

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7615193号  
(P7615193)

(45)発行日 令和7年1月16日(2025.1.16)

(24)登録日 令和7年1月7日(2025.1.7)

(51)国際特許分類	F I
H 0 4 S 7/00 (2006.01)	H 0 4 S 7/00 3 5 0
G 1 0 K 15/00 (2006.01)	G 1 0 K 15/00 M
G 1 0 K 15/12 (2006.01)	G 1 0 K 15/12

請求項の数 25 (全43頁)

(21)出願番号	特願2023-31911(P2023-31911)	(73)特許権者	500341779
(22)出願日	令和5年3月2日(2023.3.2)		フラウンホーファー・ゲゼルシャフト・
(62)分割の表示	特願2021-77240(P2021-77240)の 分割		ツール・フェルデルング・デル・アンゲ ヴァンテン・フォルシュング・アインゲ トラーゲネル・フェライン
原出願日	平成26年7月18日(2014.7.18)		ドイツ連邦共和国, 8 0 6 8 6 ミュン ヘン, ハンザシュトラーセ 2 7 ツェー
(65)公開番号	特開2023-71866(P2023-71866A)	(74)代理人	100134119
(43)公開日	令和5年5月23日(2023.5.23)		弁理士 奥町 哲行
審査請求日	令和5年3月30日(2023.3.30)	(72)発明者	フュグ・シモーネ
(31)優先権主張番号	13177361.6		ドイツ連邦共和国 9 0 5 6 2 カルヒロ イト ヴァイスガッセ 2 4
(32)優先日	平成25年7月22日(2013.7.22)	(72)発明者	プログスティーズ・ヤン
(33)優先権主張国・地域又は機関	欧州特許庁(EP)		ドイツ連邦共和国 9 0 7 6 8 フュルト サンドライテ 3 9 アー
(31)優先権主張番号	13189255.6		
(32)優先日	平成25年10月18日(2013.10.18)		
(33)優先権主張国・地域又は機関			
	最終頁に続く		最終頁に続く

(54)【発明の名称】 オーディオ信号を処理するための方法、信号処理ユニット、パイノラールレンダラ、オーディオエンコーダおよびオーディオデコーダ

(57)【特許請求の範囲】

【請求項 1】

室内インパルス応答に従ってオーディオ信号を処理するための方法であって、前記室内インパルス応答は初期部分と後期残響とを含み、前記室内インパルス応答の前記後期残響は合成残響に置き換えられ、前記方法は、

前記室内インパルス応答の前記初期部分と前記合成残響とを用いて前記オーディオ信号を別々に処理することであって、前記合成残響を用いて前記オーディオ信号を処理することが、スケーリングされた残響信号を生成することを含み、前記スケーリングが前記オーディオ信号に依存する、処理することと、

前記室内インパルス応答の前記初期部分を用いて処理された前記オーディオ信号と前記スケーリングされた残響信号とを合成することとを含む、方法。

【請求項 2】

前記スケーリングは、前記オーディオ信号の1つまたは複数の入力チャネルの状態に依存する、請求項1に記載の方法。

【請求項 3】

前記オーディオ信号の前記1つまたは複数の入力チャネルの前記状態は、入力チャネルの数と、アクティブ入力チャネルの数と、前記入力チャネル中のアクティビティとのうちの1つまたは複数を含む、請求項2に記載の方法。

【請求項 4】

10

20

前記スケーリングされた残響信号を生成することは、利得ファクタを適用することを含み、前記利得ファクタは、前記オーディオ信号の前記1つまたは複数の入力チャンネルの前記状態に基づいておよび/または前記オーディオ信号についての前記あらかじめ定義されたまたは計算される相関測度に基づいて決定される、請求項2または3に記載の方法。

【請求項5】

前記スケーリングされた残響信号を生成することは、前記オーディオ信号の前記合成残響を用いて前記オーディオ信号を処理する前に、その間にまたはその後前記利得ファクタを適用することを含む、請求項4に記載の方法。

【請求項6】

前記利得ファクタは以下のように決定され、

$$g = c_u + \cdot (c_c - c_u)$$

ただし

$c_u$  = 前記オーディオ信号についてのあらかじめ定義されたまたは計算される相関係数、  
 $c_c$  = 前記オーディオ信号の前記1つまたは複数の入力チャンネルの前記状態を示すファクタであり、 $c_u$ は、全体として無相関なチャンネルを指し、 $c_c$ は、全体として相関するチャンネルに関する、  
 請求項4または5に記載の方法。

【請求項7】

$c_u$ および $c_c$ は以下のように決定され、

【数1】

$$c_u = 10^{\frac{10 \cdot \log_{10}(K_{in})}{20}} = \sqrt{K_{in}}$$

【数2】

$$c_c = 10^{\frac{20 \cdot \log_{10}(K_{in})}{20}} = K_{in}$$

ただし

$K_{in}$  = アクティブまたは固定ダウンミックスチャンネルの数である、  
 請求項6に記載の方法。

【請求項8】

前記利得ファクタは、複数のオーディオフレームにわたって低域フィルタ処理される、  
 請求項4から7のうちのいずれか一項に記載の方法。

【請求項9】

前記利得ファクタは以下のように低域フィルタ処理され、

【数3】

$$g_s(t_i) = c_{s,old} \cdot g_s(t_i - 1) + c_{s,new} \cdot g$$

【数4】

$$c_{s,old} = e^{-\left(\frac{1}{f_s \cdot \frac{t_s}{k}}\right)}$$

10

20

30

40

50

## 【数 5】

$$C_{S,new} = 1 - C_{S,old}$$

ただし

$t_s$  = 前記低域フィルタの時定数

$t_i$  = フレーム  $t_i$  におけるオーディオフレーム

$g_s$  = 平滑化された利得ファクタ

$k$  = フレームサイズ、および

$f_s$  = サンプル周波数である、

10

請求項 8 に記載の方法。

## 【請求項 10】

前記スケーリングは、前記オーディオ信号のあらかじめ定義されたまたは計算される相関測度に依存する、請求項 1 から 9 のうちのいずれか一項に記載の方法。

## 【請求項 11】

前記スケーリングされた残響信号を生成することは、前記オーディオ信号の相関分析を含む、請求項 1 から 10 のうちのいずれか一項に記載の方法。

## 【請求項 12】

前記オーディオ信号の前記相関分析は、前記オーディオ信号のオーディオフレームについて合成相関測度を決定することを含み、前記合成相関測度は、1つのオーディオフレームの複数のチャンネル合成についての相関係数を合成することによって計算され、各オーディオフレームは1つまたは複数のタイムスロットを含む、請求項 11 に記載の方法。

20

## 【請求項 13】

前記1つのオーディオフレームの前記複数のチャンネル合成についての前記相関係数を合成することは、前記1つのオーディオフレームの複数の相関係数を平均化することを含む、請求項 12 に記載の方法。

## 【請求項 14】

前記合成相関測度を決定することは、

(i) 前記1つのオーディオフレームのチャンネルごとに全体的な平均値を計算することと、

30

(ii) 前記1つのオーディオフレームの対応するチャンネルから前記平均値を減算することによってゼロ平均オーディオフレームを計算することと、

(iii) 複数のチャンネル合成についてそれぞれの相関係数を計算することと、

(iv) 前記合成相関測度を前記それぞれの相関係数の平均値として計算することとを含む、請求項 12 または 13 に記載の方法。

## 【請求項 15】

チャンネル合成についての前記それぞれの相関係数は以下のように計算され、

## 【数 6】

$$\rho[m, n] = \left| \frac{1}{(N-1)} \cdot \frac{\sum_i \sum_j x_m[i, j] \cdot x_n[i, j]^*}{\sum_j \sigma(x_m[j]) \cdot \sigma(x_n[j])} \right|$$

40

ただし

$[m, n]$  = 相関係数、

$(x_m[j])$  = チャンネル  $m$  の1つのタイムスロット  $j$  にわたる標準偏差、

$(x_n[j])$  = チャンネル  $n$  の1つのタイムスロット  $j$  にわたる標準偏差、

$x_m, x_n$  = ゼロ平均変数、

$i [1, N]$  = 周波数帯域、

50

$j \in [1, M]$  = タイムスロット、  
 $m, n \in [1, K]$  = チャネル、  
 \* = 複素共役である、

請求項 1 2 から 1 4 のうちのいずれか一項に記載の方法。

【請求項 1 6】

前記スケーリングされた残響信号の開始を前記室内インパルス応答における初期反射から後期残響への遷移点に一致させるように前記スケーリングされた残響信号を遅延させることを含む、請求項 1 から 1 5 のうちのいずれか一項に記載の方法。

【請求項 1 7】

前記合成残響を用いて前記オーディオ信号を処理することは、前記オーディオ信号をダウンミックスすることと、前記ダウンミックスされたオーディオ信号を残響器に印加することを含む、請求項 1 から 1 6 のうちのいずれか一項に記載の方法。

10

【請求項 1 8】

コンピュータによって実行されているとき、請求項 1 から 1 7 のうちのいずれか一項に記載の方法を行うための命令を記憶したコンピュータ可読媒体。

【請求項 1 9】

オーディオ信号を受信するための入力と、  
 室内インパルス応答の初期部分に従って前記受信されたオーディオ信号を処理するための初期部分プロセッサであって、前記室内インパルス応答は前記初期部分と後期残響とを含み、前記室内インパルス応答の前記後期残響は合成残響に置き換えられる、初期部分プロセッサと、

20

前記室内インパルス応答の前記合成残響に従って前記受信されたオーディオ信号を処理するための後期残響プロセッサであって、前記後期残響プロセッサが、スケーリングされた残響信号を生成するように構成され、前記スケーリングが、前記受信されたオーディオ信号に依存する後期残響プロセッサと、

前記受信されたオーディオ信号の前記処理された初期部分と前記スケーリングされた残響信号とを出力オーディオ信号に合成するための出力とを備える、信号処理ユニット。

【請求項 2 0】

前記後期残響プロセッサは、  
 前記オーディオ信号を受信し残響信号を生成する残響器と、  
 前記残響器の入力または出力に結合され利得ファクタによって制御される利得段とを備える、請求項 1 9 に記載の信号処理ユニット。

30

【請求項 2 1】

前記オーディオ信号に依存する前記利得ファクタを生成する相関分析器を備える、請求項 2 0 に記載の信号処理ユニット。

【請求項 2 2】

前記利得段に結合された低域フィルタと、  
 前記利得段と加算器との間に結合された遅延要素とのうちの少なくとも 1 つをさらに備え、前記加算器は、前記初期部分プロセッサと前記出力とにさらに結合された、請求項 2 0 または 2 1 に記載の信号処理ユニット。

40

【請求項 2 3】

請求項 1 9 から 2 2 のうちのいずれか一項に記載の信号処理ユニットを備える、バイノーラルレンダラ。

【請求項 2 4】

コーディングより前にオーディオ信号を処理するために請求項 1 9 から 2 2 のうちのいずれか一項に記載の信号処理ユニットまたは請求項 2 3 に記載のバイノーラルレンダラを備える、前記オーディオ信号をコーディングするためのオーディオエンコーダ。

【請求項 2 5】

復号されたオーディオ信号を処理するために請求項 1 9 から 2 2 のうちのいずれか一項

50

に記載の信号処理ユニットまたは請求項 2 3 に記載のバイノーラルレンダラを備える、符号化されたオーディオ信号を復号するためのオーディオデコーダ。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、オーディオ符号化/復号の分野に関し、特に空間オーディオコーディングおよび空間オーディオ・オブジェクト・コーディング、例えば3Dオーディオ・コーデック・システムの分野に関する。本発明の実施形態は、室内インパルス応答に従ってオーディオ信号を処理するための方法に、信号処理ユニット、バイノーラルレンダラ、オーディオエンコーダおよびオーディオデコーダに関する。

10

【背景技術】

【0002】

空間オーディオ・コーディング・ツールが当技術分野でよく知られており、例えば、MPEGサラウンド規格において規格化されている。空間オーディオコーディングは、複数の元の入力、例えば、5つまたは7つの入力チャンネルから開始し、これらは、再生セットアップにおいてそれらの配置によって、例えば、左チャンネル、中央チャンネル、右チャンネル、左サラウンドチャンネル、右サラウンドチャンネルおよび低周波拡張チャンネルとして識別される。空間オーディオエンコーダは、元のチャンネルから1つまたは複数のダウンミックスチャンネルを導出し得、さらに、チャンネルコヒーレンス値のチャンネル間レベル差、チャンネル間位相差、チャンネル間時間差など、空間キューに関係するパラメトリックデータを導出し得る。元の入力チャンネルの近似バージョンである出力チャンネルを最終的に取得するために、1つまたは複数のダウンミックスチャンネルは、ダウンミックスチャンネルおよび関連するパラメトリックデータを復号するために空間キューを示すパラメトリックサイド情報とともに空間オーディオデコーダに送信される。出力セットアップにおけるチャンネルの配置は固定であり得、例えば、5.1フォーマット、7.1フォーマットなどであり得る。

20

【0003】

また、空間オーディオ・オブジェクト・コーディング・ツールが当技術分野でよく知られており、例えば、MPEG SAOC規格(SAOC=空間オーディオ・オブジェクト・コーディング)において規格化されている。元のチャンネルから開始する空間オーディオコーディングとは対照的に、空間オーディオ・オブジェクト・コーディングは、あるレンダリング再生セットアップのために自動的に専用化されないオーディオオブジェクトから開始する。そうではなく、再生シーン中のオーディオオブジェクトの配置はフレキシブルであり、例えば、あるレンダリング情報を空間オーディオ・オブジェクト・コーディング・デコーダ中に入力することによって、ユーザによって設定され得る。代替または追加として、レンダリング情報は、追加のサイド情報またはメタデータとして送信され得、レンダリング情報は、あるオーディオオブジェクトが再生セットアップ中のその位置において(例えば、時間とともに)配置されるべきである情報を含み得る。あるデータ圧縮を得るために、いくつかのオーディオオブジェクトはSAOCエンコーダを使用して符号化され、SAOCエンコーダは、あるダウンミックス情報に従ってオブジェクトをダウンミックスすることによって、入力オブジェクトから1つまたは複数のトランスポートチャンネルを計算する。さらに、SAOCエンコーダは、オブジェクトレベル差(OLD)、オブジェクトコヒーレンス値など、オブジェクト間キューを表すパラメトリックサイド情報を計算する。SAC(SAC=空間オーディオコーディング)の場合のように、オブジェクト間パラメトリックデータは、個々の時間/周波数タイルについて計算される。オーディオ信号のあるフレーム(例えば、1024または2048個のサンプル)では、フレームごとにおよび周波数帯域ごとにパラメトリックデータが提供されるように複数の周波数帯域(例えば24、32、または、64個の帯域)が考慮される。例えば、オーディオ部分が20個のフレームを有するときおよび各フレームが32個の周波数帯域に再分割されるとき、時間/周波数タイルの数は640である。

30

40

【0004】

50

3Dオーディオシステムでは、オーディオ信号が特定の室内において聴取されるかのようにオーディオ信号の空間印象を提供することが望まれることがある。そのような状況では、特定の室内の室内インパルス応答が、例えばその測定に基づいて提供され、オーディオ信号を聴取者に提示するときはそのオーディオ信号を処理するために使用される。そのような提示では、後期残響から分離された直接音および初期反射を処理することが望まれ得る。

【発明の概要】

【発明が解決しようとする課題】

【0005】

本発明の基礎をなす目的は、室内インパルス応答の初期部分と後期残響を用いてオーディオ信号を別々に処理するための承認された手法を提供し、それにより、完全なインパルス応答を用いたオーディオ信号の畳み込みの結果とできる限り知覚的に同等である結果を達成することを可能にすることである。

10

【課題を解決するための手段】

【0006】

この目的は、請求項1の方法と、請求項19の信号処理ユニットと、請求項23のバイノーラルレンダラと、請求項24のオーディオエンコーダと、請求項25のオーディオデコーダとによって達成される。

【0007】

本発明は、従来手法では、室内インパルス応答によるオーディオ信号の処理時に、初期部分と残響とに関してオーディオ信号を別々に処理する結果が、完全なインパルス応答を用いて畳み込みを適用するときの結果から逸脱するという点で問題が存在するという発明者の発見に基づく。本発明は、例えば、合成残響手法を使用するとき、残響への入力オーディオ信号の影響が十分に保持されないため、適切なレベルの残響が入力オーディオ信号とインパルス応答の両方に依存するという発明者の発見にさらに基づく。インパルス応答の影響は、入力パラメータとして既知の残響特性を使用することによって考慮され得る。入力信号の影響は、入力オーディオ信号に基づいて決定される残響のレベルを適応させるための信号依存スケールリングによって考慮され得る。この手法により、残響の知覚されるレベルは、バイノーラルレンダリングのために完全な畳み込み手法を使用するときの残響のレベルにより良く一致することが分かっている。

20

【0008】

(1) 本発明は、室内インパルス応答に従ってオーディオ信号を処理するための方法を提供し、本方法は、

室内インパルス応答の初期部分と後期残響とを用いてオーディオ信号を別々に処理することであって、後期残響を処理することが、スケールリングされた残響信号を生成することを含み、スケールリングがオーディオ信号に依存する、処理することと、

室内インパルス応答の初期部分を用いて処理されたオーディオ信号とスケールリングされた残響信号とを合成することを含む。

30

上記で説明した従来手法と比較すると、本発明の手法は、完全な畳み込み結果を計算する必要なしにまたは広範で不正確な聴覚モデルを適用する必要なしに後期残響をスケールリングすることを可能にするので、有利である。本発明の手法の実施形態は、完全な畳み込み手法での残響のようにそれが聞こえるように人工的後期残響をスケールリングするための容易な方法を提供する。スケールリングは入力信号に基づき、追加の聴覚モデルまたはターゲット残響ラウドネスは不要である。スケールリングファクタは、エンコーダ/デコーダチェーン中のオーディオ素材も時間周波数領域においてしばしば利用可能でもあるので利点であるこの時間周波数領域において導出され得る。

40

【0009】

(2) 実施形態によれば、スケールリングは、オーディオ信号の1つまたは複数の入力チャンネルの状態(例えば、入力チャンネルの数、アクティブ入力チャンネルの数および/また

50

は入力チャンネル中のアクティビティ)に依存し得る。

スケーリングは、低減した計算オーバーヘッドにより入力オーディオ信号から容易に決定され得るので、これは有利である。例えば、スケーリングは、元のオーディオ信号と比較して低減した数のチャンネルを含む、現在考慮されているダウンミックスチャンネルにダウンミックスされる元のオーディオ信号中のチャンネルの数を単に決定することによって決定され得る。代替的に、現在考慮されているダウンミックスチャンネルにダウンミックスされるアクティブチャンネル(現在オーディオフレーム中で何らかのアクティビティを示すチャンネル)の数は、残響信号をスケーリングするための基礎を形成し得る。

【0010】

(3) 実施形態によれば、(入力チャンネル状態の追加または代替として)スケーリングは、オーディオ信号のあらかじめ定義されたまたは計算される相関測度に依存する。

10

あらかじめ定義された相関測度を使用することは、プロセスにおける計算複雑さを低減するので、有利である。あらかじめ定義された相関測度は、複数のオーディオ信号の分析に基づいて経験的に決定され得る、例えば0.1~0.9の範囲内の、固定値を有し得る。一方、個々に現在処理されているオーディオ信号についてより正確な測度を取得することが望まれる場合、必要とされる追加の計算リソースにもかかわらず、相関測度を計算することが有利である。

【0011】

(4) 実施形態によれば、スケーリングされた残響信号を生成することは、利得ファクタを適用することを含み、利得ファクタは、オーディオ信号の1つまたは複数の入力チャンネルの状態に基づいておよび/またはオーディオ信号についてのあらかじめ定義されたまたは計算される相関測度に基づいて決定され、利得ファクタは、オーディオ信号の後期残響を処理する前に、その間にまたはその後適用され得る。

20

利得ファクタは、上記のパラメータに基づいて容易に計算され得、実装詳細の依存する処理チェーン中の残響器に関してフレキシブルに適用され得るので、これは有利である。

【0012】

(5) 実施形態によれば、利得ファクタは以下のように決定される。

$$g = c_u + \cdot (c_c - c_u)$$

ただし

= オーディオ信号についてのあらかじめ定義されたまたは計算される相関測度、

30

$c_u$ 、 $c_c$  = オーディオ信号の1つまたは複数の入力チャンネルの状態を示すファクタであり、 $c_u$ は、全体として無相関なチャンネルを指し、 $c_c$ は、全体として相関するチャンネルに関係する。

ファクタは、オーディオ信号中のアクティブチャンネルの数とともに経時的にスケーリングするので、これは有利である。

【0013】

(6) 実施形態によれば、 $c_u$ および $c_c$ は以下のように決定される。

【数1】

$$c_u = 10^{\frac{10 \cdot \log_{10}(K_{in})}{20}} = \sqrt{K_{in}}$$

40

【数2】

$$c_c = 10^{\frac{20 \cdot \log_{10}(K_{in})}{20}} = K_{in}$$

ただし

50

$K_{in}$  = アクティブまたは固定ダウンミックスチャンネルの数。

ファクタは、オーディオ信号中のアクティブチャンネルの数に直接依存するので、これは有利である。どのチャンネルもアクティブでない場合、残響は0でスケーリングされ、より多くのチャンネルがアクティブであれば、残響の振幅はより大きくなり得る。

【0014】

(7) 実施形態によれば、利得ファクタは、複数のオーディオフレームにわたって低域フィルタ処理され、利得ファクタは以下のように低域フィルタ処理され得る。

【数3】

$$g_s(t_i) = c_{s,old} \cdot g_s(t_i - 1) + c_{s,new} \cdot g \quad 10$$

【数4】

$$c_{s,old} = e^{-\left(\frac{1}{f_s \cdot k} \cdot \frac{t_s}{k}\right)}$$

【数5】

$$c_{s,new} = 1 - c_{s,old} \quad 20$$

ただし

$t_s$  = 低域フィルタの時定数

$t_i$  = フレーム  $t_i$  におけるオーディオフレーム

$g_s$  = 平滑化された利得ファクタ

$k$  = フレームサイズ、および

$f_s$  = サンプリング周波数。

30

スケーリングファクタについて時間とともに急激な変化が起こらないので、これは有利である。

【0015】

(8) 実施形態によれば、スケーリングされた残響信号を生成することは、オーディオ信号の相関分析を含み、オーディオ信号の相関分析は、オーディオ信号のオーディオフレームについて合成相関測度を決定することを含み得、合成相関測度は、1つのオーディオフレームの複数のチャンネル合成についての相関係数を合成することによって計算され得、各オーディオフレームは1つまたは複数のタイムスロットを含み、相関係数を合成することは、オーディオフレームの複数の相関係数を平均化することを含み得る。

相関は、1つのオーディオフレームの全体的な相関を記述する1つの単一の値によって記述され得るので、これは有利である。複数の周波数依存値を扱う必要はない。

40

【0016】

(9) 実施形態によれば、合成相関測度を決定することは、(i) 1つのオーディオフレームのチャンネルごとに全体的な平均値を計算することと、(ii) 対応するチャンネルから平均値を減算することによってゼロ平均オーディオフレームを計算することと、(iii) 複数のチャンネル合成について相関係数を計算することと、(iv) 合成相関測度を複数の相関係数の平均として計算することとを含み得る。

上述のように、フレームごとにただ1つの単一の全体的な相関値が計算され(容易な処理)、その計算は、ゼロ平均信号およびその標準偏差を同じく使用する「標準」のピアソンの相関係数と同様に行われ得るので、これは有利である。

50

## 【 0 0 1 7 】

( 1 0 ) 実施形態によれば、チャンネル合成についての相関係数は以下のように決定される。

## 【数 6】

$$\rho[m, n] = \left| \frac{1}{(N-1)} \cdot \frac{\sum_i \sum_j x_m[i, j] \cdot x_n[i, j]^*}{\sum_j \sigma(x_m[j]) \cdot \sigma(x_n[j])} \right|$$

ただし

- [ m , n ] = 相関係数、  
 ( x<sub>m</sub> [ j ] ) = チャンネル m の 1 つのタイムスロット j にわたる標準偏差、  
 ( x<sub>n</sub> [ j ] ) = チャンネル n の 1 つのタイムスロット j にわたる標準偏差、  
 x<sub>m</sub> , x<sub>n</sub> = ゼロ平均変数、  
 i [ 1 , N ] = 周波数帯域、  
 j [ 1 , M ] = タイムスロット、  
 m , n [ 1 , K ] = チャンネル、  
 \* = 複素共役。

ピアソンの相関係数のためのよく知られている式が使用され得、周波数依存および時間依存式に変換されるので、これは有利である。

## 【 0 0 1 8 】

( 1 1 ) 実施形態によれば、オーディオ信号の後期残響を処理することは、オーディオ信号をダウンミックスすることと、ダウンミックスされたオーディオ信号を残響器に印加することを含む。

例えば、残響器中での処理は、より少ないチャンネルを扱う必要があり、ダウンミックスプロセスが直接制御され得るので、これは有利である。

## 【 0 0 1 9 】

( 1 2 ) 本発明は、オーディオ信号を受信するための入力と、室内インパルス応答の初期部分に従って、受信されたオーディオ信号を処理するための初期部分プロセッサと、室内インパルス応答の後期残響に従って、受信されたオーディオ信号を処理するための後期残響プロセッサであって、後期残響プロセッサが、受信されたオーディオ信号に依存するスケーリングされた残響信号を生成するように構成またはプログラムされた後期残響プロセッサと、室内インパルス応答の初期部分を用いて処理されたオーディオ信号とスケーリングされた残響信号とを出力オーディオ信号に合成するための出力とを備える、信号処理ユニットを提供する。

## 【 0 0 2 0 】

( 1 3 ) 実施形態によれば、後期残響プロセッサは、オーディオ信号を受信し残響信号を生成する残響器と、オーディオ信号に依存する利得ファクタを生成する相関分析器と、残響器の入力または出力に結合され相関分析器によって提供される利得ファクタによって制御される利得段とを備える。

## 【 0 0 2 1 】

( 1 4 ) 実施形態によれば、信号処理ユニットは、相関分析器と利得段との間に結合された低域フィルタと、利得段と加算器との間に結合された遅延要素とのうちの少なくとも 1 つとをさらに備え、加算器は、初期部分プロセッサと出力とにさらに結合される。

## 【 0 0 2 2 】

( 1 5 ) 本発明は、本発明の信号処理ユニットを備える、バイノーラルレンダラを提供する。

## 【 0 0 2 3 】

( 1 6 ) 本発明は、コーディングより前にオーディオ信号を処理するための本発明の信号処理ユニットまたは本発明のバイノーラルレンダラを備える、オーディオ信号をコー

10

20

30

40

50

ディングするためのオーディオエンコーダを提供する。

【0024】

(17) 本発明は、復号されたオーディオ信号を処理するための本発明の信号処理ユニットまたは本発明のバイノーラルレンダラを備える、符号化されたオーディオ信号を復号するためのオーディオデコーダを提供する。

本発明の実施形態について、添付の図面に関して説明する。

【図面の簡単な説明】

【0025】

【図1】3Dオーディオシステムの3Dオーディオエンコーダの概観を示す。

【図2】3Dオーディオシステムの3Dオーディオデコーダの概観を示す。

10

【図3】図2の3Dオーディオデコーダにおいて実装され得るフォーマット変換器を実装するための一例を示す。

【図4】図2の3Dオーディオデコーダにおいて実装され得るバイノーラルレンダラの一実施形態を示す。

【図5】室内インパルス応答 $h(t)$ の一例を示す。

【図6(a)】室内インパルス応答を用いてオーディオ入力信号を処理するための異なる可能性を示し、室内インパルス応答に従って完全なオーディオ信号を処理することを示す。

【図6(b)】室内インパルス応答を用いてオーディオ入力信号を処理するための異なる可能性を示し、初期部分と後期残響部分との別々の処理を示す。

【図7】本発明の教示に従って動作する、バイノーラルレンダラのような信号処理ユニットのブロック図を示す。

20

【図8】本発明の一実施形態によるものためのバイノーラルレンダラにおけるオーディオ信号のバイノーラル処理を概略的に示す。

【図9】本発明の一実施形態による図8のバイノーラルレンダラの周波数領域残響器における処理を概略的に示す。

【発明を実施するための形態】

【0026】

次に、本発明の手法の実施形態について説明する。以下の説明は、本発明の手法がその中で実装され得る3Dオーディオ・コーデック・システムのシステム概観から始める。

【0027】

30

図1および図2は、実施形態による3Dオーディオシステムのアルゴリズムブロックを示す。より詳細には、図1は、3Dオーディオエンコーダ100の概観を示す。オーディオエンコーダ100は、随意によっては提供され得るプリレンダラ/ミキサ回路102において、入力信号を受信し、より詳細には、複数のチャンネル信号104と、複数のオブジェクト信号106と、対応するオブジェクトメタデータ108とをオーディオエンコーダ100に提供する複数の入力チャンネルを受信する。プリレンダラ/ミキサ102によって処理されたオブジェクト信号106(信号110を参照)は、SAOCエンコーダ112(SAOC=空間オーディオ・オブジェクト・コーディング)に提供され得る。SAOCエンコーダ112は、USACエンコーダ116(USAC=統合された音声およびオーディオコーディング)に提供されるSAOCトランスポートチャンネル114を生成する。さらに、信号SAOC-SI 118(SAOC-SI=SAOCサイド情報)もUSACエンコーダ116に提供される。USACエンコーダ116は、プリレンダラ/ミキサから直接的にオブジェクト信号120をさらに受信し、ならびにチャンネル信号およびプリレンダリングされたオブジェクト信号122を受信する。オブジェクトメタデータ情報108がOAMエンコーダ124(OAM=オブジェクトメタデータ)に印加され、OAMエンコーダ124は、圧縮されたオブジェクトメタデータ情報126をUSACエンコーダに提供する。USACエンコーダ116は、上述の入力信号に基づいて、128に示されているように、圧縮された出力信号mp4を生成する。

40

【0028】

図2は、3Dオーディオシステムの3Dオーディオデコーダ200の概観を示す。図1

50

のオーディオエンコーダ 100 によって生成された符号化された信号 128 (mp4) は、オーディオデコーダ 200 において、より詳細には USAC デコーダ 202 において受信される。USAC デコーダ 202 は、受信された信号 128 を、チャンネル信号 204 と、プリレンダリングされたオブジェクト信号 206 と、オブジェクト信号 208 と、SAOC トランスポートチャンネル信号 210 とに復号する。さらに、圧縮されたオブジェクトメタデータ情報 212 と、信号 SAOC - SI 214 とが USAC デコーダ 202 によって出力される。オブジェクト信号 208 はオブジェクトレンダラ 216 に提供され、オブジェクトレンダラ 216 は、レンダリングされたオブジェクト信号 218 を出力する。SAOC トランスポートチャンネル信号 210 は SAOC デコーダ 220 に供給され、SAOC デコーダ 220 は、レンダリングされたオブジェクト信号 222 を出力する。圧縮されたオブジェクトメタ情報 212 は OAM デコーダ 224 に供給され、OAM デコーダ 224 は、レンダリングされたオブジェクト信号 218 とレンダリングされたオブジェクト信号 222 とを生成するために、オブジェクトレンダラ 216 と SAOC デコーダ 220 とにそれぞれの制御信号を出力する。デコーダは、図 2 に示されているように、チャンネル信号 228 を出力するために入力信号 204、206、218 および 222 を受信するミキサ 226 をさらに備える。チャンネル信号は、230 において示されているように、ラウドスピーカ、例えば、32 チャンネルラウドスピーカに直接出力され得る。信号 228 は、チャンネル信号 228 が変換されるべき方法を示す再生レイアウト信号を制御入力として受信するフォーマット変換回路 232 に提供され得る。図 2 に示された実施形態では、変換は、信号が、234 において示されているように 5.1 スピーカシステムに提供され得るような方法で行われると仮定する。また、チャンネル信号 228 はバイノーラルレンダラ 236 に提供され得、バイノーラルレンダラ 236 は、238 において示されているような、例えばヘッドフォンのために、2つの出力信号を生成する。

#### 【0029】

本発明の一実施形態では、図 1 および図 2 に示された符号化 / 復号システムは、チャンネル信号とオブジェクト信号と (信号 104 および 106 を参照) のコーディングのための MPEG - D USAC コーデックに基づく。大量のオブジェクトをコーディングするための効率を高めるために、MPEG SAOC 技術が使用され得る。3つのタイプのレンダラが、オブジェクトをチャンネルにレンダリングするか、チャンネルをヘッドフォンにレンダリングするか、またはチャンネルを異なるラウドスピーカセットアップにレンダリングするタスクを実施し得る (図 2 を参照、参照符号 230、234 および 238)。オブジェクト信号が、SAOC を使用して明示的に送信されるかまたはパラメータ的に符号化されるとき、対応するオブジェクトメタデータ情報 108 が圧縮され (信号 126 を参照)、3D オーディオビットストリーム 128 に多重化される。

図 1 および図 2 に示された 3D オーディオシステム全体のアルゴリズムブロックについて以下でさらに詳細に説明する。

#### 【0030】

符号化の前にチャンネル + オブジェクト入力シーンをチャンネルシーンに変換するために、随意によってはプリレンダラ / ミキサ 102 が提供され得る。機能的に、プリレンダラ / ミキサ 102 は、以下で説明するオブジェクトレンダラ / ミキサと同等である。同時にアクティブなオブジェクト信号の数に基本的に依存しないエンコーダ入力における決定性信号エントロピーを保証するために、オブジェクトのプリレンダリングが望まれることがある。オブジェクトのプリレンダリングでは、オブジェクトメタデータ送信は必要とされない。エンコーダが使用するように構成されるチャンネルレイアウトに離散オブジェクト信号がレンダリングされる。各チャンネルについてのオブジェクトの重みが、関連するオブジェクトメタデータ (OAM) から取得される。

#### 【0031】

USAC エンコーダ 116 は、ラウドスピーカチャンネル信号と、離散オブジェクト信号と、オブジェクトダウンミックス信号と、プリレンダリングされた信号とのためのコアコーデックである。USAC エンコーダ 116 は MPEG - D USAC 技術に基づく。U

10

20

30

40

50

S A Cエンコーダ 1 1 6 は、入力チャンネルおよびオブジェクト割当ての幾何学的およびセマンティック情報に基づいてチャンネルおよびオブジェクトマッピング情報を作成することによって上記の信号のコーディングを処理する。このマッピング情報は、どのように入力チャンネルおよびオブジェクトが、チャンネルペア要素 ( C P E )、単一チャンネル要素 ( S C E )、低周波効果 ( L F E ) およびクワッドチャンネル要素 ( Q C E ) のような U S A C チャンネル要素にマッピングされるかを記述し、 C P E、S C E および L F E、ならびに対応する情報がデコーダに送信される。S A O C データ 1 1 4、1 1 8 またはオブジェクトメタデータ 1 2 6 のようなすべての追加のペイロードがエンコーダのレート制御において考慮される。オブジェクトのコーディングは、レンダラのためのレート / ひずみ要件およびインタラクティビティ要件に応じて、様々な方法で可能である。実施形態によれば、以下のオブジェクトコーディング変形態が可能である。

10

**【 0 0 3 2 】**

・プリレンダリングされたオブジェクト：オブジェクト信号は、符号化の前に 2 2 . 2 チャンネル信号にプリレンダリングされミックスされる。後続のコーディングチェーンは 2 2 . 2 チャンネル信号を経験する。

**【 0 0 3 3 】**

・離散オブジェクト波形：オブジェクトはモノフォニック波形としてエンコーダに供給される。エンコーダは、チャンネル信号に加えてオブジェクトを送信するために単一チャンネル要素 ( S C E ) を使用する。復号されたオブジェクトは、受信機側においてレンダリングされ、ミックスされる。圧縮されたオブジェクトメタデータ情報は受信機 / レンダラに送信される。

20

**【 0 0 3 4 】**

・パラメトリックオブジェクト波形：オブジェクトプロパティおよびそれらの互いの関係が S A O C パラメータによって記述される。オブジェクト信号のダウンミックスは、U S A C を用いてコーディングされる。パラメータ情報が一緒に送信される。ダウンミックスチャンネルの数は、オブジェクトの数と、全体的なデータレートとに応じて選定される。圧縮されたオブジェクトメタデータ情報は S A O C レンダラに送信される。

**【 0 0 3 5 】**

オブジェクト信号のための S A O C エンコーダ 1 1 2 および S A O C デコーダ 2 2 0 は M P E G S A O C 技術に基づき得る。システムは、より少ない数の送信されたチャンネルと、O L D、I O C ( オブジェクト間コヒーレンス )、D M G ( ダウンミックス利得 ) などの追加のパラメトリックデータとに基づいていくつかのオーディオオブジェクトを再作成、修正およびレンダリングすることが可能である。追加のパラメトリックデータは、すべてのオブジェクトを個々に送信するのに必要とされるよりも著しく低いデータレートを示し、それにより、コーディングが極めて効率的になる。S A O C エンコーダ 1 1 2 は、入力としてオブジェクト / チャンネル信号をモノフォニック波形として取り、( 3 D オーディオビットストリーム 1 2 8 中にパックされる ) パラメータ情報と、( 単一チャンネル要素を使用して符号化され、送信される ) S A O C トランスポートチャンネルとを出力する。S A O C デコーダ 2 2 0 は、復号された S A O C トランスポートチャンネル 2 1 0 およびパラメータ情報 2 1 4 からオブジェクト / チャンネル信号を再構成し、再生レイアウト、復元されたオブジェクトメタデータ情報に基づいて、および随意によってはユーザ対話情報に基づいて出力オーディオシーンを生成する。

30

40

**【 0 0 3 6 】**

オブジェクト・メタデータ・コーデック ( O A M エンコーダ 1 2 4 および O A M デコーダ 2 2 4 を参照 ) は、各オブジェクトについて、3 D 空間中のオブジェクトの幾何学的位置およびボリュームを指定する関連するメタデータが、時間および空間におけるオブジェクトプロパティの量子化によって効率的にコーディングされるように提供される。圧縮されたオブジェクトメタデータ c O A M 1 2 6 はサイド情報として受信機 2 0 0 に送信される。

**【 0 0 3 7 】**

50

オブジェクトレンダラ 2 1 6 は、所与の再生フォーマットに従ってオブジェクト波形を生成するために、圧縮されたオブジェクトメタデータを利用する。各オブジェクトは、そのメタデータに従ってある出力チャンネルにレンダリングされる。このブロックの出力は、部分的な結果の合計から生じる。チャンネルベースのコンテンツならびに離散ノパラメトリックオブジェクトの両方が復号される場合、チャンネルベースの波形と、レンダリングされたオブジェクト波形とはミキサ 2 2 6 によってミックスされ、その後、得られた波形 2 2 8 が出力されるか、あるいは得られた波形 2 2 8 は、バイノーラルレンダラ 2 3 6 またはラウドスピーカ・レンダラモジュール 2 3 2 のようなポストプロセッサモジュールに供給される。

#### 【 0 0 3 8 】

バイノーラル・レンダラ・モジュール 2 3 6 は、各入力チャンネルが仮想音源によって表されるようにマルチチャンネルオーディオ素材のバイノーラルダウンミックスを生成する。処理は Q M F (直交ミラーフィルタバンク) 領域中でフレームごとに行われ、バイノーラル化は、測定されたバイノーラル室内インパルス応答に基づく。

#### 【 0 0 3 9 】

ラウドスピーカレンダラ 2 3 2 は、送信されたチャンネル構成 2 2 8 と所望の再生フォーマットとの間で変換する。それは「フォーマット変換器」と呼ばれることもある。フォーマット変換器は、より低い数の出力チャンネルへの変換を実施し、すなわち、それはダウンミックスを生じる。

#### 【 0 0 4 0 】

図 3 は、フォーマット変換器 2 3 2 を実装するための一例を示す。ラウドスピーカレンダラとも呼ばれる、フォーマット変換器 2 3 2 は、送信機チャンネル構成と所望の再生フォーマットとの間で変換する。フォーマット変換器 2 3 2 は、より低い数の出力チャンネルへの変換を実施し、すなわち、それはダウンミックス (D M X) プロセス 2 4 0 を実施する。好ましくは Q M F 領域中で動作するダウンミキサ 2 4 0 は、ミキサ出力信号 2 2 8 を受信し、ラウドスピーカ信号 2 3 4 を出力する。コントローラとも呼ばれる、構成器 2 4 2 が提供され得、この構成器 2 4 2 は、制御入力として、ミキサ出力レイアウト、すなわち、ミキサ出力信号 2 2 8 によって表されるデータがそのために決定されるレイアウトを示す信号 2 4 6 と、所望の再生レイアウトを示す信号 2 4 8 とを受信する。この情報に基づいて、コントローラ 2 4 2 は、好ましくは自動的に、入力フォーマットと出力フォーマットとの所与の合成のための最適化されたダウンミックス行列を生成し、これらの行列をダウンミキサ 2 4 0 に適用する。フォーマット変換器 2 3 2 は、標準のラウドスピーカ構成、ならびに非標準のラウドスピーカ位置をもつランダム構成を可能にする。

#### 【 0 0 4 1 】

図 4 は、図 2 のバイノーラルレンダラ 2 3 6 の一実施形態を示す。バイノーラル・レンダラ・モジュールは、マルチチャンネルオーディオ素材のバイノーラルダウンミックスを提供し得る。バイノーラル化は、測定されたバイノーラル室内インパルス応答に基づき得る。室内インパルス応答は、現実の室内の音響プロパティの「指紋」と見なされ得る。室内インパルス応答は測定され、記憶され、任意の音響信号にはこの「指紋」が提供され得、それにより、聴取者において、室内インパルス応答に関連する室内の音響プロパティのシミュレーションが可能になる。バイノーラルレンダラ 2 3 6 は、頭部伝達関数またはバイノーラル室内インパルス応答 (B R I R) を使用して出力チャンネルを 2 つのバイノーラルチャンネルにレンダリングするように構成またはプログラムされ得る。例えば、モバイルデバイスでは、ヘッドフォンまたはそのようなモバイルデバイスに取り付けられたラウドスピーカのためのバイノーラルレンダリングが望まれる。そのようなモバイルデバイスでは、制約により、デコーダとレンダリングの複雑さとを制限することが必要であり得る。そのような処理シナリオでは無相関を省略することに加えて、ダウンミキサ 2 5 0 を使用して、中間ダウンミックス信号 2 5 2 への、すなわち、より低い数の出力チャンネルへのダウンミックスを最初に実施することが好適であり得、それにより、実際のバイノーラル変換器 2 5 4 のためにより低い数の入力チャンネルが得られる。例えば、2 2 . 2 チャンネル素材

10

20

30

40

50

が、ダウンミキサ 250 によって 5.1 中間ダウンミックスにダウンミックスされ得るか、または、代替的に、中間ダウンミックスが、一種の「ショートカット」モードで図 2 の SAOC デコーダ 220 によって直接計算され得る。その場合、バイノーラルレンダリングは、2.2 入力チャンネルが直接レンダリングされるべきであった場合、44 個の HRTF (頭部伝達関数) または BRIR 関数を適用するのとは対照的に、5 つの個々のチャンネルを異なる位置においてレンダリングするために 10 個の HRTF または BRIR 関数を適用するだけでよい。バイノーラルレンダリングに必要な畳み込み演算は多くの処理電力を必要とし、したがって、許容できるオーディオ品質を依然として獲得しながらこの処理電力を低減することは、モバイルデバイスにとって特に有用である。バイノーラルレンダラ 236 は、(LFE チャンネルを除く) 各入力チャンネルが仮想音源によって表されるように、マルチチャンネルオーディオ素材 228 のバイノーラルダウンミックス 238 を生成する。処理は QMF 領域中でフレームごとに行われ得る。バイノーラル化は、測定されたバイノーラル室内インパルス応答に基づき、直接音および初期反射は、QMF 領域の上で高速畳み込みを使用して擬似 FFT 領域において畳み込み手法を介してオーディオ素材に転写され得るが、後期残響は別々に処理され得る。

#### 【0042】

図 5 は、室内インパルス応答  $h(t)$  300 の一例を示す。室内インパルス応答は、3 つの成分、直接音 301、初期反射 302 および後期残響 304 を含む。このようにして、室内インパルス応答は、インパルスが発されたときの密閉型残響音響空間の反射挙動を記述する。初期反射 302 は、増加する密度を伴う個別の反射であり、個々の反射がもはや区別され得ないインパルス応答の部分は、後期残響 304 と呼ばれる。直接音 301 は、室内インパルス応答において容易に識別され得、初期反射から分離され得るが、初期反射 302 から後期残響 304 への遷移はあまり明白でない。

#### 【0043】

上記で説明したように、バイノーラルレンダラ、例えば図 2 にそれが示されているようなバイノーラルレンダラでは、室内インパルス応答に従ってマルチチャンネルオーディオ入力信号を処理するための様々な手法が知られている。

#### 【0044】

図 6 は、室内インパルス応答を用いてオーディオ入力信号を処理するための異なる可能性を示す。図 6 (a) は、室内インパルス応答に従って完全なオーディオ信号を処理することを示し、図 6 (b) は、初期部分と後期残響部分との別々の処理を示す。図 6 (a) に示されているように、入力信号 400、例えばマルチチャンネルオーディオ入力信号が受信され、プロセッサ 402 に印加され、プロセッサ 402 は、室内インパルス応答 (図 5 を参照) を用いてマルチチャンネルオーディオ入力信号 400 の完全な畳み込みを可能にするように構成またはプログラムされ、これは、図示された実施形態では、2 チャンネルオーディオ出力信号 404 を生じる。上述のように、この手法は、インパルス応答全体に畳み込みを使用することが計算量的に極めてコストがかかるので、不利であると見なされる。したがって、別の手法によれば、図 6 (b) に示されているように、図 6 (a) に関して説明した室内インパルス応答を用いて完全な畳み込みを適用することによってマルチチャンネルオーディオ入力信号全体を処理する代わりに、処理は、室内インパルス応答 300 の初期部分 301、302 (図 5 を参照) と、後期残響部分 302 とに関して分離される。より詳細には、図 6 (b) に示されているように、マルチチャンネルオーディオ入力信号 400 が受信されるが、信号は、初期部分を処理するために、すなわち、図 5 に示された室内インパルス応答 300 中の直接音 301 と初期反射 302 とに従ってオーディオ信号を処理するために、第 1 のプロセッサ 406 に並列に印加される。マルチチャンネルオーディオ入力信号 400 はまた、室内インパルス応答 300 の後期残響 304 に従ってオーディオ信号を処理するためにプロセッサ 408 に適用される。図 6 (b) に示された実施形態では、マルチチャンネルオーディオ入力信号は、マルチチャンネル信号 400 を、より低い数のチャンネルを有する信号にダウンミックスするためにダウンミキサ 410 にも印加され得る。ダウンミキサ 410 の出力は、次いでプロセッサ 408 に印加される。プロセッサ 4

10

20

30

40

50

06とプロセッサ408との出力は412において合成されて2チャンネルオーディオ出力信号404'が生成される。

【0045】

バイノーラルレンダラでは、上述のように、主に計算複雑さの低減のために、直接音および初期反射を後期残響とは別個に処理することが望まれることがある。直接音と初期反射との処理は、例えば、プロセッサ406(図6(b))によって実行される畳み込み手法によってオーディオ信号に転写され得るが、後期残響は、プロセッサ408によって行われる合成残響と置換され得る。全体的なバイノーラル出力信号404'は、その場合、プロセッサ406によって提供される畳み込み結果と、プロセッサ408によって提供される合成残響信号との合成である。

10

【0046】

この処理は従来技術文献[1]にも記載している。上記で説明した手法の結果は、完全なインパルス応答の畳み込み、図6(a)に関して説明した完全変換手法の結果とできる限り知覚的に同等であるはずである。しかしながら、オーディオ信号、またはより一般的な、オーディオ素材が、インパルス応答の直接音と初期反射部分とで畳み込まれた場合、結果として異なるチャンネルが合計されて、聴取者の一方の耳への再生信号に関連付けられた全体的な音信号が形成される。残響は、しかしながら、この全体的な信号から計算されないが、概して、元の入力オーディオ信号の1つのチャンネルまたはダウンミックスの残響信号である。本発明の発明者によって、したがって、後期残響は、プロセッサ406によって提供される畳み込み結果に適切に適合していないと判断されている。適切なレベルの残響は入力オーディオ信号と室内インパルス応答300の両方に依存することが分かっている。インパルス応答の影響は、プロセッサ408の一部であり得る残響器の入力パラメータとして残響特性を使用することによって達成され、これらの入力パラメータは、測定されたインパルス応答、例えば周波数依存残響時間および周波数依存エネルギー測度の分析から取得される。これらの測度は、概して、例えばオクターブフィルタバンク分析においてエネルギーおよびRT60残響時間を計算することによって、単一のインパルス応答から決定され得るか、または複数のインパルス応答分析の結果の平均値である。

20

【0047】

しかしながら、残響器に提供されるこれらの入力パラメータにもかかわらず、図6(b)に関して説明するような合成残響手法を使用するとき、残響への入力オーディオ信号の影響は十分に保持されないことが分かっている。例えば、合成残響テールを生成するために使用されるダウンミックスにより、入力オーディオ信号の影響は失われる。残響の得られるレベルは、したがって、特に入力信号が複数のチャンネルを含む場合、完全な畳み込み手法の結果と知覚的に同等でない。

30

【0048】

これまで、後期残響の量を完全な畳み込み手法の結果と比較するかまたはそれを畳み込み結果に一致させる既知の手法はない。後期残響の品質をレーティングするかまたはそれがどれくらい自然に聞こえるかをレーティングすることを試みるいくつかの技法がある。例えば、1つの方法では、自然な音響残響のためのラウドネス測度が定義され、これは、ラウドネスモデルを使用して残響の知覚されるラウドネスを予測する。この手法は従来技術文献[2]に記載されており、レベルはターゲット値に適合され得る。この手法の欠点は、複雑で不正確である人間の聴覚のモデルにそれが依拠することである。それはまた、完全な畳み込み結果を使用して見つけられ得る後期残響にスケールリングファクタを提供するためのターゲットラウドネスを必要とする。

40

【0049】

従来技術文献[3]に記載した別の方法では、人工的残響品質テストのための相互相関基準が使用される。しかしながら、これは、異なる残響アルゴリズムをテストするためにのみ適用可能であるが、マルチチャンネルオーディオには適用可能でなく、バイノーラルオーディオには適用可能でなく、後期残響のスケールリングを適格とするためには適用可能でない。

50

## 【 0 0 5 0 】

別の可能な手法は、スケーリングファクタとしての考慮される耳における入力チャンネルの数の使用に対するものであるが、全体的な音信号の知覚される振幅は、異なるオーディオチャンネルの相関に依存し、チャンネルの数だけに依存しないので、これは知覚的に正しいスケーリングを与えない。

## 【 0 0 5 1 】

したがって、本発明の手法によれば、入力オーディオ信号に従って残響のレベルを適応させる信号依存スケーリング方法が提供される。上述のように、残響の知覚されるレベルは、バイノーラルレンダリングのために完全な畳み込み手法を使用するときの残響のレベルと一致することが望まれ、したがって、適切なレベルの残響のための測度の決定は、良好な音品質を達成するために重要である。実施形態によれば、オーディオ信号は、室内インパルス応答の初期部分と後期残響とを用いて別々に処理され、後期残響を処理することは、スケーリングされた残響信号を生成することを含み、スケーリングはオーディオ信号に依存する。オーディオ信号の処理された初期部分と、スケーリングされた残響信号とは、出力信号に合成される。一実施形態によれば、スケーリングは、オーディオ信号の1つまたは複数の入力チャンネルの状態（例えば、入力チャンネルの数、アクティブ入力チャンネルの数および/または入力チャンネル中のアクティビティ）に依存する。別の実施形態によれば、スケーリングは、オーディオ信号についてのあらかじめ定義されたまたは計算される相関測度に依存する。代替実施形態は、1つまたは複数の入力チャンネルの状態と、あらかじめ定義されたまたは計算される相関測度との合成に基づいてスケーリングを実施し得る。

## 【 0 0 5 2 】

実施形態によれば、スケーリングされた残響信号は、オーディオ信号の1つまたは複数の入力チャンネルの状態に基づいて、あるいはオーディオ信号についてのあらかじめ定義されたまたは計算される相関測度に基づいて、あるいはそれらの合成に基づいて決定された利得ファクタを適用することによって生成され得る。

## 【 0 0 5 3 】

実施形態によれば、オーディオ信号を別々に処理することは、第1のプロセス中に室内インパルス応答300の初期反射部分301、302を用いてオーディオ信号を処理することと、第1のプロセスとは異なり別々である第2のプロセス中に室内インパルス応答300の拡散残響304を用いてオーディオ信号を処理することとを含む。第1のプロセスから第2のプロセスに変化することは遷移時間において行われる。さらなる実施形態によれば、第2のプロセスにおいて、拡散（後期）残響304は合成残響と置換され得る。この場合、第1のプロセスに適用される室内インパルス応答は初期反射部分300、302（図5を参照）のみを含んでおり、後期拡散残響304は含まれない。

## 【 0 0 5 4 】

以下で、入力オーディオ信号の相関分析に基づいて利得ファクタがそれに従って計算される本発明の手法の一実施形態についてさらに詳細に説明する。図7は、本発明の教示に従って動作する、バイノーラルレンダラのような信号処理ユニットのブロック図を示す。バイノーラルレンダラ500は、N個のチャンネルを含むオーディオ信号 $x[k]$ を入力504から受信するプロセッサ502を含む第1の分岐を備える。プロセッサ502は、バイノーラルレンダラの一部であるとき、入力信号504を処理して出力信号 $x_{c.o.n.v}[k]$ を生成する。より詳細には、プロセッサ502は、複数の記録されたバイノーラル室内インパルス応答を保持している外部データベース508からプロセッサ502に提供され得る室内インパルス応答の直接音と初期反射とを用いて、オーディオ入力信号504の畳み込みを引き起こす。プロセッサ502は、上述のように、データベース508によって提供されるバイノーラル室内インパルス応答に基づいて動作し得、それにより、ただ2つのチャンネルを有する出力信号502を生成する。出力信号506はプロセッサ502から加算器510に提供される。入力信号504は、残響器プロセッサ514とダウンミキサ516とを含む残響分岐512にさらに提供される。ダウンミックスされた入力信号は残響器514に提供され、残響器514は、それぞれデータベース518および5

20中に保持された残響RT60および残響エネルギーのような残響器パラメータに基づいて、ただ2つのチャンネルを含み得る残響器514の出力において残響信号 $r[k]$ を生成する。データベース518および520に記憶されたパラメータは、図7の破線で示されているように、適切な分析522によって、記憶されたバイノーラル室内インパルス応答から取得され得る。

#### 【0055】

残響分岐512は相関分析プロセッサ524をさらに含み、相関分析プロセッサ524は、入力信号504を受信し、その出力において利得ファクタ $g$ を生成する。さらに、残響器514と加算器510との間に結合された利得段526が提供される。利得段526は利得ファクタ $g$ によって制御され、それにより、利得段526の出力において、スケールリングされた残響信号 $r_g[k]$ を生成し、この残響信号 $r_g[k]$ は加算器510に印加される。加算器510は、初期処理部分と残響信号とを合成して、同じく2つのチャンネルを含む出力信号 $y[k]$ を提供する。随意によっては、残響分岐512は、いくつかのオーディオフレームにわたって利得ファクタを平滑化するために、プロセッサ524と利得段との間に結合された低域フィルタ528を備え得る。随意によっては、スケールリングされた残響信号が室内インパルス応答中の初期反射と残響との間の遷移に一致するように、スケールリングされた残響信号を遅延させるために、利得段526の出力と加算器510との間に遅延要素530も提供され得る。

#### 【0056】

上記で説明したように、図7は、後期残響とは別々に直接音と初期反射とを処理するバイノーラルレンダラのブロック図である。分かるように、バイノーラル室内インパルス応答の直接反射と初期反射とを用いて処理された入力信号 $x[k]$ は、信号 $x_{conv}[k]$ を生じる。この信号は、図示のように、それを残響信号成分 $r_g[k]$ に加算するために加算器510に転送される。この信号は、入力信号 $x[k]$ のダウンミックス、例えばステレオダウンミックスを残響器514に供給し、続いて、ダウンミックスの残響信号 $r[k]$ と利得ファクタ $g$ とを受信する乗算器または利得段526に供給することによって生成される。利得ファクタ $g$ は、プロセッサ524によって実行される入力信号 $x[k]$ の相関分析によって取得され、上述のように、低域フィルタ528によって時間とともに平滑化され得る。スケールリングまたは重み付けされた残響成分は、その開始が初期反射から後期残響への遷移点と一致するように、随意によっては遅延要素530によって遅延され得、このようにして、加算器510の出力において出力信号 $y[k]$ が取得される。

#### 【0057】

図7に示されたマルチチャンネルバイノーラルレンダラは合成2チャンネル後期残響を導入し、従来の手法の上記で説明した欠点を克服するために、本発明の手法によれば、合成後期残響は、知覚を完全な畳み込み手法の結果に一致させるために、利得ファクタ $g$ によってスケールリングされる。聴取者の耳における複数のチャンネルの重ね合わせ（例えば最高2.2）は相関依存である。そのような理由で、後期残響は入力信号チャンネルの相関に従ってスケールリングされ得、本発明の手法の実施形態は、後期残響の適切な振幅を決定する相関ベースの時間依存スケールリング方法を提供する。

#### 【0058】

スケールリングファクタを計算するために、相関係数に基づく相関測度が導入され、実施形態によれば、2次元時間周波数領域、例えばQMF領域において定義される。各多次元オーディオフレームについて-1と1との間の相関値が計算され、各オーディオフレームは、周波数帯域の数 $N$ と、フレームごとのタイムスロットの数 $M$ と、オーディオチャンネルの数 $A$ とによって定義される。フレームごとに耳ごとに1つのスケールリングファクタが取得される。

#### 【0059】

以下で、本発明の手法の一実施形態についてさらに詳細に説明する。まず第1に、図7の相関分析プロセッサ524において使用された相関測度への参照が行われる。相関測度は、この実施形態によれば、以下のように、2つの変数 $X$ 、 $Y$ の共分散をそれらの標準偏

10

20

30

40

50

差の積で除算することによって計算されるピアソンの積率係数（相関係数としても知られる）に基づく。

【数 7】

$$\rho_{\{X,Y\}} = \frac{E\{(X - \bar{X}) \cdot (Y - \bar{Y})\}}{\sigma_X \cdot \sigma_Y}$$

ただし

$E\{\cdot\}$  = 期待値演算子

$\{X, Y\}$  = 相関係数、

$\sigma_X, \sigma_Y$  = 変数  $X, Y$  の標準偏差

10

説明する実施形態によるこの処理は、時間周波数領域、例えば Q M F 領域における 2 次元に伝達される。この 2 次元はタイムスロットおよび Q M F 帯域である。データは時間 - 周波数領域中でしばしば符号化され、送信もされるので、この手法は妥当である。期待値演算子は、いくつかの時間および / または周波数サンプルにわたる平均演算と置換され、したがって、( 0 , 1 ) の範囲内の 2 つのゼロ平均変数  $x_m, x_n$  間の時間周波数相関測度は以下のように定義される。

【数 8】

$$\rho[m, n] = \left| \frac{1}{(N - 1)} \cdot \frac{\sum_i \sum_j x_m[i, j] \cdot x_n[i, j]^*}{\sum_j \sigma(x_m[j]) \cdot \sigma(x_n[j])} \right|$$

20

ただし

$[m, n]$  = 相関係数、

$(x_m[j])$  = チャンネル  $m$  の 1 つのタイムスロット  $j$  にわたる標準偏差、

$(x_n[j])$  = チャンネル  $n$  の 1 つのタイムスロット  $j$  にわたる標準偏差、

$x_m, x_n$  = ゼロ平均変数、

$j [1, N]$  = 周波数帯域、

$j [1, M]$  = タイムスロット、

$m, n [1, K]$  = チャンネル、

\* = 複素共役。

30

1 つのオーディオフレームの複数のチャンネル合成 ( $m, n$ ) についてのこの係数の計算の後に、 $[m, n, t_i]$  の値は、複数の相関値  $[m, n, t_i]$  の平均を取る（または複数の相関値  $[m, n, t_i]$  を平均化する）ことによって単一の相関測度  $\rho_m(t_i)$  に合成される。オーディオフレームは 3 2 個の Q M F タイムスロットを含み得、 $t_i$  はそれぞれのオーディオフレームを示すことに留意されたい。上記の処理は、1 つのオーディオフレームについて以下のように要約され得る。

( i ) 最初に、サイズ  $[N, M, K]$  を有するオーディオまたはデータフレーム  $x$  の  $k$  個のチャンネルごとの全体的な平均値

40

【数 9】

$$\bar{x}(k)$$

が計算され、実施形態によれば、すべての  $k$  個のチャンネルは残響器の 1 つの入力チャンネルにダウンミックスされる。

( i i ) 対応するチャンネルから値

【数 1 0】

50

$$\bar{x}(k)$$

を減算することによってゼロ平均オーディオまたはデータフレームが計算される。

( i i i ) 複数のチャンネル合成 ( m , n ) のそれぞれについて、定義された相関係数  $[ m , n ]$  が計算される。

( i v ) (例えばゼロ除算によって誤って計算された値を除く) 複数の相関係数  $[ m , n ]$  の平均として平均相関値が計算される。

【 0 0 6 0 】

上記で説明した実施形態によれば、スケーリングは、オーディオ信号についての計算された相関測度に基づいて決定された。例えば、個々に現在処理されているオーディオ信号について相関測度を取得することが望まれるとき、必要とされる追加の計算リソースにもかかわらず、これは有利である。

10

【 0 0 6 1 】

ただし、本発明はそのような手法に限定されない。他の実施形態によれば、相関測度を計算するのではなく、あらかじめ定義された相関測度も使用され得る。あらかじめ定義された相関測度を使用することは、プロセスにおける計算複雑さを低減するので、有利である。あらかじめ定義された相関測度は、複数のオーディオ信号の分析に基づいて経験的に決定され得る、例えば 0 . 1 ~ 0 . 9 の、固定値を有し得る。そのような場合、相関分析 5 2 4 は省略され得、利得段の利得は適切な制御信号によって設定され得る。

20

【 0 0 6 2 】

他の実施形態によれば、スケーリングは、オーディオ信号の 1 つまたは複数の入力チャンネルの状態 (例えば、入力チャンネルの数、アクティブ入力チャンネルの数および/または入力チャンネル中のアクティビティ) に依存し得る。スケーリングは、低減した計算オーバーヘッドにより入力オーディオ信号から容易に決定され得るので、これは有利である。例えば、スケーリングは、元のオーディオ信号と比較して低減した数のチャンネルを含む、現在考慮されているダウンミックスチャンネルにダウンミックスされる元のオーディオ信号中のチャンネルの数を単に決定することによって決定され得る。代替的に、現在考慮されているダウンミックスチャンネルにダウンミックスされるアクティブチャンネル (現在オーディオフレーム中で何らかのアクティビティを示すチャンネル) の数は、残響信号をスケーリングするための基礎を形成し得る。これはブロック 5 2 4 において行われ得る。

30

【 0 0 6 3 】

以下で、オーディオ信号の 1 つまたは複数の入力チャンネルの状態に基づいて、および (上記で説明したように固定であるかあるいは計算される) 相関測度に基づいて残響信号のスケーリングを決定する実施形態について詳細に説明する。そのような実施形態によれば、利得ファクタまたは利得またはスケーリングファクタ  $g$  は以下のように定義される。

【数 1 1】

$$g = c_u + \rho \cdot (c_c - c_u)$$

40

【数 1 2】

$$c_u = 10^{\frac{10 \cdot \log_{10}(K_{in})}{20}} = \sqrt{K_{in}}$$

【数 1 3】

50

$$c_c = 10^{\frac{20 \cdot \log_{10}(K_{in})}{20}} = K_{in}$$

ただし

= オーディオ信号についてのあらかじめ定義されたまたは計算される相関係数、

$c_u$ 、 $c_c$  = オーディオ信号の1つまたは複数の入力チャンネルの状態を示すファクタであり、 $c_u$ は、全体として無相関なチャンネルを指し、 $c_c$ は、全体として相関するチャンネルに関係し、

10

$K_{in}$  = アクティブ非ゼロまたは固定ダウンミックスチャンネルの数。

$c_u$ は、ダウンミックスされたチャンネルが全体として無相関である（チャンネル間依存性がない）場合に適用されるファクタである。1つまたは複数の入力チャンネルの状態のみを使用する場合、 $g = c_u$ 、および、あらかじめ定義された固定相関係数は0に設定される。 $c_c$ は、ダウンミックスされたチャンネルが全体として相関する（信号が互いの重み付けされたバージョン（+位相シフトおよびオフセット）である）場合に適用されるファクタである。1つまたは複数の入力チャンネルの状態のみを使用する場合、 $g = c_c$ 、および、あらかじめ定義された固定相関係数は1に設定される。これらのファクタは、（（アクティブ）チャンネルの数に応じて）オーディオフレーム中で後期残響の最小スケーリングと最大スケーリングとを記述する。

20

「チャンネル数」 $K_{in}$ は、実施形態によれば、以下のように定義される。どの入力チャンネルがどのダウンミックスチャンネル中に含まれるかを定義するダウンミックス行列 $Q$ （サイズ $M \times 2$ 、ただし、 $M$ はオーディオ入力素材の入力チャンネルの数であり、例えば5.1セットアップでは6つのチャンネルである）を使用して、マルチチャンネルオーディオ信号がステレオダウンミックスにダウンミックスされる。

ダウンミックス行列 $Q$ の一例は以下のとおりであり得る。

【数14】

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.7071 & 0.7071 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

30

2つのダウンミックスチャンネルの各々について、スケーリング係数は以下のように計算される。

$$g = f(c_c, c_u, avg) = c_u + avg \cdot (c_c - c_u)$$

ただし、 $avg$ は、いくつかの $K_{in} \cdot K_{in}$ チャンネル合成 $[m, n]$ についてのすべての相関係数 $[m, n]$ の平均/平均値であり、 $c_c$ 、 $c_u$ は、以下のとおりであり得るチャンネル数 $K_{in}$ に依存する。

40

・ $K_{in}$ は、現在考慮されているダウンミックスチャンネルにダウンミックスされるチャンネルの数 $k$   $[1, 2]$ （0に等しくない値を含んでいる列 $k$ 中のダウンミックス行列 $Q$ 中の行の数）であり得る。この数は、ダウンミックス行列 $Q$ が、1つの入力チャンネル構成についてあらかじめ定義され、1つのオーディオ入力信号の全長にわたって変化しないので、時不変である。

例えば5.1入力信号を考慮するとき、以下が適用される。

チャンネル1、3、4がダウンミックスチャンネル1にダウンミックスされ（上記の行列 $Q$ を参照）、

50

$K_{in}$  はフレームごとに  $= 3$  (3つのチャンネル) である。

・  $K_{in}$  は、現在考慮されているダウンミックスチャンネルにダウンミックスされるアクティブチャンネルの数  $k \in [1, 2]$  (現在オーディオフレーム中でアクティビティがあり、また、列  $k$  中のダウンミックス行列  $Q$  の対応する行が、アクティブチャンネルと  $Q$  の列  $k$  中の等しくない要素との交差において、0 チャンネルの数に等しくない値を含んでいる、入力チャンネルの数) であり得る。この数は、 $Q$  が同じままである場合でも、信号アクティビティが時間とともに変化し得るので、1つのオーディオ入力信号の全長にわたって時変であり得る。

【0064】

例えば5.1入力信号を考慮するとき、以下が適用される。

チャンネル1、3、4がダウンミックスチャンネル1にダウンミックスされ(上記の行列  $Q$  を参照)、

フレーム  $n$  では、

アクティブチャンネルはチャンネル1、2、4であり、

$K_{in}$  は交差  $\{1, 4\}$  におけるチャンネルの数であり、

$$K_{in}(n) = 2$$

フレーム  $n+1$  では、

アクティブチャンネルはチャンネル1、2、3、4であり、

$K_{in}$  は交差  $\{1, 3, 4\}$  におけるチャンネルの数であり、

$$K_{in}(n+1) = 3.$$

(あらかじめ定義されたフレーム中の)オーディオチャンネルは、それが、あらかじめ定義されたフレーム内に、プリセットしきい値を超える振幅またはエネルギーを有する場合はアクティブであると見なされ得、例えば、実施形態によれば、(あらかじめ定義されたフレーム中の)オーディオチャンネル中のアクティビティは以下のように定義され得る。

・フレーム中の(時間領域、QMF領域などにおける)信号の絶対振幅の合計または最大値が0よりも大きいか、あるいは

・フレーム中の信号エネルギー(時間領域またはQMF領域における振幅の2乗絶対値)の合計または最大値が0よりも大きい。

0の代わりに、0よりも大きい(最大エネルギーまたは振幅に対する)別のしきい値、例えば0.01のしきい値も使用され得る。

【0065】

実施形態によれば、アクティブチャンネルの数(時間変動)またはダウンミックスチャンネル中の含まれるチャンネル(0に等しくないダウンミックス行列)の固定数  $K_{in}$  に依存する各耳についての利得ファクタが提供される。ファクタは、全体として無相関な場合と全体として相関する場合との間で線形的に増加すると仮定する。全体として無相関は、チャンネル間依存性なしを意味し(相関値は0である)、全体として相関は、信号が互いの重み付けされたバージョンであることを意味する(オフセットの位相差がある場合、相関値は1である)。

【0066】

上述のように、利得またはスケーリングファクタ  $g$  は、低域フィルタ528によってオーディオフレームにわたって平滑化され得る。低域フィルタ528は、以下のように、フレームサイズ  $k$  のために  $g_s(t)$  の平滑化された利得ファクタを生じる  $t_s$  の時定数を有し得る。

【数15】

$$g_s(t_i) = c_{s,old} \cdot g_s(t_i - 1) + c_{s,new} \cdot g$$

【数16】

10

20

30

40

50

$$c_{s,old} = e^{-\left(\frac{1}{f_s \cdot k} \cdot t_s\right)}$$

【数 17】

$$c_{s,new} = 1 - c_{s,old}$$

10

ただし

$t_s$  = [ s ]での低域フィルタの時定数

$t_i$  = フレーム  $t_i$ におけるオーディオフレーム

$g_s$  = 平滑化された利得ファクタ

$k$  = フレームサイズ、および

$f_s$  = [ Hz ]でのサンプリング周波数

フレームサイズ  $k$  は、時間領域サンプル中のオーディオフレームのサイズ、例えば 2048 個のサンプルであり得る。

オーディオフレーム  $x(t_i)$  の左チャンネル残響信号は、その場合、ファクタ  $g_{s, left}(t_i)$  でスケーリングされ、右チャンネル残響信号はファクタ  $g_{s, right}(t_i)$  でスケーリングされる。スケーリングファクタは、残響器に供給されるステレオダウンミックスの左チャンネル中に存在するチャンネル（のアクティブ非 0 または総数）の数として  $K_{in}$  を用いて 1 回計算され、それにより、スケーリングファクタ  $g_{s, left}(t_i)$  が得られる。次いで、スケーリングファクタは、残響器に供給されるステレオダウンミックスの右チャンネル中に存在するチャンネル（のアクティブ非 0 または総数）の数として  $K_{in}$  を用いてもう 1 回計算され、それにより、スケーリングファクタ  $g_{s, right}(t_i)$  が得られる。残響器は、オーディオフレームのステレオ残響バージョンを戻す。残響バージョンの左チャンネル（または残響器の入力の左チャンネル）は  $g_{s, left}(t_i)$  でスケーリングされ、残響バージョンの右チャンネル（または残響器の入力の右チャンネル）は  $g_{s, right}(t_i)$  でスケーリングされる。

20

30

スケーリングされた人工的（合成）後期残響は、直接音と初期反射とを用いて処理されている信号 506 に加算されるために、加算器 510 に印加される。

上述のように、本発明の手法は、実施形態によれば、オーディオ信号のバイノーラル処理のためのバイノーラルプロセッサにおいて使用され得る。以下で、オーディオ信号のバイノーラル処理の一実施形態について説明する。バイノーラル処理は、復号された信号を、ヘッドフォンを介して聴取されたときにサラウンド音エクスペリエンスを提供するバイノーラルダウンミックス信号に変換する、デコーダプロセスとして行われ得る。

【0067】

図 8 は、本発明の一実施形態によるオーディオ信号のバイノーラル処理のためのバイノーラルレンダラ 800 の概略図を示す。図 8 はまた、バイノーラルレンダラにおける QMF 領域処理の概観を提供する。入力 802 において、バイノーラルレンダラ 800 は、処理されるべきオーディオ信号、例えば、N 個のチャンネルと 64 個の QMF 帯域とを含む入力信号を受信する。さらに、バイノーラルレンダラ 800 は、オーディオ信号の処理を制御するためのいくつかの入力パラメータを受信する。入力パラメータは、 $2 \times N$  チャンネルと 64 個の QMF 帯域とのためのバイノーラル室内インパルス応答 (BRIR) 804、BRIR 804 の初期反射部分を用いたオーディオ入力信号の畳み込みのために使用される最大帯域の指示  $K_{max}$  806、ならびに上述の残響器パラメータ 808 および 810 (RT60 および残響エネルギー) を含む。バイノーラルレンダラ 800 は、受信された BRIR 804 の初期部分を用いて入力オーディオ信号 802 を処理するための高速畳み込

40

50

みプロセッサ 812 を備える。プロセッサ 812 は、出力において、2つのチャンネルと  $K_{max}$  個の QMF 帯域とを含む初期処理信号 814 を生成する。バイノーラルレンダラ 800 は、高速畳み込みプロセッサ 812 を有する初期処理分岐のほかに、2つの残響器 816 a および 816 b を含む残響分岐をも備え、各残響器は、入力パラメータとして RT60 情報 808 と残響エネルギー情報 810 とを受信する。残響分岐は、両方とも同じく入力オーディオ信号 802 を受信するステレオ・ダウンミックス・プロセッサ 818 および 10 相関分析プロセッサ 820 をさらに含む。さらに、ステレオ・ダウンミックス・プロセッサ 818 によって提供されるダウンミックス信号 822 の利得を制御するために、ステレオ・ダウンミックス・プロセッサ 818 とそれぞれの残響器 816 a および 816 b との間に 2つの利得段 821 a および 821 b が提供される。ステレオ・ダウンミックス・プロセッサ 818 は、入力信号 802 に基づいて、2つの帯域と 64 個の QMF 帯域とを有するダウンミックス信号 822 を提供する。利得段 821 a および 821 b の利得は、相関分析プロセッサ 820 によって提供されるそれぞれの制御信号 824 a および 824 b によって制御される。利得制御されたダウンミックス信号はそれぞれの残響器 816 a および 816 b に入力され、それぞれの残響信号 826 a、826 b が生成される。初期処理信号 814 と残響信号 826 a、826 b とはミキサ 828 によって受信され、ミキサ 828 は、受信された信号を、2つのチャンネルと 64 個の QMF 帯域とを有する出力オーディオ信号 830 に合成する。さらに、本発明によれば、高速畳み込みプロセッサ 812 と残響器 816 a および 816 b とは、上記で説明したように決定される初期部分から後期残響への室内インパルス応答 804 中の遷移を示す追加の入力パラメータ 832 を受信 20 する。

#### 【0068】

バイノーラル・レンダラ・モジュール 800 (例えば、図 2 または図 4 のバイノーラルレンダラ 236) は、入力 802 として、復号されたデータストリームを有する。信号は、ISO/IEC 14496-3:2009、4.B.18.2 節に概説されその修正が ISO/IEC 14496-3:2009、8.6.4.2 節に述べられている、QMF 分析フィルタバンクによって処理される。レンダラモジュール 800 はまた、QMF 領域入力データを処理し得、この場合、分析フィルタバンクは省略され得る。バイノーラル室内インパルス応答 (BRIR) 804 は複素 QMF 領域フィルタとして表される。時間領域バイノーラル室内インパルス応答から複素 QMF フィルタ表現への変換は、ISO/IEC FDIS 23003-1:2006、アネックス B に概説されている。複素 QMF 領域では、BRIR 804 が初期反射部分 301、302 (図 5 を参照) のみを含み、かつ後期拡散残響 304 が含まれないように、BRIR 804 は一定数のタイムスロットに制限される。初期反射から後期残響への遷移点 832 は、上記で説明したように、例えば、バイノーラル処理の前処理ステップにおける BRIR 804 の分析によって決定される。QMF 領域オーディオ信号 802 および QMF 領域 BRIR 804 は、次いで、バイノーラル処理を実施するために帯域的な高速畳み込み 812 によって処理される。2チャンネル QMF 領域後期残響 826 a、826 b を生成するために QMF 領域残響器 816 a、816 b が使用される。残響モジュール 816 a、816 b は、残響の特性を適応させるために周波数依存残響時間 808 とエネルギー値 810 とのセットを使用する。残響の波形は、オーディオ入力信号 802 のステレオダウンミックス 818 に基づき、それは、マルチチャンネルオーディオ信号 802 の相関分析 820 に応じて振幅が適応的にスケールリング 821 a、821 b される。次いで、2チャンネル QMF 領域畳み込み結果 814 と、2チャンネル QMF 領域残響 816 a、816 b とは、合成 828 され、最終的に、2つの QMF 合成フィルタバンクが、ISO/IEC 14496-3:2009、4.6.18.4.2 節に概説されているようにバイノーラル時間領域出力信号 830 を計算する。レンダラはまた、QMF 領域出力データを生成することができる。その場合、合成フィルタバンクは省略される。

#### 【0069】

定義

10

20

30

40

50

バイノーラル・レンダラ・モジュール 800 中に供給されるオーディオ信号 802 は、以下では入力信号と呼ばれる。バイノーラル処理の結果であるオーディオ信号 830 は出力信号と呼ばれる。バイノーラル・レンダラ・モジュール 800 の入力信号 802 は、コーデコーダのオーディオ出力信号である（例えば図 2 の信号 228 を参照）。以下の変数定義が使用される。

【表 1】

$N_{in}$	入力チャネルの数	
$N_{out}$	出力チャネル数、 $N_{out} = 2$	
$M_{DMX}$	実数値の非負ダウンミックス係数（ダウンミックス利得）を含んでいるダウンミックス行列。 $M_{DMX}$ は次元 $N_{out} \times N_{in}$ である	10
$L$	時間領域オーディオサンプルにおいて測定されるフレーム長。	
$v$	時間領域サンプルインデックス	
$n$	QMF タイム・スロット・インデックス（サブバンド・サンプル・インデックス）	
$L_k$	QMF タイムスロットにおいて測定されるフレーム長	
$F$	フレームインデックス（フレーム番号）	
$K$	QMF 周波数帯域の数、 $K = 64$	
$k$	QMF 帯域インデックス（1 . . 64）	
$A, B, ch$	チャネルインデックス（チャネル構成のチャネル番号）	
$L_{BRIR}$	時間領域サンプル中の BRIR の反射初期部分の長さ	
$L_{BRIR}$	QMF タイムスロット中の BRIR の初期反射部分の長さ	
$N_{BRIR}$	BRIR データセット中の BRIR ペアの数	20
$L_{FFT}$	FFT 変換の長さ	
$\Re(\cdot)$	複素数値信号の実数部	
$\Im(\cdot)$	複素数値信号の虚数部	
$\mathbf{m}_{conv}$	どの入力信号チャネルが BRIR データセット中のどの BRIR ペアに属するかをシグナリングするベクトル	
$f_{max}$	バイノーラル処理のために使用される最大周波数	
$f_{max, decoder}$	コーデコーダのオーディオ出力信号中に存在する最大信号周波数	
$K_{max}$	BRIR の初期反射部分を用いたオーディオ入力信号の畳み込みのために使用される最大帯域	
$a$	ダウンミックス行列係数	
$c_{eq,k}$	帯域的なエネルギー等化ファクタ	
$\epsilon$	数値定数、 $\epsilon = 10^{-20}$	30
$d$	QMF 領域タイムスロットにおける遅延	
$\tilde{F}_{ck}^{n,k}$	周波数帯域 $k$ における疑似 FFT 領域信号表現	
$n'$	疑似 FFT 周波数インデックス	
$\tilde{\mathbf{h}}^{n,k}$	周波数帯域 $k$ における BRIR の疑似 FFT 領域表現	
$\tilde{z}_{1,conv}^{n,k}$	周波数帯域 $k$ における疑似 FFT 領域畳み込み結果	
$\tilde{z}_{2,conv}^{n,k}$	中間信号：QMF 領域における 2 チャネル畳み込み結果	
$\tilde{z}_{res}^{n,k}$	中間信号：QMF 領域における 2 チャネル残響	
$K_{ana}$	（残響器のために使用される）分析周波数帯域の数	
$f_{c,ana}$	分析周波数帯域の中心周波数	
$N_{DMX,act}$	ステレオダウンミックスの 1 つのチャネルにダウンミックスされ、実際の信号フレーム中でアクティブであるチャネルの数	40
$c_{corr}$	1 つの信号フレームについての全体的な相関係数	
$c_{corr}^{A,B}$	チャネル A、B の合成のための相関係数	
$\sigma_{\tilde{F}_{ck}^{n,k}}$	信号 $\tilde{F}_{ck}^{n,k}$ のタイムスロット $n$ の標準偏差	
$\mathbf{c}_{scale}$	2 スケーリングファクタのベクトル	
$\tilde{\mathbf{c}}_{scale}$	時間とともに平滑化される、2 スケーリングファクタのベクトル	

【0070】

処理

次に、入力信号の処理について説明する。バイノーラル・レンダラ・モジュールは、入力オーディオ信号の長さ = 2048 時間領域サンプルの連続的な重複しないフレームに作

10

20

30

40

50

用し、長さの処理された入力フレームごとに個のサンプルの1つのフレームを出力する。

【0071】

(1) 初期化および前処理

バイノーラル処理ブロックの初期化は、コアデコーダ(例えば図2の200のデコーダを参照)によって供給されるオーディオサンプルの処理が起こる前に行われる。初期化はいくつかの処理ステップからなる。

【0072】

(a) 分析値の読取り

残響器モジュール816a、816bは、入力パラメータとして残響時間808とエネルギー値810との周波数依存セットを取る。これらの値は、バイノーラル処理モジュール800の初期化においてインターフェースから読み取られる。さらに、時間領域サンプルにおける初期反射から後期残響への遷移時間832が読み取られる。値は、サンプルごとに32ビット、フロート値、リトルエンディアン順序付けで書き込まれるバイナリファイルに記憶され得る。処理のために必要とされる読取り値を以下の表に述べる。

10

【表2】

値の説明	数	データタイプ
遷移の長さ $L_{trans}$	1	整数
周波数帯域の数 $K_{ana}$	1	整数
周波数帯域の中心周波数 $f_{c, ana}$	$K_{ana}$	フロート
秒での残響時間 $RT60$	$K_{ana}$	フロート
1つのBRIRの後期残響部分のエネルギー(振幅の2乗)を表すエネルギー値	$K_{ana}$	フロート

20

【0073】

(b) BRIRの読取りおよび前処理

バイノーラル室内インパルス応答804は、左耳BRIRと右耳BRIRとを個々に記憶する2つの専用ファイルから読み取られる。BRIRの時間領域サンプルは、サンプルごとに24ビットの解像度と32個のチャンネルとを用いて整数waveファイルに記憶される。ファイル中のBRIRの順序付けは以下の表に述べられているとおりである。

30

40

50

【表 3】

チャンネル番号	スピーカラベル
1	CH M L045
2	CH M R045
3	CH M 000
4	CH LFE1
5	CH M L135
6	CH M R135
7	CH M L030
8	CH M R030
9	CH M 180
10	CH LFE2
11	CH M L090
12	CH M R090
13	CH U L045
14	CH U R045
15	CH U 000
16	CH T 000
17	CH U L135
18	CH U R135
19	CH U L090
20	CH U R090
21	CH U 180
22	CH L 000
23	CH L L045
24	CH L R045
25	CH M L060
26	CH M R060
27	CH M L110
28	CH M R110
29	CH U L030
30	CH U R030
31	CH U L110
32	CH U R110

10

20

30

## 【0074】

ラウドスピーカ位置の1つにおいて測定されるBRIRがない場合、waveファイル中の対応するチャンネルは0値を含んでいる。LFEチャンネルはバイノーラル処理のために使用されない。

## 【0075】

前処理ステップとして、バイノーラル室内インパルス応答(BRIR)の所与のセットが時間領域フィルタから複素数値QMF領域フィルタに変換される。複素数値QMF領域における所与の時間領域フィルタの実装は、ISO/IEC FDIS 23003-1: 2006、アネックスBに従って行われる。フィルタ変換のためのプロトタイプフィルタ係数は、ISO/IEC FDIS 23003-1: 2006、アネックスB、表B.1に従って使用される。1 v L t r a n s , nである複素数値QMF領域フィルタ

40

$$h_{ch}^v = \begin{bmatrix} h_{1L}^v & h_{N_{BRIR}}^v \end{bmatrix}$$

を獲得するために、1 v L t r a n s である時間領域表現

50

$$\hat{\mathbf{h}}_{\text{ch}}^{n,k} = \begin{bmatrix} \hat{h}_1^{n,k} & \text{L} & \hat{h}_{N_{\text{BRIR}}}^{n,k} \end{bmatrix}$$

が処理される。

(2) オーディオ信号処理

バイノーラル・レンダラ・モジュール 800 のオーディオ処理ブロックは、コアダコーダから  $N_{\text{in}}$  個の入力チャンネルのための時間領域オーディオサンプル 802 を取得し、 $N_{\text{out}} = 2$  個のチャンネルからなるバイノーラル出力信号 830 を生成する。

処理は、入力として以下を取る。

- ・コアダコーダからの復号されたオーディオデータ 802、
- ・BRIRセット 804 の初期反射部分の複素 QMF 領域表現、および
- ・後期残響 826 a、826 b を生成するために QMF 領域残響器 816 a、816 b によって使用される周波数依存パラメータセット 808、810、832。

【0076】

(a) オーディオ信号の QMF 分析

第1の処理ステップとして、バイノーラル・レンダラ・モジュールは、(コアダコーダから来る)  $N_{\text{in}}$  チャンネル時間領域入力信号の  $= 2048$  時間領域サンプル

$$\begin{bmatrix} y_{\text{ch},1}^v & \text{L} & y_{\text{ch},N_{\text{in}}}^v \end{bmatrix} = \mathbf{y}_{\text{ch}}^v$$

を、次元  $L_n = 32$  QMF タイムスロット (スロットインデックス  $n$ ) と  $K = 64$  周波数帯域 (帯域インデックス  $k$ ) との  $N_{\text{in}}$  チャンネル QMF 領域信号表現 802 に変換する。

ISO/IEC 14496-3:2009、4.B.18.2 節に概説されその修正が ISO/IEC 14496-3:2009、8.6.4.2 節に述べられている QMF 分析が時間領域信号

$$\mathbf{y}_{\text{ch}}^v$$

のフレーム上で実施されて、 $1 \leq v \leq L$  および  $1 \leq n \leq L_n$  である QMF 領域信号

$$\begin{bmatrix} \hat{y}_{\text{ch},1}^{n,k} & \text{L} & \hat{y}_{\text{ch},N_{\text{in}}}^{n,k} \end{bmatrix} = \hat{\mathbf{y}}_{\text{ch}}^{n,k}$$

のフレームが獲得される。

【0077】

(b) QMF 領域オーディオ信号と QMF 領域 BRIR との高速畳み込み

次に、QMF 領域オーディオ信号 802 と QMF 領域 BRIR 804 とを処理するために帯域的な高速畳み込み 812 が行われる。FFT 分析が、入力信号 802 と各 BRIR 804 との各チャンネルのための各 QMF 周波数帯域について行われ得る。

QMF 領域中の複素数値により、1つの FFT 分析は QMF 領域信号表現の実数部で行われ、1つの FFT 分析は QMF 領域信号表現の虚数部で行われる。結果が次いで合成されて、最終的な帯域的な複素数値擬似 FFT 領域信号が次のように形成され、

$$\mathbf{y}_{\text{ch}}^{n',k} = \text{FFT}(\hat{\mathbf{y}}_{\text{ch}}^{n',k}) = \text{FFT}(\Re(\hat{\mathbf{y}}_{\text{ch}}^{n',k})) + j \cdot \text{FFT}(\Im(\hat{\mathbf{y}}_{\text{ch}}^{n',k}))$$

および以下のように帯域的な複素数値 BRIR が形成され、

左耳では

10

20

30

40

50

$$\left\{ \begin{array}{l} \mathbf{h}_1^{n',k} = \text{FFT}(\hat{\mathbf{h}}_1^{n',k}) = \text{FFT}(\Re(\hat{\mathbf{h}}_1^{n',k})) + j \cdot \text{FFT}(\Im(\hat{\mathbf{h}}_1^{n',k})) \\ \mathbf{h}_2^{n',k} = \text{FFT}(\hat{\mathbf{h}}_2^{n',k}) = \text{FFT}(\Re(\hat{\mathbf{h}}_2^{n',k})) + j \cdot \text{FFT}(\Im(\hat{\mathbf{h}}_2^{n',k})) \end{array} \right.$$

右耳では

$$\left\{ \begin{array}{l} \mathbf{h}_1^{n',k} = \text{FFT}(\hat{\mathbf{h}}_1^{n',k}) = \text{FFT}(\Re(\hat{\mathbf{h}}_1^{n',k})) + j \cdot \text{FFT}(\Im(\hat{\mathbf{h}}_1^{n',k})) \\ \mathbf{h}_2^{n',k} = \text{FFT}(\hat{\mathbf{h}}_2^{n',k}) = \text{FFT}(\Re(\hat{\mathbf{h}}_2^{n',k})) + j \cdot \text{FFT}(\Im(\hat{\mathbf{h}}_2^{n',k})) \end{array} \right.$$

10

。 F F T 変換の長さは、複素数値 Q M F 領域 B R I R フィルタの長さ  $L_{\text{trans},n}$  と、Q M F 領域タイムスロット  $L_n$  におけるフレーム長とに従って決定され、したがって、  
 $L_{\text{FFT}} = L_{\text{trans},n} + L_n - 1$ 。

複素数値擬似 F F T 領域信号は、次いで複素数値擬似 F F T 領域 B R I R フィルタが掛けられて、高速畳み込み結果が形成される。入力信号のどのチャンネルが B R I R データセット中のどれ B R I R ペアに対応するかをシグナリングするためにベクトル  $m_{\text{conv}}$  が使用される。

この掛算は、 $1 \leq k \leq K_{\text{max}}$  であるすべての Q M F 周波数帯域  $k$  について帯域的に行われる。最大帯域  $K_{\text{max}}$  は、 $18 \text{ kHz}$  またはコアデコーダからのオーディオ信号中に存在する最大信号周波数のいずれかの周波数を表す Q M F 帯域によって決定される

20

$$f_{\text{max}} = \min(f_{\text{max,decoder}}, 18 \text{ kHz})$$

各 B R I R ペアを用いた各オーディオ入力チャンネルからの掛算の結果は、 $1 \leq k \leq K_{\text{max}}$  である各 Q M F 周波数帯域において合計され、それにより、中間 2 チャンネル  $K_{\text{max}}$  帯域擬似 F F T 領域信号が生じる。

$$\left\{ \begin{array}{l} \mathbf{z}_{\text{ch},1,\text{conv}}^{n',k} = \sum_{ch=1}^{ch=N_{\text{in}}} \mathbf{y}_{\text{ch},ch}^{n',k} \cdot \mathbf{h}_{1,m_{\text{conv}}}^{n',k}[ch] \end{array} \right.$$

30

および

$$\left\{ \begin{array}{l} \mathbf{z}_{\text{ch},2,\text{conv}}^{n',k} = \sum_{ch=1}^{ch=N_{\text{in}}} \mathbf{y}_{\text{ch},ch}^{n',k} \cdot \mathbf{h}_{2,m_{\text{conv}}}^{n',k}[ch] \end{array} \right.$$

は、Q M F 領域周波数帯域  $k$  における擬似 F F T 畳み込み結果

$$\left\{ \begin{array}{l} \mathbf{z}_{\text{ch},\text{conv}}^{n',k} = \left[ \mathbf{z}_{\text{ch},1,\text{conv}}^{n',k}, \mathbf{z}_{\text{ch},2,\text{conv}}^{n',k} \right] \end{array} \right.$$

40

である。

次に、帯域的な F F T 合成が行われて、畳み込み結果が Q M F 領域に逆変換され、それにより、 $1 \leq n \leq L_{\text{FFT}}$  および  $1 \leq k \leq K_{\text{max}}$  である  $L_{\text{FFT}}$  タイムスロット

$$\left\{ \begin{array}{l} \hat{\mathbf{z}}_{\text{ch},\text{conv}}^{n,k} = \left[ \hat{\mathbf{z}}_{\text{ch},1,\text{conv}}^{n,k}, \hat{\mathbf{z}}_{\text{ch},2,\text{conv}}^{n,k} \right] \end{array} \right.$$

50

をもつ中間2チャンネル  $K_{max}$  帯域 QMF 領域信号が生じる。

$L = 32$  個のタイムスロットをもつ各 QMF 領域入力信号フレームについて、 $L = 32$  個のタイムスロットをもつ畳み込み結果信号フレームが返される。残りの  $L_{FFT} - 32$  個のタイムスロットは記憶され、後続のフレームにおいてオーバーラップ加算処理が行われる。

【0078】

(c) 後期残響の生成

第2の中間信号 826 a、826 b として、

$$\hat{\mathbf{z}}_{ch,rev}^{n,k} = \begin{bmatrix} \hat{z}_{ch,1,rev}^{n,k} & \hat{z}_{ch,2,rev}^{n,k} \end{bmatrix} \quad 10$$

と呼ばれる残響信号が周波数領域残響器モジュール 816 a、816 b によって生成される。周波数領域残響器 816 a、816 b は、入力として以下を取る。

- ・入力信号の1つのフレームの QMF 領域ステレオダウンミックス 822、
- ・周波数依存残響時間 808 とエネルギー値 810 とを含んでいるパラメータセット。

周波数領域残響器 816 a、816 b は2チャンネル QMF 領域後期残響テールを返す。

周波数依存パラメータセットの最大使用帯域数が、最大周波数に応じて計算される。

【0079】

最初に、入力信号

$$\hat{\mathbf{y}}_{ch}^{n,k} \quad 20$$

の1つのフレームの QMF 領域ステレオダウンミックス 818 が行われて、入力信号チャンネルの重み付けされた合計によって残響器の入力が形成される。重み付け利得はダウンミックス行列  $M_{DMX}$  中に含まれている。重み付け利得は実数値で非負であり、ダウンミックス行列は次元  $N_{out} \times N_{in}$  である。それは、入力信号のチャンネルが2つの出力チャンネルのうちの1つにマッピングされる、非0値を含んでいる。

【0080】

左半球上のラウドスピーカ位置を表すチャンネルは左出力チャンネルにマッピングされ、右半球上にあるラウドスピーカを表すチャンネルは右出力チャンネルにマッピングされる。これらのチャンネルの信号は1の係数によって重み付けされる。正中面におけるラウドスピーカを表すチャンネルは、バイノーラル信号の両方の出力チャンネルにマッピングされる。これらのチャンネルの入力信号は次の係数によって重み付けされる。

$$a = 0.7071 \approx \frac{1}{\sqrt{2}} \quad 40$$

【0081】

さらに、ダウンミックスにおいてエネルギー等化ステップが実施される。エネルギー等化ステップは、1つのダウンミックスチャンネルの帯域的エネルギーが、このダウンミックスチャンネル中に含まれている入力信号チャンネルの帯域的エネルギーの合計に等しくなるように適応させる。このエネルギー等化は、次の実数値係数を用いた帯域的掛算によって行われる。

$$c_{eq,k} = \sqrt{\frac{P_{in}^k}{P_{out}^k + \varepsilon}}$$

【 0 0 8 2 】

ファクタ  $c_{eq,k}$  は [ 0 . 5 , 2 ] の区間に制限される。ゼロ除算を回避するために数値定数  $\varepsilon$  が導入される。ダウンミックスはまた、周波数  $f_{max}$  に帯域制限される。すべてのより高い周波数帯域における値は 0 に設定される。

10

【 0 0 8 3 】

図 9 は、本発明の一実施形態によるバイノーラルレンダラ 8 0 0 の周波数領域残響器 8 1 6 a、8 1 6 b における処理を概略的に表す。

【 0 0 8 4 】

周波数領域残響器において、入力ミキサ 9 0 0 を使用してステレオ入力のモノダウンミックスが計算される。これは、第 2 の入力チャンネル上で 9 0 ° 位相シフトを適用して非コヒーレントに行われる。

【 0 0 8 5 】

このモノ信号は、次いで、各周波数帯域  $k$  においてフィードバック遅延ループ 9 0 2 に供給され、それにより、インパルスの減衰シーケンスが作成される。その後、インパルス間の間隔中に減衰様式で信号エネルギーを分配し、出力チャンネル間に非コヒーレンスを作成する、並列 FIR 無相関化器が続く。エネルギー減衰を作り出すために減衰フィルタタップ密度が適用される。フィルタタップ位相演算は、スパースで乗算器なしの無相関化器を実装するために 4 つのオプションに限定される。

20

【 0 0 8 6 】

残響の計算の後に、チャンネル間コヒーレンス ( ICC ) 補正 9 0 4 が QMF 周波数帯域ごとに残響器モジュール中に含まれる。ICC 補正ステップでは、ICC を適応させるために周波数依存直接利得  $g_{direct}$  とクロスミックス利得  $g_{cross}$  とが使用される。

【 0 0 8 7 】

異なる周波数帯域のためのエネルギーの量と残響時間とは入力パラメータセット中に含まれている。値は、 $K = 64$  個の QMF 周波数帯域に内部でマッピングされたいくつかの周波数ポイントにおいて与えられる。

30

【 0 0 8 8 】

最終的な中間信号

$$\hat{z}_{ch,rev}^{n,k} = \left[ \hat{z}_{ch,1,rev}^{n,k}, \hat{z}_{ch,2,rev}^{n,k} \right]$$

を計算するために周波数領域残響器の 2 つのインスタンスが使用される。

$$\hat{z}_{ch,1,rev}^{n,k}$$

40

信号は残響器の第 1 のインスタンスの第 1 の出力チャンネルであり

$$\hat{z}_{ch,2,rev}^{n,k}$$

は残響器の第 2 のインスタンスの第 2 の出力チャンネルである。それらは、2 つのチャンネルと、64 個の帯域と、32 個のタイムスロットとの次元を有する最終的な残響信号フレー

50

ムに合成される。

残響器出力の正しいスケールリングを保証するために、ステレオダウンミックス 8 2 2 は、入力信号フレームの相関測度 8 2 0 に従って両方の時間においてスケールリング 8 2 1 a、bされる。スケールリングファクタは、0と1との間の相関係数  $c_{corr}$  に応じて線形的に

$$\left[ \sqrt{N_{DMX,act}}, N_{DMX,act} \right]$$

10

の区間中の値として定義され、ただし、

$$c_{corr} = \frac{1}{N_{in}^2} \cdot \sum_{A=1}^{A=N_{DMX,act}} \sum_{B=1}^{B=N_{DMX,act}} c_{corr}^{A,B}$$

および

$$c_{corr}^{A,B} = \left| \frac{1}{K-1} \cdot \frac{\sum_k \sum_n \hat{y}_{ch,A}^{n,k} \cdot \hat{y}_{ch,B}^{n,k*}}{\sum_n \sigma_{\hat{y}_{ch,A}^n} \cdot \sigma_{\hat{y}_{ch,B}^n}} \right|$$

20

【 0 0 8 9 】

ただし、

$$\sigma_{\hat{y}_{ch,A}^n}$$

30

は、チャンネル A

の1つのタイムスロット n にわたる標準偏差を意味し、演算子 { \* } は複素共役を示し、

$$\hat{y}$$

は、実際の信号フレーム中の Q M F 領域信号

$$\hat{y}$$

40

のゼロ平均バージョンである。

【 0 0 9 0 】

$c_{corr}$  は 2 回計算される。1 回は、実際の信号フレーム F においてアクティブでありステレオダウンミックスの左チャンネル中に含まれる複数のチャンネル A、B についてであり、1 回は、実際の信号フレーム F においてアクティブである、そして、ステレオダウンミックスの右チャンネル中に含まれる複数のチャンネル A、B についてである。 $N_{DMX,act}$  は、1 つのダウンミックスチャンネル A ( 0 に等しくないダウンミックス行列  $M_{DMX}$  の A 番

50

目の行中の行列要素の数)にダウンミックスされ、現在フレーム中でアクティブである、入力チャンネルの数である。

【 0 0 9 1 】

スケーリングファクタは、次いで、以下のようになる。

$$\begin{aligned} \mathbf{c}_{\text{scale}} &= [c_{\text{scale},1}, c_{\text{scale},2}] \\ &= \left[ \sqrt{N_{\text{DMX,act},1}} + c_{\text{corr}} \cdot (N_{\text{DMX,act},1} - \sqrt{N_{\text{DMX,act},1}}), \sqrt{N_{\text{DMX,act},2}} + c_{\text{corr}} \cdot (N_{\text{DMX,act},2} - \sqrt{N_{\text{DMX,act},2}}) \right] \end{aligned}$$

スケーリングファクタは、1次低域フィルタによってオーディオ信号フレームにわたって平滑化され、それにより、平滑化されたスケーリングファクタ

10

$$\mathcal{G}_{\text{scale}} = \left[ \mathcal{G}_{\text{scale},1}, \mathcal{G}_{\text{scale},2} \right]$$

が生じる。

【 0 0 9 2 】

スケーリングファクタは、同じ手段を用いた時間領域相関分析によって、第1のオーディオ入力データフレームにおいて初期化される。

【 0 0 9 3 】

第1の残響器インスタンスの入力はスケーリングファクタ

20

$$\mathcal{G}_{\text{scale},1}$$

でスケーリングされ、第2の残響器インスタンスの入力はスケーリングファクタ

$$\mathcal{G}_{\text{scale},2}$$

30

でスケーリングされる。

【 0 0 9 4 】

(d) 畳み込み結果と後期残響の合成

次に、1つQMF領域オーディオ入力フレームについての、畳み込み結果 8 1 4、

$$\hat{\mathbf{z}}_{\text{ch,conv}}^{n,k} = \left[ \hat{z}_{\text{ch},1,\text{conv}}^{n,k}, \hat{z}_{\text{ch},2,\text{conv}}^{n,k} \right]$$

、および残響器出力 8 2 6 a、8 2 6 b、

$$\hat{\mathbf{z}}_{\text{ch,rev}}^{n,k} = \left[ \hat{z}_{\text{ch},1,\text{rev}}^{n,k}, \hat{z}_{\text{ch},2,\text{rev}}^{n,k} \right]$$

40

が、2つの信号を帯域的に合計するミキシングプロセス 8 2 8 によって合成される。畳み込みは  $K_{\text{max}}$  までの帯域においてしか行われないので、 $K_{\text{max}}$  よりも高い上側帯域は、

$$\hat{\mathbf{z}}_{\text{ch,conv}}^{n,k}$$

が0であることに留意されたい。

50

後期残響出力は、ミキシングプロセスにおいて  
 $d = ((L_{trans} - 20 \cdot 64 + 1) / 64 + 0.5) + 1$  のタイムスロットの量だけ  
 遅延される。

遅延  $d$  は、妥当なタイムスロットにおける後期残響の挿入を保証するために、B R I R  
 における初期反射から後期反射への遷移時間、および  $20$  Q M F タイムスロットの残響器  
 の初期遅延、ならびに B R I R の Q M F 分析のための  $0.5$  Q M F タイムスロットの分析  
 遅延を考慮に入れる。1つのタイムスロット  $n$   
 における合成信号

$$\hat{\mathbf{z}}_{ch}^{n,k} \quad 10$$

は、

$$\hat{\mathbf{z}}_{ch,conv}^{n,k} + \hat{\mathbf{z}}_{ch,rev}^{n-d,k}$$

によって計算される。

【0095】

(e) バイノーラル Q M F 領域信号の Q M F 合成  
 Q M F 領域出力信号

20

$$\hat{\mathbf{z}}_{ch}^{n,k}$$

の  $32$  タイムスロットの1つの2チャンネルフレームは、I S O / I E C 14496-3  
 : 2009、4.6.18.4.2.節による Q M F 合成によって長さ  $L$  で2チャンネル時  
 間領域信号フレームに変換され、それにより、最終的な時間領域出力信号  $830$ 、

$$\mathbf{z}_{ch}^{o/v} = \begin{bmatrix} \mathbf{z}_{ch,1}^{o/v} L & \mathbf{z}_{ch,2}^{o/v} \end{bmatrix} \quad 30$$

が生じる。

【0096】

本発明の手法によれば、入力信号の特性を考慮に入れて合成または人工的後期残響がス  
 ケーリングされ、それにより、別々の処理によって得られる計算複雑さの低減を利用しな  
 がら出力信号の品質が改善される。また、上記の説明から分かるように、追加の聴覚モデ  
 ルまたはターゲット残響ラウドネスは不要である。

【0097】

40

本発明は、上記で説明した実施形態に限定されないことに留意されたい。例えば、上記  
 の実施形態では Q M F 領域との関連において説明したが、他の時間 - 周波数領域、例えば  
 S T F T 領域も使用され得ることに留意されたい。また、スケーリングファクタは、相関  
 が周波数帯域の全数、すなわち  $i \in [1, N]$  にわたって計算されないように周波数依存  
 様式で計算され得るが、以下のように定義されたいいくつかの  $S$  個のサブセットにおいて計  
 算される。

$$i_1 \in [1, N_1], i_2 \in [N_1 + 1, N_2], \dots, i_s \in [N_{s-1} + 1, N]$$

【0098】

また、周波数帯域にわたって平滑化が適用され得るか、または、特定のルールに応じて  
 、例えば聴覚の周波数解像度に従って帯域が合成され得る。平滑化は、例えばフレームサ

50

イズまたは聴取者の選好に応じて、様々な時定数に適合され得る。

【0099】

本発明の手法はまた、様々なフレームサイズについて適用され得、時間周波数領域中のただ1つのタイムスロットのフレームサイズさえ可能である。

【0100】

実施形態によれば、ダウンミックスのために様々なダウンミックス行列が、例えば対称ダウンミックス行列または非対称行列が使用され得る。

【0101】

相関測度は、オーディオビットストリーム中で送信されるパラメータから、例えばMPEGサラウンドまたはSACにおけるチャンネル間コヒーレンスから導出され得る。また、実施形態によれば、必要な場合、平均値計算から行列のいくつかの値、例えば、誤って計算された値または主対角線上の値、自己相関値を除外することが可能である。

【0102】

プロセスは、デコーダ側におけるバイノーラルレンダラにおいてそれを使用する代わりに、例えば低複雑度バイノーラルプロファイルを適用するとき、エンコーダにおいて行われ得る。これは、スケールリングファクタのその何らかの表現、例えばスケールリングファクタ自体、0と1との間の相関測度などを生じ、これらのパラメータは、固定ダウンストリーム行列についてエンコーダからデコーダにビットストリーム中で送信される。

【0103】

また、上記で説明した実施形態では、残響器514に続いて利得を適用することについて説明したが、他の実施形態によれば、利得はまた、残響器514の前に、または、例えば残響器514内の利得を修正することによって、残響器内で適用され得ることに留意されたい。より少ない計算が必要とされ得るので、これは有利である。

【0104】

いくつかの態様について装置のコンテキストで説明したが、これらの態様は対応する方法の説明をも表すことが明らかであり、ブロックまたはデバイスは、方法ステップ、または方法ステップの特徴に対応する。同様に、方法ステップのコンテキストで説明する態様も、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説明を表す。方法ステップの一部または全部は、例えば、マイクロプロセッサ、プログラマブル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（またはそれを使用して）実行され得る。いくつかの実施形態では、最も重要な方法ステップのうちの何らかの1つまたは複数はそのような装置によって実行され得る。

【0105】

いくつかの実装要件に応じて、本発明の実施形態はハードウェアまたはソフトウェアにおいて実装され得る。実装は、それぞれの方法が実施されるようにプログラマブル・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可能である）、電子的に読取り可能な制御信号をその上に記憶した、デジタル記憶媒体などの非一時的記憶媒体、例えばフロッピー（登録商標）ディスク、DVD、Blu-Ray（登録商標）、CD、ROM、PROM、およびEPROM、EEPROMまたはFLASH（登録商標）メモリを使用して実施され得る。したがって、デジタル記憶媒体はコンピュータ可読であり得る。

【0106】

本発明によるいくつかの実施形態は、本明細書で説明する方法のうちの1つが実施されるようにプログラマブル・コンピュータ・システムと協働することが可能である、電子的に読取り可能な制御信号を有するデータキャリアを備える。

【0107】

概して、本発明の実施形態は、プログラムコードをもつコンピュータプログラム製品として実装され得、プログラムコードは、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、方法のうちの1つを実施するように動作可能である。プログラムコードは、例えば、機械可読キャリア上に記憶され得る。

【0108】

10

20

30

40

50

他の実施形態は、機械可読キャリアに記憶された、本明細書で説明する方法のうちの1つを実施するためのコンピュータプログラムを備える。

【0109】

言い換えれば、本発明の方法の実施形態は、したがって、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、本明細書で説明する方法のうちの1つを実施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。

【0110】

本発明の方法のさらなる実施形態は、したがって、本明細書で説明する方法のうちの1つを実施するためのコンピュータプログラムをその上に記録して備える、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可読媒体）である。データキャリア、デジタル記憶媒体または記録媒体は、典型的には有形および/または非一時的である。

10

【0111】

発明方法のさらなる実施形態は、したがって、本明細書で説明する方法のうちの1つを実施するためのコンピュータプログラムを表す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、例えば、データ通信接続を介して、例えば、インターネットを介して転送されるように構成され得る。

【0112】

さらなる実施形態は、本明細書で説明する方法のうちの1つを実施するように構成またはプログラムされる処理手段、例えば、コンピュータまたはプログラマブル論理デバイスを備える。

20

【0113】

さらなる実施形態は、本明細書で説明する方法のうちの1つを実施するためのコンピュータプログラムをその上にインストールしたコンピュータを備える。

【0114】

本発明によるさらなる実施形態は、本明細書で説明する方法のうちの1つを実施するためのコンピュータプログラムを受信機に（例えば、電子的または光学的に）転送するように構成される装置またはシステムを備える。受信機は、例えば、コンピュータ、モバイルデバイス、メモリデバイスなどであり得る。装置またはシステムは、例えば、コンピュータプログラムを受信機に転送するためのファイルサーバーを備え得る。

【0115】

いくつかの実施形態では、本明細書で説明する方法の機能の一部または全部を実施するためにプログラマブル論理デバイス（例えば、フィールド・プログラマブル・ゲート・アレイ）が使用され得る。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、本明細書で説明する方法のうちの1つを実施するためにマイクロプロセッサと協働し得る。概して、方法は、好ましくはどんなハードウェア装置によっても実施される。

30

【0116】

また、以下にさらなる実施形態を説明する。

第1の実施形態は、

室内インパルス応答（300）に従ってオーディオ信号（504、802）を処理するための方法であって、前記方法は、

40

前記室内インパルス応答（300）の初期部分（301、302）と後期残響（304）とを用いて前記オーディオ信号（504、802）を別々に処理すること（502、514、812、816a、816b）であって、前記後期残響（304）を処理することが、スケーリングされた残響信号を生成することを含み、前記スケーリングが前記オーディオ信号（504、802）に依存する、処理することと、

前記室内インパルス応答の前記初期部分を用いて処理された前記オーディオ信号と前記スケーリングされた残響信号とを合成することとを含む、方法を提供する。

第2の実施形態は、

50

前記スケーリングが、前記オーディオ信号（504、802）の1つまたは複数の入力チャンネルの状態に依存する、前記1の実施形態の方法を提供する。

第3の実施形態は、

前記オーディオ信号（504、802）の前記1つまたは複数の入力チャンネルの前記状態が、入力チャンネルの数と、アクティブ入力チャンネルの数と、前記入力チャンネル中のアクティビティとのうちの1つまたは複数を含む、前記2の実施形態の方法を提供する。

第4の実施形態は、

前記スケーリングが、前記オーディオ信号（504、802）のあらかじめ定義されたまたは計算される相関測度に依存する、前記第1から第3の実施形態のうちのいずれか一の方法を提供する。

10

第5の実施形態は、

前記スケーリングされた残響信号を生成することが、利得ファクタを適用することを含み、前記利得ファクタは、前記オーディオ信号（504、802）の前記1つまたは複数の入力チャンネルの前記状態に基づいておよび/または前記オーディオ信号（504、802）についての前記あらかじめ定義されたまたは計算される相関測度に基づいて決定される、前記第1から第4の実施形態のうちのいずれか一の方法を提供する。

第6の実施形態は、

前記スケーリングされた残響信号を生成することが、前記オーディオ信号（504、802）の前記後期残響（304）を処理する前に、その間にまたはその後前記利得ファクタを適用することを含む、前記第5の実施形態の方法を提供する。

20

第7の実施形態は、

前記利得ファクタが以下のように決定され、

$$g = c_u + \cdot (c_c - c_u)$$

ただし

= 前記オーディオ信号（504、802）についてのあらかじめ定義されたまたは計算される相関測度、

$c_u$ 、 $c_c$  = 前記オーディオ信号（504、802）の前記1つまたは複数の入力チャンネルの前記状態を示すファクタであり、 $c_u$ は、全体として無相関なチャンネルを指し、 $c_c$ は、全体として相関するチャンネルに関係する、前記第5または第6の実施形態の方法を提供する。

30

第8の実施形態は、

$c_u$ および $c_c$ が以下のように決定され、

【数18】

$$c_u = 10^{\frac{10 \cdot \log_{10}(K_{in})}{20}} = \sqrt{K_{in}}$$

【数19】

$$c_c = 10^{\frac{20 \cdot \log_{10}(K_{in})}{20}} = K_{in}$$

40

ただし

$K_{in}$  = アクティブまたは固定ダウンミックスチャンネルの数である、前記第7の実施形態の方法を提供する。

第9の実施形態は、

前記利得ファクタが、複数のオーディオフレームにわたって低域フィルタ処理される、前記第5から第8の実施形態うちのいずれか一の方法を提供する。

50

第 10 の実施形態は、

前記利得ファクタが以下のように低域フィルタ処理され、

【数 20】

$$g_s(t_i) = c_{s,old} \cdot g_s(t_i - 1) + c_{s,new} \cdot g$$

【数 21】

$$c_{s,old} = e^{-\left(\frac{1}{f_s \cdot k} \right)}$$

10

【数 22】

$$c_{s,new} = 1 - c_{s,old}$$

ただし

20

$t_s$  = 前記低域フィルタの時定数

$t_i$  = フレーム  $t_i$  におけるオーディオフレーム

$g_s$  = 平滑化された利得ファクタ

$k$  = フレームサイズ、および

$f_s$  = サンプル周波数である、

前記第 9 の実施形態の方法を提供する。

第 11 の実施形態は、

前記スケールされた残響信号を生成することが、前記オーディオ信号 (504、802) の相関分析を含む、前記第 1 から第 10 の実施形態のうちのいずれか一の方法を提供する。

30

第 12 の実施形態は、

前記オーディオ信号 (504、802) の前記相関分析が、前記オーディオ信号 (504、802) のオーディオフレームについて合成相関測度を決定することを含み、前記合成相関測度は、1つのオーディオフレームの複数のチャンネル合成についての相関係数を合成することによって計算され、各オーディオフレームは1つまたは複数のタイムスロットを含む、前記第 11 の実施形態の方法を提供する。

第 13 の実施形態は、

前記相関係数を合成することが、前記オーディオフレームの複数の相関係数を平均化することを含む、前記第 12 の実施形態の方法を提供する。

第 14 の実施形態は、

40

前記合成相関測度を決定することが、

(i) 前記1つのオーディオフレームのチャンネルごとに全体的な平均値を計算することと、

(ii) 前記対応するチャンネルから前記平均値を減算することによってゼロ平均オーディオフレームを計算することと、

(iii) 複数のチャンネル合成について前記相関係数を計算することと、

(iv) 前記合成相関測度を複数の相関係数の前記平均として計算することと

を含む、前記第 11 または第 12 の実施形態の方法を提供する。

第 15 の実施形態は、

チャンネル合成についての前記相関係数が以下のように計算され、

50

【数 2 3】

$$\rho[m, n] = \left| \frac{1}{(N-1)} \cdot \frac{\sum_i \sum_j x_m[i, j] \cdot x_n[i, j]^*}{\sum_j \sigma(x_m[j]) \cdot \sigma(x_n[j])} \right|$$

ただし

- [ m , n ] = 相関係数、
- ( x<sub>m</sub> [ j ] ) = チャンネル m の 1 つのタイムスロット j にわたる標準偏差、
- ( x<sub>n</sub> [ j ] ) = チャンネル n の 1 つのタイムスロット j にわたる標準偏差、
- x<sub>m</sub> , x<sub>n</sub> = ゼロ平均変数、
- i [ 1 , N ] = 周波数帯域、
- j [ 1 , M ] = タイムスロット、
- m , n [ 1 , K ] = チャンネル、
- \* = 複素共役である、

前記第 1 1 から第 1 4 の実施形態のうちのいずれか一の方法を提供する。

第 1 6 の実施形態は、

前記スケールされた残響信号の開始を前記室内インパルス応答 ( 3 0 0 ) における初期反射から後期残響 ( 3 0 4 ) への遷移点に一致させるように前記スケールされた残響信号を遅延させることを含む、前記第 1 から第 1 5 の実施形態のうちのいずれか一の方法を提供する。

第 1 7 の実施形態は、

前記オーディオ信号 ( 5 0 4 、 8 0 2 ) の前記後期残響 ( 3 0 4 ) を処理することが、前記オーディオ信号 ( 5 0 4 、 8 0 2 ) をダウンミックスすることと、前記ダウンミックスされたオーディオ信号を残響器に印加することを含む、前記第 1 から第 1 6 の実施形態のうちのいずれか一の方法を提供する。

第 1 8 の実施形態は、

コンピュータによって実行されているとき、前記第 1 から第 1 7 の実施形態のうちのいずれか一の方法を行うための命令を記憶したコンピュータ可読媒体を含む非有形コンピュータ製品を提供する。

第 1 9 の実施形態は、

オーディオ信号 ( 5 0 4 、 8 0 2 ) を受信するための入力と、室内インパルス応答 ( 3 0 0 ) の初期部分 ( 3 0 1 、 3 0 2 ) に従って前記受信されたオーディオ信号 ( 5 0 4 、 8 0 2 ) を処理するための初期部分プロセッサと、

前記室内インパルス応答 ( 3 0 0 ) の後期残響に従って前記受信されたオーディオ信号 ( 5 0 4 、 8 0 2 ) を処理するための後期残響プロセッサであって、前記後期残響プロセッサが、スケールされた残響信号を生成するように構成され、前記スケールが、前記受信されたオーディオ信号 ( 5 0 4 、 8 0 2 ) に依存する後期残響プロセッサと、

前記受信されたオーディオ信号 ( 5 0 4 、 8 0 2 ) の前記処理された初期部分と前記スケールされた残響信号とを出力オーディオ信号に合成するための出力とを備える、信号処理ユニットを提供する。

第 2 0 の実施形態は、

前記後期残響プロセッサが、前記オーディオ信号 ( 5 0 4 、 8 0 2 ) を受信し残響信号を生成する残響器と、前記残響器の入力または出力に結合され利得ファクタによって制御される利得段とを備える、前記第 1 9 の実施形態の信号処理ユニットを提供する。

第 2 1 の実施形態は、

前記オーディオ信号 ( 5 0 4 、 8 0 2 ) に依存する前記利得ファクタを生成する相関分析器を備える、前記第 2 0 の実施形態の信号処理ユニットを提供する。

第 2 2 の実施形態は、

10

20

30

40

50

前記利得段に結合された低域フィルタと、  
 前記利得段と加算器との間に結合された遅延要素と  
 のうちの少なくとも1つをさらに備え、前記加算器は、前記初期部分プロセッサと前記出力とにさらに結合された、前記第20または第21の実施形態の信号処理ユニットを提供する。

第23の実施形態は、

前記第19から第22のうちのいずれか一の信号処理ユニットを備える、バイノーラルレンダラを提供する。

第24の実施形態は、

コーディングより前にオーディオ信号を処理するために前記第19から第22の実施形態のうちのいずれか一の信号処理ユニットまたは前記第23のバイノーラルレンダラを備える、前記オーディオ信号をコーディングするためのオーディオエンコーダを提供する。

10

第25の実施形態は、

復号されたオーディオ信号を処理するために前記第19から第22の実施形態のうちのいずれか一の信号処理ユニットまたは前記第23のバイノーラルレンダラ

を備える、符号化されたオーディオ信号を復号するためのオーディオデコーダを提供する。

上記で説明した実施形態は本発明の原理を例示するためのものにすぎない。本明細書で説明する構成および詳細の修正形態および変形形態は、他の当業者に明らかであることを理解されたい。したがって、本明細書の実施形態の記述および説明として提示される特定の

20

の詳細によっては限定されず、直下の特許請求の範囲によってのみ限定されることが意図される。

【0117】

文献

[1] M. R. Schroeder, 「Digital Simulation of Sound Transmission in Reverberant Spaces」, The Journal of the Acoustical Society of America, Vol. 47, pp. 424 - 431 (1970)、および、J. A. Moorer, 「About This Reverberation Business」, Computer Music Journal, Vol. 3, no. 2, pp. 13 - 28, MIT Press (1979)において拡張されている。

30

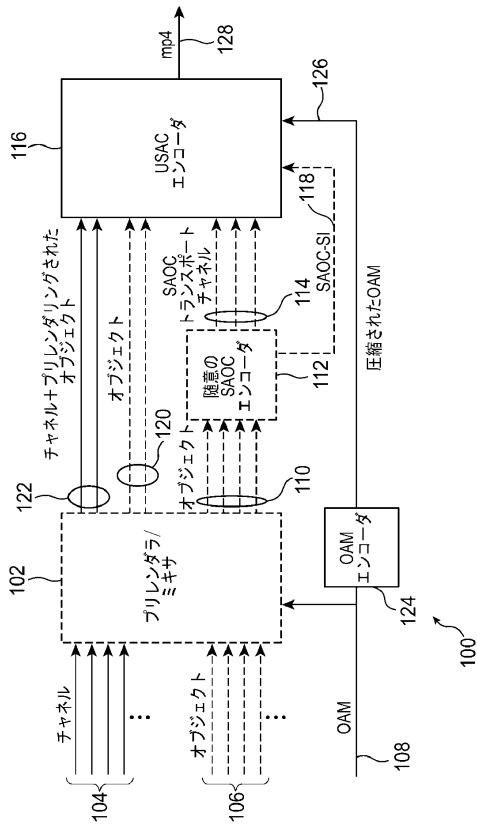
[2] Uhle, Christian, Paulus, Jouni, Herre, Jürgen: 「Predicting the Perceived Level of Late Reverberation Using Computational Models of Loudness」, Proceedings, 17th International Conference on Digital Signal Processing (DSP), 2011年7月6~8日、コルフ、ギリシャ。

[3] Czyzewski, Andrzej: 「A Method of Artificial Reverberation Quality Testing」, J. Audio Eng. Soc., Vol. 38, No. 3, 1990。

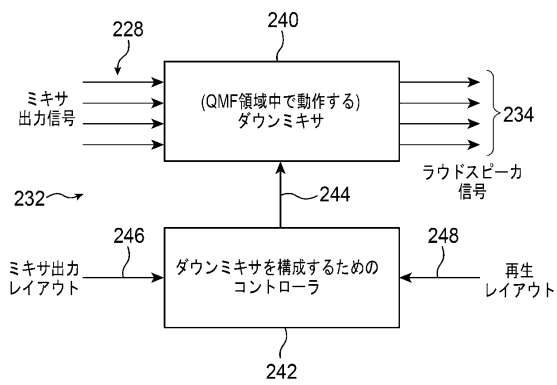
40

【図面】

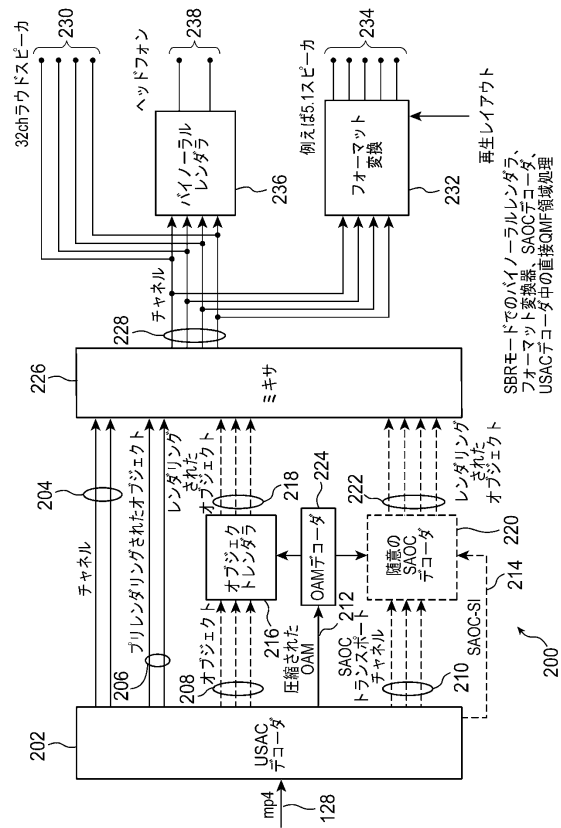
【図 1】



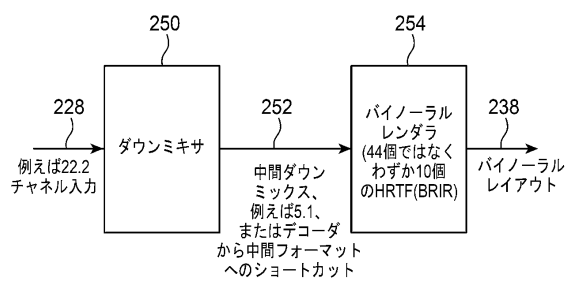
【図 3】



【図 2】



【図 4】



10

20

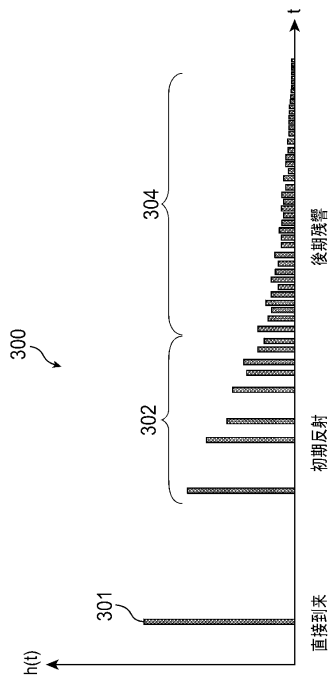
30

40

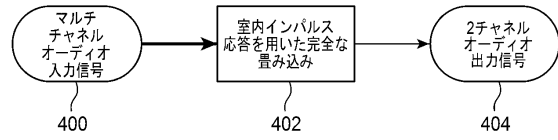
50

再生レイアウト  
SBRモードでのバイノーラルレンダラ、  
フォーマット変換器、SAOCデコーダ、  
USACデコーダ中の直接QMF領域処理

【図5】



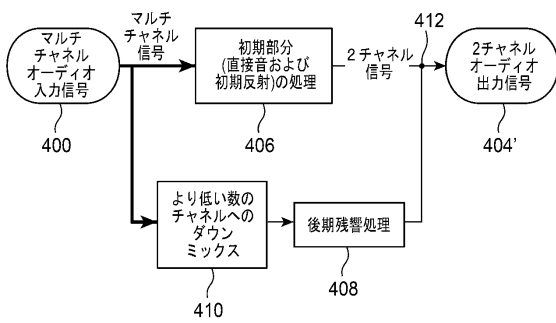
【図6(a)】



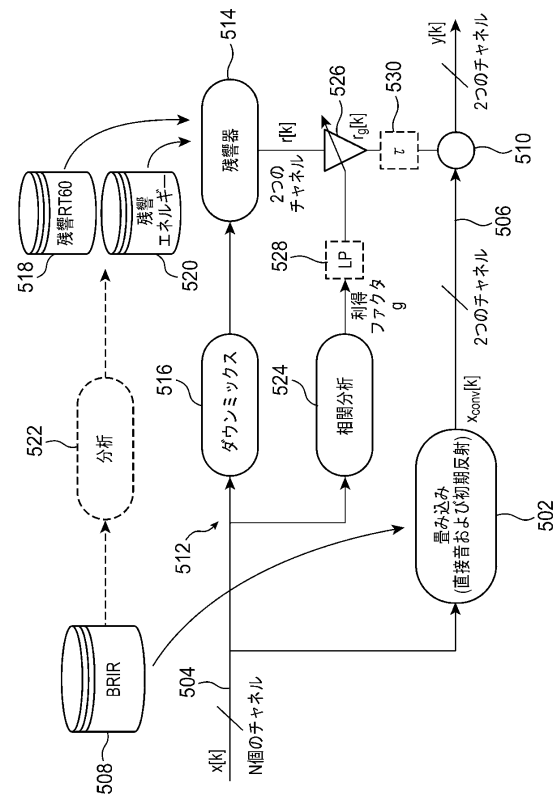
10

20

【図6(b)】



【図7】

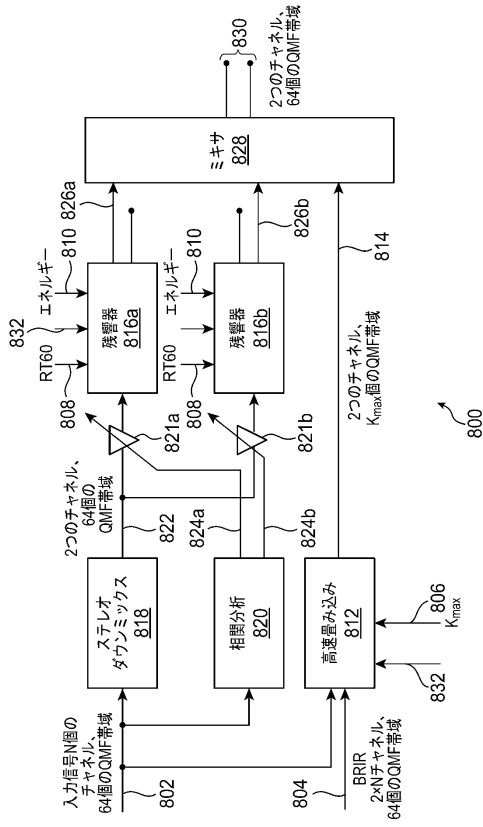


30

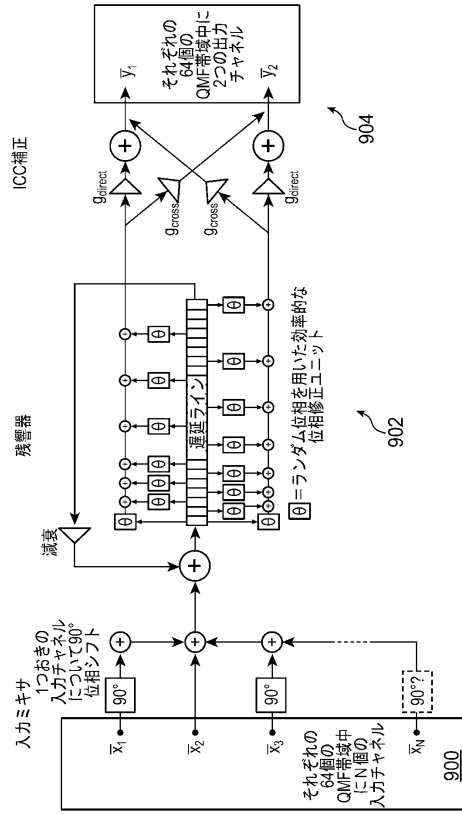
40

50

【図8】



【図9】



10

20

30

40

50

## フロントページの続き

欧州特許庁(EP)

審査官 渡邊 正宏

(56)参考文献 特許第 6 8 7 9 9 7 9 ( J P , B 2 )

特開 2 0 0 4 - 0 9 3 7 2 8 ( J P , A )

(58)調査した分野 (Int.Cl. , D B 名)

G 1 0 K 1 5 / 0 0 - 1 5 / 1 2

H 0 4 S 1 / 0 0 - 7 / 0 0