APPROVED APPROACH TO PROTECTION OF SYNTHETIC SPEECH
DATA TRANSMITTED OVER NOISY CHANNELS

Rateb, M.  A.  Abulshohoud, W.  A.  Elmaghazzy, A.
* M.T.C.  ** Armament Authority, Cairo, Egypt.

RESUME

Le caddage de parole base sur le principe de "caddage lineaire predictive" (CLP), permet la transmission simutanee de plusieurs canaux.
Les erreurs du canal, quant'il s'agisse dans une partie particulaire du Signal Code en CLP, sont un source tres efficace de la distorsion.
On expose ici une methode de la correction des erreurs pour le signal code en CLP La methode prend en consideration l'effet pondere des erreurs sur la qualite du signal, et il, est nomme "Supervision multiple de caddage". La methode proposee est realisee par une structure de caddage en reseau Elle permet la protection contre les erreurs en permettant, au meme temps, un rythme de transmission relativement eleve.

SUMMARY

LPC speech signals can achieve greater number of multiplexed channels than PCM and DM at nearly the same speech quality when TM is considered.
Channel errors falling in specific segments of the transmitted signal yielding to harmful distortion made LPC less attractive for practical communication systems application.
This paper introduces a proposed weighted FBC technique in which the LPC parameter data are protected against channel errors according to the effect of those errors on the received speech signal. This approach referred to as "Multiple-Code Supervision" yielded to a "Lattice Coding Structure" which provides the required protection at relatively high transmission rate.

1- INTRODUCTION

The quality of speech when using waveform coders significantly deteriorates if bit rate falls below 16 kb/s. On the other hand v coders achieve maximum quality at a bit rate of approximately 4.8 kb/s and a negligible quality improvement is attained by further bit-rate increase. Reduction of bit rate to the order of 2.4 kb/s and even smaller will cause a nonremarkable degradation of speech quality. The basic concept behind the linear prediction is that speech could be modelled as the outcome of a linear time-varying system excited by voiced or unvoiced speech. The voiced speech is in the quasi-periodic pulses while the unvoiced is in the noise-like form. The system could be uniquely identified as an all-pole linear system. The formulation of linear prediction analysis yield to nearly equivalent methods like covariance, autocorrelation, lattice, inverse filter, maximum likelihood, spectral estimation, etc. Attention will be devoted to the autocorrelation method which is the most suitable in case of short-time averages processing.

According to the simplified model in Fig.1, the speech samples s(n) are related to the excitation signal u(n) by the difference equation:

\[ s(n) = \sum_{k=1}^{p} a_k s(n-k) + G u(n) \]  

with p poles, u(n) is the proper input excitation, \( a_k \) is the set of predictor coefficients and G is the filter gain.

On applying the autocorrelation the waveform segment under consideration is windowed (usually by Hamming window) such that:

\[ s_n^n \begin{cases} n & 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases} \]

The autocorrelation function \( R(j) \) is defined by:

\[ R(j) = \sum_{n=0}^{N-1-l} s_n s_{n+j} \]

Then the LPC coefficients are achieved from the following normal equations:

\[ \sum_{k=1}^{p} a_k R(k-j) = -R(j), \quad p \geq l \]

The solution of normal equations (4) could be achieved either by direct iterative techniques. The Durbin's method is preferable since it is twice times faster as Levinson's and requires only 2p storage locations and \( p^2 \) operations. The Durbin's recursive procedure takes place as follows:

\[ E = R(0) \]

\[ k_j = \frac{1}{R(j)} \left[ \sum_{i=1}^{j} a_i R(j-i) \right], \quad j \leq N-1 \]

\[ a_j = a_j + k_j a_{j-1}, \quad j = 1, 2, \ldots, p \]

\[ E_j = (1-k_j^2) E_{j-1} \]

Equations (5a-5e) are recursively solved for \( j = 1, 2, \ldots, p \). The final solution is:

\[ a_j = a_j(p), \quad p \geq j \]

There exist many possible sets of parameters that can uniquely characterize the all-pole filter H(z) and its inverse A(z) where:

\[ H(z) = G \left[ 1 + \sum_{k=1}^{p} a_k z^{-k} \right] \]

\[ A(z) = G/H(z) \]

Those sets are the poles of H(z) or zeros of A(z), impulse response of either H(z) or A(z), autocorrelation coefficients, spectral coefficients of A(z) and reflection coefficients \( k_j(p) \geq j \). The last ones (the reflection or PARCOR coefficients) are the most

Fig.1 Block diagram For Speech Production.
suitable for quantization since they secure the simul-
taneous filter stability upon quantization and the natural ordering of those parameters. 

For filter stability it must be fulfilled: 

\[ |k_i| < 1 \]  \( \text{(8)} \)

which is obviously satisfied by \( (5) \)–\( (e) \). 

Pitch detection techniques utilize the time-domain, frequency-domain or time and frequency-domain proper-
ties of the speech signals. Various quantization approaches are for quantizing the LPC reflection coeffi-
cients exist. The most attractive one is the log area ratio (LAR) quantization method. The LAR coding sche-
me transforms \( k \) into a new coefficient \( q \) given by 

\[ q_i = \log \left( \frac{(1-k_i)}{(1-k_{i-1})} \right) \]

(9)

Consequently, the coding scheme uses efficiently a relatively small number of bits for each coefficient. 

However from point of view of hardware realization the piecewise linear quantization is much simpler. 

2- BASIC CONCEPTS

2-1 Considerations Affecting the Selection of Frame Structure and LPC Bit Rate. It has been found that the pitch and main parameters could be geometrically interpolated \( [3] \). This geometrical interpolation takes place linearly on a log scale. 

The PARCOR coefficients are bounded \( |k| < 1 \) and they could be pooled directly while the filter stability is maintained. Similarly the resulting log area ratio (LAR) coefficients can be interpolated yielding to a stable system if the original reflection coefficients preserve this stability. Due to the fact that the PARCOR coefficients have highly skewed distribution, LAR coefficients given by \( (9) \) are the most suitable for transmission and 5–6 bits per log area ratio are sufficient to achieve the same quality synthetic speech as obtained from coded parameters. The choice of the number of LAR coefficients is tightly connected with the variation of prediction error with the predictor order \( p \) specially for voiced speech. For the autocorrelation method it was found by \( [3] \) that the normalized mean-squared error \( \nu \) remains at a value 0.1 for \( p = 7 \) for short windows \( (N=60) \) which is our typical case. 

\[ \nu = \frac{1}{N} \sum_{n=1}^{N-1} s^2(n) / s^2(0) \]

(10)

where \( s(n) \) is the output of the prediction error filter corresponding to the speech segment \( s(n) \). Thus, select-
ing \( p=10 \) will hand over a satisfactory result. 

On the basis of all the previously-mentioned considerations the frame structure will consist of 7 LAR coeffi-
cients each of 5 bits and the last three each of 4 bits, the pitch period is represented by 6 bits, the gain by 5 bits and \( v/\nu \) by one bit. In addition a permanent synchron. bit attaches frame beginning yielding to a total sum of 60 bits per frame. When transmitting LPC speech data over multiplexed PCM systems we have two alternatives. 

The first is to decrease the LPC vocoder actual bit rate to 2 KHz then rise it to 4 KHz to allow the addition of equal number of redundant bits. 

In the second alternative the actual bit rate is doubled and the speech quality on the expense of decreasing the number of multiplexed channels by one half. Consequently, 512 channels for the first and 256 channels (including signalling and synchronization) for the second version could be mul-
ti-plexed instead of 32 channels for the European 2,048 

MB PCM systems. For the Bell T-1 carrier 1,544 Mb/s system, 384 channel for the first and 192 channel for the second version instead of 24 channels in case of PCM.

2-2 Basic Idea Of Proposed Coding Approach. 

In most cases the transmitted data inside a message are of different levels of importance. Channel errors falling in specific critical segments of the message may lead to the loss of the whole message. This is a typical case when LPC data transmission is considered. Channel errors falling into the first two PARCOR coefficients in the LAR form or in the pitch period will yield to the complete distortion of the data contained in the whole frame. \( [10] \). 

The newly proposed weighted FEC based on "Multiple Code Supervision" of the critical data segments, hands over an effective solution to this problem on the expense of the least possible redundancy. In this technique the critical segment is sharing more than one error correcting and detecting code encoding other data of lower levels of importance. The number of supervising codes is determined according to the immunity required and to the available redundancy. 

According to Fig.2.a, the pitch period \( (PP) \) with an added parity is encoded with the 6th \( (g_6) \), the 9th \( (g_9) \) and 10th \( (g_{10}) \) LAR coefficients with 5 parity checks for each to establish the codes \( (16,11) \) with minimum distance \( d=4 \) \( c_1 \), \( c_2 \) and \( c_3 \) respectively. The \( (PP) \) segment with added parity is supervised by 3 codes. 

Fig.2 Proposed coding lattice For LPC

In Fig 2-b the first LAR coefficient \( (g_1) \) is supervis-
ed by the codes \( c_1, c_6 \) and \( c_7 \) while \( (g_2) \) is supervised by \( c_2, c_6 \) and \( c_3 \). Each of these has an added parity and all codes are \( (16,11) \) with \( d=4 \), \( (g_3) \) with added parity is supervised by \( c_3, c_7 \) while \( (g_4) \) with added parity is supervised by \( c_6, c_8 \) and \( c_9 \) without added parity. The "lattice Coding Structure" introduced secures a coding strength propor-
tional to the importance of those parameters.

3- DECODING CRITERIA AND PROBABILITIES

3-1 Decoding Criteria

The selected block code with \( d=4 \) is a single error-
correcting code with an additional overall parity. This parity enables the detection of all even error patterns except those yielding to a permissible code word. 

For an arbitrary code segment having an additional parity supervised by \( k \) code blocks the following criteria are valid: 

a- Basic criteria for correct decoding 
1- No error is detected in all supervising code blocks and local parity detects no error. 
2- All code blocks detect even errors while the local parity detects no error. 
3- When at least two code blocks locate the same error inside the segment and local parity detects an error, while the rest of code blocks locate different errors.

b- Criteria for error detection 
1- Local parity detects an error while supervising codes point to different error locations inside or outside the segment. 
2- All supervising codes detect even errors and local parity detects an error.
3-2 Associated Probabilities

Assume a BSC having channel probability of error P. Let the length of the supervised code segment by \( n_s \)
and the length of the supervising code be \( n_s \), \( j = 1, 2, \ldots, R \). On the basis of the upper mentioned criterias the
probability of correct decoding will be:

\[
P_c(n_s) = (1-P)^{n_s} + n_s P (1-P)^{n_s-1} \left[ \sum_{j=1}^{R} \binom{n_s}{j} R^{1-P} n_s^{j-1} \right]
\]

The probability of error decision is simply given by:

\[
P_o(n_s) = \sum_{r=1}^{n_s} \binom{n_s}{r} P_r (1-P)^{n_s-r-1} \left[ \sum_{j=1}^{R} \binom{n_s-j}{r} R^{1-P} n_s^{j-1} \right]
\]

and the probability of error detection is:

\[
P_d(n_s) = \sum_{k=1}^{n_s} \binom{n_s-k}{2} \left[ (n_s-k)^{2r-1} \right]
\]

3-3 The Complete Decoding Algorithm

The complete decoding algorithm is interpreted by the simple flowchart Fig. 3.

Fig. 3 Simplified Flowchart of the decoding algorithm.

The received lattice coding structure is stored in a buffer and the decoding begins by the segment having the
highest supervising capabilities. The decoding is accomplished on the basis of logic introduced in 3-1. If the
decoding result will yield to a correct decision the supervising code is afterwards corrected, otherwise the following data segment is pro-
cessed. After all segments are being decoded, the erroneous parameters which have error beyond the
correction capabilities are interpolated.

4 - RESULTS ANALYSIS AND CONCLUSIONS

Fig. 4 illustrates the probabilities of correct decision \( P_c \) and of error \( P_o \) for a code segment \( 6,5 \)
supervised by three codes \( 16,11 \) which is in direct correspondence with the proposed coding structure.
The results computed on the basis of (16) and (17) show a dramatic increase in \( P_c \) and decrease in \( P_o \)
when compared with the capabilities handed over by the supervising code. This deviation increases \( P_c \) is in-
creased. However there is an implicit improvement in capabilities of the supervising code handed over by
the proposed coding lattice structure. This additional capabilities occurs when one supervising code has two
channel errors one falling inside and the other outside the supervising segment. Both errors in such cases
are correctable. The overall probability of correct decoding for this supervising code will be:

\[
P_c(n_s) = \sum_{i=0}^{n_s} \binom{n_s}{i} R^{1-P} n_s^{i-1} + n_s (n_s-n_s) P(n_s) \cdot
\]

\[
P_o(n_s) = \left( \sum_{i=0}^{n_s} \binom{n_s-i}{2} (1-P)^{n_s-2i} \right) - \sum_{i=0}^{n_s} \binom{n_s-i}{2} (1-P)^{n_s-2i}
\]

As the number of supervising code blocks increases expressions (16) and (17) will reduce to:

\[
P_c(n_s) = (1-P)^{n_s} + n_s P (1-P)^{n_s-1} = \sum_{i=0}^{n_s} \binom{n_s}{i} P^i (1-P)^{n_s-i}
\]

This means that the supervising short data segment will have the capability of an independent error corre-
correcting code of the segment length without additional redundant bits. This will yield to a substantial high drop in the probability of error.
REFERENCES


