

個人差を克服する表情認識システム構成法の予備検討

立岡, 積 / AKAMATSU, Shigeru / TATSUOKA, Tsumoru / YUI, Keigo / 赤松, 茂 / 由井, 恵吾

(出版者 / Publisher)

電子情報通信学会

(雑誌名 / Journal or Publication Title)

電子情報通信学会技術研究報告. IE, 画像工学 / 電子情報通信学会技術研究報告. IE, 画像工学

(巻 / Volume)

104

(号 / Number)

648

(開始ページ / Start Page)

81

(終了ページ / End Page)

86

(発行年 / Year)

2005-01-27

個人差を克服する表情認識システム構成法の予備検討

由井 恵吾[†] 立岡 積[‡] 赤松 茂^{†‡}

[†] 法政大学大学院工学研究科システム工学専攻

[‡] 法政大学工学部 システム制御工学科

〒184-8584 東京都小金井市梶野町3-7-2

E-mail: [†] {i03r6132, e01d5112}@k.hosei.ac.jp [‡] akamatsu@k.hosei.ac.jp

あらまし カメラで撮影された顔画像から画像処理によって抽出される数値的特徴を用いて、顔表情が表している感情のカテゴリを識別する表情認識システムにおいて、各感情を表している顔表情の数値的特徴の登録に用いる人物と表情認識の対象となるユーザとの間に、個人差に起因する表情表出の仕方の違いや、人種の違いによる表情表出スタイルの違いなどがある時には、認証精度の低下が予想される。本論文ではまず、未登録の人物や異なる人種の人物に対して生じる表情認識精度の低下傾向を定量的に明らかにする。また、そのような認識精度の低下を克服する表情認識システム構成法の一つとして、個人識別や人種判定と組み合わせることにより特定人物や人種特有の表情辞書に切り替えて認識を行う個人(人種)対応型表情認識と、不特定多数の人物や複数の人種に対応した表情辞書を用いる汎用型表情認識とについて、その効果と限界について検討を行う。

キーワード 個人差, 人種効果, NN法, 学習サンプル

A preliminary study toward building a facial expression recognition system that overcomes individual variation

Keigo YUI[†] Tsumoru TATSUOKA[‡] and Shigeru AKAMATSU^{†‡}

[†] Department of System Engineering, Graduate School of Engineering, Hosei University

[‡] Department of Systems Control Engineering, Faculty of Engineering, Hosei University

3-7-2, Kajino-cho, Koganei-shi, Tokyo, 184-8584 Japan

E-mail: [†] {i03r6132, e01d5112}@k.hosei.ac.jp, [‡] akamatsu@k.hosei.ac.jp

Abstract This paper describes a facial expression recognition system that is designed to overcome individual variation. This approach employs that the facial expression image from the camera are coded by applying Gabor filter and Principal component analysis. If there are differences between training persons and test persons in race and degree of intensity expression action, a generalization rate for recognition of expression comes down. The results show that if a general-purpose expression recognition system equips with personal or racial identification before process of expression recognition, the possibility of a system capable of overcoming individual variation facial expression could become reality.

Keyword facial expression recognition, individual variation, racial effect, nearest-neighbor, learning samples

1. はじめに

コンピュータが画像処理を用いて、人の顔の笑顔などの表情を読み取る表情認識の研究は、近年盛んに行われている。認識精度の向上のためには、いくつかの解決すべき課題があげられている[1]が、本論文ではその中の“人による表情の個人差を克服する”という課題にとりくんだ。顔という対象の形状が人物それぞれ異なるように、表情表出の仕方、例えばある感情をどの程度の大きさの顔面変形として表出するかについて個人差がある。したがって表情による感情の分類機構を土台として、これを個々の人の表情表出の特性に応

じて修正していく適応的な学習メカニズムを導入する必要がある。そこで本論文では、個々人の個性による表情表出の違い、および、集団としての人種差による表情表出の違いが、顔表情の自動認識の精度にどのように影響するかを実験によって明らかにし、これらの影響を受けにくい表情認識システムの構築に寄与する基礎データを得ることを狙いとしている。

表情の個人差への対策としては、不特定多数の人物による表情表出の顔画像から作成された識別辞書を用いた従来の汎用型表情認識システムに替わって、図1に対比して示すように、表情認識を行う前に個人識別

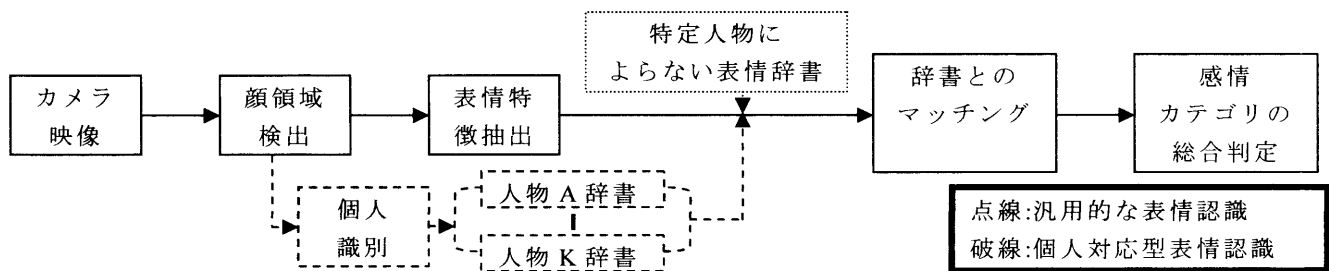


図 1 汎用型と個人対応型の表情認識システムの対比

を行って人物を特定した後、当該人物による表情画像から作成された識別辞書を用いて認識を行うという個人対応型表情認識システムが提案されている[2]。本論文ではまず、このような個人対応型表情認識システムの有効性を確認するために、共通の顔画像データベースを用いて、汎用型表情認識システムとの認識性能の比較実験を行う。

個人対応型システムでは、予め登録してある特定人物に対しては従来の汎用型システムよりも正解率が上がることが期待されるのに対し、

- 1 未登録人物には対応できない
- 2 各々の登録人物について、その表情表出時の顔を撮影して登録するのに、多くの手間と時間を要する
- 3 各個人対応の表情データを登録しておく必要があり、多数人物をユーザとして想定する場合、上記2の問題が解決されたとしても、識別辞書の容量の増大が問題となるといった点で、現実的な表情認識システムへの適用は困難と思われる。

そこで本論文では、真顔同士が互いに似通った特徴をもっている人物の顔の間では、各表情についても、2つの人物の顔は似通った特徴をもつものと仮定し、未登録人物については、真顔同士が(そのパターンを表現する特徴間の距離が小さいという意味で)類似している登録人物の表情辞書を用いて、表情認識を行うというアプローチをとりあげ、シミュレーション実験によってその有効性の評価を試みる。

さらに本論文では、人種差による表情表出の違いが顔表情の自動認識精度に与える影響を分析するために、表情辞書の作成に顔画像を提供した学習用人物と表情認識の対象となるテスト用人物の人種が異なる場合の表情認識の精度を検証することにした。具体的には、人種を日本人と外国人(実際は大部分が西洋人)とに分け、日本人の表情表出時の顔画像を学習サンプルとして作成した表情辞書を用いて、日本人ならびに外国人が表出した表情の認識実験を行う。外国人の表情顔画像にもとづいて作られた表情辞書に対しても同様に、日本人ならびに外国人それぞれについて同様の表情認識のテストを行う。また、両人種を混ぜ合わせた表情

顔画像の学習サンプルを用いて作成された表情辞書に対する認識実験の結果との比較も行う。

2. 表情認識処理

本研究では、静止画像を用いた表情分類において有力なアプローチの一つと言われている、GaborJet 特徴ベクトルに対して主成分分析を行うという手法を採用した[3]。

2.1. 顔の検出、正規化

顔画像から表情を自動認識するためには、シーン中から顔を自動検出し、顔の位置・大きさを正規化した顔画像を抽出することが不可欠である。本研究では、個人対応型表情認識システム[2]の評価においては、シーン画像からの正規化顔画像の抽出に HumanScan 社の提供する BioID software development kit (SDK)を使用している。一方、未登録人物の表情認識に類似人物の表情辞書を用いる手法の有効性、ならびに、表情辞書に表情を登録した人物と表情認識の対象となる人物との人種の違いの影響を探る実験は、大量の顔画像データベースを用いたシミュレーションによって行うため、顔画像正規化ソフトウェア FUTON システム[4]における顔の正規化機能を使用した。

2.2. 表情特徴抽出

2.2.1. Gabor Filter による特徴抽出

Gabor フィルタは、人間の視覚系における1次視覚野の特性を近似していることが知られていて、視覚パターンの局所領域の構造的な特徴を抽出することができる[3]。

本研究では Gabor フィルタに、方向のパラメータを与え、周波数と同時に方向性を抽出する。顔面で表情の特徴を捉えやすいと思われる49点において、Gabor フィルタの角度4方向・周波数窓幅4種の計16通りのフィルタを施すことにより、1画像ごとに784個の成分をもつ GaborJet 特徴が得られる。GaborJet 特徴ベクトルによる表情特徴抽出の流れを図2に示す。

GaborJet特徴ベクトルの抽出

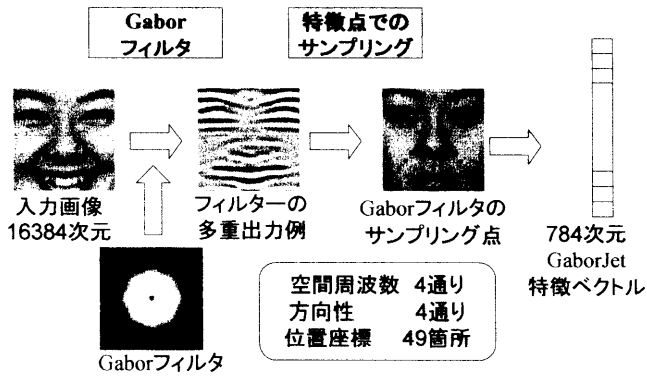


図 2 GaborJet 特徴ベクトルによる表情特徴抽出

2.2.2. 主成分分析

表情表出した顔パターンの特徴を表す GaborJet 特徴は 784 次元という極めて次元数の大きいベクトルとして得られるが、これらの多数の成分は互いに独立ではなく、強い相関をもつと考えられる。したがって、異なる表情パターンの違いを、より少数のパラメータ(低次元の特徴ベクトル)によって表現するために主成分分析による特徴ベクトルの次元圧縮を行う。今回の表情認識実験においては、後述のように計 10 種類の表情を多数の人物に表出させた表情顔画像から抽出された GaborJet 特徴ベクトルのサンプル集合に対して主成分分析を適用し、大きさの順に上位 50 個を選んだ固有値に対応する固有ベクトルが定める主成分を求めることにより、784 次元の GaborJet 特徴ベクトルは 50 次元の識別用の特徴ベクトルに次元圧縮される。

3. 個人識別

登録済みの人物に対する個人対応型表情認識システム[2]の有効性を評価する実験においては、表情認識する前に HumanScan 社の提供する個人識別(認証)エンジンを利用し、個人識別を行う[5]。

未登録人物の顔画像入力に対して、その真顔ともっとも類似している登録済み人物の真顔を求めることによって、入力顔画像の表情認識に用いる表情辞書を選択するという手法の評価は、予め用意された画像データベースを用いたシミュレーション実験によって行う。このため、BioID SDK で提供される個人識別(認証)エンジンは使用せず、個人認証への適用を念頭においてパラメータやサンプリング点を設定して得られる GaborJet 特徴[6]を用いており、これをさらに次元圧縮して得られる識別用特徴による最小距離識別法によって、もっとも類似性の高い登録人物の選択を行う。

4. 実験 1 個人対応型と汎用型表情認識システムの比較

個人対応型表情認識システム[2]の有効性を示すため、図 1 に示すフローで学習サンプルに対し“個人対応型の表情辞書”と、一般的である多人数の標準パターンを用いた“不特定多数の人物による表情辞書”を作り、NN(Nearest Neighbor)法でマッチングを行い、10 表情の識別率の比較検討を行った。

本実験の使用データベースの概要を表 1 に示す。

表 1 実験 1 での使用データベースの概要

	人数	表情の種類	1人1表情あたりのサンプル数	合計
学習サンプル	8人	10種類	5	400枚
テストサンプル	8人	10種類	100	8000枚

表情識別実験に用いた表情の種類は(真顔、喜び(開口)、喜び(閉口)、怒り(開口)、怒り(閉口)、悲しみ、恐れ、軽蔑、嫌悪、驚き)の 10 通りである。実験環境は被験者の顔から約 40cm 離れた位置に CMOS カメラを設置し顔画像を取り込み、ノートパソコン上で実装したリアルタイム表情認識システムで評価を行った。10 表情に対し“個人対応型表情辞書”と“不特定多数の人物による表情辞書”を用いて表情認識を行った場合の正解率の比較結果を図 3 に示す。

個人対応型表情認識では、驚き以外の 9 表情で、汎用型表情認識に比べ認識率が上昇している様子がわかる。また、汎用型な表情認識では正解率は 56.7%に対し、個人対応型の正解率は 74.8%という正解率向上の結果を得た。

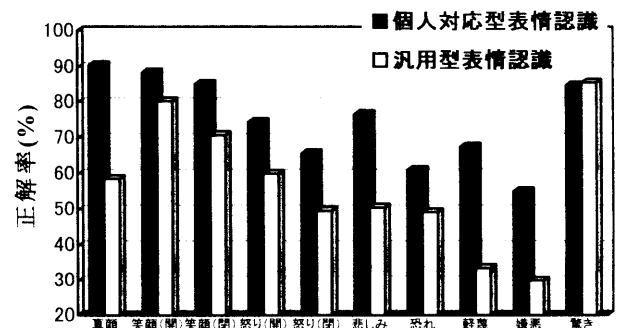


図 3 個人対応型と汎用型表情認識の正解率

5. 実験 2 未登録人物に対する個人対応型表情認識の適用の試み

個人対応型システムが未登録人物にどの程度適用

できるか評価を行った。本研究では、未登録人物に対しては、その真顔が、登録人物の中から最も類似した真顔をもっている人物の表情辞書を用いて表情の認識を行うアプローチと、不特定多数の人物による表情表出の顔画像から作成された表情辞書を用いた従来の汎用型表情認識のアプローチについて NN 法による、10 表情の識別精度の比較・検討を、ATR 表情データベース[7]を用いたシミュレーションによって行う。

本実験の使用データベースの概要を表 2 に示す。なお本実験では、表情サンプルの少なさを補うため、識別率の評価は Leave-one-out Method を用いる。

表 2 実験 2 での使用データベースの概要

	人数	表情の種類	1人1表情あたりのサンプル数	合計
表情サンプル	20人	10種類	2	400枚

表情認識の正解率の比較結果を図 4 に示す。従来の汎用型表情認識では全体の正解率は 64.0% となったのに対し、類似人物の表情辞書を用いて個人対応型表情認識を行った場合には全体平均の正解率は 60.3% であり、3.7% マイナスのレベルにとどまった。

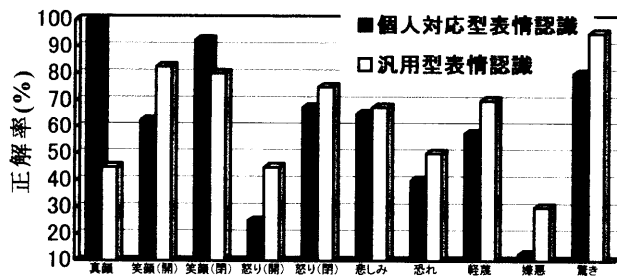


図 4 未登録人物に対する個人対応型 VS 汎用型表情認識の正解率

6. 実験 3 人種差が表情の識別精度に与える影響

人種差が表情の識別性能に及ぼす影響を探るため、本研究では、異なる人種の顔として、日本人の顔のサンプルと主に西洋人からなる外国人の顔のサンプルの 2 つを用いた。実験に使用した表情顔画像データベースの概要を表 3 に示す。本実験でも表情の識別には NN 法を用いるため、識別辞書は学習サンプルとして用意された各表情を表出している顔画像から抽出された特徴ベクトルを格納したものとなる。実験で使用する表情辞書としては、各表情を表出している日本人の顔から得られた特徴を格納した「日本人表情辞書」、外国人の顔から得られた特徴を格納した「外国人表情辞書」、両者を混

在させた「人種混合表情辞書」の 3 種類を用意した。そして、日本人の顔、ならびに、外国人の顔に対して、これら 3 種類の表情辞書を用いて NN 法によって 10 表情の認識を行った場合について、識別精度の比較・検討を行う。

本識別実験でも、比較的少数のサンプルを用いて識別性能をできるだけ精度良く評価するために、交差確認法(Leave-one-out 法)を用いて評価を行う。学習サンプルとして表 3 に示すような表情の顔データベース N 人分を使えるものとする、ここではまず一人あたり 10 表情、各 1 パターンからなる顔データ(N-1)人分、計 10(N-1)個を学習サンプルとして表情辞書を作成する。この時、残された 1 人の 10 表情、各 1 パターンずつ、計 10 個のパターンを用いて識別評価を行う。テストパターンを人物単位で順次交代させながら、上記の過程を N 回繰り返すことによって、トータルの表情認識率を求める。

表 3 実験 3 での使用データベースの概要

	人数	表情の種類	1人1表情あたりのサンプル数	合計
日本人サンプル	男女各12名	10種類	1	240枚
外国人サンプル	男女各18名	10種類	1	360枚

日本人の表情顔をテストサンプルとして、これを同人種の表情辞書(日本人表情辞書)と他人種の表情辞書(外国人表情辞書)と人種混合表情辞書を用いた NN 法によって 10 表情の識別を行った時の、表情毎の正解率を図 5 に示す。10 表情のうち、喜び(開口)・喜び(閉口)・嫌悪を除く 7 つの表情では、他人種の表情辞書を用いると同人種の表情辞書よりも正解率の低下が確認された。人種混合表情辞書を用いると 10 表情のうち、他人種の表情辞書(外国人表情辞書)を用いた場合には正解率が著しく低かった真顔・怒り(閉口)・悲しみ・恐れ・軽蔑の 5 表情においては、正解率の改善が認められた。

10 表情全体についての正解率の平均は、同人種の表情辞書を用いた場合の 70.0% が、他人種の表情辞書を用いた場合では 46.3% に低下し、人種混合表情辞書を用いた場合には 65.8% という結果になった。

また、外国人の表情顔をテストサンプルとして、これを同人種の表情辞書(外国人表情辞書)と他人種の表情辞書(日本人表情辞書)と人種混合表情辞書を用いた NN 法によって 10 表情の識別を行った時の、表情毎の正解率を図 6 に示す。10 表情のうち、怒り

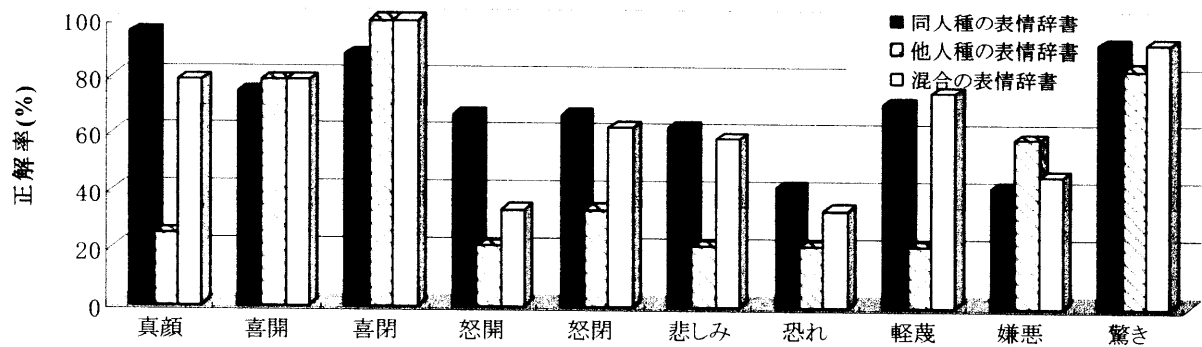


図 5 日本人の顔表情の認識(同種辞書 VS 他種辞書 VS 混合辞書の表情辞書)

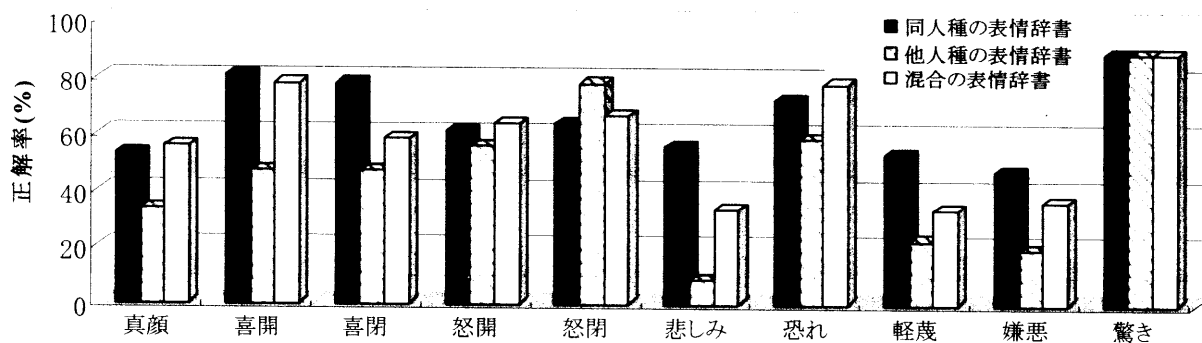


図 6 外国人の顔表情の認識(同種辞書 VS 他種辞書 VS 混合辞書の表情辞書)

(閉口)を除く 9つの表情では、他人種の表情辞書を用いると同種辞書よりも正解率の低下が確認された。人種混合表情辞書を用いると他人種の表情辞書(日本人表情辞書)を用いた場合には正解率が低かった表情において正解率の改善が認められた。

10 表情全体についての正解率の平均は、同種辞書を用いた場合の 65.3%が、他人種の表情辞書を用いた場合では 45.8%に低下し、人種混合表情辞書を用いた場合には 59.2%という結果になった。

3 種類の表情辞書を用いた場合のそれぞれについて、10 表情全体についての正解率の平均を図 7に示す。

日本人、外国人ともに、表情認識を行う際の表情辞書が他人種の表情にもとづいて設定されていた場合には、認識の正解率は大きく低下してしまう。複数の人種による表情データを集めた人種混合表情辞書を用いることで、この低下した分はかなり回復させることができるものの、しかし同一人種の表情に基づく同種辞書の表情辞書による表情認識の精度には及ばないので、被認証者として複数の人種の人物がありうる場合には、表情認識の正解率の期待値は低下することは避けられない。したがって、個人の識別によって個人対応型表情認識が考えられるように、表情の認識に先立ってユーザの一属性としての人種を推定して、その結果に応じて表情辞書を人種毎に切り替えて表情の識別を行うというアプローチも有望と考えられる。

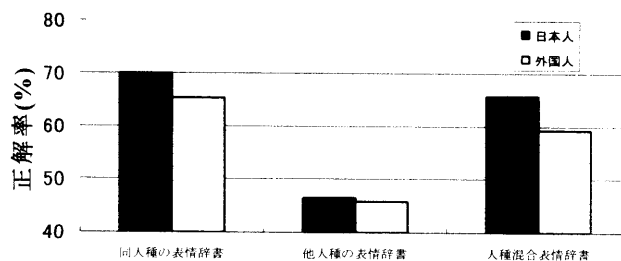


図 7 表情辞書の種類に応じた表情認識の平均正解率

7. まとめ

表情認識の性能向上のため、表情表出の個人差・人種間差に着目し、それらに影響を受けない表情認識システムの実現に向けた基礎検討を行った。

まず実験 1 では、先に提案した個人対応型表情認識システム[2]の有効性の評価のため、実装したシステムを用いて汎用型表情認識との比較を行った。被験者 8 人に対し、従来の汎用型表情認識では正解率は 56.7%となったのに対して、個人対応型表情認識では 74.8%となり、18.1%の認識率向上という結果を得た。嫌悪・軽蔑などの表情は、人それぞれ個人差が激しかったため、汎用型表情認識では高精度な認識が困難であったが、個人対応型表情認識では正解率の向上が確認された。今回の被験者は特別な訓練をした役者ではなく一般人であったので、一定した表情の表出が困難であつ

た。このため、個人対応型表情認識システムにおいても、正解率は75%程度にとどまるという結果となっている。

また個人対応型表情認識では未登録の人物に対応できないという問題を克服するため、被験者と真顔が類似している登録人物についての個人対応表情辞書を用いて未登録人物の表情認識を行うというアプローチを提案し、実験2においてLeave-one-out法によるシミュレーションによってその有効性の評価を試みた。類似人物の表情辞書を用いた個人対応型表情認識の長所としては、顔つきに応じて真顔を真顔と認識できないケースが減り、真顔に対する表情認識の正解率が汎用型表情認識より向上したことである。しかし全体としての正解率は、従来の汎用型表情認識では64.0%となったのに対し、真顔の類似人物の個人対応表情辞書を用いて表情認識を行うという提案手法では全体平均の正解率は60.3%であり、3.7%もマイナスとなってしまった。顔の表情の個人差には、表情変化の土台となる顔そのものの造りの個人差と、FACSでも規定されている個々の表情筋の動きの強弱やそのバランスの違いといった表情表出のダイナミックな特性における個人差とが含まれる。今回の提案手法では、真顔が類似している登録人物の表情辞書を選択することで、前者の個人差についてはある程度まで対応できていると思われる。しかし、顔パターンの特徴抽出レベルでは同じ表情内の個人差による変動と、異なる表情の間のパターンの違いとを区別できないため、後者の個人差については対処しきれていない。今後は表情表出のダイナミックの特性において、個人差と異なる表情間の差異とを分離できるような特徴抽出法の導入を検討する必要があると思われる。

そして実験3では、人種差が表情の識別精度に与える影響をシミュレーション実験により分析した結果、日本人の表情顔からなるテストサンプルに対しては、同人種（日本人）の表情辞書を用いた場合に正解率が70.0%、他人種（外国人）の表情辞書を用いた場合に46.3%、人種混合による表情辞書を用いた場合に65.8%という結果が得られた。外国人の表情顔からなるテストサンプルに対しては、同人種の表情辞書を用いた場合に65.3%、他人種の表情辞書を用いた場合に45.8%、人種混合による表情辞書を用いた場合に59.2%という結果になった。被認証者として異なる人種の人物がありうる場合には、最初に人種の識別を行い、その人種特有の表情辞書を適用するのが望ましいという結果が得られた。また日本人の場合、同人種の表情辞書を用いても正解率は70%に留まった。これは評価を行った日本人データベースでは、喜び（開口）などのように口の開閉を指定している表情があるのに対して、恐れ、

嫌悪など口の開閉を指定していない表情もあって、表情表出における口の開閉に関する個人差の存在が正解率の下降の主要な要因になっていると思われる。外国人の場合、同人種の表情辞書を用いても正解率は65.3%に留まったが、外国人データベースでは真顔と嫌悪、および、怒りと軽蔑の差は、人による評価でも見分けがつかないのが主要な要因になっていると思われる。

以上をまとめると、既知の人物、あるいは、人種に対してはその人物・人種に特有の表情辞書を用いて表情認識を行った方が高精度な表情認識が可能であるので、表情認識に先立って、個人識別あるいは人種の識別を行って、表情辞書の選択を行う個人対応型表情認識のアプローチが有効となる。また、個人対応の表情辞書を利用できない場合には、表情表出の類似した他の登録人物の表情辞書を代用する方法が考えられるが、その実現は、同一表情の表出に際しての個人差と、異なる表情間の差異とを区別しやすい特徴抽出法の確立を待つことにし、ここでは次善の策として、汎用型表情認識の採用の効用を明らかにした。

8. 謝辞

本研究の一部は、独立行政法人日本学術振興会の科学研究費補助金（基盤研究(B)(1) 15300076）の助成を得て行った。記して謝意を表す。

参考文献

- [1] 赤松茂, "講座 人間とコンピュータによる顔表情の認識[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-", "電子情報通信学会誌, Vol. 85, No. 10, pp. 766-771, Oct. 2002
- [2] 由井恵吾, 赤松茂, "個人対応型システムの実装と評価", 画像センシングシンポジウム予稿集, pp. 109-112 June. 2004
- [3] M. J. Lyons, J. Budynek, and S. Akamatsu, "Automatic classification of single facial images.", IEEE Trans. pattern Anal. March. Intell., vol. 21, no. 12, pp. 1357-1362, Dec. 1999
- [4] 向田茂, 蒲池みゆき, 尾田正臣, 加藤隆, 吉川左紀子, 赤松茂, 千原國宏, "操作性を考慮した顔画像合成システム:FUTON-顔認知研究のツールとしての評価-", "信学論, Vol.J85-A, No.10, pp.1126-1137, Oct. 2002
- [5] R. Frischholz: "BioID: A Multimodal Biometric Identification System", IEEE Computer, Vol. 33, No. 2, pp. 64-68 (2000)
- [6] 加藤博之, 田中健一, 赤松茂, "被認証顔の属性の多様性に対する個人認証システムのロバスト性の評価-GaborJet法とEigenface法の比較-"電子情報通信学会 画像工学研究会, Feb. 2005
- [7] 蒲池みゆき, "ATR 顔表情データベース(DB99)概要"ATRテクニカルレポートTR-H-305,(株)国際電気通信基礎技術研究所, Feb 2001
- [8] Nim Stim Emotional Face Stimuli (<http://www.sacklerinstitute.org/cornell/assaysandtools/>)