

個人モデルを用いた表情分析／合成による感情通信システム

倉石英俊 柴田義孝

東洋大学 工学部 情報工学科

{kuraishi,shibata}@sb.cs.toyo.ac.jp

筆者等はこれまで、個人モデルを用いて感情の表現、認識の個人差を解消し、より正確な感情の伝達を可能とする感情通信システムを提案している。本システムにおける分析側の個人モデルは、感情を表現するユーザ個人の表情の特徴を表すモデルであり、合成側の個人モデルは、表情から相手の感情をどう理解するか、すなわち表情の見方の特徴を表すモデルである。従って分析側のモデルを用いて表情分析を行なうことで、表情からのより正確な感情の抽出が可能となり、合成側の個人モデルを用いて表情を合成し、表示することで、より正確な感情の伝達が可能となる。本稿では、分析側の機能評価および合成側の個人モデルの作成方法について述べる。

Feeling Communication System by Facial Expression Analysis/Synthesis using Individual Models

Hidetoshi Kuraishi and Yoshitaka Shibata

Department of Information and Computer Sciences

Toyo University

{kuraishi,shibata}@sb.cs.toyo.ac.jp

We proposed a feeling communication system which enable to convey more accurate feeling by considering the difference of facial expression and the difference of feeling recognition from one facial expression using individual models. In our system, an analytical individual model is defined as a model which shows features of the individual expression. Beside, an synthetical individual model is defined as a model which shows features of facial expression which one person can understand an fellow's feeling correctly.

The analysis side can extract an actual feelings from facial expression facial expression analysis using analytical individual model, and the synthesis side can convey a more accurate feeling by facial expression synthesis using synthetical individual model and display of the synthesized facial image. In this paper, we describe about the evaluation of facial expression analysis and about synthetical individual model generation.

1 はじめに

我々日本人は他国民族に比べ感情の表現力に乏しいと言われる。また、ある民族は我々から見て表現が豊かであるためその表情がオーバーに見える。このような表情における感情表現の

違いにより、互いに表情から感情を理解する場合に認識のズレが生じ、国際的なコミュニケーションの場などにおいて相互理解にしばしば支障をきたすことがある。

我々はこのような問題を解決するために、個

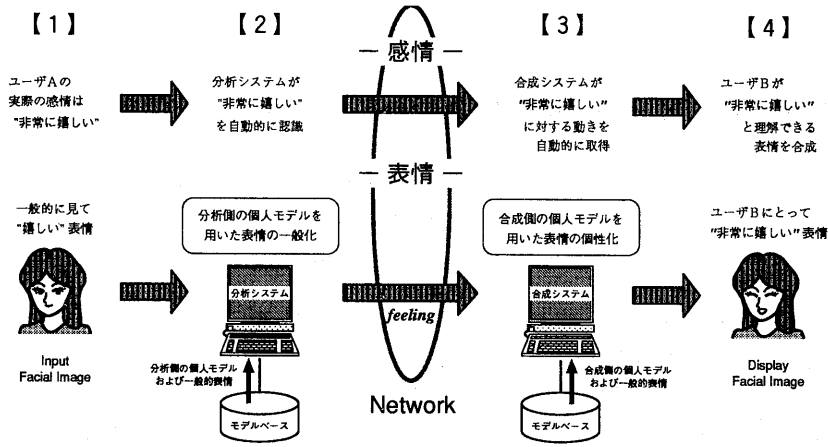


図1. 個人モデルを用いた感情通信

人モデルに基づく表情の分析および合成を行なうことにより、より正確な感情の伝達を可能とする感情通信システムを提案している [1]。

本システムは、分析側の個人モデルを用いてユーザ個人の表情からそのユーザの特徴を取り除き (表情の一般化)、感情を判別することで、より正確なユーザの感情を取り出す機能、および合成側の個人モデルを用いて一般的表情にユーザの見方を考慮した特徴を加え (表情の個性化)、その表情をコンピュータグラフィックス (CG) を用いて合成、表示する機能により正確な感情の伝達を行なうシステムであり、通常のビデオ通信システムと組み合わせることにより、初対面の人間同士やフォーマルな場でのコミュニケーションなどの相互理解が必要とされる場合においても円滑なコミュニケーションを提供する。

本稿では、2,3 章で感情通信システムの概念を述べ、4 章で分析側の個人モデルを用いた分析の機能評価および考察を行ない、5 章で合成側の個人モデルの作成方法について述べ、最後に6 章でまとめる。

2 個人モデルを用いた感情通信

感情通信システムにおいて、伝達の対象とされる感情は、喜び、怒り、悲しみ、驚き、恐れ、

嫌悪の基本的な6通りであり、各感情において、少し、普通、非常にの3段階の感情の度合を考慮している。ここで、これらの感情の表現は、必ずしも元の人間の顔に忠実な表情によって行なう必要はなく、アニメーション (表情、ジャスチャー [3]) や言葉、文字などでもよいと考えるが、本システムでは、感情を表現するものとして原顔画像による表情を用いている。

図1は、個人モデルを用いた感情通信の流れを、表現力に乏しいユーザAと感情をやや控え目に理解するユーザBがネットワークを介してテレビ電話のような通信を行なっている (図2) という例を用いて示している。

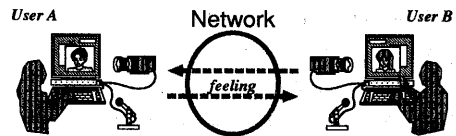


図2. 1対1通信

- 1) ユーザAは本当は「非常に嬉しい」のだが、表現力が乏しいために表情は一般的に見て「ただ嬉しい」を表現しているように見える。カメラはユーザAの「ただ嬉しい」の表情を忠実に取り込む。

- 2) 分析システムは、ユーザAの「ただ嬉しい」の表情をユーザAの分析側の個人モデルを用いて一般化し、感情を判別することにより、ユーザAの正確な感情(「非常に嬉しい」)を抽出する。そして感情情報(感情、感情の度合)を合成側に転送する。
- 3) 合成システムは、受信したユーザAの感情(「非常に嬉しい」)に対する一般的表情をユーザBの合成側の個人モデルを用いて個性化し、ユーザBがユーザAの感情を正確に理解できる表情(少しオーバーな表情)を合成する。
- 4) ユーザBは表示されたユーザAの表情から、ユーザAが「非常に嬉しい」ことを理解できる。

3 システム構成

ここでは処理モジュールに沿って処理の流れを説明する。

図3は本システムの構成を示す。分析側では各感情に対する一般的表情パターンと分析側の個人モデル、合成側では一般的動作マトリクスと合成側の個人モデル[1]があらかじめ作成されており、それらのモデルはモデルベースに格納されている。

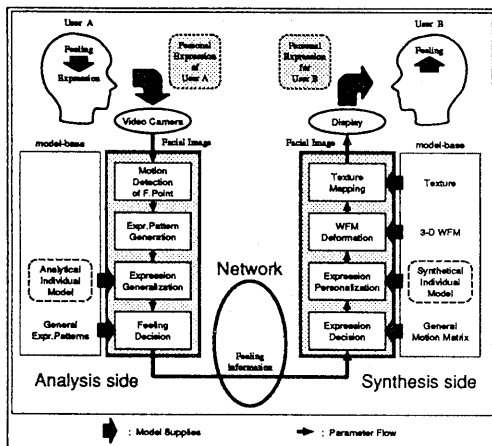


図3. システム構成

- 1) ビデオカメラによる顔画像の入力
- 2) **Motion Detection of Feature Points**
特徴点の動き検出を行なう。
- 3) **Expression Pattern Extraction**
特徴点の動きを用いて個人の表情パターンを作成する。
- 4) **Expression Generalization**
分析側の個人モデル M_{analy} をモデルベースから得、式 (1) を用いて個人の表情パターンを E_{indiv} を次式で一般化し、一般化した表情パターン E'_{gen} を得る。

$$E'_{gen} = M_{analy} \cdot E_{indiv} \quad (1)$$

- 5) **Feeling Decision**
一般化した表情パターン E'_{gen} とあらかじめモデルベースに登録されている一般的表情パターン E_{gen} とのパターンマッチングを行ない、誤差が最小となる一般的表情パターンに対する感情を抽出する。
- 6) 感情情報として、「喜び」「驚き」といった感情を表す単語と、「少し」「非常に」といった感情の度合を表す単語を合成側へ転送する。
- 7) **Expression Decision**
感情を表す単語および感情の度合を表す単語を受信し、受信した感情および感情の度合に対する表情の一般的動作マトリクス V_{gen} を得る。
- 8) **Expression Personalization**
合成側の個人モデル M_{synth} をモデルベースから得、式 (2) を用いて一般的動作マトリクス V_{gen} を個性化し、個人の動作マトリクス V_{indiv} を得る。

$$V_{indiv} = M_{synth} \cdot V_{gen} \quad (2)$$

- 9) **WFM Deformation**
動作マトリクス V_{indiv} に従って、ワイヤフレームモデル (WFM) を変形させる。
- 10) **Texture Mapping**
テクスチャをはる。
- 11) 表示。

4 分析側の機能評価

ここでは、個人モデルの有効性を評価することによって分析側の感情認識機能を評価する。

4.1 評価手順

- 1) 3次元曲面形状計測装置による学生14人の表情の3次元データの測定。「喜び」「怒り」「悲しみ」「驚き」に対する表情を作らせ、測定した。「喜び」と「驚き」については3段階(「少し」「普通」「非常に」)、「怒り」と「悲しみ」については2段階(「普通」「非常に」)の合わせて10通りの感情を表す表情を一人当たり3回ずつ測定した(無表情と合わせて一人当たり合計31表情)。14人分で表1に示す数の表情を測定した。
- 2) 各感情(例:少し嬉しい、非常に驚いた)を表す表情を平均化し、各表情に対する一般的表情パターンを作成する。
- 3) 一般的表情パターンと個人の表情パターン(測定した3つの表情パターンを平均したパターン)を用いて、すべての表情に対する分析側の個人モデルを作成する。
- 4) 測定したすべての表情を分析する。

	少し	普通	非常に
喜び	42	42	42
怒り	0	42	42
悲しみ	0	42	42
驚き	42	42	42

表1. 測定、分析した表情の数

- 5) 各感情において感情認識率 FRR (Feeling Recognition Ratio) を計算する。

$$FRR = \frac{N_{success}}{N_{analysis}} \times 100 \quad [\%] \quad (3)$$

- FRR : 感情認識率 (FRR)
 $N_{success}$: 正確に分析できた表情パターンの数
 $N_{analysis}$: 分析した表情パターンの数

- 6) 求めた感情認識率を以下の場合において比較評価する。

場合1 個人モデルを用いない場合

場合2 「普通」に対する個人モデルを用いた場合(「少し」「非常に」の表情に対しても「普通」の個人モデルを使用する)

場合3 「少し」「普通」「非常に」に対する個人モデルを用いた場合

ここで「恐れ」「嫌悪」に対する表情および「怒り」「悲しみ」の「少し」に対する表情については、これらの表情を故意に作る事が困難であったため、今回は評価を行なわなかった。

4.2 評価結果

表2,3,4はそれぞれ場合1,2,3の各感情の感情認識率を示す。

	少し	普通	非常に
喜び	7.143	7.143	14.286
怒り	-	71.429	92.857
悲しみ	-	21.429	7.143
驚き	64.286	64.286	78.571

表2. 場合1: 感情認識率 [%]

	少し	普通	非常に
喜び	33.333	100.000	71.429
怒り	-	100.000	66.667
悲しみ	-	100.000	64.286
驚き	33.333	100.000	69.048

表3. 場合2: 感情認識率 [%]

	少し	普通	非常に
喜び	97.619	100.000	95.238
怒り	-	100.000	100.000
悲しみ	-	100.000	100.000
驚き	97.619	100.000	100.000

表4. 場合3: 感情認識率 [%]

表2を見ると、場合1においては、感情によって認識率にかなりのバラツキがあるが、「怒り」

と「驚き」が「喜び」と「悲しみ」よりも比較的高い感情認識率が得られている。これは、図4に示すように、「喜び」と「驚き」の表情における特徴点の移動ベクトルが非常に近いものであったために、「喜び」の表情に対して「驚き」が、同様に「悲しみ」の表情に対して「怒り」が認識されてしまったためと考えられる。

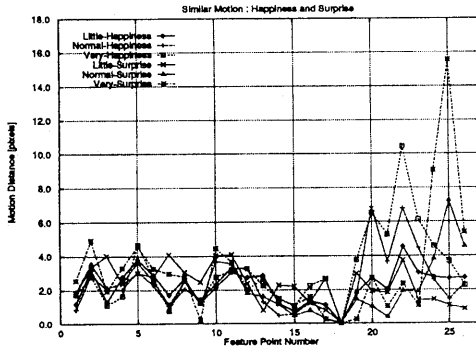


図4. 「喜び」「驚き」に対する一般的表情と個人の表情における特徴点の移動距離の比較

表3を見ると、場合2においては、場合1に比べ比較的安定した結果が得られている。「普通」に対する個人モデルを用いているので、「普通」に対しては十分な感情認識率が得られているが、「少し」に対する感情認識率は低く、「非常に」に対しても十分とは言えない。これは図5が示すように、感情の度合に対する特徴点動作において、一般的表情と個人の表情との変化が必ずしも一致しない、すなわち「少し」「普通」「非常に」のそれぞれにおける一般的表情と個人の表情との特徴点動作の差が等しくないことが原因であると考えられる。しかし図6に示すように、「普通」に対する個人モデルにより「少し」「非常に」に関するパターンマッチングによる自乗誤差は減少している。

場合3においては、表4を見てわかるように、どの感情でも十分な感情認識率が得られている。「少し」「非常に」に対する個人モデルの使用により、図6が示すように精度の高い分析が行なわれるため、場合2で全く改善されなかった「少し」に対しても正確に感情を認識す

ることができた。

以上の結果から、感情通信システムにおいて分析側の個人モデルは有効であり、「少し」「普通」「非常に」に対する個人モデルを用いることにより、感情の度合の変化にも対応でき、実験を行なった4つの感情において、分析側は個人の表情から正確な感情が認識できることが確かめられた。

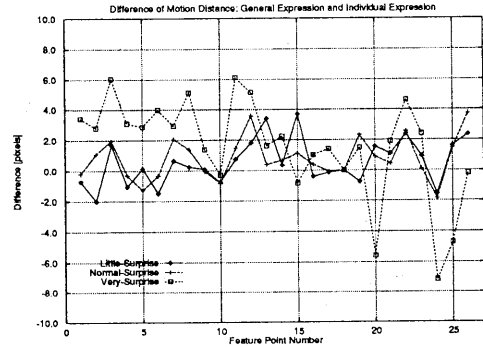


図5. 各感情の度合における、一般的表情と個人の表情との特徴点の移動距離の差の比較

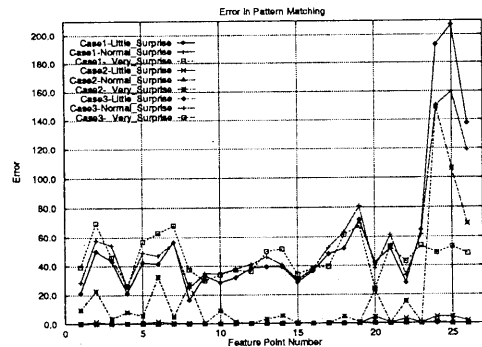


図6. パターンマッチングによる各場合の自乗誤差の比較

5 合成側の個人モデル

合成側の個人モデルは、一般的表情を受信側のユーザが相手の感情を理解できる表情に変換するモデルである。合成側の個人モデルの精度は、受信側のユーザが表示された表情を

見て、感情を理解できるか否かの評価に依存する。従って合成側の個人モデルは、以下の処理をユーザが相手の感情を正しく理解できるまで繰り返すことにより作成される。

- 1) 合成側の個人モデルの初期化 ($M_{synth}(0)$: 単位行列)。
- 2) ある感情 (例: 喜び) に対する一般的表情を個性化 (式 (4)) し、合成表示する。

$$V_{indiv}(i) = M_{synth}(i) \cdot V_{gen} \quad (4)$$

- 3) ユーザによる合成画像の評価。顔の部位の動きの変更により $V_{indiv}(i+1)$ を得る。

$$V_{indiv}(i) \rightarrow V_{indiv}(i+1)$$

- 4) 合成側の個人モデルの更新。式 (5) を満たす $M_{synth}(i+1)$ を得る。

$$V_{indiv}(i+1) = M_{synth}(i+1) \cdot V_{gen} \quad (5)$$

図7は「喜び」を表す一般的表情の合成画像である。図7は $M_{synth}(0)$ を用いて合成した表情と一致し、上記の合成側の個人モデルの作成は、この表情に対するユーザの評価から開始される。

現在は合成側の個人モデルの学習・更新を行なうツールを作成中であり、合成側の個人モデルの評価は次に行なう予定である。



図7. 「喜び」の合成画像

6 まとめ

これまで、感情通信システム の概念、分析側の機能評価および合成側の個人モデルの作成方法について述べた。本システムは感情を正確に伝える必要がある場合に有効なシステムであり、分析側および合成側の個人モデルを用いることにより、民族、年齢などによる感情表現の違いや表情からの相手の感情の認識の違いから生じる誤解を取り除き、円滑なコミュニケーションを行なうことが可能である。また個人モデルに基づく表情の分析/合成を行なうことにより、コンピュータとのより心のこもった対話を実現するアプリケーションの作成も可能であると考えられる。

今後は、より正確な表情合成システムを実装し、合成側の個人モデルの作成、評価を行なった上で、実際の通信における感情通信システムの機能評価を行なう予定である。

参考文献

- [1] 倉石英俊, 柴田義孝: “ユーザモデルを用いた感情通信システム”, マルチメディア通信と分散処理研究会, Vol.95, No.53, pp.19-24, May 1995.
- [2] K.Aizawa, H.Harashima, T.Saito, “Model-based analysis synthesis image coding (MBASIC) system for person's face”, *Signal Process., Image Com.*, vol.1, no.2, pp.139-152 (Oct. 1989)
- [3] T.Inoue, K.Okada, Y.Matsusita, “Using gestural animations to convey the feelings in asynchronous network communication”, ICOIN-J005
- [4] 崔昌石, 原島博, 武部幹, “顔の3次元モデルに基づく表情の記述と合成”, 信学論(A), J73-A, 7, pp.1270-1280 (1990-07)
- [5] P.Ekman, W.V.Friesen, “Facial Action Coding System”, Consulting Psychologists Press Inc., 577 College Arc, Palo, Alto, 1977.