10.21437/Eurospeech.2003-239

INTRODUCTION OF THE CELP STRUCTURE OF THE GSM CODER IN THE ACOUSTIC ECHO CANCELLER FOR THE GSM NETWORK

H.Gnaba (2), *M.Turki-Hadj Alouane* (2), *M.Jaidane-Saidane* (2), *P. Scalart* (1)

(1) France Télécom R&D DIH/DIPS, Technopole Anticipa Lannion,
(2) Unité de Recherche Signaux et Théorie des Systèmes − Ecole Nationale d'Ingénieurs de Tunis
ENIT BP 37, 1002 Tunis, Tunisia

helagnaba@yahoo.fr − monia.turki@enit.rnu.tn − meriem.jaidane@enit.rnu.tn − pascal.scalart@rd.francetelecom.fr

ABSTRACT

This paper presents a new structure of an Acoustic Echo Canceller (AEC) designed to operate in the Mobile Switching Center (MSC) of a GSM network. The purpose of such system is to cancel the echo for all the subscribers. Contrarily to the conventional AEC, the proposed combined AEC/CELP Predictor is able to take into account the non linearities introduced by the GSM speech coders/decoders. A short term predictor is used to model the behavior of the codecs. This new combined system presents higher performance compared to the conventional AEC.

1. INTRODUCTION

The implementation of the AEC in the GSM network allows the operator to enhance the audio quality for speaker phones as well as cell phones [1]. The MSC seems to be the obvious location for a centralized AEC.
The main problem of this AEC is the nonlinearity introduced by speech codecs, which are cascaded along the echo path as shown in figure 1, for a mobile-land call.

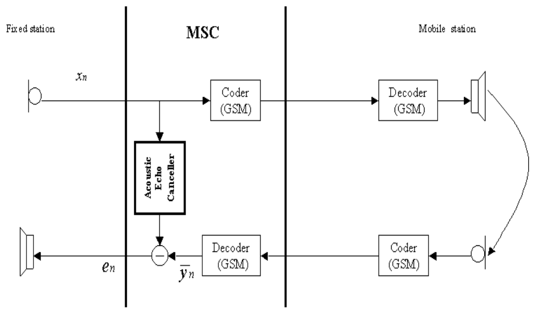

Figure 1: Centralized AEC for a mobile-land call.

Low-bit-rate codecs (such as G723, G729, GSM,..) used in GSM network to reduce the speech transmission rate, introduce severe distortions to the far-end and the near-end microphone signals [2][3].

THIS WORK WAS SUPPORTED BY FTR&D (LANNION) AND ENIT (TUNIS).

Hence, a classical linear AEC in the MSC is not sufficient and doesn't guarantee an unperceivable echo reduction [2][3][4].
On the other hand, an AEC based on non linear adaptive filtering [5] is not adequate, since the non linearities of the GSM codecs present complex characteristics and don't obey to conventional models. As solution to this problem, centralized AEC systems are recently proposed in the literature [2][3][4].
In the same way, we propose in this paper an original combined AEC/CELP Predictor based on short term speech analysis. This new system exploits the CELP structure of the GSM codecs, to further suppress the residual echo produced by the linear adaptive filter.
The remainder of this paper is organized as follows. In section 2, we present the problematic of the centralized AEC when a communication between two mobile stations is considered. Section 3 presents the preliminary approach of the proposed combined AEC system in such context. Finally, we propose a combined AEC/CELP Predictor, in order to recover a non distorted local speech. Results presented in section 4.2 show that the proposed combined AEC is very promising for high acoustic echo suppression and presents robust performance in this centralized context of echo cancellation.

2. PROBLEMATIC OF THE CENTRALIZED AEC IN A MOBILE-MOBILE COMMUNICATION

In this section, we discuss the case of a communication between two mobiles, that is more complex than the case of mobile-land call.

2.1. Practical considerations

The problem that arises in this situation (mobile-mobile call) is that we have only in the MSC coded signals that can't be processed by the AEC. So, the operation of decoding these signals is necessary.
It's important therefore to implement in the MSC (:"with bold style in figure 2") the following features,

- two GSM decoders, in order to have the decoded version of signals that we can treat.
- one GSM coder to code the residual echo before its transmission to the other mobile.

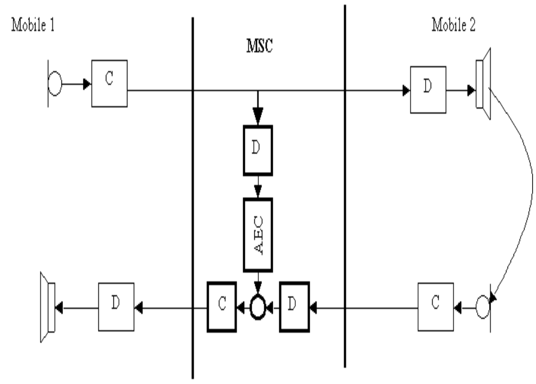

Figure 2: Centralized AEC for a communication between two mobiles.

2.2. Identification problem

In figure 3, we present the adaptive identification schema equivalent to the centralized AEC of figure 2. The filter F represents the impulse response of the echo path and H_n is the adaptive echo canceller. The presence of the codecs

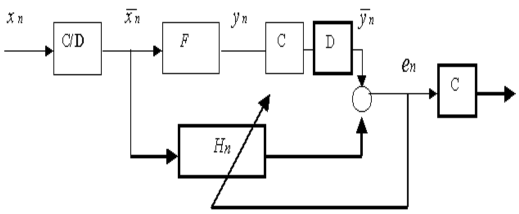

Figure 3: The adaptive identification schema of the centralized AEC.

in tandem constitutes the particularity of this structure. Indeed, the coding-decoding process of the near-end microphone signal y_n introduces a specific kind of non linearity. Therefore, the impulse response of the entire echo path to identify, consists of the linear filter F cascaded with a non-linear function that represents the coder/decoder.
In the following, the time evolution of the adaptive filter H_n is controlled by the Normalized Least Mean Squares (NLMS) algorithm.
The NLMS algorithm can be written as follows :

$$e_n = \bar{y}_n - H_{n-1}^T \cdot \bar{X}_n \tag{1}$$

$$H_n = H_{n-1} + \mu \frac{e_n}{\bar{X}_n^T \cdot \bar{X}_n} \cdot \bar{X}_n \tag{2}$$

where,
- e_n is the residual echo signal.
- $\bar{y}_n$, represents the coded/decoded microphone signal.
- $\bar{X}_n = [\bar{x}_n, ..., \bar{x}_{n-p}]^T$ is the sequence vector of the last p samples of the coded/decoded far-end signal $\bar{x}_n$.
- $H_n = [h_n^0, .., h_n^p]^T$ represents the adaptive filter vector of length p.
- $F = [f_0, .., f_q]^T$ is the vector of the q coefficients of the acoustic echo path.

Equation (1) shows that in the optimal case ($p = q$), we can only remove the echo caused by the linear part of the hole echo path. Hence, the residual echo roughly equal to the quantization noise ($y_n - \bar{y}_n$), is audible and still too high to achieve the goal of echo cancellation [2][3]. In order to reduce the residual echo given by the conventional AEC, we propose in the next section a centralized AEC that incorporates a CELP Predictor, to model in a simple manner the coder/decoder effects.

3. PRELIMINARY APPROACH OF THE COMBINED AEC/CELP PREDICTOR

3.1. A simplified model of the GSM codec

The GSM coder is an analysis by perceptual synthesis coder based on CELP (Code Excited Linear Prediction) method [6].
The GSM coding-decoding function is modeled here (figure 4) by an analysis ("prediction") of the audio signal x_n, followed by a quantization of the prediction error ϵ_n, and a synthesis ("an inverse prediction") operating on the quantized prediction error $\bar{\epsilon}_n$. The quantization noise

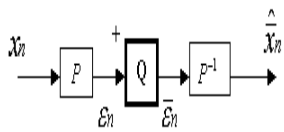

Figure 4: The simplified model of the GSM coder/decoder.

$b_n^x = \bar{x}_n - x_n$ related to the speech signal can be then approximated by the following equation,

$$b^x(z) = \bar{x}(z) - x(z) \simeq [\bar{\epsilon}(z) - \epsilon(z)] \cdot P_x^{-1}(z) \simeq \epsilon^x(z) \cdot P_x^{-1}(z) \tag{3}$$

Hence, as known in the literature related to prediction theory, the power of the prediction error ϵ_n^x is less than that of b_n^x since

$$E((\epsilon_n^x)^2) = \frac{E((b_n^x)^2)}{G^2} \tag{4}$$

where G denotes the prediction gain, that is as higher as the correlation input is higher.

3.2. Preliminary approach of the combined AEC/CELP Predictor

In the optimal case corresponding to $p = q$ and when a low NLMS step size is used, the residual echo e_n is roughly equal to $b_n^y = \bar{y}_n - y_n$, the quantization noise of y_n. According to (4) and (3), it is possible by prediction to obtain a residual echo with power close to the one of ϵ_n^y. To reach this aim, we realize a prediction of the signal $\bar{y}_n$, and the computed predictor P will operate on the residual echo e_n given by the conventional AEC filter. The proposed combined AEC/CELP predictor is presented in figure 5, where $P = [p_1, .., p_M]^T$ the vector of the M coefficients of the predictor, the new residual echo signal is $e_n^{'}$ output of the proposed combined AEC/CELP Predictor system.

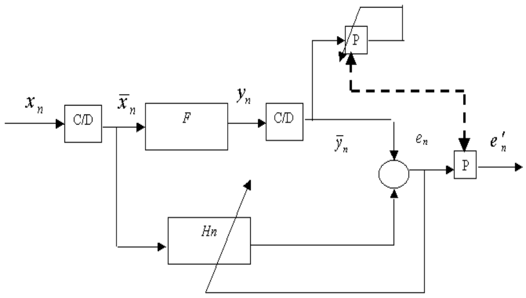

Figure 5: A preliminary approach of the combined AEC/CELP Predictor system.

3.3. Theoretical analysis

When we consider that the input of the AEC is a white sequence, and under the assumption of a statistical independence between the quantization noise b_n^y and the echo signal y_n, we show that the power of the residual echos e_n and $e_n^{'}$ are given by,

$$E(e_n^2) = E((b_n^y)^2) + E(x_n^2)(\sum_{i=p+1}^{q} f_i^2) \quad (5)$$

$$E(e_n^{'2}) = E(x_n^2)(1 + \sum_{j=1}^{M} p_j^2)(\sum_{i=p+1}^{i=q} f_i^2) + \frac{E((b_n^y)^2)}{G^2} \quad (6)$$

where $G >> 1$ is the prediction gain.
The analysis of the equations (5) and (6) show the following results:
- in the optimal case corresponding to $p = q$, the combined AEC/CELP Predictor performs always better than the conventional AEC system since $E(e_n^{'2}) = \frac{E((b_n^y)^2)}{G^2}$ is less than $E(e_n^2) = E((b_n^y)^2)$. The gain reached therefore is as proportional to the prediction gain G.
- in the undermodelling case $p < q$ (practical case), the combined AEC/CELP predictor performs better than the conventional AEC only if the G is too higher and $(\sum_{j=1}^{M} p_j^2)$ is too low in such a way that

$$E(e_n^{'2}) - E(e_n^2) = E(x_n^2)(\sum_{j=1}^{M} p_j^2)(\sum_{i=p+1}^{i=q} f_i^2) + E((b_n^y)^2)(\frac{1-G^2}{G^2})$$

is negative. These conditions are satisfied in the general case.

4. PROPOSED AEC/CELP PREDICTOR SYSTEM

4.1. Retained AEC/CELP Predictor schema

As shown in the previous section and also will be shown later, the first approach of the combined AEC/CELP Predictor system works well during echo period. However, during local speech period, it is not efficient since the predictor must operate only on the echo and not on the local speech. The predictor coefficients must turn to zero when only the local speech is transmitted. For this reason, we propose to modify the previous version by computing the predictor from the estimated echo signal z_n instead of real echo $\bar{y}_n$. The retained version of the hybrid AEC/CELP Predictor system is presented in figure 6.

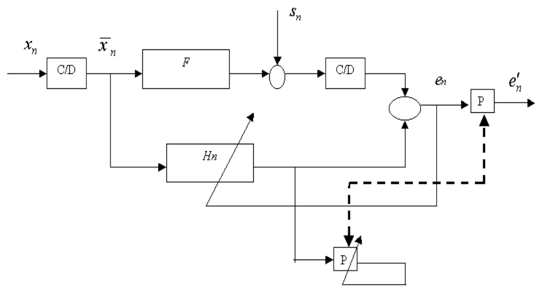

Figure 6: Retained combined AEC/CELP Predictor schema.

4.2. The performance of the AEC/CELP Predictor during echo period

In this section, we evaluate the performance of the combined AEC/CELP Predictor in the echo period where only the echo is handled by the system. The performance of the combined AEC/CELP Predictor system are evaluated in two cases : the optimal case and the adaptive case.

a) The optimal case

The aim is to evaluate the performance of the prediction process assuming that the adaptive filter H_n is constant and equal to F (the real impulse response of the acoustic echo path). The predictor is operating in an adaptive mode. Its time evolution is controlled by the NLMS algorithm. In order to test the proposed combined AEC/CELP Predictor, we consider an impulse response of the echo path F of length $q = 2000$. We compute the ERLE (Echo Return Loss Enhancement : ratio of the echo power over the residual echo power).
We plot in figure 7 the time evolution of the ERLE corresponding to three cases: with predictor of order $M = 10$ (curve (1)), with predictor of order $M = 2$ (curve(2)) and without predictor (curve (3)). The simulation results show that the proposed new combined AEC/CELP Predictor system significantly outperforms the conventional AEC system. This is valuable for $M = 2$ and $M = 10$. The gain inherent to the prediction process reaches 15 dB.
What's more, the increasing of the predictor order does not improve remarkably the ERLE. Hence, we can achieve a large attenuation of residual echo with a low predictor order. Therefore the complexity of the proposed system is not really increased compared to that of a classical AEC system.

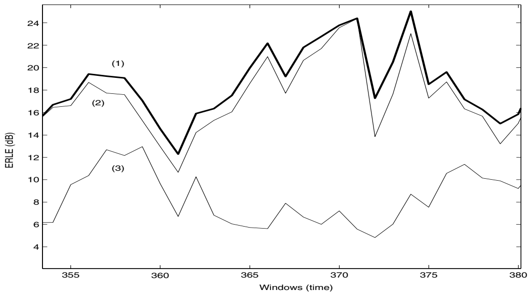

Figure 7: ERLE in the optimal case.

b) The adaptive case

In this case the length p of the AEC adaptive filter (H_n) is less than that of the echo path filter (F) q.
The time evolution of the adaptive AEC filter H_n and the adaptive predictor P_n are ensured by the NLMS algorithm. For this simulation, we consider that $p = 300$, $q = 2000$, $M = 2$. The time evolution of the ERLEs related to both cases (with predictor : curve (2) and without predictor : curve (1)) are presented in figure 8. In this case, we also notice that the proposed combined AEC system has better performance than the classical one, an improvement of 13 dB is obtained. As expected (3.3) the gain reached here is less than that corresponding to the optimal case for $M = 2$ (15 dB). The difference is not significant compared to the difference between the real echo path length, $q = 2000$, and the adaptive filter length $p = 300$.

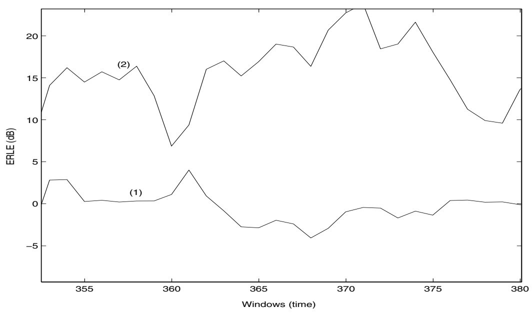

Figure 8: ERLE in the adaptive case.

4.3. The performance of the AEC/CELP Predictor during local speech period

The purpose here, is to test the behavior of the adaptive predictor when a local speech is present. For the simulation we consider a data input sequence constituted by an echo period followed by a local speech one. The performance of the first version of AEC/CELP predictor (figure 5) are compared to that of the retained one (figure 6). the time evolution of the first coefficient of the predictor is plotted in figure 9 : for curve (1) the input of the adaptive predictor is $\bar{y}_n$ and for curve (2) the input is z_n. This figure shows, that once we predict the signal z_n, the predictor coefficient approach to zero during the local speech period (as the second coefficient). Therefore we can recover a non distorted local speech. What's more, from curve (2) we notice that the coefficient switchs rapidly from its value relative to echo period to that relative to local speech period value. This is an important advantage for a real time implementation of the proposed AEC/CELP Predictor system.

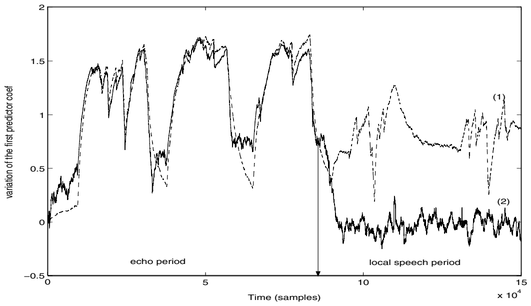

Figure 9: The variation of the first coefficient of the predictor.

5. CONCLUSION

We propose here a new combined AEC/CELP Predictor tailored to the acoustic echo cancellation in the centralized GSM context that takes into account explicitly the coder/decoder structure. It has better performances during the echo period and the local speech period than the classical AEC. It has a low complexity that allows an implementation for real time application.

REFERENCES

[1] D.J.Jones, S.D. Waston, K.G.Evans, "A network speech echo Canceller with confort noise," Eurospeech 1997.

[2] H.Gnaba, P.Scalart, M.Turki and M.Jaidane, "Combined acoustic echo canceller for the GSM network," EUSIPCO, vol. I, pp. 547-550, Sep. 2002.

[3] X. Lu and B. Champagne, "Acoustic echo cancellation over a non-linear channel," in Proc. IWAENC'01, pp. 139-142, Sep. 2001.

[4] X. Lu and B. Champagne, "Pitch analysis-based acoustic echo cancellation over a non-linear channel," EUSIPCO, vol. I, pp. 159-162, Sep. 2002.

[5] A. Stenger, L. Trautmann, R. Rabenstein, "Nonlinear acoustic echo canceller with 2nd order adaptive Volterra filters," ICASSP 99, USA, 1999.

[6] ITU-T Recommendation G729, Coding of Speech at 8 kbits/s Using Conjugate-Structure Algebraic-Code-Excited-Linear-Prediction (CS-ACELP), International Telecommunication Union, Mars. 1996.