

(19)日本国特許庁(JP)

## (12)特許公報(B2)

(11)特許番号  
特許第7008862号  
(P7008862)

(45)発行日 令和4年1月25日(2022.1.25)

(24)登録日 令和4年1月13日(2022.1.13)

(51)国際特許分類 F I  
H 0 4 S 7/00 (2006.01) H 0 4 S 7/00 3 0 0

請求項の数 35 (全29頁)

(21)出願番号	特願2021-151299(P2021-151299)	(73)特許権者	518253875
(22)出願日	令和3年9月16日(2021.9.16)		ブームクラウド 3 6 0 インコーポレイ
(62)分割の表示	特願2020-532786(P2020-532786)		テッド
原出願日	平成30年11月29日(2018.11.29)		アメリカ合衆国 9 2 0 2 4 カリフォル
(65)公開番号	特開2021-192553(P2021-192553)		ニア州 エンシニータス サウス コースト
	A)		ハイウェイ 1 0 1 6 8 7 ナンバー 3
(43)公開日	令和3年12月16日(2021.12.16)	(74)代理人	110001243
審査請求日	令和3年9月17日(2021.9.17)		特許業務法人 谷・阿部特許事務所
(31)優先権主張番号	62/599,605	(72)発明者	ザッカリー セルデス
(32)優先日	平成29年12月15日(2017.12.15)		アメリカ合衆国 9 2 0 2 4 カリフォル
(33)優先権主張国・地域又は機関	米国(US)		ニア州 エンシニータス サウス コースト
(31)優先権主張番号	16/151,201		ハイウェイ 1 0 1 6 8 7 ナンバー 3
(32)優先日	平成30年10月3日(2018.10.3)		1 1 ブームクラウド 3 6 0 インコー
	最終頁に続く		ポレイテッド内
			最終頁に続く

(54)【発明の名称】 カンファレンスのためのサブバンド空間処理およびクロストークキャンセルシステム

## (57)【特許請求の範囲】

## 【請求項1】

複数のクライアントデバイスのうちの一つのクライアントデバイスにカンファレンスを提供するための方法であって、  
 前記複数のクライアントデバイスから入力オーディオストリームを受信するステップと、  
 前記複数のクライアントデバイスのうちの前記クライアントデバイスに対して、サウンドフィールド内の前記複数のクライアントデバイスのうちの他のクライアントデバイスの空間位置を定義する配置データを決定するステップと、  
 前記配置データに従って前記他のクライアントデバイスの入力オーディオストリームを混合しパンすることによって、左混合チャンネルおよび右混合チャンネルを含む前記クライアントデバイスのための混合ストリームを生成するステップと、  
 前記クライアントデバイスに関連付けられたスピーカのタイプを決定するステップと、  
 前記混合ストリームの前記左混合チャンネルと前記右混合チャンネルにクロストーク処理を適用することによって左チャンネルと右チャンネルを含むクロストーク処理されたストリームを生成するステップであって、前記クロストーク処理は、  
 前記左混合チャンネルの少なくとも一部をフィルタリングし時間遅延させることによって、左クロストーク処理された成分を生成することと、  
 前記右混合チャンネルの少なくとも一部をフィルタリングし時間遅延させることによって、右クロストーク処理された成分を生成することと、  
 前記左混合チャンネルと前記右クロストーク処理された成分を結合することによって、前記

左チャンネルを生成することと、  
前記右混合チャンネルと前記左クロストーク処理された成分を結合することによって、前記右チャンネルを生成することと

を含み、

スピーカ-の第1のタイプが前記クライアントデバイスに関連付けられると決定すること  
にตอบสนองして、前記クロストーク処理はクロストークシミュレーションであり、

スピーカ-の第2のタイプが前記クライアントデバイスに関連付けられると決定すること  
にตอบสนองして、前記クロストーク処理はクロストークキャンセレーションである、

ステップと、

前記クライアントデバイスの左スピーカ-へ前記クロストーク処理されたストリームの前  
記左チャンネルと、前記クライアントデバイスの右スピーカ-へ前記クロストーク処理され  
たストリームの前記右チャンネルとを提供するステップと

10

を備えることを特徴とする方法。

【請求項2】

前記左クロストーク処理された成分を生成することは、左混合インバンド信号および左混  
合アウトオブバンド信号に前記左混合チャンネルを分割することを含み、

前記左混合チャンネルの前記少なくとも一部は、前記左混合インバンド信号を含み、

前記右クロストーク処理された成分を生成することは、右混合インバンド信号および右混  
合アウトオブバンド信号に前記右混合チャンネルを分割することを含み、

前記右混合チャンネルの前記少なくとも一部は、前記右混合インバンド信号を含む

20

ことを特徴とする請求項1に記載の方法。

【請求項3】

前記混合ストリームに前記クロストーク処理を適用することによって引き起こされるスペ  
クトル欠陥を調整するクロストーク補正を前記混合ストリームに適用するステップをさら  
に備えることを特徴とする請求項1に記載の方法。

【請求項4】

前記配置データを決定するステップは、ネットワークを介して前記クライアントデバイス  
から前記配置データを受信することを含むことを特徴とする請求項1に記載の方法。

【請求項5】

前記クライアントデバイスからの前記入力オーディオストリームの各々は、複数のオーデ  
ィオチャンネルを含むことを特徴とする請求項1に記載の方法。

30

【請求項6】

前記配置データに従って前記他のクライアントデバイスの前記入力オーディオストリーム  
を混合しパンすることによって、前記左混合チャンネルおよび前記右混合チャンネルを含む、  
前記クライアントデバイスのための前記混合ストリームを生成するステップは、

前記サウンドフィールド内の前記他のクライアントデバイスの前記空間位置に従って、前  
記他のクライアントデバイスの前記入力オーディオストリームの各々から左チャンネルおよ  
び右チャンネルを生成することと、

前記左混合チャンネルを生成するために前記他のクライアントデバイスの前記入力オーデ  
ィオストリームから左チャンネルを結合することと、

40

前記右混合チャンネルを生成するために前記他のクライアントデバイスの前記入力オーデ  
ィオストリームから前記右チャンネルを結合することと

を含むことを特徴とする請求項1に記載の方法。

【請求項7】

前記複数のクライアントデバイスに接続されたサーバーは、前記混合ストリームを生成し  
、前記混合ストリームを前記クライアントデバイスに提供し、

前記クライアントデバイスは、前記混合ストリームから前記クロストーク処理されたスト  
リームを生成する

ことを特徴とする請求項1に記載の方法。

【請求項8】

50

前記複数のクライアントデバイスに接続されたサーバーは、前記混合ストリームおよび前記クロストーク処理されたストリームを生成し、前記クロストーク処理されたストリームを前記クライアントデバイスに提供することを特徴とする請求項 1 に記載の方法。

【請求項 9】

前記サーバーは、前記クライアントデバイスからデバイスの説明を受信し、前記デバイスの説明に基づいて、前記クロストーク処理を適用するためのパラメータを決定することを特徴とする請求項 8 に記載の方法。

【請求項 10】

前記クライアントデバイスは、前記混合ストリームおよび前記クロストーク処理されたストリームを生成することを特徴とする請求項 1 に記載の方法。

10

【請求項 11】

前記混合ストリームの前記左混合チャンネルおよび前記右混合チャンネルにサブバンド空間処理を適用することによって、左強調チャンネルおよび右強調チャンネルを含む空間強調ストリームを生成するステップであって、前記サブバンド空間処理は、

前記左混合チャンネルおよび前記右混合チャンネルのミッド成分のミッドサブバンド成分に第 1 のゲインを適用して、強調ミッド成分を生成することと、

前記左混合チャンネルおよび前記右混合チャンネルのサイド成分のサイドサブバンド成分に第 2 のゲインを適用して、強調サイド成分を生成することと、

前記強調ミッド成分および前記強調サイド成分を使用して、前記左強調チャンネルおよび前記右強調チャンネルを生成することと

20

を含む、ステップと、

前記クライアントデバイスの左スピーカーに前記空間強調ストリームの前記左強調チャンネルと、前記クライアントデバイスの右スピーカーに前記空間強調ストリームの前記右強調チャンネルとを提供するステップと

をさらに備えることを特徴とする請求項 1 に記載の方法。

【請求項 12】

プロセッサによって実行されると、

複数のクライアントデバイスから入力オーディオストリームを受信し、

前記複数のクライアントデバイスのうちの一つのクライアントデバイスに対して、サウンドフィールド内の前記複数のクライアントデバイスのうちの他のクライアントデバイスの空間位置を定義する配置データを決定し、

30

前記配置データに従って前記他のクライアントデバイスの入力オーディオストリームを混合しパンすることによって、左混合チャンネルおよび右混合チャンネルを含む前記クライアントデバイスのための混合ストリームを生成し、

前記クライアントデバイスに関連付けられたスピーカーのタイプを決定し、

前記混合ストリームの前記左混合チャンネルと前記右混合チャンネルにクロストーク処理を適用することによって、左チャンネルおよび右チャンネルを含むクロストーク処理されたストリームを生成し、前記クロストーク処理は、

前記左混合チャンネルの少なくとも一部をフィルタリングし時間遅延させることによって、左クロストーク処理された成分を生成することと、

40

前記右混合チャンネルの少なくとも一部をフィルタリングし時間遅延させることによって、右クロストーク処理された成分を生成することと、

前記左混合チャンネルと前記右クロストーク処理された成分を結合することによって、前記左チャンネルを生成することと、

前記右混合チャンネルと前記左クロストーク処理された成分を結合することによって、前記右チャンネルを生成することと

を含み、

スピーカーの第 1 のタイプが前記クライアントデバイスに関連付けられると決定することに対応して、前記クロストーク処理はクロストークシミュレーションであり、

スピーカーの第 2 のタイプが前記クライアントデバイスに関連付けられると決定すること

50

にตอบสนองして、前記クロストーク処理はクロストークキャンセレーションであるように前記プロセッサを構成するプログラムコードを格納することを特徴とする非一時的なコンピュータ可読媒体。

【請求項 13】

前記クロストーク処理されたストリームを生成するように前記プロセッサを構成する前記プログラムコードは、前記混合ストリームの前記左混合チャンネルおよび前記右混合チャンネルにサブバンド空間処理を適用するように前記プロセッサをさらに構成するプログラムコードを含むことを特徴とする請求項 12 に記載のコンピュータ可読媒体。

【請求項 14】

実行されると、前記クライアントデバイスからデバイスの説明を受信し、前記デバイスの説明に基づいて前記クロストーク処理を適用するためのパラメータを決定するように前記プロセッサをさらに構成するプログラムコードをさらに含むことを特徴とする請求項 12 に記載のコンピュータ可読媒体。

10

【請求項 15】

実行されると、前記混合ストリームに前記クロストーク処理を適用することによって引き起こされるスペクトル障害を調整する、前記混合ストリームへのクロストーク補正を適用するように前記プロセッサをさらに構成するプログラムコードをさらに含むことを特徴とする請求項 12 に記載のコンピュータ可読媒体。

【請求項 16】

前記左クロストーク処理された成分を生成するように前記プロセッサを構成する前記プログラムコードは、

20

左混合インバンド信号および左混合アウトオブバンド信号に前記左混合チャンネルを分割し、前記左混合チャンネルの前記少なくとも一部は、前記左混合インバンド信号を含む

ように前記プロセッサをさらに構成するプログラムコードを含み、

前記右クロストーク処理された成分を生成するように前記プロセッサを構成する前記プログラムコードは、

右混合インバンド信号および右混合アウトオブバンド信号に前記右混合チャンネルを分割し、前記右混合チャンネルの前記少なくとも一部は、前記右混合インバンド信号を含む

ように前記プロセッサをさらに構成するプログラムコードを含むことを特徴とする請求項 12 に記載のコンピュータ可読媒体。

30

【請求項 17】

実行されると、前記左混合チャンネルおよび前記右混合チャンネルを含む前記混合ストリームを前記クライアントデバイスに提供するように前記プロセッサを構成するプログラムコードをさらに含むことを特徴とする請求項 12 に記載のコンピュータ可読媒体。

【請求項 18】

実行されると、ネットワークを介して前記クライアントデバイスから前記配置データを受信するように前記プロセッサを構成する命令を含む前記配置データを決定するように前記プロセッサをさらに構成するプログラムコードをさらに含むことを特徴とする請求項 12 に記載のコンピュータ可読媒体。

【請求項 19】

前記複数のクライアントデバイスからの前記入力オーディオストリームの各々は、複数のオーディオチャンネルを含むことを特徴とする請求項 12 に記載のコンピュータ可読媒体。

40

【請求項 20】

前記配置データに従って前記他のクライアントデバイスの前記入力オーディオストリームを混合しパンすることによって、前記左混合チャンネルおよび前記右混合チャンネルを含む前記クライアントデバイスのための前記混合ストリームを生成するように前記プロセッサを構成する前記プログラムコードは、

前記サウンドフィールド内の前記他のクライアントデバイスの前記空間位置に従って、前記他のクライアントデバイスの前記入力オーディオストリームの各々から前記左チャンネルおよび前記右チャンネルを生成し、

50

前記左混合チャンネルを生成するために前記他のクライアントデバイスの前記入力オーディオストリームから前記左チャンネルを結合し、  
前記右混合チャンネルを生成するために前記他のクライアントデバイスの前記入力オーディオストリームから前記右チャンネルを結合する  
ように前記プロセッサをさらに構成するプログラムコードを含むことを特徴とする請求項 1 2 に記載のコンピュータ可読媒体。

【請求項 2 1】

前記複数のクライアントデバイスに接続されたサーバーは、前記混合ストリームを生成し、前記混合ストリームを前記クライアントデバイスに提供し、  
前記クライアントデバイスは、前記混合ストリームから前記クロストーク処理されたストリームを生成する  
ことを特徴とする請求項 1 2 に記載のコンピュータ可読媒体。

10

【請求項 2 2】

前記複数のクライアントデバイスに接続されたサーバーは、前記混合ストリームおよび前記クロストーク処理されたストリームを生成し、前記クロストーク処理されたストリームを前記クライアントデバイスに提供することを特徴とする請求項 1 2 に記載のコンピュータ可読媒体。

【請求項 2 3】

前記クライアントデバイスは、前記混合ストリームおよび前記クロストーク処理されたストリームを生成することを特徴とする請求項 1 2 に記載のコンピュータ可読媒体。

20

【請求項 2 4】

複数のクライアントデバイスのうちの一つのクライアントデバイスにカンファレンスを提供するためのシステムであって、

前記複数のクライアントデバイスから入力オーディオストリームを受信し、  
前記複数のクライアントデバイスのうちの前記クライアントデバイスに対して、サウンドフィールド内の前記複数のクライアントデバイスのうちの他のクライアントデバイスの空間位置を定義する配置データを決定し、

前記配置データに従って前記複数のクライアントデバイスのうちの前記他のクライアントデバイスの入力オーディオストリームを混合しパンすることによって、左混合チャンネルおよび右混合チャンネルを含む、前記複数のクライアントデバイスのうちの前記クライアント  
デバイスのための混合ストリームを生成し、

30

前記クライアントデバイスに関連付けられたスピーカーのタイプを決定し、  
前記混合ストリームの前記左混合チャンネルと前記右混合チャンネルにクロストーク処理を適用することによって、左チャンネルおよび右チャンネルを含むクロストーク処理されたストリームを生成し、前記クロストーク処理は、

前記左混合チャンネルの少なくとも一部をフィルタリングし時間遅延させることによって、左クロストーク処理された成分を生成することと、

前記右混合チャンネルの少なくとも一部をフィルタリングし時間遅延させることによって、右クロストーク処理された成分を生成することと、

前記左混合チャンネルと前記右クロストーク処理された成分を結合することによって、前記  
左チャンネルを生成することと、

40

前記右混合チャンネルと前記左クロストーク処理された成分を結合することによって、前記  
右チャンネルを生成することと

を含み、

スピーカーの第 1 のタイプが前記クライアントデバイスに関連付けられると決定することに  
応答して、前記クロストーク処理はクロストークシミュレーションであり、

スピーカーの第 2 のタイプが前記クライアントデバイスに関連付けられると決定することに  
応答して、前記クロストーク処理はクロストークキャンセレーションである

ように構成される処理回路を備えたことを特徴とするシステム。

【請求項 2 5】

50

前記処理回路は、前記混合ストリームに前記クロストーク処理を適用することによって引き起こされるスペクトル障害を調整する、前記混合ストリームへのクロストーク補正を適用するようにさらに構成されることを特徴とする請求項 2 4 に記載のシステム。

【請求項 2 6】

前記複数のクライアントデバイスからの前記入力オーディオストリームの各々は、複数のオーディオチャンネルを含むことを特徴とする請求項 2 4 に記載のシステム。

【請求項 2 7】

前記処理回路は、前記左混合チャンネルおよび前記右混合チャンネルを含む前記混合ストリームを前記複数のクライアントデバイスのうちの前記クライアントデバイスに提供するようにさらに構成されることを特徴とする請求項 2 4 に記載のシステム。

10

【請求項 2 8】

前記処理回路は、前記混合ストリームの前記左混合チャンネルおよび前記右混合チャンネルにサブバンド空間処理を適用することによって、左強調チャンネルおよび右強調チャンネルを含む空間強調ストリームを生成し、前記サブバンド空間処理は、

前記左混合チャンネルおよび前記右混合チャンネルのミッド成分のミッドサブバンド成分に第 1 のゲインを適用して、強調ミッド成分を生成することと、

前記左混合チャンネルおよび前記右混合チャンネルのサイド成分のサイドサブバンド成分に第 2 のゲインを適用して、強調サイド成分を生成することと、

前記強調ミッド成分および前記強調サイド成分を使用して、前記左強調チャンネルおよび前記右強調チャンネルを生成することと

20

を含むようにさらに構成されることを特徴とする請求項 2 4 に記載のシステム。

【請求項 2 9】

前記左クロストーク処理された成分を生成することは、左混合インバンド信号および左混合アウトオブバンド信号に前記左混合チャンネルを分割することを含み、

前記左混合チャンネルの前記少なくとも一部は、前記左混合インバンド信号を含み、

前記右クロストーク処理された成分を生成することは、右混合インバンド信号および右混合アウトオブバンド信号に前記右混合チャンネルを分割することを含み、

前記右混合チャンネルの前記少なくとも一部は、前記右混合インバンド信号を含む

ことを特徴とする請求項 2 4 に記載のシステム。

【請求項 3 0】

30

前記配置データを決定することは、ネットワークを介して前記クライアントデバイスから前記配置データを受信することを含むことを特徴とする請求項 2 4 に記載のシステム。

【請求項 3 1】

前記配置データに従って前記他のクライアントデバイスの前記入力オーディオストリームを混合しパンすることによって、前記左混合チャンネルおよび前記右混合チャンネルを含む前記クライアントデバイスのための前記混合ストリームを生成することは、

前記サウンドフィールド内の前記他のクライアントデバイスの前記空間位置に従って、前記他のクライアントデバイスの前記入力オーディオストリームの各々から左チャンネルおよび右チャンネルを生成することと、

前記左混合チャンネルを生成するために前記他のクライアントデバイスの前記入力オーディオストリームから左チャンネルを結合することと、

40

前記右混合チャンネルを生成するために前記他のクライアントデバイスの前記入力オーディオストリームから前記右チャンネルを結合することと

を含むことを特徴とする請求項 2 4 に記載のシステム。

【請求項 3 2】

前記複数のクライアントデバイスに接続されたサーバーは、前記混合ストリームを生成し、前記混合ストリームを前記クライアントデバイスに提供し、

前記クライアントデバイスは、前記混合ストリームから前記クロストーク処理されたストリームを生成する

ことを特徴とする請求項 2 4 に記載のシステム。

50

**【請求項 3 3】**

前記複数のクライアントデバイスに接続されたサーバーは、前記混合ストリームおよび前記クロストーク処理されたストリームを生成し、前記クロストーク処理されたストリームを前記クライアントデバイスに提供することを特徴とする請求項 2 4 に記載のシステム。

**【請求項 3 4】**

前記サーバーは、前記クライアントデバイスからデバイスの説明を受信し、前記デバイスの説明に基づいて、前記クロストーク処理を適用するためのパラメータを決定することを特徴とする請求項 3 3 に記載のシステム。

**【請求項 3 5】**

前記クライアントデバイスは、前記混合ストリームおよび前記クロストーク処理されたストリームを生成することを特徴とする請求項 2 4 に記載のシステム。

10

**【発明の詳細な説明】****【技術分野】****【0001】**

本明細書で説明される主題は、オーディオ処理、特にカンファレンスクライアントデバイスの空間的オーディオ処理に関する。

**【背景技術】****【0002】**

電子デバイスは、複数のユーザー間のリモートカンファレンスを提供するために使用される。通常、ユーザーのオーディオストリームはユーザーのオーディオをキャプチャするために生成され、他のユーザーのオーディオストリームはユーザーが聞くためのオーディオを提供するために組み合わせられる。例えば、合成されたストリームは、スピーカーのモノラルストリームであり得る。ステレオスピーカーの場合、モノラルストリームは左右のスピーカーで再生される。対面カンファレンスとは異なり、モノラルストリームの参加者ごとにサウンドフィールド（音場）の空間的な感覚がないため、オーディオの区別と明瞭性が低下する。

20

**【発明の概要】****【0003】**

実施形態は、空間的オーディオを使用してクライアントデバイスにカンファレンスを提供することに関連する。いくつかの実施形態では、入力オーディオストリームはクライアントデバイスから受信される。クライアントデバイスの場合、サウンドフィールド内の他のクライアントデバイスの空間的な位置を定義する配置データが決定される。クライアントデバイスの左混合チャンネルと右混合チャンネルを含む混合ストリームは、配置データに従って他のクライアントデバイスの入力オーディオストリームを混合およびパンすることによって生成される。左強調チャンネルと右強調チャンネルを含む空間強調ストリームは、サブバンド空間処理とクロストーク処理を混合ストリームの左混合チャンネルと右混合チャンネルに適用することによって生成される。空間的に強調されたストリームがクライアントデバイスに提供される。

30

**【0004】**

いくつかの実施形態では、プロセッサによって実行されると命令を格納する非一時的なコンピュータ可読媒体はクライアントデバイスから入力オーディオストリームを受信するように構成される、クライアントデバイスの場合、サウンドフィールド内の他のクライアントデバイスの空間位置を定義する配置データを決定する。そして、配置データに従って他のクライアントデバイスの入力オーディオストリームを混合およびパンすることにより、左混合チャンネルおよび右混合チャンネルを含むクライアントデバイス用の混合ストリームを生成する。

40

**【0005】**

いくつかの実施形態では、システムはクライアントデバイスのためのカンファレンスを提供する。システムは、以下のように構成された処理回路を含む。クライアントデバイスから入力オーディオストリームを受信する。クライアントデバイスの場合、サウンドフィー

50

ルド内の他のクライアントデバイスの空間位置を定義する配置データを決定する。そして、配置データに従って他のクライアントデバイスの入力オーディオストリームを混合およびパンすることにより、左混合チャンネルおよび右混合チャンネルを含むクライアントデバイス用の混合ストリームを生成する。

【0006】

その他の側面として、要素、デバイス、システム、改善、方法、プロセス、アプリケーション、コンピュータ可読媒体、および上記のいずれかに関連するその他の技術が含まれる。

【図面の簡単な説明】

【0007】

【図1】いくつかの実施形態による、オーディオ処理システムの概略ブロック図である。 10

【図2】いくつかの実施形態による、空間混合プロセッサの概略ブロック図である。

【図3】いくつかの実施形態による、クロストークキャンセルを備えた空間強調プロセッサの概略ブロック図である

【図4】いくつかの実施形態によるサブバンド空間プロセッサの概略ブロック図である。

【図5】いくつかの実施形態によるクロストーク補正プロセッサの概略ブロック図である。

【図6】いくつかの実施形態によるクロストークキャンセルプロセッサのブロック図である。

【図7】いくつかの実施形態による、クロストークシミュレーションによる空間強調プロセッサのブロック図である。

【図8】いくつかの実施形態によるクロストークシミュレーションプロセッサのブロック図である。 20

【図9】いくつかの実施形態によるクライアントデバイス間のカンファレンスのための処理のフローチャートである。

【図10】いくつかの実施形態によるオーディオ処理システムの図である。

【図11】いくつかの実施形態によるオーディオ処理システムの図である。

【図12】いくつかの実施形態によるオーディオ処理システムの図である。

【図13】いくつかの実施形態によるコンピュータシステムの概略ブロック図である。

【0008】

図面は、および詳細な説明は、例示のみの目的のための様々な非限定的な実施形態を描写する。 30

【発明を実施するための形態】

【0009】

ここで、実施形態を詳細に参照し、その例を添付図面に示す。以下の説明は、ある特定の具体的な詳細を、様々な実施形態の徹底した理解を提供するために示す。ただし、これらの具体的な詳細なしに、記載されている実施形態を実施することができる。その他の事例では、明確な方法、手順、構成要素、回路、およびネットワークについては、実施形態の態様を不必要に曖昧にしないように詳細に説明されていない。

【0010】

実施形態は、クライアントデバイスごとに空間的オーディオ出力を使用して、クライアントデバイスにカンファレンスを提供することに関連している。例えば、オーディオ処理システムは、リアルタイム空間混合と空間強調の組み合わせを使用して、各クライアントデバイスのレンダリングデバイスタイプに適応的に最適化して、各クライアントデバイスに空間化されたオーディオを配信する。クライアントデバイスの出力ストリームを生成するために、他のクライアントデバイスからのオーディオストリームはそれぞれ、クライアントデバイスのサウンドフィールド内の空間的な位置に関連付けられる。空間的な場所に応じてオーディオストリームが混合およびパンされて、空間化されたオーディオを含む混合ストリームが生成される。次に、空間強調を空間的なオーディオ混合ストリームに適用して、サウンドフィールドの空間感覚を強調する。空間強調には、サブバンド空間処理やクロストーク処理が含まれる場合がある。クロストーク処理には、クロストークキャンセル（スピーカーなど）やクロストークシミュレーション（ヘッドフォンなど）が含まれる。 40



特に、リモート参加者の声が異なるサウンドフィールドの空間的な感覚によって、リモートカンファレンスの声の区別と明瞭さが向上する。

【0011】

カンファレンスパイプラインには、サーバー（または「ブリッジ」）とおよび2台以上のクライアントデバイスが含まれ得る。クライアントデバイスには、さまざまなサウンドレンダリングハードウェアが搭載されている場合がある。サーバーは、さまざまなタイプのハードウェアに対応し、デバイスに最適化された空間強調オーディオを提供する。

【0012】

いくつかの実施形態では、着信オーディオストリームの1つまたは複数のチャンネルがステレオオーディオストリームにミックスダウンされる。次に、各クライアントデバイスのレンダリングシステム（内蔵のラップトップスピーカ、Bluetoothスピーカ、ヘッドホンなど）に基づいて、デバイス固有の空間オーディオ強調機能が適応的に適用される。いくつかの実施形態では、各クライアントデバイスはサーバーから混合ストリームを受信し、適切な空間拡張を適用される。いくつかの実施形態では、クライアントデバイスは、空間強調のためのパラメータをサーバーに提供することができ、サーバーは、空間強調を実行して、クライアントデバイスのための空間強調ストリームを生成する。いくつかの実施形態では、クライアントデバイスは、1つまたは複数のクライアントオーディオストリームを受信し、ミックスダウンとそれに続く適切な空間強調の両方を混合ステレオストリームに適用することができる。

【0013】

（オーディオ処理システムの例）

図1は、いくつかの実施形態による、オーディオ処理システム100の概略ブロック図である。オーディオ処理システム100は、サーバー102およびクライアントデバイス130を含む。単一のクライアントデバイス130が示されているが、サーバー102は、複数のクライアントデバイス130間のカンファレンスを提供する。サーバー102は、インターネットを含むネットワークを介してクライアントデバイスに接続されてもよい。各クライアントデバイス130は、オーディオキャプチャデバイス156（例えば、マイクフォン）と、左スピーカ152および右スピーカ154とを含み得る。

【0014】

サーバー102は、空間混合プロセッサ110および空間強調プロセッサ120を含む。複数のクライアントデバイス130間のカンファレンスを容易にするために、空間混合プロセッサ110は、それぞれクライアントデバイス130から複数の入力オーディオストリーム140を受け取り、入力オーディオストリーム140のパンおよびミックスダウンをステレオオーディオストリームに適用する。ステレオオーディオストリームは、左混合チャンネル142および右混合チャンネル144を含む。入力オーディオストリーム140のそれぞれは、クライアントデバイス130のオーディオキャプチャデバイス156によって生成された、1つまたは複数のチャンネルを有するオーディオストリームであり得る。いくつかの実施形態では、オーディオストリーム140は、モノラルオーディオストリームを含み得る。

【0015】

カンファレンスの各参加者は、クライアントデバイス130に関連付けられる。各クライアントデバイス130は、空間混合プロセッサ110に送信されるオーディオストリーム140を生成することができる。オーディオストリーム140に関連して、各クライアントデバイス130は、デバイス識別データまたは配置データを含む様々なタイプのメタデータを空間混合プロセッサ110に提供することができる。デバイス識別データは、クライアントデバイス130を識別する一意の識別子である。配置データは、クライアントデバイス130のための他のクライアントデバイス130の他のオーディオストリーム140の空間位置を定義する。空間混合プロセッサ110は、他のクライアントデバイス130からのオーディオ信号140が配置データに従ってステレオサウンドフィールドにわたって分散される（例えば、パンされる）各クライアントデバイス130のための固有のス

10

20

30

40

50

テレオオーディオストリームを生成する。

【0016】

いくつかの実施形態では、オーディオストリーム140のパンニングは、オーディオストリーム140の空間分布を定義する所定の配置データに基づく。配置技術は、リモートクライアントデバイス130の等しい空間分布、または参加者メタデータ（例えば、ユーザーの役割、エンティティ/会社名、グループメンバーシップ、部屋の場所、地理的な場所、IPアドレスなど）に基づくソースのグループ化およびパンニングを含み得る。いくつかの実施形態では、各クライアントデバイス130のパンニングは、クライアントデバイス130からの制御に基づいてもよい。例えば、クライアントデバイス130のユーザーは、各遠隔参加者の所望のパン位置を指定する配置データを定義することができる。

10

【0017】

空間混合プロセッサ110は、各クライアントデバイス130について、左混合チャンネル142および右混合チャンネル144を含む固有のステレオオーディオストリームを生成する。オーディオ信号140の空間分布を定義する配置データがプログラムで定義されるか、またはユーザー定義であるかにかかわらず、他のクライアントデバイス130のオーディオ信号140は、ステレオミックスダウンにパンされる。パンニングは、振幅パンニング、遅延パンニング、バイノーラルパンニングなどのパンニング技法を含み、クライアントデバイス130のステレオストリームを生成することができる。いくつかの実施形態では、パンニングは、ゲインに応じて、最終的な混合における各要素の存在を調整することを含むことができる。距離手がかりのために空気吸収シミュレーションや室内音響シミュレーションなど、知覚的に動機づけられるその他のキューも適用され得る。

20

【0018】

いくつかの実施形態では、クライアントデバイス130の空間混合プロセッサ110の出力は、クライアントデバイス130またはリスニングコンテキストが1チャンネル再生を使用するときなどに、単一チャンネルへのモノラル低減を含み得る。

【0019】

空間強調プロセッサ120は、左混合チャンネル142および右混合チャンネル144に空間処理を適用して、左強調チャンネル146および右強調チャンネル148を生成する。空間強調は、サブバンド空間処理およびクロストーク処理を含み得る。クロストーク処理には、クロストークキャンセルやクロストークシミュレーションが含まれ得る。クライアントデバイス130のスピーカー152、154がラウドスピーカーであり、ラウドスピーカーが原因のクロストークを除去する場合に、クロストークキャンセレーションが使用され得る。スピーカー152、154がヘッドフォンでスピーカーの効果を実シミュレートする場合、クロストークシミュレーションを使用することができる。左混合チャンネル142および右混合チャンネル144に適用されるクロストーク処理がある場合、空間強調処理は、クロストーク処理の適用によって引き起こされるスペクトル欠陥を補正するクロストーク補正をさらに含み得る。空間強調プロセッサ120は、クライアントデバイス130の左スピーカー152に左強調チャンネル146を提供し、クライアントデバイス130の右スピーカー154に右強調チャンネル148を提供する。スピーカー152、154は、それぞれの出力チャンネル $O_L$ および $O_R$ をサウンドに変換する。

30

40

【0020】

いくつかの実施形態では、空間強調プロセッサ120は、サーバー102に配置される。サーバー102は、各クライアントデバイス130に対して空間強調処理の別個の過程を実行することができる。他の実施形態では、空間強調プロセッサ120は、各クライアントデバイス130に配置される。クライアントデバイス130ごとに、サーバー102は、左混合チャンネル142および右混合チャンネル144を含む混合ストリームをクライアントデバイス130の空間強調プロセッサ120に提供する。各クライアントデバイス130の空間強調プロセッサ120は、サーバー102から受信された混合チャンネルを処理して、左出力チャンネル146および右出力チャンネル148を生成する。いくつかの実施形態では、空間混合プロセッサ110は、システム100のピアツーピアカンファレンス構成

50

などのクライアントデバイス 130 内に、またはサーバーがミックスダウンを実行しない場合にも配置される。

#### 【0021】

システム 100 は、いくつかのまたは追加の構成要素を含み得る。例えば、オーディオストリームを使用したカンファレンスをビデオストリームと統合することができる。いくつかの実施形態では、システム 100 は人工現実システムであり、各クライアントデバイス 130 はヘッドマウントディスプレイを含み得る。ヘッドマウントディスプレイは、他のユーザーまたはユーザーのアバターをビデオでレンダリングすることができ、ユーザーは、サウンドフィールド内の混合ストリームのクライアントデバイス 130 の空間位置に対応するように配置される。これにより、人工現実環境のイマーシブ品質が向上する。

10

#### 【0022】

(空間混合プロセッサの例)

図 2 はいくつかの実施形態による空間混合プロセッサ 110 の概略ブロック図である。空間混合プロセッサ 110 は、空間ミキサー 205、バイノーラルフィルタ 210 および 215、左チャンネル結合器 220、右チャンネル結合器 225、および知覚シミュレータ 230 を含む。空間ミキサー 205 は、入ってくるオーディオストリーム内のエネルギーが最終的なステレオミックスダウン全体にどのように分配されるか、また関連して、最終的なミックスダウン内の要素がサウンドステージのコンテキスト内でどのように知覚されるかに関して制御を適用する。例えば、モノラルオーディオストリームの空間混合は、定出力またはリニアステレオパンニング技術を使用して実現することができる。いくつかの実施形態では、空間ミキサー 205 は、各混合信号 140 に空間混合を適用して左および右チャンネルを生成し、左チャンネルを結合して左チャンネル 220 を生成し、右チャンネルを結合して右チャンネル 222 を生成する。

20

#### 【0023】

オーディオストリームは、バイノーラルフィルタリングを使用して実現できるように、リスナーの頭部の周囲の 3D 空間にオーディオを知覚的に配置するように処理することもできる。バイノーラルフィルタ 210 と 215 のそれぞれは、リスナーが入力チャンネルの音を知覚する必要があるターゲットソースの場所を記述する頭部伝達関数 (HRTF) を適用する。バイノーラルフィルタ 210 は、左チャンネル 220 を受信し、右チャンネル 250 に関連付けられた角度位置を調整する HRTF を適用することにより、左および右の出力チャンネルを生成する。バイノーラルフィルタ 215 は、右チャンネル 252 を受信し、右チャンネル 252 に関連付けられた角度位置を調整する HRTF を適用することにより、左および右の出力チャンネルを生成する。左チャンネル結合器 220 は、バイノーラルフィルタ 210 および 215 から左チャンネルを受け取り、これらのチャンネルを結合して左チャンネル 254 を生成する。右チャンネル結合器 225 は、バイノーラルフィルタ 210 および 215 から右チャンネルを受け取り、これらのチャンネルを結合して、右チャンネル 256 を生成する。

30

#### 【0024】

また、処理を適用して、他の現実世界の知覚的手がかりをシミュレートすることもできる。知覚シミュレータ 230 は、左チャンネル 254 および右チャンネル 256 を受け取り、知覚的手がかりを 1 つまたは複数のチャンネルに適用して、左混合チャンネル 142 および右混合チャンネル 144 を生成する。チャンネルチャンネルは、例えば自由野逆距離法則 (the free-field inverse distance law) を使用して、混合内の可変距離と空気吸収の知覚を提供するようにスケージングおよびフィルタリングできる。この場合、音圧レベルは、基準距離と比較して距離の 2 倍ごとに 6 デシベル減少する (例えば、リスナーの仮想位置からの定義された距離)。高周波エネルギーが基準距離に対する距離の増加に反比例して減衰する場合、空気吸収をシミュレートすることができる。残響は、距離の認識をさらに高め、室内の音響と環境の感覚を作り出すために使用できる。

40

#### 【0025】

いくつかの実施形態では、オーディオ信号 140 は、複数 (例えば、左および右) のチャンネルを含む。空間ミキサー 205 は、オーディオ信号のチャンネルをモノラルチャンネルに結

50

合することによりミックスダウンを適用し、モノラルチャンネルに定電力またはリニアステレオパンニングを適用する。別の例では、モノラル信号への組み合わせはなく、ミックスダウンはオーディオ信号140の各チャンネルを使用して適用される。例えば、空間ミキサー205は、リスナーの周りの空間におけるその理想的な位置に基づいて、各チャンネルに異なるバイノーラルフィルタリングを適用し、結果をステレオ左および右チャンネルに混合してもよい。

【0026】

空間混合プロセッサ110には、いくつかのまたは追加の構成要素を含み得る。例えば、知覚シミュレータ230またはバイノーラルフィルタ210および215は省略できる。空間混合プロセッサ110は、振幅パンニング、遅延パンニング、バイノーラルパンニングなどを含む、オーディオ信号140の様々なタイプのミックスダウンを実行することができる。

10

【0027】

(空間強調プロセッサの例)

図3は、一実施形態による空間強調プロセッサ300の概略ブロック図である。空間強調プロセッサ300は、クロストーク処理が、スピーカー152、154がラウドスピーカーであるクライアントデバイス130のクロストークキャンセルである空間強調プロセッサ120の例である。空間強調プロセッサ300は、左入力チャンネル $X_L$ および右入力チャンネル $X_R$ を含む入力オーディオ信号 $X$ を受信する。左入力チャンネル $X_L$ は、空間混合プロセッサ110からの左出力チャンネル142であり得、右入力チャンネル $X_R$ は、空間混合プロセッサ110からの右出力チャンネル144であり得る。

20

【0028】

空間強調プロセッサ300は、入力チャンネル $X_L$ および $X_R$ を処理することにより、左強調チャンネル $L$ および右強調チャンネル $O_R$ を含む出力オーディオ信号 $O$ を生成する。左強調チャンネル $O_L$ は左強調チャンネル146に対応し、右強調チャンネル $O_R$ は右強調チャンネル148に対応する。オーディオ出力信号 $O$ は、クロストーク補正およびクロストークキャンセルを備えた入力オーディオ信号 $X$ の空間的に強調されたオーディオ信号である。

【0029】

空間強調プロセッサ300は、サブバンド空間プロセッサ310、クロストーク補正プロセッサ320、結合器360、およびクロストークキャンセルプロセッサ370を含む。空間強調プロセッサ300は、入力オーディオ入力チャンネル $X_L$ 、 $X_R$ のクロストーク補正およびサブバンド空間処理を実行し、サブバンド空間処理の結果をクロストーク補正の結果と結合し、その後、結合された信号に対してクロストークキャンセルを実行する。

30

【0030】

サブバンド空間プロセッサ310は空間周波数帯域ディバイダ340、空間周波数帯域プロセッサ345、および空間周波数帯域結合器350を含む。空間周波数帯域ディバイダ340は、入力チャンネル $X_L$ および $X_R$ ならびに空間周波数帯域プロセッサ345に結合される。空間周波数帯域ディバイダ340は、左入力チャンネル $X_L$ および右入力チャンネル $X_R$ を受け取り、入力チャンネルを空間(または「サイド」)成分 $Y_S$ および非空間(または「ミッド」)成分 $Y_M$ に処理する。例えば、左入力チャンネル $X_L$ と右入力チャンネル $X_R$ の差に基づいて空間成分 $Y_S$ を生成できる。非空間成分 $Y_M$ は、左入力チャンネル $X_L$ と右入力チャンネル $X_R$ の合計に基づいて生成できる。空間周波数帯域ディバイダ340は、空間成分 $Y_S$ と非空間成分 $Y_M$ を空間周波数帯域プロセッサ345に提供する。

40

【0031】

空間周波数帯域プロセッサ345は、空間周波数帯域ディバイダ340および空間周波数帯域結合器350に結合される。空間周波数帯域プロセッサ345は、空間周波数帯域ディバイダ340から空間成分 $Y_S$ と非空間成分 $Y_M$ を受信し、受信信号を強調する。特に、空間周波数帯域プロセッサ345は、空間成分 $Y_S$ から空間強調成分 $E_S$ を生成し、非空間成分 $Y_M$ から非空間強調成分 $E_M$ を生成する。

【0032】

50

例えば、空間周波数帯域プロセッサ 345 は、空間成分  $Y_S$  にサブバンドゲインを適用して、空間強調成分  $E_S$  を生成する。また、非空間成分  $Y_m$  にサブバンドゲインを適用して、非空間強調成分  $E_m$  を生成する。いくつかの実施形態では、空間周波数帯域プロセッサ 345 は、空間成分  $Y_S$  にサブバンド遅延を追加のまたは代替として提供して、空間強調成分  $E_S$  を生成しまた、非空間成分  $Y_m$  へのサブバンド遅延を提供して、非空間強調成分  $E_m$  を生成する。サブバンドのゲインおよび/または遅延は、空間成分  $Y_S$  と非空間成分  $Y_m$  のサブバンド ( $n$  など) によって異なる場合がある。また、同じ場合もある (2 つ以上のサブバンドの場合など)。空間周波数帯域プロセッサ 345 は、空間成分  $Y_S$  と非空間成分  $Y_m$  の異なるサブバンドのゲインおよび/または遅延を相互に調整して、空間強調成分  $E_S$  と非空間強調成分  $E_m$  を生成する。次に、空間周波数帯域プロセッサ 345 は、空間強調成分  $E_S$  と非空間強調成分  $E_m$  を空間周波数帯域結合器 350 に提供する。

10

#### 【0033】

空間周波数帯域結合器 350 は、空間周波数帯域プロセッサ 345 に結合され、さらに結合器 360 に結合される。空間周波数帯域結合器 350 は、空間周波数帯域プロセッサ 345 から空間強調成分  $E_S$  および非空間強調成分  $E_m$  を受け取り、空間強調成分  $E_S$  および非空間強調成分  $E_m$  を左空間強調チャンネル  $E_L$  および右空間強調チャンネルに  $E_R$  結合する。例えば、左空間強調チャンネル  $E_L$  は、空間強調成分  $E_S$  と非空間強調成分  $E_m$  の合計に基づいて生成できる。また、非空間強調成分  $E_m$  と拡張強調成分  $E_S$  の差に基づいて、右空間強調チャンネル  $E_R$  を生成できる。空間周波数帯域結合器 350 は、左空間強調チャンネル  $E_L$  または右空間強調チャンネル  $E_R$  を結合器 360 に提供する。サブバンド空間プロセッサ 310 の追加の詳細は図 4 に関連して以下で説明される。

20

#### 【0034】

クロストーク補正プロセッサ 320 は、クロストーク補正を実行して、クロストークキャンセルのスペクトル欠陥やアーチファクトを補正する。クロストーク補正プロセッサ 320 は、入力チャンネル  $X_L$  および  $X_R$  を受け取り、クロストークキャンセルプロセッサ 370 によって実行される非空間強調成分  $E_m$  および空間強調成分  $E_S$  の後続のクロストークキャンセルにおけるアーチファクトを補正する処理を実行する。いくつかの実施形態では、クロストーク補正プロセッサ 320 は、フィルタを適用して、左クロストーク補正チャンネル  $Z_L$  および右クロストーク補正チャンネル  $Z_R$  を含むクロストーク補正信号  $Z$  を生成することにより、非空間成分  $X_m$  および空間成分  $X_S$  に対して強調を実行し得る。他の実施形態では、クロストーク補正プロセッサ 320 は、非空間成分  $X_m$  のみに対して強調を実行することができる。クロストーク補正プロセッサ 320 に関する追加の詳細は図 5 に関連して以下で説明される。

30

#### 【0035】

結合器 360 は、左空間強調チャンネル  $E_L$  を左クロストーク補正チャンネル  $Z_L$  と結合して左強調補正チャンネル  $T_L$  を生成し、右空間強調チャンネル  $E_R$  を右クロストーク補正チャンネル  $Z_R$  と結合して右補正チャンネル  $T_R$  を生成する。結合器 360 はクロストークキャンセルプロセッサ 370 に結合され、クロストークキャンセルプロセッサ 370 に左強調補正チャンネル  $T_L$  と右強調補正チャンネル  $T_R$  を提供する。

#### 【0036】

クロストークキャンセルプロセッサ 370 は、左強調補正チャンネル  $T_L$  および右強調補正チャンネル  $T_R$  を受け取り、チャンネル  $T_L$ 、 $T_R$  に対してクロストークキャンセルを実行して、左出力チャンネル  $O_L$  および右出力チャンネル  $O_R$  を含む出力オーディオ信号  $O$  を生成する。クロストークキャンセルプロセッサ 370 に関する追加の詳細は、図 5 に関連して以下で説明される。

40

#### 【0037】

図 4 は、いくつかの実施形態による、サブバンド空間プロセッサの概略ブロック図である。空間周波数帯域ディバイダ 340 には、左入力チャンネル  $X_L$  および右入力チャンネル  $X_R$  を受信する L/R/M/S コンバータ 402 を含み、これらの入力を空間成分  $Y_m$  および  $Y_S$  変換する。

50

## 【 0 0 3 8 】

空間周波数帯域プロセッサ 3 4 5 は、非空間成分  $Y_m$  を受信し、サブバンドフィルタのセットを適用して、非空間強調サブバンド成分  $E_m$  を生成する。空間周波数帯域プロセッサ 3 4 5 は空間サブバンド成分  $Y_S$  も受信し、サブバンドフィルタのセットを適用して非空間強調サブバンド成分  $E_m$  を生成する。サブバンドフィルタには、ピークフィルタ、ノッチフィルタ、ローパスフィルタ、ハイパスフィルタ、ローシェルフフィルタ、ハイシェルフフィルタ、バンドパスフィルタ、バンドストップフィルタ、およびすべてのパスフィルタのさまざまな結合を含み得る。

## 【 0 0 3 9 】

より具体的には、空間周波数帯域プロセッサ 3 4 5 は、非空間成分  $Y_m$  の  $n$  個の周波数サブバンドのそれぞれに対するサブバンドフィルタと、空間成分  $Y_S$  の  $n$  個のサブバンドのそれぞれに対するサブバンドフィルタとを含む。  $n = 4$  サブバンドの場合、例えば、空間周波数帯域プロセッサ 3 4 5 には、非空間成分  $Y_m$  用の一連のサブバンドフィルタが含まれている。これには、サブバンド ( 1 ) 用のミッドイコライゼーション (  $E Q$  ) フィルタ 4 0 4 ( 1 )、サブバンド ( 2 ) 用のミッド  $E Q$  フィルタ 4 0 4 ( 2 )、サブバンド ( 3 ) 用のミッド  $E Q$  フィルタ 4 0 4 ( 3 )、サブバンド ( 4 ) 用のミッド  $E Q$  フィルタ 4 0 4 ( 4 ) が含まれている。各ミッド  $E Q$  フィルタ 4 0 4 は、非空間成分  $Y_m$  の周波数サブバンド部分にフィルタを適用して、非空間強調成分  $E_m$  を生成する。

## 【 0 0 4 0 】

空間周波数帯域プロセッサ 3 4 5 には、空間成分  $Y_S$  の周波数サブバンド用の一連のサブバンドフィルタが含まれている。これには、サブバンド ( 1 ) 用のサイドイコライゼーション (  $E Q$  ) フィルタ 4 0 6 ( 1 )、サブバンド ( 2 ) 用のサイド  $E Q$  フィルタ 4 0 6 ( 2 )、サブバンド ( 3 ) 用のサイド  $E Q$  フィルタ 4 0 6 ( 3 )、サブバンド ( 4 ) 用のサイド  $E Q$  フィルタ 4 0 6 ( 4 ) が含まれている。各サイド  $E Q$  フィルタ 4 0 6 は、空間成分  $Y_S$  の周波数サブバンド部分にフィルタを適用して、空間強調成分  $E_S$  を生成する。

## 【 0 0 4 1 】

非空間成分  $Y_m$  および空間成分  $Y_S$  の  $n$  個の周波数サブバンドは、それぞれ周波数の範囲に対応し得る。例えば、周波数サブバンド ( 1 ) は  $0 \sim 300 \text{ Hz}$  に対応し、周波数サブバンド ( 2 ) は  $300 \sim 510 \text{ Hz}$  に対応し、周波数サブバンド ( 3 ) は  $510 \sim 2700 \text{ Hz}$  に対応し、周波数サブバンド ( 4 ) は  $2700 \text{ Hz} \sim$  ナイキスト周波数に対応する。いくつかの実施形態では、  $n$  周波数サブバンドは重要なバンドの統合セットである。重要なバンドは色々な音楽的なジャンルからの可聴周波数サンプルのコーパスを使用して定められ得る。  $24$  バーク尺度の臨界帯域におけるミッド成分とサイド成分の長期平均エネルギー比は、サンプルから決定される。次に、同様の長期平均比を持つ連続周波数帯域をグループ化して、重要な帯域のセットを形成する。周波数サブバンドの範囲と周波数サブバンドの数は調整することができる。

## 【 0 0 4 2 】

空間周波数帯域結合器 3 5 0 は、ミッドとサイドの成分を受け取り、各成分にゲインを適用し、ミッドとサイドの成分を左右のチャンネルに変換する。例えば、空間周波数帯域結合器 3 5 0 は、非空間強調成分  $E_m$  および空間強調成分  $E_S$  を受信し、非空間強調成分  $E_m$  および空間強調成分  $E_S$  を左空間強調チャンネル  $E_L$  および右空間強調チャンネル  $E_R$  に変換する前に全体的なミッドおよびサイドゲインを実行する。

## 【 0 0 4 3 】

具体的には、空間周波数帯域結合器 3 5 0 には、グローバルミッドゲイン 4 0 8、グローバルサイドゲイン 4 1 0、およびグローバルミッドゲイン 4 0 8 およびグローバルサイドゲイン 4 1 0 に結合された  $M / S / L / R$  コンバータ 4 1 2 が含まれる。グローバルミッドゲイン 4 0 8 は非空間強調成分  $E_m$  を受信してゲインを適用し、グローバルサイドゲイン 4 1 0 は非空間強調成分  $E_S$  を受信してゲインを適用する。  $M / S / L / R$  コンバータ 4 1 2 はグローバルミッドゲイン 4 0 8 から非空間強調成分  $E_m$ 、グローバルサイドゲイン 4 1 0 から空間強調成分  $E_S$  を受信する。そして、これらの入力を左空間強調チャンネル  $E_L$

10

20

30

40

50

および右空間強調チャンネル  $E_R$  に変換する。

【 0 0 4 4 】

( クロストークキャンセルプロセッサの例 )

図 5 はいくつかの実施形態によるクロストーク補正プロセッサ 3 2 0 の概略ブロック図である。クロストーク補正プロセッサ 3 2 0 には、L / R / M / S コンバータ 5 0 2 ミッド成分プロセッサ 5 2 0、サイド成分プロセッサ 5 3 0、M / S / L / R コンバータ 5 1 4 が含まれる。

【 0 0 4 5 】

クロストーク補正プロセッサ 3 2 0 は、左右の入力チャンネル  $X_L$ 、 $X_R$  を受信し、クロストーク補正処理を実行して、左クロストーク補正チャンネル  $Z_L$  および右クロストーク補正チャンネル  $Z_R$  を生成する。チャンネル  $Z_L$ 、 $Z_R$  は、クロストークキャンセルやクロストークシミュレーションなど、クロストーク処理におけるアーチファクトやスペクトル障害を補正するために使用できる。L / R / M / S コンバータ 5 0 2 は、左入力チャンネル  $X_L$  と右入力チャンネル  $X_R$  を受信するし、入力チャンネル  $X_L$ 、 $X_R$  の非空間成分  $X_m$  と空間成分  $X_S$  を生成する。左右のチャンネルを合計して左右のチャンネルの非空間成分を生成し、サブトラクションして左右のチャンネルの空間成分を生成することができる。

10

【 0 0 4 6 】

ミッド成分プロセッサ 5 2 0 には、 $m$  ミッドフィルタ 5 4 0 ( a )、5 4 0 ( b ) ~ 5 4 0 ( m ) などの複数のフィルタ 5 4 0 が搭載されている。ここでは、各  $m$  ミッドフィルタ 5 4 0 は、非空間成分  $X_m$  の  $m$  周波数バンドの 1 つを処理する。ミッド成分プロセッサ 5 2 0 は、非空間成分  $X_m$  を処理することによって、ミッドクロストーク補正チャンネル  $Z_m$  を生成する。いくつかの実施形態によると、ミッドフィルタ 5 4 0 は、シミュレーションによるクロストーク処理を使用した非空間  $X_m$  の周波数応答プロットを使用して構成されている。また、周波数応答プロットを解析することにより、クロストーク処理のアーチファクトとして発生する周波数応答プロットのピークやトラフなどのスペクトル障害を、あらかじめ設定されたしきい値 ( 1 0 d B など ) を超えて推定することができる。これらのアーチファクトは、主にクロストーク処理における遅延し、おそらく反転した逆測信号と、対応する同側信号との和に起因し、最終的なレンダリング結果にくし形フィルタのような周波数応答を効果的にもたらず。ミッドクロストーク補正チャンネル  $Z_m$  は、ミッド成分プロセッサ 5 2 0 によって生成され、各  $m$  周波数帯域がピークまたはトラフに対応する推定ピークまたはトラフを補正できる。具体的には、クロストーク処理で適用される特定の遅延、フィルタリング周波数、およびゲインに基づいて、周波数応答でピークとトラフが上下に移動し、スペクトルの特定の領域におけるエネルギーの増幅および / または減衰を引き起こす。各ミッドフィルタ 5 4 0 は、1 つまたは複数のピークとトラフに合わせて調整するように設定できる。

20

30

【 0 0 4 7 】

サイド成分プロセッサ 5 3 0 には、 $m$  サイドフィルタ 5 5 0 ( a )、5 5 0 ( b ) ~ 5 5 0 ( m ) などの複数のフィルタ 5 5 0 が含まれる。サイド成分プロセッサ 5 3 0 は、空間成分  $X_S$  を処理することにより、サイドクロストーク補正チャンネル  $Z_S$  を生成する。いくつかの実施形態によると、シミュレーションによって空間  $X$  の周波数応答プロットとクロストーク処理を得ることができる。周波数応答プロットを解析することにより、クロストーク処理のアーチファクトとして発生する周波数応答プロットのピークやトラフなどのスペクトル障害を、あらかじめ設定されたしきい値 ( 1 0 d B など ) を超えて推定できる。サイドクロストーク補正チャンネル  $Z_S$  は、サイド成分プロセッサ 5 3 0 によって生成され、推定ピークまたはトラフを補正することができる。具体的には、クロストーク処理で適用される特定の遅延、フィルタリング周波数、およびゲインに基づいて、周波数応答でピークとトラフが上下に移動し、スペクトルの特定の領域におけるエネルギーの増幅および / または減衰を引き起こす。各サイドフィルタ 5 5 0 は、1 つまたは複数のピークおよびトラフに合わせて調整するように設定できる。いくつかの実施形態では、ミッド成分プロセッサ 5 2 0 とサイド成分プロセッサ 5 3 0 に異なる数のフィルタが含まれている場合がある。

40

50

【 0 0 4 8 】

いくつかの実施形態では、ミッドフィルタ 5 4 0 とサイドフィルタ 5 5 0 には、式 1 で定義された転送機能を持つ複二次フィルタを含むことができる。

【 0 0 4 9 】

【数 1】

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{a_0 + a_1 z^{-1} + a_2 z^{-2}} \quad \text{式 (1)}$$

【 0 0 5 0 】

$z$  は複素変数で、 $a_0$ 、 $a_1$ 、 $a_2$ 、 $b_0$ 、 $b_1$ 、および  $b_2$  はデジタルフィルタ係数である。このようなフィルタを実装する 1 つの方法は、式 2 で定義されたダイレクトフォーム I トポロジである。

【 0 0 5 1 】

【数 2】

$$Y[n] = \frac{b_0}{a_0} X[n-1] + \frac{b_1}{a_0} X[n-1] + \frac{b_2}{a_0} X[n-2] - \frac{a_1}{a_0} Y[n-1] - \frac{a_2}{a_0} Y[n-2] \quad \text{式 (2)}$$

【 0 0 5 2 】

ここで、 $X$  は入力ベクトル、 $Y$  は出力である。ほかのトポロジは最大ワード長および飽和動作に応じて、使用される。

【 0 0 5 3 】

その後、複二次を使用して、実値の入出力を持つ 2 次フィルタを実装できる。離散時間フィルタを設計するために、連続時間フィルタが設計され、双一次変換によって離散時間に変換される。さらに、中心周波数と帯域幅のシフトは、周波数歪みを使用して補正できる。

【 0 0 5 4 】

例えば、ピークフィルタは、式 3 で定義され複素平面転送機能がある。

【 0 0 5 5 】

【数 3】

$$H(s) = \frac{s^2 + s\left(\frac{A}{Q}\right) + 1}{s^2 + s\left(\frac{A}{Q}\right) + 1} \quad \text{式 (3)}$$

【 0 0 5 6 】

ここで、 $S$  は複素変数、 $A$  はピークの振幅、 $Q$  はフィルタ「品質」、そしてデジタルフィルタ係数は次のように定義される。

【 0 0 5 7 】

【数 4】

$$b_0 = 1 + \alpha A$$

$$b_1 = -2 * \cos(\omega_0)$$

$$b_2 = 1 - \alpha A$$

$$a_0 = 1 + \frac{\alpha}{A}$$

$$a_1 = -2 \cos(\omega_0)$$

$$a_2 = 1 + \frac{\alpha}{A}$$

10

20

30

40

50



【 0 0 5 8 】

ここで、 $\omega_0$  は、フィルタの中心周波数をラジアンおよび

【 0 0 5 9 】

【数 5】

$$\alpha = \frac{\sin(\omega_0)}{2Q}$$

【 0 0 6 0 】

で表したものである。

【 0 0 6 1 】

さらに、フィルタ品質  $Q$  は式 4 で定義できる。

【 0 0 6 2 】

【数 6】

$$Q = \frac{f_c}{\Delta f} \quad \text{式 (4)}$$

【 0 0 6 3 】

ここで

【 0 0 6 4 】

【数 7】

$\Delta f$

【 0 0 6 5 】

、は帯域幅、 $f_c$  は中心周波数である。

【 0 0 6 6 】

ミッドフィルタ 5 4 0 は連続にあるものとして示され、そしてサイドフィルタ 5 5 0 も連続にあるものとして示されているいくつかの実施形態では、ミッドフィルタ 5 4 0 がミッド成分  $X_m$  と平行に適用され、サイドフィルタがサイド成分  $X_s$  と平行に適用される。

【 0 0 6 7 】

M / S / L / R コンバータ 5 1 4 は、ミッドクロストーク補正チャネル  $Z_m$  とサイドクロストーク補正チャネル  $Z_s$  を受信し、左クロストーク補正チャネル  $Z_L$  と右クロストーク補正チャネル  $Z_R$  を生成する。ミッドとサイドの成分を合計して、ミッドとサイドの成分の左チャネルを生成できる。ミッドとサイドの成分を減算して、ミッドとサイドの成分の右チャネルを生成できる。

【 0 0 6 8 】

(クロストークキャンセルプロセッサの例)

図 6 は、いくつかの実施形態によるクロストークキャンセルプロセッサ 3 7 0 の概略ブロック図である。クロストークキャンセレーションプロセッサ 3 7 0 は、左強調補正チャネル  $T_L$  と右強調補正チャネル  $T_R$  を受信する。また、チャネル  $T_L$ 、 $T_R$  でクロストークキャンセルを実行して、左出力チャネル  $O_L$ 、右出力チャネル  $O_R$  を生成する。

【 0 0 6 9 】

クロストークキャンセルプロセッサ 3 7 0 には、インアウトバンドディバイダ 6 1 0、インバータ 6 2 0 および 6 2 2 逆測推定器 6 3 0 および 6 4 0、結合器 6 5 0 および 6 5 2、インアウトバンド結合器 6 6 0 が含まれる。これらの構成要素は、入力チャネル  $T_L$ 、 $T_R$  をインバンド成分およびアウトオブバンド成分に分割し、インバンド成分でクロストークキャンセレーションを実行して出力チャネル  $O_L$ 、 $O_R$  を生成する。

【 0 0 7 0 】

10

20

30

40

50

入力オーディオ信号Tを異なる周波数帯域成分に分割し、選択的成分（インバンド成分など）でクロストークキャンセルを実行することで、他の周波数帯域での劣化をなくしながら、特定の周波数帯域でクロストークキャンセルを実行できる。入力オーディオ信号Tを異なる周波数帯域に分割せずにクロストークキャンセルを実行すると、クロストークキャンセル後のオーディオ信号は、低周波数（350Hz未満など）、高周波数（12000Hz以上など）、または両方での非空間成分および空間成分で大きな減衰または増幅を示す場合がある。影響の大きい空間的手がかりの大部分が存在するインバンド（250Hz～14000Hzなど）のクロストークキャンセルを選択的に実行することで、混合内のスペクトル全体にわたって、特に非空間的な成分でバランスのとれた全体的なエネルギーを維持できる。

10

#### 【0071】

インアウトバンドディバイダ610は、入力チャンネル $T_L$ 、 $T_R$ をインバンドチャンネル $T_{L,IN}$ 、 $T_{R,IN}$ およびアウトバンドチャンネル $T_{L,OUT}$ 、 $T_{R,OUT}$ にそれぞれ分離する。特に、インアウトバンドディバイダ610は、左強調補正チャンネル $T_L$ を左インバンドチャンネル $T_{L,IN}$ 、左アウトオブバンドチャンネル $T_{L,OUT}$ に分割する。同様に、インアウトバンドディバイダ610は、右強調補正チャンネル $T_R$ を右インバンドチャンネル $T_{R,IN}$ 、および右アウトオブバンドチャンネル $T_{R,OUT}$ に分離する。各インバンドチャンネルは、例えば250Hz～14kHzなど、周波数範囲に対応する各入力チャンネルの一部を包含し得る。周波数帯域の範囲は、スピーカーのパラメータなどに応じて調整することができる。

#### 【0072】

インバーター620と逆測推定器630は、左インバンドチャンネル $T_{L,IN}$ による逆測サウンド成分を補正するために、左逆測キャンセル成分 $S_L$ を生成するように一緒に動作する。同様に、インバーター622と逆測推定器640は、右インバンドチャンネル $T_R$ のために逆測サウンド成分を補正するために、右逆測キャンセル成分 $S_R$ を生成するように一緒に動作する。

20

#### 【0073】

1つの手法では、インバーター620は、インバンドチャンネル $T_{L,IN}$ を受け取り、受け取ったインバンドチャンネル $T_{L,IN}$ の極性を反転して、反転されたインバンドチャンネル $T_{L,IN}'$ を生成する。逆測推定器630は、反転されたインバンドチャンネル $T_{L,IN}'$ を受け取り、フィルタリングを通じて、逆測音成分に対応する反転されたインバンドチャンネル $T_{L,IN}'$ の一部を抽出する。フィルタリングは、反転されたインバンドチャンネル $T_{L,IN}'$ で実行されるので、反逆測推定器630によって抽出された部分は、逆測のサウンド成分に起因するインバンドチャンネル $T_L$ の一部の逆になる。したがって、反逆測の推定器630によって抽出された部分は、左の反逆測のキャンセル成分 $S_L$ となり、これを対応するインバンドチャンネル $T_{R,IN}$ に追加することができ、インバンドチャンネル $T_{L,IN}$ による反逆測の音成分を低減する。いくつかの実施形態では、インバーター620と逆測推定器630は、異なる順序で実装される。

30

#### 【0074】

インバーター622と逆測推定器640は、インバンドチャンネル $T_R$ に関して同様の操作を行い、右側逆測キャンセル成分 $S_R$ を生成する。従って、その詳細な説明は、本明細書では簡潔さのために省略される。

40

#### 【0075】

1つの例示的実装形態では、逆測推定器825Aは、フィルタ852A、アンプ854A、及び遅延ユニット856Aを含む。フィルタ632は反転入力チャンネル $T_{L,IN}'$ を受け取り、フィルタリング機能を介して反逆測のサウンド成分に対応する反転インバンドチャンネル $T_{L,IN}'$ の一部を抽出する。フィルタ実装の例としては、中心周波数が5000～10000Hzで選択され、Qが0.5～1.0の範囲で選択されたNotchまたはハイシェルフフィルタがある。デシベル（GdB）単位のゲインは、式5から算出できる。

$$G_{dB} = -3.0 - \log 1.333(D) \quad \text{式(5)}$$

ここで、Dは、サンプル中の遅延ユニット636および646による遅延量である。例え

50

ば、サンプリングレートは48kHzである。別の実装方法としては、ローパスフィルタがあり、コーナー周波数は5000~10000Hzの範囲で選択され、Qは0.5~1.0の範囲で選択される。さらに、アンプ634は、対応するゲイン係数 $G_{L,IN}$ によって抽出部分を増幅し、遅延ユニット636は遅延機能Dに従ってアンプ634からの増幅出力を遅延させ、左側逆測キャンセル成分 $S_L$ を生成する。逆測推定器640には、フィルタ642、アンプ644、および遅延ユニット646が含まれている。このユニットは、反転されたインバンドチャンネル $T_{R,IN}$ で同様の操作を実行して、右側逆測キャンセル成分 $S_R$ を生成する。一例として、反逆測の推定630、640では、次の式に従って、左右の逆測キャンセル成分 $S_L$ 、 $S_R$ が生成される。

$$S_L = D [ G_{L, IN} * F [ T_{L, IN} ] ] \quad \text{式(6)}$$

$$S_R = D [ G_{R, IN} * F [ T_{R, IN} ] ] \quad \text{式(7)}$$

ここで、 $F [ ]$ はフィルタ関数、 $D [ ]$ は遅延関数である。

【0076】

クロストークキャンセルの設定は、スピーカのパラメータによって決定できる。例えば、2つのスピーカ280間のリスナーに対する角度に応じて、フィルタの中心周波数、遅延量、アンプゲイン、およびフィルタゲインを決定できる。いくつかの実施形態では、スピーカ角度間の値を使用して他の値を補間する。

【0077】

結合器650は、右側逆測キャンセル成分 $S_R$ と左側インバンドチャンネル $T_{L,IN}$ を結合して、左側インバンドクロストークチャンネル $U_L$ を生成する。また、結合器652は、左側逆測キャンセル成分 $S_L$ を右側インバンドチャンネル $T_{R,IN}$ に結合して、右側インバンドクロストークチャンネル $U_R$ を生成する。インアウトバンド結合器660は、左インバンドクロストークチャンネル $U_L$ とアウトバンドチャンネル $T_{L,OUT}$ を結合して、左出力チャンネル $O_L$ を生成する。また、右インバンドクロストークチャンネル $U_R$ とアウトバンドチャンネル $T_{R,OUT}$ を結合して、右出力チャンネル $O_R$ を生成する。

【0078】

したがって、左の出力チャンネル $O_L$ には逆測の音に関連してインバンドチャンネル $T_{R,IN}$ の一部の反転に対応する逆測キャンセル成分 $S_R$ が含まれる。また、右出力チャンネル $O_R$ には、逆測の音に関連して、インバンドチャンネル $T_{R,IN}$ の一部の逆に対応する左逆測キャンセル成分 $S_L$ が含まれる。この構成では、右耳に到達した右出力チャンネル $O_R$ に応じて右スピーカから出力した同側サウンド成分の波面は、左出力チャンネル $O_L$ に応じて左スピーカから出力された逆測サウンド成分の波面をキャンセルすることができる。同様に、左耳に到達した左出力チャンネル $O_L$ に応じて左スピーカから出力された同側サウンド成分の波面は、右出力チャンネル $O_R$ に応じて、右スピーカから出力された逆測サウンド成分の波面をキャンセルすることができる。これにより、空間的検出性を高めるために逆測音を低減することができる。

【0079】

(空間強調プロセッサの例)

図7は、いくつかの実施形態による空間強調プロセッサ700の概略ブロック図である。空間強調プロセッサ700は、クロストーク処理がヘッドフォンを使用するクライアントデバイス130のクロストークシミュレーションである空間強調プロセッサ120の一例である。空間強調プロセッサ700は、入力オーディオ信号 $X$ のクロストークシミュレーションを実行して、左ヘッドフォンの左出力チャンネル $O_L$ と右ヘッドフォンの右出力チャンネル $O_R$ を含む出力オーディオ信号 $O$ を生成する。左の入力チャンネル $X_L$ は空間混合プロセッサ110からの左の出力チャンネル142、右の入力チャンネル $X_R$ は空間混合プロセッサ110からの右の出力チャンネル144となることがある。左強調チャンネルの $O_L$ は左強調チャンネル146に対応し、右強調チャンネルの $O_R$ は右強調チャンネル148に対応する。

【0080】

空間強調プロセッサ700には、サブバンド空間プロセッサ310、クロストーク補正プロセッサ720、クロストークシミュレーションプロセッサ780、および結合器760

10

20

30

40

50

が含まれる。クロストーク補正プロセッサ720はチャンネル $X_L$ および $X_R$ を受け取り、クロストークシミュレーションプロセッサ780により生成された左クロストークシミュレーション信号 $W_L$ と右クロストークシミュレーション信号 $W_R$ を含むクロストークシミュレーション信号 $W$ と、サブバンド空間プロセッサ310により生成された左強調チャンネル $E_L$ と右強調チャンネル $E_R$ を含む強調信号 $E$ の結合におけるアーチファクト補正処理を実行する。クロストーク補正プロセッサ720は、左クロストーク補正チャンネル $Z_L$ および右クロストーク補正チャンネル $Z_R$ を含むクロストーク補正信号 $Z$ を生成する。クロストークシミュレーションプロセッサ780は、左クロストークシミュレーションチャンネル $W_L$ および右クロストークシミュレーションチャンネル $W_R$ を生成する。サブバンド空間プロセッサ310は、左強調チャンネル $E_L$ と右強調チャンネル $E_R$ を生成する。クロストークシミュレーションプロセッサ780の詳細については、以下の図8で説明される。

10

#### 【0081】

結合器760は、左強調チャンネル $E_L$ 、右強調チャンネル $E_R$ 、左クロストークシミュレーションチャンネル $W_L$ 、右クロストークシミュレーションチャンネル $W_R$ 、左クロストーク補正チャンネル $Z_L$ 、および右クロストーク補正チャンネル $Z_R$ を受信する。結合器760は、左強調チャンネル $E_L$ 、右クロストークシミュレーションチャンネル $W_R$ 、および左クロストーク補正チャンネル $Z_L$ を結合し、左出力チャンネル $O_L$ を生成する。結合器760は、左強調チャンネル $E_L$ 、右クロストークシミュレーションチャンネル $W_R$ 、および左クロストーク補正チャンネル $Z_L$ を結合することで、右出力チャンネル $O_R$ を生成する。

20

#### 【0082】

(クロストークシミュレーションプロセッサの例)

図8はいくつかの実施形態におけるクロストークシミュレーションプロセッサ780の概略ブロック図である。クロストークシミュレーションプロセッサ780は、ステレオヘッドフォンに出力する逆測サウンド成分を生成し、ヘッドフォンでスピーカーのようなリスニング体験を提供する。左入力チャンネル $X_L$ は、空間混合プロセッサ110からの左出力チャンネル142であり得、右入力チャンネル $X_R$ は、空間混合プロセッサ110からの右出力チャンネル144であり得る。

#### 【0083】

クロストークシミュレーションプロセッサ780には、左入力チャンネル $X_L$ を処理するために左ヘッドシャドウローパスフィルタ802、左ヘッドシャドウハイパスフィルタ824、左クロストーク遅延804、左ヘッドシャドウゲイン810が含まれている。クロストークシミュレーションプロセッサ780には、右入力チャンネル $X_R$ を処理するために右ヘッドシャドウローパスフィルタ806、右ヘッドシャドウハイパスフィルタ826、右クロストーク遅延808、右ヘッドシャドウゲイン812が含まれている。左ヘッドシャドウローパスフィルタ802および左ヘッドシャドウハイパスフィルタ824は、リスナーのヘッドを通過した後の信号の周波数応答をモデル化する左入力チャンネル $X_L$ に変調を適用する。左ヘッドシャドウハイパスフィルタ824の出力は、時間遅延を適用する左クロストーク遅延804に提供される。遅延時間は、同側サウンド成分に対して逆測サウンド成分によって横断される耳の距離を表す。左ヘッドシャドウゲイン810は、左クロストーク遅延804の出力にゲインを適用して、右左シミュレーションチャンネル $W_L$ を生成する。

30

40

#### 【0084】

同様に、右入力チャンネル $X_R$ の場合、右ヘッドシャドウローパスフィルタ806および右ヘッドシャドウハイパスフィルタ826は、リスナーのヘッドの周波数応答をモデル化する右入力チャンネル $X_R$ に変調を適する。右ヘッドシャドウハイパスフィルタ826の出力は、時間遅延を適用する右クロストーク遅延808に提供される。右ヘッドシャドウゲイン812は、右クロストーク遅延808の出力にゲインを適用して、右クロストークシミュレーションチャンネル $W_R$ を生成する。

#### 【0085】

左右の各チャンネルのヘッドシャドウローパスフィルタ、ヘッドシャドウハイパスフィルタ

50

、クロストーク遅延、およびヘッドシャドウゲインの適用は、異なる順序で実行できる。

【0086】

(カンファレンス処理の例)

図9は、いくつかの実施形態におけるクライアントデバイス間でカンファレンスを行うためのプロセス900のフローチャートである。プロセス900は、クライアントデバイス130からの入力オーディオストリームを含む、空間強調ステレオフォニックオーディオストリームをクライアントデバイス130に提供するために実行される。プロセス900は、いくつかのステップまたは追加のステップを含むことができ、ステップは異なる順序で実行することができる。プロセス900は、いくつかの実施形態による、オーディオ処理システム100の図をそれぞれが含む図9および10を参照して説明される。

10

【0087】

サーバー102(例えば、空間混合プロセッサ110を有する)は、複数のクライアントデバイス130から905の入力オーディオストリームを受信する。例えば、オーディオストリームには、モノラルオーディオストリームが含まれ得る。別の例では、1つまたは複数の入力オーディオストリームに複数のオーディオチャンネル(ステレオ、5.1チャンネルサラウンドサウンド、1次オーダーのアビソニックなど)が含まれる場合がある。入力オーディオストリームは、ネットワーク経由で複数のクライアントデバイス130から送信できる。いくつかの実施形態では、各クライアントデバイス130は、入力オーディオストリームに関連してメタデータを送信する。メタデータには、クライアントデバイスを識別するデバイス識別データ、他のクライアントデバイスの他のオーディオストリームの必要な空間的位置を定義する配置データが含まれる場合がある。

20

【0088】

サーバー102(例えば空間混合プロセッサ110を有する)は、クライアントデバイス130では、サウンドフィールド内の他のクライアントデバイス130の空間的な位置を定義する配置データ910を決定する。配置データはさまざまな方法で定義できる。例えば、配置データは、クライアントデバイスからメタデータとして受信するか、サーバー102によって事前定義されている場合がある。いくつかの実施形態では、ユーザーパントロールがない場合、クライアントデバイス130が自動的にパンされる。ユーザーは、好みに応じてオーディオをパンすることもできる。

【0089】

サーバー102(例えば、空間混合プロセッサ110を有する)は、入力オーディオストリームを配置データに従って混合およびパンすることで、915A混合ストリームを生成する。混合ストリームは、左混合チャンネルと右混合チャンネルを含むステレオオーディオ信号であり得る。例えば、空間混合プロセッサ110は、サウンドフィールド内のオーディオストリーム140の空間位置を決定することができ、オーディオストリーム140に関連付けられた空間位置に応じて、オーディオストリーム140から左チャンネルと右チャンネルを生成し得る。左右のチャンネルは、他のクライアントデバイス130のオーディオストリーム140ごとに生成される。複数のオーディオストリーム140の左側のチャンネルは左混合チャンネル142に結合され、右側のチャンネル140は右混合チャンネル144に結合される。

30

40

【0090】

図10を参照すると、複数のクライアントデバイス130A、130b、130C、および140Dがカンファレンスに参加し、サーバー110に接続されている。クライアントデバイス130Aの場合、クライアントデバイス130Aの周囲にサウンドフィールド1002が定義される。配置データに従って、他のクライアントデバイス130b、130C、および130dがサウンドフィールド1002に配置される。サウンドフィールド1002には、間隔をおいたクライアントデバイス130の均一な空間分布が示されているが、空間分布は必ずしも均一ではない。例えば、クライアントデバイス130は、ユーザーロール、エンティティ/会社名、グループメンバーシップ、カンファレンス室の場所、地理的な場所、IPアドレスなどを定義する配置データに基づいて、サウンドフィールド

50

1002内に分散され得る。

【0091】

サーバー102（空間強調プロセッサ120を搭載）は、サブバンド空間処理とクロストーク処理を混合ストリームに適用することで、空間的に強調されたストリーム920を生成する。空間的に強調されたストリームは、左強調チャンネルと右強調チャンネルを含むステレオオーディオ信号の場合がある。空間強調プロセッサ120は、さまざまなタイプの空間強調を実行できる。サブバンド空間処理の場合、空間強調プロセッサ120は、左混合チャンネルのミッドおよびサイドサブバンド成分と混合ストリームの右混合チャンネルをゲイン調整することによって、左チャンネルと右チャンネルを生成する。クロストーク処理には、クライアントデバイス130がスピーカーとヘッドフォンのどちらを使用しているかによ

10

【0092】

いくつかの実施形態では、クライアントデバイス130の混合ストリームに適用される空間強調処理は、クライアントデバイス130からのデバイス情報に基づく。デバイス情報には、スピーカーパラメータ（サイズ、周波数応答、位置、方向など）や、フィルタ値、ゲイン値、時間遅延値、空間強調効果の強度などの空間強調プロセッサの動作パラメータなど、空間強調プロセッサ120の成分のパラメータが含まれる場合がある。いくつかの実施形態では空間強調プロセッサ120の動作パラメータは、スピーカーパラメータから

20

【0093】

いくつかの実施形態では、サーバー102は、各クライアントデバイスからデバイスの説明を受信し、デバイスの説明に基づいて空間強調プロセッサ120のパラメータを決定する。デバイスの説明には、ラップトップ、サウンドバー、ヘッドフォン、スピーカー、デスクトップスピーカーなどのクライアントデバイスのタイプが表示される場合がある。さまざまなタイプのクライアントデバイスが異なるパラメータに関連付けられ、パラメータ

30

【0094】

サーバー102は、クライアントデバイス130に空間強調ストリーム925を提供する。例えば、サーバー102は、左強調チャンネルと右強調チャンネルを含む空間強調ストリームをネットワーク経由でクライアントデバイス130に送信する。左強調チャンネルは左スピーカー152に、右強調チャンネルは右スピーカー154に提供され、空間強調ストリームに応じてサウンドを再現する。

【0095】

いくつかの実施形態では、サーバー102（空間混合プロセッサ110など）は、空間強調ストリーム（または混合ストリーム）に関連して、各クライアントデバイス130にパニングメタデータを提供する。パニングメタデータには、オーディオストリームの参加者を識別する情報が含まれ得る。参加者は、ユーザーロール、エンティティ/会社名、グループメンバーシップ、部屋の場所、地理的な場所、IPアドレスによって識別できる。

40

【0096】

いくつかの実施形態では、サーバー102は、クライアントデバイス130ごとに空間強調処理を行うために、クライアントデバイス130に混合ストリームを提供する。クライアントデバイス130には、混合ストリームを使用して左右の強調チャンネルを生成する空間強調プロセッサ120を含む。図11を参照すると、各クライアントデバイス130A～130dには、空間強調プロセッサ120が含まれており、空間混合プロセッサ110からの混合ストリームを処理する。クライアントデバイス120の空間強調プロセッサ1

50

20で使用される空間拡張プロセッサ120の成分のパラメータやデバイス情報はサーバー102に送信する必要はない。前述したように、さまざまなタイプの空間強調処理が、さまざまなタイプのクライアントデバイス130およびスピーカ構成に使用され得る。

【0097】

いくつかの実施形態では、サーバー102はオーディオ信号140の混合されていないストリームをクライアントデバイス130に提供し、クライアントデバイス130は空間強調によりステレオミックスダウンされる。図12を参照すると、各クライアントデバイス130A~130dには、ミックスダウンを実行して混合ストリームを生成する空間混合プロセッサ110と、空間混合プロセッサ110からの混合ストリームを処理する空間強調プロセッサ120が含まれている。サーバー102は、クライアントデバイス130Aから130dへの接続と、他のクライアントデバイス130から各クライアントデバイス130への入力信号140の提供を仲介する。いくつかの実施形態では、ピアツーピア構成など、各クライアントデバイス130はサーバー102を使用せずに相互に接続される。ここでは、各クライアントデバイス130が他のクライアントデバイス130にオーディオストリーム140を提供する。

10

【0098】

ステップ910-925は、カンファレンスに参加しているクライアントデバイス130ごとに並行して実行することができる。クライアントデバイス130ごとに、1つまたは複数の他のクライアントデバイス130の入力オーディオストリームから異なる混合ストリームを生成できる。サーバー102には、クライアントデバイスごとに空間強調プロセッサ120の複数のインスタンスが含まれ、それぞれ異なる混合ストリームを処理する。さらに、ステップ910-925を繰り返して、カンファレンスに参加している各クライアントデバイス130にオーディオストリームの入出力を提供できる。

20

【0099】

(コンピュータの例)

図13は、一実施形態によるコンピュータ1300の概略ブロック図である。コンピュータ1300は、オーディオシステムを実装する回路の一例である。図は、チップセット1304に1つ以上のプロセッサ1302を結合したものである。チップセット1304は、メモリコントローラハブ1320とI/Oコントローラハブ1322を搭載している。メモリ1306とグラフィックスアダプタ1312はメモリコントローラハブ1320に接続され、ディスプレイデバイス1318はグラフィックスアダプタ1312に接続されている。ストレージデバイス1308、キーボード1310、ポインティングデバイス1314、およびネットワークアダプタ1316は、I/Oコントローラハブ1322に接続されている。コンピュータ1300には、さまざまなタイプの入出力デバイスが含まれ得る。コンピュータ1300の他の実施形態には、異なるアーキテクチャがある。例えば、いくつかの実施形態によると、メモリ1306はプロセッサ1302に直接結合されている。

30

【0100】

ストレージデバイス1308には、ハードドライブ、コンパクトディスク読み取り専用メモリ(CD-ROM)、DVD、ソリッドステートメモリデバイスなど、一時的にコンピュータで読み取り可能な1つ以上のストレージメディアが含まれている。メモリ1306は、プログラムコード(1つ以上の命令で構成される)とプロセッサ1302が使用するデータを保持する。プログラムコードは、図1から図11で説明されている処理面に対応し得る。

40

【0101】

ポインティングデバイス1314をキーボード1310と組み合わせて使用し、コンピュータシステム1300にデータを入力する。グラフィックスアダプタ1312は、ディスプレイデバイス1318に画像やその他の情報を表示する。いくつかの実施形態によると、ディスプレイ装置1318には、ユーザーの入力と選択を受信するためのタッチスクリーン機能が組み込まれている。ネットワークアダプタ1316は、コンピュータシステム

50

1300をネットワークに接続する。コンピュータ1300のいくつかの実施形態は、図13に示すものとは異なるものや、その他の成分を備えているものもある。

【0102】

本明細書で説明された範囲から逸脱することなく、当業者には明らかなはずの様々な修正形態、変更形態及び変形形態が、本明細書で開示された方法及び装置の配置、動作及び詳細において作製され得る。

10

20

30

40

50



【 図 面 】

【 図 1 】

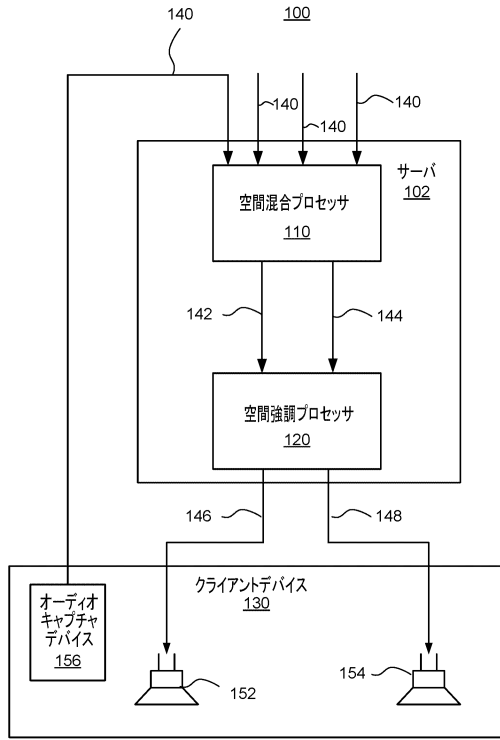


FIG. 1

【 図 2 】

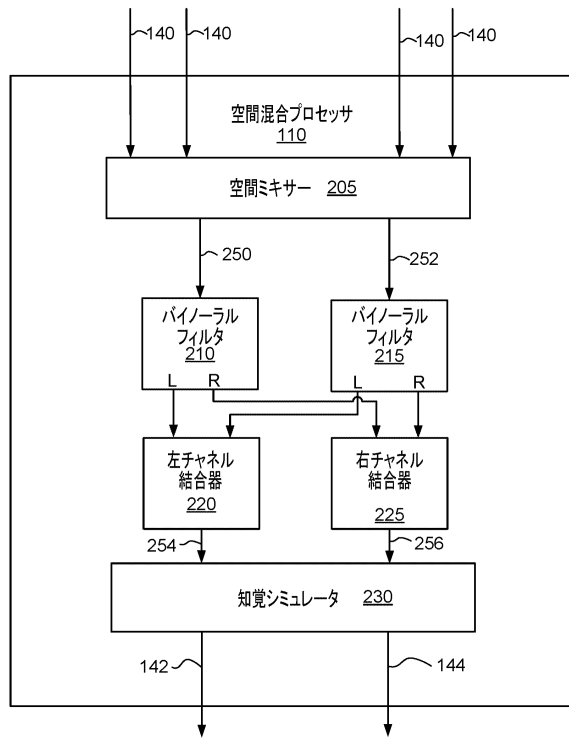


FIG. 2

【 図 3 】

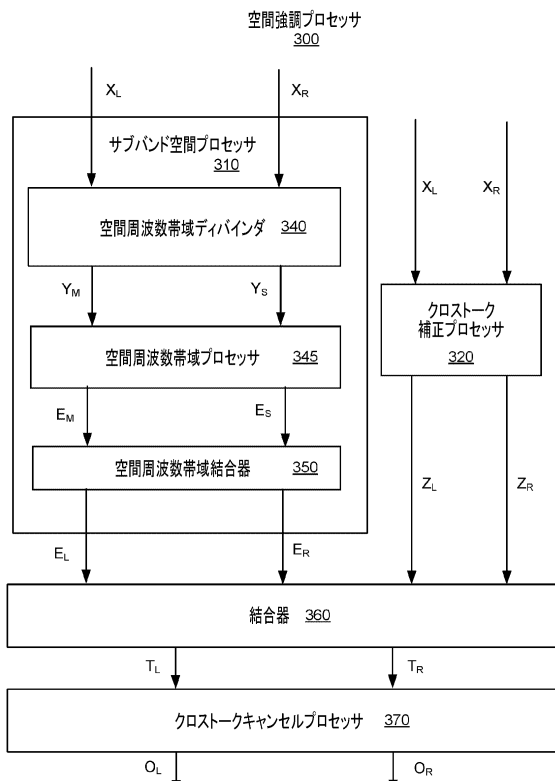


FIG. 3

【 図 4 】

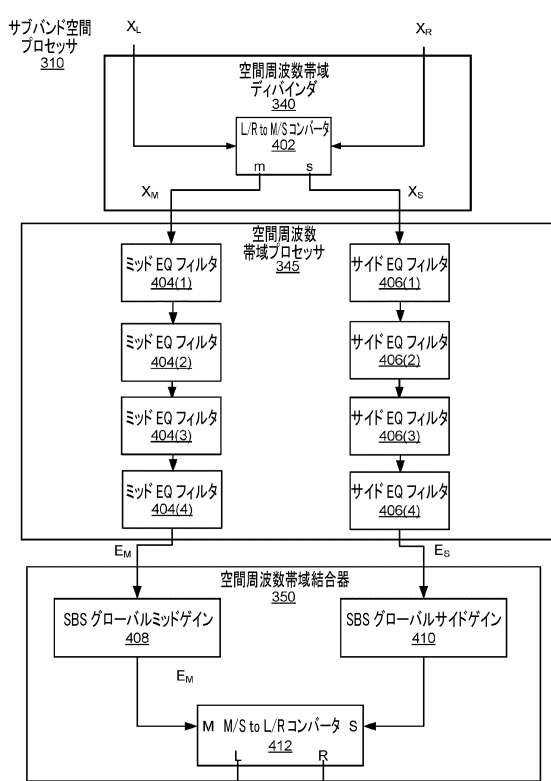


FIG. 4

10

20

30

40

50

【 図 5 】

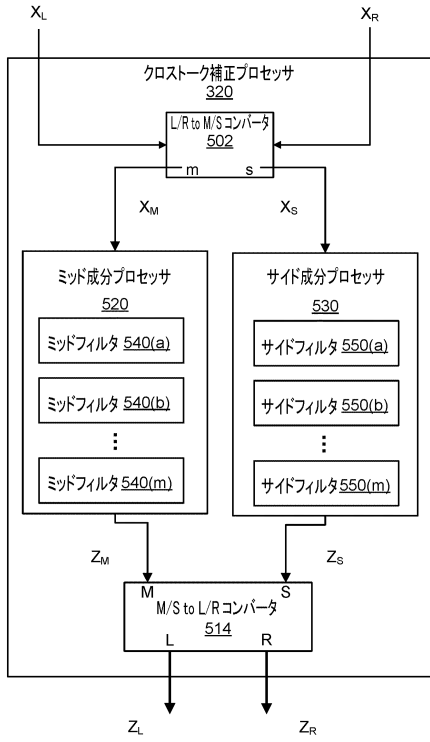


FIG. 5

【 図 6 】

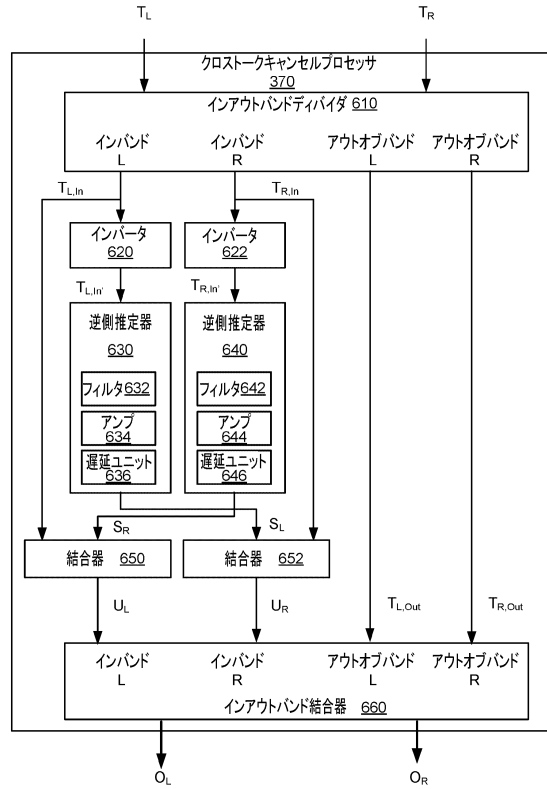


FIG. 6

【 図 7 】

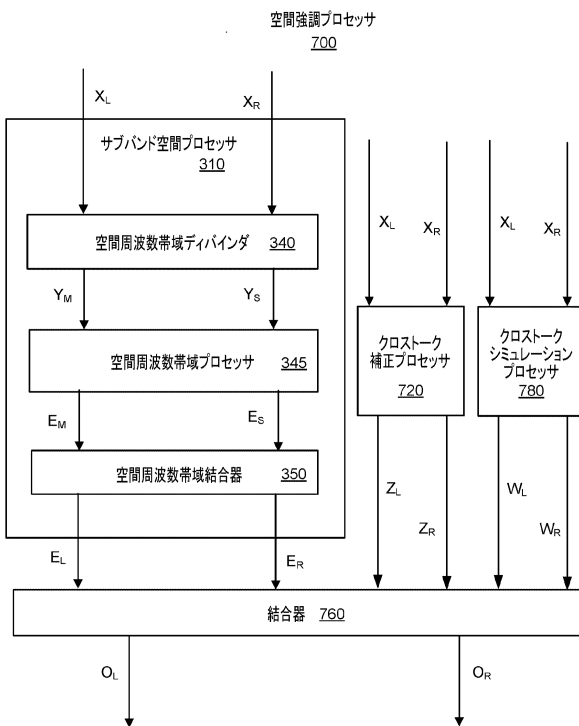


FIG. 7

【 図 8 】

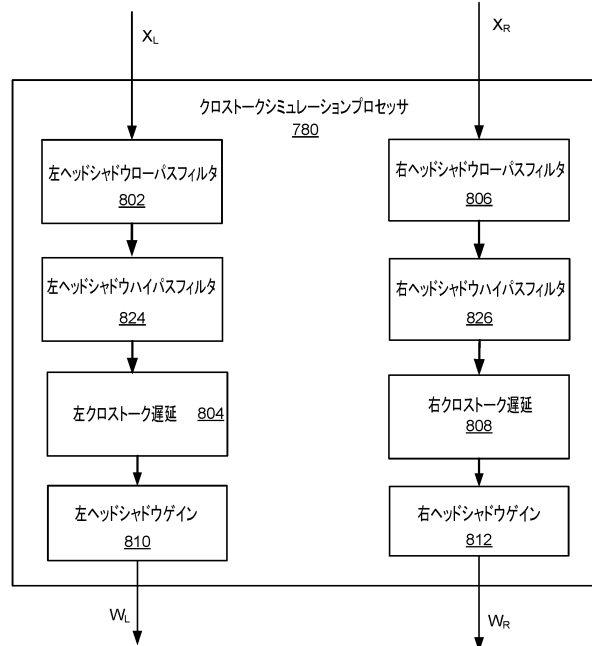


FIG. 8

10

20

30

40

50

【 図 9 】

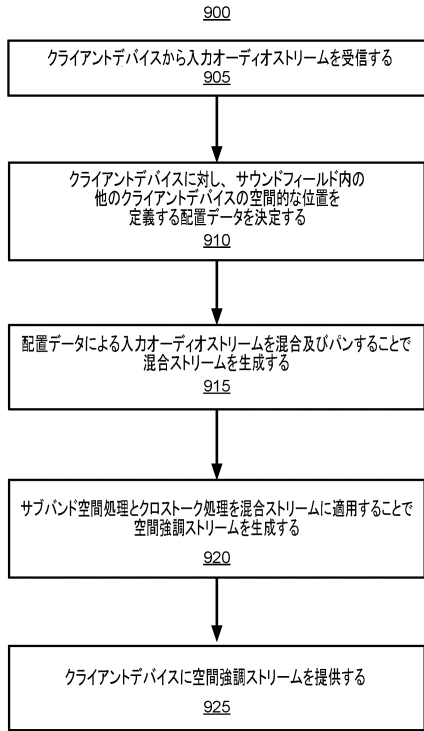


FIG. 9

【 図 10 】

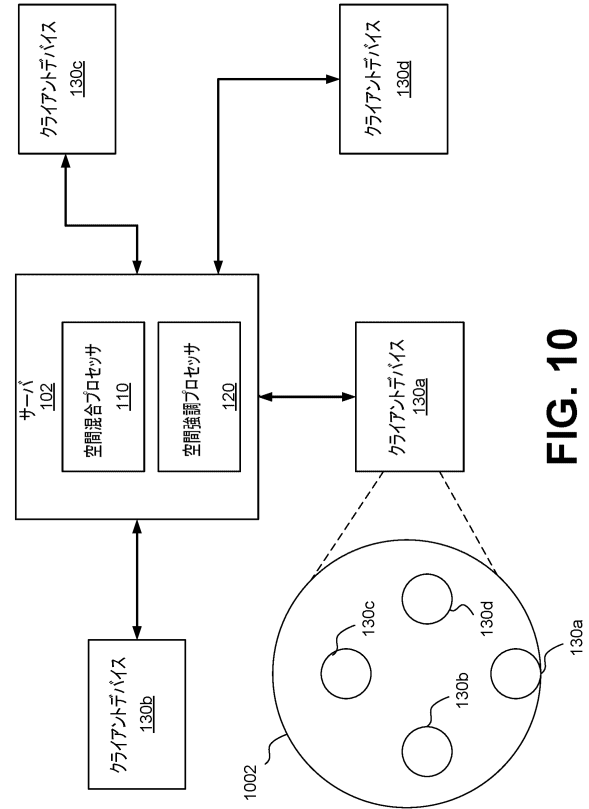


FIG. 10

【 図 11 】

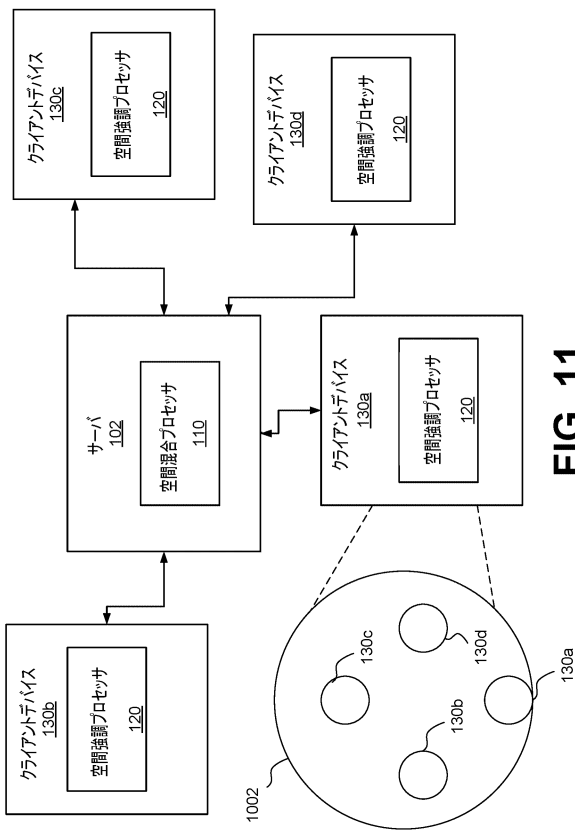


FIG. 11

【 図 12 】

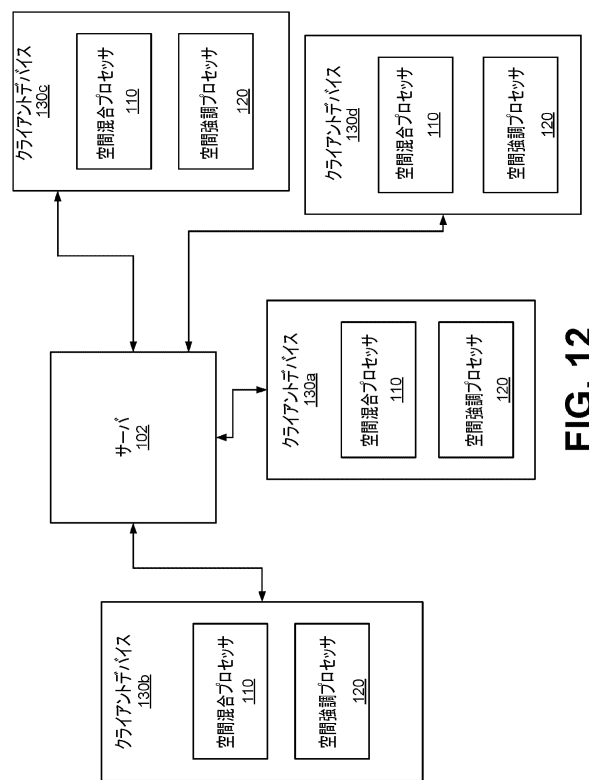


FIG. 12

10

20

30

40

50

【 図 1 3 】

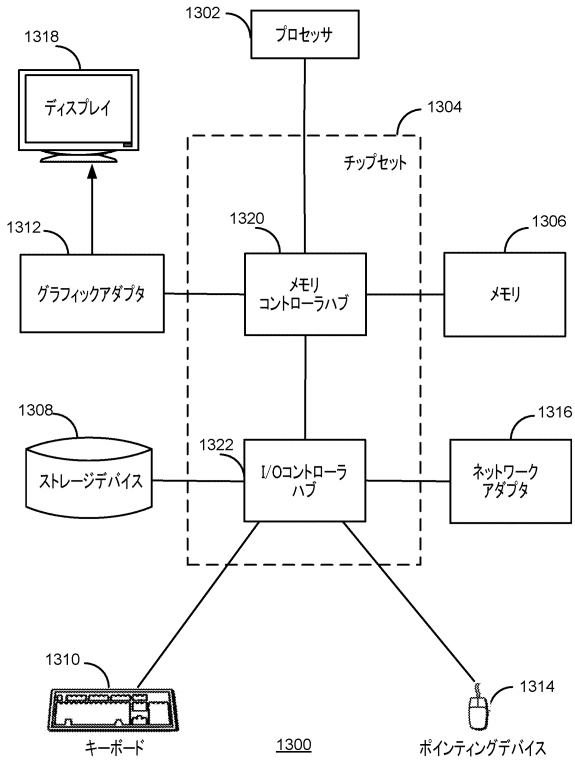


FIG. 13

10

20

30

40

50

---

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

早期審査対象出願

審査官 富澤 直樹

(56)参考文献 特開2006-025439(JP,A)  
特開平10-304498(JP,A)  
米国特許第7974418(US,B1)  
米国特許出願公開第2006/0008091(US,A1)  
米国特許出願公開第2014/0016793(US,A1)  
米国特許出願公開第2017/0230777(US,A1)  
米国特許出願公開第2017/0208411(US,A1)

(58)調査した分野 (Int.Cl., DB名)

H04S 1/00-7/00

G10L 19/00-19/26