Joint Source-Channel CELP Coding

Snejana G. PLESHKOVA-BEKIARSKA

Faculty of Communications Techniques and Technologies, Technical University, Sofia, Bulgaria

snegpl@tu-sofia.bg

Abstract. *The method of speech coding CELP is extensively used in much voice communication-, multimedia-, video conference and other systems. There are a lot of papers related to CELP coding characteristics improvement for a better speech quality after decoding. Most of the papers are dedicated to lower the rate of CELP coded speech transmission. One problem related to low rate CELP speech transmission with a good speech decoding quality is noising in the channel for transmission. The goal of this paper is to combine the advantages and the possibilities of CELP speech coding to reduce the rate of transmission and the methods of channel coding to protect the most important CELP coding parameters in each speech frame such as line prediction coefficients, excitation indexes etc.*

Keywords

Source-channel Coding, CELP Coding, Speech coding, Speech Processing.

1. Introduction

The Shannon communications coding theory is based on the assumption of separately and sequentially source coding (compression) and channel coding (error protection) [1]. This is true only in the case of asymptotically long block lengths of data. In many practical applications, such a speech coding and especially in CELP coding [2, 4, 5] this condition is not satisfied perfectly. Thus, it is the goal of this article to propose a joint source-channel CELP coding method. Its goal is also to study the proposed method and to show the effectiveness from the combination of the advantages of transmission rate reduction with CELP speech source coding and the error protection with channel coding of the CELP coded parameters.

The reason for the joint source-channel CELP coding is the presence of inter- and intra-frame redundancy in the CELP coded and quantified parameters. Also it is known, that the importance of each of these parameters for the quality of the decoded speech is different. This fact can be exploited to achieve an additional effectiveness in the joint source-channel CELP coding. The theory of joint source-channel is based on the proposition to design jointly the source and channel coding operation. This means the channel coding, which is mainly used for error protection, works not independently of the source coding for the speech in CELP method – the statistic characteristics of CELP coded and quantified parameters representation for each CELP speech coded frame.

2. Joint Source-Channel CELP Model

The proposed model for joint source-channel CELP coding is shown as a sequence of blocks in Fig.1. The speech signal source is considered as a sequence of frames, each with 30 ms duration (240 samples for 8 kHz sample frequency).

After the block of CELP Source Coder, the CELP coded parameters for one frame are presented: 10 LSP – Line Spectral Pairs; Stochastic code book index, Stochastic code book index and gain; Adaptive code book index and gain. More detailed explanation of these parameters and their bits allocation is shown in Tab. 1. Next, the Channel Coder for Error Protection is shown which can work independently or joint with the CELP Source Coder. The coded information is the input to Transmission media with a chosen type of modulation. The transmission channel depends on the type of modulation. It is also necessary to choose a type of Noise Characteristics in the channel of transmission.

	LSP	Stochastic	Adaptive
Update time	30 ms (240 samples)	30/4=7,5 ms (60 samples)	30/4=7,5 ms (60 samples)
Bit per frame	34 bits [3444433333]	Index:8+6+8+6 Gain: 5x4	Index: 9x4 Gain: 5x4

Tab. 1. Detailed explanation of CELP parameters.

In the Receiving Media the received signal is demodulated and then it is put to the Channel Decoder. At the output of the Channel Decoder there are the decoded CELP parameters. They are used in CELP Source decoder to restore the speech decoded signal. In Fig.1 one can find also a block for Speech Quality Estimation. This can be used for comparatively evaluation of speech quality with or without joint source-channel CELP coding.

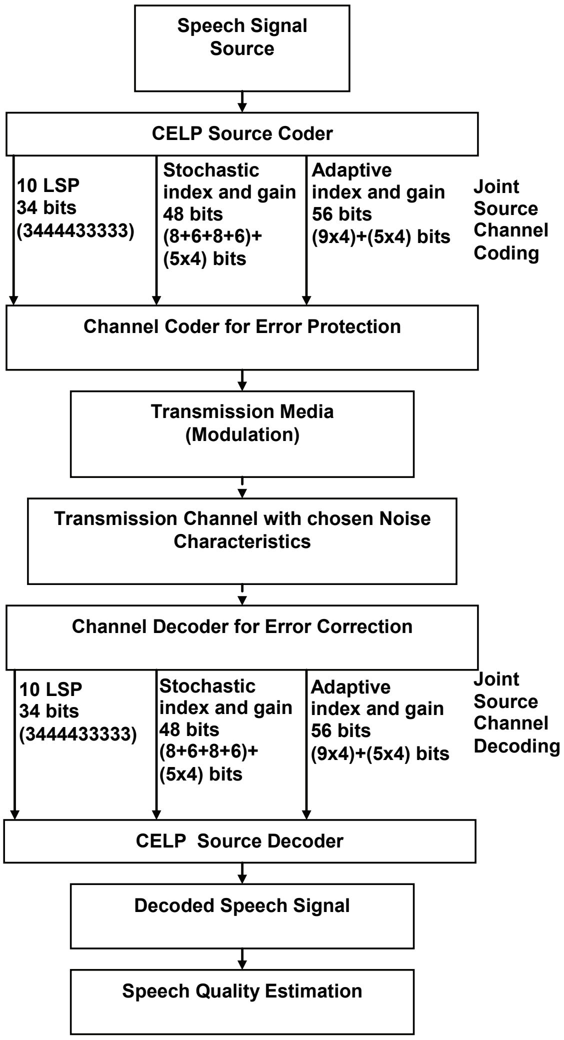

Fig. 1. Model for joint source-channel CELP coding.

3. Joint Channel Coding of CELP Parameters

The assumption that CELP coded parameters are redundant gives a reason to propose a joint source-channel coding of both these parameters, for error protection and for using their redundancy to improve the effectiveness jointly from the source and channel CELP coding in this channel coding.

It is chosen to use a family of so called rate compatible punctured convolution codes (RCPC) [3], commonly applied in channel coding. With these codes it is possible to achieve higher rate $R=k/n$ than the conventional code with rate $R=1/n$, where k and n are the number of input and output bits, respectively. These codes can be extracted easily from the basic puncturing matrix. The basic convolution code is periodically extracted from this matrix. For example to produce a punctured convolution code with rate $R=P/(P+\delta)$ and period P from a basic code with rate $R=1/n$ a punctured matrix $A(\delta)$ with dimension $(n\ x\ P)$ and $\delta\in[1,(n-1)P]$ is used. For $n=2$ basic code has a rate $R=1/2$ and for a period $P=4$ four different code rates $R=4/5$, 4/6, 4/7 or 4/8 are possible and the punctured matrix is:

$$A(\delta)_{2x2}=\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}. \qquad (1)$$

For the basic code with length 3 the generator matrix is:

$$G(D)=\left[1+D^2,1+D+D^2\right]. \qquad (2)$$

With (1) and (2) a resulting code with rate $R=2/3$ and with the following generator matrix can be performed.

$$G(D)=\begin{bmatrix}1+D & 1+D & 1\\ 0 & D & 1+D\end{bmatrix}. \qquad (3)$$

The corresponding trellis for this punctured convolution code is shown in Fig.2 for period $P=2$, output bits X0, X1 and “X”, witch means punctured bits.

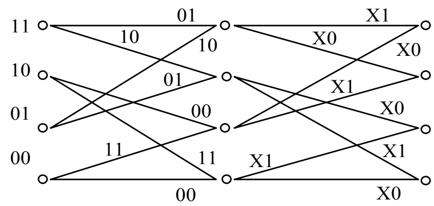

Fig. 2. The trellis for punctured convolution code.

More generally for basic code with rate $R=1/n$ and period P the punctured matrix is:

$$A(\delta)=\begin{bmatrix}a_{11}(\delta) & \dots & a_{1p}(\delta)\\ . & & \\ . & & \\ a_{n1}(\delta) & \dots & a_{1p}(\delta)\end{bmatrix} \qquad (4)$$

where

$$\{a_{ij}\}(\delta)\in\{0,1\},\ 1\le i\le n,\ 1\le j\le P^2 \text{ and } 1\le\delta\le(n-1)P.$$

The advantage of punctured convolution codes is less complexity in relation to non-punctured codes, and this is reason for choosing them in the proposition to investigate the joint source-channel CELP coding in this paper.

The starting point for channel coding in Fig.1 is after the CELP source coding, when each frame’s CELP coding parameters are calculated: 10 LPS, 4 pitch gains, 4 pitch delays, 4 codebook gains and 4 codebook indices.

There are some different possibilities to perform channel coding:

- without using the channel coding;
- independent channel coding from the source CELP coding;
- joint source–channel coding equal for all CELP coding parameters;
- joint source–channel coding unequal or partial for some of CELP coding parameters, depending on their importance for the quality of the CELP decoded speech.

Also it is necessary to decide what type or model of channel noise to use CELP coded parameters are affected from the chosen channel noise. All of these mentioned conditions and possibilities can be modeled and tested and some of them are presented shortly in this paper and especially the decoding process, which is mainly responsible for the quality of the decoded speech.

4. Joint Channel Decoding of CELP Parameters

CELP parameters are sent over a channel without memory. At the receiver, the channel decoder of CELP parameters is based on the Viterbi algorithm, which chooses the received code sequence:

$$\hat{x}^k = \left(\hat{x}_1, \hat{x}_2, ..., \hat{x}_k \right), \tag{5}$$

that minimizes

$$P_r\left(y^k \middle| \hat{x}^k\right) P_r\left(\hat{x}^k\right) \tag{6}$$

where $y^k = \left(y_1, y_2, ..., y_k \right)$ is the received sequence of length K, equal to the number of transmitted CELP parameters.

For applying the equations (5) and (6) it is necessary to define or choose the type of modulation used in transmitter and also the model of the channel noise. Here it is proposed to apply BPSK modulation with AWGN and fully interleaved Raleigh Fading channels and with noise variance $N_0/2$. For this case the minimization of:

$$\begin{aligned}
&\sum_{k=1}^{K} \left\| y_k - a_k \hat{x}_k \right\|^2 - N_0 \ln P_r\left(\hat{x}_k \right) = \\
&= \sum_{k=1}^{K} \left[\left\| y_k - a_k \hat{x}_k \right\|^2 - N_0 \ln P_r\left(\hat{x}_k \middle| \hat{x}_{k-1}, \hat{x}_{k-2}, ... \right) \right] = \\
&= \sum_{k=1}^{K} \left[\left\| y_k - a_k \hat{x}_k \right\|^2 - N_0 \ln P_r\left(\hat{u}_k \middle| \hat{u}_{k-1}, \hat{u}_{k-2}, ... \right) \right]
\end{aligned} \tag{7}$$

where a_k is the sequence of Raleigh Fading coefficients, which are assumed to be available at the decoder. If the channel is chosen as AWGN, then a_k is the all-one vector for all k; $P_r(\hat{u}^k)$ - Markov model of CELP parameters; and $\hat{u}^k$– presentation of CELP parameters as random process.

There are different possibilities to represent CELP parameters as Markov model of chosen order: 1st, 2nd etc. Usually, the first-order or second-order Markov model is chosen and each i-th CELP parameter is described as a $\{U_{i,j}\}$ random process, where j is the current frame number from the sequence of frames, defined in speech signal. If the number of CELP parameters in a frame is l, then

$$U_j = \left\{ U_{1,j}, U_{2,j}, ..., U_{k,j} \right\} \tag{8}$$

is the random representation of CELP parameters in the j-th frame.

The first-order or second-order Markov models of the CELP parameters as random process $\{U_{i,j}\}$ can be used for estimation of their entropy rate and defining of redundancy and minimize the probability from the equation (5). This equation is related to the first-order Markov model chosen for the transmitting CELP parameters:

$$P_r\left(U_j = u_j \middle| U_{j-1}, = u_{j-1}, ..., U_1 = u_1\right) = P_r\left(U_j = u_j\right) \tag{9}$$

and

$$\begin{aligned}
&P_r\left(U_{i,j} = u_{i,j} \middle| U_{i-1,j} = u_{i-1,j}, ..., U_{1,j} = u_{1,j}\right) = \\
&= P_r\left(U_{i,j} = u_{i,j} \middle| U_{i-1,j}, = u_{j-1}\right) = P_F^{(i)}\left(u_{i,j} \middle| u_{i-1,j}\right)
\end{aligned} \tag{10}$$

for $i = 1, 2, ..., l$ and $j = 1, 2, ...,$ where $P^{(i)}{}_F$ is the probability, defined for the first Markov model.

The second-order model also can be described and the advantage of this model is the possibility to use not only the intra-frame, but also the inter-frame redundancy:

$$\begin{aligned}
&P_r\left(U_{i,j} = u_{i,j} \middle| U_{j-1}, = u_{j-1,}, ..., U_1 = u_1\right) = \\
&= u_{i,j}, ..., U_{i-1,j}, j = u_{i-1,j} = \\
&= P_r\left(U_{i,j} = u_{i,j} \middle| U_{i-1} = u_{j-1,j}, U_{i,j-1} = u_{i,j-1}\right) = \\
&= P_S^{(i)}\left(u_{i,j} \middle| u_{i-1,j}, u_{i,j-1}\right),
\end{aligned} \tag{11}$$

for $i = 1, 2, ..., l$ and $j = 1, 2, ...;$ where $P^{(i)}{}_S$ is the probability, defined for the second-order model.

5. The Simulation of Joint Source-Channel CELP Coding

The proposed method of joint source-channel CELP parameters coding is simulated in Matlab and it is presented in Fig. 3.

The simulation is a combination of a Matlab program for CELP Coded Parameters, calculated for each speech frame and a Simulink model for transmission of these parameters. First for a good representation of the speech signal, used for coding, it is chosen a lot of male and female speech testing signals separated as speech testing segments, which are in standard WAVE format. The existing Matlab FS1016 algorithm as a standard of CELP coding is used to calculate the CELP coding parameters. Then these parameters are used in two ways: without or with using the joint source-channel coding method. That means to represent CELP parameters as it is described in section 3 and 4 with first or second order Markov models, and applying the rate compatible punctured convolution codes (RCPC).

The CELP parameters without or with joint source-channel coding are put in the first block of Simulink model (Fig.3) from Matlab workspace. Following the block of the Simulink model, these parameters are transmitted using BPSK modulation, AWGN Channel and BPSK demodulation and Error Rate Calculator and Indicator or Display. The demodulated CELP Parameters in the receiving part sent back to Matlab Workspace for standard CELP decoded speech signals are saved in WAVE format for comparing of their quality with the same original encoded testing speech signals.

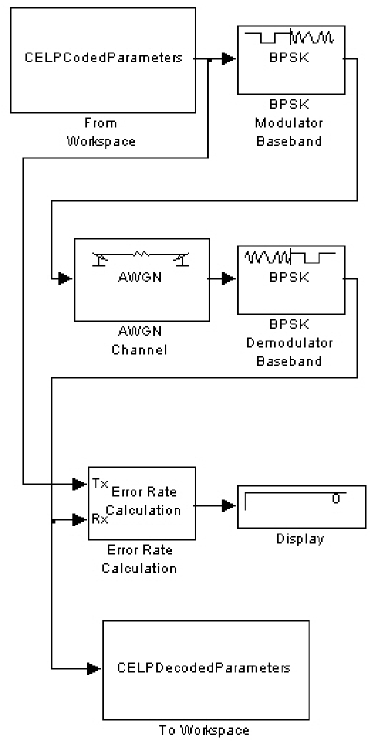

Fig. 3. The joint source-channel CELP simulation.

The results from the simulation are examined with application of standard subjective speech quality estimation method MOS (Mean Option Score) comparing original with decoded speech signals from a chosen number of listeners. The generalized conclusions from these estimations are summarized in Tab. 2 and are presented in percentages as a comparative level of estimation of the proposed method.

SNR (E_b / N_0)	Coding method	Relative speech quality
-2 dB	Without joint coding	96%
1 dB	Without joint coding	84 %
-2 dB	With joint coding	88 %
1 dB	With joint coding	97 %

Tab. 2. Relative speech quality without and with joint coding.

The block for Error rate calculation and estimation is shown in Fig. 3. The received and decoded CELP parameters are returned back to Matlab workspace, where they are used by a program for speech quality estimation comparing CELP decoded speech with and without joint source-channel coding and the speech signal before CELP coding. The subjective speech quality estimation results for male and female speech testing signals are shown in Tab. 2 also for different SNR E_b/N_0 in percentage for perfect quality (without listening quality degradation) and for some quality losses.

6. Conclusion

The results in Tab. 2 show the improvement of decoded speech quality whith the proposed joint source-channel CELP coding in the cases with larger SNR values. The briefly presented results show only a little part of the experimental studies, which are made to investigate all properties of the proposed joint source-channel CELP coding for speech signal. Of course it is necessary to continue with testing of these procedures in the direction of subjective evaluation of the joint source-channel CELP speech coding quality in comparison with traditional CELP coding in some future works.

There are many possibilities for application of the proposed method and of course mainly in audio communication systems using CELP coding. Furthermore this work is connected to the application of this method in audio visual moving robots, for which it is important to perceive and communicate using speech information. The processing and transmitting of this information can be more effective using joint source-channel CELP coding.

Acknowledgements

This work was supported by the Ministry of Science and Education of Bulgaria under Contract BY-I-302/2007 Audio-video information and communication system for active surveillance cooperating with a Mobile Security Robot.

References

[1] SHANNON, C. A mathematical theory of communication. *Bell Sys. Tech. Journal,* 1948, vol. 27, p. 379-423.

[2] CAMPBELL, J., TREMAN, T. The federal standard 1016 4800 bps CELP voice coder. *Digital Signal Processing, Academic Press*, 1991, vol.1, no.3, p.145-155.

[3] HAGENAUER, J. Rate-compatible punctured convolutional codes (RCPC Codes) and their applications. *IEEE Transactions on Communications*, 1988, vol. 36, no. 4, p.389-400.

[4] MILETIEV, R., ARNAUDOV, R. Spectrum analysis of nonuniformly sampled data for radar applications. *Frequenz*, 2007, vol. 61, no. 3-6, p. 128-132.

[5] PLESHKOVA, S., MOMCHEDJIKOV, M. Fast CELP code book search method. In *Proc. of the Int. Sci. Conf. on Energy and Information Syst. and Technol*. Sofia (Bulgaria), 2003, p.178-182.

About Authors...

Snejana PLESHKOVA-BEKIARSKA (*1964, Dobritch, Bulgaria) received M.S. degrees in Electronic and Automatic engineering from the Technical University-Sofia, Bulgaria in 1989. She is assistant professor at the same university. She published over 50 research papers. Her research interests are in the field of Neural Networks, Signal Processing and Pattern Recognitions, Digital Signal Processors for Image and Audio Processing.