

Hopfield ニューラルネットワークの性質に関する二、三の基礎的考察

京大工学部 西川 禰一 (Yoshikazu Nishikawa)

京大工学部 喜多 一 (Hajime Kita)

1.はじめに

計算機を用いた情報処理手法は飛躍的な進歩を遂げているが、画像や音声情報のパターン認識などの分野では未だ十分に実用的な処理を実現するには至っていない。これらの情報処理の困難さは、

1) 有効なアルゴリズムが発見されていないこと、2) 逐次的な処理では極めて長い処理時間を要すること、による。

これら従来の計算手法に内在するの困難を克服する一つのアイディアとして、近年、神経回路(ニューラルネットワーク)モデルを用いた情報処理が注目を集めている。¹⁾ すなわち、高等動物の脳は、それを構成する神経細胞の応答速度の低速性($O(\text{msec})$)にもかかわらず、従来の計算手法では困難とされる上記のような問題を実時間で効率よく処理していることに着目し、その計算原理を探って新しい情報処理技術として応用しようとするものである。

このような観点からいくつかの神経回路モデルが提案されているが、本稿ではその一つとして Hopfield らによって提案され、巡回セールスマン問題を例にデモンストレーションされた神経回路モデル(以下 H モデルと略称)と、その組合せ最適化計算への応用について考察する。²⁾ これは、離散的な組合せ最適化問題を常微分方程式で記述される神経回路モデルに埋め込み、その安定平衡点とし

て解を求めようとするものである。この手法は並列処理化の可能性と、それによる高速求解が期待できる点で関心を集めており、既に幾つかの具体的な例題への適用が報告されている。^{2,5-8,11-16)} しながら、手法の妥当性や限界などの基本的な点については、漠然とした期待あるいは批判¹⁷⁾が見られる程度で、詳細に検討された例は少ない。本論文では H モデルによる最適化計算の妥当性や限界について、計算機シミュレーションにより基礎的な検討を行う。

2. Hopfield 神経回路モデル

Hopfield によって提案された神経回路モデル (H モデル) には、連続時間・連続状態モデルと離散時間・離散状態モデルの 2 種類がある。

図 1 に Hopfield の連続時間・連続状態型神経回路モデル (以下、連続モデルと呼ぶ)^{2,3)} の構成を示す。回路中の各増幅器はそれぞれ計算ユニットあるいは神経細胞に対応し、図 2 に示すようなシグモイド型の非線形の入出力特性を持つ。増幅器 i の入力を u_i 、出力を v_i とし、結合コンダクタンスを T_{ij} 、バイアス電流を I_i で表すと、このシステムの挙動は次のような微分方程式で記述される。

$$\frac{du_i}{dt} = -\frac{u_i}{\tau} + \sum_j T_{ij}v_j + I_i \quad (1)$$

$$v_i = g(u_i) \equiv \frac{1 + \tanh(u_i/u_0)}{2} \quad (2)$$

ここで $u_0 > 0$ はシグモイド特性の急峻さを定めるパラメータであり、また τ は正定数である。

Hopfield はユニット間の相互結合行列 $T = \{T_{ij}\}$ が対称、すなわち $T_{ij} = T_{ji}$ の場合に、この神経回路モデルが次式で定義される「エネルギー関数 E 」を極小化するように挙動することを示した。³⁾

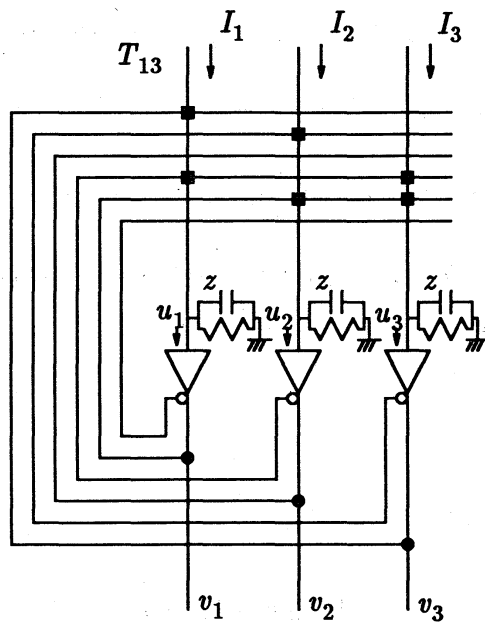


図1 Hopfield 型神経回路モデル

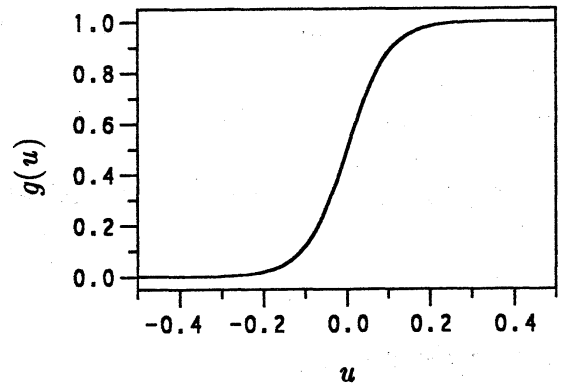


図2 増幅器のシグモイド特性

$$E = -\frac{1}{2} \sum_i \sum_j T_{ij} v_i v_j - \sum_i I_i v_i + \sum_i \frac{1}{\tau} \int_{v_0}^{v_i} g^{-1}(v) dv \quad (3)$$

ここで v_0 は $0 < v_0 < 1$ なる定数である。すなわち、エネルギー関数 E はこのシステムの Lyapunov 関数になっている。

増幅器のシグモイド特性が十分に急峻なとき、(3)式の右辺第3項は変数 v_i , $i = 1, \dots, n$ のとりうる範囲、すなわち n 次元超立方体 $[0, 1]^n$ の内部では無視できる。また、相互結合行列 T の対角成分 T_{ii} , $i = 1, \dots, n$ が0のとき、系の安定平衡点は超立方体のコーナ近傍に現れる。

相互結合 T_{ij} に対する対称性の要求について一言コメントしておこう。神経回路における同様に、多くの要素間に複雑な相互作用があるスピングラスのモデルについてはこの対称性は自然に導入されるが、神経細胞間のシナプス結合が方向性を持つ神経回路のモデルとしては不自然なものと言える。また H モデルは Cohen と Grossberg によってその安定性が議論された神経回路モデル^{9,10)}を、簡単化したものになっている。

離散時間・離散状態型モデル（以下、離散モデルと呼ぶ）⁴⁾とはユニットが0か1かの2つの状態のみをとり、その状態遷移則が離散時間によって次式のように表されるものである。

$$\begin{aligned}
 v_i(t+1) &= 1 && \text{if } \sum_{j \neq i} T_{ij} v_j + I_i > 0 \\
 &= 0 && \text{if } \sum_{j \neq i} T_{ij} v_j + I_i < 0 \\
 &= v_i(t) && \text{if } \sum_{j \neq i} T_{ij} v_j + I_i = 0
 \end{aligned} \tag{4}$$

この状態遷移則がある時刻に一つの状態変数についてのみ適用されるという仮定（非同期遷移）及び相互結合の対称性、さらに各ユニットが自己フィードバックを持たないという（ $T_{ii} = 0$ ）仮定のもとで、連続モデルと同様に、系は次式で定義されるエネルギー関数を極小化する：

$$E = -\frac{1}{2} \sum_i \sum_{j \neq i} T_{ij} v_i v_j - \sum_i I_i v_i \tag{5}$$

なお、このモデルの状態遷移則(4)を確率的にしたものがボルツマンマシン¹⁾である。

3. 神経回路モデルによる最適化計算

HopfieldらはHモデルの応用として、

- 1) 系が多く安定平衡点を持つことを利用した連想記憶装置、
- 2) 系がエネルギー関数を極小化することを利用した組合せ最適化問題の求解、

の2つを提案した。²⁾ 本稿で考察する後者は以下のような手順からなる：

- 1) 最適化すべき問題を、0と1の2値を定義域とする変数の、多変数2次関数の無制約最適化問題として表す。制約条件があれば、この範囲内で適当な罰関数として表現する。

- 2) 回路モデルがこの目的関数をエネルギー関数として持つように、回路定数を設計する。
- 3) 適当な初期点を与えて回路を動作させ、到達した安定平衡点を解とする。必要ならば初期点をいくつか与え直して計算を繰り返し、最も良い解を採用する。

Hopfield と Tank は、最適化計算には変数 v_i が 0 と 1 の中間の値をとって滑らかに遷移する連続モデルが、離散モデルよりも適していると主張している。²⁾ 彼らはこの手法の適用例として巡回セールスマン問題²⁾ や A/D コンバータ⁵⁾ 時系列信号の認識⁶⁾ などを示した。H モデルとその関連事項を図 3 にまとめる。

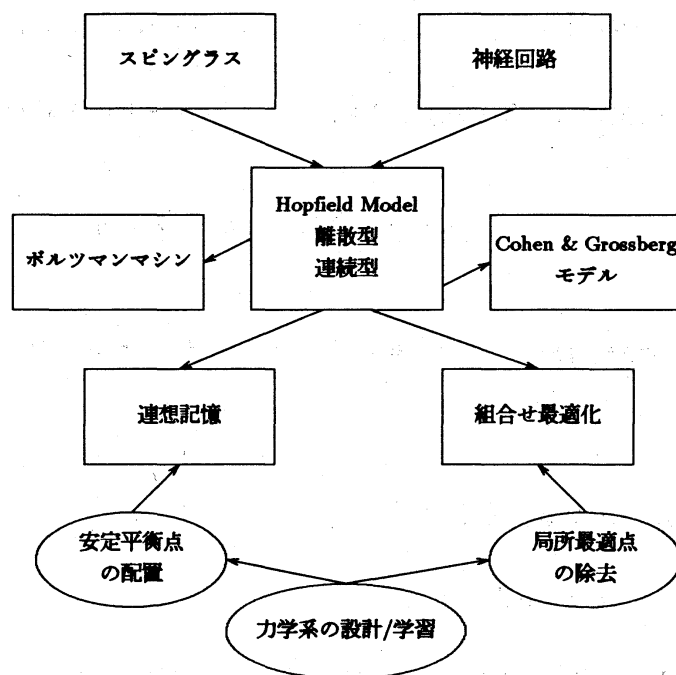


図 3 H モデルと関連事項

4. Hopfield モデルによる最適化計算の問題点

彼ら以外にも同様の計算手法を他の問題（例えばヒッチコック問題，ナップサック問題，回路ブロック配置問題，関係マッチング問題など）に適用した例が報告されている。^{7,11-16)} しかしながら，これらの適用例すべてにおいて十分に効果的な結果が得られている訳ではなく，本手法の妥当性や，その限界についての基礎的な検討が必要である。以下に検討すべき点を挙げる。

- (1) 本手法はより良い（エネルギー値の小さい）安定平衡点ほど，より広い basin（引込み領域）を持つことを前提としている。この前提は妥当であるか。また，この前提が成立しない場合には力学系はどのような様相を示すか。
- (2) 連続時間・連続状態モデルは離散時間・離散状態モデルに比べて，最適化の観点からどの程度有利か。なお相互結合 T の対角要素 T_{ii} を 0 とするような設定により，連続モデルにより得られる解は，離散モデルの局所最適解にもなっていることが上坂によって示されている。¹⁸⁾
- (3) 微分方程式(1)に与える初期点をどのように選定すればよいか。これに関して，上坂は超立方体の中心に初期点を与えることが有効であるとの予想を提出している。¹⁸⁾
- (4) 解くべき問題をどのように回路にエンコードすればよいのか。問題のエンコードの仕方によって必要となるユニット数や，生じる局所最適解の個数は変わるであろう。一般的で有効な設計法はあるのか。この点に関連して，Takedaらは数値データの表現法について検討している。⁷⁾ また，筆者らは不等式制約条件の導入方法を提案している。^{11,12)}
- (5) 回路の動作が不満足なものであったとき，学習により回路定数を改良するにはどうすればよいのか。フィードバックのある神経回路モデルの挙動を有効に制御する学習方法については，まだ十分な研究がなされていない。

本論文では上記の問題点のうち(1), (2)及び(3)についていくつかの数値実験を通じて考察する。

4. 数値実験 1

目的

3. で挙げた問題点のうち(1)を検討するために計算機シミュレーションを行った。すなわち「小さいエネルギー値を持つ安定平衡点ほど広い basin を持つ」という前提の妥当性を調べる。また、この前提が成立しない場合の具体例を得て力学系の様相を検討する。

方法

- 1) 対称で対角要素が0の行列 T を一様乱数を用いてランダムに発生させる。
- 2) 得られた各 T に対して、エネルギー関数 E の第1及び第2項からなる2次関数の鞍点が、超立方体の定められた位置を占めるようにベクトル $I=(I_1, \dots, I_n)^T$ を決定する。ここでは、鞍点を超立方体の中心に位置させた場合（鞍点中心）、及び超立方体の内部に中心を外して位置させた場合（非中心）を検討した。
- 3) このようにして得られたモデルについて、変数 $v(t)=(v_1(t), \dots, v_n(t))^T$ の初期点 $v(0)$ を超立方体の内部に一様乱数で発生させ、微分方程式を解く。そして到達したコーナ（安定平衡点）とその頻度を求める。

なお、本数値実験は行列 T と初期点 $v(0)$ に関する2重のランダム・シミュレーションであり、要求される計算時間も考慮する必要がある。ここではユニット数を8、 T を100例とし、各 T に対して1000個の初期点 $v(0)$ を用いた。

結果

実験結果を表1にまとめる。表に示されるように、行列 T の対角要素が0であるという設定のもとでも、局所最適なコーナがかなりの個数存在する場合（本シミュレーションでは最大16）がある。局所最適コーナでのエネルギー値と basin の大きさとの間には、概してかなり強い負の相関が観測された。

しかしながら、大域的に最適なコーナの basin が最大の大きさを持たない例が、鞍点を中心に置いた場合で7例、鞍点を中心から外した場合で4例観測された。このような場合に力学系がどのような様相を示しているのか興味深いところである。上述の例において観察される一つの特徴として、最大の到達率を持つコーナと最小のエネルギー値を持つコーナのほかに、ある程度大きな到達頻度を持つ安定なコーナが存在していることが挙げられる（ただし、鞍点を中心に置いた場合には、対称性から同じエネルギー値を持つ安定なコーナは2つずつ現れることに注意する必要がある）。これらの安定なコーナについて、例えばコーナ間のハミング距離などをより詳細に検討する必要がある。

実験1で得られた200例の結果のうち、典型的な例を図3及び4に示す。

表1 到達したコーナの数の分布

	連続モデル		離散モデル	
	鞍点中心	非中心	鞍点中心	非中心
最少数	2	3	2	2
最多数	16	11	12	7
平均数	8.34	6.35	5.46	3.80
標準偏差	3.22	2.27	2.32	1.26
最適コーナが最大の basin を持たない例	7例	4例	16例	12例

5. 数値実験2

目的

最適化の観点からみた連続モデルの離散モデルに対する優位性について検討する。

方法

実験1で用いた200例について同じ結合係数を持つ離散型のHモデルをシミュレーションし、実験1の結果と比較する。離散モデルのシミュレーションは、超立方体のコーナを一樣乱数により発生させ、非同期型の状態遷移則を適用して行う。遷移則を適用する変数も一樣乱数を用いて決定する。このようにして到達した（離散的に安定な）コーナごとにその到達頻度を求める。

結果

実験結果を表1に合わせて示す。離散モデルのほうが安定なコーナの個数が少なくなっているが、これは連続モデルにおけるエネルギー関数(2)式の右辺、第3項の影響、もしくは微分方程式の数値解法における収束判定条件の影響と考えられる。ただし、連続モデルでは安定に得られたコーナで、離散モデルでは不安定になったものは、いずれも到達頻度の低いものであったことを付言しておく。

安定平衡点の数だけで見ると離散モデルのほうが有利なように思えるが、最小エネルギー値のコーナへの到達頻度が最高ではない例の数では連続モデルの方が少なく、超立方体の内点を利用することの優位性が示されている。

もっとも、実際のシミュレーションでは連続モデルは微分方程式を解く必要がある分だけ、より多くの計算時間を必要とすることに注意しなければならない。両者の利点を兼ね備えた現実的な方法として、連続状態・離散時間モデルも提案されている⁸⁾

6. 数値実験3

目的

初期点の選定法について上坂の予想¹⁸⁾を検討する。

方法

実験1では、初期点を超立方体の内部に一様乱数で与えた。ここでは同じモデルについて、初

期点を

- 1) 超立方体の中心に与える（上坂予想）。
- 2) 2次関数で表されるエネルギー関数の鞍点の近傍に与える。

結果

実験3の結果を表2に示す。この結果を見れば、初期点を超立方体の中心や鞍点近傍に限定したときに概して良好な成績が得られている。しかしながら、鞍点を中心から外した場合に、超立方体の中心に与えた初期点から最適点に至らない例が11例観測された。すなわち、初期点を中心に置けば必ず最適点を見出し得るという上坂の予想は成立していない。

得られた結果について若干の考察を加える。超立方体の内部では力学系は近似的にエネルギー関数 E の第1及び第2項の勾配系であり、線形の方程式で記述される。従って、その解

$v(t) \equiv (v_1(t), \dots, v_n(t))^T$ は次の形で表される：

$$v(t) = \sum_i a_i x_i e^{\lambda_i t} + b \quad (6)$$

ここで a_i は初期点により定まる定数、 x_i は相互結合行列 T （に増幅器のゲインを掛けたもの）の固有値 λ_i に対応する固有ベクトル、また b は鞍点を表すベクトルである。

(3)式を見れば、総和中で、 T の最大固有値に対応する項が最も急速に増大することが分かる。

安定な平衡点の basin はすべて鞍点近傍にまで達していると考えられるが、この鞍点近傍に初期点を与えることは定数 a_i に小さな値を与えることであり、その結果、急速に増大する項の影響が相対的に大きくなる。すなわち初期点を鞍点近傍に限定すれば、行列 T の最小固有値に対応する固有ベクトル方向の軌道を選び出すことになる。

この方向はエネルギー値を最も効率的に減少させる方向であるが、移動可能な量は鞍点と超立方体との位置関係により決定されることに注意しなければならない。この方向の軌道を選ぶことにより最適点を得られる場合には、初期点領域を限定することにより最適点を見出す率を高められる。しかしながら、最小の固有値に対応する固有ベクトルの方向に最適点がない場合には、初期点領域を限定することは逆に最適点を得ることの妨げになる。

このような困難はモデルの大きさ n の増大に伴って深刻化するであろう。なぜならば、 n が大きくなれば、一般に中心から外れて位置する鞍点から各コーナへの距離の比が大きくなるからである。Hopfield らによる巡回セールスマン問題の適用例でも、問題が大規模になると回路パラメータの設定がクリティカルになることが報告されている。²⁾ 上の考察で示されたことが、H モデルを大規模最適化問題へ適用するに際した問題点となる。

次に超立方体の中心が最適点の basin に含まれない場合について考察する。このような状況を生じる一つの例として、エネルギー関数の鞍点から見て、相対的に絶対値の小さい正の固有値に対応する固有ベクトル方向の付近に超立方体の中心が位置している場合が考えられる。本実験で行ったように、ランダムに発生させたモデルではこのような状況の発生率は低かったが、特定の種類の問題から導かれるモデル、例えば巡回セールスマン問題を解くモデルではどのようになっているかは、検討する必要がある。

到達頻度

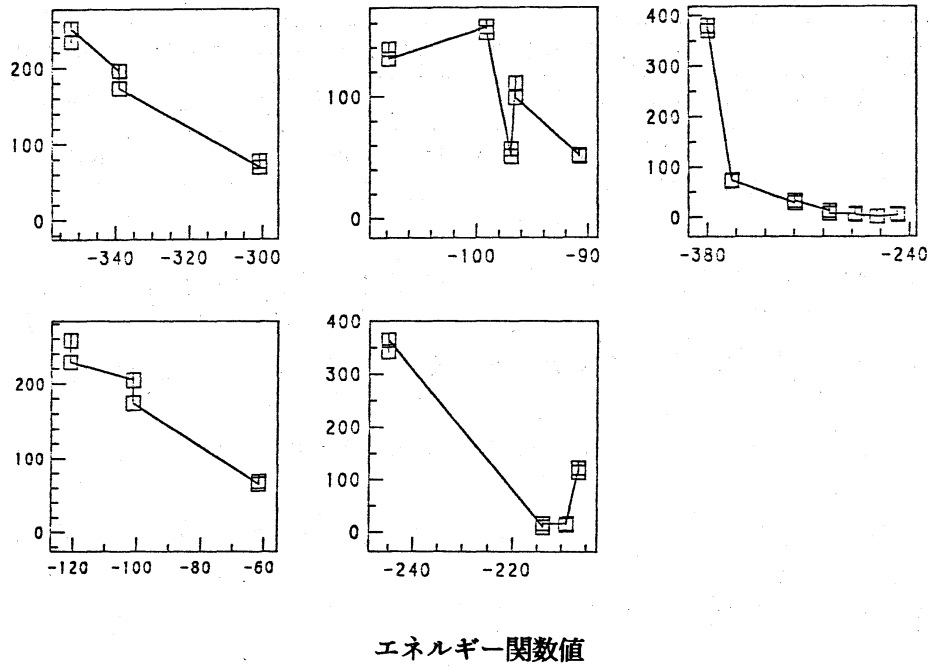


図3 コーナのエネルギー値と到達頻度の関係の例
(鞍点を超立方体の中心に置いた場合)

到達頻度

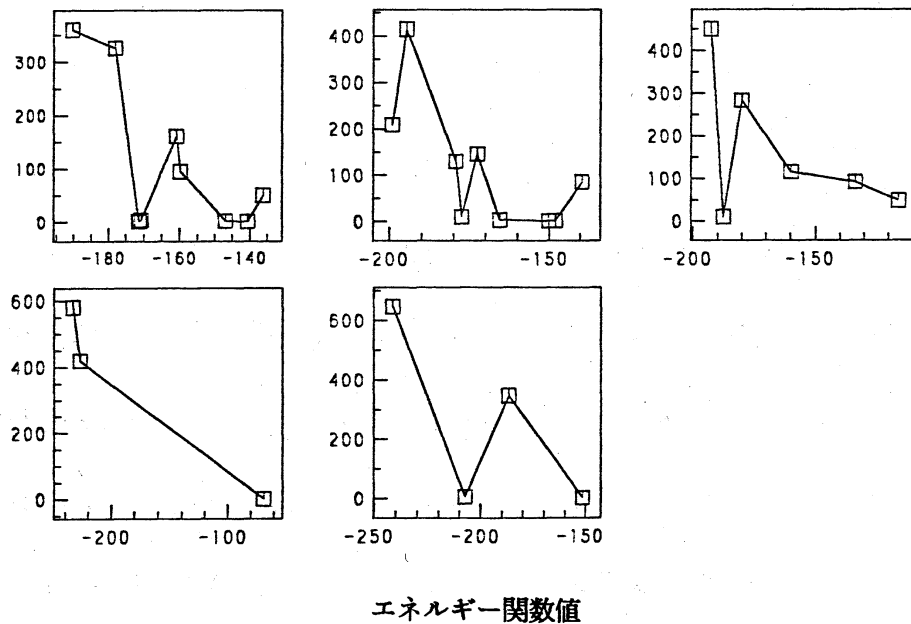


図4 コーナのエネルギー値と到達頻度の関係の例
(鞍点を超立方体の中心から外した場合)

表2 初期点領域を限定した場合

鞍点がある	中心近傍から最適点に達する比率が最大でない	20例中1例
鞍点がない	鞍点近傍から最適点に達する比率が最大でない	100例中19例
	中心から最適点に到達しない	100例中11例

以上の考察から分かるように、初期点を鞍点近傍に与えることと超立方体の中心近傍に与えることとは、エネルギー関数の形状に関連して、それぞれ異なった利点と欠点を持ち、補完的であると考えられる。表2の第2及び第3行に示した各初期点の設定手法が有効に働かない例のうち、双方に共通のものは少なく、2例のみであった。従って、これら両初期点設定法の折衷的手法も考えられるが、その有効性についてはさらに検討を必要とする。

なお、実験1で鞍点を中心から外して置いた場合に最適点が最大の basin を持たなかった4例のうち2例は、実験3で中心から最適点に到達しなかった11例に含まれている。これらの場合のエネルギー関数の形状（2次関数の固有値、固有ベクトル、鞍点の位置）と超立方体との位置関係については、より詳細に検討する必要があると思われる。

5. おわりに

本論文では Hopfield 神経回路モデルによる最適化計算に関して、その妥当性や初期点の選定法などについての基礎的検討を、計算機実験を用いて行った。その結果、この手法自体や従来提案されている初期点の選定法が、ある程度の妥当性を持つことが示された。また、それと同時に力学系の安定平衡点とその basin の大きさが必ずしも単純な関係にない例も観察された。これらの例をより詳細に分析することが、この手法の特徴や限界を明らかにするうえで有効なことと考えられる。今後、これらの点のより詳細な分析を進めるとともに、具体的な問題について、その構造を反映したパラメータを持つ H モデルについても検討する予定である。

本研究にあたって計算機シミュレーションを一部担当された京都大学工学部電気工学科・市森俊秀氏（現修士課程）に謝意を表する。

参考文献

- 1) D.E. Rumelhart, J.L. McClelland et al.: *Parallel Distributed Processing*, The MIT Press (1986).
- 2) J.J. Hopfield and D.W. Tank: *Biol. Cybern.*, Vol. 52, pp. 141-152 (1985).
- 3) J.J. Hopfield: *Proc. Natl. Acad. Sci. USA*, Vol. 81, pp. 3088-3092 (1984).
- 4) J.J. Hopfield: *Proc. Natl. Acad. Sci. USA*, Vol. 79, pp. 2554-2558 (1982).
- 5) D.W. Tank and J.J. Hopfield: *IEEE Trans.*, Vol. CAS-33, pp. 533-541 (1986).
- 6) D.W. Tank and J.J. Hopfield: *Proc. Natl. Acad. Sci. USA*, Vol. 84, pp. 1896-1900 (1987).
- 7) M. Takeda and J.W. Goodman: *Applied Optics*, Vol. 25, pp. 3033-3046 (1986).
- 8) B. Kangar-Parsi et al.: *Proc. IEEE 1st ICNN*, pp. III-785-790 (1987).
- 9) S. Grossberg (ed.): *The Adaptive Brain I*, North-Holland (1988).
- 10) Carpenter et al.: *SCIENCE*, Vol. 235, pp. 1226-1229 (1987).
- 11) 西川, 喜多, 旁: 第30回自動制御連合講演会, pp. 461-462 (1987).
- 12) 西川, 喜多, 旁: SICE 第7回自律分散システム研究会, pp. 73-78 (1988).
- 13) 坂本 他: JAACE 第32回システムと制御研究発表講演会, pp. 173-174 (1988).
- 14) 島本 他: JAACE 第32回システムと制御研究発表講演会, pp. 175-176 (1988).
- 15) 田中 他: 信学技報, MBE88-59, pp. 43-48 (1988).
- 16) 勝田 他: 信学技報, MBE88-15, pp. 35-41 (1988).
- 17) 深尾 他: JAACE 第32回システムと制御研究発表講演会, pp. 169-170 (1988).
- 18) 上坂: 信学技報, PRV88-6, pp. 7-14 (1988).