微妙な表情の生成を可能とするインタラクティブ顔画像生成システム

A System Using An Interactive Interface for Synthesizing Facial Images with Subtle Expressions

磯野 勝 宣[†], 山本 光 穂^{†*}, 正会員 尾田 政 臣^{†,††},
赤 松 茂[†]

Katsunori Isono[†], Mitsuo Yamamoto^{†*}, Masaomi Oda^{†,††} and Shigeru Akamatsu[†]

Abstract Available methods for synthesizing facial expressions synthesize basic expressions. Few methods synthesize subtle facial expressions such as a knavish smile or a cold smile. We developed a system for synthesizing subtle facial expressions. It uses an interactive interface based on a genetic algorithm. We use the five main parts of the face: eyebrows, eyes, mouth, and contours. These parts are deformed independently. Using this system, subtle facial expressions can be synthesized by iterating a deformation process for each part. Experiments showed that this system can efficiently synthesize subtle facial expressions.

キーワード: 顔表情, 表情差分, 対話型インタフェース, 遺伝的アルゴリズム, 主成分分析

1. ま え が き

任意方向・任意表情の顔画像の生成技術は、感性豊かなヒューマンインタフェースの実現だけでなく、臨場感通信・テレビ会議システムに代表される、バーチャルリアリティ技術を用いたマルチメディアコミュニケーションなど、様々なアプリケーションに応用することができる¹⁾²⁾。本論文では、ある人物の任意の表情の実画像から、それとは異なる任意の表情の顔画像や同一の表情カテゴリ内での印象の異なる多様な表情を生成する手法について検討を行う。

これまで、筆者らは、ある表情とそれと異なる任意の表情間の対応する特徴点座標の差分として表現される表情カテゴリの変化に関する知識を用いて、任意の表情の顔の3次元構造を復元することで、任意方向・任意表情の顔画像の生成を行ってきた³⁾。この手法では、表情

のサンプルデータベースから、2つの表情カテゴリ間において対応している特徴点の座標の差分をデータベースの平均として求め、それを与えられた表情の特徴点座標に加えることで、任意の表情を生成している。そのため、生成された表情は画一的なものとなり、また、必ずしもすべての人物に対して自然な画像が得られるとは限らない。

また、表情は、例えば「笑い」を考えた場合、「作り笑い」、「不敵な笑い」、「冷笑」や、言葉では表現できない微妙なニュアンスの変化など、同一の表情カテゴリであっても様々な表情があり、見る人によってその受け取り方は異なる。例えば、同じ表情を見ても、ある人はそれを「作り笑い」と感じ、またある人は「不敵な笑い」と感じてしまう。前述の手法³⁾では、このような見る側（観察者）の主観に基づく表情画像を生成することはできない。観察者の望みの顔を生成するためには、観察者自身の主観に基づく評価が必須であり、観察者、つまり人間を介在させたシステムが必要となってくる。

そこで、本論文では、平均差分を用いて生成された表情を人間を介在させたシステムにより変形・修正を行うことにより、観察者の主観に基づく表情を生成する。平均差分といった統計情報を用いた手法は、多くの人に支持される顔画像の生成が目的であるのに対し、提案手法は、個々の観察者の主観を反映した、個々の観察者のイメージにマッチした顔画像の生成を目的としている。

1999年3月23日、電子情報通信学会ヒューマン情報処理研究会で発表
1999年12月1日受付、2000年5月10日最終受付、2000年5月15日採録

[†]株式会社 ATR 人間情報通信研究所
(〒619-0288 京都府相楽郡精華町光台 2-2, TEL 0774-95-1013)

^{††}立命館大学 文学部
(〒603-8577 京都市北区等持院北町 56-1, TEL 075-465-1111)

*長岡技術科学大学

[†]ATR Human Information Research Laboratories
(2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan)

^{††}Ritsumeikan University
(56-1, Tojiin Kitamachi, Kita-ku, Kyoto 603-8577, Japan)

*Nagaoka University of Technology

観察者の主観を反映させるアプローチとしては、Facial Action Coding System (FACS) ⁴⁾ の表情変形規則に基づいて、直接ワイヤフレームモデルを変形させることで表情を生成する手法がある ⁵⁾⁶⁾。中川らのシステム ⁶⁾ では、FACS の記述単位である Action Unit (AU) をワイヤフレーム上の特定の格子点の移動として定義しなおし、AU の選択とその強度調整をインタラクティブに行うことで表情を生成している。つまり、観察者が画面を見ながら、表情パラメータを試行錯誤的に求めながら、自分のイメージに合った顔を生成していく。この手法では、各 AU ごとに細かな設定が行えるため、対象人物に対して自然で多様な表情が生成できる。

このような、観察者が直接顔を変形させながら表情を生成する手法では、あらかじめ生成したい顔のイメージを想起しておく必要がある。しかし、例えば、「不敵な笑い」、「冷笑」といった、観察者が想起したあいまいな顔イメージを描画したり、言語的に報告することは困難である。このような印象の変化は、顔の微妙な変化に起因しており、上記のような手法を用いた場合、各 AU の機能を詳しく十分に把握していなければならず、経験が必要である。つまり、上記手法では、笑い、怒り、悲しみといった、少数の特定の顕著な表情は容易に表現できるものの、その中間的で微妙な表情の変化を定量的に記述し、その表情を生成することは困難であると考えられる。

ここで、あいまいな印象語で表現されたような顔のイメージを画像として明確に表現することは困難であるが、提示された画像に対して、それが自分のイメージに合っているかどうかを判断することは容易である。そこで、画像を次々に“提示－評価－生成”し、対話型画像検索インタフェースに基づき、インタラクティブに画像を検索・生成するシステムがいくつか提案されている ^{7)~10)}。これらのシステムでは、次々に提示される画像に対して、観察者が自分のイメージに合っているかどうかの評価を行う。システムは、その評価に基づき、新たな画像を生成し、次の候補画像として提示する。この操作を繰り返すことにより、観察者のイメージに合った画像の生成を行うことができる。この手法では、観察者は画像を見て判断するだけでよく、表情のパラメータ等を意識する必要はない。また、多くの画像を見ることで、これまで自分のイメージにはなく、かつ、より自分のイメージに合った画像の発見・想起・生成を行うことができると期待される。

筆者らは、文献 8) で、3次元データを対象として、「柔和な顔」、「怖い顔」といった、感性的な印象語で表現されるイメージを生成するシステムを開発した。この手法では、限られたデータから多様な画像を生成するために、主成分分析 (PCA) を用い、選択候補画像決定方式として、遺伝的アルゴリズム (GA) ¹¹⁾ を用いた。PCA は顔画像の生成や認識技術などとして使われ、一定の成果を

収めている ^{12)~14)} こと、さらに、2次元濃淡顔画像を対象として、PCA によって求められた特徴を用いて、感性的な情報をうまく表現できることが示されている ¹⁵⁾ ことなどから、画像生成に PCA を用いた。また、大域的探索を行うことができ、局所解に陥る可能性が小さいという利点から、選択候補画像決定方式として GA を用いた。GA を用いた研究 ^{8)~10)} では、このような GA の性質と観察者の主観的評価を組合せることにより、広い探索空間から、観察者のイメージに合った画像の生成を行うことを可能にしている。

本論文では、文献 8) で提案された対話型インタフェースを用いて、これまで困難であった、観察者の主観による微妙なニュアンスの異なる表情など、多様な表情画像の生成を可能にするシステムの提案を行う。

文献 8) では、顔全体を 1 つのデータとして扱っていたため、ある程度イメージが固まった段階で、最終的に満足いく画像を得るために、操作をさらに繰り返した場合、変化させたくない部品までも変化してしまい、目だけを変えたい、口だけを変えたいといった部品レベルの変形ができなかった。このため、生成画像の満足度が下がってしまうことも考えられ、その場合、再び満足度の高い画像を得るために、さらに操作を繰り返す必要がある。これは、作業効率の低下を招くだけでなく、満足度の高い画像の生成を困難にする要因とも成り得る。さらに、小川ら ¹⁶⁾ は、部品レベルの変化だけでも表情の変化として認知できることを示した。つまり、ある部品を変化させるだけで、その表情に対して観察者が受ける印象は変化する。以上より、より満足度の高い画像を生成するためには、各部品ごとに変形できる機能が必要であると言える。

そこで、本論文では、顔を眉、目、鼻、口、輪郭の 5 つの部品に分けて考え、各部品ごとに変形・修正を行う。なお、文献 16) では、輪郭については検討されていないが、輪郭の変化も表情の重要な要因であることが予想されるので、今回、部品として輪郭も含めることとした。部品を分けて考えることにより、各部品を独立に操作することができ、例えば、目はイメージ通りだが、口はイメージに合わないといった場合など、その重要度や生成の難易度に合わせて、変化させたい部品だけを操作することが可能となる。さらに、ある部品のみを変化させた場合の印象の変化など、心理学的側面からの研究にも応用することができる。

また、今回は、観察者の主観に依存する多様な表情画像の生成を目的とし、顔の向きの変化については扱わないこととする。そのため、顔データとしては、3次元データではなく、2次元正面画像を用いることとする。

以下、2章では、主成分分析 (PCA) による顔画像生成方法について述べる。3章で、提案システムの説明を行い、4章で顔画像生成実験結果を示す。

図 1 入力データ。(a) 入力画像。対象人物の表情 A の顔画像。
(b) 形状とテクスチャ。特徴点を頂点とする三角形パッチの
集合 (形状) と画素の集合 (テクスチャ) に分けて考える。
Input data. (a) Input image. (b) Shape and texture.

2. 任意表情顔画像の表現

本論文で提案するシステムは、図 1(a) のような、任意の人物についての、表情 A (通常は中立表情) の顔画像 (表情 A 画像) が与えられた時、同じ人物の表情 B (笑いなど) の顔画像 (表情 B 画像) を生成する。表情変化は、形状の変化とそれに伴うテクスチャの変化として表現されるので、顔画像を顔の特徴点を頂点として生成される三角形パッチの集合 (形状) と、各画素の輝度値 (RGB 値) の集合、つまり、画像そのもの (テクスチャ) の 2 つに分けて考え、表情変化時の形状とテクスチャを別々に生成する。図 1(b) に、表情 A 画像から得られる、表情 A の形状 (表情 A 形状) とテクスチャ (表情 A テクスチャ) を同時に示す。表情 B の形状 (表情 B 形状) とテクスチャ (表情 B テクスチャ) が得られれば、形状にテクスチャをマッピングすることで、対象人物の表情 B 画像を得ることができる。

ここで、多様な顔画像を生成するとき、イメージの生成過程において、形状とテクスチャを同時に変化させたのでは、印象の変化が形状の変化による影響なのかテクスチャの変化による影響なのか不明確なものになってしまう。顔に対する年齢や性差などに関する心理学的研究では、形状とテクスチャの影響を分離して検討しており、表情生成においても、画像から受ける印象の変化がどちらの影響によるものなのかを分けて考える必要が出てくる。そこで、今回は、形状の変化のみによる印象の変化を考え、テクスチャは、あらかじめ生成される表情 B テクスチャのみを用い、テクスチャは一定とする。

2.1 表情形状の生成

表情 A から表情 B への表情変化を考えたとき、表情変化による顔形状の変化を文献 3) と同様に、表情 A 画像 $\mathbf{m}^A$ と表情 B 画像 $\mathbf{m}^B$ の対応している特徴点の座標の差分として捉える。この差分を差分形状と呼び、 $\delta\mathbf{m}$ で表すすると、表情 B 形状は、表情 A 形状と差分形状との和で表すことができる。

$$\mathbf{m}_i^B = \mathbf{m}_i^A + \delta\mathbf{m}_i \quad (1)$$

i : 特徴点番号

つまり、差分形状 $\delta\mathbf{m}$ をいろいろ変化させることで、多様な表情画像を生成することができる。

多様な差分形状 $\delta\mathbf{m}$ を取得するために、画像サンプルデータベースから得られる差分形状に対して PCA を行う。画像サンプルデータベースは、正面から撮影された、多数の人物の表情 A と表情 B の画像のペアが格納されており、両目尻の中点が画像の中心に位置するようにし、両目尻を結んだ直線が水平になるように回転させ、顔の位置と方向に関して正規化を行う。

ここで、文献 8) では、顔全体を 1 つのデータとして PCA を行っていたが、例えば、目はイメージ通りだが、口はイメージに合わないといった場合に対応することができなかった。そこで、今回は、顔形状を眉、目、鼻、口、輪郭の 5 つの部品に分けて考え、各部品ごとに別々に PCA を適用する。これにより、各部品を独立に変形させることが可能となる。顔全体としては、5 つの部品の組合せによって表現される。

画像サンプルデータベースから得られる各部品における差分形状に対して PCA を適用することにより、部品 K における固有ベクトル $\mathbf{u}^{(K)}$ が求まる。部品 K の画像サンプル内の平均差分形状を $\overline{\delta\mathbf{m}^{(K)}}$ 、表情 A 形状を $\mathbf{m}^{A(K)}$ とすると、部品 K における差分形状 $\delta\mathbf{m}^{(K)}$ および表情 B 形状 $\mathbf{m}^{B(K)}$ は以下の式で表される。

$$\delta\mathbf{m}^{(K)} = \sum_j \omega_j^{(K)} \mathbf{u}_j^{(K)} + \overline{\delta\mathbf{m}^{(K)}} \quad (2)$$

$$\mathbf{m}_i^{B(K)} = \mathbf{m}_i^{A(K)} + \delta\mathbf{m}_i^{(K)} \quad (3)$$

$\omega_j^{(K)}$: 重み係数

i : 特徴点番号

j : 固有ベクトル番号

K : 部品番号

重み係数 $\omega_j^{(K)}$ の値を制御することにより、部品 K の様々な差分形状が得られ、それを表情 A 形状に加えることにより多様な表情 B 形状を得ることができる。

以上より、限られた顔画像サンプルから多様な顔画像の生成が行え、また、大規模なデータベースを構築することができる。

2.2 表情テクスチャの生成

本論文では、形状の変化による影響のみを考え、テクスチャの変化による影響は考えない。表情 B 画像は、システムにより生成された表情 B 形状に表情 B テクスチャをマッピングすることによって得られる。今回、表情 B テクスチャの生成は、テクスチャ変化の個人差の影響を小さくするために、差分形状と同様に、データベース内のテクスチャの差分の平均 (平均差分テクスチャ) を取得し、それを表情 A テクスチャに加えることで生成する。

差分テクスチャは、画像サンプルデータベース内の表情 A 画像と表情 B 画像から得られるテクスチャの、対応する各画素における輝度値 (RGB 値) の差分を取る

図 2 オリジナルテクスチャと表情 B テクスチャ。オリジナルテクスチャに、画像サンプルデータベースから抽出された平均差分テクスチャを加えることで、表情 B テクスチャを得る。
Original texture and expression-B texture.

ことによって得られる。

表情間の差分テクスチャを得るには、それぞれのテクスチャの画素同士の対応が取れていなければならない。しかし、顔の形状は個人によって異なり、また、同一人物の画像であっても、表情 A と表情 B では形状が異なるため、テクスチャの画素同士の対応が取れていない。そこで、画素の対応を取るために、データベース内の表情 A の平均形状を基準形状とし、それぞれの表情のテクスチャを基準形状の各三角形パッチごとにマッピングする。なお、表情変化におけるテクスチャの変化は顔内部のみで生ずるものとし、頭髮や背景のテクスチャは変化しないものとする。

画像サンプルデータベースの表情 A、表情 B の各テクスチャを基準形状にマッピングすることにより、画素同士の対応の取れたテクスチャが得られる。さらに、各表情ごとに平均を求め、それらを平均基準テクスチャ $\overline{\mathbf{T}^A}$ 、 $\overline{\mathbf{T}^B}$ と呼ぶこととする。

平均基準テクスチャ $\overline{\mathbf{T}^A}$ 、 $\overline{\mathbf{T}^B}$ は、画素同士の対応が取れているので、対応する画素において、それぞれ輝度値の差分を取ることで、平均差分テクスチャを得ることができる。

$$\delta \mathbf{T} = \overline{\mathbf{T}^B} - \overline{\mathbf{T}^A} \quad (4)$$

得られた平均差分テクスチャ $\delta \mathbf{T}$ を対象人物の表情 A 形状にマッピングする ($\delta \hat{\mathbf{T}}$) ことで、対象人物の表情 A テクスチャ (オリジナルテクスチャ: $\mathbf{T}^{ori}$) との画素間の対応を取る。対応する画素間において、 $\delta \hat{\mathbf{T}}$ と $\mathbf{T}^{ori}$ を加えることにより表情 B テクスチャ $\mathbf{T}^E$ を得る。

$$\mathbf{T}^E = \mathbf{T}^{ori} + \delta \hat{\mathbf{T}} \quad (5)$$

なお、平均差分テクスチャを加えて生成された表情 B テクスチャを平均差分形状を加えて得られる表情 B 形状にマッピングした画像に対して、「自然な画像かどうか」の主観的評価実験をあらかじめ行い、自然な画像が得られていることを確認した。図 2 に対象人物の表情 A テクスチャと表情 B テクスチャを示す。

3. 顔画像生成システム

図 3 に任意表情顔画像生成システムを、図 4 に画像生成処理の流れをそれぞれ示す。

画像生成処理の流れは、初めに、表情 A 形状に平均差分形状を加えて得られる表情 B 画像 (基準表情顔画像) を表示する。観察者は、変形させたい部品を選択する。部品が選択されると、システムはその部品のみをランダムに変形させた 12 枚の画像を生成し提示する。これらの画像は PCA によって得られた対象部品の固有ベクトルにさまざまな重み係数を掛けた加重和によって得られる。観察者は新たに提示された画像が、自分のイメージに合っているかどうかの評価を行う。システムは、その評価に基づいて、固有ベクトルに対する新しい重み係数を GA により逐次算出する。そして、得られた重み係数を用いて再び画像を生成し、提示する。これらの「画像提示—部品選択—評価—画像生成」を観察者が満足のいく画像が現れるまで繰り返し行うことにより、観察者がイメージする顔画像により近い画像が得られるようになる。

ここで、観察者による画像の評価方法であるが、文献 8) では、提示されたすべての画像に対して評価点を与えている。しかし、この方法では、観察者の作業負荷が大きいという問題があった。そこで、本システムでは、提示画像の中から、観察者のイメージに最も合っている画像のみを選択し、その評価点を最高評価とし、それ以外の画像には、一律に低い評価点 (< 最高評価) を与えるものとする。選択された画像以外はすべて一律の評価点を与えることで、生成効率が下がる可能性があるが、すべての画像に対して評価点を与える作業にはかなりの時間を要しており、それを省くことで、結果的に、作業時間の短縮、作業効率の上昇が図られると期待される。

今回、逐次的に重み係数を決定する方法として GA を用いる。本論文で用いたアルゴリズムは以下の通りである。

(1) コード化

染色体としては、各部品の重み係数を画像生成に用いる固有ベクトル数分並べたものを用いる。今回、各部品の重み係数の変化域は $-3\sigma_j^{(K)} < \omega_j^{(K)} < 3\sigma_j^{(K)}$ とし、その変化域に充分に対応するため、各重み係数を 8 ビットで表現した。したがって、染色体の長さは各部品に対して 8 ビット $\times$ (生成に用いる固有ベクトル数) となる。ここで、 $\sigma_j^{(K)}$ は画像サンプルに対する、部品 K の第 j 固有ベクトルの重み係数の標準偏差である。

(2) 適応度

GA は適応度の高い染色体を確率的に優先的に生存させながら集団を生成し、世代を更新する。本システムでは、その評価を人間の主観的評価によって決定する。観察者によって選択された顔画像における部品の評価点を

図 3 任意表情顔画像生成システム。
Facial images with arbitrary expression synthesizing system.

7 点（最高評価），それ以外の評価点を 2 点（＜最高評価）とし，その評価点を適応度とした。

(3) 選 択

親の選択はルーレット選択方式（適応度に比例した確率で選択し，重複を許す方式）に，エリート選択方式を組合せた方法を用いる。エリートとして選ばれる画像は，観察者によって選択された画像を適用する。したがって，観察者により選択された部品はそのままの形で次世代に残ることになる。

(4) 交 叉

交叉法としては一様交叉を用いる。一様交叉は，交叉時にランダムマスクをかけ，それによって，各遺伝子座においてどちらの親の遺伝子を受け継ぐかをランダムに決定する方法である。

(5) 突然変異

各々の部品に対して，1 世代につき 1 個の割合で突然変異が起こるように，変異確率を 0.002 とした。

4. 任意表情画像生成実験

表情 A 画像（図 1(a)）および，画像サンプルデータベース内の画像に対して，特徴点の取得を行う。特徴点の取得は，複数の画像から自然な合成画像を生成するためのシステムとして ATR で開発された FUTON システム¹⁷⁾のサンプリングツールを用いて手入力で行った。特

図 4 画像生成処理の流れ。
Flowchart of image synthesizing process.

徴点は，特徴点どうしを結んだ直線が，顔の部品（眉，目，口など）の境界（エッジ）をなぞるように設定されている（図 1(b)）。FUTON システムでは，比較的安定に抽出できる特徴点を基準にして，補助線を出すなど，作業者を支援することによって，抽出基準のあいまいな特徴点についても異なる人物の顔の間で一貫性のある安

定した特徴点抽出を可能にする¹⁸⁾。さらに、特徴点の取得は同一人物がすべてのデータに対して行い、特徴点取得基準が一定となるようにした。

今回用いた画像サンプルデータベースは、20代女性61名分の「中立表情」、「笑い」のデータを格納しており、表情Aを「中立表情」とし、表情Bは「笑い」とした。なお、これらのサンプル画像に対しては、あらかじめ、主観的評価実験が行われており、表情が正しく表出されていることが確認されている¹⁹⁾。

4.1 実験方法

本論文で提案されたシステムを用いて、被験者による画像生成実験を行った。実験は、「不敵な笑い」および「寂しそうな笑い」を生成課題とした。被験者は、男女大学生16名である。

(1) 実験1

以下の3つの生成条件で「不敵な笑い」および「寂しそうな笑い」の生成を行ってもらった。中立表情の実画像は、いつでも見られるように常時提示しておいた。生成は被験者の満足のいく表情画像が生成できた時点で終了とした。画像の生成後、生成された画像をそれを生成した被験者自身によって、7段階評価を行ってもらった。

- cA: 顔を5つの部品に分けずに、顔全体を変形させながら表情を生成する。
- cP: 顔を5つの部品に分け、部品毎に変形を行いながら表情を生成する。
- cAP: 「cA」と「cP」を組合せ、初めに顔全体を変形し、ある程度イメージが固まったところで、部品毎に変形を行い表情を生成する（「cA」から「cP」への移行は被験者の判断で行う）。

(2) 実験2

本システムに関して簡単なアンケート調査を行った。

4.2 結果と考察

図1を入力画像とし、被験者A、Bによって生成された、「不敵な笑い」および「寂しそうな笑い」を図5に示す。

実験1での評価点の平均値を表1に示す。いずれの表情においても、評価点の平均値は「cAP」が最も高く、次いで「cP」、「cA」という順であった。LSD法による検定の結果、「不敵な笑い」では、「cA」と「cAP」の間に有意差が見られ ($p < 0.01$)、「寂しそうな笑い」では、「cA」と「cP」および「cA」と「cAP」の間に有意差が見られた (共に $p < 0.05$)。また、実験2でのアンケート調査において、「cA、cP、cAPの中で、最も作りやすい方法はどれか」という質問に対し、16名中12名が「cAP」と答えた (4名は「cP」)。その理由としては、「cA」は、全体のイメージは捉えやすいが、細かな修正ができない。逆に「cP」は細かな修正は可能であるが、全体のイメージが捉えにくいという意見が多く聞かれた。つまり、全体のみ、または、部品のみで生成を行う方法には、それ

図5 画像生成例。本システムを用いて生成された (i) 不敵な笑い、(ii) 寂しそうな笑い。
Example of synthesized images.

表1 評価点の平均値
Average rating value.

不敵な笑い			寂しそうな笑い		
cA	cP	cAP	cA	cP	cAP
4.625	5.125	5.688	4.750	5.563	5.688

ぞれ一長一短があり、どちらか一方の方法では、ある程度の満足度を得る画像を生成することは可能であるが、それらを組合せる方法を用いることで、より満足度の高い画像を生成することができるといえる。また、「自分のイメージ通りの表情が生成できたか」という質問に対しては、16名中14名から「ほぼイメージ通り」という回答を得ており、本システムにより、被験者のイメージにマッチした画像の生成ができたことを示している。

「表情生成時（不敵な笑い）に特に注目した部品はあったか、ある場合はその部品は何か（複数回答可）」という質問に対しては、すべての被験者が注目部品があると答え、各部品の注目度は、目：9名、口：8名であり、その他の部品は各1名であった。さらに、「表情生成が難しかった部品は何か（複数回答可）」という質問に対しても、目：10名、口：5名（その他の部品は1または0名）であった。これらより、部品によって、その重要度が異なることが確認され、重要な部品ほど、生成に時間をかけていることがうかがわれる。したがって、各部品を分割して扱うことのできる機能を追加することで、重要な部品の生成に集中でき、より満足度の高い表情生成につながると考えられる。

また、今回の実験での「cAP」条件での、最終的な結果が得られるまでの平均生成時間、平均世代数は、「不敵な笑い」では、6分、34.9世代、「寂しそうな笑い」では、

5分30秒, 35.6世代と表情による明確な差は認められなかった。

以上より、「不敵な笑い」、「寂しそうな笑い」という生成課題に対して、本システムは有効であるといえ、5つの部品を別個に変形・修正できる機能を追加することで、より満足度の高い画像の生成が可能であることがわかった。

5. む す び

本論文では、与えられた任意の表情を持つ実画像から、それとは異なり、かつ微妙な表情の生成を可能とするシステムの構築を検討した。微妙な表情変化の実現は、あらかじめ、画像サンプルデータベースから、表情カテゴリ一間の対応する特徴点の座標の差分を取得し、得られた差分を基に、与えられた画像の特徴点の座標値を変化させることによって得られる。観察者の主観に合った表情を生成するため、GAを用いた対話型インタフェースを採用し、また、顔を眉、目、鼻、口、輪郭の5つの部品に分割し、各造作毎に変形を行えるようにすることで、微妙な表情の違いを容易に生成することを可能とした。

提案システムを用いて、顔全体を変形する(cA)、部品を用いて変形する(cP)、顔全体を変形した後、部品を用いて修正を行う(cAP)の3つの生成方法で、画像の生成実験を行った結果、「cAP」が最も満足度の高い画像が得られた。顔全体の変形と部品毎の変形を組合せることで、より満足度の高い画像を得ることができ、部品毎の変形機能の追加は観察者の主観に依存する感性的イメージの生成という課題に対して有効であることがわかった。

本システムは、画像サンプルデータベースを適宜交換することにより、表情生成だけでなくモンタージュなどの類似顔の生成にも応用することができる。また、心理学的感性情報処理の研究における刺激の作成などにも用いることができると考えられる。

今回は、形状の変化による影響のみを考えたが、反対に、形状を一定とし、テクスチャデータに対してのみPCAを行うことによって、テクスチャの変化による影響を見ることもできる。さらに、形状、テクスチャ両方に対してPCAを行い、それぞれの係数を変化させ、得られた結果を組合せることによって、より多様な顔画像を生成することが可能であり、今後の課題である。

〔文 献〕

- 1) 赤松茂：“コンピュータによる顔の認識 — サーベイ —”，信学論 (D-II)，J80-D-II，8，pp. 2031-2046 (1997)
- 2) 長谷川修，森島繁生，金子正秀：“「顔」の情報処理”，信学論 (D-II)，J80-D-II，8，pp. 2047-2065 (1997)
- 3) 磯野勝宣，赤松茂：“任意表情の顔の3次元構造の生成 — 3d および画像サンプルにおける表情差分ベクトルの利用 —”，信学論 (D-II)，J82-D-II，10，pp. 1729-1738 (1999)
- 4) P. Ekman and W.V. Friesen：“Facial Action Coding System”，Consulting Psychologists Press (1978)
- 5) 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)

- 6) 中川督之，大場敏文，デロワオリビエ，金子正秀，原島博：“3次元モデルを用いたインタラクティブ顔画像合成システム”，信学ソ大，SA-10 (1996)
- 7) 尾田政臣：“人間のイメージ形成過程の特性を利用した画像検索システム”，情処学論，35，7，pp. 1449-1456 (1994)
- 8) 磯野勝宣，尾田政臣，赤松茂：“遺伝的アルゴリズムを用いた顔画像生成システム — 感性的イメージの生成 —”，信学論 (D-II)，J82-D-II，3，pp. 483-493 (1999)
- 9) C. Caldwell and V.S. Johnston：“Tracking a criminal suspect through “face-space” with a genetic algorithm”，In Proceedings of International Conference on Genetic Algorithms，pp. 416-421 (1992)
- 10) 長尾光悦，三田村保，山本雅人，大内東：“遺伝的アルゴリズムを用いた顔画像検索アルゴリズム”，平成8年度電気関係学会北海道支部連合大会，p. 18 (1996)
- 11) 北野 宏明編：“遺伝的アルゴリズム”，産業図書 (1993)
- 12) 永田明德，岡崎透，崔昌石，原島博：“主成分分析による顔画像の基底生成と空間記述”，信学論 (D-II)，J-79-D-II，7，pp. 1230-1235 (1996)
- 13) T. Vetter and N. Troje：“A separated linear shape and texture space for modeling two-dimensional images of human faces”，Technical Report No.15，Max-Planck-Institut für biologische Kybernetik (Apr. 1995)
- 14) 櫻井保志，岩佐英彦，竹村治雄，横矢直和：“距離画像を用いた顔方向と表情によらない顔認識”，画像の認識・理解シンポジウム講演論文集 I，pp. 181-186 (1996)
- 15) 尾田政臣，赤松茂，深町映夫：“あいまい顔イメージの検索に対するKL展開の適用性”，信学論 (A)，J79-A，2，pp. 288-297 (1996)
- 16) Tokuko Ogawa and Masaomi ODA：“How facial parts work in facial-expression judgement”，ATR Symposium on Face and Object Recognition '99，pp. 109-110 (1999)
- 17) 蒲池みゆき，向田茂，吉川左紀子，加藤隆，尾田政臣，赤松茂：“顔・表情認知に関する心理実験のための顔画像合成システム — FUTON system —”，信学技報，HIP97-39，pp. 73-80 (1998)
- 18) 向田茂，蒲池みゆき，赤松茂：“顔画像合成システム (FUTON system) におけるマニュアルサンプリングの評価”，信学技報，HIP99-49，pp. 13-18 (1999)
- 19) 小川徳子，尾田正臣，吉川左紀子，赤松茂：“3つの顔角度の表情画像に対する評価 — 表情画像データベースの構築 —”，信学技報，HIP97-17 (1997)

磯野 勝宣 ^{いその かつのり} 平4 慶大・理工・機械卒。平6 同大学院生体医工学専攻修士課程了。同年(株)インテック入社。平8より(株)ATR人間情報通信研究所に就任。顔イメージの計算機による表現と生成に関する研究に従事。現在、インテック・ウェブ・アンド・ゲノム・インフォマティクス(株)。

山本 光穂 ^{やまもと みつお} 平11 長岡技科大・工・電子機器卒。平10.11~平11.2(株)ATR人間情報通信研究所学外実習生。現在、同大学院修士課程在学中。生体情報計測に興味を持つ。

尾田 政臣 ^{おだ まさおみ} 昭46 北大・工・電気卒。昭51 同大学院博士課程了。同年日本電信電話公社(現NTT)武蔵野電気通信研究所入所。昭63よりATR視覚機構研究所に就任。以来、人間の図形・文字の認知特性の研究、ヒューマンインターフェースの研究に従事。現在立命館大学文学部教授。(株)ATR人間情報通信研究所客員研究員。工学博士。正会員

赤松 茂 ^{あかまつ しげる} 昭50 東大・工・計数卒。昭52 同大学院修士課程了。同年、日本電信電話公社(現NTT)入社。以来、文字認識、画像処理の研究に従事。昭60~61米国カリフォルニア大学Irvine校客員研究員。平4より(株)ATR人間情報通信研究所第二研究室長。現在に至る。平11より東京工業大学大学院総合理工学研究科客員教授を兼務。工博。平11より本学会ヒューマン情報処理研究専門委員会委員長。パターン認識のヒューマンインターフェースへの応用、感性情報処理、特に人間とコンピュータによる顔情報処理に関心をもつ。