Design and training of codebook for CELP coder

Lecturer Cornel Balint,
„Politehnica" University of Timişoara

ABSTRACT. The technique of Code Excited Linear Prediction (CELP) developed in the last years are used in voice coding system that provide a good quality at low and very low bit rate. CELP schemes with trained excitation codebook are able to accurate reproduce more complex waveform than stochastic CELP schemes. The vectors of the excitation codebook are adapted to a training speech sequence using an iterative algorithm. Comparative listening test with CELP that use stochastic codebook have shown quality improvements in objective SNR and subjective listening test, concerning transition region between unvoiced and voiced signals.

1 CELP coding of speech

CELP speech coders [GG92] employ an analysis-by-synthesis technique to characterize the excitation signal for a reconstruction filter (LPC filter).

A CELP coder, schematic illustrated in fig. 1, performs first a LPC analysis. LPC parameters are coding and sending to the CELP decoder. In coder, LPC parameters are used in order to compute the speech residual, as a difference between original speech signal x(n) and a synthetic speech signal $\hat{x}(n)$ locally generated by filtering, in LPC synthesis bloc, an excitation signal extracted from the codebook and multiplied with a gain factor. A perceptual weighting filter, based on LPC coefficients, introduce some perceptual improvements in the closed loop analysis. The search objective is finding the excitation signal from the codebook that, after filtering, produce a signal identical to the input signal.

The speech residual, after short and log term prediction (long term prediction are not represented in fig.1), are closely to a Gaussian distributed

signal, so stochastic codebook, generated by Gaussian process are used to predict the residual of speech.

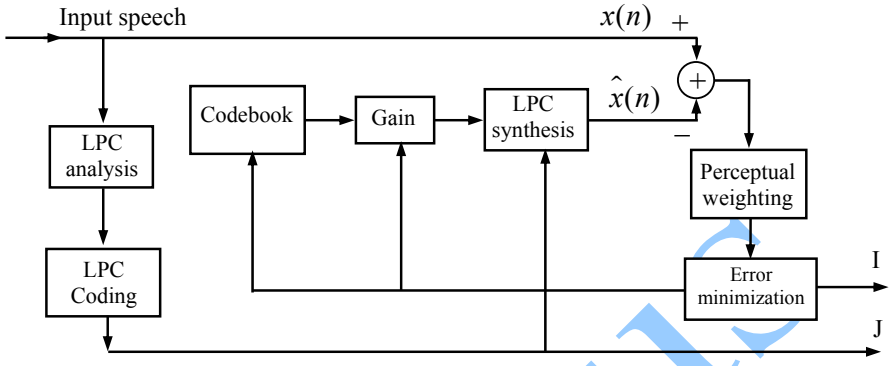

Figure 1 CELP Speech coder

In order to reduce the complexity of codebook search, some particular structure for the stochastic codebook was proposed [SA85, Soh91].

An ultimate solution is an algebraic codebook, containing sparse ternary vector, which have only 0, +1 and -1 value. Such a stochastic codebook can be deterministic organized [KB92], with some advantages concerning computing complexity.

2 Training the CELP codebook

CELP codebook, stochastic or deterministic, obtains a relatively high accuracy in stationary parts of voiced speech. Problems arise in transition region where the memory of long term predictor is not yet adapted to the changed signal characteristics.

As shown in fig. 2, CELP with trained excitation codebook can reproduce more accurate a complex excitation signal.

The proposed codebook training method the excitation vectors are adapted to a training speech sequence by applying an iterative optimization scheme based on classic LBG algorithm [LBG80], [CRL92]. According to LBG algorithm, two steps are performed successively. In the first step, the data training set are vectorial coded using the given codebook. In the second step, the centroid of all data vector assigned to the same codeword is calculated.

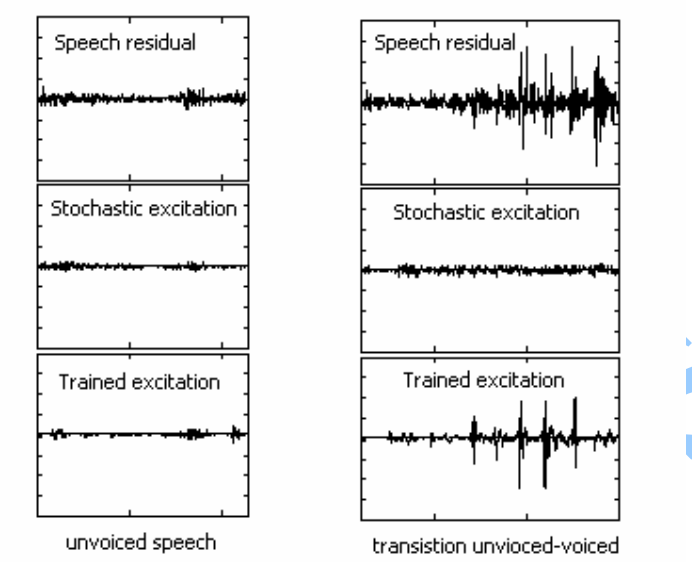

Figure 2. Speech excitation signal for unvoiced and voiced speech

In order to assure a minimum coding error in CELP coding process, in the training process is used an error measure similar to the error measure used in search process. For a given set of N-dimensional training data $\{x_i(n)\}$ the cumulated error for all input vector belonging to a region P_j of a codebook vector c_j is given by:

$$\varepsilon_j = \sum_i \frac{\sum_{n=0}^{N-1} (x_i(n) - g_i c_j(n) * h_i(n))^2}{\sum_{n=0}^{N-1} x_i^2(n)} \qquad (1)$$

where $x_i(n)$ is the i-th data vector, g_i is the optimum gain value and $h_i(n)$ denote the impulse response of the cascaded synthesis filter and weighting filter. Minimizing ε_j leads to a set of linear equation for a given value g_i, and each iteration to find the optimum codebook is followed by an iteration in order to optimize the gain value.

3 Experimental result

The proposed method for training the CELP codebook was simulated for the coding of speech signal limited to the telephone bandwidth.

A simplified CELP algorithm based on [ITU96] was used for testing a variety of codebook. The tests were made using all type of codebook: stochastic, algebraic and trained. Codebook size was selected in range 32-1024.

An improvement of SNR of 1,25 to 2,0 for the trained codebook relative to the stochastic and algebraic codebook was computed.

Comparative listening tests show a more natural sound speech.

Future experiments are oriented on training the algebraic codebook, in order to determine by training sign and position of the pulse.

References

[CRL92] **D. Cohn, E. Riskin, R. Ladner** – *Theory and Practice of Vector Quantizers Trained on Small training Sets*, Technical Report TR 92-12-08, University of Washington, 1992

[GG92] **A. Gersho, R. M. Gray** – *Vector quantization and Signal Compression* Kluwer Academic Publishers, London 1992

[ITU96] **ITU** – T Recommendation G729 – *Coding of speech at 8 kbps using conjugate-structure algebraic code excited linear prediction (CS-ACELP)* 1996

[KB92] **Yu-Hung Kao, J. S. Baras** – *A new Deterministic Codebook Structure for CELP Speech Coding,* Tehnical research report, Inastitute for System Research, 1992, http://www.isr.umd.edu

[LBG80] **Y. Linde, A. Buzo, R. M. Gray** – *An Algorithm for Vector Quantizer Design*, IEEE Transaction on Communications, vol. 28, 1980

[SA85] **M. R. Schroeder, B. S. Atal** – *Code-Excited Linear Predictive (CELP): High Quality Speech at Very Low Bit Rates,* Proc. IEEE Int. Conf. ASSP, 1985

[Soh91] **Y. Soham** – *Constrained Stochastic excitation coding of speech at 4,8 kbit/sec*, in Advanced in Speech Coding, Kluwer Academic Publishers, 1991