A REAL TIME PROCESSING TCELP CODER/DECODER AT 4.8 KBIT/SEC USING LSP SPLIT VECTOR QUANTIZATION

N. NAJA*, A. GOALIC#, A. EL KESRI* ,J.M. BOUCHER# and S. SAOUDI#

* Laboratoire d'Electronique et Communications, E.M.I., B.P. 765, Rabat-Agdal, MAROC
# Département Signal et Communication, ENST-Br., B.P. 832 - 29285 Brest - FRANCE
andre.goalic@enst-bretagne.fr

ABSTRACT

Before transmission in a narrow band channel, the speech signal has to be compressed. The Code Excited Linear Prediction Coder (CELP) makes it possible to synthesize good quality speech at low bit rates. Different speech linear predictive coding parameters can be used to design the speech spectral envelope. In our coder the Line Spectrum Pairs (LSP), belonging to the frequency domain, enable us to design the vocal track transfer function. In its first version the coder uses a CLSP (LSP cosine function) scalar quantization leading to a rate of 5.45 kbit/sec. In order to reduce the bit rate to a standard (4.8 kbit/sec), vector quantization was introduced in our coder.

1 TCELP CODER

The CELP Coder (fig. 1) is made up of three parts.

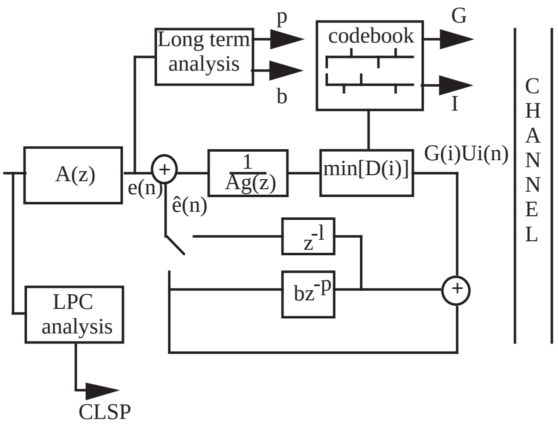

fig. 1 : TCELP Coder

Short and long term analyses make it possible to design vocal track model and vocal chords effect.

After removing short and long term redundancies, the original speech gives the residual signal. This is modelled by a waveform chosen in a codebook. Different kinds of codebook are used according to applications and real time necessities (Gaussian, stochastic, algebraic..).Table 1 presents the bit allocation for the CLSPs vector quantization coder.

Parameters	Block(ms)	Bits/frame	Rate(bits/s)
CLSP(10)	20	17 x 1	850
Glpc	20	7 x 1	350
B ltp	10	3 x 2	300
Pitch	10	7 x 2	700
Gq	5	3 x 4	600
I index	5	10 x 4	2000
	TOTAL	96	4800

Table 1 : Bit allocation

1.1 LPC Analysis

At low bit rates, speech quality mainly depends on spectral envelope design accuracy. Various speech linear predictive coding parameters have been studied to design this spectral envelope. After the PARtial CORrelation (PARCOR) parameters, the Line Spectrum Pairs (LSP) have some interesting properties in speech coding techniques. They are bounded which is interesting for computation and quantization on fixed point DSP. Belonging to the frequency domain, they have good interpolation properties useful for very low bit rate vocoders. Moreover their ordering property is a stability condition for the short term synthesis filter. An intrinsically reliable and fast algorithm (the Split Levinson algorithm) is used to compute these parameters[1]. The computation process includes the Split Levinson algorithm and the bisection method. Each CLSP parameter is computed independently of the others. Before transmission to the channel, the

CLSPs have to be quantisezed. Different quantization schemes can be used to perform the coding process : each CLSP can be independently quantized using scalar quantization or the CLPs can be vector quantized.

1.1.1 Scalar Quantization

The first version of our coder included a non uniform quantization. Each CLSP was coded on 3 bits, giving 30 bits per frame (frame length : 20 ms, 1500 bits/s for CLSP coding) . In that scheme the coder had a nominal bit rate of 5.45 kbits/s.

1.1..2 Vector Quantization

To reduce this nominal bit rate to the standard of 4.8 kbits/s, a CLSP vector quantization was chosen. To reach this aim the CLSPs have to be quantized on 17 bits per frame that is to say 850 bits/s. To make possible a real time processing we have firstly introduced a localization method to determine the nearest neighbour (nn) of an input vector to quantize[2]. This method exploits CLSP ordering property, using the Euclidean norm. This

L2 / L1	128	64	32	16	8	4	2
512	1.81	1.82	1.84	1.87	1.95	2.16	2.37
256	1.83	1.84	1.86	1.89	1.97	2.17	2.39
128	1.86	1.87	1.89	1.92	2.00	2.20	2.42
64	1.94	1.95	1.96	1.99	2.07	2.27	2.48
32	2.12	2.13	2.14	2.17	2.24	2.43	2.63
16	2.39	2.40	2.41	2.44	2.51	2.68	2.87
8	2.70	2.71	2.72	2.74	2.80	2.96	3.14
4	3.07	3.08	3.09	3.11	3.16	3.31	3.47
2	3.47	3.47	3.48	3.50	3.55	3.69	3.84

Table 2 : Spectral distorsion average (dB)

is possible if we reorder the codebook according to the ascending order of the norms of its codewords which are stored to avoid their recalculation in the coding process. Such computation will increase the nn's computing time. In the coding process, the norm is used to localize L consecutive codewords. The determination of the search area of the nn is done by dichotomy by comparing the norm of the input vector to those of the codewords which are already stored. The number L is a power of 2. Secondly to reduce requirement storage, the CLSP vector is split into two parts. The first six CLSP are coded with 10 bits and the others with 7 bits. This means building two codebooks, the first containing 1024 codewords and the second 128. Table 2 shows the spectral distorsion average for different search areas in each codebook : L1 in the 1024 codewords codebook, L2 in the 128 codewords codebook. For search areas L1 strictly greater than 32 and L2 strictly greater than 8, spectral distortion average at most increases of 0.2 dB. Such values are acceptable to real time implementation.

1.2 LTP Analysis

After the first short term prediction filter, a second one is used to remove periodic speech signal redundancy. It allows the current signal to be predicted from a linear combination from the past version signal. In our applications, in order not to increase the bit number too much, only a first order predictor is used. We only compute two parameters, the pitch p and the first order long term predictor coefficient b. Pitch p is equivalent in sample number to a delay in the range [2,..,18 ms]. It is also the ratio between sampling and fundamental frequency. For periodic signals, the predictor parameter b value is close to one. A simplified method, called "Correlation-peak-picking", is used to compute those parameters. The input speech, after removing short term redundancies, gives the glottal signal, e(n). This is used to compute LTP parameters p and b. In this method p only is the value which maximizes the correlation function.

1.3 Algebraic Codebook

The last part of the Code Excited Linear Prediction, the codebook, involves a huge amount of computation in its basic scheme. The synthesized analysis procedure consists in finding the waveform innovation in the codebook which is optimum to a subjective criterion. Each waveform (fig. 2) is scaled by an optimal gain factor and processed through the short and long term inverse filters.The difference between original and synthesized speech is processed through the perceptual weighting filter. Then the best waveform is chosen to minimize perceptual error signal energy.

To achieve a real time processing we use :

- a modified basic scheme structure [3]. The purpose of this modification only is to remove computation redundancies. The previous design can be split into two parts. One part, memoryless, is computed once a frame and a second part containing the memory process have to be computed at each excitation update, that is to say four times a frame.

- a ternary codebook each sample of which belongs to the set [-1,0,+1]. The codebook only contains 1024 waveforms of 40 samples, thus waveform index is coded on 10 bits. Such a codebook does not need to be stored, waveforms are generated from binary index development. According to the bit value the sample is set to +1 (bit value = 0), to -1 (bit value = 1). To obtain a 40 sample waveform we add three zero samples after each previous one.

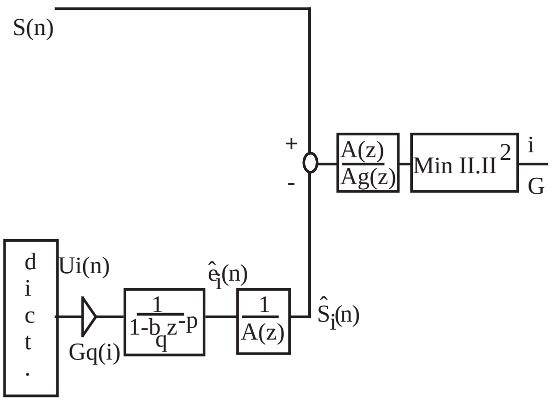

fig 2 : CELP basic scheme

2 TCELP DECODER

Compared to the coder, the TCELP decoder (fig. 3) has a very low computational complexity. It only includes in the reverse order the three coder stages.

- the codebook
- the long term filter
- the short term filter

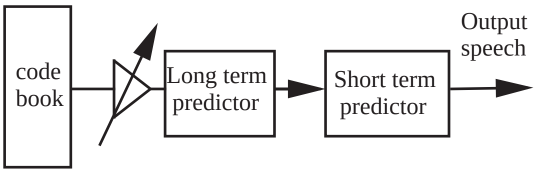

fig. 3 : TCELP decoder

3 REAL TIME PROCESSING

This coder (Scalar quantizer version) has been implemented on a DSP 56001 Motorola (clock frequency : 27 MHz) and runs in real time on our testing bench.

3.1 Coder

The speech signal is segmented in frames having a length of 20 ms. Therefore all the computing process must be achieved in less than 20 ms. It includes on the DSP implementation the three following tasks :

3.1.1 Acquisition task

using a fast interrupt, the input speech is sampled at 8 kHz and digitized by an Analog to Digital Converter(ADC).

3.1.2 Background job

the input stream (160 samples coded on 16 bits : 2560 bits per frame) is compressed to design the 23 coder parameters (96 bits). The first stage computes and quantizes the CLSP using firstly the Split Levinson algorithm [1] and secondly the fast quantization algorithm [2] with L equal to 64 for the six first CLSP codebook and 32 for the second codebook. The second stage looks after fundamental frequency in the range [56,...,500] Hz. This is done twice a frame, that is to say each 10 ms. After removing long and short term correlation, the residual signal is modelled by a waveform stored in a codebook containing 1024 waveforms. This excitation signal is renewed 4 times a frame. It is the most important part of the computational time. The real time processing is achieved using a ternary codebook.

3.1.3 Transmission task

Using a long interrupt, the compressed bit stream is sent to the decoder via the Serial Communication Interface (SCI) and the data transmission line.

3.2 Decoder

In real time environment, the speech decoder includes three tasks :

3.2.1 Reception task

This task wait for the synchronization word before reading the frame code over the transmission channel.

3.3.2 Background job

With the frame code, the 23 parameters are restored, thus allowing filter design and speech synthezisation.

3.3.3 Speaking task

Using fast interrupts, the speech samples are written via the SSI interface on the Digital-Analog-Converter (DAC), at 8 kHz, to drive the speakers.

4. PERFORMANCES

4.1 Signal to noise ratio

These performances had been computed in C language on workstation. Comparison had been performed with the coder running with the scalar quantizer (SQ) and the same coder running with the vector quantizer (VQ) for different values of the search areas L1 and L2. Table 3 summarizes the performances. According to search areas lengths,VQ (850 bits/s) performances are close to SQ (1500 bits/s) ones.

Scalar Quantizer	6.53			
Vector Quantizer L2 / L1	128	64	32	16
512	6.45	6.45	6.46	6.45
256	6.45	6.44	6.47	6.48
128	6.49	6.49	6.48	6.46
64	6.31	6.31	6.29	6.34
32	6.27	6.27	6.37	6.44

Table 3 : RSB(dB),
comparison between SQ and VQ

4.2 Time processing

As stated previously, the computational time must be lower than 20 ms. It was easy to measure on our testing-bench (SQ scheme only). The whole computational, for the coder is lower than 18 ms and for the decoder lower than 1.5 ms on a DSP56001 (clock frequency 27 Mhz). Comparison between SQ and VQ computation procedures has been performed (on DSP simulator) , VQ procedure only slightly increases computational time. Table 4 summarizes in MIPS computational requirements.
For the L1=64 and L2=32 computational time of CLSPs vector quantization is 2.5 times greater than CLSPs scalar quantization one.

4.2 Memory requirement

The main difficulty for The VQ utilization is the memory requirement. TCELP coder with CLSPs scalar quantization requires close to 9 Kwords, the design using CLSPs Vector Quantization needs more than 16.5 Kwords

in MIPS	Computational Requirement
CODER (whole)	12.5
DECODER (whole)	1
CLSPs SQ	0.05
CLSPs VQ L1=64 L2=32	0.117
CLSPs VQ L1=256 L2=64	0.37

Table 4 : Computational Requirement

CONCLUSION :

The coder first version including CLSP scalar quantization already runs in a digital underwater acoustic phone designed in our university(#)[3].The method reducing the complexity of the CLSPs vector quantization, introduced here, after integration in the TCELP system, running on DSP 56001 Motorola, will lead us to a bit rate standard (4.8 kbit/sec). The VQ design leads to a slight decrease of the signal to noise ratio and to a slight increase in computational time. The memory requirement remains the main drawback to the VQ development.

REFERENCES :

[1] A. Goalic and S. Saoudi," An intrinsically reliable and fast algorithm to compute the LSP in low bit rate CELP coding", ICASSP 95, Detroit, USA , Vol 1, pp 728-731, May 1995.
[2] N. NAJA "Construction de dictionnaires et quantification vectorielle pour les codeurs de parole LSP-CELP", thèse de doctorat de l'Université de Rennes 1, Juin 1994.
[3] A. GOALIC "Étude et réalisation de Codeur/décodeur de parole à bas débit pour la Télephonie numérique acoustique sous-marine", thèse de doctorat de l'Université de Bretagne Occidentale, Juin 1994.