

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 4 月 11 日現在

機関番号：32644

研究種目：基盤研究（C）

研究期間：2009～2011

課題番号：21500221

研究課題名（和文） 情報エンハンスメント法とその応用

研究課題名（英文） Information Enhancement for Interpreting Internal Representation in Neural Networks

研究代表者

上村 龍太郎（KAMIMRUA RYOTARO）

東海大学・情報教育センター・教授

研究者番号：80176643

研究成果の概要（和文）：

本研究では、ニューラルネットワーク（特に競合学習）が学習によって獲得した内部表現を明確に解釈する新しい方法を提案した。この解釈とは、ニューラルネットワークのすべての構成要素・要素結合の意味と機能を理解すると考える。また自然神経回路の知見により内部表現を設定し学習モデルを作成する研究も視野に入れた。研究の結果、自己組織化マップへ応用した場合、明確な特徴を抽出することができることがわかった。さらにニューロンの社会化などの新しいモデルを作る研究とつながることもわかった。

研究成果の概要（英文）：

In this study, we proposed a new method called “information enhancement” for interpreting the mechanism of neural networks. We tried to interpret all the components and all the possible combinations of the components by this information enhancement. Applied to the self-organizing maps, we found that clear class structure could be produced for interpretation. In addition, we found that the method was related to a new model for neural networks taking into account the social activities of neurons.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,400,000	420,000	1,820,000
2010年度	1,000,000	300,000	1,300,000
2011年度	1,000,000	300,000	1,300,000
年度			
年度			
総計	3,400,000	1,020,000	4,420,000

研究分野：総合領域

科研費の分科・細目：感性情報学・ソフトコンピューティング

キーワード：ニューラルネットワーク

## 1. 研究開始当初の背景

ニューラルネットはその優れた学習能力ため多くの分野で応用されている。しかし最大の問題点は、ニューラルネットワークが獲得した内部表現を理解することが困難なことにある。

解釈の問題は複雑であり、以前のように活

発に研究がおこなわれているとは言えない。しかしながら、最近形を変えた形で、内部表現の研究が盛んになってきている。代表的な例として特徴選択の例をあげる。これは、ネットワークはどの変数を選択すればよいか決定するための方法であり、解釈の立場からするとどの変数はどのような意味を持って

いるかという問題に変換される。

さて、インターネットおよび各種情報技術の発展により現在取り扱う必要のあるデータは加速度的に増加している。時には何万もの変数、特徴などを処理し解析する必要にせまられる場合がある。特徴抽出選択等の方法は、データの理解、データの視覚化、記憶領域の減少、訓練時間の短縮、処理時間の短縮、データを処理する基底にある仕組みの理解に関係しており、その必要性は差し迫ったものがある。この最後が内部表現の理解と直接関係しているが、その他の特徴は、この基底にある仕組みを理解することによって理解できると考える。

内部表現解釈の研究は、形を変えて研究が行われてきている。本研究は、これら研究成果を活かし、競合学習という既成の学習法を使う方法から、最終的には、情報エンハンスメント法と学習法が一体化した新しい学習法の提案へ移行しようとした。

## 2. 研究の目的

本研究では、ニューラルネットワーク（特に競合学習）が学習によって獲得した内部表現を明確に解釈する新しい方法を提案した。この解釈とは、ニューラルネットワークのすべての構成要素・要素結合の意味と機能を理解すると考える。また自然神経回路の知見により内部表現を設定し学習モデルを作成する研究も視野に入れた研究を行った。

ニューラルネットワークはその内部がブラックボックスと言われ、これが現実の応用への障害になっている。ニューラルネットワーク研究の初期では活発に解釈の問題がとりあげられたが、まだ決定的な解釈の方法は見出されていない。そこで、入力パターンについての相互情報量を考え、ある特定の要素に注意をおく（enhance）。この時の情報量の変化を見ることによって内部表現を理解する情報エンハンスメント法を提案した。

## 3. 研究の方法

学習方法として競合学習を取り上げ、情報エンハンスメント法が本当にネットワークの各部位の働きを理解することに役に立つかどうか検討することを中心として研究を行った。

技術的な取り組みとして次の4つの方法を挙げた。

### (1) 構成要素のエンハンスト法

ニューラルネットワークの構成要素をどのようにエンハンスするか、各種の方法を調査した。特に、ガウス分布の変形による方法を検討した。ガウスの変形は、統計力学の温度の変化に対応する。

### (2) エンハンスト対象

エンハンストする対象の問題である。単一の要素のエンハンスの方法からスタートし、これらの要素の組み合わせについてもエンハンスが可能であるかどうか検討した。ネットワークのすべての要素のすべての組み合わせの情報量を計算できる可能性を実証した。これまでの情報理論を用いた方法では、すべての要素の同時確率の計算が必要となり、計算は実質上不可能であった。この研究は、多くの変数間の関係を探求できる方法となる可能性が高い。

### (3) 情報構造

情報の構造についても検討した。情報量の種類を分類し、その間の関係を検討した。情報構造の研究は、各要素の細かい機能の理解に関係している。これまでは、単に相互情報量だけの計算が主体であったが、より細かい情報量をみることができるようになった。

### (4) 自由エネルギー

情報量の計算の加速化法として、自由エネルギーを用いて、学習を加速化する方法を提案した。自由エネルギーを用いると、相互情報量の計算は、単純な分配関数の計算に還元され、計算が非常に単純になることがわかった。

### (5) 階層型ネットワークへの拡張

学習方法として任意の方法が用いることができる。したがって、これまでに扱ってきた競合学習、SOMに加えて、良く用いられている教師付学習（BP、RBF）等へ応用し、その可能性を探求した。

## 4. 研究成果

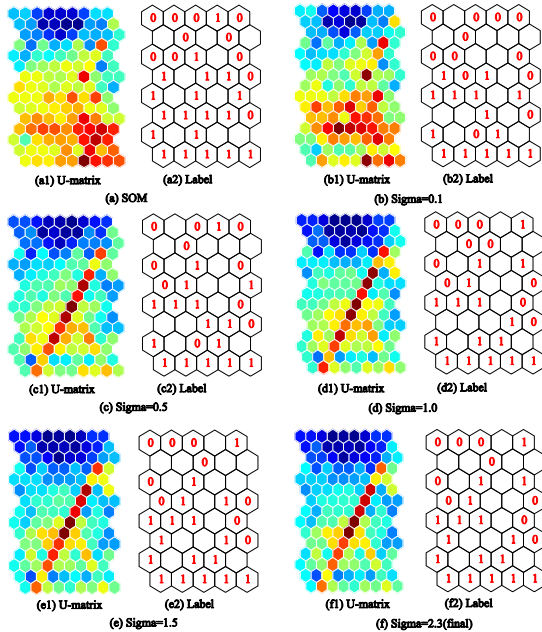
### (1) 3年間の成果の要約

ニューラルネットワークの内部表現を解釈する情報エンハンスメント法を提案した。3年間を通してエンハンスメント法は、解釈しやすい内部表現を生成するために有効であるということがわかった。研究を通して、ニューロンの社会性を考慮する新しい学習方法の提案へとつながった。内部表現の解釈に関する成果を次に要約する。

自己組織化マップへの応用を通して明確なグループの間の境界線を創りだすことに成功した。下図は、ニューラルネットワークの代表的な自己組織化マップの内部表現を図式化したものである。下図で、赤い部分がグループの境界となっている。一番左上の図は、標準の自己組織化マップを用いて得られた図である。赤い境界部分が下にあるが、分散して配置されておりグループの境界として見ることはできない。その後の図は、次第にエンハンストを高めて得られた図である。

最初は、自己組織化マップと同じような図が得られ得る。しかし、次第に赤い境界線が対角線上に現れる。あきらかに情報エンハンスメント法によって境界を明確にすることができたと言える。

境界が明確になることによってニューラルネットはどのようにデータを分類しているか理解することが容易になる。



## (2) 2009年度の成果

2009年度の研究内容は、次の3項目に要約できる、すなわち、ネットワーク構成要素のエンハンスメントの可能性、自己組織化マップへの応用である。

まず、エンハンスメントの可能性を多くの実験によって例証できた。情報エンハンスメントは、入力パターンについての相互情報量を考え、ある特定の要素に注意をおき、この時の情報量の変化を見ることによって内部表現を理解する方法である。要素として入力ユニット、競合ユニット、および入力パターンを考え、これらの重要性は、情報エンハンスメント法によって明確に測定できることがわかった。

次に、自己組織化マップに応用した場合に、測定された重要性を考慮し学習した場合、明確なクラスの境界線が表現できることがわかった。自己組織化マップは視覚化のツールとして広く用いられるようになってきている。しかし、必ずしも常に人間にとってわかりやすいマップが得られるとは限らない。エンハンスメント法を用いることによって、各構成要素の重要性を考慮したマップを生成できることが分かった。これは、より理解しやすい自己組織化マップを作ることに寄与すると考える。

2009年度は、情報エンハンスメント法が実際の問題に応用できる可能性があることを多くの例を通して実証することができたと考える

## (3) 2010年度の研究成果

2010年度の研究内容は、次の3項目に要約できる、すなわち、エンハンスメント法の変数選択への応用、成長型ネットワークへの応用、そして、エンハンスメント法による結果のより正確な評価である。

まず、エンハンスメント法の変数選択への応用である。特定の変数に注意を向ける、すなわち、ある変数をエンハンスすることによって生ずる相互情報量の変化を見ることによって、その変数の重要性を測定することができる。得られた重要な変数は、実際に学習する際に大きな役割を果たすが実験によって確かめられた。

次に、エンハンスメント法の成長型ネットワークへの応用である。成長型ネットワークでは、最小構成のネットワークからスタートし、次第にその複雑さを増して行く。エンハンスメント法を導入すると、最も重要なユニットで構成されているネットワークからスタートし、重要性の高い順番にユニットを付加させることができる。得られたネットワークは、通常のネットワークより高い能力を示した。

最後に、エンハンスメント法による結果の厳格な評価である。これまでは、評価には伝統的な評価測度を用いており、従来の方法との差異を明確にすることが困難であった。そこで最新の評価尺度、たとえば trustworthiness と continuity 等を導入し、エンハンスメント法の結果をよりよく評価することが可能となった。

## (4) 2011年度の研究成果

2011年度の研究内容は、次の2項目に要約できる。ニューロンの社会化、そして教師付き学習への拡張である。

まずニューロンの社会化である。解釈しやすい表現の研究を通して、ニューロンの個性を考慮する方法の必要性を感じた。ニューロンの個性（個別の特徴）とニューロン集団の特徴の抽出と相互作用の研究の必要性である。そこで、ニューロンの個別の特徴と集団的特徴をいかにして分離するかの研究をおこなった。この分離がある程度可能であることを確かめた。さらに、個別ニューロンと集団としてニューロン相互作用の研究にも着手した。

次に、方法の教師付き学習への拡張である。これまでの研究はすべて教師無し学習を対象としていた。このため応用の範囲が限定されていた。本年度は、教師付き学習も含めてより一般的なニューラルネットワークに適

用できる方法へと拡張した。

2011年度の研究を通して、情報エンハンスメント方法は、ニューロンの社会性を考慮する新しいニューラルネットワークの学習方法につながる可能性を見出すことができた。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計28件)

- (1) R. Kamimura, “Self-enhancement learning: target-creating learning and its application to self-organizing maps,” 査読有、Biological Cybernetics, Vol. 104, 2011, 305–338
- (2) R. Kamimura, “Constrained Information maximization by free energy minimization,” 査読有、International Journal of General Systems, Vol. 40, No.7, 2011, 701-725
- (3) R. Kamimura, “Selective information enhancement learning for creating interpretable representations in competitive learning,” 査読有 Neural Networks, Vol. 24, No.4, May, 2011, 387-405
- (4) R. Kamimura, “Comprehensibility maximization and humanly comprehensible representations,” 査読有、International Journal of General Systems, Vol. 41, No. 3, 2012, 265-287
- (5) R. Kamimura, “Explicit class structure by weighted cooperative learning,” 査読有、Lecture Notes in Computer Science, , Volume Vol. 6791, 2011, 109-116
- (6) R. Kamimura, “Cooperation control and enhanced class structure in self-organizing maps,” 査読有、Proceedings of the International Joint Conference on Neural Networks, 2011, 689 - 695
- (7) R. Kamimura, “Explicit class structure with closeness and similarity between neurons,” 査読有、Proceedings of World Congress on Nature and Biologically Inspired Computing (NaBIC), 2011, 92 - 98
- (8) R. Kamimura, “Supposed maximum information for comprehensible representations in SOM,” 査読有、Neurocomputing, Vol.74, No. 7, 2011, 1116-1134
- (9) R. Kamimura, “Structural enhanced information and its application to improved visualization of self-organizing maps,” Applied Intelligence, 査読有、Vol. 34, No.1, 2011, 102-115
- (10) R. Kamimura, “Information-theoretic enhancement learning and its application to visualization of self-organizing maps,”

Neurocomputing, 査読有、Vol.73, No.13-15, 2010, 2642-2664

- (11) R. Kamimura, “Relative relaxation and weighted information loss to simplify and stabilize feature detection,” Journal of Advanced Computational Intelligence and Intelligent Informatics, 査読有、Vol.13, No.4, 2009, 489-498
- (12) R. Kamimura, “Information enhancement for interpreting competitive learning,” International Journal of General Systems, 査読有、Vol. 39, No.7, 2010, 705-728
- (13) R. Kamimura, “Pseudo-network growing for gradual interpretation of input patterns,” LNCS(Lecture Notes in Computer Science, Neural Information Processing, Models and Applications, 査読有、Vol.6444, 2010, 375-382
- (14) R. Kamimura, “Generation of comprehensible representations by supposed maximum information,” LNCS(Lecture Notes in Computer Science, Artificial Neural Networks), 査読有、Vol. 6353, 2010, 333-342
- (15) R. Kamimura, “Information-theoretic competitive and cooperative learning for self-organizing Maps,” LNCS(Lecture Notes in Computer Science, Neural Information Processing, Models and Applications), 査読有、Vol. 6444, 2010, 423-430
- (16) R. Kamimura, “Explicit class structure produced by information-theoretic competitive and cooperative learning,” Proceedings of World Congress on Nature and Biologically Inspired Computing, 査読有、2010
- (17) R. Kamimura, “Enhancing and relaxing competitive units for feature discovery,” Neural Processing Letters, 査読有、Vol. 30, No.1, 2009, 37-57
- (18) R. Kamimura, “An information-theoretic approach to feature extraction in competitive learning,” Neurocomputing, 査読有、Vol. 72, No.10-12, 2009, 2693-2704
- (19) R. Kamimura, “Feature discovery by information loss,” Journal of Computers, 査読有 Vol.4, No.10, 2009, 943-953
- (20) R. Kamimura, “Feature detection by structural enhanced information,” LNCS(Lecture Notes in Computer Science, Advances in Neuro-Information Processing), 査読有、Vol. 5507, 2009, 171-178
- (21) R. Kamimura, “Partially enhanced competitive learning,” LNCS(Lecture Notes in Computer Science, Advances in Neuro-Information Processing), 査読有、Vol.5507, 2009, 163-170
- (22) R. Kamimura, “Collective activations to generate self-organizing maps,” LNCS(Lecture Notes in Computer Science, Advances in

Neuro-Information Processing), 査読有  
Vol.5507, 2009, 943-950

(23) R. Kamimura, "Enhanced visualization by  
combing SOM and mixture models," 査読有,  
LNCS(Lecture Notes in Computer Science,  
Advances in Neuro-Information Processing),  
Vol. 5507, 2009, 268-275

(24) R. Kamimura, "Self-enhancement learning:  
self-supervised and target-creating learning,"  
Proceedings of international joint conference of  
neural networks, 査読有, 2009, 1503-1509

(25) R. Kamimura, "Self-supervised learning by  
information enhancement: target-generating and  
spontaneous learning for competitive learning,"  
Proceedings of IEEE international conference on  
systems, man and cybernetics, 査読有, 2009,  
pp.113-119

(26) R. Kamimura, "Structural enhanced  
information to detect features in competitive  
learning," Proceedings of IEEE international  
conference on systems, man and cybernetics, 査  
読有, 2009, 3395-3401

(27) R. Kamimura, "Contradiction resolution and  
its application to self-organizing maps,"  
Proceedings of IASTED artificial intelligence  
and applications, 査読有, 2010.

(28) R. Kamimura, "Information maximization  
for variable selection in competitive learning,"  
Proceedings of IASTED artificial intelligence  
and applications, 査読有, 2010.

## 6. 研究組織

### (1) 研究代表者

上村 龍太郎 (KAMIMURA RYOTARO)  
東海大学・情報教育センター・教授  
研究者番号：80176643

### (2) 研究分担者

竹内 晴彦 (TAKEUCHI HARUHIKO)  
独立行政法人・産業技術総合研究所・ヒ  
ューマンライフテクノロジー研究部  
門・主任研究員  
研究者番号：00357401