

キャラクターフィギュアの表情変化を伴う AR コミュニケーションシステム

AR communication system using a character figure with facial expression change

吉崎 翔大†

Shota Yoshizaki

高井 昌彰‡

Yoshiaki Takai

1. はじめに

近年、拡張現実とは様々な分野で利用されており、撮影画像内の対象物を変形させるモーフィングが注目されている。一方、アニメやゲームなどの登場人物を模した玩具としてキャラクターフィギュアが社会に浸透している。本来、キャラクターフィギュアは収集・鑑賞を目的とする造形物であるが、これがカメラを向けてきた人に対して表情を変えて個々に話しかけてくるようになれば、フィギュアを起点とした新たなコミュニケーションの世界が広がるのが期待できる。そこで本研究では、このフィギュアの表情変化を顔認識と拡張現実により実現する AR コミュニケーションシステムのプロトタイプを開発した。

本稿ではシステムの概要とおおまかな処理の流れ、およびフィギュアの顔認識と変形処理の結果について述べる。

2. システム概要

本システムが対象とするフィギュアは 2.5 頭身程度で、両目が開き、眼鏡をかけていないものと仮定する。

本システムはフィギュア及びユーザの情報を管理するサーバと、カメラ入力映像からフィギュアの表情付加を行うクライアントで構成される。

発話文はデータベースにキャラクターごとに登録されたものからランダムに選択される。表情の基本となる目の形状として“閉”、“笑”、“ジト目”を定義している。これらの基本形状と動作タイミングから、“長い瞬き”、“瞬き”、“短い微笑み”、“長い微笑み”、“ジト目状態継続”の 5 種類を目の表情モーションを形成する。口の形状はア行の 5 音と“開”と“閉”を定義している。

2.1 システム処理の流れ

本システムの構成概要を図 1 に示す。ユーザは使用するフィギュアのキャラクター名を予めサーバに登録しておく。はじめに、クライアントがカメラ撮影画像からフィギュアを認識すると、サーバからキャラクター名の一覧が取得されキャラクターが選択される。またクライアントでのモーフィング処理のため、初期化時に口の開閉状態をサーバで識別してクライアントに返す。

初期化後クライアントはカメラ入力映像から顔位置と顔パーツの認識を行い、サーバから取得した発話文や表情に応じたモーフィング処理を行う。また、ユーザの指やペン先を画像から認識し、フィギュアとの相対位置をもとに、キャラクターが怒ったり喜んだりするといった表情変化を伴ったインタラクションが可能である。

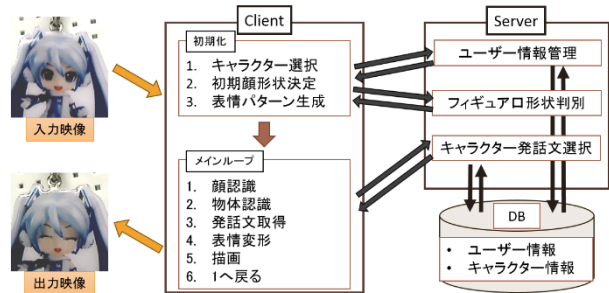


図 1 システム構成

2.2 実装環境

サーバはクライアントからの要求を待つ TCP サーバプログラムである。開発・実行環境は Windows、開発言語には C# と .NET Framework を用いた。データベースには MongoDB を用いた。またクライアントの実行環境は Windows であるが、開発言語には C++ を使い、カメラ映像の取得や画像処理・顔認識に OpenCV、描画 API に OpenGL を用いた。クライアント画面の詳細を図 2 に示す。



図 2 クライアント画面

3. フィギュア識別関連処理

本章ではフィギュアの顔に関する画像認識処理について述べる。フィギュアの顔位置識別と顔パーツ検出処理はそれぞれクライアントで行われる。

3.1 顔位置認識

フィギュアの顔の位置特定には OpenCV に実装されているカスケード分類器を用いた。本システムでは LBP 特徴量による識別器を用いた。識別器の作成はフィギュアの顔を含む画像及び画像内の顔位置情報とフィギュアの顔を含まない画像を用意することで OpenCV に付属のプログラムで作成することができる。本システムではフィギュアの顔を含む画像を 4,000 枚程度、フィギュアの顔を含まない画像を 7,000 枚程度用いて学習を行った。

† 北海道大学大学院情報科学研究科, Graduate school of Info. Sci. Technology, Hokkaido University

‡ 北海道大学情報基盤センター, Information Initiative Center, Hokkaido University

3.2 顔パーツ認識

本システムではフィギュアのみや口といった顔パーツを認識するために Constrained Local Models (CLM) [1]を用いた。このアルゴリズムは顔を頂点列で表現し、各頂点を特徴点として顔画像から位置を探索する。探索後の頂点列を、学習した頂点列形状の固有ベクトルをもとに補正することで顔パーツの位置及び輪郭を取得する。特徴点の探索の手法はいくつか提案されているが、本システムでは線形ソフトマージン SVM を用いた。

顔頂点列の定義を図 3 に示す。フィギュアの口の見えが閉じているものと開いているものとで顔頂点列が大きく異なるため、顔モデルを二種類作成した。学習にはフィギュアの顔を含む画像及び顔頂点座標情報とフィギュアの顔を含まない画像が必要である。口の開閉状態によりそれぞれ 100 枚程度のフィギュアの顔画像を用いた。顔頂点数は口が閉じている場合は 26 個、口が開いている場合は 31 個である。

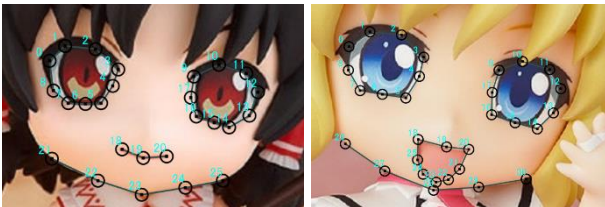


図 3 フィギュアの顔頂点定義

左図：<http://www.goodsmile.info/ja/product/2568/ねんどろいど+博麗霊夢.html> より引用・加工

右図：<http://www.goodsmile.info/ja/product/3666/ねんどろいど+アリス+マーガトロイド.html> より引用・加工

4. フィギュアインタラクション処理

4.1 物体追跡

本システムはフィギュアとのインタラクション方法として OpenCV に実装されている CamShift 法を利用した物体追跡を用いる。追跡物体の初期化においては、ユーザが物体の初期位置を明示的に指定する必要がある。追跡対象はフィギュアに含まれている色と遠い、細長い物体が望ましい。

4.2 フィギュア反応決定

物体の先端座標とフィギュアの顔との相対的な位置関係でフィギュアの反応が決定される。本システムでは“撫でる”に相当する動作として物体を顔に近づけるとキャラクターが喜ぶ表情を見せる。また、“いたずら”に相当する動作として物体を目に近づけると怒る表情を見せる。

フィギュアが喜びや怒りの表情をとる際にはそれぞれに対応する発話文を示す。それ以外の場合にはランダムな発話文と表情を見せる。各状態における発話文はキャラクター名ごとに全てデータベース内で定義されている。

5. フィギュア表情変形と描画処理

本システムでは CLM で取得した顔パーツの頂点列からモーフィング処理のためのメッシュを生成して変形することで表情変化を実現する。顔パーツのメッシュを図 4 に示す。同図左は口を閉じたフィギュアのメッシュ、右は口を開いたフィギュアのメッシュである。目と口について、そ

れぞれの輪郭の内側と外側で異なるメッシュを構築して輪郭の内部と外部を分離することで、図 5 に示すようなレイヤー構造を形成し、顔パーツ画像の局所的な変形を行う。

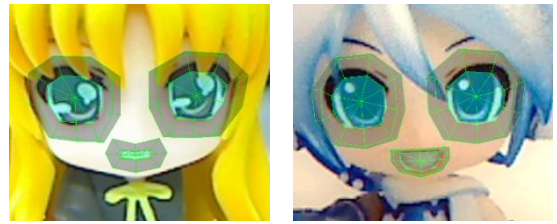


図 4 フィギュアの顔パーツメッシュ

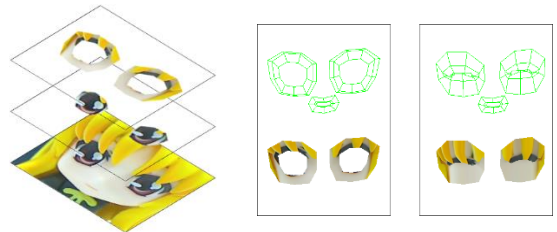


図 5 レイヤー構造と変形例

6. 動作結果

フィギュアの顔認識は照明や撮影角度の影響を受けて失敗する可能性があるが、ここでは顔認識が成功した状態の変形結果を示す。図 6 は目の変形結果である。左から“オリジナル”，“瞬き”，“笑み”，“ジト目”の表情である。図 7 は口の変形結果である。左がオリジナル，上段左からア行の 5 音，下段が“閉”と“開”の状態である。動作速度は Xeon (3.7GHz) CPU 上で約 30fps を達成している。



図 6 目の変形結果



図 7 口の変形結果

7. おわりに

本稿では顔認識と拡張現実を利用してフィギュアに表情を持たせた AR コミュニケーションシステムのプロトタイプについて述べた。認識精度及び変形品質の向上、インタラクションの拡充が今後の課題である。

参考文献

- [1] D. Cristinacce, T.F. Cootes, “Feature Detection and Tracking with Constrained Local Models”, Proceedings of the British Machine Vision Conference, Vol.3, pp.929-938 (2006).