

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7581359号
(P7581359)

(45)発行日 令和6年11月12日(2024.11.12)

(24)登録日 令和6年11月1日(2024.11.1)

(51)国際特許分類	F I	
H 0 4 N 23/60 (2023.01)	H 0 4 N 23/60	5 0 0
H 0 4 N 23/69 (2023.01)	H 0 4 N 23/69	
H 0 4 N 23/695 (2023.01)	H 0 4 N 23/695	
H 0 4 N 5/265(2006.01)	H 0 4 N 5/265	
H 0 4 N 5/272(2006.01)	H 0 4 N 5/272	
請求項の数 15 (全23頁) 最終頁に続く		

(21)出願番号	特願2022-548695(P2022-548695)	(73)特許権者	515076873
(86)(22)出願日	令和3年2月2日(2021.2.2)		ノキア テクノロジーズ オサケユイチア
(65)公表番号	特表2023-513318(P2023-513318 A)		フィンランド国, 0 2 6 1 0 エスプー, カラカーリ 7
(43)公表日	令和5年3月30日(2023.3.30)	(74)代理人	100094569
(86)国際出願番号	PCT/EP2021/052382		弁理士 田中 伸一郎
(87)国際公開番号	WO2021/160465	(74)代理人	100103610
(87)国際公開日	令和3年8月19日(2021.8.19)		弁理士 吉 田 和彦
審査請求日	令和4年8月10日(2022.8.10)	(74)代理人	100109070
(31)優先権主張番号	20157385.4		弁理士 須田 洋之
(32)優先日	令和2年2月14日(2020.2.14)	(74)代理人	100067013
(33)優先権主張国・地域又は機関	欧州特許庁(EP)		弁理士 大塚 文昭
		(74)代理人	100086771
			弁理士 西島 孝喜
		(74)代理人	
最終頁に続く			

(54)【発明の名称】 マルチメディアコンテンツを処理するための装置、コンピュータプログラム及び方法

(57)【特許請求の範囲】

【請求項 1】

装置であって、
少なくとも1つのプロセッサと、
コンピュータプログラムコードを含む少なくとも1つの非一時的メモリと、
前記少なくとも1つの非一時的メモリおよび前記コンピュータプログラムコードは、前記少なくとも1つのプロセッサと共に、前記装置に、
音源が第1のしきい値を超えるエネルギーまたはレベルを有することを判定することと、

ユーザ選択された音源である目標音源と前記音源との間に、第2のしきい値未満のエネルギーまたはレベルについての相関があることを判定することと、

前記音源のエネルギーまたはレベルが第1のしきい値を超えていること、および前記目標音源と前記音源との相関が第2のしきい値未満であることに応じて、前記音源のオーディオフォーカス減衰を識別することと、

オーディオフォーカス減衰を受ける前記音源と、対応する視覚的物体との間の対応を決定することと、

オーディオフォーカス減衰を受ける前記音源に対応する前記視覚的物体を少なくとも部分的に除外および/または修正するようにイメージの取込みを修正することと
を実行させるように構成された、装置。

【請求項 2】

前記少なくとも1つの非一時的メモリおよび前記コンピュータプログラムコードは、前記少なくとも1つのプロセッサと共に、前記装置に、

オーディオフォーカス減衰を受ける前記音源の位置を決定することと、

オーディオフォーカス減衰を受ける前記音源の前記位置と前記視覚的物体との関連付けに少なくとも部分的に基づいて、前記視覚的物体を決定することと、

決定した視覚的物体を除外するように、取り込まれたイメージを修正することとを実行させるように構成された、請求項1に記載の装置。

【請求項3】

前記音源が、複数の場所のうちのいずれか1つに配置される、請求項2に記載の装置。

【請求項4】

前記音源が、空間オーディオを使用して配置される、請求項2に記載の装置。

【請求項5】

前記目標音源がユーザ選択され、前記装置が、前記目標音源に関連付けられる目標視覚的物体の、ビューファインダディスプレイでの選択子を備える、請求項1に記載の装置。

【請求項6】

前記少なくとも1つの非一時的メモリおよび前記コンピュータプログラムコードは、前記少なくとも1つのプロセッサと共に、前記装置に、物体認識を行わせる、請求項2に記載の装置。

【請求項7】

前記少なくとも1つの非一時的メモリおよび前記コンピュータプログラムコードは、前記少なくとも1つのプロセッサと共に、前記装置に、

取り込まれたイメージを再フレーミングして、前記音源に対応する前記視覚的物体を除外することを行わせるように構成された、請求項1に記載の装置。

【請求項8】

前記少なくとも1つの非一時的メモリおよび前記コンピュータプログラムコードは、前記少なくとも1つのプロセッサと共に、前記装置に、

前記イメージの取込みを修正することが、イメージセンサアレイで検出されたイメージに適用されるクロップを変更することを行わせるように構成され、適用されるクロップが、取り込まれる前記イメージを決定する、請求項1に記載の装置。

【請求項9】

前記イメージの取込みを修正することが、前記イメージセンサアレイで検出された前記イメージに適用される前記クロップの位置を変更することを行わせるように構成される、請求項8に記載の装置。

【請求項10】

前記少なくとも1つの非一時的メモリおよび前記コンピュータプログラムコードは、前記少なくとも1つのプロセッサと共に、前記装置に、

前記イメージセンサアレイで検出された前記イメージに適用される前記クロップのサイズを維持することを行わせるように構成される、請求項9に記載の装置。

【請求項11】

前記少なくとも1つの非一時的メモリおよび前記コンピュータプログラムコードは、前記少なくとも1つのプロセッサと共に、前記装置に、

オーディオフォーカス減衰を受ける音源に関連付けられるすべての視覚的物体を除外することを行わせるように構成される、請求項1のいずれかに記載の装置。

【請求項12】

前記少なくとも1つの非一時的メモリおよび前記コンピュータプログラムコードは、前記少なくとも1つのプロセッサと共に、前記装置に、

オーディオフォーカス減衰を受ける音源に関連付けられる、選択された視覚的物体を除外することを行わせるように構成され、前記視覚的物体の前記選択が、

取り込まれたイメージ内にある、オーディオフォーカス減衰を受ける音源に関連付けられる視覚的物体の数を最小限に抑えること、

10

20

30

40

50

取り込まれたイメージから除外される、オーディオフォーカス減衰を受ける音源に関連付けられる視覚的物体の数を最大にすること、

オーディオフォーカス減衰を受け、取り込まれたイメージ内の視覚的物体に関連付けられる音源の全エネルギーを最小限に抑えること、

オーディオフォーカス減衰を受け、取り込まれたイメージから除外される視覚的物体に関連付けられる音源の全エネルギーを最大にすること、

オーディオフォーカス減衰を受ける音源に関連付けられる視覚的物体の物体認識、

オーディオフォーカス減衰を受ける前記音源に関連付けられる視覚的物体の、前記装置に対する近接、または、

オーディオフォーカス減衰を受ける前記音源に関連付けられる前記視覚的物体で占められる、取り込まれたイメージの面積を最小限に抑えること

のうちの1つまたは複数に依存する、請求項1に記載の装置。

【請求項13】

非一時的コンピュータ可読記憶媒体に記録されたコンピュータプログラムであって、

コンピュータで実行されるとき、

音源が第1のしきい値を超えるエネルギーまたはレベルを有すると判定することと、

ユーザ選択された音源である目標音源と前記音源との間に、第2のしきい値未満のエネルギーまたはレベルについての相関があることを判定することと、

前記音源のエネルギーまたはレベルが第1のしきい値を超えていること、および前記目標音源と前記音源との相関が第2のしきい値未満であることに応じて、前記音源のオーディオフォーカス減衰を識別することと、

オーディオフォーカス減衰を受ける前記音源と、対応する視覚的物体との間の対応を決定することと、

オーディオフォーカス減衰を受ける前記音源に対応する前記視覚的物体を少なくとも部分的に除外および/または修正するようにイメージの取込みを修正することと

を実施するように構成された、コンピュータプログラム。

【請求項14】

コンピュータで実行されるとき、

オーディオフォーカス減衰を受ける前記音源の位置を決定することと、

オーディオフォーカス減衰を受ける前記音源の位置と前記視覚的物体との関連付けに少なくとも部分的に基づいて、前記視覚的物体を決定することと、

決定した視覚的物体を除外するように、取り込まれたイメージを修正することと
を実施するように構成された、請求項13に記載のコンピュータプログラム。

【請求項15】

音源が第1のしきい値を超えるエネルギーまたはレベルを有することを判定することと、ユーザ選択された音源である目標音源と前記音源との間に、第2のしきい値未満のエネルギーまたはレベルについての相関があることを判定することと、

前記音源のエネルギーまたはレベルが第1のしきい値を超えていること、および前記目標音源と前記音源との相関が第2のしきい値未満であることに応じて、前記音源のオーディオフォーカス減衰を識別することと、

オーディオフォーカス減衰を受ける前記音源と、対応する視覚的物体との間の対応を決定することと、

オーディオフォーカス減衰を受ける前記音源に対応する前記視覚的物体を少なくとも部分的に除外および/または修正するようにイメージの取込みを修正することと

を含む方法。

【発明の詳細な説明】

【技術分野】

【0001】

本開示の実施形態はマルチメディアコンテンツに関する。詳細には、マルチメディアコンテンツを取り込むこと、および/または処理することに関する。

10

20

30

40

50

【背景技術】

【0002】

マルチメディアコンテンツはしばしば、専門家ではない人々によって場当たりの無計画のシナリオで記録される。たとえば、音声付きのビデオが、ポータブルカメラ、カメラ対応携帯電話、またはスポーツアクティビティカメラ上に容易に記録され得る。

【発明の概要】

【発明が解決しようとする課題】

【0003】

取り込まれ、または記録されるマルチメディアコンテンツの品質を改善することが望ましいはずである。

10

【課題を解決するための手段】

【0004】

必ずしもすべての実施形態ではないが、様々な実施形態によれば、音源のオーディオフィーカス減衰を識別し、オーディオフィーカス減衰を受ける音源と、対応する視覚的物体との間の対応を決定し、オーディオフィーカス減衰を受ける音源に対応する視覚的物体を少なくとも部分的に除外および/または修正するようにイメージの取込みを修正するための手段を備える装置が提供される。

【0005】

必ずしもすべての実施例ではないが、いくつかの実施例では、装置は、オーディオフィーカス減衰を受ける音源の位置を決定し、オーディオフィーカス減衰を受ける音源の位置に関連付けられる視覚的物体を決定し、決定した視覚的物体を除外するように、取り込まれたイメージを修正するための手段を備える。

20

【0006】

必ずしもすべての実施例ではないが、いくつかの実施例では、オーディオフィーカス減衰を受ける音源の位置を決定するための手段は、候補音源がしきい値を超えるエネルギーまたはレベルを有することを判定し、目標音源と候補音源との間の十分な相関の欠如を判定し、候補音源がしきい値を超えるエネルギーまたはレベルを有し、目標音源との間の十分な相関が欠如している場合、候補音源の位置を取得するための手段を備える。

30

【0007】

目標音源は、その代わりにユーザ選択された音源であり得る。

【0008】

ターゲティングは、たとえばユーザ選択を介して行われ得る。必ずしもすべての実施例ではないが、いくつかの実施例では、目標音源がユーザ選択され、装置は、目標音源に関連付けられる視覚的物体の、ビューファインダディスプレイ内のユーザ選択のための手段を備える。

【0009】

必ずしもすべての実施例ではないが、いくつかの実施例では、候補音源が、複数の場所、たとえば別個の場所のうちのいずれか1つに配置され得る。

40

【0010】

必ずしもすべての実施例ではないが、いくつかの実施例では、候補音源が空間オーディオを使用して配置され得る。

【0011】

必ずしもすべての実施例ではないが、いくつかの実施例では、オーディオフィーカス減衰を受ける音源の位置に関連付けられる視覚的物体を決定するための手段は、物体認識のための手段を備える。

【0012】

50

必ずしもすべての実施例ではないが、いくつかの実施例では、イメージの取込みを修正するための手段が、取り込まれたイメージを再フレーミングして、音源に関連付けられる視覚的物体を除外するように構成される。

【0013】

必ずしもすべての実施例ではないが、いくつかの実施例では、イメージの取込みを修正するための手段が、イメージセンサアレイで検出されたイメージに適用されるクロップを変更するように構成され、適用されるクロップが、取り込まれるイメージを決定する。

【0014】

必ずしもすべての実施例ではないが、いくつかの実施例では、イメージの取込みを修正するための手段が、イメージセンサアレイで検出されたイメージに適用されるクロップの位置を変更するように構成される。

10

【0015】

必ずしもすべての実施例ではないが、いくつかの実施例では、イメージの取込みを修正するための手段が、イメージセンサアレイで検出されたイメージに適用されるクロップのサイズを維持するように構成される。

【0016】

必ずしもすべての実施例ではないが、いくつかの実施例では、イメージの取込みを修正するための手段が、オーディオフィーカス減衰を受ける音源に関連付けられるすべての視覚的物体を除外するように構成される。

【0017】

必ずしもすべての実施例ではないが、いくつかの実施例では、イメージの取込みを修正するための手段が、オーディオフィーカス減衰を受ける音源に関連付けられる、選択された視覚的物体を除外するように構成され、視覚的物体の選択が、

20

取り込まれたイメージ内の、オーディオフィーカス減衰を受ける音源に関連付けられる視覚的物体の数を最小限に抑えること、

取り込まれたイメージから除外される、オーディオフィーカス減衰を受ける音源に関連付けられる視覚的物体の数を最大にすること、

オーディオフィーカス減衰を受け、取り込まれたイメージ内の視覚的物体に関連付けられる音源の全エネルギーを最小限に抑えること、

オーディオフィーカス減衰を受け、取り込まれたイメージから除外される視覚的物体に関連付けられる音源の全エネルギーを最大にすること、

30

オーディオフィーカス減衰を受ける音源に関連付けられる視覚的物体の物体認識、

オーディオフィーカス減衰を受ける音源に関連付けられる視覚的物体の、装置に対する近接、および

オーディオフィーカス減衰を受ける音源に関連付けられる視覚的物体によって占められる、取り込まれたイメージの面積を最小限に抑えること

のうちの1つまたは複数に依存する。

【0018】

必ずしもすべての実施形態ではないが、様々な実施形態によれば、コンピュータ上で実行されるとき、

40

音源のオーディオフィーカス減衰を識別することと、

オーディオフィーカス減衰を受ける音源と、対応する視覚的物体との間の対応を決定することと、

オーディオフィーカス減衰を受ける音源に対応する視覚的物体を少なくとも部分的に除外および/または修正するようにイメージの取込みを修正することと

を実施するコンピュータプログラムが提供される。

【0019】

必ずしもすべての実施形態ではないが、様々な実施形態によれば、

音源のオーディオフィーカス減衰を識別することと、

オーディオフィーカス減衰を受ける音源と、対応する視覚的物体との間の対応を決定す

50

ることと、

オーディオフォーカス減衰を受ける音源に対応する視覚的物体を少なくとも部分的に除外および/または修正するようにイメージの取込みを修正することを含む方法が提供される。

【0020】

必ずしもすべての実施形態ではないが、様々な実施形態によれば、添付の特許請求の範囲で特許請求される実施例が提供される。

【0021】

定義

音源のオーディオフォーカス減衰は、その音源に対する相対的減衰の適用である。音源のレベルが低下する。

10

【0022】

音源のオーディオフォーカス減衰は、その音源に対する相対的利得の適用である。音源のレベルが増加する。

【0023】

複数の音源のうちの1つ(または複数)のオーディオフォーカス減衰は、必然的に複数の音源のうちの他の音源のオーディオフォーカス利得となる。

【0024】

複数の音源のうちの1つ(または複数)のオーディオフォーカス利得は、必然的に複数の音源のうちの他の音源のオーディオフォーカス減衰となる。

20

【0025】

次に、添付の図面を参照しながら、いくつかの例示的实施形態が説明される。

【図面の簡単な説明】

【0026】

【図1A】本明細書で説明される主題の例示的实施形態を示す図である。

【図1B】本明細書で説明される主題の例示的实施形態を示す図である。

【図2A】本明細書で説明される主題の別の例示的实施形態を示す図である。

【図2B】本明細書で説明される主題の別の例示的实施形態を示す図である。

【図2C】本明細書で説明される主題の別の例示的实施形態を示す図である。

【図3】本明細書で説明される主題の別の例示的实施形態を示す図である。

30

【図4A】本明細書で説明される主題の例示的实施形態を示す図である。

【図4B】本明細書で説明される主題の例示的实施形態を示す図である。

【図5A】本明細書で説明される主題の例示的实施形態を示す図である。

【図5B】本明細書で説明される主題の例示的实施形態を示す図である。

【図6A】本明細書で説明される主題の例示的实施形態を示す図である。

【図6B】本明細書で説明される主題の例示的实施形態を示す図である。

【図6C】本明細書で説明される主題の例示的实施形態を示す図である。

【発明を実施するための形態】

【0027】

図1Aおよび1Bは、複数の音源12を含む音響空間10の一実施例を示す。相異なる位置にある音源12には、相異なる添字で符号が付けられる。図1Aおよび図1Bは、原点14に対して相異なる位置にある3つの音源12₁、12₂、12₃を示す。

40

【0028】

マイクロフォンなどの音声取込みデバイスが、原点14または原点14の近くに配置され得る。たとえば、音声取込みデバイスは、アレイ、たとえば1次元、2次元、または3次元アレイとして構成され得る。

【0029】

複数の音声取込みデバイスの使用により、ある音源12の、別の音源12に対する選択的相対的減衰が可能となる。

【0030】

50

この実施例では、少なくとも目標音源 12_3 、および1つまたは複数の他の非目標音源 12_1 、 12_2 を含む3つの音源 12 がある。

【0031】

音声取込み装置が、目標音源 12_3 に相対的利得を適用し、1つまたは複数の他の非目標音源 12_1 、 12_2 に相対的減衰を適用して、目標音源 12_3 に対するオーディオフィーカス 16 を生み出すように構成される。1つまたは複数の非目標音源 12_1 、 12_2 のオーディオフィーカス減衰がある。

【0032】

目標音源 12_3 は、原点 14 に対する第1の位置にあり、1つまたは複数の非目標音源 12_1 、 12_2 は、第1の位置とは異なる、原点 14 からの第2の位置にある。すなわち、目標音源 12_3 は、1つまたは複数の非目標音源 12_1 、 12_2 とは空間的に異なる。

10

【0033】

音声取込み装置は、ビームフォーミングを使用して第1の位置から来る音を分離するように構成され得る。ビームフォーミングは、相異なる音声取込みデバイスからのオーディオ信号の重みつき一次結合を使用する。オーディオ信号の重みつき一次結合は、ハードウェアもしくはソフトウェアで、またはハードウェアとソフトウェアの組合せで実施され得る。

【0034】

必ずしもすべての実施例ではないが、いくつかの実施例では、一次結合についての重みづけは適応的である。重みづけは、たとえば、いくつかの実施例では、第1の位置に向かう目標方向に応答を保つための制約下で、干渉する音源（たとえば、非目標音源 12_1 、 12_2 ）のエネルギーを最小限に抑えることによって取得され得る。これは最小分散無ひずみ応答（minimum variance distortionless response）である。

20

【0035】

必ずしもすべての実施例ではないが、いくつかの実施例では、重みづけは周波数に依存する。

【0036】

たとえば、別個のマイクロフォンのアレイからの信号を処理し、伝播する音場の空間フィルタリングを実現するためにビームフォーマが使用され得る。空間フィルタリングは、重複する周波数内容を有するが、相異なる空間位置から生じる信号を分離する。目標位置（方位）からの目標信号は減衰されないが、他の場所（方位）からの信号は減衰される。空間フィルタは、非目標信号の方向について利得のディープヌル（deep null）を生み出す。したがって、ビームフォーマは、目標方向 16 からの信号に対する空間帯域フィルタとして、および非目標方向からの信号に対する帯域消去フィルタとして動作し得る。

30

【0037】

相異なるマイクロフォンからのオーディオ信号の重みつき一次結合は、 $y(t) = w^H x(t)$ として表され得る。到来方向 j を有する複素平面波について、 $y(t) = \exp(j\omega t) \cdot r(\theta, \phi)$ であり、ただしビームフォーマ応答 $r(\theta, \phi) = w^H d(\theta, \phi)$ であり、 $d(\theta, \phi)$ はステアリングベクトルである。

40

【0038】

w^H は、 $y(t)$ が目標音源 12_3 からの著しく高い（たとえば、最大の）寄与（目標方向のオーディオフィーカス 16 ）と、非目標音源 12_1 、 12_2 （非目標方向）からの著しく低い（たとえば、最小の）寄与とを含むように選択される。 w^H は、単一の周波数帯にわたって、または別個の周波数帯にわたって別々に適用され得る。

【0039】

重み w^H は、費用関数を最小限に抑えることによって、たとえば、ある方向に対する目標利得と、別の方向に対する目標減衰（たとえば、力ゼロ利得）などの応答制約を受ける（目標とするものからの）出力分散を最小限に抑えることによって取得され得る。これに

50

より、非目標方向からの（たとえば、非目標音源 $1\ 2\ 1$ 、 $1\ 2\ 2$ からの）非目標信号からの出力に対する寄与を最小限に抑えながら、（たとえば、目標音源 $1\ 2\ 3$ からの）目標信号を保持する最小分散無ひずみ応答（MVD R）ビームフォーマが生み出される。

【0040】

重みづけおよび空間フィルタリングはリアルタイムに変化し得る。たとえば、ステアリングベクトル d （ θ 、 ϕ ）はリアルタイムに変化し得る。たとえば、重み w^H はリアルタイムで適合され得る。

【0041】

図1 Aおよび1 Bはまた、1つまたは複数の視覚的物体 $2\ 2$ を含む視覚空間 $2\ 0$ の一実施例を示す。相異なる位置にある視覚的物体 $2\ 2$ には、相異なる添字で符号が付けられる。図1 Aおよび1 Bは、原点 $1\ 4$ に対して相異なる位置にある3つの視覚的物体 $2\ 2\ 1$ 、 $2\ 2\ 2$ 、 $2\ 2\ 3$ を示す。

10

【0042】

必ずしもすべての実施例ではないが、この実施例では、各音源 $1\ 2\ 1$ 、 $1\ 2\ 2$ 、 $1\ 2\ 3$ が、対応する視覚的物体 $2\ 2\ 1$ 、 $2\ 2\ 2$ 、 $2\ 2\ 3$ を有する。

【0043】

この実施例では、音響空間 $1\ 0$ と視覚空間 $2\ 0$ が対応する。音響空間 $1\ 0$ 内の音源 $1\ 2\ 1$ 、 $1\ 2\ 2$ 、 $1\ 2\ 3$ の位置は、視覚空間 $2\ 0$ 内の視覚的物体 $2\ 2\ 1$ 、 $2\ 2\ 2$ 、 $2\ 2\ 3$ の位置と重複する。

【0044】

カメラなどの1つまたは複数のイメージ取込みデバイスが、原点 $1\ 4$ または原点 $1\ 4$ の近くに配置され得る。イメージ取込みデバイスは、たとえば、視野 $2\ 6\ 1$ によって画定される視覚的シーンを取り込むように構成され得る。イメージ取込みデバイスは、たとえば、視野 $2\ 6\ 1$ によって画定された視覚的シーンのシーケンスを含むビデオを取り込むビデオカメラであり得る。

20

【0045】

この実施例では、視覚的物体 $2\ 2\ 1$ 、 $2\ 2\ 2$ は、減衰された音源 $1\ 2\ 1$ 、 $1\ 2\ 2$ に対応し、視覚的物体 $2\ 2\ 3$ は、減衰されない音源 $1\ 2\ 3$ に対応する。減衰された音源は、たとえば消音された音源であり得る。

【0046】

図1 Bから、減衰された音源 $1\ 2\ 1$ に対応する視覚的物体 $2\ 2\ 1$ が視野 $2\ 6\ 1$ 内にはなく、したがって原点 $1\ 4$ にあるカメラによって取り込まれる視覚的シーン内にはないことを理解することができる。

30

【0047】

図1 Bから、減衰されない音源 $1\ 2\ 3$ に対応する視覚的物体 $2\ 2\ 3$ が視野 $2\ 6\ 1$ 内にあり、したがって原点 $1\ 4$ にあるカメラによって取り込まれる視覚的シーン内にあることを理解することができる。

【0048】

図1 Bから、減衰された音源 $1\ 2\ 2$ に対応する視覚的物体 $2\ 2\ 2$ が視野 $2\ 6\ 1$ 内にあり、したがって原点 $1\ 4$ にあるカメラによって取り込まれる視覚的シーン内にあることを理解することができる。

40

【0049】

この実施例では、視野 $2\ 6\ 1$ 内にある、減衰された音源 $1\ 2\ 2$ に対応する視覚的物体 $2\ 2\ 2$ は、減衰されない音源 $1\ 2\ 3$ に対応する視覚的物体 $2\ 2\ 3$ から注意をそらすものとなり得る。このことは、視野 $2\ 6\ 1$ 内にある、減衰された音源 $1\ 2\ 2$ に対応する視覚的物体 $2\ 2\ 2$ が、視覚的物体 $2\ 2\ 3$ よりも視覚的シーンの広い部分を占めるとき、または視覚的物体 $2\ 2\ 3$ が背景の物体であるときに視覚的物体 $2\ 2\ 2$ が前景の物体であるとき、または視覚的物体 $2\ 2\ 2$ が動的視覚的物体（位置を移動し、かつ/または変化する視覚的物体）であるときに特にそうである。

【0050】

50

図 2 A は視覚的シーンの一実施例を示す。視覚的物体 2 2₂ は、減衰された音源 1 2₂ に対応する。視覚的物体 2 2₂ と、視覚的物体 2 2₂ に対応する減衰された音源 1 2₂ の位置は、視野 2 6₁ 内にある。取り込まれたイメージ内に視覚的物体 2 2₂ を確認することができる。視覚的物体 2 2₃ は、減衰されない音源 1 2₃ に対応する。視覚的物体 2 2₃ と、視覚的物体 2 2₃ に対応する減衰されない音源 1 2₃ は、視野 2 6₁ 内にある。取り込まれたイメージ内で視覚的物体 2 2₃ を確認することができる。減衰されない音源 1 2₃ から取り込まれた音声が聴覚的に最も顕著であるのに対して、より大きいサイズの前景の視覚的物体 2 2₂ が視覚的に最も顕著である。より大きいサイズの前景の視覚的物体 2 2₂ は、減衰されない音源 1 2₃ に対応する視覚的物体 2 2₃ から視覚的に注意をそらす。

【 0 0 5 1 】

以下の方法 1 0 0 (図 3) が、注意をそらすことを低減する。

【 0 0 5 2 】

方法 1 0 0 は、

ブロック 1 0 2 で、音源 1 2₂ のオーディオフィーカス減衰を識別することと、

ブロック 1 0 4 で、オーディオフィーカス減衰を受ける音源 1 2₂ と、対応する視覚的物体 2 2₂ との間の対応を決定することと、

ブロック 1 0 6 で、オーディオフィーカス減衰を受ける音源 1 2₂ に対応する視覚的物体 2 2₂ を少なくとも部分的に除外および / または修正するようにイメージの取込みを修正することと

を含む。

【 0 0 5 3 】

必ずしもすべての実施例ではないが、いくつかの実施例では、ブロック 1 0 2 で、方法は、音源 1 2₂ のオーディオフィーカス減衰を自動的に識別することを含み得る。必ずしもすべての実施例ではないが、いくつかの実施例では、ブロック 1 0 4 で、方法は、オーディオフィーカス減衰を受ける音源 1 2₂ と、対応する視覚的物体 2 2₂ との間の対応を自動的に決定することを含み得る。必ずしもすべての実施例ではないが、いくつかの実施例では、ブロック 1 0 6 では、方法は、オーディオフィーカス減衰を受ける音源 1 2₂ に対応する視覚的物体 2 2₂ を少なくとも部分的に除外および / または修正するようにイメージの取込みを自動的に修正することを含み得る。

【 0 0 5 4 】

方法は、オーディオフィーカス減衰を受ける音源 1 2₂ に対応する視覚的物体 2 2₂ の取り込まれたイメージに対する視覚的影響を低減する。

【 0 0 5 5 】

方法 1 0 0 の結果の一例が図 2 B に示されている。方法は、オーディオフィーカス減衰を受ける音源 1 2₂ に対応する視覚的物体 2 2₂ を少なくとも部分的に修正するようにイメージの取込みを修正する。未修正の視覚的物体 2 2₂ が図 2 A に示されており、修正された視覚的物体 2 2₂' が図 2 B に示されている。この実施例では、(視覚的物体 2 2₃ ではなく) 未修正の視覚的物体 2 2₂ が、たとえば 5 0 ピクセルのぼけカーネル (`blurring kernel`) を受け、修正された視覚的物体 2 2₂' が生成される。実際には、オーディオフィーカス減衰を受ける音源 1 2₂ に対応する視覚的物体 2 2₂ が、たとえばぼけ、着色

【 0 0 5 6 】

方法 1 0 0 の結果の一例が図 2 C に示されている。方法は、視覚的シーンを再フレーミングして、オーディオフィーカス減衰を受ける音源 1 2₂ に対応する視覚的物体 2 2₂ を少なくとも部分的に除外するようにイメージの取込みを修正する。未修正の視覚的物体 2 2₂ が図 2 A に示されており、少なくとも部分的に除外された、修正された視覚的物体 2 2₂' が図 2 C に示されている。この実施例では、(視覚的物体 2 2₃ ではなく) 未修正の視覚的物体 2 2₂ が、取り込まれたイメージからの少なくとも部分的な除外の対象となる。これは、たとえば、パニングもしくはトラッキングおよび / またはズーミングによって取り込まれるイメージを実質的に再フレーミングすることによって達成され得る。これは、視野 2

10

20

30

40

50

6のサイズおよび/または方向を変更することによって達成され得る。

【0057】

図4Aは図1Bに類似している。図4Aは、減衰されない音源12₃に対応する視覚的物体22₃が視野26₁内にあり、したがって原点14にあるカメラによって取り込まれる視覚的シーン内にあること、および減衰された音源12₂に対応する視覚的物体22₂が視野26₁内にあり、したがって原点14にあるカメラによって取り込まれる視覚的シーン内にあることを示す。

【0058】

図4Bは図1Bと同様であるが、視野26₁が視野26₁'という新しい視野に修正されている。これは、取り込まれたイメージを再フレーミングし、減衰された音源12₂に対応する視覚的物体22₂のすべてまたは部分を除外する。

10

【0059】

したがって、図4Bは、減衰されない音源12₃に対応する視覚的物体22₃が新しい視野26₁'内にあり、したがって原点14にあるカメラによって取り込まれる視覚的シーン内にあること、および減衰された音源12₂に対応する視覚的物体22₂が新しい視野26₁'から少なくとも部分的に除外され、したがって原点14にあるカメラによって取り込まれる視覚的シーン内にもはやないことを示す。

【0060】

いくつかの例では、再フレーミングが自動的に行われる。他の例では、再フレーミングをさせるためのユーザ選択可能なオプションがユーザに与えられる。

20

【0061】

この実施例では、再フレーミングは同一の解像度を維持し、実質的にはパンまたはトラックである。他の例では、追加または代替として、再フレーミングは解像度を変更し、実質的にはズームである。

【0062】

図4Aおよび4Bの実施例では、第2のより広い視野26₂が示されている。広い視野26₂は、カメラのイメージセンサによって撮像される視覚的シーンを示す。より狭い視野26₁、26₁'は、イメージセンサによって取り込まれる視覚的シーンを示す。より狭い視野26₁、26₁'は、第2のより広い視野26₂の副部分である。図4Bに示されるフレーミングは、より狭い視野イメージを達成するように広い視野イメージをクロッピングすることによって達成される。クロップは、図4Aと図4Bの間でサイズおよび/または位置を変更し得る。クロップは、図4Aと図4Bの間で同一のサイズを保ち得る。この実施例では、イメージの取込みを修正することは、イメージセンサアレイで検出されたイメージに適用されるクロップを変更するように構成され、適用されるクロップが、取り込まれるイメージを決定する。クロップは、センサアレイのセンスル(sense1)の連続するサブセットを画定する。イメージセンサアレイで検出されたイメージに適用されるクロップの位置が変更される。イメージセンサアレイで検出されたイメージに適用されるクロップのサイズが維持され、または変更され得る。再クロッピングの効果が図5Aおよび5Bに示されている。図5Aは、イメージセンサアレイで検出されたイメージ(視野26₂)と、現在適用されているクロップ(減衰された音源12₂に対応する視覚的物体22₂を含む視野26₁)とを示す。図5Bは、イメージセンサアレイで検出されたイメージ(視野26₂)に新しく適用された新しいクロップによって画定される、取り込まれたイメージ(減衰された音源12₂に対応する視覚的物体22₂を含まない視野26₁')を示す。必ずしもすべての実施例ではないが、この実施例では、取り込まれたイメージは、減衰された音源12₂に対応する、除外される視覚的物体22₂の一部(修正された視覚的物体22₂')を保持する。

30

40

【0063】

方法100(図3)は、ブロック102で、音源12₂のオーディオフォーカス減衰を識別することを含む。このブロックは、たとえば、

オーディオフォーカス方向、たとえばビームフォーミングステアリングベクトルd(

50

、) を取得することと、

音源方向を取得することと、

たとえばオーディオフィーカス方向ではない方向に減衰される 1 つまたは複数の音源を決定することと

を含み得る。

【 0 0 6 4 】

方法 1 0 0 は、ブロック 1 0 4 で、オーディオフィーカス減衰を受ける音源 1 2 と、対応する視覚的物体 2 2 との間の対応を決定することを含む。

【 0 0 6 5 】

このブロック 1 0 4 は、

オーディオフィーカス減衰を受ける音源 1 2 の位置を決定することと、

オーディオフィーカス減衰を受ける音源 1 2 の位置に関連付けられる視覚的物体 2 2 を決定することと、

決定した視覚的物体 2 2 を除外または修正するように、取り込まれたイメージを修正することと

を含み得る。

【 0 0 6 6 】

視覚的物体 2 2 が、オーディオフィーカス減衰を受ける音源 1 2 の位置またはその近くにある場合、たとえば視覚的物体 2 2 (または視覚的物体 2 2 の任意の部分) がその位置のしきい値距離以内にある場合、視覚的物体 2 2 はその位置に関連付けられる。

【 0 0 6 7 】

必ずしもすべての実施例ではないが、いくつかの実施例では、ブロック 1 0 4 の前に、イメージが取り込まれる。次いで、ブロック 1 0 6 で、この取り込まれたイメージが修正される。必ずしもすべての実施例ではないが、いくつかの実施例では、取り込まれたイメージが、対応する視覚的物体 2 2 を識別するために処理される。

【 0 0 6 8 】

位置に関連付けられる視覚的物体 2 2 の周囲を識別し、あるいは場所に関連付けられる視覚的物体 2 2 を認識または特徴付けるためにコンピュータビジョン技法が使用され得る。コンピュータビジョン技法は、たとえば、縁部またはコーナ検出、注目点検出、特徴検出を含み得る。

【 0 0 6 9 】

位置は、たとえば 1 次元方位 (たとえば方位角)、2 次元方位 (たとえば方位角 および極角)、2 次元座標 (たとえばデカルト座標 (x , y) または極座標 (r ,))、または 3 次元座標 (たとえばデカルト座標 (x , y , z) または円柱極座標 (r , , z) または球面極座標 (r , ,)) であり得る。

【 0 0 7 0 】

たとえば、決定した視覚的物体 2 2 を修正するように、取り込まれたイメージを修正することは、たとえば図 2 B を参照しながら説明したように、取り込まれたイメージを処理することを含み得る。

【 0 0 7 1 】

たとえば、決定した視覚的物体 2 2 を修正するように、取り込まれたイメージを修正することは、たとえば図 2 C または図 4 A、4 B、5 A、5 B を参照しながら説明したように、取り込まれたイメージを再フレーミングまたはクロッピングすることを含み得る。

【 0 0 7 2 】

候補音源 (オーディオフィーカス減衰を受ける音源 1 2 であることがあり、またはそうでないことがある) 1 2 がしきい値を超えるエネルギーまたはレベルを有すること、および目標音源 1 2 と候補音源 1 2 との間の十分な相関の欠如 (しきい値未満の相関スコア) の結果として、候補音源がオーディオフィーカス減衰を受ける音源 1 2 として識別され得る。

【 0 0 7 3 】

10

20

30

40

50

オーディオフィォークラス減衰を受ける音源 1 2 の位置を決定することは、候補音源 1 2 がしきい値を超えるエネルギーまたはレベルを有すると判定することと、目標音源 1 2₃と候補音源 1 2 との間の十分な相関の欠如を判定することと、候補音源 1 2 がしきい値を超えるエネルギーまたはレベルを有し、目標音源 1 2₃との間の十分な相関が欠如している場合、候補音源 1 2 の位置を取得することとを含み得る。

【 0 0 7 4 】

しきい値はゼロまたは非ゼロしきい値であり得る。

【 0 0 7 5 】

いくつかの例では、候補音源 1 2 の位置が、たとえばビームフォーミングステアリングベクトル d (,) の量子化された値によって定義される、複数の別個の位置のうちのいずれか 1 つに特定され得る。

【 0 0 7 6 】

候補音源 1 2 の位置が、空間オーディオを使用して、たとえば前述のビームフォーマを使用して、または候補音響物体に関連付けられるメタデータから特定され得る。

【 0 0 7 7 】

必ずしもすべての実施例ではないが、いくつかの実施例では、目標音源 1 2₃がユーザ選択され得る。装置は、たとえば、ビューファインダディスプレイ内のユーザ選択のための手段を備え得る。

【 0 0 7 8 】

図 6 A は装置 5 0 の一実施例を示す。

【 0 0 7 9 】

装置 5 0 は音声取込み装置であり、音声の入力のためのオーディオ入力回路 6 0 を備える。オーディオ入力回路は複数の音声取込みデバイス 6 2 を備え、この実施例では、複数の音声取込みデバイス 6 2 はマイクロフォンである。

【 0 0 8 0 】

マイクロフォンは空間的に分散し、別個のマイクロフォンのアレイを形成する。アレイは、たとえば 1 D、2 D、または 3 D であり得る。

【 0 0 8 1 】

装置 5 0 はイメージ取込み装置であり、イメージ取込みデバイス 7 0 を備える。イメージ取込みデバイス 7 0 はカメラまたはイメージセンサであり得る。

【 0 0 8 2 】

イメージ取込みデバイス 7 0 は、たとえばビデオカメラ 7 2 であり得る。

【 0 0 8 3 】

コントローラ 8 0 は、オーディオ入力回路 6 0 からオーディオデータを受け取り、イメージ取込みデバイス 7 0 から視覚的データを受け取るように構成される。データは raw であり得、何らかの前処理がオーディオ入力回路 6 0 および / またはイメージ取込みデバイス 7 0 で行われ得る。

【 0 0 8 4 】

この実施例では、装置 5 0 はビームフォーマ装置であり得る。装置 5 0 は、オーディオ入力回路 6 0 からのオーディオデータを空間的にフィルタリングして、前述のようにオーディオフィォークラスを実現するように構成される。オーディオフィォークラスは望ましくない方向からの音を減衰する。

【 0 0 8 5 】

コントローラ 8 0 はまた、

音源 1 2 のオーディオフィォークラス減衰を識別し、

オーディオフィォークラス減衰を受ける音源 1 2 と、対応する視覚的物体 2 2 との間の対応を決定し、

オーディオフィォークラス減衰を受ける音源 1 2 に対応する視覚的物体 2 2 を少なくとも部分的に除外および / または修正するようにイメージの取込みを修正する

10

20

30

40

50

ように構成される。

【0086】

装置50はレンダリング装置であり、イメージおよび音声をマルチメディアコンテンツとして同時にレンダリングするためのユーザ出力回路40を備える。ユーザ出力回路40は、この実施例では、より多くのスピーカ、ヘッドフォンのうちの1つ、あるいはスピーカまたはヘッドフォンに対するインターフェースを備え得るオーディオ出力回路42を備える。ユーザ出力回路40は、この実施例では、たとえばより多くのディスプレイ44のうちの1つを備え得るイメージ出力回路を備える。

【0087】

装置50は、オーディオフォーカス減衰を受ける音源に対応する視覚的物体の除外/修正前のイメージをディスプレイ44上に表示するように構成され得る。

10

【0088】

装置50は、オーディオフォーカス減衰を受ける音源に対応する視覚的物体の除外/修正後のイメージをディスプレイ44上に表示するように構成され得る。

【0089】

装置50は、オーディオフォーカス減衰を受ける音源に対応する視覚的物体を少なくとも部分的に除外および/または修正する、取り込まれたイメージをディスプレイ44上に表示するように構成され得る。

【0090】

装置50は、いくつかの例では、ユーザが装置50に制御コマンドを与えることを可能にするように構成されるユーザ入力回路90を備え得る。必ずしもすべての実施例ではないが、いくつかの実施例では、ユーザ入力回路90は、タッチスクリーンディスプレイを作成するようにディスプレイ44と一体化され得る。必ずしもすべての実施例ではないが、いくつかの実施例では、ディスプレイ44またはタッチスクリーンディスプレイは、ビデオカメラ72のためのカメラビューファインダを提供する。ユーザ入力回路90は、ユーザが音源12および/または音源12に関連付けられる視覚的物体22を識別することを可能にし得る。

20

【0091】

装置50はポータブルまたはハンドポータブルデバイスであり得る。装置50は、たとえばカメラであり得、またはカメラ機能を有する携帯電話であり得る。

30

【0092】

図6Bは、コントローラ80の一実施例を示す。コントローラ80の実装はコントローラ回路としてのものであり得る。コントローラ80は、ハードウェアだけで実装され得、ファームウェアだけを含むソフトウェアのいくつかの態様を有し、またはハードウェアとソフトウェア（ファームウェアを含む）の組合せであり得る。

【0093】

図6Bに示されるように、コントローラ80は、たとえば、汎用または専用プロセッサ82によって実行されるべき、コンピュータ可読記憶媒体（ディスク、メモリなど）上に記憶され得る、そのようなプロセッサ82内のコンピュータプログラム86の実行可能命令を使用することにより、ハードウェア機能を可能にする命令を使用して実装され得る。

40

【0094】

プロセッサ82は、メモリ84から読み取り、メモリ84に書き込むように構成される。プロセッサ82はまた、データおよび/またはコマンドがそれを介してプロセッサ82によって出力される出力インターフェースと、データおよび/またはコマンドがそれを介してプロセッサ82に入力される入力インターフェースとを備え得る。

【0095】

メモリ84は、プロセッサ82内にロードされたとき、装置50の動作を制御するコンピュータプログラム命令（コンピュータプログラムコード）を含むコンピュータプログラム86を記憶する。コンピュータプログラム86のコンピュータプログラム命令は、装置が図3に示される方法を実施することを可能にする論理およびルーチンを提供する。メモ

50

リ 8 4 を読み取ることにより、プロセッサ 8 2 は、コンピュータプログラム 8 6 をロードおよび実行することができる。

【 0 0 9 6 】

したがって装置 5 0 は、
少なくとも 1 つのプロセッサ 8 2 と、
コンピュータプログラムコードを含む少なくとも 1 つのメモリ 8 4 と、
少なくとも 1 つのプロセッサ 8 2 と共に、装置 5 0 に、
音源のオーディオフォーカス減衰を識別することと、
オーディオフォーカス減衰を受ける音源と、対応する視覚的物体との間の対応を決定することと、

オーディオフォーカス減衰を受ける音源に対応する視覚的物体を少なくとも部分的に除外および/または修正するようにイメージの取込みを修正することと

を少なくとも実施させるように構成された少なくとも 1 つのメモリ 8 4 およびコンピュータプログラムコードと
を備える。

【 0 0 9 7 】

図 6 B に示されるように、コンピュータプログラム 8 6 は、任意の適切な送達機構 8 8 を介して装置 5 0 に到達し得る。送達機構 8 8 は、たとえば、機械可読媒体、コンピュータ可読媒体、非一時的コンピュータ可読記憶媒体、コンピュータプログラム製品、メモリデバイス、コンパクトディスク読取り専用メモリ (C D - R O M)、デジタルバーサタイルディスク (D V D)、ソリッドステートメモリなどの記録媒体、コンピュータプログラム 8 6 を含む、またはコンピュータプログラム 8 6 を有形に実施する製造品であり得る。送達機構は、コンピュータプログラム 8 6 を確実に転送するように構成された信号であり得る。装置 5 0 は、コンピュータプログラム 8 6 をコンピュータデータ信号として伝播または送信し得る。

【 0 0 9 8 】

装置に少なくとも以下を実施させる、または少なくとも以下を実施するためのコンピュータプログラム命令。

音源のオーディオフォーカス減衰を識別することと、
オーディオフォーカス減衰を受ける音源と、対応する視覚的物体との間の対応を決定することと、

オーディオフォーカス減衰を受ける音源に対応する視覚的物体を少なくとも部分的に除外および/または修正するようにイメージの取込みを修正すること。

【 0 0 9 9 】

コンピュータプログラム命令は、コンピュータプログラム、非一時的コンピュータ可読媒体、コンピュータプログラム製品、機械可読媒体内に含まれ得る。必ずしもすべての実施例ではないが、いくつかの実施例では、コンピュータプログラム命令は、複数のコンピュータプログラムにわたって分散し得る。

【 0 1 0 0 】

メモリ 8 4 が単一の構成要素/回路として示されているが、メモリ 8 4 は、1 つまたは複数の別々の構成要素/回路として実装され得、その一部またはすべてが一体化され/取外し可能であり得、かつ/または永続的/半永続的/動的/キャッシュされたストレージを提供し得る。

【 0 1 0 1 】

プロセッサ 8 2 が単一の構成要素/回路として示されているが、プロセッサ 8 2 は、1 つまたは複数の別々の構成要素/回路として実装され得、その一部またはすべてが一体化され/取外し可能であり得る。プロセッサ 8 2 はシングルコアまたはマルチコアプロセッサであり得る。

【 0 1 0 2 】

前述の実施例が、複数のオーディオフォーカス方向がある状況に拡張され得る。

10

20

30

40

50

【 0 1 0 3 】

前述の実施例が、オーディオフォーカス減衰を受け、候補の取り込まれたイメージ内の視覚的物体に対応する複数の音源がある状況に拡張され得る。

【 0 1 0 4 】

方法 1 0 0 は、そうした音源のそれぞれについて実施される。

【 0 1 0 5 】

方法は、

ブロック 1 0 2 で、複数の音源 1 2 のオーディオフォーカス減衰を識別することと、

ブロック 1 0 4 で、オーディオフォーカス減衰を受ける複数の音源 1 2 と、取り込まれるべきイメージ内に存在する、それぞれの複数の視覚的物体 2 2 との間の対応を決定することと、

ブロック 1 0 6 で、オーディオフォーカス減衰を受ける複数の音源 1 2 に対応するそれぞれの複数の視覚的物体 2 2 のうちの少なくとも一部を少なくとも部分的に除外および/または修正するようにイメージの取込みを修正することとを含み得る。

【 0 1 0 6 】

必ずしもすべての実施例ではないが、いくつかの実施例では、方法は、普通なら取り込まれたイメージ内にあるはずの、オーディオフォーカス減衰を受ける音源 1 2 に関連付けられるすべての視覚的物体 2 2 を全体的または部分的に除外および/または修正する。その結果、減衰された音源 1 2 は見えない。

【 0 1 0 7 】

必ずしもすべての実施例ではないが、いくつかの実施例では、イメージの取込みを修正することが、オーディオフォーカス減衰を受ける音源 1 2 に関連付けられる、選択された視覚的物体 2 2 を除外するように構成され、視覚的物体 2 2 の選択が、

取り込まれたイメージ内の、オーディオフォーカス減衰を受ける音源 1 2 に関連付けられる視覚的物体 2 2 の数を最小限に抑えること、

取り込まれたイメージから除外される、オーディオフォーカス減衰を受ける音源 1 2 に関連付けられる視覚的物体 2 2 の数を最大にすること、

オーディオフォーカス減衰を受け、取り込まれたイメージ内の視覚的物体 2 2 に関連付けられる音源 1 2 の全エネルギーを最小限に抑えること、

オーディオフォーカス減衰を受け、取り込まれたイメージから除外される視覚的物体 2 2 に関連付けられる音源 1 2 の全エネルギーを最大にすること、

オーディオフォーカス減衰を受ける音源 1 2 に関連付けられる視覚的物体 2 2 の物体認識、

オーディオフォーカス減衰を受ける音源 1 2 に関連付けられる視覚的物体 2 2 の、装置 5 0 に対する近接、

オーディオフォーカス減衰を受ける音源に関連付けられる視覚的物体 2 2 の移動または変化、

オーディオフォーカス減衰を受ける音源 1 2 に関連付けられる視覚的物体 2 2 によって占められる、取り込まれたイメージの面積を最小限に抑えること
のうちの 1 つまたは複数に依存する。

【 0 1 0 8 】

「コンピュータ可読記憶媒体」、「コンピュータプログラム製品」、「有形に実施されたコンピュータプログラム」など、または「コントローラ」、「コンピュータ」、「プロセッサ」などへの参照は、単一/マルチプロセッサアーキテクチャや順次(フォンノイマン)/並列アーキテクチャなどの様々なアーキテクチャを有するコンピュータだけではなく、フィールドプログラマブルゲートアレイ(FPGA)、特定用途向け回路(ASIC)、信号処理デバイス、他の処理回路などの専用回路をも包含すると理解されたい。コンピュータプログラム、命令、コードなどへの参照は、たとえば、プロセッサのための命令であっても、固定機能デバイス、ゲートアレイ、またはプログラマブル論理デバイスなど

10

20

30

40

50

のための構成設定であっても、ハードウェアデバイスのプログラム可能な内容などのプログラム可能プロセッサまたはファームウェアのためのソフトウェアを包含すると理解されたい。

【0109】

本願では、「回路」という用語は、以下のうちの1つまたは複数またはすべてを指すことがある。

(a) ハードウェアのみの回路実装(アナログおよび/またはデジタルのみの回路実装など)、および

(b) 以下のような(適用可能なとき)、ハードウェア回路とソフトウェアの組合せ

(i) アナログおよび/またはデジタルハードウェア回路とソフトウェア/ファームウェアの組合せ、および

(ii) 携帯電話やサーバなどの装置に様々な機能を実施させるように共に働く、ソフトウェアを有するハードウェアプロセッサ(デジタル信号プロセッサを含む)、ソフトウェア、およびメモリの任意の部分、および

(c) 動作のためにソフトウェア(たとえばファームウェア)を必要とするが、動作のためにソフトウェアが不要であるときにソフトウェアが存在しないことがある、ハードウェア回路、および/またはマイクロプロセッサやマイクロプロセッサの一部などのプロセッサ。

【0110】

回路のこの定義は、任意の請求項内を含む、本願でのこの用語のすべての使用に当てはまる。別の実施例として、本願では、回路という用語は、単なるハードウェア回路またはプロセッサおよびその付随するソフトウェアおよび/またはファームウェアの一実装をも包含する。回路という用語はまた、たとえば、特定の請求項要素に適用可能な場合、モバイルデバイス用のベースバンド集積回路、あるいはサーバ、セルラネットワークデバイス、または他のコンピューティングもしくはネットワークデバイス内の類似の集積回路をも包含する。

【0111】

図3に示されるブロックは、方法でのステップ、および/またはコンピュータプログラム86内のコードのセクションを表し得る。ブロックに対する特定の順序の例示は、ブロックについての必要とされる、または好ましい順序があることを必ずしも示唆せず、ブロックの順序および構成は変動し得る。さらに、いくつかのブロックを省略することが可能であり得る。

【0112】

構造的特徴を説明したところでは、構造的特徴が、構造的特徴の機能のうちの1つまたは複数を実施するための手段によって、そうした機能が明示的に説明されるか、それとも暗黙的に説明されるかに関わらず置換され得る。

【0113】

必ずしもすべての実施例ではないが、いくつかの実施例では、装置50は、装置50でのメモリ84内へのデータのローカル記憶を用いて、または用いずに、装置50での回路またはプロセッサによるデータのローカル処理を用いて、または用いずに、装置50からデータを通信するように構成される。

【0114】

データは、1つまたは複数のデバイスに処理済みフォーマットまたは未処理フォーマットでリモートに記憶され得る。データはクラウド内に記憶され得る。

【0115】

データは1つまたは複数のデバイスでリモートに処理され得る。データは、部分的にローカルに処理され、1つまたは複数のデバイスで部分的にリモートに処理され得る。

【0116】

データは、たとえばWi-Fi(登録商標)やBluetooth(登録商標)などの近距離無線通信を介して、または遠距離セルラ無線リンクを介してワイヤレスにリモート

10

20

30

40

50

デバイスに通信され得る。装置は、たとえばデータの通信用の無線トランシーバなどの通信インターフェースを備え得る。

【0117】

装置50は、より大規模な分散型ネットワークの部分を形成する、モノのインターネットの部分であり得る。

【0118】

データの処理は、ローカルであっても、リモートであっても、健康監視、データアグリゲーション、患者監視、生命徴候監視、または他の目的のためのものであり得る。

【0119】

データの処理は、ローカルであっても、リモートであっても、人工知能または機械学習アルゴリズムを使用し得る。データは、たとえば、機械学習ネットワークをトレーニングするための学習入力として使用され得、または機械学習ネットワークに対する照会入力として使用され得、機械学習ネットワークは応答を提供する。機械学習ネットワークは、たとえば、線形回帰、ロジスティック回帰、ベクトルサポートマシン、または単一もしくは複数隠れ層ニューラルネットワークなどの非循環機械学習ネットワーク(cyclic machine learning network)を使用し得る。

10

【0120】

データの処理は、ローカルであっても、リモートであっても出力を生成し得る。出力は装置50に通信され得、装置50は、音声出力、視覚的出力、または触覚出力などの対象に感応する出力を生成し得る。

20

【0121】

データの記録は一時的な記録のみを含み得、またはデータの記録は永続的記録を含み得、またはデータの記録は一時的記録と永続的記録の両方を含み得る。一時的記録は一時的なデータの記録を示唆する。これは、たとえば、感知またはイメージ取込み中に行われ、ダイナミックメモリで行われ、環状バッファ、レジスタ、キャッシュなどのバッファで行われ得る。永続的記録は、データが、アドレス指定可能メモリ空間から検索可能であるアドレス指定可能データ構造の形態であり、したがって削除または上書きされるまで記憶および検索され得ることを示唆するが、長期記憶が行われることがあり、または行われなことがある。イメージに関連する「取込み」という用語の使用は、イメージのデータの一時的記録に関係する。イメージに関連する「記憶」という用語の使用は、イメージのデータの永続的記録に関係する。

30

【0122】

上記の実施例では、取り込まれたイメージは記録されたイメージでもあり得る。

【0123】

ここでは、「モジュール」とは、製造業者またはユーザによって追加されることになるいくつかの部品/構成要素を除外するユニットまたは装置を指す。

【0124】

たとえば、装置50は、時間および/または空間で分離され得る、オーディオ入力回路60、イメージ取込みデバイス70、コントローラ80、およびユーザ出力回路40のうちの一つまたは複数に物理的に分離され得る。装置50の機能は、複数の異なるデバイスであるが、機能的に結合されたデバイスを使用して実装され得る。

40

【0125】

前述の実施例は、

自動車システム、遠隔通信システム、消費者電子製品を含む電子システム、分散コンピューティングシステム、音声コンテンツ、視覚的コンテンツ、およびオーディオビジュアルコンテンツ、ならびに複合現実感、媒介現実感、仮想現実感、および/または拡張現実感を含むメディアコンテンツを生成またはレンダリングするためのメディアシステム、パーソナルヘルスシステムまたはパーソナルフィットネスシステムを含むパーソナルシステム、ナビゲーションシステム、ヒューマンマシンインターフェースとも呼ばれるユーザインターフェース、セルラ、非セルラ、および光ネットワークを含むネットワーク、アドホ

50

ックネットワーク、インターネット、モノのインターネット、仮想化ネットワーク、ならびに関連するソフトウェアおよびサービスの使用可能構成要素として応用される。

【0126】

「備える」という用語は、本文書では排他的ではなく包含的な意味で使用される。すなわち、Yを備えるXに対する任意の参照は、Xがただ1つのYを備え得、または複数のYを備え得る。「備える」を排他的な意味で使用することが意図される場合、「ただ1つの...を備える」を参照することによって、または「からなる」を使用することによって、文脈でそれが明らかにされる。

【0127】

この説明では、様々な実施例に対して参照が行われた。実施例に関連する特徴または機能の説明は、そうした特徴または機能がその実施例に存在することを示す。本文での「実施例」または「たとえば」または「することができる」または「し得る」という用語の使用は、明示的に述べられたか否かに関わらず、そのような特徴または機能が、一例として説明されたか否かに関わらず、少なくとも説明された実施例に存在すること、および必ずしもそうではないが、他の実施例のうちいくつか、またはすべての他の実施例に存在し得ることを示す。したがって、「実施例」、「たとえば」、「することができる」、または「し得る」は、実施例のクラス内の特定の例を参照する。例の特性は、その例のみの特性、またはクラスの特性、またはクラス内の例のすべてではないが一部を含むクラスのサブクラスの特性であり得る。したがって、一実施例を参照するが別の実施例を参照せずに説明された特徴が、可能なら、機能する組合せの部分としてその別の実施例で使用され得るが、その別の実施例で必ずしも使用する必要のないことが暗黙的に開示される。

【0128】

先行するパラグラフで、様々な実施例を参照しながら実施形態が説明されたが、特許請求の範囲から逸脱することなく、与えられた実施例に対する修正が行われ得ることを理解されたい。

【0129】

先行する説明で説明された特徴が、上記で明示的に説明した組合せ以外の組合せで使用され得る。

【0130】

いくつかの特徴を参照しながら機能が説明されたが、そうした機能は、説明されたか否かに関わらず、他の特徴によって実施可能であり得る。

【0131】

いくつかの実施形態を参照しながら特徴が説明されたが、そうした特徴は、説明されたか否かに関わらず、他の実施形態にも存在し得る。

【0132】

「a」または「the」という用語は、本文書では排他的な意味ではなく、包含的な意味で使用される。すなわち、Yを備えるX(X comprising a/the Y)に対する任意の参照は、文脈がそうではないことを明確に示すのでない限り、Xがただ1つのYを備え得、または複数のYを備え得ることを示す。「a」または「the」を排他的な意味で使用することが意図される場合、文脈でそれが明らかにされる。ある状況では、「少なくとも1つ」または「1つまたは複数」の使用は、包含的な意味を強調するために使用され得るが、こうした用語の欠如を排他的な意味を推論するように理解するべきではない。

【0133】

請求項での特徴(特徴の組合せ)の存在は、その特徴または(特徴の組合せ)自体に対する参照、さらには実質的に同一の技術的效果(同等の特徴)を達成する特徴に対する参照である。同等の特徴は、たとえば、変形形態であり、ほぼ同様にして実質的に同一の結果を達成する特徴を含む。同等の特徴は、たとえば、ほぼ同一の結果を達成するためにほぼ同一の方式でほぼ同一の機能を実施する特徴を含む。

10

20

30

40

50

【 0 1 3 4 】

この説明では、実施例の特徴を説明するために、形容詞または形容詞句を用いて様々な実施例に対する参照が行われた。実施例に関する、特徴のそのような説明は、特徴がいくつかの例では厳密に説明したように存在し、他の例ではほぼ説明したように存在することを示す。

【 0 1 3 5 】

上記の明細書では重要であると考えられる特徴に注意を引くように努めたが、出願人は、強調が置かれているか否かに関わらず、上記で参照され、かつ/または図面に示される、特許を受けることができる特徴または特徴の組合せに関して請求の範囲を介して保護を求めることがあることを理解されたい。

10

20

30

40

50

【図面】

【図 1 A】

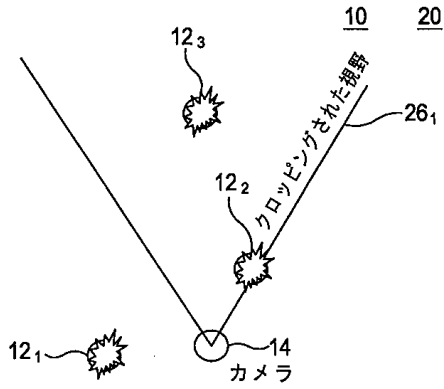


FIG. 1A

【図 1 B】

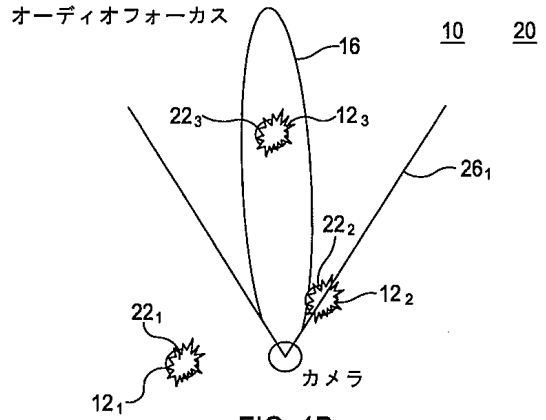


FIG. 1B

【図 2 A】

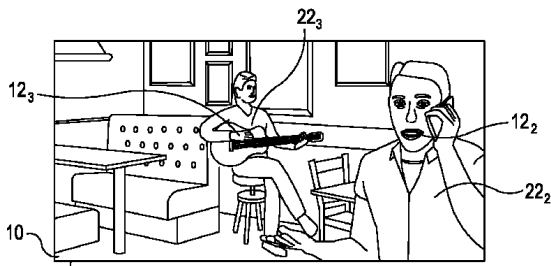


FIG. 2A

【図 2 B】

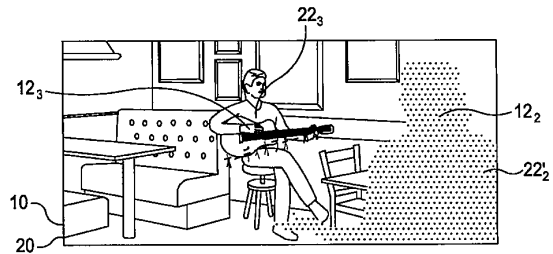


FIG. 2B

10

20

30

40

50

【図 2 C】

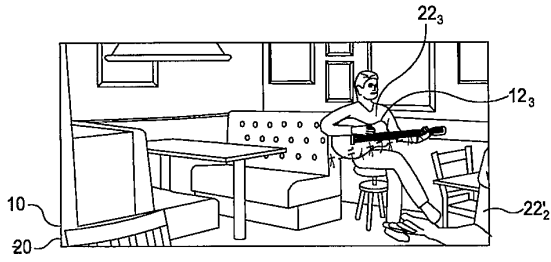


FIG. 2C

【図 3】

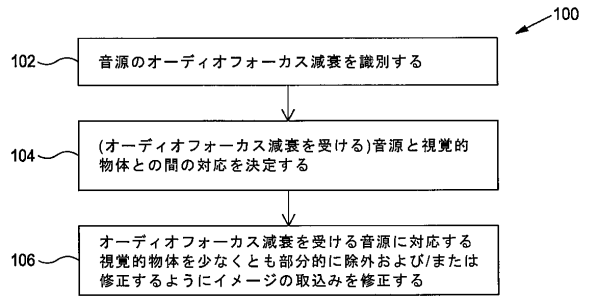


FIG. 3

10

【図 4 A】

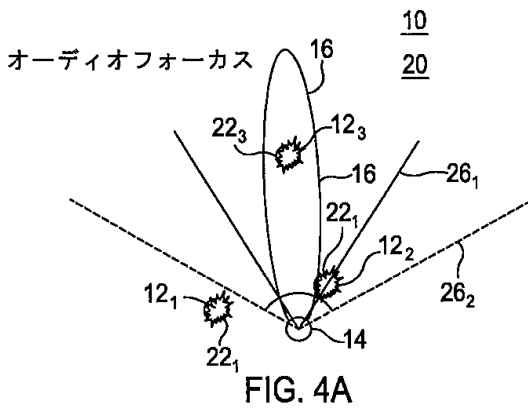


FIG. 4A

【図 4 B】

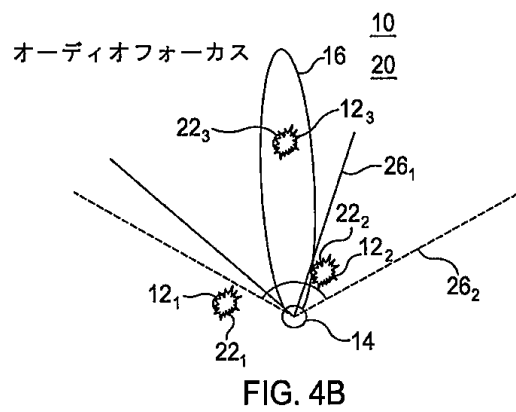


FIG. 4B

20

【図 5 A】

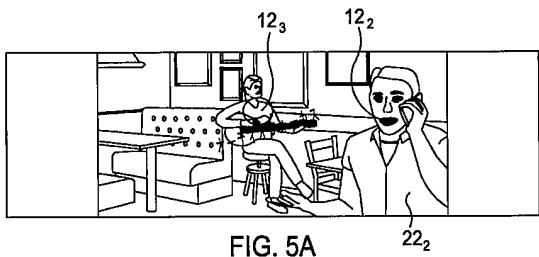


FIG. 5A

【図 5 B】

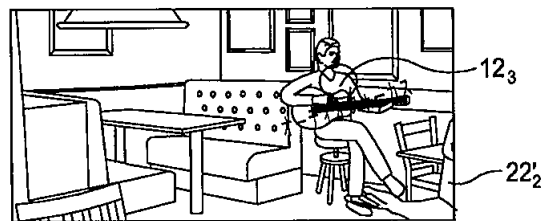


FIG. 5B

30

40

50

【図 6 A】

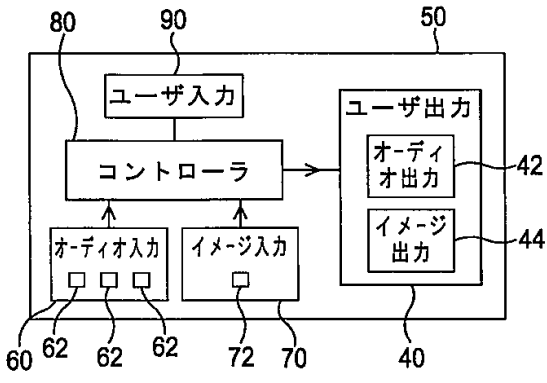


FIG. 6A

【図 6 B】

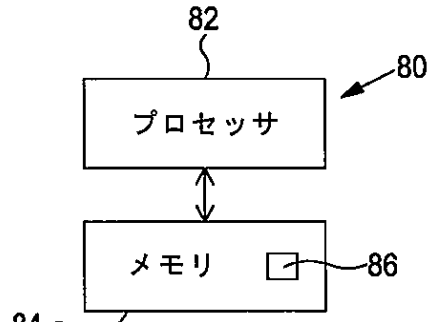


FIG. 6B

【図 6 C】

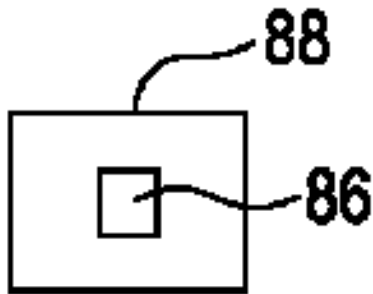


FIG. 6C

10

20

30

40

50

フロントページの続き

(51)国際特許分類

H 0 4 R 3/00 (2006.01)

F I

H 0 4 R 3/00 3 2 0

上杉 浩

(74)代理人 100120525

弁理士 近藤 直樹

(74)代理人 100139712

弁理士 那須 威夫

(74)代理人 100176418

弁理士 工藤 嘉晃

(72)発明者 レッパネン ユッシ アルトゥリ

フィンランド 3 3 5 8 0 タンペレ ミエリキンカトゥ 2 6

(72)発明者 ラークソネン ラッセ ユハニ

フィンランド 3 3 2 1 0 タンペレ ナーシリンナンカトゥ 2 3 ベー 2 8

(72)発明者 ヴィレルモ ミーカ タパニ

フィンランド 3 7 2 0 0 シウロ セルマンクヤ 5

(72)発明者 レーティニエミ アルト ユハニ

フィンランド 3 3 8 8 0 レンパーラ ロマランタティエ 1 3

審査官 藏田 敦之

(56)参考文献 特表 2 0 1 8 - 5 3 6 3 6 5 (J P , A)

特開 2 0 1 7 - 0 2 8 6 3 3 (J P , A)

米国特許出願公開第 2 0 1 5 / 0 1 4 6 0 7 8 (U S , A 1)

(58)調査した分野 (Int.Cl., D B 名)

H 0 4 N 2 3 / 6 0

H 0 4 N 2 3 / 6 9

H 0 4 N 2 3 / 6 9 5

H 0 4 N 5 / 2 6 5

H 0 4 N 5 / 2 7 2

H 0 4 R 1 / 4 0

H 0 4 R 3 / 0 0