

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6386376号
(P6386376)

(45) 発行日 平成30年9月5日(2018.9.5)

(24) 登録日 平成30年8月17日(2018.8.17)

(51) Int.Cl.		F I			
G 1 0 L	19/005	(2013.01)	G 1 0 L	19/005	
G 1 0 L	19/00	(2013.01)	G 1 0 L	19/00	3 3 0 C
H 0 4 L	1/00	(2006.01)	H 0 4 L	1/00	B
			H 0 4 L	1/00	E

請求項の数 12 (全 38 頁)

(21) 出願番号	特願2014-505075 (P2014-505075)	(73) 特許権者	503447036
(86) (22) 出願日	平成24年4月11日 (2012.4.11)		サムスン エレクトロニクス カンパニー リミテッド
(65) 公表番号	特表2014-512575 (P2014-512575A)		大韓民国・16677・キョンギード・ス ウォンシ・ヨントンク・サムスンロー ・129
(43) 公表日	平成26年5月22日 (2014.5.22)	(74) 代理人	100107766
(86) 国際出願番号	PCT/KR2012/002738		弁理士 伊東 忠重
(87) 国際公開番号	W02012/141486	(74) 代理人	100070150
(87) 国際公開日	平成24年10月18日 (2012.10.18)		弁理士 伊東 忠彦
審査請求日	平成25年11月18日 (2013.11.18)	(74) 代理人	100091214
審査番号	不服2016-18608 (P2016-18608/J1)		弁理士 大貫 進介
審査請求日	平成28年12月9日 (2016.12.9)		
(31) 優先権主張番号	10-2012-0037625		
(32) 優先日	平成24年4月11日 (2012.4.11)		
(33) 優先権主張国	韓国 (KR)		

最終頁に続く

(54) 【発明の名称】 マルチレート・スピーチ/オーディオ・コーデックのためのフレーム損失隠匿

(57) 【特許請求の範囲】

【請求項1】

コーデックの動作モードを設定し、前記動作モードがハイフレームエラーの比率の状態を考慮したモードである場合、複数のコーディングモードから選択されたコーディングモードによって、現在フレームの部分的なりだんだんデータ(partial redundant data)を少なくとも1つの隣接フレームに付加するプロセッサを含み、

前記ハイフレームエラーの比率の状態は、フレームエラーの比率が基準値より高い場合に対応し、

前記部分的なりだんだんデータのサイズは、信号特性に基づいて決定され、

前記ハイフレームエラーの比率の状態を考慮したモードでは、全体パケットのサイズを変更することなく、前記部分的なりだんだんデータが付加されるように減少したビットレートで符号化を実行する、

ことを特徴とする端末機。

【請求項2】

前記プロセッサは、

入力オーディオデータの複数のフレームそれぞれのために、複数の動作モードから前記動作モードを設定する、

ことを特徴とする請求項1に記載の端末機。

【請求項3】

前記ハイフレームエラーの比率の状態を考慮した動作モードは、

3 G P P 標準の E V S コーデックのための動作モードであり、
 前記コーデックは、E V S コーデックであり、
 前記 E V S コーデックは、少なくとも 1 つの隣接フレームからエンコーディングされたオーディオを、結合された E V S ソースビットとして、現在フレームのためのパケットで、現在フレームのエンコーディング結果に追加し、

前記隣接フレームは、一つ以上の以前フレーム及び/または一つ以上の以後フレームそれぞれのエンコーディングされたオーディオを含み、

前記結合された E V S ソースビットは、現在パケットで、R T P ペイロード部分と区分されて表現され、

前記 E V S コーデックは、エンコーディングされたオーディオである少なくとも 1 つの隣接フレームそれぞれから、個別的にオーディオをエンコーディングし、現在パケットから分離されたパケットに、少なくとも 1 つの隣接フレームそれぞれからエンコーディングされたオーディオを追加させる、

10

ことを特徴とする請求項 2 に記載の端末機。

【請求項 4】

前記コーデックは、さらに、

現在フレームに係わる設定された動作モードに、前記ハイフレームエラーの比率の状態を考慮した動作モードフラグを追加する、

ことを特徴とする請求項 3 に記載の端末機。

【請求項 5】

20

前記ハイフレームエラーの比率の状態を考慮した動作モードのフラグは、
 現在パケットの R T P ペイロード部分で、1 つのビットとして、現在パケットに表現される、

ことを特徴とする請求項 4 に記載の端末機。

【請求項 6】

前記コーデックは、さらに、

現在フレームについて選択された複数のコーディングモードを識別するコーディングモードのフラグを、現在フレームのためのパケットに追加する、

ことを特徴とする請求項 3 に記載の端末機。

【請求項 7】

30

前記コーディングモードのフラグは、

既設定の個数のビットで、現在パケットで表現される、

ことを特徴とする請求項 6 に記載の端末機。

【請求項 8】

前記コーデックは、

現在フレームに係わるコーディングモードのフラグを、他のフレームのパケットに、リダンダンシを用いて追加する、

ことを特徴とする請求項 7 に記載の端末機。

【請求項 9】

40

前記プロセッサは、

端末機外部で決定された伝送品質のうち少なくとも一つ以上、伝送の際に現在フレームがフレーム損失にさらに敏感であるとの決定、および、前記現在フレームの決定に基づいて、複数の動作モードの他の動作モードと比較して、異なるリダンダンシ、増加されたリダンダンシ、及び/または多様なリダンダンシで、動作モードを前記ハイフレームエラーの比率の状態を考慮した動作モードに設定する、ように構成されている、

ことを特徴とする請求項 1 に記載の端末機。

【請求項 10】

前記プロセッサは、

複数の使用可能なコーディング・タイプから、現在フレーム及び隣接フレームのうち少なくとも一つの決定されたコーディング・タイプに基づいて、または、複数の使用可能な

50

フレーム分類から、現在フレーム及び隣接フレームの決定されたフレーム分類のうち少なくとも一つの決定されたフレーム分類に基づいて、一つまたはそれ以上のコーディングモードのうち一つのコーディングモードに動作モードを設定する、ように構成されている、ことを特徴とする請求項 1 に記載の端末機。

【請求項 1 1】

前記複数の使用可能なコーディング・タイプは、

アンボイスされたスピーチフレームのためのアンボイスされたワイドバンド・タイプ、ボイスされたスピーチフレームのためのボイスされたワイドバンド・タイプ、ノンステーションナリ・スピーチフレームのための一般ワイドバンド・タイプ、及び向上されたフレーム除去パフォーマンスのために使用されたトランジション・ワイドバンド・タイプを含む

10

ことを特徴とする請求項 1 0 に記載の端末機。

【請求項 1 2】

前記複数の使用可能なフレーム分類は、

アンボイス、サイレンス、ノイズ、ボイスされたオフセットのためのアンボイスされたフレーム分類；アンボイスされたコンポーネントからボイスされたコンポーネントへのトランジションのためのアンボイスされたトランジション分類；ボイスされたコンポーネントからアンボイスされたコンポーネントへのトランジションのためのボイスされたトランジション分類；ボイスされたフレーム及びすでにボイスされたり、あるいはオンセット・フレームに分類された以前フレームのためのボイスされた分類；及びデコーディング器によってボイス隠匿に従うように十分に良好に設計されたボイスされたオンセットのためのオンセット分類；を含む、

20

ことを特徴とする請求項 1 0 に記載の端末機。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、オーディオ・エンコーディング/デコーディングのための技術、技法と係わる一つ以上の実施形態に係り、さらに具体的には、マルチレートスピーチと、オーディオ・コーデックとを利用して、向上されたフレームエラー損失技法で、オーディオをエンコーディング並びにデコーディングする方法及び装置に関する。

30

【背景技術】

【0 0 0 2】

エンコーディングされたスピーチまたはオーディオのフレームが伝送される間、時折損失されると予想される環境で遂行されるコーディングされたスピーチとオーディオのための伝送システムまたはデコーディング・システムは、フレーム損失を何パーセントかに制限するために考案された。

【0 0 0 3】

かようなフレーム損失を制限するため、またはフレーム損失を補償するために、フレーム損失隠匿 (F E C : frame erasure concealment) アルゴリズムは、デコーディング・システムで、スピーチやオーディオをエンコーディングしたりデコーディングするとき

40

【0 0 0 4】

かようなフレーム損失隠匿アルゴリズムは、最近、特定標準 (standard) や規格 (specification) によって作動するセルラ通信ネットワークまたは環境で活用された。ここで、標準または規格は、連結及び通信のために使用されなければならない通信プロトコル及び/またはパラメータを定義することができる。例えば、前記標準または規格は、通信プロトコル及びモバイル通信のための G S M (global system for mobile communications)、G S M / enhanced data rates for G S M evolution、A M P S (American

50

mobile phone system)、WCDMA(登録商標(wideband code division multiple access))、3G(generation)UMTS(universal mobile telecommunications system)、IMT2000(international mobile telecommunications 2000)などを含む。

【0005】

ここで、スピーチ・コーディングは、以前に可変レート(variable rate)または固定レート(fixed rate)のうちいずれか一つで遂行された。可変レートでエンコーディングするとき、ソースは、スピーチを異なる比率に分類するアルゴリズムを使用し、分類されたスピーチを既設定のビットレートそれぞれに対応してエンコーディングすることができる。代案として、探知されたボイススピーチ・オーディオが固定されたビットレートによってコーディングされなければならない場合、スピーチ・コーディングは、固定されたビットレートを利用して遂行された。

10

【0006】

例えば、かような固定レートでコーディングするコーデックは、AMR(adaptive multi-rate)及びAMR-WB(adaptive multi-rate wideband)のようなGSM/EDGEとWCDMAとの通信ネットワークのために、3GPP(3rd generation partnership project)によって開発されたマルチレート・スピーチ・コーデックを含んでもよい。かようなコーデックは、探知されたボイス情報によってスピーチをコーディングし、さらに無線インターフェースのネットワーク容量(network capacity)及び無線チャンネル条件(radio channel condition)のようなファクタに基づいて、スピーチをコーディングすることができる。ここで、マルチレートは、コーデックの動作モードに依存して使用される固定レートを意味する。

20

【0007】

例えば、AMRコーデックは、スピーチのために、4.7 kbit/sから12.2 kbit/sまで8個の使用可能なビットレートを含む。一方、AMR-WBは、スピーチのために、6.6 kbit/sから23.85 kbit/sまで9個の使用可能なビットレートを含む。AMRコーデック及びAMR-WBコーデックの規格は、それぞれ3GPP無線システムの3世代に対する技術規格である3GPP TS 26.090と3GPP TS 26.190で使用可能である。そして、AMR-WBコーデックのスピーチ感知部分は3GPP無線システムの3世代に係わる技術規格である3GPP TS 26.194技術規格で求めることができる。

30

【0008】

例えば、かようなセルラ環境で、損失(losses)は、セルラ無線リンク内での干渉、またはIP(internet protocol)ネットワーク内でのルータオーバーフローによって発生する。LTE(long term evolution)と呼ばれるEPS(enhanced packet service)のための主要無線インターフェースで、EPSと知られた3GPP無線システムの4世代技術は、現在開発中にある。例えば、図1は、スピーチメディア・コンポーネント12を有したEPS 10を図示している。ここで、ボイスデータは、AMR-WB(wideband)とAMR-NB(narrowband)によってコーディングされる。

【0009】

例えば、3GPPリリース8,9で、EPS 10は、UMTSとLTEとのボイス・コーデックによる。3GPPリリース8,9で、LTEスピーチ・コーデックを含むUMTSは、EPSによって、IMS(IP multimedia core network subsystem)のためのマルチメディア・テレフォニ・サービスと呼ばれる。UMTSは、4世代3GPP無線システムのために最初にリリースされた。IMSは、IPマルチメディア・サービスのための構造的なフレームワークである。

40

【0010】

たとえLTEが潜在的な伝送干渉の観点で開発され、セルラ・ネットワークまたは無線ネットワークに失敗したとしても、3GPPセルラ・ネットワークで伝送されるスピーチフレームは、伝送される間、一部フレーム及び/またはパケットが除去(erasure)され

50

やすい。除去は、デコーダ側面で、パケットの情報が損失されたり、あるいは使用されたりするという事を仮定するための分類 (classification) である。例えば、E P S ネットワークの場合、フレーム除去が予想される。除去されたフレームを、処理 (address) するために、デコーダは、損失されたフレームに対応する衝撃を緩和するためのフレーム損失隠匿 (F E C) アルゴリズムを遂行することができる。

【 0 0 1 1 】

いくつかの F E C アルゴリズムは、ただ損失されたフレームのように除去されたフレームの隠匿をデコーダで処理するために使用されるのみである。例えば、デコーダは、フレーム除去が発生したということを知り認識することができ、除去されたフレームの直前または直後にデコーダに達する良好な状態のフレームから除去されたフレームのコン

10

【 0 0 1 2 】

いくつかの 3 G P P セルラ・ネットワークのフレーム除去が発生された受信端 (receiving station) を識別して通知することができる能力を有している。従って、スピーチ・デコーダは、受信されたスピーチフレームが良好な状態のフレームであるか否か、または除去されたフレームと見なされるか否かということが分かる。かようなスピーチ及びオーディオの本質的特性のために、適切なフレーム損失の緩和または隠匿の技法が遂行されるのであるならば、低比率のフレーム損失は容認されるであろう。いくつかの F E C アルゴリズムは、フレーム損失があまり目立たないように損失されたパケット、サイレンス、いくつかのタイプのフェーディングアウト/フェーディングイン、またはいくつかのタイプ

20

の補間 (interpolation) をノイズに代替する。

【 0 0 1 3 】

代替的な F E C アルゴリズムのアプローチ方式は、リダンダント方式 (redundant fashion) で規格情報を伝送するエンコーダを含む。例えば、参照によって含まれた I T U - T G . 7 1 8 標準は、向上レイヤ (enhancement layer) で、コアエンコーダ出力と係わるリダンダント情報を伝送することを推薦する。向上レイヤは、コアレイヤと異なるパケットを伝送することができる。

【 発明の概要 】

【 課題を解決するための手段 】

【 0 0 1 4 】

本発明の一実施形態による端末機は、コーデックを利用して入力オーディオデータをコーディングするために、複数の動作モードから 1 つの動作モードを設定するコーディング・モード設定部と、前記動作モードがハイフレーム除去レートモード (high F E R : frame erasure rate) であるとき、複数のフレーム損失隠匿 (F E C : frame erasure concealment) モードのうちいずれか一つによって、入力オーディオデータの現在フレームをコーディングすることにより、前記入力オーディオデータをコーディングするコーデックと、を含み、前記動作モードを high F E R 動作モードに設定するやいなや、前記コーディング・モード設定部は、high F E R 動作モードに係わる既設定の F E C モードから、いずれか 1 つの F E C モードを選択し、入力オーディオデータをコーディングするとき、リダンダンシ (redundancy) を導入したり、あるいは設定された 1 つの F E C モード

30

40

に基づいて、入力オーディオデータをコーディングするようにコーデックを制御することができる。

【 0 0 1 5 】

前記端末機の前記コーディング・モード設定部は、前記入力オーディオデータを構成する複数のフレームそれぞれのために、複数の F E C モードから 1 つの F E C モードを選択することができる。

【 0 0 1 6 】

前記 high F E R 動作モードは、3 G P P 標準の E V S (enhanced voice services) コーデックのための動作モードであり、前記コーデックは、E V S コーデックであり、

50

前記EVSコーデックが現在フレームのオーディオをエンコーディングするとき、前記EVSコーデックは、少なくとも1つの隣接フレームでエンコーディングされたオーディオを、結合されたEVSソースビットとして、現在フレームのためのパケットで、現在フレームのエンコーディング結果に追加し、前記隣接フレームは、一つ以上の以前フレーム及び/または一つ以上の以後フレームそれぞれのエンコーディングされたオーディオを含み、前記結合されたEVSソースビットは、現在パケットでRTPペイロード部分と区分されて表現され、前記EVSコーデックは、エンコーディングされたオーディオである少なくとも1つの隣接フレームそれぞれから、個別的にオーディオをエンコーディングし、現在パケットから分離されたパケットに、少なくとも1つの隣接フレームそれぞれからエンコーディングされたオーディオを追加させることができる。

10

【0017】

前記複数のFECモードのうち一つ以上は、選択的に異なる固定ビットレート及び/または異なるパケットサイズによって、現在フレームと隣接フレームとをコーディングするようにコーデックを制御することができる。

【0018】

前記複数のFECモードのうち一つ以上は、同一の固定ビットレートによって、現在フレームと隣接フレームとをコーディングするようにコーデックを制御することができる。

【0019】

前記複数のFECモードのうち一つ以上は、同一のパケットサイズによって、現在フレームと隣接フレームとをエンコーディングするように制御することができる。

20

【0020】

前記複数のFECモードのうち一つ以上は、現在フレームをサブフレームに分割し、同一の固定ビットレートより低いビットレートでコーディングされたサブフレームそれぞれのコードブック・ビットの数を計算し、サブフレームのビットに係わるコードワードを定義するために使用されるそれぞれのコードブック・ビットの数と同一の固定ビットレートを利用して、サブフレームをエンコーディングするように、コーデックを制御することができる。

【0021】

前記EVSコーデックは、現在フレームのビットを、少なくとも最初のサブフレームと2番目のサブフレームとを含むサブフレームに分類したところに基づいて、現在フレームのビットのための差等的なリダンダンシ(unequal redundancy)を提供し、最初のサブフレームに分類された現在フレームのエンコーディングビットを、隣接パケットでは、2番目のサブフレームに分類して加えるように、それぞれの一つまたはそれ以上の隣接パケットに、異なる方式で追加することができる。

30

【0022】

前記EVSコーデックは、現在フレームのビットを、少なくとも最初のサブフレームと2番目のサブフレームとを含むサブフレームに分類したところに基づいて、線形予測パラメータのための差等的なリダンダンシを提供し、最初のサブフレームに分類された現在フレームの線形予測パラメータのエンコーディングビットを、隣接パケットでは、2番目のサブフレームに分類して加えるように、それぞれの一つまたはそれ以上の隣接パケットに、異なる方式で追加することができる。

40

【0023】

前記現在フレームのためのパケットは、以前フレーム及び/または以後フレームからリダンダンシ情報に含まれたFECビットと直接に連結された区分された部分を含まなくともよい。

【0024】

前記コーデックは、現在フレームに係わる設定された動作モードを、high FER動作モードとして識別するために、現在フレームのためのパケットに、high FER動作モードフラグを追加することができる。

【0025】

50

前記high F E R動作モードフラグは、現在パケットのR T Pペイロード部分で、1つのビットとして、現在パケットに表現される。

【0026】

前記コーデックは、現在フレームについて選択された複数のF E Cモードを識別するF E Cモードフラグを、現在フレームのためのパケットに追加することができる。前記F E Cモードフラグは、既設定の個数のビットで、現在パケットで表現される。代替的な一実施形態で、既設定の個数は2個でもある。前記コーデックは、現在フレームに係わるF E Cモードフラグを、異なるフレームのパケットで、リダンダンシでもってエンコーディングすることができる。

【0027】

前記high F E R動作モードは、3 G P P標準のE V S (enhanced voice services) コーデックのための動作モードであり、前記コーデックは、E V Sコーデックであり、前記E V Sコーデックは、high F E R動作モードのフラグを探知するやいなや、high F E R動作モードとして、現在フレームに係わる動作モードを識別するために、少なくとも1つの現在パケットで、high F E R動作モードフラグをデコーディングし、現在パケットから現在フレームのために選択された複数のF E Cモードを識別する現在フレームのためのF E Cモードフラグをデコーディングし、前記入力オーディオデータのコーディングは、選択されたF E Cモードによって、入力オーディオデータをデコーディングし、前記E V Sコーデックが入力オーディオデータをデコーディングするとき、現在パケットで少なくとも1つの隣接フレームからエンコーディングされたリダンダント・オーディオ (redundant audio) をパーズングし、一つ以上の以前フレーム及び/または一つ以上の以後フレームそれぞれのエンコーディングされたオーディオを現在フレームに含め、現在パケットでパーズングされたエンコーディングされたリダンダント・オーディオそれぞれに基づいて、一つ以上の以前フレーム及び/または一つ以上の以後フレームそれぞれで損失フレーム (lost frame) をデコーディングすることができる。

【0028】

前記E V Sコーデックは、入力オーディオデータ内部で、現在フレームのためのビットまたはパラメータに係わる差等的なりダンダンシに基づいて、現在フレームをデコーディングし、前記差等的なりダンダンシは、現在フレームのビットまたはパラメータを、第1カテゴリー及び第2カテゴリーに以前に分類したところに基づいて、第1カテゴリーに分類された現在フレームのビットまたはパラメータのエンコーディングビットを、隣接パケットでは、第2カテゴリーに分類してそれぞれのリダンダント情報に加えるように、それぞれの一つまたはそれ以上の隣接パケットに、異なる方式で追加するところに基づいて、前記現在フレームのコーディングは、現在フレームが損失されたとき、一つ以上の隣接パケットからデコーディングされた現在フレームのオーディオに基づいて、現在フレームをデコーディングすることを含んでもよい。

【0029】

前記high F E R動作モードは、3 G P P標準のE V Sコーデックのための動作モードであり、前記コーデックは、E V Sコーデックであり、前記E V Sコーデックは、high F E R動作モードとして、現在フレームに係わる動作モードを識別するために、少なくとも1つの現在パケットで、high F E R動作モードのフラグをデコーディングし、high F E R動作モードのフラグを探知するやいなや、現在パケットから現在フレームのために選択された複数のF E Cモードを識別する現在フレームのためのF E Cモードフラグをデコーディングし、前記入力オーディオデータのコーディングは、選択されたF E Cモードによって、入力オーディオデータをデコーディングし、前記E V Sコーデックは、入力オーディオデータ内部で、現在フレームのためのビットまたはパラメータに係わる差等的なりダンダンシに基づいて、現在フレームをデコーディングし、前記差等的なりダンダンシは、現在フレームのビットまたはパラメータを、第1カテゴリー及び第2カテゴリーに以前に分類したところに基づいて、第1カテゴリーに分類された現在フレームのビットまたはパラメータのエンコーディングビットを、隣接パケットでは、第2カテゴリーに分類し

10

20

30

40

50

てそれぞれのリダンダント情報に加えるように、それぞれの一つまたはそれ以上の隣接パケットに、異なる方式で追加し、前記現在フレームのコーディングは、現在フレームが損失されたとき、一つ以上の隣接パケットからデコーディングされた現在フレームのオーディオに基づいて、現在フレームをデコーディングすることができる。

【0030】

前記EVSコーデックは、現在フレームのビットを第1カテゴリ及び第2カテゴリに分類することにより、現在フレームのビットに係わる差等的なリダンダンシを提供し、第1カテゴリに分類された現在フレームのビットのエンコーディングビットを、隣接パケットでは、第2カテゴリに分類して加えるように、それぞれの一つまたはそれ以上の隣接パケットに、異なる方式で追加することができる。

10

【0031】

前記EVSコーデックは、現在フレームのビットを、少なくとも第1カテゴリ及び第2カテゴリに分類することにより、現在フレームの線形予測パラメータのための差等的なリダンダンシを提供し、第1カテゴリに分類された現在フレームのビットの線形予測パラメータのエンコーディングビットを、隣接パケットでは、第2カテゴリに分類して加えるように、それぞれの一つまたはそれ以上の隣接パケットに、異なる方式で追加することができる。

【0032】

前記EVSコーデックが現在フレームのオーディオをエンコーディングするとき、前記EVSコーデックは、少なくとも1つの隣接フレームでエンコーディングされたオーディオを、現在フレームのエンコーディング結果を含むエンコーディングされたソースビット部分と区別される現在フレームのためのパケットのFEC部分に追加し、前記隣接フレームは、一つ以上の以前フレーム及び/または一つ以上の以後フレームそれぞれのエンコーディングされたオーディオを含み、前記現在パケットのエンコーディングされたソースビット部分と、現在パケットのFEC部分は、現在パケットで、 RTPペイロード部分と区別されて表現され、前記EVSコーデックは、少なくとも1つの隣接フレームそれぞれに対して個別的にオーディオをエンコーディングし、少なくとも1つの隣接フレームそれぞれについてエンコーディングされたオーディオを、現在パケットから分離されたパケットに追加させることができる。

20

【0033】

前記コーデックは、少なくとも1つの隣接フレームのビットのエンコーディング結果を、現在パケットの分離されたFEC部分に追加することにより、少なくとも1つの隣接フレームのビットに係わるリダンダンシを提供することができる。前記分離されたパケット (separate packers) は、隣接しない。

30

【0034】

前記複数のFECモードのうち一つ以上は、選択的に異なる固定ビットレート及び/または異なるパケットサイズによって、現在フレームと隣接フレームとをコーディングするように、コーデックを制御することができる。

【0035】

前記複数のFECモードのうち一つ以上は、選択的に同一の固定ビットレートによって、現在フレームと隣接フレームとをコーディングするように、コーデックを制御することができる。

40

前記複数のFECモードのうち一つ以上は、同一のパケットサイズによって、現在フレームと隣接フレームとをコーディングするように、制御することができる。

【0036】

前記複数のFECモードのうち一つ以上は、現在フレームをサブフレームに分割し、同一の固定ビットレートより低いビットレートでコーディングされたサブフレームそれぞれのコードブック・ビットの数を計算し、サブフレームのビットに係わるコードワードを定義するために使用されるそれぞれのコードブック・ビットの数と同一の固定ビットレートを利用して、サブフレームをエンコーディングするようにコーデックを制御することがで

50

きる。

【0037】

前記EVSコーデックは、現在フレームのビットを、少なくとも最初のサブフレームと2番目のサブフレームとを含むサブフレームに分類したところに基づいて、現在フレームのビットのための差等的なリダンダンシを提供し、最初のサブフレームに分類された現在フレームのエンコーディングビットを、隣接パケットでは、2番目のサブフレームに分類して加えるように、それぞれの一つまたはそれ以上の隣接パケットに、異なる方式で追加することができる。

【0038】

前記EVSコーデックは、現在フレームのビットを、少なくとも最初のサブフレームと2番目のサブフレームとを含むサブフレームに分類したところに基づいて、線形予測パラメータのための差等的なリダンダンシを提供し、最初のサブフレームに分類された現在フレームの線形予測パラメータのエンコーディングビットを、隣接パケットでは、2番目のサブフレームに分類して加えるように、それぞれの一つまたはそれ以上の隣接パケットに、異なる方式で追加することができる。

10

【0039】

前記コーディング・モード設定部は、端末機外部の伝送品質のうち一つ以上、及び/または伝送過程でフレーム損失にさらに敏感であるか、あるいは入力オーディオデータの他のフレームよりさらに重要性が高い入力オーディオデータの現在フレームの決定に基づいて、端末機で活用可能なフィードバック情報の分析に基づいて、一般動作モードのための複数の動作モードのうち残っているモードを比較した他の(different)、増加した(increased)、かつ/または多様な(varied)リダンダンシで、動作モードをhigh FER動作モードに設定することができる。

20

【0040】

前記フィードバック情報は、物理的階層で伝送されたハイブリッド自動反復要請(HARQ: hybrid automatic repeat request)フィードバックであるファースト・フィードバック(FFB: Fast Feedback)情報; 物理的階層よりさらに高い階層で伝送されたネットワーク・シグナリングからフィードバックされたスロー・フィードバック(SFB: slow feedback: SFB)情報; 終端(far end)でコーデックからインバンド・シグナリングされたフィードバック(ISB: in-band feedback: ISB)情報; 及びリダンダント方式(redundant fashion)で伝送される特定クリティカル・フレーム(specific critical frame)のコーデックによる選択であるハイセンシティブフレーム(HSF: high sensitivity frame)情報のうち少なくとも一つを含んでもよい。

30

【0041】

前記端末機は、FFB情報、HARQフィードバック、SFB情報、ISB情報のうち少なくとも一つを受信し、端末外部からの伝送と係わる一つ以上の品質を決定するために、受信されたフィードバック情報を分析することができる。

【0042】

前記端末機は、パケットに受信されたフラグに基づいて、以前に遂行されるFFB情報、HARQフィードバック、SFB情報、ISB情報のうち少なくとも1つの分析結果を示す情報を受信し、前記フラグは、high FER動作モードによってエンコーディングされた現在パケットの現在フレーム、またはhigh FER動作モードでコーデックによって遂行されなければならない現在パケットのコーディングを示すことができる。

40

【0043】

前記コーディング・モード設定部は、複数の使用可能なコーディング・タイプで、現在フレーム及び/または隣接フレームの決定されたコーディング・タイプ、または複数の使用可能なフレーム分類で、現在フレーム及び/または隣接フレームの決定されたフレーム分類のうち一つに基づいて、複数のFERモードのうち一つに動作モードを設定することができる。

50

【 0 0 4 4 】

前記複数の使用可能なコーディング・タイプは、アンボイスされたスピーチフレーム (unvoiced speech frames) のためのアンボイスされたワイドバンド・タイプ (unvoiced wideband type)、ボイスされたスピーチフレーム (voiced speech frames) のためのボイスされたワイドバンド・タイプ (voiced wideband type)、ノンステーションナリ・スピーチフレーム (non-stationary speech frame) のための一般ワイドバンド・タイプ (generic wideband type)、及び向上されたフレーム除去パフォーマンス (enhanced frame erasure performance) のために使用されたトランジション・ワイドバンド・タイプ (transition wideband type) を含んでもよい。

【 0 0 4 5 】

前記複数の使用可能なフレーム分類は、アンボイス、サイレンス、ノイズ、ボイスされたオフセット (voiced offset) のためのアンボイスされたフレーム分類 (unvoiced frame classification)、アンボイスされたコンポーネントからボイスされたコンポーネントへのトランジションのためのアンボイスされたトランジション分類 (unvoiced transition classification)、ボイスされたコンポーネントからアンボイスされたコンポーネントへのトランジションのためのボイスされたトランジション分類 (voiced transition classification)、ボイスされたフレーム及びすでにボイスされたか、あるいはオンセツフレーム (onset frame) に分類された以前フレームのためのボイスされた分類 (voiced classification)、及びデコーディング器によってボイス隠匿 (voice concealment) に従うように十分に良好に設計されたボイスされたオンセットのためのオンセット分類を含んでもよい。

【 0 0 4 6 】

本発明の一実施形態によるコーディング方法は、コーデックを利用して入力オーディオデータをコーディングするために、複数の動作モードから1つの動作モードを設定する段階と、前記動作モードがハイフレーム除去レートモード (high FER: frame erasure rate) であるとき、複数のフレーム損失隠匿 (FEC: frame erasure concealment) モードのうちいずれか一つによって、入力オーディオデータの現在フレームをコーディングすることにより、前記入力オーディオデータをコーディングする段階と、を含み、前記動作モードをhigh FER動作モードに設定するやいなや、前記入力オーディオデータをコーディングする段階は、high FER動作モードに係わる既設定のFECモードから、いずれか1つのFECモードを選択し、入力オーディオデータをコーディングするとき、リダンダンシを導入したり、あるいは設定された1つのFECモードによってコーディングされた入力オーディオデータに分類されたリダンダンシ情報に基づいて、入力オーディオデータをコーディングすることができる。

【 発明の効果 】

【 0 0 4 7 】

本発明の一実施形態によれば、フレーム伝送過程で除去されたフレームに対して、効率的にフレーム損失隠匿を遂行したりまたは復元することができる。

【 図面の簡単な説明 】

【 0 0 4 8 】

【 図 1 】 本発明の一実施形態による、EVS (enhanced voice service) を含むEPS (evolved packet system) を図示した図面である。

【 図 2 A 】 本発明の一実施形態による、エンコーディング端末、一つ以上のネットワーク及びデコーディング端末を図示した図面である。

【 図 2 B 】 本発明の一実施形態による、EVSコーデックを含む端末を図示した図面である。

【 図 3 】 本発明の一実施形態による、代替パケットに提供される1つのフレームに係わるリダンダント・ビット (redundant bit) の例示を図示した図面である。

【 図 4 】 本発明の一実施形態による、2個の代替パケットに提供される1つのフレームに係わるリダンダント・ビットの例示を図示した図面である。

10

20

30

40

50

【図5】本発明の一実施形態による、フレームの packets 前後に位置した代替 packets に提供される 1 つのフレームに係わるリダンダント・ビットの例示を図示した図面である。

【図6】本発明の一実施形態による、ソースビットの異なる分類に基づいて、代替 packets でソースビットの差等的なリダンダンシを図示した図面である。

【図7】本発明の一実施形態による、差等的なリダンダンシを有する F E C 動作モードの一例を図示した図面である。

【図8】本発明の一実施形態による、同じ伝送ブロックサイズを有する high F E R 動作モードに係わる異なる F E C 動作モードを図示した図面である。

【図9】本発明の一実施形態による、C クラスビットの個数と同じ A クラスビットの個数に基づいて、差等的なリダンダンシ伝送のために使用可能な packets の 4 種サブタイプを図示した図面である。

10

【図10】本発明の一実施形態による、オンセット・フレームに、向上されたプロテクション (enhanced protection) を提供する多様な packets ・サブタイプを図示した図面である。

【図11】本発明の一実施形態による、high F E R 動作モードで、異なる F E C 動作モードを利用して、オーディオデータをコーディングする方法を図示した図面である。

【図12】本発明の一実施形態による、すべての F E C 動作モードについて、同じビットレートまたは packets サイズが維持されるか否かということに基づいた F E C フレームワークを図示した図面である。

【図13】本発明の一実施形態による、3 個の F E C 動作モードの例示を図示した図面である。

20

【図14】本発明の一実施形態による、high F E R 動作モードで、異なる F E C 動作モードを利用して、オーディオデータをデコーディングする方法を図示した図面である。

【発明を実施するための形態】

【0049】

以下、図示された図面によって、本発明の一実施形態について具体的に説明する。そして、同じ参照図面は、同じ構成要素を示す。本発明の一実施形態は、他の形態によって構成され、特定の構成要素に限定解釈されるものではなく、システムの多様な変更、修正、同一性の範囲まで包括しなければならない。そして、説明される装置及び/または方法は、従来技術に基づいて理解されもする。従って、本発明の一実施形態は、図面によって、

30

【0050】

本発明の一実施形態は、スピーチ・コーディング及びオーディオ・コーディングの技術領域と係わるものであり、エンコーディングされたスピーチまたはオーディオのフレームは、伝送過程で時折損失されもする。セルラ無線リンク (cellular radio link) での妨害 (interference)、または IP (internet protocol) ネットワークでのルータ・オーバーフロー (router overflow) のような理由で、スピーチフレームまたはオーディオフレームの損失が発生することがある。

【0051】

本発明の一実施形態は、3 G P P (3rd generation partnership project) 無線システム構造の 4 世代方式に採択される E V S (enhanced voice service) コーデックと係わるものであるが、本発明の一実施形態は、必ずしも E V S に制限されるものではない。

40

【0052】

3 G P P は、将来の無線携帯電話または無線システムのための新たなスピーチ・コーデック及びオーディオ・コーデックを標準化する過程である。E V S コーデックとして周知のこのコーデックは、E P S (enhanced packet services) として周知の 3 G P P の 4 世代ネットワークのためのエンコーディングされたビットレートの広い範囲で、スピーチ及びオーディオを効率的に圧縮するように設計された。E P S の特徴のうち一つは、L T E (long term evolution) として知られた E P S 無線インターフェース (air interf

50

ace) を介して、スピーチ及びオーディオの圧縮結果を含む全てのサービスのために、パケット基盤の伝送で使用されるものである。EVSコーデックは、パケット基盤環境で、効率的に動作するように設計される。

【0053】

EVSコーデックは、狭帯域 (narrowband) から全帯域 (full-band) に至るまでの帯域幅で、オーディオを圧縮することができ、ステレオ能力もあり、存在する3GPPコーデックのための窮極的な代替と見られる。3GPPで、新たなコーデックの動機 (motivation) は、さらに高いオーディオ帯域幅及びステレオを要求する新たなアプリケーションを除いたスピーチ・コーディング及びオーディオ・コーディング・アルゴリズムの発展 (advancement)、並びにサーキットスイッチされた環境で、パケットスイッチされた環境

10

でスピーチ及びオーディオのマイグレーション (migration) を含む。

【0054】

以前の3GPP基盤ネットワークの場合のように、EVSコーデックが動作する環境の主な様相 (aspect) は、送信機 (sender) から受信機 (receiver) へのスピーチ/オーディオフレームが伝送されるときにの損失である。これは、セルラ・ネットワークでの伝送時に予想される結果であり、かような環境で動作するように設計されたスピーチ及びオーディオの設計過程とすることができる。EVSコーデックは、スピーチのフレーム損失と、フレーム除去の衝撃とを最小化するためのアルゴリズムを含んでもよい。EPSだけではなく、レガシー3GPPセルラ・ネットワークも、一般的な条件の間、ほとんどのユーザに係わる合理的なフレーム除去の比率を維持するように設計される。

20

【0055】

図1のEVSコーデック26は、パケットが損失される環境である3GPPアプリケーションだけでなく、その後の3GPPでも使用されもする。さらに、何人のユーザは、所望のEVSより、フレーム除去の一般的な比率よりさらに高い比率を経験することができる。かような観点で、本発明は、EVSコーデックのためのhigh FER (high frame erasure rate) 動作モードを提案する。high FER動作モードは、特定環境で、追加的なフレーム損失軽減 (mitigation) を提供するために、追加的なリソース (追加的なビットレート及び/またはディレイ) を使用することができる。

【0056】

例えば、high FER動作モードは、LTEで、極限的な動作環境でのフレーム除去の比率を意味する。high FER動作モードで、10%またはそれ以上の程度でのフレーム除去の比率で、さらに優れた性能を発揮するためには、追加的なリソース (ビットレート、ディレイ) が要求されるトレードオフ (trade off) が存在する。

30

【0057】

本発明の一実施形態によれば、EVSコーデック26のhigh FER動作モードのために、FEC (frame erasure concealment) と直接に連結される。本発明の一実施形態は、特定パラメータの重要性に基づいて、スピーチフレームの多様なエンコーディングされたパラメータが、多様なリダンダンシ (redundancy) と共に伝送されるリダンダンシ方式を提案する。さらに、エンコーディングされたスピーチ部分ではない、エンコーダで生成されるFECビットは、優先化 (prioritized) され、多様なリダンダンシと共に伝送

40

される。リダンダンシは、多重パケットで、同じビットまたは全てのビットの反復を介して導出され、フレーム間またはフレーム内部で、差等的な (unequal) 方式で遂行される。

【0058】

図1はスピーチメディア・コンポーネント22の内部で、4世代3GPP方式のために、EVS (enhanced voice service) コーデック26及びボイスサービス・コーデック24を含むEPS (evolved packet system) 20を図示している。EVSコーデック26は、LTE無線インターフェースを介して、効率的に動作する。かような効率的な設計によって、多様なコーデック・フレームサイズとRTPペイロードは、LTEですでに定義された伝送ブロックサイズとマッチングされる。EVSコーデック26は、無線イン

50

ターフェース及びV O I Pネットワークでフレーム損失が発生したり発生しうる環境で動作するマルチレート及びマルチ帯域幅コーデックである。従って、本発明の一実施形態によれば、E V Sコーデック26は、フレーム損失の衝撃を低減させるためのF E C (frame erasure concealment) アルゴリズムを含む。

【0059】

オーディオ・コーディングでF E Cを利用するのは、スピーチまたはオーディオをエンコーディングしたり、あるいはエンコーディングするために使用されたスピーチ・コーデックと独立したデコーディング・システムによって遂行された。しかし、潜在的に、さらに効果的な利用のために、E V Sコーデック26のデコーダ側面の開発段階で、E V Sコーデック26で、F E Cアルゴリズムを設計するのである。

10

【0060】

エンコーダ側面で、エンコーダは、オーディオデータのスピーチをエンコーディングするために遂行されるコーデックと独立して、データに提供されたりダンダンシを有することができる。そのために、たとえ以前コーデックは、フレーム損失による品質悪化 (degradation) を減らすために、ただデコーダと係わるアルゴリズムを利用したが、本発明の一実施形態によれば、たとえシステム帯域幅の追加コストや潜在的なディレイが必要であるとしても、E V Sコーデック26のデコーダ側面の開発段階で、E V Sコーデック26のエンコーダに、F E Cアルゴリズムを採択することができる。

【0061】

本発明の一実施形態によれば、エンコーダに適用されるF E Cアルゴリズムだけでなくエラーまたはパケットの損失を隠匿するために、デコーダにも適切なF E Cアルゴリズムを適用することができる。そして、追加的なフレームエラー隠匿アルゴリズムの組み合わせが使用されもする。また、デコーダは、デコーディングされたオーディオデータの適切なタイミングを維持するために、エラーが発生したビットまたは損失されたパケットを再構成することができる。従って、E V Sコーデック26は、前述のフレーム損失隠匿だけでなく、F E Cフレームと係わる事項を遂行することができる。

20

【0062】

従って、本発明の一実施形態によれば、4世代3 G P P無線システム方式のように、エンコーダ基盤のF E Cアルゴリズムを採択することができる。そして、他の実施形態によれば、本発明は、エンコーディング動作とデコーディング動作とをそれぞれ遂行することができるエンコーダとデコーダとを含んでもよい。

30

【0063】

図2Aによれば、エンコーディング端末100、一つ以上のネットワーク140及びデコーディング端末150が図示されている。本発明の一実施形態によれば、一つ以上のネットワーク140は、E V Sコーデック26を含み、エンコーディング、デコーディングまたは変形 (transformation) を遂行することができる一つ以上の中間端末 (intermediary terminals) を含んでもよい。エンコーディング端末100は、エンコーダ側コーデック120、ユーザ・インターフェース130を含み、デコーディング端末150は、同様にデコーダ側コーデック160及びユーザ・インターフェース130を含んでもよい。

40

【0064】

図2Bは、本発明の一実施形態による、図2Aのエンコーディング端末100及びデコーディング端末150を一つまたは二ともいずれもだけではなく、一つ以上のネットワーク140内部の中間端末を代表する端末200を図示する。端末200は、マイク260のようなオーディオ入力装置と連結されたエンコーディング部205、スピーカ270のように、オーディオ出力装置と連結されたデコーディング部250、潜在的なディスプレイ230、入出力インターフェース235、中央処理装置 (C P U) 210のようなプロセッサを含んでもよい。

【0065】

C P U 210は、エンコーディング部205及びデコーディング部250と連結される。C P U 210は、エンコーディング部205とデコーディング部250との動作を

50

制御するだけでなく、端末200の他の構成要素を、エンコーディング部205とデコーディング部250との相互作用で制御することができる。本発明の一実施形態によれば、端末200は、モバイルフォン、スマートフォン、タブレットPC (personal computer) またはPDA (personal digital assistant) のようなモバイル装置でもある。そして、CPU 210は、端末の他の特徴を利用することができ、モバイルフォン、スマートフォン、タブレットPCまたはPDAでの一般的な機能のために、端末の能力 (capability) を利用することができる。

【0066】

例えば、本発明の一実施形態によれば、エンコーディング部205は、FECアルゴリズムまたはフレームワークに基づいて、デジタル的に入力オーディオをエンコーディング 10
することができる。保存されたコードブックは、適用されたFECアルゴリズムに基づいて、選択的に使用されもする。コードブックは、エンコーディング部205及びデコーディング部250のメモリに保存される。エンコーディングされたデジタルオーディオは、キャリア信号に変調されたパケットを介して伝送され、アンテナ240によって伝送されもする。また、エンコーディング・オーディオデータは、その後の再生のために、不揮発性メモリまたは揮発性メモリのようなメモリ215に保存されもする。

【0067】

他の一例として、本発明の一実施形態によれば、デコーディング部250は、FECアルゴリズムに基づいて、入力オーディオをデコーディングすることができる。デコーディング部250によってデコーディングされたオーディオは、アンテナ240から提供され 20
たり、あるいは以前にエンコーディングされたオーディオが保存されたメモリ215から獲得されもする。さらに、保存されたコードブックは、エンコーディング部205、デコーディング部250またはメモリ215に保存され、FECアルゴリズムに基づいて、選択的に使用されもする。

【0068】

前述のように、本発明の一実施形態によれば、エンコーディング部205及びデコーディング部250は、それぞれ適切なコードブック、及び適切なコーデック・アルゴリズムまたはFECアルゴリズムを保存するためのメモリを含んでもよい。エンコーディング部205及びデコーディング部250は、オーディオデータをエンコーディングしたり、あ 30
るいはデコーディングするために使用されるコーデックと共に、プロセッシング装置に含まれ、同一に使用される単一ユニット (single unit) でもある。本発明の一実施形態によれば、プロセッシング装置は、入力オーディオ、または他のオーディオ・ストリームの他の部分のために、並列的にエンコーディング・プロセッシング及び/またはデコーディング・プロセッシングを遂行することができる。

【0069】

端末200は、エンコーディング部205及び/またはデコーディング部250で遂行される複数の動作モードを選択するコーデックモード設定部255を含んでもよい。それぞれのコーデックモード設定部255それぞれは、エンコーディング部205及びデコーディング部250いずれものための1つのコーデックモード設定部255でもある。EVSコーデックは、同一の動作モードで、スピーチオーディオ及びノンスピーチ・オーディオである音楽 (music) をエンコーディングすることができる。もし入力オーディオが 40
ノンスピーチ・オーディオである場合、エンコーディング部205またはデコーディング部250は、音楽、またはさらに良質のオーディオのために設計されたコーデックのように、広帯域コーデック (wideband codec) によって、ノンスピーチ・オーディオをそれぞれエンコーディングしたり、あるいはデコーディングすることができる。

【0070】

もし入力オーディオがスピーチ・オーディオであると決定されれば、コーデックモード設定部255は、エンコーディング部205またはデコーディング部250それぞれが、オーディオデータをエンコーディングまたはデコーディングすることができるように、複数の動作モードを決定することができる。 50

【 0 0 7 1 】

もしコーデックモード設定部 2 5 5 が、high F E R 動作モードが決定されたということを感じた場合、コーデックモード設定部 2 5 5 は、high F E R 動作モードで動作するために、F E C モードのうち一つを選択することができる。たとえ動作モードが、high F E R 動作モードに設定されたために、スピーチ・コーディングのために活用可能な他の動作モードが利用されないとしても、F E C モードは、F E C フレームワークで、他のスピーチ・コーディング・モードと共に使用されもする。

【 0 0 7 2 】

コーデックモード設定部 2 5 5 は、エンコーディングされた入力パケットをパーズングし、受信されたエンコーディングされたオーディオがスピーチであるか否かを識別する情報、high F E R 動作モードが設定されているか否かを示すノンスピーチ・オーディオのための動作モード、F E R モードのために、いかなる潜在的な F E C 動作モードも抽出することができる。また、コーデックモード設定部 2 5 5 は、パーズングされた情報を、エンコーディングされた出力パケットに追加することができる。そして、かような情報は、窮極的な (ultimate) エンコーディングが遂行されるように、エンコーディング部 2 0 5 によって追加されもする。

10

【 0 0 7 3 】

本発明の一実施形態によれば、E V S コーデック 2 6 は、スピーチ・オーディオのための複数の動作モードを含んでもよい。動作モードそれぞれは、関連したエンコーディングされたビットレートを有することができる。特定モードでのビットレートに従属し、動作モードは、オーディオ帯域幅の選択を伝送たり、あるいはレガシー A M R - W B コーデックでエンコーディングされたスピーチを伝送するために多様に使用されもする。スピーチ・オーディオに係わる動作モードの例示は、以下の表 1 に図示されている。

20

【 0 0 7 4 】

L T E 無線インターフェースは、多様なサイズを有する伝送パケットで使用することができる固定された個数の伝送ブロックサイズに設計されもする。3 G P P 無線システムでは、存在する 3 G P P コーデックのために、伝送ブロックサイズよりさらに小さく設計されもする。そして、伝送ブロックサイズは、コーデックが動作するビットレートの厳格な選択を介して、E V S コーデック 2 6 によって再使用されもする。本発明の一実施形態において、E V S コーデック 2 6 は、エンドツーエンド・ディレイ (end-to-end delay) を最小化するために、スピーチを 2 0 m s フレームにエンコーディングすることができ、1つのフレームは、パケットごとに伝送される。しかし、本発明は、かような実施形態に限定されるものではない。

30

【 0 0 7 5 】

以下で図示された表 1 は、ビットレート範囲の低い部分でのスピーチ E V S コーデック・ビットレートの例示と、ビットレート・モードと結合して使用される伝送ブロックサイズを図示している。表 1 で例示された R T P ペイロードのサイズは、A M R - W B コーデックで存在する R T P ペイロードサイズに基づく。しかし、本発明の一実施形態は、表 1 の R T P ペイロードサイズに限定されるものではない。

【 0 0 7 6 】

40

【表 1】

表 1

EVS Codec Bit rate (kbps) [Ⓐ]	Bits per 20ms Frame [Ⓐ]	RTP Payload [Ⓐ]	Unused bits (one frame per packet) [Ⓐ]	LTE Transport Block Size [Ⓐ]
6.60 [Ⓐ]	132 [Ⓐ]	74 [Ⓐ]	2 [Ⓐ]	208 [Ⓐ]
7.50 [Ⓐ]	150 [Ⓐ]	74 [Ⓐ]	0 [Ⓐ]	224 [Ⓐ]
8.85 [Ⓐ]	177 [Ⓐ]	74 [Ⓐ]	5 [Ⓐ]	256 [Ⓐ]
11.10 [Ⓐ]	222 [Ⓐ]	74 [Ⓐ]	0 [Ⓐ]	296 [Ⓐ]
12.65 [Ⓐ]	253 [Ⓐ]	74 [Ⓐ]	1 [Ⓐ]	328 [Ⓐ]
14.25 [Ⓐ]	285 [Ⓐ]	74 [Ⓐ]	17 [Ⓐ]	376 [Ⓐ]
15.85 [Ⓐ]	317 [Ⓐ]	74 [Ⓐ]	1 [Ⓐ]	392 [Ⓐ]
18.25 [Ⓐ]	365 [Ⓐ]	74 [Ⓐ]	1 [Ⓐ]	440 [Ⓐ]
19.85 [Ⓐ]	397 [Ⓐ]	74 [Ⓐ]	1 [Ⓐ]	472 [Ⓐ]
23.05 [Ⓐ]	461 [Ⓐ]	74 [Ⓐ]	1 [Ⓐ]	536 [Ⓐ]
23.85 [Ⓐ]	477 [Ⓐ]	74 [Ⓐ]	1 [Ⓐ]	552 [Ⓐ]

前述のところは、固定レートコーデック、や固定レートでスピーチフレームをエンコーディングするコーデックに係わる。パケット・スイッチされた環境で動作するように、スピーチ発話（utterances）間のサイレンスまたは中止（pause）がエンコーディングされ、不連続的な方式で非常に低いレートで伝送される。

【 0 0 7 7 】

前述のように、ネットワークと、3 G P P セルラ・ネットワークとで伝送されたスピーチフレームは、伝送過程で伝送されたデータの小さい比率ほど除去される。

【 0 0 7 8 】

フレーム損失隠匿（F E C）アルゴリズムは、一般的に、2個のカテゴリに分類される。一つは、コーデック独立的 F E C アルゴリズムと、コーデック従属的 F E C アルゴリズムとである。コーデック独立的 F E C アルゴリズムは、特定コーディング・アルゴリズムの知識なしにも十分に適用され、コーデック従属的 F E C アルゴリズムほど、その結果が効率的である。コーデック従属的 F E C アルゴリズムは、開発過程で、コーデックと結合されるように設計され、一般的に、さらに効果的である。本発明の一実施形態によれば、少なくとも1つのコーデック従属的 F E C アルゴリズムを含んでもよく、コーデック従属的 F E C アルゴリズムと、コーデック独立的 F E C アルゴリズムとを含んでもよい。

【 0 0 7 9 】

フレーム損失隠匿（F E C）アルゴリズムは、2個のセットに分類される。フレーム損失隠匿（F E C）アルゴリズムは、受信機基盤の F E C アルゴリズム、及び送信機基盤の F E C アルゴリズムに分類される。受信機基盤の F E C アルゴリズムは、スピーチ・デコーダ、及び/またはデコーディング部 2 5 0 のジッタバッファに単独で位置することができる。そして、受信機基盤の F E C アルゴリズムは、デコーダのために受信機で生成されたフレーム除去フラグによって、触発になる。デコーディング部 2 5 0 のエラー隠匿（error concealment）は、サイレンス利用、ホワイトノイズ、波形置換（waveform substitution）、サンプル補間（sample interpolation）、ピッチ波形置換（pitch waveform replacement）、タイムスケール修正（time scale modification）、知識または隣接オーディオ特徴に基づいた再生（regeneration）、及び/またはモデルへのエラーまたは損失のうちいずれか1つのスピーチ特徴にマッチングされた復旧（recover）に基づいたモデルを含むデータ隠匿を含んでもよい。

【 0 0 8 0 】

ユーザがパケット損失を認知することを最小化することができるように簡単なアルゴリ

10

20

30

40

50

ズムは、除去されたフレーム、または以前良好なフレームの反復のために復元されたオーディオ (restored audio) に、サイレンスまたはノイズ代替 (noise substitution) を含んでもよい。フレーム除去の連続したストリング (continuing string) のために、デコーダは、デコーディングされたスピーチボリュームを音消去することができる。さらに向上されたアルゴリズムは、以前に受信された状態が良好なスピーチフレームの特徴を考慮し、以前に受信された状態が良好なパラメータを補間することができる。もしジッタバッファが採択されれば、補間目的のために除去されたフレームの両側面で、状態が良好なスピーチフレームを使用する機会がある。

【 0 0 8 1 】

送信機基盤の F E C アルゴリズムは、さらにリソースを消費するが、受信機基盤の F E C アルゴリズムよりさらに強力である。送信機基盤の F E C アルゴリズムは、一般的に、フレーム除去が発生した場合、損失されたフレームの再構成のために使用するリダンダント情報を、サイドチャンネルを介して伝送することができる。送信機基盤の F E C アルゴリズムの性能は、プライマリー・チャンネルからの付加情報伝送と相関関係がない。セルラ・ネットワークで、リアルタイムスピーチ・コーディング・アプリケーションのために、部分的に相関関係を除去することは、一つ以上のフレームにリダンダント情報を伝送することをディレイすることによって行われる。それは、典型的には、ディレイが制限されたシステムの伝送経路でディレイをもたらし、ディレイは、受信機にジッタバッファによって部分的に軽減される。ジッタバッファは、デコーディング部 2 5 0 に含まれる。

【 0 0 8 2 】

本発明の一実施形態によれば、受信機に提供される付加 (side) 情報またはリダンダンシ情報は、本来スピーチフレーム (全体リダンダンシ) の完璧な複写本 (copy)、またはフレームの臨界的 (critical) サブセット (部分リダンダンシ) を含んでもよい。選択的なリダンダンシは、スピーチフレームの選択されたサブセットが、付加情報と共に伝送される技術を意味する。全体スピーチフレームまたはフレームのサブセットは、選択的な方式で伝送される。

【 0 0 8 3 】

他のアプローチ方式は、スピーチを、2つの異なるコーデックでエンコーディングするのである。一つは、一般的なコーディングのために、所望のコーデックでエンコーディングするものであり、他の一つは、低いレート、低い正確度のコーデックでエンコーディングするのである。本発明の一実施形態によれば、多様なレンダリングが適用される。付加チャンネルの考慮された低いレートバージョンでエンコーディングされたスピーチが、デコーダに伝送される。

【 0 0 8 4 】

さらに、本発明の一実施形態によれば、差等的なエラー保護 (unequal error protection) が行われる。フレームの符号化されたビットは、クラスに分類される。クラス A, B, C は、除去されるビットまたはパラメータの敏感度に基づいて決定される。クラス A に属するビットまたはパラメータの除去 (erasure) は、クラス C に属するビットまたはパラメータが損失されるときより、ボイス品質にさらに大きい影響を及ぼす。符号化されたビットまたはパラメータをクラスに分類することは、フレームをサブフレームに分割することと参照される。サブフレームという用語の使用は、分類されたエンコーディングされたビットが、サブフレームそれぞれが連続的になることを要求しないということ意味する。

【 0 0 8 5 】

送信機基盤の F E C システムで、受信機は、フレーム除去を認識し、除去されたフレームのためのリダンダント付加情報が受信されているか否かを判断することができる。もし付加情報も損失された状況は、受信機基盤の F E C システムで、付加情報が損失されることと同一である。それにより、受信機基盤の F E C アルゴリズムが適用される。もしリダンダント付加情報が存在する場合、付加情報は、受信機が隠匿目的に使用することができる他の関連情報と、損失されたフレームとを隠匿するために使用される。

10

20

30

40

50

【 0 0 8 6 】

前述のように、E V Sコーデック 2 6 は、他の動作モードと区分されるhigh F E R動作モードを含んでもよい。E V Sコーデック 2 6のhigh F E R動作モードは、プライマリー動作モードではなく、ユーザが、フレーム損失が発生する一般的な状況よりさらによく経験する場合に選択される。

【 0 0 8 7 】

このメカニズムの成功と失敗は、フレームが無線インターフェースを介して首尾よく伝送された否かということのように、迅速なフィードバックを提供するのである。全体伝送経路を伴うリンク品質のフィードバックは、一般的に遅い。そして、フィードバックは、さらに高い階層通信、またはモバイルとモバイルとの通話のような場合、E V Sコーデック 2 6間に専念するバンド信号のうちいずれか一つを伴う。

10

【 0 0 8 8 】

本発明の一実施形態によれば、E V Sコーデック 2 6のhigh F E R動作モードのために、F E Cフレームワークが提供される。このフレームワークは、E V Sコーデック 2 6の固定レートモード及び帯域幅に有効である。一実施形態で、このF E Cフレームワークは、E V Sコーデック 2 6の全体固定レートモード及び帯域幅に有効である。従って、本発明の一実施形態によれば、フレームワークは、固定レートでエンコーディングされたフレームの部分的または全体的なりダングランシの伝送方法を含んでもよい。

【 0 0 8 9 】

本発明の一実施形態によれば、部分的及び全体的なりダングランシは、high F E R動作モードの間、固定されたサイズの伝送ブロックを伝送することができる。一般的な動作モードで、high F E R動作モードへの転移は、伝送ブロックサイズの変化を引き起こす。本発明の一実施形態によれば、(1)固定されたか、あるいは多様なビットレートと、固定されたサイズとの伝送ブロックと共に、部分的 (partial)、差等的 (unequal) または全体的 (full) なりダングランシを使用したり、あるいは (2)固定されたり、あるいは多様なビットレートと多様なサイズとの伝送ブロックと共に、部分的、差等的または全体的なりダングランシを使用することができる。

20

【 0 0 9 0 】

本発明の一実施形態によれば、図 1 で、E V Sコーデック 2 6のhigh F E R動作モードは、選択的なりダングランシの例示を示している。

30

【 0 0 9 1 】

以下で説明するように、E P S環境で、E V Sコーデック 2 6と相互作用する 2 種の例示がある。ここで、相互作用というのは、エンコーディング部 1 0 0がhigh F E R動作モードとして決定するか否かを判断するために、デコーディング部 1 5 0からエンコーディング部 1 0 0へのフィードバックを意味する。そして、デコーディング部 1 5 0は、フレーム除去レートをモニタリングすることにより、high F E R動作モードに入るか否かを決定することができる。

【 0 0 9 2 】

もしデコーディング部 1 5 0が、high F E R動作モードに入ると決定する場合、かような決定は、オーディオまたはスピーチの次のフレームを、high F E R動作モードでエンコーディングするように、エンコーディング部 1 0 0に伝送される。同様に、図 2 Bから分かるように、もしエンコーディング部 1 0 0及びデコーディング部 1 5 0のうちいずれか一つが受信された情報に基づいて、high F E R動作モードに入ると決定されれば、端末 2 0 0は、カンファレンス・コールまたはV O I Pセッションから、オーディオデータまたはスピーチデータをエンコーディングしたり、あるいはデコーディングすることができる。そして、端末 2 0 0は、high F E R動作モードで、次のフレームをエンコーディングすることができ、終端に位置した端末 2 0 0が、high F E Rモードで動作するように、終端に位置した端末 2 0 0に通知することができる。また、デコーダは、フレームと関連したシグナリングから、フレームがhigh F E Rモードにあるか否かが分かる。

40

【 0 0 9 3 】

50

EVSコーデック26は、4種のソースのうち一つ以上を処理された情報に基づいて、high FER動作モードに入ることができる。ここで、4種ソースは、次の通りである。
 (1) 物理的階層で伝送されたハイブリッド自動回復要請(HARQ: hybrid automatic repeat request)フィードバックであるファースト・フィードバック(FFB: fast feedback)情報、(2) 物理的階層よりさらに高い階層で伝送されたネットワーク・シグナリングからフィードバックされたスロー・フィードバック(SFB: slow feedback)情報、(3) 終端(far end)で、EVSコーデック26からインバンド・シグナリングされたフィードバック(IS: in-band feedback)情報、及び(4) リダundant方式(redundant fashion)で伝送される特定クリティカル・フレーム(specific critical frame)のEVSコーデック26による選択であるハイセンシティビティ・フレーム(HSF: high sensitivity frame: HSF)情報。ソース(1)及び(2)は、EVSコーデック26に独立的である一方、ソース(3)及び(4)は、EVSコーデック26に依存的であり、EVSコーデック26のための特定アルゴリズムを要求する。

10

【0094】

high FER動作モードに入るか否かを決定することは、high FER動作モード・アルゴリズムに基づく。本発明の一実施形態によれば、図2Bのコーディング・モード設定部255は、以下のアルゴリズム1で図示されたところによって、high FER動作モード・アルゴリズムを遂行することができる。

(アルゴリズム1)

20

Definitions	Set During Initialization
<ul style="list-style-type: none"> • SFBavg: Average error rate over Ns frames • FFBavg: Average error rate over Nf frames • ISBavg: Average error rate over Ni frames • Ts: Threshold for slow feedback error rate. • Tf: Threshold for fast feedback error rate. • Ti: Threshold for inband feedback error rate. 	<ul style="list-style-type: none"> • Ns=100 • Nf=10 • Ni=100 • Ts=20 • Tf=2 • Ti=20

30

Algorithm

```

Loop over each frame {
HFM = 0;
IF((HiOK) AND SFBavg > Ts) THEN HFM = 1;
ELSE IF ((HiOK) AND FFBavg > Tf) THEN HFM = 1;
ELSE IF ((HiOK) AND ISBavg > Ti) THEN HFM = 1;
ELSE IF ((HiOK) AND (HSF=1) THEN HFM = 1;
Update SFBavg;
Update FFBavg;
Update ISBavg;
}

```

40

前述のように、本発明の一実施形態によれば、図2Bのコーディング・モード設定部255は、4個のソースのうち一つ以上処理された分析情報に基づいて、EVSコーデック26に、high FERモードに入ることを指示することができる。ここで、ソースは、次

50

の通りである。(1)SSF情報を利用して、Nsフレームの計算された平均エラーレートから導出されたSFBavg、(2)FFB情報を利用して、Nsフレーム平均の計算された平均エラーレートから導出されたFFBavg、(3)ISB情報と、それぞれの臨界値であるTs、Tf及びTiを利用して、Nsフレームの計算された平均エラーレートから導出されたISBavg。

【0095】

それぞれの臨界値を比較した結果に基づいて、図2Bのコーディング・モード設定部255は、highFER動作モードに入るか否かということと、選択するFECモードとを決定することができる。選択されたFECモードは、表6及び表7で説明するコーディング・タイプ及びフレーム分類決定に基づく。

10

【0096】

本発明の一実施形態によれば、highFER動作モードに入るという決定に従属し、オーディオ情報またはスピーチ情報をエンコーディングするために、追加してhighFER動作モードに含まれた複数のサブモードが存在する。ここで、highFER動作モードは、複数のサブモードで動作し、小さい数のビットは、選択されたそれぞれのサブモードに係わるシグナリングのために使用される。ここで、小さい数のビットは、オーバーヘッド部分になり、潜在的に、現在または将来の4世代3GPP無線ネットワーク方式で、保有ビット(reserved bit)にもなる。

【0097】

本発明の一実施形態によれば、RTPペイロードでの1つのビットは、highFER動作モードをシグナリングするために要求される。この1つのビットは、highFERモードフラグとすることができる。例えば、既存のAMR-WBで、RTPペイロードは、4個の余分ビット(extra bit)を有し、かようなビットは、割り当てられずに保有される。さらに、highFER動作モードで、サブモードをシグナリングするために、いくつかのビットの保有が要求される。かようなビットは、FECモードフラグとすることができる。それらビットは、表3のクラスAに属するビットのためのリダンダンシと類似した方式でリダンダンシとして保護される。

20

【0098】

送信機基盤のFECアルゴリズムは、一般的に、リダンダント情報を伝送するために、付加チャンネル(side channel)を使用することができる。本発明の一実施形態によれば、EVSコーデック26のコンテキスト及びEPSで、コンテキストの使用側面で、たとえ予想されるEVSコーデックが付加チャンネルを提供しないとしても、LTE無線インターフェースで定義された伝送ブロックを効率的に使用することができる。動作モードそれぞれについて、下記表2は、最初から次に大きい(next higher)、または2番目の次に大きい(second next)伝送ブロックサイズが活用可能な追加ビットの個数を示す。本発明の一実施形態によれば、効率的な動作のために、全ての追加ビットが使用される。

30

【0099】

【表 2】

表 2

Bit rate (kbps)	Bits per Frame	RTP Payload	unused	Transport Block Size	# FEC bits if Using next TBS	# FEC bits if using 2 nd larger TBS
6.60	132	74	2	208	16	48
7.50	150	74	0	224	32	72
8.85	177	74	5	256	40	72
11.10	222	74	0	296	32	80
12.65	253	74	1	328	48	64
14.25	285	74	17	376	16	64
15.85	317	74	1	392	48	80
18.25	365	74	1	440	32	96
19.85	397	74	1	472	64	80
23.05	461	74	1	536	16	
23.85	477	74	1	552		

フレーム n と無関係なパケットに、フレーム n と係わるリダンダント・ビットまたはパラメータを送送することにより、フレーム損失の強靭性 (robustness) が遂行される。例えば、フレーム n と係わるエンコーディングされたビットは、パケット N で送送される一方、フレーム n と係わるリダンダント・ビットは、パケット N + 1 で送送される。それは、時間ダイバーシティ (time diversity) として知られている。もしパケット N が除去され、パケット N + 1 が有効に送送されるのであるならば、リダンダント・ビットは、フレーム n を隠匿したり、あるいは再構成するために使用される。

【 0 1 0 0 】

図 3 は、本発明の一実施形態による、代替パケット (alternate packet) に提供される 1 つのフレームのためのリダンダント・ビットの例示を示している。図 3 で、第 1 パケットは、EVS コーデック 2.6 で、high FER 動作モードではない一般動作モードを示す。そして、AMR-WB コーデックの RTP ペイロードのヘッダーサイズと同一に、図 3 の RTP ペイロードのヘッダーサイズは、74 ビットである。

【 0 1 0 1 】

中間パケットは、high FER 動作モードでの送送メカニズムを示す。そして、118 個の FEC ビットは、以前フレーム (n - 1) のためにパケットに含まれる。リダンダント情報が含まれた中間パケットは、送送ブロックのサイズが 472 である。3 番目のパケットは、high FER 動作モードで動作するパケットの次のところに位置する。3 番目のパケットは、再び high FER 動作モードでの送送メカニズムを示し、118 個の FEC ビットが、以前フレーム n のために、パケットに含まれる。従って、本発明の一実施形態によれば、high FER 動作モードで、少なくとも 1 つの代替パケットでのデータは、リダンダント情報を送送するために使用される。

【 0 1 0 2 】

図 4 は、本発明の一実施形態による、フレーム n のためのリダンダント・ビットが 2 個の代替パケットに提供されることを図示している。図 4 に図示されたように、それぞれのパケットは、それぞれのフレームのための EVS エンコーディングされたソースビットと、2 個の以前フレームのための FEC ビットとを含む。例えば、パケット (N + 2) は、EVS エンコーディングされたソースビット、フレーム (n + 1) のための FEC ビット、及びフレーム n のための FEC ビットを含む。他の方法として、フレーム n のためのリダンダント・ビットは、2 個の以後の (N + 1) パケットと (N + 2) パケットとを介して送送される。

【 0 1 0 3 】

10

20

30

40

50

図5は、本発明の一実施形態による、フレームnの packets の前後に位置した代替 packets に提供されるフレームnに係わるリダンダンシ・ビットの例示を図示した図面である。図5を参照すれば、packets の前後位置に存在する packets に、リダンダンシ・ビットが位置するように、エンコーダは、ディレイのための余分フレームを挿入することができる。ここで、リダンダンシ・ビット (redundancy bits) は、ターゲット・フレームに係わるEVSエンコーディングされたソースビットを含む。図5のように、デコーダで、エンコーダへの追加的なディレイがシフトされる。さらに、図5のように、シーケンスで真っ先に除去されたリダンダンシ・ビットよりは、伝送が成功したシーケンス内部で、中間に除去されたリダンダンシ・ビットの3個の除去結果 (triple erasure results) のような除去パターンがシフトされる。代替 packets は、隣接 packets とされ、追加 packets は、中間 packets の前後に位置する非連続的な (non-consecutive) packets を含む。追加 packets は、隣接 packets として参照される。

10

【0104】

さらに、他の隣接 packets で、リダンダンシ・ビットが位置し、リダンダンシ・ビットは、知覚的な重要度 (perceptual importance) に基づいて、過不足 (more or less) リダンダンシが選択的に含まれもする。

【0105】

従って、本発明の一実施形態によれば、固定ビットレートに係わるhigh FERモードは、知覚的な重要度により、さらに大きいリダンダンシ、同一のリダンダンシ、またはさらに小さいリダンダンシでエンコーディングされたスピーチビットを優先化して保護することができる差等的なリダンダンシ保護概念 (unequal redundancy protection concept) を使用することができる。例えば、本発明は、3GPPコーデックであるAMR及びAMR-WBを使用してエンコーディングされたビットをクラスに分類することができる。例えば、クラスA, B, Cで、クラスAに属するビットは、除去されるとき、最も敏感なビットを意味し、クラスCに属するビットは、除去されるとき、最も敏感ではないビットを意味する。アプリケーションが、サーキット・スイッチされた伝送 (circuit-switched transport)、または packets ・スイッチされた伝送 (packet-switched transport) を使用するか否かにより、それらビットを保護するための異なるメカニズムが存在する。

20

【0106】

本発明の一実施形態によれば、差等的なリダンダンシ保護概念は、エンコーディングされたソースビットだけではなく、追加的なFEC付加情報に拡張される。異なるクラスに属するビットは、時間ダイバーシティを利用して、リダンダンシ方式で伝送される。そして、ビットのクラスにより、リダンダンシの量が変更される。

30

【0107】

図6は、本発明の一実施形態による、ソースビットが属する異なる分類に基づいて、代替 packets に含まれたソースビットの差等的なリダンダンシを図示している。図6は、図3ないし図5に図示された方法と異なる方法を意味する。

【0108】

図6に図示されたように、ソースビットに係わる3個のカテゴリーが定義される。クラスAに属するソースビットは、3個の連続的な packets を介して、3回リダンダンシに (redundantly) 伝送される。そして、クラスBに属するソースビットは、2個の連続的な packets を介して、2回リダンダンシに伝送される。また、クラスCに属するソースビットは、1回リダンダンシに伝送される。図6で、Nは、 packets 番号を示し、nは、フレーム番号を示す。図6の例示で、同じサイズを有した packets それぞれは、RTPペイロードに追加された $3 * A + 2 * B + C$ ビットを含む。

40

【0109】

デコーディング部250のように、デコーダのジッタバッファ深 (jitter buffer depth) が十分である場合、デコーダは、クラスAに属するソースビットまたはパラメータを3回デコーディングする機会を有し、クラスBに属するソースビットまたはパラメータ

50

を2回デコーディングする機会を有し、クラスCに属するソースビットまたはパラメータを1回デコーディングする機会を有する。

【0110】

例えば、選択的な実施形態として、エンコーディングされたソースビットは、クラス(A, B)または(A, B, C, D)のように、さらに少なかったり、あるいは多いクラスに分類される。全体リダンダンシは、部分リダンダンシよりクラスCに属するビットを追加的に伝送することによって行われる。そして、さらに高い動作効率のために、クラスCに属するビットは、伝送されないこともある。そして、効率的な目標のために、クラスAに属するビットだけ伝送されもする。

【0111】

従って、本発明の一実施形態によれば、現在フレームの以前フレームまたは以後フレームである隣接フレームに、現在フレームのためのFECビットが追加して含まれる。ソースフレームのビットは、それらの知覚的な重要度のような優先度に基づいて、カテゴリ化される。最大の知覚的重要度を有したり、あるいは損失されたとき、人間の耳にさらに敏感であったり、あるいは認知されるソースフレームのビットまたはパラメータは、さらに低い知覚度を有した同じソースフレームのビットまたはパラメータよりさらに多くの隣接パケットを介してリダンダントに伝送される。

【0112】

エンコーダから導き出された付加情報は、エンコーディング・アルゴリズムの一部にもなる。以下で具体的に説明するように、付加情報は、他のビットまたはパラメータのようにリダンダントに伝送される。

【0113】

隠匿目的のために、本発明の一実施形態によるデコーダは、図3ないし図6のように、エンコーディングされたソースビットのリダンダント複写本に係わる利益だけでなく、デコーダFECアルゴリズムのために、特別に設計されたFECパラメータに係わる利益を受けることができる。一例として、ITU-Tスピーチ・コーデック標準G.718で、16個のFECビットは、コーデックの3階層から付加情報として伝送され、隠匿的に1階層が使用される。

【0114】

一例として、下記表3では、G.718コーデックと係わり、EVSコーデック26及び付加情報の6.6Kbpsモードを使用することができる。EVSコーデック26の6.6Kモードは、132個のソースビットを含む。さらに、G.718コーデックと同様に、FECビットをシグナリングするための2個のビットと、FEC付加情報のための16個のビットとを追加して定義することができる。下記下表は、本発明の一実施形態による、優先度に基づいて、EVSソースビットとFECビットとを割り当てする例を示している。

【0115】

10

20

30

【表 3】

表 3

Priority	EVS Codec 6.6K Mode Source Bits	FEC Bits
A	41 coder_type (3) ISF's (31) midISFs (4) Energy (3)	4 (G.718) frame class (2) FEC sub-mode (2)
B	43 1 st subframe pitch(8) all subframe gains (4*5) 2 nd -4 th subframes pitch(3*5)	14 (G.718) Pulse position (8) (G.718) Energy (6)
C	48 cb_bits (4*12)	-
Total	132	18

10

20

前記表 3 から分かるように、全体 (4 5 + 5 7 + 4 8) ビットが伝送される。前述のリダダンシ方法を利用すれば、各パケットは、全体 (3 A + 2 B + C = 2 9 7) ビットと、7 4 R T P ペイロード・ビットとから構成された総 3 7 1 ビットを含む。伝送ブロックの全体サイズ 3 7 6 で 5 ビットが余る。そして、他のクラス A , B , C に分類されたソースビットは、動作モードに基づいて、コーデックが C E L P (code-excited linear prediction) コーデックで動作するとき、線形予測パラメータのように、異なって分類されたスピーチのパラメータを示す。

【 0 1 1 6 】

従って、本発明の一実施形態による、一回 high F E R モードに入る場合、使用可能な帯域幅 (容量 : capacity) 及び F E C 保護 (強靭性) の程度により、使用可能なさまざまなサブモードが存在する。それらパラメータは、要求する固有したスピーチ品質の量とトレードオフ関係にある。例えば、帯域幅、品質、エラー強靭性の互いに異なる優先順位に基づいて、6 個のサブモードが存在する。下記表 4 は、多様なサブモードの属性を示している。

30

【 0 1 1 7 】

以下の例示のように、クラス A , B 及び C と表現されるソースビットのリダダンシ伝送を仮定し、献身的な (dedicated) F E C ビットがないと仮定する。さらに容易には、R T P ペイロードのサイズは、全ての例で 7 4 と仮定する。

【 0 1 1 8 】

【表 4】

表 4

Sub-mode	Bit Rate	TBS	Numerology	Features
Normal Mode	Depends on Codec Mode (12.65Kbps in example)	Depends on Codec Mode (328 in example)		Original codec mode. One of N may be selected.
1	7.5 Kbps	224	A, B, C = 14, 62, 56. 2A+B+C=150. 150+74 = 224.	Shift to 6.6K mode. Single redundancy of class A bits only. Mild robustness and lower capacity impact.
2	8.85 Kbps	256	A, B, C = 14, 62, 56. 3A+2B=166. 166+74=256.	Shift to 6.6K mode. Dual redundancy of class A bits. Single redundancy of class B bits. Drop the class C bits. Lower capacity desired and high redundancy of more critical bits.
3	11.1 Kbps	296	A, B, C = 14, 62, 56. 3A+2B+C=222. 222+74 = 296.	Shift to 6.6K mode. Dual redundancy of class A bits. Single redundancy of class B bits. No redundancy in class C bits. Higher redundancy and lower capacity than original.
4	Depends on Codec Mode (12.65Kbps in example)	Depends on Codec Mode (TBS = 328 in example)	A, B, C = 46, 30, 56. 3A+2B+C=254. 254+74=328.	Shift to 6.6K mode. Maintains original packet size. No capacity impact. Lower quality and higher robustness.
5	14.25 Kbps	376	A, B, C = 38, 38, 56. 3A+2B+2C=302. 302+74 = 376.	Shift to 6.6K mode. Full redundancy of all source bits. Dual redundancy of class A bits.
6	Depends on Codec Mode (18.25Kbps in example)	Depends on Codec Mode (TBS = 440 in example)	A, B, C = 20, 73, 160 3A+2B+C = 366 366+74 = 440	Maintain original codec mode. Add redundancy into a larger packet. Packet size depends on the original mode. Maintain high quality, higher robustness at cost of capacity.

図7は、本発明の一実施形態による、差等的なリダンダンシが適用されたFEC動作モードの例示を図示している。例えば、多くのサブモードは、high FER動作モードではないスピーチモードで遂行するように、同一のEVSコーディング・モードを使用する。当該例として、最も低いモードは、効率性目的のために選択され、high FER動作モードであるとき、強靭性及び容量の優先順位が最も高い。さらに、同じEVSコーディング・モードを使用することは、デコーダが1つのFECコーディング・モードを使用するように、FECアルゴリズムを単純化することができる。選択的に、以下で説明するように、本発明の他の実施形態は、追加的なコーディング・モードを使用することができる。

【0119】

図7から分かるように、増加されたリダンダンシを収容するように、サイズがさらに大

10

20

30

40

50

きいパケットのために、サブモード1からサブモード6にサブモード過程が増大する。

【0120】

図11は、本発明の一実施形態による、high FER動作モードの異なるFECモードを利用して、オーディオデータをコーディングする方法を図示する。図11に図示されたように、段階(1105)で、入力オーディオが分析され、入力オーディオは、スピーチ・オーディオであるか、あるいはノンスピーチ・オーディオであるかが決定される。もし入力オーディオがノンスピーチ・オーディオである場合、段階(1110)で、入力オーディオは、ノンスピーチ・コーデックでエンコーディングされたり、あるいはノンスピーチモードのEVSコーデック26でエンコーディングされる。もし入力オーディオがスピーチ・オーディオである場合、段階(1115)でbhigh FER動作モードに入るか否かを判断することができる。high FER動作モードに入るか否かを判断するのは、前述のアルゴリズム1と係わる。

10

【0121】

もし段階(1115)でhigh FER動作モードに入ると決定されていなければ、段階(1120)で、前述の表1の動作モードのうち一つが、EVSコーデック26のために選択される。段階(1120)で、一回スピーチ・エンコーディングのための動作モードが選択されれば、段階(1130)で、スピーチ・エンコーディングのために選択された動作モードによって、入力オーディオがエンコーディングされる。もし段階(1115)で、high FER動作モードに入ると決定されれば、段階(1125)で、多様なFEC動作モードのうち一つのFEC動作モードが選択される。そのために、段階(1135)で、入力オーディオは、選択されたFEC動作モードで、EVSコーデック26を利用してエンコーディングされる。

20

【0122】

同様に、図14は、本発明の一実施形態による、high FER動作モードで、異なるFECモードを使用して、オーディオデータをデコーディングする過程を図示している。段階(1405)で、受信されたパケット内部に存在するエンコーディングされたフレームが、スピーチ・オーディオまたはノンスピーチ・オーディオに基づいて、エンコーディングされているか否かを判断することができる。もしエンコーディングされたフレームが、ノンスピーチ・オーディオである場合、段階(1410)で、EVSコーデック26が適切な動作モードを利用して、ノンスピーチ・オーディオをデコーディングすることができる。

30

【0123】

もし受信されたパケットに、エンコーディングされたスピーチデータが含まれた場合、段階(1415)で、パケットは、スピーチデコーディングのための動作モードを決定するためにパーズングされる。ここで、動作モードは、フレームがhigh FER動作モードでエンコーディングされているか否かを決定することができる。例えば、high FERモードフラグが受信されたパケットに設定されておらず、フレームがhigh FER動作モードでエンコーディングされていない場合、段階(1420)で、スピーチ・デコーディングのための適切な動作モードが選択され、EVSコーデック26は、選択された動作モードで、スピーチ・デコーディングを遂行することができる。もしフレームがhigh FER動作モードでエンコーディングされたものであるならば、段階(1425)で、フレームをエンコーディングするとき、いかなるFEC動作モードが使用されたかを判断するために、パケットがパーズングされる。EVSコーデック26は、判断されたFEC動作モードに基づいて、フレームをデコーディングすることができる。

40

【0124】

ここで、本発明の一実施形態によれば、図14の方法は、段階(1405)と段階(1405)とが動作する以前、あるいは動作する間に判断する段階をさらに含む。具体的には、パケットが損失されているか否かを判断する段階がさらに含まれる。かような判断は、本発明の一実施形態による、隣接パケットに含まれたりリダンダント情報に基づいて、損失されたパケットを再構成(reconstruct)したり、あるいは損失されたパケットを隠匿

50

するために、F E Cフレームワークに基づいて、以前パケットまたは以後パケットで、リダンダント情報を使用するように、E V Sコーデック26での命令を含む。

【0125】

図7と異なる伝送ブロックサイズを代替するために、一般的な(regular)伝送モードで使用されるような複数の動作モードのために、同じ伝送ブロックサイズが維持される。かような場合、E P Sシステムが、パケットサイズの変更をシグナリングする必要のないものではなく、high F E Rモードで、多くのE V Sコーデック26の動作モードを利用する短所がないということの意味する。さらに多くのコーデックモードを使用するほど、隠匿アルゴリズムは、さらに複雑になる。

【0126】

図8は、本発明の一実施形態による、同じ伝送ブロックサイズを有したhigh F E R動作モードで、異なるF E C動作モードを図示した図面である。ここで、異なるF E C動作モードは、high F E R動作モードのサブモードとすることができる。その例として、E V Sコーデック26の12.65 K b p sは、一般的なnon-high F E R動作モードの一例として使用される。high F E R動作モードのサブモード1-4それぞれは、同じ伝送ブロックサイズ328を維持する。低いソース・コーディングの比率によって、リダンダントの増加が伴いもする。

【0127】

サーキット・スイッチされた伝送で、マルチモードA M Rコーデック及びA M R - W Bコーデックのように、他の3 G P Pコーデックによって使用される以前の方法と異なり、チャンネル条件に基づいて、さらに低いか、あるいは増加されたビットレートで、モードがスイッチされる。図8は、追加的なリダンダントまたはF E Cビットが含まれたり、あるいはフレームパケットサイズが維持されるように、異なるサブモードでビットレートが低下するところを図示している。

【0128】

図12は、本発明の一実施形態による、全てのF E C動作モードのために、同じビットレートまたはパケットサイズで維持するか否かに基づいたF E Cフレームワークを図示した図面である。図12に図示されたように、段階(1125)で、F E C動作モードが選択され、段階(1125)で、E V Sコーデック260は、選択されたF E C動作モードによって遂行される。図示されているように、段階(1125)で、段階(1220)または段階(1230)によって表現されたF E C動作モードのうち一つを直接に選択したり、あるいは段階(1210)で、同じビットレートまたは同じパケットサイズが決定されれば、段階(1220)が遂行され、他のビットレートまたは異なるパケットサイズが決定されれば、段階(1230)が遂行される。

【0129】

図7と同様に、段階(1230)が考慮される。ここで、パケットサイズは、多様に変更可能である。そして、段階(1220)で、隣接フレームから抽出されたエンコーディングされたE V Sソースビットは、現在パケットのエンコーディングされたE V Sソースビットの低減されたレートモードに追加される。具体的には、段階(1220)で、E V Sビットレートは、低いビットレート・モードに変更される。その場合、隣接フレームから抽出したソースビットは、本来の動作モードとパケットサイズを同一に維持するために追加される。段階(1220)で、E V Sビットレートは、本来の動作モードと同一に維持される。その場合、隣接フレームから抽出したソースビットは、パケットサイズと無関係に追加される。

【0130】

段階(1240)で、high F E R動作モードに入り、F E C動作モードが選択されれば、F E C付加情報は、エンコーディングされたフレームのパケットで、フラグとして反映される。high F E R動作モードは、パケット内部で、1つのビットを利用して設定され、選択されたF E C動作モードは、2~3個のビットを利用して設定される。

【0131】

10

20

30

40

50

隣接フレームから導き出された全ての情報は、リダンダンシ情報である。リダンダンシ情報は、現在パケットで伝送される。現在フレームと関連したリダンダンシ情報は、隣接した隣接パケットを介して伝送される。もし同じビットレートを維持するためには、リダンダンシ・ビットを収容するように、パケットサイズが増大させることができる。そして、同じパケットサイズを維持するために、ソースビットの個数が減少するように、コーディング・モードが変更される。

【 0 1 3 2 】

本発明の一実施形態によれば、high F E R動作モードに入った後、コードブック「robbing」を伴い、同じ伝送ブロックサイズを維持することができる。そして、コードブックは、表 4 及び図 8 のサブモード 1 と同様に、リダンダンシの小さい量を提供するとき

10

【 0 1 3 3 】

【表 5】

表 5

total bits	132 6k60	150 7k50	177 8k85	222 11k10	253 12k65	285 14k25	317 15k85	365 18k25	397 19k85	461 23k05	477 23k85
speech/audio core	0	1	1	1	1	1	1	1	1	1	1
coder type	3	3	3	3	3	3	3	3	3	3	3
ISFs	31	31	31	38	37	37	43	43	43	43	43
midISFs	4	4	3	4	4	4	5	5	5	5	5
scal. energy pred.	3	3	3	4	4	4	5	5	5	5	5
LTP filtering	0	0	0	4	4	4	4	4	4	4	4
1st subfr pitch	8	8	9	9	9	9	9	9	9	9	9
1st subfr gain	5	6	6	6	6	6	6	6	6	6	6
2nd subfr pitch	5	4	5	5	5	5	5	5	5	5	5
2nd subfr gain	5	6	6	6	6	6	6	6	6	6	6
3rd subfr pitch	5	4	5	5	5	5	5	5	5	5	5
3rd subfr gain	5	6	6	6	6	6	6	6	6	6	6
4th subfr pitch	5	4	5	5	5	5	5	5	5	5	5
4th subfr gain	5	6	6	6	6	6	6	6	6	6	6
cb_bits 1st subfr	12	12	20	28	36	44	52	64	72	88	88
cb_bits 2nd subfr	12	12	20	28	36	44	52	64	72	88	88
cb_bits 3rd subfr	12	20	20	28	36	44	52	64	72	88	88
cb_bits 4th subfr	12	20	28	36	44	52	52	64	72	88	88
HB energy	0	0	0	0	0	0	0	0	0	0	16

20

30

40

本発明の一実施形態において、もし E V S コーデック 2 6 の一般的な動作モードが、1 2 . 6 5 K b p s であるならば、high F E R動作モードに入るように、一般的な動作モードが維持される。エンコーダが、4 個のサブフレームのうち一つについて、high F E R動作モードで動作すれば、動作モードが、実際に 1 2 . 6 5 K b p s であるとしても、

50

動作モードが 8 . 8 5 K b p s で動作するように、コードブック・ビットを計算することができる。サブフレームは、フレームのオーディオを表現するフレームのビットまたはパラメータによって表現される。パラメータは、コーデックが C E L P コーデックで動作するとき、コーデックによって生成される C E L P (code-excited linear prediction) コーディングの線形予測パラメータを含む。

【 0 1 3 4 】

前述の表 5 のように、1 2 . 6 5 K b p s 動作モードによって、コードブック・ビットが計算されるのであるならば、要求される 3 6 ビットの代わりに、最初ないし 3 番目のサブフレームのビットについて、コードブックを定義するために、2 0 ビットが使用される。F E C の目的のために、コードブック「r o b b i n g」を利用することにより、1 6 ビットが節約される。F E C ビットの伝送は、同じ個数のビットが存在するために、本来の動作モードのように、同じパケットサイズで行われる。ほとんどの high F E R 動作モードのサブモードのように、かようなアプローチと関連した若干の品質劣化が存在する。

10

【 0 1 3 5 】

表 4 及び図 8 のアプローチと異なり、high F E R 動作モードのサブモードそれぞれについてソース・コーディングを行うコーデックのために、ビットレートは、順次に低下する。表 5 によれば、ビットレートが低下したビットレートである場合、ビットレートは、低下させるだけではなく、コードワードを計算する必要がない。図 8 に図示された F E C 情報は、図 1 ないし図 6 で説明されるところと類似したリダンダンシを含む。前記リダンダンシは、前記表 3 で説明された差等的なリダンダンシを含む。ここで、分割されたサブフレームは、それぞれ表 3 で、A , B または C それぞれのために使用される。ここで、さらに重要なサブフレームまたはパラメータは、他のサブフレームまたはパラメータよりさらに多くのリダンダンシを有する。

20

【 0 1 3 6 】

図 1 3 は、本発明の一実施形態による、F E C 動作モードの 3 種の例示を図示している。表 3 及び図 6 で考慮したように、フレームのビットまたはパラメータは、知覚的重要度によってクラスに分類される。従って、段階 (1 3 1 0) で、ビットを異なるクラスまたはサブフレームに分類するために、フレームは、分割されたり、あるいは分離される。そして、段階 (1 3 1 5) で、各クラスまたはサブフレームに係わるリダンダント情報は、図 6 及び図 7 のように、隣接フレームに差等的に提供される。

30

【 0 1 3 7 】

段階 (1 3 2 0) で、分割されたり、あるいは分離されたビットまたはパラメータそれぞれについて、コードブック・ビットの個数が計算される。フレームの動作モードに係わるビットレートより低いビットレートでエンコーディングされるために、ビットまたはパラメータは、クラスとサブフレームとに分類される。従って、段階 (1 3 3 0) で、計算されたコードブック・ビットの個数に基づいて、定義されたコードワードは、エンコーディングされる。

【 0 1 3 8 】

さらに、段階 (1 3 4 0) で、定義されたコードワードを考慮するとき、図 6 及び図 7 と同様に、エンコーディングされたクラスまたはサブフレームのリダンダント情報は、隣接パケットに差等的に提供される。

40

【 0 1 3 9 】

前述の図 3 ないし図 8、及び表 3 ないし表 5 の high F E R 動作モードは、スピーチフレームが、ビットのクラスまたはパラメータのクラスに分類するために利用される。ビットのクラスまたはパラメータのクラスは、除去されるビットまたはパラメータの知覚的重要度によって区分される。

【 0 1 4 0 】

しかし、G . 7 1 8 コーデック及び予想された E V S 候補コーデックを含むいくつかのスピーチ・コーデックで、入力スピーチフレームは、スピーチタイプにより、多様なコーディング・タイプにコーディングされる。G . 7 1 8 コーデック及び予想された E V S 候

50

補コーデックのいずれでも、エンコーディングされたスピーチフレームは、FEC目的のために追加して分類される。それらフレームの分類は、スピーチフレームのシーケンスで、コーディング・タイプ及びスピーチフレームの位置に基づく。

【0141】

例えば、広帯域スピーチのために、下記表6に図示されたように、G.718コーデック及び予想されたEVS候補コーデックで、4個のコーディング・タイプが使用される。

【0142】

【表6】

表 6

Coding Type	Code	Comment
Unvoiced WB	0	For unvoiced speech frames
Voiced WB	2	For purely voiced speech frames
Generic WB	4	Non-stationary speech frames
Transition WB	6	Used for enhanced frame erasure performance by limiting use of past information

G.718コーデックによれば、コーディング・タイプ情報は、付加チャンネルを介して伝送される。付加チャンネルは、予想されたEVS候補コーデックで、現在使用可能ではない。付加チャンネルの不足を克服するために、G.718コーデックのアプローチと類似した付加情報は、前述のコンセプトと、表3で説明したコンセプトとを利用して、FECビットに伝送される。特定フレームの分類タイプが隣接したフレームの分類タイプに従属すれば、5個のコーディング・タイプは、既設定の個数のビットでシグナリングされる。本発明の一実施形態によれば、表7に図示されたコーディング・タイプが図示される。

【0143】

10

20

30

【表 7】

表 7

Frame Classification	Code	Comment
Unvoiced	0	Unvoiced, silence, noise, voiced offset
Unvoiced Transition	1	Transition from unvoiced to voiced components - possible onset, but too small
Voiced Transition	2	Transition from voiced - still voiced, but with very weak voiced characteristics
Voiced	3	Voiced frame, previous frame was also voiced or ONSET
Onset	4	Voiced onset sufficiently well built to follow with a voiced concealments

10

前述のように、図 6 に図示された多様なパケット構造は、知覚的な重要度を考慮して、多様な量のリダンダンシを有したスピーチフレームを伝送するために使用される。フレームの知覚的重要度は、表 6 に図示されたコーディング・タイプ、表 7 に図示されたフレーム分類または隣接したフレームで示されるあるアルゴリズムのうちいずれか一つから決定される。そして、フレームの知覚的重要度は、隣接したフレーム間に、リダンダンシ・ビットに係わる最適のトレードオフを決定することができる。

20

【0144】

本発明の一実施形態によれば、図 6 のアプローチ方式、表 6 のコーディング・タイプ及び表 7 のフレーム分類を考慮して、コーディング・タイプまたはフレーム分類に基づいて、使用される多様な量のリダンダンシを有したスピーチフレームを伝送するように、図 6 のパケット構造が制限される。本発明の一実施形態によれば、前記制限は、クラス A の個数は、クラス C の個数と同一である。

30

【0145】

かようなアプローチによって、リダンダンシを伝送するとき使用される 4 種のサブタイプが図 9 に図示される。

【0146】

図 9 は、本発明の一実施形態による、クラス A の個数と、クラス C の個数とが同一であるという制約に基づいて、リダンダンシを伝送するとき使用されるパケットの 4 種サブタイプを図示している。

【0147】

例えば、図 9 のパケットタイプ 1 は、図 6 のリダンダンシの伝送で使用されるように、同じパケット配列である。例えば、図 6 のパケット N についてエンコーディングされたソースビット A_n 、 B_n 、 C_n 、 A_{n-1} 、 B_{n-1} 及び A_{n-2} が使用される。

40

【0148】

図 10 は、本発明の一実施形態による、オンセット・フレームに、向上された保護を提供する多様なパケット・サブタイプを図示している。

【0149】

図 9 に図示された 4 種のパケット・サブタイプから、データパケット・サブタイプを選択することにより、エンコーディングされたスピーチフレームは、それぞれのフレームに係わる知覚的重要度により、さらに高いか、あるいはさらに低いリダンダンシ保護のために選択される。図 10 は、オンセット・フレーム（隣接したフレームのコストで）の向上

50

された保護 (enhanced protection) を提供するために、多様なパケット・サブタイプが使用される。

【 0 1 5 0 】

図 1 0 の例示で、パケット (N - 1) は、オンセット・フレームを含む。オンセット・フレームは、知覚的な観点で除去されるとき、最も敏感度が高いと知られたフレームを意味する。フレーム (n - 1) のリダンダンシ保護のために、パケット N 及びパケット (N + 1) が使用される。従って、パケット N は、サブタイプ 0 が選択され、パケット (N + 1) は、サブタイプ 3 が選択される。フレーム (n - 1) の向上されたリダンダンシ保護の結果が図示される。

【 0 1 5 1 】

図 1 0 で図示されたように、フレーム (n - 1) は、パケット (N - 1)、パケット N 及びパケット (N + 1) を介して、全体的に 3 回連続的に伝送される。増加された保護は、フレーム (n - 1) 及びフレーム n の保護に係わるコストとして示される。一般的に、フレーム (n - 1) がオンセットであるならば、フレーム (n - 2) は、相対的に低い保護が必要なアンボイスされたフレームである。本発明の一実施形態によれば、2 個のシグナリングビットを伝送するために、4 個のパケット・サブタイプが使用される。例えば、表 3 に図示されたように、それらのシグナリングビットは、クラス A に属する F E C ビットのように伝送される。

【 0 1 5 2 】

前述のように、図 2 A 及び図 2 B は、F E C アルゴリズムを介して、オーディオデータをエンコーディングまたはデコーディングすることができる一つ以上の端末 2 0 0 を含む。端末 2 0 0 は、図 1 のように、E P S コーデック及び/または E V S コーデック 2 6 で行われる。代替的な環境 (alternative environment) とコーデックは、同等に使用される。

【 0 1 5 3 】

さらに、本発明の一実施形態による図 2 B の端末 2 0 0 は、ソース端末、受信機端末、エンコーディング動作とデコーディング動作とを遂行することができる中間エンコーディング/デコーディング端末、デコーディング端末 1 5 0、またはネットワーク 1 4 0 によって提供された 2 個の端末間のネットワーク経路を含む。一つ以上の実施形態によれば、端末 2 0 0 は、異なるプロトコルで異なるネットワークタイプを介して、オーディオデータを受信したり伝送することができる。ここで、異なるネットワークタイプは、有線電話通信システム、セルラ電話またはデータ通信ネットワーク、あるいは無線携帯電話またはデータ通信ネットワークを含む。本発明の一実施形態によれば、端末 2 0 0 は、V O I P アプリケーション及びシステムを含むだけではないリアルタイム・ブロードキャスト、マルチキャスト・ブロードキャスト、及び時間遅延、保存またはストリーミングされたオーディオ・アプリケーション及びシステムを介した遠隔カンファレンス・アプリケーション及びシステムを含む。エンコーディングされたオーディオデータは、その後の再生のために記録され、ストリーミングされたブロードキャストまたは保存されたオーディオデータからデコーディングされる。

【 0 1 5 4 】

本発明の一実施形態によれば、一つ以上の端末 2 0 0 は、有線携帯電話、モバイルフォン、P D A、スマートフォン、タブレット・コンピュータ、セットトップボックス、ネットワーク端末、ラップトップ・コンピュータ、デスクトップ・コンピュータ、サーバ、ルータまたはゲートウェイを含む。端末 2 0 0 は、D S P (digital signal processor)、M C U (main control unit) または C P U のようなプロセッシング装置のうち少なくとも一つを含む。

【 0 1 5 5 】

本発明の一実施形態によれば、無線ネットワークは、ブルートゥース (登録商標 (Blue tooth)) または赤外線通信のような W P A N (wireless personal area network)、無線 L A N (local area network) (I E E E 8 0 2 . 1 1 と同様)、無線大都市ネ

10

20

30

40

50

ットワーク (wireless metropolitan area network)、802.16eのようなWiMaxネットワーク、802.16eのようなWiBroネットワーク、ネットワーク、GSM (登録商標 (global system for mobile communications))、PCS (personal communications service)、及びいかなる3GPPネットワークを含む。

【0156】

有線ネットワークは、地上基盤または衛星基盤の電話ネットワーク、ケーブルTV (television)、インターネット接続、光ファイバ通信、導波路、イーサネット (登録商標) 通信ネットワーク、ISDN (integrated services digital network)、DSL (digital subscriber line) ネットワーク、HDSL (high bit rate digital subscriber line) ネットワーク、SDSL (symmetric digital subscriber line) ネットワーク、ADSL (asymmetric digital subscriber line) ネットワーク、ILECs (local exchange carriers) と係わるRADSL (rate-adaptive digital subscriber line) ネットワーク、VDSL ネット、及びスイッチされたデジタルサービス (Non-P) 及びPOTSシステムを含む。

10

【0157】

ネットワーク140と通信することができるソース端末は、ネットワーク140と通信することができる受信端末と異なる。そして、オーディオデータは、オーディオソースと、オーディオ受信機140との経路を介して、特定ポイントで、端末及び2個以上の異なるネットワークを介して通信することができる。本発明の一実施形態によれば、オーディオデータのエンコーディング、伝送、保存及び/またはデコーディングは、FEC情報を有することができる。そして、オーディオデータは、伝送プロトコルに適するパケットで包まれる。

20

【0158】

伝送プロトコルは、RTPパケットまたはHTTPパケットを支援することができる。RTPパケットまたはHTTPパケットそれぞれは、少なくとも1つのヘッダ、コンテンツ・テーブル及びペイロードデータをそれぞれ有することができる。例えば、RTPパケットまたはHTTPパケットは、それぞれTCP protocol、UDP protocol、Cyclic UDP protocol、DCCP protocol、Fiber Channel Protocol、NetBIOS protocol、Reliable Datagram Protocol、RDP、SCTP protocol、SPX (sequenced packet exchange)、SST (structured stream transport)、VSP protocol、ATM (asynchronous transfer mode)、MTP/IP (multipurpose transaction protocol)、µTP (micro transport protocol)、及び/またはLTEでもある。

30

【0159】

本発明の一実施形態によれば、デコーディング端末150とエンコーディング端末100とのQoS (quality of service) 通信を含む。QoSは、RTPまたはオーディオデータ伝送経路から外れた経路を含むいかなる経路またはプロトコルを介しても伝送される。QoSは、データパケットに含まれたエラーチェック・コードに基づいて決定される。本発明の一実施形態によれば、QoSに基づいて、FECモードを変更することができる。そして、FECモードを適用することにより、コーディング・ビットレートとコーディング・モードを変更することができる。

40

【0160】

本発明の一実施形態によれば、FEC方式を適用するか否か、及び/またはいかなるFECモードを適用するかを決定するために、QoSを比較するための一つ以上の臨界値を使用することができる。それぞれの比較のための一つ以上の臨界値が存在する。そして、QoSが、特定臨界値 (Th1) より小さいか、あるいはそれと同じであるならば、臨界値は、FECモードがさらに信頼性があるか、低下されなければならないか、または増加されなければならないかを調節する必要があるか否かを示す。そして、QoSが、特定臨界値 (Th2) より大きいか、あるいはそれと同じであるならば、臨界値は、ビットレートとFECモードとが信頼性が不足しているか、低減されなければならないか、あるいは

50

増加されなければならないかを調節する必要があるか否かを示す。ここで、臨界値 $Th1$ と $Th2$ は、同一でもある。

【0161】

本発明の一実施形態によれば、エンコーディング端末100とデコーディング端末150は、FECアプローチを利用して、オーディオデータをコーディングするために使用されるオーディオ・コーデックを含む。オーディオ・コーディングは、LPC(LAR、LSP)、WLPC、CELP、ACELP、A-law、 μ -law、ADPCM、DPCM、MDCT、bit rate control(CBR、ABR、VBR)、及び/またはsub-bandコーディングを利用した一つ以上のアルゴリズムを使用することができる。そして、FECアプローチを利用するオーディオ・コーデックは、AMR、AMR-WB(G.722.2) 10、AMR-WB+、GSM-HR、GSM-FR、GSM-EFR、G.718及びEVSコーデックを含むいかなる3GPPコーデックをも含む。本発明の一実施形態で使用されるコーデックは、以前バージョンのコーデックと、逆に相互互換性を有することができる。

【0162】

エンコーディング端末100によって生成されたエンコーディングされたオーディオデータ・パケットは、エンコーダ側の一つ以上のコーデック120によってエンコーディングされたオーディオデータを含む。エンコーディングされたオーディオデータ・パケットは、エンコーダによってダウンミックスされたモノ信号であるSWB(super wideband audio)、エンコーダによってダウンミックスされたbinaural stereo audio data、 20フルバンド(FB)オーディオ及び/またはマルチチャネル・オーディオを含む。本発明の一実施形態によれば、エンコーディング過程は、同じであるか、あるいは異なるビットレートで、異なるタイプのオーディオデータをエンコーディングすることができる。本発明の一実施形態によれば、デコーディング端末150は、エンコーディングされたオーディオデータ・パケットと同様にパージングされる。

【0163】

従って、本発明の一実施形態によれば、端末200は、通信経路で制限された、マルチレート、多様なエンコーディングまたは翻訳(translation)を行うコーデックを含む。そして、端末200は、同じサンプリング・レートまたは異なるサンプリング・レートを有する多重レイヤ、または向上されたレイヤで、スケーラブル・コーディングを行うこと 30ができる。そして、デコーダは、ジッタバッファを含む。エンコーダ側面のコーデック120は、空間パラメータ推定、及びモノまたはバイナリのダウンミキシングを含む。前記リスティングされたオーディオ・コーデックのうち一つ以上は、一つ以上の異なるオーディオデータを生成することができる。そして、デコーダ側面のコーデック150は、推定されたパラメータのデコーディングに基づいて、対応するコーデック、モノまたはバイナリのアップミキシング及び空間レンダリングを含む。

【0164】

本発明の一実施形態によれば、ある装置、システム及びユニットの説明は、一つ以上のハードウェア装置またはハードウェア・プロセッシング要素を含む。例えば、本発明の一実施形態で、説明された装置、システム及びユニットは、追加してメモリ、ハードウェア入出力伝送装置を含む。そして、装置は、物理的なシステムの構成要素と同意関係にあると見なされる。しかし、装置は、1つのデバイスに制限されたり、あるいは限定解釈されるものではない。そして、全ての説明された構成要素は、1つのそれぞれの保護範囲内に含まれもする。 40

【0165】

本発明の実施形態による方法は、多様なコンピュータ手段を介して遂行されるプログラム命令形態に具現され、コンピュータ可読媒体に記録される。前記コンピュータ可読媒体は、プログラム命令、データファイル、データ構造などを単独または組み合わせで含む。前記媒体に記録されるプログラム命令は、本発明のために特別に設計されて構成されたものであるか、コンピュータ・ソフトウェア当業者に公知されて使用可能なものでもある。 50

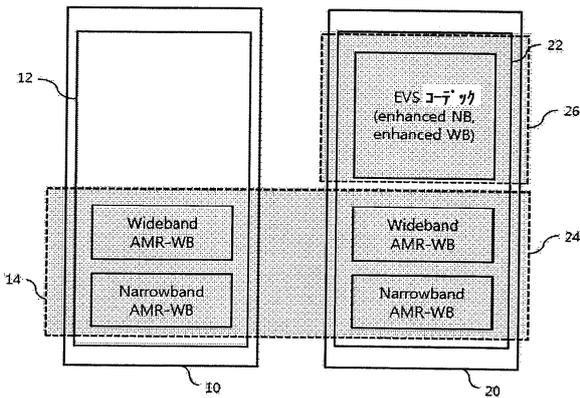
【0166】

以上のように本発明は、たとえ限定された実施形態及び図面によって説明されたにしても、本発明は、前記の実施形態に限定されるものではなく、本発明が属する分野で当業者であるならば、かような記載から多様な修正及び変形が可能であろう。

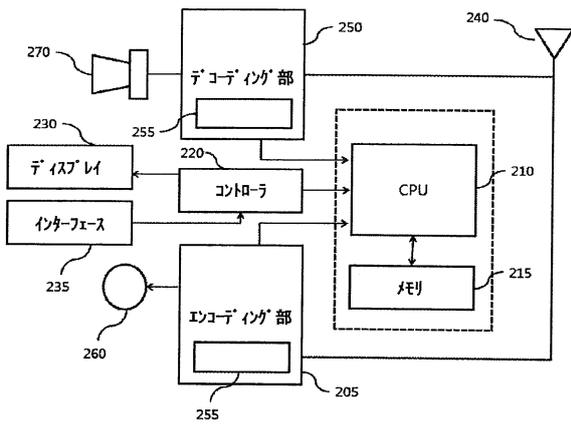
【0167】

従って、本発明の範囲は、説明された実施形態に限って決められるものではなく、特許請求の範囲だけではなく、当該特許請求の範囲と均等なものなどによっても決められるものである。

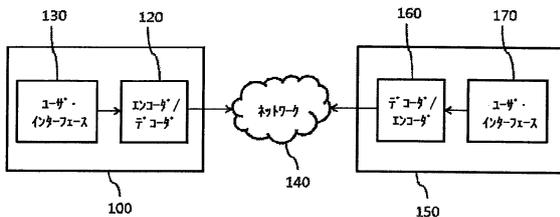
【図1】



【図2B】



【図2A】



【図3】

FIG. 3

Normal Mode	High FER Mode	High FER Mode
TBS=328	TBS=472	TBS=472
RTP Payload 74 Bits	RTP Payload 74 Bits	RTP Payload 74 Bits
Frame N EVS Encoded Source Bits (253)	Frame N-1 FEC Bits (118)	Frame N FEC Bits (118)
Frame N	Frame N EVS Encoded Source Bits (253)	Frame N+1 EVS Encoded Source Bits (253)
Frame N	Frame N	Frame N+1

【 4 】

FIG. 4

TBS=392	TBS=392	TBS=392
FEC:32 Bits from Frame N	FEC:32 Bits from Frame N-1	FEC:32 Bits from Frame N-2
FEC:32 Bits from Frame N+1	FEC:32 Bits from Frame N	FEC:32 Bits from Frame N-1
RTP Payload 74 Bits	RTP Payload 74 Bits	RTP Payload 74 Bits
EVS Encoded Source Bits 253 (+1)	EVS Encoded Source Bits 253 (+1)	EVS Encoded Source Bits 253 (+1)
Packet N+2	Packet N+1	Packet N

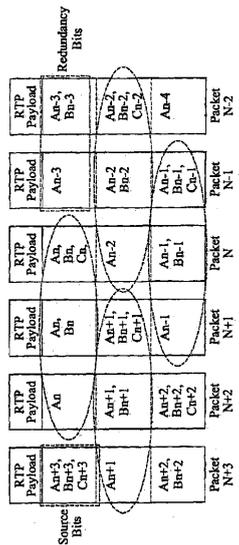
【 5 】

FIG. 5

TBS=392	TBS=392	TBS=392
FEC:32 Bits from Frame N+2	FEC:32 Bits from Frame N+1	FEC:32 Bits from Frame N
FEC:32 Bits from Frame N	FEC:32 Bits from Frame N-1	FEC:32 Bits from Frame N-2
RTP Payload 74 Bits	RTP Payload 74 Bits	RTP Payload 74 Bits
EVS Encoded Source Bits 253 (+1)	EVS Encoded Source Bits 253 (+1)	EVS Encoded Source Bits 253 (+1)
Packet N+1	Packet N	Packet N-1

【 6 】

FIG. 6



【 7 】

FIG. 7

Normal Mode 12.65Kbps	High FER Submode1 7.3Mbps	High FER Submode2 11.1Mbps	High FER Submode3 14.25Mbps	High FER Submode4 18.25Mbps	High FER Submode5 22.25Mbps
TBS=228	TBS=224	TBS=296	TBS=328	TBS=376	TBS=440
RTP Payload 74 Bits	RTP Payload 74 Bits	RTP Payload 74 Bits	RTP Payload 74 Bits	RTP Payload 74 Bits	RTP Payload 74 Bits
EVS 12.65K mode (233) bits	EVS 6.6K mode (146 bits)	EVS 6.6K mode (222 bits)	EVS 6.6K mode (254 bits)	EVS 6.6K mode (316 bits)	EVS 12.65K mode (366 bits)
Packet N	Packet N	Packet N	Packet N	Packet N	Packet N

【 8 】

FIG. 8

Normal Mode 12.65Kbps	High FER submode 1	High FER submode 2	High FER submode 3	High FER submode 4
TBS=328	TBS=328	TBS=328	TBS=328	TBS=328
RTP Payload 74 Bits	RTP Payload 74 Bits	RTP Payload 74 Bits	RTP Payload 74 Bits	RTP Payload 74 Bits
EVS 12.65K mode (253) bits	EVS 11.1K mode (222) bits	EVS 8.85K mode (177) bits	EVS 7.5K mode (150) bits	EVS 6.6K mode (132) bits
Packet N	Packet N	Packet N	Packet N	Packet N

【 9 】

FIG. 9

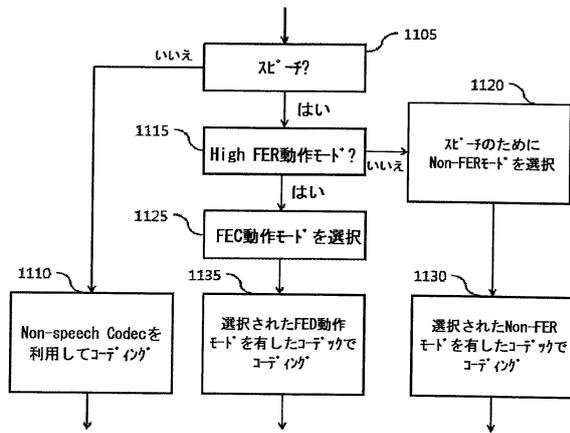
	Packet Type 0	Packet Type 1	Packet Type 2	Packet Type 3
Source	A _n	A _n	A _n	A _n
	B _n	B _n	B _n	B _n
	C _n	C _n	C _n	C _n
Redundancy	A _{n-1}	A _{n-1}	A _{n-1}	A _{n-2}
	B _{n-1}	B _{n-1}	B _{n-2}	B _{n-2}
	C _{n-1}	A _{n-2}	B _{n-2}	C _{n-2}

【図10】

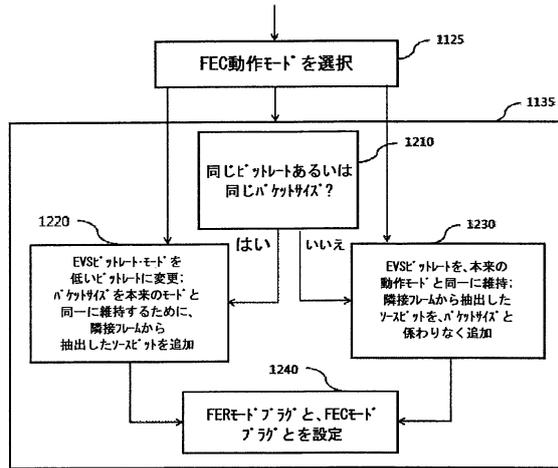
FIG. 10

	Packet N+2	Packet N+1 (type 3)	Packet N (type 0)	Packet N-1 (onset)	Packet N-2
Source	A _{N+2}	A _{N+1}	A _N	A _{N-1}	A _{N-2}
	B _{N+2}	B _{N+1}	B _N	B _{N-1}	B _{N-2}
	C _{N+2}	C _{N+1}	C _N	C _{N-1}	C _{N-2}
Redundancy		A _{N+1} B _{N+1} C _{N+1}	A _{N+2} B _{N+2} C _{N+2}		

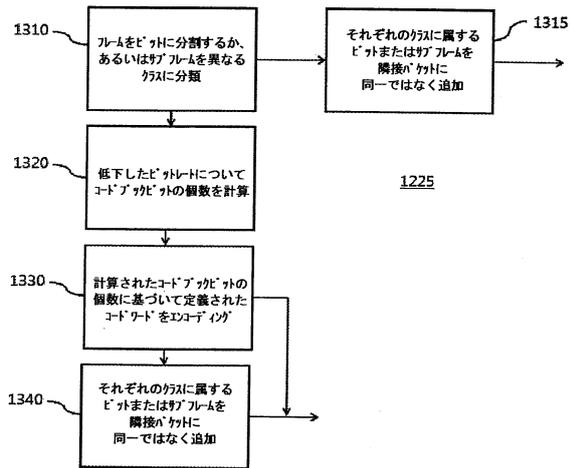
【図11】



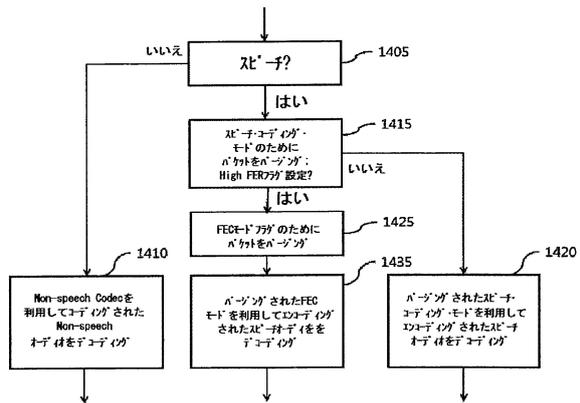
【図12】



【図13】



【図14】



フロントページの続き

(72)発明者 ソン, ホ - サン
大韓民国 448-785 キョンギ - ド ヨンイン - シ スジ - グ プンドックチョン 2 - ド
ン サムスン5 - チャ・アパート 521-1104 (番地なし)

(72)発明者 グリアー, スティーヴン クレイグ
アメリカ合衆国 テキサス 75089 ロウレット トーラー・ロード 3201

合議体

審判長 井上 信一

審判官 國分 直樹

審判官 関谷 隆一

(56)参考文献 米国特許出願公開第2005/0228651 (US, A1)
特開2002-141810 (JP, A)
米国特許出願公開第2001/0041981 (US, A1)
Johansson et al., Bandwidth efficient AMR operation for VoIP, IEEE Workshop Proceedings on Speech Coding 2002, IEEE, 2002年10月6日, pp. 150-152

(58)調査した分野(Int.Cl., DB名)

G10L19/005

H04L1/00