

# 砂時計型ニューラルネットワークを用いた特徴空間の構築

Formation of the Feature Space Using the Sandgrass-type Neural Network

小山 達矢<sup>†</sup> 大枝 真一<sup>‡</sup>  
Tatsuya Koyama<sup>†</sup> Shinichi Oeda<sup>‡</sup>

木更津工業高等専門学校 制御・情報システム工学専攻<sup>†</sup> 木更津工業高等専門学校 情報工学科<sup>‡</sup>  
Advanced DJ Engineering Course, Kisarazu National College of Technology<sup>†</sup>  
Department of Information Engineering, Kisarazu National College of Technology<sup>‡</sup>

## 1. まえがき

近年、情報機器の高度化に伴い、音声認識や画像認識をはじめとするパターン識別技術の工学的応用が期待されている。パターン識別を行うモデルとしては様々なものが研究されているが、その代表的なもののひとつにニューラルネットワーク(以下、NN)がある。

NNの応用例として、情報圧縮、特徴抽出を行うための砂時計型NNが考案されている<sup>1)</sup>。しかしながら、砂時計型NNの非線形写像によって形成される特徴空間はユーザにとって扱いづらい側面があり、またその正当性についてはほとんど言及されてこなかった。本研究では特徴空間を調べ、より利便性の高い特徴空間を形成するために、従来とは異なる砂時計型NNの学習法を提案する。具体的には、バックプロパゲーション(以下、BP)とネットワークインバージョン(Network Inversion; 以下、NI)を組み合わせた、学習中に特徴空間の方向性を指示する手法を提案する。

## 2. 砂時計型ニューラルネットワーク

砂時計型 NN は 5 層からなる階層型 NN であり、入力層、出力層と比較して中間ニューロンの数が少ないという特徴を有する。このネットワークに入力を与え、それを復元出力する恒等写像を学習させることによって、中間層に情報が圧縮され、そこから入力信号の特徴表現を抽出することがで

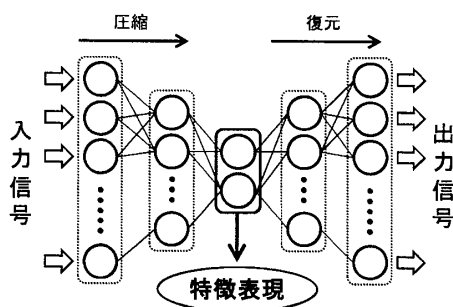


図1 砂時計型ニューラルネットワーク

きる。図1に、砂時計型 NN のモデル図を示す。

特徴空間は高次元のデータを低次元に写像するため、入力データが持っていたデータ間の距離関係は、特徴空間上では少なからず失われることになる。例えば、図2に示すような立方体状の3次元の座標群を砂時計型 NN に学習させると、2次元の特徴空間は学習初期の段階では図3のように、学習完了時には図4のようになる。図4において、立方体の頂点5から見た頂点6と頂点8は変化量が対等であるにも関わらず、特徴空間上での距離関係は大きく異なっている。この結果は、特徴空間を用いて未知データ分類を行うことの危険性を示唆する。特徴空間の利便性を高めるためには、非線形写像による“歪み”をできるだけ均一に広げるような形成手法が求められる。

## 3. 提案手法

従来の手法では、砂時計型 NN の特徴空間は恒等写像の BP 学習によって付随的に決定されるだけであった。そこで、特徴空間(中間ニューロンの出力)の方向性を指示しながら学習を行うアプローチが考えられる。そのために、以下で述べる NI を用いる。

### 3.1 ネットワークインバージョン

結合荷重を修正するBP法に対し、出力の誤差の原因は入力にあると考え、入力信号を修正する手法がNIである<sup>2)</sup>。NIでは、入力信号の修正量 $\Delta x_i$

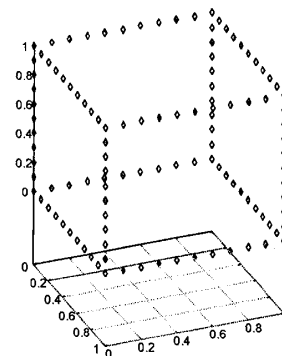


図2 入力に与えた立方体状の座標

は BP 法と同様に計算される.

$$\Delta x_i = -\eta \frac{\partial E_q}{\partial x_i} \quad (1)$$

ここで,  $\eta$  は学習係数,  $E_q$  は二乗誤差

$$E_q = \frac{1}{2} \sum (y_k - o_k)^2 \quad (2)$$

である. ただし  $y_k$  は対象出力 (教師信号),  $o_k$  は実際のネットワーク出力である.

### 3.2 DoubleBP 法の提案

中間出力を入力信号と捉えて NI を行うことで, 特徴空間の形成に目指すべき方向性を指示することができると考えられる. 中間出力の修正は, NI による修正後の中間出力を教師信号とし, 中間層以下の層に対して BP 学習を行うことで実現する. 恒等写像学習の BP と中間出力修正の BP を交互に行うことで, 効率の良い特徴空間の形成が期待できる. この手法を, 便宜上 DoubleBP と呼ぶ. DoubleBP の概念図を図 5 に示す.

### 4. 計算機実験

DoubleBP を用いて立方体状の座標群を学習させ, 特徴空間を形成する実験を行った. 恒等写像学習と中間出力修正の BP を各 1 回と考え, 10000 回の BP 学習を行う. 学習完了時に形成された特徴空間を図 6 に示す. これを見ると, 頂点 1 と頂点 7 の組以外は, ユークリッド的な距離関係を 2 次元上にうまく実現していることがわかる. 見方によれば, 特徴空間の歪みを一箇所に集中させた形と捉えることもできる. 図 4 と比較して, 従来の手法とは明確に傾向の異なる特徴空間を得ることに成功したと言える.

### 5. まとめ

本研究では, 学習中に特徴空間の方向性を指示することにより, データ間の距離関係を多く反映した特徴空間を得ることに成功した. 今後は実験を重ねることで, 得られる特徴空間の意味とその利便性を見出していく予定である. また, 提案手法を実際の未知データ分類に適用し, 分類の成功率を調べることで特徴空間の定量的な評価を行っていききたい.

### 参考文献

- 1) Malthouse, E.C.: Limitations of nonlinear PCA as performed with generic neural networks, pp.165-173, IEEE Trans. On Neural Networks, 9, 1998.
- 2) 小杉幸夫: 神経回路システム, pp.72-81, コロナ社, 1995.

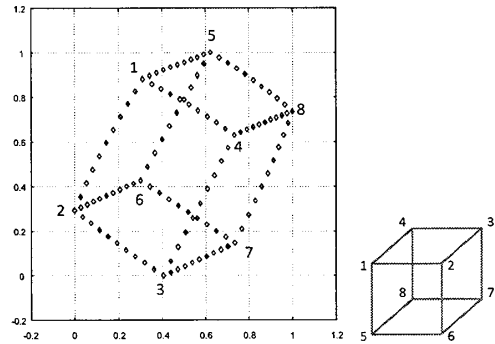


図 3 BP 学習 1000 回時の特徴空間

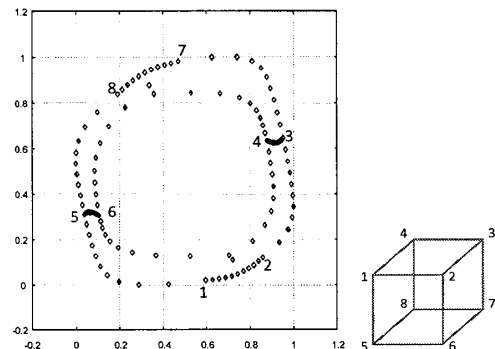


図 4 BP 学習 10000 回時の特徴空間

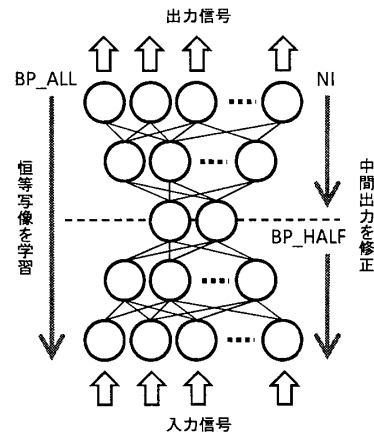


図 5 DoubleBP の概念図

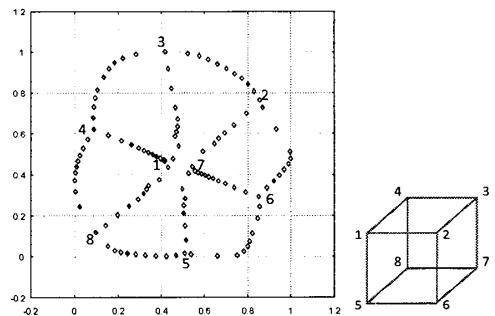


図 6 DoubleBP により形成された特徴空間