

【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 6 部門第 2 区分

【発行日】平成29年2月9日 (2017.2.9)

【公表番号】特表2015-514239(P2015-514239A)

【公表日】平成27年5月18日 (2015.5.18)

【年通号数】公開・登録公報2015-033

【出願番号】特願2015-505720(P2015-505720)

【国際特許分類】

G 1 0 L 15/10 (2006.01)

G 0 6 T 7/00 (2017.01)

G 0 6 T 7/20 (2017.01)

G 1 0 L 25/51 (2013.01)

G 1 0 L 15/00 (2013.01)

G 1 0 L 15/28 (2013.01)

G 1 0 L 15/24 (2013.01)

【F I】

G 1 0 L 15/10 5 0 0 Z

G 0 6 T 7/00 P

G 0 6 T 7/20 B

G 1 0 L 25/51 4 0 0

G 1 0 L 15/00 2 0 0 G

G 1 0 L 15/28 4 0 0

G 1 0 L 15/24 Z

【手続補正書】

【提出日】平成29年1月6日 (2017.1.6)

【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】

デバイスによって実行される方法であって、前記方法は、  
シーンにおいて記録されたオーディオから複数の音響認識特徴を計算することと、  
オブジェクトの音源のタイプを判断するために前記音響認識特徴を 1 つまたは複数のオブジェクトに対応する所定の音響認識特徴と比較することと、  
前記音源のタイプに基づいて前記オブジェクトに対応するキーポイントを選択すること  
 と、

前記選択されたキーポイントおよび前記音源のタイプに基づいて前記オブジェクトを識別することとを備える、方法。

【請求項 2】

前記シーンにおいて記録されたオーディオに基づいて、1 つまたは複数のオブジェクトに対応する 1 つまたは複数のキーポイントシグネチャを選択することと、

前記シーンの画像中の複数のキーポイントを識別することと、

前記オブジェクトを識別するために前記キーポイントを前記キーポイントシグネチャと比較することとをさらに備える、請求項 1 に記載の方法。

【請求項 3】

前記シーンにおいて記録された前記オーディオに基づいてシーン画像の一部を選択す

ることと、

前記画像の前記一部分内からのみ前記キーポイントを選択することとをさらに備える、請求項 1 に記載の方法。

【請求項 4】

前記シーンにおいて記録された前記オーディオに基づいて前記画像の一部分を選択することが、

前記オーディオからオーディオ到来方向 (DOA) を判断することと、

前記オーディオ DOA に基づいて前記画像の前記一部分を選択することを含む、請求項 3 に記載の方法。

【請求項 5】

前記オーディオ DOA を判断することが、

前記シーンに位置する複数のマイクロフォンにおいて前記オーディオを受信し、それによって複数のマイクロフォン信号を生成することと、

前記マイクロフォン信号に基づいて前記オーディオ DOA を判断することを含む、請求項 4 に記載の方法。

【請求項 6】

前記シーンのビデオ記録から複数の局所動きベクトルを計算することと、

前記局所動きベクトルを 1 つまたは複数のオブジェクトに対応する所定の局所動きベクトルのデータベースと比較することによって、および前記キーポイントを 1 つまたは複数のキーポイントシグネチャと比較することによって前記オブジェクトを識別することとをさらに備える、請求項 1 に記載の方法。

【請求項 7】

前記オブジェクトを識別することは、前記キーポイントを 1 つまたは複数のキーポイントシグネチャと比較することに基づく、請求項 1 に記載の方法。

【請求項 8】

前記音響認識特徴がメル周波数ケプストラム係数を含む、請求項 7 に記載の方法。

【請求項 9】

画像中に現れる 1 つまたは複数のオブジェクトについての範囲情報を判断することと、

前記範囲情報に基づいて前記キーポイントを分析することとをさらに備える、請求項 1 に記載の方法。

【請求項 10】

範囲情報を判断することが、オートフォーカスカメラを使用して範囲情報を判断することと、マルチカメラ画像視差推定を使用して範囲情報を判断することと、上記の任意の組合せとからなるグループから選択される、請求項 9 に記載の方法。

【請求項 11】

シーンにおいて記録されたオーディオから複数の音響認識特徴を計算するように構成されたオーディオプロセッサと、

音源のタイプに基づいてオブジェクトに対応するキーポイントを選択するように構成されたキーポイントセクタと、

前記選択されたキーポイントに基づいて前記オブジェクトを識別し、前記オブジェクトの前記音源のタイプを判断するために前記音響認識特徴を 1 つまたは複数のオブジェクトに対応する所定の音響認識特徴と比較するように構成された整合デバイスとを備える、装置。

【請求項 12】

シーンの画像中の複数のキーポイントを識別するように構成されたキーポイント検出器をさらに備え、

前記キーポイントセクタが、前記シーンにおいて記録されたオーディオに基づいて、1 つまたは複数のオブジェクトに対応する 1 つまたは複数のキーポイントシグネチャを選択するように構成され、

前記整合デバイスが、前記シーン中のオブジェクトを識別するために前記キーポイント

を前記キーポイントシグネチャと比較するように構成された、請求項 11 に記載の装置。

【請求項 13】

前記シーンにおいて記録された前記オーディオに基づいて前記シーンの画像の一部分を選択するように構成された第 1 のセクタと、

前記画像の前記一部分内からのみ前記キーポイントを選択するように構成された第 2 のセクタとをさらに備える、請求項 11 に記載の装置。

【請求項 14】

前記第 1 のセクタが、

前記オーディオからオーディオ到来方向 (DOA) を判断するように構成された検出器と、

前記オーディオ DOA に基づいて前記画像の前記一部分を選択するように構成された第 3 のセクタとを含む、請求項 13 に記載の装置。

【請求項 15】

前記検出器が、

前記オーディオを受信して、複数のマイクロフォン信号を生成するための、前記シーンに位置する複数のマイクロフォンと、

前記マイクロフォン信号に基づいて前記オーディオ DOA を判断するように構成されたオーディオプロセッサとを含む、請求項 14 に記載の装置。

【請求項 16】

前記シーンのビデオ記録から複数の局所動きベクトルを計算するように構成されたビデオプロセッサをさらに備え、

前記整合デバイスが、前記局所動きベクトルを 1 つまたは複数のオブジェクトに対応する所定の局所動きベクトルのデータベースと比較することによって、および前記キーポイントを 1 つまたは複数のキーポイントシグネチャと比較することによって前記オブジェクトを識別するように構成された、請求項 11 に記載の装置。

【請求項 17】

前記整合デバイスは、前記キーポイントを 1 つまたは複数のキーポイントシグネチャと比較することによって前記オブジェクトを識別するように構成される、請求項 11 に記載の装置。

【請求項 18】

前記音響認識特徴がメル周波数ケプストラム係数を含む、請求項 17 に記載の装置。

【請求項 19】

画像中に現れる 1 つまたは複数のオブジェクトについての範囲情報を判断するように構成された範囲検出器と、

前記範囲情報に基づいて前記キーポイントを分析するように構成されたキーポイント検出器とをさらに備える、請求項 11 に記載の装置。

【請求項 20】

前記範囲検出器が、オートフォーカスカメラと、マルチカメラアレイと、上記の任意の組合せとからなるグループから選択される検出器を含む、請求項 19 に記載の装置。

【請求項 21】

シーンにおいて記録されたオーディオから複数の音響認識特徴を計算するための手段と

、

オブジェクトの音源のタイプを判断するために前記音響認識特徴を 1 つまたは複数のオブジェクトに対応する所定の音響認識特徴と比較するための手段と、

前記音源のタイプに基づいて前記オブジェクトに対応するキーポイントを選択するための手段と、

前記選択されたキーポイントおよび前記音源のタイプに基づいて前記オブジェクトを識別するための手段とを備える、装置。

【請求項 22】

前記シーンにおいて記録されたオーディオに基づいて、1 つまたは複数のオブジェクト

に対応する 1 つまたは複数のキーポイントシグネチャを選択するための手段と、

前記シーンの画像中の複数のキーポイントを識別するための手段と、

前記シーン中の前記オブジェクトを識別するために前記キーポイントを前記キーポイントシグネチャと比較するための手段とをさらに備える、請求項 2 1 に記載の装置。

【請求項 2 3】

前記シーンにおいて記録された前記オーディオに基づいて前記シーンの画像の一部分を選択するための手段と、

前記画像の前記一部分内からのみ前記キーポイントを選択するための手段とをさらに備える、請求項 2 1 に記載の装置。

【請求項 2 4】

前記シーンにおいて記録された前記オーディオに基づいて前記画像の一部分を選択するための前記手段が、

前記オーディオからオーディオ到来方向 (DOA) を判断するための手段と、

前記オーディオ DOA に基づいて前記画像の前記一部分を選択するための手段とを含む、請求項 2 3 に記載の装置。

【請求項 2 5】

前記オーディオ DOA を判断するための手段が、

前記シーンに位置する複数のマイクロフォンにおいて前記オーディオを受信し、それによって複数のマイクロフォン信号を生成するための手段と、

前記マイクロフォン信号に基づいて前記オーディオ DOA を判断するための手段とを含む、請求項 2 4 に記載の装置。

【請求項 2 6】

前記シーンのビデオ記録から複数の局所動きベクトルを計算するための手段と、

前記局所動きベクトルを 1 つまたは複数のオブジェクトに対応する所定の局所動きベクトルのデータベースと比較することによって、および前記キーポイントを 1 つまたは複数のキーポイントシグネチャと比較することによって前記オブジェクトを識別するための手段とをさらに備える、請求項 2 1 に記載の装置。

【請求項 2 7】

前記キーポイントを 1 つまたは複数のキーポイントシグネチャと比較するための手段をさらに備える、請求項 2 1 に記載の装置。

【請求項 2 8】

前記音響認識特徴がメル周波数ケプストラム係数を含む、請求項 2 7 に記載の装置。

【請求項 2 9】

画像中に現れる 1 つまたは複数のオブジェクトについての範囲情報を判断するための手段と、

前記範囲情報に基づいて前記キーポイントを分析するための手段とをさらに備える、請求項 2 1 に記載の装置。

【請求項 3 0】

範囲情報を判断するための手段が、オートフォーカスカメラを使用して範囲情報を判断するための手段と、マルチカメラ画像視差推定を使用して範囲情報を判断するための手段と、上記の任意の組合せとからなるグループから選択される、請求項 2 9 に記載の装置。

【請求項 3 1】

シーンにおいて記録されたオーディオから複数の音響認識特徴を計算するためのコードと、

オブジェクトの音源のタイプを判断するために前記音響認識特徴を 1 つまたは複数のオブジェクトに対応する所定の音響認識特徴と比較するためのコードと、

前記音源のタイプに基づいて前記オブジェクトに対応するキーポイントを選択するためのコードと、

前記選択されたキーポイントおよび前記音源のタイプに基づいて前記オブジェクトを識別するためのコードとを備える、1 つまたは複数のプロセッサによって実行可能な命令の

セットを具備する非一時的有形コンピュータ可読媒体。

【請求項 3 2】

前記シーンにおいて記録されたオーディオに基づいて、1つまたは複数のオブジェクトに対応する1つまたは複数のキーポイントシグネチャを選択するためのコードと、

前記シーンの画像中の複数のキーポイントを識別するためのコードと、

前記シーン中の前記オブジェクトを識別するために前記キーポイントを前記キーポイントシグネチャと比較するためのコードとをさらに備える、請求項 3 1 に記載のコンピュータ可読媒体。

【請求項 3 3】

前記シーンにおいて記録された前記オーディオに基づいて画像の一部分を選択するためのコードと、

前記画像の前記一部分内からのみ前記キーポイントを選択するためのコードとをさらに備える、請求項 3 1 に記載のコンピュータ可読媒体。

【請求項 3 4】

前記シーンにおいて記録された前記オーディオに基づいて前記画像の一部分を選択するための前記コードが、

前記オーディオからオーディオ到来方向(DOA)を判断するためのコードと、

前記オーディオDOAに基づいて前記画像の前記一部分を選択するためのコードとを含む、請求項 3 3 に記載のコンピュータ可読媒体。

【請求項 3 5】

前記オーディオDOAを判断するためのコードが、

前記シーンに位置する複数のマイクロフォンにおいて前記オーディオを受信し、それによって複数のマイクロフォン信号を生成するためのコードと、

前記マイクロフォン信号に基づいて前記オーディオDOAを判断するためのコードとを含む、請求項 3 4 に記載のコンピュータ可読媒体。

【請求項 3 6】

前記シーンのビデオ記録から複数の局所動きベクトルを計算するためのコードと、

前記局所動きベクトルを1つまたは複数のオブジェクトに対応する所定の局所動きベクトルのデータベースと比較することによって、および前記キーポイントを1つまたは複数のキーポイントシグネチャと比較することによって前記オブジェクトを識別するためのコードとをさらに備える、請求項 3 1 に記載のコンピュータ可読媒体。

【請求項 3 7】

前記キーポイントを1つまたは複数のキーポイントシグネチャと比較するためのコードをさらに備える、請求項 3 1 に記載のコンピュータ可読媒体。

【請求項 3 8】

前記音響認識特徴がメル周波数ケプストラム係数を含む、請求項 3 7 に記載のコンピュータ可読媒体。

【請求項 3 9】

画像中に現れる1つまたは複数のオブジェクトについての範囲情報を判断するためのコードと、

前記範囲情報に基づいて前記キーポイントを分析するためのコードとをさらに備える、請求項 3 1 に記載のコンピュータ可読媒体。

【請求項 4 0】

範囲情報を判断するためのコードが、オートフォーカスカメラを使用して範囲情報を判断するためのコードと、マルチカメラ画像視差推定を使用して範囲情報を判断するためのコードと、上記の任意の組合せとからなるグループから選択される、請求項 3 9 に記載のコンピュータ可読媒体。