

JT-G729

8kbit/s CS-ACELPを用いた 音声符号化方式

Coding of Speech at 8kbit/s using
Conjugate-Structure Algebraic-Code-Excited
Linear-Prediction (CS-ACELP)

第 6.1 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>.....	8
1. 本標準の規定範囲.....	10
2. コーデックの概要.....	10
2. 1 符号器.....	11
2. 2 復号器.....	12
2. 3 遅延.....	13
2. 4 音声コーデックの記述.....	13
2. 5 記法.....	13
3. 符号器の機能記述.....	18
3. 1 前処理.....	18
3. 2 線形予測分析と量子化.....	18
3. 2. 1 窓かけと自己相関計算.....	20
3. 2. 2 レビンソン・ダービナルゴリズム.....	21
3. 2. 3 LPからLSPへの変換.....	21
3. 2. 4 LSP係数の量子化.....	22
3. 2. 5 LSP係数の補間.....	24
3. 2. 6 LSPからLPへの変換.....	24
3. 3 聴覚重み付け.....	25
3. 4 開ループピッチ分析.....	26
3. 5 インパルス応答の計算.....	27
3. 6 ターゲット信号の計算.....	27
3. 7 適応コードブック探索.....	28
3. 7. 1 適応コードブックベクトルの生成.....	29
3. 7. 2 適応コードブック遅延に対する符号語の計算.....	30
3. 7. 3 適応コードブック利得の計算.....	30
3. 8 固定コードブックの構造と探索.....	30
3. 8. 1 固定コードブック探索手順.....	31
3. 8. 2 固定コードブック符号語の計算.....	33
3. 9 利得の量子化.....	33
3. 9. 1 利得予測.....	34
3. 9. 2 利得量子化のためのコードブック探索.....	35
3. 9. 3 利得量子化器のための符号語の計算.....	35
3. 10 メモリの更新.....	35
4. 復号器の機能記述.....	37
4. 1 パラメータ復号処理.....	39
4. 1. 1 LPフィルタパラメータの復号.....	39
4. 1. 2 パリティビットの計算.....	39
4. 1. 3 適応コードブックベクトルの復号.....	39
4. 1. 4 固定コードブックベクトルの復号.....	40
4. 1. 5 適応および固定コードブック利得の復号.....	40
4. 1. 6 再生音声信号の計算.....	40
4. 2 後処理.....	40
4. 2. 1 長期ポストフィルタ.....	41
4. 2. 2 短期ポストフィルタ.....	42
4. 2. 3 傾き補償.....	42
4. 2. 4 適応利得制御.....	42
4. 2. 5 高域通過フィルタとアップスケーリング.....	43
4. 3 符号器と復号器の初期化.....	43
4. 4 フレーム消失補償.....	43
4. 4. 1 合成フィルタパラメータの反復.....	44
4. 4. 2 適応および固定コードブック利得の減衰.....	44
4. 4. 3 利得予測器のメモリの減衰.....	44
4. 4. 4 置換用励振信号の生成.....	44
5. CS-ACELPコーデックのビットイグザクテナ記述.....	45
5. 1 シミュレーションソフトウェアの使い方.....	45
5. 2 シミュレーションソフトウェアの構成.....	45

付属資料A 低演算量版 8kbit/s CS-ACELP 音声コーデック	
A. 1 本付属資料の規定範囲	50
A. 2 コーデックの概要説明	50
A. 2. 1 音声コーデックの記述	50
A. 2. 2 記法	50
A. 3 符号器の機能記述	51
A. 3. 1 前処理	51
A. 3. 2 線形予測分析と量子化	51
A. 3. 2. 1 窓かけと自己相関計算	51
A. 3. 2. 2 レビンソン・ダービンアルゴリズム	51
A. 3. 2. 3 LPからLSPへの変換	51
A. 3. 2. 4 LSP係数の量子化	51
A. 3. 2. 5 LSP係数の補間	51
A. 3. 2. 6 LSPからLPへの変換	51
A. 3. 3 聴覚重み付け	51
A. 3. 4 開ループピッチ分析	52
A. 3. 5 インパルス応答の計算	52
A. 3. 6 ターゲット信号の計算	53
A. 3. 7 適応コードブック探索	53
A. 3. 7. 1 適応コードブックベクトルの生成	54
A. 3. 7. 2 適応コードブック遅延に対する符号語の計算	54
A. 3. 7. 3 適応コードブック利得の計算	54
A. 3. 8 固定コードブックの構造と探索	54
A. 3. 8. 1 固定コードブック探索手順	54
A. 3. 8. 2 固定コードブック符号語の計算	54
A. 3. 9 利得の量子化	54
A. 3. 10 メモリの更新	54
A. 4 復号器の機能記述	55
A. 4. 1 パラメータ復号処理	55
A. 4. 2 後処理	55
A. 4. 2. 1 長期ポストフィルタ	55
A. 4. 2. 2 短期ポストフィルタ	55
A. 4. 2. 3 傾き補償	56
A. 4. 2. 4 適応利得制御	56
A. 4. 2. 5 高域通過フィルタとアップスケーリング	56
A. 4. 3 符号器と復号器の初期化	56
A. 4. 4 フレーム消失補償	56
A. 5 低演算量版CS-ACELPコーデックのビットイグザクトな記述	56
A. 5. 1 シミュレーションソフトウェアの使い方	56
A. 5. 2 シミュレーションソフトウェアの構成	57

付属資料B ITU-T勧告V. 70端末に適した標準JT-G729に対する無音圧縮手法	
B. 1 本付属資料の規定範囲	58
B. 2 VAD/DTX/CNGアルゴリズムの概要	58
B. 3 VADアルゴリズムの詳細記述	59
B. 3. 1 パラメータの抽出	60
B. 3. 1. 1 線スペクトル周波数(LSF)	61
B. 3. 1. 2 全帯域エネルギー	61
B. 3. 1. 3 低域エネルギー	61
B. 3. 1. 4 零交差数	61
B. 3. 2 背景雑音特性の移動平均の初期化	61
B. 3. 3 長時間の最小エネルギーの生成	63
B. 3. 4 差分パラメータの生成	63
B. 3. 4. 1 スペクトル歪 ΔS	63
B. 3. 4. 2 全帯域エネルギーの差分 ΔE_f	63
B. 3. 4. 3 低域エネルギー差分 ΔE_l	63
B. 3. 4. 4 零交差数の差分 ΔZC	63
B. 3. 5 多境界初期有音判定	63
B. 3. 6 有音判定の平滑化	64
B. 3. 7 背景雑音特性の移動平均の更新	65

B. 4	DTX/CNGアルゴリズムの詳細記述	65
B. 4. 1	DTXアルゴリズムの記述	66
B. 4. 1. 1	フレーム自己相関関数の保持	66
B. 4. 1. 2	現在のフレームタイプの計算	66
B. 4. 1. 3	LPCフィルタの比較	67
B. 4. 1. 4	エネルギーの比較	67
B. 4. 2	SIDの評定と量子化	68
B. 4. 2. 1	エネルギーの量子化	68
B. 4. 2. 2	SID LPCフィルタの推定と量子化	68
B. 4. 3	SIDビット列の記述	69
B. 4. 4	無音用符号器と無音用復号器 (CNG) の記述	69
B. 4. 5	CNGに関するフレーム消失補償	71
B. 5	無音圧縮手法のビットイグザクトな記述	72
B. 5. 1	シミュレーションソフトウェアの構成	72

付属資料C 浮動小数点演算版 8KBIT/SCS-ACELP 音声コーデック

C. 1	本付属資料の規定範囲	74
C. 2	参照すべき標準	74
C. 3	概要	74
C. 4	アルゴリズムの詳細	74
C. 5	ANSI Cコード	74

付属資料C+ 標準 JT-G 7 2 9 SCS-ACELP 音声符号化方式本体、

付属資料B、D、およびEの統合における浮動小数点演算での実現方法

C+. 1	本付属資料の規定範囲	77
C+. 2	参照すべき標準	77
C+. 3	概要	77
C+. 4	新しい機能	78
C+. 4. 1	付属資料Dにおける付属資料BのDTX動作	78
C+. 4. 2	付属資料Eにおける付属資料BのDTX動作	78
C+. 5	アルゴリズムの記述	78
C+. 5. 1	音楽検出	78
C+. 5. 1. 1	関連パラメータの計算	79
C+. 5. 1. 2	分類	82
C+. 5. 2	不連続伝送時の付属資料Dに特有な状態変数の更新	82
C+. 5. 3	不連続伝送時の付属資料Eに特有な状態変数の更新	83
C+. 5. 3. 1	付属資料Eに特有な符号器の状態変数の更新	83
C+. 5. 3. 2	付属資料Eに特有な復号器の状態変数の更新	83
C+. 6	Cソースコードの記述	83
C+. 6. 1	シミュレーションソフトウェアの使い方	83
C+. 6. 2	シミュレーションソフトウェアの構成	84

付属資料D 6.4KBIT/SCS-ACELP 音声コーデック

D. 1	本付属資料の規定範囲	87
D. 2	参照すべき標準	87
D. 3	6.4kbit/s拡張に対する符号器の概要記述	87
D. 4	ビット割り当て	88
D. 5	符号器の機能記述	88
D. 5. 1	前処理	88
D. 5. 2	線形予測分析と量子化	88
D. 5. 3	聴覚重み付け	88
D. 5. 4	開ループピッチ分析	88
D. 5. 5	インパルス応答の計算	88
D. 5. 6	ターゲット信号の計算	88
D. 5. 7	適応コードブック探索	88
D. 5. 8	固定コードブックの構造と探索	89
D. 5. 9	利得の量子化	89
D. 5. 10	メモリの更新	89

D. 6	復号器の機能記述	89
D. 6. 1	パラメータ復号処理	89
D. 6. 2	固定コードブックの後処理	89
D. 6. 3	ポストフィルタと後処理	89
D. 6. 4	フレーム消失補償	90
D. 7	ANSI Cコード	90
付属資料E 11.8KBIT/SCS-ACELP音声コーデック		
E. 1	本付属資料の規定範囲	92
E. 2	コーデックの概要	92
E. 2. 1	符号器	93
E. 2. 2	復号器	94
E. 2. 3	遅延	94
E. 2. 4	音声コーデックの記述	94
E. 3	符号器の機能記述	94
E. 3. 1	前処理	94
E. 3. 2	線形予測分析と量子化	94
E. 3. 2. 1	窓かけと自己相関計算	95
E. 3. 2. 2	レビンソン・ダービンアルゴリズム	95
E. 3. 2. 3	LPからLSPへの変換	95
E. 3. 2. 4	LSP係数の量子化	95
E. 3. 2. 5	LSP係数の補間	95
E. 3. 2. 6	LSPからLPへの変換	96
E. 3. 2. 7	バックワード/フォワード判定および切替え手順	96
E. 3. 2. 7. 1	切替え手順	96
E. 3. 2. 7. 2	グローバル定常性インジケータと高定常性インジケータ	96
E. 3. 2. 7. 3	バックワード/フォワードの判定手順	99
E. 3. 3	聴覚重み付け	99
E. 3. 4	開ループピッチ分析	100
E. 3. 5	インパルス応答の計算	100
E. 3. 6	ターゲット信号の計算	100
E. 3. 7	適応コードブック探索	100
E. 3. 8	固定コードブックの構造と探索	101
E. 3. 8. 1	フォワードLPモードにおける固定コードブック	101
E. 3. 8. 2	バックワードLPモードにおける固定コードブック	103
E. 3. 9	利得の量子化	104
E. 3. 10	メモリの更新	104
E. 4	復号器の機能記述	105
E. 4. 1	パラメータ復号手順	107
E. 4. 1. 1	フォワード/バックワード復号手順	107
E. 4. 1. 2	LPフィルタパラメータの復号	108
E. 4. 1. 2. 1	バックワードLPフィルタの計算	108
E. 4. 1. 2. 2	フォワードモード	108
E. 4. 1. 2. 3	バックワードモード	108
E. 4. 1. 3	適応コードブック遅延のパリティビットの計算	108
E. 4. 1. 4	適応コードブックベクトルの復号	108
E. 4. 1. 5	固定コードブックベクトルの復号	108
E. 4. 1. 6	適応および固定コードブック利得の復号	108
E. 4. 1. 7	再生音声信号の計算	109
E. 4. 2	後処理	109
E. 4. 2. 1	長期ポストフィルタ	109
E. 4. 2. 2	短期ポストフィルタ	109
E. 4. 2. 3	傾き補償	110
E. 4. 2. 4	適応利得制御	110
E. 4. 2. 5	高域通過フィルタとアップスケーリング	110
E. 4. 3	符号器と復号器の初期化	110
E. 4. 4	フレーム消失補償	111
E. 4. 4. 1	LPモードの反復	111
E. 4. 4. 2	合成フィルタパラメータの計算	111
E. 4. 4. 3	適応および固定コードブック利得の減衰	112
E. 4. 4. 4	利得予測器のメモリの減衰	113

E. 4. 4. 5	置換用励振信号の生成	113
E. 5	CS-ACELP符号器のビットイグザクナ記述	113
E. 5. 1	シミュレーションソフトウェアの使い方	114
E. 5. 2	シミュレーションソフトウェアの構成	114
E. 6	参照すべき標準	116
付属資料F 標準J T-G 7 2 9付属資料Dにおける付属資料BのDTX機能の実現方法		
F. 1	本付属資料の規定範囲	117
F. 2	参照すべき標準	117
F. 3	概要	117
F. 4	新しい機能	117
F. 4. 1	付属資料Dにおける付属資料BのDTX動作	117
F. 5	アルゴリズムの記述	118
F. 5. 1	不連続伝送時の付属資料Dに特有な状態変数の更新	118
F. 6	Cソースコードの記述	118
F. 6. 1	シミュレーションソフトウェアの使い方	118
F. 6. 2	シミュレーションソフトウェアの構成	118
付属資料G 標準J T-G 7 2 9付属資料Eに対する付属資料BのDTX機能の実現方法		
G. 1	本付属資料の規定範囲	122
G. 2	参照すべき標準	122
G. 3	概要	122
G. 4	新しい機能	122
G. 4. 1	付属資料Eにおける付属資料BのDTX動作	122
G. 5	アルゴリズムの記述	123
G. 5. 1	音楽検出	123
G. 5. 1. 1	関連パラメータの計算	123
G. 5. 1. 2	分類	127
G. 5. 2	不連続伝送時における付属資料Eに特有な状態変数の更新	127
G. 5. 2. 1	付属資料Eに特有な符号器の状態変数の更新	127
G. 5. 2. 2	付属資料Eに特有な復号器の状態変数の更新	127
G. 6	Cソースコードの記述	127
G. 6. 1	シミュレーションソフトウェアの使い方	128
G. 6. 2	シミュレーションソフトウェアの構成	128
付属資料H 標準J T-G 7 2 9付属資料DおよびEの切替手順に対する実現方法		
H. 1	本付属資料の規定範囲	131
H. 2	参照すべき標準	131
H. 3	概要	131
H. 4	アルゴリズムの記述	131
H. 4. 1	付属資料Eフレームにおける付属資料Dに特有な状態変数の更新	131
H. 4. 2	付属資料Dフレームにおける付属資料Eに特有な状態変数の更新	132
H. 4. 2. 1	付属資料Eに特有な符号器の状態変数の更新	132
H. 4. 2. 2	付属資料Eに特有な復号器の状態変数の更新	132
H. 5	Cソースコードの記述	132
H. 5. 1	シミュレーションソフトウェアの使い方	132
H. 5. 2	シミュレーションソフトウェアの構成	132
付属資料I 標準J T-G 7 2 9本体、付属資料B、DおよびEのCコードの統合における実現方法		
I. 1	本付属資料の規定範囲	136
I. 2	参照すべき標準	136
I. 3	概要	136
I. 4	新しい機能	136
I. 4. 1	付属資料Dにおける付属資料BのDTX動作	137
I. 4. 2	付属資料Eにおける付属資料BのDTX動作	137
I. 5	アルゴリズムの記述	137
I. 5. 1	音楽検出	137
I. 5. 1. 1	関連パラメータの計算	138

I. 5. 1. 2	分類	141
I. 5. 2	不連続伝送時の付属資料Dに特有な状態変数の更新	141
I. 5. 3	不連続伝送時の付属資料Eに特有な状態変数の更新	141
I. 5. 3. 1	付属資料Eに特有な符号器の状態変数の更新	141
I. 5. 3. 2	付属資料Eに特有な復号器の状態変数の更新	142
I. 6	Cソースコードの記述	142
I. 6. 1	シミュレーションソフトウェアの使い方	142
I. 6. 2	シミュレーションソフトウェアの構成	142

付録1 外部VAD/DTX/CNGを用いたシステムにおける標準JT-G729コーデックに対する外部同期リセット性能

概要	147	
本付録の規定範囲	147	
1. 1	はじめに	147
1. 2	実験計画	148
1. 3	試験結果	148
1. 4	結論	148
1. 5	参考文献	148

付録2 VoIPにおけるTTC標準JT-G729付属資料Bの機能拡張—オプション1

概要	150	
2. 1	本付録の規定範囲	150
2. 2	参照すべき標準	150
2. 3	略語	150
2. 4	概要	151
2. 5	VoIPにおけるJT-G729付属資料Bの問題	151
2. 6	実験方法	152
2. 6. 1	平滑化の4段階目	152
2. 6. 2	ヒステリシス	153
2. 6. 3	更新判定	153
2. 7	電子ファイル	154

付録3 VoIPにおけるTTC標準JT-G729付属資料Bの機能拡張—オプション2

概要	155	
3. 1	本付録の規定範囲	155
3. 2	報告されているTTC標準JT-G729付属資料Bの問題点に対する解決手法	155
3. 3	報告されているTTC標準JT-G729付属資料Bの問題点に対する解決例	155
3. 4	電子ファイル	158

付録 用語対照表	159
----------	-----

付録 用語解説	164
---------	-----

＜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G.729に準拠したものである。本標準の付属資料Aは1996年11月に承認されたITU-T勧告G.729 AnnexAに準拠したものである。本標準の付属資料Bは1996年10月の世界電気通信標準化会議(WTSC-96)において承認されたITU-T勧告G.729 AnnexBに準拠したものである。本標準の付属資料C、D、Eは1998年9月に承認されたITU-T勧告G.729 AnnexC、D、Eに準拠したものである。本標準の付属資料C+、F、G、H、Iは2000年2月に承認されたITU-T勧告G.729 AnnexC+、F、G、H、Iに準拠したものである。本標準の付属資料Bは2000年11月にITU-T SG16でAAPに進むことが合意され、2001年3月に承認されたCorrigenda to G.729 Annexesに準拠したものである。本標準の付録1は2001年6月に承認されたITU-T勧告G.729 Appendix Iに準拠したものである。本標準の付録2は2005年8月に承認されたITU-T勧告G.729 Appendix IIに準拠したものである。本標準の付録3は2005年8月に承認されたITU-T勧告G.729 Appendix IIIに準拠したものである。

本標準は、2002年10月および2006年4月に承認されたITU-T勧告G.729に対するインプリメンターズガイドに準拠して改定されたものである。

3. 上記国際勧告に対する追加項目等

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

上記国際勧告等との章立ての構成の相違はない。

4. 改版の履歴

版数	制定日	改版内容
第1版	平成8年 11月27日	制定
第2版	平成9年 4月23日	付属資料A、Bの追加
第3版	平成11年 4月22日	付属資料C、D、Eの追加
第4版	平成12年 11月30日	付属資料C+、F、G、H、Iの追加
第4.1版	平成13年 11月27日	付属資料Bの修正
第5版	平成14年 5月30日	付録1の追加
第5.1版	平成15年 11月27日	Cコード改定に関する記述の追加
第6版	平成18年 3月16日	付録2、3の追加
第6.1版	平成18年 11月27日	Cコード改定に関する記述の追加

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-G726、JT-G728

ITU-T勧告 : V. 70

- (2) 重み付き音声信号と誤差信号に関して、標準JT-G729本体では sw, ew として扱っているが、標準JT-G729付属資料Aでは s_w, e_w として扱っている。
- (3) 傾き補償の計算に使用する係数に関して、標準JT-G729本体および標準JT-G729付属資料Aでは1次反射係数と記述しているが、標準JT-G729付属資料Eでは1次偏自己相関係数と記述している。
- (4) JT-G729付属資料C、D、Eに対するプログラム著作権
有益な注記：本勧告の使用にあたって、ITU著作権方針の条項に従って著作権保持者からの許可が必要となる。なお、ITU著作権方針の条項は現在準備中であり、すでに承認されている勧告にも著作権問題に関連する勧告の改版が必要となる。
- (5) TTC標準JT-G729は、ITU-T勧告G. 729に準拠しており、本標準中で言及しているANSI Cコードとは、ITU-T勧告G. 729のものをさし、日本ITU協会から入手可能である。
- (6) 本標準の付属資料B、F、G、IおよびC+にて参照しているCコードの改定に関して、2002年10月および2006年4月に承認されたITU-T勧告G. 729に対するインプリメンターズガイドに記述がある。

1. 本標準の規定範囲

本標準では、共役構造一代数的符号励振線形予測（CS-ACELP）を使用して、音声信号を 8kbit/s で符号化するためのアルゴリズムについて記述する。

本アルゴリズムは、電話帯域フィルタ（TTC標準 JT-G 7 1 2 参照）により帯域制限されたアナログ入力信号を 8000Hz で標本化し、次に 16 ビット均一PCMに変換することによって得られたデジタル信号を、符号器の入力として動作するように設計されている。復号器の出力は同様な方法によってアナログ信号に戻す必要がある。TTC標準 JT-G 7 1 1 で規定される 64kbit/s PCMデータのような他の入出力形式の信号は、符号器の前で 16 ビット均一PCMへ、また復号器の後で 16 ビット均一PCMから適当な形式へ変換する必要がある。符号器から復号器へ渡されるビット列は本標準の中で規定される。

本標準は以下の章構成から成る。第 2 章では CS-ACELP アルゴリズムの概要について、第 3 章と第 4 章では CS-ACELP の符号器と復号器の動作原理についてそれぞれ述べる。第 5 章では、本アルゴリズムを 16 ビット固定小数点演算を用いて記述したシミュレーションソフトウェアについて説明する。

2. コーデックの概要

CS-ACELP は符号励振線形予測（CELP）符号化モデルに基づいたコーデックで、10ms の音声フレーム（8000 サンプル/秒の標本化周波数では 80 サンプルに相当）に対して動作する。10ms のフレーム毎に音声信号を分析し、CELP モデルで使用するパラメータ（線形予測フィルタ係数、適応コードブックと固定コードブックのインデックス、利得）を抽出する。これらのパラメータは符号化された後送信される。本コーデックで使用するパラメータのビット配分を Table 2-1/JT-G729 に示す。復号器においては、これらのパラメータは励振信号や合成フィルタのパラメータを再構築するために使用される。音声は、Figure 2-1/JT-G729 に示されているように、短期合成フィルタにこの励振信号を通すことにより再生される。短期合成フィルタは 10 次の線形予測（LP）フィルタを基に構成され、長期すなわちピッチ合成フィルタは、いわゆる適応コードブックを用いて実現される。再生音声の音質は、ポストフィルタを通すことによって改善される。

Table 2-1/JT-G729 Bit allocation of the 8 kbit/s CS-ACELP algorithm (10 ms frame).
(ITU-T G.729)

<i>Parameter</i>	<i>Codeword</i>	<i>Subframe 1</i>	<i>Subframe 2</i>	<i>Total per frame</i>
Line Spectrum Pairs	<i>L0, L1, L2, L3</i>			1 8
Adaptive-codebook delay	<i>P1, P2</i>	8	5	1 3
Pitch-delay parity	<i>P0</i>	1		1
Fixed-codebook index	<i>C1, C2</i>	1 3	1 3	2 6
Fixed-codebook sign	<i>S1, S2</i>	4	4	8
Codebook gains (stage 1)	<i>GA1, GA2</i>	3	3	6
Codebook gains (stage 2)	<i>GB1, GB2</i>	4	4	8
Total				8 0

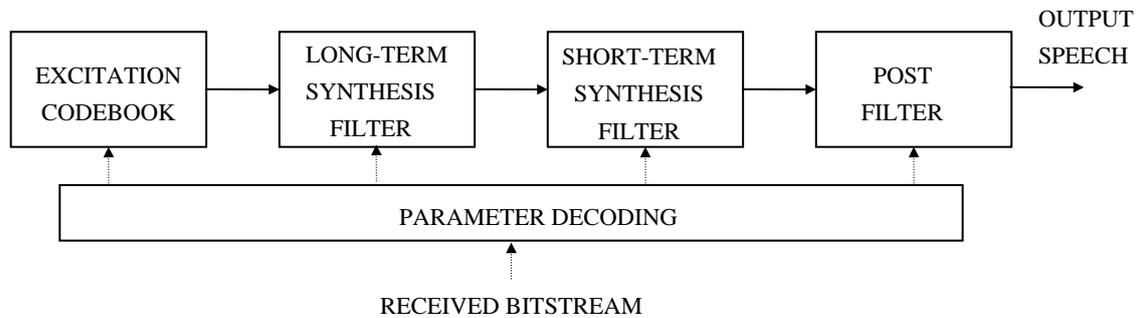


Figure 2 – 1 / JT-G729 Block diagram of conceptual CELP synthesis model.
(ITU-T G.729)

2. 1 符号器

符号化原理を Figure 2 – 2 / JT-G729 に示す。入力信号は、前処理ブロックにおいて高域通過フィルタリングおよびスケールングされる。前処理された信号は、その後の全ての分析に対する入力信号として提供される。LP 分析は、LP フィルタ係数の算出を 10ms フレームに 1 度行う。この係数は、線スペクトル対 (LSP) に変換され、そして、18 ビットの予測 2 段ベクトル量子化 (VQ) を用いて量子化される。励振信号は、入力音声と再生音声間の誤差を聴覚重み付き歪み尺度に基づいて最小化する合成による分析 (A-b-s) 法探索手順を用いて選択される。これは、聴覚重み付きフィルタで誤差信号をフィルタリングすることにより行われ、その係数は量子化されていない LP フィルタ係数を使用する。聴覚重み付けは、平坦な送話周波数応答を持つ系からの入力信号に対する性能を改善するために適応的な構成となっている。

励振信号パラメータ (固定および適応コードブックパラメータ) は、5ms (40 サンプル) 毎のサブフレーム単位で決定される。量子化された LP フィルタ係数および量子化されていない LP フィルタ係数は第 2 サブフレームに対して用いられ、第 1 サブフレームでは補間された LP フィルタ係数が用いられる (量子化されたものと量子化されていないもの両方)。開ループピッチ遅延は、聴覚重み付き音声信号を基に 10ms フレーム毎に 1 度推定される。その後、次の処理が各サブフレームに対して繰り返される。ターゲット信号 $x(n)$ は、重み付き合成フィルタ $W(z)/\hat{A}(z)$ を通して LP 残差をフィルタリングすることによって算出される。これらのフィルタの初期状態は、LP 残差信号と励振信号間の誤差をフィルタリングすることによって更新される。これは、重み付けされた音声信号から重み付き合成フィルタの零入力応答を差し引く通常の手法と等価である。そして重み付き合成フィルタのインパルス応答 $h(n)$ が算出される。さらに、閉ループピッチ分析によりターゲット $x(n)$ とインパルス応答 $h(n)$ を用いて開ループピッチ遅延の近傍を探索し、適応コードブック遅延と利得を確定する。分解能が 1/3 の分数ピッチ遅延が用いられる。ピッチ遅延は第 1 サブフレームにおいて 8 ビットで符号化され、第 2 サブフレームでは 5 ビットで差分符号化される。ターゲット信号 $x(n)$ は (フィルタリングされた) 適応コードブック寄与分を減ずることによって更新され、新たなターゲット $x'(n)$ が固定コードブック探索において最適励振信号の確定に用いられる。固定コードブック励振信号として、17 ビットの代数コードブックが用いられる。適応および固定コードブック寄与分の利得は、7 ビットでベクトル量子化される (MA 予測が固定コードブック利得に適用される)。最後に、確定された励振信号を用いてフィルタのメモリ内容を更新する。

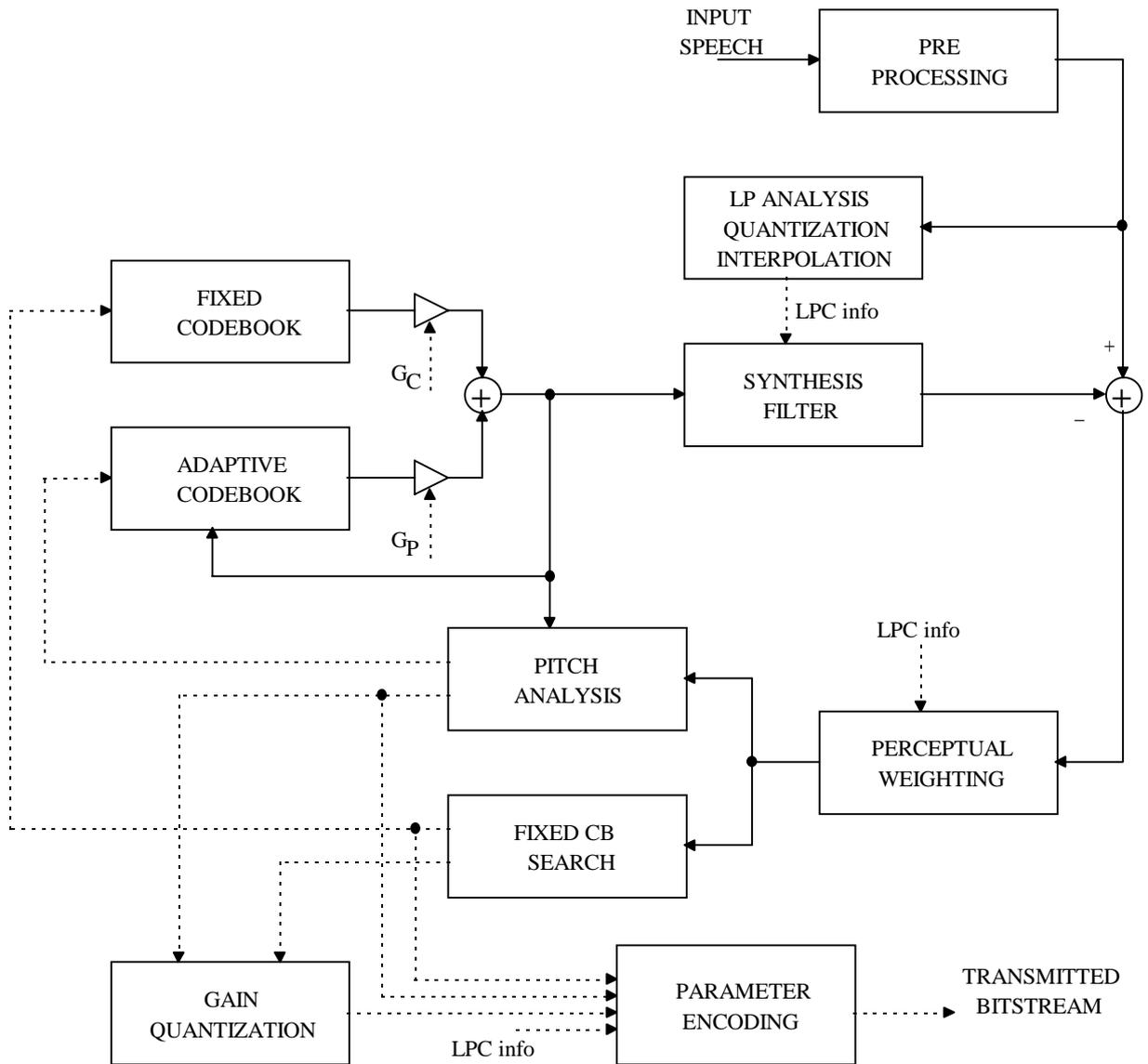


Figure 2 - 2 / JT-G729 Encoding principle of CS-ACELP encoder.
(ITU-T G.729)

2. 2 復号器

復号原理を Figure 2 - 3 / JT-G729 に示す。まず、パラメータインデックスが受信したビット列から抽出される。これらのインデックスは、10ms の音声フレームに対応したコーデックパラメータを取得するために復号される。これらのパラメータは L S P 係数、2つの分数ピッチ遅延、2つの固定コードブックベクトルおよび2つの適応および固定コードブック利得の組である。L S P 係数は補間され、そして、各サブフレームに対して L P フィルタ係数へ変換される。その後、各 5ms サブフレームに対して以下のステップが行われる。

- (1) 励振信号は、各々の利得によりスケールされた適応コードブックベクトルと固定コードブックベクトルを加算することによって生成される。
- (2) 音声は、L P 合成フィルタを通して励振信号をフィルタリングすることによって再生される。

- (3) 再生音声信号は、長期および短期合成フィルタを基にした適応ポストフィルタ、それに続く高域通過フィルタおよびスケーリング動作を含む後処理部を通過する。

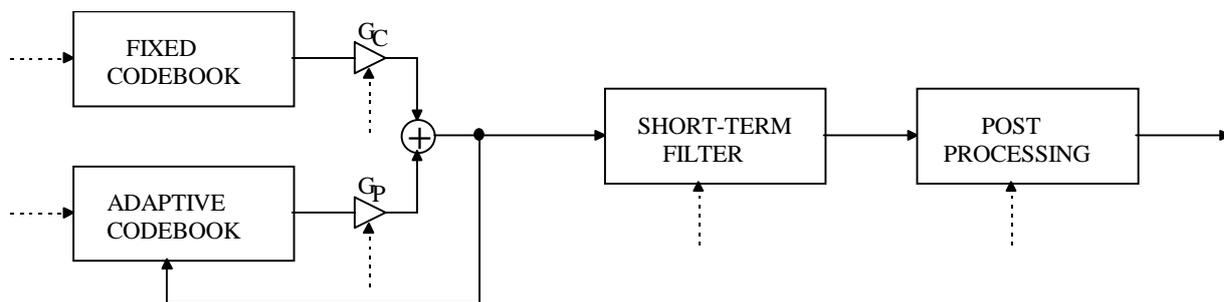


Figure 2 - 3 / JT-G729 Principle of CS-ACELP decoder.
(ITU-T G.729)

2. 3 遅延

本コーデックは、10ms のフレームを用いて、音声あるいは他のオーディオ信号を符号化する。さらに先読み 5ms が必要なので、アルゴリズム上の遅延は全体で 15ms である。本コーデックを実現するには、さらに以下の要因による遅延が加えられる。

- (1) 符号化および復号動作に必要な処理時間
- (2) 通信リンク上の伝送時間
- (3) オーディオデータと他のデータを多重化する際の多重化遅延

2. 4 音声コーデックの記述

本標準では、音声符号化アルゴリズムをビットイグザクトな固定小数点数値演算で記述している。5章に示した、本標準で必須のANSI Cコードは、このビットイグザクトな固定小数点記述を実現している。符号器(3章)と復号器(4章)の演算記述は他の方法でも実現可能であるが、その様な方法で実現されたコーデックは本標準には準拠しない可能性がある。よって、相違点がある場合、5章のANSI Cコードによるアルゴリズムの記述は、3章あるいは4章における数学的な演算記述より優先される。十分なものではないが、このANSI Cコードとともに使用できるテスト信号のセットは、新日本ITU協会より入手可能である。

2. 5 記法

本標準全般にわたり、以下に示す記法に従う。

- (1) コードブックは、カリグラフ文字で記述する(例 c)。
- (2) 時間領域の信号は、そのシンボルと丸括弧で括られたサンプル番号で記述する(例 $s(n)$)。シンボル n は、サンプル番号である。
- (3) 丸括弧で括られた上付きの添字は、時間に依存する変数に用いる(例 $g^{(m)}$)。変数 m はその前後関係よりフレーム番号、あるいはサブフレーム番号に対応している。
- (4) 再帰を示す添字は、角括弧で括られた上付きで記述する(例 $E^{[k]}$)。
- (5) 下付きの添字は、係数配列の各要素を示す。
- (6) 記号 $\hat{\cdot}$ は量子化されたパラメータを示す(例 \hat{g})。
- (7) パラメータの範囲は、角括弧で括られた値で記述する。この値は境界値を含む(例 $[0.6, 0.9]$)。

- (8) 関数 \log は、底 10 の対数を示す。
- (9) 関数 int は、切り捨てによる整数値への変換を示す。
- (10) 使用される 10 進の浮動小数点値は、16 ビット固定小数点 ANS I C での実現に使用された値を丸めたものである。

本標準全般にわたり主要なシンボルを、Table 2 – 2 / JT-G729 に示す。また主要な信号名を Table 2 – 3 / JT-G729 に、主要な変数名とその次数を Table 2 – 4 / JT-G729 に、定数を Table 2 – 5 / JT-G729 に、略語一覧を Table 2 – 6 / JT-G729 に示す。

Table 2 – 2 / JT-G729 Glossary of most relevant symbols.
(ITU-T G.729)

<i>Name</i>	<i>Reference</i>	<i>Description</i>
$1/\hat{A}(z)$	Eq. (2)	LP synthesis filter
$H_{h1}(z)$	Eq. (1)	Input high-pass filter
$H_p(z)$	Eq. (78)	Long-term postfilter
$H_f(z)$	Eq. (84)	Short-term postfilter
$H_t(z)$	Eq. (86)	Tilt-compensation filter
$H_{h2}(z)$	Eq. (91)	Output high-pass filter
$P(z)$	Eq. (46)	Pre-filter for fixed codebook
$W(z)$	Eq. (27)	Weighting filter

Table 2 – 3 / JT-G729 Glossary of most relevant signals.
(ITU-T G.729)

<i>Name</i>	<i>Reference</i>	<i>Description</i>
$c(n)$	Section 3.8	Fixed-codebook contribution
$d(n)$	Section 3.8.1	Correlation between target signal and $h(n)$
$ew(n)$	Section 3.10	Error signal
$h(n)$	Section 3.5	Impulse response of weighting and synthesis filters
$r(n)$	Section 3.6	Residual signal
$s(n)$	Section 3.1	Pre-processed speech signal
$\hat{s}(n)$	Section 4.1.6	Reconstructed speech signal
$s'(n)$	Section 3.2.1	Windowed speech signal
$sf(n)$	Section 4.2	Postfiltered output
$sf'(n)$	Section 4.2	Gain-scaled postfiltered output
$sw(n)$	Section 3.6	Weighted speech signal
$x(n)$	Section 3.6	Target signal
$x'(n)$	Section 3.8.1	Second target signal
$u(n)$	Section 3.10	Excitation to LP synthesis filter
$v(n)$	Section 3.7.1	Adaptive-codebook contribution
$y(n)$	Section 3.7.3	Convolution $v(n)*h(n)$
$z(n)$	Section 3.9	Convolution $c(n)*h(n)$

Table 2 – 4 / JT-G729 Glossary of most relevant variables.
(ITU-T G.729)

<i>Name</i>	<i>Size</i>	<i>Description</i>
g_p	1	Adaptive-codebook gain
g_c	1	Fixed-codebook gain
g_l	1	Gain term for long-term postfilter
g_s	1	Gain term for short-term postfilter
g_t	1	Gain term for tilt postfilter
G	1	Gain for gain normalization
T_{op}	1	Open-loop pitch delay
a_i	11	LP coefficients ($a_0 = 1.0$)
k_i	10	Reflection coefficients
k'_i	1	Reflection coefficient for tilt postfilter
o_i	2	LAR coefficients
ω_i	10	LSF normalized frequencies
$\hat{p}_{i,j}$	40	MA predictor for LSF quantization
q_i	10	LSP coefficients
$r(k)$	11	Auto-correlation coefficients
$r'(k)$	11	Modified auto-correlation coefficients
w_i	10	LSP weighting coefficients
\hat{l}_i	10	LSP quantizer output

Table 2 – 5 / JT-G729 Glossary of most relevant constants.
(ITU-T G.729)

<i>Name</i>	<i>Value</i>	<i>Description</i>
f_s	8000	Sampling frequency
f_0	60	Bandwidth expansion
γ_1	0.94/0.98	Weight factor perceptual weighting filter
γ_2	0.60/[0.4-0.7]	Weight factor perceptual weighting filter
γ_n	0.55	Weight factor post filter
γ_d	0.70	Weight factor post filter
γ_p	0.50	Weight factor pitch post filter
γ_t	0.90/0.2	Weight factor tilt post filter
C	Table 3-1/JT-G729	Fixed (algebraic) codebook
L0	Section 3.2.4	Moving-average predictor codebook
L1	Section 3.2.4	First stage LSP codebook
L2	Section 3.2.4	Second stage LSP codebook (low part)
L3	Section 3.2.4	Second stage LSP codebook (high part)
GA	Section 3.9	Gain codebook (first stage)
GB	Section 3.9	Gain codebook (second stage)
w_{lag}	Eq. (6)	Correlation lag window
w_{lp}	Eq. (3)	LP analysis window

Table 2 – 6 / JT-G729 Glossary of acronyms.
(ITU-T G.729)

<i>Acronym</i>	<i>Description</i>
CELP	Code-excited linear-prediction
CS-ACELP	Conjugate-structure algebraic-CELP
MA	Moving average
MSB	Most significant bit
MSE	Mean-squared error
LAR	Log area ratio
LP	Linear prediction
LSP	Line spectral pair
LSF	Line spectral frequency
VQ	Vector quantization

3. 符号器の機能記述

この章では、Figure 2-2/JT-G729 のブロックに示される符号器の各機能について記述する。詳細な信号の流れを Figure 3-1/JT-G729 に示す。

3.1 前処理

2章で記述しているように、音声符号器への入力は16ビットPCM信号であると仮定する。入力信号には、符号化処理の前に以下の2つの前処理が適用される。

- (1) 信号スケーリング
- (2) 高域通過フィルタリング

スケーリングは、固定小数点演算におけるオーバーフローを避けるため、入力を2で除算する。高域通過フィルタは、不要な低周波数成分を除去するために用いられる。このフィルタには、カットオフ周波数が140Hzの2次の極零フィルタが用いられる。スケーリングと高域通過フィルタリングは、高域通過フィルタの分子の係数を2で除算することによって結合される。その結果生成されたフィルタは、以下の式で与えられる。

$$H_{h1}(z) = \frac{0.46363718 - 0.92724705z^{-1} + 0.46363718z^{-2}}{1 - 1.9059465z^{-1} + 0.9114024z^{-2}} \quad (1)$$

$H_{h1}(z)$ によってフィルタリングされた入力信号は、 $s(n)$ として参照され、以降の全ての符号化演算で使用される。

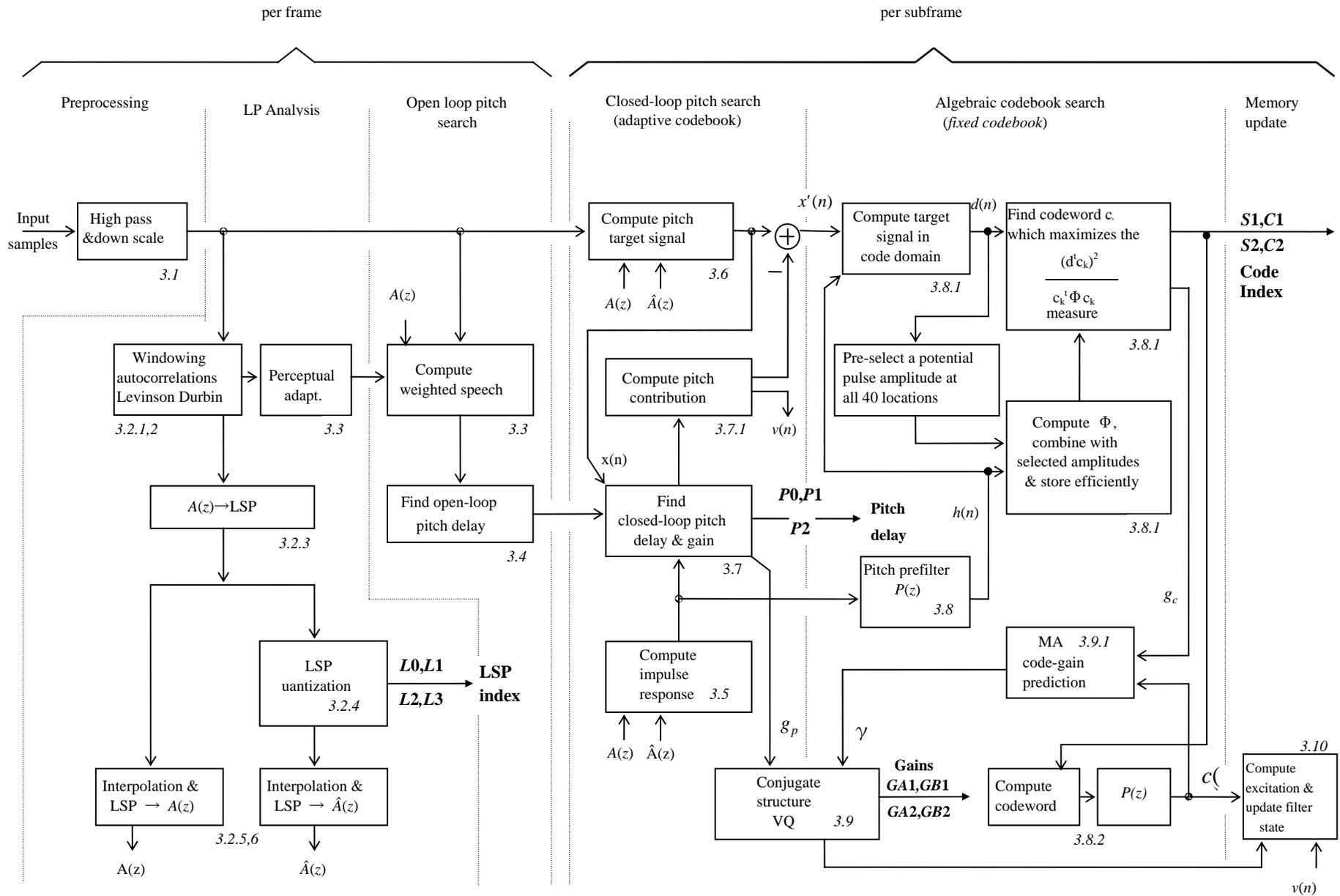
3.2 線形予測分析と量子化

短期の分析および合成フィルタは、10次の線形予測(LP)フィルタに基づいている。LP合成フィルタは、以下の通り定義される。

$$\frac{1}{\hat{A}(z)} = \frac{1}{1 + \sum_{i=1}^{10} \hat{a}_i z^{-i}} \quad (2)$$

ここで、 $\hat{a}_i, i=1, \dots, 10$ は(量子化された)線形予測(LP)係数である。短期予測分析すなわち線形予測分析は、30msの非対称窓による自己相関法を用いて、音声フレーム毎に1回実行される。80サンプル(10ms)毎に、窓かけされた音声の自己相関係数が計算され、レビンソンアルゴリズムを用いてLP係数に変換される。さらに、LP係数は、量子化および補間のためにLSP領域に変換される。量子化され補間されたフィルタ係数、および量子化されずに補間されたフィルタ係数は、再びLPフィルタ係数に変換される(サブフレーム毎の合成フィルタと重み付けフィルタを生成するため)。

Figure 3-1 / JT-G729 Signal flow at the CS-ACELP encoder.
(ITU-T G.729)



3. 2. 1 窓かけと自己相関計算

LP分析窓は2つの部分から成る。第1部はハミング窓の半分で、第2部はコサイン関数周期の4分の1である。その窓は次式で与えられる。

$$w_{lp}(n) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{399}\right) & , n = 0, \dots, 199 \\ \cos\left(\frac{2\pi(n-200)}{159}\right) & , n = 200, \dots, 239 \end{cases} \quad (3)$$

LP分析では5msの先読みが存在し、これは未来の音声フレームから40サンプルが必要とされることを意味する。これは符号器の段階で5msの余分なアルゴリズム遅延となる。LP分析窓は、過去の音声フレームからの120サンプル、現在の音声フレームからの80サンプル、未来の音声フレームからの40サンプルに適用される。窓かけ処理をFigure 3-2/JT-G729に示す。

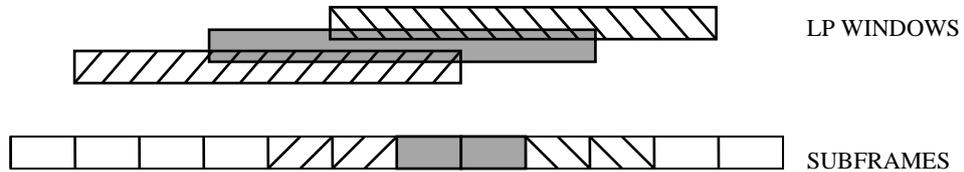


Figure 3-2/JT-G729 Windowing procedure in LP analysis. The different shading patterns identify (ITU-T G.729) corresponding excitation and LP analysis windows.

窓かけされた音声

$$s'(n) = w_{lp}(n)s(n), \quad n = 0, \dots, 239 \quad (4)$$

は、自己相関係数

$$r(k) = \sum_{n=k}^{239} s'(n)s'(n-k), \quad k = 0, \dots, 10 \quad (5)$$

の計算に使われる。

低レベル入力信号の算術上の問題を避けるために $r(0)$ の値は、 $r(0)=1.0$ を下限とする。自己相関係数に次式を乗ずることにより、60Hzの帯域幅拡張が適用される。

$$w_{lag}(k) = \exp\left[-\frac{1}{2}\left(\frac{2\pi f_0 k}{f_s}\right)^2\right], \quad k = 1, \dots, 10 \quad (6)$$

ここで $f_0=60\text{Hz}$ は帯域幅拡張、 $f_s=8000\text{Hz}$ は標準化周波数である。さらに $r(0)$ には白色雑音補正係数1.0001が乗じられ、これは-40dBの雑音を加えることを意味する。補正された自己相関係数は次式で与えられる。

$$\begin{aligned} r'(0) &= 1.0001r(0), \\ r'(k) &= w_{\text{lag}}(k)r(k), \quad k = 1, \dots, 10 \end{aligned} \quad (7)$$

3. 2. 2 レビンソン・ダービンアルゴリズム

補正された自己相関係数 $r'(k)$ は、LPフィルタ係数 $a_i, i = 1, \dots, 10$ を得るために次式で使われる。

$$\sum_{i=1}^{10} a_i r'(|i-k|) = -r'(k), \quad k = 1, \dots, 10 \quad (8)$$

式(8)の方程式の組は、レビンソン・ダービンアルゴリズムを使って解かれる。このアルゴリズムは次に示す再帰法を使う。

$$\begin{aligned} E^{[0]} &= r'(0) \\ \text{for } i &= 1 \text{ to } 10 \\ & a_0^{[i-1]} = 1 \\ & k_i = - \left[\sum_{j=0}^{i-1} a_j^{[i-1]} r'(i-j) \right] / E^{[i-1]} \\ & a_i^{[i]} = k_i \\ & \text{for } j = 1 \text{ to } i-1 \\ & a_j^{[i]} = a_j^{[i-1]} + k_i a_{i-j}^{[i-1]} \\ & \text{end} \\ & E^{[i]} = (1 - k_i^2) E^{[i-1]}, \\ \text{end} \end{aligned}$$

最終的には $a_j = a_j^{[10]}, j = 0, \dots, 10$ の解が得られる。なお、 $a_0 = 1.0$ である。

3. 2. 3 LPからLSPへの変換

LPフィルタ係数 $a_i, i = 0, \dots, 10$ は、量子化と補間を行うために線スペクトル対 (LSP) に変換される。10次のLPフィルタについてLSP係数は、以下の和および差多項式の根として定義される。

$$F_1'(z) = A(z) + z^{-11} A(z^{-1}) \quad (9)$$

$$F_2'(z) = A(z) - z^{-11} A(z^{-1}) \quad (10)$$

多項式 $F_1'(z)$ は対称で、多項式 $F_2'(z)$ は非対称である。これらの多項式のすべての根は単位円上にあり、そしてそれらは交互に並ぶということが証明されている。 $F_1'(z)$ は根 $z = -1 (\omega = \pi)$ を持ち、そして $F_2'(z)$ は根 $z = 1 (\omega = 0)$ を持つ。これらの2つの根は新しい多項式

$$F_1(z) = F_1'(z) / (1 + z^{-1}) \quad (11)$$

と

$$F_2(z) = F_2'(z) / (1 - z^{-1}) \quad (12)$$

を定義することによって消去される。

どちらの多項式も単位円 ($e^{\pm j\omega_i}$) 上に5つの共役な根を持ち、それらは以下の式として記述できる。

$$F_1(z) = \prod_{i=1,3,\dots,9} (1 - 2q_i z^{-1} + z^{-2}) \quad (13)$$

$$F_2(z) = \prod_{i=2,4,\dots,10} (1 - 2q_i z^{-1} + z^{-2}) \quad (14)$$

ここで $q_i = \cos(\omega_i)$ である。係数 ω_i は線スペクトル周波数 (L S F) で、それらは、順序性 $0 < \omega_1 < \omega_2 < \dots < \omega_{10} < \pi$ を満たしている。係数 q_i をコサイン領域の L S P 係数と呼ぶ。

多項式 $F_1(z)$ と $F_2(z)$ は両方とも対称であるので、それぞれの多項式の最初の5つの係数だけを計算すればよい。これらの多項式の係数は以下の再帰的關係によって求められる。

$$\begin{aligned} f_1(i+1) &= a_{i+1} + a_{10-i} - f_1(i), & i &= 0, \dots, 4 \\ f_2(i+1) &= a_{i+1} - a_{10-i} + f_2(i), & i &= 0, \dots, 4 \end{aligned} \quad (15)$$

ここで $f_1(0) = f_2(0) = 1.0$ である。L S P 係数は、0 から π までの間の等間隔な60の点で多項式 $F_1(z)$ と $F_2(z)$ の値を求め、極性変化を調べることで見つけられる。極性の変化は根の存在を示し、そのとき極性の変化している区間ではより良い根を求める為に、区間の分割が4回行われる。 $F_1(z)$ と $F_2(z)$ を解くためにチェビシエフ多項式が使われる。この方法により根はコサイン領域で直接的に見つけられる。多項式 $F_1(z)$ または $F_2(z)$ は $z = e^{j\omega}$ として次式のように書ける。

$$F(\omega) = 2e^{-j5\omega} C(x) \quad (16)$$

$$C(x) = T_5(x) + f(1)T_4(x) + f(2)T_3(x) + f(3)T_2(x) + f(4)T_1(x) + f(5)/2 \quad (17)$$

ここで $T_m(x) = \cos(m\omega)$ は m 次のチェビシエフ多項式であり、 $f(i), i = 1, \dots, 5$ は $F_1(z)$ もしくは $F_2(z)$ の係数であり、式(15)を使って計算される。多項式 $C(x)$ は、 $x = \cos(\omega)$ として次の再帰的關係を使って求められる。

$$\begin{aligned} & \text{for } k = 4 \text{ down to } 1 \\ & \quad b_k = 2xb_{k+1} - b_{k+2} + f(5-k) \\ & \text{end} \\ & C(x) = xb_1 - b_2 + f(5)/2 \end{aligned}$$

ただし初期値は、 $b_5 = 1, b_6 = 0$ である。

3. 2. 4 L S P 係数の量子化

L S P 係数 q_i は、周波数領域 $[0, \pi]$ において正規化された L S F で表現したものである ω_i を用いて量子化される。即ち、

$$\omega_i = \arccos(q_i), \quad i = 1, \dots, 10 \quad (18)$$

現在のフレームのL S F係数を予測するため、4次の切換え型移動平均（MA）予測を用いる。算出された係数と予測された係数の差分は、2段ベクトル量子化器を用いて量子化される。1段目では、128候補（7ビット）を持つコードブックL1を用いた10次元ベクトル量子化を行う。2段目は10ビットのベクトル量子化であり、それぞれ32候補（5ビット）を持つ2つの5次元コードブックL2,L3を用いて、配列を分割して別々にベクトル量子化を実行する。

量子化のプロセスを判りやすく表現するために、まず最初に復号のプロセスを示す。各係数はそれぞれ2つのコードブックの和を求める事によって得られる。

$$\hat{l}_i = \begin{cases} L1_i(L1) + L2_i(L2), & i = 1, \dots, 5 \\ L1_i(L1) + L3_{i-5}(L3), & i = 6, \dots, 10 \end{cases} \quad (19)$$

ここで、L1,L2,およびL3は各コードブックのインデックスである。量子化されたLP合成フィルタにおいて強い共振が起きないようにするため、隣接する係数の距離が最低でもJとなるように係数 \hat{l}_i の値を配置する。再配置処理を以下に記す。

```

for i=2,...,10
  if ( $\hat{l}_{i-1} > \hat{l}_i - J$ )
     $\hat{l}_{i-1} = (\hat{l}_i + \hat{l}_{i-1} - J) / 2$ 
     $\hat{l}_i = (\hat{l}_i + \hat{l}_{i-1} + J) / 2$ 
  end
end

```

この再配置処理は2度実行される。第1の再配置では $J=0.0012$ の値を、第2の再配置では $J=0.0006$ を用いる。再配置処理の後、現在のフレーム m に対応する量子化されたL S F係数 $\hat{\omega}_i^{(m)}$ を求める。 $\hat{\omega}_i^{(m)}$ は、過去のフレームの量子化器の出力 $\hat{l}_i^{(m-k)}$ を重み付けして和をとったものと、現在のフレームの量子化器の出力 $\hat{l}_i^{(m)}$ から得られる。

$$\hat{\omega}_i^{(m)} = \left(1 - \sum_{k=1}^4 \hat{p}_{i,k} \right) \hat{l}_i^{(m)} + \sum_{k=1}^4 \hat{p}_{i,k} \hat{l}_i^{(m-k)}, \quad i = 1, \dots, 10 \quad (20)$$

ここで、 $\hat{p}_{i,k}$ は切換え型MA予測器の係数である。いずれのMA予測器を用いるかは、識別ビットL0によって決まる。スタートアップ時の $\hat{l}_i^{(k)}$ の初期値は、 $k < 0$ を満たす全ての k について $\hat{l}_i = i\pi/11$ で与えられる。

$\hat{\omega}_i$ の計算後、次は、得られた合成フィルタの安定性をチェックする。これは、以下のようにして行う。

- (1) $\hat{\omega}_i$ を値の小さな順に並べる。
- (2) $\hat{\omega}_1 < 0.005$ ならば $\hat{\omega}_1 = 0.005$ とする。
- (3) $\hat{\omega}_{i+1} - \hat{\omega}_i < 0.0391$ ならば $\hat{\omega}_{i+1} = \hat{\omega}_i + 0.0391$, ($i = 1, \dots, 9$)とする。
- (4) $\hat{\omega}_{10} > 3.135$ ならば $\hat{\omega}_{10} = 3.135$ とする。

L S Fパラメータの符号化手順の概要を以下に示す。2つのMA予測器についてそれぞれ現フレームのL S F係数と比較して最も近い、次式の重み付き平均自乗誤差を最小にする値を求める。

$$E_{lsf} = \sum_{i=1}^{10} w_i (\omega_i - \hat{\omega}_i)^2 \quad (21)$$

重み係数 w_i は、非量子化 L S F 係数の関数として適応的に計算される。

$$\begin{aligned}
 w_1 &= \begin{cases} 1.0 & \text{if } \omega_2 - 0.04\pi - 1 > 0 \\ 10(\omega_2 - 0.04\pi - 1)^2 + 1 & \text{otherwise} \end{cases} \\
 w_i \quad (2 \leq i \leq 9) &= \begin{cases} 1.0 & \text{if } \omega_{i+1} - \omega_{i-1} - 1 > 0 \\ 10(\omega_{i+1} - \omega_{i-1} - 1)^2 + 1 & \text{otherwise} \end{cases} \\
 w_{10} &= \begin{cases} 1.0 & \text{if } -\omega_9 + 0.92\pi - 1 > 0 \\ 10(-\omega_9 + 0.92\pi - 1)^2 + 1 & \text{otherwise} \end{cases}
 \end{aligned} \tag{22}$$

さらに、重み係数 w_5 と w_6 については、それぞれ 1.2 を掛ける。

現在のフレーム m における量子化すべきベクトルは次式で与えられる。

$$l_i = \left[\omega_i^{(m)} - \sum_{k=1}^4 \hat{p}_{i,k} \hat{l}_i^{(m-k)} \right] / \left(1 - \sum_{k=1}^4 \hat{p}_{i,k} \right), \quad i = 1, \dots, 10 \tag{23}$$

まず、第 1 コードブック L1 を探索して、（重み付けされていない）平均自乗誤差を最小にするベクトル L1 を選択する。次に、第 2 段階において低次の係数を定義する第 2 コードブック L2 の探索を行う。各々の候補についてすべて、式(20)を用いて部分ベクトル $\hat{\omega}_i$ ($i = 1, \dots, 5$) を再現する。隣接要素の最小距離が 0.0012 を保証するように $\hat{\omega}_i$ の再配置を行う。式(21)の重み付き平均自乗誤差を計算し、誤差を最小とするベクトル L2 を選択する。第 1 段階で選択されたベクトル L1 と、第 2 段階で選択された低次のベクトル L2 を用いて、第 2 段階の高次の係数を、コードブック L3 から探索する。再び、隣接要素の最小距離が 0.0012 を保証するように $\hat{\omega}_i$ の再配置を行う。重み付き平均自乗誤差を最小とするベクトル L3 を選択する。この結果得られた量子化ベクトル \hat{l}_i ($i = 1, \dots, 10$) を、今度は隣接要素の最小距離が 0.0006 を保証するように再配置する。以上の処理を、L0 で定義される 2 つの MA 予測器の各々について実行する。そして、最小の重み付き平均自乗誤差をとる MA 予測器 L0 を選択する。本節の冒頭で述べたように、量子化された L S F 係数 $\hat{\omega}_i$ を求めるために、選択されたベクトル \hat{l}_i に対して 2 度の再配置処理と、（合成フィルタの）安定性チェックを行っている。

3. 2. 5 L S P 係数の補間

量子化された（および、量子化されていない）L P 係数は、第 2 サブフレームで使用される。第 1 サブフレームでは、量子化された（および、量子化されていない）L P 係数は、隣接するサブフレームにおける該当する係数の線形補間によって得られる。補間は、L S P 係数のコサイン領域において行われる。 $q_i^{(current)}$ を現在の 10ms フレームにおいて計算された L S P 係数、そして $q_i^{(previous)}$ を直前の 10ms フレームにて計算された L S P 係数とする。2 つのサブフレームで用いられる（量子化されていない）補間された L S P 係数は、各々次式で与えられる。

$$\begin{aligned}
 \text{Subframe 1: } q_i^{(1)} &= 0.5q_i^{(previous)} + 0.5q_i^{(current)}, \quad i = 1, \dots, 10 \\
 \text{Subframe 2: } q_i^{(2)} &= q_i^{(current)}, \quad i = 1, \dots, 10
 \end{aligned} \tag{24}$$

同様な補間処理を用いて、量子化された L S P 係数の補間も行う。この場合、式(24)の q_i を \hat{q}_i に置換すればよい。

3. 2. 6 L S P から L P への変換

L S P 係数は量子化および補間されたのち L P 係数 a_i に逆変換される。この変換は以下のように行う。

$F_1(z)$ および $F_2(z)$ の係数は量子化および補間されたLSP係数が既知であるため、式(13)および(14)の展開式から求まる。係数 $f_1(i), i=1, \dots, 5$ は q_i から以下の再帰式を用いて計算される。

```

for i = 1 to 5
   $f_1(i) = -2q_{2i-1}f_1(i-1) + 2f_1(i-2)$ 
  for j = i - 1 down to 1
     $f_1^{[i]}(j) = f_1^{[i-1]}(j) - 2q_{2i-1}f_1^{[i-1]}(j-1) + f_1^{[i-1]}(j-2)$ 
  end
end

```

ただし初期値は $f_1(0)=1$ および $f_1(-1)=0$ である。係数 $f_2(i)$ は q_{2i-1} を q_{2i} に置き換えることにより同様に計算される。

係数 $f_1(i)$ および $f_2(i)$ が求まったら、 $F_1'(z)$ および $F_2'(z)$ を得るために、以下のように $F_1(z)$ および $F_2(z)$ にそれぞれ $1+z^{-1}$ および $1-z^{-1}$ を乗算する。

$$\begin{aligned} f_1'(i) &= f_1(i) + f_1(i-1), & i=1, \dots, 5 \\ f_2'(i) &= f_2(i) - f_2(i-1), & i=1, \dots, 5 \end{aligned} \quad (25)$$

最終的にLSP係数は $f_1'(i)$ および $f_2'(i)$ から以下のように計算される。

$$a_i = \begin{cases} 0.5f_1'(i) + 0.5f_2'(i), & i=1, \dots, 5 \\ 0.5f_1'(11-i) - 0.5f_2'(11-i), & i=6, \dots, 10 \end{cases} \quad (26)$$

これは $A(z) = (F_1'(z) + F_2'(z))/2$ の関係があり、また $F_1'(z)$ および $F_2'(z)$ がそれぞれ対称、非対称の多項式であることから直接的に導かれる。

3. 3 聴覚重み付け

聴覚重み付けフィルタは量子化されていないLSPフィルタ係数 a_i に基づいており以下の式で与えられる。

$$W(z) = \frac{A(z/\gamma_1)}{A(z/\gamma_2)} = \frac{1 + \sum_{i=1}^{10} \gamma_1^i a_i z^{-i}}{1 + \sum_{i=1}^{10} \gamma_2^i a_i z^{-i}} \quad (27)$$

γ_1 および γ_2 の値はフィルタ $W(z)$ の周波数応答を決める。これらの値を適切に調整することによって、重み付けをより効果的にすることができる。これは γ_1 および γ_2 を入力信号のスペクトル形状の関数とすることにより行う。この適応は10msフレーム毎に1回行われるが、この適応処理を円滑にするために第1サブフレームの係数は補間によって求められる。スペクトル形状は2次の線形予測フィルタ係数から得られるが、これはレビンソン・ダービン再帰部(3.2.2節)の副産物として得られる。反射係数 k_i は以下のように対数断面積比(LAR)係数 o_i に変換される。

$$o_i = \log \frac{(1.0 + k_i)}{(1.0 - k_i)} \quad i=1, 2 \quad (28)$$

現在の 10ms フレームに対応する LAR 係数は第 2 サブフレームで使用される。第 1 サブフレームの LAR 係数は前フレームの LAR パラメータとの間を線形補間することにより得られる。補間後の 2 つのサブフレームの LAR 係数はそれぞれ以下ようになる。

$$\begin{aligned} \text{Subframe 1 : } o_i^{(1)} &= 0.5o_i^{(\text{previous})} + 0.5o_i^{(\text{current})}, & i = 1,2 \\ \text{Subframe 2 : } o_i^{(2)} &= o_i^{(\text{current})}, & i = 1,2 \end{aligned} \quad (29)$$

スペクトル包絡は、平坦 ($flat=1$) であるか、あるいは傾斜 ($flat=0$) しているかで特徴付けられる。各サブフレームにおいて LAR 係数に対して、しきい値関数を適用することによりこの特徴付けを得ることができる。頻繁に変化することを避けるため、前サブフレーム $m-1$ の $flat$ の値を考慮に入れることによりヒステリシスを持たせている。

$$flat^{(m)} = \begin{cases} 0 & \text{if } o_1^{(m)} < -1.74 \text{ and } o_2^{(m)} > 0.65 \text{ and } flat^{(m-1)} = 1 \\ 1 & \text{if } (o_1^{(m)} > -1.52 \text{ or } o_2^{(m)} < 0.43) \text{ and } flat^{(m-1)} = 0 \\ flat^{(m-1)} & \text{otherwise} \end{cases} \quad (30)$$

補間されたサブフレームのスペクトルが平坦 ($flat^{(m)} = 1$) に分類された場合、重み付け係数は $\gamma_1 = 0.94$ および $\gamma_2 = 0.6$ に設定される。スペクトルが傾斜 ($flat^{(m)} = 0$) に分類された場合には γ_1 の値は 0.98 に設定され、 γ_2 の値は LP 合成フィルタの共振の大きさに適応する。ただし 0.4 から 0.7 の範囲に制限される。もし大きな共振が現れたら γ_2 の値は上限に近い値に設定される。この適応は、現サブフレームの連続する 2 つの LSP 係数の間の最小距離に基づいて行われる。最小距離は次式で与えられる。

$$d_{min} = \min[\omega_{i+1} - \omega_i], \quad i = 1, \dots, 9 \quad (31)$$

γ_2 の値は以下のように線形な関係で計算される。

$$\gamma_2 = -6.0d_{min} + 1.0, \text{ bounded by } 0.4 \leq \gamma_2 \leq 0.7 \quad (32)$$

サブフレームにおける重み付き音声信号は次式で与えられる。

$$sw(n) = s(n) + \sum_{i=1}^{10} a_i \gamma_1^i s(n-i) - \sum_{i=1}^{10} a_i \gamma_2^i sw(n-i), \quad n = 0, \dots, 39 \quad (33)$$

重み付き音声信号 $sw(n)$ は音声フレームのピッチ遅延を推定するために使用される。

3. 4 開ループピッチ分析

最適適応コードブック遅延に対する探索の演算量を削減するため、探索範囲は遅延候補 T_{op} の近傍に限定される。 T_{op} は開ループピッチ分析によって得られる。開ループピッチ分析はフレーム(10ms)毎に行われる。開ループピッチの算出は、式(33)の重み付き音声信号 $sw(n)$ を用いて以下のように行われる。まず最初に、相関関数

$$R(k) = \sum_{n=0}^{79} sw(n)sw(n-k) \quad (34)$$

の最大値を以下の3つの区間についてそれぞれ求める。

$$i = 1: 80, \dots, 143$$

$$i = 2: 40, \dots, 79$$

$$i = 3: 20, \dots, 39$$

得られた最大値 $R(t_i)$, $i = 1, \dots, 3$, は以下の式により正規化される。

$$R'(t_i) = \frac{R(t_i)}{\sqrt{\sum_n sw^2(n-t_i)}}, \quad i = 1, \dots, 3 \quad (35)$$

3つの正規化相関の中から最終候補が選択される。ここでは、長い遅延に対応する正規化相関に重み付けを行うことによって、短い遅延の区間の値が選ばれやすくなるようにしている。最適開ループ遅延 T_{op} は以下のように決定される。

```

 $T_{op} = t_1$ 
 $R'(T_{op}) = R'(t_1)$ 
if  $R'(t_2) \geq 0.85R'(T_{op})$ 
     $R'(T_{op}) = R'(t_2)$ 
     $T_{op} = t_2$ 
end
if  $R'(t_3) \geq 0.85R'(T_{op})$ 
     $R'(T_{op}) = R'(t_3)$ 
     $T_{op} = t_3$ 
end

```

遅延区間を3つに分けて、より小さい値を選択しやすくする手順は、倍数ピッチの選択を避けるために用いられている。

3. 5 インパルス応答の計算

重み付き合成フィルタ $W(z)/\hat{A}(z)$ のインパルス応答 $h(n)$ は適応コードブックと固定コードブックの探索に必要となる。インパルス応答 $h(n)$ は零詰めされたフィルタ $A(z/\gamma_1)$ の係数から成る信号を2つのフィルタ $1/\hat{A}(z)$ と $1/A(z/\gamma_2)$ に通すことにより、サブフレーム毎に算出される。

3. 6 ターゲット信号の計算

適応コードブック探索のターゲット信号 $x(n)$ は、一般的に重み付き合成フィルタ $W(z)/\hat{A}(z) = A(z/\gamma_1)/[\hat{A}(z)A(z/\gamma_2)]$ の零入力応答を式(33)の重み付き音声信号 $sw(n)$ から減ずることによって算出される。これはサブフレーム単位で行われる。

本標準で使用されているターゲット信号算出のための等価手順は、LP残差信号 $r(n)$ を合成フィルタ $1/\hat{A}(z)$ と重み付けフィルタ $A(z/\gamma_1)/A(z/\gamma_2)$ をつなげたものを通してフィルタリングするものである。サブ

フレームの励振信号が決定された後、これらのフィルタの初期状態は残差信号と励振信号の差分信号をフィルタリングすることによって更新される。これら3つのフィルタのメモリ更新については3.10節で説明する。

ターゲットベクトル算出時に必要となる残差信号 $r(n)$ は、適応コードブック探索時にも過去の励振信号のバッファを拡張するために用いられる。これは、次節で説明されるようにサブフレーム長40より短い遅延に対する適応コードブック探索手順を簡略化することになる。このLP残差信号は次式で与えられる。

$$r(n) = s(n) + \sum_{i=1}^{10} \hat{a}_i s(n-i), \quad n = 0, \dots, 39 \quad (36)$$

3.7 適応コードブック探索

適応コードブックパラメータ（またはピッチパラメータ）とは遅延と利得のことである。ピッチフィルタを実現するために適応コードブック探索では励振信号はサブフレーム長以下の遅延の場合に繰り返される。探索中は、閉ループ探索を簡略化させるためにLP残差によって励振信号が拡張される。適応コード

ブック探索は5msサブフレーム毎に行われる。第1サブフレームにおいて分数ピッチ遅延 T_1 は $\left[19\frac{1}{3}, 84\frac{2}{3}\right]$ の範囲内では1/3の分析精度で、[85,143]の範囲内では整数のみが使用される。第2サブフレームにおいては、遅延 T_2 は常に $\left[\text{int}(T_1) - 5\frac{2}{3}, \text{int}(T_1) + 4\frac{2}{3}\right]$ の範囲内で1/3の分析精度で使用される。ここで $\text{int}(T_1)$ は第1サブフレームの分数ピッチ遅延 T_1 の整数部分である。この範囲は T_1 が含まれる遅延範囲の境界に適応している。

各サブフレームに対して最適な遅延は重み付き平均自乗誤差を最小にする閉ループピッチ分析を用いて決定される。第1サブフレームにおいて遅延 T_1 は開ループ遅延 T_{op} （3.4節参照）の周辺の6サンプルの遅延値を探索することによって求められる。この探索の境界値 t_{min} および t_{max} は以下のように定義される。

```

 $t_{min} = T_{op} - 3$ 
if  $t_{min} < 20$  then  $t_{min} = 20$ 
 $t_{max} = t_{min} + 6$ 
if  $t_{max} > 143$  then
     $t_{max} = 143$ 
     $t_{min} = t_{max} - 6$ 
end

```

第2サブフレームに対して、閉ループピッチ分析は最適な遅延 T_2 を求めるために、第1サブフレームにおいて選択されたピッチの周辺に対して行われる。この探索の境界値は $t_{min} - 2/3$ および $t_{max} + 2/3$ である、ここで t_{min} および t_{max} は T_1 から以下の手順で求められる。

```

 $t_{min} = \text{int}(T_1) - 5$ 
if  $t_{min} < 20$  then  $t_{min} = 20$ 
 $t_{max} = t_{min} + 9$ 
if  $t_{max} > 143$  then
     $t_{max} = 143$ 
     $t_{min} = t_{max} - 9$ 
end

```

閉ループピッチ探索は入力音声と再生音声間の重み付き平均自乗誤差を最小にする。これは以下の項を最大にすることによって達成される。

$$R(k) = \frac{\sum_{n=0}^{39} x(n)y_k(n)}{\sqrt{\sum_{n=0}^{39} y_k(n)y_k(n)}} \quad (37)$$

ここで $x(n)$ はターゲット信号であり、 $y_k(n)$ は遅延 k でフィルタ処理された過去の励振信号 ($h(n)$ で畳み込まれた過去の励振信号) である。探索範囲は予備選択された第 1 サブフレームの開ループピッチ T_{op} および第 2 サブフレームの T_1 の近傍に限定されることに注意しなければならない。

畳込み $y_k(n)$ は遅延 t_{min} に対して計算される。探索範囲内のその他の整数の遅延 $k=t_{min}+1, \dots, t_{max}$ に対しては再帰的な関係を用いて更新される。

$$y_k(n) = y_{k-1}(n-1) + u(-k)h(n), \quad n = 39, \dots, 0 \quad (38)$$

ここで $u(n), n = -143, \dots, 39$ は励振信号バッファであり、 $y_{k-1}(-1) = 0$ である。探索中にはサンプル $u(n), n = 0, \dots, 39$ は未知であることに注意しなければならない。またそれらの値は 40 以下のピッチ遅延のために必要である。LP 残差は式(38)を全ての遅延に対して有効にするために $u(n)$ にコピーされ、探索を簡略化させる。

もし最適な整数閉ループピッチ遅延が 85 以下のときは、 T_2 および T_1 の値を決定するために、最適な整数遅延の周辺の分数値をテストしなければならない。分数ピッチ探索は式(37)の正規化された相関を補間し、その最大値を探索することによって行われる。補間はハミング窓をかけられた標本化関数を基にした FIR フィルタ b_{12} を用いて行われる。この標本化関数は ± 11 で打ち切れ、 ± 12 で 0 になる ($b_{12}(12) = 0$)。このフィルタはオーバサンプル領域内の 3600Hz においてカットオフ周波数 (-3dB) をもつ。 $R(k)$ の $-2/3, -1/3, 0, 1/3$ および $2/3$ に対する補間値は以下の補間公式を用いて得られる。

$$R(k)_t = \sum_{i=0}^3 R(k-i)b_{12}(t+i \cdot 3) + \sum_{i=0}^3 R(k+1+i)b_{12}(3-t+i \cdot 3), \quad t = 0, 1, 2 \quad (39)$$

ここで $t=0, 1, 2$ はそれぞれ分数 $0, 1/3, 2/3$ に対応する。式(37)の相関は正確な補間を導くために範囲 $t_{min}-4, t_{max}+4$ を用いて計算する必要がある。

3. 7. 1 適応コードブックベクトルの生成

一度ピッチ遅延が決定されると、適応コードブックベクトル $v(n)$ が与えられた整数遅延 k および分数値 t における過去の励振信号 $u(n)$ を補間することによって計算される。

$$v(n) = \sum_{i=0}^9 u(n-k-i)b_{30}(t+i \cdot 3) + \sum_{i=0}^9 u(n-k+1+i)b_{30}(3-t+i \cdot 3), \quad n = 0, \dots, 39, t = 0, 1, 2 \quad (40)$$

この補間フィルタ b_{30} はハミング窓をかけられた標本化関数をもとにしている。この標本化関数は ± 29 で打ち切られ ± 30 において 0 になる ($b_{30}(30)=0$)。このフィルタはオーバサンプル領域内の 3600Hz においてカットオフ周波数 (-3dB) を持つ。

3. 7. 2 適応コードブック遅延に対する符号語の計算

第1サブフレームのピッチ遅延 T_1 は8ビットで符号化され、第2サブフレームの相対遅延は5ビットで符号化される。分数遅延 T は整数部分 $\text{int}(T)$ と分数部分 $\text{frac}/3, \text{frac}=-1,0,1$ を用いて表現される。ピッチインデックス $P1$ はこのとき以下の通りに符号化される。

$$P1 = \begin{cases} 3 \times (\text{int}(T_1) - 19) + \text{frac} - 1, & \text{if } T_1 = [19, \dots, 85], \text{frac} = [-1, 0, 1], \\ (\text{int}(T_1) - 85) + 197, & \text{if } T_1 = [86, \dots, 143], \text{frac} = 0 \end{cases} \quad (41)$$

ピッチ遅延 T_2 の値は T_1 の値に相対して符号化される。前出の補間と同じ補間を用いて、分数遅延 T_2 はその整数部分 $\text{int}(T_2)$ および分数部分 $\text{frac}/3, \text{frac}=-1,0,1$ によって表現される。これは以下の通りに符号化される。

$$P2 = 3 \times (\text{int}(T_2) - t_{\min}) + \text{frac} + 2 \quad (42)$$

ここで t_{\min} は 3. 7 節同様 T_1 から求められる。

この符号器にランダムビット誤りに対する耐性を持たせるために、第1サブフレームの遅延インデックス $P1$ からパリティビット $P0$ を計算する。このパリティビットは $P1$ の上位6ビットに XOR (排他的論理和) を施したもから生成される。復号器においてこのパリティビットは再計算され、もし再計算された値が転送された値と異なる場合は誤り補償処理が適用される。

3. 7. 3 適応コードブック利得の計算

一度適応コードブック遅延が計算されると、適応コードブック利得 g_p が以下の通りに計算される。

$$g_p = \frac{\sum_{n=0}^{39} x(n)y(n)}{\sum_{n=0}^{39} y(n)y(n)}, \quad \text{bounded by } 0 \leq g_p \leq 1.2 \quad (43)$$

ここで $x(n)$ はターゲット信号であり、 $y(n)$ はフィルタ処理された適応コードブックベクトル ($v(n)$ に対する $W(z)/\hat{A}(z)$ の零状態応答) である。このベクトルは $h(n)$ を用いて $v(n)$ に畳込みを施すことにより得られる。

$$y(n) = \sum_{i=0}^n v(i)h(n-i), \quad n = 0, \dots, 39 \quad (44)$$

3. 8 固定コードブックの構造と探索

固定コードブックは、インタリーブ単一パルス並び換え (ISPP) 設計を用いた代数的コードブック構造に基づいている。このコードブックでは、各コードブックベクトルは4つの零ではないパルスを持つ。各パルスは、振幅が+1または-1で、Table 3-1/JT-G729 で与えられる位置を取り得るものとする。

Table 3 – 1 / JT-G729 Structure of fixed codebook C.
(ITU-T G.729)

<i>Pulse</i>	<i>Sign</i>	<i>Positions</i>
i_0	$s_0: \pm 1$	$m_0: 0,5,10,15,20,25,30,35$
i_1	$s_1: \pm 1$	$m_1: 1,6,11,16,21,26,31,36$
i_2	$s_2: \pm 1$	$m_2: 2,7,12,17,22,27,32,37$
i_3	$s_3: \pm 1$	$m_3: 3,8,13,18,23,28,33,38$ 4,9,14,19,24,29,34,39

コードブックベクトル $c(n)$ は、40次元零ベクトル中の該当する位置に、対応する極性を乗じた4つの単位パルスを置くことにより構築される。

$$c(n) = s_0\delta(n-m_0) + s_1\delta(n-m_1) + s_2\delta(n-m_2) + s_3\delta(n-m_3), \quad n = 0, \dots, 39 \quad (45)$$

ここで $\delta(0)$ は単位パルスである。このコードブック特有な処理として、選択されたコードブックベクトルを適応プリフィルタ $P(z)$ に通すことにより、ピッチの高調波成分が強調され、再生音声の品質が向上する。ここでは、フィルタ

$$P(z) = 1 / (1 - \beta z^{-T}) \quad (46)$$

が使われ、 T は現在のサブフレームのピッチ遅延の整数部分であり、 β はピッチ利得である。 β の値は、直前のサブフレームで量子化された適応コードブック利得を利用することで適応的に作られる。すなわち、

$$\beta = \hat{g}_p^{(m-1)}, \quad \text{bounded by } 0.2 \leq \beta \leq 0.8 \quad (47)$$

40より小さい遅延に対しては、式(45)のコードブック $c(n)$ は以下にしたがって補正される。

$$c(n) = \begin{cases} c(n), & n = 0, \dots, T-1 \\ c(n) + \beta c(n-T), & n = T, \dots, 39 \end{cases} \quad (48)$$

この補正は、次式のようにインパルス応答 $h(n)$ を補正することによって、固定コードブック探索に組み込まれる。

$$h(n) = \begin{cases} h(n), & n = 0, \dots, T-1 \\ h(n) + \beta h(n-T), & n = T, \dots, 39 \end{cases} \quad (49)$$

3. 8. 1 固定コードブック探索手順

固定コードブックは、式(33)の重み付き入力音声 $sw(n)$ と重み付き再生音声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。閉ループピッチ探索で使用されたターゲット信号は、適応コードブック寄与分を減ずることによって更新される。それは、

$$x'(n) = x(n) - g_p y(n), \quad n = 0, \dots, 39 \quad (50)$$

ここで、 $y(n)$ は式(44)のフィルタ処理された適応コードブックベクトルであり、 g_p は式(43)の適応コードブック利得である。

行列 \mathbf{H} は、主対角成分に $h(0)$ を、それより下位の対角成分に $h(1), \dots, h(39)$ を持つ下三角テプリッツ畳込み行列として定義される。行列 $\Phi = \mathbf{H}'\mathbf{H}$ は $h(n)$ の相関を含み、その対称行列の要素は次式のように与えられる。

$$\phi(i, j) = \sum_{n=j}^{39} h(n-i)h(n-j), \quad i = 0, \dots, 39, j = i, \dots, 39 \quad (51)$$

相関信号 $d(n)$ はターゲット信号 $x'(n)$ とインパルス応答 $h(n)$ から次式により得られる。

$$d(n) = \sum_{i=n}^{39} x'(i)h(i-n), \quad n = 0, \dots, 39 \quad (52)$$

\mathbf{c}_k を k 番目の固定コードブックベクトルとするとき、コードブックは次式を最大にする k を求めることによって探索される。

$$\frac{C_k^2}{E_k} = \frac{\left(\sum_{n=0}^{39} d(n)c_k(n) \right)^2}{\mathbf{c}_k' \Phi \mathbf{c}_k} \quad (53)$$

ここで、 t は転置を意味する。

信号 $d(n)$ と行列 Φ はコードブック探索の前に計算される。これは探索処理を高速化するために、 k に依存する要素のみ各 k 毎に計算し、 k に依存しない要素は事前に求めておくということである。

代数的構造を持つコードブック \mathbf{c} は、コードブックベクトル \mathbf{c}_k が 4 つの零でないパルスしか含まないため、高速探索手順が可能である。与えられたベクトル \mathbf{c}_k の式(53)の分子の相関は次式によって与えられる。

$$C = \sum_{i=0}^3 s_i d(m_i) \quad (54)$$

ここで、 m_i は i 番目のパルスの位置、 s_i はその振幅である。式(53)の分母のエネルギーは次式によって与えられる。

$$E = \sum_{i=0}^3 \phi(m_i, m_i) + 2 \sum_{i=0}^2 \sum_{j=i+1}^3 s_i s_j \phi(m_i, m_j) \quad (55)$$

探索手順を簡略化するために、パルス振幅は信号 $d(n)$ を量子化することによって予備選択される。これは、ある位置のパルス振幅を、対応する位置の $d(n)$ の極性と同一にすることで行う。コードブック探索の前に、次のステップで行われる。まず信号 $d(n)$ が、絶対値 $|d(n)|$ と極性 $\text{sign}[d(n)]$ の 2 つに分解される。次に行列 Φ は、極性情報を含むように変形される。すなわち、

$$\phi'(i, j) = \text{sign}[d(i)]\text{sign}[d(j)]\phi(i, j), \quad i = 0, \dots, 39, j = i+1, \dots, 39 \quad (56)$$

Φ の主対角線成分は、式(55)の係数 2 が除かれるようにスケーリングされる。

$$\phi'(i,i) = 0.5\phi(i,i), \quad i = 0, \dots, 39 \quad (57)$$

式(54)の相関は、次式により与えられる。

$$C = |d(m_0)| + |d(m_1)| + |d(m_2)| + |d(m_3)| \quad (58)$$

また式(55)のエネルギーは次式によって与えられる。

$$\begin{aligned} E/2 = & \phi'(m_0, m_0) \\ & + \phi'(m_1, m_1) + \phi'(m_0, m_1) \\ & + \phi'(m_2, m_2) + \phi'(m_0, m_2) + \phi'(m_1, m_2) \\ & + \phi'(m_3, m_3) + \phi'(m_0, m_3) + \phi'(m_1, m_3) + \phi'(m_2, m_3) \end{aligned} \quad (59)$$

探索手順をさらに簡略化するために、絞り込み探索アプローチを使用する。本アプローチでは、最終ループに入る前に、事前に計算されたしきい値と比較され、しきい値を超えた時だけ最終ループに入る。入ることができるループ回数の最大値を固定することにより、探索されるコードブックの割合が低く抑えられる。そのしきい値は相関 C を基に計算される。コードブック探索の前に、最初の 3 パルスを使用して相関値の絶対値の最大値 max_3 と相関値の平均値 av_3 を求める。そのしきい値は次式で与えられる。

$$thr_3 = av_3 + K_3(max_3 - av_3) \quad (60)$$

4 番目のループは、相関値の絶対値（3 パルスによる）が thr_3 を超えた時だけ入る。ただし、 $0 \leq K_3 < 1$ である。 K_3 の値は探索するコードブックの割合を制御し、ここでは 0.4 に設定する。注意することは、これにより探索回数変動し、また更に探索処理を制御するために、2 サブフレームの区間で最終ループに入る回数がある最大値を超えないようにしている点であり、ここでは 180（サブフレームあたりのワーストケースの平均値が 90 回）に設定している。

3. 8. 2 固定コードブック符号語の計算

パルス i_0, i_1, i_2 の位置は、各々 3 ビットで符号化され、 i_3 の位置は 4 ビットで符号化される。各パルス振幅は 1 ビットで符号化される。これにより、4 パルスで合計 17 ビットとなる。パルスの極性が正の時 $s=1$ 、負の時 $s=0$ と定義することで、パルスの極性に関する符号語は、次式により得られる。

$$S = s_0 + 2s_1 + 4s_2 + 8s_3 \quad (61)$$

更に、固定コードブック符号語は、次式により得られる。

$$C = (m_0 / 5) + 8(m_1 / 5) + 64(m_2 / 5) + 512(2(m_3 / 5) + jx) \quad (62)$$

ここで、 $m_3 = 3, 8, \dots, 38$ の時 $jx = 0$ とし、 $m_3 = 4, 9, \dots, 39$ の時 $jx = 1$ とする。

3. 9 利得の量子化

適応コードブック利得（ピッチ利得）と、固定コードブック利得は、7ビットでベクトル量子化される。利得コードブック探索は、以下に示される入力音声と再生音声との間の重み付き平均自乗誤差を最小にすることで行われる。

$$E = \mathbf{x}^t \mathbf{x} + g_p^2 \mathbf{y}^t \mathbf{y} + g_c^2 \mathbf{z}^t \mathbf{z} - 2g_p \mathbf{x}^t \mathbf{y} - 2g_c \mathbf{x}^t \mathbf{z} + 2g_p g_c \mathbf{y}^t \mathbf{z} \quad (63)$$

ここで、 \mathbf{x} は求めるベクトル（3. 6節参照）であり、 \mathbf{y} は式(44)のフィルタ処理後の適応コードブックベクトルであり、 \mathbf{z} は次式のように $h(n)$ に畳込まれた固定コードブックベクトルを示す。

$$z(n) = \sum_{i=0}^n c(i)h(n-i), \quad n = 0, \dots, 39 \quad (64)$$

3. 9. 1 利得予測

固定コードブック利得 g_c は次のように表現することができる。

$$g_c = \gamma g_c' \quad (65)$$

ここで、 g_c' は過去の固定コードブックエネルギーを基に予測された利得であり、 γ は補正係数である。固定コードブック寄与分の平均エネルギーは、次のように与えられる。

$$E = 10 \log \left(\frac{1}{40} \sum_{n=0}^{39} c(n)^2 \right) \quad (66)$$

ベクトル $c(n)$ を固定コードブック利得 g_c でスケーリングした後、スケーリングされた固定コードブックのエネルギーは $20 \log g_c + E$ で与えられる。サブフレーム m における（スケーリングされた）固定コードブック寄与分から平均エネルギーを引いたものを $E^{(m)}$ とすると、次式により与えられる（単位 dB）。

$$E^{(m)} = 20 \log g_c + E - \bar{E} \quad (67)$$

ここで、 $\bar{E} = 30$ dB は固定コードブック励振信号の平均エネルギーである。利得 g_c は、 $E^{(m)}$ 、 E 、 \bar{E} の関数として、次式のように表すことができる。

$$g_c = 10^{(E^{(m)} + \bar{E} - E)/20} \quad (68)$$

予測利得 g_c' は、過去の固定コードブック寄与分の対数エネルギーから現在の固定コードブック寄与分の対数エネルギーを予測することによって求められる。4次のMA予測は次のように行われる。予測エネルギーは、次式により与えられる。

$$\tilde{E}^{(m)} = \sum_{i=1}^4 b_i \hat{U}^{(m-i)} \quad (69)$$

ここで、 $[b_1 b_2 b_3 b_4] = [0.68 0.58 0.34 0.19]$ はMA予測係数であり、 $\hat{U}^{(m)}$ は次式で定義されるサブフレーム m における予測誤差 $U^{(m)}$ を量子化したものである。

$$U^{(m)} = E^{(m)} - \tilde{E}^{(m)} \quad (70)$$

予測利得 g'_c は、式 (68) において $E^{(m)}$ をその予測値に置き換えることにより求められる。

$$g'_c = 10^{(\tilde{E}^{(m)} + \bar{E} - E)/20} \quad (71)$$

補正係数 γ は、利得予測誤差と次のような関係がある。

$$U^{(m)} = E^{(m)} - \tilde{E}^{(m)} = 20 \log(\gamma) \quad (72)$$

3. 9. 2 利得量子化のためのコードブック探索

適応コードブック利得 g_p と係数 γ は、2段の共役構造コードブックを用いてベクトル量子化される。1段目は3ビットの2次元コードブック GA から成り、2段目は4ビットの2次元コードブック GB から成る。各コードブックの1番目の要素は量子化された適応コードブック利得 \hat{g}_p を表し、2番目の要素は量子化された固定コードブック利得補正係数 $\hat{\gamma}$ を表す。コードブック GA と GB に対して、コードブックインデックス GA と GB が与えられた場合、量子化された適応コードブック利得は、

$$\hat{g}_p = \mathbf{GA}_1(\mathbf{GA}) + \mathbf{GB}_1(\mathbf{GB}) \quad (73)$$

により得られ、量子化された固定コードブック利得は、

$$\hat{g}_c = g'_c \hat{\gamma} = g'_c (\mathbf{GA}_2(\mathbf{GA}) + \mathbf{GB}_2(\mathbf{GB})) \quad (74)$$

により得られる。

この共役構造は、予備選択処理を用いることによりコードブック探索を簡単にしている。最適化ピッチ利得 g_p と固定コードブック利得 g_c は、式(63)より導かれ予備選択に用いられる。コードブック GA は8個のベクトルからなり、全般に、 g_c に対応する2番目の要素は、 g_p に対応する1番目の要素より大きな値をとる。この特性が、 g_c の値を用いた予備選択を可能とする。この予備選択処理においては、2番目の要素が g_c に近い4つのベクトルを選択する。同様に、コードブック GB は16個のベクトルからなり、 g_p に対応する1番目の要素の方が2番目の要素より大きいという特性がある。そこで、1番目の要素が g_p に近い8つのベクトルを選択する。こうして各コードブックに対して上位50%の候補ベクトルを選ぶ。選択された $4 \times 8 = 32$ の可能な2つのインデックスの組合せについて全探索を行い、式(63)の重み付き平均自乗誤差を最小にする組合せを求める。

3. 9. 3 利得量子化器のための符号語の計算

利得量子化器の符号語 GA と GB は、最適な選択に対応するインデックスから得られる。単一ビット誤りの影響を減少させるように、コードブックインデックスはマッピングされている。

3. 10 メモリの更新

次のサブフレームにおけるターゲット信号を計算するためには、合成フィルタと重み付けフィルタの状態の更新が必要となる。2つの利得が量子化された後、現在のサブフレームにおける励振信号 $u(n)$ は、

$$u(n) = \hat{g}_p v(n) + \hat{g}_c c(n), \quad n = 0, \dots, 39 \quad (75)$$

により得られる。ここで、 \hat{g}_p と \hat{g}_c は、量子化された適応コードブック利得と量子化された固定コードブック利得である。また、 $v(n)$ は過去の励振信号から補間して求めた適応コードブックベクトルであり、 $c(n)$ は高調波強調を含む固定コードブックベクトルである。フィルタの状態は、40 サンプルのサブフレームに対して、信号 $r(n) - u(n)$ （残差信号と励振信号の差分）をフィルタ $1/\hat{A}(z)$ と $A(z/\gamma_1)/A(z/\gamma_2)$ に通して、フィルタの状態を保存することにより、更新可能である。これは、3つのフィルタ演算を必要とするが、以下に示す1つのフィルタ演算で実現できる。符号器での再生音声 $\hat{s}(n)$ は、励振信号をフィルタ $1/\hat{A}(z)$ に通すことにより計算される。入力が $r(n) - u(n)$ のため、フィルタの出力は、 $e(n) = s(n) - \hat{s}(n)$ と同等である。従って、合成フィルタ $1/\hat{A}(z)$ の状態は、 $e(n), n = 30, \dots, 39$ によって与えられる。フィルタ $A(z/\gamma_1)/A(z/\gamma_2)$ の状態の更新は、聴覚重み付き誤差 $ew(n)$ を得るために誤差信号 $e(n)$ をこのフィルタに通すことにより、可能である。しかしながら、信号 $ew(n)$ は、

$$ew(n) = x(n) - \hat{g}_p y(n) - \hat{g}_c z(n) \quad (76)$$

により、同等に得られる。信号 $x(n)$ 、 $y(n)$ 、 $z(n)$ は得られているため、重み付けフィルタの状態は、 $n = 30, \dots, 39$ について、式(76)に従って $ew(n)$ を計算することにより更新される。これは、2つのフィルタ演算の節約となる。

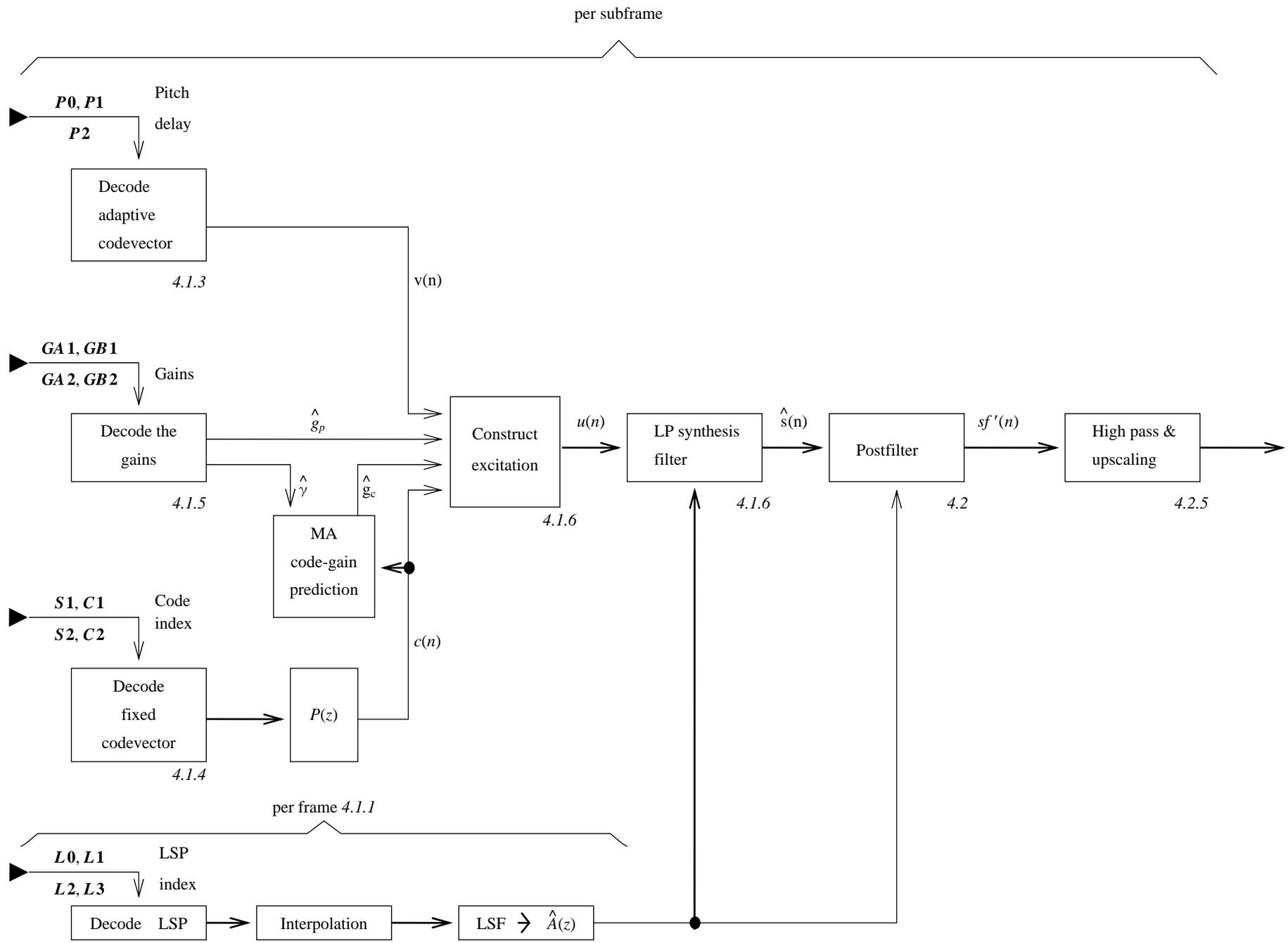
4. 復号器の機能記述

復号器の原理は2章 (Figure 2-3 / JT-G729) に示されている。最初にパラメータを復号する (LP係数、適応コードブックベクトル、固定コードブックベクトルおよび利得)。伝送されるパラメータのリストを Table 4-1 / JT-G729 に示す。これらの復号されたパラメータは、4.1節で述べるように再生音声信号を計算するために使用される。この再生信号はポストフィルタ、高域通過フィルタおよびアップスケーリングにより構成される後処理によって音質が改善される (4.2節)。4.4節では、パリティ誤りが発生した場合またはフレーム消失フラグがセットされている場合に使用される誤り補償処理について述べる。復号器の詳細な信号の流れを Figure 4-1 / JT-G729 に示す。

Table 4-1 / JT-G729 Description of transmitted parameters indices. The bitstream ordering (ITU-T G.729) is reflected by the order in the table. For each parameter the most significant bit (MSB) is transmitted first.

<i>Symbol</i>	<i>Description</i>	<i>Bits</i>
<i>L0</i>	Switched MA predictor index of LSP quantizer	1
<i>L1</i>	First stage vector of LSP quantizer	7
<i>L2</i>	Second stage lower vector of LSP quantizer	5
<i>L3</i>	Second stage higher vector of LSP quantizer	5
<i>P1</i>	Pitch delay 1st subframe	8
<i>P0</i>	Parity bit for pitch delay	1
<i>C1</i>	Fixed codebook 1st subframe	13
<i>S1</i>	Signs of fixed-codebook pulses 1st subframe	4
<i>GA1</i>	Gain codebook (stage 1) 1st subframe	3
<i>GB1</i>	Gain codebook (stage 2) 1st subframe	4
<i>P2</i>	Pitch delay 2nd subframe	5
<i>C2</i>	Fixed codebook 2nd subframe	13
<i>S2</i>	Signs of fixed-codebook pulses 2nd subframe	4
<i>GA2</i>	Gain codebook (stage 1) 2nd subframe	3
<i>GB2</i>	Gain codebook (stage 2) 2nd subframe	4

Figure 4 – 1 / JT-G729 Signal flow at the CS-ACELP decoder.
(ITU-T G.729)



4. 1 パラメータ復号処理

復号処理は次に示す順番で実行される。

4. 1. 1 LPフィルタパラメータの復号

受信したLSP量子化器のインデックス L_0 、 L_1 、 L_2 および L_3 は、量子化されたLSP係数を3. 2. 4節で述べた処理により再生するために使用される。3. 2. 5節で述べた補間処理は、2組の補間されたLSP係数(2つのサブフレームに対応)を得るために使用される。サブフレーム毎に補間されたLSP係数は、サブフレームにおける再生信号を合成するために使用されるLPフィルタ係数 a_i に変換される。

サブフレーム毎に以下のステップが繰り返される。

- (1) 適応コードブックベクトルの復号
- (2) 固定コードブックベクトルの復号
- (3) 適応および固定コードブック利得の復号
- (4) 再生音声信号の計算

4. 1. 2 パリティビットの計算

励振信号が再生される前に、適応コードブックの遅延インデックス P_1 (3. 7. 2節) からパリティビットが計算される。このビットが、伝送されてきたパリティビット P_0 に等しくないならば、伝送中にビット誤りが発生したと見なす。

P_1 にパリティ誤りが発生した場合、遅延値 T_1 には、前のフレームの遅延値 T_2 の整数部分をセットする。遅延値 T_2 は、この新しい値 T_1 を使用して4. 1. 3節に示す手順で算出される。

4. 1. 3 適応コードブックベクトルの復号

パリティ誤りが発生していないならば、受信した適応コードブックのインデックス P_1 は、ピッチ遅延 T_1 の整数部分と分数部分を求めるために使用される。整数部分 $int(T_1)$ と T_1 の分数部分 $frac$ は、以下の通り P_1 から得られる。

```
if  $P_1 < 197$ 
   $int(T_1) = (P_1 + 2) / 3 + 19$ 
   $frac = P_1 - 3 \times int(T_1) + 58$ 
else
   $int(T_1) = P_1 - 112$ 
   $frac = 0$ 
end
```

T_2 の整数部分と分数部分は、 P_2 と t_{min} から得られる。ここで、 t_{min} は T_1 から以下の手順で求められる。

```
 $t_{min} = int(T_1) - 5$ 
if  $t_{min} < 20$  then  $t_{min} = 20$ 
 $t_{max} = t_{min} + 9$ 
if  $t_{max} > 143$  then
   $t_{max} = 143$ 
   $t_{min} = t_{max} - 9$ 
end
```

ここで T_2 は、以下のように復号される。

$$\begin{aligned} \text{int}(T_2) &= (P 2 + 2)/3 - 1 + t_{\min} \\ \text{frac} &= P 2 - 2 - 3 \times ((P 2 + 2)/3 - 1) \end{aligned}$$

適応コードブックベクトル $v(n)$ は式(40)を用いて、過去の励振信号 $u(n)$ のピッチ遅延部分を補間することによって求められる。

4. 1. 4 固定コードブックベクトルの復号

受信した固定コードブックのインデックス C は、励振パルスの位置を抽出するために使用される。パルスの極性は、 S から求められる。これは、3. 8. 2 節の逆の手順により実行される。パルスの位置と極性が復号されると、固定コードブックのベクトル $c(n)$ は式(45)から再生される。ピッチ遅延 T の整数部分がサブフレーム長 40 よりも短い場合、 $c(n)$ は式(48)に基づいて補正される。

4. 1. 5 適応および固定コードブック利得の復号

受信した利得コードブックのインデックスから、適応コードブックの利得 \hat{g}_p および固定コードブックの利得補正係数 $\hat{\gamma}$ が得られる。この手順は 3. 9 節に詳細に示されている。固定コードブックの予測利得 g'_c は式(71)を用いて求められる。固定コードブックの利得は、式(74)でこの予測利得と量子化された利得補正係数の積から得られる。適応コードブックの利得は式(73)を用いて再生される。

4. 1. 6 再生音声信号の計算

式(75)の励振信号 $u(n)$ は LP 合成フィルタに入力される。サブフレームに対する再生音声信号は、以下の式で求められる。

$$\hat{s}(n) = u(n) - \sum_{i=1}^{10} \hat{a}_i \hat{s}(n-i) \quad , \quad n = 0, \dots, 39 \quad (77)$$

ここで \hat{a}_i は、現サブフレームに対する補間された LP フィルタ係数である。この再生音声信号 $\hat{s}(n)$ に次節で説明する後処理を適用する。

4. 2 後処理

後処理は、適応ポストフィルタリング、高域通過フィルタリングと信号アップスケーリングの 3 種の機能からなる。適応ポストフィルタは、長期ポストフィルタ $H_p(z)$ 、短期ポストフィルタ $H_f(z)$ 、傾き補償フィルタ $H_t(z)$ の 3 種類のフィルタを縦続接続したものであり、その後適応利得制御を行う。ポストフィルタ係数は、5ms のサブフレーム毎に更新される。ポストフィルタ処理は、以下のように実行される。最初に残差信号 $\hat{r}(n)$ の生成のため、再生音声信号 $\hat{s}(n)$ が、 $\hat{A}(z/\gamma_n)$ を通して逆フィルタリングされる。この信号は、長期ポストフィルタ $H_p(z)$ の遅延 T と利得 g_l の計算に使用する。次に信号 $\hat{r}(n)$ は、長期ポストフィルタ $H_p(z)$ と合成フィルタ $1/[g_f \hat{A}(z/\gamma_d)]$ を通してフィルタリングされる。最後に、合成フィルタ $1/[g_f \hat{A}(z/\gamma_d)]$ の出力信号は、ポストフィルタリングされた再生音声信号 $sf(n)$ を生成するために傾き補償フィルタ $H_t(z)$ に通される。その後、適応利得制御が $\hat{s}(n)$ のエネルギーに合致するように $sf(n)$ に適用される。利得制御された信号 $sf'(n)$ は、復号器の出力信号を得るために高域通過フィルタに通され、スケールされる。

4. 2. 1 長期ポストフィルタ

長期ポストフィルタは、次式で与えられる。

$$H_p(z) = \frac{1}{1 + \gamma_p g_l z^{-T}} \quad (78)$$

ここで、 T はピッチ遅延であり、 g_l は利得係数である。 g_l は上限が1であり、長期予測利得が3dBより小さければ、0にセットされる。係数 γ_p は、長期ポストフィルタリング量を制御するものであり、 $\gamma_p=0.5$ とする。長期遅延と利得は、短期ポストフィルタ（4. 2. 2節参照）の分子である $\hat{A}(z/\gamma_n)$ による音声信号 $\hat{s}(n)$ のフィルタリングによって得られる残差信号 $\hat{r}(n)$ から計算される。

$$\hat{r}(n) = \hat{s}(n) + \sum_{i=1}^{10} \gamma_n^i \hat{a}_i \hat{s}(n-i) \quad (79)$$

長期遅延は、2段階の手続きで計算される。最初の段階で、 $[\text{int}(T_1)-1, \text{int}(T_1)+1]$ の範囲から最適な整数 T_0 を選択する。ここで、 $\text{int}(T_1)$ は第1サブフレームの（伝送された）ピッチ遅延 T_1 の整数部分とする。最適な整数遅延は次式の相関を最大にする。

$$R(k) = \sum_{n=0}^{39} \hat{r}(n) \hat{r}(n-k) \quad (80)$$

第2段階では、 T_0 の周りで1/8の分解能を持つ最適な分数遅延 T を選択する。これは、擬似正規化相関最大値を与える遅延を見つけることで実行される。

$$R'(k) = \frac{\sum_{n=0}^{39} \hat{r}(n) \hat{r}_k(n)}{\sqrt{\sum_{n=0}^{39} \hat{r}_k(n) \hat{r}_k(n)}} \quad (81)$$

ここで、 $\hat{r}_k(n)$ は遅延 k における残差信号である。一旦最適な遅延 T が見つかり、対応する相関 $R'(T)$ は、 $\hat{r}(n)$ のエネルギーの平方根で正規化される。この正規化相関の平方根値が長期ポストフィルタを無効にすべきかを決定するために使用される。つまり、次式を満たした場合、 $g_l = 0$ とし、

$$\frac{R'(T)^2}{\sum_{n=0}^{39} \hat{r}(n) \hat{r}(n)} < 0.5 \quad (82)$$

上式を満たさない場合、 g_l の値は式(83)から計算される。

$$g_l = \frac{\sum_{n=0}^{39} \hat{r}(n) \hat{r}_k(n)}{\sum_{n=0}^{39} \hat{r}_k(n) \hat{r}_k(n)} \quad \text{bounded by } 0 \leq g_l \leq 1.0 \quad (83)$$

非整数遅延信号 $\hat{r}_k(n)$ は、最初長さ 33 の補間フィルタを用いて計算される。 T の選択後、 $\hat{r}_k(n)$ は長さ 129 のより長い補間フィルタを用いて再計算される。長い補間フィルタが $R'(T)$ の値を増加させる場合のみ、再計算後の信号を再計算前の信号に置き換える。

4. 2. 2 短期ポストフィルタ

短期ポストフィルタは、次式で与えられる。

$$H_f(z) = \frac{1}{g_f} \frac{\hat{A}(z/\gamma_n)}{\hat{A}(z/\gamma_d)} = \frac{1}{g_f} \frac{1 + \sum_{i=1}^{10} \gamma_n^i \hat{a}_i z^{-i}}{1 + \sum_{i=1}^{10} \gamma_d^i \hat{a}_i z^{-i}} \quad (84)$$

ここで、 $\hat{A}(z)$ は、受信した量子化 LP 逆フィルタ（復号器では LP 分析は実行されない）である。 γ_n と γ_d は短期ポストフィルタリング量を制御する係数であり、 $\gamma_n = 0.55$ 、 $\gamma_d = 0.7$ にセットされる。利得項 g_f は、フィルタ $\hat{A}(z/\gamma_n)/\hat{A}(z/\gamma_d)$ の打ち切られたインパルス応答 $h_f(n)$ で計算され、次式で与えられる。

$$g_f = \sum_{n=0}^{19} |h_f(n)| \quad (85)$$

4. 2. 3 傾き補償

フィルタ $H_i(z)$ は、短期ポストフィルタ $H_f(z)$ の傾きに対して補償し、次式で与えられる。

$$H_i(z) = \frac{1}{g_i} (1 + \gamma_i k'_i z^{-1}) \quad (86)$$

ここで、 $\gamma_i k'_i$ は傾き係数で、 k'_i は次式の $h_f(n)$ から計算される 1 次反射係数である。

$$k'_i = -\frac{\gamma_h(1)}{\gamma_h(0)}, \quad \gamma_h(i) = \sum_{j=0}^{19-i} h_f(j) h_f(j+i) \quad (87)$$

利得項 $g_i = 1 - |\gamma_i k'_i|$ は、 $H_f(z)$ において g_f の減少効果に対して補償している。積形式のフィルタ $H_f(z)H_i(z)$ は一般に利得がないことが示されている。 γ_i は、 k'_i の符号に依存して 2 つの値が与えられる。 k'_i が負の場合 $\gamma_i = 0.9$ で、 k'_i が正の場合 $\gamma_i = 0.2$ となる。

4. 2. 4 適応利得制御

適応利得制御は、再生音声信号 $\hat{s}(n)$ とポストフィルタリングされた信号 $sf(n)$ の利得差を補償するために使用される。現サブフレームに対する利得調整係数 G は、次式で計算される。

$$G = \frac{\sum_{n=0}^{39} |\hat{s}(n)|}{\sum_{n=0}^{39} |sf(n)|} \quad (88)$$

利得調整されポストフィルタリングされた信号 $sf'(n)$ は次式で与えられる。

$$sf'(n) = g^{(n)} sf(n) \quad n=0, \dots, 39 \quad (89)$$

ここで、 $g^{(n)}$ はサンプル毎に更新され、次式で与えられる。

$$g^{(n)} = 0.85g^{(n-1)} + 0.15G, \quad n=0, \dots, 39 \quad (90)$$

$g^{(-1)}$ の初期値は 1.0 が使用される。そして、新しいサブフレーム毎に $g^{(-1)}$ は、前サブフレームの $g^{(39)}$ と等しい値に設定される。

4. 2. 5 高域通過フィルタとアップスケーリング

カットオフ周波数 100Hz の高域通過フィルタは、ポストフィルタ処理された再生音声 $sf'(n)$ に適用される。このフィルタは、以下のように表される。

$$H_{h2}(z) = \frac{0.93980581 - 1.8795834z^{-1} + 0.93980581z^{-2}}{1 - 1.9330735z^{-1} + 0.93589199z^{-2}} \quad (91)$$

フィルタ処理された信号は、もとの入力信号レベルに戻すために 2 倍される。

4. 3 符号器と復号器の初期化

Table 4 – 2 / JT-G729 に示す変数を除き符号器と復号器の全ての静的変数は零に初期化される。

Table 4 – 2 / JT-G729 Description of parameters with nonzero initialization.
(ITU-T G.729)

Variable	Reference	Initial value
β	Section 3.8	0.8
$g^{(-1)}$	Section 4.2.4	1.0
\hat{l}_i	Section 3.2.4	$i\pi/11$
q_i	Section 3.2.4	$\arccos(i\pi/11)$
$\hat{U}^{(k)}$	Section 3.9.1	-14

4. 4 フレーム消失補償

誤り補償処理は、ビット列におけるフレーム消失によって起こる再生音声品質の劣化を緩和するために復号器の中に組み込まれる。この誤り補償処理は、符号化パラメータフレーム（10ms フレーム相当）の消失を検出した時に機能する。消失フレームを検出する機能は、アプリケーションに依存するため本標準では定義しない。

失われた信号の補償のためには、過去に受信した情報を基にして、現在のフレームを復元せねばならない。ここでは、失われた励振信号を同じような特徴を持つ励振信号に置き換えると同時に、そのエネルギーを徐々に減衰させる方法を用いる。これは、長期ポストフィルタ分析の 1 パラメータとして計算される長期予測利得を基にした有声/無声識別器を用いることによって実現される。長期ポストフィルタ（4. 2. 1 節参照）

は、予測利得が 3dB 以上の長期予測器を見つけだす。これは、式(82)で示すように、正規化された（自己）相関関数の自乗の値についてしきい値を 0.5 に設定することで実現される。誤り補償の処理では、3dB 以上の長期予測利得を持つ 5ms のサブフレームが少なくとも 1 つ存在するならば、その 10ms フレームは、周期的フレームと設定する。その他のフレームは、非周期的フレームと設定する。消失フレームにおいては、その設定の直前の（再生された）音声フレームの設定を継承する。なお、有声／無声の識別結果は、この再生された音声信号を基に常に更新される。

消失フレームに対して実行される特別な処理を以下に示す。

- (1) 合成フィルタパラメータの反復
- (2) 適応および固定コードブック利得の減衰
- (3) 利得予測器のメモリの減衰
- (4) 置換用励振信号の生成

4. 4. 1 合成フィルタパラメータの反復

消失フレームにおける合成フィルタでは、直前の正常フレームの LP パラメータを使用する。MA LSF 予測器のメモリは、受信した符号語 \hat{l}_i の値を含んでいる。現在のフレーム m における符号語は失われているので、これを過去の LSF パラメータ $\hat{\omega}_i$ と予測器メモリから以下の式を用いて求める。

$$\hat{l}_i = \left[\hat{\omega}_i^{(m)} - \sum_{k=1}^4 \hat{p}_{i,k} \hat{l}_i^{(m-k)} \right] / \left(1 - \sum_{k=1}^4 \hat{p}_{i,k} \right), \quad i=1, \dots, 10 \quad (92)$$

ここで、MA 予測係数 $\hat{p}_{i,k}$ は、直前の正常フレームにおける係数である。

4. 4. 2 適応および固定コードブック利得の減衰

固定コードブック利得は、過去の固定コードブック利得を減衰させたものを基にして次式により求める。

$$g_c^{(m)} = 0.98 g_c^{(m-1)} \quad (93)$$

ここで、 m はサブフレーム番号である。適応コードブック利得は、過去の適応コードブック利得を減衰させたものを基にして、次式により求める。

$$g_p^{(m)} = 0.9 g_p^{(m-1)}, \quad \text{bounded by } g_p^{(m)} < 0.9 \quad (94)$$

4. 4. 3 利得予測器のメモリの減衰

3. 9 節で述べたように、利得予測器は前もって選択された固定コードブックベクトル $c(n)$ のエネルギーを用いる。復号器での過渡的影響を避けるため、一度正常なフレームを受信したら、利得予測器のメモリをコードブックエネルギーの減衰値で更新する。現在のサブフレーム m に対する $\hat{U}^{(m)}$ は量子化された利得予測誤差の平均値を 4dB 減衰させた値に設定される。

$$\hat{U}^{(m)} = \left(0.25 \sum_{i=1}^4 \hat{U}^{(m-i)} \right) - 4.0, \quad \text{bounded by } \hat{U}^{(m)} \geq -14 \quad (95)$$

4. 4. 4 置換用励振信号の生成

使用される励振信号は周期性に依存する。もし直前に再生したフレームが周期的であれば、現在のフレームも同様に周期的であると考えられる。この場合、適応コードブックのみが用いられ、固定コードブック寄与分は零に設定される。ピッチ遅延は前のフレームのピッチ遅延の整数部分の値をとり、各連続するフレーム毎に繰り返される。過度の周期性を避けるため、遅延を次のサブフレーム毎に1ずつ、上限143まで、増大させる。適応コードブック利得は式(94)に示した減衰値をとる。

もし直前に再生したフレームが周期的でないとして識別された場合は、現在のフレームも同様に周期的でないと考えられ、適応コードブック寄与分を零にする。固定コードブック寄与分は、ランダムに選択した値をコードブックインデックスおよび極性インデックスとして使用する。乱数発生器は次の式に基づく。

$$seed = seed * 31821 + 13849 \quad (96)$$

ここで *seed* の初期値は 21845 である。固定コードブックインデックスは、次の乱数の下位 13 ビットを用いる。固定コードブックの極性は、次の乱数の下位 4 ビットを用いる。固定コードブック利得は式(93)で示した減衰値を用いる。

5. CS-ACELPコーデックのビットイグザクトな記述

CS-ACELPコーデックを16ビット固定小数点でシミュレーションしたANSI Cのソフトウェアが新日本ITU協会から入手可能である。以下の節ではシミュレーションソフトウェアの使い方や構成について概要を述べる。

5. 1 シミュレーションソフトウェアの使い方

C言語のシミュレーションソフトウェアは、符号器をシミュレーションした `coder.c` と、復号器をシミュレーションした `decoder.c` の2つのメインプログラムで構成されている。符号器は次のように実行させる。

```
coder inputfile bitstreamfile
```

入力ファイルと出力ファイルは16ビットPCM信号で構成されたサンプリングデータファイルである。ビット列ファイルは、81の16ビットワードのデータで構成され、最初のワードがフレーム消失を示し、残りの80ワードがそれぞれ1ビットのデータを含んでいる。復号器はビット列ファイルを入力とし、後処理された16ビットPCM信号で構成された出力ファイルを生成する。復号器は次のように実行させる。

```
decoder bitstreamfile outputfile
```

5. 2 シミュレーションソフトウェアの構成

固定小数点のANSI Cのシミュレーションでは、Table 5-1/JT-G729で示される2つの固定小数点データ形式のみを用いる。シミュレーションソフトウェアを実行するのを容易にするため、ループインデックス、ブール変数およびフラグは、Flag型の変数を用いる。Flag型変数のビット長が16ビットか32ビットになるかは、動作環境に依存する。

Table 5-1/JT-G729 Data Types used in ANSI C simulation.
(ITU-T G.729)

Type	Max. value	Min. value	Description
Word16	0x7fff	0x8000	Signed 2's complement 16 bit word
Word32	0x7fffffffL	0x80000000L	Signed 2's complement 32 bit word

全ての計算は予め定義された基本演算を用いて行われる。これらの基本演算を Table 5-2/JT-G729 に示す。シミュレーションソフトウェアで用いられる各種テーブルの一覧を Table 5-3/JT-G729 に示す。これらのメインプログラムはサブルーチンのライブラリを使用している。このライブラリの一覧を Table 5-4/JT-G729、5-5/JT-G729、5-6/JT-G729 に示す。

Table 5 – 2 / JT-G729 Basic operations used in ANSI C simulation.
(ITU-T G.729)

<i>Operation</i>	<i>Description</i>
Word16 satire(Word32 L_var1)	Limit to 16bits
Word16 add(Word16 var1, Word16 var2)	Short addition
Word16 sub(Word16 var1, Word16 var2)	Short subtraction
Word16 abs_s(Word16 var1)	Short absolute value
Word16 shl(Word16 var1, Word16 var2)	Short shift left
Word16 shr(Word16 var1, Word16 var2)	Short shift right
Word16 mult(Word16 var1, Word16 var2)	Short multiplication
Word32 L_mult(Word16 var1, Word16 var2)	Long multiplication
Word16 negate(Word16 var1)	Short negate
Word16 extract_h(Word32 L_var1)	Extract high
Word16 extract_l(Word32 L_var1)	Extract low
Word16 round(Word32 L_var1)	Round
Word32 L_mac(Word32 L_var3, Word16 var1, Word16 var2)	Multiply and accumulate
Word32 L_msu(Word32 L_var3, Word16 var1, Word16 var2)	Multiply and subtract
Word32 L_add(Word32 L_var1, Word32 L_var2)	Long addition
Word32 L_sub(Word32 L_var1, Word32 L_var2)	Long subtraction
Word32 L_negate(Word32 L_var1)	Long negate
Word16 mult_r(Word16 var1, Word16 var2)	Multiplication with rounding
Word32 L_shl(Word32 L_var1, Word16 var2)	Long shift left
Word32 L_shr(Word32 L_var1, Word16 var2)	Long shift right
Word16 shr_r(Word16 var1, Word16 var2)	Shift right with rounding
Word16 mac_r(Word32 L_var3, Word16 var1, Word16 var2)	Mac with rounding
Word16 msu_r(Word32 L_var3, Word16 var1, Word16 var2)	Msu with rounding
Word32 L_deposit_h(Word16 var1)	16 bit var1 into MSB part
Word32 L_deposit_l(Word16 var1)	16 bit var1 into LSB part
Word32 L_shr_r(Word32 L_var1, Word16 var2)	Long shift right with round
Word32 L_abs(Word32 L_var1)	Long absolute value
Word16 norm_s(Word16 var1)	Short norm
Word16 div_s(Word16 var1, Word16 var2)	Short division
Word16 norm_l(Word32 L_var1)	Long norm

Table 5 – 3 / JT-G729 Summary of tables.
(ITU-T G.729)

<i>File</i>	<i>Table name</i>	<i>Size</i>	<i>Description</i>
tab_hup.c	tab_hup_s	28	Upsampling filter for postfilter
tab_hup.c	tab_hup_l	112	Upsampling filter for postfilter
inter_3.c	inter_3	13	FIR filter for interpolating the correlation
pred_lt3.c	inter_3	31	FIR filter for interpolating past excitation
lspcb.tab	lspcb1	128 × 10	LSP quantizer (first stage)
lspcb.tab	lspcb2	32 × 10	LSP quantizer (second stage)
lspcb.tab	fg	2 × 4 × 10	MA predictors in LSP VQ
lspcb.tab	fg_sum	2 × 10	Used in LSP VQ
lspcb.tab	fg_sum_inv	2 × 10	Used in LSP VQ
qua_gain.tab	gbk1	8 × 2	Codebook GA in gain VQ
qua_gain.tab	gbk2	16 × 2	Codebook GB in gain VQ
qua_gain.tab	map1	8	Used in gain VQ
qua_gain.tab	imap1	8	Used in gain VQ
qua_gain.tab	map2	16	Used in gain VQ
qua_gain.tab	imap2	16	Used in gain VQ
window.tab	window	240	LP analysis window
lag_wind.tab	lag_h	10	Lag window for bandwidth expansion (high part)
lag_wind.tab	lag_l	10	Lag window for bandwidth expansion (low part)
grid.tab	grid	61	Grid points in LP to LSP conversion
inv_sqrt.tab	tabsqr	49	Lookup table in inverse square root computation
log2.tab	tablog	33	Lookup table in base 2 logarithm computation
lsp_lsf.tab	table	65	Lookup table in LSF to LSP conversion and vice versa
lsp_lsf.tab	slope	64	Line slopes in LSP to LSF conversion
pow2.tab	tabpow	33	Lookup table in 2 ^x computation
acelp.h			Prototypes for fixed codebook search
ld8k.h			Prototypes and constants
typedef.h			Type definitions

Table 5 – 4 / JT-G729 Summary of encoder specific routines.
(ITU-T G.729)

<i>Filename</i>	<i>Description</i>
acelp_co.c	Search fixed codebook
cod_ld8k.c	Encoder routine
lpc.c	LP analysis
pitch.c	Pitch search
pre_proc.c	Pre-processing (HP filtering and scaling)
pwf.c	Computation of perceptual weighting coefficients
qua_gain.c	Gain quantizer
qua_lsp.c	LSP quantizer

Table 5 – 5 / JT-G729 Summary of decoder specific routines.
(ITU-T G.729)

<i>Filename</i>	<i>Description</i>
de_acelp.c	Decode algebraic codebook
dec_gain.c	Decode gains
dec_lag3.c	Decode adaptive codebook index
dec_ld8k.c	Decoder routine
lspdec.c	LSP decoding routine
post_pro.c	Post processing(HP filtering and scaling)
pst.c	Postfilter routines

Table 5 – 6 / JT-G729 Summary of general routines.
(ITU-T G.729)

<i>Filename</i>	<i>Description</i>
basicop2.c	Basic operators
oper_32b.c	Extended basic operators
bits.c	Bit manipulation routines
dspfunc.c	Mathematical functions
filter.c	Filter functions
gainpred.c	Gain predictor
lpcfnc.c	Miscellaneous routines related to LP filter
lspgetq.c	LSP quantizer
p_parity.c	Compute pitch parity
pred_lt3.c	Generation of adaptive codebook
util.c	Utility functions

付属資料 A

(標準 J T - G 7 2 9 に対する)

低演算量版 8kbit/s C S - A C E L P 音声コーデック

(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、標準 J T - G 7 2 9 音声コーデックの低演算量版の概要記述である。本コーデックは標準 J T - G 7 2 9 本体と相互接続可能である。すなわち、低演算量版の符号器は標準 J T - G 7 2 9 本体の復号器とともに利用でき、逆もまた同じである。しかしながら、本付属資料で規定されているコーデックを実現する場合、ある状況下で本コーデックの性能が標準 J T - G 7 2 9 本体とは同等でないことを認識しておく必要がある。

低演算量版コーデックは、音声とデータを同時に利用するマルチメディア・アプリケーションのために開発されたものであるが、このアプリケーションに限定されるものではない。

コーデックの説明は標準 J T - G 7 2 9 本体に基づいている。本付属資料は、標準 J T - G 7 2 9 本体に対して、コーデックのアルゴリズムの複雑さを軽減するために必要な変更を記述している。アルゴリズムの変更の無い部分については、標準本体の該当する節が参照される。

A. 2 コーデックの概要説明

符号化/復号アルゴリズムの概要説明は標準 J T - G 7 2 9 本体に基づいている。ビット割り当ては Table 2 - 1 / J T - G 7 2 9 に示されているものと同一である。遅延も同一である (10ms の音声フレームと 5ms の先読み)。標準 J T - G 7 2 9 本体に対する主なアルゴリズム変更を以下に示す。

- (1) 聴覚重み付けフィルタには量子化された L P フィルタパラメータを使用する。すなわち、 $\gamma = 0.75$ の固定値を用い、 $W(z) = \hat{A}(z) / \hat{A}(z/\gamma)$ としている。
- (2) 開ループピッチ分析の簡略化のために、重み付き音声の関連の計算中に間引きを利用する。
- (3) $W(z) / \hat{A}(z)$ を $1 / \hat{A}(z/\gamma)$ に簡略化することにより、重み付き合成フィルタ $W(z) / \hat{A}(z)$ のインパルス応答の計算、ターゲット信号の計算、フィルタ状態の更新を簡略化する。
- (4) 適応コードブック探索を簡略化する。過去の励振信号と後ろ向きにフィルタ処理されたターゲット信号の関連の最大値を探索する (フィルタ処理された過去の励振信号のエネルギーは考慮しない)。
- (5) 固定代数的コードブック探索を簡略化する。入れ子状ループの絞り込み探索アプローチの代わりに反復的デプス・ファースト木探索アプローチを使用する。
- (6) 復号器において、整数遅延のみを用いることにより高調波ポストフィルタを簡略化する。

上記の変更について、より詳細な記述を A. 3 章と A. 4 章に示す。

A. 2. 1 音声コーデックの記述

低演算量版音声コーデックの記述はビットイグザクトな固定小数点演算で表現されている。A. 5 章に示す本付属資料で必須な A N S I C コードはこのビットイグザクトな固定小数点記述法を用いている。符号器 (A. 3 章) と復号器 (A. 4 章) の数学的な記述は他の方式でも実現可能であるが、その様な方法で実現されたコーデックは本付属資料に準拠しない可能性がある。よって、相違点がある場合、A. 5 章の A N S I C コードによるアルゴリズム記述は、A. 3 章と A. 4 章における数学的な記述より優先される。十分なものではないが、この A N S I C コードとともに利用できるテスト信号のセットは、新日本 I T U 協会から入手可能である。

A. 2. 2 記法

記法は標準 J T - G 7 2 9 の 2. 5 節と同一である。

A. 3 符号器の機能記述

本章では、Figure 3 - 1 / JT-G729 のブロックに示す符号器の変更された機能について記述する。アルゴリズムの簡略化が実行された部分を除き標準本体が本章の大部分で参照される。

A. 3. 1 前処理

3. 1 節と同一である。

A. 3. 2 線形予測分析と量子化

A. 3. 2. 1 窓かけと自己相関計算

3. 2. 1 節と同一である。

A. 3. 2. 2 レビンソン・ダービンアルゴリズム

3. 2. 2 節と同一である。

A. 3. 2. 3 LP から LSP への変換

いくつかの簡略化を除いて 3. 2. 3 節と同一である。多項式 $F_1(z)$ と $F_2(z)$ を評価するポイント数が (60 に代わり) 50 に低減され、多項式のより良い根を求めるために、極性が変化している区間の分割が 4 回に代わり 2 回行われる。

A. 3. 2. 4 LSP 係数の量子化

3. 2. 4 節と同一である。

A. 3. 2. 5 LSP 係数の補間

簡略化のため重み付けフィルタには量子化されたパラメータを使用する。量子化された LP 係数のみが補間される点を除いて、3. 2. 5 節と同一である。

A. 3. 2. 6 LSP から LP への変換

3. 2. 6 節と同一である。

A. 3. 3 聴覚重み付け

3. 3 節と異なり、聴覚重み付けフィルタは量子化された LP フィルタ係数 \hat{a}_i に基づいており以下の式で与えられる。

$$W(z) = \frac{\hat{A}(z)}{\hat{A}(z/\gamma)} \quad (\text{A.1})$$

ここで、 $\gamma = 0.75$ である。これは合成フィルタと重み付けフィルタの結合を $W(z)/A(z) = 1/\hat{A}(z/\gamma)$ に簡略化し、インパルス応答とターゲット信号の計算やフィルタ状態の更新で用いられるフィルタ処理の演算回数を低減する。 γ の値が 0.75 に固定されていることと、3. 3 節で述べられている聴覚重み付けフィルタの係数の適応処理はこの低演算量版では使用されないことに注意しなければならない。

重み付き音声信号は別の方法を用いるため、ターゲット信号の計算に使用されない (A. 3. 6 節参照)。しかしながら、重み付き音声信号 (低域通過フィルタリングしたもの) は、開ループピッチ推定の計算に使用される。低域通過フィルタリングされた重み付き音声信号は、フィルタ $\hat{A}(z)/[\hat{A}(z/\gamma)(1-0.7z^{-1})]$ に音声信号 $s(n)$ を通すことにより得られる。最初にフィルタ $A'(z) = \hat{A}(z/\gamma)(1-0.7z^{-1})$ の係数が計算され、次にサブフレームの低域通過フィルタリングされた重み付き音声信号が以下の式で計算される。

$$s_w(n) = r(n) - \sum_{i=1}^{10} a'_i s_w(n-i), \quad n = 0, \dots, 39 \quad (\text{A.2})$$

ここで、 $r(n)$ は次式で与えられる LP 残差信号である。

$$r(n) = s(n) + \sum_{i=1}^{10} \hat{a}_i s(n-i), \quad n = 0, \dots, 39 \quad (\text{A.3})$$

信号 $s_w(n)$ は音声フレームのピッチ遅延の推定のために使用される。

A. 3. 4 開ループピッチ分析

最適適応コードブック遅延に対する探索の演算量を低減するため、探索範囲は遅延候補 T_{op} の近傍に限定される。 T_{op} は開ループピッチ分析によって得られる。開ループピッチ分析はフレーム(10ms)毎に行われる。開ループピッチの算出は、式(A.2)の低域通過フィルタリングされた重み付き音声信号 $s_w(n)$ を用いて以下のように行なわれる。まず最初に、相関関数

$$R(k) = \sum_{n=0}^{39} s_w(2n) s_w(2n-k) \quad (\text{A.4})$$

の最大値を以下の3つの区間についてそれぞれ求める。

$$\begin{aligned} i = 1: & 20, \dots, 39 \\ i = 2: & 40, \dots, 79 \\ i = 3: & 80, \dots, 143 \end{aligned}$$

得られた最大値 $R(t_i), i = 1, \dots, 3$ は以下の式により正規化される。

$$R'(t_i) = \frac{R(t_i)}{\sqrt{\sum_{n=0}^{39} s_w^2(2n-t_i)}}, \quad i = 1, \dots, 3 \quad (\text{A.5})$$

3つの正規化候補の中から最終候補が選択される。ここでは、短い遅延区間の遅延が長い遅延区間の遅延の約数ならば短い遅延区間に対応する正規化相関を大きくすることによって、短い遅延区間の値が選ばれやすくなるようにしている。

ここで、次の点に注意すること。式(A.4)の相関関数の計算では偶数サンプルのみが用いられる。ただし、3番目の遅延区間[80,143]では最初に偶数遅延で相関関数を計算し、そして次にこの偶数遅延の ± 1 の遅延の値を求める。

A. 3. 5 インパルス応答の計算

重み付き合成フィルタ $W(z)/\hat{A}(z)$ のインパルス応答 $h(n)$ は、適応コードブックと固定コードブックの探索に必要となる。インパルス応答 $h(n)$ は、単位サンプルと零詰めされたサンプルから成る信号をフィルタ $1/\hat{A}(z/\gamma)$ に通すことにより、サブフレーム毎に算出される。

A. 3. 6 ターゲット信号の計算

適応コードブック探索のターゲット信号 $x(n)$ は、LP 残差信号 $r(n)$ を重み付き合成フィルタ $1/\hat{A}(z/\gamma)$ に通すことにより算出される。サブフレームの励振信号が決定された後、これらのフィルタの初期状態は A. 3. 10 節で説明されるように更新される。

ターゲットベクトル算出時に必要となる残差信号 $r(n)$ は、適応コードブック探索時にも過去の励振信号のバッファを拡張するために用いられる。LP 残差信号の計算は式(A.3)で与えられる。

A. 3. 7 適応コードブック探索

適応コードブックの構造は 3. 7 節と同様である。第 1 サブフレームにおいて、分数ピッチ遅延 T_1 は $\left[19\frac{1}{3}, 84\frac{2}{3}\right]$ の範囲内では 1/3 の分析精度で、 $[85, 143]$ の範囲内では整数のみが使用される。第 2 サブフレームにおいては、遅延 T_2 は常に $\left[\text{int}(T_1) - 5\frac{2}{3}, \text{int}(T_1) + 4\frac{2}{3}\right]$ の範囲内で 1/3 の分析精度で使用される。ここで $\text{int}(T_1)$ は第 1 サブフレームの分数ピッチ遅延 T_1 の整数部分である。この範囲は T_1 が含まれる遅延範囲の境界に適応している。

両サブフレームの探索の境界値 t_{\min} と t_{\max} は 3. 7 節と同様の範囲で決定される。閉ループピッチ探索は以下の項を最大にすることによって達成される。

$$R(k) = \frac{\sum_{n=0}^{39} x(n)y_k(n)}{\sqrt{\sum_{n=0}^{39} y_k(n)y_k(n)}} \quad (\text{A.6})$$

ここで $x(n)$ はターゲット信号であり、 $y_k(n)$ は遅延 k でフィルタ処理された過去の励振信号 ($h(n)$ で畳み込まれた過去の励振信号) である。本低演算量版では探索を簡略化するために、式(A.6)の分子のみの最大化を行う。すなわち以下の項を最大にする。

$$R_N(k) = \sum_{n=0}^{39} x(n)y_k(n) = \sum_{n=0}^{39} x_b(n)u_k(n) \quad (\text{A.7})$$

ここで $x_b(n)$ は後ろ向きにフィルタ処理されたターゲット信号 ($x(n)$ とインパルス応答 $h(n)$ の相関) であり、 $u_k(n)$ は遅延 k における過去の励振信号であり、 $u(n-k)$ を示す。探索範囲は予備選択された第 1 サブフレームの開ループピッチ T_{op} および第 2 サブフレームの T_1 の近傍に限定されることに注意しなければならない。

探索中にはサンプル $u(n)$, $n=0, \dots, 39$ は未知であることに注意しなければならない。またそれらの値は 40 以下のピッチ遅延のために必要である。LP 残差は $u(n)$ にコピーすることにより、探索を簡略化させる。

もし最適な整数ピッチ遅延が 85 以下のときは、 T_2 および T_1 の値を決定するために、最適な整数遅延の近傍の分数値をテストしなければならない。分数ピッチ探索は分数 $-1/3, 0$, および $1/3$ における過去の励振信号を補間し、式(A.7)の相関を最大にする分数を選択することにより行われる。過去の励振信号の補間は 3.

7節での定義と同様にFIRフィルタ b_{30} を用いて行われる。与えられた整数遅延 k と分数 t に対する補間された過去の励振信号は次式により与えられる。

$$u_{kt}(n) = \sum_{i=0}^9 u(n-k-i)b_{30}(t+3i) + \sum_{i=0}^9 u(n-k+1+i)b_{30}(3-t+3i), \quad n=0, \dots, 39, \quad t=0, 1, 2 \quad (\text{A.8})$$

A. 3. 7. 1 適応コードブックベクトルの生成

3. 7. 1節と同一である。

A. 3. 7. 2 適応コードブック遅延に対する符号語の計算

3. 7. 2節と同一である。

A. 3. 7. 3 適応コードブック利得の計算

3. 7. 3節と同一である。

A. 3. 8 固定コードブックの構造と探索

17ビットの代数的コードブックの構造は3. 8節と同一である。

A. 3. 8. 1 固定コードブック探索手順

パルスの極性は3. 8. 1節における説明と同様のアプローチを用いて求められる。しかし、パルスの位置についてはより効率的なアプローチで求められる。入れ子状ループ探索アプローチの代わりに、反復的デプス・ファースト木探索アプローチが用いられる。この新しいアプローチにより、探索するパルス位置の組み合わせ数がより少なくなり、演算量も一定となる。

A. 3. 8. 2 固定コードブック符号語の計算

3. 8. 2節と同一である。

A. 3. 9 利得の量子化

3. 9節と同一である。

A. 3. 10 メモリの更新

次のサブフレームにおけるターゲット信号を計算するためには、重み付き合成フィルタの状態の更新が必要となる。2つの利得が量子化された後、現在のサブフレームにおける励振信号 $u(n)$ は、

$$u(n) = \hat{g}_p v(n) + \hat{g}_c c(n), \quad n=0, \dots, 39 \quad (\text{A.9})$$

により得られる。ここで、 \hat{g}_p と \hat{g}_c は、それぞれ量子化された適応コードブック利得と量子化された固定コードブック利得である。また、 $v(n)$ は（過去の励振信号から補間して求めた）適応コードブックベクトルであり、 $c(n)$ は高調波強調を含む固定コードブックベクトルである。重み付き合成フィルタの状態は、40サンプルのサブフレームに対して、信号 $r(n) - u(n)$ （残差信号と励振信号の差分）をフィルタ $1/\hat{A}(z/\gamma)$ に通して、フィルタの状態を保存することにより、更新可能である。フィルタ演算を必要としない簡略法を次に示す。入力 $r(n) - u(n)$ によるフィルタの出力は重み付き誤差信号 $e_w(n)$ であり、これは、

$$e_w(n) = x(n) - \hat{g}_p y(n) - \hat{g}_c z(n) \quad (\text{A.10})$$

により得られる。ここで、 $x(n)$ はターゲット信号であり、 $y(n)$ と $z(n)$ は、それぞれフィルタ通過後の適応コードブックベクトルと固定コードブックベクトルである。信号 $x(n)$ 、 $y(n)$ 、 $z(n)$ は得られているため、重み付き合成フィルタの状態は、 $n=30, \dots, 39$ について、式(A.10)に従って $e_w(n)$ を計算することにより更新される。

A. 4 復号器の機能記述

復号器の原理は、Figure 2-3/JT-G729 に示されている。伝送されるパラメータは、Table 4-1/JT-G729 に示されたものと同じである。復号されたパラメータは、再生音声信号を計算するために用いられる。この再生信号は、ポストフィルタ、高域通過フィルタ、およびアップスケーリングにより構成される後処理によって音質が改善される（A. 4. 2 節）。復号器の詳細な信号の流れを、Figure 4-1/JT-G729 に示す。

復号器における唯一の違いは、A. 4. 2 節で述べるポストフィルタである。

A. 4. 1 パラメータ復号処理

4. 1 節と同一である。

A. 4. 2 後処理

後処理は、適応ポストフィルタにおけるいくつかの簡略化を除き、4. 2 節と同一である。

適応ポストフィルタは、長期ポストフィルタ $H_p(z)$ 、短期ポストフィルタ $H_f(z)$ 、および傾き補償フィルタ $H_t(z)$ の3種類のフィルタを縦続接続したものであり、その後適応利得制御を行う。長期ポストフィルタは、整数値の遅延量のみを使用することにより簡略化されている。短期ポストフィルタおよび傾き補償フィルタでは、利得項である g_f と g_t が使用されない。ポストフィルタ処理は、 $1/\hat{A}(z/\gamma_d)$ による合成フィルタ処理よりも前に傾き補償フィルタ処理が実行されることを除いて、標準JT-G729の記述と同様である。

A. 4. 2. 1 長期ポストフィルタ

長期ポストフィルタは、次式で与えられる。

$$H_p(z) = \frac{1}{1 + \gamma_p g_l z^{-T}} \quad (\text{A.11})$$

4. 2. 1 節との唯一の違いは、長期遅延 T が常に整数の遅延であり、この T は $[T_{cl} - 3, T_{cl} + 3]$ の範囲から探索することである。ここで T_{cl} は、 $T_{cl} \leq 140$ に制限された現在のサブフレームにおける（伝送された）ピッチ遅延の整数部分である。

A. 4. 2. 2 短期ポストフィルタ

短期ポストフィルタは、次式で与えられる。

$$H_f(z) = \frac{\hat{A}(z/\gamma_n)}{\hat{A}(z/\gamma_d)} = \frac{1 + \sum_{i=1}^{10} \gamma_n^i \hat{a}_i z^{-i}}{1 + \sum_{i=1}^{10} \gamma_d^i \hat{a}_i z^{-i}} \quad (\text{A.12})$$

ここで、 $\hat{A}(z)$ は受信した量子化L P逆フィルタ（復号器ではL P分析は行われぬ）である。 γ_n と γ_d は短期ポストフィルタリング量を制御する係数であり、 $\gamma_n = 0.55$ 、 $\gamma_d = 0.7$ にセットされる。

4. 2. 2節との唯一の違いは、利得係数 g_f が削除されていることである。

A. 4. 2. 3 傾き補償

フィルタ $H_i(z)$ は、短期ポストフィルタ $H_f(z)$ の傾きを補償し、次式で与えられる。

$$H_i(z) = 1 + \gamma_i k'_i z^{-1} \quad (\text{A.13})$$

ここで、 $\gamma_i k'_i$ は傾き係数で、 k'_i は次式から計算される1次反射係数である。

$$k'_i = -\frac{r_h(1)}{r_h(0)}, \quad r_h(i) = \sum_{j=0}^{21-i} h_f(j)h_f(j+i) \quad (\text{A.14})$$

ここで、 $h_f(n)$ はフィルタ $\hat{A}(z/\gamma_n)/\hat{A}(z/\gamma_d)$ の打ち切られたインパルス応答である。 $k'_i < 0$ の場合 $\gamma_i = 0.8$ 、 $k'_i \geq 0$ の場合 $\gamma_i = 0$ が用いられる。4. 2. 3節で用いられていた利得係数 g_i は削除されている。

A. 4. 2. 4 適応利得制御

現在のサブフレームに対する利得調整係数 G が、次式で計算される点を除いて、4. 2. 4節と同一である。

$$G = \sqrt{\frac{\sum_{n=0}^{39} s^2(n)}{\sum_{n=0}^{39} sf^2(n)}} \quad (\text{A.15})$$

そして、 $g^{(n)}$ は次式で与えられる。

$$g^{(n)} = 0.9g^{(n-1)} + 0.1G, \quad n = 0, \dots, 39 \quad (\text{A.16})$$

A. 4. 2. 5 高域通過フィルタとアップスケーリング

4. 2. 5節と同一である。

A. 4. 3 符号器と復号器の初期化

4. 3節と同一である。

A. 4. 4 フレーム消失補償

有声/無声識別を使用しない点を除いて、4. 4節と同一である。励振信号は常に適応コードブック寄与分と固定コードブック寄与分の加算である。

A. 5 低演算量版CS-ACELPコーデックのビットイグザクトな記述

低演算量版CS-ACELPコーデックは、Table 5-2/JT-G729で示される固定小数点の基本演算と同じセットを用いた16ビット固定小数点のANSI Cコードでシミュレーションされている。

A. 5. 1 シミュレーションソフトウェアの使い方

5. 1 節と同一である。

A. 5. 2 シミュレーションソフトウェアの構成

5. 2 節と同一である。

シミュレーションコーデックで用いられるテーブルは、ファイル `tab_1d8a.c` に存在する。このファイルは標準 J T - G 7 2 9 本体でのファイル `tab_1d8k.c` に置き換わるものである。この 2 つのファイルの相違点は、`tab_1d8a.c` に存在するテーブル `tab_hup_s`、`tab_hup_1`、そして `inter_3` が、`tab_1d8a.c` では削除されていることである。またテーブル `glid` も変更されている。

メインプログラムは、固定小数点 ANS I C シミュレーションプログラムで提供されるサブルーチンのライブラリを使用している。ほとんどのサブルーチンは、標準 J T - G 7 2 9 本体で使用されているものと同じである。変更された主なルーチンは、Table A - 5 - 1 / JT-G729 に要約されている。詳細は、ソフトウェアとともに供給される `read.me` ファイルを参照のこと。

Table A - 5 - 1 / JT-G729 Summary of the principle routines which have been changed
(ITU-T G.729)

<i>G.729 routine name</i>	<i>G.729A routine name</i>
<code>Coder_1d8k()</code>	<code>Coder_1d8a()</code>
<code>Decod_1d8k()</code>	<code>Decod_1d8a()</code>
<code>Pitch_o1()</code>	<code>Pitch_o1_fast()</code>
<code>Pitch_fr3()</code>	<code>Pitch_fr3_fast()</code>
<code>ACELP_Codebook()</code>	<code>ACELP_Code_A()</code>
<code>Post()</code>	<code>Post-Filter()</code>

付属資料B

(標準 J T - G 7 2 9 に対する)

I T U - T 勧告 V . 7 0 端末に適した標準 J T - G 7 2 9 に対する無音圧縮手法

(本付属資料は仕様の一部である。)

B. 1 本付属資料の規定範囲

一般に、音声通信では音響的な背景雑音が存在する(オフィス、自動車、街路など)。聴感上で、背景雑音を正確に表現するために必要とされるビットレートは音声の場合よりかなり低い。そのため、帯域幅の削減が重要なアプリケーションについては、音声休止区間の検出、背景雑音の発生および伝送のための手法が要求される。無音圧縮音声通信システムのブロック図を、Figure B - 1 - 1 / JT-G729 に示す。

本付属資料では、音声検出(VAD)、不連続伝送(DTX)および擬似背景雑音発生器(CNG)アルゴリズムについて詳細な記述を提供する。これらのアルゴリズムは、音声のない区間において伝送速度を下げるために使用される。本アルゴリズムは I T U - T 勧告 V . 7 0 と共に動作するように設計され、そして最適化されている。I T U - T 勧告 V . 7 0 では、I T U - T 勧告 G . 7 2 9 Annex A (TTC 標準 J T - G 7 2 9 付属資料 A と同一)の音声符号化方式の使用が指定されている。しかし、必要ならば、音質を改善するために標準 J T - G 7 2 9 本体を使用することもできる。本アルゴリズムは、標準 J T - G 7 2 9 本体および付属資料 A の両方で動作するように適合させてある。ここでは標準 J T - G 7 2 9 本体に基づいた記述をしており、付属資料 A に基づく記述との違いを B . 3 . 1 . 1 節に示す。

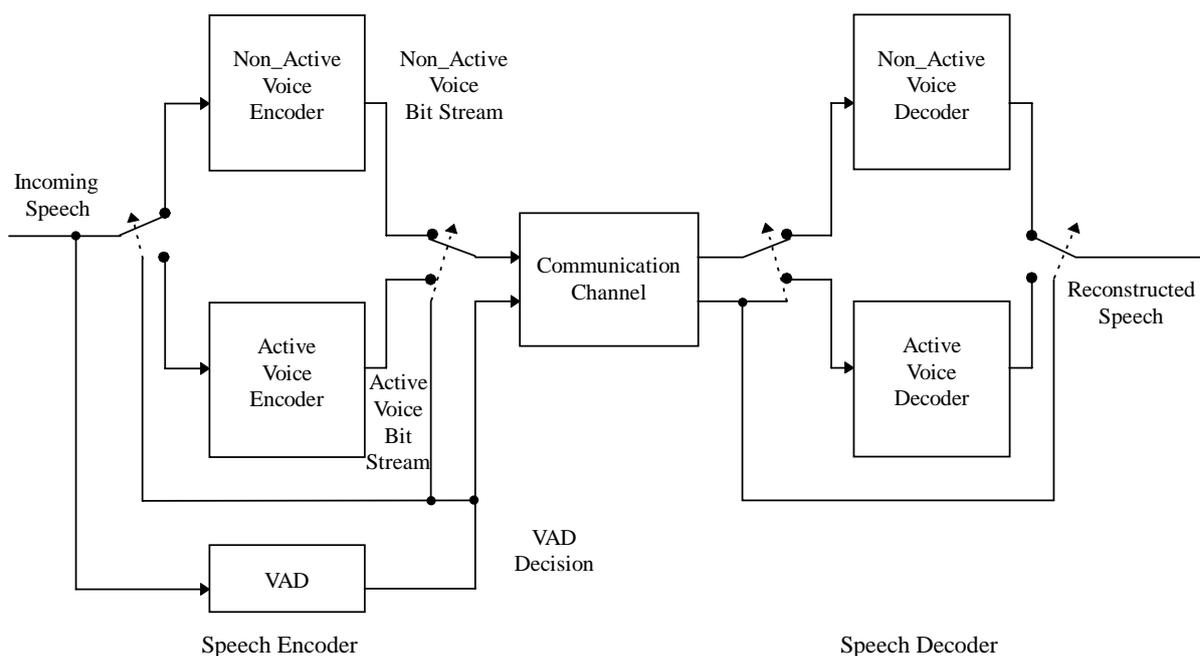


Figure B - 1 - 1 / JT-G729 Speech Communication System with VAD
(ITU-T G.729)

B. 2 VAD/DTX/CNGアルゴリズムの概要

VADアルゴリズムは、標準 J T - G 7 2 9 音声符号器と同じ 10ms フレーム毎に有音判定を行う。差分パラメータのセットが抽出され、初期判定に使用される。パラメータは全帯域のエネルギー、低域のエネルギー、零交差数およびスペクトル分布である。無音区間におけるパラメータの長時間平均は、背景雑音の特性変化に従う。差分パラメータのセットは、フレーム毎に得られる。これらは、それぞれのパラメータとそれぞれ

の長時間平均との差分量である。初期有音判定は、各差分パラメータ間の部分的な線形の判定境界を用いて得られる。最終的な有音判定は、初期判定を平滑化することによって得られる。

VADモジュールの出力は1または0で、それぞれ音声の有無を示している。もしVAD出力が1であるなら、有音フレームを符号化／復号するために標準JT-G729音声コーデックが呼び出される。一方、もしVAD出力が0であるなら、無音フレームを符号化／復号するために、ここに記すDTX/CNGアルゴリズムが用いられる。従来の音声コーデックでは、無音フレームにおける背景雑音を模擬するために擬似背景雑音（コンフォートノイズ）が用いられる。もし背景雑音が定常的でないなら、単調な擬似背景雑音の挿入では元の背景雑音の自然さを提供することはできない。そのため、無音フレームが検出されたとき、断続的に背景雑音についての若干の情報を送ることがよりよい品質を得るために望ましい。無音フレームのエネルギーとスペクトルの情報をわずか15ビットで符号化することにより、無音フレームの効率的な符号化を行うことができる。これらのビットは無音フレームが検出されても自動的に伝送されるとは限らない。もっと正確に言えば、最後に伝送された無音フレームについて感知できる変化が検出された時のみ、これらのビットが伝送される。

復号器側では、受信されたビット列が復号される。もしVAD出力が1であるなら、有音フレームを再生するために標準JT-G729復号器が呼び出される。もしVAD出力が0であるなら、無音フレームを再現するためにCNGモジュールが呼び出される。

B. 3 VADアルゴリズムの詳細記述

VAD処理のフローチャートをFigure B-3-1/JT-G729に示す。VADは、デジタル化された音声のフレームに基づいて動作する。フレームは時系列で処理され、それぞれの会話もしくは録音について、最初から連続した番号がつけられる。

第一段階において、4つの基本的な特徴が入力信号から抽出される。抽出されたパラメータは、計算効率のために有音用の符号器モジュールと無音用の符号器モジュールで共用される。パラメータは、全帯域と低域のフレームエネルギー、線スペクトル周波数(LSF)、および零交差数である。

フレーム番号が N_i より小さい場合、長時間平均の初期化が行われ、LPC分析のフレームエネルギーが15dB(式(B.1))より大きい場合、有音判定は強制的に1となる。そうでない場合には、有音判定は強制的に0となる。フレーム番号が N_i と等しい場合、背景雑音特有のエネルギー特徴量の初期化が行われる。

次の段階では、差分パラメータのセットが計算される。このセットは、現フレームのパラメータと背景雑音の特性の移動平均との間の差分量として生成される。4つの差分量が計算される。

- (1) スペクトル歪み
- (2) エネルギーの差
- (3) 低域エネルギーの差
- (4) 零交差数の差

最初の有音判定は、4つの差分量の空間における多境界判定領域を使用して次の段階で行われる。有音判定は、判定領域の結合として与えられ、無音判定は、その相補的な論理判定である。隣接する過去のフレームの判定と共に、エネルギー量の考慮が、判定を平滑化するために使用される。

移動平均は、音声に対してではなく、背景雑音に対してのみ更新されなければならない。適応的なしきい値がテストされ、更新はしきい値の基準が満たされた場合に限り行われる。

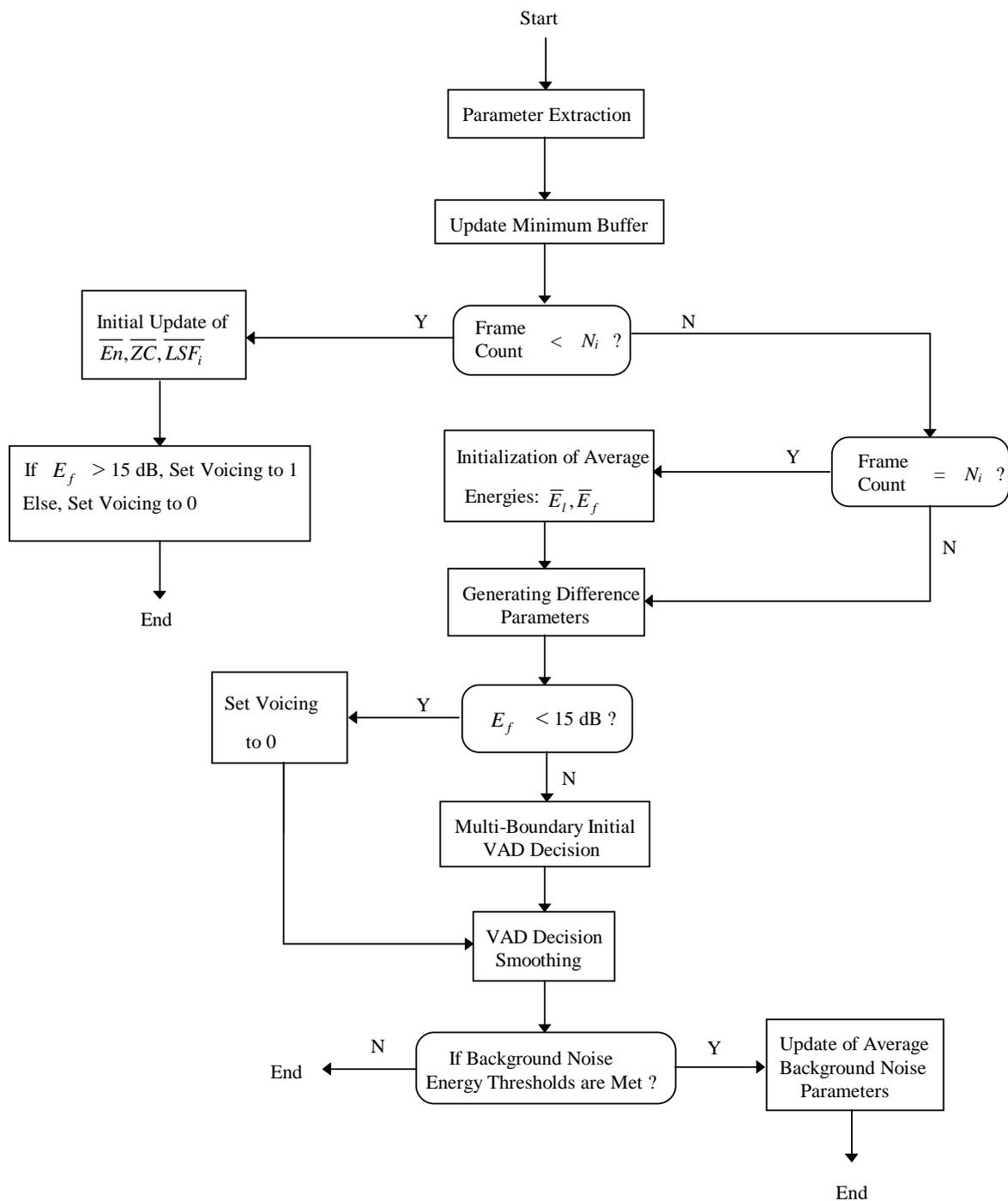


Figure B – 3 – 1 / JT-G729 VAD Flowchart
(ITU-T G.729)

B. 3. 1 パラメータの抽出

フレーム毎に音声信号からパラメータのセットが抽出される。パラメータ抽出モジュールは、VADモジュールと有音用の符号器と無音用の符号器との間で共用することができる。基本的なパラメータは自己相関係数であり、それは標準JT-G729の3. 2. 1節と同様に導出される。自己相関係数は、 $\{R(i)\}_{i=0}^q$ と表される。ここで $q=12$ である。

B. 3. 1. 1 線スペクトル周波数 (LSF)

線形予測係数が自己相関から求められ、そして標準 J T - G 7 2 9 の 3. 2. 3 節 (または付属資料 A の A. 3. 2. 3 節) で記述されているように $\{LSF_i\}_{i=1}^p$ が線形予測係数から求められる。ここで $p=10$ である。

B. 3. 1. 2 全帯域エネルギー

全帯域エネルギー E_f は、正規化された 0 次の自己相関係数 $R(0)$ の対数をとったものであり、次式で示される。

$$E_f = 10 \cdot \log_{10} \left[\frac{1}{N} R(0) \right] \quad (\text{B.1})$$

ここで $N=240$ は音声サンプルに対する L P C 分析窓のサイズである。

B. 3. 1. 3 低域エネルギー

0 から F_l Hz までの低域エネルギーは、次式で計算される。

$$E_l = 10 \cdot \log_{10} \left[\frac{1}{N} \mathbf{h}^T \mathbf{R} \mathbf{h} \right] \quad (\text{B.2})$$

ここで、 \mathbf{h} はカットオフ周波数が F_l Hz の F I R フィルタのインパルス応答であり、 \mathbf{R} は対角成分が自己相関係数であるテプリッツ自己相関行列である。

B. 3. 1. 4 零交差数

フレーム毎の正規化された零交差数 ZC は、次式で計算される。

$$ZC = \frac{1}{2M} \sum_{i=0}^{M-1} \left[\left| \text{sgn}[x(i)] - \text{sgn}[x(i-1)] \right| \right] \quad (\text{B.3})$$

ここで、 $\{x(i)\}$ は前処理された入力信号 (標準 J T - G 7 2 9 3. 1 節参照) であり、 $M=80$ である。

B. 3. 2 背景雑音特性の移動平均の初期化

最初の N_i フレームにおいて $\{\overline{LSF}_i\}_{i=1}^p$ で表される背景雑音のスペクトルパラメータは、それらのフレームの $\{LSF_i\}_{i=1}^p$ の平均として初期化される。 \overline{ZC} で表される背景雑音の零交差数の平均は、それらのフレームの零交差数 ZC の平均として初期化される。

\overline{E}_f で表される背景雑音エネルギーの移動平均、および \overline{E}_l で表される背景雑音低域エネルギーの移動平均は、以下のように初期化される。まず最初に、初期化処理は最初の N_i フレームのエネルギー E_f の平均として定義される \overline{E}_n を使用する。これらの 3 つの平均 (\overline{E}_n 、 \overline{ZC} 、および $\{\overline{LSF}_i\}_{i=1}^p$) は、エネルギー E が 15dB より大きいフレームのみ含む。次に、初期化処理は次のように続けられる。

if $\overline{E}_n \leq T_1$ then

$$\overline{E}_f = \overline{E}_n + K_0$$

$$\overline{E}_l = \overline{E}_n + K_1$$

else if $T_1 < \overline{E}_n < T_2$ then

$$\overline{E}_f = \overline{E}_n + K_2$$

$$\overline{E}_l = \overline{E}_n + K_3$$

else

$$\overline{E}_f = \overline{E}_n + K_4$$

$$\overline{E}_l = \overline{E}_n + K_5$$

定数については、Table B-3-1/JT-G729 を参照のこと。

Table B-3-1/JT-G729 Table of Constants
(ITU-T G.729)

Name	Constant	Name	Constant
N_i	32	N_1	4
N_0	128	N_2	10
K_0	0	T_1	671088640
K_1	-53687091	T_2	738197504
K_2	-67108864	T_3	26843546
K_3	-93952410	T_4	40265318
K_4	-134217728	T_5	40265318
K_5	-161061274	T_6	40265318
a_1	23488	b_1	28521
a_2	-30504	b_2	19446
a_3	-32768	b_3	-32768
a_4	26214	b_4	-19661
a_5	0	b_5	-30802
a_6	28160	b_6	-19661
a_7	0	b_7	30199
a_8	16384	b_8	-22938
a_9	-19065	b_9	-31576
a_{10}	0	b_{10}	-17367
a_{11}	22400	b_{11}	-27034
a_{12}	30427	b_{12}	29959
a_{13}	-24576	b_{13}	-29491
a_{14}	23406	b_{14}	-28087

B. 3. 3 長時間の最小エネルギーの生成

長時間の最小エネルギー E_{\min} は、直前の N_0 フレームの E_f の最小値として計算される。 N_0 は比較的大きいため、 E_{\min} は過去の短い区間での E_f の最小値を蓄積しておいた値を用いて計算される。

B. 3. 4 差分パラメータの生成

4つの差分量は、現在のフレームのパラメータと背景雑音の移動平均から計算される。

B. 3. 4. 1 スペクトル歪 ΔS

スペクトル歪量は、現在のフレームの $\{LSF_i\}_{i=1}^p$ ベクトルと背景雑音の移動平均 $\{\overline{LSF}_i\}_{i=1}^p$ との差の自乗和として得られる。

$$\Delta S = \sum_{i=1}^p (LSF_i - \overline{LSF}_i)^2 \quad (\text{B.4})$$

B. 3. 4. 2 全帯域エネルギーの差分 ΔE_f

全帯域エネルギーの差分量は、現在のフレームのエネルギー E_f と背景雑音エネルギーの移動平均 \overline{E}_f との差分として得られる。

$$\Delta E_f = \overline{E}_f - E_f \quad (\text{B.5})$$

B. 3. 4. 3 低域エネルギー差分 ΔE_l

低域エネルギーの差分量は、現在のフレームの低域エネルギー E_l と背景雑音低域エネルギーの移動平均 \overline{E}_l との差分として得られる。

$$\Delta E_l = \overline{E}_l - E_l \quad (\text{B.6})$$

B. 3. 4. 4 零交差数の差分 ΔZC

零交差数の差分量は、現在のフレームの零交差数 ZC と背景雑音の零交差数の移動平均 \overline{ZC} との差分として得られる。

$$\Delta ZC = \overline{ZC} - ZC \quad (\text{B.7})$$

B. 3. 5 多境界初期有音判定

初期有音判定は、 I_{VD} で表され、差分パラメータのベクトルが無音領域内に位置する場合は0（偽）に設定される。そうでなければ、初期有音判定は、1（真）に設定される。4次元空間における14の境界判定は、以下のように定義される。

- (1) if $\Delta S > a_1 \cdot \Delta ZC + b_1$ then $I_{VD} = 1$
- (2) if $\Delta S > a_2 \cdot \Delta ZC + b_2$ then $I_{VD} = 1$
- (3) if $\Delta E_f < a_3 \cdot \Delta ZC + b_3$ then $I_{VD} = 1$
- (4) if $\Delta E_f < a_4 \cdot \Delta ZC + b_4$ then $I_{VD} = 1$
- (5) if $\Delta E_f < b_5$ then $I_{VD} = 1$
- (6) if $\Delta E_f < a_6 \cdot \Delta S + b_6$ then $I_{VD} = 1$
- (7) if $\Delta S > b_7$ then $I_{VD} = 1$

- (8) *if* $\Delta E_f < a_8 \cdot \Delta ZC + b_8$ *then* $I_{VD} = 1$
- (9) *if* $\Delta E_f < a_9 \cdot \Delta ZC + b_9$ *then* $I_{VD} = 1$
- (10) *if* $\Delta E_f < b_{10}$ *then* $I_{VD} = 1$
- (11) *if* $\Delta E_l < a_{11} \cdot \Delta S + b_{11}$ *then* $I_{VD} = 1$
- (12) *if* $\Delta E_l > a_{12} \cdot \Delta E_f + b_{12}$ *then* $I_{VD} = 1$
- (13) *if* $\Delta E_l < a_{13} \cdot \Delta E_f + b_{13}$ *then* $I_{VD} = 1$
- (14) *if* $\Delta E_l < a_{14} \cdot \Delta E_f + b_{14}$ *then* $I_{VD} = 1$

14条件のうちの一つも「真」がなければ、 $I_{VD} = 0$ とする。定数については、Table B-3-1/JT-G729を参照のこと。

B. 3. 6 有音判定の平滑化

初期有音判定は音声信号の長時間の定常状態を反映するために平滑化（ハングオーバ）される。平滑化は4つの段階に分けて行われる。

ハングオーバが生じたことを示すフラグを v_flag と定義する。これは有音判定の平滑化が行われる前に毎回0に設定される。現在のフレーム、直前のフレーム、2つ前のフレームの平滑化された有音判定をそれぞれ S_{VD}^0 、 S_{VD}^{-1} 、 S_{VD}^{-2} とする。 S_{VD}^{-1} と S_{VD}^{-2} は1に初期化される。また初期状態では $S_{VD}^0 = I_{VD}$ に設定する。第1段階での平滑化は次のように行われる。

$$if(I_{VD} = 0) \text{ and } (S_{VD}^{-1} = 1) \text{ and } (E > \bar{E}_f + T_3) \text{ then } S_{VD}^0 = 1 \text{ and } v_flag = 1$$

第2段階での平滑化に対し、ブール変数 F_{VD}^{-1} と平滑化カウンタ C_e を定義する。 F_{VD}^{-1} は1に、 C_e は0に初期化される。直前のフレームのエネルギーを E_{-1} とする。第2段階での平滑化は次のように行われる。

$$if(F_{VD}^{-1} = 1) \text{ and } (I_{VD} = 0) \text{ and } (S_{VD}^{-1} = 1) \text{ and } (S_{VD}^{-2} = 1) \text{ and } (|E_f - E_{-1}| \leq T_4) \{$$

$$S_{VD}^0 = 1 \text{ and } v_flag = 1$$

$$C_e = C_e + 1$$

$$if(C_e \leq N_1) \{$$

$$F_{VD}^{-1} = 1$$

$$\}$$

$$else \{$$

$$F_{VD}^{-1} = 0$$

$$C_e = 0$$

$$\}$$

$$\}$$

$$else$$

$$F_{VD}^{-1} = 1$$

第3段階での平滑化に対し、雑音継続カウンタ C_s を定義する。これは0に初期化される。もし $S_{VD}^0 = 0$ ならば、 C_s は1が加算される。第3段階での平滑化は次のように行われる。

$$\begin{aligned} & \text{if } (S_{VD}^0 = 1) \text{ and } (C_s > N_2) \text{ and } (E_f - E_{-1} \leq T_5) \{ \\ & \quad S_{VD}^0 = 0 \\ & \quad C_s = 0 \\ & \quad \} \\ & \text{if } (S_{VD}^0 = 1) \ C_s = 0 \end{aligned}$$

第4段階では、もし次の条件が満足されていれば有音判定を0に設定する。

$$\text{if } ((E_f < \bar{E}_f + T_6) \text{ and } (\text{frame_count} > N_0) \text{ and } (v_flag = 0)) \text{ then } S_{VD}^0 = 0$$

B. 3. 7 背景雑音特性の移動平均の更新

背景雑音特性の移動平均はVADモジュールの最後の段階で更新される。この段階では、次の条件がテストされ条件を満たしていた場合更新が行なわれる。

$$\text{if } (E_f < \bar{E}_f + T_6) \text{ then update}$$

背景雑音特性の移動平均は1次の自己回帰（AR）手法を用いて更新される。異なったAR係数が異なった変数に対して用いられ、係数の異なったセットが録音または会話の最初や雑音特性の大幅な変化を検出した時に用いられる。

β_{E_f} を \bar{E}_f の更新のためのAR係数、 β_{E_l} を \bar{E}_l の更新のためのAR係数、 β_{ZC} を \bar{ZC} の更新のためのAR係数そして β_{LSF} を $\{LSF_i\}_{i=1}^p$ の更新のためのAR係数とする。更新条件を満たしているフレームの総数は C_n でカウントされる。 C_n の値により係数 β_{E_f} 、 β_{E_l} 、 β_{ZC} そして β_{LSF} の異なったセットが用いられる。

ARの更新は次の式に従って行われる。

$$\begin{aligned} \bar{E}_f &= \beta_{E_f} \cdot \bar{E}_f + (1 - \beta_{E_f}) \cdot E_f \\ \bar{E}_l &= \beta_{E_l} \cdot \bar{E}_l + (1 - \beta_{E_l}) \cdot E_l \\ \bar{ZC} &= \beta_{ZC} \cdot \bar{ZC} + (1 - \beta_{ZC}) \cdot ZC \\ \overline{LSF}_i &= \beta_{LSF} \cdot \overline{LSF}_i + (1 - \beta_{LSF}) \cdot LSF_i, \quad i = 1, \dots, p \end{aligned} \tag{B.8}$$

\bar{E}_f と C_n はさらに次のように更新される。

$$\begin{aligned} & \text{if } (\text{frame_count} > N_0) \text{ and } (\bar{E}_f < E_{\min}) \{ \\ & \quad \bar{E}_f = E_{\min} \\ & \quad C_n = 0 \\ & \quad \} \end{aligned}$$

B. 4 DTX/CNGアルゴリズムの詳細記述

DTX/CNGアルゴリズムは、低い平均ビットレートを保ちながら無音区間についての連続的かつスムーズな情報を提供する。

B. 4. 1 DTXアルゴリズムの記述

各々の無音フレームに対して、DTXモジュールは無音信号における変化を測定することにより、音声復号器に対して無音の更新パラメータを送信すべきかどうかを決定する。更新するかどうかの判断にはフレームエネルギーの絶対値と適応的しきい値、および、スペクトル歪量が用いられる。更新が必要なときは無音用の符号器はオリジナルの無音信号と聴感上で同様な信号を生成するのに必要な情報を送信する。この情報はエネルギーレベルとスペクトル包絡の記述から構成される。更新が必要のない場合は、無音信号は最後に受信した無音フレームのエネルギーとスペクトル形状情報に従って無音用の復号器により生成される。

しかし、2つの連続するSID (B. 4. 2節参照) フレームの間には $N_{min}=2$ フレームの最小間隔が必要である。すなわち、もしスペクトルまたはレベルの変化がSIDフレームの後の N_{min} より小さい n フレーム目で発生した場合は、SIDの送信が遅れてしまう。

送信端において、DTXモジュールはVADモジュールから、有音か無音かの情報を受け取り、符号器モジュールから80サンプルのフレームと過去の励振サンプルとの間で計算された音声信号の自己相関関数を受け取る。各々のフレームに対してDTXの判定値 $Ftyp_t$ (フレーム番号 t に対するフレームタイプ) が出力される。これは以下の手順に従う非伝送フレーム、有音フレーム、または、SIDフレームのいずれかであることを示し、値としては各々0、1、または2をとる。

B. 4. 1. 1 フレーム自己相関関数の保持

有音あるいは無音の全てのフレーム t に対して、帯域幅拡張や雑音補正 (標準JT-G 7 2 9の記述参照) を含む現在のフレーム t についての自己相関関数はメモリに保持される。フレーム t の自己相関は $r_t'(j)$, $j=0, \dots, 10$ で示される。

B. 4. 1. 2 現在のフレームタイプの計算

もし現在のフレーム t が有音フレーム ($Vad_t=1$) ならば、現在のフレームタイプ $Ftyp_t$ は1であり、通常の音声符号化処理が継続される。

他の場合には現在のフレーム t を含む直前の $N_{cur}=2$ フレーム上で計算される現在のLPCフィルタ $A_t(z)$ が最初に評価される。

N_{cur} の自己相関関数は次式で計算される。

$$R^t(j) = \sum_{i=t-N_{cur}+1}^t r_i'(j), j=0, \dots, 10 \quad (B.9)$$

$A_t(z)$ は $R^t(j)$ を入力として用いるレビンソン・ダービンの手順 (標準JT-G 7 2 9の記述参照) で計算される。このフィルタの係数は $a_t(j)$, $j=0, \dots, 10$ で示される。レビンソン・ダービンの手順は残差エネルギー E_t も提供する。これは再スケーリングされ、フレーム励振エネルギーの推定値として用いられる。

その後で現在のフレームタイプ $Ftyp_t$ は以下の方法で決定される。

- (1) 現在の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。エネルギー和を反映する変数 \bar{E} は E_t と等しい値をとり、エネルギー和に関するフレームの数 k_E は1に初期化される。

$$(Vad_{t-1}=1) \Rightarrow \begin{cases} Ftyp_t = 2 \\ \bar{E} = E_t \\ k_E = 1 \end{cases} \quad (B.10)$$

(2) その他のフレームに対しては、本アルゴリズムでは現在のパラメータを直前のS I Dパラメータと比較する。もし、現在のフィルタが直前のS I Dフィルタと大幅に異なるか、現在の励振エネルギーが直前のS I Dエネルギーと大幅に異なる場合には、フラグ $flag_chang$ は1にセットされる。そうでない場合にはフラグ $flag_chang$ は変化しない。

直前のS I Dフレームから何フレームが経過したかを示すカウンタ $count_fr$ はインクリメントされる。もし、その値が N_{min} よりも大きい場合にはS I Dフレームの送信が許される。そのとき、もし、 $flag_chang$ が1ならばS I Dフレームが送信される。その他の場合には全て現在のフレームは送信されない。

$$\left. \begin{array}{l} count_fr \geq N_{min} \\ flag_chang = 1 \end{array} \right\} \Rightarrow Ftyp_t = 2 \quad (B.11)$$

Otherwise: $Ftyp_t = 0$

S I Dフレームの場合にはカウンタ $count_fr$ とフラグ $flag_chang$ は0に再初期化される。

L P Cフィルタとエネルギーは以下の方法に従って比較される。

B. 4. 1. 3 L P Cフィルタの比較

直前のS I D L P Cフィルタは $A_{sid}(z)$ とその係数 $a_{sid}(j)$, $j=0, \dots, 10$ (本フィルタの推定法はB. 4. 2. 2節において記述される) によって表される。現在、および、直前のS I D L P Cフィルタは次式で示されるように2つのフィルタの板倉距離が与えられたしきい値を超えた場合に大幅に異なるとみなされる。

$$\sum_{j=0}^{10} R_a(j) \times R^t(j) \geq E_t \times thr1 \quad (B.12)$$

ここで、 $R_a(j)$, $j=0, \dots, 10$ はS I Dのフィルタ係数の自己相関から求められる関数であり、次式で与えられる。

$$\left\{ \begin{array}{l} R_a(j) = 2 \sum_{k=0}^{10-j} a_{sid}(k) \times a_{sid}(k+j), \quad \text{if } j \neq 0 \\ R_a(0) = \sum_{k=0}^{10} a_{sid}^2(k) \end{array} \right. \quad (B.13)$$

$thr1$ の値としては1.20226が用いられる。

B. 4. 1. 4 エネルギーの比較

k_E が最大値 $Ng=2$ までインクリメントされながらフレームエネルギーの合計が計算される。

$$\bar{E} = \sum_{i=t-k_E+1}^t E_i \quad (B.14)$$

それから \bar{E} はB. 4. 2. 1節で記述される5ビットの対数量子化器により量子化される。逆量子化された対数エネルギー E_q は逆量子化された直前のS I Dの対数エネルギー E_q^{sid} と比較される。もし、その差がしきい値 $thr2=2\text{dB}$ を超えている場合には、2つのエネルギーは大幅に異なるものと見なされる。

B. 4. 2 S I Dの評定と量子化

無音挿入記述子 (S I D) には、量子化されたフレーム励振エネルギー (即ち、S I Dフレームのための現在の量子化された励振エネルギー $Q(\bar{E})$)、および推定されたS I D L P Cフィルタに対応した量子化されたL S Pが含まれる。S I Dフレームは4つのインデックスから成る。1つのインデックスはエネルギーを示し、3つのインデックスはS I Dフレームのスペクトル部分を示す。

B. 4. 2. 1 エネルギーの量子化

エネルギー \bar{E} の量子化は、次のように行われる。まず、サブフレームの自己相関関数 $r'(j)$ に現れる帯域幅拡張と窓掛けの効果を考慮して、スケーリングファクタ $\alpha_w = 0.125$ が導入される。

利得量子化器の入力には、次の値を用いる。

$$E' = \alpha_w \times \frac{1}{k_E \times N_{cur} \times 80} \bar{E} \quad (\text{B.15})$$

エネルギー項 E' は、-12dB から 66dB の範囲での対数領域において、5ビットの不均一量子化器によって量子化される。16dB から 66dB の範囲では、2dB の均一ステップサイズが用いられる。-4dB から 16dB の範囲では、4dB のステップサイズが用いられる。-4dB 以下では、8dB のステップサイズが用いられ、量子化レベルを-12dB としている。量子化は、ダイレクトに行われ、量子化テーブルの記憶は必要としない。

エネルギー比較 (B. 4. 1. 4 節参照) は復号されたエネルギーで処理されるものであり、エネルギーの量子化は全ての無音フレームに対して行われることに注意しなければならない。

B. 4. 2. 2 S I D L P Cフィルタの推定と量子化

S I D L P Cフィルタの推定は、近接したS I Dフレームにおける雑音が、局所的に定常であるか非定常であるかを考慮する。

まず、現在のS I Dフレームに先行する N_p 個のフレームからなる過去平均フィルタ $\bar{A}_p(z)$ が計算される。これには、レビンソン・ダービン法の入力として次の自己相関総和を用いる。

$$\bar{R}_p(j) = \sum_{k=t'-N_p}^{t'} r'_k(j), \quad j = 0 \text{ to } 10 \quad (\text{B.16})$$

この総和におけるフレームの数は、 $N_p = 6$ に固定される。

フレーム番号 t' は $[t-1, t-N_{cur}]$ の範囲であり、現在のフレーム番号 t を N_{cur} でユークリッド除算したときの余りに依存する。

そこで、S I D L P Cフィルタは以下の式で与えられる。

$$A_{sid}(z) = \begin{cases} A_t(z), & \text{if distance}(A_t(z), \bar{A}_p(z)) \geq thr3 \\ \bar{A}_p(z), & \text{otherwise} \end{cases} \quad (\text{B.17})$$

しきい値 $thr3$ は 1.12202 に固定されており、現在のL P Cフィルタと過去平均フィルタの距離はB. 4. 1. 3 節 (式(B.12)) と同様に計算される。

次に、量子化のために、S I D L P CフィルタはL S F領域に変換される。L S Fは、それぞれ5ビットと4ビットからなる2段の切換え型予測ベクトル量子化 (V Q) により量子化される。L S Fベクトルの

量子化には、最適な3つのインデックスを決定することを要する。最初のインデックスは、予測器に関するものである。残りの2つのインデックスは、2段のベクトル量子化の結果であり、異なるベクトルテーブルからそれぞれ得られる。全体的な量子化手順は、標準JT-G729（3.2.4節）で与えられているものに従うが、次に示す変更を伴う。

- (1) 標準JT-G729で使用されている第2の4次MA予測器を、以下に示す第1と第2のMA予測器の線形結合に変更する。

$$p_{i,k,2} = 0.6p_{i,k,1} + 0.4p_{i,k,2}, \quad \text{where } i = 1, \dots, 10, k = 1, \dots, 4 \quad (\text{B.18})$$

- (2) 1段目のVQは、標準JT-G729で用いられるものと同様である。但し、最初の量子化器のテーブルを部分的に用いる。そのテーブルの中で該当する要素のサブセットは、32のアドレスインデックスを有する補助的な参照テーブルに格納される。さらには、2段目への入力としていくつかの候補を残しておくことにより、ディレイドデジジョン量子化が用いられる。
- (3) 1段目における複数の候補は、2段目のVQに使用される。2段目のVQは、標準JT-G729で用いられるものとは異なる。標準JT-G729が分割VQであるのに対し、完全なVQが用いられる。その上、2段目の量子化器のテーブルを部分的に用いる。該当する要素のサブセットは、2組の16の要素を有する別の補助的な参照テーブルに格納される。重み付き平均自乗誤差を最小にする場合の、予測器、1段目から得られるベクトル、そして2段目から得られるベクトルの組合せが、LSF記述子として選択される。

B.4.3 SIDビット列の記述

SIDフレームの伝送に関連するビット列を、Table B-4-1/JT-G729に示す。有音フレームの伝送に関連するビット列は、Table 4-1/JT-G729に定められている。ビット列の順番は、この表の中の順序に従う。各パラメータのMSBが、最初に伝送される。

Table B-4-1/JT-G729
(ITU-T G.729)

<i>Parameter Description</i>	<i>Bits</i>
Switched predictor index of LSF quantizer	1
First stage vector of LSF quantizer	5
Second stage vector of LSF quantizer	4
Gain(Energy)	5

B.4.4 無音用符号器と無音用復号器（CNG）の記述

復号器側では、レベルが制御された擬似白色励振信号を補間されたLPCフィルタに入力し、擬似背景雑音を発生させる。これは、復号器が復号された励振信号をフィルタリングすることにより、音声が発生する方法と同じである。励振信号のレベルとLPCフィルタ係数は、直前のSID情報から求める。SID情報のLSPを現在のLSPとし、標準JT-G729の有音フレームで行ったように直前フレームのLSPで補間することにより、サブフレームでの補間されたLPCフィルタを求める。

擬似白色励振信号 $ex(n)$ は、有音時の励振信号と同じ種類の励振信号 $ex_1(n)$ と白色ガウス励振信号 $ex_2(n)$ で構成される。標準JT-G729の励振信号 $ex_1(n)$ は、利得が小さい適応励振信号とACELP固定励振

信号から構成され、有音フレームと無音フレーム間の遷移域で品質を向上する。ガウス雑音励振信号 $ex_2(n)$ の付加は、発生する信号を白色化する。

符号器と復号器は無音区間においても同期関係を維持する必要があるので、S I Dフレームと非伝送フレームに対し双方で励振信号を発生する。

まず、現在のフレーム t の合成された励振信号に対する平均エネルギーの平方根であるターゲット励振信号 \tilde{G}_t を定める。 \tilde{G}_t は以下に示す平滑化の手順を用いて計算される。 \tilde{G}_{sid} は、復号した時に求まるS I D利得である。

$$\tilde{G}_t = \begin{cases} \tilde{G}_{sid} & \text{if } Vad_{t-1} = 1 \\ \frac{7}{8}\tilde{G}_{t-1} + \frac{1}{8}\tilde{G}_{sid} & \text{otherwise} \end{cases} \quad (\text{B.19})$$

1 フレーム 80 サンプルを 40 サンプルからなる 2 つのサブフレームに分割する。各サブフレーム毎に、以下に示すアルゴリズムを用いてC N G励振信号が合成される。

ピッチ周期を 40 から 103 の間でランダムに選ぶ。

次に、グリッド、そしてパルスの極性と位置を標準 J T - G 7 2 9 の A C E L P の符号構造に従いランダムに選択し、サブフレームの固定コードブックベクトルを組み立てる。

そして、単位利得の適応励振信号 $e_a(n)$, $n = 0, \dots, 39$ が計算される。また、選択されたサブフレームの固定励振信号は、 $e_f(n)$, $n = 0, \dots, 39$ となる。

適応利得 G_a と固定利得 G_f は、サブフレームの平均エネルギーが \tilde{G}_t^2 に等しくなるよう次式のように計算される。

$$\frac{1}{40} \sum_{n=0}^{39} (G_a \times e_a(n) + G_f \times e_f(n))^2 = \tilde{G}_t^2 \quad (\text{B.20})$$

G_f は負の値も取り得ることに注意すること。

ここで、

$$E_a = \left(\sum_{n=0}^{39} e_a^2(n) \right), \quad I = \left(\sum_{n=0}^{119} e_a(n)e_f(n) \right) \text{ and } K = 40 \times \tilde{G}_t^2$$

とする。

A C E L P 励振信号の構造によれば、

$$\sum_{n=0}^{39} e_f^2(n) = 4$$

となる。

もし適応利得 G_a をランダムに固定すれば、式(B.20)は次式のような固定利得 G_f を変数とする 2 次方程式になる。

$$G_f^2 + \frac{G_a \times I}{2} G_f + \frac{E_a \times G_a^2 - K}{4} = 0 \quad (\text{B.21})$$

この方程式が確実に解を持つように、 G_a を強制的に定める。さらに、大きな適応利得を用いないことが望ましい。この条件に対して、適応利得 G_a は以下のようにランダムに選ばれる。

$$\left[0, \text{Max} \left\{ 0.5, \sqrt{\frac{K}{A}} \right\} \right], \text{ with } A = E_a - I^2 / 4 \quad (\text{B.22})$$

式(B.21)の根のうち、絶対値が最小となる値が G_f として選ばれる。
最後に標準 J T - G 7 2 9 励振信号が次式のように求められる。

$$ex_1(n) = G_a \times e_a(n) + G_f \times e_f(n), \quad n = 0, \dots, 39 \quad (\text{B.23})$$

結合された励振信号 $ex(n)$ は、以下の方法で求める。

E_1 を $ex_1(n)$ のエネルギー、 E_2 を $ex_2(n)$ のエネルギーとする。 $ex_2(n)$ は1あるいは0の値を持つ。 E_3 を $ex_1(n)$ と $ex_2(n)$ の相互エネルギーとする。

$$\begin{aligned} E_1 &= \sum ex_1^2(n) \\ E_2 &= \sum ex_2^2(n) \\ E_3 &= \sum ex_1(n) \cdot ex_2(n) \end{aligned} \quad (\text{B.24})$$

上式において、和はサブフレームの範囲で行われる。

α と β は励振信号を合成する時の $ex_1(n)$ と $ex_2(n)$ それぞれに対する調整比とする。 α は0.6に設定され、 β は次式に示す2次方程式の解として求められる。

$$\beta^2 E_2 + 2\alpha\beta E_3 + (\alpha^2 - 1)E_1 = 0, \quad \text{with } \beta > 0 \quad (\text{B.25})$$

もし β に対して解が得られなければ、 α は1、 β は0に設定される。

CNG励振信号 $ex(n)$ は、次式のようになる。

$$ex(n) = \alpha ex_1(n) + \beta ex_2(n) \quad (\text{B.26})$$

B. 4. 5 CNGに関するフレーム消失補償

フレーム消失が復号器で検出された場合、直前のフレームタイプから消失したフレームタイプを決定する。

- (1) もし、直前のフレームが有音フレームならば、現在の消失フレームも有音フレームとみなす。
- (2) もし、直前のフレームがS I Dフレームあるいは非伝送フレームならば、現在の消失フレームを非伝送フレームとみなす。

$$\begin{cases} Ftyp_{t-1} = 1 & \Rightarrow Ftyp_t = 1 \\ Ftyp_{t-1} = 0 \text{ or } 2 & \Rightarrow Ftyp_t = 0 \end{cases} \quad (\text{B.27})$$

非伝送フレームが消失している場合、誤りは発生しない。

SIDフレームが消失している場合、以下に示す2つの状況が考えられる。

(a) もし、SIDフレームが現在の無音区間の最初のSIDフレームでなければ、直前のSIDパラメータが保持される。

(b) もし、SIDフレームが現在の無音区間の最初のSIDフレームならば、特別の保護がなされる。

上記(b)の状況は、 $Ftyp_{i-1} = 1$ かつ $Ftyp_i = 0$ によって検出できることに注意すること。この組み合わせは、直前のフレームが正常な有音フレームでないことを示しており、直前のフレームより前の幾つかのフレームも消失している場合があり得る。確かなことは、最後の正常なフレームは有音フレームであり、現在のフレームは消失したものではなく、現在の非伝送フレームの情報を提供しているSIDフレームが消失している、ということである。

上記(b)の場合、SID情報を再生するために、CNGモジュールは標準JT-G729復号器のメイン部より供給されるパラメータを以下のように用いる。

(7) 最新の有音フレームのLSPをSID LPCフィルタに用いる。

(4) 全ての有音フレーム処理中の復号器によって、エネルギー項は励振信号を用いて計算される。

SID利得 \tilde{G}_{sid} を再生するために、最新有音フレームのエネルギー項をSID利得量子化器で量子化し復号する。

最後に、励振信号を計算するランダム発生器の状態不一致を避けるために、有音フレーム毎に符号器、復号器双方で擬似ランダムシーケンスのリセットを行う。

B. 5 無音圧縮手法のビットイグザクトな記述

無音圧縮手法は、Table 5-2/JT-G729 に示した固定小数点の基本演算セットを用いた16ビット固定小数点のANSI Cコードでシミュレーションされている。このANSI Cコードは、ビットイグザクトな無音圧縮手法の固定小数点記述を反映した本標準の全てで構成されている。本標準の記述と、Cソースコードの不一致が発生した場合、Cソースコードが優先される。

B. 5. 1 シミュレーションソフトウェアの構成

5. 2節と同一である。

付属資料BのANSI CソフトウェアモジュールをTable B-5-1/JT-G729 に示す。詳細は、ソフトウェアとともに供給される read.me ファイルを参照すること。

Table B-5-1/JT-G729

(ITU-T/G.729)

<i>G.729 Annex B ANSI-C Module Name</i>	<i>Description</i>
Vad.c	VAD
Dtx.c	DTX Decision
Qsidgain.c	SID Gain Quantization
QsidLSF.c	SID LSF Quantization
Calcexc.c	CNG Excitation Calculation
Dec_sid.c	Decode SID Information
Miscel.c	Miscellaneous Calculations
<i>G.729 Annex B ANSI-C .h File Name</i>	<i>Description</i>
Vad.h	Prototype and Constants
Dtx.h	Prototype and Constants

Sid.h	Prototype and Constants
Miscel.h	Prototype and Constants

付属資料C

(標準 J T - G 7 2 9 に対する)

浮動小数点演算版 8kbit/s C S - A C E L P 音声コーデック

(本付属資料は仕様の一部である。)

C. 1 本付属資料の規定範囲

本付属資料は、標準 J T - G 7 2 9 および標準 J T - G 7 2 9 付属資料 A に対する浮動小数点演算に基づいた別の実現方法についての記述である。固定小数点仕様で規定されている標準 J T - G 7 2 9 付属資料 B と同様な特性を持つ音声検出器 (VAD)、不連続伝送/無音圧縮 (DTX) および背景雑音発生器 (CNG) の相互接続可能な浮動小数点仕様の標準化は、今後の課題である。

C. 2 参照すべき標準

(1) TTC 標準 J T - G 7 2 9

8kbit/s C S - A C E L P を用いた音声符号化方式。平成 8 年 1 1 月 2 7 日制定。

(2) TTC 標準 J T - G 7 2 9 付属資料 A

低演算量版 8kbit/s C S - A C E L P 音声コーデック。平成 9 年 4 月 2 3 日制定。

C. 3 概要

TTC 標準 J T - G 7 2 9 は、音声信号を 8kbit/s で符号化するためのアルゴリズムをビットイグザクトな固定小数点演算で規定している。また、標準 J T - G 7 2 9 付属資料 A は、標準 J T - G 7 2 9 と相互接続可能な低演算量版である。これらの仕様の詳細な記述は、ビットイグザクトな固定小数点演算を示した C コードとして新日本 I T U 協会から入手可能である。本付属資料は、標準 J T - G 7 2 9 および標準 J T - G 7 2 9 付属資料 A に対する浮動小数点演算に基づいた別の実現方法について記述ならびに定義する。

C. 4 アルゴリズムの詳細

本浮動小数点版は、標準 J T - G 7 2 9 (および標準 J T - G 7 2 9 付属資料 A) と同様のアルゴリズムステップを踏んでいる。同様に浮動小数点版のビットストリームは標準 J T - G 7 2 9 (および標準 J T - G 7 2 9 付属資料 A) のビットストリームと全く同一のものとなっている。アルゴリズムの詳細に関しては、標準 J T - G 7 2 9 (および標準 J T - G 7 2 9 付属資料 A) を参照すること。

C. 5 ANS I C コード

本付属資料にて定義されている標準 J T - G 7 2 9 (および標準 J T - G 7 2 9 付属資料 A) 浮動小数点版をシミュレートする ANS I C コードはすでに開発が終了し、本付属資料の対応物として利用可能である。ANS I C コードは、本付属資料を規定している。C コードで指定されているアルゴリズム記述は、標準 J T - G 7 2 9 本体、標準 J T - G 7 2 9 付属資料 A あるいは標準 J T - G 7 2 9 付属資料 C の本文の記述よりも優先される。現在の ANS I C コードは 1 9 9 8 年 9 月 2 5 日付け、バージョン 1.01 となっている。これらの浮動小数点版ソースコードの構成は固定小数点版ソースコードのそれと対応し、関連付けられている。TTC 標準 J T - G 7 2 3. 1 付属資料 B と同様に、typedef.h ファイルに全ての浮動小数点変数および定数を倍精度型、もしくは単精度型として定義した記述が含まれている。version.h と呼ばれるファイルは、C コードが標準 J T - G 7 2 9 または標準 J T - G 7 2 9 付属資料 A のどちらの処理を行うかを選択する際に利用できる。ソフトウェアファイル名とその概要のリストを Table C - 5 - 1 / JT-G729 から Table C - 5 - 3 / JT-G729 に示す。固定小数点版の basic_op.c、oper_32b.c、dspfunc.c、および、basic_op.h、oper_32b.h のファイルは、浮動小数点版では必要でないことに注意すること。また、float を short に変換す

るルーチンがファイル `util.c` に加えられた。

Table C – 5 – 1 / JT-G729 List of software files specific to G.729 floating point source code
(IUT-T G.729)

<i>File name</i>	<i>Description</i>	<i>File Size (in bytes)</i>
<code>coder.c</code>	main program for G.729 encoder	4591
<code>cod_ld8k.c</code>	G.729 encoder routine	19336
<code>acelp_co.c</code>	G.729 fixed codebook search	23318
<code>lpc.c</code>	G.729 LP analysis	9470
<code>lpcfunc.c</code>	miscellaneous routines related to LP filter	5470
<code>pitch.c</code>	G.729 pitch search	14279
<code>pwf.c</code>	G.729 computation of perceptual weighting coefficients	3849
<code>decoder.c</code>	main program for G.729 decoder	5235
<code>dec_ld8k.c</code>	G.729 decoder routine	9219
<code>postfil.c</code>	G.729 postfilter	23554
<code>tab_ld8k.c</code>	G.729 constants tables	33179
<code>ld8k.h</code>	G.729 prototypes and constant declarations	16238
<code>tab_ld8k.h</code>	G.729 declaration of constants tables	1675
<code>version.h</code>	used to select the G.729 (main body) mode	916

Table C – 5 – 2 / JT-G729 List of software files specific to G.729A floating point source code
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>File Size (in bytes)</i>
<code>codera.c</code>	main program for G.729A encoder	4514
<code>acelp_ca.c</code>	G.729A fixed codebook search	25238
<code>cod_ld8a.c</code>	G.729A encoder routine	18453
<code>lpca.c</code>	G.729A LP analysis	9535
<code>lpcfunca.c</code>	miscellaneous routines related to LP filter	4019
<code>pitch_a.c</code>	G.729A pitch search	12468
<code>decoder.a.c</code>	main program for G.729A decoder	5043
<code>dec_ld8a.c</code>	G.729A decoder routine	9473
<code>postfila.c</code>	G.729A postfilter	12949
<code>tab_ld8a.c</code>	G.729A tables of constants	32830
<code>ld8a.h</code>	G.729A prototypes and constant declarations	17449
<code>tab_ld8a.h</code>	declaration of G.729A constants tables	1295
<code>version.h</code>	used to select the G.729 Annex A mode	931

Table C – 5 – 3 / JT-G729 List of software files common to G.729 and G.729A floating point source code
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>File Size (in bytes)</i>
bits.c	bit manipulation routines	4644
qua_lsp.c	LSP quantizer	11953
qua_gain.c	gain quantizer	7402
cor_func.c	miscellaneous routines related to excitation computation	2603
de_acelp.c	algebraic codebook decoder	2014
dec_gain.c	gain decoder	3247
dec_lag3.c	adaptive codebook index decoder	2412
filter.c	filter functions	3634
gainpred.c	gain predictor	4363
lspdec.c	LSP decoding routine	3852
lspgetq.c	LSP quantizer	6815
p_parity.c	pitch parity computation	2036
post_pro.c	post-processing (HP filtering)	2906
pre_proc.c	pre-processing (HP filtering)	2909
pred_lt3.c	generation of adaptive codebook	2424
taming.c	pitch taming functions	3396
util.c	utility function	3292
typedef.h	data type definition (machine dependent)	1504

付属資料C+

(標準J T-G 7 2 9に対する)

標準J T-G 7 2 9 CS-ACELP音声符号化方式本体、付属資料B、D、およびEの統合における
浮動小数点演算での実現方法

(本付属資料は仕様の一部である。)

C+. 1 本付属資料の規定範囲

本付属資料は、標準J T-G 7 2 9本体、付属資料B、D、およびEに関し、浮動小数点演算を用いた統合についての記述を規定している。本付属資料は、この統合と機能的な拡張に関し、標準的な実現方法を示している。これにより、製品化を助け、標準的な音声品質と互換性を世界的に保証することができる。本統合は、製品化における要求を満足するために、以下に示すいくつかの制約に着目しつつ、これを達成している。

- (1) 付属資料Cに示した浮動小数点演算版の標準本体とビットイグザクトなこと。
- (2) プログラムコードやメモリ、演算量の増加を最小限に抑えること。
- (3) 関連する標準や付属資料におけるアプリケーションや品質との調和を図りつつ、新しい機能に対して厳しい品質を要求すること。

C+. 2 参照すべき標準

本付属資料の規定は、下記のTTC標準とその他の参照文献に含まれた規定から構成されており、それらの規定は本文中の至る所で参照されている。本付属資料の出版時点で、表記した版が制定されている。すべての標準およびその他の参照文献は改訂の対象である。従って、本付属資料のすべての利用者は、以下に示す標準およびその他の参照文献の最新版が適用できるかどうか調査することを奨励する。最新版のTTC標準のリストは、定期的な出版されている。

- [1] TTC標準J T-G 7 2 9
共役構造一代数的符号励振線形予測(CS-ACELP)を用いた8kbit/s音声符号化方式。平成8年11月27日制定。
- [2] TTC標準J T-G 7 2 9付属資料B
ITU-T勧告V. 70端末に適した標準J T-G 7 2 9に対する無音圧縮手法。平成9年4月23日制定。
- [3] TTC標準J T-G 7 2 9付属資料C
浮動小数点演算版8kbit/s CS-ACELP音声コーデック。平成11年4月22日制定。
- [4] TTC標準J T-G 7 2 9付属資料D
6.4kbit/s CS-ACELP音声コーデック。平成11年4月22日制定。
- [5] TTC標準J T-G 7 2 9付属資料E
11.8kbit/s CS-ACELP音声コーデック。平成11年4月22日制定。

C+. 3 概要

標準J T-G 7 2 9本体[1]、付属資料B[2]、D[4]、およびE[5]は、それぞれ8kbit/s CS-ACELP符号化方式、DTX機能、6.4kbit/sの低ビットレート拡張、および11.8kbit/sの高ビットレート拡張のビットイグザクトな固定小数点演算仕様を規定している。これらの仕様の詳細な記述はビットイグザクトな固定小数点演算のCコードで示されており、そのCコードは日本ITU協会から入手可能である。付属資料C[3]は、標準J T-G 7 2 9本体の別の実現方法について記述ならびに定義している。付属資料C+は、浮動小数点演算による標準J T-G 7 2 9本体、付属資料B、D、およびEの統合について、記述ならびに定義してお

り、標準 J T - G 7 2 9 付属資料 C の拡張と考えることができる。

C+. 4 新しい機能

本章では、標準本体、付属資料 B、D、および E の統合を容易にするためのアルゴリズムの修正および追加の概要を示す。アプリケーションに対応させるために、モジュールによっては、ある追加が必要であることが判っている。

C+. 4. 1 付属資料 D における付属資料 B の D T X 動作

付属資料 D において D T X 動作を提供するために付属資料 B と D の機能を統合することは、簡単なことである。付属資料 B の V A D (音声検出) や、S I D (無音挿入記述子) の符号化、C N G (擬似背景雑音発生器) は、特に修正せずにそのまま使用できる。不連続伝送の区間では、付属資料 D のポストフィルタにおいて、位相拡散のためのパラメータの更新に注意する必要がある (C+. 5. 2 節、参照)。

C+. 4. 2 付属資料 E における付属資料 B の D T X 動作

付属資料 E において D T X 動作を提供するために付属資料 B と E の機能を統合することは、若干の複雑さがある。付属資料 B の D T X 動作が 10 次の L P C 分析に基づいているため、付属資料 B の V A D 機能は、付属資料 E の 10 次のフォワード適応 L P C 分析が行われた後で、かつバックワード適応 L P C 分析が行われる前に実行される。V A D 機能が「無音」を検出した場合は、付属資料 E の L P C モードは強制的にフォワード適応 L P C となり、バックワード適応 L P C 分析はスキップされる。さらに音楽を検出するために、V A D 機能の後に補正用のモジュールを追加する必要があることが判っている。これは、ある程度の品質で音楽を伝送する能力を提供することが付属資料 E の目的の一つであり、付属資料 E の拡張されたアプリケーションにも対応させるためのものである。ちなみに、付属資料 E の開発に当たっては、音楽信号に対する性能において厳しい要求項目があった。一方、標準本体、付属資料 B および D に関しては、音楽信号に対する性能において厳しい要求項目はなかった。付属資料 B の D T X 動作を適用した場合の付属資料 E の音楽信号に対する品質を保証するために、音楽信号が継続している区間では、音楽検出機能は強制的に V A D 機能を「有音」の状態にする。これにより、音楽の区間では確実に付属資料 E の 11.8kbit/s で符号化が行われる。付属資料 B における S I D の符号化や C N G は、特に修正せずにそのまま使用できる。なお、不連続伝送の区間では、付属資料 E の L P C モード選択アルゴリズムにおけるパラメータを正しく更新することに注意する必要がある (C+. 5. 3 節、参照)。

C+. 5 アルゴリズムの記述

本章では、個々の付属資料のアルゴリズムに対し、統合を容易にするために必要な付加アルゴリズムの記述を示す。残りのすべてのモジュールは、標準本体、付属資料 B、D または E が元になっている。

C+. 5. 1 音楽検出

音楽検出は新しい機能である。これは V A D の直後に実行され、音楽区間中は、V A D を強制的に「有音」の状態にする。これは付属資料 E の動作中に限り有効であるが、そのパラメータは、統合された標準 J T - G 7 2 9 の D T X 動作中におけるビットレートモードと独立しており、常に更新される。

音楽検出アルゴリズムは、音楽信号が存在している間の、音声検出 (V A D) からの判定を訂正する。これは、付属資料 B の D T X、すなわち不連続伝送モード動作中に、付属資料 E と結合して用いられる。音楽検出は、以下のパラメータに基づいている。

- (1) *Vad_deci*、現フレームの V A D 判定
- (2) *PVad_dec*、直前フレームの V A D 判定

- (3) *Lpc_mod*、直前フレームの適応LPCが、フォワードかバックワードかいずれかを示すフラグインジケータ
- (4) *Rc*、LPC分析から得られる反射係数
- (5) *Lag_buf*、最終5フレームの修正された開ループピッチラグのバッファ
- (6) *Pgain_buf*、最終サブ5フレームの閉ループピッチ利得のバッファ
- (7) *Energy*、LPC分析からの1次自己相関係数 $R(0)$
- (8) *LEnergy*、VADモジュールからの正規化対数エネルギー
- (9) *Frm_count*、処理された信号フレーム数のカウンタ
- (10) *Rate*、音声コーデックの選択

本アルゴリズムは、2つの主要な部分からなる。

- (1) 関連パラメータの計算
- (2) パラメータに基づく分類

C+. 5. 1. 1 関連パラメータの計算

本節では、判定モジュールによって使用されるパラメータの計算について記述している。

- (1) 部分正規化残差エネルギー

$$LEnergy = 10 \log_{10} \left(\prod_{i=1}^4 (1 - Rc(i)^2) Energy / 240 \right)$$

- (2) スペクトル差分と背景雑音の部分正規化残差エネルギーの移動平均

現フレームの反射係数 Rc と背景雑音の移動平均反射係数 mRc の間のスペクトル差分量は次式で与えられる。

$$SD = \sum_{i=1}^{10} (Rc(i) - mRc(i))^2$$

移動平均 \overline{mrc} と $mLEnergy$ は、VADモジュールによって生成されたVAD判定 Vad_dec を用いて以下のとおり更新される。

```

if Vad_dec == NOISE {
     $\overline{mrc} = 0.9\overline{mrc} + 0.1rc$ 
     $mLEnergy = 0.9mLEnergy + 0.1LEnergy$ 
}

```

- (3) ピッチラグバッファ更新のための開ループピッチラグ補正

開ループピッチラグ T_{op} は、倍ピッチあるいは3倍ピッチを回避するため、以下のように補正して用いる。

$$avg_lag = \sum_{i=1}^4 Lag_buf(i) / 4$$

```

if (abs( $T_{op} / 2 - avg\_lag$ ) <= 2)
    Lag_buf(5) =  $T_{op} / 2$ 
else if (abs( $T_{op} / 3 - avg\_lag$ ) <= 2)
    Lag_buf(5) =  $T_{op} / 3$ 
else
    Lag_buf(5) =  $T_{op}$ 

```

なお、開ループピッチラグ T_{op} は変更されないこと、かつ開ループピッチ分析によって得られたものと同一であることに注意すべきである。

(4) ピッチラグの標準偏差

$$std = \sqrt{Var / 4}$$

$$\text{where } Var = \sum_{i=1}^5 (Lag_buf(i) - \mu)^2 \text{ and } \mu = \sum_{i=1}^5 (Lag_buf(i) / 5)$$

(5) ピッチ利得の移動平均

$$mPgain = 0.8mPgain + 0.2\theta, \quad \text{where } \theta = \sum_{i=1}^5 (Lag_buf(i) / 5)$$

ピッチ利得バッファ $Pgain_buf$ は、サブフレーム処理の後、もし $Vad_deci = NOISE$ ならば、ピッチ利得値 0.5 で更新され、そうでなければ、量子化されたピッチ利得で更新される。

(6) ピッチラグの滑らかさと有声強度インジケータ

ピッチラグの滑らかさと有声強度インジケータ $Pflag$ は、以下の論理的なステップを使用して生成される。

最初に、2つの中間論理フラグ $Pflag1$ と $Pflag2$ が以下のとおり得られる。

```

if (std < 1.3 and mPgain > 0.45) set Pflag1=1 else 0
if (mPgain > Thres) set Pflag2=1 else 0,
where Thres = 0.73 if Rate=G729D, otherwise Thres=0.63

```

最後に、 $Pflag$ は以下より決定される

```

if ((PVad_dec == VOICE and (Pflag1==1 or Pflag2==1)) or (Pflag2==1))
set Pflag=1 else 0

```

(7) 定常性カウンタ

1組のカウンタが定義され、以下のとおり更新される。

- (a) $count_consc_rflag$ は、2次反射係数とピッチ利得の移動平均が以下の条件を満足する場合に、連続フレーム数をカウントする。

```

if (Rc(2) < 0.45 and Rc(2) > 0 and mPgain < 0.5)
    count_consc_rflag = count_consc_rflag + 1
else
    count_consc_rflag = 0

```

- (b) *count_music* は、64 フレームの窓内で直前のフレームがバックワード適応LPCを使用しており、現フレームが（VADにより）「有音」である場合に、フレーム数をカウントする。

```

if (LPC_mod == 1 and Vad_deci == VOICE)
    cont_music = cont_music + 1

```

64 フレーム毎に、*count_music* の移動平均 *mcount_music* は、以下の記述のとおり更新および零にリセットされる。

```

if ((Frm_count mod 64) == 0){
    if (Frm_count == 64)
        mcount_music = count_music
    else
        mcount_music = 0.9mcount_music + 0.1count_music
}

```

- (c) *count_consc* は、*count_music* が零のままである場合、連続フレーム数をカウントする。

```

if (count_music == 0)
    count_consc = count_consc + 1
else
    count_consc = 0

```

```

if (count_consc > 500 or count_consc_rflag > 150) set mcount_music = 0

```

(b)における *count_music* は、関連するカウンタの更新の後、64 フレーム毎に零にリセットされる。

(c)における論理は、*count_music* の移動平均をリセットするために使用される。

- (d) *count_pflag* は、64 フレームの窓内で *Pflag* = 1 であるフレーム数をカウントする。

```

if (Pflag == 1)
    count_pflag = count_pflag + 1

```

64 フレーム毎に、*count_pflag* の移動平均 *mcount_pflag* は更新され、以下の記述のとおり零にリセットされる。

```

if ((Frm_count mod 64) == 0){
  if (Frm_count == 64)
    mcount_pflag = count_pflag
  else{
    if (count_pflag > 25)
      mcount_pflag = 0.98mcount_pflag + 0.02count_pflag
    else(count_pflag > 20)
      mcount_pflag = 0.95mcount_pflag + 0.05count_pflag
    else
      mcount_pflag = 0.9mcount_pflag + 0.1count_pflag
  }
}

```

(e) *count_consc_pflag* は、以下の条件を満足する連続フレーム数をカウントする。

```

if (count_pflag == 0)
  count_consc_pflag = count_consc_pflag + 1
else
  count_consc_pflag = 0

if (count_consc_pflag > 100 or count_consc_rflag > 150) set mcount_pflag = 0

```

count_pflag は 64 フレーム毎に零にリセットされる。(e)における論理は *count_pflag* の移動平均をリセットするために使用される。

C+. 5. 1. 2 分類

上記のパラメータ推定に基づいて、VADモジュールからのVAD判定*Vad_deci* は、もし以下の条件が満足されるのならば判定をやり直す。

```

if (Rate = G729E){
  if (SD > 0.15 and (Lenergy - mLenergy) > 4 and LLenergy > 50)
    Vad_deci = VOICE
  else if ((SD > 0.38 or (Lenergy - mLenergy) > 4) and LLenergy > 50)
    Vad_deci = VOICE
  else if ((mcount_pflag >= 10 or mcount_music >= 5 or Frm_count < 64) and LLenergy > 7)
    Vad_deci = VOICE
}

```

注意すべきことは、音楽検出機能は、メモリを最新の状態に維持するために、動作中の符号化モードに関わらず、常に呼び出されることである。ただし、VAD判定*Vad_deci* については、統合された標準J T-G 7 2 9が、11.8kbit/s (付属資料E) で動作しているときに限り変化する。また、音楽検出は、VAD判定に対し「無音」から「有音」へ判定を変えることだけを許容し、その逆はなされないことに注意すべきである。

C+. 5. 2 不連続伝送時の付属資料Dに特有な状態変数の更新

付属資料Dに特有な唯一の状態変数は、復号器における位相拡散モジュールの状態変数である(標準J T-G 7 2 9 付属資料D[4] D. 6. 2 節、参照)。無音フレームでは、擬似背景雑音発生器によって計算さ

れた利得値（標準 J T-G 7 2 9 付属資料 B[2] B. 4. 4 節、参照）を、適応および A C E L P 利得の推定値として、標準ビットレート（8kbit/s）と同じ更新手順が実行される。高レートに対する更新が標準ビットレートに対する更新と同一であることを注意すること。

C+. 5. 3 不連続伝送時の付属資料 E に特有な状態変数の更新

C+. 5. 3. 1 付属資料 E に特有な符号器の状態変数の更新

符号器において、無音フレームの場合、状態変数の更新は標準 J T-G 7 2 9 付属資料 E[5]における標準ビットレート 8kbit/s に切り替える場合と同一である。更新手順は以下の通りである。LP モードを 0 にセットし、グローバル定常性インジケータの値を減少させ、高定常性インジケータを 0 にリセットする（標準 J T-G 7 2 9 付属資料 E[5] E. 3. 2. 7. 2 節、参照）。フォワード LP フィルタからバックワード LP フィルタへのなめらかな切り替えを行うための補間係数を、最大値にリセットする（標準 J T-G 7 2 9 付属資料 E[5] E. 3. 2. 7. 1 節、参照）。この更新は低ビットレートである 6.4kbit/s へ切り替える場合も使用される。

C+. 5. 3. 2 付属資料 E に特有な復号器の状態変数の更新

復号器において、無音フレームの場合、状態変数の更新はピッチ追跡手順による計算（標準 J T-G 7 2 9 付属資料 E[5] E. 4. 4. 5 節、参照）を実行する代わりに、ピッチ遅延定常性インジケータを 0 にリセットすることを除いて、標準 J T-G 7 2 9 付属資料 E[5]におけるフォワードモードのみのレート（8kbit/s および 6.4kbit/s）に切り替える場合と同一の処理を行う。

C+. 6 C ソースコードの記述

付属資料 C+ は標準 J T-G 7 2 9 本体、付属資料 B、D および E の機能を統合したものであり、ANSI C による浮動小数点演算でシミュレーションされている。標準 J T-G 7 2 9 付属資料 C において、`typedef.h` ファイルは倍精度型または単精度型のどちらかのタイプの浮動小数点変数および定数を定義した記述を含んでいる。ANSI C コードは本付属資料の仕様を規定している。C コードによって与えられるアルゴリズム記述は標準 J T-G 7 2 9 本体、付属資料 B、C、D、E および C+ の本文の記述よりも優先される。

C+. 6. 1 シミュレーションソフトウェアの使い方

C コードは 2 つのメインプログラム `codercp.c` と `decodercp.c` からなり、それぞれ符号器と復号器のシミュレーションを行う。符号器は次のように実行される

```
codercp inputfile bitstreamfile dtx_option rate_option
```

復号器は次のように実行される

```
decodercp bitstreamfile outputfile
```

入力ファイルと出力ファイルは、8kHz で標準化された 16 ビットの PCM 信号からなるデータファイルである。ビットストリームファイルは、ビットストリームを含んだバイナリファイルである。符号化されたビットストリームのマッピングテーブルは、シミュレーションソフトウェアに含まれる。符号器に対する 2 つのオプションは、`dtx_option` と `rate_option` である。

`dtx_option` = 0 : DTX 無効

 = 1 : DTX 有効

 初期値は 0 (DTX 無効)

`rate_option` = 0 : 低レート (6.4kbit/s) を選択

 = 1 : 標準 J T-G 7 2 9 本体 (8kbit/s) を選択

= 2 : 高レート (11.8kbit/s) または、フレーム毎に 0、1、2 のレートを選択する
16 ビットワードからなるバイナリファイル `file_rate_name` を選択
初期値は 1 (8kbit/s)

C+. 6. 2 シミュレーションソフトウェアの構成

Table C+. 6-1 / JT-G729 に簡潔な記述を伴ったソフトウェアファイル名の一覧を示し、そしてファイルがどの付属資料から得られたか (標準 J T-G 7 2 9 付属資料 C のファイルに同一あるいは類似しているか、またはファイルに固定から浮動小数点の書き換えがあるか) を示す。固定小数点用のファイル `basic_op.c`、`oper_32b.c`、`dspfunc.c` および `basic_op.h`、`oper_32b.h` は浮動小数点演算には使用されない事に注意すること。標準 J T-G 7 2 9 付属資料 C において、`float` を `short` に変換するルーチンがユーティリティファイル `utilcp.c` に追加された。

Table C + . 6 – 1 / JT-G729 List of software files of integrated G.729 in floating point
(ITU-T G.729)

<i>File Name</i>	<i>Description</i>	<i>Link</i>
Gainpred.c	Gain predictor	C
Lpfuncpc.c	Miscellaneous routines related to LP filter	C+E
Cor_func.c	miscellaneous routines related to excitation computation	C
Pre_proc.c	Pre-processing (HP filtering and scaling)	C
P_parity.c	Compute pitch parity	C
Pwf.c	Computation of perceptual weighting coefficients (8 kbit/s)	C
Pred_lt3.c	Generation of adaptive codebook	C
Post_pro.c	Post processing (HP filtering and scaling)	C
Tab_ld8k.c	ROM tables	C
Ld8k.h	Function prototypes	C
Tab_ld8k.h	Extern ROM table declarations	C
Typedef.h	Data type definition (machine dependent)	C
Taming.c	Pitch instability control	C
Qsidgain.c	SID Gain Quantization	B
Qsidlsf.c	SID-LSF Quantization	B
Tab_dtx.c	ROM tables	B
Sid.h	Prototype and Constants	B
Octet.h	Octet transmission mode definition	B
Tab_dtx.h	Extern ROM table declarations	B
Pwfe.c	Computation of perceptual weighting coefficients (11.8 kbit/s)	E
Vad.c	VAD	B
Dtx.c	DTX Decision	B
Vad.h	Prototype and Constants	B
Dtx.h	Prototype and Constants	B
Calcexc.c	CNG Excitation Calculation	B
Dec_sid.c	Decode SID Information	B
Utilcp.c	Utility functions	C+B
Phdisp.c	Phase dispersion	D
Bwfw.c	Backward/forward switch selection	E
Bwfwfunc.c	Miscellaneous routines related to backward/forward switch	E
Filtere.c	Filter functions	C+E
Lpcp.c	LP analysis	C+E
Lspcdece.c	LSP decoding routines	C+E
Lspgetqe	LSP quantizer	C+E
Qua_lspe.c	LSP quantizer	C+E
Track_pi.c	Pitch tracking	E
Codercp.c	Main encoder routine	C+B+D+E

Codld8cp.c	Encoder routine	C+B+D+E
Decodep.c	Main decoder routine	C+B+D+E
Declld8cp.c	Decoder routine	C+B+D+E
Acelp_cp.c	search ACELP fixed codebook (6.4, 8, 11.8 kbit/s)	C+D+E
Dacelpcp.c	Decode algebraic codebook (6.4, 8, 11.8 kbit/s)	C+D+E
Pitchcp.c	Pitch search	C+D+E
Declagcp.c	Decode adaptive-codebook index	C+D+E
Q_gaincp.c	Gain quantizer	C+D+E
Degaincp.c	Decode gain	C+D+E
Pstpcp.c	Postfilter routines	C+B+E
Bitscp.c	Bit manipulation routines	C+B+D+E
Tabld8cp.c	ROM tables for G.729 at 6.4 and 11.8 kbit/s	D+E
Tabld8cp.h	Extern ROM declarations for G.729 at 6.4 and 11.8 kbit/s	D+E
Ld8cp.h	Constant and Function prototypes for G.729 at 6.4 and 11.8 kbit/s	D+E
Mus_dtct.c	Music detection module	New

付属資料D

(標準 J T - G 7 2 9 に対する)

6.4kbit/s C S - A C E L P 音声コーデック

(本付属資料は仕様の一部である。)

D. 1 本付属資料の規定範囲

本付属資料は、標準 J T - G 7 2 9 において過負荷状態を扱うような場合等に柔軟性をもたせるために規定する標準 J T - G 7 2 9 の低レート拡張に関するものである。標準 J T - G 7 2 9 本体と同等レベルの品質ではないが、ほとんどの場合 24kbit/s 時の標準 J T - G 7 2 6 より著しくすぐれた品質となっている。しかしながら、高レベル時の自動車雑音に対しては、本アルゴリズムはある性能制限を有している。標準 J T - G 7 2 9 本体との差分について記述している。

D. 2 参照すべき標準

- (1) T T C 標準 J T - G 7 2 9
8kbit/s C S - A C E L P を用いた音声符号化方式

D. 3 6.4kbit/s 拡張に対する符号器の概要記述

符号器は、わずかな例外を除いて、T T C 標準 J T - G 7 2 9 と同様である。変更点を以下に要約しており、以降の章でより詳細に記述している。

- (1) 標準 J T - G 7 2 9 の A C E L P コードブックを、長さが異なり（各々 16 および 32 の位置を有する）2 つの重なり合ったトラックに、2 本の極性付きパルスをもつ新 A C E L P コードブックに変更した。
- (2) 利得に対する共役構造コードブックを 6 ビットの新共役構造コードブックに置き換えた。
- (3) 第 2 サブフレームにおいてピッチ遅延の変形符号化を使用する。ビット数を 4 ビットに縮小した。分数デルタ値の不規則分布を用いることにより、デルタラグ範囲を設定している。
- (4) よりまばらな代数的コードブックの影響を軽減させるため、後処理フィルタを付加する手法を適用した。
- (5) ピッチ遅延パリティビットを除去した。

新符号器は、標準 J T - G 7 2 9 で使用される 8.0kbit/s すなわちフレームあたり 80 ビットの代わりに 6.4kbit/s すなわちフレームあたり 64 ビットで使用する。

D. 4 ビット割り当て

Table D-4-1 / JT-G729 Bit allocation for 6.4 kbit/s G.729
(ITU-T G.729)

<i>Parameter</i>	<i>Number of bits per frame (10 ms)</i>
LP parameters	18
Adaptive codebook	8 + 4
Fixed codebook	2 * 11
Gain quantizer	2 * 6
Total	64

太字は標準本体からの変更を表す。

D. 5 符号器の機能記述

D. 5. 1 前処理

標準 JT-G 7 2 9 と同一である。

D. 5. 2 線形予測分析と量子化

標準 JT-G 7 2 9 と同一である。

D. 5. 3 聴覚重み付け

標準 JT-G 7 2 9 と同一である。

D. 5. 4 開ループピッチ分析

標準 JT-G 7 2 9 と同一である。

D. 5. 5 インパルス応答の計算

標準 JT-G 7 2 9 と同一である。

D. 5. 6 ターゲット信号の計算

標準 JT-G 7 2 9 と同一である。

D. 5. 7 適応コードブック探索

基準となる符号化サブフレーム（第1サブフレーム）に対するLTP符号化は標準JT-G729本体と同一である。第2サブフレームのLTPラグの数を32から16に減少させた。整数デルタラグ値は、 $int(T1) - 5$ から $int(T1) - 2$ および $int(T1) + 1$ から $int(T1) + 4$ の範囲で使用される。ここで、 $T1$ は直前のサブフレームのLTPラグを示す。 $\frac{1}{3}$ の分析精度を有する分数ラグは、範囲 $int(T1) - 1\frac{2}{3}$ から $int(T1) + \frac{2}{3}$ の範囲で使用される。

D. 5. 8 固定コードブックの構造と探索

標準 J T - G 7 2 9 で使用している 4 本のパルスをもつコードブックを、2 つの重なり合ったトラックが 2 本の極性付きパルスをもつ A C E L P コードブックに置き換えた。トラックテーブルを Table D - 5 - 1 に示す。パルスの極性は、標準本体のようにあらかじめ設定される。パルス位置の探索には、効果的な探索をするため、全 512 ベクトルにわたって徹底的に計算される。

Table D - 5 - 1 / JT-G729 ACELP track table
(ITU-T G.729)

<i>Pulse</i>	<i>Sign</i>	<i>Positions</i>
i_0	+1/-1	1, 3, 6, 8, 11, 13, 16, 18, 21, 23, 26, 28, 31, 33, 36, 38
i_1	+1/-1	0, 1, 2, 4, 5, 6, 7, 9, 10, 11, 12, 14, 15, 16, 17, 19, 20, 21, 22, 24, 25, 26, 27, 29, 30, 31, 32, 34, 35, 36, 37, 39

D. 5. 9 利得の量子化

共役構造の利得コードブックは再設計された。サブフレームあたり 6 ビットが利得コードブックに割り当てられる。コードブックはランダム分布をもつ 0.1% のビット誤りの条件を用いて学習された。本コードブックは 32 ワードのメモリを要する。

D. 5. 10 メモリの更新

標準 J T - G 7 2 9 と同一である。

D. 6 復号器の機能記述

D. 6. 1 パラメータ復号処理

標準 J T - G 7 2 9 と同様である。パラメータ数は減少し、より小さな固定励振コードブックパラメータを使用している。

D. 6. 2 固定コードブックの後処理

付加される後処理フィルタは、まばらな励振による聴覚的悪影響を軽減させるため、復号器に適用される。本フィルタは、エネルギーがサブフレームにさらに拡散するように新しい信号に変更する。本フィルタは、主に「擬似ランダム」インパルス応答により位相を変更する。フィルタ処理は、記憶された 3 種類のインパルス応答のうちの一つを用いて、巡回畳込みにより実行される。フィルタの選択は、フィルタを通して受信された L T P 利得に基づく有声/無声判定によって制御される。3 種類のインパルス応答は、拡散量の違いに対応する。最大拡散は、フィルタを通した L T P 利得が低く、雑音に近い区間に適用される。中間拡散は、中間的 L T P 利得に対して適用される。そしてフィルタを通した L T P 利得が高い有声音に対しては、拡散が適用されない。加えて、コードブック利得に急激な増加が検出されると、始めの拡散を避ける。

D. 6. 3 ポストフィルタと後処理

標準 J T - G 7 2 9 と同一である。

D. 6. 4 フレーム消失補償

標準 J T - G 7 2 9 と同一である。

D. 7 A N S I Cコード

標準 J T - G 7 2 9 の低レート拡張に規定されている A N S I Cコードは、本付属資料の対応物として利用可能である。現在の A N S I Cコードのバージョンは、1 9 9 8年5月付け、バージョン 1.2 である。A N S I Cコードは、本付属資料を規定している。Cコードで規定されているアルゴリズムは、標準 J T - G 7 2 9 本体および標準 J T - G 7 2 9 付属資料 D に記載されている内容よりも優先される。

Table D - 7 - 1 / JT-G729 List of software files specific to G.729 lower bit rate extension encoder
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Size (Bytes)</i>
acelp-cod.c	Search fixed codebook	41657
cod-ld8kd.c	Encoder routine	27883
coderd.c	Encoder	5467
pitchd.c	Pitch search	23000
qua_g6k.c	Gain quantizer	19506

Table D - 7 - 2 / JT-G729 List of software files specific to G.729 lower bit rate extension decoder
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Size (Bytes)</i>
dec-lag3d.c	Decode adaptive-codebook index	3416
dec-ld8kd.c	Decoder routine	10930
decoderd.c	Decoder	6971
de-acelpd.c	Decode algebraic codebook	1331
dec_g6k.c	Decode gain	4443

Table D - 7 - 3 / JT-G729 List of software files specific to G.729 lower bit rate extension routines
(ITU-T G.729) common to encoder and decoder

<i>File name</i>	<i>Description</i>	<i>Size (Bytes)</i>
bitsd.c	Bit manipulation routines	4952
filterd.c	Filter functions	8540
ld8kd.h	Switching variables	7136
tab-ld8kd.c	Tables	3545
tab-ld8kd.h		710

TableD – 7 – 4 / JT-G729 List of software files specific to G.729 lower bit rate extension routines
 (ITU-T G.729) common to G.729 and G.729D

<i>File name</i>	<i>Description</i>
basic_op.h	<i>Common to G.729 main body</i>
ld8k.h	
oper_32b.h	
tab_ld8k.h	
typedef.h	
basic_op.c	
de_acelp.c	
dec_gain.c	
gainpred.c	
lpc.c	
lpcfunc.c	
lspdec.c	
lspgetq.c	
oper_32b.c	
p_parity.c	
post_pro.c	
pre_proc.c	
pred_lt3.c	
pst.c	
pwf.c	
qua_gain.c	
qua_lsp.c	
tab_ld8k.c	
util.c	

付属資料 E

(標準 J T - G 7 2 9 に対する)

11.8kbit/s C S - A C E L P 音声コーデック

(本付属資料は仕様の一部である。)

E. 1 本付属資料の規定範囲

本付属資料は、標準 J T - G 7 2 9 の高ビットレート拡張に関する概要記述である。本拡張は、背景雑音が存在する音声や音楽のような幅広い入力信号に適応するよう設計されている。

E. 2 コーデックの概要

本拡張アルゴリズムは、標準 J T - G 7 2 9 アルゴリズムへの変更、追加をできるだけ少なくするように設計されている。実際に標準 J T - G 7 2 9 へ追加されるものは、音楽信号や定常的な背景雑音に適合したバックワード L P 分析を導入した L P 部分と、11.8kbit/s へビットレート拡張するための 2 つの新しい代数的励振コードブックの設計に関わっている。このコードブックの 1 つはフォワードモードの中で使用され、もう 1 つの大きい方はバックワードモードで使用されている。ポストフィルタリングと聴覚重み付け処理への変更が僅かに存在するが、それ以外の全ての処理は標準 J T - G 7 2 9 での処理と完全に同一である。また、誤り補償においてもバックワード/フォワード L P 構造へ適合させるための変更がなされている。

2 つの L P 分析はフレーム毎に行われる。1 つは合成信号に対しバックワードで、もう 1 つは入力信号に対しフォワードで行われる。適応判定処理は最適なフィルタを選択し、必要ならば切り替え動作を行う。このアルゴリズムのフォワード L P 部は、同じ L S P 量子化手法を有する標準 J T - G 7 2 9 のフォワード L P 部と同一である。バックワード L P 分析は 30 の次数を持っており、符号器と復号器の両方で処理を行う。L P 係数は伝送されないため、余ったビットレートは代数的励振コードブックサイズを増やすために使われている。L P モードを示すためには 1 つの情報ビットが必要であり、その情報ビットはパリティビットによって保護されている。8kbit/s から 11.8kbit/s へのビットレートの変更による本拡張において、上記の 2 ビット (L P モード表示ビット + パリティビット) を除く全付加ビットは、代数的コードブックのサイズを増やすために使われている。符号器におけるパラメータのビット割り当てを Table E - 2 - 1 / J T - G 7 2 9 に示す。

バックワード/フォワード判定基準によって音声 (主にフォワードモードで符号化される) と音楽 (主にバックワードモードで符号化される) との識別を可能にする。バックワード/フォワード処理は、切り替えの数を減らすように、また必要ときに不自然さを伴わないフィルタ間のスムーズな切り替えを行なえるように設計されている。L P モードとそれに関連した情報は、ポストフィルタリングや聴覚重み付けを音楽または音声のどちらかに対しより適合させるために使用されている。これは誤り補償のためにも使われている。

次節以降に、標準 J T - G 7 2 9 の 11.8kbit/s 拡張に関する概要記述が与えられている。ただしその中では、標準 J T - G 7 2 9 アルゴリズムに対する変更や追加のみを記述する。

Table E – 2 – 1 / JT-G729 Bit allocation of the 11.8 kbit/s CS-ACELP algorithm (10 ms frame)
 (ITU-T G729) The numbers of bits corresponding to modified parts of the structure
 (compared to JT-G729) are typed in bold.

LP mode indication bit	<i>Extension at 11.8 kbit/s</i>	
	1 + 1 (parity)	
	<i>Forward</i>	<i>Backward</i>
LP filter	18	0
LTP Delay (1 st /2 nd sub-fr.)	8 + 1 (parity)/5	8 + 1 (parity)/5
EXC Codes (1 st /2 nd sub-fr.)	35/35	44/44
Gains (LTP + EXC) (1 st /2 nd sub-fr.)	7/7	7/7
Total	118	118

E. 2. 1 符号器

伝送誤りに対し十分な耐性を維持し、定常性の低い信号や特に音声の劣化を防ぎつつ（標準 J T – G 7 2 9 で使われている純粋なフォワード構造と比較して）、音楽に対してこのような高品質を達成するために混合バックワード／フォワード L P 構造と呼ばれる新しい技術が導入されている。入力信号の定常性やバックワード、フォワードフィルタの予測利得から最適な L P 分析を選択できる。

音楽信号（一般的に非常に定常的である）に対しては、バックワード L P モードが主に使用されている。L P 分析は合成信号に対して処理され、係数の伝送を必要としない。この時、次の 2 つの利点をもっている。

- (1) L P の次数は 30 まで増され、複雑な音楽信号スペクトルに、より適合させた係数となっている（標準 J T – G 7 2 9 のような L P フォワードコーデックの 10 次の L P フィルタでは、音楽に対して十分でない）。
- (2) ビットレートがより適切に配分されている。すなわち、連続する非常に類似した L P フィルタに無駄なビットレートを費やすことはない。すべての余ったビットレートは、励振コードブックのサイズを拡張するために使用される。44 ビットの代数的コードブックは、固定コードブック励振信号のために使用される。

純粋のバックワード L P 分析の弱点は、主に急激なスペクトル遷移を伴う非定常な信号や伝送誤りに対する感度に関わっている。混合バックワード／フォワード L P 構造を用いれば、スペクトル遷移が生じたときにはフォワードモードが選択され、10 個の L P 係数は符号化され伝送される。さらに、たとえバックワードモードが優位であっても、純粋のバックワード構造と比較した場合に、フォワード L P フィルタ係数の伝送が明らかに耐性を改善している。

フォワードモードにおける符号器は、励振コードブックへより多くのビットが配分された標準 J T – G 7 2 9 とほとんど同等である。35 ビットの代数的コードブックが、固定コードブック励振信号のために使用される。

E. 2. 2 復号器

最初にパラメータインデックスは、受信したビットストリームから抽出される。これらのインデックスは、10ms 音声フレームに対応した符号器のパラメータを得るために復号される。復号された1番目のパラメータは、LPモード情報とそのパリティビットである。この情報に従って、フレームはフォワード、バックワード、消失のいずれかに分類される。フォワードモードにおけるパラメータは、LSP係数、2つの分数ピッチ遅延、2つのフォワード固定コードブックベクトル、2組の適応および固定コードブック利得である。バックワードモードにおけるパラメータは、2つの分数ピッチ遅延、2つのバックワード固定コードブックベクトル、2組の適応および固定コードブック利得である。まず、バックワードLP分析が行われる。そして、もしフレームがフォワードモードであったならば、LSP係数は補間され、各サブフレームに対するLPフィルタ係数へ変換される。固定コードブック励振信号の生成を除けば、復号処理は標準JTG729の復号処理とまさに同様である。さらに、各5msサブフレームで次のような処理が行われる。

- (1) それぞれの利得によってスケールされた適応および固定コードブックベクトルを加算することによって励振信号は生成される。
- (2) フォワードあるいはバックワードどちらか一方のLP合成フィルタを用い、励振信号をフィルタリングすることによって音声再生される。
- (3) 再生された音声信号は、次のような後処理段階を通る。この段階は、長期および短期合成フィルタに基づいた適応ポストフィルタを含んでおり、高域通過フィルタリングおよびスケール処理が後に続く。標準JTG729と比較すると、ポストフィルタの重み係数は適応的になっている。

E. 2. 3 遅延

標準JTG729の2.3節と同一である。

E. 2. 4 音声コーデックの記述

本付属資料では、音声符号化アルゴリズムをビットイグザクトな固定小数点数値演算で記述している。E.5章に示す本付属資料で必須なANSI Cコードは、このビットイグザクトな固定小数点記述を実現している。符号器(E.3章)と復号器(E.4章)の数学的な記述は他の方法でも実現可能であるが、そのような方法で実現されたコーデックは本付属資料に準拠しない可能性がある。よって、相違点がある場合、E.5章のANSI Cコードによるアルゴリズムの記述は、E.3章とE.4章における数学的な記述より優先される。十分なものではないが、このANSI Cコードとともに使用できるテスト信号のセットは、新日本ITU協会より入手可能である。

E. 3 符号器の機能記述

本章では、符号器の変更された機能を記述する。アルゴリズムの変更や追加が行われた部分を除いて、本章のほとんどの部分で標準JTG729本体が参照される。

E. 3. 1 前処理

標準JTG729の3.1節と同一である。

E. 3. 2 線形予測分析と量子化

10ms フレーム周期毎に2種類のLP分析を同時に行う。一方は、標準JTG729と同一の量子化手法を用いて、入力信号に対して行われる標準JTG729と同一のフォワード分析であり、他方は、過去の合成信号に対して行われるバックワード分析である。

E. 3. 2. 1 窓かけと自己相関計算

(1) フォワードLP分析

標準J T-G 7 2 9の3. 2. 1節と同一である。

(2) バックワードLP分析

標準J T-G 7 2 8と同一のハイブリッド巡回窓かけ手法を用いる。

サンプル1は最新の合成されたフレームの最新のサンプルとし、サンプル番号*i*は最も古いサンプルが最も大きな番号となるように並べられた過去のサンプルとする。

サンプル*i*=1,...,35は、窓の非巡回部分で窓かけされる。

$$w_{lpbwd}(i) = \sin(i * c_{lpbwd}), \quad i = 1, \dots, 35 \quad \text{where } c_{lpbwd} = 0.047783$$

窓の巡回部分は次の関数により与えられる。(サンプル > 35)

$$w_{lpbwd}(i) = b_{lpbwd} * a_{lpbwd}^{(i-36)}, \quad i > 35$$

$$\text{with } a_{lpbwd} = 0.9928337491 \quad \text{and } b_{lpbwd} = \sin(36 * c_{lpbwd})$$

自己相関係数の巡回計算は、標準J T-G 7 2 8に記載されているように実行される。

フォワードLP分析に関する限り同一の白色雑音補正係数(1.0001)が、最初の自己相関係数に適用されるが、係数に適用された帯域幅拡張は60Hz(標準J T-G 7 2 9)から5Hzに縮小される。

若干の付加的スペクトル平坦化は、(E. 3. 2. 2節で計算された)LP係数に対する $\gamma_{lpbwd} = 0.98$ の重み付け関数によって行われる。

E. 3. 2. 2 レビンソン・ダービンアルゴリズム

フォワード分析とバックワード分析で同一のアルゴリズムが用いられる。標準J T-G 7 2 9のアルゴリズムと比較すると、高次LPに対処するために配列の大きさが拡張されている。

E. 3. 2. 3 LPからLSPへの変換

フォワードLPフィルタについては、標準J T-G 7 2 9の3. 2. 3節と同一である。バックワードLPフィルタについては、LSPの計算が不要である。

E. 3. 2. 4 LSP係数の量子化

フォワードLSP係数については、標準J T-G 7 2 9の3. 2. 4節と同一である。バックワードLPフィルタについては、LSPの量子化が不要である。

E. 3. 2. 5 LP係数の補間

フォワードLP分析では、標準J T-G 7 2 9の3. 2. 5節に記載されているように、量子化された(および、量子化されていない)LP係数は、第2サブフレームで使用される。第1サブフレームでは、量子化された(および、量子化されていない)フォワードLSP係数は、前フレームがフォワードモード時には、標準J T-G 7 2 9の3. 2. 5節に記載されているように補間によって得られる。前フレームがバックワ

ードモード時には、補間が行われず、第2サブフレームの量子化された（および、量子化されていない）LPフィルタが第1サブフレームでも使用される。

バックワードLP分析では、第2サブフレームでは、E. 3. 2. 2節で計算された現在のバックワードLPフィルタ A_{bwd} 、あるいは、E. 3. 2. 7. 1節に示されるような遷移フィルタが使用される。

第1サブフレームでは、LPフィルタの係数は、第2サブフレームのバックワードLPフィルタと前フレームのフィルタとの間で、標準JT-G 7 2 9と同一の補間係数（0.5,0.5）を用いて直接的に補間される。

E. 3. 2. 6 LSPからLPへの変換

フォワードLPフィルタについては、標準JT-G 7 2 9の3. 2. 6節と同一である。バックワードLPフィルタについては、変換が不要である。

E. 3. 2. 7 バックワード/フォワード判定および切替え手順

E. 3. 2. 7. 1 切替え手順

本節では、合成信号が不自然になるのを避けるために、フォワード（または、バックワード）フィルタが用いられた1つ前のフレームから、バックワード（または、フォワード）フィルタが選択された現在のフレームへの切替え方法について記述する。

(1) フォワードLPフィルタからバックワードLPフィルタへの切替え

これは一般的に信号が定常的な場合に生じる。従って、合成信号上での、聴感的に感じられる不自然なスペクトル遷移を引き起こすようなフィルタ遷移を避けることが重要である。これを実現するため、符号器側と復号器側の双方にて下記に示す補間を行う。

フレーム n にて切替えが行われるとし、 $A_{fwd}(n-1)$ をフレーム $n-1$ におけるフォワードLPフィルタ、 $A_{bwd}(n)$ をE. 3. 2. 2節で算出された現在のフレーム n におけるバックワードLPフィルタとする。フレーム $n+i$ にて用いられるLPフィルタ A は次式により与えられる。

$$A(n+i) = 0.1 * i * A_{bwd}(n+i) + (1. - 0.1 * i) * A(n+i-1), \quad 0 \leq i \leq 9$$

$$A(n+i) = A_{bwd}(n+i), \quad i \geq 10 \quad \text{with } A(n-1) = A_{fwd}(n-1)$$

10 遷移フレーム後は、バックワードフィルタがそのまま使用される。

(2) バックワードフィルタからフォワードフィルタへの切替え

これは、入力信号にスペクトル遷移が存在する場合に生じる。その場合には平滑化は行われず、フレーム n にて切替えが行われるとした場合、以下とする。

$$A(n) = A_{fwd}(n)$$

E. 3. 2. 7. 2 グローバル定常性インジケータと高定常性インジケータ

フレーム n におけるグローバル定常性インジケータ（Stat(n)と呼ぶ）は、入力信号の全体的な定常性を特徴づけるものである。それは、フレーム n においてバックワード/フォワード判定が行われた後に算出され、フィルタ間の切替え回数を削減するため、フレーム毎に行われる次のフレーム $n+1$ におけるバックワード/フォワード判定に用いられる。原理としては、入力信号の定常性に応じて漸次一方のモードを選択し、他方のモードへの切替え回数を削減するものである。

このインジケータの算出は、バックワード/フォワード判定の履歴とバックワードおよびフォワードのフ

フィルタ予測利得に基づいて行われる。それは、入力信号に対する高い定常性を表す値（32000）から低い定常性を表す値（0）までの値をとる。このインジケータは、フレーム間でゆっくりと変化する（上記のとおり値の最小値から最大値までの変化をするのに最低 80 フレーム必要とする）。

以下に示す適応化は、エネルギーが 40dB を超えるフレームに対してのみ行われる。その他のフレームは無音フレームと考えられ、Stat(n)は上限値 13000 で Stat($n-1$)と等しい。

(1) 切替え判定結果に基づく適応化の第 1 段階

n を現在のフレーム番号とし、 $N_{\text{bwd}}(n)$ をフレーム n で測定された連続するバックワードフレームのフレーム数とする。フレーム n がフォワードフレームならば、 $N_{\text{bwd}}(n)$ は 0 とする。Stat¹(n) を第 1 ステップでの定常性評価結果の出力値とする。

- (a) もし、フレーム n がバックワードからフォワードへの遷移フレーム（すなわち、フレーム $n-1$ がバックワード、フレーム n がフォワード）であり、かつ連続したバックワードフレーム数が 20 未満であるならば、

$$\text{Stat}^1(n) = \text{Stat}(n-1) - (5000 - 250 * N_{\text{bwd}}(n-1))$$

- (b) もし、フレーム n がバックワードフレームであり、

- (7) もし ($N_{\text{bwd}}(n) > 20$) ならば、

$$\text{Stat}^1(n) = \text{Stat}(n-1) + 500$$

- (i) そうでなく、もし ($N_{\text{bwd}}(n) = 20$) ならば、

$$\text{Stat}^1(n) = \text{Stat}(n-1) + 2500$$

- (c) そうでなければ、

$$\text{Stat}^1(n) = \text{Stat}(n-1)$$

(2) 予測利得に基づく適応化の第 2 段階

x をバックワード LP フィルタ予測利得とフォワード LP フィルタ予測利得との差、すなわち、

$$x = G_{\text{pred}_b} - G_{\text{pred}_f} \text{ (in dB)}$$

とすると、Stat(n)は下式で表される。

$$\text{Stat}(n) = \text{Stat}^1(n) + \Delta(x)$$

ただし、 $\Delta(x)$ は以下とする。

もし、Stat¹(n) < 13000 ならば、

$$\Delta(x) = \begin{cases} 3200 & \text{if } x > 4 \\ 2400 & \text{if } x \in]3,4] \\ 1600 & \text{if } x \in]2,3] \\ 800 & \text{if } x \in]1,2] \\ 400 & \text{if } x \in]0,1] \\ 0 & \text{if } x \in [-1,0] \\ -400 & \text{if } x \in [-2,-1[\\ -800 & \text{if } x \in [-3,-2[\\ -1600 & \text{if } x \in [-4,-3[\\ -3200 & \text{if } x \in [-4.7,-4[\\ -6400 & \text{if } x < -4.7 \end{cases}$$

そうでなければ、

$$\Delta(x) = \begin{cases} 0 & \text{if } x \geq -1 \\ -400 & \text{if } x \in [-2,-1[\\ -800 & \text{if } x \in [-3,-2[\\ -1600 & \text{if } x \in [-4,-3[\\ -3200 & \text{if } x \in [-4.7,-4[\\ -6400 & \text{if } x < -4.7 \end{cases}$$

高定常状態は、パラメータ **High_Stat** の値が 1 に設定されると共に決定される。この高定常状態は、バックワードフレームの割合がフォワードフレームの割合に比べて非常に高くなった場合に検出される。

N_{bwd} (または、 N_{fwd}) が、過去の N_{fb} フレーム ($N_{\text{fb}} = N_{\text{bwd}} + N_{\text{fwd}}$) におけるバックワード (または、フォワード) フレームのフレーム数を表すとする。最初の 100 フレームに対しては、 N_{bwd} (または、 N_{fwd}) は過去 N_{fb} フレームにおける実際のバックワード (または、フォワード) フレームのフレーム数である。そして N_{fb} が 100 に達する度ごとに、 N_{fb} , N_{bwd} , N_{fwd} は 2 で除算される。そして、下記により高定常状態の判定を行う。

(a) もし、 $N_{\text{fb}} < 10$ ならば、

$$\text{High_Stat} = 0$$

(b) そうでなく、もし、 $N_{\text{bwd}} > 4 * N_{\text{fwd}}$ ならば、

$$\text{High_Stat} = 1$$

(c) そうでなければ、

$$\text{High_Stat} = 0$$

上記手順は、フレームエネルギーが 40dB を超えるフレームに対してのみ実行される。無音フレームに対しては、 N_{fb} , N_{bwd} , N_{fwd} および **High_Stat** は更新されない。

E. 3. 2. 7. 3 バックワード／フォワードの判定手順

現フレーム n において、バックワード／フォワードの判定は順次提示される 4 つの基準を基に行われる。

(1) 予測利得による第 1 の基準

バックワード LP フィルタ、補間されたバックワード LP フィルタ（フォワードからバックワードへの遷移過程においては、通常のバックワードフィルタとは異なる）、そしてフォワード LP フィルタの予測利得（dB 値）が計算される（それぞれの値を G_{pred_b} 、 $G_{\text{pred}_{\text{int}}}$ 、 G_{pred_f} とする）。

Gap を適応判定しきい値（dB 値）とする。

第 1 段階の判定は以下の手順で行われる。

次式で示される条件が成立した場合、バックワード LP フィルタが選択される。

$$(G_{\text{pred}_{\text{int}}} > G_{\text{pred}_f} - \text{Gap}) \text{ and } (G_{\text{pred}_b} > G_{\text{pred}_f} - \text{Gap}) \text{ and } (G_{\text{pred}_b} > 0) \text{ and } (G_{\text{pred}_{\text{int}}} > 0)$$

その他の場合は、フォワード LP フィルタが選択される。

Gap 変数は定常性インジケータに従って、以下のように更新される（E. 3. 2. 7. 2 節 参照）。

$$\text{Gap}(n) = 0.0366 * (\text{Stat}(n-1)/320) + 1.0 \quad \text{Stat}(n-1) \in [0, 32000]$$

(2) グローバル定常性インジケータを用いた第 2 の基準

グローバル定常性インジケータ $\text{Stat}(n-1) \in [0, 32000]$ が 13000 以下の値を維持している間は、フォワード LP モードが選択される（第 2 段階の判定）。この基準は、低あるいは中定常性の音声や他の信号による、バックワードモードへの不必要な切替を回避する。

(3) LSP による第 3 の基準

短時間スペクトラムが安定している期間での不自然な遷移を回避するため、以下のように 2 つの連続したフォワード LP フィルタの LSP ベクトル間のユークリッド距離が計算される。次式に 2 つのベクトル間のユークリッド距離を示す。

$$d_{\text{LSP}}(n) = \| \mathbf{LSP}_n, \mathbf{LSP}_{n-1} \|^2$$

ここで、 \mathbf{LSP}_n は現フレーム n において計算されたフォワード LP フィルタの LSP ベクトルであり、 \mathbf{LSP}_{n-1} はフレーム $n-1$ において計算されたフォワード LP フィルタの LSP ベクトルである。

もし $d_{\text{LSP}}(n) < \text{Thresh}_{\text{LSP}}(n)$ であり、前フレームがバックワードモードであり、予測利得 G_{pred_b} と $G_{\text{pred}_{\text{int}}}$ が正の値であれば、バックワードからフォワードへの切替は禁止される（この場合、第 3 段階の判定としてバックワードモードが選択される）。

$\text{Thresh}_{\text{LSP}}$ は、 $\text{Stat}(n-1)$ の値に従って、フレーム毎に以下のように更新される。もし、 $\text{Stat}(n-1) = 32000$ （最大値）ならば、 $\text{Thresh}_{\text{LSP}}(n) = 0.03$ とし、それ以外では $\text{Thresh}_{\text{LSP}}(n) = 0$ とする。

(4) エネルギーによる第 4 の基準

伝送誤りに対するアルゴリズムの耐性を向上するために、エネルギーが 40dB 以下のフレームでは、フォワード LP フィルタが用いられる。

E. 3. 3 聴覚重み付け

聴覚重み付けフィルタは以下の式で与えられる。

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

(1) フォワードモードの場合

パラメータ γ_1 と γ_2 は標準 J T-G 7 2 9 の 3. 3 節に記載されているように算出される。LP フィルタ $A(z)$ は、High_Stat が 0 の場合、量子化されていないフォワードフィルタ $A_{fwd}(z)$ で、それ以外では量子化されたフォワードフィルタである。

(2) バックワードモードの場合

High_Stat が検出されない場合、LP フィルタ $A(z)$ は量子化されていないフォワードフィルタを用い、それ以外ではバックワードに計算されたフィルタである。

パラメータ γ_1 と γ_2 とは、高定常性インジケータ (High_Stat) によって定まる固定値をとる。

入力信号が高定常性な場合 (High_Stat=1)、ノイズマスキング効果が強調され

$$\begin{aligned}\gamma_{1bwdh} &= 0.98 \\ \gamma_{2bwdh} &= 0.4\end{aligned}$$

となる。

標準的な定常性 (High_Stat=0) では、

$$\begin{aligned}\gamma_{1bwdl} &= 0.9 \\ \gamma_{2bwdl} &= 0.4\end{aligned}$$

となる。

(3) 重み付き音声は標準 J T-G 7 2 9 の 3. 3 節の式(33)に示されるように算出され、フィルタの次数は選択された重み付けフィルタによって選択される (10 または 30)。

E. 3. 4 開ループピッチ分析

標準 J T-G 7 2 9 の 3. 4 節と同一である。

E. 3. 5 インパルス応答の計算

標準 J T-G 7 2 9 の 3. 5 節と同様である (LP フィルタの次数が、10 次の代わりに 30 次になる)。

E. 3. 6 ターゲット信号の計算

標準 J T-G 7 2 9 の 3. 6 節と同様である (LP フィルタの次数が、10 次の代わりに 30 次になる)。

E. 3. 7 適応コードブック探索

適応コードブック探索、適応コードブックベクトルの生成、遅延インデックス $P1$ 、 $P2$ に対する符号語の算出、適応コードブック利得の算出は、標準 J T-G 7 2 9 の 3. 7 節に記述されている手順と同一である。パリティビット $P0$ は、第 1 サブフレームの遅延インデックス $P1$ の上位 7 ビット (標準 J T-G 7 2 9 の 3. 7 節では 6 ビット) で算出される。

E. 3. 8 固定コードブックの構造と探索

E. 3. 8. 1 フォワードLPモードにおける固定コードブック

フォワードLPモードでは、35ビットの代数的コードブックが固定コードブックとして用いられる。このコードブックでは、それぞれの励振ベクトルが10本の零でないパルスを含んでいる。パルスの振幅は-1もしくは+1である。それぞれのサブフレームにおける40の位置は、2本のパルスを含んだ5つのトラックに分割される。この設計では、それぞれのトラックに対する2本のパルスが重なって、結果として振幅が-2もしくは+2の1本のパルスになる可能性がある。パルスの許された位置をTable E-3-1/JT-G729に示す。

Table E-3-1/JT-G729 Structure of fixed codebook in forward mode C_{fwd}
(ITU-T G.729)

Track	Pulses	Signs	Positions
1	p_0, p_1	$s_0, s_1: \pm 1$	0,5,10,15,20,25,30,35
2	p_2, p_3	$s_2, s_3: \pm 1$	1,6,11,16,21,26,31,36
3	p_4, p_5	$s_4, s_5: \pm 1$	2,7,12,17,22,27,32,37
4	p_6, p_7	$s_6, s_7: \pm 1$	3,8,13,18,23,28,33,38
5	p_8, p_9	$s_8, s_9: \pm 1$	4,9,14,19,24,29,34,39

標準JT-G729本体と同様に選択されたコードブックベクトルは調波構造を強調するために次のプリフィルタでフィルタリングされる。

$$P(z) = 1 / (1 - \beta z^{-T})$$

β を適応させる方法は標準JT-G729本体と同様である。

E. 3. 8. 1. 1 35ビットコードブックの探索手順

固定コードブックは重み付き入力音声と重み付き再生音声の平均自乗誤差を最小にすることによって探索される。 $c_k(n)$ はインデックス k における代数的コードベクトル、 $h(n)$ は重み付き合成フィルタのインパルス応答、そして $d(n)$ はターゲットベクトルと $h(n)$ の相関であるとする、代数的コードブックは次式で表される基準が最大になるように探索される。

$$T_k = \frac{(C_k)^2}{E_k}$$

ここで、 C は $c_k(n)$ と $d(n)$ の相関であり、そして E はフィルタされた符号ベクトル($c_k(n) * h(n)$)のエネルギーである。代数的コードベクトルは零以外のパルスをほとんど含んでいないので相関は次式で表すことができる。

$$C = \sum_{i=0}^{N_p-1} s_i d(m_i)$$

ここで、 m_i は i 番目のパルスの位置、 s_i は振幅、そして N_p ($N_p = 10$) はパルス数を示す。そして、分母のエネルギーは次式で与えられる。

$$E = \sum_{i=0}^{N_p-1} \phi(m_i, m_i) + 2 \sum_{i=0}^{N_p-2} \sum_{j=i+1}^{N_p-1} s_i s_j \phi(m_i, m_j)$$

ここで、 $\phi(i, j)$ は $h(n-i)$ と $h(n-j)$ の相関を含む。信号 $d(n)$ と相関 $\phi(i, j)$ はコードブック探索の前に計算される。

標準 J T-G 7 2 9 本体と同様に、探索手順を高速化するために閉ループ探索の外側で、いわゆるパルス振幅信号選択アプローチによってパルス振幅があらかじめ設定される。このアプローチでは、ある位置において発生しているパルスの、最も可能性の高い振幅がある補助情報信号を使って推定される。標準 J T-G 7 2 9 本体では、信号 $d(n)$ がパルスの振幅をあらかじめ選択するために使用される。本ビットレート拡張では、正規化された $d(n)$ ベクトルと正規化された長期予測残差の重み付き合計である信号 $b(n)$ が使用される。信号 $b(n)$ は次式で与えられる。

$$b(n) = d(n) / \sigma_d + e(n) / \sigma_e$$

ここで、 $e(n)$ は長期予測残差であり、 σ_d と σ_e はそれぞれ $d(n)$ と $e(n)$ の実効値である。ある位置におけるパルスの極性は先験的に対応する位置の $b(n)$ の極性と同一にされる。極性情報は標準 J T-G 7 2 9 本体と同様に最適なパルス位置の探索を開始する前に信号 $d(n)$ と $\phi(i, j)$ に組み入れられる。

最適なパルス位置は限定的な「合成による分析」探索手順を用いて決定される。使用された手順は、妥当な複雑さで巨大なコードブックを効率的に探索する一般的な縦形木探索の特殊なケースである。このアプローチで、 N_p 本の励振パルスは N_m 本のパルスによる M 個のサブセットに分割される。探索はサブセット 1 から開始し、木の m 番目のレベルでサブセット m が探索され、木構造に従って次のサブセットに進む。探索はパルスが置かれるトラックに割り当てられる順番を変更することによって繰り返される。この特殊なコードブック構造においてパルスは 2 本のパルスからなる 5 つのサブセット（木は 5 つのレベルを持つ）に分割される。

パルスの位置は以下のように決定される。

5 つのトラックそれぞれについて、 $d(n)$ の絶対値を最大にするパルス位置が見つけれられる。これらから、最も大きな結合の最大値を持つ 2 つの連続したトラック T_{k_0} と $T_{(k_0+1) \bmod 5}$ が決定される。ここでインデックス k_0 は最初のトラックに対するパルスの割り当てのために使用される。次に 2 番目に大きな結合の最大値を持つ 2 つの連続したトラック T_{k_1} と $T_{(k_1+1) \bmod 5}$ および 3 番目に大きな結合の最大値を持つ 2 つの連続したトラック T_{k_2} と $T_{(k_2+1) \bmod 5}$ が同様に決定される。

1 回目の繰り返しで、パルスは次のようにトラックに割り当てられる。パルス i_n , $n = 0, \dots, 9$ をトラック $T_{(k_0+n) \bmod 5}$, $n = 0, \dots, 9$ にそれぞれ割り当てる。

パルスは 2 本のパルスのサブセットから探索される。トラック T_{k_0} の最大値にパルス i_0 を、そしてトラック $T_{(k_0+1) \bmod 5}$ の最大値にパルス i_1 をセットすることから始める。そして次にトラック $T_{(k_0+2) \bmod 5}$ と $T_{(k_0+3) \bmod 5}$ での 8×8 の可能な位置の組み合わせ（パルス i_0 と i_1 は既知として与えられている）全てをテストすることによってパルス対 (i_2, i_3) を探索する。同様の手順が残りのパルス対 (i_4, i_5) , (i_6, i_7) および (i_8, i_9) のために、それらのそれぞれのトラックでの 8×8 の可能な位置の組み合わせをテストすることにより繰り返される。木のそれぞれのレベルにおいて、テスト基準はそのレベルにおいて利用可能なパルスのみに基づいて計算される。これは（最初のパルス対がそれらの最大値にセットされた時から）合計で $4 \times 8 \times 8$ の位置をテストした結果になる。

あと2回の繰り返しがトラックへのパルスの割り当てを変更することによって実行される（2回目の繰り返してで k_1 によって k_0 を、そして3回目の繰り返してで k_2 によって k_0 を置き換える）。10本全ての初期パルス位置が、2番目の繰り返してでトラック $T_{(k_1+n) \bmod 5}$ に、そして3番目の繰り返してでトラック $T_{(k_2+n) \bmod 5}$ に割り当てられる。上記と同一の探索手順が、これらあと2回の繰り返しのために繰り返される。3回の繰り返してに対して、テストされた位置の組み合わせの総数は $3 \times 4 \times 8 \times 8 = 768$ である。

E. 3. 8. 1. 2 35ビット固定コードブック符号語の計算

それぞれのトラックにおける2本のパルス位置は6ビットで符号化される。そしてそれぞれのトラックの最初のパルスの極性は1ビットで符号化される。2番目のパルスの極性はパルス位置の順序に基づいて無条件に決定される。それぞれのトラックでの2本のパルス（2つの位置と2つの極性）は7ビットで符号化される。それぞれのパルス位置には3ビット（8つの可能な位置）を必要とし、そしてそれぞれの極性には1ビットを必要とする。それぞれのパルス対に対して合計で8ビットになる。しかしながら、位置の組み合わせのおよそ半分が重複する事実を考慮すると、1ビット削減することが出来る。例えば、（極性が考慮されない場合）パルス1を位置aに置きパルス2を位置bに置くことは、パルス1をbに置きパルス2をaに置くことに等しい。パルスの符号化を実現することについての単純なアプローチは、ある意味で他方の極性情報が容易に推測できることから、極性情報に1ビットのみを使用し、2つの位置のために6ビットを使用することである。より良い説明のために、トラックに、位置 $p1$ と $p2$ にそれぞれ極性インデックス $s1$ と $s2$ （極性が正の場合 $s=0$ 、極性が負の場合 $s=1$ ）の2本のパルスがあると仮定する。2本のパルスのインデックスは次式で与えられる。

$$I = (p1/5) + s1 \times 8 + (p2/5) \times 16$$

もし $p1 \leq p2$ ならば、 $s2 = s1$ となり、さもなければ $s2$ は $s1$ と異なる値となる。従って、符号語を構成するとき、もし2本の極性が同じならばより小さな位置を $p1$ に、より大きな位置を $p2$ に割り当て、さもなければより大きな位置を $p1$ に、より小さな位置を $p2$ に割り当てる。

この手順は5つの7ビットのインデックスを得るまで繰り返される。

E. 3. 8. 2 バックワードLPモードにおける固定コードブック

バックワードLPモードにおいて、LPモデルに必要とされる18ビットは伝送されない。各サブフレームで節約される9ビットは、固定コードブックサイズを35ビットから44ビットへ増加させるために使用される。この44ビットのコードブックでは、各々のコードブックベクトルが12本のパルスを含む。サブフレームの位置は、Table E-3-1/JT-G729に記載されているのと同じトラック構造に分けられる。しかしながら2本のパルスが増えたので、2つの連続したトラックは2本ではなく3本のパルスを含むことができるように配置される。3本のパルスを含む連続した2つのトラックはトリプルパルストラックと呼ばれ、他方の2本のパルスを含む3つのトラックはダブルパルストラックと呼ばれる。

各ダブルパルストラック中のパルスは7ビット（35ビットコードブックと同様）で符号化され、各トリプルパルストラック中のパルスは10ビットで符号化される。最初のトリプルパルストラックのインデックスは5つの異なった値（5トラック）をとりえるため、さらに3ビットを必要とする。この結果、合計44ビットとなる（ $3 \times 7 + 2 \times 10 + 3$ ）。

E. 3. 8. 2. 1 44ビットコードブックの探索手順

コードブック探索は、木が6レベルのパルス対を持つことを除いて、35ビットコードブック探索と同様に行われる。E. 3. 8. 1. 1節に記述されている探索手順と同様である。

同様の手順をパルスの極性の前設定に使う。

初期トラック T_k と T_{k+1} は同様の方法で決定する。

12 のパルス $i_n(n=0,\dots,11)$ は、トラック $T_{(k+n)\bmod 5}(n=0,\dots,11)$ にそれぞれ割り当てられる。

パルス i_0 をトラック T_k の最大値に、パルス i_1 をトラック $T_{(k+1)\bmod 5}$ の最大値に初期設定することにより、パルスは、これら 2 本のパルスのサブセット中で探索される。そしてトラック $T_{(k+2)}$ と $T_{(k+3)}$ における 8×8 全ての可能な組み合わせを試すことでパルス対 (i_2, i_3) の探索を進める。その後はパルス対 (i_4, i_5) 、 (i_6, i_7) 、 (i_8, i_9) そして (i_{10}, i_{11}) にこの手順を繰り返し行う。その結果 $5 \times 8 \times 8$ 全ての組み合わせが試される。

あと 2 回の繰り返しにより 35 ビットコードブックと同様に、合計で $3 \times 5 \times 8 \times 8 = 960$ 回の試行を行った結果が得られる。

標準 J T-G 7 2 9 および 35 ビットフォワードコードブックと同様に、選択されたコードブックベクトルは調波構造を強調するために次のプリフィルタでフィルタリングされる。

$$P(z) = 1 / (1 - \beta z^{-T})$$

E. 3. 8. 2. 2 44 ビット固定コードブック符号語の計算

3 つのダブルパルストラックの各々の 2 本のパルスは、E. 3. 8. 1. 2 節で記述されているアプローチを用いて同様に符号化される。

トリプルパルストラックの 3 本のパルスは、3 番目のパルスの位置のために 3 ビットを加え、同じ考え方を適用することによって符号化される。3 つの位置は各々 3 ビットで符号化され、最初のパルスの極性は 1 ビットで符号化される。他の 2 本のパルスの極性は、ダブルパルストラックと同様にパルスの順序から演算する。もう一度、例を用いて説明を行う。トリプルパルストラックに含まれる 3 本のパルスが、各々極性インデックス s_1 , s_2 , s_3 を持って位置 p_1 , p_2 , p_3 に配置していると仮定する。3 本のパルスのインデックスは次式により与えられる。

$$I = (p_1/5) + s_1 \times 8 + (p_2/5) \times 16 + (p_3/5) \times 128$$

もし $p_1 \leq p_2$ ならば、 $s_2 = s_1$ となり、さもなければ s_2 は s_1 と異なる値となる。同様に、もし $p_2 \leq p_3$ ならば、 $s_3 = s_2$ となり、さもなければ s_3 は s_2 と異なる値となる。符号語を構成するとき、トラック中のパルス位置は、このパルスの極性の関係を考慮にいれて p_1 , p_2 , p_3 に割り当てられる。

各々のトラックに 1 つずつ、合計で 5 つのインデックスが返される。最初のインデックスは第 1 トリプルパルストラックのインデックスである。このインデックスは 13 ビットで符号化される。10 ビットは上述のように位置と極性のためのものであり、残り 3 ビットはトラックインデックス (0 から 4 までの値をとる) のためのものである。2 番目のインデックスは第 2 トリプルパルストラックのインデックスであり、10 ビットで符号化される。最後の 3 つのインデックスは 3 つのダブルパルストラックのインデックスであり、各々 7 ビットで符号化される。

E. 3. 9 利得の量子化

標準 J T-G 7 2 9 の 3. 9 節と同一である。

E. 3. 10 メモリの更新

標準 J T-G 7 2 9 の 3. 10 節と同一である。

E. 4 復号器の機能記述

最初にパラメータを復号する。伝送されるパラメータのリストを Table E-4-1/JT-G729 に示す。最初に復号されるパラメータはLPモード情報とそのパリティビットである。この情報により、そのフレームがフォワードかバックワードか、消失されたかを分類する。フォワードモードにおいて復号されるパラメータは、LSP係数、2つの分数ピッチ遅延、2つのフォワード固定コードブックベクトル、そして2組の適応および固定コードブック利得である。また、バックワードモードにおいて復号されるパラメータは、2つの分数ピッチ遅延、2つのバックワード固定コードブックベクトル、そして2組の適応および固定コードブック利得である。バックワードLP分析は過去の合成信号により実行され、復号されたパラメータはE. 4. 1節に記述されるように再生音声信号の算出に使用される。この再生音声信号は、E. 4. 2節に記述されるようにポストフィルタ、高域通過フィルタ、そしてアップスケーリングによって構成される後処理操作によって品質の向上がなされる。パリティ誤りが生じた場合、またはフレーム消失フラグがセットされた場合に使われる誤り補償手順についてはE. 4. 4節に記述する。

Table E – 4 – 1 / JT-G729 Description of transmitted parameters indices - The bit stream ordering is reflected by the

(ITU-T G.729) order in the table - For each parameter the Most Significant Bit(MSB) is transmitted first

Table E – 4 – 1 A / JT-G729 Description of transmitted parameters in forward mode
(ITU-T G.729)

<i>Symbol</i>	<i>Description</i>	<i>Bits</i>
<i>M0</i>	Switch LP mode	1
<i>M1</i>	Parity bit for LP mode	1
<i>L0</i>	Switched MA predictor of LSP quantizer	1
<i>L1</i>	First stage vector of quantizer	7
<i>L2</i>	Second stage lower vector of LSP quantizer	5
<i>L3</i>	Second stage higher vector of LSP quantizer	5
<i>P1</i>	Pitch delay first subframe	8
<i>P0</i>	Parity bit for pitch delay	1
<i>C0_1</i>	Track 0 fixed codebook first subframe	7
<i>C1_1</i>	Track 1 fixed codebook first subframe	7
<i>C2_1</i>	Track 2 fixed codebook first subframe	7
<i>C3_1</i>	Track 3 fixed codebook first subframe	7
<i>C4_1</i>	Track 4 fixed codebook first subframe	7
<i>GA1</i>	Gain codebook (stage 1) 1 st subframe	3
<i>GB1</i>	Gain codebook (stage 2) 1 st subframe	4
<i>P2</i>	Pitch delay second subframe	5
<i>C0_2</i>	Track 0 fixed codebook second subframe	7
<i>C1_2</i>	Track 1 fixed codebook second subframe	7
<i>C2_2</i>	Track 2 fixed codebook second subframe	7
<i>C3_2</i>	Track 3 fixed codebook second subframe	7
<i>C4_2</i>	Track 4 fixed codebook second subframe	7
<i>GA2</i>	Gain codebook (stage 1) 2 nd subframe	3
<i>GB2</i>	Gain codebook (stage 2) 2 nd subframe	4

Table E-4-1 B/JT-G729 Description of transmitted parameters indices in backward mode
(ITU-T G.729)

<i>Symbol</i>	<i>Description</i>	<i>Bits</i>
<i>M0</i>	Switch LP mode	1
<i>M1</i>	Parity bit for LP mode	1
<i>P1</i>	Pitch delay first subframe	8
<i>P0</i>	Parity bit for pitch delay	1
<i>C0_1</i>	Fixed codebook track index + pluses 0, 5 and 10 first subframe	13
<i>C1_1</i>	Fixed codebook pulses 1, 6 and 11 first subframe	10
<i>C2_1</i>	Fixed codebook pulses 2 and 7 first subframe	7
<i>C3_1</i>	Fixed codebook pulses 3 and 8 first subframe	7
<i>C4_1</i>	Fixed codebook pulses 4 and 9 first subframe	7
<i>GA1</i>	Gain codebook (stage 1) 1 st subframe	3
<i>GB1</i>	Gain codebook (stage 2) 1 st subframe	4
<i>P2</i>	Pith delay second subframe	5
<i>C0_2</i>	Fixed codebook track index + pluses 0, 5 and 10 second subframe	13
<i>C1_2</i>	Fixed codebook pulses 1, 6 and 11 second subframe	10
<i>C2_2</i>	Fixed codebook pulses 2 and 7 second subframe	7
<i>C3_2</i>	Fixed codebook pulses 3 and 8 second subframe	7
<i>C4_2</i>	Fixed codebook pulses 4 and 9 second subframe	7
<i>GA2</i>	Gain codebook (stage 1) 2 nd subframe	3
<i>GB2</i>	Gain codebook (stage 2) 2 nd subframe	4

E. 4. 1 パラメータ復号手順

標準 J T - G 7 2 9 と同様であるが、パラメータの数は多い（励振コードブックのパラメータが多くなっており、また L P モードを示すパラメータが 1 ビット増えている）。復号処理は以下の順に行われる。

E. 4. 1. 1 フォワード/バックワード復号手順

L P モード、すなわちフォワード/バックワードのどちらで行うかを復号器に示すために、1 ビットが割り当てられている。そして、L P モードのためのパリティビットと L P モードビットが比較される。もし、これらのビットが一致していなければ、そのフレームはフレーム消失とみなされ、E. 4. 4 節に記述されている手順が適用される。一致していれば、この L P モードに従って E. 3. 2. 7 節に記述されている処理と同じ切り替え手順が、合成フィルタに使用される L P フィルタを得るために復号器にて実行される。

高定常性インジケータ $High_Stat(n)$ は、E. 3. 2. 7 節に記述されているように、フレーム毎に一度計算される。

フレーム消失時に利得減衰の手順に使用されるもう 1 つの高定常性インジケータ $High_Stat2$ は、サブフレーム毎に計算される（E. 4. 3. 3 節参照）。現在のサブフレームが、少なくともバックワード処理のサブフレームが 30 回続いた後のサブフレームであるならば、 $High_Stat2$ は 1 にセットされ、それ以外の場合には 0 にセットされる。

E. 4. 1. 2 LPフィルタパラメータの復号

E. 4. 1. 2. 1 バックワードLPフィルタの計算

いかなるLPモード（フォワードまたはバックワード）においても、また、消失フレーム（E. 4. 4節参照）の場合でさえ、フレーム毎に一度バックワードLP分析が実行され、その処理は符号器のバックワードLPフィルタを得るために符号器で実行されるE. 3. 2節の処理（窓かけ処理や自己相関計算、レビンソン・ダービンアルゴリズム）と同じ手順を使用する。

E. 4. 1. 2. 2 フォワードモード

フォワードモードでは、標準JT-G 729で適用されるLPパラメータの復号手順と同一の手順を行う。LPフィルタ係数の補間手順はE. 3. 2. 5節の記述と同一である。

E. 4. 1. 2. 3 バックワードモード

過去のフレームの1つに消失フレームが含まれている場合には、E. 4. 1. 2. 1節で計算される現在のバックワードフィルタ $A_{bwd}^{(current)}$ は直接使用されず、最後に正しかったバックワードフィルタ（E. 4. 4節参照）の線形補間を行って使用する。この処理は、E. 3. 2. 5節に記述されているLPフィルタ係数の補間手順に先だって行われる。

E. 4. 1. 3 適応コードブック遅延のパリティビットの計算

励振信号が再生される前に、適応コードブックの遅延インデックス $P1$ （E. 3. 7節参照）からパリティビットが計算される。このビットが、伝送されてきたパリティビット $P0$ に等しくないならば、伝送中にビット誤りが起こったと見なす。 $P1$ にパリティ誤りが発生した場合には、遅延値 T_1 は前サブフレームで計算された遅延値に置き換えられる（E. 4. 4. 5節参照）。

E. 4. 1. 4 適応コードブックベクトルの復号

標準JT-G 729の4. 1. 3節と同一である。

E. 4. 1. 5 固定コードブックベクトルの復号

受信したコードブックインデックスはパルスの位置と極性を抽出するために使用される。これは、35ビットコードブックの時はE. 3. 8. 1. 2節、44ビットコードブックの時はE. 3. 8. 2. 2節に記述されている処理の逆の手順により実行される。パルス位置と極性が復号されると、固定コードブックベクトル $c(n)$ は以下の式により生成される。

$$c(n) = \sum_{i=0}^{N_p-1} s_i \delta(n - p_i)$$

ここで s_i はパルスの極性であり、 p_i はパルス位置である。そして N_p はパルスの数（10または12）である。ピッチ遅延の整数部がサブフレーム長40よりも小さい場合、 $c(n)$ は標準JT-G 729の式(48)と同様に補正される。

E. 4. 1. 6 適応および固定コードブック利得の復号

標準JT-G 729の4. 1. 5節と同一である。

E. 4. 1. 7 再生音声信号の計算

標準 J T-G 7 2 9 の 4. 1. 6 節と同様である (LP フィルタ係数の次数が、10 次の代わりに 30 次になる)。

E. 4. 2 後処理

標準 J T-G 7 2 9 と同様に、後処理は適応ポストフィルタ、高域通過フィルタ、信号のアップスケーリングの 3 つの機能から構成される。適応ポストフィルタは、高定常性インジケータ **High_Stat** と現在のフレームの LP モードとに応じて適応的に作られる係数 γ_p 、 γ_n および γ_d を除けば、標準 J T-G 7 2 9 のポストフィルタと同様である。高定常なバックワードフレームが 20 フレーム連続した後ではポストフィルタはかからなくなる。

E. 4. 2. 1 長期ポストフィルタ

長期ポストフィルタの動作は適応フィルタの高定常性インジケータ **High_Stat** と現在のフレームの LP モードに応じて適応的に作られる係数 γ_p の値を除いては、標準 J T-G 7 2 9 の 4. 2. 1 節と同一である。

$$H_p(z) = \frac{1}{1 + \gamma_p g_l} (1 + \gamma_p g_l z^{-T})$$

もし入力信号に高定常状態が検出される (**High_Stat**=1) と長期聴覚フィルタは徐々に平坦化される。フレーム n において、**High_Stat**=1 であり、かつ、フレームがバックワード状態であれば、

$$\begin{aligned} \gamma_p(n) &= \gamma_p(n-1) - (\gamma_{p \max} / 20) \\ \text{if } (\gamma_p(n) < 0) \text{ then } \gamma_p(n) &= 0 \end{aligned}$$

そうでなければ、フィルタは徐々に初期値 $\gamma_{p \max}$ に復帰していく。

$$\begin{aligned} \gamma_p(n) &= \gamma_p(n-1) + (\gamma_{p \max} / 20) \\ \text{if } (\gamma_p(n) > \gamma_{p \max}) \text{ then } \gamma_p(n) &= \gamma_{p \max} \end{aligned}$$

$\gamma_{p \max}$ の値は 0.25 に設定される。 $\gamma_p(n)$ が零と等しいときは、長期も短期もどちらの適応ポストフィルタもかからなくなる。

E. 4. 2. 2 短期ポストフィルタ

標準 J T-G 7 2 9 のアルゴリズムに関する変更箇所は、以下の項目のみである。

- (1) 短期聴覚重み付けフィルタ $H_f(z)$ の計算に用いられる LP フィルタは、E. 4. 1. 2 節で計算される LP フィルタである。このフィルタの係数は、フレームがフォワードモードのときは E. 4. 1. 2. 2 節で計算されるフォワード LP フィルタの 10 個の係数であり、フレームがバックワードモードのときは E. 4. 1. 2. 3 節で計算されるバックワード LP フィルタの 30 個の係数である。

$$H_f(z) = \frac{A(z/\gamma_n)}{A(z/\gamma_d)} = \frac{1 + \sum_{i=1}^{i=m_{b/f}} \gamma_n^i a_i z^{-i}}{1 + \sum_{i=1}^{i=m_{b/f}} \gamma_d^i a_i z^{-i}}$$

(2) 係数 γ_n と γ_d の値は高定常性インジケータ **High_Stat** (E. 4. 1. 1 節参照) と現フレームの LP モードがバックワードかフォワードかに応じて更新される。

もし入力信号により高定常状態が検出され (**High_Stat**=1)、現フレームがバックワードモードであれば、短期 LP ポストフィルタは徐々にポストフィルタを全くかけない方向 ($\gamma_n(n) = \gamma_d(n) = 0$) に平坦化されていく。

もしフレーム n が、 (**High_Stat** = 1 and **LP_mode**= 1) であれば、

$$\gamma_n(n) = \gamma_n(n-1) - (\gamma_{n \max} / 20)$$

$$\gamma_d(n) = \gamma_d(n-1) - (\gamma_{d \max} / 20)$$

$$\text{if } (\gamma_n(n) < 0) \text{ then } \gamma_n(n) = 0$$

$$\text{if } (\gamma_d(n) < 0) \text{ then } \gamma_d(n) = 0$$

そうでなければ、フィルタは初期値 $\gamma_{n \max}$ および $\gamma_{d \max}$ に徐々に復帰していく。

$$\gamma_n(n) = \gamma_n(n-1) + (\gamma_{n \max} / 20)$$

$$\gamma_d(n) = \gamma_d(n-1) + (\gamma_{d \max} / 20)$$

$$\text{if } (\gamma_n(n) > \gamma_{n \max}) \text{ then } \gamma_n(n) = \gamma_{n \max}$$

$$\text{if } (\gamma_d(n) > \gamma_{d \max}) \text{ then } \gamma_d(n) = \gamma_{d \max}$$

ここで、 $\gamma_{n \max} = 0.7$ であり、 $\gamma_{d \max} = 0.65$ である。

E. 4. 2. 3 傾き補償

傾き補償フィルタはインパルス応答長を 20 の代わりに 32 で 1 次偏自己相関係数を計算することを除いて標準 JT-G 729 の 4. 2. 3 節と同一である。

E. 4. 2. 4 適応利得制御

標準 JT-G 729 の 4. 2. 4 節と同一である。

E. 4. 2. 5 高域通過フィルタとアップスケーリング

標準 JT-G 729 の 4. 2. 5 節と同一である。

E. 4. 3 符号器と復号器の初期化

Table 4-2/JT-G729 と Table E-4-2/JT-G729 に示す変数を除き符号器と復号器の全ての静的変数は零に初期化される。

Table E – 4 – 2 / JT-G729 Description of parameters with non-zero initialization
(ITU-T G.729)

Variable	Reference	Initial value
Stat(-1)	E.3.2.7.2	10 000
$\gamma_p(-1)$	E.4.2.2	0.25
$\gamma_n(-1)$	E.4.2.2	0.7
$\gamma_d(-1)$	E.4.2.2	0.65
$\alpha_g^{(-1)}$	E.4.4.3	1.
$T_{sav}^{(-1)}$	E.4.4.5	30

E. 4. 4 フレーム消失補償

基本的に、不正フレーム補償の手順は標準 JT-G 7 2 9 の 4. 4 節と同様である。標準 JT-G 7 2 9 におけるものと同じ有声／無声の判定が用いられる。但し、利得の減衰の手順には、減衰係数を適応的に変えるため高定常性インジケータ High_Stat2 を考慮するという改良が加えられている。また、フレーム消失に対するバックワードフィルタの耐性を向上させるために、特別な手順が追加されている。

消失フレームに対して実行される特別な処理を以下に示す。

- (1) LPモードの反復
- (2) フォワードモードの場合には合成フィルタパラメータの反復、バックワードモードの場合には E. 4. 4 節に記述されている第2段バックワードLPフィルタの使用
- (3) 適応および固定コードブック利得の減衰
- (4) 利得予測器のメモリの減衰
- (5) 置換用励振信号の生成

E. 4. 4. 1 LPモードの反復

フレームが消失した時、LPモードに直前のフレームのLPモードを設定する。初期値は、0（フォワードモード）に設定される。

E. 4. 4. 2 合成フィルタパラメータの計算

まず、たとえフレームが消失しても、E. 4. 2. 1 節に記述されているバックワードLP分析は常に実行される点に留意のこと。

バックワードフィルタの耐性を改善するために、各フレームにおいて、第2段バックワードフィルタが計算される。このフィルタは、エラーフリー状態ではバックワードフィルタに等しいが、現フレームよりも以前にいくつかの消失フレームがあった場合には等しくはない。ここで、フレーム n における第2段バックワードフィルタを $A_{bwd}^*(n)$ とする。現フレーム n において算出されたLPフィルタを $A_{bwd}(n)$ とすると、 $A_{bwd}^*(n)$ は $A_{bwd}(n)$ と $A_{bwd}^*(n_e)$ の線形補間により求められる。ここで、 n_e は最後の消失フレームを表す（即ち、 $A_{bwd}^*(n_e)$ は最後の信頼できる第2段バックワードフィルタである）。

$$A_{bwd}^*(n) = \alpha_{lpbwd} A_{bwd}^*(n_e) + (1 - \alpha_{lpbwd}) A_{bwd}(n)$$

補間係数 α_{lpbwd} の初期値は0である。 α_{lpbwd} は、次のフレームに適用するために現フレームの最後に更新される。適応の手順を以下に示す。

- (1) 消失フレームが生じた場合には必ず α_{lpbwd} は最大値の 1 に固定される。
- (2) 正常フレーム n では次の処理が行われる。
 - (a) もし、フレーム n がフォワードモードならば、 $\alpha_{lpbwd} = 0$
 - (b) そうでなければ（フレーム n がバックワードモード）、高定常性インジケータ **High_Stat** の値に依存した量で α_{lpbwd} を減少させる。
 - (7) もし、**High_Stat** = 1 ならば、 α_{lpbwd} を 0.1 ステップで減少させる。
 - (4) そうでなければ、 α_{lpbwd} を 0.5 ステップで減少させる（高定常であればゆっくりと回復させ、それ以外では早く回復させる）。

そして、第 2 段バックワードフィルタ $A_{bwd}^*(n)$ がフレーム n の処理に用いられる。

消失フレームがフォワードモードと思われる場合には、標準 J T - G 7 2 9 の 4. 4. 1 節と同じ手順が適用される。

E. 4. 4. 3 適応および固定コードブック利得の減衰

適応および固定コードブック利得の減衰は、現サブフレーム以前の消失サブフレーム数および E. 4. 1. 1 節で与えられる第 2 定常性インジケータ **High_Stat2** に依存する。ここで、現サブフレーム m 以前に連続して消失したサブフレームの数を N_{bf} とする。減衰の手順を以下に示す。

- (1) もし、連続して消失したサブフレーム数が 2 未満（ $N_{bf} < 2$ ）ならば、

$$g_c^{(m)} = g_c^{(m-1)}$$

- (a) もし、**High_Stat2** が 1 ならば $g_p^{(m)} = 1$ 。
- (b) そうでなければ、 $g_p^{(m)} = 0.95$ とする。
- (2) それ以外（ $N_{bf} \geq 2$ ）の場合は、

$$\begin{aligned} g_p^{(m)} &= g_p^{(m-1)} * \alpha_g^{(m-1)} \\ g_c^{(m)} &= g_c^{(m-1)} * \alpha_g^{(m-1)} \end{aligned}$$

減衰係数の適応も N_{bf} および **High_Stat2** に依存している。即ち、

- (1) もし、連続して消失したサブフレーム数が 2 未満（ $N_{bf} < 2$ ）ならば、

$$\alpha_g^{(m)} = \alpha_g^{(m-1)} (= 1.)$$

- (2) それ以外（ $N_{bf} \geq 2$ ）の場合は、
 - (a) もし、**High_Stat2** = 1 ならば、

$$\text{if}(N_{bf} > 10) \alpha_g^{(m)} = \alpha_g^{(m-1)} * \alpha_g^h \text{ else } \alpha_g^{(m)} = \alpha_g^{(m-1)}$$

- (b) そうでなければ

$$\alpha_g^{(m)} = \alpha_g^{(m-1)} * \alpha_g^l$$

ここで $\alpha_g^l = 0.98$ 、 $\alpha_g^h = 0.995$ である。サブフレームが消失していない場合、 $\alpha_g^{(m)}$ は初期値 $\alpha_g^{(-1)}$ （1 に等しい）にリセットされる。

E. 4. 4. 4 利得予測器のメモリの減衰

標準 J T - G 7 2 9 の 4. 4. 3 節と同一である。

E. 4. 4. 5 置換用励振信号の生成

標準 J T - G 7 2 9 の 4. 4. 4 節に示すように、使用する励振信号は周期性に依存する。もし直前に再生したフレームが周期的であれば、現在のフレームも同様に周期的であると考えられる。この場合、適応コードブックのみが用いられ、固定コードブック寄与分は零に設定される。

誤りを含んだ（消失フレームもしくはピッチ遅延のパリティビットから誤りが検出された）サブフレーム m の適応コードベクトルのインデックスは、先行するフレームで算出および格納された分数ピッチ遅延を用いる。ここで、サブフレーム m における分数ピッチ遅延を $T^{(m)}$ とし、次のサブフレームでの誤り補償のために格納される分数ピッチ遅延を $T_{sav}^{(m)}$ とする。もし m が正常サブフレームならば $T^{(m)}$ は正常な復号値となり、そうでなければ $T^{(m)}$ は $T_{sav}^{(m-1)}$ に等しい値となる。 $T_{sav}^{(m)}$ の算出を以下に示す。

ここで整数 $stat_T^{(m)}$ を導入する。 $stat_T^{(m)}$ はピッチ遅延の定常性を表し、 $[0,7]$ の値をとる。 $stat_T^{(m)}$ の初期値は 0 である。ここで $T^{(m)}$ の整数部を $\text{int}(T^{(m)})$ とする。

(1) もし、 $|\text{int}(T^{(m)}) - \text{int}(T^{(m-1)})| < 5$ ならば、

$$stat_T^{(m)} = stat_T^{(m-1)} + 1 \quad \text{and} \quad T_{sav}^{(m)} = T^{(m)}$$

(2) そうでなければ、

(a) もし $|T_{multi} - \max(\text{int}(T^{(m)}), \text{int}(T^{(m-1)}))| < 5$ となるような $\min(\text{int}(T^{(m)}), \text{int}(T^{(m-1)}))$ の倍数 T_{multi} が存在するならば、

(7) もし、 $stat_T^{(m-1)} > 0$ ならば、

$$stat_T^{(m)} = stat_T^{(m-1)} - 1 \quad \text{and} \quad T_{sav}^{(m)} = T_{sav}^{(m-1)}$$

(b) それ以外の場合は、

$$stat_T^{(m)} = 0 \quad \text{and} \quad T_{sav}^{(m)} = T^{(m)}$$

そしてピッチ遅延の倍数もしくは約数は、信号の定常的な有声区間において推定されたピッチ遅延に置き換えられる。

適応コードブック利得は E. 4. 4. 3 節で算出された減衰した値に基づく。

もし直前に再生したフレームが周期的でないとして識別された場合は、現フレームも同様に周期的でないと考えられ、適応コードブック寄与分を零にする。固定コードブック寄与分は、ランダムに選択された 5 つのコードブックインデックスにより生成される。標準 J T - G 7 2 9 と同じ乱数発生器が用いられる。固定コードブック利得は、 E. 4. 4. 3 節に示した手順で減衰される。

E. 5 CS-ACELP 符号器のビットイグザクトな記述

16ビット固定小数点演算を用いて記述された 11.8kbit/s CS-ACELP 符号器の ANS I C コードは、新日本 I T U 協会から入手可能である。この ANS I C コードの現在の版は、1998年5月付け、バージョン 1.2 である。以下の節で、このシミュレーションソフトウェアの使い方とその構成について要約する。

E. 5. 1 シミュレーションソフトウェアの使い方

C言語のシミュレーションソフトウェアは、符号器をシミュレーションしたcoder.cと、復号器をシミュレーションしたdecoder.cの二つのメインプログラムから構成されている。符号器は、次のように実行させる。

```
coder inputfile bitstreamfile rate_option
```

復号器は、次のように実行させる

```
decoder bitstreamfile outputfile
```

入力ファイルと出力ファイルは、標準化された16ビットPCM信号で構成されているデータファイルである。符号化されたビット列のマッピングテーブルは、シミュレーションソフトウェアの中に含まれている。rate_optionは、高ビットレート拡張（11.8kbit/s）を選択する場合には1に設定し、標準JT-G729本体（8kbit/s）を選択する場合は0に設定する。デフォルト値は、0（8kbit/s）である。また、フレーム毎にレートを設定するために、0か1の値を持つ16ビットワードからなるバイナリファイルfile_rate_nameをrate_optionとして設定することもできる。

E. 5. 2 シミュレーションソフトウェアの構成

固定小数点演算を用いて記述されたANSI Cのシミュレーションソフトウェアにおいて、使用されている固定小数点データ形式と基本演算セットは、標準JT-G729に対応するソフトウェアと同一である。tab_ld8e.h（Table E-5-1/JT-G729参照）中に見られるいくつかの付加的なテーブルが追加された。

Table E-5-1/JT-G729 Summary of tables found in tab_ld8e.h
(ITU-T G.729)

Table name	Size	Description
lag_h_bwd	30	Lag window for backward LP bandwidth expansion (high part)
lag_l_bwd	30	Lag window for backward LP bandwidth expansion (low part)
bitsno_E_fwd	18	Bit allocation in forward mode
bitsno_E_bwd	16	Bit allocation in backward mode
hw	145	Backward LP analysis window
bitrates	2	Table of available bit rates
tab_log	17	Lookup table in base 2 logarithm Q.11

ファイルは、4つのグループに分類される。

- (1) 標準JT-G729に対応するソフトウェアファイルと同一なファイルである。この一覧を Table E-5-2/JT-G729に示す。
- (2) 標準JT-G729に対応するソフトウェアファイルと同様なファイル。標準JT-G729付属資料Eに対処するため、幾つかの小さな変更がなされている。この一覧を Table E-5-3/JT-G729に示す。
- (3) 標準JT-G729に対応するソフトウェアファイルを適切に修正したファイル。標準JT-G729付属資料Eを扱えるように、幾つかのソースコードラインが標準JT-G729ファイルに取り入れられた。この一覧を Table E-5-4/JT-G729に示す。
- (4) 標準JT-G729付属資料Eに特有なファイル（新たに作成されたファイル）。この一覧を Table E-5-5/JT-G729に示す。

Table E – 5 – 2 / JT-G729 List of software files identical to G.729 software
(ITU-T G.729)

<i>File name</i>	<i>Description</i>
basic_op.c	Basic operators
oper_32b.c	Extended basic operators
dspfunc.c	Mathematical functions
gainpred.c	Gain predictor
lpcfunc.c	Miscellaneous routines related to LP filter
pred_lt3.c	Generation of adaptive codebook
pre_proc.c	Pre-processing (HP filtering and scaling)
p_parity.c	Compute pitch parity
qua_gain.c	Gain quantizer
pwf.c	Computation of perceptual weighting coefficients (8 kbit/s)
pitch.c	Pitch search
util.c	Utility functions
acelp_co.c	Search fixed codebook (8 kbit/s)
post_pro.c	Post processing (HP filtering and scaling)
de_acelp.c	Decode algebraic codebook (8 kbit/s)
dec_lag3.c	Decode adaptive-codebook index
basic_op.h	Basic operators prototypes
ld8k.h	Function prototypes
oper_32b.h	Extended basic operators prototypes
tab_ld8k.c	ROM tables
tab_ld8k.h	Extern ROM table declarations
typedef.h	Data type definition (machine dependent)

Table E – 5 – 3 / JT-G729 List of software files similar to G.729 software
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Size (Bytes)</i>
qua_lspe.c	LSP quantizer	9999
filtere.c	Filter functions	3917

Table E – 5 – 4 / JT-G729 List of software files adapted from G.729 software
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Size (Bytes)</i>
codere.c	Main encoder routine	6679
cod_ld8e.c	Encoder routine	38119
decodere.c	Main decoder routine	8592
dec_ld8e.c	Decoder routine	20664
decgain.c	Decode gains	5633
pste.c	Postfilter routines	34020
bitse.c	Bit manipulation routines	6629
lspgetqe.c	LSP quantizer	5804
lpce.c	LP analysis	23394
lspdece.c	LSP decoding routing	3453

Table E – 5 – 5 / JT-G729 List of software files specific to G.729E software
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Size (Bytes)</i>
bffw.c	Backward/forward switch selection	11598
bffwfunc.c	Miscellaneous routines related to backward/forward switch selection	7416
ld8e.h	Function prototypes for G.729 Annex E	15540
pwfe.c	Computation of perceptual weighting coefficients (11.8 kbit/s)	1148
acelp_e.c	Search fixed codebook (11.8 kbit/s)	38780
deacelp_e.c	Decode algebraic codebook (11.8 kbit/s)	3542
tab_ld8e.c	ROM tables for G.729 Annex E	3313
tab_ld8e.h	Extern ROM declarations for G.729 Annex E	358
track_pi.c	Pitch tracking	3004

E. 6 参照すべき標準

- (1) TTC標準 JT-G 7 2 8

低遅延符号励振線形予測 (LD-CELP) を用いた 16kbit/s 音声符号化方式

- (2) TTC標準 JT-G 7 2 9

8kbit/s CS-ACELP を用いた音声符号化方式

付属資料 F

(標準 J T - G 7 2 9 に対する)

標準 J T - G 7 2 9 付属資料 D における付属資料 B の D T X 機能の実現方法

(本付属資料は仕様の一部である。)

F. 1 本付属資料の規定範囲

本付属資料は、標準 J T - G 7 2 9 付属資料 B と D を統合するための記述を規定するものであり、これにより付属資料 D における D T X 機能が定義される。本付属資料は、この統合と機能的な拡張に関し、標準的な実現方法を示している。これにより、製品化を助け、標準的な音声品質と互換性を世界的に保証することができる。本統合は、製品化における要求を満足するために、以下に示すいくつかの制約に着目しつつ、これを達成している。

- (1) 標準本体および各付属資料とビットイグザクトなこと。
- (2) プログラムコードやメモリ、演算量の増加を最小限に抑えること。
- (3) 関連する標準や付属資料におけるアプリケーションや品質との調和を図りつつ、新しい機能に対して厳しい品質を要求すること。

F. 2 参照すべき標準

本付属資料の規定は、下記の T T C 標準とその他の参照文献に含まれた規定から構成されており、それらの規定は本文中の至る所で参照されている。本付属資料の出版時点で、表記した版が制定されている。すべての標準およびその他の参照文献は改訂の対象である。従って、本付属資料のすべての利用者は、以下に示す標準およびその他の参照文献の最新版が適用できるかどうか調査することを奨励する。最新版の T T C 標準のリストは、定期的に出版されている。

[1] T T C 標準 J T - G 7 2 9

共役構造一代数的符号励振線形予測 (C S - A C E L P) を用いた 8kbit/s 音声符号化方式。平成 8 年 1 1 月 2 7 日制定。

[2] T T C 標準 J T - G 7 2 9 付属資料 B

I T U - T 勧告 V . 7 0 端末に適した標準 J T - G 7 2 9 に対する無音圧縮手法。平成 9 年 4 月 2 3 日制定。

[3] T T C 標準 J T - G 7 2 9 付属資料 D

6.4kbit/s C S - A C E L P 音声コーデック。平成 1 1 年 4 月 2 2 日制定。

F. 3 概要

標準 J T - G 7 2 9 本体 [1]、付属資料 B [2] および D [3] は、それぞれ 8kbit/s C S - A C E L P 符号化方式、D T X 機能、および 6.4kbit/s の低ビットレート拡張のビットイグザクトな固定小数点演算仕様を規定している。これらの仕様の詳細な記述はビットイグザクトな固定小数点演算の C コードで示されており、その C コードは日本 I T U 協会から入手可能である。付属資料 F は、標準 J T - G 7 2 9 付属資料 B と D の統合について記述ならびに定義している。

F. 4 新しい機能

本章では、付属資料 B と D の統合を容易にするためのアルゴリズムの修正および追加の概要を示す。

F. 4. 1 付属資料 D における付属資料 B の D T X 動作

付属資料 D において D T X 動作を提供するために付属資料 B と D の機能を統合することは、簡単なことで

ある。付属資料BのVAD（音声検出）や、SID（無音挿入記述子）の符号化、CNG（擬似背景雑音発生器）は、特に修正せずにそのまま使用できる。不連続伝送の区間では、付属資料Dのポストフィルタにおいて、位相拡散のためのパラメータの更新に注意する必要がある。（F. 5. 1 節、参照）

F. 5 アルゴリズムの記述

本章では、個々の付属資料のアルゴリズムに対し、統合を容易にするために必要な追加アルゴリズムの記述を示す。残りのすべてのモジュールは、標準本体、付属資料BまたはDが元になっている。

F. 5. 1 不連続伝送時の付属資料Dに特有な状態変数の更新

付属資料Dに特有な唯一の状態変数は、復号器における位相拡散モジュールの状態変数である（標準J T-G 7 2 9 付属資料D[3] D. 6. 2 節、参照）。無音フレームでは、擬似背景雑音発生器によって計算された利得値（標準J T-G 7 2 9 付属資料B[2] B. 4. 4 節、参照）を、適応およびACELP利得の推定値として標準ビットレート（8kbit/s）と同じ更新手順が、実行される。

F. 6 Cソースコードの記述

標準J T-G 7 2 9 付属資料F（付属資料BとDの統合版）は、16ビット固定小数点ANSI Cコードでシミュレーションされており、標準J T-G 7 2 9に対応するソフトウェアと同じタイプの固定小数点データ、および同じセットの固定小数点基本演算を用いている。ANSI Cコードは付属資料Fの仕様を規定している。Cコードによって与えられるアルゴリズム記述は、標準J T-G 7 2 9本体、付属資料B、D、およびFの本文の記述よりも優先される。以下に、シミュレーションコードの使い方とソフトウェアの構成について概説する。

F. 6. 1 シミュレーションソフトウェアの使い方

Cコードは2つのメインプログラム `coderf.c` と `decoderf.c` からなり、それぞれ符号器と復号器のシミュレーションを行う。符号器は以下のように実行される。

```
coderf inputfile bitstreamfile dtx_option rate_option
```

復号器は以下のように実行される。

```
decoderf bitstreamfile outputfile
```

入力ファイルと出力ファイルは、8kHzで標準化された16ビットのPCM信号からなるデータファイルである。ビットストリームファイルは、ビットストリームを含んだバイナリファイルである。符号化されたビットストリームのマッピングテーブルは、シミュレーションソフトウェアに含まれる。符号器に対する2つのパラメータは、`dtx_option` と `rate_option` である。

`dtx_option = 0` : DTX無効

 = 1 : DTX有効

 初期値は0（DTX無効）

`rate_option = 0` : 低レート（6.4kbit/s）を選択

 = 1 : 標準J T-G 7 2 9本体（8kbit/s）または、フレーム毎に0、1のレートを選択

 選択する16ビットワードからなるバイナリファイル `file_rate_name` を選択

 初期値は1（8kbit/s）

F. 6. 2 シミュレーションソフトウェアの構成

ファイルは、3つのグループに分類できる。

(1) 標準J T-G 7 2 9本体[1]、付属資料B[2]、または付属資料D[3]に対応するソフトウェアファイル

- と同一なファイル。この一覧を Table F-6-1/JT-G729 に示す。
- (2) 標準 J T-G 7 2 9 付属資料 B または 付属資料 D に対応するソフトウェアファイルから適切に修正したファイル。統合に対応するため、軽微な修正が加えられている。この一覧を Table F-6-2/JT-G729 に示す。
 - (3) 標準 J T-G 7 2 9 付属資料 B および 付属資料 D に対応するソフトウェアファイルを統合したファイル。この一覧を Table F-6-3/JT-G729 に示す。

Table F – 6 – 1 / JT-G729 List of software files identical to software files of G.729 main body, Annex B or D
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Identical to</i>
Basic_op.c	Basic operators	Main
Oper_32b.c	Extended basic operators	Main
Dspfunc.c	Mathematical functions	Main
Gainpred.c	Gain predictor	Main
Lpcfunc.c	Miscellaneous routines related to LP filter	Main
Pre_proc.c	Pre-processing (HP filtering and scaling)	Main
P_parity.c	Compute pitch parity	Main
Pwf.c	Computation of perceptual weighting coefficients (8 kbit/s)	Main
Pred_lt3.c	Generation of adaptive codebook	Main
Post_pro.c	Post processing (HP filtering and scaling)	Main
Typedef.h	Data type definition (machine dependent)	Main
Basic_op.h	Basic operators prototypes	Main
Oper_32b.h	Extended basic operators prototypes	Main
Filter.c	Filter functions	Main
Lspgetq.c	LSP quantizer	Main
De_acelp.c	ACELP decoding	Main
Lpc.c	LP analysis	B
Lspcdec.c	LSP decoding routines	B
Qua_lsp.c	LSP quantizer	B
Tab_ld8k.c	ROM tables	B
Taming.c	Pitch instability control	B
Dtx.c	DTX Decision	B
Dtx.h	Prototype and Constants	B
Qsidgain.c	SID Gain Quantization	B
Qsidlsf.c	SID-LSF Quantization	B
Tab_dtx.c	ROM tables	B
Pst.c	Postfilter routines	B
Vad.c	VAD	B
ld8k.h	Function prototypes	B
Vad.h	Prototype and Constants	B
Tab_ld8k.h	Extern ROM tables declarations	B
Sid.h	Prototype and Constants	B
Octet.h	Octet transmission mode definition	B
Tab_dtx.h	Extern ROM table declarations	B
Util.c	Utility functions	B
Pitchd.c	Pitch search	D
Declag3d.c	Decode adaptive-codebook index	D

Acelpcod.c	ACELP codebook search	D
Deacelpd.c	Decode Acelp codebook	D
Qua_g8k.c	Gain quantizer	D
Dec_g8k.c	Decode gain	D
Qua_g6k.c	Gain quantizer	D
Dec_g6k.c	Decode gain	D
Tabld8kd.c	ROM tables for G.729 at 6.4 kbit/s	D
Tabld8kd.h	Extern ROM declarations for G.729 at 6.4 kbit/s	D
ld8kd.h	Function prototypes for G.729D	D

Table F – 6 – 2 / JT-G729 List of software files adapted from software files of G.729 Annexes B and D (ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Adapted from</i>
Calcexc.c	CNG Excitation Calculation	B
Dec_sidf.c	Decode SID Information	B
Phdisp.c	Phase dispersion	D

Table F – 6 – 3 / JT-G729 List of software files integrating software files from G.729 main, Annexes B and D (ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Integrated from</i>
Coderf.c	Main encoder routine	Main+B+D
Cod_ld8f.c	Encoder routine	Main+B+D
Decoderf.c	Main decoder routine	Main+B+D
Dec_ld8f.c	Decoder routine	Main+B+D
Bitsf.c	Bit manipulation routines	Main+B+D
Ld8f.h	Constant and Function prototypes for G.729 Annex F	Main+B+D

付属資料G

(標準 J T - G 7 2 9 に対する)

標準 J T - G 7 2 9 付属資料Eに対する付属資料BのDTX機能の実現方法

(本付属資料は仕様の一部である。)

G. 1 本付属資料の規定範囲

本付属資料は、標準 J T - G 7 2 9 付属資料BとEを統合するための記述を規定するものであり、これにより付属資料EにおけるDTX機能が定義される。本付属資料は、この統合と機能的な拡張に関し、標準的な実現方法を示している。これにより、製品化を助け、標準的な音声品質と互換性を世界的に保証することができる。本統合は、製品化における要求を満足させるために、以下に示すいくつかの制約に着目しつつ、これを達成している。

- (1) 標準本体および各付属資料とビットイグザクトなこと。
- (2) プログラムコードやメモリ、演算量の増加を最小限に抑えること。
- (3) 関連する標準や付属資料におけるアプリケーションや品質との調和を図りつつ、新しい機能に対して厳しい品質を要求こと。

G. 2 参照すべき標準

本付属資料の規定は、下記のTTC標準とその他の参照文献に含まれた規定から構成されており、それらの規定は本文中の至る所で参照されている。本付属資料の出版時点で、表記した版が制定されている。すべての標準およびその他の参照文献は改訂の対象である。従って、本付属資料のすべての利用者は、以下に示す標準およびその他の参照文献の最新版が適用できるかどうか調査することを奨励する。最新版のTTC標準のリストは、定期的な出版されている。

[1] TTC標準 J T - G 7 2 9

共役構造一代数的符号励振線形予測 (CS-ACELP) を用いた 8kbit/s 音声符号化方式 平成8年11月27日制定。

[2] TTC標準 J T - G 7 2 9 付属資料B

I T U - T 勧告 V . 7 0 端末に適した標準 J T - G 7 2 9 に対する無音圧縮手法 平成9年4月22日制定。

[3] TTC標準 J T - G 7 2 9 付属資料E

11.8kbit/s CS-ACELP 音声コーデック 平成11月4月22日制定。

G. 3 概要

標準 J T - G 7 2 9 本体[1]、付属資料B[2]およびE[3]は、それぞれ8kbit/s CS-ACELP 符号化方式、DTX機能、および11.8kbit/sの高ビットレート拡張のビットイグザクトな固定小数点演算仕様を規定している。これらの仕様の詳細な記述はビットイグザクトな固定小数点演算のCコードで示されており、そのCコードは日本ITU協会から入手可能である。付属資料Gは、標準 J T - G 7 2 9 付属資料BとEの統合について記述ならびに定義している。

G. 4 新しい機能

本章では、付属資料BとEの統合を容易にするためのアルゴリズムの修正および追加の概要を示す。アプリケーションに対応させるために、モジュールによっては、ある追加が必要であることが判っている。

G. 4. 1 付属資料Eにおける付属資料BのDTX動作

付属資料EにおいてDTX動作を提供するために付属資料BとEの機能を統合することは、若干の複雑さがある。付属資料BのDTX動作が10次のLPC分析に基づいているため、付属資料BのVAD機能は、付属資料Eの10次のフォワード適応LPC分析が行われた後で、かつバックワード適応LPC分析が行われる前に実行される。VAD機能が「無音」を検出した場合は、付属資料EのLPCモードは強制的にフォワード適応LPCとなり、バックワード適応LPC分析はスキップされる。さらに音楽を検出するために、VAD機能の後に補正用のモジュールを追加する必要があることが判っている。これは、ある程度の品質で音楽を伝送する能力を提供することが付属資料Eの目的の一つであり、付属資料Eの拡張されたアプリケーションにも対応させるためのものである。ちなみに、付属資料Eの開発に当たっては、音楽信号に対する性能において厳しい要求項目があった。一方、標準本体、付属資料BおよびDに関しては、音楽信号に対する性能において厳しい要求項目はなかった。付属資料BのDTX動作を適用した場合の付属資料Eの音楽信号に対する品質を保証するために、音楽信号が継続している区間では、音楽検出機能は強制的にVAD機能を「有音」の状態にする。これにより、音楽の区間では確実に付属資料Eの11.8kbit/sで符号化が行われる。付属資料BにおけるSIDの符号化やCNGは、特に修正せずにそのまま使用できる。なお、不連続伝送の区間では、付属資料EのLPCモード選択アルゴリズムにおけるパラメータを正しく更新することに注意する必要がある（G. 5. 2節、参照）。

G. 5 アルゴリズムの記述

本章では、個々の付属資料のアルゴリズムに対し、統合を容易にするために必要な付加アルゴリズムの記述を示す。残りのすべてのモジュールは、標準本体、付属資料Bまたは付属資料Eが元になっている。

G. 5. 1 音楽検出

音楽検出は新しい機能である。これはVADの直後に実行され、音楽区間中は、VADを強制的に「有音」の状態にする。

音楽検出アルゴリズムは、音楽信号が存在している間の、音声検出（VAD）からの判定を訂正する。音楽検出は、以下のパラメータに基づいている。

- (1) *Vad_deci*、現フレームのVAD判定
- (2) *PVad_dec*、直前フレームのVAD判定
- (3) *Lpc_mod*、直前フレームの適応LPCが、フォワードかバックワードかいずれかを示すフラグインジケータ
- (4) *Rc*、LPC分析から得られる反射係数
- (5) *Lag_buf*、最終5フレームの修正された開ループピッチラグのバッファ
- (6) *Pgain_buf*、最終5サブフレームの閉ループピッチ利得のバッファ
- (7) *Energy*、LPC分析からの1次自己相関係数 $R(0)$
- (8) *Frm_count*、VADモジュールからの正規化対数エネルギー
- (9) *Frm_count*、処理された信号フレーム数のカウンタ
- (10) *Rate*、音声コーデックの選択

本アルゴリズムは、2つの主要な部分からなる。

- (1) 関連パラメータの計算
- (2) パラメータに基づく分類

G. 5. 1. 1 関連パラメータの計算

本節では、判定モジュールで使用されるパラメータの計算について記述している。

(1) 部分正規化残差エネルギー

$$Lenergy = 10 \log_{10} \left(\prod_{i=1}^4 (1 - Rc(i))^2 \right) Energy / 240$$

(2) スペクトル差分と背景雑音の部分正規化残差エネルギーの移動平均

現フレームの反射係数 Rc と背景雑音の移動平均反射係数 mRc の間のスペクトル差分量は次式で与えられる。

$$SD = \sum_{i=1}^{10} (Rc(i) - mRc(i))^2$$

移動平均 \overline{mrc} と $mLenergy$ は、VADモジュールによって生成されたVAD判定 Vad_deci を用いて以下のとおり更新される。

```
if Vad_deci == NOISE {
     $\overline{mrc} = 0.9\overline{mrc} + 0.1rc$ 
     $mLenergy = 0.9mLenergy + 0.1Lenergy$ 
}
```

(3) ピッチラグバッファ更新のための開ループピッチラグ補正

開ループピッチラグ T_{op} は、倍ピッチあるいは3倍ピッチを回避するため、以下のように補正して用いる。

$$avg_lag = \sum_{i=1}^4 Lag_buf(i) / 4$$

```
if (abs( $T_{op} / 2 - avg\_lag$ ) <= 2)
```

$$Lag_buf(5) = T_{op} / 2$$

```
else if (abs( $T_{op} / 3 - avg\_lag$ ) <= 2)
```

$$Lag_buf(5) = T_{op} / 3$$

```
else
```

$$Lag_buf(5) = T_{op}$$

なお、開ループピッチラグ T_{op} は変更されないこと、かつ開ループピッチ分析によって得られたものと同一であることに注意すべきである。

(4) ピッチラグの標準偏差

$$std = \sqrt{Var / 4}$$

$$\text{where } Var = \sum_{i=1}^5 (Lag_buf(i) - \mu)^2 \text{ and } \mu = \sum_{i=1}^5 (Lag_buf(i) / 5)$$

(5) ピッチ利得の移動平均

$$mPgain = 0.8mPgain + 0.2\theta, \quad \text{where } \theta = \sum_{i=1}^5 (Lag_buf(i) / 5)$$

ピッチ利得バッファ *Pgain_buf* は、サブフレーム処理の後、もし *Vad_deci = NOISE* ならば、ピッチ利得値 0.5 で更新され、そうでなければ、量子化されたピッチ利得で更新される。

(6) ピッチラグの滑らかさと有声強度インジケータ

ピッチラグの滑らかさと有声強度インジケータ *Pflag* は、以下の論理的なステップを使用して生成される。

最初に、2つの中間論理フラグ *Pflag1* と *Pflag2* が以下のとおり得られる。

```
if (std < 1.3 and mPgain > 0.45) set Pflag1=1 else 0
if (mPgain > Thres) set Pflag2=1 else 0,
where Thres = 0.63
```

最後に、*Pflag* は以下より決定される。

```
if ((PVad_dec == VOICE and (Pflag1==1 or Pflag2==1)) or (Pflag2==1))
set Pflag=1 else 0
```

(7) 定常性カウンタ

1組のカウンタが定義され、以下のとおり更新される。

- (a) *count_consc_rflag* は、2次反射係数とピッチ利得の移動平均が以下の条件を満足する場合に、連続フレーム数をカウントする。

```
if (Rc(2) < 0.45 and Rc(2) > 0 and mPgain < 0.5)
count_consc_rflag = count_consc_rflag + 1
else
count_consc_rflag = 0
```

- (b) *count_music* は、64 フレームの窓内で直前のフレームがバックワード適応LPCを使用しており、現フレームが(VADにより)「有音」である場合は、フレーム数をカウントする。

```
if (LPC_mod == 1 and Vad_deci == VOICE)
cont_music = cont_music + 1
```

64 フレームごとに、*count_music* の移動平均 *mcount_music* は、以下の記述のとおり更新および零にリセットされる。

```

if((Frm_count mod 64) == 0){
  if(Frm_count == 64)
    mcount_music = count_music
  else
    mcount_music = 0.9mcount_music + 0.1count_music
}

```

(c) *count_consc* は、*count_music* が零のままである場合、連続フレーム数をカウントする。

```

if (count_music==0)
  count_consc=count_consc+1
else
  count_consc=0

if (count_consc>500 or count_consc_rflag>150) set mcount_music=0

```

(b)における *count_music* は、関連するカウンタの更新の後、64 フレーム毎に零にリセットされる。

(c)における論理は、*count_music* の移動平均をリセットするために使用される。

(d) *count_pflag* は、64 フレームの窓内で *Pflag* = 1 であるフレーム数をカウントする。

```

if (Pflag==1)
  count_pflag =count_pflag+1

```

64 フレーム毎に、*count_pflag* の移動平均 *mcount_pflag* は更新され、以下の記述のとおり零にリセットされる。

```

if((Frm_count mod 64)== 0){
  if (Frm_count==64)
    mcount_pflag =count_pflag
  else{
    if (count_pflag > 25)
      mcount_pflag =0.98mcount_pflag +0.02count_pflag
    else(count_pflag > 20)
      mcount_pflag =0.95mcount_pflag +0.05count_pflag
    else
      mcount_pflag =0.9mcount_pflag +0.1count_pflag
  }
}

```

(e) *count_consc_pflag* は、以下の条件を満足する連続フレーム数をカウントする。

```

if (count_pflag == 0)
  count_consc_pflag =count_consc_pflag +1
else
  count_consc_pflag = 0

```

```
if(count_consc_pflag > 100 or count_consc_rflag > 150) set mcount_pflag = 0
```

count_pflag は 64 フレーム毎に零にリセットされる。(e)における論理は *count_pflag* の移動平均をリセットするために使用される。

G. 5. 1. 2 分類

上記のパラメータ推定に基づいて、VADモジュールからのVAD判定 *Vad_deci* は、もし以下の条件が満足されるのならば判定をやり直す。

```
if (Rate = G729E){
  if (SD > 0.15 and (Lenergy - mLenergy) > 4 and LLenergy > 50)
    Vad_deci=VOICE
  else if ((SD > 0.38 or (Lenergy - mLenergy) > 4) and LLenergy > 50)
    Vad_deci = VOICE
  else if ((mcount_pflag >= 10 or mcount_music >= 1.0938 or Frm_count < 64) and LLenergy > 7)
    Vad_deci = VOICE
}
```

注意すべきことは、音楽検出機能は、メモリを最新の状態に維持するために、動作中の符号化モードに関わらず、常に呼び出されることである。ただし、VAD判定 *Vad_deci* は、付属資料Gが 11.8kbit/s (付属資料E) で動作しているときに限り変化する。また、音楽検出は、VAD判定に対し「無音」から「有音」へ判定を変えることだけを許容し、その逆はなされないことに注意すべきである。

G. 5. 2 不連続伝送時における付属資料Eに特有な状態変数の更新

G. 5. 2. 1 付属資料Eに特有な符号器の状態変数の更新

符号器において、無音フレームの場合、状態変数の更新は標準 J T - G 7 2 9 付属資料E [3]において標準ビットレート 8kbit/s に切り替える場合と同一である。更新手順は以下の通りである。LPモードを0にセットし、グローバル定常性インジケータの値を減少させ、高定常性インジケータを0にリセットする(標準 J T - G 7 2 9 付属資料E [3] E. 3. 2. 7. 2 節、参照)。フォワードLPフィルタからバックワードLPフィルタへのなめらかな切り替えを行うための補間係数を、最大値にリセットする(標準 J T - G 7 2 9 付属資料E [3] E. 3. 2. 7. 1 節、参照)。

G. 5. 2. 2 付属資料Eに特有な復号器の状態変数の更新

復号器において、無音フレームの場合、状態変数の更新はピッチ追跡手順による計算(標準 J T - G 7 2 9 付属資料E [3] E. 4. 4. 5 節、参照)を実行する代わりに、ピッチ遅延定常性インジケータを0にリセットすることを除いて、標準 J T - G 7 2 9 付属資料E [3]におけるフォワードモードのみのレート(8kbit/s)に切り替える場合と同一の処理を行う。

G. 6 Cソースコードの記述

標準 J T - G 7 2 9 付属資料G (付属資料BとEの統合版) は、16 ビット固定小数点ANSI Cコードでシミュレーションされており、標準 J T - G 7 2 9 に対応するソフトウェアと同じタイプの固定小数点データ、および同じセットの固定小数点基本演算を用いている。ANSI Cコードは付属資料Gの仕様を規定している。Cコードによって与えられるアルゴリズム記述は、標準 J T - G 7 2 9 本体、付属資料B、E、

およびGの本文の記述よりも優先される。以下に、シミュレーションコードの使い方とソフトウェアの構成について概説する。

G. 6. 1 シミュレーションソフトウェアの使い方

Cコードは2つのメインプログラム `coderg.c` と `decoderg.c` からなり、それぞれ符号器と復号器のシミュレーションを行なう。符号器は以下のように実行される。

```
coderg inputfile bitstreamfile dtx_option rate_option
```

復号器は以下のように実行される。

```
decoderg bitstreamfile outputfile
```

入力ファイルと出力ファイルは、8kHz で標準化された 16 ビットの PCM 信号からなるデータファイルである。ビットストリームファイルは、ビットストリームを含んだバイナリファイルである。符号化されたビットストリームのマッピングテーブルは、シミュレーションソフトウェアに含まれる。符号器に対する2つのパラメータは、`dtx_option` と `rate_option` である。

`dtx_option` = 0 : DTX無効

 = 1 : DTX有効

 初期値は0 (DTX無効)

`rate_option` = 1 : 標準 JT-G 7 2 9 本体 (8kbit/s) を選択

 = 2 : 高レート (11.8kbit/s) またはフレーム毎に 1、2 のレートを選択する

 16 ビットワードからなるバイナリファイル `file_rate_name` を選択

 初期値は1 (8kbit/s)

G. 6. 2 シミュレーションソフトウェアの構成

ファイルは、4つのグループに分類できる。

- (1) 標準 JT-G 7 2 9 標準本体[1]、付属資料B[2]、または付属資料E[4]に対応するソフトウェアファイルと同一なファイル。この一覧を TableG-6-1/JT-G729 に示す。
- (2) 標準 JT-G 7 2 9 付属資料B、または付属資料Eに対応するソフトウェアファイルから適切に修正したファイル。統合に対応するため、軽微な修正が加えられている。この一覧を TableG-6-2/JT-G729 に示す。
- (3) 標準 JT-G 7 2 9 付属資料Bと付属資料Eに対応するソフトウェアファイルを統合したファイル。この一覧を TableG-6-3/JT-G729 に示す。
- (4) 本標準 JT-G 7 2 9 統合版に特有なファイル(新ファイル)。この一覧を TableG-6-4/JT-G729 に示す。

Table G – 6 – 1 / JT-G729 List of software files identical to software files of G.729 main body, Annex B or E
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Identical to</i>
Basic_op.c	Basic operators	Main
Oper_32b.c	Extended basic operators	Main
Dspfunc.c	Mathematical functions	Main
Gainpred.c	Gain predictor	Main
Lpcfunc.c	Miscellaneous routines related to LP filter	Main
Pre_proc.c	Pre-processing (HP filtering and scaling)	Main
P_parity.c	Compute pitch parity	Main
Pwf.c	Computation of perceptual weighting coefficients (8 kbit/s)	Main
Pred_lt3.c	Generation of adaptive codebook	Main
Post_pro.c	Post processing (HP filtering and scaling)	Main
Pitch.c	Pitch search	Main
Dec_lag3.c	Decode adaptive-codebook index	Main
Typedef.h	Data type definition (machine dependent)	Main
Basic_op.h	Basic operators prototypes	Main
Oper_32b.h	Extended basic operators prototypes	Main
Acelp_co.c	ACELP codebook search	Main
De_acelp.c	Decode Acelp codebook	Main
Qua_gain.c	Gain quantizer	Main
Tab_ld8k.c	ROM tables	B
Taming.c	Pitch instability control	B
Qsidgain.c	SID Gain Quantization	B
Qsidlsf.c	SID-LSF Quantization	B
Tab_dtx.c	ROM tables	B
Calcexc.c	CNG Excitation Calculation	B
Util.c	Utility functions	B
Ld8k.h	Function prototypes	B
Tab_ld8k.h	Extern ROM tables declarations	B
Dtx.h	Prototype and Constants	B
Sid.h	Prototype and Constants	B
Octet.h	Octet transmission mode definition	B
Tab_dtx.h	Extern ROM table declarations	B
Vad.h	Prototype and Constants	B
Pwfe.c	Computation of perceptual weighting coefficients	E
Filtere.c	Filter functions	E
Lspgetqe.c	LSP quantizer	E
Lspdece.c	LSP decoding routing	E
Qua_lspe.c	LSP quantizer	E
Bwfwfunc.c	Miscellaneous routines related to backward/forward switch selection	E
Ld8e.h	Function prototypes for G.729, Annex E	E
Acelp_e.c	Search fixed codebook (11.8 kbit/s)	E
Deacelp.c	Decode algebraic codebook (11.8 kbit/s)	E
Decgain.c	Decode gains	E
Tab_ld8e.c	ROM tables for G.729 at 11.8 kbit/s	E

Tab_ld8e.h	Extern ROM declarations for G.729 at 11.8 kbit/s	E
Track_pi.c	Pitch tracking	E

Table G – 6 – 2 / JT-G729 List of software files adapted from software files of G.729 main body/B/D/E (ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Adapted from</i>
Dtxg.c	DTX Decision	B
Vadg.c	VAD	B
Dec_sidf.c	Decode SID Information	B
Bwfwg.c	Backward/forward switch selection	E

Table G – 6 – 3 / JT-G729 List of software files integrating software files from G.729 main, Annex B, D or E (ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Integrated from</i>
Coderg.c	Main encoder routine	B+E
Cod_ld8g.c	Encoder routine	B+E
Decoderg.c	Main decoder routine	B+E
Dec_ld8g.c	Decoder routine	B+E
Bitsg.c	Bit manipulation routines	B+E
Lpcg.c	LP analysis	B+E
Pstg.c	Postfilter routines	B+E
Ld8g.h	Constant and Function prototypes for G.729 Annex G	B+E

Table G – 6 – 4 / JT-G729 List of software files specific to integrated G.729 Annexes B, D and E (ITU-T G.729)

<i>File name</i>	<i>Description</i>
Mus_dtct.c	Music detection module

付属資料H

(標準 J T - G 7 2 9 に対する)

標準 J T - G 7 2 9 付属資料DおよびEの切替手順に対する実現方法

(本付属資料は仕様の一部である。)

H. 1 本付属資料の規定範囲

本付属資料は、標準 J T - G 7 2 9 付属資料DおよびEの統合について、またこれによる付属資料DおよびEの切替手順の定義についての記述を規定している。本付属資料は、この統合と機能的な拡張に関し、標準的な実現方法を示している。これにより、製品化を助け、標準的な音声品質と互換性を世界的に保証することができる。本統合は、製品化における要求を満足させるために、いくつかの制約に着目しつつ、これを達成している。

- (1) 標準本体および各付属資料とビットイグザクトなこと。
- (2) プログラムコード、メモリ、演算量の増加を最小限に抑えること。
- (3) 関連する標準や付属資料におけるアプリケーションや品質との調和を図りつつ、新しい機能に対して厳しい品質を要求すること。

H. 2 参照すべき標準

本付属資料の規定は、下記の T T C 標準とその他の参照文献に含まれた規定から構成されており、それらの規定は本文中の至る所で参照されている。本付属資料の出版時点で、表記した版が制定されている。すべての標準およびその他の参照文献は改訂の対象である。従って、本付属資料の全ての利用者は、以下に示す標準およびその他の参照文献の最新版が適用できるかどうか調査することを奨励する。最新版の T T C 標準のリストは、定期的に出版されている。

[1] T T C 標準 J T - G 7 2 9

共役構造一代数的符号励振線形予測 (C S - A C E L P) を用いた 8kbit/s 音声符号化方式。平成 8 年 1 1 月 2 7 日制定。

[2] T T C 標準 J T - G 7 2 9 付属資料D

6.4kbit/s C S - A C E L P 音声コーデック。平成 1 1 年 4 月 2 2 日制定。

[3] T T C 標準 J T - G 7 2 9 付属資料E

11.8kbit/s C S - A C E L P 音声コーデック。平成 1 1 年 4 月 2 2 日制定。

H. 3 概要

標準 J T - G 7 2 9 本体[1]、付属資料D[2]およびE[3] は、それぞれ 8kbit/s C S - A C E L P 符号化方式、6.4kbit/s 低ビットレート拡張、および 11.8kbit/s 高ビットレート拡張のビットイグザクトな固定小数点演算仕様を規定している。これらの仕様の詳細な記述は、ビットイグザクトな固定小数点演算の C コードで示されており、その C コードは日本 I T U 協会から入手可能である。付属資料Hは、標準 J T - G 7 2 9 付属資料DおよびEの統合について記述ならびに定義をしている。

H. 4 アルゴリズムの記述

本章では、個々の付属資料のアルゴリズムに対し、統合を容易にするために必要な付加アルゴリズムの記述を示す。残りのすべてのモジュールは、標準 J T - G 7 2 9 本体、付属資料DまたはEが元になっている。

H. 4. 1 付属資料Eフレームにおける付属資料Dに特有な状態変数の更新

付属資料Dに特有な唯一の状態変数は、復号器における位相分散モジュールの状態変数である (標準 J T

ーG 7 2 9 付属資料D[2] D. 6. 2 節、参照)。11.8kbit/s のフレームの場合には、標準ビットレート (8kbit/s) の場合と同じ更新手順が実行される。

H. 4. 2 付属資料Dフレームにおける付属資料Eに特有な状態変数の更新

H. 4. 2. 1 付属資料Eに特有な符号器の状態変数の更新

符号器において、6.4kbit/s のフレームの場合、状態変数の更新は標準 J T - G 7 2 9 付属資料 E [3] の標準ビットレート (8kbit/s) に切り替わる場合に行われる更新と同一である。更新は以下のように行われる。L P モードが 0 にセットし、グローバル定常性インジケータの値を減少させ、高定常性インジケータが 0 にリセットする (標準 J T - G 7 2 9 付属資料 E [3] E. 3. 2. 7. 2 節、参照)。フォワード L P フィルタからバックワード L P フィルタへのなめらかな切り替えを行うための補間係数を、最大値にリセットする (標準 J T - G 7 2 9 付属資料 E [3] E. 3. 2. 7. 1 節、参照)。

H. 4. 2. 2 付属資料Eに特有な復号器の状態変数の更新

復号器において、6.4kbit/s のフレームの場合、状態変数の更新は標準 J T - G 7 2 9 付属資料 E [3] の標準ビットレート (8kbit/s) に切り替わる場合と同一の処理を行う。

H. 5 Cソースコードの記述

標準 J T - G 7 2 9 付属資料 H (付属資料 D と E の統合版) は、16 ビット固定小数点 ANS I C コードでシミュレーションされており、標準 J T - G 7 2 9 に対応するソフトウェアと同じタイプの固定小数点データ、および同じセットの固定小数点基本演算を用いている。ANS I C コードは付属資料 H の仕様を規定している。C コードによって与えられたアルゴリズム記述は、標準 J T - G 7 2 9 本体、付属資料 D、E および H の本文の記述よりも優先される。以下に、シミュレーションコードの使い方とソフトウェアの構成について概説する。

H. 5. 1 シミュレーションソフトウェアの使い方

C コードは 2 つのメインプログラム `coderh.c` と `decoderh.c` からなり、それぞれ符号器と復号器のシミュレーションを行う。符号器は以下のように実行される

```
coderh inputfile bitstreamfile rate_option
```

復号器は次のように実行される

```
decoderh bitstreamfile outputfile
```

入力ファイルと出力ファイルは、8kHz で標本化された 16 ビットの P C M 信号からなるデータファイルである。ビットストリームファイルは、ビットストリームを含んだバイナリファイルである。符号化されたビットストリームのマッピングテーブルは、シミュレーションソフトウェアに含まれる。符号器に対するパラメータは、`rate_option` である。

```
rate_option = 0 : 低レート (6.4kbit/s) を選択、  
              = 1 : 標準 J T - G 7 2 9 本体 (8kbit/s) を選択、  
              = 2 : 高レート (11.8kbit/s) または、フレーム毎に 0、1、2 のレートを選択する 16  
                  ビットワードからなるバイナリファイル file_rate_name を選択  
              初期値は 1 (8kbit/s)
```

H. 5. 2 シミュレーションソフトウェアの構成

ファイルは、3 つのグループに分類できる。

(1) 標準 J T - G 7 2 9 本体、付属資料 D または付属資料 E に対応するソフトウェアファイルと同一なフ

- ファイル。この一覧を Table H-6-1/JT-G729 に示す
- (2) 標準 J T-G 7 2 9 付属資料Dまたは付属資料Eに対応するソフトウェアファイルから適切に修正したファイル。統合に対応し状態変数の更新を扱うため、軽微な修正が加えられている。この一覧を、Table H-6-2/JT-G729 に示す。
 - (3) 標準 J T-G 7 2 9 付属資料Dまたは付属資料Eに対応するソフトウェアファイルを統合したファイル。この一覧を、Table H-6-3/JT-G729 に示す

Table H-6-1 / JT-G729 List of software files identical to software files of G.729 main body, Annex D or E (ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Identical to</i>
Basic_op.c	Basic operators	Main
Oper_32b.c	Extended basic operators	Main
Dspfnc.c	Mathematical fncions	Main
Gainpred.c	Gain predictor	Main
Lpcfnc.c	Miscellaneous routines related to LP filter	Main
Pre_proc.c	Pre-processing (HP filtering and scaling)	Main
P_parity.c	Compute pitch parity	Main
Pwf.c	Computation of perceptual weighting coefficients (8 kbit/s)	Main
Pred_lt3.c	Generation of adaptive codebook	Main
Post_pro.c	Post processing (HP filtering and scaling)	Main
Tab_ld8k.c	ROM Table	Main
Basic_op.h	Basic operators prototypes	Main
Ld8k.h	Function prototypes	Main
Oper_32b.h	Extended basic operators prototypes	Main
Tab_ld8k.h	Extern ROM table declarations	Main
Typedef.h	Data type definition (machine dependent)	Main
Taming.c	Pitch instability control	B
Qua_g8k.c	Gain quantizer	D
Qua_g6k.c	Gain quantizer	D
Tabld8kd.c	ROM tables for G.729 at 6.4 kbit/s	D
Tabld8kd.h	Extern ROM declarations for G.729 at 6.4 kbit/s	D
ld8kd.h	Function prototypes for G.729D	D
Bwfwfnc.c	Miscellaneous routines related to backward/forward switch selection	E
Filtere.c	Filter functions	E
Lpce.c	LP analysis	E
Lspcdece.c	LSP decoding routines	E
Lspgetqe.c	LSP quantizer	E
Qua_lspe.c	LSP quantizer	E
Pstpe.c	Postfilter routines	E
Track_pi.c	Pitch tracking	E
Tab_ld8e.c	ROM tables for G.729 at 11.8 kbit/s	E
Tab_ld8e.h	Extern ROM declarations for G.729 at 11.8 kbit/s	E
Util.c	Utility functions	E
pwfe.c	computation of perceptual weighting coefficients	E
ld8e.c	Function prototypes for G.729 annex E	E

Table H – 6 – 2 / JT-G729 List of software files adapted from software files of G.729 main body/D/E
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Adapted from</i>
Phsidp.c	Phase dispersion	D
Bwfw.h.c	Backward/forward switch selection	E

Table H – 6 – 3 / JT-G729 List of software files integrating software files from G.729 main, Annex D or E
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Integrated from</i>
Coder.c	Main encoder routine	D+E
Cod_ld8h.c	Encoder routine	D+E
Decoderh.c	Main decoder routine	D+E
Dec_ld8h.c	Decoder routine	D+E
Acelp_h.c	search ACELP fixed codebook (6.4, 8, 11.8 kbit/s)	D+E
Deacelph.c	Decode algebraic codebook (6.4, 8, 11.8 kbit/s)	D+E
Pitchh.c	Pitch search	D+E
Declagh.c	Decode adaptive-codebook index	D+E
Decgainh.c	Decode gain	D+E
Bitsh.c	Bit manipulation routines	D+E
Ld8h.h	Constant and Function prototypes for G.729 Annex H	D+E

付属資料 I

(標準 J T-G 7 2 9 に対する)

標準 J T-G 7 2 9 本体、付属資料 B、D および E の C コードの統合における実現方法

(本付属資料は仕様の一部である。)

I. 1 本付属資料の規定範囲

本付属資料は、標準 J T-G 7 2 9 本体、付属資料 B、D および E を統合するための記述を規定するものであり、これにより統合された C コードが定義される。本付属資料は、この統合と機能的な拡張に関し、標準的な実現方法を示している。これにより、製品化を助け、標準的な音声品質と互換性を世界的に保証することができる。本統合は、製品化における要求を満足するために、以下に示すいくつかの制約に着目しつつ、これを達成している。

- (1) 標準本体および各付属資料とビットイグザクトなこと。
- (2) プログラムコードやメモリ、演算量の増加を最小限に抑えること。
- (3) 関連する標準や付属資料におけるアプリケーションや品質との調和を図りつつ、新しい機能に対して厳しい品質を要求すること。

I. 2 参照すべき標準

本付属資料の規定は、下記の T T C 標準とその他の参照文献に含まれる規定から構成されており、それらの規定は本文中の至る所で参照されている。本付属資料の出版時点で、表記した版が制定されている。すべての標準およびその他の参照文献は改訂の対象である。従って、本付属資料のすべての利用者は、以下に示す標準およびその他の参照文献の最新版が適用できるかどうか調査することを奨励する。最新版の T T C 標準のリストは、定期的に出版されている。

[1] T T C 標準 J T-G 7 2 9

共役構造一代数的符号励振線形予測 (C S-A C E L P) を用いた 8kbit/s 音声符号化方式。平成 8 年 1 1 月 2 7 日制定。

[2] T T C 標準 J T-G 7 2 9 付属資料 B

I T U-T 勧告 V. 7 0 端末に適した標準 J T-G 7 2 9 に対する無音圧縮手法。平成 9 年 4 月 2 3 日制定。

[3] T T C 標準 J T-G 7 2 9 付属資料 D

6.4kbit/s C S-A C E L P 音声コーデック。平成 1 1 年 4 月 2 2 日制定。

[4] T T C 標準 J T-G 7 2 9 付属資料 E

11.8kbit/s C S-A C E L P 音声コーデック。平成 1 1 年 4 月 2 2 日制定。

I. 3 概要

標準 J T-G 7 2 9 本体[1]、付属資料 B[2]、D[3]、および E[4]は、それぞれ 8kbit/s C S-A C E L P 符号化方式、D T X 機能、6.4kbit/s の低ビットレート拡張、および 11.8kbit/s の高ビットレート拡張のビットイグザクトな固定小数点演算仕様を規定している。これらの仕様の詳細な記述はビットイグザクトな固定小数点演算の C コードで示されており、その C コードは日本 I T U 協会から入手可能である。付属資料 I は標準 J T-G 7 2 9 本体、付属資料 B、D、および E の統合について、記述ならびに定義している。

I. 4 新しい機能

本章では、標準本体、付属資料 B、D、および E の統合を容易にするためのアルゴリズムの修正および追加の概要を示す。アプリケーションに対応させるために、モジュールによっては、ある追加が必要であるこ

とが判っている。

I. 4. 1 付属資料Dにおける付属資料BのDTX動作

付属資料DにおいてDTX動作を提供するために付属資料BとDの機能を統合することは、簡単なことである。付属資料BのVAD（音声検出）や、SID（無音挿入記述子）の符号化、CNG（擬似背景雑音発生器）は、特に修正せずにそのまま使用できる。不連続伝送の区間では、付属資料Dのポストフィルタにおいて、位相拡散のためのパラメータの更新に注意する必要がある（I. 5. 2節、参照）。

I. 4. 2 付属資料Eにおける付属資料BのDTX動作

付属資料EにおいてDTX動作を提供するために付属資料BとEの機能を統合することは、若干の複雑さがある。付属資料BのDTX動作が10次のLPC分析に基づいているため、付属資料BのVAD機能は、付属資料Eの10次のフォワード適応LPC分析が行われた後で、かつバックワード適応LPC分析が行われる前に実行される。VAD機能が「無音」を検出した場合は、付属資料EのLPCモードは強制的にフォワード適応LPCとなり、バックワード適応LPC分析はスキップされる。さらに音楽を検出するために、VAD機能の後に補正用のモジュールを追加する必要があることが判っている。これは、ある程度の品質で音楽を伝送する能力を提供することが付属資料Eの目的の一つであり、付属資料Eの拡張されたアプリケーションにも対応させるためのものである。ちなみに、付属資料Eの開発に当たっては、音楽信号に対する性能において厳しい要求項目があった。一方、標準本体、付属資料BおよびDに関しては、音楽信号に対する性能において厳しい要求項目はなかった。付属資料BのDTX動作を適用した場合の付属資料Eの音楽信号に対する品質を保証するために、音楽信号が継続している区間では、音楽検出機能は強制的にVAD機能を「有音」の状態にする。これにより、音楽の区間では確実に付属資料Eの11.8kbit/sで符号化が行われる。付属資料BにおけるSIDの符号化やCNGは、特に修正せずにそのまま使用できる。なお、不連続伝送の区間では、付属資料EのLPCモード選択アルゴリズムにおけるパラメータを正しく更新することに注意する必要がある（I. 5. 3節、参照）。

I. 5 アルゴリズムの記述

本章では、個々の付属資料のアルゴリズムに対し、統合を容易にするために必要な追加アルゴリズムの記述を示す。残りのすべてのモジュールは、標準本体、付属資料B、DまたはEが元になっている。

I. 5. 1 音楽検出

音楽検出は新しい機能である。これはVADの直後に実行され、音楽区間中は、VADを強制的に「有音」の状態にする。これは付属資料Eの動作中に限り有効であるが、そのパラメータは、統合された標準JTG 729のDTX動作中におけるビットレートモードと独立しており、常に更新される。

音楽検出アルゴリズムは、音楽信号が存在している間の、音声検出（VAD）からの判定を訂正する。これは、付属資料BのDTX、すなわち不連続伝送モード動作中に、付属資料Eと結合して用いられる。音楽検出は、以下のパラメータに基づいている。

- (1) *Vad_deci*、フレームのVAD判定
- (2) *PVad_dec*、直前フレームのVAD判定
- (3) *Lpc_mod*、直前フレームの適応LPCが、フォワードかバックワードかいずれかを示すフラグインジケータ
- (4) *Rc*、LPC分析から得られる反射係数
- (5) *Lag_buf*、最終5フレームの修正された開ループピッチラグのバッファ
- (6) *Pgain_buf*、最終5サブフレームの開ループピッチ利得のバッファ

- (7) *Energy*、LPC分析からの1次自己相関係数 $R(0)$
- (8) *LEnergy*、VADモジュールからの正規化対数エネルギー
- (9) *Frm_count*、処理された信号フレーム数のカウンタ
- (10) *Rate*、音声コーデックの選択

本アルゴリズムは、2つの主要な部分からなる。

- (1) 関連パラメータの計算
- (2) パラメータに基づく分類

I. 5. 1. 1 関連パラメータの計算

本節では、判定モジュールによって使用されるパラメータの計算について記述している。

- (1) 部分正規化残差エネルギー

$$LEnergy = 10 \log_{10} \left(\prod_{i=1}^4 (1 - Rc(i)^2) Energy / 240 \right)$$

- (2) スペクトル差分と背景雑音の部分正規化残差エネルギーの移動平均

現フレームの反射係数 Rc と背景雑音の移動平均反射係数 mRc の間のスペクトル差分量は次式で与えられる。

$$SD = \sum_{i=1}^{10} (Rc(i) - mRc(i))^2$$

移動平均 \overline{mrc} と $mLEnergy$ は、VADモジュールによって生成されたVAD判定 Vad_deci を用いて以下のとおり更新される

```

if Vad_deci == NOISE {
     $\overline{mrc} = 0.9\overline{mrc} + 0.1rc$ 
     $mLEnergy = 0.9mLEnergy + 0.1LEnergy$ 
}

```

- (3) ピッチラグバッファ更新のための開ループピッチラグ補正

開ループピッチラグ T_{op} は、倍ピッチあるいは3倍ピッチを回避するため、以下のように補正して用いる。

```

 $avg\_lag = \sum_{i=1}^4 Lag\_buf(i) / 4$ 
if (abs( $T_{op} / 2 - avg\_lag$ ) <= 2)
     $Lag\_buf(5) = T_{op} / 2$ 
else if (abs( $T_{op} / 3 - avg\_lag$ ) <= 2)
     $Lag\_buf(5) = T_{op} / 3$ 
else
     $Lag\_buf(5) = T_{op}$ 

```

なお、開ループピッチラグ T_{op} は変更されないこと、かつ開ループピッチ分析によって得られたものと同一であることに注意すべきである。

(4) ピッチラグの標準偏差

$$std = \sqrt{Var / 4}$$

$$\text{where } Var = \sum_{i=1}^5 (Lag_buf(i) - \mu)^2 \text{ and } \mu = \sum_{i=1}^5 (Lag_buf(i) / 5)$$

(5) ピッチ利得の移動平均

$$mPgain = 0.8mPgain + 0.2\theta, \text{ where } \theta = \sum_{i=1}^5 (Lag_buf(i) / 5)$$

ピッチ利得バッファ $Pgain_buf$ は、サブフレーム処理の後、もし $Vad_deci = NOISE$ ならば、ピッチ利得値 0.5 で更新され、そうでなければ、量子化されたピッチ利得で更新される。

(6) ピッチラグの滑らかさと有声強度インジケータ

ピッチラグの滑らかさと有声強度インジケータ $Pflag$ は、以下の論理的なステップを使用して生成される。

最初に、2つの中間論理フラグ $Pflag1$ と $Pflag2$ が以下のとおり得られる。

```
if (std < 1.3 and mPgain > 0.45) set Pflag1=1 else 0
if (mPgain > Thres) set Pflag2=1 else 0,
where Thres = 0.73 if Rate=G729D, otherwise Thres=0.63
```

最後に、 $Pflag$ は以下より決定される。

```
if ((PVad_dec == VOICE and (Pflag1==1 or Pflag2==1)) or (Pflag2==1))
set Pflag=1 else 0
```

(7) 定常性カウンタ

1組のカウンタが定義され、以下のとおり更新される。

- (a) $count_consc_rflag$ は、2次反射係数とピッチ利得の移動平均が以下の条件を満足する場合に、連続フレーム数をカウントする。

```
if (Rc(2) < 0.45 and Rc(2) > 0 and mPgain < 0.5)
count_consc_rflag = count_consc_rflag + 1
else
count_consc_rflag = 0
```

- (b) $count_music$ は、64 フレームの窓内で直前のフレームがバックワード適応 LPC を使用しており、現フレームが (VADにより) 「有音」である場合は、フレーム数をカウントする。

```

if (LPC_mod == 1 and Vad_deci == VOICE)
  cont_music = cont_music + 1

```

64 フレーム毎に、*count_music* の移動平均 *mcount_music* は、以下の記述のとおり更新および零にリセットされる。

```

if ((Frm_count mod 64) == 0){
  if (Frm_count == 64)
    mcount_music = count_music
  else
    mcount_music = 0.9mcount_music + 0.1count_music
}

```

(c) *count_consc* は、*count_music* が零のままである場合、連続フレーム数をカウントする。

```

if (count_music == 0)
  count_consc = count_consc + 1
else
  count_consc = 0

if (count_consc > 500 or count_consc_rflag > 150) set mcount_music = 0

```

(b)における *count_music* は、関連するカウンタの更新の後、64 フレーム毎に零にリセットされる。

(c)における論理は、*count_music* の移動平均をリセットするために使用される。

(d) *count_pflag* は、64 フレームの窓内で *Pflag* = 1 であるフレーム数をカウントする。

```

if (Pflag == 1)
  count_pflag = count_pflag + 1

```

64 フレーム毎に、*count_pflag* の移動平均 *mcount_pflag* は更新され、以下の記述のとおり零にリセットされる。

```

if ((Frm_count mod 64) == 0){
  if (Frm_count == 64)
    mcount_pflag = count_pflag
  else{
    if (count_pflag > 25)
      mcount_pflag = 0.98mcount_pflag + 0.02count_pflag
    else if (count_pflag > 20)
      mcount_pflag = 0.95mcount_pflag + 0.05count_pflag
    else
      mcount_pflag = 0.9mcount_pflag + 0.1count_pflag
  }
}

```

(e) *count_consc_pflag* は、以下の条件を満足する連続フレーム数をカウントする。

```
if (count_pflag == 0)
    count_consc_pflag = count_consc_pflag + 1
else
    count_consc_pflag = 0

if (count_consc_pflag > 100 or count_consc_rflag > 150) set mcount_pflag = 0
```

count_pflag は 64 フレーム毎に零にリセットされる。(e)における論理は *count_pflag* の移動平均をリセットするために使用される。

I. 5. 1. 2 分類

上記のパラメータ推定に基づいて、VADモジュールからのVAD判定 *Vad_deci* は、もし以下の条件が満足されるのならば判定をやり直す。

```
if (Rate = G729E){
    if (SD > 0.15 and (Lenergy - mLenergy) > 4 and LLenergy > 50)
        Vad_deci = VOICE
    else if ((SD > 0.38 or (Lenergy - mLenergy) > 4) and LLenergy > 50)
        Vad_deci = VOICE
    else if ((mcount_pflag >= 10 or mcount_music >= 1.0938 or Frm_count < 64) and LLenergy > 7)
        Vad_deci = VOICE
}
```

注意すべきことは、音楽検出機能は、メモリを最新の状態に維持するために、動作中の符号化モードに関わらず、常に呼び出されることである。ただし、VAD判定 *Vad_deci* については、統合された標準 J T - G 7 2 9 が 11.8kbit/s (付属資料 E) で動作しているときに限り変化する。また、音楽検出は、VAD判定に対し「無音」から「有音」へ判定を変えることだけを許容し、その逆はなされないことに注意すべきである。

I. 5. 2 不連続伝送時の付属資料 D に特有な状態変数の更新

付属資料 D に特有な唯一の状態変数は、復号器における位相拡散モジュールの状態変数である (標準 J T - G 7 2 9 付属資料 D [3] D. 6. 2 節、参照)。無音フレームでは、擬似背景雑音発生器によって計算された利得値 (標準 J T - G 7 2 9 付属資料 B [2] B. 4. 4 節、参照) を、適応および ACELP 利得の推定値として、標準ビットレート (8kbit/s) と同じ更新手順が実行される。高レートに対する更新が標準ビットレートに対する更新と同一であることに注意すること。

I. 5. 3 不連続伝送時の付属資料 E に特有な状態変数の更新

I. 5. 3. 1 付属資料 E に特有な符号器の状態変数の更新

符号器において、無音フレームの場合、状態変数の更新は標準 J T - G 7 2 9 付属資料 E [4] における標準ビットレート 8kbit/s に切り替える場合と同一である。更新手順は以下の通りである。LP モードを 0 にセットし、グローバル定常性インジケータの値を減少させ、高定常性インジケータを 0 にリセットする (標準 J T - G 7 2 9 付属資料 E [4] E. 3. 2. 7. 2 節、参照)、フォワード LP フィルタからバックワード LP フィルタへのなめらかな切り替えを行うための補間係数を、最大値にリセットする (標準 J T - G 7 2 9

付属資料E[4] E. 3. 2. 7. 1節、参照)。この更新は低ビットレートである6.4kbit/sへ切り替える場合も使用される。

I. 5. 3. 2 付属資料Eに特有な復号器の状態変数の更新

復号器において、無音フレームの場合、状態変数の更新はピッチ追跡手順による計算（標準JT-G729付属資料E[4] E. 4. 4. 5節、参照）を実行する代わりに、ピッチ遅延定常性インジケータを0にリセットすることを除いて、標準JT-G729付属資料E[4]におけるフォワードモードのみのレート（8kbit/sおよび6.4kbit/s）に切り替える場合と同一の処理を行う。

I. 6 Cソースコードの記述

標準JT-G729付属資料I（標準本体、付属資料B、D、Eの統合版）は、16ビット固定小数点ANSI Cコードでシミュレーションされており、標準JT-G729に対応するソフトウェアと同じタイプの固定小数点データおよび同じセットの固定小数点基本演算を用いている。ANSI Cコードは付属資料Iの仕様を規定している。Cコードによって与えられるアルゴリズム記述は、標準JT-G729本体、付属資料B、D、EおよびIの本文の記述よりも優先される。以下に、シミュレーションコードの使い方とソフトウェアの構成について概説する。

I. 6. 1 シミュレーションソフトウェアの使い方

Cコードは2つのメインプログラム `coderi.c` と `decoderi.c` からなり、それぞれ符号器と復号器のシミュレーションを行う。符号器は以下のように実行される。

```
coderi inputfile bitstreamfile dtx_option rate_option
```

復号器は以下のように実行される。

```
decoderi bitstreamfile outputfile
```

入力ファイルと出力ファイルは、8kHzで標準化された16ビットのPCM信号からなるデータファイルである。ビットストリームファイルは、ビットストリームを含んだバイナリファイルである。符号化されたビットストリームのマッピングテーブルは、シミュレーションソフトウェアに含まれる。符号器に対する2つのパラメータは、`dtx_option` と `rate_option` である。

`dtx_option` = 0 : DTX無効

= 1 : DTX有効

初期値は0 (DTX無効)

`rate_option` = 0 : 低レート (6.4kbit/s) を選択

= 1 : 標準本体 (8kbit/s) を選択

= 2 : 高レート (11.8kbit/s) あるいはフレームバイフレームでのレート選択するため、

1, 2のいずれかを含む16ビット語のバイナリファイル `file_rate_name` を選択

初期値は1 (8kbit/s)

I. 6. 2 シミュレーションソフトウェアの構成

ファイルは、4つのグループに分類できる。

- (1) 標準JT-G729標準本体[1]、付属資料B[2]、付属資料D[3]、または付属資料E[4]に対応するソフトウェアファイルと同一なファイル。この一覧を [Table I-6-1/JT-G729](#) に示す。
- (2) 標準JT-G729付属資料B、付属資料D、付属資料Eに対応するソフトウェアファイルを適切に修正したファイル。統合に対応するため、軽微な修正が加えられている。この一覧を [Table I-6-2/JT-G729](#) に示す。ほとんどの変更は各付属資料の型宣言ファイルを1つのファイル (`ld8cp.h`) に

- 統合するためのものであり、また各付属資料のファイルにおける外部ROM宣言を1つのファイル (tabld8cp.h) に統合するためのものである。いくつかは、状態変数の更新を扱うために導かれた。
- (3) 標準JT-G 7 2 9 付属資料B、付属資料D、付属資料Eに対応するソフトウェアファイルを統合したファイル。この一覧を Table I - 6 - 3 / JT-G729 に示す。
 - (4) 本標準JT-G 7 2 9 統合版で規定したファイル (新ファイル) 。この一覧を Table I - 6 - 4 / JT-G729 に示す。

Table I – 6 – 1 / JT-G729 List of software files identical to software files of G.729 main body, Annex B, D or E
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Identical to</i>
Basic_op.c	Basic operators	Main
Oper_32b.c	Extended basic operators	Main
Dspfunc.c	Mathematical functions	Main
Gainpred.c	Gain predictor	Main
lpcfunc.c	Miscellaneous routines related to LP filter	Main
Pre_proc.c	Pre-processing (HP filtering and scaling)	Main
P_parity.c	Compute pitch parity	Main
pwf.c	Computation of perceptual weighting coefficients (8 kbit/s)	Main
Pred_lt3.c	Generation of adaptive codebook	Main
Post_pro.c	Post processing (HP filtering and scaling)	Main
Tab_ld8k.c	ROM tables	Main
Basic_op.h	Basic operators prototypes	Main
Ld8k.h	Function prototypes	Main
Oper_32b.h	Extended basic operators prototypes	Main
Tab_ld8k.h	Extern ROM table declarations	Main
Typedef.h	Data type definition (machine dependent)	Main
Taming.c	Pitch instability control	B
Qsidgain.c	SID Gain Quantization	B
Qsidlsf.c	SID-LSF Quantization	B
Tab_dtx.c	ROM tables	B
Sid.h	Prototype and Constants	B
Octet.h	Octet transmission mode definition	B
Tab_dtx.h	Extern ROM table declarations	B
Pwfe.c	Computation of perceptual weighting coefficients (11.8 kbit/s)	E

Table I – 6 – 2 / JT-G729 List of software files adapted from software files of G.729 main body/B/D/E
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Adapted from</i>
Vad.c	VAD	B
Dtx.c	DTX Decision	B
Vad.h	Prototype and Constants	B
Dtx.h	Prototype and Constants	B
Calcexc.c	CNG Excitation Calculation	B
Dec_sid.c	Decode SID Information	B
Utilcp.c	Utility functions	B
Phdisp.c	Phase dispersion	D
Bwfw.c	Backward/forward switch selection	E
Bwfwfunc.c	Miscellaneous routines related to backward/forward switch selection	E
Filtere.c	Filter functions	E
Lpccp.c	LP analysis	E
Lspcdece.c	LSP decoding routines	E
Lspgetqe.c	LSP quantizer	E
Qua_lspe.c	LSP quantizer	E
Track_pi.c	Pitch tracking	E
utilcp.c	Utility functions	B

Table I – 6 – 3 / JT-G729 List of software files integrating software files from G.729 main, Annex B, D or E
(ITU-T G.729)

<i>File name</i>	<i>Description</i>	<i>Integrated from</i>
Coderi.c	Main encoder routine	B+D+E
Codld8cp.c	Encoder routine	B+D+E
Decoderi.c	Main decoder routine	B+D+E
Declld8cp.	Decoder routine	B+D+E
Acelp_cp.c	search ACELP fixed codebook (6.4, 8, 11.8 kbit/s)	D+E
Dacelpcp.c	Decode algebraic codebook (6.4, 8, 11.8 kbit/s)	D+E
Pitchcp.c	Pitch search	D+E
Declagcp.c	Decode adaptive-codebook index	D+E
Q_gaincp.c	Gain quantizer	D+E
Degaincp.c	Decode gain	D+E
Pstpcp.c	Postfilter routines	B+E
Bitscp.c	Bit manipulation routines	B+D+E
Tabld8cp.c	ROM tables for G.729 at 6.4 and 11.8 kbit/s	D+E
Tabld8cp.h	Extern ROM declarations for G.729 at 6.4 and 11.8 kbit/s	D+E
Ld8cp.h	Constant and Function prototypes for G.729 at 6.4 and 11.8 kbit/s	D+E

Table I – 6 – 4 / JT-G729 List of software files specific to integrated G.729 Annexes B, D and E
(ITU-T G.729)

<i>File name</i>	<i>Description</i>
Mus_dtct.c	Music detection module

付録 1

(標準 J T-G 7 2 9 に対する)

外部 VAD/DTX/CNG を用いたシステムにおける標準 J T-G 7 2 9 コーデック に対する外部同期リセット性能

(本付録は参考資料であり、仕様ではない。)

概要

本付録は、標準 J T-G 7 2 9 本体、標準 J T-G 7 2 9 付属資料 A、または標準 J T-G 7 2 9 付属資料 C とともに外部の VAD/DTX/CNG を用いたシステム (例えば CME (回線多重化装置)) における外部同期リセットの能力について論じるものである。

外部同期リセットの用途として、標準 J T-G 7 2 9 本体、標準 J T-G 7 2 9 付属資料 A、または標準 J T-G 7 2 9 付属資料 C とともに外部の VAD/DTX/CNG を用いたシステムを意図している。この条件においては、VAD が用いられている場合に、背景雑音の多い環境下において可能な限り最良の音声品質を得るためには、外部同期リセットを用いることが一般に好ましい。このことは、特に積極的な VAD を用いた場合に当てはまる。外部の VAD が十分に長いハングオーバー区間を有する (すなわち消極的な VAD の) 場合、「リセットを用いない」場合と比較した外部同期リセットによる音質の改善は、ほとんど知覚できない。

本付録の規定範囲

標準 J T-G 7 2 9 付属資料 B では「固有の」(もしくは内部の) VAD/DTX/CNG メカニズムを定義しているが、いくつかのアプリケーションでは、システムあるいは複雑度の制約のために、異なるアルゴリズムの利用が要求される。このような事例では、外部の VAD/DTX/CNG アルゴリズム (すなわち、独立して動作し、符号器内部の情報を利用しないもの) が用いられる場合に、符号器と復号器との状態が著しく異なる可能性があり、それが品質を劣化させてしまう。それゆえ、このような外部の VAD/DTX/CNG アルゴリズムが用いられる場合に、符号器と復号器との同期リセットが全体的な品質に有益であり得る。本付録は、標準 J T-G 7 2 9 本体、J T-G 7 2 9 付属資料 A、または標準 J T-G 7 2 9 付属資料 C とともに外部の VAD/DTX/CNG を用いたシステム (例えば CME (回線多重化装置)) における外部同期リセットの能力について論じるものである。

1. 1 はじめに

同期リセットの定義とは、符号器の状態変数と復号器の状態変数とが、同じフレーム時刻において各々の初期値に設定されることである。

外部同期リセットの用途として、標準 J T-G 7 2 9 本体、標準 J T-G 7 2 9 付属資料 A、または標準 J T-G 7 2 9 付属資料 C とともに外部 VAD/DTX/CNG を使用するシステムを意図している。この条件では、VAD が用いられている場合に、背景雑音の多い環境下において可能な限り最良の音声品質を得るためには、外部同期リセットを用いることが一般に好ましい。このことは、特に比較的短いハングオーバー区間を持つ積極的な VAD を使用する場合に当てはまる。外部 VAD が十分に長いハングオーバー区間を有する (すなわち消極的な VAD の) 場合、外部同期リセットを追加したことによる「リセットを用いない」ものと比較した品質の改善はほとんど知覚できない。どのような場合においても、外部 VAD/DTX/CNG を用いたシステムにおける標準 J T-G 7 2 9 符号器および復号器に対して同期リセットを適用することによる品質への悪影響は考えなくてよい。逆に、標準 J T-G 7 2 9 アルゴリズムは同期外れ後に迅速な収束をするものの、同期リセットの利用が、可能な限り最良の音声品質を一般に達成し得るであろうことの証拠がある。

1. 2 実験計画

標準 J T-G 7 2 9 標準本体、標準 J T-G 7 2 9 付属資料 A、または標準 J T-G 7 2 9 付属資料 C とともに外部の VAD/DTX/CNG を用いたシステム（例えば CME（回線多重化装置））に対して標準 J T-G 7 2 9 コーデックの同期リセットを導入することによる品質への影響を調べるために、いくつかの限定的な実験が実施された。この実験は、標準 J T-G 7 2 9 付属資料 C（および標準 J T-G 7 2 9 本体）を用い、プール構成のコーデックを持つ CME 動作のシミュレーションに限定された。この CME 動作においては、符号器と復号器との間の「1 対 1 対応」は呼を通じて期待することはできない。そのため、符号器と復号器との間の同期外れが起きる。

標準 J T-G 7 2 9 コーデックへの同期リセットの導入効果を調べるために、（同期リセット有りおよび同期リセット無し）の両方式の品質を評価するためのいくつかの実験が実施された。入力レベルが通常、高、低のクリーンな音声、および異なる信号対雑音比（SNR）を持った異なるタイプの背景雑音（Babble 雑音、Hall 雑音、Car 雑音）を含んだ音声、といった様々な試験条件が用いられた。各条件において、男性と女性各 1 名ずつの話者が用いられた。一方は北米英語、他方はフランス語の、2 つのエキスパートによる受聴試験が実施された。それぞれの実験には独自の外部 VAD インジケータが用いられた。

プール構成のコーデックを持つ CME 動作をシミュレートするために、標準 J T-G 7 2 9 復号器への入力ビットストリームは、2 つの異なる標準 J T-G 7 2 9 符号器から出力された 2 つのビットストリームファイルをインターリーブすることによって構成された。インターリーブは、2 つの入力ファイルそれぞれの VAD にしたがって行われた（ファイル 1 の最初の有音部分、ファイル 2 の最初の有音部分、ファイル 1 の第 2 の有音部分、ファイル 2 の第 2 の有音部分、等）。最後に、復号器の出力ファイルは、インターリーブのやり方にしたがって 2 つの復号ファイルに分割された。同期リセットが用いられた場合は、符号器と復号器の両方がそれぞれの有音区間の先頭でリセットされ、他の場合にはリセットされなかった。

1. 3 試験結果

2 つの方式の品質に与える影響を評価するために、エキスパートによる非公式な受聴試験が復号された音声ファイルの有音区間の対比較を用いて実施された。その結果は、外部 VAD と、二つのインターリーブされたファイルにおける背景雑音の類似性とに依存したものだ。外部 VAD が十分に長いハングオーバー区間を有する（すなわち、消極的な VAD の）場合、二つのインターリーブされたファイルにおける背景雑音が同程度もしくは高 SN 比の時には、二つの方式は同様の品質であった。また、人工的な音は感じられなかった。低 SN 比の背景雑音の区間に高 SN 比の背景雑音の区間をインターリーブした場合、同期外れの後の標準 J T-G 7 2 9 アルゴリズムにおける迅速な収束のおかげで時間的に短いものではあったが、有音区間の先頭において多少人工的な音が聞こえた。より積極的な VAD を用いた場合、同期リセット方式には明確な改善が認められた。

1. 4 結論

標準 J T-G 7 2 9 コーデックに同期リセットを導入した場合の音声品質への影響を調べるために、いくつかの限定的な実験が行われた。CME での利用では、標準 J T-G 7 2 9 の符号器および復号器に同期リセットを適用することによる音質の劣化は起きないという予測の証明が確認されている。さらに、VAD が用いられている場合に、背景雑音の多い環境下において可能な限り最良の音声品質を得るためには、同期リセットの導入が一般に好ましいということがわかった。この結果は、標準 J T-G 7 2 9 とともに外部の VAD/DTX/CNG を用いた他のシステムにも拡張が可能であることが期待される。

1. 5 参考文献

- [1] TTC標準 JT-G 7 2 9 : 8kbit/s C S - A C E L P を用いた音声符号化方式 (1996)
- [2] TTC標準 JT-G 7 2 9 付属資料 A : 低演算量版 8kbit/s C S - A C E L P 音声コーデック (1997)
- [3] TTC標準 JT-G 7 2 9 付属資料 B : I T U - T 勧告 V . 7 0 端末に適した標準 JT-G 7 2 9 に対する無音圧縮手法 (1997)
- [4] TTC標準 JT-G 7 2 9 付属資料 C : 浮動小数点演算版 8kbit/s C S - A C E L P 音声コーデック (1999)

付録 2

(標準 J T - G 7 2 9 に対する)

V o I P における T T C 標準 J T - G 7 2 9 付属資料 B の機能拡張 - オプション 1

(本付録は参考資料であり、仕様ではない。)

概要

本付録は、J T - G 7 2 9 本体および付属資料 A, C または H とともに付属資料 B が V o I P に適用される際の、その付属資料 B に記述されている V A D の最適化について扱うものである。

本付録では、特に、以下の問題を扱う。

- (1) 15dB 未満のレベルで始まる入力信号に対する望ましくない性能
- (2) C N G 部で生じる耳障りな呼気音のような雑音
- (3) 高雑音条件下での V A D 性能
- (4) 誤った変数の初期化

本付録は、帯域利用効率の最適化よりも総合的な音質をより重視するような V o I P システムでの使用を意図している。

2. 1 本付録の規定範囲

J T - G 7 2 9 付属資料 B は、V A D / D T X / C N G の動作を規定しているが、アプリケーションによっては特定の制約のため別の V A D アルゴリズムの使用が要求される場合がある。特に、V o I P がそうであり、付属資料 B に記述されている現状のアルゴリズムでは、下記の動作において不適切なふるまいをする。

- (1) 15dB 未満のレベルで始まる入力信号に対する望ましくない性能
- (2) C N G 部で生じる耳障りな呼気音のような雑音
- (3) 高雑音条件下での V A D 性能劣化

更に、

- (4) 現状の付属資料 B においては誤った変数の初期化が存在する。

本付録は、現状の付属資料 B の代替として、上述の (1) から (4) の項目に対応する修正提案について記述している。

2. 2 参照すべき標準

[1] T T C 標準 J T - G 7 2 9 8 kbit/s C S - A C E L P を用いた音声符号化方式

[2] T T C 標準 J T - G 7 2 9 付属資料 A 低演算量版 8 kbit/s C S - A C E L P 音声コーデック

[3] T T C 標準 J T - G 7 2 9 付属資料 B I T U - T 勧告 V . 7 0 端末に適した標準 J T - G 7 2 9 に対する無音圧縮手法

[4] T T C 標準 J T - G 7 2 9 付属資料 C 浮動小数点演算版 8 kbit/s C S - A C E L P 音声コーデック

2. 3 略語

D T X Discontinuous Transmit Mode

D S V D digital simultaneous voice and data

S I D Silence Insertion Descriptor

S N R Signal-to-noise

V A D Voice Activity Detector

2. 4 概要

TTC標準JT-G729付属資料Bは、ITU-T勧告V. 70のDSVD (digital simultaneous voice and data) に最適化された標準JT-G729に対する無音圧縮手法である。JT-G729付属資料Bは、当初はそのDSVD向けに設計されたものであるが、VoIPに多く用いられるようになってきており、今後も利用され続けると想定される。JT-G729付属資料Bは、JT-G729 (およびその付属資料) が、音声検出器 (VAD) を用いてクラス分けされた有音と無音という2つの伝送モードで動作することを可能にする。不連続伝送モード (DTX) は、無音挿入記述子 (SID) を用いて無音フレームを代表するフレームを決定するのに用いられる。近年、JT-G729付属資料BをVoIPに用いた際の問題が報告されるようになった。

2. 5 VoIPにおけるJT-G729付属資料Bの問題

報告された問題で本付録が扱うものは次の通りである。

- (1) 15dB 未満のレベルで始まる入力信号に対する望ましくない性能
- (2) 無音フレームにおける耳障りな呼気音のような雑音
- (3) 高雑音条件下でのVAD性能

問題1 : 15dB 未満レベルで始まる入力信号に対する望ましくない性能

もし、初期化後の32フレーム (320ms) の入力フレームのエネルギーが15dB 未満である場合、JT-G729Bは、以降の15dB 以上の全ての入力フレームを有音と判定してしまう。これは、JT-G729Bを用いる効果を損なわせることになる。すなわち、JT-G729Bを動作させるための余分な演算量やメモリを消費しながら、帯域利用効率の向上に全く寄与できないことになる。

提案する解決策は、背景雑音の特性が適切に初期化されるまで、320ms 毎に初期化フレームカウンタを再スタートさせることである。

問題2 : 無音フレームにおける耳障りな呼気音のような雑音

デジタル的な無音または非常に低レベルの雑音が、129 フレーム (1.29 秒) 以上のトーン信号または他の定常信号の後に続いた場合に、最初のSIDフレームの雑音利得が非常に高いレベルとして推定されてしまう。これにより、レベルの高い呼気音のような雑音が生じ、音声品質の劣化を引き起こす。本問題は、アルカテル社やテキサスインスツルメンツ社の様々な顧客からも報告されている。提案する解決法は、有音フレームの終端にてハングオーバを導入することである。

問題3 : 高雑音条件下でのVAD性能

SNRが15dB 未満の場合、現在のVADには2つの問題がある。一つは、有音と雑音の判定結果が相互間でふらつき、SIDの更新により、帯域利用効率向上の効果が低減することである。もう一つは、有音区間中に雑音への判定が頻繁に発生することである。

上記問題に対する本付録での修正は、現状の符号器に対してはわずかであり、復号器に対しては修正は全くない。問題解決のための狙いは、雑音と判定される区間をできるだけ多くするのではなく、有音信号からの情報をできる限り失わないようにすることである。

提案する解決法は、特に低SNR ($\leq 15\text{dB}$) での有音を無音と判定する誤判定を低減することであり、背景雑音更新の条件および有音検出判定結果平滑化の条件の修正からなる。

現状の符号器に導入した3つの変更は以下の通りである。

- 現フレームと過去のフレームの間の雑音の差が非常に小さい場合に NOISE を修正するという判定処理を行わないようにする。

- 背景雑音の平均エネルギーが十分大きい場合には、VOICE から NOISE への切替の判定においてヒステリシスを導入する。VOICE から NOISE への切替の前に、少なくとも 6 フレーム連続する NOISE の検出が必要であるようにする。
- 背景雑音特性の移動平均を更新するために必要な条件を修正する。

これら 3 つの問題点に対処するために変更された C コードは、日本 I T U 協会から入手可能である。

2. 6 実験方法

これらの機能拡張の導入による影響を調べるために、実験が行われた。

問題 1 : 15dB未満のレベルで始まる入力信号の望ましくない性能

T T C 標準 J T - G 7 2 9 付属資料 B の V A D は、平均全域のエネルギー \bar{E}_f 、平均低域エネルギー \bar{E}_l 、平均零交差数 \bar{ZC} 、および平均スペクトルパラメータ $\left\{ \overline{LSF}_i \right\}_{i=1}^p$ 、といった背景雑音特性の移動平均値をもつ。 \bar{E}_f 、 \bar{E}_l 、 \bar{ZC} 、および $\left\{ \overline{LSF}_i \right\}_{i=1}^p$ は、vad_init関数の開始時にゼロに初期化される。T T C 標準 J T - G 7 2 9 付属資料 B のソースコードでは、T T C 標準 J T - G 7 2 9 付属資料 B の V A D は、入力信号の開始 32 フレームの中で 15dB より大きいエネルギー E_f を持つフレームのみを用いて、背景雑音特性の移動平均の最初の更新を行う。最初の 32 フレームすべてのエネルギーが 15dB を下回る場合、背景雑音特性の移動平均は更新されない。實際上、開始 320ms の入力は、チャンネル設定時間なので、実際の背景雑音を反映していないかもしれない。電話がチャンネル接続の最初に消音モードやオンフックの場合は、入力信号は通常 15dB を下回る。この場合、背景雑音特性の移動平均は、最初の更新をしないですべてゼロに設定されてしまう。これにより、T T C 標準 J T - G 7 2 9 付属資料 B の V A D は、このチャンネルにおいて、すべてのフレームを音声フレームであると判定してしまう。

最初の更新を現在よりもロバストにするために、呼び出しのセットアップの開始時に 32 フレームすべてのエネルギーが 15dB を下回る場合に、フレームカウンタを再スタートさせることを提案した。提案版の T T C 標準 J T - G 7 2 9 付属資料 B のソースコードは、現在の T T C 標準 J T - G 7 2 9 付属資料 A および B、および付属資料 B のテストベクトルに準じている。

問題 2 : 無音フレームにおける耳障りな呼気音のような雑音

異なるレベルを持った異なるトーン信号（特に、フランスおよび米国の呼出音）がテストされた。

問題 3 : 雑音環境下での V A D 性能

正規入力レベル、高入力レベルおよび低入力レベルの雑音のない音声、および異なる信号対雑音比 (S N R) を持った異なるタイプの背景雑音（白色雑音、バブル雑音、ファン雑音）といった、さまざまなテスト条件が用いられた。

2. 6. 1 平滑化の 4 段階目

音声検出の平滑化は 4 段階に分けられている (T T C 標準 J T - G 7 2 9 付属資料 B の B. 3. 6 節を参照のこと)。平滑化の 4 段階目は、以下のようになる。

$if((E_f < \bar{E}_f + 614) \text{ and } (Frm_count > 128) \text{ and } (v_flag = 0) \text{ and } (rc < 19661)) \text{ then } (marker = noise)$

まず、この判定文は、以前の検出結果を考慮しない単なる平滑化のための判定文であり、一方、この平滑化関数の目的は、「音声信号の長期的な定常性を反映」することである。その上、この判定文では、現在のフレームが音声フレームの場合は、背景雑音がどのようなものでも、 E と \bar{E}_f との差が常に最低 614 あるということ仮定している。実際、背景雑音のエネルギーがあまりに大きい場合は、この仮定は正しくなく、この判定は

VADにおいて誤ってしまう。

平滑化の4段階目は削除された。ブール代数 v_flag は、この判定文においてのみ利用されるものであり、削除されている。

2.6.2 ヒステリシス

雑音環境下では、差分パラメータのベクトルの係数がより小さいため、初段の多境界音声検出関数が音声と雑音とを明確に区別することがより難しくなる。それが、このような場合に初段のVADの出力がよりとっぴな値になる理由である。この振舞を回避し、より一貫した出力信号を得るために、平滑化関数の最初にヒステリシスが追加された。

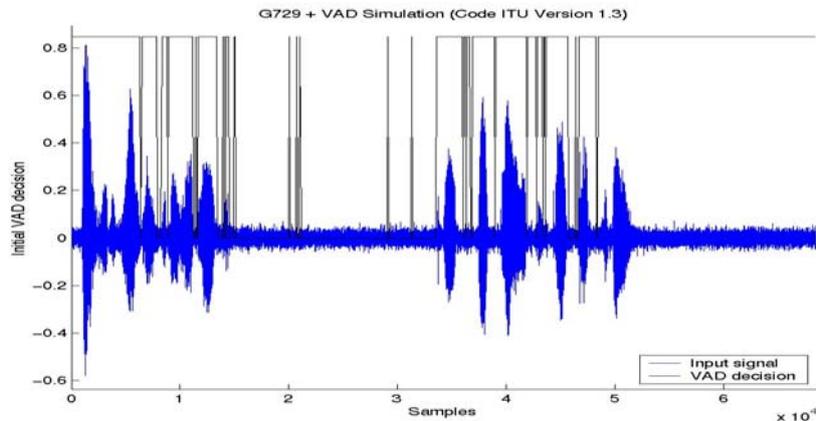
```
if(marker = VOICE)then(Count_inert = 0)
if((marker = NOISE)and(Count_inert < 6)and(MeanSE > 8000))then{
    Count_inert ++;
    marker = VOICE; }
```

両判定文間の相互作用を回避するために、平滑化の3段階目（264行目）において、等式 $Count_inert = 6$ が追加されている。

2.6.3 更新判定

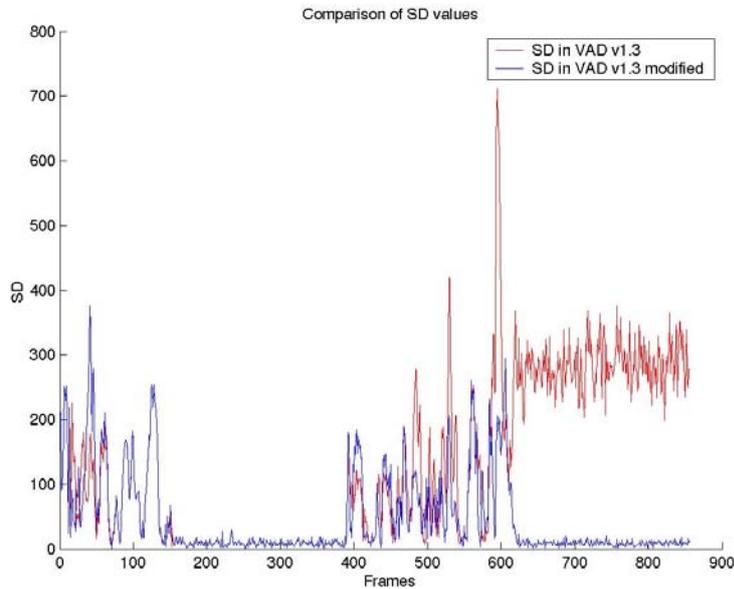
標準（B.3.7節）によると、背景雑音特性の移動平均は、 $(E_f < \bar{E}_f + T_6)$ の場合に更新される。Cソースコードでは、この条件は、 $((E_f < \bar{E}_f + 614)and(rc < 24576)and(\Delta S < 83))$ で置き換えられている。

TTC標準JT-G 7 2 9 付属資料Bに含まれるテストベクトル $tstseq3.bin$ を考えてみる。

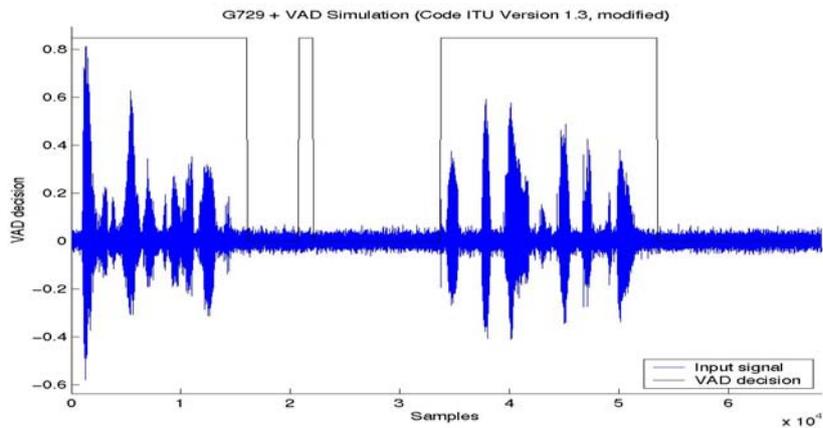


上の図は、入力信号、および初段の多境界VADによる検出結果を示す。信号の最後の部分は、雑音のみ存在する部分であるが、送信されていないと思われる。それにも関わらず、初段のVADの判定は、この雑音が長い時間継続しても、1で一定である。これは、あまりに ΔS の値が高いので、常に上記の更新の条件を満たしていないという事実によるものである。次の図は、現在の判定文での ΔS （赤線）、および $(SD < 83)$ という条件を外した ΔS （青線）である。¹

¹ ΔS は、スペクトル距離であり、CソースコードではSDという変数である。



この条件を外せば、雑音環境下にあると考えられるように、移動平均が正しく更新され、 ΔS の値が非常に小さくなる。VADアルゴリズムによる判定結果もまた正しい。



2. 7 電子ファイル

2セットの電子ファイルが日本ITU協会から入手可能である。

1セット目は、修正されたvad.cファイルであり、このファイルは、本付録の記述にしたがってTTC標準JT-G729付属資料BのVADアルゴリズムにおける変更を実装したものである。2セット目は、本付録におけるアルゴリズムをテストするために使用されたテストファイルセットである。

付録 3

(標準 J T - G 7 2 9 に対する)

V o I P における T T C 標準 J T - G 7 2 9 付属資料 B の機能拡張 - オプション 2

(本付録は参考資料であり、仕様ではない。)

概要

本付録は、J T - G 7 2 9 本体および付属資料 A, C または H とともに付属資料 B が V o I P に適用される際の、その付属資料 B に記述されている無音圧縮手法の最適化について扱うものである。

本付録は特に、T T C 標準 J T - G 7 2 9 付属資料 B に対して報告されている以下の問題点に対処している。

- (1) 呼の開始時における背景雑音の特徴パラメータの初期化、背景雑音特性が変化した場合における背景雑音の特徴パラメータの更新
- (2) 急峻なエネルギーの立下り時に呼気音のような雑音を生成する S I D パラメータの早期推定
- (3) 「無音」として長く継続する、高レベルなトーン状信号部分の識別
- (4) 特定のタイプの背景雑音に対する無音 - 有音間の頻繁な変化
- (5) 頻繁な S I D 更新フレーム。これは、このようなフレームに対するペイロードよりも、相当に大きいパケットオーバーヘッド情報を有しているため、V o I P 用途には不適切である。

本付録は、T T C 標準 J T - G 7 2 9 付属資料 B に対して報告されているこれらの問題点の解決手法を提供する。本付録は、V o I P システムを対象としており、帯域利用効率を維持しつつ、T T C 標準 J T - G 7 2 9 付属資料 B に対する品質を改善するものである。

3. 1 本付録の規定範囲

T T C 標準 J T - G 7 2 9 付属資料 B は、I T U - T 勧告 V. 7 0 適合端末に適した T T C 標準 J T - G 7 2 9 インプリメンテーションに対する無音圧縮手法を提供する。この無音圧縮手法は、音声検出 (V A D)、不連続伝送 (D T X)、無音挿入記述子 (S I D)、そして擬似背景雑音発生 (C N G) を含んでいる。この標準の用途は、V. 7 0 端末のみならず、音声パケット網 (例えば V o I P) においても、広く共通的に利用されており、そこでも、帯域を有効に利用した通信のために、不連続伝送アルゴリズムおよび音声検出が必要とされている。音声パケット網に対する T T C 標準 J T - G 7 2 9 付属資料 B の動作に関していくつかの問題点が報告されている。本付録は、T T C 標準 J T - G 7 2 9 付属資料 B において、その報告されている問題点に対する解決方法を記述しており、音声パケット網用途に適し、高い帯域利用効率を維持しつつ、音声品質を改善するものである。

3. 2 報告されている T T C 標準 J T - G 7 2 9 付属資料 B の問題点に対する解決手法

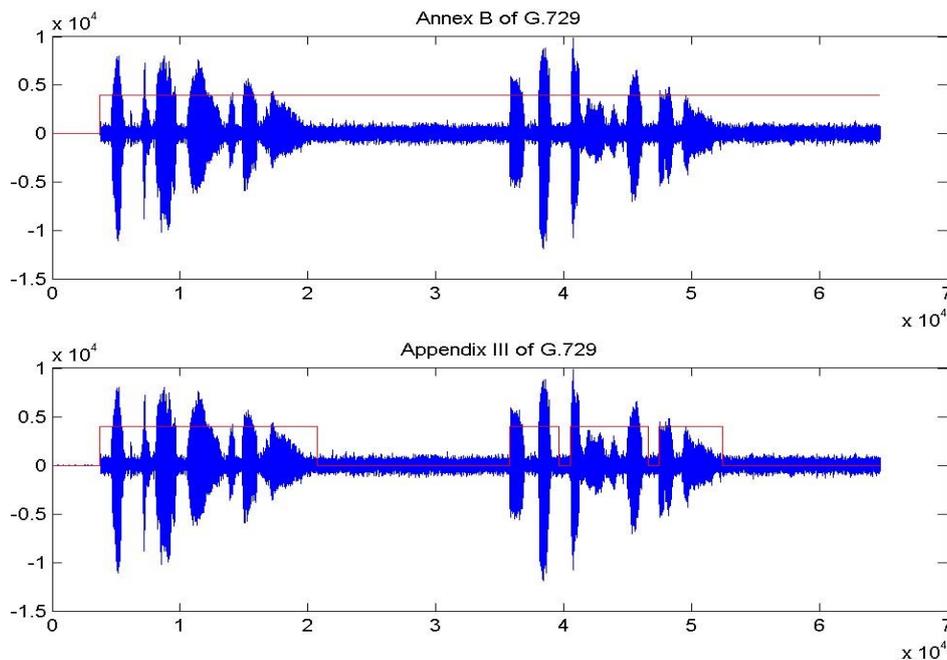
T T C 標準 J T - G 7 2 9 付属資料 B に関して報告されている問題点は、いくつかの変更で解決されている。この変更は、日本 I T U 協会から入手可能な C プログラムファイル vad.c と dtx.c の中で、V A D_V O I P_A P P_ I I I フラグに基づいている。このフラグをプロジェクトファイルあるいはメイクファイルで定義する必要がある。

3. 3 報告されている T T C 標準 J T - G 7 2 9 付属資料 B の問題点に対する解決例

パケット音声通信への適用 (例えば V o I P) において、T T C 標準 J T - G 7 2 9 付属資料 B の V A D および D T X の動作に関し、以下の問題点が報告されている。付録 3 を適用することによって、これらの問題点を解決することができる。

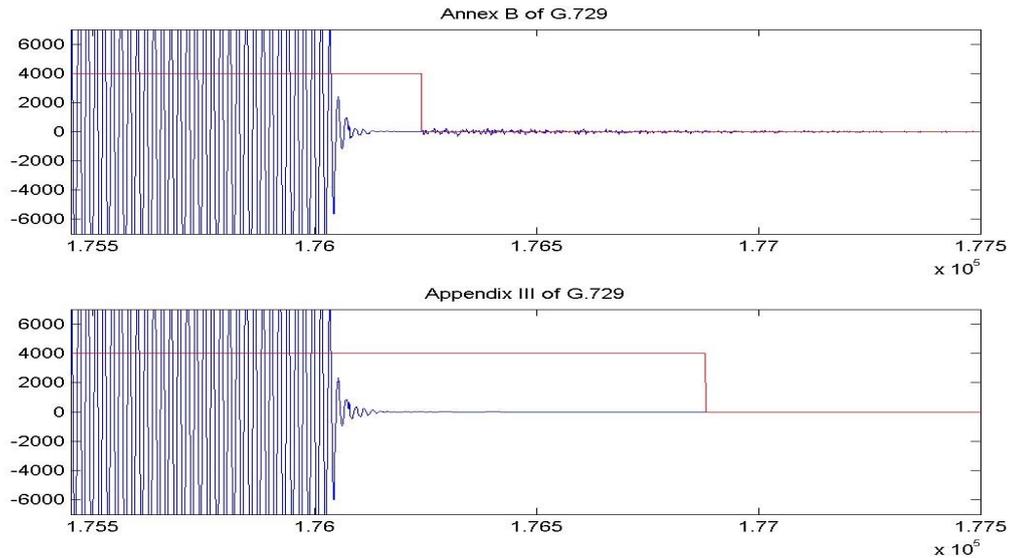
1. 初期化時の背景雑音の特徴パラメータは、最初の 320msec の間に 15dB の閾値にて計算される。これらの 2 つの制約事項によって、算出された初期化時の背景雑音特徴パラメータが、結果的に役に立たなくなること

がある。例えば、実際の音声チャンネルが確立される前に、音声パケットネットワークプロセッサがTTC標準JT-G729符号器を起動させた場合、算出された初期化時の背景雑音特徴パラメータが、結果的に役に立たなくなる。背景雑音のタイプあるいはレベルが突然変化した場合にも、同様な振る舞いが起きる可能性がある。この特徴パラメータの更新がうまくいかないために、その後が続く背景雑音フレームを「有音フレーム」と判定してしまうという結果をもたらすこともあり、結果的に、TTC標準JT-G729付属資料Bで規定される無音圧縮方式から期待される帯域利用効率を損ねてしまう。付図3-3-1/JT-G729は、一例として、TTC標準JT-G729付属資料Bにおいて背景雑音特徴パラメータが結果的に役に立たなくなるケース、およびこれをTTC標準JT-G729付録3でどのように対処しているかを示している。



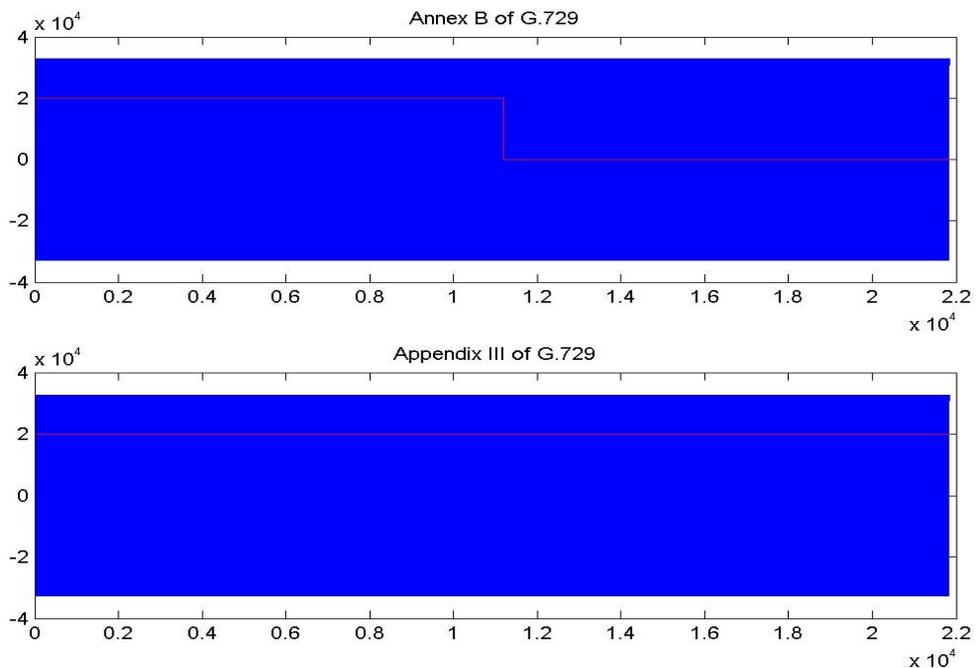
付図3-3-1/JT-G729 Example of estimation of the background noise characteristics (ITU-T G.729)

2. 信号レベルが鋭いエッジ状に高レベルから低レベル（例えば、ポーズ区間）に急激に遷移した場合、遷移直後に最初のSIDフレームの信号エネルギーとスペクトル成分を計算してしまうため、結果的に「呼吸音」のような疑似背景雑音を再生することがある。この問題点は、主にリングバックトーンやビジートーンのように、鋭い立下りエッジを有する音声帯域トーン信号（音声帯域トーン信号の鋭い立下りエッジ部分）において顕著になる。付図3-3-2/JT-G729は、一例として、TTC標準JT-G729付属資料Bにおける鋭いエッジの立下り部分を持つ問題点、およびTTC標準JT-G729付録3を適用することによる解決結果を示している。



付図 3-3-2 / JT-G729 Example of sharp-edge offset issue
(ITU-T G.729)

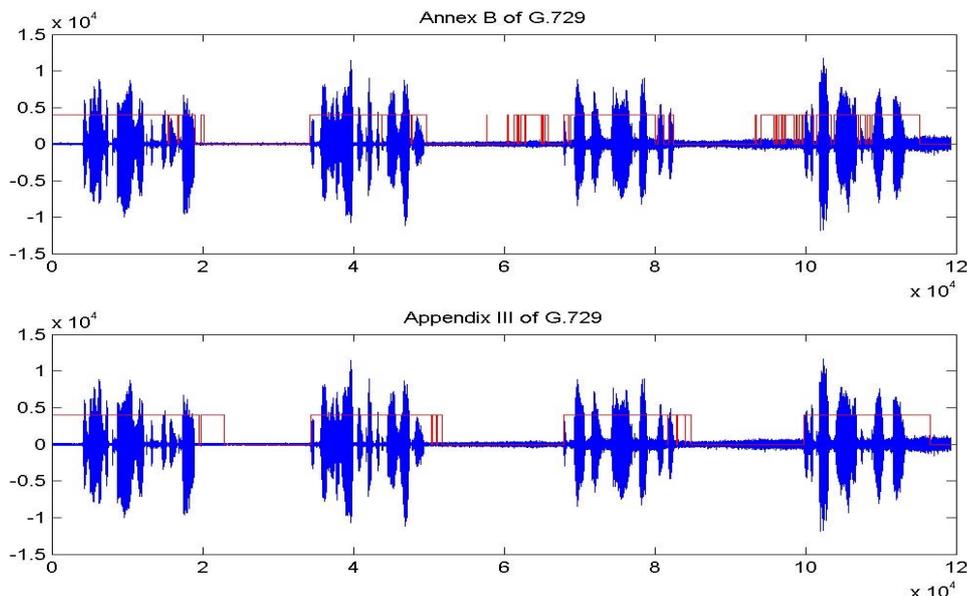
3. 長時間継続する高レベルのトーン状信号（一般には 1sec より長い信号）について、後の部分が「無音」と判定される場合がある。この場合、復号器では、長時間継続するトーン状信号の後ろの部分は、高レベルの雑音信号として誤って再生されてしまう恐れがある。付図 3-3-3 / JT-G729 は、一例として、TTC 標準 JT-G 7 2 9 付属資料 B における長時間継続する高レベルのトーン状信号の問題点の例、および TTC 標準 JT-G 7 2 9 付録 3 を適用することによる解決結果を示している。



付図 3-3-3 / JT-G729 Example of very long and high-level tonal signals issue
(ITU-T G.729)

4. ある固有の特徴量をもつ背景雑音については、8 kbit/s フルレート符号化と、低レートの無音圧縮符号化とが頻繁に切り替わる可能性がある。付図 3-3-4 / JT-G729 は、一例として、TTC 標準 JT-G 7 2 9 付

属資料Bにおける8 kbit/sフルレート符号化と、低レートの無音圧縮符号化とが頻繁に切り替わる問題点、およびTTC標準JT-G729付録3を適用することによる解決結果を示している。



付図3-3-4/JT-G729 Example of frequent changes between full-rate encoding at 8 kbit/s and low-rate silence encoding (ITU-T G.729)

5. 音声パケットネットワークへの適用において、伝送信号のビットレートは、ペイロードだけではなく、パケットのアドレスとヘッダ情報の影響も受ける。TTC標準JT-G729付属資料Bをこのようなネットワークへ適用する場合、SIDフレームのペイロードサイズ自体は小さいものの、SIDが頻繁に更新されることにより、ビットレートのかなりの割合を占有してしまう恐れがある。

付表3-3-1/JT-G729 に、一例として、TTC標準JT-G729付属資料B、及び同付録3適用時における有音フレーム、SIDフレーム、及び非伝送（NT）フレームが出現する割合をそれぞれ示す。付録3を適用することにより、生成されるSIDフレーム数が減ることが判る。

付表3-3-1/JT-G729 Example of the reduction in activeframes, SID frames and NT frames (ITU-T G.729)

Noise Type	VAD	Speech Frames	SID Frames	NT Frames	SID / Inactive
15 dB Street Noise	G.729B	4805 (46.2%)	799 (7.68%)	4796 (46.12%)	14.3%
	App III	6079 (58.45%)	275 (2.64%)	4046 (38.90%)	6.4%
15 dB Car Noise	G.729B	4986 (47.94%)	846 (8.13%)	4568 (43.92%)	15.6%
	App III	5806 (55.83%)	292 (2.81%)	4302 (41.37%)	6.4%
15 dB Babble Noise	G.729B	4991 (47.99%)	1287 (12.38%)	4122 (39.63%)	23.8%
	App III	5896 (56.69%)	561 (5.39%)	3943 (37.91%)	12.5%

3.4 電子ファイル

二つの電子ファイルが日本ITU協会から入手可能である。これは、Cソースコードvad.cおよびdtx.cが修正されたものである。正しくコンパイルするために、VAD_VOIP_APP_IIIフラグをプロジェクトファイルあるいはメイクファイルで定義する必要がある。

付録

(標準 J T - G 7 2 9 に対する)

(本付録は参考資料であり、仕様ではない。)

用語対照表

英 語	T T C 標準用語
active	有音
active frame	有音フレーム
active voice frame	有音フレーム
adaptive codebook	適応コードブック
adaptive codebook contribution	適応コードブック寄与分
analysis-by-synthesis (A-b-S)	合成による分析 (A - b - S)
attenuation factor	減衰係数
auto-regressive (AR)	自己回帰 (AR)
background acoustic noise	音響的な背景雑音
background noise	背景雑音
backward	バックワード
backward/forward decision	バックワード/フォワード判別
bad frame	不正フレーム
bit error	ビット誤り
bit-exact	ビットイグザクトな
bitstream	ビット列
calligraphic characters	カリグラフィ文字
Chebyshev polynomial	チェビシェフ多項式
circular convolution	巡回畳込み
classification	分類
closed-loop pitch analysis	閉ループピッチ分析
code-excited linear-prediction (CELP)	符号励振線形予測 (C E L P)
codebook contribution	コードブック寄与分
coder	符号器
codeword	符号語
comfort noise	擬似背景雑音
comfort noise excitation generator	擬似背景雑音発生器
comfort noise generator (CNG)	擬似背景雑音発生器 (C N G)
conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP)	共役構造一代数的符号励振線形予測 (C S - A C E L P)
correction factor	補正係数
cosine domain	コサイン領域
cut-off	カットオフ
decision region	判定領域
decoder	復号器
delay	遅延
delta lag range	デルタラグ範囲

depth-first tree search difference measure	縦形木探索 差分量
---	--------------

英 語	T T C標準用語
discontinued transmission	不連続伝送
discontinuous transmission (DTX)	不連続伝送 (D T X)
double-pulse tracks	ダブルパルストラック
DTX operation	D T X動作
encoder	符号器
erased frame	消失フレーム
error concealment	誤り補償
error free	エラーフリー
estimation	推定
Euclidean distance	ユークリッド距離
evaluation	評定
excitation	励振信号
excitation signal	励振信号
external synchronous reset	外部同期リセット
first parcor	1次偏自己相関
first reflection coefficient	1次反射係数
fixed codebook	固定コードブック
fixed codebook contribution	固定コードブック寄与分
focused search approach	絞り込み探索アプローチ
forward	フォワード
fractional delay	分数遅延
fractional lag	分数ラグ
fractional pitch	分数ピッチ
fractional pitch delay	分数ピッチ遅延
frame erasure concealment	フレーム消失補償
full version	標準 J T - G 7 2 9 本体
full band energy	全帯域エネルギー
full recommendation	標準 J T - G 7 2 9 本体
gain	利得
global stationarity indicator	グローバル定常性インジケータ
good active frame	正常な有音フレーム
hangover	ハングオーバー
harmonic components	調波構造
harmonic postfiltering	長期ポストフィルタ
high level description	概要記述
high stationarity	高定常性
high stationarity backward frames	高定常バックワードフレーム
high stationarity indicator	高定常性インジケータ
high stationarity state	高定常状態
high-pass filter	高域通過フィルタ
inactive	無音

英 語	T T C 標準用語
inactive frame	無音フレーム
industry	製品化
initial voice activity decision	初期有音判定
interleaved single-pulse permutation (ISPP)	インタリーブ単一パルス並び換え (I S P P)
interpolation	補間
interpolation factor	補間係数
interpolation factor	補間係数
iteration	繰り返し
iterative depth-first tree search	反復的デプス・ファースト木探索
lag	ラグ
less stationary signal	定常性の低い信号
Levinson-Durbin algorithm	レビンソン・ダービンアルゴリズム
line spectral frequency (LSF)	線スペクトル周波数 (L S F)
line spectral pair (LSP)	線スペクトル対 (L S P)
linear interpolation	線形補間
linear PCM	均一 P C M
linear prediction (LP) coefficients	線形予測 (L P) 係数
Log Area Ratio (LAR) coefficients	対数断面積比 (L A R) 係数
long term average	長時間平均
look-ahead	先読み
loop index	ループインデックス
low band energy	低域エネルギー
low stationarity	低定常性
lower triangular Toeplitz convolution matrix	下三角テプリッツ畳込み行列
LP analysis	L P 分析
LP filter	L P フィルタ
LP mode	L P モード
LP parameters	L P パラメータ
LSP coefficients	L S P 係数
LTP lag	L T P ラグ
mean-squared weighted error	重み付き平均自乗誤差
moving average (MA) prediction	移動平均 (M A) 予測
multi-boundary decision region	多境界判定領域
music detection	音楽検出
muting factor	減衰係数
nested loop	入れ子状ループ
nominal bit rate	標準ビットレート
non speech	無音
non-active voice frame	無音フレーム
non-active voice segment	無音区間
non-stationary	非定常な

英 語	T T C 標準用語
non-uniform quantizer	不均一量子化器
open-loop pitch analysis	開ループピッチ分析
original speech	入力音声
oversampled	オーバサンプル
partial normalized	部分正規化残差エネルギー
partial vector	部分ベクトル
perceptual weighting filter	聴覚重み付けフィルタ
phase dispersion	位相拡散
piecewise linear decision boundary	部分的な線形判定境界
pitch multiples	倍数ピッチ
pole/zero filter	極零フィルタ
post-processing, pre-processing	後処理、前処理
postfilter	ポストフィルタ
pre-filter	プリフィルタ
predictive two-stage vector quantization	予測 2 段ベクトル量子化
procedure	手順
pulse pair	パルス対
r.m.s. value	実効値
reconstructed speech	再生音声
reduced complexity version	低演算量版
residual	残差
robustness	耐性
running average	移動平均
running mean	移動平均
sampling rate	標本化周波数
scheme	手法
search	探索
second stationary indicator	第 2 定常性インジケータ
second step backward LP filter	第 2 段バックワード L P フィルタ
semi-random	擬似ランダム
separate bit L0	識別ビット L 0
short-term perceptual weighting filter	短期聴覚重み付けフィルタ
side information	補助情報
sign	極性
signal-selected pulse amplitude approach	信号選択パルス振幅アプローチ
silence compression	無音圧縮
silence compression scheme	無音圧縮手法
Silence Insertion Descriptor (SID)	無音挿入記述子 (S I D)
sinc function	標本化関数
sparse excitation	まばらな励振
Spectral envelope	スペクトル包絡

英 語	T T C 標準用語
spectral measure	スペクトル分布
spectrum transition	スペクトル推移
speech	有音
stability	安定性
stability check	安定性チェック
standard deviation	標準偏差
stationarity counter	定常性カウンタ
switched MA predictor	切換え型MA予測器
synthesis filter	合成フィルタ
target signal	ターゲット信号
tilt compensation	傾き補償
track	トラック
track table	トラックテーブル
transition filter	遷移フィルタ
transmission error	伝送誤り
tree	木
triple-pulse tracks	トリプルパルストラック
two signed pulses	2本の極性付きパルス
untransmitted frame	非伝送フレーム
valid frame	正常フレーム
vector quantization (VQ)	ベクトル量子化 (VQ)
voice activity decision	有音判定
voice activity detection (VAD)	音声検出 (VAD)
voice segment	音声区間
voicing classifier	有声/無声識別器
Voicing Strength Indicator	有声強度のインジケータ
weighted mean-squared error	重み付き平均自乗誤差
weighted MSE	重み付き平均自乗誤差
weighted speech signal	重み付き音声信号
weighted synthesis filter	重み付き合成フィルタ
weights	重み係数
zero-crossing rate	零交差数

付録

(標準 J T - G 7 2 9 に対する)

用語解説

(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 40 次元の零ベクトルに 4 本のパルスを置いている。

合成による分析 (A - b - S) 法 (analysis-by-synthesis)

信号生成モデルを用いて合成した音声と実際の信号との誤差を最小とするように、パラメータを変化させて求める分析方法。

バックワード/フォワード LP 構造 (backward/forward LP structure)

バックワード LP フィルタとフォワード LP フィルタを切り替えて使用する構成。本標準の付属資料 E では、従来のフォワード LP フィルタと分析次数の大きなバックワード LP フィルタを切り替えて使用することにより、背景雑音が存在する音声信号や音楽信号にも対応している。

符号励振線形予測 (C E L P) (code excited linear prediction)

音声信号の短期および長期相関を取り除いた残差信号をあらかじめ作成されたコードブックを用いてベクトル量子化し、このコードブックのインデックスと補助情報を用いて音声を符号化する方式。コードブックの探索は、合成による分析 (A - b - S) 法を用いる。

擬似背景雑音発生器 (C N G) (comfort noise generator (CNG))

入力音声が無音の場合、背景雑音についての若干の情報から、聴感上、違和感のない背景雑音を擬似的に生成する機構。本標準では、復号器の音声生成アルゴリズムを流用することにより、符号化側から伝送された情報 (S I D) を使用して擬似背景雑音を生成している。

共役構造一代数的符号励振線形予測 (C S - A C E L P)

(conjugate-structure algebraic-code-excited linear-prediction)

C E L P 方式を基本にした符号化方式で、主に 2 系統の共役構造からなる予測型利得ベクトル量子化、単位パルス列による代数的雑音コードブックの導入により、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。さらに、本標準では、L S P 係数のフレーム間予測を用いて符号化遅延を抑えている。

縦形木探索 (depth-first tree search)

深さ優先で木構造を探索する探索法。

不連続伝送 (DTX) (discontinuous transmission (DTX))

入力音声が無音の区間において、背景雑音についての若干の情報を断続的に伝送する方式。本標準では、無音の始まりや背景雑音が大きく変化した場合にのみ背景雑音に関する情報 (SID) を伝送している。

外部同期リセット (external synchronous reset)

符号器の状態変数と復号器の状態変数とが、同じフレーム時刻において各々の初期状態に設定されること。外部同期リセットを用いることにより、CME (回線多重化装置) 等コーデック外部のVAD/DTX/CNGアルゴリズムが用いられる場合に起こり得る符号器と復号器との同期外れによる音質の劣化を低減することができる。本標準の付録1では、本標準の本体、付属資料A、あるいは付属資料CをCMEのコーデックとして用いた場合に、有音区間の最初のフレームにおいてこの外部同期リセットを適用することにより、可能な限り最良の音質が得られることが実験によって確認されたことが述べられている。

分数ピッチ遅延 (fractional pitch delay)

適応コードブックパラメータ (またはピッチパラメータ) の1つであり、有声音の基本周期を示す。本標準では分析精度を上げるため、分数値を用いている。

グローバル定常性インジケータ (global stationarity indicator)

入力信号の全体的な定常性を特徴づける指標。本標準の付属資料EではバックワードLPフィルタを用いるかフォワードLPフィルタを用いるかの判定をするために、4つの基準の1つに用いている。

高定常性インジケータ (high stationarity indicator)

入力信号の定常性の高さを示す指標。一般に音楽信号の定常性は高く、音声信号の定常性は低い。本標準の付属資料Eでは、バックワードモード時に聴覚重み付けフィルタの重み付け係数を決定するために使用される。また、フレーム消失時に適応および固定コードブック利得の減衰係数を適応的に変えるために使用される。

レビンソン・ダービンアルゴリズム (Levinson-Durbin algorithm)

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にした時に得られる係数。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。

線形予測分析 (linear prediction analysis)

現時刻におけるサンプル値を、過去のサンプル値の線形結合として予測する線形予測モデルを用いて、現サンプルの観測値と予測値の誤差が最小になるように線形式の係数を求め、それを分析値とする音声信号の分析方法。

対数断面積比係数 (log area ratio coefficients)

レビンソン・ダービン再帰法の副産物として得られる反射係数によりスペクトル形状の特徴付けを行うことができる。スペクトル形状の特徴付けを効率的にするために、反射係数に非線形変換を施すことによって得られる係数。

移動平均 (MA) 予測 (moving average (MA) prediction)

現在および過去の入力信号系列のみを用い、それぞれに重みをかけて加え合わせることによって予測値を得る予測方式。

音楽検出 (music detection)

音楽信号が継続している区間かどうかを検出し判定する機能。本機能により、本標準の付属資料E (11.8kbit/s) に本標準の付属資料BのD T X動作を適用した場合において、音楽信号が継続している区間では確実に本標準の付属資料Eの 11.8kbit/s で符号化を行い、音楽品質を保証することができる。

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聴感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。

無音挿入記述子 (S I D) (Silence Insertion Descriptor (SID))

入力音声が無音の場合、擬似背景雑音を生成するために符号器から復号器に対して伝送される情報。本標準では無音の始まりや、背景雑音が大きく変化した場合にスペクトル形状とエネルギー (雑音の電力) に関する情報を伝送している。

無音圧縮 (silence compression)

入力音声が無音の区間において、符号化情報の伝送を行わないことにより、伝送に必要な帯域幅の削減を行う手法。本標準では無音の始まりや、背景雑音が大きく変化した場合にも若干の情報 (SID) を伝送し、復号側で擬似背景雑音を生成し、聴感上の品質を保っている。

音声検出 (VAD) (voice activity detection (VAD))

無音圧縮を行うために、入力音声を有音と無音に分類するための判定。本標準では、入力音声を 10ms 毎のフレームに分割し、フレーム単位で、有音か無音かの判定をしている。有音と無音の判定には、線スペクトル周波数、全帯域エネルギー、低域エネルギー、零交差数の 4 つのパラメータを用いている。