CELP CODE BOOK WITH INDEX PREDICTION

Snejana Pleshkova – Bekiarska, Damian Damianov

Technical University – Sofia, Bulgaria
snegpl@tu-sofia.bg ellov@abv.bg

Abstract

The code book is one of important blocks in the CELP speech coding method, with which it is possible to make the vector quantization (VQ) of each speech sub frame. This reduces the speed of transmission replacing the sub frame with the index of most similar excitation vector in the stochastic CELP code book. In the decoder the received index is used as an address of the same stochastic code book as in the coder to produce the excitation vector, with which it is prepare the speech synthesis CELP algorithm.

In this article it is proposed an approach to reduce additionally the speed of CELP transmission with a method of index prediction. The proposed method is connected with minimal additional components in CELP speech coding and decoding and some little increasing of calculations. There are present the equations describing the proposed method. Some simulations are made for feature establishment of the proposed method and some of the results of these simulations are present in comparative style to reviewing the advantages of this method.

1. Introduction

In the CELP speech coding method there are two code books: stochastic and adaptive [1]. The adaptive code book work as a long term prediction filter for pitch prediction [2]. The stochastic code book prepare a vector quantization of speech signals. In this code book there is a collection of predefined and stored excitation vectors. The comparison of each of these vectors with the corresponding vectors of current subframe vector of the original speech signal give the decision that one of these vectors is more similar to the original speech signal subframe. The address or index to the place in the stochastic code book, where this vector is stored is transmitted after the coding of the current speech frame. This operation is called code book search, because the vector the vector is find after a procedure of mean square error calculating and minimization. The transmission of the code book index instead of situated in the code book excitation vector give a suitable bit rate reduction. But after that still redundancy in speech signal rest and it is possible to search some additional means or algorithm to improve the bit rate reduction. One such a method is proposed in this article – the method of code book indexes prediction. First it is given a brief presentation of the basic operation of the CELP coder with the place of the stochastic code book and the main equations describing the code book search procedure. Then it is present the proposed modification of this part of the CELP coder, where it is added the index prediction. The proposed modification is argued with the corresponding equations. At the end of the article there are presented the comparative results of simulations and testing of the proposed method with real speech signals coded and then decoded with a standard CELP system with and without of the proposed code book index prediction.

2. Basic CELP Operations

The code books are shown in Fig. 1.

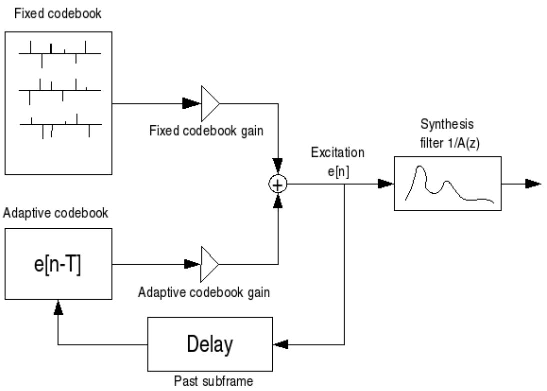

Figure 1. CELP code books

The stochastic code book is named "Fixed" because the content of the code book is constant and is preliminary filed with a lot of vectors (usually the number of vectors is 512). The "Adaptive code book" calculate the pitch T as a difference $e[n-T]$ between the current n and "Past subframe". The output vectors of the "Fixed" and "Adaptive" code books are multiplied with the corresponding gains "Fixed code book gain" and "Adaptive code book gain". These two parts are summing $\oplus$ and form of the excitation vector $e[n]$. The excitation vector $e[n]$ is then performed as the current synthesis speech subframe $\hat{s}(n)$ with the short term synthesis filter $1/A(z)$:

$$A(z) = 1 + \sum_{i=1}^{P} a_i z^{-i}, \qquad (1)$$

where

a_i are the coefficients of the speech frame analysis filter;

P – the order of the speech frame analysis filter (in CELP standard usually $P = 10$.

The procedure of code book search by means of iterative calculating and minimization of the mean square error is shown in Figure 2.

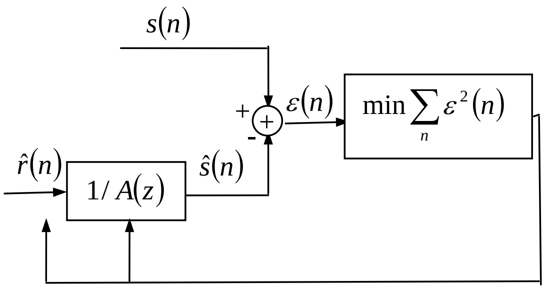

Figure 2. Code books search

The current subframe of the original speech signal $s(n)$ is compared (subtracted) with the current subframe of the synthesis speech $\hat{s}(n)$ build from the chosen excitation vector $e(n)$, as is shown in Figure 1. The result of this comparison is the error $\varepsilon(n)$:

$$\varepsilon(n) = s(n) - \hat{s}(n) \qquad (2)$$

The minimization of the mean square error:

$$\min \sum_n \varepsilon^2(n) \qquad (3)$$

is performed by choosing each time a new vector from the stochastic code book, calculating the error $\varepsilon(n)$ and then check the square error for the minimum. If these minimum is found, then the index of the corresponding vector in the stochastic code book is transmitted. It must to imply from the Figure 2, that the determined index is the output of the block, which prepare the error minimization $\min \sum_n \varepsilon^2(n)$. The sequence of the indexes follow the sequence of the subframes and is added to the rest of parameters transmitted for each frame of the analysed original speech frame. The quantity of the information transmitted for each index depend of the number of excitation vectors in the stochastic code book (usually the number of vectors is 512) and hence the number of bits transmitted for each index is 9. To reduce this quantity of transmitted information it is proposed in this article to perform an index prediction before of their transmission.

3. Stochastic Code Book Index Prediction

The place, where it is possible to add the stochastic code book index prediction is shown in Figure 3.

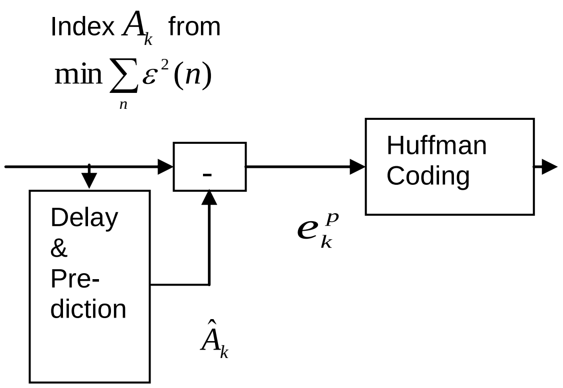

Figure 3. Stochastic code book index prediction

The input for the prediction in Figure 3 is the sequence of the indexes A_k, which is the output of the block of minimization $\min \sum_n \varepsilon^2(n)$ in Figure 2. To perform the prediction it is necessary to story some of the past values of the indexes A_k. This is done with the block named Delay & Prediction. Also in this block the stored past values are used for calculating the predicted value $\hat{A}_k$ of the current index A_k:

$$\hat{A}_k = \sum_{i=1}^{N} a_i A_{k-i} \ , \qquad (4)$$

where

N is the order of the prediction and can be chosen experimentally by controlling the decreasing of the quality of the decoded speech signals:

a_i – the coefficients of the prediction filter.

The substraction of the current index A_k and the predicted value $\hat{A}_k$ give the error e_k^p:

$$e_k^p = A_k - \hat{A}_k \qquad (5)$$

It is known from the prediction theory [3] that the variances of the prediction error e_k^p are in principles more less than the variances of the indexes A_k. Then it is evident that the transmission of the prediction error e_k^p instead of the indexes A_k can be made with less number of bits, than the transmission of indexes A_k.

To gain an additional reduction of the bit rate transmission it is proposed on the Figure 3 to use Huffman coding [4] after the stochastic code book index prediction. It is known, that Huffman code is an irregular method of coding and it is necessary to examine and define the length of the Huffman code. This is possible if there is a representative statistic of variances of stochastic code book indexes A_k and respectively of the prediction error e_k^p.

4. The Code Book Index Characteristics Analysis

The effect of the proposed method of stochastic code book index prediction can be increased if the positions of the excitation vectors placed in stochastic code book are chosen in a manner, that their correlation give the minimal variances in prediction error e_k^p. That means the values of indexes for the neighbouring subframes, chosen after minimization of the mean square error can be also neighbouring if it is possible. This condition can be prove be means of the equation for disordering of vectors in the stochastic code book:

$$D_\rho = W_\rho \sum_{i=1}^{S} \sum_{j=1}^{S} \rho^{|i-j|} d(y_i, y_j), \qquad (6)$$

where

$d(y_i, y_j)$ is Euclidean distance between $i-th$ and $j-th$ code vectors;

$\rho \in [0,1]$ – coefficient to improve the weight of vectors, which are very close in the code book.

There are some possibilities to decrease the disordering of vectors in the stochastic code book. For example: the Kohonen Neural Network [5] or other self organized Neural Networks [6], but it is the object of the future work.

5. Results of Experiments and Conclusion

Some experiments are prepared for testing the properties of the proposed method of stochastic code book index prediction. For the test are used simulated and real speech signals, which are coded and then decoded using both the standard CELP method and proposed method with stochastic code book index prediction. The generalized results as a comparative presentation of calculated segmental signal to noise ratio (SEGSNR) are present in the Table 1. These results shows that for simulated sinusoidal signals (Sin600Hz and Sin1600Hz) and real (Female and Male) speech signals there are little degradations of segmental signal to noise ratio (SEGSNR) for the proposed method with stochastic code book index prediction. It can be supposed, that this is because the order of prediction is not enough, but it is chosen small to keep the calculations little. Also it is possible to improve the segmental signal to noise ratio (SEGSNR) for the proposed method with stochastic code book index prediction if it is performed one of the above mentioned for decreasing the disorder of vectors in the stochastic code book.

Table 1

Test signals	SEGSNR for standard CELP method	SEGSNR for index prediction method
Sin600Hz	9,3452	10,0654
Sin1600Hz	10,6734	9,6453
Female	8,6733	8,6689
Male	7,8567	7,6590

References

[1] Kondoz, A.M. Digital speech coding for low bit rate communication systems. Yohn Wiley&Sons, New York, 2000.

[2] Federal Standard 1016, “Telecommunications: Analog to digital conversion of radio voice by 4800 bit/second code excited linear prediction (CELP)“, National Communication System-Office of Technology and Standards, Feb. 1991.

[3] Atal B.S. and Remde J.R., A new model of LPC excitation for producing natural-sounding speech at low bit rates. Proc. IEEE-ICASSP, Paris, France, 614-617, 1982.

[4] Cormen Thomas H., Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. Introduction to Algorithms, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Section 16.3, pp.385–392.

[5] Kohonen T., Self – Organization and Associative Memory, 2nd Edition, Springer – Verlag, 1988.

[6] Riskin E.A., Atlas L.E., Lay S.R., Ordered neural maps and their applications to data compression, Proc. IEEE Workshop on Neural Networks for Signal Processing, pp.543-551, September 1991.