

平成 21 年度（2009 年度）

卒業論文要旨集

Abstracts of the 2009 Graduation Theses

岩手県立大学ソフトウェア情報学部
Faculty of Software and Information Science
Iwate Prefectural University

対面者との相互作用を目的とした表情合成に関する研究

A study on facial expression synthesis for human interaction

インテリジェントソフトウェアシステム学講座 0312005093 鷹野拓也
指導教員：羽倉淳 榎松理樹 藤田ハミド

1 はじめに

本研究は人間と擬人化システム¹⁾の間の相互作用において感情を介したコミュニケーションを促進するための表情合成に関するものである。

これまでの擬人化システムにおける感情合成は、主に、Ekman と Friesen による FACS (Facial Action Coding System) に基づくものが一般的である²⁾。FACS では、人間の表情を「視覚的に識別できる解剖学的に独立した動作の最小単位」である AU (Action Unit) の動作の組み合わせとして表現し、これらの組み合わせによる感情表現はコンピュータによる表情合成と適合性が高いといえる。

一方で、現行の表情合成方法では、基本 6 感情に対して予め表情の表現方法を一意に定めていることが多い。従って、表出すべき感情が同一の場合、毎回同じ表情が合成されることとなる。しかし、人間同士の対話では対面者との関係や状況に応じた表情を表出するため、同じ感情を表す表情は一意に定まらない。

そこで本研究では、特定人物の人格のモデル化による再現を目的とした擬人化システム¹⁾において、次節に示すような対面者との関係や状況に応じて同一感情を表現する際に様々な表情を表出できる表情合成手法を提案する。

2 表情合成手法

本論文で対象とする擬人化システム¹⁾の表情は、CG によって再現された特定の人物の顔上にリアルタイムで逐次的に合成される。ここでは、表情合成法の概要を示す。

2.1 表情合成の概要

本システムにおける表情合成は、FACS を基に AU やその動作による皮膚の皺を 3次元の顔モデル上で再現する動作パーツを作成し、その組合せにより様々な表情を合成する。本システムでは、システムの発話内容に感情名・感情の強度を決定するモジュールを持つ。このモジュールからの指示により感情を表出する。感情の種類は基本 6 感情である。

2.2 顔モデルの制御及び AU の合成手法

3次元顔モデルは x, y, z 空間内で制御される。顔モデルの表面は多数のメッシュで構成されており、各メッシュの頂点を制御点として用いる。これらの制御点を x, y, z 軸方向に移動することにより、顔モデルの形状は変化する。この変化を用いて表情を合成するための AU を作成する。

AU を作成する際は以下の手順で行う：

- ① 図 1 に示されるように、作成する AU に含まれる各制御点の移動ベクトルを設定する。
- ② 各 AU に対して、一般の人間の表情筋の可動範囲の限界を推測し、各制御点の移動ベクトルの最大値とする。
- ③ ①, ② で設定に用いた制御点の集合を AU と関連付け登録する。

本手順により設定した AU のベクトルの最大値を変更することで、AU の動作量を変更することができる。



図 1 AU1 (内眉を上げる) の制御点のベクトル

3 対面者との関係及び状況

ここでは「対面者との関係」は、対面者のプロフィールとモデル化される人物の性格によって定まるものとする。さらに「状況」は、相互作用の目的、相互作用時の対面者の反応によって定まるものとする。

3.1 対面者のプロフィール

Ekman²⁾によれば、人は相手の立場や、文化の違いによりその関係に適した表情を表出する。よってここでは、「対面者のプロフィール」として、対面者の年齢や性別、職業といった人間の一般的な属性に注目する。本稿では年齢 (子供 (3 歳~10 歳), 大人 (11 歳~64 歳), 老人 (65 歳~)), 性別 (男, 女) を対面者のプロフィールとして扱う。

3.2 対面者の反応

人間同士の会話においては、相手の反応により自分自身の内部状態が変化し、相手との接し方が変化する³⁾。そこで、「対面者の反応」として、対面者の表情や仕草から推測されるシステムへの興味の示し方と感情を考える。さらに 4 節で示すように、システムに内部状態を与え、対面者の反応により内部状態を変化させ、それにより表情を変化させるメカニズムを付与する。

3.3 システムの性格

人は、同様の感情を抱いたとしても、その表情への表出方法は、性格によって様々であることが知られている⁴⁾。このことから、システムにモデル化の対象となる人物と同様の性格を反映させることで、対面者との関係を考慮した表情表出が可能となることが期待できる。これをシステムの性格と呼ぶ。モデルとなる人物にエニアグラム⁵⁾による性格分析を行い 9 つの性格タイプへの当てはまりの程度を図 2 で示すように、10 段階の値で表わすことができる。各性格タイプに対して、対面者との関係、状況に応じた表情を各性格タイプの特徴を基に作成することで、性格を反映した変化をシステムに与えることができると考えられる。

3.4 システムの心理状態

人は、同じ感情を表す表情でも、毎回同じ表情を出すとは考えられにくい。これは、対面者の関係や状況により本人の心理状態が変化するからである³⁾。本稿では「対人感情に伴う対人的情緒」(好意、慈愛、優越感、軽蔑、嫌悪感、恐怖心、劣等感、敬意) をシステムの心理状態とする。さらに、システムの性格はシステムの心理状態の変化に影響を与えるようモデル化を行う。



図2 エニアグラム9つのタイプの特徴グラフの例

4 対面者との関係や状況を考慮した表情合成手法

賢治システムの中には9つの心理状態が存在しており、現在滞在している心理状態に対面者から抽出された感情が入力されると心理状態が遷移する。各心理状態が遷移するためにはそれぞれ決まった条件が存在しているため、その条件を満たさないと心理状態は遷移しない。また、入力値により遷移先も異なる。各心理状態の中にはその心理状態の特徴と6感情の特徴を関連付けた表情が入っており、システムは現在滞在している心理状態に適した表情出力する。どこの心理状態に遷移するか(以降、遷移モデルと呼ぶ)、心理状態が遷移するための条件(以降、遷移条件と呼ぶ)、各心理状態の表情の表現の仕方は、システムの性格により変化する。性格は節2.3で述べたエニアグラムを用いている。図1の中心を原点とし、性格タイプとそのタイプの特徴の値を1つのベクトル(以降、性格ベクトルと呼ぶ)とし、9つの性格ベクトルを組み合わせたものをシステムの性格と考える。以下に本手法の詳細を述べる。

4.1 性格を考慮したメンタルモデル

[a] 遷移モデルについて述べる。まず、各性格タイプにおける、各心理状態からの遷移先を予め全て設定しておく。入力値は6感情なので1つの心理状態からの遷移先は6通り考えなければならない。そして、設定した遷移モデルにシステムの性格の重みを反映させる。各性格タイプの重みは以下の(1)で定義される。

- ・ x = 性格ベクトルの大きさ
- ・ y = 9つの性格ベクトルの大きさの和

$$\text{各性格タイプの重み} = \frac{x}{y} \quad (1)$$

- ① 各性格タイプにおいて、現在滞在している心理状態及び入力値が同一であったとき、遷移先が同一となる性格タイプを同グループとする。
- ② (1)より、グループ内の各性格タイプの重みの和を求める。
- ③ 各グループ内の重みの和を比較し、一番値が大きかったグループの心理状態が遷移先となる。

[b] 次に遷移条件について述べる。各心理状態全てをシステム内のオブジェクトとし、入力値に応じて心理状態の遷移が行われる。賢治システムが読むテキストの1文を感情抽出区間とし、1文読み終えた際に、抽出された感情が入力値となり、現在滞在している心理状態に入力される。この工程を何度か繰り返し、入力値の回数をカウントしていき、設定された回数分連続で入力された際に心理状態が遷移する。カウントしている際、違う入力値が入力された時は、カウント回数がリセットされる。尚、遷移条件の回数の設定方法は、プロフィールの情報を考慮して求めているため、次章にて述べる。

[c] 各心理状態における表情の表現方法について述べる。まず、各心理状態を表情で表す際に、AUの表現の方法が異なる表情が幾パターンか必要となるため、予め全て作

成しておく。各心理状態の特徴に適した表情を各感情から1つずつ選択し組み合わせるものを、心理状態における表情集合とする。そして、各性格タイプにおける各心理状態の表情の動作量を予め設定する。表情の動作量は(0~100)で表わされる。各感情における表情の動作量は(3)で求められる。"z"には"表情の動作量"が

- ・ z = 係数
 - ・ i = 性格のタイプ
- ※"z"に入る値は毎時提示する

$$\text{システムの性格の重みを考慮した値} = \sum_{i=1}^9 \frac{x_i z_i}{y} \quad (2)$$

4.2 プロファイルを考慮したメンタルモデル

遷移条件及び各心理状態における表情の表現方法はプロフィールによっても変化する。本稿で扱うプロフィールは節2.1で述べられているように年齢と性別である。プロフィール情報を考慮し、各性格タイプにおける、各心理状態へ遷移する際の遷移条件の回数(以降、プロフィール情報を含む遷移条件回数とする)を予め全て設定しておく。入力値は6感情なので、1つの心理状態からの遷移条件はそれぞれ6通り考えなければならない。そして式(2)の"z"は"プロフィール情報を含む遷移条件回数 = m"としてプロフィールを考慮した遷移条件が求められる。また、各心理状態における表情の表現方法について述べる。まず、各心理状態における表情の動作量を、プロフィールによりどの程度動作量を大きくするか(以降、動作量の割合とする)を各性格タイプ毎に予め設定しておく。尚、動作量が100を超えてしまうと、人間が表出不可能な表情になってしまうため最大値を100とする。そして、式(2)の"z"を"動作量の割合 = l"とし、システムの性格における動作量の割合を求める。求めた値を前項の[c]で導かれた各感情の動作量に乗算することでプロフィールを考慮した各心理状態における表情の表現方法が求められる。

5 評価実験

本手法を用いて実際賢治システムを利用している対面者の様子の動画と、その際に賢治システムが出力していた表情の動画を同時に提示し、不特定多数の人に見比べてもらい、賢治システムが表情を変化させながら対応している様子を見て、出力した表情がその状況に適切であったかを判定してもらい、効果を検証する予定である。

6 おわりに

本研究では、対面者の反応やプロフィールを考慮し、実在する人物の人格をシステムに再現することで、人間同士の対話のようなHCIシステムの構築を試みた。今後は評価実験を重ね、対面者のプロフィールや反応を追加すること、また、状態が遷移する条件をさらに細かくすることで、より人間らしいHCIシステムの構築が可能ではないかと考えられる。

謝辞

本研究は科研費(20300078)の助成を受けたものである。

参考文献

- 1) 中坪祐介他: 擬人化システムの利用者に応じた表情合成手法, 岩手県立大学ソフトウェア情報学部(2008)
- 2) Paul Ekman, Wallace V. Friesen, 工藤力 訳: 表情分析入門, 誠信書房(1987)
- 3) 吉森護: 人間関係の心理学ハンドブック, 北大路書房(1991)
- 4) ヘレン・パーマー, 佐々木和恵 訳: 新エニアグラム, 株式会社ディビーエス・ブリタニカ(1996)
- 5) 鈴木秀子: 9つの性格, PHP文庫(2004)

視線と頭部姿勢に着目したユーザの興味推定

Estimation of Human Interest from eyes and head posture

インテリジェントソフトウェアシステム学講座 0312006108 高橋宣裕

指導教員：羽倉淳 榎松理樹 藤田ハミド

1. はじめに

HCI システムにおいて、人の感情を認識し、感情に応じた応答を行うシステムへの要求が近年高まってきている¹⁾。人とHCIシステム間で円滑なコミュニケーションを実現することが1つの課題として挙げられる。それを完成するために、ユーザの興味の有無をシステムが認識し、興味をもたせるための応答をシステムが実現することで人間とシステム間の相互作用性が向上すると考えられる。

心理学においては、人間は興味があるものがある場合、より正確な情報を得られる位置を保つことが知られている。正確な情報を得る位置とは、対象の正面から、対象に近い位置を意味する。また、興味が無い場合は体の負担を減らすために、頭を傾け、椅子の背もたれによりかかるなどの動作をとりやすい。この知見から、興味を推定する上でユーザの頭の向き、視線のように情報を取得している方向や、ユーザと対象との距離、ユーザの顔の傾きからどのような姿勢で対象を観測しているかが興味を推定する上で重要な情報であると考えられる。

この心理学的知見に基づき、ここでは、ユーザのシステムに対する興味をシステムと対面したユーザの顔までの距離、ユーザの顔の向きと傾き、および、ユーザの視線方向から検出されるものとし、システムが自動的に推定可能にするための手法を提案する。

2. 興味推定へのアプローチ

本研究で対象とするシステムは、ユーザがシステムの前に置かれた椅子に座った状態でシステムを利用するエンターテインメントを意図した擬人化システム²⁾である。そして、このシステムはカメラから動画像をとりこむことが可能である。このようなシステムにおいては、システムの目的上、ユーザがシステムに対して積極的、好意的であり、面白い、楽しいと感じている状態を興味がある状態と認識することが望ましいと仮定する。一方で、対象を漠然と眺めている傍観といった状態は興味が無い状態として判定を行う必要があるものとする。

上述の心理学的知見とここで述べた興味がある際の行動を考察し、比較すると、システムに対して積極的、好意的な際には、身を乗り出しシステムに顔を近づけると考えられる。同様に、面白い、楽しいと感じている際には、システムの反応を見逃すまいと凝視するためにシステムを正面から捉えるものと予想される。一方、興味なく漠然とシステムを眺めている場合には、自ずと楽な姿勢に陥るものと考えられる。すなわち、上述の心理学的知見に基づく人間の反応とここで必要とする興味の有無の判断はほぼ一致していると考えられる。

そこで本研究では、システムに対するユーザの興味の有無を推定するために、ユーザの顔の向き、視線方向のように情報を取得している方向や、ユーザとシステムとの距離、ユーザの顔の傾きからどのような姿勢で対象を観測しているかに着目する。さらに、人間は興味を持っていても、

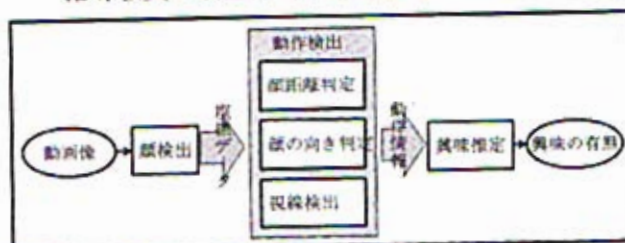


図1: システム全体像

肉体の疲労や集中力の持続性の観点から、興味がないと判断され得る行動をとることも想定される。この問題に対して、ユーザの一定時間の動作系列を観察し、そこで興味があると判定される動作回数と興味がないと判定される動作回数を比較し、最終的な興味の推定を行う。

3. 興味推定手法

システムの全体像を図1に示す。以下では各特徴の抽出手法と興味推定方法について述べる。

3-1. 顔距離判定

カメラから得られた動画像からユーザの顔を検出し、顔領域を表す4つの座標点を取得する。長方形である顔領域の座標点は左上から時計回りに (x_0, y_0) (x_1, y_1) (x_3, y_3) (x_4, y_4) とする。

顔とカメラとの距離判定では、顔とカメラの距離の変化で顔画像中の顔領域が占める面積が変化することを利用し、面積の変化からユーザと対象との距離をNear, Middle, Farの3段階で出力する。

$$A = \frac{((x_1y_2 + x_2y_3 + x_3y_4 + x_4y_1) - (x_2y_1 + x_3y_2 + x_4y_3 + x_1y_4))}{2} \quad (1)$$

顔面積Aを式(1)より求める。面積Aが閾値min以下の場合(Far)、閾値max以上の場合(Near)、2つの閾値の間の場合(Middle)とする。閾値は特定のユーザが椅子に座っている状態において前傾姿勢、後傾姿勢になるように実験的に設定する。

3-2. 顔の向き判定

顔の向き判定では、 x, y 軸方向の回転においてユーザの向いている方向に対象が存在する場合(Straight)と、対象が存在しない場合(Out Side)の2段階で出力する。顔領域の重心(図2中)と、鼻の特徴点座標との距離を用いて閾値に基づき顔の向きを判別する(図2右)。閾値は特定のユーザの顔の向きが(Out Side)になるように鼻特徴点と重心との距離を実験的に求め設定した。また、ユーザとカメラとの距離によって特徴点の動く範囲も大きくなるため、顔距離判定で得られる3つの状態毎に閾値を設定した。顔の傾きの回転では、 z 軸方向の回転においてユーザが意識的に顔を傾けている場合(On)と傾けていない場合(off)の2段階で出力する。

ユーザの顔の傾きは顔領域座標の右下と左下の2点から



図2:顔検出・顔向き判定

求める。閾値は上述の手法と同様に特定のユーザが首を意図的に傾けている状態になるように実験的に設定する。

3-3. 視線検出

ユーザが対象に視線を向けている時間(Gaze duration)、対象に視線を向けた頻度(Gaze frequency)の検出を行う。視線検出には目の特徴点の位置情報を用いる。虹彩の中心の特徴点と目の輪郭部分の特徴点との位置関係からユーザの視線を推定する。視線を向けている時間が人の視線の平均固視⁴⁾とされている250msを超えた場合(long)、それ以下の場合を(short)とする。また、観測開始から現時刻tまでに対象に視線を向けた頻度を計算し一定値以上増加した場合、視線回数が多い状態(High)とする。

3-4. 興味推定手法

ユーザの興味の有無を推定するため、2節で述べたように動作に興味がある動作と興味がない動作に分類し、抽出された動作がどちらに属するか判別を行う。本研究では興味がある動作と興味がない動作を表1のようにまとめた。そして、これらの動作の検出数を比較し、もう一方より一定回以上多く現れた興味状態をユーザの興味として推定する。また、互いの動作回数の差が小さい場合、ユーザの興味は中立として出力する。

4. 実験

本手法を用いて実際にユーザの興味が認識できることを確認するために被験者3名に対して以下のような実験を行った。被験者に18秒ごとに変化する様々なジャンルの動画を3分間、計10動画見せもらう。その後、興味のある動画となかった動画を被験者に質問し、その結果と動画を閲覧中の被験者の興味をシステムにより推定したものと比較を行った。なお、今回の顔検出、追跡ではFSE⁵⁾を用いた。FSEでは図2左のように入力された画像から目、鼻、口が収まるように顔領域を検出し入力顔画像中の位置を表す左上、右上、右下、左下の4つの座標点で表現する。また、領域に加えて鼻・目の特徴点の位置の取得もFSEを用いる。

4-1. 結果

被験者のシステムの興味推定結果を図3に示す。10通りの動画の変化について被験者の興味があったもの無かったものを図3下にあるように挙げてもらった。グラフ中の縦軸は興味に関連した動作のフレーム毎の検出数の差から求めたユーザの興味の度合いを表しており数値が高いほど、システムはユーザが興味をもっているとして認識している。ここで本システムの認識結果を表2に示す。認識率とは被験者が興味の有無があったとした区間に対するシステムの正答数の割合である。

今回の実験では被験者が動画に対して興味が無い状態はある程度正確に認識できていた。しかし、興味あり状態の認識率が個人によって大きな差がでてしまう結果となった。さらに、被験者が興味の有無のどちらも示さない中立状態においてシステムは興味有り、もしくは興味無しとして誤認識してしまうケースが多いことがわかった。

4-2 考察

今回の実験結果から被験者の興味状態が正しく認識されなかった原因として動作を検出するための閾値を特定のユーザに合わせて設定したことが挙げられる。

特定のユーザの動作情報を用いる場合、個人差による顔

表1:興味動作の分類のルール

興味あり	興味なし
FD(Far or Middle → Near)	FD(Near or Middle → Far)
FP(Straight), FD(Near)	FP(Straight), FD(Far)
FP(Straight), FD(Middle)	FP(Out side)
GD(Long)	FR(On)
GF(High)	GF(Low)

FD…顔距離, FP…顔の向き, FR…顔傾き, GD…視線量, GF…視線頻度

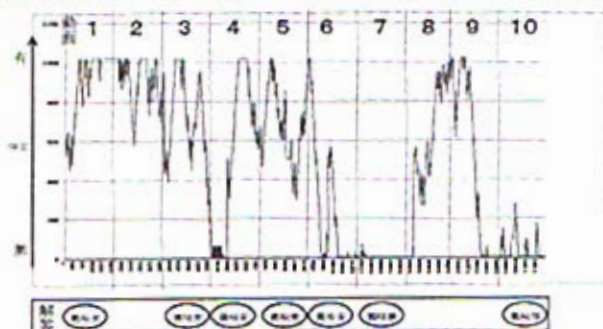


図3:実験結果

表2:認識結果

	興味有り	興味無し
被験者A	4	3
システム	2	2
認識率	50%	66%
被験者B	4	3
システム	1	2
認識率	25%	66%
被験者C	3	2
システム	2	1
認識率	66%	50%

パーツの座標位置の違いから誤差が生じるため閾値の設定方法について考え直す必要がある。また、本システム興味推定手法では検出された全ての特徴動作をルールに基づいて分類し興味推定を行ったため、ユーザの興味の有無に関連しない動作も興味の有無を含む動作として検出していると考えられる。したがって、特徴動作と興味を関係付けるのではなく特徴動作の組み合わせから興味を推定することで認識率が向上すると考えられる。

以上の結果から、今後は特徴動作抽出から得られた情報をユーザ状態として、ユーザ状態がどのような組み合わせの場合に興味が変わるか心理学的知見³⁾などに基づいてルールを作成し、それらを用いた興味推定手法の実装、実験、検証を行う。

5. おわりに

本稿では、利用者である人間の擬人化システムに対する興味の推定を目的として、顔の特徴動作、視線の動かし方に着目した興味推定手法を提案した。

参考文献

- 1) 趙晋輝, 他: 特集感性情報学, 電子情報通信学会誌, Vol.92, P911-997 (2009)
- 2) Kenji-Project: <http://www.fujita.soft.iwate-pu.ac.jp/KENJI/index.html>
- 3) P. ブル, 市河淳章 訳, 高橋超 訳: 姿勢としぐさの心理学, 北大路出版(2001)
- 4) 乾敏郎: 視覚情報処理の基礎, サイエンス社(1991)
- 5) 沖電気: Face Sensing Engine, 沖電気(2008)