

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5814341号
(P5814341)

(45) 発行日 平成27年11月17日(2015.11.17)

(24) 登録日 平成27年10月2日(2015.10.2)

(51) Int.Cl.		F I			
G 1 0 L	19/008	(2013.01)	G 1 0 L	19/008	1 0 0
G 1 0 L	19/02	(2013.01)	G 1 0 L	19/02	1 5 0
G 1 0 L	19/00	(2013.01)	G 1 0 L	19/00	3 3 0 B

請求項の数 15 (全 33 頁)

(21) 出願番号	特願2013-503104 (P2013-503104)	(73) 特許権者	510185767
(86) (22) 出願日	平成23年4月6日(2011.4.6)		ドルビー・インターナショナル・アーベー
(65) 公表番号	特表2013-525830 (P2013-525830A)		オランダ王国, セーエン アムステルダム
(43) 公表日	平成25年6月20日(2013.6.20)		ズイドースト 1101, ヘリケルベル
(86) 国際出願番号	PCT/EP2011/055374		グウェグ 1-35, アポロ ビルディン
(87) 国際公開番号	W02011/124621		グ 3エー
(87) 国際公開日	平成23年10月13日(2011.10.13)	(74) 代理人	100107766
審査請求日	平成24年11月27日(2012.11.27)		弁理士 伊東 忠重
(31) 優先権主張番号	61/322, 458	(74) 代理人	100070150
(32) 優先日	平成22年4月9日(2010.4.9)		弁理士 伊東 忠彦
(33) 優先権主張国	米国 (US)	(74) 代理人	100091214
			弁理士 大貫 進介
		(74) 代理人	100135079
			弁理士 宮崎 修

最終頁に続く

(54) 【発明の名称】 MDC Tベース複素予測ステレオ符号化

(57) 【特許請求の範囲】

【請求項1】

複素予測ステレオ符号化によりステレオ信号を提供するデコーダシステムであって、
 ダウンミックス信号と残差信号の第1の周波数領域表示に基づいて、前記ステレオ信号を生成するように構成されたアップミックス段階であって、各第1の周波数領域表示は多次元空間の第1の副空間で表された対応する信号のスペクトルコンテンツを表す第1のスペクトル成分を有するアップミックス段階を有し、前記アップミックス段階は、
 前記ダウンミックス信号の第1の周波数領域表示に基づき、前記ダウンミックス信号の第2の周波数領域表示を計算するモジュールであって、前記第2の周波数領域表示は、前記第1の副空間に含まれない多次元空間の一部を含む、前記多次元空間の第2の副空間で表された信号のスペクトルコンテンツを表す第2のスペクトル成分を有する、モジュールと、

ビットストリーム信号にエンコードされた前記ダウンミックス信号の第1と第2の周波数領域表示と、前記残差信号の第1の周波数領域表示と、複素予測係数とに基づいてサイド信号を計算する重み付け加算器と、

前記ダウンミックス信号と前記サイド信号の第1の周波数領域表示に基づいて、前記ステレオ信号を計算する和・差段階を有するアップミックス段階とを有し、

前記アップミックス段階は前記ダウンミックス信号と残差信号に、独立の帯域幅制限を適用するように構成された、デコーダシステム。

【請求項2】

適用される前記帯域幅制限は、各信号に対して、復号する最高周波数帯を示す2つのデータフィールドによりシグナリングされる、請求項1に記載のデコーダシステム。

【請求項3】

前記データフィールドのそれぞれが最大sfbの値として符号化されるMP EGビットストリームを受け取るように構成される、請求項2に記載のデコーダシステム。

【請求項4】

ビットストリーム信号に基づいて、前記ダウンミックス信号と残差信号の前記第1の周波数領域表示を提供する、前記アップミックス段階の上流に配置された逆量子化段階をさらに有する、請求項1ないし3いずれか一項に記載のデコーダシステム。

【請求項5】

前記第1のスペクトル成分は前記第1の副空間で表された実数値を有し、
前記第2のスペクトル成分は前記第2の副空間で表された虚数値を有し、
任意的に、前記第1のスペクトル成分は、離散余弦変換DCT又は修正離散余弦変換MDCTのうち一方により求められ、
任意的に、前記第2のスペクトル成分は、離散正弦変換DST又は修正離散正弦変換MDSTのうち一方により求められる、
請求項1ないし3いずれか一項に記載のデコーダシステム。

【請求項6】

前記アップミックス段階の上流に配置された少なくとも1つの時間的ノイズシェーピングTNSモジュールと、

前記アップミックス段階の下流に配置された少なくとも1つのさらなるTNSモジュールと、

(a)前記アップミックス段階の上流の前記TNSモジュール、または(b)前記アップミックス段階の下流にある前記さらなるTNSモジュールのうちいずれかを選択的にアクティブ化するセレクト装置とを有する

請求項1ないし5いずれか一項に記載のデコーダシステム。

【請求項7】

前記ダウンミックス信号は連続した時間フレームにパーティションされ、各時間フレームは複素予測係数の値に関連し、

前記ダウンミックス信号の第2の周波数領域表示を計算するモジュールは、前記複素予測係数の虚部の絶対値が時間フレームの所定の許容値より小さいことに応じて、自身を非アクティブ化して、その時間フレームに対して出力を生成しないようにするように構成された、

請求項5に記載のデコーダシステム。

【請求項8】

前記第1のスペクトル成分は変換係数の時間フレームに配置された変換係数であり、各ブロックは時間領域信号の時間セグメントへの変換の適用により生成され、

前記ダウンミックス信号の第2の周波数領域表示を計算するモジュールは、

前記第1のスペクトル成分から第1の中間成分を求め、

インパルス応答の少なくとも一部分により前記第1のスペクトル成分の組み合わせを構成して、第2の中間成分を求め、

前記第2の中間成分から前記第2のスペクトル成分を求める、ように構成された、

請求項1ないし7いずれか一項に記載のデコーダシステム。

【請求項9】

前記ダウンミックス信号の第2の周波数領域表示を計算するモジュールは、(a)第1のスペクトル成分の同時時間フレームと、

(b)第1のスペクトル成分の同時時間フレーム及び先行時間フレームと、

(c)第1のスペクトル成分の同時時間フレーム、先行時間フレーム、及び後続時間フレームとのうちの1つに基づいて第2のスペクトル成分の各時間フレームを求めるように構成された、

10

20

30

40

50

請求項 8 に記載のデコーダシステム。

【請求項 10】

前記ダウンミックス信号の第 2 の周波数領域表示を計算するモジュールは、少なくとも 2 つの時間的に隣接した、及び / 又は周波数的に隣接した第 1 のスペクトル成分の組み合わせにより決定された近似的な第 2 のスペクトル成分を有する近似的第 2 のスペクトル表示を計算するように構成されている、

請求項 1 ないし 9 いずれか一項に記載のデコーダシステム。

【請求項 11】

前記ステレオ信号は時間領域で表され、デコーダシステムはさらに、

(a) パススルー段階、又は (b) 和・差段階のいずれかとして機能でき、それにより直接及び同時符号化ステレオ入力信号の間を切り替えられる、前記逆量子化段階と前記アップミックス段階の間に配置されたスイッチングアセンブリと、

前記ステレオ信号の時間領域表示を計算するように構成された逆変換段階と、

前記逆変換段階の上流に配置され、これを、(a) 複素予測により求めたステレオ信号が前記逆変換段階に供給される、前記アップミックス段階の下流のポイントに、又は (b) 直接ステレオ符号化により求めたステレオ信号が前記逆変換段階に供給される、前記スイッチングアセンブリの下流であり、かつ前記アップミックス段階の上流であるポイントのいずれかに選択的に接続するように構成されたセレクト装置とを有する、

請求項 4 に記載のデコーダシステム。

【請求項 12】

前記ダウンミックス信号の第 2 の周波数領域表示を計算するモジュールは、

前記多次元空間の第 1 の副空間における各信号の第 1 の周波数領域表示に基づき、前記ダウンミックス信号及び / 又は彩度信号の時間領域の表示を計算する逆変換段階と、

前記信号の時間領域表示に基づき各信号の第 2 の周波数領域表示を計算する変換段階とを有し、

好ましくは、前記逆変換段階は逆修正離散余弦変換 M D C T を行い、前記変換段階は修正離散余弦変換 M D S T を行う、

請求項 4 に記載のデコーダシステム。

【請求項 13】

前記ステレオ信号は時間領域で表され、デコーダシステムはさらに、

(a) 同時ステレオ符号化に用いるパススルー段階、又は (b) 直接ステレオ符号化に用いる和・差段階のいずれかとして機能できる、前記逆量子化段階と前記アップミックス段階の間に配置されたスイッチングアセンブリと、

前記サイド信号の時間領域表示を計算する、前記アップミックス段階に配置されたさらなる逆変換段階と、

(a) 前記スイッチングアセンブリの下流かつ前記アップミックス段階の上流のポイントに接続されたさらなる和・差段階、または (b) 前記スイッチングアセンブリから得られるダウンミックス信号と前記重み付け加算器から得られるサイド信号のいずれかに選択的に接続されるように構成された、前記逆変換段階の上流に配置されたセレクト装置とを有する、

請求項 12 に記載のデコーダシステム。

【請求項 14】

複素予測ステレオ符号化により入力ステレオ信号を出力ステレオ信号にアップミックスする復号方法であって、

前記入力ステレオ信号は、ダウンミックスチャンネルと残差チャンネルの第 1 の周波数領域表示と、複素予測係数とを含み、

前記第 1 の周波数領域表示の各々は、多次元空間の第 1 の副空間で表された対応する信号のスペクトルコンテンツを表す第 1 のスペクトル成分を有し、

前記方法は、アップミックス段階により実行され、

前記ダウンミックスチャンネルの第 1 の周波数領域表示に基づき、前記ダウンミックス

10

20

30

40

50

チャンネルの第2の周波数領域表示を計算するサブステップであって、前記第2の周波数領域表示は、前記第1の副空間に含まれない多次元空間の一部を含む、前記多次元空間の第2の副空間で表された信号のスペクトルコンテンツを表す第2のスペクトル成分を有する、ステップと、

前記ダウンミックスチャンネルの第1と第2の周波数領域表示と、前記残差チャンネルの第1の周波数領域表示と、前記複素予測係数とに基づいて、サイドチャンネルを計算するステップと、を有し、

前記ダウンミックスチャンネルと残差信号には、独立の帯域幅制限を適用する、復号方法。

【請求項15】

10

コンピュータにより実行されたとき、前記コンピュータに請求項14に記載の方法を実行させるコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

ここに開示する発明は概してステレオオーディオ符号化に関し、より詳しくは周波数領域における複素予測を用いるステレオ符号化の手法に関する。

【背景技術】

【0002】

ステレオ信号の左(L)チャンネルと右(R)チャンネルの同時符号化(joint coding)により、LとRを独立に符号化するのと比較して、符号化が効率的になる。同時ステレオ符号化に対する一般的なアプローチはミッド/サイド(M/S)符号化である。ここで、ミッド(M)信号はL信号とR信号を加えることにより構成され、例えばM信号は

20

【数1】

$$M = (L + R)/2$$

により得られる。また、サイド(S)信号は2つのチャンネルLRを引くことにより構成され、例えばS信号は

30

【数2】

$$S = (L - R)/2$$

により得られる。M/S符号化の場合、L信号とR信号ではなく、M信号とS信号が符号化される。

【0003】

MPEG(Moving Picture Experts Group)のAAC(Advanced Audio Coding)標準(標準文書ISO/IEC 13818-7を参照)では、時間および周波数可変で、L/Rステレオ符号化とM/Sステレオ符号化を選択できる。このように、ステレオエンコーダは、ステレオ信号のある周波数帯域にL/R符号化を適用でき、そのステレオ信号の他の周波数帯域のエンコードにはM/S符号化が使われる(周波数可変)。さらに、エンコーダは時間的にL/R符号化とM/S符号化を切り替えられる(時間可変)。MPEG AACでは、ステレオエンコーディングは周波数領域で、より具体的にはMDCT(修正離散余弦変換)領域で行われる。これにより、周波数および時間的に可変に、L/R符号化またはM/S符号化のいずれかを適応的に選択できる。

40

【0004】

パラメトリックステレオ符号化は、ステレオオーディオ信号を、モノラル信号と、ステレオパラメータとなる少量のサイド情報として、効率的に符号化する手法である。これは

50

MPEG-4 オーディオ標準（標準文書 ISO/IEC 14496-3 を参照）の一部である。モノラル信号はどのオーディオ符号化器を用いてもエンコードできる。ステレオパラメータはモノビットストリーム (mono bit stream) の付属部分に組み込まれるので、完全に前方互換かつ後方互換となる。デコーダでは、モノラル信号が最初に復号され、その後ステレオパラメータを用いてステレオ信号が再構成される。復号されたモノ信号を無相関化した信号 (decorrelated version) は、モノ信号との相互相関がゼロである。この無相関化信号を、無相関化器 (decorrelator) により、例えば遅延ラインを含む適当なオールパスフィルタ (all-pass filter) により生成する。基本的に、無相関化信号はモノ信号と同じスペクトル的・時間的エネルギー分布を有する。モノラル信号は無相関化信号とともにアップミックスプロセスに入力される。このプロセスは、ステレオパラメータにより制御され、ステレオ信号を再構成する。さらに詳しい情報は、非特許文献 1 を参照されたい。

10

【0005】

MPEG サラウンド (MPS ; ISO/IEC 23003-1 及び非特許文献 2 を参照) は、パラメトリックステレオ符号化の原理に残差符号化の原理を組み合わせるものであり、無相関化信号を送信される残差で置き換え、知覚できる音質を改善している。残差符号化は、マルチチャンネル信号をダウンミックスし、任意的に空間的キューを抽出することにより行われる。ダウンミックスプロセスにおいて、エラー信号を表す残差信号が計算され、エンコードされて送信される。残差信号はデコーダにおいて無相関化信号の代わりになる。ハイブリッドアプローチでは、残差信号は一定の周波数帯域において、好ましくは比較的低い帯域において、無相関化信号に取って代わる。

20

【0006】

現在の MPEG Unified Speech and Audio Coding (USAC) システムは、図 1 に 2 つの例を示したが、デコーダはコアデコーダの下流に位置する複素値直交ミラーフィルタ (QMF) バンクを有する。このフィルタバンクの出力として得られる QMF 表現は、複素値であり、それゆえ 2 倍にオーバーサンプルされており、ダウンミックス信号 (すなわちミッド信号) M と残差信号 D として構成できる。これには複素値成分を有するアップミックス行列を使うことができる。(QMF 領域の) L 信号と R 信号は

【数 3】

$$\begin{bmatrix} L \\ R \end{bmatrix} = g \begin{bmatrix} 1 - \alpha & 1 \\ 1 + \alpha & -1 \end{bmatrix} \begin{bmatrix} M \\ D \end{bmatrix}$$

30

として得られる。ここで、 g は実数値のゲインファクタであり、 α は複素値の予測係数である。 α は残差信号 D のエネルギーが最小になるように選択するのが好ましい。ゲインファクタは規格化により、すなわち和信号のパワーが左右信号のパワーの和と等しくなるようにして決定できる。L 信号と R 信号それぞれの実部と虚部は互いに冗長性があり、原理的には一方を他方に基づき計算できる。しかし、後で可聴なエイリアシングアーティファクトを生ぜずにスペクトル帯域レプリケーション (SBR) デコーダを使えるという利益がある。モノ・ステレオアップミックスなどその他の時間または周波数適応的信号処理 (図示せず) と関連するアーティファクトを防止する目的で、オーバーサンプルされた信号表現の利用も同様の理由で選択される。逆 QMF フィルタリングがデコーダにおける最後の処理ステップである。信号の帯域制限的 QMF 表現により、帯域制限残差手法と「残差フィル」手法が使える。これらの手法をこのタイプのデコーダに組み込むことができる。

40

【0007】

上記の符号化構成は低ビットレートの場合、一般的には 80 kb/s 未満の場合には良く適合するが、計算複雑性の点で高ビットレートの場合には最適なものではない。より詳しく言うと、高ビットレートでは、一般的に SBR ツールは (符号化効率の改善にならないので) 使われない次に、SBR 段階を有さないデコーダでは、複素値アップミックス行列があるが故に QMF フィルタバンクを使うが、これは計算量が多く遅延を生じる (10

50

24 サンプルのフレーム長では、QMF 分析 / 合成フィルタバンクにより 961 サンプルの遅延が生じる)。これは、もっと効率的な符号化構成の必要性を明らかに示している。

【先行技術文献】

【非特許文献】

【0008】

【非特許文献 1】H. Purnhagen 著「Low Complexity Parametric Stereo Coding in MPEG-4」, Proc. of the 7th Int. Conference on Digital Audio Effects (DAFx'04), Naples, Italy, October 5-8, 2004, pages 163-168

【非特許文献 2】J. Herre et al. 著「MPEG Surround - The ISO/MPEG Standard for Efficient and Compatible Multi-Channel Audio Coding」, Audio Engineering Convention Paper 7084, 122 Convention, May 5-8, 2007

【発明の概要】

【課題を解決するための手段】

【0009】

本発明の一目的は、高ビットレートのレンジにおいても計算効率の高いステレオ符号化をする方法と装置を提供することである。

【0010】

本発明は、独立請求項に規定した、符号化及び復号をするための、それぞれコーダ及びデコーダ、コーディング及びデコーディング方法、及びコンピュータプログラム製品を提供することにより、この目的を達成する。従属項は本発明の実施形態を規定している。

【0011】

第 1 の態様では、本発明は次のシステムを提供する。すなわち、複素予測ステレオ符号化によりステレオ信号を提供するデコーダシステムであって：

ダウンミックス信号 (M) と残差信号 (D) の第 1 の周波数領域表示に基づいて、前記ステレオ信号を生成するように構成されたアップミックス段階であって、各第 1 の周波数領域表示は多次元空間の第 1 の副空間で表された対応する信号のスペクトルコンテンツを表す第 1 のスペクトル成分を有するアップミックス段階を有し、前記アップミックス段階は、

前記ダウンミックス信号の第 1 の周波数領域表示に基づき、前記ダウンミックス信号の第 2 の周波数領域表示を計算するモジュールであって、前記第 2 の周波数領域表示は、前記第 1 の副空間に含まれない多次元空間の一部を含む、前記多次元空間の第 2 の副空間で表された信号のスペクトルコンテンツを表す第 2 のスペクトル成分を有する、モジュールと、

前記ビットストリーム信号にエンコードされた前記ダウンミックス信号の第 1 と第 2 の周波数領域表示と、前記残差信号の第 1 の周波数領域表示と、複素予測係数 () とに基づいてサイド信号 (S) を計算する重み付け加算器とを有するアップミックス段階と

前記ダウンミックス信号と前記サイド信号の第 1 の周波数領域表示に基づいて、前記ステレオ信号を計算する和・差段階を有し、

前記アップミックス段階は、さらに、前記ダウンミックス信号と残差信号が前記和・差段階に直接供給されるパススルーモードで動作可能である。

【0012】

第 2 の態様では、本発明は次のシステムを提供する。すなわち、複素予測ステレオ符号化によりビットストリーム信号によりステレオ信号をエンコードするエンコーダシステムであって、

複素予測係数を推定する推定器と、

(a) 前記ステレオ信号を、前記複素予測係数の値により決定される関係を有するダウンミックス信号と残差信号の周波数領域表示に変換するように動作可能な符号化段階と、

前記符号化段階と推定器から出力を受け取り、これを前記ビットストリーム信号にエンコードするマルチプレクサとを有する。

【0013】

本発明の第3と第4の態様では、ステレオ信号をビットストリームにエンコードする方法と、ビットストリームを少なくとも1つのステレオ信号に復号する方法が提供される。各方法の技術的特徴は、それぞれエンコーダシステムとデコーダシステムの技術的特徴を同様である。第5と第6の態様では、本発明は、各方法をコンピュータで実行する命令を含むコンピュータプログラム製品を提供する。

【0014】

本発明は、MPEG U SACシステムにおける統一ステレオ符号化の優位性からの利益を受ける。これらの優位性は、SBRが一般的には利用されない、QMFベースアプローチに伴う計算上の複雑さを大幅に増大することなく、高ビットレートでも保存され、これが可能になる理由は、クリティカルサンプリングされたMDCT変換は、MPEG U SAC変換の基本であるが、ダウンミックス及び残差チャンネルの符号オーディオ帯域が同じであり、アップミックスプロセスが非相関化を含まない場合には少なくとも、本発明により、複素予測ステレオ符号化でも使える。これは、追加的なQMF変換がもはや必要ないことを意味する。QMF領域における複素予測ステレオ符号化の代表的な実施形態は、従来のL/RまたはM/Sステレオと比較して、1単位時間当たりの演算数を大幅に増やす。そのため、本発明による符号化装置は、控えめな計算負荷により高音質を提供するため、かかるビットレートで競争力があるように思われる。

【0015】

当業者は気づくように、アップミックス段階はパススルーモードでも動作可能であるという事実により、デコーダは、エンコーダ側での判断により、従来の直接符号化または同時符号化、及び複素予測符号化により、適応的に復号できる。よって、デコーダが音質レベルを従来の直接L/Rステレオ符号化や同時M/Sステレオ符号化より積極的に上げられない場合に、少なくとも、同じレベルを維持することを保証できる。よって、本発明のこの態様によるデコーダは、機能的観点から、背景技術に対して上位集合 (superset) と見なせる。

【0016】

QMFベース予測符号化ステレオに対する優位性として、(任意に小さくできる量子化誤差を除いて) 信号の完全再構成が可能である。

【0017】

このように、本発明は、複素予測による変換ベースのステレオ符号化をする符号化装置を提供する。好ましくは、本発明による装置は、複素予測ステレオ符号化に限定されず、背景技術による直接L/Rステレオ符号化や同時M/Sすれてお符号化でも動作可能であり、具体的なアプリケーションや特定の時間中に最も適した符号化方法を選択できる。

【0018】

信号のオーバーサンプリングされた表示 (例えば、複素表示) は、第1と第2のスペクトル成分を両方とも含み、本発明による複素予測の基礎として用いられ、よって、かかるオーバーサンプリングされた表示を計算するモジュールが、本発明によるエンコーダシステムとデコーダシステムに構成される。スペクトル成分は、多次元空間の第1と第2の副空間を指す。これは、有限のサンプリング周波数でサンプリングされた、所与の時間的長さ (例えば、所定の時間フレームの長さ) の、一組の時間依存関数である。この多次元空間中の関数は基底関数の有限の重み付け和により近似できることは周知である。

【0019】

当業者には明らかなように、デコーダと協働するように構成されたエンコーダは、エンコードされた信号の忠実な再生を可能とするように、予測符号化のベースとなるオーバーサンプリングされた表示を提供する等価なモジュールが備えられている。かかる等価なモジュールは、同じ又は類似したモジュールか、同じ又は類似した伝達特性を有するモジュールである。特に、エンコーダとデコーダのモジュールは、それぞれ、等価な数学的演算

10

20

30

40

50

を実行するコンピュータプログラムを実行する類似した、または非類似のユニットであってもよい。

【 0 0 2 0 】

デコーダシステムやエンコーダシステムのある実施形態では、第 1 のスペクトル成分は第 1 の副空間で表された実数値を有し、第 2 のスペクトル成分は第 2 の副空間で表された虚数値を有する。第 1 と第 2 のスペクトル成分は共に、信号の複素スペクトル表示を構成する。第 1 の副空間は第 1 の組の基底関数の線形スパンであり、第 2 の副空間は第 2 の基底関数の組の線形スパンであり、その一部は第 1 の組の基底関数とは線形独立である。

【 0 0 2 1 】

一実施形態では、複素表示を計算するモジュールは、実・虚変換、すなわち、信号の実スペクトル表示に基づき、離散時間信号のスペクトルの虚数部を計算するモジュールである。この変換は、高調波分析やヒューリスティック関係からの式など、厳密な、又は近似的な数学的關係に基づく。

10

【 0 0 2 2 】

デコーダシステム又はエンコーダシステムのある実施形態では、第 1 のスペクトル成分は、離散時間領域信号の時間・周波数領域変換により、好ましくはフーリエ変換により、例えば離散余弦変換 (DCT)、修正離散余弦変換 (MDCT)、離散正弦変換 (DST)、修正離散正弦変換 (MDCT)、高速フーリエ変換 (FFT)、素因子ベース (prime-factor-based) フーリエアルゴリズムなどにより求められる。最初の 4 つの場合には、第 2 のスペクトル成分は DST、MDST、DCT、及び MDCT によりそれぞれ求めら

20

【 0 0 2 3 】

一実施形態では、デコーダシステムは、少なくとも 1 つの時間的ノイズシェーピングモジュール (TNS モジュール、すなわち TNS フィルタ) を含み、これはアップミックス段階の上流に配置される。一般的に言って、TNS の使用により、過渡成分を有する信号の知覚される音質が改善され、TNS を有する本発明のデコーダシステムの実施形態にも当てはまる。従来の L/R 及び M/S ステレオ符号化では、TNS フィルタは周波数領域における最後の処理ステップとして、逆変換の直前に適用される。しかし、複素予測ステレオ符号化の場合には、TNS フィルタをダウンミックス信号と残差信号に、すなわちアップミックス行列の前に適用すると有利であることが多い。言い換えると、TNS は左右チャンネルの線形結合に適用され、これにはいくつかの利点がある。最初に、ある状況では、TNS が例えばダウンミックス信号に対してのみ有利であることが分かる。次に、残差信号については TNS フィルタリングは省略でき、これは利用できる帯域幅の経済的な使用を意味する。TNS フィルタ係数は、ダウンミックス信号についてだけ送信されればよい。第 2 に、ダウンミックス信号のオーバーサンプリングされた表示の計算は (例えば、複素周波数領域表示を構成するために、MDST データは MDCT データから求められる)、複素予測符号化では必要であるが、ダウンミックス信号の時間領域表示が計算可能であることを要する。これは、ダウンミックス信号が、好ましくは一様に求めた MDCT スペクトルの時間シーケンスとして利用できることを意味する。TNS フィルタが、ダウンミックス/残差表示を左/右表示に変換するアップミックス行列の後にデコーダで適用された場合、ダウンミックス信号の TNS 残差 MDCT スペクトルのシーケンスのみが得られる。これにより、対応する MDST スペクトルの効率的な計算が非常に難しくなる。特に、左/右チャンネルが特性が異なる TNS フィルタを用いている場合にそうである。

30

40

【 0 0 2 4 】

強調しておくが、MDCT スペクトルの時間シーケンスが得られるかは、複素予測符号化の基礎として機能するようにフィットした MDST 表示を得るための絶対的な基準では

50

ない。実験的な証拠に加えて、この事実は、一般的に、T N Sによりフィルタされた残差信号が低周波のフィルタされていない残差信号に近似的に対応するように、T N Sが、例えば数キロヘルツより高い高周波のみに適用されるということにより説明できる。このように、本発明は、以下に説明するように、T N Sフィルタがアップミックス段階の上流以外に配置される、複素予測ステレオ符号化をするデコーダとして実施できる。

【 0 0 2 5 】

一実施形態では、デコーダシステムは、アップミックス段階の下流に配置された少なくとも1つのさらなるT N Sモジュールを含む。セレクタ装置により、アップミックス段階の上流のT N Sモジュールまたはアップミックス段階の下流のT N Sモジュール。ある状況下では、複素周波数領域表示の計算は、ダウンミックス信号の時間領域表示が計算可能である必要はない。さらに、上記の通り、デコーダは、複素予測符号化を適用せずに、直接または同時符号化モードで選択的に動作可能であり、T N Sモジュールを従来の場所に用いる、すなわち周波数領域における最後の処理ステップの1つとして用いる方が適している。

10

【 0 0 2 6 】

一実施形態では、デコーダシステムは、ダウンミックス信号の第2の周波数領域表示を計算するモジュールを非アクティブ化することにより、処理リソース及び場合によってはエネルギーを節約するように構成されている。前記ダウンミックス信号は連続した時間ブロックにパーティションされ、各時間ブロックは複素予測係数の値に関連する。この値は、デコーダと協働するエンコーダにより各時間ブロックに対する決定により決まる。さらに、この実施形態では、ダウンミックス信号の第2の周波数領域表示を計算するモジュールは、所与の時間ブロックについて、複素予測係数の虚部の絶対値がゼロであるか、所定の許容値より小さい場合、自分自身を非アクティブ化するように構成されている。モジュールの非アクティブ化は、この時間ブロックについてダウンミックス信号の第2の周波数領域表示を計算しないことを意味する。非アクティブ化をしない場合、第2の周波数領域表示（例えば、一組のM D S T係数）にはゼロ、またはデコーダのマシンイプシロン（四捨五入単位）又はその他の好適な閾値とほぼ同じオーダーの数がかけられる。

20

【 0 0 2 7 】

前記の実施形態をさらに発展させたものでは、ダウンミックス信号がパーティションされる時間ブロックのサブレベルで処理リソースの節約が為される。例えば、時間ブロック内のかかるサブレベルは周波数帯域であり、エンコーダは時間ブロック内の各周波数帯域に対して、複素予測係数の値を決定する。同様に、第2の周波数領域表示を生成する方法は、複素予測係数がゼロであるか、大きさが許容値より小さい、時間ブロック内の周波数帯域に対する演算を抑制するように構成されている。

30

【 0 0 2 8 】

一実施形態において、前記第1のスペクトル成分は変換係数の時間ブロックに配置された変換係数であり、各ブロックは時間領域信号の時間セグメントへの変換の適用により生成される。さらに、前記ダウンミックス信号の第2の周波数領域表示を計算するモジュールは、

- ・前記第1のスペクトル成分から第1の中間成分を求め、
- ・インパルス応答の少なくとも一部により前記第1のスペクトル成分の結合を構成して第2の中間成分を求め、
- ・前記第2の中間成分から第2のスペクトル成分を求めるように構成されている。

40

この手順により、米国特許第6,980,933 B2号に、特にコラム8乃至28に、特に式41に詳細に記載されているように、第1の周波数領域表示から直接第2の周波数領域表示を計算することができる。当業者は気づくように、例えば、異なる変換が続く逆変換とは反対に、計算は、時間領域によっては実行されない。

【 0 0 2 9 】

本発明による複素予測ステレオ符号化の実施例の場合、計算の複雑さは、従来のL/RまたはM/Sステレオと比較してほんの少ししか増加しない（Q M F領域における複素予

50

測ステレオ符号化により生じる増加よりも大幅に少ない)ことが推測されている。第2のスペクトル成分の厳密な計算を含むこのタイプの実施形態では、QMFベースの実施形態により生じるより数パーセント長いだけの遅延が生じる(時間ブロックの長さは1024サンプルであると仮定し、QMF分析/合成フィルタバンクの961サンプルの遅延と比較した)。

【0030】

好適にも、少なくとも前出の実施形態の一部では、インパルス応答は、第1の周波数領域表示を求められる、より正確には、その周波数応答特性により求められる変換に適応される。

【0031】

実施形態によっては、ダウンミックス信号の第1の周波数領域表示は、1つ又はそれ以上の分析窓関数(又は、カットオフ関数、例えば矩形窓、正弦窓、カイザー・ベッセル窓など)に対して適用される変換により得られ、その一目的は、危険なノイズ音量を生じたり、スペクトルに好ましくない変化を与えたりすることなく、時間的セグメント化を実現することである。場合によっては、かかる窓関数は、部分的にオーバーラップしている。次に、好ましくは、変換の周波数応答特性は、前記の1つ又はそれ以上の分析窓関数の特性に依存する。

【0032】

周波数領域における第2の周波数領域表示の計算を特徴とする実施形態をさらに参照して、近似的な第2の周波数領域表示を用いることにより、計算負荷を減らすことができる。かかる近似は、計算の基礎とする情報に完全性を求めないことにより実現できる。例えば、米国特許第6,980,933B2号の教示によると、3つの時間ブロック、すなわち出力ブロックと同時のブロック、先行するブロック、及び後続のブロックからの第1の周波数領域データは、一ブロック中のダウンミックス信号の第2の周波数領域表示の厳密な計算に必要である。本発明による複素予測符号化を目的として、後続ブロック及び/又は先行するブロックからのデータを省略、またはゼロで置き換える(モジュールの動作が原因となる、すなわち遅延に貢献しない)ことにより、好適な近似を得られ、第2の周波数領域表示の計算が1つ又は2つの時間ブロックのみに基づくようにする。留意点として、入力データの省略は、例えば、もはや同じパワーを表さないという意味で、第2の周波数領域表示のリスキューリングを意味するが、上記の通り、エンコーダ側とデコーダ側の両方で等価な方法で計算されている限り、複素予測符号化の基礎として用いることができる。確かに、この種のリスキューリングは、予測係数値の対応する変化により、補償される。

【0033】

ダウンミックス信号の第2の周波数領域表示の一部を構成するスペクトル成分を計算するさらに他の近似方法は、第1の周波数領域表示からの少なくとも2つの成分の結合を含む。後者の成分は時間及び/又は周波数に関して隣接している。代替案として、後者の成分は、比較的少数のステップで、有限インパルス応答(FIR)フィルタリングにより結合できる。例えば、1024の時間ブロックサイズを用いるシステムでは、かかるFIRフィルタは2、3、4個等のタップを含む。この種の近似的計算方法の説明は、例えば、米国特許出願公開第2005/019783A1号に見いだすことができる。各時間ブロック境界の近傍に比較的小さい重みを与える窓関数を、例えば非矩形関数を用いた場合、時間ブロックの第2のスペクトル成分を同じ時間ブロックの第1のスペクトル成分の組み合わせのみに基づかせると都合がよいが、最も外側の成分については同量の情報は得られないことになる。かかるプラクティスにより生じる可能性のある近似誤差は、ある程度抑えることができ、または窓関数の形状により隠蔽することができる。

【0034】

時間領域ステレオ信号を出力するように設計されたデコーダの一実施形態では、直接または同時ステレオ符号化と複素予測符号化との間で切り換える可能性がある。これは次のものを備えることにより実現できる。すなわち、

- ・(信号を変化させない)パススルー段階として、または和・差変換として選択的に動

10

20

30

40

50

作可能なスイッチ；

- ・周波数・時間変換を行う逆変換段階；及び

- ・直接（または同時）符号化した信号を、または複素予測により符号化された信号を、逆変換段階に入力するセレクタ装置。

当業者が気づくように、デコーダの側にこのようなフレキシビリティがあるので、エンコーダは、従来の直接または同時符号化と、複素予測符号化とを選択する自由度を有する。よって、この実施形態は、従来の直接L/Rステレオ符号化や同時M/Sステレオ符号化の音質レベルを越えられない場合には、少なくとも、同じレベルを維持することを保証できる。よって、本実施形態によるデコーダは、関連技術に対して上位集合（superset）であるとみなすことができる。

10

【0035】

デコーダシステムの他の一群の実施形態は、時間領域を介して、第2の周波数領域表示の第2のスペクトル成分の計算を行う。より正確には、第1のスペクトル成分を求めた（又は求め得る）変換の逆変換を適用し、次に、出力として第2のスペクトル成分を有する異なる変換を行う。具体的に、逆MDCTの後にMDSTを行う。かかる実施形態では、変換と逆変換の数を減らすため、逆MDCTの出力を、MDSTと、復号システムの出力端子（場合によっては、さらにべつの処理ステップが前置されている）とに送る。

【0036】

本発明による複素予測ステレオ符号化の実施例の場合、計算の複雑さは、従来のL/RまたはM/Sステレオと比較してほんの少ししか増加しない（QMF領域における複素予測ステレオ符号化により生じる増加よりも大幅に少ない）ことが推測されている。

20

【0037】

前記パラグラフで言及した実施形態のさらなる発展として、アップミックス段階はサイド信号を処理するさらなる逆変換段階を有しても良い。そして、和・差段階に、前記さらなる逆変換段階により生成されたサイド信号の時間領域表示と、前述の逆変換により生成されたダウンミックス信号の時間領域表示とを供給する。再度述べるが、計算の複雑性の観点から、都合良く、後者の信号は、上述の和・差段階と異なる変換段階との両方に供給される。

【0038】

一実施形態では、時間領域ステレオ信号を出力するように設計されたデコーダは、直接L/Rステレオ符号化または同時M/Sステレオ符号化と、複素予測ステレオ符号化との間で切り換えが可能である。これは次のものを備えることにより実現できる。すなわち、

- ・パススルー段階として、または和・差段階として動作できるスイッチ；

- ・サイド信号の時間領域表示を計算するさらなる逆変換段階；

- ・逆変換段階を、（好ましくは、複素予測符号化により生成されたステレオ信号を復号する場合のように、スイッチがアクティブ化されパスフィルタとして機能するときに、）アップミックスの上流にあり、かつスイッチの下流にあるポイントに接続されたさらなる和・差段階に、または（好ましくは、直接符号化されたステレオ信号を復号する場合のように、スイッチがアクティブ化され、和・差段階として機能するときに、）スイッチからのダウンミックス信号と、重み付け加算器からのサイド信号との組み合わせに、接続するセレクタ装置。

30

当業者は気づくように、これはエンコーダに、従来の直接又は同時符号化と、複素予測符号化との間を選択する自由度を与え、すなわち直接又は同時ステレオ符号化と少なくとも等しい音質レベルを保証できる。

40

【0039】

一実施形態では、本発明の第2の態様によるエンコーダシステムは、残差信号の信号パワー又は平均信号パワーを低減又は最小化する目的で、複素予測係数を推定する推定器を有する。最小化はある時間にわたり、好ましくは符号化する時間セグメント又は時間ブロック又は時間フレームにわたり行われる。振幅の二乗を瞬間信号パワーの尺度とでき、振幅の二乗の一時間区間にわたる積分をその時間区間における平均信号パワーの尺度とでき

50

る。好適にも、複素予測係数は時間ブロックごとに、及び周波数帯域ごとに決定できる。すなわち、その値は、その時間ブロック及び周波数帯域における残差信号の平均パワー（すなわち、全エネルギー）を低減するように設定される。具体的に、IID、ICC及びIPD又は同様のパラメータなどのパラメトリックステレオ符号化パラメータを推定するモジュールは、当業者には知られた数学的関係により複素予測係数を計算できる出力を提供する。

【0040】

一実施形態では、エンコーダシステムの符号化段階は、直接ステレオ符号化を可能とするため、さらに、パススルー段階として機能する。直接ステレオ符号化がより高い音質を提供すると期待される状況では、これを選択することにより、エンコーダシステムは、符号化されたステレオ信号が少なくとも直接符号化と同じ音質を有することを保証できる。同様に、音質が大幅に向上しても複素予測符号化により生じる大きな計算負荷が望ましくない状況では、エンコーダシステムには、計算リソースを節約するオプションが容易に利用できる。コーダにおける同時、直接実予測符号化と、複素予測符号化との間の決定は、一般的に、レート/歪み最適化の原理に基づく。

【0041】

一実施形態では、エンコーダシステムは、第1のスペクトル成分に直接基づき（すなわち、時間領域に逆変換を適用せず、かつ信号の時間領域データを用いずに）第2の周波数領域表示を計算するモジュールを有する。上述のデコーダシステムの対応する実施形態に関して、このモジュールは、同様の構成を有する、すなわち、同様の、しかし異なる順序の処理動作を有し、エンコーダがデコーダ側の入力に適したデータを出力するように構成される。この実施形態を説明する目的で、符号化するステレオ信号は、ミッド及びサイドチャンネルを有し、又はこの構成に変換され、符号化段階は、第1の周波数領域表示を受け取るように構成されているものと仮定する。符号化段階は、ミッドチャンネルの第2の周波数領域表示を計算するモジュールを有する。（ここで参照する第1と第2の周波数領域表示は、上で定義した通りである；具体的に、第1の周波数領域表示はMDCT表示であってもよく、第2の周波数領域表示はMDST表示であってもよい。）符号化段階は、さらに、サイド信号と、ミッド信号の2つの周波数領域表示とから構成され、複素予測係数の実部と虚部により重み付けされた線形結合として、残差信号を計算する重み付け加算器を有する。ミッド信号は、または好適にもその第1の周波数領域表示は、ダウンミックス信号として直接用いられる。この実施形態では、さらに、残差信号のパワー又は平均信号パワーを最小化する目的で、推定器が複素予測係数の値を決定する。最終動作（最適化）は、フィードバック制御により、さらに必要であれば、推定器が、調整すべき現在の予測係数値により得られる残差信号を受け取るフィードバック制御により、またはフィードフォワード的に、元のステレオ信号の左/右チャンネルに又はミッド/サイドチャンネルで直接行った計算により、行われる。ミッド信号の第1と第2の周波数領域表示と、サイド信号の第1の周波数領域表示とに基づいて、複素予測係数が直接的に（特に、非反復的又は非フィードバック的に）計算されるフィードフォワード法が好ましい。留意点として、複素予測係数の決定後、各オプションで得られる品質（好ましくは、例えば信号対マスク効果を考慮した知覚的品質）を考慮して、直接、同時実予測符号化をするか、または複素予測符号化をするかの決定を行う。よって、上記のステートメントは、エンコーダにフィードバックメカニズムが存在しないという旨と解釈してはならない。

【0042】

一実施形態では、エンコーダシステムは、時間領域を介して、ミッド（すなわちダウンミックス）信号の第2の周波数領域表示を計算するモジュールを有する。この実施形態に関する実施の詳細事項は、少なくとも第2の周波数領域表示の計算に関する限り、同様であり、対応するデコーダの実施形態と同様に行うことができる。この実施形態では、符号化段階は、次のものを有する。すなわち：

- ・ステレオ信号をミッドチャンネルとサイドチャンネルに変換する和・差段階；
- ・サイドチャンネルの周波数領域表示と、ミッドチャンネルの複素値（すなわち、オー

10

20

30

40

50

バーサンプリングされた)周波数領域表示とを提供する変換段階;及び

・複素予測係数を重みとして用いる、残差信号を計算する重み付け加算器。

ここで、推定器は、残差信号を受け取り、場合によってはフィードバック制御形式で、残差信号のパワーまたは平均パワーを低減または最小化する複素予測係数を決定する。しかし、好ましくは、推定器は、符号化するステレオ信号を受け取り、それに基づいて予測係数を決定する。サイドチャンネルのクリティカルサンプリングされた周波数領域表示を用いることは、計算の経済性の観点から有利である。この実施形態では、サイドチャンネルは複素数との乗算をされないからである。好適にも、変換段階は、並列に構成されたMDC T段階とMDS T段階とを含み得る。両者は、ミッドチャンネルの時間領域表示を入力として有する。このように、ミッドチャンネルのオーバーサンプリングされた周波数領域表示と、サイドチャンネルのクリティカルサンプリングされた周波数領域表示とを生成する。

10

【0043】

留意点として、このセクションで開示した方法と装置は、通常の実験を含む当業者の能力の範囲内で適当な修正をして、2より多いチャンネルを有する信号の符号化に適用できる。かかるマルチチャンネルオペラビリティへの変更は、例えば、上で引用したJ. Herre等による論文のセクション4、5に即して行える。

【0044】

さらに別の実施形態では、上記の2つ以上の実施形態の特徴を、明らかに補完的でない限り、組み合わせられる。2つの特徴が異なるクレームに記載されていても、それらを組み合わせられないと言うわけではない。同様に、さらに別の実施形態では、所望の目的に対して必要でない、または本質的でない特徴を省略してもよい。一例として、本発明による復号システムは、処理する符号化信号が量子化されていない場合、又はアップミックス段階での処理に好適な形式にすでになっている場合、逆量子化段階無しに実施してもよい。

20

【図面の簡単な説明】

【0045】

添付した図面を参照して、次のセクションで説明する実施形態により、本発明をさらに説明する。

【図1A】背景技術によるQMFベースデコーダを示すブロック図である。

30

【図1B】背景技術によるQMFベースデコーダを示すブロック図である。

【図2】本発明の一実施形態による複素予測を有するMDC Tベースステレオデコーダシステムを示すブロック図である。復号される信号のチャンネルの複素表示は周波数領域で計算される。

【図3】本発明の一実施形態による複素予測を有するMDC Tベースステレオデコーダシステムを示すブロック図である。復号される信号のチャンネルの複素表示は時間領域で計算される。

【図4】図2のデコーダシステムの別の実施形態を示す図である。アクティブTNS段階の位置は選択可能である。

【図5】本発明の他の一態様の実施形態による、複素予測を有するMDC Tベースステレオエンコーダシステムを示すブロック図である。

40

【図6】本発明の一実施形態による複素予測を有するMDC Tベースステレオエンコーダシステムを示すブロック図である。符号化される信号のチャンネルの複素表示はその時間領域表現に基づき計算される。

【図7】図6に示したエンコーダシステムの別の実施形態を示す図である。このシステムは直接L/R符号化モードでも動作可能である。

【図8】本発明の一実施形態による複素予測を有するMDC Tベースステレオエンコーダシステムを示すブロック図である。符号化される信号のチャンネルの複素表示はその第1周波数領域表現に基づき計算される。このシステムは直接L/R符号化モードでも動作可能である。

50

【図 9】図 7 に示したエンコーダシステムの別の実施形態を示す図である。このシステムは、符号化段階の下流に配置された T N S 段階をさらに含む。

【図 10】図 2 と図 8 にラベル A で示した部分の別の実施形態を示す図である。

【図 11】図 8 に示したエンコーダシステムの別の実施形態を示す図である。このシステムは、符号化段階の上流と下流にそれぞれ配置された周波数領域修正デバイスをさらに含む。

【図 12】6 サブジェクトからの 96 kb/s におけるリスニングテスト結果を示すグラフであり、M D S T スペクトルの計算または近似のための異なる複雑性対音質トレードオフオプションを示している。ここで、ラベル「+」で示されたデータ点は隠れた基準を示す。「x」は 3.5 kHz 帯域制限アンカーを示す。「*」は U S A C による従来のステレオ (M/S または L/R) を示す。「 \square 」は、予測係数の虚部をディスエーブルした (すなわち、M D S T を必要としない実数値予測による) 複素予測による M D C T 領域ユニファイドステレオ符号化を示す。「 \triangle 」は、現在の M D C T フレームを用いて M D S T の近似値を計算する複素予測による M D C T 領域ユニファイドステレオ符号化を示す。「 \diamond 」は、現在と前の M D C T フレームを用いて M D S T の近似値を計算する複素予測による M D C T 領域ユニファイドステレオ符号化を示す。「 \circ 」は、現在と前と次の M D C T フレームを用いて M D S T を計算する複素予測による M D C T 領域ユニファイドステレオ符号化を示す。

【図 13】現在の M D C T フレームを用いて M D S T の近似値を計算する複素予測による M D C T 領域ユニファイドステレオ符号化に関する差分スコアとして図 12 のデータを示す図である。

【図 14 A】本発明の実施形態によるデコーダシステムの一実施形態を示すブロック図である。

【図 14 B】本発明の実施形態によるデコーダシステムの他の一実施形態を示すブロック図である。

【図 14 C】本発明の実施形態によるデコーダシステムのさらに他の一実施形態を示すブロック図である。

【図 15】本発明の一実施形態による復号方法を示すフローチャートである。

【図 16】本発明の一実施形態による符号化方法を示すフローチャートである。

【発明を実施するための形態】

【0046】

I. デコーダシステム

図 2 は、概略的なブロック図の形式で、少なくとも 1 つの複素予測係数値 $\hat{c}_i = \hat{c}_R + i \hat{c}_I$ を有するビットストリームを復号する復号システムを示す。ステレオ信号の M D C T 表現はダウンミックス M チャンネルと残差 D チャンネルを有する。予測係数の実部と \hat{c}_R と虚部 \hat{c}_I は量子化され、及び / または同時符号化 (coded jointly) されている。しかし、好ましくは、実部と虚部は独立かつ均一に、一般的にはステップサイズ 0.1 (無次元数) で、量子化される。M P E G 標準によると、複素予測係数に用いる周波数帯域の解像度は、スケールファクタ帯域 (sfb、すなわち同じ M D C T 量子化ステップサイズと量子化範囲を用いる一群の M D C T ライン) の解像度と同じである必要はない。特に、予測係数の周波数帯域解像度は、バークスケール (Bark scale) のように音響心理学的に妥当なものである。デマルチプレクサ 201 は、供給されるビットストリームから、これらの M D C T 表現と予測係数 (図示した制御情報の一部) を取り出すように構成されている。実際、ビットストリームには、それを予測モードと非予測モードのどちらで復号するかという命令や T N S 情報などの、複素予測係数以上の制御情報がエンコードされている。T N S 情報は、デコーダシステムの T N S (合成) フィルタにより使われる T N S パラメータの値を含む。両チャンネルなどの複数の T N S フィルタに同じ一組の T N S パラメータを用いる場合、パラメータの組のアイデンティティを示すビットの形式でこの情報を受け取る方が、二組のパラメータを別々に受け取るよりも、経済的である。例えば、2 つのオプションの音響心理学的評価に基づき、T N S をアップミックス段階の前または後に適

10

20

30

40

50

用するかの情報も含まれる。さらに、制御情報はダウンミックス信号と残差信号の個別に制限された帯域幅を示す。各チャンネルに対して、帯域幅制限より上の周波数帯域は復号されず、ゼロに設定される。場合によっては、最も高い周波数帯域のエネルギーコンテンツは小さいので、量子化されたときにすでにゼロになっている。通常のプラクティス（MPEG標準のmax_sfbパラメータを参照）では、ダウンミックス信号と残差信号の両方に、同じ帯域幅制限を用いなければならない。しかし、残差信号は、ダウンミックス信号よりも大幅に、低周波数帯域に局限されたエネルギーコンテンツを有する。そのため、残差信号に専用の帯域幅上限を課すことにより、音質を大幅に損なわずに、ビットレートの削減が可能である。例えば、これは、ビットストリームにエンコードされた、ダウンミックス信号用と残差信号用の2つの独立なmax_sfbパラメータにより調節される。

10

【0047】

この実施形態では、ステレオ信号のMDC T表現は、一定数のデータポイント（例えば、1024ポイント）、複数の一定数のデータポイントのうちの一つ（例えば、128ポイント又は1024ポイント）、または可変数のポイントを含む、連続した時間フレーム（すなわち時間ブロック）にセグメント化される。当業者には既知であるが、MDC Tはクリティカルにサンプリングされる。復号システムの出力は、図の右側部分に示したが、左Lチャンネルと右Rチャンネルを有する時間領域のステレオ信号である。逆量子化モジュール202は、復号システムに入力されたビットストリームを、必要に応じて、元のビットストリームを逆多重化後に得られるダウンミックスチャンネルと残差チャンネルのそれぞれに対応する2つのビットストリームを処理するように構成されている。逆量子化されたチャンネル信号は、変換行列

20

【数4】

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

に対応するパススルーモードで、又は変換行列

【数5】

$$\frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

30

に対応する和と差モードで動作できるスイッチングアセンブリ203に提供される。次のパラグラフでさらに説明するように、デコーダシステムは第2のスイッチングアセンブリ205を含む。両スイッチングアセンブリ203、205は、この実施形態及びこれから説明する実施形態におけるその他のほとんどのスイッチやスイッチングアセンブリのように、周波数選択的に動作可能である。これにより、例えば、関連技術として知られているように、周波数依存のL/RまたはM/S復号などの非常に様々な復号モードの復号が可能になる。よって、本発明によるデコーダは、関連技術に対して上位集合（superset）であるとみなすことができる。

40

【0048】

ここでスイッチングアセンブリ203がパススルーモードであると仮定して、この実施形態では、逆量子化されたチャンネル信号はそれぞれのTNSフィルタ204をパススルーされる。TNSフィルタ204は、復号システムの動作には本質的ではなく、パススルー要素により置き換えることもできる。この後、信号は、上流に配置されたスイッチングアセンブリ203と同じ機能を有する第2のスイッチングアセンブリ205に供給される。上記の通り入力信号を入力され、パススルーモードに設定されていると、第2のスイッチングアセンブリ205の出力はダウンミックスチャンネル信号と残差チャンネル信号である。ダウンミックス信号は、時間的に連続したMDC Tスペクトルで表されているが、ダウンミックス信号のMDS Tスペクトルを計算するように構成された実・虚変換206

50

に供給される。この実施形態では、1つのMDS Tフレームは3つのMDC Tフレームと、1つの前フレームと、1つの現在(すなわち同時)フレームと、1つの後フレームとに基づく。実・虚変換206の入力側が遅延コンポーネントを有することがシンボリックに (Z^{-1}, Z) 示されている。

【0049】

実・虚変換206から得られるダウンミックス信号のMDS T表示は、予測係数の虚部 I により重み付けされ、予測係数の実部 R と残差信号のMDC T表示により重み付けされたダウンミックス信号のMDC T表示に加えられる。2つの加算と乗算は、重み付け加算器210、211を(機能的に)構成する加算器及び乗算器により行われる。これらには、デコーダシステムにより最初に受け取られたビットストリームにエンコードされていた複素予測係数の値が供給される。複素予測係数は時間フレームごとに1つ決定される。複素予測係数は、もっと頻繁に決定してもよく、フレーム中の周波数帯域ごとに1つ決定してもよい。周波数帯域は音響心理学的に動機付けされたパーティションである。本発明の符号化システムに関して後で説明するように、複素予測係数は、それほど頻繁に決定しなくてもよい。実・虚変換206は、ダウンミックスチャンネル信号の現在のMDS Tフレームがダウンミックスチャンネル信号と残差チャンネル信号のそれぞれの同時MDC Tフレームと結合されるように、重み付け加算器と同期している。これら3つの信号の和はサイド信号 $S = Re\{M\} + D$ である。この式で、 M はダウンミックス信号のMDC T表示とMDS T表示を両方とも含み、すなわち $M = M_{MDC T} - i M_{MDS T}$ である。D = DMDC Tは実数値である。このように、ダウンミックスチャンネルとサイドチャンネルを有するステレオ信号が得られ、和差変換207は、このステレオ信号から

【数6】

$$\begin{cases} L = M + S \\ R = M - S \end{cases}$$

により左チャンネルと右チャンネルを回復する。これらの信号はMDC T領域で表される。復号システムの最後のステップでは、各チャンネルに逆MDC T209を適用して、左右ステレオ信号の時間領域表示を求める。

【0050】

実・虚変換206の可能な実装は、上記の通り、出願人の米国特許第6,980,933B2号に詳しく説明されている。上記文献に記載された式41により、変換は有限インパルス応答フィルタとして表せる。例えば、偶数のポイントに対して、

【数7】

$$S(2v) = \frac{2}{N} \sum_{p=0}^{N-1} [(-1)^{p+1} X_I(p) + X_{III}(p)] h_{I,III}(2v-p) + \frac{4}{N} \sum_{l=0}^{N-1} X_{II}(2l+1) h_{II}(2v-2l-1),$$

ここで、 $S(2v)$ は2番目のMDS Tデータポイントであり、 X_I 、 X_{II} 、 X_{III} は各フレームからのMDC Tデータであり、 N はフレーム長である。さらに、 h_I 、 h_{II} 、 h_{III} は、適用される窓関数に応じたインパルス応答であり、方形、シノソイド、カイザー・ベッセルなどの選択された窓関数ごとに、またフレーム長ごとに決定される。エネルギーコンテンツが比較的小さくて、MDS Tデータに対する貢献が比較的小さいインパルス応答を無視することにより、この計算の複雑さが低下する。この単純化の代わりに、または拡張として、インパルス応答自体を、例えば、全フレーム長 N から少ないポイントに短くしてもよい。一例として、インパルス応答長を1024ポイント(タップ)か

ら 10 ポイントに減らせる。意味があると考えられる最も極端な単純化 (truncation) は【数 8】

$$S(v) = X_{II}(v+1) - X_{II}(v-1)$$

である。その他の分かりやすいアプローチは米国特許出願公開第 2005/019783 1 A 1 号に載っている。

【0051】

計算の基礎にする入力データ量をさらに減らすことが可能である。説明のため、図では「A」で示した部分である、実・虚変換 206 とその上流の接続を、単純化した変形例により置き換えてもよい。そのうちの 2 つ、A と A を図 10 に示す。変形例 A は信号の虚数表示の近似を与える。ここで、MDS T 計算は現在フレーム及び前のフレームのみを考慮する。この段落の上記の式を参照して、 $p = 0, \dots, N-1$ に対して $X_{II}(p) = 0$ と設定することにより行う (インデックス II は後の時間フレームを示す)。変形例 A は後のフレームの MDC T スペクトルを入力として必要としないので、MDS T 計算は時間遅延を生じない。明らかに、この近似により、得られる MDS T 信号の正確性はいくぶん低下するが、この信号のエネルギーも減少することを示唆する。予測符号化の性質として、後者は I を大きくすることにより完全に補償できる。

【0052】

変形例 A を図 10 に示した。これは、現在の時間フレームの MDC T データのみを入力として用いる。変形例 A により得られる MDS T 表示は、変形例 A により得られるものより正確性が劣る。他方、変形例 A は変形例 A のようにゼロ遅延で動作し、計算の複雑性が低い。前述の通り、エンコーダシステムとデコーダシステムで同じ近似を使っている限りでは、波形符号化特性には影響はない。

【0053】

留意点として、変形例 A、A または A 、またはこれらをさらに発展させたもののどれを使うかにかかわらず、MDS T スペクトルの複素予測係数の虚部がゼロでない、すなわち $I = 0$ である部分のみを計算すればよい。実際的な状況では、これは、係数の虚部の絶対値 $|I|$ が所定閾値より大きいことを意味すると解することができる。この所定閾値は用いるハードウェアの単位の丸め (unit round-off) に関する。時間フレーム中のすべての周波数帯域の係数の虚部がゼロである場合、そのフレームについて MDS T データを計算する必要はない。よって、やはり、実・虚変換 206 は、MDS T 出力を生成しないことにより、 $|I|$ の値が非常に小さい場合に応答するように構成されている。これにより計算資源を節約できる。しかし、現在フレーム以上のフレームを使って MDS T データの 1 フレームを生成する実施形態では、非ゼロ予測係数に関連する次の時間フレームが生じた時に、実・虚変換 206 に対して十分な入力データがあるように、変換 206 の上流のユニットは、MDS T スペクトルが必要なくても動作し続けなければならない。特に、第 2 スイッチングアセンブリ 205 は、MDC T スペクトルを転送し続けなければならない。これはもちろん次の時間ブロックである。

【0054】

図 2 に戻り、スイッチングアセンブリ 203、205 が両方ともそれぞれパススルーモードに設定されていると仮定して、復号システムの機能を説明した。ここで説明するように、デコーダシステムは予測符号化されたものではない信号も復号できる。この利用のために、第 2 のスイッチングアセンブリ 205 は、和・差モード (sum-and-difference mode) に設定され、図に示したように、セクタ装置 208 は下ポジションに設定され、信号が、TNS フィルタ 204 と第 2 のスイッチングアセンブリ 205 の間のソースポイントから逆変換 209 に直接入力されるようになっている。正しい復号をするため、信号は適切にソースポイントにおいて L/R 形式を有する。それゆえ、実・虚変換に (例えば、左信号により簡潔にではなく) 常に正しいミッド (すなわち、ダウンミックス) 信号を供給するため、非予測符号化ステレオ信号の復号時には、第 2 のスイッチングアセンブリ

リ205を和・差モードに設定することが好ましい。上記の通り、予測符号化は、例えばデータレート対音質決定などに基づいて、従来の直接符号化または複数フレームの同時符号化により置き換えられる。かかる決定の結果は、いろいろな方法で、例えば各フレーム中の専用インジケータビットの値により、または予測係数値の存否により、エンコーダからデコーダに送られる。これら事実を立証すれば、第1のスイッチングアセンブリ203の役割は容易に実現できる。事実、非予測符号化モードでは、デコーダシステムは、直接(L/R)ステレオ符号化による信号と、同時(M/S)符号化による信号を両方とも処理できる。第1のスイッチングアセンブリ203をパススルーモードまたは和・差モードのいずれかで動作させることにより、直接符号化された信号とともに常にソースポイントが提供されるようにすることが可能である。明らかに、スイッチングアセンブリ203は、和・差段階で機能するとき、M/S形式の入力信号をL/R形式の出力信号に変換する(任意的なTNSフィルタ204に供給される)。

10

【0055】

デコーダシステムは、そのデコーダシステムによりある時間フレームを予測符号化モードで復号するか、非予測符号化モードで復号するかを示す信号を受け取る。非予測モードは、各フレーム中の専用インジケータビットの値により、または予測係数の存否(または値がゼロ)により、シグナリングされる。予測モードは同様にシグナリングすることができる。特に有利な実施形態は、オーバーヘッド無し of フォールバックを可能とするが、2ビットフィールドms_mask_present(MPEG-2 AAC、ISO/IEC 13818-7文書参照)の予約された第4の値を利用する。これは、時間フレームごとに送信され、次のように規定されており、次のように規定されている：

20

【表1】

表1:USACにおけるms_mask_presentの定義

値	意味
00	すべての周波数帯域のL/R符号化
01	L/RまたはM/Sを示すため1帯域あたり1シグナリングビットを使う
10	すべての周波数帯域のM/S符号化
11	予約

30

値11を「複素予測符号化」を意味すると再定義することにより、デコーダは、ビットレートを損なわずに全レガシーモードで、特にM/S及びL/R符号化モードで、動作でき、関連するフレームの複素予測符号化モードを示す信号を受け取ることができる。

【0056】

図4は、一般的構成のデコーダシステムを示し、図2に示したものと同様であるが、少なくとも2つの異なる構成を含む。最初に、図4のシステムは、アップミックス段階の上流及び/又は下流に、周波数領域修正を含む処理ステップの適用を可能にするスイッチ404、411を含む。これは、一方で、逆量子化モジュール401と第1のスイッチングアセンブリ402の下流にあり、かつアップミックス段階406、407、408、409のすぐ上流に配置された第2のスイッチングアセンブリ405の上流にある、第1のスイッチ404とともに設けられた第1組の周波数領域モディファイア403(この図ではTNS合成フィルタとして描いた)により実現される。他方、デコーダシステムは、アップミックス段階406、407、408、409の下流にあり、逆変換段階412の上流にある、第2のスイッチ411とともに設けられた第2の組の周波数領域モディファイア410を含む。有利にも、図に示したように、各周波数領域モディファイアは、上流では周波数領域モディファイアの入力側に接続され、下流では関連スイッチに接続されたパススルーラインと並行して配置されている。この構成により、周波数領域モディファイアに

40

50

は常に信号データが供給され、現在の時間フレームだけでなくより多くの時間フレームに基づいた周波数領域における処理が可能になる。第1の組の周波数領域モディファイア403または第2の組の周波数領域モディファイア410のどちらを適用するかは決定は、エンコーダによりなされ(ビットストリームで送られ)、又は予測符号化が適用されるかに基づき、又は実際の状況に適しているその他の基準に基づいてもよい。一例として、周波数領域モディファイアがTNSフィルタである場合、第1の組403はある種の信号に対する利用に有利であり、一方第2の組410は他の種類の信号に対する利用に有利である。この選択の結果がビットストリームにエンコードされている場合、デコーダシステムはTNSフィルタの各組を適宜アクティブ化する。

【0057】

10

図4に示したデコーダシステムの理解を容易にするため、明示的に留意しておくが、直接(L/R)符号化信号の復号は $=0$ (擬似L/RとL/Rは同じであり、サイドチャンネルと残差チャンネルが違わないことを示唆する)であり、第1のスイッチングアセンブリ402がパスモードであり、第2のスイッチングアセンブリが和・差モードであり、アップミックス段階の第2スイッチングアセンブリ405と和・差段階409の間で信号がM/S形式である時に行われる。この時、アップミックス段階は有効にパスするステップであるから、(各スイッチ404、411を用いて)第1の組の周波数領域モディファイアまたは第2の組の周波数領域モディファイアがアクティブ化されているかは重要ではない。

【0058】

20

図3は、図2と図4のデコーダシステムに関連して、アップミキシングに必要なMDS Tデータの供給への異なるアプローチを表す、本発明の一実施形態によるデコーダシステムを示す。すでに説明したデコーダシステムと同様に、図3のシステムは、逆量子化モジュール301、パススルーモード又は和・差モードで動作可能な第1のスイッチングアセンブリ302、及びTNS(合成)フィルタ303を有する。これらはすべてデコーダシステムの入力端から直列に配置されている。このポイントの下流にあるモジュールは、2つの第2のスイッチ305、310により選択的に利用される。これらの第2のスイッチは、図示したように、両方が上ポジションか下ポジションになるよう、同時に動作することが好ましい。デコーダシステムの出力端には、和・差段階312があり、そのすぐ上流には、各チャンネルのMDC T領域表示を時間領域表示に変換する2つの逆MDC Tモジュール306、311がある。

30

【0059】

複素予測復号では、デコーダシステムにダウンミックス/残差ステレオ信号と複素予測係数をエンコードしたビットストリームが供給され、第1のスイッチングアセンブリ302はパススルーモードに設定され、第2のスイッチ305、310は上ポジションに設定される。TNSフィルタの下流では、(逆量子化され、TNSフィルタされたMDC T)ステレオ信号の2つのチャンネルには違う処理がなされる。ダウンミックスチャンネルは、一方では、乗算器及び加算器308に供給される。乗算器及び加算器308は、予測係数の実部 R で重み付けされたダウンミックスチャンネルのMDC T表示を、残差チャンネルのMDC T表示に加算する。他方では、複数のMDC T変換モジュールの1つ306に供給される。ダウンミックスチャンネルMの時間領域表示は、逆MDC T変換モジュール306からの出力であり、最終和・差段階312とMDS T変換モジュール307の両方に供給される。このようにダウンミックスチャンネルの時間領域表示を二重で使うことは、計算の複雑性の観点から有利である。このように得られたダウンミックスチャンネルのMDS T表示は、さらに別の乗算器及び加算器309に供給される。この乗算器及び加算器309は、予測係数の虚部 I により重み付けしてから、この信号を加算器308からの線形結合出力に加える。よって、加算器309の出力はサイドチャンネル信号 $S = R e \{ M \} + D$ である。同様に、乗算器及び加算器308、309は、図2に示したデコーダシステムに結合され、ダウンミックス信号のMDC T表示とMDS T表示、残差信号のMDC T表示、及び複素予測係数値を入力とする重み付けマルチ信号加算器を形成する

40

50

。本実施形態において、このポイントの下流では、サイドチャンネル信号が最終和・差段階 3 1 2 に供給される前に、逆 M D C T 変換モジュール 3 1 1 を通る経路のみが残る。

【 0 0 6 0 】

デコーダシステムにおいて必要な同期性は、両方の逆 M D C T 変換モジュール 3 0 6、3 1 1 において適用する変換長と窓形状を同じにすることにより、実現できる。これは、周波数選択的 M / S および L / R 符号化ですでに実用されている。逆 M D C T モジュール 3 0 6 のある実施形態と、M D S T モジュール 3 0 7 のある実施形態を組み合わせると、1 フレームの遅延が生じる。そのため、5 つの任意的な遅延ブロック 3 1 3 (またはコンピュータ実施の場合にこの効果を発揮するソフトウェア命令) が設けられ、システムのうち破線の右側にある部分を、必要に応じて、左側にある部分に対して 1 フレーム遅延できる。明らかに、破線と接続ラインの間のすべての交点には遅延ブロックが設けられているが、逆 M D C T モジュール 3 0 6 と M D S T 変換モジュール 3 0 7 の間の接続は例外であり、ここでは補償を要する遅延が生じる。

10

【 0 0 6 1 】

1 つの時間フレームの M D S T データの計算には、時間領域表示の 1 フレームからのデータが必要である。しかし、逆 M D C T 変換には、1 つのフレーム (現在フレーム)、2 つの連続したフレーム (好ましくは、前のフレームと現在フレーム)、又は 3 つの連続したフレーム (好ましくは、前のフレーム、現在フレーム、及び後のフレーム) に基づく。M D C T に関連する周知の時間領域エイリアスキャンセレーション (T D A C) のため、3 フレームオプションは入力フレームの完全なオーバーラップを実現し、少なくとも時間領域エイリアスを含むフレームでは、最も (場合によっては完全に) 正確である。明らかに、3 フレーム逆 M D C T は 1 フレーム遅れで動作する。M D S T 変換への入力として近似的な時間領域表示の利用を許容することにより、この遅延を回避して、それにより、デコーダシステムの異なる部分間の遅延を補償する必要性を回避できる。2 フレームオプションでは、フレームの前半でオーバーラップ / アッド・イネープリング T D A C が行われ、エイリアスは後半にのみ存在する。1 フレームオプションでは、T D A C が無いので、エイリアスはフレーム全体で生じる。しかし、このように実現され複素予測符号化で昼間信号として用いられる M D S T 表示は、十分なクオリティを提供できる。

20

【 0 0 6 2 】

図 3 に示した復号システムは、2 つの非予測復号モードでも動作できる。直接 L / R 符号化ステレオ信号を復号するため、第 2 のスイッチ 3 0 5、3 1 0 は下ポジションに設定され、第 1 のスイッチングアセンブリ 3 0 2 はパススルーモードに設定される。このように、この信号は、和・差段階 3 0 4 の上流では、L / R 形式である。和・差段階 3 0 4 はこれを M / S 形式に変換する。この M / S 形式に逆 M D C T 変換と最終和・差演算が行われる。同時 M / S 符号化形式で提供されたステレオ信号を復号するため、第 1 のスイッチングアセンブリ 3 0 2 は和・差モードに設定され、第 1 のスイッチングアセンブリ 3 0 2 と和・差段階 3 0 4 の間で信号が L / R 形式となるようにする。L / R 形式は、T N S フィルタリングの観点から、M / S 形式よりも適している。和・差段階 3 0 4 の下流における処理は、直接 L / R 復号の場合と同じである。

30

【 0 0 6 3 】

図 1 4 (1 4 A ないし 1 4 C) は、本発明の実施形態によるデコーダを示す 3 つのブロック図である。本願に添付した他のブロック図とは異なり、図 1 4 の接続線はマルチチャンネル信号を示す。具体的に、かかる接続線は、左 / 右、ミッド / サイド、ダウンミックス / 残差、擬似左 / 擬似右のチャンネルその他の組み合わせを有するステレオ信号を送信するように構成されている。

40

【 0 0 6 4 】

図 1 4 A は、入力信号の周波数領域表示 (この図の目的において、M D C T 表示として示した) を復号するデコーダシステムを示す。デコーダシステムは、その出力として、ステレオ信号の時間領域表示を供給するように構成される。この表示は入力信号に基づき生成される。複素予測ステレオ符号化により符号化された入力信号を復号できるようにする

50

ため、デコーダシステムにはアップミックス段階 1 4 1 0 が設けられている。しかし、他のフォーマットで符号化され、場合によっては時間の経過と共に複数の符号化フォーマット間で切り替わる入力信号を、例えば複素予測符号化により符号化された時間フレームのシーケンスに直接左 / 右符号化により符号化された時間部分が続く入力信号を、処理することも可能である。異なる符号化フォーマットを処理するデコーダシステムの機能は、前記アップミックス段階 1 4 1 0 と並列に接続ライン（パススルー）を設けることにより実現される。スイッチ 1 4 1 1 により、アップミックス段階 1 4 1 0 からの出力（図の下スイッチポジション）と、接続ラインにより得られる処理されていない信号（図の上スイッチポジション）とのどちらを、さらに下流に配置されたデコーダモジュールに供給するか選択できる。この実施形態では、逆 M D C T モジュール 1 4 1 2 はスイッチの下流に配置されている。M D C T モジュール 1 4 1 2 は、信号の M D C T 表示を時間領域表示に変換する。一例として、アップミックス段階 1 4 1 0 に供給される信号は、ダウンミックス / 残差形式のステレオ信号であってもよい。次に、サイド信号を求め、（M D C T 領域で）左 / 右ステレオ信号を出力するように和・差演算を行うため、アップミックス段階 1 4 1 0 が適用される。

【 0 0 6 5 】

図 1 4 B は、図 1 4 A に示したものと同様のデコーダシステムを示す。本システムは入力信号としてビットストリームを受け取るように構成されている。ビットストリームは、最初、結合されたデマルチプレクサ及び逆量子化モジュール 1 4 2 0 により処理される。この結合されたデマルチプレクサ及び逆量子化モジュール 1 4 2 0 は、図 1 4 A に示したスイッチ 1 4 1 1 と同様の機能を果たすスイッチ 1 4 2 2 のポジションにより決まるように、第 1 の出力信号として、さらなる処理のため、マルチチャンネルステレオ信号の M D C T 表示を提供する。より正確には、スイッチ 1 4 2 2 は、デマルチプレクサ及び逆量子化からの第 1 の出力を、アップミックス段階 1 4 2 1 と逆 M D C T モジュール 1 4 2 3 により処理するか（下ポジション）、逆 M D C T モジュール 1 4 2 3 のみにより処理するか（上ポジション）決定する。結合されたデマルチプレクサ及び逆量子化モジュール 1 4 2 0 は制御情報も出力する。このケースでは、ステレオ信号に関連する制御情報は、スイッチ 1 4 2 2 の上ポジション又は下ポジションが信号の復号に適しているか、より抽象的には、そのステレオ信号をどの符号化フォーマットに復号するかを示すデータを含む。制御情報は、例えば、すでに説明したように、複素予測符号化で用いる複素予測係数の値などの、アップミックス段階の特性を調節するパラメータも含む。

【 0 0 6 6 】

図 1 4 C は、図 1 4 B に示したものと同様のエンティティに加えて、アップミックス段階 1 4 3 3 の上流と下流にそれぞれ配置された第 1 と第 2 の周波数領域修正デバイス 1 4 3 1、1 4 3 5 を有する。この図面の目的において、各周波数領域修正デバイスは T N S フィルタにより例示されている。しかし、周波数領域修正デバイスとの用語は、T N S フィルタリング以外の、アップミックス段階の前後で適用できるプロセスであると理解することもできる。周波数領域修正の例には、予測、ノイズ付加、帯域幅拡張、非線形処理が含まれる。場合によっては、処理する信号の特性及び / 又はかかる周波数領域修正デバイスの設定を含む、音響心理学的考察及び類似の理由から、前記周波数領域修正を、アップミックス段階 1 4 3 3 の下流ではなく、その上流で適用する方が有利である。他の場合には、同様の考察から、周波数領域修正の下流での位置は、上流の方が好ましい。スイッチ 1 4 3 2、1 4 3 6 により、周波数領域修正デバイス 1 4 3 1、1 4 3 5 は、制御情報に応じて、デコーダシステムが所望の構成を選択できるように、選択的にアクティブ化される。一例として、図 1 4 C は、結合されたデマルチプレクサ及び逆量子化モジュール 1 4 3 0 からのステレオ信号が、第 2 の周波数領域調整デバイス 1 4 3 5 を通らずに、最初に第 1 の周波数領域修正デバイス 1 4 3 1 により処理され、次にアップミックス段階 1 4 3 3 に供給され、最後に逆 M D C T モジュール 1 4 3 7 に直接転送される構成を示す。発明の概要欄に説明したように、この構成は、複素予測符号化におけるアップミックス後に T N S を行うオプションより好ましい。

10

20

30

40

50

【 0 0 6 7 】

I I . エンコーダシステム

本発明によるエンコーダシステムを、図5を参照して説明する。図5は、複素予測符号化により、出力ビットストリームとして、左/右(L/R)ステレオ信号を符号化するエンコーダシステムを示すブロック図である。このエンコーダシステムは、信号の時間領域または周波数領域の表示を受け取り、これをダウンミックス段階と予測係数推定器の両方に供給する。予測係数の実部と虚部は、左右チャンネルのダウンミックス及び残差チャンネルへの変換を制御するために、ダウンミックス段階に供給される。次に、ダウンミックス及び残差チャンネルは、最終的マルチプレクサMUXに供給される。信号は、周波数領域表示としてエンコーダに供給されなかった場合、ダウンミックス段階またはマルチプレクサで、かかる表示に変換される。

10

【 0 0 6 8 】

予測符号化の原理の1つは、左/右信号をミッド/サイド形式に変換することであり、すなわち

【 数 9 】

$$\begin{cases} M = (L + R)/2 \\ S = (L - R)/2 \end{cases}$$

次にこれらのチャンネル間に残っている相関を用いる、すなわち

20

【 数 1 0 】

$$S = \text{Re}\{\alpha M\} + D$$

と設定する。ここで、 α は決定する複素予測係数であり、Dは残差信号である。残差信号のエネルギー $D = S - \text{Re}\{\alpha M\}$ を最小化するために α を選択できる。エネルギーの最小化は、瞬間パワー、短期的エネルギー、又は長期的エネルギー(パワー平均)によりもたらされる。これは、離散信号の場合には、平均二乗の意味で最適化される。

【 0 0 6 9 】

予測係数の実部 α_R と虚部 α_I は量子化され、及び/または同時符号化(coded jointly)される。しかし、好ましくは、実部と虚部は独立かつ均一に、一般的にはステップサイズ0.1(無次元数)で、量子化される。MPEG標準によると、複素予測係数に用いる周波数帯域の解像度は、スケールファクタ帯域(sfb、すなわち同じMDCT量子化ステップサイズと量子化範囲を用いる一群のMDCTライン)の解像度と同じである必要はない。特に、予測係数の周波数帯域解像度は、バークスケール(Bark scale)のように音響心理学的に妥当なものである。留意点として、変換長が変わると、周波数帯域の解像度が変化する。

30

【 0 0 7 0 】

前述の通り、本発明によるエンコーダシステムは、予測ステレオ符号化を適用するか否かの自由度を有する。後者の場合は、L/R又はM/S符号化へのフォールバックを示唆する。かかる決定は、時間フレームまたはそれより細かいベースで、または時間フレーム内の周波数帯域ベースで行える。上記の通り、その決定の否定的な結果は、いろいろな方法で、例えば各フレーム中の専用インジケータビットの値により、または予測係数値の存否(またはゼロ値)により、復号エンティティに送られる。肯定的決定も同様に送られる。特に有利な実施形態は、オーバーヘッド無しフォールバックを可能とするが、2ビットフィールドms_mask_present(MPEG-2 AAC、ISO/IEC 131818-7文書参照)の予約された第4の値を利用する。これは、時間フレームごとに送信され、次のように規定されており、次のように規定されている：

40

【表 2】

表 1:USAC における ms_mask_present の定義

値	意味
00	すべての周波数帯域の L/R 符号化
01	L/R または M/S を示すため 1 帯域あたり 1 シグナリングビットを使う
10	すべての周波数帯域の M/S 符号化
11	予約

10

値 11 を「複素予測符号化」を意味すると再定義することにより、エンコーダは、ビットレートを損なわずに全レガシーモードで、特に M/S 及び L/R 符号化モードで、動作でき、有利であれば、フレームの信号複素予測符号化を示す信号を受け取ることができる。

【0071】

実質的な決定は、データレート対音質原理に基づいてもよい。(入手可能な MDCT ベースのオーディオエンコーダの場合にはよくあることだが、)音質の尺度として、エンコーダに含まれる音響心理学的モデルを用いて得られたデータを使っても良い。具体的に、エンコーダの実施形態には、予測係数のレート歪み最適化選択をするものもある。したがって、かかる実施形態では、予測ゲインの増加により残差信号の符号化のために十分なビットを節約せず、予測係数の符号化に必要なビットの使用を正当化できない場合、予測係数の虚部は、及び場合によっては実部も、ゼロに設定される。

20

【0072】

エンコーダの実施形態は、TNS 関連の情報をビットストリームにエンコードする。かかる情報は、デコーダサイドで TNS (合成) フィルタにより使われる TNS パラメータの値を含む。両方のチャンネルで同じ TNS パラメータの組を使う場合、2つの組のパラメータを別々に送信するよりも、パラメータが同じ事を示すシグナリングビットを含めると経済的である。例えば、2つのオプションの音響心理学的評価に基づき、TNS をアップミックス段階の前または後に適用するかの情報も含まれる。

30

【0073】

さらに他の任意的な一特徴として、これは複雑性とビットレートの観点から潜在的に有益なものであるが、エンコーダは残差信号の符号化のために、個別に制限された帯域幅を使うように構成される。この限界より上の周波数帯域はデコーダに送信されず、ゼロに設定される。場合によっては、最も高い周波数帯域のエネルギーコンテンツは小さいので、量子化されたときにすでにゼロになっている。通常のプラクティス(MPEG標準のmax_sfbパラメータを参照)では、ダウンミックス信号と残差信号の両方に、同じ帯域幅制限の使用を必要とする。ここで、発明者は経験的に、残差信号が、ダウンミックス信号よりも大幅に、低周波数帯域に局限されたエネルギーコンテンツを有することを見いだした。そのため、残差信号に専用の帯域幅上限を課すことにより、音質を大幅に損なわずに、ビットレートの削減が可能である。例えば、これは、ダウンミックス信号用と残差信号用の2つの独立なmax_sfbパラメータにより送信することにより実現される。

40

【0074】

指摘しておくが、図5に示したデコーダシステムを参照して予測係数、量子化とその符号化、M/S又はL/Rモードへのフォールバック、TNSフィルタリング、及び帯域幅上限などの最適な決定の問題を説明したが、同じ事が、後続の図面を参照して説明する実施形態において開示する実施形態にも等しく適用可能である。

【0075】

図6は、複素予測ステレオ符号化を行うように構成された、本発明による他のエンコーダシステムを示す。このシステムは、連続した、場合によってはオーバーラップした時間

50

フレームに分割され、左右チャンネルを含むステレオ信号の時間領域表示を入力として受け取る。和・差段階 601 は、この信号をミッドチャンネルとサイドチャンネルに変換する。ミッドチャンネルは MDC T モジュール 602 と MDS T モジュール 603 の両方に供給され、サイドチャンネルは MDC T モジュール 604 のみに供給される。予測係数水滴 605 は、上記の通り、各時間フレームに対して、及び場合によってはフレーム内の個々の周波数帯域に対して、複素予測係数の値を推定する。係数の値は、重み付け加算器 606、607 に重みとして供給される。重み付け加算器 606、607 は、ミッド信号の MDC T 及び MDS T 表示と、サイド信号の MDC T 表示の線形結合として、残差信号 D を構成する。好ましくは、複素予測係数は、それがビットストリームにエンコードされる時に用いられる同じ量子化スキームにより表された重み付け加算器 606、607 に供給される。これは、エンコーダとデコーダが両方とも同じ予測係数の値を用いるので、明らかに、より忠実な再構成を提供する。残差信号、ミッド信号（残差信号と組み合わせて現れる時にはダウンミックス信号と呼んだ方がより適切である）、及び予測係数は結合された量子化及びマルチプレクサ段階 608 に供給される。結合された量子化及びマルチプレクサ段階 608 は、これらの信号及び場合によってはさらに別の情報を出力ビットストリームとしてエンコードする。

【0076】

図 7 は、図 6 に示したエンコーダの変形例である。図の記号が同様なことから分かるように、図 7 に示したエンコーダの構成は同様であるが、直接 L/R 符号化フォールバックモードで動作するという機能が付加されている。エンコーダシステムは、結合された量子化及びマルチプレクサ段階 709 のすぐ上流に設けられたスイッチ 710 により、複素予測符号化モードとフォールバックモードとの間でアクティブ化される。スイッチ 710 が上ポジションにあると、エンコーダはフォールバックモードで動作する。ミッド・サイド信号は、MDC T モジュール 702、704 のすぐ下流のポイントから、和・差段階 705 に供給される。和・差段階 705 は、その信号を左/右信号に変換した後、スイッチ 710 に送る。スイッチ 710 は、その信号を結合された量子化及びマルチプレクサ段階 709 に接続する。

【0077】

図 8 は、本発明によるエンコーダシステムを示す図である。図 6 と図 7 に示したエンコーダシステムと異なり、この実施形態は、複素予測符号化に必要な MDS T データを MDC T データから直接的に、すなわち周波数領域における実・虚変換により求める。実・虚変換は、図 2 と図 4 のデコーダシステムに関して説明したいずれかのアプローチを適用する。忠実な復号を行えるように、デコーダの計算方法をエンコーダの計算方法と一致させることが重要である。エンコーダ側とデコーダ側とで同じ実・虚変換方法を用いる。デコーダの実施形態について、破線で囲んだ、実・虚変換 804 を有する部分 A は、これに近い変形例や、使う入力時間フレームを少なくすることにより置き換えできる。同様に、上記のいずれかの近似アプローチを用いて符号化を単純化できる。

【0078】

高レベルでは、図 8 のエンコーダシステムは、（適切に接続された）実・虚モジュールにより、図 7 の MDS T モジュールを単に置き換えることにより得られるだろう構成とは異なる構成を有する。このアーキテクチャはきれいであり、ロバストかつ計算量的に経済的に、予測符号化と直接 L/R 符号化との間の切り替え機能を実現できる。入力ステレオ信号は MDC T 変換モジュール 801 に入力され、MDC T 変換モジュール 801 は各チャンネルの周波数領域表示を出力する。これは、エンコーダシステムを予測符号化モードと直接符号化モードの間で起動するファイナルスイッチ 808 と、和・差段階 802 との両方に送られる。直接 L/R 符号化、または予測係数がゼロに設定された時間フレームで行われる同時 M/S 符号化において、この実施形態は、入力信号を MDC T 変換、量子化、及び多重化のみする。後から 2 つのステップは、システムの出力端に配置された結合された量子化及びマルチプレクサ段階 807 により行われ、ビットストリームが供給される。予測符号化では、各チャンネルは、和・差段階 802 とスイッチ 808 との間でさら

10

20

30

40

50

に処理される。実・虚変換 804 は、ミッド信号の MDC T 表示から MDST データを求め、それを予測係数推定器 803 と重み付け加算器 806 の両方に送る。図 6 と図 7 に示したエンコーダシステムと同様に、別の重み付け加算器 805 を用いて、サイド信号をミッド信号の重み付け MDC T と MDST 表示と結合し、残差チャンネル信号を構成する。残差チャンネル信号は、結合された量子化及びマルチプレクサ段階 807 により、ミッド（すなわち、ダウンミックス）チャンネル信号と予測係数とともにエンコードされる。

【0079】

ここで図 9 を参照して、エンコーダシステムの各実施形態は 1 つまたはそれ以上の TNS（分析）フィルタと結合できることを説明する。前述の通り、TNS フィルタリングをダウンミックス形式の信号に適用することは有利であることが多い。よって、図 9 に示したように、図 7 のエンコーダシステムを TNS を含むように適応させることは、結合された量子化及びマルチプレクサ段階 909 のすぐ上流に TNS フィルタ 911 を加えることにより行われる。

【0080】

右ノ残差 TNS フィルタ 911 b の替わりに、右チャンネルまたは残差チャンネルを処理するように構成された 2 つの TNS フィルタ（図示せず）をスイッチ 910 の部分のすぐ上流に設けてもよい。このように、2 つの TNS フィルタのそれぞれには、常に各チャンネル信号が供給され、現在フレームだけより多い時間フレームに基づく TNS フィルタリングが可能である。前述の通り、TNS フィルタは、周波数領域修正デバイスの一例であり、特に現在時間フレームより多いフレームの処理に基づくデバイスである。これは、TNS フィルタと同じくらいまたはそれ以上に、かかる配置から利益を得る。

【0081】

図 9 に示した実施形態の他の一代替として、選択的アクティブ化のための TNS フィルタを、各チャンネルに対して 1 つ以上のポイントで構成できる。これは、スイッチにより異なる組の TNS フィルタを接続できる、図 4 に示したデコーダシステムの構成と同様である。これにより、各時間フレームに対して、TNS フィルタリングのために最も適した段階を選択できる。特に、複素予測ステレオ符号化モードとその他の符号化モードとの間の切り替えに関して、異なる TNS ロケーションの間でスイッチすることは、有利である。

【0082】

図 11 は、図 8 のエンコーダシステムに基づく変形例であって、ダウンミックス信号の第 2 の周波数領域表示を実・虚変換 1105 により求めるものを示す。図 4 に示したデコーダシステムと同様に、このエンコーダシステムも、選択的にアクティブ化可能な周波数領域モディファイアモジュールを含み、そのうちの 1 つ 1102 はダウンミックス段階の上流に設けられ、1 つ 1109 はその下流に設けられている。周波数領域モジュール 1102、1109 は、この図では TNS フィルタにより例示したが、4 つのスイッチ 1103 a、1103 b、1109 a 及び 1109 b を用いて各信号経路に接続できる。

【0083】

III. 非装置実施形態

本発明の第 3 と第 4 の態様の実施形態を図 15 と 16 に示す。図 15 は、ビットストリームをステレオ信号に復号する方法を示し、次のステップを有する：

1. ビットストリームを入力する。
2. ビットストリームを逆量子化し、それによりステレオ信号のダウンミックスチャンネルと残差チャンネルの第 1 の周波数領域表示を求める。
3. ダウンミックスチャンネルの第 2 の周波数領域表示を計算する。
4. チャンネルの 3 つの周波数領域表示に基づき、サイドチャンネル信号を計算する。
5. ステレオ信号を、好ましくは左ノ右形式のものを、サイドチャンネルとダウンミックスチャンネルに基づき計算する。
6. こうして求めたステレオ信号を出力する。

ステップ 3 乃至ステップ 5 は、アップミキシングのプロセスと考えてもよい。ステップ 1

から6はそれぞれ、本書類の前出の部分で開示したいずれかのデコーダシステムの対応する機能と同様であり、実装に関する詳細事項は同部分から読み取ることができる。

【0084】

図16は、ステレオ信号をビットストリーム信号にエンコードする方法を示し、次のステップを有する：

1. ステレオ信号を入力する。
2. ステレオ信号を第1周波数領域表示に変換する。
3. 複素予測係数を決定する。
4. 周波数領域表示をダウンミックスする。
5. ダウンミックスチャンネルと残差チャンネルを、複素予測係数とともにビットストリームとしてエンコードする。
6. ビットストリームを出力する。

ステップ1から5はそれぞれ、本書類の前出の部分で開示したいずれかのエンコーダシステムの対応する機能と同様であり、実装に関する詳細事項は同部分から読み取ることができる。

【0085】

両方法は、ソフトウェアプログラムの形式のコンピュータ読み取り可能命令として表現でき、コンピュータで実行できる。本発明の保護の範囲は、かかるソフトウェアと、かかるソフトウェアを配布するためのコンピュータプログラム製品とに及ぶ。

【0086】

IV. 実験的評価

ここに開示した実施形態を実験的に評価した。このプロセスで得られた実験的資料の最も重要な部分を以下にまとめる。

【0087】

実験に用いた実施形態は次の特徴を有する：

(i) (時間フレームの)各MDS Tスペクトルを、現在の、前の、及び次のMDC Tスペクトルから2次元有限インパルス応答フィルタリングにより計算した。

(ii) USACステレオエンコーダからの音響心理学的モデルを用いた。

(iii) PSパラメータICC、CLD及びIPDの代わりに、複素予測係数の実部と虚部が送信を送信した。実部と虚部は、別々に処理され、 $[-3.0, 3.0]$ の範囲に制限され、0.1のステップサイズを用いて量子化される。時間微分符号化し、最終的にUSACのスケールファクタコードブックを用いてハフマン符号化する。予測係数は1スケールファクタ帯域おきに更新され、周波数解像度がMP EGサラウンド(例えば、ISO/IEC 23003-1参照)と同様になった。この量子化および符号化スキームにより、目標ビットレートが96 kb/sの一般的な構成において、ステレオサイド情報の平均ビットレートが約2 kb/sになった。

(iv) 2ビットのms_mask_presentビットストリーム要素の取り得る値は3つしかないので、現在のUSACビットストリームを壊すことなく、ビットストリームフォーマットを修正した。複素予測を示す第4の値を用いることにより、ビットを無駄にせずに、基本的なミッド/サイド符号化のフォールバックモードを許した(これについては、本開示の前のサブセクションを参照されたい)。

【0088】

ヘッドホンで再生し、サンプリングレートが48 kHzの8テストアイテムを用いたMUSHR A法によりリスニングテストを行った。各テストには3人、5人または6人の被験者が参加した。

【0089】

MDS T近似が異なることによるインパクトを評価して、これらのオプションの間にある実際的な複雑さ対音質のトレードオフを示した。結果を図12と図13に示した。前者は得られた絶対スコアを示し、後者は96 s USAC cplfに対する、すなわちMDS Tの近似値を計算するために現在のMDC Tフレームを用いた複素予測によるMDC

10

20

30

40

50

T領域統一ステレオ符号化に対する差分スコアを示す。MDC Tベース統一ステレオ符号化により実現される音質ゲインは、MDS Tスペクトルを計算するのに計算的により複雑なアプローチを用いると、増加することが分かる。テスト全体の平均を考えると、単一フレームベースシステム 96s USAC cp1fにより、従来のステレオ符号化に対して、符号化効率が大幅に上昇する。同様に、96s USAC cp3fの場合、すなわち、MDS Tを計算するために、現在の、前の、及び次のMDC Tフレームを用いる複素予測によるMDC T領域統一ステレオ符号化の場合、さらによい結果が得られる。

【0090】

V. 結語

本発明のさらなる実施形態は、上記の説明を読めば、当業者には明らかになるだろう。本明細書と図面は実施形態と実施例を開示しているが、本発明はこれらの具体的な例に制約されない。添付した特許請求の範囲で規定した本発明の範囲から逸脱することなく、多数の修正や変形をすることができる。

【0091】

留意点として、この出願で開示した方法と装置は、通常の実験を含む当業者の能力の範囲内で適当な修正をして、2より多いチャンネルを有する信号の符号化に適用できる。強調しておくが、説明した実施形態に関連して述べた信号、パラメータ、及び行列は、周波数可変または周波数不変及び/又は時間可変または時間不変であってもよい。説明した計算ステップは、周波数ごとに、または一度にすべての周波数に対して行え、すべてのエンティティは周波数選択的動作を有するように実施できる。出願の目的において、任意の量子化スキームを音響心理学モデルにより適応できる。さらに、留意点として、様々な和・差変換、すなわちダウンミックス/残差形式から擬似L/R形式への変換及びL/R-to-M/S変換及びM/S-to-L/R変換はすべて次の形式

【数11】

$$g \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

であり、単にゲインファクタgのみが変化する。よって、ゲインファクタを個別に調整することにより、復号ゲインを適切に選択することにより、符号化ゲインを補正することができる。さらに、当業者には明らかのように、偶数個の直列に配置された差・差変換は、パスルー段階に影響を及ぼし、場合によってゲインは1ではない。

【0092】

ここに開示したシステムと方法は、ソフトウェア、ファームウェア、ハードウェアまたはこれらの組み合わせとして実施できる。一部または全部のコンポーネントは、デジタルシグナルプロセッサやマイクロプロセッサにより実行されるソフトウェアとして実施でき、またはハードウェアまたは特定目的集積回路として実施できる。かかるソフトウェアは、コンピュータ読み取り可能媒体で配布可能である。コンピュータ読み取り可能媒体は、コンピュータ記憶媒体と通信媒体を含む。当業者には周知だが、コンピュータ記憶媒体には、コンピュータ読み取り可能命令、データ構造、プログラムモジュールその他のデータなどの情報を記憶するための任意の方法や技術で実施された、揮発性および不揮発性、取り外し可能および取り外し不可能媒体を含む。コンピュータ記憶媒体は、RAM、ROM、EEPROM、フラッシュメモリその他のメモリ技術、CD-ROM、デジタルバーサタイルディスク(DVD)その他の光ディスク記憶媒体、磁気カセット、磁気テープ、磁気ディスク記憶その他の磁気記憶デバイス、またはその他の、所望の情報の記憶に使える任意の媒体を含むが、これらに限定されない。さらに、当業者には知られているように、通信媒体は、一般的に、コンピュータ読み取り可能命令、データ構造、プログラムモジュール、その他の搬送波その他の伝送メカニズムなどの変調データ信号中のデータを化体し、任意の情報配信媒体を含む。

10

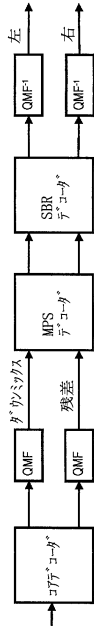
20

30

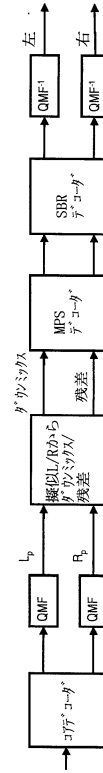
40

50

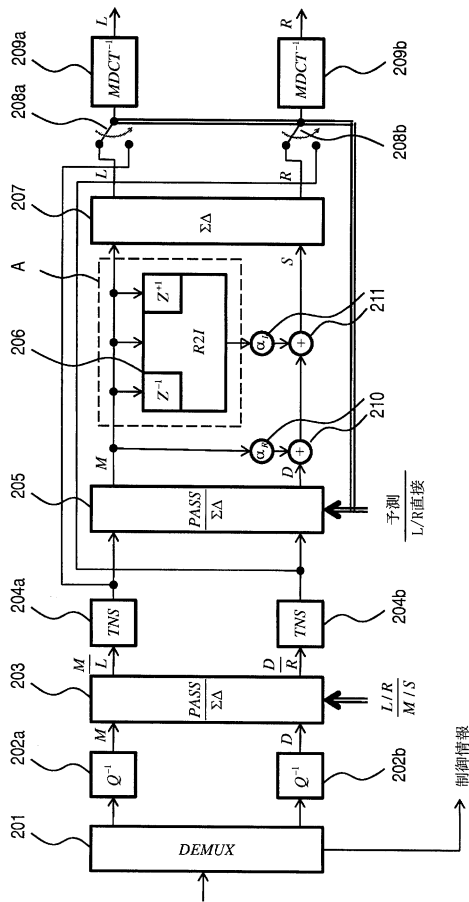
【図1A】



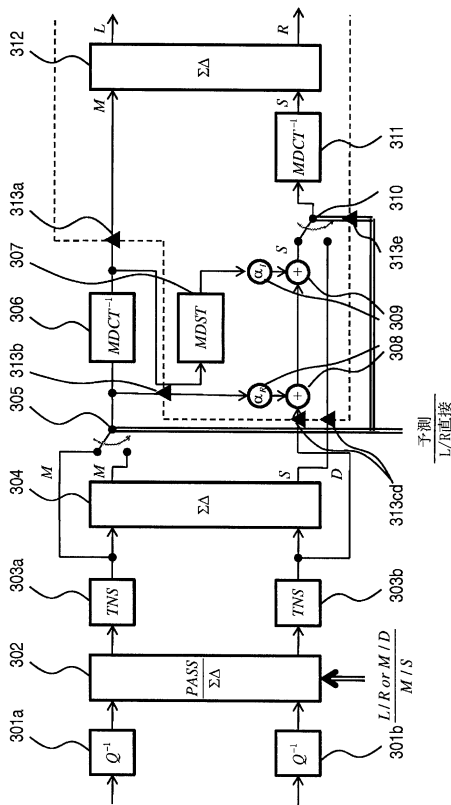
【図1B】



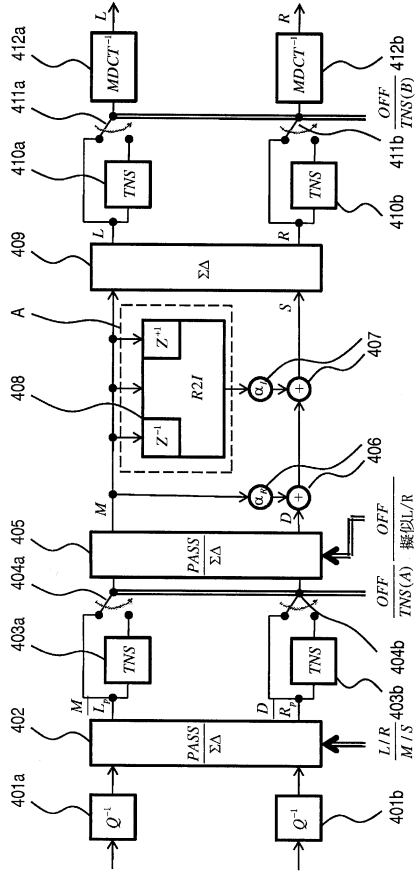
【図2】



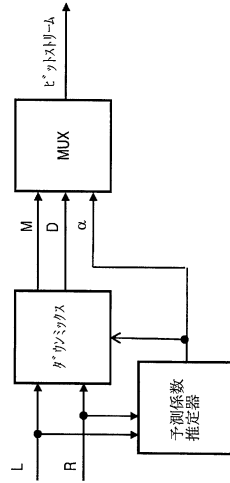
【図3】



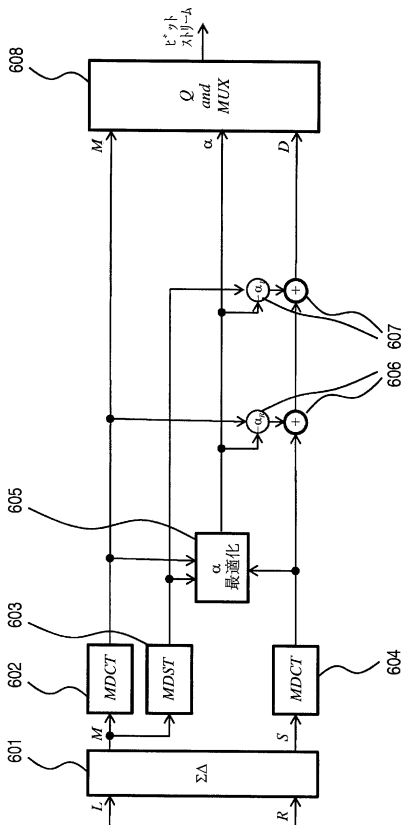
【図4】



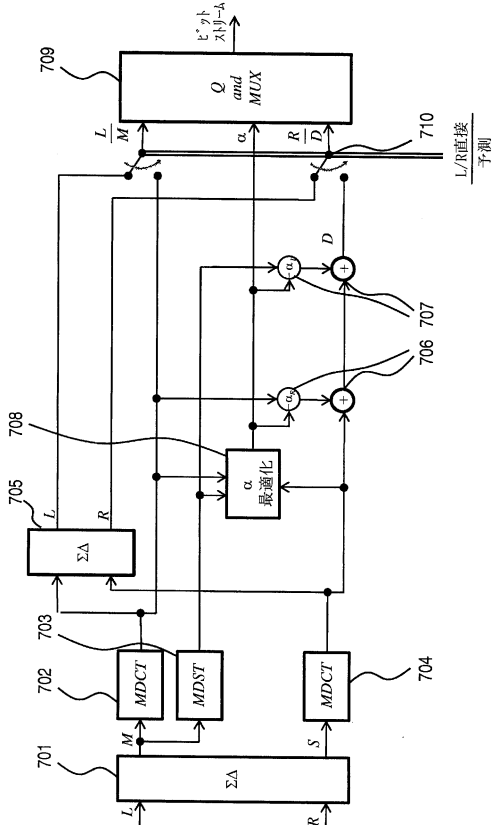
【図5】



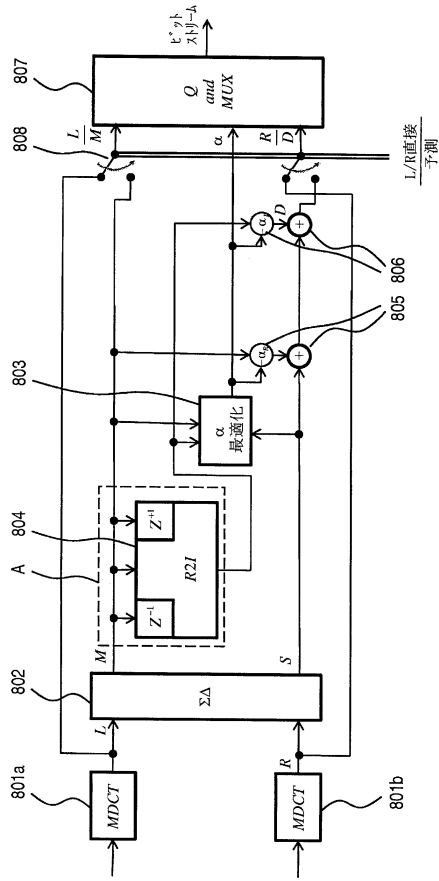
【図6】



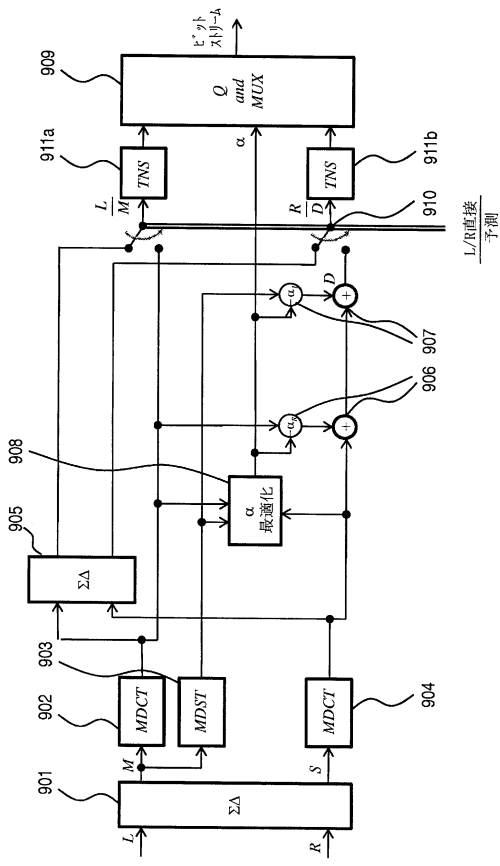
【図7】



【図 8】



【図 9】



【図 10】

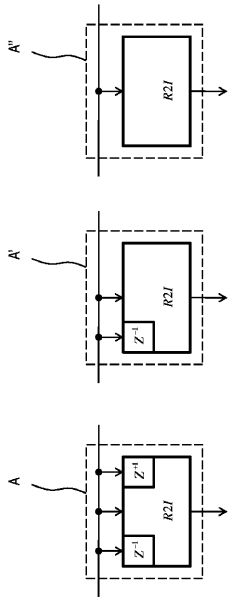
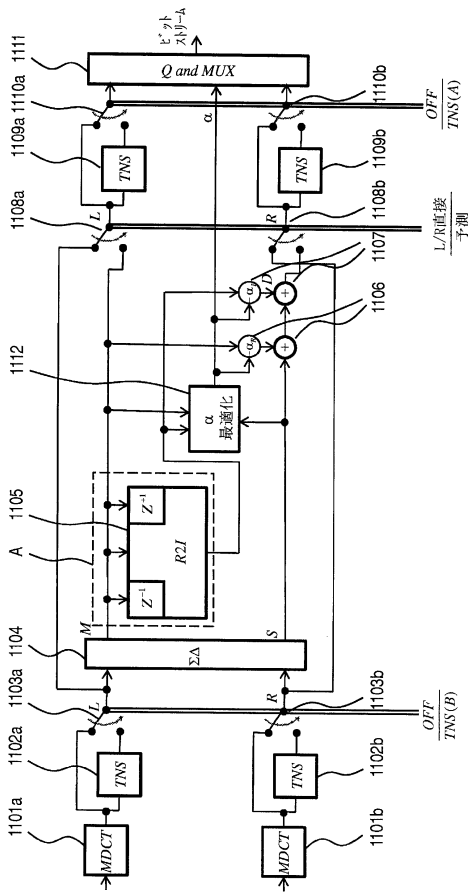
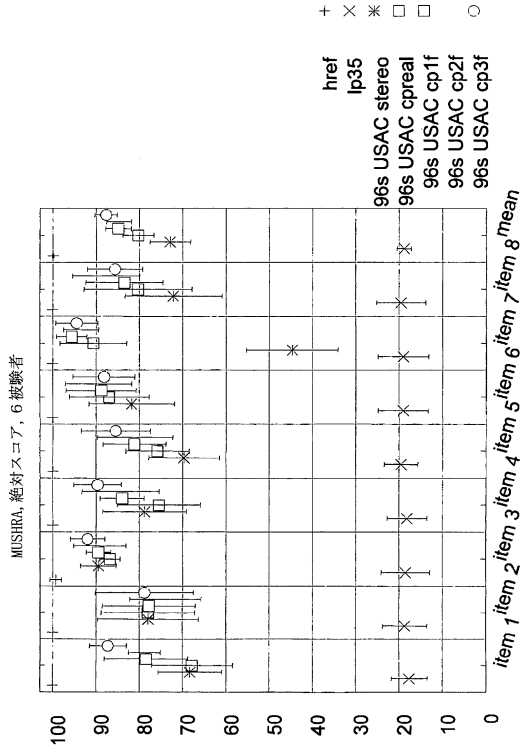


Fig. 10

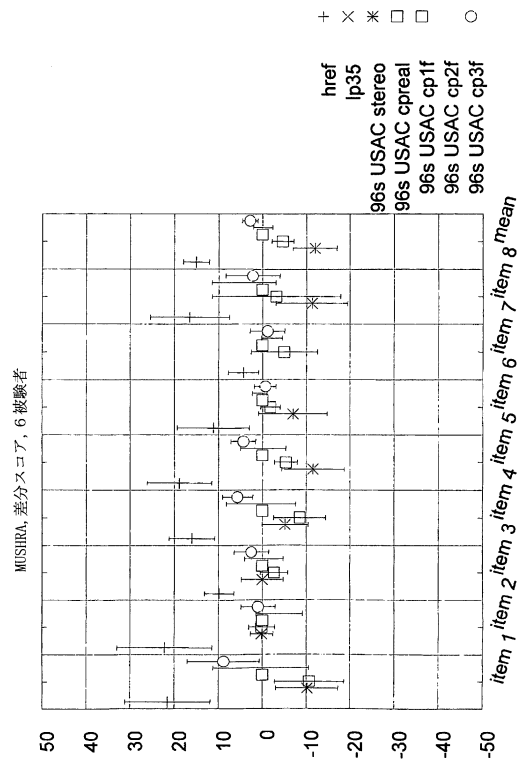
【図 11】



【 図 1 2 】



【 図 1 3 】



【 図 1 4 A 】

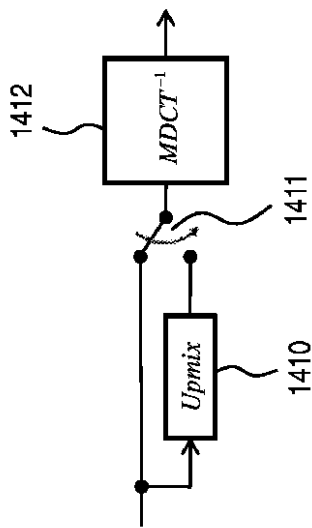
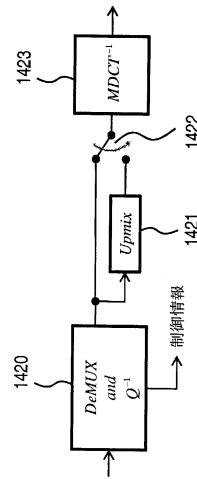
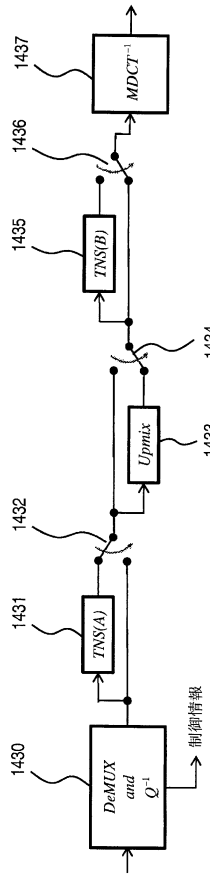


Fig. 14A

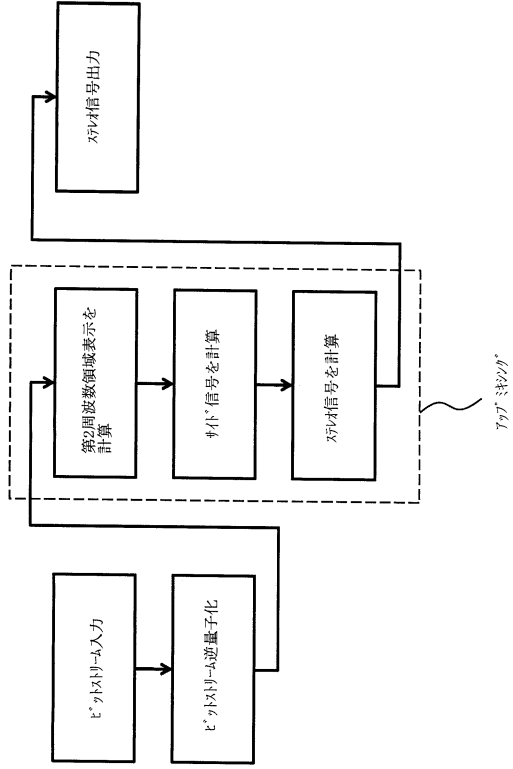
【 図 1 4 B 】



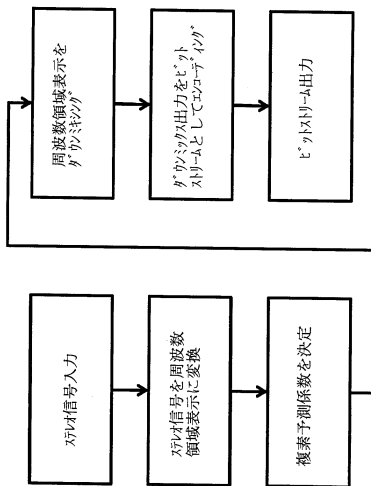
【図14C】



【図15】



【図16】



フロントページの続き

- (72)発明者 ブルンハーゲン, ヘイコ
スウェーデン王国, 113 30 ストックホルム, ガヴレガタン 12アー ドルビー スウェ
ーデン アーバー内
- (72)発明者 カルルスソン, ポンタス
スウェーデン王国, 168 32 ブロツマ, バイグマスタールヴァーゲン 3
- (72)発明者 ヴィレモーズ, ラルス
スウェーデン王国, 113 30 ストックホルム, ガヴレガタン 12アー ドルビー スウェ
ーデン アーバー内

審査官 山下 剛史

- (56)参考文献 国際公開第2009/141775(WO, A1)
国際公開第2009/049896(WO, A1)
特表2008-516275(JP, A)
特表平4-506141(JP, A)
Heiko PURNHAGEN, et al., "Technical description of proposed Unified Stereo Coding in U
SAC", ISO/IEC JTC1/SC29/WG11, 2009年10月, MPEG2009/M16921, pp.1-14
Max NEUENDORF, "WD5 of USAC", ISO/IEC JTC1/SC29/WG11, 2009年10月, MPEG2009/N1104
0, pp.1-146

(58)調査した分野(Int.Cl., DB名)

G10L 19/00 - 19/26