

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2021-111102  
(P2021-111102A)

(43) 公開日 令和3年8月2日(2021.8.2)

(51) Int. Cl.	F I	テーマコード (参考)
<b>GO6T 13/40 (2011.01)</b>	GO6T 13/40	5B050
<b>HO4N 21/235 (2011.01)</b>	HO4N 21/235	5B084
<b>HO4N 21/431 (2011.01)</b>	HO4N 21/431	5C164
<b>HO4N 21/234 (2011.01)</b>	HO4N 21/234	
<b>HO4N 21/854 (2011.01)</b>	HO4N 21/854	
審査請求 未請求 請求項の数 9 O L (全 20 頁) 最終頁に続く		

(21) 出願番号 特願2020-2259 (P2020-2259)  
(22) 出願日 令和2年1月9日(2020.1.9)

特許法第30条第2項適用申請有り 令和1年9月25日および令和1年9月26日に、Unite Tokyo o 2019にて発表。

(71) 出願人 517054833  
株式会社 Z I Z A I  
東京都渋谷区南平台町16-28 Daiwa渋谷スクエア3F  
(74) 代理人 110002273  
特許業務法人インターブレイン  
(72) 発明者 塚本 大地  
東京都渋谷区南平台町16-28 Daiwa渋谷スクエア3F 株式会社 Z I Z A I 内  
Fターム(参考) 5B050 AA10 BA08 BA12 CA07 CA08 EA24 FA02 FA05 FA13

最終頁に続く

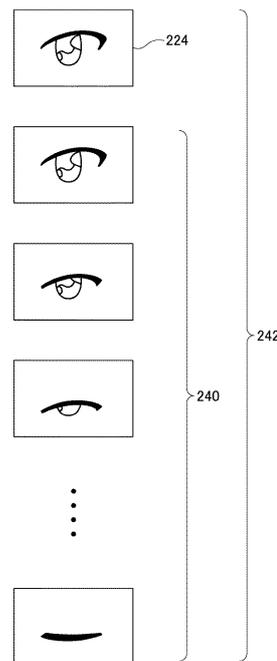
(54) 【発明の名称】 動画生成装置およびライブ通信システム

(57) 【要約】

【課題】 静止画像としての顔画像から動画像を生成する。

【解決手段】 配信者により操作される配信装置、視聴者により操作される視聴装置およびサーバを含むライブ通信システムを想定する。サーバは、ユーザにより造形されたキャラクターの静止画像としての顔画像を取得し、顔画像に含まれる特定の部分領域を変化させた画像として、複数の派生画像を自動生成し、複数の派生画像を前記視聴装置に送信する。配信装置は、配信者の撮像画像から、配信者の動作の特徴を示す動作情報を生成するとともに配信者から第1メッセージを受け付け、動作情報および第1メッセージを送信する。視聴装置は、画面にキャラクターを表示させ、動作情報に応じて複数の派生画像から表示対象となる派生画像を選択することにより、キャラクターを動作させるとともにキャラクターの動作に合わせて前記第1メッセージを出力する。

【選択図】 図12



**【特許請求の範囲】****【請求項 1】**

静止画像としての顔画像を取得する画像取得部と、  
顔画像に含まれる特定の部分領域を変化させた画像として、複数の派生画像を自動生成する派生画像生成部と、

前記複数の派生画像に基づいて、顔画像を動画表示させる画像表示部と、を備えることを特徴とする動画生成装置。

**【請求項 2】**

前記派生画像生成部は、単一の静止画像としての顔画像から複数の派生画像を自動生成することを特徴とする請求項 1 に記載の動画生成装置。

10

**【請求項 3】**

顔画像から前記部分領域を抽出する部分抽出部、を更に備え、

前記部分抽出部は、あらかじめ用意された複数の顔画像それぞれにおいて指定される部分領域を教師データとする機械学習モデルにより、前記取得された顔画像における部分領域を抽出することを特徴とする請求項 1 または 2 に記載の動画生成装置。

**【請求項 4】**

前記静止画像としての顔画像は、実在人物を撮像した顔画像ではなく、ユーザにより造形されたキャラクターの静止画像としての顔画像であることを特徴とする請求項 1 から 3 のいずれかに記載の動画生成装置。

**【請求項 5】**

20

前記部分領域は、顔画像のうち開眼状態の目に対応する領域であり、

前記派生画像生成部は、顔画像に含まれる開眼状態の目領域が閉眼状態に移行する過程を示す複数の派生画像を生成し、

前記画像表示部は、複数の派生画像を連続表示させることにより、瞬きを示す動画を表示させることを特徴とする請求項 1 から 4 のいずれかに記載の動画生成装置。

**【請求項 6】**

前記派生画像生成部は、目領域の上眼瞼を前記目領域内において複数段階を経て下降させるとともに前記上眼瞼に対応する部分を所定色によって塗色することにより複数の派生画像を生成することを特徴とする請求項 5 に記載の動画生成装置。

**【請求項 7】**

30

前記部分領域は、顔画像のうち口に対応する領域であり、

前記派生画像生成部は、顔画像に含まれる開口状態の口領域が閉口状態に移行する過程を示す複数の派生画像、または、顔画像に含まれる閉口状態の口領域が開口状態に移行する過程を示す複数の派生画像を生成し、

前記画像表示部は、複数の派生画像を連続表示させることにより、口の開閉を示す動画を表示させることを特徴とする請求項 1 から 6 のいずれかに記載の動画生成装置。

**【請求項 8】**

配信者により操作される配信装置、視聴者により操作される視聴装置およびサーバを含み、

前記サーバは、

40

ユーザにより造形されたキャラクターの静止画像としての顔画像を取得する顔画像取得部と、

前記顔画像に含まれる特定の部分領域を変化させた画像として、複数の派生画像を自動生成する派生画像生成部と、

前記複数の派生画像を前記視聴装置に送信する画像送信部と、を備え、

前記配信装置は、

配信者を撮像する撮像部と、

配信者の撮像画像から、配信者の動作の特徴を示す動作情報を生成する動作情報生成部と、

配信者から第 1 メッセージを受け付ける第 1 メッセージ取得部と、

50

前記動作情報および前記第1メッセージを送信する送信部と、を備え、  
前記視聴装置は、  
前記動作情報および前記第1メッセージを受信する受信部と、  
前記視聴装置の画面に前記キャラクタを表示させ、前記動作情報に応じて前記複数の派生画像から表示対象となる派生画像を選択することにより、前記キャラクタを動作させる視聴キャラクタ表示部と、  
前記キャラクタの動作に合わせて前記第1メッセージを出力する第1メッセージ出力部と、を備えることを特徴とするライブ通信システム。

【請求項9】

静止画像として顔画像を取得する機能と、  
顔画像に含まれる特定の部分領域を変化させた画像として、複数の派生画像を自動生成する機能と、  
前記複数の派生画像に基づいて、前記顔画像を動画表示させる機能と、をコンピュータに発揮させることを特徴とする動画生成プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、動画像を生成する技術、に関する。

【背景技術】

【0002】

スマートフォンの普及にともない、SNS（Social Networking Service）を介したコミュニケーションが一般化している。SNSでは、テキストだけでなく写真や動画像などさまざまなコンテンツを気軽に共有できる。SNSは、人と人との「絆（つながり）」を強く感じさせるツールである。

【0003】

一方、コンピュータゲームの分野においては、ユーザがキャラクタ（コンピュータ）と会話をしながらプレイするゲームがある。ユーザは、現実には存在しないキャラクタに対して強い思い入れをもつことが多い。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2012-231281号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

キャラクタは、ゲームデザイナーによってつくられた仮想的存在であり、人間のような実在感をもたない。その一方、キャラクタは、非実在性に起因する独特の魅力をもつ。本発明者らは、キャラクタ特有の魅力と人間同士特有の「リアルなつながり」の双方を取り込むことにより、新しいコミュニケーション方法を提案できるのではないかと考えた。

【0006】

具体的には、配信者（人間）と視聴者（人間）の会話を想定する。配信者は、ちょうどテレビアニメの声優のようにキャラクタの代弁者となり、キャラクタを通して視聴者と会話する。このようなコミュニケーションを想定した場合、視聴者はキャラクタと言葉以上の関わりをもつことができれば、視聴者のキャラクタに対する思い入れをいっそう強めることができると考えられる。

【0007】

配信者の化身であるキャラクタは、表情を豊かに変化させることが望ましい。また、多数の配信者が多数のキャラクタを通してサービスを提供することでサービス水準も向上していくと考えられる。本発明者らはこのような背景から、生き生きとしたキャラクタを生み出しやすい環境、いいかえれば、キャラクタの動画像を効率的に生成する仕組みが必要

10

20

30

40

50

であると考えた。

【0008】

本発明は、上記認識に基づいて完成された発明であり、その主たる目的は、静止画像としての顔画像から動画像を生成する技術、特に、生成された動画像に基づいてキャラクタと視聴者の非言語コミュニケーションを実現する技術、を提供することにある。

【課題を解決するための手段】

【0009】

本発明のある態様における動画生成装置は、静止画像としての顔画像を取得する画像取得部と、顔画像に含まれる特定の部分領域を変化させた画像として、複数の派生画像を自動生成する派生画像生成部と、複数の派生画像に基づいて、顔画像を動画表示させる画像表示部と、を備える。

10

【0010】

本発明のある態様におけるライブ通信システムは、配信者により操作される配信装置、視聴者により操作される視聴装置およびサーバを含む。

サーバは、ユーザにより造形されたキャラクタの静止画像としての顔画像を取得する顔画像取得部と、顔画像に含まれる特定の部分領域を変化させた画像として、複数の派生画像を自動生成する派生画像生成部と、複数の派生画像を視聴装置に送信する画像送信部と、を備える。

配信装置は、配信者を撮像する撮像部と、配信者の撮像画像から、配信者の動作の特徴を示す動作情報を生成する動作情報生成部と、配信者から第1メッセージを受け付ける第1メッセージ取得部と、動作情報および第1メッセージを送信する送信部と、を備える。

20

視聴装置は、動作情報および第1メッセージを受信する受信部と、視聴装置の画面にキャラクタを表示させ、動作情報に応じて複数の派生画像から表示対象となる派生画像を選択することにより、キャラクタを動作させる視聴キャラクタ表示部と、キャラクタの動作に合わせて第1メッセージを出力する第1メッセージ出力部と、を備える。

【発明の効果】

【0011】

本発明によれば、静止画像としての顔画像から豊かな表情変化をつくりだしやすくなる。

【図面の簡単な説明】

30

【0012】

【図1】ライブ通信システムのハードウェア構成図である。

【図2】ライブ通信システムによる会話サービスを説明するための模式図である。

【図3】配信装置およびサーバの機能ブロック図である。

【図4】視聴装置の機能ブロック図である。

【図5】視聴装置に表示される視聴者会話画面の画面図である。

【図6】配信装置に表示される配信者会話画面の画面図である。

【図7】配信装置および視聴装置の双方においてキャラクタを表示させる処理のシーケンス図である。

【図8】図8(a)は、キャラクタ静止画像のうち、キャラクタの顔の拡大図である。図8(b)は、キャラクタ静止画像から抽出された特定部位を示す図である。

40

【図9】キャラクタ静止画像における左目を示す図である。

【図10】左目において半眼となったときの状態を示す図である。

【図11】左目において閉眼となったときの状態を示す図である。

【図12】左目のキャラクタ動画像を示す図である。

【図13】図13(a)は、キャラクタの閉口時の顔の拡大図である。図13(b)は、キャラクタの開口時の顔の拡大図である。

【図14】派生画像の生成過程を示すフローチャートである。

【図15】左上眼瞼の低下方法の別例を説明するための図である。

【発明を実施するための形態】

50

## 【 0 0 1 3 】

図 1 は、ライブ通信システム 1 0 0 のハードウェア構成図である。

ライブ通信システム 1 0 0 においては、サーバ 1 0 2 に対して、複数の配信装置 1 0 4 a、1 0 4 b・・・1 0 4 m（以下、まとめて言うときや特に区別しないときには「配信装置 1 0 4」と総称する）および複数の視聴装置 1 0 8 a、1 0 8 b・・・1 0 8 n（以下、まとめて言うときや特に区別しないときには「視聴装置 1 0 8」と総称する）がインターネット 1 0 6 を介して接続される。配信装置 1 0 4 および視聴装置 1 0 8 としては、スマートフォンを想定しているが、ラップトップ PC などの汎用コンピュータであってもよい。

## 【 0 0 1 4 】

配信装置 1 0 4 を使用するユーザを「配信者」とよぶ。視聴装置 1 0 8 を使用するユーザを「視聴者」とよぶ。配信者がサービスの提供者となり、視聴者がサービスの享受者となる。

## 【 0 0 1 5 】

図 2 は、ライブ通信システム 1 0 0 による会話サービスを説明するための模式図である。

本実施形態においては、1 つの配信装置 1 0 4 に対して多数の視聴装置 1 0 8 が同時接続できる。以下、1 対多通信を前提として説明する。図 2 は、1 人の配信者 1 1 0 に対して、3 人の視聴者 1 1 2 a、1 1 2 b、1 1 2 c（以下、まとめて言うときや特に区別しないときには「視聴者 1 1 2」と総称する）がグループで会話をする状況を想定している。

## 【 0 0 1 6 】

配信者 1 1 0 にはあらかじめキャラクタが対応づけられる。配信装置 1 0 4 には配信者 1 1 0 に対応づけられる配信キャラクタ 1 1 4 a が表示される。視聴装置 1 0 8 にも視聴キャラクタ 1 1 4 b が表示される。配信キャラクタ 1 1 4 a と視聴キャラクタ 1 1 4 b はまったく同じキャラクタである。以下、配信キャラクタ 1 1 4 a と視聴キャラクタ 1 1 4 b を特に区別しないときには単に「キャラクタ 1 1 4」とよぶ。

## 【 0 0 1 7 】

配信者 1 1 0 が配信装置 1 0 4 の前でしゃべると、その音声データは視聴装置 1 0 8 に一斉送信される。配信装置 1 0 4 は配信者の口の動きを検出し、配信キャラクタ 1 1 4 a の口も動かす。また、これに連動して、視聴装置 1 0 8 は視聴キャラクタ 1 1 4 b の口も動かす。視聴者 1 1 2 には、配信者ではなく、あたかも視聴キャラクタ 1 1 4 b が自らの意思でしゃべっているように見える。いわば、配信者 1 1 0 は「声優」としてキャラクタ 1 1 4 の代弁者となる。以下、配信者 1 1 0 により入力される音声メッセージを「第 1 メッセージ」とよぶ。

## 【 0 0 1 8 】

視聴者 1 1 2 は、テキストにて視聴装置 1 0 8 にメッセージ（以下、「第 2 メッセージ」とよぶ）を入力する。第 2 メッセージは、配信装置 1 0 4 および視聴装置 1 0 8 に一斉送信される。視聴者 1 1 2 は、視聴キャラクタ 1 1 4 b の話を聞きながら、第 2 メッセージを入力することでキャラクタ 1 1 4（配信者 1 1 0）と会話する。

## 【 0 0 1 9 】

キャラクタ 1 1 4 の基本設定（キャラクタ 1 1 4 の容姿や表情変化の豊かさなど）と配信者 1 1 0 の声質や会話力によって、キャラクタ 1 1 4 の個性がかたちづくられる。

## 【 0 0 2 0 】

図 3 は、配信装置 1 0 4 およびサーバ 1 0 2 の機能ブロック図である。

配信装置 1 0 4 およびサーバ 1 0 2 の各構成要素は、CPU（Central Processing Unit）および各種コプロセッサなどの演算器、メモリやストレージといった記憶装置、それらを連結する有線または無線の通信線を含むハードウェアと、記憶装置に格納され、演算器に処理命令を供給するソフトウェアによって実現される。コンピュータプログラムは、デバイスドライバ、オペレーティングシステム、それらの上位層に位置する各種アプリケ

10

20

30

40

50

ーションプログラム、また、これらのプログラムに共通機能を提供するライブラリによって構成されてもよい。以下に説明する各ブロックは、ハードウェア単位の構成ではなく、機能単位のブロックを示している。

視聴装置 108 についても同様である。

【0021】

(配信装置 104)

配信装置 104 は、ユーザインタフェース処理部 116、撮像部 118、データ処理部 120、通信部 122 およびデータ格納部 124 を含む。

通信部 122 は、通信処理を担当する。撮像部 118 は、配信者 110 を撮像する。ユーザインタフェース処理部 116 は、配信者 110 からの操作を受け付けるほか、画像表示や音声出力など、ユーザインタフェースに関する処理を担当する。データ格納部 124 はキャラクタ 114 の画像データを含む各種データを格納する。データ処理部 120 は、通信部 122、ユーザインタフェース処理部 116、撮像部 118 により取得されたデータおよびデータ格納部 124 に格納されているデータに基づいて各種処理を実行する。データ処理部 120 は、ユーザインタフェース処理部 116、撮像部 118、通信部 122 およびデータ格納部 124 のインタフェースとしても機能する。

【0022】

通信部 122 は、送信部 148 と受信部 150 を含む。

【0023】

ユーザインタフェース処理部 116 は、配信者 110 からの入力を受け付ける入力部 126 と配信者 110 に各種情報を出力する出力部 128 を含む。

入力部 126 は、動作検出部 130 および第 1 メッセージ取得部 132 を含む。動作検出部 130 は、撮像部 118 を介して配信者 110 の表情の変化を検出する。第 1 メッセージ取得部 132 は、第 1 メッセージ（音声メッセージ）を取得する。

【0024】

出力部 128 は、配信キャラクタ表示部 138 および第 2 メッセージ出力部 140 を含む。

配信キャラクタ表示部 138 は配信装置 104 に配信キャラクタ 114 a を表示させる。第 2 メッセージ出力部 140 は視聴者 112 による第 2 メッセージ（テキストメッセージ）を配信装置 104 に表示させる。

【0025】

データ処理部 120 は、動作情報生成部 144 を含む。

動作情報生成部 144 は、動作検出部 130 の検出結果から「動作情報」を生成する。本実施形態における動作情報とは、配信者 110 の表情を示す特徴ベクトル、たとえば、眉の角度、目や口の開度を示す数値など、配信者 110 の動作状態、特に、表情を示す数値データである。

【0026】

(サーバ 102)

サーバ 102 は、通信部 152、データ処理部 154 およびデータ格納部 156 を含む。

通信部 152 は、通信処理を担当する。データ格納部 156 は各種データを格納する。データ処理部 154 は、通信部 152 により取得されたデータおよびデータ格納部 156 に格納されているデータに基づいて各種処理を実行する。データ処理部 154 は、通信部 152 およびデータ格納部 156 のインタフェースとしても機能する。

【0027】

通信部 152 は、送信部 210 と受信部 212 を含む。

送信部 210 は、画像送信部 214 を含む。画像送信部 214 は、キャラクタ 114 の動画像を形成するための画像群（以下、まとめて「キャラクタ動画像」とよぶ）を配信装置 104 および視聴装置 108 に送信する。受信部 212 は、画像取得部 216 を含む。画像取得部 216 は、キャラクタ 114 の顔画像を含む静止画像（以下、「キャラクタ静

10

20

30

40

50

止画像」とよぶ)を配信装置104などの外部装置から取得する。

【0028】

データ処理部154は、ライブ制御部218、部分抽出部220および派生画像生成部222を含む。

ライブ制御部218は、配信装置104と視聴装置108のコミュニケーションを仲介・統括する。部分抽出部220は、キャラクタ静止画像から、口、目などの特定部位に対応する画像領域を抽出する。派生画像生成部222は、特定部位を対象として後述の「派生画像」を生成する。キャラクタ静止画像と複数の派生画像の集合としてキャラクタ動画像が生成されるが、詳細は図12に関連して後述する。

【0029】

データ格納部156は、複数のキャラクタ114それぞれのキャラクタ動画像を格納するキャラクタ格納部158を含む。このほか、データ格納部156は後述の機械学習モデル(以下、「画像認識モデル」とよぶ)も格納する。

【0030】

図4は、視聴装置108の機能ブロック図である。

視聴装置108は、ユーザインタフェース処理部162、データ処理部166、通信部164およびデータ格納部168を含む。

通信部164は、通信処理を担当する。ユーザインタフェース処理部162は、視聴者112からの操作を受け付けるほか、画像表示や音声出力など、ユーザインタフェースに関する処理を担当する。データ格納部168はキャラクタ動画像を含む各種データを格納する。配信装置104と視聴装置108はどちらもキャラクタ動画像をそれぞれ格納する。データ処理部166は、通信部164およびユーザインタフェース処理部162により取得されたデータおよびデータ格納部168に格納されているデータに基づいて各種処理を実行する。データ処理部166は、ユーザインタフェース処理部162、通信部164およびデータ格納部168のインタフェースとしても機能する。

【0031】

通信部122は、送信部186と受信部190を含む。

【0032】

ユーザインタフェース処理部162は、視聴者112からの入力を受け付ける入力部170と視聴者112に各種情報を出力する出力部172を含む。

入力部170は、第2メッセージ取得部174およびキャラクタ選択部176を含む。第2メッセージ取得部174は、第2メッセージ(テキストメッセージ)を取得する。キャラクタ選択部176は、会話対象となるキャラクタ114の選択を受け付ける。

【0033】

出力部172は、視聴キャラクタ表示部180、第1メッセージ出力部182および第2メッセージ出力部184を含む。

視聴キャラクタ表示部180は、視聴装置108に視聴キャラクタ114bを表示させる。第1メッセージ出力部182は配信者110による第1メッセージを視聴装置108から音声出力する。第2メッセージ出力部184は視聴者112による第2メッセージを視聴装置108に画面表示させる。

【0034】

ライブ通信サービスの開始時において、サーバ102のライブ制御部218は、視聴装置108にキャラクタ選択画面(図示せず)を表示させる。視聴者112は、視聴装置108に表示されたキャラクタ選択画面において会話したい視聴キャラクタ114b(配信者110)を選ぶ。ライブ制御部218は、会話可能な状態(スタンバイ状態)にある配信者110に対応するキャラクタ114を会話可能なキャラクタとしてキャラクタ選択画面に提示する。キャラクタ選択部176は視聴キャラクタ114bの選択を受け付ける。

以下、視聴者112は視聴キャラクタ114b1(図5参照)との会話を楽しむものとして説明する。

【0035】

10

20

30

40

50

図5は、視聴装置108に表示される視聴者会話画面196の画面図である。

視聴装置108の視聴キャラクタ表示部180は、おしゃべりの相手として選ばれた視聴キャラクタ114b1を動画表示させる。視聴者112は、視聴キャラクタ114b1との会話を楽しむ。視聴者112からは配信者110の姿が見えることはない。視聴装置108の第1メッセージ出力部182は、視聴キャラクタ114b1を担当する配信者110からの音声(第1メッセージ)を出力する。視聴者112は、テキストにより視聴キャラクタ114b1へのメッセージ(第2メッセージ)を入力する。サーバ102は、第1メッセージおよび第2メッセージを配信装置104、視聴装置108の間で共有させる。

#### 【0036】

第2メッセージ出力部184はメッセージ領域198に第2メッセージをテキスト表示させる。1人の配信者110が複数の視聴者112と同時並行にて会話する方式であるため、サーバ102は各視聴者112の第2メッセージを配信装置104および複数の視聴者112に一斉配信する。視聴者112は、メッセージ領域198により自分の第2メッセージだけでなく他の視聴者112の第2メッセージも視認できる。第2メッセージ出力部184は第2メッセージを受信順に並べて表示させる。

#### 【0037】

図6は、配信装置104に表示される配信者会話画面206の画面図である。

図6に示す配信者会話画面206は、配信キャラクタ114a1(視聴キャラクタ114b1に対応)を担当する配信者110の配信装置104に表示される。配信キャラクタ表示部138は、配信キャラクタ114a1を動画表示させる。視聴キャラクタ114b1と配信キャラクタ114a1は同一キャラクタであり、動きも同じである。配信者110は、自らの分身である配信キャラクタ114a1の動きを配信装置104にて確認しながら複数の視聴者112(視聴キャラクタ114b1を会話の相手として選んだ視聴者112)と会話する。配信者110は、配信装置104に向かって音声にて第1メッセージを入力する。送信部148は第1メッセージをサーバ102に送信し、サーバ102のライブ制御部218は第1メッセージを複数の視聴装置108に配信する。

#### 【0038】

第2メッセージ出力部140は、視聴者112からの第2メッセージをメッセージ領域198に時系列順にて表示させる。配信者110はメッセージ領域198に次々に表示される第2メッセージをみながら不特定多数の視聴者112に話しかける。配信装置104のカメラ208(撮像部118)は、配信者110を撮像する。動作検出部130は、撮像画像から配信者110の表情を検出し、動作情報生成部144は動作情報を生成する。配信キャラクタ表示部138は、動作情報に基づいて配信キャラクタ114a1の動きを決定する。また、詳細は後述するが視聴キャラクタ114b1も同じ動作情報に基づいて視聴装置108にて動画表示される。

#### 【0039】

図7は、配信装置104および視聴装置108の双方においてキャラクタ114を表示させる処理のシーケンス図である。

配信装置104の動作検出部130は、配信者110の動作を検出する(S10)。本実施形態においては110の右目および左目の開眼度合い(以下、「開眼度」とよぶ)と、口の開口度合い(以下、「開口度」とよぶ)を検出する。動作情報生成部144は、開眼度および開口度を示す数値データとして動作情報を生成する(S12)。

#### 【0040】

配信装置104の送信部148は、サーバ102経由にて、各視聴装置108に動作情報を送信する(S14)。配信キャラクタ表示部138は、動作情報に基づいて配信キャラクタ114a1の動画像を生成し(S16)、配信装置104に配信キャラクタ114a1を動画表示させる(S18)。

#### 【0041】

一方、視聴装置108の視聴キャラクタ表示部180も、受信された動作情報に基づい

10

20

30

40

50

て視聴キャラクタ 1 1 4 b 1 の動画像を生成し ( S 2 0 )、視聴装置 1 0 8 に視聴キャラクタ 1 1 4 b 1 を動画表示させる ( S 2 2 )。

【 0 0 4 2 】

配信キャラクタ 1 1 4 a 1 と視聴キャラクタ 1 1 4 b 1 は、同一のキャラクタであり、かつ、同一の動作情報に基づいて動きが決まるため、配信装置 1 0 4 および視聴装置 1 0 8 それぞれにおいて配信キャラクタ 1 1 4 a 1 と視聴キャラクタ 1 1 4 b は同じように表情を変化させる。

【 0 0 4 3 】

配信装置 1 0 4 は、キャラクタ 1 1 4 の動画像 ( キャラクタ画像 ) そのものを視聴装置 1 0 8 に送信しているのではなく、キャラクタ 1 1 4 の動作を定義する動作情報のみを送信している。動画像そのものを送信するよりも送信データ量が少ないため、配信キャラクタ 1 1 4 a 1 と視聴キャラクタ 1 1 4 b 1 それぞれの動きを速やかに連動させやすい。い

10

【 0 0 4 4 】

いかえれば、動作情報が配信キャラクタ 1 1 4 a 1 の動きとして反映されるまでの時間と、同じ動作情報が視聴キャラクタ 1 1 4 b 1 の動きとして反映されるまでの時間の差が小さいため、視聴者 1 1 2 はキャラクタ 1 1 4 ( 配信者 1 1 0 ) との快適な会話を楽しむことができる。

20

【 0 0 4 5 】

第 1 メッセージ ( 音声メッセージ ) と動作情報は一体化して配信装置 1 0 4 から視聴装置 1 0 8 に送信されてもよい。この場合には、視聴キャラクタ表示部 1 8 0 は、視聴キャラクタ 1 1 4 b 1 の動作と音声メッセージがずれにくくなる。また、動画像そのものではなく、動画像を決定するための動作情報を送信することは、インターネット 1 0 6 に対する通信負荷を抑制する上でも有効である。

【 0 0 4 6 】

図 8 ( a ) は、キャラクタ静止画像 2 2 4 のうち、キャラクタ 1 1 4 の顔の拡大図である。図 8 ( b ) は、キャラクタ静止画像 2 2 4 から抽出された特定部位を示す図である。

【 0 0 4 7 】

キャラクタ静止画像 2 2 4 を生成したユーザ ( 以下、「デザイナー」とよぶ ) は、配信装置 1 0 4 などの端末装置からサーバ 1 0 2 に自作のキャラクタ静止画像 2 2 4 を送信する。サーバ 1 0 2 の画像取得部 2 1 6 は、キャラクタ静止画像 2 2 4 を取得する。

30

【 0 0 4 8 】

部分抽出部 2 2 0 はキャラクタ静止画像 2 2 4 から、特定部位の画像領域 ( 以下、「部分領域」とよぶ ) を特定する。以下、キャラクタ静止画像 2 2 4 から部分領域を特定する機能を「部分特定機能」とよぶ。

40

【 0 0 4 9 】

本実施形態の部分抽出部 2 2 0 は、キャラクタ 1 1 4 の口に対応する口領域 2 3 2、右目に対応する右目領域 2 3 0 R、左目に対応する左目領域 2 3 0 L を部分領域として特定する。以下、右目領域 2 3 0 R と左目領域 2 3 0 L をまとめていうときや特に区別しないときには単に「目領域 2 3 0」とよぶ。目領域 2 3 0 からは、更に、上眼瞼 2 2 6 と眼球部 2 2 8 を特定する。右目領域 2 3 0 R は、右上眼瞼 2 2 6 R および右眼球部 2 2 8 R を含む。左目領域 2 3 0 L は、左上眼瞼 2 2 6 L と左眼球部 2 2 8 L を含む。上眼瞼 2 2 6 は、目のうち、睫毛および上瞼皮に対応する。眼球部 2 2 8 は、目のうち、眼球および下眼瞼に対応する。

【 0 0 4 8 】

サーバ 1 0 2 のデータ格納部 1 5 6 は、部分特定機能を実現するための機械学習モデルである画像認識モデルを格納する。ライブ通信サービスの開始前に、あらかじめ多種多様なキャラクタ静止画像 2 2 4 を画像認識モデルに学習させておく。本実施形態の画像認識モデルとしては、多数の中間層を有するニューラルネットワーク・モデル ( ディープラーニング・モデル ) を想定する。

【 0 0 4 9 】

まず、あるキャラクタ静止画像 2 2 4 を画像認識モデルの入力データとして与えるとともに、キャラクタ静止画像 2 2 4 の部分領域を教師データとして指定する。画像認識モデ

50

ルは、既知の教師あり学習方式により、内部パラメータを調整することでキャラクタ静止画像 2 2 4 から部分領域を抽出する方法を学習する。別のキャラクタ静止画像 2 2 4 についても同様の処理を繰り返すことで画像認識モデルの学習を繰り返す。

#### 【 0 0 5 0 】

画像認識モデルは、回答（出力データ）と正解（教師データ）の差分に基づいて誤差逆伝播法（Backpropagation）等の修正アルゴリズムにより、内部パラメータを調整する。数万から数十万種類のキャラクタ静止画像 2 2 4 を画像認識モデルに入力し、それぞれのキャラクタ静止画像 2 2 4 について回答（教師データ）を与えて内部パラメータの修正を繰り返すことにより、画像認識モデルの部分特定機能を鍛えていく。

#### 【 0 0 5 1 】

本実施形態においては、部分抽出部 2 2 0 は、右上眼瞼 2 2 6 R、左上眼瞼 2 2 6 L、右眼球部 2 2 8 R、左眼球部 2 2 8 L、口領域 2 3 2 の 5 領域を部分領域として抽出するものとして説明する。このほかにも、鼻や眉毛、髪の毛など他の部分領域を抽出してもよい。

#### 【 0 0 5 2 】

図 9 は、キャラクタ静止画像 2 2 4 における左目 2 3 4 を示す図である。

まず、キャラクタ 1 1 4 の左目 2 3 4 を対象として、キャラクタ静止画像 2 2 4 からキャラクタ動画像を生成する方法について説明する。上述したように、キャラクタ静止画像 2 2 4 の左目 2 3 4 からは左上眼瞼 2 2 6 L および左眼球部 2 2 8 L が抽出される。キャラクタ 1 1 4 が開眼しているとき、左上眼瞼 2 2 6 L は「睫毛」のみが表示される。

#### 【 0 0 5 3 】

図 1 0 は、左目 2 3 4 において半眼となったときの状態を示す図である。

派生画像生成部 2 2 2 は、キャラクタ静止画像 2 2 4 の上眼瞼 2 2 6 の位置を下降させ、半眼となった左目の画像を自動生成する。以下、このようにキャラクタ静止画像 2 2 4 から自動生成される画像を「派生画像」とよぶ。派生画像生成部 2 2 2 は、左目領域 2 3 0 L 内において、左上眼瞼 2 2 6 L（睫毛）の位置を下方に移動させる。このとき、派生画像生成部 2 2 2 は、左上眼瞼 2 2 6 L のうち左目領域 2 3 0 L からはみ出す部分である外部分 2 5 2 は派生画像からカットする。また、派生画像生成部 2 2 2 は、左上眼瞼 2 2 6 L（睫毛）が下降したときの上部（瞼皮部 2 3 6）を所定色で塗りつぶす。本実施形態においては、派生画像生成部 2 2 2 は瞼皮部 2 3 6 を所定の肌色で塗りつぶす。

#### 【 0 0 5 4 】

このように、派生画像生成部 2 2 2 は、左上眼瞼 2 2 6 L（睫毛）を左目領域 2 3 0 L の範囲内にて下降させ、睫毛の上部である瞼皮部 2 3 6 を肌色に塗色した派生画像を生成する。このとき、瞳は動かさない。瞼皮部 2 3 6 の色は、キャラクタ 1 1 4 の肌色、たとえば、頬部分の色と同一色としてもよい。あるいは、瞼皮部 2 3 6 の色は、キャラクタ 1 1 4 の頬、額、顎、首などの複数の所定領域の色の平均色としてもよい。このようは制御方法により、派生画像生成部 2 2 2 は左目 2 3 4 における開眼状態の左目から、半眼状態の派生画像を自動的に生成する。

#### 【 0 0 5 5 】

図 1 1 は、左目 2 3 4 において閉眼となったときの状態を示す図である。

派生画像生成部 2 2 2 は、複数の開眼度に対応して複数の派生画像、いいかえれば、左上眼瞼 2 2 6 L の下降度が異なる複数の派生画像を生成する。派生画像生成部 2 2 2 は、左上眼瞼 2 2 6 L の睫毛ラインが左目領域 2 3 0 L の下部まで到達し、左目領域 2 3 0 L 全体が瞼皮部 2 3 6 に覆われる派生画像を生成することで左目 2 3 4 の閉眼を表現する。

#### 【 0 0 5 6 】

図 1 2 は、左目 2 3 4 のキャラクタ動画像 2 4 2 を示す図である。

上述したように、派生画像生成部 2 2 2 は開眼状態のキャラクタ静止画像 2 2 4 から、複数の派生画像 2 4 0 を生成する。キャラクタ静止画像 2 2 4 に派生画像 2 4 0 を重ねることで、さまざまな開眼度に対応したキャラクタ 1 1 4 の画像を用意する。1 枚のキャラクタ静止画像 2 2 4 と複数の派生画像 2 4 0 の集合により、キャラクタ動画像 2 4 2 が形

10

20

30

40

50

成される。右目（右目領域 2 3 0 R）についても同様にして、派生画像生成部 2 2 2 は複数の派生画像を生成する。口についても同様である（後述）。

【 0 0 5 7 】

1 枚のキャラクタ静止画像 2 2 4 から、右目に関する複数の派生画像 2 4 0、左目に関する複数の派生画像 2 4 0 および口に関する複数の派生画像 2 4 0 の 3 種類の画像グループが形成される。これら 3 種類の画像グループに含まれる多数枚の派生画像 2 4 0 と 1 枚のキャラクタ静止画像 2 2 4 を組み合わせることにより、キャラクタ 1 1 4 の右目の開眼度、左目の開眼度および口の開口度を多様に表現できる。

以下、キャラクタ静止画像 2 2 4 から複数の派生画像を自動生成する機能を「派生画像生成機能」とよぶ。

【 0 0 5 8 】

図 1 3 ( a ) は、キャラクタ 1 1 4 の閉口時の顔の拡大図である。図 1 3 ( b ) は、1 1 4 の開口時の顔の拡大図である。

部分抽出部 2 2 0 は、キャラクタ静止画像 2 2 4 から口領域 2 3 2 を抽出する。部分抽出部 2 2 0 は、次に、口の状態が開口または閉口のいずれであるかを判定する。開口・閉口判定も上述した画像認識モデルに基づいて判定すればよい。図 1 3 ( a ) に示すように閉口しているときには、派生画像生成部 2 2 2 は口を囲む位置に長方形の開口範囲 2 4 4 を設定する。開口範囲 2 4 4 は、キャラクタ 1 1 4 の顔幅 X 等に基づいて適正な大きさに定められる。開口範囲 2 4 4 のサイズも画像認識モデルに基づいて部分領域特定機能と同様の方式により設定されればよい。

【 0 0 5 9 】

派生画像生成部 2 2 2 は、口領域 2 3 2 を上下左右に輪状に拡げることで複数の派生画像 2 4 0 を生成する。派生画像生成部 2 2 2 は、開口時の口の中を赤系統の所定色により塗りつぶす。派生画像生成部 2 2 2 は、開口サイズが開口範囲 2 4 4 に達するまでキャラクタ 1 1 4 を開口させる（図 1 3 ( b )）。

【 0 0 6 0 】

一方、キャラクタ静止画像 2 2 4 においてキャラクタ 1 1 4 が開口しているときには、派生画像生成部 2 2 2 は、キャラクタ静止画像 2 2 4 の口を徐々に閉口させつつ複数の派生画像 2 4 0 を作成すればよい。このような制御方法により、キャラクタ静止画像 2 2 4 に基づいて閉口状態から開口状態までの複数の派生画像 2 4 0 を生成する。

【 0 0 6 1 】

複数のキャラクタ静止画像 2 2 4 を対象として機械学習させることにより、キャラクタ静止画像 2 2 4 ごとに適正サイズの開口範囲 2 4 4 を決定すればよい。別の方法として、部分抽出部 2 2 0 は、キャラクタ 1 1 4 の顔幅 X と、鼻から顎までの距離 Y を算出し、開口範囲 2 4 4 の横幅 X 1 を  $X \times n 1$ 、縦幅 Y 1 を  $Y \times n 2$  と設定してもよい。ここで、n 1 および n 2 はあらかじめ設定された所定倍率である。

【 0 0 6 2 】

図 1 4 は、派生画像の生成過程を示すフローチャートである。

サーバ 1 0 2 の画像取得部 2 1 6 は、まず、キャラクタ静止画像を取得する（S 1 0）。部分抽出部 2 2 0 は、画像認識モデル（機械学習モデル）に基づいて、まず、キャラクタ静止画像から目に該当する部分領域（目領域 2 3 0）を特定する（S 1 2）。より具体的には、図 8 ( b ) に関連して説明したように、右上眼瞼 2 2 6 R、右眼球部 2 2 8 R、左上眼瞼 2 2 6 L および左眼球部 2 2 8 L の 4 つの部分領域を特定する。

【 0 0 6 3 】

部分抽出部 2 2 0 は、キャラクタ 1 1 4 の口に対応する部分領域（口領域 2 3 2）を特定する（S 1 4）。派生画像生成部 2 2 2 は、右目領域 2 3 0 R および左目領域 2 3 0 L のそれぞれについて派生画像群を生成する（S 1 6）。また、派生画像生成部 2 2 2 は、口領域 2 3 2 について派生画像群を生成する（S 1 8）。派生画像生成部 2 2 2 は、右目、左目、口それぞれについて形成される派生画像群とキャラクタ静止画像をまとめてキャラクタ動画像として登録する。サーバ 1 0 2 の画像送信部 2 1 4 は、生成されたキャラク

10

20

30

40

50

タ動画像を配信装置 104 および視聴装置 108 の双方に送信する (S20)。

【0064】

キャラクタ動画像の送信後、このキャラクタ 114 によるライブ通信サービスが可能となる。

配信者 110 が右目を開閉させるとき、動作検出部 130 は配信者 110 の右目の開眼度を検出する。配信者 110 の右目の開眼度を示す特徴データは動作情報として、配信装置 104 からサーバ 102 に送信される。また、この動作情報はサーバ 102 から視聴装置 108 にも送信される。

【0065】

配信装置 104 の配信キャラクタ表示部 138 は、動作情報に含まれる右目の開眼度に対応する派生画像 240 を選択し、キャラクタ静止画像 224 の右目部分に選択した派生画像 240 を重畳する。たとえば、配信者 110 が右目を半眼にしたときには、配信キャラクタ表示部 138 はキャラクタ静止画像 224 の右目部分に半眼状態の派生画像 240 を重畳する。

10

【0066】

一方、視聴装置 108 の視聴キャラクタ表示部 180 も、動作情報に含まれる右目の開眼度に対応する派生画像 240 を選択し、キャラクタ静止画像 224 の右目部分に選択した派生画像 240 を重畳する。このような制御方法によれば、配信者 110 の右目の動き (開眼度) に対して、配信キャラクタ 114 a および視聴キャラクタ 114 b の右目の動き (開眼度) を追従させることができる。左目および口に関しても同様である。

20

【0067】

< 総括 >

以上、実施形態に基づいてライブ通信システム 100 について説明した。

ライブ通信システム 100 による会話サービスを盛り上げるためには、多数の配信者 110 がサービスを競い合う環境を作ることが望ましい。多数の配信者 110 がライブ通信システム 100 に参加するためには、多数のキャラクタ動画像を準備する必要がある。複数のイメージ画像の差分を補完することにより、あるイメージ画像から別のイメージ画像へ徐々に変化させていくモーフィング (morphing) 技術を適用する場合、複数のイメージ画像に相当する複数のキャラクタ静止画像 224 を準備する必要があり、手間がかかる。本実施形態においては、たった一枚のキャラクタ静止画像 224 を準備するだけで、キャラクタ 114 の目と口の変化を表現可能なキャラクタ動画像 (派生画像群) を自動的に生成できる。キャラクタのデザイナーは 1 人のキャラクタ 114 のために複数のキャラクタ静止画像 224 を作図する必要がないため、キャラクタ動画像の生産性を高めることができる。配信者 110 自身がデザイナーとなり、自ら演じてみたいキャラクタ 114 のキャラクタ静止画像 224 を自作してもよい。

30

【0068】

本実施形態においては、右目、左目および口という 3 つの特定部位について、3 種類の派生画像群が生成される。3 種類の派生画像群からそれぞれ 1 枚ずつ派生画像を選んで組み合わせることにより、たった 1 枚のキャラクタ静止画像 224 に基づいて多様な表情を作ることができる。いいかえれば、右目の開眼度、左目の開眼度および口の開口度それぞれを独立に制御できる。このあめ、キャラクタ 114 が右目を半眼にし、左目を大きく見開き、口は閉じる、といった複雑な表情であっても簡単に作成することができる。

40

【0069】

本実施形態においては、画像認識モデルにより、キャラクタ静止画像 224 から口や目などの部分領域を自動的に特定できる。特に、女の子を表すキャラクタ 114 は目が大きく睫毛を濃く描かれる傾向にあるため、上眼瞼 226 と眼球部 228 を高精度で特定しやすい。

【0070】

なお、本発明は上記実施形態や変形例に限定されるものではなく、要旨を逸脱しない範囲で構成要素を変形して具体化することができる。上記実施形態や変形例に開示されてい

50

る複数の構成要素を適宜組み合わせることにより種々の発明を形成してもよい。また、上記実施形態や変形例に示される全構成要素からいくつかの構成要素を削除してもよい。

【0071】

本実施形態においては、サーバ102が「動画生成装置」としての主要機能を担う。動画生成装置としての機能の一部は、サーバ102だけでなく、視聴装置108等の他の装置が担当してもよい。たとえば、部分特定機能、派生画像生成機能をサーバ102以外の装置が担ってもよい。また、動画表示機能は、配信装置104および視聴装置108が担ってもよいし、サーバ102自身が動画表示機能を実現してもよい。

【0072】

[変形例]

大量のキャラクタ静止画像224を画像認識モデル(機械学習モデル)に読み込ませ、画像認識モデルは回答データとして部分領域を特定する。ユーザ(教育者)は、キャラクタ静止画像224の部分領域を手動で指定し、これを教師データ(正解)として画像認識モデルに与える。画像認識モデルは、回答データと教師データの差分を最小化すべく内部調整、具体的には、ニューラルネットワークに含まれる中間ノードの重みを調整する。

【0073】

デザイナーが新たなキャラクタ静止画像224を作成したとき、画像認識モデルはこのキャラクタ静止画像224から部分領域を特定し、デザイナーに提示する。デザイナーは画像認識モデルが特定した部分領域を手動で修正してもよい。このような制御方法によれば、画像認識モデルにより部分領域を特定し、デザイナーは手動で部分領域を補正することで部分領域の抽出精度を高めることができる。

【0074】

変形例として、画像認識モデルを使用することなくデザイナーが部分領域をキャラクタ静止画像224から直接指定してもよい。たとえば、デザイナーはキャラクタ静止画像224において右上眼瞼226Rに対応する部分領域で指定し、これをサーバ102に通知してもよい。画像認識モデルによる部分特定機能を使わないためデザイナーの手間は増えるが、デザイナーは画像認識モデルに頼ることなくキャラクタ静止画像224の部分領域を精密に設定できる。

【0075】

瞬きのように、右目と左目は同時に開閉してもよい。このほかにも、ウィンクのように片目を開き、片目を閉じる動きを表現することもできる。

【0076】

動作対象となる部分領域は、目、口以外にも設定可能である。たとえば、眉毛を部分領域として特定し、配信者110の眉毛の角度を動作情報として検出してもよい。派生画像生成部222は、あらかじめ複数の眉毛の角度に対応して複数の派生画像240を生成しておく。配信キャラクタ表示部138等は配信者110の眉毛の動きに応じてキャラクタ114の眉毛の形を変化させてもよい。キャラクタ114の頬を部分領域として設定してもよい。配信者110の頬の膨らみ具合を動作情報として検出し、キャラクタ114のフェイスラインを配信者110の頬に連動させて変化させてもよい。この場合には、派生画像生成部222は、あらかじめ複数段階の頬の膨らみに対応して複数の派生画像240を生成しておけばよい。

【0077】

眉間を部分領域として特定してもよい。配信者110の眉間の皺を動作情報として検出し、配信者110の眉間に皺をしめす筋目を入れることで、配信者110とキャラクタ114の表情を連動させてもよい。

【0078】

口に関しては、開閉だけではなく、引き結ぶ、両方の口角を上げる、片方の口角を上げる、両方の口角を下げるなど複数のパターンに対応した派生画像群を生成しておき、配信キャラクタ表示部138等は、配信者110の口の形状に合わせてキャラクタ114の口を変化させてもよい。

10

20

30

40

50

## 【 0 0 7 9 】

部分抽出部 2 2 0 は、キャラクタ静止画像 2 2 4 からキャラクタ 1 1 4 の頭部全体を部分領域として抽出し、派生画像生成部 2 2 2 はキャラクタ 1 1 4 が頭部を傾げる複数の派生画像 2 4 0 を生成してもよい。動作検出部 1 3 0 は、配信者 1 1 0 の首の傾きを動作情報として検出し、配信キャラクタ表示部 1 3 8 等は配信者 1 1 0 の首の傾きに応じてキャラクタ 1 1 4 の首の傾きを変化させてもよい。

## 【 0 0 8 0 】

部分抽出部 2 2 0 は、キャラクタ静止画像 2 2 4 からキャラクタ 1 1 4 の髪に対応する部分領域を抽出してもよい。派生画像生成部 2 2 2 は、髪の揺れや乱れに対応する複数の派生画像 2 4 0 を生成してもよい。キャラクタ 1 1 4 の表示中において、配信キャラクタ表示部 1 3 8 等はキャラクタ 1 1 4 の髪を揺らせることにより、配信者 1 1 0 の動作に変化がないときでもキャラクタ 1 1 4 の髪が風に吹かれているかのような動きを表現してもよい。

10

## 【 0 0 8 1 】

部分抽出部 2 2 0 は、目領域 2 3 0 から瞳に対応する部分領域を抽出してもよい。派生画像生成部 2 2 2 は、目領域 2 3 0 内部において瞳を上下左右に移動させた派生画像を生成してもよい。配信キャラクタ表示部 1 3 8 等は、配信者 1 1 0 の瞳の位置（視線方向）を示す動作情報に基づいて、キャラクタ 1 1 4 の瞳の位置を変化させてもよい。

## 【 0 0 8 2 】

本実施形態においては、派生画像 2 4 0 の瞼皮部 2 3 6 を肌色系の所定色で塗色するとして説明した。肌色系に限らず、青系などアイシャドーを模した他の色により瞼皮部 2 3 6 を塗色してもよい。

20

## 【 0 0 8 3 】

図 1 5 は、左上眼瞼 2 2 6 L の低下方法の別例を説明するための図である。

本実施形態においては、左上眼瞼 2 2 6 L の形状を変化させることなく、左目領域 2 3 0 L 内において左上眼瞼 2 2 6 L の睫毛ラインを真っ直ぐに下降させながら複数の派生画像 2 4 0 を生成するとして説明した。変形例として、左目領域 2 3 0 L 内において左上眼瞼 2 2 6 L を下げるごとに左上眼瞼 2 2 6 L の睫毛ラインの形状を変化させてもよい。

## 【 0 0 8 4 】

具体的には、左目 2 3 4 の目頭と目尻を結ぶ曲線 2 5 0 を設定し、この曲線に合わせて左上眼瞼 2 2 6 L の睫毛ラインの形状を変化させてもよい。派生画像生成部 2 2 2 は、開眼度が小さくなるほど（閉眼に近づくほど）、曲線 2 5 0 を上に凸から下に凸に変化させる。派生画像生成部 2 2 2 は、左上眼瞼 2 2 6 L の睫毛ラインを曲線 2 5 0 に沿った形状に変化させる、このような制御方法によれば、開眼度に応じて左上眼瞼 2 2 6 L の形状が変化するため、眼球の僅かな丸みを表現できる。右目についても同様である。

30

## 【 0 0 8 5 】

部分抽出部 2 2 0 は、キャラクタ静止画像 2 2 4 におけるキャラクタ 1 1 4 の開口度を判定してもよい。部分抽出部 2 2 0 は、まず、口領域 2 3 2 を特定し、開口範囲 2 4 4 を設定する。開口範囲 2 4 4 の縦幅に比べて口領域 2 3 2 の縦幅が所定割合以上であれば開口と判定し、所定割合未満であれば閉口と判定してもよい。派生画像生成部 2 2 2 は、開口時には閉口状態に向かう複数の派生画像 2 4 0 を生成し、閉口時には開口状態に向かう複数の派生画像 2 4 0 を生成すればよい。なお、派生画像生成部 2 2 2 は、開口状態のキャラクタ静止画像 2 2 4 のみを対象として派生画像 2 4 0 を生成するとしてもよいし、閉口状態のキャラクタ静止画像 2 2 4 のみを対象として派生画像 2 4 0 を生成するとしてもよい。

40

## 【 0 0 8 6 】

部分抽出部 2 2 0 は、画像認識モデル（機械学習モデル）に基づいて開口範囲 2 4 4 を設定してもよい。変形例として、デザイナーはキャラクタ静止画像 2 2 4 において手動で開口範囲 2 4 4 を設定してもよい。このような制御方法によれば、デザイナーはキャラクタ 1 1 4 の最大開口サイズを指定できる。キャラクタ 1 1 4 をあまり大きく開口させたく

50

ないときや、キャラクタ 114 を極めて大きく開口させたいときにはデザイナーは自由に開口範囲 244 のサイズを決めることでキャラクタ 114 の最高開口度を調整すればよい。

【0087】

本実施形態においては、キャラクタ静止画像 224 は、デザイナーにより造形されたキャラクタ 114 の顔画像を含むとして説明した。変形例として、実在の人物の顔画像を含む静止画像から同様の方法により動画像を生成してもよい。あるいは、実在の人形を撮像した静止画像から同様の方法により動画像を生成してもよい。

【0088】

本実施形態においては、サーバ 102 がキャラクタ静止画像 224 を取得し、複数の派生画像 240 を生成することでキャラクタ動画像を準備するとして説明した。キャラクタ動画像の生成機能は、サーバ 102 ではなく、配信装置 104 等の端末装置が備えてもよい。具体的には、配信装置 104 が部分特定機能および派生画像生成機能を搭載してもよい。たとえば、配信者 110 はキャラクタ静止画像 224 を配信装置 104 に入力し、配信装置 104 の部分抽出部 220 および派生画像生成部 222 はキャラクタ静止画像 224 からキャラクタ動画像 242 を生成してもよい。配信者 110 は、配信装置 104 においてキャラクタ動画像 242 を動作確認し、必要に応じて、部分領域等の修正をした上で、サーバ 102 に修正後のキャラクタ動画像をアップロードしてもよい。

【0089】

本実施形態においては、キャラクタ静止画像 224 から生成した派生画像 240 をライブ通信システム 100 によるライブ通信サービスに適用するとして説明した。キャラクタ静止画像 224 から派生画像 240 を生成する技術は、ライブ通信以外のサービスにも応用可能である。たとえば、小さな子どもが描いた顔の絵（キャラクタ静止画像 224）から、派生画像 240 を生成することで、お絵描きの楽しみとともに、絵が動く楽しさを提供してもよい。

【符号の説明】

【0090】

100 ライブ通信システム、102 サーバ、104 配信装置、106 インターネット、108 視聴装置、110 配信者、112 視聴者、114 キャラクタ、114a 配信キャラクタ、114b 視聴キャラクタ、116 ユーザインタフェース処理部、118 撮像部、120 データ処理部、122 通信部、124 データ格納部、126 入力部、128 出力部、130 動作検出部、132 第1メッセージ取得部、138 配信キャラクタ表示部、140 第2メッセージ出力部、144 動作情報生成部、148 送信部、150 受信部、152 通信部、154 データ処理部、156 データ格納部、158 キャラクタ格納部、162 ユーザインタフェース処理部、164 通信部、166 データ処理部、168 データ格納部、170 入力部、172 出力部、174 第2メッセージ取得部、176 キャラクタ選択部、180 視聴キャラクタ表示部、182 第1メッセージ出力部、184 第2メッセージ出力部、186 送信部、190 受信部、196 視聴者会話画面、198 メッセージ領域、206 配信者会話画面、208 カメラ、210 送信部、212 受信部、214 画像送信部、216 画像取得部、218 ライブ制御部、220 部分抽出部、222 派生画像生成部、224 キャラクタ静止画像、226 上眼瞼、226L 左上眼瞼、226R 右上眼瞼、228 眼球部、228L 左眼球部、228R 右眼球部、230 目領域、230L 左目領域、230R 右目領域、232 口領域、234 左目、236 瞼皮部、240 派生画像、242 キャラクタ動画像、244 開口範囲、250 曲線、252 外部分

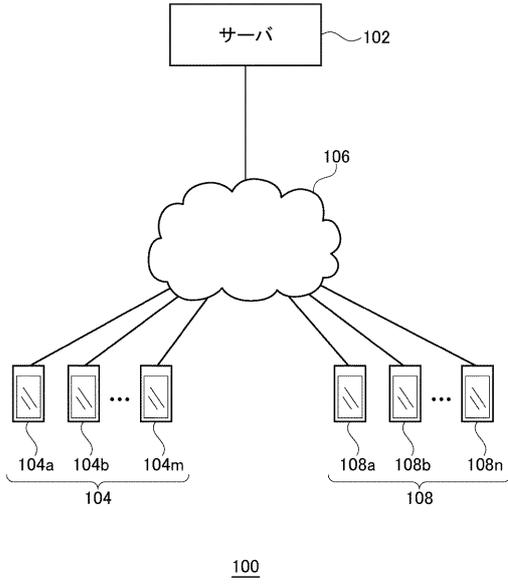
10

20

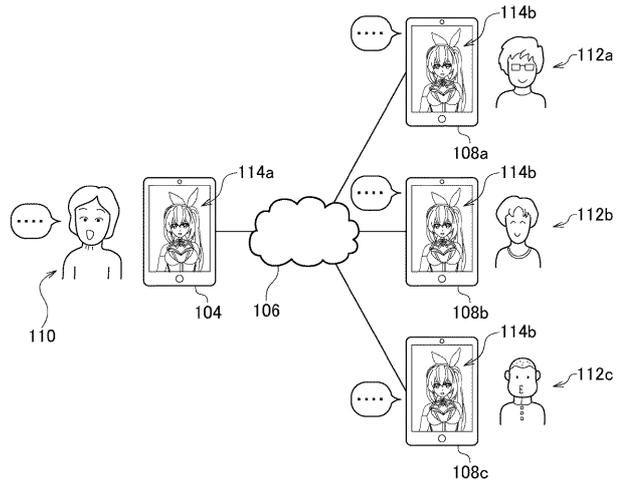
30

40

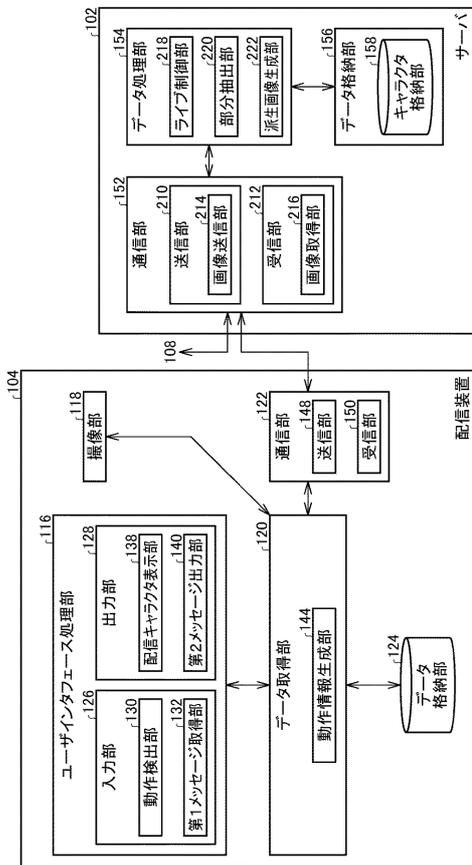
【 図 1 】



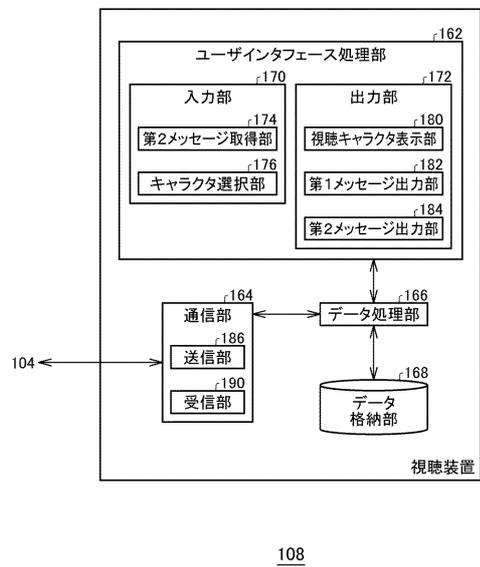
【 図 2 】



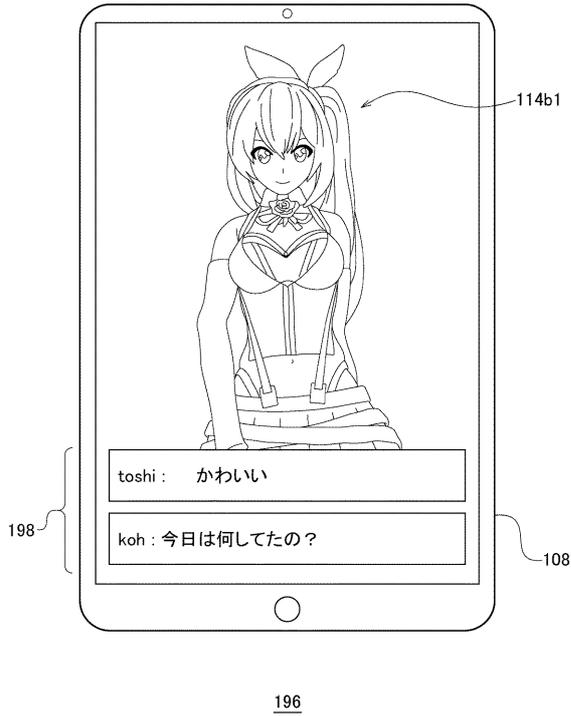
【 図 3 】



【 図 4 】



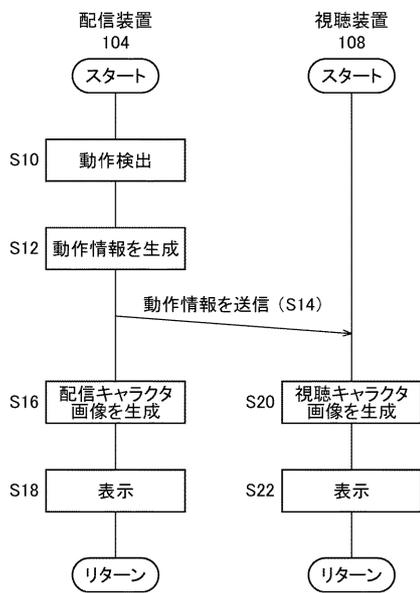
【図5】



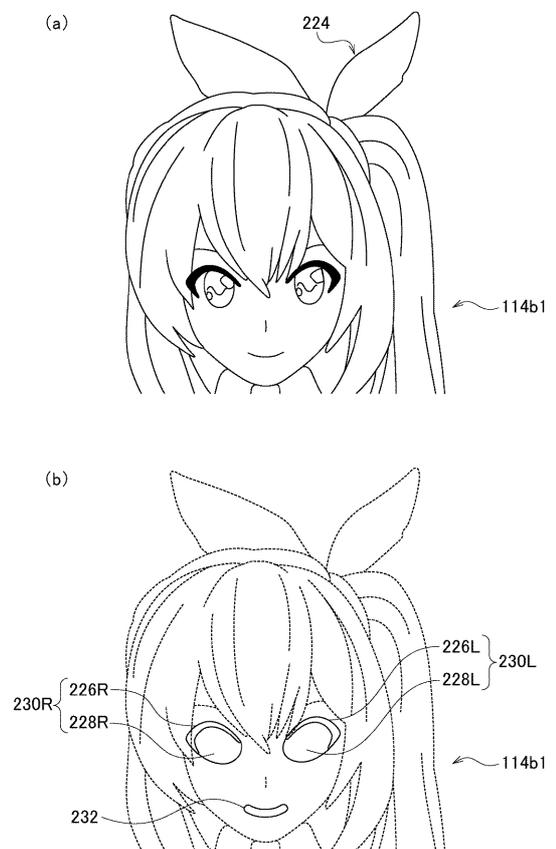
【図6】



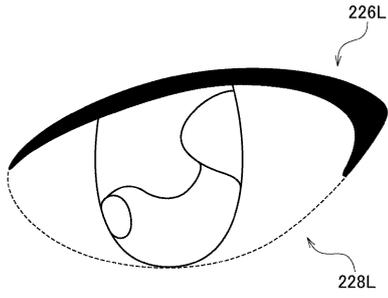
【図7】



【図8】

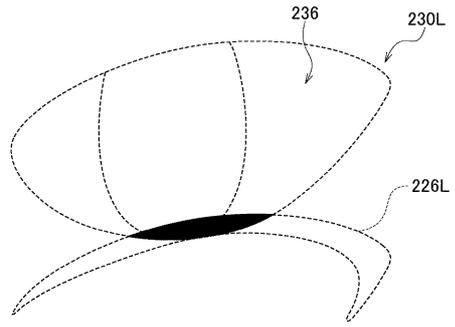


【 図 9 】



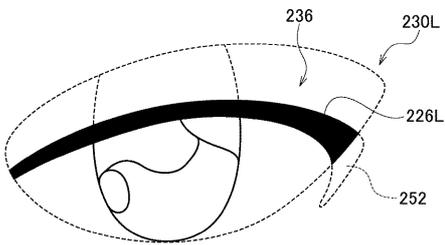
234

【 図 1 1 】



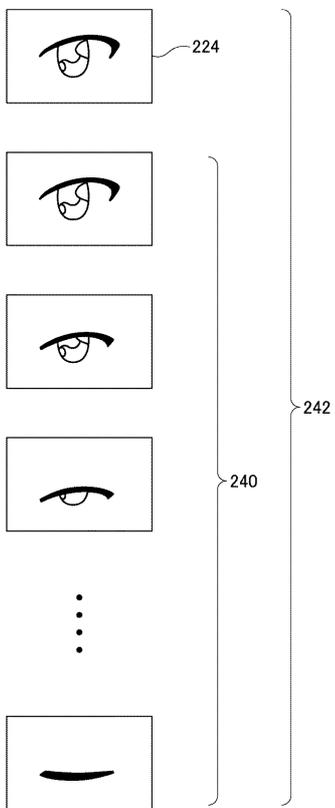
234

【 図 1 0 】

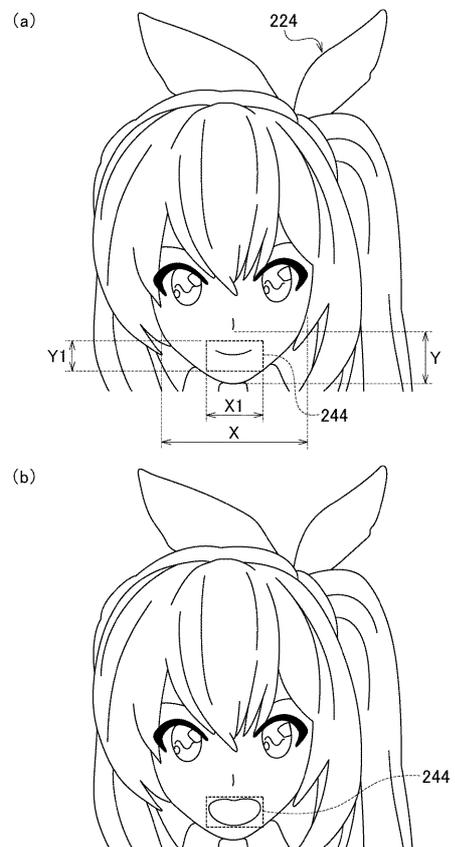


234

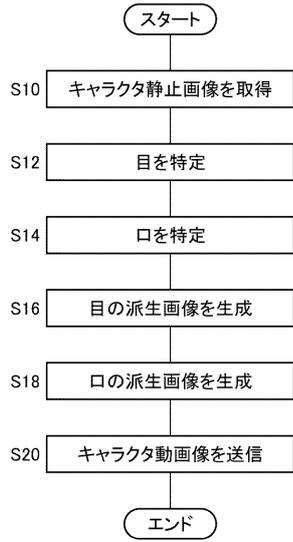
【 図 1 2 】



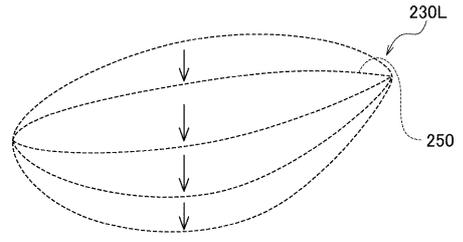
【 図 1 3 】



【図 1 4】



【図 1 5】



234

## フロントページの続き

(51)Int.Cl.	F I			テーマコード(参考)		
G 0 6 T 11/80 (2006.01)	G 0 6 T 11/80		D			
G 0 6 F 13/00 (2006.01)	G 0 6 F 13/00		6 5 0 R			

F ターム(参考) 5B084 AA02 AA12 AA18 AB07 AB13 AB31 BA03 BB02 BB15 CA07  
CB12 CB23 CE02 CE12 CF03 CF12 DB02 DB08 DC02 DC03  
DC06 DC13 DC18 EA47  
5C164 FA22 MA03S MA06S MC05P SA25S SB01P SB08P UB10S UB90P YA12