

講座 人間とコンピュータによる顔表情の認識(4・完)コンピュータによる顔表情認識の展望：人間による表情認知に学ぶ

赤松, 茂 / AKAMATSU, Shigeru

(出版者 / Publisher)

電子情報通信学会

(雑誌名 / Journal or Publication Title)

電子情報通信学会誌 / 電子情報通信学会誌

(号 / Number)

1

(開始ページ / Start Page)

54

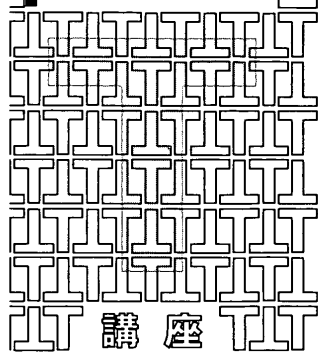
(終了ページ / End Page)

61

(発行年 / Year)

2003-01-01

Tutorial Series



人間とコンピュータによる顔表情の認識 [IV・完]

 —— コンピュータによる顔表情認識の展望：
 人間による表情認知に学ぶ ——

赤松 茂

 赤松 茂 正員 法政大学工学部システム制御工学科
 E-mail akamatsu@k.hosei.ac.jp

Recognition of Facial Expressions by Human and Computer [IV : Finish] : Toward Computer Recognition of Facial Expressions Consistent with the Perception by Human. By Shigeru AKAMATSU, Member (Faculty of Engineering, Hosei University, Koganei-shi, 184-8584 Japan).

1. ま え が き

人間が顔の表情から相手の感情を読み取っているように、コンピュータによって顔画像から感情を認識しようとする表情認識の研究は、ここ数年の間に内外で大いに活発化してきた。多くの制約を設けた実験室環境の下でという条件付きとはいえ、今日では顔画像の入力から表情が表している感情カテゴリーの認識までを自動かつほぼリアルタイムで処理できる「動くシステム」の構築に成功している事例も少なくない。しかしそれらは5~7種類の基本的感情を意識的に（しかも普段目にすることがまれなほどに強調した形で）表出した表情を入念にコントロールされた条件で顔画像として入力し、これを基本的感情に対応する表情カテゴリーに強制的に分類するという、実験目的に限定されたタスクを実現したものであった。人間が顔表情を介して自然にやり取りしている感情情報をコンピュータで読み取れるようにすることで、機械が言葉によらない人間の意図をくみ取って対応してくれるような、高度なヒューマンインタフェースを実現したい、という最終目標と照らせば、表情認識の研究はようやく緒についたところといえる。そしてヒューマンインタフェースの高度化への貢献を現実のものとする

ためには、今後更に以下のような課題への取組みが不可欠と考えられる。

- ① 人間が表出する顔の表情について、より多くの事例を対象として学習・評価を行うことによって、更に高い認識精度の達成を目指す。
- ② 一般的環境での画像入力において起り得る種々の要因（例えば照明など）による変動に対して影響を受けにくく安定な認識の実現を目指す。
- ③ 一般的な表情から感情を認識する普遍的な分類機構をベースとしながら、更に適応的な学習を行うことによって、個人による表情の個性に対応した認識の実現を目指す。

これらは、独り、顔表情からの感情の認識に限らず、正しく・安定で・柔軟な認識の実現を目指すために広くパターン認識・学習の研究全般において取り組むべき課題といえる。加えて、ここで認識対象となっている顔の表情は、高度な視覚を有するとともに社会的な存在である人間が、相手の感情や意図を視覚を通じて認識し相互のコミュニケーションを円滑に行う重要な手段としているものである。したがって、コンピュータによる表情の認識においては、単に正確さだけでなく、人間の目から見た判断の自然さが求められる。そこで、

- ④ 人間による表情認知過程の特性とよく整合し、人間の認識判断との矛盾や不自然さが少ない認識アルゴリズムの実現を目指す、

という人間に学ぶ視点も重要である。具体的な検討課題としては次のようなものが挙げられる。

(a) 表情の表出に伴う顔パターンの物理的な特徴の微妙な変化から人間は相手の感情を敏感に認識することができる。人間にとって不自然さの少ないコンピュータによる表情認識を実現するためには、顔パターンの物理的な特徴の表現法として、人間によって知覚される感情の心理量をできるだけ反映した特徴量を抽出することが望まれる。

目 次

- | | |
|--------|---|
| [I] | コミュニケーションにおける表情とコンピュータによるその自動解析 (9月号) |
| [II] | コンピュータによる顔表情認識技術 (1)
— 表情による感情の認識 — (10月号) |
| [III] | コンピュータによる顔表情認識技術 (2)
— 顔面の表情動作の認識 — (12月号) |
| [IV・完] | コンピュータによる顔表情認識の展望
— 人間による表情認知に学ぶ — (1月号) |

(b) コンピュータによって、従来のように人為的に誇張された表情ばかりでなく、日常的によく目にするような自然で中間的な表情に対しても適切な認識を行えるようにするためには、人間によって表情から認知される感情が、顔パターンの連続的に変化する表情空間の中でどのように分布しているかを明らかにすることが望まれる。

(c) 人間が相手の顔から読み取る感情は表情筋によって生じた顔面動作という狭義の表情だけで決まるものではない。例えば、視線、姿勢、ジェスチャ、更には音声や言語など表情以外のモードから得られる情報も、顔の表情から認知される感情に大きな影響を与える。また、複数のモードによる情報が時間的なずれを伴って提示される場合には、それらの記憶によって形成される文脈もまた感情の認知に様々な影響を与える。したがって顔表情の認識によって人間とコンピュータとの非言語コミュニケーションを現実のものとするためには、人間による表情の認知に関して、マルチモーダル情報の相互作用やその文脈依存性についての基本的性質を明らかにし、その知見をシステム構築に反映させることが望まれている。

以上のような課題については、人間による表情認知のメカニズムを明らかにしようと、心理学の研究者たちが中心となって、古くからその解明に取り組んできた⁽¹⁾。しかし最近になって、工学や心理学の研究者の間の共同研究を通じて、人間による表情の認知の性質を明らかにすると同時に、それらの知見との整合性の高い認識手法の開発を目指すという研究が活発化してきている。

本稿では、紙面の制約もあるので、特に上記 (a) の視点に立った取組みに的を絞り、その代表的な事例として表情の顔パターンが持つ静的な形状の情報及び表情表出時の顔面動作の動的特性に関する情報が、人間による表情の認知にどのような影響を及ぼしているかを明らかにするために、最近の研究事例を紹介することにしたい。

2. 表情による感情の分類のための心理学的妥当性のある顔パターンの特徴表現法

顔の表情が表している感情カテゴリーを静止顔画像から認識する研究については主要な取組みを本講座第2回で紹介した。その代表的アプローチとして、顔パターンの画像から様々な特徴を抽出して得られる多次元特徴量に対して主成分分析を適用することによってより低次元の特徴ベクトルに変換した後、この特徴ベクトルを用いて、最小距離識別法に代表されるよく知られたパターン認識の多クラス分類法を適用して基本的感情カテゴリー間の識別を行う幾つかの手法がこれまでに試みられてい

る。

これら各手法を特徴付けるのは、表情によって生じる顔パターンの変化をどのような多次元特徴量として数値的に表現すればいいかという問題である。最も素朴な方法は、目や口など顔に特有な造作を基準として、顔パターンの位置と大きさを正規化して得られた画像から、各画素における濃淡値を順次抽出して多次元の特徴量を得るというものである。このように各画素での濃淡値を成分とする多次元データを多くの顔画像について収集し、それらのサンプル集合の主成分分析によって求められる正規直交基底を用いて個々の顔画像の多次元データを展開し、より低次元に圧縮された識別用特徴ベクトルを求めるという手法は Eigenface 法と呼ばれている。これは顔画像によって個人識別を行う実用的なシステムの実現法の一つとして、その評価がほぼ定着している⁽²⁾。

一般に、多次元データのサンプル集合に主成分分析を適用して得られる基底を用いて個々のデータをその線形結合で表し、そのときの各基底の重み係数によって低次元の特徴ベクトルに次元圧縮する方法を適用するためには、個々の多次元データは数学的にはベクトルとしての性質を満足している必要がある⁽³⁾。これは、顔画像を記述する多次元特徴量に関していえば、その各成分が異なる顔パターンの中でそれぞれ同一の意味を表している必要があるということである。言い換えれば、多次元特徴量の任意の成分に関係する顔画像中の各画素は、異なる顔パターンの中で互いに対応付けられていなければならない。しかし、顔の位置や大きさを正規化しただけの画像から各画素における濃淡値を順番に抽出して得られる多次元特徴量は、厳密にはこのような条件を満足しているとはいえない。

その解決策として、各画素でのオプティカルフローを計算することによって異なる顔画像間で各画素の詳細な対応関係を求める試み⁽⁴⁾もあるが、これを自動的に求める計算は複雑であり誤差も大きい。そこで、異なる顔画像の間で近似的に対応関係がとれている多次元特徴量を求める次のような方法も提案されている。まず、顔の造作に対応して定義された幾つかの特徴点を画像中から抽出し、それらの位置から個々の顔の形状特徴を求める。多数の顔についてこの形状特徴を平均することによって標準顔形状を求めておき、個々の顔パターンをその形状が標準顔形状にフィットするように変換する。そして、顔の形状が規格化された画像について、各画素での濃淡値をサンプリングすることによって多次元特徴量を得る、というものである。このようにして得られた多次元特徴量は shape-free 特徴と呼ばれている⁽⁵⁾。その各成分は、近似的にはあるが、異なる顔パターン間で互いに対応がとれているため、ベクトルとしての取り扱いに適

すると考えられている。

図1は、このような shape-free 特徴を求める処理の概要を示している。

なお、この shape-free 特徴という概念は上記のようにパターン認識において特徴ベクトルが満たすべき条件に関する数学的な考察に基づいて導かれたものであるが、提案されている処理内容の本質は、顔の三次元ワイヤフレームモデルへのテクスチャマッピングによって多様な顔画像の生成を目指して行われた先行研究⁽⁶⁾において、ワイヤフレームモデルの節点の位置座標が与える顔の形状情報と、各パッチ面にマッピングされるテクスチャ情報とを分離して表現し、両者を独立な線形モーフィングモデルを用いて合成することにより新規顔画像の生成を行おうとする考え方と理論的にほぼ等価なものと考えられる。

この shape-free 特徴による認識と従来の Eigenface 法による認識との比較は顔画像による個人識別のタスクに関してまず行われた⁽⁵⁾。その結果、shape-free 特徴による認識では、顔パターンの位置合せのみを行った Eigenface 法と比べて、個人識別の精度が向上しているだけでなく、例えば、平均的な顔よりも特異性の強い顔の方が正しく認識されやすいなど、人間による顔認知に関して報告されている幾つかの特性が shape-free 特徴による認識シミュレーションにおいても再現されることが確かめられた。すなわち、顔による個人識別に関しては shape-free 特徴が心理学的妥当性の高い特徴表現となっ

ていることが示唆されている。

これに対して Calder らは、顔パターンの特徴を多次元ベクトルとして表現する幾つかの方法について、表情による感情カテゴリーの認識というタスクにおける妥当性の比較を行った。そして前記の shape-free 特徴が、表情の認識においても、単に認識精度の向上ばかりでなく、人間による表情認知特性との整合という心理学的妥当性の観点からも、従来の Eigenface 法と比べて優れていることを実験によって確認している⁽⁷⁾。以下にその概要について紹介する。

顔パターンの多次元特徴量が表情による感情カテゴリーの認識にどの程度有効かを評価するための実験データとして、ほぼ一定の撮影条件で様々な表情の正面顔を撮影したモノクロ写真からなる Ekman & Friesen の表情画像セット⁽⁸⁾から、14名の被験者について六つの基本感情を表出した表情と真顔の計7種類の顔画像を各1枚以上、計110枚を選んで使用した。

顔パターンを表現する多次元特徴量としては、表1に示す4種類を比較対象として取り上げ、各特徴を用いた場合の感情カテゴリーへの分類の精度と、人間による表情認知の傾向との整合性についての比較を行った。

表情が表している感情カテゴリーを分類する能力の優劣を比較する実験では、7種類の表情からなる前記の表情画像セットについて、表1に挙げた4種類の多次元特徴量を求めた。そしてまず主成分分析によってより低次元の特徴ベクトルに次元圧縮を行った後、特にクラス間

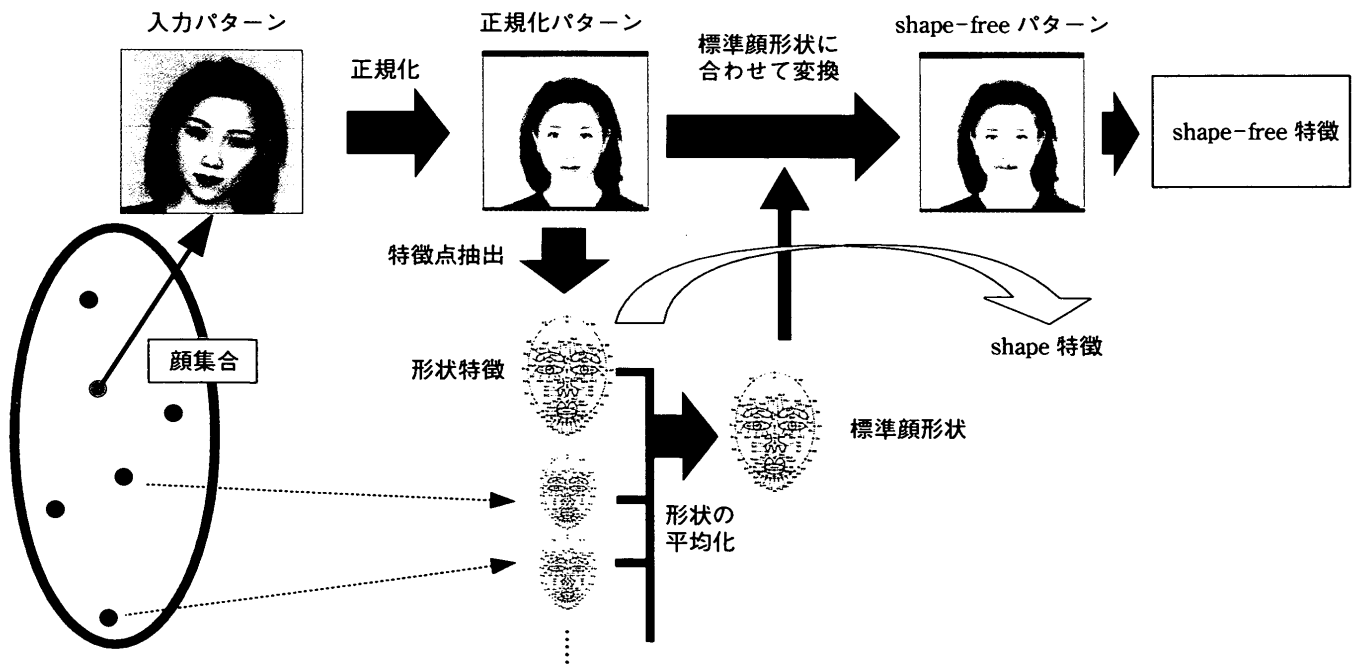


図1 顔画像からの shape-free 特徴の抽出 顔の造作に対応して定義された幾つかの特徴点を顔画像から抽出し、それらの位置を表す形状特徴 (shape 特徴と呼ぶ) を求める。多数の顔の形状特徴を平均化して得られる標準顔形状にフィットするように入力パターンをモーフィングで変換し、得られた形状が規格化された顔画像の濃淡値をサンプリングすることで shape-free 特徴が求められる。

表1 実験で比較した顔パターンを表現する多次元特徴量

特徴の種類	特徴が表している情報	特徴の次元数
original image	位置を正規化した濃淡画像	190 × 285
shape-only	31個の特徴点のx,y座標	31 × 2
shape-free	形を標準形状に規格化した濃淡画像	190 × 285
shape-free + shape	shape-only と shape-free の併用	190 × 285 + 31 × 2

判別能力が高いと判定された成分を変数として選んだ。これらの変数を用いて7種類の感情カテゴリーを分類する多クラスの線形判別分析を適用し、leave-one-out法に従ってその分類精度の評価を行った。この結果、顔パターンの形状を平均顔の持つ標準顔形状に合うように変換して個々の顔の形状を規格化した画像の濃淡値情報(shape-free特徴)は、顔による個人識別の場合と同様に、感情カテゴリーの分類においてもより高い精度による識別を可能としていること、またこの形状を規格化した濃淡値情報と特徴点の位置情報を併用した多次元特徴量(shape-free + shape特徴)を用いた場合は更に分類精度の向上が可能となることが確認された。

また、上記の線形判別分析による感情カテゴリーの認識結果を人間による感情カテゴリー判断の結果と対比させることによって、顔表情パターンに対する各多次元特徴量が人間による表情認知の傾向とどの程度整合しているかを比較した。前述のEkman & Friesenの表情画像セットを対象に線形判別分析を行った結果から、顔パターンをshape-free + shape特徴によって表現することによって、人間とほぼ同等の分類精度(約83%)が達成できることが分かった。更に、怒り—嫌悪、悲しみ—真顔など、人間でさえも誤りやすい感情カテゴリー間の判別は、このshape-free + shape特徴を用いた分類においても同様に困難であり、同種の混同が見られることが確認された。これらの実験の結果、shape-free + shape特徴による顔パターンの多次元ベクトル表現は、Eigenface法で従来用いられてきたoriginal image特徴による表現に比べて、人間による感情カテゴリー判断との類似性の高い表情認識が可能となることが期待される。

ところで、人間が表情から認知する感情については、これを離散的にカテゴリー化して考えるのではなく、感情が連続する空間内に個々の表情が配置されているとする考えも有力である⁹⁾。こうした立場に立つ研究者たちの多くは、快・不快感と覚醒水準を表す二つの次元を表情空間を規定する直交軸として提唱している。そこで認識実験に使用した表情画像セットに含まれる個々の表情について、あらかじめ主観評価実験を行い、人間によって認知される快・不快感、覚醒度の大きさを定量化した。そして個々の顔パターンについて、それぞれの多次元特

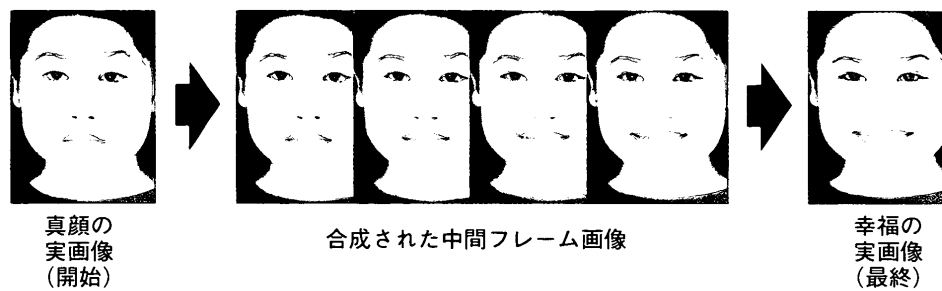
徴量を用いて算出され、表情の感情カテゴリー分類に使用される線形判別関数の値(ここでは七つの感情カテゴリーを分類するために使用された6個の判別関数値のうち、識別への寄与の大きい上位4個を選んだ)と、その表情に対する快・不快感、覚醒度の主観評定値との間の相関の大きさを調べた。その結果、shape-free + shape特徴を用いて得られた判別関数値が、主観評定値と最も高い相関を示した。言い換えると、shape-free + shape特徴に基づく顔パターンの表現が表情空間を規定しているといわれる認知的な二つの次元を最もよく説明できる可能性が示されたことになる。

以上に述べたように、顔パターンの形状をあらかじめ平均顔にそろえて規格化した後に、濃淡値をサンプリングして得られるshape-free特徴は、個々の顔の形状を表すshape特徴とも併用することによって、感情カテゴリーの認識精度の向上に寄与するだけでなく、顔表情から感情を読み取る人間の感情カテゴリー判断の傾向ともよく整合した統計的性質を持っていることが確かめられた。まだ特徴抽出の完全自動化には問題を残しているが、人間の認知に近い顔表情の認識を実現する可能性を秘めており、心理学的妥当性のある顔パターンの特徴表現法として有望視されている。

3. 人間による表情の認知・生成のダイナミクス

表情認知における顔面の動き情報の重要性に関しては、バイオロジカルモーションに関するBassiliの先駆的研究¹⁰⁾がよく知られている。これは、顔の表面に多数の光点をつけて、その動きだけを観察できるようにしたところ、観察者は光点の動き情報だけを用いても多くの表情を読み取ることができたというものである。しかし表情認知における動的情報の役割を更に詳しく明らかにする認知的研究はその後余り進展してこなかった。その理由の一つは、表情表出によって自然に創り出される顔の動的な変化を視覚心理実験で提示する刺激として統制することが極めて困難であったことが挙げられる。

これに対して蒲池らは、モーフィングと呼ばれる画像合成技術を用いて、個人や表情間での動的情報の条件を統制した表情表出の動画像を作成し、これを刺激とした心理実験を行って、顔面の動き情報が表情による感情の認知に及ぼす影響の幾つかの側面を明らかにしてい



実験に使用した動画の生成条件

表出速度	中間フレーム数	モーフィング率	提示時間
大	4枚	0, 20, 40, …100%	0.2秒
中	24枚	0, 4, 8, …100%	0.87秒
小	99枚	0, 1, 2, …100%	3.37秒

図2 表出速度を統制した表情の動画の生成 真顔と表情顔との合成割合を段階的に変化させて両者のモーフィングを行うことにより中間フレームを作成し、これらを連続的に提示して動画刺激とした。

る^{(11)・(12)}。以下にその研究成果の概要を紹介する。

表情を表出する動画の作成にあたっては、表出者の真顔を撮影した静止画像を開始画像とし、同じ人物が悲しみや喜びといった基本感情を表出したときの顔の表情を撮影した静止画像を最終画像とした。そして開始画像と最終画像の合成割合を段階的に変化させながらモーフィングすることによって、図2に示すような中間画像を順次作成し、これらを一定の時間間隔で連続的に提示することによって動画として再生される画像系列を求めた。ここで、開始画像と最終画像の間に中間フレームとして挿入する画像の計算方法とフレーム数を適当に設定することによって、表情表出の動的な特性を自在に制御することが可能となる。統制を行う動的特性のパラメータとしてここでは表情の表出速度を取り上げ、その大きさを変化させた場合にそれぞれ生成される動画を刺激として主観評定を行い、表情から認知される感情の種類や強度が表出速度にどのように影響されるのか、また、表出される感情によって表出速度が認知に及ぼす影響がどのように異なってくるかを心理実験を通じて調べた。

研究の第一歩として、真顔から幸福、悲しみ、怒り、驚きの4種類の感情を強く表出する過程の動画が刺激として使用された。また、中間フレームの画像生成では開始画像と最終画像のモーフィングによる合成割合を等間隔で変化させた場合、すなわち、真顔から表情表出のピークまでの顔面上の各点の変位が時間に対して線形となるような場合が用いられた。更に表情の表出速度としては、図2に示した大・中・小の三つの条件について、表情から認知される感情を比較する実験が行われた。

実験では大学生28名を被験者として、各被験者には、6名のモデルが4種類の表情を3段階の速度で表出した72種類の合成動画が提示され、計2セッションから

なる主観判定実験が行われた。セッション1では表情を表出している動画刺激がどの感情を表しているように見えたかを自由に回答させ、セッション2では表情を表出している動画刺激から認知される各感情の強度をそれぞれ7段階で評定させた。なお、72種の動画刺激は各セッション内でランダムに提示された。

図3はセッション1の自由記述による正答率を示す。

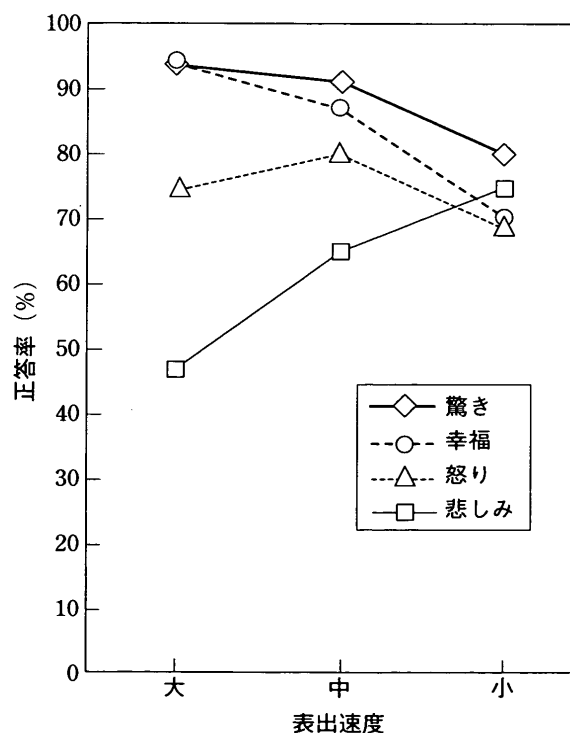


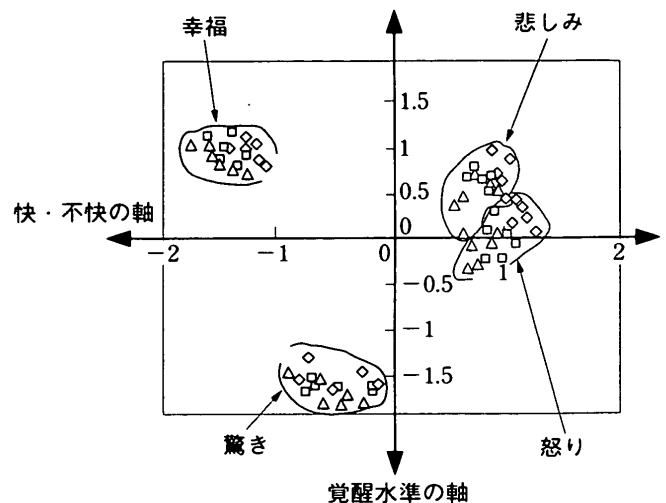
図3 表情の動画刺激が表している感情の主観判定 人が表情から感情をどれほど正確に読み取れるかは、感情の種類に応じて、また、表情の表出速度によって異なる変化を示すことが分かった。

グラフから明らかなように、表情から認知される感情と速度との関係は、表情によって異なることが確認された。例えば、「悲しみ」の表情については、表出速度が小さいほど「悲しみ」が正しく認知される割合は増大している。逆に表出速度が大きくなるにつれて正答率は下がっており、表出速度が大となる条件で提示された動画刺激に対する正答率は50%を下回っている。これに対して「驚き」の表情は表出速度が小さくなるにつれ正答率が下がっていくことがわかる。また、「幸福」や「怒り」の表情は、素早く提示された場合には「驚き」と混同される割合が大きくなり、逆にゆっくりと提示された場合には「悲しみ」と混同される割合が増大することが確かめられた。

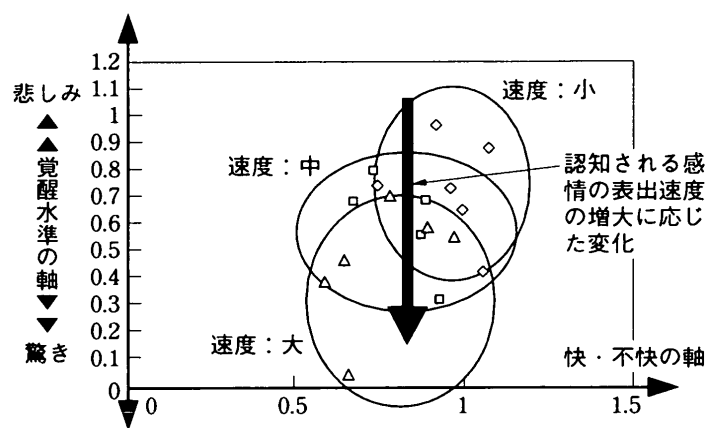
一方、セッション2においては、72種の動画刺激の各々に対して、各被験者が認知した幸福、悲しみ、怒り、驚きの4種類の感情の強度が評定結果として得られた。個々の動画刺激について全被験者による評定値の平均を求め、それらの主成分分析によって得られる第一、第二主成分を直交軸とする平面上に各々の動画刺激に対する評定値を射影した結果を図4(a)に示す。第一主成分では「幸福」と「怒り」、第二主成分では「驚き」と「悲しみ」が対極の位置に分布しており、表情空間の次元に関する先行研究⁹⁾による知見に基づいて、第一主成分が「快—不快」軸、第二主成分が「覚醒水準」軸に相当していると解釈された。

図4(b)は「悲しみ」の動画刺激に対して認知される感情の評定値が、表出速度の変化に応じてどのように分布するかを示したものである。表出速度の増大につれて、「悲しみ」の評定が「驚き」の方向、すなわち覚醒水準の変化を表す方向にシフトしていくのが観察された。

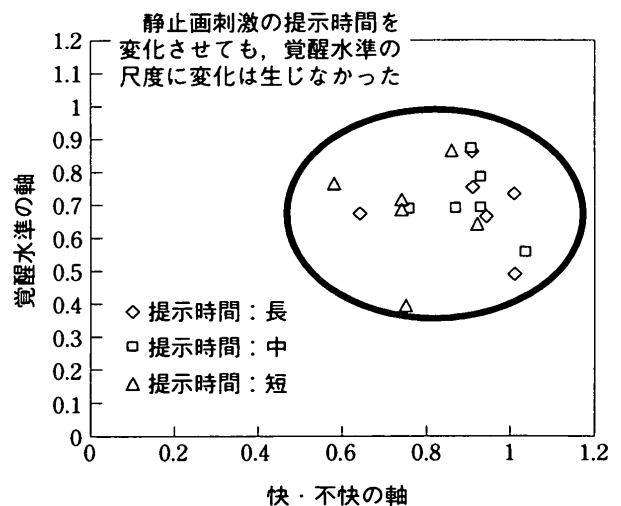
本実験での動画刺激における表情の表出速度の増大は、中間フレーム画像の削減によって実現されているため、単位時間当りに表示されるフレーム数は一定であるので、表示速度の増大につれて刺激提示時間が短縮されることは避けられない。このような刺激提示時間の変化が影響しているかどうかを明らかにする必要があるので、静止画を視覚刺激としてその提示時間を変化させた場合に認知される感情評定値の変化も調べている。図4(c)は、動画刺激での表出速度の変化(小→中→大)に対応する時間間隔(長→中→短)で「悲しみ」の静止画刺激を提示した場合に認知される感情の評定値がどのように変化したかを表している。「悲しみ」以外の表情の静止画について提示時間を変化させてみた場合も、「覚醒水準」の次元で有意な変動は確認されなかった。これらの結果から、表出速度は表情認知空間を規定する「覚醒水準」の次元に影響を与えるが、「快—不快」の次元には影響が少ないこと、更に提示時間の長短はいずれの



(a) 動画刺激に対する4種類の感情強度評定値の空間分布



(b) 「悲しみ」の表情を表出した動画刺激に対する感情の評定値の変化



(c) 「悲しみ」の表情を表出した静止画刺激に対する感情の評定値の変化

図4 様々な表情刺激に対する感情評定値の空間的分布
個々の画像刺激に対して、幸福・悲しみ・怒り・驚きの4種類の感情の強度について各被験者が評定した結果の平均値を求め、それらの主成分分析によって得られる第一・第二主成分を軸とする二次元平面上の分布を求めた。

次元にも影響を与えない、という表情認知空間のダイナミックな性質の一面が明らかになった。

人間が顔の表情を認知する際に動きの情報を利用していること、特に表情の表出速度が感情の認知に影響を及ぼすことを実験的に明らかにした上記の研究は、その本来の目的であったところの人間の「高次視覚システム」における情報処理メカニズムの解明という目標と照らして、科学的に少なからぬ価値のある成果と思われる。と同時に、顔表情から人間が読み取っている感情をコンピュータにできるだけ忠実に認識させることを目指す工学研究の視点からも、表情認識を人間の認知特性にできるだけ近づけるためには、静止画像で表される表情の形状情報だけではなく、動画像で表されるパターンの動き情報、特に速度情報にも着目していく必要があることを示唆し、表情認識の研究が今後進むべき方向を示したという点で、工学的にも本研究の意義は大きいものといえる。

さて、上記の研究^{(11),(12)}では、開始画像と最終画像の間の合成割合を時間の一次関数とし、パターンの全体にわたって均一なモーフィングを行うことによって中間フレームを生成して動画刺激を作成している。これは表情の認知にかかわる要因として動画像から抽出される速度情報を取り上げ、その認知に及ぼす影響を、速度情報を与えない静止画の場合と対比して示すことを主眼とした研究であったゆえに採用した近似であったといえる。表情表出時の顔面の移動量、形状の変化を高速度カメラを用いて調べた研究⁽¹³⁾によれば、意図的にコントロールされて表出された表情と、何らかの情動喚起によって自発的に生じた表情との間では、目、眉、唇といった各部分の動きには時間的に微妙なずれがあること、また、顔面上の各点の変位量は必ずしも時間の一次関数によって記述できるとは限らないことが確認されている。そこで、開始画像から最終画像に至る間の中間フレームでの顔面各点の変位量を時間の一次関数に固定して制御していた先行研究を拡張し、時間の三乗、三乗根、S字関数によって与えることで制御する場合も選択肢に含め、合成された表情の動きから人間に認知される感情を比較する研究⁽¹⁴⁾も進められている。その結果、顔面変位量を制御する時間関数の与え方によって、表情から人間が認知する感情の強度や表情の自然さ・わざとらしさの印象は変化し、最適な表情が得られるとされた場合の条件も表情の種類に応じて変わっていることが確認されている。今後の研究の展開によって期待される知見の蓄積は、人間の目に自然で、効果的に感情を伝える表情をコンピュータグラフィックスによって生成する技術に関して、顔表情の動作モデルのパラメータ最適化に大きく寄与するものと考えられる。

4. む す び

表情の表出によって生じる顔パターンの形状の変化やその動的な特性を記述した物理的特徴量が、人間に認知される感情の心理量とどのような関係にあるかを探っている最近の研究事例を二つほど紹介した。いずれもまだ多くの未解決の課題が残されており、画像入力された顔の表情から実際に感情の認識を行う「動くシステム」の実現にすぐに貢献するには至っていない基礎的レベルの研究ではある。しかし人間の視覚にならった高度な認識技術によって自然なインタフェースを実現するためには、このような基礎研究によって、まずは手本となる人間の視覚情報処理の特性についてのデータを蓄積することが必要であろう。と同時に、人間の情報処理メカニズムの解明そのものを目的とする科学研究の視点に立っても、科学的知見に対する応用可能性を探る工学研究との連携は、人間の情報処理に対する一つの仮説となり得る機能モデルや、その有効性を心理実験によって検証するための刺激作成ツールなどがもたらされるという点で有益なものとなるはずである。

そして顔表情の認識というテーマだけに限らず、人間のコミュニケーションにかかわる情報処理技術全般にわたって、心理学に代表される人間科学と、音声画像などの信号処理やパターン認識などの情報工学の専門家同士の連携による学際的なアプローチでの取組みが今後ますます求められてくると予想される。幸い我が国では、本学会がこのような時代のすう勢をいち早く見抜き、ヒューマンコミュニケーショングループ (HCG) を設立し、そのような学際的な情報交換や共同研究を積極的に支援する場を提供している (HCG の活動状況については URL <http://www.ieice.org/hcg/jpn/> を参照されたい)。

人間とコンピュータによる顔表情の認識という本講座のテーマは、当グループのヒューマン情報処理研究会 (HIP) 及びヒューマンコミュニケーション基礎研究会 (HCS) を中心に活発な研究発表が続けられている最もホットなテーマの一つとなっている。この講座を契機に本テーマに関心を寄せて頂き、このディスカッションの輪に加わって頂ければ、望外の喜びである。

謝 辞 本稿の執筆にあたって研究成果の画像や図面を提供して頂いた、英国 MRC Cognition and Brain Sciences Unit の Andrew Calder 博士、(株)国際電気通信基礎技術研究所の蒲池みゆき博士に感謝します。

文 献

- (1) V. Bruce, *Recognizing faces*, Lawrence Erlbaum Assoc., London, 1988.
- (2) 赤松 茂, “コンピュータによる顔の認識—サーベイ—,”

- 信学論(A), vol.J80-A, no.8, pp.1215-1230, Aug. 1997.
- (3) D. Beymer and T. Poggio, "Image representations for visual learning," Science, vol.272, pp.1905-1909, June 1996.
- (4) T. Vetter and N. Troje, "A separated linear shape and texture space for modeling two-dimensional images of human faces," Max-Planck-Institut für biologische Kybernetik, Technical Report, no.15, 1995.
- (5) I. Craw, N. Costen, T. Kato, and S. Akamatsu, "How should we represent faces for automatic recognition?," IEEE Trans. Pattern Anal. Mach. Intell., vol.21, no.8, pp.725-736, Aug. 1999.
- (6) 永田明徳, 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), vol.J79-D-II, no.7, pp.1230-1235, July 1996.
- (7) A. Calder, A.M. Burton, P. Miller, A.W. Young, and S. Akamatsu, "A principal component analysis of facial expressions," Vis. Res., vol.41, no.9, pp.1179-1208, 2001.
- (8) P. Ekman and W.V. Friesen, "Pictures of Facial affect," Human Interaction Laboratory, Univ. of California Medical Center, San Francisco, 1976.
- (9) C.A. Smith and P.C. Ellsworth, "Patterns of cognitive appraisal in emotion," J. of Personality and Social Psychology, vol.48, no.4, pp.813-838, 1985.
- (10) J.N. Bassili, "Facial motion in the perception of faces and of emotional expression," J. Experimental Psychology: Human Perception and Performance, vol.4, no.3, pp.373-379, 1978.
- (11) 蒲池みゆき, 吉川左紀子, 赤松 茂, "変化速度は表情認知に影響するか?—動画刺激を用いた顔表情認知の時間特性の解明—," 信学技報, HCS98-34, pp.17-24, Oct. 1998.
- (12) M. Kamachi, V. Bruce, S. Mukaida, J. Gyoba, S. Yoshikawa, and S. Akamatsu, "Dynamic properties influence the perception of facial expressions," Perception, vol.30, pp.875-887, July 2001.
- (13) 内田英子, 四倉達夫, 森島繁生, 山田 寛, 大谷 淳, 赤松 茂 "高速度カメラを用いた顔面表情の動的変化に関する分析," 信学技報, HIP99-76, pp.1-6, March 2000.
- (14) 尾田正臣, 磯野勝宣, "表情認知の時間関数特性についての一検討," 信学技報, HCS01-43, pp.67-72, Jan. 2002.



あかまつ しげお
赤松 茂 (正員)

昭50 東大・工・計数卒。昭52 同大学院修士課程了。同年電電公社(現 NTT)電気通信研究所入所。以来、パターン認識システムの研究に従事。平4~12 ATR 人間情報通信研究所第二研究室長。平13より法政大学教授。ヒューマン情報処理研究専門委員会顧問。工博。

国際会議

39th Design Automation Conference

主催：ACM/SIGDA, IEEE Circuits and Systems Society, the EDA Consortium

日時：2002年6月10~14日(5日間)

会場：Ernest N. Morial Convention Center (米国, ニューオーリンズ)

参加者：約9,500名(会議, 展示, チュートリアル)

主要参加国：米国, ほか

セッション数及び論文数：55セッション(うちパネル8), 6チュートリアル, 7“ハンズオン”チュートリアル, 論文数147

展示：約200社

主たるトピックス：

本会議は、EDA関連の最大の会議でEDA業界各社の展示会が併設されている。景気のせいで開催場所が西海岸でなかったことが原因かは定かではないが、参加者は昨年に比べ大幅に減ったようである。昨年から引続き会議の大きなテーマは、SOC時代に向けて

の embedded system 向けの設計ツールであった。今年の新しい試みとしては、参加者各自がワークステーションを用いてベンダツールのデモや説明を受けることができる“hands-on”チュートリアルと呼ばれる企画があった。

今年もテクニカルセッションは、Design Methods, Design Tools, Embedded Systems の3トラックに分けられ、設計事例から設計手法までEDAに関する広範囲な話題が扱われていた。そのため、業界全体の情報収集には非常に役立つと思われる反面、各自が特に興味を持つ分野に関して突っ込んだ議論をするという雰囲気は少なく感じられた。なお、キーノートや特別セッション等の講演内容を録画したビデオや発表スライド及び論文などが、<http://www.dac.com/39th/pindex.html> で公開されている。次回は、2003年の6月2日から6日まで、米国のカリフォルニア州アナハイムで開催される予定である。

(執筆者 山下 茂 正員 日本電信電話株式会社)

NTTコミュニケーション科学基礎研究所)