 1.77 dB. This conclusion was confirmed by informal listening tests. (See Appendix D and Table VII.)

TABLE VII

RESULTS OF LISTENING TESTS FOR FORWARD AND BACKWARD PREDICTION WITH MPCQ (TOTAL FOR FOUR SENTENCES)

Number of listeners	Preferred backward	Preferred forward	Undecided
12	0 (0%)	11 (92%)	1 (8%)

B. Forward Prediction with Optimal 3 Level Forward Adaptive Quantizer versus Backward Prediction with MPCQ.

Here, both schemes achieve the required 16 Kb/s bit rate. Table VII gives the SNRs for both systems.

TABLE VIII

SNR FOR KALMAN PREDICTOR WITH MPCQ AND FORWARD PREDICTOR WITH 3 LEVEL QUANTIZER

Sentence	Backward predictor.	Forward predictor.
A	16.29	14.28
E	14.51	12.66
F	14.90	11.16
G	17.85	15.93
Average	16.10	13.87

The backward adaptive system performs better than the forward adaptive system, with an average gain in SNR of approximately 2.2 dB.

The results of the listening tests confirmed this conclusion. (See Appendix D and Table IX.)

TABLE IX

RESULTS OF LISTENING TESTS FOR KALMAN PREDICTOR WITH MPCQ AND FORWARD PREDICTOR WITH OPTIMIZED 3 LEVEL QUANTIZER

Sentence	Number of listeners	Preferred backward	Preferred forward	Undecided
A	7	5 (71%)	2 (29%)	0
E	16	11 (69%)	4 (25%)	1 (6%)
F	8	8 (100%)	0	0
G	7	4 (57%)	2 (29%)	1 (14%)
Total	38	28 (74%)	8 (21%)	2 (5%)

C. Interpretation of the results.

These results show that at equal transmission rates, the system using backward adaptive prediction performs better than the system using forward adaptive prediction. This is consistent with the analytical results in [8].

Stress is brought upon the degradation of the performances for the forward adaptive system when the transmission of the predictor coefficients is accounted for. Many authors tend to neglect this additional data rate. Following this example would have resulted in comparing only the systems with MPCQ, and would have changed the conclusions of the study.

Note that the results of the listening tests were consistent with the signal to noise ratios for each language. This was not true for subjective comparisons between performances for different languages. Bi-lingual listeners always chose English or French as sounding better than German or Arabic. The SNR indicates a better performance for German and Arabic. (Up to approximately 3 dB.) Whether this is due to the character of each language or only to a more favorable noise spectrum is not known. It seems that the systems do perform differently for the harsh and guttural languages (German and Arabic) than for the smoother ones (English and French).

CHAPTER VII

CONCLUSION

This paper has examined various types of forward and backward predictors for ADPCM speech coding, as well as various adaptive quantizers. Systems have been designed using a Kalman type backward adaptive predictor and a forward adaptive predictor based on the autocorrelation of the input. For each predictor, the best quantizer was used that achieved a transmission rate of 16 Kb/s, and the performances of these two systems were compared. The two predictors were also compared using the same quantizer.

It was seen that the backward adaptive predictive scheme performs better than the forward predictive scheme, for identical transmission rates of 16 Kb/s.

The author believes that this research should be extended to other transmission rates, as well as to the more general case of a non-ideal channel. The effect of transmission errors on the backward adaptive system was studied in [12]. Little work has been done on the effect of transmission errors on the forward adaptive predictive ADPCM when errors may occur in the transmission of the predictor and quantizer parameters.

The divergence of the backward adaptive Jayant quantizers used with a forward adaptive predictor (described on p. 18 of this thesis) is also an interesting problem and could be investigated. This instability was not reported by Noll in [3], although he studied this type of configuration. It seems that the backward adaptive quantizers he used with forward adaptive prediction were not coarse enough for the system to become unstable. It therefore appears that there exists a minimum transmission rate that can be attained by an ADPCM speech coder using a forward adaptive predictor with a backward adaptive quantizer. This property could be investigated in more detail.

REFERENCES

[1] R. A. McDonald, "Signal-to-noise and idle channel performances of differential pulse code modulation systems - Particular applications to voice signals," Bell Syst. Tech. J., vol. 45, pp. 1123 - 1151, Sept. 1966.

[2] B. S. Atal and M. R. Schroeder, "Adaptive predictive coding for speech signals," Bell Syst. Tech. J., vol. 48, pp. 1973 - 1986, Oct. 1970.

[3] F. Noll, "A comparative study of various quantization schemes for speech encoding," Bell Syst. Tech. J., vol. 54, pp. 1597 - 1614, Nov. 1975.

[4] R. W. Stroh, "Optimum and adaptive differential PCM," Ph.D. dissertation, Polytech. Inst. Brooklyn, Farmingdale, N. Y., 1970.

[5] P. Cummiskey, "Adaptive differential pulse code modulation for speech processing," Ph.D. dissertation, Newark College of Engineering, Newark, N. J., 1973.

[6] J. D. Gibson, S. K. Jones and J. L. Melsa, "Sequentially adaptive prediction and coding of speech signals," IEEE Trans. Commun., vol. COM-22, pp. 1789 - 1797, Nov. 1974.

[7] J. D. Gibson, "Sequentially adaptive backward prediction in ADPCM speech coders," IEEE Trans. Commun., vol. COM-26, pp. 145 - 150, Jan. 1978.

[8] J. D. Gibson, "Comparisons and analyses of forward and backward adaptive prediction in ADPCM," Conf. Rec., 1978 Nat'l Telecommun. Conf., Birmingham, Ala., pp. 19.2.1 - 19.2.5, Dec. 3 - 6, 1978.

[9] J. Makhoul, "Linear prediction: a tutorial review," Proc. IEEE, vol. 63, pp. 560 - 580, Apr. 1975.

[10] J. Durbin, "The fitting of time-series models," Rev. Inst. Int. Statist., vol. 28, no. 3, pp. 233 - 243, 1960.

[11] S. U. H. Qureshi and G. D. Forney, Jr., "A 9.6/16 Kb/s speech digitizer," Conf. Rec., IEEE Int. Conf. on Commun., San Francisco, Calif., pp. 30.31 - 30.34, June 1975.

[12] J. D. Gibson, V. P. Berglund and L. C. Sauter, "Kalman backward adaptive predictor coefficient identification in ADPCM with PCQ," to be published.

[13] N. S. Jayant, "Adaptive quantization with a one word memory," Bell Syst. Tech. J., vol. 52, pp. 1119 - 1144, Sept. 1973.

[14] D. L. Cohn and J. L. Melsa, "A pitch compensating quantizer," Conf. Rec., 1976 IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 258 - 261, 1976.

[15] F. Jelinek and K. S. Schneider, "On variable-length-to-block coding," IEEE Trans. Inform. Theory, vol. IT-18, pp. 765 - 774, Nov. 1972.

[16] D. L. Cohn and J. L. Melsa, "The residual encoder - an improved ADPCM system for speech digitization," IEEE Trans. Commun., vol. COM-23, pp. 935 - 941, Sep. 1975.

APPENDIX A

SPEECH DATA

The Arabic, French and German sentences are translations of the English sentence. All were spoken by a single male speaker. The speech was then filtered and digitized at 8000 Hz. The sample values were stored using 12 bits.

Sent. A: (Arabic) " كأس حليب خير من قطعة خبز "

Sent. E: (English) "A glass of milk is better than a piece of bread."

Sent. F: (French) "Un verre de lait est meilleur qu'un morceau de pain."

Sent. G: (German) "Ein Glass Milch ist besser als ein Stück Brot."

The Arabic sentence can be written phonetically as follows:

"Ka'su Haleebin khayrun mina qiTεatu khubzin."

APPENDIX B

SOURCE CODE AND ACHIEVED BIT-RATE FOR THE BACKWARD PREDICTIVE - MPCQ SYSTEM

The following source code is proposed, where the quantizer output levels are numbered from 1 to 5 by decreasing frequency. (Table X.)

TABLE X

SOURCE CODE FOR BACKWARD PREDICTION - MPCQ

Message	Codeword	Message	Codeword
111	0000	114	1111 0011
112	0001	115	1111 0101
113	0010	24	1111 0110
12	0011	34	1111 1001
13	0100	25	1111 1010
14	0101	35	1111 1100
15	0110		
21	0111		
22	1000		
23	1001		
31	1010		
32	1011		
33	1100		
4	1101		
5	1110		

This code is a variable input length to block source code, as described in [15] and used in [16]. It has the advantage over Huffmann codes that channel errors do not cause long losses of synchronization.

The coder accepts a sequence of input symbols until a message is formed. It then transmits the corresponding codeword.

With this source code, the considered system achieves a transmission rate of 16.072 Kb/s, which is approximately equal to the desired 16 Kb/s.

APPENDIX C

SOURCE CODE AND ACHIEVED BIT-RATE FOR FORWARD PREDICTOR WITH 3 LEVEL QUANTIZER

A simple 3 ternary digits to 5 bits code was used, resulting in 1.667 bits per sample, or 13.33 Kb/s for the quantizer output.

The predictor coefficients are transmitted using an average of 5 bits per coefficient. The quantizer step size is transmitted using 8 bits. This results in a total transmission rate of 16.076 Kb/s

APPENDIX D

DESCRIPTION OF THE LISTENING TESTS

The listening tests comparing the performances of the two predictors as discussed in Chapter VI, paragraph A were very informal. No earphones were used, although the tests were held in a silent room. Each listener judged only the performance associated with his native language. Only the total results, summed for the four sentences, are given.

The listening tests comparing the two systems achieving the required bit-rate were more formal: earphones were used and the tests were held in the same silent room. As above, each listener only heard and judged the sentence associated with his native language. In the case of a perfectly bilingual person, a comparison was made of the performances for both languages. Male and female listeners were tested. Each person was given the following instructions: "You are about to listen to a sentence that will be spoken twice using different methods of ADPCM speech coding. Please indicate which you think sounds better. The sentences will be repeated as many times as you wish." Most listeners asked to hear the sentences 3 times or more before giving their opinion.

VITA

Louis Clyde Sauter was born in Rome, New York on September 18, 1955, the son of Mr. and Mrs. Louis C. Sauter.

Mr. Sauter has lived in France since 1965. He attended high-school in Nice, France, graduating in 1974. He has been a student of the Ecole Nationale Superieure des Telecommunications in Paris, France from which he received the diploma in 1979. In 1977, he was employed by IBM-France at La Gaude, France, where he worked in the speech coding area. Louis Sauter entered the Graduate College of Texas A&M University in Fall 1978. His permanent address is 25 rue Georges Doublet, 06100 Nice, France.