

ニューラルネットワークと KL 展開による画像圧縮

島田 恭宏・載 江濤*・塩野 充

岡山理科大学工学部情報工学科

岡山理科大学大学院工学研究科修士課程電子工学専攻*

(1993年 9月30日 受理)

1 まえがき

近年、データ圧縮の新たな手法として、ニューラル・ネットワークを用いた手法（ニューロ・コーデック）が提案されている¹⁾²⁾、ニューロ・コーデックは、階層型ネットワークをデータ圧縮に応用するものであり、同一数 α 個のユニットを持つ入力層、および、出力層、 β 個のユニットを持つ中間層から構成される階層型ネットワークに、入力パターン、および、教師信号として同一のパターンを与え、恒等写像をバックプロパゲーション・アルゴリズムで学習させるものである（図1）。中間層ユニット数 β が入力層、および、出力層のユニット数 α よりも小さい場合、階層型ネットワークは、入力層から中間層への変

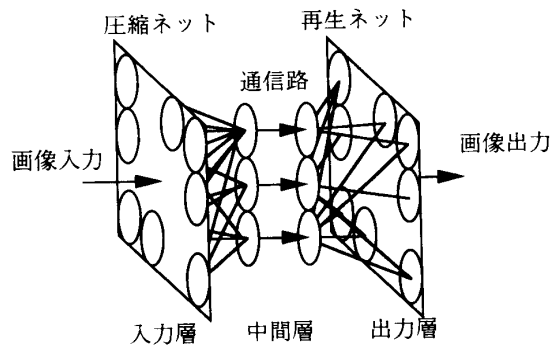


図1 ニューラルネットワークによるデータ圧縮の概念

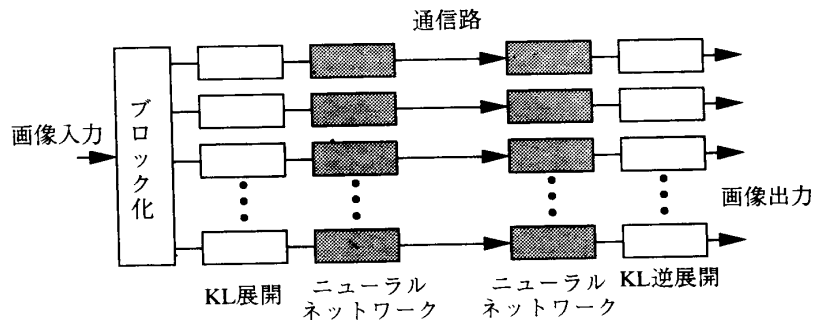


図2 実験モデル

換で、入力パターンを効率良くコーディングし、中間層から出力層への変換で入力パターンをデコーディングする。この時、入力層から中間層へのネットワークは、情報圧縮フィルタとして、中間層から出力層へのネットワークは、復号フィルタとして機能する。

ニューラルネットワークの効用は、学習によって自動的に変換規則を獲得することであり、獲得した変換規則は、特定の学習パターンのみならず、類似の特性を有する未知パターンに対しても適用できると考えられる。しかし現実的には、画像圧縮にニューラルネットワークを適用した場合、再生画像のSNを最大にするという意味で最適な変換規則を獲

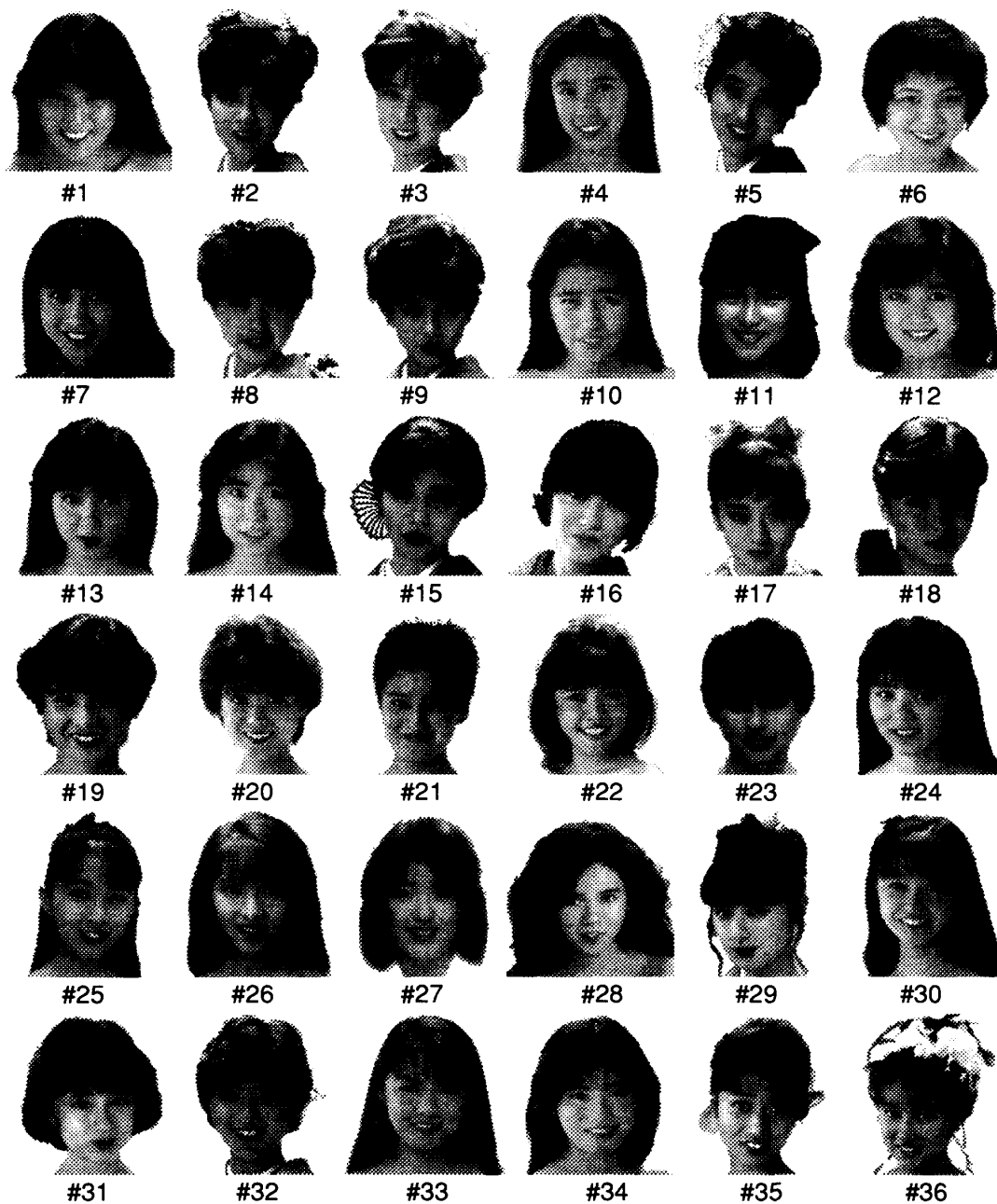


図3 実験に使用した画像

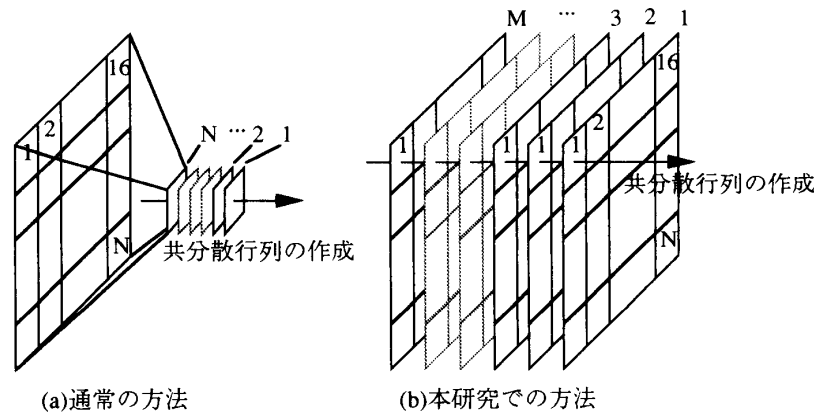


図4 共分散行列の求め方

得するためには、膨大な学習パターンと学習回数、および、これに見合うネットワーク容量を必要とし現実的ではない²⁾。よって、ネットワークが収束しやすい特徴量に画像を変換する必要がある。そこで本論文では、まず画像データから KL 展開³⁾⁻⁶⁾により特徴抽出を行う。次にそれを入力データとしたニューラルネットワークによる画像圧縮実験を行い、圧縮効率と再生画像の品質調査を行った。

図2に本論文での実験モデルを示す。処理は入力画像のブロック化、KL 展開による特徴抽出部、ニューラルネットワークによるデータ圧縮の3部から構成される。まず入力画像は、 8×8 画素の部分画像に分割され、この領域が以後の特徴抽出、圧縮処理の対象となる。特徴抽出部では、各部分画像に対し、KL 展開により展開係数を求め特徴抽出を行う。この特徴量を入力信号として圧縮部の3層構造のニューラルネットワークを駆動しデータ圧縮を行う。画像の再生は、ニューラルネットワークの出力層の出力信号を逆展開することで得る。

2 実験に使用したデータ

実験に用いた顔画像データは、イメージスキャナより読み取り綿密度100dpi, モノクロ256階調で入力した。これらの画像は、雑誌よりサンプリングしたため、背景が一様でなく、KL 展開による特徴抽出に対して顔以外のテクスチャの特徴が抽出される恐れがある。背景情報は現段階では対象としないので、手操作により顔部分のみを抽出した。これらの画像に対しては、大きさと位置の正規化は行っていない。画像の大きさは、 128×128 で36パターンを用意した。このうち30パターンまでを学習パターンとして、6サンプルを未知パターンとして用いた。図3に全パターンを示す。なお画像のブロック化は、単純に 8×8 (=64)画素毎に分割し、256個の部分画像を作成する。

3 特徴抽出

3.1 KL 展開による特徴抽出

通常、画像に対する KL 展開は、対象画像を部分画像に分割し、その標本値系列の共分散行列を作成し、この行列の固有ベクトルで直交変換を行う(図 4(a))。これによって部分画像間の濃度分布に関する特徴が抽出される。しかしこの手法では、画像毎に固有ベクトルを求める必要があり、これにより得られる特徴量は画像間で一義的に評価できず、ニューラルネットワークの入力信号作成のための特徴抽出過程としては不相当である。

本論文では、ニューラルネットワークによるデータ圧縮を行うため、類似の特性を有する画像(顔画像)を対象としている。複数画像(固有ベクトルを求めるための学習パターン群)における部分画像内で KL 展開を行うことで、画像間の冗長性を抑圧し、顔画像における特定部分のテクスチャの特徴を抽出する(図 4(b))。具体的には、まず学習パターン群すべてに対し、ブロック化(8×8画素)を行い、部分画像を作成する。次に、各学習パターンにおいて同位置に属する部分画像の集合から同分布を表現するために最適な固有ベクトル(KL 展開における変換核)を求める。これにより求められた固有ベクトルは、顔の特定部分の特徴を表現しているものと考えられ、学習パターン以外の画像に対しても有効である。実際の特徴抽出は、ここで得られた変換核に対し、部分画像を射影することによって行われる。本来、このように複数のパターン間で直交変換を用いて特徴を抽出する場合、主成分変換法⁽⁴⁾と呼ぶべきかもしれないが、文献などでも分類が曖昧なため、本論文では KL 展開と呼ぶことにした。詳細を以下に示す。

KL 展開の学習過程において、学習パターンが M (最大で30) 個あるとして、そのすべてに対し 8×8 画素のブロック化を施す。その結果、得られる p 番目の学習パターンの n 番目の部分画像 ($L=64$ 次元) を $f_n^p (p=1,2,\dots,M (<=30); n=1,2,\dots,N=256)$ で定義する。ここで、KL 展開の対象となるパターン集合は $f_n^1, f_n^2, \dots, f_n^M$ である。第 n 番目の部分画像のパターン集合より作成される共分散行列 C_n を次式で定義する。

$$C_n = \frac{1}{M} \sum_{p=1}^M (f_n^p - \mu_n)(f_n^p - \mu_n)^T \quad (1)$$

ここでは、 μ_n はパターン集合 $f_n^1, f_n^2, \dots, f_n^M$ の平均パターンである。 C_n より固有値、固有ベクトルを計算し、固有値を降順に並べ直した集合を $\{\gamma_m\} (m=1,2,\dots,L)$ とし、各固有値に対応する固有ベクトルを $\{\phi_m\}$ とする。ここで、固有ベクトルをどこまで採用するかが問題となるが、本論文では、式(2)で定義される累積寄与率 η_n を導入し、これが 1 となる位置、すなわち、原パターンを完全に再現できる位置 (K_n) まで固有ベクトルを採用し、変換核 $\{\phi_{K_n}\}$ を構成する。このように、部分画像毎に固有ベクトルを求め、それぞれに累積寄与率を求め固有ベクトルの採用を決定しているため、各部分画像で得られる特徴量の次元数は異なる。部分画像の原パターンの次元数は 64 次元であり、本研究での学習パターン数は最大でも 30 である。よって、完全に原パターンを復元するまで固有ベクトルを採用したとしても、

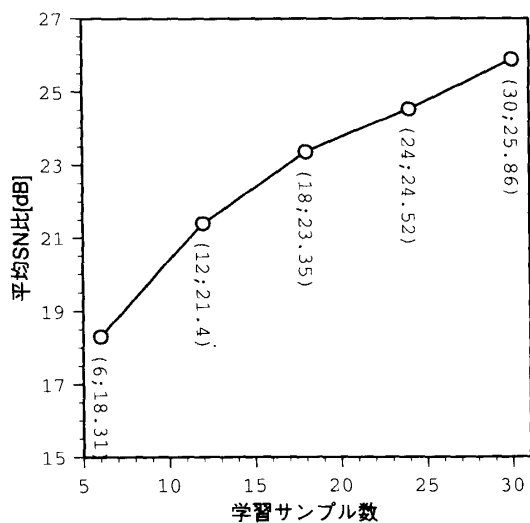


図5 学習パターン数による再生画像の品質

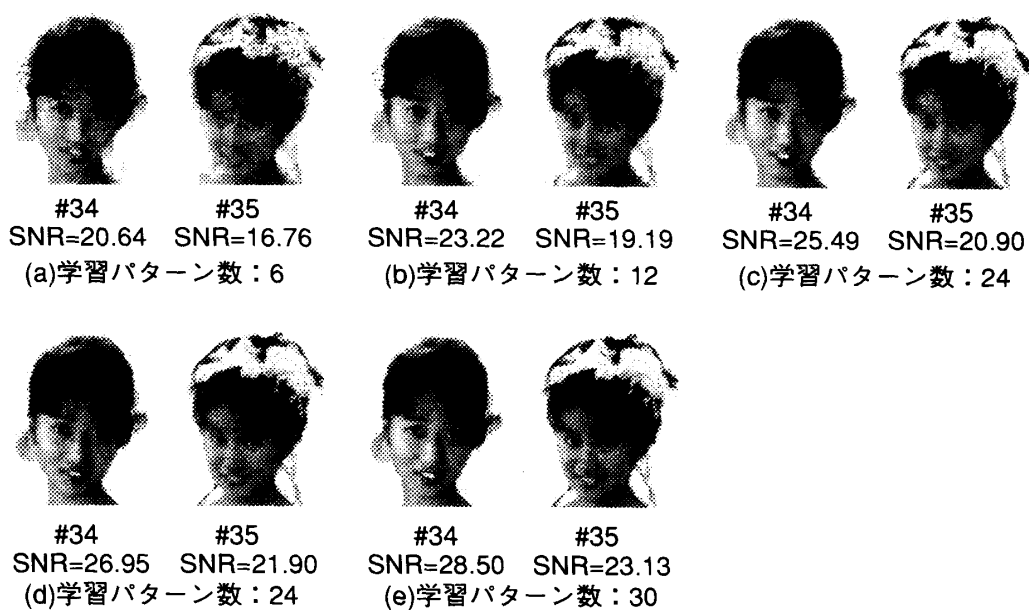


図6 特徴抽出後の画像の再生例（未知パターン）

表1 特徴抽出過程における次元縮退の効果

学習パターン数	6	12	18	24	30
縮退後次元数	956	2073	3153	4160	5250
圧縮率 [%]	5.83	12.65	19.24	25.39	32.04

最大で29の固有ベクトルを採用することで可能である。よって、特徴抽出過程による次元縮退の効果も期待できる。

$$\eta_n = \sum_{m=1}^{K_n} \gamma_m / \sum_{m=1}^L \gamma_m \quad (2)$$

この変換核を用いた特徴量の抽出は、求められた変換核 $\{\phi_{K_n}\}$ からなる K_n 行 L 列の行列を Φ_n とすると、抽出されるパターン、すなわち特徴量 g_n^p は、次式で求められる。

$$g_n^p = \Phi_n f_n^p \quad (3)$$

また、画像の再生も同様に行われる。実際の実験では、各部分画像毎に変換核 (256個) を求め、未知パターンも含め全実験パターンのすべての部分画像に対して特徴量を求めてお

表2 ネットワーク構成

入力層 ユニッ ト数	ネット ワーク 数	中間層のユニット数					
		2†	3†	4†	5†	6†	7†
7	5	3	2	1	1	1	1
8	2	4	2	2	1	1	1
9	5	4	3	2	1	1	1
10	4	5	3	2	2	1	1
11	4	5	3	2	2	1	1
12	8	6	4	3	2	2	1
13	2	6	4	3	2	2	1
14	2	7	4	3	2	2	2
15	3	7	5	3	3	2	2
16	4	8	5	4	3	2	2
17	5	8	5	4	3	3	2
18	1	9	6	4	3	3	2
19	4	9	6	4	3	3	2
20	4	10	6	5	4	3	2
21	2	10	7	5	4	3	3
22	2	11	7	5	4	3	3
23	3	11	7	5	4	3	3
24	5	12	8	6	4	4	3
25	5	12	8	6	5	4	3
26	5	13	8	6	5	4	3
27	6	13	9	6	5	4	3
28	4	14	9	7	5	4	4
29	127	14	9	7	5	4	4

(注) ただし†を有する数字は、入力層と中間層のユニット数の比を表わす。

く。

3.2 評価実験

本論文で用いた特徴抽出法では、抽出した特徴量がどの程度、原画像を表現しているかが後の圧縮、再生後の画像の品質に影響を与える。先に述べたように、学習画像に対しては完全に原パターンを再生できる位置まで固有ベクトルを採用しているゆえ問題はない。しかし未知パターンに対しては、どの程度表現できるか不明である。そこで未知パターンに対し特徴抽出後、画像を再生し、その品質を評価することにした。この時点での未知パターンの画像品質を決定する要因は、学習パターン数と考えられる。よって、学習パターン数をパラメータとして特徴抽出を行い、再生画像の評価を行う。再生画像の品質の評価尺度として、次式で示す規格化 2 乗平均誤差を用いた。これは、部分画像毎ではなく、再生画像全体で評価を行った。

$$SNR = 10 \log_{10} [E \langle F_{i,j}^2 \rangle / E \langle (F_{i,j} - G_{i,j})^2 \rangle] \quad (4)$$

ここで、 $F_{i,j}$ は原画像を、 $G_{i,j}$ は再生画像であり、 $E \langle \rangle$ は画像全体での平均を示す。実験にはワークステーション (NWS-3860) を用い、プログラムは C 言語で記述した。

3.3 評価結果

図 5 に再生画像の評価結果を示す。横軸は学習パターン数を、縦軸は未知パターンの平均 SN 比を示している。また、図 6 に未知パターンの再生画像の一例を示す。図からもわかるように、学習パターン数が多いほど、未知パターンの再生画像の品質が良いことがわかる。パターン空間上では、学習パターン数が増えるに従いパターンは広く分布することとなる。これに伴い、採用する固有ベクトルの最大数も増加する。各固有ベクトルは部分画像の特徴を表現しているものと考えられ、多様なパターンを用いることで、得られる固有ベクトルは、部分画像のテクスチャの特徴を示すものと考えられる。これによって、未知パターンに対しても良好な結果が得られるものと考ええる。

次に次元縮退の効果を表 1 に示す。採用する固有ベクトルの数は、学習パターン数に依存しており、学習パターンが少ない場合、大きな縮退の効果がみられる。これは、採用する固有ベクトルの最大値に依存しており、学習パターン数が増えるにしたがって次元縮退の効果は薄れて行く、しかし、再生画像の SN 比は先の議論でも記したように学習パターン数が増えるにしたがい改善される。本実験では、KL 展開を特徴抽出過程として位置付

表 3 ニューラルネットワークによるデータ圧縮率

ユニット比	2	3	4	5	6	7
圧縮後次元数	2981	2093	1711	1380	1202	1162
対画像圧縮率 [%]	18.19	12.77	10.44	8.42	7.34	7.09
対特徴量圧縮率 [%]	56.78	39.87	32.59	26.29	22.90	22.13

けており、圧縮過程はニューラルネットワークで行うものとした。よって、KL展開による次元縮退は、副次的な効果とみなす。そこで以後のニューラルネットワークでの圧縮実験は、SN比の最も良かった学習パターン数が30の場合の特徴量を用いて圧縮実験を行う。

4 ニューラルネットワークによる画像圧縮

4.1 ネットワーク構成

ニューラルネットワークによる圧縮部では、KL展開で求められた特徴量の次元数より低次元のデータに変換して圧縮機能を実現するため、抽出された特徴量の次元数のユニットを持つ入出力層、および、それに比べユニット数を小さく構成した中間層からなる3層ニューラルネットワークを用いた。学習則としてバックプロパゲーション則を用いる。入出力層のユニット数は、抽出された特徴量の次元数に対応する。また、中間層のユニット数は、入出力層のユニット数との比で機械的に決定した(表2参照)。ただし、特徴量の次元数(=ネットワークの入出力層のユニット数)が1~6の場合はデータが少量のため、ニューラルネットワークによる圧縮は行わない。

学習過程では、学習すべき画像データから得られた特徴量を入力層に与え、出力層には教師信号として入力層と同一の信号を与える。入出力信号間の評価関数としては2乗平均誤差を用いた。なお本論文では、中間層での量子化、および、量子化符号化から通信路符号への冗長度抑圧符号化については扱っていない。圧縮率の評価は、中間層のユニット数に対する入力層のユニット数の比、および、部分画像の次元数 L (=64)に対する中間層

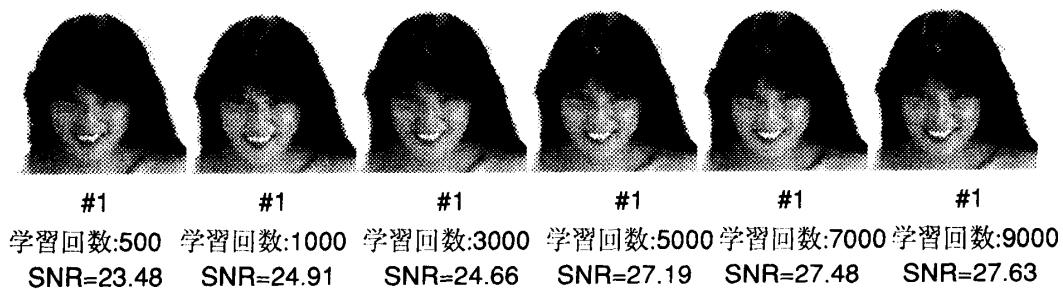


図7 再生画像例(学習パターン)



図8 再生画像例(未知パターン)

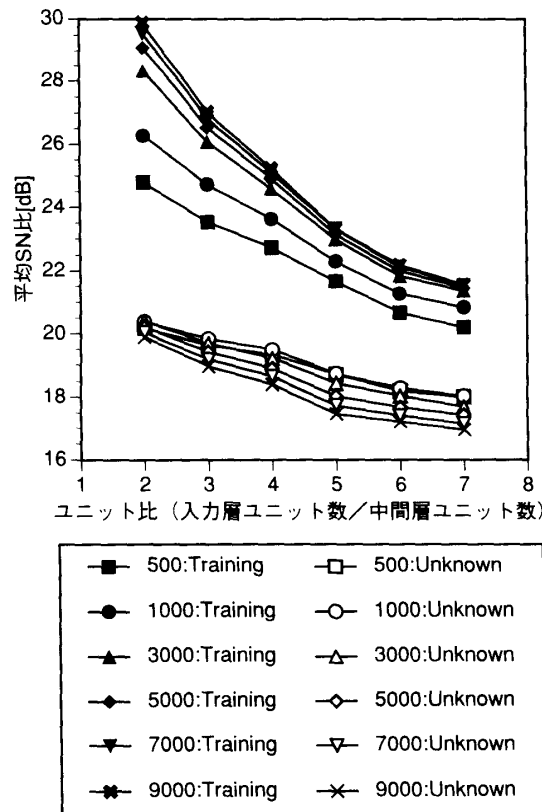


図9 中間層ユニット数による再生画像の品質

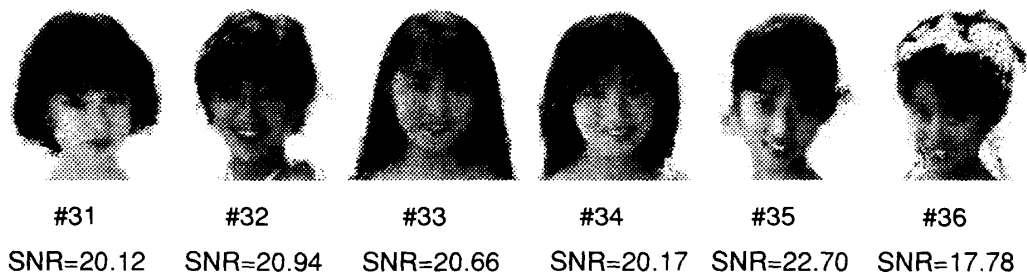


図10 平均 SN 比が最良の未知パターンの再生例 (学習回数: 1000, ユニット比: 2)

のユニット数の比を用いることとした(表3参照)。また、再生画像の評価には式(4)に示した規格化2乗平均誤差を用いた。

4.2 圧縮実験, および, 結果

ネットワークの運用形態は局所並列型とし、部分画像に対してそれぞれ独立に動作する1つのネットワークを用意した。実験には56台のワークステーション(SPARC station IPC)を用い、212のネットワークを分散させ処理を行った。ネットワークの学習は、30枚の学習パターンから得られた部分画像の特徴量を1セットとして行った。

図7に学習パターンの画生画像例を示す。この時のネットワーク構成は、表2に示した

ユニット比が2の場合である。今回の実験では、学習回数500回から実験を行った。学習初期の段階では、顔の外郭部にブロック歪が見受けられるが、学習が進むにしたがい減少している。顔の外郭部は各個人の髪形の違いや正規化していない画像を対象としていることから、各画像間でのテクスチャの差異が顕著な領域であり、変換規則を学習するためにはある程度の学習回数を必要とすることがわかる。学習回数9000回で平均 SN 比が最も良く 29.88dB を得た。次に未知パターンの再生例を図8に示す。学習パターンの場合と同様の条件下で平均19.90dBであった。画像の細部構造の劣化やブロック歪は大きいものの、各画像の特徴は再現されていることがわかる。

次に学習回数数と入力層と中間層のユニット数の比をパラメータとした場合の SN 比の変動を図9にまとめた。図10に未知パターンの平均 SN 比が最も良かった場合（ユニット比2，学習回数1000回）の再生画像例を示す。まず、学習回数に注目した場合、学習パターンについては、先の画像例からもわかるように学習が進むにしたがい SN 比の向上が認められる。しかし、未知パターンに対しては、学習回数が少ない場合が SN 比が良いことがわかる。学習回数が進むに従い、ネットワークが学習する変換規則が学習パターンに特定されてしまい、過学習の状態にあるものと予想される。入力層と中間層のユニット数の比に対する SN 比は、中間層のユニット数の減少とともに SN 比が減少している。入出力層のユニット数に対する中間層のユニット数の比は、学習データから圧縮、再生の変換規則の精度とその記憶容量を与えるものと考えられ、中間層のユニット数の減少は直接、再生画像の品質の低下につながっている。

5 む す び

本論文では、顔画像に対する KL 展開とニューラルネットワークによるデータ圧縮実験について報告を行った。特徴抽出過程における KL 展開では、主成分変換法的な手法に基づき、学習、未知パターン双方に対して特徴抽出後の再生画像は良好な結果を得ることができた。また、次元縮退の効果も大きく、ニューラルネットワークへの特徴の抽出過程としては有効である。しかし、今回の固有ベクトルの採用方法では、学習パターンが増えるにともない縮退の効果は薄れることとなり、今後検討を要する課題である。ニューラルネットワークによる圧縮実験では、特徴抽出後の再生画像で SN 比が最も良かった学習サンプル数30枚の場合に限定し、学習回数、および、中間層のユニット数をパラメータとして実験を行った。学習パターンについては、ユニット比が2，学習回数9000回の場合で平均 29.88dB という良好な結果を得ることができた。しかし、未知パターンに関しては各画像の特徴は表現しうるものの、細部の劣化やブロック歪が大きく満足できる結果ではない。また学習パターンの場合とは逆に学習が進むにしたがい未知パターンの SN 比は減少している。これはネットワークの過学習の問題と考えられ、今後、学習回数を含め、ネットワークの特性を決定するパラメータの設定について検討を要する。

最後に本研究に対し有益な御討論を頂いた本学情報工学科講師大倉 充博士、本学博士課程大学院生島田英之氏に深謝する。

参考文献

- 1) 電子情報通信学会編：“デジタル信号処理ハンドブック”，オーム社（1993）。
- 2) 曾根原登，川人光男，三宅 誠：“ニューラル・ネットワークによる画像データ圧縮（Neuro-CODEC）の検討”，信学技報，IE88-62，pp. 57-64（1988）。
- 3) エルッキ・オヤ著，小川英光，佐藤 誠訳：“パターン認識と部分空間法”，産業図書（昭和61年）。
- 4) 原島 博監修：“画像情報圧縮”，オーム社（平成3）。
- 5) 高木幹雄，下田陽久：“画像解析ハンドブック”，東京大学出版会（1991）。
- 6) 舟久保登：“視覚パターンの処理と認識”，啓学出版（1990）。

Facial Image Data Compression using K-L Expansion and a Neural Network

Yasuhiro SHIMADA, Dai Jiang TAO*, Mitsuru SHIONO

Department of Information & Computer Engineering,

Graduate School of Electronic Engineering,*

Okayama University of Science

Ridaicho 1-1, Okayama 700, Japan

(Received September 30, 1993)

In this paper, K-L expansion and the artificial neural network are applied the data compression of facial image data. This method is composed of feature extraction process and data compression process. The input data to the neural network is the linear feature vector extracted from an original image using K-L expansion. Using this method, an image data can be compressed to 18%. Some experimental results are shown with the compression ratio and S/N ratio.