

バーチャルヒューマンを用いたアバターシステムの構築

見尾昇亮[†] 西陽一郎 保坂悠太 野崎智彦 正木裕 加藤清敬

東京理科大学

1. 結論

近年、電子メールやチャット、さらにはTV電話を用いたビデオチャット等がネットワーク上でのコミュニケーション手段として盛んに活用されている。しかし、電子メールやチャット等では、文字主体のコミュニケーションであるので、表現の限界があると考えられる。また、ビデオチャット等では、不特定多数の人とコミュニケーションをとる際には、プライバシーや匿名性の欠如といった問題点が挙げられる。

そこで、本研究では、これらの問題を考慮し、3Dのバーチャルヒューマンを用いて、人間の実環境の動きを反映させた新たなアバターシステムを提案する。特徴としては、まず一点目として、従来の2Dで表現されたアバターではなく、3Dで表現されたリアルなバーチャルヒューマンをアバターとして使用している。アバターとして動きを表現することにおいて、2Dのアバターに比べ、バーチャルヒューマンは多彩な動きをもたらすことが可能となり、リアリティの向上に効果的である。さらに、二点目としては、カメラを用いて、実環境の人間の顔の向きや口の動きなどを検出し、バーチャルヒューマンに反映することで人間と同じ動作を行うことが可能となる点である。これは、カメラから取得した画像から人間の顔領域を検出し、その領域内の目や口といった各要素の動きを解析して、バーチャルヒューマンに情報を送信する。その情報を基に動きをバーチャルヒューマンに反映させることにより、顔の動きなどを含めたアバターを用いたコミュニケーションが可能となる。

結果として、同一の環境下でアバターと自然に対話しているかのように感じさせる効果を可能にする。よって、アバターを使用した円滑なコミュニケーションの向上を実現させる。

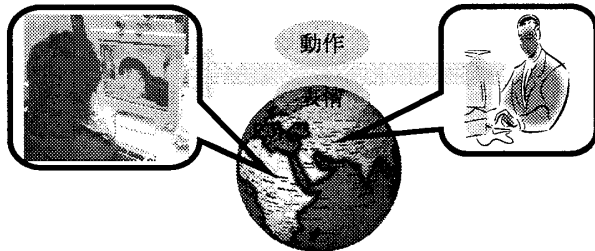


図1 本研究のコンセプト図

2. システム構成

本研究のシステム概要の構成図を図2に示す。認識部では、カメラから取得した画像を処理することにより、送信データの作成を行っている。その際、人間の顔を抽出する方法として、肌色抽出を行い、複雑な背景下においても人間の顔を認識可能としている。口の形状を検出する手法としては、画像処理することにより、口内の形状を

口の形状として、抽出し、認識を行っている。通信部では、TCP/IPプロトコルにより、認識部と描画部の通信を行っている。そして、描画部では、受信したデータを基に、3Dモデルのデータの変更、描画を行い、制御する。また、描画部では、同一環境にいるように表現するため、カメラからパソコンディスプレイ後部を撮影し、動画背景として、描画部に描画することにより、よりリアリティを高める。

図2の構成図で示すように、カメラから取得した画像を認識部で処理し、送信データを作成し、そのデータを通信部に送り、描画部へ受け渡し、描画部で受信したデータを基に3Dモデルのデータ変更・描画といった流れでシステムを構築していく。

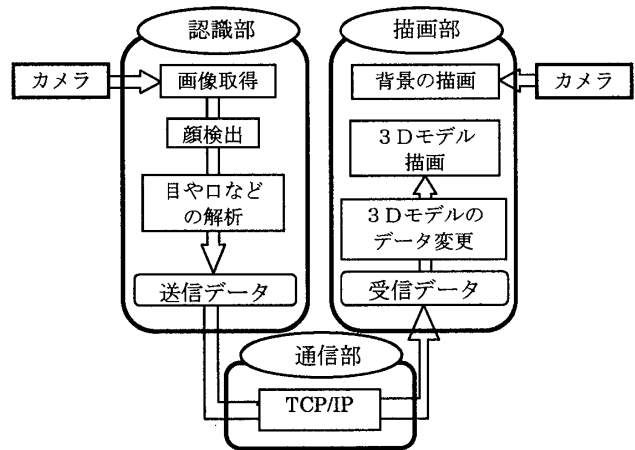


図2 アバターシステムの構成図

3. 要素技術

3-1 認識部

顔の検出部分では、カメラから取得した画像を YCrCb 情報に変換する。変換した画像の Cr 成分の画像を抽出し、2値化することにより、肌色部分、つまり、顔の領域を検出することが可能である。しかし、2値化のみでは雑音が入り込むため、平滑化処理を行い、微小な雑音を除去する。また、平滑化で除去できなかった雑音部分においては、ラベリング処理を行うことにより、面積最大部分である顔の領域を正確に抽出することが可能となる。

抽出した顔領域から口の周りの領域を選択し、その画像を HSV 情報に変換を行い、V成分を抽出し、2値化することにより、口領域を抽出することが可能である。しかし、2値化のみでは雑音が入り込むため、顔検出と同様に、平滑化、ラベリング処理を行い、雑音を除去する¹⁾。

また、顔領域を抽出した2値画像より、顔領域を面積が最小となるような矩形によって囲む。顔が正面を向いているときの、最小矩形の傾きを初期条件とし、顔が傾くことにより、最小矩形も顔と同様傾くので、顔の向きが検出される。顔がうなずく方向においては、傾きは変わらないが、最小矩形の傾きではなく、矩形形状の変化により、判別を行う。

口の動きに対しては、閉口状態と日本語単音発音時の状態を認識するために、形状特徴抽出を行う。2値化された口領域より、口の幅と高さを抽出し、そのアスペクト比により、口の状態を判別する。しかし、アスペクト比のみでは、個人差により異なる可能性があるため、面積比による判別も行う²⁾。形状特徴抽出では、計算が容易に行えるので、計算量の軽減にも最適である。

3-2 通信部

本研究では、通信プロトコルとしてTCP/IPを、TCP/IPのインタフェースとしてソケットを使用した。それによって、クライアントサーバーシステムによりシステムを構築した。クライアントサーバーシステムとは、接続する側と待機する側によって成り立つネットワークシステムである。

3-3 描面部

コミュニケーションを自然と違和感なくとることを目的とし、描面部では、バーチャルヒューマンの頭部をよりリアリティのあるものにするため、目や口などの複雑に変化する曲面に対して、必要に応じなければならない。また、人間の表情は、表情筋と呼ばれる顔の皮膚下にある筋肉の伸縮運動によって顔表面が変化し、さらにそれに伴って皺ができることで生み出される³⁾。この違和感のない表情変化を実現するには、柔軟で自由度の高い表現力があり、容易に形状を変えることができ、なおかつその処理は高速でなければならない。そこで本研究では、NURBS曲面を拡張した離散化制御点曲面⁴⁾からバーチャルヒューマンの頭部を構成する。

また、同一環境にいるように表現するため、カメラからパソコンディスプレイ後部を撮影し、バーチャルヒューマンの背景に表示させる。これにより、パソコンディスプレイの後部も表示でき、バーチャルヒューマンとの同一環境にいるように感じさせることを目指す。

4. 検証実験

本研究の動作を確認するため、システムを動作させる。二台のパソコンで通信し、リアルタイムにおいての人間とバーチャルヒューマンの動作させる。実験風景は、図4に示す。



図4 ネットワーク上でのシステム検証

また、人間の各動きに対して、バーチャルヒューマンが適切に動作しているか検証する。以下図5～7は検証実験の画像である。図の左側の人間が動き、それに対応して、右側のバーチャルヒューマンが遠隔地で通信を通して動作するか検証する。



図5 うなずいた場合



図6 右に傾いた場合



図7 左に傾いた場合

うなずいた場合、顔を傾けた場合において適切動作している様子が分かる。

5. 結論

検証実験より、カメラで撮影した顔の動きに対応して、バーチャルヒューマンの頭部も同様の動作をしていることが分かる。すなわち、カメラで撮影した顔画像の動きを解析し、認識部で作成された送信データが描面部に送られ、そのデータを基にバーチャルヒューマンが制御されているということであり、このシステムは適切に動作することが分かる。また、背景を描画することにより、より同一環境下にバーチャルヒューマンがいるように感じさせる効果が可能となった。

よって本研究のシステムにより、人間の動きに反映したリアルなバーチャルヒューマンを用いたアバターシステムを構築できた。つまり、同一の環境下でアバターと自然に対話しているかのように感じさせる効果を可能にした。

6. 今後の展望

今後は、さらに様々な人間の動きにバーチャルヒューマンに対応していく。また、本研究では、人間一人に対しての認識、解析により、バーチャルヒューマン単体を動作していたが、今後人間複数人に対しても認識、解析を行い、バーチャルヒューマンを複数それぞれの人間に対応させ、動作できるシステムを構築しようと考えている。現実の実際の環境下では複数の人間がカメラでキャプチャーされる場合も存在し、それがより自然である。複数の人間を認識し、複数のバーチャルヒューマンを描画することにより、さらに円滑に、また、さらにリアリティを向上させ、コミュニケーションの発展を目指す。

参考文献

- 1) 奈良先端科学技術大学院大学 OpenCV プログラミングブック製作チーム：“OpenCV プログラミングブック”，株式会社毎日コミュニケーションズ
- 2) 斉藤剛史，久木貢，小西亮介：“口内領域形状に基づく日本語単音の分類”，電子情報通信学会技術研究報告 vol. 106
- 3) 飯島貴志：“CGデザイナーのためのグラフィックバイブル 人体のしくみ”，ワークスコーポレーション
- 4) 加藤清敬，三上博，小山宏峰，西村敬介，川面恵司：“離散化制御点による曲面創成方法”，情報処理学会論文誌 vol. 44 No. 1 (2003)