A LOW NOISE FIXED POINT IMPLEMENTATION OF GSM SPEECH CODEC ON TMS320C25

Shu H. Leung, Chi Y. Chung and Andrew Luk

Department of Electronic Engineering
City Polytechnic of Hong Kong

ABSTRACT - This paper is to present a fixed point implementation of Regular-pulse Excitation Linear Predictive coder (RPE-LPC) that combines with long term prediction (LTP) on using a single TMS320C25. The bit rate of the coder is 13kbits/s. The coding is done by Toeplitz approximation that permits the use of lattice filter for reducing the finite wordlength effects such as coefficient sensitivity and roundoff error.

INTRODUCTION

The coding scheme based on regular-pulse excitation linear predictive coding that combines with long term predictor is being adopted by CEPT Groupe-Speciale-Mobile (GSM) as the standard of speech codec in digital mobile radio system in Europe (Vary et. al., 1988). In the encoding and decoding, the direct form structure is commonly used for synthesis and regular-pulse encoding. Such direct implementation can simplify the formulation but has very large dynamic range and poor numerical behaviour.

The most important and time-consuming part of the codec is the regular-pulse excitation encoding. The computations involve the calculation of the impulse response of the weighting synthesis filter, the covariance matrix of the impulse response and the solution of the Normal equation. The direct implementation of the equations has very poor numerical errors and heavy computations. In order to reduce the calculations in obtaining the regular pulses, the problem can be simplified by Toeplitz approximation. Not only there is improvement in the calculation and also the pulses can be solved by using numerically stable lattice algorithm.

In developing the lattice algorithm, first of all we regenerate the autocorrelation lags directly from the coefficients of the weighting filter instead from the impluse response. Then we apply the algorithm as in (Roux & Gueguen, 1977) on the Toeplitz system with spacing of four to compute a new set of reflection coefficients. The new reflection coefficients together with a set of new variables construct an efficient lattice-ladder filter for solving the Toeplitz equation.

The algorithm is implemented on a single TMS320C25 and is proved to have good numerical behaviour.

RPE-LPC ENCODER

The block diagram of the RPE-LPC encoder is subdivided in five functional blocks as shown in Figure 1: preprocessing, LPC analysis, short term analysis, long term prediction, RPE encoding.

The preprocessing is to remove the DC component and use a first order FIR filter to pre-emphasize the speech signal in order to "flatten" the speech spectrum and to reduce the signal's dynamic range for easing fixed point implementation.

The LPC analysis is to calculate the reflection coefficients from the autocorrelation lags in every 20 ms by using a fixed point computational technique as introduced in (Roux & Gueguen, 1977).

Let v_i^m denote the intermediate variables as defined by

$$v_i^m = \sum_{l=0}^{m} a_l^m \, r_{i-m} \qquad (1)$$

where a_i^m and r_i are the m-th order forward predictor and the i-th autocorrelation lag, respectively. It has been shown that the squared magnitude of the variables are bounded by r_0. The recursions for computing the reflection coefficients k_m are summarized as follows:

$$v_m^0 = r_m , \qquad m = 0, 1, \ldots, p \qquad (2)$$

$$k_{m+1} = -v_{m+1}^m \,/\, v_0^m \qquad (3)$$

$$v_0^{m+1} = v_0^m \, (1 - k_{m+1}^2) \qquad (4)$$

$$v_i^{m+1} = v_i^m + k_{m+1} \, v_{m+1-i}^m \, . \qquad (5)$$

The reflection coefficients are quantized and coded in the form of log-area ratios (LAR). The current and the previous set of decoded LAR coefficients are interpolated linearly in every 5 ms to avoid spurious transients. The speech signal is then inversely filtered by the corresponding transversal lattice filter to generate a residual signal (prediction error signal) d(t). The advantages of using the lattice filter are two folds: (1) we do not require to calculate and use the prediction filter whose coefficients have very large dynamic range; (2) the lattice filter is a low noise structure.

The long term predictor section is a closed loop circuit that uses the reconstructed excitation $d^*(t)$ to estimate the residual signal. The gain b and the delay M in the predictor are updated in every 5 ms as follows:

$$M = \arg\{ \max_i C(i) \} \qquad (6)$$

where

$$C(i) = \sum_{j=0}^{39} d(m_l + j) \, d^*(m_l + j - i) ,$$

$$m_l = m_0 + 40l , \qquad l = 0, 1, 2, 3 \qquad i = 40, \ldots, 120$$

$$d^*(t) = \text{reconstructed excitation}$$
$$= \text{long term prediction + quantized regular pulses.}$$

Regular pulse excitation encoding

Let L, N and K denote the frame size, the number of pulses and the number of phases, respectively. In the standard, they are respectively equal to 40, 13 and 3. Let h(t) denote the impulse response of the weighting filter $1/A(z/\gamma)$. The corresponding response matrix H is thus given by

$$H = \begin{bmatrix} h(0) & h(1) & \cdots & h(39) \\ 0 & h(0) & & h(38) \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & h(0) \end{bmatrix} . \qquad (7)$$

Mathematically, the response due to the memory hangover and the residual is given by

$$\mathbf{y}_0 = \mathbf{y}_m + \hat{\mathbf{d}}\, \mathbf{H} . \qquad (8)$$

where

$\mathbf{y}_m$ = response due to memory hangover

$\hat{\mathbf{d}} = \mathbf{d} - \mathbf{d}^*$, $\mathbf{d}$ = residual vector, $\mathbf{d}^*$ = residual estimated by LTP

In the implementation, the residual $\hat{d}(t)$ is inputted to the synthesis filter of lattice structure as shown in Figure 2 in which the weighting parameter γ is multiplied to the delay outputs.

The response vector $\mathbf{y}$ due to the regular pulses of phase l is given by

$$\mathbf{y}_l = \mathbf{z}_l \mathbf{H}_l, \qquad l = 0,\ 1,\ 2 \qquad (9)$$

where $\mathbf{z}_l$ is the excitation pulse vector of phase l and $\mathbf{H}_l$ is the submatrix of $\mathbf{H}$ given by

$$\mathbf{H}_l = \mathbf{M}_l\, \mathbf{H} \qquad (10)$$

where

$$M_l = \begin{cases} m_{ij} = 1 & \text{if } j = iK + l \quad , \quad l = 0,\ \ldots,\ K-1 \\ m_{ij} = 0 & \text{otherwise,} \quad 0 \le i \le N-1, \quad 0 \le j \le L-1 . \end{cases}$$

The error vector $\mathbf{e}_l$ is the difference between $\mathbf{y}_0$ and $\mathbf{y}_l$,

$$\mathbf{e}_l = \mathbf{y}_0 - \mathbf{y}_l .$$

The optimal pulse vector $\mathbf{z}_l$ is the one that minimizes the squared error $\mathbf{e}_l \mathbf{e}_l^t$ and is given by

$$\mathbf{z}_l = \mathbf{y}_0 \mathbf{H}_l^t\, [\mathbf{H}_l \mathbf{H}_l^t]^{-1} . \qquad (11)$$

A simple algorithm is proposed in (Kroon et. al., 1986) for determining the phase l: use a low order "smoother" to process the residual signal d(t) (Flanagan et. al., 1979), and then downsample the response at different phases to form K subsequences, the phase for which the subsequence has the maximum energy is selected.

The solution according to (11) is quite computational intense. The calculations can be reduced by modifying the matrix $H_1H_1^t$ to some simple forms. The Toeplitz approximation has been proved to have very close performance to the solution as obtained in (11). Besides, such approximation provides a low noise lattice implementation as discussed in next section.

Toeplitz approximation

In direct implementation of (11), we require to compute the impulse response of the weighting filter and then calculate the autocorrelation lags. In the following, we avoid the computation of the impulse response and calculate the autocorrelation sequence directly from the prediction filters.

The coefficients of the weighting filter can be obtained by finding the impulse response of the analysis filter as shown in Figure 3. Then the reflection coefficients of the weighting filter can be calculated by inversing the Levinson algorithm:

$$k_m = a_m^m \ , \qquad a_m^m = \text{m-th coefficient of the m-th order predictor}$$

$$v_0^{m-1} = v_0^m / (1 - k_m^2)$$

$$a_l^{m-1} = (a_l^m - k_m a_{m-l}^m)/(1 - k_m^2)$$

with $v_0^p = 1$.

According to equation (3), the autocorrelation sequence is recursively computed as follows:

$$r_{m+1} = \begin{cases} -k_{m+1} v_0^m - \sum_{l=1}^{m} a_l^m r_{m+1-l} \ , & \text{for } m \le p \\ - \sum_{l=1}^{m} a_l^p r_{p-l} \ , & \text{for } m > p \ . \end{cases} \qquad (12)$$

The Toeplitz matrix $H_1H_1^t$ has the first row elements given by $\{r_0, r_4, \ldots, r_{36}\}$. Let $\hat{r}_l = r_{4l}$ for $l = 0, 1, \ldots, N$. We can make use of equations (2) to (5) to calculate the new set of reflection coefficients for obtaining the solution in (11).

Let us rewrite equation (11) as

$$\hat{\mathbf{R}}\,\mathbf{z} = \mathbf{q} \ , \qquad \mathbf{q} = \mathbf{y}_0 \mathbf{H}_1^t \ . \qquad (13)$$

where the vector $\mathbf{q}$ can be obtained by decimating the response of the weighting filter at the phase 1 with the vector $\mathbf{y}_0$ as the input.

The solution $\mathbf{z}$ can be solved recursively using the following lattice-ladder algorithm. Let us define the solution at m-th stage be z_m. The recursion for computing the solution is given by

$$z_{m+1} = \begin{pmatrix} z_m \\ 0 \end{pmatrix} + k^b_{m+1} \hat{A}_{m+1} \qquad (14)$$

where $\hat{A}_m$ is the m-th order backward prediction filter. Let us define a new variable g^m_l as

$$g^m_l = \sum_{i=0}^{m} z^m_i \hat{r}_{l-i} . \qquad (15)$$

We have

$$g^{m+1}_l = g^m_l + k^b_{m+1} v^m_{m+1-l} . \qquad (16)$$

The coefficient k^b_m is calculated by

$$k^b_m = (q_m - g^{m-1}_m) / v^m_0 , \quad q_m = \text{m-th element of the vector } \mathbf{q}. \qquad (17)$$

Using the new reflection coefficients we can obtain the backward prediction filters and then compute the solution recursively by equation (14).

The complexity of the new Toeplitz approximation method is summarized as follows:

weighting filter coefficients	$3p^2/ 2$
inverse Levinson algorithm	p^2
autocorrelation sequence	Lp
calculation of $\mathbf{q}$ vector	Lp
Toeplitz solution	$3p^2/ 2$

The overall complexity is about $O(2Lp + 4p^2)$ compared to the direct implementation of $O(LN + Lp + 2p^2)$. It is apparent that the Toeplitz method requires lower computations and better numerical behaviour.

RESULTS

The new algorithm is implemented on a single TMS320C25. The average signal to noise ratio is about 13 dB for a number of Chinese utterances. The hearing test shows that the new implementation has better performance than the direct implementation of the RPE encoding.

CONCLUSION

A low noise fixed point implementation of RPE-LPC with long term predictor has been presented. The algorithm is shown to have lower computations and better numerical properties than the direct implementation. As a result, the algorithm is well suitable for implementing on low cost fixed point digital signal processor.

REFERENCES

Flanagan J. L. et. al. (1979) "Speech Coding", IEEE Trans. Commun. 27, 710-736.

Kroon P. et. al.(1986) "Regular-Pulse Excitation - A Novel Approach to Effective and Efficient Multipulse Coding of Speech", IEEE Trans. Acoust., Speech, Signal Processing 34, 1054-1063.

Le Roux J. & Gueguen C. (1977) "A Fixed Point Computation of Partial Correlation Coefficients", IEEE rans. Acoust., Speech, Signal Processing, 257-259.

Vary P. et. al., (1988) "Speech Codec for the European Mobile Radio System", Proc. IEEE Int. Conf. Acoust., Speech Signal Processing, 227-230.

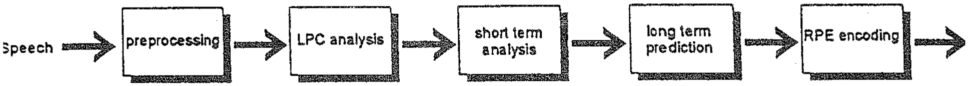

Fig. 1 The block diagram of the RPE-LPC encoder

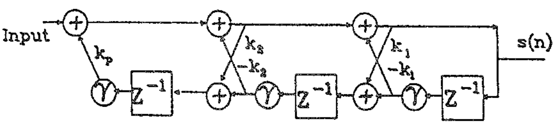

Fig. 2 Synthesis lattice filter

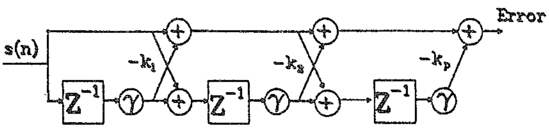

Fig. 3 Analysis lattice filter