

# 個性を反映した表情生成手法の構築

～ 顔パーツの特異的な動きの抽出 ～

An Expression Generation Using Characteristic Movements of Users:  
Extraction of Characteristic Movements in Facial Components

山平 慎吾      羽倉 淳      藤田 八ミド

岩手県立大学    ソフトウェア情報学部    ソフトウェア情報学科

**Abstract** 本研究では、擬人化エージェントに対し、個人の表情変化の特徴を反映した表情生成手法の構築を目指し、表情におけるパーツの特異的な動きを抽出する手法の構築を試みる。ここではまず、表情譜と呼ばれる顔の各観測点の座標変化とパーツ(眉、目、口など)の運動のタイミング構造のデータ、を基に平均表情モデルを作る。その上で擬人化対象となる人の表情と平均表情モデルとを比較することで、その人の個性を反映した表情変化の特徴を抽出するための手法を提案する。

## 1 はじめに

人間が個人の特定をするときは視覚的な特徴(顔、身体、身振り)認識が大きく寄与していることが知られている。これらの視覚的特徴のうちの1つ、或いは、いくつかの特徴が類似した人を見たとき似ていると感じる。中でも、顔の表情は一目で特定の個人を断定できる程、個人の特徴を反映したものであるといえる [3]。

一方、HCI に対する要求の高まりとともに、人間とコンピュータの親和性を高めるコンピュータ(を利用した)エンターテインメントへの要求も高まってきている。これらのなかに、3D スキャナーで人の顔情報を取り込み映像作品の登場人物の顔を作成するシステムがある[4]。このシステムにより、視聴者はこの作品そして作品内の登場人物により親近感を持つことができることが示されている。

しかしこのシステムではデータを採った人の顔の形状は再現できるが、顔の動きは決まりきった動きしか再現していないため、利用者は違和感を覚えることも言及されている。したがって個々の顔の動きの特徴を反映させることによって、利用者の親しみがより深まると考えられる。

一般に行動認識はパターン認識問題として扱われ、コンピュータによる人間の動作認識は、人間の動きを計測しあらかじめ定義された行動にマッチングしていくという方法が多く用いられる(例えば、[1]など)。そのためこれらの方法を用いることで、行動自体は認識可能である。しかし、行動の中に存在している個人特有の動きは除外されてしまう。表情変化の個性を認識するためには、個人特有の動きも考慮する必要があると考えられる。

そこで本論文では、被験者の表情変化の個性を反映した表情生成手法のための、被験者の表情変化の特異的な動きを抽出する手法を提案する。ここでは表情変化の特異的な動きを多くの人にみられる一般的な顔の動きとの差であると捉える。そ

こで、まず多数の被験者により、特定の表情を演じさせ、その際の顔の観測点の動きをモーションキャプチャーシステムを用いて、時系列データとして取得する。その上で、これらの時系列データから表情譜[2]を作成し(図1)、表情譜内のモードを無表情、目標表情、無表情から目標表情への推移、目標表情から無表情への推移の4つに分類することにより、被験者間に共通する動きの遷移の抽出を試みる。さらに1つのモードの平均動作時間とフレームごとの前フレームとの差分を定義し表情変化の平均値である平均表情モデルを作る。その上で、擬人化エージェントのモデルとなる個人の各フレームの差分と平均表情モデルの各フレームの差分を比較することにより、表情変化の個性を抽出する。ここでは、まず、本稿で必要となる範囲で表情譜を概説した後、平均表情モデルについて説明し、平均表情モデルの構築方法、及び、表情変化における個性の抽出方法について詳述する。

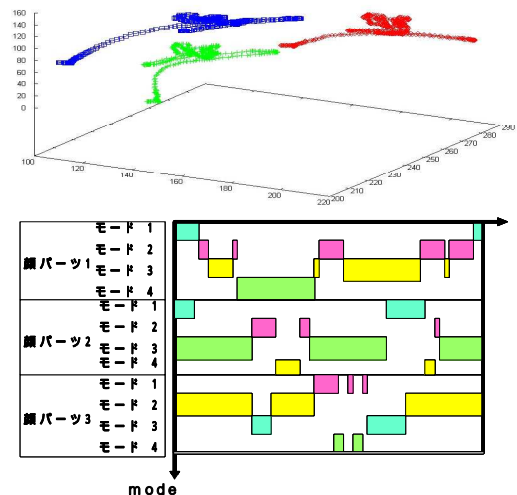


図1 表情譜の獲得 [2]

## 2 表情譜

表情認識の方法として、カメラから取得した顔画像の時系列データから表情を認識する手法として表情譜と呼ばれる手法が提案されている[2]。

そこでは、顔を独立した動きが観測される領域（顔パーツ：目、口、鼻など）に分け、表情変化を各顔パーツの動的な運動の組み合わせと考え、顔パーツの変化を静止状態もしくは収束性の運動を行う事象の時間範囲をモードという単位で表す。その上で、モード内の遷移が似ているモード同士を階層型クラスタリングによって同定しモードの数を減らしていく（図 2）。最終的にモード間の発生タイミングを用いることで、表情をより詳細に理解・分析する事が可能であることが示されている。

この手法は人間の表情変化を予め定められたカテゴリにマッチングを行わず、顔パーツの動きの時系列からボトムアップにモードを自己組織化して行き、モードを定義し表情を認識するため、他の表情認識手法よりも詳細に表情を理解・分析できると考えられる。

この手法を用い、次章では平均表情モデルについて説明する。

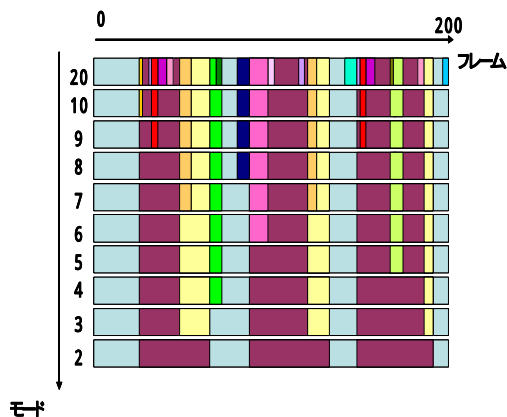


図 2 クラスタリングによるモード形成の様子[2]

## 3 平均表情モデル

ここでは、提案手法の基礎となる平均表情モデルについて述べる。

表情の変化は各顔パーツの動きによって構成されると捉え、表情変化において多くの人々に一般的に見られる顔パーツの動きのことをここでは平均表情モデルと呼ぶ。

まず表情譜内の顔パーツ毎のモードを無表情、目標表情、無表情から目標表情への推移、目標表情から無表情への推移の 4 つに分類することにより、被験者間に共通する動きの遷移を抽出する。そして各顔パーツの 1 つのモードの動作時間の平均値と、モード内の前フレームとの差分の平均値を一般的に見られる顔パーツの動きとして扱う。

ここで擬人化エージェントのモデルとなる個人のフレームごとの差分と平均表情モデルとのフレーム毎の差分を比較し個性を抽出する時、同じ差分値が出てくる可能性は低く、全てのフレームが

個性的な動きとして扱われるという問題が考えられる。そのため平均表情モデルのフレームごと差分における予測される誤差を求め、一般的な動きの範囲を定義する必要がある。この予測誤差としてフレーム毎の差分値の標準偏差を用いる。そして差分 ± 標準偏差の値の範囲を一般的な動きの範囲とする。

したがって、平均表情モデルはモードの平均動作時間とフレームごとの前フレームとの差分の平均値と標準偏差によって構成されている。

そして次章では、この平均表情モデルの作成手法と、個人の表情変化の特徴の抽出方法について述べる。

## 4 個人の表情変化の個性抽出手法

ここでの表情変化における個性とは、平均表情モデルと特定の個人の表情変化における顔パーツの動きとの違いを意味する。本節では、表情変化の個性を抽出するために必要な、顔パーツごとの動作抽出方法、平均表情モデルの構築方法、及び、これらを用いた個性の抽出方法について述べる。

### 4.1 モーションキャプチャーを利用した顔パーツの時系列データの取得

被験者から各顔パーツの動きの時系列データを取得する。ここでは、顔パーツの時系列データの取得にはモーションキャプチャーを利用する。モーションキャプチャーの利用により各顔パーツの動きを直接時系列データで入手でき、[2]で行われている画像処理の手間を省く事ができる。

また、表情変化における個性を抽出するためには顔パーツの大きな動きではなく細かな動きも計測する必要がある。そのため、モーションキャプチャーの利用により、[2]でおこなわれているような動画からの時系列データよりも詳細なデータが入手できると考えられる。そのため、ここで目的とするような表情変化の個性を認識するためにはモーションキャプチャーによる動作取得が有効だと考えられる。モーションキャプチャーを利用した各顔パーツの時系列データ取得の様子を図 3 に示す。

ここで、マーカーの位置は各顔パーツの輪郭を捉えるようにする。そのためにマーカーの数は、各眉 3 点、各目 4 点、鼻 7 点、口 8 点、顔の下半分の輪郭に 9 点としフレームレート 60fps で計測した。

### 4.2 顔パーツごとの動作抽出方法

次に、顔パーツごとの動作抽出方法について述べる。ここでは顔パーツの動きとして、現在のフレームの前フレームとの差分を使用する。そして被験者  $i$  の顔パーツ  $a$  の時刻  $t-1$  から時刻  $t$  への  $x$  座標の差分  $Sx$  と  $y$  座標の差分  $Sy$  の値は次式で求められる。

$$Sx_t^{(i,a)} = x_t^{(i,a)} - x_{t-1}^{(i,a)} \quad (1)$$

$$Sy_t^{(i,a)} = y_t^{(i,a)} - y_{t-1}^{(i,a)} \quad (2)$$

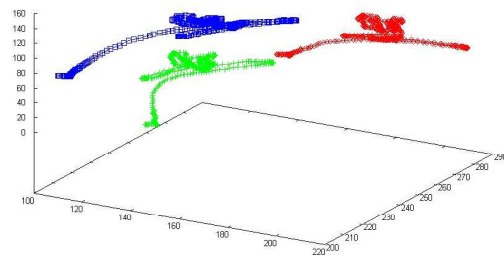
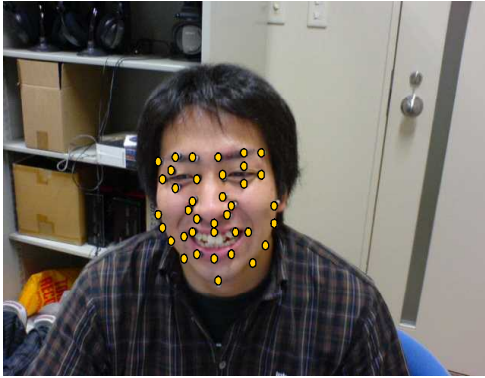


図3 顔の特徴点とモーションキャプチャーを利用した時系列入手例

#### 4.3 平均表情モデル作成手法

まず、モーションキャプチャーにより取得したデータから多くの被験者から表情変化の時系列データを集め、表情譜を作成する。しかし平均表情モデルを作る際には多くの人から表情変化の時系列データを集め表情譜を作成する必要があり、表情譜は時系列からボトムアップにモードを自己組織化していく手法であるため、時系列によって自己組織化の違いが生じる可能性がありモードの定義の違いが出てしまう。この問題点の解決するため、各被験者の時系列データを 1 つの時系列データとして扱い、同時に表情譜を作成する。この方法によって、平均表情モデル作成のための被験者から取得した時系列は、モードの自己組織化が 1 回の表情譜作成で行われるため、被験者それぞれの表情譜におけるモードの自己組織化に共通性を持たせる事ができると考える。そして今回は表情譜内のモードを、無表情、目標表情、無表情から目標表情への推移、目標表情から無表情への推移の 4 つに分類し、被験者間に共通する動きの遷移の抽出を試みる。次にモードのフレーム毎に差分の平均値と標準偏差を計算する。今回は動作の個性を抽出するため、動作部分のモードである無表情から目標表情への推移、目標表情から無表情への推移の平均動作時間とモード内のフレームごとの差分の平均値として標準偏差を計算する。

そしてモードの平均動作時間のフレーム長に一致させるため、被験者の行動モードのフレーム長が平均動作時間より長い場合は差分の少ないフレームを余ったフレームの個数分取り除き差分を再

計算しモードのフレーム長を平均表情モデルのモード動作時間に合わせる。また被験者の行動モードのフレーム長が平均動作時間より短い場合は差分の少ないフレームから順に二等分しその間の差分を再計算し、足りないフレームの個数分この作業によりフレームを補充する。ここでは表情変化の動きはフレーム間で大きな差が生じないと仮定しこの方法によりフレーム長を合わせる。前述の作業により各被験者の行動部分のモードのフレーム長が一致したためフレーム毎に平均差分と標準偏差を求め、これを平均的な表情変化の差分とする。ここで平均表情モデル作成のための被験者の総人数を  $N$ 、被験者  $i$  の顔パーツ  $a$  の時刻  $t-1$  から時刻  $t$  への  $x$  座標の平均差分  $Sx_t$  と  $y$  座標の平均差分  $Sy_t$  の値は式(1),(2)を使い次式で求められる。

$$Sx_t^{(a)} = \frac{1}{N} \sum_{i=1}^N (Sx_t^{(i,a)}) \quad (3)$$

$$Sy_t^{(a)} = \frac{1}{N} \sum_{i=1}^N (Sy_t^{(i,a)}) \quad (4)$$

そしてフレームごとの標準偏差を求める。時刻  $t$  における  $x$  座標の分散  $Vx$  と  $y$  座標の分散  $Vy$  は次式で表すことができる。

$$Vx_t^{(a)} = \frac{1}{N} \sum_{i=1}^N (x_t^{(i,a)} - Sx_t^{(a)})^2 \quad (5)$$

$$Vy_t^{(a)} = \frac{1}{N} \sum_{i=1}^N (y_t^{(i,a)} - Sy_t^{(a)})^2 \quad (6)$$

式(5)、(6)から時刻  $t$  における  $x$  座標の標準偏差  $Hx_t$  と  $y$  座標の標準偏差  $Hy_t$  は 次式で表される。

$$Hx_t^{(a)} = \sqrt{Vx_t^{(a)}} \quad (7)$$

$$Hy_t^{(a)} = \sqrt{Vy_t^{(a)}} \quad (8)$$

#### 4.3 表情変化の個性抽出手法

表情変化の個性はこの平均表情モデルとの比較をすることにより抽出する。

始めに表情変化の個性を抽出する擬人化エージェントのモデルとなる被験者の表情譜は、平均表情モデルの表情譜と同じモード数が 4 個になるように階層型クラスタリングでモードの定義をしていく。次に被験者の表情譜の無表情から目的の表情へ、目的の表情から無表情へ、という行動部分のモードの行動時間を抽出する。そしてこのモードの行動時間のフレームごとの前フレームとの差異を式(1)、(2)により求め、求めたフレームごとの差異と平均表情モデルの差異を行動開始フレームから順に比べる。表情変化の個性を抽出したい被

験者の差異が平均表情モデルの差異  $\pm$  標準偏差内の値なら、そのフレームの動きは標準的な動きとし、差異が平均表情モデルの差異  $\pm$  標準偏差の値を超えた場合そのフレームの行動を個性的な動きとして抽出する。そして平均表情モデルの行動時間を超えた被験者の行動については、平均には無い行動とし、その部分の行動も個性として抽出する。

この方法により擬人化エージェントのモデルとなる被験者の表情変化による顔パーツの個性的な動きと一般的な動きを分ける事ができ、表情におけるパーツの特異的な動きを抽出することが可能である。

#### 4 おわりに

本稿では擬人化エージェントとしてモデル化される個人の表情変化における個性を抽出する手法の提案を試みた。ここでは、表情変化における個性抽出方法として、表情譜による表情解析とその結果に基づく平均表情モデルの構築方法、及び、平均表情モデルに基づく個性の抽出方法を示した。

ここでは、特定の表情変化を行った際の個人の顔パーツの動きから個性を抽出する手法について述べたが、表情変化における個性には、顔パーツの動き出す順番、行動の終わる順番といったさらなる要因も重要な意味を持つことが考えられる。これらの要因については、今後検討する必要があると考えている。また、現時点では本手法を未だ実装した状態で検証していないため、早急に実装し、検証を行う予定である。

#### 参考文献

- [1] 西田 佳史, 相澤 洋志, 堀 俊夫, 柿倉 正義 “超音波センサを用いた対象物のセンサ化に基づく人の日常活動の認識”, 東京電機大学, 産業総合技術研究所, 2002 年.
- [2] 西山 正紘, 川嶋 宏彰, 松山 隆司, “表情譜: 顔パーツ間のタイミング構造の記述とその自動獲得”, 京都大学大学院情報学研究科, 2004 年.
- [3] 真狩 和加子, “個人の身体的特徴を反映したアバタの自動生成手法”, 奈良先端科学技術大学大学院 情報処理学専攻, 修士論文, 2000 年 2 月 14 日.
- [4] 三井・東芝, “スペースチャイルドアドベンチャー「グランオデッセイ」”, 2005 年.