

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2013-242579

(P2013-242579A)

(43) 公開日 平成25年12月5日(2013.12.5)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 O L 19/13 (2013.01)	G 1 O L 19/12 B	
G 1 O L 19/06 (2013.01)	G 1 O L 19/06 Z	
G 1 O L 19/00 (2013.01)	G 1 O L 19/00 4 O O A	

審査請求 有 請求項の数 39 O L 外国語出願 (全 79 頁)

(21) 出願番号	特願2013-141575 (P2013-141575)	(71) 出願人	595020643 クアアルコム・インコーポレイテッド QUALCOMM INCORPORATED
(22) 出願日	平成25年7月5日(2013.7.5)		
(62) 分割の表示	特願2010-512371 (P2010-512371) の分割		
原出願日	平成20年6月13日(2008.6.13)		
(31) 優先権主張番号	60/943,558	(74) 代理人	100108855 弁理士 蔵田 昌俊
(32) 優先日	平成19年6月13日(2007.6.13)		
(33) 優先権主張国	米国 (US)	(74) 代理人	100109830 弁理士 福原 淑弘
(31) 優先権主張番号	12/137,700		
(32) 優先日	平成20年6月12日(2008.6.12)	(74) 代理人	100088683 弁理士 中村 誠
(33) 優先権主張国	米国 (US)		
		(74) 代理人	100103034 弁理士 野河 信久

最終頁に続く

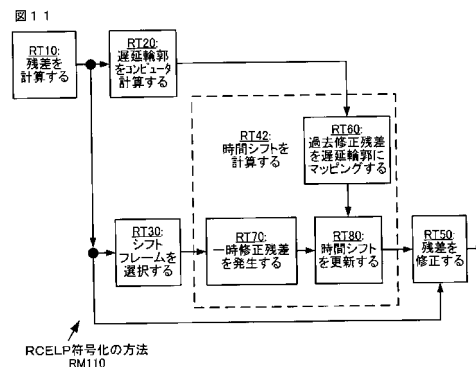
(54) 【発明の名称】 ピッチ調整コーディング及び非ピッチ調整コーディングを使用する信号符号化

(57) 【要約】 (修正有)

【課題】 R C E L P コーディング方式はピッチ調整 (pitch-regularizing) (「P R」) コーディング方式であり、信号のピッチ周期間の変動が、一般に、ピッチパルスの相対位置をより滑らかな合成遅延輪郭に一致または近似するように変化させる。

【解決手段】 オーディオ信号のフレームのピッチ調整 (P R) 符号化が非 P R 符号化中に別のフレームのセグメントに後続し且つ連続するために、時間修正を時間シフトまたはタイムワープが使用される。

【選択図】 図 1 1



【特許請求の範囲】

【請求項 1】

オーディオ信号のフレームを処理する方法であって、
 ピッチ調整（PR）コーディング方式に従って前記オーディオ信号の第 1 のフレームを符号化することと、

非 PR コーディング方式に従って前記オーディオ信号の第 2 のフレームを符号化することと、

を備え、

前記第 2 のフレームが、前記オーディオ信号中の前記第 1 のフレームに後続し且つ連続し、

前記第 1 のフレームを符号化することが、前記第 1 のフレームに基づく第 1 の信号のセグメントを、時間シフトに基づいて時間修正することを含み、前記時間修正することが、（A）前記時間シフトに従って前記第 1 のフレームの前記セグメントを時間シフトすることと、（B）前記時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを含み、

前記第 1 の信号のセグメントを時間修正することが、前記第 1 の信号の別のピッチパルスに対して前記セグメントのピッチパルスの位置を変化させることを含み、

前記第 2 のフレームを符号化することが、前記第 2 のフレームに基づく第 2 の信号のセグメントを、前記時間シフトに基づいて時間修正することを含み、前記時間修正することが、（A）前記時間シフトに従って前記第 2 のフレームの前記セグメントを時間シフトすることと、（B）前記時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを含む、方法。

【請求項 2】

前記第 1 のフレームを符号化することが、前記第 1 の信号の前記時間修正されたセグメントに基づく第 1 の符号化フレームを生成することを含み、

前記第 2 のフレームを符号化することが、前記第 2 の信号の前記時間修正されたセグメントに基づく第 2 の符号化フレームを生成することを含む、請求項 1 に記載の方法。

【請求項 3】

前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、請求項 1 に記載の方法。

【請求項 4】

前記第 1 及び第 2 の信号が重み付きオーディオ信号である、請求項 1 に記載の方法。

【請求項 5】

前記第 1 のフレームを符号化することが、前記オーディオ信号中の前記第 1 のフレームに先行する第 3 のフレームの残差からの情報に基づいて前記時間シフトを計算することを含む、請求項 1 に記載の方法。

【請求項 6】

前記時間シフトを計算することが、前記オーディオ信号の遅延輪郭に前記第 3 のフレームの前記残差のサンプルをマッピングすることを含む、請求項 5 に記載の方法。

【請求項 7】

前記第 1 のフレームを符号化することが、前記オーディオ信号のピッチ周期に関する情報に基づいて前記遅延輪郭をコンピュータ計算することを含む、請求項 6 に記載の方法。

【請求項 8】

前記 PR コーディング方式がリラクストコード励振線形予測コーディング方式であり、

前記非 PR コーディング方式が、（A）ノイズ励起線形予測コーディング方式と、（B）修正離散コサイン変換コーディング方式と、（C）プロトタイプ波形補間コーディング方式と、のうちの 1 つである、請求項 1 に記載の方法。

【請求項 9】

前記非 PR コーディング方式が修正離散コサイン変換コーディング方式である、請求項

10

20

30

40

50

1 に記載の方法。

【請求項 1 0】

前記第 2 のフレームを符号化することが、
 符号化残差を得るために前記第 2 のフレームの残差に対して修正離散コサイン変換 (M D C T) 演算を実行することと、
 復号残差を得るために前記符号化残差に基づく信号に対して逆 M D C T 演算を実行することと、
 を含み、
 前記第 2 の信号が前記復号残差に基づく、請求項 1 に記載の方法。

【請求項 1 1】

前記第 2 のフレームを符号化することが、
 前記第 2 の信号である、前記第 2 のフレームの残差を発生させることと、
 前記第 2 の信号のセグメントを時間修正することについて、符号化残差を得るために、
 前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換演算を実行することと、
 前記符号化残差に基づいて第 2 の符号化フレームを生成することと、
 を含む、請求項 1 に記載の方法。

【請求項 1 2】

前記方法が、前記オーディオ信号中の前記第 2 のフレームに後続するフレームの残差のセグメントを、前記時間シフトに従って時間シフトすることを備える、請求項 1 に記載の方法。

【請求項 1 3】

前記方法が、前記第 2 のフレームに後続する前記オーディオ信号の第 3 のフレームに基づく第 3 の信号のセグメントを、前記時間シフトに基づいて時間修正することを含み、
 前記第 2 のフレームを符号化することが、前記第 2 及び第 3 の信号の前記時間修正されたセグメントのサンプルを含むウィンドウに対して修正離散コサイン変換 (M D C T) 演算を実行することを含む、請求項 1 に記載の方法。

【請求項 1 4】

前記第 2 の信号が M 個のサンプルの長さを有し、前記第 3 の信号が M 個のサンプルの長さを有し、
 前記 M D C T 演算を実行することが、(A) 前記時間修正されたセグメントを含む、前記第 2 の信号の M 個のサンプルと、(B) 前記第 3 の信号の $3 M / 4$ 個以下のサンプルと、
 に基づく M 個の M D C T 係数の組を生成することを含む、請求項 1 3 に記載の方法。

【請求項 1 5】

前記第 2 の信号が M 個のサンプルの長さを有し、前記第 3 の信号が M 個のサンプルの長さを有し、
 前記 M D C T 演算を実行することが、(A) 前記時間修正されたセグメントを含む、前記第 2 の信号の M 個のサンプルを含み、(B) ゼロ値の少なくとも M / 8 個のサンプルのシーケンスで始まり、(C) ゼロ値の少なくとも M / 8 個のサンプルのシーケンスで終わる、
 2 M 個のサンプルのシーケンスに基づく M 個の M D C T 係数の組を生成することを含む、請求項 1 3 に記載の方法。

【請求項 1 6】

オーディオ信号のフレームを処理するための装置であって、
 ピッチ調整 (P R) コーディング方式に従って前記オーディオ信号の第 1 のフレームを符号化するための手段と、
 非 P R コーディング方式に従って前記オーディオ信号の第 2 のフレームを符号化するための手段と、
 を備え、
 前記第 2 のフレームが、前記オーディオ信号中の前記第 1 のフレームに後続し且つ連続し、

10

20

30

40

50

前記第 1 のフレームを符号化するための手段が、前記第 1 のフレームに基づく第 1 の信号のセグメントを、時間シフトに基づいて時間修正するための手段を含み、前記時間修正するための手段が、(A) 前記時間シフトに従って前記第 1 のフレームの前記セグメントを時間シフトすることと、(B) 前記時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され、

前記第 1 の信号のセグメントを時間修正するための手段が、前記第 1 の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させるように構成され、

前記第 2 のフレームを符号化するための手段が、前記第 2 のフレームに基づく第 2 の信号のセグメントを、前記時間シフトに基づいて時間修正するための手段を含み、前記時間修正するための手段が、(A) 前記時間シフトに従って前記第 2 のフレームの前記セグメントを時間シフトすることと、(B) 前記時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成される、装置。

10

【請求項 17】

前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、請求項 16 に記載の装置。

【請求項 18】

前記第 1 及び第 2 の信号が重み付きオーディオ信号である、請求項 16 に記載の装置。

【請求項 19】

前記第 1 のフレームを符号化するための手段が、前記オーディオ信号中の前記第 1 のフレームに先行する第 3 のフレームの残差からの情報に基づいて前記時間シフトを計算するための手段を含む、請求項 16 に記載の装置。

20

【請求項 20】

前記第 2 のフレームを符号化するための手段が、
前記第 2 の信号である、前記第 2 のフレームの残差を発生させるための手段と、
符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換演算を実行するための手段と、
を含み、

前記第 2 のフレームを符号化するための手段が、前記符号化残差に基づいて第 2 の符号化フレームを生成するように構成される、請求項 16 に記載の装置。

【請求項 21】

30

前記第 2 の信号のセグメントを時間修正するための手段が、前記オーディオ信号中の前記第 2 のフレームに後続するフレームの残差のセグメントを、前記時間シフトに従って時間シフトするように構成される、請求項 16 に記載の装置。

【請求項 22】

前記第 2 の信号のセグメントを時間修正するための手段が、前記第 2 のフレームに後続する前記オーディオ信号の第 3 のフレームに基づく第 3 の信号のセグメントを、前記時間シフトに基づいて時間修正するように構成され、

前記第 2 のフレームを符号化するための手段が、前記第 2 及び第 3 の信号の前記時間修正されたセグメントのサンプルを含むウィンドウに対して修正離散コサイン変換(MDCT)演算を実行するための手段を含む、請求項 16 に記載の装置。

40

【請求項 23】

前記第 2 の信号が M 個のサンプルの長さを有し、前記第 3 の信号が M 個のサンプルの長さを有し、

前記 MDCT 演算を実行するための手段が、(A) 前記時間修正されたセグメントを含む、前記第 2 の信号の M 個のサンプルと、(B) 前記第 3 の信号の $3M/4$ 個以下のサンプルと、に基づく M 個の MDCT 係数の組を生成するように構成される、請求項 22 に記載の装置。

【請求項 24】

オーディオ信号のフレームを処理するための装置であって、

ピッチ調整(PR)コーディング方式に従って前記オーディオ信号の第 1 のフレームを

50

符号化するように構成された第 1 のフレームエンコーダと、

非 P R コーディング方式に従って前記オーディオ信号の第 2 のフレームを符号化するように構成された第 2 のフレームエンコーダと、

を備え、

前記第 2 のフレームが、前記オーディオ信号中の前記第 1 のフレームに後続し且つ連続し、

前記第 1 のフレームエンコーダが、前記第 1 のフレームに基づく第 1 の信号のセグメントを、時間シフトに基づいて時間修正するように構成された第 1 の時間修正器を含み、前記第 1 の時間修正器が、(A) 前記時間シフトに従って前記第 1 のフレームの前記セグメントを時間シフトすることと、(B) 前記時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され、

前記第 1 の時間修正器が、前記第 1 の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させるように構成され、

前記第 2 のフレームエンコーダが、前記第 2 のフレームに基づく第 2 の信号のセグメントを、前記時間シフトに基づいて時間修正するように構成された第 2 の時間修正器を含み、前記第 2 の時間修正器が、(A) 前記時間シフトに従って前記第 2 のフレームの前記セグメントを時間シフトすることと、(B) 前記時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成される、装置。

【請求項 25】

前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、請求項 24 に記載の装置。

【請求項 26】

前記第 1 及び第 2 の信号が重み付きオーディオ信号である、請求項 24 に記載の装置。

【請求項 27】

前記第 1 のフレームエンコーダが、前記オーディオ信号中の前記第 1 のフレームに先行する第 3 のフレームの残差からの情報に基づいて前記時間シフトを計算するように構成された時間シフト計算器を含む、請求項 24 に記載の装置。

【請求項 28】

前記第 2 のフレームエンコーダが、

前記第 2 の信号である、前記第 2 のフレームの残差を発生させるように構成された残差発生器と、

符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換 (M D C T) 演算を実行するように構成された M D C T モジュールと、

を含み、

前記第 2 のフレームエンコーダが、前記符号化残差に基づいて第 2 の符号化フレームを生成するように構成される、請求項 24 に記載の装置。

【請求項 29】

前記第 2 の時間修正器が、前記オーディオ信号中の前記第 2 のフレームに後続するフレームの残差のセグメントを、前記時間シフトに従って時間シフトするように構成される、請求項 24 に記載の装置。

【請求項 30】

前記第 2 の時間修正器が、前記第 2 のフレームに後続する前記オーディオ信号の第 3 のフレームに基づく第 3 の信号のセグメントを、前記時間シフトに基づいて時間修正するように構成され、

前記第 2 のフレームエンコーダが、前記第 2 及び第 3 の信号の前記時間修正されたセグメントのサンプルを含むウィンドウに対して修正離散コサイン変換 (M D C T) 演算を実行するように構成された M D C T モジュールを含む、請求項 24 に記載の装置。

【請求項 31】

前記第 2 の信号が M 個のサンプルの長さを有し、前記第 3 の信号が M 個のサンプルの長さを有し、

前記 M D C T モジュールが、(A) 前記時間修正されたセグメントを含む、前記第 2 の信号の M 個のサンプルと、(B) 前記第 3 の信号の $3 M / 4$ 個以下のサンプルと、に基づく M 個の M D C T 係数の組を生成するように構成される、請求項 3 0 に記載の装置。

【請求項 3 2】

プロセッサによって実行されると前記プロセッサに、

ピッチ調整 (P R) コーディング方式に従って前記オーディオ信号の第 1 のフレームを符号化することと、

非 P R コーディング方式に従って前記オーディオ信号の第 2 のフレームを符号化することと、

を行わせる命令を備えるコンピュータ可読媒体であって、

前記第 2 のフレームが、前記オーディオ信号中の前記第 1 のフレームに後続し且つ連続し、

前記実行されると前記プロセッサに第 1 のフレームを符号化することを行わせる命令が、前記第 1 のフレームに基づく第 1 の信号のセグメントを、時間シフトに基づいて時間修正する命令を含み、前記時間修正する命令が、(A) 前記時間シフトに従って前記第 1 のフレームの前記セグメントを時間シフトする命令と、(B) 前記時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープする命令と、のうちの 1 つを含み、

前記第 1 の信号のセグメントを時間修正する命令が、前記第 1 の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させる命令を含み、

前記実行されると前記プロセッサに第 2 のフレームを符号化することを行わせる命令が、前記第 2 のフレームに基づく第 2 の信号のセグメントを、前記時間シフトに基づいて時間修正する命令を含み、前記時間修正する命令が、(A) 前記時間シフトに従って前記第 2 のフレームの前記セグメントを時間シフトする命令と、(B) 前記時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープする命令と、のうちの 1 つを含む、コンピュータ可読媒体。

【請求項 3 3】

オーディオ信号のフレームを処理する方法であって、

第 1 のコーディング方式に従って前記オーディオ信号の第 1 のフレームを符号化することと、

ピッチ調整 (P R) コーディング方式に従って前記オーディオ信号の第 2 のフレームを符号化することと、

を備え、

前記第 2 のフレームが、前記オーディオ信号中の前記第 1 のフレームに後続し且つ連続し、

前記第 1 のコーディング方式が非 P R コーディング方式であり、

前記第 1 のフレームを符号化することが、前記第 1 のフレームに基づく第 1 の信号のセグメントを、第 1 の時間シフトに基づいて時間修正することを含み、前記時間修正することが、(A) 前記第 1 の時間シフトに従って前記第 1 の信号の前記セグメントを時間シフトすることと、(B) 前記第 1 の時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを含み、

前記第 2 のフレームを符号化することが、前記第 2 のフレームに基づく第 2 の信号のセグメントを、第 2 の時間シフトに基づいて時間修正することを含み、前記時間修正することが、(A) 前記第 2 の時間シフトに従って前記第 2 の信号の前記セグメントを時間シフトすることと、(B) 前記第 2 の時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを含み、

前記第 2 の信号のセグメントを時間修正することが、前記第 2 の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させることを含み、

前記第 2 の時間シフトが、前記第 1 の信号の前記時間修正されたセグメントからの情報

に基づく、方法。

【請求項 34】

前記第 1 のフレームを符号化することが、前記第 1 の信号の前記時間修正されたセグメントに基づく第 1 の符号化フレームを生成することを含み、

前記第 2 のフレームを符号化することが、前記第 2 の信号の前記時間修正されたセグメントに基づく第 2 の符号化フレームを生成することを含む、請求項 33 に記載の方法。

【請求項 35】

前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、請求項 33 に記載の方法。

【請求項 36】

前記第 1 及び第 2 の信号が重み付きオーディオ信号である、請求項 33 に記載の方法。

【請求項 37】

前記第 2 の信号のセグメントを時間修正することが、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づいて前記第 2 の時間シフトを計算することを含み、

前記第 2 の時間シフトを計算することが、前記第 2 のフレームからの情報に基づく遅延輪郭に、前記第 1 の信号の前記時間修正されたセグメントをマッピングすることを含む、請求項 33 に記載の方法。

【請求項 38】

前記第 2 の時間シフトが、前記マッピングされたセグメントのサンプルと一時修正残差のサンプルとの間の相関に基づかれ、

前記一時修正残差が、(A) 前記第 2 のフレームの残差のサンプルと、(B) 前記第 1 の時間シフトと、に基づかれる、請求項 37 に記載の方法。

【請求項 39】

前記第 2 の信号が前記第 2 のフレームの残差であり、

前記第 2 の信号のセグメントを時間修正することが、前記第 2 の時間シフトに従って前記残差の第 1 のセグメントを時間シフトすることを含み、

前記方法が、

前記第 1 の信号の前記時間修正されたセグメントからの情報に基づいて、前記第 2 の時間シフトとは異なる第 3 の時間シフトを計算することと、

前記第 3 の時間シフトに従って前記残差の第 2 のセグメントを時間シフトすることと

を備える、請求項 33 に記載の方法。

【請求項 40】

前記第 2 の信号が前記第 2 のフレームの残差であり、

前記第 2 の信号のセグメントを時間修正することが、前記第 2 の時間シフトに従って前記残差の第 1 のセグメントを時間シフトすることを含み、

前記方法が、

前記残差の前記時間修正された第 1 のセグメントからの情報に基づいて、前記第 2 の時間シフトとは異なる第 3 の時間シフトを計算することと、

前記第 3 の時間シフトに従って前記残差の第 2 のセグメントを時間シフトすることと

を備える、請求項 33 に記載の方法。

【請求項 41】

前記第 2 の信号のセグメントを時間修正することが、前記第 2 のフレームからの情報に基づく遅延輪郭に、前記第 1 の信号の前記時間修正されたセグメントのサンプルをマッピングすることを含む、請求項 33 に記載の方法。

【請求項 42】

前記方法が、

適応コードブックバッファに前記第 1 の信号の前記時間修正されたセグメントに基づくシーケンスを記憶することと、

10

20

30

40

50

前記記憶することに続いて、前記第 2 のフレームからの情報に基づく遅延輪郭に前記適応コードブックバッファのサンプルをマッピングすることと、
を備える、請求項 3 3 に記載の方法。

【請求項 4 3】

前記第 2 の信号が前記第 2 のフレームの残差であり、前記第 2 の信号のセグメントを時間修正することが前記第 2 のフレームの前記残差をタイムワープすることを含み、

前記方法が、前記第 2 のフレームの前記タイムワープされた残差からの情報に基づいて前記オーディオ信号の第 3 のフレームの残差をタイムワープすることを備え、前記第 3 のフレームが前記オーディオ信号中の前記第 2 のフレームに連続する、請求項 3 3 に記載の方法。

10

【請求項 4 4】

前記第 2 の信号が前記第 2 のフレームの残差であり、前記第 2 の信号のセグメントを時間修正することが、(A) 前記第 1 の信号の前記時間修正されたセグメントからの情報と、(B) 前記第 2 のフレームの前記残差からの情報と、に基づいて前記第 2 の時間シフトを計算することを含む、請求項 3 3 に記載の方法。

【請求項 4 5】

前記 P R コーディング方式がリラクストコード励振線形予測コーディング方式であり、前記非 P R コーディング方式が、(A) ノイズ励起線形予測コーディング方式と、(B) 修正離散コサイン変換コーディング方式と、(C) プロトタイプ波形補間コーディング方式と、のうちの 1 つである、請求項 3 3 に記載の方法。

20

【請求項 4 6】

前記非 P R コーディング方式が修正離散コサイン変換コーディング方式である、請求項 3 3 に記載の方法。

【請求項 4 7】

前記第 1 のフレームを符号化することが、

符号化残差を得るために前記第 1 のフレームの残差に対して修正離散コサイン変換 (M D C T) 演算を実行することと、

復号残差を得るために前記符号化残差に基づく信号に対して逆 M D C T 演算を実行することと、

を含み、

前記第 1 の信号が前記復号残差に基づく、請求項 3 3 に記載の方法。

30

【請求項 4 8】

前記第 1 のフレームを符号化することが、

前記第 1 の信号である、前記第 1 のフレームの残差を発生させることと、

前記第 1 の信号のセグメントを時間修正することに続いて、符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換演算を実行することと、

前記符号化残差に基づいて第 1 の符号化フレームを生成することと、

を含む、請求項 3 3 に記載の方法。

【請求項 4 9】

40

前記第 1 の信号が M 個のサンプルの長さを有し、前記第 2 の信号が M 個のサンプルの長さを有し、

前記第 1 のフレームを符号化することが、前記時間修正されたセグメントを含む、前記第 1 の信号の M 個のサンプルと、前記第 2 の信号の $3M/4$ 個以下のサンプルと、に基づく M 個の修正離散コサイン変換 (M D C T) 係数の組を生成することを含む、請求項 3 3 に記載の方法。

【請求項 5 0】

前記第 1 の信号が M 個のサンプルの長さを有し、前記第 2 の信号が M 個のサンプルの長さを有し、

前記第 1 のフレームを符号化することが、(A) 前記時間修正されたセグメントを含む

50

、前記第 1 の信号の M 個のサンプルを含み、(B) ゼロ値の少なくとも M / 8 個のサンプルのシーケンスで始まり、(C) ゼロ値の少なくとも M / 8 個のサンプルのシーケンスで終わる、2 M 個のサンプルのシーケンスに基づく M 個の修正離散コサイン変換 (M D C T) 係数の組を生成することを含む、請求項 3 3 に記載の方法。

【請求項 5 1】

オーディオ信号のフレームを処理するための装置であって、

第 1 のコーディング方式に従って前記オーディオ信号の第 1 のフレームを符号化するための手段と、

ピッチ調整 (P R) コーディング方式に従って前記オーディオ信号の第 2 のフレームを符号化するための手段と、

を備え、

前記第 2 のフレームが、前記オーディオ信号中の前記第 1 のフレームに後続し且つ連続し、

前記第 1 のコーディング方式が非 P R コーディング方式であり、

第 1 のフレームを符号化するための前記手段が、前記第 1 のフレームに基づく第 1 の信号のセグメントを、第 1 の時間シフトに基づいて時間修正するための手段を含み、前記時間修正するための手段が、(A) 前記第 1 の時間シフトに従って前記第 1 の信号の前記セグメントを時間シフトすることと、(B) 前記第 1 の時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され

、前記第 2 のフレームを符号化するための手段が、前記第 2 のフレームに基づく第 2 の信号のセグメントを、第 2 の時間シフトに基づいて時間修正するための手段を含み、前記時間修正するための手段が、(A) 前記第 2 の時間シフトに従って前記第 2 の信号の前記セグメントを時間シフトすることと、(B) 前記第 2 の時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され

、前記第 2 の信号のセグメントを時間修正するための手段が、前記第 2 の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させるように構成され、

前記第 2 の時間シフトが、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づく、方法。

【請求項 5 2】

前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、請求項 5 1 に記載の装置。

【請求項 5 3】

前記第 1 及び第 2 の信号が重み付きオーディオ信号である、請求項 5 1 に記載の装置。

【請求項 5 4】

前記第 2 の信号のセグメントを時間修正するための手段が、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づいて前記第 2 の時間シフトを計算するための手段を含み、

前記第 2 の時間シフトを計算するための手段が、前記第 2 のフレームからの情報に基づく遅延輪郭に、前記第 1 の信号の前記時間修正されたセグメントをマッピングするための手段を含む、請求項 5 1 に記載の装置。

【請求項 5 5】

前記第 2 の時間シフトが、前記マッピングされたセグメントのサンプルと一時修正残差のサンプルとの間の相関に基づき、

前記一時修正残差が、(A) 前記第 2 のフレームの残差のサンプルと、(B) 前記第 1 の時間シフトと、に基づく、請求項 5 4 に記載の装置。

【請求項 5 6】

前記第 2 の信号が前記第 2 のフレームの残差であり、

前記第 2 の信号のセグメントを時間修正するための手段が、前記第 2 の時間シフトに従

10

20

30

40

50

って前記残差の第 1 のセグメントを時間シフトするように構成され、
前記方法が、

前記残差の前記時間修正された第 1 のセグメントからの情報に基づいて、前記第 2 の時間シフトとは異なる第 3 の時間シフトを計算するための手段と、

前記第 3 の時間シフトに従って前記残差の第 2 のセグメントを時間シフトするための手段と、

を備える、請求項 5 1 に記載の装置。

【請求項 5 7】

前記第 2 の信号が前記第 2 のフレームの残差であり、第 2 の信号のセグメントを時間修正するための手段が、(A) 前記第 1 の信号の前記時間修正されたセグメントからの情報と、(B) 前記第 2 のフレームの前記残差からの情報と、に基づいて前記第 2 の時間シフトを計算するための手段を含む、請求項 5 1 に記載の装置。

10

【請求項 5 8】

前記第 1 のフレームを符号化するための手段が、

前記第 1 の信号である、前記第 1 のフレームの残差を発生させるための手段と、

符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換演算を実行するための手段と、

を含み、

前記第 1 のフレームを符号化するための手段が、前記符号化残差に基づいて第 1 の符号化フレームを生成するように構成される、請求項 5 1 に記載の装置。

20

【請求項 5 9】

前記第 1 の信号が M 個のサンプルの長さを有し、前記第 2 の信号が M 個のサンプルの長さを有し、

前記第 1 のフレームを符号化するための手段が、前記時間修正されたセグメントを含む、前記第 1 の信号の M 個のサンプルと、前記第 2 の信号の $3M/4$ 個以下のサンプルと、に基づく M 個の修正離散コサイン変換 (M D C T) 係数の組を生成するための手段を含む、請求項 5 1 に記載の装置。

【請求項 6 0】

前記第 1 の信号が M 個のサンプルの長さを有し、前記第 2 の信号が M 個のサンプルの長さを有し、

30

前記第 1 のフレームを符号化するための手段が、(A) 前記時間修正されたセグメントを含む、前記第 1 の信号の M 個のサンプルを含み、(B) ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで始まり、(C) ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで終わる、 $2M$ 個のサンプルのシーケンスに基づく M 個の修正離散コサイン変換 (M D C T) 係数の組を生成するための手段を含む、請求項 5 1 に記載の装置。

【請求項 6 1】

オーディオ信号のフレームを処理するための装置であって、

第 1 のコーディング方式に従って前記オーディオ信号の第 1 のフレームを符号化するように構成された第 1 のフレームエンコーダと、

ピッチ調整 (P R) コーディング方式に従って前記オーディオ信号の第 2 のフレームを符号化するように構成された第 2 のフレームエンコーダと、

40

を備え、

前記第 2 のフレームが、前記オーディオ信号中の前記第 1 のフレームに後続し且つ連続し、

前記第 1 のコーディング方式が非 P R コーディング方式であり、

前記第 1 のフレームエンコーダが、前記第 1 のフレームに基づく第 1 の信号のセグメントを、第 1 の時間シフトに基づいて時間修正するように構成された第 1 の時間修正器を含み、前記第 1 の時間修正器が、(A) 前記第 1 の時間シフトに従って前記第 1 の信号の前記セグメントを時間シフトすることと、(B) 前記第 1 の時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成

50

され、

前記第 2 のフレームエンコーダが、前記第 2 のフレームに基づく第 2 の信号のセグメントを、第 2 の時間シフトに基づいて時間修正するように構成された第 2 の時間修正器を含み、前記第 2 の時間修正器が、(A) 前記第 2 の時間シフトに従って前記第 2 の信号の前記セグメントを時間シフトすることと、(B) 前記第 2 の時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され、

前記第 2 の時間修正器が、前記第 2 の信号の別のピッチパルスに対する第 2 の信号の前記セグメントのピッチパルスの位置を変化させるように構成され、

前記第 2 の時間シフトが、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づく、方法。

【請求項 6 2】

前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、請求項 6 1 に記載の装置。

【請求項 6 3】

前記第 1 及び第 2 の信号が重み付きオーディオ信号である、請求項 6 1 に記載の装置。

【請求項 6 4】

前記第 2 の時間修正器が、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づいて前記第 2 の時間シフトを計算するように構成された時間シフト計算器を含み、前記時間シフト計算器が、前記第 2 のフレームからの情報に基づく遅延輪郭に、前記第 1 の信号の前記時間修正されたセグメントをマッピングするように構成されたマッパーを含む、請求項 6 1 に記載の装置。

【請求項 6 5】

前記第 2 の時間シフトが、前記マッピングされたセグメントのサンプルと一時修正残差のサンプルとの間の相関に基づき、

前記一時修正残差が、(A) 前記第 2 のフレームの残差のサンプルと、(B) 前記第 1 の時間シフトとに基づく、請求項 6 4 に記載の装置。

【請求項 6 6】

前記第 2 の信号が前記第 2 のフレームの残差であり、

前記第 2 の時間修正器が、前記第 2 の時間シフトに従って前記残差の第 1 のセグメントを時間シフトするように構成され、

前記時間シフト計算器が、前記残差の前記時間修正された第 1 のセグメントからの情報に基づいて、前記第 2 の時間シフトとは異なる第 3 の時間シフトを計算するように構成され、

前記第 2 の時間シフト器が、前記第 3 の時間シフトに従って前記残差の第 2 のセグメントを時間シフトするように構成される、請求項 6 1 に記載の装置。

【請求項 6 7】

前記第 2 の信号が前記第 2 のフレームの残差であり、前記第 2 の時間修正器が、(A) 前記第 1 の信号の前記時間修正されたセグメントからの情報と、(B) 前記第 2 のフレームの前記残差からの情報と、に基づいて前記第 2 の時間シフトを計算するように構成された時間シフト計算器を含む、請求項 6 1 に記載の装置。

【請求項 6 8】

前記第 1 のフレームエンコーダが、

前記第 1 の信号である、前記第 1 のフレームの残差を発生させるように構成された残差発生器と、

符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換 (M D C T) 演算を実行するように構成された M D C T モジュールと、

を含み、

前記第 1 のフレームエンコーダが、前記符号化残差に基づいて第 1 の符号化フレームを

10

20

30

40

50

生成するように構成される、請求項 6 1 に記載の装置。

【請求項 6 9】

前記第 1 の信号が M 個のサンプルの長さを有し、前記第 2 の信号が M 個のサンプルの長さを有し、

前記第 1 のフレームエンコーダが、前記時間修正されたセグメントを含む、前記第 1 の信号の M 個のサンプルと、前記第 2 の信号の $3M/4$ 個以下のサンプルと、に基づく M 個の修正離散コサイン変換 (M D C T) 係数の組を生成するように構成された M D C T モジュールを含む、請求項 6 1 に記載の装置。

【請求項 7 0】

前記第 1 の信号が M 個のサンプルの長さを有し、前記第 2 の信号が M 個のサンプルの長さを有し、

前記第 1 のフレームエンコーダが、(A) 前記時間修正されたセグメントを含む、前記第 1 の信号の M 個のサンプルを含み、(B) ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで始まり、(C) ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで終わる、 $2M$ 個のサンプルのシーケンスに基づく M 個の修正離散コサイン変換 (M D C T) 係数の組を生成するように構成された M D C T モジュールを含む、請求項 6 1 に記載の装置。

【請求項 7 1】

プロセッサによって実行されると前記プロセッサに、

第 1 のコーディング方式に従って前記オーディオ信号の第 1 のフレームを符号化することと、

ピッチ調整 (P R) コーディング方式に従って前記オーディオ信号の第 2 のフレームを符号化することと、

を行わせる命令を備えるコンピュータ可読媒体であって、

前記第 2 のフレームが、前記オーディオ信号中の前記第 1 のフレームに後続し且つ連続し、

前記第 1 のコーディング方式が非 P R コーディング方式であり、

前記プロセッサによって実行されると前記プロセッサに第 1 のフレームを符号化することを行わせる命令が、前記第 1 のフレームに基づく第 1 の信号のセグメントを、第 1 の時間シフトに基づいて時間修正する命令を含み、前記時間修正する命令が、(A) 前記第 1 の時間シフトに従って前記第 1 の信号の前記セグメントを時間シフトする命令と、(B) 前記第 1 の時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープする命令と、のうちの 1 つを含み、

前記プロセッサによって実行されると前記プロセッサに第 2 のフレームを符号化することを行わせる命令が、前記第 2 のフレームに基づく第 2 の信号のセグメントを、第 2 の時間シフトに基づいて時間修正する命令を含み、前記時間修正する命令が、(A) 前記第 2 の時間シフトに従って前記第 2 の信号の前記セグメントを時間シフトする命令と、(B) 前記第 2 の時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープする命令と、のうちの 1 つを含み、

前記第 2 の信号のセグメントを時間修正する命令が、前記第 2 の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させる命令を含み、

前記第 2 の時間シフトが、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づく、コンピュータ可読媒体。

【発明の詳細な説明】

【技術分野】

【0001】

米国特許法第 1 1 9 条に基づく優先権の主張

本特許出願は、本出願の譲受人に譲渡された、2007年6月13日出願の「METHOD AND APPARATUS FOR MODE SELECTION IN A GENERALIZED AUDIO CODING SYSTEM INCLUDING MULTIPLE CODING MODES」と題する仮出願第 60 / 9 4 3 ,

10

20

30

40

50

558号の優先権を主張するものである。

【0002】

本開示は、オーディオ信号の符号化に関する。

【背景技術】

【0003】

スピーチ及び/または音楽などのオーディオ情報のデジタル技法による伝送は、特に長距離電話通信、有声音オーバーIP（VoIPとも呼ばれ、IPはインターネットプロトコルを示す）などのパケット交換電話通信、及びセルラー電話通信などのデジタル無線電話通信において普及してきた。そのような普及は、再構成されたスピーチの知覚品質を維持しながら、伝送チャンネルを介して有声音通信を転送するために使用される情報量を低減することの関心を生じている。たとえば、（特にワイヤレスシステムにおいて）利用可能なシステム帯域幅を効率的に使用することが望まれている。システム帯域幅を効率的に使用する1つの方法は、信号圧縮技法を使用することである。スピーチ信号を搬送するシステムの場合、通例、スピーチ圧縮（または「スピーチコーディング」）技法がこの目的のために使用される。

10

【0004】

人間スピーチ発生のモデルに係るパラメータを抽出することによってスピーチを圧縮するように構成されたデバイスは、しばしば、オーディオコーダ、有声音コーダ、コーデック、ボコーダ、またはスピーチコーダと呼ばれ、以下の説明では、これらの用語を互換的に使用する。オーディオコーダは概してエンコーダとデコーダとを含む。エンコーダは、一般に、デジタルオーディオ信号を、「フレーム」と呼ばれるサンプルの一連のブロックとして受信し、いくつかの係るパラメータを抽出するために各フレームを分析し、対応する一連の符号化フレームを生成するためにパラメータを量子化する。符号化フレームは、伝送チャンネル（すなわち、有線またはワイヤレスネットワーク接続）を介して、デコーダを含む受信機に送信される。代替として、符号化オーディオ信号は、後で検索及び復号するために記憶されることができる。デコーダは、符号化フレームを受信して処理し、パラメータを生成するためにそれらを逆量子化し、そして、それら逆量子化されたパラメータを使用してスピーチフレームを再現する。

20

【0005】

コード励振線形予測（「CELP」）は、元のオーディオ信号の波形を適合させようと試みるコーディング方式である。リラクストド（relaxed）CELP（「RCELP」）と呼ばれるCELPの変形態を使用して、スピーチ信号のフレーム、特に有声音フレームを符号化することが望ましい場合がある。RCELPコーディング方式では、波形適合制約は緩和される。RCELPコーディング方式はピッチ調整（pitch-regularizing）（「PR」）コーディング方式であり、信号のピッチ周期の変動（「遅延輪郭（delay contour）」とも呼ばれる）が、一般に、ピッチパルスの相対位置をより滑らかな合成遅延輪郭に一致または近似するように変化させることによって調整される。ピッチ調整により、一般に知覚品質の低下をほとんどまたはまったく伴わずにピッチ情報をより少ないビットで符号化することを可能にする。一般に、調整量を指定する情報はデコーダに送信されない。以下の文書には、RCELPコーディング方式を含むコーディングシステムが記載されている；第3世代パートナーシッププロジェクト2（「3GPP2」）文書C.S0030-0、v3.0、表題「Selectable Mode Vocoder（SMV）Service Option for Wideband Spread Spectrum Communication Systems」、2004年1月（www.3gpp.orgからオンラインで入手可能）；及び3GPP2文書C.S0014-C、v1.0、表題「Enhanced Variable Rate Codec, Speech Service Options 3, 68, and 70 for Wideband Spread Spectrum Digital Systems」、2007年1月（www.3gpp.orgからオンラインで入手可能）。プロトタイプピッチ周期（「PPP」）などのプロトタイプ波形補間（「PWI」）方式を含む、有声音

30

40

50

フレーム用の他のコーディング方式は、（たとえば、上記で参照した3GPP2文書C. S0014-Cの第4.2.4.3部に記載されているように）PRとして実装されることもできる。男性話者のピッチ周波数の通常の範囲は50または70～150または200Hzを含み、女性話者のピッチ周波数の通常の範囲は120または140～300または400Hzを含む。*

公衆交換電話網（「PSTN」）を介したオーディオ通信は、従来、帯域幅が300～3400キロヘルツ（kHz）の周波数範囲に制限されてきた。セルラー電話通信及び/またはVoIPを使用するネットワークなど、オーディオ通信用のより最近のネットワークは、同じ帯域幅制限をもたない場合があり、そのようなネットワークを使用する装置では、広帯域周波数範囲を含むオーディオ通信を送信及び受信する能力を有することが望ましい場合がある。たとえば、そのような装置では、下は50Hzまで及び/または上は7もしくは8kHzまでに及ぶ可聴周波数範囲をサポートすることが望ましい場合がある。また、そのような装置では、従来のPSTN制限外の範囲のオーディオスピーチコンテンツを有することがある、高品質オーディオまたはオーディオ/テレビ会議、音楽及び/またはテレビジョンなどのマルチメディアサービスの配信など、他の適用例をサポートすることが望ましい場合がある。

【0006】

スピーチコーダによってサポートされる範囲をより高い周波数に拡大することは、了解度を向上させることができる。たとえば、「s」や「f」などの摩擦音を区別するスピーチ信号中の情報は、大部分は高周波数にある。ハイバンド拡大は、臨場感など、復号されたスピーチ信号の他の品質を向上させることもできる。たとえば、有声母音でさえも、PSTN周波数範囲をはるかに上回るスペクトルエネルギーを有する場合がある。

【発明の概要】

【0007】

概略構成によるオーディオ信号のフレームを処理する方法は、ピッチ調整（「PR」）コーディング方式に従ってオーディオ信号の第1のフレームを符号化することと；非PRコーディング方式に従ってオーディオ信号の第2のフレームを符号化することと、を含む。この方法では、第2のフレームは、オーディオ信号中の第1のフレームに後続し且つ連続し、第1のフレームを符号化することは、第1のフレームに基づく第1の信号のセグメントを時間シフトに基づいて時間修正（time-modify）することを含み、時間修正することは、（A）時間シフト（time-shift）に従って第1の信号のセグメントを時間シフトすることと、（B）第1の時間シフトに基づいて第1の信号のセグメントをタイムワープ（time-warp）することと、のうちの1つを含む。この方法では、第1の信号のセグメントを時間修正することは、第1の信号の別のピッチパルスに対するセグメントのピッチパルスの位置を変化させることを含む。この方法では、第2のフレームを符号化することは、第2のフレームに基づく第2の信号のセグメントを時間シフトに基づいて時間修正することを含み、時間修正することは、（A）時間シフトに従って第2のフレームのセグメントを時間シフトすることと、（B）時間シフトに基づいて第2の信号のセグメントをタイムワープすることと、のうちの1つを含む。また、そのような方法でオーディオ信号のフレームを処理するための命令を有するコンピュータ可読媒体、ならびに同様の方法でオーディオ信号のフレームを処理するための装置及びシステムが説明される。

【0008】

別の概略構成に従ったオーディオ信号のフレームを処理する方法は、第1のコーディング方式に従ってオーディオ信号の第1のフレームを符号化することと；PRコーディング方式に従ってオーディオ信号の第2のフレームを符号化することと、を含む。この方法では、第2のフレームは、オーディオ信号中の第1のフレームに後続し且つ連続し、第1のコーディング方式は非PRコーディング方式である。この方法では、第1のフレームを符号化することは、第1のフレームに基づく第1の信号のセグメントを第1の時間シフトに基づいて時間修正することを含み、時間修正することは、（A）第1の時間シフトに従って第1の信号のセグメントを時間シフトすることと、（B）第1の時間シフトに基づいて

第 1 の信号のセグメントをタイムワープすることと、のうちの 1 つを含む。この方法では、第 2 のフレームを符号化することは、第 2 のフレームに基づく第 2 の信号のセグメントを第 2 の時間シフトに基づいて時間修正することを含み、時間修正することは、(A) 第 2 の時間シフトに従って第 2 の信号のセグメントを時間シフトすることと、(B) 第 2 の時間シフトに基づいて第 2 の信号のセグメントをタイムワープすることと、のうちの 1 つを含む。この方法では、第 2 の信号のセグメントを時間修正することは、第 2 の信号の別のピッチパルスに対するセグメントのピッチパルスの位置を変化させることを含み、第 2 の時間シフトは、第 1 の信号の時間修正されたセグメントからの情報に基づく。また、そのような方法でオーディオ信号のフレームを処理するための命令を有するコンピュータ可読媒体、ならびに同様の方法でオーディオ信号のフレームを処理するための装置及びシステムが説明される。

10

【図面の簡単な説明】

【0009】

【図 1】図 1 は、ワイヤレス電話システムの例を示している。

【図 2】図 2 は、パケット交換データ通信をサポートするように構成されたセルラー電話通信システムの例を示している。

【図 3 a】図 3 a は、オーディオエンコーダ A E 1 0 とオーディオデコーダ A D 1 0 とを含むコーディングシステムのブロック図を示している。

【図 3 b】図 3 b は、1 対のコーディングシステムのブロック図を示している。

【図 4 a】図 4 a は、オーディオエンコーダ A E 1 0 のマルチモード実装形態 A E 2 0 のブロック図を示している。

20

【図 4 b】図 4 b は、オーディオデコーダ A D 1 0 のマルチモード実装形態 A D 2 0 のブロック図を示している。

【図 5 a】図 5 a は、オーディオエンコーダ A E 2 0 の実装形態 A E 2 2 のブロック図を示している。

【図 5 b】図 5 b は、オーディオエンコーダ A E 2 0 の実装形態 A E 2 4 のブロック図を示している。

【図 6 a】図 6 a は、オーディオエンコーダ A E 2 4 の実装形態 A E 2 5 のブロック図を示している。

【図 6 b】図 6 b は、オーディオエンコーダ A E 2 0 の実装形態 A E 2 6 のブロック図を示している。

30

【図 7 a】図 7 a は、オーディオ信号のフレームを符号化する方法 M 1 0 のフローチャートを示している。

【図 7 b】図 7 b は、オーディオ信号のフレームを符号化するように構成された装置 F 1 0 のブロック図を示している。

【図 8】図 8 は、遅延輪郭に対してタイムワープされる前及び後の残差の例を示している。

【図 9】図 9 は、区分的修正の前及び後の残差の例を示している。

【図 10】図 10 は、R C E L P 符号化の方法 R M 1 0 0 のフローチャートを示している。

40

【図 11】図 11 は、R C E L P 符号化方法 R M 1 0 0 の実装形態 R M 1 1 0 のフローチャートを示している。

【図 12 a】図 12 a は、R C E L P フレームエンコーダ (frame encoder) 3 4 c の実装形態 R C 1 0 0 のブロック図を示している。

【図 12 b】図 12 b は、R C E L P エンコーダ R C 1 0 0 の実装形態 R C 1 1 0 のブロック図を示している。

【図 12 c】図 12 c は、R C E L P エンコーダ R C 1 0 0 の実装形態 R C 1 0 5 のブロック図を示している。

【図 12 d】図 12 d は、R C E L P エンコーダ R C 1 1 0 の実装形態 R C 1 1 5 のブロック図を示している。

50

【図13】図13は、残差発生器(residual generator) R10の実装形態R12のブロック図を示している。

【図14】図14は、RCELP符号化のための装置RF100のブロック図を示している。

【図15】図15は、RCELP符号化方法RM100の実装形態RM120のフローチャートを示している。

【図16】図16は、MDCTコーディング方式のための典型的な正弦ウィンドウ形状の3つの例を示している。

【図17a】図17は、MDCTエンコーダ34dの実装形態ME100のブロック図を示している。

10

【図17b】図17bは、MDCTエンコーダ34dの実装形態ME200のブロック図を示している。

【図18】図18は、図16に示すウィンドウ処理技法とは異なるウィンドウ処理技法の一例を示している。

【図19a】図19aは、概略構成によるオーディオ信号のフレームを処理する方法M100のフローチャートを示している。

【図19b】図19bは、タスクT110の実装形態T112のフローチャートを示している。

【図19c】図19cは、タスクT112の実装形態T114のフローチャートを示している。

20

【図20a】図20aは、MDCTエンコーダME100の実装形態ME110のブロック図を示している。

【図20b】図20bは、MDCTエンコーダME200の実装形態ME210のブロック図を示している。

【図21a】図21aは、MDCTエンコーダME100の実装形態ME120のブロック図を示している。

【図21b】図21bは、MDCTエンコーダME100の実装形態ME130のブロック図を示している。

【図22】図22は、MDCTエンコーダME120及びME130の実装形態ME140のブロック図を示している。

30

【図23a】図23aは、MDCT符号化の方法MM100のフローチャートを示している。

【図23b】図23bは、MDCT符号化のための装置MF100のブロック図を示している。

【図24a】図24aは、概略構成によるオーディオ信号のフレームを処理する方法M200のフローチャートを示している。

【図24b】図24bは、タスクT620の実装形態T622のフローチャートを示している。

【図24c】図24cは、タスクT620の実装形態T624のフローチャートを示している。

40

【図24d】図24dは、タスクT622及びT624の実装形態T626のフローチャートを示している。

【図25a】図25aは、オーディオ信号の連続フレームにMDCTウィンドウを適用することから生じる重複追加領域の一例を示している。

【図25b】図25bは、非PRフレームのシーケンスに時間シフトを適用する例を示している。

【図26】図26は、オーディオ通信用のデバイス1108のブロック図を示している。

【発明を実施するための形態】

【0010】

本明細書で説明するシステム、方法、及び装置は、マルチモードオーディオコーディン

50

グシステム、特に修正離散コサイン変換（「MDC T」）コーディング方式などの重複追加非PRコーディング方式を含むコーディングシステムにおけるPRコーディング方式と非PRコーディング方式との間の遷移中に、高い知覚品質をサポートするために使用されることができる。以下で説明する構成は、符号分割多元接続（「CDMA」）無線インターフェースを使用するように構成されたワイヤレス電話通信システム中に存在する。とはいえ、本明細書で説明する特徴を有する方法及び装置は、有線及び/またはワイヤレス（たとえば、CDMA、TDMA、FDMA、及び/またはTD-SCDMA）伝送チャンネルを介した有声音オーバーIP（「VoIP」）を使用するシステムなど、当業者に知られている広範な技術を使用する様々な通信システムのいずれにも存在できることが当業者には理解されよう。

10

【0011】

本明細書に開示する構成は、パケット交換式であるネットワーク（たとえば、VoIPなどのプロトコルに従ってオーディオ伝送を行なうように構成された有線及び/またはワイヤレスネットワーク）及び/または回線交換式であるネットワークにおける使用に適應され得るとということが明確に企図され、本明細書によって開示される。また、本明細書に開示する構成は、狭帯域コーディングシステム（たとえば、約4または5キロヘルツのオーディオ周波数範囲を符号化するシステム）での使用、ならびに全バンド（whole-band）帯域広帯域コーディングシステム及びスプリットバンド（split-band）広帯域コーディングシステムを含む、広帯域コーディングシステム（たとえば、5キロヘルツを超えるオーディオ周波数を符号化するシステム）での使用に適應され得るとということが明確に企図され、本明細書によって開示される。

20

【0012】

文脈によって明確に限定されない限り、「信号」という用語は、本明細書では、ワイヤ、バス、または他の伝送媒体上に表されたメモリ位置（またはメモリ位置の組）の状態を含む、その通常の意味のいずれをも表すのに使用される。文脈によって明確に限定されない限り、「発生（generating）」という用語は、本明細書では、コンピュータ計算（computing）または別様の生成（producing）など、その通常の意味のいずれをも表すのに使用される。文脈によって明確に限定されない限り、「計算（calculating）」という用語は、本明細書では、コンピュータ計算、評価、平滑化、及び/または複数の値からの選択など、その通常の意味のいずれをも表すのに使用される。文脈によって明確に限定されない限り、「取得」という用語は、計算、導出、（たとえば、外部デバイスからの）受信、及び/または（たとえば、記憶要素のレイからの）検索など、その通常の意味のいずれをも表すのに使用される。「備える（comprising）」という用語は、本明細書及び特許請求の範囲において使用される場合、他の要素または動作を除外するものではない。「AはBに基づく」という表現は、（特定の文脈において適切であるならば）（i）「Aは少なくともBに基づく」及び（ii）「AはBに等しい」という場合を含む、その通常の意味のいずれをも表すのに使用される。

30

【0013】

別段の指示がない限り、特定の特徴を有する装置の動作のいかなる開示も、類似の特徴を有する方法を開示する（その逆も同様）ことをも明確に意図し、特定の構成による装置の動作のいかなる開示も、類似の構成による方法を開示する（その逆も同様）ことをも明確に意図する。たとえば、別段の指定がない限り、特定の特徴を有するオーディオエンコードのいかなる開示も、類似の特徴を有するオーディオ符号化の方法を開示する（その逆も同様）ことをも明確に意図し、特定の構成によるオーディオエンコードのいかなる開示も、類似の構成によるオーディオ符号化の方法を開示する（その逆も同様）ことをも明確に意図する。

40

【0014】

文書の一部の参照によるいかなる組込みも、その部分内で言及された用語または変数の定義が文書中の他の場所に現れた場合、そのような定義を組み込んでいることをも理解されたい。

50

【 0 0 1 5 】

「コーダ」、「コーデック」、及び「コーディングシステム」という用語は、(場合によっては知覚的重み付け及び/または他のフィルタ処理操作などの1つまたは複数の前処理操作の後に)オーディオ信号のフレームを受信するように構成された少なくとも1つのエンコーダと、フレームの復号化表現を生成するように構成された対応するデコーダと、を含むシステムを示すのに互換的に使用される。

【 0 0 1 6 】

図1に示すように、ワイヤレス電話システム(たとえば、CDMA、TDMA、FDMA、及び/またはTD-SCDMAシステム)は、概して、複数の基地局(BS)12と1つまたは複数の基地局コントローラ(BSC)14とを含む無線アクセスネットワークとワイヤレスで通信するように構成された複数の移動体加入者ユニット10を含む。そのようなシステムは、概して、BSC14に結合され、従来の公衆交換電話網(PSTN)18に当該無線アクセスネットワークをインターフェースするように構成された、移動体交換センター(MSC)16をも含む。このインターフェースをサポートするために、MSCは、ネットワーク間の変換ユニットとして働くメディアゲートウェイを含むか、またはメディアゲートウェイと通信することができる。メディアゲートウェイは、異なる伝送及び/またはコーディング技法など、異なるフォーマット間で変換する(たとえば、時分割多重化(「TDM」)有声音とVoIPとの間で変換する)ように構成され、また、エコー除去、デュアルタイム多重周波数(「DTMF」)、及びトーン送信などのメディアストリーミング機能を実行するように構成されることができる。BSC14は、迂回中継線を介して基地局12に結合される。迂回中継線は、たとえば、E1/T1、ATM、IP、PPP、フレームリレー、HDSL、ADSL、またはxDSLを含む、いくつかの知られているインターフェースのいずれをもサポートするように構成されることができる。基地局12と、BSC14と、MSC16と、もしあればメディアゲートウェイとの集合は「インフラストラクチャ」とも呼ばれる。

【 0 0 1 7 】

各基地局12は、有利には、少なくとも1つのセクタ(図示せず)を含み、各セクタは、全方向性アンテナ、または基地局12から径方向に離れるある特定の方向を向いたアンテナを備える。代替として、各セクタは、ダイバーシチ受信用の2つ以上のアンテナを備えることができる。各基地局12は、有利には、複数の周波数割当てをサポートするように設計されることができる。セクタと周波数割当ての交わり(intersection)は、CDMAチャネルと呼ばれることがある。基地局12は、基地局トランシーバサブシステム(BTS)12としても知られている。代替として、「基地局」は、当業界において、BSC14と1つまたは複数のBTS12とを総称するのに使用される場合がある。BTS12は「セルサイト」12と表される場合もある。代替として、所与のBTS12の個々のセクタをセルサイトと呼ぶことがある。移動体加入者ユニット10は、一般に、セルラー及び/またはパーソナル通信サービス(「PCS」)電話、携帯情報端末(「PDA」)、及び/または移動電話機能を有する他のデバイスが含まれる。そのようなユニット10は、内蔵型のスピーカ及びマイクロホン、スピーカとマイクロホンとを含むコード付きハンドセットまたはヘッドセット(たとえば、USBハンドセット)、またはスピーカとマイクロホンとを含むワイヤレスヘッドセット(たとえば、Bluetooth Special Interest Group(ワシントン州ベルビュー)によって公表されたBluetooth(登録商標)プロトコルのバージョンを使用してユニットにオーディオ情報を通信するヘッドセット)を含むことができる。そのようなシステムは、IS-95標準の1つまたは複数のバージョン(たとえば、Telecommunications Industry Alliance(バージニア州アーリントン)によって発表されたIS-95、IS-95A、IS-95B、cdma2000)に従う使用のために構成されることができる。

【 0 0 1 8 】

次に、セルラー電話システムの典型的な動作について説明する。基地局12は、移動体

10

20

30

40

50

加入者ユニット10の組から、上りリンク信号の組を受信する。移動体加入者ユニット10は、電話通話または他の通信を行っている。所与の基地局12によって受信された各上りリンク信号は、その基地局12内で処理され、得られたデータはBSC14に転送される。BSC14は、コールリソース割当てと、基地局12間のソフトハンドオフの編成を含むモビリティ管理機能とを提供する。BSC14はまた、PSTN18とのインターフェースのための追加のルーティングサービスを提供するMSC16に受信データをルーティングする。同様に、PSTN18は、MSC16とインターフェースをとり、MSC16は、下りリンク信号の組を移動体加入者ユニット10の組に送信するように基地局12を制御するBSC14とインターフェースをとる。

【0019】

図1に示すセルラー電話通信システムの要素は、パケット交換データ通信をサポートするように構成されることもできる。図2に示すように、パケットデータトラフィックは、概して、パケットデータネットワークに接続されたゲートウェイルータに結合されたパケットデータサービスノード(PDSN)22を使用して、移動体加入者ユニット10と外部のパケットデータネットワーク24(たとえば、インターネットなどの公衆網)との間をルーティングされる。PDSN22は、1つまたは複数のBSC14にサービスを提供し、パケットデータネットワークと無線アクセスネットワークとの間のリンクとして働く、1つまたは複数のパケット制御機能(PCF)20にデータを順次ルーティングする。パケットデータネットワーク24はまた、ローカルエリアネットワーク(「LAN」)、キャンパスエリアネットワーク(「CAN」)、メトロポリタンエリアネットワーク(「MAN」)、広域ネットワーク(「WAN」)、リング型ネットワーク、スター型ネットワーク、トークンリングネットワークなどを含むように実装されることができ、ネットワーク24に接続されたユーザ端末は、PDA、ラップトップコンピュータ、パーソナルコンピュータ、ゲームデバイス(そのようなデバイスの例には、XBox及びXBox360(マイクロソフト社(ワシントン州レッドモンド))、プレイステーション3及びプレイステーション・ポータブル(ソニー(株)、日本国東京)、ならびにWii及びDS(任天堂、日本国京都)がある)、及び/またはオーディオ処理機能を有する任意のデバイスとすることができ、VoIPなど、1つまたは複数のプロトコルを使用して電話通話または他の通信をサポートするように構成されることができ、そのような端末は、内蔵型のスピーカ及びマイクロホン、スピーカとマイクロホンとを含むコード付きハンドセット(たとえば、USBハンドセット)、またはスピーカとマイクロホンとを含むワイヤレスヘッドセット(たとえば、Bluetooth Special Interest Group(ワシントン州ベルビュー)によって公表されたBluetoothプロトコルのバージョンを使用して端末にオーディオ情報を通信するヘッドセット)を含むことができる。そのようなシステムは、PSTNに決して入ることなく、(たとえば、VoIPなど、1つまたは複数のプロトコルによる)異なる無線アクセスネットワーク上の移動体加入者ユニット間、移動体加入者ユニットと非移動体ユーザ端末との間、または2つの非移動体ユーザ端末間のパケットデータトラフィックとして電話通話または他の通信を行なうように構成されることができ、移動体加入者ユニット10または他のユーザ端末は、「アクセス端末」とも呼ばれる。

【0020】

図3aは、デジタルオーディオ信号S100を(たとえば、一連のフレームとして)受信し、通信チャネルC100(たとえば、有線、光及び/またはワイヤレス通信リンク)上でオーディオデコーダAD10に送信するための対応する符号化信号S200を(たとえば、一連の対応する符号化フレームとして)生成するように構成されたオーディオエンコーダAE10を示している。オーディオデコーダAD10は、符号化オーディオ信号S200の受信されたバージョンS300を復号し、対応する出力スピーチ信号S400を合成するように構成される。

【0021】

オーディオ信号S100は、デジタル化され、パルス符号変調(「PCM」)、 μ -1

10

20

30

40

50

a w 圧伸または A 則圧伸など当技術分野で知られている様々な方法のいずれかに従って量子化された（たとえば、マイクロホンによって捕捉された）アナログ信号を表す。この信号は、ノイズ抑圧、知覚的重み付け、及び/または他のフィルタ処理操作など、アナログ及び/またはデジタル領域における他の前処理操作を受けてもよい。追加または代替として、そのような操作は、オーディオエンコーダ A E 1 0 内で実行されることができる。オーディオ信号 S 1 0 0 のインスタンス（instance）は、デジタル化され、量子化された（たとえば、一連のマイクロホンによって捕捉された）アナログ信号の組合せを表すこともできる。

【 0 0 2 2 】

図 3 b は、デジタル化オーディオ信号 S 1 0 0 の第 1 のインスタンス S 1 1 0 を受信し、通信チャンネル C 1 0 0 の第 1 のインスタンス C 1 1 0 上でオーディオデコーダ A D 1 0 の第 1 のインスタンス A D 1 0 a に送信するための、符号化信号 S 2 0 0 の対応するインスタンス S 2 1 0 を生成するように構成されたオーディオエンコーダ A E 1 0 の第 1 のインスタンス A E 1 0 a を示している。オーディオデコーダ A D 1 0 a は、符号化オーディオ信号 S 2 1 0 の受信されたバージョン S 3 1 0 を復号し、出力スピーチ信号 S 4 0 0 の対応するインスタンス S 4 1 0 を合成するように構成される。

10

【 0 0 2 3 】

図 3 b はまた、デジタルオーディオ信号 S 1 0 0 の第 2 のインスタンス S 1 2 0 を受信し、通信チャンネル C 1 0 0 の第 2 のインスタンス C 1 2 0 上でオーディオデコーダ A D 1 0 の第 2 のインスタンス A D 1 0 b に送信するための、符号化信号 S 2 0 0 の対応するインスタンス S 2 2 0 を生成するように構成されたオーディオエンコーダ A E 1 0 の第 2 のインスタンス A E 1 0 b をも示している。オーディオデコーダ A D 1 0 b は、符号化オーディオ信号 S 2 2 0 の受信されたバージョン S 3 2 0 を復号し、出力スピーチ信号 S 4 0 0 の対応するインスタンス S 4 2 0 を合成するように構成される。

20

【 0 0 2 4 】

オーディオエンコーダ A E 1 0 a 及びオーディオデコーダ A D 1 0 b（同様に、オーディオエンコーダ A E 1 0 b 及びオーディオデコーダ A D 1 0 a）は、たとえば、図 1 及び図 2 に関して上述した加入者ユニット、ユーザ端末、メディアゲートウェイ、B T S、または B S C を含む、スピーチ信号を送信及び受信するためのどんな通信デバイスにおいても一緒に使用されることができる。本明細書で説明するように、オーディオエンコーダ A E 1 0 は多数の異なる方法で実装でき、オーディオエンコーダ A E 1 0 a 及び A E 1 0 b はオーディオエンコーダ A E 1 0 の異なる実装形態のインスタンスとすることができる。同様に、オーディオデコーダ A D 1 0 は多数の異なる方法で実装でき、オーディオデコーダ A D 1 0 a 及び A D 1 0 b はオーディオデコーダ A D 1 0 の異なる実装形態のインスタンスとすることができる。

30

【 0 0 2 5 】

オーディオエンコーダ（たとえば、オーディオエンコーダ A E 1 0）は、オーディオ信号のデジタルサンプルを入力データの一連のフレームとして処理するもので、各フレームは所定数のサンプルを備える。フレームまたは（サブフレームとも呼ばれる）フレームのセグメントを処理する操作は、その入力中の 1 つまたは複数の隣接フレームのセグメントを含むこともできるとはいえ、この一連は、通常、重複しない一連として実装される。オーディオ信号のフレームは、一般に、信号のスペクトル包絡線がそのフレームにわたって比較的固定のままであることが予想できるほど十分に短い。フレームは、一般に、5 ミリ秒と 3 5 ミリ秒との間のオーディオ信号（または約 4 0 サンプルから 2 0 0 サンプルまで）に対応し、電話通信の適用例では 2 0 ミリ秒が通常のフレームサイズである。通常のフレームサイズの他の例は、1 0 ミリ秒及び 3 0 ミリ秒を含む。一般に、オーディオ信号のすべてのフレームは同じ長さを持ち、本明細書で説明する特定の例では、一様のフレーム長を仮定する。ただし、一様でないフレーム長が使用され得ることも明確に企図され、本明細書によって開示される。

40

【 0 0 2 6 】

50

20ミリ秒のフレーム長は、7キロヘルツ(kHz)のサンプリングレートでは140個のサンプルに対応し、8kHzのサンプリングレート(狭帯域コーディングシステム用の1つの典型的なサンプリングレート)では160個のサンプルに対応し、16kHzのサンプリングレート(広帯域コーディングシステム用の1つの典型的なサンプリングレート)では320個のサンプルに対応するが、特定の適用例に好適であると考えられるどんなサンプリングレートも使用され得る。スピーチコーディングに使用され得るサンプリングレートの別の例は12.8kHzであり、さらなる例は、12.8kHz~38.4kHzの範囲の他のレートを含む。

【0027】

電話通話など典型的なオーディオ通信セッションでは、各話者は約60パーセントの時間の間沈黙している。そのような適用例のためのオーディオエンコーダは、通常、スピーチまたは他の情報を含むオーディオ信号のフレーム(「アクティブフレーム」)を、バックグラウンドノイズまたは無音のみを含むオーディオ信号のフレーム(「非アクティブフレーム」)と区別するように構成される。アクティブフレーム及び非アクティブフレームを符号化するのに異なるコーディングモード及び/またはビットレートを使用するようにオーディオエンコーダAE10を実装することが望ましい場合がある。たとえば、オーディオエンコーダAE10は、アクティブフレームを符号化するために使用するビットよりも、非アクティブフレームを符号化するために使用するビットのほうが少なくなる(すなわち、ビットレートが低くなる)ように実装されることができ、オーディオエンコーダAE10のためには、異なるタイプのアクティブフレームを符号化するのに異なるビットレートを、比較的少ないスピーチ情報を含むフレームに対して選択的に使用されることができ、アクティブフレームを符号化するのに通例使用されるビットレートの例は、1フレーム当たり171ビット、1フレーム当たり80ビット、及び1フレーム当たり40ビットを含み;非アクティブフレームを符号化するのに通例使用されるビットレートの例には、1フレーム当たり16ビットを含む。セルラー電話通信システム(特に、Telecommunications Industry Association(バージニア州アーリントン)によって公表された暫定標準(IS)-95または同様の業界標準に準拠するシステム)のコンテキスト(context)では、これらの4つのビットレートは、それぞれ「フルレート」、「ハーフレート」、「1/4レート」、及び「1/8レート」とも呼ばれる。

【0028】

オーディオエンコーダAE10のためには、オーディオ信号の各アクティブフレームをいくつかの異なるタイプの1つとして分類することが望ましい場合がある。これらの異なるタイプは、有声音スピーチ(たとえば、母音を表すスピーチ)のフレーム、遷移フレーム(たとえば、単語の先頭または末尾を表すフレーム)、無声音スピーチ(たとえば、摩擦音を表すスピーチ)のフレーム、及び非スピーチ情報のフレーム(たとえば、歌唱及び/または楽器などの音楽、あるいは他のオーディオコンテンツ)を含むことができる。異なるタイプのフレームを符号化するのに異なるコーディングモードを使用するようにオーディオエンコーダAE10を実装することが望ましい場合がある。たとえば、有声音スピーチのフレームは、長期間であり(すなわち、複数のフレーム周期の間継続し)且つピッチに関連する周期構造を有する傾向があり、一般に、この長期間スペクトル特徴の記述を符号化するコーディングモードを使用して有声音フレーム(または一連の有声音フレーム)を符号化するのがより効率的である。そのようなコーディングモードの例は、コード励振線形予測(「CELP」)、プロトタイプ波形補間(「PWI」)、及びプロトタイプピッチ周期(「PPP」)を含む。一方、無声音フレーム及び非アクティブフレームは、通常、著しい長期間スペクトル特徴がなく、オーディオエンコーダは、そのような特徴を記述しようと試みないコーディングモードを使用して、これらのフレームを符号化するように構成されることができ、ノイズ励起線形予測(「NELP」)は、そのようなコーディングモードの一例である。音楽のフレームは、通常、異なるトーンの混合体を含み、

10

20

30

40

50

オーディオエンコーダは、フーリエ変換またはコサイン変換などの正弦分解に基づく方法を使用して、これらのフレーム（またはこれらのフレームに対するLPC分析演算の残差）を符号化するように構成されることができる。1つのそのような例は、修正離散コサイン変換（「MDC T」）に基づくコーディングモードである。

【0029】

オーディオエンコーダAE10、または対応するオーディオ符号化の方法は、ビットレートとコーディングモード（「コーディング方式」とも呼ばれる）の様々な組合せの中から選択するように実装されることができる。たとえば、オーディオエンコーダAE10は、有声音スピーチを含むフレーム及び遷移フレームにはフルレートCELP方式を、無声音スピーチを含むフレームにはハーフレートNELP方式を、非アクティブフレームには1/8レートNELP方式を、及び（たとえば、音楽を含むフレームを含んでいる）一般的なオーディオフレームにはフルレートMDC T方式を使用するように実装されることができる。代替として、オーディオエンコーダAE10のそのような実装形態は、有声音スピーチを含む少なくともいくつかのフレーム、特に高度有声音フレームに対してフルレートPPP方式を使用するように構成され得る。

10

【0030】

オーディオエンコーダAE10は、フルレート及びハーフレートCELP方式、及び/またはフルレート及び1/4レートPPP方式など、1つまたは複数のコーディング方式の各々に対して複数のビットレートをサポートするように構成されることができる。安定した有声音スピーチの周期を含む一連のフレームは、たとえば、少なくともフレームの一部が、知覚品質を大きく損なうことなくフルレート未満で符号化されることができるように、かなり冗長になる傾向がある。

20

【0031】

（複数のビットレート及び/またはコーディングモードをサポートするオーディオコーダを含む）マルチモードオーディオコーダは、概して、低ビットレートでの効率的なオーディオコーディングを提供する。当業者は、コーディング方式の数を増やすと、コーディング方式を選択する際の柔軟性が増し、その結果、平均ビットレートを低くできることを認識するであろう。ただし、コーディング方式の数が増えると、それに応じて全システム内の複雑さが増すことになる。所与のシステムにおいて使用される利用可能な方式の特定の組合せは、利用可能なシステムリソースと特定の信号環境とによって規定されるだろう。マルチモードコーディング技法の例は、たとえば、「VARIABLE RATE SPEECH CODING」と題する米国特許第6,691,084号、及び「ARBITRARY AVERAGE DATA RATES FOR VARIABLE RATE CODERS」と題する米国特許公開第2007/0171931号に記載されている。

30

【0032】

図4aは、オーディオエンコーダAE10のマルチモード実装形態AE20のブロック図を示している。エンコーダAE20は、コーディング方式選択器20と、複数p個のフレームエンコーダ30a~30pとを含む。p個のフレームエンコーダの各々は、それぞれのコーディングモードに従ってフレームを符号化するように構成され、コーディング方式選択器20によって生成されたコーディング方式選択信号は、現在フレームに対して所望のコーディングモードを選択するようにオーディオエンコーダAE20の1対の選択器50aと50bを制御するために使用される。コーディング方式選択器20は、選択されたビットレートで現在フレームを符号化するように、選択されたフレームエンコーダを制御するように構成されることができる。オーディオエンコーダAE20のソフトウェアまたはファームウェア実装形態は、実行のフローをフレームデコーダの1つまたは別の1つに導くためにコーディング方式指示を使用することができ、そのような実装形態は、選択器50a及び/または選択器50bの類似物を含まないことができることに留意されたい。フレームエンコーダ30a~30pのうち2つ以上（場合によってはすべて）は、（場合によっては、スピーチ及び非スピーチフレームの次数が非アクティブフレームの次数

40

50

よりも高いなど、異なるコーディング方式に対して異なる次数を有する結果を生成するように構成された) L P C 係数値の計算器及び/または L P C 残差発生器など、共通の構造を共有することができる。

【0033】

コーディング方式選択器 20 は、一般に、入力オーディオフレームを調べ、どのコーディングモードまたは方式をそのフレームに適用するかに関する決定を行うオープンループ決定モジュールを含む。このモジュールは、一般に、フレームをアクティブまたは非アクティブとして分類するように構成され、また、有声音、無声音、遷移、または一般的なオーディオなど、2 つ以上の異なるタイプのうちの 1 つとして、アクティブフレームを分類するように構成されることもできる。フレーム分類は、全体的なフレームエネルギー、2 つ以上の異なる周波数帯域の各々におけるフレームエネルギー、信号対ノイズ比(「SNR」)、周期性、及びゼロ交差レートなど、現在フレーム、及び/あるいは 1 つまたは複数の前のフレームの 1 つまたは複数の特性に基づかれることができる。コーディング方式選択器 20 は、そのような特性の値を計算するように、オーディオエンコーダ A E 20 の 1 つまたは複数の他のモジュールからそのような特性の値を受信するように、及び/またはオーディオエンコーダ A E 20 を含むデバイス(たとえば、セルラー電話)の 1 つまたは複数の他のモジュールからそのような特性の値を受信するように実装されることができる。フレーム分類は、そのような特性の値または大きさを閾値と比較すること、及び/またはそのような値の変化の大きさを閾値と比較することを含むことができる。

10

【0034】

オープンループ決定モジュールは、フレームが含んでいるスピーチのタイプに従って特定のフレームを符号化するときのビットレートを選択するように構成されることができる。そのような動作は「可変レートコーディング」と呼ばれる。たとえば、より高いビットレート(たとえば、フルレート)で遷移フレームを符号化し、より低いビットレート(たとえば、1/4 レート)で無声音フレームを符号化し、中間のビットレート(たとえば、ハーフレート)でまたはより高いビットレート(たとえば、フルレート)で有声音フレームを符号化するようにオーディオエンコーダ A D 20 を構成することが望ましい場合がある。特定のフレームに対して選択されるビットレートは、所望の平均ビットレート、(所望の平均ビットレートをサポートするために使用され得る)一連のフレームにわたるビットレートの所望のパターン、及び/または前のフレームに対して選択されたビットレートなどの基準に依存することもできる。

20

30

【0035】

コーディング方式選択器 20 はまた、オープンループ選択コーディング方式を使用する完全なまたは部分的な符号化の後に符号化性能の 1 つまたは複数の尺度が得られる、閉ループコーディング決定を実行するように実装されることができる。閉ループテストにおいて考察される性能尺度は、たとえば、SNR、PPP スピーチエンコーダなどの符号化方式における SNR 予測、予測誤差量子化 SNR、位相量子化 SNR、振幅量子化 SNR、知覚 SNR、及び定常性の尺度としての現在フレームと過去のフレームとの間の正規化相互相関を含む。コーディング方式選択器 20 は、そのような特性の値を計算するように、オーディオエンコーダ A E 20 の 1 つまたは複数の他のモジュールからそのような特性の値を受け取るように、及び/またはオーディオエンコーダ A E 20 を含むデバイス(たとえば、セルラー電話)の 1 つまたは複数の他のモジュールからそのような特性の値を受け取るように実装されることができる。性能尺度が閾値を下回る場合、ビットレート及び/またはコーディングモードは、より良い品質を与えることが期待されるものに変更されることができる。可変レートマルチモードオーディオコーデの品質を維持するために使用されることができる閉ループ分類方式の例は、「METHOD AND APPARATUS FOR MAINTAINING A TARGET BIT RATE IN A SPEECH CODER」と題する米国特許第 6,330,532 号、及び「METHOD AND APPARATUS FOR PERFORMING SPEECH FRAME ENCODING MODE SELECTION IN A VARIABLE

40

50

E R A T R E E N C O D I N G S Y S T E M」と題する米国特許第5,911,128号に記載されている。

【0036】

図4bは、対応する復号オーディオ信号S400を生成するために、受信された符号化オーディオ信号S300を処理するように構成されたオーディオデコーダAD10の実装形態AD20のブロック図を示している。オーディオデコーダAD20は、コーディング方式検出器60と、複数p個のフレームデコーダ70a~70pとを含む。デコーダ70a~70pは、フレームデコーダ70aがフレームエンコーダ30aによって符号化されたフレームを復号するように構成され、以下同様となるように、上述のオーディオエンコーダAE20のエンコーダに対応するように構成されることができる。フレームデコーダ70a~70pのうち2つ以上(場合によってはすべて)は、復号LPC係数値の組に従って構成可能な合成フィルタなど、共通の構造を共有することができる。そのような場合、フレームデコーダは、主に、復号オーディオ信号を生成するために合成フィルタを励起する励起信号を発生させるために使用する技法が異なる。オーディオデコーダAD20は、一般に、(たとえば、フォルマント周波数を強調すること、及び/またはスペクトルの谷を減衰させることによって)量子化ノイズを低減するために復号オーディオ信号S400を処理するように構成されたポストフィルタ(postfilter)をも含み、また適応利得制御を含むこともできる。オーディオデコーダAD20を含むデバイス(たとえば、セルラー電話)は、イヤホン、スピーカ、もしくは他のオーディオトランスデューサ、及び/またはデバイスの筐体内にあるオーディオ出力ジャックに出力するための、復号オーディオ信号S400からアナログ信号を生成するように設定及び構成されたデジタルアナログ変換器(DAC)を含むことができる。そのようなデバイスは、アナログ信号がジャック及び/またはトランスデューサに印加される前に、そのアナログ信号に対して1つまたは複数のアナログ処理演算(たとえば、フィルタ処理、等化、及び/または増幅)を実行するように構成されることができる。

10

20

【0037】

コーディング方式検出器60は、受信された符号化オーディオ信号S300の現在フレームに対応するコーディング方式を指示するように構成される。適切なコーディングビットレート及び/またはコーディングモードは、フレームのフォーマットによって指示されることができる。コーディング方式検出器60は、レート検出を実行するように、または多重サブレイヤなど、オーディオデコーダAD20が埋め込まれている装置の別の部分からレート指示を受けるとして構成されることができる。たとえば、コーディング方式検出器60は、ビットレートを指示するパケットタイプインジケータを多重サブレイヤから受けるように構成されることができる。代替として、コーディング方式検出器60は、フレームエネルギーなどの1つまたは複数のパラメータから符号化フレームのビットレートを判断するように構成されることができる。適用例によっては、コーディングシステムは、符号化フレームのビットレートがコーディングモードをも指示するように、特定のビットレートに対してただ1つのコーディングモードを使用するように構成される。他の場合には、符号化フレームは、フレームが符号化される際のコーディングモードを特定する1つまたは複数のビットの組などの情報を含むことができる。そのような情報(「コーディングインデックス」とも呼ばれる)は、(たとえば、他の可能なコーディングモードには無効である値を指示することによって)明示的にまたは暗示的にコーディングモードを指示することができる。

30

40

【0038】

図4bは、コーディング方式検出器60によって生成されたコーディング方式指示が、フレームデコーダ70a~70pのうち1つを選択するようにオーディオデコーダAD20の1対の選択器90aと90bを制御するために使用される一例を示している。オーディオデコーダAD20のソフトウェアまたはファームウェアの実装形態は、フレームデコーダのある1つまたは別の1つに実行のフローを導くためにコーディング方式指示を使用でき、そのような実装形態は、選択器90a及び/または選択器90bの類似物を含ま

50

なくて良いということに留意されたい。

【0039】

図5 aは、フレームエンコーダ30 a、30 bの実装形態32 a、32 bを含むマルチモードオーディオエンコーダAE20の実装形態AE22のブロック図を示している。この例では、コーディング方式選択器20の実装形態22は、オーディオ信号S100のアクティブフレームを非アクティブフレームと区別するように構成される。そのような動作は、「ボイスアクティビティ検出」とも呼ばれ、コーディング方式選択器22は、ボイスアクティビティ検出器を含むように実装され得る。たとえば、コーディング方式選択器22は、アクティブフレームに対しては（アクティブフレームエンコーダ32 aの選択を指示する）ハイであり、非アクティブフレームに対しては（非アクティブフレームエンコーダ32 bの選択を指示する）ローである、バイナリ値コーディング方式選択信号を出力するように構成されること、またはその逆に構成されることも可能である。この例では、コーディング方式選択器22によって生成されたコーディング方式選択信号は、オーディオ信号S100の各フレームがアクティブフレームエンコーダ32 a（たとえば、CELPエンコーダ）及び非アクティブフレームエンコーダ32 b（たとえば、NELPエンコーダ）のうちの選択された1つによって符号化されるように、選択器50 a、50 bの実装形態52 a、52 bを制御するために使用される。

10

【0040】

コーディング方式選択器22は、フレームエネルギー、信号対ノイズ比（「SNR」）、周期性、スペクトル分布（たとえば、スペクトル傾斜）、及び/またはゼロ交差レートなど、フレームのエネルギー及び/またはスペクトル成分の1つまたは複数の特性に基づいてボイスアクティビティ検出を実行するように構成されることができる。コーディング方式選択器22は、そのような特性の値を計算するように、オーディオエンコーダAE22の1つまたは複数の他のモジュールからそのような特性の値を受け取るように、及び/またはオーディオエンコーダAE22を含むデバイス（たとえば、セルラー電話）の1つまたは複数の他のモジュールからそのような特性の値を受け取るように実装されることができる。そのような検出は、そのような特性の値または大きさを閾値と比較すること、及び/または（たとえば、先行フレームに対する）そのような特性の変化の大きさを閾値と比較することを含むことができる。たとえば、コーディング方式選択器22は、現在フレームのエネルギーを評価し、エネルギー値が閾値よりも小さい（あるいは、それ以下である）場合にフレームを非アクティブとして分類するように構成されることができる。そのような選択器は、フレームエネルギーをフレームサンプルの平方和として計算するように構成されることができる。

20

30

【0041】

コーディング方式選択器22の別の実装形態は、低周波帯域（たとえば、300 Hz ~ 2 kHz）及び高周波帯域（たとえば、2 kHz ~ 4 kHz）の各々における現在フレームのエネルギーを評価し、各帯域のエネルギー値がそれぞれの閾値よりも小さい（あるいは、それ以下である）場合にフレームが非アクティブであることを指示するように構成される。そのような選択器は、フレームに通過帯域フィルタを適用し、フィルタ処理されたフレームのサンプルの平方和を計算することによって帯域におけるフレームエネルギーを計算するように構成されることができる。そのようなボイスアクティビティ検出動作の一例は、www.3gpp2.orgにおいてオンラインで入手可能な第3世代パートナーシッププロジェクト2（「3GPP2」）標準文書C.S0014-C、v1.0（2007年1月）の第4.7節に記載されている。

40

【0042】

追加または代替として、ボイスアクティビティ検出動作は、1つまたは複数の前のフレーム及び/または1つまたは複数の後続のフレームからの情報に基づかれることができる。たとえば、2つ以上のフレームにわたって平均化されたフレーム特性の値に基づいてフレームをアクティブまたは非アクティブとして分類するようにコーディング方式選択器22を構成することが望ましい場合がある。前のフレームからの情報（たとえば、バックグ

50

ラウンドノイズレベル、SNR)に基づく閾値を使用してフレームを分類するようにコーディング方式選択器22を構成することが望ましい場合がある。また、アクティブフレームから非アクティブフレームへのオーディオ信号S100における遷移に後続する第1のフレームのうちの1つまたは複数をアクティブとして分類するようにコーディング方式選択器22を構成することが望ましい場合もある。遷移の後にそのような様式で前の分類状態を継続する行為は、「ハングオーバー」とも呼ばれる。

【0043】

図5bは、フレームエンコーダ30c、30dの実装形態32c、32dを含むマルチモードオーディオエンコーダAE20の実装形態AE24のブロック図を示している。この例では、コーディング方式選択器20の実装形態24は、オーディオ信号S100のスピーチフレームを非スピーチフレーム(たとえば、音楽)と区別するように構成される。たとえば、コーディング方式選択器24は、スピーチフレームに対しては(CELPエンコーダなどのスピーチフレームエンコーダ32cの選択を指示する)ハイであり、非スピーチフレームに対しては(MDCTエンコーダなどの非スピーチフレームエンコーダ32dの選択を指示する)ローである、バイナリ値コーディング方式選択信号を出力するように構成されること、またはその逆に構成されることも可能である。そのような分類は、フレームエネルギー、ピッチ、周期性、スペクトル分布(たとえば、スペクトル傾斜、LPC係数、線スペクトル周波数(「LSF」))、及び/またはゼロ交差レートなど、フレームのエネルギー及び/またはスペクトル成分の1つまたは複数の特性に基づかれることができる。コーディング方式選択器24は、そのような特性の値を計算するように、オーディオエンコーダAE24の1つまたは複数の他のモジュールからそのような特性の値を受け取るように、及び/またはオーディオエンコーダAE24を含むデバイス(たとえば、セルラー電話)の1つまたは複数の他のモジュールからそのような特性の値を受け取るように実装されることができ、そのような分類は、そのような特性の値または大きさを閾値と比較すること、及び/またはそのような特性の(たとえば、先行フレームに対する)変化の大きさを閾値と比較することを含むことができる。そのような分類は、隠れマルコフモデルなどの多状態モデルを更新するために使用され得る、1つまたは複数の前のフレーム及び/または1つまたは複数の後続のフレームからの情報に基づかれることができる。

【0044】

この例では、コーディング方式選択器24によって生成されたコーディング方式選択信号は、オーディオ信号S100の各フレームがスピーチフレームエンコーダ32c及び非スピーチフレームエンコーダ32dのうちの選択された1つによって符号化されるように、選択器52a、52bを制御するために使用される。図6aは、スピーチフレームエンコーダ32cのRCELP実装形態34cと非スピーチフレームエンコーダ32dのMDCT実装形態34dとを含むオーディオエンコーダAE24の実装形態AE25のブロック図を示している。

【0045】

図6bは、フレームエンコーダ30b、30d、30e、30fの実装形態32b、32d、32e、32fを含むマルチモードオーディオエンコーダAE20の実装形態AE26のブロック図を示している。この例では、コーディング方式選択器20の実装形態26は、オーディオ信号S100のフレームを、有声音スピーチ、無声音スピーチ、非アクティブスピーチ、及び非スピーチとして分類するように構成されることができ、そのような分類は、上記のようにフレームのエネルギー及び/またはスペクトル成分の1つまたは複数の特性に基づかれることができ、そのような特性の値または大きさを閾値と比較すること、及び/またはそのような特性の(たとえば、先行フレームに対する)変化の大きさを閾値と比較することを含むことができ、1つまたは複数の前のフレーム及び/または1つまたは複数の後続のフレームからの情報に基づかれることができる。コーディング方式選択器26は、そのような特性の値を計算するように、オーディオエンコーダAE26の1つまたは複数の他のモジュールからそのような特性の値を受け取るように、及び/また

10

20

30

40

50

はオーディオエンコーダ A E 2 6 を含むデバイス（たとえば、セルラー電話）の 1 つまたは複数の他のモジュールからそのような特性の値を受けるといえるように実装されることができ。この例では、コーディング方式選択器 2 6 によって生成されたコーディング方式選択信号は、オーディオ信号 S 1 0 0 の各フレームが、有声音フレームエンコーダ 3 2 e（たとえば、C E L P またはリラクスト C E L P（「R C E L P」）エンコーダ）、無声音フレームエンコーダ 3 2 f（たとえば、N E L P エンコーダ）、非スピーチフレームエンコーダ 3 2 d、及び非アクティブフレームエンコーダ 3 2 b（たとえば、低レート N E L P エンコーダ）のうちの選択された 1 つによって符号化されるように、選択器 5 0 a、5 0 b の実装形態 5 4 a、5 4 b を制御するために使用される。

【 0 0 4 6 】

オーディオエンコーダ A E 1 0 によって生成された符号化フレームは、一般に、オーディオ信号の対応するフレームが再構成されることができるパラメータ値の組を含む。このパラメータ値の組は、一般に、周波数スペクトルにわたるフレーム内のエネルギーの分散の記述などのスペクトル情報を含む。そのようなエネルギーの分散は、フレームの「周波数包絡線」または「スペクトル包絡線」とも呼ばれる。フレームのスペクトル包絡線の記述は、対応するフレームを符号化するために使用される特定のコーディング方式に応じて異なる形態及び/または長さをもつことができる。オーディオエンコーダ A E 1 0 は、パケットのサイズ、フォーマット、及びコンテンツが、そのフレーム用に選択された特定のコーディング方式に対応するように、パラメータ値の組をパケット中に配置するように構成されたパケッタイザ（図示せず）を含むように実装されることができ。オーディオデコーダ A D 1 0 の対応する実装形態は、ヘッダ及び/または他のルーティング情報などパケット中の他の情報からパラメータ値の組を分離するように構成されたデパケッタイザ（図示せず）を含むように実装されることができ。

【 0 0 4 7 】

オーディオエンコーダ A E 1 0 などのオーディオエンコーダは、一般に、フレームのスペクトル包絡線の記述を、値の順序付きシーケンスとして計算するように構成される。いくつかの実装形態では、オーディオエンコーダ A E 1 0 は、各値が、対応する周波数における、または対応するスペクトル領域にわたる、信号の振幅または大きさを指示するように、順序付きシーケンスを計算するように構成される。そのような記述の一例は、フーリエ変換または離散コサイン変換係数の順序付きシーケンスである。

【 0 0 4 8 】

他の実装形態では、オーディオエンコーダ A E 1 0 は、スペクトル包絡線の記述を、線形予測コーディング（L P C）分析の係数値の組など、コーディングモデルのパラメータ値の順序付きシーケンスとして計算するように構成される。L P C 係数値は、「フォルマント」とも呼ばれる、オーディオ信号の共鳴を示す。L P C 係数値の順序付きシーケンスは、一般に、1 つまたは複数のベクトルとして構成され、オーディオエンコーダは、これらの値をフィルタ係数または反射係数として計算するように実装されることができ。その組中の係数値の数は、L P C 分析の「次数」とも呼ばれ、（セルラー電話などの）通信デバイスのオーディオエンコーダによって実行される L P C 分析の典型的な次数の例は、4、6、8、1 0、1 2、1 6、2 0、2 4、2 8、及び 3 2 を含む。

【 0 0 4 9 】

オーディオエンコーダ A E 1 0 の実装形態を含むデバイスは、一般に、スペクトル包絡線の記述を伝送チャネル上で、量子化された形態で（たとえば、対応するルックアップテーブルまたは「コードブック」への 1 つまたは複数のインデックスとして）伝送するように構成される。したがって、オーディオエンコーダ A E 1 0 では、L P C 係数値の組を、線スペクトル対（「L S P」）、L S F、イミタンススペクトル対（「I S P」）、イミタンススペクトル周波数（「I S F」）、ケプストラム（cepstral）係数、または対数面積比（log area ratio）の値の組など、効率的に量子化できる形態で計算することが望ましい場合がある。オーディオエンコーダ A E 1 0 は、変換及び/または量子化の前の値の順序付きシーケンスに対して、知覚的重み付けまたは他のフィルタ処理演算など、1 つま

10

20

30

40

50

たは複数の他の処理演算を行うように構成されることもできる。

【0050】

場合によっては、フレームのスペクトル包絡線の記述は、（たとえば、フーリエ変換係数または離散コサイン変換係数の順序付きシーケンス中などに）フレームの時間的情報の記述をも含む。他の場合には、パケットのパラメータの組は、フレームの時間的情報の記述をも含むことができる。時間的情報の記述の形態は、フレームを符号化するために使用される特定のコーディングモードに依存することができる。いくつかのコーディングモードの場合（たとえば、CELPまたはPPPコーディングモード、及びいくつかのMDC Tコーディングモードの場合）、時間的情報の記述は、LPCモデル（たとえば、スペクトル包絡線の記述に従って構成された合成フィルタ）を励起するためにオーディオデコーダによって使用されるべき励起信号の記述を含むことができる。励起信号の記述は、通常、フレームに対するLPC分析演算の残差に基づかれる。励起信号の記述は、一般に、（たとえば、対応するコードブックへの1つまたは複数のインデックスとして）量子化された形態でパケット中に現れ、励起信号の少なくとも1つのピッチ成分に関する情報を含むことができる。たとえば、PPPコーディングモードの場合、符号化された時間的情報は、励起信号のピッチ成分を再現するためにオーディオデコーダによって使用されるべきプロトタイプ記述を含むことができる。RCELPまたはPPPコーディングモードの場合、符号化された時間的情報は1つまたは複数のピッチ周期推定値を含むことができる。ピッチ成分に関する情報の記述は、一般に、（たとえば、対応するコードブックへの1つまたは複数のインデックスとして）量子化された形態でパケット中に現れる。

10

20

【0051】

オーディオエンコーダAE10の実装形態の様々な要素は、意図された適用例に好適であると考えられるハードウェア、ソフトウェア、及び/またはファームウェアの任意の組合せで実施されることができる。たとえば、そのような要素は、たとえば同一チップ上に、またはチップセット中の2つ以上のチップの間に存在する電子デバイス及び/または光デバイスとして製造されることができる。そのようなデバイスの一例は、トランジスタまたは論理ゲートなどの論理要素の固定またはプログラマブルなアレイであり、これらの要素のいずれも1つまたは複数のそのようなアレイとして実装されることができる。これらの要素の任意の2つ以上、さらにはすべてが、同一アレイまたは複数のアレイ内に実装されることができる。そのような1つまたは複数のアレイは、1つまたは複数のチップ内（たとえば、2つ以上のチップを含むチップセット内）に実装されることができる。同じことは、対応するオーディオデコーダAD10の実装形態の様々な要素にも当てはまる。

30

【0052】

本明細書で説明するオーディオエンコーダAE10の様々な実装形態の1つまたは複数の要素は、全体または一部を、マイクロプロセッサ、組み込みプロセッサ、IPコア、デジタル信号プロセッサ、フィールドプログラマブルゲートアレイ（「FPGA」）、特定用途向け標準製品（「ASSP」）、及び特定用途向け集積回路（「ASIC」）などの論理要素の1つまたは複数の固定的なアレイまたはプログラマブルアレイ上で実行するように構成された命令の1つまたは複数の組として実装されることもできる。オーディオエンコーダAE10の実装形態の様々な要素のいずれも、1つまたは複数のコンピュータ（たとえば、「プロセッサ」とも呼ばれる、命令の1つまたは複数の組またはシーケンスを実行するようにプログラムされた1つまたは複数のアレイを含む機械）として実施されることもでき、これらの要素の任意の2つ以上、さらにはすべてが同一のそのようなコンピュータまたは複数のコンピュータ内に実装されることができる。同じことは、対応するオーディオデコーダAD10の様々な実装形態の要素にも当てはまる。

40

【0053】

オーディオエンコーダAE10の実装形態の様々な要素は、セルラー電話など、有線及び/または無線通信のためのデバイス、またはそのような通信機能をもつ他のデバイス内に含まれることができる。そのようなデバイスは、（VoIPなどの1つまたは複数のプロトコルを使用して）回線交換及び/またはパケット交換ネットワークと通信するよう

50

に構成されることができる。そのようなデバイスは、インターリーブ、パンクチャリング、畳込みコーディング、誤り訂正コーディング、ネットワークプロトコル（たとえば、イーサネット（登録商標）、TCP/IP、cdma2000）の1つまたは複数のレイヤのコーディング、1つまたは複数の無線周波（「RF」）搬送波及び/または光搬送波の変調、及び/またはチャンネルを介した1つまたは複数の被変調搬送波の送信などの動作を、符号化フレームを搬送する信号に対して実行するように構成されることができる。

【0054】

オーディオデコーダAD10の実装形態の様々な要素は、セルラー電話など、有線及び/またはワイヤレス通信を行うためのデバイス、またはそのような通信機能をもつ他のデバイス内に含まれることができる。そのようなデバイスは、（VoIPなどの1つまたは複数のプロトコルを使用して）回線交換及び/またはパケット交換ネットワークと通信するように構成されることができる。そのようなデバイスは、デインターリーブ、デパンクチャリング、畳込みデコーディング、誤り訂正デコーディング、ネットワークプロトコル（たとえば、イーサネット、TCP/IP、cdma2000）の1つまたは複数のレイヤのデコーディング、1つまたは複数の無線周波（「RF」）搬送波及び/または光搬送波の復調、及び/またはチャンネルを介した1つまたは複数の被変調搬送波の受信などの動作を、符号化フレームを搬送する信号に対して実行するように構成されることができる。

10

【0055】

オーディオエンコーダAE10の実装形態の1つまたは複数の要素では、装置が組み込まれているデバイスまたはシステムの別の動作に関係するタスクなど、装置の動作に直接関係しないタスクを実施し、あるいは装置の動作に直接関連しない命令の他の組を実行するために使用されることが可能である。また、オーディオエンコーダAE10の実装形態の1つまたは複数の要素では、共通の構造（たとえば、異なる要素に対応するコードの部分異なる時間の実行するために使用されるプロセッサ、異なる要素に対応するタスクを異なる時間の実行するように実行される命令の組、あるいは異なる要素向けの動作を異なる時間の実行する電子デバイス及び/または光デバイスの構成）を有することが可能である。同じことは、対応するオーディオデコーダAD10の様々な実装形態の要素にも当てはまる。そのような一例では、コーディング方式選択器20及びフレームエンコーダ30a~30pは、同一プロセッサ上で実行するように構成された命令の組として実装される。そのような別の例では、コーディング方式検出器60及びフレームデコーダ70a~70pは、同一プロセッサ上で実行するように構成された命令の組として実装される。フレームエンコーダ30a~30pのうち2つ以上は、異なる時間に行われる命令の1つまたは複数の組を共有するように実装でき；同じことは、フレームデコーダ70a~70pにも当てはまる。

20

30

【0056】

図7aは、オーディオ信号のフレームを符号化する方法M10のフローチャートを示している。方法M10は、エネルギー及び/またはスペクトル特性など、上述のようなフレーム特性の値を計算するタスクTE10を含む。計算された値に基づいて、タスクTE20は、（たとえば、コーディング方式選択器20の様々な実装形態に関して上述したように）コーディング方式を選択する。タスクTE30は、符号化フレームを生成するために（たとえば、フレームエンコーダ30a~30pの様々な実装形態に関して本明細書で説明したように）選択されたコーディング方式に従ってフレームを符号化する。随意のタスクTE40は、符号化フレームを含むパケットを発生させる。方法M10は、オーディオ信号の一連のフレーム中の各々を符号化するように構成される（たとえば、繰り返される）ことができる。

40

【0057】

方法M10の実装形態の典型的な適用例では、論理要素のアレイ（たとえば、論理ゲート）は、この方法の様々なタスクの1つ、複数、さらにはすべてを実行するように構成される。タスクの1つまたは複数（場合によってはすべて）は、論理要素（たとえば、プロ

50

セッサ、マイクロプロセッサ、マイクロコントローラ、または他の有限状態機械)のアレイを含む機械(たとえば、コンピュータ)によって可読及び/または実行可能であるコンピュータプログラム製品(たとえば、ディスク、フラッシュまたは他の不揮発性メモリカード、半導体メモリチップなどの1つまたは複数のデータ記憶媒体)に埋め込まれたコード(たとえば、命令の1つまたは複数の組)として実装されることもできる。方法M10の実装形態のタスクは、2つ以上のそのようなアレイまたは機械によって実行されることもできる。これらのまたは他の実装形態では、タスクは、セルラー電話など、ワイヤレス通信用のデバイス、またはそのような通信機能をもつ他のデバイス内で実行されることのできる。そのようなデバイスは、(V o I Pなどの1つまたは複数のプロトコルを使用して)回線交換ネットワーク及び/またはパケット交換ネットワークと通信するように構成されることができる。たとえば、そのようなデバイスは、符号化フレームを受信するように構成されたRF回路を含むことができる。

【0058】

図7bは、オーディオ信号のフレームを符号化するように構成された装置F10のブロック図を示している。装置F10は、上述のようにエネルギー及び/またはスペクトル特性など、フレーム特性の値を計算するための手段FE10を含む。装置F10はまた、(たとえば、コーディング方式選択器20の様々な実装形態に関して上述したように)計算された値に基づいてコーディング方式を選択するための手段FE20をも含む。装置F10はまた、符号化フレームを生成するために、(たとえば、フレームエンコーダ30a~30pの様々な実装形態に関して本明細書で説明したように)選択されたコーディング方式に従ってフレームを符号化するための手段FE30をも含む。装置F10はまた、符号化フレームを含むパケットを発生するための随意の手段FE40をも含む。装置F10は、オーディオ信号の一連のフレーム中の各々を符号化するように構成されることができる。

【0059】

RCELPコーディング方式などのPRコーディング方式の典型的な実装形態、またはPPPコーディング方式のPR実装形態では、相関に基づかれることができるピッチ推定演算を使用して、フレームまたはサブフレームごとに、ピッチ周期が1回推定される。フレームまたはサブフレームの境界におけるピッチ推定ウィンドウを中央に置くことが望ましい場合がある。サブフレームへのフレームの典型的な分割は、1フレーム当たり3つのサブフレーム(たとえば、160サンプルフレームの非重複サブフレームの各々について53、53及び54個のサンプル)、1フレーム当たり4つのサブフレーム、ならびに1フレーム当たり5つのサブフレーム(たとえば、160サンプルフレームにおいて5つの32サンプル非重複サブフレーム)を含む。ピッチ半分、ピッチ2倍、ピッチ3倍などの誤差を回避するために、推定されたピッチ周期の間の整合性について確認することが望ましい場合もある。ピッチ推定値の更新の間に、合成遅延輪郭を生成するためにピッチ周期が補間される。そのような補間は、サンプルごとに実行されるか、またはより少ない頻度で(たとえば、2つまたは3つのサンプルごとに)実行されるか、またはより多い頻度で(たとえば、サブサンプル分解能で)実行されることができる。たとえば、上記で参照した3GPP2文書C.S0014-Cに記載されているエンハンスト可変レートコーデック(「EVRC」)は、8倍オーバーサンプリングされた合成遅延輪郭を使用する。一般に、補間は、線形補間または双線形補間であり、1つまたは複数のポリフェーズ補間フィルタまたは別の適切な技法を使用して実行されることができる。RCELPなどのPRコーディング方式は、一般に、1/4レートなどの他のレートで符号化する実装形態も可能であるが、フルレートまたはハーフレートでフレームを符号化するように構成される。

【0060】

無声音フレームとともに連続ピッチ輪郭を使用することは、バジングなどの望ましくないアーティファクトを生じることがある。したがって、無声音フレームの場合、各サブフレーム内で一定のピッチ周期を使用して、サブフレーム境界において別の一定のピッチ周期に急激に切り替えることが望ましい場合がある。そのような技法の典型的な例は、40

ミリ秒ごとに繰り返される（8 kHz サンプリングレートで）20個のサンプルから40個のサンプルの範囲にわたるピッチ周期の擬似ランダムシーケンスを使用する。上述のボイスアクティビティ検出（「VAD」）動作は、有声音フレームを無声音フレームと区別するように構成されることができ、そのような動作は、一般に、スピーチ及び/または残差の自己相関、ゼロ交差レート、及び/または第1の反射係数などのファクタに基づかれる。

【0061】

PRコーディング方式（たとえば、RCELP）は、スピーチ信号のタイムワープを行う。「信号修正」とも呼ばれるこのタイムワープ演算では、信号の特徴（たとえば、ピッチパルス）間の元の時間関係が変更されるように、異なる時間シフトが信号の異なるセグメントに適用される。たとえば、信号のピッチ周期輪郭が合成ピッチ周期輪郭に一致するように信号をタイムワープすることが望ましい場合がある。時間シフトの値は、一般に、プラス数ミリ秒からマイナス数ミリ秒の範囲内にある。フォルマントの位置を変更するのを避けることが望ましい場合があるので、PRエンコーダ（たとえば、RCELPエンコーダ）では、スピーチ信号ではなく残差を修正するのが一般的である。しかしながら、以下で特許請求される構成は、スピーチ信号を修正するように構成されたPRエンコーダ（たとえば、RCELPエンコーダ）を使用して実施されることもできるということが明確に企図され、本明細書によって開示される。

10

【0062】

最良の結果は、連続ワーピング（warping）を使用して残差を修正することによって得られるであろうことが予想できる。そのようなワーピングは、サンプルごとに実行されることができ、あるいは残差のセグメント（たとえば、サブフレームまたはピッチ周期）を圧縮及び伸張することによって実行されることができ。

20

【0063】

図8は、平坦な遅延輪郭に対してタイムワープされる前（波形A）及び後（波形B）の残差の例を示している。この例では、垂直な点線間の間隔は規則的なピッチ周期を示している。

【0064】

連続ワーピングは、非常に計算集約的なので、携帯用、組込み型、リアルタイム、及び/または電池式の適用例では実行できない場合がある。したがって、RCELPまたは他のPRエンコーダでは、時間シフトの量が各セグメントにわたって一定となるように残差のセグメントを時間シフトすることによって残差の区分的修正を行うことがより一般的である（が、以下で特許請求される構成は、連続ワーピングを使用して、スピーチ信号を修正するか、または残差を修正するように構成されたRCELPまたは他のPRエンコーダを使用して実施されることもできることが明確に企図され、本明細書によって開示される）。そのような動作は、各ピッチパルスがターゲット残差中の対応するピッチパルスに一致するようにセグメントをシフトすることによって現在の残差を修正するように構成されることができ、上記ターゲット残差は、信号の前のフレーム、サブフレーム、シフトフレーム、または他のセグメントからの修正残差に基づかれる。

30

【0065】

図9は、区分的修正の前（波形A）及び後（波形B）の残差の例を示している。この図において、点線は、太線で示したセグメントが残差の残部（rest）に関してどのようにして右にシフトされるのかを示している。（たとえば、各シフトセグメントが2つ以上のピッチパルスを含まないように）各セグメントの長さはピッチ周期よりも短いことが望ましい場合がある。セグメント境界がピッチパルスにおいて生じないようにする（たとえば、セグメント境界を残差の低エネルギー領域に制限する）ことが望ましい場合もある。

40

【0066】

区分的修正手順は、一般に、（「シフトフレーム」とも呼ばれる）ピッチパルスを含むセグメントを選択することを含む。そのような動作の一例は、上記で参照したEVR文書C.S0014-Cの第4.11.6.2節（4-95～4-99ページ）に記載され

50

ており、この節は、一例として参照により本明細書に組み込まれる。一般に、最後に修正されたサンプル（または第1の無修正のサンプル）がシフトフレームの開始として選択される。E V R Cの例では、セグメント選択動作は、シフトすべきパルス（たとえば、まだ修正されていないサブフレームの領域における第1のピッチパルス）のための現在のサブフレーム残差を探索し、このパルスの位置に対してシフトフレームの終端を設定する。シフトフレーム選択動作（及び区分的修正手順の後続の動作）が単一のサブフレームに対して数回実行されることができるよう、サブフレームは複数のシフトフレームを含むことができる。

【0067】

区分的修正手順は、一般に、残差を合成遅延輪郭に一致させる動作を含む。そのような動作の一例は、上記で参照したE V R C文書C . S 0 0 1 4 - Cの第4 . 1 1 . 6 . 3節（4 - 9 9 ~ 4 - 1 0 1ページ）に記載されており、この節は、一例として参照により本明細書に組み込まれる。この例は、（たとえば、一例として参照により本明細書に組み込まれる、上記で参照したE V R C文書C . S 0 0 1 4 - Cの第4 . 1 1 . 6 . 1節（4 - 9 5ページ）に記載されているように、）バッファから前のサブフレームの修正残差を検索し、それを遅延輪郭にマッピングすることによってターゲット残差を発生させる。この例では、一致判定演算は、選択されたシフトフレームのコピーをシフトすることによって一時修正残差を発生し、その一時修正残差とターゲット残差との間の相関に従って最適なシフトを決定し、その最適なシフトに基づいて時間シフトを計算する。時間シフトは一般に蓄積された値であり、したがって（たとえば、上記の参照により組み込まれる第4 . 1 1 . 6 . 3節の第4 . 1 1 . 6 . 3 . 4部に記載されているように）時間シフトを計算する演算は、蓄積された時間シフトを最適なシフトに基づいて更新することを含む。

【0068】

現在の残差の各シフトフレームについて、区分的修正は、シフトフレームに対応する現在の残差のセグメントに、対応する計算された時間シフトを適用することによって、達成される。そのような修正動作の一例は、上記で参照したE V R C文書C . S 0 0 1 4 - Cの第4 . 1 1 . 6 . 4節（4 - 1 0 1ページ）に記載されており、この節は、一例として参照により本明細書に組み込まれる。一般に、時間シフトは、分数である値を有し、したがって修正手順は、サンプリングレートよりも高い分解能で実行される。そのような場合、1つまたは複数の多相補間フィルタまたは別の適切な技法を使用して実行され得る線形補間または双線形補間などの補間を使用して、残差の対応するセグメントに時間シフトを適用することが望ましい場合がある。

【0069】

図10は、概略構成によるR C E L P符号化の方法R M 1 0 0（たとえば、方法M 1 0のタスクT E 3 0のR C E L P実装形態）のフローチャートを示している。方法R M 1 0 0は、現在フレームの残差を計算するタスクR T 1 0を含む。タスクR T 1 0は、一般に、オーディオ信号S 1 0 0など、（前処理されることがある）サンプリングされたオーディオ信号を受信するように構成される。タスクR T 1 0は、一般に、線形予測コーディング（「L P C」）分析演算を含むように実装され、線スペクトル対（「L S P」）などL P Cパラメータの組を生成するように構成されることができ、タスクR T 1 0は、1つまたは複数の知覚的重み付け及び/または他のフィルタ処理演算など、他の処理演算を含むこともできる。

【0070】

方法R M 1 0 0はまた、オーディオ信号の合成遅延輪郭を計算するタスクR T 2 0と、発生した残差からシフトフレームを選択するタスクR T 3 0と、選択されたシフトフレーム及び遅延輪郭から情報に基づいて時間シフトを計算するタスクR T 4 0と、計算された時間シフトに基づいて現在フレームの残差を修正するタスクR T 5 0とを含む。

【0071】

図11は、R C E L P符号化方法R M 1 0 0の実装形態R M 1 1 0のフローチャートを示している。方法R M 1 1 0は、時間シフト計算タスクR T 4 0の実装形態R T 4 2を含

10

20

30

40

50

む。タスクRT42は、前のサブフレームの修正残差を現在のサブフレームの合成遅延輪郭にマッピングするタスクRT60と、（たとえば、選択されたシフトフレームに基づいて）一時修正残差を発生させるタスクRT70と、（たとえば、一時修正残差とマッピングされた過去修正残差の対応するセグメントとの間の相関に基づいて）時間シフトを更新するタスクRT80とを含む。方法RM100の実装形態は、方法M10の実装形態内（たとえば、符号化タスクTE30内）に含められることができ、上記のように、論理要素（たとえば、論理ゲート）のレイは、その方法の様々なタスクのうちの1つ、複数、さらにはすべてを実行するように構成されることができる。

【0072】

図12aは、RCELPフレームエンコーダ34cの実装形態RC100のブロック図を示している。エンコーダRC100は、（たとえば、LPC分析演算に基づいて）現在フレームの残差を計算するように構成された残差発生器R10と、（たとえば、現在及び最近のピッチ推定値に基づいて）オーディオ信号S100の合成遅延輪郭を計算するように構成された遅延輪郭計算器R20とを含む。エンコーダRC100はまた、現在の残差のシフトフレームを選択するように構成されたシフトフレーム選択器R30と、（たとえば、一時修正残差に基づいて時間シフトを更新するために）時間シフトを計算するように構成された時間シフト計算器（time shift calculator）R40と、（たとえば、計算された時間シフトをシフトフレームに対応する残差のセグメントに適用するために）時間シフトに従って残差を修正するように構成された残差修正器（residual modifier）R50とを含む。

10

20

【0073】

図12bは、時間シフト計算器R40の実装形態R42を含むRCELPエンコーダRC100の実装形態RC110のブロック図を示している。計算器R42は、前のサブフレームの修正残差を現在のサブフレームの合成遅延輪郭にマッピングするように構成された過去修正残差マッパーR60と、選択されたシフトフレームに基づいて一時修正残差を発生させるように構成された一時修正残差発生器R70と、一時修正残差とマッピングされた過去修正残差の対応するセグメントとの間の相関に基づいて時間シフトを計算する（たとえば、更新する）ように構成された時間シフト更新器R80とを含む。エンコーダRC100及びRC110の要素の各々は、1つまたは複数のプロセッサによって実行するための論理ゲート及び/または命令の組など、対応するモジュールによって実装されることができ、オーディオエンコーダAE20などのマルチモードエンコーダは、エンコーダRC100のインスタンスまたはその実装形態を含むことができ、そのような場合、RCELPフレームエンコーダの要素の1つまたは複数（たとえば、残差発生器R10）は、他のコーディングモードを実行するように構成されたフレームエンコーダと共有されることができる。

30

【0074】

図13は、残差発生器R10の実装形態R12のブロック図を示している。発生器R12は、オーディオ信号S100の現在フレームに基づいてLPC係数値の組を計算するように構成されたLPC分析モジュール210を含む。変換ブロック220は、LPC係数値の組をLSFの組に変換するように構成され、量子化器230は、LPCパラメータSL10を生成するために（たとえば、1つまたは複数のコードブックインデックスとして）LSFを量子化するように構成される。逆量子化器240は、量子化LPCパラメータSL10から復号LSFの組を得るように構成され、逆変換ブロック250は、復号LSFの組から復号LPC係数値の組を得るように構成される。復号LPC係数値の組に従って構成された（分析フィルタとも呼ばれる）白色化フィルタ260は、LPC残差SR10を生成するためにオーディオ信号S100を処理する。残差発生器R10は、特定の適用例に好適と考えられる他の設計に従って実装されることもできる。

40

【0075】

時間シフトの値があるシフトフレームから次のシフトフレームに変化すると、間隙または重複が、シフトフレーム間の境界に生じることがあり、残差修正器R50またはタスク

50

RT50ではこの領域中の信号の一部を適宜に反復または省略することが望ましい場合がある。また、(たとえば、後続のフレームの残差に対する区分的修正手順の実行に使用されるべきターゲット残差を発生させるためのソースとして)バッファに修正残差を記憶するようにエンコーダRC100または方法RM100を実装することが望ましい場合がある。そのようなバッファは、時間シフト計算器R40(たとえば、過去修正残差マッパーR60)への、または時間シフト計算タスクRT40(たとえば、マッピングタスクRT60)への入力を提供するように構成されることができる。

【0076】

図12cは、そのような修正残差バッファR90と、バッファR90からの情報に基づいて時間シフトを計算するように構成された時間シフト計算器R40の実装形態R44と、を含むRCELPエンコーダRC100の実装形態RC105のブロック図を示している。図12dは、バッファR90のインスタンスと、バッファR90から過去修正残差を受けよう構成された過去修正残差マッパーR60の実装形態R62と、を含むRCELPエンコーダRC105とRCELPエンコーダRC110との実装形態RC115のブロック図を示している。

10

【0077】

図14は、オーディオ信号のフレームのRCELP符号化(たとえば、装置F10の手段FE30のRCELP実装形態)のための装置RF100のブロック図を示している。装置RF100は、残差RF10(たとえば、LPC残差)を発生させるための手段と、(たとえば、現在のピッチ推定値と前のピッチ推定値との間で線形補間または双線形補間を実行することによって)遅延輪郭RF20を計算するための手段とを含む。装置RF100はまた、(たとえば、次のピッチパルスの位置を特定することによって)シフトフレームRF30を選択するための手段と、(たとえば、一時修正残差とマッピングされた過去修正残差との間の相関に従って時間シフトを更新することによって)時間シフトRF40を計算するための手段と、(たとえば、シフトフレームに対応する残差のセグメントを時間シフトすることによって)残差RF50を修正するための手段とを含む。

20

【0078】

修正残差は、一般に、現在フレーム用の励起信号に対する固定のコードブックの寄与を計算するために使用される。図15は、そのような演算をサポートするための追加のタスクを含むRCELP符号化方法RM100の実装形態RM120のフローチャートを示している。タスクRT90は、前のフレームからの復号励起信号のコピーを保持する適応コードブック(「ACB」)を遅延輪郭にマッピングすることによってACBをワーブする。タスクRT100は、知覚領域におけるACB寄与を得るために、ワーブされたACBに対し、現在のLPC係数値に基づいてLPC合成フィルタを適用し、タスクRT110は、知覚領域における現在の修正残差を得るために、現在の修正残差に対し、現在のLPC係数値に基づいてLPC合成フィルタを適用する。たとえば、上記で参照した3GPP2EVR文書C.S0014-Cの第4.11.4.5節(4-84~4-86ページ)に記載されているように、タスクRT100及び/またはタスクRT110では、重み付きLPC係数値の組に基づくLPC合成フィルタを適用することが望ましい場合がある。タスクRT120は、固定のコードブック(「FCB」)探索のためのターゲットを得るために2つの知覚領域信号間の差を計算し、タスクRT130は、励起信号に対するFCB寄与を得るためにFCB探索を実行する。上記のように、論理要素(たとえば、論理ゲート)のレイは、方法RM100のこの実装形態の様々なタスクのうちの1つ、複数、さらにはすべてを実行するように構成されることができる。

30

40

【0079】

一般に、RCELPコーディング方式を含む現代のマルチモードコーディングシステム(たとえば、オーディオエンコーダAE25の実装形態を含むコーディングシステム)は、無声音フレーム(たとえば、発話摩擦音)と背景ノイズのみを含むフレームとに一般に使用されるノイズ励起線形予測(「NELP」)など、1つまたは複数の非RCELPコーディング方式をも含む。非RCELPコーディング方式の他の例は、一般により高い有

50

声音フレームに使用される、プロトタイプ波形補間（「PWI」）とプロトタイプピッチ周期（「PPP」）などその変形態とを含む。RCELPコーディング方式がオーディオ信号のフレームを符号化するために使用され、且つ、非RCELPコーディング方式がオーディオ信号の隣接フレームを符号化するために使用される場合、合成波形中に不連続性が生じることがある。

【0080】

隣接フレームからのサンプルを使用してフレームを符号化することが望ましい場合がある。そのような方法でフレーム境界にわたって符号化することは、量子化誤差、切捨て、丸め、不必要な係数の廃棄などのファクタによりフレーム間に生じることがあるアーティファクトの知覚影響を低減する傾向がある。そのようなコーディング方式の一例は、修正離散コサイン変換（「MDCT」）コーディング方式である。

10

【0081】

MDCTコーディング方式は、音楽及び他の非スピーチサウンドを符号化するために通例使用される非PRコーディング方式である。たとえば、国際標準化機構（ISO）/国際電気標準会議（IEC）文書14496-3:1999で明記された、MPEG-4第3部としても知られるAdvanced Audio Codec（「AAC」）は、MDCTコーディング方式である。上記で参照した3GPP2 EVRC文書C.S0014-Cの第4.13節（4-145～4-151ページ）は別のMDCTコーディング方式を記載しており、この節は、一例として参照により本明細書に組み込まれる。MDCTコーディング方式は、その構造がピッチ周期に基づく信号としてではなく、シヌソイドの混合として周波数領域中でオーディオ信号を符号化するものであり、歌唱、音楽、及びシヌソイドの他の混合を符号化するのにより適している。

20

【0082】

MDCTコーディング方式は、2つ以上の連続フレームにわたる（すなわち、重複する）符号化ウィンドウを使用する。フレーム長がMの場合、MDCTは、2M個のサンプルの入力に基づいてM個の係数を生成する。したがって、MDCTコーディング方式の1つの特徴は、符号化フレームを表すために必要な変換係数の数を増加することなしに、変換ウィンドウに1つまたは複数のフレーム境界にわたることを許すことである。しかしながら、そのような重複コーディング方式が、PRコーディング方式を使用して符号化されたフレームに隣接するフレームを符号化するために使用された場合、不連続性が対応する復号されたフレーム中に生じることがある。

30

【0083】

M個のMDCT係数の計算は、次のように表されることができる。

【数1】

$$X(k) = \sum_{n=0}^{2M-1} x(n)h_k(n) \quad (\text{式1})$$

【0084】

40

ここで、

【数2】

$$h_k(n) = w(n) \sqrt{\frac{2}{M}} \cos \left[\frac{(2n+M+1)(2k+1)\pi}{4M} \right] \quad (\text{式2})$$

【0085】

ただし、 $k = 0, 1, \dots, M-1$ である。関数 $w(n)$ は、一般に、（プリンセンブラッドレイ条件とも呼ばれる）条件 $w^2(n) + w^2(n+M) = 1$ を満たすウィンドウとなるように選択される。

50

【 0 0 8 6 】

対応する逆MDC T演算は、次のように表されることができる。

【 数 3 】

$$\hat{x}(n) = \sum_{k=0}^{M-1} \hat{X}(k) h_k(n) \quad (\text{式3})$$

【 0 0 8 7 】

10

$n = 0, 1, \dots, 2M - 1$ であり、ここで、

【 数 4 】

$$\hat{X}(k)$$

【 0 0 8 8 】

はM個の受信されたMDC T係数であり、

【 数 5 】

$$\hat{x}(n)$$

20

【 0 0 8 9 】

は2M個の復号サンプルである。

【 0 0 9 0 】

図16は、MDC Tコーディング方式のための典型的な正弦ウィンドウ形状の3つの例を示している。プリンセン-ブラッドレイ条件を満たすこのウィンドウ形状は、次のように表されることができる。

【 数 6 】

$$w(n) = \sin\left(\frac{n\pi}{2M}\right) \quad (\text{式4})$$

30

【 0 0 9 1 】

但し、 $0 \leq n < 2M$ であり、ここで、 $n = 0$ は、現在フレームの第1のサンプルを示す。

【 0 0 9 2 】

図に示すように、現在フレーム（フレーム p）を符号化するために使用されるMDC Tウィンドウ804は、フレーム p 及びフレーム（p + 1）にわたって0でない値を有し、他の場合は0値にされる。前のフレーム（フレーム（p - 1））を符号化するために使用されるMDC Tウィンドウ802は、フレーム（p - 1）及びフレーム p にわたって0でない値を有し、他の場合は0値にされ、次のフレーム（フレーム（p + 1））を符号化するために使用されるMDC Tウィンドウ806は同様に構成される。デコーダにおいて、復号されたシーケンスは入力シーケンスと同様に重複され、追加される。図25aは、図16に示すウィンドウ804及び806を適用することから生じる重複追加領域の一例を示している。重複追加動作は、変換によってもたらされた誤差を消去し、 $w(n)$ がプリンセン-ブラッドレイ条件を満たし、量子化誤差がない場合に）完全再構成を可能にする。MDC Tが重複ウィンドウ関数を使用するとしても、重複追加の後に、フレーム当たりの入力サンプルの数がフレーム当たりのMDC T係数の数と同じであるので、それは臨海サンプリング（critically sampled）フィルタバンクである。

40

【 0 0 9 3 】

50

図17aは、MDCTフレームエンコーダ34dの実装形態ME100のブロック図を示している。残差発生器D10は、量子化LPCパラメータ（たとえば、上記の参照により組み込まれる3GPP2 EVRC文書C.S0014-Cの第4.13節の第4.13.2部に記載されている量子化LSP）を使用して残差を発生させるように構成されることができる。代替として、残差発生器D10は、非量子化LPCパラメータを使用して残差を発生させるように構成されることができる。RCELPエンコーダRC100及びMDCTエンコーダME100の実装形態を含むマルチモードコーダでは、残差発生器R10及び残差発生器D10は同じ構造として実装されることができる。

【0094】

エンコーダME100は、（たとえば、式1に上記したように $X(k)$ の式に従って）MDCT係数を計算するように構成されたMDCTモジュールD20をも含む。エンコーダME100は、量子化された符号化残差信号S30を生成するためにMDCT係数を処理するように構成された量子化器D30をも含む。量子化器D30は、正確な関数コンピュータ計算を使用してMDCT係数の因数（factorial）コーディングを実行するように構成されることができる。代替として、量子化器D30は、たとえば、U. Mittel他の「Low Complexity Factorial Pulse Coding of MDCT Coefficients Using Approximation of Combinatorial Functions」、IEEE ICASSP 2007、I-289~I-292ページ、及び上記の参照により組み込まれる3GPP2 EVRC文書C.S0014-Cの第4.13節の第4.13.5部に記載されている近似関数計算を使用してMDCT係数の因数コーディングを実行するように構成されることができる。図17aに示すように、MDCTエンコーダME100は、（たとえば、式3に上記したように

【数7】

$$\hat{x}(n)$$

【0095】

の式に従って）量子化信号に基づいて復号サンプルを計算するように構成された随意的逆MDCT（「IMDCT」）モジュールD40をも含むことができる。

【0096】

場合によっては、オーディオ信号S100の残差に対してではなく、オーディオ信号S100に対してMDCT演算を実行することが望ましい場合がある。LPC分析は、人間のスピーチの共鳴の符号化には好適であるが、音楽などの非スピーチ信号の特徴の符号化には効率的でない場合がある。図17bは、MDCTモジュールD20が、入力としてオーディオ信号S100のフレームを受けるとして構成される、MDCTフレームエンコーダ34dの実装形態ME200のブロック図を示している。

【0097】

図16に示す標準MDCT重複方式は、変換が実行され得る前に2M個のサンプルが利用可能である必要がある。そのような方式は、コーディングシステムに対して2M個のサンプル（すなわち、現在フレームのM個のサンプル+ルックアヘッド（lookahead）のM個のサンプル）という遅延制約を効果的に加える。CELP、RCELP、NELP、PWI、及び/またはPPPなど、マルチモードコーダの他のコーディングモードは、一般に、より短い遅延制約（たとえば、現在フレームのM個のサンプル+ルックアヘッドのM/2個、M/3個、またはM/4個のサンプル）に対して動作するように構成される。現代のマルチモードコーダ（たとえば、EVRC、SMV、AMR）では、コーディングモード間の切替えは、自動的に実行され、1秒間に数回行われることさえある。特に、特定の速度でパケットを生成するエンコーダを含む送信機を必要とする回線交換適用例では、そのようなコーダのコーディングモードでは、同じ遅延で動作することが望ましい場合がある。

10

20

30

40

50

【 0 0 9 8 】

図 1 8 は、M よりも短いルックアヘッド間隔を可能にするために（たとえば、図 1 6 に示す関数 $w(n)$ の代わりに）M D C T モジュール D 2 0 によって適用されるウィンドウ関数 $w(n)$ の一例を示している。図 1 8 に示す特定の例では、ルックアヘッド間隔は $M/2$ 個のサンプルの長さであるが、そのような技法は、L 個のサンプルの任意のルックアヘッドを可能にするように実装されることができ、ここで、L は 0 から M までの任意の値を有する。この技法（上記の参照により組み込まれる 3 G P P 2 E V R C 文書 C . S 0 0 1 4 - C の第 4 . 1 3 節の第 4 . 1 3 . 4 部（4 - 1 4 7 ページ）、及び「S Y S T E M S A N D M E T H O D S F O R M O D I F Y I N G A W I N D O W W I T H A F R A M E A S S O C I A T E D W I T H A N A U D I O S I G N A L」と題する米国特許公開第 2 0 0 8 / 0 0 2 7 7 1 9 号に記載されている例では、M D C T ウィンドウは、長さ $(M - L) / 2$ の 0 パッド領域で開始及び終了し、 $w(n)$ はプリンセン - ブラッドレイ条件を満たす。そのようなウィンドウ関数の 1 つの実装形態は、次のように表されることができる。

【 数 8 】

$$w(n) = \begin{cases} 0, & 0 \leq n < \frac{M-L}{2} \\ \sin\left[\frac{\pi}{2L}\left(n - \frac{M-L}{2}\right)\right], & \frac{M-L}{2} \leq n < \frac{M+L}{2} \\ 1, & \frac{M+L}{2} \leq n < \frac{3M-L}{2} \\ \sin\left[\frac{\pi}{2L}\left(3L+n - \frac{3M-L}{2}\right)\right], & \frac{3M-L}{2} \leq n < \frac{3M+L}{2} \\ 0, & \frac{3M+L}{2} \leq n < 2M \end{cases} \quad (\text{式 5})$$

【 0 0 9 9 】

ここで、 $n = (M - L) / 2$ は現在フレーム p の第 1 のサンプルであり、 $n = (3M - L) / 2$ は次のフレーム (p + 1) の第 1 のサンプルである。そのような技法に従って符号化された信号は（量子化誤差及び数値誤差がない場合に）完全再構成性質を保持する。L = M の場合、このウィンドウ関数は、図 1 6 に示すものと同じであり、L = 0 の場合、 $M/2 \leq n < 3M/2$ ならば、 $w(n) = 1$ であり、他の場合は、重複が生じないように 0 であることに留意されたい。

【 0 1 0 0 】

P R コーディング方式と非 P R コーディング方式とを含むマルチモードコーデでは、現在のコーディングモードが P R コーディングモードから非 P R コーディングモードに（またはその逆に）切り替わるフレーム境界にわたって合成波形が連続的であるようにすることが望ましい場合がある。コーディングモード選択器は、1 秒に数回、あるコーディング方式から別のコーディング方式に切り替わることができ、それらの方式間で知覚的に滑らかな遷移を行うことが望ましい。残念ながら、P R コーディング方式と非 P R コーディング方式との間の切替えは復号信号中に可聴クリックまたは他の不連続性を生じることがあるので、調整フレームと非調整フレームとの間の境界にわたるピッチ周期は、異常に大きくなったり、小さくなったりすることがある。さらに、上記のように、非 P R コーディング方式は、連続フレームにわたる重複追加ウィンドウを使用してオーディオ信号のフレームを符号化することができ、それらの連続フレーム間の境界での時間シフトの変化を回避することが望ましい場合がある。これらの場合、P R コーディング方式によって適用された時間シフトに従って非調整フレームを修正することが望ましい場合がある。

【 0 1 0 1 】

図19aは、概略構成に従ったオーディオ信号のフレームを処理する方法M100のフローチャートを示している。方法M100は、PRコーディング方式（たとえば、RCCELLPコーディング方式）に従って第1のフレームを符号化するタスクT110を含む。方法M100は、非PRコーディング方式（たとえば、MDC Tコーディング方式）に従ってオーディオ信号の第2のフレームを符号化するタスクT210をも含む。上記のように、第1及び第2のフレームの一方または両方は、そのような符号化の前及び/または後に知覚的に重み付けされ、及び/または別の方法で処理されることができる。

【0102】

タスクT110は、時間シフトTに従って第1の信号のセグメントを時間修正するサブタスクT120を含み、ここで第1の信号は第1のフレームに基づかれる（たとえば、第1の信号は第1のフレームまたは第1のフレームの残差である）。時間修正は、時間シフトによって、またはタイムワープによって実行されることができる。一実装形態では、タスクT120は、Tの値に従って時間的に前方または後方に（すなわち、フレームまたはオーディオ信号の別のセグメントに対して）セグメント全体を移動することによって、セグメントを時間シフトする。そのような動作は、断片（fractional）時間シフトを実行するためにサンプル値を補間することを含むことができる。別の実装形態では、タスクT120は、時間シフトTに基づいてセグメントをタイムワープする。そのような動作は、Tの値に従ってセグメントのあるサンプル（たとえば、第1のサンプル）を移動することと、Tの大きさよりも小さい大きさを有する値だけセグメントの別のサンプル（たとえば、最後のサンプル）を移動することと、を含むことができる。

【0103】

タスクT210は、時間シフトTに従って第2の信号のセグメントを時間修正するサブタスクT220を含み、ここで第2の信号は第2のフレームに基づかれる（たとえば、第2の信号は、第2のフレームまたは第2のフレームの残差である）。一実装形態では、タスクT220は、Tの値に従って時間的に前方または後方に（すなわち、フレームまたはオーディオ信号の別のセグメントに対して）セグメント全体を移動することによって、セグメントを時間シフトする。そのような動作は、断片時間シフトを実行するためにサンプル値を補間することを含むことができる。別の実装形態では、タスクT220は、時間シフトTに基づいてセグメントをタイムワープする。そのような動作は、セグメントを遅延輪郭にマッピングすることを含むことができる。たとえば、そのような動作は、Tの値に従ってセグメントのあるサンプル（たとえば、第1のサンプル）を移動することと、Tの大きさよりも小さい大きさを有する値だけセグメントの別のサンプル（たとえば、最後のサンプル）を移動することと、を含むことができる。たとえば、タスクT120は、フレームまたは他のセグメントを、時間シフトTの値だけ短縮された（たとえば、Tが負の値の場合、延長された）対応する時間間隔にマッピングすることによってタイムワープすることができ、その場合、Tの値は、ワープされたセグメントの終了時に0にリセットされることができる。

【0104】

タスクT220が時間修正するセグメントは、第2の信号全体を含むことができるか、または、セグメントは、残差のサブフレーム（たとえば、初期サブフレーム）など、その信号のより短い部分とすることができる。一般に、タスクT220は、図17aに示す残差発生器D10の出力など非量子化残差信号のセグメントを（たとえば、オーディオ信号S100の逆LPCフィルタ処理の後に）時間修正する。しかしながら、タスクT220は、図17aに示す信号S40、またはオーディオ信号S100のセグメントなど、復号残差のセグメントを（たとえば、MDC T - IMD C T処理の後に）時間修正するように実装されることもできる。

【0105】

時間シフトTが、第1の信号を修正するために使用された最後の時間シフトであることが望ましい場合がある。たとえば、時間シフトTは、第1のフレームの残差の最後に時間シフトされたセグメントに適用された時間シフト、及び/または蓄積された時間シフトの

最新の更新から生じた値とすることができる。RCELPエンコーダRC100の実装形態は、タスクT110を実行するように構成されることができ、その場合、時間シフトTは、第1のフレームの符号化中にブロックR40またはブロックR80によって計算される最後の時間シフト値とすることができる。

【0106】

図19bは、タスクT110の実装形態T112のフローチャートを示している。タスクT112は、最新のサブフレームの修正残差など、前のサブフレームの残差からの情報に基づいて時間シフトを計算するサブタスクT130を含む。上述のように、RCELPコーディング方式では前のサブフレームの修正残差に基づかれるターゲット残差を発生させ、選択されたシフトフレームとターゲット残差の対応するセグメントとの間の一致に従って時間シフトを計算することが望ましい場合がある。

10

【0107】

図19cは、タスクT130の実装形態T132を含むタスクT112の実装形態T114のフローチャートを示している。タスクT132は、前の残差のサンプルを遅延輪郭にマッピングするタスクT140を含む。上述のように、RCELPコーディング方式では、前のサブフレームの修正残差を現在のサブフレームの合成遅延輪郭にマッピングすることによってターゲット残差を発生させることが望ましい場合がある。

【0108】

第2の信号を時間シフトし、第2のフレームを符号化するためのルックアヘッドとして使用される後続のフレームの任意の部分をも時間シフトするようにタスクT210を構成することが望ましい場合がある。たとえば、タスクT210では、(たとえば、MDC T及び重複ウィンドウに関して上述したように)時間シフトTを第2の(非PR)フレームの残差に適用し、第2のフレームを符号化するためのルックアヘッドとして使用される後続のフレームの残差の任意の部分にも適用することが望ましい場合がある。時間シフトTを、非PRコーディング方式(たとえば、MDC Tコーディング方式)を使用して符号化された任意の後続の連続フレームの残差に適用し、そのようなフレームに対応する任意のルックアヘッドセグメントに適用するように、タスクT210を構成することが望ましい場合もある。

20

【0109】

図25bは、2つのPRフレーム間の非PRフレームのシーケンス中の各々が、第1のPRフレームの最後のシフトフレームに適用された時間シフトによってシフトされる例を示している。この図では、実線は元のフレームの位置を経時的に示し、破線はフレームのシフトされた位置を示し、点線は元の境界とシフトされた境界との間の対応を示す。より長い垂直線はフレーム境界を示し、第1の短い垂直線は、第1のPRフレームの最後のシフトフレームの開始を示し(ピークはシフトフレームのピッチパルスを示す)、最後の短い垂直線はシーケンスの最終非PRフレーム用のルックアヘッドセグメントの終了を示す。一例では、PRフレームはRCELPフレームであり、非PRフレームはMDC Tフレームである。別の例では、PRフレームはRCELPフレームであり、非PRフレームのいくつかはMDC Tフレームであり、他の非PRフレームはNELPフレームまたはPWIフレームである。

30

40

【0110】

方法M100は、ピッチ推定値が現在の非PRフレームに利用できない場合に好適であり得る。しかしながら、ピッチ推定値が現在の非PRフレームに利用できる場合でも、方法M100を実行することが望ましい場合がある。(MDC Tウィンドウの場合など)連続フレーム間の重複追加を伴う非PRコーディング方式では、連続フレーム、任意の対応するルックアヘッド、及びフレーム間の任意の重複領域を同じシフト値だけシフトすることが望ましい場合がある。そのような整合性は、再構成されたオーディオ信号の品質の劣化を回避するのに役立つことができる。たとえば、MDC Tウィンドウなど重複領域に寄与するフレームの両方に同じ時間シフト値を使用することが望ましい場合がある。

【0111】

50

図20aは、MDC TエンコーダME100の実装形態ME110のブロック図を示している。エンコーダME110は、時間修正した残差信号S20を生成するために、残差発生器D10によって発生した残差信号のセグメントを時間修正するように構成された時間修正器(time modifier)TM10を含む。一実装形態では、時間修正器TM10は、Tの値に従って前方または後方にセグメント全体を移動することによって、セグメントを時間シフトするように構成される。そのような動作は、断片時間シフトを実行するためにサンプル値を補間することを含むことができる。別の実装形態では、時間修正器TM10は、時間シフトTに基づいてセグメントをタイムワープするように構成される。そのような動作は、遅延輪郭にセグメントをマッピングすることを含むことができる。たとえば、そのような動作は、Tの値に従ってセグメントのあるサンプル(たとえば、第1のサンプル)を移動することと、Tの大きさよりも小さい大きさを有する値だけ別のサンプル(たとえば、最後のサンプル)を移動することと、を含むことができる。たとえば、タスクT120は、フレームまたは他のセグメントを、時間シフトTの値だけ短縮された(たとえば、Tが負の値の場合、延長された)対応する時間間隔にマッピングすることによってタイムワープすることができ、その場合、Tの値は、ワープされたセグメントの終了時に0にリセットされることができる。上記のように、時間シフトTは、PRコーディング方式によって時間シフトされたセグメントに最新に適用された時間シフト、及び/またはPRコーディング方式によって蓄積された時間シフトの最新の更新から生じた値とすることができる。RC ELPエンコーダRC105とMDC TエンコーダME110の実装形態を含むオーディオエンコーダAE10の実装形態では、エンコーダME110は、時間修正された残差信号S20をバッファR90に記憶するように構成されることもできる。

【0112】

図20bは、MDC TエンコーダME200の実装形態ME210のブロック図を示している。エンコーダME200は、時間修正されたオーディオ信号S25を生成するために、オーディオ信号S100のセグメントを時間修正するように構成された時間修正器TM10のインスタンスを含む。上記のように、オーディオ信号S100は、知覚的に重み付けされ、及び/または別の方法でフィルタ処理されたデジタル信号とすることができる。RC ELPエンコーダRC105とMDC TエンコーダME210の実装形態を含むオーディオエンコーダAE10の実装形態では、エンコーダME210は、時間修正された残差信号S20をバッファR90に記憶するように構成されることもできる。

【0113】

図21aは、ノイズ注入モジュールD50を含むMDC TエンコーダME110の実装形態ME120のブロック図を示している。ノイズ注入モジュールD50は、(たとえば、上記の参照により組み込まれる3GPP2 EVRC文書C.S0014-Cの第4.13節の第4.13.7部(4-150ページ)に記載されている技法に従って)所定の周波数範囲内で量子化された符号化残差信号S30の0値要素とノイズを置換するように構成される。そのような動作は、残差線スペクトルのアンダーモデリング(undermodeling)中に生じることがあるトーンアーティファクト(tonal artifacts)の知覚を低減することによって、オーディオ品質を改善することができる。

【0114】

図21bは、MDC TエンコーダME110の実装形態ME130のブロック図を示している。エンコーダME130は、(たとえば、上記の参照により組み込まれる3GPP2 EVRC文書C.S0014-Cの第4.13節の第4.13.3部(4-147ページ)に記載されている技法に従って)残差信号S20の低周波フォルマント領域の知覚的重み付けを実行するように構成されたフォルマントエンファシスモジュールD60と、(たとえば、3GPP2 EVRC文書C.S0014-Cの第4.13節の第4.13.9部(4-151ページ)に記載されている技法に従って)知覚的重み付けを除去するように構成されたフォルマントデエンファシスモジュールD70と、を含む。

【0115】

図22は、MDC TエンコーダME120とME130の実装形態ME140のプロ

ック図を示している。MDC TエンコーダMD 1 1 0の他の実装形態は、残差発生器D 1 0と復号残差信号S 4 0との間の処理経路中に1つまたは複数の追加の動作を含むように構成されることができる。

【0 1 1 6】

図2 3 aは、概略構成に従ったオーディオ信号MM 1 0 0のフレームをMDC T符号化する方法（たとえば、方法M 1 0のタスクTE 3 0のMDC T実装形態）のフローチャートを示している。方法MM 1 0 0は、フレームの残差を発生させるタスクMT 1 0を含む。タスクMT 1 0は、一般に、オーディオ信号S 1 0 0など、（前処理されることがある）サンプリングされたオーディオ信号のフレームを受けるとして構成される。タスクMT 1 0は、一般に、線形予測コーディング（「LPC」）分析演算を含むように実装され、線スペクトル対（「LSP」）などLPCパラメータの組を生成するように構成されることができる。タスクMT 1 0は、1つまたは複数の知覚的重み付け及び/または他のフィルタ処理演算など、他の処理演算を含むこともできる。

10

【0 1 1 7】

方法MM 1 0 0は、発生された残差を時間修正するタスクMT 2 0を含む。一実装形態では、タスクMT 2 0は、Tの値に従って前方または後方にセグメント全体を移動して、残差のセグメントを時間シフトすることによって、残差を時間修正する。そのような動作は、断片時間シフトを実行するためにサンプル値を補間することを含むことができる。別の実装形態では、タスクMT 2 0は、時間シフトTに基づいて残差のセグメントをタイムワープすることによって残差を時間修正する。そのような動作は、遅延輪郭にセグメントをマッピングすることを含むことができる。たとえば、そのような動作は、Tの値に従ってセグメントのあるサンプル（たとえば、第1のサンプル）を移動することと、Tよりも小さい大きさを有する値だけ別のサンプル（たとえば、最後のサンプル）を移動することを含むことができる。時間シフトTは、PRコーディング方式によって時間シフトされたセグメントに最新に適用された時間シフト、及び/またはPRコーディング方式によって蓄積された時間シフトの最新の更新から生じた値とすることができる。RC E L P符号化方法RM 1 0 0とMDC T符号化方法MM 1 0 0の実装形態を含む符号化方法M 1 0の実装形態では、タスクMT 2 0は、（たとえば、次のフレームのためのターゲット残差を発生させるために方法RM 1 0 0によって使用することができるように）時間修正された残差信号S 2 0を修正残差バッファに記憶するように構成されることができる。

20

30

【0 1 1 8】

方法MM 1 0 0は、MDC T係数の組を生成するために、（たとえば、上記のようにX(k)のための式に従って）時間修正された残差に対してMDC T演算を実行するタスクMT 3 0を含む。タスクMT 3 0は、（たとえば、図1 6または図1 8に示すように）本明細書で説明するウィンドウ関数w(n)を適用するか、またはMDC T演算を実行するために別のウィンドウ関数またはアルゴリズムを使用することができる。方法MM 4 0は、因数コーディング、組合せ近似、切捨て、丸め、及び/または、特定の適用例に好適であると考えられる任意の他の量子化演算を使用してMDC T係数を量子化するタスクMT 4 0を含む。この例では、方法MM 1 0 0は、（たとえば、上記のように

40

【数 9】

$$\hat{x}(n)$$

【0 1 1 9】

のための式に従って）復号サンプルの組を得るために量子化係数に対してIMDC T演算を実行するように構成された随意のタスクMT 5 0をも含む。

【0 1 2 0】

方法MM 1 0 0の実装形態は、方法M 1 0の実装形態内（たとえば、符号化タスクTE 3 0内）に含められることができ、上記のように、論理要素（たとえば、論理ゲート）の

50

アレイは、その方法の様々なタスクのうちの一つ、複数、さらにはすべてを実行するように構成されることができる。方法M10が方法MM100と方法RM100の両方の実装形態を含む場合、残差計算タスクRT10と残差発生タスクMT10は、共同で演算を共有することができる（たとえば、LPC演算の順序のみが異なり）、さらには同じタスクとして実装できる。

【0121】

図23bは、オーディオ信号のフレームのMDC T符号化のための装置MF100（たとえば、装置F10の手段FE30のMDC T実装形態）のブロック図を示している。装置MF100は、（たとえば、上述のようにタスクMT10の実装形態を実行することによって）フレームFM10の残差を発生させるための手段を含む。装置MF100は、（たとえば、上述のようにタスクMT20の実装形態を実行することによって）発生した残差FM20を時間修正するための手段を含む。RC ELP符号化装置RF100とMDC T符号化装置MF100との実装を含む符号化装置F10の実装形態では、手段FM20は、（たとえば、次のフレームのためのターゲット残差を発生させるために、装置RF100によって使用することができるように）時間修正された残差信号S20を修正残差バッファに記憶するように構成されることもできる。装置MF100はまた、（たとえば、上述のようにタスクMT30の実装形態を実行することによって）MDC T係数の組を得るために時間修正された残差FM30に対してMDC T演算を実行するための手段と、（たとえば、上述のようにタスクMT40の実装形態を実行することによって）MDC T係数FM40を量子化するための手段と、を含む。装置MF100は、（たとえば、上述のようにタスクMT50を実行することによって）量子化係数FM50に対してIMDC T演算を実行するための随意の手段をも含む。

10

20

【0122】

図24aは、別の概略構成によるオーディオ信号のフレームを処理する方法M200のフローチャートを示している。方法M200のタスクT510は、非PRコーディング方式（たとえば、MDC Tコーディング方式）に従って第1のフレームを符号化する。方法M200のタスクT610は、PRコーディング方式（たとえば、RC ELPコーディング方式）に従ってオーディオ信号の第2のフレームを符号化する。

【0123】

タスクT510は、第1の時間シフトTに従って第1の信号のセグメントを時間修正するサブタスクT520を含み、ここで第1の信号は第1のフレームに基づかれる（たとえば、第1の信号は第1の（非PR）フレームまたは第1のフレームの残差である）。一例では、時間シフトTは、オーディオ信号中の第1のフレームに先行したフレームのRC ELP符号化中に計算されるような、蓄積された時間シフトの値（たとえば、最後に更新された値）である。タスクT520が時間修正するセグメントは、第1の信号全体を含むか、または、そのセグメントは、残差のサブフレーム（たとえば、最終サブフレーム）など、その信号のより短い部分とすることができる。一般に、タスクT520は、図17aに示す残差発生器D10の出力など非量子化残差信号を（たとえば、オーディオ信号S100の逆LPCフィルタ処理の後に）時間修正する。しかしながら、タスクT520は、図17aに示す信号S40、またはオーディオ信号S100のセグメントなど、復号残差のセグメントを（たとえば、MDC T - IMDC T処理の後に）時間修正するように実装されることもできる。

30

40

【0124】

一実装形態では、タスクT520は、Tの値に従って時間的に前方または後方に（すなわち、フレームまたはオーディオ信号の別のセグメントに対して）セグメント全体を移動することによって、セグメントを時間シフトする。そのような動作は、断片時間シフトを実行するためにサンプル値を補間することを含むことができる。別の実装形態では、タスクT520は、時間シフトTに基づいてセグメントをタイムワープする。そのような動作は、遅延輪郭にセグメントをマッピングすることを含むことができる。たとえば、そのような動作は、Tの値に従ってセグメントのあるサンプル（たとえば、第1のサンプル）を

50

移動することと、Tの大きさよりも小さい大きさを有する値だけセグメントの別のサンプル（たとえば、最後のサンプル）を移動することと、を含むことができる。

【0125】

タスクT520は、（たとえば、次のフレームのためのターゲット残差を発生させるために）以下に説明するタスクT620によって使用することができるように、時間修正された信号をバッファ（たとえば、修正残差バッファ）に記憶するように構成されることができる。タスクT520は、PR符号化タスクの他の状態メモリを更新するように構成されることもできる。タスクT520の1つのそのような実装形態は、適応コードブック（「ACB」）メモリへの復号残差信号S40などの復号量子化残差信号と、PR符号化タスク（たとえば、RCELP符号化方法RM120）のゼロ入力応答フィルタ状態と、を記憶する。

10

【0126】

タスクT610は、時間修正されたセグメントからの情報に基づいて第2の信号をタイムワープするサブタスクT620を含み、ここで第2の信号は第2のフレームに基づかれる（たとえば、第2の信号は、第2のPRフレームまたは第2のフレームの残差である）。たとえば、PRコーディング方式は、過去修正残差の代わりに、時間修正された（たとえば、時間シフトされた）セグメントを含む第1のフレームの残差を使用することによって、上述のように第2のフレームを符号化するように構成されたRCELPコーディング方式とすることができる。

20

【0127】

一実装形態では、タスクT620は、時間的に前方または後方に（すなわち、フレームまたはオーディオ信号の別のセグメントに対して）セグメント全体を移動することによって、第2の時間シフトをセグメントに適用する。そのような動作は、断片時間シフトを実行するためにサンプル値を補間することを含むことができる。別の実装形態では、タスクT620は、セグメントをタイムワープするもので、セグメントを遅延輪郭にマッピングすることを含むことができる。たとえば、そのような動作は、時間シフトに従ってセグメントのあるサンプル（たとえば、第1のサンプル）を移動することと、より小さい時間シフトだけセグメントの別のサンプル（たとえば、最後のサンプル）を移動することと、を含むことができる。

30

【0128】

図24bは、タスクT620の実装形態T622のフローチャートを示している。タスクT622は、時間修正されたセグメントからの情報に基づいて第2の時間シフトを計算するサブタスクT630を含む。タスクT622は、第2の信号のセグメントに（この例では、第2のフレームの残差に）第2の時間シフトを適用するサブタスクT640をも含む。

【0129】

図24cは、タスクT620の実装形態T624のフローチャートを示している。タスクT624は、オーディオ信号の遅延輪郭に時間修正されたセグメントのサンプルをマッピングするサブタスクT650を含む。上述のように、RCELPコーディング方式では、現在のサブフレームの合成遅延輪郭に前のサブフレームの修正残差をマッピングすることによってターゲット残差を発生させることが望ましい場合がある。この場合、RCELPコーディング方式は、時間修正されたセグメントを含む第1の（非RCELP）フレームの残差に基づくターゲット残差を発生させることによってタスクT650を実行するように構成されることができる。

40

【0130】

たとえば、そのようなRCELPコーディング方式は、現在フレームの合成遅延輪郭に、時間修正されたセグメントを含む第1の（非RCELP）フレームの残差をマッピングすることによって、ターゲット残差を発生させるように構成されることができる。RCELPコーディング方式は、ターゲット残差に基づいて時間シフトを計算し、上述のように、第2のフレームの残差をタイムワープするために計算された時間シフトを使用するよう

50

に構成されることもできる。図24dは、タスクT650と、時間修正されたセグメントのマッピングされたサンプルからの情報に基づいて第2の時間シフトを計算するタスクT630の実装形態T632と、タスクT640とを含む、タスクT622及びT624の実装形態T626のフローチャートを示している。

【0131】

上記のように、約300～3400HzのPSTN周波数範囲を超える周波数範囲を有するオーディオ信号を送信及び受信することが望ましい場合がある。そのような信号のコーディングに対する1つの手法は、(たとえば、拡張周波数範囲をカバーするようにPSTN範囲用のコーディングシステムをスケールリングすることによって)拡張周波数範囲全体を単一の周波数帯域として符号化する「フルバンド(full-band)」技法である。別の手法は、拡張周波数範囲中にPSTN信号からの情報を外挿すること(たとえば、PSTN範囲のオーディオ信号からの情報に基づいて、PSTN範囲を上回るハイバンド範囲用の励起信号を外挿すること)である。さらなる手法は、PSTN範囲の外部にあるオーディオ信号の情報(たとえば、3500～7000Hzまたは3500～8000Hzなどハイバンド周波数範囲用の情報)を別々に符号化する「スプリットバンド」技法である。スプリットバンドPRコーディング技法についての記述は、「TIME-WARPING FRAMES OF WIDEBAND VOCODER」と題する米国特許公開第2008/0052065号、及び「SYSTEMS, METHODS, AND APPARATUS FOR HIGH BAND TIME WARPING」と題する第2006/0282263号などの文献に記載されている。オーディオ信号の狭帯域部分とハイバンド部分の両方に方法M100及び/またはM200の実装形態を含むようにスプリットバンドコーディング技法を拡張することが望ましい場合がある。

10

20

【0132】

方法M100及び/またはM200は、方法M10の実装形態内で実行されることができ。たとえば、タスクT110及びT210(同様に、タスクT510及びT610)は、方法M10がオーディオ信号S100の連続フレームを処理するように実行するとき、タスクTE30の連続反復によって実行されることができ。方法M100及び/またはM200は、装置F10及び/または装置AE10(たとえば、装置AE20またはAE25)の実装形態によって実行されることができ。上記のように、そのような装置は、セルラー電話などの携帯型通信デバイス中に含まれることができる。そのような方法及び/または装置は、メディアゲートウェイなどインフラストラクチャ機器中に実装されることができ。

30

【0133】

説明した構成の前述の提示は、本明細書で開示した方法及び他の構造を当業者が製造または使用できるように与えたものである。本明細書で図示及び説明したフローチャート、ブロック図、状態図、及び他の構造は例にすぎず、これらの構造の他の変形態も開示の範囲内である。これらの構成に対する様々な変更が可能であり、本明細書で提示した一般的な原理は他の構成にも同様に適用できる。したがって、本開示は、上記に示した構成に限定されるものではなく、原開示の一部をなす、出願される添付の特許請求の範囲を含む、本明細書において任意の方法で開示された原理及び新規の特徴に合致する最も広い範囲を与えられるべきである。

40

【0134】

上記で参照したEVRC及びSMVコーデックに加えて、本明細書で説明するスピーチエンコーダ、スピーチ符号化の方法、スピーチデコーダ、及び/またはスピーチ復号の方法とともに使用される、またはそれらとともに使用するように適合されるコーデックの例は、文書ETSI TS 126 092 V6.0.0(欧州電気通信標準化機構(「ETSI」)、Sophia Antipolis Cedex、FR、2004年12月)に記載されている適応マルチレート(Adaptive Multi Rate)(「AMR」)スピーチコーデック;及び文書ETSI TS 126 192 V6.0.0(ETSI、2004年12月)に記載されているAMR広帯域スピーチコーデックを含む。

50

【0135】

情報及び信号は、様々な異なる技術及び技法のいずれかを使用して表すことができることを、当業者は理解されよう。たとえば、上記の説明全体にわたって言及されるデータ、命令、コマンド、情報、信号、ビット、及びシンボルは、電圧、電流、電磁波、磁界または磁性粒子、光場または光粒子、あるいはそれらの任意の組合せによって表されることができる。

【0136】

さらに、本明細書で開示した構成に関連して説明した様々な例示的な論理ブロック、モジュール、回路、及び動作は、電子ハードウェア、コンピュータソフトウェア、または両方の組合せとして実装され得ることを、当業者は理解されよう。そのような論理ブロック、モジュール、回路、及び動作は、本明細書で説明した機能を実行するように設計された、汎用プロセッサ、デジタル信号プロセッサ(「DSP」)、ASICまたはASSP、FPGAまたは他のプログラマブル論理デバイス、ディスクリートゲートまたはトランジスタロジック、個別のハードウェアコンポーネント、あるいはそれらの任意の組合せを用いて実装または実行されることができる。汎用プロセッサはマイクロプロセッサとすることができるが、代替として、プロセッサは、通常のプロセッサ、コントローラ、マイクロコントローラ、または状態機械とすることができる。プロセッサは、コンピュータ計算デバイスの組合せ、たとえば、DSPとマイクロプロセッサとの組合せ、複数のマイクロプロセッサ、DSPコアと連携する1つまたは複数のマイクロプロセッサ、あるいは任意の他のそのような構成としても実装されることができる。

10

20

【0137】

本明細書で説明した方法及びアルゴリズムのタスクは、ハードウェアで直接実施されるか、プロセッサによって実行されるソフトウェアモジュールで実施されるか、またはその2つの組合せで実施されることができる。ソフトウェアモジュールは、ランダムアクセスメモリ(「RAM」)、読取り専用メモリ(「ROM」)、フラッシュRAMなどの不揮発性RAM(「NVRAM」)、消去可能プログラマブルROM(「EPROM」)、電氣的消去可能プログラマブルROM(「EEPROM」)、レジスタ、ハードディスク、リムーバブルディスク、CD-ROM、または当技術分野で知られている任意の他の形態の記憶媒体中に存在することができる。例示的な記憶媒体は、プロセッサが記憶媒体から情報を読み取り、記憶媒体に情報を書き込むことができるように、プロセッサに結合される。代替として、記憶媒体は、プロセッサに一体化されることができる。プロセッサ及び記憶媒体は、ASIC中に存在することができる。ASICは、ユーザ端末内に存在することができる。代替として、プロセッサ及び記憶媒体は、ユーザ端末内に個別のコンポーネントとして存在することができる。

30

【0138】

本明細書で説明した構成の各々は、少なくとも部分的に、ハードワイヤード回路として、特定用途向け集積回路中に作製された回路構成として、あるいは、機械可読コードとして不揮発性記憶装置にロードされたファームウェアプログラムまたはデータ記憶媒体からロードされた、もしくはデータ記憶媒体中にロードされたソフトウェアプログラムとして実装されることができ、そのようなコードは、マイクロプロセッサまたは他のデジタル信号処理ユニットなど論理要素のアレイによって実行可能な命令である。データ記憶媒体は、(限定はしないが、ダイナミックまたはスタティックRAM、ROM、及び/またはフラッシュRAMを含む)半導体メモリ、または強誘電体メモリ、磁気抵抗メモリ、オボニックメモリ、ポリマーメモリ、もしくは位相変化メモリなどの記憶要素のアレイ;あるいは磁気ディスクまたは光ディスクなどのディスク媒体とすることができる。「ソフトウェア」という用語は、ソースコード、アセンブリ言語コード、機械コード、バイナリコード、ファームウェア、マクロコード、マイクロコード、論理要素のアレイによって実行可能な命令の任意の1つまたは複数の組またはシーケンス、及びそのような例の任意の組合せを含むものと理解されたい。

40

【0139】

50

本明細書で開示した方法 M 1 0、R M 1 0 0、M M 1 0 0、M 1 0 0、及び M 2 0 0 の実装形態は、論理要素のアレイ（たとえば、プロセッサ、マイクロプロセッサ、マイクロコントローラ、または他の有限状態機械）を含む機械によって読取り可能及び/または実行可能な命令の 1 つまたは複数の組として、（たとえば、上記に記載した 1 つまたは複数のデータ記憶媒体中で）有形に実施されることもできる。したがって、本開示は、上記に示した構成に限定されるものではなく、原開示の一部をなす、出願される添付の特許請求の範囲を含む、本明細書において任意の方法で開示された原理及び新規の特徴に合致する最も広い範囲を与えられるべきである。

【 0 1 4 0 】

本明細書で説明した装置の様々な実装形態（たとえば、A E 1 0、A D 1 0、R C 1 0 0、R F 1 0 0、M E 1 0 0、M E 2 0 0、M F 1 0 0）の要素は、たとえば、同一チップ上またはチップセット中の 2 つ以上のチップ上に存在する電子デバイス及び/または光デバイスとして作製されることができる。そのようなデバイスの一例は、トランジスタまたはゲートなど、論理要素の固定またはプログラマブルなアレイである。本明細書で説明した装置の様々な実装形態の 1 つまたは複数の要素は、全体または一部を、マイクロプロセッサ、埋込み型プロセッサ、I P コア、デジタル信号プロセッサ、F P G A、A S S P、及び A S I C など論理要素の 1 つまたは複数の固定またはプログラマブルなアレイ上で実行するように構成された命令の 1 つまたは複数の組として実装されることもできる。

【 0 1 4 1 】

本明細書で説明した装置の一実装形態の 1 つまたは複数の要素は、装置が組み込まれているデバイスまたはシステムの別の動作に関係するタスクなど、装置の動作に直接関係しないタスクを実施するため、あるいは装置の動作に直接関係しない命令の他の組を実行するために、使用することが可能である。また、そのような装置の実装形態の 1 つまたは複数の要素は、共通の構造（たとえば、異なる要素に対応するコードの部分異なる時間に実行するために使用されるプロセッサ、異なる要素に対応するタスクを異なる時間に実施するために実行される命令の組、あるいは、異なる要素向けの動作を異なる時間に実施する電子デバイス及び/または光デバイスの構成）を有することが可能である。

【 0 1 4 2 】

図 2 6 は、本明細書で説明したシステム及び方法を有するアクセス端末として使用され得るオーディオ通信のためのデバイス 1 1 0 8 の一例のブロック図を示している。デバイス 1 1 0 8 は、デバイス 1 1 0 8 の動作を制御するように構成されたプロセッサ 1 1 0 2 を含む。プロセッサ 1 1 0 2 は、方法 M 1 0 0 または M 2 0 0 の実装形態を実行するようにデバイス 1 1 0 8 を制御するように構成されることができる。デバイス 1 1 0 8 は、命令及びデータをプロセッサ 1 1 0 2 に供給するように構成され、R O M、R A M、及び/または N V R A M を含むことができる、メモリ 1 1 0 4 をも含む。デバイス 1 1 0 8 は、トランシーバ 1 1 2 0 を収容するハウジング 1 1 2 2 をも含む。トランシーバ 1 1 2 0 は、デバイス 1 1 0 8 と遠隔地との間のデータの送信及び受信をサポートする、送信機 1 1 1 0 と受信機 1 1 1 2 とを含む。デバイス 1 1 0 8 のアンテナ 1 1 1 8 は、ハウジング 1 1 2 2 に取り付けられ、トランシーバ 1 1 2 0 に電氣的に結合される。

【 0 1 4 3 】

デバイス 1 1 0 8 は、トランシーバ 1 1 2 0 によって受信された信号を検出し、信号のレベルを定量化するように構成された信号検出器 1 1 0 6 を含む。たとえば、信号検出器 1 1 0 6 は、総エネルギー、擬似ノイズチップ（pseudonoise chip）当たりのパイロットエネルギー（E b / N o と表される）、及び/または電力スペクトル密度などのパラメータの値を計算するように構成されることができる。デバイス 1 1 0 8 は、デバイス 1 1 0 8 の様々なコンポーネントを共に結合するように構成されたバスシステム 1 1 2 6 を含む。データバスに加えて、バスシステム 1 1 2 6 は、電力バス、制御信号バス、及び/またはステータス信号バスを含むことができる。デバイス 1 1 0 8 は、トランシーバ 1 1 2 0 によって受信された信号及び/またはトランシーバ 1 1 2 0 によって送信すべき信号を処理するように構成された D S P 1 1 1 6 をも含む。

10

20

30

40

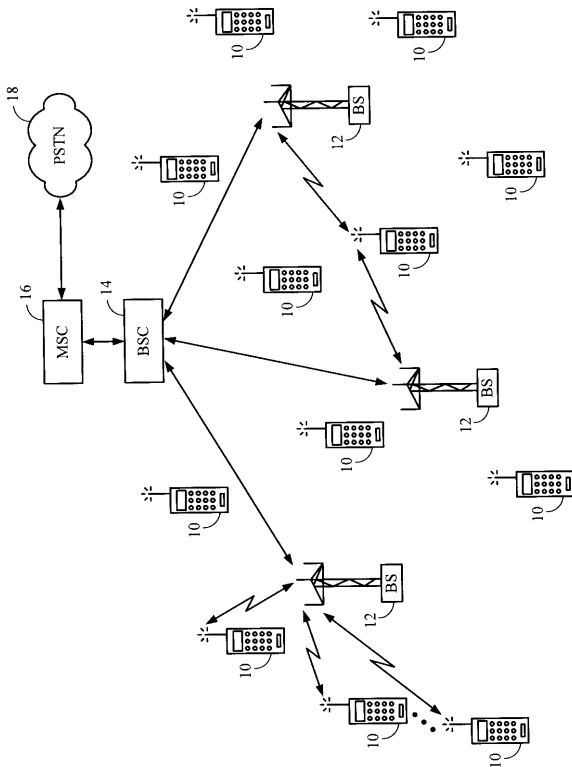
50

【0144】

この例では、デバイス1108は、いくつかの異なる状態のうちいずれか1つで動作するように構成されており、デバイスの現在の状態と、トランシーバ1120によって受信され、信号検出器1106によって検出された信号とに基づいてデバイス1108の状態を制御するように構成された状態変更器1114を含む。この例では、デバイス1108は、現在のサービスプロバイダが不十分であると判断し、異なるサービスプロバイダに転送するようにデバイス1108を制御するように構成されたシステム判断器1124をも含む。

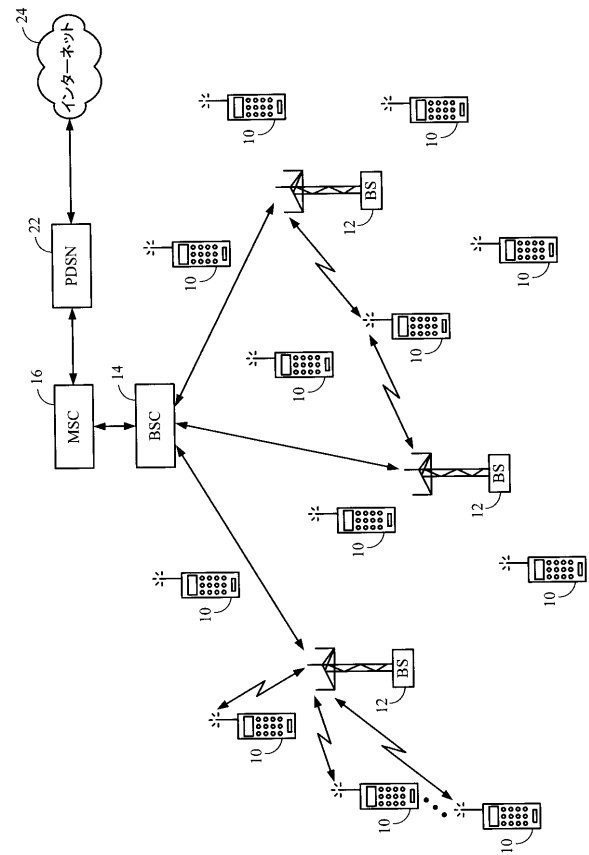
【図1】

図1

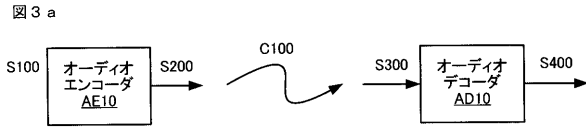


【図2】

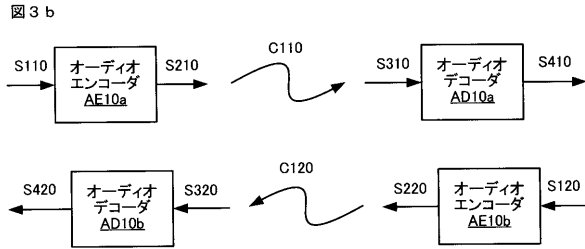
図2



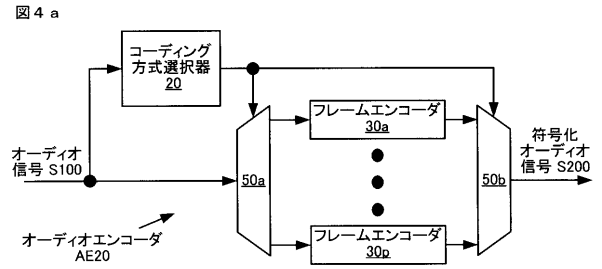
【 図 3 a 】



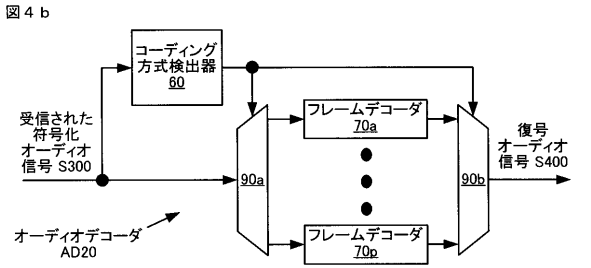
【 図 3 b 】



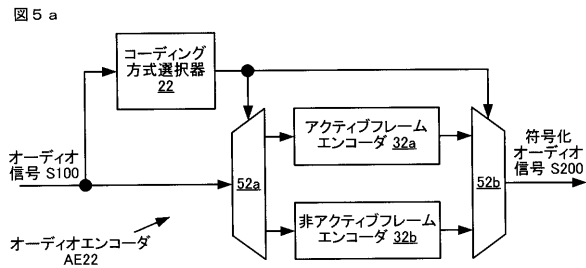
【 図 4 a 】



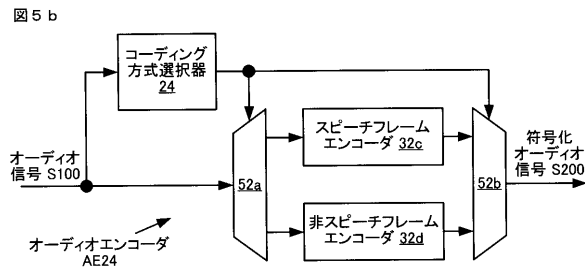
【 図 4 b 】



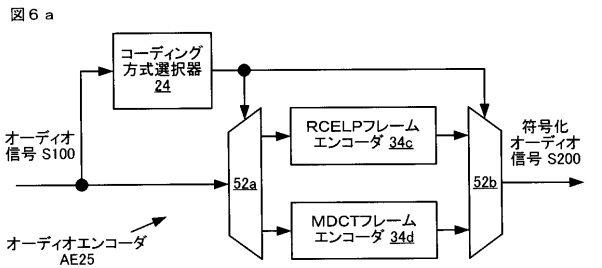
【 図 5 a 】



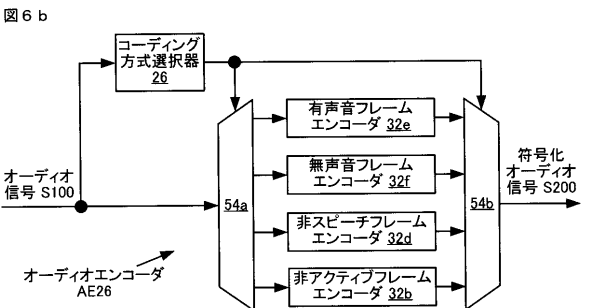
【 図 5 b 】



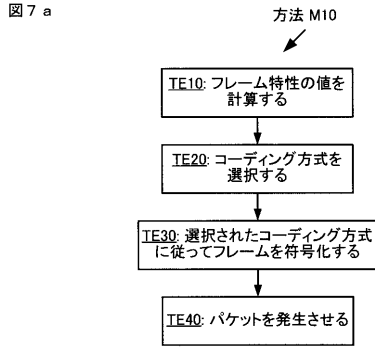
【 図 6 a 】



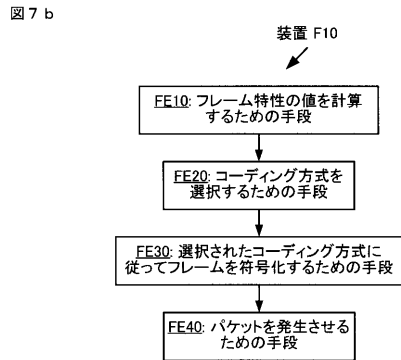
【 図 6 b 】



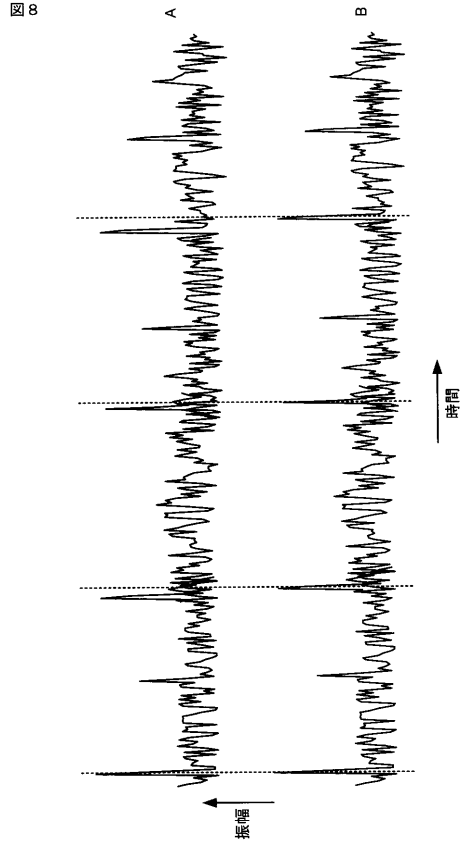
【 図 7 a 】



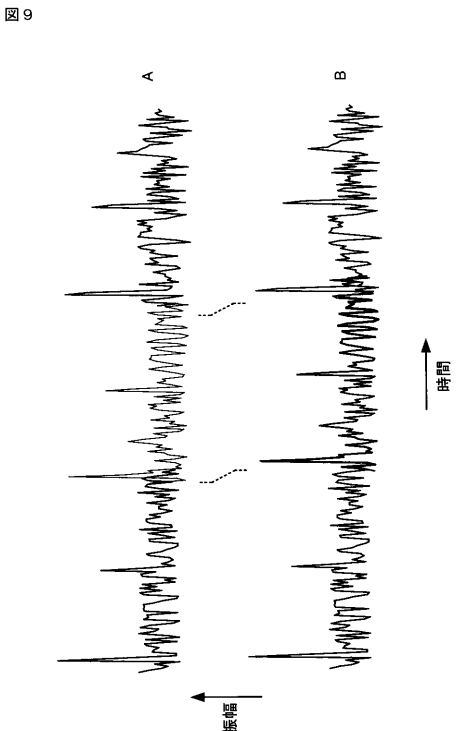
【 図 7 b 】



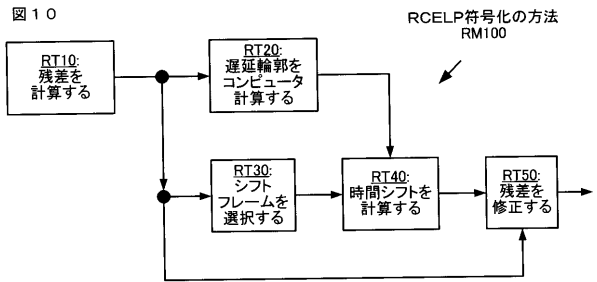
【 図 8 】



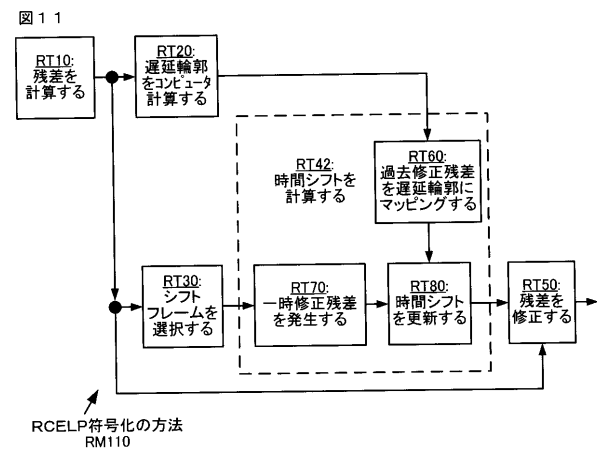
【 図 9 】



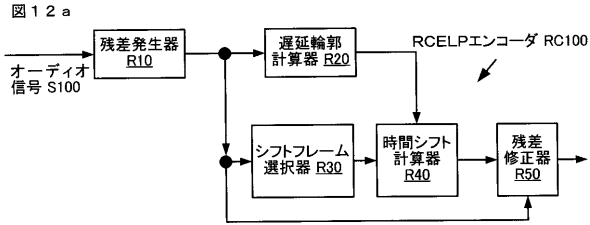
【 図 10 】



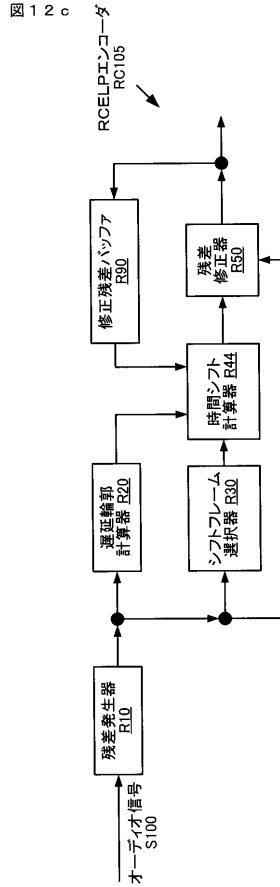
【 図 11 】



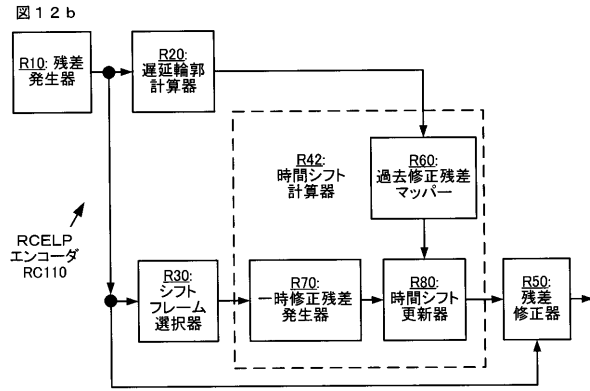
【図12a】



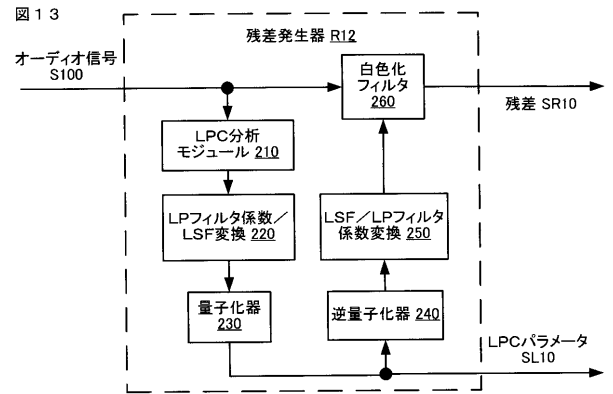
【図12c】



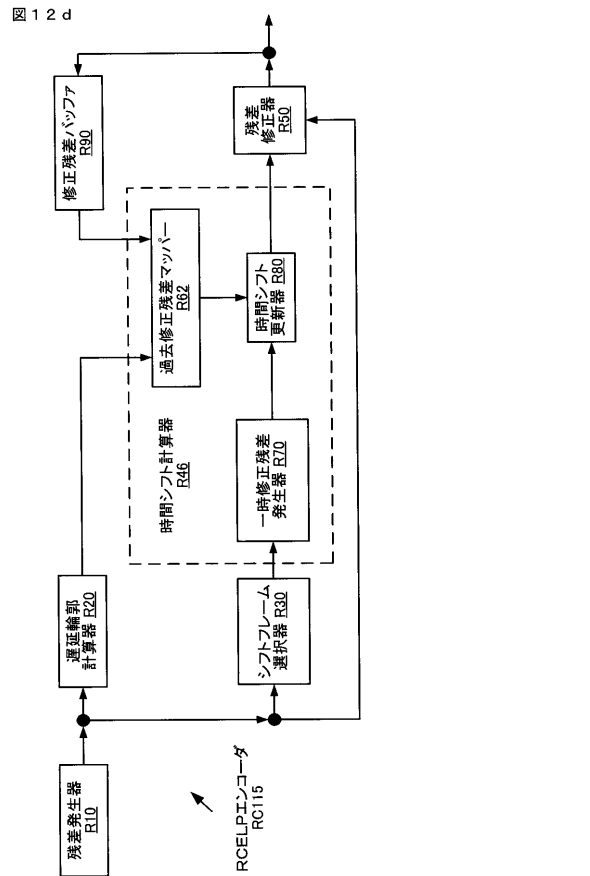
【図12b】



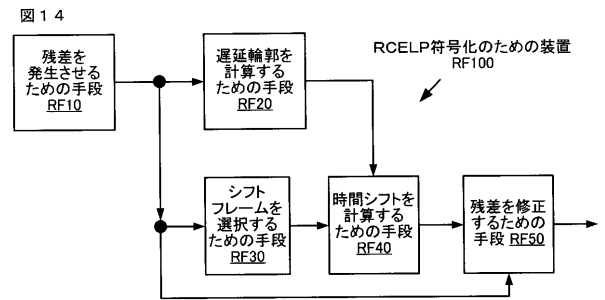
【図13】



【図12d】

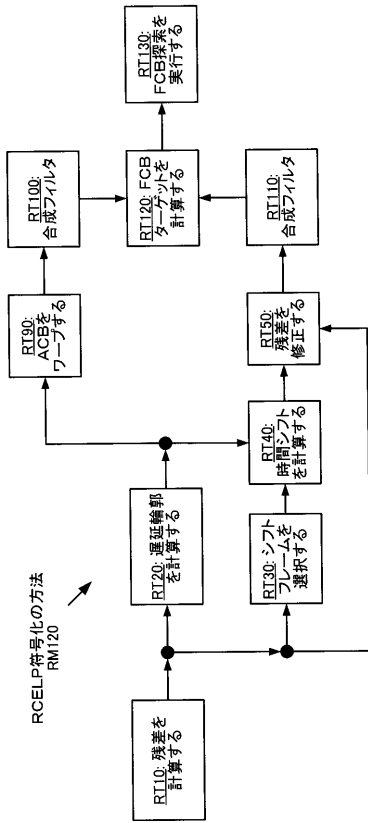


【図14】



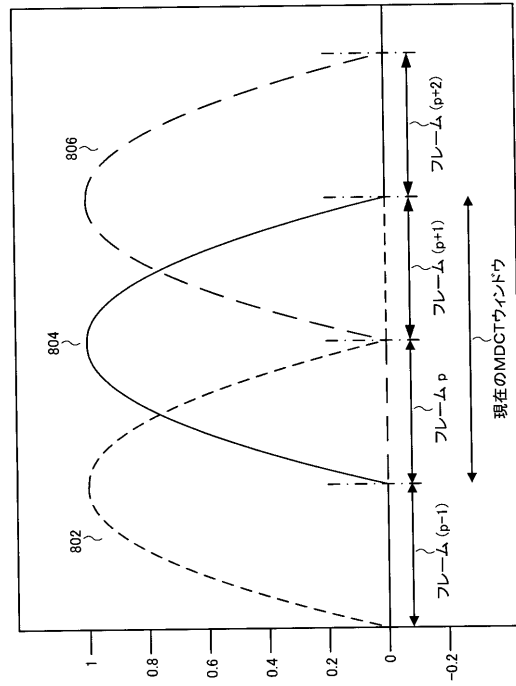
【 図 15 】

図 15



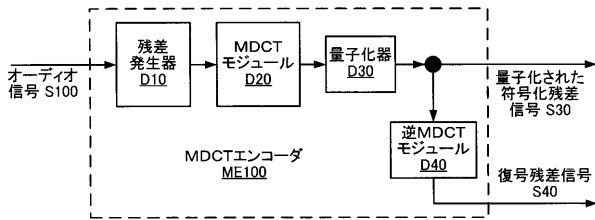
【 図 16 】

図 16



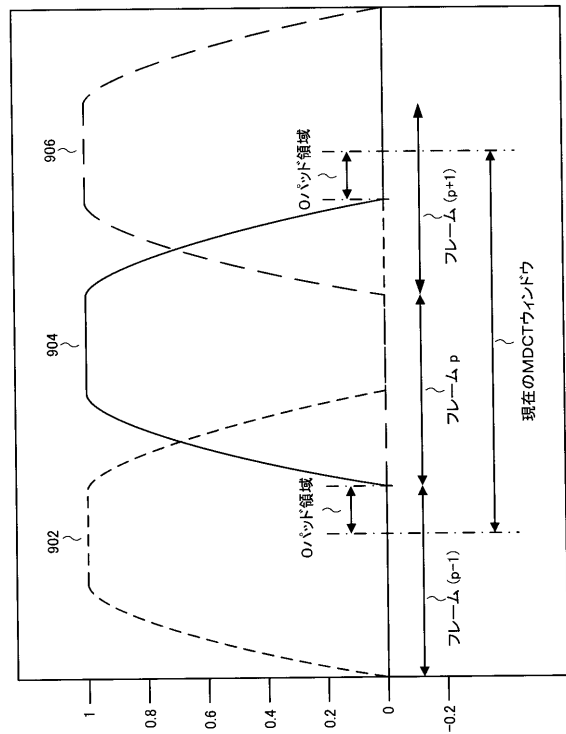
【 図 17 a 】

図 17 a



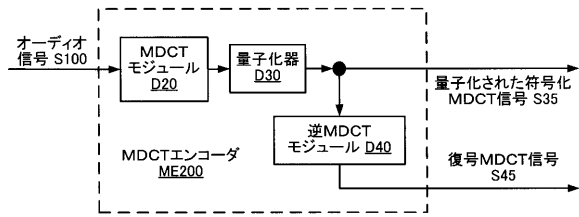
【 図 18 】

図 18



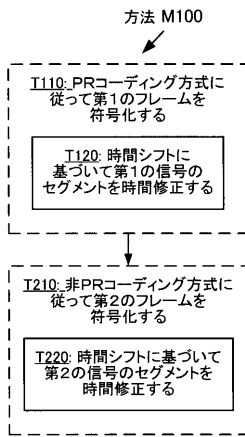
【 図 17 b 】

図 17 b



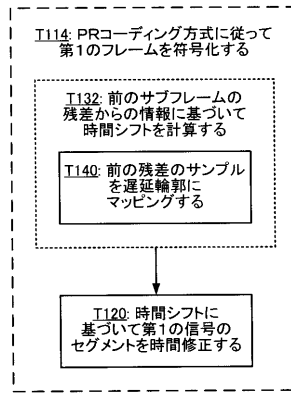
【図19a】

図19a



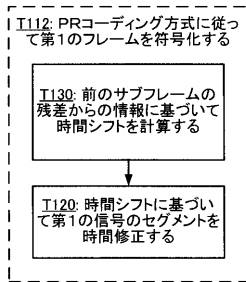
【図19c】

図19c



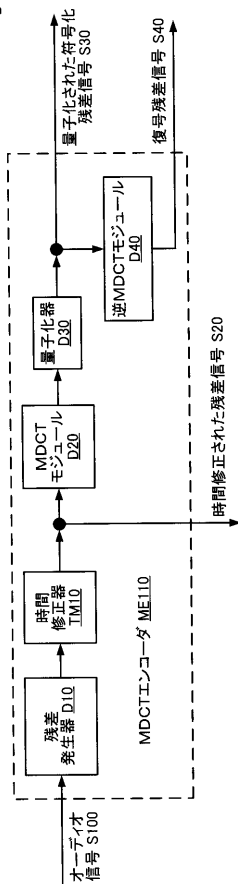
【図19b】

図19b



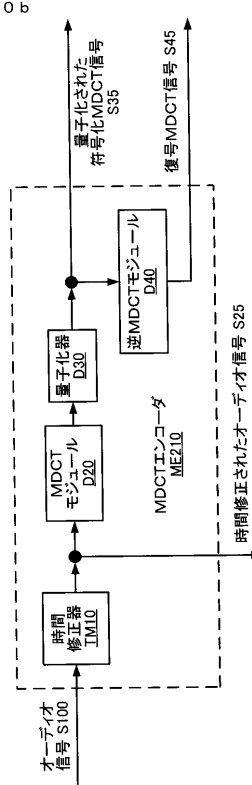
【図20a】

図20a

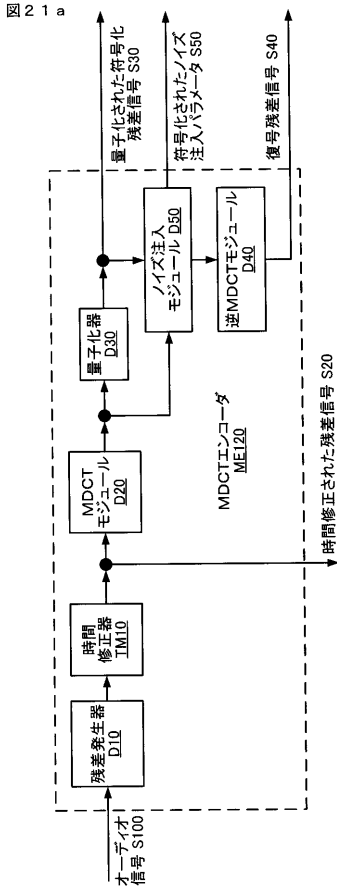


【図20b】

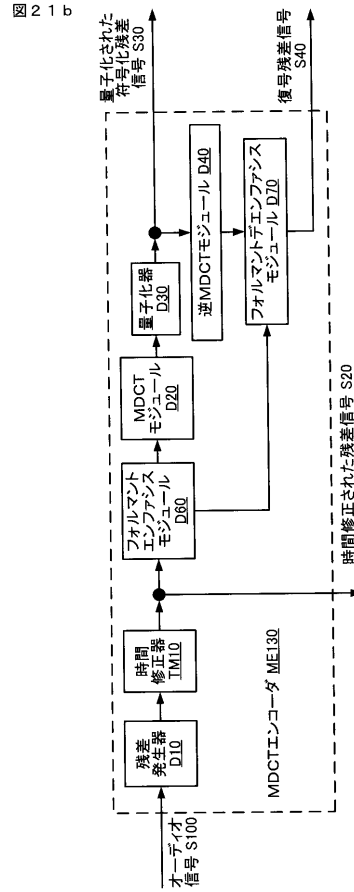
図20b



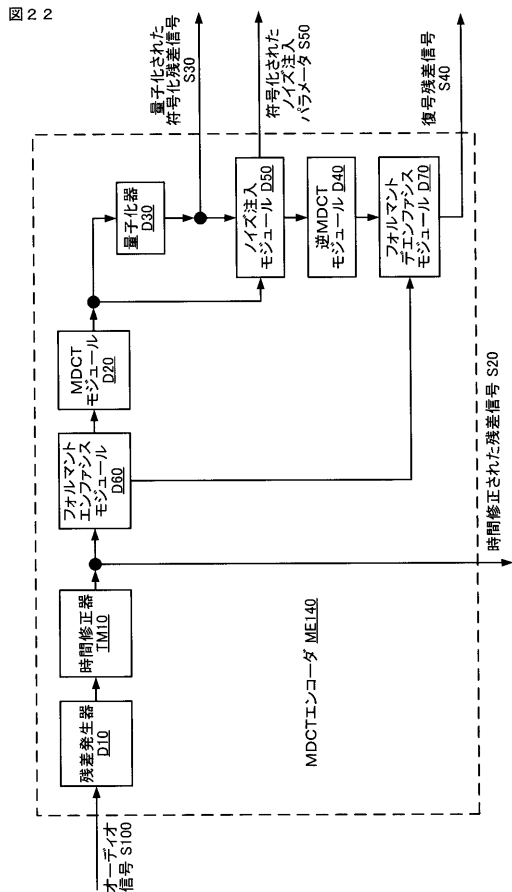
【 図 2 1 a 】



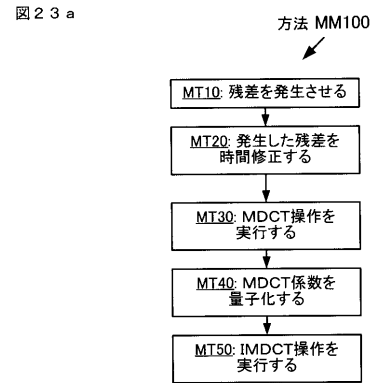
【 図 2 1 b 】



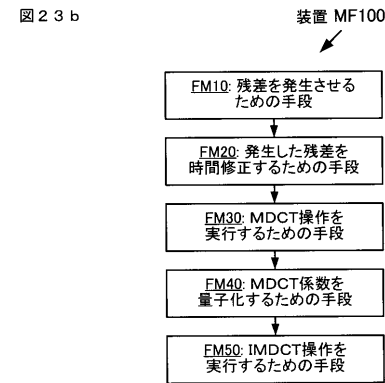
【 図 2 2 】



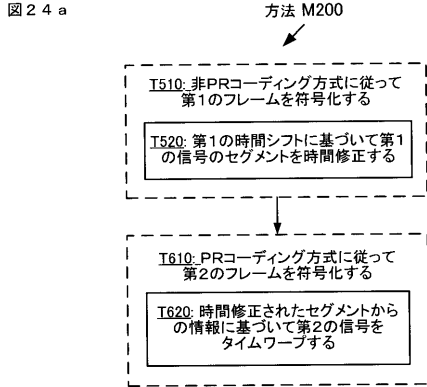
【 図 2 3 a 】



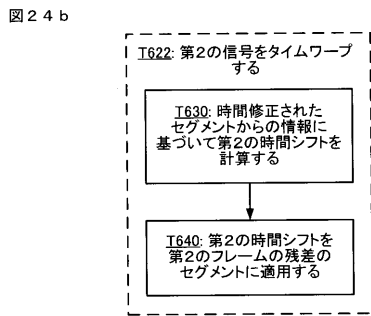
【 図 2 3 b 】



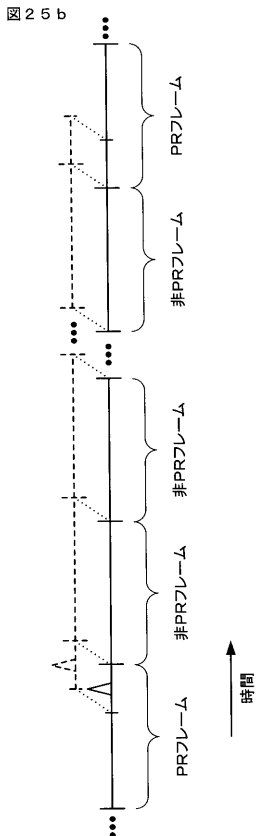
【図24a】



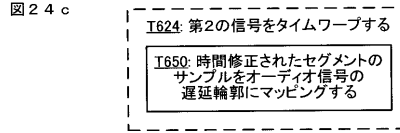
【図24b】



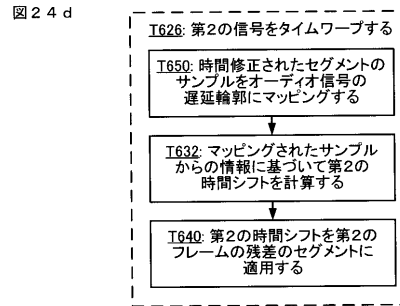
【図25b】



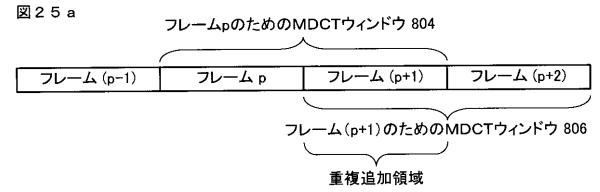
【図24c】



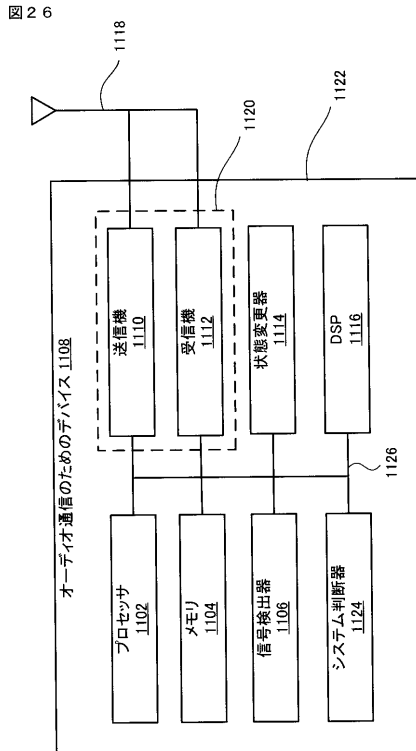
【図24d】



【図25a】



【図26】



【手続補正書】

【提出日】平成25年8月5日(2013.8.5)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

オーディオ信号のフレームを処理する方法であって、

第1のコーディング方式に従って前記オーディオ信号の第1のフレームを符号化することと、

ピッチ調整(P R)コーディング方式に従って前記オーディオ信号の第2のフレームを符号化することと、

を備え、

前記第2のフレームが、前記オーディオ信号中の前記第1のフレームに後続し且つ連続し、

前記第1のコーディング方式が非P Rコーディング方式であり、

前記第1のフレームを符号化することが、前記第1のフレームに基づく第1の信号のセグメントを、第1の時間シフトに基づいて時間修正することを含み、前記時間修正することが、(A)前記第1の時間シフトに従って前記第1の信号の前記セグメントを時間シフトすることと、(B)前記第1の時間シフトに基づいて前記第1の信号の前記セグメントをタイムワープすることと、のうちの1つを含み、

前記第2のフレームを符号化することが、前記第2のフレームに基づく第2の信号のセグメントを、第2の時間シフトに基づいて時間修正することを含み、前記時間修正することが、(A)前記第2の時間シフトに従って前記第2の信号の前記セグメントを時間シフトすることと、(B)前記第2の時間シフトに基づいて前記第2の信号の前記セグメントをタイムワープすることと、のうちの1つを含み、

前記第2の信号のセグメントを時間修正することが、前記第2の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させることを含み、

前記第2の時間シフトが、前記第1の信号の前記時間修正されたセグメントからの情報に基づく、方法。

【請求項2】

前記第1のフレームを符号化することが、前記第1の信号の前記時間修正されたセグメントに基づく第1の符号化フレームを生成することを含み、

前記第2のフレームを符号化することが、前記第2の信号の前記時間修正されたセグメントに基づく第2の符号化フレームを生成することを含む、請求項1に記載の方法。

【請求項3】

前記第1の信号が前記第1のフレームの残差であり、前記第2の信号が前記第2のフレームの残差である、請求項1に記載の方法。

【請求項4】

前記第1及び第2の信号が重み付きオーディオ信号である、請求項1に記載の方法。

【請求項5】

前記第2の信号のセグメントを時間修正することが、前記第1の信号の前記時間修正されたセグメントからの情報に基づいて前記第2の時間シフトを計算することを含み、

前記第2の時間シフトを計算することが、前記第2のフレームからの情報に基づく遅延輪郭に、前記第1の信号の前記時間修正されたセグメントをマッピングすることを含む、請求項1に記載の方法。

【請求項6】

前記第2の時間シフトが、前記マッピングされたセグメントのサンプルと一時修正残差

のサンプルとの間の相関に基づかれ、

前記一時修正残差が、(A)前記第2のフレームの残差のサンプルと、(B)前記第1の時間シフトと、に基づかれる、請求項5に記載の方法。

【請求項7】

前記第2の信号が前記第2のフレームの残差であり、

前記第2の信号のセグメントを時間修正することが、前記第2の時間シフトに従って前記残差の第1のセグメントを時間シフトすることを含み、

前記方法が、

前記第1の信号の前記時間修正されたセグメントからの情報に基づいて、前記第2の時間シフトとは異なる第3の時間シフトを計算することと、

前記第3の時間シフトに従って前記残差の第2のセグメントを時間シフトすることと

を備える、請求項1に記載の方法。

【請求項8】

前記第2の信号が前記第2のフレームの残差であり、

前記第2の信号のセグメントを時間修正することが、前記第2の時間シフトに従って前記残差の第1のセグメントを時間シフトすることを含み、

前記方法が、

前記残差の前記時間修正された第1のセグメントからの情報に基づいて、前記第2の時間シフトとは異なる第3の時間シフトを計算することと、

前記第3の時間シフトに従って前記残差の第2のセグメントを時間シフトすることと

を備える、請求項1に記載の方法。

【請求項9】

前記第2の信号のセグメントを時間修正することが、前記第2のフレームからの情報に基づく遅延輪郭に、前記第1の信号の前記時間修正されたセグメントのサンプルをマッピングすることを含む、請求項1に記載の方法。

【請求項10】

前記方法が、

適応コードブックバッファに前記第1の信号の前記時間修正されたセグメントに基づくシーケンスを記憶することと、

前記記憶することに続いて、前記第2のフレームからの情報に基づく遅延輪郭に前記適応コードブックバッファのサンプルをマッピングすることと、

を備える、請求項1に記載の方法。

【請求項11】

前記第2の信号が前記第2のフレームの残差であり、前記第2の信号のセグメントを時間修正することが前記第2のフレームの前記残差をタイムワープすることを含み、

前記方法が、前記第2のフレームの前記タイムワープされた残差からの情報に基づいて前記オーディオ信号の第3のフレームの残差をタイムワープすることを備え、前記第3のフレームが前記オーディオ信号中の前記第2のフレームに連続する、請求項1に記載の方法。

【請求項12】

前記第2の信号が前記第2のフレームの残差であり、前記第2の信号のセグメントを時間修正することが、(A)前記第1の信号の前記時間修正されたセグメントからの情報と、(B)前記第2のフレームの前記残差からの情報と、に基づいて前記第2の時間シフトを計算することを含む、請求項1に記載の方法。

【請求項13】

前記PRコーディング方式がリラックストコード励振線形予測コーディング方式であり、前記非PRコーディング方式が、(A)ノイズ励起線形予測コーディング方式と、(B)修正離散コサイン変換コーディング方式と、(C)プロトタイプ波形補間コーディング

方式と、のうちの1つである、請求項1に記載の方法。

【請求項14】

前記非PRコーディング方式が修正離散コサイン変換コーディング方式である、請求項1に記載の方法。

【請求項15】

前記第1のフレームを符号化することが、
符号化残差を得るために前記第1のフレームの残差に対して修正離散コサイン変換(MDCT)演算を実行することと、
復号残差を得るために前記符号化残差に基づく信号に対して逆MDCT演算を実行することと、
を含み、
前記第1の信号が前記復号残差に基づく、請求項1に記載の方法。

【請求項16】

前記第1のフレームを符号化することが、
前記第1の信号である、前記第1のフレームの残差を発生させることと、
前記第1の信号のセグメントを時間修正することに続いて、符号化残差を得るために、
前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換演算を実行することと、
前記符号化残差に基づいて第1の符号化フレームを生成することと、
を含む、請求項1に記載の方法。

【請求項17】

前記第1の信号がM個のサンプルの長さを有し、前記第2の信号がM個のサンプルの長さを有し、
前記第1のフレームを符号化することが、前記時間修正されたセグメントを含む、前記第1の信号のM個のサンプルと、前記第2の信号の $3M/4$ 個以下のサンプルと、に基づくM個の修正離散コサイン変換(MDCT)係数の組を生成することを含む、請求項1に記載の方法。

【請求項18】

前記第1の信号がM個のサンプルの長さを有し、前記第2の信号がM個のサンプルの長さを有し、
前記第1のフレームを符号化することが、(A)前記時間修正されたセグメントを含む、前記第1の信号のM個のサンプルを含み、(B)ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで始まり、(C)ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで終わる、 $2M$ 個のサンプルのシーケンスに基づくM個の修正離散コサイン変換(MDCT)係数の組を生成することを含む、請求項1に記載の方法。

【請求項19】

オーディオ信号のフレームを処理するための装置であって、
第1のコーディング方式に従って前記オーディオ信号の第1のフレームを符号化するための手段と、
ピッチ調整(PR)コーディング方式に従って前記オーディオ信号の第2のフレームを符号化するための手段と、
を備え、
前記第2のフレームが、前記オーディオ信号中の前記第1のフレームに後続し且つ連続し、
前記第1のコーディング方式が非PRコーディング方式であり、
第1のフレームを符号化するための前記手段が、前記第1のフレームに基づく第1の信号のセグメントを、第1の時間シフトに基づいて時間修正するための手段を含み、前記時間修正するための手段が、(A)前記第1の時間シフトに従って前記第1の信号の前記セグメントを時間シフトすることと、(B)前記第1の時間シフトに基づいて前記第1の信号の前記セグメントをタイムワープすることと、のうちの1つを実行するように構成され

、
前記第2のフレームを符号化するための手段が、前記第2のフレームに基づく第2の信号のセグメントを、第2の時間シフトに基づいて時間修正するための手段を含み、前記時間修正するための手段が、(A)前記第2の時間シフトに従って前記第2の信号の前記セグメントを時間シフトすることと、(B)前記第2の時間シフトに基づいて前記第2の信号の前記セグメントをタイムワープすることと、のうちの1つを実行するように構成され、

前記第2の信号のセグメントを時間修正するための手段が、前記第2の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させるように構成され、

前記第2の時間シフトが、前記第1の信号の前記時間修正されたセグメントからの情報に基づく、装置。

【請求項20】

前記第1の信号が前記第1のフレームの残差であり、前記第2の信号が前記第2のフレームの残差である、請求項19に記載の装置。

【請求項21】

前記第1及び第2の信号が重み付きオーディオ信号である、請求項19に記載の装置。

【請求項22】

前記第2の信号のセグメントを時間修正するための手段が、前記第1の信号の前記時間修正されたセグメントからの情報に基づいて前記第2の時間シフトを計算するための手段を含み、

前記第2の時間シフトを計算するための手段が、前記第2のフレームからの情報に基づく遅延輪郭に、前記第1の信号の前記時間修正されたセグメントをマッピングするための手段を含む、請求項19に記載の装置。

【請求項23】

前記第2の時間シフトが、前記マッピングされたセグメントのサンプルと一時修正残差のサンプルとの間の相関に基づき、

前記一時修正残差が、(A)前記第2のフレームの残差のサンプルと、(B)前記第1の時間シフトと、に基づく、請求項22に記載の装置。

【請求項24】

前記第2の信号が前記第2のフレームの残差であり、

前記第2の信号のセグメントを時間修正するための手段が、前記第2の時間シフトに従って前記残差の第1のセグメントを時間シフトするように構成され、

前記装置が、

前記残差の前記時間修正された第1のセグメントからの情報に基づいて、前記第2の時間シフトとは異なる第3の時間シフトを計算するための手段と、

前記第3の時間シフトに従って前記残差の第2のセグメントを時間シフトするための手段と、

を備える、請求項19に記載の装置。

【請求項25】

前記第2の信号が前記第2のフレームの残差であり、第2の信号のセグメントを時間修正するための手段が、(A)前記第1の信号の前記時間修正されたセグメントからの情報と、(B)前記第2のフレームの前記残差からの情報と、に基づいて前記第2の時間シフトを計算するための手段を含む、請求項19に記載の装置。

【請求項26】

前記第1のフレームを符号化するための手段が、

前記第1の信号である、前記第1のフレームの残差を発生させるための手段と、

符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換演算を実行するための手段と、

を含み、

前記第1のフレームを符号化するための手段が、前記符号化残差に基づいて第1の符号

化フレームを生成するように構成される、請求項 19 に記載の装置。

【請求項 27】

前記第 1 の信号が M 個のサンプルの長さを有し、前記第 2 の信号が M 個のサンプルの長さを有し、

前記第 1 のフレームを符号化するための手段が、前記時間修正されたセグメントを含む、前記第 1 の信号の M 個のサンプルと、前記第 2 の信号の $3M/4$ 個以下のサンプルと、に基づく M 個の修正離散コサイン変換 (MDCT) 係数の組を生成するための手段を含む、請求項 19 に記載の装置。

【請求項 28】

前記第 1 の信号が M 個のサンプルの長さを有し、前記第 2 の信号が M 個のサンプルの長さを有し、

前記第 1 のフレームを符号化するための手段が、(A) 前記時間修正されたセグメントを含む、前記第 1 の信号の M 個のサンプルを含み、(B) ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで始まり、(C) ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで終わる、 $2M$ 個のサンプルのシーケンスに基づく M 個の修正離散コサイン変換 (MDCT) 係数の組を生成するための手段を含む、請求項 19 に記載の装置。

【請求項 29】

オーディオ信号のフレームを処理するための装置であって、

第 1 のコーディング方式に従って前記オーディオ信号の第 1 のフレームを符号化するように構成された第 1 のフレームエンコーダと、

ピッチ調整 (PR) コーディング方式に従って前記オーディオ信号の第 2 のフレームを符号化するように構成された第 2 のフレームエンコーダと、

を備え、

前記第 2 のフレームが、前記オーディオ信号中の前記第 1 のフレームに後続し且つ連続し、

前記第 1 のコーディング方式が非 PR コーディング方式であり、

前記第 1 のフレームエンコーダが、前記第 1 のフレームに基づく第 1 の信号のセグメントを、第 1 の時間シフトに基づいて時間修正するように構成された第 1 の時間修正器を含み、前記第 1 の時間修正器が、(A) 前記第 1 の時間シフトに従って前記第 1 の信号の前記セグメントを時間シフトすることと、(B) 前記第 1 の時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され、

前記第 2 のフレームエンコーダが、前記第 2 のフレームに基づく第 2 の信号のセグメントを、第 2 の時間シフトに基づいて時間修正するように構成された第 2 の時間修正器を含み、前記第 2 の時間修正器が、(A) 前記第 2 の時間シフトに従って前記第 2 の信号の前記セグメントを時間シフトすることと、(B) 前記第 2 の時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され、

前記第 2 の時間修正器が、前記第 2 の信号の別のピッチパルスに対する第 2 の信号の前記セグメントのピッチパルスの位置を変化させるように構成され、

前記第 2 の時間シフトが、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づく、装置。

【請求項 30】

前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、請求項 29 に記載の装置。

【請求項 31】

前記第 1 及び第 2 の信号が重み付きオーディオ信号である、請求項 29 に記載の装置。

【請求項 32】

前記第 2 の時間修正器が、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づいて前記第 2 の時間シフトを計算するように構成された時間シフト計算器を含み、

前記時間シフト計算器が、前記第2のフレームからの情報に基づく遅延輪郭に、前記第1の信号の前記時間修正されたセグメントをマッピングするように構成されたマッパーを含む、請求項29に記載の装置。

【請求項33】

前記第2の時間シフトが、前記マッピングされたセグメントのサンプルと一時修正残差のサンプルとの間の相関に基づき、

前記一時修正残差が、(A)前記第2のフレームの残差のサンプルと、(B)前記第1の時間シフトとに基づく、請求項32に記載の装置。

【請求項34】

前記第2の信号が前記第2のフレームの残差であり、

前記第2の時間修正器が、前記第2の時間シフトに従って前記残差の第1のセグメントを時間シフトするように構成され、

時間シフト計算器が、前記残差の前記時間修正された第1のセグメントからの情報に基づいて、前記第2の時間シフトとは異なる第3の時間シフトを計算するように構成され、

前記第2の時間修正器が、前記第3の時間シフトに従って前記残差の第2のセグメントを時間シフトするように構成される、請求項29に記載の装置。

【請求項35】

前記第2の信号が前記第2のフレームの残差であり、前記第2の時間修正器が、(A)前記第1の信号の前記時間修正されたセグメントからの情報と、(B)前記第2のフレームの前記残差からの情報と、に基づいて前記第2の時間シフトを計算するように構成された時間シフト計算器を含む、請求項29に記載の装置。

【請求項36】

前記第1のフレームエンコーダが、

前記第1の信号である、前記第1のフレームの残差を発生させるように構成された残差発生器と、

符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換(MDCT)演算を実行するように構成されたMDCTモジュールと、

を含み、

前記第1のフレームエンコーダが、前記符号化残差に基づいて第1の符号化フレームを生成するように構成される、請求項29に記載の装置。

【請求項37】

前記第1の信号がM個のサンプルの長さを有し、前記第2の信号がM個のサンプルの長さを有し、

前記第1のフレームエンコーダが、前記時間修正されたセグメントを含む、前記第1の信号のM個のサンプルと、前記第2の信号の $3M/4$ 個以下のサンプルと、に基づくM個の修正離散コサイン変換(MDCT)係数の組を生成するように構成されたMDCTモジュールを含む、請求項29に記載の装置。

【請求項38】

前記第1の信号がM個のサンプルの長さを有し、前記第2の信号がM個のサンプルの長さを有し、

前記第1のフレームエンコーダが、(A)前記時間修正されたセグメントを含む、前記第1の信号のM個のサンプルを含み、(B)ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで始まり、(C)ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで終わる、 $2M$ 個のサンプルのシーケンスに基づくM個の修正離散コサイン変換(MDCT)係数の組を生成するように構成されたMDCTモジュールを含む、請求項29に記載の装置。

【請求項39】

コンピュータに、第1のコーディング方式に従ってオーディオ信号の第1のフレームを符号化させるためのコードと、

コンピュータに、ピッチ調整(PR)コーディング方式に従って前記オーディオ信号の

第2のフレームを符号化させるためのコードと、
 を備えるプログラムを記録したコンピュータ可読記録媒体であって、
 前記第2のフレームが、前記オーディオ信号中の前記第1のフレームに後続し且つ連続し、

前記第1のコーディング方式が非PRコーディング方式であり、

前記コンピュータに第1のフレームを符号化させるためのコードは、コンピュータに、前記第1のフレームに基づく第1の信号のセグメントを、第1の時間シフトに基づいて時間修正させるためのコードを含み、前記コンピュータに時間修正させるためのコードが、(A)コンピュータに、前記第1の時間シフトに従って前記第1の信号の前記セグメントを時間シフトさせるためのコードと、(B)コンピュータに、前記第1の時間シフトに基づいて前記第1の信号の前記セグメントをタイムワープさせるためのコードと、のうちの1つを含み、

前記コンピュータに第2のフレームを符号化させるためのコードは、コンピュータに、前記第2のフレームに基づく第2の信号のセグメントを、第2の時間シフトに基づいて時間修正させるためのコードを含み、前記コンピュータに時間修正させるためのコードが、(A)コンピュータに、前記第2の時間シフトに従って前記第2の信号の前記セグメントを時間シフトさせるためのコードと、(B)コンピュータに、前記第2の時間シフトに基づいて前記第2の信号の前記セグメントをタイムワープさせるためのコードと、のうちの1つを含み、

前記コンピュータに第2の信号のセグメントを時間修正させるためのコードが、コンピュータに、前記第2の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させるためのコードを含み、

前記第2の時間シフトが、前記第1の信号の前記時間修正されたセグメントからの情報に基づく、コンピュータ可読記録媒体。

【**手続補正2**】

【**補正対象書類名**】明細書

【**補正対象項目名**】0005

【**補正方法**】変更

【**補正の内容**】

【**0005**】

コード励振線形予測(「CELP」)は、元のオーディオ信号の波形を適合させようと試みるコーディング方式である。リラクストド(relaxed)CELP(「RCELP」)と呼ばれるCELPの変形態を使用して、スピーチ信号のフレーム、特に有声音フレームを符号化することが望ましい場合がある。RCELPコーディング方式では、波形適合制約は緩和される。RCELPコーディング方式はピッチ調整(pitch-regularizing)(「PR」)コーディング方式であり、信号のピッチ周期間の変動(「遅延輪郭(delay contour)」とも呼ばれる)が、一般に、ピッチパルスの相対位置をより滑らかな合成遅延輪郭に一致または近似するように変化させることによって調整される。ピッチ調整により、一般に知覚品質の低下をほとんどまたはまったく伴わずにピッチ情報をより少ないビットで符号化することを可能にする。一般に、調整量を指定する情報はデコーダに送信されない。以下の文書には、RCELPコーディング方式を含むコーディングシステムが記載されている；第3世代パートナーシッププロジェクト2(「3GPP2」)文書C.S0030-0、v3.0、表題「Selectable Mode Vocoder(SMV) Service Option for Wideband Spread Spectrum Communication Systems」、2004年1月(www.3gpp.orgからオンラインで入手可能)；及び3GPP2文書C.S0014-C、v1.0、表題「Enhanced Variable Rate Codec, Speech Service Options 3, 68, and 70 for Wideband Spread Spectrum Digital Systems」、2007年1月(www.3gpp.orgからオンラインで入手可能)。プロトタイプピ

ッチ周期（「P P P」）などのプロトタイプ波形補間（「P W I」）方式を含む、有声音フレーム用の他のコーディング方式は、（たとえば、上記で参照した3 G P P 2 文書 C . S 0 0 1 4 - C の第 4 . 2 . 4 . 3 部に記載されているように）P R として実装されることもできる。男性話者のピッチ周波数の通常の範囲は 5 0 または 7 0 ~ 1 5 0 または 2 0 0 H z を含み、女性話者のピッチ周波数の通常の範囲は 1 2 0 または 1 4 0 ~ 3 0 0 または 4 0 0 H z を含む。

公衆交換電話網（「P S T N」）を介したオーディオ通信は、従来、帯域幅が 3 0 0 ~ 3 4 0 0 キロヘルツ（k H z）の周波数範囲に制限されてきた。セルラー電話通信及び/または V o I P を使用するネットワークなど、オーディオ通信のより最近のネットワークは、同じ帯域幅制限をもたない場合があり、そのようなネットワークを使用する装置では、広帯域周波数範囲を含むオーディオ通信を送信及び受信する能力を有することが望ましい場合がある。たとえば、そのような装置では、下は 5 0 H z まで及び/または上は 7 もしくは 8 k H z までに及び可聴周波数範囲をサポートすることが望ましい場合がある。また、そのような装置では、従来の P S T N 制限外の範囲のオーディオスピーチコンテンツを有することがある、高品質オーディオまたはオーディオ/テレビ会議、音楽及び/またはテレビジョンなどのマルチメディアサービスの配信など、他の適用例をサポートすることが望ましい場合がある。

【手続補正 3】

【補正対象書類名】明細書

【補正対象項目名】0 0 1 9

【補正方法】変更

【補正の内容】

【0 0 1 9】

図 1 に示すセルラー電話通信システムの要素は、パケット交換データ通信をサポートするように構成されることもできる。図 2 に示すように、パケットデータトラフィックは、概して、パケットデータネットワークに接続されたゲートウェイルータに結合されたパケットデータサービスノード（P D S N）1 7 を使用して、移動体加入者ユニット 1 0 と外部のパケットデータネットワーク 1 9（たとえば、インターネットなどの公衆網）との間をルーティングされる。P D S N 1 7 は、1 つまたは複数の B S C 1 4 にサービスを提供し、パケットデータネットワークと無線アクセスネットワークとの間のリンクとして働く、1 つまたは複数のパケット制御機能（P C F）1 5 にデータを順次ルーティングする。パケットデータネットワーク 1 9 はまた、ローカルエリアネットワーク（「L A N」）、キャンパスエリアネットワーク（「C A N」）、メトロポリタンエリアネットワーク（「M A N」）、広域ネットワーク（「W A N」）、リング型ネットワーク、スター型ネットワーク、トークンリングネットワークなどを含むように実装されることができ。ネットワーク 1 9 に接続されたユーザ端末は、P D A、ラップトップコンピュータ、パーソナルコンピュータ、ゲームデバイス（そのようなデバイスの例には、X B o x 及び X B o x 3 6 0（マイクロソフト社（ワシントン州レッドモンド））、プレイステーション 3 及びプレイステーション・ポータブル（ソニー（株）、日本国東京）、ならびに W i i 及び D S（任天堂、日本国京都）がある）、及び/またはオーディオ処理機能を有する任意のデバイスとすることができ、V o I P など、1 つまたは複数のプロトコルを使用して電話通話または他の通信をサポートするように構成されることができ。そのような端末は、内蔵型のスピーカ及びマイクロホン、スピーカとマイクロホンとを含むコード付きハンドセット（たとえば、U S B ハンドセット）、またはスピーカとマイクロホンとを含むワイヤレスヘッドセット（たとえば、B l u e t o o t h S p e c i a l I n t e r e s t G r o u p（ワシントン州ベルビュー）によって公表された B l u e t o o t h プロトコルのバージョンを使用して端末にオーディオ情報を通信するヘッドセット）を含むことができる。そのようなシステムは、P S T N に決して入ることなく、（たとえば、V o I P など、1 つまたは複数のプロトコルによる）異なる無線アクセスネットワーク上の移動体加入者ユニット間、移動体加入者ユニットと非移動体ユーザ端末との間、または 2 つの非移

動体ユーザ端末間のパケットデータトラフィックとして電話通話または他の通信を行なうように構成されることができる。移動体加入者ユニット10または他のユーザ端末は、「アクセス端末」とも呼ばれる。

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0043

【補正方法】変更

【補正の内容】

【0043】

図5bは、フレームエンコーダ30c、30dの実装形態32c、32dを含むマルチモードオーディオエンコーダAE20の実装形態AE24のブロック図を示している。この例では、コーディング方式選択器20の実装形態24は、オーディオ信号S100のスピーチフレームを非スピーチフレーム（たとえば、音楽）と区別するように構成される。たとえば、コーディング方式選択器24は、スピーチフレームに対しては（CELPエンコーダなどのスピーチフレームエンコーダ32cの選択を指示する）ハイであり、非スピーチフレームに対しては（MDCTエンコーダなどの非スピーチフレームエンコーダ32dの選択を指示する）ローである、バイナリ値コーディング方式選択信号を出力するように構成されることが、またはその逆に構成されることも可能である。そのような分類は、フレームエネルギー、ピッチ、周期性、スペクトル分布（たとえば、スペクトル傾斜、LPC係数、線スペクトル周波数（「LSF」））、及び/またはゼロ交差レートなど、フレームのエネルギー及び/またはスペクトル成分の1つまたは複数の特性に基づかれることができる。コーディング方式選択器24は、そのような特性の値を計算するように、オーディオエンコーダAE24の1つまたは複数の他のモジュールからそのような特性の値を受け取るように、及び/またはオーディオエンコーダAE24を含むデバイス（たとえば、セルラー電話）の1つまたは複数の他のモジュールからそのような特性の値を受け取るように実装されることができ。そのような分類は、そのような特性の値または大きさを閾値と比較すること、及び/またはそのような特性の（たとえば、先行フレームに対する）変化の大きさを閾値と比較することを含むことができる。そのような分類は、隠れマルコフモデルなどの多状態モデルを更新するために使用され得る、1つまたは複数の前のフレーム及び/または1つまたは複数の後続のフレームからの情報に基づかれることができる。

【手続補正5】

【補正対象書類名】明細書

【補正対象項目名】0098

【補正方法】変更

【補正の内容】

【0098】

図18は、Mよりも短いルックアヘッド間隔を可能にするために（たとえば、図16に示す関数 $w(n)$ の代わりに）MDCTモジュールD20によって適用されるウィンドウ関数 $w(n)$ の一例を示している。図18に示す特定の例では、ルックアヘッド間隔はM/2個のサンプルの長さであるが、そのような技法は、L個のサンプルの任意のルックアヘッドを可能にするように実装されることができ。ここで、Lは0からMまでの任意の値を有する。この技法（上記の参照により組み込まれる3GPP2 EVRC文書C.S0014-Cの第4.13節の第4.13.4部（4-147ページ）、及び「SYSTEMS AND METHODS FOR MODIFYING A WINDOW WITH A FRAME ASSOCIATED WITH AN AUDIO SIGNAL」と題する米国特許公開第2008/0027719号に記載されている例）では、MDCTウィンドウは、長さ $(M-L)/2$ の0パッド領域で開始及び終了し、 $w(n)$ はプリンセン-ブラッドレイ条件を満たす。そのようなウィンドウ関数の1つの実装形態は、次のように表されることができる。

【数 8】

$$w(n) = \begin{cases} 0, & 0 \leq n < \frac{M-L}{2} \\ \sin\left[\frac{\pi}{2L}\left(n - \frac{M-L}{2}\right)\right], & \frac{M-L}{2} \leq n < \frac{M+L}{2} \\ 1, & \frac{M+L}{2} \leq n < \frac{3M-L}{2} \\ \sin\left[\frac{\pi}{2L}\left(3L+n - \frac{3M-L}{2}\right)\right], & \frac{3M-L}{2} \leq n < \frac{3M+L}{2} \\ 0, & \frac{3M+L}{2} \leq n < 2M \end{cases} \quad (\text{式5})$$

【手続補正 6】

【補正対象書類名】明細書

【補正対象項目名】0144

【補正方法】変更

【補正の内容】

【0144】

この例では、デバイス1108は、いくつかの異なる状態のうちいずれか1つで動作するように構成されており、デバイスの現在の状態と、トランシーバ1120によって受信され、信号検出器1106によって検出された信号とに基づいてデバイス1108の状態を制御するように構成された状態変更器1114を含む。この例では、デバイス1108は、現在のサービスプロバイダが不十分であると判断し、異なるサービスプロバイダに転送するようにデバイス1108を制御するように構成されたシステム判断器1124をも含む。

以下に、本願出願の当初の特許請求の範囲に記載された発明を付記する。

(1) オーディオ信号のフレームを処理する方法であって、

ピッチ調整(PR)コーディング方式に従って前記オーディオ信号の第1のフレームを符号化することと、

非PRコーディング方式に従って前記オーディオ信号の第2のフレームを符号化することと、

を備え、

前記第2のフレームが、前記オーディオ信号中の前記第1のフレームに後続し且つ連続し、

前記第1のフレームを符号化することが、前記第1のフレームに基づく第1の信号のセグメントを、時間シフトに基づいて時間修正することを含み、前記時間修正することが、

(A) 前記時間シフトに従って前記第1のフレームの前記セグメントを時間シフトすることと、(B) 前記時間シフトに基づいて前記第1の信号の前記セグメントをタイムワープすることと、のうちの1つを含み、

前記第1の信号のセグメントを時間修正することが、前記第1の信号の別のピッチパルスに対して前記セグメントのピッチパルスの位置を変化させることを含み、

前記第2のフレームを符号化することが、前記第2のフレームに基づく第2の信号のセグメントを、前記時間シフトに基づいて時間修正することを含み、前記時間修正することが、(A) 前記時間シフトに従って前記第2のフレームの前記セグメントを時間シフトすることと、(B) 前記時間シフトに基づいて前記第2の信号の前記セグメントをタイムワープすることと、のうちの1つを含む、方法。

(2) 前記第 1 のフレームを符号化することが、前記第 1 の信号の前記時間修正されたセグメントに基づく第 1 の符号化フレームを生成することを含み、

前記第 2 のフレームを符号化することが、前記第 2 の信号の前記時間修正されたセグメントに基づく第 2 の符号化フレームを生成することを含む、(1) に記載の方法。

(3) 前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、(1) に記載の方法。

(4) 前記第 1 及び第 2 の信号が重み付きオーディオ信号である、(1) に記載の方法。

(5) 前記第 1 のフレームを符号化することが、前記オーディオ信号中の前記第 1 のフレームに先行する第 3 のフレームの残差からの情報に基づいて前記時間シフトを計算することを含む、(1) に記載の方法。

(6) 前記時間シフトを計算することが、前記オーディオ信号の遅延輪郭に前記第 3 のフレームの前記残差のサンプルをマッピングすることを含む、(5) に記載の方法。

(7) 前記第 1 のフレームを符号化することが、前記オーディオ信号のピッチ周期に関する情報に基づいて前記遅延輪郭をコンピュータ計算することを含む、(6) に記載の方法。

(8) 前記 P R コーディング方式がリラクストコード励振線形予測コーディング方式であり、

前記非 P R コーディング方式が、(A) ノイズ励起線形予測コーディング方式と、(B) 修正離散コサイン変換コーディング方式と、(C) プロタイプ波形補間コーディング方式と、のうちの 1 つである、(1) に記載の方法。

(9) 前記非 P R コーディング方式が修正離散コサイン変換コーディング方式である、(1) に記載の方法。

(1 0) 前記第 2 のフレームを符号化することが、
符号化残差を得るために前記第 2 のフレームの残差に対して修正離散コサイン変換 (M D C T) 演算を実行することと、

復号残差を得るために前記符号化残差に基づく信号に対して逆 M D C T 演算を実行することと、

を含み、

前記第 2 の信号が前記復号残差に基づく、(1) に記載の方法。

(1 1) 前記第 2 のフレームを符号化することが、

前記第 2 の信号である、前記第 2 のフレームの残差を発生させることと、

前記第 2 の信号のセグメントを時間修正することに続いて、符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換演算を実行することと、

前記符号化残差に基づいて第 2 の符号化フレームを生成することと、

を含む、(1) に記載の方法。

(1 2) 前記方法が、前記オーディオ信号中の前記第 2 のフレームに後続するフレームの残差のセグメントを、前記時間シフトに従って時間シフトすることを備える、(1) に記載の方法。

(1 3) 前記方法が、前記第 2 のフレームに後続する前記オーディオ信号の第 3 のフレームに基づく第 3 の信号のセグメントを、前記時間シフトに基づいて時間修正することを含み、

前記第 2 のフレームを符号化することが、前記第 2 及び第 3 の信号の前記時間修正されたセグメントのサンプルを含むウィンドウに対して修正離散コサイン変換 (M D C T) 演算を実行することを含む、(1) に記載の方法。

(1 4) 前記第 2 の信号が M 個のサンプルの長さを有し、前記第 3 の信号が M 個のサンプルの長さを有し、

前記 M D C T 演算を実行することが、(A) 前記時間修正されたセグメントを含む、前記第 2 の信号の M 個のサンプルと、(B) 前記第 3 の信号の $3 M / 4$ 個以下のサンプルと

、に基づくM個のMDC T係数の組を生成することを含む、(13)に記載の方法。

(15) 前記第2の信号がM個のサンプルの長さを有し、前記第3の信号がM個のサンプルの長さを有し、

前記MDC T演算を実行することが、(A)前記時間修正されたセグメントを含む、前記第2の信号のM個のサンプルを含み、(B)ゼロ値の少なくともM/8個のサンプルのシーケンスで始まり、(C)ゼロ値の少なくともM/8個のサンプルのシーケンスで終わる、2M個のサンプルのシーケンスに基づくM個のMDC T係数の組を生成することを含む、(13)に記載の方法。

(16) オーディオ信号のフレームを処理するための装置であって、

ピッチ調整(PR)コーディング方式に従って前記オーディオ信号の第1のフレームを符号化するための手段と、

非PRコーディング方式に従って前記オーディオ信号の第2のフレームを符号化するための手段と、

を備え、

前記第2のフレームが、前記オーディオ信号中の前記第1のフレームに後続し且つ連続し、

前記第1のフレームを符号化するための手段が、前記第1のフレームに基づく第1の信号のセグメントを、時間シフトに基づいて時間修正するための手段を含み、前記時間修正するための手段が、(A)前記時間シフトに従って前記第1のフレームの前記セグメントを時間シフトすることと、(B)前記時間シフトに基づいて前記第1の信号の前記セグメントをタイムワープすることと、のうちの1つを実行するように構成され、

前記第1の信号のセグメントを時間修正するための手段が、前記第1の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させるように構成され、

前記第2のフレームを符号化するための手段が、前記第2のフレームに基づく第2の信号のセグメントを、前記時間シフトに基づいて時間修正するための手段を含み、前記時間修正するための手段が、(A)前記時間シフトに従って前記第2のフレームの前記セグメントを時間シフトすることと、(B)前記時間シフトに基づいて前記第2の信号の前記セグメントをタイムワープすることと、のうちの1つを実行するように構成される、装置。

(17) 前記第1の信号が前記第1のフレームの残差であり、前記第2の信号が前記第2のフレームの残差である、(16)に記載の装置。

(18) 前記第1及び第2の信号が重み付きオーディオ信号である、(16)に記載の装置。

(19) 前記第1のフレームを符号化するための手段が、前記オーディオ信号中の前記第1のフレームに先行する第3のフレームの残差からの情報に基づいて前記時間シフトを計算するための手段を含む、(16)に記載の装置。

(20) 前記第2のフレームを符号化するための手段が、

前記第2の信号である、前記第2のフレームの残差を発生させるための手段と、

符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換演算を実行するための手段と、

を含み、

前記第2のフレームを符号化するための手段が、前記符号化残差に基づいて第2の符号化フレームを生成するように構成される、(16)に記載の装置。

(21) 前記第2の信号のセグメントを時間修正するための手段が、前記オーディオ信号中の前記第2のフレームに後続するフレームの残差のセグメントを、前記時間シフトに従って時間シフトするように構成される、(16)に記載の装置。

(22) 前記第2の信号のセグメントを時間修正するための手段が、前記第2のフレームに後続する前記オーディオ信号の第3のフレームに基づく第3の信号のセグメントを、前記時間シフトに基づいて時間修正するように構成され、

前記第2のフレームを符号化するための手段が、前記第2及び第3の信号の前記時間修正されたセグメントのサンプルを含むウィンドウに対して修正離散コサイン変換(MDC

T) 演算を実行するための手段を含む、(16)に記載の装置。

(23) 前記第2の信号がM個のサンプルの長さを有し、前記第3の信号がM個のサンプルの長さを有し、

前記MDC T演算を実行するための手段が、(A)前記時間修正されたセグメントを含む、前記第2の信号のM個のサンプルと、(B)前記第3の信号の $3M/4$ 個以下のサンプルと、に基づくM個のMDC T係数の組を生成するように構成される、(22)に記載の装置。

(24) オーディオ信号のフレームを処理するための装置であって、

ピッチ調整(P R)コーディング方式に従って前記オーディオ信号の第1のフレームを符号化するように構成された第1のフレームエンコーダと、

非P Rコーディング方式に従って前記オーディオ信号の第2のフレームを符号化するように構成された第2のフレームエンコーダと、

を備え、

前記第2のフレームが、前記オーディオ信号中の前記第1のフレームに後続し且つ連続し、

前記第1のフレームエンコーダが、前記第1のフレームに基づく第1の信号のセグメントを、時間シフトに基づいて時間修正するように構成された第1の時間修正器を含み、前記第1の時間修正器が、(A)前記時間シフトに従って前記第1のフレームの前記セグメントを時間シフトすることと、(B)前記時間シフトに基づいて前記第1の信号の前記セグメントをタイムワープすることと、のうちの1つを実行するように構成され、

前記第1の時間修正器が、前記第1の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させるように構成され、

前記第2のフレームエンコーダが、前記第2のフレームに基づく第2の信号のセグメントを、前記時間シフトに基づいて時間修正するように構成された第2の時間修正器を含み、前記第2の時間修正器が、(A)前記時間シフトに従って前記第2のフレームの前記セグメントを時間シフトすることと、(B)前記時間シフトに基づいて前記第2の信号の前記セグメントをタイムワープすることと、のうちの1つを実行するように構成される、装置。

(25) 前記第1の信号が前記第1のフレームの残差であり、前記第2の信号が前記第2のフレームの残差である、(24)に記載の装置。

(26) 前記第1及び第2の信号が重み付きオーディオ信号である、(24)に記載の装置。

(27) 前記第1のフレームエンコーダが、前記オーディオ信号中の前記第1のフレームに先行する第3のフレームの残差からの情報に基づいて前記時間シフトを計算するように構成された時間シフト計算器を含む、(24)に記載の装置。

(28) 前記第2のフレームエンコーダが、

前記第2の信号である、前記第2のフレームの残差を発生させるように構成された残差発生器と、

符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換(MDC T)演算を実行するように構成されたMDC Tモジュールと、

を含み、

前記第2のフレームエンコーダが、前記符号化残差に基づいて第2の符号化フレームを生成するように構成される、(24)に記載の装置。

(29) 前記第2の時間修正器が、前記オーディオ信号中の前記第2のフレームに後続するフレームの残差のセグメントを、前記時間シフトに従って時間シフトするように構成される、(24)に記載の装置。

(30) 前記第2の時間修正器が、前記第2のフレームに後続する前記オーディオ信号の第3のフレームに基づく第3の信号のセグメントを、前記時間シフトに基づいて時間修正するように構成され、

前記第2のフレームエンコーダが、前記第2及び第3の信号の前記時間修正されたセグメントのサンプルを含むウィンドウに対して修正離散コサイン変換(MDCT)演算を実行するように構成されたMDCTモジュールを含む、(24)に記載の装置。

(31) 前記第2の信号がM個のサンプルの長さを有し、前記第3の信号がM個のサンプルの長さを有し、

前記MDCTモジュールが、(A)前記時間修正されたセグメントを含む、前記第2の信号のM個のサンプルと、(B)前記第3の信号の $3M/4$ 個以下のサンプルと、に基づくM個のMDCT係数の組を生成するように構成される、(30)に記載の装置。

(32) プロセッサによって実行されると前記プロセッサに、

ピッチ調整(PR)コーディング方式に従って前記オーディオ信号の第1のフレームを符号化することと、

非PRコーディング方式に従って前記オーディオ信号の第2のフレームを符号化することと、

を行わせる命令を備えるコンピュータ可読媒体であって、

前記第2のフレームが、前記オーディオ信号中の前記第1のフレームに後続し且つ連続し、

前記実行されると前記プロセッサに第1のフレームを符号化することを行わせる命令が、前記第1のフレームに基づく第1の信号のセグメントを、時間シフトに基づいて時間修正する命令を含み、前記時間修正する命令が、(A)前記時間シフトに従って前記第1のフレームの前記セグメントを時間シフトする命令と、(B)前記時間シフトに基づいて前記第1の信号の前記セグメントをタイムワープする命令と、のうちの1つを含み、

前記第1の信号のセグメントを時間修正する命令が、前記第1の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させる命令を含み、

前記実行されると前記プロセッサに第2のフレームを符号化することを行わせる命令が、前記第2のフレームに基づく第2の信号のセグメントを、前記時間シフトに基づいて時間修正する命令を含み、前記時間修正する命令が、(A)前記時間シフトに従って前記第2のフレームの前記セグメントを時間シフトする命令と、(B)前記時間シフトに基づいて前記第2の信号の前記セグメントをタイムワープする命令と、のうちの1つを含む、コンピュータ可読媒体。

(33) オーディオ信号のフレームを処理する方法であって、

第1のコーディング方式に従って前記オーディオ信号の第1のフレームを符号化することと、

ピッチ調整(PR)コーディング方式に従って前記オーディオ信号の第2のフレームを符号化することと、

を備え、

前記第2のフレームが、前記オーディオ信号中の前記第1のフレームに後続し且つ連続し、

前記第1のコーディング方式が非PRコーディング方式であり、

前記第1のフレームを符号化することが、前記第1のフレームに基づく第1の信号のセグメントを、第1の時間シフトに基づいて時間修正することを含み、前記時間修正することが、(A)前記第1の時間シフトに従って前記第1の信号の前記セグメントを時間シフトすることと、(B)前記第1の時間シフトに基づいて前記第1の信号の前記セグメントをタイムワープすることと、のうちの1つを含み、

前記第2のフレームを符号化することが、前記第2のフレームに基づく第2の信号のセグメントを、第2の時間シフトに基づいて時間修正することを含み、前記時間修正することが、(A)前記第2の時間シフトに従って前記第2の信号の前記セグメントを時間シフトすることと、(B)前記第2の時間シフトに基づいて前記第2の信号の前記セグメントをタイムワープすることと、のうちの1つを含み、

前記第2の信号のセグメントを時間修正することが、前記第2の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させることを含み、

前記第 2 の時間シフトが、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づく、方法。

(3 4) 前記第 1 のフレームを符号化することが、前記第 1 の信号の前記時間修正されたセグメントに基づく第 1 の符号化フレームを生成することを含み、

前記第 2 のフレームを符号化することが、前記第 2 の信号の前記時間修正されたセグメントに基づく第 2 の符号化フレームを生成することを含む、(3 3) に記載の方法。

(3 5) 前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、(3 3) に記載の方法。

(3 6) 前記第 1 及び第 2 の信号が重み付きオーディオ信号である、(3 3) に記載の方法。

(3 7) 前記第 2 の信号のセグメントを時間修正することが、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づいて前記第 2 の時間シフトを計算することを含み、

前記第 2 の時間シフトを計算することが、前記第 2 のフレームからの情報に基づく遅延輪郭に、前記第 1 の信号の前記時間修正されたセグメントをマッピングすることを含む、(3 3) に記載の方法。

(3 8) 前記第 2 の時間シフトが、前記マッピングされたセグメントのサンプルと一時修正残差のサンプルとの間の相関に基づかれ、

前記一時修正残差が、(A) 前記第 2 のフレームの残差のサンプルと、(B) 前記第 1 の時間シフトと、に基づかれる、(3 7) に記載の方法。

(3 9) 前記第 2 の信号が前記第 2 のフレームの残差であり、前記第 2 の信号のセグメントを時間修正することが、前記第 2 の時間シフトに従って前記残差の第 1 のセグメントを時間シフトすることを含み、

前記方法が、

前記第 1 の信号の前記時間修正されたセグメントからの情報に基づいて、前記第 2 の時間シフトとは異なる第 3 の時間シフトを計算することと、

前記第 3 の時間シフトに従って前記残差の第 2 のセグメントを時間シフトすることと

を備える、(3 3) に記載の方法。

(4 0) 前記第 2 の信号が前記第 2 のフレームの残差であり、前記第 2 の信号のセグメントを時間修正することが、前記第 2 の時間シフトに従って前記残差の第 1 のセグメントを時間シフトすることを含み、

前記方法が、

前記残差の前記時間修正された第 1 のセグメントからの情報に基づいて、前記第 2 の時間シフトとは異なる第 3 の時間シフトを計算することと、

前記第 3 の時間シフトに従って前記残差の第 2 のセグメントを時間シフトすることと

を備える、(3 3) に記載の方法。

(4 1) 前記第 2 の信号のセグメントを時間修正することが、前記第 2 のフレームからの情報に基づく遅延輪郭に、前記第 1 の信号の前記時間修正されたセグメントのサンプルをマッピングすることを含む、(3 3) に記載の方法。

(4 2) 前記方法が、

適応コードブックバッファに前記第 1 の信号の前記時間修正されたセグメントに基づくシーケンスを記憶することと、

前記記憶することに続いて、前記第 2 のフレームからの情報に基づく遅延輪郭に前記適応コードブックバッファのサンプルをマッピングすることと、

を備える、(3 3) に記載の方法。

(4 3) 前記第 2 の信号が前記第 2 のフレームの残差であり、前記第 2 の信号のセグメントを時間修正することが前記第 2 のフレームの前記残差をタイムワープすることを含み、

前記方法が、前記第2のフレームの前記タイムワープされた残差からの情報に基づいて前記オーディオ信号の第3のフレームの残差をタイムワープすることを備え、前記第3のフレームが前記オーディオ信号中の前記第2のフレームに連続する、(33)に記載の方法。

(44) 前記第2の信号が前記第2のフレームの残差であり、前記第2の信号のセグメントを時間修正することが、(A)前記第1の信号の前記時間修正されたセグメントからの情報と、(B)前記第2のフレームの前記残差からの情報と、に基づいて前記第2の時間シフトを計算することを含む、(33)に記載の方法。

(45) 前記PRコーディング方式がリラックスドコード励振線形予測コーディング方式であり、前記非PRコーディング方式が、(A)ノイズ励起線形予測コーディング方式と、(B)修正離散コサイン変換コーディング方式と、(C)プロトタイプ波形補間コーディング方式と、のうちの1つである、(33)に記載の方法。

(46) 前記非PRコーディング方式が修正離散コサイン変換コーディング方式である、(33)に記載の方法。

(47) 前記第1のフレームを符号化することが、
符号化残差を得るために前記第1のフレームの残差に対して修正離散コサイン変換(MDCT)演算を実行することと、

復号残差を得るために前記符号化残差に基づく信号に対して逆MDCT演算を実行することと、

を含み、

前記第1の信号が前記復号残差に基づく、(33)に記載の方法。

(48) 前記第1のフレームを符号化することが、

前記第1の信号である、前記第1のフレームの残差を発生させることと、

前記第1の信号のセグメントを時間修正することに続いて、符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換演算を実行することと、

前記符号化残差に基づいて第1の符号化フレームを生成することと、

を含む、(33)に記載の方法。

(49) 前記第1の信号がM個のサンプルの長さを有し、前記第2の信号がM個のサンプルの長さを有し、

前記第1のフレームを符号化することが、前記時間修正されたセグメントを含む、前記第1の信号のM個のサンプルと、前記第2の信号の $3M/4$ 個以下のサンプルと、に基づくM個の修正離散コサイン変換(MDCT)係数の組を生成することを含む、(33)に記載の方法。

(50) 前記第1の信号がM個のサンプルの長さを有し、前記第2の信号がM個のサンプルの長さを有し、

前記第1のフレームを符号化することが、(A)前記時間修正されたセグメントを含む、前記第1の信号のM個のサンプルを含み、(B)ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで始まり、(C)ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで終わる、 $2M$ 個のサンプルのシーケンスに基づくM個の修正離散コサイン変換(MDCT)係数の組を生成することを含む、(33)に記載の方法。

(51) オーディオ信号のフレームを処理するための装置であって、

第1のコーディング方式に従って前記オーディオ信号の第1のフレームを符号化するための手段と、

ピッチ調整(PR)コーディング方式に従って前記オーディオ信号の第2のフレームを符号化するための手段と、

を備え、

前記第2のフレームが、前記オーディオ信号中の前記第1のフレームに後続し且つ連続し、

前記第1のコーディング方式が非PRコーディング方式であり、

第 1 のフレームを符号化するための前記手段が、前記第 1 のフレームに基づく第 1 の信号のセグメントを、第 1 の時間シフトに基づいて時間修正するための手段を含み、前記時間修正するための手段が、(A) 前記第 1 の時間シフトに従って前記第 1 の信号の前記セグメントを時間シフトすることと、(B) 前記第 1 の時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され、

前記第 2 のフレームを符号化するための手段が、前記第 2 のフレームに基づく第 2 の信号のセグメントを、第 2 の時間シフトに基づいて時間修正するための手段を含み、前記時間修正するための手段が、(A) 前記第 2 の時間シフトに従って前記第 2 の信号の前記セグメントを時間シフトすることと、(B) 前記第 2 の時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され、

前記第 2 の信号のセグメントを時間修正するための手段が、前記第 2 の信号の別のピッチパルスに対する前記セグメントのピッチパルスの位置を変化させるように構成され、前記第 2 の時間シフトが、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づく、方法。

(5 2) 前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、(5 1) に記載の装置。

(5 3) 前記第 1 及び第 2 の信号が重み付きオーディオ信号である、(5 1) に記載の装置。

(5 4) 前記第 2 の信号のセグメントを時間修正するための手段が、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づいて前記第 2 の時間シフトを計算するための手段を含み、

前記第 2 の時間シフトを計算するための手段が、前記第 2 のフレームからの情報に基づく遅延輪郭に、前記第 1 の信号の前記時間修正されたセグメントをマッピングするための手段を含む、(5 1) に記載の装置。

(5 5) 前記第 2 の時間シフトが、前記マッピングされたセグメントのサンプルと一時修正残差のサンプルとの間の相関に基づき、

前記一時修正残差が、(A) 前記第 2 のフレームの残差のサンプルと、(B) 前記第 1 の時間シフトと、に基づく、(5 4) に記載の装置。

(5 6) 前記第 2 の信号が前記第 2 のフレームの残差であり、前記第 2 の信号のセグメントを時間修正するための手段が、前記第 2 の時間シフトに従って前記残差の第 1 のセグメントを時間シフトするように構成され、前記方法が、

前記残差の前記時間修正された第 1 のセグメントからの情報に基づいて、前記第 2 の時間シフトとは異なる第 3 の時間シフトを計算するための手段と、

前記第 3 の時間シフトに従って前記残差の第 2 のセグメントを時間シフトするための手段と、を備える、(5 1) に記載の装置。

(5 7) 前記第 2 の信号が前記第 2 のフレームの残差であり、第 2 の信号のセグメントを時間修正するための手段が、(A) 前記第 1 の信号の前記時間修正されたセグメントからの情報と、(B) 前記第 2 のフレームの前記残差からの情報と、に基づいて前記第 2 の時間シフトを計算するための手段を含む、(5 1) に記載の装置。

(5 8) 前記第 1 のフレームを符号化するための手段が、前記第 1 の信号である、前記第 1 のフレームの残差を発生させるための手段と、符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に対して修正離散コサイン変換演算を実行するための手段と、を含み、

前記第 1 のフレームを符号化するための手段が、前記符号化残差に基づいて第 1 の符号化フレームを生成するように構成される、(5 1) に記載の装置。

(5 9) 前記第 1 の信号が M 個のサンプルの長さを有し、前記第 2 の信号が M 個のサンプルの長さを有し、

前記第 1 のフレームを符号化するための手段が、前記時間修正されたセグメントを含む、前記第 1 の信号の M 個のサンプルと、前記第 2 の信号の $3 M / 4$ 個以下のサンプルと、に基づく M 個の修正離散コサイン変換 (M D C T) 係数の組を生成するための手段を含む、(5 1) に記載の装置。

(6 0) 前記第 1 の信号が M 個のサンプルの長さを有し、前記第 2 の信号が M 個のサンプルの長さを有し、

前記第 1 のフレームを符号化するための手段が、(A) 前記時間修正されたセグメントを含む、前記第 1 の信号の M 個のサンプルを含み、(B) ゼロ値の少なくとも $M / 8$ 個のサンプルのシーケンスで始まり、(C) ゼロ値の少なくとも $M / 8$ 個のサンプルのシーケンスで終わる、 $2 M$ 個のサンプルのシーケンスに基づく M 個の修正離散コサイン変換 (M D C T) 係数の組を生成するための手段を含む、(5 1) に記載の装置。

(6 1) オーディオ信号のフレームを処理するための装置であって、

第 1 のコーディング方式に従って前記オーディオ信号の第 1 のフレームを符号化するように構成された第 1 のフレームエンコーダと、

ピッチ調整 (P R) コーディング方式に従って前記オーディオ信号の第 2 のフレームを符号化するように構成された第 2 のフレームエンコーダと、

を備え、

前記第 2 のフレームが、前記オーディオ信号中の前記第 1 のフレームに後続し且つ連続し、

前記第 1 のコーディング方式が非 P R コーディング方式であり、

前記第 1 のフレームエンコーダが、前記第 1 のフレームに基づく第 1 の信号のセグメントを、第 1 の時間シフトに基づいて時間修正するように構成された第 1 の時間修正器を含み、前記第 1 の時間修正器が、(A) 前記第 1 の時間シフトに従って前記第 1 の信号の前記セグメントを時間シフトすることと、(B) 前記第 1 の時間シフトに基づいて前記第 1 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され、

前記第 2 のフレームエンコーダが、前記第 2 のフレームに基づく第 2 の信号のセグメントを、第 2 の時間シフトに基づいて時間修正するように構成された第 2 の時間修正器を含み、前記第 2 の時間修正器が、(A) 前記第 2 の時間シフトに従って前記第 2 の信号の前記セグメントを時間シフトすることと、(B) 前記第 2 の時間シフトに基づいて前記第 2 の信号の前記セグメントをタイムワープすることと、のうちの 1 つを実行するように構成され、

前記第 2 の時間修正器が、前記第 2 の信号の別のピッチパルスに対する第 2 の信号の前記セグメントのピッチパルスの位置を変化させるように構成され、

前記第 2 の時間シフトが、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づく、方法。

(6 2) 前記第 1 の信号が前記第 1 のフレームの残差であり、前記第 2 の信号が前記第 2 のフレームの残差である、(6 1) に記載の装置。

(6 3) 前記第 1 及び第 2 の信号が重み付きオーディオ信号である、(6 1) に記載の装置。

(6 4) 前記第 2 の時間修正器が、前記第 1 の信号の前記時間修正されたセグメントからの情報に基づいて前記第 2 の時間シフトを計算するように構成された時間シフト計算器を含み、前記時間シフト計算器が、前記第 2 のフレームからの情報に基づく遅延輪郭に、前記第 1 の信号の前記時間修正されたセグメントをマッピングするように構成されたマッパーを含む、(6 1) に記載の装置。

(6 5) 前記第 2 の時間シフトが、前記マッピングされたセグメントのサンプルと一時修正残差のサンプルとの間の相関に基づき、

前記一時修正残差が、(A) 前記第 2 のフレームの残差のサンプルと、(B) 前記第 1

の時間シフトとに基づき、(64)に記載の装置。

(66) 前記第2の信号が前記第2のフレームの残差であり、
前記第2の時間修正器が、前記第2の時間シフトに従って前記残差の第1のセグメント
を時間シフトするように構成され、

前記時間シフト計算器が、前記残差の前記時間修正された第1のセグメントからの情報
に基づいて、前記第2の時間シフトとは異なる第3の時間シフトを計算するように構成さ
れ、

前記第2の時間シフト器が、前記第3の時間シフトに従って前記残差の第2のセグメン
トを時間シフトするように構成される、(61)に記載の装置。

(67) 前記第2の信号が前記第2のフレームの残差であり、前記第2の時間修正器
が、(A)前記第1の信号の前記時間修正されたセグメントからの情報と、(B)前記第
2のフレームの前記残差からの情報と、に基づいて前記第2の時間シフトを計算するよう
に構成された時間シフト計算器を含む、(61)に記載の装置。

(68) 前記第1のフレームエンコーダが、
前記第1の信号である、前記第1のフレームの残差を発生させるように構成された残差
発生器と、

符号化残差を得るために、前記時間修正されたセグメントを含む、前記発生した残差に
対して修正離散コサイン変換(MDCT)演算を実行するように構成されたMDCTモジ
ュールと、

を含み、

前記第1のフレームエンコーダが、前記符号化残差に基づいて第1の符号化フレームを
生成するように構成される、(61)に記載の装置。

(69) 前記第1の信号がM個のサンプルの長さを有し、前記第2の信号がM個のサ
ンプルの長さを有し、

前記第1のフレームエンコーダが、前記時間修正されたセグメントを含む、前記第1の
信号のM個のサンプルと、前記第2の信号の $3M/4$ 個以下のサンプルと、に基づくM個
の修正離散コサイン変換(MDCT)係数の組を生成するように構成されたMDCTモジ
ュールを含む、(61)に記載の装置。

(70) 前記第1の信号がM個のサンプルの長さを有し、前記第2の信号がM個のサ
ンプルの長さを有し、

前記第1のフレームエンコーダが、(A)前記時間修正されたセグメントを含む、前記
第1の信号のM個のサンプルを含み、(B)ゼロ値の少なくとも $M/8$ 個のサンプルのシ
ーケンスで始まり、(C)ゼロ値の少なくとも $M/8$ 個のサンプルのシーケンスで終わる
、 $2M$ 個のサンプルのシーケンスに基づくM個の修正離散コサイン変換(MDCT)係数
の組を生成するように構成されたMDCTモジュールを含む、(61)に記載の装置。

(71) プロセッサによって実行されると前記プロセッサに、

第1のコーディング方式に従って前記オーディオ信号の第1のフレームを符号化するこ
とと、

ピッチ調整(PR)コーディング方式に従って前記オーディオ信号の第2のフレームを
符号化することと、

を行わせる命令を備えるコンピュータ可読媒体であって、

前記第2のフレームが、前記オーディオ信号中の前記第1のフレームに後続し且つ連続
し、

前記第1のコーディング方式が非PRコーディング方式であり、

前記プロセッサによって実行されると前記プロセッサに第1のフレームを符号化するこ
とを行わせる命令が、前記第1のフレームに基づく第1の信号のセグメントを、第1の時
間シフトに基づいて時間修正する命令を含み、前記時間修正する命令が、(A)前記第1
の時間シフトに従って前記第1の信号の前記セグメントを時間シフトする命令と、(B)
前記第1の時間シフトに基づいて前記第1の信号の前記セグメントをタイムワープする命
令と、のうちの1つを含み、

前記プロセッサによって実行されると前記プロセッサに第2のフレームを符号化することを
行わせる命令が、前記第2のフレームに基づく第2の信号のセグメントを、第2の時
間シフトに基づいて時間修正する命令を含み、前記時間修正する命令が、(A)前記第2
の時間シフトに従って前記第2の信号の前記セグメントを時間シフトする命令と、(B)
前記第2の時間シフトに基づいて前記第2の信号の前記セグメントをタイムワープする命
令と、のうちの1つを含み、

前記第2の信号のセグメントを時間修正する命令が、前記第2の信号の別のピッチパル
スに対する前記セグメントのピッチパルスの位置を変化させる命令を含み、

前記第2の時間シフトが、前記第1の信号の前記時間修正されたセグメントからの情報
に基づく、コンピュータ可読媒体。

【手続補正7】

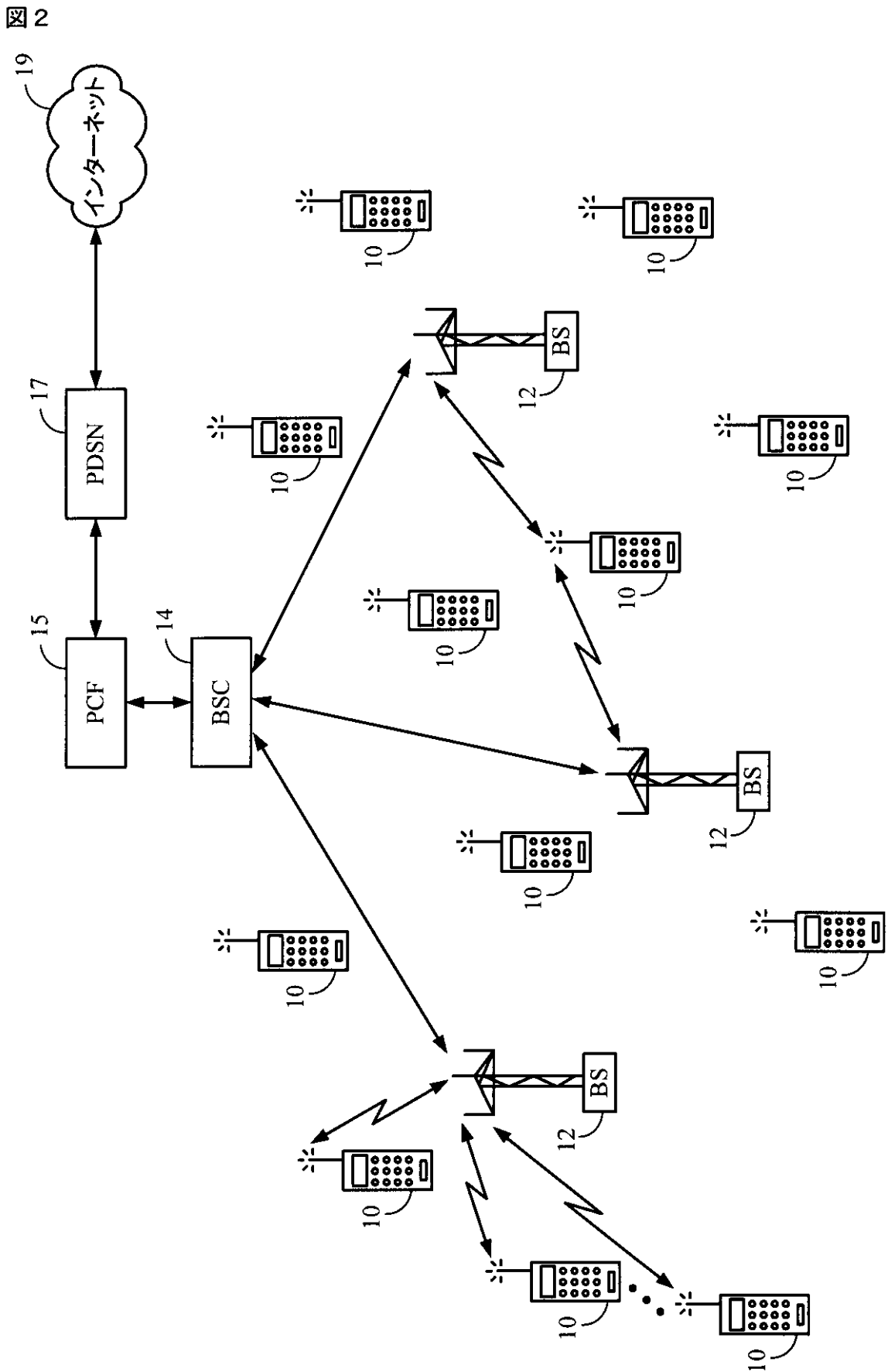
【補正対象書類名】図面

【補正対象項目名】図2

【補正方法】変更

【補正の内容】

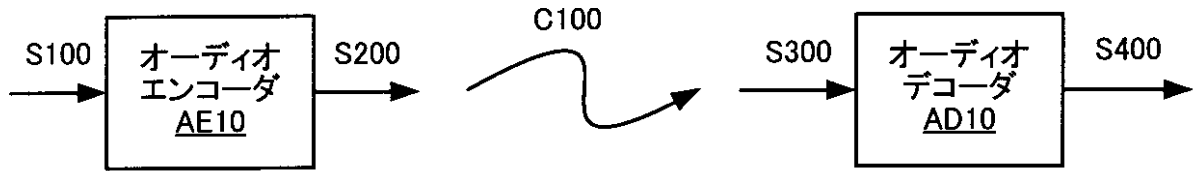
【 図 2 】



【 手続補正 8 】

- 【補正対象書類名】図面
- 【補正対象項目名】図3 a
- 【補正方法】変更
- 【補正の内容】
- 【図3 a】

図3 a



フロントページの続き

- (74)代理人 100075672
弁理士 峰 隆司
- (74)代理人 100153051
弁理士 河野 直樹
- (74)代理人 100140176
弁理士 砂川 克
- (74)代理人 100158805
弁理士 井関 守三
- (74)代理人 100172580
弁理士 赤穂 隆雄
- (74)代理人 100179062
弁理士 井上 正
- (74)代理人 100124394
弁理士 佐藤 立志
- (74)代理人 100112807
弁理士 岡田 貴志
- (74)代理人 100111073
弁理士 堀内 美保子
- (72)発明者 ビベク・ラジェンドラン
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7
7 5
- (72)発明者 アナンサパドマナブハン・エー．．カンドハダイ
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7
7 5
- (72)発明者 ベンカテシュ・クリシュナン
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7
7 5

【外国語明細書】

2013242579000001.pdf