

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7470695号
(P7470695)

(45)発行日 令和6年4月18日(2024.4.18)

(24)登録日 令和6年4月10日(2024.4.10)

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

請求項の数 16 (全27頁)

(21)出願番号	特願2021-538732(P2021-538732)	(73)特許権者	598036300
(86)(22)出願日	令和1年12月20日(2019.12.20)		テレフオンアクチーボラゲット エルエム
(65)公表番号	特表2022-515910(P2022-515910 A)		エリクソン(パブル)
(43)公表日	令和4年2月22日(2022.2.22)		スウェーデン国 ストックホルム エス - 1 6 4 8 3
(86)国際出願番号	PCT/EP2019/086877	(74)代理人	100109726
(87)国際公開番号	WO2020/144062		弁理士 園田 吉隆
(87)国際公開日	令和2年7月16日(2020.7.16)	(74)代理人	100161470
審査請求日	令和3年8月27日(2021.8.27)		弁理士 富樫 義孝
(31)優先権主張番号	62/789,617	(74)代理人	100194294
(32)優先日	平成31年1月8日(2019.1.8)		弁理士 石岡 利康
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100194320
			弁理士 藤井 亮
		(74)代理人	100150670
			弁理士 小椋 晴美

最終頁に続く

(54)【発明の名称】 仮想現実のための効率的な空間的にヘテロジーニアスなオーディオ要素

(57)【特許請求の範囲】

【請求項1】

ユーザのために空間的にヘテロジーニアスなオーディオ要素をレンダリングするための方法(1100)であって、

前記空間的にヘテロジーニアスなオーディオ要素を表す2つ以上のオーディオ信号を取得する(s1102)ことであって、前記オーディオ信号の組合せが前記空間的にヘテロジーニアスなオーディオ要素の空間像を提供する、2つ以上のオーディオ信号を取得する(s1102)ことと、

前記空間的にヘテロジーニアスなオーディオ要素に関連付けられたメタデータを取得する(s1104)ことであって、当該メタデータが前記空間的にヘテロジーニアスなオーディオ要素の空間的広がりを示す空間的広がり情報を含み、前記空間的広がり、前記空間的にヘテロジーニアスなオーディオ要素の物理的次元を表す、メタデータを取得する(s1104)ことと、

前記空間的にヘテロジーニアスなオーディオ要素の空間的広がり、前記空間的にヘテロジーニアスなオーディオ要素に対する前記ユーザの位置および/または向きを示す位置情報を使用して、前記空間的にヘテロジーニアスなオーディオ要素の修正された空間的広がりを導き出すことと、

i)取得された前記2つ以上のオーディオ信号と、i i)導き出された、前記修正された空間的広がり、i i i)前記空間的にヘテロジーニアスなオーディオ要素に対する前記ユーザの位置および/または向きを示す前記位置情報と、を使用して前記空間的にヘテ

ロジューニアスなオーディオ要素をレンダリングする (s 1 1 0 6) ことと、
を含む、方法 (1 1 0 0) 。

【請求項 2】

前記空間的にヘテロジューニアスなオーディオ要素の前記空間的広がり、前記空間的にヘテロジューニアスなオーディオ要素に対する第 1 の仮想位置または第 1 の仮想向きにおいて知覚される 1 つまたは複数の次元における前記空間的にヘテロジューニアスなオーディオ要素の知覚されるサイズに対応する、請求項 1 に記載の方法。

【請求項 3】

前記空間的にヘテロジューニアスなオーディオ要素をレンダリングすることが、前記空間的にヘテロジューニアスなオーディオ要素に対する前記ユーザの前記位置、および / または前記空間的にヘテロジューニアスなオーディオ要素の方向ベクトルに対する前記ユーザの前記向きに基づいて、前記 2 つ以上のオーディオ信号のうちの少なくとも 1 つを修正することを含む、請求項 1 または 2 に記載の方法。

10

【請求項 4】

前記メタデータが、

i) 前記 2 つ以上のオーディオ信号を取得するマイクロフォン間の間隔、デフォルト軸に対する前記マイクロフォンの向き、および前記マイクロフォンのタイプのうちの少なくとも 1 つを示すマイクロフォンセットアップ情報、

i i) 前記マイクロフォンと前記空間的にヘテロジューニアスなオーディオ要素との間の距離、および前記空間的にヘテロジューニアスなオーディオ要素の軸に対する仮想マイクロフォンの向きのうちの少なくとも 1 つを示す第 1 の関係情報、ならびに

20

i i i) 前記空間的にヘテロジューニアスなオーディオ要素に対するデフォルトの位置、および前記デフォルトの位置と前記空間的にヘテロジューニアスなオーディオ要素との間の距離のうちの少なくとも 1 つを示す第 2 の関係情報、
のうちの少なくとも 1 つをさらに含む、請求項 1 から 3 のいずれか一項に記載の方法。

【請求項 5】

導き出された、前記修正された空間的広がり、 $R E * f (d , D)$ によって決定され、 $R E$ は前記メタデータから取得された空間的広がりであり、 d は空間的にヘテロジューニアスなオーディオ要素と前記ユーザの現在の位置との間の距離であり、 D は空間的にヘテロジューニアスなオーディオ要素と前記メタデータで指定されたデフォルト位置との間の距離である、請求項 1 から 4 のいずれか一項に記載の方法。

30

【請求項 6】

前記空間的にヘテロジューニアスなオーディオ要素をレンダリングすることは、前記ユーザが移行領域にいるときに、前記 2 つ以上のオーディオ信号のうちの少なくとも 1 つに非相関信号を追加することによって、前記 2 つ以上のオーディオ信号のうちの少なくとも 1 つを修正することをさらに含む、請求項 1 から 5 のいずれか一項に記載の方法。

【請求項 7】

前記メタデータは、前記空間的にヘテロジューニアスなオーディオ要素に対する前記ユーザの位置および / または向きの変化に基づいて、前記空間的にヘテロジューニアスなオーディオ要素の空間幅を変更するかどうか、前記ユーザが前記空間的にヘテロジューニアスなオーディオ要素の一方の側から前記空間的にヘテロジューニアスなオーディオ要素の反対側に移動するとき左右のオーディオチャンネルを交換するかどうか、および、単一次元に沿って、前記空間的にヘテロジューニアスなオーディオ要素の前記空間的広がりを修正するかどうか、の少なくとも 1 つを示す空間的広がり修正情報をさらに含む、請求項 1 から 6 のいずれか一項に記載の方法。

40

【請求項 8】

前記空間的にヘテロジューニアスなオーディオ要素をレンダリングすることが、修正されたオーディオ信号を生成することを含み、

前記 2 つ以上のオーディオ信号が、前記空間的にヘテロジューニアスなオーディオ要素に対する第 1 の仮想位置および / または第 1 の仮想向きで知覚される前記空間的にヘテロジ

50

ーニアスなオーディオ要素を表し、

前記修正されたオーディオ信号が、前記空間的にヘテロジーニアスなオーディオ要素に対する第2の仮想位置および/または第2の仮想向きで知覚される前記空間的にヘテロジーニアスなオーディオ要素を表すために使用され、

前記ユーザの前記位置が前記第2の仮想位置に対応し、および/または前記ユーザの前記向きが前記第2の仮想向きに対応する、

請求項1から7のいずれか一項に記載の方法。

【請求項9】

前記2つ以上のオーディオ信号が左オーディオ信号(L)および右オーディオ信号(R)を含み、

前記空間的にヘテロジーニアスなオーディオ要素をレンダリングすることが、修正された左信号(L')および修正された右信号(R')を生成することを含み、

$[L' \ R']^T = H \times [L \ R]^T$ であり、ここで、Hは変換行列であり、

前記変換行列が、前記取得したメタデータおよび前記位置情報に基づいて決定される、請求項1から8のいずれか一項に記載の方法。

【請求項10】

前記空間的にヘテロジーニアスなオーディオ要素をレンダリングすることが、

第1の出力信号(EL)および第2の出力信号(ER)を生成することを含み、ここで、

$EL = L' * HRTFL$ であり、HRTFLが左耳の頭部伝達関数であり、

$ER = R' * HRTFR$ であり、HRTFRが右耳の頭部伝達関数である、

請求項9に記載の方法。

【請求項11】

前記空間的にヘテロジーニアスなオーディオ要素に関連付けられた前記メタデータが、

前記空間的にヘテロジーニアスなオーディオ要素の空間中心、および

前記空間的にヘテロジーニアスなオーディオ要素の方向ベクトル、

のうちの少なくとも1つを指定する、請求項1から10のいずれか一項に記載の方法。

【請求項12】

前記空間的にヘテロジーニアスなオーディオ要素をレンダリングするステップが、

1つまたは複数の修正されたオーディオ信号を生成することと、

前記修正されたオーディオ信号を含む前記オーディオ信号のバイノーラルレンダリングを行うことと、

を含む、請求項1から11のいずれか一項に記載の方法。

【請求項13】

前記空間的にヘテロジーニアスなオーディオ要素をレンダリングするステップが、

1つまたは複数の修正されたオーディオ信号を生成することと、

前記修正されたオーディオ信号を含む前記オーディオ信号を物理的なラウドスピーカ上にレンダリングすることと、

を含む、請求項1から11のいずれか一項に記載の方法。

【請求項14】

前記修正されたオーディオ信号を含む前記オーディオ信号が仮想スピーカとしてレンダリングされ、

前記仮想スピーカの位置は、前記仮想スピーカ間の角度が前記空間的にヘテロジーニアスなオーディオ要素の導き出された前記修正された空間的広がりに対応するように、前記空間的にヘテロジーニアスなオーディオ要素に対する前記ユーザの位置および/または向きの変化を検出することに対応して動的に更新される、

請求項12または13に記載の方法。

【請求項15】

ユーザのために空間的にヘテロジーニアスなオーディオ要素をレンダリングするための装置(1200)であって、

コンピュータ可読記憶媒体(1242)と、

10

20

30

40

50

前記コンピュータ可読記憶媒体に結合された処理回路(1202)であって、前記装置に、

前記空間的にヘテロジニアスなオーディオ要素を表す2つ以上のオーディオ信号を取得することであって、前記オーディオ信号の組合せが前記空間的にヘテロジニアスなオーディオ要素の空間像を提供する、2つ以上のオーディオ信号を取得することと、

前記空間的にヘテロジニアスなオーディオ要素に関連付けられたメタデータを取得することであって、当該メタデータが前記オーディオ要素の空間的広がりを示す空間的広がり情報を含み、前記空間的広がり情報は、前記空間的にヘテロジニアスなオーディオ要素の物理的次元を表す、メタデータを取得することと、

前記空間的にヘテロジニアスなオーディオ要素の空間的広がりとは、前記空間的にヘテロジニアスなオーディオ要素に対する前記ユーザの位置および/または向きを示す位置情報を使用して、前記空間的にヘテロジニアスなオーディオ要素の修正された空間的広がりを導き出すことと、

i) 取得された前記2つ以上のオーディオ信号と、i i i) 導き出された、前記修正された空間的広がりとは、i i i) 前記空間的にヘテロジニアスなオーディオ要素に対する前記ユーザの位置および/または向きを示す前記位置情報と、を使用して、前記空間的にヘテロジニアスなオーディオ要素をレンダリングすることと、

を行わせるように設定されている処理回路(1202)と、を備える、装置(1200)。

【請求項16】

請求項2から14のいずれか一項に記載の方法を実行するように設定された、請求項15に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

空間的にヘテロジニアスなオーディオ要素のレンダリングに関する実施形態が開示される。

【背景技術】

【0002】

人々は、しばしば、ある特定の表面上にまたはある特定の体積/面積内に位置する様々な音源から生成された音波の和である音を知覚する。このような表面または体積/面積は、概念的には、空間的にヘテロジニアスな性質を有する単一のオーディオ要素(すなわち、空間的広がり内に、ある特定の量の空間的な音源変動を有するオーディオ要素)と考えることができる。

【0003】

以下は、空間的にヘテロジニアスなオーディオ要素の例のリストである。

【0004】

群衆の音：規定された体積の空間内で互いに近接して立っている多くの個人によって生成され、リスナの両耳に届く音声の和。

【0005】

川の音：川の表面から生成され、リスナの両耳に届く水跳ね音の和。

【0006】

ビーチの音：ビーチの海岸線に当たる海の波によって生成され、リスナの両耳に届く音の和。

【0007】

噴水音：噴水の表面に当たる水流によって生成され、リスナの両耳に届く音の和。

【0008】

混雑した高速道路の音：多くの車によって生成され、リスナの両耳に届く音の和。

【0009】

これらの空間的にヘテロジニアスなオーディオ要素の中には、3次元(3D)空間の

10

20

30

40

50

ある特定の経路に沿ってあまり変化しない、知覚される空間的にヘテロジーニアスな性質のものがある。例えば、川のそばを歩いているリスナが知覚する川の音の性質は、リスナが川のそばを歩いても大きくは変化しない。同様に、ビーチフロントに沿って歩いているリスナによって知覚されるビーチの音の性質、または群衆の周りを歩いているリスナによって知覚される群衆の音の性質は、リスナがビーチフロントに沿って歩いても、または群衆の周りを歩いてもあまり変化しない。

【0010】

ある特定の空間的広がりを持つオーディオ要素を表現する既存の方法が存在するが、結果として得られる表現は、オーディオ要素の空間的にヘテロジーニアスな性質を維持するものではない。そのような既存の方法の1つは、モノラルオーディオオブジェクトの周囲の位置にモノラルオーディオオブジェクトの複数の複製を作成することである。モノラルオーディオオブジェクトの周囲にモノラルオーディオオブジェクトの複数の複製があると、特定のサイズを持つ空間的に均質なオーディオオブジェクトの知覚が作成される。この概念は、MPEG-H 3Dオーディオ規格の「オブジェクト拡散」および「オブジェクト発散」機能、ならびにEBUオーディオ定義モデル(ADM)規格の「オブジェクト発散」機能において使用されている。

10

【0011】

モノラルオーディオオブジェクトを使用して空間的広がりを持つオーディオ要素を表現する別のやり方は、オーディオ要素の空間的にヘテロジーニアスな性質を維持するわけではないが、2016年1月に発行された「Efficient HRTF-based Spatial Audio for Area and Volumetric Sources」と題されたIEEE Transactions on Visualization and Computer Graphics 22(4):1-1に記載されており、その全体が参照により本明細書に組み込まれる。具体的には、モノラルオーディオオブジェクトを使用して、リスナの周りの球体上にサウンドオブジェクトの面積-体積状の幾何学的形状を投影し、一对の頭部関連(HR)フィルタを用いて、球体上のサウンドオブジェクトの幾何学的投影をカバーするすべてのHRフィルタの積分として評価される音をリスナにレンダリングすることによって、空間的広がりを持つオーディオ要素を表現することができる。球体体積音源の場合、この積分は、解析解を有するが、任意の面積-体積状音源の幾何学的形状の場合、積分は、いわゆるモンテカルロ光線サンプリングを使用して、球体上に投影された音源表面をサンプリングすることによって評価される。

20

30

【0012】

既存の方法の別の1つは、モノラルオーディオ信号に加えて空間的に拡散した成分をレンダリングして、空間的に拡散した成分とモノラルオーディオ信号との組み合わせが幾分拡散したオブジェクトの知覚を作成するようにすることである。単一のモノラルオーディオオブジェクトとは対照的に、拡散オブジェクトには明確なピンポイントの位置はない。この概念は、MPEG-H 3Dオーディオ規格の「オブジェクト拡散性」機能およびEBU ADMの「オブジェクト拡散性」機能において使用されている。

【0013】

既存の方法の組み合わせも知られている。例えば、EBU ADMの「オブジェクト広がり」機能は、モノラルオーディオオブジェクトのコピーを複数作成するという概念と、拡散成分を追加するという概念と、を組み合わせている。

40

【発明の概要】

【0014】

上述したように、オーディオ要素を表現するための様々な技法が知られている。しかしながら、これらの既知の技法の大部分は、空間的に均質な性質(すなわち、オーディオ要素内に空間的な変化がない)または空間的に拡散した性質のいずれかを有するオーディオ要素しかレンダリングすることができず、これは、説得力のあるやり方で上記の例のいくつかをレンダリングするには限界が多すぎる。言い換えれば、これらの既知の技法では、明確な空間的にヘテロジーニアスな性質を持つオーディオ要素をレンダリングすることはで

50

きない。

【0015】

空間的にヘテロジーニアスなオーディオ要素の概念を作成する1つのやり方は、複数の個々のモノラルオーディオオブジェクト（基本的には個々のオーディオソース）の空間的に分散されたクラスタを作成し、複数の個々のモノラルオーディオオブジェクトを何らかのより高いレベルで（例えば、シーングラフまたはその他のグループ化メカニズムを使用して）一緒にリンクすることによるものである。しかしながら、これは、多くの場合、特に高度にヘテロジーニアスなオーディオ要素（すなわち、上記の例などの、多くの個々の音源を含むオーディオ要素）については、効率的なソリューションではない。さらに、レンダリングされるべきオーディオ要素がライブキャプチャされたコンテンツである場合、オーディオ要素を形成する複数のオーディオソースのそれぞれを別々に録音することは実現不可能または非現実的である場合もある。

10

【0016】

したがって、空間的にヘテロジーニアスなオーディオ要素の効率的な表現、および空間的にヘテロジーニアスなオーディオ要素の効率的な動的な6自由度（6DOF）レンダリングを提供するための改善された方法が必要とされている。特に、リスナによって知覚されるオーディオ要素のサイズ（例えば、幅または高さ）を、異なる聴取位置および/または向きに対応させること、および知覚される空間的性質を知覚されるサイズ内に維持することが望ましい。

【0017】

本開示の実施形態は、空間的にヘテロジーニアスなオーディオ要素の効率的な表現、および効率的かつ動的な6DOFレンダリングを可能にし、オーディオ要素のリスナに、リスナがいる仮想環境と空間的および概念的に一致した現実に近いサウンド体感を提供する。

20

【0018】

空間的にヘテロジーニアスなオーディオ要素のこの効率的かつ動的な表現および/またはレンダリングは、コンテンツ作成者にとって非常に有用であり、コンテンツ作成者は、仮想現実（VR）、拡張現実（AR）、または複合現実（MR）アプリケーションのために非常に効率的なやり方で空間的に豊富なオーディオ要素を6DOFシナリオに組み込むことができるであろう。

【0019】

本開示の一部の実施形態では、空間的にヘテロジーニアスなオーディオ要素は、組み合わせられてオーディオ要素の空間像を提供する少数（例えば、2以上であるが一般に6以下）のオーディオ信号のグループとして表される。例えば、空間的にヘテロジーニアスなオーディオ要素は、関連付けられたメタデータを有するステレオ音響信号として表現されてもよい。

30

【0020】

さらに、本開示の一部の実施形態では、レンダリングメカニズムは、空間的にヘテロジーニアスなオーディオ要素のヘテロジーニアスな空間的特性を保持しながら、空間的にヘテロジーニアスなオーディオ要素のリスナの位置および/または向きが変化するにつれ、オーディオ要素の知覚される空間的広がりが制御されたやり方で修正されるように、空間的にヘテロジーニアスなオーディオ要素の動的な6DOFレンダリングを可能にすることができる。空間的広がりのこの修正は、空間的にヘテロジーニアスなオーディオ要素のメタデータと、空間的にヘテロジーニアスなオーディオ要素に対するリスナの位置および/または向きと、に依存してもよい。

40

【0021】

一態様では、ユーザのために空間的にヘテロジーニアスなオーディオ要素をレンダリングするための方法が存在する。一部の実施形態では、本方法は、空間的にヘテロジーニアスなオーディオ要素を表す2つ以上のオーディオ信号を取得することを含み、オーディオ信号の組合せが、空間的にヘテロジーニアスなオーディオ要素の空間像を提供する。本方法はまた、空間的にヘテロジーニアスなオーディオ要素に関連付けられたメタデータを取

50

得することを含む。メタデータは、空間的にヘテロジニアスなオーディオ要素の空間的広がりを指定する空間的広がり情報を含むことができる。本方法は、i) 空間的広がり情報と、ii) 空間的にヘテロジニアスなオーディオ要素に対するユーザの位置（例えば、仮想位置）および/または向きを示す位置情報と、を使用してオーディオ要素をレンダリングすることをさらに含む。

【0022】

別の態様では、コンピュータプログラムが提供される。コンピュータプログラムは、処理回路によって実行されると、処理回路に上述した方法を実行させる命令を含む。別の態様では、キャリアが提供され、このキャリアには、コンピュータプログラムが含まれている。キャリアは、電子信号、光信号、無線信号、およびコンピュータ可読記憶媒体のうちの1つである。

10

【0023】

別の態様では、ユーザのために空間的にヘテロジニアスなオーディオ要素をレンダリングするための装置が提供される。装置は、空間的にヘテロジニアスなオーディオ要素を表す2つ以上のオーディオ信号を取得することであって、オーディオ信号の組合せが空間的にヘテロジニアスなオーディオ要素の空間像を提供する、オーディオ信号を取得することと、空間的にヘテロジニアスなオーディオ要素に関連付けられたメタデータであって、空間的にヘテロジニアスなオーディオ要素の空間的広がりを示す空間的広がり情報を含む、メタデータを取得することと、i) 空間的広がり情報と、ii) 空間的にヘテロジニアスなオーディオ要素に対するユーザの位置（例えば、仮想位置）および/または向きを示す位置情報と、を使用して、空間的にヘテロジニアスなオーディオ要素をレンダリングすることと、を行うように設定されている。

20

【0024】

一部の実施形態では、装置は、コンピュータ可読記憶媒体と、コンピュータ可読記憶媒体に結合された処理回路であって、装置に本明細書に記載された方法を実行させるように設定された、処理回路と、を備える。

【0025】

本開示の実施形態は、少なくとも以下の2つの利点を提供する。

【0026】

関連付けられた「サイズ」、「拡散」、または「拡散性」パラメータを使用してモノラルオーディオオブジェクトの「サイズ」を拡張して、結果として空間的に均質なオーディオ要素をもたらす既知のソリューションと比較して、本開示の実施形態は、明確な空間的にヘテロジニアスな性質を有するオーディオ要素の表現および6 DoFレンダリングを可能にする。

30

【0027】

空間的にヘテロジニアスなオーディオ要素を個々のモノラルオーディオオブジェクトのクラスタとして表現する既知のソリューションと比較して、本開示の実施形態に基づく空間的にヘテロジニアスなオーディオ要素の表現は、表現、トランスポート、およびレンダリングの複雑さに関してより効率的である。

【0028】

本明細書に組み込まれ、本明細書の一部を形成する添付の図面は、様々な実施形態を示す。

40

【図面の簡単な説明】

【0029】

【図1】一部の実施形態による、空間的にヘテロジニアスなオーディオ要素の表現を示す図である。

【図2】一部の実施形態による、空間的にヘテロジニアスなオーディオ要素の表現の修正を示す図である。

【図3A - C】一部の実施形態による、空間的にヘテロジニアスなオーディオ要素の空間的広がりを修正する方法を示す図である。

50

【図 4】一部の実施形態による、空間的にヘテロジーニアスなオーディオ要素をレンダリングするためのシステムを示す図である。

【図 5 A - B】一部の実施形態による仮想現実 (VR) システムを示す図である。

【図 6 A - B】一部の実施形態による、リスナの向きを決定する方法を示す図である。

【図 7 A - B】仮想スピーカの配置を修正する方法を示す図である。

【図 8】仮想スピーカの配置を修正する方法を示す図である。

【図 9】頭部伝達関数 (HRTF) フィルタのパラメータを示す図である。

【図 10】空間的にヘテロジーニアスなオーディオ要素をレンダリングするプロセスの概要を示す図である。

【図 11】一部の実施形態によるプロセスを示す流れ図である。

10

【図 12】一部の実施形態による装置のブロック図である。

【発明を実施するための形態】

【0030】

図 1 は、空間的にヘテロジーニアスなオーディオ要素 101 の表現を示す。一実施形態では、空間的にヘテロジーニアスなオーディオ要素は、ステレオオブジェクトとして表すことができる。ステレオオブジェクトは、2チャンネルステレオ (例えば、左右の) 信号および関連付けられたメタデータを含むことができる。ステレオ信号は、ステレオ音響マイクのセットアップを使用した現実のオーディオ要素 (例えば、群衆、混雑した高速道路、ビーチ) の実際のステレオ録音から、または個々の (録音または生成された) オーディオ信号をミキシング (例えば、ステレオパンニング) することによって人工的に作成したも

20

【0031】

関連付けられたメタデータは、空間的にヘテロジーニアスなオーディオ要素 101 およびその表現に関する情報を提供することができる。図 1 に示すように、メタデータは、以下の情報のうちの少なくとも 1 つまたは複数を含むことができる。すなわち、

【0032】

(1) 空間的にヘテロジーニアスなオーディオ要素の概念的な空間中心の位置 P_1 と、

【0033】

(2) 空間的にヘテロジーニアスなオーディオ要素の空間的広がり (例えば、空間幅 W) と、

30

【0034】

(3) 空間的にヘテロジーニアスなオーディオ要素を録音するために使用されるマイクロフォン 102 および 103 (仮想マイクロフォンまたは実マイクロフォンのいずれか) のセットアップ (例えば、間隔 S および向き) と、

【0035】

(4) マイクロフォン 102 および 103 のタイプ (例えば、オムニ、カーディオイド、8 の字) と、

【0036】

(5) マイクロフォン 102 および 103 と空間的にヘテロジーニアスなオーディオ要素 101 との関係、例えば、オーディオ要素 101 の表記上の中心の位置 P_1 とマイクロフォン 102 および 103 の位置 P_2 との間の距離 d 、ならびに空間的にヘテロジーニアスなオーディオ要素 101 の基準軸 (例えば、 Y 軸) に対するマイクロフォン 102 および 103 の向き (例えば、向き) と、

40

【0037】

(6) デフォルトの聴取位置 (例えば、位置 P_2) と、

【0038】

(7) P_1 と P_2 の関係 (例えば、距離 d) と、である。

【0039】

空間的にヘテロジーニアスなオーディオ要素 101 の空間的広がり、絶対サイズ (例えば、メートル単位) として、または相対サイズ (例えば、キャプチャ位置またはデフォ

50

ルトの観察位置などの参照位置に対する角度幅)として提供されてもよい。また、空間的広がり(例えば、単一の次元で空間的広がりを指定するか、またはすべての次元に対して使用される空間的広がりを指定する)単一の値として、あるいは(例えば、異なる次元に対して別々の空間的広がりを指定する)複数の値として指定されてもよい。

【0040】

一部の実施形態では、空間的広がり(例えば、噴水)の実際の物理的サイズ/寸法であってもよい。他の実施形態では、空間的広がり(例えば、リスナ)によって知覚される空間的広がりを表してもよい。例えば、オーディオ要素が海または川である場合、リスナは、海または川の全体的な幅/寸法を知覚することができず、リスナに近い海または川の一部のみを知覚することができる。このよう

10

【0041】

図2は、リスナ104の位置の動的変化に基づき、空間的にヘテロジニアスなオーディオ要素101の表現の修正を示す。図2では、リスナ104は、最初は仮想位置Aおよび最初の仮想向き(例えば、リスナ104から空間的にヘテロジニアスなオーディオ要素101への垂直方向)に位置している。位置Aは、空間的にヘテロジニアスなオーディオ要素101に対してメタデータで指定されたデフォルトの位置であってもよい(同様に、リスナ104の初期の向きは、メタデータで指定されたデフォルトの向きと等しくてもよい)。リスナの初期位置および向きがデフォルトと一致すると仮定すると、空間的にヘテロジニアスなオーディオ要素101を表すステレオ信号は、いかなる修正もなしにリスナ104に提供され得て、したがって、リスナ104は、空間的にヘテロジニアスなオーディオ要素101のデフォルトの空間的オーディオ表現を体感することになる。

20

【0042】

リスナ104が仮想位置Aから空間的にヘテロジニアスなオーディオ要素101に近い仮想位置Bに移動した場合、リスナ104の位置の変化に基づいて、リスナ104によって知覚されるオーディオ体感を変化させることが望ましい。したがって、位置Bにおいてリスナ104によって知覚される空間的にヘテロジニアスなオーディオ要素101の空間幅 W_B を、仮想位置Aにおいてリスナ104によって知覚されるオーディオ要素101の空間幅 W_A よりも広くなるように指定することが望ましい。同様に、位置Cにおいてリスナ104によって知覚されるオーディオ要素101の空間幅 W_C を、空間幅 W_A よりも狭くなるように指定することが望ましい。

30

【0043】

したがって、一部の実施形態では、リスナによって知覚される空間的にヘテロジニアスなオーディオ要素の空間的広がり(例えば、リスナ)によって知覚される空間的にヘテロジニアスなオーディオ要素に対するリスナの位置および/または向き、ならびに空間的にヘテロジニアスなオーディオ要素のメタデータ(例えば、空間的にヘテロジニアスなオーディオ要素に対するデフォルトの位置および/または向きを示す情報)に基づいて更新される。上で説明したように、空間的にヘテロジニアスなオーディオ要素のメタデータは、空間的にヘテロジニアスなオーディオ要素のデフォルトの空間的広がりに関する空間的広がり情報、空間的にヘテロジニアスなオーディオ要素の概念的な中心の位置、ならびにデフォルトの位置および/または向きを含むことができる。修正された空間的広がり(例えば、リスナ)は、デフォルトの位置およびデフォルトの向きに対するリスナの位置および向きの変化の検出に基づいて、デフォルトの空間的広がりを修正することによって取得することができる。

40

【0044】

他の実施形態では、空間的にヘテロジニアスな広がり(例えば、川、海)のあるオーディオ要素(例えば、川、海)の表現は、空間的にヘテロジニアスな広がり(例えば、川、海)のあるオーディオ要素の知覚可能な部分のみを表す。そのような実施形態では、デフォルトの空間的広がり(例えば、川、海)は、図3A~図3Cに示すように、異なるやり方で修正されてもよい。図3Aおよび図3Bに示すように、リスナ104が空間的にヘテロジニアスな広がり(例えば、川、海)のあるオーディオ要素301と一

50

緒に移動するにつれ、空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 の表現は、リスナ 104 と共に移動することができる。したがって、リスナ 104 にレンダリングされるオーディオは、基本的に、特定の軸（例えば、図 3 A の水平軸）に対するリスナ 104 の位置とは無関係である。この場合、図 3 C に示すように、リスナ 104 によって知覚される空間的広がり、リスナ 104 と空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 との間の垂直距離 d と、リスナ 104 と空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 との間の基準垂直距離 D との比較にのみ基づいて修正されてもよい。基準垂直距離 D は、空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 のメタデータから取得することができる。

【 0045 】

例えば、図 3 C を参照すると、リスナ 104 によって知覚される修正された空間的広がり、 $SE = RE * f(d, D)$ の関数に従って決定することができ、ここで、 SE は修正された空間的広がりであり、 RE は空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 のメタデータから得られるデフォルト（または基準）の空間的広がりであり、 d は空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 とリスナ 104 の現在の位置との間の垂直距離であり、 D は空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 とメタデータで指定されたデフォルトの位置との間の垂直距離であり、 f は d および D をパラメータとして有する曲線を規定する関数である。関数 f は、線形関係または非線形曲線などの多くの形状をとることができる。曲線の例を図 3 A に示す。

【 0046 】

曲線は、空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 の空間的広がりが、空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 から非常に遠い距離ではゼロに近く、ゼロに近い距離では 180 度に近いかを示すことができる。図 3 A に示すように、空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 が海などの非常に大きな現実の要素を表す場合、曲線は、リスナが海に近づくにつれ空間的広がりが徐々に増加する（リスナが海岸に到着したときに 180 度に達する）ようなものであってもよい。空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 が噴水などのより小さな現実の要素を表す場合、曲線は、空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 から遠い距離では空間的広がりが非常に狭くなるが、空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 の近くでは非常に急速に広がるように、強く非線形であってもよい。

【 0047 】

関数 f はまた、特に空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 が小さい場合、オーディオ要素のリスナの観察角度に依存してもよい。

【 0048 】

曲線は、空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 のメタデータの一部として提供されてもよく、またはオーディオレンダラに記憶または提供されてもよい。空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 の空間的広がり修正を実施することを望むコンテンツ作成者は、空間的にヘテロジェニアスな広がりのあるオーディオ要素 301 の所望のレンダリングに基づいて、曲線の様々な形状間の選択が与えられ得る。

【 0049 】

図 4 は、一部の実施形態による、空間的にヘテロジェニアスなオーディオ要素をレンダリングするためのシステム 400 を示す。システム 400 は、コントローラ 401 と、左オーディオ信号 451 用の信号修正器 402 と、右オーディオ信号 452 用の信号修正器 403 と、左オーディオ信号 451 用のスピーカ 404 と、右オーディオ信号 452 用のスピーカ 405 と、を含む。左オーディオ信号 451 および右オーディオ信号 452 は、デフォルトの位置およびデフォルトの向きにおける空間的にヘテロジェニアスなオーディオ要素を表す。図 4 には、2つのオーディオ信号、2つの修正器、および2つのスピーカのみが示されているが、これは、例示のみを目的としており、本開示の実施形態を決して

10

20

30

40

50

限定するものではない。さらに、図 4 は、システム 4 0 0 が左オーディオ信号 4 5 1 および右オーディオ信号 4 5 2 を別々に受信および修正することを示しているものの、システム 4 0 0 は、左オーディオ信号 4 5 1 および右オーディオ信号 4 5 2 の内容を含む単一のステレオ信号を受信し、左オーディオ信号 4 5 1 および右オーディオ信号 4 5 2 を別々に修正することなく、ステレオ信号を修正することができる。

【 0 0 5 0 】

コントローラ 4 0 1 は、1 つまたは複数のパラメータを受信し、修正器 4 0 2 および 4 0 3 をトリガして、受信したパラメータに基づいて左右のオーディオ信号 4 5 1 および 4 5 2 に対して修正を実行するように設定されていてもよい。図 4 に示す実施形態では、受信されるパラメータは、(1) 空間的にヘテロジニアスなオーディオ要素のリスナの位置および / または向きに関する情報 4 5 3 と、(2) 空間的にヘテロジニアスなオーディオ要素のメタデータ 4 5 4 と、である。

【 0 0 5 1 】

本開示の一部の実施形態では、情報 4 5 3 は、図 5 A に示す仮想現実 (V R) システム 5 0 0 に含まれる 1 つまたは複数のセンサから提供されてもよい。図 5 A に示すように、V R システム 5 0 0 は、ユーザによって着用されるように設定されている。図 5 B に示すように、V R システム 5 0 0 は、向き検知ユニット 5 0 1 と、位置検知ユニット 5 0 2 と、システム 4 0 0 のコントローラ 4 0 1 に結合された処理ユニット 5 0 3 と、を備えることができる。向き検知ユニット 5 0 1 は、リスナの向きの変化を検出するように設定され、検出された変化に関する情報を処理ユニット 5 0 3 に提供する。一部の実施形態では、処理ユニット 5 0 3 は、向き検知ユニット 5 0 1 によって検出された向きの検出された変化が与えられると、(何らかの座標系に対する) 絶対向きを決定する。向きおよび位置を決定するための異なるシステム、例えば灯台トラッカ (ライダ) を使用する H T C V i v e システムも存在し得る。一実施形態では、向き検知ユニット 5 0 1 は、検出された向きの変化が与えられると、(何らかの座標系に対する) 絶対向きを決定することができる。この場合、処理ユニット 5 0 3 は、向き検知ユニット 5 0 1 からの絶対向きデータと、位置検知ユニット 5 0 2 からの絶対位置データとを単純に多重化することができる。一部の実施形態では、向き検知ユニット 5 0 1 は、1 つまたは複数の加速度計および / または 1 つまたは複数のジャイロスコープを備えることができる。

【 0 0 5 2 】

図 6 A および図 6 B は、リスナの向きを決定する例示的な方法を示す。

【 0 0 5 3 】

図 6 A では、リスナ 1 0 4 のデフォルトの向きは、X 軸の方向である。リスナ 1 0 4 が X - Y 平面に対して頭を持ち上げると、向き検知ユニット 5 0 1 は、X - Y 平面に対する角度 を検出する。向き検知ユニット 5 0 1 はまた、異なる軸に対するリスナ 1 0 4 の向きの変化を検出することができる。例えば、図 6 B において、リスナ 1 0 4 が X 軸に対して頭を回転させると、向き検知ユニット 5 0 1 は、X 軸に対する角度 を検出する。同様に、リスナが X 軸の周りに頭を回転させたときに得られる Y - Z 平面に対する角度 が、向き検知ユニット 5 0 1 によって検出され得る。向き検知ユニット 5 0 1 によって検出されたこれらの角度 、 、 および は、リスナ 1 0 4 の向きを表す。

【 0 0 5 4 】

図 5 B に戻ると、向き検知ユニット 5 0 1 に加えて、V R システム 5 0 0 は、位置検知ユニット 5 0 2 をさらに備えることができる。位置検知ユニット 5 0 2 は、図 2 に示すようにリスナ 1 0 4 の位置を決定する。例えば、位置検知ユニット 5 0 2 は、リスナ 1 0 4 の位置を検出することができ、検出された位置を示す位置情報は、リスナ 1 0 4 が位置 A から位置 B に移動した場合に、空間的にヘテロジニアスなオーディオ要素 1 0 1 の中心とリスナ 1 0 4 との間の距離がコントローラ 4 0 1 によって決定され得るように、位置検知ユニット 5 0 2 を介してコントローラ 4 0 1 に提供することができる。

【 0 0 5 5 】

それに応じて、向き検知ユニット 5 0 1 によって検出された角度 、 、 および 、 な

らびに位置検知ユニット502によって検出されたリスナ104の位置がVRシステム500の処理ユニット503に提供されてもよい。処理ユニット503は、検出された角度および検出された位置に関する情報をシステム400のコントローラ401に提供することができる。1)空間的にヘテロジニアスなオーディオ要素101の絶対位置および向きと、2)空間的にヘテロジニアスなオーディオ要素101の空間的広がり、3)リスナ104の絶対位置と、が与えられると、リスナ104から空間的にヘテロジニアスなオーディオ要素101までの距離ならびにリスナ104によって知覚される空間幅を評価することができる。

【0056】

図4に戻ると、メタデータ454は、様々な情報を含むことができる。メタデータ454に含まれる情報の例は、上で提供されている。情報453およびメタデータ454を受信すると、コントローラ401は、修正器402および403をトリガして、左オーディオ信号451および右オーディオ信号452を修正する。修正器402および403は、コントローラ401から提供される情報に基づいて左オーディオ信号451および右オーディオ信号452を修正し、リスナが空間的にヘテロジニアスなオーディオ要素の修正された空間的広がりを知覚するように、修正されたオーディオ信号をスピーカ404および405に出力する。

【0057】

空間的にヘテロジニアスなオーディオ要素のレンダリング

【0058】

空間的にヘテロジニアスなオーディオ要素をレンダリングする多くのやり方が存在する。空間的にヘテロジニアスなオーディオ要素をレンダリングする1つのやり方は、オーディオチャンネルのそれぞれを仮想スピーカとして表現し、仮想スピーカをバイノーラルでリスナにレンダリングするか、またはパニング技法などを使用して物理的ラウドスピーカ上にレンダリングすることである。例えば、空間的にヘテロジニアスなオーディオ要素を表す2つのオーディオ信号は、それらが、固定位置にある2つの仮想ラウドスピーカから出力されるかのように生成することができる。しかしながら、このような設定では、2つの固定ラウドスピーカからリスナへの音響伝達時間は、リスナが移動するにつれ変化する。2つの固定ラウドスピーカから出力される2つのオーディオ信号間の相関関係および時間的関係のために、このような音響伝達時間の変化は、空間的にヘテロジニアスなオーディオ要素の空間像の深刻な色付けおよび/または歪みをもたらす。

【0059】

したがって、図7Aに示す実施形態では、リスナ104が位置Aから位置Bに移動するにつれ、仮想ラウドスピーカ701および702をリスナ104から等距離に維持しながら、仮想ラウドスピーカ701および702を動的に更新する。この概念は、仮想ラウドスピーカ701および702によってレンダリングされたオーディオが、リスナ104の視点から見て空間的にヘテロジニアスなオーディオ要素101の位置および空間的広がりに一致するように、リスナ104によって知覚されることを可能にする。図7Aに示すように、仮想ラウドスピーカ701と702との間の角度は、リスナ104の視点から見て空間的にヘテロジニアスなオーディオ要素101の空間的広がり(例えば、空間幅)に常に対応するように制御することができる。言い換えれば、位置Bでの仮想ラウドスピーカ701および702とリスナ104との間の距離が、位置Aでの仮想ラウドスピーカ701および702とリスナ104との間の距離と同じであったとしても、リスナが位置Aから位置Bに移動するにつれ、仮想ラウドスピーカ701と702との間の角度は、AからBに変化する。この角度の変化がリスナ104によって知覚される空間幅の減少に対応する。

【0060】

仮想ラウドスピーカ701および702の位置ならびに向きはまた、リスナ104の頭の姿勢に基づいて制御されてもよい。図8は、仮想ラウドスピーカ701および702が、リスナ104の頭の姿勢に基づいてどのように制御され得るかの一例を示

10

20

30

40

50

す。図 8 に示す実施形態では、リスナ 104 が頭を傾けると、仮想ラウドスピーカ 701 および 702 の位置は、ステレオ信号のステレオ幅が空間的にヘテロジーニアスなオーディオ要素 101 の高さまたは幅に対応し得るように制御される。

【0061】

本開示の他の実施形態では、仮想ラウドスピーカ 701 と 702 との間の角度は、特定の角度（例えば、+ または - 30 度の標準ステレオ角度）に固定されている場合があり、リスナ 104 によって知覚される空間的にヘテロジーニアスなオーディオ要素 101 の空間幅は、仮想ラウドスピーカ 701 および 702 から放出される信号を修正することによって変化してもよい。例えば、図 7B において、リスナ 104 が位置 A から位置 B に移動した場合であっても、仮想ラウドスピーカ 701 と 702 との間の角度は同じままである。したがって、仮想ラウドスピーカ 701 と 702 との間の角度は、リスナ 104 の修正された視点から見た空間的にヘテロジーニアスなオーディオ要素 101 の空間的広がりにはもはや対応しない。しかしながら、仮想ラウドスピーカ 701 および 702 から放出されるオーディオ信号が修正されるため、空間的にヘテロジーニアスなオーディオ要素 101 の空間的広がり、位置 B においてリスナ 104 によって異なって知覚されることになる。本方法は、リスナの位置の変化に起因して空間的にヘテロジーニアスなオーディオ要素 101 の知覚される空間的広がりが増加するとき（例えば、空間的にヘテロジーニアスなオーディオ要素 101 に近づくかまたは遠ざかるときに、あるいはメタデータが異なる観察角度に対して空間的にヘテロジーニアスなオーディオ要素に対して異なる空間的広がりを指定するときに）、望ましくないアーチファクトが生じないという利点を有する。

【0062】

図 7B に示す実施形態では、リスナ 104 によって知覚される空間的にヘテロジーニアスなオーディオ要素 101 の空間的広がり、オーディオ要素 101 の左右のオーディオ信号にリミックス操作を施すことによって制御されてもよい。例えば、修正された左右のオーディオ信号は、以下のように表すことができる。

$$L' = H_L L + H_{LR} R \text{ および } R' = H_{RL} L + H_R R$$

または
行列表記では $(L' \ R')^T = H^* (L \ R)^T$

ここで、L および R は、デフォルト表現におけるオーディオ要素 101 についてのデフォルトの左および右のオーディオ信号であり、L' および R' は、リスナ 104 の修正された位置および/または向きにおいて知覚されるオーディオ要素 101 に対する修正された左および右のオーディオ信号である。H は、デフォルトの左右のオーディオ信号を修正された左右のオーディオ信号に変換するための変換行列である。

【0063】

変換行列 H は、空間的にヘテロジーニアスなオーディオ要素 101 に対するリスナ 104 の位置および/または向きに依存してもよい。さらに、変換行列 H はまた、空間的にヘテロジーニアスなオーディオ要素 101 のメタデータに含まれる情報（例えば、オーディオ信号を録音するために使用されるマイクロフォンのセットアップに関する情報）に基づいて決定されてもよい。

【0064】

変換行列 H を実施するために、多くの異なる混合アルゴリズムおよびそれらの組合せを使用することができる。一部の実施形態では、変換行列 H は、ステレオ信号のステレオ像を広げるおよび/または狭めるために既知のアルゴリズムのうちの 1 つまたは複数によって実施されてもよい。アルゴリズムは、空間的にヘテロジーニアスなオーディオ要素のリスナが空間的にヘテロジーニアスなオーディオ要素に近づくか、または遠ざかるときに、空間的にヘテロジーニアスなオーディオ要素の知覚されるステレオ幅を修正するのに適している可能性がある。

【0065】

このようなアルゴリズムの一例は、ステレオ信号を和信号と差信号（「ミッド」信号と「サイド」信号とも呼ばれる）に分解し、これら 2 つの信号のバランスを変化させて、オーディオ要素のステレオ像の制御可能な幅を達成することである。一部の実施形態では、

空間的にヘテロジーニアスなオーディオ要素の元のステレオ表現は、すでに和差（またはミッド・サイド）フォーマットである場合があり、その場合は、上述した分解ステップは、必要でない場合がある。

【 0 0 6 6 】

例えば、図 2 を参照すると、参照位置 A において、和信号と差信号を等しい割合で混合することができ（左右の信号において差信号の極性を逆にして）、結果としてデフォルトの左および右の信号が得られる。しかしながら、位置 A よりも空間的にヘテロジーニアスなオーディオ要素 1 0 1 に近い位置 B では、和信号よりも差信号により多くの重みを与えることで、結果として、デフォルトのものよりも広い空間像が得られる。一方、位置 A よりも空間的にヘテロジーニアスなオーディオ要素 1 0 1 から離れている位置 C では、差信号よりも和信号により多くの重みを与えることで、結果として、より狭い空間像が得られる。したがって、知覚される空間幅は、和信号と差信号との間のバランスを制御することによって、リスナ 1 0 4 と空間的にヘテロジーニアスなオーディオ要素 1 0 1 との間の距離の変化に応じて制御することができる。

10

【 0 0 6 7 】

前述した技法はまた、リスナと空間的にヘテロジーニアスなオーディオ要素との間の相対角度が変化したときに、すなわち、リスナの観察角度が変化したときに、空間的にヘテロジーニアスなオーディオ要素の空間幅を修正するために使用されてもよい。図 2 は、空間的にヘテロジーニアスなオーディオ要素 1 0 1 から参照位置 A と同じ距離にあるが、異なる角度にあるユーザ 1 0 4 の位置 D を示す。図 2 に示すように、位置 D では、位置 A よりも狭い空間像が予想され得る。この異なる空間像は、和信号と差信号の相対的な比率を変化させることによってレンダリングすることができる。具体的には、位置 D に対してより少ない差信号が使用されて、結果としてより狭い像が得られる。

20

【 0 0 6 8 】

本開示の一部の実施形態では、その全体が参照により本明細書に組み込まれる米国特許第 7, 4 4 0, 5 7 5 号、米国特許出願公開第 2 0 1 0 / 0 0 4 0 2 4 3 A 1 号、および W I P O 特許公開第 2 0 0 9 1 0 2 7 5 0 A 1 号に記載されているように、非相関技法を使用して、ステレオ信号の空間幅を増大させることができる。

【 0 0 6 9 】

本開示の他の実施形態では、その全体が参照により本明細書に組み込まれる米国特許第 8, 6 6 0, 2 7 1 号、米国特許出願公開第 2 0 1 1 / 0 1 9 4 7 1 2 号、米国特許第 6, 9 2 8, 1 6 8 号、米国特許第 5, 8 9 2, 8 3 0 号、米国特許出願公開第 2 0 0 9 / 0 1 3 6 0 6 6 号、米国特許第 9, 3 9 8, 3 9 1 B 2 号、米国特許第 7, 4 4 0, 5 7 5 号、および独国特許公開第 3 8 4 0 7 6 6 A 1 号に記載されているように、ステレオ像を広げるおよび/または狭める異なる技法を使用することができる。

30

【 0 0 7 0 】

リミックス処理（上述した例示的なアルゴリズムを含む）にはフィルタリング操作が含まれる場合があるため、一般に変換行列 H は、複雑で、周波数に依存することに留意されたい。変換は、変換領域信号に対して、潜在的なフィルタリング操作（畳み込み）を含む時間領域において適用されてもよく、あるいは同様の形態で、変換領域、例えば離散フーリエ変換（DFT）または変形離散コサイン変換（MDCT）領域において適用されてもよい。

40

【 0 0 7 1 】

一部の実施形態では、空間的にヘテロジーニアスなオーディオ要素は、単一の頭部伝達関数（HRTF）フィルタ対を使用してレンダリングされてもよい。図 9 は、HRTF フィルタの方位角（ θ ）および仰角（ ϕ ）パラメータを示す。上述したように、空間的にヘテロジーニアスなオーディオ要素が左信号 L および右信号 R によって表される場合、リスナの向きおよび/または位置の変化に基づいて修正された左および右の信号は、修正された左信号 L' および修正された右信号 R' として表すことができ、ここで、 $(L' \ R')^T = H^* (L \ R)^T$ であり、H は変換行列である。これらの実施形態では、HRTF フィルタ

50

リングは、左耳オーディオ信号 E_L および右耳オーディオ信号 E_R がリスナに出力され得るように、修正された左信号 L' および修正された右信号 R' に適用される。 L および E_R は、以下のように表すことができる。

【0072】

$$E_L(\theta, \phi, x, y, z) = L'(\theta, \phi, x, y, z) * HRTF_L(\theta, \phi)$$

【0073】

$$E_R(\theta, \phi, x, y, z) = R'(\theta, \phi, x, y, z) * HRTF_R(\theta, \phi)$$

【0074】

$HRTF_L$ は、オーディオソースのリスナに対して特定の方位角 (θ) および特定の仰角 (ϕ) に位置する仮想点オーディオソースに対応した左耳 $HRTF$ フィルタである。同様に、 $HRTF_R$ は、オーディオソースのリスナに対して特定の方位角 (θ) および特定の仰角 (ϕ) に位置する仮想点オーディオソースに対応する右耳 $HRTF$ フィルタである。 x 、 y 、 z は、デフォルトの位置 (別名「デフォルト観察位置」) に対するリスナの位置を表す。特定の実施形態では、修正された左信号 L' および修正された右信号 R' は、同じ位置でレンダリングされ、すなわち、 $R = L$ および $R = L$ である。

【0075】

一部の実施形態では、アンビソニックスフォーマットが、特定の仮想ラウドスピーカセットアップのためのバイノーラルレンダリングまたはマルチチャネルフォーマットへの変換の前に、あるいはその一部として、中間フォーマットとして使用されてもよい。例えば、上述した実施形態では、修正された左および右のオーディオ信号 L' および R' は、アンビソニックス領域に変換され、次いでバイノーラルにまたはラウドスピーカ用にレンダリングされてもよい。空間的にヘテロジーニアスなオーディオ要素は、様々なやり方でアンビソニックス領域に変換されてもよい。例えば、空間的にヘテロジーニアスなオーディオ要素は、それぞれが点音源として扱われる仮想ラウドスピーカを使用してレンダリングすることができる。このような場合、仮想ラウドスピーカのそれぞれは、既知の方法を使用してアンビソニックス領域に変換され得る。

【0076】

一部の実施形態では、2016年1月に発行された「Efficient HRTF-based Spatial Audio for Area and Volumetric Sources」と題されたIEEE Transactions on Visualization and Computer Graphics 22(4):1-1に記載されているように、より高度な技法を使用して $HRTF$ を計算することができる。

【0077】

本開示の一部の実施形態では、空間的にヘテロジーニアスなオーディオ要素は、環境要素 (例えば、海もしくは川)、またはシーン内のある領域を占有する複数の物理的なエンティティから構成された概念的なエンティティ (例えば、群衆) の代わりに、複数の音源を備える単一の物理的なエンティティ (例えば、エンジン音源および排気音源を有する車) を表してもよい。上述した空間的にヘテロジーニアスなオーディオ要素をレンダリングする方法は、複数の音源を含み、明確な空間的レイアウトを有するそのような単一の物理的なエンティティにも適用可能であってもよい。例えば、リスナが車両の運転席側で車両に向かって立っていて、車両が、リスナの左側で第1の音 (例えば、車両の前側からのエンジン音) およびリスナの右側で第2の音 (例えば、車両の後側からの排気音) を生成した場合、リスナは、第1および第2の音に基づいて、車両の明確な空間的オーディオレイアウトを知覚することができる。このような場合に、リスナが車両の周りを移動して、車両の反対側 (例えば、車両の助手席側) から観察した場合でも、リスナが明確な空間的レイアウトを知覚できるようにすることが望ましい。したがって、本開示の一部の実施形態では、リスナが一方の側 (例えば、車両の運転席側) から反対側 (例えば、車両の助手席側) に移動すると、左オーディオチャンネルと右オーディオチャンネルが入れ替わる。言い換えれば、リスナが一方の側から反対側に移動するにつれ、空間的にヘテロジーニアスなオーディオ要素の空間的表現が車両の軸の周りにミラーリングされる。

10

20

30

40

50

【 0 0 7 8 】

しかしながら、リスナが一方の側から反対側に移動する瞬間に左右のチャンネルが瞬時に入れ替わる場合、リスナは、空間的にヘテロジーニアスなオーディオ要素の空間像の不連続性を知覚する可能性がある。したがって、一部の実施形態では、リスナが2つの側の間の小さな遷移領域にいる間に、少量の非相関信号が修正されたステレオミックスに追加されてもよい。

【 0 0 7 9 】

本開示の一部の実施形態では、空間的にヘテロジーニアスなオーディオ要素のレンダリングがモノラルに崩れてしまうのを防止する追加の機能が提供される。例えば、図2を参照すると、空間的にヘテロジーニアスなオーディオ要素101が、単一方向（例えば、図2の水平方向）にのみ空間的広がりを有する1次元オーディオ要素である場合、空間的にヘテロジーニアスなオーディオ要素101のレンダリングは、リスナ104が位置Eに移動すると、位置Eには空間的にヘテロジーニアスなオーディオ要素101の知覚される空間的広がりがないため、モノラルに崩れてしまう。これは、モノラルがリスナ104にとって不自然に聞こえる可能性があるため、望ましくない可能性がある。この崩壊を防止するために、本開示の実施形態では、規定された小さな領域内での空間的広がりの修正が防止されるように、位置Eの周りの空間幅または規定された小さな領域に下限を設けている。代替的または追加的に、この崩壊は、小さな遷移領域において、レンダリングされたオーディオ信号に少量の非相関信号を追加することによって防止することができる。これにより、モノラルへの不自然な崩壊が確実に生じなくなる。

【 0 0 8 0 】

本開示の一部の実施形態では、空間的にヘテロジーニアスなオーディオ要素のメタデータはまた、リスナの位置および/または向きが変化したときに、ステレオ像の異なるタイプの修正を適用すべきかどうかを示す情報を含むことができる。具体的には、特定のタイプの空間的にヘテロジーニアスなオーディオ要素については、リスナの位置および/または向きの変化に基づいて空間的にヘテロジーニアスなオーディオ要素の空間幅を変化させること、あるいはリスナが空間的にヘテロジーニアスなオーディオ要素の一方の側から空間的にヘテロジーニアスなオーディオ要素の反対側に移動するとき左右のチャンネルを入れ替えることは望ましくない場合がある。また、特定のタイプのオーディオ要素については、空間的にヘテロジーニアスなオーディオ要素の空間的広がりを1つの次元だけに沿って修正することが望ましい場合がある。

【 0 0 8 1 】

例えば、群集は、直線に沿って並ぶのではなく、通常は2D空間を占有する。したがって、空間的広がりが1次元でしか指定されていない場合、ユーザが群集の周りを移動するとき、群集の空間的にヘテロジーニアスなオーディオ要素のステレオ幅が著しく狭められるとすれば、極めて不自然になる。また、群集から来る空間的および時間的情報は、典型的にはランダムであり、あまり向き特有ではないため、群集の単一のステレオ録音は、任意の相対的なユーザ角度で群集を表現するのに完全に適している可能性がある。したがって、群集の空間的にヘテロジーニアスなオーディオ要素のためのメタデータには、群集の空間的にヘテロジーニアスなオーディオ要素のリスナの相対位置に変化があっても、群集の空間的にヘテロジーニアスなオーディオ要素のステレオ幅の修正を無効にすべきであることを示す情報が含まれていてもよい。代替的または追加的に、メタデータにはまた、リスナの相対位置に変化があった場合に、ステレオ幅の特定の修正を適用すべきであることを示す情報が含まれていてもよい。前述の情報はまた、高速道路、海、川などの巨大な現実の要素の知覚可能な部分のみを表す空間的にヘテロジーニアスなオーディオ要素のメタデータに含まれていてもよい。

【 0 0 8 2 】

本開示の他の実施形態では、特定のタイプの空間的にヘテロジーニアスなオーディオ要素のメタデータには、空間的にヘテロジーニアスなオーディオ要素の空間的広がりを指定する位置依存、方向依存、または距離依存の情報が含まれていてもよい。例えば、群衆の

10

20

30

40

50

音を表す空間的にヘテロジーニアスなオーディオ要素については、空間的にヘテロジーニアスなオーディオ要素のメタデータは、空間的にヘテロジーニアスなオーディオ要素のリスナが第1の基準点に位置するときの空間的にヘテロジーニアスなオーディオ要素の第1の特定の空間幅と、空間的にヘテロジーニアスなオーディオ要素のリスナが第1の基準点とは異なる第2の基準点に位置するときの空間的にヘテロジーニアスなオーディオ要素の第2の特定の空間幅と、を指定する情報を含むことができる。このようにして、観察角度特有の聴覚イベントはないが、観察角度特有の幅を有する空間的にヘテロジーニアスなオーディオ要素を効率的に表現することができる。

【0083】

前の段落で説明した本開示の実施形態は、1次元または2次元に沿って空間的にヘテロジーニアスな特性を有する空間的にヘテロジーニアスなオーディオ要素を使用して説明されているが、本開示の実施形態は、追加の次元のための対応するステレオ信号およびメタデータを追加することによって、3つ以上の次元に沿って空間的にヘテロジーニアスな特性を有する空間的にヘテロジーニアスなオーディオ要素に等しく適用可能である。言い換えれば、本開示の実施形態は、マルチチャンネルステレオ音響信号、すなわち、ステレオ音響パニング技法を使用するマルチチャンネル信号（したがって、ステレオ、5.1、7.x、22.2、VBAPなどを含む全スペクトル）によって表される空間的にヘテロジーニアスなオーディオ要素に適用可能である。追加的または代替的に、空間的にヘテロジーニアスなオーディオ要素は、1次元アンビソニックスBフォーマット表現で表されてもよい。

【0084】

本開示のさらなる実施形態では、空間的にヘテロジーニアスなオーディオ要素を表すステレオ音響信号は、信号の冗長性が、例えば、ジョイントステレオ符号化技法を使用することによって利用されるように符号化される。この機能は、空間的にヘテロジーニアスなオーディオ要素を複数の個々のオブジェクトのクラスタとして符号化する場合と比較して、さらなる利点を提供する。

【0085】

本開示の実施形態では、表現されるべき空間的にヘテロジーニアスなオーディオ要素は、空間的に豊富であるが、空間的にヘテロジーニアスなオーディオ要素内の様々なオーディオソースの正確な位置決めは重要ではない。しかしながら、本開示の実施形態は、1つまたは複数の重要なオーディオソースを含む空間的にヘテロジーニアスなオーディオ要素を表現するために使用することもできる。そのような場合、重要なオーディオソースは、空間的にヘテロジーニアスなオーディオ要素のレンダリングにおいて、空間的にヘテロジーニアスなオーディオ要素に重ね合わされた個々のオブジェクトとして明示的に表現されてもよい。そのような場合の例は、ある声または音が一貫して目立つ群衆（例えば、誰かがメガホンで話している）あるいは犬が吠えているビーチのシーンである。

【0086】

図10は、一部の実施形態による、空間的にヘテロジーニアスなオーディオ要素をレンダリングするプロセス1000を示す。ステップs1002は、ユーザの現在の位置および/または現在の向きを取得することを含む。ステップs1004は、空間的にヘテロジーニアスなオーディオ要素の空間的な特性付けに関する情報を取得することを含む。ステップs1006は、ユーザの現在の位置および/または現在の向きにおいて、以下の情報、すなわち、空間的にヘテロジーニアスなオーディオ要素への方向および距離、空間的にヘテロジーニアスなオーディオ要素の知覚される空間的広がり、ならびに/あるいはユーザに対する仮想オーディオソースの位置を評価することを含む。ステップs1008は、仮想オーディオソースのレンダリングパラメータを評価することを含む。レンダリングパラメータは、ヘッドホンに配信するときの仮想オーディオソースのそれぞれについてのHRフィルタの設定情報、およびラウドスピーカ設定を介して配信するときの仮想オーディオソースのそれぞれについてのラウドスピーカパニング係数を含むことができる。ステップs1010は、マルチチャンネルオーディオ信号を取得することを含む。ステップs1012は、マルチチャンネルオーディオ信号およびレンダリングパラメータに基づいて仮想オ

10

20

30

40

50

オーディオソースをレンダリングすること、およびヘッドホンまたはラウドスピーカ信号を出力することを含む。

【0087】

図11は、一実施形態によるプロセス1100を示す流れ図である。プロセス1100は、ステップs1102で開始することができる。

【0088】

ステップs1102は、空間的にヘテロジニアスなオーディオ要素を表す2つ以上のオーディオ信号を取得することを含み、オーディオ信号の組合せが空間的にヘテロジニアスなオーディオ要素の空間像を提供する。ステップs1104は、空間的にヘテロジニアスなオーディオ要素に関連付けられたメタデータを取得することを含み、メタデータは、空間的にヘテロジニアスなオーディオ要素の空間的広がりを含む。ステップs1106は、i)空間的広がり情報と、ii)空間的にヘテロジニアスなオーディオ要素に対するユーザの位置(例えば、仮想位置)および/または向きを示す位置情報と、を使用して空間的にヘテロジニアスなオーディオ要素をレンダリングすることを含む。

【0089】

一部の実施形態では、空間的にヘテロジニアスなオーディオ要素の空間的広がり、空間的にヘテロジニアスなオーディオに対して第1の仮想位置または第1の仮想向きで知覚される、1つまたは複数の次元における空間的にヘテロジニアスなオーディオ要素のサイズに対応する。

【0090】

一部の実施形態では、空間的広がり情報は、空間的にヘテロジニアスなオーディオ要素の物理的サイズまたは知覚されるサイズを指定する。

【0091】

一部の実施形態では、空間的にヘテロジニアスなオーディオ要素をレンダリングすることは、空間的にヘテロジニアスなオーディオ要素に対する(例えば、空間的にヘテロジニアスなオーディオ要素の概念的な空間中心に対する)ユーザの位置および/または空間的にヘテロジニアスなオーディオ要素の方向ベクトルに対するユーザの向きに基づいて、2つ以上のオーディオ信号のうち少なくとも1つを修正することを含む。

【0092】

一部の実施形態では、メタデータは、i)マイクロフォン(例えば、仮想マイクロフォン)間の間隔、デフォルトの軸に対するマイクロフォンの向き、および/またはマイクロフォンのタイプを示すマイクロフォンセットアップ情報、ii)マイクロフォンと空間的にヘテロジニアスなオーディオ要素との間の距離(例えば、マイクロフォンと空間的にヘテロジニアスなオーディオ要素の概念的な空間中心との間の距離)および/または空間的にヘテロジニアスなオーディオ要素の軸に対する仮想マイクロフォンの向きを示す第1の関係情報、ならびに/あるいはiii)空間的にヘテロジニアスなオーディオ要素に対する(例えば、空間的にヘテロジニアスなオーディオ要素の概念的な空間中心に対する)デフォルトの位置および/またはデフォルトの位置と空間的にヘテロジニアスなオーディオ要素との間の距離を示す第2の関係情報をさらに含む。

【0093】

一部の実施形態では、空間的にヘテロジニアスなオーディオ要素をレンダリングすることは、修正されたオーディオ信号を生成することを含み、2つ以上のオーディオ信号は、オーディオ要素に対する第1の仮想位置および/または第1の仮想向きにおいて知覚される空間的にヘテロジニアスなオーディオ要素を表し、修正されたオーディオ信号は、空間的にヘテロジニアスなオーディオ要素に対する第2の仮想位置および/または第2の仮想向きにおいて知覚される空間的にヘテロジニアスなオーディオ要素を表すために使用され、ユーザの位置が第2の仮想位置に対応し、および/またはユーザの向きが第2の仮想向きに対応する。

【0094】

10

20

30

40

50

一部の実施形態では、2つ以上のオーディオ信号は、左オーディオ信号(L)および右オーディオ信号(R)を含み、オーディオ要素をレンダリングすることは、修正された左信号(L')および修正された右信号(R')を生成することを含み、 $[L' R']^T = H \times [L R]^T$ であり、ここで、Hは変換行列であり、変換行列は、取得したメタデータおよび位置情報に基づいて決定される。

【0095】

一部の実施形態では、空間的にヘテロジニアスなオーディオ要素をレンダリングするステップは、1つまたは複数の修正されたオーディオ信号を生成することと、修正されたオーディオ信号のうちの少なくとも1つを含むオーディオ信号のバイノーラルレンダリングと、を含む。

10

【0096】

一部の実施形態では、空間的にヘテロジニアスなオーディオ要素をレンダリングすることは、第1の出力信号(E_L)および第2の出力信号(E_R)を生成することを含み、ここで、 $E_L = L * HRTFL$ であり、 $HRTFL$ は、左耳に対する頭部伝達関数(または対応するインパルス応答)であり、 $E_R = R * HRTFR$ であり、 $HRTFR$ は、右耳に対する頭部伝達関数(または対応するインパルス応答)である。2つの出力信号の生成は、インパルス応答を使用したフィルタリング演算(畳み込み)による時間領域で、または $HRTF$ の適用による離散フーリエ変換(DFT)領域などの任意の変換領域で行うことができる。

【0097】

20

一部の実施形態では、2つ以上のオーディオ信号を取得することは、複数のオーディオ信号を取得することと、複数のオーディオ信号をアンビソニックスフォーマットに変換することと、変換された複数のオーディオ信号に基づいて2つ以上のオーディオ信号を生成することと、をさらに含む。

【0098】

一部の実施形態では、空間的にヘテロジニアスなオーディオ要素に関連付けられたメタデータは、空間的にヘテロジニアスなオーディオ要素の概念的な空間中心、および/または空間的にヘテロジニアスなオーディオ要素の方向ベクトルを指定する。

【0099】

一部の実施形態では、空間的にヘテロジニアスなオーディオ要素をレンダリングするステップは、1つまたは複数の修正されたオーディオ信号を生成することと、修正されたオーディオ信号のうちの少なくとも1つを含むオーディオ信号を物理的なラウドスピーカ上にレンダリングすることと、を含む。

30

【0100】

一部の実施形態では、少なくとも1つの修正されたオーディオ信号を含むオーディオ信号は、仮想スピーカとしてレンダリングされる。

【0101】

図12は、一部の実施形態による、図4に示すシステム400を実装するための装置1200のブロック図である。図12に示すように、装置1200は、1つまたは複数のプロセッサ(P)1255(例えば、汎用マイクロプロセッサ、および/または特定用途向け集積回路(ASIC)、フィールドプログラマブルゲートアレイ(FPGA)などの1つもしくは複数の他のプロセッサ)を含むことができる処理回路(PC)1202であって、これらのプロセッサが、単一のハウジング内または単一のデータセンタ内で同じ場所に位置してもよく、あるいは地理的に分散されていてもよい、処理回路(PC)1202と、装置1200が、ネットワークインターフェース1248が接続されたネットワーク110(例えば、インターネットプロトコル(IP)ネットワーク)に接続された他のノードとの間でデータを送受信することを可能にするための送信機(Tx)1245および受信機(Rx)1247を備えるネットワークインターフェース1248と、1つもしくは複数の不揮発性記憶デバイスおよび/または1つもしくは複数の揮発性記憶デバイスを含むことができるローカル記憶ユニット(別名「データ記憶システム」)1208と、を

40

50

備えることができる。PC 1202がプログラム可能なプロセッサを含む実施形態では、コンピュータプログラム製品(CPP) 1241が提供されてもよい。CPP 1241は、コンピュータ可読命令(CRI) 1244を含むコンピュータプログラム(CP) 1243を記憶するコンピュータ可読媒体(CRM) 1242を含む。CRM 1242は、磁気媒体(例えば、ハードディスク)、光媒体、メモリデバイス(例えば、ランダムアクセスメモリ、フラッシュメモリ)などの非一時的なコンピュータ可読媒体であってもよい。一部の実施形態では、コンピュータプログラム1243のCRI 1244は、PC 1202によって実行されると、CRIが装置1200に本明細書に記載されたステップ(例えば、流れ図を参照して本明細書に記載されたステップ)を実行させるように設定されている。他の実施形態では、装置1200は、コードを必要とせずに、本明細書に記載されたステップを実行するように設定されてもよい。すなわち、例えば、PC 1202は、1つまたは複数のASICのみで構成されてもよい。したがって、本明細書に記載された実施形態の特徴は、ハードウェアおよび/またはソフトウェアで実施することができる。

10

【0102】

実施形態の概要

【0103】

A1. ユーザのために空間的にヘテロジニアスなオーディオ要素をレンダリングするための方法であって、空間的にヘテロジニアスなオーディオ要素を表す2つ以上のオーディオ信号を取得することであって、オーディオ信号の組合せが空間的にヘテロジニアスなオーディオ要素の空間像を提供する、2つ以上のオーディオ信号を取得することと、空間的にヘテロジニアスなオーディオ要素に関連付けられたメタデータを取得することであって、メタデータが空間的にヘテロジニアスなオーディオ要素の空間的広がりを示す空間的広がり情報を含む、メタデータを取得することと、i) 空間的広がり情報と、ii) 空間的にヘテロジニアスなオーディオ要素に対するユーザの位置(例えば、仮想位置)および/または向きを示す位置情報と、を使用してオーディオ信号のうち少なくとも1つを修正し、それによって少なくとも1つの修正されたオーディオ信号を生成することと、修正されたオーディオ信号を使用して、空間的にヘテロジニアスなオーディオ要素をレンダリングすることと、を含む、方法。

20

【0104】

A2. 空間的にヘテロジニアスなオーディオ要素の空間的広がりが、空間的にヘテロジニアスなオーディオ要素に対する第1の仮想位置または第1の仮想向きにおいて知覚される1つまたは複数の次元における空間的にヘテロジニアスなオーディオ要素のサイズに対応する、実施形態A1に記載の方法。

30

【0105】

A3. 空間的広がり情報が、空間的にヘテロジニアスなオーディオ要素の物理的サイズまたは知覚されるサイズを指定する、実施形態A1またはA2に記載の方法。

【0106】

A4. オーディオ信号のうち少なくとも1つを修正することが、空間的にヘテロジニアスなオーディオ要素に対する(例えば、空間的にヘテロジニアスなオーディオ要素の概念的な空間中心に対する)ユーザの位置および/または空間的にヘテロジニアスなオーディオ要素の方向ベクトルに対するユーザの向きに基づいて、オーディオ信号のうち少なくとも1つを修正することを含む、実施形態A3に記載の方法。

40

【0107】

A5. メタデータが、i) マイクロフォン(例えば、仮想マイクロフォン)間の間隔、デフォルトの軸に対するマイクロフォンの向き、および/またはマイクロフォンのタイプを示すマイクロフォン設定情報、ii) マイクロフォンと空間的にヘテロジニアスなオーディオ要素との間の距離(例えば、マイクロフォンと空間的にヘテロジニアスなオーディオ要素の概念的な空間中心との間の距離)および/または空間的にヘテロジニアスなオーディオ要素の軸に対する仮想マイクロフォンの向きを示す第1の関係情報、ならびに/あるいはiii) 空間的にヘテロジニアスなオーディオ要素に対する(例えば、空

50

間的にヘテロジーニアスなオーディオ要素の概念的な空間中心に対する) デフォルトの位置、および/またはデフォルトの位置と空間的にヘテロジーニアスなオーディオ要素との間の距離を示す第2の関係情報をさらに含む、実施形態A1からA4のいずれか一項に記載の方法。

【0108】

A6. 2つ以上のオーディオ信号が、空間的にヘテロジーニアスなオーディオ要素に対する第1の仮想位置および/または第1の仮想向きにおいて知覚される空間的にヘテロジーニアスなオーディオ要素を表し、修正されたオーディオ信号が、空間的にヘテロジーニアスなオーディオ要素に対する第2の仮想位置および/または第2の仮想向きにおいて知覚される空間的にヘテロジーニアスなオーディオ要素を表すために使用され、ユーザの位置が第2の仮想位置に対応し、および/またはユーザの向きが第2の仮想向きに対応する、実施形態A1からA5のいずれか一項に記載の方法。

10

【0109】

A7. 2つ以上のオーディオ信号が左オーディオ信号(L)および右オーディオ信号(R)を含み、修正されたオーディオ信号が修正された左信号(L')および修正された右信号(R')を含み、 $[L' \ R']^T = H \times [L \ R]^T$ であり、ここでHは変換行列であり、変換行列が、取得したメタデータおよび位置情報に基づいて決定される、実施形態A1からA6のいずれか一項に記載の方法。

【0110】

A8. 空間的にヘテロジーニアスなオーディオ要素をレンダリングすることが、第1の出力信号(E_L)および第2の出力信号(E_R)を生成することを含み、ここで、 $E_L = L^* H R T F_L$ であり、 $H R T F_L$ は左耳の頭部伝達関数(または対応するインパルス応答)であり、 $E_R = R^* H R T F_R$ であり、 $H R T F_R$ は右耳の頭部伝達関数(または対応するインパルス応答)である、実施形態A7に記載の方法。

20

【0111】

A9. 2つ以上のオーディオ信号を取得することが、複数のオーディオ信号を取得することと、複数のオーディオ信号をアンビソニックフォーマットに変換することと、変換された複数のオーディオ信号に基づいて2つ以上のオーディオ信号を生成することと、をさらに含む、実施形態A1からA8のいずれか一項に記載の方法。

【0112】

A10. 空間的にヘテロジーニアスなオーディオ要素に関連付けられたメタデータが、オーディオ要素の概念的な空間中心および/または空間的にヘテロジーニアスなオーディオ要素の方向ベクトルを指定する、実施形態A1からA9のいずれか一項に記載の方法。

30

【0113】

A11. 空間的にヘテロジーニアスなオーディオ要素をレンダリングするステップが、少なくとも1つの修正されたオーディオ信号を含むオーディオ信号のバイノーラルレンダリングを含む、実施形態A1からA10のいずれか一項に記載の方法。

【0114】

A12. 空間的にヘテロジーニアスなオーディオ要素をレンダリングするステップが、少なくとも1つの修正されたオーディオ信号を含むオーディオ信号を物理的なラウドスピーカ上にレンダリングすることを含む、実施形態A1からA10のいずれか一項に記載の方法。

40

【0115】

A13. 少なくとも1つの修正されたオーディオ信号を含むオーディオ信号が仮想スピーカとしてレンダリングされる、実施形態A11またはA12に記載の方法。

【0116】

本開示の様々な実施形態が本明細書に記載されているが(付録がある場合はそれも含む)、それらは限定ではなく、単なる例として提示されていることを理解されたい。したがって、本開示の広さおよび範囲は、上述した例示的な実施形態のいずれによっても限定されるべきではない。さらに、本明細書で別段の指示がない限り、さもなければ文脈によっ

50

て明らかに矛盾しない限り、上述した要素のすべての可能な変形形態における任意の組合せは、本開示によって包含される。

【 0 1 1 7 】

さらに、上述され、図面に示されたプロセスは、一連のステップとして示されているが、これは単に説明のために行われたものである。したがって、いくつかのステップが追加されてもよく、いくつかのステップが省略されてもよく、ステップの順序が並び替えられてもよく、いくつかのステップが並行して実行されてもよいことが企図される。

10

20

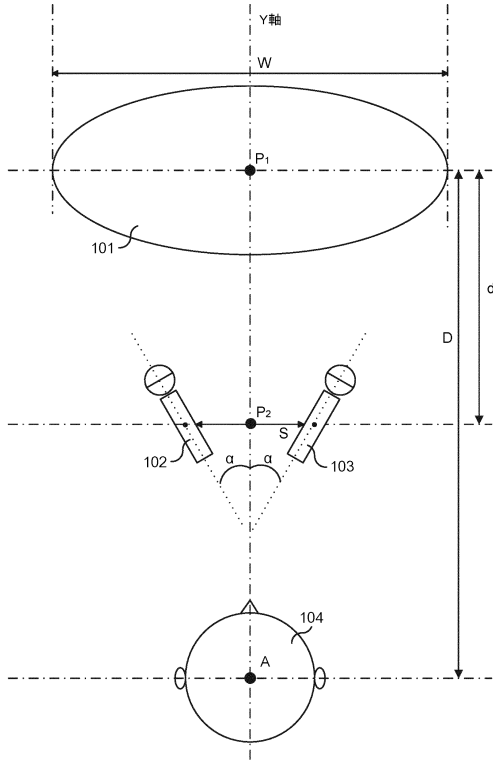
30

40

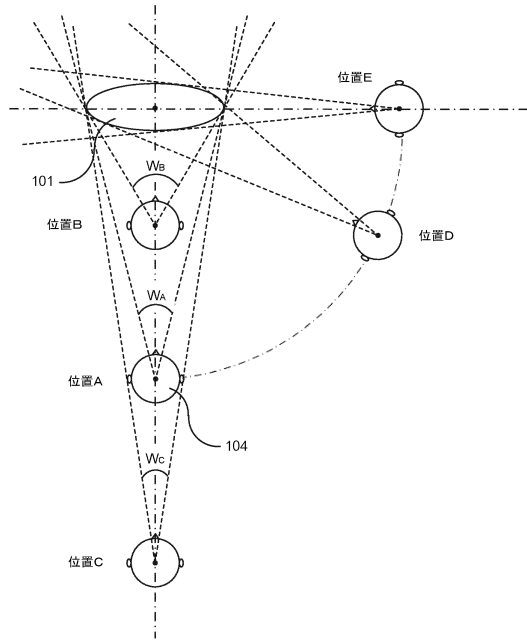
50

【図面】

【図 1】



【図 2】



10

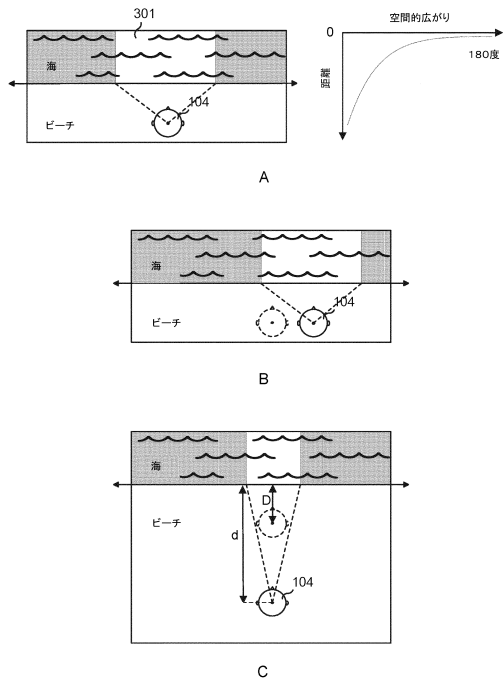
20

30

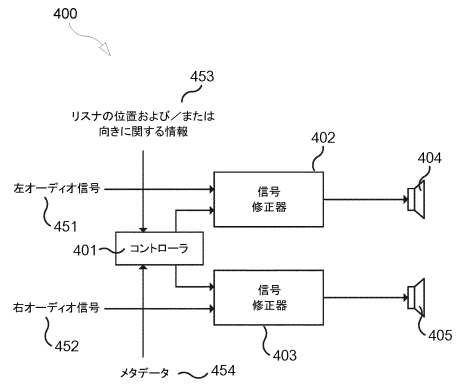
40

50

【図3A - C】



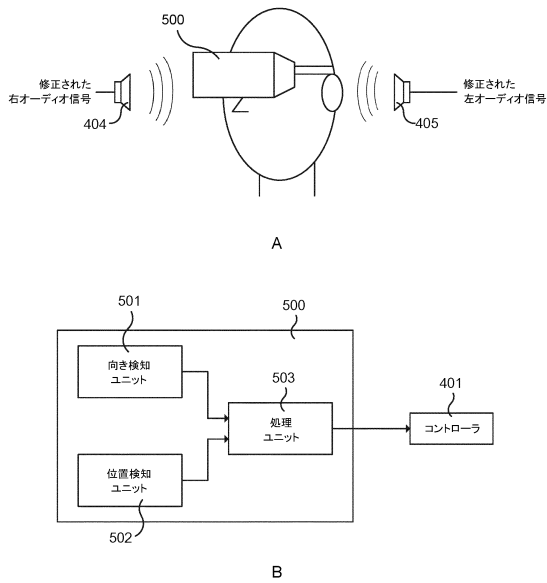
【図4】



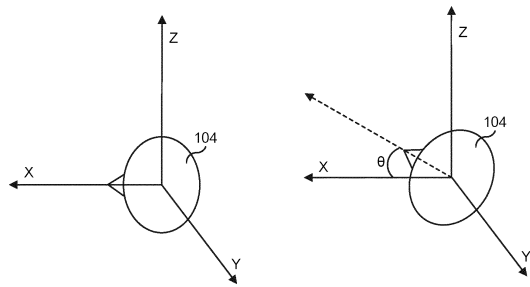
10

20

【図5A - B】



【図6A】

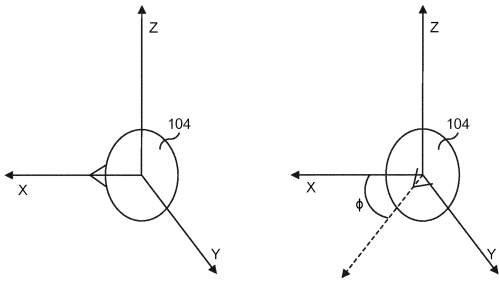


30

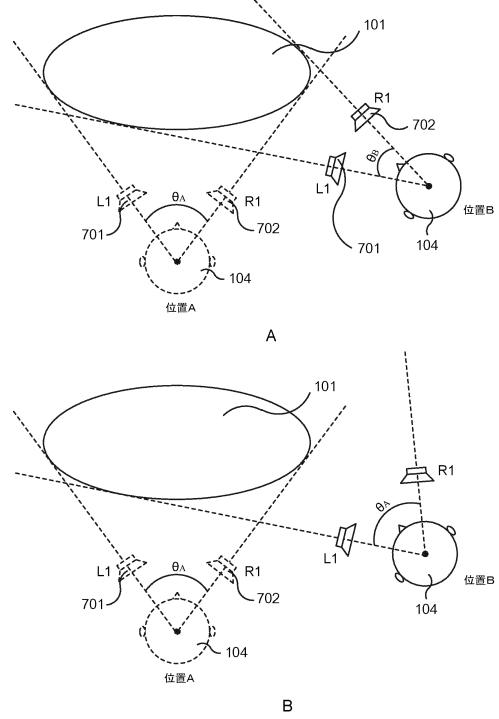
40

50

【 図 6 B 】



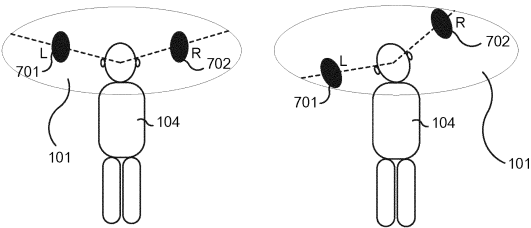
【 図 7 A - B 】



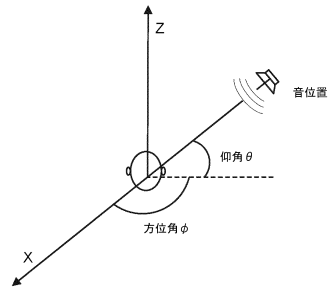
10

20

【 図 8 】



【 図 9 】

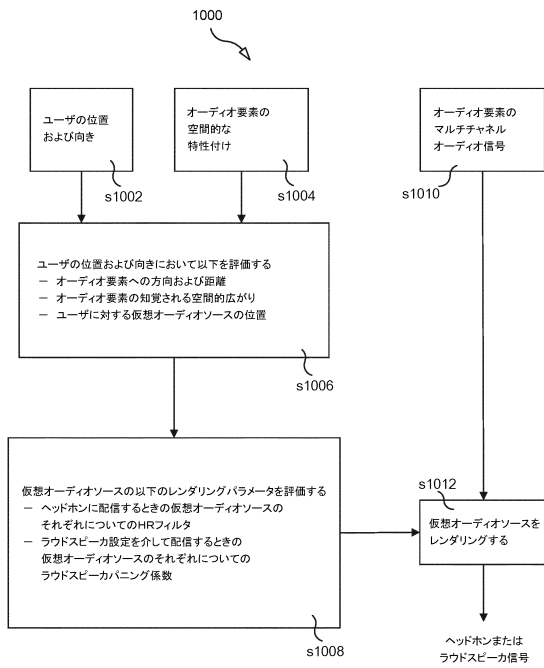


30

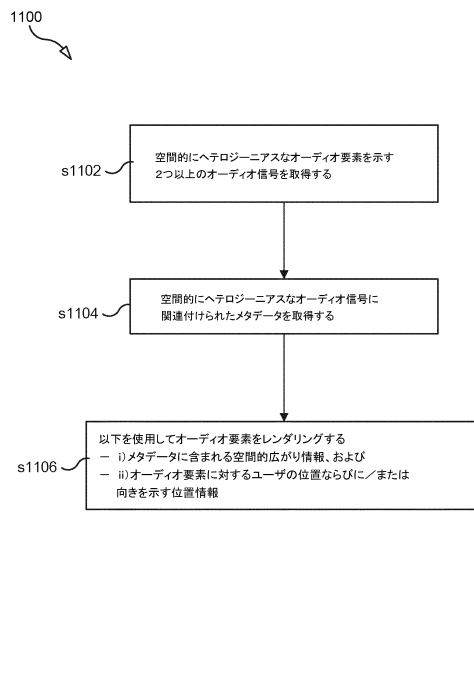
40

50

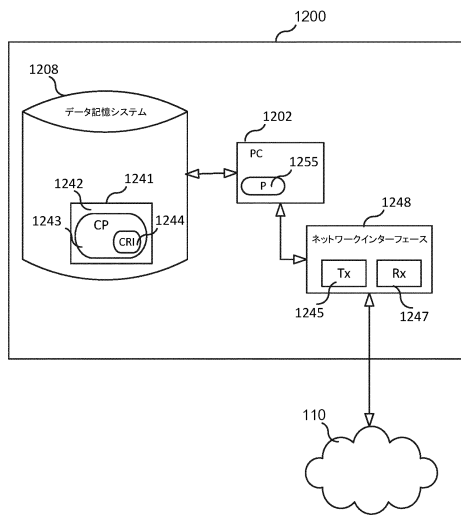
【図 10】



【図 11】



【図 12】



フロントページの続き

- (72)発明者 ファルク, トミ
スウェーデン国 エスエー - 1 6 3 4 7 スパンガ, ガングブリンケン 1 8
- (72)発明者 カールソン, エルレンドゥル
スウェーデン国 エスエー - 7 5 6 5 2 ウプサラ, ハイムダルズヴェーゲン 4 ベー
- (72)発明者 チャン, メンチウ
スウェーデン国 エスエー - 1 1 3 5 6 ストックホルム, ビルガー ヤールスガータン 1 0 9 アー
- (72)発明者 ヤンソン トフゴード, トマス
スウェーデン国 エスエー - 7 5 7 5 6 ウプサラ, スクッグスヴェーゲン 3 7
- (72)発明者 デ ブルーイン, ウェルネル
スウェーデン国 エスエー - 1 1 3 6 6 ストックホルム, ヘルシングガータン 6 1
- 審査官 佐久 聖子
- (56)参考文献 国際公開第 2 0 1 8 / 1 9 7 7 4 8 (W O , A 1)
米国特許出願公開第 2 0 1 8 / 0 1 0 9 9 0 1 (U S , A 1)
米国特許出願公開第 2 0 1 8 / 0 0 6 8 6 6 4 (U S , A 1)
国際公開第 2 0 1 8 / 1 5 0 7 7 4 (W O , A 1)
国際公開第 2 0 1 7 / 1 1 0 8 8 2 (W O , A 1)
- (58)調査した分野 (Int.Cl., D B 名)
H 0 4 S 1 / 0 0 - 7 / 0 0