Publicly	**PAS 0001-7**
Available	**Version: 1.0.3**
Specification	**Date: 15 October 1997**

Source: TETRAPOL Forum　　Work Item No: 0001

Key word: TETRAPOL

TETRAPOL Specifications;
Part 7: Codec

TETRAPOL FORUM

TETRAPOL Secretariat

Postal address: BP 40 78392 Bois d'Arcy CEDEX - FRANCE
Tel.: +33 1 34 60 55 88 - Fax: +33 1 30 45 28 35

Contents

Foreword

This document is the Publicly Available Specification (PAS) of the TETRAPOL land mobile radio system, which shall provide digital narrow band voice, messaging, and data services. Its main objective is to provide specifications dedicated to the more demanding PMR segment: the public safety. These specifications are also applicable to most PMR networks.

This PAS is a multipart document which consists of:

Part 1	General Network Design
Part 2	Radio Air interface
Part 3	Air Interface Protocol
Part 4	Gateway to X.400 MTA
Part 5	Dispatch Centre interface
Part 6	Line Connected Terminal interface
Part 7	**Codec**
Part 8	Radio conformance tests
Part 9	Air interface protocol conformance tests
Part 10	Inter System Interface
Part 11	Gateway to PABX, ISDN, PDN
Part 12	Network Management Centre interface
Part 13	User Data Terminal to System Terminal interface
Part 14	System Simulator
Part 15	Gateway to External Data Terminal
Part 16	Security
Part 18	Base station to Radioswitch interface
Part 19	Stand Alone Dispatch Position interface

1. Scope

This part gives a general description of the TETRAPOL speech CODEC which is based on the RPCELP algorithm.

It describes the algorithm used to transform an input frame of 20 ms speech samples into an encoded block of 120 bits and the corresponding inverse transform. The sampling frequency is 8 kHz leading to a bit rate of 6 kbit/s for the encoded bitstream.

2. Normative references

This PAS incorporates by dated and undated references, provisions from other publications. These normative references are cited at the appropriate places in the text and the publications are listed hereafter. For dated references, subsequent amendments to or revisions of any of these publications apply to this PAS only when incorporated in it by amendment or revision. For undated references the latest edition of the publication referred to applies.

[1] PAS 0001-1: "Tetrapol Specifications; General Network Design".

[2] PAS 0001-2: "Tetrapol Specifications; Radio Air interface".

3. Abbreviations

3.1 Abbreviations

For the purposes of this PAS the following abbreviations apply:

CELP	Code Excited Linear Prediction
RPCELP	Regular Pulse Code Excited Linear Prediction
LP	Linear Prediction
LTP	Long Term Predictor (or Long Term Prediction)
FIR	Finite Impulse Response
COD	Speech Coder.
DEC	Speech Decoder.
PAS	Public Available Specification
PMR	Private Mobile Radiocommunications
PCM	Pulse Code Modulation
SP	Stochastic Parameters
ST	System Terminal

4. Principles of the TETRAPOL codec

This part is the description of the coding and decoding speech algorithm of TETRAPOL which is based on the CELP technology. In the following, the general CELP model, the RPCELP algorithm and its implementation for the TETRAPOL codec are described.

4.1 General principle of the CELP model

The codec is based on the code-excited linear predictive (CELP) coding model (figure 1). A 10th order linear prediction (LP), or short-term, synthesis filter is used which is given by:

$$\frac{1}{A(z)} = \frac{1}{1+\sum_{i=1}^{p} a_i z^{-i}} \qquad (1)$$

where a_i, $i = 1,\dots,p$ are the linear prediction (LP) parameters, and $p = 10$ is the predictor order. The long-term, or pitch, synthesis filter is given by:

$$\frac{1}{B(z)} = \frac{1}{1-b_0 z^{-T_0}} \qquad (2)$$

where T_0 is the pitch lag and b_0 is the pitch gain. The pitch synthesis filter is implemented using the so-called adaptive codebook approach.

In the CELP speech synthesis model, the excitation signal at the input of the short-term LP synthesis filter is constructed by adding two excitation vectors from adaptive and fixed codebooks. The speech is synthesized by feeding the two properly chosen vectors from these codebooks through the short-term synthesis filter. The optimum excitation sequence is chosen from a codebook using an analysis-by-synthesis search procedure, in which the error between the original and synthesized speech is minimized according to a perceptually weighted distortion measure.

The perceptual weighting filter used in the analysis-by-synthesis search technique is given by:

$$W(z) = \frac{A(z/\gamma_1)}{A(z/\gamma_2)}, \tag{3}$$

where $A(z)$ is generally the unquantized LP filter, and $0 < \gamma_2 < \gamma_1 \leq 1$ are the perceptual weighting factors. The weighting filter uses the unquantized LP parameters, while the formant synthesis filter uses the quantified ones.

The coder operates on speech frames of T_f ms corresponding to M samples at the sampling frequency of 8 kHz. The parameters of the CELP model (LP filter coefficients, adaptive and stochastic codebooks indices and gains) are extracted from the speech signal on each M samples. These parameters are encoded and transmitted. At the decoder, these parameters are decoded and the speech signal is synthesized by filtering the reconstructed excitation signal through the LP synthesis filter.

4.2 General principle of RPCELP algorithm

The RPCELP algorithm is a CELP-like algorithm with a fixed excitation codebook containing ternary vectors in which the pulses (± 1) are placed each D samples. By choosing this codebook structure and the values $\gamma_1 = 1.0$ and $\gamma_2 = 0.8$ for the perceptual weighting factors, we reduce significantly the computational load of the original CELP model and we obtain a good perceptual masking effect.

Figure 2 shows the structure of the RPCELP algorithm. To reduce the computational load, we first obtain the long term excitation parameters. These parameters are obtained, for each sub-frame, by minimising the following error expression:

$$E_{LTP} = \left\| \mathbf{e_f} \right\| = \left\| \mathbf{Hd} - b_0 \mathbf{Hd'_{T_0}} \right\|, \tag{4}$$

where $\mathbf{H}$ is the lower triangular Toeplitz convolution matrix, with $h(0)$ on the diagonal and $h(1),\ldots,h(N)$ on the lower diagonals (where N is the length of the sub-frame in samples), $\mathbf{d}$ the residual signal obtained by filtering the input speech by $A(z)$ and $\mathbf{d'_{T_0}}$ is the past quantized residual for the delay T_0. The value of b_0 that minimized the expression (4) is:

$$b_0 = \frac{\mathbf{d}^t \mathbf{H^t H d'_{T_0}}}{\left(\mathbf{d'_{T_0}}\right)^t \mathbf{H^t H d'_{T_0}}}. \tag{5}$$

By replacing the optimal value of b_0 in the expression (4) we obtain:

$$E_{LTP} = \left\| \mathbf{Hd} \right\|^2 - \frac{\left(\mathbf{d}^t \mathbf{H^t H d'_{T_0}}\right)^2}{\left(\mathbf{d'_{T_0}}\right)^t \mathbf{H^t H d'_{T_0}}}. \tag{6}$$

The expression (6) is minimized by maximizing the normalized correlation term:

$$C_{LTP} = \frac{\left(\mathbf{d}^t \mathbf{H^t H d'_{T_0}}\right)^2}{\left(\mathbf{d'_{T_0}}\right)^t \mathbf{H^t H d'_{T_0}}} . \tag{7}$$

The optimal delay T_0^{opt} that maximize the expression (7) is found by filtering all the vectors of the adaptive codebook (which contains the quantized past samples of the residual signal) using the matrix $\mathbf{H}$. To avoid this computational load a simplified version of the expression (7) is used:

$$C'_{LTP} = \frac{\left(\mathbf{d}^t \mathbf{H^t H d'_{T_0}}\right)^2}{\left(\mathbf{d'_{T_0}}\right)^t \mathbf{d'_{T_0}}} . \tag{8}$$

By this way only one filtering operation is performed i.e. the filtering of $\mathbf{d}$ by $\mathbf{H^t H}$. When the optimal delay T_0^{opt} is found, the optimal gain is obtained using the expression (5) and replacing T_0 by T_0^{opt} .

Then the long term parameters contribution is subtracted from the LP residual to obtain the target to be modelled by the short term excitation quantizer:

$$\mathbf{e} = \mathbf{d} - b_0 \mathbf{d'_{T_0^{opt}}} . \tag{9}$$

The stochastic parameters are obtained minimising the following expression:

$$C_{SP} = \frac{\left(\mathbf{e}^t \mathbf{H^t H c_k}\right)^2}{\mathbf{c_k}^t \mathbf{H^t H c_k}} \tag{10}$$

where $\mathbf{c_k}$ is the stochastic vector to be found. This expression is defined by an analysis close to the one used for LTP parameters. The optimal gain for the stochastic vector is defined by the expression:

$$G_k = \frac{\mathbf{e}^t \mathbf{H^t H c_k}}{\mathbf{c_k}^t \mathbf{H^t H c_k}} . \tag{11}$$

If the samples of the vectors $\mathbf{c_k}$ follow a particular rule, the computation load of C_{SP} can be significantly reduced. For this reason, the stochastic codebook contains ternary vectors in which the pulses (± 1) are placed each D samples. These D samples are equal to zero. The position of the first pulse different from zero (that is noted in the following « phase of the vector » or ph) should be into the interval [0, D -1]. We observe that if the $h(i)$ autocorrelation function noted $R(n)$ is equal to zero for the samples $n = i \cdot D + ph$ the denominator of expression (10) is simply equal to $QR(0)$ where Q is the number of pulses different from zero. The sign of each pulse is simply chosen by selecting the sign of the pulse at the position of the vector $\mathbf{x}$ obtained by calculating the expression:

$$\mathbf{x}^t = \mathbf{e}^t \mathbf{H^t H} . \tag{12}$$

By this way, the numerator of the expression (10) is maximized by searching the phase of the vector $\mathbf{c_k}$ that maximizes the term:

$$MAG(ph_k, D) = \sum_{i=1}^{Q-1} \left| x(iD + ph_k) \right| \tag{13}$$

where ph_k is the phase of the vector $\mathbf{c_k}$ and D the decimation factor of the codebook. The expression (13) is calculated for each k and each D and the vector $\mathbf{c_k}$ that maximizes the term:

$$C'_{SP} = \frac{\left(MAG(ph_k, D)\right)^2}{Q} \tag{14}$$

The last step is the calculation of the optimal stochastic gain according to the expression:

$$G_k = \frac{MAG(ph_k, D)}{Q}. \tag{15}$$

4.3 Implementation of the TETRAPOL codec

This section is the description of the implementation of the RPCELP algorithm used in TETRAPOL. The coder operates on speech frames of 20 ms corresponding to 160 samples at the sampling frequency of 8 kHz. The input samples are companded using a 8 bits/A-law format and transformed into a block of 120 bits. At the decoder, this block is decoded and the speech signal is synthesized leading to a frame of 160 samples converted to a 13 bits uniform PCM format.

Figure 3 shows a simplified block diagram of the TETRAPOL coder. The LP analysis of order $p = 10$ is performed on each frame of 160 samples, weighted by the rectangular window, using the classical autocorrelation method. The working frame is delayed by 10 ms. The LP parameters are quantized in the LAR domain, using 37 bits.

Each frame is divided into 3 sub-frames having a size of 56, 48, 56 samples each. For each sub-frame, the LP parameters are obtained by interpolation of the LAR coefficients of the current and previous frame using the following weights: (7/8, 1/8) (1/2, 1/2) (1/8, 7/8).

For each sub-frame, the LTP parameters (gain and optimal delay) are calculated. The algorithm used is described in the section 5.2. The lag resolution is 1/3. The lag and its fractional part are coded using 8 bits. The optimal gains for the long term predictor are encoded using 3 bits with an optimal quantizer.

The stochastic codebook is structured in 4 sub-codebooks each with a different decimation factor (D) respectively: 8, 12, 15 and N (where N is the size of the current sub-frame 56 or 48). Each decimation factor leads to a different number of pulses for each vector (Q) according to the following expression:

$$Q = \lfloor N / D \rfloor \tag{16}$$

where $\lfloor x \rfloor$ is equal to the largest integer less than or equal to x. Each pulse can be positive or negative. The possible Q and phase (ph) values corresponding to the decimation factors are shown in table 1.

Table 1. Decimation factor and number of pulses corresponding.

Decimation factor D	Number of pulses Q	Phase values ph
8	7 (or 6)	[0,...,7]
12	4	[0,...,11]
15	3	[0,...,14]
56 (or 48)	1	[0,...,55] (or [0,...,47])

Table 1. Decimation factor and number of pulses corresponding.

Each sub-codebook is searched to find the optimal codeword according to the algorithm described in section 5.2. The optimal vector, which is characterized by its pulse number (or decimation factor), phase and the sign of each pulse, is coded using 10 bits (or 9 bits for a 48 bits sub-frame) for the phase and the sign and 2 bits for the decimation factor. The optimal gain found with the algorithm described in section 5.2 is coded using 5 bits.

The bit allocation is summarized in table 2.

Parameters		**Number of bits/sub-frame**			**Number of bits/frame**
		sf1	**sf2**	**sf3**	
LP coefficients		-	-	-	37
LTP	lag	8	8	8	24
	gain	3	3	3	9
Stochastic	decimation	2	2	2	6
	signs+phase	10	9	10	29
	gain	5	5	5	15
Total		28	27	28	120

Table 2. Bit allocation of the 6 kbit/s TETRAPOL encoder

Figure 4 shows a simplified block diagram of the TETRAPOL decoder. This decoder extracts the different parameters described previously in this section from the 120 bits block and the speech signal is synthesized by filtering the reconstructed excitation signal through the LP synthesis filter.

5. Functional description of the encoder

The *LP analysis section* of the encoder comprises the following 5 sub-blocks:

Segmentation and scaling of the input signal (6.1)
Autocorrelation and bandwidth expansion (6.2)
Leroux-Gueguen algorithm (6.3)
Transformation of reflection coefficients to LARs (6.4)
Coding and decoding of LARs (6.5)

The *short term filtering section* comprises the following 4 sub-blocks:

Interpolation of LARs(6.6)
Transformation of LARs into reflection coefficients (6.7)
Temporal shift of the window (6.8)
Short term analysis filtering (6.9)

The *LTP section* comprises 6 sub-blocks each working on sub-segments (6.10) of the short term residual signal:

Calculation of the perceptual filter (6.11)
Calculation of the LTP parameters (6.12)
Coding and decoding of the LTP lags (6.13)
Coding and decoding of the LTP gains (6.14)
Long term analysis filtering (6.15)
Long term synthesis filtering (6.16)

The *excitation encoding section* comprises the following 4 sub-blocks:

Optimal stochastic vector selection (6.17)
Quantization of the stochastic parameters (6.18)
Gain control procedure (6.19)
Bit ranking in the bitstream (6.20)

5.1 Segmentation and scaling of the input signal

The speech signal $S(k)$ is divided into frames having a length of T_f = 20 ms (160 samples). An LP analysis of order $p = 10$ is performed for each frame. The input speech samples are in a 8 bit A-law companded format. They are expanded to the 13-bit uniform format.

5.2 Autocorrelation and bandwidth expansion

The first $p+1$ = 11 values of the autocorrelation function are calculated by:

$$\mathrm{Autocor}(k) = \sum_{i=1}^{159} S(i)S(i-k)\ , \qquad k = 0,1,\cdots,10 \tag{17}$$

A bandwidth expansion procedure is then performed for the autocorrelation function:

$$\mathrm{Autocor}'(k) = \mathrm{Autocor}(k) \cdot B(k) \qquad k = 0,1,\cdots,10 \tag{18}$$

The window used is a binomial window ensuring a band expand of 90 Hz. The exact values are given in table 3.

5.3 Leroux-Gueguen algorithm

The reflection coefficients are calculated using Leroux-Gueguen algorithm.

5.4 Transformation of the reflection coefficients into LARs

The reflection coefficients $r(i)$, i = 1,...,10, previously calculated, are in the range [-1,1]. Due to the favourable quantization characteristics, the reflection coefficients are converted into Log-Area-Ratios (LAR) which are defined as follows:

$$\mathrm{LAR}_{\mathrm{def}}(i) = \log_{10}\left(\frac{1+r(i)}{1-r(i)}\right) \tag{19}$$

In order to reduce the complexity and since it is the companding of this transformation that is of importance, the following segmented approximation is used:

$$\begin{array}{ll} \text{if } \left(\left|r(i)\right| < 0.675\right) & LAR(i) = r(i) \\ \text{if } \left(0.675 \le \left|r(i)\right| < 0.950\right) & LAR(i) = sign(r(i)) \cdot \left(2 \cdot \left|r(i)\right| - 0.675\right) \\ \text{if } \left(0.950 \le \left|r(i)\right| \le 1.00\right) & LAR(i) = sign(r(i)) \cdot \left(8 \cdot \left|r(i)\right| - 6.375\right) \end{array} \tag{20}$$

5.5 Coding and decoding of LARs

The Log Area Ratios $LAR(i)$ have different dynamic ranges and different distribution densities. For this reason, the transformed coefficients LARs are limited and quantized using different tables obtained for each of theses coefficients. A dichotomy algorithm is used in order to find the index $LAR_c(i)$ and the quantized value $LAR''(i)$ for each of the coefficients $LAR(i)$, i = 1,...,10.

5.6 Interpolation of LARs

To avoid spurious transient which may occur if the LP filter coefficients are changed abruptly, two subsequent sets of LARs are interpolated. The interpolation procedure is described in the following table:

k	$LAR_j^{'}(i)$
0...55	$0.875 \cdot LAR_{J-1}^{''}(i) + 0.125 \cdot LAR_J^{'}(i)$
56...103	$0.500 \cdot LAR_{J-1}^{''}(i) + 0.500 \cdot LAR_J^{'}(i)$
104...159	$0.125 \cdot LAR_{J-1}^{''}(i) + 0.875 \cdot LAR_J^{'}(i)$

Table 4. Interpolation rules for LARs coefficients.

where J is the number of the current analysis frame.

5.7 Transformation of LARs into reflection coefficients

The inverse transformation which leads to reflection coefficients is the following:

$$\begin{array}{lll} \text{if } \left(\left|LAR^{'}(i)\right| < 0.675\right) & r^{'}(i) = LAR^{'}(i) & \\ \text{if } \left(0.675 \le \left|LAR^{'}(i)\right| < 1.225\right) & r^{'}(i) = sign(LAR^{'}(i)) \cdot \left(0.5 \cdot \left|LAR^{'}(i)\right| + 0.3375\right) & (21) \\ \text{if } \left(1.225 \le \left|LAR^{'}(i)\right| \le 1.625\right) & r^{'}(i) = sign(LAR^{'}(i)) \cdot \left(0.125 \cdot \left|LAR^{''}(i)\right| + 0.796875\right) & \end{array}$$

5.8 Temporal shift of the window

The working window $s(i)$ and the analysis window $S(i)$ are temporally shifted by 10 ms. The following operations are then performed:

For $i = 0 \ to \ 79$
$s(i) = sprec(i)$

For $i = 80 \ to \ 159$
$s(i) = S(i - 80)$ (22)

For $i = 0 \ to \ 79$
$sprec(i) = S(i + 80)$

For the first frame, $sprec$ is set to 0.

5.9 Short term analysis filtering

The short term analysis filter is implemented using a lattice structure in order to obtain the short term residual signal $d(k)$:

For k = 0,..., $N-1$
$tmp_1[0,k] = s(k)$
$tmp_2[0,k] = s(k)$
For i = 1,...,10
$tmp_1[i,k] = tmp_1[i-1,k] + r^{'}(i) \cdot tmp_2[i-1,k-1]$ (23)
$tmp_2[i,k] = tmp_2[i-1,k-1] + r^{'}(i) \cdot tmp_1[i-1,k]$
$d(k) = tmp_1[10,k]$

5.10 Sub-segmentation

Each input frame of the short term residual contains 160 samples, corresponding to 20 ms. The long term correlation is evaluated three times per frame, which means that the current residual frame is divided into three sub-frames having a size of 56, 48 and 56 samples each. For convenience in the following, we note j = 0,1,2 the sub-frame number.

5.11 Calculation of the perceptual filter

For each of the three sub-frames, a perceptual filter is calculated. The coefficients of this FIR filter are the autocorrelation coefficients of the impulse response of the filter $1/A(z/\gamma)$, where $A(z)$ is the quantized version of the short term analysis filter and $\gamma = 0.8$. This procedure is implemented in four steps.

1. The first step is the evaluation of the impulse response of the filter $1/A(z/\gamma)$. The first 20 coefficients of the impulse response of $1/A(z)$ ($h_0(i)$, $i = 0,\cdots 19$), are obtained by filtering a unit impulse using the algorithm described in section 7.3. For each sub-frame, the reflection coefficients obtained for the second sub-frame (interpolation (1/2,1/2)) are selected to model $1/A(z)$.

2. The impulse response of $1/A(z/\gamma)$ ($h_0(i)$, $i = 0,\cdots 19$), is calculated by:

$$h(i) = \gamma^i \cdot h_0(i), \qquad i = 0,\cdots,19 \tag{24}$$

3. The third step is the calculation of the first 8 values of the autocorrelation function:

$$R(k) = \sum_{i=k}^{19} h(i)h(i-k), \qquad k = 0,\cdots,7 \tag{25}$$

4. The last step is the final calculation of the FIR perceptual filter, $R_0(i)$, i = 0,...,14, obtained by:

$$R_0(i) = R(|7-i|)/R(0), \qquad i = 0,\cdots,14 \tag{26}$$

5.12 Calculation of the LTP parameters

For each of the three sub-frames, a long term correlation lag T_{0_j} and an associated gain factor b_{0_j}, j = 0,...,2 are determined. The determination of these parameters is implemented in six steps. In the following, we note N the size of the used sub-frame (N = 56 or 48).

1. The perceptual filter $R_0(i)$, i = 0,...,14, is applied to each sub-frame of the short term residual signal $d(k)$, k = 0,..., N-1. This filtered signal will be noted $d_f(k)$, k = 0,..., N-1, and is obtained by:

$$d_f(k) = \sum_{i=0}^{14} R_0(i) \cdot d(k+7-i),$$

with: j-1 index of the previous sub-frame (27)

$$d(k+7-i) = 0 \qquad \text{for}(k+7-i) \geq N$$
$$d(k+7-i) = d_{j-1}(N+7+k-i) \qquad \text{for}(k+7-i) < 0$$

The *Lavltp* term, used by the gain control procedure, is also calculated:

$$Lavltp = \mathbf{d}^t \mathbf{d_f} \tag{28}$$

2. The second step is the evaluation of the cross-correlation term $R_j(\lambda)$ of the current sub-segment of short term residual signal $d_f(k)$, k = 0,..., N -1, and the previous samples of the reconstructed short term residual signal $d'(k)$, k = -104,...,-1.

$$R_j(\lambda)=\sum_{i=0}^{N-1} d_f(i)\cdot d'(i-\lambda)\ ,$$
$$\text{with: } \lambda = 20,\cdots,104 \qquad (29)$$
$$d'(i-\lambda)=d'(i-2\cdot\lambda) \qquad \text{if}(i-\lambda)\ge 0$$
$$d'(i-2\cdot\lambda)=d'(i-3\cdot\lambda) \qquad \text{if}(i-2\cdot\lambda)\ge 0$$

3. The third step is the evaluation of the energy $E_j(\lambda)$ for each sub-frame of the reconstructed short term residual signal $d'(k)$, k = -104,...,-1:

$$E_j(\lambda)=\sum_{i=0}^{N-1}\left(d'(i-\lambda)\right)^2\ ,$$
$$\text{with: } \lambda = 20,\cdots,104 \qquad (30)$$
$$d'(i-\lambda)=d'(i-2\cdot\lambda) \qquad \text{if}(i-\lambda)\ge 0$$
$$d'(i-2\cdot\lambda)=d'(i-3\cdot\lambda) \qquad \text{if}(i-2\cdot\lambda)\ge 0$$

4. The next step is to find the optimal lag T_{0_j}, calculated by:

$$C_j\left(T_{0_j}\right)=\max\left(R_j(\lambda)^2/E_j(\lambda)\right),\qquad \text{with } \lambda = 20,\cdots,104 \qquad (31)$$

   To reduce the computation load, the maximum of this expression is found by cross product, and the optimal values $R(j)$ and E_j are stored at each step of the selection.

5. This step is the evaluation of the expression $C_j(\lambda)$ for non integer values of the lag. This evaluation procedure is limited within the interval [l_{min},l_{max}], obtained according to:

   l_{min} = T_{0_j} -1 and l_{max} = T_{0_j} +1

   If (l_{min} < 20) then l_{min} = 20 and l_{max} = 21 (32)

   If (l_{min} > 104) then l_{min} = 103 and l_{max} = 104

   The reconstructed residual sample $d'(k)$, k = -104,...,-1, is then up-sampled by a ratio of 3 in order to obtain the sequences $d_1'(k)$ and $d_2'(k)$, k = -104,...,-1, corresponding to the interpolated signal with phase 1/3 and 2/3. Two interpolation filters $H_1(i)$ and $H_2(i)$, i = 0,...9, are used to calculate these signals (see table 5):

$$d_1'(k)=\sum_{i=0}^{9} H_1(i)\cdot d'(k+4-i)$$
$$d_2'(k)=\sum_{i=0}^{9} H_2(i)\cdot d'(k+4-i)$$
$$\text{with: } k = -104,\ldots,-1 \qquad (33)$$
$$d'(k+4-i)=0 \qquad \text{for}(k+4-i)\ge 0$$

The expressions $R_j(\lambda)$ and $E_j(\lambda)$ are estimated, for values of λ in the range [-l_{min}, l_{max}], by using the interpolated sequences $d_1'(k)$ and $d_2'(k)$. For the lag value $\lambda = l + 1/3$, $d_1'(k)$ is selected, whereas for the lag value $\lambda = l + 2/3$, it is $d_2'(k)$ which is selected (l is an integer within the interval [-l_{min}, l_{max}]). At last, the optimal lag is evaluated according to:

$$C_j(T_{0_j}) = \max\left(R_j(\lambda)^2 / E_j(\lambda)\right) \qquad \text{with } \lambda = l_{min}, l_{min}+1/3, \cdots, l_{max}-1/3, l_{max} \tag{34}$$

6. The last step is the evaluation of the following expression:

$$S_j(T_{0_j}) = \sum_{i=0}^{N-1} d_l'(i - T_{0_j}) \cdot d_{l_f}'(i - T_{0_j}) \, , \tag{35}$$

with: l = 0,1,2 depending on the lag value T_{0_j}

$d_0'(k) = d'(k) \qquad k = -104, \cdots, -1$

$d_{l_f}'()$ is the filtered version of $d_l'()$ with the perceptual filter $R_0(i)$, $i = 0, \cdots, 14$

$S_j(T_{0_j})$ is needed to find the gain factor:

$$b_{0_j} = R_j(T_{0_j}) / S_j(T_{0_j}). \tag{36}$$

5.13 Coding and decoding of the LTP lags

The lag T_{0_j} can have values within the set of values (20, 20.33, 20.66, ..., 104.66) and so shall be coded using 8 bits with:

$$T_{c0_j} = 3\, T_{0_j} - 60 \tag{37}$$

At the receiving end, the decoding of these values will restore the actual lags:

$$T_{0_j} = 20 + T_{c0_j}/3. \tag{38}$$

5.14 Coding and decoding of the LTP gains

The long term prediction gains b_{0_j} are encoded with 3 bits each, according to the following rules:

$$\begin{array}{lll} \text{If } \left(b_{0_j} \le DLB(0)\right) & \text{then } b_{c0_j} = 0 \text{ and } T_{c0_j} = 255 & \\ \text{If } \left(DLB(i) < b_{0_j} \le DLB(i+1)\right) & \text{then } b_{c0_j} = i & \text{for } i = 0,\ldots,5 \\ \text{If } \left(b_{0_j} > DLB(6)\right) & \text{then } b_{c0_j} = 7 & \end{array} \tag{39}$$

where $DLB(i)$, i = 0,...,7 denotes the decision levels of the quantizer, and b_{c0_j} represents the coded gain value.

The decoding rule is:

$$\begin{array}{ll} \text{If } T'_{c0_j} = 255 & \text{then } b_{0_j}' = 0 \\ \text{else } b_{0_j}' = QLB(b_{c0_j}) & \text{for } j = 0,\ldots,3 \end{array} \tag{40}$$

where $QLB(i)$, i = 0,...,7 denotes the gain quantizing levels, and b'_{0_j} represents the decoded gain value. The quantization values, DLB and QLB are given in tables 6.

5.15 Long term analysis filtering

The short term residual signal $d(k)$, k = 0,..., N -1 is processed on each sub-frame. From each of the three sub-frames (j = 0,...,3) of short term residual samples, an estimate $d''(k)$, k = 0,..., N -1 of the long term contribution signal is subtracted to give the long term residual signal $e(k)$, k = 0,..., N -1:

$$e(k) = d(k) - d''(k) \qquad k = 0,..., N-1. \tag{41}$$

Prior to this subtraction, the estimated samples $d''(k)$ are computed from the vector of the adaptive codebook corresponding to the optimal lag T'_{0_j} and weighted with the sub-segment LTP gain b'_{0_j} ':

$$d''(k) = b'_{0_j} \cdot d'_l\left(k - T'_{0_j}\right) \qquad l = 0,1,2. \tag{42}$$

5.16 Long term synthesis filtering

The stochastic excitation signal $e'(k)$, k = 0,..., N -1 is calculated for each sub-frame. The estimate $d''(k)$, k = 0,..., N -1 of the signal is added to each sub-frame to update the adaptive codebook. The signal $d'(k)$, k = -104,...,-1 is calculated according to the following equations:

$$\begin{aligned} d'(-104+k) &= d'(N-104+k) && k = 0,...,103-N \\ d'(-N+1+k) &= e'(k) + d''(k) && k = 0,..., N. \end{aligned} \tag{43}$$

5.17 Optimal stochastic vector selection

For each sub-frame, the signal $e(k)$, k = 0,..., N is perceptually filtered using the coefficients $R_0(i)$, i = 0,...,14. The resulting signal is noted $x(k)$, k = 0,..., N. The $Lapltp$ term, used in the gain control procedure, is calculated by the following equation:

$$Lapltp = \mathbf{e}^t \mathbf{x}. \tag{44}$$

The optimal vector $\mathbf{c_k}$ is found by searching, for each decimation factor, the phase that maximizes the term:

$$MAG(ph_k, D) = \sum_{i=1}^{Q-1} \left| x(iD + ph_k) \right| \tag{45}$$

where ph_k is the phase of the vector $\mathbf{c_k}$ (ph_k = 0,..., $D-1$) and D the decimation factor of the codebook. At last, the chosen optimal vector is the one which maximizes the expression:

$$C'_{SP} = \frac{(MAG(ph_k, D))^2}{Q} \tag{46}$$

Then the optimal stochastic gain is calculated according to the following expression:

$$G_k = \frac{MAG(ph\ \ , D)}{Q}. \tag{47}$$

where ph, Q and D are respectively the phase, the number of pulses and the decimation factor of the optimal vector.

At the last step the signs of the pulses of the optimal vector are chosen equal to the signs of the signal $x(k)$ for $k = i \cdot D + ph$, i = 0,...,$Q-1$.

5.18 Quantization of the stochastic parameters

The phase and index of the optimal vector are jointly coded using the following procedure:

1. the optimal vector phase is coded using 3, 4, 4 or 6 bits depending on the decimation factor of the optimal vector
2. the index I_1 of the optimal vector is binary coded using 7 (or 6), 4, 3 or 1 bit with the following algorithm:

I_1 = 0
For i = 0,...,$Q-1$
    if $x(i \cdot D + ph) > 0$ (48)
        $I_1 = 2 \cdot I_1 + 1$
    else
        $I_1 = 2 \cdot I_1$

3. At last, the global index obtained by concatenation of the phase and index bitstreams is coded using at the maximum 10 bits (or 9 bits for the second sub-frame).

The optimal gain G_k is quantized using 5 bits according to algorithm of section 6.14. The decision table $DLBG$ and the quantization table $QLBG$ are given respectively in table 7 an table 8. The quantized gain is noted $G_1^{'}$ and the corresponding index is noted G_{1c}.

5.19 Gain control procedure

In order to emphasize the adaptive codebook contribution in the final excitation signal, the quantized gain $G_1^{'}$ is modified when the signal is highly predictable. The modification is performed using the following algorithm:

If ($Lapltp \leq 0$ or $Lavltp \leq 0$)
    $G_1^{'}$ = 0 and G_{1c} = 0
else if ($Lapltp < Lavltp$)
    tmp = $Lapltp / Lavltp$
    k =31; G_{opt} = $QLBG(31)$
    while $\left(G_{opt} / G_1^{'}\right)^2 > tmp$
        $k = k$ -1
        G_{opt} = $QLBG(k)$
    If tmp > 0.5   (49)
        k =31; G_{new} = $QLBG(31)$
        while $tmp^5 \cdot Lavltp / Q < G_{new}^2$
            $k = k$ -1
            G_{new} = $QLBG(k)$
    else
        k =31; G_{new} = $QLBG(31)$

while $Lavltp/(32 \cdot Q) < G_{new}^2$
$k = k$ -1
$G_{new} = QLBG(k)$
$G_1' = G_{new}$ and $G_{1c} = k$

The stochastic excitation signal is obtained by multiplying the optimal vector by this gain G_1'.

5.20 Bit ranking in the bitstream

The different parameters of the encoded speech and their individual bits have unequal importance with respect to the subjective quality. Before being submitted to the channel encoder, the bits have to be rearranged in order to protect only the twenty most important bits.

6. Functional description of the decoder

The *decoder section* comprises the following 5 sub-blocks:

Decoding of the excitation (7.1)
Long term synthesis (7.2)
Short term synthesis filtering (7.3)

6.1 Decoding of the excitation

The excitation signal is obtained after decoding of the parameters I_1, ph, D and G_1' using the following equations:

$\mathbf{e_r} = \mathbf{0}$
For i = 0,..., $Q-1$
$index = i \cdot D + ph$
if $[(I_1 >> Q-1-i) \& 1] \neq 0$ (50)
$e_r(index) = 1$
else
$e_r(index) = -1$

6.2 Long term synthesis

The $\mathbf{e_r}$ vector is used in the long term predictor as described in section 6.15 and 6.16 in order to obtain the short term residual vector $\mathbf{d_r}$.

6.3 Short term synthesis filtering

The LAR coefficients are decoded and interpolated as described in section 6.6. The reflection coefficients are obtained following the algorithm described in section 6.7. Then the output signal $s_r(k)$ is obtained by filtering the signal $d_r(k)$, k = 0,..., $N-1$ through the lattice implementation of the short term synthesis filter as describe below:

For k = 0,..., $N-1$
$tmp[0,k] = d_r(k)$
For i = 1,...,10
$tmp[i,k] = tmp[i-1,k] - r_r'(11-i) \cdot v_{10-i}(k-1)$ (51)
$v_{11-i}(k) = v_{10-i}(k-1) - r_r(11-i) \cdot tmp[i,k]$
$v_0(k) = tmp[10,k]$
$s_r(k) = tmp[10,k]$

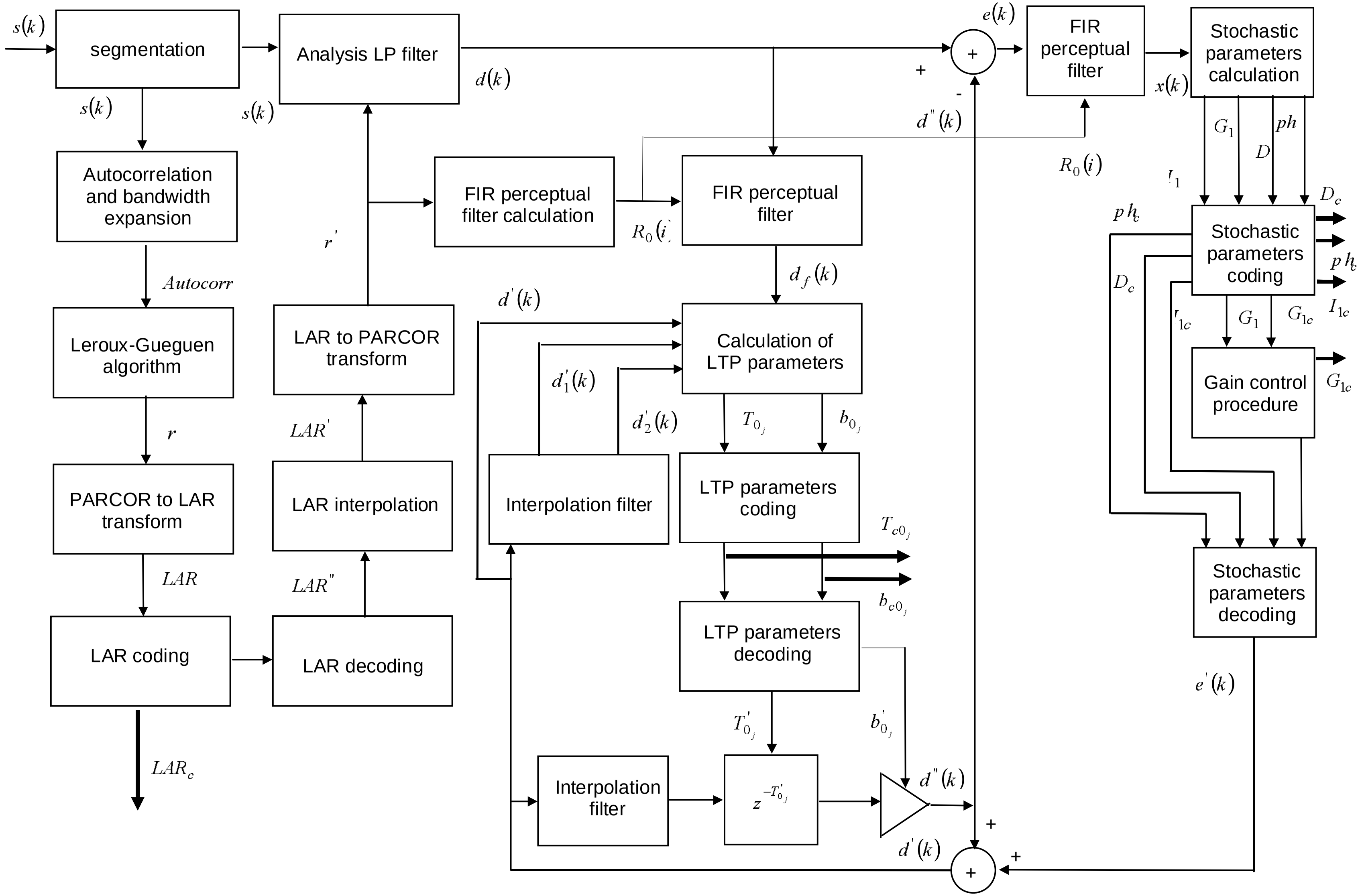

Figure 1. Simplified block diagram of the TETRAPOL coder.

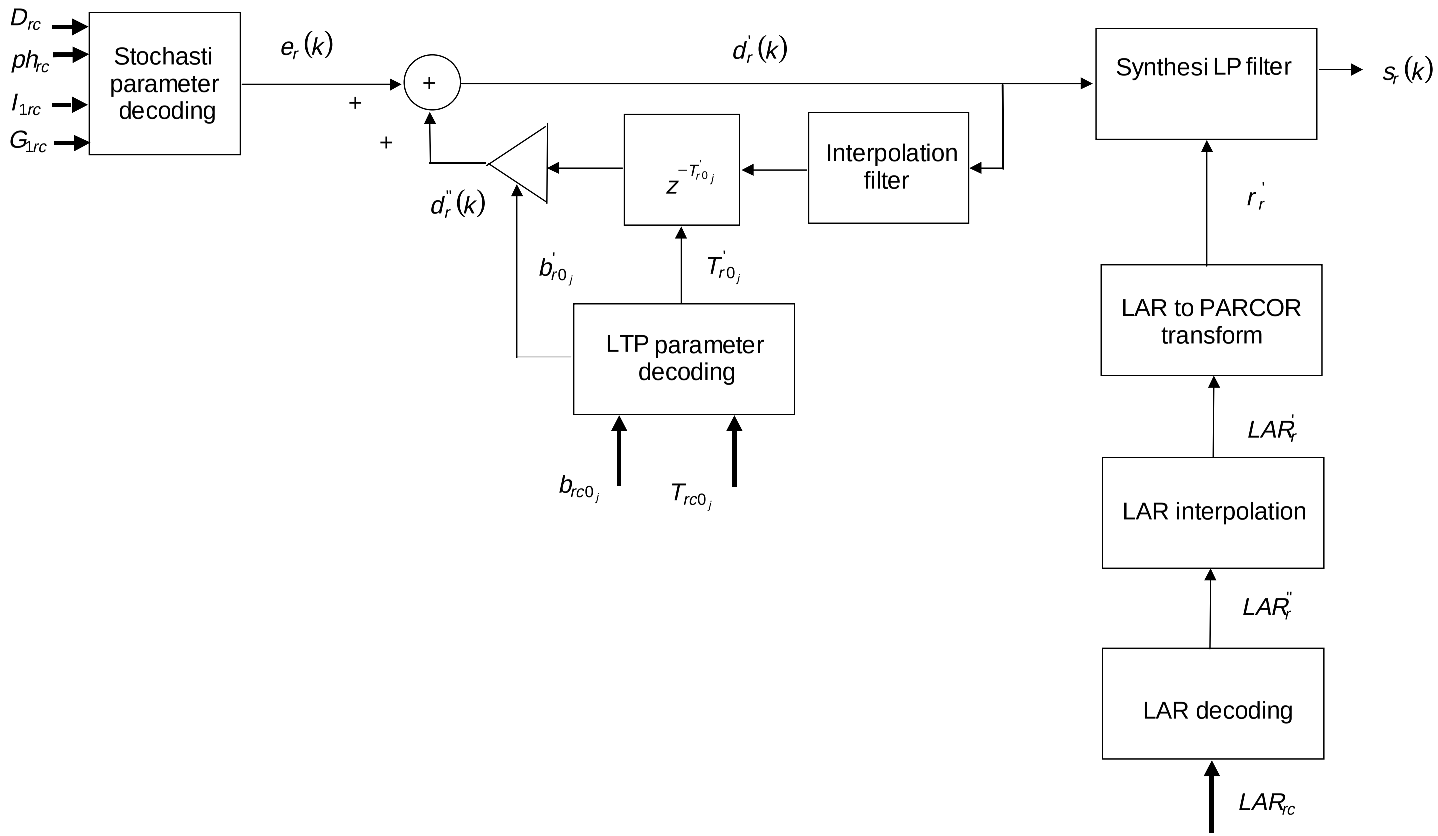

Figure 2. Simplified block diagram of the TETRAPOL decoder.

k	$B(k)$	k	$B(k)$
0	1.0000	6	0.9839
1	0.9995	7	0.9781
2	0.9982	8	0.9716
3	0.9959	9	0.9641
4	0.9928	10	0.9559
5	0.9888		

Table 3. Binomial window for the bandwidth expansion

i	$H_1(i)$	$H_2(i)$
0	521 / 8192	484 / 8192
1	-677 / 8192	-616 / 8192
2	968 / 8192	847 / 8192
3	-1694 / 8192	-1355 / 8192
4	6775 / 8192	3387 / 8192
5	3387 / 8192	6775 / 8192
6	-1355 / 8192	-1694 / 8192
7	847 / 8192	968 / 8192
8	-616 / 8192	-677 / 8192
9	484 / 8192	521 / 8192

Table 5. Interpolation filters

i	$DLB(i)$	$QLB(i)$
0	0.150	0.30
1	0.375	0.45
2	0.525	0.60
3	0.670	0.74
4	0.800	0.86
5	0.920	0.98
6	1.060	1.14
7		1.23

Table 6. Decision levels and quantization values for the LTP lag

i	$DLBG(i)$	i	$DLBG(i)$
0	3	16	255
1	7	17	319
2	15	18	383
3	23	19	447
4	31	20	511
5	39	21	639
6	47	22	767
7	55	23	895
8	63	24	1023
9	79	25	1279
10	95	26	1535
11	111	27	1791
12	127	28	2047
13	159	29	2559
14	191	30	3071
15	223		

Table 7. Decision values of the stochastic excitation gain

i	$QLBG(i)$	i	$QLBG(i)$
0	0	16	239
1	5	17	287
2	11	18	351
3	19	19	415
4	27	20	479
5	35	21	575
6	43	22	703
7	51	23	831
8	59	24	959
9	71	25	1151
10	87	26	1407
11	103	27	1663
12	119	28	1919
13	143	29	2303
14	175	30	2815
15	207	31	3583

Table 8. Quantization values of the stochastic excitation gain

7. History

Document history		
Date	**Status**	**Comment**
08 November 1995	First version	Version 0.0.1
29 March 1996	Update following review (editorial)	Version 0.1.1
30 April 1996	TETRAPOL Forum approval	Version 1.0.0
31 July	Formatting	Version 1.0.1
16 December 1996	Converted to word 6	Version 1.0.2
15 October 1997	Completion of the document.	Version 1.0.3