

## 個人モデルに基づく感情通信システムの研究

林耕平 片平静穂 戴瑩 柴田義孝

岩手県立大学 ソフトウェア情報学部

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、コンピュータネットワーク上での人間同士のコミュニケーションにおいて相手の感情を正しく伝達することが出来る感情通信システムの提案を行ってきた。本システムは表情の分析側と合成側に分けられる。分析側の個人モデルは、感情を表現するユーザの表情表出の際の特徴を表し、合成側の個人モデルは、表情から感情を理解する際の特徴、つまり表情の受け止め方を表すモデルである。これらの使用により本システムの分析側ではユーザの正しい感情の抽出が可能になり、また合成側ではユーザが正しく感情を認識することが出来る表情の合成が可能になる。本稿では、これまでの感情通信システムを見直し、より正確な感情の伝達を可能にするシステムの設計について述べる。

## Research of feeling communication system based on individual model

Kouhei Hayashi, Shizuho Katahira, Ying Dai and Yoshitaka Shibata

Faculty of Software and Information Science, Iwate Prefectural University

In this paper, a feeling communication system which enables users convey feelings more accurately by modeling individual differences of facial expressions as "Individual Model" is proposed. Through the use of "Individual Model", this communication system extracts and transmits the real meaning of a user's facial expression to other users. A suitable facial expression is then synthesized and presented on an individual basis to other users, thus helping users to bridge cultural barriers.

### 1. はじめに

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、ある国民は我々から見て感情表現が豊かであるために感情がオーバーに感じられる。このような感情表現及びその認識の違いは、互いの感情を理解しようとする場合に正しい感情の認識を阻害する可能性がある。

本研究では、感情表現及びその認識のずれに起因するコミュニケーション阻害の問題を解決するために感情表現及びその認識の個人差を表す個人モデルを定義し、これに基づく表情の分析・合成を行うことでずれを吸収し正確な感情の伝達を可能とする感情通信システムの開発を行った。

本稿では2章で感情通信システムの概念について述べ、3章では分析パート、4章で合成パートについて述べる。5章では個人モデルについて説明する。6章では実装と評価について述べる。最後に7章でまとめを行い、今後の課題を述べる。

### 2. 感情通信システム

#### 2.1. システム概要<sup>[1]</sup>

本研究で提案する感情通信システムはコンピュータネットワーク上での電子会議システム等において個人モデルに基づいて互いの持つ感情をより正確に伝達し円滑なコミュニケーションを支援することを目的としている。

本システムは、ユーザの表情を分析し感情を抽出する「分析パート」と、分析パートから受信した感情情報をもとにユーザの表情を合成する「合成パート」から構成される。図1は感情通信システムが個人モデルを用いて感情を伝達する過程について示している。

ここではユーザ A は普通の人よりも感情を控えめに表現し、ユーザ B は相手の感情を控えめに受けとめると仮定する。

[1] 分析側ユーザ A の感情表現は控えめであり、ユーザ A が「非常に嬉しい」と感じていて

も、表出される表情は一般的に見て「嬉しい」表情である。

- [2] ユーザ A の「嬉しい」表情に分析側の個人モデルを適用することにより、一般的な「非常に嬉しい」表情が得られる。
- [3] 合成側ユーザ B の合成側個人モデルを用いて一般的に「非常に嬉しい」表情を個性化し、ユーザ B が「非常に嬉しい」と理解できる表情、つまり一般的に見てオーバーな「非常に嬉しい」表情を得る。
- [4] オーバーな「非常に嬉しい」表情を合成、表示する。合成側ユーザ B は控えめに理解するので、この表情を見て「非常に嬉しい」という感情を理解する。

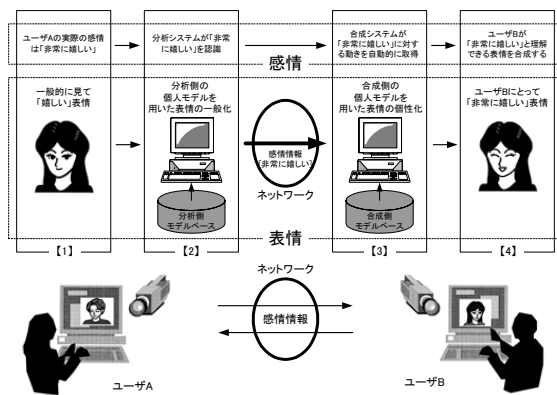


図 1. 個人モデルを用いた感情通信の例

この様にユーザ A のそのままの表情をユーザ B が見たとするならば、ユーザ A の感情を「少し嬉しい」と受けとめるところを、上記のようなプロセスにより、ユーザ B はユーザ A の真の感情である「非常に嬉しい」を正しく理解することができる。

## 2.2. システム構成

本システムは図 2 に示すように「分析パート」と「合成パート」から構成され、分析パートではカメラから取り込んだユーザの顔画像を分析し、分析側の個人モデルに基づいて本当の感情の抽出を行う。合成パートでは、分析パートで抽出された感情を合成側の個人モデルに基づいて表情に変換し、ユーザが感情を正しく理解できる表情を CG で合成する。分析パートで抽出された感情値は、Action Unit (AU)、感情のレベルとともにネットワークを通じて合成パートに送信される。

AU とは、表情記述法 Facial Action Coding System (FACS) に定義されている、解剖学的に独立し且つ視覚的に識別可能な表情動作の最小単位である。AU は合計 44 個あり、これらの組み

合わせにより人間のあらゆる表情が記述可能であるといわれている。

AU	動作
01	眉の内側を上げる
02	眉の外側を上げる
04	眉を下げる
05	上まぶたを上げる

表 1. AU の例

図 2 は本システムのシステム構成を表す。この章では、システムを構成するモジュールについて簡単に説明する。各モジュールの詳しい説明は 3 章以下で行う。

### 2.2.1. 分析側の構成

分析パートを構成する各モジュールの主な動きは以下の通りである。Facial expression change extraction では表情変化の検出を行う。Facial feature extraction では特徴抽出を行い、以下の処理を含む。Optical flow 計算、Histogram 化、Filter 処理を行い、動きの特徴を抽出したテンプレートを出力する。また AU leveling において動きの大きさから AU のレベル決定を行い、AU レベル、感情レベルを出力する。AU Recognition では、テンプレートデータを受け取り、ニューラルネットワークによる AU の認識を行う。Facial Expression Recognition では AU をもとに表情の認識を行い、感情を出力する。Model Base には個人モデルである AU パターンと表情パターンの 2 種類があり、それぞれ AU の認識、表情の認識に利用する。その後、合成パートへ感情情報と AU とそのレベルを送信する。

### 2.2.2. 合成側の構成

合成パートを構成する各モジュールの主な動きは以下の通りである。

Feeling - AU Converter では分析パートから受信した感情情報を AU に変換する。Muscle Model Generator では、表情筋をシミュレートすることで各 AU ごとの動きの情報を計算する。その内部の Muscle Motion Generator では、動作する筋肉の決定を行い、Muscle Model Simulator で筋肉の動作の計算を行う。その後 Graphic Engine にて WFM の変形を行う。このモジュールでは WFM Deformation で WFM の変形を行い、Texture Mapping にて WFM にテクスチャを貼り付ける。その後 Display へ出力する。Model Base には各モジュールで利用する個人モデルが格納されており、各モジュールへ提供する。

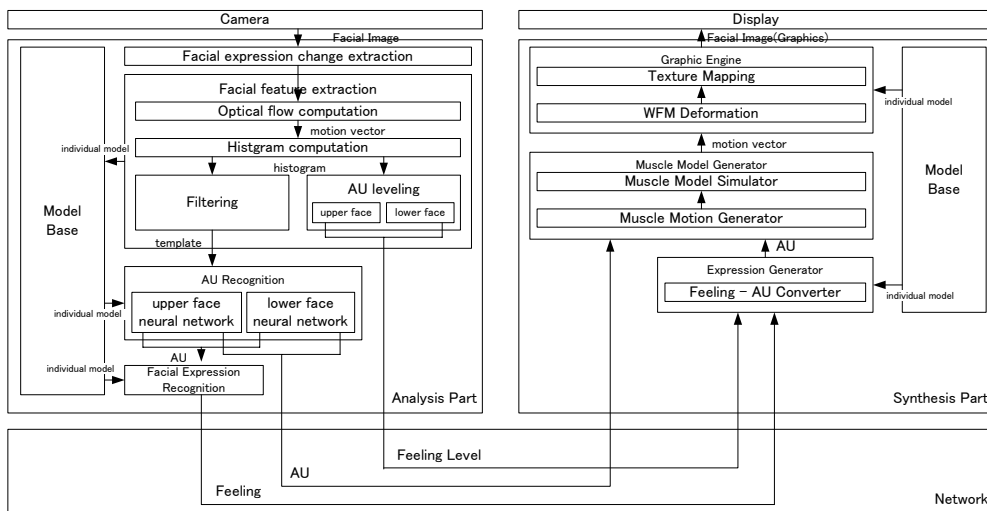


図 2. システム構成

### 3 分析パート

#### 3.1. 表情変化の検出<sup>2)</sup>

表情の変化を検出するために、RGB カラー画像を YIQ 表現に変換し、橙色からシアンまでの色調を含む I 成分を利用して人の肌の部分を抽出する。また連続する 2 枚の顔画像の差分をとり、動きのある部分を抽出する。この I 成分画像と差分画像を組み合わせることで、顔表情の変化を検出することができる。

#### 3.2. 顔表情の特徴抽出

##### 3.2.1. オプティカルフローの計算

顔表情に変化のあった連続画像のオプティカルフローにより動きの特徴を計算する。オプティカルフローは連続した 2 枚の画像を使用し、フレームから次のフレームへのある画像点が移動する動きの方向とその速度を同時に二次元ベクトルで表現できる。時間  $t$  の時のグレイレベルを  $f(x, y, t)$  とし、時間  $t + dt$  の時の移動点  $(x + dx, y + dy)$  におけるグレイレベルが同じであるとすると、次式が成り立つ。

$$f(x + dx, y + dy, t + dt) = f(x, y, t) \quad (1)$$

上式を展開すると

$$-\frac{\partial f}{\partial t} = u \frac{\partial f}{\partial x} + v \frac{\partial f}{\partial y} \quad (2)$$

となり、 $u$  は  $x$  (水平) 方向の速さ、 $v$  は  $y$  (垂直) 方向の速さであり、この  $u, v$  フィールドの分布を以下のようにヒストグラム化する。

##### 3.2.2. ヒストグラム化

顔の動きの特徴は主に目の周りとお口の周りに表れるため顔画像を上下に分割し、 $u$  フィールドを  $x$  軸に  $v$  フィールドを  $y$  軸にとりヒストグラム化する。

このとき右・下方向の動きをプラスに、左・上方向の動きをマイナスに計算したものを縦軸にとり、横軸に画像の  $x, y$  座標をとる。図 3 は「唇両端を横に引く」という動きの特徴を示した画像についての  $y$  軸のヒストグラム例である。

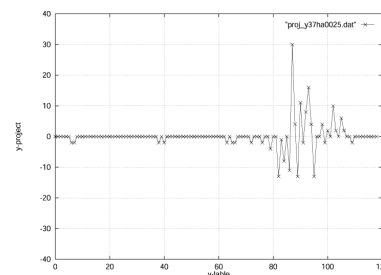


図 3. lip stretcher y ヒストグラム例

##### 3.2.3. AU レベルの決定

感情の度合いを示す要素として AU の動きの大きさと表情の大きさを定義する。まず各ヒストグラムから動きの大きさを 3 段階で求め、AU の動きの大きさとする。そして顔表情に表れた AU レベルの平均から表情のレベルを求め、感情を「やや」、「普通」、「かなり」で表す。

##### 3.2.4. フィルター処理

次に動きの特徴を抽出するため、フィルター処理を行い、以下の手順でヒストグラムをテンプレート化する。

- 背景部分や小さな動きなどのノイズを除去
- プラス、マイナス方向それぞれの極大値、次極大値を検出し記録。尚、2 点間の座標が近い場合、さらに次の値を検索する。

- その他の値を除去。

図4は「唇両端を横に引く」という動きの特徴表した画像のy軸テンプレート例であり，上唇の周囲が上に上がり，下唇の周囲がやや下に下がる特徴を示している。

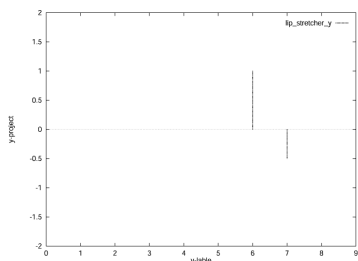


図4. lip stretcher y テンプレート例

### 3.3. AU の認識

顔の上半分と下半分の AU を扱う二つのニューラルネットワークを用い、それぞれ三層のバックプロパゲーションアルゴリズムを使って AU の認識を行う。

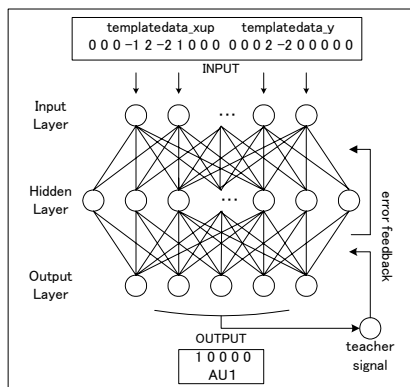


図5. 顔上半分の AU の認識を行うニューラルネットワーク

図5は顔上半分のネットワーク構造を示したもので，入力層には 20 ユニットあり，テンプレートデータが入力される．出力層には 5 ユニットありそれぞれが 1 つの AU に対応している．予めテンプレートデータとそれに対する AU を組にした学習データを使用し学習を行う．これによりそれぞれのネットワークの入力層に  $x$  軸，  $y$  軸のテンプレートデータが入力されるとニューラルネットワークにより計算が行われ，出力層から AU が出力されることになる。

### 3.4. 表情の認識

連想記憶モデルにより AU から表情の認識を行う．連想記憶モデルは  $x^k$  と  $y^k$  が対で記憶され

ているとき，連想ベクトル  $y^k$  がキーベクトル  $x^k$  から想起できる。

$$[y^1, y^2, \dots, y^k] = M[x^1, x^2, \dots, x^k] \quad (3)$$

連想ベクトルは「喜び」，「悲しみ」，「怒り」，「恐怖」，「嫌悪」，「驚き」の基本 6 表情についてそれぞれ 1 つの記憶パターンを 1 つのベクトルで表現し，パターンを 1, 0 にコード化したもので，  $y = (y_1, y_2, \dots, y_n)^T$  で表す．表 2 は基本 6 表情と連想ベクトルの対応表である．また AU の有無を 1, 0 のベクトル  $x = (x^1, x^2, \dots, x^{44})$  で表したものがキーベクトルとなる．そして連想ベクトル  $Y$  と AU 発生率  $X$  から  $\|y - MX\|^2$  が最小となる記憶行列  $M$  が概算される．ここで，出力を 1, 0 のベクトルで表すため量子関数を

$$\Psi(u) = \begin{cases} 1 & \text{if } u > 0; \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

と定義すると，記憶行列とキーベクトルによる想起は次のように表すことができる。

$$z = \Psi(Mx - \theta) \quad (5)$$

閾値  $\theta$  により  $z$  が連想ベクトルに最も近いものを認識結果とし，その感情を出力する。

attribute value	association vector
喜び Happiness (H)	(100000) <sup>T</sup>
悲しみ Sadness (SA)	(010000) <sup>T</sup>
驚き Surprise (SR)	(001000) <sup>T</sup>
恐怖 Fear (F)	(000100) <sup>T</sup>
怒り Anger (A)	(000010) <sup>T</sup>
嫌悪 Disgust (D)	(000001) <sup>T</sup>

表 2. 連想ベクトル

## 4. 合成パート

ここでは表情の合成を実現するための手法を述べた後，2.2 で簡単に説明したモジュール構成に沿って，合成パートにおいて分析パートから受信した感情情報からユーザが相手の感情を正しく理解できる表情を合成する過程を説明する。

### 4.1. 表情の合成を実現する手法

#### 4.1.1. 感情の分類<sup>[3]</sup>

本研究では感情の分類に，感情を「驚き (surprise)，恐怖 (fear)，嫌悪 (disgust)，怒り (anger)，喜び (happiness)，悲しみ (sadness)」の基本 6 感情と呼ばれる 6 つに分類する手法を用いている．この手法は現在感情に関する研究でもっともよく用いられている。

また，これらの感情から引き起こされる表情を基本 6 表情と呼ぶ。

#### 4.1.2 顔の3領域

顔表情は、表情筋と呼ばれる顔の筋肉の収縮および弛緩によって、皮下組織や皮膚表面の変形として表出される。表情筋の収縮および弛緩により変形される部分により、顔は互いに独立した動作が可能な以下の3つの領域に分割することが出来る。

- ① 上部：眉と額
- ② 中部：目と鼻の付け根
- ③ 下部：口と頬

また、これらの3領域に表出する基本6表情と中立(無表情)の表情の組み合わせによって複雑な感情を表す表情や感情の隠蔽を示す表情を表すことが出来る。

図6に顔の3領域を利用した表情「軽蔑」の合成について示す。「軽蔑」の表情を合成する場合は、顔の上部・中部が中立の表情、下部に嫌悪の表情を表出させることで表現することができる。

顔の領域	基本表情
上部	中立
中部	中立
下部	嫌悪



図.6 顔の3領域を利用した表情合成の例(軽蔑)

### 4.2. 表情の合成

#### 4.2.1. 表情の生成

分析パートから受信した感情情報を表情(AUの組み合わせ)に変換する。この際モデルベースから提供される合成側ユーザの個人モデルから基本6表情のAUの組み合わせを読み取り、顔の3領域に基本6表情をそれぞれ当てはめて感情に対する表情のAUの組み合わせを決定する。

#### 4.2.2. 表情筋のシミュレーション<sup>[4]</sup>

人間の表情は表情筋(顔面筋)および口の周りの咀嚼筋と呼ばれる筋肉の収縮・弛緩により引き起こされる。本研究では、これらの筋肉をモデル化し、シミュレーションすることで顔面の形状を変化させる筋肉モデル法を「Muscle Model Generator」で用いる。

これらの表情筋は2章で述べた各AUにそれぞれ対応しており、Expression Generatorにて表情を表出するためのAUの組み合わせが決定すると、Muscle Motion Generatorにて動作させる表情筋を決定することができる。その後Muscle Model Simulatorで感情のレベルを考慮した上で各表情

筋に対応する顔モデル(WFM)の頂点の動きを計算し、グラフィックエンジンにデータを送る。

#### 4.2.3. グラフィックエンジン

筋肉モデル生成部から受信した情報を元にWFMの変形を行う。この際、モデルベースからはユーザの顔モデルの情報とテクスチャが提供される。WFMの変形は表情筋に基づいて行われるため、自然な表情の表出が可能である。

## 5. 個人モデル

本研究では、個人モデルに基づいた表情の認識・合成を行うことで正しい感情の伝達を可能にする。個人モデルとは人間の表情の個人差をモデルとして表したものである。個人モデルには分析側用の個人モデルと合成側用の個人モデルがあり、分析側個人モデルは個性的な表情から一般的な表情を得るために用いる。これにより表情を表出したユーザの正しい感情が得られる。また、合成側個人モデルは一般的な感情から個性的な表情を作るために用いる。これにより表情認識の際の個人差を埋めることが可能になる。

### 5.1 分析側個人モデル

分析側の個人モデルには、AUパターンと表情パターンの2種類存在する。そこであらかじめ、基本6表情の連続画像を記録し、オプティカルフロー計算、ヒストグラム化、テンプレート化処理を行い個人モデルの作成を行う。AUパターンはAUの動きを表したテンプレートデータとそれに対するAU名を対で記録したものでAUの認識において学習データとして利用する。また表情パターンはユーザーの表情を構成するAUの組み合わせを示した各表情におけるAU発生率を個人モデルとして表情の認識に利用する。

### 5.2 合成側個人モデル

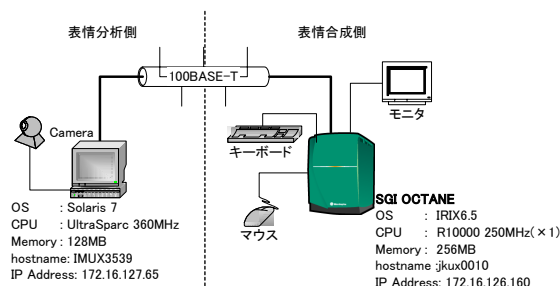
合成側の個人モデルも大きく分けて2種類存在する。ひとつは表情に対する個人モデルで、もうひとつは顔モデルに関する個人モデルである。本研究では特に表情に関する個人モデルに注目しており、これはAUとそのレベルをもって記述する。具体的には基本6表情表出時のAUの組み合わせを顔の上、中、下の3つに分割した3領域ごとに記述し、感情情報-AU変換部にて個人モデルをロードして表情の合成を行う。一般的な表情も個人モデルの一つと考える、この際のAUの組み合わせはEkmanの定義したAUの組み合わせを用いる。また、表情の表出度合いの調整を行うために個人モデルには表情表出度合い

の傾向も記述する。表情の表出度合いは標準の表出度合いからのプラスとマイナスで記述する。

顔モデルに関する個人モデルには個人ごとのWFMとテクスチャが提供される。

## 6. 実装と評価

本研究におけるプロトタイプシステムを、SGI社製グラフィックスワークステーションOCTANE上に実装している。(図7)



開発言語はC言語で、3DCGの合成にはOpenGLを使用している。このプロトタイプシステムの分析側においてはカメラから顔画像を取り込み、表情変化の検出を行うモジュールについて実装ができていないため、現在は連続画像における前後二枚を入力として特徴抽出、AU認識、表情認識を行い、感情情報とAU、およびそのレベルを出力している。またAUの認識においては、表情の表出において使用頻度の高い18のAUについて、表情の認識では、基本6表情とは一部異なるが「喜び」、「安心」、「不安」、「嫌悪」、「苦痛」、「驚き」の表情についての認識が実装されており、その他のAU、表情については今後実装を進めていく予定である。

合成側では現在基本6表情の表出と拡張10表情、それらの表情ごとにAction Unitによる眼球の制御、顔の角度の制御が可能である。また、現在は顔モデルのワイヤフレームの変形をキーノードと呼ばれる頂点の移動により実現している。キーノードは顔の中で特に動きの大きい頂点であり、各Action Unitごとにキーノードを設定できる。

現段階では、本システムの評価を分析パートと合成パートで分けて行う。

分析処理の性能評価項目として、AUの認識率と表情の認識率が挙げられる。AUの認識率はオプティカルフローを計算し、ベクトル表現にした画像とAUの認識処理の出力したAUの一致から求める、また表情の認識率は被験者の感情と表情の認識処理の出力した表情の一致から求め、性能評価を行う。

合成パートにおける評価は分析システムから分析側ユーザの正しい感情が送信されているものとして行う。本システムでの評価項目としては、個人モデルを用いた場合の感情の伝達率(精度)と個人モデルを用いない場合の感情の伝達率を比べ、個人モデルの有効性を立証する。



## 7. まとめと今後の課題

本稿では感情を正しく伝達するコミュニケーションシステムについて、実現方法と現段階での実装状況について述べた。本システムの利用により、相手の感情を間違えて受けとめることが減少し、スムーズなコミュニケーションを図ることができる。

今後の課題としては、分析パートから合成パートまでの一連の処理を考えた実装を行うことが挙げられる。その後、作成したプロトタイプシステムを用いて評価を行う予定である。また、現在はキーノードを用いている部分に、表情筋をシミュレートする手法を用いるための筋肉モデル生成部の詳細な設計、実装を行うことが挙げられる。

## 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学会, Vol.95, No.53, pp.19-24, May 1995.
- [2] Y.Dai, Y.Shibata, T.Ishii, K.Hashimoto, K.Katamachi, K.Noguchi, N.Kakizaki, D.Cai, "AN ASSOCIATE MEMORY MODEL OF FACIAL EXPRESSIONS AND ITS APPLICATION IN FACIAL EXPRESSION RECOGNITION OF PATIENTS ON BED" IEEE proceeding of ICME, pp772-775, (August. 2001)
- [3] P・Ekman, W・V・Friesen 著 工藤 力 訳 編: 表情分析入門, 誠信書房 (1987)
- [4] 楊大昭: "バーチャル空間における顔表情の動的合成に関する研究", 京都大学大学院エネルギー科学研究科 修士論文 (2000)