

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5758902号
(P5758902)

(45) 発行日 平成27年8月5日 (2015.8.5)

(24) 登録日 平成27年6月12日 (2015.6.12)

(51) Int.Cl.

F I

G 1 O L 19/008 (2013.01)
 G 1 O L 19/00 (2013.01)
 H O 4 S 3/00 (2006.01)
 G 1 O L 19/02 (2013.01)

G 1 O L 19/008
 G 1 O L 19/00 4 O O Z
 H O 4 S 3/00 Z
 G 1 O L 19/02 1 5 O

請求項の数 21 (全 40 頁)

(21) 出願番号 特願2012-533643 (P2012-533643)
 (86) (22) 出願日 平成22年10月15日 (2010.10.15)
 (65) 公表番号 特表2013-507664 (P2013-507664A)
 (43) 公表日 平成25年3月4日 (2013.3.4)
 (86) 国際出願番号 PCT/EP2010/065503
 (87) 国際公開番号 W02011/045409
 (87) 国際公開日 平成23年4月21日 (2011.4.21)
 審査請求日 平成24年6月7日 (2012.6.7)
 (31) 優先権主張番号 61/252, 298
 (32) 優先日 平成21年10月16日 (2009.10.16)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 10171459.0
 (32) 優先日 平成22年7月30日 (2010.7.30)
 (33) 優先権主張国 欧州特許庁 (EP)

(73) 特許権者 591037214
 フラウンホッフアーゲーゼルシャフト ツ
 ァ フェルダールング デア アンゲヴァ
 ンテン フォアシュンク エー. ファオ
 ドイツ連邦共和国 80686 ミュンヘ
 ン ハンザシュトラッセ 27 ツェー
 (74) 代理人 100079577
 弁理士 岡田 全啓
 (74) 代理人 100167966
 弁理士 扇谷 一
 (72) 発明者 ヘレ ユールゲン
 ドイツ連邦共和国 91054 ブッケン
 ホフ ハーラーシュトラッセ 24

最終頁に続く

(54) 【発明の名称】 ダウンミックス信号表現と、ダウンミックス信号表現に関するパラメトリックサイド情報に基づくアップミックス信号表現の提供に対して、平均値を用いて、1つ以上の調整されたパラメー

(57) 【特許請求の範囲】

【請求項 1】

ダウンミックス信号表現 (2 1 0 ; 4 2 0) と前記ダウンミックス信号表現に係るパラメトリックサイド情報 (2 1 2 ; 4 2 2) に基づくアップミックス信号表現 (2 2 0 ; 4 3 0 a ~ 4 3 0 M) の提供に対して、1つ以上の調整されたパラメータ (1 2 0 ; 2 5 2 ; 3 5 2 ; \tilde{R} ; \tilde{T}) を提供する装置であって、

複数の入力パラメータ (1 1 0 ; 2 1 4 ; 3 3 7 ; R ; T) を受信し、それに基づいて、1つ以上の調整されたパラメータ (1 2 0 ; 2 5 2 ; 3 5 2 ; \tilde{R} ; \tilde{T}) を提供するよう
に構成された、パラメータ調整器を備え、前記パラメータ調整器は、前記アップミックス
信号表現の提供に対して、非最適なパラメータの使用によって生じる前記アップミックス
信号表現の歪みが、最適パラメータから予め定められた偏差よりも大きく偏移している前
記複数の入力パラメータうちの少なくとも1つ以上に対して低減されるように、前記複数の
の入力パラメータ (1 1 0 ; 2 1 4 ; 3 3 7 ; R ; T) の平均値 (1 3 2 ; \bar{R} ; \bar{T}) に従っ
て、前記1つ以上の調整されたパラメータを提供するように構成され、

オーディオオブジェクトの、前記アップミックス信号表現 (2 2 0 ; 4 3 0 a ~ 4 3 0 M) の1つ以上のチャンネルに対する所望の貢献度を記述する複数のユーザ指定されたレンダリング係数 (2 1 4 ; R) を受信し、前記調整されたパラメータとして、1つ以上の調整されたレンダリング係数 (2 5 2 ; \tilde{R}) を提供するよう
に構成された、
装置 (1 0 0 ; 2 5 0 ; 3 5 0 ; 4 4 0 ; 4 5 0) 。

【請求項2】

前記パラメータ調整器は、前記複数の入力パラメータの加重平均である平均値に従って、前記1つ以上の調整されたパラメータを提供するように構成された、請求項1に記載の装置 (1 0 0 ; 2 5 0 ; 3 5 0 ; 4 4 0 ; 4 5 0) 。

【請求項3】

前記パラメータ調整器は、前記1つ以上の調整されたパラメータが、前記平均値から、対応する受信されたパラメータよりも小さく偏移するように、前記1つ以上の調整されたパラメータを提供するように構成された、請求項1または2に記載の装置 (1 0 0 ; 2 5 0 ; 3 5 0 ; 4 4 0 ; 4 5 0) 。

【請求項4】

前記パラメータ調整器は、前記入力パラメータとして、複数のレンダリング係数 (2 1 4 ; R) を受信し、

前記パラメータ調整器は、複数のオーディオオブジェクトに係るレンダリング係数を通じた平均 (\bar{R}) を演算するように構成され、

前記パラメータ調整器は、前記複数のオーディオオブジェクトに係るレンダリング係数を通じた平均からの、調整されたレンダリング係数の偏差が限定されるように、前記調整されたレンダリング係数 (2 5 2 ; \tilde{R}) を提供するよう
に構成された、

請求項1～3のいずれかに記載の装置 (1 0 0 ; 2 5 0 ; 4 4 0) 。

【請求項5】

前記パラメータ調整器は、前記レンダリング係数を通じた平均 (\bar{R}) に従って決定される許容差の範囲内にあるレンダリング係数 (214; R) を不変のままにし、
 前記許容差の上側境界値 ($\Lambda \bar{R}$) より大きいレンダリング係数 (214; R) を、前記許容差の範囲内であって、前記上側境界値より小さいまたは等しい値に選択的にセットし、
 前記許容差の下側境界値 ($\frac{1}{\Lambda} \bar{R}$) より小さいレンダリング係数 (214; R) を、前記許容差の範囲内であって、前記下側境界値より大きいまたは等しい値に選択的にセットする
 ように構成された、
 請求項4に記載の装置 (100; 250; 440)。

10

【請求項6】

前記パラメータ調整器は、前記レンダリング係数を通じた平均に従って決定される許容差の外側にあるレンダリング係数を、反復的に前記許容差内に持ってくるために、それぞれの反復において、前記レンダリング係数を通じた平均 (\bar{R}) からの最大偏差 ($R_{d,max}$) を含む前記レンダリング係数のそれぞれの1つ ($R(i_{max})$) を反復的に選択し、前記レンダリング係数の選択された1つ ($R(i_{max})$) を、前記レンダリング係数を通じた平均 (\bar{R}) に近づく演算を行うように構成された、
 請求項4に記載の装置 (100; 250; 440)。

20

【請求項7】

前記パラメータ調整器は、前記レンダリング係数のそれぞれの1つ ($R(i_{max})$) を反復的に選択し、前記レンダリング係数の選択された1つを前記レンダリング係数を通じた平均に近づく演算を、全てのレンダリング係数が適用可能な許容差内にあるように調整されるまで繰り返すように構成された、
 請求項6に記載の装置 (100; 250; 440)。

【請求項8】

前記ダウンミックス信号表現 (210; 420) の1つ以上のチャンネルの、前記アップミックス信号表現 (220; 430a~430M) の1つ以上のチャンネルへのマッピングを記述する、1つ以上の変換符号化係数 (337; T) を受信するように構成され、

30

前記調整されたパラメータとして、1つ以上の調整された変換符号化係数 (352; \tilde{T}) を提供するように構成された、
 請求項1または2に記載の装置 (100; 350; 450)。

【請求項9】

前記パラメータ調整器は、前記入力パラメータとして、変換符号化係数 (337; T) の時間シーケンスを受信するように構成され、

前記パラメータ調整器は、複数の変換符号化係数に従って時間平均 (\bar{T}) を演算するよう
 に構成され、

40

前記パラメータ調整器は、前記調整された変換符号化係数の、前記時間平均からの偏差が限定されるように、前記調整された変換符号化係数 (352; \tilde{T}) を提供するように構成された、
 請求項8に記載の装置 (100; 350; 450)。

【請求項10】

前記パラメータ調整器は、前記時間平均 (\bar{T}) に従って決定される許容差の範囲内にある変換符号化係数 ($337; T$) を、不変のままにし、

前記許容差の上側境界値 ($\Lambda \bar{T}$) より大きい変換符号化係数を、前記許容差の範囲内であって、前記許容差の上側境界値より小さいまたは等しい値に選択的にセットし、

前記許容差の下側境界値 ($1/\Lambda \bar{T}$) より小さい変換符号化係数を、前記許容差の範囲内であって、前記許容差の下側境界値より大きいまたは等しい値に選択的にセットするよう
に構成された、

請求項 9 に記載の装置 ($100; 350; 450$)。

10

【請求項 11】

前記パラメータ調整器は、前記変換符号化係数 ($337; T$) のシーケンスの再帰的ローパスフィルタリングを用いて、前記時間平均 (\bar{T}) を算出するように構成された、
請求項 9 または 10 に記載の装置 ($100; 350; 450$)。

【請求項 12】

前記パラメータ調整器は、前記調整されたパラメータの所定の 1 つが、前記複数の入力パラメータの平均値 ($132; \bar{R}; \bar{T}; \bar{X}$) と 1 つ以上の許容パラメータ ($\Lambda_{R-}; \Lambda_{R+}; \Lambda_{T-}; \Lambda_{T+}; \Lambda_{X-}; \Lambda_{X+}$) に従って境界が定められる許容差の範囲内にあるように、そして、入力パラメータと、対応する調整されたパラメータとの偏差が最小化され、または予め定められた最大許容範囲内に保持されるように、前記 1 つ以上の調整されたパラメータの所定の 1 つを提供するように構成された、
請求項 1 ～ 11 のいずれかに記載の装置 ($100; 250; 350; 440; 450$)。

20

【請求項 13】

前記パラメータ調整器は、前記入力パラメータの調整されたバージョンを取得するために、前記複数の入力パラメータの平均値 ($132; \bar{R}; \bar{T}; \bar{X}$) に従って境界が定められる前記許容差の外側にあることがわかった入力パラメータを、前記許容差の上側境界値 ($\Lambda \bar{R}; \Lambda \bar{T}; \Lambda \bar{X}; \bar{X} + \Lambda_{X+}$) または下側境界値 ($\frac{1}{\Lambda} \bar{R}; \frac{1}{\Lambda} \bar{T}; \frac{1}{\Lambda} \bar{X}; \bar{X} - \Lambda_{X-}$) に、選択的にセットするように構成された、
請求項 12 に記載の装置 ($100; 250; 350; 440; 450$)。

30

【請求項 14】

前記パラメータ調整器は、前記平均値に従って境界が定められる許容差の外側にあると決定された入力パラメータを、反復的に前記許容差内に持ってくるために、それぞれの反復において、前記平均値 ($132; \bar{R}; \bar{T}; \bar{X}$) からの最大偏差を含む前記入力パラメータのそれぞれの 1 つ ($R(i_{\max}); X_{i^*}$) を反復的に選択し、前記入力パラメータの選択された 1 つを前記平均値に近づける演算を行うように構成された、
請求項 12 に記載の装置 ($100; 250; 350; 440; 450$)。

40

【請求項 15】

前記パラメータ調整器は、前記入力パラメータの選択された 1 つ ($R(i_{\max}); X_{i^*})$ を前記平均値に近づける演算において使用される修正ステップサイズを、前記入力パラメータの選択された 1 つと前記平均値との差異の予め定められた分数になるように、選択するように構成された、
請求項 14 に記載の装置 ($100; 350; 450$)。

50

【請求項 16】

ダウンミックス信号表現 (210; 420) とパラメトリックサイド情報 (212; 422) に基づいて、アップミックス信号表現 (220; 430a~430M) を提供する装置であって、

複数の入力パラメータ (110; 214; 337; R; T) に基づいて、1つ以上の調整されたパラメータ (120; 252; 352; \tilde{R} ; \tilde{T}) を提供する、請求項1~15のいずれかに記載の装置 (100; 250; 350; 440; 450) と、

前記ダウンミックス信号表現および前記パラメトリックサイド情報に基づいて、前記アップミックス信号表現を取得するように構成された、信号処理器 (230; 330) とを備え、

前記1つ以上の調整されたパラメータを提供する装置は、前記信号処理器の1つ以上の処理パラメータ (110; 214; 337; R; T) を調整するように構成された、装置 (200; 300; 410)。

10

【請求項 17】

前記信号処理器 (230) は、前記アップミックス信号表現の1つ以上のチャンネルに対するオーディオオブジェクトの貢献度を記述する、調整されたレンダリング係数 (252; \tilde{R}) に従って、前記アップミックス信号表現 (220; 430a~430M) を提供するように構成され、

前記1つ以上の調整されたパラメータを提供する装置 (100; 250; 440) は、前記入力パラメータとして、複数のユーザ指定のレンダリングパラメータ (214; R) を受信し、それに基づいて、前記信号処理器による使用のために、1つ以上の調整されたレンダリングパラメータ (252; \tilde{R}) を提供するように構成された、

請求項16に記載された、装置 (200; 300; 410)。

20

【請求項 18】

前記1つ以上の調整されたパラメータを提供する装置 (100; 350; 450) は、前記入力パラメータとして、混合マトリックスの1つ以上の混合マトリクス要素 (337; T) を受信し、それに基づいて、前記信号処理器 (330) による使用のために、前記混合マトリックスの1つ以上の調整された混合マトリクス要素 (352; \tilde{T}) を提供するように構成され、

前記信号処理器 (330) は、前記ダウンミックス信号表現の1つ以上のオーディオチャンネル信号の、前記アップミックス信号表現の1つ以上のオーディオチャンネル信号へのマッピングを記述する前記混合マトリックスの前記調整された混合マトリクス要素 (352; \tilde{T}) に従って、前記アップミックス信号表現 (220; 430a~430M) を提供するように構成された、

請求項16に記載された、装置 (300; 410)。

30

40

【請求項 19】

前記信号処理器は、MPEGサラウンド任意ダウンミックスゲイン値を取得するように構成され、

前記1つ以上の調整されたパラメータを提供する装置は、前記入力パラメータとして、複数の任意ダウンミックスゲイン値を受信し、複数の調整された任意ダウンミックスゲイン値を提供するように構成された、

請求項16に記載された、装置 (200; 300; 410)。

【請求項 20】

ダウンミックス信号表現と、前記ダウンミックス信号表現に関するパラメトリックサイド情報に基づくアップミックス信号表現の提供に対して、1つ以上の調整されたパラメ

50

ータを提供する方法であって、

オーディオオブジェクトの、前記アップミックス信号表現(220; 430a~430M)の1つ以上のチャンネルに対する所望の貢献度を記述する複数のユーザ指定されたレンジング係数(214; R)を含む複数のパラメータを受信するステップと、

前記受信されたパラメータに基づいて、非最適なパラメータの使用によって生じるアップミックス信号表現の歪みが、最適パラメータから予め定められた偏差より大きく偏移している前記複数のパラメータのうちの少なくとも1つ以上に対して低減されるように、複数のパラメータの平均値に従って、前記1つ以上の調整されたパラメータを提供するステップと、

を備えた、方法。

10

【請求項21】

コンピュータプログラムがコンピュータ上で動作するときに、請求項20に記載された方法を実行する、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明に係る実施形態は、ダウンミックス信号表現と、ダウンミックス信号表現に係るパラメトリックサイド情報に基づくアップミックス信号表現の提供に対して、1つ以上の調整されたパラメータを提供する装置に関する。

20

【0002】

本発明に係る他の実施形態は、ダウンミックス信号表現とパラメトリックサイド情報に基づいてアップミックス信号表現を提供する装置に関する。

【0003】

本発明に係る他の実施形態は、ダウンミックス信号表現と、ダウンミックス信号表現に係るパラメトリックサイド情報に基づくアップミックス信号表現の提供に対して、1つ以上の調整されたパラメータを提供する方法に関する。

【0004】

本発明に係る他の実施形態は、前記方法を実行するためのコンピュータプログラムに関する。

30

【0005】

本発明に係るいくつかの実施形態は、MPEG SAOCにおける歪み制御のためのパラメータ制限スキームに関する。

【背景技術】

【0006】

オーディオ処理、オーディオ伝送およびオーディオ記憶の技術において、聴覚インプレッションを改善するために、マルチチャンネルコンテンツを取り扱うという増大する要望がある。マルチチャンネルオーディオコンテンツの使用は、ユーザに対して有意の進歩をもたらす。例えば、娯楽アプリケーションにおいて、改善されたユーザ満足度をもたらす三次元聴覚インプレッションを取得することができる。しかしながら、マルチチャンネルオーディオコンテンツは、また、マルチチャンネルオーディオ再生を用いて話者了解度を改善することができるので、専門の環境、例えば電話会議アプリケーションにおいて有用である。

40

【0007】

しかしながら、また、マルチチャンネルアプリケーションによって生じる過度なリソース負荷を回避するために、オーディオ品質とビットレート要求条件との良好なトレードオフを有することが望ましい。

【0008】

最近、ビットレートの効率的な伝送のためのパラメトリック技術および/または多重の

50

オーディオオブジェクトを含むオーディオシーンの記憶、例えば、バイノーラルキュー符号化（I型）（例えば、非特許文献1を参照）、ジョイントソース符号化（例えば、非特許文献2を参照）、およびMPEG空間オーディオオブジェクト符号化（SAOC）（例えば、非特許文献3, 4, 5を参照）が提案されている。

【0009】

受信サイドでのユーザ対話性と共に、このような技術は、極端なオブジェクトレンダリングが実行される場合に、出力信号の低いオーディオ品質に結果として導く可能性がある（例えば、特許文献1を参照）。

【0010】

これらの技術は、所望の出力オーディオシーンを、波形マッチングによるよりもむしろ知覚的に復元することを目指している。

【0011】

図8は、このようなシステム（ここでは、MPEG SAOC）のシステム概要を示す。図8に示されるMPEG-SAOCシステム800は、SAOCエンコーダ810とSAOCデコーダ820を備える。SAOCエンコーダ810は、例えば、時間ドメイン信号として、または時間周波数ドメイン信号（例えば、フーリエタイプ変換の変換係数のセットの形の、またはQMFサブバンド信号の形の）として表すことができる複数のオブジェクト信号 $x_1 \sim x_N$ を受信する。SAOCエンコーダ810は、通常は、オブジェクト信号 $x_1 \sim x_N$ に関するダウンミックス係数 $d_1 \sim d_N$ も受信する。ダウンミックス係数の分離したセットは、ダウンミックス信号の各チャンネルに対して利用することができる。SAOCエンコーダ810は、通常は、オブジェクト信号 $x_1 \sim x_N$ に関するダウンミックス係数 $d_1 \sim d_N$ に従って結合することによって、ダウンミックス信号のチャンネルを取得するように構成される。通常、ダウンミックスチャンネルは、オブジェクト信号 $x_1 \sim x_N$ より少ない。SAOCデコーダ820の側でのオブジェクト信号の分離（または分離処理）を（少なくとも近似的に）可能とするため、SAOCエンコーダ810は、1つ以上のダウンミックス信号（ダウンミックスチャンネルとして示される）812と、サイド情報814の両方を提供する。サイド情報814は、デコーダ側でのオブジェクト特有の処理を可能とするため、オブジェクト信号 $x_1 \sim x_N$ の特性を記述する。

【0012】

SAOCデコーダ820は、1つ以上のダウンミックス信号812とサイド情報814の両方を受信するように構成される。また、SAOCデコーダ820は、通常は、所望のレンダリングセットアップを記述するユーザ対話情報および/またはユーザ制御情報822を受信するように構成される。例えば、ユーザ対話情報/ユーザ制御情報822は、スピーカセットアップと、オブジェクト信号 $x_1 \sim x_N$ を提供するオブジェクトの所望の空間配置を記述することができる。

【0013】

10

20

30

S A O Cデコーダ820は、例えば、複数の復号化されたアップミックスチャンネル信号 $\hat{y}_1 \sim \hat{y}_M$ を提供するように構成される。アップミックスチャンネル信号は、例えば、マルチスピーカのレンダリング構成の個々のスピーカに関係することができる。S A O Cデコーダ820は、例えば、1つ以上のダウンミックス信号812とサイド情報814に基づいてオブジェクト信号 $x_1 \sim x_N$ を、少なくとも近似的に、復元するように構成され、それにより復元されたオブジェクト信号820bを取得する、オブジェクトセパレータ820aを備えることができる。しかしながら、復元されたオブジェクト信号820bは、例えば、サイド情報814が、ビットレートの制約のため、完全な復元に対してかなり充分ではないので、オリジナルのオブジェクト信号 $x_1 \sim x_N$ からいくぶん偏移している可能性がある。S A O Cデコーダ820は、復元されたオブジェクト信号820cとユーザ対話情報／ユーザ制御情報822を受信し、それに基づいて、アップミックスチャンネル信号 $\hat{y}_1 \sim \hat{y}_M$ を提供するように構成することができる、混合器820cを更に備えることができる。混合器820cは、ユーザ対話情報／ユーザ制御情報822を用い、個々の復元されたオブジェクト信号820bのアップミックスチャンネル信号 $\hat{y}_1 \sim \hat{y}_M$ に対する貢献度を決定するように構成することができる。ユーザ対話情報／ユーザ制御情報822は、例えば、個々の復元されたオブジェクト信号822のアップミックスチャンネル信号 $\hat{y}_1 \sim \hat{y}_M$ に対する貢献度を決定する、レンダリングパラメータ（レンダリング係数としても示される）を備えることができる。

【0014】

しかしながら、多くの実施形態において、図8においてオブジェクトセパレータ820aで示されるオブジェクト分離と、図8において混合器820cで示される混合が、1つのシングルステップで実行される点に注意しなければならない。この目的のため、1つ以上のダウンミックス信号812のアップミックスチャンネル信号 $\hat{y}_1 \sim \hat{y}_M$ への直接マッピングを記述する全体のパラメータを演算することができる。これらのパラメータは、サイド情報とユーザ対話情報／ユーザ制御情報820に基づいて演算することができる。

【0015】

ここで図9a、9b、9cを参照して、ダウンミックス信号表現とオブジェクト関連サイド情報に基づいてアップミックス信号表現を取得する異なる装置が記載される。オブジェクト関連サイド情報は、ダウンミックス信号に関係するサイド情報の例である点に注意しなければならない。図9aは、S A O Cデコーダ920を備えるM P E G S A O Cシステム900の概略ブロック図である。S A O Cデコーダ920は、分離した機能ブロックとして、オブジェクトデコーダ922と混合器／レンダラー926を備える。オブジェクトデコーダ922は、ダウンミックス信号表現（例えば、時間ドメインにおいてまたは時間周波数ドメインにおいて表現される1つ以上のダウンミックス信号の形の）と、オブジェクト関連のサイド情報（例えば、オブジェクトメタデータの形の）に従って、複数の復元されたオブジェクト信号924を提供する。混合器／レンダラー926は、複数のN個のオブジェクトに関係する復元されたオブジェクト信号924を受信し、それとレンダリング情報に基づいて、1つ以上のアップミックスチャンネル信号928を提供する。S A O Cデコーダ920において、オブジェクト信号924の抽出は、オブジェクト復号化機能の混合／レンダリング機能からの分離を可能とする混合／レンダリングから分離して実行されるが、比較的高い計算量をもたらす。

【0016】

ここで図9bを参照して、S A O Cデコーダ950を備える他のM P E G S A O Cシステム930が簡単に述べられる。S A O Cデコーダ950は、ダウンミックス信号表現（例えば、1つ以上のダウンミックス信号の形の）と、オブジェクト関連サイド情報（例えば、オブジェクトメタデータの形の）に従って、複数のアップミックスチャンネル信号958を提供する。S A O Cデコーダ950は、オブジェクト復号化と混合／レンダリングの分離なしの合同の混合プロセスにおいてアップミックスチャンネル信号958を取得するように構成され、前記合同のアップミックスプロセスに対するパラメータがオブジェクト関連サイド情報とレンダリング情報の両方に依存する、複合されたオブジェクトデコーダおよび混合器／レンダラーを備える。合同のアップミックスプロセスは、オブジェク

10

20

30

40

50

ト関連サイド情報の一部であるとみなされるダウンミックス情報にも依存する。

【 0 0 1 7 】

上記を要約すると、アップミックスチャンネル信号 9 2 8、9 5 8 の提供は、1 つのステッププロセスまたは 2 つのステッププロセスにおいて実行することができる。

【 0 0 1 8 】

ここで図 9 c を参照して、M P E G S A O C システム 9 6 0 が記載される。S A O C システム 9 6 0 は、S A O C デコーダよりもむしろ、S A O C M P E G サラウンドトランスコーダ 9 8 0 を備える。

【 0 0 1 9 】

S A O C M P E G サラウンドトランスコーダは、オブジェクト関連サイド情報（例えば、オブジェクトメタデータの形の）と、オプションとして、1 つ以上のダウンミックス信号とレンダリング情報に関係する情報を受信するように構成された、サイド情報トランスコーダ 9 8 2 を備える。サイド情報トランスコーダは、また、受信されたデータに基づいて、M P E G サラウンドサイド情報（例えば、M P E G サラウンドビットストリームの形の）を提供するように構成される。したがって、サイド情報トランスコーダ 9 8 2 は、オブジェクトエンコーダから受信されるオブジェクト関連（パラメトリック）サイド情報を、レンダリング情報とオプションとして 1 つ以上のダウンミックス信号のコンテンツについての情報を考慮に入れて、チャンネル関連（パラメトリック）サイド情報に変換するように構成される。

10

【 0 0 2 0 】

オプションとして、S A O C M P E G サラウンドトランスコーダ 9 8 0 は、例えば、ダウンミックス信号表現によって記述された 1 つ以上のダウンミックス信号を操作し、操作されたダウンミックス信号表現 9 8 8 を取得するように構成することができる。しかしながら、S A O C M P E G サラウンドトランスコーダ 9 8 0 の出力ダウンミックス信号表現 9 8 8 が S A O C M P E G サラウンドトランスコーダの入力ダウンミックス信号表現と同じであるように、ダウンミックス信号操作器 9 8 6 を省略することができる。ダウンミックス信号操作器 9 8 6 は、例えば、いくつかのレンダリング配列において存在する可能性がある、チャンネル関連 M P E G サラウンドサイド情報 9 8 4 が S A O C M P E G サラウンドトランスコーダ 9 8 0 の入力ダウンミックス信号表現に基づいて所望の聴覚インプレッションを提供することを可能にしない場合に、用いることができる。

20

30

【 0 0 2 1 】

したがって、S A O C M P E G サラウンドトランスコーダ 9 8 0 は、S A O C M P E G サラウンドトランスコーダ 9 8 0 へのレンダリング情報入力に従ってオーディオオブジェクトを表現する複数のアップミックスチャンネル信号を、M P E G サラウンドビットストリーム 9 8 4 とダウンミックス信号表現 9 8 8 を受信する M P E G サラウンドデコーダを用いて生成することができるように、ダウンミックス信号表現 9 8 8 と M P E G サラウンドビットストリーム 9 8 4 を提供する。

【 0 0 2 2 】

上記を要約すると、S A O C 符号化されたオーディオ信号を復号化するために異なるコンセプトを用いることができる。場合によっては、ダウンミックス信号表現とオブジェクト関連パラメトリックサイド情報に従ってアップミックスチャンネル信号（例えば、アップミックスチャンネル信号 9 2 8、9 5 8）を提供する、S A O C デコーダが用いられる。このコンセプトに対する実施例は、図 9 a と 9 b に見ることができる。あるいは、S A O C 符号化されたオーディオ情報は、所望のアップミックスチャンネル信号を提供するために M P E G サラウンドデコーダによって用いることができる、ダウンミックス信号表現（例えばダウンミックス信号表現 9 8 8）とチャンネル関連サイド情報（例えば、チャンネル関連 M P E G サラウンドビットストリーム 9 8 4）を取得するために変換することができる。

40

【 0 0 2 3 】

図 8 においてシステム概要が与えられる、M P E G S A O C システム 8 0 0 において

50

、一般的な処理が周波数選択的方法で行われ、各周波数バンド内で以下のように記述することができる。

N個の入力オーディオオブジェクト信号 $x_1 \sim x_N$ は、S A O C エンコーダ処理の一部としてダウンミックスされる。モノラルダウンミックスに対して、ダウンミックス係数は、 $d_1 \sim d_N$ で示される。加えて、S A O C エンコーダ 8 1 0 は、入力オーディオオブジェクトの特性を記述するサイド情報 8 1 4 を抽出する。M P E G S A O C に対して、お互いに関するオブジェクトパワーの関係は、このようなサイド情報の最も基本的な形である。

ダウンミックス信号 8 1 2 とサイド情報 8 1 4 は、送信され、および/または、記憶される。この目的に対して、ダウンミックスオーディオ信号は、M P E G 1 のレイヤ I I または I I I (「. m p 3」としても知られる)、M P E G アドバンストオーディオコーディング (A A C) またはその他のオーディオコードのような周知の知覚的オーディオコードを用いて圧縮することができる。

●受信側において、S A O C デコーダ 8 2 0 は、概念的には、送信されたサイド情報 8 1 4 (および、当然、1つ以上のダウンミックス信号 8 1 2) を用いて、オリジナルのオブジェクト信号を復元(「オブジェクト分離」)しようとする。これらの近似されたオブジェクト信号(復元されたオブジェクト信号 8 2 0 bとしても示される)は、次に、レンダリングマトリクスを用いて、M個のオーディオ出力チャンネルによって表現される目標シーン(例えば、アップミックスチャンネル信号 $\hat{y}_1 \sim \hat{y}_M$ によって表現することができる)に混合される。モノラル出力に対して、レンダリングマトリクス係数は、 $r_1 \sim r_N$ で与えられる。

事実上、オブジェクト信号の分離は、分離ステップ(オブジェクト分離器 8 2 0 a によって示される)と混合ステップ(混合器 8 2 0 c によって示される)の両方がしばしば計算量において莫大な減少に結果としてなる単一の変換符号化ステップに結合されるので、ほとんど実行されない(または決して実行されない)。

【0024】

このようなスキームは、伝送ビットレート(N個の離散オブジェクトオーディオ信号または離散システムの代わりに、少しのダウンミックスチャンネルといくつかのサイド情報を送信することが必要なだけである)と計算量(処理複雑度は、主にオーディオオブジェクトの数よりむしろ出力チャンネル数に関係する)の両方に関して、大いに効率的であることが分かっている。受信端のユーザに対する更なる利益は、ユーザ選択(モノラル、ステレオ、サラウンド、バーチャル化されたヘッドホン再生、その他)のレンダリングセットアップを選択する自由度と、ユーザ対話性の特徴を含み、レンダリングマトリクス、従って出力シーンは、ユーザによって、意志、個人的嗜好または他の基準に従って設定し、対話的に変更することができる。例えば、1つの空間エリアに固まっている1つのグループから話し手を位置決めし、他の残りの話し手からの識別を最大化することが可能である。この対話性は、デコーダ・ユーザインターフェースを提供することによって達成される。

【0025】

各送信されたオブジェクトに対して、その相対レベルと、(非モノラルレンダリングに対して)レンダリングの空間位置を調整することができる。これは、ユーザが付随するグラフィカルユーザインターフェイス(G U I)のスライダの位置を変える(例えば、object level = +5dB, object position = -30deg)ように、リアルタイムに発生することができる。

【0026】

しかしながら、アップミックス信号表現(例えばアップミックスチャンネル信号 $\hat{y}_1 \sim \hat{y}_M$)の提供に対するパラメータのデコーダ側の選択が、場合によっては聞き取れる劣化をもたらすことが分かっている。

【0027】

10

20

30

40

このような状況から鑑みて、本発明の目的は、アップミックス信号表現（例えば、アップミックスチャンネル信号 $\hat{y}_1 \sim \hat{y}_M$ の形の）を提供するときに、聞き取れる歪みを低減する、あるいは回避することさえ可能にするコンセプトを構築することである。

【先行技術文献】

【非特許文献】

【0028】

【非特許文献1】C. Faller および F. Baumgarte、「バイノーラルキュー符号化 第2部：スキームおよびアプリケーション」、IEEE Trans. on Speech and Audio Proc., vol. 11, No. 6、2003年11月

【非特許文献2】C. Faller、「オーディオソースのパラメトリックジョイント符号化」 10
、第120回AES大会、予稿集6752、パリ、2006年

【非特許文献3】J. Herre, S. Disch, J. Hilpert, O. Hellmuth、「SACからSAOC 空間オーディオのパラメトリック符号化における最近の成果」、第22回英国AES会議、ケンブリッジ、英国、2007年4月

【非特許文献4】J. Engdegaerd, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hoelzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen、「空間オーディオオブジェクト符号化(SAOC) パラメトリックオブジェクトベースのオーディオ符号化に関するやがて公開されるMPEG標準」、第124回AES大会、予稿集7377、アムステルダム、2008年

【非特許文献5】ISO/IEC、「MPEGオーディオ技術 第2部：空間オーディオオブジェクト符号化(SAOC)」、ISO/IEC JTC1/SC29/WG11(MPEG) FCD23003-2 20

【非特許文献6】EBU技術勧告：「中間オーディオ品質の主観的リスニングテストのためのMUSHRA EBU法」、文書B/AIM022、1999年10月

【非特許文献7】ISO/IEC JTC1/SC29/WG11(MPEG)、文書N10843、「ISO/IEC23003-2に関する研究：200x年空間オーディオオブジェクト符号化(SAOC)」、第89回MPEGミーティング、ロンドン、英国、2009年7月

【特許文献】

【0029】

【特許文献1】米国特許出願61/173,456、歪みを回避するオーディオ信号処理の方法、装置およびコンピュータプログラム

【発明の概要】

【0030】

上記課題は、ダウンミックス信号表現と、前記ダウンミックス信号表現に係るパラメトリックサイド情報に基づくアップミックス信号表現の提供に対して、1つ以上の適応されたパラメータを提供する装置によって解決される。装置は、1つ以上のパラメータ（それは、いくつかの実施形態において、入力パラメータとすることができる）を受信し、それに基づいて、1つ以上の調整されたパラメータを提供するように構成された、パラメータ調整器を備える。パラメータ調整器は、非最適パラメータの使用によって生じるアップミックス信号表現の歪みが、少なくとも最適パラメータから予め定められた偏差以上偏移しているパラメータ（または入力パラメータ）に対して低減されるように、複数のパラメータ値（それは、いくつかの実施形態において、入力パラメータ値とすることができる）の平均値に従って、1つ以上の調整されたパラメータを提供するように構成される。

【0031】

本発明に係るこの実施形態は、歪みはしばしば平均値からの過剰な偏差によって生ずるので、複数の入力パラメータ値の平均値が、ダウンミックス信号表現とダウンミックス信号表現に係るパラメトリックサイド情報に基づくアップミックス信号表現の提供に用いられるパラメータの調整を可能にする意味のある量を構成するという考えに基づいている。平均値の使用は、平均値(average value)（時には、中間値(mean value)として

30

40

50

も示される)からのこのような過剰な偏差を回避するために、1つ以上のパラメータの調整を可能にし、従って極端に劣化したオーディオ品質を回避するという可能性をもたらす。

【0032】

上述された実施形態は、S A O C デコーダ/トランスコーダがパラメータの調整に必要な全情報を備えるので、全ての処理を完全にS A O C デコーダ/トランスコーダ内で行うことができるレンダリングされたS A O C シーンの主観的音質を保護するコンセプトを提供する。また、パラメータ値と平均値との大きな偏差は、通常は聞き取れる歪みに結果としてなるのに対して、パラメータ値と平均値との偏差の制限は、通常は良好な聴覚インプレッションに結果としてなることが分かっているので、上述の実施形態は、レンダリングされたシーンの知覚されたオーディオ品質の複雑な尺度の明示の計算を含まない。このように、上述された実施形態は、アップミックス信号表現の提供に対して考慮されるパラメータを適切に調整するために、特に効率的なメカニズム、すなわち、平均値の使用を提供する。

10

【0033】

好ましい実施形態において、装置のパラメータ調整器は、複数のパラメータ値の加重平均である平均値に従って、1つ以上の調整されたパラメータを提供するように構成される。

t は異なるパラメータ値に対して異なる重みを割り当てることが可能であるので、加重平均を使用することは高度な自由度を提供する。しかしながら、パラメータ値に対して同じ重みを割り当てることも可能である。

20

【0034】

好ましい実施形態において、装置のパラメータ調整器は、1つ以上の調整されたパラメータが、平均値から、対応する受信されたパラメータよりも小さく偏移するように、1つ以上の調整されたパラメータを提供するように構成される。調整されたパラメータを、平均値の近くに持ってくることによって、またはさらに平均値に等しくセットすることによって、歪みの有意の低減を達成することができる。

【0035】

好ましい実施形態において、装置は、オーディオオブジェクトの、アップミックス信号表現の1つ以上のチャンネルに対する貢献度を記述する1つ以上のレンダリング係数(レンダリングパラメータとしても示される)を受信するように構成される。この場合、装置は、好ましくは、調整されたパラメータとして、1つ以上の調整されたレンダリング係数を提供するように構成される。入力パラメータ値の役割をする複数のレンダリングパラメータの平均値に従ってレンダリングパラメータを調整することは、過剰な聞き取れる歪みを回避する適切に調整されたレンダリングパラメータを取得する可能性をもたらすことが分かっている。

30

【0036】

好ましい実施形態において、パラメータ調整器は、入力パラメータとして、複数のレンダリング係数を受信するように構成される。この場合、パラメータ調整器は、複数のオーディオオブジェクトに関するレンダリング係数を通じた平均を計算するように構成される。また、パラメータ調整器は、調整されたレンダリング係数の、複数のオーディオオブジェクトに関するレンダリング係数を通じた平均からの偏差が限定されるように、調整されたレンダリング係数を提供するように構成される。本発明に係るこの実施形態は、調整されたレンダリング係数の、複数のオーディオオブジェクトに関するレンダリング係数を通じた平均からの偏差が限定される場合に、非最適レンダリングパラメータの使用によって生じるアップミックス信号表現の歪みは、少なくとも最適レンダリングパラメータから予め定められた偏差よりも大きく偏移しているレンダリングパラメータに対して、通常は低減されるという発見に基づいている。このように、簡単なメカニズム、すなわち、調整されたレンダリング係数の、複数のオーディオオブジェクトに関するレンダリング係数を通じた平均からの偏差が限定されるようなレンダリング係数の調整が、過剰な聞き取れる歪みを回避することを可能とする。

40

50

【0037】

好ましい実施形態において、パラメータ調整器は、レンダリング係数を通じた平均に従って決定される許容差の範囲内にあるレンダリング係数を不変のままにし、許容差の上側境界値よりも大きいレンダリング係数を上側境界値より小さいまたは等しい値に選択的にセットし、許容差の下側境界値よりも小さいレンダリング係数を下側境界値より大きいまたは等しい値に選択的にセットするように構成される。したがって、レンダリング係数を調整するために、平均値から大きく異なる非最適レンダリングパラメータの使用によって生じるアップミックス信号表現の過剰な歪みを回避する調整されたレンダリング係数を得ることを依然として可能とする、非常に簡単なメカニズムが確立される。

【0038】

好ましい実施形態において、パラメータ調整器は、それぞれの反復において、レンダリング係数を通じた平均からの最大偏差を含むレンダリング係数のそれぞれの1つを反復的に選択し、レンダリング係数の選択された1つを、レンダリング係数を通じた平均の近くに持ってくるように構成される。したがって、レンダリング係数を通じた平均に従って決定される許容差の外側にあるレンダリングパラメータは、反復的に許容差内に持ってこられる。このように、レンダリングパラメータは、非最適レンダリングパラメータの使用によって生じるアップミックス信号表現の歪みが、通常は低減されるように（少なくとも、最適レンダリングパラメータから、予め定められた偏差より大きく偏移している入力レンダリングパラメータに対して）、平均値に従って調整される。

【0039】

好ましい実施形態において、パラメータ調整器は、レンダリング係数のそれぞれの1つの反復的な選択と、選択された1つのレンダリング係数の反復的な修正を、全てのレンダリングパラメータが適用可能な許容差の範囲内にあるように調整されるまで繰り返すように構成される。したがって、アップミックス信号表現における聞き取れる歪みが十分小さく保たれることが確保される。

【0040】

好ましい実施形態において、装置は、ダウンミックス信号表現の1つ以上のチャンネルの、アップミックス信号表現の1つ以上のチャンネルへのマッピングを記述する1つ以上の変換符号化係数を受信するように構成される。この場合、装置は、調整されたパラメータとして、1つ以上の調整された変換符号化係数を提供するように構成される。本発明に係るこの実施形態は、変換符号化係数の平均値からの大きな偏差が通常は聞き取れる歪みを生じるので、変換符号化パラメータは、平均値に従う調整に対しても適切であるという発見に基づいている。したがって、平均値に従う変換符号化パラメータの調整または制限によって、非最適変換符号化パラメータの使用によって生じるアップミックス信号表現の歪みを、（少なくとも、最適変換符号化パラメータから予め定められた偏差より大きく偏移している入力変換符号化パラメータに対して）低減することが可能である。

【0041】

好ましい実施形態において、パラメータ調整器は、入力パラメータとして、変換符号化係数（変換符号化パラメータとしても示される）の時間シーケンスを受信するように構成される。この場合、パラメータ調整器は、複数の変換符号化係数に従って時間平均（temporal mean）（temporal averageとしても示される）を演算するように構成される。また、パラメータ調整器は、調整された変換符号化係数の時間平均からの偏差が限定されるように、調整された変換符号化係数を提供するように構成される。あらためて、非最適変換符号化係数の使用によって生じるアップミックス信号表現の過剰な聞き取れる歪みを回避する簡単なメカニズムが構築される。

【0042】

好ましい実施形態において、パラメータ調整器は、時間平均（それは平均値を構成する）に従って決定される許容差の範囲内にある変換符号化係数を不変のままにするように構成される。また、パラメータ調整器は、許容差の上側境界値よりも大きい変換符号化係数を、許容差の上側境界値よりも小さいまたは等しい値に選択的にセットし、許容差の下側

10

20

30

40

50

境界値よりも小さい変換符号化係数を、許容差の下側境界値よりも大きいまたは等しい値に選択的にセットするように構成される。したがって、変換符号化係数は、非最適変換符号化の使用によって生じるアップミックス信号表現の歪みを、少なくとも最適変換符号化係数から予め定められた偏差よりも大きく偏差している変換符号化係数に対して低減することを可能とする明確に定められた許容差内に持ってくるることができる。許容差は、時間平均が用いられるので、適応的な方法で選択される。このコンセプトは、変換符号化係数の大きな時間変化は、通常は聞き取れる歪みをもたらす、それ故にある程度に制限しなければならないという発見に基づいている。

【0043】

好ましい実施形態において、パラメータ調整器は、時間平均を、一連の変換符号化係数の再帰的ローパスフィルタリングを用いて算出するように構成される。このコンセプトは、変換符号化係数の長期の進化を考慮に入れた非常に明確に定められた時間平均をもたらすことを示している。また、一連の変換符号化係数のこのような再帰的ローパスフィルタリングは、少ない計算労力とメモリ要求条件を低減することを助ける記憶労力によって遂行することができることが分かっている。特に、長期の期間に対する変換符号化係数の履歴を記憶することなく、意味のある時間平均を取得することが可能である。

【0044】

好ましい実施形態において、パラメータ調整器は、調整されたパラメータの所定の1つが、複数の入力パラメータの平均値と1つ以上の許容差パラメータに従って境界が定められる許容差の範囲内にあるように、そして、入力パラメータと、対応する調整されたパラメータとの偏差が最小化されるまたは予め定められた最大許容範囲内に保持されるように、1つ以上の調整されたパラメータの所定の1つを提供するように構成される。良好な聴覚インプレッションをもたらす調整されたパラメータは、入力パラメータと、対応する調整されたパラメータとの過度に大きな差異を回避する目的を考慮に入れながら、調整されたパラメータを許容差に限定することによって、取得することができることが分かっている。したがって、非最適パラメータの使用によって生じるアップミックス信号表現の歪みを、入力パラメータによって定められる所望の聴覚設定を不必要に妥協することなく低減することができる。

【0045】

好ましい実施形態において、パラメータ調整器は、入力パラメータの調整されたバージョンを取得するために、複数の入力パラメータ値の平均値に従って境界が定められる許容差の外側にあることがわかった入力パラメータを、許容差の上側境界値または下側境界値に選択的にセットするように構成される。

【0046】

他の好ましい実施形態において、パラメータ調整器は、(平均値に従って境界が定められる)許容差の外側にある入力パラメータを、許容差内に反復的に持ってくるために、それぞれの反復において、平均値からの最大偏差を含む入力パラメータのそれぞれ1つを反復的に選択し、入力パラメータの選択された1つを平均値の近くに持ってくるように構成される。

【0047】

好ましい実施形態において、パラメータ調整器は、入力パラメータの選択された1つを平均値の近くに持ってくるために使用されるステップサイズを、入力パラメータの選択された1つと平均値の差異の予め定められた分数になるように選択するように構成される。

【0048】

本発明に係る他の実施形態は、ダウンミックス信号表現とパラメトリックサイド情報に基づいてアップミックス信号表現を提供する装置を構築する。前記装置は、前に述べられたような、1つ以上の入力パラメータに基づいて1つ以上の調整されたパラメータを提供する装置を備える。アップミックス信号表現を提供する装置は、また、ダウンミックス信号表現とパラメトリックサイド情報に基づいてアップミックス信号表現を取得するように構成された信号処理器を備える。1つ以上の調整されたパラメータを提供する装置は、信

10

20

30

40

50

号処理器の1つ以上の処理パラメータの、例えば、信号処理器に入力されるレンダリングパラメータの、または、アップミックス信号表現を取得するために、信号処理器において演算され、信号処理器によって適用される変換符号化パラメータの、調整されたバージョンを提供するように構成される。

【0049】

この実施形態は、信号処理器によって適用され、信号処理器に入力されるかまたはさらに信号処理器において算出されるかのいずれかであり、平均値に基づく上述のパラメータ調整から利益を得ることができる多数のパラメータがあるという発見に基づいている。信号処理器は、通常は、パラメータのセット（例えば、異なるオーディオオブジェクトに係るレンダリング係数のセット、または時間において異なるインスタンスに係る変換符号化係数のセット）が良くバランスしている場合に、そのような値のセットの個々の値が平均値からの過度に大きい偏差を含まないように、小さい歪みで、良い品質のアップミックス信号表現を提供することが分かっている。このように、1つ以上の調整されたパラメータを提供する装置を、アップミックス信号表現を提供する装置と組み合わせて適用することによって、発明コンセプトの利益を実現することができる。

【0050】

好ましい実施形態において、信号処理器は、オーディオオブジェクトの、アップミックス信号表現の1つ以上のチャンネルに対する貢献度を記述する調整されたレンダリング係数に従ってアップミックス信号表現を提供するように構成される。1つ以上の調整されたパラメータを提供する装置は、入力パラメータとして、複数のユーザ指定のレンダリングパラメータを受信し、それに基づいて、信号処理器による使用のために（好ましくは信号処理器に）、1つ以上の調整されたレンダリングパラメータを提供するように構成される。1つ以上の調整されたパラメータを提供する装置を用いて取得することができる良くバランスしたレンダリングパラメータは、通常は良い聴覚インプレッションに結果としてなることが分かっている。

【0051】

他の実施形態において、1つ以上の調整されたパラメータを提供する装置は、1つ以上の入力パラメータとして、混合マトリクスの1つ以上の混合マトリクス要素を受信し、それに基づいて、信号処理器による使用のために、混合マトリクスの1つ以上の調整された混合マトリクス要素を提供するように構成される。この場合、信号処理器は、ダウンミックス信号表現の1つ以上のオーディオチャンネル信号（例えば、時間ドメイン表現の形でまたは時間周波数ドメイン表現の形で表された）の、アップミックス信号表現の1つ以上のオーディオチャンネル信号上へのマッピングを記述する混合マトリクスの調整された混合マトリクス要素に従って、アップミックス信号表現を提供するように構成される。混合マトリクス要素は、また、例えば、混合マトリクス要素の時間的变化が制限されているという点で、平均値によく適合しなければならないことが分かっている。

【0052】

本発明に係る他の実施形態において、オーディオ処理器は、MPEGサラウンド任意ダウンミックスゲイン値を取得するように構成される。この場合、1つ以上の調整されたパラメータを提供する装置は、入力パラメータとして、複数の任意ダウンミックスゲイン値を受信し、複数の調整された任意ダウンミックスゲインを提供するように構成される。任意ダウンミックスゲイン値に対する調整されたパラメータを提供する装置のアプリケーションは、また、良好な聴覚インプレッションに結果としてなり、聞き取れる歪みを制限することを可能にすることが分かっている。

【0053】

本発明に係る更なる実施形態は、1つ以上の調整されたパラメータを提供する方法およびコンピュータプログラムを構築する。前記実施形態は、上述の装置と同じ知見に基づき、発明の装置に関して本願明細書において述べられた構成および機能のいずれかによって拡張することができる。

【図面の簡単な説明】

【 0 0 5 4 】

【図 1】本発明の実施形態に係る 1 つ以上の調整されたパラメータを提供する装置の概略ブロック図を示す。

【図 2】本発明の実施形態に係るアップミックス信号表現を提供する装置の概略ブロック図を示す。

【図 3】本発明の他の実施形態に係るアップミックス信号表現を提供する装置の概略ブロック図を示す。

【図 4】間接制御および直接制御を用いたパラメータ制限スキーム概略表現を示す。

【図 5 a】リスニングテスト条件を表すテーブルを示す。

【図 5 b】リスニングテストのオーディオ項目を表すテーブルを示す。

【図 6】テストされた極端なレンダリング条件を表すテーブルを示す。

【図 7】異なるパラメータ制限スキーム (P L S) に対する M U S H R A リスニングテスト結果のグラフィック表現を示す。

【図 8】参照用 M P E G S A O C システムの概略ブロック図を示す。

【図 9 a】分離したデコーダおよび混合器を用いた参照用 S A O C システムの概略ブロック図を示す。

【図 9 b】統合したデコーダおよび混合器を用いた参照用 S A O C システムの概略ブロック図を示す。

【図 9 c】S A O C M P E G トランスコーダを用いた参照用 S A O C システムの概略ブロック図を示す。

【図 1 0】どの変換符号化係数が提案されたパラメータ制限スキームによって修正することができるかを記述するテーブルを示す。

【発明を実施するための形態】

【 0 0 5 5 】

1 . 図 1 に係る 1 つ以上の調整されたパラメータを提供する装置

【 0 0 5 6 】

以下に、ダウンミックス信号表現と、ダウンミックス信号表現に関するパラメトリックサイド情報に基づくアップミックス信号表現の提供に対して、1 つ以上の調整されたパラメータを提供する装置が記載される。図 1 は、このような装置 1 0 0 の概略ブロック図である。

【 0 0 5 7 】

装置 1 0 0 は、1 つ以上の入力パラメータ 1 1 0 を受信し、それに基づいて、1 つ以上の調整されたパラメータ 1 2 0 を提供するように構成される。装置 1 0 0 は、1 つ以上の入力パラメータ 1 1 0 を受信し、それに基づいて、1 つ以上の調整されたパラメータ 1 2 0 を提供するように構成された、パラメータ調整器 1 3 0 を備える。パラメータ調整器 1 3 0 は、非最適パラメータ (例えば、1 つ以上の入力パラメータ 1 1 0) の使用によって生じるアップミックス信号表現の歪みが、少なくとも最適パラメータから予め定められた偏差以上偏移している入力パラメータ (例えば、入力パラメータ 1 1 0) に対して低減されるように、複数の入力パラメータ値の平均値 1 3 2 に従って、1 つ以上の調整されたパラメータ 1 2 0 を提供するように構成される。例えば、パラメータ調整器 1 3 0 は、1 つ以上の調整されたパラメータ 1 2 0 が、1 つ以上の入力パラメータ 1 1 0 よりも、最適パラメータ (それは、歪みのないアップミックス信号表現に結果としてなる) に「近い」 (より小さい歪みを生じるという意味において) という効果を有することができる。

【 0 0 5 8 】

この目的のため、パラメータ調整器 1 3 0 は、平均値演算を実施し、関連する入力パラメータ 1 1 0 のセット (例えば、共通の時間インターバルに関する入力パラメータ、または異なる時間インスタンスに関する同じパラメータタイプの入力パラメータ) の平均値 1 3 2 (例えば、時間平均またはオブジェクト間平均として) を取得する。装置 1 0 0 の動作に関して、平均値 1 3 2 はパラメータを調整するために意味のある量であることが分かっているので、1 つ以上の入力パラメータ 1 1 0 に基づく 1 つ以上の調整されたパラ

10

20

30

40

50

メータ 1 2 0 の提供が平均値 1 3 2 に従ってなされる点に注意しなければならない。特に、適度なパラメータ（平均値に関して）は、通常は適度な歪みをもたらすことが分かっている。

【 0 0 5 9 】

更なる詳細が引き続いて記載される。

【 0 0 6 0 】

2 . 図 2 に係るアップミックス信号表現を提供する装置

【 0 0 6 1 】

以下に、図 2 に係るアップミックス信号表現を提供する装置が記載される。図 2 は、オーディオ信号デコーダとみなすことができる装置 2 0 0 の概略ブロック図を示す。例えば、装置 2 0 0 は、S A O C デコーダまたは S A O C トランスコーダの機能を備えることができる。

【 0 0 6 2 】

装置 2 0 0 は、ダウンミックス信号表現 2 1 0 とパラメトリックサイド情報 2 1 2 を受信するように構成される。また、装置 2 0 0 は、ユーザ指定のレンダリングパラメータ 2 1 4 を受信するように構成される。装置は、アップミックス信号表現 2 2 0 を提供するように構成される。

【 0 0 6 3 】

ダウンミックス信号表現 2 1 0 は、例えば、1 チャンネルのオーディオ信号または 2 チャンネルのオーディオ信号の表現とすることができる。ダウンミックス信号表現 2 1 0 は、例えば、時間ドメイン表現または符号化された表現とすることができる。いくつかの実施形態では、ダウンミックス信号表現 2 1 0 は、ダウンミックス信号表現 2 1 0 の 1 つ以上のチャンネルがスペクトル値の引き続くセットによって表現される、時間 周波数ドメイン表現とすることができる。

【 0 0 6 4 】

アップミックス信号表現 2 2 0 は、例えば、時間ドメイン表現または時間 周波数ドメイン表現の形の、個々のオーディオチャンネルの表現とすることができる。あるいは、アップミックス信号表現 2 2 0 は、ダウンミックス信号表現と、チャンネル関連サイド情報、例えば、M P E G サラウンドサイド情報の両方を含む符号化された表現とすることができる。

【 0 0 6 5 】

ユーザ指定のレンダリングパラメータ 2 1 4 は、複数のオーディオオブジェクトの、アップミックス信号表現 2 2 0 の 1 つ以上のチャンネルに対する所望の貢献度を記述するレンダリングマトリクスエントリーの形で提供することができる。あるいは、ユーザ指定のレンダリングパラメータ 2 1 4 は、例えば、オーディオオブジェクトの所望のレンダリング位置とレンダリング量を特定する、他のいかなる適当な形でも提供することができる。

【 0 0 6 6 】

装置 2 0 0 は、ダウンミックス信号表現 2 1 0 とパラメトリックサイド情報 2 1 2 に基づいてアップミックス信号表現 2 2 0 を提供するように構成された、信号処理器 2 3 0 を備える。信号処理器 2 3 0 は、ダウンミックス信号表現 2 1 0 に基づいてアップミックス信号表現 2 2 0 を提供するために、再混合機能 2 3 2 を備える。例えば、再混合機能 2 3 2 は、アップミックス信号表現 2 2 0 の 1 つ以上のチャンネルを取得するために、ダウンミックス信号表現 2 1 2 の複数のチャンネルを線形に結合するように構成することができる。この再混合において、ダウンミックス信号表現 2 1 0 のチャンネルの、アップミックス信号表現 2 2 0 のチャンネルに対する貢献度は、混合マトリクス G の混合マトリクス要素によって決定することができ、混合マトリクス G の第 1 の次元（例えば、列の数）はアップミックス信号表現 2 2 0 のチャンネル数によって決定することができ、混合マトリクス G の第 2 の次元（例えば、行の数）はダウンミックス信号表現 2 1 0 のチャンネル数で決定することができる。

【 0 0 6 7 】

例えば、再混合プロセス 2 3 2 は、ダウンミックス信号表現 2 1 0 の 1 つ以上のチャンネルのスペクトル値を含む 1 つ以上のベクトルを、混合マトリクス G と掛けることによって、アップミックス信号表現 2 2 0 の 1 つ以上のチャンネルに關係するスペクトル値を含む 1 つ以上のベクトルを提供するために用いることができる。

【 0 0 6 8 】

信号処理器 2 3 0 は、また、混合マトリクス G (または同様に、その要素) を提供する、混合パラメータ演算 2 3 6 を備えることができる。混合マトリクス要素は、混合パラメータ演算 2 3 6 によって、パラメトリックサイド情報 2 1 2 と修正されたレンダリングパラメータ 2 5 2 に従って決定される。混合マトリクス G の混合マトリクス要素は、例えば、アップミックス信号表現 2 2 0 の 1 つ以上のチャンネルが、ダウンミックス信号表現 2 1 0 の 1 つ以上のチャンネルによって表されるオーディオオブジェクトを記述するように、修正されたレンダリングパラメータ 2 5 2 によって提供される。この目的のため、例えば、オブジェクトレベル差情報 O L D、オブジェクト間相關情報 I O C、ダウンミックスゲイン情報 D M G および (オプションとして) ダウンミックスチャンネルレベル差情報 D C L D を含むパラメトリックサイド情報 2 1 2 は、混合パラメータ演算 2 3 6 によって評価される。オブジェクトレベル差情報は、例えば、周波数バンドワイズに、複数のオーディオオブジェクト間のレベル差を記述することができる。同様に、オブジェクト間相關情報は、例えば、周波数バンドワイズに、複数のオーディオオブジェクト間の相關を記述することができる。ダウンミックスゲイン情報と (オプションの) ダウンミックスチャンネルレベル差情報は、オーディオオブジェクト信号を複数のオーディオオブジェクトからダウンミックス信号表現の 1 つ以上のチャンネルに結合するために実行され、ダウンミックス信号表現 2 1 0 のチャンネルよりも通常は多いオーディオオブジェクトが存在するダウンミックスを記述することができる。

【 0 0 6 9 】

したがって、混合パラメータ演算 2 3 6 は、パラメトリックサイド情報 2 1 2 と修正されたレンダリングパラメータ 2 5 2 に基づいて期待された統計的特性を含むアップミックス信号表現 2 2 0 を取得するために、混合マトリクス要素がどのように選択されなければならないかについて評価することができる。

【 0 0 7 0 】

信号処理器 2 3 0 は、パラメトリックサイド情報 2 1 2 を受信し、修正されたサイド情報と、再混合プロセスによって提供される関連する再混合されたダウンミックス信号表現が所望のオーディオシーンを記述するように、修正されたサイド情報 (例えば、M P E G サラウンドサイド情報) を提供するように構成された、サイド情報修正あるいはサイド情報変換 2 4 0 をオプションとして備えることができる。

【 0 0 7 1 】

要約すると、信号処理器 2 3 0 は、例えば、ダウンミックス信号表現 2 1 0 が 1 つ以上のダウンミックス信号 8 1 2 の役割を持ち、パラメトリックサイド情報 2 1 2 がサイド情報 8 1 4 の役割を持ち、アップミックス信号表現 2 2 0 が出力チャンネル信号 $\hat{y}_1 \sim \hat{y}_M$ に相当する、S A O C デコーダ 8 2 0 の機能を満たすことができる。

【 0 0 7 2 】

あるいは、信号処理器 2 3 0 は、分離したデコーダおよび混合器 9 2 0 の機能を備えることができ、ダウンミックス信号表現 2 1 0 は 1 つ以上のダウンミックス信号の役割を持つことができ、パラメトリックサイド情報 2 1 2 はオブジェクトメタデータの役割を持つことができ、アップミックス信号表現 2 2 0 は 1 つ以上の出力チャンネル信号 9 2 8 の役割を持つことができる。

【 0 0 7 3 】

あるいは、信号処理器 2 3 0 は、統合化されたデコーダおよび混合器 9 5 0 の機能を備えることができ、ダウンミックス信号表現 2 1 0 は 1 つ以上のダウンミックス信号の役割を持つことができ、パラメトリックサイド情報 2 1 2 はオブジェクトメタデータの役割を持つことができ、アップミックス信号表現 2 2 0 は 1 つ以上の出力チャンネル信号 9 5 8

10

20

30

40

50

の役割を持つことができる。

【0074】

あるいは、信号処理器230は、S A O C M P E Gサラウンドトランスコーダ980の機能を備えることができ、ダウンミックス信号表現210は1つ以上のダウンミックス信号の役割を持つことができ、パラメトリックサイド情報212はオブジェクトメタデータの役割を持つことができ、アップミックス信号表現はM P E Gサラウンドビットストリーム984と組み合わせられるときに1つ以上のダウンミックス信号988に相当することができる。

【0075】

いずれにせよ、修正されたレンダリングパラメータ252は、ユーザ対話/制御情報822またはレンダリング情報の役割を持つことができる。

【0076】

装置200は、また、調整されたレンダリングパラメータを提供する装置250を備える。調整されたレンダリングパラメータを提供する装置250は、ユーザ指定のレンダリングパラメータ214を受信し、それに基づいて、修正されたレンダリングパラメータ252を提供する。装置250は、通常は、異なるオーディオオブジェクトに関する複数のユーザ指定のレンダリングパラメータを通じた平均値を算出し、平均値を取得するように構成される。また、装置250は、平均値に従ってレンダリングパラメータ制限を実行し、ユーザ指定のレンダリングパラメータ214を制限することによって、修正されたレンダリングパラメータ252を取得するように構成される。修正されたレンダリングパラメータ252が制限される許容差は、ユーザ指定のレンダリングパラメータ214の1つ以上が平均値から大きな偏差を含む場合であっても、修正されたレンダリングパラメータ252の、平均値からの大きな偏差が回避されるように、通常は平均値に従って決定される。このように、異なるオーディオオブジェクトに関するレンダリングパラメータ間の大きな差異は聞き取れるアーチファクトに結果としてなるが、制限されたオブジェクト間偏差を含む修正されたレンダリングパラメータ252は低歪のアップミックス信号表現に結果としてなるので、アップミックス信号表現220内の過剰な歪みは、通常は回避される。

【0077】

ここで、調整されたレンダリング係数を提供する装置250は、1つ以上の調整されたパラメータを提供する装置100と同じ全体機能を備えることができ、ユーザ指定のレンダリングパラメータ214は1つ以上の入力パラメータ110の役割を持つことができ、調整されたレンダリングパラメータ252は1つ以上の調整されたパラメータ120の役割を持つことができる点に注意しなければならない。

【0078】

修正されたレンダリングパラメータ252の提供に関する詳細は、図4を参照して後述される。

【0079】

3. 図3に係るアップミックス信号表現を提供する装置

【0080】

以下に、本発明の他の実施形態に係るアップミックス信号表現を提供する装置が、そのような装置300の概略ブロック図を示す図3を参照して記述される。

【0081】

本願明細書において、同一または等価である信号を記載するために同一の参照番号が用いられるように、装置300は、通常は、装置200と同じタイプの入力信号を受信し、同じタイプの出力信号を提供する。要約すると、装置300は、ダウンミックス信号表現210、パラメトリックサイド情報212、およびユーザ指定のレンダリングパラメータ214を受信し、装置300は、それに基づいて、アップミックス信号表現220を提供する。

【0082】

装置 300 は、信号処理器 230 に対して機能において実質的に等価とすることができる、信号処理器 330 を備える。信号処理器 330 は、ダウンミックス信号表現に基づいて再混合されたオーディオチャンネル信号を提供するという点で、信号処理器 230 の再混合機能 232 と同一である、再混合機能 332 を備える。しかしながら、再混合 332 は、混合パラメータ演算から直接取得される混合マトリクスよりむしろ、調整された混合マトリクスを用いる。

【0083】

信号処理器 330 は、また、信号処理器 230 の混合パラメータ演算 236 に対して機能において同一とすることができる、混合パラメータ演算 336 を備える。したがって、混合パラメータ演算 336 は、パラメトリックサイド情報 212 とユーザ指定のレンダリングパラメータ 214 を受信し、それに基づいて、混合マトリクス G (または、同等に、337 によって示される混合マトリクス G の混合マトリクス要素) を提供する。

10

【0084】

信号処理器 330 は、オプションとして、また、サイド情報修正 240 と機能が、同一であるサイド情報修正 338 を備える。

【0085】

加えて、装置 300 は、調整された混合マトリクス要素を提供する装置 350 を備える。装置 350 は、信号処理器 330 の一部であってもよく、一部でなくてもよい。装置 350 は、混合パラメータ演算 336 によって提供される混合マトリクス 337, G (または、同等に、その混合マトリクス要素) を受信し、それに基づいて、調整された混合マトリクス 352, G' (または、同等に、その調整された混合マトリクス要素) を提供するように構成される。例えば、周波数バンド毎に、そしてオーディオフレーム毎に、1 セットの混合マトリクス要素と、1 セットの調整された混合マトリクス要素を提供することができる。言い換えれば、混合マトリクス G と修正された混合マトリクス G' は、フレームワイズの処理が選択された場合、ダウンミックス信号表現 210 のオーディオフレーム毎に一度更新することができる。しかしながら、更新インターバルは、場合によって異なってもよい。また、異なる周波数バンドに対して、多重の混合マトリクスと調整された混合マトリクス G, G' がある必要はない。

20

【0086】

しかしながら、装置 350 は、混合パラメータ演算 336 によって提供される混合マトリクス 337 の混合マトリクス要素に基づいて、調整された混合マトリクス 352 の調整された混合マトリクス要素を提供するように構成される。例えば、処理は、所定の混合マトリクス位置の一連の調整された混合マトリクス要素が、同じ混合マトリクス位置での混合マトリクス 337 の一連の混合マトリクス要素に依存するが、異なる混合マトリクス位置での混合マトリクス要素から独立することができるように、混合マトリクス (または調整された混合マトリクス) の位置毎に個別に実行することができる。

30

【0087】

調整された混合マトリクス要素を提供する装置 350 は、混合マトリクス 337 に基づいて演算される 1 つ以上の平均値 (例えば、1 つ以上のマトリクス位置個々の平均値) に従って、調整された混合マトリクス 352 の 1 つ以上の調整された混合マトリクス要素を提供するように構成される。調整された混合マトリクス 352 の調整された混合マトリクス要素を提供する装置 350 は、好ましくは、所定の混合マトリクス位置での混合マトリクス要素の時間上の平均値を算出するように構成される。このように、所定の混合マトリクス位置に対して、平均値 (好ましくは、しかしながら必然的ではなく、例えば、浮動平均または準無限インパルス応答平均値または再帰的ローパスフィルタリングまたは時間平均に対してよく知られた類似する数値演算によって得られる平均値のような時間的平均値) を、所定の混合マトリクス位置の一連の混合マトリクス要素に基づいて演算することができる。有限インパルス応答平均値または (準) 無限インパルス応答平均値 (例えば、再帰的ローパスフィルタリングまたは時間平均に対してよく知られた類似する数値演算を用いて取得された) とすることができるそのような平均値 (average value) (mean value

40

50

としても示される)を取得するために、例えば、ダウンミックス信号表現 2 1 0 の所定のチャンネルの、混合マトリクス要素が複数のオーディオフレームに関係するアップミックス信号表現 2 2 0 の所定のチャンネルへの貢献度を記述する一連の混合マトリクス要素を用いることができる。(ダウンミックス信号表現 2 1 0 の所定のチャンネルの、アップミックス信号表現 2 2 0 の所定のチャンネルへの貢献度を記述する)所定の混合マトリクス位置の現在の調整された混合マトリクス要素は、装置 3 5 0 によって、所定の混合マトリクス位置に関係する平均値に従って定められる許容差に制限することができる。

【0088】

したがって、調整された混合マトリクス要素は、例えば、同じ混合マトリクス位置での前の混合マトリクス要素の平均(有限インパルス応答平均または無限インパルス応答平均)で決定される許容差に限定されるので、混合マトリクス要素の過剰な時間変動は回避される。調整された混合マトリクス 3 5 2 の調整された混合マトリクス要素のこのような限定は、通常は、少なくとも非最適なユーザ指定のレンダリングパラメータが最適なユーザ指定のレンダリングパラメータから予め定められた偏差より大きく偏移している場合に、非最適パラメータ(例えば、非最適なユーザ指定のレンダリングパラメータ)の使用によって生じるアップミックス信号 2 2 0 の歪みの制限をもたらすことが分かっている。

【0089】

ここで、調整された混合マトリクス要素を提供する装置 3 5 0 は、1つ以上の調整されたパラメータを提供する装置 1 0 0 と同じ全体の機能を備えることができ、混合マトリクス 3 3 7 の混合マトリクス要素は1つ以上の入力パラメータ 1 1 0 の役割を持つことができ、調整された混合マトリクス 3 5 2 の調整された混合マトリクス要素は1つ以上の調整されたパラメータ 1 2 0 の役割を持つことができる点に注意しなければならない。

【0090】

4. 図4に係るパラメータ制限スキーム

【0091】

以下に、本発明に係るパラメータ制限スキームが、そのようなパラメータ制限スキームの概略表現を示す図4を参照して記載される。

【0092】

図4は、パラメータ制限スキームのアプリケーションを、S A O C デコーダ 4 1 0 と組合せて示す。しかしながら、パラメータ制限スキームは、例えば、S A O C トランスコーダのような、オーディオデコーダまたはオーディオトランスコーダの異なるタイプと組合せて適用することができる。

【0093】

S A O C デコーダ 4 1 0 は、ダウンミックス 4 2 0 と S A O C ビットストリーム 4 2 2 を受信する。また、S A O C デコーダは、1つ以上の出力チャンネル 4 3 0 a ~ 4 3 0 M を提供する

【0094】

(a)で示される第1の実施において、パラメータ制限スキーム 4 4 0 は、間接制御を実施する。パラメータ制限スキーム 4 4 0 は、入力レンダリングマトリクス R、例えば、ユーザ指定のレンダリングマトリクスを受信し、それに基づいて、調整されたレンダリングマトリクス \tilde{R} を S A O C デコーダに提供する。この場合、S A O C デコーダは、上述の

ように、混合マトリクス G の導出のために調整されたレンダリングマトリクス \tilde{R} を用いる。パラメータ制限スキーム 4 4 0 は、また、許容差の境界を決定することができるパラメータ Λ_{R-} 、 Λ_{R+} を受信することができる。

【0095】

代わりにまたは加えて、第2のパラメータ制限スキーム450を適用することができる。第2のパラメータ制限スキームは、変換符号化パラメータ T を受信し、それに基づいて、調整された変換符号化パラメータ \tilde{T} を提供する。変換符号化パラメータ T は、S A O Cデコーダ410において演算することができ、調整された変換符号化パラメータ \tilde{T} は、S A O Cデコーダ410によって適用することができる。例えば、変換符号化パラメータ T は、前に述べたように、混合マトリクス G の混合マトリクス要素に相当することができ、調整された変換符号化パラメータ \tilde{T} は、調整された混合マトリクス G の調整された混合マトリクス要素に相当することができる。

10

【0096】

パラメータ制限スキーム450は、許容差の境界を決定することができる1つ以上のパラメータ T_- , T_+ を受信することができる。

【0097】

4.1 概要

【0098】

以下に、歪み制御のためのパラメータ制限スキームを通じて、概要が与えられる。

【0099】

一般的なS A O C処理は、時間/周波数選択的方法で遂行され、以下に記載される。

20

【0100】

S A O Cエンコーダは、いくつかの入力オーディオオブジェクト信号の音響心理学的特性（例えば、オブジェクトのパワー関係および相関）を抽出し、次に、それらを複合されたモノラルまたはステレオチャンネルにダウンミックスする（それは、例えば、ダウンミックス信号表現として示すことができる）。このダウンミックス信号と抽出されたサイド情報は、周知の知覚オーディオコードを用いて、圧縮されたフォーマットで送信される（または記憶される）。受信側では、S A O Cデコーダは、概念的に、送信されたサイド情報（例えば、オブジェクトレベル差情報O L D、オブジェクト間相関情報I O C、ダウンミックスゲイン情報D M Gおよびダウンミックスチャンネルレベル差情報D C L D）を用いて、オリジナルのオブジェクト信号（すなわち、分離したダウンミックスオブジェクト）を復元しようと試みる。これらの近似されたオブジェクト信号は、次に、レンダリングマトリクス（通常は、異なるオーディオオブジェクトの、アップミックス信号表現の異なるチャンネルへの貢献度を記述する）を用いて、目標シーンに混合される。レンダリングマトリクスは、各送信されたオーディオオブジェクトとアップミックスセットアップスピーカに対して特定された相対レンダリング係数 $R C$ （またはオブジェクトゲイン）から構成される。これらのオブジェクトゲインは、全ての分離された/レンダーされたオブジェクトの空間位置を決定する。事実上、分離と混合は単一の複合された処理ステップにおいて実行され、それは計算量の莫大な低減に結果としてなるので、オブジェクト信号の分離はめったに実行されない（または更に決して実行されない）。単一の複合された処理ステップは、例えば、オブジェクト分離と分離されたオブジェクトの混合の組合せを記述する変換符号化係数を用いて実行することができる。

30

40

【0101】

このスキームは、伝送ビットレート（それは、多数の個別のオブジェクトオーディオ信号の代わりに、1つまたは2つのダウンミックスチャンネルと、加えていくつかのサイド情報を送信することを必要とするだけである）と、計算量（処理複雑度は、オーディオオブジェクト数よりもむしろ出力チャンネル数に主に関係する）の両方に関して、大いに効率的であることが分かっている。

【0102】

S A O Cデコーダは、オブジェクトゲインと他のサイド情報を、レンダーされた出力オーディオシーン（または、更なる復号化演算、例えば、通常は多重チャンネルM P E Gサ

50

ラウンドレンダリングに対して前処理されたダウンミックス信号)に対して、対応する信号をつくるためにダウンミックス信号に適応される変換符号化係数(TC)に、直接的に変換(パラメトリックレベルで)する。

【0103】

レンダーされた出力シーンの主観的に知覚されたオーディオ品質は、特許文献1に記述されるように、歪み制御尺度あるいはDCMのアプリケーションによって改善することができる。この改善は、目標レンダリング設定の適度な動的修正を受け入れる代価で達成することができる。レンダリング情報の修正は、特定の環境下で不自然な音響呈色と時間変動アーチファクトに結果としてなる可能性がある時間および周波数可変の性質を有する。

10

【0104】

特許文献1に記載された歪み制御尺度(DCM)の変形例として、本発明に係る実施形態は、オーディオアーチファクト(音響呈色、時間変動、その他)の低減にフォーカスし、同時に自然な音響品質を保持する、多数のパラメータ制限スキームを使用する。

【0105】

本願明細書に記載された提案されたパラメータ制限スキームのコンセプトは、音響心理学的モデルに基づく複雑なアルゴリズムを用いて算出される歪み尺度に基づいてレンダリング係数(RC)を調整することはしない。その代わりに、提案されたパラメータ制限スキームのコンセプトは、低い計算量と構成上の複雑度を示し、それ故にSAOC技術への統合化に対して魅力的である。にもかかわらず、それらは、また、お互いに補足することでより良好な全体の出力品質を達成するために、特許文献1に記載されたスキームと都合よく組合せることができる。

20

【0106】

全体のSAOCシステムの範囲内で、パラメータ制限スキームは、2つの方法でSAOCデコーダ処理チェーンに組み込むことができる。例えば、そのパラメータ制限スキームは、図4において変形例(a)として示されるように、レンダリング係数(RC)を制御することによってSAOC出力の間接的な(外部の)修正のためのフロントエンドに位置付けることができる。あるいは、固有の変換符号化係数(TC)は、図4において変形例(b)として示されるように、係数がダウンミックス信号に適用され、出力アップミックスチャンネル信号を生成する前に、SAOCデコーダのバックエンドにおいて直接的に(内部的に)修正される。

30

【0107】

4.2 間接制御

【0108】

以下に、間接制御のコンセプトが更に詳細に述べられる。

【0109】

間接制御法の基礎をなす前提は、歪みレベルと、RCのオブジェクト平均化された値からの偏差との関係を考慮する。これは、RCによって、他のオブジェクトに関する特定のオブジェクトに、特別な減衰/ブーストが適用されればされるほど、SAOCデコーダ/トランスコーダによって、伝送されたダウンミックス信号の積極的な修正が実行されるという知見に基づいている。言い換えれば、「オブジェクトゲイン」値の偏差がお互いと比較して高ければ高いほど、容認できない歪みが起こる機会が高い(同一のダウンミックス係数と仮定して)。これは、RCの、全てのオブジェクト全体のRCの平均(例えば、平均レンダリング値)からの偏差を調べることによって、テストすることができる。【0110】

40

引き続く記述は、一般性の喪失なしに、全てのオブジェクトに対して単一のダウンミックスゲインを有するモノラルダウンミックスを考慮する構成に基づいている。(異なるおよび/または動的なオブジェクトゲインを有する)非自明なダウンミックスの場合、アルゴリズムは適切に修正することができる。加えて、RCは、表記を簡単にするため、周波

50

数不変であると仮定される。

【 0 1 1 1 】

オブジェクトインデックスを有する係数 $R(i)$ によって表されたユーザ指定のレンダリングシナリオに基づいて、P L S は、S A O C レンダリングエンジンによって実際に用いられる修正された R C 値 $\tilde{R}(i)$ を生成することによって、極端なレンダリング値を防止する。それらは、次の関数として導き出すことができる。

$$\tilde{R}(i) = F_R(R(i), \Lambda),$$

ここで、 Λ は、P L S 制御パラメータ（すなわち、閾値）である。P L S 制御パラメータは、許容差パラメータとみなすことができる。

10

【 0 1 1 2 】

レンダリング係数 $R(i)$ の、平均化されたレンダリング値 \bar{R} （例えば算術平均）からの偏差 $R_d(i)$ は、次のように取得することができる。

$$R_d(i) = \frac{R(i)}{\bar{R}},$$

ここで、

$$\bar{R} = \frac{1}{N_{ob}} \sum_{i=1}^{N_{ob}} R(i).$$

20

【 0 1 1 3 】

したがって、 $R_d(i)$ は、レンダリング係数 $R(i)$ と平均化されたレンダリング値 \bar{R} の比率である。平均化されたレンダリング値 \bar{R} は、レンダリング係数 $R(i)$ のオーディオオブジェクトインデックス i を有するオーディオオブジェクトを通じて平均化された平均値である。

30

【 0 1 1 4 】

制限された偏差 $\tilde{R}_d(i)$ は、次のような一定の許容差 Λ の範囲に限定される。

$$\tilde{R}_d(i) = \Lambda \quad \text{for} \quad R_d(i) > \Lambda,$$

$$\tilde{R}_d(i) = \frac{1}{\Lambda} \quad \text{for} \quad R_d(i) < \frac{1}{\Lambda}.$$

【 0 1 1 5 】

これは、例えば、特定の予め定められた値よりむしろ入力された R C から動的に演算される基準値、例えば \bar{R} と関連して実行される R C 制限演算に対応する点に注意すべきである。

40

【 0 1 1 6 】

記載された P L S アプローチに対して、最適な解法は、所定の R C $R(i)$ と修正された（制限された） $\tilde{R}(i)$ との差異が最小化される最小化問題として定式化することができる。

$$\|\tilde{R}(i) - R(i)\| \rightarrow \min.$$

【 0 1 1 7 】

50

以下に、調整されたパラメータとみなすことができる調整されたレンダリング係数 $\tilde{R}(i)$ を提供するいくつかのアルゴリズム的解法が記載される。

【 0 1 1 8 】

以下の2つのアルゴリズム的解法は、次式に示すように、許容範囲の外側にあるそれらのレンダリング値の偏差に基づいている。

$$R_{d,out}(i) = R_d(i) \quad \text{for} \quad R_d(i) > \Lambda, \text{ or } R_d(i) < \frac{1}{\Lambda}.$$

【 0 1 1 9 】

10

4 . 2 . 1 ワンステップ解法

【 0 1 2 0 】

次式によって許容範囲の外側の全てのレンダリング値を制限する簡単で速いワンステップ解法を用いることができる。

$$\tilde{R}(i) = \Lambda \bar{R} \quad \text{for} \quad R_{d,out}(i) > \Lambda,$$

$$\tilde{R}(i) = \frac{\bar{R}}{\Lambda} \quad \text{for} \quad R_{d,out}(i) < \frac{1}{\Lambda}.$$

【 0 1 2 1 】

20

対照的に、このようなレンダリング値 $\tilde{R}(i)$ に対して、許容範囲の内側のレンダリング値は、次のように、影響されないで残すことができる。

$$\tilde{R}(i) = R(i)$$

【 0 1 2 2 】

4 . 2 . 2 反復解法

【 0 1 2 3 】

関連する偏差 $R_{d,out}(i)$ を有する範囲外のレンダリング値が徐々に制限される他の直接的な方法を使用することができる。このアルゴリズムの各反復において、最大限度のレンダリング偏差 $R_{d,max}$ は、次のように定められる。

30

$$R_{d,max} = \max \{ R_{d,out}(i) \} \quad \text{for} \quad R_d > \Lambda,$$

$$R_{d,max} = \min \{ R_{d,out}(i) \} \quad \text{for} \quad R_d < \frac{1}{\Lambda}.$$

【 0 1 2 4 】

対応するレンダリング係数は、次のように限定される。

$$\tilde{R}(i) = (1-\lambda)R(i) + \lambda\bar{R}, \quad \lambda \in (0,1).$$

40

【 0 1 2 5 】

この処理は、全ての値が許容範囲の内側となるまで、または予め定められた反復回数によって実行することができる。

【 0 1 2 6 】

したがって、各反復において、偏差 $R_{d,out}(i_{\max})$ (例えば、平均値 \bar{R} からの) が、最大値 $R_{d,\max}$ をとるレンダリング係数 $R(i_{\max})$ が選択される。言い換えれば、それぞれの反復におけるレンダリング係数を通じた平均 \bar{R} からの最大偏差 (偏差値 $R_{d,out}$ に関して) を含むレンダリング係数 $R(i_{\max})$ が選択される。加えて、上述の $R(i)$ と \bar{R} の線形結合 (それは、 $i = i_{\max}$ に対して選択的に適用することができる) を用いて、選択されたレンダリング係数 $R(i_{\max})$ がレンダリング係数を通じた平均の近くに持ってこられる。反復手順の各ステップにおいて、異なるレンダリング係数が反復アルゴリズムの異なるステップにおいて修正されるように、平均値からの最大偏差を有するレンダリング係数の新規な選択を実行することができる。言い換えれば、 i_{\max} は、通常はいずれの反復においても更新される。また、平均値は、オプションとして、反復アルゴリズムのいずれのステップに対しても、前に修正されたレンダリング係数を考慮して再演算することができる。

10

【 0 1 2 7 】

4 . 3 直接制御

【 0 1 2 8 】

直接制御法の基礎をなす前提は、歪みレベルと、TCの時間平均された値からの偏差との関係を考慮する。これは、他のオブジェクトに関する特定のオブジェクトに対して、特別な減衰 / ブーストが適用されればされるほど、TCによって送信されたダウンミックス信号の積極的な修正が、SAOCデコーダ / トランスコーダによって実行されるという知見に基づいている。言い換えれば、TCの値が異常に大きい場合、SAOCアルゴリズムは、小さいパワーを有するオブジェクト信号を、大きなブーストを適用することによって、大きいパワーを有する他のオブジェクト信号によって支配される出力内に修正することを試みると結論づけることができる。逆にいえば、TCが異常に小さい場合、SAOCアルゴリズムは、大きいパワーを有するオブジェクト信号を、大きな減衰を適用することによって、小さいパワーを有する他のオブジェクト信号によって支配される出力内に修正することを試みると結論づけることができる。いずれの場合においても、SAOC出力において、容認できないほど低い信号品質を生じる高いリスクがある。このように、中心的なアイデアは、TCの、平均値からの大きな偏差を防止することである。

20

30

【 0 1 2 9 】

このPLSは、SAOC信号パラメータ (例えばOLD、IOC) への全ての従属と変換符号化 / 復号化プロセスの発見的要素を含むので、時間および周波数可変とみなすことができる。

【 0 1 3 0 】

引き続き記述は、一般性の喪失なしに、モノラルアップミックスを考慮する構成に基づいている。

【 0 1 3 1 】

周波数インデックス k を有するSAOC出力TC $T(k)$ に基づいて、PLSは、それら (例えば、許容差の外側の変換符号化係数) を、次に実際のSAOCレンダリングプロセスによって用いられる修正されたTC値と交換することによって、TCの極端な値を防止する。修正されたTC値 $\tilde{T}(k)$ は、次の関数によって導き出すことができる。

40

$$\tilde{T}(k) = F_T(T(k), \Lambda),$$

ここで、 Λ は、PLS制御パラメータ (すなわち、閾値) である。PLS制御パラメータは、許容差パラメータとみなすことができる。

【 0 1 3 2 】

50

T C は時間可変であるので、再帰的ローパスフィルタが平均を算出するために適用される。

$$\bar{T}_n(k) = \mu T_n(k) + (1 - \mu) \bar{T}_{n-1}(k).$$

【 0 1 3 3 】

平均 \bar{T} は、平均値とみなされ、再帰的ローパスフィルタリングのアプリケーションによって、個々の変換符号化値の重みが導入される。

【 0 1 3 4 】

ここで、n は、T C の時間インデックスを表し、 $\mu \in (0,1]$ は、平均化パラメータである

10

。修正された T C 値 $\tilde{T}(k)$ に対する許容範囲は、次のように定められる。

$$\frac{\bar{T}(k)}{\Lambda} \leq \tilde{T}(k) \leq \Lambda \bar{T}(k).$$

【 0 1 3 5 】

これは、特定の予め定義された値よりもむしろ T C から動的に演算される基準値に関連して実行される T C 制限演算に対応する点に注意すべきである。

【 0 1 3 6 】

20

記載された P L S アプローチに対して、最適な解法は、与えられた T C $T(k)$ と、修正された（制限された）T C $\tilde{T}(k)$ との差異が最小化される最小化問題として定式化することができる。

$$\|\tilde{T}(k) - T(k)\| \rightarrow \min.$$

【 0 1 3 7 】

以下に、この問題に対する可能な解法アルゴリズムが記載される。

【 0 1 3 8 】

30

4 . 3 . 1 解法アルゴリズム

【 0 1 3 9 】

修正された T C 値 $\tilde{T}(k)$ は、次のように取得することができる。

$$\tilde{T}(k) = \Lambda \bar{T}(k) \quad \text{for} \quad T(k) > \Lambda,$$

$$\tilde{T}(k) = \frac{\bar{T}(k)}{\Lambda} \quad \text{for} \quad T(k) < \frac{1}{\Lambda}.$$

【 0 1 4 0 】

40

4 . 3 . 2 変換符号化係数の例

【 0 1 4 1 】

上述の変換符号化係数に対するパラメータ制限スキームは、例えば、上で述べた S A O C デコーダおよびトランスコーダにおいて用いられる異なる変換符号化係数に適用することができる。

【 0 1 4 2 】

例えば、変換符号化係数に対するパラメータ制限スキームは、装置 300 の信号処理器 330 において用いられる混合マトリクス G のパラメータを制限するために適用することができる。この場合、マトリクス G の所定のマトリクス位置での混合マトリクス要素は、変換符号化係数 $T(k)$ (ここで、 k は周波数インデックスである) の代わりをすることができる。混合マトリクス G' の対応する混合マトリクス要素は、調整された変換符号化係数 $\tilde{T}(k)$ に対応することができる。変換符号化パラメータ制限スキームは、例えば、混合マトリクスの異なるマトリクス位置に個別に適用することができる。例えば、混合マトリクス G が混合マトリクス要素 g_{11} 、 g_{12} 、 g_{21} および g_{22} を含み、調整された混合マトリクス G' が対応するマトリクス要素 g_{11}' 、 g_{12}' 、 g_{21}' および g_{22}' を含む場合、調整された混合マトリクス要素 g_{11}' ($n0$) は、シーケンス $g_{11}(1) \sim g_{11}(n0)$ から導き出すことができる。調整された混合マトリクス G' の他の混合マトリクス要素 g_{12}' 、 g_{21}' および g_{22}' に対して、等価な導出を用いることができる。

10

【0143】

図10の表は、全てのSAOC動作モードに対して、提案されたパラメータ制限スキームによって修正、例えば、制限することができる変換符号化係数のリストを提供する。図10の表は、第1カラム1010において、異なるSAOCモードを示す。図10の表は、更に、第2カラム1020において、提案されたパラメータ制限スキームによって、どのパラメータを修正する(例えば、制限する)ことができるかを示す。第3カラム1030は、非特許文献7のMPEG SAOCのFCD文書の対応する節の参照表示を示す。要約すると、図10の表は、全てのSAOC動作モードに対して、提案されたパラメータ制限スキームによって修正する(例えば、制限する)ことができる変換符号化係数のリストを、MPEG SAOCのFCD文書の対応する節を参照して示す。

20

【0144】

4.4 制限された相対偏差に対するパラメータ制限スキームの一般化された定式化

【0145】

上述のPLSに対して、一般化された定式化が存在する。この定式化は、次のように、一般的なパラメータ変数 \tilde{X}_i に対して、以下の最小化問題の形で表すことができる。

30

$$\begin{cases} \frac{\bar{X}_i}{\Lambda} \leq \tilde{X}_i \leq \Lambda \bar{X}_i, \\ \|\tilde{X}_i - X_i\| \rightarrow \min. \end{cases}$$

【0146】

ここで、 X_i の値は、最初に与えられ、「リファレンス」値 \bar{X}_i は、修正された変数 \tilde{X}_i の関数として、 $\bar{X}_i = F(\tilde{X}_i)$ のように推定することができる。

40

【0147】

上記において、パラメータ変数 X_i は、例えば、 $R(i)$ または $T(i)$ と同一とすることができる。同様に、調整されたパラメータ変数 \tilde{X}_i は、調整されたレンダリング係数 $\tilde{R}(i)$ または調整された変換符号化係数 $\tilde{T}(i)$ と同一とすることができる。変数 X_i 、 \tilde{X}_i は、また、例えば、混合マトリクス要素 $g_{mn}(i)$ および $g_{mn}'(i)$ に相当することができる。

【0148】

以下に、2つの解法アルゴリズムが述べられる。

50

【 0 1 4 9 】

一般に、このような最小化問題の正確な解を取得する解析的アプローチは、計算上大変な労力を要する。にもかかわらず、依然として P L S 目的に適するサブオプティマルな結果を提供する簡単で速い代替方法が存在する。2つのこのような簡単なアプローチがここで記載される。

【 0 1 5 0 】

4 . 4 . 1 ワンステップ解法

【 0 1 5 1 】

$\bar{X}_i \approx F(\tilde{X}_i)$ という仮定に基づくワンステップ解法は、次式によって、許容範囲の外側にある全ての値をその内側にあるように制限する。

10

$$\tilde{X}_i = \Lambda \bar{X}_i \quad \text{for} \quad X_i > \Lambda,$$

$$\tilde{X}_i = \frac{\bar{X}_i}{\Lambda} \quad \text{for} \quad X_i < \frac{1}{\Lambda}.$$

【 0 1 5 2 】

許容範囲（それは、許容差とみなすことができる）の内側にある値は、例えば、不変のままとすることができる。

【 0 1 5 3 】

4 . 4 . 2 反復解法

20

【 0 1 5 4 】

反復解法は、各ステップにおいて、1つの選択された範囲外の値 X_{i*} を \tilde{X}_{i*} に修正する。

$$\tilde{X}_{i*} = (1-\lambda)X_{i*} + \lambda \bar{X} \quad \text{with} \quad \lambda \in (0,1).$$

【 0 1 5 5 】

例えば、処理インデックス i^* は、次の条件を用いて選択することができる。

$$X_{i*} = \max\left(\frac{X_i}{\bar{X}}\right) \quad \text{and} \quad \frac{X_i}{\bar{X}} > \Lambda, \text{ or}$$

30

$$X_{i*} = \min\left(\frac{X_i}{\bar{X}}\right) \quad \text{and} \quad \frac{X_i}{\bar{X}} < \frac{1}{\Lambda}.$$

【 0 1 5 6 】

反復の数は、特定の値にセットするかまたはアルゴリズムから暗黙に導き出すことができる。

【 0 1 5 7 】

全てのこれらの方法は、上述のように、R C と T C を制限するために適用することができる点に注意しなければならない。

【 0 1 5 8 】

40

4 . 5 一般化された線形定式化

【 0 1 5 9 】

上述された P L S に対する一般化された線形定式化が存在する。前のセクションにおいて、一般的なパラメータ X_i の偏差は、比率 $\frac{X_i}{\bar{X}}$ として記載されている。対照的に、それは

、次のように、一般的なパラメータ変数 \tilde{X}_i に対して、以下の最小化問題に結果として導く、 $\|X_i - \bar{X}_i\|$ として定義することができる。

$$\begin{cases} (\bar{X}_i - \Lambda_{X-}) \leq \tilde{X}_i \leq (\bar{X}_i + \Lambda_{X+}), \\ \|\tilde{X}_i - X_i\| \rightarrow \min. \end{cases} \quad 10$$

【 0 1 6 0 】

ここで、 X_i の値は、最初に与えられ、「リファレンス」値 \bar{X}_i は、修正された変数 \tilde{X}_i の関数として、 $\bar{X}_i = F(\tilde{X}_i)$ のように推定することができる。

【 0 1 6 1 】

以下に、この問題に対する 2 つの解法アルゴリズムが記載される。

【 0 1 6 2 】

一般に、このような最小化問題の正確な解を取得する解析的アプローチは、計算上大変な労力を要する。にもかかわらず、依然として P L S 目的に適するサブオプティマルな結果を提供する簡単で速い代替方法が存在する。2 つのこのような簡単なアプローチがここで記載される。

【 0 1 6 3 】

4 . 5 . 1 ワンステップ解法

【 0 1 6 4 】

$\bar{X}_i \approx F(X_i)$ というという仮定に基づくワンステップ解法は、次式によって、許容範囲の外側にある全ての値をその内側にあるように制限する。

$$\tilde{X}_i = \min\left(\max\left(X_i, \bar{X}_i - \Lambda_{X-}\right), \bar{X}_i + \Lambda_{X+}\right).$$

【 0 1 6 5 】

4 . 5 . 2 反復解法

【 0 1 6 6 】

反復解法は、各ステップにおいて、 X_{i*} が許容範囲の外側にある場合に、1 つの選択された値 X_{i*} を \tilde{X}_{i*} に修正する。

$$X_{i*} > \bar{X}_{i*} \quad \text{and} \quad \|X_{i*} - \bar{X}_{i*}\| > \|X_{i*} - \Lambda_{X+}\| \Rightarrow \tilde{X}_{i*} = X_{i*} - S, \quad 40$$

$$X_{i*} < \bar{X}_{i*} \quad \text{and} \quad \|X_{i*} - \bar{X}_{i*}\| < \|X_{i*} - \Lambda_{X-}\| \Rightarrow \tilde{X}_{i*} = X_{i*} + S.$$

【 0 1 6 7 】

例えば、処理インデックス i^* は、条件： $\|X_{i*} - \bar{X}_{i*}\| \geq \|X_i - \bar{X}_i\|$ と、 $S = \lambda \|X_{i*} - \bar{X}_{i*}\|$ (ここで、 $\lambda \in (0,1)$) のような修正ステップサイズ値を用いて選択することができる。反復の数は、特定の値にセットするかまたはアルゴリズムから暗黙に導き出すことができる。

【 0 1 6 8 】

このアルゴリズムは、許容範囲を用いる、すなわち、それは動的に変化する（ X_{i*} に從属して）、フレキシブルな方法を提供する。

【 0 1 6 9 】

全てのこれらの方法は、上述のように、RCとTCを制限するために適用することができる点に注意しなければならない。

【 0 1 7 0 】

あるいは、以下のアルゴリズムを用いることができる。

If $X_{i*} > \bar{X}_{i*}$ and $\|X_{i*} - \bar{X}_{i*}\| > \Lambda_{X+}$ then

$$\tilde{X}_{i*} = X_{i*} - S$$

そして、

if $X_{i*} < \bar{X}_{i*}$ and $\|X_{i*} - \bar{X}_{i*}\| > \Lambda_{X-}$ then

$$\tilde{X}_{i*} = X_{i*} + S.$$

【 0 1 7 1 】

このバージョンのアルゴリズムは、固定の（静的な）許容範囲 x_- , x_+ を用いる。

【 0 1 7 2 】

4 . 6 更なる注釈

【 0 1 7 3 】

上述のように、全てのこれらの方法は、レンダリング係数と変換符号化係数を制限するために適用することができる点に注意しなければならない。

【 0 1 7 4 】

5 . 多重チャンネルのダウンミックス / アップミックスシナリオへのパラメータ制限スキームのアプリケーション

【 0 1 7 5 】

モノラルのダウンミックス / モノラルアップミックスシナリオの単一のTC PLS（例えば、直接制御）は、ダウンミックス / アップミックスチャンネルのいかなる組合せも考慮するTCマトリクスに拡張する。従って、直接制御は、各TCに対して個々に適用することができる。RC PLS（例えば間接制御）に対する多重チャンネルのアップミックスシナリオは、例えば、全ての個々のレンダリング係数が独立に処理される簡単な多重のモノラルアプローチにおいて実現することができる。

【 0 1 7 6 】

6 . リスニングテスト結果

【 0 1 7 7 】

6 . 1 テスト計画および項目

【 0 1 7 8 】

主観的リスニングテストは、提案された歪み制御尺度（DCM）コンセプトの知覚的パフォーマンスを評価し、それを通常のSAOC参照モデル（SAOC RM）復号化処理と比較するために行われた。

【 0 1 7 9 】

テスト計画は、提案されたパラメータ制限スキームの直接および間接の制御アプローチの個々のアプリケーションのケースならびにそれらの組み合わせを含む。通常の（パラメータ制限スキームPLSによって処理されていない）SAOCデコーダの出力信号は、SAOCのベースラインパフォーマンスを実証するために、試験に含まれる。加えて、ダウンミックス信号に対応する平凡なレンダリングのケースが、リスニングテストにおいて比較の目的で用いられる。

【 0 1 8 0 】

図 5 a の表は、リスニングテスト条件を記載する。

【 0 1 8 1 】

現行のリスニングテストに対して、極端なレンダリング条件に対する典型的なおよび最もクリチカルなアーチファクトタイプを表現する 4 つの項目が、提案募集 (C f P) のリスニングテスト素材から選択された。

【 0 1 8 2 】

図 5 b の表は、リスニングテストのオーディオ項目を記載する。

【 0 1 8 3 】

図 6 の表に係るレンダリングオブジェクトゲインは、考慮されるアップミックスシナリオに対して適用された。

10

【 0 1 8 4 】

提案された P L S は、通常の S A O C ビットストリームおよびダウンミックス (S A O C エンコーダサイドでのいかなる P L S 関連アクティビティも必要ない) を用いて動作し、残余情報を中継しないので、対応する S A O C ダウンミックス信号に対してコアコードは適用されなかった。

【 0 1 8 5 】

全てのテスト項目と考慮されるレンダリング条件に対して、次のように、P L S に対してグローバルな設定がとられる。

$$\Lambda_{\{R-,R+\}} = \Lambda_{\{T-,T+\}} = 6$$

20

【 0 1 8 6 】

6 . 2 テスト方法

【 0 1 8 7 】

主観的リスニングテストは、高品質リスニングができるように設計された音響的に隔離されたリスニングルームで行われた。再生は、ヘッドホン (L a k e P e o p l e の D / A コンバータと S T A X の S R M モニタを有する S T A X S R L a m d a P r o) を用いてなされた。

【 0 1 8 8 】

テスト方法は、中間品質オーディオの主観的評価のための隠されたリファレンスとアンカーを有する多重励振 (M U S H R A) 法 (非特許文献 6) に基づいて、空間オーディオ検証試験において用いられる手順に準拠した。テスト方法は、提案された D C M コンセプトの知覚的パフォーマンスを評価するために、ぴったりあわせて修正された。採用されたテスト方法に従って、リスナーは、以下のリスニングテスト指令に従って全てのテスト条件をお互いに比較するように命じられた。

30

【 0 1 8 9 】

各オーディオ項目に対して、

最初に、あなたがシステムユーザとして達成することを望む所望のサウンドミックスの記述を読んで下さい。

項目「BlackCoffee」： サウンドミックス内のソフトなホーンセクションサウンド

項目「Fanta4」：サウンドミックス内の大きなドラムサウンド

40

項目「LovePop」：サウンドミックス内のソフトなストリングセクションサウンド

項目「Audition」：ソフトな音楽と大きなボーカルサウンド

次に、以下の両方を記述する 1 つの共通の等級を用いて信号を等級分けして下さい。

所望のサウンドミックスの目的を達成する

全体のシーンのサウンド品質 (歪み、アーチファクト、不自然さ... を考慮する)

【 0 1 9 0 】

合計 9 人のリスナーは、実行された試験の各々に参加した。全ての被検者は、経験豊かなリスナーとみなすことができる。テスト条件は、各テスト項目と各リスナーに対して自動的にランダム化された。主観的応答は、コンピュータベースの M U S H R A プログラムによって、0 から 1 0 0 にわたるスケールで記録された。テスト下の項目間の瞬時スイッ

50

チングが可能とされた。

【 0 1 9 1 】

6 . 3 リスニングテスト結果

【 0 1 9 2 】

取得されたリスニングテスト結果を示す図面に関する簡単な概要は、解説において見ることができる。これらのプロットは、全てのリスナーを通じた項目毎の平均 M U S H R A 等級と、全ての評価された項目を通じた統計的平均値を、関連する 9 5 % の信頼区間と共に示す。

【 0 1 9 3 】

行われたリスニングテストの結果に基づいて、以下の知見をなすことができる。行われた全てのリスニングテストに対して、取得された M U S H R A スコアは、通常

10

の S A O C R M システムと比較して、全体の統計的平均値の意味で、提案された P L S 機能が良好なパフォーマンスを提供することを証明している。通常

の S A O C デコーダ (考慮された極端なレンダリング条件に対して大きなオーディオアーチファクトを示す) によって生成された全ての項目の品質は、所望のレンダリングシナリオを全く満たさないダウンミックとス同一のレンダリング設定の品質と比較して、わずかに高く等級分けされる点に注意しなければならない。それ故、提案された P L S は、全ての考慮されるリスニングテストシナリオに対して、主観的信号品質のかなりの改善に導くと結論づけることができる。また、最も有望な制限システムは、 R C と T C の P L S の両方の組合せから成ると結論づけることができる。

20

【 0 1 9 4 】

リスニングテスト結果に関する詳細は、図 7 の図解図において見ることができる。

【 0 1 9 5 】

7 . 実施変形例

【 0 1 9 6 】

いくつかの態様が装置の局面において記載されてきたが、これらの態様は、1つのブロックまたはデバイスが1つの方法ステップまたは方法ステップの特徴に対応する、対応する方法の記述をも表していることは明らかである。同様に、方法ステップの局面において記載された態様は、対応する装置の対応するブロックまたはアイテムまたは特徴の記述をも表している。いくつかまたは全ての方法ステップは、例えば、マイクロプロセッサ、プログラム可能なコンピュータまたは電子回路のようなハードウェア装置によって (または用いて) 実行することができる。いくつかの実施形態において、いくつかの1つ以上の最も重要な方法ステップは、このような装置によって実行することができる。

30

【 0 1 9 7 】

発明の符号化されたオーディオ信号は、デジタル記憶媒体上に記憶することができる、または、無線伝送媒体のような伝送媒体またはインターネットのような有線伝送媒体上を送信することができる。

【 0 1 9 8 】

特定の実施要求に従って、本発明の実施形態は、ハードウェアにおいてまたはソフトウェアにおいて実施することができる。実施は、その上に格納される電子的に読み込み可能な制御信号を有し、それぞれの方法が実行されるようにプログラム可能なコンピュータシステムと協働する (または協働することができる) デジタル記憶媒体、例えばフロッピー (登録商標) ディスク、DVD、ブルーレイ、CD、ROM、PROM、EPROM、EEPROM またはフラッシュメモリを用いて実行することができる。従って、デジタル記憶媒体はコンピュータ読取可能とすることができる。

40

【 0 1 9 9 】

本発明に係るいくつかの実施形態は、電子的に読み込み可能な制御信号を有し、本願明細書に記載された方法の1つが実行されるように、プログラム可能なコンピュータシステムと協働することができる、データキャリアを含む。

【 0 2 0 0 】

50

一般に、本発明の実施形態は、コンピュータプログラム製品がコンピュータ上で動作するとき、本発明の方法の1つを実行するために動作可能であるプログラムコードを有するコンピュータプログラム製品として実施することができる。プログラムコードは、例えば、機械読取可能なキャリア上に記憶することができる。

【0201】

他の実施形態は、機械読取可能なキャリア上に記憶され、本願明細書に記載された方法の1つを実行するためのコンピュータプログラムを含む。

【0202】

言い換えれば、本発明の方法の実施形態は、それ故に、コンピュータプログラムがコンピュータ上で動作するとき、本願明細書に記載された方法の1つを実行するためのプログラムコードを有するコンピュータプログラムである。

10

【0203】

本発明の方法の更なる実施形態は、それ故に、本願明細書に記載された方法の1つを実行するためのコンピュータプログラムがその上に記録されたデータキャリア（またはデジタル記憶媒体またはコンピュータ読取可能媒体）である。データキャリア、デジタル記憶媒体または記録された媒体は、通常は有形および/または非遷移的である。

【0204】

本発明の方法の更なる実施形態は、それ故に、本願明細書に記載された方法の1つを実行するためのコンピュータプログラムを表現するデータストリームまたは信号のシーケンスである。データストリームまたは信号のシーケンスは、データ通信接続、例えばインターネットを介して伝送されるように構成することができる。

20

【0205】

更なる実施形態は、本願明細書に記載された方法の1つを実行するように構成され、または適合された処理手段、例えばコンピュータ、またはプログラマブルロジックデバイスを含む。

【0206】

更なる実施形態は、本願明細書に記載された方法の1つを実行するためのコンピュータプログラムをインストールしたコンピュータを含む。

【0207】

いくつかの実施形態では、プログラマブルロジックデバイス（例えばフィールドプログラマブルゲートアレイ）を、本願明細書に記載された方法の機能の一部または全部を実行するために用いることができる。いくつかの実施形態では、フィールドプログラマブルゲートアレイは、本願明細書に記載された方法の1つを実行するために、マイクロプロセッサと協働することができる。一般に、方法は、好ましくはいかなるハードウェア装置によっても実行される。

30

【0208】

上記した実施形態は、単に本発明の原理に対して説明したものである。本願明細書に記載された構成および詳細の修正および変更は、他の当業者にとって明らかであると理解される。本発明は、それ故に、特許クレームのスコープのみによって制限され、本願明細書の実施形態の記述および説明によって提供された特定の詳細によって制限されないことを意図する。

40

【0209】

8．結論

【0210】

本発明に係る実施形態は、オーディオデコードにおける歪み制御に対して、パラメータ制限スキームを構築する。本発明に係るいくつかの実施形態は、所望の再生セットアップ（例えば、モノラル、ステレオ、5.1、他）の選択と、個人的な嗜好または他の基準に従ってレンダリングマトリクスを制御することによる所望の出力レンダリングシーンの対話式リアルタイム修正のためのユーザインターフェース手段を提供する、空間オーディオオブジェクト符号化（SAOC）に焦点を合わせている。しかしながら、提案された方法

50

をパラメトリック技術に一般的に適合させることは、直接的な作業である。

【0211】

ダウンミックス / 分離 / 混合ベースのパラメトリックアプローチのため、レンダーされたオーディオ出力の主観的品質はレンダリングパラメータ設定に依存する。ユーザ選択のレンダリング設定を選択する自由度は、全体の音響シーン内のオブジェクトの極端なゲイン操作のような、不適切なオブジェクトレンダリングオプションを選択するユーザのリスクを引き起こす。

【0212】

商用製品に対して、悪い音響品質および / またはオーディオアーチファクトを生じるとは、ユーザインターフェースのいかなる設定に対しても、なんとしても容認できない。生成されたS A O Cオーディオ出力の過剰な歪みを制御するために、レンダーされたシーンの知覚的な品質の尺度を演算し、この尺度（および他の情報）に基づいて、実際に適用されたレンダリング係数を修正するというアイデアに基づく、いくつかの計算上の尺度が記述されている（特許文献1参照）。

10

【0213】

本発明は、次のようなレンダーされたS A O Cシーンの主観的音響品質を保護する代替のアイデアを構築する。

全ての処理がS A O Cデコーダ / トランスコーダの中で完全に行われる

レンダーされた音響シーンの知覚されたオーディオ品質の複雑な尺度の明示の計算を含まない

20

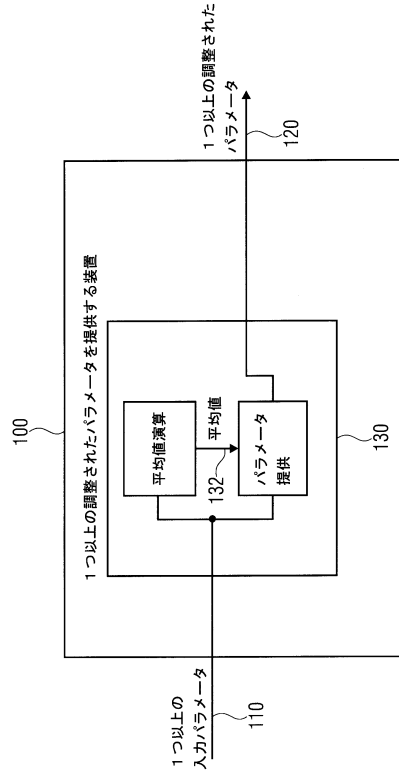
【0214】

これらのアイデアは、このように、S A O Cデコーダ / トランスコーダのフレームワーク内で、構造的に簡単で極めて効率的な方法で実施することができる。提案された歪み制御メカニズム（DCM）は、S A O Cデコーダ、すなわち、レンダリング係数（RC）および変換符号化係数（TC）に固有のパラメータを制限することを目的とするので、本書面の全体にわたって、パラメータ制限スキーム（PLS）と呼ばれる。

【0215】

しかしながら、パラメータ制限スキームは、いかなる異なるオーディオデコーダに対しても同様に適用することができる。

【図 1】



【図 2】

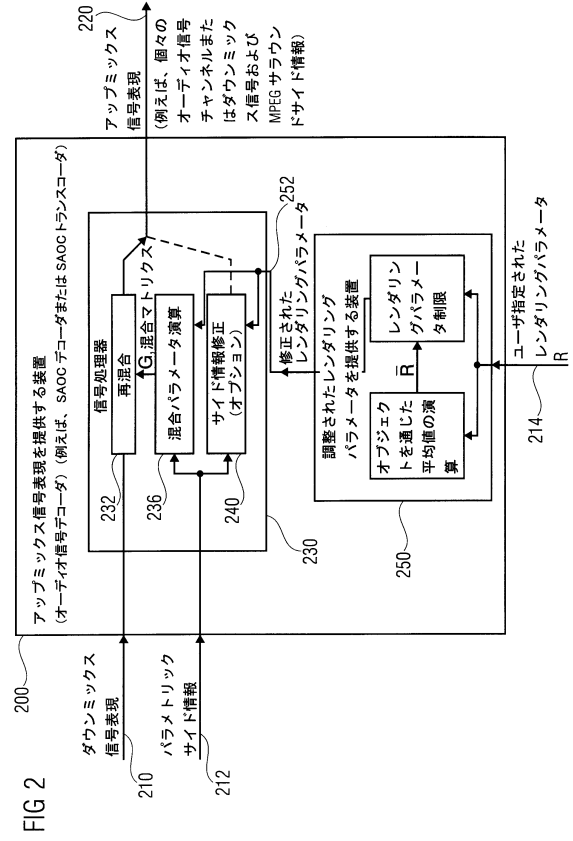


FIG 1

FIG 2

【図 3】

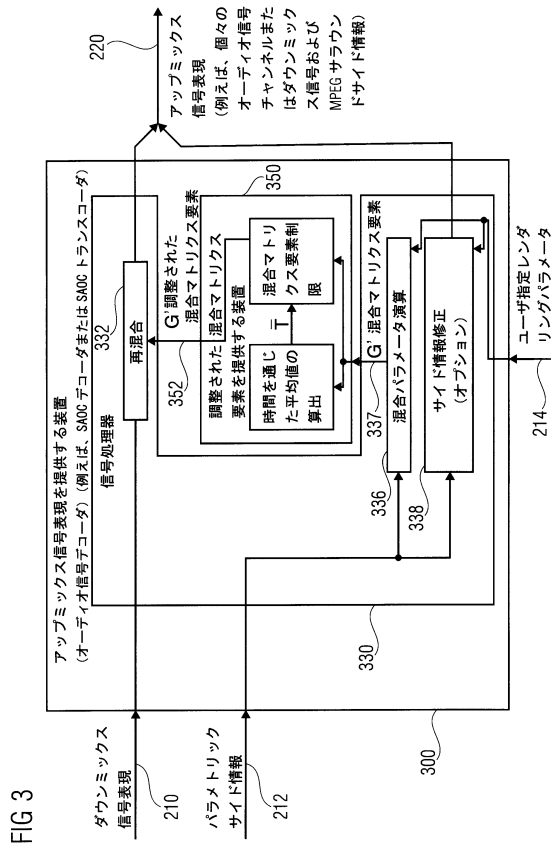
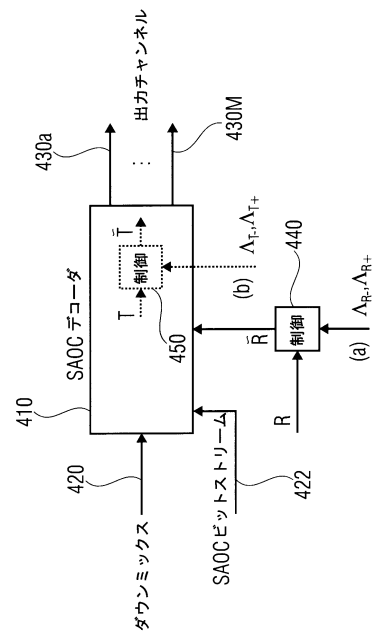


FIG 3

【図 4】



パラメータ制御スキーム - (a) 間接および (b) 直接制御

FIG 4

【 図 5 a 】

表－リスニングテスト条件

コード名	説 明
"DMX"	通常の SAOC デコーダの平凡なダウンミックス類似のレンダリング信号
"noLim"	通常の (PLS によって処理されない) SAOC デコーダの出力
"TC"	TC 直接制御 PLS を有する SAOC デコーダの出力
"RO"	RC 間接制御 PLS を有する SAOC デコーダの出力
"RO_TC_1"	統合された TC/RC 制御 PLS を有する SAOC デコーダの出力

FIG 5A

【 図 6 】

表－テストされた極端なレンダリング条件

リスニング項目	オブジェクト#	オーディオオブジェクトの説明	レンダリング係数
"BlackCoffee"	1	Brass 2	0.0001
	2	Brass 1	0.0001
	3	Organ	3.1623
	4	Drums&Bass	3.1623
	5	Percussion	3.1623
"Fanta4"	1	Vocals	0.2512
	2	Bass	0.2512
	3	Beat	2.0
"LovePop"	1	Drums	2.5119
	2	Bass	2.5119
	3	Electric Guitar	2.5119
	4	Acoustic Guitar	2.5119
	5	Strings	0.0001
"Audition"	1	BG Vocals	0.1778
	2	Bass	0.1778
	3	Drums Reverb	0.1778
	4	Kick Drum	0.1778
	5	Lead Guitar	0.1778
	6	Lead Vocals Double	3.1623
	7	Lead Vocals	3.1623
	8	Drums Overhead	0.1778

FIG 6

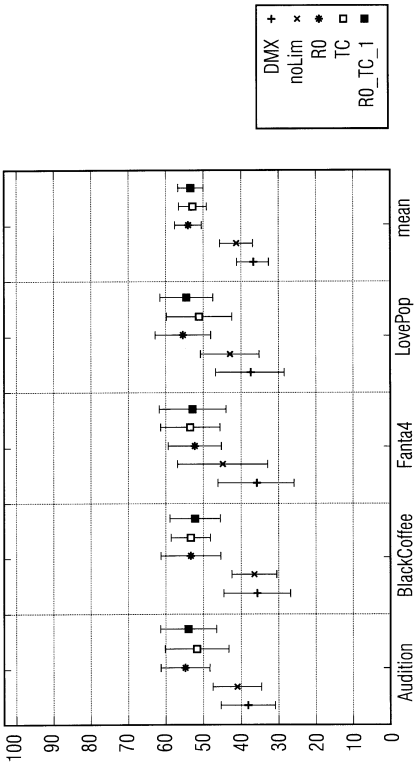
【 図 5 b 】

表－リスニングテストのオーディオ項目

リスニング項目	レンダリングのタイプとマトリクス	継続時間
"BlackCoffee"	サウンドミックス内のソフトなホーンセクションサウンド	7 秒
"Fanta4"	サウンドミックス内の強いドラムサウンド	7 秒
"LovePop"	サウンドミックス内のソフトなストリングセクションサウンド	7 秒
"Audition"	ソフトな音楽と強いボーカルサウンド	1 3 秒

FIG 5B

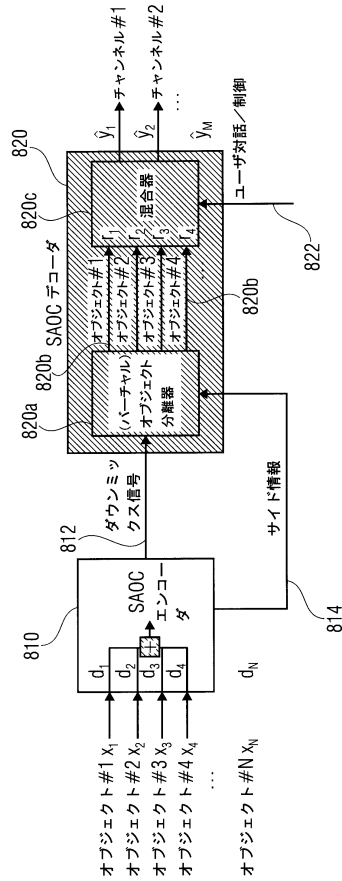
【 図 7 】



図－PLS に対する MUSHRA リスニングテスト結果

FIG 7

【図 8】



MPEG-SAOC システム概要

FIG 8

【図 9 a】

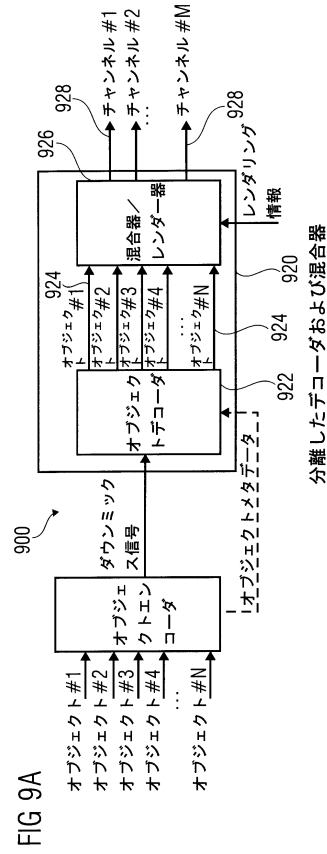


FIG 9A

【図 9 b】

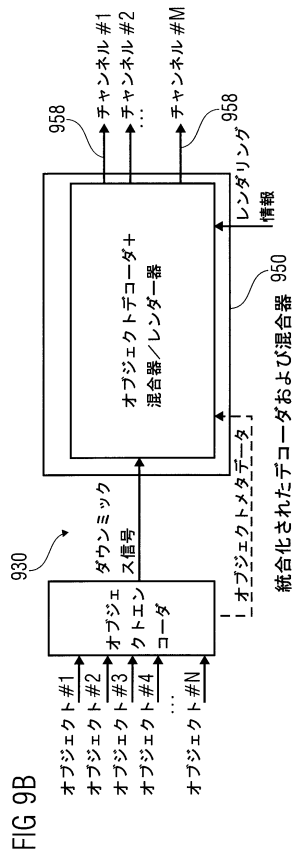


FIG 9B

【図 9 c】

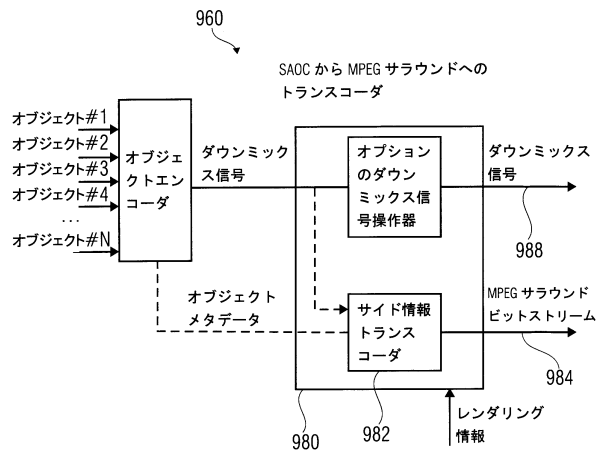


FIG 9C

1010	1020	1030
SAOC モード	修正された係数	参考資料
"X-1-1" 復号化モード	混合マトリクス係数 ($G^{(m)}$)	6.6.1.3
"X-1-2" 復号化モード	混合マトリクス係数 ($G^{(m)}$)	6.6.1.2
"X-1-0" 復号化モード	混合マトリクス係数 ($G^{(m)}$)	6.6.1.1
"X-1-5" 変換符号化モード	任意ダウンミックスゲイン係数 ($ADG^{(m)}$)	6.5.2.4
"X-2-1" 復号化モード	混合マトリクス係数 ($G^{(m)}$)	6.6.1.6
"X-2-2" 復号化モード	混合マトリクス係数 ($G^{(m)}$)	6.6.1.5
"X-2-0" 復号化モード	混合マトリクス係数 ($G^{(m)}$)	6.6.1.4
"X-2-5" 変換符号化モード	予測マトリクス係数 (C_3)	6.5.3.3.1

FIG 10

フロントページの続き

(31)優先権主張番号 61/369,256

(32)優先日 平成22年7月30日(2010.7.30)

(33)優先権主張国 米国(US)

(72)発明者 ファルヒ コルネリア

オーストリア共和国 6063 ルム フィンケンベルク 2

(72)発明者 テレンチエフ レオン

ドイツ連邦共和国 91056 エアランゲン アム オイローパカナル 36 アパルトマーン
11

審査官 金田 孝之

(56)参考文献 国際公開第2008/100067(WO, A1)

米国特許出願公開第2010/0076774(US, A1)

(58)調査した分野(Int.Cl., DB名)

G10L 19/00 - 19/26

H04S 1/00

H04S 3/00 - 3/02

H04S 5/00 - 5/02

H04S 7/00

(54)【発明の名称】ダウンミックス信号表現と、ダウンミックス信号表現に関するパラメトリックサイド情報に基づくアップミックス信号表現の提供に対して、平均値を用いて、1つ以上の調整されたパラメータを提供する装置、方法およびコンピュータプログラム