

## 研究課題別評価

1. 研究課題名： 学習によるシーン理解の研究

2. 研究者氏名： 長谷川修

3. 研究の狙い：

近年、視覚、聴覚、発話、表情やジェスチャの表出、といった機能を兼ね備えた、マルチモーダル対話インタフェースの研究開発が盛んである。こうしたインタフェースはヒューマノイドロボットなどに実装され、人と対話しながら実環境で人と豊かに共存する人工システムの実現が目指されている。

しかし、これまでに実現された対話インタフェースを知能情報処理の観点から眺めると、対話機能は概ねシナリオの決まった限定的なものであり、予め想定された(組み込まれた)範囲内では有効に機能するが、その範囲を超えると破綻してしまう。すなわち人間のように環境や他者との相互作用を通じて新奇な経験を蓄積・学習し、徐々に環境に適応したり、成長・発達して新たな言語能力や身体運動能力を身につけるといったことは実現されていない。

このように従来の対話インタフェースの機能が限定的である主要な要因のひとつに、従来のパターン認識技術の枠組みが「固く」、柔軟性に欠ける点がある。一般に既存のパターン認識の枠組みでは、まず学習用のパターンが用意され、「学習過程」としてそれらのパターンを精度良く識別できるよう学習器のパラメータが調整(学習)される。次いでパラメータの調整が完了すると、学習器は「認識過程」に移行し、学習用パターンに類似するパターンが識別・認識されるようになる。すなわち従来のパターン認識の枠組みには「学習過程」と「認識過程」の区別があり、これらは完全に分離していて学習過程が終了するまで認識を行うことはできない。このため通常学習後の学習器に新奇なパターンを追加的に学習させようとする、既に学習したパターンも含めて全ての学習過程をやり直さなければならない。また学習用のパターンにノイズが混入していると、学習そのものがうまく行かないこともある。

こうした問題は、システムを実環境に置いたときに必要となる、システムが実環境で出会うノイズで膨大なパターン情報(視覚や聴覚など)から本質的な情報を取り出し、それらを経験や知識として追加的に蓄積・学習してゆくといった機能の実現を考える上で大きな障害となる。そしてそれ故に、従来のパターン認識技術を用いた対話インタフェースは、設計時に想定された範囲内においては比較的良く機能するが、想定された範囲を超えると破綻してしまうこととなる。本研究ではこの問題の本質的な解決を図る。

4. 研究成果：

本研究では上記の問題の解決のため、まず独自の「自己増殖型ニューラルネットワーク」を提

案するとともに、これをコア技術として利用する「持続的発達学習メカニズム」を提案・実現した。さらには、このメカニズムを用いた持続的学習・成長・発達型のマルチモーダル対話インタフェースをヒューマノイドロボット上に実装した。

本研究で提案する自己増殖型ニューラルネットワークは、“Growing Neural Gas” と呼ばれる競合学習の一手法に独自の改良を加えたもので、

1. 刺激が加わると刺激量に応じて適応的にニューロンが増減し刺激の分布を近似
2. 刺激に混入するノイズを無視し刺激の本質的な(密度の高い)部分のみを追加的に抽出可能
3. 刺激の本質的な(密度の高い)部分が何個あったか(刺激に含まれるクラスタ数)を出力可能といった世界的に見ても他に類を見ない顕著な特長を持つ。

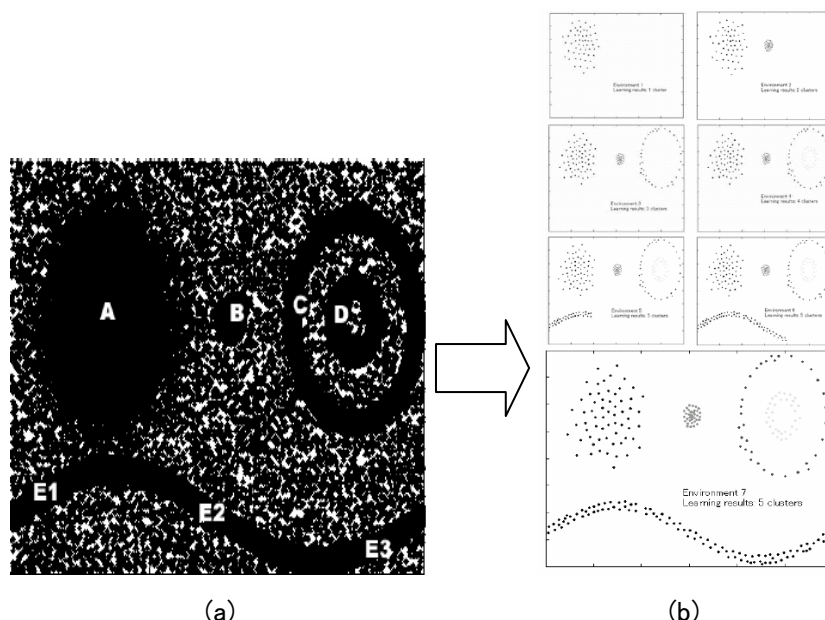


図1:(a)システムへの入力データ(10%のノイズを含んでいる), (b)処理結果(ノイズを無視し、(a)に含まれる本質的な領域のみを追加的に抽出している。)

図1に、筆者らの自己増殖型ニューラルネットワークを、2次元のノイズが混入したデータに適用した結果を示す。提案するネットワークは高次元のデータに対しても機能するが、ここでは図示を可能とするため、2次元のデータを用いた。図1(a)は非線形な入力データの全容であり、ランダムに10%のノイズを混入してある。このデータを図中のA,B,C,D,E1,E2,E3の順に、追加的にニューラルネットに与えた。図1(b)は処理結果である。図1(b)の左上より右下に向けて、A, B, C, D, E1, E2, E3のデータが順に与えられたときのニューロンの増殖の様子を示している。図示されるように、ランダムノイズが混入しながら追加的に与えられたデータに対し、ノイズを除去しながらデータの本質的な(密度の高い)部分だけにニューロンが反応(増殖)し、それらの部分を抽出している。

ここで、抽出された領域はクラスタリングされており、処理の最終段階((b)の最下段)において

は、システムは入力データには5つのクラスター(密度の高い部分)があったと出力している。この手法は、実データを用いたいくつかの実験、具体的には複数人の顔画像の追加学習やベクトル量子化の問題などにも適用し、有効性を確認している。こうした高度に柔軟な追加学習を可能とした手法は、世界的に見ても他に例がない。この成果は、ニューラルネットワーク研究分野における代表的な国際学術誌 Neural Networks 誌に採録された。

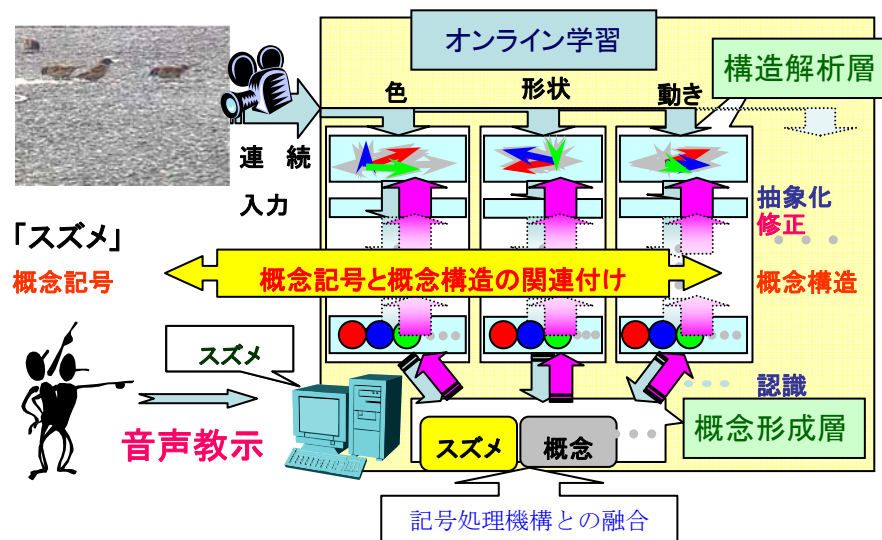


図2: 持続的発達学習メカニズムの概要(「スズメ」の概念の形成例)

図2に、上記の自己増殖型ニューラルネットワークをコア技術として利用する「持続的発達学習メカニズム」の概要を示す。ここでは実環境から視覚(色、形状、動き)のパターン情報を恒常的に入力するとともに、自己増殖型ニューラルネットワークを階層的に用いて自己組織的に入力パターン情報の集約的概念シンボル化を進める(構造解析層)。獲得した概念シンボルは、人間との相互作用(準教師あり学習による教示)に基づき言語情報(音声パターンをシンボル化した情報)と結合する(概念形成層)。ここで獲得された概念シンボルに誤りがあると人間からの教示をトップダウン情報として構造解析層を修正する。以上の過程は全て並列に稼働し、獲得した色や形、物の名前、動きなどの概念は直ちに認識に利用できる。すなわちこの学習メカニズムには従来のパターン認識手法に見られる「学習過程」「認識過程」の区別がなく、新奇な情報を恒常的・追加的に学習しながら同時に認識も行うことができる。また入力パターンの表現層とネットワークで繋がった概念シンボルに言語による記号情報を付与することで、人工知能における本質的課題である「記号接地問題」の解決を図ることができる。さらにはパターンから自己組織的に獲得した概念シンボルをシンボルベースの推論機構等と結合することで、パターン処理と記号処理の融合を図ることも可能である。

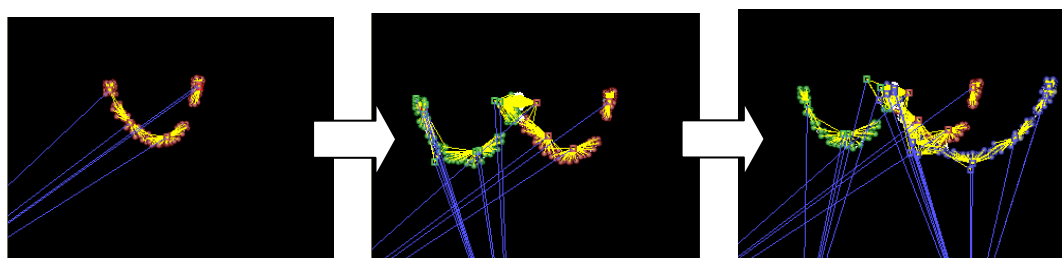
本研究の関連研究には、Growing Neural Gas (GNG) や Growing Neural Gas with Utility (GNG-U) などの研究があるが、これらはいずれも増殖するノード数に制約(上限)があるほか、ノ

イズに対する耐性がない。また入力データに含まれるクラスタの数を出力することもできない。一般的なクラスタリング手法としてはk-means法やEM法があるが、これらは算出するクラスタの数を予め決めておかねばならない。教師無しのクラスタリング手法としてはSOM(自己組織化マップ)があるが、SOMは事前にマップのサイズを決めておく必要がある。提案手法ではそうした必要がない。

知能情報処理・マルチモーダル対話インタフェースの観点からは、これまでに、ノイズーな実環境中で環境や人間と相互作用しながらオンラインで持続的・追加的に知識や身体運動機能を獲得し、成長・発達するシステムを実現した事例は世界的に見ても存在しない。

以下では、筆者らの自己増殖型NNを時系列データの学習に用いた事例を示す。まず自己増殖型NNを接続したカメラに対して2つの物体を見せ、それらの間で「近づける」「遠ざける」「またぐ」の動きを複数回ずつ順に提示した。自己増殖型NNには2つの物体間の「動き」の軌跡が時系列に入力されるようにした。その結果、入力された動きの軌跡に沿ってニューロンが増殖し、「近づける」「遠ざける」「またぐ」の代表的(平均的)な軌跡が自己組織的・追加的に学習された。図3に学習した3つの動きの軌跡を示す。

ここで強調したい点は、自己増殖型NNでは入力される時系列データの長さに応じて適応的に獲得される軌跡の長さが調節されている点である。具体的には、図3(c)で最後に獲得された「またぐ」は軌道の長さが「近づける」よりも長かったため、自己増殖型NNはより多くの状態数で軌跡を表現した。こうした時系列データの長さに応じた適応的な状態数の決定は、一般に隠れマルコフモデル(HMM)では困難である。今後は、上記の自己増殖型NNの機能をベースに、追加学習が可能で、かつ状態数を自律的・適応的に調節・決定することのできる新たな時系列データの学習・認識メカニズムを確立し、その優位性をHMMと比較して示すことが目標である。



(a) 「遠ざける」を学習 (b) 「近づける」を追加学習 (c) 「またぐ」を追加学習

図3: ANGを用いた時系列データ(動き)の追加学習に関する予備実験結果

図4に、提案する「持続的発達学習メカニズム」をヒューノイドロボットに実装した例を示す。ここでは図3に示した「動き」の概念の獲得機能を利用し、その対話的学習・認識・生成を実現した。具体的には、まずロボットの前に2つの物体を置き、「近づける」と発話しながら双方を近づける動きを複数回見せる。するとロボットは、この教示データから「近づける」という言葉の「意味」に対応した動き(軌跡)を学習した。学習後は、ロボットは「近づける」と命令されると様々な位置関係に置

かれた2つの物体を近づける動きを生成することができた。こうした一連の学習・認識・生成は、「遠ざける」「またぐ」の動きに対しても同様に(追加的に)行うことができた。こうした、提示と発話による教示(やって見せつつ言って聞かせる)を通じた学習・認識・生成の研究は、今後さらに重要度が増すと考えられる。

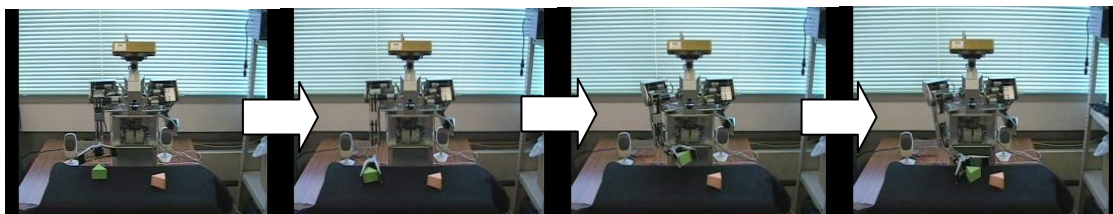


図4: 自己増殖型NNにより学習した「動き」の概念を用いたロボットの制御(近づける)

この他、72種類の物体を順にロボットに見せ(ここでは「動き」は含めていない)、その色・形・名称をオンラインで対話的に追加学習させることを可能としたシステムを実現し、その内容を電子情報通信学会のPRMU研究会で発表して大きな反響を得た。この論文の内容は、一部を拡充して英訳し、IEEE の論文誌(SMC)に投稿中である。

さらに本研究では、上記の研究を統合的に発展させ、視聴覚情報を相補的に用いた文法の獲得も行った。具体的には、例えばロボットに「赤色ーリンゴー黄色ーレモンー右ー置く」といった発話を聞かせると同時に、その様子を実演してみせる。ロボットは見せられた状況(視覚情報)を自らの内部(メンタルモデル)に取り込むとともに、言語で発話された内容と見せられた状況とを内部的に関連付けながら発話を解析し、そこに含まれる文法構造を徐々に獲得する。ただし赤色、リンゴ、黄色、レモン、置く、といった単語(名詞・形容詞・動詞)の意味は予め個別に学習(グラウンディング)させる。この研究のポイントは、従来研究に見られるように多くの文章のセットを予め人間が用意してバッチ的に学習させるのではなく、逐次的に与えられる少ない文章からでも視覚情報を援用して対話的に文法の獲得を可能とした点にある。なおここでは日本語で教える例を示したが、日本語に限らず、英語等の他言語からでも文法の獲得が可能である。

総じて本研究では、乳幼児が養育者や環境とのマルチモーダルな相互作用から単語のグラウンディング、文法の獲得、文章の理解・生成などへと発達するのと同じように、持続的・段階的に発達する機械学習メカニズム(持続的発達学習メカニズム)のプロトタイプを構築・実現した。

## 5. 自己評価:

本研究では、人工システムを実環境に置いたときに必要となる、システムが実環境で出会うノイズで膨大なパターン情報(視覚や聴覚など)から本質的な情報を取り出し、それらを経験や知識として追加的に蓄積・学習してゆくといった機能を、独自の自己増殖型ニューラルネットを提案・導入することにより実現した。これにより、乳幼児が養育者や環境とのマルチモーダルな相互作用から単語のグラウンディング、文法の獲得、文章の理解・生成などへと発達するのと同じように、

持続的・段階的に発達する機械学習メカニズムのプロトタイプを世界で初めて構築した。今後こうした研究の重要性は広く理解され、本研究の持つ意義や先駆性、独創性も認知されるようになると思う。

本研究に関連して発表した論文は、2005年12月現在、採録済みもしくは発表済みのものだけで、論文誌論文15件(うち国際10件、国内5件)、国際会議論文24件、国内会議論文20件になる。本研究では、学術的に積極的に発表することを目標の一つに掲げたが、その目標はある程度達成できたと思う。

#### 6. 研究総括の見解:

人の赤ちゃんが知識を持たない状態から徐々に知的に発達するように、視覚や聴覚を介した人との相互作用を通じ、視覚的概念や言語を複合的に学習して発達するロボットを世界で初めて構築したことは、大きな成果であり、高く評価できる。このロボットの学習機能には、独自に研究開発した「自己増殖型ニューラルネットワーク」を用いた「持続的発達学習メカニズム」を導入している。いずれも先駆性、独創性に優れた技術である。さらに実験の結果、対話による教示・育成を通じ、ロボットに人間の2歳児程度の知能を持たせることに成功したことも評価に値する。

今後、自己増殖型ニューラルネットワーク (Adaptive Neural Gas : ANG) の機能をさらに充実させるとともに、ANGが Self Organizing Map (SOM)、Adaptive Resonance Theory (ART)、Expectation Maximization (EM)法などと並び称される標準的な機械学習手法の一つとなるよう努力されたい。

#### 7. 主な論文等:

##### 論文

1. Shen Furoo and Osamu Hasegawa, "An Incremental Network for On-line Unsupervised Classification and Topology Learning", Neural Networks, (2005) in press
2. Shen Furoo and Osamu Hasegawa, "An Adaptive Incremental LBG for Vector Quantization", Neural Networks (2005) in press
3. 豊田 崇弘, 長谷川 修, "高次局所自己相関特徴の拡張", 画像電子学会誌, Vol.34, No.4, pp.390-397 (2005)
4. Shen Furoo, Osamu Hasegawa : "A fast no search fractal image coding method", Signal Processing: Image Communication, vol.19, pp.393-404, (2004)
5. 長谷川修、金出武雄: "一般道路映像中の移動物体の識別・色の推定と特定対象の検出", 情報処理学会論文誌、Vol.44, No.7, pp.1795-1807, (2003)

##### 国際会議発表論文

1. Shen Furoo and Osamu Hasegawa : "An On-line Learning Mechanism for Unsupervised Classification and Topology Representation", IEEE International Conference on Computer

- Vision and Pattern Recognition (CVPR05), San Diego, pp.651-656, (2005)
2. Aram Kawewong and Osamu Hasegawa, "3D Texture Classification by Using Pre-testing Stage and Reliability Table", IEEE Proc. International Conference on Image Processing (ICIP05), (2005) accepted
  3. Shen Furao, Osamu Hasegawa : "An Incremental k-means Clustering Algorithm with Adaptive Distance Measure", International Conference on Neural Information Processing, (ICONIP2005), (accepted)
  4. Takahiro Toyoda and Osamu Hasegawa : "Texture Classification Using Extended Higher Order Local Autocorrelation Features", The 4th international workshop on texture analysis and synthesis (Texture2005), pp.131-136 (2005)
  5. Shen Furao, Osamu Hasegawa : "A Self-organized Growing Network for On-line Unsupervised Learning", IEEE International Joint Conference on Neural Networks (IJCNN2004), CD-ROM ISBN 0-7803-8360-5, Vol.1, pp.11-16 (2004)

#### 国内会議発表論文

1. 小島量、長谷川修：“ヒューマノイドロボット上の自己増殖型ニューラルネットワークを用いた視聴覚情報からの能動的・追加的概念獲得”，信学技報 PRMU2005-57, pp.35-40, (2005-9)
2. Shogo Okada and Osamu Hasegawa : "Unsupervised Learning of Dynamic Patterns for Recognition and Generation", 計測自動制御学会 システムインテグレーション部門講演会, (2004)
3. Shen Furao and Osamu Hasegawa : "A self-controlled incremental method for vector quantization", 電子情報通信学会 PRMU 研究会, PRMU2004-65, pp.93-100 / 情報処理学会 CVIM 研究会, 2004-CVIM-145, 93-100 (2004)
4. 神谷祐樹、申富饒、長谷川修：“実世界パターン情報の非停止型学習・認識手法の提案”，第 18 回人工知能学会全国大会, CD-ROM 2H1-03, (2004)
5. 豊田崇弘、長谷川修：“局所自己相関マスクの拡張とその応用に関する一考察”，第 10 回画像センシングシンポジウム(SSII04), pp.307-310, (2004)

#### 解説・総説

長谷川修：シーン理解のためのパターン認識、O plus E, Vol.25, No.10, pp.1130-1136 (2003)

#### 招待講演

北陸先端科学技術大学院大学情報科学研究科セミナー(平成17年9月13日)

受賞

計測自動制御学会 システムインテグレーション部門講演会(SI2004) ベストセッション講演賞:対象論文 Shogo Okada and Osamu Hasegawa :“Unsupervised Learning of Dynamic Patterns for Recognition and Generation”