

AAM を利用した表情の模倣

福井 竜一[†] 桂田 浩一[†] 入部 百合絵[†] 新田 恒雄[†]

[†]豊橋技術科学大学 大学院工学研究科

1. はじめに

映画やテレビゲームに向け、キャラクターの表情を CG 合成するソフトが数多く開発されている[1][2]. こうしたソフトでは、表情をモーションキャプチャした上で、CG 技術によりユーザの表情を模倣させるものが多い. 本研究では、CG による表情模倣ではなく、コンピュータ内の一枚の顔画像をユーザの表情に合わせて変える (模倣する) 方法を提案する. 表情を模倣するには、まずユーザの表情を認識した後、コンピュータ内の顔画像に表情を付与する必要がある. 従来の研究[3][4]では、顔の特徴点 (目・眉・口等) を検出した後、特徴点を顔モデル上で移動させ、表情を模倣するものが多い. しかしこの手法では、眉間のしわといった顔表面の変化を模倣することができない. そこで我々は、Active Appearance Model (以下 AAM) [5]を利用して、顔の細部を含めたより詳細な表情を模倣する手法を提案する.

AAM は、顔の表情や個人性、照明状況などを複合的に表すパラメータ群からなる顔画像合成モデルである. パラメータを変動させることで様々な顔画像を合成できるが、AAM には表情以外の成分も含まれるため、パラメータの変動によって個人性なども変化する. すなわち表情を模倣するには、表情のみを示す成分を取得する必要がある. そこで本研究では、ある人物の表情変化に対応するパラメータ変動を集め、これらを主成分分析することで表情成分を抽出し、これを利用して表情模倣を実現する.

2. AAM を利用した表情の模倣

2.1. AAM の構築

AAM は、画像の特徴点情報と輝度情報をそれぞれ主成分分析して得た形状モデル、輝度モデルを、さらに主成分分析して得た顔画像合成モデルである. 以下では AAM の構築法について述べる.

まず AAM を構築するトレーニングデータとして、各顔画像に数十点の特徴点を付与する. これら特徴点のベクトル x を正規化した上で主成分分析し、形状モデルを構築する. 続いて、顔画像から輝度情報 g を取得し、これも正規化した上で主成分分析し、輝度モデルを構築する. 最後に、形状モデルと輝度モデルを統合して主成分分析すると、特徴点のベクトル x と輝度情報 g は次に示す(1), (2)式のように表される.

$$x = \bar{x} + Q_s c \quad (1)$$

$$g = \bar{g} + Q_g c \quad (2)$$

\bar{x} , \bar{g} はそれぞれ平均形状、平均明度で、 Q_s , Q_g はパラメータ c から形状、輝度に変換する行列である. 式中のパラメータ c を変動させることで様々な顔画像が合成可能になる.

Facial expression mimic using AAM

Ryuichi Fukui[†], Kouichi Katsurada[†], Yurie Iribe[†], Tsuneo Nitta[†]

[†] Toyohashi University of Technology

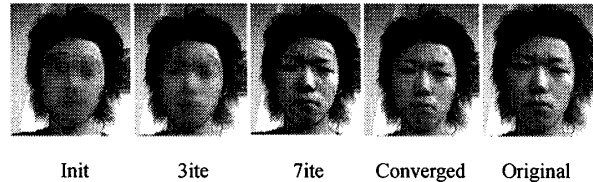


図1 フィッティング操作

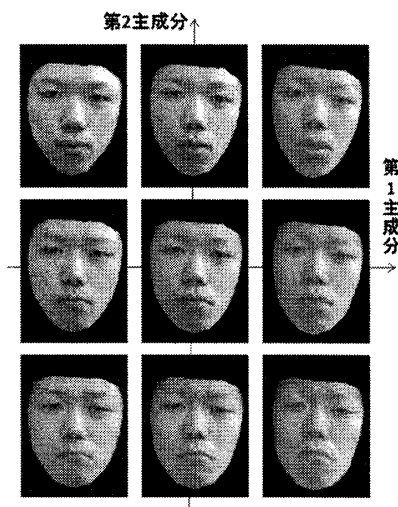


図2 表情パラメータを変動させた時の表情変化 ($-\sigma \sim \sigma$)

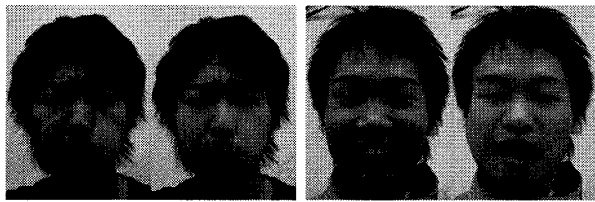
2.2. フィッティング

未知の顔画像に対して AAM を適用して近似画像を合成することで、その顔画像を生成するパラメータを探求することができる. この操作をフィッティングと呼ぶ. パラメータ探索には、最急降下法を用いた. また、最適化に用いる誤差関数として、対象の画像と合成した画像との輝度の差を用いた. 図 1 に、ある顔画像に対してフィッティング操作をした例を示す. 最初のステップではガウスフィルタをかけて AAM で合成した画像をぼかしている. これは、誤差関数を滑らかにし、局所解に収束しにくくするためである.

2.3. 表情成分の抽出

通常の AAM ではパラメータを変動させると、表情変化だけではなく、個人間で異なる骨格や筋肉のつき方も変化する. これは各パラメータが個人性や表情を変化させ得る要素を複合的に持つことに依る. このため、顔が持つ表情情報のみを得ることができず、表情模倣が困難になる. そこで、AAM のパラメータから表情変化のみを表す成分を抽出する方法を検討する.

まず、ある人物の平常顔を合成するパラメータをフィッティングから求める. 続いて、同じ人物の別の表情顔を合成するパラメータも同様に求め、平常顔のパラメータとの差分をとる. この差分は、その人物の表情の動き



原画像 合成画像 原画像 合成画像
(輝度誤差 8.6 形状誤差 6.5) (輝度誤差 13.0 形状誤差 5.0)

図3 合成画像の例

を表現している。こうした表情の動きを収集し、主成分分析することで、個人間で共通な表情の動きを抽出する。このときの主成分得点ベクトルを表情パラメータと呼ぶ。表情パラメータを変動させ、顔画像を合成した様子を図 2 に示す。

2.4. 表情の模倣

ある人物の平常顔と表情顔の差から表情パラメータを求め、その表情パラメータを他人の平常顔パラメータに加えることで、元の人物と同じ表情の顔画像を合成できる。この手法を利用して表情模倣を実現する。

3. 評価実験

3.1. 実験 1：抽出した表情成分の検証

正しく表情成分が抽出できていることを確認するため、表情画像から表情パラメータを算出し、算出したパラメータをもとに合成した表情画像と原画像を比較した。実験用の画像としてソフトピアジャパン顔画像データベース[6]収録の顔画像 300 枚と研究室の男性 11 人の表情画像各 30 枚を用いた。画像はすべて 640×320 のグレースケール画像(256 階調)である。まず、モデルの構築にはソフトピアジャパン顔画像データベースの 300 枚と研究室の男性 10 人分、表情成分の抽出にはモデル構築に利用した研究室の男性 10 人分のデータを利用し、残りの 1 人分のデータをテストデータとして交差確認法で評価した。

実合成した表情画像と原画像における輝度の平均二乗誤差、および特徴点座標の平均二乗誤差を表 1 に示す。表 1 に示す通り、輝度値の平均二乗誤差が 8.25、座標の平均二乗誤差が 5.38 となった。図 3 左に示すように誤差が平均値付近のものは画像の合成に成功しており、表情成分がうまく抽出できていることが確認できた。一方で、図 3 右の画像のように合成に失敗しているものも見受けられた。特に、口を開いている表情画像ではうまく合成できないものが多かった。これは、口内の輝度値が歯並び等に大きく左右され、個人ごとの差が大きいためが原因と考える。改善案として、個人ごとに異なる口内モデルを作成することなどが考えられる。

3.2. 実験 2：模倣精度の検証

提案手法を用いて表情を模倣したときの精度をアンケート調査から検証した。実験では、表情パラメータを他の人物に適用して得た合成画像について二つのアンケートを行った。アンケート 1 では同じ表情をしていたかど

表 1 実験 1 の結果

	平均値	標準偏差
輝度の平均二乗誤差	8.25	1.83
形状の平均二乗誤差	5.38	1.54

表 2 アンケート結果

	平均値
アンケート 1	3.46
アンケート 2	3.57

うかを、アンケート 2 では、合成画像とその人物が同じ人物かどうかを、被験者に 5 段階 (5 が最高) で答えさせた。アンケート用の画像には、実験 1 で使用した研究室の男性 10 人の表情画像から無作為に 5 枚ずつを選んだ。この 50 枚の画像を 25 枚ずつに分け、8 人の被験者にアンケートを依頼した。

結果を表 2 に示す。それぞれのアンケートから概ね良好な結果が得られたことが分かる。しかし、一部の画像に関しては良い結果が得られていない。これは、実験 1 で誤差が大きかった画像も含まれていたためである。実際に、実験 1 で輝度の誤差が平均値より小さい画像に限ると、アンケート 1 は平均 3.7、アンケート 2 は平均 3.9 と良好な結果が得られた。一方、実験 1 で形状の誤差が平均以下のものについては、おおよそ表 2 の結果と同様であり、形状の誤差が表情の模倣精度に与える影響が輝度ほど大きくないことが分かった。このことから、模倣精度を向上させるには輝度の誤差を小さくする必要がある。

4. おわりに

本稿では、主成分分析で抽出した表情成分を利用した表情の模倣法を検討した。提案手法を適用することで、多くの表情に対して模倣が成功することが確認できた。一方、一部の表情については十分に模倣できず、特に口が開いている表情に関して良い結果が得られなかった。また、表情の模倣精度の向上には輝度の誤差を小さくすることが重要なことがわかった。今後、これらの課題に対処していきたい。

参考文献

- 1) Pendulum's AlterEgo, <http://www.studiopendulum.com/alterego/index.html>
- 2) Mova contour, <http://www.mova.com/>
- 3) Breazeal, C., Buchsbaum, D., Gray, J., Blumberg, B.: Learning From and about Others: Towards Using Imitation to Bootstrap the Social Competence of Robots, Article Life, vol. 11(2005)
- 4) Ghys, C., Taron, M., Paragios, N., Basclé, B., Komodakis, N.: Expression Mimicking: from 2D Monocular Sequences to 3D Animations, Advances in Visual Computing: Third International Symposium, ISVC 2007, Proceedings, pp621-630(2008)
- 5) Cootes, T.F., Edwards, G.J., and Taylor, C.J.: Active Appearance models, IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6), pp681-685(2001)
- 6) ソフトピアジャパン顔画像データベース <http://www.softopia.or.jp/rd/facedb.html>