

Low complexity forward error correction for CELP-type speech coding over erasure channel transmission

Nadir Benamirouche¹ · Bachir Boudraa² · Domingo López-Oller³ · José L. Pérez-Córdoba³

Received: 4 March 2016 / Accepted: 16 August 2016 / Published online: 29 August 2016
© Springer Science+Business Media New York 2016

Abstract One of the well-known problems of Code-Excited Linear Prediction (CELP)-type codec is its vulnerability to a frame erasure. When a frame is erased, the inter-frame dependency introduced by the Long Term Prediction causes a desynchronization of the Adaptive Codebook (ACB) which introduces in its turn an error propagation through the correctly received frames. In this paper, we propose a media-specific Forward Error Correction (FEC) method using a Pitch-Pulse Codebook (PPCB)-based approach to model the ACB contribution for voiced frame (frame onset) determined under Zero Crossing Rate constraint. The PPCB uses a single pulse optimized by Multipulse Maximum Likelihood Quantization algorithm to model the pitch-like contribution at the encoder side while the quantized version of that pulse will be sent as FEC information to resynchronize the ACB at the decoder side after a frame erasure. Through this approach a noticeable improvement of the synthesis speech quality is achieved

under adverse channel conditions with the advantage of low computational complexity while the legacy bit-rate of the codec is kept unchanged.

Keywords VoIP · Pitch pulse · Forward error correction · Error propagation · ACB resynchronization

1 Introduction

The last few years have witnessed an increasing use of IP networks. This has led to considerable interest in the integration of voice services, such as Voice transmission over IP networks (VoIP). Nowadays, speech services are interesting since they offer an easy and economical alternative for providing a reliable speech transmission over IP (Chandra and Ray 2015). Although, the Quality of Service (QoS) of IP networks has greatly improved over the recent years, unfortunately, the QoS of VoIP does not yet provide toll-quality voice equivalent to that offered by the traditional public switched telephone network (Toral-Cruz et al. 2013). Indeed, another critical issue for media streaming applications such as VoIP, is its vulnerability to end-to-end performance (Bhebhe and Parkkali 2011; Gupta et al. 2015). In fact, voice traffic is very sensitive to issues such as congestion (jitter), delay (buffer) and packet losses (Singh et al. 2014). Hence, the missing packets have to be concealed at the decoder side using packet loss concealment (PLC) techniques.

A wide variety of packets loss concealment techniques have been proposed as solutions for the above problems in order to mitigate the effect of frame erasures, especially in the context of voice over packet networks (Kang and Kim 2011; Kheddar and Boudraa 2015; Kim et al. 2013; Kuo et al. 2013; López-Oller et al. 2014; Ma et al. 2014;

✉ Nadir Benamirouche
benam_nadir@yahoo.fr

Bachir Boudraa
bboudraa@usthb.dz

Domingo López-Oller
domingolopez@ugr.es

José L. Pérez-Córdoba
jlpc@ugr.es

¹ Laboratoire de Génie Electrique, Faculté de Technologie, Université de Bejaia, 06000 Bejaia, Algeria

² Faculty of Electronics and Computer Science, University of Sciences and Technology H. B., 16000 Algiers, Algeria

³ Department of Signal Theory, Networking and Communications, University of Granada, 18071 Granada, Spain

Mehran 2011; Merazka 2013; Miralavi et al. 2011; Oh et al. 2012; Park et al. 2010; Perkins et al. 2001; Toyoshima and Shimamura 2014). Traditionally these techniques have been classified on sender-driven and receiver-based methods (Perkins et al. 2001). In the former, many sender-based repair techniques have been proposed in the literature, including retransmission (Kuo et al. 2013; Oh et al. 2012), and interleaving packets (Kheddar and Boudraa 2015; Mehran 2011). In the latter, the receiver-based techniques are applied to recover lost packet at the decoding stage by means of extrapolation (Merazka 2013) or model-based speech estimations (Ma et al. 2014; Miralavi et al. 2011; Toyoshima and Shimamura 2014).

The robustness against frame erasures can also be significantly improved by adding redundancy at the packetization level to alleviate the problem of inter-frame dependency through Forward Error Correction (FEC) approaches (Al-Rousan and Nawasrah 2012; Andersen et al. 2002; Assem et al. 2013; Carmona et al. 2008; Casu et al. 2015; Gomez et al. 2010, 2011; Jalil et al. 2015; Liu et al. 2011; Merazka et al. 2014; Nagano and Ito 2013; Silveira and Silva 2012). Some of them, adapt FEC mechanisms for voice calls based on Reed Solomon codes (Assem et al. 2013; Silveira and Silva 2012) or parity codes (Casu et al. 2015; Merazka et al. 2014) to recover the lost packets. These mechanisms select the optimum codes from a family of codes and added as extra bit-rate to improve the conversational call quality. In Jalil et al. (2015) an improved conventionnel sender-driven method based on Automatic Repeat Request (ARQ) which deals with a FEC technique was proposed in order to reduce the number of retransmitted packet and to conceal the effects of packet losses.

In regards to the inter-frame dependency, most of recent speech transmission codecs are based on Code-Excited Linear Prediction (CELP) paradigm as TS; ITU Rec (1996), and they are essentially relying on the Long Term Prediction (LTP) to encode the current speech frame through the past frame parameters (Taleb 2011). Unfortunately, the LTP or Adaptive Codebook (ACB) used in the CELP-type codecs can dramatically reduce the speech quality in the presence of frame erasures (López-Oller et al. 2014). Since the speech parameters are not efficiently estimated by the embedded concealment technique, it is reported that this mismatch on the obtained excitation causes a desynchronization on the ACB which introduces an error propagation through the correctly received frames (Gomez et al. 2010). However, few works have been dedicated to reducing the error propagation caused by the ACB. Andersen et al. (2002) developed a frame independent-based speech codec, named Internet low bit-rate (iLBC) but at a cost of a considerable increase in bit-rate. A hybrid solution was presented in Carmona et al. (2008),

where a CELP-type codec is combined with the iLBC codec. The idea is to use iLBC to force an intra-frame coding for a given number of frames. Thus, the iLBC-coded frames can assure resynchronization in the case of packet loss. In the same context, a very attractive approach for decoder resynchronization was proposed in Gomez et al. (2011). This method is based on multipulse description of the previous excitation where additional information was sent for every subframe (SF). The main drawback of these methods is the introduced increase in the bit-rate. Obviously, sending redundancy information increases the bandwidth requirements and, therefore, the loss rate.

In this paper, we propose an efficient ACB resynchronization technique based on low complexity media-specific FEC optimization. The proposed method is based on a dynamically adding of FEC information to tackle the ACB desynchronization problem. Through this approach a significant improvement of the synthesis speech quality is achieved while the legacy bit-rate of the codec is kept unchanged with the advantage of low computational complexity.

The proposed method consists on replacing the ACB memory at the encoder side using a Pitch-Pulse Codebook (PPCB)-based approach to model the pitch-like contribution involved in building the CELP excitation (i.e., total excitation) signal for frame onset (voiced frames) determined under Zero Crossing Rate (ZCR) constraint (Jalil et al. 2013; Nath and Kalita 2014). Furthermore, the proposed method exploited the subframe correlation to limit the sending of the PPCB parameters only for the first two subframes (SF_0 and SF_1) of voiced frame, since the pitch component of the subsequent subframes (SF_2 and SF_3) can be estimated using the previous two subframes. Hence, at the decoder side when the previous frame is erased, the received PPCB parameters are exploited to generate an alternative excitation to replace the corrupted ACB excitation which contributes to form the total excitation. The resynchronization algorithm further exploits the formed total excitation to update the ACB memory. Then, the LTP contribution for the subsequent subframes can be regenerated using the resynchronized ACB memory. The proposed method can be implemented in virtually any CELP-type speech codec. The standard G.723.1 ITU Rec (1996) is used as a baseline codec.

This paper is organized as follows. In Sect. 2, we present the ACB resynchronization using the pitch pulse codebook-based approach and the applied algorithm for its optimization. Subsequently, in Sect. 3, we describe the experimental framework applied to generate lossy packet channels, the used speech database and objective quality measure to assess the quality of the proposed method. In Sect. 4, we discuss the effectiveness of the proposed method, where the obtained relevant results with the

proposed FEC method are shown. Finally, conclusions of this work are presented in Sect. 5.

2 Pitch pulse codebook-based approach for error propagation recovery

2.1 Excitation search in a CELP-type codec

We assume the well-known CELP-type speech production paradigm, where speech is generated by passing an excitation signal through a synthesis filter (Taleb 2011; Anselam and Pillai 2014). In this model, a speech signal, $s(n)$, is described by two components: firstly, a predictable signal that contains the vocal tract information, $\tilde{s}(n)$, as

$$\tilde{s}(n) = \sum_{j=1}^{\infty} a_j(n) \cdot s(n-j). \tag{1}$$

where $\tilde{s}(n)$, $s(n)$ and a_j , are the predicted speech signal, the original speech signal and the linear prediction coefficients respectively. Secondly, a residual signal that contains the excitation information is defined,

$$e(n) = s(n) - \tilde{s}(n). \tag{2}$$

Under this model, a segment of synthesized speech for each subframe is obtained by filtering an error signal (2), by means of a short-term linear prediction (LP) filter, $1/A(z)$. After removing the contribution of the LP filter memory, the new version of the error signal $\hat{e}(n)$ can be expressed as follows,

$$\begin{aligned} \hat{e}(n) &= x(n) - \hat{x}(n), \\ &= x(n) - \sum_{j=1}^{N-1} h(j) \cdot \hat{e}(n-j). \end{aligned} \tag{3}$$

where $x(n)$ indicates the target signal once the contribution of the LP filter memory has been removed, $\hat{x}(n)$ is the synthesized one, $h(n)$ the impulse response of the LP filter and N is the subframe length. Similar for most CELP-type codecs, the excitation signal consists again of two components, the LTP contribution which supplies the pitch-like component by placing a shifted repetition of past excitation samples, while the second excitation component is the Fixed Codebook (FCB) contribution also known as innovative vector $e_f(n)$, which is considered as a set of pulses with different positions. Formally, the excitation signal, $\hat{e}(n)$, is generally obtained as

$$\begin{aligned} \hat{e}(n) &= \sum_{j=-(l-1)/2}^{(l+1)/2} b(j) \cdot e(n - (T+j)) + g_f \cdot e_f(n). \\ &= e_a(n) + g_f \cdot e_f(n). \end{aligned} \tag{4}$$

where T , $b(j)$, l and g_f are pitch lag, LTP filter gains, LTP filter length and fixed vector gain respectively. The goal of the FCB contribution $e_f(n)$ is to model the residual signal remaining after removing the long-term redundancy, where the recursive part in (4) is generally known as ACB excitation, $e_a(n)$. The term adaptive comes from the fact that the codebook is constantly filled. The content of the ACB for the current subframe is simply the excitation of the previous subframes. Accordingly, the search for the best CELP excitation parameters is performed in the perceptually weighted domain using a perceptual filter, $W(z)$, allowing the search procedure to take into account psychoacoustic properties of human auditory perception. Figure 1 illustrates the analysis by synthesis mechanism followed in the CELP-type paradigm.

2.2 Subframe correlation impact on the pitch lag

To evaluate the impact of the subframe correlation on the pitch lag value, we investigated the first two subframes (SF_0 and SF_1) intervention for supplying the pitch-like component of subsequent subframes (SF_2 and SF_3). To this end, a study of the pitch-lag value occurrence has been performed for every coded subframe. To figure out this feature, Fig. 2 shows an histogram of pitch lag values occurred during all the used speech sequences. As we can see on this figure, most of pitch-lag values are concentrated between 20 and 120. In other words, most of pitch-like contribution for the last-two subframes (SF_2 and SF_3) depends only on the first-two subframes (SF_0 and SF_1). The exploitation of this feature offers for us the opportunity to insert the redundant information for only the first two subframes.

2.3 Voiced subframe classification

In the context of speech signal classification into voiced/unvoiced parts, a zero crossing in speech signal is used to occur if successive samples have different algebraic signs. The rate at which zero crossing occurs is a simple measure of the frequency content of a signal (Jalil et al. 2013; Nath and Kalita 2014). Figure 3 shows an illustration of zero crossing rate (bottom graph) for a sequence of speech signal (top graph).

Let consider the ZCR. It is a measure of number of times in a given subframe that the amplitude of the weighted speech signal $x_w(n)$, passes through a value of zero, such as

$$ZCR = \frac{1}{N} \sum_{n=0}^{N-1} |sng[x_w(n)] - sng[x_w(n-1)]|. \tag{5}$$

N is the number of subframe samples, where

Fig. 1 Analysis by synthesis coding

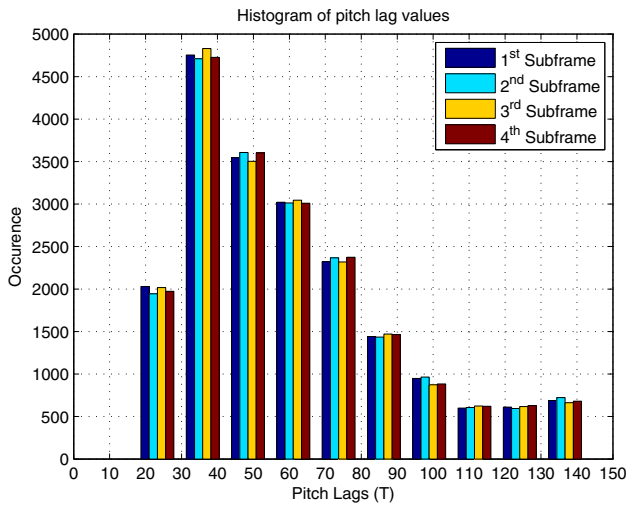
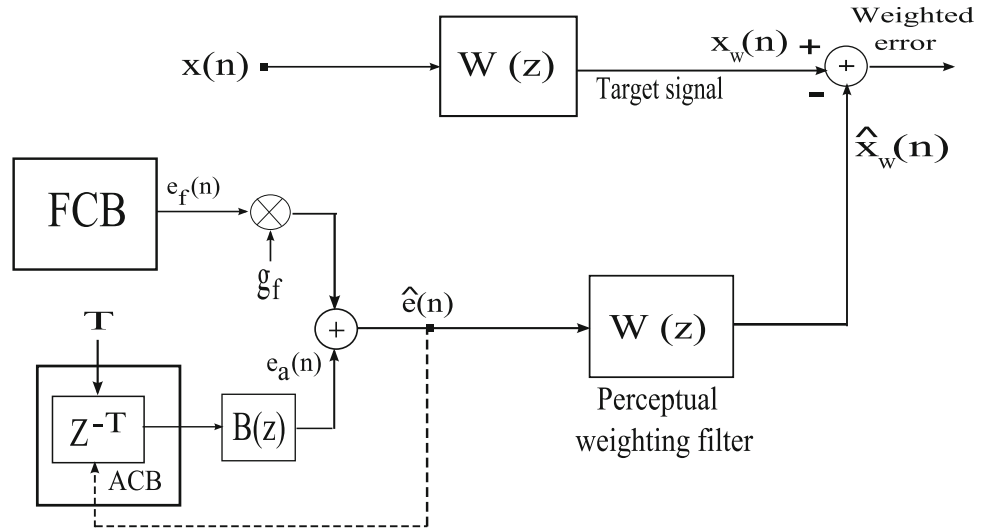


Fig. 2 Histogram of pitch lag values

$$\begin{aligned}
 \text{sgn}[x_w(n)] &= +1, & x_w(n) &\geq 0; \\
 &= -1, & x_w(n) &< 0.
 \end{aligned}
 \tag{6}$$

In fact, constraining the ZCR allows us to limit the sending of FEC parameters for only subframes (SFs) judged as voiced (i.e., important subframes). To evaluate the ZCR constraint influence, we measured the Perceptual Evaluation of Speech Quality (PESQ) scores of the synthesized speech as a function of three network conditions of 6, 10 and 13 % of packet loss rate, respectively. The results are given in Fig. 4. In regard to the obtained graphs in Fig. 4, we can notice remarkable PESQ score improvement around ZCR equal 0.35. However, the constraint/decision is applied as follows. The ZCR value of each subframe is compared to the set threshold $ZCR = 0.35$. If the value of ZCR is less than or equal to 0.35, implies the corresponding subframe is judged voiced, otherwise, the subframe is unvoiced.

2.4 Modification of the fixed codebook

Commonly, most of CELP-type codecs use a fixed codebook (FCB) defined as a set of pulses with different positions to model the remaining signal after removing LTP contribution. The modification of the FCB is done only when the LTP contribution is judged high relative the ZCR constraint measured in the perceptually weighted domain. In order to deal with the binary payload introduced by coding the FEC parameters during voiced subframes, the term $e_f(n)$ in (4) is modified. Thus, the number of FCB pulses, M , is reduced to $(M - \alpha)$ and the new expression of the fixed codebook excitation is

$$e_f(n) = \sum_{i=1}^{M-\alpha} g_f \cdot \delta(n - m_i) \tag{7}$$

where $(1 \leq \alpha \leq 3)$ is the number of the subtracted pulses from FCB codebook and M is a legacy number of FCB pulses which is greater than α . Table 1 lists the gained bits from FCB reduction with respect to the different values of reduced pulse α from the first two subframe (SF_0 and SF_1) contribution.

2.5 The search for the optimal pitch pulse

Figure 5 shows the modified encoder scheme and how the proposed PPCB-based approach fits within a typical CELP encoder. The proposed PPCB-based approach aims to achieve a fast ACB resynchronization and prevent the error propagation through the correctly received frames. To this end, we employed the ZCR as a constraint for voiced frame detection. When the imposed constraint is satisfied, a modification on the FCB is introduced by subtracting pulses from the first two subframe while the gained bits from this modification will be exploited by the FEC intervention. To optimize the FEC parameters (pitch pulse

Fig. 3 Zero crossing rate representation for a speech signal segment

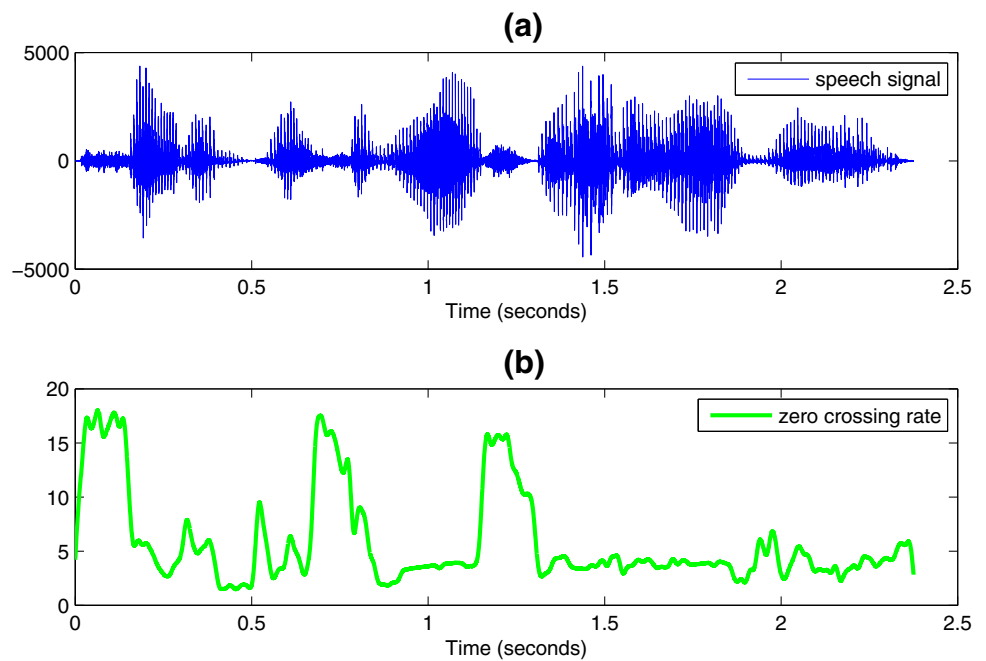


Fig. 4 PESQ scores in function of ZCR values with respect to different channel erasure conditions (a), (b) and (c) of 6, 10 and 13 % of loss rate, respectively

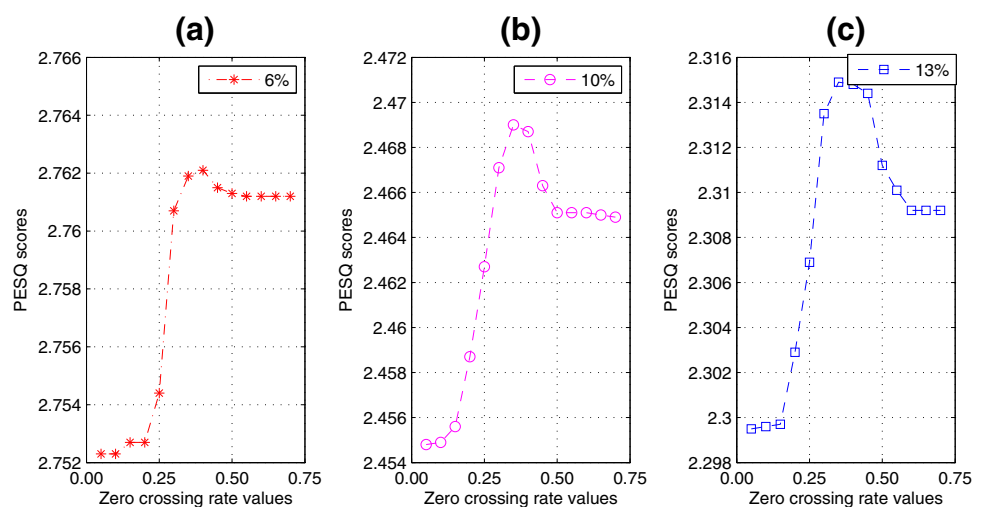


Table 1 Gained bits from the first two subframes (SF_0 and SF_1) with respect to the different cases of reduced-FCB pulse number ($M - \alpha$)

Reduced FCB ($M - \alpha$)	Number of subtracted pulses	Number of gained bits
($M - 1$)	2 Pulse/frame	27 bit/frame
($M - 2$)	4 Pulse/frame	32 bit/frame
($M - 3$)	6 Pulse/frame	38 bit/frame

position and gain), we applied the Multipulse Maximum likelihood Quantization (MP-MLQ) algorithm (ITU Rec 1996). Throughout the following steps, the search for the PPCB parameters is described in details.

2.5.1 Pulse position and amplitude optimization

The optimal pitch pulse position and the optimal gain are those that minimize the quadratic error in the weighted

domain between a single-pulse excitation named, PPCB excitation, and the CELP excitation signal, $\hat{e}(n)$. Since the pitch pulse search is a subframe-based optimization, the pulse position must be placed on one of two grids: even positions (Grid 0) or odd positions (Grid 1) given in Table 2.

It shall be noticed that one bit is used to specify which of the grids is used. As it can be seen from this table, we dispose of 60 pulse positions in which to place the best pulse position. The PPCB pulse gain for each subframe can

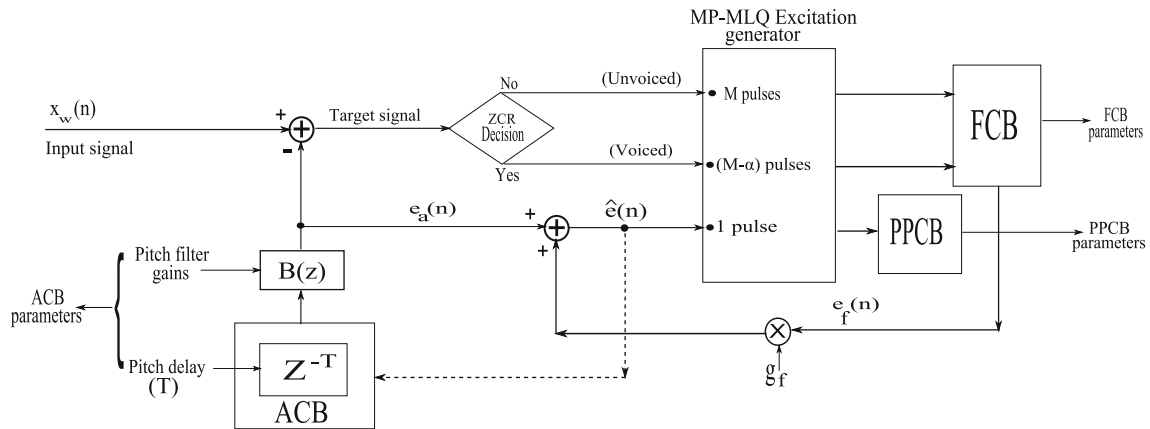


Fig. 5 Block diagram of the modified encoder

take different amplitudes which are selected from the 24 values of FCB gain table, and the pulse sign is specified separately with one bit.

Let consider ϕ_{eh} the cross-correlation between the weighted CELP excitation $\hat{e}(n)$ and the weighted impulse response of the synthesis filter $h_w(n)$, so that

$$\begin{aligned} \phi_{eh}[m] &= \sum_{n=0}^{N-1} \hat{e}(n) \cdot h_w(n - m), \\ &= \sum_{n=m}^{N-1} \hat{e}(n) \cdot h_w(n - m). \end{aligned} \tag{8}$$

Similarly the autocorrelation of the weighted impulse response, ϕ_{hh} , is given by

$$\begin{aligned} \phi_{hh}[m] &= \sum_{n=0}^{N-1} h_w(n) \cdot h_w(n - m), \\ &= \sum_{n=m}^{N-1} h_w(n) \cdot h_w(n - m). \end{aligned} \tag{9}$$

Hence, the best possible pulse position m of amplitude g_m can be found once the quadratic error between this single pulse signal and the CELP excitation signal, $\hat{e}_w(n)$, is minimized. The quadratic error in the weighted domain, Δ_w , is defined as

$$\begin{aligned} \Delta_w &= \sum_{n=0}^{N-1} (h_w(n) * \hat{e}(n) - g_m \delta(n - m))^2, \\ &= \sum_{n=0}^{N-1} (\hat{e}_w(n) - g_m h_w(n - m))^2, \\ &= E_e - 2 \cdot g_m \cdot \phi_{eh}[m] + g_m^2 \cdot \phi_{hh}[0]. \end{aligned} \tag{10}$$

where, E_e is the excitation signal energy thus, the weighted quadratic error in (10) is minimized with respect to the gain g_m , by solving the partial derivative $\frac{\partial \Delta_w}{\partial g_m} = 0$, that is

$$\frac{\partial \Delta_w}{\partial g_m} = -2 \cdot \phi_{eh}[m] + 2 \cdot g_m \cdot \phi_{hh}[0] = 0. \tag{11}$$

Hence, the optimal gain is given by

$$g_{opt} = \frac{\phi_{eh}[m]}{\phi_{hh}[0]}. \tag{12}$$

The choice of the optimal quantized gain g_m is performed over a fixed set of amplitude codebook, but allowing g_m to take on either sign. To reduce the search complexity, a quantized estimate of the gain is used and the searching procedure is done over gain amplitudes codebook nearby the estimated gain. If the gain which minimizes (10) is g_{opt} , the error using another gain value can be expressed as

$$\Delta_w = E_e - 2 \cdot g_m \cdot \phi_{eh}[m] + (g_m - g_{opt})^2 \cdot \phi_{hh}[0]. \tag{13}$$

The quantized gain that minimizes the mean square error is that value closest to g_{opt} .

2.5.2 Optimization of pulse amplitude

Hence, the best PPCB pulse position which gives the maximum reduction in squared error is found as

$$\begin{aligned} m_{opt} &= \max_m (2 \cdot g_{opt} \cdot \phi_{eh}[m] - g_{opt}^2 \cdot \phi_{hh}[0]), \\ &= \max_m \left(\frac{\phi_{eh}[m]}{\phi_{hh}[0]} \right), \\ &= \arg \max_m [\phi_{eh}[m]]. \end{aligned} \tag{14}$$

Table 2 Pulse positions

Grid 0										Grid 1										
0	2	4	6	8	10	12	14	16	18	—	1	3	5	7	9	11	13	15	17	19
20	22	24	26	28	30	32	34	36	38	—	21	23	25	27	29	31	33	35	37	39
40	42	44	46	48	50	52	54	56	58	—	41	43	45	47	49	51	53	55	57	59

Once the best position m_{opt} is found, the optimal gain g_{opt} for that pulse is given by (12). Then the quantized value of gain \hat{g}_m nearest g_{opt} is found. For MP-MLQ in ITU Rec (1996), the gain codebook contains $C_g[i]$ elements with $0 \leq i \leq 24$. Hence, the index of the quantized amplitude is found as

$$i_g = \min_i (|g_{opt}| - C_g[i]), \tag{15}$$

$$= \arg \min_i (C_g[i] \phi_{hh}[0] - |\phi_{eh}[m]|).$$

The gain search that is used for every pulse is limited to quantized gain values near $C_g[i]$, [see algorithm in ITU Rec (1996)]. And then, the best pulse position will be found for this gain value.

2.5.3 Pulse position search

Given the estimated quantized gain, the error given in (10) becomes,

$$\Delta_w = E_e \mp 2 \cdot C_g[i] \cdot \phi_{eh}[m] + C_g^2[i] \cdot \phi_{hh}[0]. \tag{16}$$

where the upper sign is used if the pulse is positive and the lower sign is used if the pulse is negative. The position that gives the lowest squared error is

$$m_{opt} = \arg \max_m (\phi_{eh}[m]). \tag{17}$$

The sign of the pulse is determined by the sign of $\phi_{eh}[m_{opt}]$, and

$$g_{m_{opt}} = \text{sign}(\phi_{eh}[m_{opt}]) \cdot C_g[i]. \tag{18}$$

The PPCB memory is the excitation that can be used to build the excitation signal for the current frame (or subframe), which can be introduced as FEC information for the purpose of resynchronization. However, the PPCB parameters which will be sent are referring to the optimized pulse positions with its respective gains in the PPCB memory. Note that the vector length corresponding to the PPCB memory is the same as used for the ACB memory. Since the maximum allowable pitch lag value, T_{max} is greater than the subframe length, N , hence two bits are used to specify the organization of the optimized pulses in the PPCB memory.

2.5.4 Multipulse-based PPCB excitation

To assess the performance of the PPCB-based approach according to the included pulses for modeling the LTP contribution in CELP excitation $\hat{e}(n)$, we performed a comparison of reconstructed PPCB excitation signals of three, two and one pulse, respectively. Thus, the new expression of the weighted quadratic error in (10)

under multipulse-based PPCB excitation can be formulated as,

$$\Delta_w = \sum_{n=0}^{N-1} \left(\hat{e}_w(n) - \sum_{k=1}^3 g_{m_k} h_w(n - m_k) \right)^2. \tag{19}$$

where, k , is the index of each pulse to be set, ($k = 1, 2, 3$), while all the pulses have the same amplitude with different signs. Once the position and gain of the first pulse have been found, the effect of that pulse can be subtracted from the excitation signal $\hat{e}(n)$,

$$\hat{e}'(n) = \hat{e}(n) - g_{m_{opt}} \cdot \delta(n - m_{opt}) * h_w(n), \tag{20}$$

$$= \hat{e}(n) - g_{m_{opt}} \cdot h_w(n - m_{opt}).$$

Thus, $\hat{e}'(n)$ is the new CELP excitation signal after removing the first optimal pulse contribution thus, the cross-correlation ϕ_{eh} in (8) can be updated as,

$$\phi_{e'h}[m] = \sum_{n=m}^{N-1} \hat{e}'(n) \cdot h_w(n - m),$$

$$= \phi_{eh}[n] - g_{m_{opt}} \cdot \sum_{n=m}^{N-1} h_w(n - m) \cdot h_w(n - m_{opt}),$$

$$= \phi_{eh}[n] - g_{m_{opt}} \cdot \phi_{hh}[\lceil n - m_{opt} \rceil]. \tag{21}$$

With the updated cross-correlation, the next pulse can be placed. As each pulse is placed, its corresponding position is marked as occupied to prevent a subsequent pulse being placed in the same position.

In regards to the obtained results, Fig. 6 shows the excitation signals of the correct ACB excitation given by the standard in (a), compared with the PPCB excitation signals relative to the included pulses. As can be observed on (b), some secondary pulses are appeared with high amplitudes relative to the corresponding ones in (a). Accordingly, the apparition of such pulses may not be reliable due to the risk of introducing artifacts on the synthesized speech signal. Similarly, in (c), the use of pair of pulses provides a decrease in pitch pulse amplitude and introduces a secondary pulse around this latter with an amplitude almost equal to the principal pulse. In contrast, in (d), when a single pulse is used, we noticed an increase in PPCB excitation amplitudes which is approaching the ACB excitation amplitudes and absence of secondary pulses with high amplitudes. Otherwise, these undesirable effects may arise from the MP-MLQ algorithm due from using the same amplitude for all the optimized pulses. In spite of this, in our proposed method, we use a single pulse to model only the principal pitch pulse to avoid introducing changes on its contour.

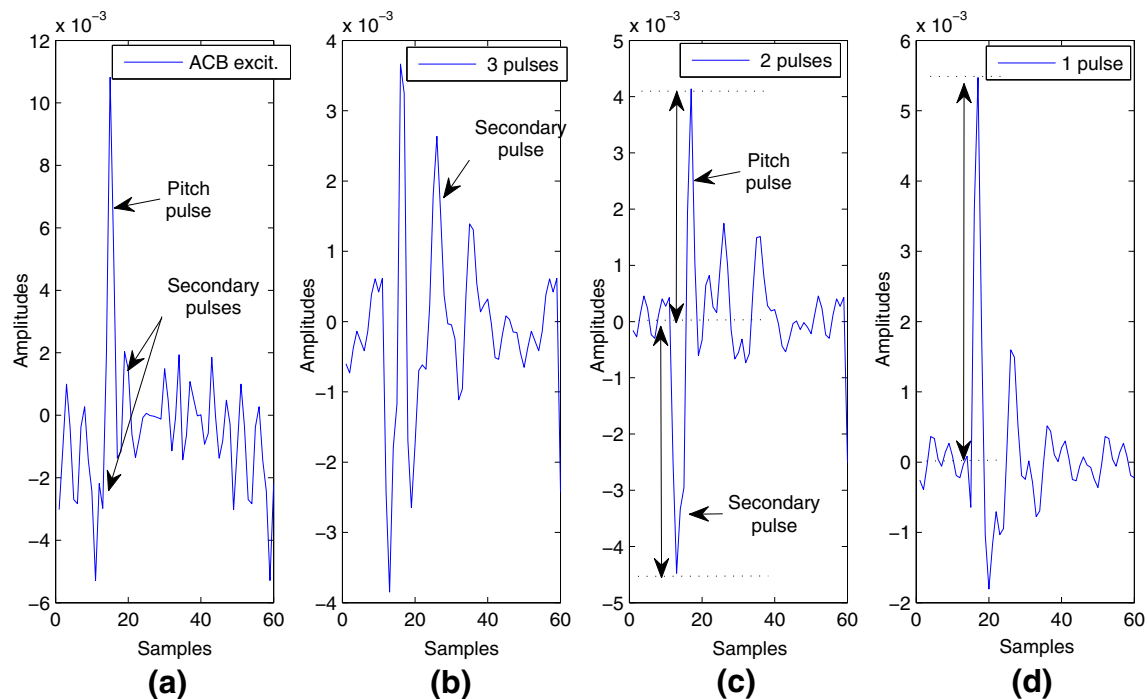


Fig. 6 Comparison of the excitation signals during the coding procedure. (a) correct ACB excitation given by the standard codec; (b), (c) and (d) Excitation signals of PPCB intervention including 3, 2 and 1 pulse, respectively

2.6 Description of the ACB resynchronization procedure

We assume that the received ACB parameters: pitch delay, T , and gains of the pitch filter, $B(z)$, are correct since they are based on good ACB memory (i.e., past excitation samples). Thus, after a frame erasure, the resynchronization algorithm uses the received pitch lag to localise the concerned pulse in PPCB memory while the pitch filter gains are used to shape the contour around this principal pitch pulse.

In Fig. 7, we summarize the resynchronization procedure of the ACB memory at the decoder side after a frame erasure. At this stage, the generated excitation using the received PPCB parameters (pulse position and gain), shifted by the pitch lag, T , filtered by the LP synthesis filter impulse response, $H(z)$, then scaled by the pitch filter gains of, $B(z)$, and added to the FCB excitation is used to replace the corrupted ACB excitation while the total excitation $\hat{e}(n)$ is immediately used to update the ACB memory.

2.6.1 Involving the ACB parameters to shape the pitch pulse contour

An experiment of resynchronization stage is performed to demonstrate the utility of involving the pitch lag, T , to localize the optimal pulse into the PPCB memory and the role of the pitch filter gains of $B(z)$ to shape the pitch pulse

contour. Note that this experiment is realized under error-free conditions. Figures 8 and 9, show a comparison between correct ACB and PPCB excitation signals for the first two subframes (SF_0 and SF_1).

Through both Figs. 8 and 9, we can see the pulse position similarities between the optimized pulses in PPCB memory (bottom graphs (c) and (g) on the presented figures) and the pitch pulses in the ACB memory (top graphs in (a) and (e) on both figures). Moreover, these similarities are saved even if the ACB parameters have been involved to generate the PPCB-based excitation (bottom graphs (d) and (h) on the presented figures respectively). Note that the main feature of the pitch-like contribution is not limited to the principal pitch pulse but also include the surrounding secondary pulses.

2.6.2 Updating the ACB memory during the resynchronization

When the pitch pulse is resynchronized into the first two subframe excitations, it is more convenient to use every built CELP excitation from the resynchronization procedure to update the corrupted ACB memory then the LTP contribution for the subsequent subframes will be generated from the updated ACB memory. To illustrate the effectiveness of the followed approach through sending the PPCB parameters for the first two subframes only (with updating ACB memory) instead of four subframes (without

Fig. 7 Block diagram of the modified decoder

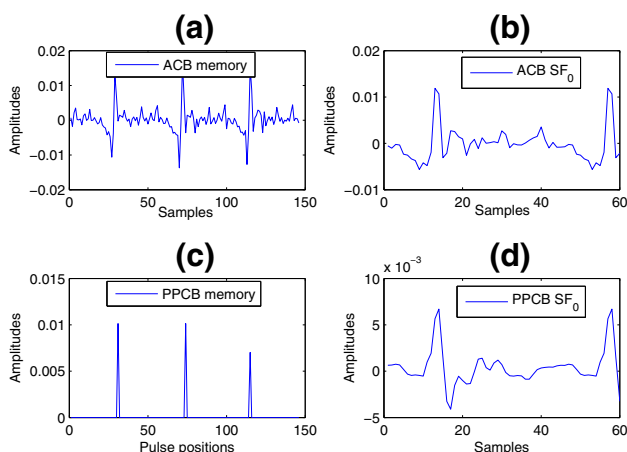
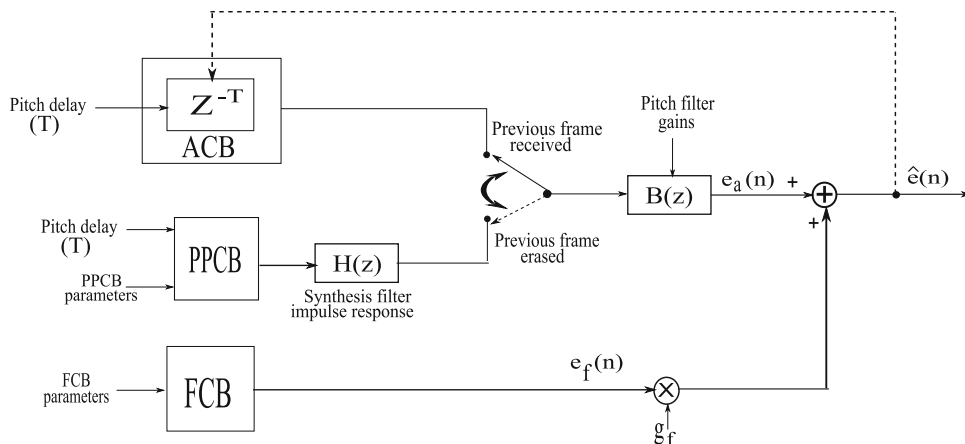


Fig. 8 comparison between correct ACB and PPCB excitation signals of the first subframe (SF_0) involving same ACB parameters (pitch lag $T = 44$, and Pitch filter gains of $B(z) = [0.2068 \ 0.0958 \ 0.6172 \ 0.2946 \ -0.2676]$)

updating ACB memory), we also performed a comparison between correct ACB excitations and excitations built from two PPCB memory cases, without updating and with updating, respectively, corresponding to SF_2 and SF_3 . Note that these introduced cases are not considered as part of the resynchronization algorithm but they have been included only to illustrate both cases of PPCB intervention relative to 4 SFs/frame and 2 SFs/frame, respectively. However, the two cases are realized as follows.

- Case 1: The LTP contribution (or PPCB excitation) for SF_2 and SF_3 , is also obtained from the optimized pitch pulse involving the ACB parameters, and the resulting total excitation, $\hat{e}(n)$, is not considered to update the PPCB memory.
- Case 2: The total excitation, $\hat{e}(n)$, built from the PPCB intervention during the resynchronization of SF_0 and SF_1 , is constantly used to update the PPCB memory, while the LTP contribution (PPCB excitation) for the

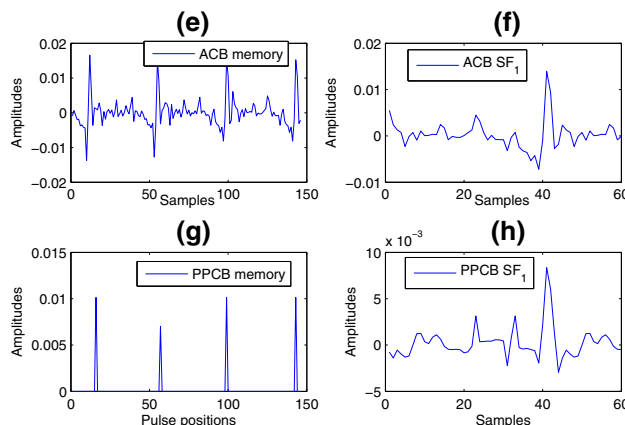


Fig. 9 Comparison between correct ACB and PPCB excitation signals of the second subframe (SF_1) involving same ACB parameters (pitch lag $T = 43$, and Pitch filter gains of $B(z) = [-0.0068 \ 0.0157 \ 0.0421 \ 0.8882 \ 0.0086]$)

last two subframes is generated by exploiting the updated PPCB memory.

Figure 10 shows the considered two cases relative to the third subframe (SF_2). As can be seen from the obtained plots, the PPCB-based excitation in (d) of case 2, shows more waveform similarities with the good ACB excitation in (b) compared with that obtained in (d) of case 1.

Similarly, Fig. 11 shows the fourth subframe (SF_3) under both of cases. As expected, in (d) of case 2, we noticed more improvement of the PPCB excitation waveform where the obtained excitation matches better to the good ACB excitation in (b) compared to that obtained in (d) of case 1.

3 Experimental framework

In this section, the applied experimental framework to generate lossy packet channels, the speech database considered during the evaluation and the used objective quality

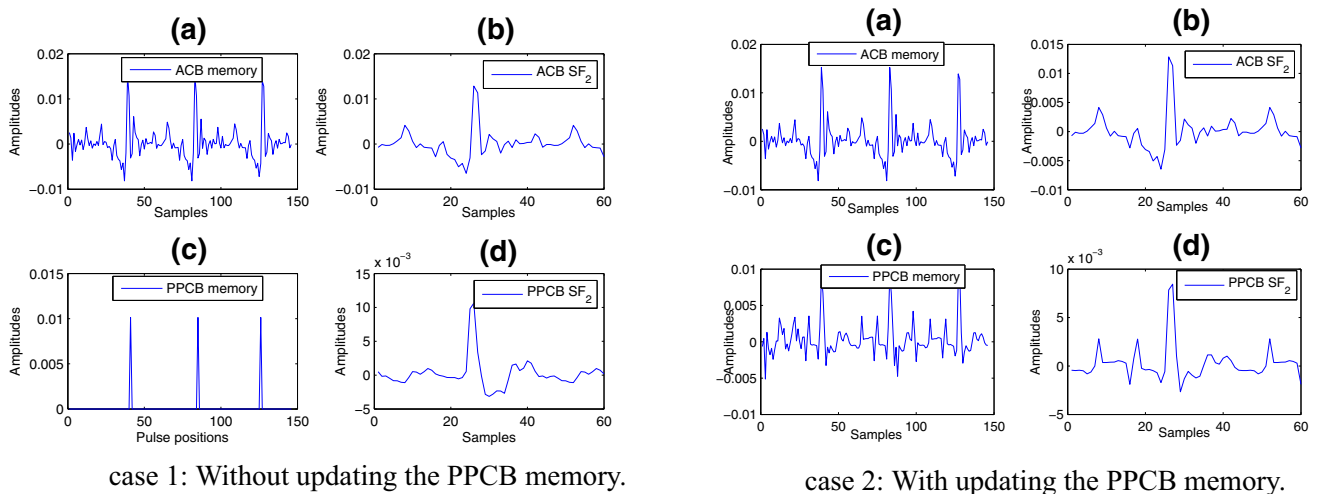


Fig. 10 comparison between ACB and PPCB excitation signals of the third subframe (SF_2) involving same ACB parameters (pitch lag $T = 44$, and Pitch filter gains of $B(z) = [-0.0068 \ 0.0157 \ 0.0421 \ 0.8882 \ 0.0086]$)

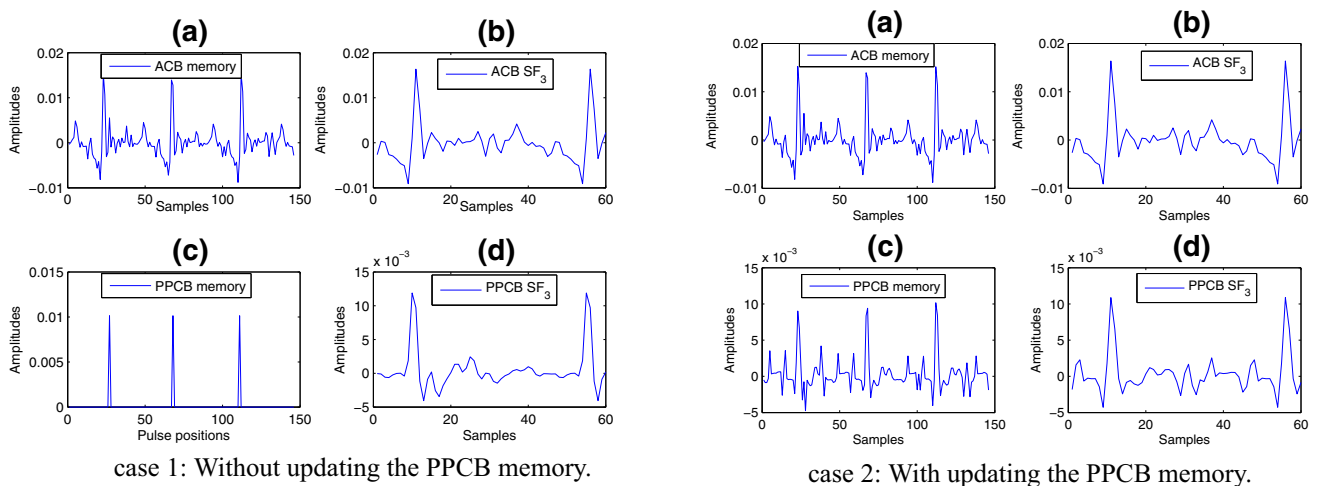


Fig. 11 comparison between ACB and PPCB excitation signals of the fourth subframe (SF_3) involving same ACB parameters (pitch lag $T = 45$, and Pitch filter gains of $B(z) = [0.0581 \ 0.2232 \ 0.9009 \ -0.1403 \ 0.0445]$)

measure to assess the quality of the proposed method is described.

3.1 Objective quality evaluation

For performance evaluation, we have considered an objective test performed by means of the ITU Perceptual Evaluation of Speech Quality standard (PESQ) (Perceptual Evaluation of Speech Quality 2001). In order to provide an objective quality measure, PESQ is applied over a subset of the well-known TIMIT database which contains wideband recordings from 630 speakers of eight major dialects of American English (Lamel et al. 1986; Garofolo 0000). To this end, testing and training utterances from TIMIT database are down-sampled to 8 kHz

and their lengths artificially extended to approximately 14 s. A total of 80 utterances are used from both considered sets and for each utterance, the PESQ algorithm provides a score within a range from -0.5 (bad) to 4.5 (excellent). In order to obtain an overall score for each channel condition, the score of each sentence is weighted by its relative length.

3.2 Transmission and channel model

During these simulations, packet loss rates of 6, 8, 10, 13, 16, 18, 20, 21 and 23 % were generated by the Gilbert-Elliot model defined in Jiang and Schulzrinne (2000). Under these packet loss channels, the burst of consecutive frame losses is varying from 1 up to 3 frames.

4 Experimental results

Table 3, lists the obtained PESQ results corresponding to the standard G.723.1 under the modified-FCB cases ($M - \alpha$, $\alpha = 0, 1, 2$ and 3), with respect to the packet loss conditions. The results show a very small variation of PESQ scores although the value of α is increased from 0 up to 3. The obtained results allow us to prove again that under the imposed constraint, the FCB pulse reduction does not really affect the quality of the synthesized speech signal. Thus, this feature offers for us the opportunity to exploit the gained bits from the FCB modification to code the PPCB parameters.

Table 4, shows the obtained PESQ results from the PPCB-based approach intervention relative to the number of included pulses for modeling the pitch-like contribution which resides in the CELP excitation. As it is mentioned previously, increasing the number of pulse does not really improve the efficiency of the proposed method. This effect can be assigned to the pulse optimization procedure which uses an optimal gain for all the included pulses. In fact, the appearance of secondary pulses around the principal pitch pulse may affect its contour and introduce artifacts on the synthesized speech signal. As can be seen on Table 4, despite of some fluctuations in PESQ scores, globally the average of PESQ values remain the same for one and two pulse cases except a very slight decrease in case of three pulses which can be attributed to the speech distortion introduced on the synthesized signal.

To assess the performance of our methodology for ACB resynchronization, we carried out an experiment where we compared our own proposed method with another approach in the literature. Namely, we consider the method proposed by Gomez et al. (2011) for reducing the error propagation caused by the ACB. Throughout this comparison, we refer to that approach as the reference method. In the reference method the resynchronization pulse is optimized for every subframe either voiced or unvoiced. It is important to note that during the simulation of the reference method, the optimized pulse position and gain are unquantized to deal with the reference conditions reported in Gomez et al. (2011) and further may help in forming a valid comparison.

Therefore, Fig. 12 shows the obtained PESQ results related to the performance of the PPCB-based method

relative to those obtained from legacy standard G723.1 codec (ITU Rec 1996). In addition, the results from a complete LTP restoration (Restore memory) and from the reference method are also shown in this figure. As can be seen from this figure, the PPCB method offers a significant improvement of the PESQ quality while it is approaching the quality of complete memory restoration compared with the legacy G723.1 codec.

Furthermore, we can also notice a slight improvement of the PESQ quality is achieved by the proposed method relative to the reference method. Meanwhile, the improvement obtained by the PPCB approach can be justified, on the one hand, by the imposed constraint which takes into account only the voiced frames to replace the ACB excitation after a frame erasure. On the other hand, the use of the updated ACB memory just after the first two subframe resynchrony to generate the ACB contribution for the subsequent subframes allows the decoder to prevent more efficiently the error propagation.

To illustrate the quality difference between both considered cases of PPCB intervention (2 SFs/frame and 4 SFs/frame), mentioned previously. Table 5 shows the obtained PESQ results at different erasure channel conditions with its respective average values. As expected, a slight improvement in PESQ quality is recorded when combining the ZCR constraint with the ACB memory update during the first two subframes resynchronization. However, this combination offers a better decoder resynchronization over all the subframes following the erasure.

As can be seen from Fig. 13, the synthesized signal corresponding to the supplied decoder with PPCB-based approach given in (c) shows much waveform (pitch pulses) similarities than the signal obtained by the standard in (b). For comparison, the error signals obtained with the standard encoder only and with the PPCB intervention are plotted in (d) and (e), respectively. Both error signals are computed with the original signal as a reference. It appears clearly that the error is much reduced with the PPCB-based approach intervention than with the standard decoder.

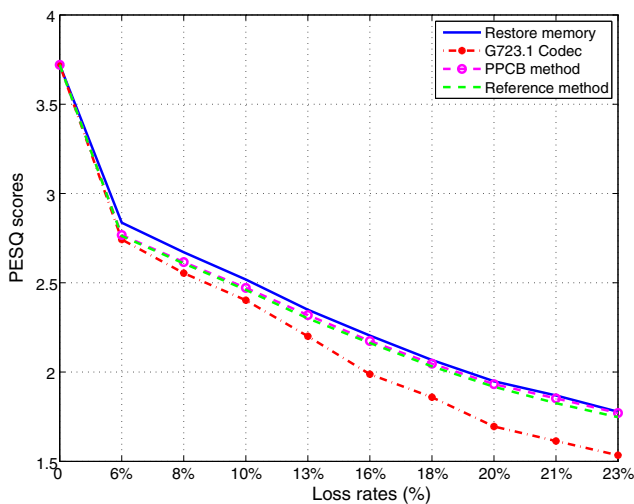
Table 6, lists a comparison of bit-rates resulting from the proposed PPCB-based approach and the reference method relative to the number of included pulses. The quantization of the pulse position used in the proposed method is based

Table 3 The obtained PESQ Results of the standard G.723.1 under the ZCR constraint relative to the number of subtracted pulses α from the M FCB pulses at multiple packet loss ratio conditions

Subtracted pulse	Packet loss ratio									
	6 %	8 %	10 %	13 %	16 %	18 %	20 %	21 %	23 %	
$\alpha = 0$	2.7421	2.5537	2.4022	2.2004	1.9880	1.8592	1.6954	1.6138	1.5343	
$\alpha = 1$	2.7400	2.5518	2.4004	2.1983	1.9859	1.8571	1.6933	1.6117	1.5320	
$\alpha = 2$	2.7381	2.5499	2.3983	2.1964	1.9837	1.8552	1.6916	1.6099	1.5301	
$\alpha = 3$	2.7325	2.5437	2.3921	2.1902	1.9878	1.8504	1.6861	1.6076	1.5282	

Table 4 PESQ Results obtained by PPCB-based approach relative to the number of used pulse to model the pitch pulse at multiple loss rate conditions

Number of pulse	Packet loss ratio									Average
	6 %	8 %	10 %	13 %	16 %	18 %	20 %	21 %	23 %	
1 pulse	2.7677	2.6162	2.4711	2.3191	2.1742	2.0472	1.9320	1.8528	1.7705	2.2168
2 pulses	2.7677	2.6164	2.4711	2.3192	2.1743	2.0471	1.9320	1.8528	1.7703	2.2168
3 pulses	2.7676	2.6163	2.4709	2.3189	2.1741	2.0472	1.9317	1.8526	1.7695	2.2165

**Fig. 12** PESQ scores obtained over 15 min of speech from test TIMIT database, for G723.1 codec relative to PPCB method intervention compared with the reference method and a complete ACB memory restoration under different lossy channel conditions

on the combinatorial coding (Blake and Mullin 2014). Thus, the total required bits for coding a single pulse is 11 bits, while the coding of two and three pulses requires 19 bits and 25 bits, respectively.

Accordingly, Table 6 shows that the proposed PPCB method does not include any increase into the legacy bit-rate either for one pulse or two pulses. Despite the efficiency of the PPCB approach is essentially based on a single pulse but the modification of the FCB under the ZCR constraint offers the opportunity to add more than one pulse without effecting the legacy bit-rate. In return, for the reference method, a small increase in bit-rate is needed in case of one pulse whereas when two pulses are used, the bit-rate is doubled. Actually, the resulting bit-rates shown

in Table 6 confirm the benefit of our proposed method in its tradeoff between constraining the PPCB intervention and exploiting the subframe correlations to keep the legacy bit-rate of the codec unchanged.

5 Conclusions

We have presented a low complexity FEC method to resynchronize the pitch pulse in CELP codec to prevent the error propagation issue after a frame erasure. The proposed method exploits the CELP-type paradigm properties and takes benefit from the high temporal subframe correlations to tackle the problem of inter-frame dependency. The efficiency of the proposed method resides in limiting the FEC intervention only when the LTP contribution is judged high relative to the ZCR constraint. Thus, the combination between the subframe correlations and the ZCR constraint offers the opportunity to keep the legacy bit-rate of the codec unchanged. Likewise, the proposed method applies the MP-MLQ algorithm to optimize the single pulse amplitude and position which models the pitch pulse within CELP excitation signal. Therefore, at the decoder side when the previous frame is erased, this optimized pulse is supposed to be only filtered by the synthesis filter response $H(z)$, shifted by the pitch lag T , scaled by pitch filter gains of $B(z)$ and added to the reduced fixed vector in order to shape the total CELP excitation signal. However, the resynchronization procedure is performed with respect to the first two subframes (SF_0 and SF_1), since the LTP parameters of the subsequent subframes (SF_2 and SF_3) can be easily predicted using the first two subframes. Accordingly, the aim behind the proposed method is to offer a low computational complexity resynchronization procedure and

Table 5 The obtained PESQ Results of PPCB-based approach intervention according to the considered subframes, 2 SFs/frame or 4 SFs/frame, at multiple packet loss rate conditions

Considered SFs	Packet loss ratio									Average
	6 %	8 %	10 %	13 %	16 %	18 %	20 %	21 %	23 %	
2 SFs/frame	2.7677	2.6162	2.4711	2.3191	2.1742	2.0472	1.9320	1.8528	1.7705	2.2168
4 SFs/frame	2.7633	2.6109	2.4657	2.3121	2.1792	2.0429	1.9303	1.8485	1.7656	2.2132

Fig. 13 Example of comparison of the speech signals in the synthesis domain. (a) The original speech signal; (b) Speech signal obtained by the standard G.723.1 codec; (c) Speech signal obtained when using the PPCB intervention; (d) and (e) Error signals for the standard and modified decoders

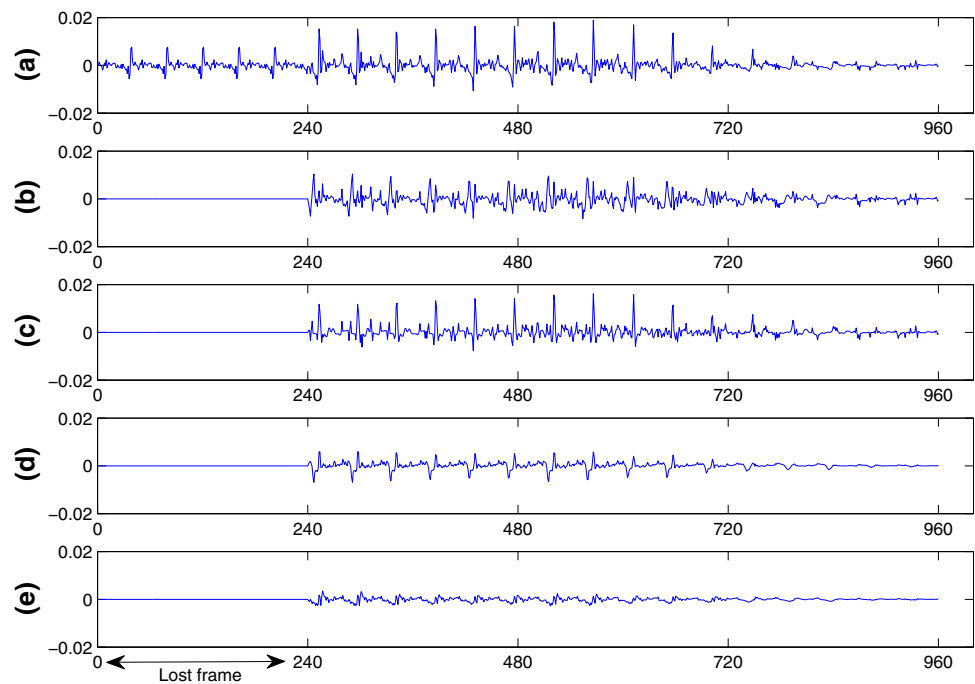


Table 6 Comparison of introduced bit-rate on G723.1 codec relative to the introduced pulses for both considered methods, the proposed PPCB method and reference method, respectively

Method	Number of pulse	Required bit/frame	Added bit-rate
PPCB method	1	∅	∅
	2	∅	∅
Reference method	1	16	0.52 kbit/s
	2	32	1.04 kbit/s

without any increase into the legacy bit-rate. Furthermore, the objective quality tests under channel erasure conditions have shown the suitability of the proposed method which deals with the interoperability of the used standard. The ITU PESQ algorithm which was used to provide objective quality scores revealed that our technique clearly outperforms the standard G723.1 codec and improves its robustness against error propagation effects. Finally, the speech quality evaluation confirmed that the PPCB based approach achieves a significant increase in PESQ quality which is approaching a complete ACB memory restoration. However, this method could be extended to other CELP-type codec which will be considered in future work. Also, we think that other approach may benefit from this method and can contribute to a better robustness against error propagation.

References

- Al-Rousan, M., & Nawasrah, A. (2012). Adaptive FEC technique for multimedia applications over the internet. *Journal of Emerging Technologies in Web Intelligence*, 4(2), 142–147.
- Andersen, S. V., Kleijn, W. B., Hagen, R., Linden, J., Murthi, M. N., & Skoglund, J. (2002). iLBCA linear predictive coder with robustness to packet losses. *Proceedings of IEEE Speech Coding Workshop, Tsukuba* (pp. 23–25).
- Anselm, A. S., & Pillai, S. S. (2014). Performance evaluation of code excited linear prediction speech coders at various bit rates. In *IEEE 2014 International Conference on Computation of Power, Energy, Information and Communication (ICCEPIC)* (pp. 93–98).
- Assem, H., Malone, D., Dunne, J., & O’Sullivan, P. (2013). A new adaptive redundancy control algorithm for VoIP applications. In *IEEE Global Communications Conference (GLOBECOM)* (pp. 1323–1328).
- Bhebbhe, L., & Parkkali, R. (2011). VoIP performance over HSPA with different VoIP clients. *Wireless Personal Communications*, 58(3), 613–626.
- Blake, I. F., & Mullin, R. C. (2014). *An introduction to algebraic and combinatorial coding theory*. London: Academic Press.
- Carmona, J. L., Pérez-Córdoba, J. L., Peinado, A. M., Gomez, A. M., & Gonzalez, J. A. (2008). A scalable coding scheme based on interframe dependency limitation. *IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 4805–4808).
- Casu, F., Cabrera, J., Jaureguizar, F., & Garca, N. (2015). A protection scheme for multimedia packet streams in bursty packet loss networks based on small block low-density parity-check codes. *EURASIP Journal on Wireless Communications and Networking*, 1, 1.

- Chandra, M., & Ray, M. (2015). Comparative study of PCM, LPC, and CELP speech coders used for VoIP applications. *Intelligent computing, communication and devices* (pp. 579–587). India: Springer.
- Garofolo, J. S. The Structure and Format of the DARPA TIMIT™, CD-ROM Prototype, Documentation of DARPA TIMIT.
- Gomez, A. M., Carmona, J. L., Peinado, A., & Sanchez, V. (2010). A multipulse-based forward error correction technique for robust CELP-coded speech transmission over erasure channels. *IEEE Transaction on Audio, Speech, Language Process*, 18(6), 1258–1268.
- Gomez, A. M., Carmona, J. L., Peinado, A., & Sanchez, V. (2011). One-pulse fec coding for robust celp-coded speech transmission over erasure channels. *IEEE Transactions on Multimedia*, 13(5), 894–904.
- Gupta, V., Dharmaraja, S., & Arunachalam, V. (2015). Stochastic modeling for delay analysis of a VoIP network. *Annals of Operations Research*, 233(1), 171–180.
- ITU Rec., (1996). G.723.1, Dual rate speech coder for multimedia communication transmitting at 5.3kbit/s and 6.3kbit/s.
- Jalil, S., Abbad, M., & El Azouzi, R. (2015). Hybrid FEC/ARQ schemes for real-time traffic in wireless networks. In: *2015 International Conference on Wireless Networks and Mobile Communications (WINCOM)* (pp. 1–6).
- Jalil, M., Butt, F. A., & Malik, A. (2013). Short-time energy, magnitude, zero crossing rate and autocorrelation measurement for discriminating voiced and unvoiced segments of speech signals. In *IEEE International Conference In Technological Advances in Electrical, Electronics and Computer Engineering (TAECE)* (pp. 208–212).
- Jiang, W., & Schulzrinne, H. (2000). Modeling of packet loss and delay and their effect on real-time multimedia service quality. In *Proceedings of NOSSDAV 2000*.
- Kang, J. A., & Kim, H. K. (2011). An adaptive packet loss recovery method based on real-time speech quality assessment and redundant speech transmission. *International Journal of Innovative Computing, Information and Control*, 7(12), 6773–6783.
- Kheddar, H., & Boudraa, B. (2015). Implementation of interleaving methods on MELP 2.4 coder to reduce packet loss in the Voice over IP (VoIP) transmission. *International Journal of Engineering Research and Applications*, 5(3), 1–4.
- Kim, B. H., Kim, H. G., Jeong, J., & Kim, J. Y. (2013). VoIP receiver-based adaptive playout scheduling and packet loss concealment technique. *IEEE Transactions on Consumer Electronics*, 59(1), 250–258.
- Kuo, C. F., Tseng, H. W., & Pang, A. C. (2013). A fragment-based retransmission scheme with quality-of-service considerations for wireless networks. *Wireless Communications and Mobile Computing*, 13(16), 1450–1463.
- Lamel, L., Kassel, R., & Seneff, S. (1986). Speech database development: Design and analysis of the acoustic-phonetic corpus. In *Proceedings on Speech Recognition Workshop (DARPA)* (pp. 100–110).
- Liu, J., Zhao, S., Wang, J., & Kuang, J. (2011). FEC-based packet loss recovery for AVS-M audio codec. In: *2011 IEEE International Conference on Multimedia Technology (ICMT)* (pp. 3069–3072).
- López-Oller, D., Gomez, A. M., & Pérez-Córdoba, J. L. (2014). Source-based error mitigation for speech transmissions over erasure channels. In *2014 22nd European Signal Processing Conference (EUSIPCO)* (pp. 1242–1246).
- Ma, Z., Martin, R., Guo, J., & Zhang, H. (2014). Nonlinear estimation of missing δ LSF parameters by a mixture of Dirichlet distributions. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 6929–6933).
- Mehran, F. (2011). Interleaver choice in serially concatenated codes for soft real-time wireless speech transmission applications. In *IEEE 2011 19th Iranian Conference on Electrical Engineering (ICEE)* (pp. 1–6).
- Merazka, F. (2013). Packet loss concealment by interpolation for speech over IP network services. In *2013 Constantinides International Workshop on Signal Processing (CIWSP 2013)* (pp. 1–4).
- Merazka, F. (2014). A comparison of packet loss concealment and control for voice transmission over IP network services. In *IEEE 9th International Symposium: Communication Systems, Networks & Digital Signal Processing (CSNDSP)* (pp. 497–501).
- Miralavi, S. R., Ghorshi, S., Mortazavi, M., & Choupan, J. (2011). Packet loss replacement in voip using a recursive low-order autoregressive model-based speech. In *8th International Multi-Conference on Systems, Signals and Devices (SSD)* (pp. 1–4).
- Nagano, T., & Ito, A. (2013). A Packet Loss Recovery of G.729 speech using discriminative model and N-gram. In *2013 Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing* (pp. 267–270).
- Nath, D., & Kalita, S. K. (2014). An effective age detection method based on short time energy and zero crossing rate. In *IEEE 2nd International Conference In Business and Information Management (ICBIM)* (pp. 99–103).
- Oh, S. M., & Kim, J. H. (2012). Application-aware retransmission design for VoIP services in BWA networks. In *IEEE 14th International Conference on Advanced Communication Technology (ICACT)* (pp. 122–131).
- Park, N. I., Kim, H. K., Jung, M. A., Lee, S. R., & Choi, S. H. (2010). A packet loss concealment algorithm robust to burst packet loss using multiple codebooks and comfort noise for CELP-type speech coders. *Communication and Networking* (pp. 138–147). Berlin, Heidelberg: Springer.
- Perceptual Evaluation of Speech Quality (PESQ). (2001). An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs. *ITU-T P.862 Recommendation*.
- Perkins, C., Hodson, O. & Hardman, V. (2001). A survey of packet loss recovery techniques for streaming audio. *Readings in Multimedia Computing and Networking*, 607–615.
- Silveira, F., & Silva, E. D. S. (2012). Predicting packet loss statistics with hidden Markov models for FEC control. *Computer Networks*, 56(2), 628–641.
- Singh, H. P., Singh, S., Singh, J., & Khan, S. A. (2014). VoIP: State of art for global connectivity a critical review. *Journal of Network and Computer Applications*, 37, 365–379.
- Taleb, A., (2011). Low-complexity code excited linear prediction encoding. U.S. Patent No. 8,000,967. Washington, DC: U.S. Patent and Trademark Office.
- Toral-Cruz, H., Pathan, A. S. K., & Ramirez Pacheco, J. C. (2013). Accurate modeling of VoIP traffic QoS parameters in current and future networks with multifractal and Markov models. *Mathematical and Computer Modelling*, 2832–2845(57), 11.
- Toyoshima, M. & Shimamura, T. (2014). Packet loss concealment for VoIP based on pitch waveform replication and linear predictive coding. In *IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)* (pp. 89–92).
- TS 26.190, Adaptive multi-rate-wideband (AMR-WB) speech codec: Transcoding functions, *3GPP Tech. Spec*.