

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6392497号
(P6392497)

(45) 発行日 平成30年9月19日(2018.9.19)

(24) 登録日 平成30年8月31日(2018.8.31)

(51) Int.Cl.

F I

H O 4 N 21/2343 (2011.01)

H O 4 N 21/2343

H O 4 N 21/44 (2011.01)

H O 4 N 21/44

H O 4 N 7/15 (2006.01)

H O 4 N 7/15 1 7 O

G O 6 F 13/00 (2006.01)

G O 6 F 13/00 6 5 O R

請求項の数 25 (全 19 頁)

(21) 出願番号 特願2013-38804 (P2013-38804)
 (22) 出願日 平成25年2月28日(2013.2.28)
 (65) 公開番号 特開2013-243646 (P2013-243646A)
 (43) 公開日 平成25年12月5日(2013.12.5)
 審査請求日 平成28年2月26日(2016.2.26)
 (31) 優先権主張番号 2012902108
 (32) 優先日 平成24年5月22日(2012.5.22)
 (33) 優先権主張国 オーストラリア(AU)

(73) 特許権者 511307096
 コモンウェルス サイエントフィック
 アンド インダストリアル リサーチ オ
 ーガニゼーション
 オーストラリア国 アクトン エーシーテ
 ィー 2601, クルーニーズ ロス
 ストリート
 (74) 代理人 100068755
 弁理士 恩田 博宣
 (74) 代理人 100105957
 弁理士 恩田 誠
 (74) 代理人 100142907
 弁理士 本田 淳

最終頁に続く

(54) 【発明の名称】 ビデオを生成するためのシステムおよび方法

(57) 【特許請求の範囲】

【請求項 1】

画像を生成する方法であって、

第1の画像を第1のコンピューティング装置で受信すること、

前記第1のコンピューティング装置のデータインタフェースで前記第1の画像をサーバに送信すること、

前記第1の画像の態様に対応する形状データを前記データインタフェースで前記サーバから受信すること、

少なくとも前記形状データおよびアバターデータに基づいて、一次出力画像を前記第1のコンピューティング装置のプロセッサによって生成すること、

前記データインタフェースでメッセージを前記サーバに送信することであって、前記メッセージは、第2のコンピューティング装置の詳細および前記アバターデータを含む、前記メッセージを送信すること、

前記第2のコンピューティング装置による前記一次出力画像の生成のために前記形状データおよび前記アバターデータを、前記第2のコンピューティング装置の詳細を用いて前記サーバから前記第2のコンピューティング装置に送信すること

を備える方法。

【請求項 2】

前記形状データが、前記第1の画像において描写されるユーザの顔の3Dモデルを含み、前記一次出力画像を生成することが、前記3Dモデルを前記アバターデータに適用する

10

20

ことを含む、請求項 1 に記載の方法。

【請求項 3】

前記形状データが、A A M（アクティブ外観モデル）パラメータを含み、前記一次出力画像を生成することが、前記 A A M パラメータをアバターデータに適用することを含む、請求項 1 に記載の方法。

【請求項 4】

前記一次出力画像を生成することが、前記アバターデータに応じた前記形状データへのマッピング機能を適用することをさらに含む、請求項 1 ～ 3 のいずれか一項に記載の方法。

【請求項 5】

前記アバターデータに関連する識別子を前記データインタフェースで受信することをさらに備える請求項 1 ～ 4 のいずれか一項に記載の方法。

【請求項 6】

前記識別子を前記サーバに送信することをさらに備え、
前記サーバが、前記形状データと、前記識別子に関連するサーバアバターデータとに基づいて、サーバ出力画像を生成する、請求項 5 に記載の方法。

【請求項 7】

前記サーバ出力画像が、前記一次出力画像と同一である、請求項 6 に記載の方法。

【請求項 8】

前記サーバは、前記第 1 のコンピューティング装置によって前記一次出力画像を生成するために用いられるアルゴリズムとは異なるアルゴリズムに従って前記サーバ出力画像を生成する、請求項 6 に記載の方法。

【請求項 9】

前記アバターデータを前記サーバに送信することをさらに備え、
前記サーバアバターデータが、前記サーバに送信されたアバターデータを含む、請求項 6 ～ 8 のいずれか一項に記載の方法。

【請求項 10】

前記第 2 のコンピューティング装置の詳細を用いて、前記サーバ出力画像を前記サーバから前記第 2 のコンピューティング装置のユーザに送信すること
をさらに備える請求項 6 ～ 9 のいずれか一項に記載の方法。

【請求項 11】

前記一次出力画像を生成することが、前記第 1 の画像を修正することを含む、請求項 1 ～ 10 のいずれか一項に記載の方法。

【請求項 12】

サーバ出力画像をソーシャルメディアサイトまたはビデオ共有サイトにアップロードする要求を、前記データインタフェースで前記サーバに送信することであって、前記ソーシャルメディアサイトまたは前記ビデオ共有サイトの詳細および選択されたアバターを含む前記要求を前記サーバに送信すること、

前記ソーシャルメディアサイトまたは前記ビデオ共有サイトの詳細を用いて、前記サーバから前記ソーシャルメディアサイトまたは前記ビデオ共有サイトに前記サーバ出力画像を送信すること、
をさらに備える請求項 6 ～ 11 のいずれか一項に記載の方法。

【請求項 13】

前記第 1 の画像が、ビデオシーケンスの画像を含む、請求項 1 ～ 12 のいずれか一項に記載の方法。

【請求項 14】

前記第 1 の画像および第 2 の画像がビデオシーケンスの画像を含むときに、前記第 2 の画像を前記第 1 のコンピューティング装置で受信すること、

前記第 2 の画像を前記データインタフェースで前記サーバに送信すること、

前記第 2 の画像の態様に対応するさらなる形状データを前記データインタフェースで受

10

20

30

40

50

信すること、

少なくとも前記さらなる形状データおよび前記アバターデータに基づいて、第2の一次出力画像を前記プロセッサによって生成すること、
をさらに備える請求項1～13のいずれか一項に記載の方法。

【請求項15】

前記第1の画像および第2の画像がビデオシーケンスの画像を含むときに、前記第2の画像を前記第1のコンピューティング装置で受信すること、

前記第2の画像を前記データインタフェースで前記サーバに送信すること、

前記第2の画像の態様に対応するさらなる形状データを前記データインタフェースで受信すること、

10

少なくとも前記さらなる形状データおよび第2のアバターデータに基づいて、第2の一次出力画像を前記プロセッサによって生成すること、
をさらに備える請求項1～13のいずれか一項に記載の方法。

【請求項16】

前記第2のアバターデータが、前記アバターデータを含む、請求項15に記載の方法。

【請求項17】

前記第1の画像が、前記第2の画像を受信する前に前記サーバに送信される、請求項14～16のいずれか一項に記載の方法。

【請求項18】

音声データを前記第1のコンピューティング装置で受信すること、

20

前記音声データを前記データインタフェースで前記サーバに送信すること、

前記音声データに関係する音声パラメータを前記データインタフェースで受信すること

、
前記音声パラメータおよび前記アバターデータに基づいて、出力音声を前記第1のコンピューティング装置のプロセッサによって生成すること、
をさらに備える請求項1～17のいずれか一項に記載の方法。

【請求項19】

前記出力音声を生成することが、前記音声データにフィルタを適用することを含む、請求項18に記載の方法。

【請求項20】

30

前記アバターデータが、アバター画像を含む、請求項1～19のいずれか一項に記載の方法。

【請求項21】

画像を生成するためのシステムであって、

第1のコンピューティング装置を備え、該第1のコンピューティング装置が、

データインタフェースと、

前記データインタフェースに接続されたプロセッサと、

前記プロセッサに接続されたメモリと

を含み、前記メモリが、前記プロセッサによって実行可能な命令として、

第1の画像を受信する命令、

40

前記第1の画像を前記データインタフェースでサーバに送信する命令、

前記第1の画像の態様に対応する形状データを前記データインタフェースで受信する命令、

少なくとも前記形状データとアバターデータとに基づいて出力画像を生成する命令

前記データインタフェースでメッセージを前記サーバに送信する命令であって、前記メッセージは、第2のコンピューティング装置の詳細および前記アバターデータを含む、前記メッセージを送信する命令、

前記第2のコンピューティング装置による前記出力画像の生成のために前記形状データおよび前記アバターデータを前記第2のコンピューティング装置の詳細を用いて前記サーバから前記第2のコンピューティング装置に送信する命令

50

を含む、システム。

【請求項 2 2】

前記プロセッサに接続された表示スクリーンをさらに含み、

前記メモリが、前記出力画像を前記表示スクリーン上に表示するための命令をさらに含む、請求項 2 1 に記載のシステム。

【請求項 2 3】

前記プロセッサに接続されたカメラをさらに含み、

前記メモリが、前記カメラから前記第 1 の画像を受信するための命令をさらに含む、請求項 2 1 または 2 2 に記載のシステム。

【請求項 2 4】

画像を生成および送信するためのシステムであって、

第 1 のコンピューティング装置を備え、該第 1 のコンピューティング装置が、

データインタフェースと、

前記データインタフェースに接続されたプロセッサと、

前記プロセッサに接続されたメモリと

を含み、前記メモリが、前記プロセッサによって実行可能な命令として、

第 1 の画像を受信する命令、

前記第 1 の画像を前記データインタフェースでサーバに送信する命令、

前記第 1 の画像の態様に対応する形状データを前記データインタフェースで受信する命令、および

少なくとも前記形状データとアバタデータとに基づいて出力画像を生成する命令

前記データインタフェースでメッセージを前記サーバに送信する命令であって、前記メッセージは、第 2 のコンピューティング装置の詳細および前記アバタデータを含む、前記メッセージを送信する命令、

を含み、前記サーバが、

サーバデータインタフェースと、

前記サーバデータインタフェースに接続されたサーバプロセッサと、

前記サーバプロセッサに接続されたサーバメモリと

を含み、前記サーバメモリが、前記サーバプロセッサによって実行可能な命令として、

前記第 1 の画像を受信する命令、

前記第 1 の画像の態様に応じた前記形状データを前記サーバプロセッサによって生成する命令、

前記形状データを前記サーバデータインタフェースで前記第 1 のコンピューティング装置に送信する命令、

前記サーバデータインタフェースで前記メッセージを前記第 1 のコンピューティング装置から受信する命令であって、前記メッセージは、前記第 2 のコンピューティング装置の詳細および前記アバタデータを含む、前記メッセージを受信する命令、

前記第 2 のコンピューティング装置による前記出力画像の生成のために前記形状データおよび前記アバタデータを前記第 2 のコンピューティング装置の詳細を用いて前記第 2 のコンピューティング装置に送信する命令

を含む、システム。

【請求項 2 5】

前記サーバプロセッサに接続されたサーバメモリが、前記サーバプロセッサによって実行可能な命令として、

少なくとも前記形状データと前記アバタデータとに基づいてサーバ出力画像を前記サーバプロセッサで生成する命令、

前記サーバデータインタフェースで前記サーバ出力画像を第 2 のコンピューティング装置に前記第 2 のコンピューティング装置の詳細を用いて送信する命令

をさらに含む、請求項 2 4 に記載のシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像生成に関し、特に限定されないが、ビデオシーケンス用の画像の生成に関する。

【背景技術】

【0002】

今ではデジタルカメラは多数の機種が存在しており、ユーザ生成コンテンツ（UGC）も、ますます一般的になってきている。ユーザは、自分自身のビデオを記録し、ビデオを編集しないまま単にアップロードもしくは送信したり、アップロードの前にビデオを修正もしくは編集したりすることができる。

10

【0003】

従来技術の画像処理ソフトウェアにより、ユーザは、写真にフィルタリングを適用して特殊効果を与えることができる。かかる効果の例には、色あせた画像や光漏れを通じて、またはフィルタを追加して特定の色を向上させることによって行われる、古いカメラのシミュレーティングが含まれる。

【0004】

従来技術の画像処理ソフトウェアに関する問題は、フィルタがコンテンツ非依存性であるということである。フィルタリングされた画像は、最初は個別化されているように見えるが、いくつかの画像に同じフィルタが適用されることによって、それらの画像が、元々見たものよりも、実際に互いにより類似して見えるようになる。

20

【0005】

また、従来技術の個別化されたビデオ生成ソフトウェアによれば、ユーザは、自分自身の顔（または別のユーザの顔）をビデオシーケンスに追加することができる。その場合、ビデオは、少なくとも表面的レベルでは、そのユーザにカスタマイズされているように見える。

【0006】

従来技術のビデオ生成ソフトウェアに関する問題は、ユーザによる入力制限されているため、結果として表現の多様性が制限されてしまうことにある。

その対極にあるのは、映画用のコンピュータ生成アニメーションにおいて用いられるものなど、マニュアルビデオおよび画像の編集および生成ソフトウェアである。かかるソフトウェアによって、ユーザは、任意のタイプのビデオまたは画像を生成することが可能になるが、それは、一般に複雑で時間がかかり、したがって一般ユーザには適していない。

30

従来技術に関する先行技術文献として以下を開示する。

【特許文献1】国際公開第2003/073322A1号パンフレット

【特許文献2】米国特許出願公開第2008/0306951号明細書

【発明の概要】

【発明が解決しようとする課題】

【0007】

本発明の一部の実施形態の目的は、上記の従来技術に対する改善および利点をユーザに提供すること、および/または従来技術の上記欠点の1つまたは複数を克服および軽減すること、および/または有用な商業上の選択肢を提供することにある。

40

【課題を解決するための手段】

【0008】

第1の態様によれば、本発明は、画像を生成する方法であって、

第1の画像を第1のコンピューティング装置で受信すること、

第1のコンピューティング装置のデータインタフェースで第1の画像をサーバに送信すること、

第1の画像の態様に対応する形状データをデータインタフェースでサーバから受信すること、

少なくとも形状データおよびアバターデータに基づいて、一次出力画像を第1のコンピ

50

ューティング装置のプロセッサによって生成すること、
を備える方法に関する。

【 0 0 0 9 】

好ましくは、形状データは、画像において描写されるユーザの顔の 3 D モデルを含む。
任意選択的に、形状データは、画像において描写される数人のユーザの顔の 3 D モデルを
含む。一次出力画像を生成することは、形状データをアバターデータに適用することを含む。

【 0 0 1 0 】

代替として、形状データは、アクティブ外観モデル (A A M : Active Appearance Mode
l) パラメータを含み、一次出力画像を生成することは、 A A M パラメータをアバターデ
ータに適用することを含む。

10

【 0 0 1 1 】

好ましくは、一次出力画像を生成することは、アバターデータに応じた形状データへの
マッピング機能を適用することをさらに含む。

好ましくは、方法は、アバターデータに関連する識別子をデータインタフェースで受信
することをさらに備える。

【 0 0 1 2 】

好ましくは、方法は、

識別子をサーバに送信することをさらに備え、

サーバは、形状パラメータと、識別子に関連するサーバアバターデータとに基づいて、
サーバ出力画像を生成する。

20

【 0 0 1 3 】

代替として、方法は、

アバターデータをサーバに送信することをさらに備え、

サーバアバターデータは、サーバに送信されたアバターデータを含む。

【 0 0 1 4 】

特定の実施形態によれば、サーバ出力画像は、一次出力画像と同一である。代替として
、サーバは、第 1 のコンピューティング装置によって一次出力画像を生成するために用い
られるアルゴリズムとは異なるアルゴリズムに従ってサーバ出力画像を生成する。

【 0 0 1 5 】

30

好ましくは、方法は、

サーバ出力画像を第 2 のコンピューティング装置のユーザに送信する要求をデータイン
タフェースでサーバに送信することであって、第 2 のコンピューティング装置のユーザの
詳細な連絡先を含む当該要求をサーバに送信すること、

ユーザの詳細な連絡先を用いて、サーバから第 2 のコンピューティング装置のユーザに
サーバ出力画像を送信すること、
をさらに備える。

【 0 0 1 6 】

好ましくは、方法は、

一次出力画像を第 2 のコンピューティング装置のユーザと共有する要求をデータインタ
フェースでサーバに送信することであって、第 2 のコンピューティング装置のユーザの詳
細な連絡先および選択されたアバターを含む当該要求をサーバに送信すること、

40

第 2 のコンピューティング装置によって一次出力画像を生成するために、ユーザの詳細
な連絡先を用いて、形状パラメータおよび選択されたアバターをサーバから第 2 のコンピ
ューティング装置のユーザに送信すること、
をさらに備える。

【 0 0 1 7 】

好ましくは、選択されたアバターを送信することは、選択されたアバターに関連する識
別子を送信することを含む。

好ましくは、一次出力画像を生成することは、第 1 の画像を修正することを含む。

50

【 0 0 1 8 】

好ましくは、方法は、

サーバ出力画像をソーシャルメディアサイトまたはビデオ共有サイトにアップロードする要求をデータインタフェースでサーバに送信することであって、ソーシャルメディアサイトまたはビデオ共有サイトの詳細および選択されたアバターを含む当該要求をサーバに送信すること、

ソーシャルメディアサイトまたはビデオ共有サイトの詳細を用いて、サーバからソーシャルメディアサイトまたはビデオ共有サイトにサーバ出力画像を送信すること、
をさらに備える。

【 0 0 1 9 】

特定の実施形態によれば、第 1 の画像は、符号化される。

好ましくは、第 1 の画像は、ビデオシーケンスの画像を含む。

好ましくは、方法は、

第 1 の画像および第 2 の画像がビデオシーケンスの画像を含むとき、第 2 の画像を第 1 のコンピューティング装置で受信すること、

第 2 の画像をデータインタフェースでサーバに送信すること、

第 2 の画像の態様に対応するさらなる形状データをデータインタフェースで受信すること、

少なくともさらなる形状データとアバターデータとに基づいて第 2 の一次出力画像をプロセッサによって生成すること、

をさらに備える。

【 0 0 2 0 】

好ましくは、方法は、

第 1 の画像および第 2 の画像がビデオシーケンスの画像を含むとき、第 2 の画像を第 1 のコンピューティング装置で受信すること、

第 2 の画像をデータインタフェースでサーバに送信すること、

第 2 の画像の態様に対応するさらなる形状データをデータインタフェースで受信すること、

少なくともさらなる形状データと第 2 のアバターデータとに基づいて第 2 の一次出力画像をプロセッサによって生成すること、

をさらに備える。

【 0 0 2 1 】

好ましくは、第 2 のアバターデータは、アバターデータを含む。

好ましくは、第 1 の画像は、第 2 の画像を受信する前にサーバに送信される。

好ましくは、方法は、

音声データを第 1 のコンピューティング装置で受信すること、

音声データをデータインタフェースでサーバに送信すること、

音声データに関係する音声パラメータをデータインタフェースで受信すること、

音声パラメータとアバターデータとに基づいた出力音声第 1 のコンピューティング装置のプロセッサによって生成すること、

をさらに備える。

【 0 0 2 2 】

好ましくは、出力音声を生成することは、音声データにフィルタを適用することを含む。

好ましくは、アバターデータは、アバター画像を含む。

【 0 0 2 3 】

第 2 の態様によれば、本発明は、画像を生成するためのシステムであって、

第 1 のコンピューティング装置を備え、該第 1 のコンピューティング装置が、

データインタフェースと、

データインタフェースに接続されたプロセッサと、

前記プロセッサに接続されたメモリと
を含み、メモリが、プロセッサによって実行可能な命令として、
第1の画像を受信するための命令、
第1の画像をデータインタフェースでサーバに送信するための命令、
第1の画像の態様に対応する形状データをデータインタフェースで受信するための命令、および
少なくとも形状データとアバターデータとに基づいて出力画像を生成するための命令を含む、システムに関する。

【0024】

好ましくは、システムは、
プロセッサに接続された表示スクリーンをさらに含み、
メモリは、出力画像を表示スクリーン上に表示するための命令をさらに含む。

10

【0025】

好ましくは、システムは、
プロセッサに接続されたカメラをさらに含み、
メモリは、第1の画像をカメラから受信するための命令をさらに含む。

【0026】

第3の態様によれば、本発明は、画像を生成および送信するためのシステムであって、
パーソナルコンピューティング装置を備え、該パーソナルコンピューティング装置が、
データインタフェースと、
データインタフェースに接続されたプロセッサと、
プロセッサに接続されたメモリと
を含み、メモリが、プロセッサによって実行可能な命令として、
画像を受信するための命令、
画像をデータインタフェースでサーバに送信するための命令、
画像の態様に対応する形状データをデータインタフェースで受信するための命令、および
少なくとも形状データとアバターデータとに基づいて出力画像を生成するための命令を含む、サーバが、

20

サーバデータインタフェースと、
サーバデータインタフェースに接続されたサーバプロセッサと、
サーバプロセッサに接続されたサーバメモリと
を含み、サーバメモリが、サーバプロセッサによって実行可能な命令として、
画像を受信するための命令、
画像の態様に応じた形状データをサーバプロセッサによって生成するための命令、
形状データをサーバデータインタフェースで第1のコンピューティング装置に送信するための命令、

30

少なくとも形状データとアバターデータとに基づいてサーバ出力画像をサーバプロセッサで生成するための命令、および

サーバ出力画像をサーバデータインタフェースで第2のコンピューティング装置に送信するための命令、
を含む、システムに関する。

40

【0027】

好ましくは、システムは、複数のパーソナルコンピューティング装置をさらに含み、サーバ出力画像は、複数のパーソナルコンピューティング装置からの画像に係る形状データに基づいて生成される。

【図面の簡単な説明】

【0028】

【図1】本発明の実施形態による画像生成システムを示す図である。

【図2】本発明の実施形態による図1のシステムにおける第1のコンピューティング装置

50

と対話するユーザを示す図である。

【図3】図3(a)は本発明の実施形態による図1のシステムにおける形状データの正面図、図3(b)は形状データの側面図である。

【図4】本発明の実施形態による図1のシステムにおけるアバター選択スクリーンのスクリーンショットを示す図である。

【図5】本発明の実施形態による画像生成システムを示す図である。

【図6】本発明の実施形態による第1のコンピューティング装置とサーバとの間の通信を示すメッセージフロー図である。

【図7】本発明の実施形態による画像生成方法を示すフロー図である。

【図8】本発明の実施形態によるコンピューティング装置を示す図である。

10

【発明を実施するための形態】

【0029】

本発明の理解を助けるために、および当業者が本発明を実施できるようにするために、本発明の好ましい実施形態の一例を、添付の図面を参照して以下に説明する。当業者は、図面に示されるような構成要素のレイアウトからの小さな逸脱が、本発明の開示した実施形態の適切な機能を損なうものでないことを理解し得る。

【0030】

本発明の実施形態は、画像およびビデオ生成システムならびに方法を包含する。本発明の要素は、簡潔な概略形態で図面に示されており、これらの図面は、以下の説明により当業者に明らかとなる過度の詳細によって本開示を混乱させることがないように、本発明の実施形態の理解に必要な特定の詳細のみを示す。

20

【0031】

本明細書において、第1および第2、左および右、前および後ろ、上端および底部などの形容詞は、その形容詞によって表現される特定の相対的位置または順序を必ずしも要求するものでなく、一つの要素または方法ステップを、別の要素または方法ステップと区別して定義するためにのみ用いる。「包含する」または「含む」などの用語は、要素または方法ステップの排他的な集まりを定義するために用いていない。むしろ、このような用語は単に、本発明のある特定の実施形態に含まれる要素または方法ステップの最小限の集まりを定義する。

【0032】

30

本明細書における先行技術に対する参照は、先行技術が共通の一般的知識の一部を形成していることを認知または示唆するものでなく、また、そのように解釈されるべきではない。

【0033】

一態様によれば、本発明は、画像を生成する方法であって、第1の画像を第1のコンピューティング装置で受信すること、第1の画像を第1のコンピューティング装置のデータインタフェースでサーバに送信すること、第1の画像の態様に対応する形状データをサーバからデータインタフェースで受信すること、少なくとも形状データとアバターデータとに基づいて一次出力画像を第1のコンピューティング装置のプロセッサによって生成することを備える方法に関する。

40

【0034】

本発明のある実施形態における利点は、限られた処理能力の装置上で、ユーザがアバターをアニメ化する能力を含む。アニメーションは低遅延で生じさせることができ、アバターは動的に変更することができる。そして、得られたアニメ化されたアバターの画像またはビデオは、優れた帯域効率で、別のユーザ、ソーシャルメディア、またはビデオ共有サイトに送信することができる。

【0035】

以下で説明する実施形態は、画像の生成に言及する。当業者によって容易に理解されるように、画像には、ビデオシーケンスの画像を含み得る。さらに、実施形態は、必ずしも各画像を単一の独立した画像として扱わずに、ビデオ生成を支援するように容易に拡張す

50

ることができる。例えば、タイミング情報をビデオシーケンスの各画像に関連付けることができ、これにより、タイミング情報を任意の出力ビデオシーケンスに適用することができる。

【0036】

なお、アバターデータという用語は、動作もしくは表情を転送できる任意のタイプの画像データおよび/または音声データを示すために用いられる。アバターは、漫画のキャラクターなどの人工のキャラクターに基づくか、または実在の人物の画像を含むことができる。さらに、アバターは、動物、もしくはエイリアンなどの空想生物など、人間以外のキャラクターに基づくか、または顔もしくは他の身体部分を上に置いた無生物もしくは空想上の産物を含むことができる。

10

【0037】

図1は、本発明の実施形態による、画像を生成するためのシステム100を示す。

システム100は、第1のコンピューティング装置105およびサーバ110を含む。第1のコンピューティング装置105およびサーバ110は、データ通信ネットワーク115を介して結合される。

【0038】

第1のコンピューティング装置105は、カメラ120、表示スクリーン125、およびマイクロホン130を含み、それら全てがプロセッサ(図示せず)に接続され、プロセッサはデータインタフェース(図示せず)に接続される。以下で説明するように、第1のコンピューティング装置105はさらに、プロセッサに接続され、本発明の方法を実行するためにプロセッサによって実行可能な命令を含むメモリ(図示せず)を含む。

20

【0039】

第1のコンピューティング装置105は、携帯電話、携帯情報端末(PDA)、タブレットコンピュータ、または他の類似装置などの下位装置とすることができる。サーバは、例えば専用コンピューティング装置とすることができ、幾つかの第1のコンピューティング装置105のために動作する場合には有利となる。

【0040】

第1のコンピューティング装置105は、画像またはビデオをサーバ110に送信し、サーバ110は、画像またはビデオを処理して、形状パラメータを第1のコンピューティング装置105に返信する。次に、第1のコンピューティング装置105は、形状パラメータに従ってアバターモデルをアニメ化する。

30

【0041】

アバターモデルのアニメーションは、携帯電話などの下位装置のグラフィック処理ユニット(GPU)に特に適切化される。アニメーションは、ユーザの表情をアバターの表情に変換すること、または身体形状または動作を変換することを含み得る。これは、例えば、ユーザの表情を含むようにアバターをアニメ化すること、特定の方法でアバターを飛び跳ねさせたり動作させたりすること、または手話などの特定の動作を行わせることを含み得る。

【0042】

アバターモデルは、第1のコンピューティング装置105か、サーバ110に配置することができる。あるいは、アバターモデルは、サーバ110および第1のコンピューティング装置105にそれぞれ部分的に配置することができる。また、アバターモデルは、第1のコンピューティング装置105とサーバ110との間で送受信することができる。

40

【0043】

図2は、本発明の実施形態に従って、第1のコンピューティング装置105と対話するユーザ205を示す。

ユーザ205の複数の画像が、カメラ120によって捕捉される。複数の画像のうちの第1の画像が捕捉されると、その第1の画像がサーバ110に送信されて直ちに処理される。すなわち、第2の画像が捕捉されるのを待つことなく、またはビデオが完全に記録されるのを待つことなく、第1の画像がサーバ110に送信されて直ちに処理される。

50

【 0 0 4 4 】

ネットワークが遅いかまたは利用不可の場合に、第 1 のコンピューティング装置 1 0 5 によって捕捉された画像を保存して、ネットワークが利用可能となるかまたは十分に速くなった場合にサーバ 1 1 0 に送信するようにしてもよい。

【 0 0 4 5 】

本発明の代替実施形態（図示せず）によれば、画像は、第 1 のコンピューティング装置 1 0 5 に送信される、すなわちダウンロードされたビデオシーケンスからなる。当業者によって容易に理解されるように、画像は、第 1 のコンピューティング装置 1 0 5 によって捕捉される必要はない。

【 0 0 4 6 】

サーバ 1 1 0 は、第 1 の画像の態様に対応する形状データを生成する。画像態様としては、ユーザの顔、身体形状、または任意の他の適切な態様とすることができる。ユーザの顔の場合、形状データは、画像に基づくユーザの顔の 3 D モデルを含むものであってもよい。あるいは、形状データは、アクティブ外観モデル（Active Appearance Model: A A M）のパラメータを含むものであってもよい。さらに、形状パラメータは、相対的なカメラ位置を含むものであってもよい。

【 0 0 4 7 】

サーバ 1 1 0 は、形状データを第 1 のコンピューティング装置 1 0 5 に送信し、第 1 のコンピューティング装置 1 0 5 は、少なくとも形状データおよびアバターデータに基づいて出力画像を生成する。

【 0 0 4 8 】

アバターデータには、例えば、形状データに従って変換されたアバター画像を含むことができる。

出力画像は、3 D モデル、A A M パラメータ、または他の形状データをアバターデータに適用して生成することを含む。また、出力画像は、アバターデータに応じた形状データへのマッピング機能を適用して生成することを含み得る。また、出力画像は、2 0 1 2 年 3 月 2 1 日出願の「METHOD AND SYSTEM FOR FACIAL EXPRESSION TRANSFER」（表情転送方法およびシステム）なる名称の国際出願 P C T / A U 2 0 1 2 / 0 0 0 2 9 5 号明細書に説明されているような複数のソースアバターマッピング機能に従って生成することができる。サーバ 1 1 0 は、画像、アバターデータ、およびトレーニングデータに基づいて、ソースアバターマッピング機能を生成することができる。ソースアバターマッピング機能は、第 1 のコンピューティング装置 1 0 5 に送信されて使用される。

【 0 0 4 9 】

図 3（a）は、本発明の実施形態による形状データ 3 0 0 の正面図を示し、図 3（b）は、形状データ 3 0 0 の側面図を示す。

形状データは、ユーザの顔の特徴に対応する複数の 3 D 点 3 0 5 を含む。なお、特徴には、目、口、眉、あごの形状、または任意の他の特徴を含むことができる。

【 0 0 5 0 】

図 4 は、本発明の実施形態によるアバター選択スクリーンのスクリーンショット 4 0 0 を示す。第 1 のコンピューティング装置 1 0 5 およびサーバ 1 1 0 は、複数のアバターに対応するアバターデータを含むことができ、アバター選択スクリーンは、複数のアバターからアバターを選択するために用いることができる。

【 0 0 5 1 】

アバター選択スクリーンは、アバター出力画像 4 0 5 および複数のアバター選択ボタン 4 1 0 を含む。アバター選択ボタン 4 1 0 を用いてアバターを選択すると、出力画像 4 0 5 が、選択されたアバターに従って更新される。

【 0 0 5 2 】

アバター選択スクリーンは、さらに、再生ボタン 4 1 5、早送りボタン 4 2 0、および巻き戻しボタン 4 2 5 を介した媒体制御機能を含む。再生ボタン 4 1 5、早送りボタン 4

10

20

30

40

50

20、および巻き戻しボタン425は、ビデオシーケンスの再生を制御するために用いることができる。本発明の特定の実施形態によれば、アバター選択ボタン410を用いて、再生中にアバターを変更することができる。

【0053】

この場合、入力ビデオシーケンスは、第1の画像および第2の画像を含む複数の画像を含むことができ、これらの画像は、第1のコンピューティング装置105によって受信される。

【0054】

上記のように、第1の画像は、第1のコンピューティング装置105によってサーバ110に送信される。サーバ110は、第1の画像の態様に対応する形状データを第1のコンピューティング装置105に送信する。次に、第1のコンピューティング装置105は、第1の選択されたアバターの形状データとアバターデータとに基づいて、一次出力画像を生成する。

【0055】

続いて、第2の画像が、第1のコンピューティング装置105によってサーバ110に送信される。サーバ110は、第2の画像の態様に対応するさらなる形状データを第1のコンピューティング装置105に送信する。次に、第1のコンピューティング装置105は、そのさらなる形状データと、新しいアバターが選択された場合には第2の選択されたアバターデータまたは第1の選択されたアバターデータとに基づいて、さらなる一次出力画像を生成する。

【0056】

したがって、一次出力画像およびさらなる一次出力画像によって出力ビデオシーケンスが形成される。出力ビデオシーケンスは、選択されたアバター（単数か複数）を含み、例えばアバター選択スクリーン上で再生することができる。

【0057】

図5は、本発明の実施形態による、画像を生成するためのシステム500を示す。

システム500は、第1のコンピューティング装置505、サーバ510、および第2のコンピューティング装置515を含む。第1のコンピューティング装置505およびサーバ510は、図1の第1のコンピューティング装置105およびサーバ110に類似している。

【0058】

第1のコンピューティング装置505は、上記のように、例えば第1のコンピューティング装置505のカメラ120から画像を受信し、それをサーバ510に送信する。次に、サーバ510は形状パラメータを生成し、その形状パラメータを第1のコンピューティング装置505に送信する。

【0059】

これにより、図4を参照して上述したように、第1のコンピューティング装置505のユーザは、異なるアバターのデータに従って出力画像を生成し、アバターを選択することができる。

【0060】

ユーザは、アバターを選択した後で、出力画像を別のユーザに転送することを選択してもよい。この場合、第2のユーザの詳細な連絡先および/または第2のユーザ装置515の詳細なアドレスとともに、選択されたアバターを含むメッセージが、第1のコンピューティング装置505からサーバ510に送信される。

【0061】

各アバターは、固有の識別子に関連付けることができる。この場合、第1のコンピューティング装置505は、固有の識別子をサーバ510に送信するだけでよい。

特定の実施形態によれば、サーバ510は、識別子に関連するサーバアバターデータと形状パラメータとに基づいてサーバ出力画像を生成する。

【0062】

サーバアバターデータは、第１のコンピューティング装置５０５のアバターデータと同一であってもよいし、またはサーバ５１０に利用可能な任意の追加の処理パワーに対応する追加の詳細を含んでもよい。同様に、サーバ５１０は、より正確なまたはより現実的なサーバ出力画像を生成するために、第１のコンピューティング装置５０５によって用いられるアルゴリズムとは異なるアルゴリズムを適用して一次出力画像を生成してもよい。

【００６３】

次に、サーバ５１０は、サーバ出力画像を第２のコンピューティング装置５１５に送信する。

他の実施形態によれば、サーバ５１０は、選択されたアバターの詳細、例えば選択されたアバターを識別する識別子を、形状パラメータとともに第２のコンピューティング装置５１５に送信する。次に、第２のコンピューティング装置５１５は、形状パラメータおよび選択されたアバターに従って出力画像を描画することができる。

【００６４】

さらに別の実施形態によれば、サーバ５１０は、サーバ出力画像を、ソーシャルネットワークングサイトに、または、ファイル、画像もしくはビデオ共有サイトに、あるいは、別のサイトにアップロードする。

【００６５】

出力画像の生成は、独立した出力画像を生成することを含み得る。または、出力画像の生成は、第１の画像を修正することを含み得る。なお、第１の画像を符号化してもよく、また、後続の画像を第１の画像に対して差分符号化してもよい。

【００６６】

代替実施形態（図示せず）によれば、サーバ出力画像および／または一次出力画像は、いくつかの画像からの形状データに基づいて生成される。したがって、出力画像は、数人のユーザの表情など、いくつかの画像からの態様を組み合わせることができる。同様に、いくつかの画像からの形状データは、数人のユーザの表情など、数人のユーザに関連するデータを含むことができる。

【００６７】

図６は、本発明の実施形態による、第１のコンピューティング装置１０５、５０５と、サーバ１１０、５１０との間の通信のメッセージシーケンス６００を示す。

ステップ６０５において、第１の画像が、第１のコンピューティング装置１０５、５０５で受信される。第１の画像は、カメラまたは任意の他のソースからのものとすることができる。

【００６８】

ステップ６１０において、第１の画像は、サーバ１１０、５１０に送信されて処理される。

ステップ６１５において、第１の画像が処理されるのを待たずに、第２の画像がサーバ１１０、５１０に送信されて処理される。これによって、累積遅延でもビデオシーケンスの長さに対応する遅延でもなく単一の往復時間遅延が追加されるだけなので、ビデオのほぼリアルタイム処理が可能になる。

【００６９】

ステップ６２０において、第１の画像がサーバ１１０、５１０によって処理され、ステップ６２５において、形状パラメータがサーバ１１０、５１０から第１のコンピューティング装置１０５、５０５に送信される。次に、ステップ６３０において、第１のコンピューティング装置１０５、５０５は出力画像を生成し表示する。

【００７０】

ステップ６３５において、第２の画像がサーバ１１０、５１０によって処理される。

図７は、本発明の実施形態による、画像を生成する方法７００を示す。

ステップ７０５において、第１の画像が第１のコンピューティング装置によって受信される。第１の画像は、カメラから直接受信されるか、または別の手段によって受信することができる。

【 0 0 7 1 】

ステップ 7 1 0 において、第 1 の画像がサーバに送信される。

ステップ 7 1 5 において、第 2 の画像が第 1 のコンピューティング装置によって受信される。ステップ 7 2 0 において、第 2 の画像がサーバに送信される。

【 0 0 7 2 】

ステップ 7 2 5 において、第 1 の画像に関する形状データがサーバから送信され、第 1 のコンピューティング装置によって受信される。ステップ 7 3 0 において、第 1 のコンピューティング装置は、形状データとアバタデータとに基づいて第 1 の一次出力画像を生成する。

【 0 0 7 3 】

ステップ 7 3 5 において、第 2 の画像に関する形状データがサーバから送信され、第 1 のコンピューティング装置によって受信される。ステップ 7 4 0 において、第 1 のコンピューティング装置は、第 2 の画像の形状データとアバタデータとに基づいて第 2 の一次出力画像を生成する。

【 0 0 7 4 】

ステップ 7 2 5 および 7 3 0 は、ステップ 7 1 5 および 7 2 0 の前に実行することができる。しかしながら、システムの全体的な遅延を低減するためには、第 2 の画像が第 1 のコンピューティング装置によって受信されたときに第 2 の画像をサーバに送信して処理するのが有利である。

【 0 0 7 5 】

本発明の特定の実施形態によれば、方法は、さらに、スピーチなどの音声データを受信および処理することを含む。音声データは、有利にはマイクロホン 1 3 0 によって捕捉される。この音声データは、音声データを画像またはビデオデータと正確に同期させるためにタイミング情報を含んでもよい。

【 0 0 7 6 】

音声データは、上記の画像と同様の方法で、データインタフェース上でサーバに送信されて処理される。次に、サーバ 1 1 0 , 5 1 0 は、音声データを処理して、音声データに関する音声パラメータを生成する。音声パラメータは、第 1 のコンピューティング装置 1 0 5 , 5 0 5 に送信され、出力音声は、音声パラメータおよびアバタデータに基づいて生成される。

【 0 0 7 7 】

音声パラメータは、音声データをフィルタリングするために用いられるフィルタパラメータを含むことができる。

図 8 は、本発明の実施形態によるコンピューティング装置 8 0 0 を示す。第 1 および第 2 のコンピューティング装置 1 0 5 , 5 0 5 , 5 1 5 , およびサーバ 1 1 0 , 5 1 0 は、コンピューティング装置 8 0 0 と同じか、または類似したものとすることができる。同様に、図 7 の方法 7 0 0 は、コンピューティング装置 8 0 0 を用いて実行することができる。

【 0 0 7 8 】

コンピューティング装置 8 0 0 は、中央処理装置 8 0 2 と、システムメモリ 8 0 4 と、システムバス 8 0 6 とを含む。システムバス 8 0 6 は、システムメモリ 8 0 4 と中央処理装置 8 0 2 との接続等、様々なシステム要素の接続を行う。システムバス 8 0 6 は、メモリバスまたはメモリコントローラ、周辺機器用バス、および様々なバスアーキテクチャのいずれかを有したローカルバスを含む種々のバス構造のいずれかであってよい。システムメモリ 8 0 4 の構造は、当業者に周知であり、読み取り専用メモリ (R O M) に記憶された基本入出力システム (B I O S) とともに、ランダムアクセスメモリ (R A M) に記憶されたオペレーティングシステム、アプリケーションプログラム、およびプログラムデータなどの 1 つまたは複数のプログラムモジュールを含み得る。

【 0 0 7 9 】

コンピューティング装置 8 0 0 はまた、データの読み出しおよび書き込みを行うための

10

20

30

40

50

様々なインタフェースユニットおよびドライブも含み得る。データは、例えば、カメラと表示スクリーンとの間の変位を含むことができる。

【 0 0 8 0 】

具体的には、コンピューティング装置 8 0 0 は、ハードディスクドライブ 8 1 2 およびリムーバブルメモリドライブ 8 1 4 をシステムバス 8 0 6 にそれぞれ接続するハードディスクインタフェース 8 0 8 およびリムーバブルメモリインタフェース 8 1 0 を含む。リムーバブルメモリドライブ 8 1 4 の例には、磁気ディスクドライブおよび光ディスクドライブが含まれる。ドライブおよびそれらに関連するデジタル多用途ディスク (D V D) 8 1 6 等のコンピュータ可読媒体は、コンピュータ可読命令、データ構造、プログラムモジュール、およびコンピュータシステム 8 0 0 用の他のデータの揮発性記憶装置を提供する。単に例示を目的とし、コンピューティング装置 8 0 0 が幾つかの同様のドライブを含み得るという理解の下に、1つのハードディスクドライブ 8 1 2 および1つのリムーバブルメモリドライブ 8 1 4 を示している。さらに、コンピューティング装置 8 0 0 は、他の種類のコンピュータ可読媒体とインタフェースをとるためのドライブを含み得る。

【 0 0 8 1 】

コンピューティング装置 8 0 0 は、デバイスをシステムバス 8 0 6 に接続するための追加インタフェースを含んでもよい。図 8 は、装置をシステムバス 8 0 6 に接続するために使用可能なユニバーサル・シリアル・バス (U S B) インタフェース 8 1 8 を示している。例えば、IEEE 1394 インタフェース 8 2 0 を使用して、追加のデバイスをコンピューティング装置 8 0 0 に接続することが可能である。追加のデバイスの例には、画像もしくはビデオを受像するためのカメラ、または音声を記録するためのマイクロホンが含まれる。

【 0 0 8 2 】

コンピューティング装置 8 0 0 は、1つまたは複数のリモートコンピュータ、あるいは、サーバ、ルータ、ネットワークパーソナルコンピュータ、ピアデバイスまたは他の共通のネットワークノード、無線電話、または無線携帯情報端末等の他のデバイスへの論理接続を用いたネットワーク環境において動作可能である。コンピューティング装置 8 0 0 は、システムバス 8 0 6 をローカル・エリア・ネットワーク (L A N) 8 2 4 に接続するネットワークインタフェース 8 2 2 を含む。ネットワーク環境は、オフィス、企業規模のコンピュータネットワークおよびホームコンピュータシステムにおいて一般的なものである。

【 0 0 8 3 】

インターネット等の広域ネットワーク (W A N) は、例えばシリアルポートインタフェース 8 2 6 に接続されたモデムユニットを介して、または L A N 8 2 4 を介して、コンピューティング装置によってアクセスすることもできる。

【 0 0 8 4 】

画像および/またはビデオの送信は、L A N 8 2 4、W A N、またはそれらの組み合わせを用いて実行することができる。

図示し説明したネットワーク接続は例示的なものであり、コンピュータ間の通信リンクを確立する他の方法を使用可能であることが理解される。TCP / IP、フレームリレー、イーサネット (登録商標)、FTP、HTTPなどの様々な周知なプロトコルのいずれかの存在が仮定され、コンピューティング装置 8 0 0 は、ユーザが、例えばウェブベースのサーバからデータを検索できるようにするクライアント・サーバ構成で動作することができる。

【 0 0 8 5 】

コンピューティング装置 8 0 0 の動作は、様々な異なるプログラムモジュールによって制御することができる。プログラムモジュールの例は、特定のタスクを実行するか、または特定の抽象データ型を実行するルーチン、プログラム、オブジェクト、コンポーネント、およびデータ構造である。本発明はまた、携帯端末、マルチプロセッサシステム、マイクロプロセッサベースまたはプログラム可能家庭用電化製品、ネットワーク PC、ミニコ

ンピュータ、メインフレームコンピュータ、携帯情報端末などを含む他のコンピュータシステム構成で実施してもよい。さらに、本発明はまた、通信ネットワークを介して接続されるリモート処理装置によってタスクが実行される分散コンピューティング環境において実施してもよい。分散コンピューティング環境において、プログラムモジュールは、ローカルおよびリモートメモリ記憶装置の両方に配置することができる。

【0086】

要約すれば、本発明の特定の実施形態における利点は、限られた処理能力の装置でユーザがアバターをアニメーション化することができることを含む。アニメーションを低遅延で生じさせることができ、かつアバターを動的に変更することができる。そして、得られたアニメ化されたアバターの画像またはビデオを優れた帯域効率で別のユーザに送信することができる。

10

【0087】

本発明の様々な実施形態における上記の説明は、当業者への説明のために提供される。本発明を網羅的に記載すること、または本発明を単一の開示された実施形態に限定することを意図したものではない。上記教示を受けた当業者には、本発明の多数の代替形態および変更形態が明らかとなる。従って、一部の代替実施形態を具体的に説明したが、他の実施形態は、当業者によって明らかとなる、または比較的容易に開発され得る。従って、本明細書は、上記した本発明の全ての代替形態、変形形態、および変更形態、および本発明の思想および範囲に包含される他の実施形態を含む。

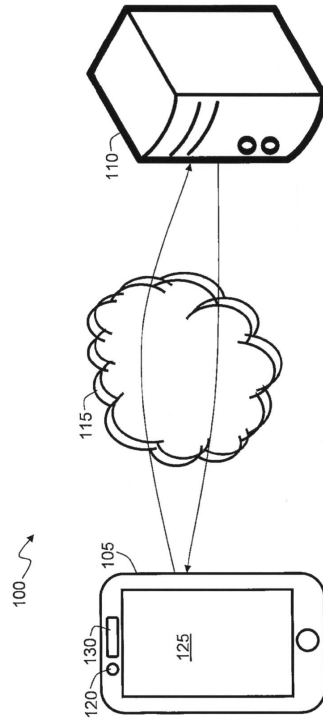
【符号の説明】

20

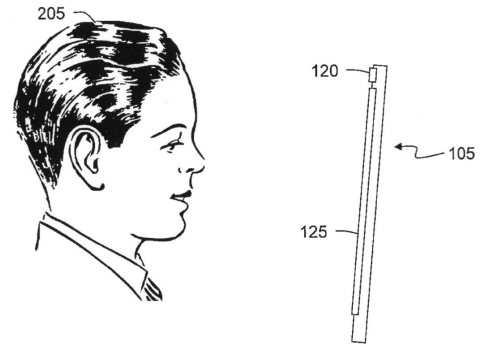
【0088】

100, 500: システム、105, 505: 第1のコンピューティング装置、110, 510: サーバ、115: データ通信ネットワーク、120: カメラ、125: 表示スクリーン、300: 形状データ、305: 3D点、400: スクリーンショット、405: アバター出力画像、410: アバター選択ボタン、515: 第2のコンピューティング装置、800: コンピューティング装置。

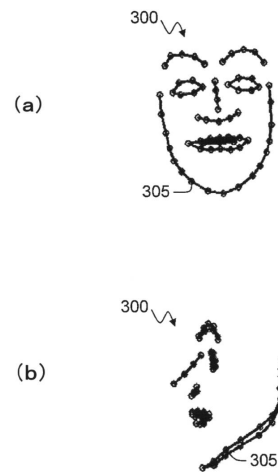
【図 1】



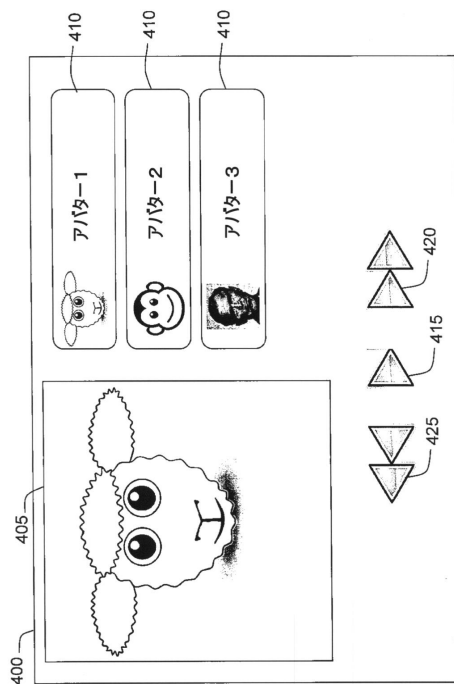
【図 2】



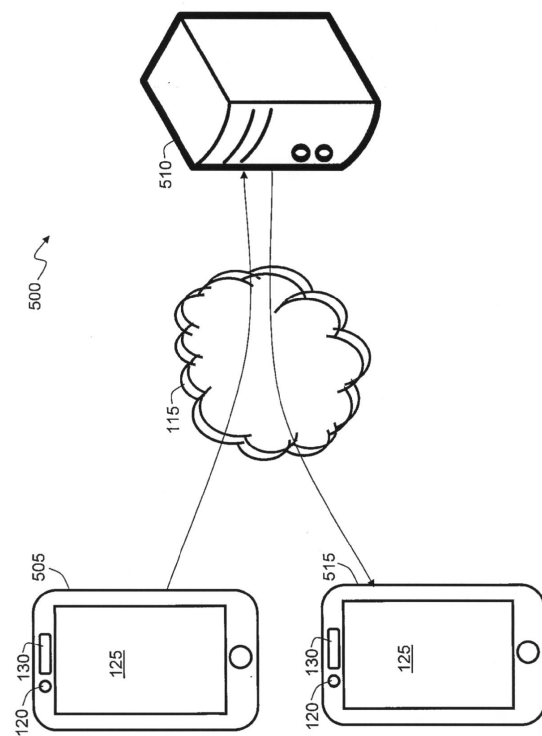
【図 3】



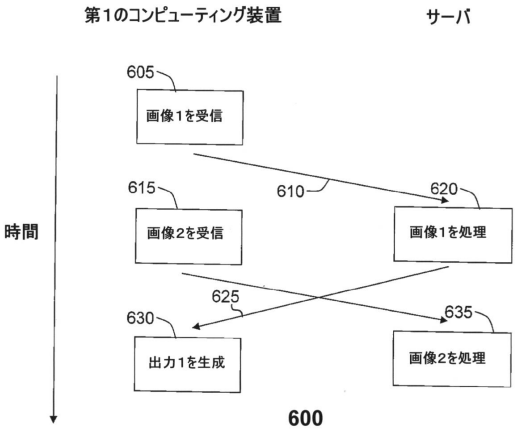
【図 4】



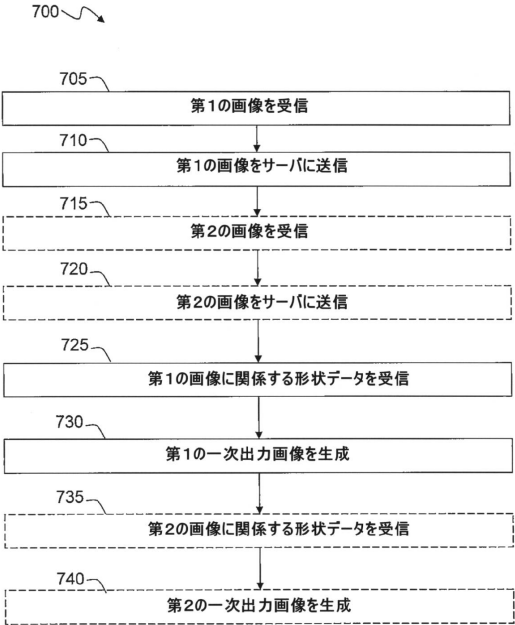
【図 5】



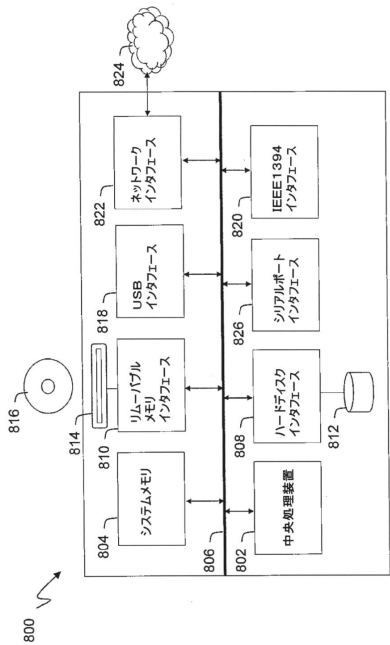
【図 6】



【図 7】



【図 8】



 フロントページの続き

- (72)発明者 サイモン ルーシー
 オーストラリア国 2612 オーストラリアン キャピタル テリトリー キャンベル ライム
 ストーン アベニュー
- (72)発明者 ゴータム テンドルカール
 オーストラリア国 2612 オーストラリアン キャピタル テリトリー キャンベル ライム
 ストーン アベニュー

審査官 古川 哲也

- (56)参考文献 特開2006-040228(JP,A)
 特表2010-517427(JP,A)
 特開2003-016475(JP,A)
 特開2004-248145(JP,A)
 特開2005-277989(JP,A)
 特開2002-288686(JP,A)
 国際公開第03/073322(WO,A1)
 WEBER, Mario et al., Mobile Entertainment Using Personal Avatars, Proceedings of the 8
 th International Conference on Telecommunications 2005 - ConTEL 2005, 米国, IEEE, 20
 05年 6月15日, p.341-346, ISBN: 953-184-081-4

(58)調査した分野(Int.Cl., DB名)

H04N	21/00	-	21/858
H04N	7/14	-	7/15
G06F	13/00		
G06T	1/00		
G06T	15/00		
IEEE	Xplore		