

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第7部門第3区分

【発行日】平成28年8月12日(2016.8.12)

【公表番号】特表2016-513410(P2016-513410A)

【公表日】平成28年5月12日(2016.5.12)

【年通号数】公開・登録公報2016-028

【出願番号】特願2015-558105(P2015-558105)

【国際特許分類】

H 04 S	3/00	(2006.01)
G 10 L	19/008	(2013.01)
G 10 L	21/0272	(2013.01)
H 04 R	3/00	(2006.01)
H 04 R	1/40	(2006.01)
H 04 S	5/02	(2006.01)
H 04 N	5/91	(2006.01)

【F I】

H 04 S	3/00	Z
G 10 L	19/008	2 0 0
G 10 L	21/0272	1 0 0 Z
H 04 R	3/00	3 2 0
H 04 R	1/40	3 2 0 A
H 04 S	5/02	D
H 04 N	5/91	C
H 04 N	5/91	Z

【手続補正書】

【提出日】平成28年6月22日(2016.6.22)

【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】

マルチチャネルオーディオデータを生成する方法であって、

デバイスによってキャプチャされたオーディオデータを解析することであって、ここにおいて、前記オーディオデータを解析することは、1つまたは複数のオーディオオブジェクトを識別し、前記1つまたは複数のオーディオオブジェクトを記述するオーディオメタデータを生成するために前記オーディオデータの聴覚的シーン解析を行うことを備える、解析することと、

前記オーディオデータの前記キャプチャと同時に前記デバイスによってキャプチャされたビデオデータを解析することであって、ここにおいて、前記ビデオデータを解析することは、1つまたは複数のビデオオブジェクトを識別し、前記1つまたは複数のビデオオブジェクトを記述するビデオメタデータを生成するために前記ビデオデータの視覚的シーン解析を行うことを備え、前記オーディオメタデータと前記ビデオメタデータとが共通のテキストフォーマットを使用する、解析することと、

前記1つまたは複数のオーディオオブジェクトのうちの少なくとも1つを前記1つまたは複数のビデオオブジェクトのうちの少なくとも1つと関連付けることと、

前記オーディオオブジェクトのうちの前記少なくとも1つをレンダリングするとき、前

記オーディオオブジェクトのうちの前記少なくとも1つを記述する前記オーディオメタデータと前記ビデオオブジェクトのうちの前記関連付けられている1つを記述する前記ビデオメタデータとの間の相関のレベルに基づいて前記オーディオオブジェクトの前記少なくとも1つの拡散性を少なくとも部分的に調整することによって、前記1つまたは複数のビデオオブジェクトのうちの前記少なくとも1つとの前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つの前記関連付けに基づいて、前記オーディオデータから前記マルチチャネルオーディオデータを生成することと
を備える方法。

【請求項2】

前記オーディオメタデータは、対応するオーディオオブジェクトの位置と、形状と、速度と、位置の信頼水準とのうちの1つまたは複数を備え、

前記ビデオメタデータは、前記対応するオーディオオブジェクトの位置と、形状と、速度と、位置の信頼水準とのうちの1つまたは複数を備える、請求項1に記載の方法。

【請求項3】

前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つを前記1つまたは複数のビデオオブジェクトのうちの前記少なくとも1つと関連付けることは、

前記1つまたは複数のオーディオオブジェクトの各々を、あるタイプのオーディオオブジェクトとして分類することと、

前記1つまたは複数のビデオオブジェクトの各々を、あるタイプのビデオオブジェクトとして分類することと、

前記オーディオオブジェクトのうちの前記少なくとも1つの前記タイプが前記ビデオオブジェクトのうちの前記少なくとも1つと同じタイプであることを決定することと、

前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つの前記タイプが前記1つまたは複数のビデオオブジェクトのうちの前記少なくとも1つと同じタイプであることの前記決定に応答して、前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つを前記1つまたは複数のビデオオブジェクトのうちの前記少なくとも1つと関連付けることを備える、請求項1に記載の方法。

【請求項4】

前記マルチチャネルオーディオデータを生成することは、

前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つの前記オーディオメタデータと前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つと関連付けられている前記1つまたは複数のビデオオブジェクトのうちの前記少なくとも1つの前記ビデオメタデータとの間の前記相関のレベルを決定することと、

前記決定された相関のレベルに基づいて、前記1つまたは複数のビデオオブジェクトのうちの前記少なくとも1つが関連付けられている前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つに関する複合メタデータを生成することと、

前記相関のレベルに基づいて、前記オーディオオブジェクトのうちの前記少なくとも1つの前記拡散性を調整することと、

前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つに関して生成された前記複合メタデータに基づいて、前記マルチチャネルオーディオデータの1つまたは複数の前景チャネル内に前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つをレンダリングすることとを備える、請求項1に記載の方法。

【請求項5】

前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つは、前記1つまたは複数のオーディオオブジェクトのうちの第1のものを備え、

前記方法は、さらに、前記1つまたは複数のオーディオオブジェクトのうちの第2のものが前記1つまたは複数のビデオオブジェクトのうちのいずれにも関連付けられていないことを決定することを備え、

前記マルチチャネルオーディオデータを生成することは、前記オーディオオブジェクトのうちの前記第2のものが前記マルチチャネルオーディオデータの1つまたは複数の背景

チャネル内で発生するように、前記マルチチャネルオーディオデータを生成することを備える、請求項 1 に記載の方法。

【請求項 6】

前記マルチチャネルオーディオデータを生成することは、前記オーディオオブジェクトのうちの前記第 2 のものが前記マルチチャネルオーディオデータの前記 1 つまたは複数の背景チャネル内で拡散したオーディオオブジェクトとして発生するように、前記マルチチャネルオーディオデータを生成することを備える、請求項 5 に記載の方法。

【請求項 7】

前記 1 つまたは複数のビデオオブジェクトのうちの前記少なくとも 1 つは、前記 1 つまたは複数のビデオオブジェクトのうちの第 1 のものを備え、

前記方法は、さらに、

前記ビデオオブジェクトのうちの前記 1 つまたは複数のうちの第 2 のものが前記 1 つまたは複数のオーディオオブジェクトのうちのいずれにも関連付けられていないことを決定することと、

前記ビデオオブジェクトのうちの前記 1 つまたは複数のうちの前記第 2 のものが前記 1 つまたは複数のオーディオオブジェクトのうちのいずれにも関連付けられていないことを決定することに応答して、前記 1 つまたは複数のビデオオブジェクトのうちの前記第 2 のものに関連付けられているであろう基準オーディオオブジェクトをオーディオライブラリから取得することと、

前記マルチチャネルオーディオデータの少なくとも一部を生成するために、前記 1 つまたは複数のビデオオブジェクトのうちの前記第 2 のものに基づいて前記基準オーディオオブジェクトをレンダリングすることとを備える、請求項 1 に記載の方法。

【請求項 8】

前記オーディオメタデータは、前記ビデオメタデータを定義するために使用されるテキストフォーマットと共にテキストフォーマットで定義される、請求項 1 に記載の方法。

【請求項 9】

前記マルチチャネルオーディオデータを生成することは、

前記オーディオオブジェクトのうちの前記少なくとも 1 つを記述する前記オーディオメタデータと前記ビデオオブジェクトのうちの前記関連付けられている 1 つを記述する前記ビデオメタデータとの間の百分率差の関数として前記相関のレベルを決定することと、

前記決定された相関のレベルが信頼区間の外側にあるとき、前記マルチチャネルオーディオデータの複数のチャネルにまたがって広がる、拡散するオーディオオブジェクトとして前記オーディオオブジェクトのうちの前記少なくとも 1 つをレンダリングすることとを備える、請求項 1 に記載の方法。

【請求項 10】

携帯電話上で行われる、請求項 1 に記載の方法。

【請求項 11】

実行されたとき、デバイスの 1 つまたは複数のプロセッサに、

前記デバイスによってキャプチャされたオーディオデータを解析することであって、ここにおいて、前記オーディオデータを解析することは、1 つまたは複数のオーディオオブジェクトを識別し、前記 1 つまたは複数のオーディオオブジェクトを記述するオーディオメタデータを生成するために前記オーディオデータの聴覚的シーン解析を行うことを備える、解析することと、

前記オーディオデータの前記キャプチャと同時に前記デバイスによってキャプチャされたビデオデータを解析することであって、前記ビデオデータを解析することは、1 つまたは複数のビデオオブジェクトを識別し、前記 1 つまたは複数のビデオオブジェクトを記述するビデオメタデータを生成するために前記ビデオデータの視覚的シーン解析を行うことを備え、前記オーディオメタデータと前記ビデオメタデータとが共通のテキストフォーマットを使用する、解析することと、

前記 1 つまたは複数のオーディオオブジェクトのうちの少なくとも 1 つを前記 1 つまた

は複数のビデオオブジェクトのうちの少なくとも1つと関連付けることと、

前記オーディオオブジェクトのうちの前記少なくとも1つをレンダリングするとき、前記オーディオオブジェクトのうちの前記少なくとも1つを記述する前記オーディオメタデータと前記ビデオオブジェクトのうちの前記関連付けられている1つを記述する前記ビデオメタデータとの間の相関のレベルに基づいて前記オーディオオブジェクトの前記少なくとも1つの拡散性を少なくとも部分的に調整することによって、前記1つまたは複数のビデオオブジェクトのうちの前記少なくとも1つとの前記1つまたは複数のオーディオオブジェクトのうちの前記少なくとも1つの前記関連付けに基づいて、前記オーディオデータからマルチチャネルオーディオデータを生成することとをさせる命令が記憶された非一時的コンピュータ可読記憶媒体。

【請求項12】

マルチチャネルオーディオデータを生成するように構成されるデバイスであって、
オーディオデータを記憶するように構成されるメモリと、

前記オーディオデータを、オーディオオブジェクトを識別し、前記オーディオオブジェクトを記述するオーディオメタデータを生成するために前記オーディオデータの聴覚的シーン解析を少なくとも部分的に行うことにより、解析し、ビデオデータを、ビデオオブジェクトを識別し、前記ビデオオブジェクトを記述するビデオメタデータを生成するために前記ビデオデータの視覚的シーン解析を少なくとも部分的に行うことにより解析し、前記オーディオオブジェクトと前記ビデオオブジェクトとを関連付け、前記オーディオオブジェクトのうちの前記少なくとも1つをレンダリングするとき、前記オーディオオブジェクトを記述する前記オーディオメタデータと前記ビデオオブジェクトを記述する前記ビデオメタデータとの間の相関のレベルに基づいて前記オーディオオブジェクトの前記少なくとも1つの拡散性を少なくとも部分的に調整することによって、前記ビデオオブジェクトへの前記オーディオオブジェクトの前記関連付けに基づいて、前記オーディオデータから前記マルチチャネルオーディオデータを生成するように構成される1つまたは複数のプロセッサと

を備えるデバイス。

【請求項13】

前記オーディオメタデータは、前記オーディオオブジェクトのサイズと前記オーディオオブジェクトの位置とを備える、請求項12に記載のデバイス。

【請求項14】

前記ビデオメタデータは、前記ビデオオブジェクトのサイズと前記ビデオオブジェクトの位置とを備える、請求項12に記載のデバイス。

【請求項15】

前記1つまたは複数のプロセッサは、サイズと位置とのうちの1つまたは複数を備える複合メタデータを少なくとも部分的に生成するように構成される、請求項12に記載のデバイス。

【請求項16】

前記オーディオメタデータは位置メタデータを含み、

前記ビデオメタデータは位置メタデータを含み、

前記複合メタデータを生成することは、

前記相關のレベルを決定するために、前記オーディオメタデータの前記位置メタデータを前記ビデオメタデータの前記位置メタデータと比較することと、

前記相關のレベルが信頼しきい値を超えているかどうかの決定に基づいて、前記複合メタデータの位置メタデータを生成することとを備える、請求項15に記載のデバイス。

【請求項17】

前記オーディオオブジェクトは、第1のオーディオオブジェクトを備え、ここにおいて、前記ビデオオブジェクトは、第1のビデオオブジェクトを備え、ここにおいて、前記1つまたは複数のプロセッサは、

前記第1のオーディオオブジェクトのデータ構成要素と前記第1のビデオオブジェクト

のデータ構成要素との第1の比較に基づいて、前記第1のビデオオブジェクトの相対物と関連付けられている前記第1のオーディオオブジェクトを識別することと、

第2のオーディオオブジェクトのデータ構成要素と第2のビデオオブジェクトのデータ構成要素との第2の比較に基づいて、前記第2のビデオオブジェクトの相対物と関連付けられていない前記第2のオーディオオブジェクトを識別することとを行いうように構成され、

ここにおいて、前記プロセッサは、さらに

第1のゾーン内に前記第1のオーディオオブジェクトをレンダリングすることと、

第2のゾーン内に前記第2のオーディオオブジェクトをレンダリングすることと、

前記第1のゾーン内の前記レンダリングされた第1のオーディオオブジェクトと、前記第2のゾーン内の前記レンダリングされた第2のオーディオオブジェクトとを組み合わせることに基づいて前記オーディオ出力信号を生成することとを行いうように構成される、

請求項1_2に記載のデバイス。

【請求項1_8】

前記第1のオーディオオブジェクトの前記データ構成要素は位置とサイズとのうちの1つを備える、請求項1_7に記載のデバイス。

【請求項1_9】

前記第1のビデオオブジェクトの前記データ構成要素は位置とサイズとのうちの1つを備える、請求項1_7に記載のデバイス。

【請求項2_0】

前記第2のオーディオオブジェクトの前記データ構成要素は位置とサイズとのうちの1つを備える、請求項1_7に記載のデバイス。

【請求項2_1】

前記第2のビデオオブジェクトの前記データ構成要素は位置とサイズとのうちの1つを備える、請求項1_7に記載のデバイス。

【請求項2_2】

前記第1のゾーンおよび第2のゾーンは、オーディオ前景内の異なるゾーン、またはオーディオ背景内の異なるゾーンである、請求項1_7に記載のデバイス。

【請求項2_3】

前記第1のゾーンおよび第2のゾーンは、オーディオ前景内の同じゾーン、またはオーディオ背景内の同じゾーンである、請求項1_7に記載のデバイス。

【請求項2_4】

前記第1のゾーンはオーディオ前景内にあり、前記第2のゾーンはオーディオ背景内にある、請求項1_7に記載のデバイス。

【請求項2_5】

前記第1のゾーンはオーディオ背景内にあり、前記第2のゾーンはオーディオ前景内にある、請求項1_7に記載のデバイス。

【請求項2_6】

前記1つまたは複数のプロセッサは、

前記第1の比較が信頼区間の外部にあるかどうかを決定することと、

前記第1の比較が前記信頼区間の外部にあるかどうかの前記決定に基づいて、前記第1のオーディオオブジェクトの前記データ構成要素と前記第1のビデオオブジェクトの前記データ構成要素とを重み付けすることとを行いうようにさらに構成される、請求項1_7に記載のデバイス。

【請求項2_7】

前記1つまたは複数のプロセッサは、前記第1のオーディオオブジェクトデータの前記データ構成要素と前記第1のビデオオブジェクトの前記データ構成要素とを平均化するよう構成される、請求項2_6に記載のデバイス。

【請求項2_8】

前記1つまたは複数のプロセッサは、前記第1の比較と前記第2の比較とのうちの1つ

または複数に基づいて異なるビット数を割り当てるようにさらに構成される、請求項 1 7 に記載のデバイス。