REAL TIME 800 BITS/S LINEAR PREDICTIVE VOCODER
BASED UPON VECTOR QUANTIZATION

Denis ROCHETTE (*), Gilles MATHEVON (*), Alain ALBARELLO (*)

ABSTRACT

This paper describes both simulation and real time implementation of an 800 bits/s LP vocoder using vector quantization. A summary of the software algorithms is presented, including the codebook generation by a threshold algorithm and the fast nearest neighbour search algorithm. The hardware configuration using the DSP TMS 32010 is described. Informal listening tests and Diagnostic Rhyme Tests indicate that the speech quality at 800 bits/s is very close to that achieved at 2400 bits/s with LPC-10 standard vocoders.

INTRODUCTION

Over the past ten years, linear prediction (ref 1) has made possible the design of low complexity, rather good quality, and fully compatible vocoders (using the LPC-10 2400 bits/s standard algorithm, ref 2). However transmission at low bit rates over perturbed HF channels requires to protect the digitized speech with an error detecting and correcting code. This is not possible with the 2400 bits/s vocoders since very few bits are available for this purpose. On the other hand, the speech signal can be digitized at very low bit rates with vector quantization of the spectral information (ref 3). With this technique, the 10 reflexion coefficients are usually encoded using a 10-bit codebook rather than the 41-bit scalar quantization at 2400 bits/s. Thus speech can be digitized at rates as low as 800 bits/s (ref 4) which enables transmission at 2400 bits/s with redundancy 3. Hence implementation of an 800 bits/s vocoder is of great interest in many military applications provided that the speech quality is very close to that of standard LPC-10 vocoders.

Vector quantization requires a training algorithm and a fast nearest neighbour search algorithm for real time encoding of the the LP filters: these algorithms are summarized in the next section (complete description is available in ref 5). Then the hardware configuration is described. Finally, performance results are presented.

ALGORITHMS

A few years ago, it was shown (ref 6) that a simple threshold algorithm could design as efficient codebooks of LP filters as those designed with the well-known LBG algorithm (ref 7), from the average distortion viewpoint. Such an algorithm has been developed as follows: given a training

(*) THOMSON-CSE/DTC, CelluleTraitement du Signal
B.P. 156, 92231 GENNEVILLIERS CEDEX
set and a distance threshold $S$, the first training vector is chosen as the first codeword. Then, assuming that $N$ training vectors have been processed and have generated $P$ codewords, the $(N+1)$th training vector will be chosen as the $(P+1)$th codeword only if the distance to his nearest neighbour among the $P$ previous codewords is less than $S$.

The likelihood ratio is used as the distortion measure since it can be fast computed directly from the autocorrelation coefficients. The LPC-10 standard algorithm (basis for the 800 bits/s vocoder) was performed on phonetically balanced sentences pronounced by 10 speakers (5 males, 5 females), providing 15000 training vectors of 11 autocorrelation coefficients. A 10-bit codebook (1024 codewords) was designed using the threshold algorithm.

Because of the high computation cost of a full search algorithm, a fast nearest neighbour search algorithm was designed. First of all, the codebook is binary tree structured up to level 5. Each node of the tree is associated with the equation of an hyperplane (in the normalized autocorrelation space). The choice between the two branches starting at that node depends on the sign of the equation. Level 5 is reached after 5 such decisions. This level is composed of 32 nodes, each of them related to the numbers of 32 codewords. From every sub-level are selected 8 "super codewords" that will be called centroids. Then the 24 numbers of "ordinary" codewords are replaced by the numbers of nearest centroids of neighbour sub-levels. At last, the 40 nearest codewords of centroids are computed. The algorithm can be summarized as follows:

Step 1: progression to level 5 of the binary tree (5 signs of hyperplanes to compute).
Step 2: first full search among the 32 (8 + 24) centroids whose numbers belong to the sub-level reached at step 1.
Step 3: second full search among the 40 nearest codewords of the centroid found at step 2.

This algorithm requires to compute only 72 likelihood ratios and the signs of 5 hyperplanes. In 70% of the cases, it provides the nearest codeword (in the sense of a full search); in 90% of the cases, it provides one of the 3 nearest codewords.

For each 22.5 ms frame, the spectral information is coded with 10 bits. Differential coding of pitch and gain operating on blocks of 3 frames enables a transmission rate of 800 bits/s (Table 1).

HARDWARE

The hardware of the vocoder (fig. 2) consists of 3 standard boards designed at Thomson-CSF/DTC for signal processing applications:
- 1 processor board (using the DSP Texas Instruments TMS 32010).
- 1 conversion board (300-3300 Hz analog band-pass filtering, A/D and D/A conversions, 8 KHZ sampling).
- 1 memory board (the codebook storage requires 23 K words of the 32 K EPROM words available in this board).

For testing, analysis and synthesis are performed back to back by that single processor unit. Full duplex transmission between two units have also been achieved.

The overall processing time (analysis and synthesis) of a 22.5 ms speech frame is 20 ms. Only 2 ms are required for vector quantization of the spectrum.

RESULTS

The intelligibility of the 800 bits/s vocoder was subjectively evaluated by the Diagnostic Rhyme Test (DRT). The DRT involved 3 speakers (2 males, 1 female). The average score was 90.6 % which means a slight degradation of the intelligibility compared to the standard LPC-10 2400 bits/s vocoder (94.1 %) which was implemented with the same hardware configuration.

According to many informal listening tests, the speech quality at 800 bits/s is very close to that achieved at 2400 bits/s (the difference is slight but noticeable).

Theses results demonstrate that vector quantization can be implemented in real time with equipment of low complexity, and that speech quality comparable to standard LPC-10 2400 bits/s can be achieved at very low bit rates, thus enabling many applications for HF transmission.

ACKNOWLEDGEMENT

This work was partly supported by Direction des Recherches, Etudes et Techniques (DRET) de la Direction Generale de l'Armement.

REFERENCES

5. D. ROCHE'TTE, "Etude et realisation d'un vocodeur a
dictionnaire LPC 800 bits/s". These Docteur-Ingenieur,
Institut National Polytechnique, Grenoble (1986).
6. M. DABOUZ & L. MICLET, "Experiences en transmission de la
parole a faible debit par vocodeur a classification",
Quantizer Design", IEEE Transactions COM-28, pp. 84-95
(1980).

![Block diagram of the 800 bits/s vocoder](image1)

![Hardware configuration](image2)

<table>
<thead>
<tr>
<th>2400 BITS/S</th>
<th>800 BITS/S</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1 FRAME)</td>
<td>(3 FRAMES)</td>
</tr>
<tr>
<td>SPECTRUM</td>
<td>41</td>
</tr>
<tr>
<td>PITCH</td>
<td>7</td>
</tr>
<tr>
<td>GAIN</td>
<td>5</td>
</tr>
<tr>
<td>SYNCHRO.</td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td>54</td>
</tr>
</tbody>
</table>

**Delay:** 7 FRAMES (4 + 3) = 157.5 ms

Table 1. Allocation of bits for 2400 and 800 b/s vocoders