The present invention relates to a coding method and a word coder known as an analytical coder with linear prediction. The invention relates more particularly to methods and word coders of the kind specified with excitation by excitation vector, often referred to by the English language abbreviation CELP, which must be distinguished from the coding methods of multi-pulse excitation linear prediction analytical coding (MPLPC) an example of which is given in the document EP-A-0 195 487, to which reference can be made.

Analytical coding by linear prediction with vectorial excitation provides an advantageous solution to the problem of word transmission over a narrow band channel, for example, transmission between and to moving vehicles over a channel of 12.5 kHz, which reduces the output available to approximately 8 kbits/sec; in this latter case the output assigned to the transmission of the parameters representing the word signal is reduced to approximately 6 kbits/sec, since a portion of the global output must be assigned to the transmission of an error correction code.

Word coders are already known having linear prediction and vectorial excitation which can be used with a low binary output, normally between a quarter of a bit and a half bit per word sample. An embodiment of such a coder can be found more particularly in the Article by SCHROEDER and ATAL "Code Excited Linear Prediction (CELP); high quality speech at very low bit rates", proc. ICASSP, March 1985.

Fig. 1 shows a basic circuit of such a coder 10. The word signal is applied to the coder via a digitizing chain. In the embodiment shown in Fig. 1, the chain comprises, starting from a microphone 12, a low-pass filter 14 limiting the pass band to approximately 4000 Hz, and a sampler-coder 16. The sampler takes word

samples at a rate which is, for example, 8 kHz and provides successive samples, regrouped by decoders occupying time windows of a predetermined duration, for example, 20 msec.

The coder 10 converts the word signal into a coded signal having a lower rate of bits and transmitted to the transmission equipment by a multiplexer 18 which receives for each pattern the indexes $\underline{k}$ of the optimum excitation vectors $c_k$, the associated gains $G_k$ and coefficients identifying prediction parameters for each of the blocks forming the pattern, each occupying a sub-window.

The coder shown by way of example in Fig. 1 uses synthesis analysis: the word spectrum in each window is modelled by a linear predictive filter whose coefficients are variable in time. The residual signal, obtained by subtraction, forms the subject of a vectorial quantification, using a wave shape dictionary. Referring to Fig. 1, the dictionary 20 contains K+1 excitation vectors $c_k$ (with k=0, ..., k, ..., K) and is applied to an amplifier 22 of gain $G_k$.

Normally the excitation vectors stored in the dictionary 20 are selected either empirically, taking into account statistical data on the language, or at random, or else on the basis of classical binary numerical codes, such as the Golay codes.

The Article by SCHROEDER et al mentioned hereinbefore proposes, for example, a dictionary containing 1024 excitation vectors each formed by 40 samples. This number of vectors lies between the minimum, below which excitation would be poorly represented, and the maximum, beyond which the number of bits left free would be inadequate to transmit the parameters of the predictors.

The output of the amplifier 22 is applied to a predictive synthesis filter formed by a long-term

predictor filter 24, intended to introduce the periodicity of the long-term signal, and a short-term predictive filter 26. The output $\hat{S}n$ of the predictive filter, which represents an estimation synthesis of the word signal, is applied to the subtractive input of a subtracter 28 which receives at its additive input the sampled and digitized word signal Sn.

Once the respective transfer functions 1/B(z) and 1/A(z) if the filters 24 and 26 have been calculated and quantified, the coding operation consists in determining the optimum innovation sequence $c_k$ and the gain $G_k$ for each word pattern by a process of synthesis analysis. For each of the coding sequences $c_k$ the synthesis signal obtained $\hat{S}_k$ is compared with the signal of origin S and the difference signal obtained in the subtracter 28 is processed in a perceptual weighting filter 30 having a transfer function W(z) whose function is to attenuate those frequencies for which the errors have less importance from the perceptual aspect and on the other hand to amplify those frequencies for which the errors have more importance from the perceptual aspect.

A circuit 32 searches for the coding sequence for which the energy contained in the weighted error signal $e_k$ for a sub-window is minimal; this sequence is selected for the current block, whereafter the optimum gain $G_k$ is calculated.

Classically the function A(z) of the short-term predictive filter 26 takes the form:

$$A(z) = 1 - \sum_{i=1}^{p} a(i).z^{-i} \qquad (1)$$

In this formula, which uses the classical z notation, the coefficients a(i) form the linear prediction parameters. Their number is generally between 8 and 16 for windows of 20 msec.

As for the transfer function B(z), this can take the form $1-bz^{-T}$ and introduce a delay T of between 40 and 120 samples.

The perceptual weighting filter 30 has a transfer function W(z) which generally takes the form:

$$W(z) = A(z)/A(z/\gamma) \text{ with } \gamma = 0.8 \quad (2)$$

In spite of its advantage, the coding method set forth hereinbefore cannot in practice be put into effect in real time, because of the enormous volume of calculations required to search for the optimum innovation sequence (i.e., the optimum excitation vector) by K+1 successive loop passes, each pass forming the filtering of an excitation, by filters having coefficients variable in time.

A CELP coding method as set forth in the preamble of Claim 1 is also known (IEEE Journal on selected areas in communications, Vol. 6, No. 2, February 1988, pages 353-363); it is an object of the present invention to provide a coding method with linear prediction and excitation by coding vectors of the kind specified which meets practical requirements more satisfactorily than the prior art methods, more particularly reducing by at least one order of magnitude the volume of calculations to be performed for the coding of a segment.

To this end the invention proposes more particularly a word coding method with linear prediction and vectorial excitation as set forth in the characterizing part of Claim 1.

Since each coding sequence is formed by a number of equi-distance pulses separated by zeros, advantageously binary ones - i.e., use is made of an excitation by sequences of regular pulses (RPCELP) -, the duration of searching for the optimum sequence is reduced to a very considerable extent, more particularly if a suitable selection is made of the characteristics of the percept-

ual weighting filter.

Other features of the invention are defined in Claims 2 to 6.

The invention will be better understood from the following description of particular exemplary non-limiting embodiments thereof, with reference to the accompanying drawings, wherein:

- Fig. 1 already mentioned, is a circuit diagram of a known word coder with linear prediction and vectorial excitation,

- Fig. 2 which is similar to Fig. 1, is a variant of the diagram showing a possible construction of the coder illustrated in Fig. 1, which can be simplified to form a first embodiment of the invention,

- Figs. 3, 4 and 5 are diagrams showing successive developments of the coder illustrated in Fig. 2,

- Fig. 6, which is similar to Fig. 5, shows in a more complete manner an embodiment of the invention which also reduces the volume of calculation,

- Fig. 7 shows a possible distribution of coding sequences in the dictionary, and

- Fig. 8 shows another possible make-up of the dictionary.

In the word coder shown diagrammatically in Fig. 2 (in which like elements to those in Fig. 1 have like references) a perceptual weighting filter 30, disposed at the output of the subtracter 28 (Fig. 1), is connected to the two input branches of the subtracter in the form of filters 34 and 36 having a transfer function $1/A(z/\gamma)$. A filter 33 of transfer function $A(z)$ and the filter 36, having the same transfer function as the filter 34, are therefore disposed in cascade on the branch assigned to the signal of origin $S(n)$.

The filtering of all the vectors by the synthesis filter, having a transfer function $1/A(z/\gamma)$, whose

coefficients vary in time, represents an enormous volume of calculations. This volume is very considerably reduced according to a first feature of the invention by adopting a perceptual weighting filter having a small number of coefficients fixed in time and selected as a function of the average characteristics of the word over a long interval of time. The perceptual weighting filter therefore has a transfer function W'(z) which can be written:

$$W'(z) = A(z)/C(z/\gamma)$$

where $C(z/\gamma)$ is the transfer function of a short-term word predictor, for example of the form:

$$C(z/\gamma) = 1 - \sum_{i=1}^{2} a(i).\gamma^{i_z - i} \qquad (3)$$

The transfer functions of the components 34 and 36 (Fig. 2) therefore become $1/C(z/\gamma)$.

Another embodiment of the invention, which can be combined with the first one, is shown more clearly by considering the successive conversions made to the circuit illustrated in Fig. 2 to arrive thereat.

First of all, as indicated in Figs. 3 and 4, the contribution of the store to the long-term predictive filter 24, having a transfer function $1/B(z)$, and to the weighted short-term predictive filter, having a transfer function $1/A(z/\gamma)$, is subtracted from the signal of origin, which has undergone weighting to obtain a signal $x_n$, before the start of searching in the vector dictionary 20. This operation is performed (Fig. 3) by means of a subtracter (38) which receives solely the store component of the long term predictive filter 24.

Thus, during the procedure of searching for the optimum vector, each vector $\vec{c}_k$ is treated solely by the weighted synthesis filter 34.

It will now be shown how the volume of calculation

can be further considerably reduced, with reference to Fig. 4. In Fig. 4 each of the filters 34 and 36 is shown broken down into a filter 34a or 36a of transfer function $1/A(z/\gamma)$, without store, and a filter 34b or 36b corresponding solely to the contribution of the store terms.

During the search for the optimum vector $\vec{c}_k$, each vector $\vec{c}_k$ amplified with the gain $G_k$ is no longer treated except by the weighted synthesis filter without store $1/A(z/\gamma)$ which delivers at the output a signal $z(n)$. If a tilde is used to identify the magnitudes without store and the following notation is adopted:

$r$ - the residual signal after subtraction of the effects of the long-term predictor 24,

$x$ - the original signal whose long-term redundancy has been removed in the subtracter 38 and which has been weighted by $W(z)$,

$z_k$ - the synthesized signal

$x^o$ and $z^o$ - the contributions of the stores of the filters to the calculation of $x$ and $z$,

we can write:

$$x = Hr + x^o$$

$$\tilde{x} = Hr$$

$$z_k = G_k . H\, c_k + z^o$$

The operation of filtering by the filter 34a, without store, is expressed hereinabove by the convolution of two finite sequences, represented by the product of a matrix and a vector:

$$\tilde{z}_k = G_k\, H\, c_k \qquad (4)$$

where H is a lower triangular matrix LxL (L being the common length of the sequences) whose elements are drawn from the pulse response $h(i)$ of $1/A(z/\gamma)$, having the form:

$$H = \begin{vmatrix} h(0) & 0 & 0 & \ldots\ldots\ldots & 0 \\ h(1) & h(0) & 0 & \ldots\ldots\ldots & 0 \\ h(2) & h(1) & h(0) & \ldots\ldots\ldots & 0 \\ & & & & \\ & & & & \\ h(L-1) & h(L-2) & h(L-3) & \ldots\ldots\ldots & h(0) \end{vmatrix} \quad (5)$$

and which merges with that of $1/A(z/\gamma)$.

The vector x' at the input of the subtracter 28, after the subtraction of the store effects, can itself be written:

$$x' = Hr + x^o - z^o$$

The energy of the weighted error $E_k$ for the vector of index k (with $0 \leq k \leq K$) can be written:

$$E(K) = ||x' - z_k||^2 = ||x' - G_k H c_k||^2 \quad (6)$$

The process of searching for the optimum innovation sequence (index k of vector $\vec{c}_k$ and amplification gain $G_k$ comprises two stages which derive from equation (6), taking into account the known fact (J.P. ADOUL et al "Fast CELP coding based on algebraic codes", Proc. ICASSP, April 1987) that minimising the energy $E_k$ amounts to maximising a scalar product Pω:

- searching for the index k for which the scalar product Pω(k) is maximum:

$$P\omega(k) = x'^t H c_k)/||H c_k|| \quad (7)$$

- calculation of the corresponding gain $G_k$:

$$G_k = P\omega(k)/||H c_k||$$

Clearly, it is quicker to calculate a scalar product than to search for a Euclidian distance, so that the circuit shown in Fig. 3 already in itself enables the volume of calculation to be reduced.

The following stage of the step consists in eliminating the store terms - i.e., the operations shown

diagrammatically at 34a and 36a -, to arrive at the construction shown in Fig. 5.

As in the case of Fig. 2, an important simplification consists in substituting fixed synthesis filters having a transfer function $1/C(z/\gamma)$ for the filters 34a and 36a having a function $1/A(z/\gamma)$, something which again amounts to adopting a perceptual weighting filter having the form $W'(z) = A(z)/C(z/\gamma)$. All that must then be done is to perform a repetitive filtering operation using 34a, to the extent to which the excitation vectors are stored in the dictionary 20, on the one hand in the prefiltered state to apply them directly toward the circuit for maximising the scalar product 38, and on the other hand in the original form for application to the gain (Gk) amplifier 22. The simplification is shown immediately by a comparison with the classical processes of searching for the minimum.

Still another way in which the invention can be put into effect uses a modified criterion for evaluating the error to be minimised. The patterns of samples each occupying a window are applied in succession; consequently, the pulse response of the weighted synthesis filter for a pattern (or a block) takes place on the following pattern (or following block). To eliminate this effect, use is made of the damping of the filters, to whose input not a sequence formed solely by L samples is applied, but a sequence formed by L samples and J zeros, J being so selected that the pulse response of the synthesis filter $W(z)/A(z)$ is practically zero after J samples. A value J=10 is as a rule adequate to enable the damping of the filters to eliminate the representative terms of their store. In that case the pulse response matrix becomes a rectangular "band" type matrix with (L+J)xL terms of the type:

$$H = \begin{vmatrix} h(0) & 0 & 0 & \ldots\ldots\ldots & 0 \\ h(1) & h(0) & 0 & \ldots\ldots\ldots & 0 \\ h(2) & h(1) & h(0) & \ldots\ldots\ldots & 0 \\ \\ h(J-1) & h(J-2) & \ldots\ldots & h(0)0\ldots\ldots & 0 \\ \\ 0 & & & & \\ \\ 0 & & & & h(0) \\ \\ 0 & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & & h(j-1) \end{vmatrix} \qquad (10)$$

The matrix $H^tH = R$ is then a Toeplitz symmetrical matrix constructed from the autocorrelation R(i) of the pulse response h(n). $H^t$ denotes the transposed value of H.

The store error which appears in the representative equation of x' is then small enough to be considered to be zero, and the equation (7) can be written

$$P\omega(k) = (x^t\ H\ c_k)/||H\ c_k|| = y^t\ c_k/||H\ c_k|| \qquad (11)$$

The vector $y^t = r^t\ h^t\ H$ can be precisely calculated once per pattern by a filtering operation, using an adaptive filter whose coefficients are the autocorrelation terms R(i).

To put this process into effect in the case of a word signal which has been sampled at 8 kHz and whose samples are distributed in patterns of 160 samples each occupying 20 msec, more particularly each pattern, after filtering at 33 (Fig. 5) can be broken up into four

blocks of L=40 samples, which are successively applied to the filter 36a, in each case followed by J=10 zeros.

A will then be calculated for each pattern, whilst k and $G_k$ will be calculated for each block.

One particularly advantageous method in this case consists in using sequences of pulses having a length L and a regular structure formed by q equi-distant pulses separated by D-1 zeros, the first pulse occupying one of the positions 0 to D-1, and the number of sequences being such that all these positions are successively occupied. In this way a satisfactory representation can be given of the phase information in the excitation signal. Fig. 7 shows by way of example four sequences (for k=0, 1, 2 and 3) which are identical, except that they correspond to D=4 different phases. It may be considered that the dictionary is formed by a basic set of K/D sequences, with a zero phase and three successive shifts - i.e., a total of K sequences.

Excitation by regular excitation sequences reduces the number of operations to be performed, since many of the products to be effected are zero, one of the factors being a zero whose position is known for each sample. Calculations can be further simplified by making the sequences solely binary samples which can have only the values +1, -1 (and 0), as shown in Fig. 8, since in that case all the sequences contain the same energy; searching for the optimum sequence is performed with purely scalar products and amounts to searching for that binary vector which gives the best result. In this respect it may be noted that the document EP-A-0 195 487 relates to a MPLPC coding method wherein in succession an optimum pulse phase is determined, whereafter a search is made for the optimum amplitude of all the sequence-forming pulses amongst discrete values, quantified, for example, over 3 bits.

In the case of the modified criterion and excitation by regular sequences, and more particularly in the case of sequences formed by binary samples, and with the supplementary condition that the autocorrelation is normalised and has zero terms whose distance apart corresponds to the non-zero samples of the sequences, the terms H $c_k$ all become equal and we have:

$$||H\ c_k||^2 = ||d_m||^2 \qquad (12)$$

where $d_m$ denotes one of the sequences (to the number of K/D) resulting from the decimation of the components of the K vectors by elimination of the zeros; the sequence $d_m$ for $0 \leq K \leq 3$ is given in Fig. 7 by way of example.

If the sequences are normalised, the searching procedure is limited to searching for that sequence for which the scalar product $P(k) = y^t.c_k$ is maximum.

The conditions required for the applicability of formula (11) can more particularly be obtained:

- either by adopting a fixed filter R such that R(iD) is zero for $i > 0$,
- or by adopting a filter with variable coefficients, but whose finite pulse responds (FPR) is truncated for indexes of samples higher than D.

In that case the coder has the basic construction shown in Fig. 6. One single filtering operation is performed on the word signal pattern by filter 33. The tested sequence $c_k$, in a form which no longer needs to be prefiltered, is applied to circuit 32 for calculating the scalar product $c_k{}^t.y$ and determining the maximum, for which an index selection order is delivered to 40. The sequence $c_k$ amplified at 22 is applied to long-term predictor 24, represented with a single coefficient $\underline{b}$. The term $\underline{r}$ is formed by subtracting the output of the long-term predictor 24 from the output of the filter 34 on the word path, in subtracter 38. Filter 42, which

receives the residue r, has a fixed response R(z) represented by a symmetrical Toeplitz matrix.

Searching for the optimum vector can then be performed by a reduced number of multiplication and addition operations, with the reservation that the response may be truncated if the filter is variable, for example, by the following step if the regular excitation vectors are binary:

- determination of the phase giving a maximum value to M(p):

$$M(p) = \sum_{i=0}^{q-1} |y(p+iD)|$$

- then, from the vectors having the phase adopted, selection of the vector $d_m$ such that

$y^t.c_k = M(p)$

that is to say:

$d_m(i)$ = sign of $y(p+iD)$

for $i = 0, \ldots, q-1$

Once the optimum vector has been selected, the gain $G_k$ to be adopted is deduced directly therefrom, since $||H\ c_k||^2$ is equal to a constant value q, whatever the value of k may be, in the case of binary vectors which all have the same norm.

This process reduces the number of calculations required in a ratio which is typically of three orders of magnitude approximately in comparison with the conventional CELP process, whatever the length L selected for the word blocks may be.

The data to be transmitted by the multiplexer 18 will be:

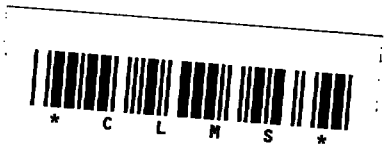
- the single coefficient b and the period T (corresponding to the periodicity of the word signal) of the

long-term predictive filter 24, once or several times per window, the coefficients $\underline{a}$ of the filter 33, having a transfer function A(z), once per window,

- the index $\underline{k}$ of the optimum vector and the corresponding gain $G_k$, once per block corresponding to a sub-window of, for example, 40 samples.

With a view to transmission the gain $G_k$ is quantified in a quantifier 46. Each signal pattern is broken down into a number of blocks; an intermediate store 48 must be interposed between the components 33 and 44.

It should moreover be noted that since excitation is binary and regular, it is not very sensitive to transmission errors: an error modifying the value of a bit modifies the vector only locally. The small number of phase bits can be protected by a correctional code.




## CLAIMS

1. A word coding method, with linear prediction and vectorial excitation, allowing the coding of word signals put in the form of digitized samples distributed in patterns, wherein:

a signal pattern is represented on the one hand by prediction parameters and on the other hand by a succession of excitation vectors contained in a dictionary (20) and by amplification gains ($G_k$) of said vectors, the vectors adopted being determined by searching (32) for the minimum energy of an error signal obtained by subtracting each vector in its turn, which has previously be subjected to filtering, from the pattern of the word signal; and,

before subtraction, subjecting each word signal pattern to a short-term analytical filtering (33) and to a perceptual weighted synthesis filtering (36) and subjecting the amplified vector to a long-term predictive filtering (24) and to the same perceptual weighted synthesis filtering (34,36) as the word signal,

characterized in that all the excitation vectors are formed by the same number of equi-distant pulses separated by zeros.

2. A method according to Claim 1, characterized in that the pulses separated by zeros are binary.

3. A method according to Claims 1 or 2, characterized in that each search (32) for the minimum energy of the error signal is performed by subjecting to filtering (34,36 or 34a,36a) an assembly comprising, in addition to the real samples of a block forming a fraction of the pattern, zero samples in an adequate number for the predictive filtering pulse response corresponding to the last real sample to be substantially damped, the filtering (34,36 or 34a,36a) being performed without storing from one block to the other.

4. A method according to Claims 1 or 2, characterized in that after being subjected to a short-term analytical predictive filtering A(z), each word signal pattern is applied to the additive input of a subtracter (38) which receives at its subtractive input the contribution from the store of the long-term predictive filter (24),

the output of the subtracter is subjected to a filtering (42),

the scalar product (32) of the filtered output of the subtracter and of each unamplified sequence is calculated in its turn, while searching for that sequence for which the scalar product is the maximum.

5. A method according to Claim 4, characterized in that the filtering (42) is with fixed coefficients.

6. A word coding method, with linear prediction and vectorial excitation, allowing the coding of word signals put in the form of digitised samples distributed in patterns, wherein:

each block forming a fraction of signal pattern is represented by one of the vectors contained in a dictionary (20), by a vector amplification gain ($G_k$) and by prediction parameters, the vector adopted being determined by searching the minimum energy of an error signal obtained by subtracting each vector in its turn, previously subjected to a filtering, from the word signal pattern; and,

before subtraction, subjecting each word signal pattern to a short-term analytical filtering A(z), characterized in that

the result of the subtraction (38) is subjected to a perceptual weighted synthesis filtering (36) with coefficients fixed in time and

the excitation vector, stored in a precalculated form and filtered, is subjected to a filtering by a perceptual weighted synthesis filter (34) $1/\tilde{C}(z/\gamma)$ which

* C L M S *

is fixed and without store, all the excitation vectors being formed by the same number of equi-distant pulses, separated by zeros.