

砂時計型ニューラルネットワークを用いた センサ情報融合による内部表現の自動獲得

Learning Internal Representation for Multi Sensor Fusion using Autoencoder Neural Networks

柳沼義典^{*}
Yoshinori YAGINUMA

木本 隆^{*}
Takashi KIMOTO

山川 宏^{**}
Hiroshi YAMAKAWA

^{*}富士通株式会社
FUJITSU LIMITED

^{**} RWCPつくば研究センタ
Real World Computing Partnership

Abstract -- For applying a system to real-world, it is important that the system can integrate from many kinds of information and obtain the essence with the system's purpose from that information. In this paper, at three-dimensional object recognition, we propose an autoencoder neural network model obtaining the essence for recognition with multi sensor fusion. We also verified the effectiveness of the proposed model by simulation exercises of a doll recognition problem, that uses camera image segmenting by that's hue and saturation.

1. はじめに

人間は、複数種類の感覚器官から得られる多種多様かつ膨大な情報から、問題に応じて必要な情報を選択適用することで、現実世界に適応的に生活している。そして、以前に行った事例を学習により内部に蓄え効率的に参照することで、膨大な情報の実時間処理を可能としている。

人間のような情報処理機構を工学的に実現しようとした場合、複数センサから得られる多種多様で曖昧かつ不完全な情報を統合し、問題に応じて必要とされる情報を自動的に抽出し、システム内部にコンパクトな表現として蓄えるしくみが重要となる。本報告では、情報圧縮の特長を持つ砂時計型ニューラルネットワークに着目し、これを用いたセンサ情報融合モデルを提案する。

以前より砂時計型ニューラルネットワークは、画像圧縮の手法として研究されてきた[1]。最近では、モデル獲得や物体認識問題において、画像情報から人の性別・表情などの識別に関する研究[2,3]や、Jordanらのモジュール構造型ネットワークと組合せた3次元物体のモデル表現に適用する研究[4]などがある。また、よりコンパクトな内部表現の自動獲得の研究[5]なども報告されている。さらに、砂時計型ニューラルネットワークを階層的に積み上げることで高次の概念を抽出する研究[6]や入力情報に欠損値がある場合の学習および欠損値推定方法の研究[7]なども報告されている。

本報告では、砂時計型ニューラルネットワークを用いて複数センサ情報を融合し、認識に必要な内部表現を自動獲得するセンサ情報融合モデルを提案し、その有効性を3次元物体認識実験により示す。

まず砂時計型ニューラルネットワークの性質を説明し、さまざまな視点から視覚センサにより観測した3次元物体の射影像を学習する計算機シミュレーション実験

を行い、獲得された内部表現を解析する。

つぎに、砂時計型ニューラルネットワークを用いたセンサ情報融合モデルを提案し、人形を対象物体として、視覚センサからHSI値により切り出すことで生成した複数の色画像情報を融合することで対象物体を認識する認識実験を行い、提案したモデルの有効性を示すとともに、物体認識のための適切な内部表現が自動的にニューラルネットワーク内に獲得されていることを示す。

2. 砂時計型ニューラルネットワークの内部表現

2-1 砂時計型ニューラルネットワーク

砂時計型ニューラルネットワークとは、入力層と出力層に同じデータを提示して恒等写像を学習させる3層以上の階層型のニューラルネットワークである。この際に、中間層ユニット数を入出力層よりも少なくするのでその形状から砂時計型と呼ばれる(図1)。

砂時計型ニューラルネットワークでは入力層から中間層までを逆モデル、中間層から出力層までを順モデルと見ることができる。つまり、入力層から中間層までで入力情報から本質的に重要な情報を抽出し、中間層から出力層までで抽出された情報から入力情報を復元する。一般に逆モデルの生成は不良設定問題(入力に対して出力が一意に決定できない問題)であり教師データの設定が難しいのだが、砂時計型ニューラルネットワークでは、恒等写像を学習して内部に逆モデルと順モデルを生成することで、不良設定問題を意識することなく逆問題を解くことが可能である。

砂時計型ニューラルネットワークの中間層に現れる内部表現とは、提示したデータのうちの変動部分である。変動部分が、内部表現としてコンパクトで連続的な空間にマッピングされる。一方、固定部分はその砂時計型ニューラルネットワーク内での順モデル部分の重みとし

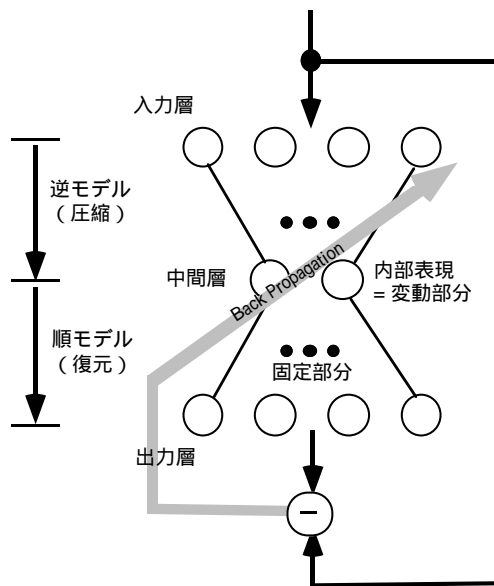


図1 砂時計型ニューラルネットワーク

て作成される。すなわち、砂時計型ニューラルネットワークは変動部分と固定部分を分離し、変動部分の抽出を行うことが可能である。

砂時計型ニューラルネットワークの性質をまとめると以下ようになる。

(1). 情報がコンパクトに表現

砂時計型ニューラルネットワークでは中間層が絞り込まれているために、その中間層部分に入力情報のうちの必要な情報がコンパクトに表現される。ここではこれを、砂時計型ニューラルネットワークの持つ内部表現と呼ぶ。

(2). 内部表現は連続空間上で表現

連続関数であるニューラルネットワークは、補間機能という特長を持つ。これは、サンプル点としていくつかの情報を学習させると、そのサンプル点間を連続的に補う機能である。この補間機能により内部表現は連続空間上で表現される。

(3). 教師データが不要

砂時計型ニューラルネットワークは恒等写像を学習するために、ニューラルネットワークの学習に必要な教師データが不要であるというメリットがある。例えば物体認識問題に適用する場合には、対象物体を提示するだけでその物体を表わす情報が砂時計型ニューラルネットワークの内部表現として抽出されることが期待される。

2-2. 内部表現獲得シミュレーション

砂時計型ニューラルネットワークを用いて、物体を提示した時に作成される内部表現について計算機シミュレーションを行なった。対象物体として5角

形を用い、それを中心とした半球状に沿って視点を変更させた時の対象物体の見えをデータとして与えた。 $\theta = 0^\circ \dots 360^\circ$ まで 30° おき、 $\phi = 0^\circ \dots 60^\circ$ まで 30° おきによって、合計で36データを学習データとして用いた(図2)。入力は五角形の頂点の座標値(x, y)で合計10入力である。用いた砂時計型ニューラルネットワークは5層の階層型ニューラルネットワークで、各層のユニット数は、それぞれ(10-10-2-10-10)である。

100,000回の学習後、 θ と ϕ とも 5° おきにデータを与えた時の中間層2個の出力結果を図3に示す。図3より、 θ の2次元が中間層上で連続した空間として綺麗に表現されていること、また、未学習の点が補間されていることがわかる。このように、砂時計型ニューラルネットワークを用いることで、与えられた情報のうちの変動部分である「視点の変更」が、絞り込まれた中間層上にコンパクトで連続な表現として抽出されることが確認された。

3. 内部表現を自動獲得するセンサ情報融合モデル

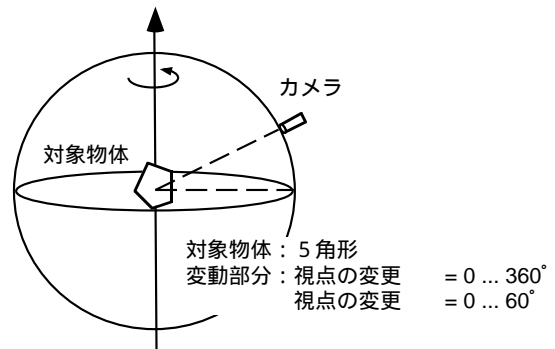


図2 シミュレーション条件

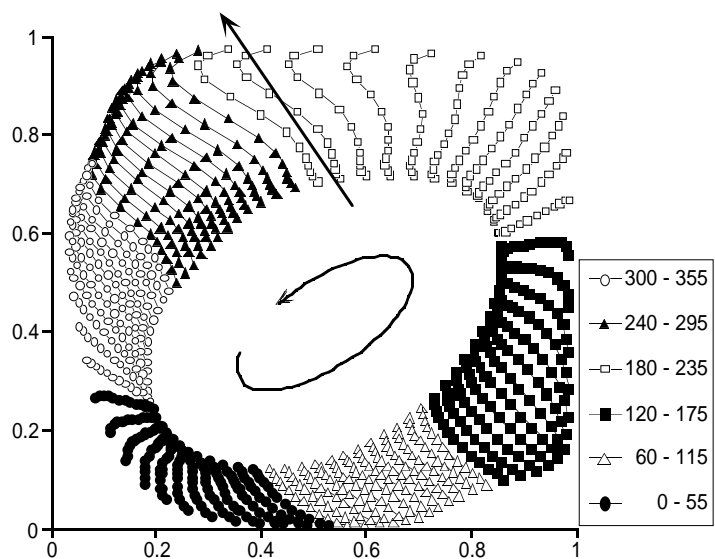


図3 抽出された内部表現

複数センサから得られる情報を融合して、内部表現を自動的に獲得するセンサ情報融合モデルを提案する(図4)。センサ1、センサ2...など、複数センサからの情報を同一の砂時計型ニューラルネットワークに入力して学習させることで、それらセンサ情報の融合結果が砂時計型ニューラルネットワークの中間層に内部表現として獲得される。

現実環境においては、その時々々の環境やシステム内部状態・目的に応じて、有用な情報を提供できるセンサと不完全で曖昧な情報しか得られないセンサが頻りに変わりうる。このような環境では、複数センサ情報入力を融合するセンサ情報融合モデルは有効である。つまり、状況に応じて有用な情報を提供できるセンサが異なるような、センサ情報の不連続性から連続的な内部表現を生成できると期待されるからである。センサ情報融合モデルを用いることにより、現実環境に適用可能なロバストなシステム構築が可能となる。

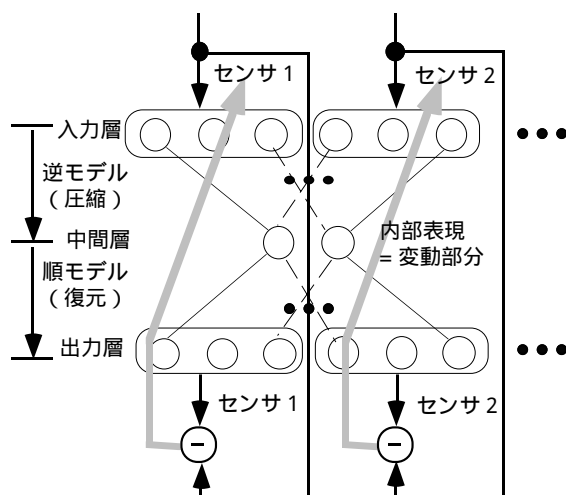


図4 センサ情報融合モデル

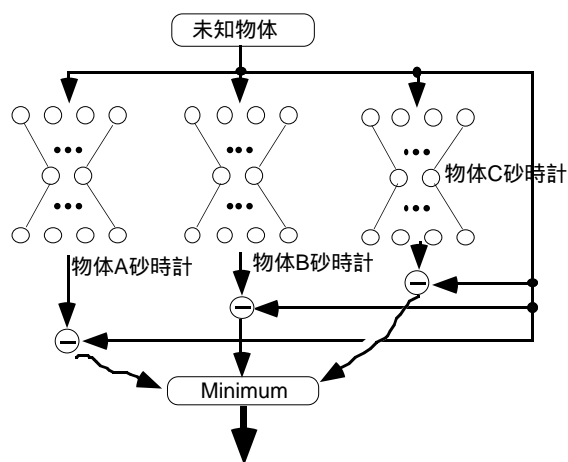


図5 物体認識問題への適用

4. 物体認識実験

4-1. センサ情報融合モデルの物体認識への適用

砂時計型ニューラルネットワークを用いたセンサ情報融合モデルを、物体認識問題に適用する方法を図5に示す。これは、一つの砂時計型ニューラルネットワークに一つの物体を学習させ、学習後、提示された物体を各砂時計型ニューラルネットワークに入力し、その誤差が最も小さな砂時計型ニューラルネットワークが示す物体を該当物体とみなす方法である。この場合、各物体のサンプル点を学習した砂時計型ニューラルネットワークが、十分に各物体を汎化しているかどうか重要となる。

4-2. 実画像による物体認識

対象物体に人形を用い、視覚センサから色切り出しにより生成した複数画像情報を砂時計型ニューラルネットワークを用いて融合することで、物体認識のための内部表現を自動獲得する実験を行った。

図6に対象物体(上)と比較物体(下)、図7に実験システムを示す。ここでは、変動分としては1次元の視点の変更(回転)を用いた。すなわちターンテーブルに乗せた対象物体を、0°から360°まで30°ずつ回転させてカメラにて撮影し、その実画像を、それが持つHSI値の固まりで切出した後(図8)、大きなセルに分割して各セルごとに各HSI値固まりの割合を入力として砂時計モデルに学習させた(表1)。

用いた砂時計型ニューラルネットワークは、各層のユニット数がそれぞれ(36-16-2-16-36)の5層の階層型ニューラルネットワークである。100,000回の学習後、中間層に生成された内部表現を図9に示す。図9より、1次元の回転を表す円が内部表現として獲得されている。また、同物体未学習データ(30°おきの学習データの間の点)を学習後の砂時計モデルに入力したところ、内部表現値はちょうど学習データの間に収まっており、自動獲得された内部表現上で補間されていることが確認される。さらに、汎化性能を調べるために、対象物体(上)学習済みネットワークに比較物体(下)のデータを入力してその2乗誤差を調べた結果が図10である。図10より、同物体未学習データに対して、未知物体データの2乗誤差値は平均で3倍程度違い、学習した物体かそうでないかを砂時計型ニューラルネットワークの2乗誤差値から識別することが可能であることが確認された。

5. まとめ

表1 実験条件

実画像	320 x 240 pixel
セル	4 x 3 個 (80pixelの正方形)
HSI値の固まり	3種類 (よって入力36個)
学習データ	30° おき (12パターン)

砂時計型ニューラルネットワークの性質を検証し、3次元物体の射影像を学習する計算機シミュレーション実験を行い、砂時計型ニューラルネットワーク内部に提示された情報の変動部分が自動的にコンパクトに獲得されることを示した。そして、この砂時計型ニューラルネットワークを用いたセンサ情報融合モデルを提案し、複数色画像情報を融合する物体認識実験を行い、物体認識のための適切な内部表現が自動的にニューラルネットワーク内部に獲得されていることを示し、提案したモデルの有効性を確認した。

今後は、行動を予測するニューラルネットワークと階層的に組み合わせることで、能動的にセンシングしながら物体を認識する能動的物体認識モデルへと発展させていきたいと考えている。

謝辞

なお本研究の一部は、科学技術庁の科学技術振興調整費による「センサフュージョンの基盤的技術の開発に関する研究」の一環として行なったものである。

6. 参考文献

1. Cottrell, G. W., Muno, P., and Zipser, D., "Image Compression by Back Propagation. An Example of Extensional Programmins", ICS Report 8702, Institute for Cognitive Science, UCSD, 1987.
2. Cottrell, G. W. and Metcalfe, J., "EMPATH: Face, Emotion, and Gender Recognition Using Holons", NIPS3, pp. 564-572, 1991.
3. DeMers, D. and Cottrell, G. W., "Non-Linear Dimensionality Reduction", NIPS5, pp. 580-587, 1993.
4. 鈴木、安藤、"モジュール学習による3次元物体の認識と類別", 信学技報, NC93-62, pp. 59-66, 1993.
5. 野田、"過負荷学習を用いた恒等写像学習による内部表現獲得", 日本神経回路学会第5回全国大会講演論文集, pp. 80-81, 1994.
6. 山川、木本、柳沼、"階層化された砂時計型ニューラルネットによる自律的な内部表現獲得", 第11回ファジィシステムシンポジウム講演予定, 1995
7. 木本、柳沼、山川、"不完全なデータに対応する砂時計型ネットワーク", 信学技報, NC94-62, pp.17-24, 1995.

連絡先

柳沼義典、木本隆
 〒211 神奈川県川崎市中原区上小田中1015
 株式会社 富士通研究所マルチメディアシステム研究所
 テクノロジ研究部門知識処理研究部
 TEL:044-754-2667 FAX:044-754-2664
 E-mail: yaginuma@flab.fujitsu.co.jp
 kimoto@flab.fujitsu.co.jp

山川宏
 〒305 茨城県つくば市竹園1-6-1 つくば三井ビル16F
 RWC Pつくば研究センタ 能動知能研究室
 Tel 0298-53-1664 Fax 0298-53-1652
 E-mail: yamakawa@trc.rwcp.or.jp



図6 対象物体(上)と比較物体(下)

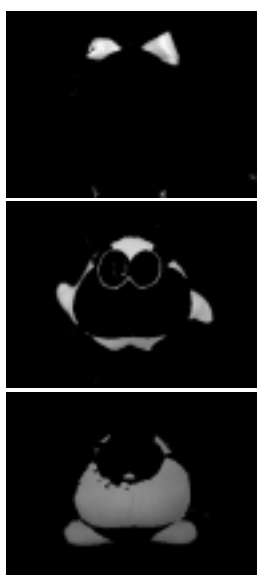


図8 色抽出画像

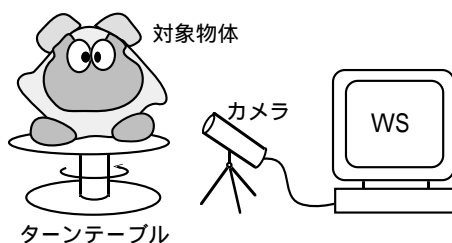


図7 実験システム

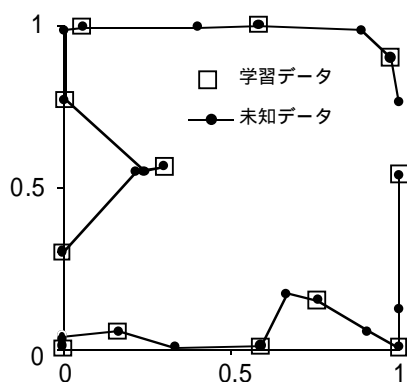


図9 学習後の砂時計ニューロの内部表現

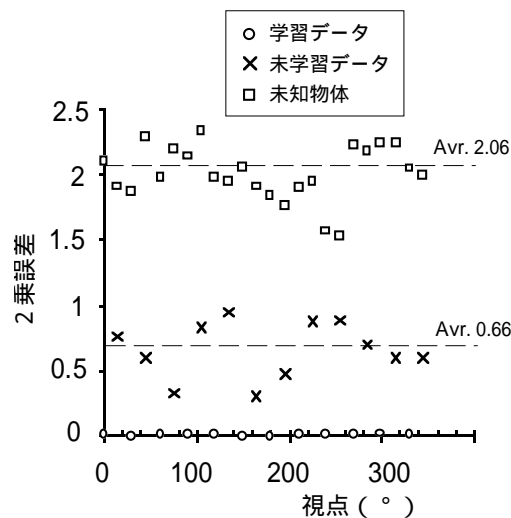


図10 学習後の砂時計ニューロの誤差値