

目の分類を行うニューラルネットの検討

Neural Network for Classification of Eye's Shape

磯 俊樹 崎田一貴 大塚作一 小杉 信
Toshiki ISO Kazutaka SAKITA Sakuichi OHTSUKA Makoto KOSUGI

NTTヒューマンインタフェース研究所
NTT Human Interface Laboratories

あらし 主観評価実験により、目の形態についての分類を行ったのち、目の物理的特徴量を入力とし、主観評価量を出力するような3層のニューラルネットの検討を行った。その結果、主観評価量のようなばらつきのある量でも学習できること、学習後のネットの重みから得られた物理的特徴量と主観評価量の関係は重回帰分析による結果と類似していることが分かった。また、学習データが少ない場合でもK近傍内挿学習法における近傍選択距離尺度を相対的基準ではなく、絶対的基準にすることで汎化能力が向上することが分かった。

Abstract Classification of eye's shape by the three layered back propagation network(BPN) is discussed. Eighteen physical features of eye's shape are used as input data, and three features obtained from subjective tests are used as output data. As for eye's shape classification, similar results are obtained between BPN and the conventional multiple regression analysis. In case of using limited learning data, unknown data can be classified by using k-neighbor interpolation training, where the decision for choice of neighbor data is based on the criterion of absolute standard of sample data.

1. はじめに

顔の分類・識別は、知的画像通信やセキュリティシステムにおける個人識別[1]などの分野で適用が期待できる。しかし、現状では、顔の特徴量を抽出する方法は確立されておらず、特徴量を基に適応的な学習を行い、顔画像を分類・識別する手法もほとんど見あたらない。また、人の主観が加味されている手法も少ない。

一方、認識・制御の分野で注目されているニューラルネットは、学習データに応じてネットの重みを修正していくので高い学習能力を有している。また、発火関数に非線形関数を用いるので線形関数を用いた場合に比べて複雑な領域に分けることが可能である。これらの特徴を生かすことにより、従来の手法を越えた顔の分類・識別能力が期待できる。例えば、各顔部品の分類・識別にサブニューラルネット用い、それらのネットを並列かつ多層に組み合わせることで、柔軟で効率的な分類が行える可能性がある(図1)。また、個々の顔部品のサブネットで行う分類に主観による分類尺度を用いることにより、人間に近い判断を行う顔画像分類・識別システムの可能性も期待できる。

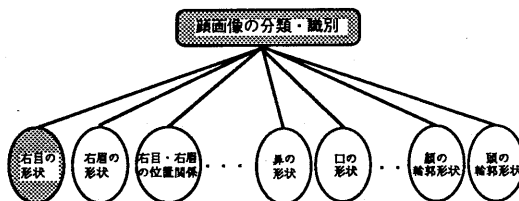


図1：階層化による顔画像分類・識別の例

今回の検討は、顔画像分類・識別システムの検討の第一段階として、まず、顔の分類・認識の際に強い影響を持つ、目の形態分類を目標とした。そして、人間の主観的な分類特性に近い出力を示すようなニューラルネットを構築するために、次のような検討を行ったので報告する。

- ・主観評価実験に基づく目の分類。
- ・ニューラルネットによる目の分類。
- ・ニューラルネットを用いた目の分類の汎化性。

2. 主観評価実験に基づく目の分類[2]

人間が顔画像をどのような特徴に基づいて識別して

いるかを明らかにすることにより、顔画像認識システムの性能向上が図れる可能性がある。しかし、人間が顔画像を認識する際、目の情報が重要であることは既に明らかにされている[3]が、目に関してどのような特徴が心理的に大きな要因であるかについては充分検討されていない。そこで、人間がどのような観点から目を識別しているかを心理実験により調べ、その結果に基づいて目に関する分類項目を抽出した。

2.1 実験概要

目の特徴に対する直感的な判断を端的な言葉で表現することは困難である。そこで、形容詞群を用いて主観評価実験を行い、主成分分析[4]により目の特徴に関する心理要因を抽出した。

2.2 形容詞の抽出

目に関する形容詞群を網羅する必要がある。そこで、形容詞辞典、分類用語辞典、人相学の本[5]から目に関する形容詞(色関係を除く)の抽出を行った。その結果、次に示す45種類の形容詞を選択した。

- | | | |
|-------------|-------------|------------|
| 1.上目 | 2.下目 | 3.伏せ目 |
| 4.上がり目 | 5.下がり目 | 6.垂れ目 |
| 7.吊り目 | 8.出目 | 9.大目玉 |
| 10.どんぐり目 | 11.くぼんだ目 | 12.細目 |
| 13.切れ長の目 | 14.二重瞼 | 15.とろんとした目 |
| 16.ぱっちりした目 | 17.ぎょろぎょろ目 | 18.淫らな目 |
| 19.陰湿な目 | 20.きらめく目 | 21.虚ろな目 |
| 22.優しい目 | 23.みずみずしい目 | 24.暗い目 |
| 25.鋭い目 | 26.張り裂けそうな目 | 27.涼しい目 |
| 28.卑しそうな目 | 29.和やかな目 | 30.血走っている目 |
| 31.熱心な目 | 32.無邪気な目 | 33.威厳のある目 |
| 34.すがすがしい目 | 35.神秘的な目 | 36.腫が大きい目 |
| 37.腫のはっきり目 | 38.大きい目 | 39.三角形の目 |
| 40.上瞼のゆるんだ目 | 41.たるんでいる目 | 42.狼のような目 |
| 43.丸目 | 44.小さい目 | 45.こせついた目 |

2.3 顔写真の主観評価実験

目の形態に関して平均的な主観評価を得るために10名の被験者(20代の女性)で主観評価実験を行った。

被写体：20～40代で、目・眉が髪の毛等で隠れていない顔の全体像を選んだ。102枚の白黒写真を用いた。

実験内容：目の形態に関する形容詞群それぞれについて、“よくあてはまる”、“ややあてはまる”、“あてはまらない”の3段階で評価を行った。被験者数は10名とし、評価時間は無制限とした。10

人の被験者の評点を平均した後、主成分分析を行った。

実験結果と考察：

表1：主成分分析結果

	変数名	因子負荷量	変数名	因子負荷量	寄与率
第1主成分	こせついた目	0.76	瞳のはっきり目	-0.92	31.2%
	暗い目	0.74	瞳が大きい目	-0.90	
	陰湿な目	0.74	ぱっちりした目	-0.90	
	小さい目	0.66	大きい目	-0.86	
	卑しそうな目	0.60	きらめく目	-0.82	
	淫らな目	0.54	どんぐり目	-0.80	
			丸目	-0.77	
			大目玉	-0.77	
			みずみずしい目	-0.74	
			ぎよろぎよろ目	-0.62	
			すがすがしい目	-0.60	
			二重瞼	-0.60	
			優しい目	-0.58	
		和やかな目	-0.58		
		無邪気な目	-0.53		
第2主成分	上がり目	0.72	下がり目	-0.81	15.2%
	吊り目	0.68	垂れ目	-0.78	
	鋭い目	0.64	とろんとした目	-0.78	
	狼のような目	0.59	たるんでいる目	-0.66	
			上瞼ゆるんだ目	-0.60	
			虚ろな目	-0.55	
			和やかな目	-0.53	
			優しい目	-0.53	
第3主成分	ぎよろぎよろ目	0.59			8.8%
	出目	0.53			

主成分分析の結果(表1)、3種類の有意な主成分が求められ、これらの累積寄与率は55%である。第1主成分：“ぱっちりした目”、“どんぐり目”といったような目の大きさ・丸さを表す要素が多く、またこれらは、“すがすがしい目”、“陰湿な目”といった目に対する“好感度”をあらわすものと強く関係があると考えられる。また、目の周囲部の状

態を表す“二重瞼”と目の全体の状態を表す“目の大きさ・丸さ”が同時に表れていることから、両者が強く関わっていることが示唆される。

第2主成分：“吊り目”に代表されるように目の“鋭さの度合い”を表していると考えられる。

第3主成分：“ぎよろぎよろ目”、“出目”に代表されるように目の“迫力度合い”を表していると考えられる。

また、上記3主成分を比較すると、第1主成分の寄与率が他の2成分に比べて2倍以上大きく、好感度が目に対する形態印象に強い影響を与えていることがわかる。一方で、主成分分析によって得られた各主成分の寄与率があまり大きな値でないことから、人間は言葉では表現できないような特徴で目を分類・識別していると考えられる。

本実験では、有意な成分である“好感度度合い”、“鋭さの度合い”、“迫力度合い”の3成分で目を分類していくことにする。

3. ニューラルネットによる目の分類

目の物理的特徴量を入力とし、人間の主観評価量を出力するようなニューラルネットの構築を行った[6]。また、その分類能力・特性を重回帰分析による結果と比較した。

3.1 入力要素となる物理的特徴量の抽出

顔画像データから顔の特徴量を画像処理により自動抽出する手法は確立されていない。そこで本実験では、まず、図2に示す目の特徴点をマニュアルで顔画像から抽出し、その特徴点に基づいて算出した18種の特徴量を入力ユニットとして用いることにした。

[A] 目の線画表現に対応する物理的特徴量

- 1) 目の長さ $|\vec{P}_4\vec{P}_4|/|\vec{Q}_1\vec{Q}_1|$
- 2) 目の幅 $|\vec{P}_8\vec{P}_8|/|\vec{Q}_1\vec{Q}_1|$
- 3) 目の縦横比 1)と2)の比 $|\vec{P}_8\vec{P}_8|/|\vec{P}_4\vec{P}_4|$
- 4) 目全体の面積 P_0, P_2, P_4, P_8 で囲まれる閉曲線内の画素数
- 5) 黒目の面積 P_2, P_12, P_8, P_11 で囲まれる閉曲線内の画素数
- 6) 黒目と白目の面積比 5)と4)の比
- 7) 目の傾き $\arccos((\vec{P}_4\vec{Q}_2 + \vec{Q}_2\vec{Q}_2)/|\vec{P}_4\vec{P}_4| |\vec{Q}_1\vec{Q}_1|)/\pi$
- 8) 黒目の位置 $|\vec{P}_0\vec{P}_0|/|\vec{Q}_1\vec{Q}_1|$

9) 目尻の傾き $\arccos((\overrightarrow{P_3P_5} \cdot \overrightarrow{Q_0P_0}) / |\overrightarrow{P_3P_5}| |\overrightarrow{Q_0P_0}|) / \pi$

10) 目尻の長さ $|\overrightarrow{P_3P_5}| / |\overrightarrow{Q_0P_0}|$

11) 上左目蓋の曲率 P0, P1, P2でできる曲率半径

12) 上右目蓋の曲率 P2, P3, P4でできる曲率半径

13) 下目蓋の曲率 P4, P8, P0でできる曲率半径

[B] 目の濃淡表現に対応する物理的特徴量

1) 黒目の平均濃度 PMの近傍濃度の平均

2) 白目の平均濃度 PH, PT各々の近傍濃度の平均

3) 上目蓋の平均濃度 P0, P1, P2, P3, P4各々の近傍濃度の平均

4) 下目蓋の平均濃度 P4, P6, P8, P10, P0各々の近傍濃度の平均

5) 黒目と白目の平均濃度比 1) と 2) の比

なお、物理的特徴量は次のように規格化して用いた。

- ・長さはすべて顔幅で規格化。
- ・面積は上記で定義した領域の面積で規格化。
- ・傾きの角度は、全て π で規格化。
- ・曲率、濃度とも各々の中で、最大のものを0.9、最小のものを0.1として規格化。

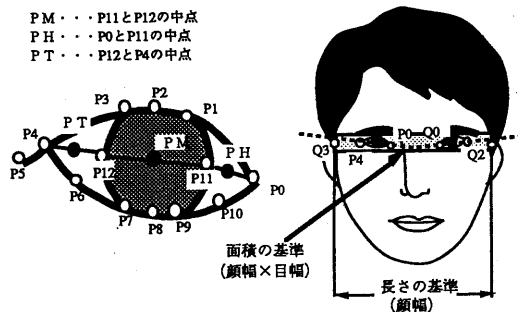


図2：目の特徴点

3.2 ニューラルネットの構成

3層のネットを以下のように構成し、学習則にはバックプロパゲーション[7]を用いた(図3)。

入力層・・・物理的特徴量18個

中間層・・・1~20個

出力層・・・主成分得点3成分

(なお、これ以降の実験では、一名の被験者が143枚のデータを主観評価して得られた結果を用いることとする。)

学習ステップ・・・0.4

慣性係数・・・0.9

データ数・・・学習用データ 20~105個、
未学習用データ 20枚

(なお、ここでの学習用データ、未学習用データの選別は、ウォード法[4]を用いたクラスター分析[4]の結果に基づいた。)

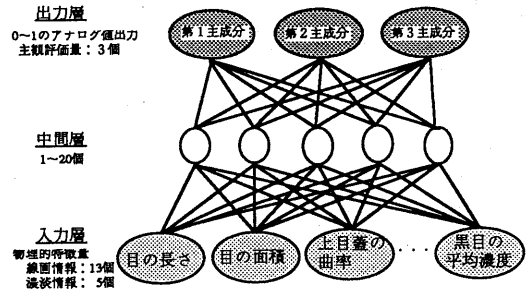


図3：ニューラルネットの構成

また、比較のため、目の物理的特徴量と主観評価量との対応付けを重回帰分析[4]を用いても行った。そして、分析で用いなかったデータについてもこの結果を適用し、実際の主観評価量との誤差を調べる。

なお、重回帰分析の際には情報量規準AIC[4]による変数選択法を用いた。

3.3 中間ユニット数を変えた場合の検討

実験内容：入力に物理的特徴量18個、出力に主観評価量3個を用いた場合について最適な中間ユニットの数を決める。

実験結果と考察：中間ユニット数を多くした方が学習データを精度良く学習できたが、反面、未学習データの誤差との差が大きくなり、汎化能力が落ちた(図4)。結局、中間ユニット数を2個にした場合が、学習データ、未学習データの誤差も抑えられ汎化性の良いネットが構成をすることができた。また、この時、約2000回の学習で物理的特徴量と主観評価量との対応付けができた。(なお、学習回数は、学習データを一通り学習して1回とする。) これより、”入力ユニット数18個、中間ユニット数2個、出力ユニット数3個”の構成でこれ以降の検討を進めることにした。

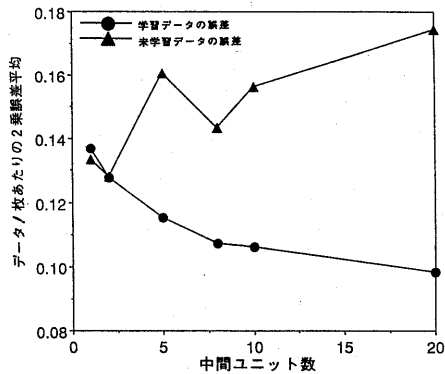


図4：学習データと未学習データの誤差(中間ユニット数変化)

3.4 学習後のネットの重み係数についての検討
 105個のデータについて学習を終えたネットの入力層と中間層間の重み係数を調べると(図6：図中の白の四角形は正の重みを、枠の四角形は負の重みをあらわしており、四角形の大きさは、重みの絶対値を表す。)、目の大きさや長さに関する入力ユニットとの結び付きが強く関わっている中間ユニットがあった。また、この中間ユニットは中間層と出力層との間では第1主成分を表している出力ユニットとの結び付きが強かった。このことは、目の大きさや長さに関する入力ユニットと第1主成分を表す出力ユニットが強い結び付きがあることを示していると考えられる。また、同様に目の傾きの入力ユニットとの結び付きが強い中間ユニットは、第2主成分を表している出力ユニットとの結び付きが強かったことから、目の傾きに関するユニットと第2主成分を表す出力ユニットも強い関係にあると考えられる。

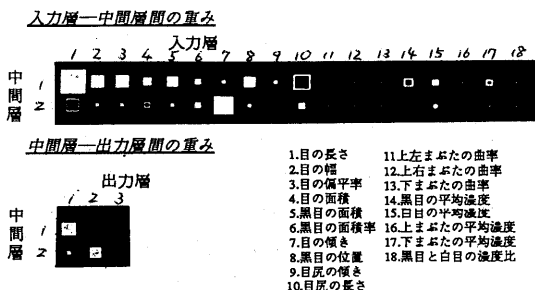


図6：学習後のネットの重み係数

また、このとき重回帰分析によって得られた物理的

特徴量と主観評価量との関係は表2のようになり、

第1主成分：目の幅・目全体の面積等の大きさを表す物理的特徴量との結び付きが強く、また、白目・黒目の濃度等との結び付きも強い。

第2主成分：目の傾きに関連するものと結び付きが強い。

ことが分かった。これは、学習後のネットの重み係数から得られるものと同様な結果を示した。これより、学習データ数が十分ある場合での本実験のニューラルネットは、重回帰分析による手法と類似した動作をしていたものと考えられる。

また、第3主成分に関しては、他の2つの成分に比べ特に強い関係を得ることはできなかった。これは、第3主成分を表現する場合、今回用いた物理的特徴量以外の別な特徴を用いる必要を示している。

	変数	F値(危険率)	重相関係数
第1主成分	目の幅	26.615(0.00000)	0.72670
	目の縦横比	9.770(0.00237)	
	黒目の面積	3.313(0.07194)	
	目全体の面積	4.547(0.03562)	
	黒目の面積比	5.622(0.01980)	
	目尻の傾き	5.875(0.01729)	
	下瞼の曲率	5.025(0.02736)	
第2主成分	黒目の濃度	7.235(0.00848)	0.53952
	白目の濃度	14.790(0.00022)	
	目の長さ	16.039(0.00012)	
	目全体の面積	12.033(0.00077)	
第3主成分	目の傾き	23.654(0.00000)	0.32176
	白目の濃度	6.555(0.01196)	
	下瞼の濃度	6.578(0.01180)	
		3.531(0.06312)	

3.5 学習データ数を変えた場合の検討

実験内容：学習データ数を変えた時の学習データ、未学習データの誤差の差を調べることにより、このネットに必要な学習データ数を検討する。

学習データ数については、20、50、80、105個の場合について行う。

実験結果と考察：学習データ数を多くするに従って、

学習データと未学習データとの誤差の差が小さくなる(図5)。今回の場合、80個ぐらいのデータ数で学習データと未学習データの誤差の比も1.1倍程度となり、かなりの汎化性を得ることができた。この時のデータ数は、未知変数(各層間の重み係数の数)のほぼ2倍の学習データ数に相当する。このような傾向は、重回帰分析を用いた場合もほぼ同様な結果を示した。ただし、学習データ数が20個の場合については、ニューラルネットを用いた方が学習データと未学習データの誤差の差が少なかった。

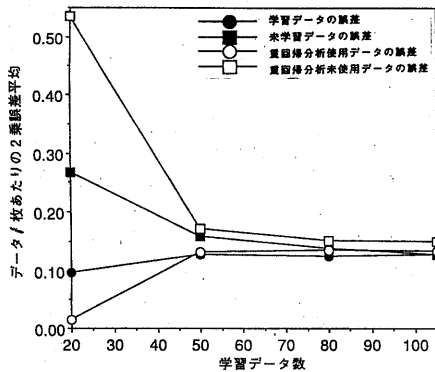


図5：学習データと未学習データの誤差(学習データ数変化)

4. ニューラルネットを用いた目の分類の汎化性

ニューラルネットで汎化性を持った分類を行うためには、やはり対象の性質を良く表す学習データの数が多く必要になる。本実験では、約80種(本ネットのリンク総数の約2倍の数)の学習データを選ぶことで十分学習を行うことができた。しかし、本研究で教師信号として用いた主観評価量などは、統計的に信頼できるものを大量に得ることが難しいものである。また、データが少ない場合には、データの母集団の性質を表すような学習データを選んでも、汎化性を備えた分類を学習することは難しい。従って、学習データ数が汎山取れない場合に汎化性を向上させるために何等かの工夫が必要となる。そのための手法として、重みやデータ空間に何等らかの拘束条件をつけることが考えられる。ここでは、前者の手法として重みの絶対値制限法、後者の手法としてK近傍内挿学習法[9]を取り上げ、その効果を検証した。

4.1 重みの変化幅に制限を設けた場合[8]

あまり数の多くないデータを用いて学習させた場合、学習データだけに適合したネットができる可能性が高い。これは、重み係数の値を自由に取ることを許したためである。従って、重み係数の大きさに制限を加えることで、特定なデータだけに適したネットの構築を防げる可能性がある。

実験内容：重みの絶対値がある値 W_0 を越えて変化する場合、その重み変化を行わない制限を設けることとする。

なお、重み係数の制限値 W_0 は、0.2, 0.5, 0.8の3段階について行った。

実験結果と考察：重み係数に制限を加えた方が学習データと未学習データの誤差の差を押えることができた。しかし、学習データについても誤差が大きくなっている。従って、重み係数に対する拘束条件の加え方に何等かの工夫を加える必要がある。

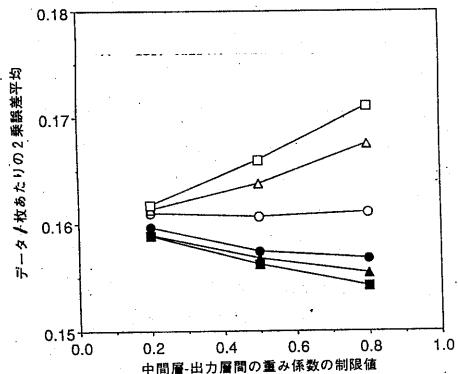
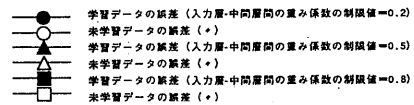


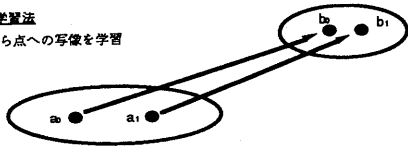
図7：学習データと未学習データの誤差(重み制限)

4.2 K近傍内挿学習法(KNN法)を用いた場合

従来のBP学習法は、点から点への写像を学習する。すなわち、各サンプルデータ間の途中の点における写像がどうなるか全く保証がない。この問題を改善するために、各サンプル点とK個の近傍点を結ぶK本の線分について近傍内挿学習を行うK近傍内挿学習法が提案されている[9]。この学習則は、網構造から網構造への写像を学習させる。このため過学習がおこりにくい。

従来のNN学習法

点から点への写像を学習



K近傍内挿学習

網構造から網構造への写像を学習

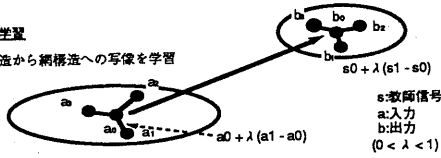


図8：K近傍内挿学習法の概念

また、実際にKNI T法を用いることによる従来の一般化δルールとの相違点は、

$$\delta = -\frac{\partial E}{\partial o_j} o_j (1 - o_j)$$

において

$$\frac{\partial E}{\partial o_j} = -(s_0 - b_0) + (s'_1 \frac{\partial f}{\partial a} (a_0) a'_1)$$

と第2項が加わるだけである。

なお、近傍選択距離尺度は、

$$D = (1 - w) \cdot d_{in} + w \cdot d_{out}$$

で定義され、あるサンプル点の近傍としてDが小さいものからK個だけ選択される。しかし、学習データが密にない場合、この選択方法を用いるとかなり異なるデータを近傍とする可能性がある。そこで今回では、近傍を選ぶ尺度を近いものから選ぶといった相対的基準ではなく、絶対的尺度（ある絶対距離以下のデータを近傍とするといったもの）を用いて行った。

実験内容：近傍選択距離尺度Dがしきい値D₀以内にある点のみを近傍として選択する。なお、D₀は、0.05, 0.10, 0.15, 0.20, 0.30に、入力距離と出力距離の比重を0.2, 0.5, 0.8に変えて行った。

実験結果と考察：今回のケースでは、

(近傍選択尺度) = 0.15、

(入力空間と出力空間の比重) = 0.5

の場合が学習データの誤差と未学習データの差が最小になり最も良い結果を示した(図9、10)。これ以外の値の時ではあまり有効な結果が得られなかった。これは、学習データ数が少ない場合、近傍として選ぶ基準を適切に決めないと、本来、近傍とは成り得ない点まで近傍にしてしまう危険があることを示している。従って、今回のようなケースでは、

各サンプル点ごとに近いもの順に個数分の近傍点を選ぶといった相対的な基準ではなく、絶対的な基準を設けてそれを満たすものを近傍点として選んだ方が有効であることを確認した。

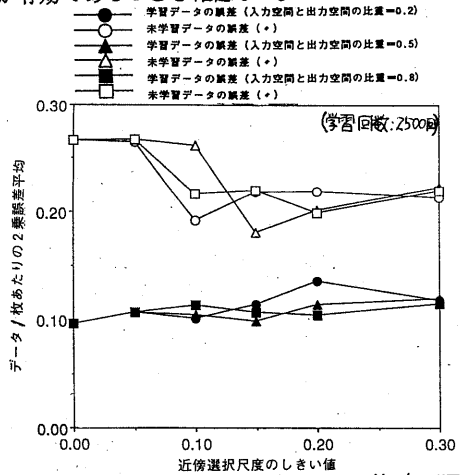


図9：学習データと未学習データの誤差 (KNIT法)

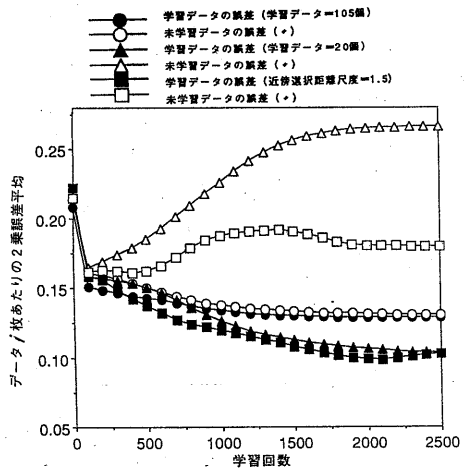


図10：学習回数による誤差変化

5. むすび

3層のニューラルネットにより目の形態分類を行った。そして、主観評価量のようにばらつきのある教師信号に対しても、重回帰分析を用いた場合と同様な分類を行うことができた。また、学習データが少ない場合でもK近傍内挿学習法を用いることで汎化能力が向上することを確認した。

謝辞

日頃、御指導頂く小林視覚情報研究部長に深く感謝致します。また、貴重な御意見を頂いた基礎研究所河原主幹員、そして視覚情報研究部の皆様に深く感謝致します。

参考文献

- [1] 萩原、増田：“パターンマッチングを主体にした顔画像による個人ID”、信学技報、PRU 8846(1988)
- [2] 磯、大塚、小杉：“目の特徴表現に関する心理要因の検討”、1990年信学春季全大D-543(1990)
- [3] 南：“顔の識別技術”、計測と制御、Vol 25, No.8, pp.707-713
- [4] 柳井、高木：“多変量解析ハンドブック”、現代数学社(1988)
- [5] 小笠原：“人相・手相・姓名判断”、新星出版(1984)
- [6] 磯、崎田、小杉：“ニューラルネットを用いた目の形状分類に関する検討”、1990年信学秋季全大D-406(1990)
- [7] Rumelhart, et al.：“Parallel Distributed Processing Volume 1,2”、The MIT Press(1986)
- [8] 立石、山崎：“手書数字認識における階層型ニューラルネットワークの中間層に関する考察”、情報学論、Vol.30, No.10, pp.1281-1288(1989)
- [9] 川端：“k-近傍内挿学習による音韻認識”、音響春季全大、2-P-21(1990)