

【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 3 区分
 【発行日】令和 2 年 12 月 17 日 (2020.12.17)

【公表番号】特表 2020-507141 (P2020-507141A)
 【公表日】令和 2 年 3 月 5 日 (2020.3.5)
 【年通号数】公開・登録公報 2020-009
 【出願番号】特願 2019-521695 (P2019-521695)
 【国際特許分類】

G 0 6 T 13/40 (2011.01)
 H 0 4 N 5/232 (2006.01)
 H 0 4 N 13/344 (2018.01)
 H 0 4 N 13/293 (2018.01)
 G 1 0 L 15/10 (2006.01)
 G 0 6 T 13/20 (2011.01)

【F I】

G 0 6 T 13/40
 H 0 4 N 5/232 2 9 0
 H 0 4 N 13/344
 H 0 4 N 13/293
 G 1 0 L 15/10 5 0 0 Z
 G 0 6 T 13/20 5 0 0

【手続補正書】
 【提出日】令和 2 年 11 月 6 日 (2020.11.6)
 【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】

発話中のユーザの顔を動画にするためのウェアラブルシステムであって、前記ウェアラブルシステムは、

画像を捕捉するように構成された内向きに面した結像システムであって、前記内向きに面した結像システムは、1 つ以上のカメラを備え、前記 1 つ以上のカメラは、前記ウェアラブルシステムが前記ユーザによって装着されているとき、ユーザの顔の眼球周囲領域が前記内向きに面した結像システムによって観察可能であり、前記ユーザの下側顔が前記内向きに面した結像システムによって観察不可能であるように位置付けられている、内向きに面した結像システムと、

前記ユーザの発話を受信するように構成されたオーディオセンサと、
 ハードウェアプロセッサと
 を備え、

前記ハードウェアプロセッサは、

前記内向きに面した結像システムを介して、前記ウェアラブルシステムが前記ユーザによって装着されているとき、前記ユーザの前記眼球周囲領域の画像を入手することと、

少なくとも部分的に前記ユーザの前記眼球周囲領域の前記画像に基づいて、前記ユーザの少なくとも前記眼球周囲領域の眼球周囲外形をエンコードする眼球周囲顔パラメータを生成することと、

前記オーディオセンサによって、前記ユーザによって発話されたオーディオストリーム

を入手することと、

前記オーディオストリームにおける音素を識別することと、

前記ユーザを含まないグループの人々に関連付けられる画像を用いて生成された基本モデルにアクセスすることと、

少なくとも部分的に前記基本モデルと前記ユーザの前記眼球周囲領域の前記画像とに基づいてマッピングをカスタマイズすることであって、前記マッピングの入力は、前記音素と前記ユーザの前記眼球周囲領域の前記画像とを備え、前記マッピングの出力は、前記ユーザの前記下側顔の外形をエンコードする下側顔パラメータを備え、前記下側顔パラメータは、前記音素と前記ユーザの前記眼球周囲領域の前記画像との分析から推測される、ことと、

前記マッピングを前記ユーザの前記眼球周囲領域の前記画像に適用し、前記下側顔パラメータを生成することと、

前記眼球周囲顔パラメータと前記下側顔パラメータとを組み合わせ、3次元(3D)顔モデルに関連付けられた完全な顔パラメータを生成することと、

少なくとも部分的に前記完全な顔パラメータに基づいて、前記ユーザの顔の動画を生成することと

を行うようにプログラムされている、ウェアラブルシステム。

【請求項2】

前記3D顔モデルは、変形可能線形モデルを備え、前記眼球周囲顔パラメータおよび前記下側顔パラメータは、前記ユーザが発話しているときの前記顔の変形を記述する、請求項1に記載のウェアラブルシステム。

【請求項3】

前記完全な顔パラメータを生成するために、前記ハードウェアプロセッサは、前記3D顔モデルを更新し、前記下側顔パラメータまたは前記眼球周囲顔パラメータのうちの少なくとも1つに対する更新を反映するようにプログラムされている、請求項2に記載のウェアラブルシステム。

【請求項4】

前記マッピングの前記入力、眼特定の情報、身体移動、または心拍数のうちの少なくとも1つをさらに含む、請求項1に記載のウェアラブルシステム。

【請求項5】

前記眼特定情報は、前記ユーザの眼姿勢、瞳孔拡張状態、眼色、または眼瞼状態のうちの少なくとも1つを含む、請求項4に記載のウェアラブルシステム。

【請求項6】

前記下側顔パラメータは、前記オーディオストリームにおける音素を視覚的に記述する口形素をエンコードする、請求項1に記載のウェアラブルシステム。

【請求項7】

前記マッピングをカスタマイズするために、前記ハードウェアプロセッサは、少なくとも部分的に前記眼球周囲領域の前記画像に基づいて、前記ユーザの顔の皮膚テクスチャを推測するようにプログラムされ、前記ユーザの顔の前記動画は、前記顔の前記皮膚テクスチャを組み込んでいる、請求項1に記載のウェアラブルシステム。

【請求項8】

前記内向きに面した結像システムは、眼カメラを備え、前記内向きに面した結像システムによって入手された前記眼球周囲領域の前記画像は、第1の眼のための前記眼球周囲領域の画像を備えている、請求項1に記載のウェアラブルシステム。

【請求項9】

前記完全な顔パラメータを生成するために、前記ハードウェアプロセッサは、

前記内向きに面した結像システムによって入手された前記眼球周囲領域の前記画像に基づいて、第2の眼のための眼球周囲顔パラメータを決定することと、

前記第2の眼のための前記眼球周囲顔パラメータを前記完全な顔パラメータの中に組み込むことと

を行うようにプログラムされている、請求項 8 に記載のウェアラブルシステム。

【請求項 10】

前記完全な顔パラメータが適用され、前記ユーザの顔の前記動画を生成するようにするために、前記ハードウェアプロセッサは、命令を複合現実ディスプレイを備えているウェアラブルデバイスに通信するようにプログラムされており、前記命令は、前記完全な顔パラメータが適用され、前記 3D 顔モデルを中立位置から変動させるようにする、請求項 1 に記載のウェアラブルシステム。

【請求項 11】

発話中のユーザの顔を動画にする方法であって、前記方法は、

画像を捕捉するように構成された内向きに面した結像システムによって入手されたユーザの眼球周囲領域の画像にアクセスすることであって、前記内向きに面した結像システムは、1つ以上のカメラを備え、前記1つ以上のカメラは、ウェアラブルシステムが前記ユーザによって装着されているとき、前記ユーザの顔の眼球周囲領域が前記内向きに面した結像システムによって観察可能であり、前記ユーザの下側顔が前記内向きに面した結像システムによって観察不可能であるように位置付けられている、ことと、

少なくとも部分的に前記画像に基づいて、前記ユーザの少なくとも前記眼球周囲領域の眼球周囲外形をエンコードする眼球周囲顔パラメータを決定することと、

オーディオセンサによって入手されたユーザによって発話されたオーディオストリームにアクセスすることと、

前記オーディオストリームにおける音素を識別することと、

前記ユーザを含まないグループの人々に関連付けられる画像を用いて生成された基本モデルにアクセスすることと、

少なくとも部分的に前記基本モデルと前記ユーザの前記眼球周囲領域の前記画像とに基づいてマッピングをカスタマイズすることであって、前記マッピングの入力は、前記音素と前記ユーザの前記眼球周囲領域の前記画像とを備え、前記マッピングの出力は、前記ユーザの前記下側顔の外形をエンコードする下側顔パラメータを備え、前記下側顔パラメータは、前記音素と前記ユーザの前記眼球周囲領域の前記画像との分析から推測される、ことと、

前記マッピングを前記画像に適用し、前記下側顔パラメータを生成することと、

前記眼球周囲顔パラメータと前記下側顔パラメータとを組み合わせ、3次元(3D)顔モデルに関連付けられた完全な顔パラメータを生成することと、

少なくとも部分的に前記完全な顔パラメータに基づいて、完全な顔画像を生成することと

を含む、方法。

【請求項 12】

前記 3D 顔モデルは、変形可能線形モデルを備え、前記眼球周囲顔パラメータおよび前記下側顔パラメータは、前記ユーザが発話しているときの前記顔の変形を記述する、請求項 11 に記載の方法。

【請求項 13】

前記完全な顔パラメータを生成するために、前記ハードウェアプロセッサは、前記 3D 顔モデルを更新し、前記下側顔パラメータまたは前記眼球周囲顔パラメータのうちの少なくとも 1 つに対する更新を反映するようにプログラムされている、請求項 12 に記載の方法。

【請求項 14】

前記完全な顔パラメータは、前記ユーザの顔に関連付けられた動画を決定するために、眼特定の情報と組み合わせられる、請求項 11 に記載の方法。

【請求項 15】

前記眼特定の情報は、前記ユーザの眼姿勢、瞳孔拡張状態、眼色、または眼瞼状態のうちの少なくとも 1 つを含む、請求項 14 に記載の方法。

【請求項 16】

前記下側顔パラメータは、前記オーディオストリームにおける音素を視覚的に記述する口形素をエンコードする、請求項 1 1 に記載の方法。

【請求項 1 7】

前記完全な顔画像は、前記ユーザの皮膚テクスチャをさらに組み込んでおり、前記ユーザの前記皮膚テクスチャは、少なくとも部分的に前記内向きに面した結像システムによって入手された前記画像に基づいて決定される、請求項 1 1 に記載の方法。

【請求項 1 8】

前記マッピングは、眼球周囲顔パラメータが下側顔パラメータに関連付けられている可能性を含み、前記下側顔パラメータは、それらが閾値基準に達していることの決定に応答して、前記完全な顔画像を生成するために選択される、請求項 1 1 に記載の方法。

【請求項 1 9】

前記画像は、静止画像またはビデオフレームのうちの少なくとも 1 つを含む、請求項 1 1 に記載の方法。

【請求項 2 0】

前記完全な顔画像を複合現実環境内にレンダリングするように頭部搭載型ディスプレイに命令することをさらに含む、請求項 1 1 に記載の方法。

【請求項 2 1】

前記基本モデルは、前記グループの人々の画像に基づいて生成される 3 D 変形可能線形モデルを備えている、請求項 1 に記載のウェアラブルシステム。

【請求項 2 2】

前記基本モデルは、前記グループの人々の画像に基づいて生成される 3 D 変形可能線形モデルを備えている、請求項 1 1 に記載の方法。

【請求項 2 3】

前記ユーザの前記下側顔は、前記ウェアラブルシステムが前記ユーザによって装着されているとき、前記ウェアラブルシステムに関連付けられる任意のカメラによって観察されない、請求項 1 に記載のウェアラブルシステム。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0 0 0 5

【補正方法】変更

【補正の内容】

【0 0 0 5】

本明細書に説明される主題の 1 つ以上の実装の詳細が、付随の図面および以下の説明に記載される。他の特徴、側面、および利点は、説明、図面、ならびに請求項から明白となるであろう。本概要または以下の発明を実施するための形態のいずれも、本発明の主題の範囲を定義または限定することを主張するものではない。

本願明細書は、例えば、以下の項目も提供する。

(項目 1)

発話中のユーザの顔を動画にするためのウェアラブルシステムであって、前記ウェアラブルシステムは、

ユーザの眼球周囲領域を結像するように構成された内向きに面した結像システムと、

前記ユーザの発話を受信するように構成されたオーディオセンサと、

ハードウェアプロセッサと

を備え、

前記ハードウェアプロセッサは、

前記内向きに面した結像システムを介して、前記ユーザの前記眼球周囲領域の画像を入手することと、

少なくとも部分的に前記画像に基づいて、前記ユーザの少なくとも前記眼球周囲領域の眼球周囲外形をエンコードする眼球周囲顔パラメータを生成することと、

前記オーディオセンサによって、前記ユーザによって発話されたオーディオストリームを入手することと、

前記オーディオストリームにおける音素を識別することと、

マッピングにアクセスすることであって、前記マッピングの入力は、前記音素と前記ユーザの前記眼球周囲領域の前記画像とを備え、前記マッピングの出力は、前記ユーザの少なくとも下側顔の下側顔外形をエンコードする下側顔パラメータを備え、前記下側顔は、前記ユーザによって装着されているとき、前記内向きに面した結像システムによって観察されていない、ことと、

前記マッピングを前記画像に適用し、下側顔パラメータを生成することと、

前記眼球周囲顔パラメータと前記下側顔パラメータとを組み合わせ、3次元(3D)顔モデルに関連付けられた完全な顔パラメータを生成することと、

少なくとも部分的に前記完全な顔パラメータに基づいて、前記ユーザの顔の動画を生成することと

を行うようにプログラムされている、ウェアラブルシステム。

(項目2)

前記3D顔モデルは、変形可能線形モデルを備え、前記眼球周囲顔パラメータおよび前記下側顔パラメータは、前記ユーザが発話しているときの前記顔の変形を記述する、項目1に記載のウェアラブルシステム。

(項目3)

前記完全な顔パラメータを生成するために、前記ハードウェアプロセッサは、前記3D顔モデルを更新し、前記下側顔パラメータまたは前記眼球周囲顔パラメータのうちの少なくとも1つに対する更新を反映するようにプログラムされている、項目2に記載のウェアラブルシステム。

(項目4)

前記マッピングの前記入力、眼特定の情報、身体移動、または心拍数のうちの少なくとも1つをさらに含む、項目1に記載のウェアラブルシステム。

(項目5)

前記眼特定の情報は、前記ユーザの眼姿勢、瞳孔拡張状態、眼色、または眼瞼状態のうちの少なくとも1つを含む、項目4に記載のウェアラブルシステム。

(項目6)

前記下側顔パラメータは、前記オーディオストリームにおける音素を視覚的に記述する口形素をエンコードする、項目1に記載のウェアラブルシステム。

(項目7)

前記ハードウェアプロセッサは、少なくとも部分的に前記眼球周囲領域の前記画像に基づいて、前記ユーザの顔の皮膚テクスチャを推測するようにさらにプログラムされ、前記ユーザの顔の前記動画は、前記顔の前記皮膚テクスチャを組み込んでいる、項目1に記載のウェアラブルシステム。

(項目8)

前記内向きに面した結像システムは、眼カメラを備え、前記内向きに面した結像システムによって入手された前記眼球周囲領域の前記画像は、第1の眼のための前記眼球周囲領域の画像を備えている、項目1に記載のウェアラブルシステム。

(項目9)

前記完全な顔パラメータを生成するために、前記ハードウェアプロセッサは、

前記内向きに面した結像システムによって入手された前記眼球周囲領域の前記画像に基づいて、第2の眼のための眼球周囲顔パラメータを決定することと、

前記第2の眼のための前記眼球周囲顔パラメータを前記完全な顔パラメータの中に組み込むことと

を行うようにプログラムされている、項目8に記載のウェアラブルシステム。

(項目10)

前記完全な顔パラメータが適用され、前記ユーザの顔の前記動画を生成するようにする

ために、前記ハードウェアプロセッサは、命令を複合現実ディスプレイを備えているウェアラブルデバイスに通信するようにプログラムされており、前記命令は、前記完全な顔パラメータが適用され、前記3D顔モデルを中立位置から変動させるようにする、項目1に記載のウェアラブルシステム。

(項目11)

発話中のユーザの顔を動画にする方法であって、前記方法は、
ユーザの眼球周囲領域を結像するように構成された内向きに面した結像システムによって入手された画像にアクセスすることと、
少なくとも部分的に前記画像に基づいて、前記ユーザの少なくとも前記眼球周囲領域の眼球周囲外形をエンコードする眼球周囲顔パラメータを決定することと、
オーディオセンサによって入手されたユーザによって発話されたオーディオストリームにアクセスすることと、
前記オーディオストリームにおける音素を識別することと、
マッピングにアクセスすることであって、前記マッピングの入力は、前記音素と前記ユーザの前記眼球周囲領域の前記画像とを備え、前記マッピングの出力は、前記ユーザの少なくとも下側顔の下側顔外形をエンコードする下側顔パラメータを備えている、ことと、
前記マッピングを前記画像に適用し、下側顔パラメータを生成することと、
前記眼球周囲顔パラメータと前記下側顔パラメータとを組み合わせ、3次元(3D)顔モデルに関連付けられた完全な顔パラメータを生成することと、
少なくとも部分的に前記完全な顔パラメータに基づいて、完全な顔画像を生成することと
を含む、方法。

(項目12)

前記3D顔モデルは、変形可能線形モデルを備え、前記眼球周囲顔パラメータおよび前記下側顔パラメータは、前記ユーザが発話しているときの前記顔の変形を記述する、項目11に記載の方法。

(項目13)

前記完全な顔パラメータを生成するために、前記ハードウェアプロセッサは、前記3D顔モデルを更新し、前記下側顔パラメータまたは前記眼球周囲顔パラメータのうちの少なくとも1つに対する更新を反映するようにプログラムされている、項目12に記載の方法。

(項目14)

前記完全な顔パラメータは、前記ユーザの顔に関連付けられた動画を決定するために、眼特定の情報と組み合わせられる、項目11に記載の方法。

(項目15)

前記眼特定の情報は、前記ユーザの眼姿勢、瞳孔拡張状態、眼色、または眼瞼状態のうちの少なくとも1つを含む、項目14に記載の方法。

(項目16)

前記下側顔パラメータは、前記オーディオストリームにおける音素を視覚的に記述する口形素をエンコードする、項目11に記載の方法。

(項目17)

前記完全な顔画像は、前記ユーザの皮膚テクスチャをさらに組み込んでおり、前記ユーザの前記皮膚テクスチャは、少なくとも部分的に前記内向きに面した結像システムによって入手された前記画像に基づいて決定される、項目11に記載の方法。

(項目18)

前記マッピングは、眼球周囲顔パラメータが下側顔パラメータに関連付けられている可能性を含み、前記下側顔パラメータは、それらが閾値基準に達していることの決定に応答して、前記完全な顔画像を生成するために選択される、項目11に記載の方法。

(項目19)

前記画像は、静止画像またはビデオフレームのうちの少なくとも1つを含む、項目11

に記載の方法。

(項目 2 0)

前記完全な顔画像を複合現実環境内にレンダリングするように頭部搭載型ディスプレイに命令することをさらに含む、項目 1 1 に記載の方法。