

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7397810号  
(P7397810)

(45)発行日 令和5年12月13日(2023.12.13)

(24)登録日 令和5年12月5日(2023.12.5)

(51)国際特許分類

F I

H 0 4 S 7/00 (2006.01) H 0 4 S 7/00 3 0 0

請求項の数 6 (全29頁)

|                   |                               |          |                                  |
|-------------------|-------------------------------|----------|----------------------------------|
| (21)出願番号          | 特願2020-568524(P2020-568524)   | (73)特許権者 | 514108838                        |
| (86)(22)出願日       | 令和1年6月12日(2019.6.12)          |          | マジック リープ, インコーポレイテッド             |
| (65)公表番号          | 特表2021-527354(P2021-527354 A) |          | M a g i c L e a p, I n c .       |
| (43)公表日           | 令和3年10月11日(2021.10.11)        |          | アメリカ合衆国 フロリダ 3 3 3 2 2 ,         |
| (86)国際出願番号        | PCT/US2019/036710             |          | プランテーション, ウェスト サンライズ             |
| (87)国際公開番号        | WO2019/241345                 |          | ブルバード 7 5 0 0                    |
| (87)国際公開日         | 令和1年12月19日(2019.12.19)        |          | 7 5 0 0 W S U N R I S E B L V D  |
| 審査請求日             | 令和4年5月25日(2022.5.25)          |          | , P L A N T A T I O N, F L 3 3 3 |
| (31)優先権主張番号       | 62/684,093                    | (74)代理人  | 2 2 U S A                        |
| (32)優先日           | 平成30年6月12日(2018.6.12)         |          | 100078282                        |
| (33)優先権主張国・地域又は機関 | 米国(US)                        | (74)代理人  | 弁理士 山本 秀策                        |
|                   |                               | (74)代理人  | 100113413                        |
|                   |                               |          | 弁理士 森下 夏樹                        |
|                   |                               | (74)代理人  | 100181674                        |
|                   |                               |          | 弁理士 飯田 貴敏                        |

最終頁に続く

(54)【発明の名称】 仮想音場の効率的レンダリング

(57)【特許請求の範囲】

【請求項1】

方法であって、前記方法は、  
仮想環境のモデルを決定することであって、前記仮想環境は、直接音源と、反射音源とを備える、ことと、  
前記仮想環境の空間構成を決定することであって、前記空間構成は、少なくともユーザ場所、前記直接音源に対応する直接音源場所、前記反射音源に対応する反射音源場所、および仮想スピーカ場所を備える、ことと、  
前記ユーザ場所、前記直接音源場所、前記反射音源場所、および前記仮想スピーカ場所のうちの1つ以上のものに関連付けられた1つ以上の信号を決定することと、  
前記仮想環境内の音源の数が所定の閾値を越えるかどうかを決定することと、  
前記音源の数が前記所定の閾値を越えないという決定に従って、前記1つ以上の信号に関連付けられたエネルギー検出器を介してエネルギーレベルを検出することと、  
前記音源の数が前記所定の閾値を越えるという決定に従って、前記1つ以上の信号に関連付けられた前記エネルギー検出器をバイパスして前記1つ以上の信号を頭部関連伝達関数（H R T F）処理回路に通すことによって、前記1つ以上の信号をデコードすることと、  
前記1つ以上の信号に基づいて、オーディオ信号をレンダリングすることと、  
前記1つ以上の信号に関連付けられたエネルギーレベルを決定することと、  
前記エネルギーレベルがエネルギー閾値未満であるかどうかを決定することと、  
前記エネルギーレベルが前記エネルギー閾値未満ではないという決定に従って、前記1

つ以上の信号のH R T F処理を実施することと、

前記エネルギーレベルが前記エネルギー閾値未満であるという決定に従って、前記1つ以上の信号の前記H R T F処理の実施を見合わせることに

を含む、方法。

【請求項2】

前記1つ以上の信号をデコードすることは、

前記エネルギーレベルが前記エネルギー閾値未満ではないという決定に従って、1つ以上の処理ブロックを実施することと、

前記エネルギーレベルが前記エネルギー閾値未満であるという決定に従って、前記処理ブロックのうちの1つ以上のものを選択的にバイパスすることであって、前記処理ブロックのうちの前記1つ以上のものは、1つ以上の非アクティブ仮想スピーカに関連付けられている、ことと

を含む、請求項1に記載の方法。

【請求項3】

仮想環境のモデルを決定することは、

前記直接音源および前記反射音源から1つ以上の音信号を受信することと、

前記1つ以上の音信号を修正し、ドップラ効果をシミュレートすることと、

前記1つ以上の音信号に遅延を追加することと、

複数の仮想スピーカを横断して前記1つ以上の音信号をパンすることと

を含み、

前記1つ以上の信号をデコードすることは、

直接音源、反射音源、またはユーザの移動に関連付けられた1つ以上の仮想音を決定すること

を含む、請求項1に記載の方法。

【請求項4】

システムであって、前記システムは、

オーディオ信号をユーザに提供するように構成されるウェアラブル頭部デバイスと、

方法を実行するように構成される1つ以上のプロセッサと

を備え、

前記方法は、

仮想環境のモデルを決定することであって、前記仮想環境は、直接音源と、反射音源とを備える、ことと、

前記仮想環境の空間構成を決定することであって、前記空間構成は、少なくともユーザ場所、前記直接音源に対応する直接音源場所、前記反射音源に対応する反射音源場所、および仮想スピーカ場所を備える、ことと、

前記ユーザ場所、前記直接音源場所、前記反射音源場所、および前記仮想スピーカ場所のうちの1つ以上のものに関連付けられた1つ以上の信号を決定することと、

前記仮想環境内の音源の数が所定の閾値を越えるかどうかを決定することと、

前記音源の数が前記所定の閾値を越えないという決定に従って、前記1つ以上の信号に関連付けられたエネルギー検出器を介してエネルギーレベルを検出することと、

前記音源の数が前記所定の閾値を越えるという決定に従って、前記1つ以上の信号に関連付けられた前記エネルギー検出器をバイパスして前記1つ以上の信号を頭部関連伝達関数（H R T F）処理回路に通すことによって、前記1つ以上の信号をデコードすることと、

前記1つ以上の信号に基づいて、オーディオ信号をレンダリングすることと

を含み、

前記方法はさらに、

前記1つ以上の信号に関連付けられたエネルギーレベルを決定することと、

前記エネルギーレベルがエネルギー閾値未満であるかどうかを決定することと、

前記エネルギーレベルが前記エネルギー閾値未満ではないという決定に従って、前記1つ以上の信号のH R T F処理を実施することと、

10

20

30

40

50

前記エネルギーレベルが前記エネルギー閾値未満であるという決定に従って、前記 1 つ以上の信号の前記 H R T F 処理の実施を見合わせることに  
を含む、システム。

【請求項 5】

前記 1 つ以上の信号をデコードすることは、

前記エネルギーレベルが前記エネルギー閾値未満ではないという決定に従って、1 つ以上の処理ブロックを実施することと、

前記エネルギーレベルが前記エネルギー閾値未満であるという決定に従って、前記処理ブロックのうちの 1 つ以上のものを選択的にバイパスすることであって、前記処理ブロックのうちの前記 1 つ以上のものは、1 つ以上の非アクティブ仮想スピーカに関連付けられている、ことと

10

を含む、請求項 4 に記載のシステム。

【請求項 6】

仮想環境のモデルを決定することは、

前記直接音源および前記反射音源から 1 つ以上の音信号を受信することと、

前記 1 つ以上の音信号を修正し、ドップラ効果をシミュレートすることと、

前記 1 つ以上の音信号に遅延を追加することと、

複数の仮想スピーカを横断して前記 1 つ以上の音信号をパンすることと

を含み、

前記 1 つ以上の信号をデコードすることは、

20

直接音源、反射音源、またはユーザの移動に関連付けられた 1 つ以上の仮想音を決定すること

を含む、請求項 4 に記載のシステム。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願の相互参照)

本願は、参照することによってその全体として本明細書に組み込まれる 2018 年 6 月 12 日に出版された米国仮特許出願第 62/684,093 号の利益を主張する。

(技術分野)

30

【0002】

本開示は、空間オーディオレンダリングおよび関連付けられるシステムに一般に関する。より具体的に、本開示は、仮想スピーカベースの空間オーディオシステムの効率を高めるためのシステムおよび方法に関する。

【背景技術】

【0003】

仮想環境は、コンピューティング環境において普遍的であり、ビデオゲーム（仮想環境が、ゲーム世界を表し得る）、マップ（仮想環境が、ナビゲートされるべき地形を表し得る）、シミュレーション（仮想環境が、実環境をシミュレートし得る）、デジタルストーリーテリング（仮想キャラクターが、仮想環境内で互いに相互作用し得る）、および多くの他の用途において使用を見出している。現代のコンピュータユーザは、概ね快適に仮想環境を知覚し、それと相互作用する。しかしながら、ユーザの仮想環境の体験は、仮想環境を提示するための技術によって限定され得る。例えば、従来のディスプレイ（例えば、2D ディスプレイ画面）およびオーディオシステム（例えば、固定スピーカ）は、人を引き付け、現実的で、かつ没入型の体験を作成するように、仮想環境を実現することが可能でないこともある。

40

【0004】

仮想現実（「VR」）、拡張現実（「AR」）、複合現実（「MR」）、および関連技術（集合的に、「XR」）は、XR システムのユーザにコンピュータシステム内のデータによって表される仮想環境に対応する感覚情報を提示する能力を共有する。そのようなシ

50

システムは、仮想視覚およびオーディオキューを現実の視界および音と組み合わせることによって、独自に強調された没入感および臨場感を提供することができる。故に、音が、ユーザの実環境内で自然に、かつユーザの予期する音と一貫して発生しているように思われるようにXRシステムのユーザにデジタル音を提示することが、望ましくあり得る。概して言えば、ユーザは、仮想音がそれらが聞こえる実環境の音響特性を帯びるであろうと予期する。例えば、大きいコンサートホール内のXRシステムのユーザは、XRシステムの仮想音が大きい洞窟に似た音質を有することを予期し、逆に、小さいアパートメント内のユーザは、音が、より減衰され、近く、直接的であることを予期するであろう。加えて、ユーザは、仮想音が遅延を伴わずに提供されるであろうと予期する。

【0005】

10

他の技法の中でもとりわけ、アンビソニックスおよび非アンビソニックスが、空間オーディオを発生させるために使用され得る。多数の音源オブジェクトに関して、アンビソニックスおよび非アンビソニックスは、その設計およびアーキテクチャにより、空間オーディオをレンダリングする効率的な方法であり得る。これは、反射がモデル化される場合、特に当てはまり得る。アンビソニックスおよび非アンビソニックスマルチチャネルベースの空間オーディオシステムは、いくつかのステップを通してオーディオ信号をレンダリングし得る。例示的ステップは、源毎のエンコードステップ、固定オーバーヘッド音場デコードステップ、および／または固定スピーカ仮想化ステップを含むことができる。1つ以上のハードウェアコンポーネントが、ステップを実施し得る。

【0006】

20

オーディオ信号をレンダリングするための第1の方法において、各音源は、それ自身の有限インパルス応答(FIR)フィルタの対を有することができる。そのようなシステムにおいて、音の知覚される位置は、FIRフィルタのフィルタ係数を変化させることによって変化させられる。いくつかの実施形態において、各音は、複数(例えば、2対)のFIRフィルタを使用し得る。各対が、2つのフィルタを使用し得る(すなわち、4つのFIRフィルタ)。音が仮想環境の周囲を移動するとき、FIRフィルタは、クロスフェードされることができる。いくつかの実施形態において、4つのFIRフィルタが、各音のために使用され得る。

【0007】

30

オーディオ信号をレンダリングするための第2の方法において、仮想スピーカパンニングが、固定数の仮想スピーカを使用して実装され得る。各音源は、固定仮想スピーカにわたってパンされ得る。いくつかの実施形態において、複数(例えば、2つ)のFIRフィルタが、各仮想スピーカのために使用され得る。仮想スピーカパンニングは、ある用途に関して効率的であり得、ごくわずかな算出リソースを使用し得る。

【0008】

いくつかの実施形態において、ある方法は、同時に再生される音の数に応じて、他の方法と比較して高い効率を有し得る。例えば、30個の音が、同時に再生され得る。4つのFIRフィルタが各音源のために使用される場合、120個のFIRフィルタ(30個の音源×音源あたり4つのFIRフィルタ=120個のFIRフィルタ)が、第1の方法のために要求され得る。2つのFIRフィルタが各仮想スピーカのために使用される場合、32個のみのFIRフィルタが、第2の方法のために要求され得る(16個の仮想スピーカ×仮想スピーカあたり2つのFIRフィルタ=32個のFIRフィルタ)。

40

【0009】

別の例として、1つのみの音が、再生され得る。第1の方法は、4つのみのFIRフィルタ(1つの音源×音源あたり4つのFIRフィルタ=4つのFIRフィルタ)を要求し得る一方、第2の方法は、32個のFIRフィルタ(16個の仮想スピーカ×仮想スピーカあたり2つのFIRフィルタ=32個のFIRフィルタ)を要求し得る。

【0010】

上記の例を通して例証されるように、第1の方法は、少数の音に関して有益であり得、第2の方法は、多数の音に関して有益であり得る。故に、所与の時間における音源の数に

50

基づいて、効率を高めるオーディオシステムおよび方法が、所望され得る。

【発明の概要】

【課題を解決するための手段】

【0011】

オーディオ信号をレンダリングするオーディオシステムおよび方法が、開示され、システムは、修正された仮想スピーカパンニングを使用する。オーディオシステムは、固定数Fの仮想スピーカを含み得、修正された仮想スピーカパンニングは、固定仮想スピーカのサブセットPを動的に選択および使用し得る。各音源は、仮想スピーカのサブセットPにわたってパンされ得る。いくつかの実施形態において、複数（例えば、2つ）のFIRフィルタが、サブセットPの各仮想スピーカのために使用され得る。仮想スピーカのサブセットPは、音源への近接性等の1つ以上の因子に基づいて選択され得る。仮想スピーカのサブセットPは、アクティブスピーカと称され得る。

10

【0012】

修正された仮想スピーカパンニング方法は、例として上で開示される第1および第2の方法と比較されることができる。3つの音が、同時に再生され、オーディオシステムが、16個の固定仮想スピーカを有する場合、第1の方法は、12個のFIRフィルタ（3つの音源×音源あたり4つのFIRフィルタ＝12個のFIRフィルタ）を要求し得、第2の方法は、32個のFIRフィルタ（16個の仮想スピーカ×仮想スピーカあたり2つのFIRフィルタ＝32個のFIRフィルタ）を要求し得る。一方、修正された仮想スピーカパンニング方法は、サブセットPの一部としてアクティブ仮想スピーカであるように3つの仮想スピーカを動的に選択し得る。修正された仮想スピーカパンニング方法は、6つのFIRフィルタ、すなわち、アクティブ各仮想スピーカのために2つのFIRフィルタ（3つの仮想スピーカ×2つのFIRフィルタ＝6つのFIRフィルタ）を要求し得る。本発明は、例えば、以下を提供する。

20

（項目1）

オーディオ信号を空間的にレンダリングする方法であって、前記方法は、  
空間モデラを使用して、仮想環境をモデル化することと、  
空間エンコーダを使用して、複数の仮想スピーカにわたって前記空間モデラからの信号を分配することと、  
内部空間表現を使用して、前記仮想環境の空間構成を表すことと、  
デコーダ／バーチャライザを使用して、前記内部空間表現からの信号をデコードすることと、  
デコーダ／バーチャライザを使用して、前記デコードされた信号の中に仮想音を導入することと、  
前記デコーダ／バーチャライザ内の非アクティブ仮想スピーカに関連付けられた1つ以上の処理ブロックを選択的にバイパスすることと、  
前記デコーダ／バーチャライザからの信号を組み合わせることと、  
前記組み合わせられた信号を前記オーディオ信号として出力することと  
を含む、方法。

30

（項目2）

音場デコーダからの前記信号に関連付けられたエネルギーレベルを決定することと、  
前記検出されたエネルギーレベルの各々がエネルギー閾値より小さいかどうかを決定することと  
をさらに含み、  
前記1つ以上の処理ブロックの前記選択的バイパスは、前記仮想スピーカのうちの少なくとも1つの前記検出されたエネルギーレベルが前記エネルギー閾値より小さいという決定に従って、前記音場デコーダからの前記対応する信号の頭部関連伝達関数（HRTF）処理をバイパスすることを含み、  
前記音場デコーダは、前記デコーダ／バーチャライザ内に含まれる、項目1に記載の方法。

40

50

(項目 3)

前記仮想スピーカのうちの少なくとも 1 つの前記検出されたエネルギーレベルが前記エネルギー閾値より小さくないという決定に従って、前記音場デコーダからの前記対応する信号の H R T F 処理を実施することをさらに含む、項目 2 に記載の方法。

(項目 4)

音源の数が所定の音源閾値以上であるかどうかを決定することをさらに含み、  
前記 1 つ以上の処理ブロックの前記選択的バイパスは、前記音源の数が前記所定の音源閾値以上であるとき、複数の検出器をバイパスし、音場デコーダからの信号を複数の H R T F ブロックに直接通すことを含み、  
前記複数の検出器および前記複数の H R T F ブロックは、前記デコーダ／バーチャライザ内に含まれる、項目 1 に記載の方法。

10

(項目 5)

前記音源の数が前記所定の音源閾値以上でないという決定に従って、前記音場デコーダからの信号を前記複数の検出器に直接通すことをさらに含む、項目 4 に記載の方法。

(項目 6)

各音源の場所を決定することと、  
前記複数の仮想スピーカのうちのどれが前記それぞれの音源に近接して位置しているかを決定することと  
をさらに含む、項目 1 に記載の方法。

20

(項目 7)

前記複数の仮想スピーカのうちのどれが前記それぞれの音源に近接して位置しているかの前記決定は、全ビデオフレームにおいて実施される、項目 6 に記載の方法。

(項目 8)

前記デコーダ／バーチャライザ内の前記 1 つ以上の処理ブロックの前記選択的バイパスは、前記デコーダ／バーチャライザ内の前記それぞれの音源に近接して位置していない少なくとも 1 つのスピーカに関連付けられた前記 1 つ以上の処理ブロックの全てをバイパスすることを含む、項目 6 に記載の方法。

(項目 9)

回転／平行移動表現を使用して、前記オーディオ信号に関連付けられた移動の表現を導入することと、  
前記回転／平行移動表現からの信号の振幅が所定の振幅閾値以上であるかどうかを決定することと  
をさらに含み、

30

前記デコーダ／バーチャライザ内の前記 1 つ以上の処理ブロックの前記選択的バイパスは、前記回転／平行移動表現からの前記信号の振幅が前記所定の振幅閾値以上でないとき、音場デコーダおよび複数の H R T F ブロックをバイパスすることを含む、  
前記音場デコーダおよび前記複数の H R T F ブロックは、前記デコーダ／バーチャライザ内に含まれる、項目 1 に記載の方法。

(項目 10)

前記回転／平行移動表現からの前記信号の振幅が前記所定の振幅閾値以上であるという決定に従って、  
前記回転／平行移動表現からの信号をデコードすることと、  
頭部関連伝達関数 (H R T F) を決定し、それを前記デコードされた信号に適用することと  
をさらに含む、項目 9 に記載の方法。

40

(項目 11)

前記複数の仮想スピーカは、第 1 の時間において、前記非アクティブ仮想スピーカとアクティブ仮想スピーカとを含み、前記第 1 の時間における前記アクティブ仮想スピーカのうちの少なくとも 1 つは、信号が処理されている間の第 2 の時間において、非アクティブとして指定される、項目 1 に記載の方法。

50

(項目 1 2)

システムであって、前記システムは、  
オーディオ信号をユーザに提供するように構成されたウェアラブル頭部デバイスと、  
前記オーディオ信号を空間的にレンダリングするように構成された回路と  
を備え、  
前記回路は、  
仮想環境をモデル化するように構成された空間モデラと、  
複数の仮想スピーカにわたって前記空間モデラからの信号を分配するように構成された  
空間エンコーダと、  
前記仮想環境の空間構成を表すように構成された内部空間表現と、  
前記内部空間表現からの信号をデコードし、前記デコードされた信号の中に仮想音を導  
入するように構成されたデコーダ／バーチャライザと  
を含み、  
前記デコーダ／バーチャライザは、  
前記オーディオ信号に関連付けられた移動の表現を導入するように構成された回転／平  
行移動表現と、  
前記回転／平行移動表現からの信号をデコードするように構成可能な音場デコーダと、  
複数の頭部関連伝達関数（H R T F）ブロックであって、前記複数の H R T F ブロック  
は、その入力信号に対応する H R T F を決定し、前記その入力信号に前記対応する H R T  
F を適用するように構成されている、複数の H R T F ブロックと、  
前記複数の H R T F ブロックからの信号を組み合わせ、前記オーディオ信号を出力する  
ように構成された複数のコンバイナと  
を含む、システム。

10

20

(項目 1 3)

前記音場デコーダから信号を受信し、前記音場デコーダからの前記信号に関連付けられ  
たエネルギーレベルを決定するように構成された複数の検出器と、  
前記決定されたエネルギーレベルがエネルギー閾値より小さくないとき、前記信号を前  
記音場デコーダから前記複数の H R T F ブロックに通すように構成された複数の第 1 のス  
イッチと  
をさらに備えている、項目 1 2 に記載のシステム。

30

(項目 1 4)

第 2 のスイッチをさらに備え、前記第 2 のスイッチは、  
前記音場デコーダから前記信号を受信することと、  
前記音場デコーダからの前記信号を直接前記複数の検出器または前記複数の H R T F ブ  
ロックに選択的に通すことと  
を行うように構成されている、項目 1 3 に記載のシステム。

(項目 1 5)

音場デコード決定をさらに備え、前記音場デコード決定は、  
前記回転／平行移動表現からの信号の振幅が所定の振幅閾値より大きいかどうかを決定  
することと、  
前記回転／平行移動表現からの前記信号の振幅が前記所定の振幅閾値より大きいという  
決定に従って、前記回転／平行移動表現からの前記信号を前記音場デコーダに通すことと  
を行うように構成されている、項目 1 2 に記載のシステム。

40

【図面の簡単な説明】【0 0 1 3】

【図 1】図 1 は、いくつかの実施形態による、例示的ウェアラブルシステムを図示する。

【0 0 1 4】

【図 2】図 2 は、いくつかの実施形態による、例示的ウェアラブルシステムと併用され得  
る例示的ハンドヘルドコントローラを図示する。

【0 0 1 5】

50

【図 3】図 3 は、いくつかの実施形態による、例示的ウェアラブルシステムと併用され得る例示的補助ユニットを図示する。

【0016】

【図 4】図 4 は、いくつかの実施形態による、例示的ウェアラブルシステムに関する例示的機能ブロック図を図示する。

【0017】

【図 5 A】図 5 A は、いくつかの実施形態による、例示的空間オーディオシステムのブロック図を図示する。

【0018】

【図 5 B】図 5 B は、いくつかの実施形態による、図 5 A のシステムを動作させるための例示的方法のフローを図示する。

10

【0019】

【図 5 C】図 5 C は、いくつかの実施形態による、例示的デコーダ／バーチャライザを動作させるための例示的方法のフローを図示する。

【0020】

【図 6】図 6 は、いくつかの実施形態による、音源およびスピーカの例示的構成を図示する。

【0021】

【図 7 A】図 7 A は、いくつかの実施形態による、複数の検出器を含む例示的デコーダ／バーチャライザのブロック図を図示する。

20

【0022】

【図 7 B】図 7 B は、いくつかの実施形態による、図 7 A のデコーダ／バーチャライザを動作させるための例示的方法のフローを図示する。

【0023】

【図 8 A】図 8 A は、いくつかの実施形態による、例示的デコーダ／バーチャライザのブロック図を図示する。

【0024】

【図 8 B】図 8 B は、いくつかの実施形態による、図 8 A のデコーダ／バーチャライザを動作させるための例示的方法のフローを図示する。

【0025】

30

【図 9】図 9 は、いくつかの実施形態による、音源およびスピーカの例示的構成を図示する。

【0026】

【図 10 A】図 10 A は、いくつかの実施形態による、アクティブスピーカを含むシステムにおいて使用される例示的デコーダ／バーチャライザのブロック図を図示する。

【0027】

【図 10 B】図 10 B は、いくつかの実施形態による、図 10 A のデコーダ／バーチャライザを動作させるための例示的方法のフローを図示する。

【発明を実施するための形態】

【0028】

40

以下の例の説明において、本明細書の一部を形成し、例証として、実践され得る具体的例が示される付随の図面が、参照される。他の例も、使用され得、構造変更が、開示される例の範囲から逸脱することなく、行われ得ることを理解されたい。

【0029】

(例示的ウェアラブルシステム)

【0030】

図 1 は、ユーザの頭部上に装着されるように構成された例示的ウェアラブル頭部デバイス 100 を図示する。ウェアラブル頭部デバイス 100 は、頭部デバイス（例えば、ウェアラブル頭部デバイス 100）、ハンドヘルドコントローラ（例えば、下で説明されるハンドヘルドコントローラ 200）、および／または補助ユニット（例えば、下で説明され

50

る補助ユニット300)等の1つ以上のコンポーネントを備えているより広範なウェアラブルシステムの一部であり得る。いくつかの例において、ウェアラブル頭部デバイス100は、仮想現実、拡張現実、または複合現実システムまたは用途のために使用されることができる。ウェアラブル頭部デバイス100は、ディスプレイ110Aおよび110B(左および右透過性ディスプレイと、直交瞳拡大(OPE)格子セット112A/112Bおよび射出瞳拡大(EPE)格子セット114A/114B等、ディスプレイからユーザの眼に光を結合するための関連付けられるコンポーネントとを備え得る)等の1つ以上のディスプレイと、スピーカ120Aおよび120B(それぞれ、つるアーム122Aおよび122B上に搭載され、ユーザの左および右耳に隣接して位置付けられ得る)等の左および右音響構造と、赤外線センサ、加速度計、GPSユニット、慣性測定ユニット(IMU)(例えば、IMU126)、音響センサ(例えば、マイクロホン150)等の1つ以上のセンサと、直交コイル電磁受信機(例えば、左つるアーム122Aに搭載されるように示される受信機127)と、ユーザから離れるように向けられた左および右カメラ(例えば、深度(飛行時間)カメラ130Aおよび130B)と、ユーザに向かって向けられた左および右眼カメラ(例えば、ユーザの眼移動を検出するため)(例えば、眼カメラ128および128B)とを備えていることができる。しかしながら、ウェアラブル頭部デバイス100は、本発明の範囲から逸脱することなく、任意の好適なディスプレイ技術およびセンサまたは他のコンポーネントの任意の好適な数、タイプ、または組み合わせを組み込むことができる。いくつかの例において、ウェアラブル頭部デバイス100は、ユーザの音声によって発生させられるオーディオ信号を検出するように構成されている1つ以上のマイクロホン150を備え得、そのようなマイクロホンは、ユーザの口に隣接してウェアラブル頭部デバイス内に位置付けられ得る。いくつかの例において、ウェアラブル頭部デバイス100は、他のウェアラブルシステムを含む他のデバイスおよびシステムと通信するために、ネットワークング特徴(例えば、Wi-Fi能力)を組み込み得る。ウェアラブル頭部デバイス100は、バッテリー、プロセッサ、メモリ、記憶ユニット、または種々の入力デバイス(例えば、ボタン、タッチパッド)等のコンポーネントをさらに含み得るか、または、1つ以上のそのようなコンポーネントを備えているハンドヘルドコントローラ(例えば、ハンドヘルドコントローラ200)または補助ユニット(例えば、補助ユニット300)に結合され得る。いくつかの例において、センサは、ユーザの環境に対する頭部搭載型ユニットの座標の組を出力するように構成され得、入力をプロセッサに提供し、同時位置特定およびマッピング(SLAM)プロシージャおよび/またはビジュアルオドメトリアルゴリズムを実施し得る。いくつかの例において、ウェアラブル頭部デバイス100は、下でさらに説明されるように、ハンドヘルドコントローラ200および/または補助ユニット300に結合され得る。

#### 【0031】

図2は、例示的ウェアラブルシステムの例示的モバイルハンドヘルドコントローラコンポーネント200を図示する。いくつかの例において、ハンドヘルドコントローラ200は、ウェアラブルヘッドデバイス100および/または下で説明される補助ユニット300と有線または無線通信し得る。いくつかの例において、ハンドヘルドコントローラ200は、ユーザによって保持されるべきハンドル部分220と、上面210に沿って配置される1つ以上のボタン240とを含む。いくつかの例において、ハンドヘルドコントローラ200は、光学追跡標的として使用するために構成され得、例えば、ウェアラブル頭部デバイス100のセンサ(例えば、カメラまたは他の光学センサ)は、ハンドヘルドコントローラ200の位置および/または向きを検出するように構成されることができ、それは、転じて、ハンドヘルドコントローラ200を保持するユーザの手の位置および/または向きを示し得る。いくつかの例において、ハンドヘルドコントローラ200は、プロセッサ、メモリ、記憶ユニット、ディスプレイ、または上で説明されるもの等の1つ以上の入力デバイスを含み得る。いくつかの例において、ハンドヘルドコントローラ200は、1つ以上のセンサ(例えば、ウェアラブル頭部デバイス100に関して上で説明されるセンサまたは追跡コンポーネントのうちのいずれか)を含む。いくつかの例において、セン

10

20

30

40

50

サは、ウェアラブル頭部デバイス100に対する、またはウェアラブルシステムの別のコンポーネントに対するハンドヘルドコントローラ200の位置または向きを検出することができる。いくつかの例において、センサは、ハンドヘルドコントローラ200のハンドル部分220内に位置付けられ得、および／またはハンドヘルドコントローラに機械的に結合され得る。ハンドヘルドコントローラ200は、例えば、ボタン240の押された状態、またはハンドヘルドコントローラ200の位置、向き、および／または運動（例えば、IMUを介して）に対応する1つ以上の出力信号を提供するように構成されることができる。そのような出力信号は、ウェアラブル頭部デバイス100のプロセッサへの入力、補助ユニット300への入力、またはウェアラブルシステムの別のコンポーネントへの入力として使用され得る。いくつかの例において、ハンドヘルドコントローラ200は、音（例えば、ユーザの発話、環境音）を検出し、ある場合、検出された音に対応する信号をプロセッサ（例えば、ウェアラブル頭部デバイス100のプロセッサ）に提供するために、1つ以上のマイクロホンを含むことができる。

10

#### 【0032】

図3は、例示的ウェアラブルシステムの例示的補助ユニット300を図示する。いくつかの例において、補助ユニット300は、ウェアラブル頭部デバイス100および／またはハンドヘルドコントローラ200と有線または無線通信し得る。補助ユニット300は、ウェアラブル頭部デバイス100および／またはハンドヘルドコントローラ200（ディスプレイ、センサ、音響構造、プロセッサ、マイクロホン、および／またはウェアラブル頭部デバイス100またはハンドヘルドコントローラ200の他のコンポーネントを含む）等のウェアラブルシステムの1つ以上のコンポーネントを動作させるためのエネルギーを提供するために、バッテリーを含むことができる。いくつかの例において、補助ユニット300は、プロセッサ、メモリ、記憶ユニット、ディスプレイ、1つ以上の入力デバイス、および／または上で説明されるもの等の1つ以上のセンサを含み得る。いくつかの例において、補助ユニット300は、補助ユニットをユーザに取り付けるためのクリップ310（例えば、ユーザによって装着されるベルト）を含む。ウェアラブルシステムの1つ以上のコンポーネントを格納するために補助ユニット300を使用する利点は、そのように行うことが、大きいまたは重いコンポーネントが、（例えば、ウェアラブル頭部デバイス100内に格納される場合）ユーザの頭部に搭載されるのではなく、または（例えば、ハンドヘルドコントローラ200内に格納される場合）ユーザの手によって持ち運ばれるのではなく、大きく重い物体を支持するために比較的に良好に適しているユーザの腰部、胸部、または背部上で持ち運ばれることを可能にし得ることである。これは、バッテリー等の比較的に重いまたは嵩張るコンポーネントに関して特に有利であり得る。

20

30

#### 【0033】

図4は、上で説明される、例示的ウェアラブル頭部デバイス100と、ハンドヘルドコントローラ200と、補助ユニット300とを含み得る等、例示的ウェアラブルシステム400に対応し得る例示的機能ブロック図を示す。いくつかの例において、ウェアラブルシステム400は、仮想現実、拡張現実、または複合現実用途のために使用され得る。図4に示されるように、ウェアラブルシステム400は、ここでは「トーテム」と称される（および上で説明されるハンドヘルドコントローラ200に対応し得る）例示的ハンドヘルドコントローラ400Bを含むことができ、ハンドヘルドコントローラ400Bは、トーテム／ヘッドギヤ6自由度（6DOF）トーテムサブシステム404Aを含むことができる。ウェアラブルシステム400は、（上で説明されるウェアラブルヘッドギヤデバイス100に対応し得る）例示的ウェアラブル頭部デバイス400Aも含むことができ、ウェアラブル頭部デバイス400Aは、トーテム／ヘッドギヤ6DOFヘッドギヤサブシステム404Bを含む。例において、6DOFトーテムサブシステム404Aおよび6DOFヘッドギヤサブシステム404Bは、ウェアラブル頭部デバイス400Aに対するハンドヘルドコントローラ400Bの6つの座標（例えば、3つの平行移動方向におけるオフセットおよび3つの軸に沿った回転）を決定するために協働する。6自由度は、ウェアラブル頭部デバイス400Aの座標系に対して表され得る。3つの平行移動オフセットは、

40

50

そのような座標系内におけるX、Y、およびZオフセット、平行移動行列、またはある他の表現として表され得る。回転自由度は、ヨー、ピッチ、およびロール回転の列、ベクトル、回転行列、四元数、またはある他の表現として表され得る。いくつかの例において、ウェアラブル頭部デバイス400A内に含まれる1つ以上の深度カメラ444（および／または1つ以上の非深度カメラ）および／または1つ以上の光学標的（例えば、上で説明されるようなハンドヘルドコントローラ200のボタン240またはハンドヘルドコントローラ内に含まれる専用光学標的）は、6DOF追跡のために使用されることができる。いくつかの例において、ハンドヘルドコントローラ400Bは、上で説明されるようなカメラを含むことができ、ヘッドギヤ400Aは、カメラと併せた光学追跡のための光学標的を含むことができる。いくつかの例において、ウェアラブル頭部デバイス400Aおよびハンドヘルドコントローラ400Bの各々は、3つの直交して向けられたソレノイドの組を含み、それらは、3つの区別可能な信号を無線で送信および受信するために使用される。受信するために使用されるコイルの各々において受信される3つの区別可能な信号の相対的大きさを測定することによって、ウェアラブル頭部デバイス400Aに対するハンドヘルドコントローラ400Bの6DOFが、決定され得る。いくつかの例において、6DOFトーテムサブシステム404Aは、向上した正確度および／またはハンドヘルドコントローラ400Bの高速移動に関するよりタイムリーな情報を提供するために有用である慣性測定ユニット（IMU）を含むことができる。

#### 【0034】

拡張現実または複合現実用途を伴ういくつかの例において、座標をローカル座標空間（例えば、ウェアラブル頭部デバイス400Aに対して固定される座標空間）から慣性座標空間に変換すること、または環境座標空間に変換することが、望ましくあり得る。例えば、そのような変換は、ウェアラブル頭部デバイス400Aのディスプレイが、ディスプレイ上の固定位置および向きにおいて（例えば、ウェアラブル頭部デバイス400Aのディスプレイにおける同一の位置において）ではなく、仮想オブジェクトを実環境に対する予期される位置および向きにおいて提示する（例えば、ウェアラブル頭部デバイス400Aの位置および向きにかかわらず、前方に面した実椅子に座っている仮想人物）ために必要であり得る。これは、仮想オブジェクトが、実環境内に存在する（かつ、例えば、ウェアラブル頭部デバイス400Aが、シフトおよび回転するにつれて、実環境内に不自然に位置付けられて見えない）という錯覚を維持することができる。いくつかの例において、座標空間の間の補償変換が、慣性または環境座標系に対するウェアラブル頭部デバイス400Aの変換を決定するために、（例えば、同時位置特定およびマッピング（SLAM）および／またはビジュアルオドメトリブロッカを使用して）深度カメラ444からの画像を処理することによって決定されることができる。図4に示される例において、深度カメラ444は、SLAM／ビジュアルオドメトリブロック406に結合されることができ、画像をブロック406に提供することができる。SLAM／ビジュアルオドメトリブロック406実装は、この画像を処理し、次いで、頭部座標空間と実座標空間との間の変換を識別するために使用され得るユーザの頭部の位置および向きを決定するように構成されているプロセッサを含むことができる。同様に、いくつかの例において、ユーザの頭部姿勢および場所に関する情報の追加の源が、ウェアラブル頭部デバイス400AのIMU409から取得される。IMU409からの情報は、SLAM／ビジュアルオドメトリブロック406からの情報と統合され、向上した正確度および／またはユーザの頭部姿勢および位置の高速調節に関するよりタイムリーな情報を提供することができる。

#### 【0035】

いくつかの例において、深度カメラ444は、ウェアラブル頭部デバイス400Aのプロセッサ内に実装され得る手のジェスチャトラッカ411に3D画像を供給することができる。手のジェスチャトラッカ411は、例えば、深度カメラ444から受信された3D画像を手のジェスチャを表す記憶されたパターンに合致させることによって、ユーザの手のジェスチャを識別することができる。ユーザの手のジェスチャを識別する他の好適な技法も、明らかであろう。

## 【0036】

いくつかの例において、1つ以上のプロセッサ416は、ヘッドギヤサブシステム404B、IMU409、SLAM／ビジュアルオドメトリブロック406、深度カメラ444、マイクロホン（図示せず）、および／または手のジェスチャトラッカ411からのデータを受信するように構成され得る。プロセッサ416は、制御信号を6DOFトーテムシステム404Aに送信し、それから受信することもできる。プロセッサ416は、ハンドヘルドコントローラ400Bが繋がれていない例等において、6DOFトーテムシステム404Aに無線で結合され得る。プロセッサ416は、視聴覚コンテンツメモリ418、グラフィカル処理ユニット（GPU）420、および／またはデジタル信号プロセッサ（DSP）オーディオ空間化装置422等の追加のコンポーネントとさらに通信し得る。DSPオーディオ空間化装置422は、頭部関連伝達関数（HRTF）メモリ425に結合され得る。GPU420は、画像毎に変調された光424の左源に結合される左チャンネル出力と、画像毎に変調された光426の右源に結合される右チャンネル出力とを含むことができる。GPU420は、立体視画像データを画像毎に変調された光424、426の源に出力することができる。DSPオーディオ空間化装置422は、オーディオを左スピーカ412および／または右スピーカ414に出力することができる。DSPオーディオ空間化装置422は、プロセッサ416から、ユーザから仮想音源（例えば、ハンドヘルドコントローラ400Bを介して、ユーザによって移動させられ得る）への方向ベクトルを示す入力を受信することができる。方向ベクトルに基づいて、DSPオーディオ空間化装置422は、対応するHRTFを決定することができる（例えば、HRTFにアクセスすることによって、または複数のHRTFを補間することによって）。DSPオーディオ空間化装置422は、次いで、決定されたHRTFを仮想オブジェクトによって発生させられた仮想音に対応するオーディオ信号等のオーディオ信号に適用することができる。これは、複合現実環境内の仮想音に対するユーザの相対的位置および向きを組み込むことによって、すなわち、その仮想音が、実環境内の実音である場合に聞こえるであろうもののユーザの予期に合致する仮想音を提示することによって、仮想音の信憑性および現実性を向上させることができる。

## 【0037】

図4に示されるもの等のいくつかの例において、プロセッサ416、GPU420、DSPオーディオ空間化装置422、HRTFメモリ425、およびオーディオ／視覚的コンテンツメモリ418のうちの1つ以上は、補助ユニット400C（上で説明される補助ユニット320に対応し得る）内に含まれ得る。補助ユニット400Cは、バッテリー427を含み、そのコンポーネントを給電し得、および／または、それは、電力をウェアラブル頭部デバイス400Aおよび／またはハンドヘルドコントローラ400Bに供給し得る。そのようなコンポーネントをユーザの腰部に搭載され得る補助ユニット内に含むことは、ウェアラブル頭部デバイス400Aのサイズおよび重量を限定することができ、それは、次に、ユーザの頭部および頸部の疲労を低減させることができる。

## 【0038】

図4は、例示的ウェアラブルシステム400の種々のコンポーネントに対応する要素を提示するが、これらのコンポーネントの種々の他の好適な配置も、当業者に明白であろう。例えば、補助ユニット400Cに関連付けられているような図4に提示される要素は、代わりに、ウェアラブル頭部デバイス400Aまたはハンドヘルドコントローラ400Bに関連付けられ得る。さらに、いくつかのウェアラブルシステムは、ハンドヘルドコントローラ400Bまたは補助ユニット400Cを完全に無くし得る。そのような変更および修正は、開示される例の範囲内に含まれるとして理解されるべきである。

## 【0039】

（複合現実環境）

## 【0040】

全ての人々のように、複合現実システムのユーザは、実環境の中に存在し、すなわち、ユーザによって知覚可能である「実世界」の3次元部分およびその内容全ての中に存在し

10

20

30

40

50

ている。例えば、ユーザは、その通常の間人感覚、すなわち、視覚、聴覚、触覚、味覚、嗅覚を使用して実環境を知覚し、実環境内でその自身の身体を移動させることによって実環境と相互作用する。実環境内の場所は、座標空間内の座標として説明されることができ、例えば、座標は、緯度、経度、および海面に対する高度、基準点からの3つの直交する次元における距離、または他の好適な値を含むことができる。同様に、ベクトルは、座標空間における方向および大きさを有する品質を説明することができる。

#### 【0041】

コンピューティングデバイスは、例えば、デバイスに関連付けられたメモリ内に仮想環境の表現を維持することができる。本明細書に使用されるように、仮想環境は、3次元空間のコンピュータ表現である。仮想環境は、任意のオブジェクト、アクション、信号、パラメータ、座標、ベクトル、またはその空間に関連付けられた他の特性の表現を含むことができる。いくつかの例において、コンピューティングデバイスの回路（例えば、プロセッサ）は、仮想環境の状態を維持および更新することができ、すなわち、プロセッサは、第1の時間に、仮想環境に関連付けられたデータおよび／またはユーザによって提供される入力に基づいて、第2の時間における仮想環境の状態を決定することができる。例えば、仮想環境内のオブジェクトが、ある時間における第1の座標に位置し、あるプログラムされた物理的パラメータ（例えば、質量、摩擦係数）を有し、ユーザから受信された入力（例えば、力が、ある方向ベクトルにおいてオブジェクトに加えられるべきであると示す場合、プロセッサは、運動学の法則を適用し、基本的力学を使用してその時間におけるオブジェクトの場所を決定することができる。プロセッサは、仮想環境についての既知の任意の好適な情報および／または任意の好適な入力を使用し、ある時間における仮想環境の状態を決定することができる。仮想環境の状態を維持および更新することにおいて、プロセッサは、任意の好適なソフトウェアを実行することができ、任意の好適なソフトウェアは、仮想環境内の仮想オブジェクトの作成および削除に関連するソフトウェア、仮想環境内の仮想オブジェクトまたはキャラクタの挙動を定義するためのソフトウェア（例えば、スクリプト）、仮想環境内の信号（例えば、オーディオ信号）の挙動を定義するためのソフトウェア、仮想環境に関連付けられたパラメータを作成および更新するためのソフトウェア、仮想環境内のオーディオ信号を発生させるためのソフトウェア、入力および出力を取り扱うためのソフトウェア、ネットワーク動作を実装するためのソフトウェア、アセットデータ（例えば、経時的に仮想オブジェクトを移動させるためのアニメーションデータ）を適用するためのソフトウェア、または多くの他の可能性を含む。

#### 【0042】

ディスプレイまたはスピーカ等の出力デバイスは、仮想環境の任意または全ての側面をユーザに提示することができる。例えば、仮想環境は、ユーザに提示され得る仮想オブジェクト（無生物オブジェクト、人物、動物、光等の表現を含み得る）を含み得る。プロセッサは、仮想環境の表示（例えば、原点座標、視軸、および錐台を伴う「カメラ」に対応する）を決定し、ディスプレイに、その表示に対応する仮想環境の視認可能な場面をレンダリングすることができる。任意の好適なレンダリング技術が、この目的のために使用され得る。いくつかの例において、視認可能な場面は、仮想環境内のいくつかの仮想オブジェクトのみを含み、ある他の仮想オブジェクトを除外し得る。同様に、仮想環境は、1つ以上のオーディオ信号としてユーザに提示され得るオーディオ側面を含み得る。例えば、仮想環境内の仮想オブジェクトが、オブジェクトの場所座標から生じる音を発生させ得る（例えば、仮想キャラクタが、発話し、または効果音を引き起こし得る）；または、仮想環境は、特定の場所に関連付けられることも、そうではないこともある音楽的キューまたは周囲音に関連付けられ得る。プロセッサが、「聴者」座標に対応するオーディオ信号（例えば、仮想環境内の音の複合物に対応し、聴者座標における聴者に聞こえるであろうオーディオ信号をシミュレートするために混合および処理されたオーディオ信号）を決定し、1つ以上のスピーカを介してユーザにオーディオ信号を提示することができる。

#### 【0043】

仮想環境は、コンピュータ構造としてのみ存在するので、ユーザは、その通常の感覚を

10

20

30

40

50

使用して仮想環境を直接知覚することができない。代わりに、ユーザは、例えば、ディスプレイ、スピーカ、触覚出力デバイス等によって、ユーザに提示されるような仮想環境を間接的にのみ知覚することができる。同様に、ユーザは、仮想環境に直接触れること、それを操作すること、または別様にそれと相互作用することができないが、入力デバイスまたはセンサを介して、仮想環境を更新するためにデバイスまたはセンサデータを使用し得るプロセッサに入力データを提供することができる。例えば、カメラセンサは、ユーザが仮想環境内のオブジェクトを移動させようとしていることを示す光学データを提供することができ、プロセッサは、そのデータを使用し、オブジェクトに仮想環境内でそれに応じて応答させることができる。

【0044】

(デジタル反響および環境オーディオ処理)

【0045】

XRシステムは、原点座標を伴う音源において生じ、システムにおける向きベクトルの方向に進行するように思われるオーディオ信号をユーザに提示することができる。ユーザは、それらが、音源の原点座標から生じ、向きベクトルに沿って進行する実オーディオ信号であるかのように、これらのオーディオ信号を知覚し得る。

【0046】

ある場合、オーディオ信号は、それらが、仮想環境内のコンピュータ信号に対応し、必ずしも、実環境内の実音に対応するわけではないという点で、仮想と見なされ得る。しかしながら、仮想オーディオ信号は、人間の耳によって検出可能な実オーディオ信号として、例えば、図1におけるウェアラブル頭部デバイス100のスピーカ120Aおよび120Bを介して発生させられたものとして、ユーザに提示されることができる。

【0047】

下で開示される実施形態の利点は、低減させられたネットワーク帯域幅、低減させられた電力消費、低減させられた算出複雑性、および低減させられた算出遅延を含む。これらの利点は、処理リソース、ネットワーキングリソース、バッテリー容量、および物理的サイズおよび重量が、多くの場合、限られているウェアラブルシステムを含むモバイルシステムに特に顕著であり得る。

【0048】

ARと同程度に動的な環境内で、システムは、オーディオ信号を連続的にレンダリングし得る。仮想スピーカの全てを使用してオーディオ信号をレンダリングすることは、高算出能力、大量の処理、高ネットワーク帯域幅、高電力消費等に特につながり得る。したがって、1つ以上の因子に基づいて固定仮想スピーカの一部を動的に選択し、使用するために、修正された仮想スピーカパンニングを使用することが、所望され得る。

【0049】

(例示的空間オーディオシステム)

【0050】

図5Aは、いくつかの実施形態による、例示的空間オーディオシステムのブロック図を図示する。図5Bは、図5Aのシステムを動作させるための例示的方法のフローを図示する。

【0051】

空間オーディオシステム500は、空間モデラ510と、内部空間表現530と、デコーダ／バーチャライザ540Aとを含み得る。空間モデラ510は、直接経路部分512と、1つ以上の反射部分520（随意）と、空間エンコーダ526とを含み得る。空間モデラ510は、仮想環境をモデル化するように構成され得る。直接経路部分512は、直接源514と、随意に、ドップラ516とを含み得る。直接源514は、オーディオ信号を提供するように構成され得る（プロセス550のステップ552）。ドップラ516は、直接源514から信号を受信し得、その入力信号の中にドップラ効果を導入するように構成され得る（ステップ554）。例えば、ドップラ516は、音源、システムのユーザ、または両方の運動に対して変化するように音源のピッチを変化させ得る（例えば、ピッ

10

20

30

40

50

チシフト)。

【0052】

反射部分520は、音リフレクタ522と、随意のドップラ516と、遅延524とを含み得る。音リフレクタ522は、その信号内に反射を導入するように構成され得る(ステップ556)。導入される反射は、環境の1つ以上の特性を表し得る。反射部分520内のドップラ516は、音リフレクタ522から信号を受信し得、その入力信号の中にドップラ効果を導入するように構成され得る(ステップ558)。遅延524は、ドップラ516から信号を受信し得、遅延を導入するように構成され得る(ステップ560)。

【0053】

空間エンコーダ526は、直接経路部分512および反射部分520から信号を受信し得る。いくつかの実施形態において、直接経路部分512から空間エンコーダ526への信号は、直接経路部分512のドップラ516からの出力信号であり得る。いくつかの実施形態において、反射部分520から空間エンコーダ526への信号は、反射部分520の遅延524からの出力信号であり得る。

【0054】

空間エンコーダ526は、1つ以上のM方向パン528を含み得る。いくつかの実施形態において、空間エンコーダ526によって受信される各入力は、独自の528に関連付けられ得る。「パンニング」は、複数のスピーカ、複数の場所、または両方にわたって信号を分配することを指し得る。M方向パン528は、複数の数の仮想スピーカにわたってその入力信号を分配するように構成され得る(ステップ562)。例えば、M方向パン528は、全てのM個の仮想スピーカにわたってその入力信号を分配することができる。例えば、図5Aに示されるように、Mは、4に等しくあり得、各M方向パン528は、4つの仮想スピーカにわたってその入力信号を分配するように構成され得る。図は、4つの仮想スピーカを有するシステムを図示するが、本開示の例は、任意の数の仮想スピーカを含むことができる。

【0055】

一例として、自動車システムが、左および右スピーカを含み得る。そのようなシステムにおける音は、各スピーカのために1つ、2つに音を分割することによって、自動車における左および右スピーカの間でパンされ得る。各スピーカのスケールンボリュームが、2つのスピーカの構成に従って設定され得、結果は、左および右スピーカに送信され得る。

【0056】

別の例として、サラウンド音システムが、6つのスピーカ等の複数のスピーカを含み得る。そのようなシステムにおける音は、6つのスピーカの間でステレオとしてパンされ得る。音は、6つ(自動車システム例におけるような2つの代わりに)に分割され得、各スピーカのスケールンボリュームが、6つのスピーカの構成に従って設定され得、結果は、6つのスピーカに送信され得る。

【0057】

例えば、第1のM方向パン528が、直接経路512のドップラ516の出力を受信し得、他のM方向パン528が、反射部分520の出力を受信し得る。各M方向パン528は、それが複数の出力にわたって分配され得るように、その入力信号を分割することができる。したがって、各M方向パン528は、入力より大きい数の出力を有し得る。

【0058】

空間モデラ510は、信号を内部空間表現530に出力し得る(ステップ564)。いくつかの実施形態において、空間モデラ510からの出力は、各M方向パン528の出力を含むことができる。内部空間表現530は、仮想環境の空間構成を表すように構成され得る(ステップ566)。一例示的表現は、ユーザ、音源、および仮想スピーカの相対的場所を表すことを含むことができる。いくつかの実施形態において、内部空間表現530は、システム500のユーザの頭部姿勢回転、頭部姿勢平行移動、音場デコード、1つ以上の頭部関連伝達関数(HRTF)、またはそれらの組み合わせを表す1つ以上の信号を出力し得る。いくつかの実施形態において、内部空間表現530は、非アンビソニックス

10

20

30

40

50

マルチチャネルベースのシステム、アンビソニックス／波動場ベースのシステム等の表現であり得る。一例示的アンビソニックス／波動場ベースのシステムは、高次アンビソニックス（H O A）であり得る。

【 0 0 5 9 】

内部空間表現 5 3 0 は、その信号 5 5 2 をデコーダ／バーチャライザ 5 4 0 A に出力し得る（ステップ 5 6 8）。デコーダ／バーチャライザ 5 4 0 は、その入力信号をデコードし、仮想音を信号の中に導入し得る（ステップ 5 7 0）。ステップ 5 7 0 は、複数のサブステップを含むことができ、下でより詳細に議論される。システムは、次いで、デコーダ／バーチャライザ 5 4 0 からの信号を左スピーカに出力され得る左信号 5 0 2 L として、かつ右スピーカに出力され得る右信号 5 0 2 R として出力する（ステップ 5 8 0）。10

【 0 0 6 0 】

システム 5 0 0 は、任意の数の異なるタイプのデコーダ／バーチャライザ 5 4 0 を含み得る。一例示的デコーダ／バーチャライザ 5 4 0 A が、図 5 A に示される。他の例示的デコーダ／バーチャライザ 5 4 0 が、下で議論される。

【 0 0 6 1 】

デコーダ／バーチャライザ 5 4 0 A は、回転／平行移動表現 5 4 2 と、音場デコーダ 5 4 4 と、1 つ以上の H R T F 5 4 6 と、1 つ以上のコンバイナ 5 4 8 とを含み得る。図 5 C は、ステップ 5 7 0 - 1 と称され得る例示的デコーダ／バーチャライザを動作させるための例示的方法のフローを図示する。回転／平行移動表現 5 4 2 は、内部空間表現 5 3 0 から信号を受信し得、オーディオ信号に関連付けられた移動の表現を導入するように構成され得る。例えば、移動は、音源、ユーザ、または両方のものであり得る（ステップ 5 7 2）。回転／平行移動表現 5 4 2 は、信号を音場デコーダ 5 4 4 に出力することができる。音場デコーダ 5 4 4 は、回転／平行移動表現 5 4 2 から信号を受信し得、信号をデコードするように構成され得る（ステップ 5 7 4）。各 H R T F 5 4 6 は、音場デコーダ 5 4 4 から信号を受信し得る。各 H R T F 5 4 6 は、その入力信号に対応する H R T F を決定し、それを信号に適用するように構成され得る（ステップ 5 7 6）。1 つ以上の H R T F 5 4 6 は、スピーカバーチャライザと集合的に称され得る。いくつかの実施形態において、H R T F 5 4 6 は、有限インパルス応答（F I R）フィルタ処理のために構成され得る。各コンバイナ 5 4 8 は、H R T F 5 4 6 から信号を受信し、組み合わせ得る（ステップ 5 7 8）。20

【 0 0 6 2 】

いくつかの実施形態において、デコーダ／バーチャライザ 5 4 0 A は、「ベースライン」処理オーバーヘッドを表し得る。ベースライン処理オーバーヘッドは、複合体であり、各仮想スピーカのために H R T F 処理を適用するための行列計算および長い F I R フィルタを伴い得る。

【 0 0 6 3 】

コンバイナ 5 4 8 からの出力は、システム 5 0 0 からの出力信号であり得る。いくつかの実施形態において、システム 5 0 0 からの出力信号 5 0 2 は、左および右スピーカ（例えば、図 1 のスピーカ 1 2 0 A および 1 2 0 B）のためのオーディオ信号であり得る。

【 0 0 6 4 】

いくつかのインスタンスにおいて、再生のための音源の数が多いとき、図 5 A の空間オーディオシステムは、有益であり得る。しかしながら、いくつかのインスタンスにおいて、再生のための音源の数が少ないとき、図 5 A の空間オーディオシステムは、有益でないこともある。再生のための音源の数が、少ないときの状況のために効率的な方法で、図 5 A のシステム 5 0 0 等の非アンビソニックスマルチチャネルベースの空間オーディオシステムまたはアンビソニックスベースの空間オーディオシステムの効率を利用することが、望ましくあり得る。

【 0 0 6 5 】

音場合成およびデコーディングを使用して空間化の効率を改良する方法が、存在し得る。第 1 の方法は、低エネルギースピーカ検出およびカリングを通してであり得る。低エネ50

ルギースピーカ検出およびカリングにおいて、非アンビソニックスマルチチャネルベースの空間オーディオシステムの仮想スピーカチャンネルまたはアンビソニックスベースの空間オーディオシステムのアンビソニックス／音場チャンネルのエネルギー出力が、所定の閾値より小さい場合、仮想スピーカチャンネルからの信号の処理は、実施されない。いくつかの実施形態において、システムは、例えば、音場デコーディングがその所与の仮想スピーカからの信号に対して実施される前、所与の仮想スピーカの出力が所定の閾値より大きいかどうかを決定し得る。低エネルギースピーカ検出およびカリングは、下でより詳細に議論される。

#### 【0066】

音場合成およびデコーディングを使用して空間化の効率を改良するための第2の方法は、源幾何学形状ベースの仮想スピーカカリングであり得る。源幾何学形状ベースの仮想スピーカカリングにおいて、デコーダ／バーチャライザ処理は、選択的に無効にされることができる。選択的無効化（または選択的有効化）は、ユーザ／聴者に対する音源の場所に基づくことができる。源幾何学形状ベースの仮想スピーカカリングは、下でより詳細に議論される。

#### 【0067】

第3の方法は、低エネルギースピーカ検出およびカリング技法を源－仮想スピーカ結合技法と組み合わせることであり得る。

#### 【0068】

空間モデラ510は、オーディオ信号を処理するために必要とされる動作の回数を表示する算出複雑性を有し得る。算出複雑性は、MにNを乗算したものに比例し得、Mは、音源（直接源および随意的反射を含む）の数に等しくあり得、Nは、アンビソニック音場を表すために必要とされるチャンネルの数に等しくあり得る。いくつかの実施形態において、Nは、 $(O+1)^2$ に等しくあり得、式中、Oは、使用されるアンビソニックスの次数である。

#### 【0069】

デコーダ／バーチャライザ540は、 $nVS$ に比例する算出複雑性を有し得、 $nVS$ は、仮想スピーカの数である。各スピーカの算出能力は、高くあり得、それは、概してFIRフィルタの対から成り得、それらは、高速フーリエ変換（FFT）または逆FFT（IFFT）を用いて典型的に実装され、それらの両方は、コンピュータ的に高コストなプロセスであり得る。

#### 【0070】

（例示的低エネルギー出力検出およびカリング方法）

#### 【0071】

いくつかの実施形態において、いくつかの仮想スピーカが、信号入力エネルギーを殆どまたは全く有していないこともある：例えば、空間オーディオシステムが、少数の音源を有するとき。スピーカ仮想化処理は、コンピュータ的に高コストな（例えば、CPU集約的）プロセスであり得る。例えば、音源が、ゼロ度方位に（例えば、ユーザの正面に直接）位置する場合、90度～270度方位に（例えば、ユーザの後方に）位置する仮想スピーカからの信号に、エネルギーが、殆どまたは全く存在しないこともある。低エネルギー信号は、音源の知覚される場所に対して重要な効果を有しないこともあり、したがって、低エネルギー信号に対してスピーカ仮想化処理を実施すること、および／または対応する仮想スピーカの特性を決定することは、コンピュータ的に非効率的であり得る。

#### 【0072】

要求される算出リソースを減らすために、低エネルギー出力検出およびカリング方法を採用するシステムは、音場デコーダとHRTFとの間に位置する検出器を含むことができる。代替として、検出器は、マルチチャネル出力とHRTFとの間に位置し得る。検出器は、1つ以上の仮想スピーカからの1つ以上のオーディオ信号に関連付けられた1つ以上のエネルギーレベルを検出するように構成され得る。

#### 【0073】

10

20

30

40

50

仮想スピーカ  $V_n$  から発する信号のエネルギーレベルが、エネルギー閾値  $\alpha$  未満である場合、信号は、低エネルギー信号と見なされ得る。オーディオ信号に関連付けられた検出されたエネルギーレベルが、エネルギー閾値  $\alpha$  未満であることに従って、HRTFブロックおよび低エネルギー信号のその処理は、バイパスされ得る。

【0074】

信号のエネルギーレベルの決定は、任意の数の技法を使用し得る。例えば、RMSアルゴリズムが、そのエネルギーを測定するために、仮想スピーカにルーティングされる信号に適用され得る。従来のオーディオコンプレッサによるそれらに類似する時間によって使用されるそれらに類似する「アタック」および「リリース」時間が、スピーカの信号が突然「ポップイン」および「ポップアウト」することを防ぐために使用され得る。

10

【0075】

図6は、いくつかの実施形態による、音源およびスピーカの例示的構成を図示する。システム600は、音源620と、複数のスピーカとを含み得る。複数のスピーカ622は、1つ以上のアクティブ仮想スピーカ622Aと、1つ以上の非アクティブ仮想スピーカ622Bとを含み得る。アクティブ仮想スピーカ622Aは、その信号が、所与の時間にHRTF546によって処理されるものであり得る。非アクティブ仮想スピーカ622Bは、例えば、その信号が、以前の時間にすでに処理されたので、または、仮想スピーカ622Bからの信号が処理を必要としないとシステムが決定しているので、その信号が、HRTF546によって処理される必要がないものであり得る。Mは、再生される音源の数を指し得、Nは、システム内の仮想スピーカの数を指し得る。図は、単一の音源を図示するが、本開示の例は、任意の数の音源を含むことができる。図は、8つの音源を図示するが、本開示の例は、16個 ( $N=16$ ) 等の任意の数の源を含むことができる。

20

【0076】

一例として、システム600は、図に示されるように、単一 ( $M=1$ ) の音源620と、8つの仮想スピーカ622とを含むことができる。所与のインスタンスにおいて、エネルギーの大部分が、3つのみの仮想スピーカにわたって出力され得る。すなわち、システム600は、第1の時間において3つのアクティブ仮想スピーカを有し得る。例えば、仮想スピーカ622A-1、622A-2、および622-3は、アクティブ仮想スピーカであり得る。いくつかの実施形態において、アクティブ仮想スピーカ622Aは、音源620に最も近いそれらであり得る。加えて、システム600は、5つの非アクティブ仮想スピーカ622Bを含み得る。システム600は、5つの非アクティブ仮想スピーカの各々からのエネルギーレベルが、エネルギー閾値より小さいと決定し得、そのような決定に従って、5つの非アクティブ仮想スピーカ622Bからの信号のHRTF処理をバイパスし得る。

30

【0077】

システム600は、アクティブ仮想スピーカの各々からのエネルギーレベルがエネルギー閾値より小さくないことも決定し得、そのような決定に従って、3つのアクティブ仮想スピーカ622Aからの信号のHRTF処理を実施し得る。

【0078】

システム600は、図5Aに示されるように、2つの信号、すなわち、(右信号502Rおよび左信号502L等の) 右スピーカのために1つ、左スピーカのために1つを出力し得る。HRTF処理をバイパスすることによるHRTF動作の回数の低減は、非アクティブ仮想スピーカの数にシステムから出力される信号の数を乗算したものに等しくあり得る。図6の例において、5つの信号のHRTF処理が、バイパスされるので、10回 (5つの非アクティブ仮想スピーカ  $\times$  2つの出力信号) のHRTF動作が、節約され得る。

40

【0079】

別の例として、システムが、13個が非アクティブ仮想スピーカである16個の仮想スピーカを含む場合、節約されるHRTF動作の回数は、26回 (16個の仮想スピーカ  $\times$  2つの出力信号) に等しくあり得る。

【0080】

50

図 7 A は、いくつかの実施形態による、複数の検出器を含む例示的デコーダ／バーチャライザのブロック図を図示する。図 7 B は、いくつかの実施形態による、図 7 A のデコーダ／バーチャライザを動作させるための例示的方法のフローを図示する。いくつかの実施形態において、下で議論されるように、デコーダ／バーチャライザ 5 4 0 A（図 5 A に示される）の代わりに、デコーダ／バーチャライザ 5 4 0 B が、システム 5 0 0 内に含まれ得る。ステップ 5 7 0 - 1（図 5 C に示される）の代わりに、ステップ 5 7 0 - 2 が、プロセス 5 5 0 内に含まれ得る。

#### 【0081】

デコーダ／バーチャライザ 5 4 0 B は、回転／平行移動表現 5 4 2 と、音場デコーダ 5 4 4 と、1 つ以上の検出器 7 1 0 と、1 つ以上のスイッチ 7 1 2 と、1 つ以上の H R T F 5 4 6 と、1 つ以上のコンバイナ 5 4 8 とを含むことができる。デコーダ／バーチャライザ 5 4 0 B は、内部空間表現 5 3 0（図 5 A に示されるような）から信号 5 5 2 を受信することができる。回転／平行移動表現 5 4 2 は、内部空間表現 5 3 0 から信号を受信し得、音源、ユーザ、または両方の移動の表現を導入するように構成され得る（ステップ 7 7 2）。回転／平行移動表現 5 4 2 は、信号を音場デコーダ 5 4 4 に出力することができる。音場デコーダ 5 4 4 は、回転／平行移動表現 5 4 2 から信号を受信することができ、信号をデコードするように構成され得る（ステップ 7 7 4）。音場デコーダ 5 4 4 は、信号を検出器 7 1 0 に出力することができる。

#### 【0082】

検出器 7 1 0 は、音場デコーダ 5 4 4 から信号を受信し得、その入力信号のエネルギーレベルを決定するように構成され得る（ステップ 7 7 6）。各検出器 7 1 0 は、独自のスイッチ 7 1 2 に結合され得る。（音場デコーダ 5 4 4 からの）入力信号のエネルギーレベルがエネルギー閾値以上である場合（ステップ 7 7 8）、スイッチ 7 1 2 は、ループを閉にし、それによって、（検出器 7 1 0 からの）その入力信号をスイッチが結合される H R T F 5 4 6 にルーティングすることができる（ステップ 7 8 0）。各 H R T F は、対応する H R T F を決定し、それを信号に適用する（ステップ 7 8 2）。

#### 【0083】

入力信号のエネルギーレベルがエネルギー閾値より小さい場合、スイッチ 7 1 2 は、（検出器 7 1 0 からの）その入力信号が対応する H R T F 5 4 6 に結合されないように、開にすることができる。したがって、対応する H R T F 5 4 6 は、バイパスされ得る（ステップ 7 8 4）。

#### 【0084】

H R T F 5 4 6 からの信号は、コンバイナ 5 4 8 に出力されることができる（ステップ 7 8 6）。コンバイナ 5 4 8 は、H R T F 5 4 6 からの信号を組み合わせる（例えば、追加する、集約する等）ように構成されることができる。H R T F 5 4 6 をバイパスしたそれらの信号は、コンバイナ 5 4 8 によって組み合わせられない。コンバイナ 5 4 8 からの出力は、システム 5 0 0 からの出力信号であり得る。いくつかの実施形態において、システム 5 0 0 からの出力信号 5 0 2 は、左および右スピーカ（例えば、図 1 のスピーカ 1 2 0 A および 1 2 0 B）のためのオーディオ信号であり得る。

#### 【0085】

いくつかの実施形態において、各検出器 7 1 0 は、仮想スピーカに対応する独自の信号に結合されることができる。このように、各仮想スピーカ 6 2 2 の処理は、独立して実施されることができる（すなわち、6 2 2 A - 1 等の 1 つのスピーカの処理は、6 2 2 B 等の別のスピーカの処理に影響を及ぼすことなく行われることができる）。

#### 【0086】

いくつかの実施形態において、デコーダ／バーチャライザ 5 4 0 のタイプは、音源の数に依存し得る。例えば、音源の数が、所定の音源閾値より小さいか、またはそれに等しい場合、図 7 A のデコーダ／バーチャライザ 5 4 0 B が、システム 5 0 0 内に含まれ得る。そのようなインスタンスにおいて、音場デコーダ 5 4 4 からの信号は、検出器 7 1 0 に入力され得る。

10

20

30

40

50

## 【0087】

音源の数が、所定の音源閾値より大きい場合、図5Aのデコーダ／バーチャライザ540Aが、システム内に含まれ得る。そのようなインスタンスにおいて、音場デコーダ544からの信号は、HRTF546に入力され得る。

## 【0088】

いくつかの実施形態において、システムは、検出器およびそのエネルギーレベル検出を  
実行すべきか、バイパスすべきかを選択し得るデコーダ／バーチャライザ540を含み得  
る。図8Aは、いくつかの実施形態による、例示的デコーダ／バーチャライザのブロック  
図を図示する。図8Bは、いくつかの実施形態による、図8Aのデコーダ／バーチャライ  
ザを動作させるための例示的方法のフローを図示する。いくつかの実施形態において、デ  
コーダ／バーチャライザ540A（図5Aに示される）およびデコーダ／バーチャライザ  
540B（図7Aに示される）の代わりに、デコーダ／バーチャライザ540Cが、シス  
テム500内に含まれ得る。ステップ570-1（図5Cに示される）の代わりに、ステ  
ップ570-3が、プロセス550内に含まれ得る。

10

## 【0089】

デコーダ／バーチャライザ540Cは、上で議論されるデコーダ／バーチャライザ540Bと同様、回転／平行移動表現542と、音場デコーダ544と、1つ以上の検出器710と、1つ以上の第1のスイッチ712と、1つ以上のHRTF546と、1つ以上のコンバイナ548とを含むことができる。ステップ872、874、および882は、上で議論されるステップ772、774、および782に対応して類似し得る。

20

## 【0090】

デコーダ／バーチャライザ540Cは、第2のスイッチ814も含み得る。第2のスイ  
ッチ814は、音場デコーダ544から検出器710および第1のスイッチ712への第  
1のループを開または閉にするように構成されることができる。加えて、または代替とし  
て、第2のスイッチ814は、検出器710および第1のスイッチ712をバイパスする  
システム500からの第2のループを開または閉にするように構成されることができる。  
いくつかの実施形態において、第2のスイッチ814は、信号を検出器710に直接通す  
こと（第1のループ）またはHRTF546に直接通すこと（第2のループ）の間で選択  
するように構成されている、双方向スイッチであり得る。

## 【0091】

例えば、システムは、音源の数が所定の音源閾値以上かどうかを決定することができる  
（ステップ876）。音源の数が、所定の音源閾値以上である場合、第2のスイッチ814は、第2のループを閉にし、音場デコーダ544からの信号をHRTF546に直接通すことができる（ステップ878）。各HRTF546は、次いで、対応するHRTFを決定し、それを信号に適用する（ステップ880）。音源の数が、数において上回るとき、信号が低エネルギーレベルを有する可能性は、低減させられ得る。

30

## 【0092】

一方、音源の数が、所定の音源閾値より小さい場合、信号は、低エネルギーレベルを有  
する可能性が高く、したがって、第2のスイッチ814は、第1のループを閉にし、音場  
デコーダ544からの信号を検出器710に直接通すことができる（ステップ882）。  
検出器710は、音場デコーダ544から信号を受信し得、その入力信号のエネルギーレ  
ベルを決定するように構成され得る（ステップ884）。（音場デコーダ544からの）  
入力信号のエネルギーレベルが、エネルギー閾値以上である場合（ステップ886）、ス  
イッチ712は、ループを閉にし、それによって、（検出器710からの）その入力信号  
を、スイッチが結合されるHRTF546にルーティングすることができる（ステップ8  
88）。入力信号のエネルギーレベルが、エネルギー閾値より小さい場合、スイッチ712は、（検出器710からの）その入力信号が、対応するHRTF546に結合されないように、開にし、HRTF546がバイパスされるようにすることができる（ステップ890）。

40

## 【0093】

50

H R T F 5 4 6 からの信号は、コンバイナ 5 4 8 に出力されることができる（ステップ 8 9 2）。

【0 0 9 4】

いくつかの実施形態において、1つ以上のエネルギー閾値検出は、エネルギーにตอบสนองしてアクティブであり得る。いくつかの実施形態において、1つ以上のエネルギー閾値検出は、振幅にตอบสนองしてアクティブであり得、従来のアタック、リリース時間等を受け得る。

【0 0 9 5】

（例示的源幾何学形状ベースのスピーカカリング方法）

【0 0 9 6】

源幾何学形状ベースの仮想スピーカカリングは、C P U消費を低減させるための別の方法であり得る。いくつかの実施形態において、源幾何学形状ベースの仮想スピーカカリングは、デコーダ／バーチャライザ処理（例えば、図 5 A のデコーダ／バーチャライザ 5 4 0 A、図 7 A のデコーダ／バーチャライザ 5 4 0 B、図 8 A のデコーダ／バーチャライザ 5 4 0 C 等）を選択的に無効にすることを含むことができる。いくつかの実施形態において、選択的無効化（または選択的有効化）は、ユーザ／聴者に対する音源の場所に基づくことができる。いくつかの実施形態において、デコーダ／バーチャライザ処理の選択的無効化は、デコーダ／バーチャライザの処理ブロックの全てをバイパスするステップを含むことができる。

【0 0 9 7】

源幾何学形状ベースの仮想スピーカカリングにおいて、アンビソニック出力が、計算されることができる。アンビソニック出力が、かなりの量のエネルギーがデコードされることを要求する場合、リアルタイムエネルギー検出方法等の（より少ないC P U消費を要求する）より単純な方法を使用することが、有益であり得る。加えて、いくつかの実施形態において、リアルタイムエネルギー検出方法は、より少ない頻度で計算を実施することができる。

【0 0 9 8】

図 9 は、いくつかの実施形態による、音源およびスピーカの例示的構成を図示する。システム 9 0 0 は、音源 9 2 0 と、複数のスピーカを含み得る。図 6 のシステム 6 0 0 と比較して、音源 9 2 0 は、図 6 の音源 6 2 0 の第 1 の位置と異なり得る第 2 の位置に位置し得る。複数のスピーカ 9 2 2 は、1つ以上のアクティブ仮想スピーカ 9 2 2 A と、1つ以上の非アクティブ仮想スピーカ 9 2 2 B と、1つ以上の非アクティブ仮想スピーカ 9 2 2 C とを含み得る。アクティブ仮想スピーカ 9 2 2 A および非アクティブ仮想スピーカ 9 2 2 B は、それぞれ、図 6 のアクティブ仮想スピーカ 6 2 2 A および非アクティブ仮想スピーカ 6 2 2 B に対応して類似し得る。

【0 0 9 9】

非アクティブ仮想スピーカ 9 2 2 C は、仮想スピーカ 9 2 2 C が、第 1 の時間にアクティブであるが、その信号が、第 2 の時間（例えば、リングアウト周期）に処理されている点において、非アクティブ仮想スピーカ 9 2 2 B と異なり得る。図 9 の例において、音源 9 2 0 は、第 1 の位置（例えば、仮想スピーカ 9 2 2 C に近接する）から第 2 の位置（例えば、仮想スピーカ 9 2 2 に近接しない）に移動していることもある。音源の移動に起因して、2つの仮想スピーカは、第 2 の時間にそれらの中に混合する音源をもはや有しないこともある。2つの仮想スピーカのフィルタ処理に起因して、2つの仮想スピーカは、フィルタ処理を適切に完了させるために、続くフレーム（例えば、第 2 の時間）のためにアクティブである必要があり得る。

【0 1 0 0】

いくつかの実施形態において、システムは、アクティブ仮想スピーカを使用するシステム内にデコーダ／バーチャライザ 5 4 0 を含み得る。図 1 0 A は、いくつかの実施形態による、アクティブスピーカを含むシステムにおいて使用される例示的デコーダ／バーチャライザのブロック図を図示する。図 1 0 B は、いくつかの実施形態による、図 1 0 A のデコーダ／バーチャライザを動作させるための例示的方法のフローを図示する。いくつかの

10

20

30

40

50

実施形態において、デコーダ／バーチャライザ５４０Ａ（図５Ａに示される）、デコーダ／バーチャライザ５４０Ｂ（図７Ａに示される）、およびデコーダ／バーチャライザ５４０Ｃ（図８Ａに示される）の代わりに、デコーダ／バーチャライザ５４０Ｄが、システム５００内に含まれ得る。ステップ５７０－１（図５Ｃに示される）、ステップ５７０－２（図７Ｂに示される）、およびステップ５７０－３（図８Ｂに示される）の代わりに、ステップ５７０－４が、プロセス５５０内に含まれ得る。

【０１０１】

デコーダ／バーチャライザ５４０Ｃは、上で議論されるデコーダ／バーチャライザ５４０Ｂおよびデコーダ／バーチャライザ５４０Ｃと同様、音場デコーダ５４４と、１つ以上のＨＲＴＦ５４６と、１つ以上のコンバイナ５４８とを含むことができる。ステップ１０７２、１０７６、１０７８、および１０８０は、上で議論されるステップ８７２、８７４、および７８２に対応して類似し得る。

【０１０２】

デコーダ／バーチャライザ５４０Ｄは、回転／平行移動表現１０４２と、音場デコード決定１０４４とも含み得る。回転／平行移動表現１０４２は、内部空間表現５３０から信号を受信し得、音源、ユーザ、または両方の移動の表現を導入するように構成され得る（ステップ１０７２）。移動の表現は、音源９２０の方位／高度も考慮し得る。回転／平行移動表現５４２は、信号を音場デコーダ決定１０４４に出力することができる。

【０１０３】

音場デコーダ決定１０４４は、回転／平行移動表現１０４２から信号を受信し得、「顕著な」出力を有する信号を決定し、それらの信号を音場デコーダ５４４に通すように構成され得る（ステップ１０７４）。顕著な出力は、知覚される音に影響を及ぼすであろう出力であり得る。例えば、顕著な出力は、所定の振幅閾値以上である振幅を有するオーディオ信号であり得る。音場デコーダ５４４は、顕著な出力を有する音場デコーダ決定１０４４からの信号を受信し得、信号をデコードするように構成され得る（ステップ１０７６）。いくつかの実施形態において、音場デコーダ１０４４は、顕著な出力を有する音場デコーダ決定１０４４からの信号を受信し得る。各ＨＲＴＦ５４６は、音場デコーダ５４４から信号を受信し得る。各ＨＲＴＦ５４６は、その入力信号に対応するＨＲＴＦを決定し、それを信号に適用するように構成され得る（ステップ１０７８）。１つ以上のＨＲＴＦ５４６は、スピーカバーチャライザと集合的に称され得る。各コンバイナ５４８は、ＨＲＴＦ５４６から信号を受信し、組み合わせ得る（ステップ１０８０）。

【０１０４】

いくつかの実施形態において、顕著な出力を有していない（例えば、所定の振幅閾値未満の振幅を有する）それらのオーディオ信号は、音場デコーダ５４４に通されないこともある。したがって、顕著な出力を有していないオーディオ信号上の音場デコーダ５４４およびＨＲＴＦ５４６は、バイパスされ得る。

【０１０５】

例示的源幾何学形状ベースのスピーカカリング方法は、音源の位置（例えば、Ｘ、Ｙ、Ｚ場所）に基づいて、アクティブ仮想スピーカであるように仮想スピーカを指定することができる。音源の場所は、源オブジェクトの場所を表し得る。システムは、各音源の場所を決定し、それぞれの音源に近接して位置する仮想スピーカを決定し得る。いくつかの実施形態において、音源に近接して位置する仮想スピーカの決定は、例えば、全ビデオフレームの開始時に（ビデオフレームレートベースのアプローチで）実施され得る。ビデオフレームレートベースのアプローチは、サンプルレートベースのアプローチ等の他のアプローチより少ない算出を要求し得る。

【０１０６】

音源は、例えば、ビデオフレームレートベースのアプローチ計算およびアンビソニックデコード式に基づいて、特定の仮想スピーカに大きく寄与し得る。上で議論されるように、デコードされた場合にエネルギーに殆どまたは全く寄与しない仮想スピーカは、対応するアンビソニックデコードおよびデコードされるアンビソニックチャンネルのＨＲＴＦ処

10

20

30

40

50

理をバイパスされ得る。いくつかの実施形態において、システムは、バイパスされる任意の処理ブロックを無効にし得る。

#### 【0107】

指定方法を実行するための例示的擬似コードは、以下であり得る：

```
For each sound source, S and decode channel n
Enable[n] |= f(sourcePosition Vector3, sourceOrientation
Vector3, ListenerPosition Vector3, ListenerOrientation Vector3, VirtualSpeakerPosition[n] Vector3).
```

10

(アンビソニック／音場例)

```
For each Ambisonic Decode Channel
If (Enable[n]) {
AmbisonicDecode(n)
Virtualize(n)
}
```

(マルチチャネル例)

```
For each Channel
If (Enable[n]) {
Virtualize(n)
}
```

20

#### 【0108】

上記の擬似コードに関して、変数 `sourcePosition` は、音源の位置を指し得、`sourceOrientation` は、音源の向きを指し得、`ListenerPosition` は、ユーザ／聴者の位置を指し得、`ListenerOrientation` は、ユーザ／聴者の向きを指し得、`VirtualSpeakerPosition` は、仮想スピーカの位置を指し得、`AmbisonicDecode` は、アンビソニックデコーディングを実施する関数を指し得、`Virtualize` は、仮想化を行う関数を指し得る。

#### 【0109】

30

上記の擬似コードに関して、各音源 `S` およびデコードチャンネル `n` のために、デコードチャンネル `n` は、音源 `S` の位置、音源 `S` の向き、ユーザ／聴者の位置、ユーザ／聴者の向き、および仮想スピーカの位置等の1つ以上の因子に基づいて有効にされ得る。依然として上記の擬似コードを参照すると、各アンビソニックデコードチャンネルのために、チャンネルが、有効化される場合、システムは、`AmbisonicDecode` 関数および `Virtualize` 関数を実行し得る。

#### 【0110】

擬似コードは、各仮想スピーカのために「リングアウト」期間を提供することによって強化され得る。例えば、源がビデオフレーム中、位置において移動した場合、仮想スピーカが、その中に混合するいかなる音源ももはや有しないことが決定され得る。しかしながら、仮想スピーカのフィルタ処理に起因して、その仮想スピーカは、フィルタ処理を適切に完了させるために、続くフレームのためのアクティブスピーカである必要があり得る。

40

#### 【0111】

本開示の例は、全てのアクティブな音源を使用し、「顕著な」出力（例えば、知覚される音場に影響を及ぼすであろう出力）を有するデコードされた音場出力を決定することを含むことができる。知覚される音場に影響を及ぼすであろうアンビソニックまたは非アンビソニックマルチチャネル出力が、デコードされ得る。さらに、いくつかの実施形態において、それらの検出される出力に対応する `HRTF546` のみが、処理される。音源の数が少ないか、または、多数であるが、互いに近い場合、合成的に発生させられたアン

50

ビソニック音場または非アンビソニックマルチチャネルレンダリングのための大きなCPU節約が、あり得る。

【0112】

(源幾何学形状ベースの仮想スピーカカリング方法と低エネルギー出力検出およびカリング方法との例示的方法組み合わせ)

【0113】

いくつかの実施形態において、源幾何学形状ベースの仮想スピーカカリングと低エネルギー出力検出およびカリングとの両方が、CPU消費をさらに低減させるために、連続的に使用され得る。上で説明されるように、源幾何学形状ベースの仮想スピーカカリングは、例えば、ユーザ／聴者に対する音源の場所に基づいて、例えば、仮想スピーカ処理を選択的に無効にすることを含み得る。低エネルギー出力検出およびカリングは、例えば、音場デコーディングまたはマルチチャネル出力とHRTF処理との間に信号エネルギー／レベル検出器を設置することを含み得る。源幾何学形状ベースの仮想スピーカカリングの出力／結果は、低エネルギー出力検出およびカリングに入力され得る。

【0114】

上で説明されるシステムおよび方法に関して、システムおよび方法の要素は、適宜、1つ以上のコンピュータプロセッサ（例えば、CPUまたはDSP）によって実装されることができる。本開示は、これらの要素を実装するために使用されるコンピュータプロセッサを含むコンピュータハードウェアの任意の特定の構成に限定されない。ある場合、複数のコンピュータシステムが、上で説明されるシステムおよび方法を実装するために採用されることができる。例えば、第1のコンピュータプロセッサ（例えば、マイクロホンに結合されるウェアラブルデバイスのプロセッサ）が、入力マイクロホン信号を受信し、それらの信号の初期処理（例えば、上で説明されるもの等の信号調整および／またはセグメント化）を実施するために利用されることができる。第2の（おそらく、よりコンピュータ的に強力な）プロセッサが、次いで、それらの信号の発話セグメントに関連付けられた確率値の決定等のよりコンピュータ的に集約的な処理を実施するために利用されることができる。クラウドサーバ等の別のコンピュータデバイスが、発話認識エンジンをホストすることができ、それに入力信号が、最終的に提供される。他の好適な構成も、明白になり、本開示の範囲内である。

【0115】

開示される例は、付随の図面を参照して完全に説明されたが、種々の変更および修正が、当業者に明らかであろうことに留意されたい。例えば、1つ以上の実装の要素は、組み合わせられ、削除され、修正され、または補完され、さらなる実装を形成し得る。そのような変更および修正は、添付される請求項によって定義されるような開示される例の範囲内に含まれるとして理解されるべきである。

10

20

30

40

50

【図面】  
【図 1】

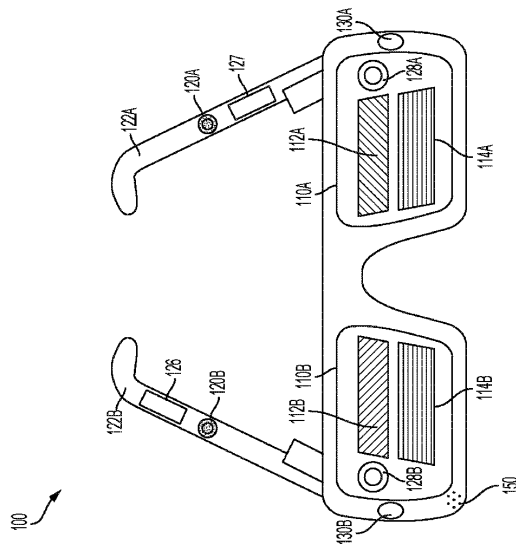


FIG. 1

【図 2】

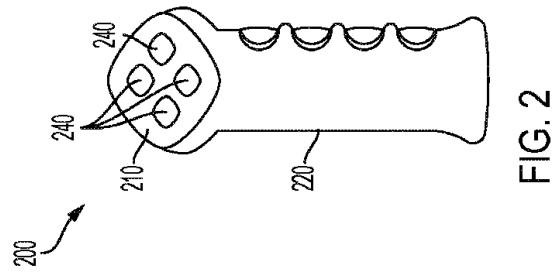


FIG. 2

【図 3】

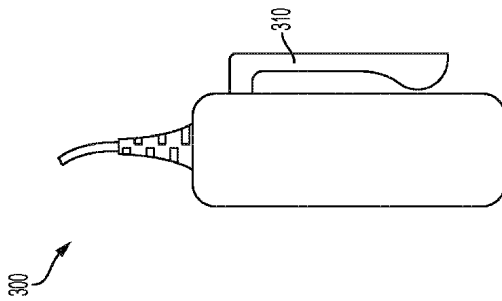


FIG. 3

【図 4】

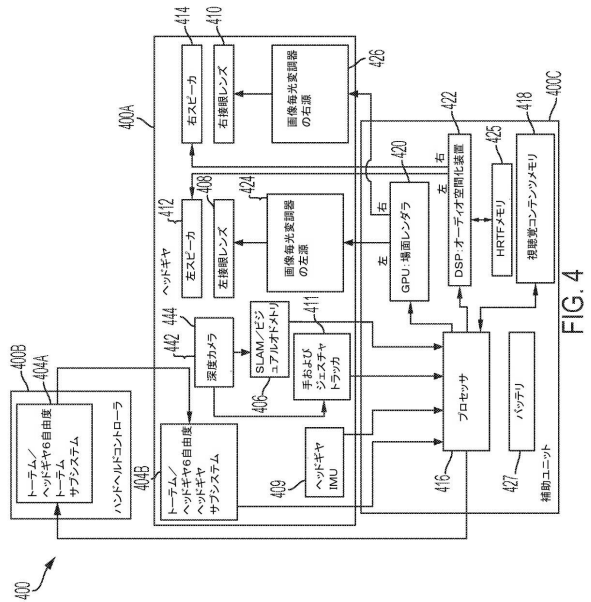


FIG. 4

【図 5 A】

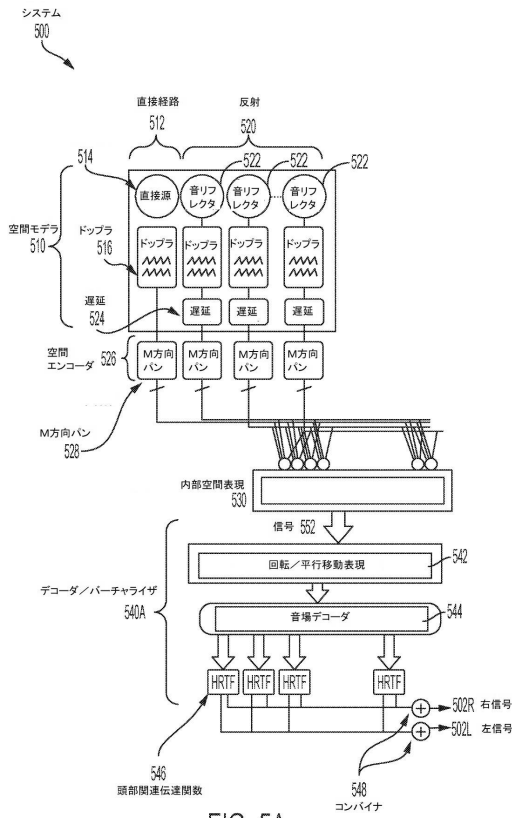


FIG. 5A

【図 5 B】

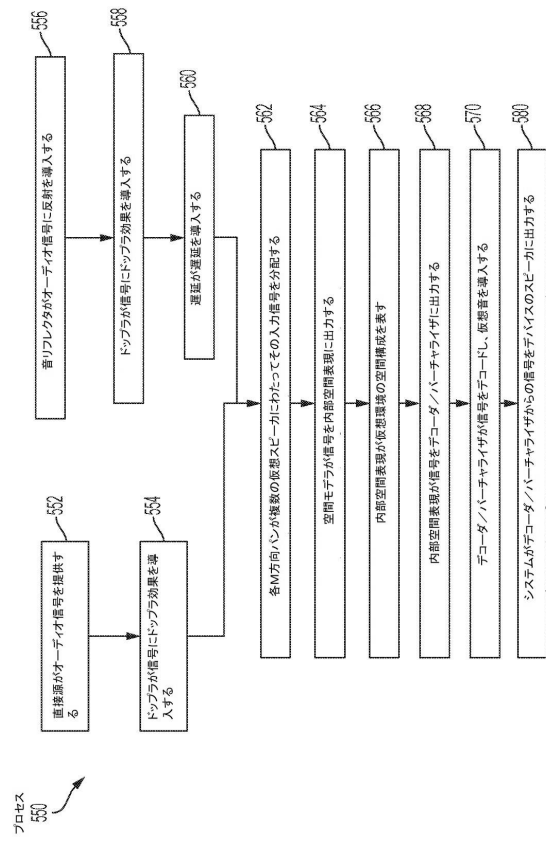


FIG. 5B

【図 5 C】

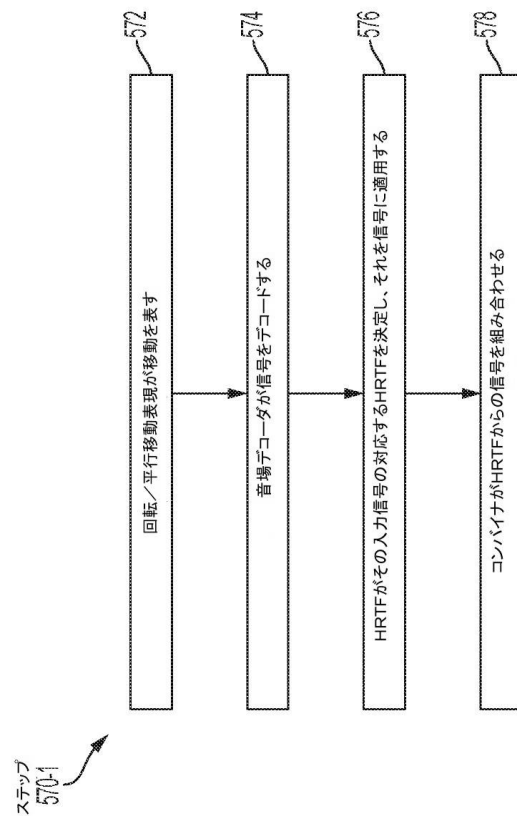


FIG. 5C

【図 6】

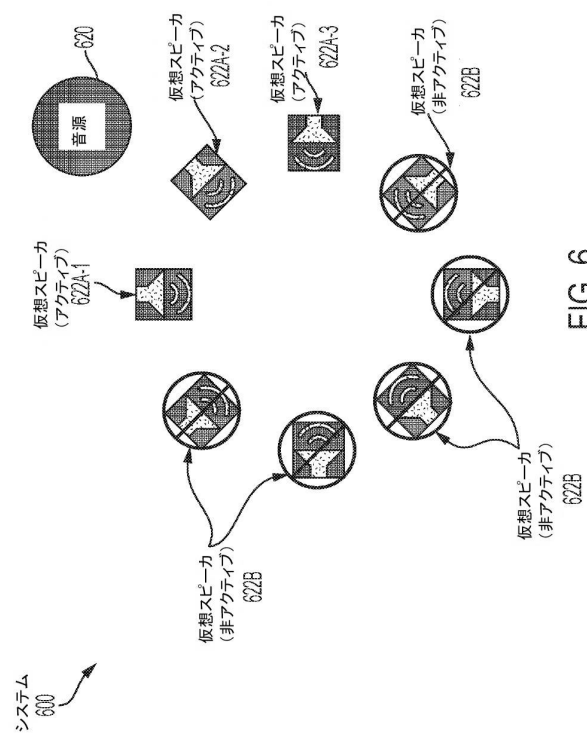
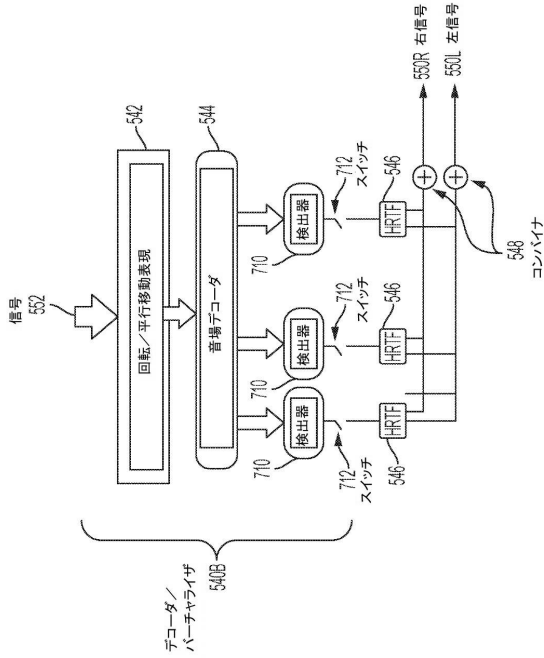
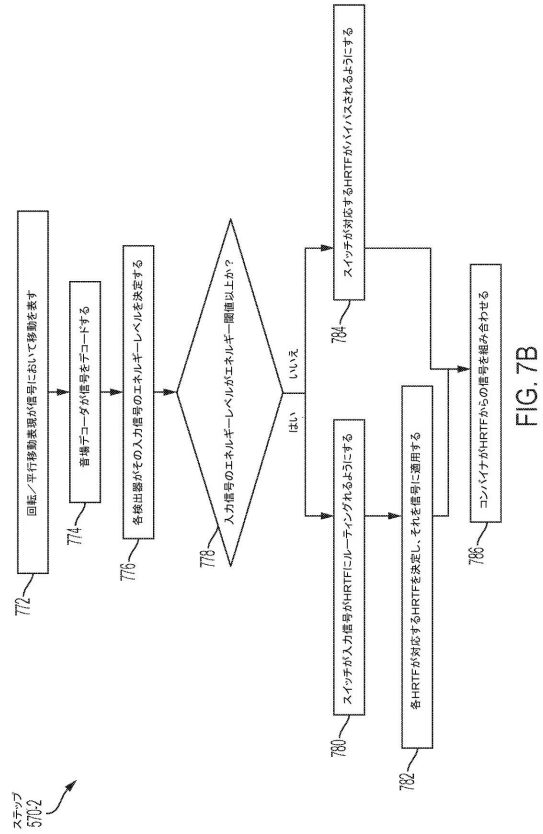


FIG. 6

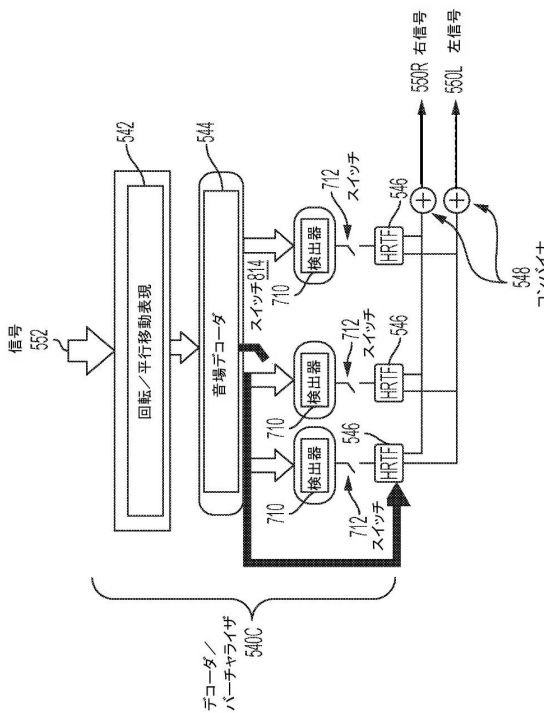
【図 7 A】



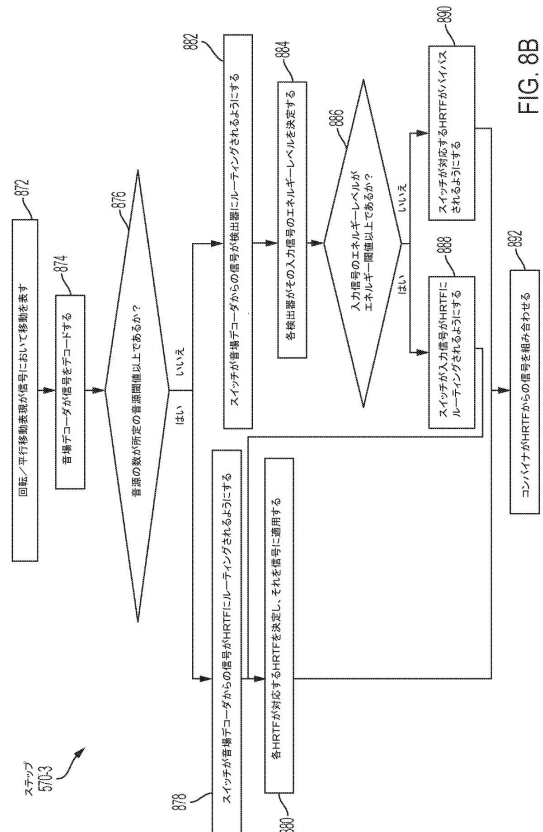
【図 7 B】



【図 8 A】



【図 8 B】



10

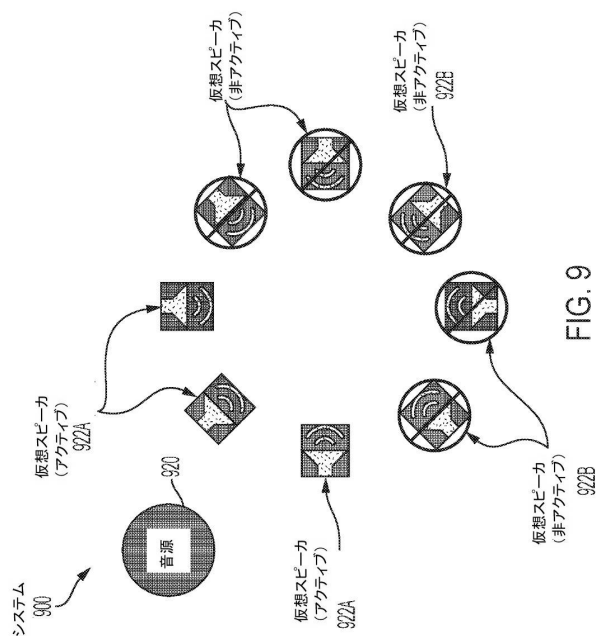
20

30

40

50

【図 9】



【図 10 A】

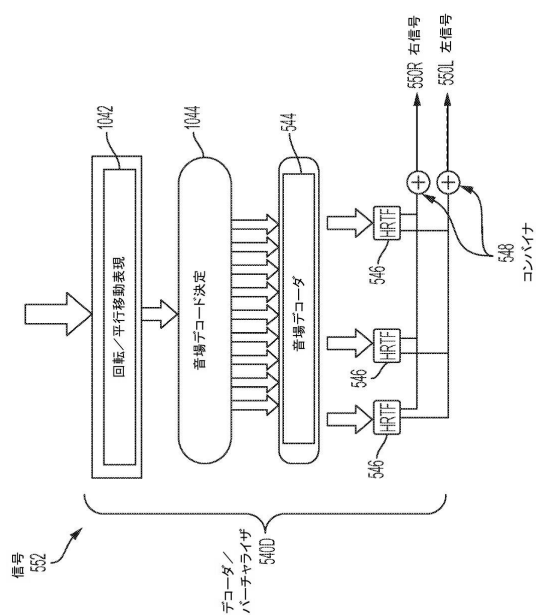


FIG. 10A

【図 10 B】

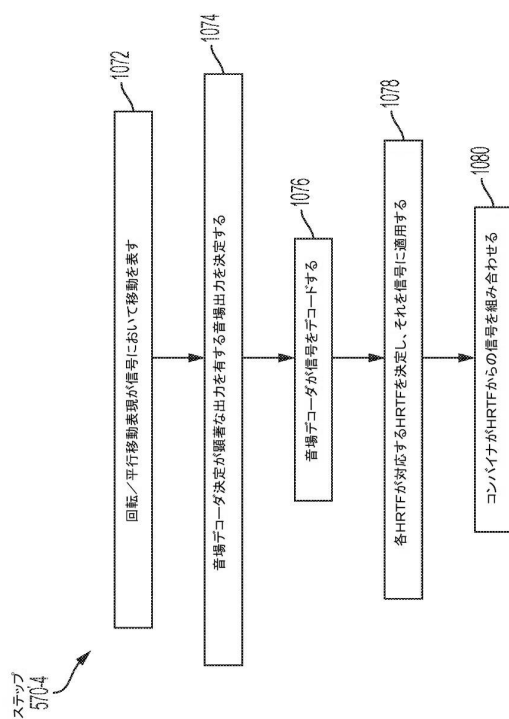


FIG. 10B

## フロントページの続き

- (74)代理人 100181641  
弁理士 石川 大輔
- (74)代理人 230113332  
弁護士 山本 健策
- (72)発明者 シュミット, ブライアン ロイド  
アメリカ合衆国 フロリダ 33322, プランテーション, ダブリュー. サンライズ ブール  
バード 7500, マジック リープ, インコーポレイテッド 気付
- (72)発明者 ディッカー, サミュエル チャールズ  
アメリカ合衆国 フロリダ 33322, プランテーション, ダブリュー. サンライズ ブール  
バード 7500, マジック リープ, インコーポレイテッド 気付
- 審査官 堀 洋介
- (56)参考文献 国際公開第2018/053047 (WO, A1)  
国際公開第2017/142759 (WO, A1)  
特開2017-188873 (JP, A)
- (58)調査した分野 (Int.Cl., DB名)  
H04S 1/00- 7/00