

キャラクターフィギュアに表情変化と発話動作を付加する拡張現実システム

吉崎 翔大† 高井 昌彰‡

北海道大学大学院情報科学研究科† 北海道大学情報基盤センター‡

1. はじめに

アニメや漫画などの登場人物を模したキャラクターフィギュアが社会に浸透している。キャラクターは登場作品のシーンの中で多様に表情を変化させているが、フィギュアはそのキャラクターの性格や振舞いをよく反映させるような表情に固定されている。このフィギュアが現実世界で目を向けてきた人に対して表情を変えながら話しかけるようになれば、フィギュアを介した仮想のキャラクターとの新たなコミュニケーションの世界を広げることができる。

そこで、本研究ではフィギュアのリアルタイムな表情変化を顔認識とモーフィングを組み合わせた拡張現実(AR)で実現するシステムのプロトタイプを構築した。カメラを向けて覗くことで、フィギュアが目の瞬きなどの表情変化や発話内容と同期した口形状変化を行う様子が見られる。さらに、フィギュアの頭をなでる動作で表情変化を誘起させるインタラクションも可能である。

本論文ではプロトタイプシステムの構成と実装および動作結果について述べる。

2. 関連研究

目印となるマーカ画像の位置に3DCGのキャラクターを表示して操作する技術はARToolKit, VuforiaなどのライブラリやUnityといったフレームワークにより容易に利用することが可能になっており、フィギュアとARを組み合わせたコンシューマ製品はこれまでに数多く発売されてきている。しかしそのほとんどはフィギュアの近くに所定のマーカを配置して、予め設定されたCGコンテンツを重畳表示するものである。

また、フィギュアにプロジェクションマッピングを行い、顎の影及び前髪によって生じる額への影の変化を疑似的に生じさせることでフィギュアの頷きと首振り動作の表現を試みた研究[1]があるが、フィギュアの表情そのものを視覚的に変化させるものではない。

3. システムの要件

本研究で構築するシステムの必須要件を3つ設定した。第一の要件は実体を持ったフィギュアを起点としたAR表示を行うことである。3DCGを重畳するだけの簡易なARでは、キャラクターに直接触れるといった触覚の欠如に加え、実世界の光源環境とテクスチャの質感の再現など、光学的整合性の実現に問題がある。カメラで映し出された実物のフィギュア画像をもとにすることで、これらの問題の解消が考えられる。

第二の要件はリアルタイム性を確保することである。カメラからの入力を即座に反映させることはARアプリケーションに対する重要な要件の1つであり、特にインタラクションにおいては欠かせない要素である。

第三の要件は利用者のフィギュアに対する直接的なアクションを可能にすることである。すなわち、キャラクターが映し出されているディスプレイをタッチすることやその他のデバイスからの入力に因らないことである。実世界のフィギュアに対する直接的接触でフィギュアの表情を変化させることにより、リアリティのあるインタラクションを実現する。

4. システムの設計・構成

本システムはリアルタイムにAR表示を行うクライアントと、発話内容やフィギュアの種類などを識別するサーバから構成される。サーバはキャラクターの発話文を保存したデータベースであると同時に発話文の選択処理を行う。また、発話文選択のためにフィギュアのキャラクター識別の機能を有する。一方、クライアントは入力装置としてカメラを、出力装置としてディスプレイとスピーカを有するデバイスである。図1にシステムの構成を示す。システムによりフィギュアの顔が変形し、発話文が音声と画面へのテキスト表示で出力される。



図1 システムの構成と入出力

撮影対象: 「SNOW MIKU 2014 雪祭初音ミク」入場記念チャーム
初音ミク ©Crypton Feature Media, Inc.

Augmented reality system to make face expression change and lip sync of a character figure

†Yoshizaki Shota, Hokkaido University, Graduate School of Information Science and Technology

‡Takai Yoshiaki, Hokkaido University Information Initiative Center

目の表情として、閉じる、笑み、ジト目の3種類を備えている。また、口形状のパターンは母音の5音の他、閉じる、開く、の計7種類である。

目の表情はシステムによってランダムに選択されるが、口形状は発話文と同期するように母音を選択して変化する。発話文はキャラクターごとに登録されたデータベースから選択される。

5. システムの実装

5.1. 実装環境

サーバは C# で実装した。またクライアントは画像処理ライブラリに OpenCV、描画 API に OpenGL を用いて C++ で実装した。サーバとクライアントは TCP 通信で接続される。

5.2. 表情変形

表情変形は以下に示す3段階で実行される。

- ① 目と口の輪郭を形成する顔パーツ頂点取得
- ② 顔パーツ頂点から二次元メッシュ生成
- ③ メッシュ変形

5.3. フィギュアの顔認識

システムはカメラ画像からフィギュアの顔位置を取得する必要がある。これには OpenCV に実装されている LBP 特徴量を用いたカスケード型分類器を使用した。フィギュアの顔データを収集して学習することで分類器を作成する。

5.4. フィギュアの顔パーツ認識

顔領域から目、口、顎の輪郭を取得するために CLM(Constrained Local Models)を利用する。CLM は顔パーツを、パーツの輪郭線を構成する頂点列で定義する。まず顔領域の画像内で各頂点の最も適した位置の探索を行う。この探索処理は短時間で終了させるため、結果の精度は必ずしも十分ではない。そのため学習済みのパーツ形状をもとにした補正処理を次に行う。各頂点での探索には線形 SVM を用いる。パーツ形状の学習は PCA による多数のフィギュアの顔パーツ形状の特徴抽出である。

5.5. 顔パーツメッシュ生成・変形

図2にメッシュの生成例を示す。メッシュは輪郭の内側と外側にそれぞれ生成される。外側のメッシュのみを変形して内側のメッシュを覆うことで、まぶたと唇の変形を表現する。

5.6. インタクション処理

撮影画像内の特定の物体を追跡し、フィギュアに対するアクションとして認識するため、

CamShift法を用いた。これにより撮影画像内の特定の色領域を追跡する。ユーザは追跡対象物体として指先やペン先を指定する。追跡対象物体とフィギュアの顔との相対位置関係によってフィギュアが予め定められた反応を示す。本システムでは、撫でるに相当する動作として頭に物体が近づくと喜び、いたずらに相当する動作として目に近づけると嫌がる動作を登録している。

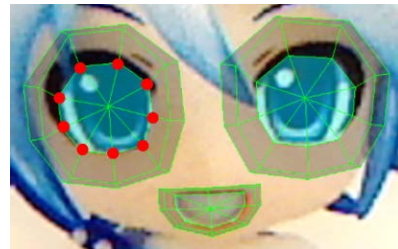


図2 顔パーツメッシュ例 (赤点:右目取得頂点)
撮影対象:同図1

6. 実行結果

本システムは Web カメラの入力を解像度 640×480 ピクセルに設定した状態で、Core-i5 6200U を搭載したノート PC にて 30fps を実現した。

表情変形にともなう顔パーツ形状の変化を図3に示す。



図3 顔パーツ変形結果

撮影対象
(上):ねんどろいど ぶち 魔法少女リリカルなのは The MOVIE 1st
フェイト・テスタロッサ (私服)
フェイト・テスタロッサ ©NANOHA The MOVIE 1st PROJECT
(下):同図1

7. 今後の課題

ユーザのシステム利用時の負担を軽減するためにはスマートフォン等のモバイル端末での実装が課題である。そのために認識処理の更なる高速・軽量化が必要である。また、実際に音声が付加して発話動作を行った場合のユーザ評価を求めるアンケート調査を行う予定である。

参考文献

- [1] Yoshihisa Ishihara, Kazuki Kobayashi, and Seiji Yamada. 2016. Behavioral Expression Design onto Manufactured Figures. (HAI '16). ACM, New York, NY, USA, 243-244.