

JT-G723.1

マルチメディア通信伝送のための 5.3 および 6.3kbit/sデュアルレート 音声符号化方式

Dual rate speech coder for multimedia
communications transmitting at 5.3 and 6.3kbit/s

第 2.3 版

2006 年 11 月 27 日制定

社団法人
情報通信技術委員会

THE TELECOMMUNICATION TECHNOLOGY COMMITTEE

本書は、（社）情報通信技術委員会が著作権を保有しています。

内容の一部又は全部を（社）情報通信技術委員会の許諾を得ることなく複製、転載、改変、転用及びネットワーク上での送信、配布を行うことを禁止します。

目 次

| | |
|----------------------------|----|
| <参考> | 5 |
| 1. 概 要 | 7 |
| 1.1 範 囲 | 7 |
| 1.2 ビットレート | 7 |
| 1.3 可能な入力信号 | 7 |
| 1.4 遅 延 | 7 |
| 1.5 音声コーデックの記述 | 7 |
| 2. 符号器原理 | 8 |
| 2.1 概 要 | 8 |
| 2.2 フレーム化 | 8 |
| 2.3 高域通過フィルタ | 9 |
| 2.4 L P C分析 | 10 |
| 2.5 L S P量子化器 | 10 |
| 2.6 L S P復号器 | 11 |
| 2.7 L S P補間器 | 12 |
| 2.8 ホルマント聴覚重み付けフィルタ | 12 |
| 2.9 ピッチ推定 | 13 |
| 2.10 サブフレーム処理 | 13 |
| 2.11 高調波ノイズシェーピング | 13 |
| 2.12 インパルス応答の計算 | 14 |
| 2.13 零入力応答とリングング補正 | 15 |
| 2.14 ピッチ予測器 | 15 |
| 2.15 高レート励振 (MP-MLQ) | 15 |
| 2.16 低レート励振(ACELP) | 17 |
| 2.17 励振信号復号器 | 19 |
| 2.18 ピッチ情報の復号 | 20 |
| 2.19 メモリの更新 | 21 |
| 2.20 ビット割当て | 21 |
| 2.21 符号器の初期化 | 23 |
| 3. 復号器の原理 | 24 |
| 3.1 概 要 | 24 |
| 3.2 L S P復号器 | 24 |
| 3.3 L S P補間器 | 24 |
| 3.4 ピッチ情報の復号 | 25 |
| 3.5 励振信号復号器 | 25 |
| 3.6 ピッチポストフィルタ | 25 |
| 3.7 L P合成フィルタ | 27 |
| 3.8 ホルマントポストフィルタ | 27 |
| 3.9 利得調整ユニット | 28 |
| 3.10 フレーム補間の取り扱い | 28 |
| 3.10.1 L S Pの補間 | 28 |
| 3.10.2 残差の補間 | 29 |

| | |
|---|----|
| 3.11 復号器の初期化..... | 29 |
| 4. ビット列のパッキング..... | 30 |
| 5. ANS I Cコード..... | 32 |
| 6. 記号解説..... | 33 |
| 付属資料A..... | 35 |
| A 1. 本付属資料の規定範囲..... | 35 |
| A 2. VADの記述..... | 36 |
| A2.1 適応イネーブルフラグの計算..... | 36 |
| A2.2 逆フィルタ..... | 37 |
| A2.3 フィルタ後エネルギーの計算..... | 37 |
| A2.4 雑音レベルの計算..... | 37 |
| A2.5 しきい値の計算..... | 37 |
| A2.6 VADの決定..... | 38 |
| A2.7 VADハングオーバーの付加..... | 38 |
| A2.8 VADの初期化..... | 38 |
| A 3. CNGの概要..... | 38 |
| A 4. CNG符号器の詳細..... | 39 |
| A4.1 フレームの自己相関関数の計算..... | 39 |
| A4.2 現フレームタイプ $Ftyp_t$ の計算..... | 40 |
| A4.3 平均エネルギーの量子化..... | 41 |
| A4.4 S I Dパラメータの計算と符号化..... | 42 |
| A4.5 CNG励振の計算..... | 43 |
| A4.6 L S Pの補間と更新..... | 44 |
| A4.7 C O D-CNGの初期化..... | 45 |
| A 5. 復号器部の記述..... | 45 |
| A5.1 D E C-CNGの記述..... | 45 |
| A5.2 CNGに関するフレーム消失補償..... | 47 |
| A5.3 D E C-CNGの初期化..... | 48 |
| A 6. ビット列のパッキング..... | 48 |
| A 7. 記号解説..... | 49 |
| A 8. ビット精度の固定小数点演算Cソースコード..... | 50 |
| 付属資料B..... | 51 |
| 1. 概 要..... | 51 |
| 2. アルゴリズム記述..... | 51 |
| 3. ANS I Cコード..... | 51 |
| 付属資料C..... | 53 |
| 1. 概 要..... | 53 |
| 1.1 範 囲..... | 53 |
| 1.2 ビットレート..... | 53 |
| 1.3 遅 延..... | 53 |
| 1.4 チャネルコーデックの記述..... | 53 |
| 1.5 制御プロトコルTTC標準J T-H 2 4 5へのインタフェース..... | 54 |
| 2. チャネル符号器..... | 54 |

| | | |
|-------|-----------------------|----|
| 2.1 | 有音モードのためのチャンネル符号器の適応化 | 54 |
| 2.1.1 | 利得インデックスの伸長 | 54 |
| 2.1.2 | LPC再構成 | 54 |
| 2.1.3 | 伝送路に適応化されたビット列のフォーマット | 57 |
| 2.2 | 有音モードのためのビット感度分類 | 59 |
| 2.3 | SIDフレームのためのビット感度分類 | 61 |
| 2.4 | CRC符号器 | 62 |
| 2.5 | 畳み込み符号器 | 62 |
| 2.5.1 | チャンネルコーデックのビットレート割り当て | 63 |
| 2.5.2 | トレリス適応の終了 | 65 |
| 2.5.3 | 情報ビット列の符号化 | 65 |
| 2.5.4 | 構成ビットの符号化 | 65 |
| 3. | チャンネル復号器 | 66 |
| 3.1 | 構成ビットの復号化 | 66 |
| 3.2 | 有音モードのための音声復号器の適応 | 66 |
| 3.2.1 | 利得インデックスの圧縮 | 66 |
| 3.2.2 | LPCの構成 | 67 |
| 3.3 | 誤り解析 | 67 |
| 3.4 | 音声パラメータのフォーマット | 67 |
| 3.5 | 誤り補償 | 70 |
| 3.6 | 誤り解析フラグの能力 | 70 |
| 4. | 固定小数点演算Cソースコード | 70 |
| 付録1 | 用語対照表 | 71 |
| 付録2 | 用語解説 | 74 |

<参考>

1. 英文記述の適用レベル

適用レベル：E3

本標準の本文および付属資料の図表に英文記述を含んでいる。

2. 国際標準等との関連

本標準は、1996年3月に承認されたITU-T勧告G723.1に準拠したものである。また、本標準の付属資料A, B, Cは1996年11月に承認されたITU-T勧告G723.1 ANNEX-A, B, Cに準拠したものである。本標準第2.1版は、2002年2月にITU-T SG16にて承認されたインプリメンターズガイドに基づく誤記等を修正したものである。本標準第2.2版は、2002年10月にITU-T SG16にて承認されたインプリメンターズガイドに基づく誤記等を修正したものである。

本標準本体および付属資料A, B, Cは、2006年4月にITU-T SG16でAAPに進むことが合意され、2006年5月に承認されたITU-T勧告G.723.1(05/06)に準拠したものである。

3. 上記国際勧告に対する追加項目

3.1 オプション選択項目

なし

3.2 ナショナルマター項目

なし

3.3 その他

- (1) 本標準は、上記ITU-T勧告に対し、先行している項目はない。
- (2) 本標準は、上記ITU-T勧告に対し、追加した項目はない。
- (3) 本標準は、上記ITU-T勧告に対し、削除した項目はない。
- (4) 本標準は、上記ITU-T勧告に対し、変更した項目はない。

3.4 原勧告と章立ての構成の相違はない。

上記国際標準との章立ての構成の相違はない。

4. 改訂の履歴

| 版数 | 制定日 | 改版内容 |
|-------|-------------|--|
| 第1版 | 1996年11月27日 | 制定 |
| 第2版 | 1997年4月23日 | 付属資料A, B, Cの追加 |
| 第2.1版 | 2002年4月25日 | 2.9節の誤記等を修正 |
| 第2.2版 | 2002年11月27日 | A.4.2節の誤記等を修正、およびCコード改定に関する記述の追加 |
| 第2.3版 | 2006年11月27日 | ITU-T勧告G.723.1改定に伴う改定(4章 Table 4-1の誤記訂正) |

5. 工業所有権

本標準に関わる「工業所有権の実施の権利に係る確認書」の提出状況は、TTCホームページでご覧になります。

6. その他

(1) 参照している勧告、標準等

TTC標準 : JT-G711, JT-G712, JT-H245, JT-H324

ITU-T勧告

(2) 本標準の付属資料A、およびBにて参照しているCコードの改定に関して、2002年10月に承認されたITU-T勧告G.723.1に対するインプリメンターズガイドに記述がある。

1. 概要

1.1 範囲

本標準は、超低ビットレートのマルチメディアサービスに関する音声あるいは他のオーディオ信号成分を圧縮するために使用する符号化方式について記述している。本コーデックの設計において、想定された主なアプリケーションは、ITU-T勧告H. 324ファミリの一部を成す超低ビットレートテレビ電話である。

1.2 ビットレート

本コーデックのビットレートは、5.3 と 6.3kbit/s の2種類である。高ビットレートはより高音質になる。低ビットレートも高音質であるが、システム設計者に付加的な柔軟性を提供する。上記の2種類のビットレートは符号器と復号器において必須である。2種類のビットレート間の切り換えは、30ms のフレーム毎に可能である。間欠的な伝送と無音区間を雑音で埋めるという可変レート操作に対するオプションも可能とする。

1.3 可能な入力信号

本コーデックは限られた複雑さで上記ビットレートで高音質な音声を実現できるように最適化してある。音楽や他のオーディオ信号は音声のように忠実には表現されないが、本コーデックで圧縮伸長することは可能である。

1.4 遅延

本コーデックは音声あるいは他のオーディオ信号を 30ms フレーム単位で符号化する。さらに、7.5ms の先読みがあるため、結果的に全演算遅延は 37.5ms となる。本コーデックの実現装置化と演算により付加される全遅延は以下のものがある。

- (1) 符号器と復号器に必要なデータ処理時間
- (2) 通信リンク上の伝送時間
- (3) 多重化プロトコルに対する付加的なバッファリング遅延

1.5 音声コーデックの記述

本標準の音声符号化アルゴリズムは、ビットイグザクトな固定小数点演算で記述される。5章で示されるANSI Cコードは、本標準の必須部分であり、ビットイグザクトな固定小数点演算を反映させたものである。2章および3章でそれぞれ示される符号器と復号器の数学的記述は、他の固定小数点演算でも実装が可能であるが、その場合本標準に準拠しない実装のコーデック装置になってしまう可能性がある。したがって、5章のCコードによるアルゴリズム記述と2章および3章の数学的記述との間に相違点がある場合、5章のCコードによるアルゴリズム記述が優先される。このCコードおよびテストシーケンス（非網羅的であるが）は、日本ITU協会から入手可能である。

2. 符号器原理

2.1 概要

本コーデックでは、符号器の入力信号として、まず初めにアナログ入力信号を電話帯域フィルタ（TTC標準JT-G712）に通し、次に8000Hzで標本化し、さらに16ビット均一PCM変換することにより得られたデジタル信号を使用するように設計されている。復号器の出力は同様の方法でアナログ信号に戻される。64kbit/s PCMデータに関するTTC標準JT-G711に記述されているような入力/出力特性は、符号化の前に16ビット均一PCMに変換するべきものであり、または復号化の後に16ビット均一PCMから特定のフォーマットに変換するべきものである。符号器から復号器へのビット列は本標準の中で規定される。

符号器は合成による分析(A-b-S)法を用いた線形予測符号化にもとづいており、聴覚重み付き誤差信号が最小になるように実行される。符号器では240サンプルを1つのブロック（フレーム）として扱う。標本化周波数が8kHzの場合、1ブロックは30msとなる。各ブロックはDC成分を除去するためにまず高域通過フィルタに通され、60サンプルごとに4つのサブフレームに分割される。サブフレームごとに未処理の入力信号から10次の線形予測係数(LPC)フィルタが計算される。最後のサブフレームのLPCフィルタは予測分割ベクトル量子化器(PSVQ)を用いて量子化される。量子化されていないLPC係数は短期聴覚重み付けフィルタを構成するために使用され、聴覚重み付き音声信号は入力フレームごとに短期聴覚重み付けフィルタを通すことにより得られる。

開ループピッチ周期 L_{OL} は、2つのサブフレーム（120サンプル）ごとに、重み付き音声信号から算出される。ピッチ推定は120サンプルのブロックごとに実行される。ピッチ周期の探索範囲は18から142サンプルまでである。

これ以後、音声はサブフレーム（60サンプル）ごとに処理される。

直前に計算された推定ピッチ周期を用いて、高調波ノイズシェーピングフィルタを構成する。LPC合成フィルタ、ホルマント聴覚重み付けフィルタと高調波ノイズシェーピングフィルタの結合は、インパルス応答を得るために使用される。インパルス応答は後の計算に使用される。

開ループピッチ予測器は、推定ピッチ周期 L_{OL} とインパルス応答を用いて計算される。ここでは5次のピッチ予測器が使用される。推定開ループ推定ピッチ近傍の差分値としてピッチ周期が計算される。ピッチ予測器の寄与分は初期ターゲットベクトルから差し引かれる。ピッチ周期と差分値は復号器に伝送される。

最後に、励振信号の非周期成分の近似をする。高ビットレートには最尤量子化型マルチパルス(MP-M LQ)励振が使用され、低ビットレートには代数的コードブック励振(ACELP)が使用される。

符号器のブロック図をFigure 2-1/JT-G723.1に示す。

2.2 フレーム化

| | | |
|-------------------|---------------------|--------------------------------|
| File : LBCCODEC.C | Procedure : main() | Reads 240 samples input frames |
| File : CODER.C | Procedure : Coder() | Performs sub-frame division |

符号器は連続的な音声サンプル $y[n]$ を240サンプルのフレーム $s[n]$ にバッファリングする事によって音声の処理を行う。推定ピッチの計算のため、フレームは120サンプルごとの2つに分割される。各部分はさらに2つに分割され、最終的に各フレームは60サンプルごとの4つのサブフレームに分割される。

2.3 高域通過フィルタ

| | | |
|-------------------|----------------------|---------------------------|
| File : UTIL_LBC.C | Procedure : Rem_Dc() | Performs high pass filter |
|-------------------|----------------------|---------------------------|

このブロックは入力信号 $s[n]$ からDC成分を除去する。フィルタの伝達関数を次式に示す。

$$H(z) = \frac{1 - z^{-1}}{1 - \frac{127}{128} z^{-1}} \quad (1)$$

このフィルタの出力は $x[n]_{n=0...239}$ である。

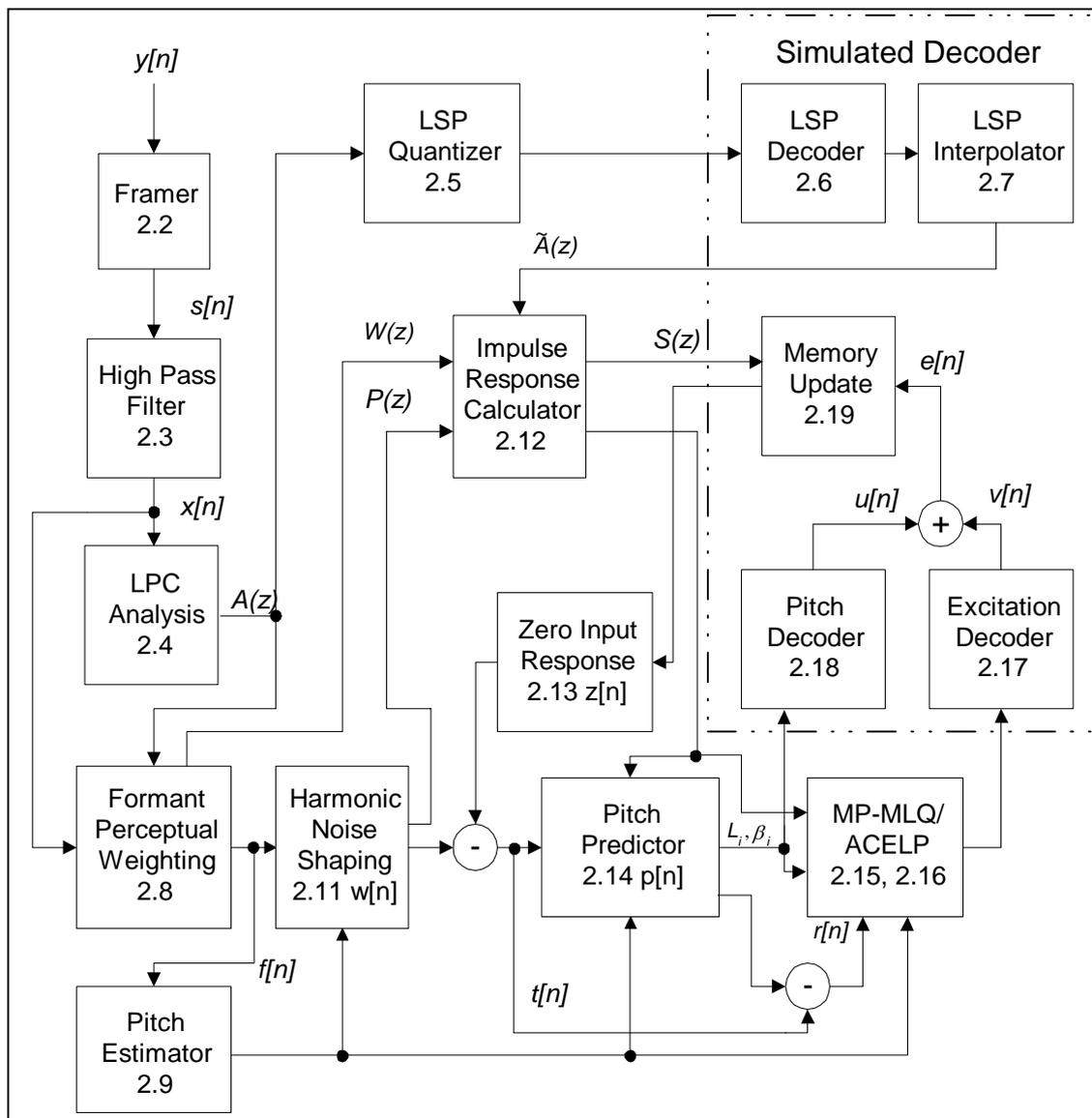


Figure 2 - 1 / JT-G723. 1
(ITU-T G. 723. 1)

Block diagram of the speech coder.
For each block the corresponding section number is indicated.

2.4 L P C分析

| | | |
|------------|----------------------|---------------------------------------|
| File:LPC.C | Procedure:Comp_Lpc() | Performs LPC coefficients calculation |
| File:LPC.C | Procedure:Durbin() | Levinson-Durbin recursion |

L P C分析は信号 $x[n]$ に対し以下のように実行される。この分析には 10 次の線形予測 (L P) 分析が用いられる。サブフレームごとに 180 サンプルの窓がサブフレームの中心におかれる。それらのサンプルに対してハミング窓が適用される。11 個の自己相関係数が窓がけされた信号から計算される。白色雑音補正係数 (1025/1024) が式 $R[0] = R[0](1 + 1 / 1024)$ として適用される。それ以外の 10 個の自己相関係数には二項窓係数表の値を乗じる。(この表を含め他の表の数値はCコードに記述されている。) 線形予測係数 (L P C) は通常のレビンソン・ダービン再帰法を用いて計算される。全ての入力フレームに対して、4組のL P C (各サブフレームごとに1組) が計算される。このL P Cは短期聴覚重み付けフィルタの構成に用いる。このL P C合成フィルタは次式で定義される。

$$A_i(z) = \frac{1}{1 - \sum_{j=1}^{10} a_{ij}z^{-j}}, 0 \leq i \leq 3 \quad (2)$$

ここで i はサブフレーム番号を示し、0 から 3 の値を持つ。

2.5 L S P量子化器

| | | |
|------------|---------------------|----------------------------------|
| File:LSP.C | Procedure:AtoLsp() | Converts LPC to LSP coefficients |
| File:LSP.C | Procedure:Lsp_Qnt() | LSP vector quantization |
| File:LSP.C | Procedure:Lsp_Svq() | LSP sub-vectors quantization |

まず、わずかな帯域幅の拡張 (7.5Hz) を実行する。次に、計算中のL Pフィルタ $A_3(z)$ のL P係数を、予測分割ベクトル量子化器を用いて以下のように量子化する。

- (1) L P係数 $\{a_j\}_{j=1..10}$ を、単位円に沿って探索し、零点との交差を補間することにより、L S P係数 $\{p'_j\}_{j=1..10}$ に変換する。
- (2) 長期DC成分 p_{DC} をL S P係数 p' から除去し、新しいDC除去済みL S Pベクトル p を得る。
- (3) 時間 n (フレーム) に、1次の固定予測器 ($b = (12 / 32)$) を、直前に復号されたL S Pベクトル \tilde{p}_{n-1} に適用し、DC除去済み予測L S Pベクトル \bar{p}_n 、残差L S P誤差ベクトル e_n を得る。

$$p_n^T = [p_{1,n} p_{2,n} \cdots p_{10,n}] \quad (3.1)$$

$$\bar{p}_n^T = [\bar{p}_{1,n} \bar{p}_{2,n} \cdots \bar{p}_{10,n}] \quad (3.2)$$

$$\bar{p}_n = b[\tilde{p}_{n-1} - p_{DC}] \quad (3.3)$$

$$e_n = p_n - \bar{p}_n \quad (3.4)$$

- (4) 量子化されていないL S Pベクトル \mathbf{p}'_n 、量子化されたL S Pベクトル $\tilde{\mathbf{p}}_n$ 、残差L S P誤差ベクトル \mathbf{e}_n は各々、次元 3,3,4 の3つのサブベクトルに分割される。 m 番目のサブベクトルはそれぞれ8ビットのコードブックを用いて量子化される。誤差評価値 $E_{l,m}$ を最小にする最適なコードブックのインデックス l が選択される。

$$\mathbf{p}'_m{}^T = [p'_{1+3m} p'_{2+3m} \cdots p'_{K_m+3m}], K_m = \begin{cases} 3, m=0 \\ 3, m=1 \\ 4, m=2 \end{cases} \quad (4.1)$$

$$\tilde{\mathbf{p}}_{l,m}{}^T = [\tilde{p}_{1,l,m} \tilde{p}_{2,l,m} \cdots \tilde{p}_{K_m,l,m}], \begin{matrix} 0 \leq m \leq 2 \\ 1 \leq l \leq 256 \end{matrix} \quad (4.2)$$

$$\mathbf{p}' = \mathbf{p} + \mathbf{p}_{DC} \quad (4.3)$$

$$\tilde{\mathbf{p}}_{l,m} = \bar{\mathbf{p}}_m + \mathbf{p}_{DC_m} + \mathbf{e}_{l,m}, \begin{matrix} 0 \leq m \leq 2 \\ 1 \leq l \leq 256 \end{matrix} \quad (4.4)$$

$$E_{l,m} = (\mathbf{p}'_m - \tilde{\mathbf{p}}_{l,m})^T \mathbf{W}_m (\mathbf{p}'_m - \tilde{\mathbf{p}}_{l,m}), \begin{matrix} 0 \leq m \leq 2 \\ 1 \leq l \leq 256 \end{matrix} \quad (4.5)$$

ここで $e_{l,m}$ は m 番目の分割残差L S Pコードブックの l 番目のエントリで、 \mathbf{W}_m は量子化されていないベクトル \mathbf{p}'_n から決定され、次式で定義される重みを持つ対角重み付け行列である。

$$w_{j,j} = \frac{1}{\min\{p'_j - p'_{j-1}, p'_{j+1} - p'_j\}}, 2 \leq j \leq 9$$

$$w_{1,1} = \frac{1}{p'_2 - p'_1} \quad (5)$$

$$w_{10,10} = \frac{1}{p'_{10} - p'_9}$$

- (5) ここで選択されたインデックスを通信チャネルへ伝送する。

2.6 L S P復号器

| | | |
|------------|---------------------|-----------------------------|
| File:LSP.C | Procedure:Lsp_Inq() | Inverse quantization of LSP |
|------------|---------------------|-----------------------------|

L S P係数の復号を以下の手順に従って行う。

- (1) 最初に3つのサブベクトル $\{\mathbf{e}_{m,n}\}_{m=0..2}$ を復号し10次のベクトル $\tilde{\mathbf{e}}_n$ を形成する。
- (2) 予測ベクトル $\bar{\mathbf{p}}_n$ に復号されたベクトル $\tilde{\mathbf{e}}_n$ およびDCベクトル \mathbf{p}_{DC} を加算して、復号されたL S Pベクトル $\tilde{\mathbf{p}}_n$ 得る。
- (3) 復号されたL S Pベクトル $\tilde{\mathbf{p}}_n$ の配置が、以下の条件に適合することを保証するために安定性チェックを行う。

$$\tilde{p}_{j+1,n} - \tilde{p}_{j,n} \geq \Delta_{\min}, 1 \leq j \leq 9 \quad (6)$$

Δ_{\min} は 31.25Hz である。もし $\tilde{\boldsymbol{p}}_j$ と $\tilde{\boldsymbol{p}}_{j+1}$ に対する安定性チェックにおいて式(6)による条件を満たさないならば $\tilde{\boldsymbol{p}}_j$ および $\tilde{\boldsymbol{p}}_{j+1}$ を以下のように補正する。

$$\tilde{p}_{avg} = (\tilde{p}_j + \tilde{p}_{j+1}) / 2 \quad (7.1)$$

$$\tilde{p}_j = \tilde{p}_{avg} - \Delta_{\min} / 2 \quad (7.2)$$

$$\tilde{p}_{j+1} = \tilde{p}_{avg} + \Delta_{\min} / 2 \quad (7.3)$$

この補正は式(6)の条件を満たすまで繰り返される。もし 10 回の繰り返しの後にも、この安定性の条件を満たさないときは直前の L S P ベクトルを使用する。

2.7 L S P 補間器

| | | |
|------------|---------------------|----------------------------------|
| File:LSP.C | Procedure:Lsp_Int() | LSP interpolator |
| File:LSP.C | Procedure:LsptoA() | Converts LSP to LPC coefficients |

サブフレームごとに、復号された L S P ベクトル $\tilde{\boldsymbol{p}}_n$ と直前の L S P ベクトル $\tilde{\boldsymbol{p}}_{n-1}$ に対して線形補間を行う。4 つの補間ベクトル $\{\tilde{\boldsymbol{p}}_i\}_{i=0..3}$ を L P C ベクトル $\{\tilde{\boldsymbol{a}}_i\}_{i=0..3}$ に変換する。

$$\tilde{\boldsymbol{p}}_{ni} = \begin{cases} 0.75\tilde{\boldsymbol{p}}_{n-1} + 0.25\tilde{\boldsymbol{p}}_n, i = 0 \\ 0.5\tilde{\boldsymbol{p}}_{n-1} + 0.5\tilde{\boldsymbol{p}}_n, i = 1 \\ 0.25\tilde{\boldsymbol{p}}_{n-1} + 0.75\tilde{\boldsymbol{p}}_n, i = 2 \\ \tilde{\boldsymbol{p}}_n, i = 3 \end{cases} \quad (8)$$

$$\tilde{\boldsymbol{a}}_i^T = [\tilde{a}_{i1}\tilde{a}_{i2}\dots\tilde{a}_{i10}]^T, 0 \leq i \leq 3 \quad (9)$$

量子化された L P C フィルタ $\tilde{A}_i(z)$ を復号された音声信号を生成するために用いる。それは次式で定義される。

$$\tilde{A}_i(z) = \frac{1}{1 - \sum_{j=1}^{10} \tilde{a}_{ij} z^{-j}}, 0 \leq i \leq 3 \quad (10)$$

2.8 ホルマント聴覚重み付けフィルタ

| | | |
|------------|------------------------|---|
| File:LPC.C | Procedure:Wght_Lpc() | Computes perceptual filter coefficients |
| File:LPC.C | Procedure>Error_Wght() | Applies perceptual weighting filter |

サブフレームごとにホルマント聴覚重み付けフィルタを構成する。このフィルタには量子化されていない L P C 係数 $\{\tilde{\boldsymbol{a}}_{ij}\}_{j=1..10}$ を用いる。このフィルタは以下の伝達関数を持つ。

$$W_i(z) = \frac{1 - \sum_{j=1}^{10} a_{ij} z^{-j} \gamma_1^j}{1 - \sum_{j=1}^{10} a_{ij} z^{-j} \gamma_2^j}, \quad 0 \leq i \leq 3 \quad (11)$$

ここで $\gamma_1 = 0.9$ であり、 $\gamma_2 = 0.5$ である。入力音声フレーム $\{x[n]\}_{n=0..239}$ は4つのサブフレームに分割され、サブフレームごとにフィルタ $W_i(z)$ をかけ、重み付き出力音声信号 $\{f[n]\}_{n=0..239}$ を得る。

2.9 ピッチ推定

| | | |
|----------------|-------------------------|----------------------------|
| File:EXC_LBC.C | Procedure:Estim_Pitch() | Open loop pitch estimation |
|----------------|-------------------------|----------------------------|

フレームごとに2つのピッチ、つまり前半2つのサブフレームについて1つと後半2つのサブフレームについて1つのピッチが推定される。開ループピッチ周期推定 L_{OL} をホルマント聴覚重み付き音声信号 $f[n]$ を用いて計算する。次式に示す相互相関 $C_{OL}(j)$ を評価値とする最大化法を用いて、ピッチ周期を決定する。

$$C_{OL}(j) = \frac{\left(\sum_{n=0}^{119} f[n] \cdot f[n-j] \right)^2}{\sum_{n=0}^{119} f[n-j] \cdot f[n-j]}, \quad 18 \leq j \leq 142 \quad (12)$$

相互相関 $C_{OL}(j)$ を最大にするインデックス j を2つのサブフレームの開ループピッチ推定として選択する。最適インデックスの探索において、倍数ピッチが選択されるのを避けるため、より小さいピッチ周期を優先させる。 $C_{OL}(j)$ の最大値は $j=18$ から探索を開始する。探索によって検出された最大値 $C_{OL}(j)$ に対して、その値を直前の最適な最大値 $C_{OL}(j')$ と比較する。もしインデックス j と j' の差が18未満であり $C_{OL}(j) > C_{OL}(j')$ ならば、新しい最大値として選択し、もし2つのインデックスの差が18以上であり、 $C_{OL}(j) > C_{OL}(j')$ で、かつその差が1.25dBよりも大きい場合のみ新しい最大値として選択する。

2.10 サブフレーム処理

これ以降、すべての計算ブロックはサブフレームごとに実行する。

2.11 高調波ノイズシェーピング

| | | |
|----------------|---------------------|---|
| File:EXC_LBC.C | Procedure:Comp_Pw() | Computes harmonic noise filter coefficients |
| File:EXC_LBC.C | Procedure:Filt_Pw() | Applies harmonic noise filter |

符号化音声の品質改善のために高調波ノイズシェーピングフィルタを構成する。このフィルタは次式で表される。

$$P_i(z) = 1 - \beta z^{-L} \quad (13)$$

このフィルタに対しての最適ラグ L は以下の評価値 $C_{PW}(j)$ を最大にするラグである。ただし、自乗する

前に評価値の分子 $N(j)$ の値が正になるもののみを考える。

$$N(j) = \sum_{n=0}^{59} f[n] \cdot f[n-j] \quad (14.1)$$

$$C_{pw}(j) = \frac{(N(j))^2}{\sum_{n=0}^{59} f[n-j] \cdot f[n-j]}, \quad L_1 \leq j \leq L_2 \quad (14.2)$$

ここで $L_1 = L_{OL} - 3$ 、 $L_2 = L_{OL} + 3$ である。この最大値は C_L として定義される。最適フィルタ利得 G_{opt} は次式で表される。

$$G_{opt} = \frac{\sum_{n=0}^{59} f[n]f[n-L]}{\sum_{n=0}^{59} f[n-L]f[n-L]} \quad (15)$$

G_{opt} は $[0,1]$ の範囲内に限定される。重み付き音声信号 $\{f[n]\}_{n=0..59}$ のエネルギー E は次式で与えられる。

$$E = \sum_{n=0}^{59} f^2[n] \quad (16)$$

このとき、高調波ノイズシェーピングフィルタ $P(z)$ の係数 β は次式で与えられる。

$$\beta = \begin{cases} 0.3125G_{opt}, & \text{if } -10\log_{10}\left(1 - \frac{C_L}{E}\right) \geq 2.0 \\ 0.0, & \text{otherwise} \end{cases} \quad (17)$$

高調波ノイズシェーピングフィルタ係数を計算した後、ホルマント聴覚重み付き音声信号 $f[n]$ に対し $P(z)$ を用いてフィルタ処理を行いターゲットベクトル $w[n]$ を得る。

$$w[n] = f[n] - \beta f[n-L], \quad 0 \leq n \leq 59 \quad (18)$$

2.12 インパルス応答の計算

| | | |
|--------------|-----------------------|------------------------------|
| File : LPC.C | Procedure : Comp_Ir() | Impulse response computation |
|--------------|-----------------------|------------------------------|

閉ループ分析のために、次に示す結合フィルタ $S_i(z)$ を用いる。

$$S_i(z) = \tilde{A}_i(z) \cdot W_i(z) \cdot P_i(z), \quad 0 \leq i \leq 3 \quad (19)$$

ここで $S_i(z)$ の各成分は、式(10)、式(11)、式(13)に定義される。このフィルタのインパルス応答が計算され、後に $\{h_i[n]\}_{n=0..59, i=0..3}$ として参照される。

2.13 零入力応答とリングング補正

| | | |
|--------------|-------------------------|------------------------------|
| File : LPC.C | Procedure : Sub_Ring () | Performs ringing subtraction |
|--------------|-------------------------|------------------------------|

結合フィルタ $S_i(z)$ の零入力応答を、入力信号のサンプル値が全て 0 としてフィルタ出力を計算することで得る。零入力応答は $\{z[n]\}_{n=0..59}$ として表される。リングング補正は高調波重み付きベクトル $\{w_i[n]\}_{n=0..59}$ から零入力応答を減算することにより求める。得られるベクトルは $t[n] = w[n] - z[n]$ として定義される。

2.14 ピッチ予測器

| | | |
|------------------|--------------------------|--|
| File : EXC_LBC.C | Procedure : Find_Acbk() | Adaptive codebook contribution. Calls Get_Rez() and Decod_Acbk() |
| File : EXC_LBC.C | Procedure : Get_Rez() | Gets residual from the excitation buffer |
| File : EXC_LBC.C | Procedure : Decod_Acbk() | Decodes the adaptive codebook contribution |

ピッチ予測寄与分は、通常の適応コードブック寄与分とみなされる。ピッチ予測器は5次のピッチ予測器である（式(41.2)を参照せよ）。サブフレーム0と2については、閉ループピッチラグは、 ± 1 の範囲で適当な開ループピッチラグの近傍から選ばれ、7ビットで符号化される。（開ループピッチラグはけっして伝送されないことに注意。）サブフレーム1と3については、閉ループピッチラグは2ビットを用いて差分として符号化されるが、直前のサブフレームのラグとの差異は、 -1 、 0 、 $+1$ または $+2$ の値を取りうる。量子化され復号されたピッチラグの値は、これ以降 L_i として参照される。ピッチ予測器の利得は、高ビットレートでは 85 または 170 エントリの2つのコードブックを、低ビットレートでは 170 エントリのコードブックを用いてベクトル量子化される。170 エントリのコードブックは両方のレートとも同じである。高レートでは、サブフレーム0、1について開ループピッチ L_0 が 58 未満またはサブフレーム2、3について L_2 が 58 未満であれば 85 エントリのコードブックをピッチ利得の量子化に用いる。そうでなければ、ピッチ利得は 170 エントリのコードブックを用いる。ピッチ予測器の寄与分 $\{p[n]\}_{n=0..59}$ はターゲットベクトル $\{t[n]\}_{n=0..59}$ から残差信号 $\{r[n]\}_{n=0..59}$ を求めるために減算される。

$$r[n] = t[n] - p[n] \quad (20)$$

2.15 高レート励振 (MP-MLQ)

| | | |
|------------------|-------------------------|---|
| File : EXC_LBC.C | Procedure : Find_Fcbk() | Fixed codebook contribution |
| File : EXC_LBC.C | Procedure : Find_Best() | Residual signal quantization |
| File : EXC_LBC.C | Procedure : Gen_Trn() | Generates a train of Dirac functions |
| File : EXC_LBC.C | Procedure : Fcbk_Pack() | Combinatorial coding of pulse positions |

残差信号 $\{r[n]\}_{n=0..59}$ は、新しいターゲットベクトルとしてMP-MLQブロックへ転送される。このブロックはこのベクトルの量子化を行う。量子化プロセスはターゲットベクトル $r[n]$ を $r'[n]$ で近似することである。

$$r'[n] = \sum_{j=0}^n h[j] \cdot v[n-j], \quad 0 \leq n \leq 59 \quad (21)$$

ここで $v[n]$ はインパルス応答 $h[n]$ を持つ結合フィルタ $S(z)$ に対する励振信号で、次式で定義される。

$$v[n] = G \sum_{k=0}^{M-1} \alpha_k \delta[n-m_k], \quad 0 \leq n \leq 59 \quad (22)$$

ここで G は利得ファクタで、 $\delta[n]$ はディラック関数である。 $\{\alpha_k\}_{k=0 \dots M-1}$ と $\{m_k\}_{k=0 \dots M-1}$ は、それぞれ極性 (± 1) とディラック関数の位置をあらわし、 M はパルス数、この数は偶数サブフレームでは6、奇数サブフレームでは5である。パルスの位置に関してはある種の制限がある。パルスの位置がすべて奇数であるか偶数であるかのいずれかである。これはグリッドビットとして表される。誤差信号 $err[n]$ の自乗平均を最小化する未知のパラメータ、 G 、 $\{\alpha_k\}_{k=0 \dots M-1}$ と $\{m_k\}_{k=0 \dots M-1}$ を推定することが課題である。

$$err[n] = r[n] - r'[n] = r[n] - G \sum_{k=0}^{M-1} \alpha_k h[n-m_k] \quad (23)$$

パラメータの推定と量子化プロセスは合成による分析 (A-b-S) 法に基づいている。 G_{max} パラメータは以下のように推定され量子化される。まず、インパルス応答 $h[n]$ と新しいターゲットベクトル $r[n]$ の相互相関関数 $d[j]$ が算出される。

$$d[j] = \sum_{n=j}^{59} r[n] \cdot h[n-j], \quad 0 \leq j \leq 59 \quad (24)$$

利得推定値は次式で与えられる。

$$G_{max} = \frac{\max\{d[j]\}_{j=0 \dots 59}}{\sum_{n=0}^{59} h[n] \cdot h[n]} \quad (25)$$

続いて、利得の推定値 G_{max} を対数量子化器で量子化する。このスカラ利得量子化器は2つのレートで共通であり 3.2dB 間隔の 24 ステップで構成されている。この量子化値 \tilde{G}_{max} の近傍から、付加利得値が $[\tilde{G}_{max} - 3.2, \tilde{G}_{max} + 6.4]$ の範囲で選ばれる。これらの利得値に関して、その極性とパルス位置が順次最適化される。この最適化処理を奇数および偶数グリッド両方に対し繰り返す。最後に $err[n]$ の最小自乗平均を求めることにより量子化されたパラメータの組合せを選択する。パルス位置と利得の最適な組み合わせが伝送される。組み合わせ符号 $\left[\frac{30}{M} \right]$ はパルス位置の伝送に用いられる。なおその上、固定コードブックの符号語数は2のべき乗になっていないため、3ビットの付加的な情報が、13ビットのインデックスを形成する4つのサブフレームに対する組み合わせ符号の4つの最上位ビット (MSB) が組み合わせられて保存される。Cコードでは、情報がパッキングされる過程が詳細に記述されている。

短かいピッチ周期を持つ音声の品質を改善するため、次の付加的な処理を用いる。上に述べた量子化のなかで、サブフレーム 0 と 1 について L_0 が 58 未満あるいはサブフレーム 2 および 3 について L_2 が 58 未満の場合は、単一のディラック関数を用いる代わりにピッチインデックス L_0 または L_2 の周期を持つディラック関数列を、それぞれの位置 m_k に対し用いる。残差信号を再生するために、ディラック関数列あるいは単一のディラック関数のどちらを選択するかは、平均自乗誤差の計算に基づいて行われる。最小の平均自乗誤差を与える構成を選択し、そのパラメータインデックスを伝送する。

2.16 低レート励振(ACELP)

| | | |
|------------------|------------------------------|--|
| File : EXC_LBC.C | Procedure : search_T0() | Pitch synchronous excitation |
| File : EXC_LBC.C | Procedure : ACELP_LBC_code() | Computes innovative vector |
| File : EXC_LBC.C | Procedure : Cor_h() | Correlations of impulse response |
| File : EXC_LBC.C | Procedure : Cor_h_X() | Correlation of target vector with impulse response |
| File : EXC_LBC.C | Procedure : D4i64_LBC() | Algebraic codebook search |
| File : EXC_LBC.C | Procedure : G_code() | Computes innovation vector gain |

17 ビットの代数的コードブックを固定コードブック励振 $v[n]$ を得るために用いる。それぞれの固定コードベクトルは多くとも 4 個の非零パルスを持っている。この 4 個のパルスは次の表に示す極性と位置と見なすことができる。

Table 2-1 / JT-G723.1 ACELP excitation codebook
(ITU-T G.723.1)

| sign | positions |
|---------|---------------------------------|
| ± 1 | 0, 8, 16, 24, 32, 40, 48, 56 |
| ± 1 | 2, 10, 18, 26, 34, 42, 50, 58 |
| ± 1 | 4, 12, 20, 28, 36, 44, 52, (60) |
| ± 1 | 6, 14, 22, 30, 38, 46, 54, (62) |

すべてのパルスの位置は（奇数の位置に置くために）全てを同時に 1 だけずらすことができるが、そのためには追加の 1 ビットが必要である。最後の 2 パルスのそれぞれの最後の位置は、サブフレーム境界の外側に落ちる。これは、パルスが存在しないことを意味する。

それぞれのパルス位置は 3 ビットで、それぞれの極性が 1 ビットで符号化される。これにより 4 個のパルスに対し合計 16 ビット要ることがわかる。さらに、追加の 1 ビットがシフト結果の符号化に用いられるため 17 ビットコードブックとなる。

コードブックは、次式に示すような重み付き音声信号 $r[n]$ と重み付き合成音声信号との間の平均自乗誤差を、最小化することによって探索される。

$$E_{\xi} = \|r - GHv_{\xi}\|^2 \quad (26)$$

ここで r は、重み付け合成フィルタの零入力応答とピッチ寄与分を減算した後の重み付き音声からなるターゲットベクトルで、 G はコードブック利得、 \mathbf{v}_ξ はインデックス ξ における代数的符号語で、 H は重み付け合成フィルタ $S_i(z)$ のインパルス応答である $h(n)$ を用いた、対角が $h(0)$ 、低位の対角が $h(1), \dots, h(L-1)$ の下三角テプリッツ畳込み行列である。

最適符号語は下記の項を最大化するものとして表すことが出来る。

$$\tau_\xi = \frac{C_\xi^2}{\varepsilon_\xi} = \frac{(\mathbf{d}^T \mathbf{v}_\xi)^2}{\mathbf{v}_\xi^T \Phi \mathbf{v}_\xi} \quad (27)$$

ここで $\mathbf{d} = \mathbf{H}^T \mathbf{r}$ はターゲットベクトル $r[n]$ とインパルス応答 $h[n]$ の相関である。さらに、 $\Phi = \mathbf{H}^T \mathbf{H}$ はインパルス応答の共分散行列である。ベクトル \mathbf{d} と行列 Φ をコードブック探索に先だって計算する。ベクトル \mathbf{d} の各要素は次式で計算される。

$$d(j) = \sum_{n=j}^{59} r[n] \cdot h[n-j], \quad 0 \leq j \leq 59 \quad (28)$$

そして、対称行列 $\Phi(i, j)$ は次式で計算される。

$$\Phi(i, j) = \sum_{n=j}^{59} h[n-i] \cdot h[n-j], \quad \begin{matrix} j \geq i \\ 0 \leq i \leq 59 \end{matrix} \quad (29)$$

注記：必要な要素だけを計算し、探索スピードを高めるための効率的記憶が配慮されている。

コードブックの代数的構造は、励振ベクトル \mathbf{v}_ξ が4個の非零パルスを持っているために、非常に高速な探索を可能にしている。探索は、それぞれのパルス位置に対応した4重のループにより実行される。ここに、それぞれのループにおいて新しいパルスの寄与分が加算される。式(27)の相関は次式で与えられる。

$$C = \alpha_0 d[m_0] + \alpha_1 d[m_1] + \alpha_2 d[m_2] + \alpha_3 d[m_3] \quad (30)$$

ここで m_k は k 番目のパルスの位置で、 α_k はその極性 (± 1) である。式(27)における偶数パルス位置のコードベクトルに対するエネルギーは次式で与えられる。

$$\begin{aligned} \varepsilon &= \Phi(m_0, m_0) \\ &+ \Phi(m_1, m_1) + 2\alpha_0 \alpha_1 \Phi(m_0, m_1) \\ &+ \Phi(m_2, m_2) + 2[\alpha_0 \alpha_2 \Phi(m_0, m_2) + \alpha_1 \alpha_2 \Phi(m_1, m_2)] \\ &+ \Phi(m_3, m_3) + 2[\alpha_0 \alpha_3 \Phi(m_0, m_3) + \alpha_1 \alpha_3 \Phi(m_1, m_3) + \alpha_2 \alpha_3 \Phi(m_2, m_3)] \end{aligned} \quad (31)$$

奇数パルス位置のコードベクトルについて、式(27)のエネルギーは、奇数位置のパルスを時間的に1サンプル早くシフトして得られる等価的偶数パルス位置のエネルギーによって近似する。探索過程を簡略化するために、関数 $d[j]$ と $\Phi(m_1, m_2)$ を変形する。この簡略化は(コードブック探索に先立ち) 次のように実行する。まず、信号 $s[j]$ が定義され、さらに信号 $d'[j]$ が構成される。

$$\begin{aligned}
s[2j] &= s[2j+1] = \text{sign}(d[2j]) && \text{if } |d[2j]| > |d[2j+1]| \\
s[2j] &= s[2j+1] = \text{sign}(d[2j+1]) && \text{otherwise}
\end{aligned} \tag{32}$$

そして、信号 d' は $d'[j] = d[j]s[j]$ で与えられる。続いて、行列 Φ は極性情報を含めることにより変形する。つまり、 $\Phi(i, j) = s[i]s[j]\Phi(i, j)$ となる。その結果、式(30)の相関は以下のように与えられる。

$$C = d'[m_0] + d'[m_1] + d'[m_2] + d'[m_3] \tag{33}$$

さらに式(31)のエネルギーは以下で与えられる。

$$\begin{aligned}
\varepsilon &= \Phi'(m_0, m_0) \\
&+ \Phi'(m_1, m_1) + 2\Phi'(m_0, m_1) \\
&+ \Phi'(m_2, m_2) + 2[\Phi'(m_0, m_2) + \Phi'(m_1, m_2)] \\
&+ \Phi'(m_3, m_3) + 2[\Phi'(m_0, m_3) + \Phi'(m_1, m_3) + \Phi'(m_2, m_3)]
\end{aligned} \tag{34}$$

さらに簡略化するため、絞り込み探索アプローチと呼ばれる手法を用いる。このアプローチにおいて、事前に計算したしきい値を用いて最後のループに入る前にテストし、このしきい値を超えたときのみこのループに入る。ループに入る最大回数が固定されるため低い比率のコードブック探索ですむ。しきい値は相関 C を基に計算する。コードブック探索の前に最初の3パルスによる絶対値最大の相関係数 max_3 と相関係数の平均値 av_3 を求める。しきい値は次式で与えられる。

$$thr_3 = av_3 + (max_3 - av_3) / 2 \tag{35}$$

第4のループには、(3つのパルスによる)相関の絶対値が thr_3 を超えている時だけ入る。これにより、探索回数変動することになるが、(4サブフレームに対し)最後のループ回数がある最大値、ここでは600回に設定されていて、この回数を超えないように制御されている。(最悪値の平均はサブフレーム当たり150回である。最初の3回のループのオーバーヘッドを除けば、これはコードブックの150×8エントリのコードブック探索と見ることができる。)

コードブックの特徴は、60未満のピッチ遅延に対し、LTPピッチ予測利得ベクトルのインデックス $PGInd_i$ に依存したピッチ寄与分が、符号に加算されることである。最適な代数的符号 $v[n]$ が決定された後、それは $v[n] \leftarrow v[n] + \beta(PGInd_i)v[n - L_i - \varepsilon(PGInd_i)]$ に従って修正される。 $\beta(PGInd_i)$ と $\varepsilon(PGInd_i)$ の値はテーブル化され、 L_i は整数ピッチ周期である。もし、 $L_i < 60$ であれば、コードブック探索に先立ち、インパルス応答は同様の手法によって変更される。

信号列 $v[n]$ を得た後の最後のステップは、利得 G の量子化である。利得は高レート励振と同様の方法によって量子化される。これは、利得量子化テーブルをステップ化し、 $|G - \tilde{G}_j|$, $0 \leq j \leq 23$ を最小化するインデックス $MGIInd_i$ を選択することによって実行される。

2.17 励振信号復号器

| | | |
|------------------|-------------------------|----------------------------------|
| File : EXC_LBC.C | Procedure : Fcbk_Unpk() | Decode fixed codebook excitation |
|------------------|-------------------------|----------------------------------|

パルスの復号は以下のように実行する。

- (1) 最初に次式を用いて最大利得 G_{max} のインデックスを求める。

$$MGInd_i = GInd_i - PGInd_i \cdot GSize \quad (36)$$

ここで $GSize = 24$ は \tilde{G} テーブルのサイズ、 $PGInd_i$ は 2.18 節で得られる。

- (2) 高レートでは、パルスの位置は組合せ $\binom{30}{M}$ を用いて復号する。ここで M は 6 または 5 のいずれかである。低レートでは、位置インデックスを直接、復号する。
- (3) グリッド位置 (偶数/奇数) はグリッドビットから求める。
- (4) パルスの極性は極性ビットから求める。
- (5) 高レートでは、条件 $L_i < 58$ が満たされたときのみパルス列ビットを復号する。
- (6) 復号されたパラメータからベクトル $v[n]$ を再構成する。
- (7) 最後にピッチ寄与分 $u[n]$ とパルス寄与分 $v[n]$ を加算し、励振ベクトル $e[n]$ を得る。

2.18 ピッチ情報の復号

| | | |
|------------------|--------------------------|--|
| File : EXC_LBC.C | Procedure : Get_Rez() | Gets residual from the excitation buffer |
| File : EXC_LBC.C | Procedure : Decod_Acbk() | Decodes the adaptive codebook contribution |

ピッチ情報の復号は以下のように実行する。

- (1) 最初に偶数サブフレームについてピッチ予測器のラグを復号する。

$$L_i = PInd_i + 18, \quad i = 0, 2 \quad (37)$$

- (2) 奇数サブフレームについてピッチ予測器のラグを以下のように復号する。

$$L_i = L_{i-1} + \Delta_i, \quad i = 1, 3 \quad (38)$$

ここで $\Delta_i \in \{-1, 0, +1, +2\}$ である。

- (3) i 番目のサブフレームのピッチ予測器の利得ベクトルを利得インデックス $GInd_i$ から求める。低レートでは、このインデックスはピッチ予測器の利得ベクトルとパルス列の利得インデックスの情報も含む。この場合、ピッチ利得インデックス $PGInd_i$ は次のように求められる。

$$PGInd_i = \lfloor GInd_i / GSize \rfloor, \quad i = 0 \dots 3 \quad (39)$$

ここで $\lfloor x \rfloor$ は x を超えない最大の整数を表す。高レートでは、条件 $L_i \geq 58$ を満たすときに式(39)と同様な方法でインデックスを求める。この場合、 $PGInd_i$ は 170 エントリの利得ベクトルコードブックに対するポインタである。その他の場合、このインデックスは 85 エントリの利得ベクトルコードブックに対するポインタであり、インパルス列ビットについての付加的な情報を含んでいる。この場合、ピッチ利得インデックスは次のように求められる。

$$PGInd_i = \lfloor GInd_i \& 0x7FF / GSize \rfloor, \quad i = 0 \dots 3 \quad (40)$$

ピッチ予測器ラグと利得ベクトルはこれらのインデックスから復号され、以下に記述したようにピッチ寄与分 $u[n]$ の抽出に用いられる。最初に信号 $e[n]$ が定義される。

$$\begin{aligned} e'[0] &= e[-L_i - 2] \\ e'[1] &= e[-L_i - 1] \\ e'[n] &= e[(n \bmod L_i) - L_i], \quad 2 \leq n \leq 63 \end{aligned} \tag{41.1}$$

ここで **mod** は剰余演算を意味する。次に

$$u[n] = \sum_{j=0}^{j=4} \beta_{ij} e'[n + j], \quad 0 \leq n \leq 59 \tag{41.2}$$

を得る。

2.19 メモリの更新

| | | |
|--------------|-------------------------|---------------|
| File : LPC.C | Procedure : Upd_Ring () | Memory update |
|--------------|-------------------------|---------------|

次のサブフレームの符号化に進む前に行う i 番目のサブフレームでの最後の処理は、合成フィルタ $\tilde{A}_i(z)$ 、ホルマント聴覚重み付けフィルタ $W_i(z)$ 、および高調波ノイズシェーピングフィルタ $P_i(z)$ のメモリ更新である。これを実行するため、結合フィルタ $S_i(z)$ の完全な応答を、再生された励振系列をこのフィルタに通すことにより計算する。励振フィルタリングが終わった時点で、結合フィルタのメモリは保存され、次の音声ベクトル符号化時の零入力応答の計算に使用される。

2.20 ビット割当て

| | | |
|-------------------|-------------------------|-------------------|
| File : UTIL_LBC.C | Procedure : Line_Pack() | Bitstream packing |
|-------------------|-------------------------|-------------------|

この節では高レートと低レートについてのビット割当てを示す。2つのレートでの主な違いはパルスの位置と振幅の符号化にある。また、低レートでは、170 エントリのコードブックが常に長期予測器の利得ベクトルの符号化に使用される。

Table 2 – 2 / JT-G723.1 Bit Allocation of the 6.3 kbit/s coding algorithm
(ITU-T G.723.1)

| Parameters coded | sub-frame 0 | sub-frame 1 | sub-frame 2 | sub-frame 3 | Total |
|-------------------------|-------------|-------------|-------------|-------------|------------|
| LPC indices | | | | | 24 |
| Adaptive code-book lags | 7 | 2 | 7 | 2 | 18 |
| All the gains combined | 12 | 12 | 12 | 12 | 48 |
| Pulse positions | 20 | 18 | 20 | 18 | 73* |
| Pulse signs | 6 | 5 | 6 | 5 | 22 |
| Grid index | 1 | 1 | 1 | 1 | 4 |
| TOTAL: | | | | | 189 |

*Notes: By using the fact that the number of codewords in the fixed codebook is not a power of 2, 3 additional bits are saved by combining the 4 MSB of each pulse position index into a single 13 bit word.

Table 2 – 3 / JT-G723.1 Bit Allocation of the 5.3 kbit/s coding algorithm
(ITU-T G.723.1)

| Parameters coded | sub-frame 0 | sub-frame 1 | sub-frame 2 | sub-frame 3 | Total |
|-------------------------|-------------|-------------|-------------|-------------|------------|
| LPC indices | | | | | 24 |
| Adaptive code-book lags | 7 | 2 | 7 | 2 | 18 |
| All the gains combined | 12 | 12 | 12 | 12 | 48 |
| Pulse positions | 12 | 12 | 12 | 12 | 48 |
| Pulse signs | 4 | 4 | 4 | 4 | 16 |
| Grid index | 1 | 1 | 1 | 1 | 4 |
| TOTAL: | | | | | 158 |

Table 2 – 4 / JT-G723.1 list of transmitted parameters
(ITU-T G.723.1)

| Name | Transmitted parameters | high rate # bits | low rate # bits |
|-------|---|------------------|-----------------|
| LPC | LSP VQ index | 24 | 24 |
| ACL0 | Adaptive CodeBook Lag | 7 | 7 |
| ACL1 | Differential Adaptive CodeBook Lag | 2 | 2 |
| ACL2 | Adaptive CodeBook Lag | 7 | 7 |
| ACL3 | Differential Adaptive CodeBook Lag | 2 | 2 |
| GAIN0 | Combination of adaptive and fixed gains | 12 | 12 |
| GAIN1 | Combination of adaptive and fixed gains | 12 | 12 |
| GAIN2 | Combination of adaptive and fixed gains | 12 | 12 |
| GAIN3 | Combination of adaptive and fixed gains | 12 | 12 |
| POS0 | Pulse positions index | 20* | 12 |
| POS1 | Pulse positions index | 18* | 12 |
| POS2 | Pulse positions index | 20* | 12 |
| POS3 | Pulse positions index | 18* | 12 |
| PSIG0 | Pulse sign index | 6 | 4 |
| PSIG1 | Pulse sign index | 5 | 4 |
| PSIG2 | Pulse sign index | 6 | 4 |
| PSIG3 | Pulse sign index | 5 | 4 |
| GRID0 | Grid index | 1 | 1 |
| GRID1 | Grid index | 1 | 1 |
| GRID2 | Grid index | 1 | 1 |
| GRID3 | Grid index | 1 | 1 |

*Notes: The 4 MSB of these codewords are combined to form a 13 bit index, MSBPOS.

2.21 符号器の初期化

| | | |
|----------------|--------------------------|----------------------|
| File : CODER.C | Procedure : Init_Coder() | Coder initialization |
|----------------|--------------------------|----------------------|

符号器の静的変数は、前フレームのLSPベクトルがLSPのDCベクトル p_{DC} に初期化される場合を除き、全て0に初期化する。

3. 復号器の原理

3.1 概要

復号器の演算もフレームごとに実行する。最初に量子化されたLPCインデックスを復号し、復号器はLPC合成フィルタを構成する。全てのサブフレームについて適応コードブック励振信号と固定コードブック励振信号の両方が復号され、合成フィルタへ入力される。適応ポストフィルタはホルマントおよび前方-後方ピッチポストフィルタから構成される。励振信号はピッチポストフィルタに入力され、その出力は合成フィルタへ入力され、次にその出力はホルマントポストフィルタへ入力される。利得調整ユニットはホルマントポストフィルタの入力レベルでのエネルギーを維持する。

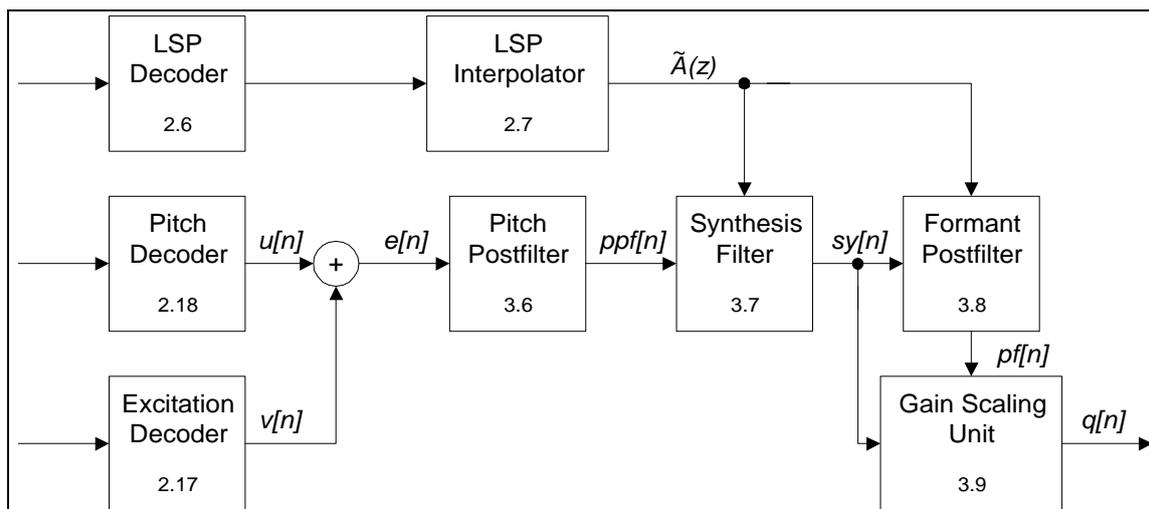


Figure 3-1 / JT-G723.1 Block diagram of the speech decoder. For each block the corresponding (ITU-T G.723.1) number is indicated.

3.2 LSP復号器

| | | |
|--------------|-----------------------|-----------------------------|
| File : LSP.C | Procedure : Lsp_Inq() | Inverse quantization of LSP |
|--------------|-----------------------|-----------------------------|

このブロックは2.6節と同様である。

3.3 LSP補間器

| | | |
|--------------|-----------------------|----------------------------------|
| File : LSP.C | Procedure : Lsp_Int() | LSP interpolator |
| File : LSP.C | Procedure : LsptoA() | Converts LSP to LPC coefficients |

このブロックは2.7節と同様である。

3.4 ピッチ情報の復号

| | | |
|------------------|--------------------------|--|
| File : EXC_LBC.C | Procedure : Get_Rez() | Gets residual from the excitation buffer |
| File : EXC_LBC.C | Procedure : Decod_Acbk() | Decodes the adaptive codebook contribution |

このブロックは2.1 8 節と同様である。

3.5 励振信号復号器

| | | |
|------------------|-------------------------|----------------------------------|
| File : EXC_LBC.C | Procedure : Fcbk_Unpk() | Decode fixed codebook excitation |
|------------------|-------------------------|----------------------------------|

このブロックは2.1 7 節と同様である。

3.6 ピッチポストフィルタ

| | | |
|------------------|------------------------|---|
| File : EXC_LBC.C | Procedure : Comp_Lpf() | Computes pitch postfilter parameters |
| File : EXC_LBC.C | Procedure : Find_F() | Forward cross correlation maximization |
| File : EXC_LBC.C | Procedure : Find_B() | Backward cross correlation maximization |
| File : EXC_LBC.C | Procedure : Get_Ind() | Gain computation |
| File : EXC_LBC.C | Procedure : Filt_Lpf() | Pitch postfiltering |

合成された信号の品質を改善するためにピッチポストフィルタを用いる。ピッチポストフィルタ処理はサブフレーム毎に行なわれるが、その実現に当たり、フレーム全体にわたる励振信号 $\{e[n]\}_{n=0..239}$ が生成され、保持されている必要がある。ピッチ周期成分のSN比を向上させることで音声品質を改善する。処理は以下の方法で行なわれる。ポストフィルタ処理された信号 $\{ppf[n]\}_{n=0..59}$ は、復号された励振信号 $\{e[n]\}_{n=0..59}$ から次式によって得られる。

$$\begin{aligned} ppf[n] &= g_p \cdot \{e[n] + g_{ip}(w_f \cdot g_f \cdot e[n + M_f] + w_b \cdot g_b \cdot e[n - M_b])\} \\ &= g_p \cdot ppf'[n] \end{aligned} \quad (42)$$

ここで $e[n]$ は復号された励振信号である。利得 g_p, g_f, g_b と遅延 M_f, M_b の計算は前方および後方の相互相関分析に基づく。 w_f, w_b は(0,0),(0,1),(1,0)の値をもつ重み係数である。相互相関を最大化する遅延が選択される。前方ピッチラグに対する相互相関は次式で与えられる。

$$C_f = \sum_{n=0}^{59} e[n]e[n + M_f], \quad M_l \leq M_f \leq M_2 \quad (43.1)$$

後方ピッチラグに対する相互相関は次式で与えられる。

$$C_b = \sum_{n=0}^{59} e[n]e[n - M_b], \quad M_1 \leq M_b \leq M_2 \quad (43.2)$$

ここで $M_1 = L_i - 3$, $M_2 = L_i + 3$ であり、 $\{L_i\}_{i=0,2}$ はサブフレーム 1 と 3 について受信したピッチラグである。 L_0 は前半 2 つのサブフレームに、 L_2 は後半 2 つのサブフレームにそれぞれ用いられる。この時、最大化された相関値のどちらかが負の場合、もしくは $n \in [0..59]$ を満たす n に対してサンプル値 $e[n + M_f]$ が利用できない場合、各々に対応する重み係数と遅延は 0 に設定される。

ここで次の 4 つの場合が考えられる。

- (0) 両方の最大値が負の場合はピッチポストフィルタの係数は算出しない。
- (1) 前方の最大値のみ正の場合はその値を選択する。
- (2) 後方の最大値のみ正の場合はその値を選択する。
- (3) 両方の最大値が正の場合は、寄与分の大きいほうを選択する。

それぞれの場合における操作手順を以下に示す。(1)、(2)、(3)の場合、最適なピッチ遅延 (M_f または M_b) に対応する信号のエネルギー (T_{en} , D_f ないし D_b) を式(44.1)、(44.2)、(44.3)により算出する。

$$D_f = \sum_{n=0}^{59} e[n + M_f]e[n + M_f] \quad (44.1)$$

$$D_b = \sum_{n=0}^{59} e[n - M_b]e[n - M_b] \quad (44.2)$$

$$T_{en} = \sum_{n=0}^{59} e[n]e[n] \quad (44.3)$$

前方エネルギーは次式で与えられる。

$$E_f = \sum_{n=0}^{59} (e[n] - g_f e[n + M_f])^2 \quad (45.1)$$

後方エネルギーは次式で与えられる。

$$E_b = \sum_{n=0}^{59} (e[n] - g_b e[n - M_b])^2 \quad (45.2)$$

$\frac{C^2}{D}$ を最大化することで E が最小化される。(3)の場合、前方か、後方かの選択は $\frac{C_f^2}{D_f}$ と $\frac{C_b^2}{D_b}$ の大きい方

を選ぶことにより決まる。ここでの予測利得は $-10 \log_{10}(1 - \frac{C^2}{DT_{en}})$ に等しい。もしこの利得が 1.25dB より小さければ、寄与は無視できる程度であると判断してピッチポストフィルタは使用しない。一方、ピッチポストフィルタを使用する場合、最適な利得は次式で得られる。

$$g = \frac{C}{D} \quad (46)$$

音声符号器のビットレートに応じた重み付け係数 γ_{lp} を最適利得に乗じる。係数の値は高レートの場合 0.1875、低レートの場合 0.25 とする。最後に、スケーリング利得 g_p は次式で計算される。

$$g_p = \sqrt{\frac{\sum_{n=0}^{59} e^2[n]}{\sum_{n=0}^{59} (ppf'[n])^2}} \quad (47)$$

式(47)の分母の値が分子より小さい場合、利得は1とする。

3.7 LP合成フィルタ

| | | |
|--------------|--------------------|----------------------------------|
| File : LPC.C | Procedure : Synt() | Synthesizes reconstructed speech |
|--------------|--------------------|----------------------------------|

ピッチポストフィルタをかけた残差 $ppf[n]$ から合成音声信号 $sy[n]$ を得るために、10次のLP合成フィルタ $\tilde{A}_i(z)$ を用いる。

$$sy[n] = ppf[n] + \sum_{j=1}^{10} \tilde{a}_{ij} sy[n-j] \quad (48)$$

3.8 ホルマントポストフィルタ

| | | |
|-------------------|-----------------------|------------------------------------|
| File : LPC.C | Procedure : Spf() | Formant postfiltering |
| File : UTIL_LBC.C | Procedure : Comp_En() | Computes synthesized signal energy |

通常のARMA型ポストフィルタを用いる。短期ポストフィルタの伝達関数は次式で与えられる。

$$k = \frac{\sum_{n=l}^{59} sy[n]sy[n-l]}{\sum_{n=0}^{59} sy[n]sy[n]} \quad (49.1)$$

$$k_1 = \frac{3}{4} k_{old} + \frac{1}{4} k \quad (49.2)$$

$$F(z) = \frac{1 - \sum_{i=1}^{10} \tilde{a}_i \lambda_1^i z^{-i}}{1 - \sum_{i=1}^{10} \tilde{a}_i \lambda_2^i z^{-i}} (1 - 0.25k_1 z^{-1}) \quad (49.3)$$

ここで $\lambda_1=0.65$ 、 $\lambda_2=0.75$ とする。 k は合成音声信号 $sy[n]$ から推定された1次の自己相関係数で、 k_{old} は直前のサブフレームにおける k_1 の値である。 $sy[n]$ を入力信号とするホルマントポストフィルタの出力としてポストフィルタ出力信号 $pf[n]$ を得る。

3.9 利得調整ユニット

| | | |
|-------------------|---------------------|--|
| File : UTIL_LBC.C | Procedure : Scale() | Gain adjustment of postfiltered signal |
|-------------------|---------------------|--|

ここでは合成音声ベクトル $\{s[n]\}_{n=0..59}$ とポストフィルタ出力ベクトル $\{f[n]\}_{n=0..59}$ の2つのベクトルを受け取る。まず次式を使って振幅比 g_s が計算される。

$$g_s = \sqrt{\frac{\sum_{n=0}^{59} sy^2[n]}{\sum_{n=0}^{59} pf^2[n]}} \quad (50)$$

分母が0の場合、 g_s は1とする。

次に以下の式をそれぞれ用い、利得 $g[n]$ を更新し、ポストフィルタ出力信号 $pf[n]$ をスケールして出力ベクトル $q[n]$ を得る。

$$g[n] = (1 - \alpha)g[n-1] + \alpha g_s \quad (51)$$

$$q[n] = pf[n] \cdot g[n] \cdot (1 + \alpha) \quad (52)$$

ここで α は 1/16 である。

3.10 フレーム補間の取り扱い

| | | |
|------------------|-------------------------|------------------------------|
| File : EXC_LBC.C | Procedure : Comp_Info() | Computes interpolation index |
| File : EXC_LBC.C | Procedure : Regen() | Current frame regeneration |

本コーデックは、所定のフレーム消失に対して頑強に設計されている。フレーム消失に対する誤り補償の方策は復号器に含まれている。しかしこの方策は、現フレームのビット列が消失したという通知を外部から受けた場合に実行される。それはコーデックがバースト誤りを想定して設計されており、ランダムビット誤りに対する誤り訂正の仕組みを備えていないためである。フレーム消失が起こると、復号器は通常の復号化動作をやめ、フレーム消失誤り補償モードに切替わる。フレーム補間はLSP係数と残差信号に対して独立に行なわれる。

3.10.1 LSPの補間

フレーム補間モードでのLSP係数の復号は以下の手順で実行される。

- (1) ベクトル \tilde{e}_n は零とする。
- (2) 復号したLSPベクトル \tilde{p}_n を構成するため、予測ベクトル \bar{p}_n をDCベクトル p_{DC} と共にベクトル \tilde{e}_n に加算する。 \bar{p}_n の生成には、符号器とは異なる固定予測器 ($b_e = (23/32)$) を用いる。以降は2.6節と同様にLSPの復号を続けて行なう。ただし、 Δ_{\min} は31.25Hzではなく62.5Hzとする。

3.10.2 残差の補間

残差の補間は、消失フレームの直前の正しいフレームの状態によって、2通りの方法で実行する。フレームは有声/無声識別器でチェックされる。

識別は相互相関関数の最大化に基づく。フレームの最後 120 サンプルに対し $L_2 \pm 3$ で相互相関を求め、相関値が最大となるインデックスが補間インデックスの候補として選ばれる。次に最適なベクトルに対する予測利得を調べ、この利得が 0.58dB より大きい場合、そのフレームは有声、そうでない場合は無声と判定する。

識別結果が無声の場合は 0 を、有声の場合は推定されたピッチの値を返す。もし無声のフレームの場合にはサブフレーム 2 と 3 に対する利得の平均を保持しておく。

現フレームが消失し、かつ、前フレームが無声であると識別されていた場合、現フレームの励振信号は一樣乱数発生器によって生成される。この乱数発生器の出力は、事前に計算した利得値でスケーリングされる。

有声の場合、現フレームは、識別器の出力値に等しい周期を持った周期的励振信号を使って再生成される。

もしフレーム消失の状態が次の 2 フレームまで続く場合は、再生成されたベクトルを 1 フレームあたり 2.5dB ずつ減衰する。3 フレーム補間した後は、出力を完全に無音とする。

3.11 復号器の初期化

| | | |
|----------------|--------------------------|------------------------|
| File : DECOD.C | Procedure : Init_Decod() | Decoder initialization |
|----------------|--------------------------|------------------------|

復号器の静的変数は以下の例外を除き、すべて 0 に初期化する。

- (1) 前フレームの L S P ベクトルは L S P の DC ベクトル P_{DC} に初期化する。
- (2) ポストフィルタの利得 $g[-1]$ は 1 に初期化する。

4. ビット列のパッキング

凡例：伝送されるパラメータの各ビットは PAR(x)_By と表す。

ここで PAR はパラメータの名称、x はそのパラメータがサブフレームに関連した場合にそのサブフレーム番号を示し、y は 0(LSB)から始まり MSB までのビット位置を表す。PARx_By.. PARx_Bz は伝送されるビット位置が y から z までの範囲であることを表す。使用しないビットは UB と名付ける(値は 0 とする)。RATEFLAG_B0 は現フレームが高レート(0)、低レート(1)のどちらが使われているかを示す。同様に VADFLAG_B0 は現フレームが有声であるか(0)、無声であるか(1)を示す。この RATEFLAG と VADFLAG の両方もが 1 に設定される組み合わせは、将来の用途に予約されている。各オクテットは Table 4-1 / JT-G723.1、Table 4-2 / JT-G723.1 に並んだ順に伝送される。それぞれのオクテット中のビット順序は、MSB を左側に、LSB を右側にして並べている。

Table 4-1 / JT-G723.1 Word bit packing for the high bit rate codec.
(ITU-T G723.1)

| TRANSMITTED OCTETS | PARx_By,.... |
|--------------------|--|
| 1 | LPC_B5...LPC_B0, VADFLAG_B0, RATEFLAG_B0 |
| 2 | LPC_B13...LPC_B6 |
| 3 | LPC_B21...LPC_B14 |
| 4 | ACL0_B5...ACL0_B0, LPC_B23...LPC_B22 |
| 5 | ACL2_B4...ACL2_B0, ACL1_B1, ACL1_B0, ACL0_B6 |
| 6 | GAIN0_B3...GAIN0_B0, ACL3_B1, ACL3_B0, ACL2_B6, ACL2_B5 |
| 7 | GAIN0_B11...GAIN0_B4 |
| 8 | GAIN1_B7...GAIN1_B0 |
| 9 | GAIN2_B3...GAIN2_B0, GAIN1_B11...GAIN1_B8 |
| 10 | GAIN2_B11...GAIN2_B4 |
| 11 | GAIN3_B7...GAIN3_B0 |
| 12 | GRID3_B0, GRID2_B0, GRID1_B0, GRID0_B0, GAIN3_B11...GAIN3_B8 |
| 13 | MSBPOS_B6...MSBPOS_B0, UB |
| 14 | POS0_B1, POS0_B0, MSBPOS_B12...MSBPOS_B7 |
| 15 | POS0_B9...POS0_B2 |
| 16 | POS1_B1, POS1_B0, POS0_B15...POS0_B10 |
| 17 | POS1_B9...POS1_B2 |
| 18 | POS2_B3...POS2_B0, POS1_B13...POS1_B10 |
| 19 | POS2_B11...POS2_B4 |
| 20 | POS3_B3...POS3_B0, POS2_B15...POS2_B12 |
| 21 | POS3_B11...POS3_B4 |
| 22 | PSIG0_B5...PSIG0_B0, POS3_B13, POS3_B12 |
| 23 | PSIG2_B2...PSIG2_B0, PSIG1_B4...PSIG1_B0 |
| 24 | PSIG3_B4...PSIG3_B0, PSIG2_B5...PSIG2_B3 |

Table 4 – 2 / JT-G723.1 Word bit packing for the low bit rate codec.
(ITU-T G723.1)

| TRANSMITTED OCTETS | PAR _x _B _y ,.... |
|--------------------|--|
| 1 | LPC_B5...LPC_B0, VADFLAG_B0, RATEFLAG_B0 |
| 2 | LPC_B13...LPC_B6 |
| 3 | LPC_B21...LPC_B14 |
| 4 | ACL0_B5...ACL0_B0, LPC_B23...LPC_B22 |
| 5 | ACL2_B4...ACL2_B0, ACL1_B1, ACL1_B0, ACL0_B6 |
| 6 | GAIN0_B3...GAIN0_B0, ACL3_B1, ACL3_B0, ACL2_B6, ACL2_B5 |
| 7 | GAIN0_B11...GAIN0_B4 |
| 8 | GAIN1_B7...GAIN1_B0 |
| 9 | GAIN2_B3...GAIN2_B0, GAIN1_B11...GAIN1_B8 |
| 10 | GAIN2_B11...GAIN2_B4 |
| 11 | GAIN3_B7...GAIN3_B0 |
| 12 | GRID3_B0, GRID2_B0, GRID1_B0, GRID0_B0, GAIN3_B11...GAIN3_B8 |
| 13 | POS0_B7...POS0_B0 |
| 14 | POS1_B3...POS1_B0, POS0_B11...POS0_B8 |
| 15 | POS1_B11...POS1_B4 |
| 16 | POS2_B7...POS2_B0 |
| 17 | POS3_B3...POS3_B0, POS2_B11...POS2_B8 |
| 18 | POS3_B11...POS3_B4 |
| 19 | PSIG1_B3...PSIG1_B0, PSIG0_B3...PSIG0_B0 |
| 20 | PSIG3_B3...PSIG3_B0, PSIG2_B3...PSIG2_B0 |

5. ANS I Cコード

本デュアルレート符号器・復号器 16 ビット固定小数点演算でシミュレーションできるANS I Cコードは、日本ITU協会から入手できる。このCコードの含まれる全てのファイルを以下の表に示す。

Table 5-1 / JT-G723.1 List of software file name.
(ITU-T G723.1)

| File name | Description |
|------------|---|
| TYPEDEF.H | data type definition is machine dependent |
| CST_LBC.H | definition of constants for G.723.1 |
| LBCCODEC.C | main program for G.723.1 speech codecs |
| LBCCODEC.H | functions prototypes |
| CODER.C | G.723.1 speech encoder for the two bit rates |
| CODER.H | functions prototypes |
| DECOD.C | G.723.1 speech decoder for the two bit rates |
| DECOD.H | functions prototypes |
| LPC.C | linear predictive analysis |
| LPC.H | functions prototypes |
| LSP.C | line spectral pair related functions, quantizer |
| LSP.H | functions prototypes |
| EXC_LBC.C | adaptive and fixed (MP-MLQ, ACELP) excitation |
| EXC_LBC.H | functions prototypes |
| UTIL_LBC.C | miscellaneous functions (HPF, pack, unpack, I/O...) |
| UTIL_LBC.H | functions prototypes |
| TAB_LBC.C | tables of constants |
| TAB_LBC.H | external declaration for constant tables |
| BASOP.C | fixed point arithmetic and logical operation |
| BASOP.H | functions prototypes |

6. 記号解説

本文中で用いられている数式記号と、それらの簡単な解説を以下の表に示す。

Table 6 – 1 / JT-G723.1(1/2) Glossary of symbols in the text.
(ITU-T G723.1)

| Symbol | Description |
|----------------------|---|
| $y[j]$ | input speech samples |
| $s[j]$ | input speech frame of 240 samples |
| $x[j]$ | highpass filtered speech frame |
| $R[j]$ | autocorrelation function, $n=0,1,\dots,10$ |
| a_i | LPC coefficient vector of sub-frame i |
| \tilde{a}_i | quantized LPC coefficient vector of sub-frame i |
| p' | unquantized LSP vector |
| p | DC-removed LSP vector |
| ${}^{p_{DC}} p_{DC}$ | long-term DC vector of LSP values |
| \tilde{p}_n | decoded LSP vector for frame n |
| \bar{p}_n | DC-removed predicted LSP vector |
| e_n | residual LSP error vector for frame n |
| \tilde{e}_n | quantized value of e_n |
| W_n | diagonal weighting matrix for LSP quantization |
| γ_1, γ_2 | weights for perceptual weighting filter, 0.9, 0.5 |
| W_i | formant perceptual weighting filter for sub-frame i |
| $f[n]$ | formant perceptually weighted speech |
| L_{OL} | open loop pitch period estimate |
| $C_{OL}(j)$ | open loop pitch estimate crosscorrelation criterion function |
| $C_{PW}(j)$ | harmonic noise shaping pitch estimate crosscorrelation criterion function |
| β | harmonic noise shaping filter gain |
| L | optimal lag for harmonic noise shaping filter |
| G_{opt} | optimal gain for harmonic noise shaping filter |
| E | energy of weighted speech signal |
| P_i | harmonic noise shaping filter for sub-frame i |
| S_i | combined harmonic and formant weighting and synthesis filters for sub-frame i |
| $h[n]$ | impulse response of combined filter |
| $z[n]$ | zero input response of combined filter |
| $w[n]$ | harmonic noise weighted speech |
| $t[n]$ | target vector |
| $p[n]$ | pitch predictor contribution vector |
| $r[n]$ | residual signal vector |
| $r'[n]$ | filtered excitation vector |
| $v[n]$ | fixed codebook excitation vector |
| M | number of pulses |
| α_k | sign of pulse k |

Table 6 – 1 / JT-G723.1(2/2) Glossary of symbols in the text.
(ITU-T G723.1)

| Symbol | Description |
|------------------------|---|
| m_k | position of pulse k |
| $d[n]$ | crosscorrelation function of $h[n]$ and $r[n]$ |
| G_{\max} | estimated gain of pulses for high rate |
| L_i | pitch lag for sub-frame i |
| H | lower triangular Toeplitz convolution matrix with diagonals $h[n]$ |
| Φ | covariance matrix formed by $H^T H$ |
| max_3 | maximum correlation of 1st 3 pulses for low rate |
| av_3 | average correlation of 1st 3 pulses for low rate |
| thr_3 | threshold for correlation of 1st pulses for low rate |
| $MGIInd_i$ | maximum gain index of sub-frame i |
| $GInd_i$ | gain index of sub-frame i |
| $PInd_i$ | pitch index of sub-frame i |
| $PGInd_i$ | pitch lag index of sub-frame i |
| $GSide$ | size of excitation gain codebook, 24 |
| \tilde{G} | quantized gain |
| β_{ij} | pitch predictor gain vector |
| $u[n]$ | adaptive codebook excitation vector |
| $e[n]$ | decoded combined excitation vector |
| $ppf[n]$ | pitch postfiltered excitation signal |
| M_f, M_b | optimal forward and backward pitch postfilter lags |
| $ppf'[n]$ | unnormalized pitch postfiltered excitation signal |
| γ_{tp} | gain weighting factor, 0.1875 or 0.25 |
| g_p | pitch postfilter scaling gain |
| g_f, g_b | optimal forward and backward gains for pitch postfilter |
| C_f, C_b | forward and backward excitation crosscorrelations |
| D_f, D_b | forward and backward excitation energies |
| E_f, E_b | forward and backward energies for pitch postfilter |
| T_{en} | energy of excitation signal |
| $sy[n]$ | LSP synthesized speech signal |
| $pf[n]$ | formant postfiltered signal |
| $q[n]$ | output speech signal |
| $g[n]$ | formant postfiltered gain signal |
| k_1 | interpolated reflection coefficient for tilt compensation filter |
| λ_1, λ_2 | weights for formant postfilter, 0.65, 0.75 |
| g_s | amplitude ratio of sy and pf vector |
| α | 0.0625 in gain scaling unit |
| b_e | fixed predictor for LSP interpolation during frame erasure concealment, 23/32 |

付属資料 A

(J T - G 7 2 3 . 1 に対する)

マルチメディア通信伝送のための 5.3 および 6.3kbit/s 音声符号化方式に適用される無音圧縮方式
(本付属資料は仕様の一部である。)

A 1. 本付属資料の規定範囲

本付属資料は T T C 標準 J T - G 7 2 3 . 1 の音声コーデックのために設計された無音圧縮方式について記述する。無音圧縮技術は音声の合間の無音区間における伝送ビットレートを減らすために使用される。不連続伝送を可能にするシステムは音声検出 (VAD) アルゴリズムおよび無音区間に合成雑音を挿入する擬似背景雑音発生 (CNG) アルゴリズムが基礎となっている。CNG は伝送が停止したときに生じる雑音に対する変調を避けるために必要である。すなわち有音区間に存在していた音響的な背景雑音が突然消失すると、不快な雑音に対する変調が会話の明瞭度を低下させることさえありうる。

VAD の目的は音声の有無を確実に検出し、検出情報を CNG アルゴリズムに伝達することである。通常 VAD アルゴリズムは、検出をより確実にするために、また断続的な検出をさけるために情報の連続した数フレームに基づいて検出を行う。VAD はそのあとに引き続き行われる音声コーデックによる符号化および擬似背景雑音発生器による擬似背景雑音の挿入と同じ 30ms の音声フレームに基づいて処理しなければならない。VAD アルゴリズムの出力は CNG アルゴリズムへと引き渡される。

音声検出において最も困難なことは様々な範囲の背景雑音環境が存在することである。VAD はきわめて小さな信号対雑音比の状態においても音声を検出できなければならない。音声の一部が雑音に埋もれた場合、単純なレベル検出手法を使用して音声と雑音を識別することは不可能である。音声と雑音との識別は入力信号のスペクトル特性を考察することによってのみ可能となる。これをおこなうため VAD は逆フィルタを有し、フィルタ係数は雑音だけの区間において CNG から得る。VAD のさらなる詳細は 2 章で述べる。

CNG アルゴリズムの目的は全体的な伝送コストを可能な限り小さくおさえながら、実際の背景雑音に似た雑音を作り出すことである。送信端において CNG アルゴリズムは各フレームで VAD から与えられる有音情報を使用して受信端で合成雑音を合成するために必要な符号化パラメータを計算する。この符号化パラメータは無音挿入記述子 (S I D) フレームを構成するが、このフレームは有音フレームよりも少ないビットである必要があり無音区間に伝送される。

CNG アルゴリズムの主な特徴は S I D フレームの伝送が周期的でないということである。すなわち各無音フレームにおいてアルゴリズムは現在の無音フレームと直前の S I D フレームとを比較することにより、S I D フレームを送出するか否かを判断する。このようなことから S I D フレームの伝送は雑音のパワースペクトルが変化したときに限られる。

無音フレームの区間、擬似背景雑音は復号器において短期合成フィルタに擬似白色励振を入力することにより合成される。擬似背景雑音の特徴付けをするためのパラメータは L P C 合成フィルタの係数と励振信号のエネルギーである。符号器において各 S I D フレームでアルゴリズムは一組の L P C パラメータを計算し、符号器 L S P 量子化器を使用して 24 ビットで量子化する。また励振エネルギーを評価しそれを 6 ビットで量子化する。これにより符号化された S I D フレームはビットレートをあわせるためと D T X 情報のための 2 ビットを含め 4 バイトとなる。

CNG アルゴリズムの注目すべき特徴は各 S I D フレームで周囲雑音のスペクトルを評価する方法である。この方法では入力信号が瞬時的に定常か非定常かを考慮に入れている。

最終的に、励振は T T C 標準 J T - G 7 2 3 . 1 コーデックの高レート励振と一致する。固定励振はスペクトル成分が乏しいので、より白色雑音形の励振を得るために長期励振も使用される。アルゴリズムはランダムに長期パラメータ (遅延および利得) および固定コードブックパラメータ (グリッド、パルス位置および符号) を選択する。2 サブフレーム毎にアルゴリズムは伝送された S I D のエネルギーをもとに全体的エネ

ルギを知るために固定励振の利得を計算する。

励振の計算は符号器と復号器の同期を維持するため双方で行う。

受信器において処理を簡略化するため、擬似背景雑音発生に発生した雑音は音声信号でないことからこの区間は高調波ポストフィルタを停止させる。

TTC標準JT-G723.1 付属資料Aに記述されたVAD/DTX/CNG方式に対するテスト結果はTTC標準JT-G723.1 付属資料Aに対する付録として今後発行される予定である。

A 2. VADの記述

本節ではTTC標準JT-G723.1 音声コーデックに使用される有音検出器(VAD)について記述する。VADの機能は音声符号器によって作られた各30msフレームが音声を含むか否かを検出することである。フレーム t におけるVADの判定結果は3章およびFigureA-3-1/JT-G723.1に記述するように $Vadt$ と名付けられ、 $Ftypt$ を計算するCOD-CNGブロックの入力となる。VADアルゴリズムの性能は聞き取れる音声欠落の総和と検出した有音の比率によって特徴づけられる。

VADは基本的にはエネルギー検出器である。逆フィルタを通った信号のエネルギーはしきい値と比較される。しきい値を越えれば必ず音声を検出される。しきい値は2段階の手順で計算される。第1段階では雑音レベルが直前の値とフィルタを通した信号のエネルギーに基づいて更新される。第2段階では対数近似によって雑音レベルからしきい値が計算される。

ハングオーバは音声バーストに続く最初の僅かな無音フレームを音声であるとみなすことである。ハングオーバは低レベル音声の欠落を除去するために使用される。ハングオーバはスパイク雑音に対して作用するのを避けるため、ある持続時間を超える音声バーストに対してのみ付加される。

A2.1 適応イネーブルフラグの計算

適応イネーブルフラグは現フレーム t に対して Aen_t で表され、VADの雑音レベルを確実に音声のない区間のみ適応させるために使用される。これは背景雑音あるいは無音は有音信号でも正弦波でもないという事実に基づいている。

—有音/無音検出

直前のフレームおよび現フレームの開ループピッチ遅延は有音であるかどうかの検査に使用される。

$L_{OL}^j, j = 0, 1, 2, 3$ の4つの値に着目する。最小遅延 $L_{OL}^{\min} = \text{Min}(L_{OL}^j, j = 0, 1, 2, 3)$ が最初に計算される。カウンタ $pc \in [1, 2, 3, 4]$ は $L_{OL}^{\min} (\pm 3)$ の倍数の近傍にどれだけの遅延 L_{OL}^j があるかを示す。 pc が4に等しければ信号は有音であるとみなされる。

—正弦波検出

以下に示す正弦波検出器はTTC標準JT-G723.1 符号器のLPC分析に含まれる。 $k_i'[2]$ をフレーム t における $i = 0, \dots, 3$ の各サブフレームでダービン再帰法によって計算された第2反射係数とする。もし最後の15個の値のうち少なくとも14個について $k_i'[2] \geq 0.95$ であるならば正弦波が検出された($SinD = 1$)。その他の場合は $SinD = 0$ とする。

—適応イネーブルフラグの計算

$$\begin{cases} Aen_t = Aen_{t-1} + 2 & \text{if } pc = 4 \text{ or } SinD = 1 \\ Aen_t = Aen_{t-1} - 1 & \text{Otherwise} \end{cases}$$

Aen_t は0から6の範囲に制限される。

A.2.2 逆フィルタ

入力信号フレーム $\{s[n]\}_{n=60..239}$ は係数 $\{a_{no}[j]\}_{j=1..10}$ を持つ F I R フィルタ $A_{no}(z)$ により逆フィルタを施される。このフィルタは C N G ブロックで計算され、現在の背景雑音に関連する L P C フィルタの算出に用いられる。

$$e'_i[n] = s[n] + \sum_{j=1}^{10} a_{no}[j] \cdot s[n-j], n = 60 \rightarrow 239 \quad (A1)$$

ここで $e'_i[n]$ は逆フィルタを施された信号である。

A.2.3 フィルタ後エネルギーの計算

エネルギー Enr_t は現フレームの逆フィルタを施された信号から計算される。

$$Enr_t = \frac{1}{180} \sum_{n=60}^{239} e'^2_t[n] \quad (A2)$$

A.2.4 雑音レベルの計算

フレーム t における雑音レベル $Nlev_t$ は、直前の値、直前のエネルギー Enr_{t-1} 、そして適応イネーブルフラグ Aen_t をもとに更新される。この更新手順は、ゆっくりした立ち上がりおよび速い立ち下がりにより特徴づけられる。フレーム t における雑音レベルのダイナミックレンジは $[Nlev_{min}, Nlev_{max}]$ に制限される。

- (1) もし、 $Nlev_{t-1} > Enr_{t-1}$ ならば雑音レベルは最初にクリッピングされる

$$Nlev_t = \begin{cases} 0.25 \cdot Nlev_{t-1} + 0.75 \cdot Enr_{t-1} & \text{if } Nlev_{t-1} > Enr_{t-1} \\ Nlev_{t-1} & \text{otherwise} \end{cases} \quad (A3)$$

- (2) もし、適応がイネーブルならば $Nlev_t$ は増加させられる。それ以外の場合少量だけ減少させられる。

$$Nlev_t = \begin{cases} 1.03125 \times Nlev_t & \text{if } Aen_t = 0 \\ 0.9995 \times Nlev_t & \text{otherwise} \end{cases} \quad (A4)$$

$$\text{with } \begin{cases} Nlev_{min} = 128 \\ Nlev_{max} = 131071 \end{cases}$$

A.2.5 しきい値の計算

フレーム t におけるノイズレベル $Nlev_t$ としきい値 Thr との関係は次式の対数近似により定められる。

$$Thr = \begin{cases} 5.012 & \text{if } Nlev = 128, \\ 10^{0.7 - 0.05 \log_2 \frac{Nlev}{128}} & \text{if } 128 < Nlev < 16384 \\ 2.239 & \text{if } Nlev \geq 16384 \end{cases} \quad (A5)$$

A2.6 VADの決定

VADの決定はしきい値 Thr と現在のエネルギー Enr_t の比較を基に行われる。

$$Vad_t = \begin{cases} 1, & Enr_t \geq Thr \\ 0, & Enr_t < Thr \end{cases} \quad (A6)$$

A2.7 VADハングオーバーの付加

6フレームのハングオーバーが、音声が入力が2フレーム以上バースト ($VAD_t = 1$)したときに限り加えられる。

A2.8 VADの初期化

以下の変数を除いて、VADアルゴリズムの全ての静的変数は0に初期化される。

$$\begin{aligned} Nlev_{-1} &= 1024 \\ Enr_{-1} &= 1024 \\ L_{OL}^j &= 1 \quad j = 0,1 \\ L_{OL}^j &= 60 \quad j = 2,3 \end{aligned} \quad (A7)$$

A3. CNGの概要

このアルゴリズムは符号器と復号器に対して2ブロックに分けられ、それぞれ COD-CNGおよび DEC-CNGと呼ばれる。符号器(Figure A-3-1/JT-G723.1 参照)に対してCOD-CNGブロックは60サンプルのサブフレーム毎に音声信号の自己相関関数、過去の励振サンプル、直前フレームのLSPを使用する。

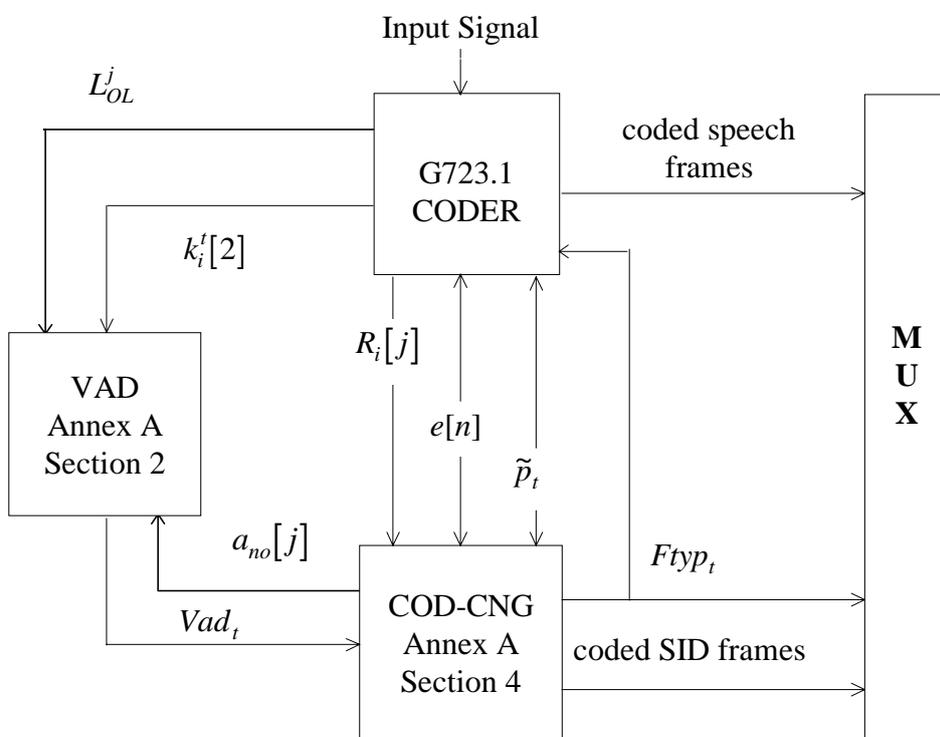


Figure A-3-1/JT-G723.1 Block diagram of the encoder with VAD/CN

(ITU-T G.723.1)

無音フレームに対してCOD-CNGは符号器の局部復号器と相手方の復号器を同期させるためにCNG励振を計算する。TTC標準JT-G723.1構造におけるLSPの予測符号化のために、無音フレームの間のLSP係数に対して、同じ入力/出力の更新が行われる。COD-CNGは符号化されたSIDフレームと、最終的に決定された F_{typ_t} （フレーム t のフレームタイプ）を非伝送フレーム、有音フレーム、SIDフレームに対応する0、1、2のいずれかの値として出力する。

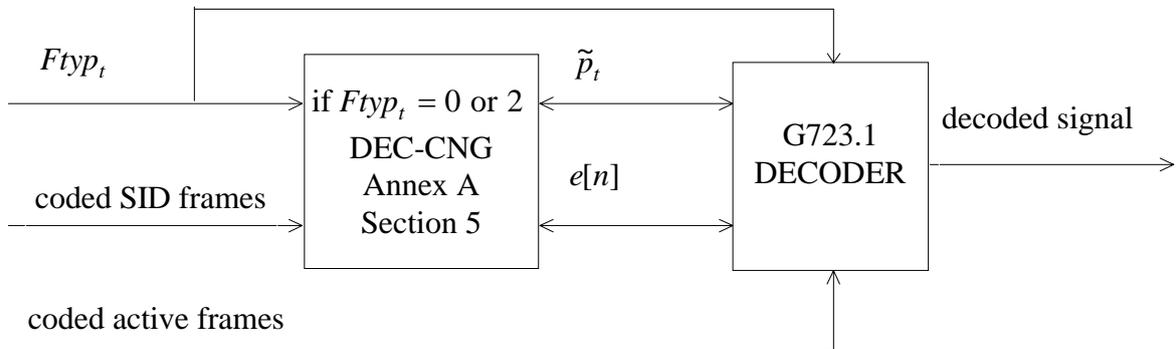


Figure A-3-2 / JT-G723.1 Block diagram of the decoder with VAD/DTX
(ITU-T G.723.1)

受信側において(Figure A-3-1 / JT-G723.1 参照)、DEC-CNGブロックは入力情報 F_{typ_t} が0または2(非伝送/SID)に等しい無音フレームに対してのみ処理を行なう。DEC-CNGはSIDおよび非伝送フレームのためにSIDフレームを、現在のLSPおよび励振をCOD-CNGと同じ計算方法を用いて復号する。

その後TTC標準JT-G723.1復号器は擬似背景雑音をCNG励振とLSPを用いて合成する。

A 4. CNG符号器の詳細

240サンプルのフレーム(有音または無音)毎に、COD-CNGブロックはVADおよび符号器からのデータを処理し、 F_{typ_t} 情報および復号されたSIDフレームをFigure A-4-1 / JT-G723.1およびA 4. 1節からA 4. 7節に記述された内容に従って生成する。

A4.1 フレームの自己相関関数の計算

| | | |
|------------------|-------------------------|---------------------------------|
| File : COD_CNG.C | procedure: Update_Acf() | Update autocorrelation function |
|------------------|-------------------------|---------------------------------|

フレーム t (有音または無音)毎に自己相関係数 $R_i[j]$ 、 $j = 0$ から10(TTC標準JT-G723.1の2.4節に記述されている復号器の中で計算される)のインデックス $i = 0$ から3までの4サブフレームにわたっての和が取られる。現フレーム t における計算された自己相関関数は以下で与えられる。

$$R^t[j] = \sum_{i=0}^3 R_i[j], \text{ for } j = 0 \text{ to } 10 \quad (\text{A8})$$

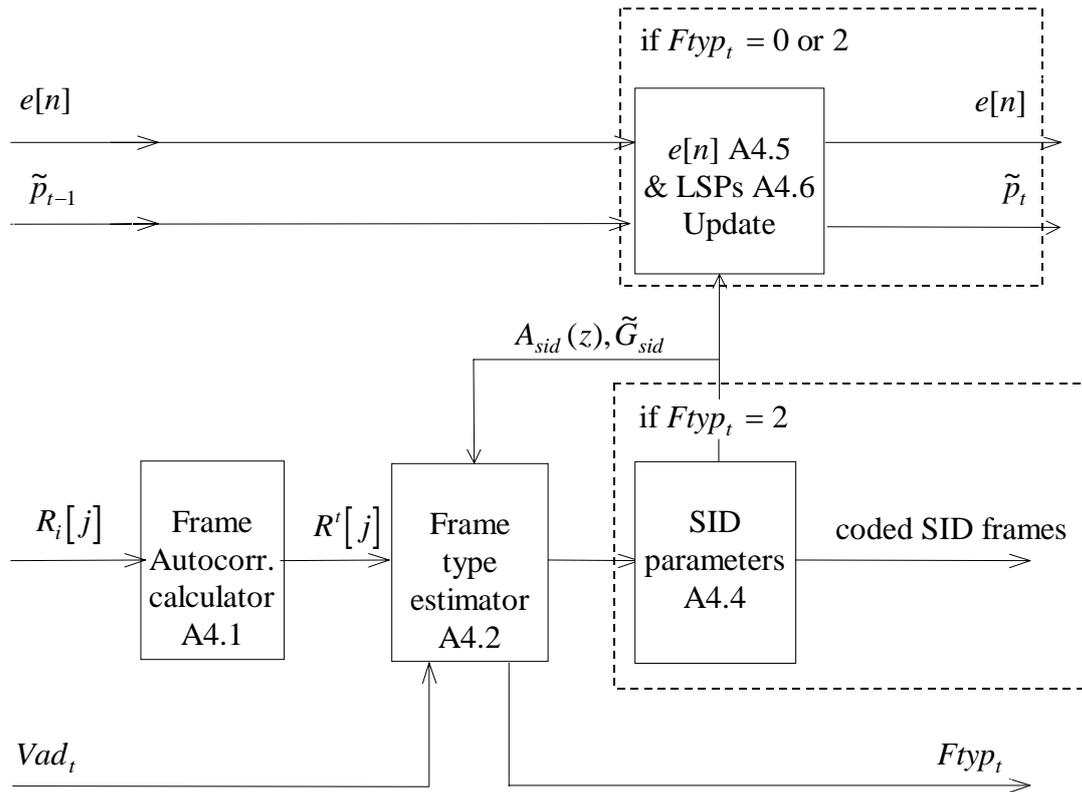


Figure A-4-1/JT-G723.1 Block diagram of the CNG at the encoder part
(ITU-T G.723.1)

A4.2 現フレームタイプ $Ftyp_t$ の計算

| | | |
|------------------|----------------------|-----------------------------|
| File : COD_CNG.C | procedure: Cod_CNG() | COD_CNG main body |
| File : COD_CNG.C | procedure: LPCDiff() | Itakura distance comparison |
| File : LPC.C | procedure : Durbin() | Levinson-Durbin recursion |

もし現フレーム t が有音フレーム ($VAD_t=1$)ならば $Ftyp_t = 1$ となりその他の処理は行なわれない。その他の場合、SID/非伝送フレームの決定は以下の手順に従って行なわれる。

現フレーム t のLPCフィルタ $A_t(z)$ は $R'[j]$ を入力としたダービン手順 (TTC標準JT-G723.1の2.4節参照)により計算される。

$A_t(z)$ の係数は $a_t[j]$, $j=1, \dots, 10$ と記される。ダービン手順ではフレーム励振エネルギーを算出する際に使用される残差エネルギー E_t も算出される。

次に現フレームタイプ $Ftyp_t$ は以下の方法に従って決定される。

- (1) もし現フレームが無音区間の最初の無音フレームならば、そのフレームはSIDフレームとして選択される。エネルギーを反映する変数 \bar{E} は E_t に、和に含まれるフレーム数 k_E は1に初期化される。

$$(Vad_{t-1} = 1) \Rightarrow \begin{cases} Ftyp_t = 2 \\ \bar{E} = E_t \\ k_E = 1 \end{cases} \quad (A9)$$

- (2) 無音区間の最初のフレームではなく、かつ現在のフィルタが直前の S I Dフィルタと有意に違う場合、もしくは現在の励振エネルギーが直前の S I Dエネルギーと有意に違う場合は、そのフレームは S I D ($Ftyp_t = 2$)として選択される。
- (3) それ以外の場合、すなわち現フレームが無音区間の最初のフレームでなく、かつ現在の L P Cフィルタおよび励振エネルギーが S I Dのものと類似している場合、そのフレームは非伝送($Ftyp_t = 0$)となる。

L P Cフィルタおよびエネルギーは以下の方法に従って比較される。

L P Cフィルタの比較

現在の L P Cフィルタおよび S I Dフィルタは二つのフィルタの間の板倉距離が与えられたしきい値より大きい場合に有意に違うとみなされる、これは以下のように表現される。

$$\sum_{j=0}^{10} R_a[j] \times R'[j] > E_t \times thr1 \quad (A10)$$

ここで $R_a[j]$, $j = 0, \dots, 10$ は S I Dフィルタの係数の自己相関係数から計算される関数であり、以下で与えられる。

$$\begin{cases} R_a[j] = 2 \sum_{k=0}^{10-j} a_{sid}[k] \times a_{sid}[k+j] & \text{if } j \neq 0 \\ R_a[0] = \sum_{k=0}^{10} a_{sid}[k]^2 \end{cases} \quad (A11)$$

(ここで $a_{sid}[0] = 1$)

$thr1$ の値として 1.2136 が用いられる。

エネルギーの比較

k_E は最初に最大値 3 まで増加させられ、フレームエネルギーの和 $\bar{E} = \sum_{i=t-k_E+1}^t E_i$ が計算される。

その後 \bar{E} は A 4. 3 節に記述されている 6 ビットの擬似対数量子化器によって量子化される。符号化利得インデックス $GInd_t$ は直前に符号化された S I D利得インデックス $GInd_{sid}$ と比較される。もしその差がしきい値 $thr2 = 3$ より大きいならば、二つのエネルギーは有意に違うとみなされる。

A4.3 平均エネルギーの量子化

| | | |
|-------------------|--------------------------|-------------------|
| File : UTIL_CNG.C | Procedure: Qua_SidGain() | Quantize Sid Gain |
| File : UTIL_CNG.C | Procedure: Dec_SidGain() | Decode Sid Gain |

量子化手順は、エネルギーの和 \bar{E} に対して実行する。復号化は平均エネルギー平方根の復号値に対応する利得を与える。サブフレームの自己相関関数 $R_i[j]$ に含まれる窓掛けと帯域幅拡張の影響を考慮するためにスケーリング係数 $\alpha_w = 2.70375$ を導入する。利得量子化器の入力に使用する値を次式に示す。

$$G = \alpha_w \times \sqrt{\frac{1}{k_E \times 240} \bar{E}}, \text{ bounded in } [0, 352].$$

量子化器は、擬似対数量子化器であり、[0, 352]の領域をそれぞれインデックス $isg = 0, 1, 2$ 、長さ $N[isg] = 16, 16, 32$ 、分解能 2, 4, 8 を持つ3つのセグメントに分割されている。
セグメント isg に対する復号化値を $G_{isg}[j]$ 、 $j = 0, \dots, N[isg]-1$ とすると、これらの値は次式で与えられる。

$$G_{isg}[j] = G_{isg}[0] + j \times 2^{(isg+1)} \quad (A12)$$

手続きは G を含むセグメントのインデックス isg と G の近似値 G_{isg} (i_s) のインデックス i_s を計算するために G^2 を使用する。現時点の量子化インデックスは次式で与えられる。

$$GInd_t = 16 \times isg + i_s \quad (A13)$$

復号化は次式により実行される。

$$Q^{-1}(GInd_t) = G_{isg}[0] + (GInd_t - \lfloor isg / 16 \rfloor) \times 2^{isg+1} \quad (A14)$$

ここで、 $\lfloor x \rfloor$ は x を超えない最大の整数とする。

A4.4 S I Dパラメータの計算と符号化

| | | |
|------------------|----------------------------------|------------------------------------|
| File : COD_CNG.C | Procedure: Cod_Cng() | COD_CNG main body |
| File : COD_CNG.C | Procedure: ComputePastAvFilter() | Computes past average filter |
| File : COD_CNG.C | Procedure: LpcDiff() | Itakura distance comparison |
| File : COD_CNG.C | Procedure: CalcRc() | Compute function RC from LPC |
| File : LPC.C | Procedure : Durbin() | Levinson-Durbin recursion |
| File : LSP.C | Procedure : AtoLsp() | Converts LPC coefficients into LSP |
| File : LSP.C | Procedure: Lsp_Qnt() | LSP quantization |
| File : LSP.C | Procedure: Lsp_Inq() | LSP inverse quantization |

現フレームがS I Dフレームの場合、S I Dパラメータが計算され量子化される。これらのパラメータは次のS I Dフレームまで続く無音フレームに対するS I D決定に利用されることに注意すること。

S I D L P Cフィルタ $A_{sid}(z)$ の計算とVAD L P Cフィルタ $A_{no}(z)$ の更新

最初に、現フレームの直前3フレームからなる過去平均L P Cフィルタ $\bar{A}_p(z)$ が、次に示す自己相関関数を入力としたダービン手順を使用して見積もられる。

$$\bar{R}_p[j] = \sum_{k=t-3}^{t-1} R^k[j], \text{ for } j = 0 \text{ to } 10 \quad (A15)$$

自己相関関数 $R^k[j]$ は式(A8)で計算されたものを累積したものである。

過去平均LPCフィルタ係数は、 $\bar{a}_p[j]$, $j = 1, \dots, 10$ で表される。

A 2 節で使われたVAD雑音LPCフィルタが、適応イネーブルフラグ Aen_t が許可されている時にだけ $\bar{a}_p[j]$ に更新される。

$$\text{if } Aen_t = 0 \text{ then } a_{no}[j] = \bar{a}_p[j], j = 1, 2, \dots, 10 \quad (\text{A16})$$

$$\text{Then } A_{sid}(z) = \begin{cases} A_t(z) & \text{if the distance between } A_t(z) \text{ and } \bar{A}_p(z) \text{ is } \geq thr1 \\ \bar{A}_p(z) & \text{otherwise} \end{cases} \quad \text{see eq. (A10)}$$

現LPCフィルタ $A_t(z)$ と平均過去LPCフィルタ $\bar{A}_p(z)$ の距離は、A 4.2 節と同様に計算される。

新しいSIDLPCフィルタ係数 $a_{sid}[j]$, $j = 1, \dots, 10$ は、LSPに変換され、符号器のLSP24ビット量子化手続き(TTC標準JT-G723.1 2.5節参照)により量子化される。復号値を \tilde{p}_{sid} と表す。

SID利得

SID利得の量子化値は次式で与えられる。

$$GInd_{sid} = GInd_t \quad (\text{A17})$$

復号値は、 \tilde{G}_{sid} で表す。

A4.5 CNG励振の計算

| | | |
|-------------------|----------------------------|--------------------------------|
| File : UTIL_CNG.C | Procedure: Calc_Exc_Rand() | Computation of the excitation |
| File : UTIL_CNG.C | Procedure: random_number() | Random number generation |
| File : UTIL_CNG.C | Procedure: distG() | Used to select excitation Gain |
| File : UTIL_LBC.C | Procedure : Sqrt_lbc() | Square root |
| File : UTIL_LBC.C | Procedure : Rand_lbc() | Pseudo-random sequence |

励振信号の更新は、SIDフレームと非伝送フレーム両者に対して行われる。

最初に現フレーム t の合成励振として得られる平均エネルギーの自乗平均としてターゲット励振利得 \tilde{G}_t を定義する。 \tilde{G}_t は次の平均化処理により計算される。

$$\tilde{G}_t = \begin{cases} \tilde{G}_{sid} & \text{if } Vad_{t-1} = 1 \\ \frac{7}{8}\tilde{G}_{t-1} + \frac{1}{8}\tilde{G}_{sid} & \text{otherwise} \end{cases} \quad (\text{A18})$$

フレームの240サンプルが120サンプル毎の2ブロックに分割される。それぞれのブロックは60サンプルの2つのサブフレームから構成される。それぞれのブロックに対して、CNG励振サンプルが次のアルゴリズムで合成される。

最初に、2サブフレームのLTPパラメータが選択される。

- (1) 第1サブフレームに対するピッチ遅延が区間[123, 143]でランダムに選ばれる。
- (2) 第2サブフレーム利得ベクトルインデックスが170エントリの利得コードブックの最初の50ベクトルに相当する[0, 49]からランダムに選ばれる。
- (3) 第2サブフレームの遅延オフセットは最初のブロックに対して0、第2ブロックに対して3が取られる。

次に、2サブフレームの固定コードブックベクトルが高レート固定励振パターンに対応するグリッド、パルスの極性、位置のランダム選択法により組み立てられる。

そして、唯一の固定励振利得がブロックの2サブフレームに対して計算される。

現ブロックの適応励振ベクトルは $u[n]$ 、 $n=0, \dots, 119$ 、固定励振ベクトルは $v[n]$ 、 $n=0, \dots, 119$ と示される。固定励振利得は、ターゲットエネルギー \tilde{G}_t^2 に最も近いブロック平均エネルギーを生じる値 Gf を計算することによって得られる。

$$\text{select } Gf \text{ such that } \left| \frac{1}{120} \sum_{n=0}^{119} (u[n] + Gf \times v[n])^2 - \tilde{G}_t^2 \right| \text{ minimum} \quad (\text{A19})$$

Gf は負の値をとり得ることに注意すること。

まず、以下の係数を持つ $C(X) = aX^2 + 2bX + c$ を定義する。

$$a = \left(\sum_{n=0}^{119} v[n]^2 \right), \quad b = \left(\sum_{n=0}^{119} u[n]v[n] \right), \quad c = \left(\sum_{n=0}^{119} u[n]^2 - 120\tilde{G}_t^2 \right)$$

次に式 $C(X) = 0$ を考える。

もし、識別式がゼロまたは負のときは、 $Gf = -\frac{b}{a}$ が選択される。さもなければ、2つの根が計算され、最小の絶対値を持つ方が選択される。

次に、 Gf は $Gf \leq 5000$ に制限される。

最後に、ブロックCNG励振が次式で構成される。

$$e[n] = u[n] + Gf \times v[n], n = 0 \text{ to } 119. \quad (\text{A20})$$

A.4.6 LSPの補間と更新

| | | |
|------------------|----------------------|-------------------|
| File : COD_CNG.C | Procedure: Cod_Cng() | COD_CNG main body |
| File : LSP.C | Procedure: Lsp_Int() | LSP Interpolator |

SID フレームと非伝送フレーム両方に対して、LPC 係数の補間は \tilde{p}_{sid} とCOD-CNGに供給される過去のLSPベクトル \tilde{p}_{t-1} から計算される。LSPの更新も $\tilde{p}_t = \tilde{p}_{sid}$ として行われる。

A4.7 COD-CNGの初期化

フレームの自己相関関数、ターゲット励振利得、VAD情報、CNG励振計算に使用するランダム発生器の種に対して以下の初期化を実施しなければならない。

$$\left\{ \begin{array}{l} R^k[j] = 0 \text{ for } j = 0, \dots, 10 \text{ and } k = -1, -2, -3 \\ \tilde{G}_{-1} = 0 \\ Vad_{-1} = 1 \\ rseed = 12345 \end{array} \right.$$

COD-CNGの他の静的変数に対して初期化は必要ない。

A 5. 復号器部の記述

受信側において、DEC-CNG処理がSIDフレームと非伝送フレームに行われ、合成擬似背景雑音が生産される。フレーム消失を扱うために開発された手続きが以下に記述されている。

A5.1 DEC-CNGの記述

| | | |
|-------------------|----------------------------|--------------------------------|
| File : DEC_CNG.C | Procedure: Dec_Cng() | DEC_CNG main body |
| File : UTIL_CNG.C | Procedure: Calc_Exc_Rand() | Computation of the excitation |
| File : UTIL_CNG.C | Procedure: random_number() | Random number generation |
| File : UTIL_CNG.C | Procedure: distG() | Used to select excitation Gain |
| File : UTIL_LBC.C | Procedure : Sqrt_lbc() | Square root |
| File : UTIL_LBC.C | Procedure : Rand_lbc() | Pseudo-random sequence |
| File : LSP.C | Procedure: Lsp-Inq() | LSP inverse quantization |
| File : LSP.C | Procedure: Lsp_Int() | LSP Interpolator |
| File : UTIL_CNG.C | Procedure: Qua_SidGain() | Quantize Sid Gain |
| File : UTIL_CNG.C | Procedure: Dec_SidGain() | Decode Sid Gain |

Figure A-5-1 / JT-G723.1 に、復号器部の擬似背景雑音生成器の概要図を示す。

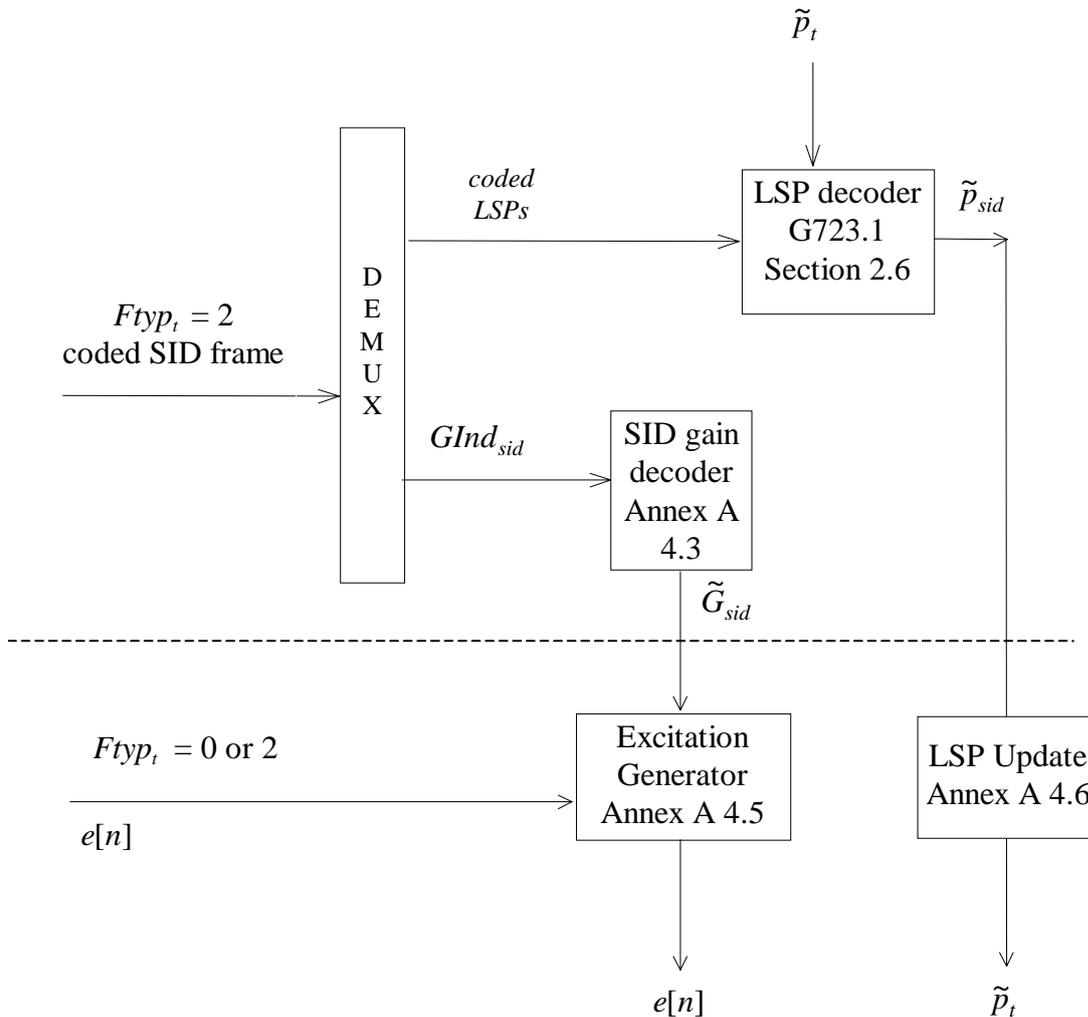


Figure A-5-1 / JT-G723.1 Block diagram of the CNG at the decoder part (ITU-T G.723.1)

復号器がSIDフレームを受信したとき、DEC-CNG処理でSIDパラメータを復号する。

SIDフレームと非伝送フレームの両者に対し、DEC-CNGモジュールは、復号化されたSIDパラメータを使用し、LSPと、復号器合成モジュールによって合成される擬似背景雑音の励振信号を計算する。受信機で提供されるフレーム情報のCNGタイプ F_{typ_t} (フレーム t に対する) は、符号器のCOD-CNGで計算された値と同じである。

- (1) $F_{typ_t} = 2$ のとき、SIDフレームパラメータは、LSPに対して \tilde{p}_{sid} 、復号化利得に対して \tilde{G}_{sid} として復号化される。
- (2) $F_{typ_t} = 0$ のとき、SID情報が消失していなかったことを検証するために $F_{typ_{t-1}}$ がテストされる (A 5. 2節参照)。もし $F_{typ_{t-1}} = 1$ であれば、TTC標準JT-G 7 2 3. 1復号器の最終有音フレーム処理で計算されたエネルギー項 Enr が量子化され、スケーリングファクター α_w がない場合を除いて、3. 1節の平均エネルギーと同じ手続きにより復号化される。復号化された値は、 \tilde{G}_{sid} に再度格納される。

次に、どちらの場合においても、CNG励振がA 4.5節COD-CNGに記述した手続きに従って計算される。新たなLSPベクトル \tilde{p}_{sid} が補間LPC係数の計算に使用され、LSP更新が $\tilde{p}_t = \tilde{p}_{sid}$ として実行される。

A5.2 CNGに関するフレーム消失補償

| | | |
|----------------|--------------------|----------------|
| File : DECOD.C | Procedure: Decod() | Frame decoding |
|----------------|--------------------|----------------|

復号器でフレーム消失を検出した時には、消失したフレームのタイプは前のフレームのタイプに依存する。

- (1) 前のフレームが有音の時には、現在の消失フレームも有音と考える。
- (2) 前のフレームがSIDあるいは非伝送フレームの時には、現フレームは非伝送フレームと考える。

$$\begin{cases} Ftyp_{t-1} = 1 & \Rightarrow Ftyp_t = 1 \\ Ftyp_{t-1} = 0 \text{ or } 2 & \Rightarrow Ftyp_t = 0 \end{cases} \quad (A21)$$

非伝送フレームが消失した時には、誤りなしと扱う。

SIDフレームが消失した時には、二つの可能性がある。

- (1) 現在の無音区間の最初のSIDフレームでない時には、過去のSIDパラメータが保持される。
- (2) 無音区間の最初のSIDフレームの時には、特別な保護が行われる。

A5.1節で述べたように、この場合には $Ftyp_{t-1} = 1$ and $Ftyp_t = 0$ と検出される。この状態の組み合わせは、前のフレームが有効な有音フレームだったとは限らない。すなわち、前のフレームまでの数フレームは消失していたかもしれない。確かなことは、最後の有効なフレームは有音のフレームであり、現在のフレーム消失はなく、現在の非伝送フレームに対する情報を与えると思われるSIDフレームは失われていることである。

SID情報を補うために、DEC-CNGはTTC標準JT-G723.1の復号器によって与えられる以下の主要なパラメータを使う。

- (1) 最後の明らかに有音だったフレームのLSPsが \tilde{p}_{sid} に対して使われる。
- (2) 最後の明らかに有音だったフレームの最後の120サンプルに渡る残差信号の補間処理(TTC標準JT-G723.1の3.10.2節参照)で復号器で計算されるエネルギー Enr が、5.1章で記述されている方法により \tilde{G}_{sid} を補償するために使われる。

最後に、駆動信号の計算に使用されるランダム発生の不一致を避けるため、符号器と復号器の双方において、それぞれの有音フレームで $rseed = 12345$ なる擬似ランダム系列のリセットを行う。

A5.3 DEC-CNGの初期化

以下の変数だけが初期化される。

$$\left\{ \begin{array}{l} \tilde{G}_{sid} = 0 \\ \tilde{p}_{sid} = LSP\ DC\ vector\ p_{DC} \\ Vad_{-1} = 1 \\ rseed = 12345 \end{array} \right.$$

A6. ビット列のパッキング

以下の表は、TTC標準JT-G723.1の4章で用いられている記述によるSIDフレームのビット列を示す。

Table A-6-1 / JT-G723.1 Bit packing for SID frames
(ITU-T G723.1)

| TRANSMITTED OCTETS | PARx_By, ... |
|--------------------|--|
| 1 | LPC_B5 ... LPC_B0, VADFLAG_B0, RATEFLAG_B0 |
| 2 | LPC_B13 ... LPC_B6 |
| 3 | LPC_B21 ... LPC_B14 |
| 4 | GAIN_B5 ... GAIN_B0, LPC_B23, LPC_B22 |

A 7. 記号解説

| | |
|-------------------|---|
| $a_{no}[j]$ | noise LPC filter coefficients |
| L_{ol} | preceding frame and current frame open loop pitch delays |
| pc | pitch delays counter for voicing estimation |
| Aen_t | adaptation enable flag |
| $e'_i[n]$ | noise-inverse filtered input signal for frame t |
| Enr_t | noise-inverse filtered input signal energy for frame t |
| $Nlev_t$ | noise level at frame t |
| $Nlev_{min}$ | minimum bound on $Nlev_t$ |
| $Nlev_{max}$ | maximum bound on $Nlev_t$ |
| Thr | adapted threshold for VAD decision |
| $k'_i[2]$ | second reflection coefficient for subframe i in frame t |
| $SinD$ | sine wave detection flag (1 : sine detected, 0 : else) |
| $e[n]$ | decoded combined excitation vector |
| $R_i[j]$ | autocorrelation function for subframe $i, j = 0, 1, \dots, 10$ |
| \tilde{p}_i | decoded LSP vector for frame t |
| $R'[j]$ | cumulated autocorrelation function for frame $t, j = 0, 1, \dots, 10$ |
| Vad_t | VAD decision for frame t (1 : active, 0 : non active) |
| $Ftyp_t$ | CNG type of frame t (0 : untransmitted, 1 : active, 2 : SID) |
| a_t | LPC coefficient vector for frame t |
| E_t | Residual energy for frame t |
| \bar{E} | Sum of residual energies |
| k_E | number of energies involved in the summation |
| thr_1 | threshold for Itakura distance |
| thr_2 | threshold for energies distance |
| $R_n[j]$ | modified autocorrelation of LPC coefficients |
| $Gind_{sid}$ | SID gain index |
| \tilde{G}_{sid} | decoded SID gain |
| G | excitation gain used at the SID quantizer input |
| $Gind_t$ | gain index for frame t |
| isg | SID gain quantizer segment index |
| $N[isg]$ | SID gain quantizer segment length |
| $G_{isg}[j]$ | gain decoded values of segment $isg, j=0, 1, \dots, N[isg]-1$ |
| i_s | gain index relative to the segment |
| α_w | energy scaling factor |
| a_{sid} | SID LPC coefficient vector |
| \bar{a}_p | past average LPC filter coefficients |
| $\bar{R}_p[j]$ | sum of past autocorrelation functions |
| \tilde{p}_{sid} | decoded SID LSP vector |
| $u[n]$ | adaptive codebook excitation vector |
| $v[n]$ | fixed codebook excitation vector |
| \tilde{G}_t | target excitation gain for excitation synthesis |
| a, b, c | coefficients of energy minimization equation |
| $C(X)$ | energy minimization equation |
| Gf | fixed codebook gain for CNG excitation synthesis |
| $rseed$ | random generator seed |

A 8. ビット精度の固定小数点演算 C ソースコード

無音区間の圧縮の全処理は、ビット精度の固定小数点演算 ANSI C ソースコードに含まれている。上述の記述と C ソースコードの間の不一致に関しては、C ソースコードが正しいと考える。この C ソースコードは TTC 標準 JT-G 7 2 3. 1 として日本 ITU 協会から入手できるコードの一部に相当する。

付属資料B

(JT-G 7 2 3. 1に対する)

マルチメディア通信伝送のための 5.3 および 6.3kbit/s 音声符号化方式に対する浮動小数点演算仕様

(本付属資料は仕様の一部である。)

1. 概要

TTC標準JT-G 7 2 3. 1はマルチメディア通信アプリケーションのためのデュアルレート 6.3 と 5.3kbit/s の音声コーデックのビット精度の固定小数点演算仕様を提供する。この仕様の厳密な詳細は日本ITU協会から入手できるビット精度の固定小数点演算Cコードに記述されている。この付録は浮動小数点演算Cソースコードに含まれているTTC標準JT-G 7 2 3. 1のいずれかのコーデックの装置の実現法について記述している。本標準の装置実現を容易にするために、この浮動小数点演算仕様のデジタルテストベクトルのセットも日本ITU協会より入手可能である。このテストベクトルが通ることは本標準に従っていることの必要条件ではあるが十分条件ではないことに注意すること。

2. アルゴリズム記述

TTC標準JT-G 7 2 3. 1の浮動小数点演算版は固定小数点演算版と同一のアルゴリズム手順を有する。同様に、ビット列もTTC標準JT-G 7 2 3. 1のビット列と一致する。詳細はTTC標準JT-G 7 2 3. 1の本文の2章、3章、4章および6章を参照のこと。

3. ANS I Cコード

本デュアルレート符号器・復号器を浮動小数点演算でシミュレーションできるANS I Cコードは、日本ITU協会から入手できる。このCコードに含まれるすべてのファイルを以下の表に示す。個々のCの関数名はTTC標準JT-G 7 2 3. 1の本文中の名前と同一である。

Table B – 1 – 1 / JT-G723.1 list of software filenames
(ITU-T G.723.1)

| File name | Description |
|------------|---|
| TYPEDEF2.H | data type definition is machine dependent |
| CST2.H | definition of constants for G.723.1 |
| LBCCODE2.C | main program for G.723.1 speech codecs |
| LBCCODE2.H | function prototypes |
| CODER2.C | G.723.1 speech encoder for the two bit rates |
| CODER2.H | function prototypes |
| DECOD2.C | G.723.1 speech decoder for the two bit rates |
| DECOD2.H | function prototypes |
| LPC2.C | linear predictive analysis |
| LPC2.H | function prototypes |
| LSP2.C | line spectral pair related functions, quantizer |
| LSP2.H | function prototypes |
| EXC2.C | adaptive and fixed (MP-MLQ, ACELP) excitation |
| EXC2.H | function prototypes |
| UTIL2.C | miscellaneous functions (HPF, pack, unpack, I/O...) |
| UTIL2.H | function prototypes |
| TAB2.C | tables of constants |
| TAB2.H | external declaration for constant tables |

付属資料C

(JT-G 7 2 3. 1に対する)

マルチメディア通信伝送のための5.3および6.3kbit/s音声符号化方式を適用した無線システムへの
応用に関する可変レートチャンネルコーデック

(本付属資料は仕様の一部である。)

1. 概要

1.1 範囲

本標準は、TTC標準JT-G 7 2 3. 1の3種類のビットレートの音声コーデックに使用可能なチャンネルコーデックについて記述している。チャンネルコーデックはビットレート可変が可能で、移動通信におけるマルチメディアに対するTTC標準JT-H 3 2 4ファミリーの一部として設計されている。この標準により、TTC標準JT-G 7 2 3. 1は、いかなる有線または無線伝送システムに適応できる。本標準では、インタリーブやバーストフォーマッティングのような伝送システムに依存する機能については論じない。

1.2 ビットレート

チャンネルコーデックがサポートするビットレート幅は0.7kbit/sから14.3kbit/sまでである。チャンネルコーデックはTTC標準JT-G 7 2 3. 1における高ビットレート、低ビットレート及び不連続伝送の3種類のモードすべてをサポートする。

伝送システムやアプリケーションに依存して適用されるチャンネルコーデックのビットレートは本標準では規定しない。

1.3 遅延

本標準で記述されているチャンネル符号器とチャンネル復号器による付加的な遅延は発生しない。

1.4 チャンネルコーデックの記述

チャンネルコーデックはバンクチャド畳込み符号を基本としている。与えられたマザーコードから多数のビットレートを生み出すことにより可変であることを特徴としている。情報ビットの中のそれぞれのクラスの主観的な重要さに基づいて、利用できるチャンネルコーデックのビットレートは最適なビットクラスに割り当てられる。このビット割り当ては、符号器及び復号器で既知なアルゴリズムに基づいている。TTC標準JT-G 7 2 3. 1のレートまたは利用できるチャンネルコーデックのビットレートによってシステムコントロール信号が変わるたびに、チャンネルコーデックを新しい音声サービス構成に適するようにこのアルゴリズムが実行される。

もし非常に低いチャンネルコーデックのビットレートのみが利用可能な時には、主観的に最も敏感なビットがまず保護される。チャンネルコーデックのビットレートが増加した時には、付加的なビットはまずさらなる情報ビットを保護するために使用され、次にすでに保護されているクラスの保護をさらに高めるために使用される。

チャンネルコーデックの特殊な構成は、主観的に敏感なビットに対して畳込み符号を使用せず、誤り検出符号のみを使用して、伝送誤りに強いTTC標準JT-G723.1伝送を可能とする。この構成でのチャンネルコーデックのビットレートは1秒あたり667ビットである。

チャンネル符号化のアプリケーションよりも優先的に伝送誤りに対してより強くするために音声パラメータはチャンネル適応レイヤにおいて部分的に修正される。以下の二種類のパラメータによりチャンネルコーデックを構成し、制御している。

- (1) TTC標準JT-G723.1構成ビット：音声符号器のモードを制御する信号
- (2) チャンネルコーデック構成ビット：チャンネルコーデックのビットレートによりチャンネルコーデックを制御する信号

1.5 制御プロトコルTTC標準JT-H245へのインタフェース

音声サービスに利用できるチャンネルコーデックのビットレートは、1ビットのチャンネルコーデックの構成ビットを用いることにより制御プロトコルTTC標準JT-H245から与えられる。チャンネルコーデック構成ビットとTTC標準JT-G723.1構成ビットは保護され、保護された情報ビット列とともに復号器へ伝送される。情報ビット列の復号化よりも優先して構成ビットは復号され、それによってチャンネル及び音声復号器が構成される。

2. チャンネル符号器

2.1 有音モードのためのチャンネル符号器の適応化

本節では、TTC標準JT-G723.1の音声パラメータを、伝送誤りに対してより強い修正されたビット列に変換する方法を述べる。

2.1.1 利得インデックスの伸長

音声パラメータのフォーマットは、Table 4-1, 4-2/JT-G723.1に指定されている。各サブフレームの12ビット総利得インデックス(GAIN_x_By)は、2つの個別利得インデックスを圧縮したインデックスである。これらのインデックスは、音声復号化部分(TTC標準JT-G723.1の2.17、2.18節)に記述されている方法で伸長される。その結果、圧縮されたインデックスと比較すると1ビット付加されて、1サブフレーム毎に1つの8ビット利得インデックス(AGAIN_x_By)と1つの5ビット利得インデックス(FGAIN_x_By)が得られる。この伸長処理によるオーバーヘッドは、誤り検出と誤り補償を目的とするチャンネル復号器で利用される。

2.1.2 LPC再構成

24ビットLPCインデックス(LPC_{Bx}, x=0, 23)は、TTC標準JT-G723.1の2.6節に記述されているように3つの8ビットサブベクトルから成る。

$$\begin{aligned} e_0 &= \{LPC_B7, LPC_B6, LPC_B5, LPC_B4, LPC_B3, LPC_B2, LPC_B1, LPC_B0\} \\ e_1 &= \{LPC_B15, LPC_B14, LPC_B13, LPC_B12, LPC_B11, LPC_B10, LPC_B9, LPC_B8\} \\ e_2 &= \{LPC_B23, LPC_B22, LPC_B21, LPC_B20, LPC_B19, LPC_B18, LPC_B17, LPC_B16\} \end{aligned}$$

サブベクトルは、次式により再構成される。

$$e_m^R = \text{ReorderTab}_m[e_m] \quad m = 0,1,2$$

LPC再構成表は、最初のサブベクトルについては Table C-2-1 a/JT-G723.1 に、2番目については Table C-2-1 b/JT-G723.1 に、3番目については Table C-2-1 c/JT-G723.1 に示される。表に記載されている (i, j) の値は、 $e_m = 16 \cdot j + i$ によって決定される。得られた出力は、サブベクトルの送信インデックスとなる。

Table C-2-1 a/JT-G723.1: Reordering table for m=0 ($\text{ReorderTab}_0[e_0]$)
(ITU-T G.723.1)

| | i=0 | i=1 | i=2 | i=3 | i=4 | i=5 | i=6 | i=7 | i=8 | i=9 | i=10 | i=11 | i=12 | i=13 | i=14 | i=15 |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|
| j=0 | 82 | 91 | 190 | 191 | 189 | 36 | 187 | 32 | 38 | 39 | 185 | 112 | 166 | 175 | 116 | 120 |
| j=1 | 34 | 35 | 122 | 41 | 40 | 43 | 42 | 56 | 57 | 90 | 168 | 85 | 74 | 170 | 234 | 174 |
| j=2 | 169 | 172 | 178 | 182 | 184 | 179 | 181 | 180 | 37 | 186 | 44 | 33 | 159 | 183 | 188 | 155 |
| j=3 | 253 | 252 | 147 | 154 | 246 | 165 | 218 | 139 | 63 | 160 | 157 | 146 | 145 | 144 | 177 | 176 |
| j=4 | 244 | 131 | 148 | 129 | 128 | 161 | 219 | 135 | 34 | 203 | 200 | 206 | 207 | 204 | 205 | 254 |
| j=5 | 212 | 222 | 213 | 220 | 221 | 141 | 216 | 88 | 38 | 137 | 136 | 217 | 133 | 132 | 201 | 197 |
| j=6 | 196 | 76 | 77 | 243 | 192 | 195 | 193 | 119 | 18 | 108 | 121 | 249 | 247 | 245 | 130 | 240 |
| j=7 | 241 | 235 | 69 | 68 | 210 | 226 | 227 | 224 | 25 | 79 | 211 | 250 | 251 | 255 | 198 | 242 |
| j=8 | 194 | 127 | 45 | 46 | 117 | 125 | 124 | 2 | 67 | 3 | 101 | 72 | 0 | 13 | 97 | 65 |
| j=9 | 66 | 98 | 96 | 158 | 248 | 228 | 229 | 93 | 31 | 109 | 100 | 110 | 111 | 75 | 73 | 115 |
| j=10 | 114 | 23 | 94 | 95 | 6 | 22 | 7 | 230 | 5 | 4 | 99 | 81 | 20 | 21 | 83 | 113 |
| j=11 | 102 | 14 | 103 | 15 | 12 | 16 | 52 | 17 | 19 | 1 | 29 | 8 | 28 | 30 | 31 | 11 |
| j=12 | 9 | 126 | 24 | 18 | 26 | 25 | 27 | 86 | 87 | 64 | 71 | 70 | 153 | 152 | 173 | 162 |
| j=13 | 167 | 164 | 89 | 80 | 238 | 208 | 232 | 233 | 15 | 236 | 239 | 92 | 202 | 209 | 199 | 143 |
| j=14 | 142 | 140 | 223 | 84 | 78 | 171 | 151 | 156 | 37 | 150 | 149 | 214 | 106 | 107 | 10 | 53 |
| j=15 | 47 | 54 | 55 | 104 | 105 | 123 | 48 | 58 | 59 | 49 | 50 | 51 | 62 | 60 | 61 | 63 |

Table C-2-1 b/JT-G723.1: Reordering table for m=1 ($ReorderTab_1[e_1]$)
(ITU-T G.723.1)

| | i=0 | i=1 | i=2 | i=3 | i=4 | i=5 | i=6 | i=7 | i=8 | i=9 | i=10 | i=11 | i=12 | i=13 | i=14 | i=15 |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|
| j=0 | 122 | 150 | 201 | 213 | 214 | 212 | 246 | 244 | 136 | 135 | 130 | 134 | 140 | 216 | 218 | 146 |
| j=1 | 209 | 208 | 110 | 153 | 156 | 158 | 152 | 144 | 147 | 145 | 199 | 197 | 193 | 204 | 205 | 249 |
| j=2 | 248 | 127 | 185 | 172 | 174 | 173 | 179 | 169 | 168 | 181 | 182 | 177 | 183 | 161 | 178 | 167 |
| j=3 | 163 | 162 | 160 | 165 | 232 | 234 | 235 | 233 | 237 | 236 | 238 | 239 | 78 | 155 | 154 | 75 |
| j=4 | 74 | 69 | 68 | 71 | 70 | 73 | 72 | 139 | 123 | 114 | 59 | 58 | 113 | 170 | 137 | 141 |
| j=5 | 186 | 133 | 151 | 132 | 148 | 159 | 196 | 149 | 187 | 189 | 191 | 157 | 198 | 190 | 121 | 126 |
| j=6 | 125 | 120 | 57 | 112 | 171 | 63 | 45 | 62 | 47 | 46 | 99 | 109 | 98 | 103 | 124 | 102 |
| j=7 | 107 | 106 | 129 | 128 | 131 | 89 | 118 | 105 | 104 | 95 | 91 | 90 | 94 | 86 | 83 | 82 |
| j=8 | 138 | 119 | 117 | 88 | 116 | 92 | 37 | 36 | 38 | 32 | 96 | 87 | 81 | 39 | 93 | 49 |
| j=9 | 48 | 23 | 22 | 21 | 20 | 29 | 8 | 56 | 55 | 44 | 175 | 60 | 61 | 53 | 43 | 143 |
| j=10 | 115 | 176 | 180 | 142 | 166 | 184 | 188 | 207 | 206 | 231 | 252 | 195 | 194 | 219 | 217 | 221 |
| j=11 | 223 | 211 | 215 | 210 | 192 | 200 | 164 | 226 | 202 | 203 | 230 | 227 | 229 | 253 | 255 | 224 |
| j=12 | 225 | 243 | 228 | 220 | 240 | 241 | 242 | 222 | 245 | 251 | 250 | 247 | 254 | 111 | 108 | 77 |
| j=13 | 76 | 79 | 101 | 100 | 30 | 15 | 97 | 52 | 54 | 41 | 50 | 51 | 40 | 33 | 42 | 35 |
| j=14 | 34 | 13 | 12 | 14 | 85 | 84 | 25 | 31 | 27 | 80 | 17 | 67 | 66 | 16 | 9 | 28 |
| j=15 | 24 | 65 | 11 | 10 | 26 | 64 | 19 | 18 | 4 | 1 | 0 | 5 | 7 | 6 | 3 | 2 |

Table C-2-1 c/JT-G723.1: Reordering table for m=2 ($ReorderTab_2[e_2]$)
(ITU-T G.723.1)

| | i=0 | i=1 | i=2 | i=3 | i=4 | i=5 | i=6 | i=7 | i=8 | i=9 | i=10 | i=11 | i=12 | i=13 | i=14 | i=15 |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|
| j=0 | 243 | 154 | 190 | 170 | 189 | 169 | 140 | 188 | 171 | 168 | 166 | 164 | 165 | 216 | 146 | 132 |
| j=1 | 242 | 234 | 158 | 144 | 252 | 147 | 148 | 13 | 74 | 75 | 19 | 77 | 76 | 79 | 78 | 70 |
| j=2 | 71 | 198 | 199 | 72 | 12 | 73 | 220 | 150 | 151 | 131 | 204 | 133 | 205 | 230 | 229 | 225 |
| j=3 | 200 | 201 | 226 | 227 | 196 | 192 | 197 | 195 | 202 | 203 | 136 | 143 | 142 | 175 | 207 | 206 |
| j=4 | 137 | 138 | 139 | 174 | 173 | 172 | 240 | 228 | 209 | 224 | 238 | 239 | 128 | 130 | 145 | 250 |
| j=5 | 236 | 237 | 255 | 248 | 253 | 232 | 153 | 179 | 181 | 180 | 235 | 233 | 152 | 156 | 184 | 178 |
| j=6 | 185 | 162 | 183 | 182 | 163 | 160 | 161 | 167 | 157 | 159 | 155 | 134 | 186 | 191 | 135 | 187 |
| j=7 | 141 | 254 | 177 | 251 | 176 | 33 | 244 | 245 | 61 | 60 | 34 | 35 | 32 | 249 | 126 | 41 |
| j=8 | 62 | 59 | 58 | 63 | 57 | 56 | 42 | 43 | 25 | 29 | 40 | 125 | 120 | 24 | 26 | 127 |
| j=9 | 123 | 124 | 122 | 48 | 121 | 46 | 47 | 105 | 104 | 109 | 108 | 28 | 116 | 112 | 113 | 247 |
| j=10 | 246 | 36 | 54 | 149 | 55 | 4 | 52 | 53 | 49 | 37 | 39 | 51 | 50 | 107 | 211 | 215 |
| j=11 | 20 | 21 | 17 | 3 | 5 | 223 | 222 | 0 | 231 | 129 | 218 | 9 | 217 | 219 | 194 | 221 |
| j=12 | 193 | 210 | 38 | 2 | 1 | 212 | 14 | 15 | 241 | 44 | 208 | 8 | 45 | 11 | 10 | 110 |
| j=13 | 68 | 69 | 103 | 100 | 213 | 214 | 16 | 106 | 18 | 111 | 22 | 23 | 6 | 7 | 64 | 65 |
| j=14 | 27 | 118 | 117 | 102 | 31 | 30 | 97 | 114 | 115 | 96 | 88 | 89 | 90 | 91 | 99 | 94 |
| j=15 | 80 | 82 | 83 | 119 | 98 | 101 | 92 | 95 | 66 | 67 | 84 | 93 | 85 | 81 | 86 | 87 |

2.1.3 伝送路に適応化されたビット列のフォーマット

利得伸長とLPC再構成後の高、低レート音声パラメータのフォーマットは、Table C-2-2 a/JT-G723.1 と Table C-2-2 b/JT-G723.1 に示される。(UBは“未使用ビット”を表す)

Table C-2-2 a / JT-G723.1: Format of decompressed high rate codec speech parameters
(ITU-T G.723.1)

| Channel Encoder Octets | PARx_By |
|---------------------------|---|
| 1 | R_LPC_B5...R_LPC_B0, VAD, RATE |
| 2 | R_LPC_B13...R_LPC_B6 |
| 3 | R_LPC_B21...R_LPC_B14 |
| 4 | ACL0_B5...ACL0_B0, R_LPC_B23, R_LPC_B22 |
| 5 | ACL2_B4...ACL2_B0, ACL1_B1, ACL1_B0, ACL0_B6 |
| 6 | AGAIN0_B3...AGAIN0_B0, ACL3_B1, ACL3_B0, ACL2_B6, ACL2_B5 |
| 7 | AGAIN1_B3...AGAIN1_B0, AGAIN0_B7...AGAIN0_B4 |
| 8 | AGAIN2_B3...AGAIN2_B0, AGAIN1_B7...AGAIN1_B4 |
| 9 | AGAIN3_B3...AGAIN3_B0, AGAIN2_B7...AGAIN2_B4 |
| 10 | FGAIN0_B3...FGAIN0_B0, AGAIN3_B7...AGAIN3_B4 |
| 11 | FGAIN2_B1, FGAIN2_B0, FGAIN1_B4...FGAIN1_B0, FGAIN0_B4 |
| 12 | FGAIN3_B4...FGAIN3_B0, FGAIN2_B4...FGAIN2_B2 |
| 13 | MSBPOS_B3...MSBPOS_B0, GRID3_B0, GRID2_B0, GRID1_B0, GRID0_B0 |
| 14 | MSBPOS_B11...MSBPOS_B4 |
| 15 | POS0_B6...POS0_B0, MSBPOS_B12 |
| 16 | POS0_B14...POS0_B7 |
| 17 | POS1_B6...POS1_B0, POS0_B15 |
| 18 | POS2_B0, POS1_B13...POS1_B7 |
| 19 | POS2_B8...POS2_B1 |
| 20 | POS3_B0, POS2_B15...POS2_B9 |
| 21 | POS3_B8...POS3_B1 |
| 22 | PSIG0_B2...PSIG0_B0, POS3_B13...POS3_B9 |
| 23 | PSIG1_B4...PSIG1_B0, PSIG0_B5...PSIG0_B3 |
| 24 | PSIG3_B1, PSIG3_B0, PSIG2_B5...PSIG2_B0 |
| 25 | UB, UB, UB, UB, UB, PSIG3_B4...PSIG3_B2 |

Table C-2-2 b/JT-G723.1: Format of decompressed low rate codec speech parameters
(ITU-T G.723.1)

| Channel Encoder Octets | PARx_By |
|---------------------------|---|
| 1 | R_LPC_B5...R_LPC_B0, VAD, RATE |
| 2 | R_LPC_B13...R_LPC_B6 |
| 3 | R_LPC_B21...R_LPC_B14 |
| 4 | ACLO_B5...ACLO_B0, R_LPC_B23, R_LPC_B22 |
| 5 | ACL2_B4...ACL2_B0, ACL1_B1, ACL1_B0, ACLO_B6 |
| 6 | AGAIN0_B3...AGAIN0_B0, ACL3_B1, ACL3_B0, ACL2_B6, ACL2_B5 |
| 7 | AGAIN1_B3...AGAIN1_B0, AGAIN0_B7...AGAIN0_B4 |
| 8 | AGAIN2_B3...AGAIN2_B0, AGAIN1_B7...AGAIN1_B4 |
| 9 | AGAIN3_B3...AGAIN3_B0, AGAIN2_B7...AGAIN2_B4 |
| 10 | FGAIN0_B3...FGAIN0_B0, AGAIN3_B7...AGAIN3_B4 |
| 11 | FGAIN2_B1, FGAIN2_B0, FGAIN1_B4...FGAIN1_B0, FGAIN0_B4 |
| 12 | FGAIN3_B4...FGAIN3_B0, FGAIN2_B4...FGAIN2_B2 |
| 13 | POS0_B3...POS0_B0, GRID3_B0, GRID2_B0, GRID1_B0, GRID0_B0 |
| 14 | POS0_B11...POS0_B8, POS0_B7...POS0_B4 |
| 15 | POS1_B7...POS1_B4, POS1_B3...POS1_B0, |
| 16 | POS2_B3...POS2_B0, POS1_B11...POS1_B8 |
| 17 | POS2_B11...POS2_B8, POS2_B7...POS2_B4 |
| 18 | POS3_B7...POS3_B4, POS3_B3...POS3_B0 |
| 19 | PSIG0_B3...PSIG0_B0, POS3_B11...POS3_B8 |
| 20 | PSIG2_B3...PSIG2_B0, PSIG1_B3...PSIG1_B0 |
| 21 | UB, UB, UB, UB, PSIG3_B3...PSIG3_B0 |

2.2 有音モードのためのビット感度分類

Table C-2-2 a/JT-G723.1 と Table C-2-2 b/JT-G723.1 に指定される伸長された音声パラメータのビットは、高レートについては Table C-2-3a/JT-G723.1、低レートについては Table C-2-3b/JT-G723.1 によって再配置される。その番号は、構成されたビット列における相当する情報ビットの位置を示す。主観的に最も敏感なビットは、構成されたビット列の位置“0”に配置される。

構成されたビット列は、高レートのコーデックについては、

$$i(0), i(1), \dots, i(192)$$

で、低レートのコーデックについては、

$$i(0), i(1), \dots, i(162)$$

で示される。

Table C-2-3 a/JT-G723.1: High rate codec bit ordering table
(ITU-T G.723.1)

| Octet | Bit position in ordered bitstream | | | | | | | |
|-------|-----------------------------------|-----|-----|-----|-----|-----|-----|------|
| 1 | 175 | 180 | 189 | 190 | 191 | 192 | VAD | RATE |
| 2 | 98 | 73 | 107 | 154 | 167 | 168 | 169 | 170 |
| 3 | 30 | 17 | 16 | 31 | 48 | 55 | 49 | 71 |
| 4 | 6 | 4 | 0 | 2 | 11 | 26 | 10 | 14 |
| 5 | 5 | 1 | 3 | 12 | 27 | 24 | 60 | 8 |
| 6 | 44 | 66 | 62 | 82 | 25 | 61 | 9 | 7 |
| 7 | 45 | 67 | 63 | 83 | 78 | 36 | 50 | 40 |
| 8 | 46 | 68 | 64 | 84 | 79 | 37 | 51 | 41 |
| 9 | 47 | 69 | 65 | 85 | 80 | 38 | 52 | 42 |
| 10 | 56 | 99 | 159 | 185 | 81 | 39 | 53 | 43 |
| 11 | 161 | 187 | 20 | 57 | 100 | 160 | 186 | 19 |
| 12 | 22 | 59 | 102 | 162 | 188 | 21 | 58 | 101 |
| 13 | 35 | 54 | 70 | 72 | 179 | 178 | 177 | 176 |
| 14 | 13 | 15 | 23 | 28 | 29 | 32 | 33 | 34 |
| 15 | 128 | 132 | 146 | 155 | 163 | 171 | 181 | 18 |
| 16 | 76 | 86 | 90 | 94 | 103 | 108 | 112 | 116 |
| 17 | 129 | 133 | 147 | 156 | 164 | 172 | 182 | 74 |
| 18 | 183 | 87 | 91 | 95 | 104 | 109 | 113 | 117 |
| 19 | 114 | 118 | 130 | 134 | 148 | 157 | 165 | 173 |
| 20 | 184 | 75 | 77 | 88 | 92 | 96 | 105 | 110 |
| 21 | 115 | 119 | 131 | 135 | 149 | 158 | 166 | 174 |
| 22 | 136 | 124 | 120 | 89 | 93 | 97 | 106 | 111 |
| 23 | 151 | 141 | 137 | 125 | 121 | 144 | 150 | 140 |
| 24 | 127 | 123 | 145 | 152 | 142 | 138 | 126 | 122 |
| 25 | UB | UB | UB | UB | UB | 153 | 143 | 139 |

Table C-2-3 b/JT-G723.1: Low rate codec bit ordering table
(ITU-T G.723.1)

| Octet | Bit position in ordered bitstream | | | | | | | |
|-------|-----------------------------------|-----|-----|-----|-----|-----|-----|------|
| 1 | 152 | 153 | 158 | 159 | 160 | 161 | VAD | RATE |
| 2 | 69 | 64 | 70 | 91 | 145 | 140 | 147 | 146 |
| 3 | 24 | 15 | 14 | 25 | 46 | 50 | 47 | 63 |
| 4 | 4 | 6 | 0 | 2 | 11 | 18 | 10 | 13 |
| 5 | 7 | 1 | 3 | 12 | 19 | 16 | 48 | 8 |
| 6 | 42 | 59 | 55 | 65 | 17 | 49 | 9 | 5 |
| 7 | 43 | 60 | 56 | 66 | 26 | 30 | 34 | 38 |
| 8 | 44 | 61 | 57 | 67 | 27 | 31 | 35 | 39 |
| 9 | 45 | 62 | 58 | 68 | 28 | 32 | 36 | 40 |
| 10 | 51 | 87 | 141 | 154 | 29 | 33 | 37 | 41 |
| 11 | 143 | 156 | 21 | 52 | 88 | 142 | 155 | 20 |
| 12 | 23 | 54 | 90 | 144 | 157 | 22 | 53 | 89 |
| 13 | 100 | 128 | 96 | 104 | 151 | 150 | 149 | 148 |
| 14 | 132 | 112 | 116 | 136 | 108 | 120 | 124 | 92 |
| 15 | 109 | 121 | 125 | 93 | 101 | 129 | 97 | 105 |
| 16 | 102 | 130 | 98 | 106 | 133 | 113 | 117 | 137 |
| 17 | 134 | 114 | 118 | 138 | 110 | 122 | 126 | 94 |
| 18 | 111 | 123 | 127 | 95 | 103 | 131 | 99 | 107 |
| 19 | 83 | 79 | 75 | 71 | 135 | 115 | 119 | 139 |
| 20 | 85 | 81 | 77 | 73 | 84 | 80 | 76 | 72 |
| 21 | UB | UB | UB | UB | 86 | 82 | 78 | 74 |

2.3 SIDフレームのためのビット感度分類

SIDフレームのフォーマットは、TTC標準JT-G723.1付属資料Aに記載されている。このフレームのビットは、Table C-2-3 c/JT-G723.1に従って再配置される。SIDフレームでは、利得伸長もLPC再構成も行われない。その番号は、構成されたビット列における相当する情報ビットの位置を示す。主観的に最も敏感なビットは、構成されたビット列の位置“0”に配置される。

構成されたビット列は、

$$i(0), i(1), \dots, i(29)$$

で示される。

Table C-2-3 c/JT-G723.1: Bit ordering table for SID frames
(ITU-T G.723.1)

| Octet | Bit position in ordered bitstream | | | | | | | |
|-------|-----------------------------------|----|----|----|----|----|-----|------|
| 1 | 21 | 22 | 23 | 24 | 25 | 26 | VAD | RATE |
| 2 | 13 | 14 | 15 | 16 | 17 | 18 | 20 | 19 |
| 3 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 4 | 0 | 1 | 2 | 29 | 28 | 27 | 3 | 4 |

2.4 CRC符号器

主観的に最も敏感なビットは、チャネル復号器において誤り検出に用いられる5つのパリティビットによって保護される。構成されたビット列の最初の CRC_{WIN} ビットは、CRC窓の中に存在する。

Low rate codec: $CRC_{WIN} = 34$

High rate codec: $CRC_{WIN} = 44$

SID frames: $CRC_{WIN} = 30$

次の巡回符号生成多項式が適用される。

$$g(D) = D^5 + D^2 + 1$$

5つのパリティビット $p(0) - p(4)$ は、 CRC_{WIN} の構成された情報ビットの直後に配置される。

CRC符号化後のビット列のフォーマットは、次の通りである。

Low rate codec: $bs(n) = i(0), i(1), \dots, i(33), p(0), p(1), p(2), p(3), p(4), i(34), i(35), \dots, i(161); n=0 \dots 166$

High rate codec: $bs(n) = i(0), i(1), \dots, i(43), p(0), p(1), p(2), p(3), p(4), i(44), i(45), \dots, i(192); n=0 \dots 197$

SID frame: $bs(n) = i(0), i(1), \dots, i(29), p(0), p(1), p(2), p(3), p(4); n=0 \dots 34$

2.5 畳み込み符号器

情報ビット $bs(n)$ は、原多項式によって定義される、パンクチャ畳み込み符号を用いて符号化される。

$$gcc^0 = D^4 + D + 1$$

$$gcc^1 = D^4 + D^3 + D^2 + 1$$

$$gcc^2 = D^4 + D^2 + D + 1$$

マザーコードを周期1/2でパンクチャすることにより、2/4種類の間接レートを生成する。

$$\frac{12}{36} \leq r \leq \frac{12}{13}$$

パンクチャ行列は、Table C-2-4/JT-G723.1 に与えられる。

Table C-2-4/JT-G723.1: Puncturing tables (all values in hexadecimal representation)
(ITU-T G.723.1)

| Rate r | 12/13 | 12/14 | 12/15 | 12/16 | 12/17 | 12/18 | 12/19 | 12/20 | 12/21 | 12/22 | 12/23 | 12/24 |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $P_r(0)$ | D6F | D7F | D7F | D7F | DFE | FFF |
| $P_r(1)$ | 690 | 690 | 691 | 695 | 695 | 695 | 69D | 6DD | 6DF | 7DF | 7FF | FFF |
| $P_r(2)$ | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 |

| Rate r | 12/25 | 12/26 | 12/27 | 12/28 | 12/29 | 12/30 | 12/31 | 12/32 | 12/33 | 12/34 | 12/35 | 12/36 |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $P_r(0)$ | FFF |
| $P_r(1)$ | FFF |
| $P_r(2)$ | 001 | 009 | 109 | 309 | 329 | 729 | 72D | 72F | 7AF | 7BF | 7FF | FFF |

2.5.1 チャンネルコーデックのビットレート割り当て

本節では、各情報ビットに対する保護の程度を決定するリソース割り当てアルゴリズムを示す。

このアルゴリズムは、少なくとも次のイベントが起こる毎に、チャンネル符号器とチャンネル復号器で実行される。

- (1) リセット
- (2) TTC標準JT-G723.1のために割り当てられたチャンネルコーデックのビットレート変更
- (3) TTC標準JT-G723.1ソースコーデックのビットレート変更

次のレート割り当て手順が実行される。

- (1) 構成されたビット列は、5つのビットクラス $c[i]$, $i=0..4$ に分類される。ここで、 $c[0]$ は最も敏感なビットで、 $c[4]$ は最も鈍感なビットとなる。
- (2) 重み付け係数 $w[i]$, $i=0..4$ は、各クラスに割り当てられ、割り当てアルゴリズムの制御を行う。
- (3) クラス0は、 $c[0]$ -5情報ビットとクラスの終わりに配置される5つのパリティ検査符号ビットを含む。

もしチャンネルビットが有効でない場合 ($B=0$)、チャンネルビット列は伸長されたJT-G723.1ビット列と5つのパリティ検査ビットを含む。

- (1) フレーム毎のチャンネルコーデックビットの最大値は、すべてのクラスが保護比率 $r = 1/3$ となった時に達成される。

$$B_{\max} = 2 \cdot \left(\sum_{i=0}^{i=4} c[i] \right) + 4$$

もし

$$B \leq [0] + 8$$

で、フレーム毎のチャンネルコーデックビットが有効な場合、これらはクラス0のビット上に均一に分配され、次式が適用される。

$$r[0] = \frac{12}{\text{INT}\left(12 * \left(1 + \frac{B}{c[0] + 4}\right)\right)}$$

(INT means: next lower integer value).

- (2) もし $B > c[0] + 8$ で、フレーム毎のチャンネルコーデックビットが有効な場合、最初の処理として、ビットクラスの間レート $r'[i]$, $i=0 \dots 4$ は、次式により決定される。

$$r'[i] = \frac{12}{\text{MIN} \left[36, \text{NINT} \left(12 * \left(1 + \frac{B * w[i]}{c[i]} \right) \right) \right]}$$

最終レートは、次の条件により決定される。

$$r[i] = \begin{cases} 1 & \text{if } r'[i] > \frac{12}{18} \\ r'[i] & \text{otherwise} \end{cases}$$

- (3) 上記アルゴリズムで利用できるビット数より割り当てが少ないならば、2回目の処理として、すべての利用できるビットが割り当てられるまで、初めの無保護クラスからのビットは最後の保護クラスへと移される。
- (4) 利用できるビット数より割り当てが多いならば、2回目の処理として、許される最大のチャンネルコーデックのビットレートになるまで、最後の保護クラスからのビットは、初めの無保護クラスへと移される。
- (5) 本来の配置において、クラス3はテールビットを含む。このクラスが上式による無保護を受けるならば、テールビットは最後の保護クラスに添付される。

2.5.1.1 有音フレームの制御パラメータの設定

$c[i], w[i]$ は、Table C-2-5 a / JT-G723.1 の通りである。

Table C-2-5 a / JT-G723.1: Settings of control parameters for rate allocation (active speech frames)
(ITU-T G.723.1)

| i | G.723.1-6.3 | | | | | G.723.1-5.3 | | | | |
|------|-------------|------|------|----------|------|-------------|------|------|----------|------|
| | 0 | 1 | 2 | 3 | 4 | 0 | 1 | 2 | 3 | 4 |
| c[i] | 44+5crc | 44 | 46 | 47+4tail | 12 | 34+5crc | 40 | 40 | 40+4tail | 8 |
| w[i] | 0.26 | 0.29 | 0.24 | 0.21 | 0.00 | 0.24 | 0.31 | 0.24 | 0.21 | 0.00 |

2.5.1.2 SIDフレームの制御パラメータの設定

$c[i], w[i]$ は、Table C-2-5 b / JT-G723.1 の通りである。

Table C-2-5 b / JT-G723.1: Settings of control parameters for rate allocation (SID frames)
(ITU-T G.723.1)

| i | G.723.1-SID | | | | |
|------|---------------|------|------|------|------|
| | 0 | 1 | 2 | 3 | 4 |
| c[i] | 30+5crc+4tail | 0 | 0 | 0 | 0 |
| w[i] | 1.0 | 0.00 | 0.00 | 0.00 | 0.00 |

2.5.2 トレリス適応の終了

最終保護ビットのエラー率を引き下げるには

$$bs(Nlast - n); \quad n = 0 \dots 9;$$

は

$$bs(Nlast - n19 + n); \quad n = 0 \dots 9.$$

と交換する。

$Nlast$ はビット列中の最後の保護情報ビットである。

この処理はクラス 1 から 3 の中で 19 以上の情報ビットが保護されている場合に限って実行される。

2.5.3 情報ビット列の符号化

保護されている ($r[i] < 1$) 各ビット $bs(j)$ は以下の手順に従って符号化される。

- (1) クラス i の情報化ビット $bs(j)$ を 3 つの生成多項式 g_{CC}^n で符号化する。この結果は 3 つの中間チャンネルビット $u'(n)$ に出力される ($n = 0, 1, 2$)。
- (2) $u'(n)$ のパンクチャは $P_{r[i]}(n)$ によって次のように決まる。

$P_{r[i]}(n)$ のビット p が 0 なら $u'(n)$ はパンクチャされ、 $P_{r[i]}(n)$ のビット p が 0 でないなら $u'(n)$ はチャンネルビット列 $u(m)$ に挿入され m を増加する。

p は情報ビットカウンタ j によって増加するモジュロ 12 のカウンタである。

保護ビットの符号化の後、チャンネルのビット列は列 $u(m)$, $m = 0 \dots M_p - 1$ に収容される。

非保護ビットは保護チャンネルビット列 ($m = M_p$ で始まり、 $m = M_{All} - 1$ で終わる) の最後に配置される。

2.5.4 構成ビットの符号化

1 ビットのチャンネルコーデックの構成ビットは 2 ビットの音声コーデックの構成ビットに Table C-2-6/JT-G723.1 のように付加される。これによって T T C 標準 J T - G 7 2 3. 1 の音声コーデックモード (高レート、低レート、SID) のそれぞれに、異なる 2 つのチャンネルコーデックビットレートの 1 つを付加することができる。チャンネルコーデックのビットレートの選択は伝送系に依存するので、本標準の主題ではない。

Table C-2-6/JT-G723.1: G.723.1 and channel codec configuration bits
(ITU-T G.723.1)

| Octet | Configuration Bits | | | | | | | |
|-------|--------------------|----|----|----|----|-----|-----|------|
| 1 | UB | UB | UB | UB | UB | CCC | VAD | RATE |

下記の生成多項式による (13,5) ブロック符号が構成ビットの誤り保護のために使われる。

$$g_{cb}(D) = D^{10} + D^8 + D^5 + D^4 + D^2 + D + 1$$

13 個の構成ビット ($ucb[n]$, $n=0 \dots 12$) は符号化情報ビットよりも前に伝送される。

Table C-2-7/JT-G723.1 はパックされたチャネルビット列を示す (M_{All} の値によって未使用のビット"UB"が最後のオクテットにある)。

Table C-2-7/JT-G723.1: Channel bitstream packing
(ITU-T G.723.1)

| Transmitted Octet | Channel Bit | | | | | | | | |
|-------------------|-------------|--------|------------|------------------|------------------|------------------|------------------|------------------|--|
| 1 | ucb[7] | ucb[6] | ucb[5] | ucb[4] | ucb[3] | ucb[2] | ucb[1] | ucb[0] | |
| 2 | u[2] | u[1] | u[0] | ucb[12] | ucb[11] | ucb[10] | ucb[9] | ucb[8] | |
| 3 | u[10] | u[9] | u[8] | u[7] | u[6] | u[5] | u[4] | u[3] | |
| $M_p/8+2$ | ... | ... | u[M_p] | u[M_p-1] | u[M_p-2] | ... | ... | ... | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| $M_{All}/8+2$ | UB | UB | UB | u[$M_{All}-1$] | u[$M_{All}-2$] | u[$M_{All}-3$] | u[$M_{All}-4$] | u[$M_{All}-5$] | |

保護されたビット列のビットレートは $(M_{All} + 13) \cdot \frac{1000}{30}$ bit/s である。

3. チャネル復号器

3.1 構成ビットの復号化

最初の13ビットの復号化から音声及びチャネルコーデックの構成を得る。2.5.1節に述べられたレート割り当てアルゴリズムは、初期化、又は配列ビットの1つが変化した場合に実行される。構成ビットの復号化が信頼性に欠ける場合は、バッドフレームの表示(3.3節参照)がセットされる。

3.2 有音モードのための音声復号器の適応

チャネルコーデックの音声パラメータフォーマットはTTC標準JT-G723.1の音声パラメータフォーマットに適応される。

3.2.1 利得インデックスの圧縮

符号化処理が高レート、低レートどちらのモードでなされようとも、TTC標準JT-G723.1の2.1.7節、2.1.8節に述べられた利得圧縮処理は、音声復号器で要求されるビット列フォーマットを作成するために実行されなければならない。サブフレーム当たり1つの8ビット(AGAINx_By)利得インデックスと1つの5ビット(FGAINx_By)利得インデックスとが12ビットの利得インデックスに圧縮される。復号化された適応的コードブックの利得インデックス、又は復号化された固定励振利得インデックスのどちらかで、禁止されたインデックスが生じた場合は、禁止インデックス表示フラグ(FII)がセットされ、インデックスは適切な有効値で置き換えられる。禁止された遅延インデックスが復号化された場合(TTC標準JT-G723.1の2.1.8節参照)は、FIIフラグがセットされる。その後、復号化された遅延は適切な有効値で置き換える。

3.2.2 LPC の構成

| Speech Decoder Octets | PAR _x _By,... |
|--------------------------|--------------------------|
| 1 | LPC_B5...LPC_B0,VAD,RATE |
| 2 | LPC_B13...LPC_B6 |
| 3 | LPC_21...LPC_B14 |

LPC の 3 つのサブベクトルが 2.1.2 節の逆テーブルを適用して構成される。

$$e_m = ReorderTab_m^{inv} \left[e_m^R \right] \quad m = 0,1,2$$

3.3 誤り解析

受信された TTC 標準 JT-G.723.1 音声パラメータの信頼性を表示する以下の 2 つのフラグは、チャンネル復号器の中で生成される。

- (1) バッドフレーム表示 (BFI) は、CRC 窓 (CRC_{WIN}) のビットの中に訂正されていないエラーがあることを示す。
- (2) 誤りフレーム表示 (EFI) は、CRC 符号の領域外に訂正されていないエラーがあることを示す。このフラグは CRC ウィンドウ以外の情報ビットが保護されている場合に参照されるだけである。

この 2 つのフラグは音声パラメータ及び FII フラグと共に、音声復号器で使えるように生成されなければならない。受信された SID フレームの場合は BFI フラグのみが決定される。

3.4 音声パラメータのフォーマット

音声パラメータは TTC 標準 JT-G 7 2 3. 1 の Table 4 - 1 / JT-G723.1 及び Table 4 - 2 / JT-G723.1 のフォーマットで、最後にエラーフラグを付加して、チャンネル復号器から音声復号器に伝送される。このフォーマットは、高レートモードは Table C - 3 - 1 a / JT-G723.1 に、低レートモードは Table C - 3 - 1 b / JT-G723.1 に、SID フレーム用は Table C - 3 - 1 c / JT-G723.1 にそれぞれ示される。

Table C – 3 – 1 a / JT-G723.1: Decoder bitstream packing for high rate mode
(ITU-TG.723.1)

| Speech Decoder Octets | PARx_By, |
|--------------------------|--|
| 1 | LPC_B5...LPC_B0, VAD, RATE |
| 2 | LPC_B13...LPC_B6 |
| 3 | LPC_B21...LPC_B14 |
| 4 | ACL0_B5...ACL0_B0, LPC_B23, LPC_B22 |
| 5 | ACL2_B4...ACL2_B0, ACL1_B1, ACL1_B0, ACL0_B6 |
| 6 | GAIN0_B3...GAIN0_B0, ACL3_B1, ACL3_B0, ACL2_B6, ACL2_B5 |
| 7 | GAIN0_B11...GAIN0_B4 |
| 8 | GAIN1_B7...GAIN1_B0 |
| 9 | GAIN2_B3...GAIN2_B0, GAIN1_B11...GAIN1_B8 |
| 10 | GAIN2_B11...GAIN2_B4 |
| 11 | GAIN3_B7...GAIN3_B0 |
| 12 | GRID3_B0, GRID2_B0, GRID1_B0, GRID0_B0, GAIN3_B11...GAIN3_B8 |
| 13 | MSBPOS_B6...MSBPOS_B0, UB |
| 14 | POS0_B1, POS0_B0, MSBPOS_B12...MSBPOS_B7 |
| 15 | POS0_B9...POS0_B2 |
| 16 | POS1_B2, POS1_B0, POS0_B15...POS0_B10 |
| 17 | POS1_B10...POS1_B3 |
| 18 | POS2_B3...POS2_B0, POS1_B13...POS1_B11 |
| 19 | POS2_B11...POS2_B4 |
| 20 | POS3_B3...POS3_B0, POS2_B15...POS2_B12 |
| 21 | POS3_B11...POS3_B4 |
| 22 | PSIG0_B5...PSIG0_B0, POS3_B13, POS3_B12 |
| 23 | PSIG2_B2...PSIG2_B0, PSIG1_B4...PSIG1_B0 |
| 24 | PSIG3_B4...PSIG3_B0, PSIG2_B5...PSIG2_B3 |
| 25 | UB , UB , UB , UB , UB , FII , EFI , BFI |

Table C – 3 – 1 b / JT-G723.1: Decoder bitstream packing for low rate mode
(ITU-T G.723.1)

| Speech Decoder Octets | PARx_By, |
|-----------------------|--|
| 1 | LPC_B5...LPC_B0, VAD, RATE |
| 2 | LPC_B13...LPC_B6 |
| 3 | LPC_B21...LPC_B14 |
| 4 | ACL0_B5...ACL0_B0, LPC_B23, LPC_B22 |
| 5 | ACL2_B4...ACL2_B0, ACL1_B1, ACL1_B0, ACL0_B6 |
| 6 | GAIN0_B3...GAIN0_B0, ACL3_B1, ACL3_B0, ACL2_B6, ACL2_B5 |
| 7 | GAIN0_B11...GAIN0_B4 |
| 8 | GAIN1_B7...GAIN1_B0 |
| 9 | GAIN2_B3...GAIN2_B0, GAIN1_B11...GAIN1_B8 |
| 10 | GAIN2_B11...GAIN2_B4 |
| 11 | GAIN3_B7...GAIN3_B0 |
| 12 | GRID3_B0, GRID2_B0, GRID1_B0, GRID0_B0, GAIN3_B11...GAIN3_B8 |
| 13 | POS0_B7...POS0_B0 |
| 14 | POS1_B3...POS1_B0, POS0_B11...POS0_B8 |
| 15 | POS1_B11...POS1_B4 |
| 16 | POS2_B7...POS2_B0 |
| 17 | POS3_B3...POS3_B0, POS2_B11...POS2_B8 |
| 18 | POS3_B11...POS3_B4 |
| 19 | PSIG1_B3...PSIG1_B0, PSIG0_B3...PSIG0_B0 |
| 20 | PSIG3_B3...PSIG3_B0, PSIG2_B3...PSIG2_B0 |
| 21 | UB , UB , UB , UB , UB , FII , EFI , BFI |

Table C – 3 – 1 c / JT-G723.1: Decoder bitstream packing for SID frames
(ITU-T G.723.1)

| Speech Decoder Octets | PARx_By, |
|-----------------------|--|
| 1 | LPC_B5...LPC_B0, VAD, RATE |
| 2 | LPC_B13...LPC_B6 |
| 3 | LPC_B21...LPC_B14 |
| 4 | GAIN_B5 ... GAIN_B0, LPC_B23, LPC_B22 |
| 5 | UB , BFI |

3.5 誤り補償

チャンネル復号器及び利得圧縮処理で生成されたフラグは、音声復号器では誤り補償の目的で用いられる。

3.6 誤り解析フラグの能力

本節では、バッドフレーム検出率(PD_{BFI})と誤りフレーム検出率(PD_{EFI})の形で誤り解析に求められる最低限の能力について規定する。Table C-3-2/JT-G723.1の内容がすべての伝送チャンネルに求められる。

Table C-3-2/JT-G723.1: Detection performance of BFI and EFI flags
(ITU-T G.723.1)

| | G.723.1-6. 3 | G.723.1-5.3 | G.723.1-SID |
|------------|-----------------|-------------|-------------|
| PD_{BFI} | 0.99 | 0.99 | 0.99 |
| PD_{EFI} | 0.65 | 0.65 | --- |

4. 固定小数点演算 C ソースコード

チャンネル符号器のすべての詳細は、固定小数点演算 ANSI-C ソースコードの形式で記述されている。上記の記述と C ソースコードとの間に何らかの矛盾があった場合は、リファレンスとしては C ソースコードが採用されるべきである。この C ソースコードは T T C 標準 J T - G 7 2 3 . 1 として日本 I T U 協会から入手できるコードの一部である。組み上げたチャンネル符号器のビットの正確さを C ソースコードと比較するためのテストベクトルが含まれている。チャンネル符号器の完全な C ソースコードがファイル `chan_enc.c` に含まれている。チャンネル復号器の C ソースコードは、ファイル `chan_dec.c` に含まれているが、実現例として提供されているだけである。

付録 1 用語対照表

(標準 J T - G 7 2 3 . 1 に対する)

(本付録は参考資料であり、仕様ではない。)

| 英語 | T T C 標準用語 |
|---------------------------------------|-------------------------|
| adaptive postfilter | 適応ポストフィルタ |
| algebraic codebook | 代数的コードブック |
| algebraic codebook excitation (ACELP) | 代数的コードブック励振 (A C E L P) |
| ACELP(algebraic codebook excitation) | A C E L P (代数的符号励振) |
| Analysis-by-Synthesis | 合成による分析 (A-b-S) 法 |
| autocorrelation coefficient | 自己相関係数 |
| backward energy | 後方エネルギー |
| bandwidth expansion | 帯域幅の拡張 |
| binomial window coefficients | 二項窓係数 |
| Bitstream | ビット列 |
| bit stream packing | ビット列のパッキング |
| closed loop pitch lag | 閉ループピッチラグ |
| combinatorial codes | 組み合わせ符号 |
| combined filter | 結合フィルタ |
| Dirac function | ディラック関数 |
| Entry | エントリ |
| error criterion | 誤差評価値 |
| excitation decoder | 励振復号器 |
| fixed codebook excitation | 固定コードブック励振 |
| focused search approach | 絞り込み探索アプローチ |
| formant perceptual weighting filter | ホルマント聴覚重み付けフィルタ |
| formant postfilter | ホルマントポストフィルタ |
| forward energy | 前方エネルギー |
| frame erasure | フレーム消失 |
| frame erasure concealment | フレーム消失誤り補償 |
| frame interpolation | フレーム補間 |
| framer | フレーム化 |
| gain scaling unit | 利得スケーリング部 |
| good quality | 高品質 |
| greater quality | より高品質 |
| grid bit | グリッドビット |
| harmonic noise shaping filter | 高調波ノイズシェーピングフィルタ |
| impulse response | インパルス応答 |
| lag | ラグ |
| Levinson-Durbin recursion | レビンソン・ダービン再帰法 |
| linear PCM | 均一 P C M |
| linear predictive (LP) analysis | 線形予測 (L P) 分析 |

| 英語 | T T C 標準用語 |
|---|--------------------------|
| long term | 長期 |
| look ahead | 先読み |
| lower triangular Toeplitz convolution matrix | 下三角テプリッツ畳込み行列 |
| LPC analysis | L P C 分析 |
| LP coefficients | L P 係数 |
| LSP coefficients | L S P 係数 |
| LSP interpolator | L S P 補間器 |
| LSP quantizer | L S P 量子化器 |
| mean square | 自乗平均 |
| mean square error | 平均自乗誤差 |
| multi-pulse maximum likelihood quantization (MP-MLQ) excitation | 最尤量子化型マルチパルス (MP-MLQ) 励振 |
| MP-MLQ (multi-pulse maximum likelihood quantization) excitation | MP-MLQ (最尤量子化型マルチパルス) 励振 |
| non-zero pulses | 非零パルス |
| open loop pitch period estimate | 開ループピッチ周期推定 |
| perceptual weighting filter | 聴覚重み付けフィルタ |
| perceptually weighted speech | 聴覚重み付き音声 |
| pitch contribution | ピッチ寄与分 |
| pitch estimation | ピッチ推定 |
| pitch lag | ピッチラグ |
| pitch postfilter | ピッチポストフィルタ |
| pitch predictor | ピッチ予測器 |
| pulse contribution | パルス寄与分 |
| Predictive Split Vector Quantizer (PSVQ) | 予測分割ベクトル量子化器 (P S V Q) |
| PSVQ (Predictive Split Vector Quantizer) | P S V Q (予測分割ベクトル量子化器) |
| residual interpolation | 残差の補間 |
| residual LSP error vector | 残差 LSP 誤差ベクトル |
| residual signal | 残差信号 |
| ringing subtraction | リングング補正 |
| sampling rate | 標準化周波数 |
| short term | 短期 |
| signs | 極性 |
| split residual LSP code-book | 分割誤差 L S P コードブック |
| stability check | 安定度チェック |
| visual telephony | テレビ電話 |
| voiced/unvoiced classifier | 有声/無声識別器 |
| weighted speech | 重み付き音声 |
| weights | 重み係数 |
| white noise correlation factor | 白色雑音補正係数 |
| zero input response | 零入力応答 |

| 英語 | TTC標準用語 |
|--|---|
| 標準JT-G723.1 付属資料A | |
| Active background noise comfort noise Comfort Noise Generator(CNG) frame erasure Hangover Inactive Itakura distance noise spikes pseudo-random sequence pseudo-white excitation random generator Silence Insertion Descriptor(SID) speech clipping Voice Activity Detection(VAD) | 有音 背景雑音 擬似背景雑音 擬似背景雑音発生器 フレーム消失 ハングオーバ 無音 板倉距離 スパイク雑音 擬似ランダム系列 擬似白色励振 ランダム発生器 無音挿入記述子 音声欠落 音声検出 |
| 標準JT-G723.1 付属資料C | |
| error concealment error detection code forbidden index indication Interleaving mother code Puncture punctured convolutional codes | 誤り補償 誤り検出符号 禁止インデックス表示 インタリーブ マザーコード パンクチャ パンクチャド畳み込み符号 |

付録2 用語解説

(標準 J T - G 7 2 3 . 1 に対する)

(本付録は参考資料であり、仕様ではない。)

代数的コードブック (algebraic codebook)

励振信号を単位パルスの位置と極性のみで表現したコードブック。本標準では 60 次元のベクトルに 4 本のパルスを置き、残りの要素は零で埋められている。

代数的コードブック励振 (algebraic codebook excitation)

代数的コードブックを雑音コードブックとして、C E L P 方式の符号化方式の励振信号を生成することにより、音質の改善とともにメモリサイズの低減、コードブック探索の効率化を実現した方式。

レビンソン・ダービン再帰法 (Levinson-Durbin recursion)

対象とする信号の相関係数から再帰的に線形予測係数と反射係数を求める効率的な方法。

線スペクトル対 (line spectral pair)

偏自己相関分析の声門における整合終端の条件を完全反射として得られる係数。量子化特性および補間特性にすぐれており、効率的な符号化ができる。

最尤量子化型マルチパルス(MP-MLQ)励振 (Multi-Pulse Maximum Likelihood Quantization (MP-MLQ) excitation)

C E L P 符号化方式の励振信号を、複数の単位パルスを置いたベクトルで近似する方式。本標準では 60 次元のベクトルに 5 ないし 6 本のパルスを置く。ターゲットベクトルは、合成による分析(A-b-S)法に基づき、利得ファクタ、単位パルスの極性と位置を推定して近似される。

聴覚重み付けフィルタ (perceptual weighting filter)

音声信号のホルマント領域での量子化雑音は、マスキング効果により小さく感じられる。この効果を利用して、聴感上の雑音感を低減するためのフィルタ。

ポストフィルタ (postfilter)

復号された音声信号の朝刊上の品質を向上させるために、復号器の出力側に設けられたフィルタ。

予測分割ベクトル量子化器(PSVQ) (Predictive Split Vector Quantization)

パラメータベクトルの時系列に対して予測を行なったのち、その残差ベクトルを複数のサブベクトルに分割し、それぞれのサブベクトルに応じたコードブックを用いてベクトル量子化する量子化器。本標準では L S P 係数ベクトルの量子化にこの手法が適用される。

(付属資料A, B, Cに対する)

擬似背景雑音 (comfort noise)

音声コーデックの不連続伝送による背景雑音の不連続性を避けるために、間欠的に送信されてきた一部の符号化パラメータから復号器側で発生させる擬似的な背景雑音である。

畳み込み符号 (convolutional codes)

符号化処理が逐次的な畳み込み演算によって生成される符号。これに対しブロック符号と呼ぶ構成は、一定の長さの系列ごとに符号化され、他のブロックとは独立であるが、畳み込み符号では、先行するブロックの情報にも依存する。

誤り検出符号 (error detection code)

デジタル情報の伝送の際に通信路で生じた誤りを検出することを目的として、通報に適切な冗長性を付加して構成される符号である。

パンクチャド畳み込み符号 (punctured convolutional codes)

符号化率(1/n)を高めるために、畳み込み符号器の出力から一定のビット数を抜き取った畳み込み符号である。

音声検出 (Voice Activity Detection)

音声コーデックの不連続伝送を行うために、連続した入力信号から音声区間のみを検出するものである。