ニューラルネットワークを利用した分解能の異なるモザイク顔による男女および個人識別

正会員 河 合 秀 夫[†], 正会員 田 村 進 一^{††}

Man-and-Woman and Individual Classifications by Mosaic Facial Images with Different Resolutions Using Neural Network

Hideo Kawai[†] and Shinichi Tamura^{††}

あらまし バックプロパゲーション法を用いて、髪を除いた8×8, 16×16, 32×32 モザイク顔画像から男女および個人識別の実験を行った。その結果、顔かどうかすら判別しにくい8×8 未知画像から87%の高率で男女識別が行えた。

1. ま え が き

顔は視覚パターンの分類というきわめて重要な問題を提起する対象のひとつである。通常、顔を見るだけでその人が男であるか女であるかは容易に識別することができる。また、知人を見分ける場合においても、ある定まった位置からの正面顔だけでなく、多少角度が変わっても、距離が少し離れていたりしても、ある程度の識別を行うことができる。これは日常生活を通して、顔から男女および個人(知人)の特徴をあらゆる角度から幾度となしに感じとっている(学習している)からだと考えられる。

最近、この学習という観点からニューラルネットワークを利用して、顔からの個人識別や表情等の解析に関する研究が各方面で行われている。例えば小杉は、ニューラルネットワーク手法では最も代表的なバックプロパゲーション法(BP法)を用いて、髪を除く顔画像を12×12のモザイクに分割して入力を行い、学習後の中間層から男女の特徴や顔の構造的な特徴を解析している¹⁾。

今回、我々もBP法を用いて、髪を除いた顔を8×

8, 16×16, 32×32の分解能の異なるモザイク画像で男女および個人識別実験を行った。そして、BP法の識別能力と人間の識別能力との比較を行った。その結果、人間より優れた識別能力を発揮する面があることがわかった。

2. 実 験 方 法

高校卒業アルバムの中から男女各50人の正面向きで、髪を除いた全体顔をテレビカメラにより160×160画素サイズ、濃度階調256レベルで画像処理装置EXCEL-IIに入力する(入力画像と呼ぶ)。BPネット(ワーク)に入力するため、EXCEL-IIにより入力画像を32×32画素サイズ(32×32モザイク画像と呼ぶ)に縮小する。この操作により1モザイクの濃度は入力画像の5×5画素の平均値になる。16×16モザイク画像も同様、入力画像を16×16画素に縮小し、これを32×32モザイク画像と同一サイズになるように2倍に拡大する。この時の1モザイクの濃度は入力画像の10×10画素分の平均値である。8×8モザイク画像も、16×16と同一方法で作成した。したがって、1画素1ユニットとして入力層に加わるユニット数は各モザ

キーワード ニューラルネットワーク 顔画像 男女識別 個人識別, モザイク画像

1991年7月 1991年テレビジョン学会年次大会で発表

1991年11月11日受付

† 大阪電気通信大学 短期大学部(〒572 寝屋川市初町18-8 TEL 0720-20-4564)

†† 大阪大学 医学部(〒565 吹田市山田丘2-2 TEL 06-877-5111)

† Osaka Electro-Communication Junior College (18-8, Hatsucho, Neyagawa shi, Osaka 572, Japan)

†† Osaka University Medical School (2-2 Yamadaoka Suita shi, Osaka 565, Japan)

イク画像とも 1024 (32×32) である。なお、BP ネットは位置ずれ、縮小拡大等の変化には弱い特徴があるので、画像入力時にガイドラインを設定し、できる限り一定状態で入力するようにした。中間層は 10 ユニット、出力層は 2 (男女識別) および 10 (個人識別) の 3 層構造とした。なお、今回使用したニューラルネットワークは、日本電気のパーソナルニューロコンピュータ Neuro-07 を使用した。パソコンは PC 9801-RA2 である。

3. 実験結果および検討

Harmon (1973) は、人間の認識能力に関し「少なくとも 8 段階の灰色の明暗があれば、わずか 16×16 の正方形の集まりからでも顔が認識できる」と述べている²⁾。図 1 から 16×16 までは、目を細めたり、少し離れて見ると、原画の様相を感じ取ることができるが、8×8 になると、とうてい困難である。表 1 は、男女 15 人、計 30 人に対する内挿パターン (学習パターン) でテストした男女識別の結果である。各モザイク画像 (32×32, 16×16, 8×8) のそれぞれにおける学習はすべて 100% 収束している (— で示されている)。学習されたネットワークを学習ネットと呼ぶ。興味あることは、すべての結果において高い認識率が得られていることである。8×8 モザイク画像は人にとって、顔であることすら認識することが困難な画像

である。空間周波数的には、低周波成分の画像といえる。32×32 のモザイク画像においては目、鼻、口の輪郭が、ある程度は認識することができ、いわば低周波+高周波成分の画像といえる。しかし、この 2 つのモザイク画像間 (32×32 モザイク画像の学習ネットを使用して 8×8 をテストした場合と、8×8 の学習ネットを使用して 32×32 をテストした場合) の認識率は、ともに 90% の高い認識率が得られている。このことから考えて、各モザイク画像の学習ネットは、低周波成分の特徴が多分に学習された構造のものとなっていると考えられる。

表 2 は表 1 の学習ネットを使用して外挿 (学習パターン以外) の男女 30 人による男女識別結果である。32×32 学習ネットを使用して、32×32 をテストした場合 (90%) と、8×8 をテストした場合 (87%) の認

表 1 内挿パターンによる男女識別結果
Classification result of man-and-woman by learning patterns.

モザイク画像別学習ネット (学習は男女各15人、計30人)	テストに用いた内挿 30 人 (男女各 15 人) のモザイク画像別認識率 (%)		
	32×32	16×16	8×8
32×32	—	100	90
16×16	97	—	97
8×8	90	93	—

原画

32 × 32

16 × 16

8 × 8

図 1 モザイク顔画像の例
An example of mosaic facial images.

表 2 外挿パターンによる男女識別結果

Classification result of man-and-woman by facial patterns excluding learning patterns.

モザイク画像別学習 ネット (学習は男女各15人、 計30人)	テストに用いた内挿 30 人 (男女各 15 人) のモザイク画像別認識率 (%)		
	32×32	16×16	8×8
32×32	90	90	87
16×16	83	83	83
8×8	67	70	87

識率の差は3%であり、この結果においても表1同様、低周波成分の特徴が多分に学習されていることがわかる。すなわち、テストに用いた未知顔画像の低周波成分の特徴が男女識別の大きな要素となっていることを意味している。しかし、8×8の学習ネットにおいては、32×32のテスト結果が67%と内挿の場合と比べて23%も悪くなっている。また、8×8と32×32の学習ネットで8×8をテストした場合の認識率は、ともに87%である。なお、16×16の学習ネットを使用した場合のテスト結果においては、各サイズのモザイクとも安定して高い認識率が得られた。Harmonの人間の認識能力との関連性を考察する上でも、興味深い結果である。以上の結果は次のように解釈される。

基本的には男女識別情報は8×8画像中にほとんど入っている。32×32画像にはそれに加えて、基本的な識別にはあまり重要でない“雑音状”の高周波成分が含まれている。学習時にはこの“雑音”の効果を発揮し、識別面がパターン点すれすれに設定されるのではなくて雑音分だけ余裕をもって設定される。したがって、その分だけ外挿データに対する汎化能力が増す。このことは、学習時にはある程度分解能の高い顔で学習ネットを構築しておけば、テスト顔としては、多少ピンボケの顔や、あまりめだたないメガネを掛けた顔ぐらいであれば充分識別可能である事実をよく説明できる。8×8画像は、人間には顔であることすら認識できない低分解能画像であるにもかかわらず、BPネットは外挿データに対して87%の高率で男女識別ができる能力には驚かされる。

表3は、男女各5人、計10人の個人の識別の結果である。カテゴリー数の多い(2→10)分だけ全体に数値が低いが、表1とよく似た傾向の結果が得られている。すなわち、個人識別の場合も8×8画像中の低周波成分が主要な個人識別情報となっていることがわかる。表3の個人識別の場合も、人間には顔かどうか定

表 3 個人識別結果

Individual classification result.

モザイク画像別学習 ネット (学習は男女各5人、 計10人)	テストに用いた内挿 10 人 (男女各 5 人) のモザイク画像別認識率 (%)		
	32×32	16×16	8×8
32×32	—*	100	90
16×16	100	—	100
8×8	70	80	—

図 2 表1において8×8の学習したネット上に32×32のモザイク画像を追加学習した学習軌跡
A locus learned mosaic images of 32×32 in addition to network learned by 8×8 in Table 1.

かでない8×8画像による識別において高い正解率を示している。ただし、これはBPネットの内挿データに対する認識能力が高いことを示してはいるが、外挿データに対する高い認識能力を必ずしも保証しているわけではない。残念ながら、外挿データに対する認識実験はデータの都合上まだ行っていない。

図2は表1で行った男女識別(内挿)において、100%収束後の8×8の学習ネット上に32×32のモザイク画像を追加学習させた学習軌跡である。低周波成分の特徴を持った学習ネットの上に高周波成分の特徴のみを追加していると思われるが、最初(最初の初期ウエイトは乱数値)から32×32を学習させるに必要な回数を1として、追加学習に要する回数はほぼ1/2である。図3に示す個人識別(表3の*における学習曲線)においては追加学習の回数は1/4であり、表3の結果から察するに、追加学習で全特徴の30%(70%→100%)に相当する高周波成分の特徴を学習したものと考えられる。また、図4は表3の* (学習、テストとも32×32のモザイク画像での実験)の学習において、中間層の数を10から20にした場合の学習軌跡の比較である。

図3 表3において8×8の学習したネット上に32×32のモザイク画像を追加学習した学習軌跡
A locus learned mosaic images of 32×32 in addition to network learned by 8×8 in Table 3.

図4 表3において中間層のユニット数を変えた学習軌跡
A learning locus when number of hidden units in Table 3 is changed.

表1の男女識別の場合においては、図には示さないが、中間層の数を倍にしても学習回数にはさほど影響が見られなかったが、10人の個人識別になると、図に示すように少し早い収束結果が得られた。

4. む す び

BP法を利用して、髪を除いた顔を8×8、16×16、32×32の分解能の異なるモザイク画像に変換して男女および個人識別の実験を行った。その結果、次のような結果が得られた。

(1) 人間では顔を認識するのに16×16以上の分解能がなければ不可能とされているが、BP法においては、8×8の分解能の画像で、男女識別(外挿の男女各15人、計30人のテスト)で87%、個人識別(内挿の男女各5人、計10人)では100%の認識率が得られた。

(2) 8×8、16×16、32×32のいずれかのモザイク画像で学習したネットを使用して、それと異なる分解能のモザイク画像を入力してテストした結果、男女識別では内挿で90%以上、外挿では67%以上、個人識別では70%以上の認識率が得られた。その結果を解析してみると、学習されたネットの特徴として、基本的には低周波成分の特徴が多分に学習され、高周波成分の特徴はそれの上に追加されるような形となって存在しているように思われる。なお、16×16以上の分解能の学習ネットを使用しての各モザイクに対する認識結果においては、ほぼ安定した高認識率が得られた。このことは(1)で述べた人間の認識能力とBP法との関連性を検討する上においても興味ある結果である。

(3) 8×8で学習したネット上に32×32のモザイク画像を追加学習した場合と、最初から32×32のモザイク画像を学習した場合とで、収束するまでの学習回数を調べた結果、前者の方が後者に比べ男女識別(内挿)で1/2、個人識別(内挿)で1/4の回数で収束した。この結果からも、(2)で述べた高周波成分が追加学習されたことの一端がうかがわれる。

今後の課題としては、ニューラルネットを用いて、

- (1) 高周波成分のみでの識別、例えば32×32と8×8の差分画像での識別
- (2) ラプラシアン画像とモザイク画像間での識別
- (3) 顔の中央部分を利用するの識別

などを行う予定である。

なお、実験に際しては「日本電気技術情報システム開発(株)」の援助を得た。

【参 考 文 献】

- 1) 小杉 信：“ニューラルネットを用いた顔画像識別の一検討”，テレビ技報，14，50，pp. 7-12 (Sep. 1990)
- 2) Vicki Bruce(著)，吉川左紀子(訳)：“顔の認知と情報処理”，pp. 59-64，サイエンス社 (1990)

かわい ひでお
河合 秀夫 昭和46年，立命館大学理工学部卒業。同年，大阪大学基礎工学部教務職員。61年，大阪電気通信大学短期大学講師，現在，同大助教授。コンピュータビジョン，医用画像処理，神経回路網の研究に従事。工学博士。正会員。

たむら しんいち
田村 進一 昭和46年，大阪大学大学院博士課程修了。同年，同大基礎工学部助手，現在，同大医工学部教授。コンピュータビジョン，画像処理，神経回路網などの研究に従事。工学博士。正会員。