

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7589226号
(P7589226)

(45)発行日 令和6年11月25日(2024.11.25)

(24)登録日 令和6年11月15日(2024.11.15)

(51)国際特許分類		F I			
H 0 4 R	3/00 (2006.01)	H 0 4 R	3/00	3 1 0	
H 0 4 S	7/00 (2006.01)	H 0 4 S	7/00	3 3 0	
G 1 0 K	15/02 (2006.01)	G 1 0 K	15/02		
G 0 6 F	3/16 (2006.01)	G 0 6 F	3/16	6 1 0	
G 0 6 F	3/01 (2006.01)	G 0 6 F	3/01	5 1 0	
請求項の数 21 (全151頁)					
(21)出願番号 特願2022-506086(P2022-506086)			(73)特許権者 507236292		
(86)(22)出願日 令和2年7月27日(2020.7.27)			ドルビー ラボラトリーズ ライセンシン		
(65)公表番号 特表2022-542388(P2022-542388			グ コーポレイション		
A)			アメリカ合衆国 9 4 1 0 3 カリフォル		
(43)公表日 令和4年10月3日(2022.10.3)			ニア州 サンフランシスコ マーケット		
(86)国際出願番号 PCT/US2020/043769			ストリート 1 2 7 5		
(87)国際公開番号 WO2021/021752			(73)特許権者 510185767		
(87)国際公開日 令和3年2月4日(2021.2.4)			ドルビー・インターナショナル・アーベ		
審査請求日 令和5年5月8日(2023.5.8)			ー		
(31)優先権主張番号 P201930702			アイルランド ディー 0 2 ヴィーケイ 6		
(32)優先日 令和1年7月30日(2019.7.30)			0 ダブリン グランド カナル ドックラ		
(33)優先権主張国・地域又は機関			ンズ サー ジョン ロジャーソンズ キー		
スペイン(ES)			7 7 ブロック シー		
(31)優先権主張番号 62/880,114			(74)代理人 100107766		
(32)優先日 令和1年7月30日(2019.7.30)			弁理士 伊東 忠重		
最終頁に続く			最終頁に続く		
(54)【発明の名称】 オーディオ装置の協調					

(57)【特許請求の範囲】

【請求項1】

複数のオーディオ装置を有するオーディオ環境のためのオーディオ・セッション管理方法であって、前記オーディオ・セッション管理方法は：

第1のアプリケーションを実装する第1の装置から、オーディオ・セッション・マネージャを実装する装置によって、第1のオーディオ・セッションのための第1のルートを開始するための第1のルート開始要求を受領する段階であって、前記第1のルート開始要求は第1のオーディオソースおよび第1のオーディオ環境宛先を示し、前記第1のオーディオ環境宛先は、前記オーディオ環境における少なくとも第1の人物に対応し、前記第1のオーディオ環境宛先は、オーディオ装置を示さない、段階と；

前記オーディオ・セッション・マネージャを実装している装置によって、前記第1のルート開始要求に対応する第1のルートを確立する段階であって、前記第1のルートを確立することは：

前記オーディオ環境における少なくとも前記第1の人物の第1の位置を判別し；

前記第1のオーディオ・セッションの第1のステージについて、前記オーディオ環境の前記複数のオーディオ装置のうちの少なくとも1つのオーディオ装置を判別し；

前記第1のオーディオ・セッションを開始またはスケジューリングすることを含む、段階とを含む、

前記第1のルート開始要求は、前記オーディオ環境の少なくとも第1のエリアを第1のルート・ソースとして示す、または前記第1のルート開始要求は、少なくとも第1のサービス

を前記第1のルート・ソースとして示す、
オーディオ・セッション管理方法。

【請求項2】

前記第1のルート開始要求は、第1のオーディオ・セッション優先度を含む、請求項1に記載のオーディオ・セッション管理方法。

【請求項3】

請求項1または2に記載のオーディオ・セッション管理方法であって、前記第1のルート開始要求は第1の接続性モードを含む、オーディオ・セッション管理方法。

【請求項4】

前記第1の接続性モードが、同期的接続性モード、トランザクション接続性モード、またはスケジュールされた接続性モードである、請求項3に記載のオーディオ・セッション管理方法。

【請求項5】

請求項1ないし4のうちいずれか一項に記載のオーディオ・セッション管理方法であって、前記第1のルート開始要求は、少なくとも前記第1の人物から確認応答が要求されるかどうかの指示を含む、オーディオ・セッション管理方法。

【請求項6】

請求項1ないし5のうちいずれか一項に記載のオーディオ・セッション管理方法であって、前記第1のルート開始要求は、第1のオーディオ・セッション目標を含む、オーディオ・セッション管理方法。

【請求項7】

前記第1のオーディオ・セッション目標は、了解度、オーディオ品質、空間的忠実度、可聴性、非可聴性またはプライバシーのうちの一つまたは複数を含む、請求項6に記載のオーディオ・セッション管理方法。

【請求項8】

請求項1ないし7のうちいずれか一項に記載のオーディオ・セッション管理方法であって、前記第1のルートのための第1の永続的固有オーディオ・セッション識別子を決定し、前記第1の永続的固有オーディオ・セッション識別子を前記第1の装置に送信する段階をさらに含む、オーディオ・セッション管理方法。

【請求項9】

請求項1ないし8のうちいずれか一項に記載のオーディオ・セッション管理方法であって、前記第1のルートを確認することは、前記環境内の少なくとも1つの装置に、前記第1のルートに対応する、前記第1のオーディオ・セッションについての少なくとも第1のメディア・ストリームを確認させることを含む、前記第1のメディア・ストリームは第1のオーディオ信号を含む、オーディオ・セッション管理方法。

【請求項10】

前記第1のオーディオ信号が第1のレンダリングされたオーディオ信号にレンダリングされるようにすることをさらに含む、請求項9に記載のオーディオ・セッション管理方法。

【請求項11】

請求項10に記載のオーディオ・セッション管理方法であって、さらに：前記オーディオ・セッションの前記第1のステージについての前記第1の人物の第1の配向を決定する段階をさらに含む、前記第1のオーディオ信号が第1のレンダリングされたオーディオ信号にレンダリングされるようにすることは：

前記第1の人物の前記第1の位置および前記第1の配向に対応する第1の基準空間モードを決定し；

前記第1の基準空間モードに対応する前記オーディオ環境におけるラウドスピーカーの第1の相対的アクティブ化を決定することを含む、
オーディオ・セッション管理方法。

【請求項12】

請求項11に記載のオーディオ・セッション管理方法であって、さらに：

前記第1のオーディオ・セッションの第2のステージについての前記第1の人物の第2の位置または第2の配向の少なくとも一方を決定し；

前記第2の位置または前記第2の配向のうちの少なくとも一方に対応する第2の基準空間モードを決定し；

前記第2の基準空間モードに対応する前記オーディオ環境におけるラウドスピーカーの第2相対的アクティブ化を決定することを含む、オーディオ・セッション管理方法。

【請求項13】

請求項10ないし12のうちいずれか一項に記載のオーディオ・セッション管理方法であって、さらに：

第2のアプリケーションを実装する第2の装置から、前記オーディオ・セッション・マネージャを実装する前記装置によって、第2のオーディオ・セッションのための第2のルートを開始するための第2のルート開始要求を受領する段階であって、前記第2のルート開始要求は第2のオーディオソースおよび第2のオーディオ環境宛先を示し、前記第2のオーディオ環境宛先は、前記オーディオ環境における少なくとも第2の人物に対応し、前記第2のオーディオ環境宛先は、オーディオ装置を示さない、段階と；

前記オーディオ・セッション・マネージャを実装している前記装置によって、前記第2のルート開始要求に対応する第2のルートを確認する段階であって、前記第2のルートを確認することは：

前記オーディオ環境における少なくとも前記第2の人物の第1の位置を判別し；

前記第2のオーディオ・セッションの第1のステージについて、前記オーディオ環境の前記複数のオーディオ装置のうちの少なくとも1つのオーディオ装置を判別し；

前記第2のオーディオ・セッションを開始することを含む、段階とを含む、オーディオ・セッション管理方法。

【請求項14】

前記第2のルートを確認することは、前記第2のルートに対応する少なくとも第2のメディア・ストリームを確認することを含み、前記第2のメディア・ストリームは、第2のオーディオ信号を含む、請求項13に記載のオーディオ・セッション管理方法。

【請求項15】

前記第2のオーディオ信号が第2のレンダリングされたオーディオ信号にレンダリングされるようにすることをさらに含む、請求項14に記載のオーディオ・セッション管理方法。

【請求項16】

少なくとも部分的には前記第2のオーディオ信号、前記第2のレンダリングされたオーディオ信号またはその特性のうちの少なくとも1つに基づいて前記第1のオーディオ信号についてのレンダリング・プロセスを修正して、修正された第1のレンダリングされたオーディオ信号を生成することをさらに含む、請求項15に記載のオーディオ・セッション管理方法。

【請求項17】

前記第1のオーディオ信号についての前記レンダリング・プロセスを修正することは：

第1のオーディオ信号の前記レンダリングを、前記第2のレンダリングされたオーディオ信号のレンダリング位置から離れるようにワープさせること、または

前記第2のオーディオ信号または前記第2のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスに応答して、前記第1のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスを修正することを含む、

請求項16に記載のオーディオ・セッション管理方法。

【請求項18】

前記第1のルート開始要求は、前記オーディオ環境の少なくとも第1のエリアを第1のルート宛先としてさらに示す、請求項1ないし17のうちいずれか一項に記載のオーディオ・セッション管理方法。

【請求項19】

請求項 1 ないし 18 のうちいずれか一項に記載の方法を実行するように構成された装置。
【請求項 20】

請求項 1 ないし 18 のうちいずれか一項に記載の方法を実行するように構成されたシステム。

【請求項 21】

請求項 1 ないし 18 のうちいずれか一項に記載の方法を実行するよう一つまたは複数の装置を制御するためのコンピュータ・プログラムを記憶している一つまたは複数の非一時的媒体。

【発明の詳細な説明】

【技術分野】

10

【0001】

関連出願への相互参照

本願は、米国仮特許出願第62/992,068号（2020年3月19日出願）、米国仮特許出願第62/949,998号（2019年12月18日出願）、欧州特許出願第19217580.0号（2019年12月18日出願）、スペイン特許出願第P201930702号（2019年7月30日出願）、米国仮特許出願第62/971,421号（2020年2月7日出願）、米国仮特許出願第62/705,410号（2020年6月25日出願）、米国仮特許出願第62/880,114号（2019年7月30日出願）、米国仮特許出願第62/705,351号（2020年6月23日出願）、米国仮特許出願第62/880,115号（2019年7月30日出願）、米国仮特許出願第62/705,143号（2020年6月12日出願）、米国仮特許出願第62/880,118号（2019年7月30日）、米国仮特許出願第16/929,215号（2020年7月15日出願）、米国仮特許出願第62/705,883号（2020年7月20日出願）、米国仮特許出願第62/880,121号（2019年7月30日出願）、米国仮特許出願第62/705,884号（2020年7月20日出願）の優先権を主張するものである。これらの各出願は、ここにその全体が参照により組み込まれる。

20

技術分野

本開示は、スマート・オーディオ装置を含んでいてもよいオーディオ装置を協調させ（オーケストレーションし）、実装するためのシステムおよび方法に関する。

【背景技術】

【0002】

スマート・オーディオ装置を含むが、それに限定されないオーディオ装置は、広く配備され、多くの家庭で一般的な機能となりつつある。オーディオ装置を制御するための既存のシステムおよび方法は利点を提供するが、改善されたシステムおよび方法が望ましいであろう。

30

【0003】

記法および名称

特許請求の範囲を含め、本開示全体を通じて、「スピーカー」および「ラウドスピーカー」は、単一のスピーカー・フィールドによって駆動される任意の放音トランスデューサ（またはトランスデューサの集合）を表すために同義で使用される。ヘッドフォンの典型的なセットは、2つのスピーカーを含む。スピーカーは、単一の共通スピーカー・フィールドまたは複数のスピーカー・フィールドによって駆動されうる複数のトランスデューサ（たとえば、ウーファーおよびツイーター）を含むように実装されうる。いくつかの例では、スピーカー・フィールドは、異なるトランスデューサに結合された異なる回路分枝において異なる処理を受けることができる。

40

特許請求の範囲を含め、本開示全体を通じて、信号またはデータ「に対して」動作を実行するという表現（たとえば、信号またはデータのフィルタリング、スケーリング、変換、または利得の適用）は、広い意味で使用され、信号またはデータに対して該動作を直接実行すること、または信号またはデータの処理されたバージョンに対して（たとえば、該動作の実行前に予備的なフィルタリングまたは前処理を受けた該信号のバージョンに対して）該動作を実行することを示す。

【0004】

50

特許請求の範囲を含む本開示全体を通じて、「システム」という表現は、広い意味で装置、システム、またはサブシステムを示すために使用される。たとえば、デコーダを実装するサブシステムがデコーダ・システムと称されることがあり、そのようなサブシステムを含むシステム（たとえば、複数の入力に応答してX個の出力信号を生成するシステムであって、そのサブシステムが入力のうちのM個を生成し、他のX-M個の入力は外部ソースから受領されるシステム）もデコーダ・システムと称することもできる。

【0005】

特許請求の範囲を含む本開示全体を通じて、用語「プロセッサ」は、データ（たとえば、オーディオ、ビデオまたは他の画像データ）に対して動作を実行するために、プログラム可能なまたは他の仕方で（たとえば、ソフトウェアまたはファームウェアを用いて）構成可能なシステムまたは装置を示すために広い意味で使用される。プロセッサの例は、フィールドプログラマブルゲートアレイ（または他の構成可能な集積回路またはチップセット）、オーディオまたは他の音声データに対してパイプライン処理を実行するようにプログラムされたおよび／または他の仕方で構成されたデジタル信号プロセッサ、プログラマブルな汎用プロセッサまたはコンピュータ、およびプログラマブルなマイクロプロセッサチップまたはチップセットを含む。

【0006】

特許請求の範囲を含む本開示全体を通じて、用語「結合する」または「結合され」は、直接的または間接的接続を意味するために使用される。よって、第1の装置が第2の装置に結合する場合、その接続は、直接接続を通じて、または他の装置および接続を介した間接接続を通じてでありうる。

【0007】

本明細書で使用される場所では、「スマート装置」とは、Bluetooth、Zigbee、近接場通信、Wi-Fi、光忠実度（Li-Fi）、3G、4G、5Gなどのさまざまな無線プロトコルを介して、一つまたは複数の他の装置（またはネットワーク）と通信するように一般的に構成された電子装置であって、ある程度対話的におよび／または自律的に動作することができるものである。スマート装置のいくつかの顕著なタイプは、スマートフォン、スマートカー、スマートサーモスタット、スマートドアベル、スマートロック、スマート冷蔵庫、タブレットとタブレット、スマートウォッチ、スマートバンド、スマートキーチェーン、スマート・オーディオ装置である。「スマート装置」という用語は、人工知能のようなユビキタスコンピューティングのある特性を示す装置を指すこともある。

【0008】

本明細書で使用される場所では、「スマート・オーディオ装置」という表現は、単一目的のオーディオ装置または多目的のオーディオ装置（たとえば、バーチャル・アシスタント機能の少なくともいくつかの側面を実装するオーディオ装置）のいずれかであるスマート装置を示す。単一目的のオーディオ装置は、少なくとも1つのマイクロフォン（および、任意的には少なくとも1つのスピーカーおよび／または少なくとも1つのカメラを含むかまたはそれに結合される）を含むかまたはそれに結合される装置（たとえば、テレビ（TV）または携帯電話）であって、大部分がまたは主として単一目的を達成するように設計されたものである。たとえば、テレビは、典型的には、番組素材からオーディオを再生することができる（また、再生することができると考えられる）が、ほとんどの場合、現代のテレビは、何らかのオペレーティングシステムを実行し、その上でテレビ視聴アプリケーションを含むアプリケーションがローカルに動作する。同様に、携帯電話のオーディオ入出力は多くのことをするかもしれないが、これらは電話上で動作するアプリケーションによってサービスされる。この意味で、スピーカーおよびマイクロフォンを有する単一目的のオーディオ装置は、しばしば、スピーカーおよびマイクロフォンを直接使用するローカル・アプリケーションおよび／またはサービスを実行するように構成される。いくつかの単一目的の諸オーディオ装置が、ゾーンまたはユーザー構成されたエリアにわたるオーディオの再生を達成するよう、グループ化するように構成されうる。

【0009】

多目的オーディオ装置の一つの一般的なタイプは、バーチャル・アシスタント機能の少なくともいくつかの側面を実装するオーディオ装置であるが、バーチャル・アシスタント機能の他の側面は、多目的オーディオ装置が通信するように構成されている一つまたは複数のサーバーのような一つまたは複数の他の装置によって実装されてもよい。そのような多目的オーディオ装置は、本明細書では「バーチャル・アシスタント」と称されることがある。バーチャル・アシスタントは、少なくとも1つのマイクロフォンを含むまたはそれに結合される（および、任意的には、少なくとも1つのスピーカーおよび／または少なくとも1つのカメラを含むまたはそれに結合される）装置（たとえば、スマートスピーカーまたは音声アシスタント統合装置）である。いくつかの例では、バーチャル・アシスタントは、ある意味ではクラウドで可能にされる、または他の仕方で完全にはバーチャル・アシスタント自体の中または上には実装されていないアプリケーションのために複数の装置（そのバーチャル・アシスタントとは異なる）を利用する能力を提供することができる。言い換えると、バーチャル・アシスタント機能の少なくともいくつかの側面、たとえば、音声認識機能は、バーチャル・アシスタントがインターネットなどのネットワークを介して通信することができる一つまたは複数のサーバーまたは他の装置によって（少なくとも部分的に）実装されてもよい。バーチャル・アシスタントどうしは、時に、たとえば離散的で、条件付きで定義された仕方で、協働することがある。たとえば、2以上のバーチャル・アシスタントは、そのうちの一つ、たとえば、ウェイクワードを聞いたことに最も自信があるバーチャル・アシスタントがそのワードに応答するという意味で、協働することができる。接続された諸バーチャル・アシスタントは、いくつかの実装では、一種のコンステレーションを形成することができ、これは、バーチャル・アシスタントであってもよい（またはそれを実装してもよい）1つのメイン・アプリケーションによって管理されてもよい。

【0010】

ここで、「ウェイクワード」とは、任意の音（たとえば、人間によって発声された単語、または何らかの他の音）を意味するために広義で使用され、スマート・オーディオ装置は、その音の検出（「聞く」）（スマート・オーディオ装置に含まれるかまたはそれに結合される少なくとも1つのマイクロフォン、または少なくとも1つの他のマイクロフォンを使用する）に応答して、覚醒するように構成される。この文脈において、「覚醒」とは、装置が音声コマンドを待つ（すなわち、音声コマンドがあるかどうか傾聴する）状態に入ることを表す。いくつかの事例では、本明細書において「ウェイクワード」と称されるものは、複数の単語、たとえば、フレーズを含んでいてもよい。

【0011】

ここで、「ウェイクワード検出器」という表現は、リアルタイムの音声（たとえば、発話）特徴とトレーニングされたモデルとの間の整列を連続的に探すよう構成された装置（または装置を構成するための命令を含むソフトウェア）を表す。典型的には、ウェイクワードが検出された確率が所定の閾値を超えることがウェイクワード検出器によって判別されるときは常に、ウェイクワード・イベントがトリガーされる。たとえば、閾値は、誤受理率と誤拒否率との間の合理的な妥協を与えるように調整された所定の閾値であってもよい。ウェイクワード・イベントに続いて、装置は、それがコマンドを待ち受け、受け取ったコマンドをより大きな、より計算集約的な認識器に渡す状態（「覚醒した」状態または「注視」状態と呼ばれてもよい）にはいってもよい。

【発明の概要】

【課題を解決するための手段】

【0012】

あるクラスの実施形態では、オーディオ装置（スマート・オーディオ装置を含むことができる）は、連続階層オーディオ・セッション・マネージャ（CHASM）を使用して協調させられる。いくつかの開示された実装では、CHASMの少なくともいくつかの側面は、本明細書で「スマートホームハブ」と称するものによって実装することができる。いくつかの例によれば、CHASMはオーディオ環境の特定の装置によって実装されてもよい。い

くつかの例では、CHASMは、少なくとも部分的には、オーディオ環境の一つまたは複数の装置によって実行されるソフトウェアを介して実装される。いくつかの実施形態では、装置（たとえば、スマートオーディオ装置）は、本稿で時に、発見可能な日和見的にオーケストレーションされる分散オーディオ・サブシステム（Discoverable Opportunistically Orchestrated Distributed Audio Subsystem、DOODAD）と称されるネットワーク接続可能な要素またはサブシステム（たとえば、ネットワーク接続可能なメディア・エンジンおよび装置属性記述子）を含み、複数の（たとえばより多数の）装置（たとえば、スマート・オーディオ装置またはDOODADSを含む他の装置）が、CHASMによって集合的に管理されるか、またはオーケストレーションされた機能を達成する別の仕方（たとえば、最初に購入したときに既知であったまたは装置のために意図されていたもの）に取って代わる）で実施される。ここで、開発のアーキテクチャーと、CHASM可能なオーディオ・システムのオーディオ機能を表現し制御するのに適した制御言語の両方について述べる。また、ここではオーケストレーションの言語を説明し、オーディオの装置（またはルート）を直接参照することなく、集合的オーディオ・システムをアドレッシングする基本的な要素と相違点を述べる。また、永続的なセッション、宛先、優先順位付け、およびオーディオをルーティングして確認応答を求めることについて説明する。これらは、オーケストレーションおよびオーディオを人や場所へ／からルーティングすることに特有なものである。

【0013】

本開示の諸側面は、開示された方法またはそのステップのいずれかの実施形態を実行するように構成された（たとえば、プログラムされた）システムと、開示された方法またはそのステップのいずれかの実施形態を実行するためのコード（たとえば、実行するために実行可能なコード）を記憶する、データの非一時的記憶（たとえば、ディスクまたは他の有形記憶媒体）を実装する有形の非一時的コンピュータ読み取り可能媒体とを含む。たとえば、いくつかの実施形態は、プログラム可能な汎用プロセッサ、デジタル信号プロセッサ、またはマイクロプロセッサであって、ソフトウェアまたはファームウェアでプログラムされ、かつ/または他の仕方で、開示された方法またはそのステップの一つまたは複数に従って、データに対して多様な操作のいずれかを実行するように構成されたものであってもよく、またはそれらを含むことができる。そのような汎用プロセッサは、入力装置、メモリ、および、それに対して呈されたデータに回答して開示された方法（またはそのステップ）のより多くを実行するようにプログラムされた（および／または別様に構成された）処理サブシステムを含むコンピュータ・システムであってもよく、または、それを含んでいてもよい。

【0014】

本開示の少なくともいくつかの側面は、方法によって実施することができる。いくつかの例において、本方法は、少なくとも部分的に、本明細書に開示されたもののような制御システムによって実装されてもよい。いくつかのそのような方法は、オーディオ環境のオーディオ・システムのオーディオ・セッション管理を含んでいてもよい。

【0015】

いくつかのそのような方法は、オーディオ・セッション・マネージャとオーディオ・システムの少なくとも第1のスマート・オーディオ装置との間の第1のスマート・オーディオ装置通信リンクを確立することを含んでいてもよい。いくつかの例では、第1のスマート・オーディオ装置は、単一目的のオーディオ装置または多目的のオーディオ装置のいずれかであるか、またはそれらを含む。いくつかのそのような例では、第1のスマート・オーディオ装置は、一つまたは複数のラウドスピーカーを含む。いくつかのそのような方法は、オーディオ・セッション・マネージャと第1のアプリケーションを実行する第1のアプリケーション装置との間の第1のアプリケーション通信リンクを確立することを含む。

【0016】

いくつかのそのような方法は、オーディオ・セッション・マネージャによって、第1のスマート・オーディオ装置の第1のメディア・エンジンの一つまたは複数の第1のメディア

・エンジン機能を決定することを含む。いくつかの例では、第1のメディア・エンジンは、第1のスマート・オーディオ装置によって受領された一つまたは複数のオーディオメディア・ストリームを管理し、第1のメディア・エンジンのサンプル・クロックに従って一つまたは複数のオーディオメディア・ストリームに対して第1のスマート・オーディオ装置信号処理を実行するように構成されている。

【0017】

いくつかのそのような例では、この方法は、第1のアプリケーションから第1のアプリケーション制御信号を、オーディオ・セッション・マネージャによって、第1のアプリケーション通信リンクを介して受領することを含む。いくつかのそのような方法は、第1のメディア・エンジン能力に従って第1のスマート・オーディオ装置を制御することを含む。いくつかの実装によれば、制御は、第1のスマート・オーディオ装置通信リンクを介して第1のスマート・オーディオ装置に送信される第1のオーディオ・セッション管理信号を介して、オーディオ・セッション・マネージャによって行われる。いくつかのそのような例では、オーディオ・セッション・マネージャは、第1のメディア・エンジン・サンプル・クロックを参照することなく、第1のスマート・オーディオ装置に第1のオーディオ・セッション管理信号を送信する。いくつかの実装では、第1のアプリケーション通信リンクは、第1のアプリケーション装置からの第1のルート開始要求に応答して確立されてもよい。

【0018】

いくつかの例によれば、第1のアプリケーション制御信号は、第1のメディア・エンジン・サンプルクロックを参照することなく、第1のアプリケーションから送信されてもよい。いくつかの例では、第1のオーディオ・セッション管理制御信号は、第1のスマート・オーディオ装置に、第1のメディア・エンジンの制御をオーディオ・セッション・マネージャに委任させることができる。

【0019】

いくつかの例によれば、オーディオ・セッション・マネージャまたは第1のスマート・オーディオ装置以外の装置が、第1のアプリケーションを実行するために構成されてもよい。しかしながら、場合によっては、第1のスマート・オーディオ装置が、第1のアプリケーションを実行するために構成されてもよい。いくつかの例では、第1のスマート・オーディオ装置は、特定の目的のオーディオ・セッション・マネージャを含むことができる。

【0020】

いくつかのそのような例によれば、オーディオ・セッション・マネージャは、第1のスマート・オーディオ装置通信リンクを介して特定目的のオーディオ・セッション・マネージャと通信することができる。いくつかのそのような例によれば、オーディオ・セッション・マネージャは、特定目的のオーディオ・セッション・マネージャから一つまたは複数の第1のメディア・エンジン能力を得ることができる。

【0021】

いくつかの実装によれば、オーディオ・セッション・マネージャは、アプリケーションが第1のスマート・オーディオ装置上で実行されているか、他の装置上で実行されているかにかかわらず、第1のメディア・エンジンを制御するすべてのアプリケーションのためのゲートウェイとして機能する可能性がある。

【0022】

いくつかのそのような方法は、第1のオーディオ源に対応する少なくとも第1のオーディオ・ストリームを確立することを含んでいてもよい。第1のオーディオ・ストリームは、第1のオーディオ信号を含んでいてもよい。いくつかのそのような例では、少なくとも第1のオーディオ・ストリームを確立することは、第1のスマート・オーディオ装置通信リンクを介して第1のスマート・オーディオ装置に送信される第1のオーディオ・セッション管理制御信号を介して、第1のスマート・オーディオ装置に、少なくとも第1のオーディオ・ストリームを確立させることを含むことができる。

【0023】

いくつかの例では、そのような方法は、第1のオーディオ信号を第1のレンダリングされ

10

20

30

40

50

たオーディオ信号にレンダリングさせるレンダリング・プロセスを含むこともできる。いくつかの例では、レンダリング・プロセスは、第1のオーディオ・セッション管理制御信号に応答して、第1のスマート・オーディオ装置によって実行されてもよい。

【0024】

いくつかのそのような方法は、第1のオーディオ・セッション管理制御信号を介して、第1のスマート・オーディオ装置に、第1のスマート・オーディオ装置と、オーディオ環境の一つまたは複数の他のスマート・オーディオ装置のそれぞれとの間のスマート・オーディオ装置間通信リンクを確立させることを含んでいてもよい。いくつかのそのような方法は、第1のスマート・オーディオ装置に、スマート・オーディオ装置間通信リンク（単数または複数）を介して、生のマイクロフォン信号、処理されたマイクロフォン信号、レンダリングされたオーディオ信号および／または未レンダリングのオーディオ信号を前記一つまたは複数の他のスマート・オーディオ装置に送信させることを含む。

【0025】

いくつかの例では、そのような方法はまた、オーディオ・セッション・マネージャと家庭オーディオ・システムの少なくとも第2のスマート・オーディオ装置との間に第2のスマート・オーディオ装置通信リンクを確立することを含む。第2のスマート・オーディオ装置は、単一目的のオーディオ装置または多目的のオーディオ装置のいずれかであってもよいし、それを含んでいてもよい。第2のスマート・オーディオ装置は、一つまたは複数のマイクロフォンを含んでいてもよい。いくつかのそのような方法は、オーディオ・セッション・マネージャによって、第2のスマート・オーディオ装置の第2のメディア・エンジンの一つまたは複数の第2のメディア・エンジン能力を決定することを含んでいてもよい。第2のメディア・エンジンは、一つまたは複数のマイクロフォンからマイクロフォン・データを受領し、第2のスマート・オーディオ装置信号処理を該マイクロフォン・データに対して実行するように構成することができる。いくつかのそのような方法は、第2のスマート・オーディオ装置通信リンクを介して第2のスマート・オーディオ装置に送信される第2のオーディオ・セッション・マネージャ制御信号を介して、オーディオ・セッション・マネージャによって第2のメディア・エンジン機能に従って第2のスマート・オーディオ装置を制御することを含んでいてもよい。

【0026】

いくつかのそのような例によれば、第2のスマート・オーディオ装置を制御することは、第2のスマート・オーディオ装置に、第2のスマート・オーディオ装置と第1のスマート・オーディオ装置との間のスマート・オーディオ装置間通信リンクを確立させることを含む。いくつかの例では、第2のスマート・オーディオ装置を制御することは、第2のスマート・オーディオ装置に、処理済および／または未処理のマイクロフォン・データを第2のメディア・エンジンから第1のメディア・エンジンに、スマート・オーディオ装置間通信リンクを介して送信させることを含む。

【0027】

いくつかの例では、第2のスマート・オーディオ装置を制御することは、オーディオ・セッション・マネージャによって、第1のアプリケーション通信リンクを介して、第1のアプリケーションから第1のアプリケーション制御信号を受領し、第1のアプリケーション制御信号に従って第2のオーディオ・セッション・マネージャ制御信号を決定することを含むことができる。

【0028】

代替的または追加的に、いくつかのオーディオ・セッション管理方法は、第1のアプリケーションを実装する第1の装置から、オーディオ・セッション・マネージャを実装する装置によって、第1のオーディオ・セッションのための第1のルートを開始するための第1のルート開始要求を受領することを含む。いくつかの例では、第1のルート開始要求は、第1のオーディオソースと第1のオーディオ環境宛先とを示し、前記第1のオーディオ環境宛先は、オーディオ環境における少なくとも第1の人物と対応するが、第1のオーディオ環境宛先は、オーディオ装置を示すものではない。

【0029】

いくつかのそのような方法は、オーディオ・セッション・マネージャを実装する装置によって、第1のルート開始要求に対応する第1のルートを確認することを含むものがある。いくつかの例によれば、第1のルートを確認することは、オーディオ環境内の少なくとも第1の人物の第1の位置を決定すること、第1のオーディオ・セッションの第1のステージに対して少なくとも1つのオーディオ装置を決定すること、および第1のオーディオ・セッションを開始またはスケジュールすることを含む。

【0030】

いくつかの例によれば、第1のルート開始要求は、第1のオーディオ・セッション優先度を含んでいてもよい。いくつかの例では、第1のルート開始要求は第1の接続性モードを含むことができる。たとえば、第1の接続性モードは、同期的接続性モード、トランザクション接続性モード、またはスケジュールされた接続性モードであってもよい。

10

【0031】

いくつかの実装では、第1のルート開始要求は、少なくとも第1の人物から確認応答が要求されるかどうかの指示を含んでいてもよい。いくつかの例では、第1のルート開始要求は、第1のオーディオ・セッション目標を含む可能性がある。たとえば、第1のオーディオ・セッション目標は、了解度、オーディオ品質、空間的忠実度、可聴性、非可聴性、および／またはプライバシーを含みうる。

【0032】

いくつかのそのような方法は、第1のルートについて第1の永続的な固有のオーディオ・セッション識別子を決定することを含んでいてもよい。そのような方法は、第1の永続的な固有のオーディオ・セッション識別子を第1の装置に送信することを含んでいてもよい。

20

【0033】

いくつかの例によれば、第1のルートを確認することは、環境内の少なくとも1つの装置に、第1のルートに対応する少なくとも第1のメディア・ストリームを確認させることを含む。第1のメディア・ストリームは、第1のオーディオ信号を含む。いくつかのそのような方法は、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングさせることを含んでいてもよい。

【0034】

いくつかのそのような方法は、オーディオ・セッションの第1のステージに対する第1の人物の第1の配向を決定することを含むことができる。いくつかのそのような例によれば、第1のオーディオ信号を第1にレンダリングされたオーディオ信号にレンダリングさせることは、第1の人物の第1の位置および第1の方向に対応する第1の基準空間モードを決定することと、第1の基準空間モードに対応するオーディオ環境におけるラウドスピーカの第1の相対的アクティブ化を決定することを含むことができる。

30

【0035】

いくつかのそのような方法は、第1のオーディオ・セッションの第2のステージについて第1の人物の第2の位置および／または第2の方向を決定することを含むものがある。いくつかのそのような方法は、第2の位置および／または第2の方向に対応する第2の基準空間モードを決定し、第2の基準空間モードに対応するオーディオ環境におけるラウドスピーカの第2の相対的アクティブ化を決定することを含むことができる。

40

【0036】

いくつかの例によれば、方法は、第2のアプリケーションを実装する第2の装置から、オーディオ・セッション・マネージャを実装する装置によって、第2のオーディオ・セッションのための第2のルートを開始するための第2のルート開始要求を受領することを含む。第2のルート開始要求は、第2のオーディオソースと第2のオーディオ環境宛先を示す可能性がある。第2のオーディオ環境宛先は、オーディオ環境内の少なくとも第2の人物と対応してもよい。いくつかの例では、第2のオーディオ環境宛先がオーディオ装置を示さない。

【0037】

50

いくつかのそのような方法は、オーディオ・セッション・マネージャを実装する装置によって、第2のルート開始要求に対応する第2のルートを確認することを含んでいてもよい。いくつかの実装では、第2のルートを確認することは、オーディオ環境内の少なくとも第2の人物の第1の位置を決定し、第2のオーディオ・セッションの第1のステージのために少なくとも1つのオーディオ装置を決定し、第2のオーディオ・セッションを開始することを含みうる。いくつかの例において、第2のルートを確認することは、第2のルートに対応する少なくとも第2のメディア・ストリームを確認することを含みうる。第2のメディア・ストリームは第2のオーディオ信号を含む。いくつかのそのような方法は、第2のオーディオ信号を第2のレンダリングされたオーディオ信号にレンダリングさせることを含む。

【0038】

いくつかのそのような方法は、少なくとも部分的に第2のオーディオ信号、第2のレンダリングされたオーディオ信号またはそれらの特性のうちの少なくとも1つに基づいて、第1のオーディオ信号についてのレンダリング・プロセスを修正して、修正された第1のレンダリングされたオーディオ信号を生成することを含む。いくつかの例によれば、第1のオーディオ信号についてのレンダリング・プロセスを修正することは、第2のレンダリングされたオーディオ信号のレンダリング位置から離れるように、第1のオーディオ信号のレンダリングをワープさせることを含むことができる。代替的または追加的に、第1のオーディオ信号のレンダリング・プロセスを修正することは、第2のオーディオ信号または第2のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスにตอบสนองして、第1のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスを修正することを含んでいてもよい。

【0039】

いくつかの例では、第1のルート開始要求は、第1のルート・ソース〔源〕または第1のルート宛先として、少なくともオーディオ環境の第1のエリアを示す可能性がある。いくつかの実装では、第1のルート開始要求は、少なくとも第1のサービス（たとえば、音楽提供サービスまたはポッドキャスト提供サービスのようなオンラインコンテンツ提供サービス）を第1のオーディオソースとして示すことができる。

【0040】

本明細書に記載された動作、機能および／または方法の一部または全部は、一つまたは複数の非一時的媒体に記憶された命令（たとえば、ソフトウェア）に従って一つまたは複数の装置によって実行されうる。そのような非一時的媒体は、ランダムアクセスメモリ（RAM）装置、読み出し専用メモリ（ROM）装置などを含むが、これらに限定されない、本明細書に記載されたもののようなメモリ装置を含んでいてもよい。よって、本開示に記載される主題のいくつかの革新的な側面は、その上に記憶されるソフトウェアを有する一つまたは複数の非一時的媒体において実装されることができる。

【0041】

たとえば、ソフトウェアは、オーディオ環境のオーディオ・システムのためのオーディオ・セッション管理に関わる一つまたは複数の方法を実行するために一つまたは複数の装置を制御するための命令を含むことができる。いくつかのそのような方法は、オーディオ・セッション・マネージャとオーディオ・システムの少なくとも第1のスマート・オーディオ装置との間の第1のスマート・オーディオ装置通信リンクを確認することを含む。いくつかの例では、第1のスマート・オーディオ装置は、単一目的のオーディオ装置または多目的のオーディオ装置のいずれかであるか、またはそれらを含む。いくつかのそのような例では、第1のスマート・オーディオ装置は、一つまたは複数のラウドスピーカーを含む。いくつかのそのような方法は、オーディオ・セッション・マネージャと第1のアプリケーションを実行する第1のアプリケーション装置との間の第1のアプリケーション通信リンクを確認することを含む。

【0042】

いくつかのそのような方法は、オーディオ・セッション・マネージャによって、第1のスマート・オーディオ装置の第1のメディア・エンジンの一つまたは複数の第1のメディア

10

20

30

40

50

・エンジン機能を決定することを含む。いくつかの例では、第1のメディア・エンジンは、第1のスマート・オーディオ装置によって受領された一つまたは複数のオーディオメディア・ストリームを管理し、第1のメディア・エンジン・サンプルクロックに従って一つまたは複数のオーディオメディア・ストリームについて第1のスマート・オーディオ装置信号処理を実行するように構成されている。

【0043】

いくつかのそのような例では、この方法は、第1のアプリケーションから第1のアプリケーション制御信号を、オーディオ・セッション・マネージャによって、第1のアプリケーション通信リンクを介して受領することを含む。いくつかのそのような方法は、第1のメディア・エンジン能力に従って第1のスマート・オーディオ装置を制御することを含む。いくつかの実装によれば、該制御は、第1のスマート・オーディオ装置通信リンクを介して第1のスマート・オーディオ装置に送信される第1のオーディオ・セッション管理信号を介して、オーディオ・セッション・マネージャによって行われる。いくつかのそのような例では、オーディオ・セッション・マネージャは、第1のメディア・エンジン・サンプル・クロックを参照することなく、第1のスマート・オーディオ装置に第1のオーディオ・セッション管理制御信号を送信する。

【0044】

いくつかの実装では、第1のアプリケーション通信リンクは、第1のアプリケーション装置からの第1のルート開始要求に応答して確立されてもよい。いくつかの例によれば、第1のアプリケーション制御信号は、第1のメディア・エンジン・サンプルクロックを参照することなく、第1のアプリケーションから送信されてもよい。いくつかの例では、第1のオーディオ・セッション管理制御信号は、第1のスマート・オーディオ装置に、第1のメディア・エンジンの制御をオーディオ・セッション・マネージャに委任させることができる。

【0045】

いくつかの例によれば、オーディオ・セッション・マネージャまたは第1のスマート・オーディオ装置以外の装置が、第1のアプリケーションを実行するために構成されてもよい。しかしながら、場合によっては、第1のスマート・オーディオ装置が、第1のアプリケーションを実行するために構成されてもよい。

【0046】

いくつかの例では、第1のスマート・オーディオ装置は、特定目的のオーディオ・セッション・マネージャを含むことができる。いくつかのそのような例によれば、オーディオ・セッション・マネージャは、第1のスマート・オーディオ装置通信リンクを介して特定目的のオーディオ・セッション・マネージャと通信することができる。

【0047】

いくつかのそのような例によれば、オーディオ・セッション・マネージャは、特定目的のオーディオ・セッション・マネージャから一つまたは複数の第1のメディア・エンジン能力を得ることができる。

【0048】

いくつかの実装によれば、オーディオ・セッション・マネージャは、アプリケーションが第1のスマート・オーディオ装置上で実行されているか、他の装置上で実行されているかにかかわらず、第1のメディア・エンジンを制御するすべてのアプリケーションのためのゲートウェイとして機能する可能性がある。

【0049】

いくつかのそのような方法は、第1のオーディオ源に対応する少なくとも第1のオーディオ・ストリームを確立することを含んでいてもよい。第1のオーディオ・ストリームは、第1のオーディオ信号を含んでいてもよい。いくつかのそのような例では、少なくとも第1のオーディオ・ストリームを確立することは、第1のスマート・オーディオ装置通信リンクを介して第1のスマート・オーディオ装置に送信される第1のオーディオ・セッション管理制御信号を介して、第1のスマート・オーディオ装置に、少なくとも第1のオーディオ・ストリームを確立させることを含むことができる。

【0050】

いくつかの例では、そのような方法は、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングさせるレンダリング・プロセスを含むこともできる。いくつかの例では、レンダリング・プロセスは、第1のオーディオ・セッション管理制御信号に応答して、第1のスマート・オーディオ装置によって実行されてもよい。

【0051】

いくつかのそのような方法は、第1のオーディオ・セッション管理制御信号を介して、第1のスマート・オーディオ装置に、第1のスマート・オーディオ装置と、オーディオ環境の一つまたは複数の他のスマート・オーディオ装置のそれぞれとの間のスマート・オーディオ装置間通信リンクを確立させることを含んでいてもよい。いくつかのそのような方法は、第1のスマート・オーディオ装置に、スマート・オーディオ装置間通信リンク（単数または複数）を介して、生のマイクロフォン信号、処理されたマイクロフォン信号、レンダリングされたオーディオ信号および／または未レンダリングのオーディオ信号を前記一つまたは複数の他のスマート・オーディオ装置に送信させることを含む。

【0052】

いくつかの例では、そのような方法はまた、オーディオ・セッション・マネージャと家庭オーディオ・システムの少なくとも第2のスマート・オーディオ装置との間に第2のスマート・オーディオ装置通信リンクを確立することを含む。第2のスマート・オーディオ装置は、単一目的のオーディオ装置または多目的のオーディオ装置のいずれかであってもよいし、それを含んでいてもよい。第2のスマート・オーディオ装置は、一つまたは複数のマイクロフォンを含んでいてもよい。いくつかのそのような方法は、オーディオ・セッション・マネージャによって、第2のスマート・オーディオ装置の第2のメディア・エンジンの一つまたは複数の第2のメディア・エンジン能力を決定することを含んでいてもよい。第2のメディア・エンジンは、一つまたは複数のマイクロフォンからマイクロフォン・データを受領し、第2のスマート・オーディオ装置信号処理を該マイクロフォン・データに対して実行するように構成することができる。いくつかのそのような方法は、第2のスマート・オーディオ装置通信リンクを介して第2のスマート・オーディオ装置に送信される第2のオーディオ・セッション・マネージャ制御信号を介して、オーディオ・セッション・マネージャによって第2のメディア・エンジン機能に従って第2のスマート・オーディオ装置を制御することを含んでいてもよい。

【0053】

いくつかのそのような例によれば、第2のスマート・オーディオ装置を制御することは、第2のスマート・オーディオ装置に、第2のスマート・オーディオ装置と第1のスマート・オーディオ装置との間のスマート・オーディオ装置間通信リンクを確立させることを含む。いくつかの例では、第2のスマート・オーディオ装置を制御することは、第2のスマート・オーディオ装置に、処理済および／または未処理のマイクロフォン・データを第2のメディア・エンジンから第1のメディア・エンジンに、スマート・オーディオ装置間通信リンクを介して送信させることを含む。

【0054】

いくつかの例では、第2のスマート・オーディオ装置を制御することは、オーディオ・セッション・マネージャによって、第1のアプリケーション通信リンクを介して、第1のアプリケーションから第1のアプリケーション制御信号を受領し、第1のアプリケーション制御信号に従って第2のオーディオ・セッション・マネージャ制御信号を決定することを含むことができる。

【0055】

代替的または追加的に、ソフトウェアは、オーディオ環境のオーディオ・システムについてのオーディオ・セッション管理に関わる一つまたは複数の他の方法を実行するよう一つまたは複数の装置を制御するための命令を含んでいてもよい。いくつかのそのようなオーディオ・セッション管理方法は、第1のアプリケーションを実装する第1の装置から、オーディオ・セッション・マネージャを実装する装置によって、第1のオーディオ・セッ

10

20

30

40

50

ョンのための第1のルートを開始するための第1のルート開始要求を受領することを含む。いくつかの例では、第1のルート開始要求は、第1のオーディオソースと第1のオーディオ環境宛先とを示し、前記第1のオーディオ環境宛先は、オーディオ環境における少なくとも第1の人物と対応するが、第1のオーディオ環境宛先は、オーディオ装置を示すものではない。

【0056】

いくつかのそのような方法は、オーディオ・セッション・マネージャを実装する装置によって、第1のルート開始要求に対応する第1のルートを確認することを含むものがある。いくつかの例によれば、第1のルートを確認することは、オーディオ環境内の少なくとも第1の人物の第1の位置を決定すること、第1のオーディオ・セッションの第1のステージに対して少なくとも1つのオーディオ装置を決定すること、および第1のオーディオ・セッションを開始またはスケジュールすることを含む。

10

【0057】

いくつかの例によれば、第1のルート開始要求は、第1のオーディオ・セッション優先度を含んでいてもよい。いくつかの例では、第1のルート開始要求は第1の接続性モードを含むことができる。たとえば、第1の接続性モードは、同期的接続性モード、トランザクション接続性モード、またはスケジュールされた接続性モードであってもよい。

【0058】

いくつかの実装では、第1のルート開始要求は、少なくとも第1の人物から確認応答が要求されるかどうかの指示を含んでいてもよい。いくつかの例では、第1のルート開始要求は、第1のオーディオ・セッション目標を含む可能性がある。たとえば、第1のオーディオ・セッション目標は、了解度、オーディオ品質、空間的忠実度、可聴性、非可聴性、および／またはプライバシーを含みうる。

20

【0059】

いくつかのそのような方法は、第1のルートについて第1の永続的な固有のオーディオ・セッション識別子を決定することを含んでいてもよい。そのような方法は、第1の永続的な固有のオーディオ・セッション識別子を第1の装置に送信することを含んでいてもよい。

【0060】

いくつかの例によれば、第1のルートを確認することは、環境内の少なくとも1つの装置に、第1のルートに対応する少なくとも第1のメディア・ストリームを確認させることを含みうる。第1のメディア・ストリームは、第1のオーディオ信号を含む。いくつかのそのような方法は、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングさせることを含んでいてもよい。

30

【0061】

いくつかのそのような方法は、オーディオ・セッションの第1のステージに対する第1の人物の第1の配向を決定することを含むことができる。いくつかのそのような例によれば、第1のオーディオ信号を第1にレンダリングされたオーディオ信号にレンダリングさせることは、第1の人物の第1の位置および第1の方向に対応する第1の基準空間モードを決定することと、第1の基準空間モードに対応するオーディオ環境におけるラウドスピーカの第1の相対的アクティブ化を決定することを含むことができる。

40

【0062】

いくつかのそのような方法は、第1のオーディオ・セッションの第2のステージについて第1の人物の第2の位置および／または第2の方向を決定することを含むものがある。いくつかのそのような方法は、第2の位置および／または第2の方向に対応する第2の基準空間モードを決定し、第2の基準空間モードに対応するオーディオ環境におけるラウドスピーカの第2の相対的アクティブ化を決定することを含むことができる。

【0063】

いくつかの例によれば、方法は、第2のアプリケーションを実装する第2の装置から、オーディオ・セッション・マネージャを実装する装置によって、第2のオーディオ・セッションのための第2のルートを開始するための第2のルート開始要求を受領することを含みう

50

る。第2のルート開始要求は、第2のオーディオソースと第2のオーディオ環境宛先を示す可能性がある。第2のオーディオ環境宛先は、オーディオ環境内の少なくとも第2の人物と対応してもよい。いくつかの例では、第2のオーディオ環境宛先がオーディオ装置を示さない。

【0064】

いくつかのそのような方法は、オーディオ・セッション・マネージャを実装する装置によって、第2のルート開始要求に対応する第2のルートを確認することを含んでいてもよい。いくつかの実装では、第2のルートを確認することは、オーディオ環境内の少なくとも第2の人物の第1の位置を決定し、第2のオーディオ・セッションの第1のステージのために少なくとも1つのオーディオ装置を決定し、第2のオーディオ・セッションを開始することを含みうる。いくつかの例において、第2のルートを確認することは、第2のルートに対応する少なくとも第2のメディア・ストリームを確認することを含みうる。第2のメディア・ストリームは第2のオーディオ信号を含む。いくつかのそのような方法は、第2のオーディオ信号を第2のレンダリングされたオーディオ信号にレンダリングさせることを含む。

【0065】

いくつかのそのような方法は、少なくとも部分的に第2のオーディオ信号、第2のレンダリングされたオーディオ信号またはそれらの特性のうちの少なくとも1つに基づいて、第1のオーディオ信号についてのレンダリング・プロセスを修正して、修正された第1のレンダリングされたオーディオ信号を生成することを含む。いくつかの例によれば、第1のオーディオ信号についてのレンダリング・プロセスを修正することは、第2のレンダリングされたオーディオ信号のレンダリング位置から離れるように、第1のオーディオ信号のレンダリングをワープさせることを含むことができる。代替的または追加的に、第1のオーディオ信号のレンダリング・プロセスを修正することは、第2のオーディオ信号または第2のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスにตอบสนองして、第1のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスを修正することを含んでいてもよい。

【0066】

いくつかの例では、第1のルート開始要求は、第1のルート・ソース〔源〕または第1のルート宛先として、少なくともオーディオ環境の第1のエリアを示す可能性がある。いくつかの実装では、第1のルート開始要求は、少なくとも第1のサービス（たとえば、音楽提供サービスまたはポッドキャスト提供サービスのようなオンラインコンテンツ提供サービス）を第1のオーディオソースとして示すことができる。

【0067】

いくつかの実施において、装置（またはシステム）は、インターフェース・システムおよび制御システムを含んでいてもよい。制御システムは、一つまたは複数の汎用の単一チップまたは複数チッププロセッサ、デジタル信号プロセッサ（DSP）、特定用途向け集積回路（ASIC）、フィールドプログラマブルゲートアレイ（FPGA）、または他のプログラマブル論理装置、離散的ゲートまたはトランジスタ論理、離散的ハードウェアコンポーネント、またはそれらの組み合わせを含んでいてもよい。

【0068】

いくつかの実装において、制御システムは、本明細書に開示された方法の一つまたは複数を実装するために構成されてもよい。いくつかのそのような方法は、オーディオ環境のオーディオ・システムについてのオーディオ・セッション管理に関与することがある。いくつかのそのような例によれば、制御システムは、本明細書でオーディオ・セッション・マネージャと称されるものを実装するように構成することができる。

【0069】

いくつかのそのような方法は、オーディオ・セッション・マネージャ（たとえば、オーディオ・セッション・マネージャを実装する装置）とオーディオ・システムの少なくとも第1のスマート・オーディオ装置との間の第1のスマート・オーディオ装置通信リンクを確認することを含んでいてもよい。いくつかの例では、第1のスマート・オーディオ装置は

、単一目的のオーディオ装置または多目的のオーディオ装置のいずれかであるか、またはそれらを含む。いくつかのそのような例では、第1のスマート・オーディオ装置は、一つまたは複数のラウドスピーカーを含む。いくつかのそのような方法は、オーディオ・セッション・マネージャと第1のアプリケーションを実行する第1のアプリケーション装置との間の第1のアプリケーション通信リンクを確立することを含む。

【0070】

いくつかのそのような方法は、オーディオ・セッション・マネージャによって、第1のスマート・オーディオ装置の第1のメディア・エンジンの一つまたは複数の第1のメディア・エンジン機能を決定することを含む。いくつかの例では、第1のメディア・エンジンは、第1のスマート・オーディオ装置によって受領された一つまたは複数のオーディオメディア・ストリームを管理し、第1のメディア・エンジンのサンプル・クロックに従って一つまたは複数のオーディオメディア・ストリームに対して第1のスマート・オーディオ装置信号処理を実行するように構成されている。

10

【0071】

いくつかのそのような例では、この方法は、第1のアプリケーションから第1のアプリケーション制御信号を、オーディオ・セッション・マネージャによって、第1のアプリケーション通信リンクを介して受領することを含む。いくつかのそのような方法は、第1のメディア・エンジン能力に従って第1のスマート・オーディオ装置を制御することを含む。いくつかの実装によれば、制御は、第1のスマート・オーディオ装置通信リンクを介して第1のスマート・オーディオ装置に送信される第1のオーディオ・セッション管理信号を介して、オーディオ・セッション・マネージャによって行われる。いくつかのそのような例では、オーディオ・セッション・マネージャは、第1のメディア・エンジン・サンプル・クロックを参照することなく、第1のスマート・オーディオ装置に第1のオーディオ・セッション管理信号を送信する。

20

【0072】

いくつかの実装では、第1のアプリケーション通信リンクは、第1のアプリケーション装置からの第1のルート開始要求に応答して確立されてもよい。いくつかの例によれば、第1のアプリケーション制御信号は、第1のメディア・エンジン・サンプルクロックを参照することなく、第1のアプリケーションから送信されてもよい。いくつかの例では、第1のオーディオ・セッション管理制御信号は、第1のスマート・オーディオ装置に、第1のメディア・エンジンの制御をオーディオ・セッション・マネージャに委任させることができる。

30

【0073】

いくつかの例によれば、オーディオ・セッション・マネージャまたは第1のスマート・オーディオ装置以外の装置が、第1のアプリケーションを実行するために構成されてもよい。しかしながら、場合によっては、第1のスマート・オーディオ装置が、第1のアプリケーションを実行するために構成されてもよい。

【0074】

いくつかの例では、第1のスマート・オーディオ装置は、特定の目的のオーディオ・セッション・マネージャを含むことができる。いくつかのそのような例によれば、オーディオ・セッション・マネージャは、第1のスマート・オーディオ装置通信リンクを介して特定目的のオーディオ・セッション・マネージャと通信することができる。いくつかのそのような例によれば、オーディオ・セッション・マネージャは、特定目的のオーディオ・セッション・マネージャから一つまたは複数の第1のメディア・エンジン能力を得ることができる。

40

【0075】

いくつかの実装によれば、オーディオ・セッション・マネージャは、アプリケーションが第1のスマート・オーディオ装置上で実行されているか、他の装置上で実行されているかにかかわらず、第1のメディア・エンジンを制御するすべてのアプリケーションのためのゲートウェイとして機能する可能性がある。

【0076】

50

いくつかのそのような方法は、第1のオーディオ源に対応する少なくとも第1のオーディオ・ストリームを確立することを含んでいてもよい。第1のオーディオ・ストリームは、第1のオーディオ信号を含んでいてもよい。いくつかのそのような例では、少なくとも第1のオーディオ・ストリームを確立することは、第1のスマート・オーディオ装置通信リンクを介して第1のスマート・オーディオ装置に送信される第1のオーディオ・セッション管理制御信号を介して、第1のスマート・オーディオ装置に、少なくとも第1のオーディオ・ストリームを確立させることを含むことができる。

【0077】

いくつかの例では、そのような方法は、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングさせるレンダリング・プロセスを含むこともできる。いくつかの例では、レンダリング・プロセスは、第1のオーディオ・セッション管理制御信号に応答して、第1のスマート・オーディオ装置によって実行されてもよい。

【0078】

いくつかのそのような方法は、第1のオーディオ・セッション管理制御信号を介して、第1のスマート・オーディオ装置に、第1のスマート・オーディオ装置と、オーディオ環境の一つまたは複数の他のスマート・オーディオ装置のそれぞれとの間のスマート・オーディオ装置間通信リンクを確立させることを含んでいてもよい。いくつかのそのような方法は、第1のスマート・オーディオ装置に、スマート・オーディオ装置間通信リンク（単数または複数）を介して、生のマイクロフォン信号、処理されたマイクロフォン信号、レンダリングされたオーディオ信号および／または未レンダリングのオーディオ信号を前記一つまたは複数の他のスマート・オーディオ装置に送信させることを含むうる。

【0079】

いくつかの例では、そのような方法はまた、オーディオ・セッション・マネージャと家庭オーディオ・システムの少なくとも第2のスマート・オーディオ装置との間に第2のスマート・オーディオ装置通信リンクを確立することを含むうる。第2のスマート・オーディオ装置は、単一目的のオーディオ装置または多目的のオーディオ装置のいずれかであってもよいし、それを含んでいてもよい。第2のスマート・オーディオ装置は、一つまたは複数のマイクロフォンを含んでいてもよい。いくつかのそのような方法は、オーディオ・セッション・マネージャによって、第2のスマート・オーディオ装置の第2のメディア・エンジンの一つまたは複数の第2のメディア・エンジン能力を決定することを含んでいてもよい。第2のメディア・エンジンは、一つまたは複数のマイクロフォンからマイクロフォン・データを受領し、第2のスマート・オーディオ装置信号処理を該マイクロフォン・データに対して実行するように構成することができる。いくつかのそのような方法は、第2のスマート・オーディオ装置通信リンクを介して第2のスマート・オーディオ装置に送信される第2のオーディオ・セッション・マネージャ制御信号を介して、オーディオ・セッション・マネージャによって第2のメディア・エンジン機能に従って第2のスマート・オーディオ装置を制御することを含んでいてもよい。

【0080】

いくつかのそのような例によれば、第2のスマート・オーディオ装置を制御することは、第2のスマート・オーディオ装置に、第2のスマート・オーディオ装置と第1のスマート・オーディオ装置との間のスマート・オーディオ装置間通信リンクを確立させることを含むうる。いくつかの例では、第2のスマート・オーディオ装置を制御することは、第2のスマート・オーディオ装置に、処理済および／または未処理のマイクロフォン・データを第2のメディア・エンジンから第1のメディア・エンジンに、スマート・オーディオ装置間通信リンクを介して送信させることを含むうる。

【0081】

いくつかの例では、第2のスマート・オーディオ装置を制御することは、オーディオ・セッション・マネージャによって、第1のアプリケーション通信リンクを介して、第1のアプリケーションから第1のアプリケーション制御信号を受領し、第1のアプリケーション制御信号に従って第2のオーディオ・セッション・マネージャ制御信号を決定することを含

10

20

30

40

50

むことができる。

【0082】

代替的または追加的に、いくつかのオーディオ・セッション管理方法は、第1のアプリケーションを実装する第1の装置から、オーディオ・セッション・マネージャを実装する装置によって、第1のオーディオ・セッションのための第1のルートを開始するための第1のルート開始要求を受領することを含む。いくつかの例では、第1のルート開始要求は、第1のオーディオソースと第1のオーディオ環境宛先とを示し、前記第1のオーディオ環境宛先は、オーディオ環境における少なくとも第1の人物と対応するが、第1のオーディオ環境宛先は、オーディオ装置を示すものではない。

【0083】

いくつかのそのような方法は、オーディオ・セッション・マネージャを実装する装置によって、第1のルート開始要求に対応する第1のルートを確認することを含むものがある。いくつかの例によれば、第1のルートを確認することは、オーディオ環境内の少なくとも第1の人物の第1の位置を決定すること、第1のオーディオ・セッションの第1のステージに対して少なくとも1つのオーディオ装置を決定すること、および第1のオーディオ・セッションを開始またはスケジュールすることを含む。

【0084】

いくつかの例によれば、第1のルート開始要求は、第1のオーディオ・セッション優先度を含んでいてもよい。いくつかの例では、第1のルート開始要求は第1の接続性モードを含むことができる。たとえば、第1の接続性モードは、同期的接続性モード、トランザクション接続性モード、またはスケジュールされた接続性モードであってもよい。

【0085】

いくつかの実装では、第1のルート開始要求は、少なくとも第1の人物から確認応答が要求されるかどうかの指示を含んでいてもよい。いくつかの例では、第1のルート開始要求は、第1のオーディオ・セッション目標を含む可能性がある。たとえば、第1のオーディオ・セッション目標は、了解度、オーディオ品質、空間的忠実度、可聴性、非可聴性、および／またはプライバシーを含みうる。

【0086】

いくつかのそのような方法は、第1のルートについて第1の永続的な固有のオーディオ・セッション識別子を決定することを含んでいてもよい。そのような方法は、第1の永続的な固有のオーディオ・セッション識別子を第1の装置に送信することを含んでいてもよい。

【0087】

いくつかの例によれば、第1のルートを確認することは、環境内の少なくとも1つの装置に、第1のルートに対応する少なくとも第1のメディア・ストリームを確認させることを含む。第1のメディア・ストリームは、第1のオーディオ信号を含む。いくつかのそのような方法は、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングさせることを含んでいてもよい。

【0088】

いくつかのそのような方法は、オーディオ・セッションの第1のステージに対する第1の人物の第1の配向を決定することを含むことがある。いくつかのそのような例によれば、第1のオーディオ信号を第1にレンダリングされたオーディオ信号にレンダリングさせることは、第1の人物の第1の位置および第1の方向に対応する第1の基準空間モードを決定することと、第1の基準空間モードに対応するオーディオ環境におけるラウドスピーカの第1の相対的アクティブ化を決定することとを含むことができる。

【0089】

いくつかのそのような方法は、第1のオーディオ・セッションの第2のステージについて第1の人物の第2の位置および／または第2の方向を決定することを含むものがある。いくつかのそのような方法は、第2の位置および／または第2の方向に対応する第2の基準空間モードを決定し、第2の基準空間モードに対応するオーディオ環境におけるラウドスピーカの第2の相対的アクティブ化を決定することを含むことができる。

10

20

30

40

50

【0090】

いくつかの例によれば、方法は、第2のアプリケーションを実装する第2の装置から、オーディオ・セッション・マネージャを実装する装置によって、第2のオーディオ・セッションのための第2のルートを開始するための第2のルート開始要求を受領することを含みうる。第2のルート開始要求は、第2のオーディオソースと第2のオーディオ環境宛先を示す可能性がある。第2のオーディオ環境宛先は、オーディオ環境内の少なくとも第2の人物と対応してもよい。いくつかの例では、第2のオーディオ環境宛先がオーディオ装置を示さない。

【0091】

いくつかのそのような方法は、オーディオ・セッション・マネージャを実装する装置によって、第2のルート開始要求に対応する第2のルートを確認することを含んでいてもよい。いくつかの実装では、第2のルートを確認することは、オーディオ環境内の少なくとも第2の人物の第1の位置を決定し、第2のオーディオ・セッションの第1のステージのために少なくとも1つのオーディオ装置を決定し、第2のオーディオ・セッションを開始することを含みうる。いくつかの例において、第2のルートを確認することは、第2のルートに対応する少なくとも第2のメディア・ストリームを確認することを含みうる。第2のメディア・ストリームは第2のオーディオ信号を含む。いくつかのそのような方法は、第2のオーディオ信号を第2のレンダリングされたオーディオ信号にレンダリングさせることを含む。

【0092】

いくつかのそのような方法は、少なくとも部分的に第2のオーディオ信号、第2のレンダリングされたオーディオ信号またはそれらの特性のうちの少なくとも1つに基づいて、第1のオーディオ信号についてのレンダリング・プロセスを修正して、修正された第1のレンダリングされたオーディオ信号を生成することを含む。いくつかの例によれば、第1のオーディオ信号についてのレンダリング・プロセスを修正することは、第2のレンダリングされたオーディオ信号のレンダリング位置から離れるように、第1のオーディオ信号のレンダリングをワープさせることを含むことができる。代替的または追加的に、第1のオーディオ信号のレンダリング・プロセスを修正することは、第2のオーディオ信号または第2のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスにตอบสนองして、第1のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスを修正することを含んでいてもよい。

【0093】

いくつかの例では、第1のルート開始要求は、第1のルート・ソース〔源〕または第1のルート宛先として、少なくともオーディオ環境の第1のエリアを示す可能性がある。いくつかの実装では、第1のルート開始要求は、少なくとも第1のサービス（たとえば、音楽提供サービスまたはポッドキャスト提供サービスのようなオンラインコンテンツ提供サービス）を第1のオーディオソースとして示すことができる。

【0094】

本明細書に記載される主題の一つまたは複数の実施の詳細は、添付の図面および以下の説明に記載される。他の特徴、側面、および利点は、明細書、図面、および特許請求の範囲から明らかになる。以下の図の相対的な寸法は、縮尺通りに描かれない場合があることに留意されたい。

【図面の簡単な説明】

【0095】

【図1A】オーディオ環境における人物およびスマート・オーディオ装置の例を示す。

【図1B】図1Aのシナリオの修正版の図である。

【図1C】本開示のある実施形態による。

【図2A】従来のシステムのブロック図である。

【図2B】図2Aに示されている装置の修正バージョンの例を示す。

【図2C】ある開示された実装の例のブロック図である。

【図2D】連続的階層オーディオ・セッション・マネージャ（CHASM）と対話する複数

10

20

30

40

50

のアプリケーションの例を示す。

【図 3 A】一例による、図 1A の装置 101 の詳細を示すブロック図である。

【図 3 B】一例による、図 1B の実装の詳細を示す。

【図 3 C】オーディオ環境の 2 つのオーディオ装置をオーケストレーションする CHASM の例を示すブロック図である。

【図 4】別の開示された実施形態を示すブロック図である。

【図 5】いくつかの実装に従った音声セッション管理方法のブロックを含むフロー図である。

【図 6】本開示のさまざまな側面を実施可能な装置の構成要素の例を示すブロック図である。

【図 7】一例による CHASM のブロックを示すブロック図である。

【図 8】一例による図 7 に示すルーティング・テーブルの詳細を示す。

【図 9 A】オーケストレーションの言語におけるルート開始要求のコンテキストフリー文法の例を示す。

【図 9 B】オーディオ・セッション目標の例を与える。

【図 10】一例による、ルートを修正する要求のフローを示す。

【図 11 A】ルートを修正する要求についてのフローの追加例を示す。

【図 11 B】ルートを修正する要求についてのフローの追加例を示す。

【図 11 C】ルートを削除するためのフローの例を示す。

【図 12】いくつかの実装に従ったオーディオ・セッション管理方法のブロックを含むフロー図である。

【図 13】いくつかの実装によるオーディオ・セッション管理方法のブロックを含むフロー図である。

【図 14】いくつかの実装に従ったオーディオ・セッション管理方法のブロックを含むフロー図である。

【図 15】いくつかの実装による、オーディオ環境に新たに導入される一つまたは複数のオーディオ装置のための自動セットアップ・プロセスのブロックを含むフロー図である。

【図 16】いくつかの実装による、バーチャル・アシスタント・アプリケーションをインストールするためのプロセスのブロックを含むフロー図である。

【図 17】いくつかの実装によるオーディオ・セッション管理方法のブロックを含むフロー図である。

【図 18 A】ある実施形態の最小バージョンのブロック図である。

【図 18 B】追加的な特徴を有する別の（より高機能な）実施形態を示す。

【図 19】図 6、図 18A または図 18B に示されるもののような装置またはシステムによって実行される方法の一例を概説するフロー図である。

【図 20】接続された居住空間のフロアプランの例を示す。

【図 21】接続された居住空間のフロアプランの例を示す。

【図 22】空間的な音楽ミックスと音声アシスタント応答の同時再生を提供するマルチストリーム・レンダラーの例を示す。

【図 23】空間的な音楽ミックスと音声アシスタント応答の同時再生を提供するマルチストリーム・レンダラーの例を示す。

【図 24】空間的な音楽ミックスが、パーティーの多数の人のために居間およびキッチン内のすべてのスピーカーを横断して最適に再生される例示的な出発点を示す。

【図 25】赤ん坊が寝室で寝ようとしているところの例を示す。

【図 26】追加的なオーディオ・ストリームの再生の例を示す。

【図 27】図 18A に示されるマルチストリーム・レンダラーの周波数／変換領域の例を示す。

【図 28】図 18B に示されるマルチストリーム・レンダラーの周波数／変換領域の例を示す。

【図 29】この例の居住空間である聴取環境のフロアプランを示す。

10

20

30

40

50

【図30】図29に示される居住空間における複数の異なる聴取位置および配向のために基準空間モードにおいて空間的オーディオを柔軟にレンダリングする例を示す。

【図31】図29に示される居住空間における複数の異なる聴取位置および配向のために基準空間モードにおいて空間的オーディオを柔軟にレンダリングする例を示す。

【図32】図29に示される居住空間における複数の異なる聴取位置および配向のために基準空間モードにおいて空間的オーディオを柔軟にレンダリングする例を示す。

【図33】図29に示される居住空間における複数の異なる聴取位置および配向のために基準空間モードにおいて空間的オーディオを柔軟にレンダリングする例を示す。

【図34】2人の聴取者が聴取環境の異なる位置にある場合の基準空間モード・レンダリングの例を示す。

10

【図35】聴取者の位置と配向に関するユーザー入力を受領するためのGUIの例を示す。

【図36】環境内の3つのオーディオ装置間の幾何学的関係の例を示す。

【図37】図36に示す環境における3つのオーディオ装置間の幾何学的関係の別の例を示す。

【図38】図36および図37に示される三角形の両方を、対応するオーディオ装置および環境の他の特徴なしで示す。

【図39】3つのオーディオ装置によって形成される三角形の内角を推定する例を示す。

【図40】図6に示されるような装置によって実行されうる方法の一例を概説するフロー図である。

【図41】環境内の各オーディオ装置が複数の三角形の頂点である例を示す。

20

【図42】順整列プロセスの一部の例を示す。

【図43】順整列プロセス中に発生したオーディオ装置位置の複数の推定値の例を示す。

【図44】逆整列プロセスの一部の例を示す。

【図45】逆整列プロセス中に発生したオーディオ装置位置の複数の推定値の例を示す。

【図46】推定されたオーディオ装置位置と実際のオーディオ装置位置との比較を示す。

【図47】図6に示されるような装置によって実行されうる方法の一例を概説するフロー図である。

【図48A】図47のいくつかのブロックの例を示す。

【図48B】聴取者角度配向データを決定することのさらなる例を示す。

【図48C】聴取者角度配向データを決定することのさらなる例を示す。

30

【図48D】図48Cを参照して説明した方法に従った、オーディオ装置座標についての適切な回転を決定する例を示す。

【図49】本開示のさまざまな側面を実装できるシステムの構成要素の例を示すブロック図である。

【図50】A、B、Cは、再生制限閾値および対応する周波数の例を示す。

【図51】A、Bは、ダイナミックレンジ圧縮データの例を示すグラフである。

【図52】聴取環境の空間ゾーンの例を示す。

【図53】図52の空間ゾーン内のラウドスピーカーの例を示す。

【図54】図53の空間ゾーンおよびスピーカー上に重ねられた公称空間位置の例を示す。

【図55】本明細書に開示されたもののような装置またはシステムによって実行されうる方法の一例を概説するフロー図である。

40

【図56A】いくつかの実施形態に従って実装可能なシステムの例を示す。

【図56B】いくつかの実施形態に従って実装可能なシステムの例を示す。

【図57】ある実施形態による環境（たとえば、家庭）で実装されるシステムのブロック図である。

【図58】図57のモジュール5701の例示的实施形態の要素のブロック図である。

【図59】図57のモジュール5701の別の例示的实施形態（図59では5900とラベル付けされている）およびその動作のブロック図である。

【図60】別の例示的な実施形態のブロック図である。

【発明を実施するための形態】

50

【0096】

多くの実施形態が開示される。当業者には、それらをどのように実施するかが、本開示から明らかになるであろう。

【0097】

現在、デザイナーは、オーディオ装置を、エンターテインメント、通信、および情報サービスのブレンドとなりうるオーディオのための単一のインターフェース・ポイントとして考えている。通知および音声制御のためにオーディオを使用することは、視覚的または物理的侵害を回避する利点がある。拡大する装置ランドスケープは、我々の1対の耳をめぐって競合するシステムが増えたことで、フラグメント化されている。ウェアラブルで拡張されたオーディオが利用可能になり始めたため、理想的などこにでもあるオーディオ・パーソナル・アシスタントを可能にすることに向けて事態が収束しつつあるようには見えず、我々のまわりにある多数の装置を、シームレスな捕捉、接続性、通信のために使用することは可能ではなかった。

10

【0098】

諸装置をブリッジするためのサービスを開発し、位置、コンテキスト、コンテンツ、タイミング、およびユーザー選好をよりよく管理することが有用であろう。規格、インフラストラクチャー、APIのセットが一緒になって、ユーザーの周囲の1つのオーディオ空間への統合されたアクセスへのよりよいアクセスを可能にできる。我々は、基本的なオーディオ入力出力を管理し、特定のアプリケーションへのオーディオ装置の接続性を許容する、オーディオ装置のための一種のオペレーティングシステムを考えている。この考え方と設計は、たとえば、改善の迅速で有機的な開発を可能にし、他者のために、装置に依存しないオーディオ接続性を提供するサービスを提供するために、対話的オーディオ・トランスポートの足場を作り出す。

20

【0099】

オーディオ対話のスペクトルは、リアルタイム通信、非同期チャット、アラート、文字起こし、履歴、アーカイブ、音楽、推薦、リマインダー、プロモーション、コンテキスト認識アシストを含む。ここでは、統一されたアプローチを容易にし、インテリジェントなメディア・フォーマットを実装できるプラットフォームを開示する。プラットフォームは、ユビキタス・ウェアラブル・オーディオを含むかまたは実装することができ、および／またはユーザーの位置特定、単一または複数の（たとえば、集合的な）オーディオ装置を最善の使用のために選択すること、識別情報、プライバシー、適時性、ジオロケーションおよび／または転送、記憶、検索およびアルゴリズム実行のためのインフラストラクチャーを管理することができる。本開示のいくつかの側面は、識別情報、優先度（ランク）およびユーザー選好の尊重、たとえば聞くことの望ましさおよび聞こえることの価値を管理することを含む。望まれないオーディオのコストは高い。我々は、「オーディオのインターネット」がセキュリティおよび信頼の不可欠な要素を提供または実装しうると考えている。

30

【0100】

単一目的のオーディオ装置および多目的のオーディオ装置のカテゴリーは、厳密には直交するものではないが、オーディオ装置（たとえば、スマートオーディオ装置）のスピーカーおよびマイクロフォンは、スマート・オーディオ装置によって可能にされる、またはスマート・オーディオ装置に接続される（またはスマート・オーディオ装置によって実装される）機能に割り当てられてもよい。しかしながら、典型的には、オーディオ装置のスピーカーおよび／またはマイクロフォンは、個別に（オーディオ装置とは別に）考えれば、集合体に追加されうるという意味は存在しない。

40

【0101】

ここでは、ローカルオーディオ装置（それぞれがスピーカーおよびマイクロフォンを含むことがある）が広告され、ローカルオーディオ装置のどの一つとも無関係な抽象的な意味で存在する集合的オーディオ・プラットフォームのために利用可能にされる、オーディオ装置接続性のカテゴリーについて述べる。我々はまた、集合的オーディオ装置オーケス

50

トレーションおよび利用のこの概念を実現することに向けた設計アプローチおよびステップの集合を実装する、少なくとも1つの発見可能な日和見的オーケストレーションされた分散オーディオ装置サブシステム（DOODAD）を含む実施形態を説明する。

【0102】

本開示のいくつかの実施形態の適用および結果を実証するために、図1Aを参照して、簡単な例を説明する。

【0103】

図1Aは、オーディオ環境における、人物およびスマート・オーディオ装置の例を示す。この例では、オーディオ環境100は家庭オーディオ環境である。

【0104】

図1Aのシナリオでは、人物（102）は、マイクロフォンを用いてユーザーの音声（103）を取り込むことができ、かつスピーカー再生（104）が可能なスマート・オーディオ装置（101）を有する。ユーザーは、装置101からかなり離れたところで話することがあり、そのことは、装置101の複信（duplex）能力を制限する。

【0105】

図1Aでは、ラベル付けされた要素は以下の通りである：

- ・100. スマート・オーディオ装置の使用の例示的シナリオを示すオーディオ環境。椅子に座っている人102が、装置101と対話する。
- ・101. スピーカーを通じたオーディオ再生およびマイクロフォンからのオーディオ捕捉ができるスマート・オーディオ装置。
- ・102. 装置101を使用するオーディオ体験に参加する人〔人物〕（ユーザーまたは聴取者とも呼ばれる）。
- ・103. 装置101に話す人102によって発声される音。
- ・104. 装置101のスピーカーから再生されるオーディオ。

【0106】

よって、この例では、図1Aは、スマート・オーディオ装置（101）のある自宅にいる人（102）の図であり、スマート・オーディオ装置はこの場合は、通信用の電話装置である。装置101は、人102によって聞かれるオーディオを出力することができ、また人102から音（103）を取り込むこともできる。しかしながら、装置101は、人102からある程度離れているので、装置101上のマイクロフォンは、装置101によって出力されているオーディオより大きく人102を聴くという課題を有する。これはエコー問題として知られている課題である。この問題を解決する確立された技術は、典型的には、エコーキャンセルを使用する。これは、全二重活動（二重通話、すなわち、両方向での同時音声）によって厳しく制限される処理の一形態である。

【0107】

図1Bは、図1Aのシナリオの修正版の図である。この例では、第2のスマート・オーディオ装置（105）もオーディオ環境100内にある。スマート・オーディオ装置105は、人102の近くに位置するが、オーディオを出力する機能のために設計されている。いくつかの例では、スマート・オーディオ装置105は、少なくとも部分的に、バーチャル・アシスタントを実装することができる。

【0108】

（図1Aの）人物102は、今や第2のスマート・オーディオ装置105を入手したが、装置105（図1Bに示される）は、第1の装置（101）の目的とは完全に別個の特定の目的（106）を実行することができるだけである。この例では、2つのスマート・オーディオ装置（101および105）は、ユーザーと同じ音響空間にあるにもかかわらず、情報を共有し、体験をオーケストレーションすることができない。

【0109】

図1Bにおいて、ラベル付けされた要素は以下の通りである：

- ・100～104. 図1A参照。
- ・105. スピーカーを通じたオーディオ再生およびマイクロフォンからのオーディオ取り

込みができる、地下的なスマート・オーディオ装置。

・106. 装置105のスピーカーを通じたオーディオの再生

電話(101)からのオーディオ通話をこのスマート・オーディオ装置(105)にペアリングまたはシフトすることは可能でありうるが、これは、ユーザーの介入および詳細な構成がなければ、これまでは不可能であった。よって、図1Bに示されたシナリオは、2つの独立したオーディオ装置のものであり、それぞれが非常に特定のアプリケーションを実行する。この例では、スマート・オーディオ装置105は、装置100よりも最近購入された。スマート・オーディオ装置(105)は、その特定の目的のために購入されたものであり、箱から取り出した状態では、その特定の目的のためにのみ有用であり、オーディオ環境100内にすでに存在しており、通信装置として使用されていた装置(101)にただちに価値を付加するものではない。

10

【0110】

図1Cは、本開示のある実施形態による。この例では、スマート・オーディオ装置105は、スマート・オーディオ装置100よりも最近購入された。

【0111】

図1Cの実施形態では、スマート・オーディオ装置101および105は、オーケストレーション可能である。このオーケストレーションの例では、スマート・オーディオ装置(101)がそのスピーカーから音声104を再生するので、スマート・オーディオ装置101に関わる電話通話のためには、スマート・オーディオ装置(105)のほうが、人102の声(103B)をピックアップするのによい位置にある。

20

【0112】

図1Cでは、ラベル付けされた要素は以下の通りである:

・100~104. 図1A参照。

・105. 図1B参照。

・103B. 人102によって発せられる音は、スマート・オーディオ装置(105)のマイクロフォンのほうがユーザーにより近いので、スマート・オーディオ装置(105)によってよりよく捕捉される。

【0113】

図1Cでは、新しいスマート・オーディオ装置105が何らかの仕方(その例を本明細書に記載される)で検出され、装置105内のマイクロフォンがスマート・オーディオ装置(101)上で動作していたアプリケーションをサポートするように機能することができる。図1Cの新しいスマート・オーディオ装置105は、ある意味では、(本開示のいくつかの側面に従って)図1Cの装置101と協調させられるまたはオーケストレーションされ、図1Cに示されている状況については、スマート・オーディオ装置105の人102への近接性が検出され、日和見的に、よりよいマイクロフォンとして認識される。図1Cでは、オーディオ104は、相対的により遠いスピーカー電話101からくるが、電話101に送信されるオーディオ103Bは、ローカルなスマート・オーディオ装置105によって捕捉される。いくつかの実施形態では、異なるスマート・オーディオ装置の構成要素のこの日和見的な使用は、電話101および/または電話通話を提供するために使用されるアプリケーションがスマート・オーディオ装置105のルーティングの複雑さおよび能力を知ることなく、可能にされる。むしろ、そのようないくつかの例では、そのような能力の発見、ルーティング、および利用のために階層的システムが実装されてもよい。

30

【0114】

後に、抽象的な連続階層オーディオ・セッション・マネージャ(CHASM)の概念をより詳細に記述する。そのいくつかの実装は、アプリケーションが装置の管理、装置接続性、同時装置使用、および/または装置の平準化〔レベリング〕およびチューニングの完全な詳細を知る必要なしに、オーディオ機能をアプリケーションに提供することができる。ある意味で、このアプローチは、アプリケーションを正常に実行している(そして少なくとも1つのスピーカーと少なくとも1つのマイクロフォンを有する)装置が、オーディオ体験の制御を放棄していることになる。しかしながら、室内のスピーカーや、より大切なこ

40

50

とにマイクロフォンの数が人の数をはるかに上回る状況では、オーディオの多くの問題の解決は、関連する人物に最も近い場所に装置を位置特定するステップを含むことがわかる。その装置は、そのようなアプリケーションのために通常使用される装置ではないかもしれない。

【0115】

オーディオ・トランスデューサー（スピーカーとマイクロフォン）についての1つの考え方は、オーディオが人の口からアプリケーションに来るためのルート中の1ステップと、アプリケーションから人の耳までのルート中の戻りステップを実装できるというものである。この意味で、ユーザーからオーディオを送達または捕捉する必要のあるいかなるアプリケーションも、諸装置の日和見的な利点を取り、オーディオを出力するまたは入力オーディオを得るために任意の装置にあるオーディオ・サブシステムとインターフェースをもつことにより、改善できる（または、少なくとも、悪化させられない）ことができることがわかる。そのような決定およびルーティングは、いくつかの例においては、装置およびユーザーが動いたり、あるいは利用可能になったり、あるいはシステムから除去される際に、連続的な仕方で行われてもよい。この点において、本明細書に開示されている連続的階層オーディオ・セッション・マネージャ（CHASM）は有用である。いくつかの実装では、発見可能な日和見的にオーケストレーション〔組織化〕された分散オーディオ装置サブシステム（DOODAD）が、集合体としてCHASMに含められることができ、あるいはDOODADは集合体としてCHASMと共に使用できる。

【0116】

いくつかの開示された実施形態は、人および場所へ／からオーディオをルーティングするように設計された集合的オーディオ・システムの概念を実装する。これは、一般に、ある装置からオーディオを入力および出力し、次いで諸装置を集合的に管理することに関係する従来の「装置中心的」設計からの逸脱である。

【0117】

図2A～図2Dを参照して、次に、本開示のいくつかの例示的な実施形態を説明する。通信機能をもつ装置から始める。この場合、たとえば、ドアベルのオーディオ・インターコム装置を考える。この装置は、アクティブ化されると、ローカル装置から何らかのリモートユーザーへの全二重オーディオ・リンクを作成するローカル・アプリケーションを起動する。このモードにおける装置の基本的な機能は、スピーカー、マイクロフォンを管理し、本明細書で「リンク」と呼ぶことができる双方向ネットワーク・ストリームを通じてスピーカー信号およびマイクロフォン信号を中継することである。

【0118】

図2Aは、従来のシステムのブロック図である。この例では、スマート・オーディオ装置200Aの動作は、アプリケーション（205A）がメディア（212）および制御情報（213）をメディア・エンジン（201）へ／から送ることを含む。メディア・エンジン201およびアプリケーション205Aは、いずれも、スマート・オーディオ装置200Aの制御システムによって実装されうる。この例によれば、メディア・エンジン201は、オーディオ入力（203）および出力（204）の管理を受け持ち、信号処理および他のリアルタイム・オーディオ・タスクのために構成されてもよい。アプリケーション205Aはまた、他のネットワーク入出力接続性（210）を有してもよい。

【0119】

図2Aでは、ラベル付けされた要素は以下の通りである：

- ・200A. 特定目的のスマート・オーディオ装置。
- ・201. アプリケーション205Aから入ってくるリアルタイム・オーディオ・メディア・ストリームの管理と、マイクロフォン入力およびスピーカー出力についての信号処理を受け持つメディア・エンジン。信号処理の例は、音響エコーキャンセル、自動利得制御、ウェイクワード検出、リミッタ、ノイズ抑制、動的ビームフォーミング、音声認識、損失のあるフォーマットへのエンコード／デコード、音声活動検出および他の分類器などを挙げる。この文脈において、語句「リアルタイム」は、たとえば、オーディオのブロックの

処理が、装置200Aによって実装されるアナログ - デジタル変換器 (ADC) からのオーディオのブロックをサンプリングするのに要する時間内に完了されるべきことをいう。たとえば、ある特定の実装では、オーディオのブロックは、10~20msの長さであってもよく、48000サンプル/秒でサンプリングされた480~960個の連続するサンプルを含んでいてもよい。

・203. マイクロフォン入力。複数のADCによってメディア・エンジン (201) にインターフェースされた音響情報を感知することができる一つまたは複数のマイクロフォンからの入力。

・204. スピーカー出力。複数のデジタル - アナログ変換器 (DAC) および／または増幅器によってメディア・エンジン (201) にインターフェースされる、音響エネルギーを再生することができる一つまたは複数のスピーカーからの入力。

・205A. 装置200A上で動作するアプリケーション (「アプリ」)。ネットワークから来るまたはネットワークに行くメディアを処理し、メディア・エンジン201との間でメディア・ストリームを送受信することを受け持つ。この例では、アプリ205Aはまた、メディア・エンジンによって送受信される制御情報を管理する。アプリ205Aの例は、以下を含む：

○インターネットに接続し、マイクロフォンからウェブ・サービスにパケットをストリーミングするウェブカメラの制御ロジック；

○電話番号をダイヤルしたり、連絡先リストを閲覧したり、音量を変更したり、通話を開始および終了したりできるようにするタッチスクリーンを介して、ユーザーとインターフェースをとる会議用電話；

○楽曲のライブラリからの音楽の再生を可能にする、スマートスピーカー内の音声駆動型アプリケーション。

・210. 装置200Aをネットワークに (たとえば、Wi-Fi、イーサネット、または4G/5Gセルラー無線を介してインターネットに) 接続する任意的なネットワーク接続。ネットワークはストリーミング・メディア・トラフィックを搬送してもよい。

・212. メディア・エンジン (201) との間でストリーミングされるメディア。たとえば、特定目的のテレビ会議装置内のアプリ205Aは、ネットワーク (210) からリアルタイムトランスポートプロトコル (RTP) パケットを受信し、ヘッダを除去し、G.711ペイロードを処理および再生のためにメディア・エンジン201に転送することができる。いくつかのそのような例では、アプリ205Aは、メディア・エンジン201からG.711ストリームを受領し、ネットワーク (210) を通じた上流への送達のためにRTPパケットをパッキングすることを受け持つことができる。

・213. メディア・エンジン (201) を制御するためにアプリ205Aへ／から送られる制御信号。たとえば、ユーザーがユーザーインターフェース上のボリュームアップボタンを押すと、アプリ205Aは、再生信号 (204) を増幅するために、制御情報をメディア・エンジン201に送信する。特定目的の装置200Aでは、メディア・エンジン (201) の制御は、外部からメディア・エンジンを制御する能力なしで、ローカル・アプリ (205A) からのみ行われる。

【0120】

図2Bは、図2Aに示されている装置の修正バージョンの例を示す。オーディオ装置200Bは、第2のアプリケーションを実行することができる。たとえば、第2のアプリケーションは、セキュリティシステムのためにドアベル装置からオーディオを連続的にストリーミングするものであってもよい。この場合、同じオーディオ・サブシステム (図2Aを参照して説明される) が、単一方向のみのネットワーク・ストリームを制御する第2のアプリケーションと共に使用できる。

【0121】

この例では、図2Bは、特定目的オーディオ・セッション・マネージャ (Specific Purpose Audio Session Manager, SPASM) を介して情報を (装置200Bの) メディア・エンジン202に送信する2つのアプリケーションまたは「アプリ」 (アプリ205Bおよび

10

20

30

40

50

アプリ206)をホスティングすることができる特定目的のスマート・オーディオ装置(200B)のブロック図である。SPASMをアプリ205Bおよび206へのインターフェースとすると、ネットワークメディアは、今や、メディア・エンジン202に直接流れることができる。ここで、メディア210は、第1のアプリ(206)へ／からのメディアであり、メディア211は、第2のアプリ(205B)へ／からのメディアである。

【0122】

ここで、SPASM(または特定目的オーディオ・セッション・マネージャ)という用語を使用して、装置が提供するように製造された単一タイプの機能性のためにオーディオ・チェーンを実装するように構成された(装置の)要素またはサブシステムを表す。SPASMは、装置の動作モードの変更を実施するために(たとえば、オーディオ・システム全体を解体することによって)再構成される必要があるかもしれない。たとえば、ほとんどのラップトップのオーディオは、SPASMとして、またはSPASMを使用して実装され、SPASMは、特定の機能のために任意の所望の単一目的のオーディオ・チェーンを実装するように構成される(そして再構成可能である)。

【0123】

図2Bにおいて、ラベル付けされた要素は以下の通りである:

- ・200B. 2つのアプリ(アプリ205Bおよびアプリ206)をホスティングする特定目的オーディオ・セッション・マネージャ(SPASM)207Bを備えたスマート・オーディオ装置。
- ・202~204. 図2A参照。
- ・205B,206. ローカル装置200Bで動作するアプリケーション。
- ・207B. オーディオ処理を管理し、装置200Bのメディア・エンジン(202)の能力を公開することを受け持つ特定目的オーディオ・セッション・マネージャ。各アプリ(206または205B)とSPASM 207Bとの間の描画は、異なるアプリがどのようにして異なるオーディオ機能を使用することを望むかを示す(たとえば、アプリは異なるサンプルレートまたは異なる数の入力および出力を必要とする可能性がある)。これらのオーディオ機能はすべて、SPASMが異なるアプリのために公開および管理する。SPASMの限界は、SPASMが特定目的のために設計され、それが知っている動作を実行できるだけであるということである。
- ・210. 第1のアプリ(205B)のためにネットワークへ／からストリーミングされるメディア情報。SPASM(207B)は、メディアのフローが、メディア・エンジン(202)に直接ストリーミングされることを許容する。
- ・211. 第2のアプリ(206)のためにネットワークにストリームされるメディア情報。この例では、アプリ206は受領すべきメディア・ストリームを持たない。
- ・214. メディア・エンジンの機能を管理するためにSPASM(207B)およびメディア・エンジン(202)へ／から送られる制御情報。
- ・215,216. アプリ(205B,206)およびSPASM(207B)へ／から送られる制御情報。

【0124】

図2Bの装置200Bの別個のサブシステムとしてSPASM 207Bを含めることは、設計において人工的なステップを実装するように思えるかもしれない。実際、単一目的のオーディオ装置設計の観点から不必要な作業と思われるものが含まれている。CHASMの価値の多く(後述)は、ネットワーク効果として有効にされており、これは(ある意味で)ネットワーク内のノード数の2乗としてスケーリングされる。しかしながら、スマート・オーディオ装置にSPASM(たとえば、図2BのSPASM 207B)を含めることは、以下を含む利点と価値を有する:

- ・SPASM 207Bによる制御の抽象化により、複数のアプリケーションを同じ装置上で実行することがより容易になる;
- ・SPASM 207Bは、オーディオ装置に密接に結合されており、ネットワーク・ストリーム接続性を直接SPASM 207Bにもたらすことによって、ネットワーク上のオーディオ・データと物理的な入出力サウンドとの間のレイテンシーを低減する。たとえば、SPASM

207Bは、装置ドライバ/データリンク層により近い、スマート・オーディオ装置200B内の下位層（たとえば、下位のOSI層またはTCP/IP層）に、または物理的ハードウェア層に存在してもよい。SPASM 207Bが上位層で実装された、たとえば、装置オペレーティングシステム内で実行されるアプリケーションとして実装されたとしたら、そのような実装は、オーディオ・データが下位層からオペレーティングシステムを通じてアプリケーション層に戻ってコピーされる必要があるため、レイテンシー・ペナルティを受ける可能性が高い。そのような実装の潜在的により悪い特徴は、レイテンシーが可変であったり予測不可能であったりしうることである；

・この設計は、これで、アプリケーション・レベルに先立って、より低いオーディオ・レベルでより大きな相互接続性のための準備ができています。

【0125】

オペレーティングシステムがSPASMを実行するスマート・オーディオ装置では、いくつかの例では、多くのアプリがスマート・オーディオ装置のスピーカーおよびマイクロフォンへの共有アクセスを得ることができる。オーディオ・ストリームを送受信する必要のないSPASMを導入することによって、いくつかの例によれば、メディア・エンジンが制御ロジックから分離されているので、メディア・エンジンは、非常に低いレイテンシーのために最適化されうる。SPASMを有する装置は、アプリケーションが追加的なメディアのストリーム（たとえば、図2Bのメディア・ストリーム210および211）を確立することを可能にする。この利点は、メディア・エンジンの機能をSPASMの制御ロジックから分離することから帰結する。この構成は、メディア・エンジンがアプリケーション固有でありスタンドアローンである図1Aおよび図2Aに示されている状況とは対照的である。これらの例では、装置は、たとえば図2Bに示すように、SPASMを含むことによって可能にされる追加的な装置への/からの追加的な低レイテンシー接続性を有するようには設計されていない。そのようないくつかの例では、図1Aおよび2Aに示されるような装置がスタンドアローンであるように設計される場合、オーケストレーションを提供するために、たとえば単にアプリケーション205Aを更新することは、実際的ではない可能性がある。しかしながら、いくつかの例では、装置200Bは、オーケストレーション準備ができていように設計される。

【0126】

図2Cを参照して、次に、（スマート・オーディオ装置の）SPASM自体が宣伝し、制御を許容する、いくつかの開示された実施形態の側面を説明する。これは、ネットワーク内の他の装置および/またはシステムが、装置のオーディオ能力をよりよく理解するためにプロトコルを利用することができるようにすることができ、また、セキュリティおよび使用能力の観点から適用可能な場合には、直接その装置へのオーディオ・ストリームの接続が、スピーカーから再生されるか、またはマイクロフォンから供給されることができるようにすることができる。この場合、装置から連続的にオーディオをストリーミングする能力を設定するための第2のアプリケーションは、メディア・エンジン（たとえば、202）を制御して、たとえば、上述のサーベillanceストリーム（たとえば、211）をストリームアウトするために、ローカルに実行されるアプリケーションをもつ必要がないことがわかる。

【0127】

図2Cは、1つの開示された実装のブロック図である。この例では、2つ以上のアプリのうちの少なくとも1つのアプリ（たとえば、図2Cのアプリ205B）は、スマート・オーディオ装置200C以外の装置によって、たとえば、クラウドベースのサービスを実装する一つまたは複数のサーバーによって、スマート・オーディオ装置200Cが存在するオーディオ環境内の別の装置によって、などで実装される。よって、別のコントローラ（この例では、図2CのCHASM 208C）が、オーディオ体験を管理するために必要とされる。この実装では、CHASM 208Cは、オーディオ機能を有するリモート・アプリとスマート・オーディオ装置（200C）との間のギャップをブリッジするコントローラである。さまざまな実施形態では、CHASM（たとえば、図2CのCHASM 208C）は、装置として、または装

10

20

30

40

50

置のサブシステム（たとえば、ソフトウェアで実装される）として実装されてもよく、ここで、CHASMである、またはCHASMを含む装置は、一つまたは複数の（たとえば、多くの）スマート・オーディオ装置とは異なる。しかしながら、いくつかの実装では、CHASMは、潜在的にオーディオ環境の一つまたは複数の装置によって実行されうるソフトウェアを介して実装されてもよい。いくつかのそのような実装では、CHASMは、潜在的にオーディオ環境の一つまたは複数のスマート・オーディオ装置によって実行されうるソフトウェアを介して実装されてもよい。図2Cにおいて、CHASM 208Cは、装置200Cの音声入力（203）および出力（204）を制御するメディア・エンジン（202）へのアクセスを得るために、装置200CのSPASM（すなわち、図2CのSPASM 207C）と協調する。

【0128】

本明細書では、「CHASM」という用語を使用して、いくつかの（たとえばある集合の）装置（これはスマート・オーディオ装置を含みうるが、これに限定されない）が利用可能となれるマネージャ（たとえば、オーディオ・セッション・マネージャ、たとえば進行中のオーディオ・セッション・マネージャである、またはそれを実装する装置）を表す。いくつかの実装によれば、CHASMは、少なくとも1つのソフトウェア・アプリケーションについて、ルーティングおよび信号処理を連続的に（少なくとも、本明細書で「ルート」と称するものが実装される時間の間）調整することができる。アプリケーションは、特定の実装に依存して、オーディオ環境のいずれの装置に実装されてもよく、実装されなくてもよい。言い換えると、CHASMは、オーディオ環境内の一つまたは複数の装置によって実行されている一つまたは複数のソフトウェア・アプリケーション、および／またはオーディオ環境の外部の一つまたは複数の装置によって実行されている一つまたは複数のソフトウェア・アプリケーションのために、オーディオ・セッション・マネージャ（本明細書では「セッション・マネージャ」とも呼ばれる）を実装するか、またはオーディオ・セッション・マネージャとして構成されてもよい。ソフトウェア・アプリケーションは、本明細書では「アプリ」と呼ばれることがある。

【0129】

いくつかの例では、CHASMを使用した結果として、オーディオ装置は、そのオーディオ装置の作成者および／または製造者によって想定されていなかった目的のために使用されることになりうる。たとえば、スマート・オーディオ装置（少なくとも1つのスピーカーおよびマイクロフォンを含む）は、スマート・オーディオ装置がオーディオ環境内の一つまたは複数の他のオーディオ装置にスピーカー・フィード信号および／またはマイクロフォン信号を提供するモードに入ることができる。なぜなら、アプリ（たとえば、スマート・オーディオ装置とは異なる他の装置に実装される）は、オーディオ環境の2つ以上のオーディオ装置からのスピーカーおよび／またはマイクロフォンを含むことができるすべての利用可能なスピーカーおよび／またはマイクロフォン（または、CHASMによって選択される利用可能なスピーカーおよび／またはマイクロフォンのグループ）を見つけ、使用するように、CHASM（スマート・オーディオ装置に結合される）に要求するからである。このような多くの実装では、アプリケーションは装置、スピーカーおよび／またはマイクロフォンを選択する必要がない。なぜなら、CHASMはこの機能を提供するからである。いくつかの例では、アプリケーションは、どの特定のオーディオ装置がアプリケーションによってCHASMに提供されたコマンドを実装するのに関与しているかを知らなくてもよい（たとえば、CHASMがアプリケーションに示さなくてもよい）。

【0130】

図2Cでは、ラベル付けされた要素は以下の通りである：

- ・200C. CHASM（208C）を通じてローカルなアプリ（206）およびリモートのアプリ（205B）を実行するセッション・マネージャである特定目的のスマート・オーディオ装置。
- ・202～204. 図2A参照。
- ・205B. たとえば、CHASM 208Cがインターネットを介して通信するように構成されているサーバー上で、またはオーディオ環境の他の装置（装置200Cとは別の装置、たと

例えば、携帯電話のような別のスマート・オーディオ装置) 上で、装置200Cからリモートで実行されるアプリ。アプリ205Bは、いくつかの例において、第1の装置上に実装されてもよく、CHASM 208Cは第2の装置上に実装されてもよく、ここで、第1の装置および第2の装置の両方は、装置200Cとは異なる。

・206. 装置200C上でローカルに実行されるアプリ。

・207C. SPASM。これは、メディア・エンジン(202)とのインターフェースをもつことに加えて、CHASM 208Cからの制御入力を管理することができる。

・208C. 連続階層オーディオ・セッション・マネージャ(CHASM)。これは、アプリ(205B)が装置200Cのメディア・エンジン(202)、入力(203)、および出力(204)のオーディオ機能を利用できるようにする。この例では、CHASM 208Cは、SPASM 207Cからメディア・エンジン(202)の少なくとも部分的な制御を得ることによって、SPASM(207C)を介してそうするように構成される。

・210~211. 図2C参照。

・217. メディア・エンジンの機能を管理するためにSPASM(207B)およびメディア・エンジン(202)へ/から送られる制御情報。

・218. ローカル・アプリ206を実装するために、ローカル・アプリ206とSPASM(207C)へ/からの制御情報。いくつかの実装では、そのような制御情報は、本明細書に開示されているオーケストレーションの言語のようなオーケストレーションの言語に従うことができる。

・219. メディア・エンジン202の機能性を制御するためのCHASM(208C)およびSPASM(207C)へ/からの制御情報。そのような制御情報は、いくつかの場合には、制御情報217と同じか、または同様であってもよい。しかしながら、いくつかの実装では、いくつかの例では装置固有の詳細がSPASM 207Cに委任されうるので、制御情報219は、より低いレベルの詳細を有することがある。

・220. アプリ(205B)とCHASM(208C)の間の制御情報。いくつかの例では、この制御情報は、本明細書において「オーケストレーションの言語」と称されるものであってもよい。

【0131】

制御情報217は、たとえば、SPASM 207Cからメディア・エンジン202への制御信号を含むことができ、これらの制御信号は、出力スピーカー・フィードの出力レベル、たとえば、デシベルで指定された利得調整、または線形スカラー値を調整する効果を有する。出力スピーカー・フィードに適用される等化曲線の変更など。いくつかの例では、SPASM 207Cからメディア・エンジン202への制御情報217は、出力スピーカー・フィードに適用される等化曲線を変化させる効果を有する制御信号を含むことができる。かかる変化は、たとえば、パラメトリックに(基本フィルタ段の直列組み合わせとして)記述されたまたは個別の周波数における利得値の列挙として表にされた新しい等化曲線を提供することによる。いくつかの例では、SPASM 207Cからメディア・エンジン202への制御情報217は、複数のオーディオ源フィードを出力ラウドスピーカーフィードにレンダリングするアップミックスまたはダウンミックス・プロセスを変更する効果を有する制御信号を含むことができる。かかる変更は、たとえば、ソースフィードをラウドスピーカーフィードに組み合わせるために使用される混合行列を提供することによる。いくつかの例では、SPASM 207Cからメディア・エンジン202への制御情報217は、たとえばオーディオ・コンテンツのダイナミックレンジを変更するなど、出力スピーカー・フィードに適用されるダイナミクス処理を変更する効果を有する制御信号を含むことができる。

【0132】

いくつかの例では、SPASM 207Cからメディア・エンジン202への制御情報217は、メディア・エンジンに提供されている一組のメディア・ストリームに対する変更を示すことができる。いくつかの例では、SPASM 207Cからメディア・エンジン202への制御情報217は、他のメディア・エンジンまたはメディア・コンテンツの他のソース(たとえば、クラウドベースのストリーミング・サービス)でメディア・ストリームを確立または終

了する必要性を示すことができる。

【0133】

いくつかの例では、制御情報217は、ウェイクワード検出情報のようなメディア・エンジン202からSPASM 207Cへの制御信号を含んでいてもよい。そのようなウェイクワード検出情報は、いくつかの場合には、ウェイクワード信頼値、または、確からしいウェイクワードが検出されたことを示すメッセージを含んでいてもよい。いくつかの例において、ウェイクワード信頼値は、時間期間毎に1回（たとえば、100ms毎に1回、150ms毎に1回、200ms毎に1回など）送信されうる。

【0134】

いくつかの例では、メディア・エンジン202からの制御情報217は、SPASM、CHASMまたは他の装置（たとえば、クラウドベースのサービスの装置）が、どのコマンドが発せられているかを決定するための復号（たとえば、ビット復号）を実行することを可能にする音声認識電話の確率を含んでいてもよい。いくつかの例では、メディア・エンジン202からの制御情報217は、SPLメーターからの音圧レベル（SPL）情報を含むことができる。いくつかのそのような例によれば、SPL情報は、ある時間間隔で、たとえば、毎秒1回、0.5秒毎に1回、N秒毎またはミリ秒毎に1回などで送信されうる。いくつかのそのような例において、CHASMは、複数の装置を横断してSPLメーターの読みに相関があるかどうかを決定するように、たとえば、装置が同じ部屋にあるかどうかを決定するために、および／または同じ音を検出するために、構成されてもよい。いくつかの例によれば、メディア・エンジン202からの制御情報217は、メディア・エンジンに利用可能なメディア・ストリームとして存在するマイクロフォン・フィードから導出される情報、たとえば、バックグラウンドノイズの推定、到来方向（DOA）情報の推定、音声活動検出（voice activity detection）を通じた音発話の存在の指示、現在のエコーキャンセル性能などを含むことができる。いくつかのそのような例では、DOA情報は、オーディオ環境においてオーディオ装置の音響マッピングを実行し、いくつかのそのような例では、オーディオ環境の音響マップを作成するように構成された上流のCHASM（または別の装置）に提供されてもよい。いくつかのそのような例において、DOA情報は、ウェイクワード検出イベントと関連付けられてもよい。いくつかのそのような実装では、DOA情報は、音響マッピングを実行して、ウェイクワードを発するユーザーを位置特定するように構成された上流のCHASM（または他の装置）に提供されてもよい。

【0135】

いくつかの例では、メディア・エンジン202からの制御情報217は、ステータス情報、たとえば、どのアクティブ・メディア・ストリームが利用可能であるかに関する情報、線形時間メディア・ストリーム内の時間位置（たとえば、テレビ・プログラム、映画、ストリーミングされるビデオ）、現在のネットワーク性能に関連する情報、たとえばアクティブ・メディア・ストリームに関連するレイテンシー、信頼性情報（たとえば、パケット損失統計）などを含んでいてもよい。

【0136】

図2Cの装置200Cの設計は、本開示の側面に従ってさまざまな仕方で拡張されうる。図2CのSPASM 207Cの機能は、ローカル・メディア・エンジン202を制御するための一組のフックまたは機能を実装することであることがわかるであろう。よって、装置200Cを、オーディオ装置を接続し、ストリームをネットワークし、信号処理を実行し、装置200Cのローカル・アプリケーションおよびオーディオ・セッション・マネージャの両方からの構成コマンドに応答することができるメディア・エンジンに、より近い（たとえば、装置200Aまたは装置200Bの機能よりも近い）ものとみなすことができる。この場合、オーディオ・セッション・マネージャを支援するために、装置が自身に関する情報（たとえば、メモリ装置に格納され、オーディオ・セッション・マネージャにとって利用可能な情報）を持っていることが重要である。この情報の単純な例は、スピーカーの数、スピーカーの能力、ダイナミクス処理情報、マイクロフォンの数、マイクロフォンの配置および感度に関する情報などを含む。

10

20

30

40

50

【0137】

図2Dは、CHASMと対話する複数のアプリケーションの例を示す。図2Dに示される例では、スマート・オーディオ装置200Dに関わる機能を提供するために、スマート・オーディオ装置にローカルなアプリ（たとえば、スマート・オーディオ装置200Dのメモリに記憶されるアプリ206）を含むすべてのアプリが、CHASM 208Dとインターフェースする必要がある。この例では、CHASM 208Dがインターフェースすることを引き受けたので、スマート・オーディオ装置200Dは、その属性207Dを広告するか、またはその属性をCHASM 208Dに利用可能にするだけでよく、SPASMはもはや必要ではない。これにより、CHASM 208Dは、アプリのためにオーディオ・セッションなどの体験をオーケストレーションするための主要コントローラとなる。

10

【0138】

図2Dでは、ラベル付けされた要素は以下の通りである：

- ・200D. ローカル・アプリ206を実装するスマート・オーディオ装置。CHASM 208Bはローカルアプリ（206）とリモートアプリ（205B）を実行する。
- ・202～204. 図2A参照。
- ・205B. スマート・オーディオ装置200Dからリモートに（言い換えると、スマート・オーディオ装置200D以外の装置で）実行される（この例では、CHASM 208Dからもリモートで実行される）アプリケーション。いくつかの例では、アプリケーション205Bは、CHASM 208Dがたとえばインターネットまたはローカルネットワークを介して通信するように構成される装置によって実行されてもよい。いくつかの例では、アプリケーション205Bは、別のスマート・オーディオ装置など、オーディオ環境の別の装置に格納されてもよい。いくつかの実装によれば、アプリケーション205Bは、セルラー電話のような、オーディオ環境中に／から移送されてもよいモバイル装置に格納されうる。
- ・206. スマート・オーディオ装置200D上でローカルに実行されるが、それについて制御情報（223）がCHASM 208Dに、および／またはCHASM 208Dから送られるアプリケーション。
- ・207D. 属性記述子。CHASM 208Dがメディア・エンジン202を管理する務めを受けると、スマート・オーディオ装置200Dは、単純な属性記述子をSPASMに置き換えることができる。この例では、記述子207Dは、入力および出力の数、可能なサンプルレート、および信号処理コンポーネントなど、メディア・エンジン202の能力をCHASMに対して示す。いくつかの例では、記述子207Dは、スマート・オーディオ装置200Dの一つまたは複数のラウドスピーカーに対応するデータ、たとえば、一つまたは複数のラウドスピーカーのタイプ、サイズおよび数を示すデータ、一つまたは複数のラウドスピーカーの能力に対応するデータ、メディア・エンジン202が、一つまたは複数のラウドスピーカーによってオーディオ・データが再生される前にオーディオ・データに適用するダイナミクス処理に関連するデータなどを示すことができる。いくつかの例では、記述子207Dは、メディア・エンジン202（または、より一般的には、スマート・オーディオ装置200Dの制御システム）が、スマート・オーディオ装置200Dおよび／またはオーディオ環境の他のオーディオ装置によって再生されるべきオーディオ・データをレンダリングするなどの、環境におけるオーディオ装置の協調に関連する機能を提供するように構成されているかどうか、CHASM 208Dの機能性を（たとえば、該装置のメモリに記憶されているCHASMソフトウェアに従って）現在提供している装置がオフにされるか、または他の仕方でも機能するのを停止した場合に、スマート・オーディオ装置200Dの制御システムが、CHASM 208Dの機能性を実装できるかどうか、などを示してもよい。
- ・208D. この例では、CHASM 208Dは、メディア・エンジン202と対話するための、すべてのアプリ（ローカルかリモートかにかかわらず）のためのゲートウェイとして機能する。ローカルアプリ（たとえば、アプリ206）でさえ、CHASM 208Dを介してローカル・メディア・エンジン202へのアクセスを得る。いくつかの例では、CHASM 208Dは、たとえば、無線ルータ、スマートスピーカー等に格納されたCHASMソフトウェアを介して、オーディオ環境の単一の装置においてのみ実装することができる。しかしながら、

20

30

40

50

いくつかの実装では、オーディオ環境の複数の装置が、CHASM機能の少なくともいくつかの側面を実装するように構成されてもよい。いくつかの例では、オーディオ環境の一つまたは複数のスマート・オーディオ装置など、オーディオ環境内の一つまたは複数の他の装置の制御システムは、CHASM 208Dの機能性を現在提供している装置がオフにされる、または他の仕方では機能を停止する場合に、CHASM 208Dの機能性を実装することができてよい。

・210～211. 図2C参照。

・221. CHASM 208Dとメディア・エンジン202との間で送られる（たとえば、送受信される）制御情報。

・222. メディア・エンジン202の能力を示すためにCHASMに送信される属性記述子207Dからのデータ。

・223. ローカル・アプリ206とCHASM 208との間で送られる制御情報。

・224. リモート・アプリ205BとCHASM 208Dとの間で送られる制御情報。

【0139】

次に、追加的な実施形態について説明する。いくつかのそのような実施形態を実装するために、初期に、単一の装置（通信装置など）が、特定目的のために設計され、コード化される。そのような装置の一例は、図1Cのスマート・オーディオ装置101であり、図3Cに示すように実装することができる。背景として、図1Aおよび図1Bの装置101の実装も説明される。

【0140】

図3Aは、一例による図1Aの装置101の詳細を示すブロック図である。ユーザーの声103がマイクロフォン303によって捕捉されると、ローカル・アプリ308Aは、装置101のネットワーク・インターフェースを介して受領したネットワーク・ストリーム317を管理し、メディア・ストリーム341を管理し、制御信号340をメディア・エンジン301Aに提供することを受け持つ。

【0141】

図3Aでは、ラベル付けされた要素は以下の通りである：

・101,103～104. 図1A参照。

・301A. アプリ308Aから入ってくるリアルタイムのオーディオメディア・ストリームを管理することを受け持つメディア・エンジン。

・303. マイクロフォン。

・304. スピーカー。

・308A. ローカル・アプリ。

・317. ネットワークへ／からのメディア・ストリーム。

・340. アプリ308Aおよびメディア・エンジン301Aへ／から送られる制御情報。

・341. アプリ308Aへ／から送られるメディア・ストリーム。

【0142】

図3Bは、一例による、図1Bの実装の詳細を示す。この例では、図1Bの装置105および図1Bの装置101の実装が示される。この場合、両方の装置は、SPASMの抽象化によって制御される汎用のまたは柔軟なメディア・エンジンが存在するという意味で、「オーケストレーションの準備ができています（orchestration ready）」ように設計される。

【0143】

図3Bでは、第2の装置105の出力106は、第1の装置101の出力104とは無関係であり、装置101のマイクロフォン303への入力103は、装置105の出力106を潜在的に捕捉する。この例では、装置105および101がオーケストレーションされた仕方では動作する方法はない。

【0144】

図3Bにおいて、ラベル付き要素は以下の通りである：

・101,103～106. 図1B参照。

・301,303-304. 図3A参照。

- ・ 302. 装置105のメディア・エンジン。
 - ・ 305. 装置105のマイクロフォン。
 - ・ 306. 装置105のラウドスピーカー。
 - ・ 308B. 装置101のローカル・アプリ。
 - ・ 312B. 装置101のためのSPASM。
 - ・ 314B. 装置105のためのSPASM。
 - ・ 317. ネットワークへ／からのメディア・ストリーム。
 - ・ 320. 装置105のためのローカル・アプリ。
 - ・ 321. アプリ308BとSPASM 312Bとの間で送られる制御情報。
 - ・ 322. SPASM 312Bとメディア・エンジン301との間で送られる制御情報。
 - ・ 323. アプリ320とSPASM 314Bとの間で（へ／から）送られる制御情報。
 - ・ 324. SPASM 314Bとメディア・エンジン302との間で（へ／から）送られる制御情報。
 - ・ 325. ネットワークからメディア・エンジン302へのメディア・ストリーム。
- 【0145】

図3Cは、オーディオ環境の2つのオーディオ装置をオーケストレーションするCHASMの例を示すブロック図である。前述の議論に基づいて、図3Bのシステムが使用される状況は、CHASMが装置101および装置105の修正バージョン上で動作するか、または別の装置上で動作する場合のCHASM（たとえば、図3Cの実施形態におけるような）によって、よりよく管理されることが理解されるべきである。CHASM 307を使用すると、いくつかの例では、電話機101上のアプリケーション308は、そのオーディオ装置101の直接制御をあきらめ、すべてのオーディオ制御を、CHASM 307に委ねる。いくつかのそのような例によれば、マイクロフォン305からの信号は、マイクロフォン303からの信号よりも少ないエコーを含むことがある。CHASM 307は、いくつかの例において、マイクロフォン305が人102により近いとの推定に基づいて、マイクロフォン305からの信号がマイクロフォン303からの信号よりも少ないエコーを含むと推論することができる。いくつかのそのような場合には、CHASMは、ネットワーク・ストリームとして装置105上のマイクロフォン305からの未加工のマイクロフォン信号または処理済みのマイクロフォン信号を電話装置101にルーティングすることによって利点を得ることができ、これらのマイクロフォン信号は、人102にとってより良好な音声通信経験を達成するために、ローカルマイクロフォン303からの信号よりも優先して使用されてもよい。

【0146】

いくつかの実装によれば、CHASM 307は、たとえば、人102の位置をモニターすること、装置101および／または装置105の位置をモニターすること等により、これがいまだ最良の構成であることを保証することができる。いくつかの例において、CHASM 307は、これがいまだ最良の構成であることを保証することができる。そのようないくつかの例によれば、CHASM 307は、低レート（たとえば、低ビットレート）データおよび／またはメタデータの交換を介して、これがいまだ最良の構成であることを保証することができる。たとえば、少量の情報が装置間で共有されるだけで、人102の位置を追跡できる。もし低ビットレートで装置間で情報が交換されているなら、帯域幅の制限についての考慮はそれほど問題ではなくなるかもしれない。装置間で交換されうる低ビットレート情報の例は、マイクロフォン信号から導出される情報を含むが、それに限定されない。これはたとえば、「follow me（ついてこい）」実装を参照して後述される。どの装置のマイクロフォンがより高い発話対エコー比を有するかを決定する際に有用であるかもしれない低ビットレート情報の一例は、ある時間期間、たとえば最後の1秒の間に、オーディオ環境内の複数のオーディオ装置のそれぞれのローカル・ラウドスピーカーによって放射された音によって引き起こされたSPLの推定値である。スピーカーからより多くのエネルギーを放射するオーディオ装置は、スピーカーによって引き起こされるエコーにかぶさって、オーディオ環境内の他の音を捕捉するのが、より少ない可能性が高い。どの装置のマイクロフォンがより高い発話対エコー比を有するかを決定する際に有用でありうる低ビットレート情

10

20

30

40

50

報の別の例は、各装置の音響エコー・キャンセラのエコー予測におけるエネルギーの量である。予測されたエコーエネルギーが大きいことは、オーディオ装置のマイクロフォンがエコーによって圧倒される可能性が高いことを示す。いくつかのそのような場合には、音響エコー・キャンセラがキャンセルできないいくらかのエコーがあるかもしれない（音響エコー・キャンセラがこの時点ですでに収束していると仮定する）。いくつかの例において、マイクロフォン305に関する問題、またはマイクロフォン305が存在しないことの情報に何かが提供するとしたら、CHASM 307は、ローカルマイクロフォン303からのマイクロフォン信号を使用して再開するよう装置101を制御することが、連続的に準備できている。

【0147】

図3Cは、装置101上で動作するアプリ308を装置105と協調させるために使用される基礎となるシステムの例を示す。この例では、CHASM 307は、ユーザーの声103Bを装置105のマイクロフォン305に捕捉させ、捕捉されたオーディオ316を装置101のメディア・エンジン301において使用させ、一方、ラウドスピーカー出力104は、第1の装置101から来る。これにより、装置101および105を横切ってユーザー102のための体験をオーケストレーションする。

【0148】

図3Cでは、ラベル付けされた要素は以下の通りである：

- ・ 101, 103B, 104, 105. 図1C参照。
- ・ 301-306. 図3B参照。
- ・ 307. CHASM。
- ・ 309. CHASM 307とメディア・エンジン302との間の制御情報。
- ・ 310. CHASM 307とメディア・エンジン301との間の制御情報。
- ・ 311. CHASM 307とアプリ308との間の（へ／からの）制御情報。
- ・ 312C. 装置101の装置属性記述子。
- ・ 313. 装置属性記述子312CからCHASM 307への制御情報および／またはデータ。
- ・ 314C. 装置105の装置属性記述子。
- ・ 315. 装置属性記述子314CからCHASM 307への制御情報および／またはデータ。
- ・ 316. 装置メディア・エンジン302から装置メディア・エンジン301へのメディア・ストリーム。
- ・ 317. ネットワークからメディア・エンジン301へ／からのメディア・ストリーム。

【0149】

いくつかの実施形態では、DOODADがCHASM（たとえば、CHASM 307）に含まれ、スマート・オーディオ装置と対話する（たとえば、スマート・オーディオ装置のそれぞれとの間で制御情報を送受信する）場合、および／またはDOODADが提供されて（たとえば、装置、たとえば装置101および105のサブシステムとして；ここで、これらの装置は、CHASMを実装する装置（たとえば、CHASM 307を実装する装置）とは別個）、CHASM（たとえば、以下に説明する図4のCHASM 401、またはCHASM 307）とともに動作する場合）、（一つまたは複数のスマート・オーディオ装置のそれぞれにおける）SPASMの必要性は、スマート・オーディオ装置がそのオーディオ能力を（CHASMに）広告する動作によって置き換えられ、アプリケーションは、オーディオ機能については、単一の抽象的な制御点（たとえば、CHASM）に委ねる。

【0150】

図4は、別の開示された実施形態を示すブロック図である。図4の設計は、いくつかの実装ではアプリケーションが直接選択または制御する必要がなく、場合によっては、どの特定のオーディオ装置が、アプリケーションに関連する機能の実行に関与するか、そのようなオーディオ装置の特定の機能などに関する情報が提供されない可能性があるという重要な抽象化を導入する。

【0151】

図4は、3つの別個の物理的なオーディオ装置420、421、および422を含むシステムの

10

20

30

40

50

ブロック図である。この例では、各装置は、発見可能な日和見的にオーケストレーションされる分散オーディオ・サブシステム (Discoverable Opportunistically Orchestrated Distributed Audio Subsystem、DOODAD) を実装し、アプリケーション (410～412) を実行しているCHASM 401によって制御される。この例によれば、CHASM 401は、アプリケーション410～412のそれぞれのためのメディア要件を管理するように構成される。

【0152】

図4では、ラベル付けされた要素は以下の通りである：

- ・400. 3つの異なる装置にまたがる、オーケストレーションされたオーディオ・システムの例。装置420、421および422はそれぞれ、DOODAD (DOODAD) を実装する。DOODADを実装する装置420、421、および422のそれぞれは、それ自身、DOODADと呼ばれることがある。この例では、図4のDOODADである、または、図4のDOODADを実装する装置は、それ自身では関連するアプリケーションを実装しないが、装置101はアプリケーション308を実装するという意味で、各DOODADは、図3Cのスマート・オーディオ装置101とは異なるスマート・オーディオ装置である (または、かかるスマート・オーディオ装置によって実装される)。

- ・401. CHASM。

- ・410～412. アプリ。この例ではオーディオ要件が異なる諸アプリ。いくつかの例では、アプリ410～412のそれぞれは、オーディオ環境の装置、または時々オーディオ環境内に位置するモバイル装置に格納され、かかる装置によって実行可能であってもよい。

- ・420～422. それぞれ発見可能な日和見的にオーケストレーションされる分散オーディオサブシステム (DOODAD) を実装するスマート・オーディオ装置。各DOODADは別々の物理的なスマート・オーディオ装置内で動作する。

- ・430,431,432. アプリ410、411、412の間で (へ／から) CHASM 401に送られる制御情報。

- ・433,434,435. DOODADS 420～422およびCHASM 401へ／から送られる制御情報。

- ・440,441,442. メディア・エンジン。

- ・450,451,452. 装置属性記述子。この例ではDOODADS 420-422がこれをCHASM 401に提供するように構成されている。

- ・460～462. ラウドスピーカー。

- ・463～465. マイクロフォン。

- ・470. メディア・エンジン442からネットワークへのメディア・ストリーム。

- ・471. メディア・エンジン441からメディア・エンジン442へのメディア・ストリーム。

- ・472. メディア・エンジン441からメディア・エンジン (440) へのメディア・ストリーム。

- ・473. ネットワークを介してメディアを提供するクラウドベースのサービス、たとえば、データセンターの一つまたは複数のサーバーがインターネットを介してメディア・エンジン441に提供するクラウドベースのサービスからのメディア・ストリーム。

- ・474. メディア・エンジン441からクラウドベースのサービスへのメディア・ストリーム。

- ・477. 一つまたは複数のクラウドベースのサービス。これは、一つまたは複数の音楽ストリーミング・サービス、映画ストリーミング・サービス、テレビ番組ストリーミング・サービス、ポッドキャスト・プロバイダーなどを含みうる。

【0153】

図4は、複数のオーディオ装置がオーディオ体験を生成するためのルーティングの可能性を持つシステムのブロック図である。この例によれば、メディア・ストリーム473に対応するオーディオ・データは、スマート・オーディオ装置421のメディア・エンジン441によって受領されており、スマート・オーディオ装置421は、メディア・ストリーム471をメディア・エンジン442に提供し、メディア・ストリーム472をメディア・エンジン440に提供している。いくつかのそのような実装によれば、メディア・エンジン441は、本

明細書で「メディア・エンジン・サンプル・クロック」と称されることがあるものの一例である、メディア・エンジン441のサンプル・クロックに従って、メディア・ストリーム473を処理することができる。

【0154】

いくつかのそのような例において、CHASM 401は、制御情報434を介して、メディア・ストリーム473の取得および処理に関連するプロセスに関する命令および情報をメディア・エンジン441に提供してもよい。そのような命令および情報は、本明細書で「オーディオ・セッション管理制御信号」と称されることがあるものの例である。

【0155】

しかしながら、いくつかの実装では、CHASM 401は、メディア・エンジン441のメディア・エンジン・サンプル・クロックを参照することなく、オーディオ・セッション管理制御信号を送信することができる。そのような例は、CHASM 401が、たとえば、オーディオ環境のオーディオ装置へのメディアの伝送を同期させる必要がないので、潜在的に有利である。代わりに、いくつかの実装では、そのような同期は、前述の例のスマート・オーディオ装置421のような別の装置に委ねられてもよい。

【0156】

いくつかのそのような実装によれば、CHASM 401は、アプリケーション410、アプリケーション411、またはアプリケーション412からの制御情報430、431、または432に応答して、メディア・ストリーム473を取得および処理することに関連するオーディオ・セッション管理制御信号をメディア・エンジン441に提供していてもよい。そのような制御情報は、本明細書で「アプリケーション制御信号」と称されることがあるものの例である。いくつかの実装によれば、アプリケーション制御信号は、メディア・エンジン441のメディア・エンジン・サンプル・クロックを参照することなく、アプリケーションからCHASM 401に送信されてもよい。

【0157】

いくつかの例では、CHASM 401は、メディア・エンジン441に、オーディオ処理情報を提供することができる。オーディオ処理情報は、それに応じてメディア・ストリーム473に対応するオーディオを処理するための命令と共に、レンダリング情報を含むが、これに限定されない。しかしながら、いくつかの実装では、CHASM 401を実装する装置（または、本明細書の他所に記載されているようなスマートホームハブの機能性のような、類似の機能性を実装する装置）は、少なくともいくつかのオーディオ処理機能性を提供するように構成することができる。いくつかの例を下記に与える。いくつかのそのような実装では、CHASM 401は、オーディオ・データを受領して処理し、処理された（たとえば、レンダリングされた）オーディオ・データをオーディオ環境のオーディオ装置に提供するように構成することができる。

【0158】

図5は、いくつかの実装に従ったオーディオ・セッション管理方法のブロックを含むフロー図である。方法500のブロックは、本明細書に記載される他の方法と同様に、必ずしも示される順序で実行されるわけではない。いくつかの実装では、方法500のブロックの一つまたは複数が同時に実行されてもよい。たとえば、いくつかの例において、ブロック505および510は、並行して実行されてもよい。さらに、方法500のいくつかの実装は、図示および／または説明されるよりも多いまたは少ないブロックを含んでいてもよい。方法500のブロックは、たとえば、図6に示されかつ以下に説明される制御システム610などの制御システム、または他の開示された制御システムの例の1つであってもよい（または、それらを含んでいてもよい）一つまたは複数の装置によって実行されてもよい。

【0159】

いくつかの実装によれば、方法500のブロックは、少なくとも部分的に、本明細書でオーディオ・セッション・マネージャと称されるもの、たとえばCHASMを実装している装置によって実行されてもよい。いくつかのそのような例において、方法500のブロックは、図2C、図2D、図3Cおよび図4を参照して上述したCHASM 208C、CHASM 208D、

10

20

30

40

50

CHASM 307および／またはCHASM 401によって、少なくとも部分的に実行されてもよい。より具体的には、いくつかの実装では、方法500のブロックにおいて言及される「オーディオ・セッション・マネージャ」の機能は、少なくとも部分的には、CHASM 208C、CHASM 208D、CHASM 307、および／またはCHASM 401によって実行されうる。

【0160】

この例によれば、ブロック505は、第1のアプリケーションを実行する第1のアプリケーション装置とオーディオ環境のオーディオ・セッション・マネージャとの間に第1のアプリケーション通信リンクを確立することを含む。いくつかの例では、第1のアプリケーション通信リンクは、Zigbee、AppleのBonjour（ランデブー）、Wi-Fi、Bluetooth、Bluetooth Low Energy（Bluetooth LE）、5G、4G、3G、一般パケット無線サービス（General Packet Radio Service、GPRS）、Amazon Sidewalk、RF24L01チップ内のNordicのカスタムプロトコルなど、オーディオ環境内で使用するのに適した任意の適切な無線通信プロトコルを介して作ることができる。いくつかの例では、第1のアプリケーション通信リンクは、「ハンドシェイク」プロセスに応答して確立されてもよく、これは、いくつかの例では、第1のアプリケーション装置によってオーディオ・セッション・マネージャを実装している装置に送信される「ハンドシェイク開始」を介して開始されてもよい。いくつかの例では、第1のアプリケーション通信リンクは、本明細書において第1のアプリケーション装置から「ルート開始要求」と称されるものに応答して確立されてもよい。便宜上、第1のアプリケーション装置からのルート開始要求は、ルート開始要求が「第1のアプリケーション装置」に対応することを示すために、本明細書では「第1のルート開始要求」と称されることがある。言い換えると、「第1の」という用語は、特定の実装に依存して、この文脈において時間的な意味を持つことも持たないこともある。

【0161】

そのような一例において、第1のアプリケーション通信リンクは、図4のアプリケーション410が実行されている装置とCHASM 401との間で確立されてもよい。いくつかのそのような例において、第1のアプリケーション通信リンクは、アプリケーション410が実行されている装置から第1のルート開始要求をCHASM 401が受け取ることに応答して確立されてもよい。アプリケーション410が実行されている装置は、たとえば、オーディオ環境のスマート・オーディオ装置であってもよい。場合によっては、アプリケーション410が実行されている装置は、携帯電話であってもよい。アプリケーション410は、CHASM 401を介してメディア、たとえば、音楽、テレビ番組、映画などにアクセスするために使用されうる。場合によっては、メディアは、クラウドベースのサービスを介してストリーミングするために利用可能であってもよい。

【0162】

「ルート（route）」が意味するもののさまざまな例が以下に詳細に説明される。一般に、ルートは、オーディオ・セッション・マネージャによって管理されるオーディオ・セッションのパラメータを示す。ルート開始要求は、たとえば、オーディオソースおよびオーディオ環境宛先を示すことができる。オーディオ環境宛先は、場合によっては、オーディオ環境内の少なくとも1人の人物と対応することができる。いくつかの例では、オーディオ環境宛先は、オーディオ環境のエリアまたはゾーンに対応することができる。

【0163】

しかしながら、ほとんどの場合、オーディオ環境宛先は、オーディオ環境においてメディアを再生することに関与するいかなる特定のオーディオ装置をも示すものではない。代わりに、アプリケーション（アプリケーション410のような）は、たとえば特定のタイプのメディアがオーディオ環境内の特定の人物に利用可能にされるべきであるというルート開始要求を提供することがある。さまざまな開示された実装では、オーディオ・セッション・マネージャは、どのオーディオ装置がルートに関与するかを決定する、たとえば、どのオーディオ装置がメディアに関連するオーディオ・データの取得、レンダリング、および再生に関与するかを決定することを受け持つ。いくつかの実装では、オーディオ・セッション・マネージャは、（たとえば、メディアの意図された受領者である人物が位置を変更

したとの判定に応答して) ルートに関与するオーディオ装置が変わったかどうかを判定する、対応するデータ構造を更新する、などを受け持つ。詳細な例が下記に与えられる。

【0164】

この例では、ブロック510は、オーディオ・セッション・マネージャによって、第1のアプリケーション通信リンクを介して、第1のアプリケーション制御信号を第1のアプリケーションから受領することを含む。再び図4を参照すると、いくつかの例において、アプリケーション制御信号は、アプリ410とCHASM 401との間で(へ／から)送られる制御情報430に対応しうる。いくつかの例では、第1のアプリケーション制御信号は、オーディオ・セッション・マネージャ(たとえば、CHASM 401)がルートを開始した後に送信されうる。しかしながら、場合によっては、第1のアプリケーション制御信号が第1のルート開始要求に対応してもよい。いくつかのそのような例では、ブロック505および510は、少なくとも部分的に、同時並行して実行されてもよい。

10

【0165】

この例によれば、ブロック515は、オーディオ・セッション・マネージャと少なくともオーディオ環境の第1のスマート・オーディオ装置との間に第1のスマート・オーディオ装置通信リンクを確立することを含む。この例では、第1のスマート・オーディオ装置は、単一目的のオーディオ装置または多目的のオーディオ装置のいずれかであるか、またはそれらを含む。この実装によれば、第1のスマート・オーディオ装置は、一つまたは複数のラウドスピーカーを含む。

【0166】

20

いくつかの例では、上述のように、第1のアプリケーション制御信号および／または第1のルート開始要求は、ルートに関与するいかなる特定のオーディオ装置を示すものでもない。いくつかのそのような例によれば、方法500は、ブロック515に先立って、オーディオ環境のどのオーディオ装置が少なくとも初期にルートに関与するかを決定する(たとえば、オーディオ・セッション・マネージャによって) プロセスを含んでいてもよい。

【0167】

たとえば、図4のCHASM 401は、オーディオ装置420、421、および422が、少なくとも初期には、ルートに関与することを決定することができる。図4に示される例では、ブロック515の第1のスマート・オーディオ装置通信リンクが、スマート・オーディオ装置421とオーディオ・セッション・マネージャ(この例では、CHASM 401)の間に確立されてもよい。第1のスマート・オーディオ装置通信リンクは、CHASM 401とスマート・オーディオ装置421との間に示される図4の破線に対応してもよく、それを通じて制御情報434が送信される。いくつかのそのような例では、第1のスマート・オーディオ装置通信リンクは、Apple Airplay、Miracast、Blackfire、Bluetooth 5、リアルタイムトランスポートプロトコル(Real-time Transport Protocol、RTP)などのような、オーディオ環境内での使用に適した任意の適切な無線通信プロトコルを介して作ることができる。

30

【0168】

図5に示した例では、ブロック520は、オーディオ・セッション・マネージャによって、第1のスマート・オーディオ装置の第1のメディア・エンジンの一つまたは複数の第1のメディア・エンジン機能を決定することを含む。この例によれば、第1のメディア・エンジンは、第1のスマート・オーディオ装置によって受領された一つまたは複数のオーディオメディア・ストリームを管理し、第1のメディア・エンジン・サンプル・クロックに従って一つまたは複数のオーディオメディア・ストリームについて第1のスマート・オーディオ装置信号処理を実行するように構成される。上述の例では、ブロック520は、スマート・オーディオ装置421からメディア・エンジン441の一つまたは複数の能力に関する情報を、たとえば、装置属性記述子451をCHASM 401に提供することによって、CHASM 401が受け取ることを含むことができる。いくつかの実装によれば、ブロック520は、スマート・オーディオ装置421の一つまたは複数のラウドスピーカーの能力に関する情報をCHASM 401が受領することを含んでいてもよい。いくつかの例では、CHASM 401は

40

50

、たとえば、方法500のブロック505、510および／または515の前に、この情報の一部または全部をあらかじめ決定していてもよい。

【0169】

この例によれば、ブロック525は、第1のスマート・オーディオ装置通信リンクを介して第1のスマート・オーディオ装置に送信される第1のオーディオ・セッション管理制御信号を介して、オーディオ・セッション・マネージャによって、第1のメディア・エンジン能力に従って第1のスマート・オーディオ装置を制御することを含む。いくつかの例によれば、第1のオーディオ・セッション管理制御信号は、第1のスマート・オーディオ装置に、第1のメディア・エンジンの制御をオーディオ・セッション・マネージャに委任させることができる。この例では、オーディオ・セッション・マネージャは、第1のメディア・エンジン・サンプル・クロックを参照することなく、第1のスマート・オーディオ装置に第1のオーディオ・セッション管理制御信号を送信する。いくつかのそのような例では、第1のアプリケーション制御信号は、第1のメディア・エンジン・サンプル・クロックを参照することなく、第1のアプリケーションからオーディオ・セッション・マネージャに送信されうる。

【0170】

ブロック525の一例では、CHASM 401はメディア・エンジン441を制御してメディア・ストリーム473を受領することができる。いくつかのそのような例では、第1のオーディオ・セッション管理制御信号を介して、CHASM 401は、メディア・ストリーム473を受領することができるウェブサイトに対応するユニバーサル・リソース・ロケータ（URL）と、メディア・ストリーム473を開始するための命令とを、メディア・エンジン441に提供することができる。いくつかのそのような例によれば、CHASM 401はまた、メディア・ストリーム471をメディア・エンジン442に提供し、メディア・ストリーム472をメディア・エンジン440に提供するための命令を、第1のオーディオ・セッション管理制御信号を介して、メディア・エンジン441に提供してもよい。

【0171】

いくつかのそのような例において、CHASM 401は、メディア・エンジン441に、第1のオーディオ・セッション管理制御信号を介して、オーディオ処理情報を提供していてもよい。該オーディオ処理情報は、それに応じてメディア・ストリーム473に対応するオーディオを処理する命令と共に、レンダリング情報を含むがこれに限定されない。たとえば、CHASM 401は、たとえばスマート・オーディオ装置420が左チャンネルに対応するスピーカー・フィード信号を受領し、スマート・オーディオ装置421が中央チャンネルに対応するスピーカー・フィード信号を再生し、スマート・オーディオ装置422が右チャンネルに対応するスピーカー・フィード信号を受領するという指示をメディア・エンジン441に与えていてもよい。

【0172】

レンダリングの種々の他の例が本明細書に開示されており、そのいくつかは、CHASM 401または別のオーディオ・セッション・マネージャが種々のタイプのオーディオ処理情報をスマート・オーディオ装置に伝達することを含んでいてもよい。たとえば、いくつかの実装では、オーディオ環境の一つ以上の装置は、質量中心振幅パン（Center of Mass Amplitude Panning、CMAP）および／または柔軟仮想化（Flexible Virtualization、FV）のような柔軟なレンダリングを実装するように構成することができる。いくつかのそのような実装では、柔軟レンダリングを実装するように構成された装置は、一組のオーディオ装置位置、推定される現在の聴取者位置、推定される現在の聴取者配向を与えられることができる。柔軟レンダリングを実装するように構成された装置は、一組のオーディオ装置の位置、推定された現在の聴取者位置、推定された現在の聴取者配向に従って、環境内の一組のオーディオ装置のオーディオをレンダリングするように構成されてもよい。いくつかの詳細な例を下記に記述する。

【0173】

図4を参照して説明した方法500の前述の例では、オーディオ・セッション・マネージャ

または第1のスマート・オーディオ装置以外の装置が、第1のアプリケーションを実行するために構成される。しかしながら、たとえば、図2Cおよび図2Dを参照して上述したように、第1のスマート・オーディオ装置のいくつかの例は、第1のアプリケーションを実行するために構成することができる。

【0174】

いくつかのそのような例、たとえば、図2Cを参照して上述したような例によれば、第1のスマート・オーディオ装置は、特定目的のオーディオ・セッション・マネージャを含むことができる。いくつかのそのような実装では、オーディオ・セッション・マネージャは、第1のスマート・オーディオ装置通信リンクを介して特定目的のオーディオ・セッション・マネージャと通信することができる。いくつかの例では、オーディオ・セッション・マネージャは、特定目的のオーディオ・セッション・マネージャから一つまたは複数の第1のメディア・エンジン能力を取得することができる。いくつかのそのような例によれば、オーディオ・セッション・マネージャは、アプリケーションが第1のスマート・オーディオ装置上で実行されているか、または別の装置上で実行されているかにかかわらず、第1のメディア・エンジンを制御するすべてのアプリケーションについてのゲートウェイとして動作することができる。

【0175】

上述のように、いくつかの例では、方法500は、第1のオーディオソースに対応する少なくとも第1のオーディオ・ストリーム（たとえば、図4のメディア・ストリーム473）を確立することを含んでいてもよい。いくつかの例では、第1のオーディオソースは、音楽ストリーミング・サービス、テレビ番組および／または映画ストリーミング・サービスなどのクラウドベースのメディア・ストリーミング・サービスを提供するように構成された一つまたは複数のサーバーなどであってもよい。第1のオーディオ・ストリームは、第1のオーディオ信号を含んでいてもよい。いくつかのそのような実装形態では、少なくとも第1のオーディオ・ストリームを確立することは、第1のスマート・オーディオ装置通信リンクを介して第1のスマート・オーディオ装置に送信される第1のオーディオ・セッション管理制御信号を介して、第1のスマート・オーディオ装置に、少なくとも第1のオーディオ・ストリームを確立させることを含むことができる。

【0176】

いくつかの例では、方法500は、第1のオーディオ信号を第1にレンダリングされたオーディオ信号にレンダリングされるようにするレンダリング・プロセスを含むことができる。いくつかのそのような実装では、レンダリング・プロセスは、第1のオーディオ・セッション管理制御信号に応答して、第1のスマート・オーディオ装置によって実行されてもよい。上述の例では、メディア・エンジン441は、第1のオーディオ・セッション管理制御信号に応答して、メディア・ストリーム473に対応するオーディオ信号をスピーカー・フィード信号にレンダリングしてもよい。

【0177】

いくつかの例によれば、方法500は、第1のオーディオ・セッション管理制御信号を介して、第1のスマート・オーディオ装置に、第1のスマート・オーディオ装置と、オーディオ環境の一つまたは複数の他のスマート・オーディオ装置のそれぞれとの間のスマート・オーディオ装置間通信リンクを確立させることを含む。図4を参照して上述した例では、メディア・エンジン441は、有線または無線のスマート・オーディオ装置間通信リンクをメディア・エンジン440および442と確立することができる。図3Cを参照して上述した例では、メディア・エンジン302は、有線または無線のスマート・オーディオ装置間通信リンクを確立して、メディア・ストリーム316をメディア・エンジン301に提供することができる。

【0178】

いくつかの例において、方法500は、第1のスマート・オーディオ装置に、スマート・オーディオ装置間通信リンク（単数または複数）を介して、未加工マイクロフォン信号、処理済みマイクロフォン信号、レンダリング済みオーディオ信号または未レンダリングオ

10

20

30

40

50

オーディオ信号のうちの一つまたは複数の一つまたは複数の他のスマート・オーディオ装置に送信させることを含むことができる。図4を参照して上述した例では、スマート・オーディオ装置間通信リンクを使用して、レンダリングされたオーディオ信号またはレンダリングされていないオーディオ信号を、メディア・ストリーム471およびメディア・ストリーム472を介して提供することができる。いくつかのそのような例では、メディア・ストリーム471は、メディア・エンジン442のためのスピーカー供給信号を含んでいてもよく、メディア・ストリーム472は、メディア・エンジン440のためのスピーカー供給信号を含んでいてもよい。図3Cを参照して上述した例では、メディア・エンジン302は、メディア・ストリーム316を介して、未加工のマイクロフォン信号または処理済みのマイクロフォン信号をメディア・エンジン301に提供することができる。

10

【0179】

いくつかの例によれば、方法500は、オーディオ・セッション・マネージャと少なくともオーディオ環境の第2のスマート・オーディオ装置との間に第2のスマート・オーディオ装置通信リンクを確立することを含む。いくつかのそのような例では、第2のスマート・オーディオ装置は、単一目的のオーディオ装置または多目的のオーディオ装置であってもよい。場合によっては、第2のスマート・オーディオ装置は、一つまたは複数のマイクロフォンを含んでいてもよい。そのような方法は、オーディオ・セッション・マネージャによって、第2のスマート・オーディオ装置の第2のメディア・エンジンの一つまたは複数の第2のメディア・エンジン能力を決定することを含む。第2のメディア・エンジンは、たとえば、一つまたは複数のマイクロフォンからマイクロフォン・データを受領し、第2のスマート・オーディオ装置信号処理をマイクロフォン・データに対して実行するように構成することができる。

20

【0180】

たとえば、図3Cを参照すると、「第1のスマート・オーディオ装置」はスマート・オーディオ装置101であってもよい。いくつかのそのような例によれば、「第2のスマート・オーディオ装置」は、スマート・オーディオ装置105であってもよい。「第1のスマート・オーディオ装置通信リンク」は、制御信号310を提供するために使用することができ、「第2のスマート・オーディオ装置通信リンク」は、制御信号309を提供するために使用することができる。CHASM 307は、少なくとも部分的には、装置属性記述子314cに基づいて、メディア・エンジン302の一つまたは複数のメディア・エンジン能力を決定することができる。

30

【0181】

いくつかのそのような方法は、第2のスマート・オーディオ装置通信リンクを介して第2のスマート・オーディオ装置に送信される第2のオーディオ・セッション・マネージャ制御信号を介して、オーディオ・セッション・マネージャによって第2のメディア・エンジン能力に従って第2のスマート・オーディオ装置を制御することを含む。いくつかの例では、第2のスマート・オーディオ装置を制御することは、第2のスマート・オーディオ装置に、第2のスマート・オーディオ装置と第1のスマート・オーディオ装置との間にスマート・オーディオ装置間通信リンク（たとえば、メディア・ストリーム316を提供するために使用されるスマート・オーディオ装置間通信リンク）を確立させることを含むことができる。いくつかのそのような例は、第2のスマート・オーディオ装置に、スマート・オーディオ装置間通信リンクを介して、第2のメディア・エンジンから第1のメディア・エンジンに、処理済みまたは未処理のマイクロフォン・データ（たとえば、マイクロフォン305からの処理済みまたは未処理のマイクロフォン・データ）の少なくとも1つを送信させることを含む。

40

【0182】

いくつかの例では、第2のスマート・オーディオ装置を制御することは、オーディオ・セッション・マネージャによって、第1のアプリケーション通信リンクを介して、第1のアプリケーションからの第1のアプリケーション制御信号を受領することを含む。図3Cの例では、CHASM 307は、アプリケーション308から制御信号311を受領する。アプリ

50

ケーション308は、この場合の電話アプリケーションである。いくつかのそのような例は、第1のアプリケーション制御信号に従って第2のオーディオ・セッション・マネージャ制御信号を決定することを含んでいてもよい。たとえば、再び図3Cを参照すると、CHASM 307は、アプリケーション308からの制御信号311に従って提供される電話会議のために発話対エコー比 (speech to echo ratio、SER) を最適化するように構成することができる。CHASM 307は、人102の発話を捕捉するために、マイクロフォン303の代わりにマイクロフォン305を使用することによって、テレビ会議についてのSERが改善できると判断することができる(図1C参照)。この判断は、いくつかの例において、人102の位置の推定に基づくことができる。オーディオ環境における人の位置および／または配向を推定するためのいくつかの詳細な例が本明細書に開示される。

【0183】

図6は、本開示のさまざまな側面を実装可能な装置の構成要素の例を示すブロック図である。本明細書で提供される他の図と同様に、図6に示される要素のタイプおよび数は、単に一例として提供されるに過ぎない。他の実装は、要素のより多くの、より少ない、および／または異なるタイプおよび数を含みうる。いくつかの例によれば、装置600は、本明細書に開示された方法の少なくとも一部を実行するように構成されたスマート・オーディオ装置であってもよく、または、それを含んでいてもよい。他の実装では、装置600は、ラップトップコンピュータ、セルラー電話、タブレット装置、スマートホームハブなどのような、本明細書に開示された方法の少なくともいくつかを実行するように構成された別の装置であってもよく、またはこれを含むこともできる。いくつかのそのような実装では、装置600は、サーバーであってもよく、またはサーバーを含んでいてもよい。いくつかのそのような実装では、装置600は、本明細書でCHASMと称されるものを実装するように構成されてもよい。

【0184】

この例では、装置600は、インターフェース・システム605および制御システム610を含む。インターフェース・システム605は、いくつかの実装において、ソフトウェア・アプリケーションを実行している、または実行するように構成されている一つまたは複数の装置と通信するように構成されてもよい。そのようなソフトウェア・アプリケーションは、本明細書では時に、「アプリケーション」または単に「アプリ」と称されることがある。インターフェース・システム605は、いくつかの実装では、アプリケーションに関する制御情報および関連したデータを交換するように構成されてもよい。

【0185】

インターフェース・システム605は、いくつかの実装において、オーディオ環境の一つまたは複数の他の装置と通信するように構成されてもよい。オーディオ環境は、いくつかの例では、家庭オーディオ環境であってもよい。インターフェース・システム605は、いくつかの実装では、制御情報および関連データをオーディオ環境の諸オーディオ装置と交換するように構成されてもよい。制御情報および関連データは、いくつかの例において、装置600が通信のために構成される一つまたは複数のアプリケーションに関連しうる。インターフェース・システム605は、いくつかの実装において、オーディオ・データを受領するように構成されてもよい。オーディオ・データは、オーディオ環境の少なくとも一部のスピーカーによって再生されるようにスケジュールされたオーディオ信号を含むことができる。オーディオ・データは、一つまたは複数のオーディオ信号および関連する空間データを含んでいてもよい。空間データは、たとえば、チャンネル・データおよび／または空間メタデータを含んでいてもよい。インターフェース・システム605は、レンダリングされたオーディオ信号を環境のラウドスピーカーのセットの少なくともいくつかのラウドスピーカーに提供するように構成されてもよい。インターフェース・システム605は、いくつかの実装では、環境内の一つまたは複数のマイクロフォンからの入力を受領するように構成されてもよい。

【0186】

インターフェース・システム605は、一つまたは複数のネットワーク・インターフェー

スおよび／または一つまたは複数の外部装置インターフェース（一つまたは複数のユニバーサルシリアルバス（USB）インターフェースなど）を含んでいてもよい。いくつかの実装形態によれば、インターフェース・システム605は、一つまたは複数の無線インターフェースを含んでいてもよい。インターフェース・システム605は、一つまたは複数のマイクロフォン、一つまたは複数のスピーカー、表示システム、タッチセンサーシステム、および／またはジェスチャーセンサーシステムのような、ユーザーインターフェースを実装するための一つまたは複数の装置を含んでいてもよい。いくつかの例では、インターフェース・システム605は、制御システム610と、図6に示される任意的なメモリ・システム615などのメモリ・システムとの間の一つまたは複数のインターフェースを含んでいてもよいが、制御システム610は、いくつかの例では、メモリ・システムを含んでいてもよい。

10

【0187】

制御システム610は、たとえば、汎用の単一チップまたはマルチチップ・プロセッサ、デジタル信号プロセッサ（DSP）、特定用途向け集積回路（ASIC）、フィールド・プログラマブル・ゲート・アレイ（FPGA）、または他のプログラマブル・ロジック装置、離散的ゲートまたはトランジスタ・ロジック、および／または離散的ハードウェア構成要素を含んでいてもよい。

【0188】

いくつかの実装では、制御システム610は、2つ以上の装置内に存在してもよい。たとえば、制御システム610の一部分は、本明細書に示される環境の1つの中の装置内に存在してもよく、制御システム610の別の一部分は、サーバー、モバイル装置（たとえば、スマートフォンまたはタブレットコンピュータ）など、環境外の装置内に存在してもよい。他の例では、制御システム610の一部分は、本明細書に示される環境の1つの中の装置内に存在してもよく、制御システム610の別の一部分は、環境の一つまたは複数の他の装置内に存在してもよい。たとえば、制御システムの機能性は、環境の複数のスマート・オーディオ装置に分散されてもよく、または、オーケストレーション装置（たとえば、本明細書ではスマートホームハブと呼ぶことがあるもの）および環境の一つまたは複数の他の装置によって共有されてもよい。インターフェース・システム605はまた、いくつかのそのような例では、2つ以上の装置に存在してもよい。

20

【0189】

いくつかの実装では、制御システム610は、少なくとも部分的に、本明細書に開示された方法を実行するために構成されてもよい。いくつかの例によれば、制御システム610は、オーディオ・セッション管理方法を実装するように構成されてもよい。

30

【0190】

本明細書に記載された方法の一部または全部は、一つまたは複数の非一時的媒体に記憶された命令（たとえば、ソフトウェア）に従って一つまたは複数の装置によって実行される。そのような非一時的媒体は、ランダムアクセスメモリ（RAM）装置、読み出し専用メモリ（ROM）装置などを含むが、これらに限定されない、本明細書に記載されたもののようなメモリ装置を含んでいてもよい。一つまたは複数の非一時的媒体は、たとえば、図6に示される任意的なメモリ・システム615および／または制御システム610に存在してもよい。よって、本開示に記載された主題の種々の革新的な側面は、ソフトウェアを格納した一つまたは複数の非一時的媒体に実装することができる。ソフトウェアは、たとえば、オーディオ・セッション管理方法を実装するために少なくとも1つの装置を制御するための命令を含むことができる。ソフトウェアは、いくつかの例において、オーディオ・データを取得、処理および／または提供するために、オーディオ環境の一つまたは複数のオーディオ装置を制御するための命令を含むことができる。ソフトウェアは、たとえば、図6の制御システム610のような制御システムの一つまたは複数の構成要素によって実行可能であってもよい。

40

【0191】

いくつかの例では、装置600は、図6に示される任意的なマイクロフォン・システム620を含んでいてもよい。任意的なマイクロフォン・システム620は、一つまたは複数のマ

50

イクロフォンを含んでいてもよい。いくつかの実装では、一つまたは複数のマイクロフォンは、スピーカー・システムのスピーカー、スマート・オーディオ装置等のような別の装置の一部であってもよく、または関連付けられてもよい。いくつかの例では、装置600は、マイクロフォン・システム620を含まなくてもよい。しかしながら、いくつかのそのような実装では、それにもかかわらず、装置600は、インターフェース・システム610を介して、オーディオ環境内の一つまたは複数のマイクロフォンのためのマイクロフォン・データを受領するように構成されてもよい。

【0192】

いくつかの実装によれば、装置600は、図6に示される任意的なスピーカー・システム625を含んでいてもよい。任意的なスピーカー・システム625は、一つまたは複数のスピーカーを含んでいてもよい。本明細書では、ラウドスピーカーを「スピーカー」と呼ぶことがある。いくつかの例では、任意的なラウドスピーカー・システム625の少なくともいくつかのラウドスピーカーは、任意に配置することができる。たとえば、任意的なラウドスピーカー・システム625の少なくともいくつかのスピーカーは、ドルビー5.1、ドルビー5.1.2、ドルビー7.1、ドルビー7.1.4、ドルビー9.1、浜崎22.2等のような、標準的に処方されたスピーカー・レイアウトに対応しない位置に配置されてもよい。いくつかのそのような例では、任意的なラウドスピーカー・システム625の少なくともいくつかのラウドスピーカーは、スペースに都合のよい位置（たとえば、ラウドスピーカーを収容するスペースがある位置）に配置されてもよいが、いかなる標準の指定されたラウドスピーカー・レイアウトにも配置されていない。

【0193】

いくつかの実装では、装置600は、図6に示される任意的なセンサー・システム630を含んでいてもよい。任意的なセンサー・システム630は、一つまたは複数のカメラ、タッチセンサー、ジェスチャーセンサー、動き検出器などを含んでいてもよい。いくつかの実装によれば、任意的なセンサー・システム630は、一つまたは複数のカメラを含んでいてもよい。いくつかの実装では、カメラは自立型カメラであってもよい。いくつかの例において、任意的なセンサー・システム630の一つまたは複数のカメラは、単一目的のオーディオ装置またはバーチャル・アシスタントであってもよいスマート・オーディオ装置内に存在してもよい。いくつかのそのような例では、任意的なセンサー・システム630の一つまたは複数のカメラが、テレビ、携帯電話、またはスマートスピーカー内に存在してもよい。いくつかの例では、装置600は、センサー・システム630を含まなくてもよい。しかしながら、いくつかのそのような実装では、それにもかかわらず、装置600は、インターフェース・システム610を介して、オーディオ環境内の一つまたは複数のセンサーについてのセンサー・データを受領するように構成されてもよい。

【0194】

いくつかの実装では、装置600は、図6に示される任意的な表示システム635を含んでいてもよい。任意的な表示システム635は、一つまたは複数の発光ダイオード（LED）ディスプレイなどの一つまたは複数のディスプレイを含んでいてもよい。場合によっては、任意的な表示システム635は、一つまたは複数の有機発光ダイオード（OLED）ディスプレイを含んでいてもよい。装置600が表示システム635を含むいくつかの例では、センサー・システム630は、表示システム635の一つまたは複数のディスプレイに近接して、タッチセンサーシステムおよび／またはジェスチャーセンサーシステムを含んでいてもよい。いくつかのそのような実装によれば、制御システム610は、表示システム635を制御して、一つまたは複数のグラフィカル・ユーザー・インターフェース（GUI）を提示するように構成されてもよい。

【0195】

いくつかの例によれば、装置600は、スマート・オーディオ装置であってもよく、またはスマート・オーディオ装置を含んでいてもよい。いくつかのそのような実装においては、装置600は、ウェイクワード検出器であってもよく、またはそれを（少なくとも部分的に）実装してもよい。たとえば、装置600は、バーチャル・アシスタントであってもよく

、またはそれを（少なくとも部分的に）実装してもよい。

【0196】

ここで図4に戻ると、いくつかの例によれば、図4のシステムは、CHASM 401がアプリケーション410、411、および412に抽象化を提供するように実装することができ、その結果、アプリケーション410、411、および412は、オーケストレーションの言語でルートを発行することによって、プレゼンテーション〔呈示〕（たとえば、オーディオ・セッション）を達成することができる。この言語を使用して、アプリケーション410、411、および412は、制御リンク430、431、および432を通じてCHASM 401にコマンドすることができる。

【0197】

上述の図1A、図1B、および図1Cを参照すると、我々は、アプリケーションおよび結果として生じる状況の、オーケストレーションの言語での記述は、図1Aの場合と図1Cの場合との間で同一であろうと考えている。

【0198】

構文の例およびオーケストレーションの言語の例を説明するために、まず、図1Aおよび図1Cの状況を考えるいくつかの例を提供する。

【0199】

いくつかの例では、本明細書でオーケストレーション言語の「ルート」と呼ばれるものは、メディア・ソース（オーディオソースを含むが、これに限定されない）およびメディア宛先の指示を含むことができる。メディア・ソースとメディア宛先は、たとえば、アプリケーションによってCHASMに送信されるルート開始要求において指定することができる。いくつかの実装によれば、メディア宛先は、オーディオ環境宛先であってもよいし、オーディオ環境宛先を含んでいてもよい。オーディオ環境宛先は、いくつかの場合には、少なくとも一部の時間には、オーディオ環境にいる少なくとも1人の人と対応することができる。いくつかの例では、オーディオ環境宛先は、オーディオ環境の一つまたは複数のエリアまたはゾーンに対応しうる。オーディオ環境ゾーンのいくつかの例がここに開示される。しかしながら、オーディオ環境宛先は、一般に、ルートに関与するオーディオ環境の特定のオーディオ装置は含まないであろう。オーケストレーションの言語（ルートを確立するためにアプリケーションから必要とされる詳細を含むが、これに限定されない）をより一般化することによって、ルート実装の詳細はCHASMによって決定され、必要に応じて更新される可能性がある。

【0200】

ルートは、いくつかの例においては、オーディオ・セッション優先度、接続性モード、一つまたは複数のオーディオ・セッション目標または基準などの他の情報を含むことができる。いくつかの実装では、ルートは、対応するコードまたは識別子を有し、これは、本明細書では、オーディオ・セッション識別子と称することができる。場合によっては、オーディオ・セッション識別子は、永続的な固有のオーディオ・セッション識別子であってもよい。

【0201】

上述の初期例において、対応する「ルート」は、次のものを含んでいてもよい：高優先度の同期モードでの人X（たとえば、図1Cのユーザー102）からネットワーク（たとえば、インターネット）へのルート；高優先度の同期モードでのネットワークから人Xへのルート。そのような用語は、「電話通話を人Xに接続する」ことを望みうることを述べる自然言語と似ているが、この用語は（装置101を含む図1Aの状況に関して）以下のことを述べるのとは全く異なる：

この装置マイク（すなわち、装置101のマイクロフォン）を処理（ノイズ/エコー除去）に接続；

処理出力（ノイズ/エコー）をネットワークに接続；

ネットワークを処理入力に接続（ダイナミクス）；

処理出力（ダイナミクス）をこの装置スピーカーに接続；

10

20

30

40

50

処理出力（ダイナミクス）を処理入力（参照）に接続。

【0202】

この時点で、そのようなコマンドのリストに従って電話アプリケーションを実行することは、装置の詳細および必要な処理（エコーおよびノイズの除去）に関わり、装置105を導入する場合（換言すれば、電話アプリケーションの実行が、装置105および装置101を含む図1Cのコンテキストで実行される場合）、たとえば下記のように、完全に変更する必要があることに気づくかもしれない：

その装置マイク（すなわち、装置105のマイクロフォン）をネットワークに接続；

ネットワークを入力処理（ノイズ/エコー除去）に接続；

処理出力（ノイズ/エコー）をネットワークに接続；

ネットワークを処理入力（ダイナミクス）に接続；

処理出力（ダイナミック）をこの装置スピーカー（すなわち、装置101のスピーカー）に接続；

処理出力（ダイナミクス）を処理入力（参照）に接続。

【0203】

信号処理をどこで行うのが最良か、信号をどのように接続するか、および一般にユーザー（既知または未知の位置にある可能性がある）にとって最良の結果は何かの詳細は、いくつかの例では、限定された数の使用ケースについてあらかじめ計算できる最適化に関わるが、多数の装置および／または多数の同時オーディオ・セッションについては管理不可能になる可能性がある。我々は、スマート・オーディオ装置のよりよい接続性、能力、知識、および制御を許容しうる、基礎となるフレームワークを提供し（装置をオーケストレーションするまたは協調させることによることを含む）、装置を制御するためのポータブルで効果的な構文を作成することがより良いことを認識した。

【0204】

いくつかの開示された実施形態は、設計において有効でありかつ極めて一般的でもあるアプローチおよび言語を使用する。オーディオ装置を、特定のエンドポイント・オーディオ装置ではなく、ルートの一部として考えるときに（たとえば、システムが、本明細書に記載されたSPASMではなく、本明細書に記載されたCHASMを含む実施形態において）最もよく理解される言語の具体的な諸側面がある。いくつかの実施形態の諸側面は、ルート指定構文、永続的な固有セッション識別子、および送達、確認応答および／または品質の連続的な観念のうちの一つまたは複数を含む。

【0205】

ルート指定構文（すべての発行されたルートが明示的または暗示的な要素を持つ必要性を扱う）は、以下を含む：

○ソース（人/装置/自動化決定）、よって含意される権限；

○この望ましいオーディオ・ルーティングが、すでに進行中であるまたは後に来る可能性のある他のオーディオと比較して、どの程度重要であるかという点での優先度；

○宛先（理想的には、人または人の集合であり、場所に一般化できる可能性がある）；

○同期的、トランザクション、またはスケジュールされた、の範囲がある接続性モード；

○メッセージが確認応答されなければならない程度、または送達における信頼度の観点での確実性の要件；および／または

○聞かれているコンテンツの最も重要な側面（了解性、質、空間的忠実度、一貫性、あるいはおそらくは非可聴性）の感覚。この最後のポイントは、ネガティブ・ルートの概念を含む。そこでは、聞き、聞かれるだけでなく、聞かれない、および／または聞こえないものを制御することにも関心がある。いくつかのそのような例は、以下に詳細に説明する「赤ん坊を起こすな」実装のように、オーディオ環境の一つまたは複数の領域が比較的静かであることを保証することを含む。そのような他の例としては、たとえば、一つまたは複数の近くのスピーカーで「白色雑音」を再生すること、一つまたは複数の近くのスピーカーにおける一つまたは複数の他のオーディオ・コンテンツの再生レベルを上げることなどによって、オーディオ環境内の他の人が内密の会話を耳にすることを防ぐことを含

みうる。

【0206】

永続的固有セッション識別子の諸側面は、以下のものを含みうる。いくつかの実施形態のある重要な側面は、いくつかの例では、ルートに対応するオーディオ・セッションは、完了するまで、または、そうでなければ閉じられるまで、永続的であるということである。たとえば、これにより、システムは、進行中のオーディオ・セッションを（たとえば、CHASMを介して）追跡し、ルーティングを変更するためにオーディオ・セッションを終了または削除することができ、接続性のどの個別のセットが変更されなければならないかをアプリケーションが決定する必要がない。ひとたび作成されると、永続的固有セッション識別子は、メッセージまたはポーリング駆動管理をシステムが実装できるようにすることができ、制御およびステータスの諸側面を含んでいてもよい。たとえば、オーディオ・セッションまたはルートのコントロールは、以下のものであってもよく、またはそれを含んでいてもよい：

- ・終了する；
- ・宛先を動かす；
- ・優先度の増減

オーディオ・セッションまたはルートについて問い合わせられる可能性のあるものには、以下のものであってもよく、またはそれを含んでいてもよい：

- ・それは適切な位置にあるか；
- ・競合する優先事項の中で、表明されている目標がどの程度うまく実施されているか；
- ・オーディオを聞いた/確認応答したユーザーの感覚または自信はどのくらいか；
- ・どのような品質であるか（たとえば、忠実度、空間的、了解性、情報、注意、一貫性、または非可聴性という異なる目標に対して）
- ・望むなら、オーディオ装置が使用されている実際のルート・レイヤーまで照会する。

【0207】

送達、確認応答および品質の連続的な観念の諸側面は、以下を含みうる。ネットワーキング・ソケットのアプローチ（およびセッション層）について何らかの感覚があるかもしれないが、オーディオ・ルーティングは、特に、同時にルーティングまたはキューイングされる可能性のあるオーディオ活動の数などを考慮する場合、全く異なるかもしれない。また、宛先が少なくとも1人である可能性があり、場合によっては、ルーティングで通る可能性のあるオーディオ装置に対するその人の位置が不確実である可能性があるため、非常に連続的な信頼度の感覚を持つことが有用である。ネットワーキングは、到着する可能性のあるデータグラムと到着することが保証されたストリームのどちらかであるリンクを含んでいてもよいし、それに関連してもよい。オーディオの場合、物事が聞かれる、または聞かれないかもしれないという感覚、および／または我々が誰かを聞くことができる、または聞くことができないという感覚がありうる。

【0208】

これらの項目はオーケストレーション言語のいくつかの実施形態で導入されるものであり、単純なネットワーキングのいくつかの側面を有することがある。これに加えて（いくつかの実施形態では）、プレゼンテーション層およびアプリケーション層がある（たとえば、「電話コール」の例示的応用を実装する際に使用するため）。

【0209】

オーケストレーション言語の実施形態は、セッション開始プロトコル（SIP）および／またはメディアサーマークアップ言語（MSML）に関連する側面を有してもよい（たとえば、装置中心で、連続的かつ自律的に、オーディオ・セッションの現在のセットに基づいてルーティングを適応させる）。SIPは、音声、ビデオ、および／またはメッセージングアプリケーションを含みうるセッションの開始、維持、および終了に使用される信号伝達プロトコルである。いくつかの場合には、SIPは、インターネット電話の通信セッションを信号伝達および制御するため、たとえば音声通話、ビデオ通話、プライベートIP電話システムのため、インターネットプロトコル（IP）ネットワーク上のインスタントメッ

セージングのため、携帯電話通話のためなどに、使用されうる。SIPは、メッセージのフォーマットおよび参加者の通信のシーケンスを定義するテキストベースのプロトコルである。SIPは、ハイパーテキスト転送プロトコル（HTTP）および単純メール転送プロトコル（SMTP）の要素を含む。SIPで確立された通話は、場合によっては複数のメディア・ストリームを含むことがあるが、SIPメッセージのペイロードとしてデータを交換するアプリケーション（たとえば、テキストメッセージングアプリケーション）には別個の諸ストリームは必要ない。

【0210】

MSMLは、コメント要請（RFC）5707に記載されている。MSMLは、IPメディア・サーバー上の各種サービスを制御するために使用される。MSMLによれば、メディア・サーバーは、リアルタイムトランスポートプロトコルのメディア・ストリームのようなメディア・ストリームを制御および／または操作することに特化したアプライアンスである。MSMLによれば、アプリケーション・サーバーはメディア・サーバーとは別に構成であり、通話接続を確立および中止するように構成される。MSMLによれば、アプリケーション・サーバーは、SIPまたはIPを介して制御「トンネル」を確立するように構成されており、アプリケーション・サーバーは、メディア・サーバーとの間で、MSMLでコーディングされた要求および応答を交換するためにそれを使用する。

【0211】

MSMLは、マルチメディア・セッションがメディア・サーバー上でどのように相互作用するかを定義し、個々のユーザーまたはユーザー・グループにサービスを適用するために使用されうる。MSMLは、ビデオレイアウトやオーディオ混合などのメディア・サーバー会議機能の制御を制御し、サイドバー会議やパーソナルミックスを作成し、メディア・ストリームの属性を設定するなどのために使用することができる。

【0212】

いくつかの実施形態は、特定のコマンドを発行することによって、ユーザーがオーディオ装置の一群を制御することを可能にする必要がない。しかしながら、いくつかの実施形態は、装置自体を参照することなく、アプリケーション層のすべての所望の呈示を効果的に達成することができることが考えられている。

【0213】

図7は、一例によるCHASMのブロックを示すブロック図である。図7は、図4に示されるCHASM 401の例を示す。図7は、オーケストレーション言語を使用して複数のアプリからルートを受領し、それらのルートに関する情報をルーティング・テーブル701に格納するCHASM 401を示している。図7の要素は、以下を含む：

401: CHASM;

430: オーケストレーション言語を使用する第1のアプリケーション（図4のアプリケーション410）からのコマンドおよびCHASM 401からの応答;

431: オーケストレーション言語を使用する第2のアプリケーション（図4のアプリケーション411）からのコマンドおよびCHASM 401からの応答;

432: オーケストレーション言語を使用する第3のアプリケーション（図4のアプリケーション412）からのコマンドおよびCHASM 401からの応答;

703: オーケストレーション言語を使用する、図4に示されていない追加的なアプリケーションからのコマンドおよびCHASM 401からの応答;

701: CHASM 401によって維持されるルーティング・テーブル;

702: 現在のルーティング情報に基づいて複数のオーディオ装置を連続的に制御する、本明細書でオーディオ・セッション・マネージャとも呼ばれる最適化器;

435: CHASM 401から第1のオーディオ装置（図4のオーディオ装置420）へのコマンドおよび第1のオーディオ装置からの応答;

434: CHASM 401から第2のオーディオ装置（図4のオーディオ装置421）へのコマンドおよび第2のオーディオ装置からの応答;

435: CHASM 401から第3のオーディオ装置（図4のオーディオ装置422）へのコマン

10

20

30

40

50

ドおよび第3のオーディオ装置からの応答。

【0214】

図8は、一例による図7に示すルーティング・テーブルの詳細を示す。図8の要素は、以下を含む：

701: CHASMによって維持されるルートのテーブル。この例によれば、各ルートには以下のフィールドがある：

- ・IDまたは「永続的固有セッション識別子」
- ・どのアプリケーションがルートを要求したかの記録；
- ・源；
- ・宛先。この例では、一または複数の人または位置を含みうるが、オーディオ装置を含まない；
- ・優先度；
- ・接続性モード。この例では、同期的モード、スケジュールされたモード、トランザクション・モードを含むモードのリストから選択される；
- ・確認応答が必要とされるかどうかの指示；
- ・どのオーディオ品質側面（単数または複数）を優先するか。これは、本明細書ではオーディオ・セッション目標とも呼ばれる。いくつかの例では、オーディオ・セッション・マネージャまたは優先順位付け器702は、オーディオ・セッションの目標（単数または複数）に従ってオーディオ・セッションを最適化する；

801: ルーティング・テーブル701内のルートは、アプリ410によって要求され、ID 50を割り当てられている。このルートは、アレックス（宛先）が優先度4でSpotifyを聞きたがっていることを指定する。この例では、優先度は整数値で、最も高い優先度は1である。接続性モードは同期的モードであり、それはこの例では実行中であることを意味する。この場合、アレックスが対応する音楽が彼女に提供されているかどうかを確認または確認応答する必要はない。この例では、唯一の指定されたオーディオ・セッション目標は音楽品質である；

802: ルーティング・テーブル701内のルートは、アプリ811によって要求され、ID 51を割り当てられている。アンガスは、優先度4でタイマーアラームを聞くことになっている。このオーディオ・セッションは、CHASM 401によって記憶されるが、ルーティング・テーブル701には示されていない将来の時間のためにスケジュールされる。この例では、アンガスはアラームを聞いたことを確認応答する必要がある。この例では、唯一の指定されたオーディオ・セッション目標は、アンガスがアラームを聞く可能性を高めるための可聴性である；

803: ルーティング・テーブル701内のルートは、アプリ410によって要求され、ID 52を割り当てられている。宛先は「赤ん坊」であるが、基礎になるオーディオ・セッション目標は、オーディオ環境における赤ん坊の近くでの非可聴性である。よって、これは「赤ん坊を起こすな!」実装の一例であり、その詳細な例を後述する。このオーディオ・セッションの優先度は2である（ほとんど何よりも重要）。接続性モードは同期的（進行中）である。起こされない赤ん坊からの確認応答は不要である。この例では、唯一の指定されたオーディオ・セッション目標は、赤ん坊の位置で音が聞こえないことである。

【0215】

804: ルーティング・テーブル701内のルートは、アプリ411によって要求され、ID 53を割り当てられている。この例では、アプリ411は電話アプリである。この場合、ジョージが電話している。ここで、オーディオ・セッションの優先度は3である。接続性モードは同期的（進行中）である。ジョージが依然として電話通話に出ていることの確認応答は必要ない。たとえば、ジョージは、ジョージの準備ができたなら、バーチャル・アシスタントに通話を終了するように依頼するつもりであってもよい。この例では、唯一の指定されたオーディオ・セッション目標は、発話了解性である。

【0216】

805: ルーティング・テーブル701内のルートは、アプリ412によって要求され、ID

54を割り当てられている。この例では、オーディオ・セッションの基礎となる目的は、配管工が玄関にいて、リチャードと話す必要があることをリチャードに伝えることである。接続性モードはトランザクション・モードである：他のオーディオ・セッションの優先度を考慮して、できるだけ早くリチャードに該メッセージを再生する。この例では、リチャードは赤ん坊を寝かせたばかりで、リチャードはまだベビールームにいる。より高い優先度を有するルート803に鑑み、CHASMのオーディオ・セッション・マネージャは、リチャードがベビールームを離れるまで待ってから、ルート805に対応するメッセージが配達されるようにする。この例では、確認応答が必要とされる；この例では、リチャードは口頭で、リチャードがメッセージを聞き、配管工に会いに行く途中であることを確認する必要がある。いくつかの例によれば、リチャードが指定された時間内に確認応答しない場合、CHASMのオーディオ・セッション・マネージャは、リチャードが応答するまで、オーディオ環境のすべてのオーディオ装置（一部の例ではベビールームのオーディオ装置を除く）にこのメッセージを提供させることができる。この例では、唯一の指定されたオーディオ・セッション目標は、リチャードがメッセージを聞き、理解するための、発話了解度である。

【0217】

806: ルーティング・テーブル701内のルートは、火災警報システム・アプリ810によって要求され、ID 55を割り当てられている。このルートの基礎になる目的は、ある種の下で（たとえば、煙検知センサーからの応答に応じて）、家から避難するよう火災警報を鳴らすことである。このルートは、最高の可能な最優度をもつ：赤ん坊を起こすことさえ許容できる。接続性モードは同期的である。確認応答は不要である。この例では、唯一の指定されたオーディオ・セッション目標は「可聴性」である。この例によれば、オーディオ環境内のすべての人がアラームを聞き、避難することを保証するために、CHASMはオーディオ環境のすべてのオーディオ装置を制御して、アラームを大音量で再生する。

【0218】

一部の実装では、オーディオ・セッション・マネージャ（たとえば、CHASM）は、一つまたは複数のメモリ構造内の各ルートに対応する情報を維持する。いくつかのそのような実装によれば、オーディオ・セッション・マネージャは、オーディオ環境における変化する条件（たとえば、人がオーディオ環境における位置を変える）および／またはオーディオ・セッション・マネージャ702からの制御信号に応じて、各ルートに対応する情報を更新するように構成されてもよい。たとえば、ルート801を参照すると、オーディオ・セッション・マネージャは、以下の情報を含む、またはそれに対応するメモリ構造を1つ記憶し、更新することができる：

【表1】

ID	宛先位置	宛先配向	オーディオ装置
50	居間のカウチ	テレビを向く	a, b, c, d, e

表 1

【0219】

表1に示されている情報は、例を提供するために、人間が読むことができるフォーマットである。オーディオ・セッション・マネージャがそのような情報（たとえば、宛先位置と宛先配向）を記憶するために使用する実際のフォーマットは、特定の实装に依存して、人間によって理解可能であっても、なくてもよい。

【0220】

この例では、オーディオ・セッション・マネージャは、ルート801の宛先であるアレックスの位置と配向を監視し、ルート801にオーディオ・コンテンツを提供することのどのオーディオ装置が関わるかを決定するように構成される。そのようないくつかの例によれば、オーディオ・セッション・マネージャは、以下に詳細に記載される方法に従って、オ

オーディオ装置位置、人物の位置および人物の配向を決定するように構成されてもよい。表1の情報が変化した場合、いくつかの実装では、オーディオ・セッション・マネージャは、対応するコマンド/制御信号を、ルート801についてのメディア・ストリームからのオーディオをレンダリングしている装置に送信し、表1を介して描かれたようなメモリ構造を更新する。

【0221】

図9Aは、オーケストレーションの言語におけるルート開始要求のコンテキストフリー文法の例を示す。いくつかの例では、図9Aは、あるルートについての、アプリケーションからCHASMへの要求の文法を表すことができる。ルート開始要求は、ユーザーが、たとえば、携帯電話上のアプリケーションに対応するアイコンを選択することによって、音声コマンドを介してなどで、アプリケーションを選択し、これと対話したことに応じてトリガーされてもよい。

【0222】

この例では、要素901は、要素902A、902B、902C、902Dと組み合わせて、ルート・ソースを定義することを可能にする。要素902A、902B、902Cおよび902Dによって示されるように、この例では、ルート・ソースは、一または複数の人、サービスおよびオーディオ環境位置であってもよく、またはそれらを含むことができる。サービスは、たとえば、クラウドベースのメディア・ストリーミング・サービス、外のドアベルから、またはドアベルに関連するオーディオ装置からのオーディオ・フィードを提供する家庭内サービスなどであってもよい。実装によっては、URL（SpotifyのURLなど）、サービスの名前、自宅のドアベルのIPアドレスなどに応じてサービスが指定されてもよい。オーディオ環境位置は、いくつかの実装では、以下に説明するオーディオ環境ゾーンと対応してもよい。いくつかの例では、オーディオ環境位置ソースは、ゾーン内の一つまたは複数のマイクロフォンに対応しうる。要素902Dのカンマは、複数のソースが指定される可能性があることを示す。たとえば、ルート要求は、「ロジャー、マイケルからのルート」または「Spotifyからのルート」または「キッチンからのルート」または「ロジャーおよびキッチンからのルート」などを示すことができる。

【0223】

この例では、要素903は、要素904A、904B、904Cおよび904Dと組み合わせて、ルート宛先の定義を可能にする。この実装では、ルート宛先は、一または複数の人、サービスおよびオーディオ環境位置であってもよく、または、それらを含んでもよい。たとえば、ルート要求は、「デーヴィッドへのルート」または「キッチンへのルート」または「デッキへのルート」または「ロジャーおよびキッチンへのルート」などを示すことができる。

【0224】

この例では、ルートごとに1つの接続性モードのみが選択されうる。この実装によれば、接続性モード・オプションは、同期的、スケジュールされた、またはトランザクションであるが、いくつかの実装では、ルート毎に複数の接続性モードが選択されてもよい。たとえば、そのような実装では、ルート開始要求は、あるルートがスケジュールされたルートおよびトランザクション・ルートの両方でありうることを示す可能性がある。たとえば、ルート開始要求は、スケジュールされた時間にメッセージがデーヴィッドに送達されるべきであり、デーヴィッドがメッセージに返答すべきであることを示してもよい。図9Aには示されていないが、いくつかの実装では、特定のメッセージが、ルート開始要求に含まれてもよい。たとえば、事前に録音されたメッセージである。

【0225】

この例では、オーディオ・セッション目標は「特質（trait）」と呼ばれる。この例によれば、一つまたは複数のオーディオ・セッション目標は、品質907と一つまたは複数の特質908Aの組み合わせを介して、ルート開始要求で示されてもよい。コンマ908Bは、この例によれば、一つまたは複数の特質が指定できることを示す。しかしながら、代替の実装では、ルート開始要求において、1つのオーディオ・セッション目標のみが示されてもよ

い。

【0226】

図9Bは、オーディオ・セッション目標の例を与えている。この例によれば、「特質」リスト908Aは、重要な一つまたは複数の品質の指定を可能にする。いくつかの実装では、ルート開始要求は、たとえば降順で、複数の特質を指定してもよい。たとえば、ルート開始要求は（品質＝了解度、空間的忠実度）を指定してもよい。つまり、了解度が最も重要な特質であり、空間的忠実度がそれに続く。ルート開始要求は（品質＝可聴性）を指定してもよい。つまり、唯一の音声セッション目標は、人々がたとえば警報を聞くことができることである。

【0227】

ルート開始要求は、（品質＝非可聴性）を指定してもよい。つまり、唯一のオーディオ・セッション目標は、ルート宛先として指定された人（たとえば、赤ん坊）が、オーディオ環境において再生されるオーディオを聞かないことである。これは、「赤ん坊を起こすな」実装のためのルート開始要求の例である。

【0228】

別の例では、ルート開始要求は（品質＝可聴性、プライバシー）を指定してもよい。これは、たとえば、一次オーディオ・セッション目標は、ルートの宛先として指定された人が送達されるオーディオを聞くことであるが、二次オーディオ・セッション目標は、ルートに従って送達および／または交換されるオーディオを他の人が聞くことができる程度を制限することであることを意味する。本明細書の他の箇所に記載されているように、後者のオーディオ・セッション目標は、ルート宛先とオーディオ環境内の1人以上の他の人との間で白色雑音または他のマスキングノイズを再生すること、オーディオ環境内の1人以上の他の人の近くで再生される他のオーディオの音量を増加させることなどによって達成することができる。

【0229】

ここで図9Aに戻ると、この例では、ルート開始要求は、要素909および910を介して優先度を指定することができる。いくつかの例において、優先度は、有限個の整数（たとえば、3、4、5、6、7、8、9、10など）のうちの整数を介して示されてもよい。いくつかの例において、1は最も高い優先度を示すことができる。

【0230】

この例によれば、ルート開始要求は、要素911を介した確認応答を任意的に指定することができる。たとえば、ルート開始要求は、「リチャードが夕食の準備ができたと言っていることをマイケルに伝え、確認応答を受け取る」ことを示してもよい。応答して、いくつかの例では、オーディオ・セッション・マネージャは、マイケルの位置を決定しようと試みることがある。たとえば、CHASMはマイケルがガレージにいると推測するかもしれない。なぜなら、そこでマイケルの声が最後に検出されたからである。よって、オーディオ・セッション・マネージャは、ガレージ内の一つまたは複数のスピーカーを介して「夕食の準備ができました。このメッセージを聞いたことを確認してください」というアナウンスを生じてもよい。マイケルが応答した場合、オーディオ・セッション・マネージャは該応答をリチャードに報告/再生させることができる。ガレージアナウンスに対するマイケルからの応答がない場合（たとえば、10秒後）、オーディオ・セッション・マネージャは、マイケルにとって2番目に可能性が高い位置でアナウンスを行わせることができる。該位置はたとえば、マイケルが多くの時間を費やす場所、または以前のガレージでの発声の前にマイケルの声が聞こえた最後の場所などである。その場所がマイケルの寝室であるとする。マイケルの寝室でのアナウンスに対してマイケルからの応答がない場合（たとえば、10秒後）、オーディオ・セッション・マネージャは、「赤ん坊を起こすな」などの他の制約条件のもとで、環境の多くのスピーカーにアナウンスを行わせてもよい。

【0231】

図10は、一例に従った、ルートを修正する要求のフローを示している。ルート修正要求は、たとえば、アプリケーションによって送信され、オーディオ・セッション・マネージ

10

20

30

40

50

によって受領されてもよい。ルート修正要求は、たとえば、ユーザーがアプリケーションを選択し、これと対話したことに応じてトリガーされてもよい。

【0232】

この例では、ID 1002は、ルート開始要求に応答して、オーディオ・セッション・マネージャが以前にアプリに提供したであろう永続的固有オーディオ・セッション番号またはコードを指す。この例によれば、接続性モードの変更が、要素1003および要素1004A、1004Bまたは1004Cを介してなされてもよい。あるいはまた、素子1004A、1004B、および1004Cは、接続性モードの変更が望まれない場合、バイパスされてもよい。

【0233】

この例によれば、一つまたは複数のオーディオ・セッション目標は、要素1005、1006A、および1006Bを介して変更されてもよい。あるいはまた、要素1005、1006A、および1006Bは、オーディオ・セッション目標変更が望まれない場合にはバイパスされてもよい。

【0234】

この例では、ルート優先度は、要素1007および1008を介して変更されてもよい。あるいはまた、要素1007および1008は、ルート優先度変更が望まれない場合にはバイパスされてもよい。

【0235】

この例によれば、要素1009または要素1011は、確認応答要求変更を行うために使用されてもよい。たとえば、要素1009は、確認応答が以前にルートについて要求されなかった場合に、確認応答が追加されうることを示す。反対に、要素1011は、確認応答が以前にルートについて要求されていた場合に、確認応答が除去されうることを示す。要素1010のセミコロンは、ルートを修正する要求の終了を示す。

【0236】

図11Aおよび11Bは、ルートを修正する要求についてのフローの追加的な例を示す。図11Cは、ルートを削除するためのフローの例を示す。ルート修正または削除要求は、たとえば、アプリケーションによって送信され、オーディオ・セッション・マネージャによって受領されてもよい。ルート修正要求は、たとえば、ユーザーがアプリケーションを選択し、これと対話したことに応じてトリガーされてもよい。

【0237】

図11Aおよび11Bにおいて、「シンク」は、ルート宛先を指す。本明細書に開示されている他のフロー図と同様に、図11A～11Cに示されている動作は、必ずしも示された順序で実行されるものではない。たとえば、いくつかの実装では、ルートIDは、フローにおいて早期に、たとえばフローの先頭において指定されてもよい。

【0238】

図11Aは、ソースまたは宛先を追加するためのフロー1100Aを示す。いくつかの例では、一つまたは複数のソースまたは宛先が追加されてもよい。この例では、ルート・ソースは、要素1101および1102Aを介して追加されてもよい。この例によれば、ルート宛先は、要素1101および1102Bを介して追加されてもよい。この例では、ルート・ソースまたは宛先が追加されるルートは、要素1103および1104を介して示される。この例によれば、人が、要素1105Aを介してソースまたは宛先として追加されてもよい。この例では、サービスが、要素1105Bを介してソースまたは宛先として追加されてもよい。この例によれば、位置が、要素1105Cを介してソースまたは宛先として追加されてもよい。要素1106は、一つまたは複数のソースまたは宛先を追加するためのフローの終了を示す。

【0239】

図11Bは、ソースまたは宛先を除去するためのフロー1100Bを示す。場合によっては、一つまたは複数のソースまたは宛先が除去されてもよい。この例では、ルート・ソースは、要素1107および1108Aを介して除去されてもよい。この例によれば、ルート宛先は、要素1107および1108Bを介して除去されてもよい。この例では、ルート・ソースまたは宛先が除去されるルートは、要素1109および1110を介して示される。この例によれば、

10

20

30

40

50

人が、要素1111Aを介して、ソースまたは宛先として除去されてもよい。この例では、サービスが、要素1111Bを介してソースまたは宛先として除去されてもよい。この例によれば、位置が、要素1111Cを介してソースまたは宛先として除去されてもよい。要素1112は、一つまたは複数のソースまたは宛先を除去するためのフローの終了を示す。

【0240】

図11Cは、ルートを削除するためのフローを示す。ここで、要素1113が削除を示す。要素1114を介して指定されたルートIDは、削除されるべきルートを示す。要素1115は、一つまたは複数のソースまたは宛先を除去するためのフローの終了を示す。

【0241】

図12は、いくつかの実装に従ったオーディオ・セッション管理方法のブロックを含むフロー図である。この例によれば、方法1200は、複数のオーディオ装置を有するオーディオ環境についてのオーディオ・セッション管理方法である。方法1200のブロックは、本明細書に記載される他の方法と同様に、必ずしも示される順序で実施されるわけではない。いくつかの実装では、方法1200の一つまたは複数のブロックは、同時に実行されてもよい。さらに、方法1200のいくつかの実装は、図示および／または説明されるよりも多いまたは少ないブロックを含んでいてもよい。方法1200のブロックは、一つまたは複数の装置（図6に示されかつ以下に説明される制御システム610などの制御システム、または他の開示された制御システム例の1つであってもよい（またはそれを含んでいてもよい））によって実行されてもよい。

【0242】

この例では、ブロック1205は、第1のアプリケーションを実装する第1の装置から、オーディオ・セッション・マネージャ（たとえば、CHASM）を実装する装置によって、第1のオーディオ・セッションのための第1のルートを開始するための第1のルート開始要求を受領することを含む。この例によれば、第1のルート開始要求は、第1のオーディオソースと第1のオーディオ環境宛先を示す。ここで、第1のオーディオ環境宛先は、オーディオ環境における少なくとも第1の人物と対応している。ただし、この例では、第1のオーディオ環境宛先はオーディオ装置を示さない。

【0243】

いくつかの例によれば、第1のルート開始要求は、第1のルート・ソースまたは第1のルート宛先として、オーディオ環境の少なくとも第1のエリアを示してもよい。場合によっては、第1のルート開始要求は、少なくとも第1のサービスを、第1のオーディオソースとして示すことができる。

【0244】

この実装では、ブロック1210は、オーディオ・セッション・マネージャを実装する装置によって、第1のルート開始要求に対応する第1のルートを確認することを含む。この例では、第1のルートを確認することは、オーディオ環境内の少なくとも第1の人の第1の位置を決定し、第1のオーディオ・セッションの第1のステージについて少なくとも1つのオーディオ装置を決定し、および第1のオーディオ・セッションを開始またはスケジュールすることを含む。

【0245】

いくつかの例によれば、第1のルート開始要求は、第1のオーディオ・セッション優先度を含んでいてもよい。いくつかの例では、第1のルート開始要求は第1の接続性モードを含むことができる。第1の接続性モードは、たとえば、同期的接続モード、トランザクション接続性モード、またはスケジュールされた接続性モードであってもよい。いくつかの例では、第1のルート開始要求は複数の接続性モードを示してもよい。

【0246】

いくつかの実装では、第1のルート開始要求は、少なくとも第1の人から確認応答が要求されるかどうかの指示を含んでいてもよい。いくつかの例では、第1のルート開始要求は、第1のオーディオ・セッション目標を含んでいてもよい。第1のオーディオ・セッション目標は、たとえば、了解度、オーディオ品質、空間的忠実度および／または非可聴性を含

10

20

30

40

50

むことができる。

【0247】

本明細書の他の箇所に記載されているように、いくつかの実装では、ルートは、関連するオーディオ・セッション識別子を有してもよく、それは、いくつかの実装では、永続的固有オーディオ・セッション識別子であってもよい。よって、方法1200のいくつかの実装は、第1のルートについての第1の永続的固有オーディオ・セッション識別子を（たとえば、オーディオ・セッション・マネージャによって）決定し、第1の永続的固有オーディオ・セッション識別子を第1の装置（第1のアプリケーションを実行している装置）に送信することを含んでいてもよい。

【0248】

いくつかの実装では、第1のルートを確立することは、環境中の少なくとも1つの装置に、第1のルートに対応する少なくとも第1のメディア・ストリームを確立させることを含んでいてもよく、第1のメディア・ストリームは第1のオーディオ信号を含む。方法1200のいくつかの実装は、第1のオーディオ信号を第1にレンダリングされたオーディオ信号にレンダリングさせることを含んでいてもよい。いくつかの例では、方法1200は、オーディオ・セッション・マネージャが、オーディオ環境の別の装置に、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングさせることを含んでいてもよい。しかしながら、いくつかの実装では、オーディオ・セッション・マネージャは、第1のオーディオ信号を受領し、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングするように構成されうる。

【0249】

本明細書の他の箇所に記載されているように、いくつかの実装においては、オーディオ・セッション・マネージャ（たとえば、CHASM）は、オーディオ環境における1人以上の人々の位置および／または配向、オーディオ環境におけるオーディオ装置の位置などといった、オーディオ環境の条件をモニターすることができる。たとえば、「赤ん坊を起こすな」の使用事例については、オーディオ・セッション・マネージャ（たとえば、図7の最適化器702）は、赤ん坊がどこにいるかを決定するか、または少なくとも推定することができる。オーディオ・セッション・マネージャは、関連するアプリケーションから「オーケストレーションの言語」で送信されたユーザーからの明示的な陳述（「赤ん坊を起こすな。赤ん坊は寝室1にいる」など）のために、赤ん坊がどこにいるを知ることができる。代替的または追加的に、オーディオ・セッション・マネージャは、赤ん坊がどこにいるかを、以前の明示的な入力に基づいて、または赤ん坊の泣き声の以前の検出に基づいて、決定することができる（たとえば、以下に説明されるように）。いくつかの例では、オーディオ・セッション・マネージャは、この制約条件を受け取り（たとえば、「非可聴性」オーディオ・セッション目標を介して）、該制約条件を実施してもよい。これはたとえば、赤ん坊の位置の音圧レベルが閾値デシベルレベル（たとえば、50dB）未満であることを保証することによる。

【0250】

方法1200のいくつかの例は、オーディオ・セッションの第1ステージに対する第1の人の第1の配向を決定することを含んでいてもよい。いくつかのそのような例によれば、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングさせることは、第1の人の第1の位置および第1の配向に対応する第1の基準空間モードを決定することと、第1の基準空間モードに対応するオーディオ環境におけるラウドスピーカーの第1の相対的アクティブ化を決定することとを含むことができる。以下に、いくつかの詳細な例を示す。

【0251】

場合によっては、オーディオ・セッション・マネージャは、第1の人が位置および／または配向を変更したと判断してもよい。方法1200のいくつかの例は、第1の人の第2の位置または第2の配向のうちの少なくとも1つを決定し、第2の位置または第2の配向のうちの少なくとも1つに対応する第2の基準空間モードを決定し、第2の基準空間モードに対応

10

20

30

40

50

するオーディオ環境におけるラウドスピーカーの第2の相対的アクティブ化を決定することを含んでいてもよい。

【0252】

本開示の他の箇所に記載されているように、オーディオ・マネージャは、場合によっては、一度に複数のルートを確認し、実施することを任されることがある。方法1200のいくつかの例は、第2のアプリケーションを実装する第2の装置から、オーディオ・セッション・マネージャを実装する装置によって、第2のオーディオ・セッションのための第2のルートを開始するための第2のルート開始要求を受領することを含んでいてもよい。第1のルート開始要求は、第2のオーディオソースと第2のオーディオ環境宛先を示してもよい。いくつかの例では、第2のオーディオ環境宛先は、オーディオ環境内の少なくとも第2の人と対応することができる。しかしながら、場合によっては、第2のオーディオ環境宛先が、第2のルートに関連する特定のオーディオ装置を示さなくてもよい。

10

【0253】

方法1200のいくつかのそのような例は、オーディオ・セッション・マネージャを実装する装置によって、第2のルート開始要求に対応する第2のルートを確認することを含んでいてもよい。場合によっては、第2のルートを確認することは、オーディオ環境内の少なくとも第2の人の第1の位置を決定し、第2のオーディオ・セッションの第1のステージについて少なくとも1つのオーディオ装置を決定し、第2のオーディオ・セッションを開始することを含みうる。

【0254】

いくつかの例によれば、第2のルートを確認することは、第2のルートに対応する少なくとも第2のメディア・ストリームを確認することを含んでいてもよい。第2のメディア・ストリームは、第2のオーディオ信号を含んでいてもよい。方法1200のいくつかのそのような例は、第2のオーディオ信号を第2のレンダリングされたオーディオ信号にレンダリングさせることを含んでいてもよい。

20

【0255】

方法1200のいくつかの例は、少なくとも部分的に、第2のオーディオ信号、第2のレンダリングされたオーディオ信号またはそれらの特性のうちの少なくとも1つに基づいて、第1のオーディオ信号のレンダリング・プロセスを修正して、修正された第1のレンダリングされたオーディオ信号を生成することを含んでいてもよい。第1のオーディオ信号のレンダリング・プロセスを修正することは、たとえば、第1のオーディオ信号のレンダリングを、第2のレンダリングされたオーディオ信号のレンダリング位置から離れるようにワープさせることを含むことができる。代替的または追加的に、第1のオーディオ信号のレンダリング・プロセスを修正することは、第2のオーディオ信号または第2のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスに応答して、第1のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスを修正することを含んでいてもよい。

30

【0256】

図13は、いくつかの実装によるオーディオ・セッション管理方法のブロックを含むフロー図である。この例によれば、方法1300は、複数のオーディオ装置を有するオーディオ環境についてのオーディオ・セッション管理方法である。方法1300のブロックは、本明細書に記載される他の方法と同様に、必ずしも示される順序で実行されるわけではない。いくつかの実装では、方法1300の一つまたは複数のブロックは、同時に実行されてもよい。さらに、方法1300のいくつかの実装は、図示および／または説明されるよりも多いまたは少ないブロックを含んでいてもよい。方法1300のブロックは、たとえば、図6に示され、以下に説明される制御システム610などの制御システム、または他の開示された制御システム例の1つ、であってもよい（または、それを含んでいてもよい）一つまたは複数の装置によって実行されてもよい。

40

【0257】

この例では、ブロック1305は、第1のアプリケーションを実装する第1の装置から、オ

50

オーディオ・セッション・マネージャ（たとえば、CHASM）を実装する装置によって、第1のオーディオ・セッションのための第1のルートを開始するための第1のルート開始要求を受領することを含む。この例によれば、第1のルート開始要求は、第1のオーディオソースと第1のオーディオ環境宛先を示す。ここで、第1のオーディオ環境宛先は、少なくともオーディオ環境の第1の領域に対応する。ただし、この例では、第1のオーディオ環境宛先はオーディオ装置を示さない。

【0258】

いくつかの例によれば、第1のルート開始要求は、オーディオ環境内の少なくとも第1の人を、第1のルート・ソースまたは第1のルート宛先として示すことができる。場合によっては、第1のルート開始要求は、少なくとも第1のサービスを、第1のオーディオソースとして示すことができる。

10

【0259】

この実装では、ブロック1310は、オーディオ・セッション・マネージャを実装する装置によって、第1のルート開始要求に対応する第1のルートを確立することを含む。この例では、第1のルートを確立することは、第1のオーディオ・セッションの第1のステージのために、オーディオ環境の第1のエリアにある少なくとも1つのオーディオ装置を決定し、第1のオーディオ・セッションを開始またはスケジュールすることを含む。

【0260】

いくつかの例によれば、第1のルート開始要求は、第1のオーディオ・セッション優先度を含んでいてもよい。いくつかの例では、第1のルート開始要求は第1の接続性モードを含むことができる。第1の接続性モードは、たとえば、同期的接続性モード、トランザクション接続性モード、またはスケジュールされた接続性モードであってもよい。いくつかの例では、第1のルート開始要求が複数の接続性モードを示してもよい。

20

【0261】

いくつかの実装では、第1のルート開始要求は、少なくとも第1の人からの確認応答が要求されるかどうかの指示を含んでいてもよい。いくつかの例では、第1のルート開始要求は、第1のオーディオ・セッション目標を含んでいてもよい。第1のオーディオ・セッション目標は、たとえば、了解度、オーディオ品質、空間的忠実度および／または非可聴性を含むことができる。

【0262】

方法1300のいくつかの実装は、第1のルートについての第1の永続的固有オーディオ・セッション識別子を決定し（たとえば、オーディオ・セッション・マネージャによって）、第1の永続的固有オーディオ・セッション識別子を第1の装置（第1のアプリケーションを実行している装置）に送信することを含んでいてもよい。

30

【0263】

いくつかの実装では、第1のルートを確立することは、環境中の少なくとも1つの装置に、第1のルートに対応する少なくとも第1のメディア・ストリームを確立させることを含んでいてもよく、第1のメディア・ストリームは第1のオーディオ信号を含む。方法1300のいくつかの実装は、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングさせることを含んでいてもよい。いくつかの例では、方法1300は、オーディオ・セッション・マネージャが、オーディオ環境の別の装置に、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングさせることを含んでいてもよい。しかしながら、いくつかの実装では、オーディオ・セッション・マネージャは、第1のオーディオ信号を受領し、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングするように構成されてもよい。

40

【0264】

本明細書の他の箇所に記載されているように、いくつかの実装では、オーディオ・セッション・マネージャ（たとえば、CHASM）は、オーディオ環境内の一つまたは複数のオーディオ装置の位置など、オーディオ環境の状態をモニターすることができる。

【0265】

50

方法1300のいくつかの例は、第1の時点において、オーディオ環境の第1のエリア内の複数のオーディオ装置の各オーディオ装置の第1の位置を自動的に決定する、第1のラウドスピーカー自動位置特定プロセスを実行することを含んでいてもよい。いくつかのそのような例では、レンダリング・プロセスは、少なくとも部分的に、各オーディオ装置の第1の位置に基づいてもよい。いくつかのそのような例は、第1のルートに関連するデータ構造内に各オーディオ装置の第1の位置を記憶することを含んでいてもよい。

【0266】

場合によっては、オーディオ・セッション・マネージャは、第1のエリア内の少なくとも1つのオーディオ装置が、変化した位置をもつと判断することがある。いくつかのそのような例は、変更された位置を自動的に決定し、少なくとも部分的には該変更された位置に基づいてレンダリング・プロセスを更新する、第2のラウドスピーカー自動位置特定プロセスを実行することを含みうる。いくつかのそのような実装は、第1のルートに関連するデータ構造において、変更された位置を記憶することを含んでいてもよい。

10

【0267】

場合によっては、オーディオ・セッション・マネージャは、少なくとも1つの追加的なオーディオ装置が第1のエリアに移動されたと判断することがある。いくつかのそのような例は、追加的なオーディオ装置の追加的なオーディオ装置位置を自動的に決定し、少なくとも部分的には、該追加的なオーディオ装置位置に基づいて、レンダリング・プロセスを更新する第2のスピーカー自動位置特定プロセスを実行することを含みうる。そのような実装のいくつかは、第1のルートに関連するデータ構造において追加的なオーディオ装置位置を格納することを含んでいてもよい。

20

【0268】

本明細書の他の箇所に記載されているように、いくつかの例において、第1のルート開始要求は、第1のルート・ソースまたは第1のルート宛先として、少なくとも第1の人を示すことができる。方法1300のいくつかの例は、オーディオ・セッションの第1ステージについての第1の人の第1の配向を決定することを含んでいてもよい。いくつかのそのような例によれば、第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングさせることは、第1の人の第1の位置および第1の配向に対応する第1の基準空間モードを決定し、第1の基準空間モードに対応するオーディオ環境におけるラウドスピーカーの第1の相対的作動を決定することを含むことができる。以下に、いくつかの詳細な例を示す。

30

【0269】

場合によっては、オーディオ・セッション・マネージャは、第1の人が位置および／または配向を変更したと判断してもよい。方法1300のいくつかの例は、第1の人の第2の位置または第2の配向のうちの少なくとも1つを決定し、第2の位置または第2の配向のうちの少なくとも1つに対応する第2の基準空間モードを決定し、第2の基準空間モードに対応するオーディオ環境内のラウドスピーカーの第2の相対的アクティブ化を決定することを含んでいてもよい。

【0270】

本開示の他の箇所に記載されているように、オーディオ・マネージャは、場合によっては、一度に複数のルートを確認し、実装することを任されることがある。方法1300のいくつかの例は、第2のアプリケーションを実装する第2の装置から、オーディオ・セッション・マネージャを実装する装置によって、第2のオーディオ・セッションのための第2のルートを開始するための第2のルート開始要求を受領することを含んでいてもよい。第1のルート開始要求は、第2のオーディオソースと第2のオーディオ環境宛先を示してもよい。いくつかの例では、第2のオーディオ環境宛先は、オーディオ環境内の少なくとも第2の人と対応してもよい。しかしながら、場合によっては、第2のオーディオ環境宛先が、第2のルートに関連するいかなる特定のオーディオ装置をも示さなくてもよい。

40

【0271】

方法1300のいくつかのそのような例は、オーディオ・セッション・マネージャを実装

50

する装置によって、第2のルート開始要求に対応する第2のルートを確認することを含んでもよい。場合によっては、第2のルートを確認することは、オーディオ環境内の少なくとも第2の人の第1の位置を決定し、第2のオーディオ・セッションの第1のステージについて少なくとも1つのオーディオ装置を決定し、第2のオーディオ・セッションを開始することを含みうる。

【0272】

いくつかの例によれば、第2のルートを確認することは、第2のルートに対応する少なくとも第2のメディア・ストリームを確認することを含んでもよい。第2のメディア・ストリームは、第2のオーディオ信号を含んでもよい。方法1300のいくつかのそのような例は、第2のオーディオ信号を第2のレンダリングされたオーディオ信号にレンダリングさせることを含んでもよい。

10

【0273】

方法1300のいくつかの例は、少なくとも部分的に、第2のオーディオ信号、第2のレンダリングされたオーディオ信号またはそれらの特性のうちの少なくとも1つに基づいて、第1のオーディオ信号のレンダリング・プロセスを修正して、修正された第1のレンダリングされたオーディオ信号を生成することを含んでもよい。第1のオーディオ信号のレンダリング・プロセスを修正することは、たとえば、第1のオーディオ信号のレンダリングを、第2のレンダリングされたオーディオ信号のレンダリング位置から離れるようにワープさせることを含むことができる。代替的または追加的に、第1のオーディオ信号のレンダリング・プロセスを修正することは、第2のオーディオ信号または第2のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスにตอบสนองして、第1のレンダリングされたオーディオ信号のうちの一つまたは複数のもののラウドネスを修正することを含んでもよい。

20

【0274】

図14は、いくつかの実装に従ったオーディオ・セッション管理方法のブロックを含むフロー図である。この例によれば、方法1400は、複数のオーディオ装置を有するオーディオ環境についてのオーディオ・セッション管理方法である。方法1400のブロックは、本明細書に記載される他の方法と同様に、必ずしも示される順序で実行されるわけではない。いくつかの実装では、方法1400の一つまたは複数のブロックは、同時に実行されてもよい。さらに、方法1400のいくつかの実装は、図示および／または説明されるよりも多いまたは少ないブロックを含んでもよい。方法1400のブロックは、たとえば、図6に示され、以下に説明される制御システム610などの制御システム、または他の開示された制御システム例の1つ、であってもよい（または、それを含んでもよい）一つまたは複数の装置によって実行されてもよい。

30

【0275】

この例では、ブロック1405において、図4のアプリケーション410は、オーケストレーションの言語を使用して、CHASM 401に命令する。ブロック1405は、たとえば、アプリケーション410がCHASM 401にルート開始要求またはルート修正要求を送信することを含むことができる。

【0276】

40

この例によれば、CHASM 401は、アプリケーション410から受け取った命令にตอบสนองして、最適なメディア・エンジン制御情報を決定する。この例では、最適なメディア・エンジン制御情報は、少なくとも部分的には、オーディオ環境内の聴取者の位置、オーディオ環境内のオーディオ装置の利用可能性、およびアプリケーション410からの命令に示されているオーディオ・セッション優先度に基づいている。いくつかの例では、最適なメディア・エンジン制御情報は、たとえば、関連するオーディオ装置によって共有される装置属性記述子を介して、CHASM 401によって決定されるメディア・エンジン能力に少なくとも部分的に基づいてもよい。いくつかの例によれば、最適なメディア・エンジン制御情報は、少なくとも部分的には、聴取者の配向に基づくことができる。

【0277】

50

この例では、ブロック415は、制御情報を一つまたは複数のオーディオ装置メディア・エンジンに送信することを含む。制御情報は、図5を参照して上述したオーディオ・セッション管理制御信号に対応する。

【0278】

この例によれば、ブロック1420は、CHASM 401がオーディオ環境内の状態およびこの特定のルートに関するアプリケーション410からの可能なさらなる通信をモニターし、ルート優先度の変化、オーディオ装置の位置の変化、聴取者の位置の変化など、何らかの有意な変化があったかどうかを判断することを表す。もしそうであれば、プロセスはブロック1410に戻り、ブロック1410の諸プロセスが、新しいパラメータに従って実行される。もしそうでなければ、CHASM 401はブロック1420のモニタリングプロセスを継続する。

10

【0279】

図15は、いくつかの実装に従った、オーディオ環境に新たに導入される一つまたは複数のオーディオ装置のための自動セットアップ・プロセスのブロックを含むフロー図である。この例では、一部または全部のオーディオ装置が新規である。方法1500のブロックは、本明細書に記載される他の方法と同様に、必ずしも示される順序で実行されるわけではない。いくつかの実装では、方法1500の一つまたは複数のブロックは、同時に実行されてもよい。さらに、方法1500のいくつかの実装は、図示および／または説明されるよりも多いまたは少ないブロックを含んでいてもよい。方法1500のブロックは、たとえば、図6に示され、以下に説明される制御システム610などの制御システム、または他の開示された制御システム例の1つ、であってもよい（または、それを含んでいてもよい）一つまたは複数の装置によって実行されてもよい。

20

【0280】

この例では、新しいオーディオ装置は、ブロック1505において開梱され、電源投入される。ブロック1510の例では、新しいオーディオ装置のそれぞれは、他のオーディオ装置を探し、特にオーディオ環境のCHASMを探すために、発見〔ディスカバリー〕モードに入る。既存のCHASMが発見された場合、新しいオーディオ装置は、CHASMと通信し、各新しいオーディオ装置の能力に関する情報をCHASMと共有する、などするように構成されてもよい。

【0281】

しかしながら、この例によれば、既存のCHASMは発見されない。よって、ブロック1510のこの例では、新しいオーディオ装置の1つがCHASMとして自身を構成する。この例では、最も多くの利用可能な計算能力および／または最大の接続性を有する新しいオーディオ装置が、新しいCHASM 401として自身を構成する。

30

【0282】

この例では、ブロック1515において、新しい非CHASMオーディオ装置はすべて、新しく任命されたCHASM 401である他の新しいオーディオ装置と通信する。この例によれば、新しいCHASM 401は、この例の図4のアプリケーション412である「セットアップ」アプリケーションを起動する。この場合、セットアップ・アプリケーション412は、たとえば、オーディオおよび／または視覚的プロンプトを介してユーザーと対話して、セットアップ・プロセスを通じてユーザーをガイドするように構成される。

40

【0283】

この例によれば、ブロック1520において、セットアップ・アプリケーション412は、オーケストレーションの言語でCHASM 401に命令を送信する。これは、「すべての新しい装置をセットアップする」ことを示し、最も高いレベルの優先度を有する。

【0284】

この例では、ブロック1525において、CHASM 401は、セットアップ・アプリケーション412からの命令を解釈（インタープリット）し、新しい音響マッピング校正が必要であると判断する。この例によれば、音響マッピング・プロセスはブロック1525で開始され、この例では図4のメディア・エンジン440、441および442である新しい非CHASMオ

50

オーディオ装置のメディア・エンジンとCHASM 401との間の通信を介して、ブロック1530において完了する。本明細書で使用される用語「音響マッピング」は、オーディオ環境のすべての発見可能なラウドスピーカー位置の推定を含む。音響マッピング・プロセスは、たとえば、以下に詳細に記載されるようなラウドスピーカー自動位置特定プロセスを含むことができる。いくつかの例では、音響マッピング・プロセスは、ラウドスピーカー能力情報および／または個々のラウドスピーカー・ダイナミクス処理情報を発見するプロセスを含むことができる。

【0285】

この例によれば、ブロック1535において、CHASM 401は、セットアップ・プロセスが完了したことの確認をアプリケーション412に送る。この例では、ブロック1540において、アプリケーション412は、セットアップ・プロセスが完了したことをユーザーに示す。

【0286】

図16は、いくつかの実装に従った、バーチャル・アシスタント・アプリケーションをインストールするためのプロセスのブロックを含むフロー図である。いくつかの例において、方法1700は、方法1500のセットアップ・プロセスの後に実行されてもよい。この例では、方法1600は、図4に示されるオーディオ環境のコンテキストでバーチャル・アシスタント・アプリケーションをインストールすることを含む。方法1600のブロックは、本明細書に記載される他の方法と同様に、必ずしも示される順序で実行されるわけではない。いくつかの実装では、方法1600の一つまたは複数のブロックは、同時に実行されてもよい。さらに、方法1600のいくつかの実装は、図示および／または説明されるよりも多いまたは少ないブロックを含んでいてもよい。方法1600のブロックは、たとえば、図6に示され、以下に説明される制御システム610などの制御システム、または他の開示された制御システム例の1つ、であってもよい（または、それを含んでいてもよい）一つまたは複数の装置によって実行されてもよい。

【0287】

この例では、「バーチャル・アシスト連携（Virtual Assisting Liaison）」またはVALと呼ばれる新しいアプリケーション411が、ブロック1605でユーザーによってインストールされる。いくつかの例によれば、ブロック1605は、インターネットを介して一つまたは複数のサーバーから、アプリケーション411を携帯電話のようなオーディオ装置にダウンロードすることを含んでいてもよい。

【0288】

この実装によれば、ブロック1610において、アプリケーション411は、CHASM 401に対して、オーケストレーションの言語で、新しいウェイクワード「Hey Val」（やあ、ヴァル）を継続的に聴くことを、最高の優先度で、永続的なオーディオ・セッションとして行うように命令する。この例では、ブロック1615において、CHASM 401は、アプリケーション411からの命令を解釈し、メディア・エンジン440、441、および442に、ウェイクワード「Hey Val」があるかどうか傾聴しているように自分のウェイクワード検出器を構成し、ウェイクワード「Hey Val」が検出されるときは常にCHASM 401にコールバックを発するように命令する。この実装では、ブロック1620において、メディア・エンジン440、441、および442は、ウェイクワードがあるかどうか傾聴し続ける。

【0289】

この例では、ブロック1625において、CHASM 401は、メディア・エンジン440および441から、ウェイクワード「Hey Val」が検出されたことを示すコールバックを受領する。応答して、CHASM 401は、メディア・エンジン440、441、442に、ウェイクワードが初期に検出された後の閾値時間期間（この例では5秒）の間、コマンドがあるかどうか傾聴するように、また、コマンドが検出された場合には、コマンドが検出されたエリア内のオーディオのボリュームを「ダックさせる（duck）」または下げるように命令する。

【0290】

この例によれば、ブロック1630において、メディア・エンジン440、441、および44

10

20

30

40

50

2はすべて、コマンドを検出し、CHASM 401に、検出されたコマンドに対応する発話オーディオ・データおよび確率を送信する。この例では、ブロック1630において、CHASM 401は、検出されたコマンドに対応する発話オーディオ・データおよび確率をアプリケーション411に転送する。

【0291】

この実装では、ブロック1635において、アプリケーション411は、検出されたコマンドに対応する発話オーディオ・データおよび確率を受領し、これらのデータを処理のためにクラウドベースの発話認識アプリケーションに転送する。この例では、ブロック1635において、クラウドベースの発話認識アプリケーションは、この例ではコマンドに対応する一つまたは複数の単語を含む、発話認識プロセスの結果を、アプリケーション411に送信する。ここで、ブロック1635において、アプリケーション411は、オーケストレーションの言語で、CHASM 401に、発話認識セッションを終了するように命令する。この例によれば、CHASM 401は、メディア・エンジン440、441、および442に対して、コマンドがあるかどうか聴取するのを停止するように命令する。

【0292】

図17は、いくつかの実装によるオーディオ・セッション管理方法のブロックを含むフロー図である。この例によれば、方法1700は、図4のオーディオ環境において音楽アプリケーションを実装するためのオーディオ・セッション管理方法である。場合によっては、方法1700は、方法1500のセットアップ・プロセス後に実行されてもよい。いくつかの例において、方法1700は、図16を参照して上述したバーチャル・アシスタント・アプリケーションをインストールするプロセスの前または後に実行されてもよい。方法1700のブロックは、本明細書に記載される他の方法と同様に、必ずしも示される順序で実行されるわけではない。いくつかの実装では、方法1700の一つまたは複数のブロックは、同時に実行されてもよい。さらに、方法1700のいくつかの実装は、図示および／または説明されるよりも多いまたは少ないブロックを含んでいてもよい。方法1700のブロックは、たとえば、図6に示され、以下に説明される制御システム610などの制御システム、または他の開示された制御システム例の1つ、であってもよい（または、それを含んでいてもよい）一つまたは複数の装置によって実行されてもよい。

【0293】

この例では、ブロック1705において、ユーザーは、オーディオ環境内のある装置上で実行されている音楽アプリケーションに入力を提供する。この場合、音楽アプリケーションは、図4のアプリケーション410である。この例によれば、アプリケーション410は、スマートフォン上で動作しており、入力は、タッチセンサーシステムおよび／またはジェスチャーセンサーシステムのようなスマートフォンのユーザーインターフェースを介して提供される。

【0294】

この例によれば、ブロック1710において、アプリケーション410は、スマートフォンを介してアプリケーション410と対話しているユーザーへのクラウドベースの音楽サービスからのルートを開始するように、この例ではオーケストレーション言語でのルート開始要求を介して、CHASM 401に命令する。この例では、ルート開始要求は、クラウドベースの音楽サービスの、ユーザーの現在のお気に入りプレイリストを使用しての最高の音楽再生品質をオーディオ・セッション目標とし、確認応答が要求されず、優先度4で、同期的モードを示す。

【0295】

この例では、ブロック1715において、CHASM 401は、ブロック1710で受領した命令に従って、オーディオ環境のどのオーディオ装置がルートに関与するかを決定する。この決定は、少なくとも部分的には、最高の音楽再生品質をオーディオ・セッション目標とし、どのオーディオ装置が現在利用可能であるか、利用可能なオーディオ装置の能力、およびユーザーの推定される現在位置に基づいてもよい。いくつかの例において、ブロック1715の決定は、少なくとも部分的に、ユーザーの推定される現在の配向に基づいてもよ

10

20

30

40

50

い。いくつかの実装では、ブロック1715で、名目または初期聴取レベルが選択されてもよい。このレベルは、少なくとも部分的には、一つまたは複数のオーディオ装置に対するユーザーの推定される近接性、ユーザーのエリア内の周囲ノイズレベルなどに基づいてもよい。

【0296】

この例によれば、ブロック1720において、CHASM 401は、アプリケーション410によって要求されたルートに対応するメディア・ビットストリームを得るために、選択されたオーディオ装置のメディア・エンジン（この例ではメディア・エンジン441）に制御情報を送信する。この例では、CHASM 401は、クラウドベースの音楽プロバイダーのHTTPアドレス、たとえば、クラウドベースの音楽プロバイダーによってホストされる特定のサーバーのHTTPアドレスをメディア・エンジン441に提供する。この実装によれば、ブロック1725において、メディア・エンジン441は、クラウドベースの音楽プロバイダーからメディア・ビットストリームを、この例では一つまたは複数の割り当てられたサーバー位置から取得する。

【0297】

この例では、ブロック1730は、ブロック1725で得られたメディア・ストリームに対応する音楽の再生を含む。この例によれば、CHASM 401は、少なくともラウドスピーカー461が、いくつかの例ではラウドスピーカー460および／またはラウドスピーカー462も、音楽の再生に関与することを決定している。いくつかのそのような例では、CHASM 401は、メディア・ストリームからのオーディオ・データをレンダリングして、レンダリングされたスピーカー・フィード信号をメディア・エンジン440および／またはメディア・エンジン442に提供するための命令をメディア・エンジン441に提供している。

【0298】

図18Aは、ある実施形態の最小バージョンのブロック図である。描かれているのは、N個のプログラム・ストリームであり（ $N \geq 2$ ）、第1のものが空間的であると明示的にラベル付けされている。対応するオーディオ信号の集まりは対応するレンダラーを通じてフィードされ、レンダラーはそれぞれ、M個の任意に離間されたラウドスピーカー（ $M \geq 2$ ）の共通セットを通じた、対応するプログラム・ストリームの再生のために、個々に構成されている。レンダラーは、本明細書では「レンダリング・モジュール」と称されることもある。レンダリング・モジュールおよびミキサー〔混合器〕1830aは、ソフトウェア、ハードウェア、ファームウェア、またはそれらの何らかの組み合わせを介して実装されてもよい。この例では、レンダリング・モジュールおよびミキサー1830aは、図6を参照して上述した制御システム610の一例である制御システム610aを介して実装されてもよい。いくつかの実装によれば、レンダリング・モジュールおよびミキサー1830aの機能は、少なくとも部分的には、本明細書でオーディオ・セッション・マネージャと称されるもの、たとえば、CHASMを実装している装置からの命令に従って実現されてもよい。いくつかのそのような例において、レンダリング・モジュールおよびミキサー1830aの機能性は、少なくとも部分的には、図2C、図2D、図3Cおよび図4を参照して上述したCHASM 208C、CHASM 208D、CHASM 307および／またはCHASM 401からの命令に従って、実装されてもよい。代替例においては、オーディオ・セッション・マネージャを実装している装置が、レンダリング・モジュールおよびミキサー1830aの機能性も実装してもよい。

【0299】

図18Aに示される例では、N個のレンダラーのそれぞれは、M個のラウドスピーカーフィードのセットを出力する。該ラウドスピーカーフィードは、M個のラウドスピーカーを通じた同時再生のために、すべてのN個のレンダラーにわたって合計される。この実装によれば、聴取環境内のM個のラウドスピーカーのレイアウトに関する情報は、ラウドスピーカー・ブロックからフィードバックする破線によって示されるように、すべてのレンダラーに提供され、その結果、レンダラーは、それらのスピーカーを通じた再生のために適正に構成される。このレイアウト情報は、具体的な実装に依存して、一つまたは複数のスピーカー自体から送信されても、そうでなくてもよい。いくつかの例によれば、レイア

ウト情報は、聴取環境におけるMラウドスピーカーのそれぞれの相対位置を決定するように構成された一つまたは複数のスマートスピーカーによって提供されてもよい。そのような自動位置特定方法のいくつかは、到来方向 (direction of arrival、DOA) 法または到達時間 (time of arrival、TOA) 法に基づくことができる。他の例では、このレイアウト情報は、別の装置によって決定されてもよく、および／またはユーザーによって入力されてもよい。いくつかの例では、聴取環境内のM個のラウドスピーカーの少なくともいくつかの能力に関するラウドスピーカー仕様情報が、すべてのレンダラーに提供されてもよい。そのようなラウドスピーカー仕様情報は、インピーダンス、周波数応答、感度、電力定格、個々のドライバの数および位置などを含むことができる。この例によれば、追加的なプログラム・ストリームの一つまたは複数のレンダリングからの情報が一次空間的ストリームのレンダラーにフィードされ、それにより、前記レンダリングが前記情報の関数として動的に修正されうる。この情報は、レンダリング・ブロック2ないしNから上方のレンダリング・ブロック1に戻る破線によって表される。

【0300】

図18Bは、追加的な特徴を有する別の（より能力のある）実施形態を示す。この例では、レンダリング・モジュールおよびミキサー1830bは、図6Aを参照して上述した制御システム610の一例である制御システム610bを介して実装されうる。いくつかの実装によれば、レンダリング・モジュールおよびミキサー1830bの機能は、少なくとも部分的には、本明細書でオーディオ・セッション・マネージャと称されるもの、たとえば、CHASMを実装している装置からの命令に従って実現されてもよい。いくつかのそのような例において、レンダリング・モジュールおよびミキサー1830bの機能性は、少なくとも部分的には、図2C、図2D、図3Cおよび図4を参照して上述したCHASM 208C、CHASM 208D、CHASM 307および／またはCHASM 401からの命令に従って、実装されてもよい。代替例においては、オーディオ・セッション・マネージャを実装している装置が、レンダリング・モジュールおよびミキサー1830bの機能性も実装してもよい。

【0301】

図18Bでは、すべてのN個のレンダラーの間で上下に移動する破線は、N個のレンダラーのうちの任意のものが残りのN-1個のレンダラーのうちの任意のものの動的修正に寄与するという考えを表す。換言すれば、N個のプログラム・ストリームのうちのいずれか1つのレンダリングは、残りのN-1個のプログラム・ストリームのいずれかの一つまたは複数のレンダリングの組み合わせの関数として、動的に修正されうる。さらに、プログラム・ストリームのうちの任意の一つまたは複数は空間ミックスであってもよく、任意のプログラム・ストリームのレンダリングは、それが空間的であるか否かにかかわらず、他のプログラム・ストリームのうちの任意のものの関数として動的に修正されてもよい。ラウドスピーカー・レイアウト情報は、たとえば上述のように、N個のレンダラーに提供されてもよい。いくつかの例では、ラウドスピーカー仕様情報は、N個のレンダラーに提供されてもよい。いくつかの実装では、マイクロフォン・システム620aは、聴取環境内にK個のマイクロフォンのセットを含んでいてもよい ($K \geq 1$)。いくつかの例では、マイクロフォンは、ラウドスピーカーのうちの前記一つまたは複数に取り付けられてもよく、または関連付けられてもよい。これらのマイクロフォンは、実線で表される捕捉されたオーディオ信号と、破線で表される追加的な構成情報（たとえば、それらの位置）の両方を、N個のレンダラーのセットにフィードバックすることができる。次いで、N個のレンダラーのうちの任意のものが、この追加的なマイクロフォン入力の関数として、動的に修正されうる。さまざまな例が本明細書において提供される。

【0302】

マイクロフォン入力から導出され、その後、N個のレンダラーのうちの任意のものを動的に修正するために使用される情報の例は、以下を含むが、これらに限定されない：

- ・システムのユーザーによる特定の単語または句の発話の検出。
- ・システムの一または複数のユーザーの場所の推定。
- ・聴取空間内の特定の位置での、N個のプログラム・ストリームの組み合わせの任意のも

のラウドネスの推定値。

・聴取環境におけるバックグラウンドノイズのような他の環境音のラウドネスの推定値。

【0303】

図19は、図6、図18Aまたは図18Bに示されるような装置またはシステムによって実行されうる方法の一例を概説するフロー図である。方法1900のブロックは、本明細書に記載される他の方法と同様に、必ずしも示される順序で実行されるわけではない。さらに、そのような方法は、図示および／または記載されているよりも多いまたは少ないブロックを含んでいてもよい。方法1900のブロックは、図6、図18Aおよび図18Bに示される制御システム610、制御システム610aまたは制御システム610bのような制御システムまたは他の開示される制御システム例の一つであってもよい（またはそれを含んでいてもよい）一つまたは複数の装置によって実行されてもよい。いくつかの実装によれば、方法1900のブロックは、少なくとも部分的には、本明細書でオーディオ・セッション・マネージャと称されるもの、たとえば、CHASMを実装している装置からの命令に従って実行されてもよい。いくつかのそのような例において、方法1900のブロックは、少なくとも部分的には、図2C、図2D、図3Cおよび図4を参照して上述したCHASM 208C、CHASM 208D、CHASM 307および／またはCHASM 401からの命令に従って、実装されてもよい。代替例においては、オーディオ・セッション・マネージャを実装している装置が、方法1900のブロックも実装してもよい。

10

【0304】

この実装では、ブロック1905は、インターフェース・システムを介して、第1のオーディオ・プログラム・ストリームを受領することを含む。この例では、第1のオーディオ・プログラム・ストリームは、環境の少なくともいくつかのスピーカーによって再生されるようにスケジュールされた第1のオーディオ信号を含む。ここで、第1のオーディオ・プログラム・ストリームは、第1の空間データを含む。この例によれば、第1の空間データは、チャンネル・データおよび／または空間メタデータを含む。いくつかの例では、ブロック1905は、制御システムの第1のレンダリング・モジュールが、インターフェース・システムを介して第1のオーディオ・プログラム・ストリームを受領することを含む。

20

【0305】

この例によれば、ブロック1910は、環境のスピーカーを介した再生のために第1のオーディオ信号をレンダリングし、第1のレンダリングされたオーディオ信号を生成することを含む。方法1900のいくつかの例は、たとえば上述のように、ラウドスピーカー・レイアウト情報を受領することを含む。方法1900のいくつかの例は、たとえば上述のように、ラウドスピーカー仕様情報を受領することを含む。いくつかの例では、第1のレンダリング・モジュールは、少なくとも部分的に、ラウドスピーカー・レイアウト情報および／またはラウドスピーカー仕様情報に基づいて、第1のレンダリングされたオーディオ信号を生成してもよい。

30

【0306】

この例では、ブロック1915は、インターフェース・システムを介して第2のオーディオ・プログラム・ストリームを受領することを含む。この実装では、第2のオーディオ・プログラム・ストリームは、環境の少なくともいくつかのスピーカーによって再生されるようにスケジュールされた第2のオーディオ信号を含む。この例によれば、第2のオーディオ・プログラム・ストリームは、第2の空間データを含む。第2の空間データは、チャンネル・データおよび／または空間メタデータを含む。いくつかの例では、ブロック215は、制御システムの第2のレンダリング・モジュールが、インターフェース・システムを介して第2のオーディオ・プログラム・ストリームを受領することを含む。

40

【0307】

この実装によれば、ブロック1920は、環境のスピーカーを介した再生するために第2のオーディオ信号をレンダリングし、第2のレンダリングされたオーディオ信号を生成することを含む。いくつかの例では、第2のレンダリング・モジュールは、受領されたラウドスピーカー・レイアウト情報および／または受領されたスピーカー仕様情報に少なくとも

50

部分的に基づいて、第2のレンダリングされたオーディオ信号を生成することができる。

【0308】

場合によっては、環境の一部またはすべてのスピーカーが任意に位置されうる。たとえば、環境の少なくともいくつかのスピーカーは、ドルビー5.1、ドルビー7.1、浜崎22.2等のような標準で規定されたスピーカー・レイアウトに対応しない位置に配置されてもよい。いくつかのそのような例では、環境の少なくともいくつかのスピーカーは、環境の家具、壁などに関して便利な位置（たとえば、スピーカーを収容するスペースがある位置）に配置されてもよいが、いかなる標準で規定されたスピーカー・レイアウトにない位置であってもよい。

【0309】

よって、いくつかの実装ブロック1910またはブロック1920は、任意に配置されたスピーカーへの柔軟なレンダリングに関わってもよい。いくつかのそのような実装は、質量中心振幅パン（CMAP）、柔軟仮想化（FV）、またはその両方の組み合わせに関わってもよい。高レベルからは、これらの技法は、それぞれが関連付けられた所望される知覚される空間位置をもつ一つまたは複数のオーディオ信号のセットを、2つ以上のスピーカーのセットを通じた再生のためにレンダリングする。ここで、セットのスピーカーの相対的アクティブ化は、それらのスピーカーを通じて再生される前記オーディオ信号の知覚された空間位置のモデルと、オーディオ信号の所望される知覚される空間位置の、スピーカーの位置に対する近接性との関数である。モデルは、オーディオ信号が、その意図された空間位置の近くで聴取者に聞こえることを保証し、近接性の項は、この空間的印象を達成するためにどのスピーカーが使用されるかを制御する。特に、近接性の項は、オーディオ信号の所望される知覚される空間位置に近いスピーカーのアクティブ化を優遇する。CMAPとFVの両方について、この関数関係は、空間的側面について1つと近接性について1つの2つの項の和として書かれたコスト関数から便利に導出される：

【数1】

$$C(\mathbf{g}) = C_{\text{spatial}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) + C_{\text{proximity}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) \quad (1)$$

【0310】

ここで、集合

【数2】

$$\{\vec{s}_i\}$$

〔 $\{\rightarrow s_i\}$ とも書く。以下、このようにベクトルの矢印を前に書くことがある〕は、M個のラウドスピーカーのセットの位置を示し、 $\rightarrow o$ は、オーディオ信号の所望される知覚される空間位置を示し、 $\mathbf{g}$ は、スピーカー・アクティブ化のM次元ベクトルを示す。CMAPについては、該ベクトル中の各アクティブ化は、スピーカー当たりの利得を表し、FVについては、各アクティブ化は、フィルタを表す（この第2の場合では、 $\mathbf{g}$ は、等価に、特定の周波数における複素数値のベクトルとみなすことができ、フィルタを形成するために複数の周波数にわたって異なる $\mathbf{g}$ が計算される）。アクティブ化の最適ベクトルは、諸アクティブ化を通じてコスト関数を最小化することによって求められる：

【数3】

$$\mathbf{g}_{\text{opt}} = \min_{\mathbf{g}} C(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) \quad (2a)$$

【0311】

コスト関数のある種の定義では、 $\mathbf{g}_{\text{opt}}$ の成分間の相対的なレベルは適切であるが、上記の最小化の結果として生じる最適な諸アクティブ化の絶対的なレベルを制御することは困難である。この問題に対処するために、諸アクティブ化の絶対的なレベルが制御されるよ

うに、 g_{opt} のその後の正規化が実行されてもよい。たとえば、単位長さを有するベクトルの正規化が望ましいことがあり、これは、一般的に使用される一定パワー・パン則に整合する：

【数 4】

$$\bar{g}_{opt} = \frac{g_{opt}}{\|g_{opt}\|} \quad (2b)$$

【0312】

柔軟レンダリング・アルゴリズムの厳密な挙動は、コスト関数の2つの項 $C_{spatial}$ および $C_{proximity}$ の具体的な構築によって支配される。CMAPについては、 $C_{spatial}$ は、一組のラウドスピーカーから再生されるオーディオ信号の知覚される空間位置を、それらのラウドスピーカーの位置に、それらの関連付けられたアクティブ化利得（ベクトル g の要素）によって重み付けした質量の中心：

【数 5】

$$\vec{o} = \frac{\sum_{i=1}^M g_i \vec{s}_i}{\sum_{i=1}^M g_i} \quad (3)$$

に配置するモデルから導出される。

【0313】

次いで、式3は、操作されて、所望されるオーディオ位置と、アクティブなラウドスピーカーによって生成されるオーディオ位置との間の誤差の二乗を表す空間コストにされる：

【数 6】

$$C_{spatial}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) = \|(\sum_{i=1}^M g_i) \vec{o} - \sum_{i=1}^M g_i \vec{s}_i\|^2 = \|\sum_{i=1}^M g_i (\vec{o} - \vec{s}_i)\|^2 \quad (4)$$

【0314】

FVでは、コスト関数の空間項は異なる仕方で定義される。ここでは、目標は、聴取者の左耳と右耳におけるオーディオ・オブジェクト位置 $\rightarrow o$ に対応するバイノーラル応答 b を生成することである。概念的には、 b は、フィルタ（各耳について1つのフィルタ）の 2×1 ベクトルであるが、より便利には、特定の周波数における複素数値の 2×1 ベクトルとして扱われる。特定の周波数でこの表現で続けると、所望されるバイノーラル応答が、オブジェクト位置によるHRTFインデックスの集合から取得されうる：

【数 7】

$$\mathbf{b} = HRTF\{\vec{o}\} \quad (5)$$

【0315】

同時に、ラウドスピーカーによって聴取者の耳のところに生成された 2×1 のバイノーラル応答 e は、 $2 \times M$ の音響伝達行列 H に、複素スピーカー・アクティブ化値の $M \times 1$ ベクトル g を乗じたものとしてモデル化される：

【数 8】

$$\mathbf{e} = H\mathbf{g} \quad (6)$$

【0316】

音響伝達行列 H は、聴取者位置に対するラウドスピーカー位置の集合 $\{\rightarrow s_i\}$ に基づいてモデル化される。最後に、コスト関数の空間成分は、所望されるバイノーラル応答（式5）とラウドスピーカーによって生成されるバイノーラル応答との間の誤差の二乗として定義

される（式6）：

【数9】

$$C_{spatial}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) = (\mathbf{b} - \mathbf{H}\mathbf{g})^*(\mathbf{b} - \mathbf{H}\mathbf{g}) \quad (7)$$

【0317】

便利なことに、式4および式7で定義されるCMAPおよびFVのコスト関数の空間項は、両方とも、スピーカー・アクティブ化 g の関数として、行列の二次形式に再構成できる：

【数10】

$$C_{spatial}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) = \mathbf{g}^* \mathbf{A} \mathbf{g} + \mathbf{B} \mathbf{g} + \mathbf{C} \quad (8)$$

ここで、 $\mathbf{A}$ は $M \times M$ の正方行列、 $\mathbf{B}$ は $1 \times M$ ベクトル、 $\mathbf{C}$ はスカラーである。行列 $\mathbf{A}$ は階数2であり、よって、 $M > 2$ の場合、空間誤差項がゼロに等しい無数のスピーカー・アクティブ化 g が存在する。コスト関数の第2項 $C_{proximity}$ を導入すると、この不定性が除去され、他の可能な解決策と比べて知覚的に有益な特性を有する特定の解が得られる。CMAPおよびFVの両方について、 $C_{proximity}$ は、所望されるオーディオ信号位置 $\vec{o}$ から離れた位置 $\vec{s}_i$ をもつスピーカーのアクティブ化が、所望される位置に近い位置をもつスピーカーのアクティブ化よりも大きなペナルティを受けるように構築される。この構築は、所望されるオーディオ信号の位置に近接したスピーカーのみが顕著にアクティブ化される、疎なスピーカー・アクティブ化の最適な集合を与え、實際上、スピーカーのセットのまわりの聴取者の動きに対して知覚的により堅牢なオーディオ信号の空間的な再現をもたらす。

【0318】

この目的に向け、コスト関数の第2項 $C_{proximity}$ は、スピーカー・アクティブ化の絶対値の2乗の、距離で重み付けされた和として定義されうる。これは、以下のように、行列形式で簡潔に表現される：

【数11】

$$C_{proximity}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) = \mathbf{g}^* \mathbf{D} \mathbf{g} \quad (9a)$$

ここで、

【数12】

$\mathbf{D}$

は、所望されるオーディオ位置と各スピーカーとの間の距離ペナルティの対角行列：

【数13】

$$\mathbf{D} = \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_M \end{bmatrix}, \quad d_i = distance(\vec{o}, \vec{s}_i) \quad (9b)$$

である。

【0319】

距離ペナルティ関数は多くの形をとることができるが、下記は有用なパラメータ化である。

【数14】

$$distance(\vec{o}, \vec{s}_i) = \alpha d_0^2 \left(\frac{\|\vec{o} - \vec{s}_i\|}{d_0} \right)^\beta \quad (9c)$$

ここで、

【数 1 5】

$$\|\vec{o} - \vec{s}_i\|$$

は、所望されるオーディオ位置とスピーカー位置との間のユークリッド距離であり、 α および β は調整可能なパラメータである。パラメータ α はペナルティのグローバルな強さを示し、 d_0 は距離ペナルティの空間的な広がりに対応し（ d_0 程度の距離にある、またはより遠くのスピーカーがペナルティを受ける）、 β は、距離 d_0 におけるペナルティの始まりの突然さを表す。

10

【0 3 2 0】

式 8 と 9a で定義されるコスト関数の 2 つの項を組み合わせると、全体的なコスト関数が得られる。

【数 1 6】

$$C(g) = g^*Ag + Bg + C + g^*Dg = g^*(A + D)g + Bg + C \quad (10)$$

このコスト関数の

【数 1 7】

g

20

に関する導関数を 0 とおいて、

【数 1 8】

g

について解くと、最適なスピーカー・アクティブ化の解が得られる：

【数 1 9】

$$g_{opt} = \frac{1}{2}(A + D)^{-1}B \quad (11)$$

30

【0 3 2 1】

一般に、式 11 の最適解は、値が負であるスピーカー・アクティブ化を生じることがある。柔軟レンダラーの CMAP 構築については、そのような負のアクティブ化は望ましくないことがあり、よって、式 11 は、すべての正のアクティブ化が正のままであるという条件のもとに、最小化されてもよい。

【0 3 2 2】

柔軟レンダリング方法（いくつかの実施形態に従って実装される）をワイヤレススマートスピーカー（または他のスマート・オーディオ装置）の集合とペアリングすることにより、きわめて能力が高く、使いやすい空間オーディオ・レンダリング・システムを与えることができる。そのようなシステムとの相互作用を考えると、システムの使用中に生じる他の目的のために最適化するために、空間的レンダリングに対する動的な修正が望ましいことがありうるということが明らかになる。この目的を達成するために、あるクラスの実施形態は、既存の柔軟レンダリング・アルゴリズムを、レンダリングされるオーディオ信号の一つまたは複数の属性、スピーカーの集合、および／または他の外部入力に依存する一つまたは複数の追加的な動的に構成可能な機能を用いて、補強する。いくつかの実施形態によれば、式 1 で与えられる既存の柔軟レンダリングのコスト関数は、次のように、これらの一つまたは複数の追加の従属関係を用いて補強される。

40

【数 2 0】

50

$$C(\mathbf{g}) = C_{\text{spatial}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) + C_{\text{proximity}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) + \sum_j C_j(\mathbf{g}, \{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j) \quad (12)$$

【 0 3 2 3 】

式12において、項

【数 2 1】

$$C_j(\mathbf{g}, \{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j)$$

10

は、追加的なコスト項を表し、

【数 2 2】

$$\{\hat{o}\}$$

は、レンダリングされる（たとえば、オブジェクトベースのオーディオ・プログラムの）オーディオ信号の一つまたは複数の属性の集合を表し、

【数 2 3】

$$\{\hat{s}_i\}$$

20

は、それを通じてオーディオがレンダリングされるスピーカーの一つまたは複数の属性の集合を表し、

【数 2 4】

$$\{\hat{e}\}$$

は、一つまたは複数の追加的な外部入力を表す。各項

【数 2 5】

$$C_j(\mathbf{g}, \{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j)$$

30

は、

【数 2 6】

$$\{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j$$

によって表される、オーディオ信号、スピーカー、および／または外部入力の一つまたは複数の属性の組み合わせに関する、アクティブ化gの関数としてのコストを返す。集合

【数 2 7】

$$\{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j$$

40

が、少なくとも、

【数 2 8】

$$\{\hat{o}\}, \{\hat{s}_i\}, \text{または} \{\hat{e}\}$$

のいずれかからの1つのみの要素を含むことが理解されるべきである。

50

【0324】

【数29】

 $\{\hat{o}\}$

の例は、以下を含むが、これらに限定されない：

- ・オーディオ信号の所望される知覚される空間位置；
- ・オーディオ信号のレベル（可能性としては時間変化する）；および／または
- ・オーディオ信号のスペクトル（可能性としては時間変化する）。

【0325】

【数30】

 $\{\hat{s}_i\}$

の例は、以下を含むが、これらに限定されない：

- ・聴取スペース内のラウドスピーカーの位置；
- ・ラウドスピーカーの周波数応答；
- ・ラウドスピーカーの再生レベル制限；
- ・リミッタ利得などスピーカー内のダイナミクス処理アルゴリズムのパラメータ；
- ・各スピーカーから他のスピーカーへの音響伝達の測定または推定；
- ・スピーカー上のエコー・キャンセラ性能の尺度；および／または
- ・スピーカーの、互いの相対的な同期。

【0326】

【数31】

 $\{\hat{e}\}$

の例は、以下を含むが、これらに限定されない：

- ・再生空間内の1人以上の聴取者または話者の位置；
- ・各ラウドスピーカーから聴取位置までの音響伝達の測定または推定；
- ・話者からラウドスピーカーの集合までの音響伝達の測定または推定；
- ・再生空間内の何らかの他のランドマークの位置；および／または
- ・各スピーカーから再生空間における何らかの他のランドマークへの音響伝達の測定または推定。

【0327】

式12で定義された新しいコスト関数を用いて、式2aおよび2bで前述したように、gに関する最小化および可能な事後正規化を通じて、アクティブ化の最適な集合を見つけることができる。

【0328】

式9aおよび9bで定義されている近接性コストと同様に、新しいコスト関数の項

【数32】

$$C_j(\mathbf{g}, \{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j)$$

を、スピーカー・アクティブ化の絶対値の2乗の重み付けされた和として表現することも便利である：

【数33】

$$C_j(\mathbf{g}, \{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j) = \mathbf{g}^* \mathbf{W}_j(\{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j) \mathbf{g}, \quad (13a)$$

10

20

30

40

50

ここで、 w_j は、項 j についてスピーカ- i をアクティブ化することに関連するコストを記述する重み

【数 3 4】

$$w_{ij} = w_{ij} \left(\{ \{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\} \}_j \right)$$

の対角行列である：

【数 3 5】

$$\mathbf{W}_j = \begin{bmatrix} w_{1j} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_{Mj} \end{bmatrix} \quad (13b) \quad 10$$

【 0 3 2 9】

式13aおよび13bを、式10で与えられたCMAPおよびFVコスト関数の行列二次形式バージョンと組み合わせることにより、式12で与えられた（いくつかの実施形態の）一般拡張されたコスト関数の潜在的に有益な実装がもたらされる：

【数 3 6】

$$C(\mathbf{g}) = \mathbf{g}^* \mathbf{A} \mathbf{g} + \mathbf{B} \mathbf{g} + \mathbf{C} + \mathbf{g}^* \mathbf{D} \mathbf{g} + \sum_j \mathbf{g}^* \mathbf{W}_j \mathbf{g} = \mathbf{g}^* (\mathbf{A} + \mathbf{D} + \sum_j \mathbf{W}_j) \mathbf{g} + \mathbf{B} \mathbf{g} + \mathbf{C} \quad (14) \quad 20$$

【 0 3 3 0】

新しいコスト関数項のこの定義では、全体的なコスト関数は行列二次形式のままであり、アクティブ化の最適な集合 $\mathbf{g}_{opt}$ は式14の微分を通じて見出すことができ、次のようになる。

【数 3 7】

$$\mathbf{g}_{opt} = \frac{1}{2} (\mathbf{A} + \mathbf{D} + \sum_j \mathbf{W}_j)^{-1} \mathbf{B} \quad (15)$$

【 0 3 3 1】

重み項 w_{ij} のそれぞれを、ラウドスピーカ-のそれぞれについての与えられた連続的なペナルティ値

【数 3 8】

$$p_{ij} = p_{ij} \left(\{ \{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\} \}_j \right)$$

の関数として考えることは有用である。ある例示的实施形態では、このペナルティ値は、（レンダリングされるべき）オブジェクトから考慮されるラウドスピーカ-までの距離である。別の例示的实施形態では、このペナルティ値は、所与のラウドスピーカ-がいくつかの周波数を再生することができないことを表す。このペナルティ値に基づいて、重み項は次のようにパラメータ化できる：

【数 3 9】

$$w_{ij} = \alpha_j f_j \left(\frac{p_{ij}}{\tau_j} \right) \quad (16)$$

ここで、 α_j は、（重み項のグローバルな強度を考慮に入れる）プレファクターを表し、 τ_j は、ペナルティ閾値を表し（その近くで、またはそれを超えるところで重み項が重要となる）、 $f_j(x)$ は単調増加関数を表す。たとえば、

10

20

30

40

50

【数 4 0】

$$f_j(x) = x^{\beta_j}$$

では、重み項は、次のような形をもつ：

【数 4 1】

$$w_{ij} = \alpha_j \left(\frac{p_{ij}}{\tau_j} \right)^{\beta_j} \quad (17)$$

ここで、 α_j 、 β_j 、 τ_j は、ペナルティのグローバルな強さ、ペナルティの始まりの突然性、ペナルティの広がりそれぞれ示す調整可能なパラメータである。これらの調整可能な値を設定する際には、コスト項 C_j の、他の任意の追加的なコスト項ならびに C_{spatial} および $C_{\text{proximity}}$ に対する相対的な効果が、望ましい成果を達成するために適切であるように、注意を払うべきである。たとえば、大雑把な目安として、ある特定のペナルティがはっきりと他のペナルティより支配的であることを望むなら、その強度を2番目に大きいペナルティ強度の約10倍に設定することが適切でありうる。

【0 3 3 2】

すべてのラウドスピーカーがペナルティを課される場合、後処理において、すべての重み項から最小のペナルティを差し引いて、スピーカーのうちの少なくとも1つがペナルティを課されないようにすることがしばしば便利で：

【数 4 2】

$$w_{ij} \rightarrow w'_{ij} = w_{ij} - \min_i(w_{ij}) \quad (18)$$

【0 3 3 3】

上述したように、本明細書に記載される新しいコスト関数項（および他の実施形態に従って使用される同様の新しいコスト関数項）を使用して実現できる多くの可能な使用事例がある。次に、3つの例を用いて、より具体的な詳細を説明する。すなわち、オーディオを聴取者または話者に向けて移動させる、オーディオを聴取者または話者から遠ざける、オーディオをランドマークから遠ざける。

【0 3 3 4】

第1の例では、ここでは「引力」と呼ばれるものが、オーディオをある位置に向けて引っぱるために使用される。その位置は、いくつかの例では、聴取者または話者の位置、ランドマーク位置、家具位置などであってもよい。本明細書では、この位置は「引力位置」または「アトラクター位置」と称されることがある。本明細書で使用される場所では、「引力」とは、引力位置により近接した近傍において、相対的により高いラウドスピーカー・アクティブ化を優遇する因子である。この例によれば、重み w_{ij} は式17の形をとり、連続的ペナルティ値 p_{ij} は、 i 番目のスピーカーの、固定したアトラクター位置

【数 4 3】

$$\vec{l}_j$$

からの距離によって与えられ、閾値 τ_j は、すべてのスピーカーにわたるこれらの距離の最大値によって与えられる：

【数 4 4】

$$p_{ij} = \|\vec{l}_j - \vec{s}_i\|, \text{ and} \quad (19a)$$

$$\tau_j = \max_i \|\vec{l}_j - \vec{s}_i\| \quad (19b)$$

【0335】

オーディオを聴取者または話者に向けて「引っ張る」使用事例を例解すると、具体的に $\alpha_j = 20$ 、 $\beta_j = 3$ に設定し、

【数45】

$$\vec{l}_j$$

$\{\rightarrow l_j\}$ を180度の聴取者／話者の位置に対応するベクトルに設定する。 α_j 、 β_j および

【数46】

$$\vec{l}_j$$

10

のこれらの値は単に例である。他の実装では、 α_j は1～100の範囲であってもよく、 β_j は1～25の範囲であってもよい。

【0336】

第2の例では、「反発力」は、聴取者の位置、話者の位置またはランドマークの位置、家具の位置などの他の位置であってもよい位置から離れるようにオーディオを「押す」ために使用される。この位置は、本明細書では、「反発力位置」または「反発位置」と称されることがある。本明細書では、「反発力」とは、反発力位置に近いほど、相対的により低いスピーカー・アクティブ化を促進する要因である。この例によれば、固定した反発位置

20

【数47】

$$\vec{l}_j$$

に対して p_{ij} および τ_j を式19における引力と同様に定義する：

【数48】

$$p_{ij} = \max_i \|\vec{l}_j - \vec{s}_i\| - \|\vec{l}_j - \vec{s}_i\|, \quad (19c)$$

$$\tau_j = \max_i \|\vec{l}_j - \vec{s}_i\| \quad (19d)$$

30

【0337】

オーディオを聴取者または話者から遠ざける使用事例を例解すると、具体的に $\alpha_j = 5$ 、 $\beta_j = 2$ に設定し、

【数49】

$$\vec{l}_j$$

$\{\rightarrow l_j\}$ を180度の聴取者／話者の位置に対応するベクトルに設定する。 α_j 、 β_j および $\rightarrow l_j$ のこれらの値は単に例である。

40

【0338】

ここで図19に戻ると、この例では、ブロック1925は、少なくとも部分的には、第2のオーディオ信号、第2のレンダリングされたオーディオ信号またはその特性のうちの少なくとも1つに基づいて、第1のオーディオ信号についてのレンダリング・プロセスを修正して、修正された第1のレンダリングされたオーディオ信号を生成することに関わる。レンダリング・プロセスを修正するさまざまな例が本明細書に開示されている。レンダリングされた信号の「特性」は、たとえば、無音環境での、または一つまたは複数の追加的なレンダリングされた信号の存在下での、意図された聴取位置における推定されたまたは測定されたラウドネスまたは可聴性を含みうる。特性の他の例は、前記信号のレンダリングに関連するパラメータ、たとえば関連するプログラム・ストリームの構成信号の意図された

50

空間位置、信号がレンダリングされるラウドスピーカーの位置、構成信号の意図された空間位置の関数としてのラウドスピーカーの相対的アクティブ化、および、前記レンダリングされた信号を生成するために利用されるレンダリング・アルゴリズムに関連する任意の他のパラメータまたは状態を含む。いくつかの例では、ブロック225は、第1のレンダリング・モジュールによって実行されてもよい。

【0339】

この例によれば、ブロック1930は、少なくとも部分的には、第1のオーディオ信号、第1のレンダリングされたオーディオ信号またはその特性のうちの少なくとも1つに基づいて、第2のオーディオ信号についてのレンダリング・プロセスを修正して、修正された第2のレンダリングされたオーディオ信号を生成することに関わる。いくつかの例では、ブロック1930は、第2のレンダリング・モジュールによって実行されてもよい。

10

【0340】

いくつかの実装では、第1のオーディオ信号についてのレンダリング・プロセスを修正することは、第2のレンダリングされたオーディオ信号のレンダリング位置から離れるように第1のオーディオ信号のレンダリングをワープさせること、および／または、第2のオーディオ信号または第2のレンダリングされたオーディオ信号のうち一つまたは複数のラウドネスに応答して、第1のレンダリングされたオーディオ信号のうち一つまたは複数のラウドネスを修正することを含みうる。代替的または追加的に、第2のオーディオ信号についてのレンダリング・プロセスを修正することは、第1のレンダリングされたオーディオ信号のレンダリング位置から離れるように第2のオーディオ信号のレンダリングをワープさせること、および／または、第1のオーディオ信号または第1のレンダリングされたオーディオ信号のうちの一つまたは複数のラウドネスに応答して、第2のレンダリングされたオーディオ信号のうちの一つまたは複数のラウドネスを修正することを含みうる。いくつかの例は、図3以下を参照して以下に提供される。

20

【0341】

しかしながら、他のタイプのレンダリング・プロセス修正は、本開示の範囲内にある。たとえば、いくつかの例において、第1のオーディオ信号または第2のオーディオ信号についてのレンダリング・プロセスを修正することは、スペクトル修正、可聴性ベースの修正またはダイナミックレンジ修正を実行することに関わってもよい。これらの修正は、具体的な例に依存して、ラウドネスに基づくレンダリング修正に関連しても、しなくてもよい。たとえば、一次空間ストリームが開放的なプランの居住エリアでレンダリングされ、料理のヒントで構成される二次ストリームが隣接するキッチンでレンダリングされる上述した場合においては、料理のヒントがキッチンにおいて可聴なままであることを保証することが望ましいことがありうる。これは、干渉する第1の信号がない場合に、キッチンにおけるレンダリングされた料理のヒントのストリームについてのラウドネスがどうなるかを推定し、次いで、キッチンにおける第1の信号の存在下で、該ラウドネスを推定し、最後に、複数の周波数にわたる両方のストリームのラウドネスおよびダイナミックレンジを動的に修正して、キッチンにおける第2の信号の可聴性を保証することによって達成できる。

30

【0342】

図19に示される例では、ブロック1935は、少なくとも、修正された第1のレンダリングされたオーディオ信号と、修正された第2のレンダリングされたオーディオ信号とを混合して、混合されたオーディオ信号を生成することに関わる。ブロック1935は、たとえば、図18Bに示されるミキサー1830bによって実行されてもよい。

40

【0343】

この例によれば、ブロック1940は、混合オーディオ信号を環境の少なくともいくつかのスピーカーに提供することに関わる。方法1900のいくつかの例は、スピーカーによる混合オーディオ信号の再生に関わる。

【0344】

図19に示されるように、いくつかの実装は、2つより多いレンダリング・モジュールを提供してもよい。いくつかのそのような実装は、N個のレンダリング・モジュールを提供

50

してもよい。ここで、Nは、2より大きい整数である。よって、いくつかのそのような実装は、一つまたは複数の追加的なレンダリング・モジュールを含んでいてもよい。いくつかのそのような例では、前記一つまたは複数の追加的なレンダリング・モジュールのそれぞれは、インターフェース・システムを介して追加的なオーディオ・プログラム・ストリームを受領するように構成されてもよい。追加的なオーディオ・プログラム・ストリームは、環境の少なくとも1つのスピーカーによって再生されるようにスケジュールされた追加的なオーディオ信号を含んでいてもよい。いくつかのそのような実装は、追加的なオーディオ信号を、環境の少なくとも1つのスピーカーを介した再生のためにレンダリングし、追加的なレンダリングされたオーディオ信号を生成し、少なくとも部分的には、第1のオーディオ信号、第1のレンダリングされたオーディオ信号、第2のオーディオ信号、第2のレンダリングされたオーディオ信号またはそれらの特性のうちの少なくとも1つに基づいて、追加的なオーディオ信号についてのレンダリング・プロセスを修正して、修正された追加的なレンダリングされたオーディオ信号を生成することに関わってもよい。いくつかのそのような例によれば、混合モジュールは、修正された追加的なレンダリングされたオーディオ信号を、少なくとも前記修正された第1のレンダリングされたオーディオ信号および前記修正された第2のレンダリングされたオーディオ信号と混合して、前記混合されたオーディオ信号を生成するようにさらに構成されてもよい。

【0345】

図6および図18Bを参照して上記したように、いくつかの実装は、聴取環境における一つまたは複数のマイクロフォンを含むマイクロフォン・システムを含んでいてもよい。いくつかのそのような例では、第1のレンダリング・モジュールは、マイクロフォン・システムからの第1のマイクロフォン信号に少なくとも部分的に基づいて、第1のオーディオ信号についてのレンダリング・プロセスを修正するように構成されてもよい。「第1のマイクロフォン信号」は、具体的な実装に依存して、単一のマイクロフォンから、または2つ以上のマイクロフォンから受領されうる。いくつかのそのような実装では、第2のレンダリング・モジュールは、少なくとも部分的に第1のマイクロフォン信号に基づいて、第2のオーディオ信号についてのレンダリング・プロセスを修正するように構成されてもよい。

【0346】

図18Bを参照して上述したように、いくつかの例では、一つまたは複数のマイクロフォンの位置が既知であってもよく、制御システムに提供されてもよい。いくつかのそのような実装によれば、制御システムは、第1のマイクロフォン信号に基づいて第1の音源位置を推定し、第1の音源位置に少なくとも部分的に基づいて、第1のオーディオ信号または第2のオーディオ信号のうちの少なくとも1つについてのレンダリング・プロセスを修正するようにさらに構成されてもよい。第1の音源位置は、たとえば、既知の位置を有する3つ以上のマイクロフォンまたはマイクロフォンのグループのそれぞれからのDOAデータに基づいて、三角測量プロセスに従って推定されてもよい。代替的または追加的に、第1の音源位置は、2つ以上のマイクロフォンからの受領された信号の振幅に従って推定されてもよい。最高振幅の信号を発生するマイクロフォンは、第1の音源位置に最も近いと想定されてもよい。いくつかのそのような例では、第1の音源位置は、最も近いマイクロフォンの位置に設定されてもよい。いくつかのそのような例において、第1の音源位置は、ゾーンの位置に関連付けられてもよく、ここで、ゾーンは、ガウス・ミキサー・モデルのような事前トレーニングされた分類器を通じて、2つ以上のマイクロフォンからの信号を処理することによって選択される。

【0347】

いくつかのそのような実装では、制御システムは、第1のマイクロフォン信号が環境ノイズに対応するかどうかを判定するように構成されてもよい。いくつかのそのような実装は、第1のマイクロフォン信号が環境ノイズに対応するかどうか少なくとも部分的に基づいて、第1のオーディオ信号または第2のオーディオ信号の少なくとも一方についてのレンダリング・プロセスを修正することに関わってもよい。たとえば、制御システムが、第1のマイクロフォン信号が環境ノイズに対応すると判定する場合、第1のオーディオ信号ま

たは第2のオーディオ信号についてのレンダリング・プロセスを修正することは、意図される聴取位置におけるノイズの存在下での信号の知覚されるラウドネスが、ノイズの不在時の信号の知覚されるラウドネスと実質的に等しくなるよう、レンダリングされたオーディオ信号のレベルを増加させることに関わってもよい。

【0348】

いくつかの例では、制御システムは、第1のマイクロフォン信号が人間の声に対応するかどうかを判定するように構成されてもよい。そのようないくつかの実装は、第1のマイクロフォン信号が人間の声に対応するかどうかに少なくとも部分的に基づいて、第1のオーディオ信号または第2のオーディオ信号の少なくとも1つについてのレンダリング・プロセスを修正することに関わってもよい。たとえば、制御システムが、第1のマイクロフォン信号がウェイクワードのような人間の音に対応すると判定する場合、第1のオーディオ信号または第2のオーディオ信号についてのレンダリング・プロセスを修正することは、第1の音源位置からより遠いスピーカーによって再生されるレンダリングされたオーディオ信号のラウドネスと比較して、第1の音源位置近くのスピーカーによって再生されるレンダリングされたオーディオ信号のラウドネスを減少させることを含んでもよい。第1のオーディオ信号または第2のオーディオ信号についてのレンダリング・プロセスを修正することは、代替的に、第1の音源位置から離れるように関連プログラム・ストリームの構成信号の意図された位置をワープさせる、および／または、第1の音源位置からより多いスピーカーと比較して、第1の音源位置の近くでスピーカーを使用することにペナルティを課すように、レンダリング・プロセスを修正することを含んでもよい。

【0349】

いくつかの実装では、制御システムが、第1のマイクロフォン信号が人間の声に対応すると判定する場合、制御システムは、第1の音源位置とは異なる、環境のある位置に近い一つまたは複数のスピーカーにおいて第1のマイクロフォン信号を再生するように構成されてもよい。いくつかのそのような例では、制御システムは、第1のマイクロフォン信号が子供の泣き声に対応するかどうかを判定するように構成されてもよい。いくつかのそのような実装によれば、制御システムは、親、親戚、保護者、保育サービス提供者、教師、看護師などのケア担当者の推定された位置に対応する、環境の位置に近い一つまたは複数のスピーカーにおいて、第1のマイクロフォン信号を再生するように構成されてもよい。いくつかの例では、ケア担当者の推定位置を推定するプロセスは、「〈ウェイクワード〉、赤ん坊を起こすな」などの音声コマンドによってトリガーされてもよい。制御システムは、3つ以上のローカルなマイクロフォンなどによって提供されるDOA情報に基づく三角測量により、バーチャル・アシスタントを実装している最も近いスマート・オーディオ装置の位置に応じて、スピーカー（ケア担当者）の位置を推定することができる。いくつかの実装によれば、制御システムは、ベビールーム（および／またはその中の聴取装置）の位置についての事前知識を有し、よって、適切な処理を実行することができる。

【0350】

いくつかのそのような例によれば、制御システムは、第1のマイクロフォン信号がコマンドに対応するかどうかを判定するように構成されてもよい。制御システムが、第1のマイクロフォン信号がコマンドに対応すると判定する場合、場合によっては、制御システムは、コマンドへの返答を決定し、該返答を再生するよう、第1の音源位置に近い少なくとも1つのスピーカーを制御するように構成されてもよい。いくつかのそのような例では、制御システムは、該返答を再生するよう、第1の音源位置に近い少なくとも1つのスピーカーを制御した後に、第1のオーディオ信号または第2のオーディオ信号についての未修正のレンダリング・プロセスに戻るように構成されてもよい。

【0351】

いくつかの実装では、制御システムは、コマンドを実行するように構成されてもよい。たとえば、制御システムは、コマンドに従ってオーディオ装置、テレビ、家電製品などを制御するように構成されたバーチャル・アシスタントであってもよく、または、それを含んでもよい。

【0352】

図6、図18A、および図18Bに示される最小限の、およびより能力の高いマルチストリーム・レンダリング・システムのこの定義により、多数の有用なシナリオについて、複数のプログラム・ストリームの同時再生の動的管理が達成されうる。ここで、図20および21を参照して、いくつかの例が記述される。

【0353】

まず、居間での空間的な映画サウンドトラックと、接続されたキッチンでの料理のヒントとの同時再生に関わる、先に論じた例を調べる。空間的な映画サウンドトラックは、上述の「第1のオーディオ・プログラム・ストリーム」の例であり、料理のヒントのオーディオは、上述の「第2のオーディオ・プログラム・ストリーム」の例である。図20および図21は、接続された居住空間のフロアプランの例を示す。この例では、居住空間2000は、左上に居間、中央下にキッチン、右下に寝室を含む。居住空間にわたって分布する四角および円2005a~2005hは、スペースに都合のよい位置に配置されているが、標準で規定されたレイアウトには準拠しない（任意に配置された）8つのラウドスピーカーのセットを表す。図20では、空間的な映画サウンドトラックのみが再生されており、居間2010およびキッチン2015内のすべてのラウドスピーカーが、ラウドスピーカーの能力およびレイアウトを考慮して、テレビ2030に面してカウチ2025に座った聴取者2020aのまわりに、最適化された空間的再生を生成するために利用される。この映画サウンドトラックの最適な再生は、アクティブなラウドスピーカーの範囲内にあるクラウド2035aによって視覚的に表現される。

【0354】

図21では、料理のヒントが同時にレンダリングされ、キッチン2015内の単一のラウドスピーカー2005gを通じて第2の聴取者2020bのために再生される。この第2のプログラム・ストリームの再生は、ラウドスピーカー2005gから出るクラウド2040によって視覚的に表される。これらの料理のヒントが、図20に示されるように、映画サウンドトラックのレンダリングへの修正なく同時に再生されたとしたら、キッチン2015内またはその近くのスピーカーから発せられる映画サウンドトラックからのオーディオは、第2の聴取者が料理のヒントを理解する能力を妨げることになる。その代わりに、この例では、空間的な映画サウンドトラックのレンダリングは、料理のヒントのレンダリングの関数として動的に修正される。具体的には、映画サウンドトラックのレンダリングは、料理のヒントのレンダリング位置（キッチン2015）の近くのスピーカーから離れてシフトされ、このシフトは、キッチン近くのスピーカーから押し離された、図21における、より小さなクラウド2035bによって視覚的に表される。映画サウンドトラックがまだ再生中に料理のヒントの再生が停止する場合、いくつかの実装では、映画サウンドトラックのレンダリングは、動的に、図20に見られるもとの最適な構成に、シフトして戻ることができる。空間的な映画サウンドトラックのレンダリングにおけるそのような動的なシフトは、多数の開示された方法を通じて達成されうる。多くの空間的なオーディオ・ミックスは、聴取空間における特定の位置で再生されるように設計された複数の構成要素オーディオ信号を含む。たとえば、ドルビー5.1および7.1のサラウンドサウンド・ミックスは、それぞれ6信号および8信号からなり、これらは、聴取者のまわりの規定された正準位置にあるスピーカーで再生されることが意図されている。オブジェクトベースのオーディオ・フォーマット、たとえばドルビー・アトモスは、構成要素オーディオ信号と、オーディオがレンダリングされることが意図される聴取空間内の、可能性としては時間変化する3D位置を記述する関連メタデータで構成される。空間的な映画サウンドトラックのレンダラーが、ラウドスピーカーの任意のセットに関して、個々のオーディオ信号を任意の位置でレンダリングすることができるとすると、図20および図21に示されるレンダリングへの動的シフトは、空間ミックス内のオーディオ信号の意図された位置をワープさせることによって達成されうる。たとえば、オーディオ信号に関連する2Dまたは3D座標は、キッチン内のスピーカーの位置から離れるように押しやられ、あるいは居間の左上隅に向かって引っ張られる。そのようなワープの結果は、キッチン近くのスピーカーはそれほど使用されなくなるというもの

である。空間ミックスのオーディオ信号のワープされた位置が、今やこの位置からより離れているためである。この方法は、第2のオーディオ・ストリームを第2の聴取者にとってより理解しやすいものにするという目標を達成するが、それは、第1の聴取者のための映画サウンドトラックの意図された空間バランスを著しく変更するという代償を伴う。

【0355】

空間的レンダリングへの動的シフトを達成するための第2の方法は、柔軟レンダリング・システムを使用することによって実現されうる。いくつかのそのような実装では、柔軟レンダリング・システムは、上述のように、CMAP、FVまたは両方のハイブリッドであってもよい。いくつかのそのような柔軟なレンダリング・システムは、すべての構成要素信号が意図された位置から来ていると知覚されるように空間ミックスを再生しようと試みる。ミックスの各信号についてそうする一方で、いくつかの例では、その信号の所望される位置に近接したラウドスピーカーのアクティブ化が優先される。いくつかの実装では、レンダリングの最適化に、追加的な項が動的に加えられてもよく、これは、他の基準に基づいてある種のラウドスピーカーの使用にペナルティを課す。今の例では、「反発力」と呼ばれるものがキッチンの位置に動的に配置されて、この位置の近くにあるラウドスピーカーの使用に高いペナルティを課し、空間的な映画サウンドトラックのレンダリングを効果的に遠ざけることができる。本明細書中で使用されるところでは、用語「反発力」は、聴取環境の特定の位置またはエリアにおける比較的低いスピーカー・アクティブ化に対応する因子を指しうる。言い換えれば、「反発力」という句は、「反発力」に対応する特定の位置またはエリアから相対的により遠いスピーカーのアクティブ化を優遇する因子を指してもよいが、いくつかのそのような実装によれば、レンダラーは、まだ、残っている、より低いペナルティのスピーカーを用いて、ミックスの意図された空間バランスを再現しようと試みてよい。このように、この技法は、単にミックスの構成要素信号の意図された位置をワープさせる方法と比較して、レンダリングの動的シフトを達成するための、より優れた方法であると考えられうる。

【0356】

空間的な映画サウンドトラックのレンダリングをキッチンにおける料理のヒントから遠ざけるようにシフトさせる上述したシナリオは、図18Aに示されるマルチストリーム・レンダラーの最小バージョンを用いて達成されうる。しかしながら、図18Bに示されている、より能力の高いシステムを採用することによって、該シナリオへの改善が実現されうる。空間的な映画のサウンドトラックのレンダリングをシフトすることは、キッチンにおける料理のヒントの了解性を改善するが、映画サウンドトラックは、依然として、キッチンにおいて顕著に可聴である場合がある。両方のストリームの瞬間的な状態によっては、料理のヒントが、映画サウンドトラックによってマスクされてしまうことがある。たとえば、映画サウンドトラックにおける音量の大きな瞬間が、料理のヒントにおける音量の小さな瞬間をマスクする。この問題に対処するために、空間的な映画サウンドトラックのレンダリングの関数として、料理のヒントのレンダリングに対する動的修正が追加されてもよい。たとえば、干渉信号の存在下でその知覚されるラウドネスを維持するために、周波数および時間にわたってオーディオ信号を動的に変更する方法が実行されてもよい。このシナリオでは、キッチン位置におけるシフトされた映画サウンドトラックの知覚されるラウドネスの推定値が生成され、干渉信号としてそのようなプロセスに供給されてもよい。次いで、料理のヒントの時間変化および周波数変化するレベルが、その知覚されるラウドネスをこの干渉よりも上に維持するよう、動的に修正されうる。それにより、第2の聴取者にとっての了解性をより良好に維持することができる。キッチン内の映画サウンドトラックの必要なラウドネス推定値は、サウンドトラックのレンダリングのスピーカー・フィード、キッチン内またはその近くのマイクロフォンからの信号、またはそれらの組み合わせから生成されうる。料理のヒントの知覚されるラウドネスを維持するプロセスは、一般に、料理のヒントのレベルを増大させ、場合によっては、全体的なラウドネスが不快なほど高くなる場合がある。この問題に対処するために、さらに別のレンダリング修正が採用されてもよい。干渉する空間的な映画サウンドトラックは、キッチンにおける、ラウドネス

10

20

30

40

50

修正された料理のヒントの音量が大きくなりすぎることに応じて、動的に小さくされてもよい。最後に、何らかの外部ノイズ源が、両方のプログラム・ストリームの可聴性に同時に干渉する可能性がある。たとえば、調理中にキッチンにおいてブレンダー〔料理用ミキサー〕が使用されることがある。居間とキッチンの両方におけるこの環境ノイズ源のラウドネスの推定値は、レンダリング・システムに接続されたマイクロフォンから生成されてもよい。この推定値は、たとえば、キッチンにおけるサウンドトラックのラウドネスの推定値に加えられてもよく、それは、料理のヒントのラウドネス修正に影響を及ぼす。同時に、居間内のサウンドトラックのレンダリングは、この環境ノイズの存在下で居間におけるサウンドトラックの知覚されるラウドネスを維持するために、環境ノイズ推定値の関数として追加的に修正されてもよく、それにより、居間にいる聴取者にとっての可聴性をよりよく維持することができる。

10

【0357】

わかるように、開示されたマルチストリーム・レンダラーのこの例示的な使用事例は、2つのプログラム・ストリームの同時再生を最適化するために、2つのプログラム・ストリームに対する多数の相互接続された修正を採用する。要約すると、ストリームに対するこれらの修正は、次のように列挙できる：

●空間的な映画サウンドトラック

○キッチンにおいてレンダリングされる料理のヒントの関数として、空間的レンダリングがキッチンから離れるようにシフトされる

○キッチンにおいてレンダリングされる料理のヒントのラウドネスの関数としての、ラウドネスの動的な低下

20

○キッチンからの干渉するブレンダー・ノイズの居間でのラウドネスの推定値の関数として、ラウドネスを動的に増加させる

●料理のヒント

○キッチンにおける映画サウンドトラックとブレンダー・ノイズの両方の組み合わせられたラウドネス推定値の関数として、ラウドネスを動的に増加させる。

【0358】

開示されたマルチストリーム・レンダラーの第2の例示的な使用事例は、ユーザーによる何らかの問い合わせに対するスマート音声アシスタントの応答とともに、音楽のような空間的なプログラム・ストリームの同時再生に関わる。既存のスマートスピーカーでは、一般に、再生は単一の装置を通じたモノラル再生またはステレオ再生に制約されており、音声アシスタントとの対話は、典型的には、以下のステージで構成される：

30

1) 音楽再生

2) ユーザーが音声アシスタントのウェイクワードを発声する

3) スマートスピーカーがウェイクワードを認識し、かなりの量、音楽を小さくする（ダッキングする（duck））

4) ユーザーがスマートアシスタントに対してコマンドを発声する（すなわち「次の曲を再生」）。

5) スマートスピーカーがコマンドを認識し、何らかの音声応答（すなわち「了解、次の曲を再生します」）をダックした音楽に混合して、スピーカーを通じて再生することでこれを確認し、次いでコマンドを実行する。

40

6) スマートスピーカーが音楽をもとのラウドネスに戻す。

【0359】

図22および23は、空間的な音楽ミックスおよび音声アシスタント応答の同時再生を提供するマルチストリーム・レンダラーの例を示す。多数のオーケストレーションされたスマートスピーカー上で空間的オーディオを再生する場合、いくつかの実施形態は、上記の一連のイベントに対する改良を提供する。具体的には、空間ミックスは、音声アシスタントからの応答を中継するために適切に選択されたスピーカーのうちの一つまたは複数から離れるようにシフトされてもよい。音声アシスタント応答のためにこの空間を作ることは、上記に挙げられた現状技術に比べて、空間ミックスがそれほど小さくされなくてもよい

50

、または全く小さくされなくてもよいことを意味する。図22および図23は、このシナリオを示す。この例では、修正された一連のイベントは次のように発生する：

- 1) 空間的な音楽プログラム・ストリームが、図22のユーザー・クラウド2035cのための多数のオーケストレーションされたスマートスピーカーを通じて再生される。
- 2) ユーザー2020cが音声アシスタントのウェイクワードを発声する。
- 3) 一つまたは複数のスマートスピーカー（たとえば、スピーカー2005dおよび／またはスピーカー2005f）が、該一つまたは複数のスマートスピーカーに付随するマイクロフォンからの関連する録音を使用して、ウェイクワードを認識し、ユーザー2020cの位置またはユーザー2020cがどのスピーカー（単数または複数）に最も近いかを決定する。
- 4) 空間的な音楽ミックスのレンダリングは、音声アシスタント応答プログラム・ストリームがその位置の近くでレンダリングされることを予期して、前のステップで決定された位置から離れるようにシフトされる（図23のクラウド2035d）。
- 5) ユーザーが、スマートアシスタントに対して（たとえば、スマートアシスタント／バーチャル・アシスタント・ソフトウェアを実行するスマートスピーカーに対して）にコマンドを発声する。
- 6) スマートスピーカーがコマンドを認識し、対応する応答プログラム・ストリームを合成し、ユーザーの位置の近くで該応答をレンダリングする（図23のクラウド2340）。
- 7) 空間的な音楽プログラム・ストリームのレンダリングは、音声アシスタント応答が完了すると、もとの状態にシフトして戻る（図22のクラウド2035c）。

【0360】

空間的な音楽ミックスおよび音声アシスタント応答の同時再生を最適化することに加えて、空間的な音楽ミックスのシフトは、ステップ5においてスピーカーのセットが聴取者を理解する能力を向上させることもできる。これは、音楽が、聴取者の近くのスピーカーから外れるようにシフトされ、それにより、関連するマイクロフォンの音声対その他の比率が改善されるためである。

【0361】

空間的な映画ミックスおよび料理のヒントを用いた前のシナリオについて説明したものと同様に、現在のシナリオは、音声アシスタント応答の関数として空間ミックスのレンダリングをシフトさせることによって提供されるものを超えて、さらに最適化される。それ自体では、空間ミックスをシフトさせることは、音声アシスタント応答をユーザーにとって完全に理解しやすくするのに十分ではないことがありうる。単純な解決策は、現在の現状技術で要求されるほどではないが、それでも空間ミックスを一定の量だけ小さくすることである。あるいはまた、音声アシスタント応答プログラム・ストリームのラウドネスは、応答の可聴性を維持するために、空間的な音楽ミックス・プログラム・ストリームのラウドネスの関数として、動的に増大させることができる。拡張として、空間的な音楽ミックスのラウドネスも、応答ストリームに対するこのブースト・プロセスが大きくなりすぎる場合、動的にカットされてもよい。

【0362】

図24、図25、および図26は、開示される複数ストリーム・レンダラーの第3の例を示す。この例は、空間的な音楽ミックス・プログラム・ストリームおよびコンフォートノイズ・プログラム・ストリームの同時再生を管理し、同時に、赤ん坊が隣の部屋で寝たままであることを確認しようとし、それでいて赤ん坊が泣けば聞こえるようにすることに関わる。図24は、空間的な音楽ミックス（クラウド2035eによって表される）が、パーティーにおける多数の人々のために、居間2010およびキッチン2015内のすべてのスピーカーにわたって最適に再生されている開始点を示す。図25では、赤ん坊2510は、右下に描かれた隣接する寝室2505で眠ろうとしている。これを確実にするために、パーティーの人々にとっては妥当な経験を維持しつつ、空間的な音楽ミックスは、クラウド2035fによって示されているように、寝室の中での漏れを最小限にするよう寝室から動的にシフトされる。同時に、心地よいホワイトノイズを含む第2のプログラム・ストリーム（クラウド2540によって表される）が、ベビールーム内のスピーカー2005hから再生され、隣接する部屋

の音楽からの残りの漏れがあればそれをマスクする。完全なマスキングを保証するために、いくつかの例では、このホワイトノイズ・ストリームのラウドネスは、ベビールームに漏れ込む空間的な音楽のラウドネスの推定値の関数として動的に修正されうる。この推定値は、空間的な音楽のレンダリングのスピーカー・フィード、ベビールーム内のマイクロフォンからの信号、またはそれらの組み合わせから生成されてもよい。また、空間的な音楽ミックスのラウドネスは、大きすぎようになる場合には、ラウドネス修正されたノイズの関数として、動的に減衰させられてもよい。これは、第1のシナリオの空間的な映画ミックスと料理のヒントとの間のラウドネス処理と類似している。最後に、ベビールーム内のマイクロフォン（たとえば、いくつかの実装ではスマートスピーカーであってもよいスピーカー2005hに付随するマイクロフォン）が、赤ん坊からのオーディオを記録するように構成されてもよく（空間的な音楽およびホワイトノイズからピックアップされうる音は打ち消す）、これらの処理されたマイクロフォン信号の組み合わせは、次いで、第3のプログラム・ストリームとして機能しうる。該第3のプログラム・ストリームは、泣いていることが検出された場合（機械学習を通じて、パターンマッチングアルゴリズムを介してなど）、居間2010にいる、親または他のケア担当者でありうる聴取者2020dの近くで同時に再生されてもよい。図26は、クラウド2650をもって、この追加的なストリームの再生を描いている。この場合、図25のクラウド2035fの形状に対するクラウド2035gの修正された形状によって示されるように、空間的な音楽ミックスは、赤ん坊の泣き声を再生する親の近くのスピーカーから離れるようにさらにシフトされてもよい。赤ん坊の泣き声のプログラム・ストリームは、赤ん坊の泣き声が聴取者2020dに可聴なままであるように、空間的な音楽ストリームの関数としてラウドネス修正されてもよい。この例の中で考慮される3つのプログラム・ストリームの同時再生を最適化する相互接続された修正は、以下のように要約されうる：

【0363】

●居間における空間的な音楽ミックス

○ベビールーム内への伝播を減らすため、空間的レンダリングがベビールームから離れるようにシフトされる

○ベビールーム内でレンダリングされるホワイトノイズのラウドネスの関数としての、ラウドネスの動的な低減

○赤ん坊の泣き声が親の近くのスピーカーにレンダリングされることに応じて、空間的レンダリングは、親から離れるようにシフトされる

●ホワイトノイズ〔白色雑音〕

○ベビールームにしみこむ音楽ストリームのラウドネスの推定値の関数としての、ラウドネスの動的な増加

●赤ん坊の泣き声の録音

○親または他のケア担当者の位置での音楽ミックスのラウドネス推定値の関数としての、ラウドネスの動的な増加。

【0364】

次に、上述の実施形態のいくつかがどのように実装されうるかの例を説明する。

【0365】

図18Aでは、各レンダリング・ブロック1…Nは、前述のCMAP、FV、またはハイブリッド・レンダラーのような、任意の単一ストリーム・レンダラーの同一の諸インスタンスとして実装されてもよい。このようにしてマルチストリーム・レンダラーを構成することは、いくつかの便利で有用な特性を有する。

【0366】

第一に、レンダリングがこの階層的な配置で行われ、単一ストリーム・レンダラー・インスタンスのそれぞれが周波数／変換領域（たとえばQMF）で動作するように構成されている場合、ストリームの混合も周波数／変換領域で発生する可能性があり、逆変換はMチャンネルについて、1回実行される必要があるだけである。これは、N×M回の逆変換を実行して時間領域で混合するのに対して、著しい効率改善である。

【0367】

図27は、図18Aに示されるマルチストリーム・レンダラーの周波数／変換領域の例を示す。この例では、各プログラム・ストリームが、レンダリング・モジュール1～Nのうちの対応するものによって受領される前に、プログラム・ストリーム1～Nのそれぞれに対して直交ミラー解析フィルタバンク（QMF）が適用される。この例によれば、レンダリング・モジュール1～Nは、周波数領域において動作する。ミキサー1830aがレンダリング・モジュール1～Nの出力を混合した後、逆合成フィルタバンク2735aは、該混合を時間領域に変換し、時間領域の混合されたスピーカー・フィード信号をラウドスピーカー1～Mに提供する。この例では、直交ミラーフィルタバンク、レンダリング・モジュール1～N、ミキサー1830a、および逆フィルタバンク2735aは、制御システム610cの構成要素である。

10

【0368】

図28は、図18Bに示されるマルチストリーム・レンダラーの周波数／変換領域の例を示す。図17のように、各プログラム・ストリームが、レンダリング・モジュール1～Nのうちの対応するものによって受領される前に、プログラム・ストリーム1～Nのそれぞれに対して直交ミラー解析フィルタバンク（QMF）が適用される。この例によれば、レンダリング・モジュール1～Nは、周波数領域において動作する。この実装では、マイクロフォン・システム620bからの時間領域マイクロフォン信号は、直交ミラーフィルタバンクにも提供され、レンダリング・モジュール1～Nは、周波数領域においてマイクロフォン信号を受領する。ミキサー1830bがレンダリング・モジュール1～Nの出力を混合した後、逆フィルタバンク2735bは、該混合を時間領域に変換し、時間領域の混合されたスピーカー・フィード信号をラウドスピーカー1～Mに提供する。この例では、直交ミラーフィルタバンク、レンダリング・モジュール1～N、ミキサー1830b、および逆フィルタバンク2735bは、制御システム610dの構成要素である。

20

【0369】

図29を参照して、実施形態の別の例を説明する。本明細書で提供される他の図と同様に、図29に示される要素のタイプおよび数は、単に例として与えられている。他の実装は、より多くの、より少ない、および／または異なるタイプおよび数の要素を含みうる。図29は、この例における居住空間である、聴取環境のフロアプランを示す。この例によれば、環境2000は、左上の居間2010、中央下のキッチン2015、および右下のベッドルーム2505を含む。居住空間にわたって分布する四角および円は、スペースに都合のよい位置に配置されているが、標準で規定されたレイアウトには準拠しない（任意に配置された）一組のラウドスピーカー2005a～2005hを表し、そのうちのいくつかは、いくつかの実装ではスマートスピーカーであってもよい。いくつかの例では、ラウドスピーカー2005a～2005hは、一つまたは複数の開示された実施形態を実装するために協調させられてもよい。たとえば、いくつかの実施形態では、ラウドスピーカー2005a～2005hは、いくつかの例ではCHASMであってもよいオーディオ・セッション・マネージャを実装している装置からのコマンドに従って協調させられてもよい。いくつかのそのような例において、開示された柔軟レンダリング機能を含むがこれに限定されない開示されたオーディオ処理は、少なくとも部分的には図2C、図2D、図3Cおよび図4を参照して上述したCHASM 208C、CHASM 208D、CHASM 307および／またはCHASM 401からの命令に従って、実装されてもよい。いくつかの実装によれば、方法1900のブロックは、少なくとも部分的には、本明細書でオーディオ・セッション・マネージャと称されるもの、たとえば、CHASMを実装している装置からの命令に従って実行されてもよい。いくつかのそのような例において、方法1900のブロックは、少なくとも部分的には、図2C、図2D、図3Cおよび図4を参照して上述したCHASM 208C、CHASM 208D、CHASM 307および／またはCHASM 401からの命令に従って、実装されてもよい。代替例においては、オーディオ・セッション・マネージャを実装している装置が、方法1900のブロックも実装してもよい。この例では、環境2000は、環境全体に分散されたカメラ2911a～2911eを含む。いくつかの実装では、環境2000内の一つまたは複数のスマート・オーディオ装置も、一つまたは複数のカメラを含んでいてもよい。前記一つまたは複数のスマート・オーディオ装置は、単一

30

40

50

目的のオーディオ装置またはバーチャル・アシスタントであってもよい。いくつかのそのような例では、一つまたは複数のカメラ（これは、図6の任意的なセンサー・システム630のカメラであってもよい）は、テレビ2030の中または表面上、携帯電話の中、またはラウドスピーカー2005b、2005d、2005eもしくは2005hの一つまたは複数のようなスマートスピーカー内に存在してもよい。カメラ2911a~2911eは、本開示において提示される聴取環境のすべての図に示されているわけではないが、それでも、環境2000を含むがそれに限定されない聴取環境のそれぞれは、いくつかの実装において一つまたは複数のカメラを含んでいてもよい。

【0370】

図30、図31、図32、および図33は、図29に示される居住空間における複数の異なる聴取位置および配向についての基準空間モードにおいて空間的オーディオを柔軟にレンダリングする例を示している。図30~図33は、4つの例示的な聴取位置におけるこの能力を示す。各例において、人3020aのほうを指す矢印3005は、（人3020aが面している）前方サウンドステージの位置を表す。各例において、矢印3010aは左サラウンド・フィールドを表し、矢印3010bは右サラウンド・フィールドを表す。

【0371】

図30では、居間のカウチ2025に座っている人3020aのために、基準空間モードが決定され（たとえば、オーディオ・セッション・マネージャを実装している装置によって）、空間的オーディオが柔軟にレンダリングされている。図30に示される例では、居間2010およびキッチン2015にあるラウドスピーカーの全部が、与えられたラウドスピーカー能力およびレイアウトで、聴取者3020aのまわりにあるオーディオ・データの最適化された空間的再生を作り出すために使用される。この最適な再生は、アクティブなラウドスピーカーの限界内にあるクラウド3035によって視覚的に表現されている。

【0372】

いくつかの実装によれば、オーディオ・セッション・マネージャを実装するように構成された制御システム（図6の制御システム610のような）は、図6のインターフェース・システム605のようなインターフェース・システムを介して受領される基準空間モード・データに従って、基準空間モードの想定された傾聴位置および／または想定された配向を決定するように構成されてもよい。いくつかの例を以下に説明する。いくつかのそのような例では、基準空間モード・データは、マイクロフォン・システム（図6のマイクロフォン・システム120のような）からのマイクロフォン・データを含んでいてもよい。

【0373】

いくつかのそのような例では、基準空間モード・データは、ウェイクワードおよび音声コマンド、たとえば「[ウェイクワード]、テレビを前方サウンドステージにしてください」に対応するマイクロフォン・データを含んでいてもよい。代替的または追加的に、マイクロフォン・データは、たとえば到来方向（DOA）データを介して、ユーザーの声の音に応じてユーザーの位置を三角測量するために使用されてもよい。たとえば、3つ以上のラウドスピーカー2005a~2005eは、マイクロフォン・データを使用して、DOAデータを介して、人3020aの声の音に応じて、居間のカウチ2025に座っている人3020aの位置を三角測量してもよい。人3020aの配向は、人3020aの位置に従って想定されてもよい：人3020aが図30に示される位置にいる場合、人3020aは、テレビジョン2030に面していると想定されてもよい。

【0374】

代替的または追加的に、人3020aの位置および配向は、カメラシステム（図6のセンサー・システム130など）からの画像データに従って決定されてもよい。

【0375】

いくつかの例では、人3020aの位置および配向は、グラフィカル・ユーザー・インターフェース（GUI）を介して得られたユーザー入力に従って決定されてもよい。いくつかのそのような例によれば、制御システムは、人3020aが人3020aの位置および配向を入力することを許容するGUIを提示するために、表示装置（たとえば、セルラー電話の表示装置

10

20

30

40

50

)を制御するように構成されてもよい。

【0376】

図31では、居間の読書椅子3115に座っている人3020aのために、基準空間モードが決定され、空間的オーディオが柔軟にレンダリングされている。図32では、キッチンカウンター330の隣に立っている人3020aのために、基準空間モードが決定され、空間的オーディオが柔軟にレンダリングされている。図33では、朝食テーブル340に座っている人3020aのために、基準空間モードが決定され、空間オーディオが柔軟にレンダリングされている。矢印3005によって示されるように、前方サウンドステージ配向は、必ずしも環境2000内の特定のラウドスピーカーに対応しないことが観察されうる。聴取者の位置および配向が変わるにつれて、空間ミックスのさまざまな成分をレンダリングするための諸スピーカーの役割も変わる。

10

【0377】

図30～図33のいずれかの人3020aについて、その人は、図示された位置および配向のそれぞれについて意図されたように、空間ミックスを聞く。しかしながら、この経験は、空間内の追加的な聴取者にとって最適ではない可能性がある。図34は、2人の聴取者が聴取環境の異なる位置にいる場合の基準空間モード・レンダリングの例を示す。図12Eは、カウチにいる人3020aおよびキッチンに立っている人3020bのための基準空間モード・レンダリングを示す。レンダリングは、人3020aにとって最適であるが、人3020bは、その位置を考えると、ほとんどサラウンド・フィールドからの信号を聞き、前方サウンドステージはほとんど聞こえない。この場合、および複数の人々が空間にいて、予測不可能な仕方

20

【0378】

図35は、聴取者の位置および配向に関するユーザー入力を受領するためのGUIの例を示す。この例によれば、ユーザーは、いくつかの可能な聴取位置および対応する配向を前もって特定している。各位置および対応する方向に対応するラウドスピーカー位置は、セットアップ・プロセス中にすでに入力され、記憶されている。いくつかの例が本明細書に記載される。オーディオ装置自動位置特定プロセスの詳細な例は後述する。たとえば、聴取環境レイアウトGUIが提供されていてもよく、ユーザーは、可能な聴取位置およびスピーカー位置に対応する位置にタッチし、該可能な聴取位置に名前を付けるように促されていてもよい。この例では、図35に示された時点で、ユーザーは、仮想ボタン「居間カウチ」をタッチすることによって、ユーザーの位置に関するユーザー入力をすでにGUI 3500に与えている。L字型カウチ2025のため、2つの可能な前向き位置があるので、ユーザーは、ユーザーがどの方向に面しているかを示すようにプロンプトを受けている。

30

【0379】

図36は、環境内の3つのオーディオ装置間の幾何学的関係の例を示している。この例では、環境3600は、テレビジョン3601、ソファ3603および5つのオーディオ装置3605を含む部屋である。この例によれば、オーディオ装置3605は、環境3600の位置1～5にある。この実装では、各オーディオ装置3605は、少なくとも3つのマイクロフォンを有するマイクロフォン・システム3620と、少なくとも1つのスピーカーを含むスピーカー・システム3625とを含む。いくつかの実装では、各マイクロフォン・システム3620は、マイクロフォンのアレイを含む。いくつかの実装によれば、オーディオ装置3605のそれぞれは、少なくとも3つのアンテナを含むアンテナシステムを含んでいてもよい。

40

【0380】

本明細書に開示される他の例と同様に、図36に示される要素のタイプ、数、および配置は、単に例である。他の実装は、異なるタイプ、数および配置の要素をもつことがあり、たとえば、より多数またはより少数のオーディオ装置3605、異なる位置にあるオーディ

50

オ装置3605などを有していてもよい。

【0381】

この例では、三角形3610aは、位置1、2および3に頂点を有する。ここで、三角形3610aは、辺12、23aおよび13aを有する。この例によれば、辺12と23との間の角度は θ_2 であり、辺12と13aとの間の角度は θ_1 であり、辺23aと13aとの間の角度は θ_3 である。これらの角度は、のちにより詳細に記載されるように、DOAデータに従って決定されてもよい。

【0382】

いくつかの実装では、三角形の辺の相対的な長さのみが決定されてもよい。代替的な実装では、三角形の辺の実際の長さが推定されてもよい。いくつかのそのような実装によれば、三角形の辺の実際の長さは、TOAデータに従って、たとえばある三角形頂点に位置するオーディオ装置によって生成され、別の三角形頂点に位置するオーディオ装置によって検出される音の到達時間に従って、推定されうる。代替的または追加的に、三角形の辺の長さは、ある三角形頂点に位置するオーディオ装置によって生成され、別の三角形頂点に位置するオーディオ装置によって検出される電磁波によって推定されてもよい。たとえば、三角形の辺の長さは、ある三角形頂点に位置するオーディオ装置によって生成され、別の三角形頂点に位置するオーディオ装置によって検出される電磁波の信号強度によって推定されうる。いくつかの実装では、三角形の辺の長さは、電磁波の検出された位相シフトに従って推定されてもよい。

【0383】

図37は、図36に示される環境における3つのオーディオ装置の間の幾何学的関係の別の例を示している。この例では、三角形3610bは、位置1、3および4に頂点を有する。ここで、三角形3610bは、辺13b、14、および34aを有する。この例によれば、辺13bと14との間の角度は θ_4 であり、辺13bと34aとの間の角度は θ_5 であり、側面34aと14との間の角度は θ_6 である。

【0384】

図36および図37を比較することによって、三角形3610aの辺13aの長さは、三角形3610bの辺13bの長さと同じ長さであることが観察されうる。いくつかの実装では、ある三角形（たとえば、三角形3610a）の辺の長さが正しいと想定されてもよく、隣接する三角形によって共有される辺の長さは、この長さに制約される。

【0385】

図38は、図36および図37に示された三角形の両方を、対応するオーディオ装置および環境の他の特徴なしに、示す。図38は、三角形3610aおよび3610bの辺の長さおよび角度配向の推定値を示す。図38に示される例では、三角形3610bの辺13bの長さは、三角形3610aの辺13aと同じ長さに制約される。三角形3610bの他の辺の長さは、結果として得られる辺13bの長さの変化に比例して、スケーリングされる。結果として得られる三角形3610b'は、三角形3610aに隣接して、図38に示されている。

【0386】

いくつかの実装によれば、三角形3610aおよび3610bに隣接する他の三角形の辺の長さは、環境3600内のすべてのオーディオ装置位置が決定されるまで、みな、同様の仕方決定されうる。

【0387】

オーディオ装置位置のいくつかの例は次のようになりうる。各オーディオ装置は、環境内の他のすべてのオーディオ装置によって生成された音に基づいて、環境（たとえば、部屋）内の他のすべてのオーディオ装置のDOAを報告してもよい（たとえば、CHASMのようなオーディオ・セッション・マネージャを実装している装置からの命令に従って）。i番目のオーディオ装置のデカルト座標は

【数50】

$$\mathbf{x}_i = [x_i, y_i]^T$$

10

20

30

40

50

として表されてもよい。ここで、上付きのTはベクトル転置を示す。環境内のM個のオーディオ装置が与えられ、 $i=\{1\cdots M\}$ である。

【0388】

図39は、3つのオーディオ装置によって形成される三角形の内角を推定する例を示す。この例では、オーディオ装置はi、j、kである。装置iから観測される、装置jから発する音源のDOAは θ_{ji} と表されてもよい。装置iから観測される、装置kから発する音源のDOAは θ_{ki} と表されてもよい。図39に示される例では、 θ_{ji} および θ_{ki} は軸3905aから測られ、軸3905aの配向は任意であり、軸3905aはたとえば、オーディオ装置iの配向に対応してもよい。三角形2110の内角aは、 $a=\theta_{ki}-\theta_{ji}$ と表されてもよい。内角aの計算は、軸3905aの配向に依存しないことが観察されうる。

10

【0389】

図39に示される例では、 θ_{ij} および θ_{kj} は軸3905bから測られ、軸3905bの配向は任意であり、軸3905bはオーディオ装置jの配向に対応してもよい。三角形3910の内角bは、 $b=\theta_{ij}-\theta_{kj}$ と表されてもよい。同様に、この例において、 θ_{jk} および θ_{ik} は軸3905cから測られる。三角形3910の内角cは、 $c=\theta_{jk}-\theta_{ik}$ と表されてもよい。

【0390】

測定誤差がある場合、 $a+b+c\neq 180^\circ$ である。各角度を他の2つの角度から予測し、たとえば次のようにして平均することにより、堅牢性が改善できる。

【数51】

$$\tilde{a}=0.5(a+\operatorname{sgn}(a)(180-|b+c|)).$$

20

【0391】

いくつかの実装では、エッジ長(A,B,C)は、正弦法則を適用することによって（スケール誤差を除いて）計算されうる。いくつかの例では、1つのエッジ長は、1のような任意の値を割り当てられてもよい。たとえば、 $A=1$ とし、頂点

【数52】

$$\hat{\mathbf{x}}_a=[0,0]^T$$

30

を原点に置くことによって、残りの2つの頂点の位置は次のように計算されうる：

【数53】

$$\hat{\mathbf{x}}_b=[A\cos a,-A\sin a]^T,\hat{\mathbf{x}}_c=[B,0]^T$$

しかしながら、任意の回転は受け容れ可能でありうる。

【0392】

いくつかの実装によれば、三角形パラメータ化のプロセスは、環境中の3つのオーディオ装置のあらゆる可能なサブセットについて繰り返されてもよい。それらのサブセットは、サイズ

40

【数54】

$$N=\binom{M}{3}$$

のスーパーセット ζ において列挙される。いくつかの例では、 T_l がlの三角形を表してもよい。実装に依存して、三角形は、いかなる特定の順序で列挙されなくてもよい。三角形は、DOAおよび／または辺長の推定値における可能性のある誤差のため、重なり合う可能性があり、完全に整列しない可能性がある。

【0393】

50

図40は、図6に示されるような装置によって実行されうる方法の一例を概説するフロー図である。方法4000のブロックは、本明細書に記載される他の方法と同様に、必ずしも示される順序で実行されるわけではない。さらに、そのような方法は、図示および／または記載されているよりも多いまたは少ないブロックを含んでいてもよい。この実装では、方法4000は、環境におけるスピーカーの位置を推定することを含む。方法4000のブロックは、図6に示される装置600であってもよい（または、それを含んでいてもよい）一つまたは複数の装置によって実行されてもよい。いくつかの実装によれば、方法4000のブロックは、少なくとも部分的には、オーディオ・セッション・マネージャ（たとえばCHASM）を実装している装置によって、および／またはオーディオ・セッション・マネージャを実装している装置からの命令に従って実行されてもよい。いくつかのそのような例において、方法4000のブロックは、少なくとも部分的には、図2C、図2D、図3Cおよび図4を参照して上述したCHASM 208C、CHASM 208D、CHASM 307および／またはCHASM 401からの命令に従って、実行されてもよい。いくつかの実装によれば、方法4000のブロックは、図15を参照して上述した方法1500のセットアップ・プロセスの一部として実行されてもよい。

【0394】

この例では、ブロック4005は、複数のオーディオ装置の各オーディオ装置について到来方向（DOA）データを取得することを含む。いくつかの例では、複数のオーディオ装置は、図36に示されるオーディオ装置3605の全部のような、環境内のオーディオ装置の全部を含んでいてもよい。

【0395】

しかしながら、場合によっては、複数のオーディオ装置は、環境内のすべてのオーディオ装置のサブセットのみを含んでいてもよい。たとえば、複数のオーディオ装置は、環境内のすべてのスマートスピーカーを含むが、環境内の他のオーディオ装置の一つまたは複数は含まないのでもよい。

【0396】

DOAデータは、特定の実装に依存して、さまざまな仕方で取得されうる。いくつかの事例では、DOAデータを決定することは、前記複数のオーディオ装置のうちの少なくとも1つのオーディオ装置についてDOAデータを決定することを含んでいてもよい。たとえば、DOAデータを決定することは、前記複数のオーディオ装置のうちの単一オーディオ装置に対応する複数のオーディオ装置マイクロフォンの各マイクロフォンからマイクロフォン・データを受領し、少なくとも部分的には該マイクロフォン・データに基づいて、単一オーディオ装置についてのDOAデータを決定することを含んでいてもよい。代替的または追加的に、DOAデータを決定することは、前記複数のオーディオ装置のうちの単一のオーディオ装置に対応する一つまたは複数のアンテナからアンテナデータを受領し、該アンテナデータに少なくとも部分的に基づいて、単一のオーディオ装置についてのDOAデータを決定することを含んでいてもよい。

【0397】

いくつかのそのような例では、単一のオーディオ装置自体がDOAデータを決定してもよい。いくつかのそのような実装によれば、前記複数のオーディオ装置の各オーディオ装置は、それ自身のDOAデータを決定してもよい。しかしながら、他の実装では、ローカルまたはリモート装置でありうる他の装置が、環境中の一つまたは複数のオーディオ装置についてのDOAデータを決定してもよい。いくつかの実装によれば、サーバーが、環境中の一つまたは複数のオーディオ装置についてのDOAデータを決定してもよい。

【0398】

この例によれば、ブロック4010は、DOAデータに基づいて複数の三角形のそれぞれについての内角を決定することを含む。この例では、前記複数の三角形の各三角形は、前記オーディオ装置のうちの3つのオーディオ装置のオーディオ装置位置に対応する頂点を有する。いくつかのそのような例は、上述されている。

【0399】

図41は、環境内の各オーディオ装置が複数の三角形の頂点である例を示している。各三角形の辺は、2つのオーディオ装置3605の間の距離に対応する。

【0400】

この実装では、ブロック4015は、各三角形の各辺の辺長を決定することを含む。（三角形の辺は、本明細書では「エッジ」と呼ばれることもある。）この例によれば、辺長は、少なくとも部分的には、内角に基づいている。場合によっては、辺長は、三角形の第1の辺の第1の長さを決定し、三角形の内角に基づいて三角形の第2の辺および第3の辺の長さを決定することによって計算されてもよい。いくつかのそのような例が上述されている。

【0401】

いくつかのそのような実装によれば、第1の長さを決定することは、第1の長さを所定の値に設定することを含んでいてもよい。しかしながら、第1の長さの決定は、いくつかの例では、到達時間データおよび／または受信信号強度データに基づいてもよい。到達時間データおよび／または受信信号強度データは、いくつかの実装では、環境内の第2のオーディオ装置によって検出される、環境内の第1のオーディオ装置からの音波に対応してもよい。代替的または追加的に、到達時間データおよび／または受信信号強度データは、環境内の第2のオーディオ装置によって検出される、環境内の第1のオーディオ装置からの電磁波（たとえば、電波、赤外線波など）に対応しうる。

【0402】

この例によれば、ブロック4020は、第1のシーケンス内の前記複数の三角形のそれぞれを整列させる順整列（forward alignment）プロセスを実行することを含む。この例によれば、順整列プロセスは、順整列行列を生成する。

【0403】

いくつかのそのような例によれば、三角形は、たとえば図38に示され、上述されるように、エッジ (x_i, x_j) が隣接エッジに等しくなるような仕方では整列することが期待される。

【数55】

ε

〔 ε と書くこともある〕が、サイズ

【数56】

$$P = \binom{M}{2}$$

のすべてのエッジの集合であるとする。いくつかのそのような実装では、ブロック4020は、 ε を通じてたどっていき、前に整列されたエッジのものと一致するようにエッジを強制することによって、諸三角形の共通エッジを正順に整列させることを含んでいてもよい。

【0404】

図42は、順整列プロセスの一部の例を与えている。図42において太字で示されている番号1から5は、図36、図37、図41に示されているオーディオ装置位置に対応している。図42に示され、本明細書に記載される順整列プロセスのシーケンスは、単なる例である。

【0405】

この例では、図38のように、三角形3610bの辺13bの長さは、三角形3610aの辺13aの長さとは一致するように強制される。結果として得られる三角形3610b'は、図42に示されており、同じ内角が維持されている。この例によれば、三角形3610cの辺13cの長さも、三角形3610aの辺13aの長さとは一致するように強制される。結果として得られる三角形3610c'は、図42に示されており、同じ内角が維持されている。

【0406】

次に、この例では、三角形3610dの辺34bの長さは、三角形3610b'の辺34aの長さとは一致するように強制される。さらに、この例では、三角形3610dの辺23bの長さは、三角形3610aの辺23aの長さとは一致するように強制される。結果として得られる三角形3610d

10

20

30

40

50

'は、図42に示されており、同じ内角が維持されている。いくつかのそのような例によれば、図41に示される残りの三角形は、三角形3610b、3610c、および3610dと同じ仕方で処理されうる。

【0407】

順整列プロセスの結果は、データ構造に記憶されてもよい。いくつかのそのような例によれば、順整列プロセスの結果は、順整列行列に格納されてもよい。たとえば、順整列プロセスの結果は、行列

【数57】

$$\vec{X} \in \mathbb{R}^{3N \times 2}$$

10

に格納されてもよく、ここで、Nは、三角形の総数を示す。

【0408】

DOAデータおよび／または初期の辺長決定が誤差を含む場合、オーディオ装置位置の複数の推定値が生じる。これらの誤差は、一般に、順整列プロセスの間に増加する。

【0409】

図43は、順整列プロセスの間に生じたオーディオ装置位置の複数の推定値の例を示している。この例では、順整列プロセスは、頂点として7つのオーディオ装置位置を有する三角形に基づく。ここで、三角形は、DOA推定値における加法的誤差のため、完全には整列しない。図43に示される番号1～7の位置は、順整列プロセスによって生成される推定オーディオ装置位置に対応する。この例では、「1」とラベル付けされたオーディオ装置位置推定値は一致するが、オーディオ装置6および7についてのオーディオ装置位置推定値は、より大きな差を示している。このことは、数字6と7が上に位置する比較的大きなエリアによって示される。

20

【0410】

図40に戻ると、この例では、ブロック4025は、第1のシーケンスの反転である第2のシーケンスにおいて、前記複数の三角形のそれぞれを整列する逆整列プロセスを含む。いくつかの実装によれば、逆整列プロセスは、以前と同様に、ただし逆順で、εをたどっていくことを含む。代替的な例では、逆整列プロセスは、正確には、順整列プロセスの動作のシーケンスの逆でなくてもよい。この例によれば、逆整列プロセスは、逆整列行列を生成し、これは、本明細書では、

30

【数58】

$$\vec{X} \in \mathbb{R}^{3N \times 2}$$

と表わされてもよい。

【0411】

図44は、逆整列プロセスの一部の例を示す。図26に太字で示されている番号1～5は、図36、図37、および図41に示されているオーディオ装置位置に対応している。図44に示され、本明細書に記載される逆整列プロセスのシーケンスは、単に一例である。

40

【0412】

図44に示される例では、三角形3610eは、オーディオ装置位置3、4、および5に基づいている。この実装では、三角形3610eの辺長（または「エッジ」）は正しいと仮定され、隣接する三角形の辺長はそれらと一致するように強制される。この例によれば、三角形3610fの辺45bの長さは、三角形3610eの辺45aの長さと同じように強制される。結果として得られる三角形3610f'は、内角が同じままであり、図44に示されている。この例では、三角形3610cの辺35bの長さは、三角形3610eの辺35aの長さと同じように強制される。結果として得られる三角形3610c"は、内角が同じままであり、図44に示されている。いくつかのそのような例によれば、図5に示されている残りの三角形は、逆整列プロセスが残りの三角形のすべてを含めてしまうまで、三角形3610cおよび3610fと同様

50

に処理されうる。

【0413】

図45は、逆整列プロセス中に生じたオーディオ装置位置の複数の推定値の例を示す。この例では、逆整列プロセスは、図43を参照して上述した、同じ7つのオーディオ装置位置を頂点としてもつ三角形に基づく。図45に示される番号1から7の位置は、逆整列プロセスによって生成される推定オーディオ装置位置に対応する。ここでもまた、これらの三角形は、DOA推定値の加法的誤差のため、完全には整列しない。この例では、6および7とラベル付けされたオーディオ装置位置推定値は一致するが、オーディオ装置1および2についてのオーディオ装置位置推定値は、より大きな差を示す。

【0414】

図40に戻ると、ブロック4030は、少なくとも部分的には、順整列行列の値および逆整列行列の値に基づいて、各オーディオ装置位置の最終的な推定値を生成することを含む。いくつかの例では、各オーディオ装置の位置の最終的な推定値を生成することは、順整列行列を並進およびスケールリングして、並進およびスケールリングされた順整列行列を生成することと、逆整列行列を並進およびスケールリングして、並進およびスケールリングされた逆整列行列を生成することとを含むことができる。

【0415】

たとえば、並進およびスケールリングは、重心を原点に移動させ、単位フロベニウスノルムを強制すること、たとえば

【数59】

$$\bar{\mathbf{X}} = \bar{\mathbf{X}} / \|\bar{\mathbf{X}}\|_2^F \quad \text{および} \quad \hat{\mathbf{X}} = \hat{\mathbf{X}} / \|\hat{\mathbf{X}}\|_2^F$$

によって固定される。

【0416】

いくつかのそのような例によれば、各オーディオ装置位置の最終的な推定値を生成することは、並進およびスケールリングされた順整列行列および並進およびスケールリングされた逆方向整列行列に基づいて回転行列を生成することを含んでもよい。回転行列は、各オーディオ装置についての複数の推定されたオーディオ装置位置を含んでもよい。順方向整列と逆方向整列との間の最適な回転は、たとえば、特異値分解によって見つけることができる。いくつかのそのような例では、回転行列を生成することは、並進およびスケールリングされた順整列行列および並進およびスケールリングされた逆整列行列に対して、たとえば次のように特異値分解を実行することを含みうる：

【数60】

$$\mathbf{U}\Sigma\mathbf{V} = \bar{\mathbf{X}}^T \hat{\mathbf{X}}$$

【0417】

上記の式において、Uは左特異ベクトルを表し、Vは行列

【数61】

$$\bar{\mathbf{X}}^T \hat{\mathbf{X}}$$

の右特異ベクトルをそれぞれ表す。Σは特異値の行列を表す。上記の式は回転行列 $\mathbf{R} = \mathbf{V}\mathbf{U}^T$ を与える。行列積 $\mathbf{V}\mathbf{U}^T$ は、

【数62】

$\mathbf{R}\hat{\mathbf{X}}$ が $\bar{\mathbf{X}}$ と整列するよう最適に回転される

ような回転行列を生成する。

【0418】

いくつかの例によれば、回転行列 $R = VU^T$ を決定した後、諸整列は、たとえば次のように平均されてもよい：

【数63】

$$\vec{X} = 0.5(\vec{X} + R\vec{X})$$

【0419】

いくつかの実装では、各オーディオ装置位置の最終推定値を生成することは、各オーディオ装置の推定オーディオ装置位置を平均して、各オーディオ装置位置の最終推定値を生成することも含む。DOAデータおよび／または他の計算が有意な誤りを含む場合でさえ、さまざまな開示された実装は堅牢であることが証明された。たとえば、

【数64】

$$\vec{X}$$

は、複数の三角形からの重複する頂点のため、同じノードの $(N-1)(N-2)/2$ 個の推定値を含む。共通のノードを通じた平均をとると、最終的な推定値

【数65】

$$\hat{X} \in \mathbb{R}^{M \times 3}$$

が得られる。

【0420】

図46は、推定されたオーディオ装置位置と実際のオーディオ装置位置との比較を示す。図28に示される例では、オーディオ装置位置は、図43および図45を参照して上述した順方向および逆方向の整列プロセスの間に推定された位置に対応する。これらの例では、DOA推定における誤差は15度の標準偏差を有していた。それにもかかわらず、各オーディオ装置位置の最終的な推定値（それぞれが図46の「x」で表される）は、実際のオーディオ装置位置（それぞれが図46の円で表される）とよく一致する。

【0421】

前述の議論の多くは、オーディオ装置の自動位置特定（auto-location）に関わる。以下の議論は、上記で簡単に説明した、聴取者の位置および聴取者の角度配向を決定するいくつかの方法に基づいて敷衍する。上述の説明において、用語「回転（rotation）」は、本質的に、以下の説明において用語「配向（orientation）」が使用されるのと同じように使用されている。たとえば、上述の「回転」は、最終的なスピーカー幾何のグローバルな回転を指してもよく、図40以下を参照して上述したプロセスの間の個々の三角形の回転ではない。このグローバルな回転または配向は、聴取者の角度配向を参照して、たとえば、聴取者が見ている方向によって、聴取者の鼻が向いている方向によって、などで、解決されてもよい。

【0422】

聴取者位置を推定するためのさまざまな満足できる方法が以下に記載される。しかしながら、聴取者角度配向の推定は困難であることがある。以下に、いくつかの関連する方法について詳述する。

【0423】

聴取者位置および聴取者角度配向を決定することは、位置特定されたオーディオ装置を聴取者に対して配向させるなど、いくつかの望ましい特徴を可能にすることができる。聴取者の位置と角度配向を知ることは、聴取者に対して、たとえば環境内のどのスピーカーが前方にあり、どれが後方にあり、どれが（もしあれば）中央近くにあるかなどの決定を

許容する。

【0424】

オーディオ装置位置と聴取者の位置と方向との間で相関をとった後、いくつかの実装は、オーディオ装置の位置データ、オーディオ装置の角度配向データ、聴取者の位置データ、聴取者の角度配向データをオーディオ・レンダリング・システムに提供することに関わってもよい。代替的または追加的に、いくつかの実装は、オーディオ装置の位置データ、オーディオ装置の角度配向データ、聴取者位置データ、および聴取者角度配向データに少なくとも部分的に基づいたオーディオ・データ・レンダリング・プロセスに関わってもよい。

【0425】

図47は、図6に示されるような装置によって実行されうる方法の一例を概説するフロー図である。方法4700のブロックは、本明細書に記載される他の方法と同様に、必ずしも示される順序で実施されるわけではない。さらに、そのような方法は、図示および／または記載されているよりも多いまたは少ないブロックを含んでいてもよい。この例では、方法4700のブロックは、図6に示される制御システム610であってもよく（または、それを含んでいてもよい）制御システムによって実行される。上述したように、いくつかの実装では、制御システム610は単一の装置内に存在してもよく、他の実装では、制御システム610は2つ以上の装置内に存在してもよい。いくつかの実装によれば、方法4700のブロックは、少なくとも部分的には、オーディオ・セッション・マネージャ（たとえばCHASM）を実装している装置によって、および／またはオーディオ・セッション・マネージャを実装している装置からの命令に従って実行されてもよい。いくつかのそのような例において、方法4700のブロックは、少なくとも部分的には、図2C、図2D、図3Cおよび図4を参照して上述したCHASM 208C、CHASM 208D、CHASM 307および／またはCHASM 401からの命令に従って、実行されてもよい。いくつかの実装によれば、方法4700のブロックは、図12を参照して上述した方法1200のプロセス、たとえば、第1の人の第1の位置および第1の配向を決定し、第1の人の第2の位置および第2の配向を決定することなどの一部として実行されてもよい。

【0426】

この例では、ブロック4705は、環境内の複数のオーディオ装置の各オーディオ装置について到来方向（DOA）データを取得することに関わる。いくつかの例では、複数のオーディオ装置は、図36に示されるオーディオ装置3605のすべてなど、環境内のオーディオ装置のすべてを含んでいてもよい。

【0427】

しかしながら、場合によっては、複数のオーディオ装置は、環境内のすべてのオーディオ装置のサブセットのみを含んでいてもよい。たとえば、複数のオーディオ装置は、環境内のすべてのスマートスピーカーを含んでいてもよいが、環境内の他のオーディオ装置の一つまたは複数は含まなくてもよい。

【0428】

DOAデータは、特定の実装に依存して、さまざまな仕方で取得されうる。いくつかの事例では、DOAデータを決定することは、前記複数のオーディオ装置のうちの少なくとも1つのオーディオ装置についてDOAデータを決定することを含んでいてもよい。いくつかの例では、DOAデータは、試験信号を再生するよう環境中の複数のラウドスピーカーの各ラウドスピーカーを制御することによって得られてもよい。たとえば、DOAデータを決定することは、複数のオーディオ装置のうちの単一のオーディオ装置に対応する複数のオーディオ装置マイクロフォンの各マイクロフォンからマイクロフォン・データを受領し、少なくとも部分的に、マイクロフォン・データに基づいて単一のオーディオ装置についてのDOAデータを決定することに関わってもよい。代替的または追加的に、DOAデータを決定することは、複数のオーディオ装置のうちの単一のオーディオ装置に対応する一つまたは複数のアンテナからアンテナデータを受領し、アンテナデータに少なくとも部分的に基づいて単一のオーディオ装置についてのDOAデータを決定することに関わってもよい。

【0429】

いくつかのそのような例では、単一のオーディオ装置自体がDOAデータを決定してもよい。いくつかのそのような実装によれば、複数のオーディオ装置の各オーディオ装置は、それ自身のDOAデータを決定してもよい。しかしながら、他の実装では、ローカルまたはリモート装置でありうる他の装置が、環境中の一つまたは複数のオーディオ装置についてのDOAデータを決定しうる。いくつかの実装によれば、サーバーが、環境中の一つまたは複数のオーディオ装置についてのDOAデータを決定してもよい。

【0430】

図47に示される例によれば、ブロック4710は、制御システムを介して、DOAデータに少なくとも部分的に基づいてオーディオ装置位置データを生成することを含む。この例では、オーディオ装置位置データは、ブロック4705で参照される各オーディオ装置についてオーディオ装置位置の推定値を含む。

10

【0431】

オーディオ装置位置データは、たとえば、デカルト座標系、球面座標系または円筒座標系などの座標系の座標であってもよい（またはそれを含んでもよい）。この座標系は、本明細書ではオーディオ装置座標系と称されることがある。いくつかのそのような例では、オーディオ装置座標系は、環境中のオーディオ装置のうちの1つを参照して配向されてもよい。他の例では、オーディオ装置座標系は、環境中の2つのオーディオ装置の間の線によって定義される軸を参照して配向されてもよい。しかしながら、他の例では、オーディオ装置座標系は、テレビジョン、部屋の壁などの環境の別の部分を参照して配向されてもよい。

20

【0432】

いくつかの例では、ブロック4710は、図40を参照して上述したプロセスに関わってもよい。いくつかのそのような例によれば、ブロック4710は、DOAデータに基づいて複数の三角形のそれぞれについての内角を決定することに関わってもよい。いくつかの事例では、複数の三角形の各三角形は、3つのオーディオ装置のオーディオ装置位置に対応する頂点を有していてもよい。いくつかのそのような方法は、少なくとも部分的に、前記内角に基づいて、各三角形の各辺についての辺長を決定することに関わってもよい。

【0433】

いくつかのそのような方法は、第1のシーケンスにおいて前記複数の三角形のそれぞれを整列させる順整列プロセスを実行して、順整列行列を生成することに関わってもよい。いくつかのそのような方法は、第1のシーケンスの逆方向である第2のシーケンスにおいて前記複数の三角形のそれぞれを整列して、逆方向整列行列を生成する逆方向整列プロセスを実行することに関わってもよい。いくつかのそのような方法は、少なくとも部分的には、順整列行列の値および逆方向整列行列の値に基づいて、各オーディオ装置の位置の最終的な推定値を生成することに関わってもよい。しかしながら、方法4700のいくつかの実装では、ブロック4710は、図40を参照して上述したもの以外の方法を適用することに関わってもよい。

30

【0434】

この例では、ブロック4715は、制御システムを介して、環境内の聴取者位置を示す聴取者位置データを決定することに関わる。聴取者位置データは、たとえば、オーディオ装置座標系を参照した位置データであってもよい。しかしながら、他の例では、座標系は、聴取者を参照して、またはテレビジョン、部屋の壁等の環境の一部を参照して配向されてもよい。

40

【0435】

いくつかの例では、ブロック4715は、聴取者に（たとえば、環境中の一つまたは複数のラウドスピーカーからのオーディオ・プロンプトを介して）一つまたは複数の発声を行うよう促し、DOAデータに従って聴取者位置を推定することを含んでもよい。DOAデータは、環境内の複数のマイクロフォンによって取得されたマイクロフォン・データに対応してもよい。マイクロフォン・データは、マイクロフォンによる前記一つまたは複数の

50

発声の検出に対応してもよい。少なくとも一部のマイクロフォンは、ラウドスピーカーと共位置 (co-located) であってもよい。いくつかの例によれば、ブロック4715は、三角測量プロセスを含んでいてもよい。たとえば、ブロック4715は、たとえば、図48Aを参照して後述するように、オーディオ装置を通過するDOAベクトル間の交点を見つけることによって、ユーザーの声を三角測量することを含んでいてもよい。いくつかの実装によれば、ブロック4715 (または方法4700の別の動作) は、聴取者位置が決定された後に、オーディオ装置座標系および聴取者座標系の原点を共位置にする (co-locating) ことを含んでいてもよい。オーディオ装置座標系と聴取者座標系の原点を共位置にすることは、オーディオ装置位置をオーディオ装置座標系から聴取者座標系に変換することを含んでいてもよい。

10

【0436】

この実装によれば、ブロック4720は、制御システムを介して、聴取者角度配向を示す聴取者角度配向データを決定することを含む。聴取者角度配向データは、たとえば、オーディオ装置座標系のような聴取者位置データを表すために使用される座標系を参照して作成されてもよい。いくつかのそのような例では、聴取者角度配向データは、オーディオ装置座標系の原点および／または軸を参照して作成されてもよい。

【0437】

しかしながら、いくつかの実装では、聴取者角度配向データは、聴取者位置によって定義された軸と、テレビジョン、オーディオ装置、壁などの環境中の別の点とを参照して作成されてもよい。いくつかのそのような実装では、聴取者位置は、聴取者座標系の原点を定義するために使用されてもよい。いくつかのそのような例では、聴取者角度配向データは、聴取者座標系の軸を参照して作成されてもよい。

20

【0438】

ブロック4720を実施するためのさまざまな方法が、本明細書に開示される。いくつかの例によれば、聴取者角度配向は、聴取者観察方向に対応しうる。いくつかのそのような例では、たとえば聴取者がテレビジョンのような特定の物体を見ていると仮定することによって、聴取者の観察方向は、聴取者位置データを参照して推定されうる。いくつかのそのような実装では、聴取者観察方向は、聴取者位置およびテレビジョン位置に従って決定されてもよい。代替的または追加的に、聴取者観察方向は、聴取者位置およびテレビのサウンドバーの位置に従って決定されてもよい。

30

【0439】

しかしながら、いくつかの例では、聴取者の観察方向は、聴取者入力に従って決定されてもよい。いくつかのそのような例によれば、聴取者入力、聴取者が保持する装置から受領された慣性センサー・データを含んでいてもよい。聴取者は、環境中の位置、たとえば、聴取者が向いている方向に対応する位置を指すために該装置を使用してもよい。たとえば、聴取者は、該装置を使用して、鳴っているスピーカー (音を再生しているスピーカー) を指し示すことができる。よって、そのような例では、慣性センサー・データは、鳴っているスピーカーに対応する慣性センサー・データを含んでいてもよい。

【0440】

いくつかのそのような事例では、聴取者入力は、聴取者によって選択されたオーディオ装置の指示を含んでいてもよい。オーディオ装置の指示は、いくつかの例において、選択されたオーディオ装置に対応する慣性センサー・データを含んでいてもよい。

40

【0441】

しかしながら、他の例では、オーディオ装置の指示は、聴取者の一つまたは複数の発声 (たとえば、「テレビは今私の正面にあります」、「スピーカー2が今私の正面にあります」など) に従って行うことができる。聴取者の一つまたは複数の発声に応じて聴取者角度配向データを決定する他の例は、以下に記載される。

【0442】

図47に示される例によれば、ブロック4725は、制御システムを介して、聴取者位置および聴取者角度配向に対する、各オーディオ装置についてのオーディオ装置角度配向を示

50

す、オーディオ装置角度配向データを決定することを含む。いくつかのそのような例によれば、ブロック4725は、聴取者位置によって定義された点のまわりのオーディオ装置座標の回転を含んでいてもよい。いくつかの実装では、ブロック4725は、オーディオ装置位置データをオーディオ装置座標系から聴取者座標系に変換することを含んでいてもよい。いくつかの例が後述される。

【0443】

図48Aは、図47のいくつかのブロックの例を示している。いくつかのそのような例によれば、オーディオ装置位置データは、オーディオ装置座標系4807を参照しての、オーディオ装置1~5のそれぞれについてのオーディオ装置位置の推定値を含む。この実装では、オーディオ装置座標系4807は、オーディオ装置2のマイクロフォンの位置を原点とするデカルト座標系である。ここで、オーディオ装置座標系4807のx軸は、オーディオ装置2のマイクロフォンの位置とオーディオ装置1のマイクロフォンの位置との間の線4803に対応する。

【0444】

この例では、聴取者位置は、カウチ1903に座っていることが示されている聴取者4805に（たとえば、環境4800a内の一つまたは複数のラウドスピーカーからのオーディオ・プロンプトを介して）、一つまたは複数の発声4827を行うように促し、到着時間（TOA）データに従って聴取者位置を推定することによって決定される。TOAデータは、環境内の複数のマイクロフォンによって取得されたマイクロフォン・データに対応する。この例では、マイクロフォン・データは、オーディオ装置1~5の少なくとも一部（たとえば、3つ、4つまたは5つすべて）のマイクロフォンによる前記一つまたは複数の発声4827の検出に対応する。

【0445】

代替的または追加的に、オーディオ装置1~5の少なくとも一部（たとえば、2つ、3つ、4つまたは5つすべて）のマイクロフォンによって提供されるDOAデータに従った聴取者位置。いくつかのそのような例によれば、聴取者位置は、DOAデータに対応する直線4809a、4809b等の交点に従って決定されてもよい。

【0446】

この例によれば、聴取者位置は、聴取者座標系4820の原点に対応する。この例では、聴取者角度配向データは、聴取者の頭部4810（および／または聴取者の鼻4825）とテレビジョン3601のサウンドバー4830との間の直線4813aに対応する聴取者座標系4820のy'軸によって示される。図48Aに示される例では、線4813aはy'軸に平行である。したがって、角度 θ は、y軸とy'軸との間の角度を表す。この例では、図21のブロック2925は、聴取者座標系4820の原点のまわりの、オーディオ装置座標の、角度 θ による回転を含んでいてもよい。よって、オーディオ装置座標系4807の原点は、図48Aではオーディオ装置2と対応するように示されているが、いくつかの実装は、聴取者座標系4820の原点のまわりのオーディオ装置座標の角度 θ の回転の前に、オーディオ装置座標系4807の原点を、聴取者座標系4820の原点と共位置にすることを含む。この共位置化は、オーディオ装置座標系4807から聴取者座標系4820への座標変換によって実行されうる。

【0447】

サウンドバー4830および／またはテレビジョン3601の位置は、いくつかの例では、サウンドバーに音を発させ、DOAおよび／またはTOAデータに従ってサウンドバーの位置を推定することによって決定されてもよく、該DOAおよび／またはTOAデータは、オーディオ装置1~5の少なくとも一部（たとえば、3つ、4つまたは5つすべて）のマイクロフォンによる前記音の検出に対応する。代替的または追加的に、サウンドバー4830および／またはテレビジョン1901の位置は、ユーザーにテレビのところまで歩いて行くよう促し、DOAおよび／またはTOAデータによってユーザーの発話を位置特定することによって決定されてもよく、該DOAおよび／またはTOAデータは、オーディオ装置1~5の少なくとも一部（たとえば、3つ、4つまたは5つすべて）のマイクロフォンによる前記音の検出に対応する。そのような方法は、三角測量に関わってもよい。そのような例は、サウンド

バー4830および／またはテレビジョン1901が関連するマイクロフォンを持たない状況において有益でありうる。

【0448】

サウンドバー4830および／またはテレビジョン3601が関連するマイクロフォンを有する他のいくつかの例では、サウンドバー4830および／またはテレビジョン3601の位置は、本明細書に開示されたDOA法などのTOAまたはDOA法に従って決定することができる。いくつかのそのような方法によれば、マイクロフォンは、サウンドバー4830と共位置にされてもよい。

【0449】

いくつかの実装によれば、サウンドバー4830および／またはテレビジョン3601は、関連するカメラ4811を有してもよい。制御システムは、聴取者の頭部4810（および／または聴取者の鼻4825）の画像を捕捉するように構成されてもよい。いくつかのそのような例では、制御システムは、聴取者の頭部4810（および／または聴取者の鼻4825）とカメラ4811との間の直線4813aを決定するように構成されてもよい。聴取者角度配向データは、直線4813aに対応してもよい。代替的または追加的に、制御システムは、直線4813aとオーディオ装置座標系のy軸との間の角度 θ を決定するように構成されてもよい。

【0450】

図48Bは、聴取者角度配向データを決定するさらなる例を示す。この例によれば、聴取者位置は、図47のブロック2115においてすでに決定されている。ここで、制御システムが、環境4800bのスピーカーを制御して、オーディオ・オブジェクト4835を環境4800b内の多様な位置にレンダリングしている。いくつかのそのような例では、制御システムは、スピーカーに、オーディオ・オブジェクト4835が聴取者4805のまわりを回転するように感じられるような仕方でオーディオ・オブジェクト4835をレンダリングさせることができる。かかるレンダリングはたとえば、オーディオ・オブジェクト4835が聴取者座標系4820の原点のまわりに回転するように感じられるようにオーディオ・オブジェクト4835をレンダリングすることによる。この例では、曲がった矢印4840は、オーディオ・オブジェクト4835が聴取者4805のまわりを回転するときの、オーディオ・オブジェクト4835の軌跡の一部を示す。

【0451】

いくつかのそのような例によれば、聴取者4805は、いつオーディオ・オブジェクト4835が聴取者4805が向いている方向にあるかを示すユーザー入力を提供することができる（たとえば、「止まれ」と言って）。いくつかのそのような例では、制御システムは、聴取者位置とオーディオ・オブジェクト4835の位置との間の直線4813bを決定するように構成されてもよい。この例では、直線4813bは、聴取者4805が向いている方向を示す、聴取者座標系のy'軸に対応する。代替的な実装では、聴取者4805は、いつオーディオ・オブジェクト4835が環境の前方、環境のテレビ位置、オーディオ装置位置などにあるかを示すユーザー入力を提供してもよい。

【0452】

図48Cは、聴取者角度配向データを決定するさらなる例を示す。この例によれば、聴取者位置は、図47のブロック2115においてすでに決定されている。ここで、聴取者4805は、ハンドヘルド装置4845をテレビ3601またはサウンドバー4830のほうに向けることによって、聴取者4805の観察方向に関する入力を提供するためにハンドヘルド装置4845を使用している。ハンドヘルド装置4845および聴取者の腕の破線の輪郭は、聴取者4805がハンドヘルド装置4845をテレビ3601またはサウンドバー4830のほうへ向けていた時点より前の時点で、聴取者4805がハンドヘルド装置4845を、この例ではオーディオ装置2のほうへ向けていたことを示す。他の例では、聴取者4805は、ハンドヘルド装置4845をオーディオ装置1のような別のオーディオ装置に向けていてもよい。この例によれば、ハンドヘルド装置4845は、オーディオ装置2とテレビ3601またはサウンドバー4830との間の角度 α を決定するように構成される。該角度 α は、オーディオ装置2と聴取者4805の観察方向との間の角度を近似する。

10

20

30

40

50

ハンドヘルド装置4845は、いくつかの例では、慣性センサー・システムと、環境4800cのオーディオ装置を制御している制御システムと通信するように構成された無線インターフェースとを含むセルラー電話であってもよい。いくつかの例では、ハンドヘルド装置4845は、必要な機能を実行するようにハンドヘルド装置4845を制御するように構成されたアプリケーションまたは「アプリ」を実行していてもよい。機能の実行は、たとえば、ユーザー・プロンプトを（たとえば、グラフィカル・ユーザー・インターフェースを介して）提供することによって、ハンドヘルド装置3045が所望の方向を指し示していることを示す入力を受領することによって、対応する慣性センサー・データを保存することおよび／または対応する慣性センサー・データを環境3000cのオーディオ装置を制御している制御システムに送信することなどによる。

【0453】

この例によれば、制御システム（これは、ハンドヘルド装置3045の制御システム、または環境3000cのオーディオ装置を制御する制御システムであってもよい）は、慣性センサー・データ、たとえばジャイロスコプデータに従って、直線3013cおよび3050の配向を決定するように構成される。この例では、直線3013cは軸 y' に平行であり、聴取者角度配向を決定するために使用されうる。いくつかの例によれば、制御システムは、オーディオ装置2と聴取者3005の観察方向との間の角度 α に応じて、聴取者座標系3020の原点のまわりのオーディオ装置座標についての適切な回転を決定することができる。

【0454】

図48Dは、図48Cを参照して説明した方法に従った、オーディオ装置座標についての適切な回転を決定する例を示す。この例では、オーディオ装置座標系4807の原点は、聴取者座標系4820の原点と共位置にされる。オーディオ装置座標系4807と聴取者座標系4820の原点を共位置にすることは、聴取者位置が決定される2115のプロセス後に可能となる。オーディオ装置座標系4807と聴取者座標系4820の原点を共位置にすることは、オーディオ装置位置をオーディオ装置座標系4807から聴取者座標系4820に変換することに関してもよい。角度 α は、図48Cを参照して上述したように決定されている。よって、角度 α は、聴取者座標系4820におけるオーディオ装置2の所望の配向に対応する。この例では、角度 β は、オーディオ装置座標系4807におけるオーディオ装置2の配向に対応する。この例では $\beta - \alpha$ である角度 θ は、オーディオ装置座標系4807の y 軸を聴取者座標系4820の y' 軸に整列させるのに必要な回転を示す。

【0455】

いくつかの実装では、図47の方法は、少なくとも部分的には対応するオーディオ装置位置、対応するオーディオ装置角度配向、聴取者位置データおよび聴取者角度配向データに基づいて、環境中のオーディオ装置の少なくとも1つを制御することを含んでいてもよい。

【0456】

たとえば、いくつかの実装は、オーディオ装置位置データ、オーディオ装置角度配向データ、聴取者位置データ、聴取者角度配向データをオーディオ・レンダリング・システムに提供することを含むことができる。いくつかの例では、オーディオ・レンダリング・システムは、図6の制御システム610のような制御システムによって実装されてもよい。いくつかの実装は、オーディオ装置位置データ、オーディオ装置角度配向データ、聴取者位置データ、聴取者角度配向データに少なくとも部分的に基づいて、オーディオ・データ・レンダリング・プロセスを制御することを含んでいてもよい。いくつかのそのような実装は、レンダリング・システムにラウドスピーカー音響能力データを提供することを含んでいてもよい。ラウドスピーカー音響能力データは、環境の一つまたは複数のラウドスピーカーに対応しうる。ラウドスピーカー音響能力データは、一つまたは複数のドライバの配向、ドライバの数、または一つまたは複数のドライバのドライバ周波数応答を示してもよい。いくつかの例では、ラウドスピーカー音響能力データはメモリから取り出され、次いでレンダリング・システムに提供されてもよい。

【0457】

あるクラスの実施形態は、複数の協調させられた（オーケストレーションされた）スマ

10

20

30

40

50

ート・オーディオ装置のうちの少なくとも1つ（たとえば、全部または一部）によって、再生のためにオーディオのレンダリングおよび／または該オーディオの再生の方法を含む。たとえば、ユーザーの家庭において（システム内に）存在するスマート・オーディオ装置の集合は、スマート・オーディオ装置の全部または一部（すなわち、全部または一部のスピーカー（単数または複数））による再生のためのオーディオの柔軟なレンダリングを含む、多様な同時使用事例を扱うために、オーケストレーションされうる。レンダリングおよび／または再生に対する動的修正を必要とする、上記システムとの多くの対話が考えられる。そのような修正は、空間的忠実度に焦点を当ててもよいが、必ずしもそうではない。

【0458】

いくつかの実施形態は、再生のためのレンダリングおよび／または協調させられる（オーケストレーションされる）複数のスマート・オーディオ装置のスピーカーのよる再生を実装する。

【0459】

いくつかの実施形態（たとえば、レンダリング・システムまたはレンダラーまたはレンダリング方法または再生システムもしくはは方法）は、再生のためにオーディオをレンダリングする、および／または一組のスピーカーの一部または全部のスピーカー（たとえば、各アクティブ化されたスピーカー）による再生のためのシステムおよび方法に関する。いくつかの実施形態では、スピーカーは、スマート・オーディオ装置であってもよい協調させられる（オーケストレーションされる）一組のオーディオ装置のスピーカーである。

【0460】

スマート・オーディオ装置の集合のスマート・オーディオ装置による（または別の一組のスピーカーによる）再生のための空間的オーディオ・ミックスのレンダリング（またはレンダリングおよび再生）（たとえば、オーディオのストリームまたはオーディオの複数のストリームのレンダリング）を実行するというコンテキストでは、（たとえば、スマート・オーディオ装置内の、またはスマート・オーディオ装置に結合されている）スピーカーのタイプが変わる可能性があり、よってスピーカーの対応する音響能力は非常に大きく変化する可能性がある。たとえば、図29に示されるオーディオ環境2000のある実装では、ラウドスピーカー2005d、2005fおよび2005hは、単一の0.6インチスピーカーを有するスマートスピーカーであってもよい。この例では、ラウドスピーカー2005b、2005c、2005e、および2005fは、2.5インチのウーファおよび0.8インチのツイーターを有するスマートスピーカーであってもよい。この例によれば、ラウドスピーカー2005gは、5.25インチ・ウーファー、3つの2インチ中距離スピーカー、および1.0インチ・ツイーターを備えたスマートスピーカーであってもよい。ここで、ラウドスピーカー2005aは、16個の1.1インチ・ビームドライバと2個の4インチ・ウーファーを有するサウンドバーであってもよい。よって、スマートスピーカー2005dおよび2005fの低周波能力は、環境2000内の他のラウドスピーカー、特に4インチまたは5.25インチのウーファーを有するラウドスピーカーよりもかなり低いであろう。

【0461】

図49は、本開示のさまざまな側面を実装できるシステムの構成要素の例を示すブロック図である。本明細書で提供される他の図と同様に、図49に示される要素のタイプおよび数は、単に例として提供されている。他の実装は、より多くの要素、より少ない要素、および／または異なるタイプおよび数の要素を含みうる。

【0462】

この例によれば、システム4900は、スマートホームハブ4905と、ラウドスピーカー4925a～4925mとを含む。この例では、スマートホームハブ4905は、図6に示され、上述された制御システム610のインスタンスを含む。いくつかの例では、システム4900の機能は、少なくとも部分的には、オーディオ・セッション・マネージャ、たとえば、図2CのCHASM 208C、図2DのCHASM 208D、図3CのCHASM 307および／または図4のCHASM 401からの命令に従って、提供されてもよい。オーディオ・セッション・マネージ

10

20

30

40

50

ャは、いくつかの事例では、スマートホームハブ4905以外の装置によって実装されてもよい。しかしながら、いくつかの例においては、オーディオ・セッション・マネージャは、スマートホームハブ4905によって実装されてもよい。この実装によれば、制御システム610は、聴取環境ダイナミクス処理構成データ・モジュール4910、聴取環境ダイナミクス処理モジュール4915、およびレンダリング・モジュール4920を含む。聴取環境ダイナミクス処理構成データ・モジュール4910、聴取環境ダイナミクス処理モジュール4915、およびレンダリング・モジュール4920のいくつかの例を以下に説明する。いくつかの例において、レンダリング・モジュール4920は、レンダリングおよび聴取環境ダイナミクス処理の両方のために構成されてもよい。

【0463】

スマートホームハブ4905とラウドスピーカー4925a~4925mとの間の矢印によって示唆されるように、スマートホームハブ4905は、図6に示され、上述されたインターフェース・システム605のインスタンスも含む。いくつかの例によれば、スマートホームハブ4905は、図2に示される環境200の一部であってもよい。いくつかの事例では、スマートホームハブ4905は、スマートスピーカー、スマートテレビ、セルラー電話、ラップトップなどによって実装されうる。いくつかの実装では、スマートホームハブ4905は、ソフトウェアによって、たとえば、ダウンロード可能なソフトウェア・アプリケーションまたは「アプリ」のソフトウェアを介して実装されてもよい。いくつかの事例では、スマートホームハブ4905は、ラウドスピーカー4925a-mのそれぞれにおいて実装されて、すべてが並列に動作して、モジュール4920からの同じ処理されたオーディオ信号を生成してもよい。いくつかのそのような例によれば、各ラウドスピーカーにおいて、レンダリング・モジュール4920は、次いで、各ラウドスピーカーまたはラウドスピーカーのグループに関連する一つまたは複数のスピーカー・フィードを生成してもよく、これらのスピーカー・フィードを各スピーカー・ダイナミクス処理モジュールに提供してもよい。

【0464】

いくつかの事例では、ラウドスピーカー4925a~4925mは、図29のラウドスピーカー305a~305hを含んでいてもよい。他の例では、ラウドスピーカー4925a~4925mは、他のラウドスピーカーであってもよく、または他のラウドスピーカーを含んでいてもよい。よって、この例では、システム4900は、M個のラウドスピーカーを含み、ここで、Mは、2より大きい整数である。

【0465】

スマートスピーカーは、多くの他のパワー付きスピーカーと同様に、典型的には、スピーカーが歪むことを防止するために、何らかのタイプの内部ダイナミクス処理を用いる。そのようなダイナミクス処理には、しばしば信号制限閾値（たとえば、周波数にわたって可変である制限閾値）が関連し、信号レベルは動的にそれより下に保持される。たとえば、ドルビーオーディオ処理（Dolby Audio Processing、DAP）オーディオ後処理スイートにおけるいくつかのアルゴリズムの1つであるドルビーのオーディオレギュレータは、そのような処理を提供する。いくつかの事例では、典型的にはではないが、スマートスピーカーのダイナミクス処理モジュールを介して、ダイナミクス処理は、一つまたは複数のコンプレッサ、ゲート、エキスパンダ、ダッカ（ducker）等を適用することにも関わってもよい。

【0466】

よって、この例では、ラウドスピーカー4925a~4925mのそれぞれは、対応するスピーカー・ダイナミクス処理（dynamics processing、DP）モジュールA~Mを含む。スピーカー・ダイナミクス処理モジュールは、聴取環境のそれぞれの個々のラウドスピーカーについて、個々のラウドスピーカー・ダイナミクス処理構成データ（loudspeaker dynamics processing configuration data）を適用するように構成される。スピーカー・DPモジュールAは、たとえば、ラウドスピーカー4925aに適した個々のラウドスピーカー・ダイナミクス処理構成データを適用するように構成される。いくつかの例では、個々のラウドスピーカー・ダイナミクス処理構成データは、個々のラウドスピーカーの一つま

10

20

30

40

50

たは複数の能力に対応してもよい。たとえば、特定の周波数範囲内で、特定のレベルのオーディオ・データを、認識可能な歪みなしに再生するラウドスピーカーの能力である。

【0467】

空間的オーディオが、それぞれ潜在的には異なる再生限界をもつ不均質なスピーカー（たとえば、スマート・オーディオ装置のスピーカー、またはスマート・オーディオ装置に結合されたスピーカー）の集合にまたがってレンダリングされる場合、全体的なミックスに対してダイナミクス処理を実行する際には注意が必要である。簡単な解決策は、空間的ミックスを各参加スピーカーのスピーカー・フィールドにレンダリングし、次いで、各スピーカーに関連するダイナミクス処理モジュールが、そのスピーカーの限界に従って、その対応するスピーカー・フィールドに対して、独立して作用することを許容することである。

【0468】

このアプローチは、各スピーカーを歪まないようにするが、ミックスの空間バランスを知覚的にわずらわしい仕方で動的にシフトさせることがある。たとえば、図29を参照して、テレビ番組がテレビ2030に示されており、対応するオーディオがオーディオ環境2000のラウドスピーカーによって再生されているとする。テレビ番組の間、静止物体（工場の重機ユニットなど）に関連するオーディオは、オーディオ環境2000内の特定の位置にレンダリングされることが意図されているとする。さらに、ラウドスピーカー4925bのほうがベース範囲の音を再生する能力が実質的に大きいため、ラウドスピーカー4925dに関連付けられたダイナミクス処理モジュールが、ベース範囲のオーディオのレベルを、ラウドスピーカー4925bに関連付けられたダイナミクス処理モジュールよりも実質的に大きく低下させるとする。静止物体に関連する信号の音量が変動する場合、音量が大きくなると、ラウドスピーカー4925dに関連するダイナミクス処理モジュールは、ベース範囲のオーディオのレベルを、ラウドスピーカー4925bに関連するダイナミクス処理モジュールによって同じオーディオのレベルが低下させられるよりも、実質的に大きく低下させる。このレベル差は、静止物体の見かけ上の位置を変化させる。よって、改善された解決策が望ましいであろう。

【0469】

本開示のいくつかの実施形態は、スマート・オーディオ装置の集合（たとえば協調させられるスマート・オーディオ装置の集合）のスマート・オーディオ装置のうちの少なくとも1つ（たとえば、全部または一部）および／またはスピーカーの別の集合のスピーカーのうちの少なくとも1つ（たとえば、全部または一部）による再生のために、空間的オーディオ・ミックスのレンダリング（またはレンダリングおよび再生）（たとえば、オーディオのストリームまたはオーディオの複数ストリームのレンダリング）のためのシステムおよび方法である。いくつかの実施形態は、そのようなレンダリング（たとえば、スピーカー・フィールドの生成を含む）およびレンダリングされたオーディオの再生（たとえば、生成されたスピーカー・フィールドの再生）のための方法（またはシステム）である。そのような実施形態の例は以下のとおり。

【0470】

オーディオ処理のためのシステムおよび方法は、少なくとも2つのスピーカー（たとえば、スピーカーの集合のスピーカーのうちの全部または一部）による再生のために、オーディオをレンダリングする（たとえば、オーディオのストリームまたはオーディオの複数のストリームをレンダリングすることによって、空間的オーディオ・ミックスをレンダリングする）ことを含んでいてもよく、下記によることを含む：

(a) 個別のラウドスピーカー・ダイナミクス処理構成データ（たとえば、個々のラウドスピーカーの制限閾値（再生制限閾値））を組み合わせ、それにより複数のラウドスピーカーのための聴取環境ダイナミクス処理構成データ（組み合わせられた閾値など）を決定する；

(b) 複数のラウドスピーカーのための聴取環境ダイナミクス処理構成データ（たとえば、組み合わせられた閾値）を使用して、オーディオ（たとえば、空間的オーディオ・ミックスを示すオーディオのストリーム）に対するダイナミクス処理を実行して、処理されたオ

10

20

30

40

50

オーディオを生成する；

(c) 処理されたオーディオをスピーカー・フィールドにレンダリングする。

【0471】

いくつかの実装によれば、プロセス(a)は、図49に示される聴取環境ダイナミクス処理構成データ・モジュール4910のようなモジュールによって実行されてもよい。スマートホームハブ4905は、インターフェース・システムを介して、M個のラウドスピーカーのそれぞれについて個々のラウドスピーカー・ダイナミクス処理構成データを得るように構成されてもよい。この実装では、個々のラウドスピーカー・ダイナミクス処理構成データは、複数のラウドスピーカーの各ラウドスピーカーについての個々のラウドスピーカー・ダイナミクス処理構成データセットを含む。いくつかの例によれば、一つまたは複数のラウドスピーカーのための個々のラウドスピーカー・ダイナミクス処理構成データは、前記一つまたは複数のラウドスピーカーの一つまたは複数の能力に対応する。この例では、個々のラウドスピーカー・ダイナミクス処理構成データセットのそれぞれは、少なくとも1つのタイプのダイナミクス処理構成データを含む。いくつかの例では、スマートホームハブ4905は、各ラウドスピーカー4925a~4925mに問い合わせることによって、個々のラウドスピーカー・ダイナミクス処理構成データセットを得るように構成されてもよい。他の実装では、スマートホームハブ4905は、メモリに記憶されている、以前に取得された個々のラウドスピーカー・ダイナミクス処理構成データセットのデータ構造に問い合わせることによって、個々のラウドスピーカー・ダイナミクス処理構成データセットを取得するように構成されてもよい。

【0472】

いくつかの例では、プロセス(b)は、図49の聴取環境ダイナミクス処理モジュール4915のようなモジュールによって実行されてもよい。プロセス(a)および(b)のいくつかの詳細な例を以下に説明する。

【0473】

いくつかの例では、プロセス(c)のレンダリングは、図49のレンダリング・モジュール4920またはレンダリング・モジュール4920'のようなモジュールによって実行されてもよい。いくつかの実施形態では、オーディオ処理は、下記に関わる：

(d) 各ラウドスピーカーについての個別のラウドスピーカー・ダイナミクス処理設定データに従って、レンダリングされたオーディオ信号に対してダイナミクス処理を実行する(たとえば、対応するスピーカーに関連付けられた再生制限閾値に従ってスピーカー・フィールドを制限し、それにより、制限されたスピーカー・フィールドを生成する)。プロセス(d)はたとえば、図49に示されるダイナミクス処理モジュールA~Mによって実行されてもよい。

【0474】

スピーカーは、スマート・オーディオ装置の集合のスマート・オーディオ装置のうちの少なくとも1つ(たとえば、全部または一部)の(またはそれに結合された)スピーカーであってもよい。いくつかの実装では、ステップ(d)において制限されたスピーカー・フィールドを生成するために、ステップ(c)において生成されたスピーカー・フィールドは、ダイナミクス処理の第2段によって(たとえば、各スピーカーの関連するダイナミクス処理システムによって)処理されて、たとえば、スピーカー・フィールドを、スピーカーを通じた最終的な再生の前に生成してもよい。たとえば、スピーカー・フィールド(またはそのサブセットもしくは一部)は、スピーカーのそれぞれの異なるもののダイナミクス処理システム(たとえば、スマート・オーディオ装置のダイナミクス処理サブシステム。ここで、スマート・オーディオ装置は、それらのスピーカーのうちの関連するものを含む、またはそれに結合されている)。前記各ダイナミクス処理システムから出力される処理されたオーディオは、スピーカーのうちの関連するものためのスピーカー・フィールドを生成するために使用されてもよい。スピーカー固有のダイナミクス処理(すなわち、各スピーカーについて独立に実行されるダイナミクス処理)に続いて、処理された(たとえば、動的に制限された)スピーカー・フィールドは、スピーカーを駆動して音声の再生を引き起こす

ために使用されうる。

【0475】

ダイナミクス処理の第1段（ステップ（b））は、ステップ（a）および（b）が省略されステップ（d）から生じるダイナミクス処理された（たとえば、制限された）スピーカー・フィールドがもとのオーディオにตอบสนองして（ステップ（b）で生成された処理されたオーディオにตอบสนองしてではなく）生成された場合に生じるであろう知覚的にわずらわしい空間バランスのシフトを低減するように設計されうる。これは、ミックスの空間バランスにおける望ましくないシフトを防止しうる。ステップ（c）からのレンダリングされたスピーカー・フィールドに対して作用するダイナミクス処理の第2段は、どのスピーカーも歪まないことを保証するように設計されてもよい。ステップ（b）のダイナミクス処理は、必ずしも信号レベルがすべてのスピーカーの閾値未満に低下したことを保証しないことがありうるためである。個別のラウドスピーカー・ダイナミクス処理構成データを組み合わせること（たとえば、第1段（ステップ（a））における閾値の組み合わせ）は、いくつかの例では、諸スピーカーにわたって（たとえば、スマート・オーディオ装置にわたって）個別のラウドスピーカー・ダイナミクス処理構成データ（たとえば、制限閾値）を平均する、または諸スピーカーにわたって（たとえば、スマート・オーディオ装置にわたって）個別のラウドスピーカー・ダイナミクス処理構成データ（たとえば、制限閾値）の最小を取るステップに関わる（たとえば、含む）。

10

【0476】

いくつかの実装では、ダイナミクス処理の第1段（ステップ（b））が、空間的ミックスを示すオーディオ（たとえば、少なくとも1つのオブジェクトチャンネルおよび任意的には少なくとも1つのスピーカーチャンネルをも含む、オブジェクトベースのオーディオ・プログラムのオーディオ）に対して作用する場合、この第1段は、諸空間ゾーンの使用を通じたオーディオ・オブジェクト処理のための技法に従って実装されうる。そのような場合、各ゾーンに関連する組み合わせられた個別のラウドスピーカー・ダイナミクス処理構成データ（たとえば、組み合わせられた制限閾値）は、個々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、個々のスピーカー制限閾値）の重み付けされた平均によって（または重み付けされた平均として）導出されてもよく、この重み付けは、少なくとも部分的には、各スピーカーの前記ゾーンへの空間的近接性および／または前記ゾーン内の位置によって与えられてもよく、または決定されてもよい。

20

30

【0477】

ある例示的实施形態では、複数 M 個のスピーカー（ $M \geq 2$ ）を想定し、ここで各スピーカーは変数 i によってインデックス付けされる。各スピーカー i には、周波数変化する再生制限閾値 $T_i[f]$ が関連付けられている。ここで、変数 f は、閾値が指定される周波数の有限集合へのインデックスを表す。（周波数の集合のサイズが1である場合、対応する単一の閾値はブロードバンドであると見なされ、周波数範囲全体にわたって適用されることに注意。）これらの閾値は、各スピーカーによって、特定の目的のためにオーディオ信号を閾値を下回るよう制限するために、それ自身の独立したダイナミクス処理機能において利用される。特定の目的は、スピーカーが歪むのを防ぐ、またはスピーカーがその近傍で好ましくないとみなされる何らかのレベルを超えて再生することを防止するなどである。

40

【0478】

図50A、50Bおよび50Cは、再生制限閾値および対応する周波数の例を示す。示される周波数の範囲は、たとえば、平均的な人間にとって聴取可能な周波数の範囲（たとえば、20Hz～20kHz）にわたることができる。これらの例では、再生制限閾値は、グラフ5000a、5000b、および5000cの垂直軸によって示され、垂直軸はこれらの例では「レベル閾値」とラベル付けされている。再生制限／レベルの閾値は、垂直軸上の矢印の方向に増加する。再生制限／レベル閾値は、たとえば、デシベルで表すことができる。これらの例では、グラフ5000a、5000b、および5000cの水平軸は周波数を示し、周波数は水平軸上の矢印の方向に増加する。曲線5000a、5000b、および5000cによって示される再生制限閾値は、たとえば、個々のラウドスピーカーのダイナミクス処理モジュールによって実

50

装されうる。

【0479】

図50Aのグラフ5000aは、周波数の関数として再生制限閾値の第1の例を示す。曲線5005aは、対応する各周波数値についての再生制限閾値を示す。この例では、ベース周波数 f_b において、入力レベル T_i で受領される入力オーディオは、出力レベル T_o でダイナミクス処理モジュールによって出力される。ベース周波数 f_b は、たとえば、60~250Hzの範囲であってもよい。しかしながら、この例では、高音周波数 f_t において、入力レベル T_i で受領される入力オーディオは、同じレベルの入力レベル T_i で、ダイナミクス処理モジュールによって出力される。高音周波数 f_t は、たとえば、1280Hzより上の範囲内であってもよい。よって、この例では、曲線5005aは、ベース周波数について、高音周波数よりも有意に低い閾値を適用するダイナミクス処理モジュールに対応する。そのようなダイナミクス処理モジュールは、ウーファをもたないラウドスピーカー（たとえば、図29のラウドスピーカー2005d）に適してもよい。

【0480】

図50Bのグラフ5000bは、周波数の関数として再生制限閾値の第2の例を示す。曲線5005bは、図50Aに示される同じベース周波数 f_b において、入力レベル T_i で受信される入力オーディオが、より高い出力レベル T_o でダイナミクス処理モジュールによって出力されることを示す。よって、この例では、曲線5005bは、曲線5005aほど低いベース周波数についての閾値を適用しないダイナミクス処理モジュールに対応する。そのようなダイナミクス処理モジュールは、少なくとも小さなウーファ（たとえば、図29のスピーカー2005b）を有するスピーカーに適している。

【0481】

図50Cのグラフ5000cは、周波数の関数として再生制限閾値の第2の例を示す。曲線5005c（この例では直線である）は、図50Aに示される同じベース周波数 f_b において、入力レベル T_i で受領される入力オーディオが、同じレベルでダイナミクス処理モジュールによって出力されることを示す。よって、この例では、曲線5005cは、ベース周波数を含む広範囲の周波数を再生することができるラウドスピーカーに適切でありうるダイナミクス処理モジュールに対応する。簡単のため、ダイナミクス処理モジュールは、示されるすべての周波数について同じ閾値を適用する曲線5005dを実装することによって、曲線5005cを近似することができるが見て取れるであろう。

【0482】

空間的オーディオ・ミックスは、重心振幅パン（Center of Mass Amplitude Panning、CMAP）、柔軟仮想化（Flexible Virtualization、FV）、または本明細書に開示されるCMAPおよびFVの組み合わせなどのレンダリング・システムを使用して、複数のスピーカーのためにレンダリングされうる。空間的オーディオ・ミックスの構成要素から、レンダリング・システムは、複数のスピーカーのそれぞれのために1つのスピーカー・フィールドを生成する。いくつかの以前の例では、スピーカー・フィールドは、その後、各スピーカーの関連付けられたダイナミクス処理機能によって、閾値 $T_i[f]$ を用いて独立して処理された。本開示の恩恵がなければ、この記述されたレンダリング・シナリオは、レンダリングされた空間的オーディオ・ミックスの知覚される空間的バランスにおけるわずらわしいシフトを生じさせることがある。たとえば、聴取領域の右側などにある、M個のスピーカーのうちの1つが、他のスピーカーよりもはるかに能力が低く（たとえば、ベース範囲のオーディオをレンダリングする能力）、よって、そのスピーカーのための閾値は、少なくとも特定の周波数範囲では、他のスピーカーの閾値よりも有意に低くてもよい。再生中、このスピーカーのダイナミクス処理モジュールは、右側の空間的ミックスの成分のレベルを、左側の成分よりも、著しく低下させるであろう。聴取者は、空間的ミックスの左右バランスの間のそのような動的なシフトに非常に敏感であり、結果が非常にわずらわしいと感ずることがありうる。

【0483】

この問題に対処するために、いくつかの例において、聴取環境の個々のスピーカーの個

々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、再生制限閾値）が組み合わせられて、聴取環境のすべてのラウドスピーカーについての聴取環境ダイナミクス処理構成データを作成する。次いで、聴取環境ダイナミクス処理構成データを利用して、まず、スピーカー・フィールドにレンダリングする前に、空間的オーディオ・ミックス全体のコンテキストにおいてダイナミクス処理を実行することができる。ダイナミクス処理のこの第1段は、ただ1つの独立したスピーカー・フィールドではなく、空間的ミックス全体へのアクセスを有するので、処理は、ミックスの知覚される空間バランスに対してわずらわしいシフトを付与しない仕方で実行されうる。個々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、再生制限閾値）は、個々のスピーカーの独立したダイナミクス処理機能のいずれかによって実行されるダイナミクス処理の量をなくすまたは低減する仕方で、組み合わせられてもよい。

10

【0484】

聴取環境ダイナミクス処理構成データを決定する一例では、個々のスピーカーについての個々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、再生制限閾値）は、ダイナミクス処理の第1段における空間的ミックスのすべての成分に適用される、聴取環境ダイナミクス処理構成データ（たとえば、周波数変化する再生制限閾値

【数66】

$$\bar{T}[f]$$

20

）の単一の集合に組み合わせられてもよい。いくつかのそのような例によれば、制限はすべての成分で同じであるため、ミックスの空間バランスが維持されうる。個々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、再生制限閾値）を組み合わせる1つの方法は、すべてのスピーカー*i*にわたる最小を取ることである：

【数67】

$$[f] = \min_i(T_i[f])$$

Equation (20)

【0485】

そのような組み合わせは、各スピーカーの個々のダイナミクス処理の動作を本質的になくす。なぜなら、空間的ミックスは、最初に、すべての周波数において、最も能力が低いスピーカーの閾値を下回るように制限されるからである。しかしながら、そのような戦略は過度に積極的である可能性がある。多くのスピーカーは、それらに対応できるよりも低いレベルで再生し、すべてのスピーカーの組み合わせられた再生レベルは、好ましくないほど低い場合がある。たとえば、図50Aに示されたベース範囲における閾値が、図50Cについての閾値に対応するラウドスピーカーに適用されたとしたら、後者のスピーカーの再生レベルは、ベース範囲において不必要なまでに低くなる。聴取環境ダイナミクス処理構成データを決定する代替的な組み合わせは、聴取環境のすべてのスピーカーにわたって個々のラウドスピーカー・ダイナミクス処理構成データの平均（アベレージ）を取ることである。たとえば、再生制限閾値のコンテキストにおいて、平均は次のように決定されうる：

30

【数68】

$$\bar{T}[f] = \text{mean}_i(T_i[f])$$

Equation (21)

40

【0486】

この組み合わせでは、ダイナミクス処理の第1段がより高いレベルに制限されるため、最小を取ることに比べ、全体的な再生レベルが増大することがあり、それにより、より能力の高いスピーカーがより大音量で再生できるようになる。個々の制限閾値が平均値を下回るスピーカーについては、その独立したダイナミクス処理機能が、必要であれば、関連するスピーカーのフィールドを制限することができる。しかしながら、ダイナミクス処理の

50

第1段は、いくつかの初期制限が空間的ミックスに対して実行されているので、この制限の要件を減らしている可能性がある。

【0487】

聴取環境ダイナミクス処理構成データを決定するいくつかの例によれば、チューニング・パラメータを通じて個々のラウドスピーカー・ダイナミクス処理構成データの最小と平均との間を補間する調整可能な組み合わせを生成することができる。たとえば、再生制限閾値のコンテキストにおいて、補間は次のように決定されうる：

【数69】

$$\bar{T}[f] = \alpha \text{mean}_i(T_i[f]) + (1 - \alpha) \min_i(T_i[f]) \quad \text{Equation (22)}$$

10

【0488】

個々のラウドスピーカー・ダイナミクス処理構成データの他の組み合わせが可能であり、本開示はそのような組み合わせすべてをカバーすることが意図されている。

【0489】

図51Aおよび51Bは、ダイナミックレンジ圧縮データの例を示すグラフである。グラフ5100aおよび5100bでは、デシベルでの入力信号レベルが水平軸に、デシベルでの出力信号レベルが垂直軸に示されている。他の開示された例と同様に、特定の閾値、比、および他の値は、単に例として示されており、限定するものではない。

【0490】

図51Aに示される例では、出力信号レベルは閾値未満の入力信号レベルに等しく、この例では-10dBである。他の例は、異なる閾値、たとえば、-20dB、-18dB、-16dB、-14dB、-12dB、-8dB、-6dB、-4dB、-2dB、0dB、2dB、4dB、6dBなどに関わってもよい。閾値より上では、圧縮比のさまざまな例が示されている。N:1の比は、閾値より上では、出力信号レベルが入力信号のNdB増加毎に1dB増加することを意味する。たとえば、10:1の圧縮比（線5105e）は、閾値より上では、出力信号レベルは、入力信号が10dB増加するごとに1dBだけ増加することを意味する。1:1の圧縮比（線5105a）は、閾値より上であっても、出力信号レベルがいまだ入力信号レベルと同じであることを意味する。線5105b、5105c、および5105dは、3:2、2:1、および5:1の圧縮比に対応する。他の実装は、2.5:1、3:1、3.5:1、4:3、4:1などのように、異なる圧縮比を提供することができる。

20

【0491】

図51Bは、「ニー」（knee）の例を示しており、これは、この例では0dBである閾値またはその付近で圧縮比がどのように変化するかを制御する。この例によれば、「硬い」ニーを有する圧縮曲線は、2つの直線セグメント、すなわち閾値までの直線セグメント5110aおよび閾値より上の直線セグメント5110bから構成される。硬いニーのほうが実装するのは簡単だが、アーチファクトを引き起こすことがある。

30

【0492】

図51Bでは、「柔らかい」ニーの一例も示されている。この例では、柔らかいニーは10dBにまたがる。この実装によれば、10dBのスパンの上下で、柔らかいニーを有する圧縮曲線の圧縮比は、硬いニーを有する圧縮曲線の圧縮比と同じである。他の実装は、「柔らかい」ニーのさまざまな他の形状を提供することができ、それらはより多いまたはより少ないデシベルにわたることもあり、スパンの上で異なる圧縮比を示すことなどもある。

40

【0493】

他のタイプのダイナミックレンジ圧縮データは、「アタック」データおよび「リリース」データを含むことができる。アタックは、圧縮比によって決定される利得に達するために、コンプレッサ〔圧縮器〕が、たとえば入力における増大したレベルに応答して利得を減少させる期間である。コンプレッサについてのアタック時間は、一般に、25ミリ秒から500ミリ秒の範囲であるが、他のアタック時間も実用可能である。リリースは、コンプレッサが、たとえば低下した入力レベルに応答して、圧縮比によって決定される出力利得（

50

または、入力レベルが閾値を下回った場合には入力レベル) に到達するために、利得を増加させる期間である。リリース時間は、たとえば、25ミリ秒～2秒の範囲であってもよい。

【0494】

よって、いくつかの例において、個々のラウドスピーカー・ダイナミクス処理構成データは、複数のラウドスピーカーの各ラウドスピーカーについて、ダイナミックレンジ圧縮データセットを含むことができる。ダイナミックレンジ圧縮データセットは、閾値データ、入出力比データ、アタック・データ、リリース・データおよび／またはニー・データを含むことができる。これらのタイプの個々のラウドスピーカー・ダイナミクス処理構成データの一つまたは複数を組み合わせて、聴取環境ダイナミクス処理構成データを決定することができる。再生制限閾値の組み合わせに関して上述したように、いくつかの例では、ダイナミックレンジ圧縮データが平均されて、聴取環境ダイナミクス処理構成データを決定することができる。いくつかの事例では、ダイナミックレンジ圧縮データの最小値または最大値が、聴取環境ダイナミクス処理構成データ（たとえば、最大圧縮比）を決定するために使用されてもよい。他の実装では、たとえば、式(22)を参照して上述したようなチューニング・パラメータを介して、個々のラウドスピーカー・ダイナミクス処理のためのダイナミックレンジ圧縮データの最小と平均との間を補間する調整可能な組み合わせを作成することができる。

10

【0495】

上述のいくつかの例では、聴取環境ダイナミクス処理構成データの単一の集合（たとえば、組み合わせられた閾値

20

【数70】

$$\bar{T}[f]$$

の単一の集合) が、ダイナミクス処理の第1段における空間的ミックスのすべての成分に適用される。そのような実装は、ミックスの空間的バランスを維持することができるが、他の望ましくないアーチファクトを与えることがある。たとえば、隔離された空間領域内の空間的ミックスの非常に音量の大きな部分がミックス全体の音量を下げさせる場合に、「空間的ダッキング (spatial ducking)」が生じることがある。この音量の大きな成分から空間的に離れている、当該ミックスのより音量の小さな他の成分は、不自然に小さいと知覚されることがある。たとえば、音量の小さな背景音楽が、空間的ミックスのサラウンド・フィールドにおいて、組み合わせられた閾値

30

【数71】

$$\bar{T}[f]$$

よりも低いレベルで再生されていることがあり、よって、ダイナミクス処理の第1段によって空間的ミックスの制限は実行されない。次いで、空間的ミックスの前方（たとえば、映画のサウンドトラックのスクリーン上）に音量の大きな銃声が瞬間的に導入されることがあり、ミックスの全体的なレベルが組み合わせられた閾値を超えて上昇する。この瞬間、ダイナミクス処理の第1段は、ミックス全体のレベルを閾値

40

【数72】

$$\bar{T}[f]$$

より下に下げる。音楽が銃声とは空間的に離れているので、これは、音楽の連続的な流れにおける不自然なダッキングとして知覚されうる。

【0496】

そのような問題に対処するために、いくつかの実装は、空間的ミックスの異なる「空間

50

ゾーン」に対する独立したまたは部分的に独立したダイナミクス処理を許容する。空間ゾーンは、空間的ミックス全体がレンダリングされる空間領域のサブセットと考えられてもよい。以下の議論の多くは、再生制限閾値に基づくダイナミクス処理の例を提供するが、これらの概念は、他のタイプの個々のラウドスピーカー・ダイナミクス処理構成データおよび聴取環境ダイナミクス処理構成データにも等しく適用される。

【0497】

図52は、聴取環境の空間ゾーンの例を示す。図52は、空間的ミックスの領域（正方形全体によって表される）の例を示しており、それが前方、中央、およびサラウンドの3つの空間ゾーンに細分されている。

【0498】

図52の空間ゾーンは、硬い境界で描かれているが、実際には、ある空間ゾーンから別の空間ゾーンへの遷移を連続的なものとして扱うことが有益である。たとえば、正方形の左エッジの中央に位置する空間的ミックスの成分は、そのレベルの半分が前方ゾーンに割り当てられ、半分がサラウンドゾーンに割り当てられてもよい。空間的ミックスの各成分からの信号レベルは、この連続的な仕方で、各空間ゾーンに割り当てられ、蓄積されうる。すると、ダイナミクス処理機能は、各空間ゾーンについて独立に、ミックスからそれに割り当てられた全体的な信号レベルに対して作用することができる。空間的ミックスの各成分について、各空間ゾーンからのダイナミクス処理の結果（たとえば、周波数毎の時間変化する利得）がその後組み合わされて、その成分に適用されてもよい。いくつかの例において、空間ゾーン結果のこの組み合わせは、各成分について異なり、各ゾーンへのその特定の成分の割り当ての関数である。最終的な結果は、類似の空間ゾーン割り当てを有する空間的ミックスの成分が、類似のダイナミクス処理を受けるが、空間ゾーン間の独立性は許容されるというものである。空間ゾーンは、有利には、左右の不均衡のような好ましくない空間シフトを防止する一方で、空間的に独立した処理を許容する（たとえば、上述の空間的ダッキングのような他のアーチファクトを低減するため）ように選択されうる。

【0499】

空間ゾーンごとに空間的ミックスを処理する技法は、本開示のダイナミクス処理の第1段において有利に使用されうる。たとえば、諸スピーカー*i*にわたる個々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、再生制限閾値）の異なる組み合わせが、各空間ゾーンについて計算されてもよい。組み合わされたゾーン閾値の集合は

【数73】

$$\bar{T}_j[f]$$

によって表されてもよく、ここで、インデックス*j*は複数の空間ゾーンのうちの1つを指す。ダイナミクス処理モジュールは、各空間ゾーン上で独立して、その関連付けられた閾値

【数74】

$$\bar{T}_j[f]$$

を用いて動作してもよく、結果は、上述の技法に従って空間的ミックスの構成要素成分に戻して適用されうる。

【0500】

それぞれが関連付けられた所望の空間位置（可能性としては時間変化する）を有する、*K*個の個々の構成要素信号 $x_k[t]$ の合計から構成される空間信号がレンダリングされることを考える。ゾーン処理を実装するための1つの具体的な方法は、各オーディオ信号 $x_k[t]$ がゾーン*j*にどれだけ寄与するかを記述する時間変化するパン利得 $\alpha_{kj}[t]$ を、ゾーンの位置に関するオーディオ信号の所望の空間位置の関数として計算することに関わる。これらのパン利得は、有利には、利得の2乗の和が1に等しいことを要求するパワー保存パン則に従うように設計されうる。これらのパン利得から、ゾーン信号 $s_j[t]$ は、構成要素信号にそのゾ

ーンについてのそれらのパン利得によって重み付けしたものの和として計算されうる：

【数 7 5】

$$s_j[t] = \sum_{k=1}^K \alpha_{kj}[t] x_k[t] \quad \text{Equation (23)}$$

次いで、各ゾーン信号は、ゾーン閾値

【数 7 6】

$$\bar{T}_j[f]$$

10

によってパラメータ化されたダイナミクス処理関数DPによって独立して処理され、周波数および時間変化するゾーン修正利得 G_j を生成する：

【数 7 7】

$$G_j[f, t] = DP\{s_j[t], \bar{T}_j[f]\} \quad \text{Equation (24)}$$

次いで、周波数および時間変化する修正利得は、ゾーン修正利得を、その信号の、諸ゾーンののためのパン利得に比例して組み合わせることによって、各個々の構成要素信号について計算されうる：

【数 7 8】

$$G_k[f, t] = \sqrt{\sum_{j=1}^J (\alpha_{kj} G_j[f, t])^2} \quad \text{Equation (25)}$$

20

これらの信号修正利得 G_k は、次いで、たとえば、フィルタバンクを使用して、各構成要素信号に適用されて、ダイナミクス処理された構成要素信号

【数 7 9】

$$\hat{x}_k[t]$$

30

を生成してもよい。該ダイナミクス処理された構成要素信号が、その後、これをスピーカ一信号にレンダリングされうる。

【0 5 0 1】

各空間ゾーンについての個々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、スピーカー再生制限閾値）の組み合わせは、多様な仕方で実行されうる。一例として、空間ゾーン再生制限閾値

【数 8 0】

$$\bar{T}_j[f]$$

40

は、空間ゾーンおよびスピーカーに依存する重み付け $w_{ij}[f]$ を使用して、スピーカー再生制限閾値 $T_i[f]$ の重み付けされた和として計算されうる：

【数 8 1】

$$\bar{T}_j[f] = \sum_i w_{ij}[f] T_i[f] \quad \text{Equation (26)}$$

【0 5 0 2】

同様の重み付け関数は、他のタイプの個々のラウドスピーカー・ダイナミクス処理構成データにも適用されうる。有利には、空間ゾーンの組み合わせられた個々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、再生制限閾値）は、その空間ゾーンに関連

50

する空間的ミックスの再生成分に最も寄与するスピーカーの個々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、再生制限閾値）に向けてバイアスされてもよい。そのようなバイアスかけは、いくつかの例では、周波数 f についてそのゾーンに関連する空間的ミックスの成分をレンダリングすることについての各スピーカーの寄与に応じて、重み $w_{ij}[f]$ を設定することによって達成することができる。

【0503】

図53は、図52の空間ゾーン内のラウドスピーカーの例を示している。図53は、図52の同じゾーンを示しているが、空間的ミックスをレンダリングするのに寄与する5つの例示的なラウドスピーカー（スピーカー1、2、3、4、5）の位置が重ねられている。この例では、ラウドスピーカー1、2、3、4、5はダイヤ形で表されている。この特定の例では、スピーカー1は中央ゾーンのレンダリング、スピーカー2および5は前方ゾーン、スピーカー3および4はサラウンドゾーンを主に受け持つ。スピーカーの空間ゾーンへのこの概念的な1対1のマッピングに基づいて重み $w_{ij}[f]$ を生成することができるが、空間的ミックスの空間ゾーンベースの処理と同様に、より連続的なマッピングのほうが好ましいことがある。たとえば、スピーカー4は前方ゾーンに非常に近く、スピーカー4と5の間に位置するオーディオ・ミックスの成分は（概念的な前方ゾーンではあるが）主にスピーカー4と5の組み合わせによって再生される可能性が高いであろう。よって、スピーカー4の個々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、再生制限閾値）が、サラウンドゾーンと同様に前方ゾーンの組み合わせされた個々のラウドスピーカー・ダイナミクス処理構成データ（たとえば、再生制限閾値）に寄与することは、意味がある。

【0504】

この連続的なマッピングを達成する一つの方法は、空間ゾーン j に関連する成分をレンダリングする際の各スピーカー i の相対的寄与を記述するスピーカー参加値に等しい重み $w_{ij}[f]$ を設定することである。そのような値は、スピーカーにレンダリングすることを受け持つレンダリング・システム（たとえば、上述のステップ（c）から）および各空間ゾーンに関連する一つまたは複数の公称空間位置の集合から直接導出されてもよい。公称空間位置のこの集合は、各空間ゾーン内の位置の集合を含んでいてもよい。

【0505】

図54は、図53の空間ゾーンおよびスピーカーに重ねられた公称空間位置の例を示している。公称位置は、番号付きの円で示されている。すなわち、前方ゾーンには正方形の上のコーナーに位置する2つの位置が関連付けられ、中央ゾーンには正方形の上の中央にある単一の位置が関連付けられ、サラウンドゾーンには正方形の下のコーナーに位置する2つの位置が関連付けられている。

【0506】

空間ゾーンについてのスピーカー参加値を計算するために、そのゾーンに関連する公称位置のそれぞれは、その位置に関連するスピーカー・アクティブ化を生成するために、レンダラーを通じてレンダリングされてもよい。これらのアクティブ化は、たとえば、CMA Pの場合は各スピーカーについての利得であってもよく、FVの場合は各スピーカーについて所与の周波数における複素数値であってもよい。次に、各スピーカーおよびゾーンについて、これらのアクティブ化は、空間ゾーンに関連する各公称位置にわたって累積されて、値 $g_{ij}[f]$ を生成してもよい。この値は、空間ゾーン j に関連した公称位置の集合全体をレンダリングするためのスピーカー i の全アクティブ化を表す。最後に、空間ゾーンにおけるスピーカー参加値は、諸スピーカーにわたるこれらのすべての累積されたアクティブ化の和によって正規化された累積アクティブ化として計算されてもよい。その後、前記重みは、このスピーカー参加値に設定されてもよい：

【数82】

$$w_{ij}[f] = \frac{g_{ij}[f]}{\sum_i g_{ij}[f]} \quad \text{Equation (27)}$$

10

20

30

40

50

上述の正規化は、すべてのスピーカー*i*にわたる $w_{ij}[f]$ の和が1に等しいことを保証し、これは、式8の重みについての望ましい属性である。

【0507】

いくつかの実装によれば、スピーカーの参加値を計算し、これらの値の関数として閾値を組み合わせるための上述のプロセスは、静的プロセスとして実行されてもよい。ここで、結果として得られる組み合わせられた閾値は、環境中のスピーカーのレイアウトおよび能力を決定するセットアップ手順の間に一度計算される。そのようなシステムでは、いったんセットアップされると、個々のラウドスピーカーのダイナミクス処理構成データと、レンダリング・アルゴリズムが所望のオーディオ信号位置の関数としてラウドスピーカーをアクティブ化する仕方との両方が、静的なままであると想定されうる。しかしながら、ある種のシステムでは、これらの側面の両方が時間とともに、たとえば再生環境における条件の変化にตอบสนองして、変化することがあり、よって、そのような変動を考慮に入れるために、連続的なまたはイベントトリガー式のいずれかで、上述のプロセスに従って組み合わせられた閾値を更新することが望ましいことがありうる。

【0508】

CMAPおよびFVレンダリング・アルゴリズムは両方とも、聴取環境の変化にตอบสนองして、一つまたは複数の動的に構成可能な機能に適合するように、拡張されてもよい。たとえば、図53に関して、スピーカー3の近くに位置する人が、スピーカーに関連付けられたスマートアシスタントのウェイクワードを発することができ、それにより、システムを、人からのその後のコマンドを聞く準備ができた状態にすることができる。ウェイクワードが発される間に、システムは、ラウドスピーカーに関連付けられたマイクロフォンを使って、前記人の位置を決定することができる。この情報を用いて、システムは、次いで、スピーカー3上のマイクロフォンがその人をよりよく聞き取れるように、スピーカー3から再生されるオーディオのエネルギーを他のスピーカーに転じる（divert）ことを選択することができる。そのようなシナリオでは、図53のスピーカー2が、ある時間期間にわたって、スピーカー3の役割を本質的に「引き継いで」もよく、結果として、サラウンドゾーンについてのスピーカー参加値は著しく変化し、スピーカー3の参加値は減少し、スピーカー2の参加値は増加する。ゾーン閾値は、変化したスピーカー参加値に依存するので、その後再計算されてもよい。レンダリング・アルゴリズムへのこれらの変更に対して代替的または追加的に、スピーカー3の制限閾値は、スピーカーが歪むのを防ぐように設定された公称値よりも下に下げられてもよい。これは、スピーカー3から再生される残りのオーディオが、人を傾聴するマイクロフォンへの干渉を引き起こすと決定された何らかの閾値を超えて増加しないようにすることができる。ゾーン閾値もまた個々のスピーカー閾値の関数であるため、この場合にも更新されうる。

【0509】

図55は、本明細書に開示されたもののような装置またはシステムによって実施されうる方法の一例を概説するフロー図である。方法5500のブロックは、本明細書に記載された他の方法と同様に、必ずしも示された順序で実行されるわけではない。いくつかの実装では、方法5500の一つまたは複数のブロックが同時に実行されてもよい。さらに、方法5500のいくつかの実装は、図示および／または説明されるよりも多いまたは少ないブロックを含んでいてもよい。方法5500のブロックは、図6に示されて上述した制御システム610のような制御システム、または他の開示された制御システムの例の1つであってもよい（またはそれを含んでいてもよい）一つまたは複数の装置によって実行されてもよい。いくつかの例では、方法5500は、少なくとも部分的には、オーディオ・セッション・マネージャ、たとえば、図2CのCHASM 208C、図2DのCHASM 208D、図3CのCHASM 307または図4のCHASM 401からの命令に従って、実行されてもよい。

【0510】

この例によれば、ブロック5505は、制御システムによって、インターフェース・システムを介して、聴取環境の複数のラウドスピーカーのそれぞれについて個々のラウドスピーカー・ダイナミクス処理構成データを取得することに関わる。この実装では、個々のラ

ウドスピーカー・ダイナミクス処理構成データは、複数のラウドスピーカーの各ラウドスピーカーについての個々のラウドスピーカー・ダイナミクス処理構成データセットを含む。いくつかの例によれば、一つまたは複数のラウドスピーカーのための個々のラウドスピーカー・ダイナミクス処理構成データは、前記一つまたは複数のラウドスピーカーの一つまたは複数の能力に対応する。この例では、個々のラウドスピーカー・ダイナミクス処理構成データセットの各データセットは、少なくとも1つのタイプのダイナミクス処理構成データを含む。

【0511】

いくつかの事例では、ブロック5505は、聴取環境の複数のラウドスピーカーのそれぞれから個々のラウドスピーカー・ダイナミクス処理構成データセットを取得することに関わってもよい。他の例では、ブロック5505は、メモリに記憶されたデータ構造から個々のラウドスピーカー・ダイナミクス処理構成データセットを取得することに関わってもよい。たとえば、個々のラウドスピーカー・ダイナミクス処理構成データセットは、たとえば各ラウドスピーカーについてのセットアップ手順の一部として以前に取得されて、データ構造に格納されていてもよい。

【0512】

いくつかの例によれば、個々のラウドスピーカー・ダイナミクス処理構成データセットは、独自仕様 (proprietary) であってもよい。いくつかのそのような例では、個々のラウドスピーカー・ダイナミクス処理構成データセットは、類似の特性を有するスピーカーについての個々のラウドスピーカー・ダイナミクス処理構成データに基づいて、以前に推定されたものであってもよい。たとえば、ブロック5505は、複数のスピーカーおよび該複数のスピーカーのそれぞれについての対応する個々のラウドスピーカー・ダイナミクス処理構成データセットを示すデータ構造から、最も類似したスピーカーを決定するスピーカー・マッチング・プロセスに関わってもよい。スピーカー・マッチング・プロセスは、たとえば、一つまたは複数のウーファ、ツイーターおよび／またはミッドレンジ・スピーカーのサイズの比較に基づいてもよい。

【0513】

この例では、ブロック5510は、制御システムによって、複数のラウドスピーカーのための聴取環境ダイナミクス処理構成データを決定することに関わる。この実装によれば、聴取環境ダイナミクス処理構成データの決定は、複数のラウドスピーカーの各ラウドスピーカーについての個々のラウドスピーカー・ダイナミクス処理構成データセットに基づく。聴取環境ダイナミクス処理構成データを決定することは、ダイナミクス処理構成データセットの個々のラウドスピーカー・ダイナミクス処理構成データを、たとえば、一つまたは複数のタイプの個々のラウドスピーカー・ダイナミクス処理構成データの平均を取ることによって組み合わせることに関わってもよい。いくつかの事例では、聴取環境ダイナミクス処理構成データを決定することは、一つまたは複数のタイプの個々のラウドスピーカー・ダイナミクス処理構成データの最小値または最大値を決定することに関わってもよい。いくつかのそのような実装によれば、聴取環境ダイナミクス処理構成データを決定することは、一つまたは複数のタイプの個々のラウドスピーカー・ダイナミクス処理構成データの最小値または最大値と平均値との間を補間することに関わってもよい。

【0514】

この実装では、ブロック5515は、制御システムによって、インターフェース・システムを介して、一つまたは複数のオーディオ信号および関連する空間データを含むオーディオ・データを受領することに関わる。たとえば、空間データは、オーディオ信号に対応する意図された知覚された空間位置を示してもよい。この例では、空間データはチャンネル・データおよび／または空間メタデータを含む。

【0515】

この例では、ブロック5520は、制御システムによって、聴取環境ダイナミクス処理構成データに基づいてオーディオ・データに対してダイナミクス処理を実行して、処理されたオーディオ・データを生成することに関わる。ブロック5520のダイナミクス処理は、

10

20

30

40

50

一つまたは複数の再生制限閾値、圧縮データなどを適用することを含むがそれに限定されない、本明細書に開示されている本開示のダイナミクス処理方法のいずれかに関わってもよい。

【0516】

ここで、ブロック5525は、複数のラウドスピーカーの少なくとも一部を含むラウドスピーカーの集合を介した再生のために、制御システムによって、処理されたオーディオ・データをレンダリングして、レンダリングされたオーディオ信号を生成することに関わる。いくつかの例では、ブロック5525は、CMAPレンダリング・プロセス、FVレンダリング・プロセス、または両者の組み合わせを適用することに関わってもよい。この例では、ブロック5520は、ブロック5525の前に実行される。しかしながら、上述のように、ブロック5520および／またはブロック5510は、少なくとも部分的に、ブロック5525のレンダリング・プロセスに基づいていてもよい。ブロック5520および5525は、図49の聴取環境ダイナミクス処理モジュールおよびレンダリング・モジュール4920を参照して上述したようなプロセスを実行することに関わってもよい。

10

【0517】

この例によれば、ブロック930は、インターフェース・システムを介して、レンダリングされたオーディオ信号をラウドスピーカーの集合に提供することに関わる。一例では、ブロック930は、スマートホームハブ4905によって、そのインターフェース・システムを介して、レンダリングされたオーディオ信号をラウドスピーカー4925a～4925mに提供することに関わってもよい。

20

【0518】

いくつかの例では、方法5500は、レンダリングされたオーディオ信号が提供されるラウドスピーカーの集合の各ラウドスピーカーについての個々のラウドスピーカー・ダイナミクス処理構成データに従って、レンダリングされたオーディオ信号に対してダイナミクス処理を実行することに関わってもよい。たとえば、再び図49を参照すると、ダイナミクス処理モジュールA～Mは、ラウドスピーカー4925a～4925mについての個々のラウドスピーカー・ダイナミクス処理構成データに従って、レンダリングされたオーディオ信号に対してダイナミクス処理を実行することができる。

【0519】

いくつかの実装では、個々のラウドスピーカー・ダイナミクス処理構成データは、複数のラウドスピーカーの各ラウドスピーカーについての再生制限閾値データセットを含んでいてもよい。いくつかのそのような例では、再生制限閾値データセットは、複数の周波数のそれぞれについての再生制限閾値を含んでいてもよい。

30

【0520】

聴取環境ダイナミクス処理構成データを決定することは、いくつかの事例では、複数のラウドスピーカーにわたる最小の再生制限閾値を決定することに関わってもよい。いくつかの例では、聴取環境ダイナミクス処理構成データを決定することは、複数のラウドスピーカーにわたる平均された再生制限閾値を得るために再生制限閾値を平均することに関わってもよい。いくつかのそのような例では、聴取環境ダイナミクス処理構成データを決定することは、複数のラウドスピーカーにわたる最小の再生制限閾値を決定し、最小の再生制限閾値と平均された再生制限閾値との間を補間することに関わってもよい。

40

【0521】

いくつかの実装によれば、再生制限閾値を平均することは、再生制限閾値の重み付けされた平均を決定することに関わってもよい。いくつかのそのような例では、重み付けされた平均は、制御システムによって実装されるレンダリング・プロセスの特性、たとえばブロック5525のレンダリング・プロセスの特性に少なくとも部分的に基づいてよい。

【0522】

いくつかの実装では、オーディオ・データに対してダイナミクス処理を実行することは、空間ゾーンに基づいていてもよい。空間ゾーンのそれぞれは、聴取環境のサブセットに対応する。

50

【0523】

いくつかのそのような実装によれば、ダイナミクス処理は、空間ゾーンのそれぞれについて別々に実行されてもよい。たとえば、聴取環境ダイナミクス処理構成データを決定することは、空間ゾーンのそれぞれについて別々に実行されてもよい。たとえば、複数のラウドスピーカーにわたるダイナミクス処理構成データセットを組み合わせることは、一つまたは複数の空間ゾーンのそれぞれについて別々に実行されてもよい。いくつかの例では、一つまたは複数の空間ゾーンのそれぞれについて別々に、複数のラウドスピーカーにわたるダイナミクス処理構成データセットを組み合わせることは、少なくとも部分的には、一つまたは複数の空間ゾーンにわたる所望のオーディオ信号位置に応じた、レンダリング・プロセスによるラウドスピーカーのアクティブ化に基づいていてもよい。

10

【0524】

いくつかの例では、一つまたは複数の空間ゾーンのそれぞれについて別々に、複数のラウドスピーカーにわたるダイナミクス処理構成データセットを組み合わせることは、少なくとも部分的には、一つまたは複数の空間ゾーンのそれぞれにおける各ラウドスピーカーについてのラウドスピーカー参加値に基づいていてもよい。各ラウドスピーカー参加値は、少なくとも部分的には、一つまたは複数の空間ゾーンのそれぞれの中の一つまたは複数の公称空間位置に基づいてもよい。公称空間位置は、いくつかの例では、ドルビー5.1、ドルビー5.1.2、ドルビー7.1、ドルビー7.1.4またはドルビー9.1のサラウンドサウンド・ミックス内のチャンネルの標準位置に対応してもよい。いくつかのそのような実装では、各ラウドスピーカー参加値は、少なくとも部分的には、一つまたは複数の空間ゾーンのそれぞれの中の一つまたは複数の公称空間位置のそれぞれにおけるオーディオ・データのレンダリングに対応する各ラウドスピーカーのアクティブ化に基づいている。

20

【0525】

いくつかのそのような例によれば、再生制限閾値の重み付けされた平均は、少なくとも部分的には、オーディオ信号の空間ゾーンへの近接性の関数としての、レンダリング・プロセスによるラウドスピーカーのアクティブ化に基づいていてもよい。いくつかの事例では、重み付けされた平均は、少なくとも部分的には、各空間ゾーンにおける各ラウドスピーカーについてのラウドスピーカー参加値に基づいてもよい。いくつかのそのような例では、各ラウドスピーカー参加値は、少なくとも部分的には、各空間ゾーン内の一つまたは複数の公称空間位置に基づいてもよい。たとえば、公称空間位置は、ドルビー5.1、ドルビー5.1.2、ドルビー7.1、ドルビー7.1.4、またはドルビー9.1のサラウンドサウンド・ミックス内のチャンネルの標準的な位置に対応してもよい。いくつかの実装では、各ラウドスピーカー参加値は、少なくとも部分的には、各空間ゾーン内の一つまたは複数の公称空間位置のそれぞれにおけるオーディオ・データのレンダリングに対応する各ラウドスピーカーの起動に基づいてもよい。

30

【0526】

図56Aおよび56Bは、いくつかの実施形態に従って実装可能なシステムの例を示す。図56Bは、図56Aにおけるユーザーの位置5601が、図56Bにおけるユーザーの位置113と異なるという点で、図56Aと異なる。図56Aおよび56Bにおいて、ラベル付けされた要素は、以下の通りである：

40

5607: ゾーン1;

5612: ゾーン2;

5601: ゾーン1におけるユーザー（話者）位置;

5602: 直接ローカル音声（ユーザーが発声）;

5603: ゾーン1に配置されたスマート・オーディオ装置（たとえば、音声アシスタント装置）内の複数のスピーカー;

5604: ゾーン1に位置するスマート・オーディオ装置（たとえば、音声アシスタント装置）における複数のマイクロフォン;

5605: ゾーン1に位置する家電製品、たとえばランプ;

5606: ゾーン1に位置する家電製品における複数のマイクロフォン;

50

5613: ゾーン2におけるユーザー（話者）位置;

5608: ゾーン2に位置するスマート・オーディオ装置（たとえば、音声アシスタント装置）内の複数のスピーカー;

5609: スマート・オーディオ装置（たとえば、ゾーン2に位置する音声アシスタント装置）における複数のマイクロフォン;

5610: ゾーン2に位置する家電製品（冷蔵庫など）;

5611: ゾーン2に位置する家電製品における複数のマイクロフォン。

【0527】

図57は、ある実施形態による環境（たとえば、家庭）において実装されるシステムのブロック図である。システムは、ユーザー位置を追跡するために「フォロー・ミー〔ついてこい〕」機構を実装する。図57において、ラベル付けされた要素は以下の通りである:

5701: 入力を受け、（入力に応答して）判別された活動（たとえば、入力5706Aによって示される）のために使用する最良のマイクロフォンおよびラウドスピーカーについて決定するように構成されたサブシステム（モジュールまたは「フォロー・ミー」モジュールと呼ばれることもある）;

5701A: 決定された活動のために使用するシステムの最良のラウドスピーカー（単数または複数）および／またはユーザー（たとえば、話者）が現在位置するゾーン（すなわち、ゾーンマップ5703によって示されるゾーンのうちの1つ）に関する決定（モジュール5701で決定される）を示すデータ;

5701B: 決定された活動のために使用するシステムの最良のマイクロフォン（単数または複数）および／またはユーザーが現在位置するゾーン（すなわち、ゾーンマップ5703によって示されるゾーンのうちの1つ）に関する決定（モジュール5701で決定される）を示すデータ;

5702: ユーザー位置特定サブシステム（モジュール）。たとえば環境のあるゾーン内のユーザー（たとえば、話者、たとえば図56Aまたは56Bのユーザー）の位置を決定するように構成されている。いくつかの実施形態では、サブシステム5702は、ユーザーのゾーンを推定するように構成される（たとえば、マイクロフォン5705のうちの少なくともいくつかから導出される複数の音響特徴に従って）。いくつかのかかる実施側面において、目標は、ユーザーの正確な幾何学的位置を推定することではなく、ユーザーが位置する離散的なゾーンの（たとえば、重いノイズおよび残留エコーの存在下での）堅牢な推定を形成することである;

5702A: モジュール5702によって決定され、モジュール5701に対して呈される、ユーザー（話者）の現在位置を示す情報（データ）;

5703: ゾーンマップサブシステム。これは、システムの環境の諸ゾーン（たとえば、システムが図56Aおよび56Bの環境内にある場合、図56Aおよび56Bの諸ゾーン）を示すゾーンマップ、および諸ゾーン内の位置によってグループ化されたシステムのすべてのマイクロフォンおよびラウドスピーカーのリストを提供する。いくつかの実装では、サブシステム5703は、ゾーンマップを示すデータを格納するメモリであるか、または、それを含む。いくつかの例によれば、サブシステム5701、5702および／または5703の機能は、本明細書でSPASMと称されるもの（たとえば、図2CのSPASM 207Cを参照）によって提供されてもよい;

5703A: （ゾーンマップの）少なくとも1つのゾーンに関する情報（データ）、およびゾーンマップのかかる各ゾーン（たとえば、ゾーンの少なくともサブセットのそれぞれ）に含まれる複数のマイクロフォンおよびラウドスピーカーに関する情報。（システムのいくつかの実装においては）モジュール5701および／またはモジュール5702に対して呈される;

【0528】

5704: マイクロフォン5705の出力の前処理を実行するように結合され構成された前処理サブシステム。サブシステム5704は、一つまたは複数のマイクロフォン前処理サブシステム（たとえば、エコー管理サブシステム、ウェイクワード検出器、および／または発

10

20

30

40

50

話認識サブシステムなど)を実装してもよい。よって、サブシステム5704は、本明細書において「メディア・エンジン」と称されうるものの構成要素の一例である(たとえば、図4のメディア・エンジン440、441、および442を参照)；

5704A: サブシステム5704によって生成され、サブシステム5704から出力される前処理されたマイクロフォン信号；

5705: 複数のマイクロフォン(たとえば、図56Aおよび56Bのマイクロフォン5604、5606、5609、および5611を含む)；

【0529】

5706: 少なくとも1つの現在のオーディオ活動(たとえば、複数の現在進行中のオーディオ活動)を実装するように結合され構成されたサブシステム。そのようなオーディオ活動(便宜上、本明細書では「活動」と呼ばれることもある)のそれぞれは、音の検出(少なくとも1つのマイクロフォンを使って)および／または音の生成(少なくとも1つのラウドスピーカーから音を放射することによる)を含む。そのようなオーディオ活動の例は、音楽再生(たとえば、サブシステム5707によるレンダリングのためのオーディオを提供するステップを含む)、ポッドキャスト(たとえば、サブシステム5707によるレンダリングのためのオーディオを提供するステップを含む)、および／または電話通話(たとえば、サブシステム5707によるレンダリングのためのテレビ会議オーディオを提供すること、およびサブシステム5704に提供される各マイクロフォン信号を処理および／または送信することを含む)を含むが、これらに限定されない。よって、サブシステム5706は、本明細書において「ソフトウェア・アプリケーション」、「アプリケーション」、「アプリ」として、またはソフトウェア・アプリケーション、アプリケーション、またはアプリを実行するために構成される装置として言及されうるものの一例である(たとえば、図4のアプリケーション410、411、および412を参照)；

5706A: サブシステム5706によって生成され、サブシステム5706からモジュール5701に対して呈される、サブシステム5706によって実装される現在進行中の活動(単数または複数)に関する情報(データ)；

【0530】

5707: システムの少なくとも1つの現在の活動の実行中に生成される、または他の方法で提供される、オーディオをレンダリングする(たとえば、スピーカー5708を駆動するためのスピーカー・フィードを生成することによって)ように結合および構成されるマルチチャンネルスピーカー・レンダラー・サブシステム。たとえば、サブシステム5707は、データ5701Aに従って、スピーカー5708(異なるスマート・オーディオ装置内に実装され、または、それに結合されてもよい)のサブセットによる再生のためにオーディオをレンダリングするように実装されてもよい。該レンダリングは、関連するラウドスピーカーによって発せられる音が、ユーザーの現在位置(たとえば、ゾーン)において、ユーザーによって(たとえば、明確に、または、最良のもしくはは所望の仕方)で知覚可能であるようにする。よって、サブシステム5707は、本明細書において「メディア・エンジン」と称されうるものの構成要素の一例である(たとえば、図4のメディア・エンジン440、441、および442を参照)；

【0531】

5708: 複数のラウドスピーカー(たとえば、図56Aおよび56Bの5603および5608を含む)；

5709: ユーザー(たとえば、話者、たとえば図56Aまたは56Bのユーザー)からの音声コマンド。これは、システムの典型的な実装においては、サブシステム5704から出力され、モジュール5701に提供される。

【0532】

要素5701、5702、および5703(または要素5702および5703)は、図57システムのユーザー位置特定および活動制御サブシステムと総称されうる。

【0533】

図57のシステム(およびいくつかの他の実施形態)の要素は、スマート・オーディオ装

10

20

30

40

50

置内に実装されてもよく、またはスマート・オーディオ装置に結合されてもよい。たとえば、ラウドスピーカー5708の全部もしくは一部、および／またはマイクロフォン5705の全部もしくは一部は、一つまたは複数のスマート・オーディオ装置内に実装されてもよく、または、それに結合されていてもよく、あるいは少なくとも一部のマイクロフォンおよびラウドスピーカーは、Bluetooth送信機/受領機（たとえば、スマートフォン）に接続されたBluetooth装置内に実装されてもよい。また、たとえば、図57システムの一つまたは複数の他の要素（たとえば、要素5701、5702、5703、5704、および5706の全部または一部）（および／または、以下に説明する図60システムの要素5701、5702、5703、5704、5706、および6011の全部または一部）は、スマート・オーディオ装置内に実装されてもよく、またはスマート・オーディオ装置に結合されてもよい。そのような例示的な実施形態では、「フォロー・ミー」モジュール5701（および他のシステム要素）は、システムの少なくとも1つのマイクロフォンによって検出される音（ユーザーによって発生される）にตอบสนองしてユーザーの位置を追跡することによって、スマート・オーディオ装置を協調させる（オーケストレーションする）ように動作する。たとえば、そのような協調は、システムの要素によって発生される音のレンダリング、および／またはシステムのマイクロフォンの出力の処理、および／またはシステムによって実施される少なくとも1つの活動（たとえば、システムの要素5706によって、たとえば、図60の活動マネージャ6011またはシステムの別の活動マネージャを制御することによって）の調整を含む。

【0534】

典型的には、サブシステム5702および5703は、緊密に統合される。サブシステム5702は、（たとえば、非同期マイクロフォンとして実装される）マイクロフォン5705の全部または一部（たとえば、2つ以上）の出力を受領してもよい。サブシステム5702は、いくつかの例ではシステムのスマート・オーディオ装置に実装される、分類器を実装してもよい。他の例では、分類器は、マイクロフォンと通信するように結合および構成されたシステムの別のタイプの装置（たとえば、オーディオを提供するように構成されていないスマート装置）によって実装されてもよい。たとえば、マイクロフォン5705の少なくとも一部は、いかなるスマート・オーディオ装置にも含まれないが、分類器としてサブシステム5702を実装する装置と通信するように構成される離散的マイクロフォン（たとえば、家電製品内のもの）であってもよく、分類器は、各マイクロフォンの出力信号から導出される複数の音響特徴に応じて、ユーザーのゾーンを推定するように構成されてもよい。いくつかのそのような実施形態において、目標は、ユーザーの正確な幾何学的位置を推定することではなく、（たとえば、ひどい雑音および残留エコーの存在下で）離散的ゾーンの堅牢な推定を形成することである。

【0535】

ここで、環境におけるオブジェクト、ユーザー、または話者の「幾何学的位置」（前述および以下の説明で言及される）という表現は、システム環境全体を基準とする（たとえば、環境内のどこかにその原点を有するデカルト座標系または極座標系に従う）、または環境内の特定の装置（たとえば、スマートオーディオ装置）を基準とする（たとえば、該装置をその原点とするデカルト座標系または極座標系に従う）、座標系（たとえば、GPS座標を基準とする座標系）に基づく位置をいう。いくつかの実装では、サブシステム5702は、マイクロフォン5705の幾何学的位置を参照することなく、環境内のユーザーの位置の推定値を決定するように構成される。

【0536】

「フォロー・ミー」モジュール5701は、いくつかの入力（5702A、5703A、5706A、および5709のうちの一つまたは複数）にตอบสนองして動作し、出力5701A、および5701Bのうちの一方または両方を生成するように結合および構成される。次に、入力の例をより詳細に説明する。

【0537】

入力5703Aは、ゾーンマップの各ゾーン（時には、音響ゾーンと呼ばれる）に関する情報を示すことができ、これは、限定されるものではないが、各ゾーン内に位置するシステ

10

20

30

40

50

ムの装置（たとえば、スマート装置、マイクロフォン、ラウドスピーカーなど）のリスト、各ゾーンの寸法（たとえば、幾何学的位置特定ユニットと同じ座標系での）、環境に関するおよび／または他のゾーンに関する各ゾーンの幾何学的位置（たとえば、キッチン、居間、寝室など）、システムの各装置の幾何学的位置（たとえば、それぞれのゾーンに関する、および／または、それらの装置のお互いのものに関する）、および／または、各ゾーンの名前のうちの一つまたは複数を含む。

【0538】

入力5702Aは、ユーザー（話者）が位置する音響ゾーン、そのようなゾーン内の話者の幾何学的位置、および話者がそのようなゾーン内にいた期間のうちの全部または一部に関するリアルタイム情報（データ）であってもよく、またはそれを含んでもよい。入力5702Aは、また、前の文に記された情報のいずれかの正確さまたは正確さ、および／または（たとえば、パラメータNが構成可能であるとして、過去N時間内の）話者の動きの履歴に関して、ユーザー位置モジュール5702による信頼度を含んでもよい。

10

【0539】

入力5709は、ユーザー（話者）によって発せられる音声コマンド、または2つ以上の音声コマンドであってもよく、各音声コマンドは、前処理サブシステム5704によって検出されている（たとえば、「フォロー・ミー」モジュール5701の機能性に関連するまたは関連しないコマンド）。

【0540】

モジュール5701の出力5701Aは、話者の現在の（たとえば、直近に決定された）音響ゾーンに応じて処理を適応させるための、レンダリング・サブシステム（レンダラー）5707への命令である。モジュール5701の出力5701Bは、話者の現在の（たとえば、直近に決定された）音響ゾーンに応じて処理を適応させるための、前処理サブシステム5704への命令である。

20

【0541】

出力5701Aは、話者の現在の音響ゾーンに関する話者の幾何学的位置、ならびに話者に関するラウドスピーカー5708のそれぞれの幾何学的位置および距離を示すことができ、たとえば、レンダラー5707に、システムによって実装される関連する活動のために可能な最良の方法でレンダリングを実行させる。可能な最良の方法は、活動およびゾーンに依存することがあり、また任意的に、話者の事前に決定された（たとえば、記録された）選好にも依存することがある。たとえば、活動が映画であり、話者が居間にいる場合、出力5701Aは、レンダラー5707に、映画館のような体験のためにできるだけ多くのラウドスピーカーを使用して映画のオーディオを再生するように命令することができる。活動が音楽またはポッドキャストであり、話者がキッチンまたは寝室にいる場合、出力5701Aは、レンダラー5707に、より親密な体験のために、最も近いラウドスピーカーのみで音楽をレンダリングするように命令することができる。

30

【0542】

出力5701Bは、サブシステム5704によって使用されるためのマイクロフォン5705の一部または全部（すなわち、出力がサブシステム5704によって無視されるべきではなく、代わりに、使用されるべき（すなわち、処理されるべき）であるマイクロフォン）のソートされたリスト、およびユーザー（話者）に対する各そのようなマイクロフォンの幾何学的位置を示してもよい。いくつかの実施形態では、サブシステム5704は、各マイクロフォンの話者からの距離（出力5701Bによって示される）、利用可能であれば、各マイクロフォンについてのウェイクワード・スコア（すなわち、そのマイクロフォンがユーザーによって発せられたウェイクワードを聞いた確からしさ）、各マイクロフォンの信号対雑音比（すなわち、環境ノイズおよび／またはそのマイクロフォンから捕捉されたオーディオ再生に対して、話者によって発せられた発話がどれだけ大きい）、または前述の2つ以上の組み合わせのうちの一つまたは複数によって決定される方法で、マイクロフォン5705の一部または全部の出力を処理してもよい。ウェイクワード・スコアおよび信号対雑音比は、前処理サブシステム5704によって計算されてもよい。

40

50

電話通話のようないくつかのアプリケーションでは、サブシステム5704は、（リストによって示されるように）マイクロフォン5705のうちの最良のものの出力のみを使用してよく、またはリストからの複数のマイクロフォンからの信号でビームフォーミングを実装してもよい。（たとえば）分散式の発話認識器または分散式のウェイクワード検出器などのいくつかのアプリケーションを実装するために、サブシステム5704は、マイクロフォン5705のうち複数のもの（たとえば、出力5701Bによって示されるソートされたリストから決定される；ここで、ソートは、たとえば、ユーザーに近い順であってもよい）の出力を使用してよい。

【0543】

いくつかの例示的なアプリケーションでは、サブシステム5704は（モジュール5701および5702と共に）、出力5701Bを使用して（すなわち、少なくとも部分的に出力5701Bに応答して）、（たとえば、ウェイクワードに続くコマンドをより良く認識するために）より効果的にユーザーのゾーンからの音をピックアップしようと試みる、マイクロフォン選択または適応ビームフォーミング方式を実装する。そのようなシナリオでは、モジュール5702は、さまざまな仕方のうちのいずれかにおけるユーザーゾーン決定を改善するためのユーザーゾーン予測の品質に関するフィードバックとして、サブシステム5704の出力5704Aを使用してよい。これには、以下が含まれるが、これらに限定されない：

ウェイクワードに続く音声コマンドの誤認識につながる予測にペナルティを課す。たとえば、ユーザーが（たとえば、「アマンダ、やめろ!」のような反対コマンドを発声することにより）コマンドへの音声アシスタントの応答を遮る結果となるユーザーゾーン予測は、ペナルティを課されてもよい；

発話認識器（サブシステム5704によって実装される）がコマンドを正常に認識したことの低い信頼度につながる予測にペナルティを課す；

（サブシステム5704によって実装される）第2パスの（second-pass）ウェイクワード検出器が、高い信頼度で遡及的にウェイクワードを検出することができない結果となる予測にペナルティを課す；および／または

ウェイクワードの信頼度の高い認識および／またはユーザーの音声コマンドの正しい認識につながる予測を強化する（reinforcing）。

【0544】

図58は、図57のモジュール5701の例示的实施形態の素子のブロック図である。図58において、ラベル付けされた要素は、以下の通りである：

図57のシステムの要素（図2および図3で同じようにラベル付けされている）；

5804：少なくとも1つの特定の種類の音声コマンド5709を認識し、（音声コマンド5709が特定の認識された種類のものであることを認識することに応答して）モジュール5803に対して指示を呈するように結合および構成されるモジュール；

5803：出力信号5701Aおよび5701B（または、いくつかの実装では、信号5701Aまたは信号5701Bのうちの1つのみ）を生成するように結合および構成されるモジュール。

5709：話者からの音声コマンド

【0545】

図58の実施形態では、「フォロー・ミー」モジュール5701は、以下のように動作するように構成される。話者からの音声コマンド5709（たとえば、サブシステム5706が電話通話を実装している間に発声された、「アマンダ、通話をここに移してくれ」）に応答して、レンダラー5707および／またはサブシステム5704がしかるべく使うための、変更されたセットのラウドスピーカー（出力5701Aによって示される）および／またはマイクロフォン（出力5701Bによって示される）を決定する。

【0546】

図58のように実装されるモジュール5701では、ユーザー位置特定モジュール5702またはサブシステム5704（両方とも図57に示される）は、話者の直接ローカル音声からコマンドを認識する単純なコマンドおよび制御モジュールであってもよく、またはそれを含んでいてもよい（すなわち、サブシステム5704からモジュール5702に提供されるマイク

10

20

30

40

50

ロフォン信号5704Aがそのようなローカル音声を示す、またはコマンド5709が、モジュール5702およびモジュール5701に提供される)。たとえば、図57の前処理サブシステム5704は、音声コマンド（マイクロフォン5705のうち一つまたは複数のものの出力によって示される）を認識し、モジュール5702およびモジュール5701に出力5709（そのようなコマンドを示す）を提供するように結合され構成される単純なコマンドおよび制御モジュールを含んでいてもよい。

【0547】

モジュール5701の図58の実装の例では、モジュール5701は、下記によって、話者からの音声コマンド5709（たとえば、「通話をここに移す」）に応答するように構成される：

ゾーンマッピングの結果としての話者の位置（入力5702Aによって示される）を知って、現在の話者の音響ゾーン情報（出力5701Aによって示される）に従ってレンダラー5707に命令することにより、レンダラーは、そのレンダリング構成を変更して、話者の現在の音響ゾーンのために最良のスピーカー（単数または複数）を使用することができる；および／または

ゾーンマッピングの結果としての話者の位置（入力5702Aによって示される）を知って、前処理モジュール5704に、現在の話者音響ゾーン情報（出力5701Bによって示される）に従って、最良のマイクロフォン（単数または複数）の出力のみを使用するように命令する。

【0548】

モジュール5701の図58の実装の例において、モジュール5701は、以下のように動作するように構成される：

1. 音声コマンドを待つ（5709）；
2. 音声コマンド5709を受領すると、（モジュール5804において）受領されたコマンド5709が所定の特定のタイプ（たとえば、「[活動]をここに移す」または「フォロー・ミー」のうちの1つ；ここで「[活動]」は、システムによって（たとえば、サブシステム5706によって）現在実装されている活動のいずれかを意味する）であるかどうかを判断する；
3. 音声コマンドが特定のタイプのものでない場合、音声コマンドを無視する（それにより、あたかも無視されている音声コマンドが受領されなかったかのように、出力信号5701Aおよび／または出力信号5701Bが、モジュール5803によって生成される）。
4. 音声コマンドが特定のタイプである場合、（モジュール5803内で）出力信号5701Aおよび／または出力信号5701Bを生成し、現在の音響ゾーン（ユーザー位置モジュール5702によって検出され、入力5702Aによって示される）に従って、それらの処理を変更するようにシステムの他の要素に命令する。

【0549】

図59は、図57のモジュール5701の別の例示的实施形態（図59では5900とラベル付けされている）およびその動作のブロック図である。図59において、ラベル付けされた要素は、以下の通りである：

5900: 「フォロー・ミー」モジュール；

図57のシステムの要素（図2および図4で同じようにラベル付けされている）；

モジュール5900の要素5803および5804（図58のモジュール5701の対応する要素と同様にラベル付けされる）；

5801: 話者（たとえば、ユーザー）の過去の経験からの学習された選好を示すデータのデータベース。データベース5801は、非一時的な仕方でデータを格納するメモリとして実装されてもよい；

5801A: 話者の過去の経験からの学習された選好に関するデータベース5801からの情報（データ）；

5802: 学習モジュールであって、入力5709および／または5706Aのうちの一つまたは複数、および／または出力5701Aおよび5701Bのうちの一つまたは両方（モジュール5803によって生成される）に応答してデータベース5801を更新するように結合および構

10

20

30

40

50

成される；

5802A: 話者の選好に関する更新された情報（データ）（モジュール5802によって生成され、記憶のためにデータベース5801に提供される）；

5806: 決定された話者位置における信頼度を評価するように結合され構成されたモジュール；

5807: 決定された話者位置が新しい位置であるかどうかを評価するように結合および構成されたモジュール、および

5808: ユーザーの確認（たとえば、ユーザー位置の確認）を要求するように結合され構成されたモジュール。

【0550】

モジュール5900は、話者の過去の経験に基づいて使用するために最良のラウドスピーカーおよびマイクロフォンについて自動的に決定するように構成されているという点で、図59のフォロー・ミー・モジュール5900は、図58のフォロー・ミー・モジュール5701の例示的实施形態に対する拡張を実装する。

【0551】

図57のモジュール5701が図59のモジュール5900として実装されると、図57の前処理サブシステム5704は、音声コマンド（マイクロフォン5705の一つまたは複数の出力によって示される）を認識し、モジュール5702およびモジュール5900の両方に出力5709（認識されたコマンドを示す）を提供するように結合され、構成される単純なコマンドおよび制御モジュールを含んでもよい。より一般的には、ユーザー位置特定モジュール5702またはサブシステム5704（両方とも図57に示される）は、話者の直接ローカル音声からコマンドを認識するように構成されるコマンドおよび制御モジュールであってもよく、またはそれを実装してもよく（たとえば、サブシステム5704からモジュール5702に提供されるマイクロフォン信号5704Aは、そのようなローカル音声を示す、または認識された音声コマンド5709が、サブシステム5704からモジュール5702およびモジュール5900に提供される）、モジュール5702は、認識されたコマンドを使用して、話者の位置を自動的に検出するように構成される。

【0552】

図59の実施形態では、ゾーンマップ5703と共にモジュール5702は、音響ゾーン・マッパーを実装してもよい（モジュール5702は、ゾーンマップ5703と共に動作するように結合され、構成されてもよく、またはゾーンマップ5703と一体化されてもよい）。いくつかの実装では、ゾーン・マッパーは、Bluetooth装置または他の無線周波数ビーコンの出力を使用して、ゾーン内の話者の位置を決定することができる。いくつかの実装では、ゾーン・マッパーは、自身のシステム内に履歴情報を保持し、（図59のモジュール5900への提供、または図57のモジュール5701の別の実施形態への提供のために）出力5702Aを生成して、話者の位置における確率的信頼性を示してもよい。話者の位置が正しく決定された確率は、（モジュール5900の）モジュール5806によって使用されて、ラウドスピーカー・レンダラーの鋭敏さ（acuity）に影響を与えてもよい（たとえば、モジュール5900が、データ5801Aによって示される、話者がその位置から話していることの他の事例を見たことがあるために、モジュール5806が、話者位置について十分に確信している場合、出力5701Aにより、レンダラー5707が、関連するオーディオを、よりフォーカスした仕方ではレンダリングするようにする）。

【0553】

逆に、モジュール5900が、話者が特定の位置において以前に位置特定されたことを認識せず、モジュール5806が、話者位置における不十分な信頼度（たとえば、所定の閾値を下回る信頼度）を有する場合、モジュール5806は、出力5701Aを生成させ、レンダラー5707が、より一般的な近傍において知覚されるように、関連するオーディオをレンダリングするようにしてもよい。

【0554】

図59の実装では、話者からのコマンド5709は、モジュール5900に、出力5701Aおよ

10

20

30

40

50

び／または出力5701Bを生成させて、現在のスピーカーおよび／またはマイクロフォンの新しいセットを示し、よって、使用中の現在のスピーカーおよび／またはマイクロフォンをオーバーライドさせてもよい。これはたとえば図58の例示的实施形態のようなものである。音響ゾーン内の話者の位置（たとえば、入力5702Aによって示される）、話者が実際に決定されたゾーン内にいることの信頼度（モジュール5806によって決定される）、現在進行中の活動（すなわち、たとえば入力5706Aによって示される、図57のサブシステム5706によって実装される活動）、および過去の学習された経験（たとえば、データ5801Aによって示される）に依存して、モジュール5900は、決定された進行中の活動のために、現在使用されているラウドスピーカーおよび／またはマイクロフォンを変更する自動決定を行う。いくつかの実装において、システムが、そのような自動決定について十分に自信がない場合（たとえば、モジュール5806が、所定の閾値を超えない、決定された話者位置における信頼度を有する場合）、話者からの位置の確認の要求を発行してもよい（たとえば、モジュール5806は、モジュール5808をして、出力5701Aにより、要求を発行させることができる）。この要求は、話者に最も近いスピーカーからの音声プロンプトの形であってもよい（たとえば、「キッチンに移動したことに気づきました。ここで音楽を再生したいですか?」というプロンプト）。

【0555】

図59のモジュール5900は、レンダラー5707の構成およびサブシステム5704がどのマイクロフォンを使うかに関する自動決定を行うように構成され、これは、諸音響ゾーン内の話者の動き、および任意的には、過去の経験（データベース5801内のデータによって示される）に基づく。そうするために、モジュール5900は、上述のコマンドおよび制御モジュール（前処理サブシステム5704またはモジュール5702によって実装される）からの、話者の直接のローカル音声によって示されるコマンドを示す入力（たとえば、コマンド5709）、ならびに、話者の位置を示す情報（たとえば、モジュール5702によって生成される入力5702A）を考慮してもよい。

【0556】

モジュール5900によって決定が行われた後（すなわち、先に決定されたラウドスピーカーおよび／またはマイクロフォンのセットに変化を生じさせる出力5701Aおよび／または出力5701Bが生成される）、学習モジュール5802は、将来においてより良好に自動的に決定される結果を確実にするための努力として、データ5802Aをデータベース5801に格納することができ、ここで、データ5802Aは、決定が満足できるものであった（たとえば、話者が、決定を手動でオーバーライドしなかった）か、または不満足なものであった（たとえば、話者が、音声コマンドを発することによって決定を手動でオーバーライドした）かを示すことができる。

【0557】

より一般的には、出力5701Aおよび／または出力5701Bの生成（たとえば、更新）は、少なくとも1つの以前の活動（出力5701Aおよび／または5701Bの生成の前に、たとえば進行中のオーディオ活動の前に、生じた活動）から学習モジュール5802（および／またはある実施形態の別の学習モジュール）によって決定された、学習された経験（たとえば、ユーザーの学習された選好）を示すデータ（たとえば、データベース5801からのデータ）に回答して、進行中のオーディオ活動の時点で実行されてもよい。たとえば、学習された経験は、現在の進行中のオーディオ活動の間に存在するものと同じまたは類似の条件下で呈された以前のユーザーコマンドから決定されてもよく、出力5701Aおよび／または出力5701Bは、そのような学習された経験を示すデータ（たとえば、データベース5801からのデータ）に基づく確率的信頼度に従って更新されうる（たとえば、モジュール5900が、学習された経験に基づいて、ユーザーの好みについて十分に自信を持っている場合、更新された出力5701Aがレンダラー5707に、関連するオーディオをよりフォーカスされた仕方レンダリングさせるという意味で、ラウドスピーカー・レンダラー5707の鋭敏さに影響を与えるために）。

【0558】

学習モジュール5802は、同じ入力（モジュール5900に提供される）および／または特徴の各セットに応答して行われた（および／または各セットを有する）最も最近の正しい決定の単純なデータベースを実装してもよい。このデータベースへの入力は、現在のシステム活動（たとえば、入力5706Aによって示される）、現在の話者音響ゾーン（入力5702Aによって示される）、以前の話者音響ゾーン（これも入力5702Aによって示される）、および同じ状況における以前の決定が正しかったかどうかに関する指示（たとえば、音声コマンド5709によって示される）であってもよく、またはそれらを含んでもよい。代替的に、モジュール5802は、話者がシステムの状態を自動的に変更することを望む確率を有する状態マップを実装することができ、それぞれの過去の決定は、正しいものも正しくないものも、そのような確率マップに加えられる。代替的に、モジュール5802は、モジュール5900の入力の全部または一部に基づいて学習するニューラルネットワークとして実装されてもよく、その出力は、出力5701Aおよび5701Bを生成するために使用される（たとえば、ゾーン変更が要求されるか否かを、レンダラー5707および前処理モジュール5704に命令するために使用される）。

【0559】

図57システム（モジュール5701が図59のモジュール5900として実装される）によって実行される処理の例示的なフローは、以下の通りである：

1. 話者は音響ゾーン1（たとえば、図56Aの要素5607）におり、アンソニーとの電話通話を開始する；
2. ユーザー位置特定モジュール5702およびフォロワー・ミー・モジュール5900は、話者がゾーン1内にあることを知っており、モジュール5900は、出力5701Aおよび5701Bを生成して、前処理モジュール5704がそのゾーンのために最良のマイクロフォン（単数または複数）を使用するようにし、レンダラー5707がそのゾーンのために最良のスピーカー構成を使用するようにする；
3. 話者が音響ゾーン2（たとえば、図56Bの要素5612）に移動する；
4. ユーザー位置特定モジュール5702は、話者の音響ゾーンの変化を検出し、該変化を示すようにモジュール5900への入力5702Aを呈する；
5. モジュール5900は、過去の経験（すなわち、データベース5801が示すデータ）から、現在のような状況において、話者が、話者の移動に際して、電話通話を新しい音響ゾーンに移すように求めたことを想起する。短時間の後、通話に移されるべきであるという自信が、設定した閾値を超え（モジュール5806によって判別される）、モジュール5900は、前処理サブシステム5704に、マイクロフォン構成を新しい音響ゾーンに変更するように命令し、また、レンダラー5707に、そのラウドスピーカー構成を調整して、新しい音響ゾーンについて最良の体験を提供するように命令する。
6. 話者は、音声コマンド5709を発声することによって自動決定をオーバーライドすることをせず（よって、モジュール5804は、学習モジュール5802およびモジュール5803に対するそのようなオーバーライドを示さない）、学習モジュール5802は、データ5802Aをデータベース5801に格納させ、モジュール5900が、この場合、正しい決定をし、同様の将来のケースについてそのような決定を強化する。

【0560】

図60は、別の例示的な実施形態のブロック図である。図60では、ラベル付けされた要素は以下の通りである：

図57のシステムの要素（図57および図60において同一にラベル付けされている）；

6011: 活動マネージャ。これは、サブシステム5706およびモジュール5701に結合され、システムが実装される環境（たとえば、家庭）内外での話者の活動の知識を有する。活動マネージャ6011は、本明細書においてオーディオ・セッション・マネージャ（その一部の例は、本明細書においてCHASMと称される（たとえば、図2CのCHASM 208C、図2DのCHASM 208D、図3CのCHASM 307および図4のCHASM 401を参照））と称されるものの一例である；

6012: 活動マネージャ6011に結合されたスマートフォン（本明細書では話者と呼ばれ

10

20

30

40

50

ることもあるシステムのユーザーのスマートフォン)、および該スマートフォンに接続されたBluetoothヘッドセット;

5706B: サブシステム5706によって実装される現在進行中の活動(単数または複数)(および/または、システムが実装される環境を超えた話者の活動)に関する情報(データ)。これは、活動マネージャ6011および/またはサブシステム5706によって生成され、モジュール5701への入力として提供される。

【0561】

図60のシステムでは、「フォロー・ミー」モジュール5701の出力5701Aおよび5701Bは、活動マネージャ6011、ならびにレンダラー5707および前処理サブシステム5704に対する命令であり、これらのそれぞれに、話者の現在の音響ゾーン(たとえば、話者が位置すると判断される新しい音響ゾーン)に従って処理を適応させる。

【0562】

図60のシステムでは、モジュール5701は、入力5706B(およびモジュール5701に提供される他の入力)に回答して、出力5701Aおよび/または出力5701Bを生成するように構成される。モジュール5701の出力5701Aは、レンダラー5707(および/または活動マネージャ6011)に、話者の現在の(たとえば、新たに決定された)音響ゾーンに従って処理を適応させるように指示する。モジュール5701の出力5701Bは、前処理サブシステム5704(および/または活動マネージャ6011)に、話者の現在の(たとえば、新しく決定された)音響ゾーンに従って処理を適応させるように指示する。

【0563】

図60システムによって実装される処理の例示的なフローは、要素6012が家屋内または家屋外のいずれでも動作可能であるほかはシステムが家屋内で実装され、モジュール5701が図59のモジュール5900のように実装されると想定する。例示的なフローは、以下の通りである:

1. 話者は散歩のために屋外に出ており、スマートフォン要素6012でアンソニーからの電話呼び出しを受ける;
2. 話者は、通話の途中で、家に入り、音響ゾーン1(たとえば、図56Aの要素5607)に入り、要素6012のブルートゥース(登録商標)ヘッドセットをオフにする;
3. ユーザー位置特定モジュール5702およびモジュール5701は、話者が音響ゾーン1に入ったことを検出し、モジュール5701は、(入力5706Bから)話者が(サブシステム5706によって実装されている)電話通話中であること、および要素6012のBluetoothヘッドセットがオフにされていることを知る;
4. モジュール5701は、過去の経験から、話者が、現在と同様の状況で、通話を新しい音響ゾーンに移すように求めたことを記憶している。短時間の後、通話が移動されるべきであるという確信が閾値を超えて高まり、モジュール5701は、活動マネージャ6011に対して、(適切な出力5701Aおよび/または5701Bを呈することによって)通話がスマートフォン要素6012から家庭内に実装されている図60のシステムの装置に移されるべきであることを命令し;モジュール5701は、(適切な出力5701Bを呈することによって)前処理サブシステム5704に対して、マイクロフォン構成を新しい音響ゾーンに変更するように命令し;モジュール5701は、(適切な出力5701Aを主張することによって)レンダラー5707に対しても、そのラウドスピーカー構成を調整して、新しい音響ゾーンについて最良の体験を提供するように命令する。
5. 話者は、音声コマンドを発声することによって自動決定(モジュール5701によってなされた)をオーバーライドせず、モジュール5701の学習モジュール(5802)は、モジュール5701がこの場合に正しい決定をしたことを示すデータを記憶し、これは、同様の将来の場合のためにこのような決定を補強する際に使用される。

【0564】

他の実施形態は、以下を含みうる:

環境内に複数のスマート・オーディオ装置を含むシステムを制御する方法であって、前記システムは、一つまたは複数のマイクロフォンのセット(たとえば、各マイクロフォン

10

20

30

40

50

は、前記環境内の前記スマート・オーディオ装置のうちの少なくとも1つに含まれる、またはそれと通信するように構成される）と、一つまたは複数のラウドスピーカーのセットとを含み、前記環境は、複数のユーザーゾーンを含み、前記方法は：少なくとも部分的には前記マイクロフォンの出力信号から、前記環境におけるユーザーの位置の推定値を決定するステップを含み、前記推定値は、前記ユーザーゾーンのうちのどれにユーザーが位置しているかを示す、方法；

複数のスマート・オーディオ装置にまたがるオーディオ・セッションを管理する方法であって、進行中のオーディオ活動のための現在使用されているマイクロフォンおよびラウドスピーカーのセットを、ユーザーの要求またはユーザーによって発せられたその他の音に応答して変更するステップを含む、方法；

複数のスマート・オーディオ装置にまたがるオーディオ・セッションを管理する方法であって、少なくとも1つの以前の経験に基づいて（たとえば、ユーザーの過去の経験からの少なくとも1つの学習された選好に基づいて）進行中のオーディオ活動のための、現在使用されているマイクロフォンおよびラウドスピーカーのセットを変更するステップを含む、方法。

【0565】

いくつかの実施形態の諸側面は、以下の箇条書き実施形態（enumerated example embodiment、EEE）を含む：

EEE1.

オーディオ信号ルーティングのためのより低いレベルの制御を発行することができる単一の階層的システムを通じて集合的に動作する複数のオーディオ装置（たとえば、スマートオーディオ装置）を含む装置の集合的システム（コンステレーション）においてオーディオを制御する方法であって：

a. 制御可能な装置のコンステレーションへの適用のための、単一のインターフェース・ポイントがあり；

b. この単一の接触点との対話は、装置に関する具体的詳細には関わらず；

c. 前記対話は、以下を含む複数の明示的または暗示的パラメータを含む

i. ソース、

ii. 宛先、

iii. 優先度；

d. システムは、これらのルート（たとえば、CHASMへの要求）のそれぞれについて固有の永続的識別子を追跡し、任意的に、それから複数の属性（たとえば、連続的な性質の属性）を照会することもでき；

e. システムは、オーディオ装置の制御を実行するために、利用可能な情報（たとえば、少なくともいくつかの、または任意の、またはすべての利用可能な現在および過去の情報）を連続的に利用する。

EEE2.

前記複数のパラメータが、モード（たとえば、同期）をも含む、EEE1に記載の方法。

EEE3.

前記複数のパラメータが、品質（たとえば、オーディオを送達することの目標、たとえば、了解度）も含む、先行する請求項のいずれかに記載の方法。

EEE4.

前記複数のパラメータが、固執度（insistence）（たとえば、どのくらい確認されていると知りたいか（how much you want to know it is confirmed））をも含む、先行する請求項のいずれかに記載の方法。

EEE5.

前記複数の属性が、それ（たとえば、オーディオ）がどのくらいよく聞こえるか（たとえば進行中）（たとえば、信頼度）を含む、先行する請求項に記載の方法。

EEE6.

前記複数の属性は、対話があった度合い（確認応答）を含む、先行する請求項のいずれ

10

20

30

40

50

かに記載の方法。

EEE7.

前記複数のパラメータが可聴性を含む、先行する請求項のいずれかに記載の方法。

EEE8.

前記複数のパラメータが可聴性の欠如を含む、先行する請求項のいずれかに記載の方法。

EEE9.

前記複数のパラメータが了解度を含む、先行する請求項のいずれかに記載の方法。

EEE10.

前記複数のパラメータが、了解度の欠如（たとえば、マスキング、「静音円錐（cone of silence）」）を含む、先行する請求項のいずれかに記載の方法。

10

EEE11.

前記複数のパラメータが、空間的忠実度（たとえば、定位性能）を含む、先行する請求項のいずれかに記載の方法。

EEE12.

前記複数のパラメータが一貫性を含む、先行する請求項のいずれかに記載の方法。

EEE13.

前記複数のパラメータが、忠実度（たとえば、符号化歪みの欠如）を含む、先行する請求項のいずれかに記載の方法。

EEE14.

ルートが単一の宛先のみを有することができる（ユニキャスト）、先行する請求項のいずれかの方法を実装するように構成されたシステム。

20

EEE15.

請求項EEE1～EEE13のいずれかの方法を実装するように構成されたシステムであって、経路が複数の宛先を有してもよい（マルチキャスト）、システム。

EEE16.

複数のオーディオ装置を有するオーディオ環境のためのオーディオ・セッション管理方法であって、前記オーディオ・セッション管理方法は：

第1のアプリケーションを実装する第1の装置から、オーディオ・セッション・マネージャを実装する装置によって、第1のオーディオ・セッションのための第1のルートを開始するための第1のルート開始要求を受領する段階であって、前記第1のルート開始要求は第1のオーディオソースおよび第1のオーディオ環境宛先を示し、前記第1のオーディオ環境宛先は、前記オーディオ環境の少なくとも第1の領域に対応し、前記第1のオーディオ環境宛先は、オーディオ装置を示さない、段階と；

30

前記オーディオ・セッション・マネージャを実装している装置によって、前記第1のルート開始要求に対応する第1のルートを確立する段階であって、前記第1のルートを確立することは：

前記第1のオーディオ・セッションの第1のステージについて、前記オーディオ環境の前記第1のエリアにある少なくとも1つのオーディオ装置を判別し；

前記第1のオーディオ・セッションを開始またはスケジューリングすることを含む、段階とを含む、方法。

40

EEE17.

前記第1のルート開始要求は、第1のオーディオ・セッション優先度を含む、EEE16に記載のオーディオ・セッション管理方法。

EEE18.

EEE16またはEEE17のオーディオ・セッション管理方法であって、前記第1のルート開始要求は第1の接続性モードを含む、方法。

EEE19.

前記第1の接続性モードが、同期的接続性モード、トランザクション接続性モード、またはスケジュールされた接続性モードである、EEE18に記載のオーディオ・セッション管

50

理方法。

EEE20.

EEE16～EEE19のいずれか1つのオーディオ・セッション管理方法であって、前記第1のルート開始要求は、第1の人物を示し、少なくとも前記第1の人物から確認応答が要求されるかどうかの指示を含む、オーディオ・セッション管理方法。

EEE21.

EEE16～EEE20のいずれか1つのオーディオ・セッション管理方法であって、前記第1のルート開始要求は、第1のオーディオ・セッション目標を含む、方法。

EEE22.

前記第1のオーディオ・セッション目標は、了解度、オーディオ品質、空間的忠実度または非可聴性のうちの一つまたは複数を含む、EEE21に記載のオーディオ・セッション管理方法。

10

EEE23.

EEE16～EEE22のいずれか1つのオーディオ・セッション管理方法であって、前記第1のルートのための第1の永続的固有オーディオ・セッション識別子を決定し、前記第1の永続的固有オーディオ・セッション識別子を前記第1の装置に送信する段階をさらに含む、オーディオ・セッション管理方法。

EEE24.

EEE16～EEE23のいずれか1つのオーディオ・セッション管理方法であって、前記第1のルートを確立することは、前記環境内の少なくとも1つの装置に、前記第1のルートに対応する少なくとも第1のメディア・ストリームを確立させることを含み、前記第1のメディア・ストリームは第1のオーディオ信号を含む、オーディオ・セッション管理方法。

20

EEE25.

前記第1のオーディオ信号を第1のレンダリングされたオーディオ信号にレンダリングされるようにするレンダリング・プロセスをさらに含む、EEE24に記載のオーディオ・セッション管理方法。

EEE26.

EEE25のオーディオ・セッション管理方法であって、さらに：

前記オーディオ環境の前記第1の領域における複数のオーディオ装置の各オーディオ装置の第1の位置を自動的に決定する第1のラウドスピーカー自動位置特定プロセスを実行し、前記レンダリング・プロセスは、少なくとも部分的には各オーディオ装置の前記第1の位置に基づく、段階と；

30

各オーディオ装置の前記第1の位置を、前記第1のルートに関連付けられたデータ構造に記憶する段階とを含む、

オーディオ・セッション管理方法。

EEE27.

EEE25のオーディオ・セッション管理方法であって、さらに：

前記第1の領域にある少なくとも1つのオーディオ装置が変更された位置を有することを判別する段階と；

前記変更された位置を自動的に判別する第2のラウドスピーカー自動位置特定プロセスを実行する段階と；

40

少なくとも部分的には前記変更された位置に基づいて前記レンダリング・プロセスを更新する段階と；

前記変更された位置を、前記第1のルートに関連付けられた前記データ構造に格納する段階とを含む、

オーディオ・セッション管理方法。

EEE28.

EEE25のオーディオ・セッション管理方法であって、さらに：

少なくとも1つの追加的なオーディオ装置が前記第1の領域に移動されたことを判別する段階と；

50

前記追加的なオーディオ装置の追加的なオーディオ装置位置を自動的に決定する第2のラウドスピーカー自動位置特定プロセスを実行する段階と；

少なくとも部分的には前記追加的なオーディオ装置位置に基づいて、前記レンダリング・プロセスを更新する段階と；

前記追加的なオーディオ装置位置を前記第1のルートに関連付けられた前記データ構造に格納する段階とを含む、

オーディオ・セッション管理方法。

EEE29.

EEE16～EEE28のいずれか1つのオーディオ・セッション管理方法であって、前記第1のルート開始要求は、少なくとも第1の人物を、第1のルート・ソースまたは第1のルート宛先として示す、方法。

10

EEE30.

EEE16～EEE29のいずれか1つのオーディオ・セッション管理方法であって、前記第1のルート開始要求は、少なくとも第1のサービスを前記第1のオーディオソースとして示す、方法。

EEE31.

EEE16～EEE30のいずれか1つの方法を実行するように構成された装置。

EEE32.

EEE16～EEE30のいずれかのメソッドを実行するように構成されたシステム。

EEE33.

20

ソフトウェアがエンコードされた一つまたは複数の非一時的媒体であって、前記ソフトウェアは、EEE16～EEE30のうちいずれかの方法を実行するよう一つまたは複数の装置を制御するための命令を含む、媒体。

【0566】

いくつかの開示される実装は、開示された方法の一部または全部を実行するように構成された（たとえば、プログラムされた）システムまたは装置と、開示された方法の一部または全部またはそのステップを実装するためのコードを記憶している有形のコンピュータ読み取り可能媒体（たとえば、ディスク）とを含む。いくつかの開示されたシステムは、プログラム可能な汎用プロセッサ、デジタル信号プロセッサ、またはマイクロプロセッサであって、開示される方法の一部または全部またはそのステップの実装を含む、データに対する多様な操作のいずれかを実行するようにソフトウェアまたはファームウェアでプログラムされ、および／または他の仕方で構成されたものであってもよく、またはそれらを含むことができる。そのような汎用プロセッサは、入力装置、メモリ、および、それに呈されたデータに応答して開示された方法の一部または全部（またはそのステップ）を実行するようにプログラムされた（および／または他の仕方で構成された）処理サブシステムを含むコンピュータ・システムであってもよく、または、それを含んでいてもよい。

30

【0567】

いくつかの実施形態は、開示された方法の一部または全部の実行を含む、オーディオ信号に対して必要な処理を実行するように構成された（たとえば、プログラムされた、および他の方法で構成された）構成可能な（たとえば、プログラム可能な）デジタル信号プロセッサ（DSP）として実装されてもよい。代替的または追加的に、いくつかの実施形態（またはその要素）は、開示された方法の一部または全部を含む多様な動作のいずれかを実行するようにソフトウェアまたはファームウェアでプログラムされた、および／または他の仕方で構成された汎用プロセッサ（たとえば、パーソナルコンピュータ（PC）または他のコンピュータ・システムまたはマイクロプロセッサであって、入力装置およびメモリを含んでいてもよい）として実装されてもよい。代替的または追加的に、いくつかの実施形態の要素は、開示された方法の一部または全部を実行するように構成された（たとえば、プログラムされた）汎用プロセッサまたはDSPとして実装され、システムはまた、他の要素（たとえば、一つまたは複数のラウドスピーカーおよび／または一つまたは複数のマイクロフォン）を含んでいてもよい。開示された方法の一部または全部を実行するように構

40

50

成された汎用プロセッサは、いくつかの例では、入力装置（たとえば、マウスおよび／またはキーボード）、メモリ、およびディスプレイ装置に結合されてもよい。

【0568】

本開示のいくつかの実装は、開示される方法の一部または全部またはそのステップを実行するためのコード（たとえば実行するために実行可能なコード）を記憶しているコンピュータ読み取り可能媒体（たとえば、ディスクまたは他の有形記憶媒体）であってもよく、あるいはそれを含んでいてもよい。

【0569】

個別的な実施形態および用途が本明細書に記載されているが、本明細書に記載されている実施形態および用途の多くの変形が、本明細書に示され、記載され特許請求される素材の範囲から逸脱することなく可能であることは、当業者には明らかであろう。ある種の実装が示され説明されたが、本開示は、説明され示された特定の実施形態または説明された特定の方法に限定されないことが理解されるべきである。

10

20

30

40

50

【図面】

【図 1 A】

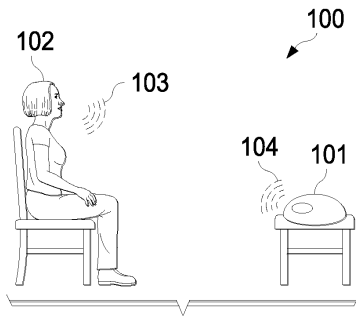


Figure 1A

【図 1 B】

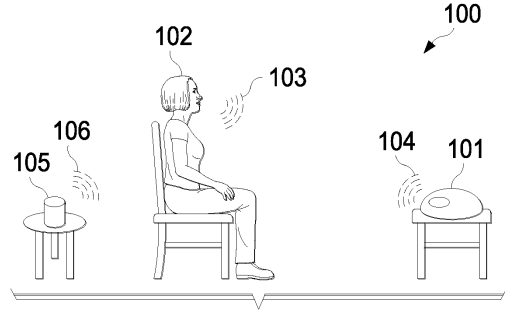


Figure 1B

【図 1 C】

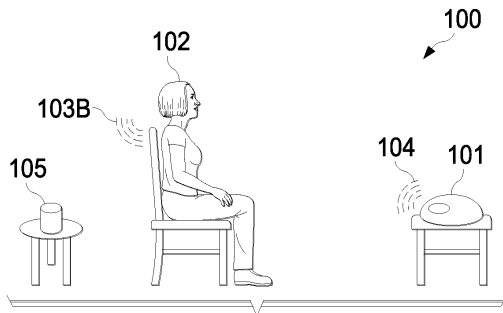


Figure 1C

【図 2 A】

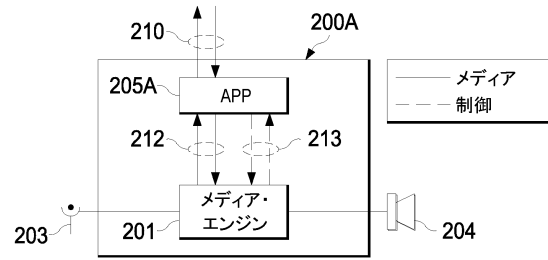


Figure 2A

10

20

30

40

50

【図 2 B】

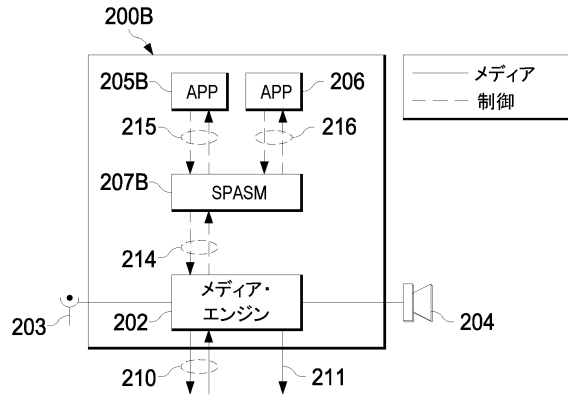


Figure 2B

【図 2 C】

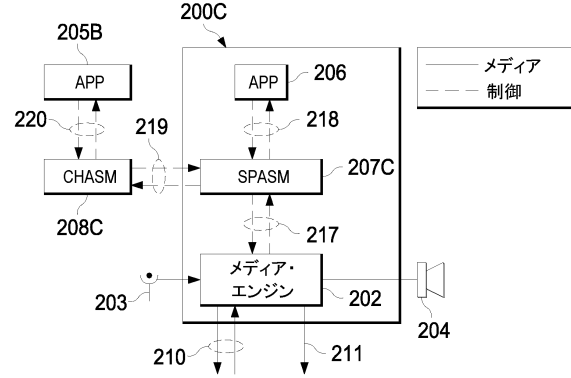


Figure 2C

【図 2 D】

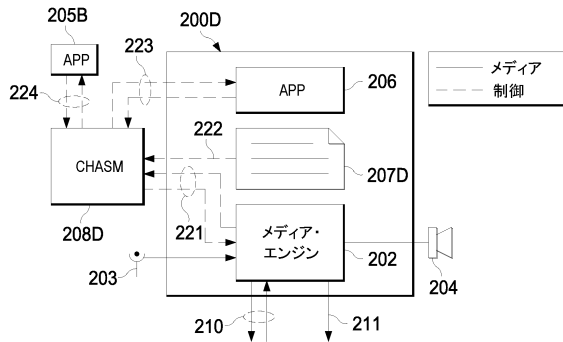


Figure 2D

【図 3 A】

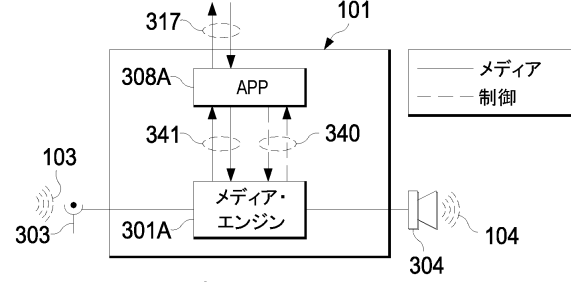


Figure 3A

10

20

30

40

50

【図 3 B】

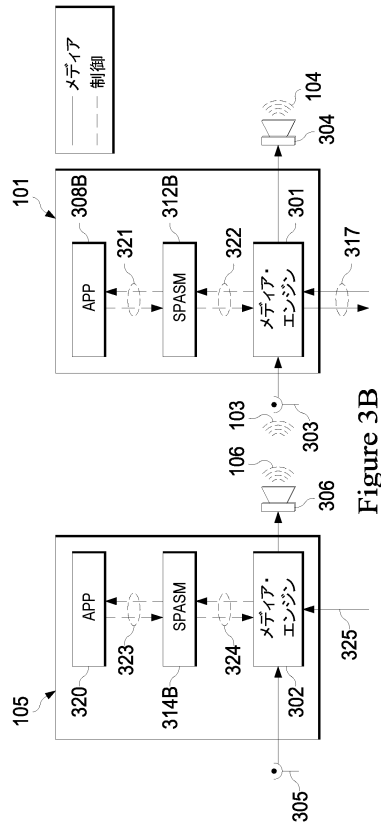


Figure 3B

【図 3 C】

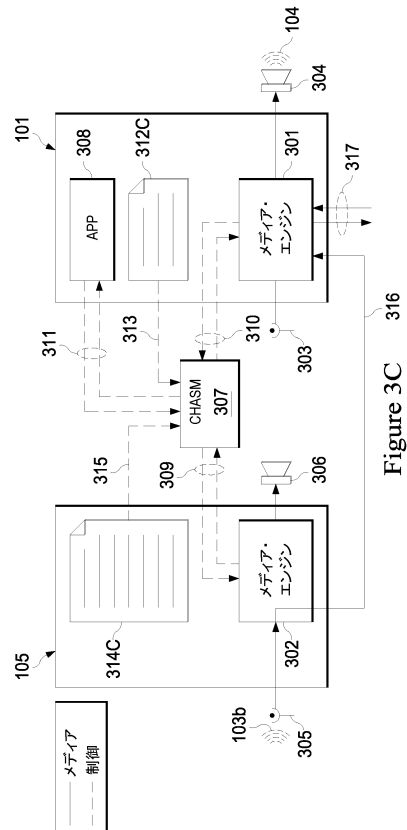


Figure 3C

【図 4】

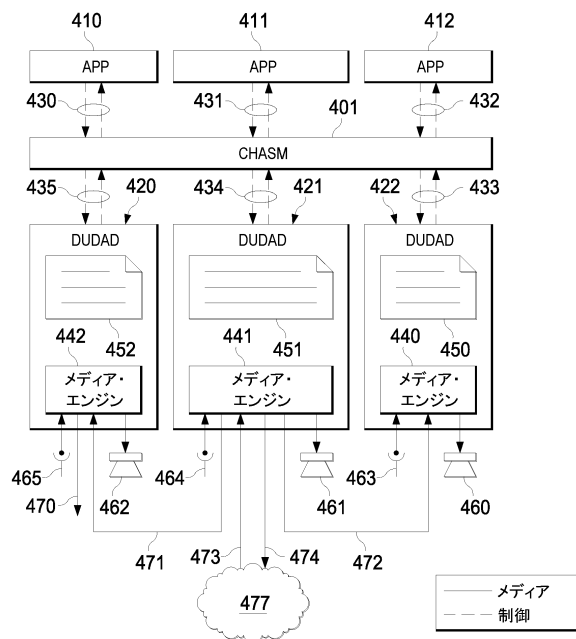


Figure 4

【図 5】

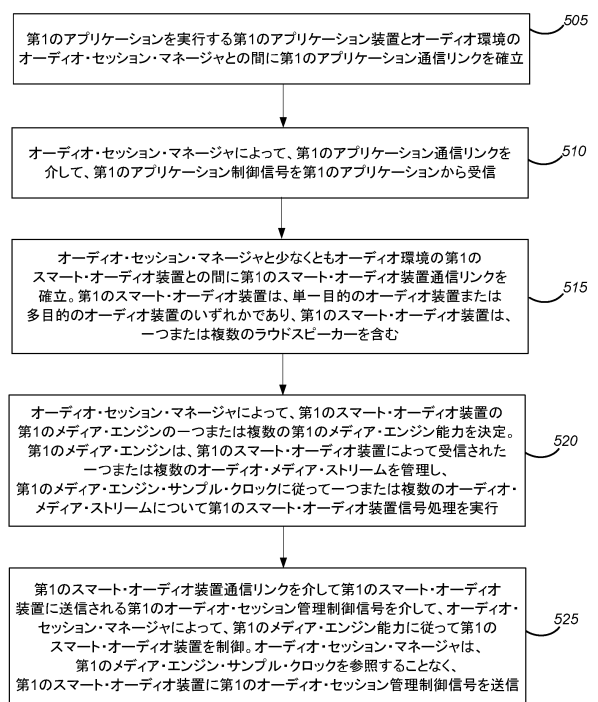


Figure 5

【図 6】

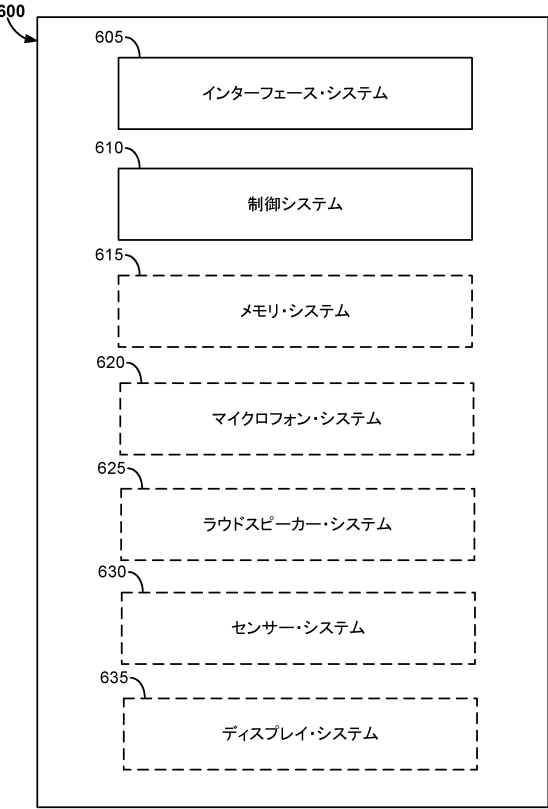


Figure 6

【図 7】

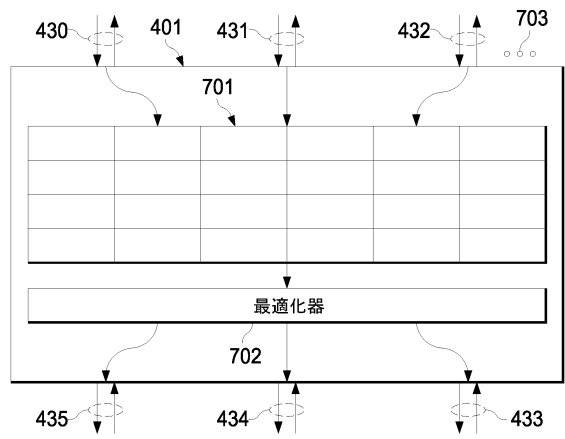


Figure 7

【図 8】

701

ID	APP	ソース	宛先	優先度	モード	ACK.	品質
50	410	SPOTIFY	アレックス	4	同期	NO	音楽
51	811	タイマー	アンガス	4	スケジュール	YES	可聴性
52	410	-	赤ん坊	2	同期	NO	非可聴性
53	411	電話	ジョージ	3	同期	NO	了解度
54	412	メッセージ	リチャード	3	トランザクション	YES	了解度
55	810	火災警報	全員	1	同期	NO	可聴性

801 802 803 804 805 806

Figure 8

【図 9 A】

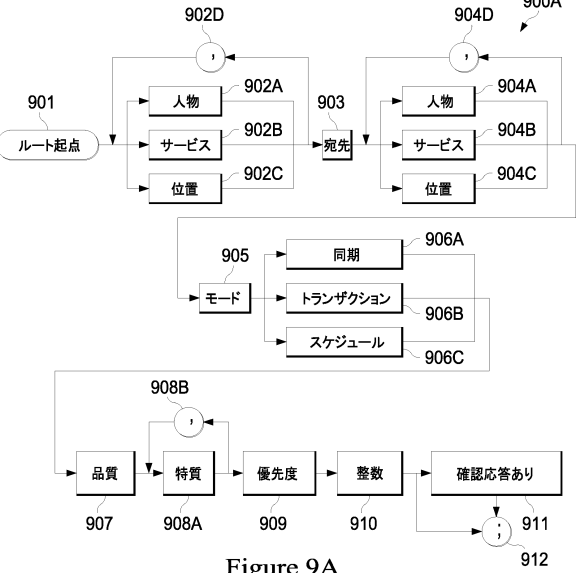


Figure 9A

10

20

30

40

50

【図 9 B】

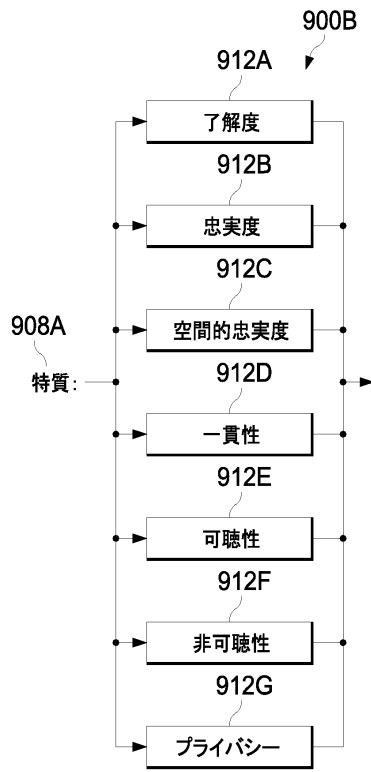


Figure 9B

【図 10】

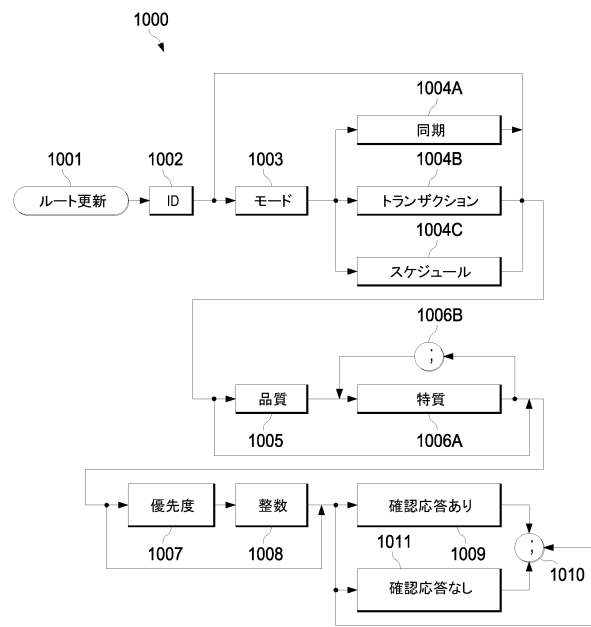


Figure 10

【図 11 A】

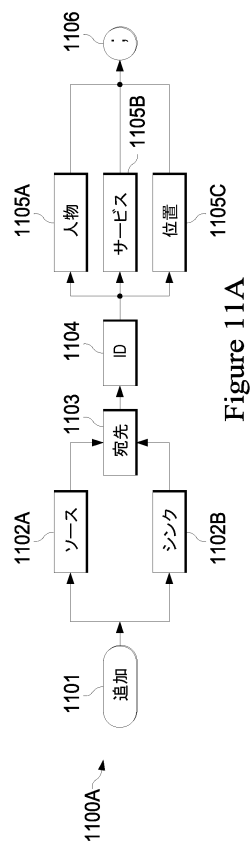


Figure 11A

【図 11 B】

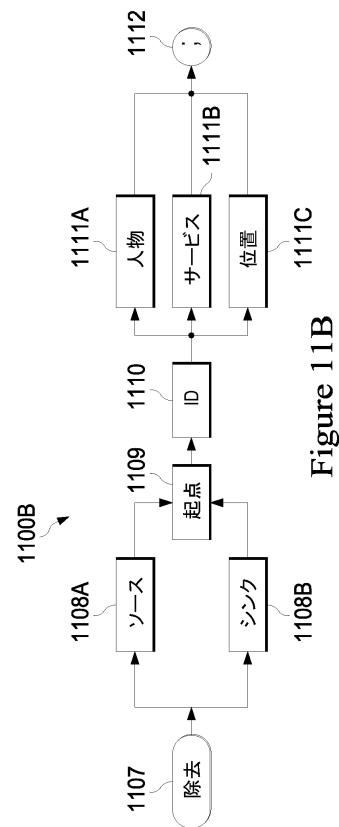


Figure 11B

10

20

30

40

50

【図 11 C】

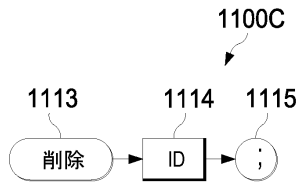


Figure 11C

【図 12】

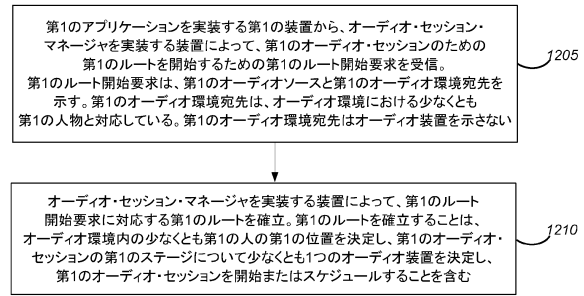


Figure 12

【図 13】

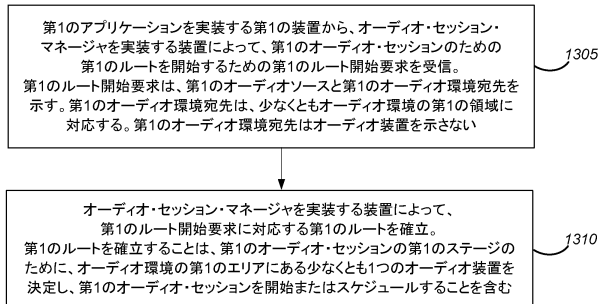


Figure 13

【図 14】

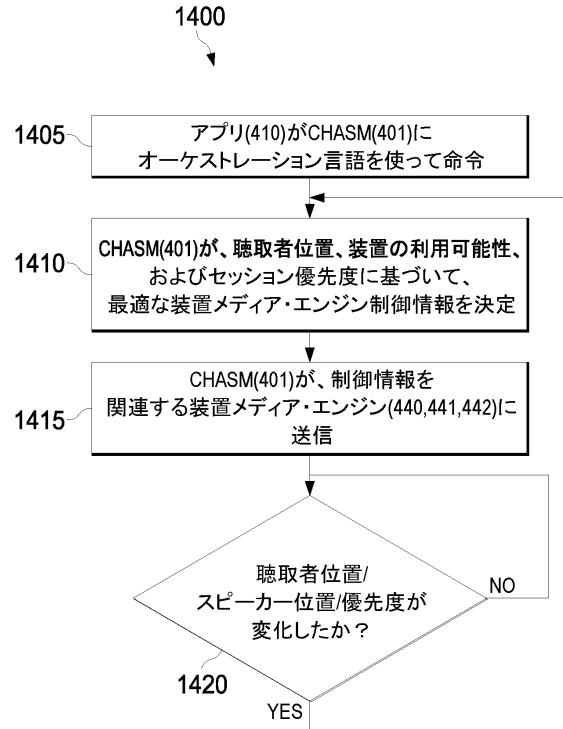


Figure 14

【図 15】

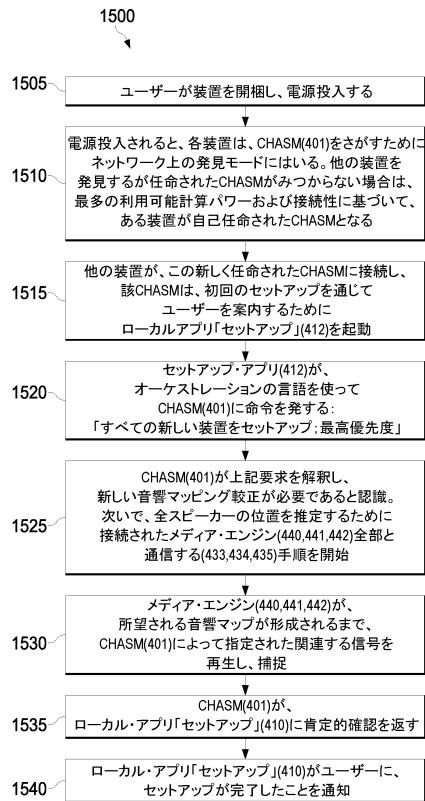


Figure 15

【図 16】

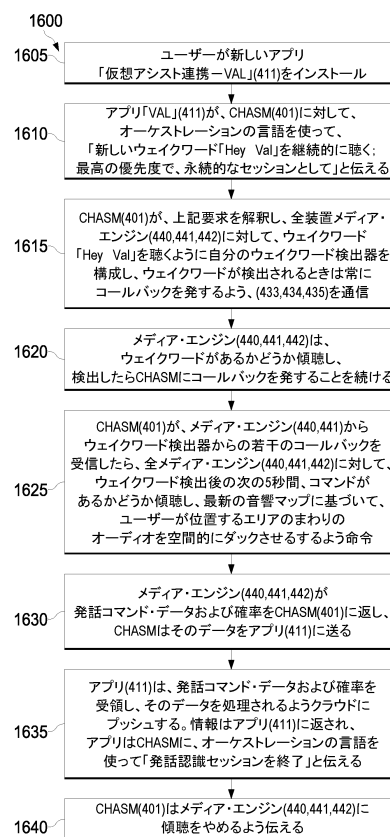


Figure 16

【図 17】

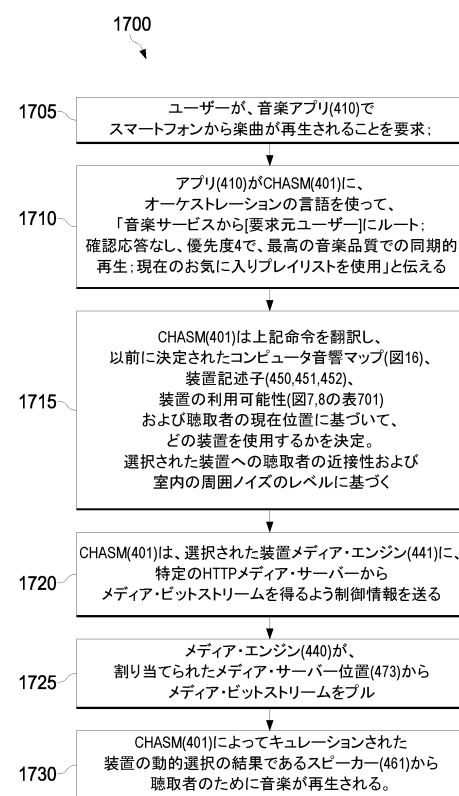


Figure 17

【図 18 A】

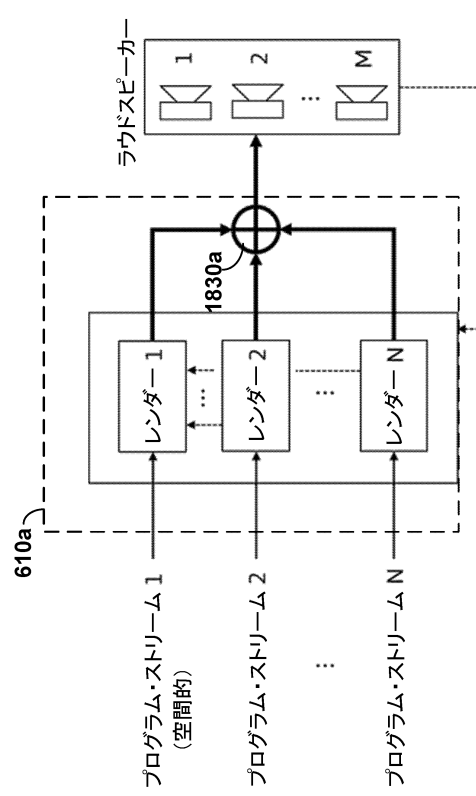


Figure 18A

10

20

30

40

50

【図 18 B】

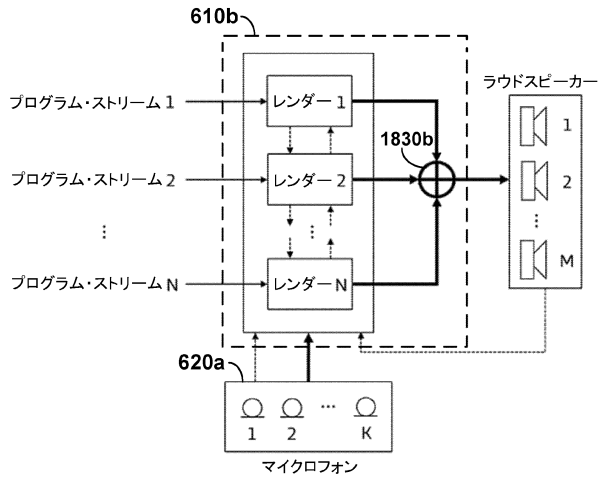


Figure 18B

【図 19】

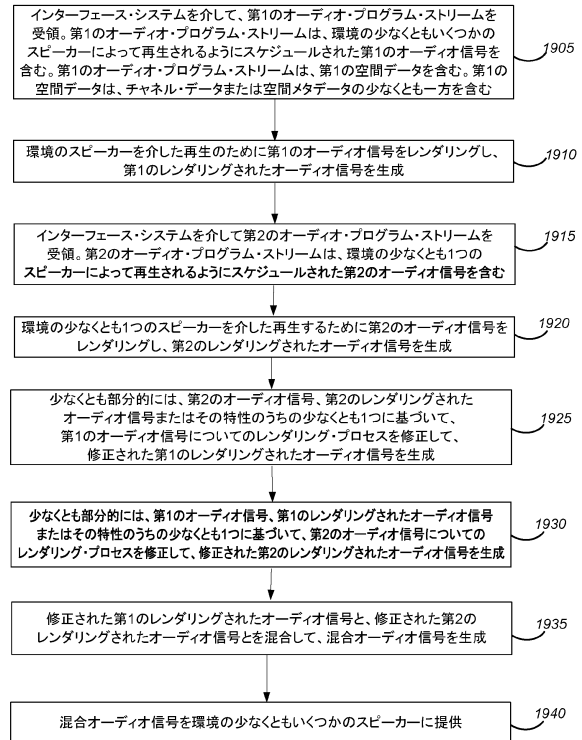


Figure 19

【図 20】

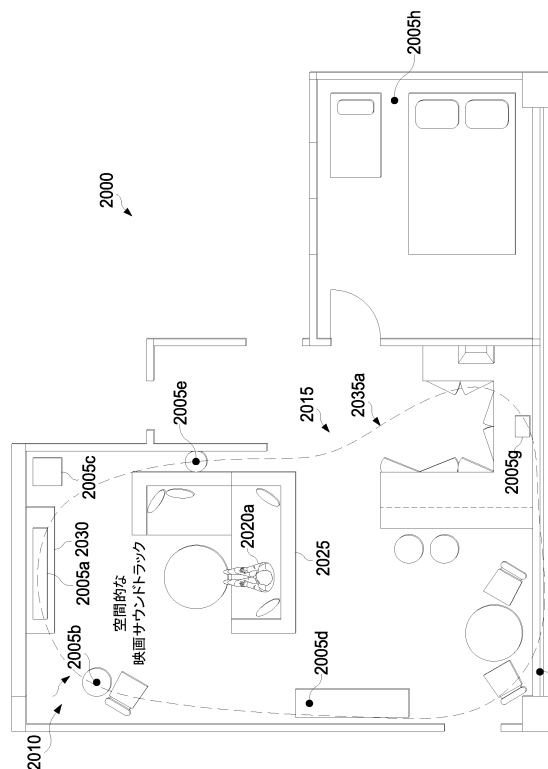


Figure 20

【図 21】

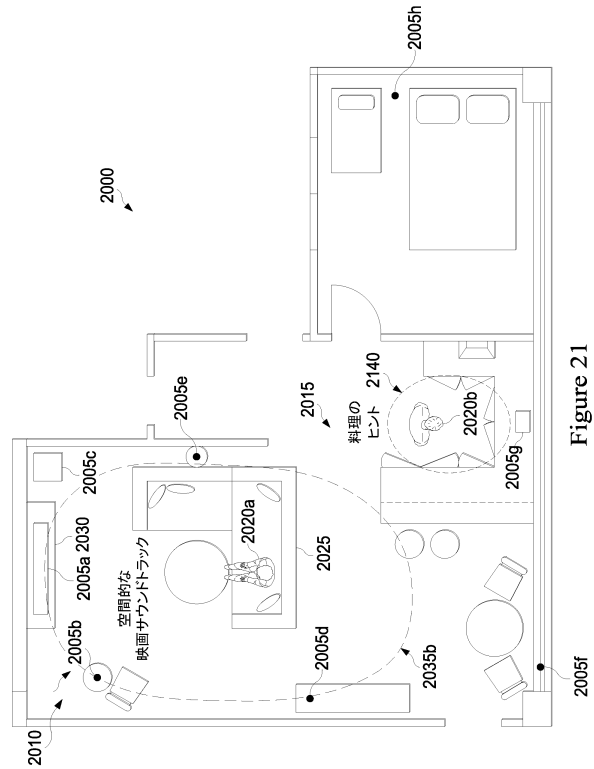


Figure 21

10

20

30

40

50

【図 2 2】

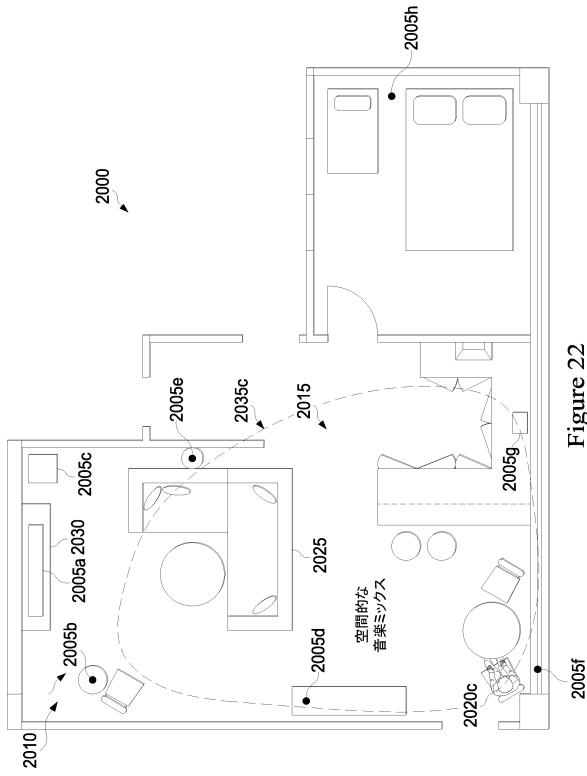


Figure 22

【図 2 3】

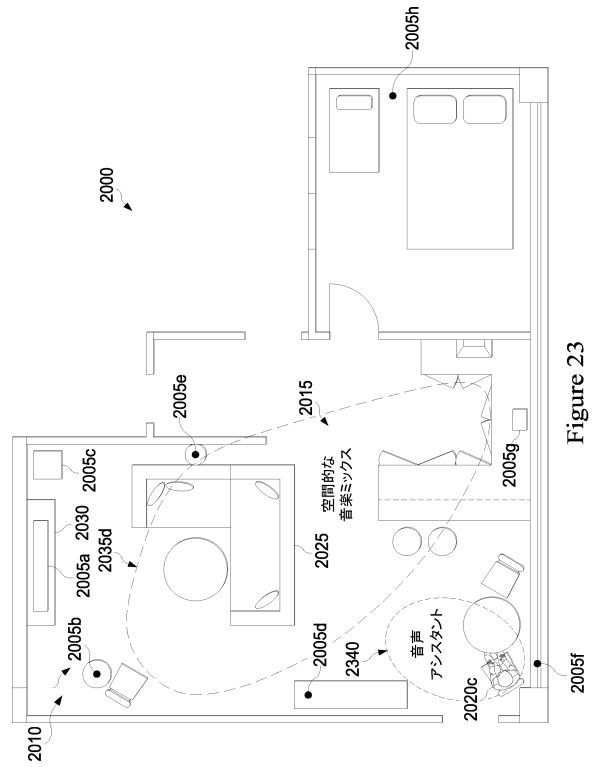


Figure 23

【図 2 4】

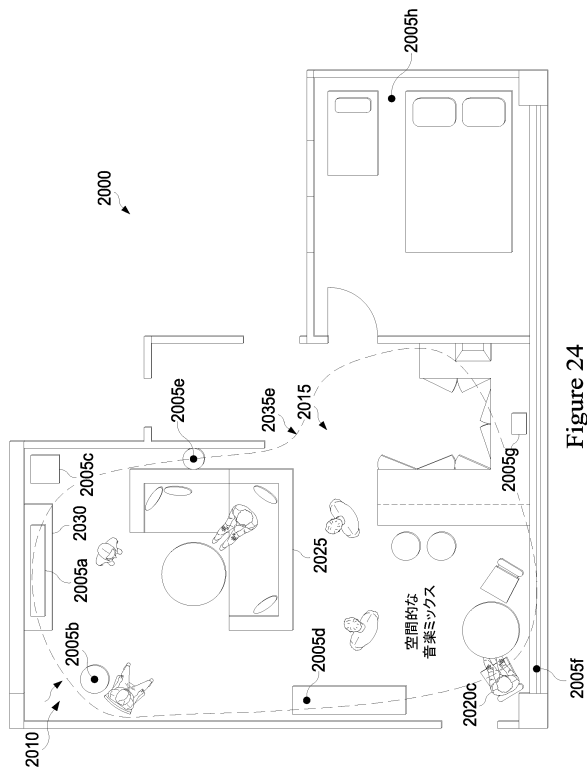


Figure 24

【図 2 5】

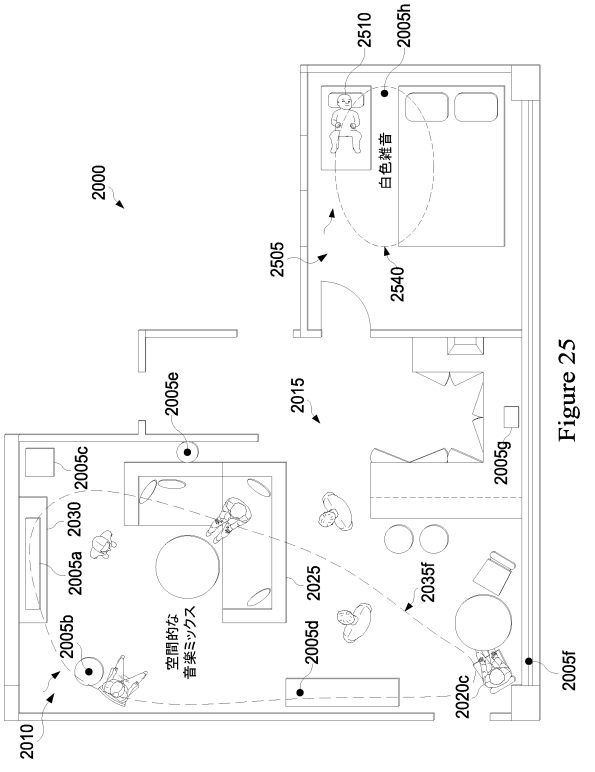


Figure 25

10

20

30

40

50

【図 26】

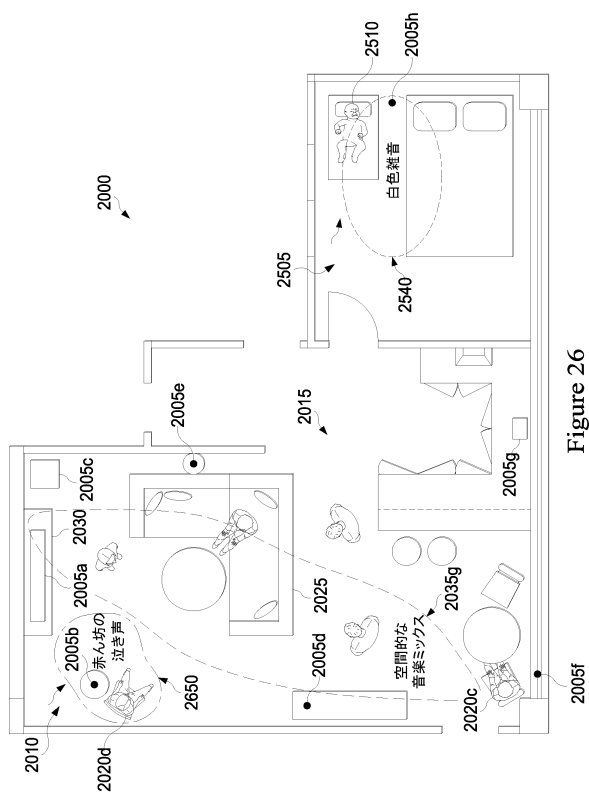


Figure 26

【圖 27】

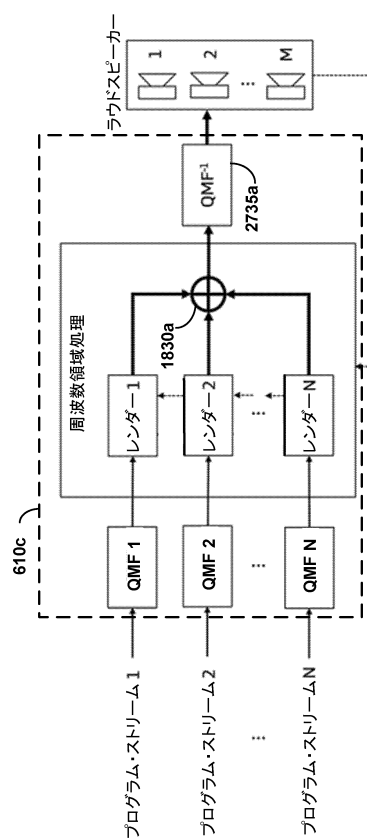


Figure 27

【図 28】

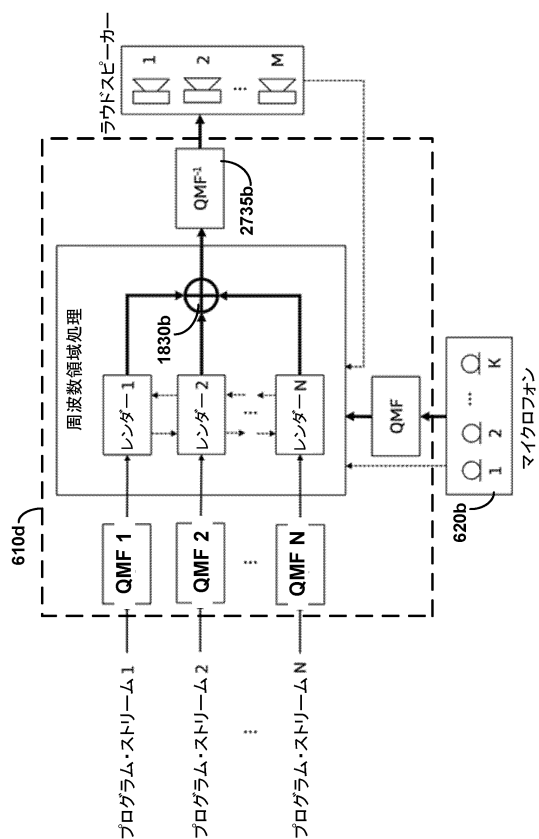


Figure 28

【図 29】

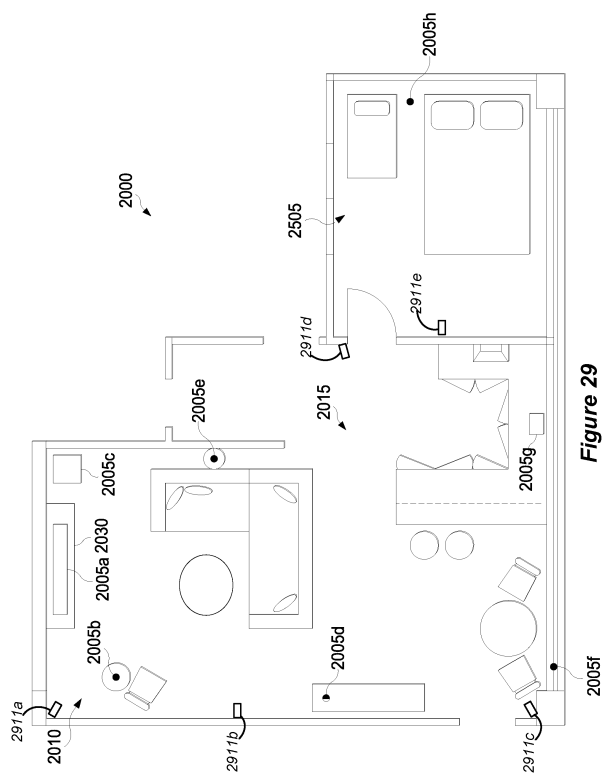
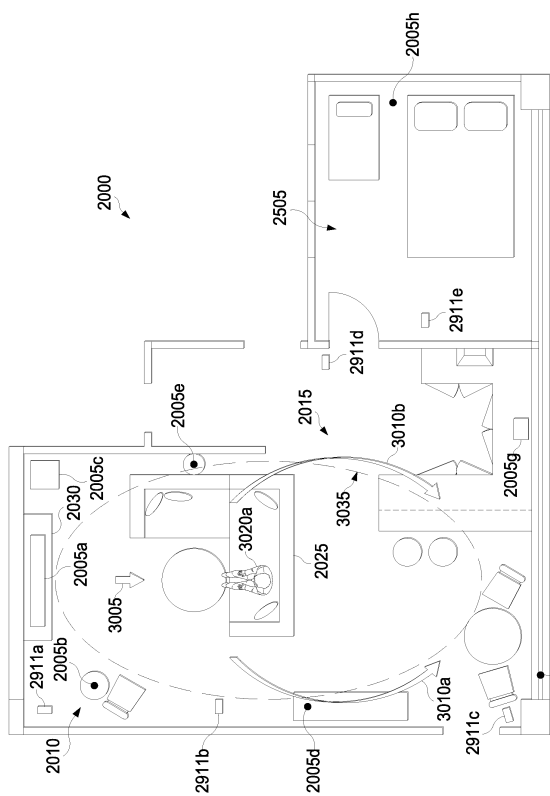
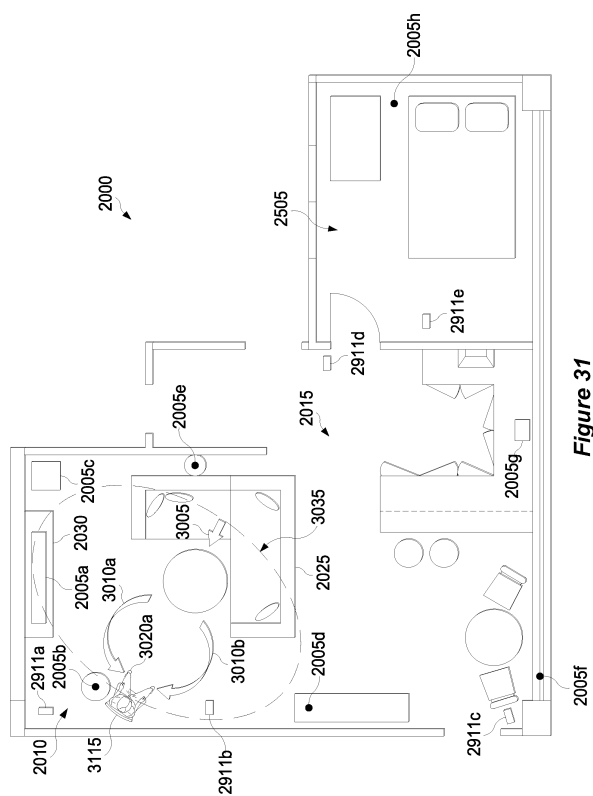


Figure 29

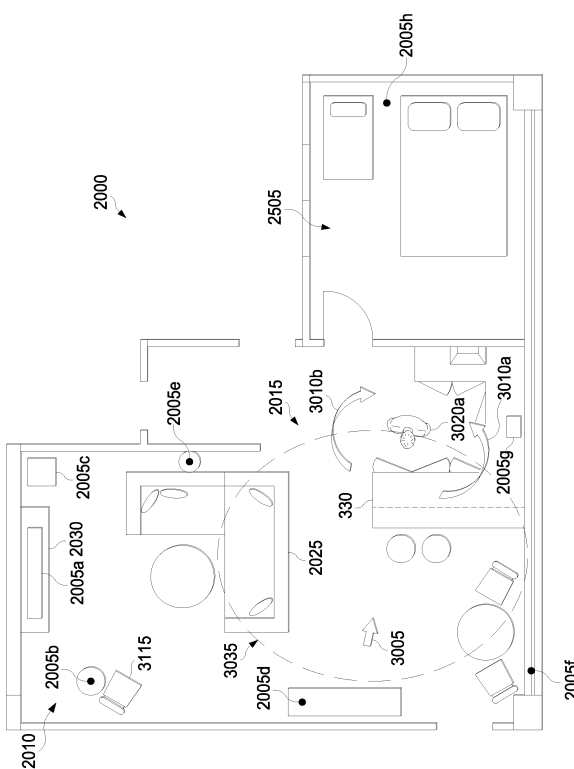
【図 30】



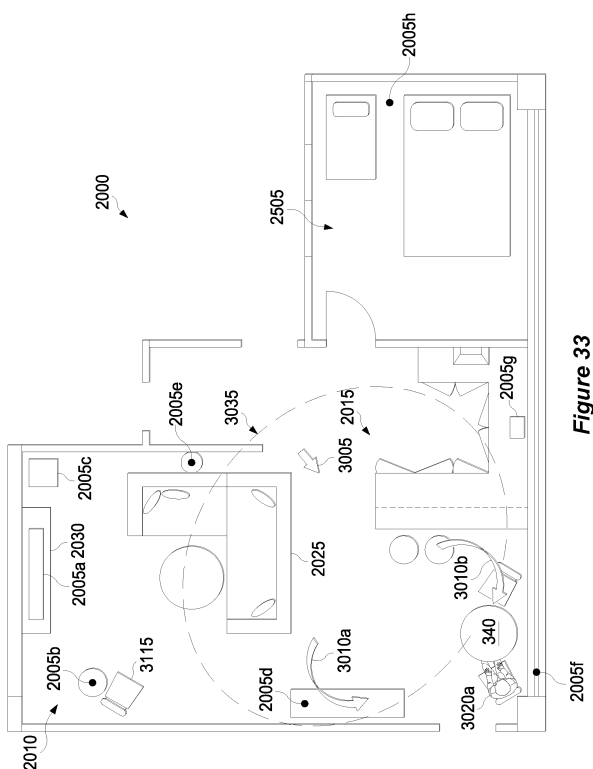
【図 3 1】



【図 3 2】



【図 3 3】



【図 3 4】

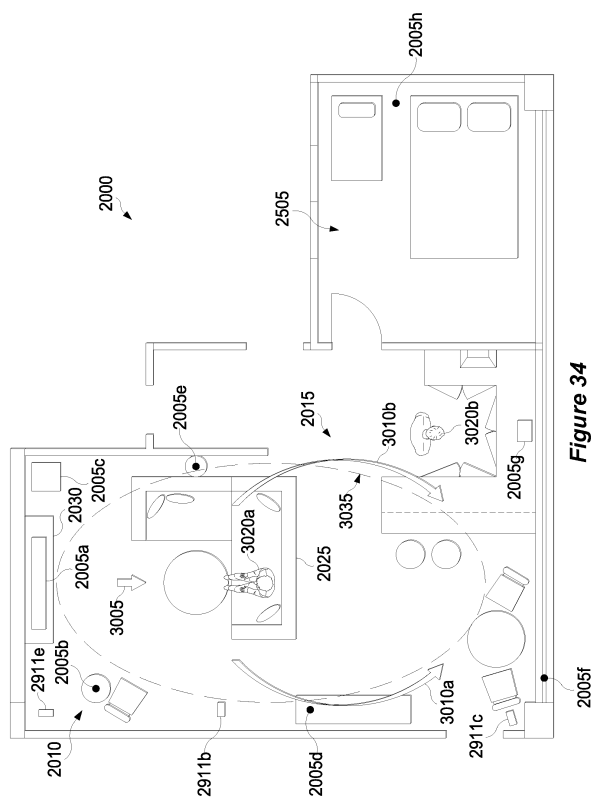


Figure 34

【図 3 5】

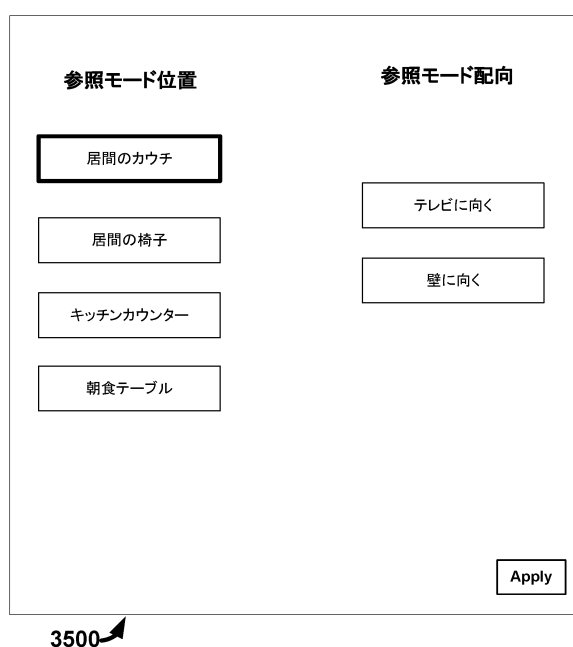


Figure 35

【図 3 6】

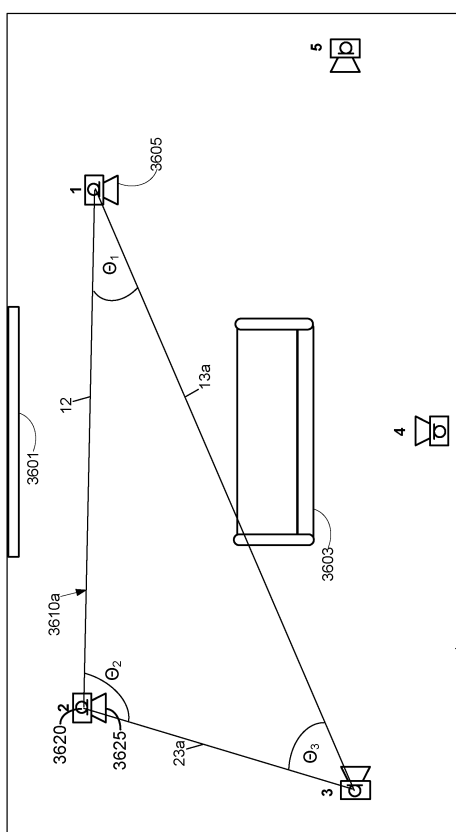


Figure 36

【図 3 7】

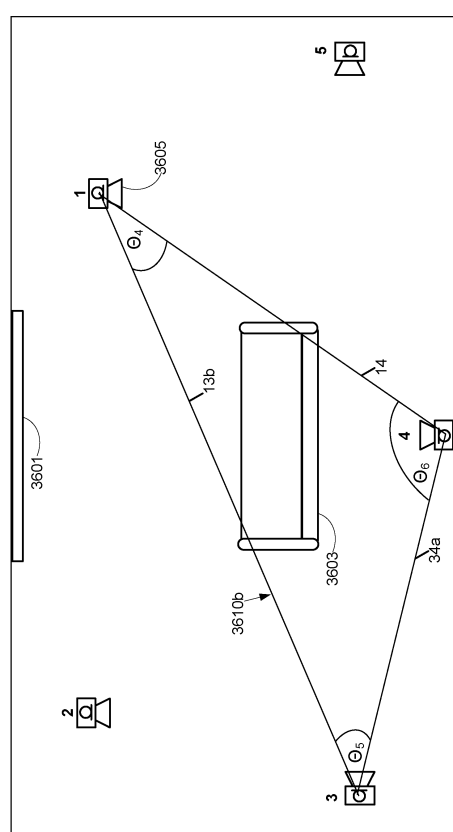


Figure 37

【図 38】

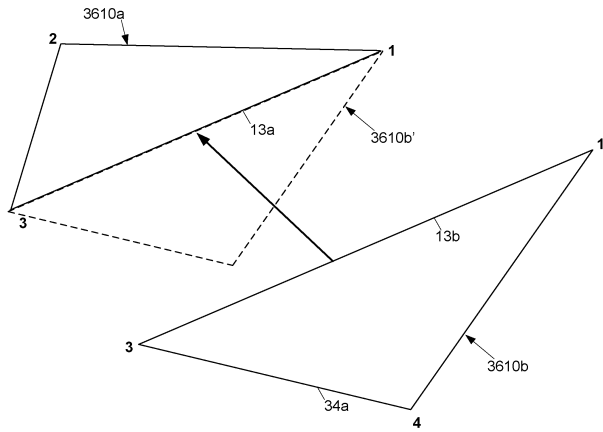


Figure 38

【図 39】

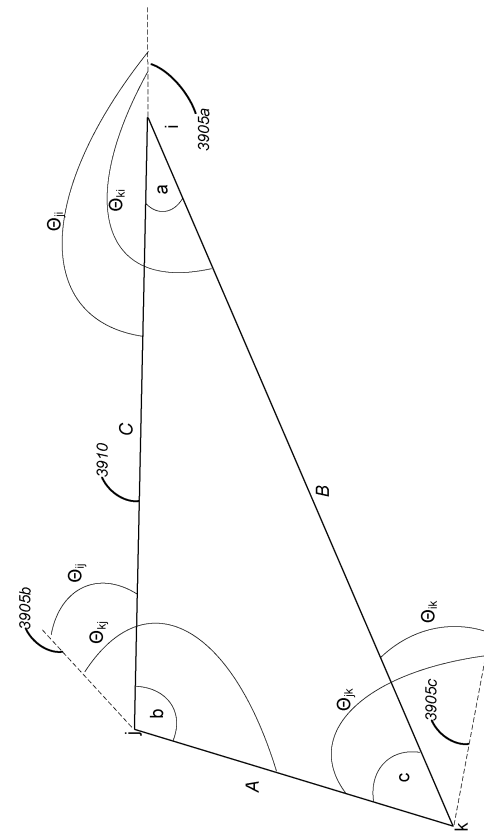
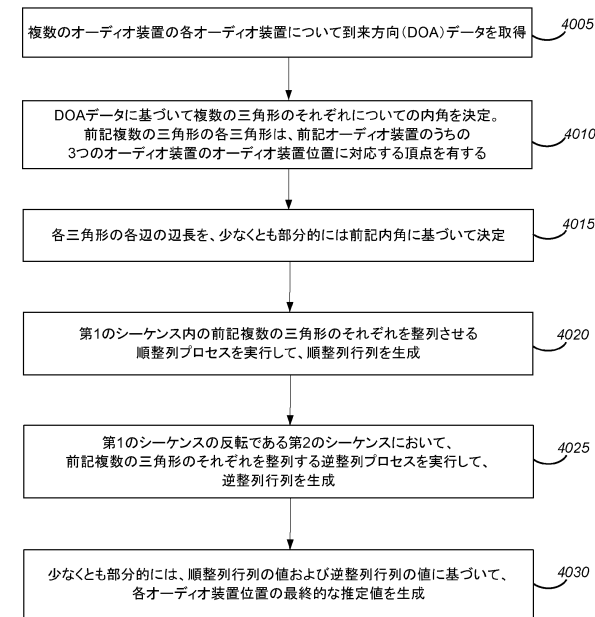


Figure 39

【図 40】



4000

Figure 40

【図 41】

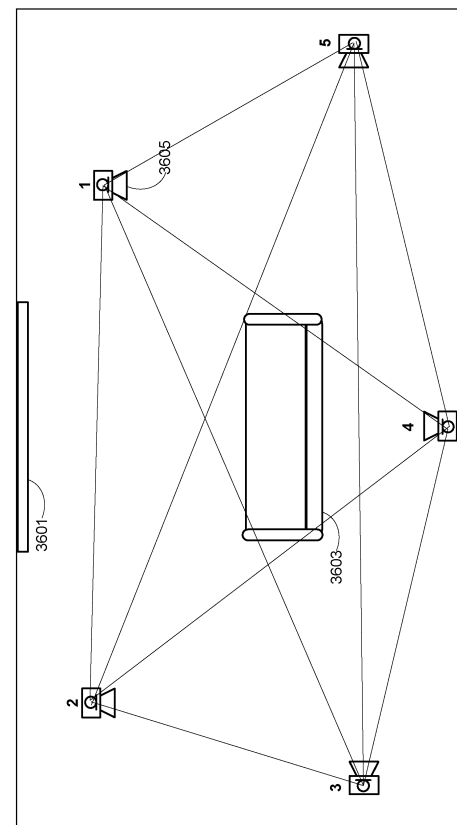


Figure 41

10

20

30

40

50

【図 4 2】

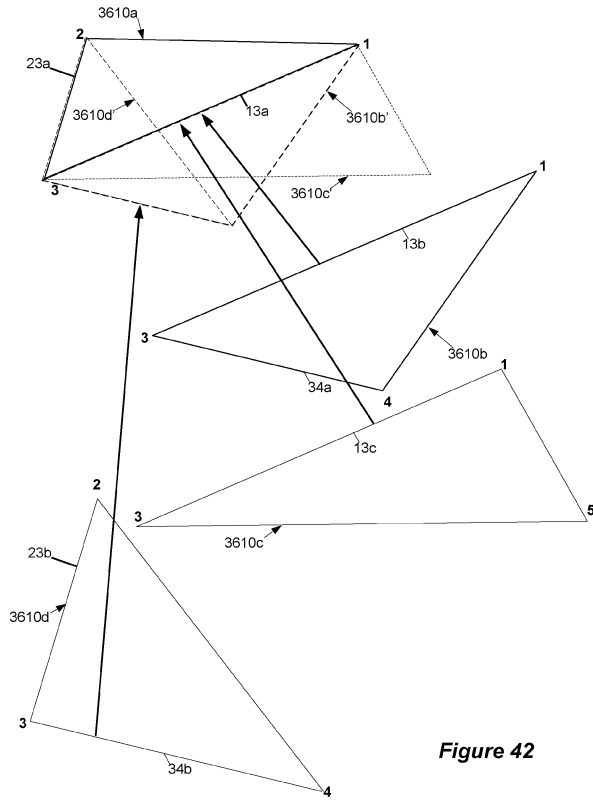


Figure 42

【図 4 3】

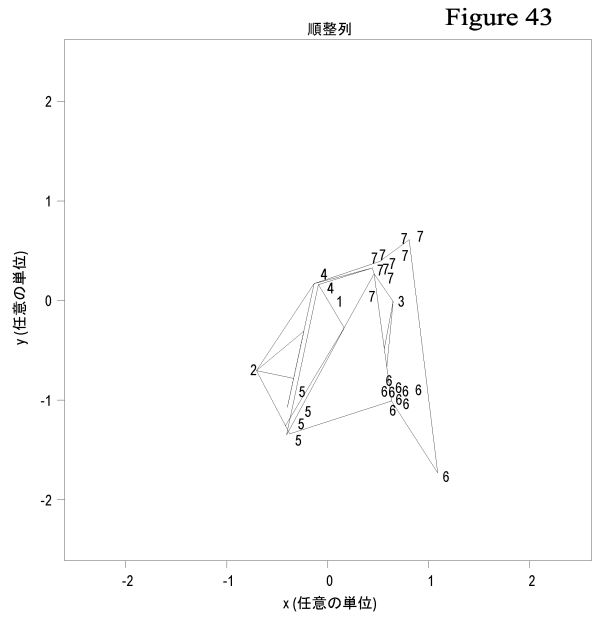


Figure 43

【図 4 4】

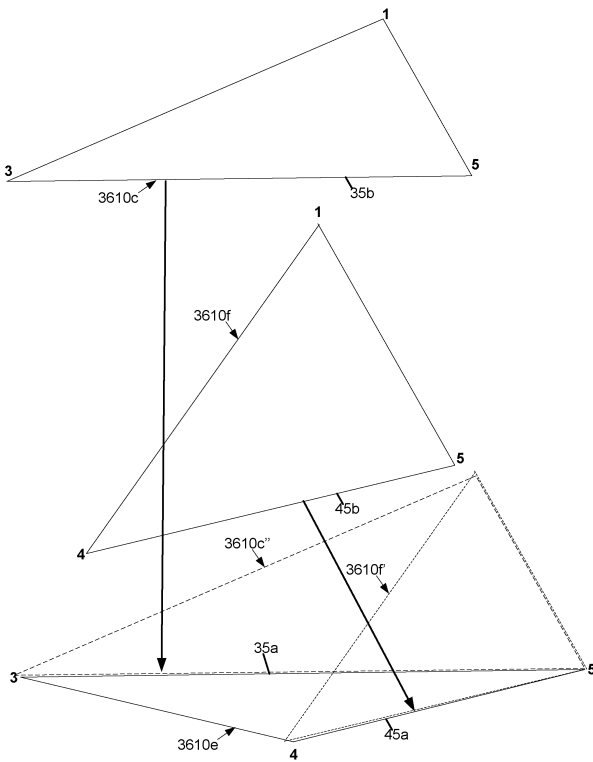


Figure 44

【図 4 5】

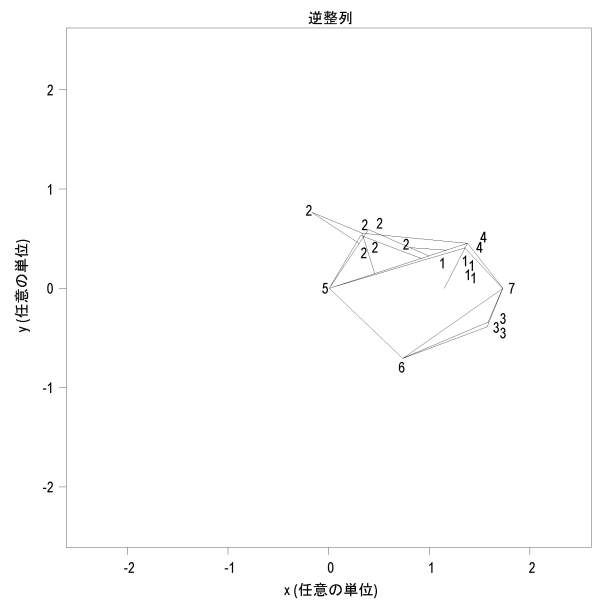


Figure 45

10

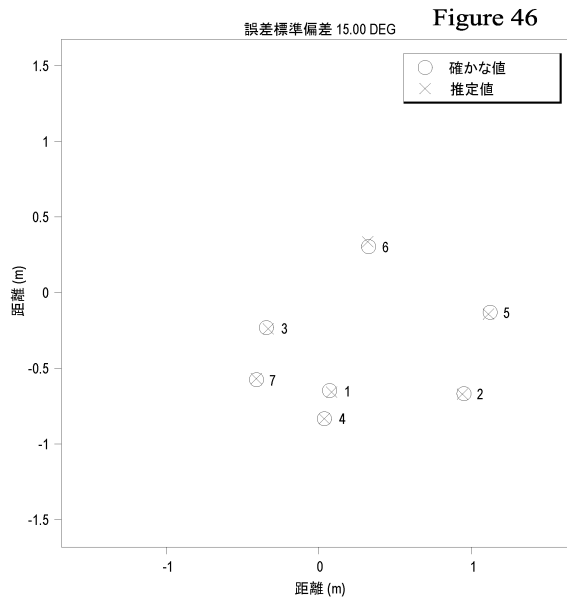
20

30

40

50

【図 4 6】



【図 4 7】

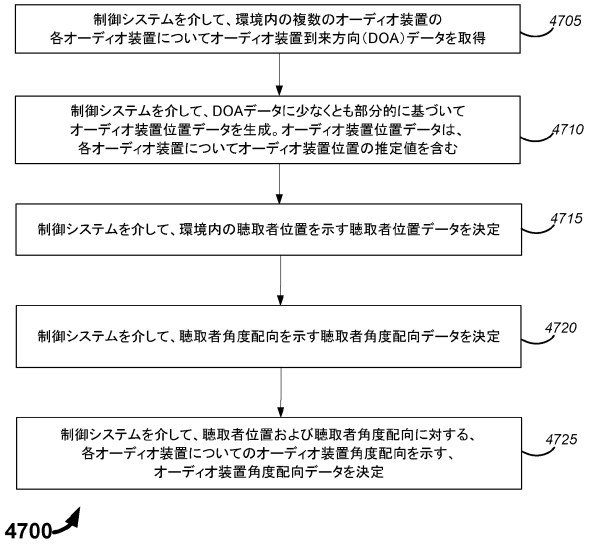
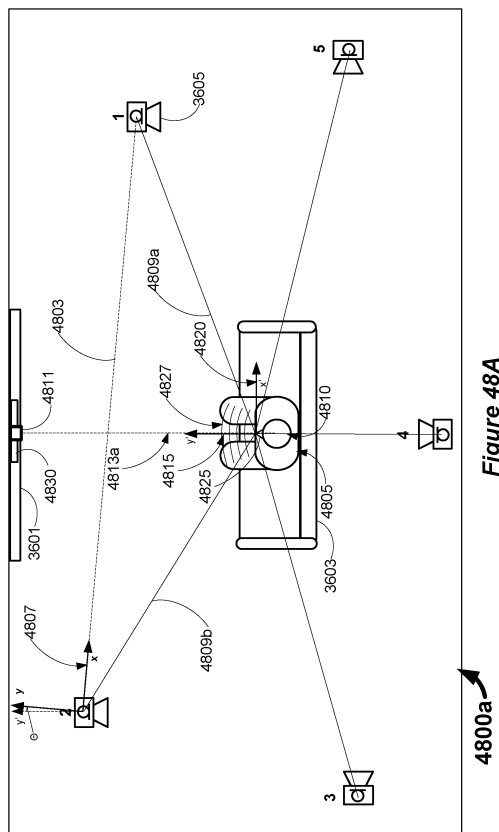
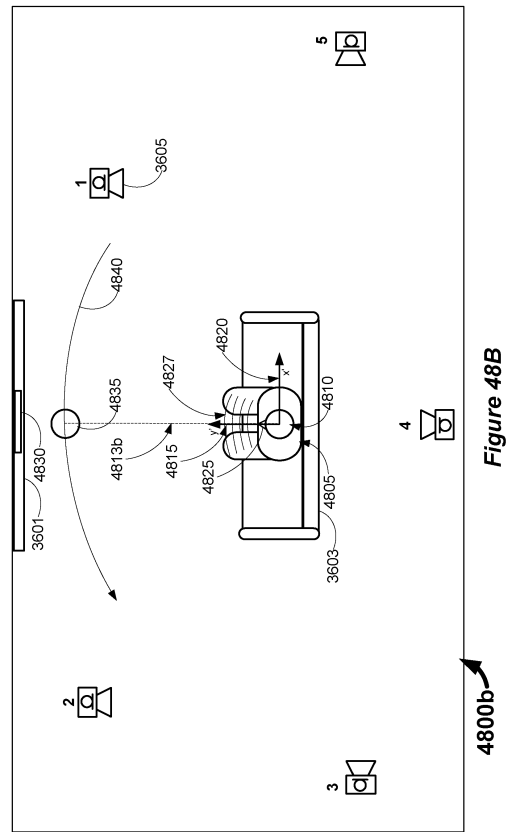


Figure 47

【図 4 8 A】



【図 4 8 B】



10

20

30

40

50

【図 48 C】

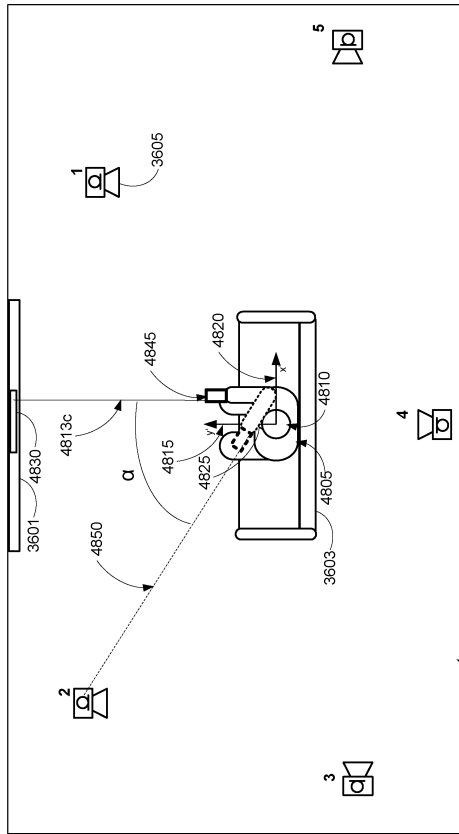


Figure 48C

【図 48 D】

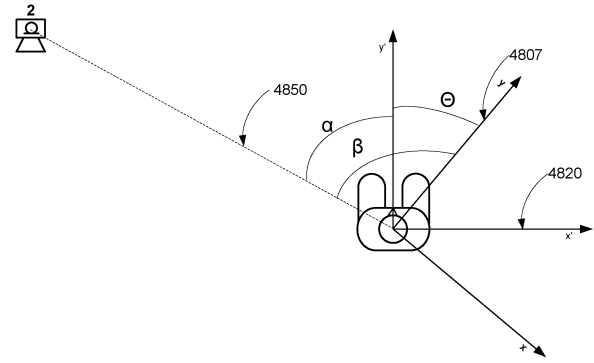


Figure 48D

【図 49】

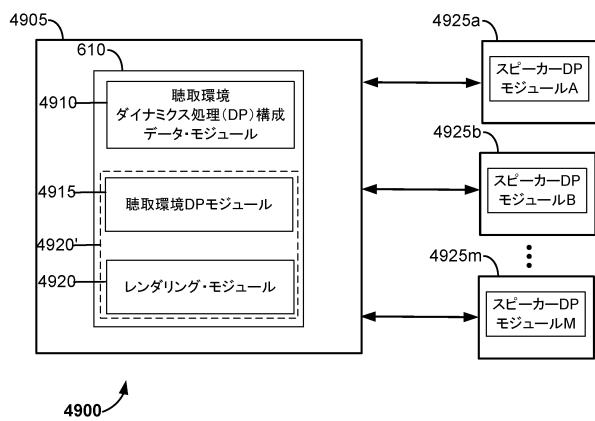


Figure 49

【図 50】

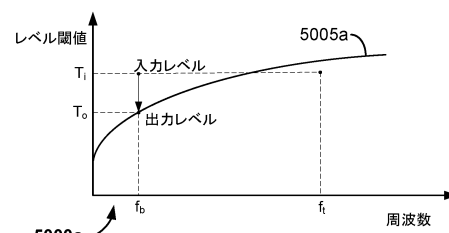


Figure 50A

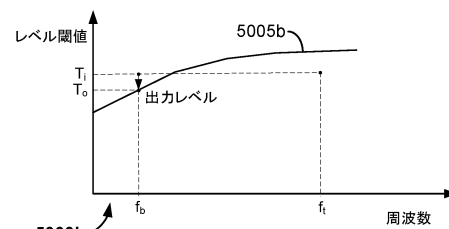


Figure 50B

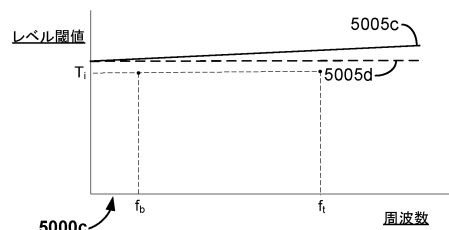


Figure 50C

10

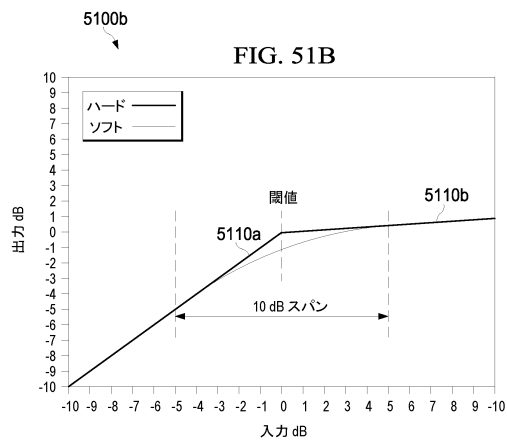
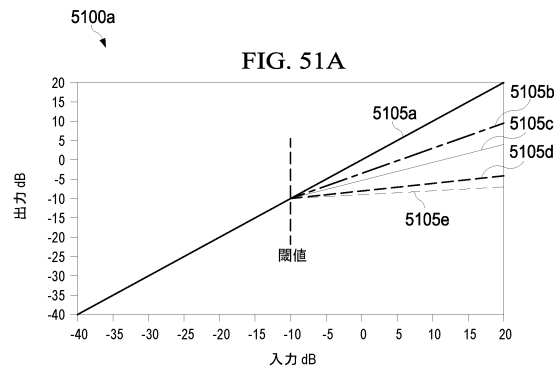
20

30

40

50

【図 5 1】



【図 5 2】



Figure 52

【図 5 3】



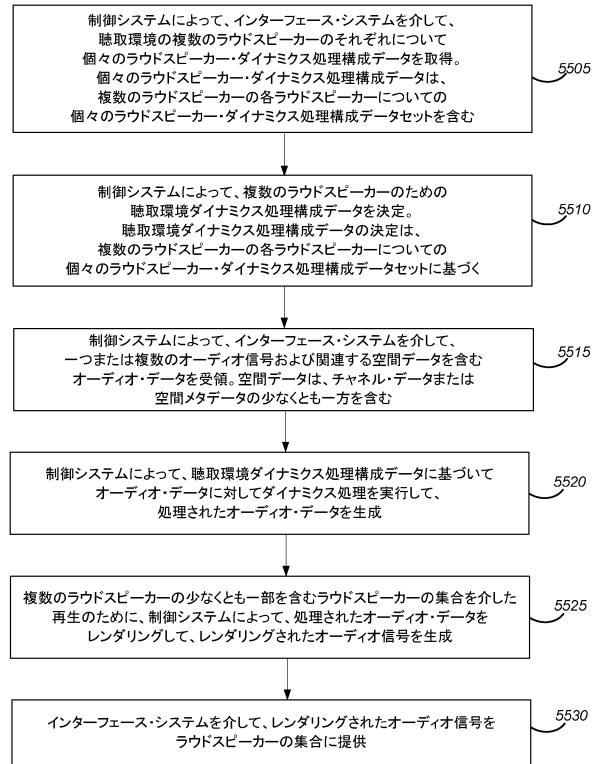
Figure 53

【図 5 4】



Figure 54

【図 5 5】



5500

Figure 55

【図 5 6 A】

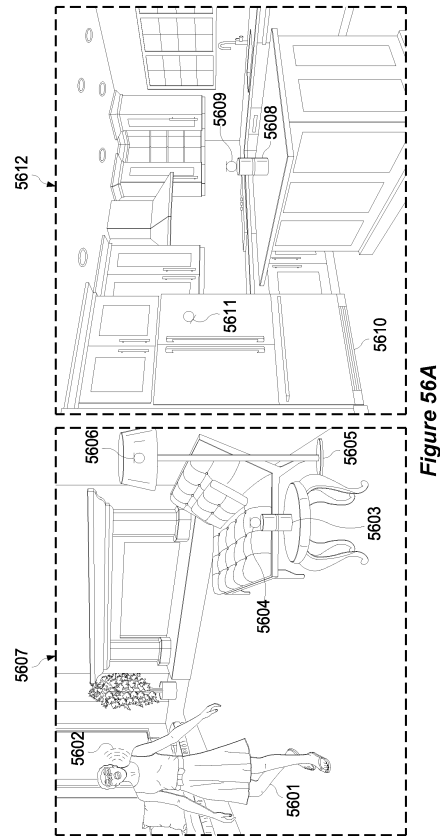


Figure 56A

【図 5 6 B】

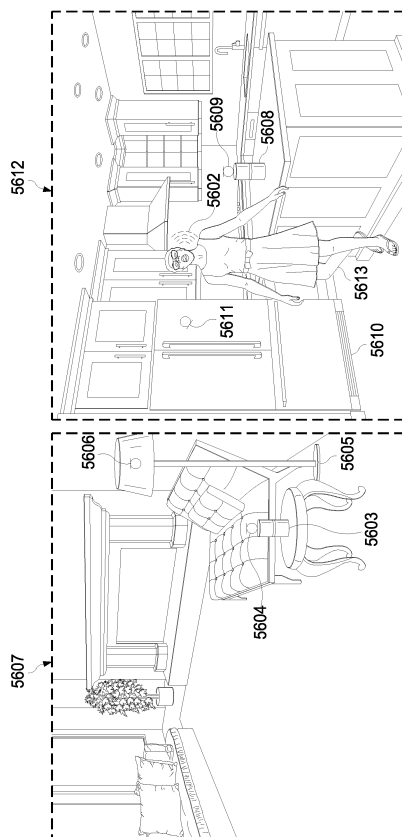


Figure 56B

【図 5 7】

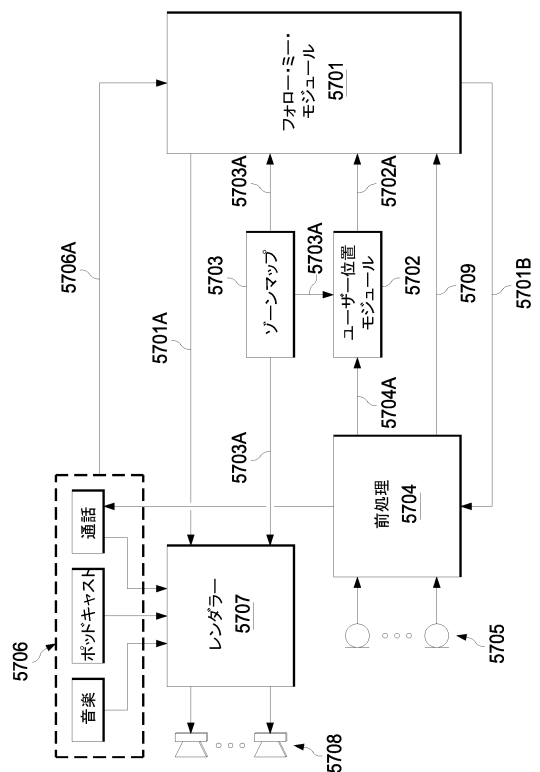


Figure 57

10

20

30

40

50

【図 58】

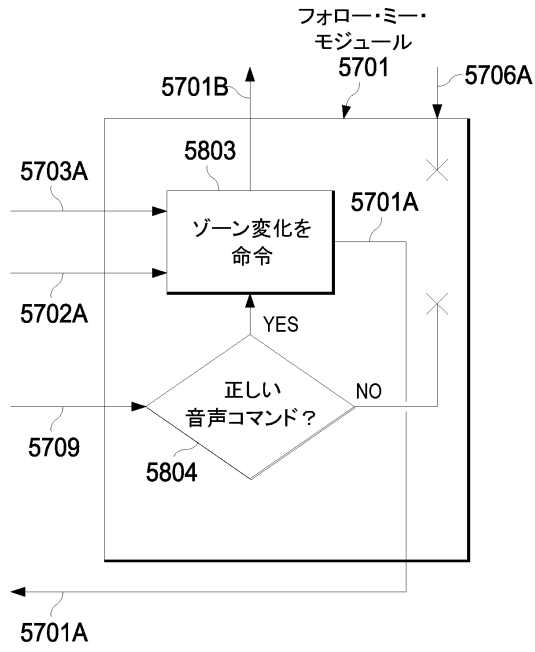


Figure 58

【図 59】

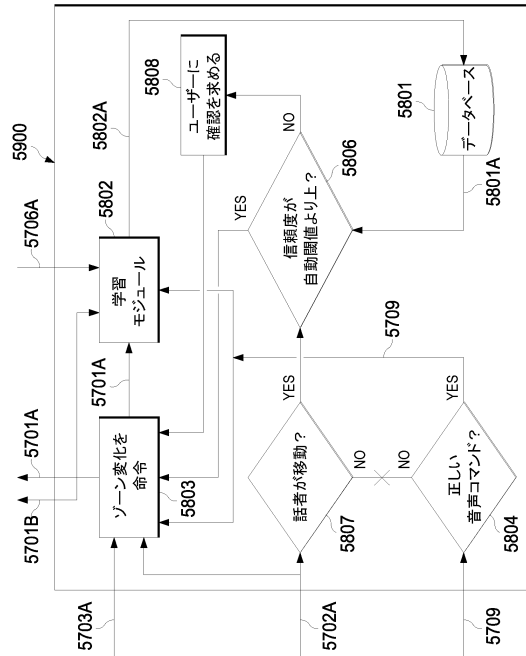


Figure 59

【図 60】

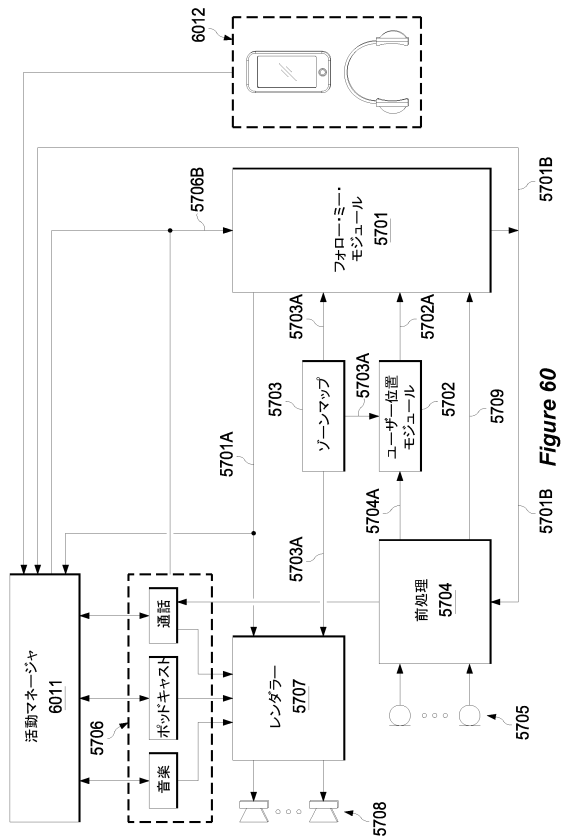


Figure 60

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)
- (31)優先権主張番号 62/880,115
(32)優先日 令和1年7月30日(2019.7.30)
(33)優先権主張国・地域又は機関
米国(US)
- (31)優先権主張番号 62/880,118
(32)優先日 令和1年7月30日(2019.7.30)
(33)優先権主張国・地域又は機関
米国(US)
- (31)優先権主張番号 62/880,121
(32)優先日 令和1年7月30日(2019.7.30)
(33)優先権主張国・地域又は機関
米国(US)
- (31)優先権主張番号 19217580.0
(32)優先日 令和1年12月18日(2019.12.18)
(33)優先権主張国・地域又は機関
欧州特許庁(EP)
- (31)優先権主張番号 62/949,998
(32)優先日 令和1年12月18日(2019.12.18)
(33)優先権主張国・地域又は機関
米国(US)
- (31)優先権主張番号 62/971,421
(32)優先日 令和2年2月7日(2020.2.7)
(33)優先権主張国・地域又は機関
米国(US)
- (31)優先権主張番号 62/992,068
(32)優先日 令和2年3月19日(2020.3.19)
(33)優先権主張国・地域又は機関
米国(US)
- (31)優先権主張番号 62/705,143
(32)優先日 令和2年6月12日(2020.6.12)
(33)優先権主張国・地域又は機関
米国(US)
- (31)優先権主張番号 62/705,351
(32)優先日 令和2年6月23日(2020.6.23)
(33)優先権主張国・地域又は機関
米国(US)
- (31)優先権主張番号 62/705,410
(32)優先日 令和2年6月25日(2020.6.25)
(33)優先権主張国・地域又は機関
米国(US)
- (31)優先権主張番号 16/929,215
(32)優先日 令和2年7月15日(2020.7.15)
(33)優先権主張国・地域又は機関
米国(US)
- (31)優先権主張番号 62/705,883
(32)優先日 令和2年7月20日(2020.7.20)
(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/705,884

(32)優先日 令和2年7月20日(2020.7.20)

(33)優先権主張国・地域又は機関

米国(US)

(74)代理人 100070150

弁理士 伊東 忠彦

(74)代理人 100135079

弁理士 宮崎 修

(72)発明者 ディキンズ, グレン エヌ.

オーストラリア 2 2 2 6 ニューサウスウェールズ州 コモ ダヴリーズ ロード 1 6

(72)発明者 カートライト, リチャード ジェイ.

アメリカ合衆国 9 4 1 0 3 カリフォルニア州 サンフランシスコ マーケット ストリート 1 2
7 5 ドルビー ラボラトリーズ インコーポレイテッド 内

(72)発明者 グナワン, デイヴィッド

アメリカ合衆国 9 4 1 0 3 カリフォルニア州 サンフランシスコ マーケット ストリート 1 2
7 5 ドルビー ラボラトリーズ インコーポレイテッド 内

(72)発明者 ハインズ, クリストファー グレアム

オーストラリア 2 0 6 0 ニューサウスウェールズ州 シドニー ラヴェンダー ストリート 6 4

(72)発明者 トーマス, マーク アール. ピー.

アメリカ合衆国 9 4 1 0 3 カリフォルニア州 サンフランシスコ マーケット ストリート 1 2
7 5 ドルビー ラボラトリーズ インコーポレイテッド 内

(72)発明者 ゼーフェルト, アラン ジェイ.

アメリカ合衆国 9 4 1 0 3 カリフォルニア州 サンフランシスコ マーケット ストリート 1 2
7 5 ドルビー ラボラトリーズ インコーポレイテッド 内

(72)発明者 ランドー, ジョシュア ビー.

アメリカ合衆国 9 4 1 0 3 カリフォルニア州 サンフランシスコ マーケット ストリート 1 2
7 5 ドルビー ラボラトリーズ インコーポレイテッド 内

(72)発明者 ディオニシオ, カルロス エドゥアルド メダーリャ

アメリカ合衆国 9 4 1 0 3 カリフォルニア州 サンフランシスコ マーケット ストリート 1 2
7 5 ドルビー ラボラトリーズ インコーポレイテッド 内

(72)発明者 アルテアガ, ダニエル

アメリカ合衆国 9 4 1 0 3 カリフォルニア州 サンフランシスコ マーケット ストリート 1 2
7 5 ドルビー ラボラトリーズ インコーポレイテッド 内

審査官 大野 弘

(56)参考文献 特表2016-502192 (JP, A)

特開2019-121827 (JP, A)

特開2017-123644 (JP, A)

(58)調査した分野 (Int.Cl., DB名)

H 0 4 R 3 / 0 0

H 0 4 S 7 / 0 0

G 1 0 K 1 5 / 0 2

G 0 6 F 3 / 1 6

G 0 6 F 3 / 0 1