

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7578219号  
(P7578219)

(45)発行日 令和6年11月6日(2024.11.6)

(24)登録日 令和6年10月28日(2024.10.28)

|                   |                               |          |                         |         |
|-------------------|-------------------------------|----------|-------------------------|---------|
| (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   |
| G 1 0 L           | 15/10 (2006.01)               | G 1 0 L  | 15/10                   | 2 0 0 W |
| G 1 0 L           | 25/84 (2013.01)               | G 1 0 L  | 25/84                   |         |
| H 0 4 S           | 7/00 (2006.01)                | H 0 4 S  | 7/00                    | 3 0 0   |
| 請求項の数 28 (全125頁)  |                               |          |                         |         |
| (21)出願番号          | 特願2022-506084(P2022-506084)   | (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2022-542387(P2022-542387 A) |          | グ コーポレイション              |         |
| (43)公表日           | 令和4年10月3日(2022.10.3)          |          | アメリカ合衆国 9 4 1 0 3 カリフォル |         |
| (86)国際出願番号        | PCT/US2020/043696             |          | ニア州 サンフランシスコ マーケット      |         |
| (87)国際公開番号        | WO2021/021707                 | (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 | スペイン(ES)                      |          | ンズ サー ジョン ロジャーソンズ キー    |         |
| (31)優先権主張番号       | 62/880,111                    | (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】 複数のスピーカーを通じた複数のオーディオ・ストリームの再生の管理

(57)【特許請求の範囲】

【請求項1】

インターフェース・システム；  
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；および  
制御システム

を有するオーディオ処理システムであって：

前記制御システムは：

前記インターフェース・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  
ムを受領する段階であって、前記第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  
ーディオ信号を含み、前記第1のオーディオ・プログラム・ストリームは、チャンネル・デ  
ータまたは空間メタデータの少なくとも一方を含む第1の空間データを含む、段階と；

前記環境の前記スピーカーを介した再生のために前記第1のオーディオ信号をレンダ  
リングして、第1のレンダリングされたオーディオ信号を生成する段階とを実行するよう  
に構成された第1のレンダリング・モジュールと；

前記インターフェース・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  
ムを受領する段階であって、前記第2のオーディオ・プログラム・ストリームは、前記環  
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第2のオ  
ーディオ信号を含む、段階と；

前記環境の少なくとも1つのスピーカーを介した再生のために前記第2のオーディオ信

号をレンダリングして、第2のレンダリングされたオーディオ信号を生成する段階とを実行するように構成された第2のレンダリング・モジュールとを有しており、

前記第1のレンダリング・モジュールは、前記マイクロフォン・システムからの第1のマイクロフォン信号と、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れらの特性のうちの少なくとも1つとに少なくとも部分的に基づいて、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ように構成されており；

前記第2のレンダリング・モジュールは、前記マイクロフォン・システムからの前記第1のマイクロフォン信号と、前記第1のオーディオ信号、前記第1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つとに少なくとも部分的に基づいて、前記第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、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ように構成されており；

前記制御システムはさらに、

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と前記修正された第2のレンダリングされたオーディオ信号とを混合して、混合されたオーディオ信号を生成するように構成された混合モジュールを有しており、

前記制御システムはさらに、

前記第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；

前記第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し；

前記第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、前記第1のオーディオ信号または前記第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段階であって、前記第1のオーディオ信号についてのレンダリング・プロセスを修正することは、前記第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前記第1の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、前記第1の音源位置の近くのスピーカーによって再生される前記第1のレンダリングされたオーディオ信号のラウドネスを低減することを含む、段階を実行し；

前記混合されたオーディオ信号を、前記環境の少なくともいくつかのスピーカーに提供するようにさらに構成されている、  
オーディオ処理システム。

【請求項2】

一つまたは複数の追加的なレンダリング・モジュールをさらに有しており、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：

前記インターフェース・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段階であって、前記追加的なオーディオ・プログラム・ストリームは、前記環境の少なくとも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む、段階と；

前記追加的なオーディオ信号を、前記環境の少なくとも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て、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段階と；

少なくとも部分的には、前記第1のオーディオ信号、前記第1のレンダリングされたオーディオ信号、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れらの特性のうちの少なくとも1つに基づいて、前記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段階とを実行するように構成されており、

前記混合モジュールは、前記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いる、

請求項1に記載のオーディオ処理システム。

【請求項3】

前記第1のオーディオ信号についてのレンダリング・プロセスを修正することは：

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、  
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；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ラウドネスを修正すること

のうちの一つまたは複数を実行することを含む、

請求項2に記載のオーディオ処理システム。

【請求項4】

前記第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：

前記第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；または

前記第1のオーディオ信号または前記第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に応答して、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ラウドネスを修正すること

のうちの一つまたは複数を実行することを含む、

請求項1ないし3のうちいずれか一項に記載のオーディオ処理システム。

【請求項5】

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からなる修正のリストか  
らの一つまたは複数の修正を実行することを含む、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処理システム。

【請求項6】

前記制御システムは：

前記第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し；

前記第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に少なくとも部分的に  
基づいて、前記第1のオーディオ信号または前記第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ようにさらに構成されている、

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処理システム。

【請求項7】

前記制御システムは：

前記第1のマイクロフォン信号がウェイクワードに対応することを判別し；

前記ウェイクワードへの返答を決定し；

前記返答を再生するよう前記第1の音源位置の近くの少なくとも1つのスピーカーを制御  
するように構成されている、

請求項1ないし6のうちいずれか一項に記載のオーディオ処理システム。

【請求項8】

前記制御システムは：

前記第1のマイクロフォン信号がコマンドに対応することを判別し；

前記コマンドへの返答を決定し；

前記返答を再生するよう前記第1の音源位置の近くの少なくとも1つのスピーカーを制御  
し；

前記コマンドを実行するように構成されている、

請求項1ないし6のうちいずれか一項に記載のオーディオ処理システム。

【請求項9】

前記制御システムは、前記返答を再生するよう前記第1の音源位置の近くの少なくとも1  
つのスピーカーを制御した後に、前記第1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さらに構成されている、請求項8に記載のオーディオ処理シ  
ステム。

【請求項10】

前記制御システムは、再生される第1のオーディオ・プログラム・ストリームまたは再生される第2のオーディオ・プログラム・ストリームについて、少なくとも部分的に前記第1のマイクロフォン信号に基づいて、ラウドネス推定値を導出するように構成されてもよい、請求項1ないし9のうちいずれか一項に記載のオーディオ処理システム。

【請求項11】

前記制御システムは、少なくとも部分的には前記ラウドネス推定値に基づいて、前記第1のオーディオ信号または前記第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いる、請求項10に記載のオーディオ処理システム。

【請求項12】

前記ラウドネス推定値は、知覚されるラウドネス推定値であり、前記レンダリング・プロセスを修正することは、干渉信号の存在下で前記第1のオーディオ信号または前記第2のオーディオ信号の知覚されるラウドネスを維持するために、前記第1のオーディオ信号または前記第2のオーディオ信号の少なくとも一方を変更することを含む、請求項11に記載のオーディオ処理システム。

【請求項13】

前記制御システムは：

前記第1のマイクロフォン信号が人間の声に対応することを判別し；

前記第1の音源位置とは異なる、前記環境のある位置に近い一つまたは複数のスピーカにおいて前記第1のマイクロフォン信号を再生するように構成されている、  
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処理システム。

【請求項14】

前記制御システムは、前記第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さらに構成されており、前記環境の前記位置は、介護者の推定された位置に対応する、請求項13に記載のオーディオ処理システム。

【請求項15】

前記制御システムは、再生される第1のオーディオ・プログラム・ストリームまたは再生される第2のオーディオ・プログラム・ストリームについてのラウドネス推定値を導出するように構成されており、前記制御システムは、少なくとも部分的には前記ラウドネス推定値に基づいて、前記第1のオーディオ信号または前記第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さらに構成されている、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処理システム。

【請求項16】

前記ラウドネス推定値は、知覚されるラウドネス推定値であり、前記レンダリング・プロセスを修正することは、干渉信号の存在下で知覚されるラウドネスを維持するために、前記第1のオーディオ信号または前記第2のオーディオ信号の少なくとも一方を変更することを含む、請求項15に記載のオーディオ処理システム。

【請求項17】

前記第1のオーディオ信号をレンダリングすることまたは前記第2のオーディオ信号をレンダリングすることの少なくとも一方は、任意に配置されたスピーカへの柔軟なレンダリングを含む、請求項1ないし16のうちいずれか一項に記載のオーディオ処理システム。

【請求項18】

前記柔軟なレンダリングは、質量中心振幅パンまたは柔軟仮想化を含む、請求項17に記載のオーディオ処理システム。

【請求項19】

オーディオ処理方法であって：

第1のレンダリング・モジュールによって、第1のオーディオ・プログラム・ストリームを受領する段階であって、前記第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み、前記第1のオーディオ・プログラム・ストリームは、チャンネル・デー

10

20

30

40

50

タまたは空間メタデータの少なくとも一方を含む第1の空間データを含む、段階と；

前記第1のレンダリング・モジュールによって、前記環境の前記スピーカーを介した再生のために前記第1のオーディオ信号をレンダリングして、第1のレンダリングされたオーディオ信号を生成する段階と；

第2のレンダリング・モジュールによって、第2のオーディオ・プログラム・ストリームを受領する段階であって、前記第2のオーディオ・プログラム・ストリームは、前記環境の少なくとも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第2のオーディオ信号を含む、段階と；

前記第2のレンダリング・モジュールによって、前記環境の少なくとも1つのスピーカーを介した再生のために前記第2のオーディオ信号をレンダリングして、第2のレンダリングされたオーディオ信号を生成する段階と；

前記第1のレンダリング・モジュールによって、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と、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れらの特性のうちの少なくとも1つとに少なくとも部分的に基づいて、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段階と；

前記第2のレンダリング・モジュールによって、前記マイクロフォン・システムからの前記第1のマイクロフォン信号と、前記第1のオーディオ信号、前記第1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つとに少なくとも部分的に基づいて、前記第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、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段階と；

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と前記修正された第2のレンダリングされたオーディオ信号とを混合して、混合されたオーディオ信号を生成する段階とを含み、

当該方法はさらに：

前記第1のマイクロフォン信号に基づいて第1の音源位置を推定する段階と；

前記第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段階と；

前記第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、前記第1のオーディオ信号または前記第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段階であって、前記第1のオーディオ信号についてのレンダリング・プロセスを修正することは、前記第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前記第1の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、前記第1の音源位置の近くのスピーカーによって再生される前記第1のレンダリングされたオーディオ信号のラウドネスを低減することを含む、段階と；

前記混合されたオーディオ信号を、前記環境の少なくともいくつかのスピーカーに提供する段階とを含む、オーディオ処理方法。

【請求項 20】

前記第1のオーディオ信号についてのレンダリング・プロセスを修正することは：

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、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；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ラウドネスを修正すること  
のうちの一つまたは複数を実行することを含む、

請求項 19 に記載のオーディオ処理方法。

【請求項 21】

前記第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：

前記第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；または

前記第1のオーディオ信号または前記第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に応答して、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ラウドネスを修正すること  
のうちの一つまたは複数を実行することを含む、  
請求項19または20に記載のオーディオ処理方法。

【請求項22】

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からなる修正のリストからの一つまたは複数の修正を実行することを含む、請求項19ないし21のうちいずれか一項に記載のオーディオ処理方法。

【請求項23】

前記第1のレンダリング・モジュールによって、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、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、前記第2のレンダリング・モジュールによって、少なくとも部分的に前記第1のマイクロフォン信号に基づいて、前記第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をさらに含む、請求項19ないし22のうちいずれか一項に記載のオーディオ処理方法。

【請求項24】

オーディオ処理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ソフトウェアが記憶されている一つまたは複数の非一時的媒体であって、前記オーディオ処理方法は：

第1のレンダリング・モジュールによって、第1のオーディオ・プログラム・ストリームを受領する段階であって、前記第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み、前記第1のオーディオ・プログラム・ストリームは、チャンネル・データまたは空間メタデータの少なくとも一方を含む第1の空間データを含む、段階と；

前記第1のレンダリング・モジュールによって、前記環境の前記スピーカーを介した再生のために前記第1のオーディオ信号をレンダリングして、第1のレンダリングされたオーディオ信号を生成する段階と；

第2のレンダリング・モジュールによって、第2のオーディオ・プログラム・ストリームを受領する段階であって、前記第2のオーディオ・プログラム・ストリームは、前記環境の少なくとも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第2のオーディオ信号を含む、段階と；

前記第2のレンダリング・モジュールによって、前記環境の少なくとも1つのスピーカーを介した再生のために前記第2のオーディオ信号をレンダリングして、第2のレンダリングされたオーディオ信号を生成する段階と；

前記第1のレンダリング・モジュールによって、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と、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れらの特性のうちの少なくとも1つとに少なくとも部分的に基づいて、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段階と；

前記第2のレンダリング・モジュールによって、前記マイクロフォン・システムからの前記第1のマイクロフォン信号と、前記第1のオーディオ信号、前記第1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つとに少なくとも部分的に基づいて、前記第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、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段階と；

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と前記修正された第2のレンダリングされたオーディオ信号とを混合して、混合されたオーディオ信号を生成する段階とを含む、

前記オーディオ処理方法はさらに：

10

20

30

40

50

前記第1のマイクロフォン信号に基づいて第1の音源位置を推定する段階と；  
前記第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段階と；  
前記第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少なくとも部分的に基づいて、前記第1のオーディオ信号または前記第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段階であって、前記第1のオーディオ信号についてのレンダリング・プロセスを修正することは、前記第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前記第1の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、前記第1の音源位置の近くのスピーカーによって再生される前記第1のレンダリングされたオーディオ信号のラウドネスを低減することを含む、段階と；

前記混合されたオーディオ信号を、前記環境の少なくともいくつかのスピーカーに提供する段階とを含む、  
一つまたは複数の非一時的な媒体。

【請求項 2 5】

前記第1のオーディオ信号についてのレンダリング・プロセスを修正することは：  
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、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；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ラウドネスを修正すること  
のうちの一つまたは複数を実行することを含む、  
請求項 2 4 に記載の一つまたは複数の非一時的な媒体。

【請求項 2 6】

前記第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：  
前記第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；または  
前記第1のオーディオ信号または前記第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に応答して、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ラウドネスを修正すること  
のうちの一つまたは複数を実行することを含む、  
請求項 2 4 または 2 5 に記載の一つまたは複数の非一時的な媒体。

【請求項 2 7】

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からなる修正のリストからの一つまたは複数の修正を実行することを含む、請求項 2 4 ないし 2 6 のうちいずれか一項に記載の一つまたは複数の非一時的な媒体。

【請求項 2 8】

前記オーディオ処理方法が、前記第1のレンダリング・モジュールによって、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、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、前記第2のレンダリング・モジュールによって、少なくとも部分的に前記第1のマイクロフォン信号に基づいて、前記第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をさらに含む、請求項 2 4 ないし 2 7 のうちいずれか一項に記載の一つまたは複数の非一時的な媒体。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

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

10

20

30

40

50

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62/880,111号（2019年7月30日出願）、米国仮特許出願第62/704,754号（2020年5月27日出願）、米国仮特許出願第62/705,896号（2020年7月21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）および米国仮特許出願第62/705,143号（2020年6月12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。

#### 【0002】

##### 技術分野

本開示は、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、アクティブ化されている各スピーカー）によるオーディオの再生および再生のためのレンダリングのための、システムおよび方法に関する。

#### 【背景技術】

#### 【0003】

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。

#### 【0004】

##### 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。

#### 【0005】

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。

#### 【0006】

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。

#### 【0007】

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

10

20

30

40

50

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。

【0008】

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。

【0009】

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。

【0010】

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。

【0011】

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、

10

20

30

40

50

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。

10

#### 【0012】

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。

#### 【0013】

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。

20

#### 【発明の概要】

#### 【課題を解決するための手段】

#### 【0014】

いくつかの実施形態は、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少なくとも1つ（たとえば、全部または一部）による、オーディオの複数のストリームの再生を管理するための方法に関わる。

#### 【0015】

あるクラスの実施形態は、複数の協調させられる（オーケストレーションされる）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）による再生を管理する方法を含む。たとえば、ユーザーの家庭において（システム内に）存在する一組のスマートオーディオ装置が、スマートオーディオ装置の全部または一部による（すなわち、全部または一部のスマートオーディオ装置のスピーカーによる）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る。

40

#### 【0016】

（たとえば、多様な同時使用事例を扱うために家庭において）スマートオーディオ装置をオーケストレーションすることは、相互接続された一組のスピーカーを通じた、一つまたは複数のオーディオ・プログラム・ストリームの同時再生に関わってもよい。たとえば、ユーザーは、一組のスピーカー（たとえば、一組のスマートオーディオ装置に含まれるか、または、一組のスマートオーディオ装置によって制御されるもの）を通じて映画の

50

トモス・サウンドトラック（または、他のオブジェクトベースのオーディオ・プログラム）を聴いていてもよく、その際、ユーザーは、関連するスマートオーディオ装置（たとえば、スマートアシスタント）に対してコマンド（たとえば、ウェイクワードおよびそれに続くコマンド）を発声してもよい。この場合、プログラムの空間的呈示（たとえば、アトモス・ミックス）を話者（話しているユーザー）の位置から離れるようにワープさせ（wrap）、スマートオーディオ装置の（たとえば、音声アシスタントの）の対応する応答の再生を、話者に近いスピーカーへの応答に差し向けるように、システムによるオーディオ再生は（いくつかの実施形態に従って）修正されてもよい。これは、コマンド（または対応するウェイクワード）の検出に応答して、オーディオ・プログラム・コンテンツの再生の音量を単に減少させることに比べて、重要な利点を提供することができる。同様に、ユーザーは、同じプログラム（たとえば、アトモス・サウンドトラック）が隣接する開放された居住空間で再生されている間に、キッチンで料理のヒントを得るためにスピーカーを使用することを望むことがありうる。この場合、いくつかの実施形態によれば、プログラム（たとえば、アトモス・サウンドトラック）の再生は、キッチンから離れるようにワープされることができ、料理のヒントは、キッチン付近またはキッチン内のスピーカーから再生されることができ、さらに、キッチンで再生される料理のヒントは、居住空間からしみ出してくることがありうるプログラム（たとえば、アトモス・サウンドトラック）のどれよりも大きく、キッチンにいる人によって聞かれるように、動的に（いくつかの実施形態に従って）調整されることができ、

【0017】

いくつかの実施形態は、上述の例示的な使用事例および考えられる他の多数を実装するように構成された複数ストリーム・レンダリング・システムである。あるクラスの実施形態では、オーディオ・レンダリング・システムは、複数の任意に配置されたラウドスピーカーを通じた同時再生のために複数のオーディオ・プログラム・ストリームをレンダリングする（および／またはそれらのストリームを同時に再生する）するように構成されてもよく、前記プログラム・ストリームの少なくとも1つは空間ミックスであり、前記空間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）は、一つまたは複数の追加的なプログラム・ストリームの同時再生（またはレンダリングおよび再生）に応答して（または関連して）動的に修正される。

【0018】

いくつかの実施形態の諸側面は、開示された方法のいずれかの実施形態またはそのステップ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、および開示された方法のいずれかの実施形態またはそのステップを実行するためのコード（たとえば、実行するためにコード実行可能なコード）を記憶する、データの非一時的記憶を実装する有形の非一時的な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の記憶媒体）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、開示された方法の実施形態またはそのステップを含む、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た、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、それに対して呈されたデータに応答して、開示された方法の実施形態（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、または、それを含んでもよい。

【0019】

本開示の少なくともいくつかの側面は、装置を介して実装されてもよい。たとえば、一つまたは複数の装置が、少なくとも部分的に、本明細書に開示された方法を実施することができてもよい。いくつかの実装では、装置は、インターフェース・システムおよび制御システムを有するオーディオ処理システムであるか、またはこれを含む。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

10

20

30

40

50

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デバイス、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。

#### 【0020】

いくつかの実装形態では、制御システムは、少なくとも2つのレンダリング・モジュールを含むか、または実装する。いくつかの例によれば、制御システムは、 $N$ が2よりも大きい整数であるとして、 $N$ 個のレンダリング・モジュールを含んでいてもよく、または実装してもよい。

#### 【0021】

いくつかの例では、第1のレンダリング・モジュールは、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ように構成される。場合によっては、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。いくつかの例では、第1のオーディオ・プログラム・ストリームは、チャンネル・データおよび／または空間メタデータを含む、第1の空間データを含む。いくつかの例によれば、第1のレンダリング・モジュールは、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ように構成される。

#### 【0022】

いくつかの実装では、第2のレンダリング・モジュールは、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ように構成される。いくつかの例では、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。いくつかの例では、第2のオーディオ・プログラム・ストリームは、チャンネル・データおよび／または空間メタデータを含む、第2の空間データを含む。いくつかの例によれば、第2のレンダリング・モジュールは、環境のスピーカーを介した再生のために第2のオーディオ信号をレンダリングして、第2のレンダリングされたオーディオ信号を生成するように構成される。

#### 【0023】

いくつかの例によれば、第1のレンダリング・モジュールは、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れらの特性のうちの少なくとも1つ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・プロセスを修正し、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ように構成される。いくつかの例によれば、第2のレンダリング・モジュールはさらに、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れらの特性のうちの少なくとも1つに少なくとも部分的に基づいて、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、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ように構成される。

#### 【0024】

いくつかの実装では、オーディオ処理システムは、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ように構成された混合モジュールを含む。いくつかの例では、制御システムは、混合されたオーディオ信号を、環境の少なくともいくつかのスピーカーに提供するようにさらに構成される。

#### 【0025】

いくつかの例によれば、オーディオ処理システムは、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

10

20

30

40

50

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、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するように構成されてもよい。いくつかの例では、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、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ように構成されてもよい。いくつかのそのような例において、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。

10

#### 【0026】

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を含みうる。代替的または追加的に、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ラウドネスを修正することを含んでいてもよい。

20

#### 【0027】

いくつかの例によれば、第2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を含んでいてもよい。代替的または追加的に、第2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んでいてもよい。いくつかの実装によれば、第1のオーディオ信号および／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および／またはダイナミックレンジ修正を実行することを含んでいてもよい。

30

#### 【0028】

いくつかの例では、オーディオ処理システムは、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。いくつかのそのような例では、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。

40

#### 【0029】

いくつかの例によれば、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。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環境ノイズに対応するかどうかを判定し、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

50

ようにさらに構成されてもよい。

【0030】

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し、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ように構成されてもよい。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第1の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の音源位置の近くのスピーカーによって再生される第1のレンダリングされたオーディオ信号のラウドネスを低減することを含むことができる。

10

【0031】

いくつかの例によれば、制御システムは、第1のマイクロフォン信号がウェイクワードに対応することを判別し、ウェイクワードへの返答を決定し、該返答を再生するために第1の音源位置の近くの少なくとも1つのスピーカーを制御するように構成されてもよい。いくつかの例では、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ことを判別し、コマンドへの返答を決定し、該返答を再生するために第1の音源位置の近くの少なくとも1つのスピーカーを制御し、コマンドを実行するように構成されてもよい。いくつかの例によれば、制御システムは、返答を再生するために第1の音源位置の近くの少なくとも1つのスピーカーを制御した後に、第1のオーディオ信号についての未変更のレンダリング・プロセスに戻るようにさらに構成されてもよい。

20

【0032】

いくつかの実装では、制御システムは、再生される第1のオーディオ・プログラム・ストリームおよび／または再生される第2のオーディオ・プログラム・ストリームについて、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、ラウドネス推定値を導出するように構成されてもよい。いくつかの例によれば、制御システムは、少なくとも部分的にはラウドネス推定値に基づいて、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。場合によっては、ラウドネス推定値は、知覚されるラウドネス推定値であってもよい。いくつかのそのような例によれば、レンダリング・プロセスを修正することは、干渉信号の存在下で第1のオーディオ信号および／または第2のオーディオ信号の知覚されるラウドネスを維持するために、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を変更することを含んでいてもよい。

30

【0033】

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ことを判別し、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によれば、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さらに構成されてもよい。いくつかのそのような例では、環境の前記位置は、介護者の推定された位置に対応してもよい。

40

【0034】

いくつかの例によれば、制御システムは、再生される第1のオーディオ・プログラム・ストリームおよび／または再生される第2のオーディオ・プログラム・ストリームについてのラウドネス推定値を導出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、少なくとも部分的にはラウドネス推定値に基づいて、第1のオーディオ信号および／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ようにさらに構成されてもよい。いくつかの例によれば、ラウドネス推定値は、知覚されるラウドネス推定値であってもよい。レンダリング・プロセスを修正することは、干渉信号の存在下で知覚されるラウドネスを維持するために、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を変更することを含んでいてもよい。

50

## 【0035】

いくつかの実装では、第1のオーディオ信号をレンダリングすることおよび／または第2のオーディオ信号をレンダリングすることは、任意に配置されたスピーカーへの柔軟なレンダリングを含むことができる。いくつかのそのような例では、柔軟なレンダリングは、質量中心振幅パン（Center of Mass Amplitude Panning）または柔軟仮想化（Flexible Virtualization）を含んでいてもよい。

## 【0036】

本開示の少なくともいくつかの側面は、一つまたは複数のオーディオ処理方法によって実装されてもよい。いくつかの例では、本方法は、少なくとも部分的に、本明細書に開示される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、第1のレンダリング・モジュールによって、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み、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。いくつかの例では、第1のオーディオ・プログラム・ストリームは、チャネル・データおよび／または空間メタデータを含む、第1の空間データを含む。いくつかのそのような方法は、第1のレンダリング・モジュールによって、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。

## 【0037】

いくつかのそのような方法は、第2のレンダリング・モジュールによって、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。いくつかの例では、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のレンダリング・モジュールによって、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第2のオーディオ信号をレンダリングして、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を含む。

## 【0038】

いくつかのそのような方法は、第1のレンダリング・モジュールによって、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・プロセスを修正し、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。いくつかのそのような方法は、第2のレンダリング・モジュールによって、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れらの特性のうちの少なくとも1つに少なくとも部分的に基づいて、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、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を含む。いくつかのそのような方法は、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し、混合されたオーディオ信号を、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を含む。

## 【0039】

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、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の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の一つまたは複数のラウドネスを修正することを含んでいてもよい。

## 【0040】

いくつかの例では、第2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

10

20

30

40

50

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んでいてもよい。

【0041】

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、スペクトル修正、可聴性ベースの修正および／またはダイナミックレンジ  
修正を実行することを含んでいてもよい。

【0042】

いくつかの方法は、第1のレンダリング・モジュールによって、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ことを含んでいてもよい。いくつかの方  
法は、第2のレンダリング・モジュールによって、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  
ることを含んでいてもよい。

【0043】

いくつかの方法は、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  
オ信号のうちの一つについて、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  
ちの一つについて、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一つにつ  
いてのレンダリング・プロセスを修正することを含んでいてもよい。

【0044】

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  
を有する非一時的媒体において実装されることができる。第1のレンダリング・モジュ  
ールによって、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み、第  
1のオーディオ・プログラム・ストリームは、環境の少なくとも一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第1のオーディオ信号を含む、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  
例では、第1のオーディオ・プログラム・ストリームは、チャンネル・データおよび／  
または空間メタデータを含む、第1の空間データを含む。いくつかのそのような方法  
は、第1のレンダリング・モジュールによって、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。

【0045】

いくつかのそのような方法は、第2のレンダリング・モジュールによって、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。いくつかの例では、第2のオー  
ディオ・プログラム・ストリームは、環境の少なくとも一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のレンダリング・モジュールによって、環境の少なくとも一つのスピー  
カーを介した再生のために第2のオーディオ信号をレンダリングして、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を含む。

【0046】

いくつかのそのような方法は、第1のレンダリング・モジュールによって、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一  
つ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  
ロセスを修正し、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。いくつかのそのような方法は、第2のレンダリング・モジュールによって、  
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れらの特性  
のうちの一つに少なくとも部分的に基づいて、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、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  
生成することを含む。

なくとも1つに少なくとも部分的に基づいて、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、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を含む。いくつかのそのような方法は、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し、混合されたオーディオ信号を、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を含む。

【0047】

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、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の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の一つまたは複数のラウドネスを修正することを含んでいてもよい。

【0048】

いくつかの例では、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んでいてもよい。

【0049】

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、スペクトル修正、可聴性ベースの修正および／またはダイナミックレンジ修正を実行することを含んでいてもよい。

【0050】

いくつかの方法は、第1のレンダリング・モジュールによって、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ことを含んでいてもよい。いくつかの方法は、第2のレンダリング・モジュールによって、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ことを含んでいてもよい。

【0051】

いくつかの方法は、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ことを含んでいてもよい。

【0052】

本明細書に記載される主題の一つまたは複数の実装の詳細は、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であろう。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】

【0053】

【図1A】本開示のさまざまな側面を実装できる装置の構成要素の例を示すブロック図である。

【図1B】ある実施形態の最小バージョンのブロック図である。

【図2A】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2B】図1A、図1Bまたは図2A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る方法の一例を概説するフロー図である。

【図2C】スピーカー・アクティブ化の例示的なセットを示す図である。

【図 2 D】スピーカー・アクティブ化およびオブジェクト・レンダリング位置の例示的なセットを示す図である。

【図 2 E】図 1A 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。

【図 2 F】ある例示的实施形態におけるスピーカー・アクティブ化のグラフである。

【図 2 G】ある例示的实施形態におけるオブジェクト・レンダリング位置のグラフである。

【図 2 H】ある例示的实施形態におけるスピーカー・アクティブ化のグラフである。

【図 2 I】ある例示的实施形態におけるオブジェクト・レンダリング位置のグラフである。

【図 2 J】ある例示的实施形態におけるスピーカー・アクティブ化のグラフである。

【図 2 K】ある例示的实施形態におけるオブジェクト・レンダリング位置のグラフである。

10

【図 3 A】接続された居住空間のフロアプランの例を示す。

【図 3 B】接続された居住空間のフロアプランの例を示す。

【図 4 A】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 4 B】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 5 A】開示されたマルチストリーム・レンダラーの第 3 の例示的な使用事例を示す。

【図 5 B】開示されたマルチストリーム・レンダラーの第 3 の例示的な使用事例を示す。

【図 5 C】開示されたマルチストリーム・レンダラーの第 3 の例示的な使用事例を示す。

【図 6】図 1B 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。

20

【図 7】図 2A 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 8】オーディオ・ストリーム・ラウドネス推定器を有するマルチストリーム・レンダリング・システムの実装を示す。

【図 9 A】複数のレンダリングされたストリームのクロスフェージングのために構成されたマルチストリーム・レンダリング・システムの例を示す。

【図 9 B】例示的实施形態におけるスピーカー・アクティブ化を示す点のグラフである。

【図 10】一例によるスピーカー・アクティブ化を示す点の間の三線形補間のグラフである。

【図 11】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。

【図 12 A】図 11 に示されてい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、基準空間モードにおいて、空間オーディオを柔軟にレンダリングする例を示す図の一つである。

30

【図 12 B】図 11 に示されてい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、基準空間モードにおいて、空間オーディオを柔軟にレンダリングする例を示す図の一つである。

【図 12 C】図 11 に示されてい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、基準空間モードにおいて、空間オーディオを柔軟にレンダリングする例を示す図の一つである。

【図 12 D】図 11 に示されてい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、基準空間モードにおいて、空間オーディオを柔軟にレンダリングする例を示す図の一つである。

40

【図 12 E】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。

【図 13 A】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グラフィカルユーザーインターフェース (GUI) の例を示す。

【図 13 B】ある例示的实施形態による分散式の空間的レンダリング・モードを示す。

【図 14 A】一例による、部分的に分散された空間的レンダリング・モードを示す。

【図 14 B】一例による、完全に分散された空間的レンダリング・モードを示す。

【図 15】2D 平面上の質量中心振幅パン (CMAP) および柔軟仮想化 (FV) レンダリング・システムについての例示的なレンダリング位置を示す。

50

【図16A】図15に表される分散式空間モードと図16Dに表される分散式空間モードとの間の中間的な分散式空間モードのさまざまな例を示す図の一つである。

【図16B】図15に表される分散式空間モードと図16Dに表される分散式空間モードとの間の中間的な分散式空間モードのさまざまな例を示す図の一つである。

【図16C】図15に表される分散式空間モードと図16Dに表される分散式空間モードとの間の中間的な分散式空間モードのさまざまな例を示す図の一つである。

【図16D】完全に分散されたレンダリング・モードを達成するために、図15のレンダリング・ポイントのすべてに適用されるワーピングの例を示す。

【図17】ユーザーがレンダリング・モードを選択できるGUIの例を示す。

【図18】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。

【図19】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20】図19に示される環境内の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21A】図19および図20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21B】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22】図1A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23】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。

【図24】順整列プロセスの一部の例を示す。

【図25】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26】逆整列プロセスの一部の例を示す。

【図27】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28】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29】図1A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30A】図29のいくつかのブロックの例を示す。

【図30B】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30C】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30D】図30C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31】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32】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33】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34】聴取環境の空間ゾーンの例を示す。

【図35】図34の空間ゾーン内のラウドスピーカーの例を示す。

【図36】図35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37】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。

【図38A】図2Cおよび図2Dの例に対応するラウドスピーカー参加値の例を示す。

【図38B】図2Cおよび図2Dの例に対応するラウドスピーカー参加値の例を示す。

【図38C】図2Cおよび図2Dの例に対応するラウドスピーカー参加値の例を示す。

【図39A】図2Fおよび図2Gの例に対応するラウドスピーカー参加値の例を示す。

【図39B】図2Fおよび図2Gの例に対応するラウドスピーカー参加値の例を示す。

【図39C】図2Fおよび図2Gの例に対応するラウドスピーカー参加値の例を示す。

【図40A】図2Hおよび図2Iの例に対応するラウドスピーカー参加値の例を示す。

【図40B】図2Hおよび図2Iの例に対応するラウドスピーカー参加値の例を示す。

【図40C】図2Hおよび図2Iの例に対応するラウドスピーカー参加値の例を示す。

10

20

30

40

50

【図 4 1 A】図 2J および図 2K の例に対応するラウドスピーカー参加値の例を示す。

【図 4 1 B】図 2J および図 2K の例に対応するラウドスピーカー参加値の例を示す。

【図 4 1 C】図 2J および図 2K の例に対応するラウドスピーカー参加値の例を示す。

【図 4 2】この例における居住空間である環境の図である。

【0054】

さまざまな図面における同様の参照番号および指示は、同様の要素を示す。

【発明を実施するための形態】

【0055】

柔軟レンダリング (flexible rendering) は、任意の数の任意に配置されたスピーカーを通じて空間的オーディオをレンダリングするための技法である。スマートオーディオ装置 (たとえば、スマートスピーカー) の家庭での普及に伴い、消費者が、スマートオーディオ装置を用いて、オーディオの柔軟なレンダリングと、そのレンダリングされたオーディオの再生とを実行できるようにする柔軟レンダリング技術を実現する必要がある。

【0056】

柔軟レンダリングを実現するために、質量中心振幅パン CEAP (Center of Mass Amplitude Panning, CMAP) および柔軟仮想化 (Flexible Virtualization, FV) を含むいくつかの技術が開発されている。これらの技術の両方は、レンダリング問題を、コスト関数最小化の問題としてキャストする。コスト関数は、レンダラーが達成しようとしている所望される空間的印象をモデル化する第 1 項と、スピーカーをアクティブ化することにコストを割り当てる第 2 項の 2 つの項から構成される。今日まで、この第 2 項は、レンダリングされているオーディオの所望される空間位置に近接したスピーカーのみがアクティブ化される疎な解決策を作り出すことに焦点を当ててきた。

【0057】

いくつかの実施形態は、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少なくとも 1 つ (たとえば、全部または一部) (またはスピーカーの別の集合のスピーカーの少なくとも 1 つ (たとえば、全部または一部) による)、オーディオの複数のストリームの再生を管理するための方法に関わる。

【0058】

あるクラスの実施形態は、複数の協調させられる (オーケストレーションされる) 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) による再生を管理する方法を含む。たとえば、ユーザーの家庭において (システム内に) 存在する一組のスマートオーディオ装置が、スマートオーディオ装置の全部または一部による (すなわち、全部または一部のスマートオーディオ装置のスピーカーによる) 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。

【0059】

(たとえば、多様な同時使用事例を扱うために家庭において) スマートオーディオ装置をオーケストレーションすることは、相互接続された一組のスピーカーを通じた、一つまたは複数のオーディオ・プログラム・ストリームの同時再生に関わってもよい。たとえば、ユーザーは、一組のスピーカーを通じて映画のアトモス・サウンドトラック (または、他のオブジェクトベースのオーディオ・プログラム) を聴いていてもよいが、その際、ユーザーは、関連するスマートアシスタント (または他のスマートオーディオ装置) に対してコマンドを発声してもよい。この場合、アトモス・ミックスの空間的呈示を話者 (話しているユーザー) の位置から離れ、最も近いスマートオーディオ装置から離れるようにワープさせ、一方、スマートオーディオ装置の (音声アシスタントの) の対応する応答の再生を、話者の位置のほうに同時にワープさせるように、システムによるオーディオ再生は (いくつかの実施形態に従って) 修正されてもよい。これは、コマンド (または対応するウェイクワード) の検出に応答して、オーディオ・プログラム・コンテンツの再生の音量を単に減少させることに比べて、重要な利点を提供することができる。同様に、ユーザーは、同じアトモス・サウンドトラックが隣接する開放された居住空間で再生されている間

10

20

30

40

50

に、キッチンで料理のヒントを得るためにスピーカーを使用することを望むことがありうる。この場合、いくつかの例によれば、アトモス・サウンドトラックは、キッチンから離れるようにワープされることができ、および／またはアトモス・サウンドトラックの一つまたは複数のレンダリングされる信号のラウドネスが、料理のヒントのサウンドトラックの一つまたは複数のレンダリングされた信号のラウドネスに応答して修正されることができ。さらに、いくつかの実装では、キッチンで再生される料理のヒントは、居住空間からしみ出してくることがありうるアトモス・サウンドトラックのどれよりも大きく、キッチンにいる人によって聞かれるように、動的に調整されることができる。

#### 【0060】

いくつかの実施形態は、上述の例示的な使用事例および考えられる他の多数を実装するように構成された複数ストリーム・レンダリング・システムに関わる。あるクラスの実施形態では、オーディオ・レンダリング・システムは、複数の任意に配置されたラウドスピーカーを通じて複数のオーディオ・プログラム・ストリームを同時に再生するように構成されてもよく、前記プログラム・ストリームの少なくとも1つは空間ミックスであり、前記空間ミックスのレンダリングは、一つまたは複数の追加的なプログラム・ストリームの同時再生に応答して（または関連して）動的に修正される。

#### 【0061】

いくつかの実施形態では、マルチストリーム・レンダラーは、上記に述べたシナリオおよび複数のオーディオ・プログラム・ストリームの同時再生が管理されなければならない多くの他の事例を実装するように構成されてもよい。マルチストリーム・レンダリング・システムのいくつかの実装は、以下の動作を実行するように構成されてもよい：

●複数のオーディオ・プログラム・ストリームを、複数の任意に配置されたラウドスピーカーを通じて同時にレンダリングおよび再生する。前記プログラム・ストリームの少なくとも1つが空間ミックスである。

○「プログラム・ストリーム」という用語は、全体として一緒に聴取されることを意図された一つまたは複数のオーディオ信号の集まりを指す。例は、音楽の選択、映画サウンドトラック、ポッドキャスト、ライブ・ボイス・コール、スマートアシスタントからの合成音声応答などを含む。

○空間ミックスは、聴取者の左耳と右耳で異なる（モノより多くの）信号を送達することを意図したプログラム・ストリームである。空間ミックスのオーディオ・フォーマットの例は、ステレオ、5.1および7.1サラウンドサウンド、ドルビー・アトモスなどのオブジェクト・オーディオ・フォーマット、およびアンビソニックスを含む。

○プログラム・ストリームのレンダリングは、特定の知覚的印象を達成するために、関連する一つまたは複数のオーディオ信号を、複数のラウドスピーカーにわたって能動的に分配するプロセスを指す。

●前記少なくとも1つの空間ミックスのレンダリングを、前記追加的なプログラム・ストリームのうちの一つまたは複数のレンダリングの関数として動的に修正する。空間ミックスのレンダリングに対するそのような修正の例は、以下を含むが、これらに限定されない：

○前記一つまたは複数の追加的なプログラム・ストリームのうちの少なくとも1つのレンダリングに関連するラウドスピーカーの相対的アクティブ化の関数として、複数のラウドスピーカーの相対的アクティブ化を修正する。

○前記一つまたは複数の追加的なプログラム・ストリームのうちの少なくとも1つのレンダリングの空間特性の関数として、空間ミックスの意図された空間バランスをワープさせること。

○前記一つまたは複数の追加的なプログラム・ストリームのうちの少なくとも1つのラウドネスまたは可聴性の関数として、空間ミックスのラウドネスまたは可聴性を修正する。

#### 【0062】

図1Aは、本開示のさまざまな側面を実装できる装置の構成要素の例を示すブロック図である。いくつかの例によれば、装置100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、スマート

10

20

30

40

50

オーディオ装置を含んでいてもよい。他の実装では、装置100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等のような、本明細書に開示される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、またはこれを含んでいてもよい。いくつかのそのような実装では、装置100は、サーバーであってもよく、サーバーを含んでいてもよい。いくつかの実装では、装置100は、本明細書で「オーディオ・セッション・マネージャ」と称されることがあるものを実装するように構成されてもよい。

#### 【0063】

この例では、装置100は、インターフェース・システム105および制御システム110を含む。インターフェース・システム105は、いくつかの実装では、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「アプリケーション」または単に「アプリ」と呼ばれることがある。インターフェース・システム105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データを交換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105は、いくつかの実装では、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105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、いくつかの例では、装置100がそれを用いて通信するように構成されるところの一つまたは複数のアプリケーションに関連してもよい。

#### 【0064】

インターフェース・システム105は、いくつかの実装では、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。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オーディオ信号を含んでいてもよい。オーディオ・プログラム・ストリームは、チャンネル・データおよび／または空間メタデータのような空間データを含んでいてもよい。インターフェース・システム105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入力を受領するように構成されてもよい。

#### 【0065】

インターフェース・システム105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。いくつかの実装によれば、インターフェース・システム105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105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、ディスプレイ・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105は、制御システム110と図1Aに示される任意的なメモリ・システム115のようなメモリ・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、いくつかの事例では、制御システム110がメモリ・システムを含んでいてもよい。

#### 【0066】

制御システム110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論理装置、離散的ゲートまたはトランジスタ論理、および／または離散的ハードウェアコンポーネントを含んでいてもよい。

#### 【0067】

いくつかの実装では、制御システム110は、2つ以上の装置に存在してもよい。たとえば、制御システム110の一部は、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

10

20

30

40

50

もよく、制御システム110の別の一部は、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、環境の外にある装置内に存在してもよい。他の例では、制御システム110の一部は、本明細書に示される環境の1つ中の装置内に存在してもよく、制御システム110の別の一部は、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は、環境の複数のスマートオーディオ装置にわたって分散されてもよく、または、オーケストレーション装置（たとえば、本明細書においてスマートホームハブと称されることがあるもの）および環境の一つまたは複数の他の装置によって分担されてもよい。インターフェース・システム105も、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。

【0068】

いくつかの実装では、制御システム110は、少なくとも部分的には、本明細書に開示される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110は、複数のスピーカーを通じたオーディオの複数のストリームの再生を管理する方法を実装するように構成されてもよい。

【0069】

本明細書に記載される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てもよい。そのような非一時的媒体は、ランダムアクセスメモリ（RAM）デバイス、読み出し専用メモリ（ROM）デバイスなどを含むがそ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。前記一つまたは複数の非一時的媒体は、たとえば、図1Aに示される任意的なメモリ・システム115および／または制御システム110に存在してもよい。よって、本開示に記載された主題のさまざまな革新的な側面は、ソフトウェアを記憶している一つまたは複数の非一時的媒体において実装できる。ソフトウェアは、たとえば、オーディオ・データを処理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んでいてもよい。ソフトウェアは、たとえば、図1Aの制御システム110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。

【0070】

いくつかの例では、装置100は、図1Aに示される任意的なマイクロフォン・システム120を含んでいてもよい。任意的なマイクロフォン・システム120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。いくつかの実装では、マイクロフォンの一つまたは複数のスピーカー・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、または別の装置と関連付けられてもよい。いくつかの例では、装置100はマイクロフォン・システム120を含まなくてもよいが、いくつかのそのような実装では、装置100は、それでも、インターフェース・システム110を介してオーディオ環境における一つまたは複数のマイクロフォンについてのマイクロフォン・データを受領するように構成されてもよい。

【0071】

いくつかの実装によれば、装置100は、図1Aに示される任意的なラウドスピーカー・システム125を含んでいてもよい。任意的なスピーカー・システム125は、一つまたは複数のスピーカーを含んでいてもよい。ラウドスピーカーは、本稿では「スピーカー」とも称される。いくつかの例では、任意的なスピーカー・システム125の少なくともいくつかのスピーカーは、任意に配置されうる。たとえば、任意的なスピーカー・システム125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、どの規格で規定されたラウドスピーカー・レイアウトにも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、任意的なスピーカー・システム125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、どの規格に規定されたラウドスピーカー・レイアウトにもない位置であってもよい。いくつかの例では、装置100は、ラウドスピーカー・システム125を含まなくてもよい。

10

20

30

40

50

## 【0072】

いくつかの実装では、装置100は、図1Aに示される任意的なセンサー・システム129を含んでいてもよい。任意的なセンサー・システム129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129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。いくつかの例では、任意的なセンサー・システム129の一つまたは複数のカメラは、単一目的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129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に存在してもよい。いくつかの例では、装置100は、センサー・システム129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、装置100は、それでも、インターフェース・システム110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・データを受信するように構成されてもよい。

10

## 【0073】

いくつかの実装では、装置100は、図1Aに示される任意的な表示システム135を含んでいてもよい。任意的な表示システム135は、一つまたは複数の発光ダイオード（LED）ディスプレイのような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135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100が表示システム135を含むいくつかの例では、センサー・システム129は、タッチセンサーシステムおよび／または表示システム135の一つまたは複数のディスプレイに近接する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110は、表示システム135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。

20

## 【0074】

いくつかのそのような例によれば、装置100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では、装置100は、ウェイクワード検出器であってもよく、または、ウェイクワード検出器を含んでいてもよい。たとえば、装置100は、バーチャル・アシスタントであってもよく、またはバーチャル・アシスタントを含んでいてもよい。

30

## 【0075】

図1B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、「レンダリング・モジュール」と称されてもよい。レンダリング・モジュールおよびミキサー〔混合器〕130a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130aは、図1Aを参照して上述した制御システム110の一例である制御システム110aを介して実装される。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うる。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

40

50

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。

【0076】

図2A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130bは、図1Aを参照して上述した制御システム110の一例である制御システム110bを介して実装される。このバージョンでは、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の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120a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。

【0077】

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。

【0078】

図2Bは、図1A、図1Bまたは図2A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る方法の一例を概説するフロー図である。方法200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200のブロックは、図1A、図1Bおよび図2Aに示される制御システム110、制御システム110aまたは制御システム110b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。

【0079】

この実装では、ブロック205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205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。

10

【0080】

この例によれば、ブロック210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200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200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。

【0081】

20

この例では、ブロック215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。

【0082】

この実装によれば、ブロック220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。

30

【0083】

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。

40

【0084】

よって、いくつかの実装ブロック210またはブロック220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

50

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。C<sub>MAP</sub>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】

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

10

【0085】

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。C<sub>MAP</sub>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：

20

【数 3】

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

【0086】

30

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、 $\mathbf{g}_{\text{opt}}$ 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】

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

40

【0087】

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 $C_{\text{spatial}}$ および $C_{\text{proximity}}$ の具体的な構築によって支配される。C<sub>MAP</sub>については、 $C_{\text{spatial}}$ 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 $\mathbf{g}$ の要素）によって重み付けした質量の中心：

【数 5】

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

50

に配置するモデルから導出される。

【0088】

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】

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

【0089】

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 $2 \times 1$ 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 $2 \times 1$ 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】

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

【0090】

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 $2 \times 1$ 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】

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

【0091】

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】

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

【0092】

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、 $A$ は $M \times M$ の正方行列、 $B$ は $1 \times M$ ベクトル、 $C$ はスカラーである。行列 $A$ 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 $\rightarrow o$ から離れた位置 $\rightarrow s_i$ 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

10

20

30

40

50

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。

【0093】

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】

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

10

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】

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

20

である。

【0094】

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】

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

30

ここで、

【数15】

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

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、 $\alpha$ および $\beta$ は調整可能なパラメータである。パラメータ $\alpha$ 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）、 $\beta$ 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。

40

【0095】

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16】

$$C(\mathbf{g}) = \mathbf{g}^* \mathbf{A} \mathbf{g} + \mathbf{B} \mathbf{g} + \mathbf{C} + \mathbf{g}^* \mathbf{D} \mathbf{g} = \mathbf{g}^* (\mathbf{A} + \mathbf{D}) \mathbf{g} + \mathbf{B} \mathbf{g} + \mathbf{C} \quad (10)$$

このコスト関数の

50

【数 1 7】

$g$

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】

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

【0 0 9 6】

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 0 9 7】

図 2C および図 2D は、スピーカー・アクティブ化およびオブジェクト・レンダリング位置の例示的なセットの例示的な集合を示す図である。これらの例では、スピーカー・アクティブ化およびオブジェクト・レンダリング位置は、4、64、165、−87、および−4 度のスピーカー位置に対応する。図 2C は、これらの特定のスピーカー位置についての式 11 に対する最適解を構成するスピーカー・アクティブ化 245a、250a、255a、260a および 265a を示す。図 2D は、個々のスピーカー位置を、スピーカー・アクティブ化 245a、250a、255a、260a および 265a にそれぞれ対応する四角 267、270、272、274 および 275 としてプロットしたものである。図 2D はまた、多数の可能なオブジェクト角度についての理想的なオブジェクト位置（言い換えれば、オーディオ・オブジェクトがレンダリングされるべき位置）をドット 276a として、それらのオブジェクトについての対応する実際のレンダリング位置を、点線 279a によって理想的なオブジェクト位置に接続されたドット 278a として示している。

【0 0 9 8】

あるクラスの実施形態は、複数の協調させられる（オーケストレーションされる）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）による再生のためにオーディオをレンダリングする方法に関わる。たとえば、ユーザーの家庭（におけるシステム）にある一組のスマートオーディオ装置が、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てもよい。かかる使用事例は、スマートオーディオ装置の全部または一部による（すなわち、全部または一部のスピーカーによる）再生のための、オーディオの（ある実施形態に従った）レンダリングを含む。システムとの多くの対話が考えられており、それはレンダリングに対する動的な修正を要求する。そのような修正は、必ずではないが、空間的忠実度に焦点を当てていてもよい。

【0 0 9 9】

いくつかの実施形態は、スマートオーディオ装置の集合の少なくとも 1 つ（たとえば、全部または一部）のスマートオーディオ装置による再生のために（またはスピーカーの別の集合の少なくとも 1 つ（たとえば、全部または一部）のスピーカーによる再生のために）オーディオをレンダリングするための方法である。レンダリングは、コスト関数の最小化を含んでいてもよく、コスト関数は、少なくとも 1 つの動的なスピーカー・アクティブ化項を含む。そのような動的なスピーカー・アクティブ化項の例は、以下を含む（ただし、これらに限定されない）：

- ・一または複数の聴取者へのスピーカーの近接性；
- ・引力または反発力に対するスピーカーの近接性；
- ・何らかの位置（たとえば、聴取者位置またはベビールーム）に関するスピーカーの可聴性；
- ・スピーカーの能力（周波数応答、歪み）；
- ・他のスピーカーに対するスピーカーの同期；
- ・ウェイクワード性能；および
- ・エコー・キャンセラ性能。

【0100】

動的なスピーカー・アクティブ化項は、多様な挙動の少なくとも1つを有効にしうる。かかる挙動は、特定のスマートオーディオ装置から離れたオーディオの空間的提示を歪めて、そのマイクロフォンが話者をより良好に聞くことができるようにする、または該スマートオーディオ装置のスピーカーから二次オーディオ・ストリームがより良好に聞こえるようにすることを含む。

【0101】

いくつかの実施形態は、複数の協調させられる（オーケストレーションされる）スマートオーディオ装置のスピーカーによる再生のためのレンダリングを実装する。他の実施形態は、スピーカーの別の集合のスピーカー（単数または複数）による再生のためのレンダリングを実装する。

【0102】

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。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。

【数20】

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

【0103】

式12において、項

【数21】

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

は、追加的なコスト項を表し、

【数22】

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

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】

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

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】

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

10

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】

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

は、

【数 2 6】

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

20

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】

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

が、少なくとも、

【数 2 8】

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

30

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。

【0 1 0 4】

【数 2 9】

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

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）。

40

【0 1 0 5】

【数 3 0】

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

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；

50

- ・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、互いとの相対的な同期。

【0106】

【数31】

{ $\hat{e}$ }

10

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。

【0107】

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。

20

【0108】

図2Eは、図1Aに示される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280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。方法280のブロックは、図1Aに示される制御システム110のような制御システムであってもよい（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。

【0109】

30

この実装では、ブロック285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、オーディオ・データを受領することに関わる。この例では、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む。この実装によれば、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る空間位置を示す。いくつかの事例において、意図された知覚される空間位置は、明示的であってもよく、たとえば、ドルビー・アトモス位置メタデータのような位置メタデータによって示されてもよい。他の事例では、意図された知覚される空間位置は暗黙的であってもよく、たとえば、意図された知覚される空間位置は、ドルビー5.1、ドルビー7.1、または他のチャンネルベースのオーディオ・フォーマットに従ってチャンネルに関連付けられる想定される位置であってもよい。いくつかの例では、ブロック285は、制御システムのレンダリング・モジュールが、インターフェース・システムを介してオーディオ・データを受領することに関わる。

40

【0110】

この例によれば、ブロック290は、環境の一組のラウドスピーカーを介した再生のために、制御システムによってオーディオ・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。この例では、オーディオ・データに含まれる一つまたは複数のオーディオ信号のそれぞれをレンダリングすることは、コスト関数を最適化することによって、環境内の一組のラウドスピーカーの相対的アクティブ化を決定することに関わる。この例によれば、コストは、環境中の一組のラウドスピーカーで再生されたときのオーディオ信号の知覚される空間位置のモデルの関数である。この例では、コス

50

トはまた、オーディオ信号の意図された知覚される空間位置の、一組のラウドスピーカーのうちの各ラウドスピーカーの位置への近接性の指標の関数でもある。この実装では、コストは、一つまたは複数の追加的な動的に構成可能な機能の関数でもある。この例では、動的に構成可能な機能は、以下のうちの一つまたは複数に基づく：一または複数の聴取者へのラウドスピーカーの近接性；引力位置へのラウドスピーカーの近接性、ここで、引力は、引力位置により近いラウドスピーカーの、相対的により高いアクティブ化を優遇する因子である；反発力の位置へのラウドスピーカーの近接性、ここで、反発力は、反発力位置により近いラウドスピーカーの、相対的により低いアクティブ化を優遇する因子である；各ラウドスピーカーの、環境中の他のラウドスピーカーに比した能力；ラウドスピーカーの、他のラウドスピーカーに対する同期；ウェイクワード性能；またはエコー・キャンセラ性能。

10

#### 【0111】

この例において、ブロック295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環境の前記一組のラウドスピーカーの少なくともいくつかのラウドスピーカーに提供することに関わる。

#### 【0112】

いくつかの例によれば、知覚される空間位置のモデルは、聴取者の左耳および右耳におけるオーディオ・オブジェクト位置に対応するバイノーラル応答を生成することができる。代替的または追加的に、知覚される空間位置のモデルは、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、前記一組のラウドスピーカーの位置に、ラウドスピーカーの関連付けられたアクティブ化利得によって重み付けしたものの質量中心に置くことができる。

20

#### 【0113】

いくつかの例では、前記一つまたは複数の追加的な動的に構成可能な機能は、少なくとも部分的には、前記一つまたは複数のオーディオ信号のレベルに基づくことができる。いくつかの事例では、前記一つまたは複数の追加的な動的に構成可能な機能は、少なくとも部分的には、前記一つまたは複数のオーディオ信号のスペクトルに基づくことができる。

#### 【0114】

方法280のいくつかの例は、スピーカー・レイアウト情報を受領することに関わる。いくつかの例では、前記一つまたは複数の追加的な動的に構成可能な機能は、少なくとも部分的には、環境中の各ラウドスピーカーの位置に基づくことができる。

30

#### 【0115】

方法280のいくつかの例は、ラウドスピーカー仕様情報を受領することに関わる。いくつかの例では、前記一つまたは複数の追加的な動的に構成可能な機能は、少なくとも部分的には、各ラウドスピーカーの能力に基づくことができ、該能力は、周波数応答、再生レベル限界または一つまたは複数のラウドスピーカー・ダイナミクス処理アルゴリズムのパラメータのうちの一つまたは複数を含むことができる。

#### 【0116】

いくつかの例によれば、前記一つまたは複数の追加的な動的に構成可能な機能は、少なくとも部分的には、各ラウドスピーカーから他のラウドスピーカーへの音響伝達の測定または推定に基づくことができる。代替的または追加的に、前記一つまたは複数の追加的な動的に構成可能な機能は、少なくとも部分的には、環境中の一つまたは複数の人の聴取者またはスピーカーの位置に基づくことができる。代替的または追加的に、前記一つまたは複数の追加的な動的に構成可能な機能は、少なくとも部分的には、各ラウドスピーカーから聴取者またはスピーカー位置への音響伝達の測定または推定に基づくことができる。音響伝達の推定値は、たとえば、少なくとも部分的には、各ラウドスピーカーと聴取者またはスピーカー位置との間に存在しうる壁、家具または他の物体に基づいてもよい。

40

#### 【0117】

代替的または追加的に、前記一つまたは複数の追加的な動的に構成可能な機能は、少なくとも部分的には、環境中の一つまたは複数の非ラウドスピーカー・オブジェクトまたは

50

ランドマークのオブジェクト位置に基づくことができる。いくつかのそのような実装では、前記一つまたは複数の追加的な動的に構成可能な機能は、少なくとも部分的には、各ラウドスピーカーからオブジェクト位置またはランドマーク位置への音響伝達の測定または推定に基づくことができる。

#### 【0118】

柔軟なレンダリングを実現するために一つまたは複数の適切に定義された追加的なコスト項を採用することによって、多くの新しい有用な挙動が達成されうる。下記に挙げるすべての例示的な挙動は、望ましくないと思なされるある種の条件の下で、ある種のラウドスピーカーにペナルティを与えるという形で作成されている。最終的な結果は、これらのラウドスピーカーが、オーディオ信号の前記集合の空間的レンダリングにおいて、より少なくアクティブ化されるということである。これらの多くの場合、空間的レンダリングの修正とは無関係に、望ましくないラウドスピーカーを単に小さくすることを考えてもよいが、そのような戦略は、オーディオ・コンテンツの全体的なバランスを著しく劣化させる可能性がある。ミックスのある種の成分は、たとえば、完全に聞こえなくなることがある。他方、開示された実施形態では、レンダリングのコア最適化にこれらのペナルティ付与を統合することにより、レンダリングは、適応し、よりペナルティの低い残りのスピーカーを用いて可能な限り最良の空間レンダリングを実行することができる。これは、はるかにエレガントで、適応性があり、効果的な解決策である。

#### 【0119】

例示的な使用事例は、以下を含むが、それらに限定されない。

#### 【0120】

●聴取エリアのまわりで、よりバランスのとれた空間的呈示を提供する

○空間的オーディオは、意図された聴取エリアからほぼ同じ距離であるラウドスピーカーを通じて最も良く提示されることが分かった。コストは、聴取エリアまでのラウドスピーカーの平均距離よりも著しく近い、または離れているラウドスピーカーがペナルティを受け、それによりそのアクティブ化が低減されるように構築されてもよい。

#### 【0121】

●オーディオを聴取者または話者から遠ざける、または聴取または話者に向かって近づける

○システムのユーザーが、システムの、またはシステムに付随するスマート音声アシスタントに話しかけようとしている場合、話者により近いラウドスピーカーにペナルティを与えるコストを作ることが有益である。このようにして、これらのラウドスピーカーは、より少なくアクティブ化され、関連するマイクロフォンが話者をよりよく聞くことができるようになる。

○聴取空間内の他の聴取者のための再生レベルを最小化する、単一の聴取者のための、より内密な体験を提供するために、聴取者の位置から遠いスピーカーは大きなペナルティを受けてもよい。それにより、聴取者に最も近いスピーカーのみが最も顕著にアクティブ化される。

#### 【0122】

●オーディオをランドマーク、ゾーン、またはエリアから遠ざける、またはランドマーク、ゾーン、またはエリアに近づける

○聴取空間の近傍におけるある種の位置、たとえばベビールーム、ベビーベッド、オフィス、読書エリア、勉強エリアなどは、センシティブであると考えられてもよい。そのような場合、この位置、ゾーンまたはエリアに近いスピーカーの使用にペナルティを与えるコストが構築されてもよい。

○あるいはまた、上記と同じ場合（または類似の場合）について、スピーカーのシステムは、特に、スピーカーのうちの1つ（取り付けられたまたは付随するマイクロフォンを備えたもの）がベビールーム自体内に存在する場合に、各スピーカーからベビールームへの音響伝達の測定値を生成していてもよい。この場合、スピーカーのベビールームへの物理的な近接性を使うのではなく、ベビールームへの測定された音響伝達が高いスピーカ

10

20

30

40

50

ーの使用にペナルティを課すコストが構築されてもよい。および／または

【0123】

●スピーカーの能力の最適な使用

○異なるラウドスピーカーの能力は、著しく異なることがある。たとえば、ある人気のスマートスピーカーは、限られた低域能力をもつ単一の1.6インチのフルレンジ・ドライバを含むだけである。他方は、別のスマートスピーカーは、ずっと能力の高い3インチのウーファーを含む。これらの能力は、一般に、スピーカーの周波数応答に反映され、よって、スピーカーに関連する応答の集合は、コスト項において利用されうる。ある特定の周波数では、周波数応答によって測定されるところにより他のスピーカーに比して能力が劣るスピーカーはペナルティを受け、よって、より低い程度にアクティブ化される。いくつかの実装では、そのような周波数応答値は、スマート・ラウドスピーカーに記憶されてもよく、次いで、柔軟レンダリングを最適化することを受け持つ計算ユニットに報告されうる。

10

【0124】

○多くのスピーカーは複数のドライバを含んでおり、それぞれが異なる周波数範囲を再生することを受け持つ。たとえば、ある人気のスマートスピーカーは、低周波のためのウーファーと高周波のためのツイーターを含むツウウェイ設計である。典型的には、そのようなスピーカーは、フルレンジの再生オーディオ信号を適切な周波数範囲に分割し、それぞれのドライバに送るためのクロスオーバー回路を含む。あるいはまた、そのようなスピーカーは、各個々のドライバに柔軟レンダラー再生アクセスを提供するとともに、周波数応答のような各個々のドライバの能力に関する情報を提供することができる。上述のようなコスト項を適用することによって、いくつかの例において、柔軟レンダラーは、異なる周波数におけるそれらの相対的な能力に基づいて、2つのドライバ間のクロスオーバーを自動的に構築することができる。

20

【0125】

○周波数応答の上述の使用例は、スピーカーの固有の能力に焦点を当てているが、聴取環境に置かれたスピーカーの能力を正確には反映していない可能性がある。ある種の場合には、意図される聴取位置で測定されたスピーカーの周波数応答は、何らかの較正手順を通じて利用可能であってもよい。そのような測定値は、スピーカーの使用をより最適化するために、あらかじめ計算された応答の代わりに使用されてもよい。たとえば、ある種のスピーカーは、本質的に、特定の周波数で非常に能力があるかもしれないが、その配置（たとえば、壁または家具の後ろ）のために、意図された聴取位置において非常に限定された応答を生じることがありうる。この応答を捕捉し、適切なコスト項に入力される測定値は、そのようなスピーカーの有意なアクティブ化を防止することができる。

30

【0126】

○周波数応答は、ラウドスピーカーの再生能力の1つの側面に過ぎない。多くの小型ラウドスピーカーは、再生レベルが高くなるにつれて歪み始め、その後、特に低周波数の場合には可動域限界（excursion limit）に達する。そのような歪みを減らすために、多くのラウドスピーカーは、再生レベルを、周波数にわたって可変でありうるいくつかの制限閾値未満に制約するダイナミクス処理（dynamics processing）を実装する。あるスピーカーがこれらの閾値に近接している、または閾値にあり、柔軟レンダリングに参加している他のスピーカーはそうでない場合、制限となるスピーカーの信号レベルを低減し、このエネルギーを他の、負担がより少ないスピーカーに向けることは理にかなっている。そのような挙動は、関連するコスト項を適正に構成することによって、いくつかの実施形態に従って自動的に達成できる。そのようなコスト項は、以下の一つまたは複数に関わってもよい：

40

- ・ラウドスピーカーの制限閾値に関連したグローバル再生音量の監視。たとえば、音量レベルがその制限閾値により近いラウドスピーカーは、より大きなペナルティが課されてもよい；

- ・可能性としては周波数にわたって変化する動的な信号レベルの、やはり可能性と

50

しては周波数にわたって変化するラウドスピーカーの制限閾値との関連での、監視。たとえば、監視された信号レベルがその制限閾値により近いラウドスピーカーは、より大きなペナルティが課されてもよい；

- ・制限利得などの、ラウドスピーカーのダイナミクス処理のパラメータの直接監視。いくつかのそのような例において、パラメータがより強い制限を示すラウドスピーカーは、より大きなペナルティが課されてもよい；および／または、

- ・ラウドスピーカーが線形範囲で動作しているかどうかを判定するための、増幅器によってラウドスピーカーに送達されている実際の瞬間的な電圧、電流、電力の監視。たとえば、より低い直線性で動作しているラウドスピーカーが、より大きなペナルティが課されてもよい。

【0127】

○統合されたマイクロフォンおよび対話的な音声アシスタントを有するスマートスピーカーは、典型的には、何らかのタイプのエコーキャンセレーションを使用して、録音マイクロフォンによってピックアップされる、スピーカーから再生されるオーディオ信号のレベルを低減する。この低減が大きいほど、スピーカーがスペース内の話者を聞いて理解する可能性が高くなる。エコー・キャンセラの残差が一貫して高い場合、これはスピーカーが、エコー経路の予測が困難になる非線形領域に駆動されていることの指標でありうる。そのような場合、信号エネルギーをそのスピーカーから逸らすことが理にかなっており、よって、エコー・キャンセラ性能を考慮したコスト項が有益でありうる。そのようなコスト項は、付随するエコー・キャンセラが貧弱な性能を発揮しているスピーカーに、高いコストを割り当ててもよい。

【0128】

○複数のラウドスピーカーで空間的オーディオをレンダリングするときに予測可能な像形成（イメージング）を達成するためには、一般に、一組のラウドスピーカーでの再生が、時間を通じて合理的に同期されることが必要である。有線ラウドスピーカーの場合、これは当然のことであるが、無線ラウドスピーカーが多数ある場合、同期は困難であり、最終結果は可変であることがある。そのような場合、各ラウドスピーカーがターゲットとの同期の相対的な度合いを報告することが可能であることがあり、この度合いが同期コスト項に入力されてもよい。いくつかのそのような例では、より低い同期度のラウドスピーカーは、より大きなペナルティを課され、よってレンダリングから除外されうる。さらに、ある種のタイプのオーディオ信号、たとえば、拡散的または非方向性であることが意図されるオーディオ・ミックスの成分に対しては、厳密な同期が必要とされない場合がある。いくつかの実装では、成分は、メタデータを用いてそのようなものとしてタグ付けされてもよく、同期コスト項は、ペナルティが低減されるように修正されてもよい。

【0129】

次に、実施形態の例を記載する。

【0130】

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】

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

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】

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

ここで、 $\mathbf{W}_j$ は、項jについてスピーカー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】

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

10

【 0 1 3 1】

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】

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

【 0 1 3 2】

20

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 1 3 3】

重み項 $w_{ij}$ のそれぞれを、ラウドスピーカー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

30

【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ーまでの距離である。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ー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】

40

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

ここで、 $\alpha_j$ 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、 $\tau_j$ 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、

【数 4 0】

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

50

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、 $\alpha_j$ 、 $\beta_j$ 、 $\tau_j$ 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をも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 $C_{\text{spatial}}$ 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 1 3 4】

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る：

【数 4 2】

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

【0 1 3 5】

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 1 3 6】

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張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 $\tau_j$ 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】

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

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

【0 1 3 7】

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 4 5】

 $\vec{l}_j$ 

( $\rightarrow l_j$ ) 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。 $\alpha_j$ 、 $\beta_j$ および $\rightarrow l_j$ のこれらの値は単に例である。他の実装では、 $\alpha_j$ は1～100の範囲であってもよく、 $\beta_j$ は1～25の範囲であってもよい。

図2Fは、例示的实施形態におけるスピーカー・アクティブ化のグラフである。この例では、図2Fは、図1および図2からの同じスピーカー位置についてのコスト関数に対する最適解を構成するスピーカー・アクティブ化245b、250b、255b、260b、および265bを示し、 $w_{ij}$ によって表される引力を加えたものである。図2Gは、ある例示的实施形態におけるオブジェクト・レンダリング位置のグラフである。この例では、図2Gは、多数の可能なオブジェクト角度についての対応する理想的なオブジェクト位置276bと、点線279bによって理想的なオブジェクト位置276bに接続された、それらのオブジェクトについての対応する実際のレンダリング位置278bとを示している。固定位置

【数 4 6】

 $\vec{l}_j$ 

に向かう実際のレンダリング位置1135bの曲がった (skewed) 配向は、コスト関数への最適解に対するアトラクター重み付けの影響を示す。

【0 1 3 8】

第2および第3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。いくつかの例では、反発力は、オフィスエリア、読書エリア、ベッドまたは寝室エリア（たとえば、ベビーベッドまたは寝室）などの聴覚環境のエリアまたはゾーンから離れるようにオーディオを押すために使用されてもよい。いくつかのそのような例によれば、特定の位置が、ゾーンまたはエリアを代表するものとして使用されてもよい。たとえば、乳児のベッドを表す位置は、乳児の頭部の推定位置、乳児に対応する推定音源位置などでありうる。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

【数 4 7】

 $\vec{l}_j$ 

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

【数 4 8】

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

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

【0 1 3 9】

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 4 9】

 $\vec{l}_j$

$\{\rightarrow l_j\}$  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。 $\alpha_j$ 、 $\beta_j$ および $\rightarrow l_j$ のこれらの値は単に例である。上記のように、いくつかの例では、 $\alpha_j$ は1～100の範囲であってもよく、 $\beta_j$ は1～25の範囲であってもよい。図2Hは、例示的实施形態におけるスピーカー・アクティブ化のグラフである。この例によれば、図2Hは、前の諸図からの同じスピーカー位置についてのコスト関数に対する最適解を構成するスピーカー・アクティブ化245c、250c、255c、260c、および265cを示し、 $w_{ij}$ によって表される反発力を加えたものである。図2Iは、ある例示的实施形態におけるオブジェクト・レンダリング位置のグラフである。この例では、図2Iは、多数の可能なオブジェクト角度についての理想的なオブジェクト位置276cと、点線279cによって理想的なオブジェクト位置276cに接続された、それらのオブジェクトについての対応する実際のレンダリング位置278cとを示している。固定位置

【数50】

$\vec{l}_j$

から遠ざかる実際のレンダリング位置235cの曲がった (skewed) 配向は、コスト関数への最適解に対する反発体重み付けの影響を示す。

【0140】

第3の例示的な使用事例は、睡眠中のベビールームへのドアのような、音響的に敏感なランドマークからオーディオを遠ざけるように「押す」ことである。前の例と同様に、 $\rightarrow l_j$ を、180度のドア位置（プロットの下部中央）に対応するベクトルに設定する。より強い反発力を達成し、音場を主要聴取空間の前方部に完全に偏らせるために、我々は $\alpha_j=20$ 、 $\beta_j=5$ に設定した。図2Jは、ある例示的实施形態におけるスピーカー・アクティブ化のグラフである。ここでもまた、この例では、図2Jは、スピーカー位置の同じ集合への最適解を構成するスピーカー・アクティブ化245d、250d、255d、260dおよび265dを示し、より強い反発力を加えている。図2Kは、例示的实施形態におけるオブジェクト・レンダリング位置のグラフである。ここでもまた、この例では、図2Kは、多数の可能なオブジェクト角度についての理想的なオブジェクト位置276dと、点線279dによって理想的なオブジェクト位置276dに接続された、それらのオブジェクトについての対応する実際のレンダリング位置278dとを示している。実際のレンダリング位置278dの曲がった (skewed) 配向は、コスト関数への最適解に対する、より強い反発体重み付けの影響を示す。

【0141】

ここで図2Bに戻ると、この例では、ブロック225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む。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。

【0142】

この例によれば、ブロック230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230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。

【0143】

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。

【0144】

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。

【0145】

図2Bに示される例では、ブロック235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235は、たとえば、図2Aに示されるミキサー130bによって実行されてもよい。

【0146】

この例によれば、ブロック240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200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。

【0147】

図2B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

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。

#### 【0148】

図1Aおよび図2A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。

#### 【0149】

図2A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。

#### 【0150】

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に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。

#### 【0151】

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

10

20

30

40

50

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。

10

#### 【0152】

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。

20

#### 【0153】

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。

30

#### 【0154】

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。

40

#### 【0155】

図1A、図1B、および図2A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3Aおよび3Bを参照して、いくつかの例が記述される。

#### 【0156】

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

50

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3Aおよび図3B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300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305a~305h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3A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310およびキッチン315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330に面してカウチ325に座った聴取者320a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335aによって視覚的に表現される。

10

#### 【0157】

図3Bでは、料理のヒントが同時にレンダリングされ、キッチン315内の単一のラウドスピーカー305gを通じて第2の聴取者320b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305gから出るクラウド340によって視覚的に表される。これらの料理のヒントが、図3A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315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315）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3Bにおける、より小さなクラウド335b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3A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3Aおよび図3B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。

20

30

40

#### 【0158】

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

50

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。

#### 【0159】

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1B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2A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

10

20

30

40

50

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。

【0160】

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

○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。

【0161】

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：

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。

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。

【0162】

図4Aおよび4B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、または全く小さくされなくてもよいことを意味する。図4Aおよび図4B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4Aのユーザー・クラウド335c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320c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305dおよび／またはスピーカー305f）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320cの位置またはユーザー320c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4Bのクラウド335d)。

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4Bのクラウド440)。

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4Aのクラウド335c)。

#### 【0163】

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。

#### 【0164】

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。

#### 【0165】

図5A、図5B、および図5C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5Aは、空間的な音楽ミックス(クラウド335e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310およびキッチン315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5Bでは、赤ん坊510は、右下に描かれた隣接する寝室505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335f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540によって表される)が、ベビールーム内のスピーカー305h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30

10

20

30

40

50

5h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310にいる、親または他のケア担当者でありうる聴取者320dの近くで同時に再生されてもよい。図5Cは、クラウド550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5Bのクラウド335fの形状に対するクラウド335g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320d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：

【0166】

●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。

【0167】

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。

【0168】

図1B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。

【0169】

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、 $N \times M$ 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。

【0170】

図6は、図1B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630a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635a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ル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630

10

20

30

40

50

a、および逆フィルタバンク635aは、制御システム110cの構成要素である。

【0171】

図7は、図2A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6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120b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630b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635b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630b、および逆フィルタバンク635bは、制御システム110dの構成要素である。

【0172】

周波数領域における階層的アプローチの別の利点は、各オーディオ・ストリームの知覚されるラウドネスの計算、および、他のオーディオ・ストリームの一つまたは複数を動的に修正する際にこの情報を使用することにある。この実施形態を例解するために、図3Aおよび図3Bを参照して上述した上記の例を検討する。この場合、2つのオーディオ・ストリーム（N=2）、空間的な映画サウンドトラック、および料理のヒントがある。また、K個のマイクロフォンのうちの一つまたは複数によって拾われる、キッチン内のブレンダーによって生成される環境ノイズがあることがある。

【0173】

各オーディオ・ストリームsが個別にレンダリングされ、各マイクロフォンiが捕捉され、周波数領域に変換された後、ソース励起信号 $E_s$ または $E_i$ が計算できる。これは、各オーディオ・ストリームsまたはマイクロフォン信号iの知覚されるラウドネスの時間変化する推定値としてはたらく。この例では、これらのソース励起信号は、レンダリングされたストリームまたは捕捉されたマイクロフォンから、オーディオ・ストリームについては変換係数 $X_s$ 、またはマイクロフォン信号については $X_i$ を介して、c個のラウドスピーカーのために時間tにわたってb個の周波数帯域について計算され、周波数依存の時定数 $\lambda_b$ を用いて平滑化される；

【数51】

$$E_s(b, t, c) = \lambda_b E_s(b, t - 1, c) + (1 - \lambda_b) |X_s(b, t, c)|^2 \quad (20a)$$

$$E_i(b, t, c) = \lambda_b(b) E_i(b, t - 1, c) + (1 - \lambda_b) |X_i(b, t, c)|^2 \quad (20b)$$

【0174】

生のソース励起は、特定の位置での各ストリームの知覚されるラウドネス推定値である。空間的ストリームについては、その位置は、図3Bのクラウド335bの中央にあるが、料理のヒントのストリームについては、それはクラウド340の中央にある。マイクロフォンによってピックアップされるブレンダー・ノイズについての位置は、たとえば、ブレンダー・ノイズの源に最も近いマイクロフォン（単数または複数）の特定の位置に基づいてよい。

【0175】

生のソース励起は、各ターゲット・オーディオ・ストリームの聴取位置において、ノイズとしてどのくらい知覚可能であるか推定するために、それらによって修正されるオーディオ・ストリーム（単数または複数）の聴取位置に変換されなければならない。たとえば、オーディオ・ストリーム1が映画サウンドトラックであり、オーディオ・ストリーム2が料理のヒントである場合、

【数52】

10

20

30

40

50

$$\hat{E}_{12}$$

〔以下、 $\hat{E}_{12}$ のように $\hat{\phantom{x}}$ を前に書くことがある〕は、変換された（ノイズ）励起となる。この変換は、各周波数帯域 $b$ についての各ラウドスピーカー $c$ の関数として、ソース・オーディオ・ストリーム $s$ からターゲット・オーディオ・ストリーム $x$ への可聴性スケール因子 $A_{xs}$ またはマイクロフォン $i$ からターゲット・オーディオ・ストリーム $x$ への $A_{xi}$ を適用することによって計算される。 $A_{xs}$ および $A_{xi}$ の値は、時間とともに変化する距離比または実際の可聴性の推定値を使用して、決定されうる。

【数 5 3】

$$\hat{E}_{xs}(b, t, c) = A_{xs}(b, t, c)E_s(b, t, c) \quad (21a)$$

$$\hat{E}_{xi}(b, t, c) = A_{xi}(b, t, c)E_i(b, t, c) \quad (21b)$$

【0 1 7 6】

式13aにおいて、 $\hat{E}_{xs}$ は、マイクロフォン入力を参照することなく、ソース・オーディオ・ストリームについて計算された生ノイズ励起を表す。式13bにおいて、 $\hat{E}_{xi}$ は、マイクロフォン入力を参照して計算された生のノイズ励起を表す。この例によれば、生のノイズ励起 $\hat{E}_{xs}$ または $\hat{E}_{xi}$ は、次いで、ストリーム1～ $N$ 、マイクロフォン1～ $K$ 、および出力チャンネル1～ $M$ にわたって合計され、ターゲット・ストリーム $x$ についての全ノイズ推定値が得られる：

【数 5 4】

$$\hat{E}_x(b, t) = \sum_{c=1}^M \left( \sum_{s=1}^N \hat{E}_{xs}(b, t, c) + \sum_{i=1}^K \hat{E}_{xi}(b, t, c) \right) \quad (22)$$

【0 1 7 7】

いくつかの代替的な実装によれば、式14の項

【数 5 5】

$$\sum_{i=1}^K \hat{E}_{xi}(b, t, c)$$

を省略することによって、マイクロフォン入力を参照することなく、全ノイズ推定値を得ることができる。

【0 1 7 8】

この例では、ターゲット・ストリームをあまりに急に修正することによって引き起こされる知覚可能なアーチファクトを回避するために、全体の生のノイズ推定値が平滑化される。この実装によれば、平滑化は、オーディオ圧縮器と同様に、速いアタックと遅いリリースを使用する概念に基づく。ターゲット・ストリーム $x$ についての平滑化されたノイズ推定値

【数 5 6】

$$\bar{E}_x$$

〔以下、 $\bar{E}_x$ のように $\bar{\phantom{x}}$ を前に書くことがある〕は、この例では、次のように計算される：

【数 5 7】

10

20

30

40

50

$$\bar{E}_x(b, t) = \lambda(b, t) \hat{E}_x(b, t) + (1 - \lambda(b, t)) \bar{E}_x(b, t - 1) \quad (23)$$

$$\lambda(b, t) = \begin{cases} \lambda_{fast}, & \hat{E}_x(b, t) > \bar{E}_x(b, t) \\ \lambda_{slow}, & \hat{E}_x(b, t) \leq \bar{E}_x(b, t) \end{cases} \quad (24)$$

【0179】

ひとたびストリームxについて完全なノイズ推定値  $\bar{E}_x(b, t)$  を得たら、以前に計算されたソース励起信号  $E_x(b, t, c)$  を再利用して、ターゲット・オーディオ・ストリームxがノイズより大きく可聴なままであることを保証よう、ターゲット・オーディオ・ストリームxに適用するための時間変化する利得  $G_x(b, t, c)$  の集合を決定することができる。これらの利得は、多様な技法のいずれかを用いて計算できる。

【0180】

ある実施形態では、人間のラウドネス知覚におけるさまざまな非線形性をモデル化し、知覚されるラウドネスの時間変化する周波数分布を記述する特定ラウドネス信号を計算するために、ラウドネス関数  $L\{\cdot, \cdot\}$  が励起に適用されることができ。ノイズ推定値とレンダリングされたオーディオ・ストリームxについての励起に  $L\{\cdot, \cdot\}$  を適用すると、各信号の特定ラウドネス (specific loudness) についての推定値が得られる：

【数58】

$$L_{xn}(b, t) = L\{\bar{E}_x(b, t)\} \quad (25a)$$

$$L_x(b, t, c) = L\{E_x(b, t, c)\} \quad (25b)$$

【0181】

式17aにおいて、 $L_{xn}$ は、ノイズの特定ラウドネスについての推定値を表し、式17bにおいて、 $L_x$ は、レンダリングされたオーディオ・ストリームxの特定ラウドネスについての推定値を表す。これらの特定ラウドネス信号は、信号が孤立して聴取されたときの知覚されるラウドネスを表す。しかしながら、2つの信号が混合されると、マスキングが発生する可能性がある。たとえば、ノイズ信号がストリームx信号よりもはるかに大きい場合には、ノイズ信号はストリームx信号をマスクし、それにより、孤立して聴取されるその信号の知覚されるラウドネスに対して、その信号の知覚されるラウドネスを減少させる。この現象は、2つの入力をとる部分ラウドネス関数  $PL\{\cdot, \cdot\}$  でモデル化されうる。第1の入力は関心対象の信号の励起であり、第2の入力は競合する（ノイズ）信号の励起である。この関数は、競合する信号の存在下での関心対象の信号の知覚されるラウドネスを表す部分特定ラウドネス信号  $PL$  を返す。次いで、ノイズ信号の存在下でのストリームx信号の部分特定ラウドネス (partial specific loudness) は、周波数帯域b、時間t、およびラウドスピーカcにわたる励起信号から直接計算されうる：

【数59】

$$PL_x(b, t, c) = PL\{E_x(b, t, c), \bar{E}_x(b, t)\} \quad (26)$$

【0182】

ノイズの存在下でオーディオ・ストリームx信号の可聴性を維持するために、式8aおよび8bに示されるように、オーディオ・ストリームxがノイズより大きく可聴になるまでラウドネスを増加させるよう、オーディオ・ストリームxに適用する利得  $G_x(b, t, c)$  を計算することができる。あるいはまた、ノイズが別のオーディオ・ストリームsからのものであれば、2組の利得を計算することができる。そのような一例では、第1のもの  $G_x(b, t, c)$  は、

オーディオ・ストリーム $x$ に適用されて、そのラウドネスを増大させ、第2のもの $G_s(b, t)$ は、競合するオーディオ・ストリーム $s$ に適用されてそのラウドネスを低減させる。それにより、それらの利得の組み合わせが音声ストリーム $x$ の可聴性を保証する。これは式9aおよび9bに示される。両方の組の式において、 $\bar{PL}_x(b, t, c)$ は、補償利得の適用後のノイズの存在下でのソース信号の部分特定ラウドネスを表す。

【数60】

$$\bar{PL}_x(b, t, c) = PL\{G_x(b, t, c)^2 E_x(b, t, c), \bar{E}_x(b, t)\} \quad (27a)$$

ここで、

$$L_x(b, t, c) - \bar{PL}_x(b, t, c) < L_x(b, t, c) - PL_x(b, t, c) \quad (27b)$$

$$\bar{PL}_x(b, t, c) = PL\{G_x(b, t, c)^2 E_x(b, t, c), G_s(b, t)^2 \bar{E}_x(b, t)\} \quad (28a)$$

やはりここで、

$$L_x(b, t, c) - \bar{PL}_x(b, t, c) < L_x(b, t, c) - PL_x(b, t, c) \quad (28b)$$

【0183】

實際上、やはり可聴アーチファクトを避けるために、生の利得は、オーディオ・ストリームに適用される前に、平滑化関数 $S\{\cdot\}$ を使用して、周波数にわたってさらに平滑化される。 $\bar{G}_x(b, t, c)$ および $\bar{G}_s(b, t)$ は、ターゲット・オーディオ・ストリーム $x$ および競合するオーディオ・ストリーム $s$ についての最終的な補償利得を表す。

【数61】

$$\bar{G}_x(b, t, c) = S\{G_x(b, t, c)\} \quad (29a)$$

$$\bar{G}_s(b, t) = S\{G_s(b, t)\} \quad (29b)$$

【0184】

ある実施形態では、これらの利得は、オーディオ・ストリームのすべてのレンダリングされた出力チャンネルに直接適用されうる。別の実施形態では、それらは、その代わりに、レンダリングされる前にオーディオ・ストリームのオブジェクトに適用されてもよい。これはたとえば、参照により本明細書に組み込まれる米国特許出願公開第2019/0037333A1号に記載の方法を使用する。これらの方法は、オーディオ・オブジェクトの空間メタデータに基づいて、複数のあらかじめ定義されたチャンネル・カバレッジ・ゾーンのそれぞれに関連する各オーディオ・オブジェクトについてのパン係数を計算することを含む。オーディオ信号は、計算されたパン係数およびオーディオ・オブジェクトに基づいて、あらかじめ定義されたチャンネル・カバレッジ・ゾーンに関して、諸サブミックスに変換されてもよい。サブミックスのそれぞれは、あらかじめ定義されたチャンネル・カバレッジ・ゾーンの1つに関連して、前記複数のオーディオ・オブジェクトの成分の和を示してもよい。サブミックス利得は、サブミックスのそれぞれにオーディオ処理を適用することによって生成されてもよく、各オーディオ・オブジェクトに適用されるオブジェクト利得を制御してもよい。オブジェクト利得は、各オーディオ・オブジェクトについてのパン係数と、あらかじめ定義されたチャンネル・カバレッジ・ゾーンのそれぞれに関連するサブミックス利得との関数であってもよい。利得をオブジェクトに適用することには、特にストリームの他の処理と組み合わせた場合に、いくつかの利点がある。

【0185】

図8は、オーディオ・ストリーム・ラウドネス推定器を有するマルチストリーム・レンダリング・システムの実装を示す。この例によれば、図8のマルチストリーム・レンダリング・システムは、たとえば式12a~21bに記載されるようなラウドネス処理、および補

償利得適用をも各単一ストリーム・レンダラーにおいて実装するように構成される。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および2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および2のそれぞれに直交ミラーフィルタバンク(QMF)が適用される。代替的な例では、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のそれぞれに直交ミラーフィルタバンク(QMF)が適用されてもよい。この例によれば、レンダリング・モジュール1および2は、周波数領域において動作する。この実装において、ラウドネス推定モジュール805aは、たとえば式12a~17bを参照して上述したように、プログラム・ストリーム1についてのラウドネス推定値を計算する。同様に、この例では、ラウドネス推定モジュール805bは、プログラム・ストリーム2について

10

#### 【0186】

この実装において、マイクロフォン・システム120c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、それにより、ラウドネス推定モジュール805cは、周波数領域においてマイクロフォン信号を受領する。この実装において、ラウドネス推定モジュール805cは、たとえば式12b~17aを参照して上述したように、マイクロフォン信号についてのラウドネス推定値を計算する。この例では、ラウドネス処理モジュール810は、たとえば式18-21bに記載されるようなラウドネス処理、および補償利得適用を各単一ストリーム・レンダリング・モジュールのために実装するように構成される。この実装では、ラウドネス処理モジュール810は、プログラム・ストリーム1のオーディオ信号およびプログラム・ストリーム2のオーディオ信号を、一つまたは複数の干渉信号の存在下でそれらの知覚されるラウドネスを維持するために変更するように構成される。いくつかの事例では、制御システムは、マイクロフォン信号が環境ノイズに対応し、プログラム・ストリームがそれより上に上げられるべきであると判定することができる。しかしながら、いくつかの例では、制御システムは、マイクロフォン信号が、スマートオーディオ装置および／または1人以上の聴取者によって聞かれる必要がありうるウェイクワード、コマンド、子供の泣き声、または他のそのようなオーディオに対応すると判定することがありうる。いくつかのそのような実装では、ラウドネス処理モジュール810は、プログラム・ストリーム1の干渉オーディオ信号および／またはプログラム・ストリーム2のオーディオ信号の存在下でそれらの知覚されるラウドネスを維持するために、マイクロフォン信号を変更するように構成されてもよい。ここで、ラウドネス処理モジュール810は、レンダリング・モジュール1および2に適切な利得を提供するように構成される。ミキサー630c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635cが、該混合を時間領域に変換し、時間領域での混合されたスピーカフィール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630c、および逆フィルタバンク635cは、制御システム110eの構成要素である。

20

30

#### 【0187】

図9Aは、複数のレンダリングされたストリームのクロスフェージングのために構成されたマルチストリーム・レンダリング・システムの例を示す。いくつかのそのような実施形態では、レンダリング構成が動的に変更されるときに、スムーズな体験を提供するために、複数のレンダリングされたストリームのクロスフェージングが使用される。一例は、図4Aおよび図4Bを参照して上述したような、音楽のような空間的プログラム・ストリームの、聴取者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の同時再生の前述の使用事例である。この場合、図9Aに示されているように、代替的な空間的レンダリング構成を有する余分な単一ストリーム・レンダラーをインスタンス化し、それらの間で同時にクロスフェードすることが有用である。

40

#### 【0188】

この例では、プログラム・ストリーム1がレンダリング・モジュール1aおよび1bによって受領される前に、QMFがプログラム・ストリーム1に適用される。同様に、プログラム

50

・ストリーム2がレンダリング・モジュール2aおよび2bによって受領される前に、プログラム・ストリーム2にQMFが適用される。いくつかの事例では、レンダリング・モジュール1aの出力は、ウェイクワードの検出前のプログラム・ストリーム1の所望される再生に対応してもよく、一方、レンダリング・モジュール1bの出力は、ウェイクワードの検出後のプログラム・ストリーム1の所望される再生に対応してもよい。同様に、レンダリング・モジュール2aの出力は、ウェイクワードの検出前のプログラム・ストリーム2の所望される再生に対応してもよく、レンダリング・モジュール2bの出力は、ウェイクワードの検出後のプログラム・ストリーム2の所望される再生に対応してもよい。この実装では、レンダリング・モジュール1aおよび1bの出力はクロスフェード・モジュール910aに与えられ、レンダリング・モジュール2aおよび2bの出力はクロスフェード・モジュール910bに与えられる。クロスフェード時間は、たとえば、数百ミリ秒から数秒の範囲であってもよい。

10

#### 【0189】

ミキサー630dがクロスフェード・モジュール910aおよび910bの出力を混合した後、逆フィルタバンク635dは、該混合を時間領域に変換し、時間領域での混合されたスピーカーフィール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、クロスフェード・モジュール、ミキサー630dおよび逆フィルタバンク635dは、制御システム110fの構成要素である。

#### 【0190】

いくつかの実施形態では、単一ストリーム・レンダラー1a、1b、2a、および2bのそれぞれにおいて使用されるレンダリング構成をあらかじめ計算することが可能でありうる。これは、空間構成がしばしば先験的に知られており、システムの他の動的側面に依存しないので、スマート音声アシスタントのような使用事例にとって特に便利で効率的である。他の実施形態では、レンダリング構成をあらかじめ計算することは可能でない、または望ましくない場合があり、その場合、各単一ストリーム・レンダラーのための完全な構成は、システムが稼働している間に動的に計算されなければならない。

20

#### 【0191】

いくつかの実施側面の諸側面は、以下を含む：

1. 複数の任意に配置されたラウドスピーカーを通じて複数のオーディオ・プログラム・ストリームを同時に再生するオーディオ・レンダリング・システムであって、前記プログラム・ストリームの少なくとも1つは空間ミックスであり、前記空間ミックスのレンダリングは、一つまたは複数の追加的なプログラム・ストリームの同時再生に応答して動的に修正される、システム。
2. 請求項1に記載のシステムであって、前記複数のオーディオ・プログラム・ストリームのうちのいずれかの前記レンダリングが、残りの複数のオーディオ・プログラム・ストリームのうちの任意の一つまたは複数の組み合わせの関数として動的に修正されうる、システム。
3. 前記修飾は、以下のうちの一つまたは複数を含む、請求項1または2に記載のシステム・前記一つまたは複数の追加的なプログラム・ストリームのうちの少なくとも1つのレンダリングに関連するラウドスピーカーの相対的アクティブ化の関数として、前記複数のラウドスピーカーの相対的アクティブ化を修正する；  
・前記一つまたは複数の追加的なプログラム・ストリームのうちの少なくとも1つのレンダリングの空間特性の関数として、空間ミックスの意図された空間バランスをワープさせること；または  
・前記一つまたは複数の追加的なプログラム・ストリームのうちの少なくとも1つのラウドネスまたは可聴度の関数として、空間ミックスのラウドネスまたは可聴度を修正すること。
4. 一つまたは複数のマイクロフォン入力関数として前記レンダリングを動的に修正することをさらに含む、請求項1または2に記載のシステム。
5. 請求項4に記載のシステムであって、前記レンダリングを修正するために使用される前

30

40

50

記マイクロフォン入力から導出される情報は、以下のうちの一つまたは複数を含む

- ・システムのユーザーによる特定のフレーズの発声の検出；
- ・システムの1人以上のユーザーの位置の推定値；
- ・聴取空間の特定の位置における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のラウドネスの推定値；または
- ・聴取環境における他の環境音、たとえば背景ノイズのラウドネスの推定値。

#### 【0192】

複数のスピーカー（たとえば、一組のオーケストレーションされたスマートオーディオ装置のスピーカー）を通じた複数のオーディオ・ストリームの再生を管理するための本発明のシステムおよび方法の実施形態の他の例は、以下を含む：

1. 複数の任意に配置されたラウドスピーカー（たとえば、一組のオーケストレーションされたスマートオーディオ装置のスピーカー）を通じて複数のオーディオ・プログラム・ストリームを同時に再生するオーディオ・システム（たとえば、オーディオ・レンダリング・システム）であって、前記プログラム・ストリームの少なくとも1つは空間ミックスであり、前記空間ミックスのレンダリングは、一つまたは複数の追加的なプログラム・ストリームの同時再生にตอบสนองして（または関連して）動的に修正される、オーディオ・システム。
2. 前記空間ミックスに対する前記修正は、以下のうちの一つまたは複数を含む、請求項1に記載のシステム：
  - ・空間ミックスのレンダリングを、前記一つまたは複数の追加的なストリームのレンダリング位置から離れるようにワープさせる、または
  - ・前記一つまたは複数の追加的なストリームのラウドネスに応じて空間ミックスのラウドネスを修正する。
3. 請求項1に記載のシステムであって、前記空間ミックスのレンダリングを、一つまたは複数のマイクロフォン入力（すなわち、一つまたは複数のスマートオーディオ装置、たとえば一組のオーケストレーションされたスマートオーディオ装置の一つまたは複数のマイクロフォンによって捕捉される信号）の関数として、動的に修正するステップをさらに含むシステム。
4. 前記一つまたは複数のマイクロフォン入力のうちの少なくとも1つは、人間の声を含む（を示す）請求項3に記載のシステム。任意的に、レンダリングは、声のソース（人間）の決定された位置にตอบสนองして動的に修正される。
5. 前記一つまたは複数のマイクロフォン入力のうちの少なくとも1つは、環境ノイズを含む、請求項3に記載のシステム。
6. 前記空間的ストリームまたは前記一つまたは複数の追加的なストリームのラウドネスの推定値は、前記一つまたは複数のマイクロフォン入力のうちの少なくとも1つから導出される、請求項3に記載のシステム。

#### 【0193】

（いくつかの実施形態に従った）動的コスト柔軟レンダリングを実装する際の実際的な考慮事項の1つは、計算量である。場合によっては、オブジェクト位置（メタデータによって示されてもよい、レンダリングされる各オーディオ・オブジェクトについての位置）が1秒間に何度も変化する可能性があることを考慮すると、各オーディオ・オブジェクトについて各周波数帯域についての固有のコスト関数をリアルタイムで解くことは実現可能ではないことがありうる。メモリを代償として計算量を減らす代替的なアプローチは、すべての可能なオブジェクト位置の三次元空間をサンプリングするルックアップテーブルを使用することである。サンプリングは、すべての次元において同じである必要はない。図9Bは、ある例示的实施形態におけるスピーカー・アクティブ化を示す点のグラフである。この例では、xおよびy次元は15点でサンプリングされ、z次元は5点でサンプリングされる。他の実装は、より多くのサンプルまたはより少ないサンプルを含んでもよい。この例によれば、各点は、CMAPまたはFV解決策についてのMスピーカー・アクティブ化を表す。

10

20

30

40

50

## 【0194】

実行時に、各スピーカーについての実際のアクティブ化を決定するために、いくつかの例では、直近の8点のスピーカー・アクティブ化の間の三線形補間 (tri-linear interpolation) が使用されてもよい。図10は、一例によるスピーカー・アクティブ化を示す点の間の三線形補間のグラフである。この例では、逐次的な線形補間のプロセスは、上面内の各対の点を補間して第1および第2の補間点1005aおよび1005bを決定し、下面内の各対の点を補間して第3および第4の補間点1010aおよび1010bを決定し、第1および第2の補間点1005aおよび1005bを補間して上面内の第5の補間点1015を決定し、第3および第4の補間点1010aおよび1010bを補間して下面内の第6の補間点1020を決定し、第5および第6の補間点1015および1020を補間して上面と下面の間の第7の補間点1025を決定することを含む。三線形補間は有効な補間方法であるが、当業者は、三線形補間が、本開示の諸側面を実装する際に使用されうる1つの可能な補間方法に過ぎず、他の例は、他の補間方法を含みうることを理解するであろう。

10

## 【0195】

たとえば音声アシスタントのための音響空間を作るために反発力を使用されている上述の第1の例では、もう一つの重要な概念は、反発力のないレンダリング・シーンから反発力のあるシーンへの移行である。なめらかな遷移を作り出し、音場が動的に歪められる印象を与えるために、反発力のない以前の一組のスピーカー・アクティブ化と反発力のある新しい一組のスピーカー・アクティブ化の両方が計算され、時間期間にわたって補間される。

20

## 【0196】

ある実施形態に従って実装されるオーディオ・レンダリングの例は、オーディオ・レンダリング方法であって：

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有する一つまたは複数のオーディオ信号の集合を、2つ以上のラウドスピーカーの集合を通じてレンダリングすることを含む、ラウドスピーカーの集合の相対的アクティブ化は、それらのラウド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る空間位置のモデル、前記オーディオ・オブジェクトの所望される知覚される空間位置のラウドスピーカーの位置への近接性、およびオーディオ信号の前記集合の少なくとも一つまたは複数の属性、ラウドスピーカーの前記集合の一つまたは複数の属性、または一つまたは複数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の関数である、方法である。

30

## 【0197】

図11を参照して、実施形態のさらなる例を説明する。本明細書で提供される他の図と同様に、図11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11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1100は、左上の居間1110、中央下のキッチン1115、および右下のベッドルーム1122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1105a～1105h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1105a～1105h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。この例では、環境1100は、環境全体に分散されたカメラ1111a～1111eを含む。いくつかの実装では、環境1100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、任意的なセンサー・システム130の一つまたは複数のカメラは、テレビ1130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1105b、1105d、1105eもしくは1105h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1111a～1111eは、本開示において提示される環境1100のすべての図に示され

40

50

ているわけではないが、それでも、環境1100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。

【0198】

図12A、図12B、図12C、および図12Dは、図11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12A～図12D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1220aのほうを指す矢印1205は、（人1220a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1210aは左サラウンド・フィールドを表し、矢印1210bは右サラウンド・フィールドを表す。

【0199】

図12Aでは、居間のカウチ1225に座っている人1220a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。いくつかの実装によれば、制御システム（図1Aの制御システム110のような）は、図1Aのインターフェース・システム105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1A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。

【0200】

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1105a～1105e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1220aの声の音に応じて、居間のカウチ1225に座っている人1220aの位置を三角測量してもよい。人1220aの配向は、人1220aの位置に従って想定されてもよい：人1220aが図12Aに示される位置にいる場合、人1220aは、テレビジョン1130に面していると想定されてもよい。

【0201】

代替的または追加的に、人1220aの位置および配向は、カメラシステム（図1A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。

【0202】

いくつかの例では、人1220a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1220aが人1220a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）を制御するように構成されてもよい。

【0203】

図13A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。下記にいくつかの例が記載される。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13A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 1300に与えている。L字型カウチ1225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。

10

20

30

40

50

## 【0204】

図12Bでは、居間の読書椅子1215に座っている人1220a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12Cでは、キッチンカウンター1230の隣に立っている人1220a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12Dでは、朝食テーブル1240に座っている人1220a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1205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1100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

## 【0205】

図12A～12Dのいずれかの人1220a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12E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1220aおよびキッチンに立っている人1220bのための基準空間モード・レンダリングを示す。この例では、レンダリングは、人1220aにとって最適でありうるが、人1220b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。

20

## 【0206】

この場合、および複数の人々が空間にいて、予測不可能な仕方で動き回る可能性がある他の場合（たとえば、パーティー）では、そのような分散された聴衆にとってより適切なレンダリング・モードが必要とされる。図13Bは、ある例示的实施形態による分散式の空間的レンダリング・モードを示す。分散式空間モードのこの例では、前方サウンドステージは、居間や、カウチにいる聴取者の前方の位置のみからではなく、聴取空間全体にわたって一様にレンダリングされる。前方サウンドステージのこの分散は、クラウド1335を周回する複数の矢印1305dによって表され、すべての矢印1305dは同じ長さ、またはほぼ同じ長さを有する。矢印1305dの意図された意味は、図示された複数の聴取者（人1220a～1220f）が、その位置にかかわらず、ミックスのこの部分を同じくらいよく聞くことができるということである。しかしながら、この一様分布がミックスのすべての成分に適用されたとすると、ミックスのすべての空間的側面が失われることになり、人1220a～1220fは本質的にモノラル・オーディオを聞くことになる。ある程度の空間性を維持するために、それぞれ矢印1210aおよび1210bで表されるミックスの左および右のサラウンド成分は、依然として空間的にレンダリングされる。（多くの場合、左側と右側のサラウンド、左と右の後方サラウンド、オーバーヘッド、およびこの空間内の空間位置をもつ動的なオーディオ・オブジェクトが存在しうる。矢印1210aおよび1210bは、これらすべての可能性の左と右の部分を表すように意図されている。）知覚される空間性を最大化するために、これらの成分が空間化されるエリアは、これまで前方サウンドステージのみによって占有されていた空間を含む、聴覚空間全体をより完全にカバーするように拡大される。サラウンド成分がレンダリングされるこの拡大されたエリアは、図13Bに示された比較的長い矢印1210aおよび1210bを、図12Aに示された比較的短い矢印1210aおよび1210bと比較することによって理解されうる。さらに、図12Aに示された矢印1210aおよび1210bは、基準空間モードにおけるサラウンド成分を表し、ほぼ人1220aの側面から聴取環境の背面まで延び、聴取環境の前方ステージ・エリアには延びていない。

30

40

## 【0207】

この例では、前方サウンドステージの一様な分布と、サラウンド成分の拡大された空間化とを実装する際に、基準空間モードについてのレンダリングと比較して、これらの成分の知覚されるラウドネスが大部分維持されるように、注意が払われる。目標は、ミックス中の各成分の相対的なレベルを維持しながら、複数の人のために最適化するよう、これらの成分の空間的印象をシフトさせることである。たとえば、前方サウンドステージが、そ

50

の様な分布の結果として、サラウンド成分に対して2倍の大きさになったとしたら、望ましくないであろう。

#### 【0208】

例示的实施形態のさまざまな基準レンダリング・モードと分散レンダリング・モードとの間で切り換えるために、いくつかの例においては、ユーザーは、オーケストレーションされたスピーカーのシステムに付随する音声アシスタントと対話することができる。たとえば、基準空間モードでオーディオを再生するために、ユーザーは、音声アシスタントのためのウェイクワード（たとえば、「Listen Dolby（聞いて、ドルビー）」）に続いて、コマンド、「私のために[コンテンツ名を挿入]を再生してください」または「パーソナルモードで[コンテンツ名を挿入]を再生してください」を発声してもよい。すると、システムに付随するさまざまなマイクロフォンからの記録に基づいて、システムは、ユーザーの位置および配向、またはいくつかの所定のゾーンのうちユーザーに最も近いものを自動的に決定し、この決定された位置に対応する基準モードでオーディオの再生を始めてもよい。分散空間モードでオーディオを再生するために、ユーザーは異なるコマンド、たとえば「分散モードで[コンテンツ名を挿入]を再生してください」を発声してもよい。

#### 【0209】

代替的または追加的に、システムは、他の入力に基づいて、基準モードと分散モードとの間で自動的に切り換えるように構成されてもよい。たとえば、システムは、空間内の聴取者の数とその位置を自動的に決定する手段を有してもよい。これは、たとえば、付随するマイクロフォンから空間内の音声活動をモニターすることによって、および／または一つまたは複数のカメラ等の他の関連するセンサーの使用を通じて達成されうる。この場合、システムは、図12Eに示されるような基準空間モードと、図13Bに示されるような完全に分散した空間モードとの間でレンダリングを連続的に変化させる機構を備えるように構成されてもよい。この連続体上でレンダリングが設定される点は、たとえば空間において報告される人数の関数として計算されてもよい。

#### 【0210】

図12A、図14A、および図14Bは、この挙動を示す。図12Aでは、システムは、テレビに面してカウチにいる単一の聴取者（人1220a）だけを検出し、よって、レンダリング・モードは、この聴取者位置および配向のための基準空間モードに設定される。図14Aは、一例による部分的に分散された空間的レンダリング・モードを示す。図14Aでは、追加的な2人（人1220eおよび1220f）が、人1220aの後ろに検出され、レンダリング・モードは、基準空間モードと完全分散空間モードとの間の点に設定される。ここで描かれているところでは、前方サウンドステージ（矢印1305a、1305b、および1305c）のいくらかが追加的な聴取者（人1220eおよび1220f）のほうに引き戻されているが、それでも、基準空間モードの前方サウンドステージの位置のほうに、より多くの強調がある。この強調は、図14Aにおいて、矢印1205と、矢印1305bおよび1305cの長さと比較して比較的長い矢印1305aの長さによって示される。また、サラウンド・フィールドは、矢印1210aおよび1210bの長さおよび位置によって示されるように、基準空間モードの前方サウンドステージの位置に向けて部分的に拡張されるだけである。

#### 【0211】

図14Bは、一例による完全分散空間レンダリング・モードを示す。いくつかの例において、システムは、空間全体にまたがる多数の聴取者（人1220a、1220e、1220f、1220g、1220hおよび1220i）を検出していることがあり、システムは、レンダリング・モードを自動的に、完全分散空間モードに設定していてもよい。他の例では、レンダリング・モードは、ユーザー入力に従って設定されていてもよい。完全分散空間モードは、図14Bにおいて、矢印1305dの一樣または実質的に一樣な長さ、ならびに矢印1210aおよび1210bの長さおよび位置によって、示される。

#### 【0212】

前述の諸例では、分散レンダリング・モードにおいて、より一樣な分散を用いてレンダリングされる空間ミックスの部分が前方サウンドステージとして指定される。多くの空間

ミックスの文脈では、これは理にかなっている。なぜなら、伝統的なミックスの実施は、典型的には、映画についてのダイアログや音楽についてのリードボーカル、ドラム、ベースなど、ミックスの最も重要な部分を、前方サウンドステージに配置するからである。これは、ほとんどの5.1および7.1サラウンドサウンド・ミックス、ならびにドルビー・プロロジックまたはドルビー・サラウンドなどのアルゴリズムを使用して5.1または7.1にアップミックスされたステレオ・コンテンツに対して成り立つ。ここで、前方サウンドステージは、左、右、および中央のチャンネルによって与えられる。これは、オーディオ・データが $y < 0.5$ の $(x, y)$ 空間位置を示す空間メタデータに従って前方サウンドステージとして指定されうる、ドルビー・アトモスのような、多くのオブジェクトベースのオーディオ・ミックスにも当てはまる。しかしながら、オブジェクトベースのオーディオでは、ミキシング・エンジニアは、オーディオを3D空間のどこにでも配置する自由がある。特に、オブジェクトベースの音楽では、ミキシング・エンジニアは伝統的なミキシング規範から離れ、リードボーカルのような、ミックスの重要な部分と考えられるものを、頭上のような非伝統的な位置に配置するようになりつつある。そのような場合、ミックスのどの成分が、分散レンダリング・モードのためのより分散した空間的態様でレンダリングするのに適しているかを決定するための単純な規則を構築することは困難になる。オブジェクトベースのオーディオは、3D空間のどこで信号がレンダリングされるべきかを記述する、その構成要素の各オーディオ信号に関連付けられたメタデータをすでに含んでいる。上述の問題に対処するために、いくつかの実装では、コンテンツ作成者が、分散レンダリング・モードにおけるより分散した空間レンダリングに適しているとして、特定の信号にフラグ付けすることを許容する追加的なメタデータが追加されてもよい。レンダリング中に、システムはこのメタデータを使用して、より分散されたレンダリングが適用されるミックスの成分を選択してもよい。これは、コンテンツ作成者に、特定のコンテンツについて分散レンダリング・モードがどのように聞こえるかに対する制御を与える。

#### 【0213】

いくつかの代替的な実装では、制御システムは、より空間的に分散された仕方でレンダリングされるべきオーディオ・データの一つまたは複数の要素を識別するために、コンテンツ・タイプ分類器を実装するように構成されてもよい。いくつかの例では、コンテンツ・タイプ分類器は、オーディオ・データがより空間的に分散された仕方でレンダリングされるべきかどうかを決定するために、コンテンツ・タイプ・メタデータ（たとえば、オーディオ・データがダイアログ、ボーカル、パーカッション、ベースなどであることを示すメタデータ）を参照してもよい。いくつかのそのような実装によれば、より空間的に分散された仕方でレンダリングされるコンテンツ・タイプ・メタデータは、たとえば、表示装置上に表示されるGUIを介したユーザー入力に従って、ユーザーによって選択可能であってもよい。

#### 【0214】

空間的オーディオ・ミックスの一つまたは複数の要素を、基準空間モードよりも空間的に分散された仕方でレンダリングするために使用される厳密な機構は、異なる実施形態間で変化してもよく、本開示は、すべてのそのような機構をカバーすることが意図されている。1つの例示的な機構は、聴取空間を通じて、より一様に分散された、複数の関連するレンダリング位置をもつ、そのような各要素の複数のコピーを生成することに関わる。いくつかの実装では、分散空間モードについてのレンダリング位置および／またはレンダリング位置の数は、ユーザー選択可能であってもよいが、他の実装では、分散空間モードについてのレンダリング位置および／またはレンダリング位置の数は、事前設定されてもよい。いくつかのそのような実装では、ユーザーは、分散空間モードのためのいくつかのレンダリング位置を選択してもよく、それらのレンダリング位置は、たとえば、聴取環境を通じて均等に離間して、事前設定されてもよい。次いで、システムは、これらのコピーのすべてを、もとの意図された位置にあるもとの単一要素ではなく、一組の分散された位置においてレンダリングする。いくつかの実装によれば、すべてのコピーの組み合わせられたレンダリングに関連する知覚されるレベルが、基準レンダリング・モードにおけるも

との単一要素のレベルと同じか、または実質的に同じ（たとえば、2dB、3dB、4dB、5dB、6dBなどの閾値数のデシベルの範囲内）になるように、コピーのレベルが修正されてもよい。

#### 【0215】

CMAPまたはFV柔軟レンダリング・システムのいずれかのコンテキストでは、または両システムのハイブリッドでは、よりエレガントな機構が実装されうる。これらのシステムでは、空間ミックスの各要素は、空間内の特定の位置でレンダリングされ、各要素には、想定される固定位置、たとえば5.1または7.1サラウンドサウンド・ミックス内のあるチャンネルの正準位置、またはドルビー・アトモスのようなオブジェクトベースのオーディオの場合のような時間変化する位置が関連付けられていてもよい。

#### 【0216】

図15は、2D平面上のCMAPおよびFVレンダリング・システムについての例示的なレンダリング位置を示す。それぞれの小さな、番号を付けられた円は、例示的なレンダリング位置を表し、レンダリング・システムは、円1500の上または円1500内のどこにでも、空間ミックスの要素をレンダリングすることができる。L、R、C、Lss、Rss、Lrs、およびRrsとラベル付けされた円1500上の位置は、この例における7.1サラウンドミックスの7つのフルレンジ・チャンネルの固定した正準レンダリング位置を表す：左（L）、右（R）、中央（C）、左側方サラウンド（Lss）、右側方サラウンド（Rss）、左後方サラウンド（Lrs）、および右後方サラウンド（Rrs）。この文脈では、L、R、およびC近傍のレンダリング位置は、前方サウンドステージとみなされる。基準レンダリング・モード（reference rendering mode）（本明細書では「基準空間モード（reference spatial mode）」とも呼ばれる）については、聴取者は、Cレンダリング位置に対向する大きな円の中心に位置すると想定される。さまざまな聴取位置および配向について参照レンダリングを示す図12A～12Dのいずれについても、図15の中心を聴取者の上に重ね合わせることを概念化することができる。ここで、C位置が前方サウンドステージ（矢印1205）の位置と整列し、図15の円1500がクラウド1235を囲むように、図15は、追加的に回転され、スケールリングされる。結果として得られる整列は、図12A～12Dの任意のスピーカーが図15の任意のレンダリング位置のいずれかに比較的近接していることを記述する。いくつかの実装では、CMAPおよびFVレンダリング・システムの両方について、特定の位置で空間ミックスの要素をレンダリングする際のスピーカーの相対的アクティブ化を大きく支配しているのは、この近接性である。

#### 【0217】

スタジオ内で空間的オーディオが混合される場合、スピーカーは一般に、聴取位置のまわりに均一な距離で配置される。ほとんどの場合、結果として得られる円または半球の範囲内にスピーカーはない。オーディオが「部屋の中」（たとえば、図15の中心）に置かれると、レンダリングは、「どこでもないサウンド」を達成するために、周上のすべてのスピーカーの発射に向かう傾向がある。CMAPおよびFVレンダリング・システムでは、スピーカー・アクティブ化を支配するコスト関数の近接ペナルティ項を変更することによって、同様の効果を達成することができる。特に、図15の円1500の周上のレンダリング位置について、近接ペナルティ項は、所望されるレンダリング位置から離れたスピーカーの使用に対して、フルにペナルティを課す。よって、意図されたレンダリング位置近傍のスピーカーのみが実質的にアクティブ化される。所望されるレンダリング位置が円の中心（半径ゼロ）に向かって移動すると、近接ペナルティ項はゼロに減少し、その結果、中心では、どのスピーカーも優先されない。半径ゼロでのレンダリング位置についての対応する結果は、聴取空間にわたるオーディオの完全に一樣な知覚された分布であり、これは、まさに、最も分散した空間レンダリング・モードにおける、ミックスのある種の要素についての所望される結果でもある。

#### 【0218】

半径ゼロでのCMAPおよびFVシステムのこの挙動を考慮すると、空間ミックスの任意の要素の、より空間的に分散されたレンダリングが、その意図した空間位置をゼロ半径点に

10

20

30

40

50

向かってワープさせることによって達成されうる。このワープは、もとの意図された位置とゼロ半径との間で連続的にされてもよく、それによって、基準空間モードとさまざまな分散空間モードとの間の自然な、連続的な制御を提供する。図16A、16B、16C、および16Dは、さまざまな分散空間レンダリング・モードを達成するために、図15のレンダリング・ポイントのすべてに適用されるワープの例を示す。図16Dは、完全に分散されたレンダリング・モードを達成するために、図15のレンダリング・ポイントのすべてに適用されるそのようなワープの例を示す。L、R、C点（前方サウンドステージ）が半径ゼロに折りたたまれ、それにより完全に一樣な仕方でのレンダリングを保証することがわかる。さらに、LssおよびRssレンダリング・ポイントは、空間化されたサラウンド・フィールド（Lss、Rss、Lbs、およびRbs）が聴取エリア全体を取り囲むように、円の周に沿ってもとの前方サウンドステージに向かって引っ張られている。このワープは、レンダリング空間全体に適用され、図15のレンダリング・ポイントのすべてが、7.1正準位置のワープに相応する図16Dの新しい位置にワープされていることがわかる。図16Dにおいて参照される空間モードは、本明細書において「最も分散した空間モード（most distributed spatial mode）」または「完全に分散した空間モード（fully distributed spatial mode）」と称されるものの一例である。

#### 【0219】

図16A、図16Bおよび図16Cは、図15に表された分散空間モードと図16Dに表された分散空間モードとの間の中間的な分散空間モードのさまざまな例を示す。図16Bは、図15に表される分散空間モードと図16Dに表される分散空間モードとの間の中間点を表す。図16Aは、図15に表される分散空間モードと図16Bに表される分散空間モードとの間の中間点を表す。図16Cは、図16Bに表される分散空間モードと図16Dに表される分散空間モードとの間の中間点を表す。

#### 【0220】

図17は、ユーザーがレンダリング・モードを選択しうるGUIの例を示す。いくつかの実装によれば、制御システムは、GUI 1700または類似のGUIをディスプレイ上に表示するよう、表示装置（たとえば、セルラー電話）を制御してもよい。表示装置は、センサー・システム（たとえば、タッチセンサーシステム、またはディスプレイに近接する（たとえば、ディスプレイの上またはディスプレイの下にある）ジェスチャーセンサーシステム）を含んでいてもよい。制御システムは、センサー・システムからのセンサー信号の形で、GUI 1700を介してユーザー入力を受領するように構成されてもよい。センサー信号は、GUI 1700の要素に対応するユーザータッチまたはジェスチャーに対応してもよい。

#### 【0221】

この例によれば、GUIは、ユーザーがレンダリング・モードを選択するためにそれを用いて対話することができる場所の仮想スライダー1701を含む。矢印1703によって示されるように、ユーザーは、スライダーをトラック1707に沿っていずれかの方向に移動させることができる。この例では、線1705は、本明細書に開示されている基準空間モードの1つなど、基準空間モードに対応する仮想スライダー1701の位置を示している。他の実装は、仮想ノブまたはダイヤルのように、ユーザーがそれを用いて対話することができる場所のGUI上の他のフィーチャーを提供してもよい。いくつかの実装によれば、基準空間モードを選択した後、制御システムは、図13Aに示されているようなGUI、またはユーザーが基準空間モードのための聴取者位置および配向を選択することを許容する別のそのようなGUIを提示してもよい。

#### 【0222】

この例では、線1725は、図13Bに示される分散空間モードのような、最も分散した空間モードに対応する仮想スライダー1701の位置を示す。この実装によれば、線1710、1715、および1720は、中間的な空間モードに対応する仮想スライダー1701の位置を示す。この例では、線1710の位置は、図16Aのような中間的な空間モードに対応する。ここで、線1715の位置は、図16Bのような中間的な空間モードに対応する。この実施において、線1720の位置は、図16Cのような中間的な空間モードに対応する。この例によれば

、ユーザーは、選択されたレンダリング・モードを実装するように制御システムに命令するために、「適用」ボタンと対話する（たとえば、タッチする）ことができる。

#### 【0223】

しかしながら、他の実装は、ユーザーが前述の分散空間モードのうちの1つを選択するための他の仕方を提供してもよい。いくつかの例によれば、ユーザーは、音声コマンド、たとえば、「半分散モードで[コンテンツ名を挿入]を再生してください」を発声してもよい。「半分散モード」は、図17のGUI 1700における線1715の位置によって示される分散モードに対応してもよい。いくつかのそのような例によれば、ユーザーは、音声コマンド、たとえば、「1/4分散モードで[コンテンツ名を挿入]を再生してください」を発声してもよい。「1/4分散モード」は、線1710の位置によって示される分散モードに対応してもよい。

#### 【0224】

図18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行されうる方法の一例を概説するフロー図である。方法1800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1800の一つまたは複数のブロックが同時並行して実行されてもよい。さらに、方法1800のいくつかの実装は、図示および／または記載されるよりも多数または少数のブロックを含んでいてもよい。方法1800のブロックは、一つまたは複数の装置（図1Aに示されて上述された制御システム110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つであってもよく、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。

#### 【0225】

この実装では、ブロック1805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。この例では、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す。ここで、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。

#### 【0226】

この例では、ブロック1810は、制御システムによってレンダリング・モードを決定することを含む。レンダリング・モードを決定することは、場合によっては、インターフェース・システムを介してレンダリング・モード指示を受領することを含んでいてもよい。レンダリング・モード指示を受領することは、たとえば、音声コマンドに対応するマイクロフォン信号を受領することに関わってもよい。いくつかの例では、レンダリング・モード指示を受領することは、グラフィカル・ユーザー・インターフェースを介して、ユーザー入力に対応するセンサー信号を受領することを含んでいてもよい。センサー信号は、たとえば、タッチセンサー信号および／またはジェスチャーセンサー信号であってもよい。

#### 【0227】

いくつかの実装では、レンダリング・モード指示を受領することは、聴取エリア内の人数の指示を受領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、制御システムは、少なくとも部分的には、聴取エリア内の人数に基づいてレンダリング・モードを決定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、聴取エリア内の人数の指示は、マイクロフォン・システムからのマイクロフォン・データおよび／またはカメラシステムからの画像データに基づいてもよい。

#### 【0228】

図18に示される例によれば、ブロック1815は、制御システムによって、ブロック1810で決定されたレンダリング・モードに従って、環境の一組のラウドスピーカーを介した再生のために、オーディオ・データをレンダリングし、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。この例では、オーディオ・データのレンダリングは、環境内の一組のラウドスピーカーの相対的なアクティブ化を決定することを含む。ここで、レンダリング・モードは、基準空間モードと一つまたは複数の分散空間モードとの間で可変であ

10

20

30

40

50

る。この実装では、基準空間モードは、想定される聴取位置および配向を有する。この例によれば、前記一つまたは複数の分散空間モードにおいて、オーディオ・データの一つまたは複数の要素は、それぞれ基準空間モードにおけるよりも空間的に分散された仕方でレンダリングされる。この例では、前記一つまたは複数の分散空間モードにおいて、オーディオ・データの残りの要素の空間的位置は、それらが基準空間モードよりも完全に、環境のレンダリング空間にまたがるように、ワープされる。

#### 【0229】

いくつかの実装では、基準空間モードにおけるよりも、より空間的に分散された仕方でオーディオ・データの前記一つまたは複数の要素をレンダリングすることは、前記一つまたは複数の要素のコピーを生成することを含みうる。そのような実装のいくつかは、環境にわたる分散された一組の位置で、すべてのコピーを同時にレンダリングすることを含んでいてもよい。

10

#### 【0230】

いくつかの実装によれば、レンダリングは、CMAP、FV、またはそれらの組み合わせに基づいてもよい。基準空間モードにおけるよりも、より空間的に分散された仕方で、オーディオ・データの前記一つまたは複数の要素をレンダリングすることは、ゼロ半径に向けて前記一つまたは複数の要素のそれぞれのレンダリング位置をワープさせることを含んでいてもよい。

#### 【0231】

この例では、ブロック1820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、レンダリングされたオーディオ信号を、環境の前記一組のラウドスピーカーのうちの少なくともいくつかに提供することを含む。

20

#### 【0232】

いくつかの実装によれば、レンダリング・モードは、基準空間モードから最も分散した空間モードまでの範囲のレンダリング・モードの連続体から選択可能であってもよい。いくつかのそのような実装では、制御システムは、インターフェース・システムを介して受領された基準空間モード・データに従って、基準空間モードの想定される聴取位置および／または配向を決定するようにさら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、基準空間モード・データは、マイクロフォン・システムからのマイクロフォン・データおよび／またはカメラシステムからの画像データを含んでいてもよい。いくつかのそのような例では、基準空間モード・データは、音声コマンド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、基準空間モード・データは、聴取環境における人の一つまたは複数の発声の位置に対応するマイクロフォン・データを含んでいてもよい。いくつかのそのような例では、基準空間モード・データは、聴取環境における人の位置および／または配向を示す画像データを含んでいてもよい。

30

#### 【0233】

しかしながら、いくつかの事例では、装置またはシステムは、表示装置および該表示装置に近接するセンサー・システムを含んでいてもよい。制御システムは、グラフィカル・ユーザー・インターフェースを提示するように表示装置を制御するように構成されてもよい。基準空間モード・データの受領は、グラフィカル・ユーザー・インターフェースを介してユーザー入力に対応するセンサー信号を受領することを含んでいてもよい。

40

#### 【0234】

いくつかの実装によれば、より空間的に分散された仕方でそれぞれレンダリングされたオーディオ・データの前記一つまたは複数の要素は、前方サウンドステージ・データ、音楽ボーカル、ダイアログ、ベース、パーカッション、および／または他のソロまたはリード楽器に対応しうる。場合によっては、前方サウンドステージ・データは、ドルビー5.1、ドルビー7.1、またはドルビー9.1フォーマットで受領された、またはこれにアップミックスされたオーディオ・データの左、右、または中央信号を含んでいてもよい。いくつかの例では、前方サウンドステージ・データはドルビー・アトモス・フォーマットで受領された、(x,y)空間位置を示す空間メタデータを有するオーディオ・データを含んでいても

50

よく、 $y < 0.5$ のである。

#### 【0235】

場合によっては、オーディオ・データは、該オーディオ・データのどの要素が、より空間的に分散された仕方ではレンダリングされるべきかを示す空間分散メタデータを含んでいてもよい。いくつかのそのような例では、制御システムは、空間分散メタデータに従って、より空間的に分散された仕方ではレンダリングされるべきオーディオ・データの前記一つまたは複数の要素を識別するように構成されてもよい。

#### 【0236】

代替的または追加的に、制御システムは、より空間的に分散された仕方ではレンダリングされるべきオーディオ・データの前記一つまたは複数の要素を識別するために、コンテンツ・タイプ分類器を実装するように構成されてもよい。いくつかの例では、コンテンツ・タイプ分類器は、オーディオ・データが、より空間的に分散された仕方ではレンダリングされるべきかどうかを決定するために、コンテンツ・タイプ・メタデータ（たとえば、オーディオ・データがダイアログ、ボーカル、パーカッション、ベースなどであることを示すメタデータ）を参照してもよい。いくつかのそのような実装によれば、より空間的に分散された仕方ではレンダリングされるべきコンテンツ・タイプ・メタデータは、たとえば、表示装置上に表示されるGUIを介したユーザー入力に従って、ユーザーによって選択可能でありうる。代替的または追加的に、コンテンツ・タイプ分類器は、レンダリング・システムと組み合わせて、オーディオ信号に対して直接動作してもよい。たとえば、分類器は、オーディオ信号を解析し、それらが、より空間的に分散された仕方ではレンダリングするのに適切と考えられうるいずれかのコンテンツ・タイプ（ボーカル、リードギター、ドラムなど）に属するかどうかを判断するように、多様なコンテンツ・タイプ上でトレーニングされたニューラルネットワークを使用して実装されうる。そのようないくつかの分類は、連続的かつ動的な仕方では実行されてもよく、結果として得られる分類結果は、連続的かつ動的な仕方では、より空間的に分散された仕方ではレンダリングされる一組の信号を調整してもよい。いくつかのそのような実装は、当該技術分野で公知の方法に従ってそのような動的な分類システムを実装するために、ニューラルネットワークのような技術の使用を含んでいてもよい。

#### 【0237】

いくつかの例では、前記一つまたは複数の分散空間モードのうちの少なくとも1つは、少なくとも1つの要素の空間位置に対して時間変化する修正を適用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、時間変化する修正は、周期的修正であってもよい。たとえば、周期的修正は、聴取環境の周のまわりの一つまたは複数のレンダリング位置を公転させることを含む。いくつかのそのような実装によれば、周期的修正は、環境において再生される音楽のテンポ、環境において再生される音楽のビート、または環境において再生されるオーディオ・データの一つまたは複数の他の特徴を含んでいてもよい。たとえば、いくつかのそのような周期的修正は、2つ、3つ、4つ、またはそれ以上のレンダリング位置の間で交替することを含む。該交替は、環境において再生される音楽のビートに対応してもよい。いくつかの実装では、周期的修正は、ユーザー入力に従って、たとえば、一つまたは複数の音声コマンドに従って、GUIなどを介して受領されたユーザー入力に従って、などの仕方では選択可能でありうる。

#### 【0238】

図19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1900は、テレビジョン1901、ソファ1903および5つのオーディオ装置1905を含む部屋である。この例によれば、オーディオ装置1905は、環境1900の位置1〜5にある。この実装では、各オーディオ装置1905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1920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1925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1920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1905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。

## 【0239】

本明細書に開示される他の例と同様に、図19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1905、異なる位置にあるオーディオ装置1905などを有していてもよい。

## 【0240】

この例では、三角形1910aは、位置1、2および3に頂点を有する。ここで、三角形1910a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 $\theta_2$ であり、辺12と13aとの間の角度は $\theta_1$ であり、辺23aと13aとの間の角度は $\theta_3$ 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。

10

## 【0241】

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。

20

## 【0242】

図20は、図19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1910bは、位置1、3および4に頂点を有する。ここで、三角形1910b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 $\theta_4$ であり、辺13bと34aとの間の角度は $\theta_5$ であり、側面34aと14との間の角度は $\theta_6$ である。

## 【0243】

図11および図12を比較することによって、三角形1910aの辺13aの長さは、三角形1910bの辺13bの長さと等しいはず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1910a）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。

30

## 【0244】

図21Aは、図19および20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21Aは、三角形1910aおよび1910bの辺の長さおよび角度配向の推定値を示す。図21Aに示される例では、三角形1910bの辺13bの長さは、三角形1910aの辺13aと同じ長さに制約される。三角形1910b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1910b'は、三角形1910aに隣接して、図21Aに示されている。

40

## 【0245】

いくつかの実装によれば、三角形1910aおよび1910bに隣接する他の三角形の辺の長さは、環境1900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。

## 【0246】

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。i番目のオーディオ装置のデカルト座標は

50

## 【数 6 2】

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

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、 $i = \{1 \dots M\}$ である。

## 【0 2 4 7】

図21B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 $\theta_{ji}$ 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 $\theta_{ki}$ と表されてもよい。図21Bに示される例では、 $\theta_{ji}$ および $\theta_{ki}$ は軸2105aから測られ、軸2105aの配向は任意であり、軸2105a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2105aの配向に依存しないことが観察されうる。

10

## 【0 2 4 8】

図21Bに示される例では、 $\theta_{ij}$ および $\theta_{kj}$ は軸2105bから測られ、軸2105bの配向は任意であり、軸2105bはオーディオ装置jの配向に対応してもよい。三角形2110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、 $\theta_{jk}$ および $\theta_{ik}$ は軸2105cから測られる。三角形2110の内角cは、 $c = \theta_{jk} - \theta_{ik}$ と表されてもよい。

20

## 【0 2 4 9】

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 6 3】

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

## 【0 2 5 0】

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

30

## 【数 6 4】

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

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 6 5】

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

しかしながら、任意の回転は受け容れ可能でありうる。

40

## 【0 2 5 1】

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

## 【数 6 6】

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

のスーパーセット $\mathcal{L}$ 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

50

、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。

【0252】

図22は、図1A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2200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2200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2200のブロックは、図1Aに示される装置100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。

10

【0253】

この例では、ブロック2205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19に示されるオーディオ装置1905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。

【0254】

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。

20

【0255】

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。

30

【0256】

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。

【0257】

この例によれば、ブロック2210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。

40

【0258】

図23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1905の間の距離に対応する。

【0259】

この実装では、ブロック2215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

50

、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。

【0260】

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。

10

【0261】

この例によれば、ブロック2220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。

【0262】

いくつかのそのような例によれば、三角形は、たとえば図21A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整列することが期待される。

【数67】

20

$\varepsilon$

{ $\varepsilon$ と書くこともある}が、サイズ

【数68】

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

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2220は、 $\varepsilon$ 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。

30

【0263】

図24は、順整列プロセスの一部の例を与えている。図24において太字で示されている番号1から5は、図1、図2、図5に示されているオーディオ装置位置に対応している。図24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。

【0264】

この例では、図21Aのように、三角形1910bの辺13bの長さは、三角形1910a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1910b'は、図24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1910cの辺13cの長さも、三角形1910a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1910c'は、図24に示されており、同じ内角が維持されている。

40

【0265】

次に、この例では、三角形1910dの辺34bの長さは、三角形1910b'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1910dの辺23bの長さは、三角形1910a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1910d'は、図24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5に示される残りの三角形は、三角形1910b、1910c、および1910dと同じ仕方で処理されうる。

【0266】

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

50

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 6 9】

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

に格納されてもよく、ここで、Nは、三角形の総数を示す。

【0 2 6 7】

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。

【0 2 6 8】

図25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25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。

【0 2 6 9】

図22に戻ると、この例では、ブロック2225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はなくてもよい。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、

【数 7 0】

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

と表わされてもよい。

【0 2 7 0】

図26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19、図21、および図23に示されているオーディオ装置位置に対応している。図26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。

【0 2 7 1】

図26に示される例では、三角形1910e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1910e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1910fの辺45bの長さは、三角形1910e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1910f'は、内角が同じままであり、図26に示されている。この例では、三角形1910cの辺35bの長さは、三角形1910e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1910c"は、内角が同じままであり、図26に示されている。いくつかのそのような例によれば、図23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1910cおよび1910fと同様に処理される。

【0 2 7 2】

図27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25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27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。

【0273】

図22に戻ると、ブロック2230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。

10

【0274】

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71】

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

によって固定される。

【0275】

20

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いてもよい。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いてもよい。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72】

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

30

【0276】

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73】

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

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、

40

【数74】

$\bar{\mathbf{R}}\bar{\mathbf{X}}$  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。

【0277】

いくつかの例によれば、回転行列 $\mathbf{R} = \mathbf{V}\mathbf{U}^T$ 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75】

50

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

【0278】

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を含む。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76】

$$\vec{X}$$

10

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77】

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

が得られる。

【0279】

20

図28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17および図19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28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28の円で表される）とよく一致する。

【0280】

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14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。

30

【0281】

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。

40

【0282】

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が、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。

【0283】

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

50

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。

#### 【0284】

図29は、図1A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2900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2900のブロックは、図1Aに示される制御システム110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110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110は2つ以上の装置内に存在してもよい。

#### 【0285】

この例では、ブロック2905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27に示されるオーディオ装置1905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。

#### 【0286】

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を含まなくてもよい。

#### 【0287】

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。

#### 【0288】

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。

#### 【0289】

図29に示される例によれば、ブロック2910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2905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。

#### 【0290】

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

10

20

30

40

50

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いてもよい）。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。

#### 【0291】

いくつかの例では、ブロック2910は、図22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2910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。

#### 【0292】

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2900のいくつかの実装では、ブロック2910は、図22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。

#### 【0293】

この例では、ブロック2915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。

#### 【0294】

いくつかの例では、ブロック2915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いてもよい。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2915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2915は、たとえば、図30A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2915（または方法2900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。

#### 【0295】

この実装によれば、ブロック2920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。

【0296】

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。

【0297】

ブロック2920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。

【0298】

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・データを含んでもよい。

【0299】

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・データを含んでもよい。

【0300】

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。

【0301】

図29に示される例によれば、ブロック2925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2925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もよい。いくつかの実装では、ブロック2925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もよい。いくつかの例が後述される。

【0302】

図30Aは、図29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3007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3007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3007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3003に対応する。

【0303】

10

20

30

40

50

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3005に（たとえば、環境3000a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3027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3027の検出に対応する。

#### 【0304】

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3009a、3009b等の交点に従って決定されてもよい。

#### 【0305】

この例によれば、聴取者位置は、聴取者座標系3020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3010（および／または聴取者の鼻3025）とテレビジョン101のサウンドバー3030との間の直線3013aに対応する聴取者座標系3020のy'軸によって示される。図30Aに示される例では、線3013aはy'軸に平行である。したがって、角度 $\theta$ は、y軸とy'軸との間の角度を表す。この例では、図29のブロック2925は、聴取者座標系3020の原点のまわりの、オーディオ装置座標の、角度 $\theta$ 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3007の原点は、図30A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3020の原点のまわりのオーディオ装置座標の角度 $\theta$ の回転の前に、オーディオ装置座標系3007の原点を、聴取者座標系3020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3007から聴取者座標系3020への座標変換によって実行される。

#### 【0306】

サウンドバー3030および／またはテレビジョン1901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3030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3030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。

#### 【0307】

サウンドバー3030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有する他のいくつかの例では、サウンドバー3030および／またはテレビジョン1901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3030と共位置にされてもよい。

#### 【0308】

いくつかの実装によれば、サウンドバー3030および／またはテレビジョン1901は、関連するカメラ3011を有してもよい。制御システムは、聴取者の頭部3010（および／または聴取者の鼻3025）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3010（および／または聴取者の鼻3025）とカメラ3011との間の直線3013a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

10

20

30

40

50

は、直線3013a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3013aとオーディオ装置座標系のy軸との間の角度 $\theta$ を決定するように構成されてもよい。

#### 【0309】

図30B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29のブロック2915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3000bのスピーカーを制御して、オーディオ・オブジェクト3035を環境3000b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3035が聴取者3005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3035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3035が聴取者座標系3020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3035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3040は、オーディオ・オブジェクト3035が聴取者3005のまわりを回転するときの、オーディオ・オブジェクト3035の軌跡の一部を示す。

#### 【0310】

いくつかのそのような例によれば、聴取者3005は、いつオーディオ・オブジェクト3035が聴取者3005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3035の位置との間の直線3013bを決定するように構成されてもよい。この例では、直線3013bは、聴取者3005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3005は、いつオーディオ・オブジェクト3035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。

#### 【0311】

図30C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29のブロック2915においてすでに決定されている。ここで、聴取者3005は、ハンドヘルド装置3045をテレビ1901またはサウンドバー3030のほうに向けることによって、聴取者3005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3045を使用している。ハンドヘルド装置3045および聴取者の腕の破線の輪郭は、聴取者3005がハンドヘルド装置3045をテレビ1901またはサウンドバー3030のほうへ向けていた時点より前の時点で、聴取者3005がハンドヘルド装置3045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3005は、ハンドヘルド装置3045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3045は、オーディオ装置2とテレビ1901またはサウンドバー3030との間の角度 $\alpha$ を決定するように構成される。該角度 $\alpha$ 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を近似する。

ハンドヘルド装置3045は、いくつかの例では、慣性センサー・システムと、環境3000c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3045は、必要な機能を実行するようにハンドヘルド装置3045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。

#### 【0312】

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ープデータに従って、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 $\alpha$ 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。

#### 【0313】

図30Dは、図30C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3007の原点は、聴取者座標系3020の原点と共位置にされる。オーディオ装置座標系3007と聴取者座標系3020の原点を共位置にすることは、聴取者位置が決定される2915のプロセス後に可能となる。オーディオ装置座標系3007と聴取者座標系3020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3007から聴取者座標系3020に変換することに関わってもよい。角度 $\alpha$ は、図30Cを参照して上述したように決定されている。よって、角度 $\alpha$ は、聴取者座標系3020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 $\beta$ は、オーディオ装置座標系3007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 $\theta$ は、オーディオ装置座標系3007の $y$ 軸を聴取者座標系3020の $y'$ 軸に整列させるのに必要な回転を示す。

#### 【0314】

いくつかの実装では、図29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。

#### 【0315】

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1Aの制御システム110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。

#### 【0316】

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。

#### 【0317】

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

ブ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。図3Aに示されるオーディオ環境の一例では、ラウドスピーカー305d、305fおよび305h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305b、305c、305e、および305f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305g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305a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305dおよび305fの低周波能力は、環境200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。

10

#### 【0318】

図31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31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むうる。

#### 【0319】

この例によれば、システム3100は、スマートホームハブ3105と、ラウドスピーカー3125a~3125mとを含む。この例では、スマートホームハブ3105は、図1Aに示され、上述された制御システム110のインスタンスを含む。この実装によれば、制御システム110は、聴取環境ダイナミクス処理構成データ・モジュール3110、聴取環境ダイナミクス処理モジュール3115、およびレンダリング・モジュール3120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3110、聴取環境ダイナミクス処理モジュール3115、およびレンダリング・モジュール3120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3120'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。

20

#### 【0320】

スマートホームハブ3105とラウドスピーカー3125a~3125m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3105は、図1Aに示され、上述されたインターフェース・システム105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3105は、図3Aに示される環境300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3105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3105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3105は、ラウドスピーカー3125a-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3120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3120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。

30

40

#### 【0321】

いくつかの事例では、ラウドスピーカー3125a~3125mは、図3A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3125a~3125m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3100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。

#### 【0322】

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。

50

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。

#### 【0323】

よって、この例では、ラウドスピーカー3125a~3125m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3125a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。

#### 【0324】

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。

#### 【0325】

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3Aを参照して、テレビ番組がテレビ330に示されており、対応するオーディオが環境300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、環境300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305b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と、ラウドスピーカー305d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305b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305d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305b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必要である。

#### 【0326】

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。

【0327】

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)) を組み合わせて、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 (組み合わされた閾値など) 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、組み合わされた閾値) 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ドにレンダリングする。

【0328】

いくつかの実装によれば、プロセス (a) は、図31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3110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3105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しうる。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3105は、各ラウドスピーカー3125a~3125m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3105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。

【0329】

いくつかの例では、プロセス (b) は、図31の聴取環境ダイナミクス処理モジュール3115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。

【0330】

いくつかの例では、プロセス (c) のレンダリングは、図31のレンダリング・モジュール3120またはレンダリング・モジュール3120'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ドを制限し、それにより、制限されたスピーカーフィードを生成する)。プロセス (d) はたとえば、図31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。

## 【0331】

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。

## 【0332】

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）。

## 【0333】

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。

## 【0334】

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

10

20

30

40

50

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。

#### 【0335】

図32A、32Bおよび32C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3200a、3200b、および3200c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3200a、3200b、および3200c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3200a、3200b、および3200c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装されうる。

#### 【0336】

図32Aのグラフ3200a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3205a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3205a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3Aのラウドスピーカー305d）に適してもよい。

#### 【0337】

図32Bのグラフ3200b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3205bは、図32A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3205bは、曲線3205a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3Aのスピーカー305b）を有するスピーカーに適している。

#### 【0338】

図32Cのグラフ3200c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3205c（この例では直線である）は、図32A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3205c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3205dを実装することによって、曲線3205cを近似することができるが見て取れるであろう。

#### 【0339】

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）または柔軟仮想化（Flexible Virtualization、FV）などの既知のレンダリング・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。

【0340】

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）が組み合わされて、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、組み合わされてもよい。

【0341】

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78】

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

）の単一の集合に組み合わされてもよい。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79】

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

Equation (30)

【0342】

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、それらが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32Aに示されたベース範囲における閾値が、図32C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：  
【数 8 0】

$$\bar{T}[f] = \text{mean}_i(T_i[f]) \quad \text{Equation (31)}$$

#### 【 0 3 4 3 】

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。

#### 【 0 3 4 4 】

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 8 1】

$$\bar{T}[f] = \alpha \text{mean}_i(T_i[f]) + (1 - \alpha) \min_i(T_i[f]) \quad \text{Equation (32)}$$

#### 【 0 3 4 5 】

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。

#### 【 0 3 4 6 】

図33Aおよび33B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3300aおよび3300b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。

#### 【 0 3 4 7 】

図33A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3305e）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3305a）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3305b、3305c、および3305d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。

#### 【 0 3 4 8 】

図33B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3310aおよび閾値より上の直線セグメント3310b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。

【0349】

図33B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10 dB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。

【0350】

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（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。

【0351】

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（32）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。

【0352】

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

【数82】

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

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

【数83】

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

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さ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

【数 8 4】

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

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。

【0 3 5 3】

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。

【0 3 5 4】

図34は、聴取環境の空間ゾーンの例を示す。図34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。

【0 3 5 5】

図34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。

【0 3 5 6】

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 8 5】

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

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 8 6】

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

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。

【0357】

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87】

$$s_j[t] = \sum_{k=1}^K \alpha_{kj}[t] x_k[t] \quad \text{Equation (33)}$$

次いで、各ゾーン信号は、ゾーン閾値

【数88】

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

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89】

$$G_j[f, t] = DP\{s_j[t], \bar{T}_j[f]\} \quad \text{Equation (34)}$$

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90】

$$G_k[f, t] = \sqrt{\sum_{j=1}^J (\alpha_{kj} G_j[f, t])^2} \quad \text{Equation (35)}$$

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91】

$$\hat{x}_k[t]$$

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、これをスピーカー信号にレンダリングされうる。

【0358】

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、多様な仕方で行なわれる。一例として、空間ゾーン再生制限閾値

【数92】

$\bar{T}_j[f]$ 

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 9 3】

$$\bar{T}_j[f] = \sum_i w_{ij}[f] T_i[f] \quad \text{Equation (36)}$$

【 0 3 5 9 】

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。これは、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。

【 0 3 6 0 】

図35は、図34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35は、図34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）に寄与することは、意味がある。

【 0 3 6 1 】

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。

【 0 3 6 2 】

図36は、図35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。

【 0 3 6 3 】

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 9 4】

$$w_{ij}[f] = \frac{g_{ij}[f]}{\sum_i g_{ij}[f]} \quad \text{Equation (37)}$$

10

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36の重みについての望ましい属性である。

【0 3 6 4】

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。

20

【0 3 6 5】

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35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35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。

30

40

【0 3 6 6】

図37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3700のブロックは、本明細書に記載された

50

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3700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3700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3700のブロックは、図1Aに示されて上述した制御システム110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。

#### 【0367】

この例によれば、ブロック3705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。

#### 【0368】

いくつかの事例では、ブロック3705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3705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。

#### 【0369】

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3705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。

#### 【0370】

この例では、ブロック3710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。

## 【0371】

この実装では、ブロック3715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。

## 【0372】

この例では、ブロック3720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3720のダイナミクス処理は、10

## 【0373】

ここで、ブロック3725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3725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、20

## 【0374】

この例によれば、ブロック3730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3730は、スマートホームハブ3105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3125a～3125mに提供することに関わってもよい。30

## 【0375】

いくつかの例では、方法3700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31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3125a～3125m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。

## 【0376】

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。40

## 【0377】

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50

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。

【0378】

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3725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もよい。

【0379】

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。

【0380】

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もよい。

【0381】

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。

【0382】

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。

【0383】

いくつかの実装によれば、処理されたオーディオ・データをレンダリングすることは、一つまたは複数の動的に構成可能な機能に従って、ラウドスピーカーの集合の相対的なアクティブ化を決定することに関わってもよい。いくつかの例は、図10以下を参照して以下に記載される。一つまたは複数の動的に構成可能な機能は、オーディオ信号の一つまたは

複数の属性、ラウドスピーカーの集合の一つまたは複数の属性、または一つまたは複数の外部入力に基づいていてもよい。たとえば、一つまたは複数の動的に構成可能な機能は、一つまたは複数の聴取者に対するラウドスピーカーの近接性；引力位置に対するラウドスピーカーの近接性（引力とは、引力位置に対する、より近い近接性において相対的に、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）；反発力位置に対するラウドスピーカーの近接性（反発力とは、反発力位置に対する、より近い近接性において、相対的により低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）；環境中の他のラウドスピーカーに対する各ラウドスピーカーの能力；他のラウドスピーカーに対するラウドスピーカーの同期性；ウェイクワード性能；またはエコー・キャンセラの性能に基づいていてもよい。

10

#### 【0384】

スピーカーの相対的アクティブ化は、いくつかの例では、スピーカーを通じて再生された場合のオーディオ信号の知覚される空間位置のモデルのコスト関数、オーディオ信号の意図された知覚される空間位置の、スピーカー位置への近接性の尺度、および一つまたは複数の動的に構成可能な機能に基づいていてもよい。

#### 【0385】

いくつかの例では、コスト関数の最小化（少なくとも1つの動的スピーカー・アクティブ化項を含む）は、スピーカーのうちの少なくとも1つのスピーカーの非アクティブ化（そのような各スピーカーが関連するオーディオ・コンテンツを再生しないという意味で）と、スピーカーのうちの少なくとも1つのスピーカーのアクティブ化（そのような各スピーカーがレンダリングされたオーディオ・コンテンツの少なくとも一部を再生するという意味で）につながることもありうる。動的スピーカー・アクティブ化項（単数または複数）は、特定のスマートオーディオ装置から離れたところでのオーディオの空間的提示を歪めることを含む、多様な挙動のうちの少なくとも1つを可能にしうる。それにより、マイクロフォンが話者の声をよりよく聞くことができ、あるいはスマートオーディオ装置のスピーカー（単数または複数）から二次オーディオ・ストリームがよりよく聞こえる。

20

#### 【0386】

いくつかの実装によれば、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。いくつかの事例では、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・データまたはニー・データのうちの一つまたは複数を含んでいてもよい。

30

#### 【0387】

上述のように、いくつかの実装では、図37に示される方法3700の少なくともいくつかのブロックが省略されてもよい。たとえば、いくつかの実装では、ブロック3705および3710は、セットアップ・プロセスの間に実行される。聴取環境ダイナミクス処理構成データが決定された後、いくつかの実装では、ステップ3705および3710は、聴取環境のスピーカーのタイプおよび／または配置が変化しない限り、「ランタイム」動作中に再度実行されることはない。たとえば、いくつかの実装では、いずれかのラウドスピーカーが追加されたまたは切り離されたか、いずれかのラウドスピーカー位置が変化したか、などを決定するために、初期チェックがあってもよい。もしそうであれば、ステップ3705および3710が実施されてもよい。もしそうでなければ、ステップ3705および3710は、ブロック3715～3730に関わってもよい「ランタイム」操作の前に再度実行されなくてもよい。

40

#### 【0388】

図38A、38B、および38Cは、図2Cおよび図2Dの例に対応するスピーカー参加値の例を示す。図38A、38Bおよび38Cにおいて、角度-4.1は図2Dのスピーカー位置272に対応し、角度4.1は図2Dのスピーカー位置274に対応し、角度-87は図2Dのスピーカー位置267に対応し、角度63.6は図2Dのスピーカー位置275に対応し、角度165.4は図2Dのスピーカー位置270に対応する。これらのスピーカー参加値は、図34～図37を参照して記載された空間ゾーンに関する「重み付け」の例である。これらの例によれば、図38A、

50

38Bおよび38Cに示されるラウドスピーカー参加値は、図34に示される空間ゾーンのそれぞれにおける各ラウドスピーカーの参加に対応する：図38Aに示されるラウドスピーカー参加値は、中央ゾーンにおける各ラウドスピーカーの参加に対応し、図38Bに示されるラウドスピーカー参加値は、前方左および右ゾーンにおける各ラウドスピーカーの参加に対応し、図38Cに示されるラウドスピーカー参加値は、後方ゾーンにおける各ラウドスピーカーの参加に対応する。

【0389】

図39A、39Bおよび39Cは、図2Fおよび図2Gの例に対応するラウドスピーカー参加値の例を示す。図39A、39Bおよび39Cにおいて、角度-4.1は図2Dのスピーカー位置272に対応し、角度4.1は図2Dのスピーカー位置274に対応し、角度-87は図2Dのスピーカー位置267に対応し、角度63.6は図2Dのスピーカー位置275に対応し、角度165.4は図2Dのスピーカー位置270に対応する。これらの例によれば、図39A、39Bおよび39Cに示されるラウドスピーカー参加値は、図34に示される各空間ゾーンにおける各ラウドスピーカーの参加に対応する：図39Aに示されるラウドスピーカー参加値は、各ラウドスピーカーの中央ゾーンにおける参加に対応し、図39Bに示されるラウドスピーカー参加値は、各ラウドスピーカーの前方左および右のゾーンにおける参加に対応し、図39Cに示されるラウドスピーカー参加値は、各ラウドスピーカーの後方ゾーンにおける参加に対応する。

【0390】

図40A、40Bおよび40Cは、図2Hおよび図2Iの例に対応するラウドスピーカー参加値の例を示す。これらの例によれば、図40A、40Bおよび40Cにおいて示されるラウドスピーカー参加値は、図34に示される各空間ゾーンにおける各ラウドスピーカーの参加に対応する。図40Aにおいて示されるラウドスピーカー参加値は、中央ゾーンにおける各ラウドスピーカーの参加に対応し、図40Bにおいて示されるラウドスピーカー参加値は、前方左および右ゾーンにおける各ラウドスピーカーの参加に対応し、図40Cにおいて示されるラウドスピーカー参加値は、後方ゾーンにおける各ラウドスピーカーの参加に対応する。

【0391】

図41A、41Bおよび41Cは、図2Jおよび図2Kの例に対応するラウドスピーカー参加値の例を示す。これらの例によれば、図41A、41Bおよび41Cにおいて示されるラウドスピーカー参加値は、図34に示される各空間ゾーンにおける各ラウドスピーカーの参加に対応する。図41Aにおいて示されるラウドスピーカー参加値は、中央ゾーンにおける各ラウドスピーカーの参加に対応し、図41Bにおいて示されるラウドスピーカー参加値は、前方左および右ゾーンにおける各ラウドスピーカーの参加に対応し、図41Cにおいて示されるラウドスピーカー参加値は、後方ゾーンにおける各ラウドスピーカーの参加に対応する。

【0392】

図42は、この例における生活空間である環境の図である。この例によれば、生活空間は、オーディオインタラクションのためのスマートオーディオ装置（装置1.1）、オーディオ出力のためのスピーカー（1.3）、および制御可能な照明（ライト）（1.2）のセットを含む。一例では、装置1.1のみがマイクロフォンを含んでおり、そのためウェイクワード・コマンドを発するユーザー（1.4）がどこにいるのかを知る。さまざまな方法を使用して、これらの装置から集合的に情報が得られて、ウェイクワードを発する（たとえば、話す）ユーザーの位置推定値（たとえば、微細な粒度の位置推定）を提供することができる。

【0393】

そのような居住空間には、人がタスクや活動を行ったり、または閾を越えたりする自然な活動ゾーンの集合がある。これらのアクションエリア（ゾーン）は、インターフェースの他の側面を支援するために、ユーザーの位置（たとえば、不確かな位置を決定する）またはユーザーのコンテキストを推定するための努力があるかもしれない場所である。

【0394】

図1の例では、重要なアクションエリアは以下の通りである：

1. キッチンシンクおよび調理エリア（生活空間の左上領域）；
2. 冷蔵庫のドア（シンクと調理エリアの右）；

3. ダイニングエリア（居住空間の左下領域）；
4. 居住空間のオープンエリア（シンクおよび調理エリアおよびダイニングエリアの右）；
5. TVカウチ（オープンエリアの右）；
6. テレビ自体；
7. テーブル；
8. ドアエリアまたは入口（居住空間の右上領域）。

#### 【0395】

しばしば、アクションエリアに合うよう、同じような位置にある同じような数のライトがあることは明らかである。ライトの一部または全部は、個々に制御可能なネットワーク接続されたエージェントであってもよい。いくつかの実施形態によれば、オーディオは、スピーカー（および／または装置（1.1）のうちの一つまたは複数のスピーカー）のうちの一つまたは複数による（本発明の方法の任意の実施形態に従った）再生のために（たとえば、図42のシステムの装置1.1のうちの1つまたは他の装置によって）レンダリングされる。

#### 【0396】

あるクラスの実施形態は、複数の協調させられる（オーケストレーションされる）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）によって、再生のためにオーディオをレンダリングするおよび／またはオーディオを再生する方法である。たとえば、ユーザーの家庭において（システムにおいて）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による（すなわち、全部または一部のスマートオーディオ装置のスピーカーによる）再生のためのオーディオの柔軟なレンダリングを含む、多様な同時の使用事例を処理するために、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要求する、本システムとの多くのインタラクションが考えられる。このような修正は、空間的忠実性に焦点を当ててもよいが、必ずではない。

#### 【0397】

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

#### 【0398】

いくつかの実施形態（たとえば、レンダリング・システムもしくはレンダラーもしくはレンダリング方法、または再生システムもしくは方法）は、一組のスピーカーの一部もしくは全部のスピーカー（すなわち、各アクティブ化されているスピーカー）による、再生のためのオーディオのレンダリングおよび／または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において、スピーカーは、スマートオーディオ装置の協調させられた（オーケストレーションされた）集合のスピーカーである。

#### 【0399】

多くの実施形態が技術的に可能である。当業者には、本開示から、それらをどのように実装するかは明らかであろう。いくつかの実施形態がここに記載される。

#### 【0400】

本開示のいくつかの側面は、開示された方法の一つまたは複数の例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つまたは複数の例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。たとえば、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、開示される方法またはそのステップの実施形態を含む、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つまたは複数の例（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／または他の仕方で作られた）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。

#### 【0401】

いくつかの実施形態は、開示された方法の一つまたは複数の例の実行を含む、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、および他の方法で作られた）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。あるいはまた、開示された方法の実施形態（またはその要素）は、開示された方法の一つまたは複数の例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／または他の仕方で作ら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。あるいはまた、本発明のシステムのいくつかの実施形態の要素は、開示された方法の一つまたは複数の例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）を含む。開示された方法の一つまたは複数の例を実行するように構成された汎用プロセッサは、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。

#### 【0402】

本開示の別の側面は、開示される方法の一つまたは複数の例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）である。

#### 【0403】

さまざまな特徴および側面は、以下の箇条書き例示の実施形態（enumerated example embodiments、EEE）から理解されるであろう。

#### 【0404】

EEE1. 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つのスマートオーディオ装置の少なくとも2つのスピーカーによる再生のためのオーディオのレンダリングのための方法であって、前記オーディオは一つまたは複数のオーディオ信号であり、各オーディオ信号は、関連付けられた所望される知覚される空間位置をもち、スピーカーの前記集合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上で再生される前記オーディオ信号の知覚される空間位置のモデル、前記オーディオ信号の所望される知覚される空間位置の、前記スピーカーの位置への近さ、および前記オーディオ信号の少なくとも一つまたは複数の属性、スピーカーの前記集合の一つまたは複数の属性、または一つまたは複数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の関数である、方法。

#### 【0405】

EEE2. 前記追加的な動的に構成可能な機能が、スピーカーの一または複数の聴取者への近さ；スピーカーの引力または反発力への近さ；スピーカーの、何らかの位置に関する可聴性；スピーカーの能力；スピーカーの、他のスピーカーに対する同期；ウェイワード性能；またはエコー・キャンセラ性能のうちの少なくとも1つを含む。

#### 【0406】

EEE3. 請求項EEE1または2に記載の方法であって、前記レンダリングは、コスト関数の最小化を含み、前記コスト関数は、少なくとも1つの動的スピーカー・アクティブ化項を含む、方法。

#### 【0407】

EEE4. スピーカーの集合のうちの少なくとも2つのスピーカーによる再生のためのオーディオのレンダリングのための方法であって、前記オーディオは一つまたは複数のオーデ

ィオ信号であり、各オーディオ信号は、関連付けられた所望される知覚される空間位置をもち、スピーカーの前記集合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上で再生される前記オーディオ信号の知覚される空間位置のモデル、前記オーディオ信号の所望される知覚される空間位置の、前記スピーカーの位置への近さ、および前記オーディオ信号の少なくとも一つまたは複数の属性、スピーカーの前記集合の一つまたは複数の属性、または一つまたは複数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の関数である、方法。

【0408】

EEE5. 前記追加的な動的に構成可能な機能が、スピーカーの一または複数の聴取者への近さ；スピーカーの引力または反発力への近さ；スピーカーの、何らかの位置に関する可聴性；スピーカーの能力；スピーカーの、他のスピーカーに対する同期；ウェイクワード性能；またはエコー・キャンセラ性能のうちの少なくとも1つを含む。

10

【0409】

EEE6. 請求項EEE4または5に記載の方法であって、前記レンダリングは、コスト関数の最小化を含み、前記コスト関数は、少なくとも1つの動的スピーカー・アクティブ化項を含む、方法。

【0410】

EEE7. オーディオ・レンダリング方法であって：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集合を、2つ以上のラウドスピーカーの集合にレンダリングすることを含み、ラウドスピーカーの前記集合の相対的アクティブ化は、それらのラウドスピーカー上で再生される前記オーディオ信号の知覚される空間位置のモデル、前記オーディオ・オブジェクトの所望される知覚される空間位置の、前記ラウドスピーカーの位置への近さ、および少なくともオーディオ信号の前記集合の一つまたは複数の属性、ラウドスピーカーの前記集合の一つまたは複数の属性、または一つまたは複数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の関数である。

20

【0411】

EEE1A. オーディオ処理方法であって：

第1のオーディオ・プログラム・ストリームを受領する段階であって、前記第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み、前記第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 (spatial data) を含み、前記第1の空間データは、チャンネル・データまたは空間メタデータ (spatial metadata) のうちの少なくとも1つを含む、段階と；

30

前記環境のスピーカーを介した再生するために前記第1のオーディオ信号をレンダリングして、第1のレンダリングされたオーディオ信号を生成する段階と；

前記環境の少なくとも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第2のオーディオ信号を含む第2のオーディオ・プログラム・ストリームを受領する段階と；

前記環境の少なくとも1つのスピーカーを介した再生のために前記第2のオーディオ信号をレンダリングして、第2のレンダリングされたオーディオ信号を生成する段階と；

40

少なくとも部分的に前記第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れら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段階と；

少なくとも部分的に前記第1のオーディオ信号、前記第1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つに基づいて前記第2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2のレンダリングされたオーディオ信号を生成する段階と；

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と前記修正された第2のレンダ

50

リングされたオーディオ信号を混合して、混合オーディオ信号を生成する段階と；

前記混合オーディオ信号を前記環境の少なくともいくつかのスピーカーに提供する段階とを含む、方法。

【0412】

EEE2A. 請求項EEE1Aに記載の方法であって、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、以下のうちの一つまたは複数を実行することを含む方法：

前記第1のオーディオ信号のレンダリングを、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一つまたは複数のもののラウドネスを修正すること。

【0413】

EEE3A. 請求項EEE1Aまたは請求項EEE2Aに記載の方法であって、さらに：

第1のマイクロフォンから第1のマイクロフォン信号を受領する段階と；

少なくとも部分的に前記第1のマイクロフォン信号に基づいて、前記第1のオーディオ信号または前記第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段階とを含む、

方法。

【0414】

EEE4A. 請求項EEE3Aに記載の方法であって、さらに：

前記第1のマイクロフォン信号に基づいて第1の音源位置を推定する段階と；

少なくとも部分的に前記第1の音源位置に基づいて、前記第1のオーディオ信号または前記第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段階とを含む、

方法。

【0415】

EEE5A. 請求項EEE3Aまたは請求項EEE4Aに記載の方法であって、さらに：

前記第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段階と；

少なくとも部分的には、前記第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基づいて、前記第1のオーディオ信号または前記第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段階とを含む、

方法。

【0416】

EEE6A. 請求項EEE3AないしEEE5Aのうちいずれか一項に記載の方法であって、さらに：

前記第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段階と；

少なくとも部分的には、前記第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に基づいて、前記第1のオーディオ信号または前記第2のオーディオ信号のうち少なくとも1つについてのレンダリング・プロセスを修正する段階とを含む、

方法

EEE7A. 少なくとも部分的に前記第1のマイクロフォン信号に基づいて、再生された第1のオーディオ・プログラム・ストリームまたは再生された第2のオーディオ・プログラム・ストリームについてのラウドネス推定値を導出する段階をさらに含む、請求項EEE3Aないし6Aのうちいずれか一項に記載の方法。

【0417】

EEE8A. 少なくとも部分的に前記ラウドネス推定値に基づいて、前記第1のオーディオ信号または前記第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段階をさらに含む、請求項EEE7Aに記載の方法。

## 【0418】

EEE9A. 請求項EEE8Aに記載の方法であって、前記ラウドネス推定値は知覚ラウドネス推定値であり、前記レンダリング・プロセスを修正することは、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、前記第1のオーディオ信号または前記第2のオーディオ信号のうち少なくとも1つを変更することを含む、方法。

## 【0419】

EEE10A. 請求項EEE4Aに記載の方法であって、さらに：

前記第1のマイクロフォン信号が人間の声に対応することを判別する段階と；

前記第1の音源位置とは異なる前記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前記第1のマイクロフォン信号を再生する段階とを含む、方法。

10

## 【0420】

EEE11A. さらに、前記第1のマイクロフォン信号が子供の泣き声に対応することを判別する段階を含み、前記環境の位置は、介護者の推定された位置に対応する、請求項EEE10Aに記載の方法。

## 【0421】

EEE12A. 請求項EEE10Aに記載の方法であって、前記第1のオーディオ信号についてのレンダリング・プロセスを修正することは、前記環境の位置に近い一つまたは複数のスピーカーにおける前記修正された第1のレンダリングされたオーディオ信号のラウドネスを低減する段階を含む、方法。

20

## 【0422】

EEE13A. 請求項EEE1Aないし10Aのうちいずれか一項に記載の方法であって、前記第1のオーディオ信号をレンダリングすることまたは前記第2のオーディオ信号をレンダリングすることのうちの少なくとも1つは、任意に位置されたスピーカーへの柔軟なレンダリングに関わる、方法。

## 【0423】

EEE14A. 前記柔軟なレンダリングは、質量中心振幅パンまたは柔軟仮想化に関わる、請求項EEE13Aに記載の方法。

## 【0424】

EEE1B. 環境のスピーカーを制御する方法であって：

30

第1のオーディオ・プログラム・ストリームを受領する段階であって、前記第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み、前記第1のオーディオ・プログラム・ストリームは、第1の空間データを含み、前記第1の空間データは、チャンネル・データまたは空間メタデータの少なくとも1つを含む、段階と；

前記環境のスピーカーを介した同時再生のために前記第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段階と；

第1のマイクロフォンから第1のマイクロフォン信号を受領する段階と；

前記第1のマイクロフォン信号に基づいて第1の音源位置を推定する段階と；

少なくとも部分的に前記第1のマイクロフォン信号に基づいて前記第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段階であって、前記第1のオーディオ信号についてのレンダリング・プロセスを修正することは：

40

前記第1のマイクロフォン信号に基づいて、前記第1のオーディオ信号のレンダリングを前記第1の音源位置から離れるようにワープさせること、または

少なくとも部分的には前記第1の音源位置に基づいて、前記第1のレンダリングされたオーディオ信号のラウドネスを修正すること

のうちの一つまたは複数を実行することに関わる、段階と；

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、前記環境のスピーカーの少なくとも一部に提供する段階とを含む、

50

方法。

【0425】

EEE2B. 前記第1のマイクロフォン信号が環境ノイズに対応することを判別することをさらに含み、前記第1のオーディオ信号についてのレンダリング・プロセスを修正することは、前記第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前記第1の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、前記第1の音源位置の近くのスピーカーによって再生される第1のレンダリングされたオーディオ信号のラウドネスを増加させることを含む、EEE1Bに記載の方法。

【0426】

EEE3B. さらに、前記第1のマイクロフォン信号が人間の声に対応することを判別する段階を含み、前記第1のオーディオ信号についてのレンダリング・プロセスを修正することは、前記第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前記第1の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、前記第1の音源位置の近くのスピーカーによって再生される前記第1のレンダリングされたオーディオ信号のラウドネスを低減することを含む、EEE1Bに記載の方法。

10

【0427】

EEE4B. 請求項EEE3Bの方法であって、さらに：

前記第1のマイクロフォン信号がウェイクワードに対応することを判別する段階と；

ウェイクワードへの返答を決定する段階と；

前記返答を再生するよう、前記第1の音源位置の近くの少なくとも1つのスピーカーを制御する段階とを含む、

20

方法。

【0428】

EEE5B. 請求項EEE3Bに記載の方法であって、さらに：

前記第1のマイクロフォン信号がコマンドに対応することを判別する段階と；

前記コマンドに対する返答を決定する段階と；

前記返答を再生するよう、前記第1の音源位置の近くの少なくとも1つのスピーカーを制御する段階と；

前記コマンドを実行する段階とを含む、

方法。

30

【0429】

EEE6B. 前記返答を再生するよう、前記第1の音源位置の近くの少なくとも1つのスピーカーを制御した後に、前記第1のオーディオ信号についての修正されていないレンダリング・プロセスに戻す段階をさらに含む、請求項EEE5Bに記載の方法。

【0430】

EEE7B. 請求項EEE1Bに記載の方法であって、さらに：

前記第1のマイクロフォン信号が人間の声に対応することを判別する段階と；

前記環境の位置の近くの一つまたは複数のスピーカーにおいて前記第1のマイクロフォン信号を再生する段階とを含む、

方法。

40

【0431】

EEE8B. 前記第1のマイクロフォン信号が子供の泣き声に対応することを判別する段階であって、前記環境の位置が介護者の推定位置に対応する、段階をさらに含む、請求項EEE7Bに記載の方法。

【0432】

EEE9B. 前記第1のオーディオ信号についてのレンダリング・プロセスを修正することが、前記第2の位置の近くの一つまたは複数のスピーカーにおける修正された第1のレンダリングされたオーディオ信号のラウドネスを低減することに関わる、請求項EEE7Bに記載の方法。

【0433】

50

本開示の個別的な実施形態および本開示の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、本明細書に記載され特許請求される開示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。本開示のある種の形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。

【図面】

【図 1 A】

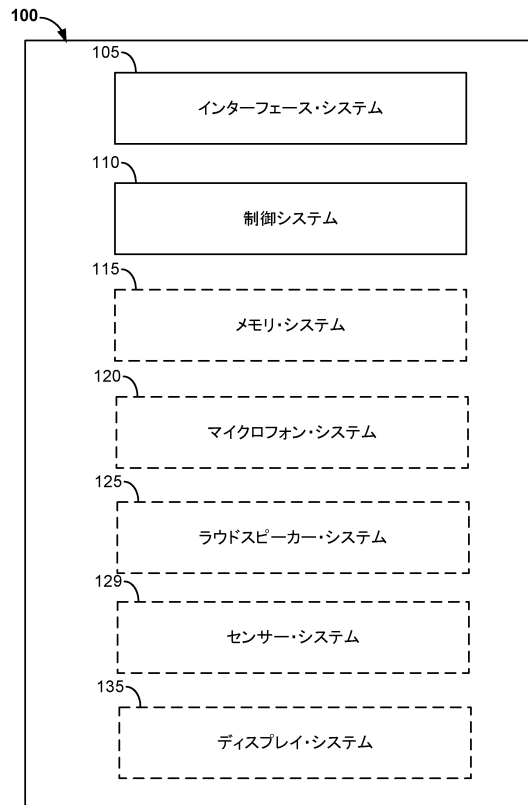


Figure 1A

【図 1 B】

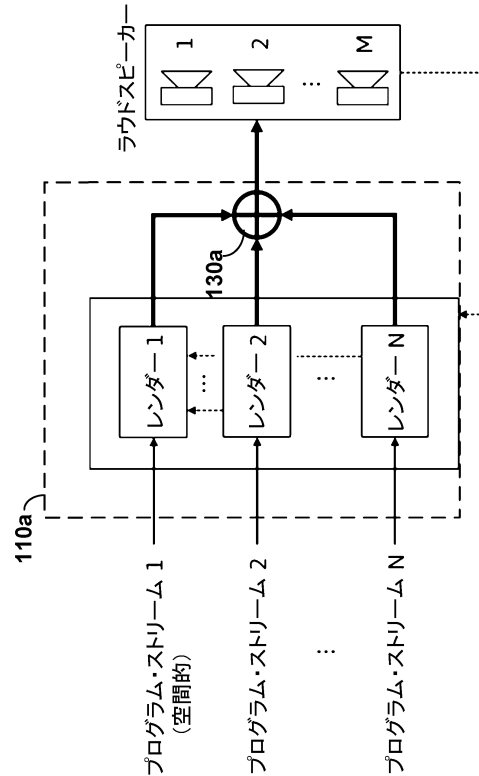


Figure 1B

10

20

30

40

50

【図 2 A】

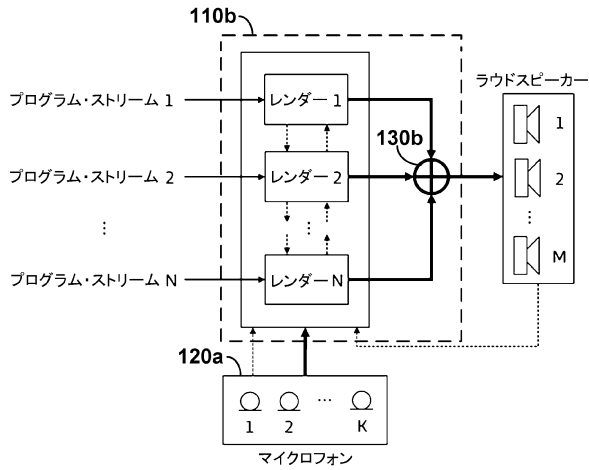


Figure 2A

【図 2 B】

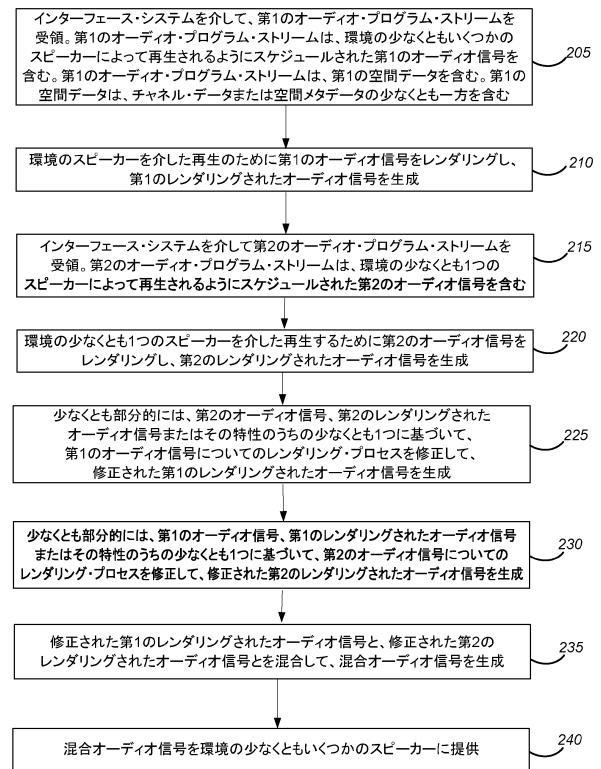


Figure 2B

【図 2 C】

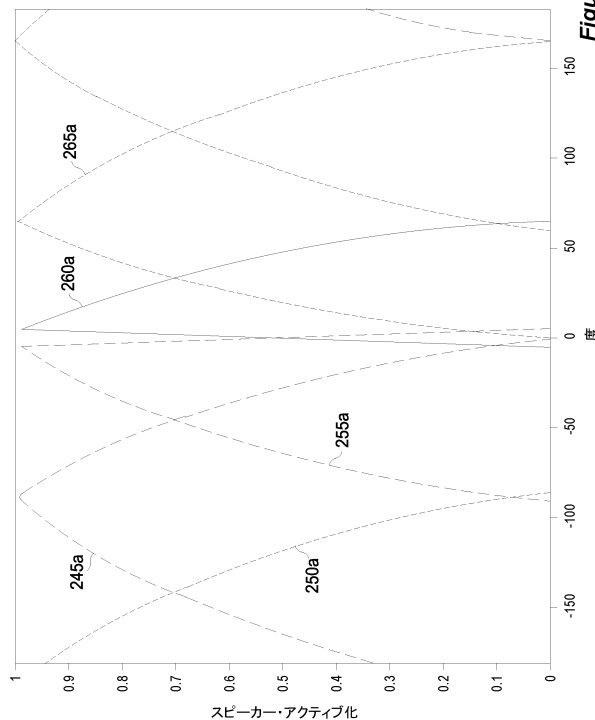


Figure 2C

【図 2 D】

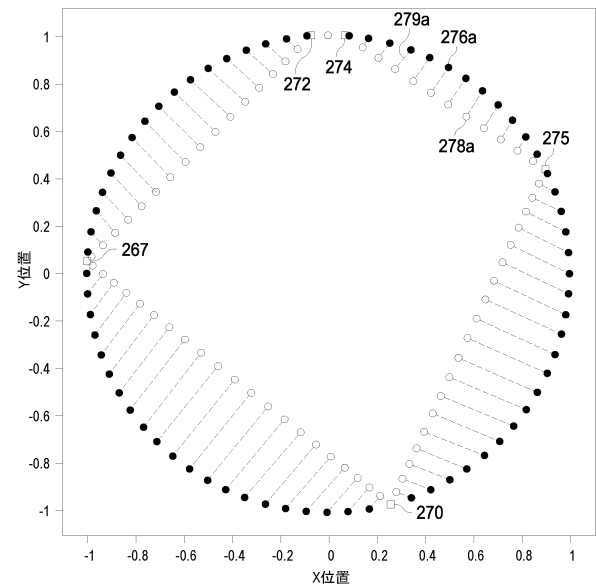


Figure 2D

10

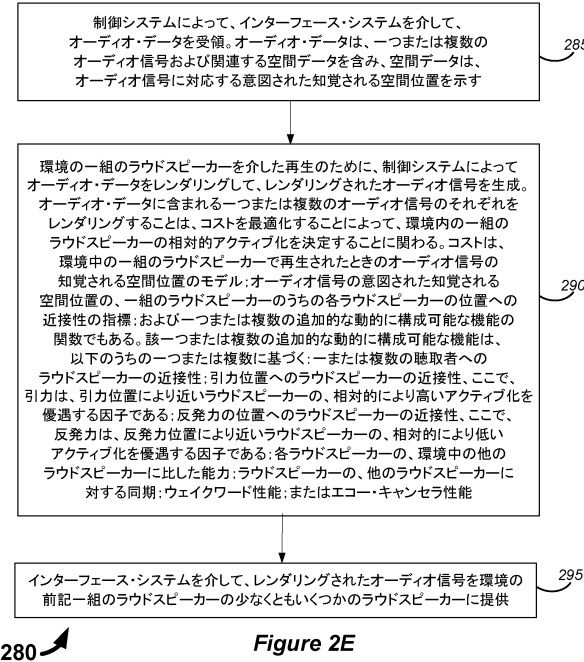
20

30

40

50

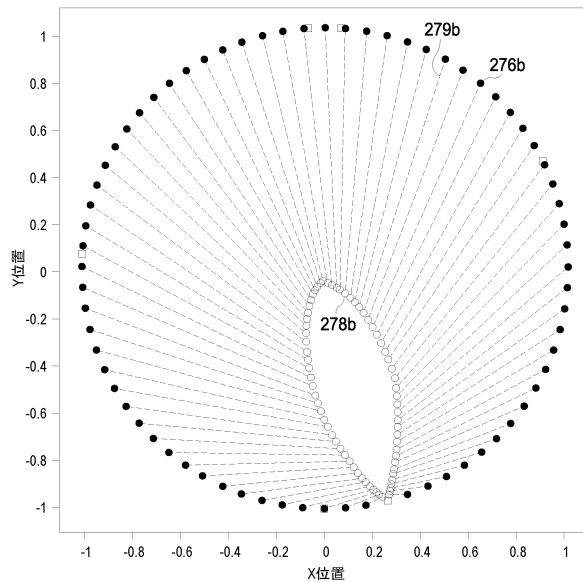
【図 2 E】



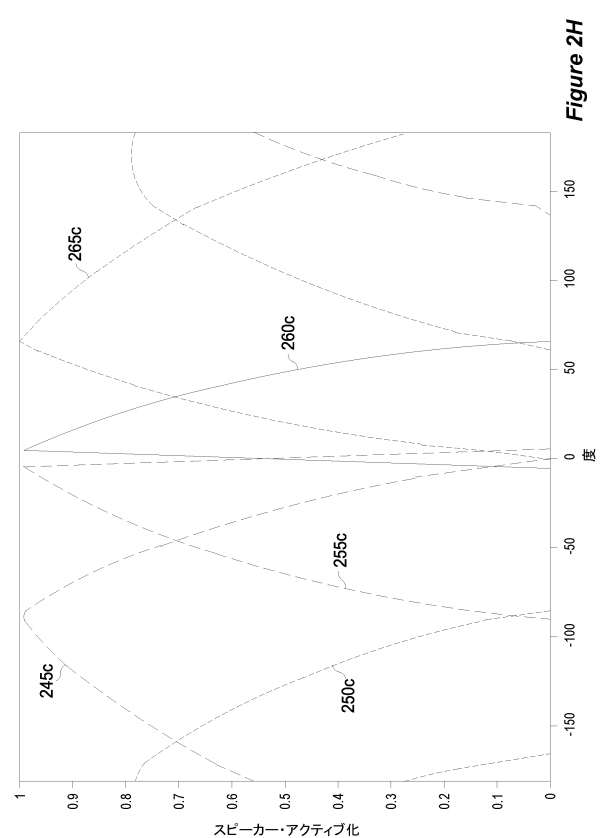
【図 2 F】



【図 2 G】



【図 2 H】



10

20

30

40

50

【図 2 I】

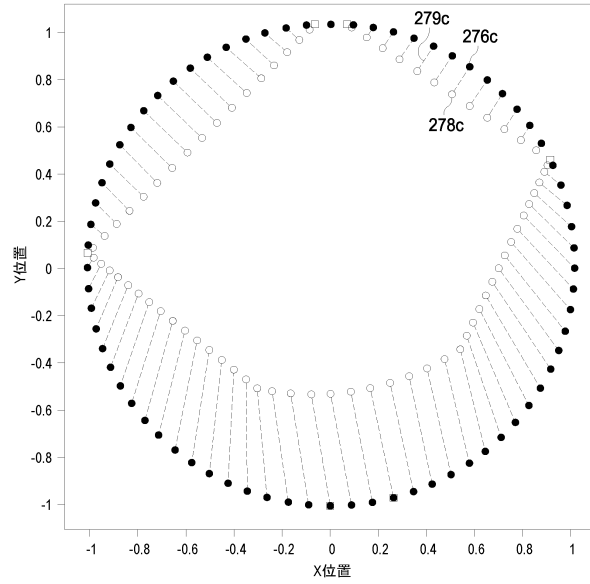


Figure 2I

【図 2 J】

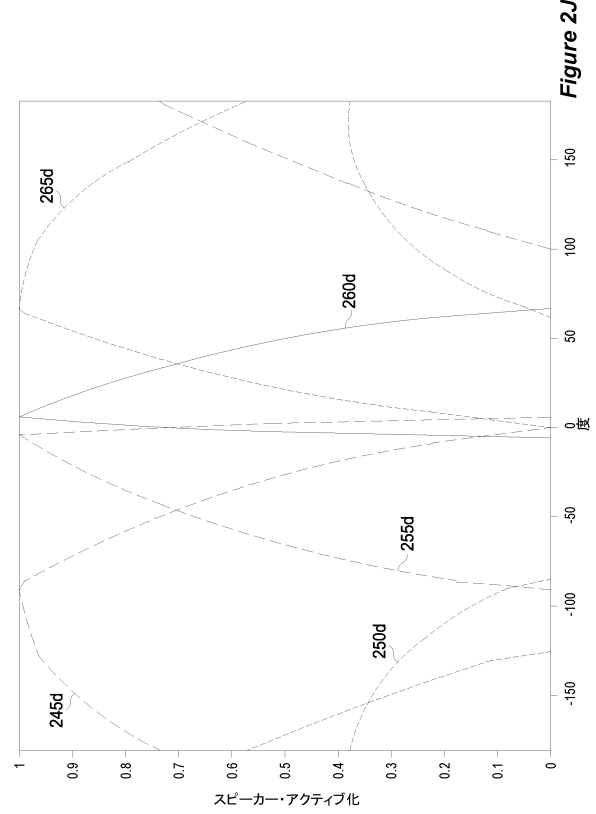


Figure 2J

【図 2 K】

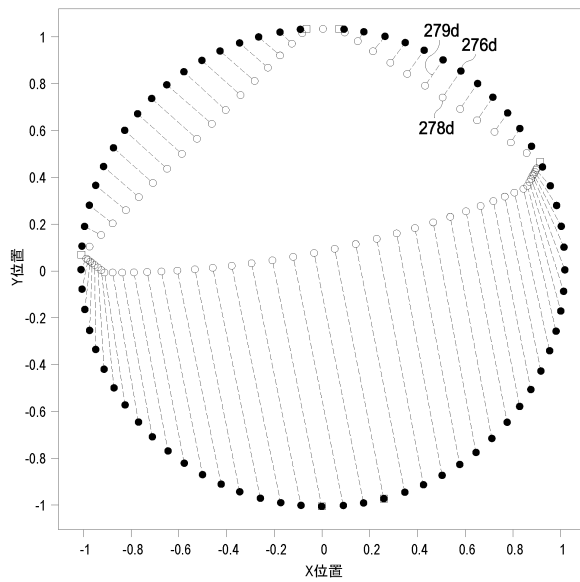


Figure 2K

【図 3 A】

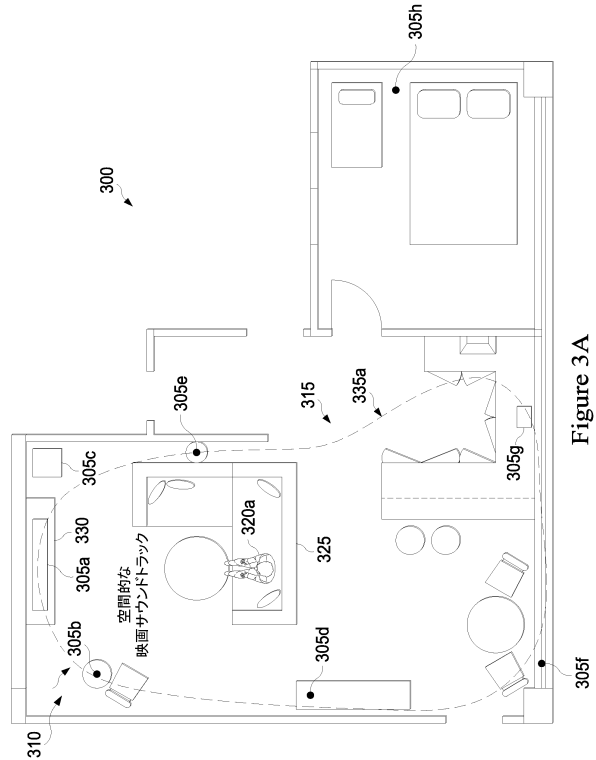


Figure 3A

10

20

30

40

50

【図 3 B】

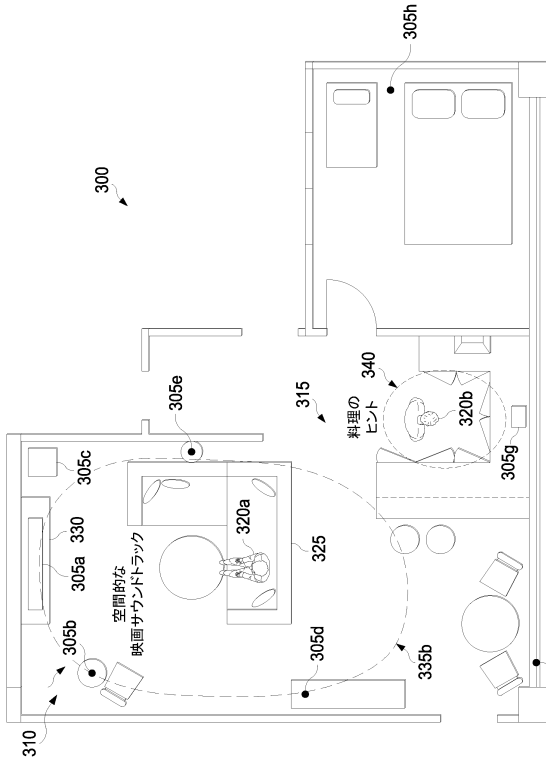


Figure 3B

【図 4 A】

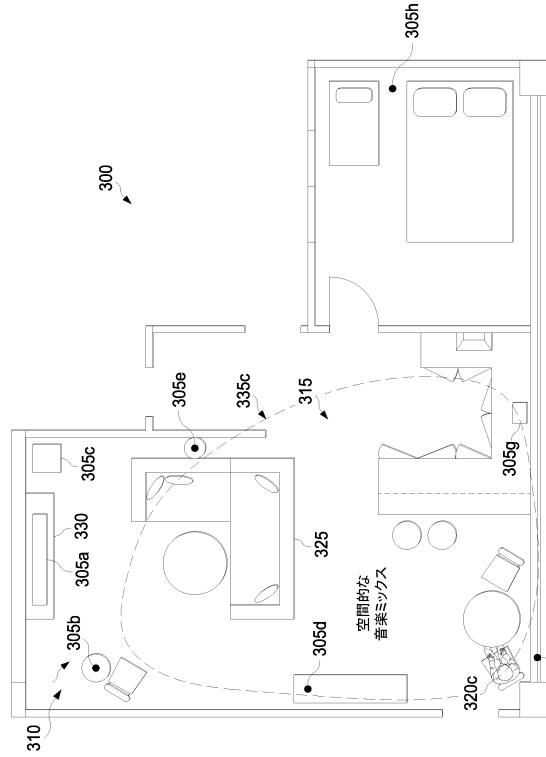


Figure 4A

【図 4 B】

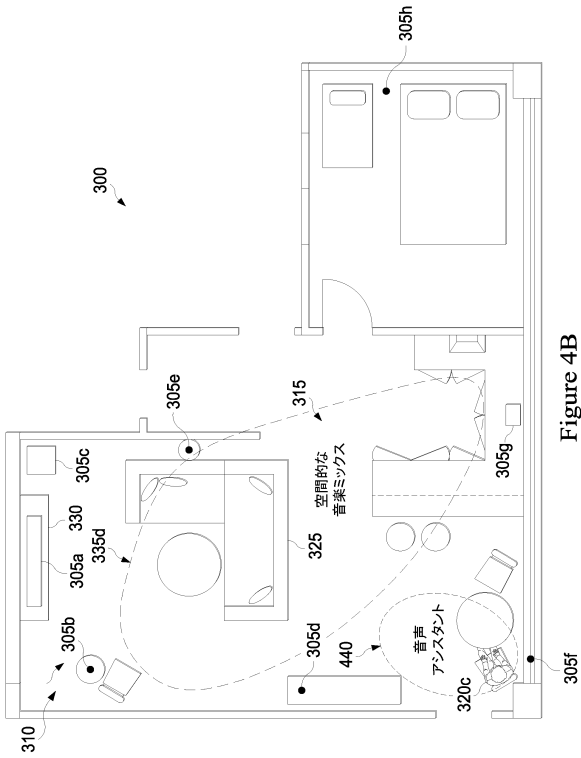


Figure 4B

【図 5 A】

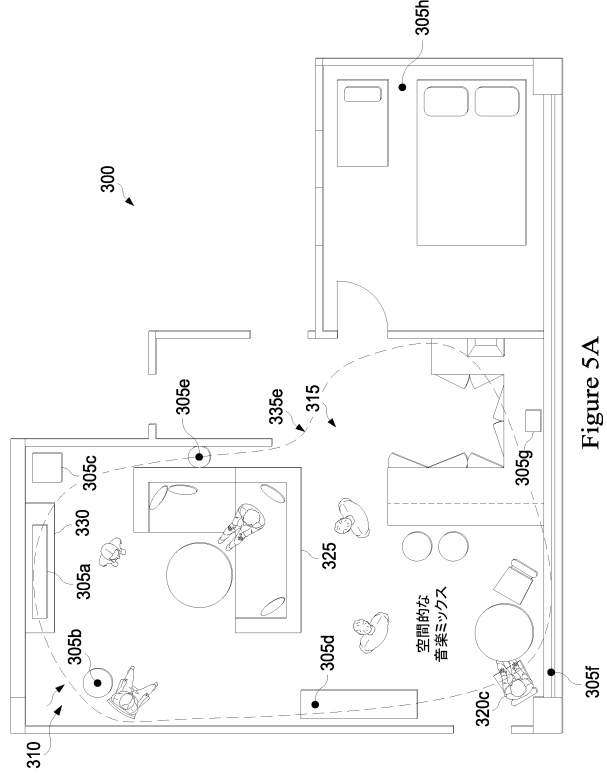


Figure 5A

10

20

30

40

50

【図 5 B】

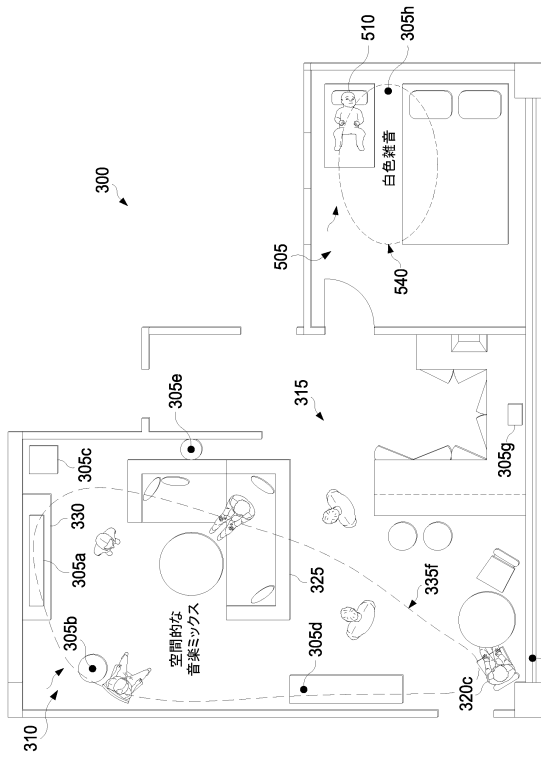


Figure 5B

【図 5 C】

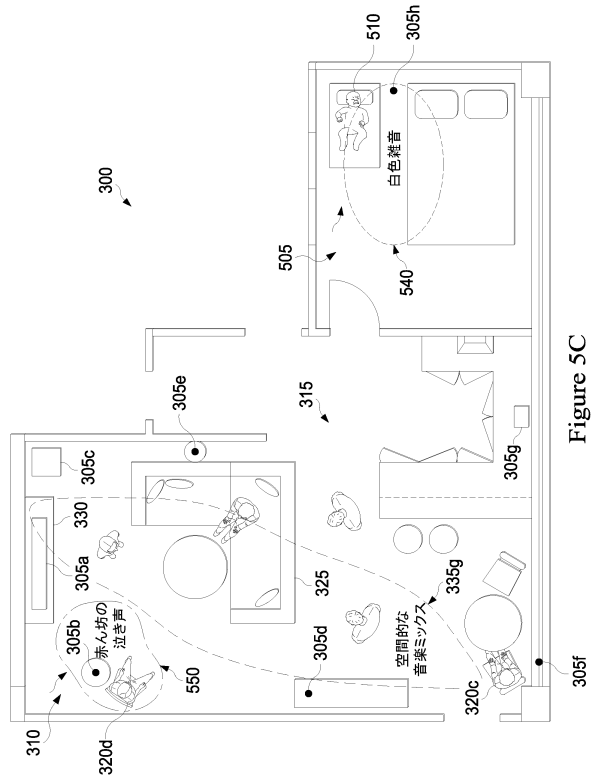


Figure 5C

【図 6】

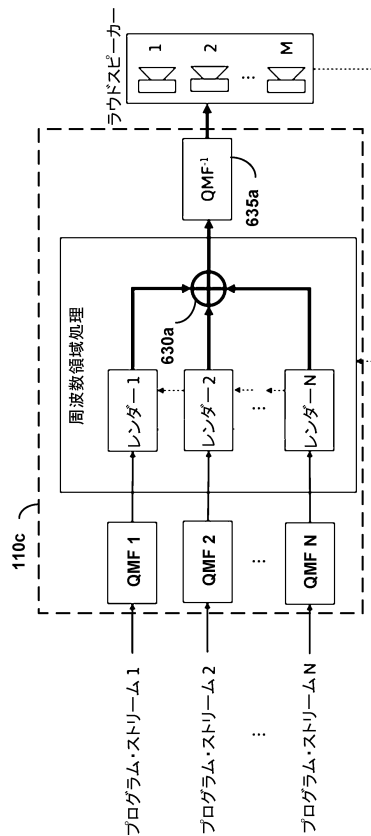


Figure 6

【図 7】

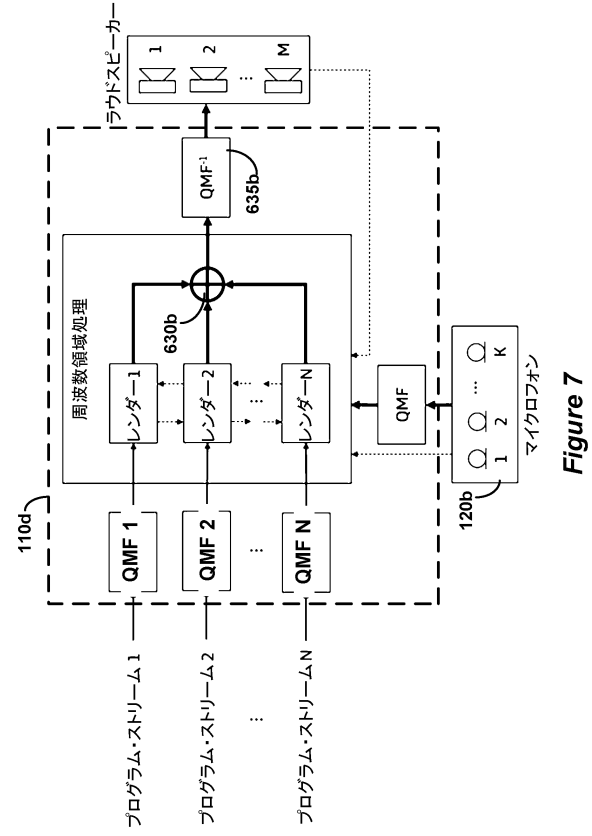


Figure 7

10

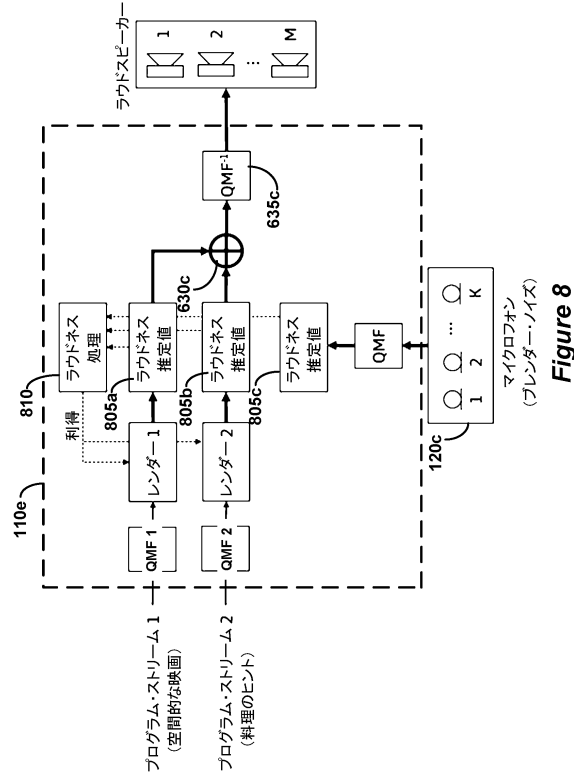
20

30

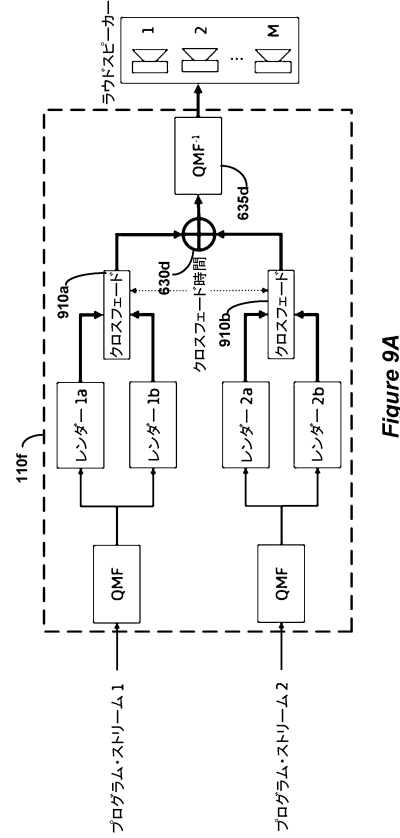
40

50

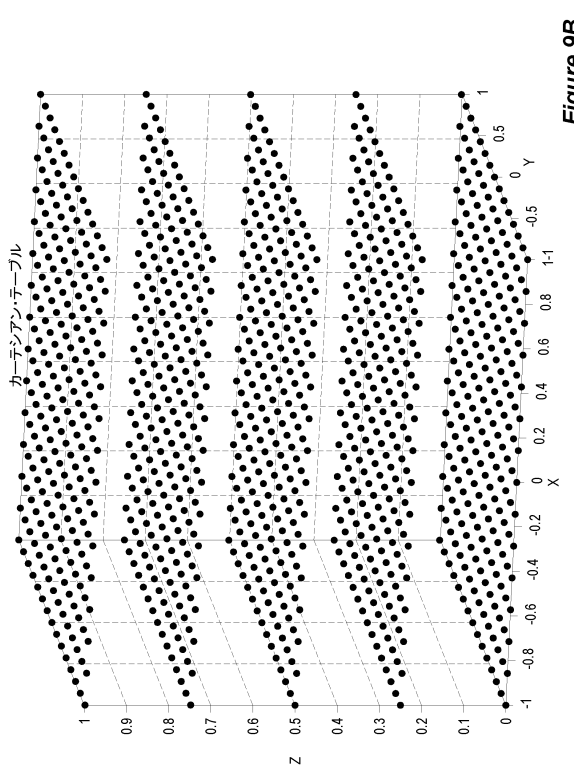
【図 8】



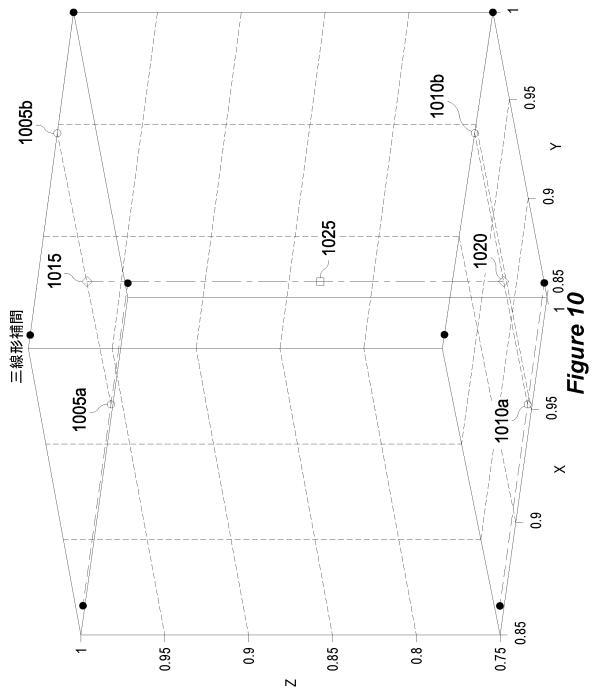
【図 9 A】



【図 9 B】



【図 10】



【図 1 1】

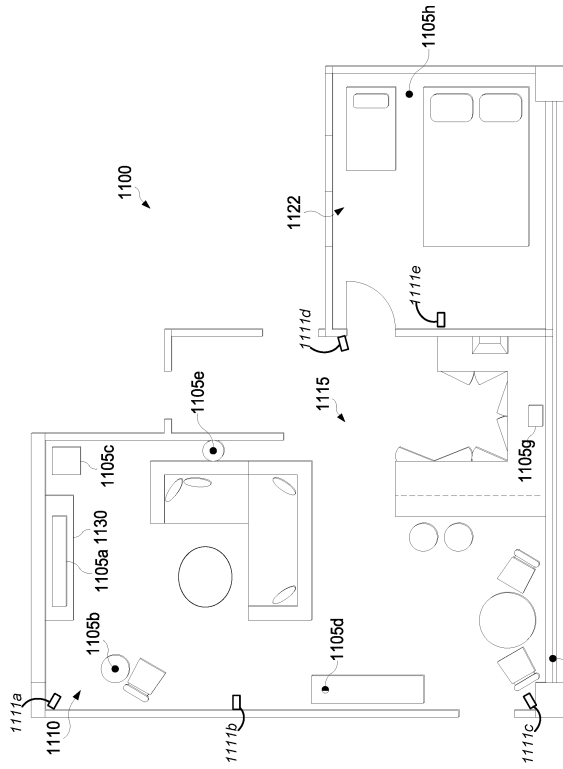


Figure 11

【図 1 2 A】

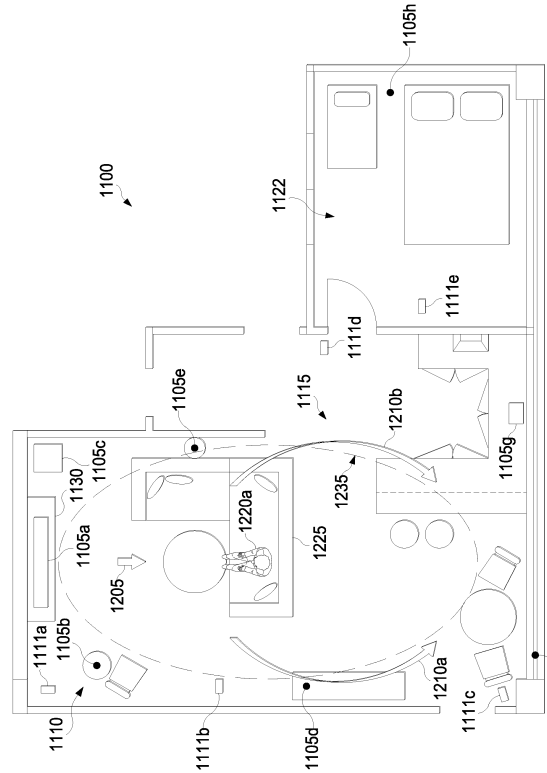


Figure 12A

【図 1 2 B】

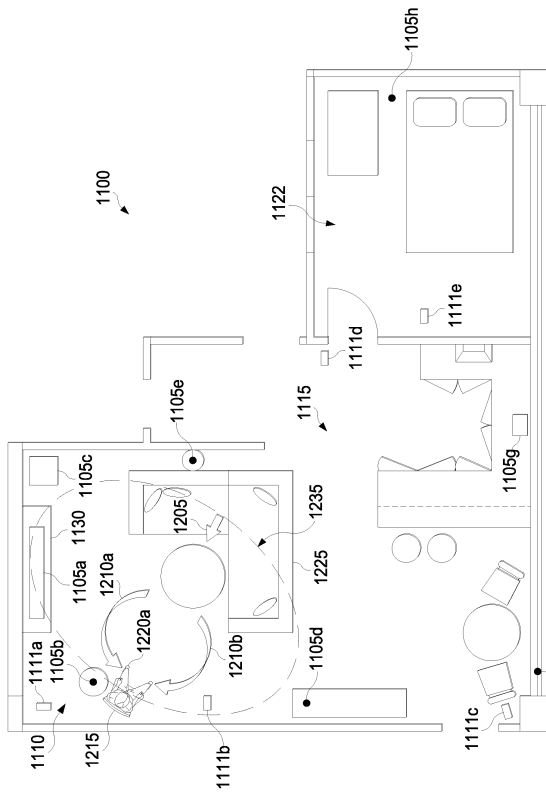


Figure 12B

【図 1 2 C】

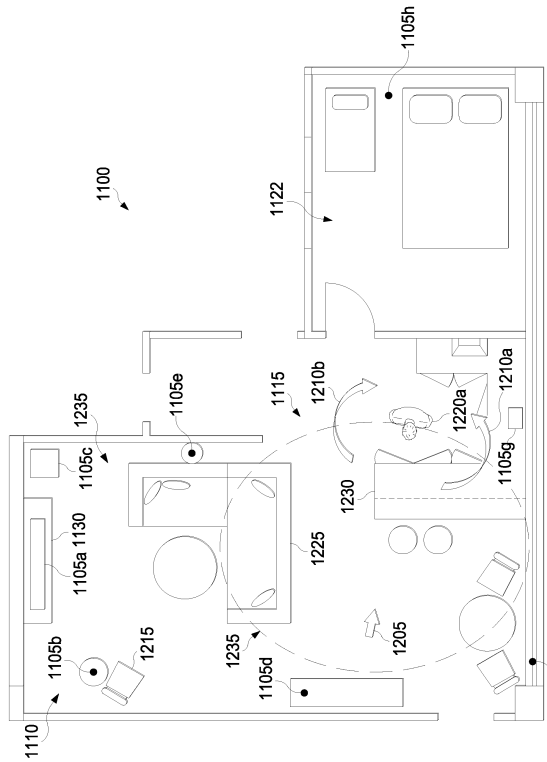


Figure 12C

10

20

30

40

50

【図 1 2 D】

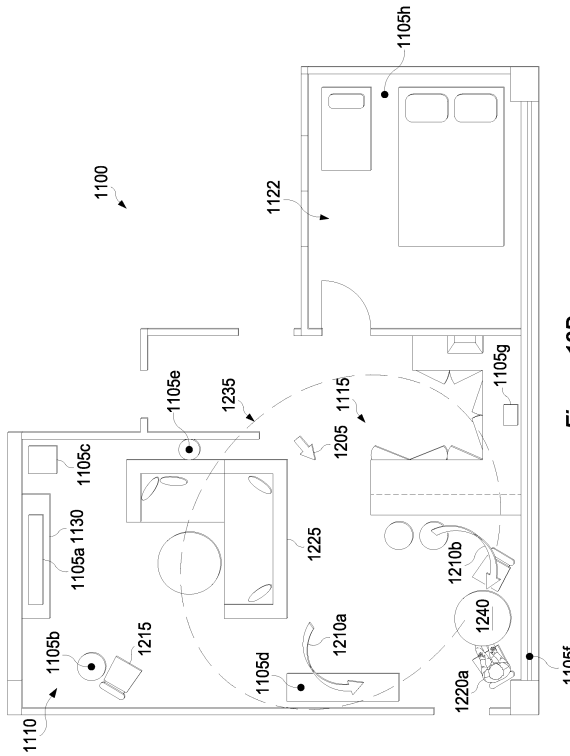


Figure 12D

【図 1 2 E】

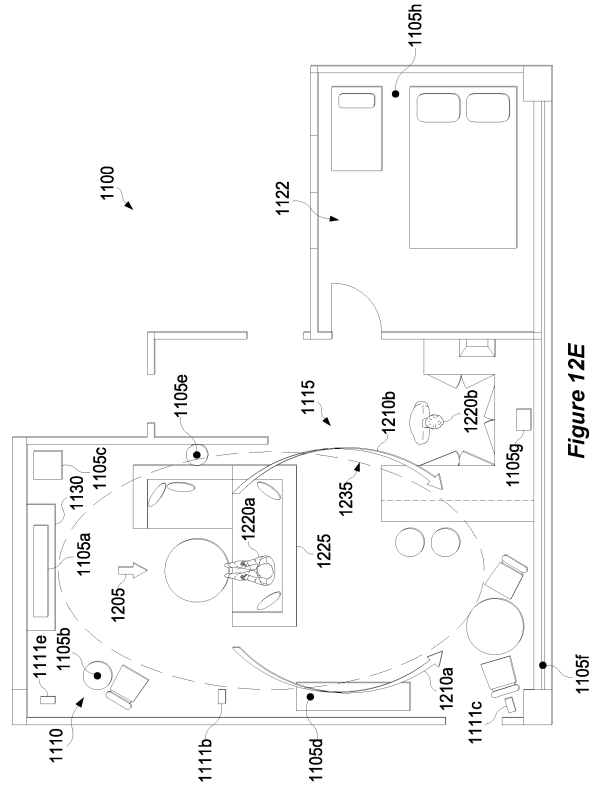


Figure 12E

【図 1 3 A】

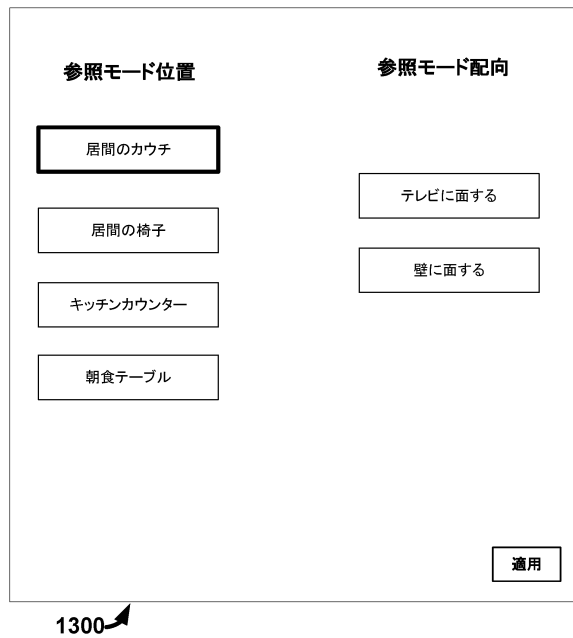


Figure 13A

【図 1 3 B】

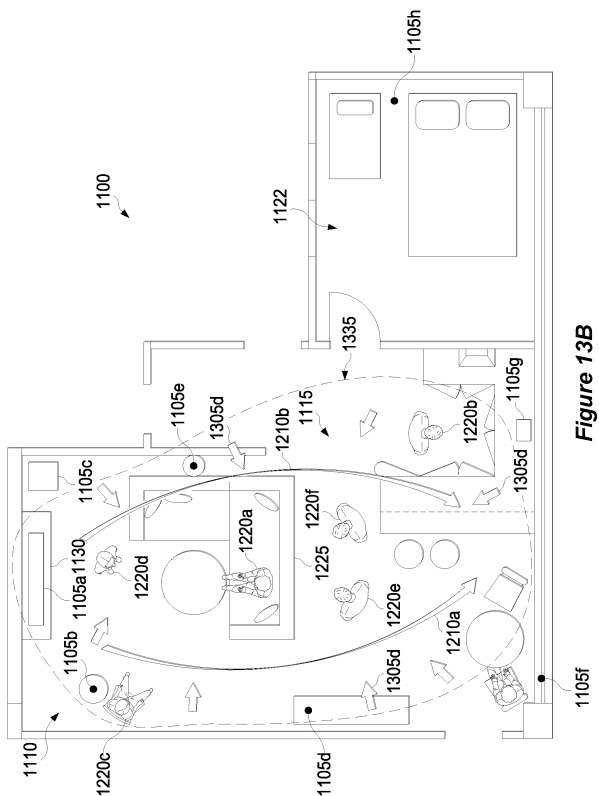


Figure 13B

10

20

30

40

50

【図 14 A】

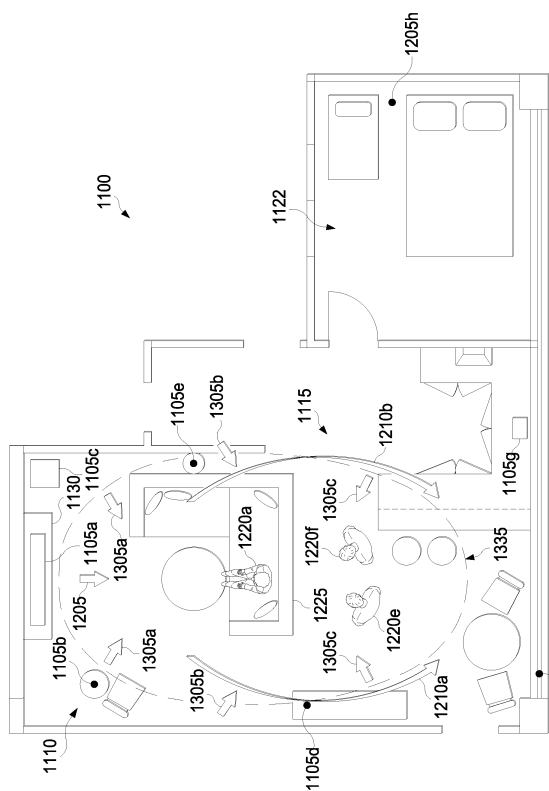


**Figure 14A**

【図 14 B】

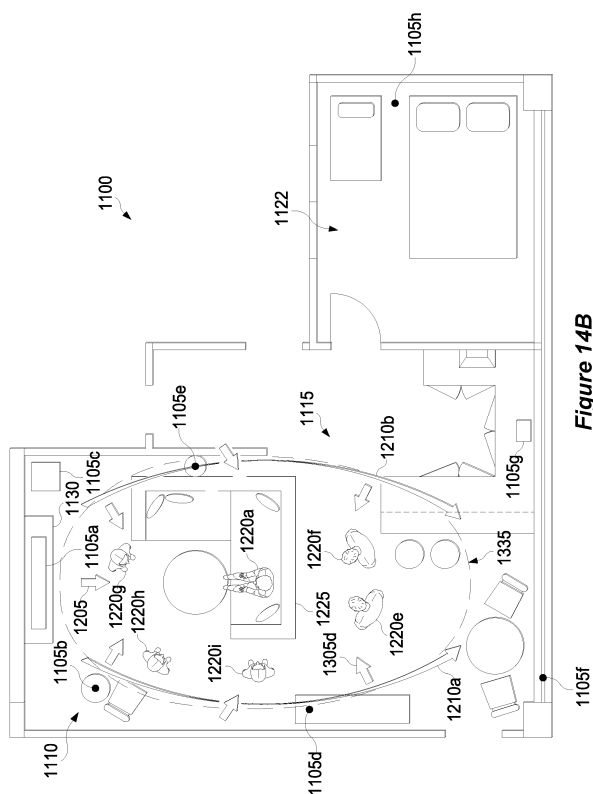


**Figure 14B**

【図 15】

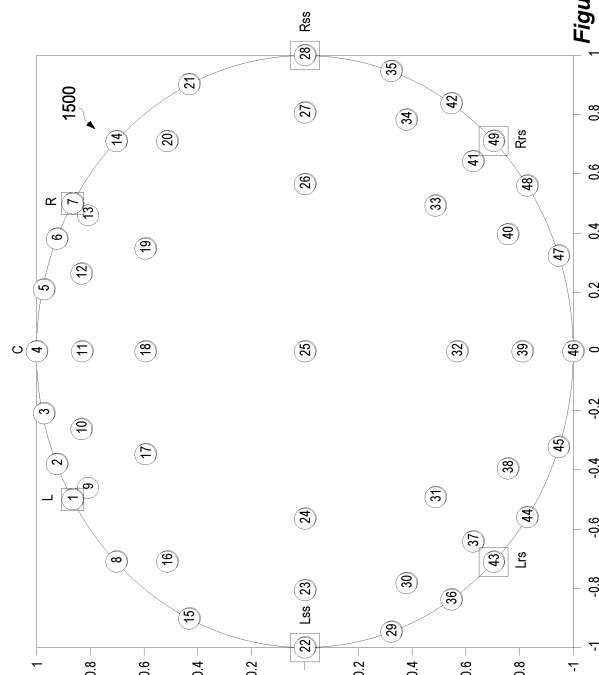


**Figure 15**

【図 16 A】

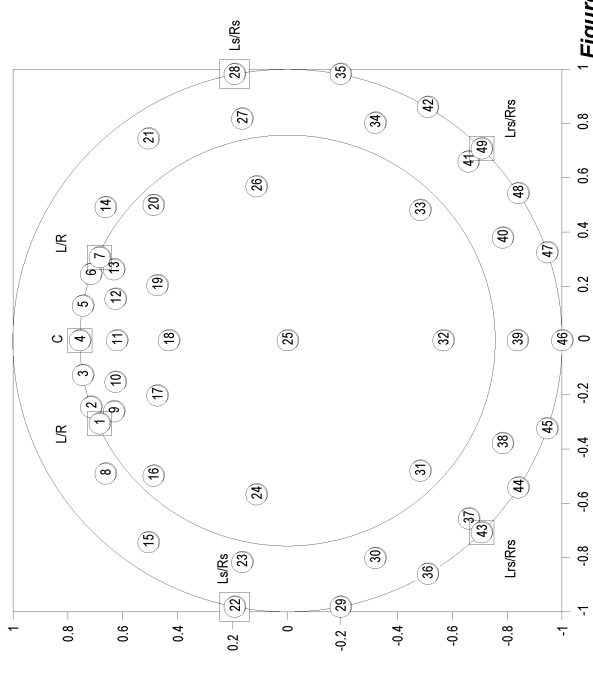


**Figure 16A**

【図 16 B】

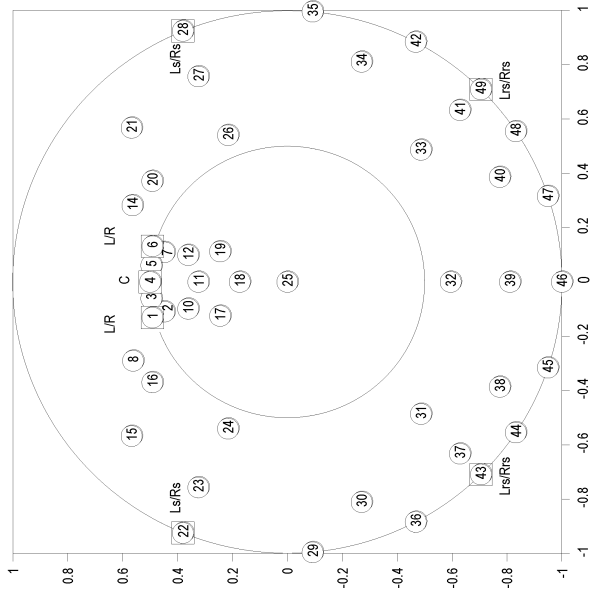


Figure 16B

【図 16 C】

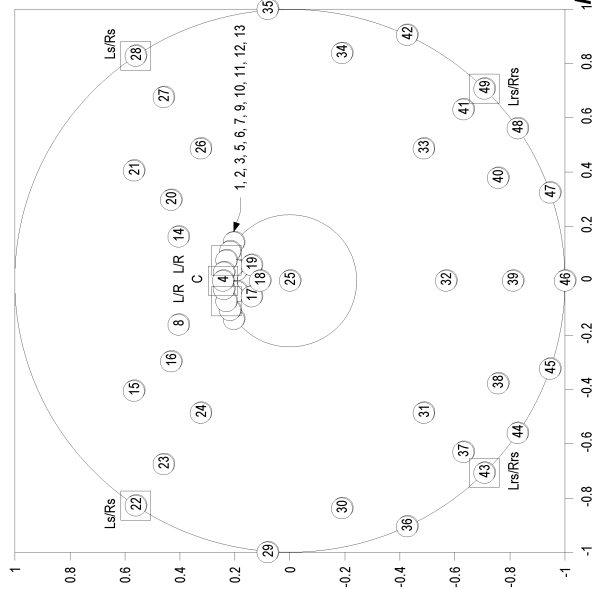


Figure 16C

【図 16 D】

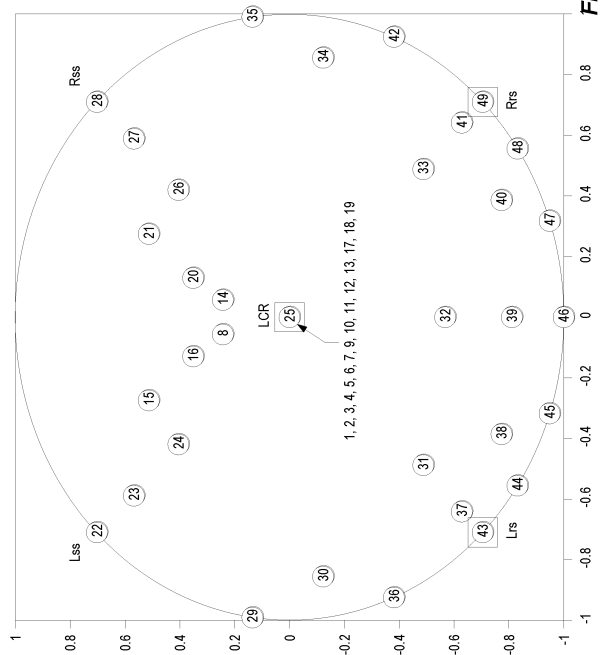


Figure 16D

【図 17】

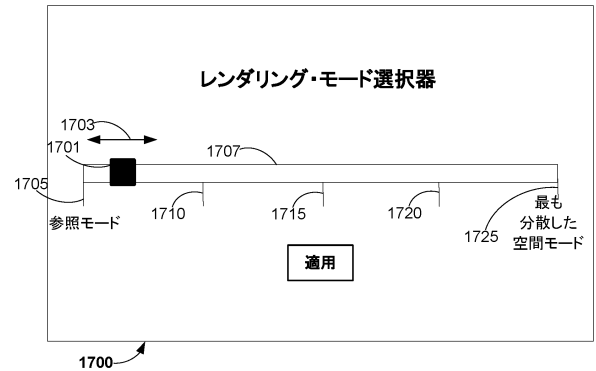


Figure 17

10

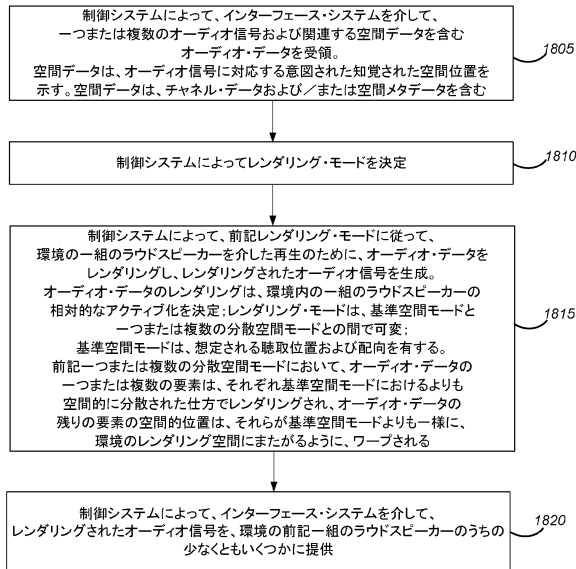
20

30

40

50

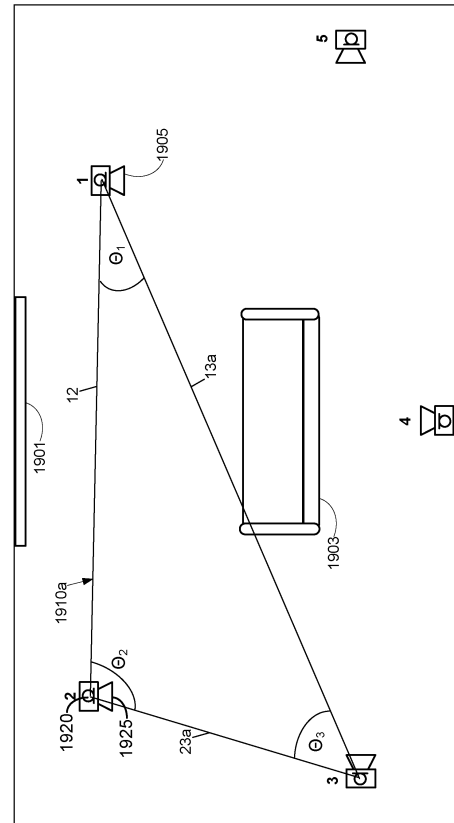
【図 18】



1800

Figure 18

【図 19】



10

20

【図 20】

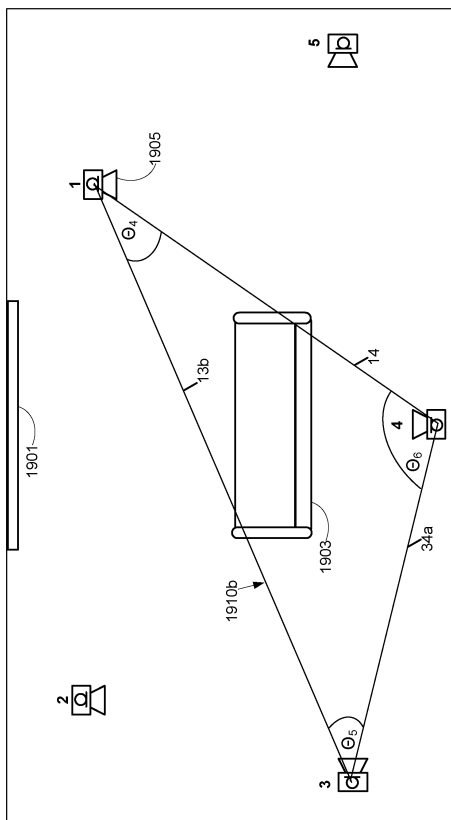


Figure 20

【図 21 A】

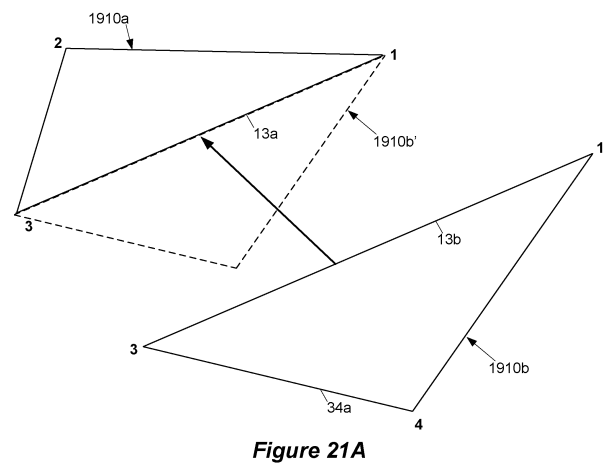


Figure 21A

30

40

50

【図 2 1 B】

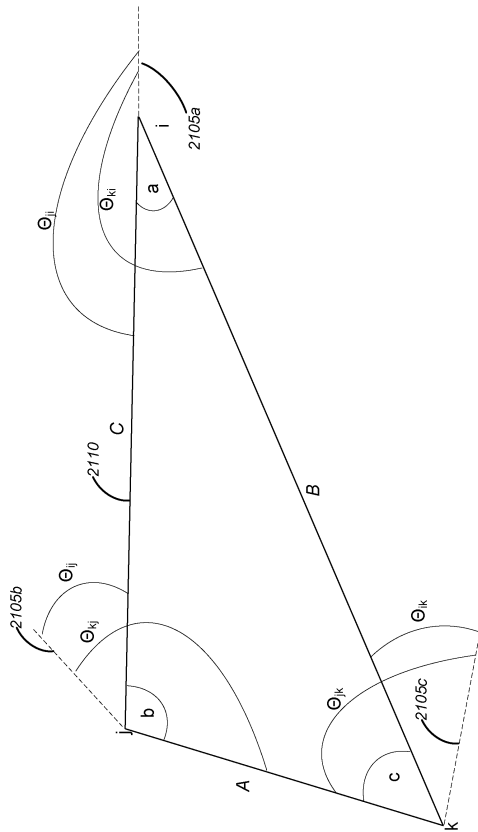
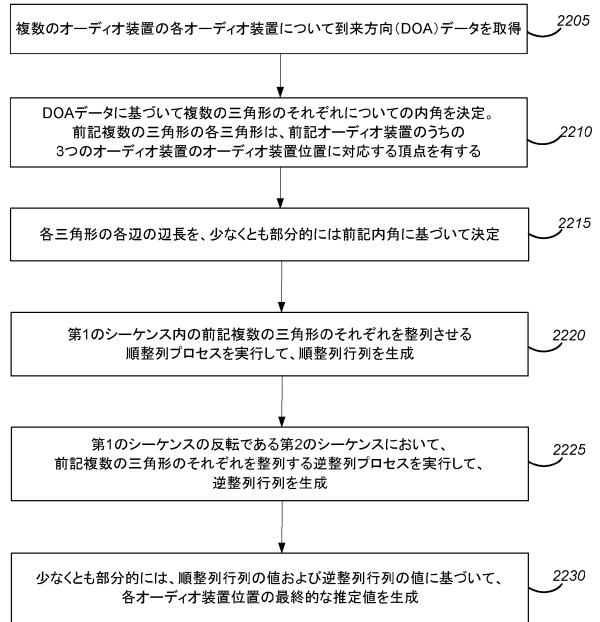


Figure 21B

【図 2 2】



2200

Figure 22

【図 2 3】

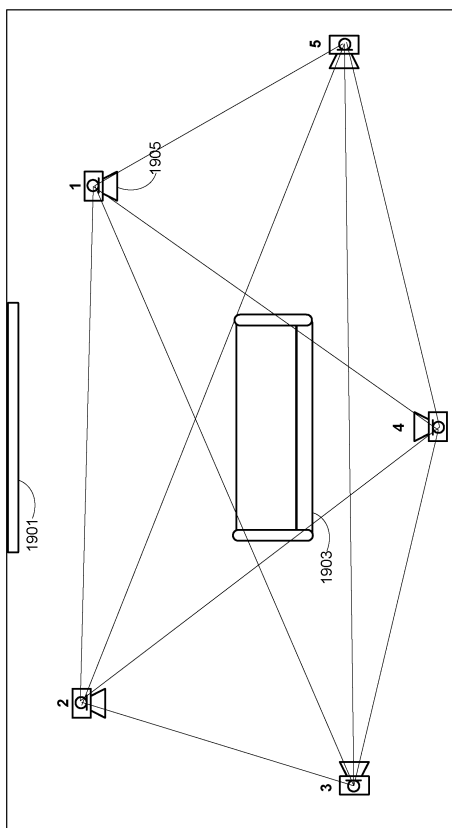


Figure 23

【図 2 4】

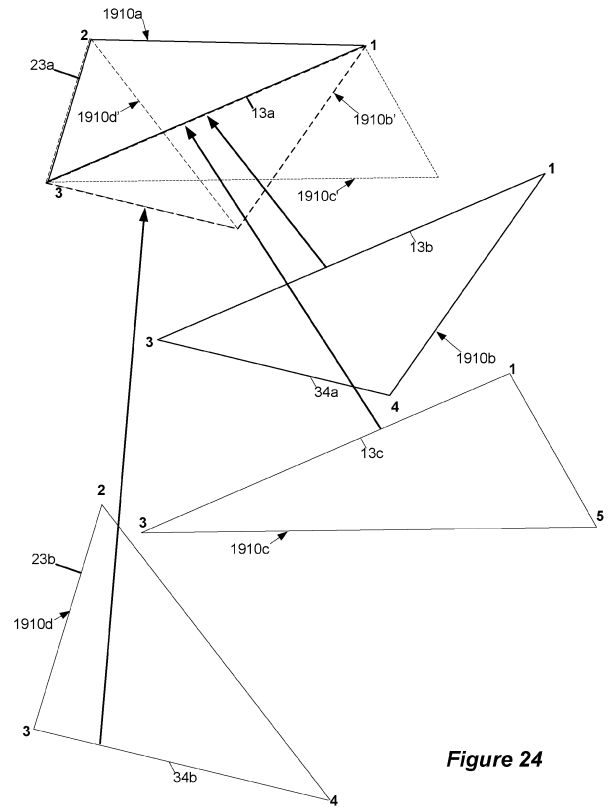


Figure 24

10

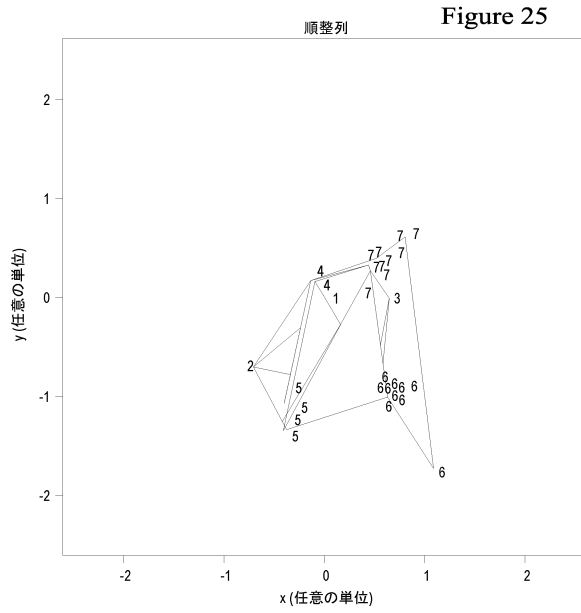
20

30

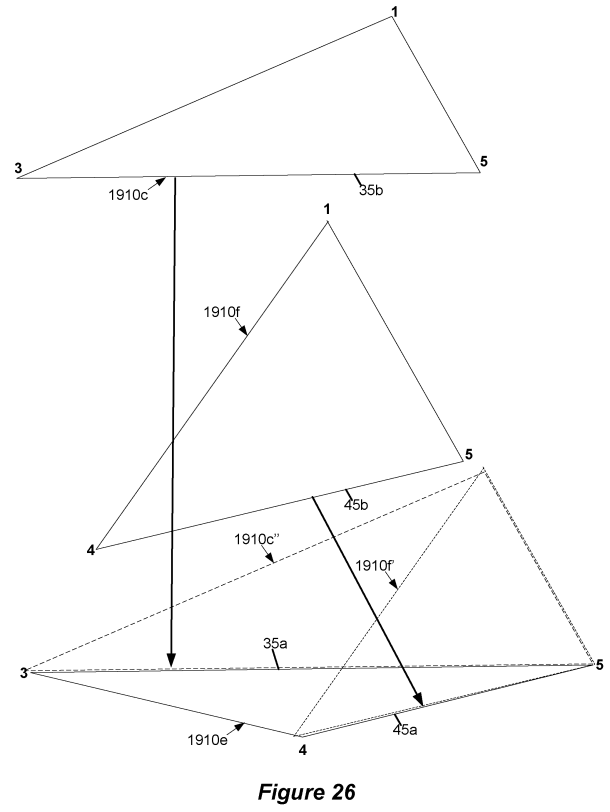
40

50

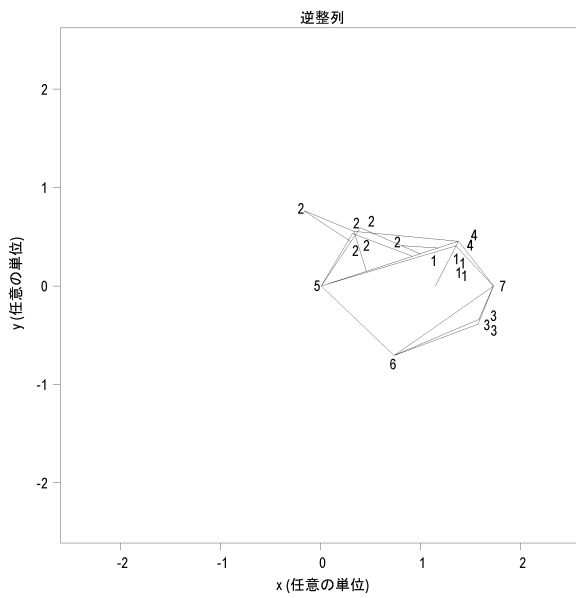
【図 2 5】



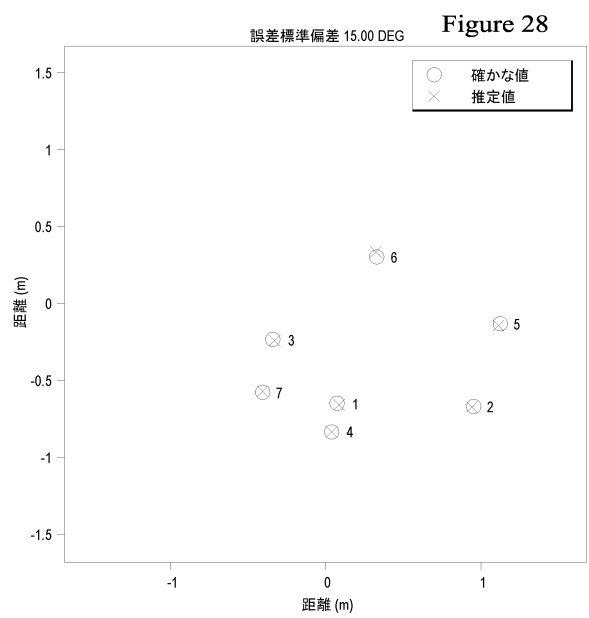
【図 2 6】



【図 2 7】



【図 2 8】



10

20

30

40

50

【圖 29】

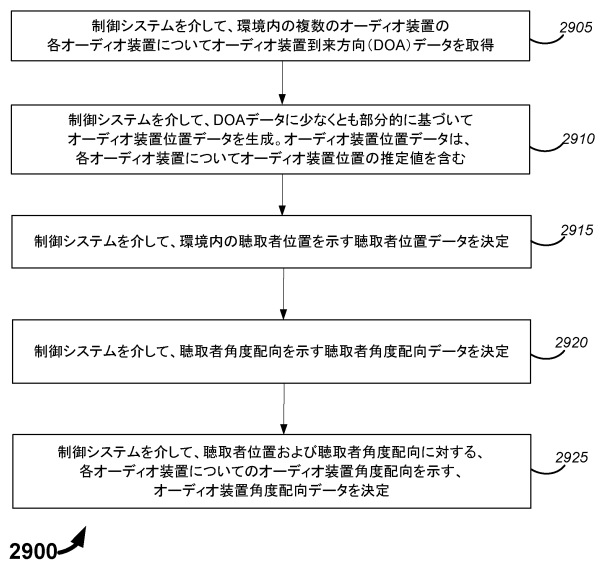
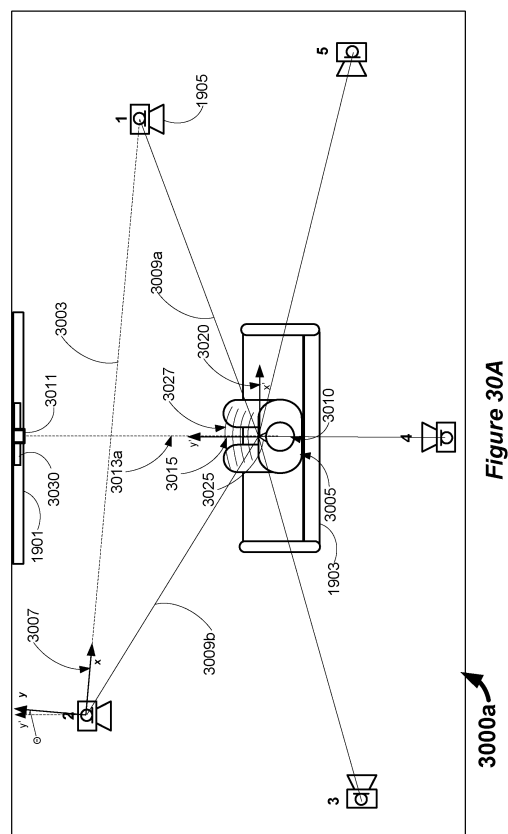
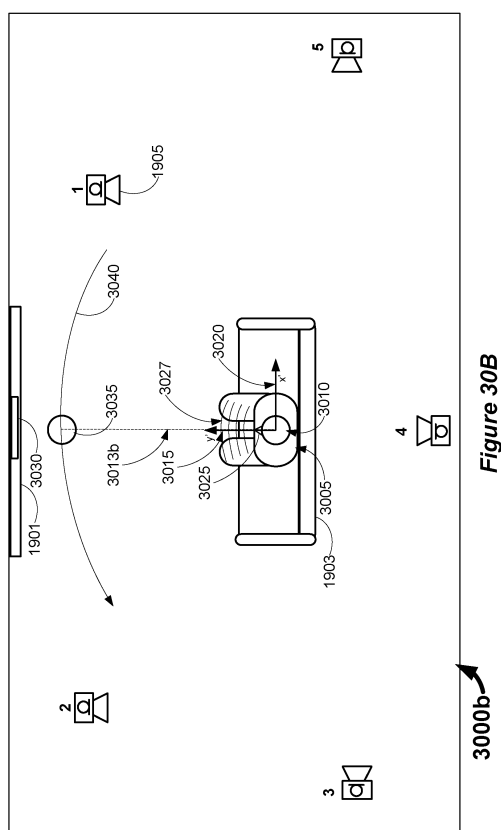


**Figure 29**

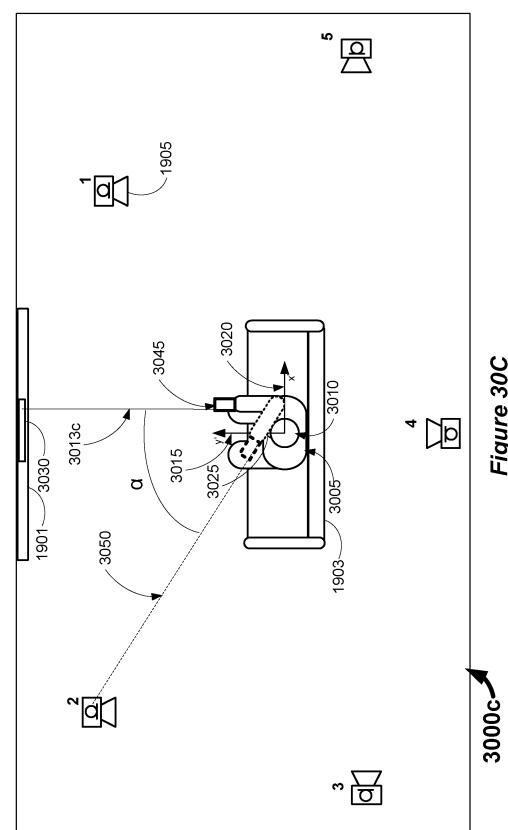
【図 30 A】



【図 30 B】



【図 30 C】



【図 30D】

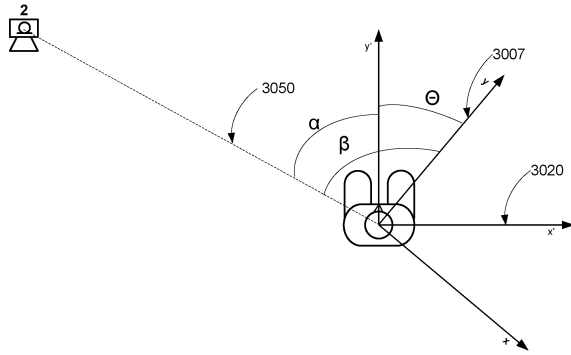


Figure 30D

【図 31】

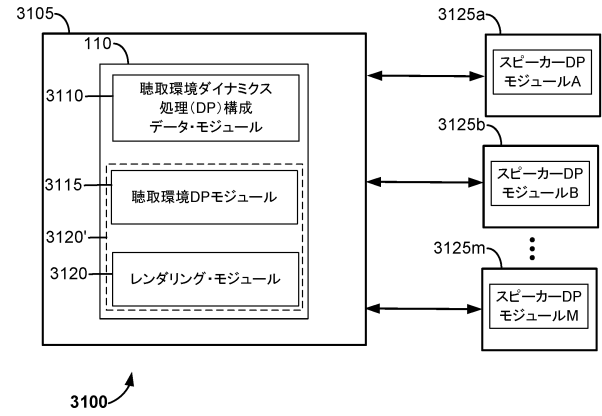


Figure 31

【図 32】

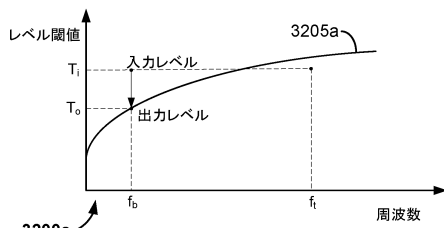


Figure 32A

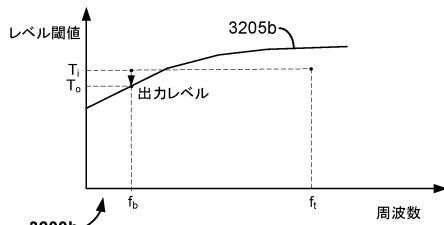


Figure 32B

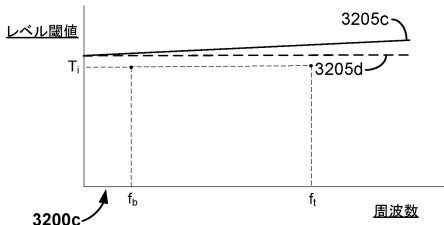


Figure 32C

【図 33】

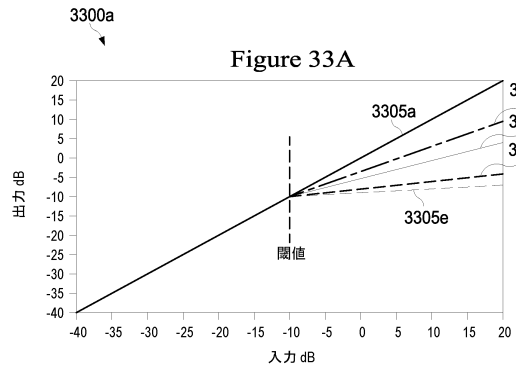


Figure 33A

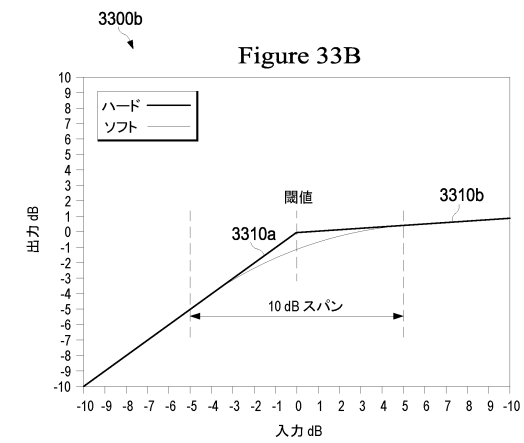


Figure 33B

10

20

30

40

50

【図 3 4】



Figure 34

【図 3 5】



Figure 35

【図 3 6】



Figure 36

【図 3 7】

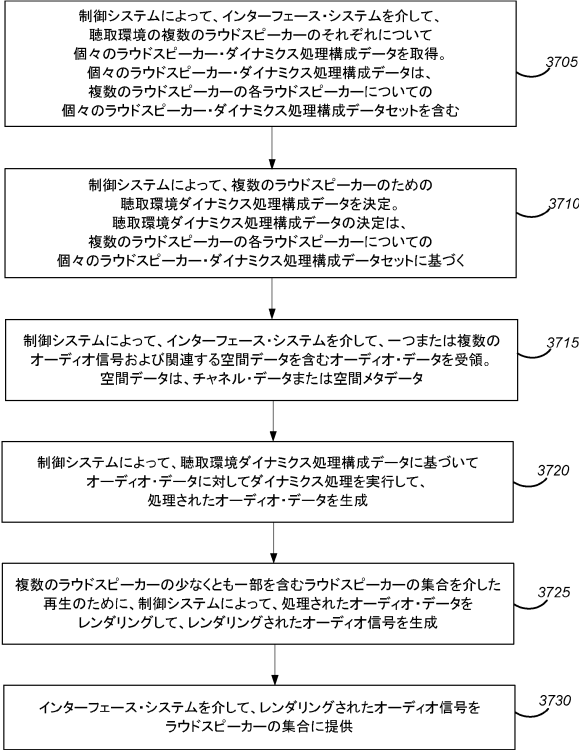


Figure 37

3700

10

20

30

40

50

【図 3 8 A】

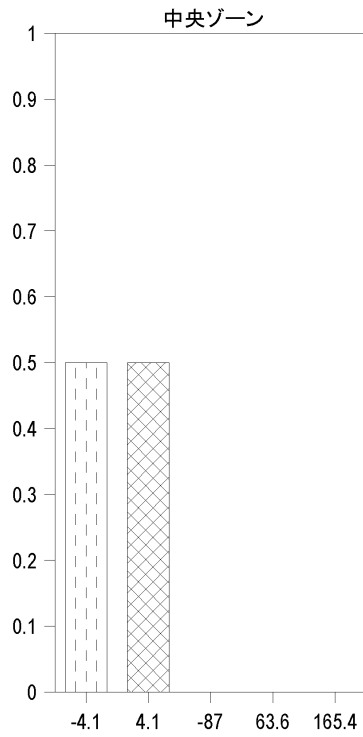


FIG. 38A

【図 3 8 B】

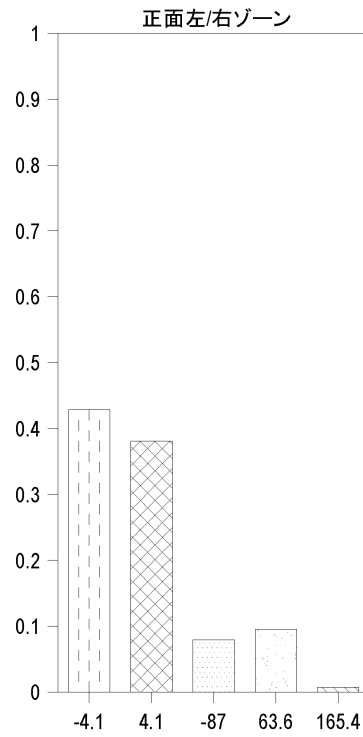


FIG. 38B

【図 3 8 C】

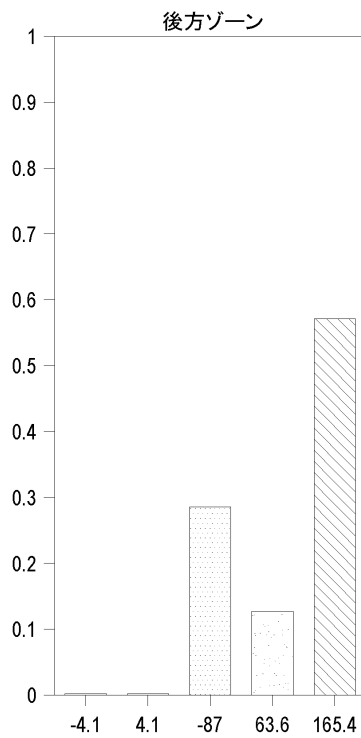


FIG. 38C

【図 3 9 A】

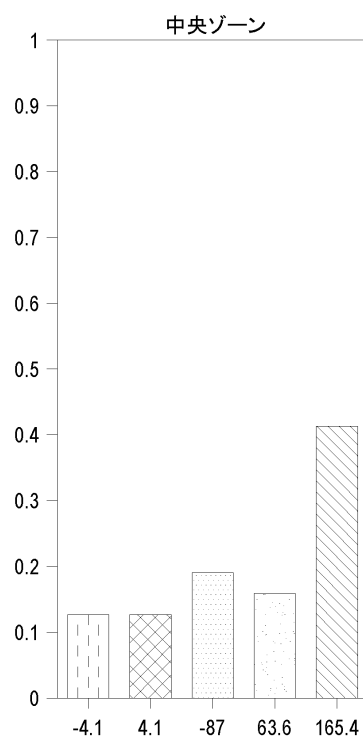


FIG. 39A

10

20

30

40

50

【図 3 9 B】

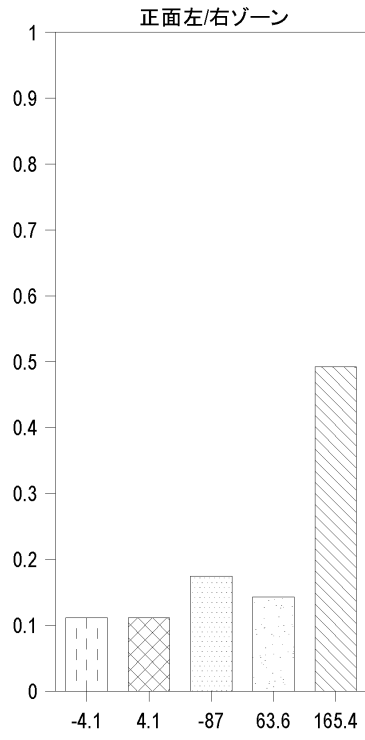


FIG. 39B

【図 3 9 C】

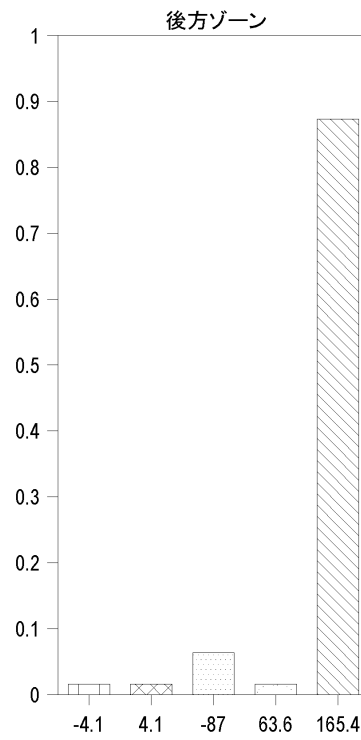


FIG. 39C

【図 4 0 A】

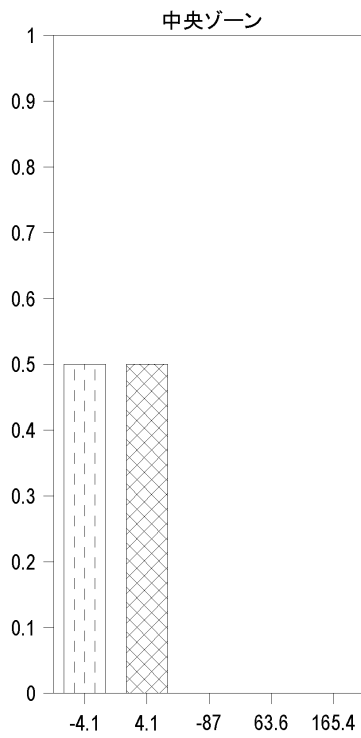


FIG. 40A

【図 4 0 B】

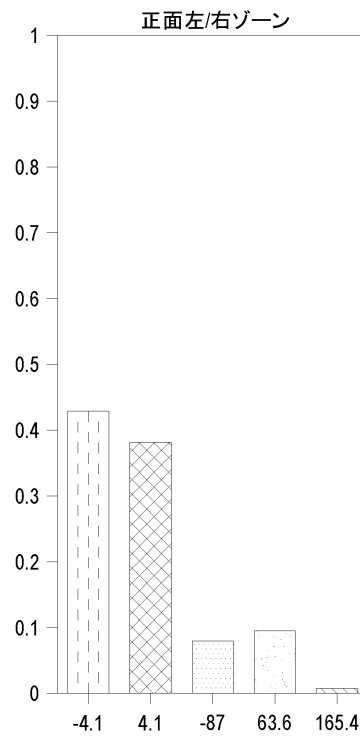


FIG. 40B

10

20

30

40

50

【図 4 0 C】

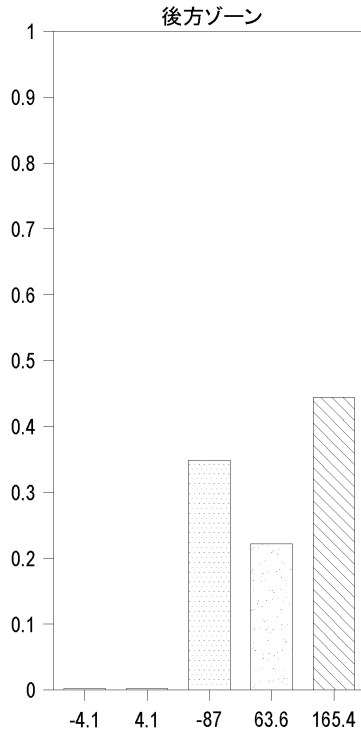


FIG. 40C

【図 4 1 A】

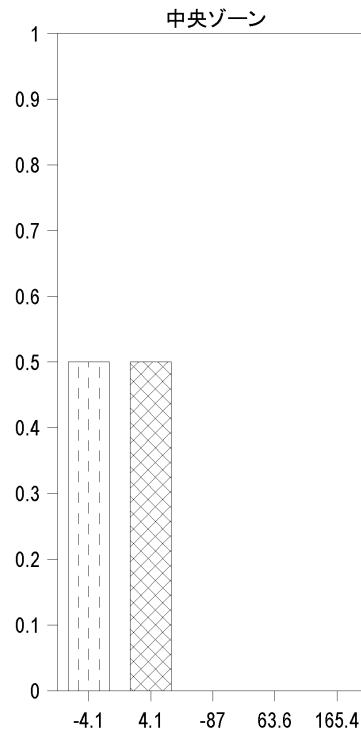


FIG. 41A

【図 4 1 B】

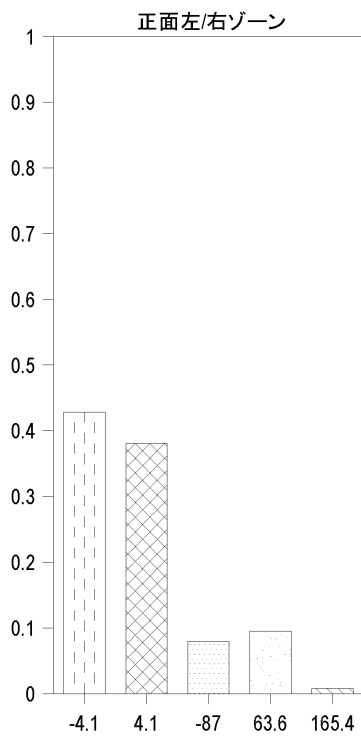


FIG. 41B

【図 4 1 C】

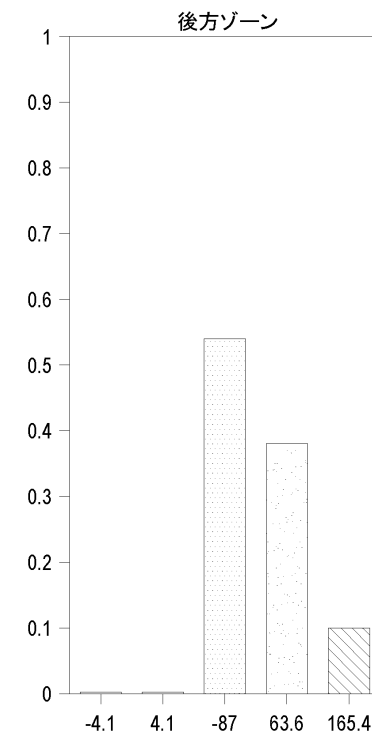


FIG. 41C

10

20

30

40

50

【図 4 2】

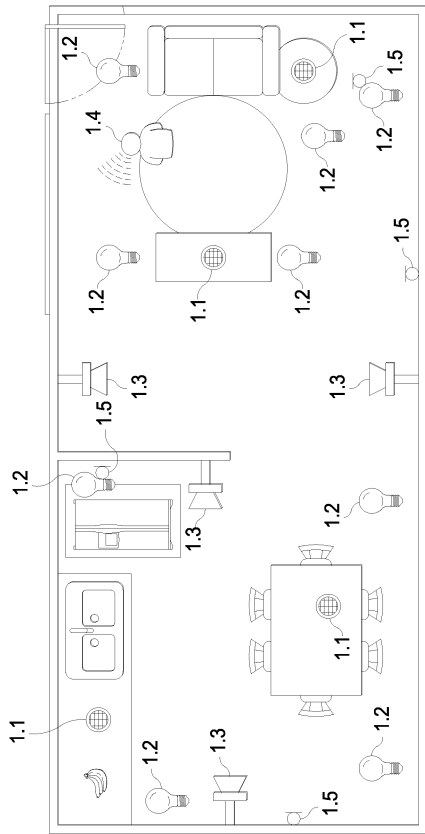


Figure 42

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 62/880,114  
(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 62/880,115  
(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 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 62/704,754  
(32)優先日 令和2年5月27日(2020.5.27)  
(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 62/705,896  
(32)優先日 令和2年7月21日(2020.7.21)  
(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 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 アルデアガ, 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
- (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 ディキンス, グレン エヌ.  
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 松崎 孝大
- (56)参考文献 特開平 0 6 - 3 1 0 9 6 2 ( J P , A )  
特開 2 0 0 6 - 1 1 5 3 6 4 ( J P , A )  
特開 2 0 1 6 - 0 0 5 2 0 7 ( J P , A )  
米国特許出願公開第 2 0 1 7 / 0 2 8 0 2 6 4 ( U S , A 1 )  
米国特許出願公開第 2 0 1 8 / 0 1 0 8 3 5 1 ( U S , A 1 )
- (58)調査した分野 (Int.Cl., D B 名)  
H 0 4 R 3 / 0 0  
G 1 0 L 1 5 / 1 0  
G 1 0 L 2 5 / 8 4  
H 0 4 S 7 / 0 0