

【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】令和 2 年 7 月 2 日 (2020.7.2)

【公表番号】特表 2019-523902 (P2019-523902A)
 【公表日】令和 1 年 8 月 29 日 (2019.8.29)
 【年通号数】公開・登録公報 2019-035
 【出願番号】特願 2018-561492 (P2018-561492)
 【国際特許分類】

G 1 0 L 19/00 (2013.01)

G 1 0 L 19/008 (2013.01)

H 0 4 N 21/854 (2011.01)

【F I】

G 1 0 L 19/00 1 0 0

G 1 0 L 19/008 2 0 0

H 0 4 N 21/854

【手続補正書】

【提出日】令和 2 年 5 月 21 日 (2020.5.21)

【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】

オリジナルの音声要素及びビデオ要素を有するコンテンツからプレゼンテーションを生成する装置であって、

前記コンテンツは複数のシーンを含み、各々のシーンは、音声特性を有する各々の環境に関連しており、

前記装置は、オーディオプロセッサ、オーサリングツール及びエンコーダを備え、

前記オーディオプロセッサは、前記各々の環境を示す環境パラメータ及び前記音声要素を受け、前記オーディオプロセッサは、前記環境パラメータの音響特性に基づいて、前記オリジナルの音声要素を修正された音声要素に修正するように構成されており、

前記オーサリングツールは、前記修正された音声要素及び前記オリジナルのビデオ要素を受け、対応する符号化係数を生成し、

前記エンコーダは、前記オリジナルのビデオ要素及び前記修正された音声要素を受け、前記符号化係数に基づいて符号化コンテンツを生成する、ことを特徴とする装置。

【請求項 2】

前記装置は、前記プレゼンテーションが提供される少なくとも 1 つの劇場に対応する劇場音響パラメータを保存する音響メモリをさらに備え、

前記オーディオプロセッサは、前記修正された音声要素を生成するため、前記劇場音響パラメータを受けるように構成されている、ことを特徴とする請求項 1 に記載の装置。

【請求項 3】

前記劇場音響パラメータを検知するように適応された音響環境センサをさらに備え、

前記音響環境センサにより検知された信号は前記音響メモリに保存される、ことを特徴とする請求項 2 に記載の装置。

【請求項 4】

前記エンコーダは、VR デバイスにおけるプレゼンテーションのために、VR (仮想現実) 符号化コンテンツを生成するように構成されており、

前記ビデオ要素は、仮想音響環境特性を有する仮想環境で少なくとも第１の仮想シーンを定義し、

前記オーディオプロセッサは、前記仮想音響環境特性を受けるように、かつ仮想シーンのために構成された音声トラックを定義する係数を生成するように、構成されている、ことを特徴とする請求項１に記載の装置。

【請求項５】

前記第１の仮想シーンにおいて定義されたＶＲ環境を分析するように構成された環境アナライザをさらに備え、

前記環境アナライザは、音響メモリに保存するための仮想環境音響を生成する、ことを特徴とする請求項４に記載の装置。

【請求項６】

前記エンコーダは、受け取ったコンテンツ及びＡＲデバイスにおけるプレゼンテーションのために現在の環境から得られたコンテンツに基づいて、ＡＲ（拡張現実）符号化コンテンツを生成するように構成されており、

前記ビデオ要素は、仮想及び現実の要素を有する少なくとも第１のＡＲシーンを定義し、

前記オーディオプロセッサは、仮想及び現実の音声要素並びに前記第１のＡＲシーンの音響特性を受けるように構成されており、

前記オーディオプロセッサは、前記第１のＡＲシーンの音響特性と一致するように構成された音声信号の記述的係数を生成する、ことを特徴とする請求項１に記載の装置。

【請求項７】

前記オーディオプロセッサは、前記第１のＡＲシーンに関連する環境信号を含む音声信号の記述的係数を生成する、ことを特徴とする請求項６に記載の装置。

【請求項８】

視聴者の現在位置に関連する現実環境を検出する現実環境検出器と、

前記現実環境の音響パラメータを決定するように構成された音響環境検出器と、をさらに備える、ことを特徴とする請求項６に記載の装置。

【請求項９】

前記現実環境検出器はＧＰＳを含む、ことを特徴とする請求項８に記載の装置。

【請求項１０】

前記現実環境検出器は、

前記現在位置に関連するイメージを作成するように適応されたカメラと、

前記音響パラメータを決定するために前記イメージを分析するアナライザと、を含む、ことを特徴とする請求項８に記載の装置。

【請求項１１】

前記現実環境検出器は、

試験信号を生成するサウンドジェネレータと、

前記サウンド検出器からの反響したサウンドを検出するように構成されたマイクと、を含み、

前記音響環境検出器は、前記反響したサウンドを分析するように適応されている、ことを特徴とする請求項８に記載の装置。

【請求項１２】

符号化されたプレゼンテーションを生成するための方法であって、

複数のシーンを有するとともに、音声要素及びビデオ要素を含むコンテンツを受けるステップと、

前記プレゼンテーションを含む局所的な環境の音響特性及び前記複数のシーンの１つの仮想位置を決定するステップと、

前記音響特性に基づいて修正された音声要素を生成するように前記音声要素を修正するステップと、

前記ビデオ要素及び前記修正された音声要素の少なくとも一方に基づいて、符号化係数

を生成するステップと、

符号化されたプレゼンテーションを生成するように前記符号化係数を用いて前記コンテンツを符号化するステップと、

を含み、

前記符号化されたプレゼンテーションは、前記ビデオ要素及び前記修正された音声要素から導出される、ことを特徴とする方法。

【請求項 13】

前記プレゼンテーションを劇場において提供するステップを含み、

前記局所的な環境は前記劇場を含む、ことを特徴とする請求項 12 に記載の方法。

【請求項 14】

前記プレゼンテーションは、パーソナル V R デバイスを用いた視聴者用の V R プレゼンテーションであり、

前記決定するステップは、前記プレゼンテーションにおいて仮想シーンの仮想音響特性を決定することを含む、ことを特徴とする請求項 12 に記載の方法。

【請求項 15】

前記音声要素を修正するステップは、付加的なサウンドを加えることを含み、

前記付加的なサウンドは、前記仮想シーンに関連する音響環境に関連している、ことを特徴とする請求項 14 に記載の方法。

【請求項 16】

前記プレゼンテーションは、パーソナル A R デバイスを用いた視聴者用の A R プレゼンテーションであり、

前記 A R プレゼンテーションは、仮想及び現実シーンエレメントを有する A R シーンを含む、ことを特徴とする請求項 12 に記載の方法。

【請求項 17】

前記決定するステップは、前記仮想シーンエレメントの仮想音響特性を決定することを含む、ことを特徴とする請求項 16 に記載の方法。

【請求項 18】

前記決定するステップは、前記現実シーンエレメントの現実の音響特性を決定することを含む、ことを特徴とする請求項 16 に記載の方法。

【請求項 19】

前記音声要素を修正するステップは、前記音声要素に付加的なサウンドを加えることを含み、

前記付加的なサウンドは、前記仮想シーンエレメントに関連する音響環境に関連している、ことを特徴とする請求項 16 に記載の方法。

【請求項 20】

前記音声要素を修正するステップは、前記音声要素に付加的なサウンドに加えることを含み、

前記付加的なサウンドは、前記現実シーンエレメントに関連する音響環境に関連している、ことを特徴とする請求項 16 に記載の方法。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0029

【補正方法】変更

【補正の内容】

【0029】

一実施例では、前記現実環境検出器は、試験信号を生成するサウンドジェネレータと、前記サウンドジェネレータからの反響したサウンドを検出するように構成されたマイクと、を含む。前記音響環境検出器は、前記反響したサウンドを分析するように適応されている。

【手続補正 3】

【補正対象書類名】明細書

【補正対象項目名】0030

【補正方法】変更

【補正の内容】

【0030】

本発明の他の態様において、符号化された3Dプレゼンテーションを生成するための方法について説明する。前記方法は、いくつかの3Dシーンを有するとともに、音声及び視覚トラックを含む3Dコンテンツを受けるステップと、前記3Dプレゼンテーションが存在する局所的な環境の音響特性及び前記3Dシーンの1つの仮想位置を決定するステップと、前記音響特性に基づいて修正されたトラックを生成するように前記音声トラックを修正するステップと、符号化された3Dプレゼンテーションを生成するように、前記音声空間パラメータに少なくとも部分的に基づいて符号化係数を用いて前記3Dコンテンツを符号化するステップと、を含む。前記符号化された3Dプレゼンテーションは、視覚要素及び音声要素を含む。音声要素は、前記修正された音声トラックから導出される。

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0048

【補正方法】変更

【補正の内容】

【0048】

ARについて、展開される一つ又は複数のビデオトラッキングマップは、コンピュータビジョンなどのHMDに含まれる深度検出手段又はモーショントラッキング、オブジェクト又は顔認識ソフトウェア、加速度計、高度計、GPS受信機、電子テープメジャー、レーザ距離計、レーザ又はデジタル音声測定装置、又はジャイロセンサから生成され得る。トラッキングマップは、ARオーバレイ、又は、音声位置調整を含む、ARオブジェクト又はグラフィックと現実世界オブジェクトとを混合したものを生成するために用いられ得る。

【手続補正5】

【補正対象書類名】明細書

【補正対象項目名】0049

【補正方法】変更

【補正の内容】

【0049】

ある例では、関心の対象であるオブジェクトの音声位置及び視覚位置が対応し、音声及び視覚オブジェクトの双方のための3D空間情報が記録され、これにより、音声オブジェクトが関心の対象である視覚オブジェクトにマッピングされ得る。本発明の一実施例では、深度位置（例えば、Z軸に沿った位置）だけが付与される。他の実施例では、3D体積における音声オブジェクトの位置（例えば、全ての3軸X、Y及びZに沿った）は、任意の視覚オブジェクトの位置から離れて付与される。図1Aを参照すると、コンパレータ100は、音声及びビデオトラッキングマップを受け、当該マップが収束しているかを判断する。単一の視覚オブジェクト及び単一の音声が存在する場合、上記の判断は単純である。しかし、マルチ視覚オブジェクト及びマルチ音声オブジェクトが存在する場合、各々の音声オブジェクトを各々の視覚オブジェクトに割り当てるため、手動制御又は他の手段を用いてもよい。特定が完了すると、構成要素103により音声及び視覚オブジェクトを常に追跡するために、自動トラッキングソフトウェアが提供される。

【手続補正6】

【補正対象書類名】明細書

【補正対象項目名】0058

【補正方法】変更

【補正の内容】

【 0 0 5 8 】

ステップ 1 7 0 では、関心の対象である音声オブジェクトが視覚オブジェクトに関連付けられているかについての確認が行われる。関連付けられている場合、ステップ 1 8 0 において、対応する視覚オブジェクト（俳優 A , B）の位置が連続的に追跡され（ステップ 1 8 0）、対応する音声キューが生成される（ステップ 1 9 0）。音声オブジェクトが、視覚オブジェクトに対応していない（例えば、俳優 B が最初の位置にいるか、俳優 A が位置 A 1 にいる）場合、遮られたオブジェクト、スクリーン外のオブジェクト又は隠れたオブジェクトに適用可能な特定の規則を用いて種々の計算が行われる（ステップ 2 0 0）。例えば、隠れたオブジェクト（例えば、壁の後ろで動く俳優）の動きは、対応する音声オブジェクトを生成するために補間され使用され得る。この時点において、以前に受けた環境パラメータに基づいて、付加的なサウンドエフェクトが得られる。ステップ 2 1 0 において、音声及び視覚パラメータはオーサリングツール 2 2 に提供される。例えば、位置キュー（一次元、二次元又は三次元における）が音声要素に付加され得る。代替例として、音声 3 D ポジションマップ、テーブル又は他の便利な形態として表示される音声キューが別のフィールドに提供される。また、新たなサウンドエフェクトがオーサリングツールに提供され、これにより、ステップ 2 2 0 において生成された係数は、各々の音響環境に応じて修正されたサウンドトラックを定義する。

【 手 続 補 正 7 】

【 補 正 対 象 書 類 名 】 明 細 書

【 補 正 対 象 項 目 名 】 0 0 6 1

【 補 正 方 法 】 変 更

【 補 正 の 内 容 】

【 0 0 6 1 】

次いで、オーディオプロセッサ 2 1 は、ログ 2 0 及び / 又はアナライザ 3 0 から得られる情報並びに手動又は自動で生成された付加的な制御信号及び環境パラメータを使用し、関心の対象である種々の音声オブジェクトの位置を定義する一セットの音声パラメータを生成する。これらの音声パラメータは、オーサリングツール 2 2 に提供される。また、オーサリングツールは、各々のフレーム又は一群のフレーム又は視覚オブジェクトの位置に関する対応するビデオ情報を、ビデオプロセッサ 2 3 から受ける。

【 手 続 補 正 8 】

【 補 正 対 象 書 類 名 】 明 細 書

【 補 正 対 象 項 目 名 】 0 0 7 5

【 補 正 方 法 】 変 更

【 補 正 の 内 容 】

【 0 0 7 5 】

一実施例では、エンコーダ 7 0 は、深度知覚キューを用いて音声信号を修正するだけでなく、関連のある逆変換を含むメタデータを作成する。これにより、下流側の信号プロセッサ（図示せず）は、2 D 視聴用にオリジナルの（又は、プロセスに損失がある場合にはオリジナルに近い）最適化されていないデジタル 3 D 映画音声ストリームを導出することができる。第 2 の符号化 3 D コンテンツとともにメタデータが保存される。

【 手 続 補 正 9 】

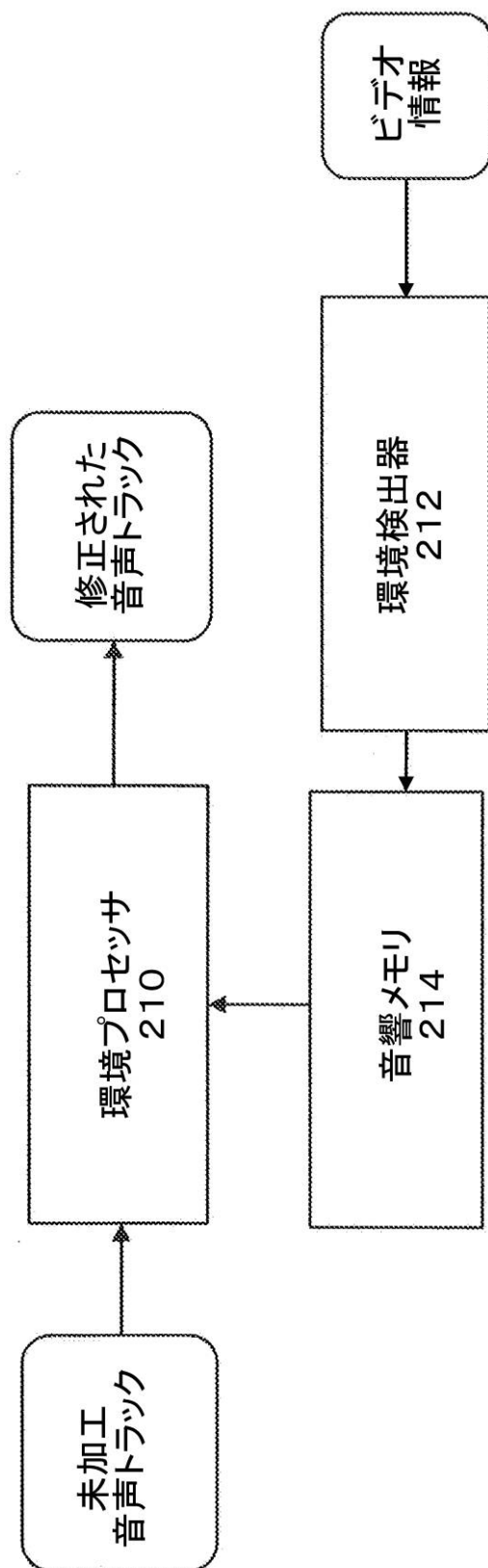
【 補 正 対 象 書 類 名 】 図 面

【 補 正 対 象 項 目 名 】 図 6 B

【 補 正 方 法 】 変 更

【 補 正 の 内 容 】

【図 6 B】



【手続補正 10】

【補正対象書類名】図面

【補正対象項目名】図 6 C

【補正方法】変更

【補正の内容】

【図 6 C】

