

(19) 日本国特許庁(JP)

## (12) 公表特許公報(A)

(11) 特許出願公表番号

特表2019-527956

(P2019-527956A)

(43) 公表日 令和1年10月3日(2019.10.3)

(51) Int.Cl.

**H04S 7/00** (2006.01)  
**H04R 3/00** (2006.01)  
**A63F 13/215** (2014.01)  
**A63F 13/54** (2014.01)  
**A63F 13/53** (2014.01)

F 1

H04S 7/00  
H04R 3/00  
A63F 13/215  
A63F 13/54  
A63F 13/53

300  
320  
13/215  
13/54  
13/53

テーマコード(参考)

5D162  
5D220

審査請求 未請求 予備審査請求 有 (全 55 頁) 最終頁に続く

(21) 出願番号 特願2019-501595 (P2019-501595)  
(86) (22) 出願日 平成29年5月25日 (2017.5.25)  
(85) 翻訳文提出日 平成31年2月26日 (2019.2.26)  
(86) 國際出願番号 PCT/US2017/034522  
(87) 國際公開番号 WO2018/013237  
(87) 國際公開日 平成30年1月18日 (2018.1.18)  
(31) 優先権主張番号 62/363,104  
(32) 優先日 平成28年7月15日 (2016.7.15)  
(33) 優先権主張国・地域又は機関  
米国(US)  
(31) 優先権主張番号 15/238,591  
(32) 優先日 平成28年8月16日 (2016.8.16)  
(33) 優先権主張国・地域又は機関  
米国(US)

(71) 出願人 595020643  
クアアルコム・インコーポレイテッド  
QUALCOMM INCORPORATED  
アメリカ合衆国、カリフォルニア州 92  
121-1714、サン・ディエゴ、モア  
ハウス・ドライブ 5775  
(74) 代理人 100108855  
弁理士 蔡田 昌俊  
(74) 代理人 100109830  
弁理士 福原 淑弘  
(74) 代理人 100158805  
弁理士 井関 守三  
(74) 代理人 100112807  
弁理士 岡田 貴志

最終頁に続く

(54) 【発明の名称】仮想、拡張、および複合現実

## (57) 【要約】

仮想音を出力するための方法は、1つまたは複数のマイクロフォンにおいて、ある環境中のオーディオ信号を検出することを含む。方法はまた、プロセッサにおいて、オーディオ信号の音源の位置を決定することと、オーディオ信号に基づいて環境の1つまたは複数の音響特性を推定することと、を含む。方法は、1つまたは複数の音響特性に基づいて環境に仮想音を挿入することをさらに含む。仮想音は、音源の位置から生成される音の1つまたは複数のオーディオ特徴を有する。

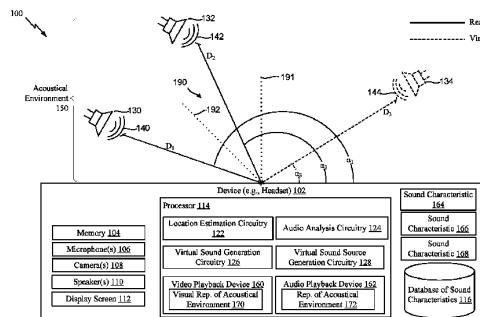


FIG. 1

**【特許請求の範囲】****【請求項 1】**

仮想音を出力するための装置であって、前記装置は、  
ある環境中のオーディオ信号を検出するように構成された 1 つまたは複数のマイクロフォンと、

前記 1 つまたは複数のマイクロフォンに結合されたプロセッサと、  
を備え、前記プロセッサは、

前記オーディオ信号の音源の位置を決定することと、

前記オーディオ信号に基づいて前記環境の 1 つまたは複数の音響特性を推定することと、

前記 1 つまたは複数の音響特性に基づいて前記環境に仮想音を挿入することと、前記仮想音は、前記音源の前記位置から生成される音の 1 つまたは複数のオーディオ特徴を有する、

を行うように構成される、装置。

**【請求項 2】**

前記プロセッサは、前記環境への前記仮想音の挿入の前に、前記仮想音を生成するよう  
にさらに構成される、請求項 1 に記載の装置。

**【請求項 3】**

前記プロセッサは、  
前記 1 つまたは複数の音響特性をメモリ中の 1 つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

を行うようにさらに構成される、請求項 1 に記載の装置。

**【請求項 4】**

前記プロセッサは、  
前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類することと、  
前記特定のオブジェクトの仮想イメージを生成することと、  
前記環境中の前記音源の前記位置において、前記仮想イメージを挿入することと、  
を行うようにさらに構成される、請求項 1 に記載の装置。

**【請求項 5】**

前記プロセッサに結合されたカメラをさらに備え、前記カメラは、前記環境をキャプチャするように構成される、請求項 1 に記載の装置。

**【請求項 6】**

前記カメラは、防犯カメラを含む、請求項 5 に記載の装置。

**【請求項 7】**

前記カメラ、前記 1 つまたは複数のマイクロフォン、および前記プロセッサは、ヘッドセットに統合される、請求項 5 に記載の装置。

**【請求項 8】**

前記ヘッドセットは、前記環境の視覚的描写を表示するように構成される、請求項 7 に記載の装置。

**【請求項 9】**

前記プロセッサは、  
前記オーディオ信号の音特性を決定することと、  
前記音特性に基づいて、前記音源が前記環境の第 1 の区域に位置するか、または前記環境の第 2 の区域に位置するかを決定することと、ここにおいて、前記第 1 の区域は、前記第 2 の区域よりも前記 1 つまたは複数のマイクロフォンに近い、  
前記オーディオ信号の到来方向を推定することと、ここにおいて、前記位置は、前記音源に関連付けられた区域および前記到来方向に基づく、

を行うようにさらに構成される、請求項 1 に記載の装置。

10

20

30

40

50

**【請求項 10】**

前記音特性は、前記オーディオ信号の少なくとも1つの残響特性を備える、請求項9に記載の装置。

**【請求項 11】**

前記少なくとも1つの残響特性は、前記オーディオ信号の直接音対残響音比を備える、請求項10に記載の装置。

**【請求項 12】**

仮想音を出力するための方法であって、前記方法は、  
1つまたは複数のマイクロフォンにおいて、ある環境中のオーディオ信号を検出することと、  
プロセッサにおいて、前記オーディオ信号の音源の位置を決定することと、  
前記オーディオ信号に基づいて前記環境の1つまたは複数の音響特性を推定することと、  
前記1つまたは複数の音響特性に基づいて前記環境に仮想音を挿入することと、前記仮想音は、前記音源の前記位置から生成される音の1つまたは複数のオーディオ特徴を有する、  
を備える、方法。

**【請求項 13】**

前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、請求項12に記載の方法。

**【請求項 14】**

前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、  
前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、  
をさらに備える、請求項12に記載の方法。

**【請求項 15】**

前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類することと、  
前記特定のオブジェクトの仮想イメージを生成することと、  
前記環境中の前記音源の前記位置において、前記仮想イメージを挿入することと、  
をさらに備える、請求項12に記載の方法。

**【請求項 16】**

カメラを介して前記環境をキャプチャすることをさらに備える、請求項12に記載の方法。

**【請求項 17】**

前記カメラは、防犯カメラを含む、請求項16に記載の方法。

**【請求項 18】**

前記カメラ、前記1つまたは複数のマイクロフォン、および前記プロセッサは、ヘッドセットに統合される、請求項16に記載の方法。

**【請求項 19】**

前記ヘッドセットにおいて前記環境の視覚的描写を表示することをさらに備える、請求項18に記載の方法。

**【請求項 20】**

前記音源の前記位置を決定することは、  
前記オーディオ信号の音特性を決定することと、  
前記音特性に基づいて、前記音源が前記環境の第1の区域に位置するか、または前記環境の第2の区域に位置するかを決定することと、ここにおいて、前記第1の区域は、前記第2の区域よりも前記1つまたは複数のマイクロフォンに近い、  
前記オーディオ信号の到来方向を推定することと、ここにおいて、前記位置は、前記音源に関連付けられた区域および前記到来方向に基づく、

を備える、請求項 1 2 に記載の方法。

【請求項 2 1】

前記音特性は、前記オーディオ信号の少なくとも 1 つの残響特性を備える、請求項 2 0 に記載の方法。

【請求項 2 2】

前記少なくとも 1 つの残響特性は、前記オーディオ信号の直接音対残響音比を備える、請求項 2 1 に記載の方法。

【請求項 2 3】

仮想音を出力するための命令を備える非一時的コンピュータ可読媒体であって、前記命令は、プロセッサによって実行されたとき、前記プロセッサに、

ある環境中のオーディオ信号の音源の位置を決定することと、前記オーディオ信号は、1 つまたは複数のマイクロフォンにおいて検出される、

前記オーディオ信号に基づいて前記環境の 1 つまたは複数の音響特性を推定することと、

前記 1 つまたは複数の音響特性に基づいて前記環境に仮想音を挿入することと、前記仮想音は、前記音源の前記位置から生成される音の 1 つまたは複数のオーディオ特徴を有する、

を備える動作を行わせる、非一時的コンピュータ可読媒体。

【請求項 2 4】

前記動作は、前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、請求項 2 3 に記載の非一時的コンピュータ可読媒体。

【請求項 2 5】

前記動作は、

前記 1 つまたは複数の音響特性をメモリ中の 1 つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、請求項 2 3 に記載の非一時的コンピュータ可読媒体。

【請求項 2 6】

前記動作は、

前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類することと、前記特定のオブジェクトの仮想イメージを生成することと、

前記環境中の前記音源の前記位置において、前記仮想イメージを挿入することと、

をさらに備える、請求項 2 3 に記載の非一時的コンピュータ可読媒体。

【請求項 2 7】

前記動作は、カメラに、前記環境をキャプチャさせることをさらに備える、請求項 2 3 に記載の非一時的コンピュータ可読媒体。

【請求項 2 8】

前記カメラは、防犯カメラを含む、請求項 2 7 に記載の非一時的コンピュータ可読媒体。

【請求項 2 9】

仮想音を出力するための装置であって、前記装置は、  
ある環境中のオーディオ信号を検出するための手段と、

前記環境中のオーディオ信号の音源の位置を決定するための手段と、

前記オーディオ信号に基づいて前記環境の 1 つまたは複数の音響特性を推定するための手段と、

前記 1 つまたは複数の音響特性に基づいて前記環境に仮想音を挿入するための手段と、前記仮想音は、前記音源の前記位置から生成される音の 1 つまたは複数のオーディオ特徴を有する、

を備える、装置。

10

20

30

40

50

**【請求項 3 0】**

前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類するための手段と、

前記特定のオブジェクトの仮想イメージを生成するための手段と、

前記環境中の前記音源の前記位置において、前記仮想イメージを挿入するための手段と、

、  
をさらに備える、請求項 2 9 に記載の装置。

**【請求項 3 1】**

音を出力するための装置であって、前記装置は、

メモリと、

10

前記メモリに結合されたプロセッサと

を備え、前記プロセッサは、

環境の 1 つまたは複数の位置において 1 つまたは複数の音響特性を決定することと、  
特定のオブジェクトを示すユーザ選択を受信することと、

特定の位置のユーザインジケーションを受信することと、

前記 1 つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の 1 つ  
または複数のオーディオ特徴を決定することと、

前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入することと、前記  
仮想音は、前記 1 つまたは複数のオーディオ特徴を有する、

を行うように構成される、装置。

20

**【請求項 3 2】**

前記プロセッサは、前記環境への前記仮想音の挿入の前に、前記仮想音を生成するよう  
にさらに構成される、請求項 3 1 に記載の装置。

**【請求項 3 3】**

前記プロセッサは、

前記 1 つまたは複数の音響特性をメモリ中の 1 つまたは複数のエントリと比較すること  
と、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮  
想音を導出することと、

を行うようにさらに構成される、請求項 3 1 に記載の装置。

30

**【請求項 3 4】**

前記プロセッサに結合されたウェアラブルセンサをさらに備え、前記ウェアラブルセン  
サは、

前記特定の位置を検出することと、

前記特定の位置の前記ユーザインジケーションを生成することと、

前記プロセッサに、前記特定の位置の前記ユーザインジケーションを送ることと

を行うように構成される、請求項 3 1 に記載の装置。

**【請求項 3 5】**

前記ウェアラブルセンサは、ユーザの手に装着されるようにさらに構成される、請求項  
3 4 に記載の装置。

40

**【請求項 3 6】**

前記プロセッサに結合された 1 つまたは複数のスピーカと、前記 1 つまたは複数のスピ  
ーカは、前記環境の異なる位置においてオーディオ信号を生成するように構成され、

前記プロセッサに結合された 1 つまたは複数のマイクロフォンと、前記 1 つまたは複数  
のマイクロフォンは、各オーディオ信号を検出するように構成され、

ここにおいて、前記プロセッサは、前記 1 つまたは複数のスピーカによって生成される  
検出されたオーディオ信号に基づいて、前記 1 つまたは複数の位置において前記 1 つまたは  
複数の音響特性を決定するように構成される、

をさらに備える、請求項 3 1 に記載の装置。

**【請求項 3 7】**

50

前記 1 つまたは複数のスピーカは、ユーザの手に装着されるようにさらに構成される、請求項 3 6 に記載の装置。

【請求項 3 8】

音を出力するための方法であって、前記方法は、  
環境の 1 つまたは複数の位置において 1 つまたは複数の音響特性を決定することと、  
特定のオブジェクトを示すユーザ選択を受信することと、  
特定の位置のユーザインジケーションを受信することと、  
前記 1 つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の 1 つまたは複数のオーディオ特徴を決定することと、

前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入することと、前記仮想音は、前記 1 つまたは複数のオーディオ特徴を有する、

を備える、方法。

【請求項 3 9】

前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、請求項 3 8 に記載の方法。

【請求項 4 0】

前記 1 つまたは複数の音響特性をメモリ中の 1 つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、請求項 3 8 に記載の方法。

【請求項 4 1】

ウェアラブルセンサを使用して前記特定の位置を検出することと、  
前記ウェアラブルセンサにおいて、前記特定の位置の前記ユーザインジケーションを生成することと、

プロセッサに、前記ウェアラブルセンサからの前記特定の位置の前記ユーザインジケーションを送ることと

をさらに備える、請求項 3 8 に記載の方法。

【請求項 4 2】

前記ウェアラブルセンサは、ユーザの手に装着されるように構成される、請求項 4 1 に記載の方法。

【請求項 4 3】

1 つまたは複数のスピーカを使用して、前記環境の異なる位置においてオーディオ信号を生成することと、

1 つまたは複数のマイクロフォンを使用して、各オーディオ信号を検出することと、  
ここにおいて、前記 1 つまたは複数の位置における前記 1 つまたは複数の音響特性は、  
前記 1 つまたは複数のスピーカによって生成される検出されたオーディオ信号に基づいて決定される、

をさらに備える、請求項 3 8 に記載の方法。

【請求項 4 4】

前記 1 つまたは複数のスピーカは、ユーザの手に装着されるように構成される、請求項 4 3 に記載の方法。

【請求項 4 5】

音を出力するための命令を備える非一時的コンピュータ可読媒体であって、前記命令は、  
プロセッサによって実行されたとき、前記プロセッサに、

環境の 1 つまたは複数の位置において 1 つまたは複数の音響特性を決定することと、

特定のオブジェクトを示すユーザ選択を受信することと、

特定の位置のユーザインジケーションを受信することと、

前記 1 つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の 1 つまたは複数のオーディオ特徴を決定することと、

10

20

30

40

50

前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入することと、前記仮想音は、前記1つまたは複数のオーディオ特徴を有する、  
を備える動作を行わせる、非一時的コンピュータ可読媒体。

【請求項46】

前記動作は、前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、請求項45に記載の非一時的コンピュータ可読媒体。

【請求項47】

前記動作は、

前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、10

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、請求項45に記載の非一時的コンピュータ可読媒体。

【請求項48】

前記特定の位置の前記ユーザインジケーションは、ウェアラブルセンサから受信される、請求項45に記載の非一時的コンピュータ可読媒体。

【請求項49】

前記ウェアラブルセンサは、ユーザの手に装着されるように構成される、請求項48に記載の非一時的コンピュータ可読媒体。

【請求項50】

前記動作は、

1つまたは複数のスピーカを使用して、前記環境の異なる位置においてオーディオ信号の生成を開始することと、20

1つまたは複数のマイクロフォンを使用して、各オーディオ信号を検出することと、

ここにおいて、前記1つまたは複数の位置における前記1つまたは複数の音響特性は、前記1つまたは複数のスピーカによって生成される検出されたオーディオ信号に基づいて決定される、

をさらに備える、請求項45に記載の非一時的コンピュータ可読媒体。

【請求項51】

前記1つまたは複数のスピーカは、ユーザの手に装着されるようにさらに構成される、請求項50に記載の非一時的コンピュータ可読媒体。30

【請求項52】

音を出力するための装置であって、前記装置は、

環境の1つまたは複数の位置において1つまたは複数の音響特性を決定するための手段と、

特定のオブジェクトを示すユーザ選択を受信するための手段と、

特定の位置のユーザインジケーションを受信するための手段と、

前記1つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定するための手段と、

前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入するための手段と、前記仮想音は、前記1つまたは複数のオーディオ特徴を有する、40

を備える、装置。

【請求項53】

前記環境への前記仮想音の挿入の前に、前記仮想音を生成するための手段をさらに備える、請求項52に記載の装置。

【請求項54】

前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較するための手段と、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出するための手段と、50

をさらに備える、請求項 5 2 に記載の装置。

【請求項 5 5】

前記特定の位置を検出することと、前記特定の位置の前記ユーザインジケーションを生成することと、前記ユーザインジケーションを受信するための前記手段に前記特定の位置の前記ユーザインジケーションを送ることと、を行うための手段をさらに備える、請求項 5 2 に記載の装置。

【請求項 5 6】

前記特定の位置を検出するための前記手段は、ユーザの手に装着されるように構成される、請求項 5 5 に記載の装置。

【請求項 5 7】

前記環境の異なる位置においてオーディオ信号を生成するための手段と、各オーディオ信号を検出するための手段と、

ここにおいて、前記 1 つまたは複数の位置における前記 1 つまたは複数の音響特性は、オーディオ信号を生成するための前記手段によって生成される検出されたオーディオ信号に基づいて決定される、

をさらに備える、請求項 5 2 に記載の装置。

【請求項 5 8】

前記オーディオ信号を生成するための前記手段は、ユーザの手に装着されるように構成される、請求項 5 7 に記載の装置。

【発明の詳細な説明】

【関連出願の相互参照】

【0 0 0 1】

[0001] 本特許出願は、「VIRTUAL, AUGMENTED, AND MIXED REALITY」と題する、2016年7月15日付で出願された米国仮特許出願第62/363,104号からの利益を主張し、かつ2016年8月16日付で出願された米国特許出願第15/238,591号の優先権を主張し、これらは、その全体が参照により明示的に組み込まれている。

【技術分野】

【0 0 0 2】

[0002] 本開示は、概して、拡張現実 (augmented reality) に関する。

【背景技術】

【0 0 0 3】

[0003] 仮想現実アプリケーションが急速に普及してきている。例えば、様々なデバイスは、ユーザが仮想環境を可聴かつ視覚的に体験することが可能な特徴を含み得る。例として、ユーザは、ビデオゲームをするためのデバイスを使用し得る。デバイスのディスプレイスクリーンは、ビデオゲームに関連付けられた仮想オブジェクトをユーザに表示し得、デバイスのスピーカは、ビデオゲームに関連付けられた音（例えば、仮想音 (virtual sounds)）をユーザに出力し得る。本明細書で使用されるとき、「仮想オブジェクト」は、仮想現実アプリケーション（例えば、ビデオゲーム）を介してユーザが見ることができるが、そうでない場合には、仮想現実アプリケーションなしにユーザが見ることができない（例えば、「実世界 (real-world)」ではユーザが見ることができない）オブジェクトに対応する。「仮想音」は、仮想現実アプリケーションを介してユーザが聞くことができるが、そうでない場合には、仮想現実アプリケーションなしにユーザが聞くことができない（例えば、「実世界」ではユーザが聞くことができない）音に対応する。

【0 0 0 4】

[0004] 特定のシナリオでは、ある人は、実世界の体験を十分に楽しむことができない可能性がある。非限定的例として、ある人が比較的遠くにいるハチドリを見ている場合に、その人は、そのハチドリが出した音（例えば、ブンブンという音 (humming noises)）を聞こうと試みるが、人とハチドリとの距離が、その人がハチドリからのブンブンという音を聞くことを妨げ得る。

【発明の概要】

10

20

30

40

50

## 【0005】

[0005] 1つの実装によると、仮想音を出力するための装置は、ある環境中のオーディオ信号を検出するように構成された1つまたは複数のマイクロフォンを含む。装置はまた、1つまたは複数のマイクロフォンに結合されたプロセッサを含む。プロセッサは、オーディオ信号の音源の位置を決定するように構成される。プロセッサは、オーディオ信号に基づいて環境の1つまたは複数の音響特性 (acoustical characteristics) を推定するようにさらに構成される。プロセッサはまた、1つまたは複数の音響特性に基づいて環境に仮想音を挿入するように構成される。仮想音は、音源の位置から生成される音の1つまたは複数のオーディオ特徴 (audio properties) を有する。

## 【0006】

[0006] 別の実装によると、仮想音を出力するための方法は、1つまたは複数のマイクロフォンにおいて、ある環境中のオーディオ信号を検出することを含む。方法はまた、プロセッサにおいて、オーディオ信号の音源の位置を決定することと、オーディオ信号に基づいて環境の1つまたは複数の音響特性を推定することと、を含む。方法は、1つまたは複数の音響特性に基づいて環境に仮想音を挿入することをさらに含む。仮想音は、音源の位置から生成される音の1つまたは複数のオーディオ特徴を有する。

10

## 【0007】

[0007] 別の実装によると、非一時的コンピュータ可読媒体は、仮想音を出力するための命令を含む。命令は、プロセッサによって実行されたとき、プロセッサに、オーディオ信号の音源の位置を決定することと、オーディオ信号に基づいて環境の1つまたは複数の音響特性を推定することと、を含む動作を行わせる。動作は、1つまたは複数の音響特性に基づいて環境に仮想音を挿入することをさらに含む。仮想音は、音源の位置から生成される音の1つまたは複数のオーディオ特徴を有する。

20

## 【0008】

[0008] 別の実装によると、仮想音を出力するための装置は、ある環境中のオーディオ信号を検出するための手段を含む。装置はまた、オーディオ信号の音源の位置を決定するための手段と、オーディオ信号に基づいて環境の1つまたは複数の音響特性を推定するための手段と、を含む。装置は、1つまたは複数の音響特性に基づいて環境に仮想音を挿入するための手段をさらに含む。仮想音は、音源の位置から生成される音の1つまたは複数のオーディオ特徴を有する。

30

## 【0009】

[0009] 別の実装によると、音を出力するための装置は、メモリと、メモリに結合されたプロセッサとを含む。プロセッサは、環境の1つまたは複数の位置において1つまたは複数の音響特性を決定するように構成される。プロセッサは、特定の位置のユーザインジケーションを受信することと、特定のオブジェクトを示すユーザ選択を受信することと、を行うようにさらに構成される。プロセッサはまた、1つまたは複数の音響特性に基づいて、特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定するように構成される。プロセッサは、特定のオブジェクトに関連付けられた仮想音を環境に挿入することを行うようにさらに構成される。仮想音は、1つまたは複数のオーディオ特徴を有する。

40

## 【0010】

[0010] 別の実装によると、音を出力するための方法は、環境の1つまたは複数の位置において1つまたは複数の音響特性を決定することを含む。方法は、特定の位置のユーザインジケーションを受信することと、特定のオブジェクトを示すユーザ選択を受信することと、をさらに含む。方法はまた、1つまたは複数の音響特性に基づいて、特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定することを含む。方法は、特定のオブジェクトに関連付けられた仮想音を環境に挿入することをさらに含む。仮想音は、1つまたは複数のオーディオ特徴を有する。

## 【0011】

[0011] 別の実装によると、非一時的コンピュータ可読媒体は、音を出力するための命

50

令を含む。命令は、プロセッサによって実行されたとき、プロセッサに、環境の1つまたは複数の位置において1つまたは複数の音響特性を決定することを含む動作を行わせる。動作は、特定の位置のユーザインジケーションを受信することと、特定のオブジェクトを示すユーザ選択を受信することと、をさらに含む。動作はまた、1つまたは複数の音響特性に基づいて、特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定することを含む。動作は、特定のオブジェクトに関連付けられた仮想音を環境に挿入することをさらに含む。仮想音は、1つまたは複数のオーディオ特徴を有する。

#### 【0 0 1 2】

[0012] 別の実装によると、音を出力するための装置は、環境の1つまたは複数の位置において1つまたは複数の音響特性を決定するための手段を含む。装置は、特定のオブジェクトを示すユーザ選択を受信するための手段と、特定の位置のユーザインジケーションを受信するための手段と、をさらに含む。装置はまた、1つまたは複数の音響特性に基づいて、特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定するための手段を含む。装置は、特定のオブジェクトに関連付けられた仮想音を環境に挿入するための手段をさらに含む。仮想音は、1つまたは複数のオーディオ特徴を有する。

10

#### 【図面の簡単な説明】

#### 【0 0 1 3】

【図1】[0013] 図1は、実世界環境に仮想音および仮想イメージを組み込むように動作可能なシステムである。

【図2】[0014] 図2は、実世界環境に仮想音および仮想イメージを組み込むように動作可能なデバイスの位置に関する、異なる区域における実世界音源と仮想音源とを描く。

20

【図3】[0015] 図3は、実世界環境に仮想音および仮想イメージを組み込むように動作可能なヘッドセットの視点からの拡張現実シーンを描く。

【図4 A】[0016] 図4 Aは、1つまたは複数の検出された音に基づいて、あるシーンに1つまたは複数の仮想オブジェクトを挿入する方法の例を描く。

【図4 B】[0017] 図4 Bは、1つまたは複数の検出された音に基づいて、あるシーンに1つまたは複数の仮想オブジェクトを挿入する方法の別の例を描く。

30

【図5】[0018] 図5は、ある環境に1つまたは複数の仮想音を挿入する例を描く。

【図6】[0019] 図6は、ある音響環境の表現を拡張する方法のフローチャートである。

【図7】[0020] 図7は、ある音響環境の表現を拡張する別の方法のフローチャートである。

40

【図8】[0021] 図8は、人工音を生成する方法のフローチャートである。

【図9】[0022] 図9は、実世界環境で仮想オブジェクトを生成する方法のフローチャートである。

【図10】[0023] 図10は、空間的にフィルタされた音を生成する方法のフローチャートである。

50

【図11】[0024] 図11は、人工音を出力する方法のフローチャートである。

【図12】[0025] 図12は、仮想音を生成する方法のフローチャートである。

【図13】[0026] 図13は、図1～図13に関して説明される技法を行うように動作可能な構成要素を含むデバイスを例示する。

#### 【発明の詳細な説明】

#### 【0 0 1 4】

[0027] 図1を参照すると、実世界環境に仮想音および仮想イメージを組み込むように動作可能なシステム100が示されている。システム100は、音響環境150中に位置するデバイス102を含む。1つの実装によると、デバイス102は、仮想現実シーン、複合現実シーン、または拡張現実シーンを生成するように構成されるヘッドセットを含む。本明細書で使用されるとき、「仮想現実シーン」は、仮想構成要素（例えば、1つまたは複数の仮想音および/または1つまたは複数の仮想オブジェクト）を含むシーンである。通常、仮想現実シーンに、実世界構成要素（例えば、実世界の音または実世界のオブジェクト）は存在しない。本明細書で使用されるとき、「拡張現実シーン」は、1つまたは

50

複数の仮想構成要素および1つまたは複数の実世界構成要素を含むシーンである。本明細書で使用されるとき、「複合現実シーン」は、1つまたは複数の仮想構成要素および1つまたは複数の実世界構成要素を含むシーンである。通常、複合現実シーンでは、仮想構成要素は、人間の感覚を使用して仮想構成要素と実世界構成要素とを区別することが難しくなるように、実世界構成要素の特徴に類似する特徴（例えば、特性）を有している。例えば、複合現実シーンは、仮想構成要素と実世界構成要素との比較的「スムーズ」かつシームレスな融合を含んでいる。

#### 【0015】

[0028] デバイス102は、メモリ104、1つまたは複数のマイクロフォン106、または1つまたは複数のカメラ108、1つまたは複数のスピーカ110、ディスプレイスクリーン112、プロセッサ114、および音特性のデータベース116を含む。デバイス102の構成要素は、互いに結合され得る。非限定的例として、1つまたは複数のマイクロフォン106は、プロセッサ114に結合され得、メモリ104はプロセッサ114に結合され得、1つまたは複数のカメラ108はプロセッサ114等に結合され得る。本明細書で使用されるとき、「結合される」は、「通信可能に結合される」、「電気的に結合される」、または「物理的に結合される」、およびそれらの組み合わせを含み得る。2つのデバイス（または構成要素）は、1つまたは複数の他のデバイス、構成要素、ワイヤ、バス、ネットワーク（例えば、ワイヤードネットワーク、ワイヤレスネットワーク、あるいはそれらの組み合わせ）などを介して、直接または間接的に結合され得る（例えば、通信可能に結合されるか、電気的に結合されるか、または物理的に結合される）。電気的に結合された2つのデバイス（または構成要素）は、同じデバイスまたは異なるデバイス中に含まれ得、非限定的例として、素子（electronics）、1つまたは複数のコネクタ、または誘導結合を介して結合され得る。いくつかの実装では、電気通信などの通信可能に結合された2つのデバイス（または構成要素）は、1つまたは複数のワイヤ、バス、ネットワークなどを介して、直接または間接的に電気信号（デジタル信号またはアナログ信号）を送受信し得る。

10

20

30

#### 【0016】

[0029] デバイス102に含まれる構成要素は、例示のためだけのものであり、限定されるものであると解釈されない。1つの実装によると、デバイス102は、追加の（または、より少ない）構成要素を含み得る。非限定的例として、デバイス102はまた、ボリュームアジャスター、1つまたは複数のバイオフィードバックセンサまたは他のセンサ（例えば、加速度計）、1つまたは複数の通信コンポーネント（例えば、ワイヤレス通信のための無線周波数（RF）トランシーバ）などを含み得る。プロセッサ114は、位置推定回路122、オーディオ分析回路124、仮想音生成回路126、仮想音源生成回路128、ビデオ再生デバイス160、およびオーディオ再生デバイス162を含む。プロセッサ114に含まれる構成要素は、例示のためだけのものであり、限定されるものであると解釈されない。いくつかの実装によると、プロセッサ114中の2つ以上の構成要素は、単一の構成要素または単一の処理ユニットに組み込まれ得る。

#### 【0017】

[0030] 音響環境150は、「実際の」（例えば、実世界の）オブジェクトおよび実際の音を含む。例えば、音響環境150は、音源130および音源132を含む。各音源130および132は、音を生成するオブジェクトであり得る。音源の非限定的例は、動物、人、自動車、機械などである。音源130は、角度位置（ $\theta_1$ ）に従った、デバイス102からの距離（ $D_1$ ）にあり得る。音源132は、角度位置（ $\theta_2$ ）に従った、デバイス102からの距離（ $D_2$ ）にあり得る。いくつかの実装では、角度位置がある角度座標（例えば、115度）を含むが、他の実装では、角度位置は、第1の角度座標191（例えば、90度）と第2の角度座標192（例えば、135度）との間の領域190などの2つの角度座標間の領域を含み得る。音源130は、オーディオ信号140を生成するように構成され得、音源132は、オーディオ信号142を生成するように構成され得る。下記で説明されるように、各オーディオ信号140、142は、1つまたは複数の音特性

40

50

を有し得る。

【0018】

[0031] プロセッサ 114 の位置推定回路 122 は、音源 130、132 の視覚的パラメータを決定するように構成され得る。例として、1つまたは複数のカメラ 108 は、音響環境 150 の視覚的表現 170 をキャプチャし得る。1つまたは複数のカメラ 108 は、オーディオキャプチャデバイスに近接し（例えば、1つまたは複数のマイクロフォン 106 に近接し）得る。始めに（例えば、下記で説明されるように、複合現実アプリケーションの追加の前に）、音響環境 150 の視覚的表現 170 は、音響環境 150 中に実オブジェクト（real object）を描き得る。例えば、音響環境 150 の視覚的表現 170 は、始めに音源 130 および音源 132 を描き得る。

10

【0019】

[0032] ビデオ再生デバイス 160 は、1つまたは複数のカメラ 108 が音響環境 150 の視覚的表現 170 をキャプチャしたことに応答して、音響環境 150 の視覚的表現 170 を処理（例えば、レンダリング）し得る。音響環境 150 の視覚的表現 170 がビデオ再生デバイス 160 によって処理された後、位置推定回路 122 は、レンダリングに基づいて、デバイス 102 からの音源 130 の距離（ $D_1$ ）と、音源 130 の角度位置（ $\theta_1$ ）とを決定するための位置推定技法を使用し得る。位置推定回路 122 はまた、レンダリングに基づいて、デバイス 102 からの音源 130 の距離（ $D_2$ ）と、音源 130 の角度位置（ $\theta_2$ ）とを決定するための位置推定技法を使用し得る。ビデオ再生デバイス 160 は、ディスプレイスクリーン 112 上に音響環境 150 の視覚的表現 170 を表示し得る。ディスプレイスクリーン 112 上に音響環境 150 の視覚的表現 170 を表示することは、図 3 に関してより詳細に説明される。

20

【0020】

[0033] プロセッサ 114 のオーディオ分析回路 124 は、オーディオ信号 140、142 の音特性 164、166 を決定するように構成され得る。例として、1つまたは複数のマイクロフォン 106 は、オーディオ信号 140、142 を検出するように構成され得る。オーディオ分析回路 124 は、1つまたは複数のマイクロフォン 106 がオーディオ信号 140 を検出したことに応答して、オーディオ信号 140 の音特性 164 を決定し得る。例えば、オーディオ分析回路 124 は、オーディオ信号 140 の少なくとも 1 つの残響特性を決定し得る。1つの実装によると、少なくとも 1 つの残響特性は、オーディオ信号 140 の直接音対残響音比（D R R : direct-to-reverberation ratio）を含み得る。オーディオ分析回路 124 は、音特性のデータベース 116 に（位置推定回路 122 によって決定された音源 130 の対応する視覚的パラメータに沿って）オーディオ信号 140 の音特性 164 を記憶し得る。音特性のデータベース 116 は、音特性と音源位置情報とを関連付ける。音特性 164 が残響特性として説明されているが、他の実装では、音特性 164 は、1つまたは複数の他の残響特性、室内インパルス応答（R I R : room impulse response）、頭部伝達関数（H R T F : head-related transfer function）、1つまたは複数の他の特性、あるいはそれらの組み合わせを含み得る。

30

【0021】

[0034] 同様の例では、オーディオ分析回路 124 は、1つまたは複数のマイクロフォン 106 がオーディオ信号 142 を検出したことに応答して、オーディオ信号 142 の音特性 166 を決定し得る。例えば、オーディオ分析回路 124 は、オーディオ信号 142 の少なくとも 1 つの残響特性を決定し得る。1つの実装によると、少なくとも 1 つの残響特性は、オーディオ信号 142 の D R R を含み得る。オーディオ分析回路 124 は、音特性のデータベース 116 に（位置推定回路 122 によって決定された音源 132 の対応する視覚的パラメータに沿って）オーディオ信号 142 の音特性 166 を記憶し得る。

40

【0022】

[0035] 1つの実装によると、音特性 164、166 は、音響環境 150 の特性に基づいて影響され得る。プロセッサ 114 は、音響環境 150 が屋内環境、屋外環境、部分的に囲まれた環境（partially enclosed environment）などに対応するかどうかを決定し得

50

る。音特性 164、166 の反響構成要素は、音響環境 150 の特性によって、音源 130、132 がエレベータ内にあるときと比較すると、音源 130、132 が野原 (open field) にあるときなどに、変化し得る。

#### 【0023】

[0036] オーディオ再生デバイス 162 は、音響環境 150 のオーディオ表現 172 を生成し得る。始めに ( 例えれば、下記で説明されるように、複合現実アプリケーションの追加の前に ) 、音響環境 150 のオーディオ表現 172 は、オーディオ信号 140 に関連付けられた音と、オーディオ信号 142 に関連付けられた音とを含み得る。音響環境 150 のオーディオ表現 172 は、1つまたは複数のスピーカ 110 を使用して出力され得る。例えば、オーディオ再生デバイス 162 は、音響環境 150 のオーディオ表現 172 を、1つまたは複数のスピーカ 110 ( 例えれば、ヘッドフォン ) を使用してユーザのヘッドセット ( 例えれば、デバイス 102 ) に提供し得る。

10

#### 【0024】

[0037] オーディオ表現 172 においてオーディオ信号 140 に関連付けられた音は、角度位置 (  $\alpha_1$  ) に基づく到來方向を有するように、かつ距離 (  $D_1$  ) に基づくボリューム ( 例えれば、音レベル ) を有するように生成され得る。例えれば、距離 (  $D_1$  ) が比較的長い場合、オーディオ信号 140 に関連付けられた音のボリュームは比較的低くなり得る。そのため、オーディオ再生デバイス 162 は、音源 130 と音特性 164 との視覚的パラメータに基づいて、オーディオ信号 140 に関連付けられた音を生成し得る。他の実装では、オーディオ再生デバイス 162 は、オーディオ信号 140 を、オーディオ表現 172 においてオーディオ信号 140 に関連付けられた音を生成するために 1つまたは複数のマイクロフォン 106 によって検出されたときに使用し得る。例えれば、オーディオ再生デバイスは、オーディオ信号 140 を 1つまたは複数のマイクロフォン 106 によって検出されたときに「再生する」ことによって、オーディオ信号 140 に関連付けられた音を生成し得る。オーディオ表現 172 中のオーディオ信号 142 に関連付けられた音は、角度位置 (  $\alpha_2$  ) に基づく到來方向を有するように、かつ距離 (  $D_2$  ) に基づくボリューム ( 例えれば、音レベル ) を有するように生成され得る。そのため、オーディオ再生デバイス 162 は、音源 132 と音特性 166 との視覚的パラメータに基づいて、オーディオ信号 142 に関連付けられた音を生成し得る。他の実装では、オーディオ再生デバイス 162 は、オーディオ信号 142 を 1つまたは複数のマイクロフォン 106 によって検出されたときに「再生する」ことによって、オーディオ信号 142 に関連付けられた音を生成し得る。

20

#### 【0025】

[0038] 1つの実装によると、オーディオ信号 140、142 の 1つまたは複数は、オーディオデバイスのユーザが開始した再生 ( user-initiated playback ) に基づいて生成され得る。非限定的例として、音源 130 はラウドスピーカを含み得、ユーザが開始した再生は、そのラウドスピーカからのオーディオ信号 140 の生成を含み得る。例として、図 5 中の第 1 の例 500 を参照すると、ラウドスピーカ 502 は、ユーザの手の上に置かれ得る。ラウドスピーカ 502 は、ラウドスピーカ 502 を使用して、環境の異なる位置においてオーディオ信号を生成し得る。ヘッドセットは、ラウドスピーカ 502 によって生成される音に基づいて、音響特性 ( 例えれば、室内インパルス応答 ( RIR ) および頭部伝達関数 ( HRTF ) ) を決定し得る。例えれば、ヘッドセットは、ラウドスピーカ 502 によって生成された各オーディオ信号を検出するように構成された 1つまたは複数のマイクロフォンを含み得る。ユーザが自分の手を動かし得るので、ラウドスピーカ 502 は、音響特性を決定するために環境の異なる位置に動き得る。ヘッドセットもまた、環境内の仮想音を生成するために、RIR および HRTF に基づいてフィルタを更新し得る。

40

#### 【0026】

[0039] 環境の異なる位置においてヘッドセットが音響特性を決定した後、ユーザは、特定のオブジェクトを示すユーザ選択をヘッドセットに提供し得る。例として、第 2 の例 520 を参照すると、ユーザは、ピアノ 504 のユーザ選択をヘッドセットに提供し得る。ユーザはまた、環境の特定の位置のユーザインジケーションを提供し得る。例えれば、ユ

50

ーザは、ユーザの指にウェアラブルセンサを装着し得る。ウェアラブルセンサは、特定の位置を示し得る。ヘッドセットは、異なる位置における音響特性に基づいて、特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定し得る。例えば、ヘッドセットは、ウェアラブルセンサによって示された位置に関連付けられたオーディオ特徴を決定するために、R I R および H R T F を使用し得る。ヘッドセットは、特定のオブジェクト（例えば、ピアノ 5 0 4）に関連付けられた仮想音（例えば、ピアノ音）を生成し得る。仮想音は1つまたは複数のオーディオ特徴を有し得、ヘッドセットは、仮想音に基づいて音信号を出力し得る。

#### 【0027】

[0040] 追加として、ユーザは、別の特定のオブジェクトを示すユーザ選択をヘッドセットに提供し得る。例として、第2の例 5 2 0 を参照すると、ユーザは、バイオリン 5 0 6 のユーザ選択をヘッドセットに提供し得る。ユーザはまた、ウェアラブルセンサを介して第2の特定の位置のユーザインジケーションを提供し得る。ヘッドセットは、異なる位置における音響特性に基づいて、第2の特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定し得る。ヘッドセットは、バイオリン 5 0 6 に関連付けられた第2の仮想音（例えば、バイオリン音）を生成し得る。仮想音は1つまたは複数のオーディオ特徴を有し得、ヘッドセットは、仮想音に基づいて音信号を出力し得る。 10

#### 【0028】

[0041] 1つの実装によると、特定の時間の量が経過した（例えば、フィルタ音響を生成するのに十分な時間が経過した）場合に、対応する仮想オブジェクトの位置にユーザの手が無くとも、オブジェクト音は残り（例えば、継続的にプレイされ）得る。例えば、ヘッドセットがユーザの手の場所で特性（例えば、残響特性、R I R、および H R T F）を決定した場合、ユーザが自分の手を動かした後の場所で仮想音が継続的にプレイされ得る。そのため、図1に戻って参照すると、オーディオ分析回路 1 2 4 は、ユーザの手の場所に関連付けられたオーディオ再生に基づいて、オーディオ信号 1 4 0 の音特性 1 6 4 を生成し得る。 20

#### 【0029】

[0042] 1つの実装によると、オーディオ信号 1 4 0、1 4 2 の1つまたは複数は、オーディオデバイスのユーザが開始した再生なしに生成され得る。非限定的例として、音源 1 3 2 は、オーディオ信号 1 4 2 を生成する動物（またはユーザの管理下にない別のオブジェクト）を含み得る。 30

#### 【0030】

[0043] プロセッサ 1 1 4 は、1つまたは複数の仮想構成要素（例えば、仮想オブジェクトおよび／または仮想音）を生成することと、1つまたは複数の仮想構成要素を、複合または拡張視覚的シーンを生成するためにディスプレイスクリーン 1 1 2 で表示される音響環境 1 5 0 の視覚的表現 1 7 0 に適用することと、1つまたは複数の仮想構成要素を、複合または拡張オーディオを生成するために1つまたは複数のスピーカ 1 1 0 に出力され得た、音響環境 1 5 0 のオーディオ表現 1 7 2 に適用することと、を行うように構成され得る。例えば、仮想音生成回路 1 2 6 は、仮想音 1 4 4 を生成するように構成され得、仮想音源生成回路 1 2 8 は、仮想音源 1 3 4 を生成するように構成され得る。 40

#### 【0031】

[0044] 例として、仮想音源生成回路 1 2 8 は、仮想音源 1 3 4 を生成し得、ビデオ再生デバイス 1 6 0 は、視覚的表現 1 7 0 に仮想音源 1 3 4 を挿入することによって、音響環境 1 5 0 の視覚的表現 1 7 0 を修正するように構成され得る。修正された視覚的表現 1 7 0 は、図3に関してさらに詳細に説明されるように、ディスプレイスクリーン 1 1 2 に表示され得る。修正された視覚的表現 1 7 0 に基づいて、位置推定回路 1 2 2 は、仮想音源 1 3 4 の視覚的パラメータを決定するように構成され得る。例として、位置推定回路 1 2 2 は、デバイス 1 0 2 からの仮想音源 1 3 4 の距離 ( $D_3$ ) と、仮想音源 1 3 4 の角度位置 ( $\theta_3$ ) とに対応する視覚的パラメータを決定するための位置推定技法を使用し得る。 50

## 【0032】

[0045] 仮想音生成回路126は、仮想音144が仮想音源134から来たことをデバイス102のユーザによって感知され得るような方法で、それぞれ、オーディオ信号140、142の音特性164、166のうちの1つまたは複数に基づいて仮想音144を生成するように構成され得る。例として、仮想音源134は鳥であり得、仮想音144は、鳥のさえずり音であり得る。鳥のさえずり音は、音データベースから導出され、かつ仮想音144が仮想音源134から生じているというユーザの認識を高めるために、仮想音生成回路126によって修正され得る。例えば、仮想音生成回路126は、オーディオ信号140、142の残響特性（例えば、音特性のデータベース116に記憶された音特性164、166）に基づいて、仮想音144の1つまたは複数の残響特性を決定し得る。例として、仮想音生成回路126は、仮想音源134の位置（例えば、距離( $D_3$ )および角度位置( $\theta_3$ )）を、音特性のデータベース中の音特性164、166に関連付けられた視覚的パラメータと比較し得る。距離( $D_2$ )が距離( $D_3$ )に実質的に類似する（substantially similar）と仮想音生成回路126が決定した場合、仮想音生成回路126は、オーディオ信号142の残響特性に実質的に類似する仮想音144についての残響特性を生成し得る。10

## 【0033】

[0046] 図2に関してより詳細に説明されるように、仮想音生成回路126は、音響環境150の第1の区域または音響環境150の第2の区域などの1つまたは複数の音響区域に音源130、132が位置するかどうかを決定し得る。仮想音144の音特性168は、仮想音源134が音源130と同じ区域に位置する場合、オーディオ信号140の音特性164に実質的に類似し得る。仮想音144の音特性168は、仮想音源134が音源132と同じ区域に位置する場合、オーディオ信号142の音特性166に実質的に類似し得る。仮想音144の音特性168は、仮想音源134が音源130とは異なる区域に位置する場合、オーディオ信号140の音特性164とは異なり得る。仮想音144の音特性168は、仮想音源134が音源132とは異なる区域に位置する場合、オーディオ信号142の音特性166とは異なり得る。20

## 【0034】

[0047] 仮想音生成回路126はまた、仮想音144に関連付けられた仮想音源134の位置に基づいて、仮想音144の到来方向を決定するように構成され得る。例えば、仮想音生成回路126は、仮想音源134の角度位置( $\theta_3$ )を（位置推定回路122から）決定し得る。角度位置( $\theta_3$ )に基づいて、仮想音生成回路126は、デバイス102のユーザが、仮想音源134の方向から仮想音144が来ているかのように仮想音144を聞くように、仮想音144をパン（pan）し得る。そのため、仮想音生成回路126は、1つまたは複数の残響特性（例えば、音特性168）と到来方向とにに基づいて、仮想音144を生成し得る。30

## 【0035】

[0048] オーディオ再生デバイス162は、オーディオ表現172に仮想音源134を挿入することによって、音響環境150のオーディオ表現172を修正するように構成され得る。音響環境150の修正されたオーディオ表現172は、1つまたは複数のスピーカ110において出力され得る。例えば、オーディオ再生デバイス162は、ヘッドセットの1つまたは複数のラウドスピーカにおいて（仮想音144に基づいて）音信号を出力し得る。40

## 【0036】

[0049] 上述されるように、鳥のさえずり音は、音データベースから導出され得る。音データベースから導出された鳥のさえずり音は、オーディオ信号のデジタル表現を含み得る。仮想音生成回路126は、空間的にフィルタされたオーディオファイルを生成するために、音特性164、166に基づいてオーディオ信号のデジタル表現を空間的にフィルタするように構成され得る。仮想音源生成回路128は、空間的にフィルタされたオーディオファイルに基づいて空間的にフィルタされたオーディオ信号を生成すること、および50

1つまたは複数のスピーカ 110 に空間的にフィルタされたオーディオ信号を送ることを行いうように構成され得る。1つまたは複数のスピーカ 110 は、空間的にフィルタされたオーディオ信号を、空間的にフィルタされた音として投影 (project) するように構成され得る。空間的にフィルタされた音は、仮想音 144 を含み得る。1つの実装によると、仮想音 144 は、コンピュータによって生成される音を含む。

#### 【0037】

[0050] 1つの実装によると、1つまたは複数のカメラ 108 は、音響環境 150 の視覚的描写などの視覚的シーンをキャプチャするように構成され得る。仮想音源生成回路 128 が音響環境 150 の視覚的表現 170 ( 例えば、視覚的シーン ) に仮想音源 134 のイメージを挿入した後、位置推定回路 122 は、視覚的シーンにおける仮想音源 134 の位置を決定し得る。位置は、1つまたは複数のカメラ 108 からの仮想音源 134 の距離 (  $D_3$  ) と、1つまたは複数のカメラ 108 に関する仮想音源 134 の角度位置 (  $\theta_3$  ) とを示し得る。1つの実装によると、位置推定回路 122 は、視覚的シーンに対応し得る深度マップに少なくとも部分的に基づいて距離 (  $D_3$  ) を決定し得る。上述されるように、図 2 に関して、仮想音生成回路 126 は、仮想音源 134 の位置に基づいて仮想音 144 の1つまたは複数の音特性 ( 例えば、音特性 168 ) を決定し得る。仮想音生成回路 126 は、1つまたは複数の音特性に基づいて仮想音 144 を生成し得、オーディオ再生デバイス 162 は、仮想音 144 を含むように音響環境 150 の音響表現 172 を修正することによって、仮想音 144 に基づいて ( 1つまたは複数のスピーカ 110 において ) 音信号を出力し得る。

10

20

30

40

50

#### 【0038】

[0051] 図 1 のシステム 100 は、ディスプレイスクリーン 112 において視覚的複合現実シーンを生成するために、視覚的オブジェクトが音響環境 150 の視覚的表現 170 に挿入されることを可能にする。例えば、音響環境 150 に存在していないオブジェクトが、ユーザの満足 ( user enjoyment ) を高めるために、音響環境 150 の視覚的表現 170 に事実上 ( virtually ) 挿入され得る。追加として、システム 100 は、1つまたは複数のスピーカ 110 において、複合現実を生成するために仮想音が音響環境のオーディオ表現 172 に挿入されることを可能にし得る。1つの実装によると、仮想音は、仮想音源 134 などの仮想オブジェクトに追加され得る。1つの実装によると、仮想音は、音源 130 および / または音源 132 などの実世界オブジェクトに追加され得る。実世界オブジェクトに仮想音を追加することは、ユーザが比較的遠くにあるオブジェクトから音を「聞く」ことを ( 例えば、そうでない場合には、仮想現実アプリケーションなしにユーザに聞こえない可能性があるオブジェクトからのサウンドを聞くことを ) 可能にし得る。

#### 【0039】

[0052] 図 2 は、図 1 のデバイス 102 の位置に関する、異なる区域での実世界音源と仮想音源との例を描く。第 1 の区域 202 と第 2 の区域 204 とを含む複数の区域を含むように図 1 の音響環境 150 が例示されている。他の実装によると、音響環境 150 は、2 よりも多い区域を含み得る。区域 202、204 は、デバイス 102 のマイクロフォン ( 複数を含む ) 106 に位置するかまたはその近くに位置する、中心点を有する同心円を含み得る。例えば、マイクロフォン ( 複数を含む ) 106 は、マイクロフォンの円形アレイを含み得、ここで、各マイクロフォンは、異なる方向でオーディオをキャプチャするように設置されている。第 1 の区域 202 は、第 2 の区域 204 よりもデバイス 102 に近い。図 2 は同心円の形で 2 つの区域 202、204 を描いているが、本明細書で説明される技法は、異なる形状を有する区域を使用して利用可能であり得る。非限定的例として、区域 202、204 の各々は、デバイス 102 に位置する中心点を有する長方形部分を含み得る。

#### 【0040】

[0053] デバイス 102 のプロセッサ 114 は、特定の音源が第 1 の区域 202 に位置するかまたは第 2 の区域 204 に位置するかを、角度位置および距離に基づいて決定するように構成され得る。例えば、オーディオ分析回路 124 は、第 1 の音源 130 とデバイ

ス 1 0 2 ( 例えば、マイクロフォン ( 複数を含む ) 1 0 6 ) との間の距離 (  $D_1$  ) に基づき、かつ ( 例えば、マイクロフォン ( 複数を含む ) 1 0 6 に関連する ) 第 1 の音源 1 3 0 の第 1 の角度位置に基づく、第 1 の区域 2 0 2 に第 1 の音源 1 3 0 が位置すると決定し得る。同様の方法で、プロセッサ 1 1 4 は、第 2 の音源 1 3 2 とデバイス 1 0 2 との間の距離 (  $D_2$  ) に基づき、かつ ( 例えば、マイクロフォン ( 複数を含む ) 1 0 6 に関連する ) 第 2 の音源 1 3 2 の第 2 の角度位置に基づく、第 2 の区域 2 0 4 に第 2 の音源 1 3 2 が位置すると決定し得る。図 1 に関連して説明されるように、プロセッサ 1 1 4 は、第 1 のオーディオ信号 1 4 0 に対応する第 1 の音特性 1 6 4 と、第 2 のオーディオ信号 1 4 2 に対応する第 2 の音特性 1 6 6 とを決定し得る。

【 0 0 4 1 】

[0054] プロセッサ 1 1 4 は、仮想音源 1 3 4 とデバイス 1 0 2 との間の距離 (  $D_3$  ) に基づき、かつ仮想音源 1 3 4 の第 3 の角度位置 (  $\theta_3$  ) に基づく、第 2 の区域 2 0 4 に位置するソース ( 仮想音源 1 3 4 ) から仮想音信号 1 4 4 が発生すると決定し得る。仮想音信号 1 4 4 を生成するために、プロセッサ 1 1 4 は、仮想音のデータベースから仮想ソース ( 例えば、鳥のさえずり ) に対応する音信号を導出し得、第 2 の区域 2 0 4 からの仮想音 1 4 4 の音特性 ( 複数を含む ) が第 2 の音源 1 3 2 からの実世界音 ( オーディオ信号 1 4 2 ) の音特性 ( 複数を含む ) をまねるように、導出された音信号に第 2 の音特性 1 6 6 を適用し得る。代替として、仮想音源 1 3 4 が第 2 の区域 2 0 4 中ではなく第 1 の区域 2 0 2 中にあると決定された場合、プロセッサ 1 1 4 は、代わりに、第 1 の区域 2 0 4 からの仮想音 1 4 4 の音特性 ( 複数を含む ) が音源 1 3 0 からの実世界音 ( オーディオ信号 1 4 0 ) の音特性 ( 複数を含む ) をまねるように、導出された音信号に第 1 の音特性 1 6 4 を適用し得る。

【 0 0 4 2 】

[0055] 同じ区域中の 1 つまたは複数の他の音源の測定された ( 例えば、感知された、計算された、推定されたなどの ) 音特性に基づいて仮想音の 1 つまたは複数の音特性を選択することによって、仮想音は、仮想音の音特性を決定するための別の技法を使用することと比較すると、低減された計算の複雑さとともに、ユーザによってよりリアルに感知され得る。例えば、仮想音の音特性の決定は、仮想音の距離および到来方向に基づいて、および音響環境の 1 つまたは複数の特徴にさらにに基づいて ( 例えば、天井または壁からの距離、およびそれらがあろうとなかろうと、反射性あるいは吸収性の構造または材料からの距離、およびそれらがあろうとなかろうと ) 、音データの記憶されたテーブルにアクセスすることによって、他の音信号とは関係なく行われ得る。類似する空間位置 ( 例えば、同じ区域 ) から来る音が類似する音特性 ( 例えば、同じ残響特性 ) を有するという近似性 ( approximation ) および実世界音の測定された音特性を使用することによって、上述されたテーブルベースのアプローチ ( table-based approach ) と比較して、よりリアルな仮想音が生成され得る。

【 0 0 4 3 】

[0056] 図 3 は、実世界環境に仮想音および仮想イメージを組み込むように動作可能なヘッドセットの視点からの複合現実シーンを描く。デバイス 1 0 2 は、左ディスプレイスクリーン 3 0 2 、右ディスプレイスクリーン 3 0 3 、マイクロフォン 3 1 0 、3 1 1 、3 1 2 、および 3 1 3 ( 例えば、マイクロフォンのアレイ ) 、ラウドスピーカ 3 2 0 ~ 3 2 3 ( 例えば、スピーカのアレイ ) を有するヘッドセットとしての透視図をユーザ側から描かれている。ディスプレイスクリーン 3 0 2 ~ 3 0 3 は、一括して、図 1 のディスプレイスクリーン 1 1 2 に対応し、マイクロフォン 3 1 0 ~ 3 1 3 は、図 1 のマイクロフォン 1 0 6 に対応し、ラウドスピーカ 3 2 0 ~ 3 2 3 は、図 1 のスピーカ 1 1 0 に対応し得る。デバイス 1 0 2 は、第 2 の木 3 4 2 よりもデバイス 1 0 2 に近い第 1 の木 3 4 0 を含む環境にある。その環境はまた、第 1 の音源 3 3 0 と第 2 の音源 3 3 2 とを含む。

【 0 0 4 4 】

[0057] デバイス 1 0 2 は、デバイス 1 0 2 の装着者に、仮想現実体験、複合現実体験、または複合現実体験のうちの 1 つまたは複数を提供するように構成され得る。例えば、

10

20

30

40

50

左ディスプレイスクリーン 302 は、ステレオビジョンを可能にするために、ユーザの左目に左シーンを表示し得、右ディスプレイスクリーン 303 は、ユーザの右目に右シーンを表示し得る。いくつかの実装では、ディスプレイスクリーン 302 および 303 は不透明であり、1つまたは複数の組み込まれた仮想オブジェクト（例えば、仮想ソース 354）を用いて（例えば、第1の音源 330 のイメージ 350 および第2の音源 332 のイメージ 352 を含む）視覚的シーンの表現を生成する。いくつかの実装では、ディスプレイ 302 および 303 は、視覚的シーン上に1つまたは複数の組み込まれた仮想オブジェクトをオーバーレイする。いくつかの実装では、デバイス 102 は、三次元（3D）ビューリングを提供し得るかまたは二次元ビューリング（2D）を提供し得る2つのディスプレイ 302～303 ではなく単一のディスプレイを含み得る。

10

#### 【0045】

[0058] デバイス 102 は、仮想オブジェクトを用いて視覚的環境を拡張するか、仮想音を用いた音響環境を拡張するか、あるいはそれらの組み合わせを行うように構成され得る。第1の例として、デバイス 102 は、実世界音源で発生しているものとしてユーザによって感知されるべき仮想音を生成し得る。デバイス 102 は、第2の木 342 が仮想音（例えば、ゲームアプリケーションにおける歌う木）のソースとして使用されるべきであることを決定し得る。例えば、第2の木 342 は、ユーザに基づいて、例えば、ユーザの指さし（user pointing）、ユーザの目の動きまたは注視点を追跡すること、またはユーザの言語認識（例えば、第2の木 342 を識別する）を識別することによって、選択され得る。

20

#### 【0046】

[0059] 第1の実装では、デバイス 102 は、第1の音源 330 からの第1の音（例えば、犬の吠え声）を検出し、第2の音源 332 からの第2の音（例えば、人の話し声）を検出し得る。デバイス 102 は、第1の距離 340 と第1の音源 330 に対応する到来方向とを決定し得、かつ第2の距離 342 と第2の音源 332 に対応する到来方向とを決定し得る。デバイス 102 は、第1の音の第1の音特性と、第2の音の第2の音特性とを決定し得る。デバイス 102 は、第2の木 342 への距離が第2の距離 342 のものよりも第1の距離 340 に比較的近いことを（または、図2で説明されるように、第2の木 342 が第1の音源 330 と同じ区域にあり、かつ第2の音源 332 とは異なる区域にあることを）、（例えば、深度マップによって）決定し得る。デバイス 102 は、第1の音源 330 からの第1の音の第1の音特性に基づいて、仮想音（例えば、歌う木の声）の音特性を選択し得る。

30

#### 【0047】

[0060] 第1の例の第2の実装では、デバイス 102 のユーザは、第2の木 342 に、またはその近くに、音源（例えば、ラウドスピーカ）を配置し得る。デバイス 102 は、音源から受信されるオーディオ信号に基づいて（例えば、選択されたオーディオ信号のユーザが開始した再生によって）、第1の音特性を決定し得る。デバイス 102 は、仮想音（例えば、歌う木の声）の音特性として第1の音特性を使用し得る。

#### 【0048】

[0061] 第1の例の第3の実装では、デバイス 102 は、音響環境において実世界の音の音特性を使用しない、仮想音の音特性のテーブルベースの決定を実装し得る。デバイス 102 は、第2の木 342 までの距離および方向を推定し、1つまたは複数の音響条件（例えば、デバイス 102 が閉鎖された空間にあるか、または解放された空間にあるか）を推定し、仮想音の音特性を生成するために1つまたは複数のテーブルルックアップ動作および計算を開始し得る。

40

#### 【0049】

[0062] 第2の例として、デバイス 102 は、実世界音のソースとしてユーザに表示されるべき仮想音源を生成し得る。例えば、第3の音源 334 は、部分的にまたは完全に隠れているか、曖昧であるか、またはそうでない場合、視覚的に感知するのが困難であり得る。デバイス 102 は、第3の音源 334 からのオーディオ信号を検出し得、その検出さ

50

れた音の1つまたは複数の特性に基づいて第3の音源334の位置を推定し得る。第3の音源334の位置を推定した後、デバイス102は、音のソースをユーザが視覚的に認識することができるよう、ディスプレイスクリーン302～303上に仮想音源354を追加し得る。

【0050】

[0063] 第2の例の第1の実装では、デバイス102は、図1のデータベース116に記憶された特性を有する、第1の音源330からおよび第2の音源332からのオーディオ信号などの音響環境の他のオーディオ信号の音特性と、検出されたオーディオ信号の1つまたは複数の音特性とを少なくとも部分的に比較することによって、第3の音源334の位置を推定し得る。デバイス102から第3の音源334までの距離は、その比較に基づいておよび距離340と342に基づいて推定され得る(例えば、第3の音源334までの距離は、検出された音の音特性との類似性に基づいて、音源への距離またはそれを含む区域の領域を使用して推定され得る)。到來方向は、デバイス102の異なるマイクロフォン310～313におけるオーディオ信号の位相差などによって、デバイス102によって推定され得る。

10

【0051】

[0064] 第2の例の第2の実装では、デバイス102は、再生されたオーディオ信号の音特性と、検出されたオーディオ信号の1つまたは複数の音特性とを少なくとも部分的に比較することによって、第3の音源334の位置を推定し得る。例えば、デバイス102は、1つまたは複数の音源から受信された1つまたは複数のオーディオ信号に基づいて(例えば、シーン内の1つまたは複数の位置におけるラウドスピーカからの選択されたオーディオ信号のユーザが開始した再生によって)、音特性を決定し得、データベース116に記憶し得る。

20

【0052】

[0065] 第3の例では、デバイス102は、仮想音と、仮想音のソースとしてユーザに表示されるべき仮想音源とを生成し得る。例えば、デバイス102は、仮想音源として第3の音源334を追加し、第3の音源334からの仮想音を追加することによって、音響および視覚的環境を拡張し得る。デバイス102は、(例えば、ゲームのシナリオに基づいて)第3の音源334の位置を選択し得、第3の音源334が実世界の視覚的シーンにあるかのように現れるよう、ディスプレイ302～303中の適切な位置に仮想音源354の視覚的表現を表示し得る。デバイス102は、上述された実装のうちの1つまたは複数を使用して(例えば、1つまたは複数の同様の場所に位置する(similarly-located)実世界音源からのオーディオ信号に基づいて)、仮想音源334から仮想音の1つまたは複数の音特性を選択し得る。

30

【0053】

[0066] 図4Aを参照すると、1つまたは複数の検出された音に基づいて、あるシーンに仮想オブジェクトを挿入する例が示されている。図4Aは、第1のシーン400と第2のシーン420とを描く。第1のシーン400は、ヘッドセット402のディスプレイスクリーン404を通して見られるような、複合現実処理していない環境の視覚的描写である。第2のシーン420は、ヘッドセット402のディスプレイスクリーン404を通して見られるような、複合現実処理を用いた環境の視覚的描写である。

40

【0054】

[0067] ヘッドセット402は、図1のデバイス102に対応し得る。例えば、ヘッドセット402は、デバイス102と同様の構成要素を含み得、かつデバイス102に実質的に類似する方法で動作し得る。ヘッドセット402は、1つまたは複数のディスプレイスクリーン404と1つまたは複数のマイクロフォン406とを含む。1つまたは複数のディスプレイスクリーン404は、図1のディスプレイスクリーン112、図3のディスプレイスクリーン302、303、あるいはそれらの組み合わせに対応し得る。1つまたは複数のマイクロフォン406は、図1の1つまたは複数のマイクロフォン106に対応し得る。

50

## 【0055】

[0068] 1つまたは複数のマイクロフォン406は、シーン400、420中の音を検出し得る。例えば、1つまたは複数のマイクロフォン406は、鳥の音410（例えば、さえずり）、人間の音声412（例えば、話し声）、および猿の音414を検出し得る。鳥の音410の音源の位置は、図1に関して説明されるオーディオ分析技法を使用して決定され得る。例えば、ヘッドセット402内のプロセッサ（図示せず）は鳥の音410を識別し、その鳥の音410が第1のシーン400の左上の方の位置から来ることを決定し得る。人間の音声412の音源の位置は、図1に関して説明されるオーディオ分析技法を使用して決定され得る。例えば、ヘッドセット402内のプロセッサは人間の音声412を識別し、その人間の音声412が第1のシーン400の中心の方の位置から来ることを決定し得る。猿の音の音源の位置は、図1に関して説明されるオーディオ分析技法を使用して決定され得る。例えば、ヘッドセット402内のプロセッサは猿の音414を識別し、その猿の音414が第1のシーン400の右上の方の位置から来ることを決定し得る。

## 【0056】

[0069] 音410、412、414は1つまたは複数のマイクロフォン406によって検出可能であるが、音源は、ヘッドセット402のカメラでは見ることができない。例えば、鳥の音410を出している鳥は、木にある葉によってカメラから隠れている可能性があり、人間の音声412を出している人間は、霧でカメラから隠れている可能性があり、猿の音414を出している猿は、別の木にある葉によってカメラから隠れている可能性がある。

## 【0057】

[0070] ヘッドセット402は、第2のシーン420を生成するために、図1の複合現実処理技法を第1のシーン400に適用し得る。例えば、プロセッサは、図1の仮想音源生成回路128と実質的に類似する方法で処理し得、鳥の音410が生成される場所に仮想の鳥430を挿入し得る。同様に、プロセッサは、人間の音声412が生成される場所に仮想の人間432を挿入し得、猿の音414が生成される場所に仮想の猿434を挿入し得る。仮想の鳥430、仮想の人間432、および仮想の猿434は、図1に関して説明される複合現実処理技法を使用して、1つまたは複数のディスプレイスクリーン404を通して第2のシーン420に表示され得る。

## 【0058】

[0071] そのため、仮想オブジェクト430、432、434は、ユーザ体験を改善するために、複合現実処理技法を使用してシーンに挿入され得る。例えば、ユーザが音を聞くことができるが音に関連した音源を見ることができない場合、ヘッドセット402は、ユーザ体験を改善するために音が生成された場所に近接する位置に、（1つまたは複数のディスプレイスクリーン404を通して見ることができる）仮想ソースを挿入し得る。

## 【0059】

[0072] 図4Bを参照すると、1つまたは複数の検出された音に基づいて、あるシーンに仮想オブジェクトを挿入する別の例が示されている。図4Bは、デバイス452（例えば、防犯カメラ）によってキャプチャされるシーンを描く。例えば、デバイス452は、そのシーンの視覚的描写をキャプチャし得る。1つの実装によると、デバイス452は、図1のデバイス102に対応し（またはそれに含まれ）得る。例えば、デバイス452は、デバイス102と同様の構成要素を含み得、かつデバイス102に実質的に類似する方法で動作し得る。

## 【0060】

[0073] デバイス452によってキャプチャされたシーンは、ベビーベッド448を含み得る。デバイス452は、オーディオ信号を検出するように構成された1つまたは複数のマイクロフォン（図示せず）を含み得る。例えば、1つまたは複数のマイクロフォンは、乳児の音声450を検出し得る。デバイス452内のプロセッサは、オーディオ信号の音源の位置を決定し得る。例えば、デバイス452内のプロセッサは、オーディオ信号の音源の位置を決定し得る。例として、プロセッサは、乳児の音声450の音特性を決定し

得る。音特性は、直接音対残響音比（D R R）などの残響特性を含み得る。音特性に基づいて、プロセッサは、デバイス452からの音源の距離を決定し得る。例えば、プロセッサは、音源が第1の区域（例えば、近距離場区域）に位置するか、または第2の区域（例えば、遠距離場区域）に位置するかを決定し得る。プロセッサはまた、乳児の音声450の到来方向を推定し得る。音源の位置は、音源に関連付けられた区域および到来方向に基づき得る。例えば、到来方向は、1つまたは複数のマイクロフォンからの音源の方向を示し得、音源に関連付けられた区域は、音源が1つまたは複数のマイクロフォンからどれだけ遠いかを示し得る。

#### 【0061】

[0074] デバイス452は、乳児の音声に基づいて環境の1つまたは複数の音響特性を推定し得る。1つの実装によると、デバイスは、1つまたは複数の音響特性に基づいて仮想音を生成し得る。例えば、デバイス452は、音源の位置から、音生成の1つまたは複数のオーディオ特徴を有する、仮想の乳児の音声462を生成し得る。プロセッサは、仮想の乳児の音声462に基づいて、離れた位置に音信号を出力し得る。例えば、ディスプレイスクリーン490は、デバイス452（例えば、防犯カメラ）とは異なる場所に位置し得る。例として、デバイス452は、家の1つの部屋に位置し得、ディスプレイスクリーン490は、家の別の部屋に位置し得る。デバイス490は、ディスプレイスクリーン490を含む、家の部屋に位置する1つまたは複数のスピーカに、仮想の乳児の音声462を出力し得る。

#### 【0062】

[0075] 1つの実装によると、デバイス452は、オーディオ信号に基づいて、音源を特定のオブジェクトに分類するように構成され得る。例えば、デバイス452は、乳児の音声450の音源を、乳児に分類し得る。デバイス452は、特定のオブジェクトの仮想イメージを生成し得る。例えば、デバイス452は、仮想の乳児460を生成し得る。デバイス452はまた、環境の視覚的描写に仮想イメージを挿入し得る。例として、デバイス452は、ディスプレイスクリーン490において、視覚的描写に仮想の乳児460を挿入し得る。仮想の乳児は、音源の位置に対応する視覚的描写中の特定の場所に位置し得る。例えば、仮想の乳児460は、ベビーベッド448に（例えば、乳児の音声450が生成される場所に）位置し得る。

#### 【0063】

[0076] 図6を参照すると、音響環境の表現を拡張する方法600を図示するフローチャートが描かれている。方法600は、図1のデバイス102などの、マイクロフォンおよびプロセッサを含むデバイスにおいて行われ得る。

#### 【0064】

[0077] 602において、オーディオ信号は、マイクロフォンにおいて検出される。例えば、オーディオ信号は、デバイス102のマイクロフォン（複数を含む）106において検出された図1のオーディオ信号140または142に対応し得る。

#### 【0065】

[0078] 604において、オーディオ信号の音特性は、プロセッサにおいて決定され得る。例えば、図1のプロセッサ114は、オーディオ信号140に基づいて、第1の音特性164を検出し得る。606において、仮想音は、オーディオ信号の音特性に基づいて生成され、608において、仮想音は、オーディオ再生デバイスでの再生のために音響環境の表現に挿入される。例えば、図1の仮想音生成回路126は、仮想音を表現し、かつオーディオ再生デバイス162での再生のために音響環境の表現172に追加されるデータを生成し得る。音響環境の表現に仮想音を挿入することは、拡張、仮想、または複合現実ヘッドセットの1つまたは複数のラウドスピーカ（例えば、イヤフォン）で仮想音を出力することを含み得る。

#### 【0066】

[0079] いくつかの実装では、音特性は、オーディオ信号の少なくとも1つの残響特性を含み得る。少なくとも1つの残響特性は、オーディオ信号の直接音対残響音比（D R R

10

20

30

40

50

)を含み得る。仮想音の1つまたは複数の残響特性は、オーディオ信号の少なくとも1つの残響特性に基づいて決定され得、仮想音の到来方向は、仮想音に関連付けられた仮想音源の位置に基づいて推定され得る。仮想音は、1つまたは複数の残響特性と到来方向とにに基づいて生成され得る。

【0067】

[0080] 代替的にまたは追加として、方法600は、オーディオ信号の音源が音響環境の第1の区域に位置するか、音響環境の第2の区域に位置するかを、音特性に基づいて決定することを含み得る。第1の区域(例えば、図2の第1の区域202)は、第2の区域(例えば、図2の第2の区域204)よりもマイクロフォンに近い可能性がある。仮想音に関連付けられた仮想音源が第1の区域に位置するか、または第2の区域に位置するかに関する決定がなされ得る。仮想音の1つまたは複数の特性は、音源の位置と仮想音源の位置とにに基づき得る。

10

【0068】

[0081] 例えば、仮想音の特性は、音源が第1の区域に位置しかつ仮想音源が第1の区域に位置する場合に、オーディオ信号の音特性に実質的に類似し得、仮想音の特性は、音源が第1の区域に位置しかつ仮想音源が第2の区域に位置する場合に、オーディオ信号の音特性と異なり得る。別の例では、仮想音の特性は、音源が第2の区域に位置しかつ仮想音源が第2の区域に位置する場合に、オーディオ信号の音特性に実質的に類似し得、仮想音の特性は、音源が第2の区域に位置しかつ仮想音源が第1の区域に位置する場合に、オーディオ信号の音特性と異なり得る。

20

【0069】

[0082] オーディオ信号は、オーディオデバイスのユーザが開始した再生に基づいて生成され得る。例えば、オーディオ信号は、仮想音の仮想音源の位置で、またはその近くで、音響環境中に設置されたラウドスピーカまたは他の音生成器を使用して生成され得る。オーディオ信号は、マイクロフォン(複数を含む)によってキャプチャされ、位置に対応しかつリアルさを高めるために仮想音に適用されることができる音特性(例えば、到来方向、残響特性など)を決定するように処理され得る。例えば、マイクロフォン(複数を含む)でキャプチャされたオーディオ信号とラウドスピーカを展開する(played out)オーディオ信号との差が識別され、その位置からマイクロフォン(複数を含む)までの音の移動特性を決定するために使用され得る。

30

【0070】

[0083] 代替的に、オーディオ信号は、オーディオデバイスのユーザが開始した再生なしに生成され得る。例えば、オーディオ信号は、仮想音の仮想音源の位置で、またはその近くで、音響環境中の音生成要素によって生成され得る。1つまたは複数のオーディオ信号は、マイクロフォン(複数を含む)によってキャプチャされた音中で検出され得、オーディオ信号の1つまたは複数のソースの位置が推定され得る。仮想音のソースの位置に近接するまたはそれとコロケートされたソースからのオーディオ信号の音特性は、仮想音の音特性を生成するために使用され得る。

【0071】

[0084] 特定の実装では、オーディオ再生デバイスは、図3に関連して説明されるように、仮想現実シーン、拡張現実シーン、または複合現実シーンのうちの少なくとも1つを生成するように構成されたヘッドセットに組み込まれる。音響環境の視覚的表現は、1つまたは複数のカメラを使用してキャプチャされ、ヘッドセットに表示され得、仮想音に関連付けられる仮想音源は、仮想表現に挿入され得る。他の実装では、オーディオ再生デバイスは、視覚ディスプレイ(visual display)を用いてヘッドセットに実装されていない可能性があり、代わりに、別のデバイス(例えば、モバイルフォンまたは音楽プレイヤーデバイス)に組み込まれ得る。

40

【0072】

[0085] 図7を参照すると、音響環境の表現を拡張する方法700を図示するフローチャートが描かれている。方法700は、図1のデバイス102などの、マイクロフォンお

50

およびプロセッサを含むデバイスにおいて行われ得る。図6と比較すると、図7の方法の実装は、受信されるオーディオ信号の音特性に基づいていない仮想音の特性を決定し得る。

【0073】

[0086] 702において、視覚的シーンは、1つまたは複数のカメラを使用してキャプチャされる。1つまたは複数のカメラは、オーディオキャプチャデバイスに近接する。オーディオキャプチャデバイスは、仮想現実シーン、拡張現実シーン、または複合現実シーンのうちの少なくとも1つを生成するように構成されたヘッドセットに組み込まれ得る。

【0074】

[0087] 704において、視覚的シーン中の仮想音源の位置が決定される。位置は、1つまたは複数のカメラからの仮想音源の距離と、1つまたは複数のカメラに関する仮想音源の角度位置とを示す。

10

【0075】

[0088] 706において、仮想音の1つまたは複数の音特性は、仮想音源の位置に基づいて決定される。1つまたは複数の音特性は、仮想音の少なくとも1つの反響特性に含まれ得る。いくつかの実装では、仮想音の1つまたは複数の特性は、音響環境の特性にさらに基づく。例えば、音響環境の特性が決定され得る。音響環境の特性は、音響環境が屋内環境にあるか、または屋外環境にあるかを示し得る。いくつかの実装では、仮想音源の位置は、位置情報を有する音特性に関連するデータベース中の位置情報と比較され得、仮想音の1つまたは複数の音特性は、その比較に基づいて決定され得る。

20

【0076】

[0089] 708において、仮想音は、1つまたは複数の音特性に基づいて生成され、710において、仮想音は、オーディオ再生デバイスでの再生のために音響環境の表現に挿入される。

【0077】

[0090] 方法700は、視覚的シーンの表現に仮想音源を挿入することと、視覚的シーンの表現を表示することとを含み得る。仮想音を挿入することは、ヘッドセットの1つまたは複数において仮想音に対応する音信号を出力することを含み得る。

30

【0078】

[0091] 図8は、仮想オーディオ信号を生成する方法800の例を描く。方法800は、図1のデバイス102などのオーディオ再生デバイスを含むデバイスにおいて行われ得る。

【0079】

[0092] 802において、方法800は、人工音（例えば、仮想音）に関連付けられるべきオブジェクトを選択することを含む。例えば、図3に関連して説明されるように、第2の木342が選択され得る。オブジェクトは、ユーザ入力に基づいて、例えば、ユーザの指でオブジェクトを指すこと、オブジェクトをユーザの目の動きまたは注視点を追跡すること、またはユーザの言語認識（例えば、オブジェクトに仮想音を追加するためのコマンドを識別する）を識別することによって、選択され得る。

40

【0080】

[0093] 804において、オブジェクトの位置が決定される。オブジェクトの位置は、深度マップまたは他の視覚的処理技法に基づいて決定され得る。オブジェクトの位置は、図2に関連して説明されるように、特定の区域に対応する。

【0081】

[0094] 806において、特定の区域に関連付けられた1つまたは複数の音反響パラメータと、特定の区域に関連付けられた1つまたは複数の到來方向（DOA）とが決定される。例えば、音反響パラメータとDOAパラメータとは、特定の区域からの、実世界で記録された音に基づいて決定され得る。別の例では、音反響パラメータとDOAパラメータとは、特定の区域からの、実世界で再生された音に基づいて決定され得る。別の例では、音区域反響パラメータは、オブジェクトの視覚的深度およびDOA情報に基づき、あらかじめコンパイルされた音響ルックアップテーブルに基づく。

50

## 【0082】

[0095] 808において、オーディオフィルタは、1つまたは複数の音反響パラメータと1つまたは複数のDOAパラメータに基づいて生成され得る。810において、オーディオフィルタは、人工音を生成するために、オブジェクトに関連付けられた「クリーンな」音信号に適用され得る。例えば、オーディオフィルタは、図1の仮想音源生成回路128によって生成および適用され得る。

## 【0083】

[0096] 812において、人工音は、図1のスピーカ（複数を含む）110または図3のラウドスピーカ320～323などにおけるヘッドセットで出力され得る。例えば、人工音は、（例えば、図3のラウドスピーカ320～321における）ヘッドセットイヤピース（*headset earpieces*）で再生され得る。

10

## 【0084】

[0097] 図9は、実世界環境で仮想オブジェクトを生成する方法900の一例を描く。方法900は、図1のデバイス102などのビデオ再生デバイスを含むデバイスにおいて行われ得る。

## 【0085】

[0098] 902において、方法900は、音響環境中の音を検出することを含む。例えば、図1を参照すると、1つまたは複数のマイクロフォン106は、音響環境150中のオーディオ信号を検出し得る。

20

## 【0086】

[0099] 904において、音の1つまたは複数の音反響パラメータが決定され得、かつ音の到来方向（DOA）が決定され得る。例えば、図1を参照すると、オーディオ分析回路124は、音の1つまたは複数の反響パラメータを決定し得る。1つの実装によると、オーディオ分析回路124は、音のDRRを決定し得る。追加として、位置推定回路122は、音のDOAを決定し得る。例えば、位置推定回路122は、音の角度位置を決定し得る。

20

## 【0087】

[00100] 906において、音の音源の位置は、1つまたは複数の音反響パラメータとDOAに基づいて決定され得る。位置は、音源の深度と、音源の方向とを示し得る。例えば、図1を参照すると、位置推定回路は、1つまたは複数の音反響パラメータとDOAに基づいて音源の位置を決定し得る。1つの実装によると、DRRおよび角度位置は、音の深度および方向を決定するためにプロセッサによって使用され得る。

30

## 【0088】

[00101] 908において、音に関連付けられた仮想オブジェクトの視覚的特性が決定され得る。視覚的オブジェクトは、音響環境の音、位置、および視覚的表現に基づき得る。例えば、図1を参照すると、プロセッサ114は、仮想オブジェクトが「実際の視覚的シーン」に「融合」するように、仮想オブジェクトのためのシェーディングスキーム、仮想オブジェクトのためのカラースキーム、仮想オブジェクトのサイズスキーム、あるいはそれらの組み合わせを決定し得る。これらスキームは、位置に基づき得る。非限定的例として、音源の位置が比較的遠いとプロセッサ114が決定した場合、プロセッサ114は、比較的小さいサイズに対応する視覚的特性を選択し得る。

40

## 【0089】

[00102] 910において、仮想オブジェクトは、視覚的特性に基づいて生成され得る。例えば、図1を参照すると、仮想音源生成回路128は、908において説明された視覚的特性（例えば、異なるスキーム）に基づいて仮想オブジェクトを生成し得る。

## 【0090】

[00103] 912において、仮想オブジェクトは、音響環境の視覚的表現に挿入され得る。例えば、ビデオ再生デバイス160は、音響環境150の視覚的表現170に仮想オブジェクトを挿入し得る。

## 【0091】

50

[00104] 図9の方法900は、仮想複合現実シーンを生成するために、仮想オブジェクトがシーンの視覚的表現に挿入されることを可能にし得る。例えば、「実世界」で表示されない仮想オブジェクトは、ユーザの満足を高めるために、視覚的表現に事実上挿入され得る。

#### 【0092】

[00105] 図10は、空間的にフィルタされた音を生成する方法1000の例を描く。方法1000は、図1のデバイス102などのオーディオ再生デバイスを含むデバイスにおいて行われ得る。

#### 【0093】

[00106] 1002において、方法1000は、マイクロフォンにおいて第1のオーディオ信号を検出することを含む。第1のオーディオ信号は、音響環境中の音源によって生成され得る。例えば、オーディオ信号は、デバイス102のマイクロフォン（複数を含む）106において検出された図1のオーディオ信号140または142に対応し得る。1004において、方法1000はまた、第1のオーディオ信号の特性を決定することを含む。例えば、図1を参照すると、プロセッサ114は、オーディオ信号140に基づいて音特性164を決定し得る。

10

#### 【0094】

[00107] 1006において、方法1000はまた、空間的にフィルタされたオーディオファイルを生成するために、特性に基づいて第2のオーディオ信号のデジタル表現を空間的にフィルタすることを含む。例えば、図1を参照すると、仮想音生成回路126は、空間的にフィルタされたオーディオファイルを生成するために、音特性164に基づいてオーディオ信号のデジタル表現を空間的にフィルタし得る。1008において、方法1000はまた、空間的にフィルタされたオーディオ信号をスピーカに送ることを含む。空間的にフィルタされたオーディオ信号は、空間的にフィルタされたオーディオファイルに基づき得る。例えば、図1を参照すると、仮想音源生成回路128は、空間的にフィルタされたオーディオに基づいて空間的にフィルタされたオーディオ信号を生成し得、空間的にフィルタされたオーディオ信号を1つまたは複数のスピーカ110に送り得る。

20

#### 【0095】

[00108] 1010において、方法1000はまた、スピーカにおいて空間的にフィルタされたオーディオ信号を、空間的にフィルタされた音として投影することを含む。例えば、図1を参照すると、1つまたは複数のスピーカ110は、空間的にフィルタされたオーディオ信号を、空間的にフィルタされた音として投影し得る。空間的にフィルタされた音は、仮想音144を含み得る。1つの実装によると、仮想音144は、コンピュータによって生成される音を含む。

30

#### 【0096】

[00109] 図11は、出力音の方法1100の例を描く。方法1100は、図1のデバイス102などのオーディオ再生デバイスを含むデバイスにおいて行われ得る。

#### 【0097】

[00110] 1102において、方法1100は、環境の1つまたは複数の位置において1つまたは複数の音響特性を決定することを含む。例えば、図5を参照すると、ラウドスピーカ502は、ユーザの手に装着されるように構成され、環境の異なる位置においてオーディオ信号をし得る。例として、ユーザが自分の手を動かし得、ラウドスピーカ502は、ユーザの手が位置する場所でオーディオ信号を生成し得る。ユーザのヘッドセットは、ラウドスピーカによって投影されたオーディオ信号を検出するように構成された1つまたは複数のマイクロフォンを含み得る。検出されたオーディオ信号に基づいて、ヘッドセットは、環境の異なる位置において1つまたは複数の音響特性を検出し得る。

40

#### 【0098】

[00111] 1104において、方法1100はまた、特定のオブジェクトを示すユーザ選択を受信することを含む。例えば、図5を参照すると、ヘッドセットは、ユーザがピアノ504を選択したことを示すユーザ選択を受信し得る。方法1100はまた、1106

50

において、特定の位置のユーザインジケーションを受信することを含む。ウェアラブルセンサは、特定の位置を検出することと、特定の位置のユーザインジケーションを生成することと、プロセッサ（例えば、ヘッドセット）に特定の位置のユーザインジケーション送ることと、を行うように構成され得る。例えば、図5を参照すると、ヘッドセットは、ユーザが特定の位置（例えば、ユーザの顔の約2フィート前方）を指さしているとのユーザインジケーション（ユーザの手におけるウェアラブルセンサを介する）を受信し得る。

#### 【0099】

[00112] 1108において、方法1100はまた、1つまたは複数の音響特性に基づいて、特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定することを含む。例えば、図5を参照すると、ヘッドセットは、ユーザの顔の2フィート前方に生成される音のオーディオ特徴を決定し得る。例として、ラウドスピーカ502がユーザの顔の約2フィート前方にあったとき、ヘッドセットは、ラウドスピーカ502から生成される音のオーディオ特徴を決定し得る。

10

#### 【0100】

[00113] 1110において、方法1100はまた、特定のオブジェクトに関連付けられた仮想音を環境に挿入することを含む。仮想音は、1つまたは複数のオーディオ特徴を有する。例えば、図5を参照すると、ヘッドセットは、ラウドスピーカがユーザの顔の約2フィート前方にあったとき、ラウドスピーカ502によって生成される音のオーディオ特徴を含む仮想のピアノ音を生成し得る。別の例では、ヘッドセットは、1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較し得る。各エントリは、異なるサウンドに関連付けられ得る。ヘッドセットは、環境への仮想音の挿入の前に、比較に基づいて特定のエントリから仮想音を導出し得る。仮想音が生成された（または、メモリから導出された）後、ヘッドセットは、環境に仮想音（例えば、ピアノ音）を挿入し得る。

20

#### 【0101】

[00114] 図12は、仮想音を生成する方法1200の例を描く。方法1200は、図1のデバイス102などのオーディオ再生デバイスを含むデバイスにおいて行われ得る。

#### 【0102】

[00115] 1202において、方法1200は、1つまたは複数のマイクロフォンにおいて、ある環境中のオーディオ信号を検出することを含む。例えば、図1を参照すると、1つまたは複数のマイクロフォン106は、オーディオ信号140を検出し得る。1204において、方法1200はまた、プロセッサにおいて、オーディオ信号の音源の位置を決定することを含む。例えば、図1を参照すると、デバイス102は、オーディオ信号140の音源130の位置を決定し得る。例えば、デバイスは、オーディオ信号140の音特性を決定し得る。音特性に基づいて、デバイス102は、音源130が音響環境150の第1の区域に位置するか、または音響環境150の第2の区域に位置するかを決定し得る。第1の区域は、第2の区域よりも近くにある可能性がある。

30

#### 【0103】

[00116] 1206において、方法1200はまた、オーディオ信号に基づいて環境の1つまたは複数の音響特性を推定することを含む。例えば、図1を参照すると、オーディオ分析回路124は、検出されたオーディオ信号140に基づいて音響環境150の1つまたは複数の音響特性を推定し得る。

40

#### 【0104】

[00117] 1208において、方法1200はまた、1つまたは複数の音響特性に基づいて環境に仮想音を挿入することを含む。仮想音は、音源の位置から生成される音の1つまたは複数のオーディオ特徴を有する。例えば、図1を参照すると、仮想音源生成回路128は、音響環境の音響特性に基づいて仮想音を生成し得る。別の例では、プロセッサ114は、1つまたは複数の音響特性をメモリ104中の1つまたは複数のエントリと比較し得る。各エントリは、異なる仮想音に関連付けられ得る。プロセッサ114はまた、環境への仮想音の挿入の前に、比較に基づいて特定のエントリから仮想音を導出し得る。仮想音が生成された（または、メモリから導出された）後、仮想音生成回路126は、音響

50

環境 150 に仮想音を挿入し得る。

【0105】

[00118] 図 13 を参照すると、デバイス 102 のブロック図が描かれる。特定の実装では、デバイス 102 は、プロセッサ 114（例えば、CPU）を含む。プロセッサ 114 は、位置推定回路 122、オーディオ分析回路 124、仮想音生成回路 126、仮想音源生成回路 128、ビデオ再生デバイス 160、およびオーディオ再生デバイス 162 を含む。

【0106】

[00119] デバイス 102 は、プロセッサ 114 に結合されるメモリ 104 を含む。追加として、音特性のデータベース 116 は、プロセッサ 114 に結合され（例えば、それによってアクセス可能であり）得る。デバイス 102 はまた、トランシーバ 1341 を介してアンテナ 1342 に結合されたワイヤレスインターフェース 1340 を含む。デバイス 102 は、ディスプレイコントローラ 1326 に結合されたディスプレイスクリーン 112 を含み得る。1つまたは複数のスピーカ 110、1つまたは複数のマイクロフォン 106、あるいは両方は、コーダ / デコーダ（CODEC）1334 に結合され得る。CODEC 1334 は、デジタル - アナログコンバータ（DAC）1302 およびアナログ - デジタルコンバータ（ADC）1304 を含み得る。特定の実装では、CODEC 1334 は、1つまたは複数のマイクロフォン 106 からアナログ信号を受信し、アナログ - デジタルコンバータ 1304 を使用してデジタル信号にコンバートし得る。CODEC 1334 は、プロセッサ 114 からデジタル信号を受信し得、CODEC 1334 は、デジタル - アナログコンバータ 1302 を使用してデジタル信号をアナログ信号にコンバートし得、1つまたは複数のスピーカ 110 にアナログ信号を提供し得る。

10

20

30

40

【0107】

[00120] メモリ 104 は、図 6 ~ 図 10 の方法 600 ~ 1000 のうちの 1つまたは複数のような、本明細書で開示された方法および処理を行うために、プロセッサ 114、デバイス 102 の別の処理ユニット、あるいはそれらの組み合わせによって実行可能な命令 1368 を含み得る。本明細書で開示される装置 / システムの 1つまたは複数の構成要素は、1つまたは複数のタスク、あるいはそれらの組み合わせを行うための命令（例えば、命令 1368）を実行するプロセッサによって、専用ハードウェア（例えば、回路）を介して実装され得る。例として、メモリ 104、またはプロセッサ 114 の 1つまたは複数の構成要素は、ランダムアクセスメモリ（RAM）、磁気抵抗ランダムアクセスメモリ（MRAM）、スピントランジスタ（STT-MRAM : spin-torque transfer MRAM）、フラッシュメモリ、読み取り専用メモリ（ROM）、プログラマブル読み取り専用メモリ（PROM）、消去可能なプログラマブル読み取り専用メモリ（EEPROM（登録商標））、レジスタ、ハードディスク、リムーバブルディスク、またはコンパクトディスク読み取り専用メモリ（CD-ROM）、などのメモリデバイスであり得る。メモリデバイスは、コンピュータ（例えば、CODEC 1334 におけるプロセッサ、プロセッサ 114、および / またはデバイス 102 中の別の処理ユニット）によって実行されるとき、コンピュータに、本明細書で説明される方法の 1つまたは複数の少なくとも一部分を行わせ得る命令（例えば、命令 1368）を含み得る。例として、メモリ 104、またはプロセッサ 114 のうちの 1つまたは複数の構成要素は、コンピュータ（例えば、CODEC 1334 におけるプロセッサ、プロセッサ 114、および / または別の処理ユニット）によって実行されるとき、コンピュータに、本明細書で開示される方法のうちの少なくとも 1つまたは複数の部分を行わせる命令（例えば、命令 1368）を含む非一時的コンピュータ可読媒体であり得る。

30

【0108】

[00121] 特定の実施実装では、デバイス 102 は、モバイル局モデム（MSM）などの、システムインパッケージまたはシステムオンチップデバイス 1322 に含まれ得る。特定の実装では、プロセッサ 114、ディスプレイコントローラ 1326、メモリ 104

50

、 C O D E C 1 3 3 4 、およびワイヤレスインターフェース 1 3 4 0 は、システムインパッケージまたはシステムオンチップのデバイス 1 3 2 2 に含まれる。特定の実装では、タッチスクリーンおよび / またはキーパッドなどの入力デバイス 1 3 3 0 、および電源 1 3 4 4 が、システムオンチップデバイス 1 3 2 2 に結合されている。さらに、特定の実装では、図 1 3 に例示されるように、ディスプレイスクリーン 1 1 2 、入力デバイス 1 3 3 0 、 1 つまたは複数のスピーカ 1 1 0 、 1 つまたは複数のマイクロフォン 1 0 6 、アンテナ 1 3 4 2 、 1 つまたは複数のカメラ 1 0 8 、および電源 1 3 4 4 は、システムオンチップのデバイス 1 3 2 2 の外部にある。しかしながら、ディスプレイスクリーン 1 3 2 8 、 1 つまたは複数のカメラ 1 0 8 、入力デバイス 1 3 3 0 、 1 つまたは複数のスピーカ 1 1 0 、 1 つまたは複数のマイクロフォン 1 0 6 、アンテナ 1 3 4 2 、および電源 1 3 4 4 の各々は、インターフェースまたはコントローラなどのシステムオンチップデバイス 1 3 2 2 の構成要素に結合され得る。例示的実施例では、デバイス 1 0 2 は、ヘッドセット、モバイル通信デバイス、スマートフォン、セルラフォン、ラップトップコンピュータ、コンピュータ、タブレットコンピュータ、パーソナルデジタルアシスタント、ディスプレイデバイス、テレビ、ゲーム機、音楽プレイヤ、無線、デジタルビデオプレイヤ、光学ディスクプレイヤ、チューナ、カメラ、ナビゲーションデバイス、デコーダシステム、エンコーダシステム、自動車または航空機などの有人または無人車両内のデバイス、あるいはこれらの任意の組み合わせに対応する。

10

## 【 0 1 0 9 】

[00122] 説明された実装に関連して、仮想音を生成するための第 1 の装置は、ある環境中のオーディオ信号を検出するための手段を含む。例えば、オーディオ信号を検出するための手段は、図 1 および図 1 3 の 1 つまたは複数のマイクロフォン 1 0 6 、 1 つまたは複数の他のセンサ、あるいはそれらの組み合わせを含み得る。

20

## 【 0 1 1 0 】

[00123] 第 1 の装置はまた、ある環境中のオーディオ信号の音源の位置を決定するための手段を含み得る。例えば、音源の位置を決定するための手段は、図 1 および図 1 3 の位置推定回路 1 2 2 、図 1 および図 1 3 のプロセッサ、 1 つまたは複数の他のデバイス、あるいはそれらの組み合わせを含み得る。

## 【 0 1 1 1 】

[00124] 第 1 の装置はまた、オーディオ信号に基づいて環境の 1 つまたは複数の音響特性を推定するための手段を含み得る。例えば、音響特性を推定するための手段は、図 1 および図 1 3 のオーディオ分析回路 1 2 4 、図 1 および図 1 3 のプロセッサ、 1 つまたは複数の他のデバイス、あるいはそれらの組み合わせを含み得る。

30

## 【 0 1 1 2 】

[00125] 第 1 の装置はまた、 1 つまたは複数の音響特性に基づいて環境に仮想音を挿入するための手段を含み得る。仮想音は、音源の位置から生成する音の 1 つまたは複数のオーディオ特徴を有し得る。例えば、環境に仮想音を挿入するための手段は、図 1 および図 1 3 の仮想音生成回路 1 2 6 、図 1 および図 1 3 のプロセッサ、 1 つまたは複数の他のデバイス、あるいはそれらの組み合わせを含み得る。

## 【 0 1 1 3 】

[00126] 説明された実装に関連して、音を出力するための第 2 の装置は、環境の 1 つまたは複数の位置において 1 つまたは複数の音響特性を決定するための手段を含む。例えば、 1 つまたは複数の音響特性を決定するための手段は、図 1 および図 1 3 のオーディオ分析回路 1 2 4 、図 1 および図 1 3 のプロセッサ、 1 つまたは複数の他のデバイス、あるいはそれらの組み合わせを含み得る。

40

## 【 0 1 1 4 】

[00127] 第 2 の装置はまた、特定のオブジェクトを示すユーザ選択を受信するための手段を含み得る。例えば、ユーザ選択を受信するための手段は、ユーザインターフェース（例えば、図 1 3 の入力デバイス 1 3 3 0 ）、 1 つまたは複数の他のデバイス、あるいはそれらの組み合わせを含み得る。

50

## 【0115】

[00128] 音装置はまた、特定の位置のユーザインジケーションを受信するための手段を含み得る。例えば、ユーザインジケーションを受信するための手段は、ユーザインターフェース（例えば、図13の入力デバイス1330）、1つまたは複数の他のデバイス、あるいはそれらの組み合わせを含み得る。

## 【0116】

[00129] 音装置はまた、1つまたは複数の音響特性に基づいて、特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定するための手段を含み得る。例えば、オーディオ特徴を決定するための手段は、図1および図13のオーディオ分析回路124、図1および図13のプロセッサ、1つまたは複数の他のデバイス、あるいはそれらの組み合わせを含み得る。

10

## 【0117】

[00130] 第2の装置はまた、特定のオブジェクトに関連付けられた仮想音を環境に挿入するための手段を含み得る。仮想音は、1つまたは複数のオーディオ特徴を有し得る。例えば、仮想音を環境に挿入するための手段は、図1および図13の仮想音生成回路126、図1および図13のプロセッサ、1つまたは複数の他のデバイス、あるいはそれらの組み合わせを含み得る。

## 【0118】

[00131] 第2の装置はまた、特定の位置を検出することと、特定の位置のユーザインジケーションを生成することと、ユーザインジケーションを受信するための手段に特定の位置のユーザインジケーションを送ることと、を行うための手段を含み得る。例えば、検出、生成、および送るための手段は、ウェアラブルセンサ、1つまたは複数の他のデバイス、あるいはそれらの組み合わせを含み得る。

20

## 【0119】

[00132] 説明された実装に関連して、音響環境の表現を拡張するための第3の装置は、オーディオ信号を検出するための手段を含む。例えば、検出するための手段は、図1および図13の1つまたは複数のマイクロフォン106、1つまたは複数の他のセンサ、あるいはそれらの組み合わせを含み得る。

## 【0120】

[00133] 第3の装置はまた、オーディオ信号の音特性を決定するための手段を含み得る。例えば、音特性を決定するための手段は、図1および図13のオーディオ分析回路124、図1および図13のプロセッサ114、1つまたは複数の他のデバイス、非一時的コンピュータ可読記憶媒体におけるプロセッサ実行命令、あるいはそれらの組み合わせを含み得る。

30

## 【0121】

[00134] 第3の装置はまた、オーディオ信号の音特性に基づいて仮想音を生成するための手段を含み得る。例えば、仮想音を生成するための手段は、図1および図13の仮想音生成回路126、図1および図13のプロセッサ114、1つまたは複数の他のデバイス、非一時的コンピュータ可読記憶媒体におけるプロセッサ実行命令、あるいはそれらの組み合わせを含み得る。

40

## 【0122】

[00135] 第3の装置はまた、オーディオ再生デバイスによって再生のために仮想音に基づいて音信号を出力するための手段を含み得る。例えば、出力するための手段は、図1および図13の1つまたは複数のスピーカ110、1つまたは複数の他のセンサ、あるいはそれらの組み合わせを含み得る。

## 【0123】

[00136] 説明された実装に関連して、仮想音を生成するための第4の装置は、視覚的シーンをキャプチャするための手段を含み得る。キャプチャするための手段は、オーディオキャプチャデバイスに近接し得る。例えば、キャプチャするための手段は、図1および図13の1つまたは複数のカメラ108、1つまたは複数の他のセンサ、あるいはそれら

50

の組み合わせを含み得る。

【0124】

[00137] 第4の装置はまた、視覚的シーン中の仮想音源の位置を決定するための手段を含み得る。位置は、キャプチャするための手段からの仮想音源の距離と、キャプチャするための手段に関する仮想音源の角度位置とを示し得る。例えば、仮想音源の位置を決定するための手段は、図1および図13の位置推定回路122、図1および図13のプロセッサ114、1つまたは複数の他のデバイス、非一時的コンピュータ可読記憶媒体におけるプロセッサ実行命令、あるいはそれらの組み合わせを含み得る。

【0125】

[00138] 第4の装置はまた、仮想音源の位置に基づいて仮想音の1つまたは複数の音特性を決定するための手段を含み得る。例えば、1つまたは複数の音特性を決定するための手段は、図1および図13のオーディオ分析回路124、図1および図13の仮想音生成回路126、図1および図13の音特性のデータベース116、図1および図13のプロセッサ114、1つまたは複数の他のデバイス、非一時的コンピュータ可読記憶媒体におけるプロセッサ実行命令、あるいはそれらの組み合わせを含み得る。

10

【0126】

[00139] 第4の装置はまた、1つまたは複数の音特性に基づいて仮想音を生成するための手段を含み得る。例えば、仮想音を生成するための手段は、図1および図13の仮想音生成回路126、図1および図13のプロセッサ114、1つまたは複数の他のデバイス、非一時的コンピュータ可読記憶媒体におけるプロセッサ実行命令、あるいはそれらの組み合わせを含み得る。

20

【0127】

[00140] 第4の装置はまた、仮想音に基づいて音信号を出力するための手段を含み得る。例えば、出力するための手段は、図1および図13の1つまたは複数のスピーカ110、1つまたは複数の他のセンサ、あるいはそれらの組み合わせを含み得る。

【0128】

[00141] 説明された実装に関連して、第5の装置は、第1のオーディオ信号を検出するための手段を含む。第1のオーディオ信号は、音響環境中の音源によって生成され得る。例えば、検出するための手段は、図1および図13の1つまたは複数のマイクロフォン106、1つまたは複数の他のセンサ、あるいはそれらの組み合わせを含み得る。

30

【0129】

[00142] 第5の装置はまた、第1のオーディオ信号の特性を決定するための手段を含み得る。例えば、特性を決定するための手段は、図1および図13のオーディオ分析回路124、図1および図13のプロセッサ114、1つまたは複数の他のデバイス、非一時的コンピュータ可読記憶媒体におけるプロセッサ実行命令、あるいはそれらの組み合わせを含み得る。

【0130】

[00143] 第5の装置は、空間的にフィルタされたオーディオファイルを生成するために、特性に基づいて第2のオーディオ信号のデジタル表現を空間的にフィルタするための手段を含む。例えば、空間的にフィルタするための手段は、図1および図13の仮想音生成回路126、図1および図13のプロセッサ114、1つまたは複数の他のデバイス、非一時的コンピュータ可読記憶媒体におけるプロセッサ実行命令、あるいはそれらの組み合わせを含み得る。

40

【0131】

[00144] 第5の装置はまた、空間的にフィルタされた音を投影するための手段を含み得る。空間的にフィルタされた音は、投影するための手段に送られる空間的にフィルタされたオーディオ信号に基づき得、空間的にフィルタされたオーディオ信号は、空間的にフィルタされたオーディオファイルに基づき得る。例えば、投影するための手段は、図1および図13の1つまたは複数のスピーカ110、1つまたは複数の他の音出力デバイス、あるいはそれらの組み合わせを含み得る。

50

## 【0132】

[00145] 当業者は、本明細書で開示された態様に関連して説明された、様々な例示的な論理ブロック、構成、モジュール、回路、およびアルゴリズムステップが、電子ハードウェア、ハードウェアプロセッサのような処理デバイスによって実行されるコンピュータソフトウェア、または両方の組み合わせとして実装され得ることをさらに理解するだろう。様々な例示的な構成要素、ブロック、構成、モジュール、回路、およびステップが、概して、それらの機能の観点から上記に説明されている。そのような機能をハードウェアとして実施するか、ソフトウェアとして実施するかは、特定の適用例および全体的なシステムに課される設計制約に依存する。当業者は、説明した機能を特定の適用例ごとに様々な方法で実施し得るが、そのような実施の決定は、本開示の範囲からの逸脱を生じるものと解釈すべきではない。

10

## 【0133】

[00146] 本明細書の開示に関して説明された方法またはアルゴリズムのステップは、直接ハードウェアにおいて、プロセッサによって実行されるソフトウェアモジュールにおいて、またはその2つの組合せにおいて具現化され得る。ソフトウェアモジュールは、ランダムアクセスメモリ(RAM)、磁気抵抗ランダムアクセスメモリ(MRAM)、スピノン注入MRAM(STM-MRAM)、フラッシュメモリ、読み取り専用メモリ(ROM)、プログラマブル読み取り専用メモリ(PROM)、消去可能なプログラマブル読み取り専用メモリ(EPROM)、電気的に消去可能なプログラマブル読み取り専用メモリ(EEPROM)レジスタ、ハードディスク、リムーバブルディスク、またはコンパクトディスク読み取り専用メモリ(CD-ROM)のようなメモリデバイスに存在し得る。例示的なメモリデバイスは、プロセッサがこのメモリデバイスから情報を読み取り、このメモリデバイスに情報を書き込むことができるようプロセッサに結合される。代替として、メモリデバイスは、プロセッサに一体化され得る。プロセッサおよび記憶媒体は、ASIC中に常駐し得る。ASICは、コンピューティングデバイスまたはユーザ端末に存在し得る。代替として、プロセッサおよび記憶媒体は、コンピューティングデバイスまたはユーザ端末内の離散構成要素として存在し得る。

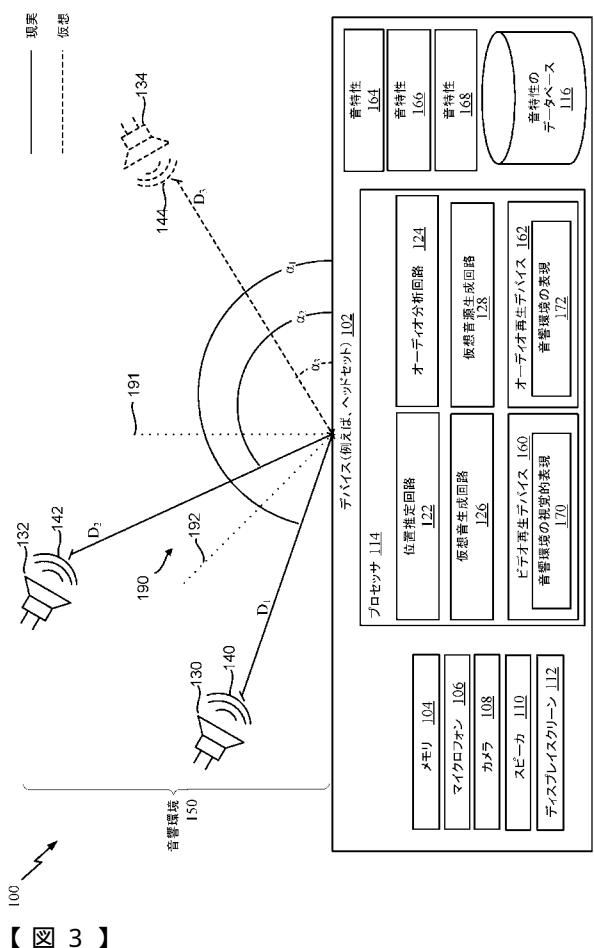
20

## 【0134】

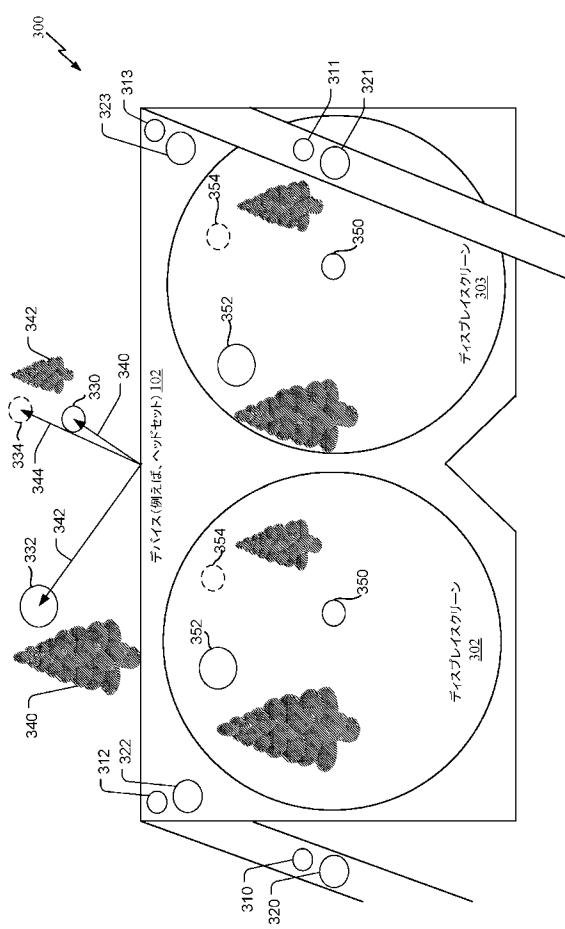
[00147] 開示された実施形態の先の説明は、当業者が、開示された態様形態を製造または使用することを可能にするために提供される。これらの態様に対する様々な修正は、当業者に対して容易に明らかであり、本明細書で定義される原理は、本開示の範囲から逸脱することなく他の態様に適用され得る。よって、本開示は、本明細書で示される態様に限定されることを意図するものではなく、下記の特許請求の範囲で定義されるような原理および新規な特徴と一致し得る最も広い範囲を与えられるべきである。

30

【 义 1 】



【 义 3 】



【 図 2 】

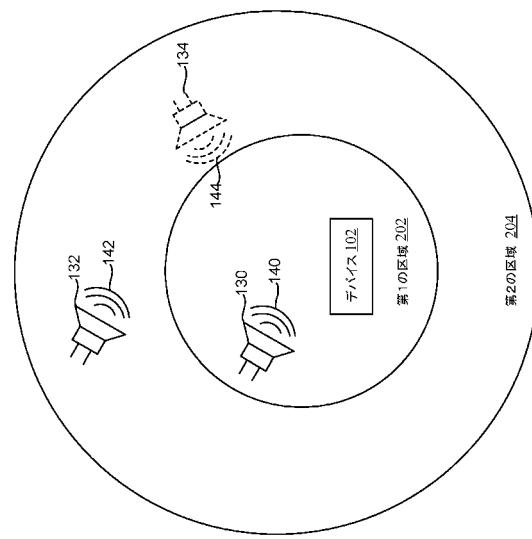


FIG. I

FIG. 2

【 図 4 A 】

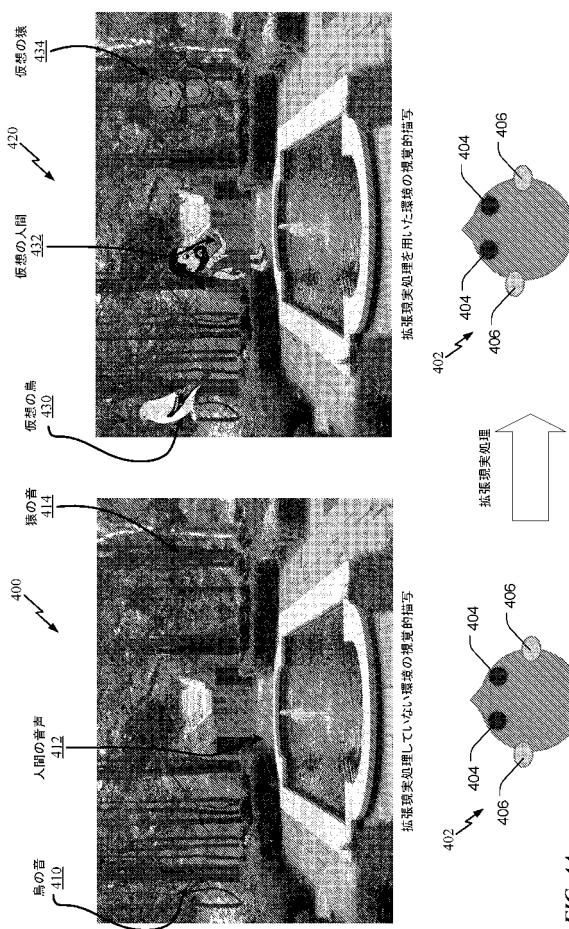
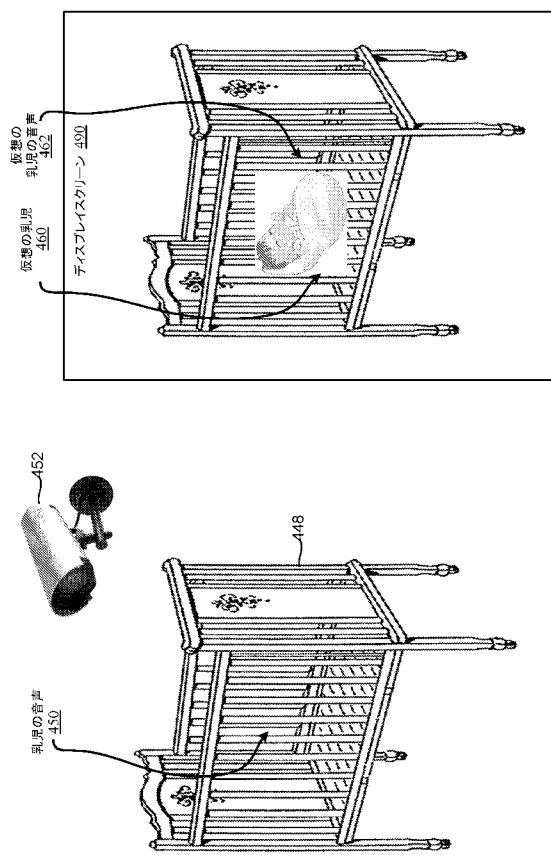


FIG. 4A

【図 4B】



【図 5】

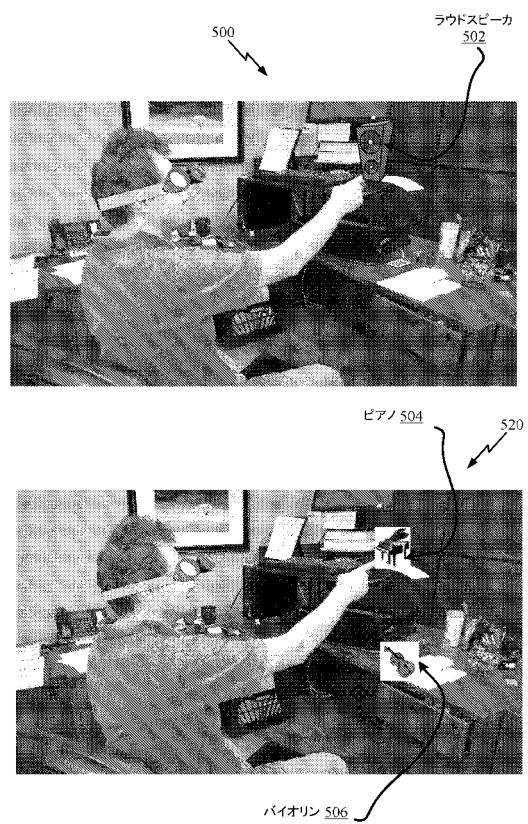


FIG. 4B

【図 7】

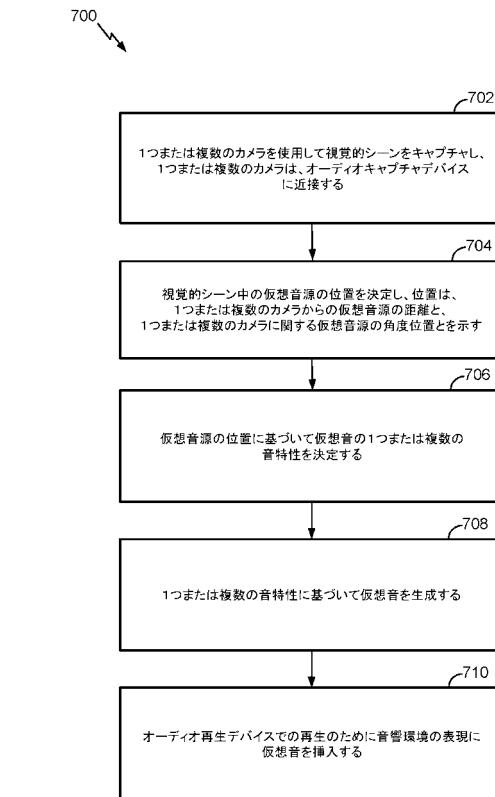


FIG. 6

FIG. 7

【図 8】

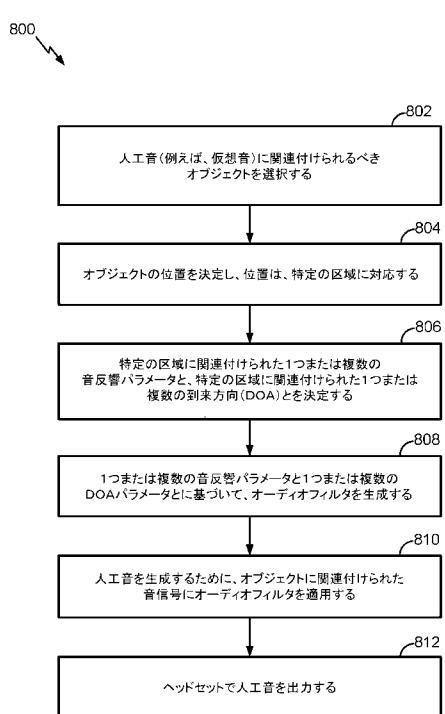


FIG. 8

【図 9】

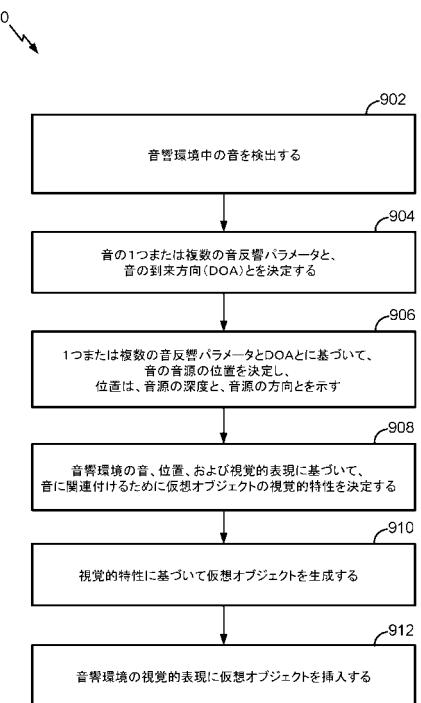


FIG. 9

【図 10】

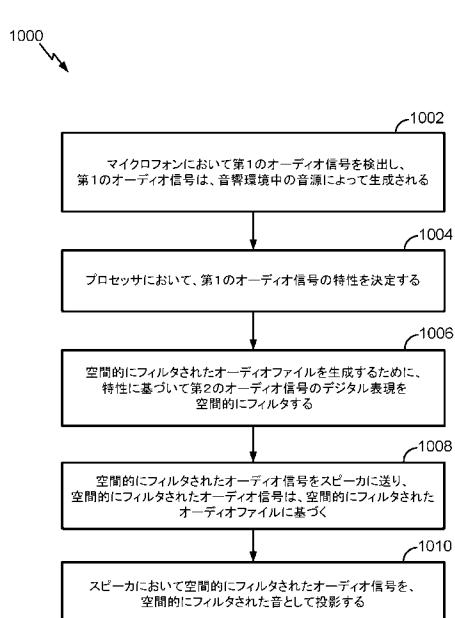


FIG. 10

【図 11】

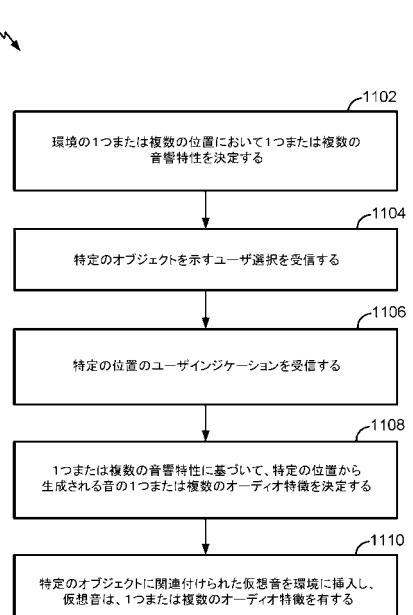


FIG. 11

【図12】

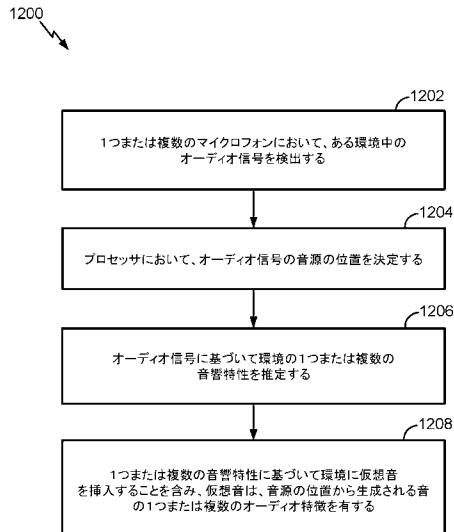


FIG. 12

【図13】

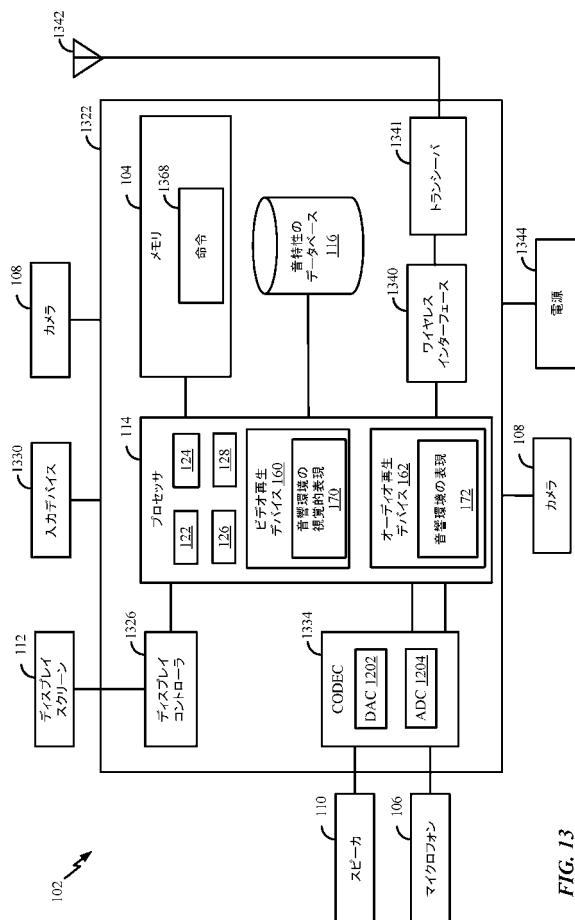


FIG. 13

## 【手続補正書】

【提出日】平成31年3月12日(2019.3.12)

## 【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

## 【補正の内容】

【特許請求の範囲】

## 【請求項1】

仮想音を出力するための装置であつて、前記装置は、  
ある環境中のオーディオ信号を検出するように構成された1つまたは複数のマイクロフォンと、

前記1つまたは複数のマイクロフォンに結合されたプロセッサと、  
を備え、前記プロセッサは、

前記オーディオ信号の音源の位置を示す第1の視覚的パラメータを決定することと、  
前記オーディオ信号に基づいて前記環境の1つまたは複数の音響特性を推定することと、

前記第1の視覚的パラメータ、前記仮想音の視覚的パラメータ、および前記1つまたは複数の音響特性に基づいて前記環境に仮想音を挿入することと、前記仮想音は、前記音源の前記位置から生成される音の1つまたは複数のオーディオ特徴を有する、  
を行つよう構成される、装置。

## 【請求項2】

前記第1の視覚的パラメータは、前記環境の視覚的表現において前記音源の前記位置を示し、前記プロセッサは、

前記環境への前記仮想音の挿入の前に、前記仮想音を生成することと、

前記第1の視覚的パラメータと前記仮想音の前記視覚的パラメータとの比較に基づいて前記環境に前記仮想音を挿入することと  
を行うようにさらに構成される、請求項1に記載の装置。

**【請求項3】**

前記プロセッサは、  
前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、  
前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、  
を行うようにさらに構成される、請求項1に記載の装置。

**【請求項4】**

前記プロセッサは、  
前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類することと、  
前記特定のオブジェクトの仮想イメージを生成することと、  
前記環境中の前記音源の前記位置において、前記仮想イメージを挿入することと、  
を行うようにさらに構成される、請求項1に記載の装置。

**【請求項5】**

前記プロセッサに結合されたカメラをさらに備え、前記カメラは、前記環境をキャプチャするように構成される、請求項1に記載の装置。

**【請求項6】**

前記カメラは、防犯カメラを含む、請求項5に記載の装置。

**【請求項7】**

前記カメラ、前記1つまたは複数のマイクロフォン、および前記プロセッサは、ヘッドセットに統合される、請求項5に記載の装置。

**【請求項8】**

前記ヘッドセットは、前記環境の視覚的描写を表示するように構成される、請求項7に記載の装置。

**【請求項9】**

前記プロセッサは、  
前記オーディオ信号の音特性を決定することと、  
前記音特性に基づいて、前記音源が前記環境の第1の区域に位置するか、または前記環境の第2の区域に位置するかを決定することと、ここにおいて、前記第1の区域は、前記第2の区域よりも前記1つまたは複数のマイクロフォンに近い、  
前記オーディオ信号の到来方向を推定することと、ここにおいて、前記位置は、前記音源に関連付けられた区域および前記到来方向に基づく、  
を行うようにさらに構成される、請求項1に記載の装置。

**【請求項10】**

前記音特性は、前記オーディオ信号の少なくとも1つの残響特性を備える、請求項9に記載の装置。

**【請求項11】**

前記少なくとも1つの残響特性は、前記オーディオ信号の直接音対残響音比を備える、請求項10に記載の装置。

**【請求項12】**

仮想音を出力するための方法であって、前記方法は、  
1つまたは複数のマイクロフォンにおいて、ある環境中のオーディオ信号を検出することと、  
プロセッサにおいて、前記オーディオ信号の音源の位置を示す第1の視覚的パラメータを決定することと、  
前記オーディオ信号に基づいて前記環境の1つまたは複数の音響特性を推定することと、  
、

前記第1の視覚的パラメータ、前記仮想音の視覚的パラメータ、および前記1つまたは複数の音響特性に基づいて前記環境に仮想音を挿入することと、前記仮想音は、前記音源の前記位置から生成される音の1つまたは複数のオーディオ特徴を有する、を備える、方法。

【請求項13】

前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、請求項12に記載の方法。

【請求項14】

前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、請求項12に記載の方法。

【請求項15】

前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類することと、前記特定のオブジェクトの仮想イメージを生成することと、

前記環境中の前記音源の前記位置において、前記仮想イメージを挿入することと、をさらに備える、請求項12に記載の方法。

【請求項16】

カメラを介して前記環境をキャプチャすることをさらに備える、請求項12に記載の方法。

【請求項17】

前記カメラは、防犯カメラを含む、請求項16に記載の方法。

【請求項18】

前記カメラ、前記1つまたは複数のマイクロフォン、および前記プロセッサは、ヘッドセットに統合される、請求項16に記載の方法。

【請求項19】

前記ヘッドセットにおいて前記環境の視覚的描写を表示することをさらに備える、請求項18に記載の方法。

【請求項20】

前記音源の前記位置を決定することは、

前記オーディオ信号の音特性を決定することと、

前記音特性に基づいて、前記音源が前記環境の第1の区域に位置するか、または前記環境の第2の区域に位置するかを決定することと、ここにおいて、前記第1の区域は、前記第2の区域よりも前記1つまたは複数のマイクロフォンに近い、

前記オーディオ信号の到来方向を推定することと、ここにおいて、前記位置は、前記音源に関連付けられた区域および前記到来方向に基づく、

を備える、請求項12に記載の方法。

【請求項21】

前記音特性は、前記オーディオ信号の少なくとも1つの残響特性を備える、請求項20に記載の方法。

【請求項22】

前記少なくとも1つの残響特性は、前記オーディオ信号の直接音対残響音比を備える、請求項21に記載の方法。

【請求項23】

仮想音を出力するための命令を備える非一時的コンピュータ可読媒体であって、前記命令は、プロセッサによって実行されたとき、前記プロセッサに、

ある環境中のオーディオ信号の音源の位置を示す第1の視覚的パラメータを決定することと、前記オーディオ信号は、1つまたは複数のマイクロフォンにおいて検出される、

前記オーディオ信号に基づいて前記環境の1つまたは複数の音響特性を推定することと

、前記第1の視覚的パラメータ、前記仮想音の視覚的パラメータ、および前記1つまたは複数の音響特性に基づいて前記環境に仮想音を挿入することと、前記仮想音は、前記音源の前記位置から生成される音の1つまたは複数のオーディオ特徴を有する、を備える動作を行わせる、非一時的コンピュータ可読媒体。

【請求項24】

前記動作は、前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、請求項23に記載の非一時的コンピュータ可読媒体。

【請求項25】

前記動作は、

前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、請求項23に記載の非一時的コンピュータ可読媒体。

【請求項26】

前記動作は、

前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類することと、

前記特定のオブジェクトの仮想イメージを生成することと、

前記環境中の前記音源の前記位置において、前記仮想イメージを挿入することと、

をさらに備える、請求項23に記載の非一時的コンピュータ可読媒体。

【請求項27】

前記動作は、カメラに、前記環境をキャプチャさせることをさらに備える、請求項23に記載の非一時的コンピュータ可読媒体。

【請求項28】

前記カメラは、防犯カメラを含む、請求項27に記載の非一時的コンピュータ可読媒体。

。

【請求項29】

仮想音を出力するための装置であって、前記装置は、

ある環境中のオーディオ信号を検出するための手段と、

前記環境中のオーディオ信号の音源の位置を示す第1の視覚的パラメータを決定するための手段と、

前記オーディオ信号に基づいて前記環境の1つまたは複数の音響特性を推定するための手段と、

前記第1の視覚的パラメータ、前記仮想音の視覚的パラメータ、および前記1つまたは複数の音響特性に基づいて前記環境に仮想音を挿入するための手段と、前記仮想音は、前記音源の前記位置から生成される音の1つまたは複数のオーディオ特徴を有する、を備える、装置。

【請求項30】

前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類するための手段と、

前記特定のオブジェクトの仮想イメージを生成するための手段と、

前記環境中の前記音源の前記位置において、前記仮想イメージを挿入するための手段と、

をさらに備える、請求項29に記載の装置。

【請求項31】

音を出力するための装置であって、前記装置は、

メモリと、

前記メモリに結合されたプロセッサと

を備え、前記プロセッサは、

環境の 1 つまたは複数の位置において 1 つまたは複数の音響特性を決定することと、特定のオブジェクトを示すユーザ選択を受信することと、

特定の位置のユーザインジケーションに関連付けられた第 1 の視覚的パラメータを決定することと、

前記 1 つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の 1 つまたは複数のオーディオ特徴を決定することと、

前記第 1 の視覚的パラメータまたは前記仮想音の視覚的パラメータのうちの少なくとも 1 つに基づいて、および前記 1 つまたは複数の音響特性に基づいて、前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入することと、前記仮想音は、前記 1 つまたは複数のオーディオ特徴を有する、

を行うように構成される、装置。

#### 【請求項 3 2】

前記プロセッサは、

ユーザの手において、前記特定の位置を位置センサから受信することと、

前記ユーザの手において、スピーカに基づいて前記 1 つまたは複数の音響特性を決定することと、

を行うようにさらに構成される、請求項 3 1 に記載の装置。

#### 【請求項 3 3】

前記プロセッサは、

前記 1 つまたは複数の音響特性をメモリ中の 1 つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

を行うようにさらに構成される、請求項 3 1 に記載の装置。

#### 【請求項 3 4】

前記プロセッサに結合されたウェアラブルセンサをさらに備え、前記ウェアラブルセンサは、

前記特定の位置を検出することと、

前記特定の位置の前記ユーザインジケーションを生成することと、

前記プロセッサに、前記特定の位置の前記ユーザインジケーションを送ることと

を行うように構成される、請求項 3 1 に記載の装置。

#### 【請求項 3 5】

前記ウェアラブルセンサは、ユーザの手に装着されるようにさらに構成される、請求項 3 4 に記載の装置。

#### 【請求項 3 6】

前記プロセッサに結合された 1 つまたは複数のスピーカと、前記 1 つまたは複数のスピーカは、前記環境の異なる位置においてオーディオ信号を生成するように構成され、

前記プロセッサに結合された 1 つまたは複数のマイクロフォンと、前記 1 つまたは複数のマイクロフォンは、各オーディオ信号を検出するように構成され、

ここにおいて、前記プロセッサは、前記 1 つまたは複数のスピーカによって生成される検出されたオーディオ信号に基づいて、前記 1 つまたは複数の位置において前記 1 つまたは複数の音響特性を決定するように構成される、

をさらに備える、請求項 3 1 に記載の装置。

#### 【請求項 3 7】

前記 1 つまたは複数のスピーカは、ユーザの手に装着されるようにさらに構成される、請求項 3 6 に記載の装置。

#### 【請求項 3 8】

音を出力するための方法であって、前記方法は、

環境の 1 つまたは複数の位置において 1 つまたは複数の音響特性を決定することと、

特定のオブジェクトを示すユーザ選択を受信することと、

特定の位置のユーザインジケーションに関連付けられた第1の視覚的パラメータを決定することと、

前記1つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定することと、

前記第1の視覚的パラメータまたは前記仮想音の視覚的パラメータのうちの少なくとも1つに基づいて、および前記1つまたは複数の音響特性に基づいて、前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入することと、前記仮想音は、前記1つまたは複数のオーディオ特徴を有する、

を備える、方法。

【請求項39】

前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、請求項38に記載の方法。

【請求項40】

前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、請求項38に記載の方法。

【請求項41】

ウェアラブルセンサを使用して前記特定の位置を検出することと、

前記ウェアラブルセンサにおいて、前記特定の位置の前記ユーザインジケーションを生成することと、

プロセッサに、前記ウェアラブルセンサからの前記特定の位置の前記ユーザインジケーションを送ることと

をさらに備える、請求項38に記載の方法。

【請求項42】

前記ウェアラブルセンサは、ユーザの手に装着されるように構成される、請求項41に記載の方法。

【請求項43】

1つまたは複数のスピーカを使用して、前記環境の異なる位置においてオーディオ信号を生成することと、

1つまたは複数のマイクロフォンを使用して、各オーディオ信号を検出することと、

ここにおいて、前記1つまたは複数の位置における前記1つまたは複数の音響特性は、前記1つまたは複数のスピーカによって生成される検出されたオーディオ信号に基づいて決定される、

をさらに備える、請求項38に記載の方法。

【請求項44】

前記1つまたは複数のスピーカは、ユーザの手に装着されるように構成される、請求項43に記載の方法。

【請求項45】

音を出力するための命令を備える非一時的コンピュータ可読媒体であって、前記命令は、プロセッサによって実行されたとき、前記プロセッサに、

環境の1つまたは複数の位置において1つまたは複数の音響特性を決定することと、

特定のオブジェクトを示すユーザ選択を受信することと、

特定の位置のユーザインジケーションに関連付けられた第1の視覚的パラメータを決定することと、

前記1つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定することと、

前記第1の視覚的パラメータまたは前記仮想音の視覚的パラメータのうちの少なくとも1つに基づいて、および前記1つまたは複数の音響特性に基づいて、前記特定のオブジェ

クトに関連付けられた仮想音を前記環境に挿入することと、前記仮想音は、前記1つまたは複数のオーディオ特徴を有する、

を備える動作を行わせる、非一時的コンピュータ可読媒体。

#### 【請求項46】

前記動作は、前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、請求項45に記載の非一時的コンピュータ可読媒体。

#### 【請求項47】

前記動作は、

前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、請求項45に記載の非一時的コンピュータ可読媒体。

#### 【請求項48】

前記特定の位置の前記ユーザインジケーションは、ウェアラブルセンサから受信される、請求項45に記載の非一時的コンピュータ可読媒体。

#### 【請求項49】

前記ウェアラブルセンサは、ユーザの手に装着されるように構成される、請求項48に記載の非一時的コンピュータ可読媒体。

#### 【請求項50】

前記動作は、

1つまたは複数のスピーカを使用して、前記環境の異なる位置においてオーディオ信号の生成を開始することと、

1つまたは複数のマイクロフォンを使用して、各オーディオ信号を検出することと、

ここにおいて、前記1つまたは複数の位置における前記1つまたは複数の音響特性は、前記1つまたは複数のスピーカによって生成される検出されたオーディオ信号に基づいて決定される、

をさらに備える、請求項45に記載の非一時的コンピュータ可読媒体。

#### 【請求項51】

前記1つまたは複数のスピーカは、ユーザの手に装着されるようにさらに構成される、請求項50に記載の非一時的コンピュータ可読媒体。

#### 【請求項52】

音を出力するための装置であって、前記装置は、

環境の1つまたは複数の位置において1つまたは複数の音響特性を決定するための手段と、

特定のオブジェクトを示すユーザ選択を受信するための手段と、

特定の位置のユーザインジケーションに関連付けられた第1の視覚的パラメータを決定するための手段と、

前記1つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定するための手段と、

前記第1の視覚的パラメータまたは前記仮想音の視覚的パラメータのうちの少なくとも1つに基づいて、および前記1つまたは複数の音響特性に基づいて、前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入するための手段と、前記仮想音は、前記1つまたは複数のオーディオ特徴を有する、

を備える、装置。

#### 【請求項53】

前記環境への前記仮想音の挿入の前に、前記仮想音を生成するための手段をさらに備える、請求項52に記載の装置。

#### 【請求項54】

前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較するため

の手段と、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出するための手段と、

をさらに備える、請求項 5 2 に記載の装置。

【請求項 5 5】

前記特定の位置を検出することと、前記特定の位置の前記ユーザインジケーションを生成することと、前記ユーザインジケーションを受信するための前記手段に前記特定の位置の前記ユーザインジケーションを送ることと、を行うための手段をさらに備える、請求項 5 2 に記載の装置。

【請求項 5 6】

前記特定の位置を検出するための前記手段は、ユーザの手に装着されるように構成される、請求項 5 5 に記載の装置。

【請求項 5 7】

前記環境の異なる位置においてオーディオ信号を生成するための手段と、

各オーディオ信号を検出するための手段と、

ここにおいて、前記 1 つまたは複数の位置における前記 1 つまたは複数の音響特性は、オーディオ信号を生成するための前記手段によって生成される検出されたオーディオ信号に基づいて決定される、

をさらに備える、請求項 5 2 に記載の装置。

【請求項 5 8】

前記オーディオ信号を生成するための前記手段は、ユーザの手に装着されるように構成される、請求項 5 7 に記載の装置。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0 1 3 4

【補正方法】変更

【補正の内容】

【0 1 3 4】

[00147] 開示された実施形態の先の説明は、当業者が、開示された態様形態を製造または使用することを可能にするために提供される。これらの態様に対する様々な修正は、当業者に対して容易に明らかであり、本明細書で定義される原理は、本開示の範囲から逸脱することなく他の態様に適用され得る。よって、本開示は、本明細書で示される態様に限定されることを意図するものではなく、下記の特許請求の範囲で定義されるような原理および新規な特徴と一致し得る最も広い範囲を与えられるべきである。

以下に本願の出願当初の特許請求の範囲に記載された発明を付記する。

[ C 1 ]

仮想音を出力するための装置であって、前記装置は、

ある環境中のオーディオ信号を検出するように構成された 1 つまたは複数のマイクロフォンと、

前記 1 つまたは複数のマイクロフォンに結合されたプロセッサと、

を備え、前記プロセッサは、

前記オーディオ信号の音源の位置を決定することと、

前記オーディオ信号に基づいて前記環境の 1 つまたは複数の音響特性を推定することと、

前記 1 つまたは複数の音響特性に基づいて前記環境に仮想音を挿入することと、前記仮想音は、前記音源の前記位置から生成される音の 1 つまたは複数のオーディオ特徴を有する、

を行うように構成される、装置。

[ C 2 ]

前記プロセッサは、前記環境への前記仮想音の挿入の前に、前記仮想音を生成するよう

にさらに構成される、C 1 に記載の装置。

[ C 3 ]

前記プロセッサは、

前記 1 つまたは複数の音響特性をメモリ中の 1 つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

を行うようにさらに構成される、C 1 に記載の装置。

[ C 4 ]

前記プロセッサは、

前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類することと、

前記特定のオブジェクトの仮想イメージを生成することと、

前記環境中の前記音源の前記位置において、前記仮想イメージを挿入することと、

を行うようにさらに構成される、C 1 に記載の装置。

[ C 5 ]

前記プロセッサに結合されたカメラをさらに備え、前記カメラは、前記環境をキャプチャするように構成される、C 1 に記載の装置。

[ C 6 ]

前記カメラは、防犯カメラを含む、C 5 に記載の装置。

[ C 7 ]

前記カメラ、前記 1 つまたは複数のマイクロフォン、および前記プロセッサは、ヘッドセットに統合される、C 5 に記載の装置。

[ C 8 ]

前記ヘッドセットは、前記環境の視覚的描写を表示するように構成される、C 7 に記載の装置。

[ C 9 ]

前記プロセッサは、

前記オーディオ信号の音特性を決定することと、

前記音特性に基づいて、前記音源が前記環境の第 1 の区域に位置するか、または前記環境の第 2 の区域に位置するかを決定することと、ここにおいて、前記第 1 の区域は、前記第 2 の区域よりも前記 1 つまたは複数のマイクロフォンに近い、

前記オーディオ信号の到来方向を推定することと、ここにおいて、前記位置は、前記音源に関連付けられた区域および前記到来方向に基づく、

を行うようにさらに構成される、C 1 に記載の装置。

[ C 10 ]

前記音特性は、前記オーディオ信号の少なくとも 1 つの残響特性を備える、C 9 に記載の装置。

[ C 11 ]

前記少なくとも 1 つの残響特性は、前記オーディオ信号の直接音対残響音比を備える、C 10 に記載の装置。

[ C 12 ]

仮想音を出力するための方法であって、前記方法は、

1 つまたは複数のマイクロフォンにおいて、ある環境中のオーディオ信号を検出することと、

プロセッサにおいて、前記オーディオ信号の音源の位置を決定することと、

前記オーディオ信号に基づいて前記環境の 1 つまたは複数の音響特性を推定することと、

前記 1 つまたは複数の音響特性に基づいて前記環境に仮想音を挿入することと、前記仮想音は、前記音源の前記位置から生成される音の 1 つまたは複数のオーディオ特徴を有する、

を備える、方法。

[ C 1 3 ]

前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、C 1 2 に記載の方法。

[ C 1 4 ]

前記 1 つまたは複数の音響特性をメモリ中の 1 つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、C 1 2 に記載の方法。

[ C 1 5 ]

前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類することと、前記特定のオブジェクトの仮想イメージを生成することと、

前記環境中の前記音源の前記位置において、前記仮想イメージを挿入することと、をさらに備える、C 1 2 に記載の方法。

[ C 1 6 ]

カメラを介して前記環境をキャプチャすることをさらに備える、C 1 2 に記載の方法。

[ C 1 7 ]

前記カメラは、防犯カメラを含む、C 1 6 に記載の方法。

[ C 1 8 ]

前記カメラ、前記 1 つまたは複数のマイクロフォン、および前記プロセッサは、ヘッドセットに統合される、C 1 6 に記載の方法。

[ C 1 9 ]

前記ヘッドセットにおいて前記環境の視覚的描写を表示することをさらに備える、C 1 8 に記載の方法。

[ C 2 0 ]

前記音源の前記位置を決定することは、

前記オーディオ信号の音特性を決定することと、

前記音特性に基づいて、前記音源が前記環境の第 1 の区域に位置するか、または前記環境の第 2 の区域に位置するかを決定することと、ここにおいて、前記第 1 の区域は、前記第 2 の区域よりも前記 1 つまたは複数のマイクロフォンに近い、

前記オーディオ信号の到来方向を推定することと、ここにおいて、前記位置は、前記音源に関連付けられた区域および前記到来方向に基づく、

を備える、C 1 2 に記載の方法。

[ C 2 1 ]

前記音特性は、前記オーディオ信号の少なくとも 1 つの残響特性を備える、C 2 0 に記載の方法。

[ C 2 2 ]

前記少なくとも 1 つの残響特性は、前記オーディオ信号の直接音対残響音比を備える、C 2 1 に記載の方法。

[ C 2 3 ]

仮想音を出力するための命令を備える非一時的コンピュータ可読媒体であって、前記命令は、プロセッサによって実行されたとき、前記プロセッサに、

ある環境中のオーディオ信号の音源の位置を決定することと、前記オーディオ信号は、1 つまたは複数のマイクロフォンにおいて検出される、

前記オーディオ信号に基づいて前記環境の 1 つまたは複数の音響特性を推定することと、

前記 1 つまたは複数の音響特性に基づいて前記環境に仮想音を挿入することと、前記仮想音は、前記音源の前記位置から生成される音の 1 つまたは複数のオーディオ特徴を有する、

を備える動作を行わせる、非一時的コンピュータ可読媒体。

[ C 2 4 ]

前記動作は、前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、C 2 3 に記載の非一時的コンピュータ可読媒体。

[ C 2 5 ]

前記動作は、

前記 1 つまたは複数の音響特性をメモリ中の 1 つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、C 2 3 に記載の非一時的コンピュータ可読媒体。

[ C 2 6 ]

前記動作は、

前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類することと、

前記特定のオブジェクトの仮想イメージを生成することと、

前記環境中の前記音源の前記位置において、前記仮想イメージを挿入することと、

をさらに備える、C 2 3 に記載の非一時的コンピュータ可読媒体。

[ C 2 7 ]

前記動作は、カメラに、前記環境をキャプチャさせることをさらに備える、C 2 3 に記載の非一時的コンピュータ可読媒体。

[ C 2 8 ]

前記カメラは、防犯カメラを含む、C 2 7 に記載の非一時的コンピュータ可読媒体。

[ C 2 9 ]

仮想音を出力するための装置であって、前記装置は、

ある環境中のオーディオ信号を検出するための手段と、

前記環境中のオーディオ信号の音源の位置を決定するための手段と、

前記オーディオ信号に基づいて前記環境の 1 つまたは複数の音響特性を推定するための手段と、

前記 1 つまたは複数の音響特性に基づいて前記環境に仮想音を挿入するための手段と、前記仮想音は、前記音源の前記位置から生成される音の 1 つまたは複数のオーディオ特徴を有する、

を備える、装置。

[ C 3 0 ]

前記オーディオ信号に基づいて、前記音源を特定のオブジェクトに分類するための手段と、

前記特定のオブジェクトの仮想イメージを生成するための手段と、

前記環境中の前記音源の前記位置において、前記仮想イメージを挿入するための手段と、

をさらに備える、C 2 9 に記載の装置。

[ C 3 1 ]

音を出力するための装置であって、前記装置は、

メモリと、

前記メモリに結合されたプロセッサと

を備え、前記プロセッサは、

環境の 1 つまたは複数の位置において 1 つまたは複数の音響特性を決定することと、

特定のオブジェクトを示すユーザ選択を受信することと、

特定の位置のユーザインジケーションを受信することと、

前記 1 つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の 1 つまたは複数のオーディオ特徴を決定することと、

前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入することと、前記

仮想音は、前記 1 つまたは複数のオーディオ特徴を有する、  
を行うように構成される、装置。

[ C 3 2 ]

前記プロセッサは、前記環境への前記仮想音の挿入の前に、前記仮想音を生成するよう  
にさらに構成される、C 3 1 に記載の装置。

[ C 3 3 ]

前記プロセッサは、

前記 1 つまたは複数の音響特性をメモリ中の 1 つまたは複数のエントリと比較すること  
と、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮  
想音を導出することと、

を行うようにさらに構成される、C 3 1 に記載の装置。

[ C 3 4 ]

前記プロセッサに結合されたウェアラブルセンサをさらに備え、前記ウェアラブルセン  
サは、

前記特定の位置を検出することと、

前記特定の位置の前記ユーザインジケーションを生成することと、

前記プロセッサに、前記特定の位置の前記ユーザインジケーションを送ることと  
を行うように構成される、C 3 1 に記載の装置。

[ C 3 5 ]

前記ウェアラブルセンサは、ユーザの手に装着されるようにさらに構成される、C 3 4  
に記載の装置。

[ C 3 6 ]

前記プロセッサに結合された 1 つまたは複数のスピーカと、前記 1 つまたは複数のスピ  
ーカは、前記環境の異なる位置においてオーディオ信号を生成するように構成され、

前記プロセッサに結合された 1 つまたは複数のマイクロフォンと、前記 1 つまたは複数  
のマイクロフォンは、各オーディオ信号を検出するように構成され、

ここにおいて、前記プロセッサは、前記 1 つまたは複数のスピーカによって生成される  
検出されたオーディオ信号に基づいて、前記 1 つまたは複数の位置において前記 1 つまたは  
複数の音響特性を決定するように構成される、

をさらに備える、C 3 1 に記載の装置。

[ C 3 7 ]

前記 1 つまたは複数のスピーカは、ユーザの手に装着されるようにさらに構成される、  
C 3 6 に記載の装置。

[ C 3 8 ]

音を出力するための方法であって、前記方法は、

環境の 1 つまたは複数の位置において 1 つまたは複数の音響特性を決定することと、  
特定のオブジェクトを示すユーザ選択を受信することと、

特定の位置のユーザインジケーションを受信することと、

前記 1 つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の 1 つま  
たは複数のオーディオ特徴を決定することと、

前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入することと、前記仮  
想音は、前記 1 つまたは複数のオーディオ特徴を有する、

を備える、方法。

[ C 3 9 ]

前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、C  
3 8 に記載の方法。

[ C 4 0 ]

前記 1 つまたは複数の音響特性をメモリ中の 1 つまたは複数のエントリと比較すること  
と、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、C 3 8 に記載の方法。

[ C 4 1 ]

ウェアラブルセンサを使用して前記特定の位置を検出することと、

前記ウェアラブルセンサにおいて、前記特定の位置の前記ユーザインジケーションを生成することと、

プロセッサに、前記ウェアラブルセンサからの前記特定の位置の前記ユーザインジケーションを送ることと

をさらに備える、C 3 8 に記載の方法。

[ C 4 2 ]

前記ウェアラブルセンサは、ユーザの手に装着されるように構成される、C 4 1 に記載の方法。

[ C 4 3 ]

1つまたは複数のスピーカを使用して、前記環境の異なる位置においてオーディオ信号を生成することと、

1つまたは複数のマイクロフォンを使用して、各オーディオ信号を検出することと、

ここにおいて、前記1つまたは複数の位置における前記1つまたは複数の音響特性は、前記1つまたは複数のスピーカによって生成される検出されたオーディオ信号に基づいて決定される、

をさらに備える、C 3 8 に記載の方法。

[ C 4 4 ]

前記1つまたは複数のスピーカは、ユーザの手に装着されるように構成される、C 4 3 に記載の方法。

[ C 4 5 ]

音を出力するための命令を備える非一時的コンピュータ可読媒体であって、前記命令は、プロセッサによって実行されたとき、前記プロセッサに、

環境の1つまたは複数の位置において1つまたは複数の音響特性を決定することと、

特定のオブジェクトを示すユーザ選択を受信することと、

特定の位置のユーザインジケーションを受信することと、

前記1つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定することと、

前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入することと、前記仮想音は、前記1つまたは複数のオーディオ特徴を有する、

を備える動作を行わせる、非一時的コンピュータ可読媒体。

[ C 4 6 ]

前記動作は、前記環境への前記仮想音の挿入の前に、前記仮想音を生成することをさらに備える、C 4 5 に記載の非一時的コンピュータ可読媒体。

[ C 4 7 ]

前記動作は、

前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較することと、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出することと、

をさらに備える、C 4 5 に記載の非一時的コンピュータ可読媒体。

[ C 4 8 ]

前記特定の位置の前記ユーザインジケーションは、ウェアラブルセンサから受信される、C 4 5 に記載の非一時的コンピュータ可読媒体。

[ C 4 9 ]

前記ウェアラブルセンサは、ユーザの手に装着されるように構成される、C 4 8 に記載

の非一時的コンピュータ可読媒体。

[ C 5 0 ]

前記動作は、

1つまたは複数のスピーカを使用して、前記環境の異なる位置においてオーディオ信号の生成を開始することと、

1つまたは複数のマイクロフォンを使用して、各オーディオ信号を検出することと、

ここにおいて、前記1つまたは複数の位置における前記1つまたは複数の音響特性は、前記1つまたは複数のスピーカによって生成される検出されたオーディオ信号に基づいて決定される、

をさらに備える、C 4 5に記載の非一時的コンピュータ可読媒体。

[ C 5 1 ]

前記1つまたは複数のスピーカは、ユーザの手に装着されるようにさらに構成される、C 5 0に記載の非一時的コンピュータ可読媒体。

[ C 5 2 ]

音を出力するための装置であって、前記装置は、

環境の1つまたは複数の位置において1つまたは複数の音響特性を決定するための手段と、

特定のオブジェクトを示すユーザ選択を受信するための手段と、

特定の位置のユーザインジケーションを受信するための手段と、

前記1つまたは複数の音響特性に基づいて、前記特定の位置から生成される音の1つまたは複数のオーディオ特徴を決定するための手段と、

前記特定のオブジェクトに関連付けられた仮想音を前記環境に挿入するための手段と、前記仮想音は、前記1つまたは複数のオーディオ特徴を有する、

を備える、装置。

[ C 5 3 ]

前記環境への前記仮想音の挿入の前に、前記仮想音を生成するための手段をさらに備える、C 5 2に記載の装置。

[ C 5 4 ]

前記1つまたは複数の音響特性をメモリ中の1つまたは複数のエントリと比較するための手段と、各エントリは、異なる仮想音に関連付けられ、

前記環境への前記仮想音の挿入の前に、前記比較に基づいて特定のエントリから前記仮想音を導出するための手段と、

をさらに備える、C 5 2に記載の装置。

[ C 5 5 ]

前記特定の位置を検出することと、前記特定の位置の前記ユーザインジケーションを生成することと、前記ユーザインジケーションを受信するための前記手段に前記特定の位置の前記ユーザインジケーションを送ることと、を行うための手段をさらに備える、C 5 2に記載の装置。

[ C 5 6 ]

前記特定の位置を検出するための前記手段は、ユーザの手に装着されるように構成される、C 5 5に記載の装置。

[ C 5 7 ]

前記環境の異なる位置においてオーディオ信号を生成するための手段と、

各オーディオ信号を検出するための手段と、

ここにおいて、前記1つまたは複数の位置における前記1つまたは複数の音響特性は、オーディオ信号を生成するための前記手段によって生成される検出されたオーディオ信号に基づいて決定される、

をさらに備える、C 5 2に記載の装置。

[ C 5 8 ]

前記オーディオ信号を生成するための前記手段は、ユーザの手に装着されるように構成

される、C 5 7 に記載の装置。

## 【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No PCT/US2017/034522									
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> INV. G06F3/01 H04S7/00 G06F3/16 ADD.											
According to International Patent Classification (IPC) or to both national classification and IPC											
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) G06F H04S											
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched											
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data											
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left; padding: 2px;">Category*</th> <th style="text-align: left; padding: 2px;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="text-align: left; padding: 2px;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td style="text-align: center; padding: 2px;">X</td> <td style="text-align: left; padding: 2px;">US 2016/088417 A1 (KIM HONGKOOK [KR] ET AL) 24 March 2016 (2016-03-24)</td> <td style="text-align: left; padding: 2px;">1,2,4-8, 12,13, 15-19, 23,24, 26-32, 34,38, 39,41, 45,46, 48,52, 53,55</td> </tr> <tr> <td style="text-align: center; padding: 2px;">Y</td> <td style="text-align: left; padding: 2px;">abstract paragraph [0002] - paragraph [0015] paragraph [0026] - paragraph [0075]; figures 1-9</td> <td style="text-align: left; padding: 2px;">3,9-11, 14, 20-22, 25,33, 35-37, 40, 42-44, 47, 49-51, -/-/       </td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	X	US 2016/088417 A1 (KIM HONGKOOK [KR] ET AL) 24 March 2016 (2016-03-24)	1,2,4-8, 12,13, 15-19, 23,24, 26-32, 34,38, 39,41, 45,46, 48,52, 53,55	Y	abstract paragraph [0002] - paragraph [0015] paragraph [0026] - paragraph [0075]; figures 1-9	3,9-11, 14, 20-22, 25,33, 35-37, 40, 42-44, 47, 49-51, -/-/
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.									
X	US 2016/088417 A1 (KIM HONGKOOK [KR] ET AL) 24 March 2016 (2016-03-24)	1,2,4-8, 12,13, 15-19, 23,24, 26-32, 34,38, 39,41, 45,46, 48,52, 53,55									
Y	abstract paragraph [0002] - paragraph [0015] paragraph [0026] - paragraph [0075]; figures 1-9	3,9-11, 14, 20-22, 25,33, 35-37, 40, 42-44, 47, 49-51, -/-/									
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.		<input checked="" type="checkbox"/> See patent family annex.									
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed											
*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family											
Date of the actual completion of the international search  13 October 2017		Date of mailing of the international search report  25/10/2017									
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer  Hasnas, Sergiu									

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US2017/034522

## Box No. II Observations where certain claims were found unsearchable (Continuation of Item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
  
  
  
  
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
  
  
  
  
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box No. III Observations where unity of invention is lacking (Continuation of Item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
  
  
  
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

## Remark on Protest

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No
PCT/US2017/034522

## C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 2 866 217 A1 (GN STORE NORD AS [DK]) 29 April 2015 (2015-04-29)  abstract paragraph [0001] paragraph [0011] - paragraph [0033] paragraph [0069] - paragraph [0075]; figures 1, 4 paragraph [0114] - paragraph [0123]; figures 9, 1 ----- DURAND R BEGAULT: "3-D Sound for Virtual Reality and Multimedia", NASA/TM-2000-000000, XX, XX, 1 January 2000 (2000-01-01), pages 1-246, XP002199910, page 69 - page 90; figures 3.1-3.17 ----- US 2015/003616 A1 (MIDDLEMISS SIMON [GB] ET AL) 1 January 2015 (2015-01-01) abstract paragraph [0001] - paragraph [0007] paragraph [0039] - paragraph [0045]; figures 4A-4D ----- WO 02/073287 A2 (CANON KK [JP]; YONEZAWA HIROKI [JP]; OKUNO YASUHIRO [JP]; MORITA KENJI) 19 September 2002 (2002-09-19) abstract page 1, line 6 - line 16 page 8, line 6 - page 10, line 1; figures 1, 2 ----- US 2015/092965 A1 (UMMINGER FREDERICK [US] ET AL) 2 April 2015 (2015-04-02)  abstract paragraph [0002] paragraph [0036] - paragraph [0044]; figures 1A, 1B paragraph [0080] - paragraph [0092]; figures 6A-7 -----	54,56-58 3,14,25, 33,40, 47,54 9-11, 20-22 9-11, 20-22 35,42, 49,56 36,37, 43,44, 50,51, 57,58

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2017/034522

Patent document cited in search report	Publication date	Patent family member(s)			Publication date
US 2016088417	A1 24-03-2016	KR 20160005695	A	15-01-2016	
		US 2016088417	A1	24-03-2016	
		WO 2014178479	A1	06-11-2014	
EP 2866217	A1 29-04-2015	EP 2866217	A1	29-04-2015	
		US 2015117664	A1	30-04-2015	
US 2015003616	A1 01-01-2015	NONE			
WO 02073287	A2 19-09-2002	JP 4708590	B2	22-06-2011	
		JP 2002271817	A	20-09-2002	
		WO 02073287	A2	19-09-2002	
US 2015092965	A1 02-04-2015	US 2015092965	A1	02-04-2015	
		WO 2015048551	A2	02-04-2015	

International Application No. PCT/ US2017/034522

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

**1. claims: 1-8, 12-19, 23-34, 38-41, 45-48, 52-55**

An apparatus for outputting virtual sound, wherein different virtual sounds are retrieved from a memory based on one or more acoustical characteristics of the environment.

---

**2. claims: 9-11, 20-22**

An apparatus for outputting virtual sound, wherein a zone of the environment where an audio source emitting an audio signal is located is determined based on a sound characteristic of the audio signal.

---

**3. claims: 35, 42, 49, 56**

An apparatus for outputting virtual sound, wherein an indicated location of a virtual sound is detected by using a wearable sensor worn on a user's hand.

---

**4. claims: 36, 37, 43, 44, 50, 51, 57, 58**

An apparatus for outputting virtual sound, wherein one or more speakers generate audio signals at different locations of the environment and wherein acoustical characteristics are determined for multiple locations based on said audio signals.

---

## フロントページの続き

(51) Int.CI.	F I	テーマコード(参考)
<b>A 6 3 F 13/213 (2014.01)</b>	A 6 3 F 13/213	
<b>A 6 3 F 13/655 (2014.01)</b>	A 6 3 F 13/655	
<b>A 6 3 F 13/25 (2014.01)</b>	A 6 3 F 13/25	

(81)指定国・地域 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, T, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, R, O, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ

(74)代理人 100184332

弁理士 中丸 慶洋

(72)発明者 ピッサー、エリック

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

(72)発明者 キム、レ-フン

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

(72)発明者 ペリ、ラグービール

アメリカ合衆国、カリフォルニア州 90007、ロサンゼルス、オーク・ストリート 2134

F ターム(参考) 5D162 AA13 CB07 CC08 CC22 CD11 DA02 DA06 EG02

5D220 BC05