

First SICE Symposium on Computational Intelligence

September 30, 2011, Kyoto

第1回コンピューターショナル・インテリジェンス研究会

「クリフォードニューロコンピューティングを中心として」

講演論文集

期 日：2011年9月30日(金)

会 場：京都工芸繊維大学

**SICE®**

主 催：計測自動制御学会 システム・情報部門

企 画：ニューラルネットワーク部会

協 賛：情報処理学会，システム制御情報学会，電子情報通信学会，電気学会，  
日本神経回路学会，日本機械学会，人工知能学会，日本知能情報ファジィ  
学会，ヒューマンインタフェース学会，Japan Chapter of IEEE  
Computational Intelligence Society，Japan Chapter of IEEE Systems, Man,  
and Cybernetics Society

後援：京都工芸繊維大学

著作権 © 2011

公益社団法人計測自動制御学会(SICE)  
〒113-0033 東京都文京区本郷 1-35-28-303

カタログ番号 11 PG 0009

著作権は、計測自動制御学会がもっている  
ので、個人の使用のための複製以外の  
目的で掲載の記事の一部または全文を  
複製する場合には、著作権者に許可を求  
め規定の複製料を支払うこと。

発行日：2011年9月30日

発行者：公益社団法人計測自動制御学会 システム・情報部門  
ニューラルネットワーク部会

# First SICE Symposium on Computational Intelligence

September 30, 2011, Kyoto

## 第1回コンピューターショナル・インテリジェンス研究会 「クリフォードニューロコンピューティングを中心として」

最近のニューラルネットワークやコンピューターショナル・インテリジェンスに関する研究・技術の発展には著しいものがあります。このような状況を鑑み、計測自動制御学会システム・情報部門では、新たな研究成果の発表、研究交流の場として「コンピューターショナル・インテリジェンス研究会」を開催することにしました。第1回の今回はサブテーマを「クリフォードニューロコンピューティングを中心として」として、広くこの分野の学生・研究者・実務者の交流をはかりたいと存じます。趣旨は以下に示しますが、関係者多数の研究会への参加をお願いいたします。

この研究会を発展させることにより、この分野の新たなパラダイムを切り拓きたいと存じます。

### 趣旨：

近年、実数で表現されていたニューラルネットワークを複素数値化など高次元化したニューラルネットワークのモデルが提案され、その情報処理能力、学習法や応用などに関する研究が盛んに行われています。本研究会では、このように豊かな表現能力を持ち高度な計算知能を実現できるものとして非常に期待されている高次元の表現を用いたニューロコンピューティングを取り上げます。複素表現、四元数表現、さらにはそれらを一般化、包含するクリフォード代数表現 (Geometric Algebra) を用いたニューロコンピューティングの基礎理論から最新の応用まで様々な問題について議論し、その可能性、研究の将来動向を探ります。本研究会は、以上のようなテーマを中心としますが、これに限らず関連する研究、周辺の研究も広く取り上げたいと思います。たとえば、ニューロに限らず、複素、四元数など高次元情報処理、クリフォード代数情報処理、量子情報処理なども含むものとします。ちなみに、最近話題となっている小惑星探査機「はやぶさ」の姿勢情報は、四元数で表現されました。

企画担当 黒江 康明 (京都工芸繊維大学)  
新田 徹 (産業技術総合研究所)

# 目 次

[挨拶] 9:50-10:00 ニューラルネットワーク部会主査 見浪 護 (岡山大学)

[セッション 1] 10:00-11:00 司会：黒江 康明 (京都工芸繊維大学)

[1] 同時摂動を用いた高次元ニューラルネットワークの学習

○山田 貴博, 前田 裕 (関西大学) (1)

[2] 高次元連想記憶モデルとその基本特性

○磯川 悌次郎, 西村 治彦, 松井 伸之 (兵庫県立大学) (5)

[3] 高次元信号に対する広域線形推定法

○新田 徹 (産業技術総合研究所) (11)

[セッション 2] 11:10-12:30 司会：新田 徹 (産業技術総合研究所)

[4] リカレントクリフォードニューラルネットワークのモデルとダイナミックス

○黒江 康明 (京都工芸繊維大学) (15)

[5] Non-constant bounded holomorphic functions of hyperbolic numbers

○Eckhard Hitzer (福井大学) (23)

[6] Conformal Geometric Algebra を用いた近似方法の提案とその応用

○ファン ミン トウン, 橘 完太, 吉川 大弘, 古橋 武 (名古屋大学) (29)

[7] 断熱的量子計算におけるハミルトニアン変化の高速化に関する考察

○金城 光永, 十川 雄一郎, 山下 大輔, 佐藤 茂雄, 島袋 勝彦 (琉球大学) (37)

[セッション 3] 13:30-15:10 司会：松井 伸之 (兵庫県立大学)

[8] リアプノフ法で保証された 3次元追跡 Eye-Vergence ビジュアルサーボ実験の周波数応答

○于 福佳, 松本 紘明, 宋 薇, 見浪 護, 矢納 陽 (岡山大学) (41)

[9] 魚捕獲ロボットのためのニューラルネットワーク組み込み型微分方程式によるカオスの生成とその検討

○伊藤 雄矢, 友野 高志, 見浪 護, 矢納 陽 (岡山大学) (49)

[10] 多モード情報を統合する複数の複素 SOM による地雷概念形成

○江尻 礼聡, 廣瀬 明 (東京大学) (57)

[11] 位相感受型ニューラルネットワークを用いたミリ波イメージングシステム

○小野島 昇吾, 廣瀬 明 (東京大学) (63)

[12] 複素行列因子分解と音響経路推定に基づく自動採譜

○池内 亮太, 池田 和司 (奈良先端科学技術大学院大学) (67)

[セッション 4] 15:20-17:00 司会：前田 裕 (関西大学)

[13] 定数項を用いた複素連想記憶

○北原 倫理, 小林 正樹 (山梨大学) (73)

[14] 時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリ

吉田 明生, ○長名 優子 (東京工科大学) (77)

[15] 複素多層パーセプトロンの探索空間と探索法

○鈴木 真矢, 中野 良平 (中部大学) (85)

[16] 複素ネットワークインバージョンによる逆問題解法と正則化

○中村 恭介, 小川 毅彦 (拓殖大学) (93)

[17] 量子ビット遺伝的アルゴリズムの基本性能評価

○村本 憲幸, 磯川 悌次郎, 松井 伸之 (兵庫県立大学) (97)

17:10～ 懇親会 京都工芸繊維大学 KITHOUSE オルタス



# 同時摂動を用いた高次元ニューラルネットワークの学習

○山田貴博 前田裕 (関西大学)

## Learning Via Simultaneous Perturbation Method for High-dimensional Neural Network

\* T. Yamada and Y. Maeda (Kansai University)

**Abstract**—Usually, the back-propagation learning rule is widely used also for high-dimensional neural networks. In this paper, we propose a learning method for quaternion neural networks using the simultaneous perturbation method. Learning process of the proposed method is simpler than the back-propagation. Comparison between the back-propagation method and the proposed simultaneous perturbation learning rule is made for some test problems. Simplicity of the proposed method results in faster learning speed.

**Key Words:** Simultaneous perturbation method, Quaternion neural networks, learning

### 1 はじめに

ニューラルネットワークに複素数や4元数を用いた高次元ニューラルネットワークが注目されている。高次元ニューラルネットワークは主として90年代に提案され、画像処理などへの応用が行われている。

複素数を拡張した数体系である4元数については3次元空間の変換を簡潔に表現することができるため人工衛星の姿勢制御やコンピュータグラフィックでの応用が行われている。この4元数ニューラルネットワークの重みおよびしきい値の学習には、実数値を用いる通常のニューラルネットワークで用いられるバックプロパゲーションを拡張した4元数バックプロパゲーション学習則が提案されている<sup>1)2)</sup>。

これに対し、本研究では、4元数ニューラルネットワークの学習に、確率的勾配法として知られる同時摂動最適化法を用いることを提案する。学習速度とその収束率について検討した。

### 2 4元数の定義

4元数は W.R.Hamilton によって 1843 年に発見された4次元の数である。4元数は複素数を拡張した概念であり、4元数全体を表す集合  $H$  は以下のようにあらわされる。

$$H = \{X | X = x_1 + x_2i + x_3j + x_4k\} \quad (1)$$

ここで、 $i^2 = j^2 = k^2 = ijk = -1$ 、 $ij = -ji = k$ 、 $jk = -kj = i$ 、 $ki = -ik = j$  である。4元数は積に対して結合法則を満たし、和に対して分配法則を満たす<sup>3)</sup>。

### 3 4元数ニューラルネットワーク

本研究では、入力信号、荷重、閾値、出力信号がすべて4元数であるニューラルネットワークについて考える。ニューロン  $l$  の内部ポテンシャル  $Z_l$  は

$$z_l = \sum_{i=1}^n x_i w_{li} - \theta_l \quad (2)$$

と定義する。ここで、 $Z_l$  は、ある層の  $l$  番目のニューロンへの入力値、 $w_{li}$  は全層の  $i$  番目のニューロンとこの  $l$  番目のニューロン間の荷重、 $x_i$  は前層の  $i$  番目のニュー

ロンの出力値、 $\theta_l$  は閾値である。出力信号  $f$  は次のよう定義される。

$$f(z_l) = f(x_1) + f(x_2)i + f(x_3)j + f(x_4)k \quad (3)$$

$$Z_l = x_1 + x_2i + x_3j + x_4k \quad (4)$$

$$f(x_l) = \frac{1}{1 + \exp(-x_l)} \quad (5)$$

### 4 4元数バックプロパゲーション

4元数バックプロパゲーションを適用するニューラルネットワークを考える。ここでは3章で定義した4元数ニューロンだけを用いて3層のネットワークを構成する。構成したネットワークを Fig.1 に示す。

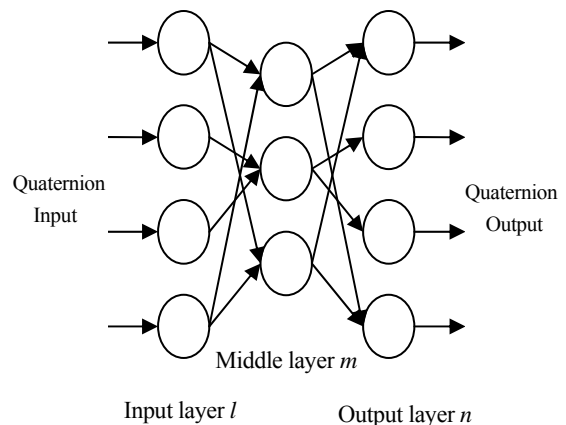


Fig.1 Quaternion Neural Network.

ニューロン  $l$  と中間ニューロン  $m$  との間の荷重を表す4元数を

$$w_{ml} = w_{ml}^a + w_{ml}^b i + w_{ml}^c j + w_{ml}^d k (\in H)$$

とする。ここで、 $H$  は4元数全体の集合を表す。また、中間ニューロン  $m$  と出力ニューロン  $n$  との間の荷重を示す4元数を

$$v_{nm} = v_{nm}^a + v_{nm}^b i + v_{nm}^c j + v_{nm}^d k (\in H)$$

中間ニューロン  $m$  の閾値を表す4元数を

$$\theta_m = \theta_m^a + \theta_m^b i + \theta_m^c j + \theta_m^d k (\in H)$$

出力ニューロン  $n$  の閾値を表す4元数を

$$\gamma_n = \gamma_n^a + \gamma_n^b i + \gamma_n^c j + \gamma_n^d k (\in H)$$

とする。

$$I_l = I_l^a + I_l^b i + I_l^c j + I_l^d k (\in \mathbf{H})$$

は入力ニューロン  $l$  への入力信号を表す 4 元数で、

$$H_m = H_m^a + H_m^b i + H_m^c j + H_m^d k (\in \mathbf{H})$$

$$O_n = O_n^a + O_n^b i + O_n^c j + O_n^d k (\in \mathbf{H})$$

は、それぞれ、中間ニューロン  $m$ 、出力ニューロン  $n$  の出力値を表す 4 元数とする。

$$\Delta_n = \Delta_n^a + \Delta_n^b i + \Delta_n^c j + \Delta_n^d k = T_n - O_n (\in \mathbf{H})$$

を  $O_n$  と出力ニューロン  $n$  に対する教師信号

$$T_n = T_n^a + T_n^b i + T_n^c j + T_n^d k (\in \mathbf{H})$$

との誤差とする。パターン  $p$  に対する 2 乗誤差を

$$E_p = (1/2) \sum_{n=1}^N |\Delta_n|^2 \quad (6)$$

と定義する。ここで、 $N$  は出力ニューロンの総数、

$$|x| = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}$$

$$x = x_1 + x_2 i + x_3 j + x_4 k (\in \mathbf{H})$$

である。

#### 4.1 学習アルゴリズム

3 章で構成した 4 元数バックプロパゲーション学習則のパラメータ修正量を示す。  $\Delta x$  はパラメータ  $x$  の修正量を表す。

$$\Delta v_{nm} = \bar{H}_m \Delta \gamma_n \quad (7)$$

$$\Delta \gamma_n = \varepsilon \{ \Delta_n^a (1 - O_n^a) O_n^a + \Delta_n^b (1 - O_n^b) O_n^b i + \Delta_n^c (1 - O_n^c) O_n^c j + \Delta_n^d (1 - O_n^d) O_n^d k \} \quad (8)$$

$$\Delta w_{ml} = \bar{I}_l \Delta \theta_m \quad (9)$$

$$\begin{aligned} \Delta \theta_m = & (1 - H_m^a) H_m^a \cdot \text{Re} [ \sum_n \Delta \gamma_n \bar{v}_{nm} ] \\ & + (1 - H_m^b) H_m^b \cdot \text{Im}^i [ \sum_n (\Delta \gamma_n \bar{v}_{nm}) ] i \\ & + (1 - H_m^c) H_m^c \cdot \text{Im}^j [ \sum_n (\Delta \gamma_n \bar{v}_{nm}) ] j \\ & + (1 - H_m^d) H_m^d \cdot \text{Im}^k [ \sum_n (\Delta \gamma_n \bar{v}_{nm}) ] k \quad (10) \end{aligned}$$

ここで、

$$\bar{x} = x_1 - x_2 i - x_3 j - x_4 k$$

$$\text{Re}[x] = x_1$$

$$\text{Im}^i[x] = x_2$$

$$\text{Im}^j[x] = x_3$$

$$\text{Im}^k[x] = x_4$$

である<sup>2)</sup>。

#### 5 同時摂動学習則

同時摂動最適化法は、差分近似の拡張として、パラメータの次元を増やしても、評価関数に対する観測回

数を増やすことなく、勾配を推定する手法として考案された確率的な勾配法である<sup>4)</sup>。また、アルゴリズムの簡便性からニューラルネットワークの学習則への適用が提案されており、学習機能を有するニューラルネットワークのハードウェア化とともに有用性が示されている。

$w (\in \mathbf{R}^M)$  をパラメータベクトル、 $J$  を評価関数とすると、符号ベクトルを用いた同時摂動による最適化のアルゴリズムはつぎのようになる。

$$w_{t+1} = w_t - \alpha \Delta w_t \quad (11)$$

$$\Delta w_t^n = \frac{J(w_t + c s_t) - J(w_t)}{c s_t^n} \quad (n=1, \dots, M) \quad (12)$$

$c (> 0)$  は、すべての要素に共通の摂動の大きさを表わす。 $\alpha$  は正の値の実数である。また、 $s_t$  および  $s_{t,i}$  は符号ベクトルとその第  $i$  要素を表し、この要素は +1 あるいは -1 の値を取るものとする。

同時摂動最適化法では、すべてのパラメータベクトル  $w$  のすべての要素に同時に  $+c$  あるいは  $-c$  の摂動を加える。摂動を加えた場合と加えない場合の評価関数に対する 2 回の計算のみで、その点における勾配を推定することができる。パラメータの次元が大きくなった場合にも、関数の二つの値のみから勾配ベクトルの推定値を求めることができる。このため、高次元のパラメータを持つ最適化問題では、差分近似と比べた場合、この手法のほうが明らかに有効である。

$w$  を閾値も含めた荷重ベクトル、 $J$  を評価関数と見なすと、この手法は 4 元数ニューラルネットワークの学習則と考えることができる。この場合、すべての荷重としきい値に摂動を加えた場合の評価関数の値と摂動がない場合の評価関数の値のみから、すなわち、2 回の 4 元数ニューラルネットワークの前向きの動作のみから、その点における勾配を計算することができることになる。これを用いて、荷重としきい値の更新を行うことができる。

荷重、閾値などの学習を行うパラメータを  $w$  とするとパラメータの修正量  $\Delta w$  は次式で表せる。

$$\begin{aligned} \Delta w_t = & \frac{J(w_t + c s_t) - J(w_t)}{c} \text{Re}[s_t] + \frac{J(w_t + c s_t) - J(w_t)}{c} \text{Im}^i[s_t] i \\ & + \frac{J(w_t + c s_t) - J(w_t)}{c} \text{Im}^j[s_t] j + \frac{J(w_t + c s_t) - J(w_t)}{c} \text{Im}^k[s_t] k \quad (13) \end{aligned}$$

ここで  $J$  は評価関数である二乗誤差

$$E = \sum_{p=1}^P E_p \quad (14)$$

である。また、 $s$  は +1、もしくは -1 の値を持つ 4 元数

である。Fig. 2 に同時摂動学習則で4元数ニューラルネットワークの学習を行うときのフローチャートを示す。

まず、符号ベクトルをランダムに+1あるいは-1かに設定します。次に各入力パターンごとに摂動なしでのニューラルネットワークの計算を行い、パラメータに摂動を加えその値を用いて、同じように各パターンのニューラルネットワークの計算を行います。これらの計算で得られた値を(13)式を用いてパラメータの修正量を求めます。このように提案手法は、その手順が簡単で、動作速度やハードウェアでの実現で有用であると考えられる<sup>5)</sup>。

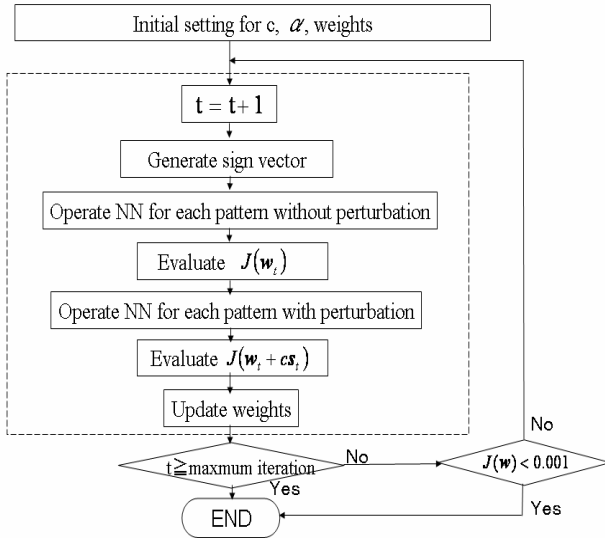


Fig.2 Flowchart.

## 6 同時摂動学習則とバックプロパゲーション学習則の比較

同時摂動学習則と4元数バックプロパゲーション学習則の比較を行うために、入力データを反転させる反転問題と入力データの原点からの距離を二分の一にする縮小問題を取り上げた。最大学習回数を、反転、縮小問題ともに10000回と設定した。100回の試行を行い、最大学習回数までに評価関数値が0.01以下になった場合正しく収束したと考え、この場合の収束率と平均収束回数を求めた。荷重と閾値の初期値は-1~+1の範囲の一様分布の乱数で生成した。各学習における摂動および修正係数は予備実験を通して、適切な値を事前に求めた。3.07GHzで動作するCore i7 950を用いて、Windows XP上のMatlabでの実験を行った。

### 6.1 反転問題

入力されたデータを反転して出力する問題を考える。反転問題における入力データのパターンとそれに対応する教師信号をTable 1に示す。測定には入力層、中間層、出力層が、それぞれ、1の4元数ニューラルネットワークを用いた。

### 6.2 縮小問題

入力されたデータを一定の比率で縮小し、これを出力する問題について考える。縮小問題におけるニューロンへの入力データは反転問題の入力データのパターンと同じものを用いた。縮小問題においても反転問題と同じ1-1-1の4元数ニューラルネットワークを用いた。縮小の比率を0.5として、実験を行った。

Table 1 Pattern of reversing problem.

	Input	Teaching signals
1	(0,0,0,0)	(1,1,1,1)
2	(1,0,0,0)	(0,1,1,1)
3	(0,1,0,0)	(1,0,1,1)
4	(0,0,1,0)	(1,1,0,1)
5	(0,0,0,1)	(1,1,1,0)
6	(1,1,0,0)	(0,0,1,1)
7	(0,1,1,0)	(1,0,0,1)
8	(0,0,1,1)	(1,1,0,0)
9	(1,0,1,0)	(0,1,0,1)
10	(1,0,0,1)	(0,1,1,0)
11	(0,1,0,1)	(1,0,1,0)
12	(1,1,1,0)	(0,0,0,1)
13	(1,1,0,1)	(0,0,1,0)
14	(1,0,1,1)	(0,1,0,0)
15	(0,1,1,1)	(1,0,0,0)
16	(1,1,1,1)	(0,0,0,0)

## 7 実験結果

反転問題、縮小問題における同時摂動学習則と4元数バックプロパゲーション学習則の結果をそれぞれTable 2、Table 3に示す。

Table 2 Reversing problem.

	Simultaneous perturbation	Back-propagation
Perturbation $c$	0.00001	—
Modification coefficient $\alpha$	0.4	1.0
Convergence rate (%)	85	100
Average iteration for convergence	743.55	354.64

Average for 100 times trials

Table 3 Reduction problem.

	Simultaneous perturbation	Back-propagation
Perturbation $c$	0.00001	—
Modification coefficient $\alpha$	0.3	0.8
Convergence rate (%)	98	100
Average iteration for convergence	1142.86	540.38

Average for 100 times trials

これより、4元数バックプロパゲーション学習則の方が収束率、収束平均回数ともに優れている結果となった。いずれの問題においても、収束のために必要とされる学習回数は、同時摂動学習則では、バックプロパゲーション法の約2倍であることがわかる。

入力層、中間層、出力層が1の四元数ニューラルネットワークにおいて、同時摂動学習則と4元数バックプロパゲーションそれぞれ10000回の学習にかかる計算時間とCPU TIMEを計測した。測定結果をTable4に示す。この結果より、10000回の修正にかかる計算時間は同時摂動学習則の方が低いことがわかる。つまり、一回の学習に要するCPU TIMEは、同時摂動学習則では、バックプロパゲーション法の約1/2であることがわかる。

Table 4 CPU time.

	Simultaneous perturbation	Back-propagation
CPU Time	754.09	1440.39

## 8 まとめ

同時摂動学習則を用いた4元数ニューラルネットワークについて提案し、簡単な問題を通して、バックプロパゲーション法との比較を行った。学習平均収束回数、収束率の点では、いずれの問題とも同時摂動学習則はバックプロパゲーション法に劣るものの、1回の修正にかかる計算時間は同時摂動学習則がバックプロパゲーション法より少ない。

以上、同時摂動学習則は4元数ニューラルネットワークの学習に対しても適用可能で、4元数バックプロパゲーション法とほぼ同等の学習性能を有することが分かった。

また、ニューラルネットワークのある種の応用においては、バックプロパゲーション法を直接用いることが困難な場合がある。このような場合、同時摂動学習則の適用により、高次ニューラルネットワークの活用範囲が広がると考えられる。

## 参考文献

- 1) 新田徹:複素バックプロパゲーション学習, 情報処理学会論文誌, **32-10**, 1319/1329 (1991)
- 2) Tohru NITTA, Masaru TANAKA: Current Status of Research on Neural Networks with High-dimensional Parameters, 電子技術総合研究所調査報告, **28**, 48/50 (1994)
- 3) 吉田英司: 四元数ニューラルネットワークの性質について, 電子情報通信学会技術研究報告, **107-157**, 29/34 (2007)
- 4) 山田貴博: 同時摂動を用いた複素ニューラルネットワークの学習則, インテリジェントシステムシンポジウム **20-99** (2010)
- 5) 前田裕: 同時摂動最適化法とその応用, システム/制御/情報, **52-2**, 47/53 (2008)

# 高次元連想記憶モデルとその基本特性

○磯川 悌次郎 西村 治彦 松井 伸之 (兵庫県立大学)

## Fundamental Properties on Hypercomplex-valued Associative Memory

\*Teijiro Isokawa, Haruhiko Nishimura, and Nobuyuki Matsui (University of Hyogo)

**Abstract**— Associative memories by Hopfield-type recurrent neural networks with quaternionic algebra, called quaternionic Hopfield neural network, are introduced in this paper. The variables in the network are represented by quaternions of four dimensional hypercomplex numbers. The neuron model, the energy function, and the rules for embedding patterns into the network are presented.

**Key Words:** Quaternion, Hopfield neural network, Multistate, Hebbian rule, Projection rule

### 1 はじめに

近年、複素数をはじめとする多元数体系に基づくニューラルネットワーク (NN) に関する研究が盛んに行われている。最も基本的な多元数体系である複素数を導入した複素 NN においては、その基礎理論について様々な研究が行われている<sup>1)</sup>。また、複素数は信号処理をはじめとする様々な工学問題に用いられているため、複素 NN はこれらの問題との親和性が高く、様々な応用システムが構築されている。

さらに、より高次元の数体系である四元数を NN に導入する試みについても行われている。四元数は4成分からなる超複素数であり、三次元空間における幾何学変換の記述に適しているために、物理学やコンピュータグラフィックスの分野で用いられている。四元数 NN の研究としては、単一四元数ニューロンの計算能力の評価<sup>2)</sup>、階層型ネットワークとその学習アルゴリズムの提案<sup>3, 4, 5)</sup>などが挙げられる。また、時系列予測、剛体制御<sup>3)</sup>、カラーナイトビジョン<sup>6, 7)</sup>などの工学問題に対する四元数ネットワークの応用も行われている。さらに、四元数領域において局所解析可能な活性化関数を持つ適応フィルタを提案し、カオス回路の出力を予測する試みが行われている<sup>8)</sup>。

ホップフィールド型(リカレント)NNに関しては、連続時間・連続状態 NN でのエネルギー関数の存在条件<sup>9)</sup>が示されており、さらに具体的なモデリングに向けて離散時間型 NN の検討が進められているのが現状である<sup>10, 11, 12, 11, 13, 14)</sup>。四元数を用いたホップフィールド型 NN の特性を明らかにすることにより、カラー画像あるいは3次元画像を対象とした連想記憶やパターン認識システムの構成が可能になると考えられる。

本論文では、著者らがこれまでに検討を行ってきた四元数に基づくホップフィールド型ニューラルネットワークモデルならびにその連想記憶への応用について述べる。これらのモデルは、離散時間型のモデルであり、状態変数としては、離散状態、連続状態、位相表現を用いた多値状態がある。以下では、四元数の表現、演算規則などを紹介した後に、上記のネットワークモデルおよびその安定条件、ならびにネットワークにパターンを記録するための学習則について述べる。

### 2 準備

#### 2.1 四元数の定義および演算規則

四元数は複素数を拡張した数体系であり、1つの実数と3つの虚数からなる。3つの虚数単位を  $i, j, k$  とすると四元数  $x$  は

$$x = x^{(e)} + x^{(i)}i + x^{(j)}j + x^{(k)}k \quad (1)$$

と表される。ここで、 $x^{(e)}, x^{(i)}, x^{(j)}, x^{(k)}$  は実数であり、四元数  $x$  の各成分を表す。したがって、四元数全体の集合  $H$  は  $1, i, j, k$  を基底とする四元数ベクトル空間を構成する。また四元数は、スカラー部分  $x^{(e)}$  とベクトル部分  $\vec{x} = \{x^{(i)}, x^{(j)}, x^{(k)}\}$  に分けることにより、

$$x = (x^{(e)}, x^{(i)}, x^{(j)}, x^{(k)}) = (x^{(e)}, \vec{x}) \quad (2)$$

とも表記される。四元数  $x (x \in H)$  の共役な四元数  $x^* (x^* \in H)$  は

$$\begin{aligned} x^* &= (x^{(e)}, -\vec{x}) \\ &= x^{(e)} - x^{(i)}i - x^{(j)}j - x^{(k)}k \end{aligned} \quad (3)$$

と定義される。

四元数の虚数単位間には次式の Hamilton 関係が成り立つ。

$$\begin{aligned} i^2 = j^2 = k^2 = ijk = -1, \\ ij = -ji = k, jk = -kj = i, ki = -ik = j \end{aligned} \quad (4)$$

ここで、 $ij \neq ji$  が示すように、四元数は非可換の数体系である。

以上の定義に基づいて、四元数  $p = (p^{(e)}, \vec{p})$  と  $q = (q^{(e)}, \vec{q})$  間の演算は以下ようになる。加減演算は、

$$\begin{aligned} p \pm q &= (p^{(e)} \pm q^{(e)}, \vec{p} \pm \vec{q}) \\ &= (p^{(e)} \pm q^{(e)}, p^{(i)} \pm q^{(i)}, p^{(j)} \pm q^{(j)}, p^{(k)} \pm q^{(k)}) \end{aligned} \quad (5)$$

となる。四元数  $p$  と  $q$  の積  $p \otimes q$  は

$$p \otimes q = (p^{(e)}q^{(e)} - \vec{p} \cdot \vec{q}, p^{(e)}\vec{q} + q^{(e)}\vec{p} + \vec{p} \times \vec{q}) \quad (6)$$

となる。ここで、 $\vec{p} \cdot \vec{q}$  と  $\vec{p} \times \vec{q}$  はそれぞれ3次元ベクトル  $\vec{p}$  と  $\vec{q}$  の内積と外積を表す。四元数の積の共役と共役な四元数の積の間には次の関係が成立する。

$$(p \otimes q)^* = q^* \otimes p^* \quad (7)$$

四元数  $\boldsymbol{x}$  のノルム  $|\boldsymbol{x}|$  は

$$|\boldsymbol{x}| = \sqrt{\boldsymbol{x} \otimes \boldsymbol{x}^*} = \sqrt{x^{(e)2} + x^{(i)2} + x^{(j)2} + x^{(k)2}} \quad (8)$$

と定義される。

スカラー  $a = (a, \vec{0})$  と四元数  $\boldsymbol{x}$  の間の積は次のようになる。

$$\begin{aligned} a\boldsymbol{x} &= (ax^{(e)}, a\vec{x}) \\ &= (ax^{(e)}, ax^{(i)}, ax^{(j)}, ax^{(k)}) \end{aligned} \quad (9)$$

## 2.2 四元数の位相表現

直交座標系における複素数  $\boldsymbol{c} = c^{(e)} + ic^{(i)}$  について極座標表示を行うと、振幅  $r$  と位相  $\theta$  を用いて、 $\boldsymbol{c} = r \cdot e^{i\theta}$  と表すことができる。ここで、 $r = \sqrt{c^{(e)2} + c^{(i)2}}$  および  $\theta = \tan^{-1} c^{(i)}/c^{(e)}$  である。

同様にして、直交座標系における四元数についても、対応する位相表現を定義することが可能である。四元数における位相表現はいくつか提案されているが、本論文では、文献<sup>15, 16)</sup>において用いられている表現方法を採用する。四元数  $\boldsymbol{x}$  は、振幅  $|\boldsymbol{x}|$  および3つの位相変数  $\varphi, \theta, \psi$  により表現される。ここで、これらの位相変数は、 $-\pi \leq \varphi < \pi$ ,  $-\pi/2 \leq \theta < \pi/2$ ,  $-\pi/4 \leq \psi \leq \pi/4$  を満たす。これらの変数を用いることにより、四元数  $\boldsymbol{x}$  は

$$\boldsymbol{x} = |\boldsymbol{x}| e^{i\varphi} \otimes e^{k\psi} \otimes e^{j\theta} \quad (10)$$

と表される。ここで、 $e^{i\varphi}, e^{j\theta}, e^{k\psi}$  は、

$$\begin{aligned} e^{i\varphi} &= \cos \varphi + i \sin \varphi, & e^{j\theta} &= \cos \theta + j \sin \theta, \\ e^{k\psi} &= \cos \psi + k \sin \psi \end{aligned} \quad (11)$$

である。

## 3 ホップフィールド型四元数ニューラルネットワーク

本節では、四元数に基づくニューロンモデルならびにホップフィールド型 NN を示す。これらのモデルでは、入出力値、閾値、結合荷重全てが四元数で表現されており、時間に関しては離散時間を採用している。

### 3.1 離散時間・離散状態モデル

まず、最も基本的なモデルである離散状態モデルを示す。四元数ニューロン  $p$  の出力値  $\boldsymbol{x}_p$  は次式で定義される。

$$\boldsymbol{s}_p(t) = \sum_q \boldsymbol{w}_{pq} \otimes \boldsymbol{x}_q(t) - \boldsymbol{\theta}_p \quad (12)$$

$$\boldsymbol{x}_p(t+1) = \boldsymbol{f}(\boldsymbol{s}_p(t)) \quad (13)$$

ここで、 $\boldsymbol{s}_p$  と  $\boldsymbol{\theta}_p$  はそれぞれニューロン  $p$  の活動電位と閾値を表し、 $\boldsymbol{x}_q$  と  $\boldsymbol{w}_{pq}$  はそれぞれニューロン  $q$  の出力値とニューロン  $q$  からニューロン  $p$  への結合荷重を表す。これらのパラメータは全て四元数である。このモデルでは、ニューロン  $p$  の出力値を決定する活性化関数  $\boldsymbol{f}$  として次式を導入する。

$$\boldsymbol{f}(\boldsymbol{s}) = f^{(e)}(s^{(e)}) + f^{(i)}(s^{(i)})\boldsymbol{i} + f^{(j)}(s^{(j)})\boldsymbol{j} + f^{(k)}(s^{(k)})\boldsymbol{k} \quad (14)$$

それぞれの成分における実活性化関数としては、次式に示される同一の符号関数を用いる。

$$\begin{aligned} f^{(e)}(s) &= f^{(i)}(s) = f^{(j)}(s) = f^{(k)}(s) \\ &= \begin{cases} 1 & \text{for } s \geq 0 \\ -1 & \text{for } s < 0 \end{cases} \end{aligned} \quad (15)$$

すなわち、このニューロンモデルでは、ニューロンの出力値における各成分は +1 または -1 の値を取る。この場合、各ニューロンは  $2^4 = 16$  通りの状態を取ることになる。

次にこのニューロンモデルを用いてホップフィールド型 NN を構成する。  $N$  ニューロンから構成されるネットワークのエネルギー関数は次式で与えられる。

$$\begin{aligned} E(t) &= -\frac{1}{2} \sum_{p=1}^N \sum_{q=1}^N \boldsymbol{x}_p^*(t) \otimes \boldsymbol{w}_{pq} \otimes \boldsymbol{x}_q(t) \\ &\quad + \frac{1}{2} \sum_{p=1}^N (\boldsymbol{\theta}_p^* \otimes \boldsymbol{x}_p(t) + \boldsymbol{x}_p^*(t) \otimes \boldsymbol{\theta}_p) \end{aligned} \quad (16)$$

結合荷重に関して  $\boldsymbol{w}_{pq} = \boldsymbol{w}_{qp}^*$  である場合、このエネルギー値は実数値をとる。さらに、自己結合  $\boldsymbol{w}_{pp}$  が実数 ( $\boldsymbol{w}_{pp} = \boldsymbol{w}_{pp}^* = (w_{pp}^{(e)}, \vec{0})$ ) かつ非負の値 ( $w_{pp}^{(e)} \geq 0$ ) を取る場合、このエネルギー値はネットワークの状態変化に対して単調減少する<sup>10, 11)</sup>。

### 3.2 離散時間・連続状態モデル

前節のニューロンモデルに対して、活性化関数を変更することにより連続状態モデルを作ることができる。ここでは、ニューロン  $p$  の出力値を決定する活性化関数  $\boldsymbol{f}$  として2種類の関数を導入する。1つは、入力された四元数の各成分に対して独立に実活性化関数を適用する方法であり、次式で定義される。

$$\boldsymbol{f}_1(\boldsymbol{s}) = f_1^{(e)}(s^{(e)}) + f_1^{(i)}(s^{(i)})\boldsymbol{i} + f_1^{(j)}(s^{(j)})\boldsymbol{j} + f_1^{(k)}(s^{(k)})\boldsymbol{k} \quad (17)$$

それぞれの成分における実活性化関数としては、次式に示す同一のシグモイド関数を用いる。

$$f_1^{(e)}(s) = f_1^{(i)}(s) = f_1^{(j)}(s) = f_1^{(k)}(s) = \tanh(s/\epsilon) \quad (18)$$

ここで、 $\epsilon > 0$  はシグモイド関数の傾きを決定する係数である。さらに、もう1つの活性化関数として、複素 NN における活性化関数として文献<sup>17)</sup>において提案された関数を四元数に拡張する。その関数は次式で定義される。

$$\boldsymbol{f}_2(\boldsymbol{s}) = \frac{a\boldsymbol{s}}{1 + |\boldsymbol{s}|} \quad (19)$$

ここで、 $a$  は実定数である。

前節と同様に、このニューロンモデルを用いてネットワークを構成する。  $N$  ニューロンから構成されるネットワークのエネルギー関数は次式で与えられる。

$$\begin{aligned} E(t) &= -\frac{1}{2} \sum_{p=1}^N \sum_{q=1}^N \boldsymbol{x}_p^*(t) \otimes \boldsymbol{w}_{pq} \otimes \boldsymbol{x}_q(t) \\ &\quad + \frac{1}{2} \sum_{p=1}^N (\boldsymbol{\theta}_p^* \otimes \boldsymbol{x}_p(t) + \boldsymbol{x}_p^*(t) \otimes \boldsymbol{\theta}_p) \end{aligned}$$

$$+ \sum_{p=1}^N G(\mathbf{x}_p(t)) \quad (20)$$

ここで、 $G(\mathbf{x}(t))$  は次式を満たす実関数である。

$$\frac{\partial G(\mathbf{x})}{\partial x^{(\alpha)}} = g^{(\alpha)}(x^{(e)}, x^{(i)}, x^{(j)}, x^{(k)}) \quad (\alpha = \{e, i, j, k\}) \quad (21)$$

また、 $g(\mathbf{x})$  は、 $f(\mathbf{x})$  の逆関数であり、次のように表される。

$$\begin{aligned} \mathbf{g}(\mathbf{x}) &= \mathbf{f}^{-1}(\mathbf{x}) \\ &= g^{(e)}(x^{(e)}, x^{(i)}, x^{(j)}, x^{(k)}) \\ &\quad + g^{(i)}(x^{(e)}, x^{(i)}, x^{(j)}, x^{(k)})\mathbf{i} \\ &\quad + g^{(j)}(x^{(e)}, x^{(i)}, x^{(j)}, x^{(k)})\mathbf{j} \\ &\quad + g^{(k)}(x^{(e)}, x^{(i)}, x^{(j)}, x^{(k)})\mathbf{k} \end{aligned} \quad (22)$$

この式は、式 (12) より、

$$\mathbf{s}_p(t) = \mathbf{f}^{-1}(\mathbf{x}_p(t+1)) = \mathbf{g}(\mathbf{x}_p(t+1)) \quad (23)$$

を満たす。前節のモデルと同様、結合荷重に関する条件  $w_{pq} = w_{qp}^*$  ならびに自己結合に関する条件 ( $w_{pp} = w_{pp}^* = (w_{pp}^{(e)}, \vec{0})$ ) を必要とする。

このネットワークにおいて、活性化関数として  $f_1$  を用いた場合は、 $w_{rr}^{(e)} > -2\epsilon$  を満たす時に、エネルギー値は状態変化に対して単調減少する。また、活性化関数として  $f_2$  の場合は、 $a > 0$  である時にエネルギー値は単調減少する。すなわち、これらの条件がこのネットワークの安定条件となる<sup>12)</sup>。

### 3.3 離散時間・多相状態モデル

多相ニューロンモデルにおけるニューロン  $p$  の状態  $\mathbf{u}_p$  は、4次元空間にある単位球面上の点として表現される。すなわち、振幅については  $|\mathbf{u}_p| = 1$  として固定され、次式の通り位相変数のみにより表現される。

$$\begin{aligned} \mathbf{u}_p &= e^{i\varphi_p} \otimes e^{k\psi_p} \otimes e^{j\theta_p} \\ &= \mathbf{q}^{(\varphi_p)} \otimes \mathbf{q}^{(\psi_p)} \otimes \mathbf{q}^{(\theta_p)} \end{aligned} \quad (24)$$

ここで、 $\mathbf{q}^{(\varphi)} = e^{i\varphi}$ 、 $\mathbf{q}^{(\psi)} = e^{k\psi}$ 、 $\mathbf{q}^{(\theta)} = e^{j\theta}$  の表現を導入している。時刻  $t$  におけるニューロン  $p$  の内部状態を  $\mathbf{h}_p(t)$  は次式として定義される。

$$\begin{aligned} \mathbf{h}_p(t) &= \sum_q w_{pq} \otimes \mathbf{u}_q(t) \\ &= \sum_q w_{pq} \otimes e^{i\varphi_q(t)} \otimes e^{k\psi_q(t)} \otimes e^{j\theta_q(t)} \\ &= \sum_q w_{pq} \otimes \mathbf{q}^{(\varphi_q)}(t) \otimes \mathbf{q}^{(\psi_q)}(t) \otimes \mathbf{q}^{(\theta_q)}(t) \end{aligned} \quad (25)$$

ここで、 $w_{pq} \in \mathbf{H}$  はニューロン  $q$  から  $p$  への結合荷重である。

このニューロンにおける活性化関数として、多相複素 NN<sup>18, 19)</sup> において用いられている活性化関数<sup>20)</sup> を四元数に拡張する。時刻  $(t+1)$  におけるニューロン  $p$  の状態を

$$\mathbf{u}_p(t+1) = qsign(\mathbf{h}_p(t)) \quad (26)$$

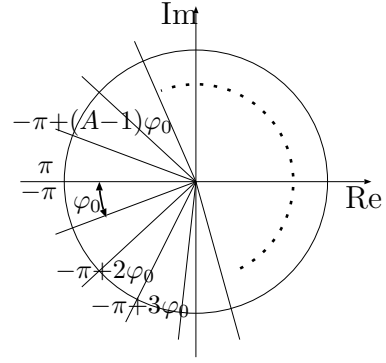


Fig. 1:  $A$  を量子化単位とした場合の複素多値活性化関数  $csign(\cdot)$

と定義する。ここで、

$$qsign(\mathbf{u}) = csign_A(\mathbf{q}^{(\varphi)}) \otimes csign_B(\mathbf{q}^{(\psi)}) \otimes csign_C(\mathbf{q}^{(\theta)}) \quad (27)$$

である。ニューロン状態の更新は、 $\mathbf{u}$  の各要素すなわち  $\mathbf{q}^{(\varphi)}$ 、 $\mathbf{q}^{(\psi)}$ 、 $\mathbf{q}^{(\theta)}$  ごとに行われる。四元数ニューロンにおける活性化関数  $qsign(\cdot)$  は、多相複素ニューロンにおいて用いられている  $csign(\cdot)$  関数の積として表現される。 $\mathbf{q}^{(\varphi)}$  を更新するために用いられる  $csign_A(\cdot)$  は、

$$csign_A(\mathbf{q}^{(\varphi)}) \equiv \begin{cases} e^{i(-\pi+0\cdot\varphi_0)} = -e^0 & \text{for } -\pi \leq \arg \mathbf{q}^{(\varphi)} < -\pi + \varphi_0 \\ -e^{i\varphi_0} & \text{for } -\pi + \varphi_0 \leq \arg \mathbf{q}^{(\varphi)} < -\pi + 2\varphi_0 \\ -e^{i2\varphi_0} & \text{for } -\pi + 2\varphi_0 \leq \arg \mathbf{q}^{(\varphi)} < -\pi + 3\varphi_0 \\ \vdots & \\ -e^{i(A-1)\varphi_0} & \text{for } -\pi + (A-1)\varphi_0 \leq \arg \mathbf{q}^{(\varphi)} < -\pi + A\varphi_0 \end{cases} \quad (28)$$

と定義する。ここで、 $\varphi_0$  は量子化単位であり、 $\varphi_0 = 2\pi/A$  を満たす。図 1 に  $csign_A$  関数を示す。この活性化関数により、ニューロンの部分状態 ( $\varphi$  成分) は  $A$  分割された量子化状態であると言える。

$\mathbf{q}^{(\psi)}$  を更新するための活性化関数  $csign_B(\cdot)$  および  $\mathbf{q}^{(\theta)}$  を更新するための関数  $csign_C(\cdot)$  についても同様に定義することができる。

$$csign_B(\mathbf{q}^{(\psi)}) \equiv \begin{cases} e^{k(-\frac{\pi}{4}+0\cdot\psi_0)} & \text{for } -\frac{\pi}{4} \leq \arg \mathbf{q}^{(\psi)} < -\frac{\pi}{4} + \psi_0 \\ \vdots & \\ e^{k(-\frac{\pi}{4}+(B-1)\cdot\psi_0)} & \text{for } -\frac{\pi}{4} + (B-1)\psi_0 \leq \arg \mathbf{q}^{(\psi)} \leq -\frac{\pi}{4} + B\psi_0 \end{cases} \quad (29)$$

$$c\text{sign}_C(\mathbf{q}^{(\theta)}) \equiv \begin{cases} e^{j(-\frac{\pi}{2}+0\cdot\theta_0)} \\ \text{for } -\frac{\pi}{2} \leq \arg \mathbf{q}^{(\theta)} \\ < -\frac{\pi}{2} + \theta_0 \\ \vdots \\ e^{j(-\frac{\pi}{2}+(C-1)\cdot\theta_0)} \\ \text{for } -\frac{\pi}{2} + (C-1)\theta_0 \\ \leq \arg \mathbf{q}^{(\theta)} < -\frac{\pi}{2} + C\theta_0 \end{cases} \quad (30)$$

$\psi_0 = \pi/2B$  および  $\theta = \pi/C$  は  $\psi, \theta$  成分の量子化単位であり、分割数はそれぞれ、 $B$  および  $C$  となる。

本研究でのニューロンにおいては、ニューロンの状態は非同期的に更新される。ある時刻において更新されるニューロンは1つのみであり、また、そのニューロンの位相状態についても全ての位相変数を同時には更新しないものとする。すなわち、時刻  $t$  においてニューロン  $r$  が更新される場合、その更新は次式にて表される。

$$\mathbf{u}_p(t+1) = \begin{cases} \mathbf{q}^{(\varphi_p)}(t) \otimes \mathbf{q}^{(\psi_p)}(t) \otimes \mathbf{q}^{(\theta_p)}(t) = \mathbf{u}_p(t) \\ \text{for } p \neq r \\ \mathbf{q}^{(\varphi_p)}(t) \otimes \mathbf{q}^{(\psi_p)}(t+1) \otimes \mathbf{q}^{(\theta_p)}(t+1) \\ \text{or} \\ \mathbf{q}^{(\varphi_p)}(t+1) \otimes \mathbf{q}^{(\psi_p)}(t+1) \otimes \mathbf{q}^{(\theta_p)}(t) \\ \text{for } p = r \end{cases} \quad (31)$$

$N$  ニューロンから構成されるネットワークのエネルギー関数は次式で与えられる。

$$E(t) = -\frac{1}{2} \sum_{p=1}^N \sum_{q=1}^N \mathbf{u}_p^*(t) \otimes \mathbf{w}_{pq} \otimes \mathbf{u}_q(t) \quad (32)$$

この関数について、 $\mathbf{w}_{pq} = \mathbf{w}_{qp}^*$  を満たす場合、エネルギー値は実数値となり、自己結合  $\mathbf{w}_{pp}$  が非負の実数値を取る必要がある。

ネットワークにおいてある1つのニューロンが状態変化を行い、その際のニューロンの  $\varphi, \psi, \theta$  成分の位相変化を  $\Delta\varphi, \Delta\psi, \Delta\theta$  とする。これらの位相変化が大きくない場合、すなわち、 $|\Delta\varphi| < \varphi_0, |\Delta\psi| < \psi_0, |\Delta\theta| < \theta_0$  という条件を満たす場合、エネルギーの変化量が非正値の値を取る。すなわち、この条件下において、ネットワークの状態変化にともないエネルギーは単調減少する<sup>13)</sup>。

## 4 ホップフィールド型四元数ニューラルネットワークにおける学習スキーム

### 4.1 Hebb 学習則

最初に射影学習の基礎となる Hebb 学習則について説明する。記録パターン  $\{\epsilon^\mu\}$  が与えられた時、ニューロン  $p$  と  $q$  の間の結合荷重を次式を用いて決定する。

$$\mathbf{w}_{pq} = \frac{1}{4N} \sum_{\mu=1}^{n_p} \epsilon_p^\mu \otimes \epsilon_q^{\mu*} \quad (33)$$

ここで、 $\epsilon_p^\mu$  は記録パターン  $\mu$  におけるニューロン  $p$  の状態であり、 $n_p$  は記録パターン数である。この式で計算される結合荷重は、 $\mathbf{w}_{pq} = \mathbf{w}_{qp}^*$  および  $\mathbf{w}_{pp} \geq 0$  の

条件を満たす。また、各記録パターンが互いに直交している場合において、それらはネットワーク状態における安定点となることが示されている<sup>11)</sup>。

### 4.2 射影学習則

前節の Hebb 学習則においては、各記名パターン  $\{\epsilon^\mu\}$  が互いに直交している、すなわち  $\mu, \nu = 1, \dots, n_p$  について、

$$\sum_{q=1}^N \epsilon_q^{\mu*} \otimes \epsilon_q^\nu = 4N \delta_{\mu,\nu} = 4N (\delta_{\mu,\nu}^{(e)}, \vec{0}),$$

が成立していなければならない。ただし、 $\delta_{\mu,\nu}^{(e)}$  は Kronecker delta を表す。

射影学習則<sup>21, 22, 23)</sup> は、上記条件を満たさない非直交パターンを直交パターンに射影することによりパターンの記録を行う。四元数 NN における射影学習則は、以下に示す行列  $\{\mathbf{Q}_{\mu\nu}\}$  を導入することにより射影を行う。

$$\mathbf{Q}_{\mu\nu} = \frac{1}{N} \sum_p \epsilon_p^{\mu*} \otimes \epsilon_p^\nu \quad (34)$$

これにより、ネットワークの結合荷重  $\tilde{\mathbf{w}}$  は、次式で計算される。

$$\tilde{\mathbf{w}}_{pq} = \frac{1}{N} \sum_{\nu,\mu} \epsilon_p^\mu \otimes (\mathbf{Q}^{-1})_{\mu\nu} \otimes \epsilon_q^{\nu*} \quad (35)$$

この学習則により形成される結合荷重を持つネットワークにおいても、Hebb 学習則と同様に各記録パターンが安定点となる。記録パターン  $\epsilon^\sigma$  が入力された場合におけるニューロン  $p$  の内部状態  $\tilde{\mathbf{h}}_p$  は、

$$\begin{aligned} \tilde{\mathbf{h}}_p &= \sum_{q=1}^N \tilde{\mathbf{w}}_{pq} \otimes \epsilon_q^\sigma \\ &= \frac{1}{N} \sum_{\mu,\nu} \epsilon_p^\mu \otimes (\mathbf{Q}^{-1})_{\mu\nu} \otimes \sum_q \epsilon_q^{\nu*} \otimes \epsilon_q^\sigma \\ &= \sum_{\mu,\nu} \epsilon_p^\mu \otimes (\mathbf{Q}^{-1})_{\mu\nu} \mathbf{Q}_{\nu\sigma} \\ &= \sum_{\mu} \epsilon_p^\mu \otimes (\mathbf{Q}^{-1} \mathbf{Q})_{\mu\sigma} \\ &= \sum_{\mu} \epsilon_p^\mu \delta_{\mu\sigma} \\ &= \epsilon_p^\sigma \end{aligned} \quad (36)$$

と表される。これにより、記録パターンが入力された場合、ニューロン状態が更新されても同じ記録パターン状態になることから、記録パターンはネットワーク状態における安定点であることが示される。

### 4.3 局所逐次学習法

局所逐次学習法<sup>24)</sup> は、前節で説明した射影学習則の一実装法である。この学習法では誤差逆伝搬法などの学習法と同様に、記録パターンを繰り返し提示する



ことにより結合荷重を徐々に形成してゆく。局所逐次学習法は、射影学習則において必要とされた疑逆行列  $Q^{-1}$  を計算する必要がなく、また各記録パターンに対する引き込み領域の大きさを制御できるという特徴がある。

四元数ホップフィールド型 NN における局所逐次学習法においては、ネットワークの結合荷重は以下の式により更新される。

$$\mathbf{w}_{pq}^{new} = \mathbf{w}_{pq}^{old} + \delta \mathbf{w}_{pq}, \quad (37)$$

$$\delta \mathbf{w}_{pq} = \frac{1}{4N} \epsilon_p^\mu \otimes \epsilon_q^{\mu*} \quad (38)$$

この学収則においても、結合荷重の更新を繰り返し行うことにより、記録パターンがネットワークにおいて安定な状態となる (詳細な証明については文献<sup>14)</sup>を参照)。

## 5 まとめ

本論文では、四元数に基づくホップフィールド型ニューラルネットワークモデルを紹介した。これらのモデルは離散時間型のモデルであり、それぞれのモデルに対して、ネットワーク中のニューロンの状態変化にともない、ネットワークのエネルギーが単調減少することを示した。しかし、記録パターン状態の引き込み領域や安定度をはじめとする特性について多くの調査すべき点が残されている。四元数ホップフィールドネットワークについて、想起特性などを詳細に調べ、さらに画像処理などへの応用を図ることが今後の課題として挙げられる。

**謝辞** 本研究の一部は日本学術振興会の科学研究補助金 (若手研究 (B)21700259 および基盤研究 (C)23500286) の支援により行われた。

## 参考文献

- 1) A. Hirose, editor: *Complex-Valued Neural Networks: Theories and Application, Innovative Intelligence*, **5**, World Scientific Publishing (2003)
- 2) T. Nitta: A Solution to the 4-bit Parity Problem with a Single Quaternary Neuron, *Neural Information Processing - Letters and Reviews*, **5-2**, 33/39 (2004)
- 3) P. Arena, L. Fortuna, G. Muscato, and M. G. Xibilia: *Neural Networks in Multidimensional Domains, Lecture Notes in Computer Science*, **234**, Springer-Verlag (1998)
- 4) T. Nitta: An Extension of the Back-propagation Algorithm to Quaternions, In *Proceedings of International Conference on Neural Information Processing (ICONIP'96)*, **1**, 247/250 (1996)
- 5) N. Matsui, T. Isokawa, H. Kusamichi, F. Peper, and H. Nishimura: Quaternion Neural Network with Geometrical Operators, *Journal of Intelligent & Fuzzy Systems*, **15-3-4** 149/164 (2004)
- 6) H. Kusamichi, T. Isokawa, N. Matsui, Y. Ogawa, and K. Maeda: A New Scheme for Color Night Vision by Quaternion Neural Network, In *Proceedings of the 2nd International Conference on Autonomous Robots and Agents (ICARA2004)*, 101/106 (2004)
- 7) T. Isokawa, N. Matsui, and H. Nishimura: Quaternionic Neural Networks: Fundamental Properties and Applications, In T. Nitta, editor, *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters*, chapter XVI, 411/439, Information Science Reference (2009)
- 8) B. C. Ujang, C. C. Took, and D. P. Mandic: Quaternion-valued nonlinear adaptive filtering, *IEEE Transactions on Neural Networks*, **22-8**, 1193/1206 (2011)
- 9) M. Yoshida, Y. Kuroe, and T. Mori: Models of Hopfield-type Quaternion Neural Networks and Their Energy Functions, *International Journal of Neural Systems*, **15-1-2**, 129/135 (2005)
- 10) T. Isokawa, H. Nishimura, N. Kamiura, and N. Matsui: Fundamental Properties of Quaternionic Hopfield Neural Network, In *Proceedings of 2006 International Joint Conference on Neural Networks*, 610/615 (2006)
- 11) T. Isokawa, H. Nishimura, N. Kamiura, and N. Matsui: Associative Memory in Quaternionic Hopfield Neural Network, *International Journal of Neural Systems*, **18-2**, 135/145 (2008)
- 12) T. Isokawa, H. Nishimura, N. Kamiura, and N. Matsui: Dynamics of Discrete-Time Quaternionic Hopfield Neural Networks, In *Proceedings of 17th International Conference on Artificial Neural Networks*, 848/857 (2007)
- 13) T. Isokawa, H. Nishimura, A. Saitoh, N. Kamiura, and N. Matsui: On the Scheme of Quaternionic Multistate Hopfield Neural Network, In *Proceedings of Joint 4th International Conference on Soft Computing and Intelligent Systems and 9th International Symposium on advanced Intelligent Systems (SCIS & ISIS 2008)*, 809/813 (2008)
- 14) T. Isokawa, H. Nishimura, and N. Matsui: An Iterative Learning Scheme for Multistate Complex-Valued and Quaternionic Hopfield Neural Networks, In *Proceedings of International Joint Conference on Neural Networks (IJCNN2009)*, 1365/1371 (2009)
- 15) T. Bülow: *Hypercomplex Spectral Signal Representations for the Processing and Analysis of Images*, PhD thesis, Christian-Albrechts-Universität zu Kiel (1999)
- 16) T. Bülow and G. Sommer: Hypercomplex Signals—A Novel Extension of the Analytic Signal to the Multi-dimensional Case, *IEEE Transactions on Signal Processing*, **49-11**, 2844/2852 (2001)
- 17) G. M. Georgiou and C. Koutsougeras: Complex domain backpropagation, *IEEE Transactions on Circuits and Systems II*, **39-5**, 330/334 (1992)
- 18) N. N. Aizenberg, Yu. L. Ivaskiv, and D. A. Pospelov: About one generalization of the threshold function, *Doklady Akademii Nauk SSSR (The Reports of the Academy of Sciences of the USSR)*, **196-6**, 1287/1290 (1971) (in Russian)
- 19) I. N. Aizenberg, N. N. Aizenberg, and J. Vandewalle: *Multi-Valued and Universal Binary Neurons - Theory, Learning and Applications -*, Kluwer Academic Publishers (2000)
- 20) S. Jankowski, A. Lozowski, and J. M. Zurada: Complex-Valued Multistate Neural Associative Memory, *IEEE Transactions on Neural Networks*, **7-6**, 1491/1496 (1996)
- 21) T. Kohonen: *Self-Organization and Associative Memory*, Springer (1984)
- 22) L. Personnaz, I. Guyon, and G. Dreyfus: Collective Computational Properties of Neural Networks: New Learning Mechanisms, *Phys. Rev. A*, **34**, 4217/4228 (1986)
- 23) Dong-Liang Lee: Improvements of complex-valued hopfield associative memory by using generalized projection rules, *IEEE Transaction on Neural Networks*, **17-5**, 1341/1347 (2006)
- 24) S. Diederich and M. Opper: Learning of Correlated Patterns in Spin-Glass Networks by Local Learning Rules, *Phys. Rev. Lett.*, **58**, 949/952 (1987)



# 高次元信号に対する広域線形推定法

○新田徹 (産業技術総合研究所)

## Widely Linear Estimation for High-Dimensional Signals

\*T. Nitta (National Institute of Advanced Industrial Science and Technology (AIST))

**Abstract**— In this paper, we formulate a Clifford-valued widely linear estimation framework. Clifford number is a hypercomplex number that generalizes real, complex numbers, quaternions, and higher dimensional numbers. And also, as a first step, we will give a theoretical foundation for a quaternion-valued widely linear estimation framework. The estimation error obtained with the quaternion-valued widely linear estimation method is proven to be smaller than that obtained using the usual quaternion-valued linear estimation method.

**Key Words:** Signal processing, Quaternion, Estimation

### 1 はじめに

広域線形推定法 (widely linear estimation method) は、複素値データを使った推定問題に有効であることが数理的に証明されている<sup>1)</sup>。広域線形推定法では、複素パラメータのみならず、その複素共役パラメータをも使用する。そのことは、Picinbonoらによって導入された、いわゆる、拡張複素統計量 (augmented complex statistics) を使うことを意味する。現在までに、通信<sup>2,3)</sup> や適応フィルタ<sup>4)</sup> などに適用されている。

広域線形推定法は、さらに4元数の場合に拡張されている<sup>5,6)</sup>。それはすべての統計量を利用した4元数データに対する推定法となっている。4元数は複素数を拡張した4次元の数であり、W. R. Hamiltonにより1843年に発見された。4元数は今までにロボット工学、コンピュータ・ビジョン、ニューラルネットワーク、信号処理、通信などの分野に応用された (たとえば、文献<sup>7,8)</sup>)。

本稿では、クリフォード数信号を対象とした広域線形推定法を定式化する。また、4元数版の広域線形推定法の数理的基礎を与える。つまり、4元数版の広域線形推定法により得られた推定誤差は、通常の4元数線形推定法を用いて得られた推定誤差よりも小さいことを証明する。

### 2 クリフォード代数

本章では、クリフォード代数 (幾何代数とも呼ばれる) について簡単に述べる<sup>9)</sup>。

クリフォード代数  $Cl_{p,q}$  は複素数体、4元数体を高次元に拡張したものであり、 $2^n$  個の基底を持つ。添数  $p, q$  は、 $p+q=n$  を満たし、クリフォード代数の性質を規定する。クリフォード代数では、一般に乗法は非可換である。

$n=2, p=0, q=2$  の場合、基底の数は  $2^2=4$  であり、 $Cl_{0,2}$  は4元数体に対応する。クリフォード代数を理解するためには、4元数が役に立つかもしれない。4元数は  $\mathbf{R}$  上で定義され、次の式を満たす3つ組  $i, j, k$  から成る虚部を持つ：

$$\begin{aligned} ij &= -ji = k, \quad jk = -kj = i, \quad ki = -ik = j, \\ i^2 &= j^2 = k^2 = ijk = -1. \end{aligned} \quad (1)$$

ここで、 $\mathbf{R}$  は実数の集合である。4元数  $q$  は次のよう

に書ける：

$$q = a + bi + cj + dk \in \mathbf{Q}, \quad a, b, c, d \in \mathbf{R}. \quad (2)$$

ここで、 $\mathbf{Q}$  は4元数の集合を表わす。4元数  $q$  の共役4元数は

$$q^* = a - bi - cj - dk \quad (3)$$

で定義される。また、4元数  $q \in \mathbf{Q}$  のノルムは、

$$\begin{aligned} |q| &= \sqrt{qq^*} \\ &= \sqrt{a^2 + b^2 + c^2 + d^2}. \end{aligned} \quad (4)$$

で与えられる。一般に、任意の4元数  $p, q \in \mathbf{Q}$  に対して、 $pq \neq qp$  である。

次にクリフォード代数について述べる。 $\mathbf{R}^{n+1}$  を基底  $\{e_0, \dots, e_n\}$  を持つ空間とし、 $p \in \{0, \dots, n\}, q \stackrel{\text{def}}{=} n-p$  とする。また、乗法に関して次の規則が成り立つと仮定する。

$$e_0 e_i = e_i e_0 = e_i \quad (i = 1, \dots, n), \quad (5)$$

$$e_i e_j = -e_j e_i \quad (i \neq j; i, j = 1, \dots, n), \quad (6)$$

$$e_0^2 = e_1^2 = \dots = e_p^2 = 1, \quad (7)$$

$$e_{p+1}^2 = \dots = e_{p+q}^2 = -1. \quad (8)$$

このとき、クリフォード代数  $Cl_{p,q}$  の  $2^n$  個の基底

$$\begin{aligned} e_0; e_1, \dots, e_n; e_1 e_2, \dots, e_{n-1} e_n; e_1 e_2 e_3, \dots; \\ \dots; e_1 e_2 \dots e_n \end{aligned} \quad (9)$$

が得られる。ただし、 $e_0$  は単位元である。

加法および実数との乗法は成分毎に行われる。たとえば、 $x = a_3 e_3 + a_9 e_9 \in Cl_{9,3}$  と  $y = b_9 e_9 + b_{32} e_3 e_2 \in Cl_{9,3}$  に対して、

$$\begin{aligned} x + y &= (a_3 e_3 + a_9 e_9) + (b_9 e_9 + b_{32} e_3 e_2) \\ &= a_3 e_3 + (a_9 + b_9) e_9 + b_{32} e_3 e_2 \end{aligned} \quad (10)$$

であり、 $\beta \in \mathbf{R}$  と  $x = a_{14} e_1 e_4 + a_{361} e_3 e_6 e_1 \in Cl_{9,3}$  に対して、

$$\begin{aligned} \beta x &= \beta (a_{14} e_1 e_4 + a_{361} e_3 e_6 e_1) \\ &= \beta a_{14} e_1 e_4 + \beta a_{361} e_3 e_6 e_1 \end{aligned} \quad (11)$$

である。さらに、次の条件が成り立つと仮定する。

$$e_1 e_2 \cdots e_n \neq \pm 1 \quad \text{if } p - q \equiv 1 \pmod{4}. \quad (12)$$

このようにして得られた代数をクリフォード代数  $Cl_{p,q}$  と呼ぶ。

クリフォード代数  $Cl_{p,q}$  において共役は次のように定義される。まず、任意の  $x \in Cl_{p,q}$  を

$$x = [x]_0 + [x]_1 + \cdots + [x]_n \quad (13)$$

と書く。ここで、

$$[x]_k \stackrel{\text{def}}{=} \sum_{\substack{A \in P_n \\ |A|=k}} x_A e_A, \quad (14)$$

$P_n$  は  $\{1, \dots, n\}$  のすべての部分集合から成る集合、

$$e_{i_1 i_2 \cdots i_p} \stackrel{\text{def}}{=} e_{i_1} e_{i_2} \cdots e_{i_p}, \quad (15)$$

式(15)の左辺における  $e_{i_1 i_2 \cdots i_p}$  の添数  $i_1 i_2 \cdots i_p$  は、集合  $\{i_1, i_2, \dots, i_p\}$  を意味する。このとき、任意の  $x \in Cl_{p,q}$  に対して、そのクリフォード共役  $x^* \in Cl_{p,q}$  は次のように与えられる。

$$x^* = [x]_0 - [x]_1 - [x]_2 + [x]_3 + [x]_4 - \cdots, \quad (16)$$

つまり、

$$([x]_k)^* = [x]_k \quad \text{for } k \equiv 0, 3 \pmod{4}, \quad (17)$$

$$([x]_k)^* = -[x]_k \quad \text{for } k \equiv 1, 2 \pmod{4}. \quad (18)$$

### 3 広域線形推定モデル

本章では、複素広域線形推定モデルと4元数広域線形推定モデルを述べた後に、クリフォード広域線形推定モデルを定式化する。

#### 3.1 複素広域線形推定モデル

$y \in \mathcal{C}$  を複素確率変数、 $\mathbf{x} \in \mathcal{C}^N$  を複素確率ベクトルとし、 $\mathbf{x}$  を観測して、 $y$  を推定するという問題を考える。ここで、 $\mathcal{C}$  は複素数の集合、 $N$  は自然数の集合である。つまり、 $y$  は真の値、 $\mathbf{x}$  は観測値を表わしている。

複素線形平均自乗推定 (Complex-Valued Linear Mean Square Estimation, 複素 LMSE) の枠組みにおいては、

$$\hat{y}_L = \mathbf{h}^H \mathbf{x}, \quad (19)$$

という形の推定値を求める。ここで、 $\mathbf{h} \in \mathcal{C}^N$  であり、 $H$  は複素共役転置を表わす。このとき、問題の目的は、推定誤差  $E|y - \hat{y}_L|^2$  を最小にするようなパラメータ  $\mathbf{h} \in \mathcal{C}^N$  を見つけることである。

一方、複素広域線形推定 (Complex-Valued Widely Linear Mean Square Estimation, 複素 WLMSE) の枠組みにおける問題は次のように書ける。

$$\hat{y} = \mathbf{h}^H \mathbf{x} + \mathbf{g}^H \mathbf{x}^*, \quad (20)$$

と定義された推定値  $\hat{y} \in \mathcal{C}$  を考える。ここで、 $\mathbf{g}, \mathbf{h} \in \mathcal{C}^N$  であり、 $v^* \stackrel{\text{def}}{=} a - bi$  は  $v = a + bi \in \mathcal{C}$  の複素共役

である。このとき、問題の目的は、推定誤差  $E|y - \hat{y}|^2$  を最小にするようなパラメータ  $\mathbf{g}, \mathbf{h} \in \mathcal{C}^N$  を求めることである。

Picinbono らは、複素 WLMSE の数理的な基礎を与えた<sup>1)</sup>。つまり、複素 WLMSE により得られる推定誤差は、通常の複素 LMSE により得られる推定誤差よりも小さいことを証明した： $E|y - \hat{y}_L|^2 \geq E|y - \hat{y}|^2$ 。ここで、等号は例外的な場合にだけ成り立つ。

#### 3.2 4元数広域線形推定モデル

4元数広域線形推定モデルは、3.1節で述べた複素 WLMSE モデルの自然な拡張である。 $y \in \mathcal{Q}$  を真の値を表わす4元数値確率変数、 $\mathbf{x} \in \mathcal{Q}^N$  を観測値を表わす4元数値確率ベクトルとする。

4元数線形平均自乗推定 (Quaternion-Valued Linear Mean Square Estimation, 4元数 LMSE) の枠組みにおいては、

$$\hat{y}_L = \mathbf{h}^H \mathbf{x}, \quad (21)$$

という形の推定値を求める。ここで、 $\mathbf{h} \in \mathcal{Q}^N$ 。  $N$  は自然数、 $H$  は4元数共役転置である。

4元数広域線形平均自乗推定 (Quaternion-Valued Widely Linear Mean Square Estimation, 4元数 WLMSE) の枠組みは次のとおりである。まず、

$$\hat{y} = \mathbf{h}^H \mathbf{x} + \mathbf{g}^H \mathbf{x}^*, \quad (22)$$

なる推定値  $\hat{y} \in \mathcal{Q}$  を考える。ここで、 $\mathbf{g}, \mathbf{h} \in \mathcal{Q}^N$ 、 $N$  は自然数、 $H$  は4元数共役転置を表わし、 $v^* \stackrel{\text{def}}{=} a - bi - cj - dk$  は  $v = a + bi + cj + dk \in \mathcal{Q}$  の4元数共役である。このとき、問題の目的は、 $E|y - \hat{y}|^2$  を最小化するパラメータ  $\mathbf{g}, \mathbf{h} \in \mathcal{Q}^N$  を求めることである。

Took と Mandic は、4元数 WLMSE に基づいて、4元数適応フィルターに対する拡張4元数最小平均自乗 (augmented quaternion least mean squares, AQLMS) アルゴリズムを導出し、ローレンツ・アトラクター、実世界風予測、データ・フージョンに関するコンピュータ・シミュレーションによってその有効性を確かめた<sup>5)</sup>。つまり、コンピュータ実験によって、4元数 WLMSE が通常の4元数 LMSE に比べて優れていることが確かめられた。しかしながら、4元数 WLMSE による推定誤差が通常の4元数 LMSE よりも優れているとの数理的な証明はこれまでに行われていない。複素数の場合にはそのような数理的証明は Picinbono らによって既に行われた<sup>1)</sup>。つまり、複素 WLMSE による推定誤差は、通常の複素 LMSE の推定誤差よりも小さいことが数理的に証明された。

#### 3.3 クリフォード広域線形推定モデル

本節では、複素広域線形推定モデルと4元数広域線形推定モデルの一般化である、クリフォード広域線形推定モデルを定式化する。

$y \in Cl_{p,q}$  をクリフォード数値確率変数、 $\mathbf{x} \in Cl_{p,q}^N$  をクリフォード数値確率ベクトルとする。ここで、 $N$  は自然数である。観測された  $\mathbf{x}$  から真の値  $y$  を推定することを考える。クリフォード線形平均自乗推定 (Clifford-Valued Linear Mean Square Estimation, クリフォード LMSE) の枠組において、問題は

$$\hat{y}_L = \mathbf{h}^H \mathbf{x}, \quad (23)$$

という形の推定値を求めることである。ここで、 $\mathbf{h} \in Cl_{p,q}^N$  であり、 $H$  はクリフォード共役転置である。

一方、クリフォード広域平均自乗推定 (Clifford-Valued Widely Linear Mean Square Estimation, クリフォード WLMSE) は、4 元数 WLMSE の自然な拡張として、次のように定式化できる。推定値  $\hat{y}$  を次のように定義する。

$$\hat{y} = \mathbf{h}^H \mathbf{x} + \mathbf{g}^H \mathbf{x}^*, \quad (24)$$

ここで、 $\mathbf{g}, \mathbf{h} \in Cl_{p,q}^N$  であり、 $v^*$  は  $v \in Cl_{p,q}$  のクリフォード共役である。このとき、問題の目的は、 $E|y - \hat{y}|^2$  を最小化するようなパラメータ  $\mathbf{g}, \mathbf{h} \in Cl_{p,q}^N$  を求めることである。

#### 4 4 元数 WLMSE の数理的基礎

本節では、3.3 節で定式化したクリフォード WLMSE の性質を調べる第 1 歩として、4 元数 WLMSE が通常の 4 元数 LMSE よりも良い結果をもたらすことを数理的に示す。主要な結果は次のとおりである：4 元数 WLMSE を使った場合の推定誤差は、例外的な場合を除いて、通常の 4 元数 LMSE を使った場合の推定誤差よりも小さい。この結果を得るのに、文献<sup>1)</sup>と同様の方法を用いた。ただし、4 元数の乗算は非可換であることを考慮する必要があった (任意の  $x, y \in \mathbf{Q}$  に対して、一般に、 $xy \neq yx$  である)。

まず、

$$X \stackrel{\text{def}}{=} \left\{ Z(\omega) = \mathbf{h}^H \mathbf{x}(\omega) + \mathbf{g}^H \mathbf{x}^*(\omega) \mid \mathbf{g}, \mathbf{h} \in \mathbf{Q}^N \right\}. \quad (25)$$

と定義する。 $X$  は 4 元数確率変数から成る集合であり、線型空間である。そして、内積  $\langle z, w \rangle \stackrel{\text{def}}{=} E[zw^*]$  による 4 元数値ヒルベルト空間  $Y = \{z(\omega) \in \mathbf{Q}\}$  のヒルベルト部分空間である。このとき、真の値  $y \in Y$ 、観測値  $\mathbf{x} \in \mathbf{Q}^N$ 、推定値  $\hat{y} \in X$  (式 (22)) に対して、次の式が成り立つ：

$$(y - \hat{y}) \perp \mathbf{x}, \quad (26)$$

$$(y - \hat{y}) \perp \mathbf{x}^*, \quad (27)$$

ここで、 $u \perp v$  は、 $v$  のすべての要素が  $u$  と内積  $\langle \cdot, \cdot \rangle$  に関して直交していることを意味する ( $u \in \mathbf{Q}, v \in \mathbf{Q}^N$ )。式 (26) と式 (27) から、次の式が得られる。

$$E[\mathbf{x}y^*] = E[\mathbf{x}\hat{y}^*], \quad (28)$$

$$E[\mathbf{x}^*y^*] = E[\mathbf{x}^*\hat{y}^*]. \quad (29)$$

よって、式 (22)、式 (28)、式 (29) から、次の式が成り立つことがわかる。

$$\Gamma_1 \mathbf{h} + C\mathbf{g} = r, \quad (30)$$

$$C^H \mathbf{h} + \Gamma_2 \mathbf{g} = s^*, \quad (31)$$

ここで、 $\Gamma_1 \stackrel{\text{def}}{=} E[\mathbf{x}\mathbf{x}^H], \Gamma_2 \stackrel{\text{def}}{=} E[\mathbf{x}^*\mathbf{x}^T], C \stackrel{\text{def}}{=} E[\mathbf{x}\mathbf{x}^T], r \stackrel{\text{def}}{=} E[\mathbf{x}y^*], s \stackrel{\text{def}}{=} E[y\mathbf{x}]$  である。そして、式 (30) と式 (31) から、

$$\mathbf{g} = (\Gamma_2 - C^H \Gamma_1^{-1} C)^{-1} \cdot (s^* - C^H \Gamma_1^{-1} r), \quad (32)$$

$$\mathbf{h} = (\Gamma_1 - C \Gamma_2^{-1} C^H)^{-1} \cdot (r - C \Gamma_2^{-1} s^*). \quad (33)$$

が得られる。

このとき、4 元数 WLMSE による推定誤差  $\varepsilon_{WL}$  は、式 (22)、式 (30)、式 (31) から、

$$\begin{aligned} \varepsilon_{WL}^2 &\stackrel{\text{def}}{=} E|y - \hat{y}|^2 \\ &= E|y|^2 - (r^H \mathbf{h} + s^T \mathbf{g}) \end{aligned} \quad (34)$$

となる。また、4 元数 LMSE による推定誤差  $\varepsilon_L$  は、式 (21) を使って、

$$\begin{aligned} \varepsilon_L^2 &\stackrel{\text{def}}{=} E|y - \hat{y}_L|^2 \\ &= E|y|^2 - r^H \Gamma_1^{-1} r \end{aligned} \quad (35)$$

であることが分かる。このとき、4 元数 WLMSE による推定誤差と 4 元数 LMSE による推定誤差の差  $\delta\varepsilon^2$  は、式 (30)、式 (32)、式 (34)、式 (35) から、

$$\begin{aligned} \delta\varepsilon^2 &\stackrel{\text{def}}{=} \varepsilon_L^2 - \varepsilon_{WL}^2 \\ &= (s^* - C^H \Gamma_1^{-1} r)^H \cdot (\Gamma_2 - C^H \Gamma_1^{-1} C)^{-1} \\ &\quad \cdot (s^* - C^H \Gamma_1^{-1} r) \end{aligned} \quad (36)$$

と計算される。 $\Gamma_2 - C^H \Gamma_1^{-1} C$  は非負値行列だから、式 (36) は非負である。さらに、式 (36) は、次の条件のうちのいずれかが成り立つ時にだけ 0 となる。

$$s^* - C^H \Gamma_1^{-1} r = \mathbf{0}, \quad (37)$$

$$\hat{y} = y. \quad (38)$$

式 (37) は例外的な場合であり、式 (38) は真の値  $y$  が確率 1 で推定されたことを意味する (これは滅多に起こらない)。

#### 5 結論

本稿では、クリフォード広域線形推定モデルを定式化した。また、4 元数広域線形推定法の数理的基礎を与えた。つまり、4 元数広域線形推定法により得られる推定誤差は、例外的な場合を除いて、通常の 4 元数線形推定法により得られる推定誤差よりも厳密に小さいことを証明した。今後は、クリフォード広域線形推定法の解析を進めていく予定である。

謝辞 質問に快く答えて下さった B. Picinbono 教授 (the Laboratoire des Signaux et Systèmes, Supélec, Gifsur Yvette, France) に感謝します。

#### 参考文献

- 1) Picinbono, B., Chevalier, P.: Widely Linear Estimation with Complex Data. IEEE Trans. Signal Processing, 43(8), 2030/2033 (1995)
- 2) Gerstacker, H., Schober, R., Lampe, A.: Receivers with Widely Linear Processing for Frequency-Selective Channels. IEEE Trans. Communications, 51(9), 1512/1523 (2003)
- 3) Schober, R., Gerstacker, W. H., Lampe, L. H.-J.: Data-Aided and Blind Stochastic Gradient Algorithms for Widely Linear MMSE MAI Suppression for DS-CDMA. IEEE Trans. Signal Processing, 52(3), 746/756 (2004)

- 4) Mandic, D. P., Goh, V. S. L.: Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models. John Wiley and Sons Ltd (2009)
- 5) Took, C. C., Mandic, D. P.: The Quaternion LMS Algorithm for Adaptive Filtering of Hypercomplex Processes. *IEEE Trans. Signal Processing*, 57(4), 1316/1327 (2009)
- 6) Took, C. C., Mandic, D. P.: 'A quaternion widely linear adaptive filter', *IEEE Trans. Signal Processing* 58(8), 4427/4431 (2010).
- 7) Nitta, T.: A Quaternary Version of the Back-Propagation Algorithm. In *Proc. IEEE Int. Conf. on Neural Networks, ICNN'95-Perth*, Nov. 27-Dec. 1, Vol.5, 2753/2756 (1995)
- 8) Arena, P., Fortuna, L., Muscato, G., Xibilia, M. G.: *Neural Networks in Multidimensional Domains. Lecture Notes in Control and Information Sciences 234*, Springer (1998)
- 9) Gürlebeck, K., Habetha, K, Sprößig, W.: *Holomorphic Functions in the Plane and N-Dimensional Space*. Birkhäuser (2008).

# リカレントクリフォードニューラルネットワークのモデルとダイナミクス

○黒江康明 (京都工芸繊維大学)

## Models of Recurrent Clifford Neural Networks and Their Dynamics

\*Y. Kuroe (Kyoto Institute of Technology)

**Abstract**— Recently, models of neural networks in the real domain have been extended into the high dimensional domain such as the complex number domain and quaternion number domain, and several high-dimensional models have been proposed. These extensions are generalized by introducing Clifford algebra (geometric algebra). In this paper we extend conventional real-valued models of recurrent neural networks into the domain defined by Clifford algebra and discuss their dynamics. We present models of fully connected recurrent neural networks, which are extensions of the real-valued Hopfield type neural networks to the domain defined by Clifford algebra. We study dynamics of the models from the point view of existence conditions of an energy function. We derive existence conditions of an energy function for some classes of the Hopfield type Clifford neural networks.

**Key Words:** Clifford algebra, Recurrent neural network, Hopfield neural network, Dynamics, Energy function

### 1 はじめに

近年、複素数(複素信号)を扱う事ができる複素数値化されたニューラルネットワーク(以後 NN)のモデルがいくつか提案され、その情報処理能力、学習法や応用などに関する研究が盛んに行われている<sup>1,2)</sup>。また最近では、さらに高次元化された四元数、八元数などを用いた NN のモデルも提案され、活発に研究が行われている。これら高次元の表現は、Clifford algebra (あるいは geometric algebra) 表現として一般化される。近年、Clifford algebra は様々な幾何学的問題の強力なかつ実用的な表現や解法を与えるものとして非常に注目されており、様々なサイエンスや工学の分野への応用の研究が盛んになってきている<sup>14,15)</sup>。そのため、Clifford algebra 表現を用いたニューロコンピューティング<sup>3)</sup>においても、豊かな表現能力を持ち高度な計算知能を実現できるものとして非常に期待されている。本稿は、Clifford algebra 表現を用いたリカレント NN のモデルを示し、そのダイナミクスについて議論する。

一般に NN の性質を決めるのは、その結合重みと非線形関数である活性化関数である。特に高次元 NN、すなわち、Clifford algebra 表現された NN においては、非線形関数に多様性があるため、活性化関数の選び方により様々なモデルが考えられ、多彩な性質を持つモデルが構築できる。特にリカレント NN は、本質的にダイナミカルシステムとしての振る舞いが問題となるため、活性化関数の選び方によりどのようなダイナミクスを示すかが興味をもたれ、それを調べることにより様々な新たな応用が考えられ、応用範囲が広がると期待できる。そこで本稿では、高次元リカレント NN として、Clifford algebra 表現を用いた 3 種類の Hopfield 型連続時間 NN のモデルを示す。またそのダイナミクスとして、エネルギー関数の存在条件について議論する。筆者は、これまで複素数、四元数などで表現される Hopfield 型 NN のエネルギー関数の存在条件とその応用に関する研究を行ってきた<sup>8,9,10,11)</sup>。本稿ではそれら結果を基にして、いくつかのクラスの Clifford algebra 表現された NN のエネルギー関数の存在条件について議論する<sup>12,13)</sup>。実際にとりあげるのは、二次元

の 3 つのクラスの Hopfield 型 NN と 4 次元の 1 つのクラスの Hopfield 型 NN である。前者の 3 つのクラスの表現は、それぞれ hyperbolic 数、複素数、dual 数表現と同型な表現となっている。また後者の 1 つクラスの表現は、四元数表現と同型な表現となっている。

### 2 Clifford Algebra 表現

本稿では、実数体  $\mathbb{R}$  上で定義される有限次元の Clifford algebra (geometric algebra) を対象とする。本節ではその概要を説明する<sup>1)</sup>。

#### 2.1 定義

$\mathbb{R}^{p,q,r}$  を実数体  $\mathbb{R}$  上の  $(p+q+r)$  次元のベクトル空間とし、 $\mathbb{R}^{p,q,r}$  に二次形式(可換なスカラー積)  $*$ :  $\mathbb{R}^{p,q,r} \times \mathbb{R}^{p,q,r} \rightarrow \mathbb{R}$  が定義されているとする。また  $\mathbb{R}^{p,q,r}$  の基底の集合を  $\mathbb{R}^{p,q,r}$  とすると、 $\mathbb{R}^{p,q,r}$  は次のように定義される。

$$\mathbb{R}^{p,q,r} := \{e_1, \dots, e_p, e_{p+1}, \dots, e_{p+q}, e_{p+q+1}, \dots, e_{p+q+r}\} \subset \mathbb{R}^{p,q,r} \quad (1)$$

ここで  $\{e_i\}$  は次の性質を持つとする<sup>2)</sup>。

$$e_i * e_j = \begin{cases} +1, & 1 \leq i = j \leq p, \\ -1, & p < i = j \leq p+q, \\ 0, & p+q < i = j \leq p+q+r, \\ 0, & i \neq j \end{cases} \quad (2)$$

このように、二次形式(スカラー積)を持つベクトル空間は二次空間(quadratic space)と呼ばれる。

$\mathbb{R}^{p,q,r}$  上の Clifford algebra を  $\mathbb{G}(\mathbb{R}^{p,q,r})$ 、あるいは単に  $\mathbb{G}_{p,q,r}$  と表すことにする。 $\mathbb{G}_{p,q,r}$  は、Clifford product (あるいは Algebraic product (幾何積))と呼ばれる積  $\circ$  を導入することにより、以下のように定義される。

#### [Geometric Algebra $\mathbb{G}_{p,q,r}$ の定義]

$\mathbb{G}_{p,q,r}$  は、次の性質を満たすとき、 $\mathbb{R}^{p,q,r}$  上の Clifford algebra (geometric algebra) と呼ばれる。 $\mathbb{G}_{p,q,r}$  は、実数

<sup>1</sup>Clifford algebra) については、多くの解説書や書籍が出版されている。詳細については、たとえば文献<sup>4,5)</sup>を参照されたい。

<sup>2</sup>この基底は正準基底と呼ばれる。

体  $\mathbb{R}$  とベクトル空間  $\mathbb{R}^{p,q,r}$  を, その部分空間として含む.

- $\mathbb{G}_{p,q,r}$  は, 加算  $+$  およびスカラー ( $\alpha \in \mathbb{R}$ ) との積に関してベクトル空間をなす.
- $\mathbb{G}_{p,q}$  に次の性質を満たす積  $\circ$  が存在する.

1.  $\mathbb{G}_{p,q}$  は, 積  $\circ$  に対し閉じている. すなわち次を満たす.

$$\mathbf{a} \circ \mathbf{b} \in \mathbb{G}_{p,q,r}, \quad \forall \mathbf{a}, \mathbf{b} \in \mathbb{G}_{p,q,r}.$$

2. 結合則:

$$(\mathbf{a} \circ \mathbf{b}) \circ \mathbf{c} = \mathbf{a} \circ (\mathbf{b} \circ \mathbf{c}), \quad \forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{G}_{p,q,r}.$$

3. 分配則:

$$\mathbf{a} \circ (\mathbf{b} + \mathbf{c}) = \mathbf{a} \circ \mathbf{b} + \mathbf{a} \circ \mathbf{c}, \quad \forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{G}_{p,q,r}.$$

4. スカラーとの積:

$$\alpha \circ \mathbf{a} = \mathbf{a} \circ \alpha = \alpha \mathbf{a}, \quad \forall \mathbf{a} \in \mathbb{G}_{p,q,r}, \forall \alpha \in \mathbb{R}.$$

- もし  $\mathbf{a} \in \mathbb{R}^{p,q,r} \subset \mathbb{G}_{p,q,r}$  なら

$$\mathbf{a} \circ \mathbf{a} = \mathbf{a} * \mathbf{a} \in \mathbb{R} \quad (3)$$

を満たす.

以上の定義から分かるように, Clifford algebra  $\mathbb{G}_{p,q,r}$  において, Clifford product  $\circ$  に関して可換であることは要請していない.

## 2.2 基本的な性質と基底 (Algebraic Basis)

Clifford algebra  $\mathbb{G}_{p,q,r}$  の要素は,  $\mathbb{R}^{p,q,r}$  の要素がベクトルと呼ばれるのに対し, マルチベクトル (multivector) と呼ばれる. 任意のマルチベクトル  $\mathbf{a}, \mathbf{b} \in \mathbb{G}_{p,q,r}$  に対し, Clifford product  $\mathbf{a} \circ \mathbf{b}$  は, その対称な部分と反対称な部分の和として

$$\mathbf{a} \circ \mathbf{b} = \frac{1}{2}(\mathbf{a} \circ \mathbf{b} + \mathbf{b} \circ \mathbf{a}) + \frac{1}{2}(\mathbf{a} \circ \mathbf{b} - \mathbf{b} \circ \mathbf{a}).$$

と表される<sup>3</sup>. 特に  $\mathbf{a}$  と  $\mathbf{b}$  をベクトル, すなわち  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^{p,q,r}$  ならば, (3) 式より次の関係が成立する.

$$(\mathbf{a} + \mathbf{b}) \circ (\mathbf{a} + \mathbf{b}) = (\mathbf{a} + \mathbf{b}) * (\mathbf{a} + \mathbf{b})$$

$$\Leftrightarrow \mathbf{a} \circ \mathbf{a} + \mathbf{a} \circ \mathbf{b} + \mathbf{b} \circ \mathbf{a} + \mathbf{b} \circ \mathbf{b} = \mathbf{a} * \mathbf{a} + 2\mathbf{a} * \mathbf{b} + \mathbf{b} * \mathbf{b}$$

$$\Leftrightarrow \frac{1}{2}(\mathbf{a} \circ \mathbf{b} + \mathbf{b} \circ \mathbf{a}) = \mathbf{a} * \mathbf{b}$$

したがって, この場合対称な成分はスカラー積に一致する. ここで, 対称な成分を

$$\mathbf{a} \wedge \mathbf{b} := \frac{1}{2}(\mathbf{a} \circ \mathbf{b} - \mathbf{b} \circ \mathbf{a}),$$

と表すことにすると

$$\mathbf{a} \circ \mathbf{b} = \mathbf{a} * \mathbf{b} + \mathbf{a} \wedge \mathbf{b}.$$

<sup>3</sup>この式の第 1 項は anticommutator product と呼ばれ, 第 2 項は commutator product と呼ばれる

と表せる. 積  $\wedge$  は, outer product あるいは wedge product と呼ばれる. 特に,  $\mathbb{R}^{p,q,r}$  の基底ベクトル  $\mathbf{e}_i, \mathbf{e}_j$  に対しては, (2) 式より  $\mathbf{e}_i * \mathbf{e}_j = 0$  ( $i \neq j$ ) が成立するので,  $\mathbf{e}_i \circ \mathbf{e}_j$  は,

$$\mathbf{e}_i \circ \mathbf{e}_j = \mathbf{e}_i \wedge \mathbf{e}_j$$

となり,

$$\mathbf{e}_i \circ \mathbf{e}_j = -\mathbf{e}_j \circ \mathbf{e}_i. \quad (4)$$

が成立する.

ここで Clifford algebra  $\mathbb{G}_{p,q,r}$  の基底 (algebraic basis と呼ばれる) の生成について説明する. 以下では簡単のため Clifford product である  $\mathbf{a} \circ \mathbf{b}$  を,  $\circ$  を省略して  $\mathbf{ab}$  と表す. また結合則を満たすので,  $(\mathbf{a} \circ \mathbf{b}) \circ \mathbf{c}$  あるいは  $\mathbf{a} \circ (\mathbf{b} \circ \mathbf{c})$  を単に  $\mathbf{abc}$  と表す. また  $\prod$  で, 複数の要素の Clifford product を表す. たとえば,  $\prod_{i=1}^3 \mathbf{a}_i = \mathbf{a}_1 \circ \mathbf{a}_2 \circ \mathbf{a}_3$ . である.  $\mathbb{G}_{p,q,r}$  の基底の生成には, basis blade と呼ばれるベクトル空間  $\mathbb{R}^{p,q,r}$  の相異なる任意の数の基底の Clifford product が重要な役割を果たす. まず basis blade の定義を与える.  $\mathbb{A}$  を順序集合とし,  $\mathbb{A}[i]$  を  $\mathbb{A}$  の第  $i$  番目の要素とする. すなわち, たとえば, もし  $\mathbb{A} = \{2, 3, 1\}$  ならば,  $\mathbb{A}[2] = 3$  である. このとき,  $\mathbb{G}_{p,q,r}$  の basis blade は,  $\mathbb{A}$  を  $\mathbb{A} \subset \{1, 2, \dots, p+q+r\}$  として次のように定義される.

$$\mathbf{e}_{\mathbb{A}} = \prod_{i=1}^{|\mathbb{A}|} \mathbb{R}^{p,q,r}[\mathbb{A}[i]] \quad (5)$$

ここで,  $|\mathbb{A}|$  は集合  $\mathbb{A}$  の要素の数を表す. 各 basis blade  $\mathbf{e}_{\mathbb{A}}$  における Clifford product により演算される要素の数, すなわち  $|\mathbb{A}|$  は, basis blade の grade と呼ばれる. たとえば  $\mathbb{A} = \{2, 3, 1\}$  ならば,  $\mathbf{e}_{\mathbb{A}} = \mathbf{e}_2 \mathbf{e}_3 \mathbf{e}_1$  で, この grade は 3 である.

(1) 式で表される  $\mathbb{R}^{p,q,r}$  の基底は  $p+q+r$  個あるので, これらの全ての組み合わせの Clifford product うち独立な組み合わせの Clifford product の数は  $2^{p+q+r}$  個である.  $\mathbb{G}_{p,q,r}$  の基底は, これら独立な  $2^{p+q+r}$  個の basis blade から構成される. ここで,  $\mathbb{I} = \{1, 2, \dots, p+q+r\}$  とし,  $\mathcal{P}[\mathbb{I}]$  を  $\mathbb{I}$  のべき集合, また  $\mathcal{P}_0[\mathbb{I}]$  を順序づけられた  $\mathbb{I}$  のべき集合とする. たとえば,  $\mathbb{I} = \{1, 2, 3\}$  とすると,

$$\mathcal{P}_0[\mathbb{I}] = \{\{\emptyset\}, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

である. ただし,  $\emptyset$  は空集合である. これより,  $\mathbb{G}_{p,q,r}$  の基底  $\overline{\mathbb{G}}_{p,q,r}$  は, 次のように定義される<sup>4</sup>.

$$\overline{\mathbb{G}}_{p,q,r} := \{\mathbf{e}_{\mathbb{A}} : \mathbb{A} \in \mathcal{P}_0[\mathbb{I}]\}$$

ただし,  $\mathbf{e}_{\emptyset} = 1 \in \mathbb{R}$  としている. たとえば,  $p+q+r = 3$  として  $\mathbb{G}_3 := \mathbb{G}_{p,q,r}$  と表すと, この正準基底  $\overline{\mathbb{G}}_3$  は次のようになる.

$$\overline{\mathbb{G}}_3 = \{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1 \mathbf{e}_2, \mathbf{e}_1 \mathbf{e}_3, \mathbf{e}_2 \mathbf{e}_3, \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3\}$$

Clifford algebra  $\mathbb{G}_{p,q,r}$  の要素, すなわちマルチベクトルは, 以上のように定義された基底の線形結合として

<sup>4</sup>この基底は特に, 正準基底 (canonical algebraic basis) と呼ばれる.



表される. すなわち,  $\mathbf{a} \in \mathbb{G}_{p,q,r}$  は,  $a^{(i)} \in \mathbb{R}$  として次のように表される.

$$\mathbf{a} = \sum_{i=1}^{2^{p+q+r}} a^{(i)} \overline{\mathbb{G}}_{p,q,r}[i] \quad (6)$$

また,  $\mathbf{a} \in \mathbb{G}_{p,q,r}$  の絶対値 (modulus)  $|\mathbf{a}|$  を次のように定義する.

$$|\mathbf{a}| = \left( \sum_{i=1}^{2^{p+q}} a^{(i)2} \right)^{1/2}$$

### 2.3 Clifford Algebra で表現された Hopfield 型ニューラルネットワークのモデル

本節では, Clifford algebra で表現されたリカレント NN のモデルを示す. 特に, 実数で表現された連続時間 Hopfield 型 NN のモデルを拡張したものを 3 種類示す. 1 目モデルは, 以下のような微分方程式で表されるモデルである.

$$\begin{cases} \tau_i \frac{du_i}{dt} = -u_i + \sum_{j=1}^n w_{ij} v_j + b_i \\ v_i = f(u_i) \quad (i = 1, 2, \dots, n) \end{cases} \quad (7)$$

ここで,  $n$  はネットワークを構成するニューロンの個数,  $u_i, v_i$  は, 時刻  $t$  における  $i$  番目のニューロンの状態と出力,  $b_i$  は  $i$  番目のニューロンのしきい値,  $w_{ij}$  は  $j$  番目のニューロンから  $i$  番目のニューロンへの結合荷重,  $\tau_i$  は  $i$  番目のニューロンの時定数である. このモデルにおいて,  $u_i, v_i, b_i, w_{ij}$  はすべて Geometric Algebra  $\mathbb{G}_{p,q,r}$  の要素 (マルチベクトル) である, すなわち  $u_i \in \mathbb{G}_{p,q,r}, v_i \in \mathbb{G}_{p,q,r}, b_i \in \mathbb{G}_{p,q,r}, w_{ij} \in \mathbb{G}_{p,q,r}$  である. 時定数  $\tau_i$  のみ正の実数, すなわち  $\tau_i \in \mathbb{R}, \tau_i > 0$  とする. また,  $w_{ij}$  と  $v_j$  の積は,  $\mathbb{G}_{p,q,r}$  における Clifford product である. また,  $f(\cdot)$  はニューロンの活性化関数であり,  $\mathbb{G}_{p,q,r}$  から  $\mathbb{G}_{p,q,r}$  へ写像する非線形な関数である. また時間微分  $du_i/dt$  は,  $u_i$  の各要素ごとの時間微分, すなわち

$$\frac{d}{dt} u_i(t) := \sum_{i=1}^{2^{p+q}} \frac{d}{dt} u^{(i)}(t) \overline{\mathbb{G}}_{p,q}[i]$$

と定義している. なお, 前節では  $\mathbb{G}_{p,q,r}$  の要素 (マルチベクトル) をすべて太活字 (bold face) で表していたが, このモデルにおいては普通字体で表しており, 本稿の以下ではすべてこのように表す.

Geometric algebra  $\mathbb{G}_{p,q,r}$  の Clifford product は, 一般には可換でないので, 結合重み係数とニューロンの出力との積の順序を逆にしたモデルは (7) と異なるモデルとなる. したがって, 二つ目のモデルとして, 次の微分方程式で表されるモデルを考える.

$$\begin{cases} \tau_i \frac{du_i}{dt} = -u_i + \sum_{j=1}^n v_j w_{ij} + b_i \\ v_i = f(u_i) \quad (i = 1, 2, \dots, n) \end{cases} \quad (8)$$

ここで,  $u_i, v_i, w_{ij}, b_i, \tau_i, f$  などの意味は (7) と同様とする.

また, ニューロンの出力に結合荷重を両側から掛けた場合も (7) とは異なるモデルとなり, この種のモデルとして次のようなモデルを考える.

$$\begin{cases} \tau_i \frac{du_i}{dt} = -u_i + \sum_{j=1}^n w_{ij}^* v_j w_{ij} + b_i \\ v_i = f(u_i) \quad (i = 1, 2, \dots, n) \end{cases} \quad (9)$$

ここで,  $u_i, v_i, w_{ij}, b_i, \tau_i, f$  などの意味は (7) と同様とする. また結合加重  $w_{ij}^*$  は, 一般には結合加重  $w_{ij}$  と異なる任意の Clifford algebra  $\mathbb{G}_{p,q,r}$  の要素でよいが,  $w_{ij}^*$  を  $w_{ij}$  に involution 作用させたものとするのが Clifford algebra においては特に意味を持つ. involution とは二度作用させると元に戻る作用素, すなわち  $(w^*)^* = w$  となる作用素で, Clifford algebra においては inversion, reversion, conjugation などがある. たとえば複素数や四元数における共役も, その一つである. 以上の 3 つのモデルにおいて, (7) 式のモデルが, Hopfield が実数のモデルとして提案した NN モデルの自然な拡張となっている.

## 3 エネルギー関数の存在条件

### 3.1 エネルギー関数の定義

リカレント NN のダイナミクスに関する重要な問題に, エネルギー関数の存在性の問題がある. Hopfield は (7) 式を実数の NN とした場合, すなわち  $u_i, v_i, b_i, w_{ij}$  をすべて  $u_i \in \mathbb{R}, v_i \in \mathbb{R}, b_i \in \mathbb{R}, w_{ij} \in \mathbb{R}$  とし, 活性化関数  $f$  を  $f: \mathbb{R} \rightarrow \mathbb{R}$  とした場合のモデルを対象として, エネルギー関数が存在するための条件とエネルギー関数を示した<sup>6)</sup>. すなわち,  $E(\mathbf{x}): \mathbb{R}^n \rightarrow \mathbb{R}$  なるある関数をエネルギー関数として定義し, 次のような性質を持つことを示した. 実数の (7) 式のモデルにおいて, 重み行列  $W = \{w_{ij}\}$  が対称行列  $W^T = W$  かつ活性化関数  $f(\cdot)$  が連続微分可能で有界かつ単調増加ならば, NN はエネルギー関数を持ち, このエネルギー関数の NN の解に沿っての時間微分が非正となる. このことより, この NN の解は必ずエネルギー関数を減少させるように動作し, 解は  $t \rightarrow \infty$  である平衡点へ収束することが示せる. また彼らはさらに, このエネルギー関数を利用して NN により最適化問題を解く方法を提案し, 実際に巡回セールスマン問題が解けることを示した<sup>7)</sup>. この研究は脳を模倣した NN で実際に最適化問題を解けることを示したのものとして, 各方面に大きなインパクトを与えた. その後, エネルギー関数を用いた NN の定性的な性質の解析やその結果を利用したさまざまな応用, たとえば連想記憶の設計への応用, 各種の最適化問題の応用など非常に活発に研究されてきた. (7), (8) および (9) 式で表される 3 種類の Clifford algebra で表された NN に対しても同様の性質を持つエネルギー関数が定義できれば, これを用いて, 実数の NN でなされてきた種々の研究が Clifford algebra 表現された NN にも可能となり, NN の応用範囲も一層広がると考えられる. 従ってエネルギー関数が存在する条件を明らかにしておくことは重要である. 筆者らはすでに複素数, さらに四元数で表現される NN に対してこの問題の研究をやってきており<sup>8,9,10,11)</sup>, これらの結果は, Clifford algebra で表現される NN の観点から再構築できる<sup>12,13)</sup>. 本稿では, (7), (8) および (9) 式で表され

る3種類の Clifford algebra  $\mathbb{G}_{p,q,r}$  で表された NN のうち特に、 $p+q+r=1$  の場合の  $\mathbb{G}_{1,0,0}$ ,  $\mathbb{G}_{0,1,0}$  および  $\mathbb{G}_{0,0,1}$  で表現された二 (2<sup>1</sup>) 次元の Hopfield 型 NN, さらに  $p+q+r=2$  場合の一つである  $\mathbb{G}_{0,2,0}$  で表現された四 (2<sup>2</sup>) 次元 Hopfield 型 NN に対して、エネルギー関数とその存在条件について議論する。

さて、リカレント NN のダイナミクスを論ずるためにはまず、活性化関数  $f(\cdot)$  として、どのような性質を持つ関数を用いるかが問題となる。実数のニューラルネットワークでは、シグモイド関数のような滑らかで有界な非線形関数が通常用いられる。また、複素 NN においては、活性化関数に有界性を課すと、滑らかな関数として正則な関数 (複素微分可能な関数) は選べない。これは有界で複素平面全体で正則な関数は定数に限るというリュービルの定理による。このため文献 8) では、滑らかさとして正則性の代わりに、複素非線形関数の実数部および虚数部が、それぞれ実数部、虚数部に関して連続偏微分可能な関数を活性化関数として選び、エネルギー関数が存在する条件を導き、具体的にエネルギー関数を示している。

リュービルの定理が Clifford algebra  $\mathbb{G}_{p,q,r}$  に対して常に成立するとは限らないが、本稿では研究の第一歩として、(7), (8) および (9) 式のモデルに対し、活性化関数に複素 NN に対する条件と同様な条件を考えてエネルギー関数の存在条件を議論することにする。

Clifford algebra 上の非線形関数を  $2^{p+q+r}$  個の実数値関数  $f^{(i)}: \mathbb{R}^{2^{p+q+r}} \rightarrow \mathbb{R}$  を用いて次のように表現する。すなわち、 $f(u)$  の  $u$  を (6) 式に従って

$$u = \sum_{i=1}^{2^{p+q+r}} u^{(i)} \overline{\mathbb{G}}_{p,q,r}[i] \quad (10)$$

と表し、 $f(u)$  を

$$f(u) = \sum_{i=1}^{2^{p+q+r}} f^{(i)}(u^{(1)}, u^{(2)}, \dots, u^{(2^{p+q+r})}) \overline{\mathbb{G}}_{p,q,r}[i] \quad (11)$$

と表す。たとえば、 $p+q+r=2$  とした  $\mathbb{G}_2 := \mathbb{G}_{p,q,r}$  においては

$$\overline{\mathbb{G}}_2 = \{1, e_1, e_2, e_1 e_2\}$$

となるので、 $f(u)$  は次のように表わされる。

$$\begin{aligned} f(u) = & f^{(0)}(u^{(0)}, u^{(1)}, u^{(2)}, u^{(3)}) \\ & + f^{(1)}(u^{(0)}, u^{(1)}, u^{(2)}, u^{(3)}) e_1 \\ & + f^{(2)}(u^{(0)}, u^{(1)}, u^{(2)}, u^{(3)}) e_2 \\ & + f^{(3)}(u^{(0)}, u^{(1)}, u^{(2)}, u^{(3)}) e_1 e_2 \end{aligned} \quad (12)$$

このように表現される活性化関数  $f(\cdot)$  に対し、次を仮定する。

- (i)  $f^{(l)}(\cdot)$  は  $u^{(m)}$  ( $l, m = 0, 1, \dots, 2^{p+q+r}$ ) に関して連続偏微分可能である。
- (ii)  $f(\cdot)$  は有界な関数である、すなわち  $|f(\cdot)| \leq M$  を満たす  $M > 0$  が存在する。

ここで、活性化関数  $f(\cdot)$  の点  $u$  におけるヤコビ行列を  $\mathbf{J}_f(u) = \{\alpha_{lm}(u)\} \in \mathbb{R}^{2^{p+q+r} \times 2^{p+q+r}}$  と表すことに

する。ただし、 $\alpha_{lm}$  は次のように定義される。

$$\alpha_{lm}(u) = \left. \frac{\partial f^{(l)}}{\partial u^{(m)}} \right|_u \quad (13)$$

実数の Hopfield 型 NN のエネルギー関数に倣って、(7), (8) 式および (9) 式の NN モデルに対しエネルギー関数を次のように定義する。

**定義 1** 次の条件を満足する関数  $E(\cdot)$  を、Clifford algebra 表現されたニューラルネットワーク ( $\mathcal{N}$ ) のエネルギー関数と呼ぶ。ただし、( $\mathcal{N}$ ) は NN のモデルの式番号、すなわち ( $\mathcal{N}$ ) = (7), ( $\mathcal{N}$ ) = (8), あるいは ( $\mathcal{N}$ ) = (9) とする。

- (i) 関数  $E(\cdot)$  は、 $\mathbb{G}_{p,q,r}$  から  $\mathbb{R}$  への下に有界な写像である。
- (ii)  $E(\cdot)$  の NN の解軌道に沿った時間微分を  $\left. \frac{dE}{dt} \right|_{(\mathcal{N})}$  と表すと、 $\left. \frac{dE}{dt} \right|_{(\mathcal{N})} \leq 0$  を満たす。また、その時間微分が 0 となるのはネットワークの平衡点においてのみである、すなわち  $\left. \frac{dE}{dt} \right|_{(\mathcal{N})} = 0$  となるのは、 $\frac{dv_i}{dt} = 0$  ( $i = 1, 2, \dots, n$ ) が成り立つときであり、またそのときに限る

なお、(7) 式のモデルが実数のモデルである場合、次の関数がエネルギー関数として提案されている。

$$\begin{aligned} E(\mathbf{v}) = & -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} v_i v_j - \sum_{i=1}^n b_i v_i \\ & + \sum_{i=1}^n \int_0^{v_i} f^{-1}(\rho) d\rho \end{aligned} \quad (14)$$

ここで  $\mathbf{v} = [v_1, v_2, \dots, v_n] \in \mathbb{R}^n$  であり、 $f^{-1}$  は  $f$  の逆関数である。

### 3.2 Clifford Algebra $\mathbb{G}_{1,0,0}$ , $\mathbb{G}_{0,1,0}$ および $\mathbb{G}_{0,0,1}$ で表現された NN に対する条件

Clifford algebra  $\mathbb{G}_{1,0,0}$ ,  $\mathbb{G}_{0,1,0}$  および  $\mathbb{G}_{0,0,1}$  の正準基底は次のように表される

$$\overline{\mathbb{G}}_{p,q,r} = \{1, e_1\}.$$

ここで  $\mathbb{G}_{1,0,0}$  に対しては  $e_1 e_1 = 1$ ,  $\mathbb{G}_{0,1,0}$  に対しては  $e_1 e_1 = -1$ ,  $\mathbb{G}_{0,0,1}$  に対しては  $e_1 e_1 = 0$  である。これより、 $\mathbb{G}_{1,0,0}$  は hyperbolic 数と同型、 $\mathbb{G}_{0,1,0}$  は複素数と同型、 $\mathbb{G}_{0,0,1}$  は dual 数と同型ということが分かる。 $\mathbb{G}_{1,0,0}$ ,  $\mathbb{G}_{0,1,0}$  あるいは  $\mathbb{G}_{0,0,1}$  の要素は次のように表される。

$$x = x^{(0)} + x^{(1)} e_1. \quad (15)$$

Clifford algebra  $\mathbb{G}_{1,0,0}$ ,  $\mathbb{G}_{0,1,0}$  あるいは  $\mathbb{G}_{0,0,1}$  においては、Clifford product は可換であるので、(7), (8) および (9) 式で表されるモデルは本質的に等価であり、したがって以下では (7) 式で表現されるモデルを対象として議論する。このモデルに対し、ネットワークの結合重み係数と活性化関数に関して次のような仮定を設ける。

仮定 1 Clifford algebra  $\mathbb{G}_{1,0,0}$ ,  $\mathbb{G}_{0,1,0}$  あるいは  $\mathbb{G}_{0,0,1}$  で表現された (7) 式の NN の結合重み係数は次の条件を満足する。すなわち、 $\mathbb{G}_{1,0,0}$  で表現されたモデルに対しては、次式を満足する。

$$w_{ji} = w_{ij} \quad (i, j = 1, 2, \dots, n). \quad (16)$$

$\mathbb{G}_{0,1,0}$  で表現されたモデルに対しては、次式を満足する。

$$w_{ji} = w_{ij}^* \quad (i, j = 1, 2, \dots, n) \quad (17)$$

ここで \* は、 $w = x^{(0)} + x^{(1)}e_1 \in \mathbb{G}_{0,1,0}$  に対して  $w^* = x^{(0)} - x^{(1)}e_1$  と定義している。 $\mathbb{G}_{0,0,1}$  で表現されたモデルに対しては、次式を満足する。

$$w_{ji} = w_{ij} \quad \text{かつ} \quad w_{ij}^{(1)} = 0 \quad (i, j = 1, 2, \dots, n) \quad (18)$$

ここで  $w_{ij} = w_{ij}^{(0)} + w_{ij}^{(1)}e_1$  である。

Clifford algebra  $\mathbb{G}_{1,0,0}$ ,  $\mathbb{G}_{0,1,0}$  および  $\mathbb{G}_{0,0,1}$  で表現された (7) 式の NN の活性化関数は次のように表される。

$$f(u) = f^{(0)}(u^{(0)}, u^{(1)}) + f^{(1)}(u^{(0)}, u^{(1)})e_1,$$

ここで、 $u = u^{(0)} + u^{(1)}e_1$  である。この活性化関数に対し次の仮定をおく。

仮定 2 Clifford algebra  $\mathbb{G}_{1,0,0}$ ,  $\mathbb{G}_{0,1,0}$  あるいは  $\mathbb{G}_{0,0,1}$  で表現された (7) 式の NN の活性化関数  $f(\cdot)$  は次の条件を満足する。すなわち、 $f(\cdot)$  は 1 対 1 の関数であり、任意の  $u \in \mathbb{G}_{1,0,0}$ ,  $u \in \mathbb{G}_{0,1,0}$  あるいは  $u \in \mathbb{G}_{0,0,1}$  に対して、

$$\left\{ \begin{array}{l} \text{(i)} \quad \frac{\partial f^{(0)}}{\partial u^{(0)}} > 0, \\ \text{(ii)} \quad \frac{\partial f^{(0)}}{\partial u^{(1)}} = \frac{\partial f^{(1)}}{\partial u^{(0)}}, \\ \text{(iii)} \quad \frac{\partial f^{(0)}}{\partial u^{(0)}} \frac{\partial f^{(1)}}{\partial u^{(1)}} - \frac{\partial f^{(0)}}{\partial u^{(1)}} \frac{\partial f^{(1)}}{\partial u^{(0)}} > 0 \end{array} \right. \quad (19)$$

が成り立つ。

ここで、仮定 2 を満足する関数  $f$  の逆関数を  $g = f^{-1}$  とし、 $v = f(u)$  に対して  $u = g(v)$  を

$$g(v) = g^{(0)}(v^{(0)}, v^{(1)}) + g^{(1)}(v^{(0)}, v^{(1)})e_1 \quad (20)$$

と表す。逆関数  $g$  に関して次の補題が成立する。

補題 1 関数  $f$  に対して仮定 2 が成立しているとき、その逆関数  $g$  に対し、次式を満足する関数  $G(\cdot) : \mathbb{G}_{1,0,0} \rightarrow \mathbb{R}$ ,  $\mathbb{G}_{0,1,0} \rightarrow \mathbb{R}$  あるいは  $\mathbb{G}_{0,0,1} \rightarrow \mathbb{R}$  が存在する。

$$\left\{ \begin{array}{l} \frac{\partial G}{\partial v^{(0)}} = g^{(0)}(v^{(0)}, v^{(1)}) \\ \frac{\partial G}{\partial v^{(1)}} = g^{(1)}(v^{(0)}, v^{(1)}) \end{array} \right. \quad (21)$$

この補題は、関数  $G(v)$  を次のように定義することにより証明できる。

$$G(v) := \int_0^{v^{(0)}} g^{(0)}(\rho, 0) d\rho + \int_0^{v^{(1)}} g^{(0)}(v^{(0)}, \rho) d\rho \quad (22)$$

以上に定義した (22) 式の関数  $G(\cdot)$  を用いて Clifford algebra  $\mathbb{G}_{1,0,0}$ ,  $\mathbb{G}_{0,1,0}$  および  $\mathbb{G}_{0,0,1}$  で表現された (7) 式で表される NN に対し、エネルギー関数が次のように構成できる。 $\mathbb{G}_{1,0,0}$  と  $\mathbb{G}_{0,0,1}$  で表されるモデルに対しては、エネルギー関数は次のように構成できる。

$$E(v) = - \sum_{i=1}^n \sum_{j=1}^n \left\{ \frac{1}{2} \mathcal{S}c(v_i w_{ij} v_j + 2b_i v_i) - G(v_i) \right\} \quad \dots \dots \dots (23)$$

ここで  $v = [v_1, v_2, \dots, v_n]^T \in \mathbb{G}_{1,0,0}^n$  あるいは  $v = [v_1, v_2, \dots, v_n]^T \in \mathbb{G}_{0,0,1}^n$  であり、 $\mathcal{S}c(\cdot)$  は、 $x \in \mathbb{G}_{p,q,r}$  に対し、 $\mathcal{S}c(x) = x^{(0)}$  と定義される。すなわち  $\mathcal{S}c(\cdot)$  は、 $\mathbb{G}_{p,q,r}$  の要素のスカラ部を取り出す作用素である。 $\mathbb{G}_{0,1,0}$  で表されるモデルに対しては、エネルギー関数は次のよう構成できる。

$$E(v) = - \sum_{i=1}^n \sum_{j=1}^n \left\{ \frac{1}{2} \mathcal{S}c(v_i^* w_{ij} v_j + 2b_i^* v_i) - G(v_i) \right\} \quad \dots \dots \dots (24)$$

ここで  $v = [v_1, v_2, \dots, v_n]^T \in \mathbb{G}_{0,1,0}^n$  である。以上のエネルギー関数は、(7) 式の NN に対して定義 1 を満足する。すなわち、以下の定理が成立する。

定理 1 Clifford algebra  $\mathbb{G}_{1,0,0}$ ,  $\mathbb{G}_{0,1,0}$  あるいは  $\mathbb{G}_{0,0,1}$  で表現された (7) 式の NN に対して、結合重み係数と活性化関数が仮定 1, 2 を満足するとき、定義 1 を満たすエネルギー関数が存在する。

この定理は次のように証明することができる。すなわち、(23) 式あるいは (24) 式の関数の、(7) 式の NN の解に沿った時間微分を補題 1 を用いて計算することにより、(23) 式あるいは (24) 式の関数がエネルギー関数の定義である定義 1 の条件を満たすことを示すことができる。

以上、Clifford algebra  $\mathbb{G}_{1,0,0}$ ,  $\mathbb{G}_{0,1,0}$  あるいは  $\mathbb{G}_{0,0,1}$  で表現された NN にエネルギー関数が存在するための持つため十分条件を与えた。存在条件は、ネットワークの結合重み係数と活性化関数に関するものである。仮定 2 を満たす活性化関数の例としては次のようなものがある。

$$f(u) = \frac{u}{1 + |u|} \quad (25)$$

$$f(u) = \tanh(u^{(0)}) + \tanh(u^{(1)})e_1 \quad (26)$$

### 3.3 Clifford Algebra $\mathbb{G}_{0,2,0}$ で表現された NN に対する条件

先に述べたように、Clifford algebra  $\mathbb{G}_{0,2,0}$  は四元数代数  $\mathbb{H}$  と同型である。まずこれを見てみよう。 $\mathbb{G}_{0,2,0}$  の正準基底  $\overline{\mathbb{G}}_{0,2,0}$  は、

$$\overline{\mathbb{G}}_{0,2,0} = \{1, e_1, e_2, e_1 e_2\}$$

であり、 $x \in \mathbb{G}_{0,2,0}$  は次のように表される。

$$x = x^{(0)} + x^{(1)}e_1 + x^{(2)}e_2 + x^{(3)}e_1 e_2 \quad (27)$$

基底どうしの積を表にして Table 1 に示す。

Table 1: Multiplication Table for Clifford Algebra  $\mathbb{G}_{0,2,0}$

	1	$e_1$	$e_2$	$e_1e_2$
1	1	$e_1$	$e_2$	$e_1e_2$
$e_1$	$e_1$	-1	$e_1e_2$	$-e_2$
$e_2$	$e_2$	$-e_1e_2$	-1	$e_1$
$e_1e_2$	$e_1e_2$	$e_2$	$-e_1$	-1

一方、四元数は次のように定義される超複素数の一種である。

$$x = x^{(0)} + ix^{(1)} + jx^{(2)} + kx^{(3)} \quad (28)$$

ここで、 $x^{(0)}, x^{(1)}, x^{(2)}, x^{(3)}$  は実数、すなわち  $\mathbb{R}$  の要素であり、 $\{i, j, k\}$  は次の関係を満たす虚数単位である。

$$\begin{aligned} i^2 = -1, \quad j^2 = -1, \quad k^2 = -1, \\ ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j \end{aligned} \quad (29)$$

四元数の積演算はこれらの関係に従って定義され、非可換演算となる。これより、 $e_1$  を四元数の虚数単位  $i$ 、 $e_2$  を四元数の虚数単位  $j$ 、 $e_1e_2$  を四元数の虚数単位  $k$  に対応させると Clifford algebra  $\mathbb{G}_{0,2,0}$  と四元数代数  $\mathbb{H}$  は同型となることがわかる。本項では、Clifford algebra  $\mathbb{G}_{0,2,0}$  表現された (7), (8) および (9) 式の NN モデルを対象として、エネルギー関数の存在条件について議論する<sup>10)</sup>。ただし、(9) 式のモデルにおける  $*$  は次のように定義する。すなわち、 $w \in \mathbb{G}_{0,2,0}$  を

$$w = w^{(0)} + w^{(1)}e_1 + w^{(2)}e_2 + w^{(3)}e_1e_2$$

と表し、 $w^*$  を次のように定義する。

$$w^* = w^{(0)} - w^{(1)}e_1 - w^{(2)}e_2 - w^{(3)}e_1e_2 \quad (30)$$

すなわち、 $w^*$  は四元数においての  $w$  の共役四元数に相当する。

Clifford algebra  $\mathbb{G}_{0,2,0}$  表現された NN (7), (8) あるいは (9) 式に対して、定義 1 を満たす関数を構成するために、ネットワークの結合重み係数と活性化関数に関して次のような仮定を設ける。

**仮定 3** Clifford algebra  $\mathbb{G}_{0,2,0}$  で表現された (7), (8) あるいは (9) 式の NN の結合重み係数は次式を満足する。

$$w_{ji} = w_{ij}^* \quad (i, j = 1, 2, \dots, n) \quad (31)$$

ここで、 $*$  は (30) 式の定義と同じである。

**仮定 4** Clifford algebra  $\mathbb{G}_{0,2,0}$  で表現された (7), (8) あるいは (9) 式の NN の活性化関数  $f(\cdot)$  は次の条件を満足する。

- (i)  $f(\cdot)$  は 1 対 1 の関数である。
- (ii) 任意の  $u \in \mathbb{G}_{0,2,0}$  に対して、 $f(\cdot)$  のヤコビ行列  $\mathbf{J}_f(u)$  は対称である。
- (iii) 任意の  $u \in \mathbb{G}_{0,2,0}$  に対して、 $f(\cdot)$  のヤコビ行列  $\mathbf{J}_f(u)$  は正定値である。

仮定 4 の条件 (i) と活性化関数の有界性から、関数  $f(\cdot)$  の逆関数  $f^{-1}(\cdot) : \mathbb{G}_{0,2,0} \rightarrow \mathbb{G}_{0,2,0}$  が存在する。ここでは、この逆関数を  $g = f^{-1}$  と定義し、 $u = g(v)$  を、四つの実数値関数  $g^{(l)}(\cdot) : \mathbb{R}^4 \rightarrow \mathbb{R}$  ( $l = 0, 1, 2, 3$ ) を用いて次のように表す。

$$\begin{aligned} g(v) = & g^{(0)}(v^{(0)}, v^{(1)}, v^{(2)}, v^{(3)}) \\ & + g^{(1)}(v^{(0)}, v^{(1)}, v^{(2)}, v^{(3)})e_1 \\ & + g^{(2)}(v^{(0)}, v^{(1)}, v^{(2)}, v^{(3)})e_2 \\ & + g^{(3)}(v^{(0)}, v^{(1)}, v^{(2)}, v^{(3)})e_1e_2 \\ & \dots \dots \dots \end{aligned} \quad (32)$$

この時、逆関数  $g(\cdot)$  に対して、以下の補題が成立する。

**補題 2** 関数  $f(\cdot)$  に対して、仮定 4 が成立しているとき、

$$\frac{\partial G}{\partial v^{(l)}} = g^{(l)}(v^{(0)}, v^{(1)}, v^{(2)}, v^{(3)}) \quad (l = 0, 1, 2, 3) \quad (33)$$

を満たす関数  $G(\cdot) : \mathbb{G}_{0,2,0} \rightarrow \mathbb{R}$  が存在する。

この補題は、関数  $G(v)$  を

$$\begin{aligned} G(v) := & \int_0^{v^{(0)}} g^{(0)}(\rho, 0, 0, 0) d\rho \\ & + \int_0^{v^{(1)}} g^{(1)}(v^{(0)}, \rho, 0, 0) d\rho \\ & + \int_0^{v^{(2)}} g^{(2)}(v^{(0)}, v^{(1)}, \rho, 0) d\rho \\ & + \int_0^{v^{(3)}} g^{(3)}(v^{(0)}, v^{(1)}, v^{(2)}, \rho) d\rho \\ & \dots \dots \dots \end{aligned} \quad (34)$$

と定義することにより証明できる。詳細は文献<sup>10)</sup>を参照されたい。

ここで、(34) の関数  $G(\cdot)$  を用いて (7), (8) あるいは (9) 式の Clifford algebra  $\mathbb{G}_{0,2,0}$  表現された NN それぞれに対し、エネルギー関数を構成する。(7) 式の NN が仮定 3, 4 を満たすとき、この NN に対しエネルギー関数を次のように構成できる。

$$E_{(7)}(v) = -\sum_{i=1}^n \sum_{q=j}^n \left\{ \frac{1}{2} Sc(v_i^* w_{ij} v_j + 2b_i^* v_i) - G(v_i) \right\} \quad (35)$$

また、(8) 式の NN が仮定 3, 4 を満たすとき、この NN にエネルギー関数を次のように構成できる。

$$E_{(8)}(v) = -\sum_{i=1}^n \sum_{j=1}^n \left\{ \frac{1}{2} Sc(v_i^* v_j w_{ij} + 2b_i^* v_i) - G(v_i) \right\} \quad (36)$$

(9) 式の NN が仮定 3, 4 を満たすとき、この NN にエネルギー関数を次のように構成できる。

$$E_{(9)}(v) = -\sum_{i=1}^n \sum_{j=1}^n \left\{ \frac{1}{2} Sc(v_i^* w_{ij}^* v_j w_{ij} + 2b_i^* v_i) - G(v_i) \right\} \quad (37)$$

ここで、 $\mathbf{v} = [v_1, v_2, \dots, v_n]^T \in \mathbb{G}_{0,2,0}^n$  としている。これらの関数  $E_{(7)}, E_{(8)}, E_{(9)}$  はそれぞれ、(7), (8), (9) 式の NN に対して定義 1 を満足する。すなわち、以下の定理が成立する。

**定理 2** Clifford algebra  $\mathbb{G}_{0,2,0}$  で表現された (7), (8) あるいは (9) 式の NN に対して、結合重み係数と活性化関数が仮定 3, 4 を満足するとき、定義 1 を満たすエネルギー関数がそれぞれ存在する。

この定理は次のようにして証明することができる。すなわち、(35), (36), (37) 式の関数の、それぞれ (7), (8), (9) 式のネットワークの解軌道に沿った時間微分を補題 2 を用いて計算することにより、これらの関数がエネルギー関数の定義である定義 1 の条件を満たすことを示すことができる。

以上、Clifford algebra  $\mathbb{G}_{0,2,0}$  表現された 3 種類の NN に対して、エネルギー関数を構成し、エネルギー関数が存在するための条件を示した。存在条件はネットワークの結合荷重と活性化関数に関するものである。仮定 4 を満足する活性化関数の例として次の関数がある。

$$f(u) = \frac{u}{1 + |u|} \quad (38)$$

$$f(u) = \tanh(u^{(0)}) + \tanh(u^{(1)})\mathbf{e}_1 + \tanh(u^{(2)})\mathbf{e}_2 + \tanh(u^{(3)})\mathbf{e}_1\mathbf{e}_2 \quad (39)$$

#### 4 おわりに

近年、ニューラルネットワークモデルを複素数値化、四元数値化など高次元化し、その情報処理能力、学習法や応用などの研究が盛んに行われている。これら高次元の表現は Clifford algebra (geometric algebra) 表現として一般化される。本稿では、Clifford algebra 表現されたリカレント NN のモデルを示し、そのダイナミクスについて議論した。特に 3 種類の Hopfield 型連続時間 NN のモデルを示した。またこれらのモデルの中のあるクラスの Clifford algebra 表現された NN のダイナミクスとして、エネルギー関数の存在条件に関して議論した。具体的には、二次元の 3 つのクラスの Hopfield 型 NN と 4 次元の 1 つのクラスの Hopfield 型 NN に対するエネルギー関数の存在条件を議論した。さらに一般のクラスの Clifford algebra で表現された Hopfield 型 NN に対しても、エネルギー関数の存在条件が本稿の結果を拡張することにより導出できると考えられ、今後の課題としたい。また、リユービルの定理が成立しない Clifford algebra で表現される Hopfield 型 NN のモデルの構築とそのエネルギー関数の存在条件の導出も今後の課題である。

Clifford algebra は、様々な現実の問題に対し、それら構造を柔軟に取り込むことが可能で、豊かな問題表現能力および情報表現能力を有している。ニューロコンピューティングにおいても Clifford algebra を導入することにより、新たな研究のパラダイムが拓けると期待される。

#### 参考文献

- 1) A. Hirose (ed.): Complex-Valued Neural Networks Theorisis and Applications, World Scientific, (2003)
- 2) T. Nitta (ed.): Complex-Valued Neural Networks Utilizing High-Dimentional Parameters, IGI Global, (2009)

- 3) S. Buchholz: A Theory of Neural Computation with Clifford Algebra, Ph.D. Thesis, University. of Kiel, (2005)
- 4) P. Lounesto: Clifford Algebras and Spinors 2nd Edition, Cambridge Univ. Press, (2001)
- 5) Christian Perwass: Geometric Algebra with Applications in Engineering, Springer-Verlag, (2009)
- 6) J. J. Hopfield: Neurons with graded response have collective computational properties like those of two-state neurons; *Proc. Natl. Acad. Sci. USA*, Vol.81, 3088/3092 (1984)
- 7) J. J. Hopfield and D. W. Tank: Neural' computation of decisions in optimization problems; *Biol. Cybern.*, Vol.52, 141/152 (1985)
- 8) 橋本, 黒江, 森: 複素ニューラルネットワークのエネルギー関数について; システム制御情報学会論文誌, Vol.15, No.10, 559/565 (2002)
- 9) Y. Kuroe, M. Yoshida and T. Mori: On Activation Functions for Complex-Valued Neural Networks - Existence of Energy Functions -; *Artificial Neural Networks and Neural Information Processing - ICANN/ICONIP 2003*, Okyay Kaynak et. al.(Eds.), Lecture Notes in Computer Science, 2714, 985/992, Springer, (2003)
- 10) M. Yoshida, Y. Kuroe and T. Mori: Models of Hopfield-Type Quaternion Neural Networks and Their Energy Functions; *International Journal of Neural Systems*, Vol.15, Nos.1 & 2, 129/135 (2005)
- 11) 谷川, 黒江, 飯間: ホップフィールド型高次元ニューラルネットワークのモデルとそのエネルギー関数; 計測自動制御学会第 37 回知能システムシンポジウム資料, 13/18 (2010)
- 12) Y. Kuroe: Models of Clifford Recurrent Neural Networks and Their Dynamics; *Proceedings of 2011 International Joint Conference on Neural Networks*, 1035/1041 (2011)
- 13) Y. Kuroe, S. Tanigawa and H. Iima: Models of Hopfield-type Clifford Neural Networks and Their Energy Functions - Hyperbolic and Dual Valued Networks -, *Proceedings of ICONIP 2011, Lecture Notes in Computer Science 7062*, Springer, (2011) (to appear)
- 14) L. Dorst, D. Fontijne and S. Mann: *Geometric Algebra for Computer Science An object-oriented Approach to Geometry*, Morgan Kaufmann Publisher, (2007)
- 15) E. Bayro-Corrochano and G. Scheuermann (Eds.): *Geometric Algebra Computing in Engineering and Computer Science*, Springer-Verlag,(2010)



# Non-constant bounded holomorphic functions of hyperbolic numbers – Candidates for hyperbolic activation functions

\* Eckhard Hitzer (University of Fukui)

**Abstract**– The Liouville theorem states that bounded holomorphic complex functions are necessarily constant. Holomorphic functions fulfill the so-called Cauchy-Riemann (CR) conditions. The CR conditions mean that a complex  $z$ -derivative is independent of the direction. Holomorphic functions are ideal for activation functions of complex neural networks, but the Liouville theorem makes them useless. Yet recently the use of hyperbolic numbers, lead to the construction of hyperbolic number neural networks. We will describe the Cauchy-Riemann conditions for hyperbolic numbers and show that there exists a new interesting type of bounded holomorphic functions of hyperbolic numbers, which are not constant. We give examples of such functions. They therefore substantially expand the available candidates for holomorphic activation functions for hyperbolic number neural networks.

**Keywords:** Hyperbolic numbers, Liouville theorem, Cauchy-Riemann conditions, bounded holomorphic functions

## 1 Introduction

For the sake of mathematical clarity, we first carefully review the notion of holomorphic functions in the two number systems of complex and hyperbolic numbers.

The Liouville theorem states that bounded holomorphic complex functions  $f : \mathbb{C} \rightarrow \mathbb{C}$  are necessarily constant [1]. Holomorphic functions are functions that fulfill the so-called Cauchy-Riemann (CR) conditions. The CR conditions mean that a complex  $z$ -derivative

$$\frac{df(z)}{dz}, \quad z = x + iy \in \mathbb{C}, \quad x, y \in \mathbb{R}, \quad ii = -1, \quad (1)$$

is independent of the direction with respect to which the incremental ratio, that defines the derivative, is taken [5]. Holomorphic functions would be ideal for activation functions of complex neural networks, but the Liouville theorem means that careful measures need to be taken in order to avoid poles (where the function becomes infinite).

Yet recently the use of hyperbolic numbers

$$z = x + hy, \quad h^2 = 1, \quad x, y \in \mathbb{R}, \quad h \notin \mathbb{R}. \quad (2)$$

lead to the construction of hyperbolic number neural networks. We will describe the generalized Cauchy-Riemann conditions for hyperbolic numbers and show that there exist bounded holomorphic functions of hyperbolic numbers, which are not constant. We give a new example of such a function. They are therefore excellent candidates for holomorphic activation functions for hyperbolic number neural networks [2, 3]. In [3] it was shown, that hyperbolic number neural networks allow to control the angle of the decision

boundaries (hyperplanes) of the real and the unipotent  $h$ -part of the output. But Buchholz argued in [4], p. 114, that

Contrary to the complex case, the hyperbolic logistic function is bounded. This is due to the absence of singularities. Thus, in general terms, this seems to be a suitable activation function. Concretely, the following facts, however, might be of disadvantage. The real and imaginary part have different squashing values. Both component functions do only significantly differ from zero around the lines<sup>1</sup>  $x = y$  ( $x > 0$ ) and  $x = y$  ( $x < 0$ ).

Complex numbers are isomorphic to the Clifford geometric algebra  $Cl_{0,1}$  which is generated by a single vector  $e_1$  of negative square  $e_1 = -1$ , with algebraic basis  $\{1, e_1\}$ . The isomorphism  $\mathbb{C} \cong Cl_{0,1}$  is realized by mapping  $i \mapsto e_1$ .

Hyperbolic numbers are isomorphic to the Clifford geometric algebra  $Cl_{1,0}$  which is generated by a single vector  $e_1$  of positive square  $e_1 = +1$ , with algebraic basis  $\{1, e_1\}$ . The isomorphism between hyperbolic numbers and  $Cl_{1,0}$  is realized by mapping  $h \mapsto e_1$ .

## 2 Complex variable functions

We follow the treatment given in [5]. We assume a complex function given by an absolute convergent

<sup>1</sup>Note that we slightly correct the two formulas of Buchholz, because we think it necessary to delete  $e_1$  in Buchholz' original  $x = ye_1$  ( $x > 0$ ), etc.

power series.

$$w = f(z) = f(x + iy) = u(x, y) + iv(x, y), \quad (3)$$

where  $u, v : \mathbb{R}^2 \rightarrow \mathbb{R}$  are real functions of the real variables  $x, y$ . Since  $u, v$  are obtained in an algebraic way from the complex number  $z = x + iy$ , they cannot be arbitrary functions but must satisfy certain conditions. There are several equivalent ways to obtain these conditions. Following Riemann, we state that a function  $w = f(z) = u(x, y) + iv(x, y)$  is a function of the complex variable  $z$  if its derivative is independent of the direction (in the complex plane) with respect to which the incremental ratio is taken. This requirement leads to two partial differential equations, named after Cauchy and Riemann (CR), which relate  $u$  and  $v$ .

One method for obtaining these equations is the following. We consider the expression  $w = u(x, y) + iv(x, y)$  only as a function of  $z$ , but not of  $\bar{z}$ , i.e. the derivative with respect to  $\bar{z}$  shall be zero. First we perform the bijective substitution

$$x = \frac{1}{2}(z + \bar{z}), \quad y = -i\frac{1}{2}(z - \bar{z}), \quad (4)$$

based on  $z = x + iy, \bar{z} = x - iy$ . For computing the derivative  $w_{,\bar{z}} = \frac{dw}{d\bar{z}}$  with the help of the chain rule we need the derivatives of  $x$  and  $y$  of (4)

$$x_{,\bar{z}} = \frac{1}{2}, \quad y_{,\bar{z}} = \frac{1}{2}i. \quad (5)$$

Using the chain rule we obtain

$$\begin{aligned} w_{,\bar{z}} &= u_{,x}x_{,\bar{z}} + u_{,y}y_{,\bar{z}} + i(v_{,x}x_{,\bar{z}} + v_{,y}y_{,\bar{z}}) \\ &= \frac{1}{2}u_{,x} + \frac{1}{2}iu_{,y} + i\left(\frac{1}{2}v_{,x} + \frac{1}{2}iv_{,y}\right) \\ &= \frac{1}{2}[u_{,x} - v_{,y} + i(v_{,x} + u_{,y})] \stackrel{!}{=} 0. \end{aligned} \quad (6)$$

Requiring that both the real and the imaginary part of (6) vanish we obtain the *Cauchy-Riemann conditions*

$$u_{,x} = v_{,y}, \quad u_{,y} = -v_{,x}. \quad (7)$$

Functions of a complex variable that fulfill the CR conditions are functions of  $x$  and  $y$ , but they are only functions of  $z$ , not of  $\bar{z}$ .

It follows from (7), that both  $u$  and  $v$  fulfill the *Laplace equation*

$$u_{,xx} = v_{,yy} = v_{,xy} = -u_{,yy} \Leftrightarrow u_{,xx} + u_{,yy} = 0, \quad (8)$$

and similarly

$$v_{,xx} + v_{,yy} = 0. \quad (9)$$

The Laplace equation is a simple example of an elliptic partial differential equation. The general theory of solutions to the Laplace equation is known as potential theory. The solutions of the Laplace equation

are called *harmonic* functions and are important in many fields of science, notably the fields of electromagnetism, astronomy, and fluid dynamics, because they can be used to accurately describe the behavior of electric, gravitational, and fluid potentials. In the study of heat conduction, the Laplace equation is the steady-state heat equation [6].

*Liouville's theorem* [1] states, that any bounded holomorphic function  $f : \mathbb{C} \rightarrow \mathbb{C}$ , which fulfills the CR conditions is constant. Therefore for complex neural networks it is not very meaningful to use holomorphic functions as activation functions. If they are used, special measures need to be taken to avoid poles in the complex plane. Instead separate componentwise (split) real scalar functions for the real part  $g_r : \mathbb{R} \rightarrow \mathbb{R}, u(x, y) \mapsto g_r(u(x, y))$ , and for the imaginary part  $g_i : \mathbb{R} \rightarrow \mathbb{R}, v(x, y) \mapsto g_i(v(x, y))$ , are usually adopted. Therefore a standard split activation function in the complex domain is given by

$$g(u(x, y) + iv(x, y)) = g_r(u(x, y)) + ig_i(v(x, y)). \quad (10)$$

### 3 Hyperbolic numbers

Hyperbolic numbers are also known as *split-complex* numbers. They form a two-dimensional commutative algebra. The canonical hyperbolic system of numbers is defined [5] by

$$z = x + hy, \quad h^2 = 1, \quad x, y \in \mathbb{R}, \quad h \notin \mathbb{R}. \quad (11)$$

The *hyperbolic conjugate* is defined as

$$\bar{z} = x - hy. \quad (12)$$

Taking the hyperbolic conjugate corresponds in the isomorphic algebra  $Cl_{1,0}$  to taking the main involution (grade involution), which maps  $1 \mapsto 1, e_1 \mapsto -e_1$ .

The *hyperbolic invariant* (corresponding to the Lorentz invariant in physics for  $y = ct$ ), or *modulus*, is defined as

$$z\bar{z} = (x + hy)(x - hy) = x^2 - y^2, \quad (13)$$

which is not positive definite.

Hyperbolic numbers are fundamentally different from complex numbers. Complex numbers and quaternions are division algebras, every non-zero element has a unique inverse. Hyperbolic numbers do not always have an inverse, but instead there are idempotents and divisors of zero.

We can define the following *idempotent basis*

$$n_1 = \frac{1}{2}(1 + h), \quad n_2 = \frac{1}{2}(1 - h), \quad (14)$$



which fulfills

$$\begin{aligned}
n_1^2 &= \frac{1}{4}(1+h)(1+h) = \frac{1}{4}(2+2h) = n_1, \\
n_2^2 &= n_2, \quad n_1 + n_2 = 1, \\
n_1 n_2 &= \frac{1}{4}(1+h)(1-h) = \frac{1}{4}(1-1) = 0, \\
\bar{n}_1 &= n_2, \quad \bar{n}_2 = n_1.
\end{aligned} \tag{15}$$

The inverse basis transformation is simply

$$1 = n_1 + n_2, \quad h = n_1 - n_2. \tag{16}$$

Setting

$$z = x + hy = \xi n_1 + \eta n_2, \tag{17}$$

we get the corresponding coordinate transformation

$$x = \frac{1}{2}(\xi + \eta), \quad y = \frac{1}{2}(\xi - \eta), \tag{18}$$

as well as the inverse coordinate transformation

$$\xi = x + y \in \mathbb{R}, \quad \eta = x - y \in \mathbb{R}. \tag{19}$$

The hyperbolic conjugate becomes, due to (15), in the idempotent basis

$$\bar{z} = \xi \bar{n}_1 + \eta \bar{n}_2 = \eta n_1 + \xi n_2. \tag{20}$$

In the idempotent basis, using (20) and (15), the hyperbolic invariant becomes multiplicative

$$\begin{aligned}
z\bar{z} &= (\xi n_1 + \eta n_2)(\eta n_1 + \xi n_2) \\
&= \xi\eta(n_1 + n_2) = \xi\eta = x^2 - y^2.
\end{aligned} \tag{21}$$

In the following we consider the product and quotient of two hyperbolic numbers  $z, z'$  both expressed in the idempotent basis  $\{n_1, n_2\}$

$$zz' = (\xi n_1 + \eta n_2)(\xi' n_1 + \eta' n_2) = \xi\xi' n_1 + \eta\eta' n_2, \tag{22}$$

and

$$\begin{aligned}
\frac{z}{z'} &= \frac{\xi n_1 + \eta n_2}{\xi' n_1 + \eta' n_2} = \frac{z\bar{z}'}{z'\bar{z}} \\
&= \frac{(\xi n_1 + \eta n_2)(\eta' n_1 + \xi' n_2)}{(\xi' n_1 + \eta' n_2)(\eta' n_1 + \xi' n_2)} \\
&= \frac{(\xi\eta' n_1 + \eta\xi' n_2)(\eta' n_1 + \xi' n_2)}{\xi'\eta'} \\
&= \frac{\xi}{\xi'} n_1 + \frac{\eta}{\eta'} n_2.
\end{aligned} \tag{23}$$

Because of (23) it is not possible to divide by  $z'$  if  $\xi' = 0$ , or if  $\eta' = 0$ . Moreover, the product of a hyperbolic number with  $\xi = 0$  (on the  $n_2$  axis) times a hyperbolic number with  $\eta = 0$  (on the  $n_1$  axis) is

$$(\xi n_1 + 0n_2)(0n_1 + \eta n_2) = \xi\eta n_1 n_2 = 0, \tag{24}$$

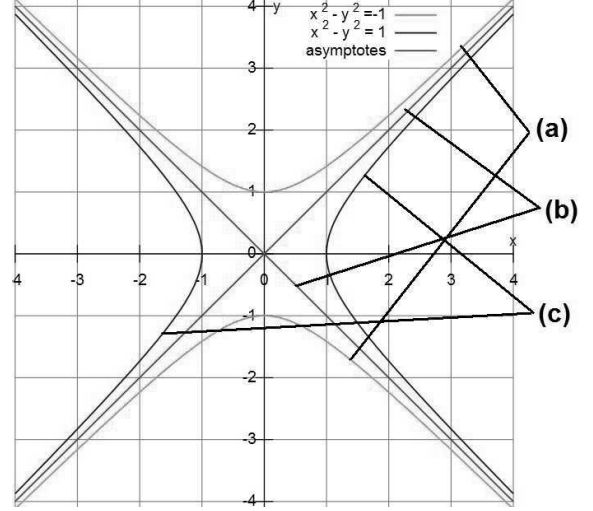


Figure 1: The hyperbolic number plane [9] with horizontal  $x$ -axis and vertical  $yh$ -axis, showing: (a) Hyperbolas with modulus  $z\bar{z} = -1$  (green). (b) Straight lines with modulus  $z\bar{z} = 0 \Leftrightarrow x^2 = y^2$  (red), i.e. divisors of zero. (c) Hyperbolas with modulus  $z\bar{z} = 1$  (blue).

due to (15). We repeat that in (24) the product is zero, even though the factors are non-zero. The numbers  $\xi n_1, \eta n_2$  along the  $n_1, n_2$  axis are therefore called *divisors of zero*. The divisors of zero have no inverse.

The hyperbolic plane with the diagonal lines of divisors of zero (b), and the pairs of hyperbolas with constant modulus  $z\bar{z} = 1$  (c), and  $z\bar{z} = -1$  (a) is shown in Fig. 1.

## 4 Hyperbolic number functions

We assume a hyperbolic number function given by an absolute convergent power series

$$\begin{aligned}
w &= f(z) = f(x + hy) = u(x, y) + hv(x, y), \\
h^2 &= 1, \quad h \notin \mathbb{R}.
\end{aligned} \tag{25}$$

where  $u, v : \mathbb{R}^2 \rightarrow \mathbb{R}$  are real functions of the real variables  $x, y$ . An example of a hyperbolic number function is the exponential function

$$\begin{aligned}
e^z &= e^{x+hy} = e^x e^{hy} = e^x (\cosh y + h \sinh y) \\
&= u(x, y) + hv(x, y),
\end{aligned} \tag{26}$$

with

$$u(x, y) = e^x \cosh y, \quad v(x, y) = e^x \sinh y. \tag{27}$$

Since  $u, v$  are obtained in an algebraic way from the hyperbolic number  $z = x + hy$ , they cannot be arbitrary functions but must satisfy certain conditions.

There are several equivalent ways to obtain these conditions. A function  $w = f(z) = u(x, y) + hv(x, y)$  is a function of the hyperbolic variable  $z$ , if its derivative is independent of the direction (in the hyperbolic plane) with respect to which the incremental ratio is taken. This requirement leads to two partial differential equations, so called *generalized Cauchy-Riemann* (GCR) conditions, which relate  $u$  and  $v$ .

To obtain the GCR conditions we consider the expression  $w = u(x, y) + hv(x, y)$  only as a function of  $z$ , but not of  $\bar{z} = x - hy$ , i.e. the derivative with respect to  $\bar{z}$  shall be zero. First we perform the bijective substitution

$$x = \frac{1}{2}(z + \bar{z}), \quad y = h\frac{1}{2}(z - \bar{z}), \quad (28)$$

based on  $z = x + hy$ ,  $\bar{z} = x - hy$ . For computing the derivative  $w_{,\bar{z}} = \frac{dw}{d\bar{z}}$  with the help of the chain rule we need the derivatives of  $x$  and  $y$  of (28)

$$x_{,\bar{z}} = \frac{1}{2}, \quad y_{,\bar{z}} = -\frac{1}{2}h. \quad (29)$$

Using the chain rule we obtain

$$\begin{aligned} w_{,\bar{z}} &= u_{,x}x_{,\bar{z}} + u_{,y}y_{,\bar{z}} + h(v_{,x}x_{,\bar{z}} + v_{,y}y_{,\bar{z}}) \\ &= \frac{1}{2}u_{,x} - \frac{1}{2}hu_{,y} + h\left(\frac{1}{2}v_{,x} - \frac{1}{2}hv_{,y}\right) \\ &= \frac{1}{2}[u_{,x} - v_{,y} + h(v_{,x} - u_{,y})] \stackrel{!}{=} 0. \end{aligned} \quad (30)$$

Requiring that both the real and the  $h$ -part of (30) vanish we obtain the GCR conditions

$$u_{,x} = v_{,y}, \quad u_{,y} = v_{,x}. \quad (31)$$

Functions of a hyperbolic variable that fulfill the GCR conditions are functions of  $x$  and  $y$ , but they are only functions of  $z$ , not of  $\bar{z}$ . Such functions are called (hyperbolic) holomorphic functions.

It follows from (31), that  $u$  and  $v$  fulfill the wave equation

$$u_{,xx} = v_{,yy} = v_{,xy} = u_{,xy} \Leftrightarrow u_{,xx} - u_{,yy} = 0, \quad (32)$$

and similarly

$$v_{,xx} - v_{,yy} = 0. \quad (33)$$

The wave equation is an important second-order linear partial differential equation for the description of waves – as they occur in physics – such as sound waves, light waves and water waves. It arises in fields like acoustics, electromagnetics, and fluid dynamics. The wave equation is the prototype of a hyperbolic partial differential equation [7].

Let us compute the partial derivatives  $u_{,x}, u_{,y}, v_{,x}, v_{,y}$  for the exponential function  $e^z$  of (26):

$$\begin{aligned} u_{,x} &= e^x \cosh y, & u_{,y} &= e^x \sinh y, \\ v_{,x} &= e^x \sinh y = u_{,y}, & v_{,y} &= e^x \cosh y = u_{,x}. \end{aligned} \quad (34)$$

We clearly see that the partial derivatives (34) fulfill the GCR conditions (31) for the exponential function  $e^z$ , as expected by its definition (26). The exponential function  $e^z$  is therefore a manifestly holomorphic hyperbolic function, but it is not bounded.

In the case of holomorphic hyperbolic functions the GCR conditions do not imply a Liouville type theorem like for holomorphic complex functions. This can most easily be demonstrated with a counter example

$$\begin{aligned} f(z) &= u(x, y) + hv(x, y), \\ u(x, y) = v(x, y) &= \frac{1}{1 + e^{-x}e^{-y}}. \end{aligned} \quad (35)$$

The function  $u(x, y)$  is pictured in Fig. 2.

Let us verify that the function  $f$  of (35) fulfills the GCR conditions

$$\begin{aligned} u_{,x} &= \frac{-1}{(1 + e^{-x}e^{-y})^2}(-e^{-x}e^{-y}) \\ &= \frac{e^{-x}e^{-y}}{(1 + e^{-x}e^{-y})^2}, \end{aligned} \quad (36)$$

where we repeatedly applied the chain rule for differentiation. Similarly we obtain

$$u_{,y} = v_{,x} = v_{,y} = \frac{e^{-x}e^{-y}}{(1 + e^{-x}e^{-y})^2}. \quad (37)$$

The GCR conditions (31) are therefore clearly fulfilled, which means that the hyperbolic function  $f(z)$  of (35) is holomorphic. Since the exponential function  $e^{-x}$  has a range of  $(0, \infty)$ , the product  $e^{-x}e^{-y}$  also has values in the range of  $(0, \infty)$ . Therefore the function  $1 + e^{-x}e^{-y}$  has values in  $(1, \infty)$ , and the components of the function  $f(z)$  of (35) have values

$$0 < \frac{1}{1 + e^{-x}e^{-y}} < 1. \quad (38)$$

We especially have

$$\lim_{x, y \rightarrow -\infty} \frac{1}{1 + e^{-x}e^{-y}} = 0, \quad (39)$$

and

$$\lim_{x, y \rightarrow \infty} \frac{1}{1 + e^{-x}e^{-y}} = 1. \quad (40)$$

The function (35) is *representative* for how to turn any real neural node activation function  $r(x)$  into holomorphic hyperbolic activation function via

$$f(x) = r(x + y)(1 + h). \quad (41)$$

We note that in [3, 4] another holomorphic hyperbolic activation function was studied, namely

$$f'(z) = \frac{1}{1 + e^{-z}}, \quad (42)$$

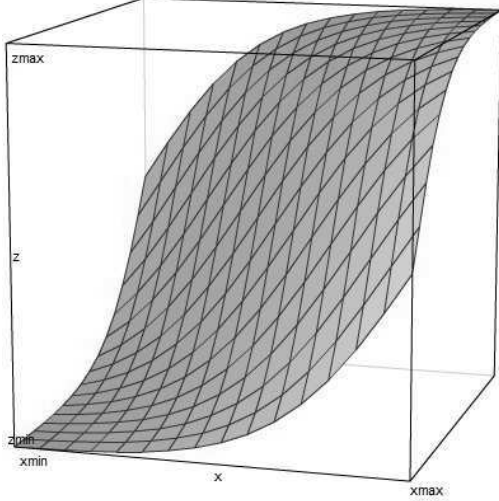


Figure 2: Function  $u(x, y) = 1/(1 + e^{-x}e^{-y})$ . Horizontal axis  $-3 \leq x \leq 3$ , from left corner into paper plane  $-3 \leq y \leq 3$ . Vertical axis  $0 \leq u \leq 1$ . (Figure produced with [8].)

but compare the quote from [4], p. 114, given in the introduction. The split activation function used in [2]

$$f''(x, y) = \frac{1}{1 + e^{-x}} + h \frac{1}{1 + e^{-y}}, \quad (43)$$

is clearly not holomorphic, because the real part  $u = 1/(1 + e^{-x})$  depends only on  $x$  and not on  $y$ , and the  $h$ -part  $v = 1/(1 + e^{-y})$  depends only on  $y$  and not on  $x$ , thus the GCR conditions (31) can not be fulfilled.

## 5 Geometric interpretation of multiplication of hyperbolic numbers

In order to geometrically interpret the product of two complex numbers, it proves useful to introduce polar coordinates in the complex plane. Similarly, for the geometric interpretation of the product of two hyperbolic numbers, we first introduce *hyperbolic polar coordinates* for  $z = x + hy$  with radial coordinate

$$\rho = \sqrt{|z\bar{z}|} = \sqrt{|x^2 - y^2|}. \quad (44)$$

The hyperbolic polar coordinate transformation [5] is then given as

1.  $x^2 > y^2, x > 0$ :

$$\theta = \operatorname{artanh}(y/x), \quad z = \rho e^{h\theta},$$

i.e. the quadrant in the hyperbolic plane of Fig. 1 limited by the diagonal idempotent lines, and including the positive  $x$ -axis (to the right).

2.  $x^2 > y^2, x < 0$ :

$$\theta = \operatorname{artanh}(y/x), \quad z = -\rho e^{h\theta},$$

i.e. the quadrant in Fig. 1 including the negative  $x$ -axis (to the left).

3.  $x^2 < y^2, y > 0$ :

$$\theta = \operatorname{artanh}(x/y), \quad z = h\rho e^{h\theta},$$

i.e. the quadrant in Fig. 1 including the positive  $y$ -axis (top).

4.  $x^2 < y^2, y < 0$ :

$$\theta = \operatorname{artanh}(x/y), \quad z = -h\rho e^{h\theta},$$

i.e. the quadrant in Fig. 1 including the negative  $y$ -axis (bottom).

The product of a constant hyperbolic number (assuming  $a_x^2 > a_y^2, a_x > 0$ )

$$a = a_x + ha_y = \rho_a e^{h\theta_a},$$

$$\rho_a = \sqrt{a_x^2 - a_y^2}, \quad \theta_a = \operatorname{artanh}(a_y/a_x), \quad (45)$$

with a hyperbolic number  $z$  (assuming  $x^2 > y^2, x > 0$ ) in hyperbolic polar coordinates is

$$az = \rho_a e^{h\theta_a} \rho e^{h\theta} = \rho_a \rho e^{h(\theta + \theta_a)}. \quad (46)$$

The geometric interpretation is a scaling of the modulus  $\rho \rightarrow \rho_a \rho$  and a hyperbolic rotation (movement along a hyperbola)  $\theta \rightarrow \theta + \theta_a$ .

In the physics of Einstein's special relativistic space-time [11, 12], the hyperbolic rotation  $\theta \rightarrow \theta + \theta_a$  corresponds to a Lorentz transformation from one inertial frame with constant velocity  $\tanh \theta$  to another inertial frame with constant velocity  $\tanh(\theta + \theta_a)$ . Neural networks based on hyperbolic numbers (dimensionally extended to four-dimensional space-time) should therefore be ideal to compute with electromagnetic signals, including satellite transmission.

## 6 Conclusion

We have compared complex numbers and hyperbolic numbers, as well as complex functions and hyperbolic functions. We saw that according to Liouville's theorem bounded complex holomorphic functions are necessarily constant, *but* non-constant bounded hyperbolic holomorphic functions exist. One such function has already been studied in [3, 4]. We have studied a promising example of a hyperbolic holomorphic function

$$f(z) = \frac{1 + h}{1 + e^{-x-y}}, \quad (47)$$

in some detail. The distinct notions of idempotents and divisors of zero, special to hyperbolic numbers, were introduced. After further introducing hyperbolic polar coordinates, a geometric interpretation of the hyperbolic number multiplication was given.

Hyperbolic neural networks offer, compared to complex neural networks, therefore the advantage of *suitable bounded non-constant hyperbolic holomorphic activation functions*. It would certainly be of interest to study convergence, accuracy and decision boundaries of hyperbolic neural networks with the activation function (35), similar to [3, 4].

## Acknowledgment

I want to acknowledge God [13]:

In the beginning was the Word<sup>2</sup>, and the Word was with God, and the Word was God. He was with God in the beginning. Through him all things were made; without him nothing was made that has been made. In him was life, and that life was the light of all mankind.

I want to thank my dear family, as well as T. Nitta and Y. Kuroe.

## References

- [1] K. Guerlebeck et al, *Holomorphic Functions in the Plane and n-dimensional Space*, Birkhauser, 2008, chp. 7.3.3.
- [2] S. Buchholz, G. Sommer, *A hyperbolic multi-layer perceptron*, Proceedings of the International Joint Conference on Neural Networks, Como, Italy, vol. 2, 129/133 (2000).
- [3] T. Nitta, S. Buchholz, *On the Decision Boundaries of Hyperbolic Neurons*, Proceedings of the International Joint Conference on Neural Networks, IJCNN'08-HongKong, June 1-6, 2973/2979(2008).
- [4] S. Buchholz, PhD Thesis, *A Theory of Neural Computation with Clifford Algebras*, University of Kiel, 2005.
- [5] F. Catoni et al, *The Mathematics of Minkowski Space-Time*, Birkhauser, 2008.
- [6] Laplace's equation, Wikipedia, accessed 24 August 2011, [http://en.wikipedia.org/wiki/Laplace's\\_equation](http://en.wikipedia.org/wiki/Laplace's_equation)
- [7] Wave equation, Wikipedia, accessed 24 August 2011, [http://en.wikipedia.org/wiki/Wave\\_equation](http://en.wikipedia.org/wiki/Wave_equation)
- [8] Online 3D function grapher, <http://www.livephysics.com/ptools/online-3d-function-grapher.php?>
- [9] Split-complex number, Wikipedia, accessed 29 August 2011, [http://en.wikipedia.org/wiki/Split-complex\\_number](http://en.wikipedia.org/wiki/Split-complex_number)
- [10] Notes of collaboration with H. Ishi, Feb. 2011, p. 15.
- [11] C. Doran and A. Lasenby, *Geometric Algebra for Physicists*, Cambridge University Press, Cambridge (UK), 2003.
- [12] E. Hitzer, *Relativistic Physics as Application of Geometric Algebra*, in K. Adhav (ed.), Proceedings of the International Conference on Relativity 2005 (ICR2005), University of Amravati, India, January 2005, 71/90(2005).
- [13] The Bible, New International Version (NIV), The Gospel according to John, chapter 1, verses 1-4, <http://www.biblegateway.com/>
- [14] Strong's Bible lexicon entry G3056 for *logos*, available online at Blue Letter Bible. <http://www.blueletterbible.org/lang/lexicon/lexicon.cfm?Strongsg3056&t=KJV>

---

<sup>2</sup>Greek term: *logos*. Note: A Greek philosopher named Heraclitus first used the term Logos around 600 B.C. to designate the divine reason or plan which coordinates a changing universe. [14]

# Conformal Geometric Algebra を用いた近似方法の提案とその応用

○ファンミントゥン (名古屋大学) 橘完太 (工学院大学)  
吉川大弘 (名古屋大学) 古橋武 (名古屋大学)

## A proposal of Approximation Method using Conformal Geometric Algebra and its Applications

\*M. T. PHAM (Nagoya University), K. Tachibana (Kogakuin University),  
T. Yoshikawa (Nagoya University) and T. Furuhashi (Nagoya University)

**Abstract**— Approximation is one of the most useful methods for understanding data distribution. However, most conventional approximation methods using Gaussian distribution or uniform distribution do not pay sufficient attention to the geometric properties of data. This study focuses Geometric algebra (GA), which is a generalization of complex numbers and quaternions able to describe spatial objects and the relations between them. This paper uses conformal GA (CGA), which is a part of GA, to transform a vector in a real Euclidean vector space into a conformal vector in a CGA space and presents a proposed new approximation method using conformal vectors. In particular, this paper shows that the proposed method was able to express various data distributions, such as those based on hyper-spheres (-planes), circles (lines), or arcs. Then, this paper shows that the proposed method is able to apply to clustering problems and three-dimensional reconstruction using combination with particle filter.

**Key Words:** Conformal Geometric Algebra, Approximation, Clustering, Three-dimensional Reconstruction

### 1 はじめに

従来の多くの近似手法では、空間的な特徴を含むデータ (以下、幾何データ) に対しても幾何的な性質を考慮せずに近似を行う。例えば、ある  $m$  次元のデータ分布に関して、回転や平行移動を持つ幾何ベクトルが計測された場合、多くの従来手法では、 $m$  個の実数、あるいは  $m$  次元の実ベクトル  $x \in \mathbf{R}^m$  として扱う。ところがこの方法では、ベクトル間における回転や平行移動などの幾何的な関係は十分に抽出されない。そのため、これら従来手法では幾何的な性質を見つけれない。

本稿では、幾何データとして回転や平行移動を持つ幾何ベクトルが観測された場合、それら回転や平行移動の性質を考慮した新たな近似手法を提案する。本手法は、空間ベクトルや空間ベクトル間の関係を簡潔に記述できる Geometric Algebra (GA) <sup>1), 2)</sup> に着目し、GA の一部となる Conformal GA (CGA) <sup>3)</sup> の幾何空間を用いた近似方法である。GA の一部とみなすことができる複素数や四元数については、これらをそれぞれ用いた画像処理・時系列信号処理や色付画像処理・多次元時系列信号処理などの、空間的な処理を要する応用での成功例が多く挙げられる <sup>4), 5), 6), 7), 8), 9)</sup>。また、GA 値入出力関数の学習手法 <sup>10)</sup> および GA による特徴抽出 <sup>11)</sup> も提案され、その有効性が示されている。

提案手法では、 $m$  次元空間におけるベクトルに対して、データ分布の超球 (平面) への近似を行う。CGA を用いた空間変換を行い、データベクトルを CGA 空間のポイント  $Q$  とし、超球  $S$  を CGA 空間の conformal ベクトルとする。CGA の性質から、任意のポイント  $Q$  が球  $S$  の表面上にある場合、内積  $S \cdot Q$  は 0 であることが容易に確認できる。この特徴を利用して最小二乗法を用いたエラー関数を定義し、近似超球 (平面) の推定を行う。

本稿では超球の中心が事前に分かる場合及び超球 (超平面) の中心が事前に分からない場合 <sup>12)</sup> をそれぞれの適用先に考察する。超球の中心が事前に分かる場合の適用例として、パーティクルフィルタ <sup>13)14)</sup> を用いた単カメラからの物体の関節位置追跡問題に提案手法を適用する。従来手法では、関節の回転を考慮する場合、回転速度に相当するクォータニオンを用いる <sup>15)16)17)</sup>。しかし本稿で扱う関節推定問題では、関節の回転半径が分からないため、クォータニオンを用いても、パーティクルの座標の予測精度が初期のパーティクル座標に左右される。提案手法では、近似手法により回転軌道を通る球を推定できるため、高い追跡精度が期待される。そして、超球 (超平面) の中心が事前に分からない場合の適用例として、クラスタリングに提案手法を適用する。提案手法は  $m$  次元における  $m$  次元の超球 (平面) のデータだけではなく  $m - k$  次元の超球 (平面) や円弧などの豊富な表現に近似可能であるため、従来のクラスタリングと比べ幾何的なクラスタの分類精度の向上が期待される。

本稿の構成は以下である。第 2 節では、提案手法として、実空間ベクトルデータを CGA 空間の conformal ベクトルに変換を行い、CGA 空間における conformal ベクトル間の内積を用いた超球への近似手法について述べる。そして、提案手法の応用例として単カメラから物体の関節位置を追跡するための新しいパーティクルフィルタと新しい幾何的クラスタリングを提案する。第 3 節では、回転中の関節位置の推定のシミュレーション実験と幾何的データにおけるクラスタリング、及び提案手法の性能を評価する。最後に第 4 節でまとめを述べる。

### 2 Conformal Geometric Algebra

Conformal Geometric Algebra (CGA) は、クリフォード代数と呼ばれる Geometric Algebra (GA) に

含まれ、代数構造の一種である。GA は  $(p + q)$  個の正規直交基底  $\mathcal{O} = \{\mathbf{e}_1, \dots, \mathbf{e}_p, \mathbf{e}_{p+1}, \dots, \mathbf{e}_{p+q}\}$  を、 $\mathbf{e}_i^2 = +1, \forall i \in \{1, \dots, p\}$ , および  $\mathbf{e}_i^2 = -1, \forall i \in \{p+1, \dots, p+q\}$  となるように定義する。本稿では、 $\mathcal{O}$  により定められる空間を  $\mathcal{G}_{p,q}$  で表わす。例えば、実ベクトル空間  $\mathbf{R}^m$  の場合は  $\mathcal{G}_{m,0}$  で表現することができる。

CGA 空間は実ベクトル空間  $\mathbf{R}^m$  を拡張し、 $m+2$  個の基底ベクトル  $\{\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{e}_+, \mathbf{e}_-\}$  により定めることができる。ただし、 $\mathbf{e}_+$  および  $\mathbf{e}_-$  は以下により定義される。

$$\begin{aligned} \mathbf{e}_+^2 &= \mathbf{e}_+ \cdot \mathbf{e}_+ = 1, \\ \mathbf{e}_-^2 &= \mathbf{e}_- \cdot \mathbf{e}_- = -1, \end{aligned} \quad (1)$$

$$\mathbf{e}_+ \cdot \mathbf{e}_- = \mathbf{e}_+ \cdot \mathbf{e}_i = \mathbf{e}_- \cdot \mathbf{e}_i = 0, \forall i \in \{1, \dots, m\}.$$

つまり、CGA 空間は  $\mathcal{G}_{m+1,1}$  により表現することができる。さらに CGA では、ベクトル  $\mathbf{e}_0$  および  $\mathbf{e}_\infty$  を  $\mathbf{e}_+$  と  $\mathbf{e}_-$  により定義する。

$$\mathbf{e}_0 = \frac{1}{2}(\mathbf{e}_- - \mathbf{e}_+), \mathbf{e}_\infty = (\mathbf{e}_- + \mathbf{e}_+). \quad (2)$$

Eq. (1) と Eq. (2) から次の関係式が得られる。

$$\begin{aligned} \mathbf{e}_0 \cdot \mathbf{e}_0 &= \mathbf{e}_\infty \cdot \mathbf{e}_\infty = 0, \\ \mathbf{e}_0 \cdot \mathbf{e}_\infty &= \mathbf{e}_\infty \cdot \mathbf{e}_0 = -1, \\ \mathbf{e}_0 \cdot \mathbf{e}_i &= \mathbf{e}_\infty \cdot \mathbf{e}_i = 0, \forall i \in \{1, \dots, m\}. \end{aligned} \quad (3)$$

次に Hestenes<sup>3)</sup> により提案された空間変換を用いる。CGA 空間でのポイント  $P \in \mathcal{G}_{m+1,1}$  は実ベクトル  $\mathbf{x} = \sum_i x_i \mathbf{e}_i \in \mathbf{R}^m$  を用い、以下の式により変換される。

$$P = \mathbf{x} + \frac{1}{2}\|\mathbf{x}\|^2 \mathbf{e}_\infty + \mathbf{e}_0. \quad (4)$$

ここで、同じポイントどうしの内積  $P \cdot P$  は 0 であることが分かる。そして、 $m$  次元における超球は CGA により表現できる。

$$\begin{aligned} S &= P - \frac{1}{2}r^2 \mathbf{e}_\infty \\ &= \mathbf{x} + \frac{1}{2}\{\|\mathbf{x}\|^2 - r^2\} \mathbf{e}_\infty + \mathbf{e}_0. \end{aligned} \quad (5)$$

$S$  は実ベクトル空間  $\mathbf{R}^m$  で中心  $\mathbf{x}$ 、半径  $r$  の超球を表す conformal ベクトルである。任意のポイント  $Q$  が球  $S$  の表面上にある場合、内積  $S \cdot Q$  は 0 となることが容易に確認できる。また、Eq. (4) 及び Eq. (5) から、CGA 空間  $\mathcal{G}_{m+1,1}$  でのポイントは半径  $r = 0$  の球であることが分かる。

CGA 空間では、平面  $L$  も表現できる、

$$L = \mathbf{n} + d\mathbf{e}_\infty. \quad (6)$$

$\mathbf{n}$  は実ベクトル空間  $\mathbf{R}^m$  における平面  $\mathbf{n} \cdot \mathbf{x} + d = 0$  の直交ベクトルであり、 $d$  はスカラー係数である。任意のポイント  $Q$  が平面  $L$  の表面上にある場合、内積  $S \cdot L$  は 0 である。

一般的に、 $\mathcal{G}_{n+1,1}$  における conformal ベクトル  $S$  は以下によって表記される。

$$S = \mathbf{s} + s_\infty \mathbf{e}_\infty + s_0 \mathbf{e}_0. \quad (7)$$

$\mathbf{s} = \sum_i^m s_i \mathbf{e}_i$  は実ベクトル空間  $\mathbf{R}^m$  におけるベクトルであり、 $s_\infty$  および  $s_0$  はそれぞれベクトル  $\mathbf{e}_\infty$  と  $\mathbf{e}_0$  のスカラー係数である。

### 3 提案手法とその応用

#### 3.1 幾何的データセットから超球の近似手法の提案

本稿では、CGA 空間でのポイント  $P$  とベクトル  $S$  との距離がそれらの内積であることに着目する。Eq. (3)、Eq. (4) および Eq. (7) から、 $P$  と  $S$  の内積を以下により計算する。

$$\begin{aligned} d(P, S) &\propto P \cdot S \\ &= \left( \mathbf{x} + \frac{1}{2}\|\mathbf{x}\|^2 \mathbf{e}_\infty + \mathbf{e}_0 \right) \cdot (\mathbf{s} + s_\infty \mathbf{e}_\infty + s_0 \mathbf{e}_0) \\ &= \mathbf{x} \cdot \mathbf{s} - s_\infty - \frac{1}{2}\|\mathbf{x}\|^2 s_0. \end{aligned} \quad (8)$$

ここで、 $d(P, S) = 0$  の場合、

$$\begin{aligned} \mathbf{x} \cdot \mathbf{s} - s_\infty - \frac{1}{2}\|\mathbf{x}\|^2 s_0 &= 0 \\ \Leftrightarrow \begin{cases} \mathbf{x} \cdot \mathbf{s} - s_\infty = 0 & (s_0 = 0), \\ \|\mathbf{x} - \frac{1}{s_0} \mathbf{s}\|^2 = \frac{\|\mathbf{s}\|^2 - 2s_\infty s_0}{s_0^2} & (s_0 \neq 0). \end{cases} \end{aligned} \quad (9)$$

つまり Eq. (9) は、 $d(P, S) = 0$  のときポイント  $P$  が超平面上にある、または超球上にあることを意味する。本稿ではこの性質を利用し、回転や平行移動のような幾何的データを超球（超平面）に近似する手法を提案する。

提案手法では、 $m$  次元空間におけるベクトル集合  $\mathcal{X} = \{\mathbf{x}_j = \sum_i x_{j,i} \mathbf{e}_i \in \mathbf{R}^m, j = 1, \dots, n\}$  が与えられたとき、Eq. (4) を用いて空間変換を行う。そのとき、CGA 空間  $\mathcal{G}_{m+1,1}$  でのポイント集合  $\mathcal{P} = \{P_j = \mathbf{x}_j + \frac{1}{2}\|\mathbf{x}_j\|^2 \mathbf{e}_\infty + \mathbf{e}_0, j = 1, \dots, n\}$  が得られる。また、ポイント集合  $\mathcal{P}$  の近似超球  $S = \mathbf{s} + s_\infty \mathbf{e}_\infty + s_0 \mathbf{e}_0 \in \mathcal{G}_{m+1,1}$  を推定するために、最小二乗法を用いる。ここでは、以下のエラー関数を用いて  $S$  の推定を行う。

$$\begin{aligned} E &= \sum_{j=1}^n d^2(P_j, S) \\ &= \sum_{j=1}^n \left( \mathbf{x}_j \cdot \mathbf{s} - s_\infty - \frac{1}{2}\|\mathbf{x}_j\|^2 s_0 \right)^2. \end{aligned} \quad (10)$$

ここで、Eq. (9) から、 $\mathbf{s}$  をスカラー倍した場合、 $s_\infty$  および  $s_0$  も同じ倍数の等価式を得ることができる。つまり、 $\mathbf{s}$  の長さを固定しても、近似超球の推定結果が正しく得られる。本稿では  $\|\mathbf{s}\|^2 = 1$  の制約条件のもとで  $E$  の最適化を行う。したがって、この最適化問題は以下のようなになる。

$$\begin{aligned} \min & \sum_{j=1}^n \left( \mathbf{x}_j \cdot \mathbf{s} - s_\infty - \frac{1}{2}\|\mathbf{x}_j\|^2 s_0 \right)^2, \\ \text{s.t.} & \|\mathbf{s}\|^2 = 1. \end{aligned} \quad (11)$$

近似超球の中心  $\mathbf{o}$  が事前に分かると想定しているため、Eq. (11) から超球  $S$  を容易に推定できる。

$$S = \mathbf{o} + s_\infty \mathbf{e}_\infty + \mathbf{e}_0. \quad (12)$$

ただし、

$$s_\infty = \frac{1}{n} \sum_{j=1}^n \left( \mathbf{x}_j \cdot \mathbf{o} - \frac{1}{2} \|\mathbf{x}_j\|^2 \right). \quad (13)$$

一般の問題では、近似超球の中心が事前に分からない場合、ラグランジュの未定乗数法を用いた最適化を行うことにより、中心を推定することが可能である。ラグランジュ関数は次式ようになる。

$$L = \sum_{k=1}^n \left( \mathbf{x}_k \cdot \mathbf{s} - s_\infty - \frac{1}{2} \|\mathbf{x}_k\|^2 s_0 \right)^2 + \lambda (\|\mathbf{s}\|^2 - 1). \quad (14)$$

Eq. (14) が最適である条件は以下のようになる。

$$\frac{\partial L}{\partial \mathbf{s}} = 2 \sum_{k=1}^n \left( \mathbf{x}_k \cdot \mathbf{s} - s_\infty - \frac{1}{2} \|\mathbf{x}_k\|^2 s_0 \right) \mathbf{x}_k + 2\lambda \mathbf{s} = 0, \quad (15)$$

$$\frac{\partial L}{\partial s_\infty} = -2 \sum_{k=1}^n \left( \mathbf{x}_k \cdot \mathbf{s} - s_\infty - \frac{1}{2} \|\mathbf{x}_k\|^2 s_0 \right) = 0, \quad (16)$$

$$\frac{\partial L}{\partial s_0} = - \sum_{k=1}^n \left( \mathbf{x}_k \cdot \mathbf{s} - s_\infty - \frac{1}{2} \|\mathbf{x}_k\|^2 s_0 \right) \|\mathbf{x}_k\|^2 = 0. \quad (17)$$

Eq. (16) および Eq. (17) から、以下の関係式が得られる。

$$s_\infty = \mathbf{f}_\infty \cdot \mathbf{s} \quad (18)$$

$$\frac{1}{2} s_0 = \mathbf{f}_0 \cdot \mathbf{s} \quad (19)$$

ただし、

$$\mathbf{f}_\infty = \frac{- \sum_{k=1}^n \|\mathbf{x}_k\|^4 \sum_{k=1}^n \mathbf{x}_k + \sum_{k=1}^n \|\mathbf{x}_k\|^2 \sum_{k=1}^n \|\mathbf{x}_k\|^2 \mathbf{x}_k}{\sum_{k=1}^n \|\mathbf{x}_k\|^2 \sum_{k=1}^n \|\mathbf{x}_k\|^2 - \sum_{k=1}^n \|\mathbf{x}_k\|^4 \sum_{k=1}^n 1}, \quad (20)$$

$$\mathbf{f}_0 = \frac{\sum_{k=1}^n \|\mathbf{x}_k\|^2 \sum_{k=1}^n \mathbf{x}_k - \sum_{k=1}^n 1 \sum_{k=1}^n \|\mathbf{x}_k\|^2 \mathbf{x}_k}{\sum_{k=1}^n \|\mathbf{x}_k\|^2 \sum_{k=1}^n \|\mathbf{x}_k\|^2 - \sum_{k=1}^n \|\mathbf{x}_k\|^4 \sum_{k=1}^n 1}. \quad (21)$$

ここで、

$$\mathbf{f}(\mathbf{x}) = \mathbf{x} - \mathbf{f}_\infty - \|\mathbf{x}\|^2 \mathbf{f}_0 \in \mathbf{R}^m \quad (22)$$

とおくと、ラグランジュ関数  $L$  (Eq. 14) は以下の式に変換することができる。

$$L = \mathbf{s}^T \mathbf{A} \mathbf{s} + \lambda (\|\mathbf{s}\|^2 - 1), \quad (23)$$

ただし、

$$\mathbf{A} = \sum_{k=1}^n \mathbf{f}(\mathbf{x}_k) \mathbf{f}^T(\mathbf{x}_k). \quad (24)$$

従って、最適化の条件 (Eq. (15)) は以下の固有値問題に帰着することができる。

$$\mathbf{A} \mathbf{s} = \lambda \mathbf{s}. \quad (25)$$

ここで、Eq. (25) から得られた固有ベクトル  $\mathbf{s}$  と固有値  $\lambda$  のそれぞれは近似超球および、データポイント  $P$  と対応する超球  $S$  との距離の分散を表わす。  $\lambda$  が小さければ小さいほど、得られる超球はデータとよりフィッティングすることを表わす。

### 3.2 パーティクルフィルタを用いた関節の追跡

本稿では、提案した超球の近似手法を用いて、単カメラの動画から物体の関節を追跡し、関節の3次元座標を推定することを応用有効例として示す。ここで、物体追跡のテクニックとして近年よく使われているパーティクルフィルタ<sup>13)14)</sup>を用いる。パーティクルフィルタとは、多数のパーティクルを用いて、前状態からの予測と現在の観測情報から、現在の状態を推定する手法である。パーティクルフィルタはすべての時間  $t$  について、以下の2ステップにより、パーティクルのパラメータ  $\theta$  の更新を繰り返す。

**予測：** すべてのパーティクル  $j$  に対して、過去のパラメータから時刻  $t$  のパラメータ  $\theta_j^t$  を予測する。

$$\theta_j^t = \delta(\theta^{t-1}, \dots, \theta^1).$$

**更新：** 制約条件または観測情報  $\mathbf{y}^t$  から、全てのパーティクルの尤度  $L_j = p(\mathbf{y}^t | \theta_j^t, \dots, \theta_j^1)$  を計算し、小さい尤度を持つパーティクルを削除し、大きい尤度を持つパーティクルを増やす。次に、すべてのパラメータに対してノイズを加える操作を行う。

本稿では、全ての時刻  $t$  に対する親関節の位置  $\mathbf{o}^t \in \mathbf{R}^3$  が分かる前提で、親関節の中心に回転中のターゲット関節の3次元座標を、単カメラの2次元動画から推定する。ただし、親関節とターゲット関節との距離は分からないとする。従来手法では、関節の回転を考慮しない場合、各パーティクル  $j$  の平行移動速度  $\mathbf{v}_j^t$  を用いて関節の座標  $\mathbf{x}_j^{t+t} = \mathbf{x}_j^t + \mathbf{v}_j^t$  を予測する。本稿では、関節の回転を考慮し、回転速度と相当するクォータニオン  $\mathbf{q}_j^t$  を用いる。このとき予測座標は  $\mathbf{x}_j^{t+t} = \mathbf{q}_j^t (\mathbf{x}_j^t - \mathbf{o}^t) \mathbf{q}_j^t + \mathbf{o}^t$  となる。しかし、この関節推定問題では、関節の回転半径が分からないため、クォータニオンを用いてもパーティクルの座標の予測は現在の座標に左右される。その結果、予測が不安定となる。逆に、回転半径を正しく推定できれば、クォータニオンを用いなくても、パーティクルの座標におけるターゲット関節とカメラの焦点を通る直線と、回転軌道を通る球の交差する点により、ターゲット関節の位置座標を推定できる。

ここでは、CGA及びパーティクルフィルタを用いて、回転軌道を通る球を推定する。まず、パーティクルの座標をターゲット関節とカメラの焦点を通る直線上になるように制約を加える。

$$\mathbf{x}_j^t = \tilde{\mathbf{o}}^t + \alpha_j^t \mathbf{l}^t \quad (26)$$

Table 1: クォータニオンを持つパーティクルフィルタと提案手法によるパーティクルフィルタ

	QPF	Proposed Method
固有パラメータ	$\mathbf{x}_j^t, \mathbf{q}_j^t$	$\mu^t, \alpha_j^t$
予測式	$\mathbf{x}_j^{t+1} = \mathbf{q}_j^t (\mathbf{x}_j^t - \mathbf{o}^t) \mathbf{q}_j^t + \mathbf{o}^t$	$\mathbf{x}_j^{t+1} = \tilde{\mathbf{o}}^{t+1} + \mu^t \alpha_j^t \mathbf{1}^{t+1}$
尤度	$L_j = \ \mathbf{x}_j^{t+1} - \tilde{\mathbf{o}}^{t+1} - ((\mathbf{x}_j^{t+1} - \tilde{\mathbf{o}}^{t+1}) \cdot \mathbf{1}) \mathbf{1}\ $	$L_j = \ \tilde{\mathbf{x}}_j^t \tilde{\mathbf{x}}_j^{t-1} (\tilde{\mathbf{x}}_j^t)^{-1} - \tilde{\mathbf{x}}_j^{t+1}\ $
パーティクル数	100	100
淘汰率	5%	5%

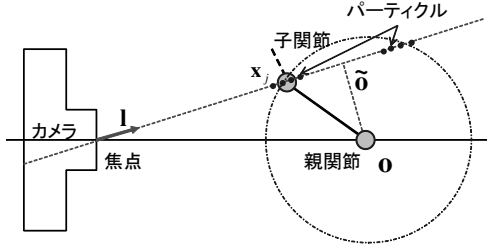


Fig. 1: カメラ及び物体関節の配置イメージ

ただし  $\mathbf{l}^t$  は、時間  $t$  でターゲット関節とカメラの焦点を通る直線と平行する単位ベクトル ( $\|\mathbf{l}^t\| = 1$ ) であり、 $\tilde{\mathbf{o}}^t = (\mathbf{l}^t \cdot \mathbf{o}^t) \mathbf{l}^t$  は親関節  $\mathbf{o}^t$  におけるその直線への写像である。  $\mathbf{l}^t$  および  $\mathbf{o}^t$  は時間  $t$  で観測できるとする。また  $\alpha_j^t$  はパーティクル  $j$  の固有パラメータである。図 1 に、カメラと物体関節、及びパーティクルの配置イメージを示す。

ここで、ある時間  $t$  ですべてのパーティクルがターゲット関節を追跡できると仮定すると、各パーティクルと親関節との距離の期待値はターゲット関節の回転半径となる。このとき、提案した近似手法により、回転軌道を通る球をすべてのパーティクルの座標から推定できる。また、追跡が安定したとき、時刻  $t$  と時刻  $t+1$  の推定された球は同じとなるため、

$$\begin{aligned} S &= \arg \min_{S^t} \sum_j S^t \cdot \mathbf{x}_j^t \\ &= \arg \min_{S^{t+1}} \sum_j S^{t+1} \cdot \mathbf{x}_j^{t+1} \end{aligned} \quad (27)$$

が得られる。パーティクルの予測式を  $\alpha_j^{t+1} = \mu^t \alpha_j^t$  とすると、式 (27) から  $\mu^t$  を容易に求められる。パーティクルの予測座標は  $\mathbf{x}_j^{t+1} = \tilde{\mathbf{o}}^{t+1} + \mu^t \alpha_j^t \mathbf{1}^{t+1}$  となる。物体の回転加速度が小さいと仮定すると、パーティクルの尤度は

$$L_j = \|\tilde{\mathbf{x}}_j^t \tilde{\mathbf{x}}_j^{t-1} (\tilde{\mathbf{x}}_j^t)^{-1} - \tilde{\mathbf{x}}_j^{t+1}\| \quad (28)$$

により算出される。ただし、 $\tilde{\mathbf{x}}_j^t = \mathbf{x}_j^t - \mathbf{o}^t$  である。

### 3.3 CGA を用いたクラスタリングの提案

本節では、回転および平行移動の混合を含まれるデータにおいて、円 (線) と球 (平面) および超球 (超平面) によりデータを分類するクラスタリング手法の提案を応用例として示す。まず、クラスタ  $c$  (データラベル  $y = c$ ) における  $m$  次元ベクトル集合  $\mathcal{X}_c = \{\mathbf{x}_k = \sum_i x_{k,i} \mathbf{e}_i \in \mathbf{R}^m \mid y_k = c, k \in \{1, \dots, n\}\}$  において、Eq. (18) と Eq. (19) および Eq. (25) から、 $m$  個の超球  $S_{c,l} = \mathbf{s}_{c,l} + s_{\infty,c,l} \mathbf{e}_\infty + s_{0,c,l} \mathbf{e}_0, l \in \{1, \dots, m\}$  および、データポイント  $P = \mathbf{x} + \frac{1}{2} \|\mathbf{x}\|^2 \mathbf{e}_\infty + \mathbf{e}_0$  と対応す

る超球  $S_{c,l}$  との距離  $d(P, S_{c,l})$  の分散  $\lambda_{c,l}$  を得ることができる。そして、距離  $d(P, S_{c,l})$  はガウス分布に従うと仮定した場合、以下の確率密度関数が得られる。

$$p(\mathbf{x} \mid \lambda_{c,l}) = \frac{1}{\sqrt{2\pi\lambda_{c,l}}} \exp\left(-\frac{d^2(P, S_{c,l})}{2\lambda_{c,l}}\right). \quad (29)$$

また、各固有値に対応する距離は独立であることを仮定するとベクトル  $\mathbf{x}$  はクラスタ  $c$  に属する確率密度関数は以下により定義することができる。

$$p(\mathbf{x} \mid c) = \prod_l^m p(\mathbf{x} \mid \lambda_{c,l}). \quad (30)$$

Eq. (30) を用いた提案手法のアルゴリズムは以下である。

**Step1** クラスタ数  $C$  を決定する。

**Step2** データラベル  $y_k \in \{1, \dots, C\}$  をランダムで初期化する。

**Step3** 全てのクラスタ  $\mathcal{X}_c = \{\mathbf{x}_k \mid y_k = c\}$  に対して、確率密度関数  $p(\mathbf{x} \mid c)$  を推定する。

**Step4** データラベル  $y_k = \arg \max_c p(\mathbf{x} \mid c) P(c)$  により  $y_k$  を更新する。ただし、 $P(c) = \frac{|\mathcal{X}_c|}{n}$  はクラスタ  $c$  である確率を表わす。  $|\mathcal{X}_c|$  はクラスタ  $c$  におけるデータ数である。

**Step5** Step4 でデータラベル  $y_k$  の更新がなければアルゴリズムが終了する。その以外 Step3 に戻る。

## 4 実験と考察

### 4.1 3次元物体の関節の推定

本節の実験では、3次元物体の関節の推定を行う。問題を簡単化するため、固定された親関節と全ての方向に回転可能な子関節だけを持つ物体を対象とする。回転クォータニオンを持つパーティクルフィルタ、及び提案手法によるパーティクルフィルタを用いて回転中の子関節の座標を推定する。表 1 に、それぞれの手法のパラメータと式を示す。本実験では、カメラの焦点座標を原点  $(0, 0, 0)$  とし、親関節の座標を  $\mathbf{o} = (0, 0, 20)$  とする。子関節と親関節の距離を 5 で固定とする。回転速度と角度をランダムに変えた 5 回の試行を行った。

表 2 に、関節と最も近いパーティクルとの距離、及び関節とパーティクルの重心との距離の平均を示す。q Best と CGA Best はそれぞれ、回転クォータニオンを持つパーティクルフィルタ、及び提案手法による関節と最も近いパーティクルとの距離の平均を示す。q Avg と CGA Avg はそれぞれ、回転クォータニオンを持つ



パーティクルフィルタ、及び提案手法による関節と全パーティクルの重心との距離の平均を示す。表 2 から、回転クォータニオンを持つパーティクルフィルタと比べ、提案手法により高い追跡精度が得られたことがわかる。一対の標本による平均の検定により、これらの平均値の間に有意差が確認された。

Table 2: 関節と最も近いパーティクルとの距離、及び関節と全パーティクルの重心との距離の平均

	q Best	CGA Best	q Avg	CGA Avg
試行 1	1.02	0.11	3.88	2.48
試行 2	0.95	0.02	4.64	3.35
試行 3	1.1	0.03	4.12	1.81
試行 4	1.33	0.3	3.01	2.31
試行 5	1.08	0.1	4.13	2.26

図 2 に、1 回の試行での関節の追跡精度を示す。図 2 において、横軸は時間を示し、縦軸は関節との距離を示す。関節の回転軌道を通る球と、関節とカメラの焦点を通る直線との交差点は通常 2 点あるため、パーティクルは 2 つのグループに分けられる。1 つのグループは関節付近に配置され、もう 1 つのグループは残りの交差点付近に配置される。そして、関節とカメラの焦点を通る直線が、関節の回転軌道を通る球との接線となると、全てのパーティクルが関節付近の 1 つのグループに集まる。そのため、図 2 で示すように、関節と全パーティクルの重心との距離は回転に応じて周期的に変動する。また関節と最も近いパーティクルとの距離は、回転クォータニオンを持つパーティクルフィルタと比べ、提案手法が関節を安定的に追跡できることが確認できた。

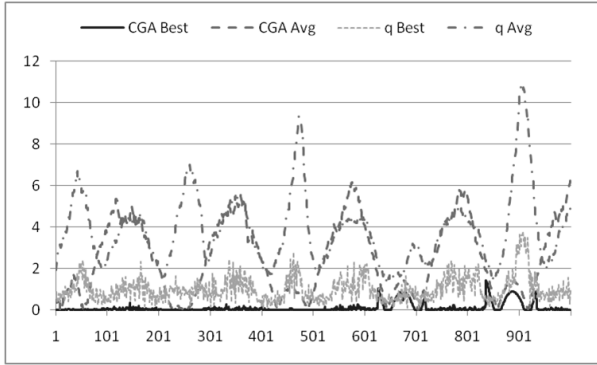


Fig. 2: 関節の追跡精度の例

## 4.2 クラスタリングの実験結果

本節では 3 種類の 2 次元データと 1 種類の 3 次元データを用いてクラスタリングを行う。k-means<sup>18)</sup> と hard C Spherical Shells (HCSS)<sup>19)</sup> および提案手法の 3 つのクラスタリング手法を用いてデータ分類を行う。Fig. (3) は本実験でクラスタリングを行う際使用された幾何的データを示す。Fig. (3) の (a)(b)(c) のそれぞれは二つの円が交差し合う 2 次元データ、円一つと直線一つの 2 次元データおよび、二つの直線が交差し合う 2 次元データである。Fig. (3) の (d) は球一つと円一つおよび線一つの 3 次元データを示す。Fig. (3) の (a)(b)(c) に対してクラス数を 2 とし、Fig. (3) の (d) に対して

クラス数を 3 とする。

ここで、クラスタリングの初期値依存問題に対して、各データにおいてラベルの初期化をランダムな 5 回の試行でそれぞれのクラスタリング手法の性能評価を行う。また、性能評価をするために、以下の alignment 関数<sup>20)</sup> を用いる。

$$A(Y_t, Y_e) = \frac{\sum_{k,l} y_{kl;t} y_{kl;e}}{\sqrt{\sum_{k,l} y_{kl;t}^2} \sqrt{\sum_{k,l} y_{kl;e}^2}} \in [0, 1]. \quad (31)$$

ただし、 $Y_t = [y_{kl;t}]$  と  $Y_e = [y_{kl;e}]$  のそれぞれはオリジナルのデータラベルと、クラスタリング手法による推定されたデータラベルに基づき作られた  $n$  行  $n$  列の対称行列である。 $y_{ij}$  はインスタンス  $k$  とインスタンス  $l$  とのラベル差を表わし、以下のより定義される。

$$y_{kl} = \begin{cases} 1 & (y_k = y_l), \\ 0 & (y_k \neq y_l). \end{cases} \quad (32)$$

Table (3) は k-means, HCSS および提案手法により 5 回試行の alignment の平均値と最大値を示す。 $\bar{A}$  は alignment の平均値であり、 $A_{max}$  は alignment の最大値である。Table (3) から、k-means または HCSS のクラスタリング結果と比べ、提案手法は高い alignment 値を得られたことを確認できる。

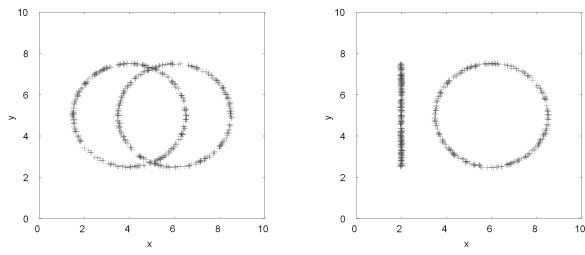
Table 3: Average and maximum alignments of k-means, HCSS and proposed method.

	k-means		HCSS		proposed	
	$\bar{A}$	$A_{max}$	$\bar{A}$	$A_{max}$	$\bar{A}$	$A_{max}$
2 circle	0.52	0.52	0.88	0.97	0.99	0.99
ten	0.74	0.74	0.70	0.99	0.90	1.00
cross	0.50	0.50	0.51	0.52	0.74	0.96
3d	0.38	0.39	0.47	0.66	1.00	1.00

本節では 5 回の試行のうち Eq. (31) の alignment 関数が最も高かったクラスタリング結果を用い、可視化し考察を行う。Fig. (4) と Fig. (5) および Fig. (6) のそれぞれは Fig. (3) で示したデータに対して k-means と HCSS および提案手法により 5 回のクラスタリングのうち、Eq. (31) の alignment 関数が最も高かった結果を示す。

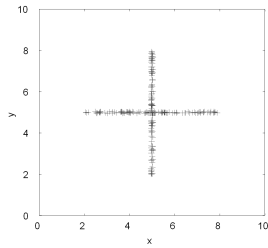
k-means によるクラスタリングに対して各クラスはユークリッド空間におけるベクトルを代表点として近似を行うため、球や円又は直線のような幾何的データについて Fig. (4) のクラスタリング結果によるデータ分布がうまく分類できないと示されている。

そして k-means の欠点に対して、円の複数個を持つ 2 次元データの場合、HCSS によるクラスタリング結果 (Fig. (5) の (a)) はうまく分類できたことを示している。しかし、HCSS では直線や平面に対するデータはうまくクラスタリングできない。直線や平面は半径が無限大の円や球の一部であるに関わらず、HCSS による近似方法はうまく推定されない。そのため、直線が含まれるデータに対して分類できない結果 (Fig. (5) の (b)(c)(d)) になった。Fig. (5) の (b) から、データ

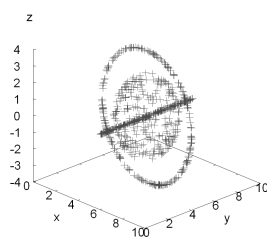


(a) 2 circle

(b) ten

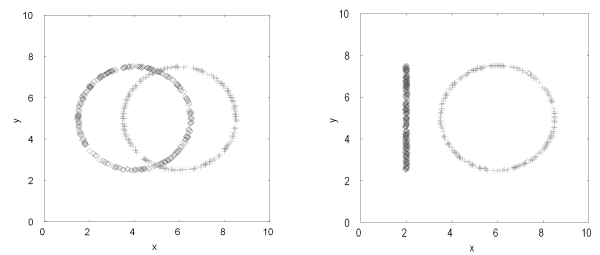


(c) cross



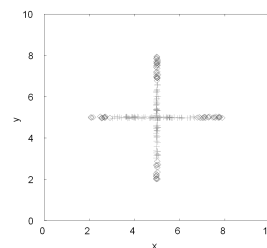
(d) 3d

Fig. 3: Geometric data for clustering.

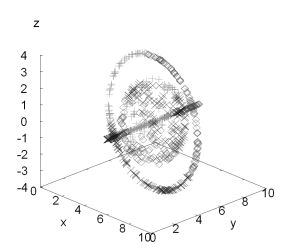


(a) 2 circle

(b) ten

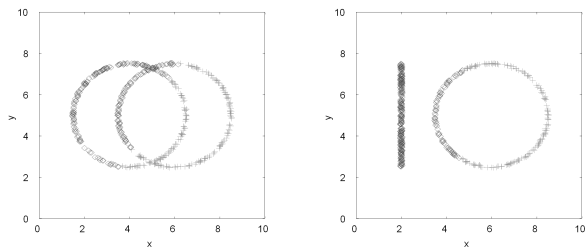


(c) cross



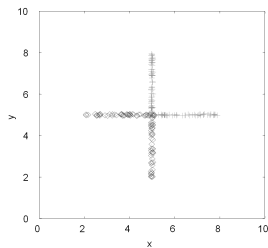
(d) 3d

Fig. 5: Clustering result using hard-C Spherical Shells method

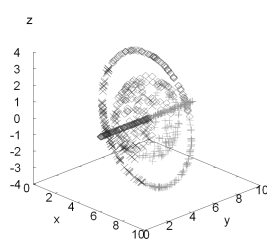


(a) 2 circle

(b) ten

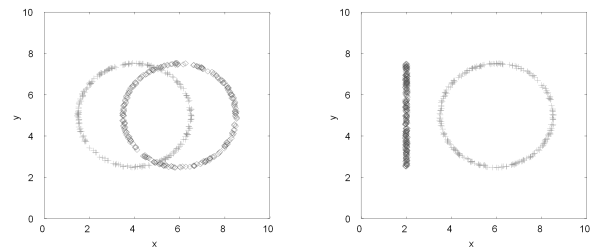


(c) cross



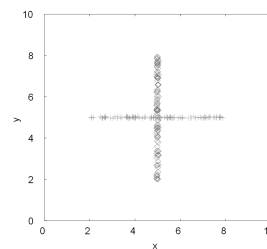
(d) 3d

Fig. 4: Clustering result using k-means method

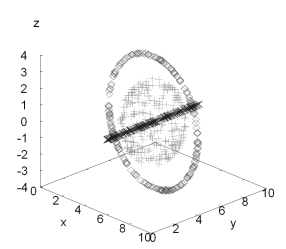


(a) 2 circle

(b) ten



(c) cross



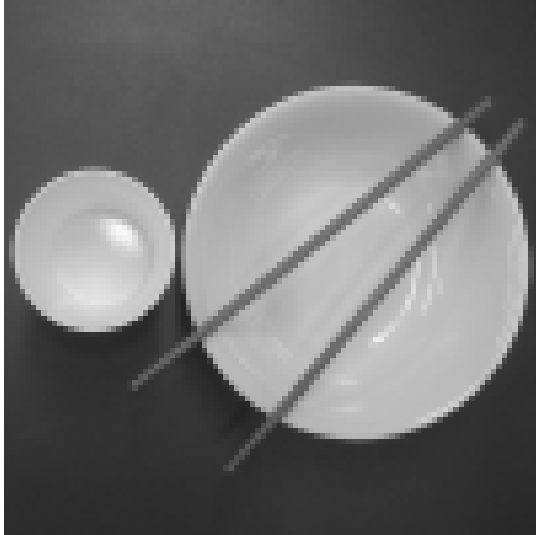
(d) 3d

Fig. 6: Clustering result using proposed method

はHCSSによって交差し合う2つの円で近似されたことを確認できる。また、Fig. (5)の(c)から、データはHCSSによって中心を共有する交差し合わない2つの円で近似されたことを確認できる。つまり、HCSSは直線に対して近似しようとならないことが分かる。同様にFig. (5)の(d)でのうまく分類できない結果に対して説明できるのである。

ここで、HCSSの欠点に対して、提案手法ではCGA

空間のベクトルが球・円と平面・直線の両方を表現することができるため、その結果直線や円や球が含まれるデータに対してうまくクラスタリングできる。Fig. (6)は提案手法によるクラスタリングがうまくできることを示す。Fig. (6)の(a)(b)(c)のそれぞれは提案手法により交差し合う二つの円のクラスタ、円一つのクラスタと直線一つのクラスタおよび、交差し合う二つ



(a) Original image



(b) Clustering result

Fig. 7: Clustering result for real image using the proposed method.

の直線のクラスタで分けられ、正しく分類できる. Fig. (6) の (d) は球と円および直線の3つクラスタにより分けられ、正しく分類できる.

ここで、実際に提案したクラスタリング手法を実画像に適用する. 図7は実画像に適用有効例の結果を示す. 図7(a)は $100 \times 100$ ピクセルのオリジナル画像であり, 図7(b)はLaplacianフィルタを用いたエッジ抽出及び2値化処理後のクラスタリング結果を示す. クラスタリング結果により提案手法を用いたクラスタリングはクラスタ数 $C = 4$ のとき, 画像から箸2本とコップ一つ及び茶碗一つを抽出できることを示し, 実データへの適用有効性を示す.

## 5 おわりに

本稿では, Conformal Geometric Algebra を用いて実空間ベクトルを Conformal ベクトルに変換し,  $m$  次元における超球を近似する手法を提案した. また, 提案した近似手法を用いて, 親関節を中心に回転中のターゲット関節の3次元座標を, 単カメラの2次元動画から推定するためのパーティクルフィルタを提案した. 提案したパーティクルフィルタは近似手法により, 回転軌道を通る球と推定できるため, 高い追跡精度が期待される. シミュレーション実験では, 回転クォータニオンを持つパーティクルフィルタ, 及び提案手法によるパーティクルフィルタを用いて, 回転中の子関節の座標を推定した. 従来手法と比べ, 提案手法は回転中の関節を高い精度かつ安定的に追跡することができることを示した. そして, 本稿では提案近似手法を用いて新しい幾何的クラスタリング手法を提案した. 提案クラスタリング手法は幾何的なデータに対して従来のクラスタリングと比べ優越な性能が人工データ及び実画像データにより確認できた.

## 参考文献

- 1) C. Doran and A. Lasenby, Geometric algebra for physicists, Cambridge University Press (2003)
- 2) D. Hestenes, New foundations for classical mechanics, Dordrecht (1986)
- 3) D. Hestenes and G. Sobczyk, Clifford Algebra to Geometric Calculus: A unified language for mathematics and physics, Reidel (1984)
- 4) I. Sekita, T. Kurita, and N. Otsu, Complex Autoregressive Model for Shape Recognition, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 14, No. 4 (1992)
- 5) A. Hirose, Complex-Valued Neural Networks: Theories and Applications, Series on Innovative Intelligence, Vol. 5 (2006)
- 6) N. Matsui, T. Isokawa, H. Kusamichi, F. Peper, and H. Nishimura, Quaternion neural network with geometrical operators, Journal of Intelligent and Fuzzy Systems, Volume 15, Numbers 3-4, 149/164 (2004)
- 7) S. Buchholz and N. Le Bihan, Optimal separation of polarized signals by quaternionic neural networks, 14th European Signal Processing Conference, EUSIPCO 2006, September 4-8, Florence, Italy (2006)
- 8) T. Nitta, An Extension of the Back-Propagation Algorithm to Complex Numbers, Neural Networks, Volume 10, Number 8, 1391/1415 (1997)
- 9) E. Hitzer, Quaternion Fourier Transform on Quaternion Fields and Generalizations, Advances in Applied Clifford Algebras, 17(3), 497/517 (2007)

- 10) G. Sommer, Geometric Computing with Clifford Algebras, Springer (2001)
- 11) M. T. Pham, K. Tachibana, E. M. S. Hitzer, T. Yoshikawa, and T. Furuhashi, Classification and Clustering of Spatial Patterns with Geometric Algebra, AGACSE 2008 Leipzig (2008)
- 12) M. T. Pham, K. Tachibana, T. Yoshikawa, and T. Furuhashi, Clustering Method for Geometric Data based on Approximation using Conformal Geometric Algebra, Fuzz IEEE 2011 Taipei (2011)
- 13) M. Isard and A. Blake, CONDENSATION-conditional density propagation for visual tracking. International Journal of Computer Vision, Volume 29, No 1, 5/28 (1998)
- 14) M. Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Transactions on Signal Processing, 50(2), 174/188 (2002)
- 15) Joao Luis Marins, Xiaoping Yun, Eric R. Bachmann, Robert B. McGhee, and Michael J. Zyda. An extended Kalman filter for quaternion-based orientation estimation using MARG sensors. In Proceedings of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003/2011 (2001)
- 16) Joseph J. LaViola Jr. A comparison of unscented and extended Kalman filtering for estimating quaternion motion. In Proceedings of the 2003 American Control Conference, volume 3, 2435/2440 (2003)
- 17) Yaakov Oshman and Avishy Carmi. Attitude estimation from vector observations using genetic-algorithm-embedded quaternion particle filter. Journal of Guidance, Control, and Dynamics, 29(4) (2006)
- 18) J. B. MacQueen, Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, 281/297 (1967)
- 19) R. Krishnapuram, O. Nasraoui and H. Frigui, The fuzzy c spherical shells algorithms: a new approach, IEEE Trans. on Neural Networks, vol.3, no.5, 663/671 (1992)
- 20) N. Cristianini, J. Kandola, A. Elisseeff, and J. Shawe-Taylor, On kernel target alignment. Journal of Machine Learning Research (2002)

# 断熱的量子計算におけるハミルトニアン変化の高速化に関する考察

金城光永 (琉球大学) 十川雄一郎 山下大輔 (琉球大学大学院理工学研究科)  
佐藤茂雄 (東北大学電気通信研究所) 島袋勝彦 (琉球大学)

## A study on increase in speed of Hamiltonian changes on adiabatic quantum computation

\*M. Kinjo (University of the Ryukyus)

Y. Sogawa D. Yamashita (Graduate School of Eng. and Sci., University of the Ryukyus)  
S. Sato (R.I.E.C., Tohoku University) K. Shimabukuro (University of the Ryukyus)

**Abstract**— The adiabatic quantum computation (AQC), proposed by Farhi et al., can be applied to various problems including non-deterministic polynomial time problems. And the neuromorphic adiabatic quantum computation (NAQC), the AQC with an efficient design method for a final Hamiltonian in consideration of the analogy with an artificial neural network, can be applied to optimization problems if its cost function can be expressed in a quadratic form. Its successful operations have been confirmed by numerical simulations. However, the general method of deciding an appropriate parameter  $T$  for a target problem is no exist. In this paper, we propose an appropriate  $T$  for 4-quantum bit system by controlling the interaction amplitudes between quantum bits.

**Key Words:** Adiabatic Quantum Computation, Neuromorphic Adiabatic Quantum Computation, Local Adiabatic Evolution, Neural Network

## 1 はじめに

量子効果を積極的に応用した量子コンピュータは、現在主流の汎用デジタルコンピュータと同等の計算を行うことができるだけでなく、重ね合わせの原理を利用した超並列計算など、汎用デジタルコンピュータでは実行困難な量子計算アルゴリズムを有し、その応用が期待されている<sup>1)</sup>。代表的な量子計算アルゴリズムは、特定用途向けとして、因数分解アルゴリズム<sup>2)</sup>やデータベース検索アルゴリズム<sup>3)</sup>が挙げられるが、より汎用性の高いものとして断熱的量子計算<sup>4)</sup> (Adiabatic quantum computation; AQC) も挙げられる。我々は、その AQC の汎用性をさらに高めるため、脳を工学的に模倣した人工ニューラルネットワークとのアナロジーを基に、断熱的变化の制約<sup>5)</sup>を除いた、ニューロ様断熱的量子計算 (Neuromorphic adiabatic quantum computation; NAQC) を提案した<sup>6, 7)</sup>。この手法は、シナプス結合を量子ビット間の相互作用と見做すことによって得られる量子システムの Hamiltonian を基に量子計算を行うことによって、組合せ最適化問題を解くことを可能とする。ただし、Hopfield 型ニューラルネットワークと同様に、対象問題に対する 2 次形式の評価関数が必要である。

断熱的量子計算では、対象問題を解くために最終 Hamiltonian を設計し、初期 Hamiltonian の断熱的な時間変化を利用するが、その時間変化を従来の等速度ではなく加速度を持った変化にすることで、計算時間を高速化できる可能性がある。しかしながら、実際の計算過程を考慮すると、計算途中で幾度も観測を行いながら計算を進める必要があり、その実現は困難であると考えられる。

そこで本稿では、別の観点から相互作用強度に着目し、その制御によって計算の速度向上を図る。いくつか

の問題について考察し、その可能性について検討する。

## 2 断熱的量子計算 (AQC)

AQC は、時間に依存する Schrödinger 方程式、

$$i\hbar \frac{\partial}{\partial t} \psi = H\psi, \quad (1)$$

において、Hamiltonian を次式、

$$H(t) = \left(1 - \frac{t}{T}\right) H_I + \frac{t}{T} H_F, \quad (2)$$

に従って断熱変化させることで計算を行う<sup>4)</sup>。ここで、 $H_I$  及び  $H_F$  はそれぞれ初期および最終ハミルトニアンである。 $H_I$  はその基底状態がすべての状態の重ね合わせ状態となるように、

$$|\psi(0)\rangle = \frac{1}{\sqrt{2^N}} \sum_{m=0}^{2^N-1} |m\rangle, \quad (3)$$

で与えられる。ここで  $N$  は量子ビット数であり、 $|m\rangle$  は  $m$  番目の固有ベクトルを表す。また  $H_F$  は基底状態の表現が対象問題の解を満たすように選択される。今、 $t=0$  の量子システムの状態が  $H_I$  の基底状態から始まるとすると、式 (3) で示されるようにすべての状態が初期状態  $|\psi(0)\rangle$  として設定される。 $T$  は断熱的变化の速度に関するパラメータであり、AQC の実行時間である。なお、十分に大きな  $T$  を選択することで、その変化は断熱的となる。断熱理論<sup>5)</sup>によれば、次の条件式、

$$\frac{|\langle \psi_1 | \frac{dH}{dt} | \psi_0 \rangle|}{\{\min(g(t))\}^2} \ll 1, \quad (4)$$

が満足されるならば，断熱的变化中の量子状態は常に基底状態近辺に停滞していることが保障されている．ここで，

$$g(t) = E_1(t) - E_0(t), \quad (5)$$

また， $E_i$  は  $i$  番目のエネルギー固有値である．したがって，エネルギー散逸が無く，かつエネルギー準位の交差が無いという制約が満たされるならば， $t = T$  のとき最終状態  $|\psi(T)\rangle$  は  $H_F$  の基底状態近辺にあり，その基底状態が対象問題の解を満たすように選択されているので，量子状態の観測の結果，最適解を得ることができる．

### 2.1 相互作用を利用した Hamiltonian の構成法

AQC では， $H_F$  をどのように対象問題へ対応させて構成するかが大きな課題である．その解決手法のひとつとして，脳を工学的に模倣した人工ニューラルネットワークを参考に，量子ビット間の相互作用を利用した Hamiltonian の構成方法がある<sup>6, 7)</sup>．これは，人工ニューロンの状態  $-1$  と  $1$  を，それぞれ量子状態の  $|0\rangle$  と  $|1\rangle$  に対応させ，また，ニューロン間の興奮性または抑制性結合を，量子ビット間の相互作用に見立てることにより，その相互作用を積極利用した構成法である．その構成法は， $N$  量子ビットの仮想システムに対して，次式，

$$H_F \equiv E \cdot I + A \cdot \sum_{ij, i \neq j}^N \{H_{ij}(w_{ij})^\dagger + H_{ij}(w_{ij})\}, \quad (6)$$

で定義される．ここで， $E$  および  $A$  はそれぞれ相互作用が無い場合のエネルギーおよび量子ビット間の相互作用強度を示す．また， $I$  は  $2^N \times 2^N$  の単位行列であり， $w_{ij}$  は  $j$  番目ニューロンから  $i$  番目ニューロンへのシナプス荷重値を示す．さらに  $H_{ij}(w_{ij})$  は，

$$H_{ij}(w_{ij}) = \begin{cases} a_i a_j & w_{ij} > 0 \\ a_i^\dagger a_j & w_{ij} \leq 0 \end{cases}, \quad (7)$$

で表されるような量子ビット間の相互作用を示す．なお， $a_i$  および  $a_i^\dagger$  はそれぞれ  $i$  番目の量子ビットに対する消滅および生成演算子のような演算子であり，

$$a = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad (8)$$

で表される．本稿では  $E = 0$  であるような単純な場合を考える．

### 2.2 局所断熱変化 (Local Adiabatic Evolution)

今，Hamiltonian の時間変化の式を，より一般的に，

$$H(t) = (1 - s(t)) H_I + s(t) H_F, \quad (9)$$

と表す．なお，従来の AQC では  $s(t) = t/T$  である．ここで，従来の AQC では式 (4) により， $T$  に関する条件式，

$$\frac{\left| \frac{ds}{dt} \langle \psi_1 | \frac{dH}{ds} | \psi_0 \rangle \right|}{\{\min(g(t))\}^2} = \frac{1}{T} \frac{\left| \langle \psi_1 | \frac{dH}{ds} | \psi_0 \rangle \right|}{\{\min(g(t))\}^2} \ll 1,$$

すなわち，

$$T \gg \frac{\left| \langle \psi_1 | \frac{dH}{ds} | \psi_0 \rangle \right|}{\{\min(g(t))\}^2}, \quad (10)$$

が導出できる．これより，第 1 励起状態と基底状態のエネルギー差  $g(t)$  の最小値が小さいと， $T$  は非常に大きな値となる．一方， $s(t)$  を非線形関数として選択すると， $s(t)$  の時間変化に関する新しい条件式，

$$\left| \frac{ds}{dt} \right| \ll \frac{g^2(t)}{\left| \langle \psi_1 | \frac{dH}{ds} | \psi_0 \rangle \right|}, \quad (11)$$

が得られる．これより， $g(t)$  の最小値に縛られることなく  $s(t)$  を設計できる．すなわち，Hamiltonian の時間変化を一定ではなく可変とすることで， $H_F$  への到達時間を短くすることができる．これを局所断熱変化<sup>8)</sup> と呼び，Roland らによって提案された．

## 3 Hamiltonian の時間変化によるエネルギー差の変化

前節で述べた AQC で実際に計算を行う場合，対象問題によって変化する  $H_F$  に応じて  $T$  あるいは  $s(t)$  を正確に設定することは一般的に困難である． $T$  に関しては，これまで必要以上に大きな  $T$  を設定するか，試行錯誤により最適値を発見していた．また， $s(t)$  に関しては，事前に解析的に得られる場合は希であり，一般的な手法は見つかっていない．すなわち，第 1 励起状態と基底状態のエネルギー差  $g(t)$  について， $g(0)$  は求まるが，それ以降の時間でリアルタイムに  $g(t)$  を測定しながら Hamiltonian を時間変化させることは困難である．ここで  $g(0)$  は  $H_I$  のエネルギー差を意味し，これは解析的に求めることができる．

そこで本稿では，相互作用強度  $A$  の変化が  $g(t)$  に与える影響を調査し， $g(t)$  の最小値を操作することで，AQC の高速化が可能かどうか考察する．

なお，以下の数値シミュレーションにおいて，便宜上  $\hbar = 1$  とした．

### 3.1 対象とする 4 量子ビットのネットワーク

本稿では簡単な例として Fig. 1(a) ~ (f) で示される 4 量子ビットのネットワークを考える．Fig. 1 に示される丸印は量子ビットを示し，それらを繋ぐ実線および破線はそれぞれ興奮性および抑制性の相互作用を示す．Fig. 1(a) および (b) は，量子ビット間の相互作用の有無は同一だが，すべて興奮性か，一部抑制性を含むかの違いを持つ．同様な関係が Fig. 1(c) および (d)，Fig. 1(e) および (f) にある．なお，Fig. 1(a) から (f) のそれぞれのネットワークに対応するシナプス荷重値行列  $W_a$  から  $W_f$  は，

$$W_a = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad (12)$$

$$W_b = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad (13)$$

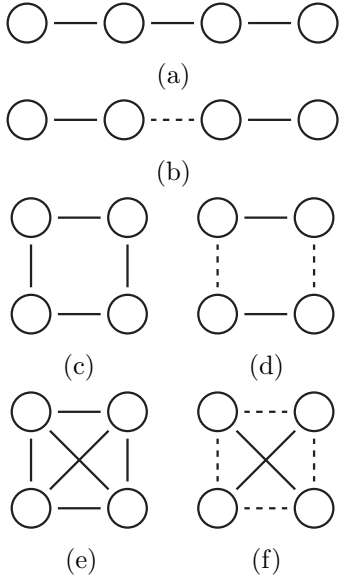


Fig. 1: 4-qubit network model. The solid and dashed line denote the excitatory and inhibitory interaction, respectively.

$$W_c = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad (14)$$

$$W_d = \begin{pmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}, \quad (15)$$

$$W_e = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \quad (16)$$

$$W_f = \begin{pmatrix} 0 & -1 & 1 & -1 \\ -1 & 0 & -1 & 1 \\ 1 & -1 & 0 & -1 \\ -1 & 1 & -1 & 0 \end{pmatrix}, \quad (17)$$

である．これらの荷重値行列からそれぞれ対応する  $H_F$  を構成し，相互作用強度  $A$  とエネルギー差  $g(t)$  の比較検討を行う．

### 3.2 相互作用の性質が異なる場合

まず，Fig. 1(a) および (b) のモデルについて，興奮性相互作用のみの場合と，抑制性相互作用を含む場合の， $A$  と  $g(t)$  の最小値の関係を Fig. 2 に示す．ここで，記号  $\circ$  および  $+$  は，それぞれモデル (a) および (b) の結果を示す．これより， $A$  を従来の 1 から大きくすると， $g(t)$  の最小値も大きくなり，AQC の高速化が可能となることが分かる．また， $A$  を変化させても両者の違いは無いと言える．なお，この性質はモデル (c) と (d)，およびモデル (e) と (f) についても同様であった．さらに， $g(t)$  の最小値は増加したとしても 2 が上限となっている．その上限は，初期 Hamiltonian である  $H_I$  の第 1 励起状態と基底状態のエネルギー差に等しい．すなわち，

$$\min\{g(t)\} = g(0) \quad (18)$$

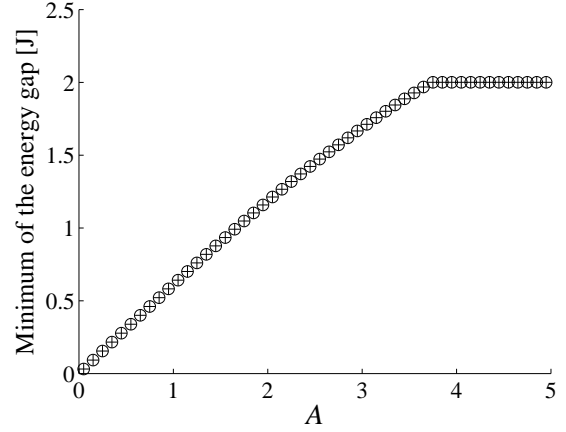


Fig. 2: Minimum of  $g(t)$  as a function of  $A$ .  $\circ$  and  $+$  denote the network model (a) and (b) cases, respectively.

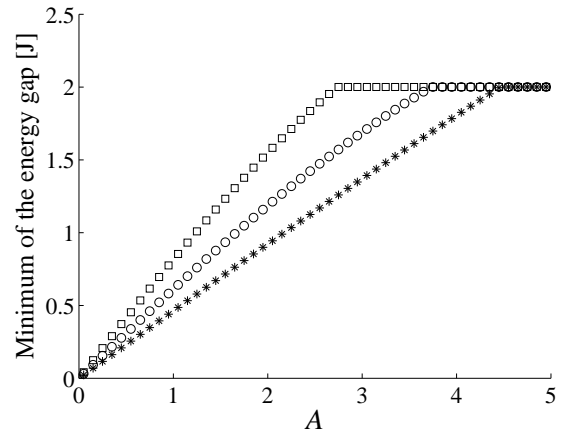


Fig. 3: Minimum of  $g(t)$  as a function of  $A$ .  $\circ$ ,  $\square$ , and  $*$  denote the network model (a), (c), and (e) cases, respectively.

となり， $g(0)$  は解析的に求められるので，試行錯誤することなく AQC の実行時間  $T$  を決定できる．

### 3.3 相互作用の箇所が異なる場合

前述の相互作用の性質が異なっても同様の結果となることを踏まえて，興奮性相互作用のみのモデル (a)，(c)，および (e) について，モデルが異なる（相互作用の箇所が異なる）場合の， $A$  と  $g(t)$  の最小値の関係を Fig. 3 に示す．ここで，記号  $\circ$ ， $\square$ ，および  $*$  は，それぞれモデル (a)，(c)，および (e) の結果を示す．これより，前述の結果と同様に  $A$  を増加させると， $g(t)$  の最小値も増加して，AQC の高速化が可能となることが分かり，その上限も 2 となっている．ただし，上限に到達する際の  $A$  はモデルにより異なっている．モデル (a)，(c)，および (e) の相互作用の箇所は，それぞれ 3，4，および 6 カ所と，徐々に複雑になっていると考えられるが，その複雑さと， $g(t)$  最小値の上限到達時の  $A$  に必ずしも相関はない．

## 4 エネルギー準位の時間変化

前述のモデル (a) において，相互作用強度  $A=1, 2$ ，および 4 の際の，時間  $t$  に対する第 1 励起状態および基底状態のエネルギーの遷移を，それぞれ Fig. 4, 5，



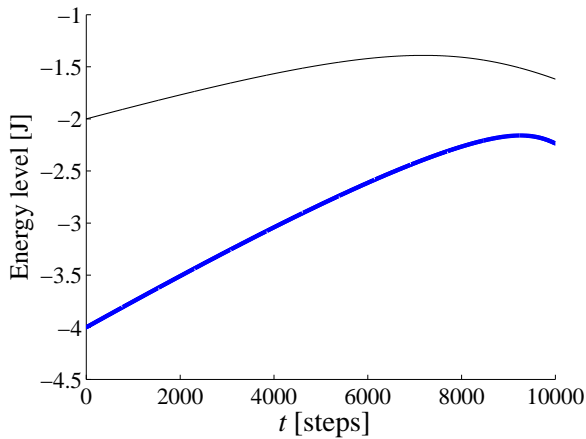


Fig. 4: The ground (thick line) and the 1st excited state (thin line) energy level changes with  $A = 1$ .

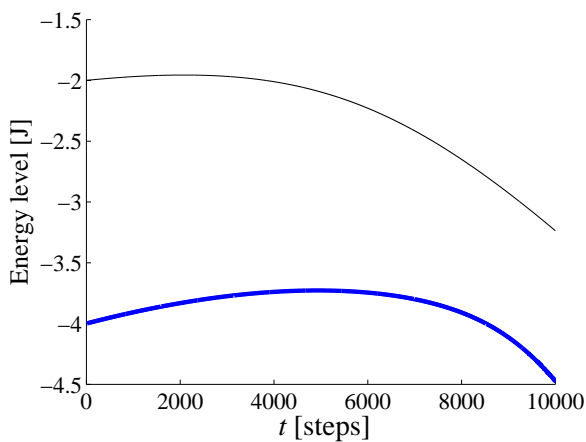


Fig. 5: The ground (thick line) and the 1st excited state (thin line) energy level changes with  $A = 2$ .

および6に示す．なお，ここではエネルギー準位の遷移を確認する目的で，実行時間  $T$  を非常に大きな値  $T = 10000$  とした．Fig. 4, 5, および6のいずれの場合もエネルギー交差はなく，AQCの制約条件のひとつを満足している．また，Fig. 4において， $g(t)$ の最小値がAQC終了間際にあるため，実際にAQCを実行する際，試行錯誤にて適切な  $T$  を調べる必要がある．相互作用強度  $A$  が2倍となったFig. 5においても， $g(T)$ が大きくなったとはいえ，Fig. 4と同様に適切な  $T$  を調べる必要がある．しかしながら，Fig. 6については， $g(0)$ が最小値となっており，その値から  $T$  を決定することで，Fig. 4および5より早く計算を終えることができる．

## 5 まとめ

断熱的量子計算を実際に行う際，Hamiltonianの断熱変化時間の設定が問題となるが，4量子ビットのネットワークにおいて，量子ビット間の相互作用強度を強くすることで計算の高速化を図ることができることを確認した．その際，相互作用の興奮性および抑制的な性質には依存しないことがわかった．今後の課題として，この結果を定量的に評価するために，異なる量子ビット数や異なるネットワークモデルでの調査を行うことが挙げられる．

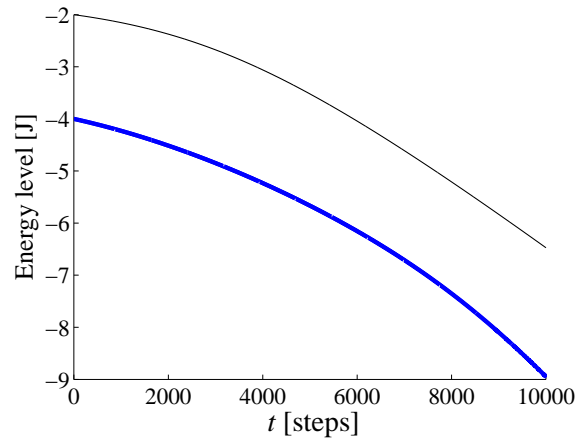


Fig. 6: The ground (thick line) and the 1st excited state (thin line) energy level changes with  $A = 4$ .

## 参考文献

- 1) M. A. Nielsen and I. L. Chuang : Quantum Computation and Quantum Information, Cambridge University Press, Cambridge (2000).
- 2) P. W. Shor : Polynomial-time algorithm for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Comput.*, **26**, 1484/1509 (1997).
- 3) L. K. Grover : A fast quantum mechanical algorithm for database search, *Proc. Twenty-Eighth Annual ACM Symp. on the Theory of Computing*, 212/219 (1996).
- 4) E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda : A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem, *Science*, **292**, 472/475 (2001).
- 5) A. Messiah : Quantum Mechanics, 747/755, Dover, New York (1999).
- 6) S. Sato, M. Kinjo, and K. Nakajima : An Approach for Quantum Computing using Adiabatic Evolution Algorithm, *Jpn. J. Appl. Phys.*, **42**, 7169/7173 (2003).
- 7) M. Kinjo, S. Sato, Y. Nakamiya, and K. Nakajima : Neuromorphic quantum computation with energy dissipation, *Physical Review A*, **72**, 052328 (2005).
- 8) J. Roland and N. J. Cerf : Quantum search by local adiabatic evolution, *Physical Review A*, **65**, 042308 (2002).



# リアプノフ法で保証された3次元追跡 Eye-Vergence ビジュアルサーボ実験の周波数応答

○于 福佳、松本 紘明、矢納 陽、見浪 護（岡山大学）  
宋 薇（上海大学）

## Frequency Response Experiments of Lyapunov-guaranteed 3-D Visual Servoing Experiments with Eye-Vergence

\*Fujia Yu Hiroaki Matsumoto Akira Yanou Mamoru Minami (Okayama University)  
Wei Song (Shanghai University)

**Abstract**– In this paper, we use on-line 1-step genetic algorithm (1-step GA) to recognize the 6-D object in a much bigger space comparing with other recognition methods, then we guarantee the convergence of this method by Lyapunov theorem. But visual servoing methods for hand-eye configuration have been presented so far seems to be vulnerable for tracking ability since it may lose a moving target. Our proposal to solve this problem is that the controller for visual servoing of the hand and the eye-vergence should be separated independantly based on decoupling each other. Base on this prerequisite the eye-vergence system to track target object in camera in view sight (trackability) can be much faster than conventional visual servoing with fixed cameras.

**Key Words:** Visual servoing, 1-step GA, Lyapunov theorem

### 1 Introduction

Comparing with the visual servoing systems whose cameras are static to the world, the hand-eye visual servoing systems are easy to become unstable because of the disturb from the motion of the end-effector. Otherwise, the hand-eye visual servoing systems can change its angle of view easily. In our research, we use the hand-eye system with two cameras.

On the other hand, a fixed-hand-eye system has some disadvantages, making the observing ability deteriorated depending on the relative geometry of the camera and the target. Such as: the robot cannot observe the object well when it is near the cameras (Fig. 1 (a)), small intersection of the possible sight space of the two cameras (Fig. 1 (b)), and the image of the object cannot appear in the center of both cameras, so we could not get clear image information of target and its periphery, reducing the pose measurement accuracy (Fig. 1 (c)). To solve the problems above, in this paper, we give the cameras an ability to rotate themselves to see target at center of the images. There is no research using such rotatable hand-eye system as far as we know. Thus it is possible to change the pose of the cameras in order to observe the object better, as it is shown in Fig. 2, enhancing the measurement accuracy in trigonometric calculation and peripheral distortion of camera lens by observing target at the center of lens. Moreover, recent researches on visual servoing are limited generally in a swath of tracking an object while keeping a certain constant distance<sup>3), 4), 5)</sup>. But the final objective of visual servoing lies in approaching the end-effector to a target and then work on it, like grasping. In this case, the desired relation between the cameras and the object is time varying, so such rotational camera system in Fig. 2 is required to keep suitable viewpoint all the time during the visual servoing application.

In visual servoing application, it is important to keep the

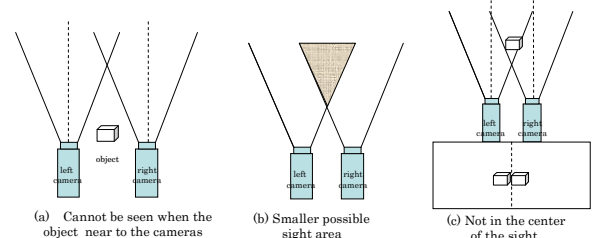


Fig. 1: Disadvantage of fix camera system

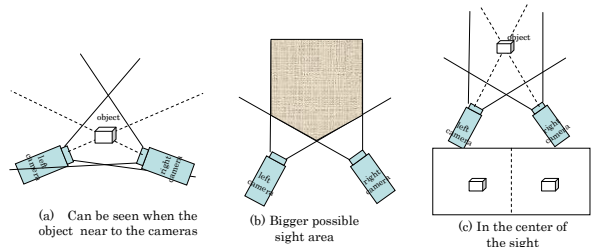


Fig. 2: Advantage of Eye-vergence system

object in the visual eye sight to make the visual feed back to keep stable closed-loop dynamical motion. If the camera lose the sight of target, its pose cannot be measured, that means, the visual feedback is cut, and the robot may fall in some unexpected motion, being dangerous. As it is shown in Fig. 3 (a), in visual servoing system the cameras can keep staring at the object at first in (a), but when the target moves so fast that the manipulator can not catch up the speed of the target because of the big mass of whole manipulator itself, then the object may disappear in the sight of the cameras, resulting in that the visual feedback of the system is cut as shown in (b), loosing feedback information that appears most dangerous. So in visual servoing system it is very important to keep the camera tracking

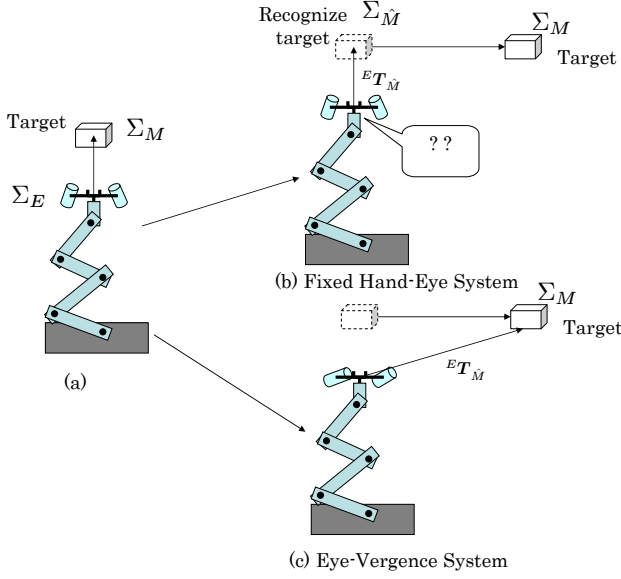


Fig. 3: Dynamics advantage of Eye-vergence system

the target. A system with high tracking ability also has better security and validity. To realize this stable tracking ability against quick and unknown motion of the target, we propose to control the cameras and the manipulator separately. Because of the small mass and inertia moment of the cameras, it can track the target better, as in Fig. 3 (c), like animals tracks target with eye motion before rotate their heads to the target to improve dynamical tracking ability.

To evaluate the observation of the camera, we put forward a concept of trackability. This concept has been used in <sup>6)</sup>, where trackability is defined as a kinematic function of singular value of Jacobian matrix connecting hand's velocities and angular joint velocities, ignoring the relationship between the hand and the target objects, including the both dynamical motion of the target and the manipulator, which seems to be essential for evaluating the eye-vergence visual servoing. Then we define a new concept of trackability to evaluate our visual servoing.

As shown in Fig.4, the proposed method includes two loops: a loop for conventional visual servoing that direct a manipulator toward a target object and an inner loop for active motion of binocular camera for accurate and broad observation of the target object. We set relatively high gain to the eye-vergence controller to put the priority to the 3D pose tracking to improve the system trackability.

## 2 On-line Evolutionary Recognition

### 2.1 3-D Model-based Matching

We use a model-based matching method to recognize a target object in a 3-D searching area. A solid models is located in  $\Sigma_E$ , its position and orientation are determined by six parameters,  $\psi = [r^T, \epsilon^T]^T$ , where  $r = [x, y, z]^T$ ,  $\epsilon = [\epsilon_1, \epsilon_2, \epsilon_3]^T$ . Here, the target's orientation is represented by unit quaternion <sup>1)</sup>, which has an advantage that can represent the orientation of a rigid body without sin-

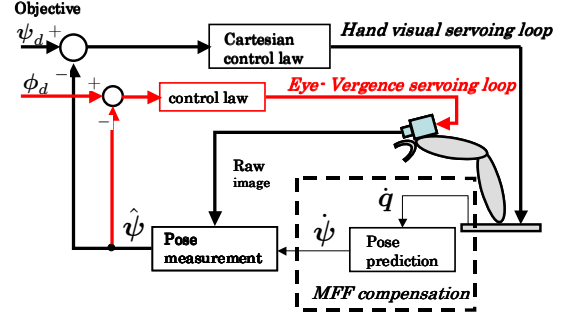


Fig. 4: Hand & Eye Visual servo system

gularities, when  $-\pi < \theta < \pi$  ( $\theta$  is defined below). The unit quaternion, is defined as

$$Q = \{\eta, \epsilon\}, \quad (1)$$

where

$$\eta = \cos \frac{\theta}{2}, \quad \epsilon = \sin \frac{\theta}{2} \mathbf{k}, \quad (2)$$

here,  $\mathbf{k}(\|\mathbf{k}\| = 1)$  is the rotation axis and  $\theta$  is the rotation angle.  $\eta$  is called the scalar part of the quaternion while  $\epsilon$  is called the vector part of the quaternion. They are constrained by

$$\eta^2 + \epsilon^T \epsilon = 1. \quad (3)$$

In (3)  $\eta$  can be calculated by  $\epsilon$ , so we just use three parameters  $\epsilon$  to represent an orientation.

The left and right input images from the stereo cameras are directly matched by the left and right searching models, which are projected from 3-D model onto 2-D image plane. The matching degree of the model to the target can be estimated by a correlation function between them as  $F(\psi)$  by using the color information of the target. Please refer to <sup>9)</sup> for a detailed definition of  $F(\psi)$ . When the searching models fit to the target objects being imaged in the right and left images,  $F(\psi)$  gives the maximum value. Therefore the 3-D object's position/orientation measurement problem can be converted to a searching problem of  $\psi$  that maximizes  $F(\psi)$ . We solve this optimization problem by 1-step GA method that will be explained in the next section.

### 2.2 GA-based On-line Recognition "1-step GA"

Theoretically optimal pose  $\psi_{max}(t)$  that gives the highest peak of  $F(\psi(t))$  is defined as

$$\psi_{max}(t) = \{\psi(t) \mid \max_{\psi \in \mathbf{L}} F(\psi(t))\} \quad (4)$$

where  $\mathbf{L}$  represents 6-DoF searching space of  $x, y, z, \epsilon_1, \epsilon_2, \epsilon_3$ .

An individual of GA is defined as  $\psi_i^j(t)$ , which means the  $i$ -th gene ( $i = 1, 2, \dots, p$ ) in the  $j$ -th generation, to search  $\psi_{max}(t)$ . Denote  $\psi_{max}^{GA}(t)$  to be the maximum among the  $p$  genes of  $\psi_i^j(t)$  in GA process,

$$\psi_{max}^{GA}(t) = \{\psi_i^j(t) \mid \max_{\psi_i^j \in \mathbf{L}} F(\psi_i^j(t))\}. \quad (5)$$

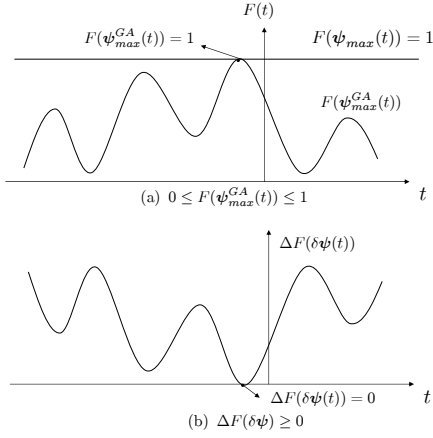


Fig. 5:  $\Delta F(\delta\psi(t))$  is positive definite

In fact we cannot always guarantee the best individual of GA  $\psi_{max}^{GA}(t)$  should coincide with the theoretically optimal pose  $\psi_{max}(t)$ , because the number of GA's individuals is not infinite. The difference between  $\psi_{max}(t)$  and  $\psi_{max}^{GA}(t)$  is denoted as

$$\delta\psi(t) = \psi_{max}(t) - \psi_{max}^{GA}(t). \quad (6)$$

And the difference between  $F(\psi_{max}(t))$  and  $F(\psi_{max}^{GA}(t))$  is denoted as

$$\Delta F(\delta\psi(t)) = F(\psi_{max}(t)) - F(\psi_{max}^{GA}(t)), \quad (7)$$

Since  $F(\psi_{max}(t)) \geq F(\psi_{max}^{GA}(t))$ , we have

$$\Delta F(\delta\psi(t)) \geq 0. \quad (8)$$

Based on the definition of  $\Delta F(\delta\psi(t))$  in (7), in this research, we let GA work in the following way:

- (a) GA evolves to minimize  $\Delta F(\delta\psi(t))$ .
- (b) The elitist individual of GA is preserved at every generation (elitist gene preservation strategy).
- (c)  $\psi_{max}^{GA}(t)$  does keep the same value in the evolving when the evolved new gene with different value gives the same value of  $\Delta F$ .

Here, we present two assumptions.

**[Assumption 1]**  $\Delta F(\delta\psi(t))$  is positive definite.

This means the distribution of  $F(\psi(t))$  satisfies  $\Delta F(\delta\psi(t)) = 0$  if and only if  $\delta\psi(t) = 0$ , which indicates  $\Delta F(\delta\psi(t)) = 0$  has a sole minimum at  $\delta\psi(t) = 0$  over the searching space  $\mathbf{L}$ , even though  $\Delta F$  is multi-peak distribution having peaks and bottoms with limited number. When the model overlap the target object in the image, then the situation can make  $\Delta F$  have the sole minimum in  $\mathbf{L}$ .  $0 \leq F(\psi(t)) \leq 1$ , since  $F(\psi(t))$  is normalized to be less than 1 and negative value to be set as zero by a definition of correlation function  $F(\psi(t))$ <sup>2</sup>. So the fitness function is always less than 1 except only one point which means the  $\psi_{max}^{GA}(t)$  can express the target object's pose, as shown in Fig. 5(a). From (7), we can see when  $\psi_{max}^{GA}(t) = 1$ ,  $\Delta F(\delta\psi(t)) = 0$  (Fig. 5(b)), which means

that only in this case,  $\psi_{max}^{GA}(t)$  can express the actual pose of the target object.

**[Assumption 2]**  $\dot{F}(\psi_{max}^{GA}(t)) \geq 0$ .

This means GA evolves itself to get a bigger fitness function value ( $\dot{F}(\psi_{max}^{GA}(t)) > 0$ ) or keep a same value ( $\dot{F}(\psi_{max}^{GA}(t)) = 0$ ). It is not only an assumption but also the character of GA if the target object is static, because the elitist individual is preserved in every generation of GA. However, when the target object is moving,  $\dot{F}(\psi_{max}^{GA}(t)) \geq 0$  will indicate that the convergence speed to the target in the dynamical images should be faster than the moving speed of the target object. Furthermore, with the pose tracking in dynamic scene being input at a certain video rate, this assumption means that  $F(\psi_{max}^{GA}(t))$  have the tendency of approaching to  $F(\psi_{max}(t))$ , and  $\psi_{max}^{GA}(t)$  moves toward  $\psi_{max}(t)$  in each period of the input image, or keeps a distance to  $\psi_{max}(t)$ . Since in this paper we think that the object's motion is enough slow comparing the calculation speed of GA's evolving to find  $F(\psi_{max}^{GA}(t))$  from the view point that the one image be input every input video period and evolving iterations in input video period are enough to catch up with the  $F(\psi_{max}^{GA}(t))$  being stationary during the input video period.

Differentiating (7) by time t, we have

$$\Delta \dot{F}(\delta\psi(t)) = \dot{F}(\psi_{max}(t)) - \dot{F}(\psi_{max}^{GA}(t)). \quad (9)$$

We defined  $F(\psi_{max}(t)) = 1$  representing that the true pose of the target object gives the highest peak. Therefore, the time differentiation of  $F(\psi_{max}(t))$  will be  $\dot{F}(\psi_{max}(t)) = 0$ . Thus, from (9) and [Assumption 2], we have

$$\Delta \dot{F}(\delta\psi(t)) = -\dot{F}(\psi_{max}^{GA}(t)) \leq 0. \quad (10)$$

$\psi_{max}^{GA}(t)$  represents current best GA solution. [Assumption 2] means GA can change its best gene  $\psi_{max}^{GA}(t)$  to always reduce the value of  $\Delta F$  regardless of dynamic image or static one, which indicates that the convergence speed to the target in the dynamically continuous images should be faster than the moving speed of the target object.

We cannot guarantee that the above two assumptions always hold, since they depend on some factors such as object's shape, object's speed, definition of  $F(\psi(t))$ , parameters of GA and viewpoint for observing, lightening environment, computer's performance et al. However, we can make efforts to improve the environment and correlation function and so on. Providing above two assumptions be satisfied, (8) and (10) hold, then  $\Delta F(\delta\psi(t))$  is so-called Lyapunov function. The objective here is to verify that  $\delta\psi(t)$  asymptotically stable, resulting in it converges to  $\mathbf{0}$  by using the Lyapunov function of  $\Delta F(\delta\psi(t))$ , meaning  $\psi_{max}^{GA}(t) \rightarrow \psi_{max}(t)$ , ( $t \rightarrow \infty$ ), and the following shows how to verify it.

Since  $\Delta \dot{F}(\delta\psi(t))$  is only negative semi-definite, in the view of LaSalle theorem,  $\delta\psi(t)$  asymptotically converges to the invariant set of the solutions  $\delta\psi$  satisfying  $\Delta \dot{F}(\delta\psi(t)) = 0$ . Considering the following expression,

$$\Delta \dot{F}(\delta\psi(t)) = \frac{\partial \Delta F}{\partial \delta\psi} \cdot \delta\dot{\psi} = 0, \quad (11)$$

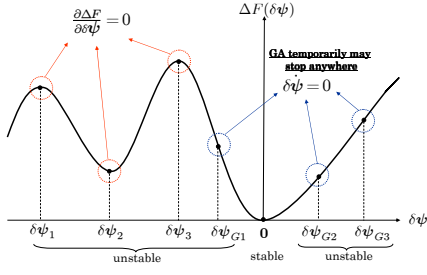


Fig. 6: The invariant set of the solutions of  $\Delta\dot{F}(\delta\psi(t)) = 0$ .

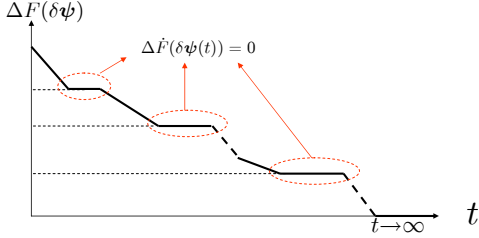


Fig. 7: The changing of  $\Delta F(\delta\psi(t))$  with respect to time  $t$  in the whole GA's evolution.

the first part  $\partial\Delta F/\partial\delta\psi$  describes partial differentiation of  $\Delta F$  with respect to  $\delta\psi$ , implying steepest descending direction of  $\Delta F$  in the space of  $\delta\psi$ ; the second part  $\delta\dot{\psi}$  describes the difference between the moving speed of the target object and the evolution speed of the best gene of GA, by the definition in (6).

Equation (11) shows the invariant set of the solutions of  $\Delta\dot{F}(\delta\psi(t)) = 0$  includes (1):  $P_1$ , the solution set of  $\partial\Delta F/\partial\delta\psi = 0$ ; (2):  $P_2$ , the solution set of  $\delta\dot{\psi} = 0$ ; and (3):  $P_3$ , the solution set satisfying  $\partial\Delta F/\partial\delta\psi \neq \mathbf{0}$ ,  $\delta\dot{\psi} \neq \mathbf{0}$ , but their inner product is 0.

As shown in Fig. 6,  $P_1$  includes the points of  $\delta\psi$  that give the local maximum or minimum values of the function  $\Delta F$  including  $\mathbf{0}$ . The number of these points is finite by [Assumption 1] denoted by  $p$ , that is

$$P_1 = \{\mathbf{0}, \delta\psi_1, \delta\psi_2, \dots, \delta\psi_{p-1}\}. \quad (12)$$

Concerning (1), the evolving process of GA may stay temporarily at the same  $\Delta F$  value. If the target object is static, it means the best gene of GA stops at some moments for the reason that the limited individuals of GA could not improve a current solution that gives a smaller fitness function value  $\Delta F$  during some generations. And when the target object is moving,  $\delta\dot{\psi} = 0$  means at these moments that the evolution speed of the best gene of GA is equal to the moving speed of the target object, by (6). The number of these points is assumed to be possibly finite, denoted by  $q$ . Thus, we describe the set of  $P_2$  as

$$P_2 = \{\mathbf{0}, \delta\psi_{G1}, \delta\psi_{G2}, \dots, \delta\psi_{G(q-1)}\}. \quad (13)$$

Notice that there is another solution set of  $\delta\dot{\psi}$ :  $P_3$ . In this case, the vector of  $\partial\Delta F/\partial\delta\psi$  is vertical to the vector of  $\delta\dot{\psi}$  since the calculation  $(\Delta F/\delta\psi) \cdot \delta\dot{\psi}$  in (11) means inner cross product, which means GA evolves in the direction that keeps a same fitness function value  $\Delta F$ . This

GA's evolution way is forbidden in this research for the GA work rule (c) that we have stated above. Then,  $P_3$  is null. So the invariant set that  $\delta\psi(t)$  asymptotically converges to is

$$P = P_1 \cup P_2. \quad (14)$$

Here,  $\delta\psi_1, \delta\psi_2, \dots, \delta\psi_{p-1}$  in  $P_1$  are all unstable since  $\delta F(\delta\psi_i) > 0$  ( $i = 1, 2, \dots, p-1$ ), and only  $\delta\psi = \mathbf{0}$  is stable from [Assumption 1], since when  $t \rightarrow \infty$  there should always remain the possibility to get out of local maximum/minimum of  $\delta\psi_1 \dots \delta\psi_{p-1}$ . And all the points in  $P_2$  except the point  $\mathbf{0}$  are unstable because GA has possibility to get out of these points by its evolving nature. Therefore,  $\mathbf{0}$  is the only stable point in the invariant set of  $P$ , that is,  $\delta\psi(t)$  will finally converges to  $\mathbf{0}$ . The image of the changing of  $\Delta F(\delta\psi(t))$  with respect to time  $t$  in the whole GA's evolution is shown in Fig.7.

The above verification shows  $\delta\psi(t) \rightarrow \mathbf{0}$ , which means

$$\psi_{max}^{GA}(t) \rightarrow \psi_{max}(t), \quad (t \rightarrow \infty) \quad (15)$$

Let  $t_\epsilon$  denotes a convergence time, then

$$|\delta\psi(t)| = |\psi_{max}(t) - \psi_{max}^{GA}(t)| \leq \epsilon, \quad (\epsilon > 0, t \geq t_\epsilon) \quad (16)$$

In (16),  $\epsilon$  is tolerable extent that can be considered as a observing error. Thus, it is possible to realize real-time optimization, because  $\psi_{max}^{GA}(t)$  can be assumed to be in the vicinity of the theoretically optimal  $\psi_{max}(t)$  after  $t_\epsilon$ .

Above discussion is under the condition of continuous time. Here, when we consider evolution time of each generation of GA denoted by  $\Delta t$ . The GA's evolving process is described as

$$\psi_i^j(t) \xrightarrow{\text{evolve}} \psi_i^{j+1}(t + \Delta t). \quad (17)$$

Obviously, this time-discrete evolution with the interval of time  $\Delta t$  may enlarge the recognition error  $\delta\psi(t)$ . Should this undesirable influence of  $\Delta t$  be considered, the tolerable pose error  $\epsilon$  will expand to  $\epsilon'$  as,

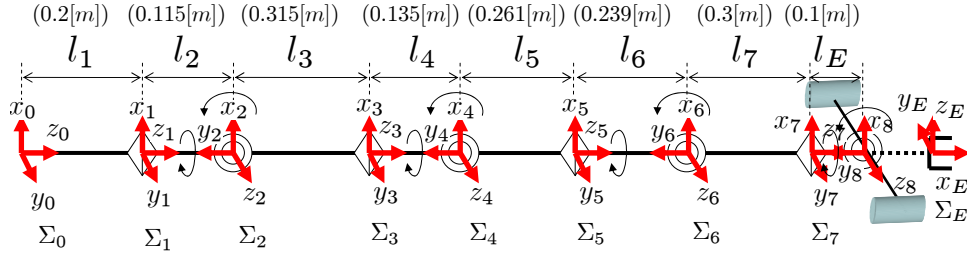
$$|\delta\psi(t)| \leq \epsilon', \quad (\epsilon' > \epsilon > 0). \quad (18)$$

Since the GA process to recognize the target's pose at the current time is executed at least one time with the period of  $\Delta t$  as the current quasi-optimal pose  $\psi_{max}^{GA}(t)$  is output synchronously, we named this on-line recognition method as "1-step GA". We have confirmed that the above real-time optimization problem could be solved by "1-step GA" through several experiments to recognize swimming fish <sup>7)</sup>, human face <sup>3)</sup> and rectangular solid block <sup>8)</sup>.

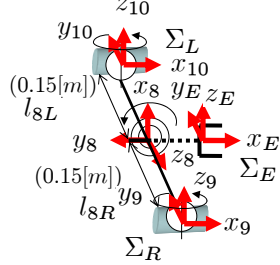
### 3 Hand & Eye Visual Servoing

#### 3.1 Experiment circumstance

The Mitsubishi PA-10 robot arm is a 7-DoF robot arm manufactured by Mitsubishi Heavy Industries, as shown in Fig. 8 (a). Two rotatable cameras with two pan angles and one sharing tilt angle mounted on the end-effector are FCB-1X11A manufactured by Sony Industries (Fig.



(a) Sketch map of the manipulator



(b) Sketch map of camera system

Fig. 8: Sketch map of the eye-vergence system

8 (b)). The frame frequency of stereo cameras is set as 33fps. The image processing board, CT-3001, receiving the image from the CCD camera is connected to the DELL WORKSTATION PWS650 (CPU: Xeon, 2.00 GHz) host computer.

### 3.2 Desired-trajectory generation

As shown in Fig. 9, the world coordinate frame is denoted by  $\Sigma_W$ , the target coordinate frame is denoted by  $\Sigma_M$ , and the desired and actual end-effector coordinate frame is denoted by  $\Sigma_{Ed}$ ,  $\Sigma_E$  separately. The desired relative relation between the target and the end-effector is given by Homogeneous Transformation as  ${}^{Ed}T_M$ , the relation between the target and the actual end-effector is given by  ${}^E T_M$ , then the difference between the desired end-effector pose  $\Sigma_{Ed}$  and the actual end-effector pose  $\Sigma_E$  is denoted as  ${}^E T_{Ed}$ ,  ${}^E T_{Ed}$  can be described by:

$${}^E T_{Ed}(t) = {}^E T_M(t) {}^{Ed} T_M^{-1}(t) \quad (19)$$

(19) is a general deduction that satisfies arbitrary object motion  ${}^W T_M(t)$  and arbitrary visual servoing objective  ${}^{Ed} T_M(t)$ . However, the relation  ${}^E T_M(t)$  is only observed by cameras using the on-line model-based recognition method and 1-step GA<sup>3), 9)</sup>. Let  $\Sigma_{\hat{M}}$  denote the detected object, there always exist an error between the actual object  $\Sigma_M$  and the detected one  $\Sigma_{\hat{M}}$ . So in visual servoing, (19) will be rewritten based on  $\Sigma_{\hat{M}}$  that includes the error  ${}^M T_{\hat{M}}$ , as

$${}^E T_{Ed}(t) = {}^E T_{\hat{M}}(t) {}^{Ed} T_{\hat{M}}^{-1}(t), \quad (20)$$

where  ${}^E T_{\hat{M}} = {}^E T_M$  determined by the given visual servoing objective. Differentiating (20) with respect to time yields

$${}^E \dot{T}_{Ed}(t) = {}^E \dot{T}_{\hat{M}}(t) {}^{\hat{M}} T_{Ed}(t) + {}^E T_{\hat{M}}(t) {}^{\hat{M}} \dot{T}_{Ed}(t), \quad (21)$$

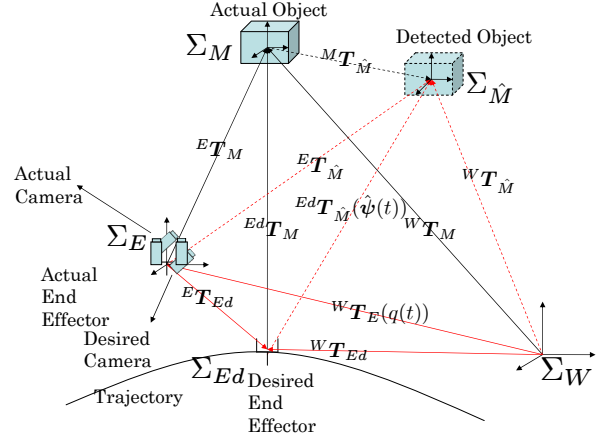


Fig. 9: Motion of the end-effector and object

Differentiating Eq. (21) with respect to time again

$${}^E \ddot{T}_{Ed}(t) = {}^E \ddot{T}_{\hat{M}}(t) {}^{\hat{M}} T_{Ed}(t) + 2 {}^E \dot{T}_{\hat{M}}(t) {}^{\hat{M}} \dot{T}_{Ed}(t) + {}^E T_{\hat{M}}(t) {}^{\hat{M}} \ddot{T}_{Ed}(t), \quad (22)$$

Where  ${}^{\hat{M}} T_{Ed}$ ,  ${}^{\hat{M}} \dot{T}_{Ed}$ ,  ${}^{\hat{M}} \ddot{T}_{Ed}$  are given as the desired visual servoing objective.  ${}^E T_{\hat{M}}$ ,  ${}^E \dot{T}_{\hat{M}}$ ,  ${}^E \ddot{T}_{\hat{M}}$  can be observed by cameras. As shown in Fig. 9, there are two errors that we have to decrease to  $\mathbf{0}$  in the visual servoing process. First one is the error between the actual object and the detected one  ${}^M T_{\hat{M}}$ , and the other is the error between the desired end-effector and the actual one  ${}^E T_{Ed}$ . In our research, the error of  ${}^M T_{\hat{M}}$  is decreased by on-line recognition method of 1-step GA, MFF compensation method and the eye-vergence camera system, and the error of  ${}^E T_{Ed}$  can be decreased by the hand visual servoing controller.

### 3.3 Hand & Eye Visual Servoing Controller

The block diagram of our proposed hand & eye-vergence visual servoing controller is shown in Fig. 4. The hand-visual servoing is the outer loop. A detailed block diagram of hand visual servoing control is depicted in Fig.10. Based on the above analysis of the desired-trajectory generation, the desired hand velocity  ${}^W \dot{r}_d$  is calculated as,

$${}^W \dot{r}_d = \mathbf{K}_{P_p} {}^W \mathbf{r}_{E,Ed} + \mathbf{K}_{V_p} {}^W \dot{r}_{E,Ed}, \quad (23)$$

where  ${}^W \mathbf{r}_{E,Ed}$ ,  ${}^W \dot{r}_{E,Ed}$  are given by transforming  ${}^E T_{Ed}$  and  ${}^E \dot{T}_{Ed}$  from  $\Sigma_E$  to  $\Sigma_W$ .  $\mathbf{K}_{P_p}$  and  $\mathbf{K}_{V_p}$  are positive definite matrix to determine PD gain.

The desired hand angular velocity  ${}^W \omega_d$  is calculated as,

$${}^W \omega_d = \mathbf{K}_{P_o} {}^W \mathbf{R}_E^E \Delta \epsilon + \mathbf{K}_{V_o} {}^W \omega_{E,Ed}, \quad (24)$$



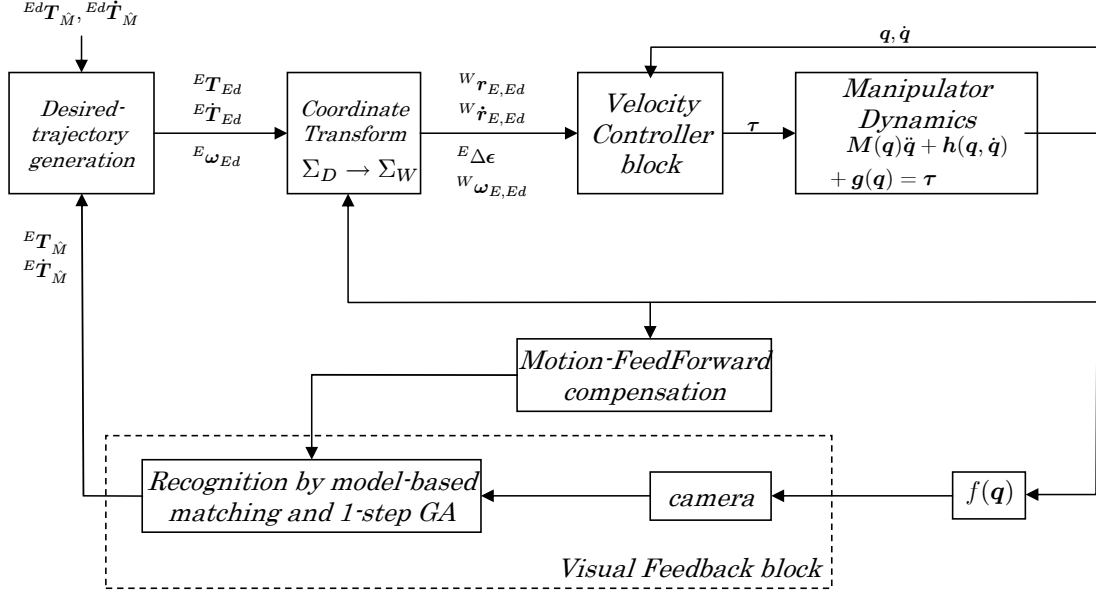


Fig. 10: Block diagram of the hand visual servoing system

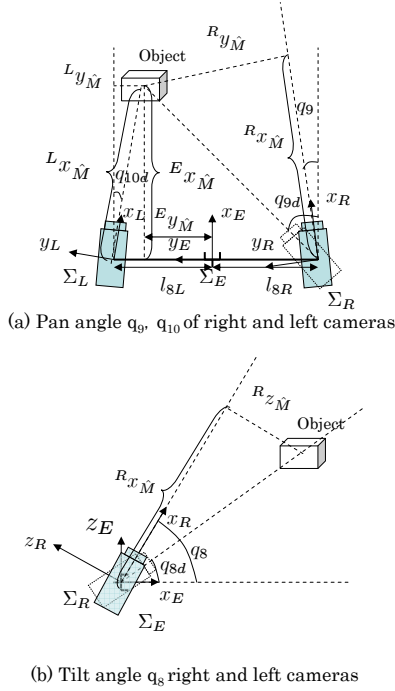


Fig. 11: Calculation of tilt and pan angles

where  $E \Delta \epsilon$  is the quaternion error that from the recognition result directly, and  $W \omega_{E,Ed}$  can be calculated by transforming  $E T_{Ed}$  and  $E \dot{T}_{Ed}$  from  $\Sigma_E$  to  $\Sigma_W$ . Also,  $K_{P_o}$  and  $K_{V_o}$  are suitable feedback matrix gains.

The desired joint variable  $\dot{q}_d$  is obtained by

$$\dot{q}_d = J^+(q) \begin{bmatrix} W \dot{\tau}_d \\ W \omega_d \end{bmatrix}. \quad (25)$$

where  $J^+(q)$  is the pseudo inverse matrix of  $J(q)$ , and  $J^+(q) = J^T(JJ^T)^{-1}$ . The hardware control system of

the velocity-based servo system of PA10 is expressed as

$$\tau = K_{SP}(\dot{q}_d - \dot{q}) + K_{SI} \int_0^t (\dot{q}_d - \dot{q}) dt \quad (26)$$

where  $K_{SP}$  and  $K_{SI}$  are symmetric positive definite matrix to determine PI gain.

The eye-vergence visual servoing is the inner loop of the visual servoing system shown in Fig. 4. In this paper, we use two pan-tilt cameras for eye-vergence visual servoing. Here, the positions of cameras are supposed to be fixed on the end-effector. For camera system,  $q_8$  is tilt angle,  $q_9$  and  $q_{10}$  are pan angles, and  $q_8$  is common for both cameras. As it is shown in Fig. 11,  $E x_M, E y_M, E z_M$  express position of the detected object in the end-effector coordinate. The desired angle of the camera joints are calculated by:

$$q_{8d} = \text{atan2}(E z_M, E x_M) \quad (27)$$

$$q_{9d} = \text{atan2}(l_{8R} + E y_M, E x_M) \quad (28)$$

$$q_{10d} = \text{atan2}(-l_{8L} + E y_M, E x_M) \quad (29)$$

where  $l_{8L} = l_{8R} = 120[\text{mm}]$  that is the camera location. We set the center line of the camera as the  $x$  axis of each camera coordinate. Then the controller of eye-visual servoing is given by

$$\dot{q}_8 = K_{P_T}(q_{8d} - q_8) + K_{D_T}(\dot{q}_{8d} - \dot{q}_8), \quad (30)$$

$$\dot{q}_9 = K_{P_C}(q_{9d} - q_9) + K_{D_C}(\dot{q}_{9d} - \dot{q}_9), \quad (31)$$

$$\dot{q}_{10} = K_{P_C}(q_{10d} - q_{10}) + K_{D_C}(\dot{q}_{10d} - \dot{q}_{10}). \quad (32)$$

where  $K_{P_T}, K_{D_T}, K_{P_C}, K_{D_C}$  are positive control gain.

#### 4 experiment of hand & eye-vergence visual servoing

To verify the effectiveness of the proposed hand & eye visual servoing system, we conduct the experiment of visual servoing to a 3D marker that is composed of a red

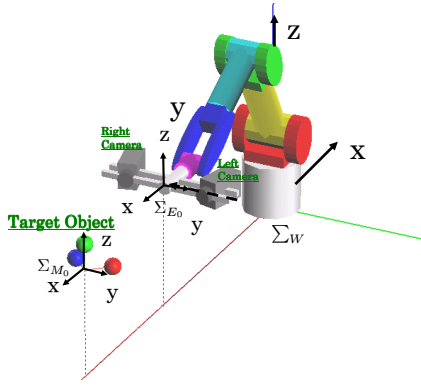


Fig. 12: Object and the visual-servoing system

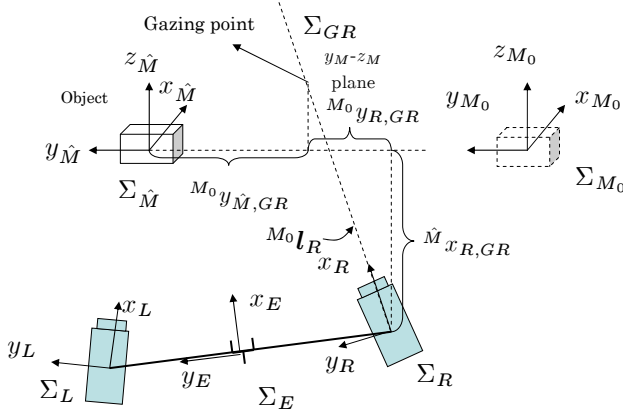


Fig. 13: Cameras' and End Effector's gazing point

ball, a green ball and a blue ball as Fig. 12. The radiuses of these three balls are set as 30[mm].

#### 4.1 experiment condition

The initial hand pose is defined as  $\Sigma_{E_0}$ , while the initial object pose is defined as  $\Sigma_{M_0}$ , and the homogeneous transformation matrix from  $\Sigma_W$  to  $\Sigma_{M_0}$  is:

$${}^W\mathbf{T}_{M_0} = \begin{bmatrix} 0 & 0 & -1 & -1410[mm] \\ 1 & 0 & 0 & 0[mm] \\ 0 & -1 & 0 & 355[mm] \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (33)$$

The target object move according to the following time function

$${}^{M_0}\psi_M = [0, {}^{M_0}y_M(t), 0, 0, 0, 0]^T \quad (34)$$

$${}^{M_0}y_M(t) = -200 \sin(\omega t)[mm] \quad (35)$$

here,  $\omega$  is the angular velocity of the motion of the object.

The relation between the object and the desired end-effector is set as:

$${}^{E_d}\psi_M = [800[mm], 0, 0, 0, 0, 0] \quad (36)$$

Here, to compare the trackability of the eye-vergence system and fixed camera system, we define a concept of gazing point. As it is shown in Fig. 13 the intersection of the gazing line of right camera and the  $y_{M_0}$ - $z_{M_0}$  plane is defined as the gazing point. The relative relation between

$\Sigma_{M_0}$  and  $\Sigma_R$  is given by Homogeneous Transformation as  ${}^{M_0}\mathbf{T}_R$ ,  ${}^{M_0}\mathbf{T}_R$  conclude the rotation matrix  ${}^{M_0}\mathbf{R}_R$  and the position vector  ${}^{M_0}\mathbf{p}_R$ , and the rotation matrix  ${}^{M_0}\mathbf{R}_R$  can be written as  $[{}^{M_0}\mathbf{x}_R, {}^{M_0}\mathbf{y}_R, {}^{M_0}\mathbf{z}_R]$ . The direction of  ${}^{M_0}\mathbf{l}_R$  in Fig. 13 is same to the direction of  $x_R$ , and  ${}^{M_0}\mathbf{l}_R$  can be expressed as:

$${}^{M_0}\mathbf{l}_R = {}^{M_0}\mathbf{p}_R + k_R {}^{M_0}\mathbf{x}_R \quad (37)$$

here  $k_R$  is a scalar variable. The gazing point of the right camera expressed in  $\Sigma_{M_0}$  is  ${}^{M_0}\mathbf{p}_{GR} = [0, {}^{M_0}y_{GR}, {}^{M_0}z_{GR}]^T$ . For  ${}^{M_0}\mathbf{l}_R = {}^{M_0}\mathbf{p}_{GR}$  in  $x$  direction,  $({}^{M_0}\mathbf{p}_R)_x + k_R({}^{M_0}\mathbf{x}_R)_x = 0$ . And usually  $({}^{M_0}\mathbf{x}_R)_x \neq 0$ ,  $k_R$  can be calculated by  $k_R = -({}^{M_0}\mathbf{p}_R)_x / ({}^{M_0}\mathbf{x}_R)_x$ , and the  $y, z$  coordinate of the gazing point in  $\Sigma_{M_0}$  can be calculated by:

$${}^{M_0}y_{GR} = ({}^{M_0}\mathbf{p}_R)_y + k_R({}^{M_0}\mathbf{x}_R)_y \quad (38)$$

$${}^{M_0}z_{GR} = ({}^{M_0}\mathbf{p}_R)_z + k_R({}^{M_0}\mathbf{x}_R)_z \quad (39)$$

The target object's motion is given by (34), (35), because the motion of the target object  $M$  is parallel to the  $y_{M_0}$ , we take  ${}^{M_0}y_M(t)$  as the input, and the gazing point of the right camera  ${}^{M_0}y_{GR}(t)$  as the response. And define the concept of trackability by the frequency response of  ${}^{M_0}y_{GR}(t)$ , the trackability of the left camera can be defined in the same way.

#### 4.2 Experiment Results

In Fig. 14, we show the result of our experiment, we change the  $\omega$  in (34) from 0.01 to 1.256 and get the data of the gazing point of the cameras of eye-vergence system and the gazing point of the end-effector of the fixed camera system separately, we do the experiment 10 times at every  $\omega$  we selected, and use the average delay time and the amplitude to draw the frequency response curve. The amplitude-frequency curve and the delay frequency curve are shown in Fig.14 (a) and Fig.14 (b). Here, for the fixed camera  $A = {}^{M_0}y_M(t)$ ,  $B = {}^{M_0}y_{GE}(t)$ . For the right camera of Eye-Vergence system  $A = {}^{M_0}y_M(t)$ ,  $B = {}^{M_0}y_{GR}(t)$ , for the left camera  $A = {}^{M_0}y_M(t)$ ,  $B = {}^{M_0}y_{GL}(t)$ . In this two figures the abscissa axes are  $\omega$ . In (a), (b), we sign the angular velocity when  $\omega = 0.314, 0.628, 1.256$ , and show the position of the gazing point of the cameras in eye-vergence experiment and the position of the gazing point of the end-effector in fixed camera experiment in (c), (d), (e). From (a), (b) we can see that the fixed-camera system cannot track the target object when  $\omega$  is faster than 0.628 so in (e), there is only the data of the cameras and the target object. From Fig. 14 (a) we can see the data of the cameras and the end-effector all become smaller as  $\omega$  increases but the curve of the fixed camera system is always below the curves of the cameras, which means that delay of the fixed camera system is bigger than the eye-vergence system, from (b) the the curve of the fixed camera system is also below the curves of the cameras, we can see that the amplitude of the eye-vergence system is more closed to the target object than the fixed camera system, so from (a) and (b) we can get the conclu-

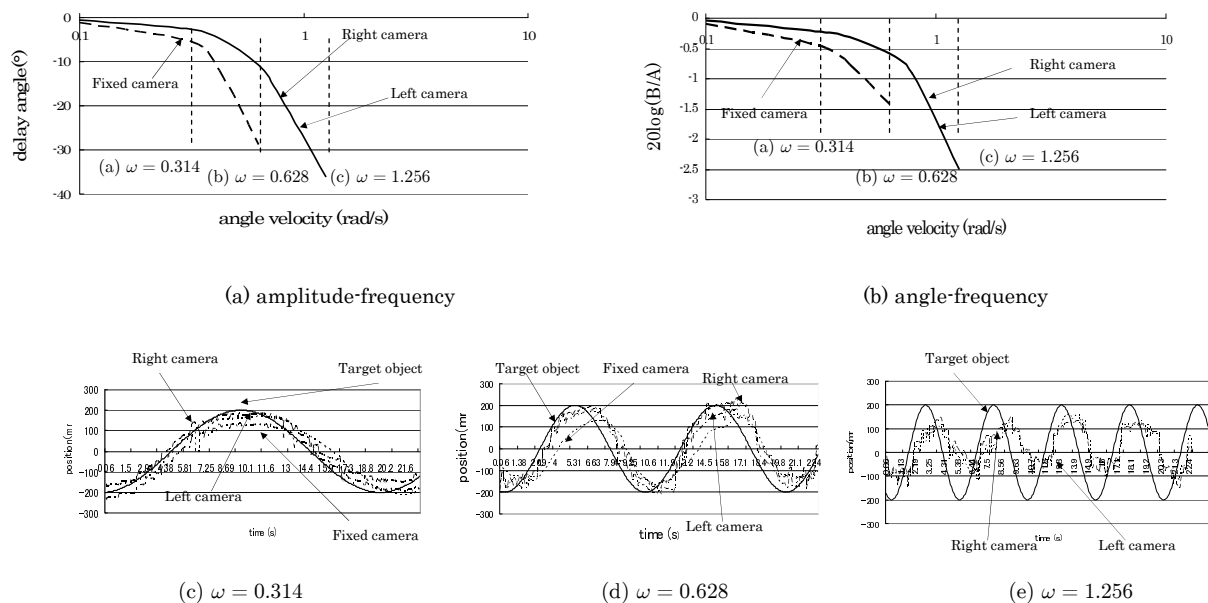


Fig. 14: Comparison of Cameras' and End-effector's Trackabilities by Frequency Response

sion that the eye-vergence system has the better trackability than the fixed-camera system. To be understood easily, we show the position of the gazing point of the cameras in eye-vergence experiment and the position of the gazing point of the end-effector in fixed camera experiment in (c), (d), (e). and  $M_0 \dot{y}_M(0) = -200[mm/s]$ , while the target object moved from static, so it cannot move stably at first, we use the data when the target object's motion became stable. From the figures it is also easily to see that comparing with the fixed camera system, the eye-vergence system can track the target object better.

## 5 Conclusion

In this paper, we put forward a new concept to evaluate the observation ability on a moving object of visual servoing system, and introduce the importance of it. Then we introduce the recognition method using "1-step GA" and our eye-vergence system. To check the trackability of eye-vergence visual servoing system, we did some experiments of eye-vergence system and fixed camera system separately. In the experiments we compared the amplitude-frequency and phase-frequency curves of the gazing point of the cameras of the eye-vergence system and the fixed camera system by moving object in different angular velocities, and get the conclusion that the trackability and stability of the eye-vergence system is better than that of the fixed-camera system by analyze the experiment data.

- 1) B.Siciliano and L.Villani: *Robot Force Control*, ISBN 0-7923-7733-8.
- 2) L. Reznik and V. Kreinovich Eds, *Soft Computing in Measurement and Information Acquisition*, ISBN 3-540-00246-4.
- 3) W. Song, M. Minami, Y. Mae and S. Aoyagi, "On-line Evolutionary Head Pose Measurement by Feed-

forward Stereo Model Matching", IEEE Int. Conf. on Robotics and Automation (ICRA) 4394/4400 (2007).

- 4) Omar Tahri and Francois Chaumette, "Point-Based and Region-Based Image Moments for Visual Servoing of Planar Objects", IEEE Tran. on Robotics, vol. 21, no. 6, Dec (2005).
- 5) Tarek Hamel and Robert Mahony, "Visual Servoing of an Under-Actuated Dynamic Rigid-Body System: An Image-Based Approach", IEEE Trans. on Robotics and Automation, VOL. 18, NO. 2, APRIL (2002).
- 6) Tomoyuki Shiozaki, Toshiyuki Murakami. "Trackability Based Motion Control in Mobile Hand-Eye System". SICE-ICASE International Joint Conference 5304/5309 (2006).
- 7) H. Suzuki, M. Minami, "Visual Servoing to catch fish Using Global/local GA Search", IEEE/ASME Transactions on Mechatronics, Vol.10, Issue 3, 352/357 (2005).
- 8) Wei Song, Mamoru Minami, and Seiji Aoyagi, "Evolutionary Pose Measurement by Stereo Model Matching" Journal of Advanced Computational Intelligence and Intelligent Informatics(JACIII), vol.9, No.2, 150/158 (2005)
- 9) W. Song, M. Minami, S. Aoyagi, "On-line Stable Evolutionary Recognition Based on Unit Quaternion Representation by Motion-Feedforward Compensation", International Journal of Intelligent Computing in Medical Sciences and Image Processing (IC-MED) Vol. 2, No. 2, 127/139 (2007).



# 魚捕獲ロボットのためのニューラルネットワーク組み込み型微分方程式によるカオスの生成とその検討

○伊藤雄矢 友野高志 見浪護 矢納陽 (岡山大学)

## Multiple Chaos Generation and Its Examination by Neural-Network-Differential-Equation for Intelligence Fish-Catching

\*Y. Ito, T. Tomono, M. Minami and A. Yanou (Okayama University)

**Abstract**— To make fish-catching robot intelligent more than fishes, we thought something that goes beyond the fishes' adapting intelligence would be required. Here we propose a chaos-generator comprising Neural-Network-Differential-Equation (NNDE) and an evolving mechanism to generate plural differential equations as many as possible that can yield different kind of chaos. We believe that the fish could not be adaptive to escape from chasing net with chaotic motion. In this paper we introduce chaos-generating system by NNDE and explore a relation between Lyapunov number and chaos generated by a variety of NNDE's coefficient.

**Key Words:** Neural network, Chaos, Genetic algorithm

### 1 緒言

近年では、画像情報を用いてマニピュレータによりターゲットを補足するビジュアルサーボイングの研究が行われている<sup>1)2)</sup>。機械知能の新しい傾向<sup>3)</sup>として、ロボット工学分野に応用されたようなAIと異なり、知能制御システムのような他の研究分野に広がっている。また、動物界はロボット工学者たちによって機械知能のための概念として使われてきた。動物の振る舞いや知能を研究する目的として、動物と機械の相互作用のモデルが提案された<sup>4)</sup>。機械知能の特徴は、変わりゆく環境の中で、どのように振舞うべきかをセンサーからの入力を使用し、障害物回避などといった行動を学習することである。よく知られているように、知能ロボットの知的レベルは比較的高度であり、この知能を測ることは重要である。実用的で体系的に測る方法としてMIQ(Machine Intelligence Quotient)<sup>5)</sup>がある。これは人と機械の協同作業をシステム化したものである。我々のシステムでは、知能の程度を魚とロボットを競争させることにより評価している。我々の過去の研究で、魚の感情的な振る舞いを調べ、魚の状態に反応する適応能力を持つロボットが考えられた。魚が逃げる戦略を見つけたとしても、ロボットが魚を捕獲することができれば、提案した手法によってそのシステムが魚より賢くなったといえる。この知能競争はロボットに知能を持たせる方法の一つとして有効であると考えられる。これより、機械知能に動物が持つひらめきを用いるだけではなく、動物の知能を越えるロボットを開発する。

本研究室では、進化的アルゴリズム<sup>6)</sup>、ビジュアルサーボイング、マニピュレータに取り付けられたCCDカメラからの画像により物体を認識する方法などを研究している<sup>7)</sup>。魚捕獲システムをFig.1に示す。また、実時間による認識能力を高めた注視GA<sup>8)</sup>により、マニピュレータに取り付けられた捕獲用網による魚の捕獲に成功している。泳いでいる魚を追従しているとき、単純に速度制御でビジュアルサーボイングするのでは、魚の捕獲には有効ではないことがわかった。魚は突然泳ぐ方向を変えたり、追ってくる網に対して逃げる戦略をとったりするため、魚の捕獲は困難になる。それらの

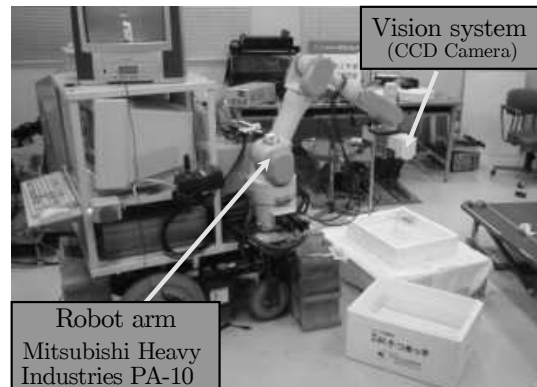


Fig. 1: Fish Catching system (PA10)

行動は、感情的要因もしくは魚が本来持っている知能であると見なす事ができる。魚の捕獲実験により魚の行動を観察すると、魚は捕獲を続けるうちに網から逃れるためプールの隅を泳ぐことがわかった。これは魚の捕獲にとって深刻な問題である。なぜなら、もし魚がプールの隅のみを泳げば、網は隅には侵入できず、魚を捕獲できないからである。効果的な知能とは、この魚の戦略の対処法を思いつくものでなくてはならない。魚とロボットの知能競争により魚が学習していくことを観察すると、魚の捕獲/リリース実験を行っている中での捕獲数の減少速度を“Fish Intelligent Quotient”(FIQ)<sup>9)</sup>と定義することができる。この評価手法により、魚が生まれつきもっている知能とロボットの人工知能とを比較することができると思われる。

ところで、カオス信号は、神経系、眼球の動き、心臓の拍動周期など動物の体内に存在していることが知られている。本報告では、細胞組織の中の信号トランスファーから得られたカオスモデルを採用する<sup>10)11)</sup>。具体的には、魚の捕獲システムの欠点を補足するため、ロボットのダイナミクスにカオスを組み込んだ。この試みに対して、魚がカオス軌道にも適応し、カオス軌道を見逃すようになることが分かった。そこで、魚が常に新しい障害に対する回避戦略を考えていることに着目した。これは、ロボットが魚の知能を上回るには、

無数の捕獲動作が必要であることを示している。このような捕獲動作を生成するため本研究では、任意の非線形関数を表現できるニューラルネットワーク<sup>12)13)</sup>を用いた関数表現を微分方程式に組み込むことで、無数のカオス軌道を生成することの出来るニューラルネットワーク組み込み型微分方程式を提案する。また、生成した軌道のカオス性をリアプノフ指数、初期値敏感性、ポアンカレリターンマップにおいて確認し、さらなるカオス軌道生成のため、ニューラルネットワークの重み係数の値に関して検討を行う。

## 2 魚の追跡と捕獲

魚の認識の問題と位置／姿勢の検出は、魚の入力画像とマッチングモデルの相関関数である  $F_{ss}(\phi)$  を最大にするような  $\phi$  の探索問題に変換される。ある対象物を動画像 33[fps] の中で認識するためには、その認識システムは実時間性を持たなければならない。すなわち、その探索モデルは連続した入力 RAW 画像内の魚に収束しなければならない。動画像による進化的認識のプロセスは、次の方法で実現される。Model-based Matching 法はビデオレートで連続して入力される RAW 画像一つに対して一回だけ適合される。これを 1-Step GA とよぶ。動画像における目標物へのモデルの収束速度は魚の泳ぐ速度よりも早くなくてはならない。実時間で入力された画像から魚の位置を最も高いゲインとして表示し、これを魚の位置として表す。魚の位置認識問題は、1-Step GA を用いて、 $F_{ss}(\phi)$  を最大にする  $\phi$  を探索することで解決した。また、Fig.2 に示したブロック図を用いたシステムを使用し、泳ぐ魚に対して追跡と捕獲の実験を行っている。カメラと魚の距離は 450[mm] である。プールの大きさは、300 (width)×400 (length)×100 (depth) [mm]、ネットの大きさは 80×100 [mm] とした。ロボットは、魚が網の中心から 86×66 [mm] 内にいると判断したとき網を持ち上げ、魚を捕まえる。前述の実時間認識システムは魚の形を知っているものとしている。 $i$  番目の制御期間のときのハンドに与える速度は次のように計算される。

$$\dot{r}_d^i = K_P \Delta r^i + K_V (\Delta r^i - \Delta r^{i-1}) \quad (1)$$

$\Delta r^i$  は、1-Step GA により実時間で得られた魚の位置とカメラ中心との位置をベクトルで表したものである。 $K_P$  と  $K_V$  は、PD ゲインで一定の正の値で示される行列である。ここで、カオスを式 (1) に加えるため、 $\Delta r^i$  を次式のように再定義する。

$$\Delta r^i = k_1 \cdot \Delta r_{fish}^i + k_2 \cdot \Delta r_{chaos}^i \quad (2)$$

$\Delta r_{fish}^i$  は、魚の位置の追跡誤差、 $\Delta r_{chaos}^i$  は、カオスの振動を表す。そして、ハンドの運動パターンは、 $k = [k_1 \ k_2]$  の値で決められる。 $k = [1 \ 0]$  の場合は、網は魚を追いかける。 $k = [0 \ 1]$  の場合はカオス運動を行う条件を満足し、網はカオス運動を行う。

## 3 魚捕獲の問題点

追従と捕獲システムの信頼性を調べるために、複数の魚を捕獲し続ける実験を 30 分間行った。5 匹の魚 (体長およそ 40[mm]) をプールの中に放ち、一匹捕まえる度にまた同じプールに逃がすことにした。この実験の結果を Fig.3 に示す。縦軸に 5 分間に捕まえた魚の数を、

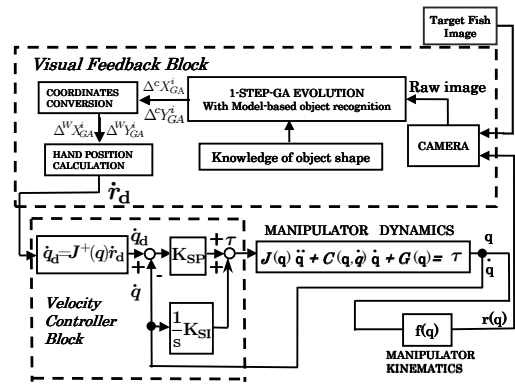


Fig. 2: Block diagram of the controller

横軸は捕獲していた時間を示している。捕獲を続けていけば魚が疲れることにより、魚の捕獲率は向上していくことが期待された。しかし、驚くことに多くの魚は徐々に捕獲されなくなっていくたのである。この理由として、魚が本来持っている学習能力、もしくは、感情的状態変化が考えられる。たとえば、魚は、Fig.4(a)のように円状にある適度な距離を保って逃げるといった方法とる。また、Fig.4(b)のように網と反対側の隅に位置しとどまるといった行動をとる。さらに、Fig.4(c)のように網が進入できないプールの隅から出てこないといった行動も行う。このような問題を解決するため、さらに知的な追従システムと、効果的に魚を捕まえるシステムを開発する必要がある。そのために、カオスを使用する。

## 4 Fish intelligence quotient

魚が網から逃げる戦略を学習する速度を数値により評価する。そのために、魚の捕獲数が減少・増加していく傾向を最小二乗法に適用した。Fig.3 に示されている  $y = -0.486t + 20.7$  は、五分毎の魚の捕獲数から最小二乗法により算出した。ロボットの性能に基づいて魚の知能を評価するとき、減少する速度を表す係数  $-0.486$  は、その魚群の適応速度もしくは学習速度を表しているといえる。この係数を “Fish Intelligence Quotient” (FIQ) と呼ぶこととする。この値が負の値ならば、魚群と比べてロボットのほうが知能指数が低いことを示し、0 ならば同等、正の値ならばロボットのほうがより知能指数が高いことを意味する。この FIQ は、今後の本研究における魚捕獲ロボットの知能向上を定量的に評価できるという点だけではなく、動物の知能行動を基準としてロボットの知能を議論するという視点において重要な意味を持つと考える。

## 5 カオスの有用性

1982 年に、植物細胞と軟体動物のニューロンを周期的な電流で刺激すると、細胞が不規則な興奮をし、カオスの性質を示すことが実験的に明らかにされた。また、1984 年にはイカの細胞体から延びている突起状の構造で、神経細胞において信号の出力を担う軸索でも周期的な電流刺激に対するカオス的応答が明らかにされている。これらの研究から、カオスはバイオロジーに関連することが明らかになった。1980 年代後半になると、神経系の機能とカオスとの関係性が議論されるようになった。Mpitosos らは、ナマコの運動ニューロンのリ

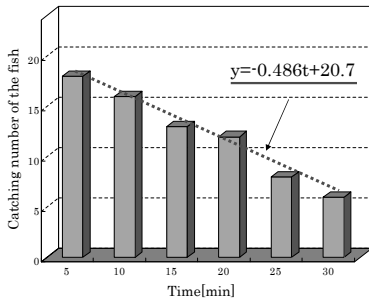


Fig. 3: Result of catching number

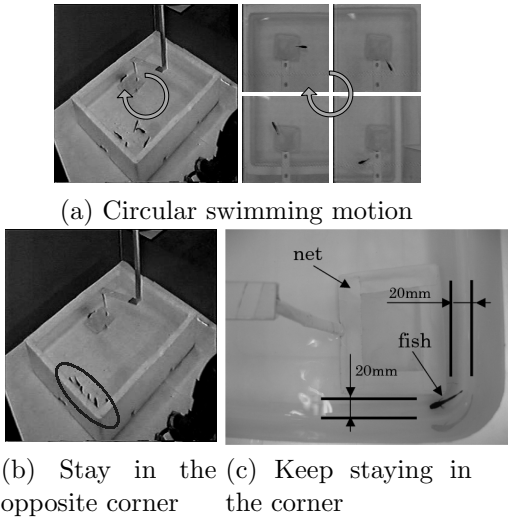


Fig. 4: Fish motion

ズミックな発火パターンを調べ、運動のリズムと関係している連続放電の周波数変動がカオスであることを示した。以上より、カオスは生物の行動の中に存在していると考えられる。これは、生体中の神経細胞が興奮するかどうかは刺激信号によって決められ、それはカオスの理論に従っているからである。このことから、動物の行動や戦略はカオスから得られると推測でき、魚の捕獲に応用できると考える。それらに対し、カオスの一つである神経系の電気的特性を表す Hodgkin-Huxley モデル (H-H モデルと略す) を簡略化した BVP (Bonhoeffer-van der Pol) モデルを用いることにより魚を隅から誘い出し捕獲する方法が提案された。この提案は実験により、カオスには魚を隅から誘い出す又は脅かすという効果が認められ、魚の知能を超えるためにカオスを用いることは有効であると確認されている。

## 6 ニューラルネットワーク組み込み型微分方程式

レスラーモデルやローレンツモデルなどは3つの微分方程式で表され、その解が三次元の相空間内の軌跡として表されるカオスアトラクターとなることが知られている。この微分方程式をニューラルネットワークで表現する。ニューラルネットワークは任意な非線形関数を表現できると言われており、その関数表現を微分方程式に組み込むことで、非線形微分方程式を多様な形に変化させることができる。この方法は、さまざまな非線形関数を表現できるため、多くのカオス軌道

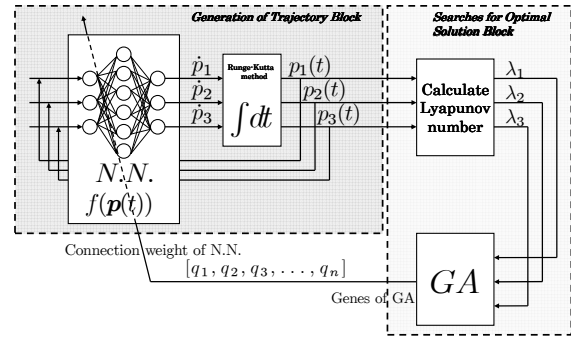


Fig. 5: Block diagram of Chaos Generation

を生成できる可能性を持つ。そこで、入力層を3つ、中間層を6つ、出力層を3つとしたニューラルネットワークを用いる。ニューラルネットワークの出力を式(3)のように定義する。

$$\dot{\mathbf{p}}(t) = \mathbf{f}(\mathbf{p}(t)). \quad (3)$$

ただし、 $\mathbf{p}(t) = [p_1(t), p_2(t), p_3(t)]^T$  とする。次に N.N. から出力された値をルンゲクッタ法を用いて解く。このとき得られた値を解とすることで、微分方程式を表現する。これをニューラルネットワーク組み込み型微分方程式と呼ぶこととする。そのブロックダイアグラムを Fig.5 に示す。

## 7 カオス性の判定基準

不規則な振動がカオスであるか判断する基準がないため、ニューラルネットワーク組み込み型微分方程式によって生成される軌道がカオスであるかの判定には複数の指標を用いる。以下は、カオスの特徴を判定する基準である。

### 7.1 リアプノフ指数

カオス軌道の性質としてリアプノフ指数というものがある。リアプノフ指数とは、力学系においてごく接近した軌道の広がり具合を表す指標であり、以下の式で表されることが知られている。

$$\lambda = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \log |f'(x_i)| \quad (4)$$

$\lambda > 0$  のとき、近接した軌道間の距離は、指数関数的に増大することになる。このような軌道の系を不安定な系という。

### 7.2 ポアンカレリターンマップ

ニューラルネットワークに基づく非線形関数(3)による軌道が、カオス軌道であるかをポアンカレリターンマップを用いて評価する。まず、Fig.6 に示す3次元の閉曲線を考える。P が示すこの軌道と交差する平面“A”がポアンカレ断面として定義されている。

交わる点をそれぞれ  $P_n, P_{n+1}, P_{n+2}, \dots$  とし、A における交差する点の  $x$  座標の値を  $x_n, x_{n+1}, x_{n+2}, \dots$  としたポアンカレリターンマップを Fig.7 に示す。Fig.7 に示される軌道は閉曲線が引き伸ばしと折りたたみの構造を持つことを表している。この構造はカオスの基本的な性質である。Fig.7 の左側では  $dx_{n+1}/dx_n > 1$

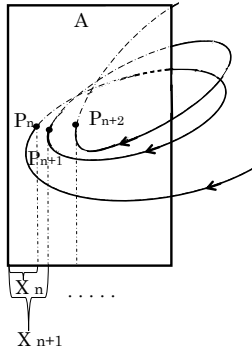


Fig. 6: Poincare section

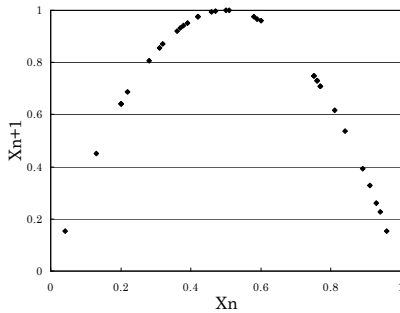


Fig. 7: Poincare return map

の傾向が見られ、右側では  $dx_{n+1}/dx_n < -1$  の傾向が見られる。これは左側が引き伸ばし、右側が折たたみの現象であることを表す。

### 7.3 初期値敏感性

カオスには、初期値がほんの少し異なると、結果が大きく異なるという特徴がある。一般的には、バタフライ効果という名前が知られている。

## 8 カオス生成システム

Fig.5にGAとリアプノフ数を用いたカオス発見手法のブロックダイアグラムを示す。N.N.組込型微分方程式により得られた軌道より、リアプノフ数  $L$  を求めることができる。ここで、 $L = [\lambda_1, \lambda_2, \lambda_3]^T$  とする。このリアプノフ数を用いてGAの進化を行う。GAの進化の手法はエリート保存戦略を用いている。評価関数は次のように設定する。

$$g = k_1 \cdot \lambda_1 - k_2 \cdot |\lambda_2| - k_3 \cdot \lambda_3. \quad (5)$$

この評価関数はカオスのリアプノフスペクトラムの性質を考慮している。ここでは、三次元相空間内のカオスアトラクターを想定しているため、リアプノフ数は3個存在する。リアプノフスペクトラムは  $(+, 0, -)$  となる。括弧内はリアプノフスペクトラムの符号を示している。つまり、 $\lambda_1$  が正の場合に  $g$  は大きくなり、 $\lambda_2$  は正負のどちらの値となっても  $g$  が小さくなり、 $\lambda_3$  が負の値となった場合、 $g$  が大きくなるように設定している。また、 $k_1, k_2, k_3$  は重み係数である。GAの遺伝子はFig. 8のように与え、探索するN.N.結合重み  $q = [q_1, q_2, \dots, q_n]^T$  はGAの遺伝子一つ一つに対応している。遺伝子は二進法で表されているので、十進法

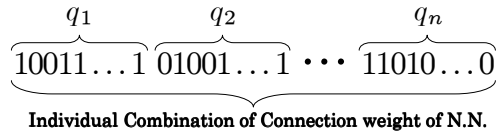


Fig. 8: Gene of GA

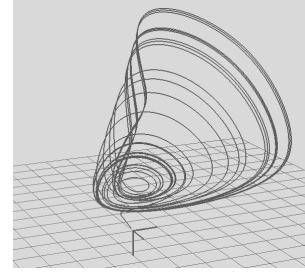


Fig. 9: Generated chaos trajectory 01

に変換し、その値を0~1の値に縮小し結合重みとして用いる。そしてリアプノフ数を求め、評価関数からGAを進化させ、これを繰り返す。これにより、カオスのリアプノフスペクトラムを満足する軌道をGAによって探索し、カオス軌道の生成を行う。

## 9 生成した軌道のカオス性の検証

現在までに4種類のカオス軌道を生成した。それらにカオス01、カオス02、カオス03、カオス04と名づけた。以下で、見つけた軌道のカオス性を検証する。

### 9.1 カオス01

#### 9.1.1 リアプノフ指数

リアプノフ指数はそれぞれ  $\lambda_1 = 0.014585$ ,  $\lambda_2 = -0.003314$ ,  $\lambda_3 = -0.165381$  となった。これらはカオスのリアプノフスペクトラム  $(+, 0, -)$  と一致している。

#### 9.1.2 初期値敏感性

わずかに異なる初期値を与えた2つの軌道の推移をFig.10, Fig.11, Fig.12に示す。 $(x_1(t), y_1(t), z_1(t))$ の軌道は初期値を  $x_1(0) = 1.00, y_1(0) = 1.00, z_1(0) = 1.00$  とし、 $(x_2(t), y_2(t), z_2(t))$ は初期値を  $x_2(0) = 1.01, y_2(0) = 1.01, z_2(0) = 1.01$  としている。 $x_1, x_2$ の推移はFig.10,  $y_1, y_2$ の推移はFig.11,  $z_1, z_2$ の推移はFig.12に示す。Fig.10~12から2つの軌道は800秒付近からずれが生じ、最終的には異なる値を取るようになる。これはカオスの特徴である初期値敏感性を表している。

#### 9.1.3 ポアンカレリターンマップ

カオス01のポアンカレリターンマップをFig.13に示す。Fig.13において1次元写像が確認でき、左側では引き伸ばし、右側で折たたみの現象が確認できる。これらはカオスの現象を引き起こす重要な特徴である。したがって、カオス01のカオス性が、リアプノフ指数・初期値敏感性・ポアンカレリターンマップの観点から確認できた。

### 9.2 カオス02~カオス04

カオス02, カオス03, カオス04の軌道をFig.14, Fig.15, Fig.16に示す。これらの軌道もカオス01と同様に、リアプノフ指数, 初期値敏感性, ポアンカレリ

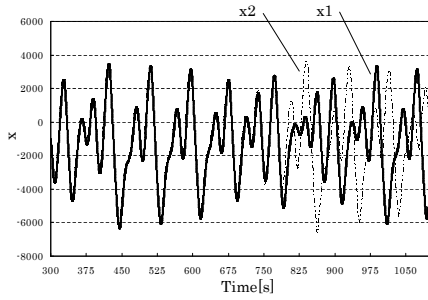


Fig. 10: Generated trajectory 01 of x (300[s] to 1100[s])

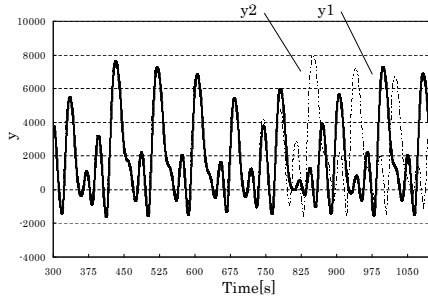


Fig. 11: Generated trajectory 01 of y (300[s] to 1100[s])

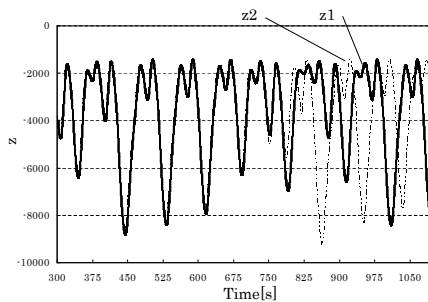


Fig. 12: Generated trajectory 01 of z (300[s] to 1100[s])

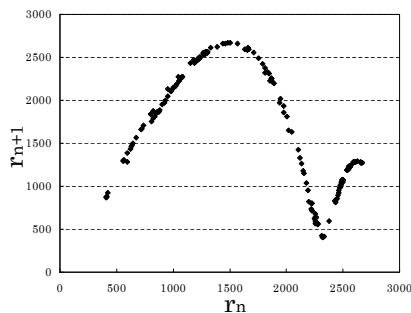


Fig. 13: Poincare return map of Chaos01

ターンマップにおいてカオス性を確認することができた。

## 10 ニューラルネットワークの重み係数と生成される軌道

9.2で示したカオス 03 とカオス 04 のニューラルネットワークの重み係数は非常に似ており、Fig.17に示すように  $q_1$  のみが異なることが分かった。そこで、重み係数  $q_1$  を  $-1 \sim 1$  の範囲で 0.01 刻みで変化させ、

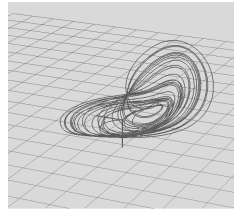


Fig. 14: Generated chaos trajectory 02

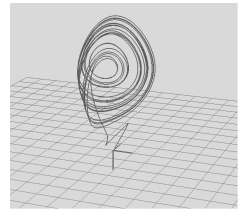


Fig. 15: Generated chaos trajectory 03

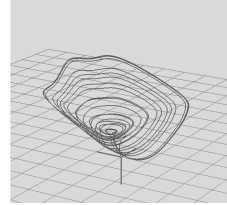


Fig. 16: Generated chaos trajectory 04

生成される軌道について検証を行なった。その結果を Table 1 に示す。5つの区間それぞれの代表的な例として、 $q_1 = -0.50, 0.00, 0.30, 0.50, 0.90$  のときの軌道を Fig.18~22 に示す。また、カオスらしい軌道が確認されたときのリアプノフ指数と適合度の値を Fig.23, 24, 25, 26 に示す。カオスらしい軌道のリアプノフ指数は、カオスのリアプノフスペクトラム (+, 0, -) と一致していることが分かる。また、9.2で示したカオス 03 とカオス 04 の  $q_1$  の値はそれぞれ、0.829098955,  $-0.108415351$  となっており、カオスらしい軌道が生成された区間であることから、この区間ではカオス軌道が生成されているのではないかと考えられる。このことから、生成される軌道がカオスとなる区間とカオスとならない区間が存在することが分かった。Fig.23より、 $\lambda_1$  は重み係数が大きくなるにつれてわずかだが大きくなる傾向にあることが分かる。Fig.24より、 $\lambda_2$  は重み係数が大きくなるにつれて、ばらつきが生じている。Fig.25より、 $\lambda_3$  は重み係数が大きくなるにつれて、より小さい値を取るようにになっている。Fig.26より、適合度は重み係数が大きくなるにつれて大きな値をとるようにになっているが、式(5)において  $k_1 = 8$ ,  $k_2 = 3$ ,  $k_3 = 1/3$  としているため、適合度の値が大きくなっていること的主要要因は  $\lambda_1$  の値が大きくなることであると考えられる。これは、軌道がカオスとなるためには最大リアプノフ指数が正という条件が重要であることに基づいている。また、カオス軌道となる区間では  $\lambda_1$  と  $\lambda_2$  の値は、ランダムのように値が変化していることが Fig.23,24 から確認できるが、Fig.27,28 に示すように、カオス軌道とならない区間である  $-1 \leq q_1 \leq -0.28$  と  $0.85 \leq q_1 \leq 1$  における  $\lambda_1$  と  $\lambda_2$  は、連続して変化しているように思われる。なお、 $0.17 \leq q_1 \leq 0.4$  の区間でランダムのように変化しているのは、Fig.20に示しているように、ある程度カオスのような軌道を描いた後発散する軌道であるからだと考える。一方、Fig.18, Fig.22 のようなカオス軌道とはまったく異なる軌道を生成する区間では、上で述べたように  $\lambda_1$  と  $\lambda_2$  が連続的に変化する傾向が見られる。しかし、この傾向は  $\lambda_3$  には見られなかった。このことから、重み係数を変化させた際に、 $\lambda_1$  と  $\lambda_2$  の値がランダムのように変化している区間を

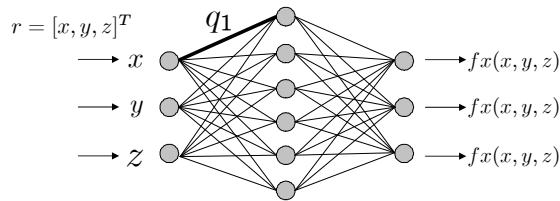


Fig. 17: Neural network for nonlinear function generation

Table 1:  $q_1$  and generated trajectories

$q_1$	Generated trajectories
$-1 \leq q_1 \leq -0.28$	Not chaos
$-0.27 \leq q_1 \leq 0.16$	Looks like chaos
$0.17 \leq q_1 \leq 0.4$	Not chaos
$0.41 \leq q_1 \leq 0.84$	Looks like chaos
$0.85 \leq q_1 \leq 1$	Not chaos

探索すればカオス軌道を見つけることが出来るのではないかと考えられる。また、これらのことは  $q_1$  以外のニューラルネットワークの重み係数にもあてはまると考えらる。ニューラルネットワークの重み係数の組み合わせとカオス生成との関連が判明すれば、さらなるカオス軌道の生成につながると思う。

## 11 結言

本論文では、魚の知能を上回るロボット知能の生成のため、ニューラルネットワークと GA を利用したカオス生成システムであるニューラルネットワーク組み込み型微分方程式を提案した。今回は、ニューラルネットワークの重み係数  $q_1$  の値を変化させることでカオス軌道を生成することが出来るか確認したが、今後は、 $q_1$  以外の重み係数を変更した場合でもカオス軌道を生成することができるか確認する。また、複数のカオス軌道を使用した魚の捕獲実験はまだ行っていないため、実際に捕獲実験を行うことで複数のカオスを使用することの有用性を検証する必要がある。

## 参考文献

- 1) R. Kelly, "Robust Asymptotically Stable Visual Servoing of Planar Robots", *IEEE Trans. Robot. Automat.*, vol.12, no.5, 759/766, (1996)
- 2) P.Y. Oh, and P.K. Allen, "Visual servoing by partitioning degrees of freedom", *IEEE Trans. Robot. Automat.*, vol.17, 1/17, (2001)
- 3) T. Fukuda, and K. Shimojima, "Intelligent Control for Robotics", *Computational Intelligence*, 202/215, (1995)
- 4) M. Bohlen, "A robot in a cage-exploring interactions between animals and robots", *CIRA.*, 214/219, (1999)
- 5) Hee-Jun Park, Byung Kook Kim, Kye Young Lim, "measuring the machine intelligence quotient (MIQ) of human-machine cooperative systems", *IEEE Trans.*, vol.31, 89/96, (2001)
- 6) M. Minami, H. Suzuki, J. Agbanhan, T. Asakura, "Visual Servoing to Fish and Catching Using Global/Local GA Search" *Int. Conf. on Advanced Intelligent Mechatronics*, Proc, 183/188, (2001)
- 7) M. Minami, J. Agubanhhan, and T. Asakura, "Manipulator Visual Servoing and Tracking of Fish using Genetic Algorithm", *Int. J. of Industrial Robot*, Vol.29, No.4, 278/289, (1999)

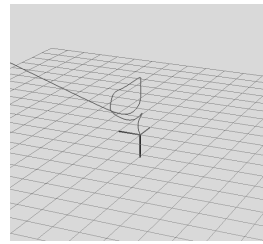


Fig. 18:  $q_1 = -0.50$

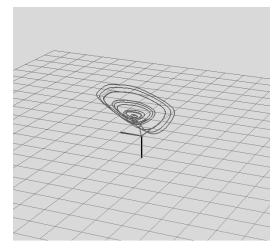


Fig. 19:  $q_1 = 0.00$

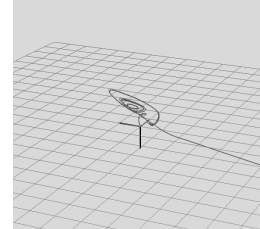


Fig. 20:  $q_1 = 0.30$

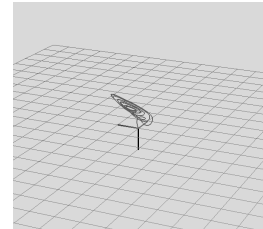


Fig. 21:  $q_1 = 0.50$

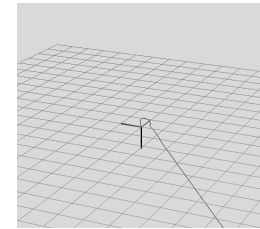


Fig. 22:  $q_1 = 0.90$

- 8) H. Suzuki, M. Minami, J. Agbanhan, "Fish Catching by Robot Using Gazing GA Visual Servoing", *Transaction of the Japan Society of Mechanical Engineers*, C-68-668, 1198/1206, (2002)
- 9) Jun Hirao and Mamoru Minami: "Intelligence Comparison between Fish and Robot using Chaos and Random", *International Conference on Advanced Intelligent Mechatronics*, 552/557, (2008)
- 10) K. Aihara, "Chaos in Neural System", 126/151, (1993)
- 11) R. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane", *Biophys.J.1*, 445/466, (1961)
- 12) C. T. Lin and C. S. Lee, "Neural Fuzzy Systems", *Englewood Cliffs, NJ:Prentice Hall PTR*, (1996)
- 13) Limin Peng and Peng-Yung Woo, "Neural-Fuzzy Control System for Robotic Manipulators", *IEEE Control Systems Magazine*, 53/63, (2002)
- 14) M. Suzuki, "Evolutionary acquisition of complex behaviors through intelligent composite motion control", *Computational Intelligence in Robot. Automat. CIRA*, vol.27-30, 231/236, (2005)
- 15) R. Endo, J. Hirao and M. Minami, "Intelligent Chaos Fish-Catching Based on Neural-Network-Differential-Equation", *SICE Annual Conference*, 1957/1962, (2010)
- 16) William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, "NUMERICAL RECIPES in C", *The Art Scientific Computing Second Edition*, (2002)
- 17) M. Minami, J. Hirao, "Intelligence comparison between fish and robot using chaos and random", *Int. Conf. on Advanced Intelligent Mechatronics*, Proc., 552/557, (2008)
- 18) Jun Hirao, Mamoru Minami, Yasushi Mae and Jingyu Gao, "Emergence of Robotic Intelligence by Chaos for Catching Fish", *SICE Annual Conference 2007*, 969/975, (2007)

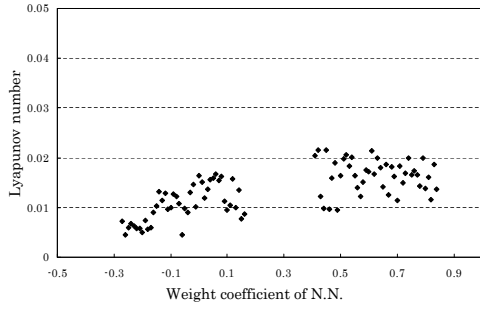


Fig. 23: Lyapunov number  $\lambda_1$

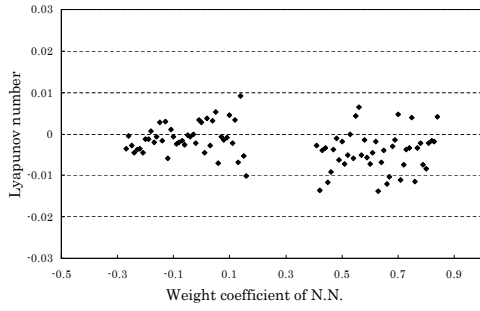


Fig. 24: Lyapunov number  $\lambda_2$

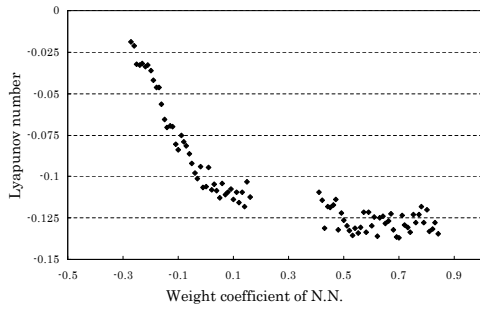


Fig. 25: Lyapunov number  $\lambda_3$

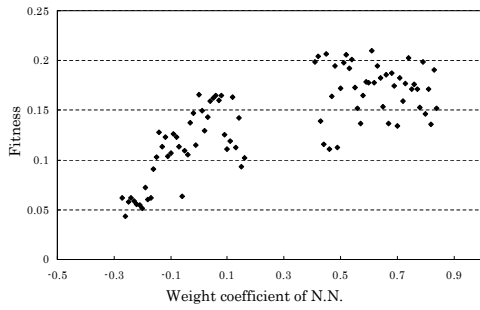


Fig. 26: Fitness  $g$

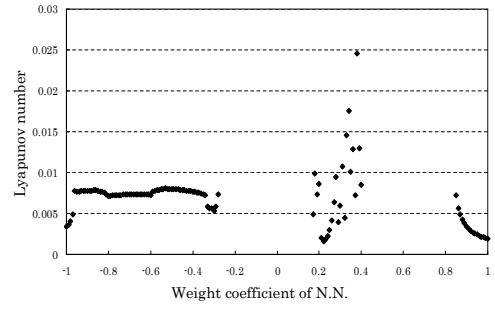


Fig. 27: Lyapunov number  $\lambda_1$  (not chaos)

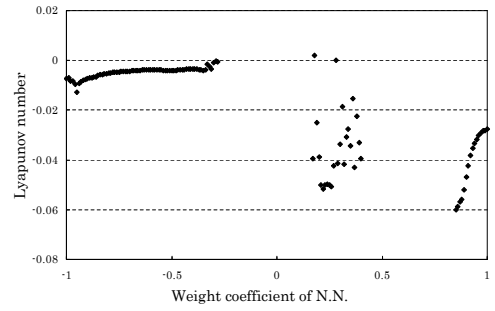


Fig. 28: Lyapunov number  $\lambda_2$  (not chaos)





# 多モード情報を統合する複数の複素 SOM による地雷概念形成

○江尻 礼聡 廣瀬 明 (東京大学)

## Generation of landmine concept among multiple complex-valued SOMs integrating multimodal information

\*A. Ejiri and A. Hirose (University of Tokyo)

**Abstract**— We generate concepts with integration of multimodal information. There have been many reports on concept generation in various fields such as linguistics and robotics. We propose a concept generation method using multiple self-organizing maps (SOMs) for a landmine visualization system where we integrate two modal information (space- and frequency- domain texture) by considering their mutual information. We also propose the use of similarity indices, instead of the mutual information, to reduce the calculation cost.

**Key Words:** Complex-valued SOM, integration of multimodal information

### 1 はじめに

ここ数年、ヒューマノイドロボットの研究が盛んである。2足歩行などのさまざまな機能を付与している。ヒューマノイドに求められているものは、人間と同程度、もしくはそれ以上に柔軟に様々なことを行えることである。そのため、作りこまなくても学習によって様々なことができるようになっていくロボットが、これから求められていくと思われる。特にコミュニケーションの学習に焦点をあてた場合、言語獲得が重要となる。言語獲得のためには、認知言語学の観点から考えると、概念獲得についての研究を行わなくてはならないと考えられる。本研究では、概念獲得に焦点をあて、複数モダリティから概念構築をするための、相互情報量を考慮した自己組織化マップについて研究を行い、これを利用した地雷探知システムを構築する。また地雷探知システムの実用化のため、計算コストの低減にも取り組む。

#### 1.1 ロボット分野における概念構築

何をもちて概念とよび、何をもちて概念構築を定義するかは議論の分かれるところではあるが、現在の研究は、おおまかに3種類に分けられる。どのように環境を捉えるのか、というものと、捉えることによってえられた情報をどのように蓄積するか、そして、どのように蓄積したデータ同士のつながりをみつけていくのかというものの3種類である。どのように環境をとらえるのか？という研究は、どのように事物を特徴量に落とすのか、その方法を探すという形で研究が行われている。情報をどのように蓄積していくのかという研究は、どのように蓄積したデータ同士のつながりをみつけていくのかというものと同時に研究されている場合が多い。特に概念構築の研究では、得られた情報の蓄え方+シンボルグラウンディングと考える方法が一般的となっている。Roy たちの研究では、情報の構築はしていないが、音声と画像の対応付けに相互情報量を用いている<sup>1) 2)</sup>。彼らの研究では、文字をつかっておらず、対応付けの正誤を与えなくても良いことから、新しい言葉にも対応できる柔軟な方法であると考えられる。本稿では、この Roy らの考え方をベースとし、相互情報量を用いて対応を探していくのと同時に、相互情報量によって同一とみなされたデータのまとまりを構築していく方法を提案する。

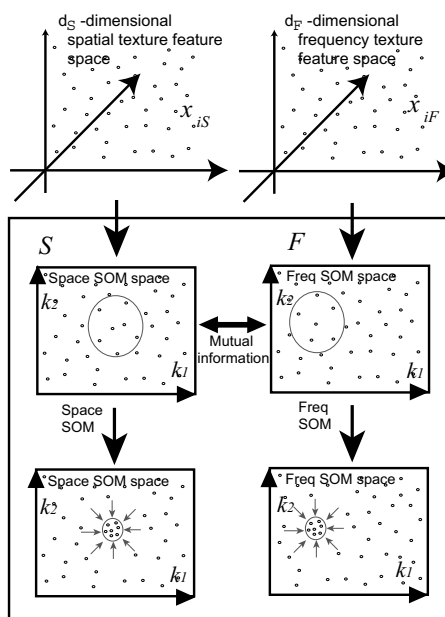


Fig. 1: 複数モード情報間の情報量に応じてダイナミクスが変調される複数 SOM による概念形成の様子を表す模式図<sup>5)</sup>。

### 2 概念形成と多モード情報統合

#### 2.1 基本的なアイデア

今回提案する多モード情報を統合する複数の複素 SOM による構築の全体の流れを Fig.1 に示す。各特徴空間を、いったん SOM<sup>3) 4)</sup> によって低次元にマッピングを行う。その際、対応している箇所を相互情報量をみることにより検出する。その相互情報量により学習の近傍関数を変えることによって、対応している箇所が、一箇所に集まってくるようにマッピングをしていく。これにより、一方のモダリティがもう一方の教師の代わりとなるため、教師なしの構築が可能となる。

通常の SOM のダイナミクスは、次のように表現できる。すなわち、時刻  $t$  に対する SOM 空間での位置  $k$  における参照ベクトル (ニューロンの荷重ベクトル)  $w_k(t)$  が、入力信号ベクトル  $x_i$  によって次のように更

新される。

$$\mathbf{w}_k(t+1) = \mathbf{w}_k(t) + \Theta(\mathbf{k}, \mathbf{k}_c; t) \alpha(t) (\mathbf{x}_i - \mathbf{w}_k(t)) \quad (1)$$

ただし、各変数と関数は次の通りである。

$$\mathbf{k}_c = \arg \min_{\mathbf{k}} |\mathbf{x}(t) - \mathbf{w}_k(t)| \quad (\text{勝者ニューロンの位置}) \quad (2)$$

$$\alpha(t) \equiv \alpha(0) \left(1 - \frac{t}{T}\right) \quad (\text{学習係数}) \quad (3)$$

$$\Theta(\mathbf{k}, \mathbf{k}_c; t) = \exp\left(-\frac{|\mathbf{k} - \mathbf{k}_c|^2}{2\delta(t)}\right) \quad (\text{近傍関数}) \quad (4)$$

$$\delta(t) \equiv \delta(0) \left(1 - \frac{t}{T}\right) \quad (\Theta \text{ の鋭さの逆数}) \quad (5)$$

また、 $T$  は自己組織化の最大時間（最大繰り返し回数）である。

しかし、われわれはそれぞれのモードの各入力信号の相関値により近傍関数の広さが変わるようにすることで、自己組織化ダイナミクスを変調する。その結果、互いに相棒を持つ情報がより強くクラスタリングを起こすようにできる。たとえば空間領域テクスチャ特徴量と周波数領域テクスチャ特徴量を考える地雷システムの場合、空間特徴量のどこからどこまでを一つの同じカテゴリとしてみるか、という判断を、周波数特徴量との相互情報量を観測することによって行える。周波数特徴量の側も、どこまでを同一データとしてみるのかを空間特徴量から判断することができる。

## 2.2 相互情報量によって変調される SOM

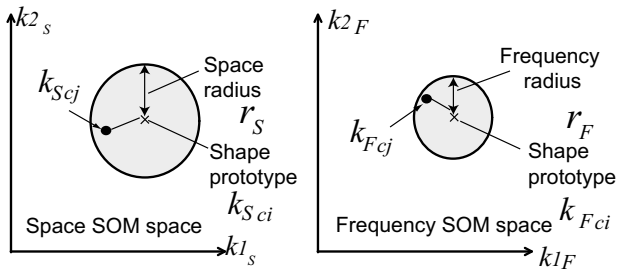


Fig. 2: 相互情報量を最大にする領域の決定の方法<sup>5)</sup>。

複数モードの情報間の関係性を考慮したクラスタリングを実現する、相互情報量を用いた SOM を提案する。なお詳細は、文献<sup>5)</sup>を参照されたい。

まず複数の特徴量（後述の地雷可視化の実験では、空間領域テクスチャ特徴量と周波数領域テクスチャ特徴量）をそれぞれ扱う 2 つの SOM を用意する。SOM は自己組織化をそれぞれ行うが、自己組織化の際の近傍関数の決定に、相互情報量を用いることによって、複数モードの情報間の関係性を反映させる。相互情報量  $I(S; F)$  は、以下の式によって計算される。

$$I(S; F) = \sum_{d_S=0,1} \sum_{d_F=0,1} p(d_S, d_F) \log \frac{p(d_S, d_F)}{p(d_S)p(d_F)} \quad (6)$$

$$p(d_S) = \frac{m(d_S)}{M} \quad (7)$$

$$p(d_F) = \frac{m(d_F)}{M} \quad (8)$$

$$p(d_S, d_F) = \frac{m(d_S, d_F)}{M^2} \quad (9)$$

$$d_S = \begin{cases} 1, & \text{if } |\mathbf{k}_{Sci} - \mathbf{k}_{Scj}| \leq r_S \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

$$d_F = \begin{cases} 1, & \text{if } |\mathbf{k}_{Fci} - \mathbf{k}_{Fcj}| \leq r_F \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

ただし、 $m(d_S)$  と  $m(d_F)$ （あるいは短く  $m(d_{S/F})$ ）は  $d_{S/F}$  を与えるデータ数、 $M$  は全データ数 すなわち集合  $\{\mathbf{x}_1 = [\mathbf{x}_{S1}, \mathbf{x}_{F1}], \mathbf{x}_2 = [\mathbf{x}_{S2}, \mathbf{x}_{F2}], \dots, \mathbf{x}_m = [\mathbf{x}_{Sm}, \mathbf{x}_{Fm}], \dots, \mathbf{x}_M = [\mathbf{x}_{SM}, \mathbf{x}_{FM}]\}$  の要素数、 $d_S(\mathbf{k}, \mathbf{k}_c)$ 、 $d_F(\mathbf{k}, \mathbf{k}_c)$  は勝者から更新の当該ニューラルネットワークまでの距離である。 $r_S, r_F$  を変えてゆき、 $I(S; F)$  が最大値  $I_{max}$  もになるものを探す。この  $I_{max}$  の値によって、SOM の自己組織化の近傍関数を次のように変えてゆく。

相互情報量  $I_{max}$  が閾値  $H$  以上になった場合、そのときの中心から  $r_S, r_F$  以内の範囲に含まれる入力データ  $\mathbf{x}_a$  に関する学習について、(5) の近傍関数  $\delta(t)$  を次の  $\tilde{\delta}(t)$  に変更する。

$$\tilde{\delta}(t) = \begin{cases} \delta(0) \left(1 - \frac{t}{T}\right)^2 & \text{for } \mathbf{k}_{S/Fcj} \in \text{circle of } \mathbf{k}_{S/Fci} \\ \delta(0) \left(1 - \frac{t}{T}\right) & \text{otherwise} \end{cases} \quad (12)$$

すなわち、(12) の上段の場合、近傍は鋭くなる。換言すれば、特徴量  $\mathbf{x}_j$  に対する勝者  $\mathbf{k}_{S/Fcj}$  がその領域に入る場合、かつ空間特徴量  $\mathbf{x}_S$  と周波数特徴量  $\mathbf{x}_F$  の相互情報量  $I_{max}$  が閾値  $H$  を越えている場合、近傍関数  $\Theta(\mathbf{k}, \mathbf{k}_c, t)$  の中の  $\tilde{\delta}(t)$  を小さくする。その結果として、特徴量  $\mathbf{x}_j$  の学習による素子の変化が鈍くなり、相互情報量の高い状態を維持することが容易になる。なお、局所的な状況に応じて  $H$  を変化させればより良い結果が得られると期待できるが、ここでは簡単に  $H$  を固定しておく。

## 3 実験

ここでは提案する複数 SOM によって実際に地雷概念の形成の実験を行った結果を報告する。

### 3.1 前処理と特徴量の抽出

Fig3 に地雷可視化システムが行う信号処理の全体の流れを示す。フロントエンドによって得られたデータ  $z(l_x, l_y, f_n)$  から、小さい空間ウィンドウ ( $L \times L$  ピクセル) の中にあるデータに関する局所的なテクスチャ特徴量ベクトル  $\mathbf{K}$  を抽出する。なお、このテクスチャ特徴量ベクトル  $\mathbf{K}$  の抽出の詳細は、文献<sup>6)</sup>などに記す。

### 3.2 実験結果

特徴量ベクトルは局所ウィンドウの数だけあり、この実験では  $M=360$  個ある。1 つの特徴量ベクトル  $\mathbf{x}_m = [\mathbf{x}_{Sm}, \mathbf{x}_{Fm}]$  は、ある 1 つの局所領域のテクスチャを表している。これらの特徴量のうち、 $\mathbf{x}_{Sm}$  が空間領域特徴量処理用の SOM（空間 SOM とよぶことにする）に、

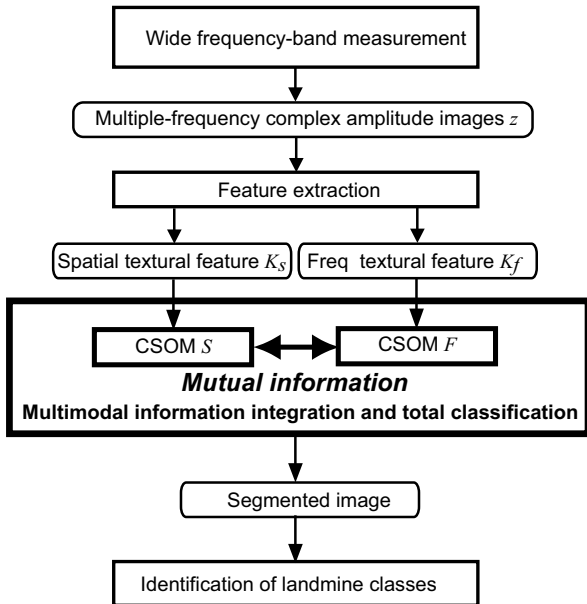


Fig. 3: 地雷可視化のための信号処理の全体の流れ。

$x_{Fm}$  が周波数領域特徴量処理用の SOM (周波数 SOM) にそれぞれ入力される。そして、SOM の自己組織化のダイナミクスによって、それぞれの SOM のニューロンが自己組織化する。その際、空間 SOM に入力される情報と周波数 SOM に入力される情報の間に相関があれば、それに関する自己組織化を調整する。

一方、従来の地雷可視化システムでは、 $x_m = [x_{Sm}, x_{Fm}]$  を一つのベクトルとしてまとめて扱い、1つの SOM によってそのまま自己組織化させていた。その場合でも、多くの場合に有効な自己組織化が行われ、地雷を表すテクスチャが適応的に区分化されて、地雷可視化が行われた。しかし本来、 $x_{Sm}$  と  $x_{Fm}$  は性質の異なる情報であり、このように合成することが必ずしも最適な扱い方であるとはいえない。むしろ、これらの情報の間の相関のあり方によって、自然に統合されることが望ましい。本提案は、そのような発想に基づく。

Fig.4~5 は、自己組織化が行われた後で、360 個のデータを入力したときに発火するニューロン (勝者ニューロン) の SOM 空間での位置を示したものである。ここでは、地雷可視化がやや難しい場合のデータを用いた。図 4 は、空間 SOM と周波数 SOM との間に特に相互作用はなく、従来どおりのダイナミクスによって自己組織化が行われた場合の結果である。この実験では、われわれは事前に地雷位置を把握している。△は地雷位置でのと特徴量に対する発火ニューロンの位置であり、●はそれ以外の主に土の部分での特徴量に対する発火位置である。これらを分別すれば、△が集まっていることがわかるが、分別されていない状態で地雷とそれ以外を区別することはあまり容易ではない。従来型の単純統合による扱いの方が優れているといえる。(ただし、SOM の方式自体が異なるので、単純比較はできない。)

一方、Fig.5 は、相互情報量を用いる本提案方式による結果である。地雷による△がそれ以外の●と分離されてゆくことが確認される。入力された空間領域テク

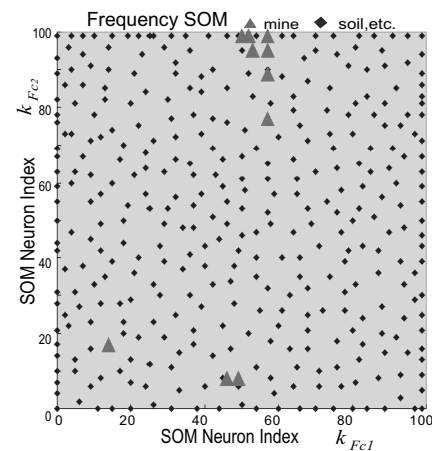
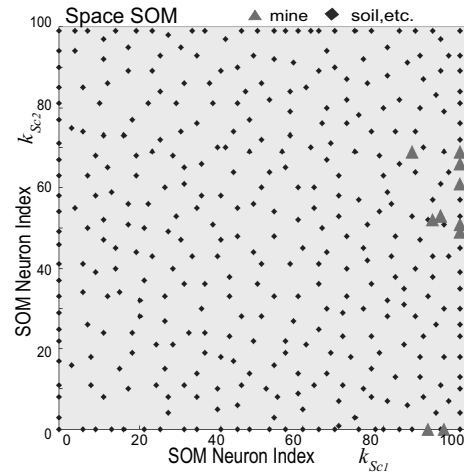


Fig. 4: 従来型 SOM で別々に処理した場合の全入力特徴量に対する勝者の分布。SpaceSOM (上) Frequency SOM (下) での勝者。△: 地雷位置で得られた入力特徴量に対する発火、●: それ以外の位置で得られた入力特徴量に対する発火。

スチャ特徴量と周波数領域特徴量を統合し、相関を観測しながら「地雷」という概念が形成されている、と解釈することが可能である。△領域に混じっている●も見られるが、これは石や土塊によるものかと考えられる。今後の検討が必要である。地雷以外の領域もやや集しつつある点も興味深い。以上のことから、提案方式が地雷の概念を自ら創出し、それを区分してゆくことが示された。

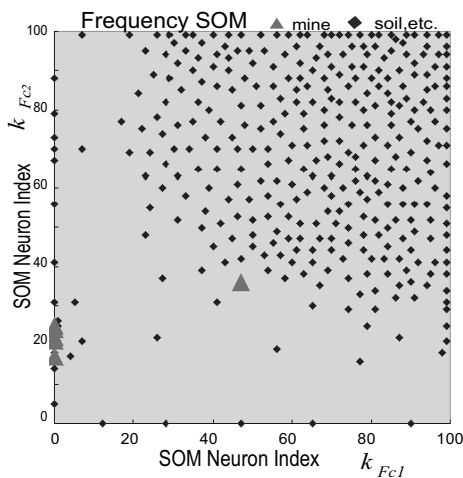
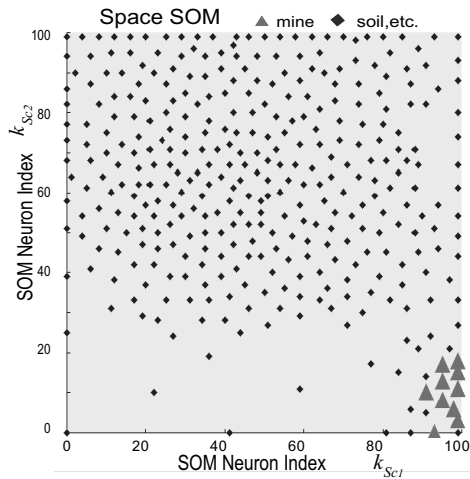


Fig. 5: MuSOM で情報統合しながら処理した場合の全入力特徴量に対する勝者の分布。SpaceSOM (上), Frequency SOM (下) での勝者。△:地雷位置で得られた入力特徴量に対する発火、●:それ以外の位置で得られた入力特徴量に対する発火。

## 4 地雷探知システムの実用化

前節では提案手法による地雷の概念形成の実験の結果を報告した。しかしこの地雷可視化システムを実用化するには問題が残っている。自己組織化後の相互情報量の計算コストは SOM の素子数に対して大きく増大し、実験で用いたような素子数  $100 \times 100$  の SOM など 500 回の自己組織化を行うと、非常に時間が掛かる。これは最大相互情報量  $I_{max}$  を決定するとき  $r_S, r_F$  を変えて計算を行うことで、計算量が膨大になっているためである。

### 4.1 類似度係数

提案手法では複数の SOM の自己組織化の状況の対応関係を相互情報量により数値化し、その値をみることで次の自己組織化に反映させたが、類似度係数によって複数の SOM の自己組織化の進行状況をモニターすることも可能である。この類似度係数を用いた評価法は相互情報量に比べ計算コストが掛からないため、素子数が多くなった場合の相互情報量の代案として考え

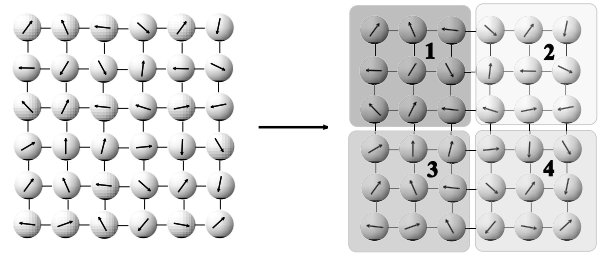


Fig. 6: 類似度係数の算出

られる。類似度係数は各 SOM 空間を Fig.6 のように複数に領域に分割し、領域内で発火した勝者ニューロンをその領域の特徴として扱うことで算出される。ある領域の特徴が、別の SOM 空間の領域の特徴と似ている場合に類似度係数は高くなり、該当する領域内の素子の自己組織化は変調される。

### 4.2 類似度係数と相互情報量の比較実験

類似度係数と相互情報量の評価を比較検証するため、素子数 10 の Ring 型 CSOM を用意し地雷可視化の実験と同様に空間領域と周波数領域のテクスチャ特徴量から地雷の概念を形成させる。このとき Ring 型 CSOM の素子の一つ一つが、10 分割された多素子の二次元 SOM の一つの領域として考えられる。この実験では類似度係数と相互情報量の純粋な評価能力を比較するため、これらの値による自己組織化の近傍関数の変調は行っていない。近傍関数の変調を行うためには、綿密な閾値の調整が必要だからである。実験には 3 種類の類似度係数を用いた。各類似度係数は以下のように算出される。

$$C_{j_{ij}} = \frac{C_{ij}}{S_i + F_j + C_{ij}} \quad \text{jaccard の群集係数}$$

$$C_{o_{ij}} = \frac{C_{ij}}{\sqrt{S_i} \sqrt{F_j}} \quad \text{落合の } C_o \text{ 指数}$$

$$C_{s_i} = \frac{2C_{ij}}{S_i + F_j} \quad \text{sorensen の } C_s \text{ 係数}$$

ただし  $C_{j_{ij}}, C_{o_{ij}}, C_{s_{ij}}$  は空間 SOM の  $i$  番目の素子と周波数 SOM の  $j$  番目の素子の類似度である。また  $S_i$  は空間 SOM の  $i$  番目の素子が勝者となった入力データの数、 $F_j$  は周波数 SOM の  $j$  番目の素子が勝者となった入力データの数、 $C_{ij}$  は空間 SOM の  $i$  番目の素子と周波数 SOM の  $j$  番目の素子が共通して勝者となった入力データの数である。最大自己組織化回数  $T$  まで自己組織化を終えた SOM 空間の素子同士の類似度係数は Fig.7~Fig.9 のようになった。X 軸、Y 軸が各 SOM の素子の番号を表す。グラフの点  $(x, y) = (i, j)$  の Z 座標は空間領域 SOM の  $i$  番目の素子と周波数領域 SOM の  $j$  番目の素子の類似度を表している。各類似度係数の間には顕著な違いは見られなかった。

次に相互情報量を算出する。相互情報量の算出は式 (6) に基づくが、今回は比較のため  $r_S = 0, r_F = 0$  で計算し、類似度係数と同様に各 SOM 空間の全て素子同士の相互情報量を算出した。その結果を Fig.10 に示す。各種類似度係数と似た形のグラフが得られた。相互情報量は一部の素子は高く出る一方で、それ以外の

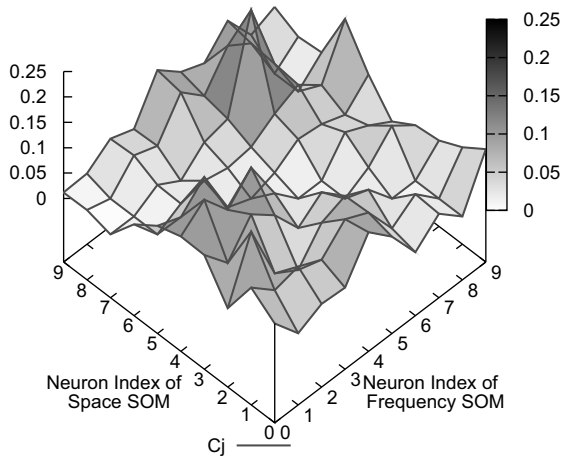


Fig. 7: jaccard の類似度係数

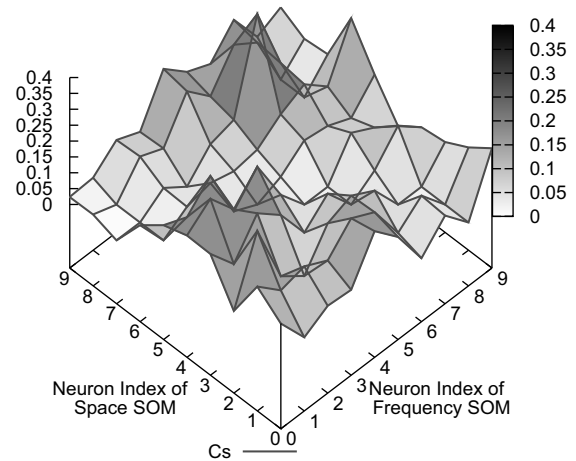


Fig. 9: Sorensen の  $C_s$  係数

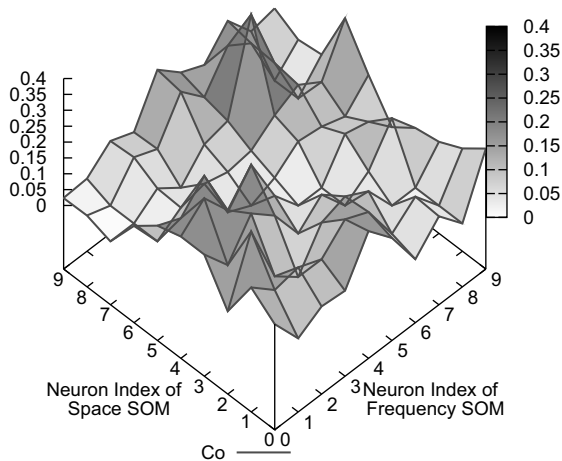


Fig. 8: 落合の  $C_0$  指数

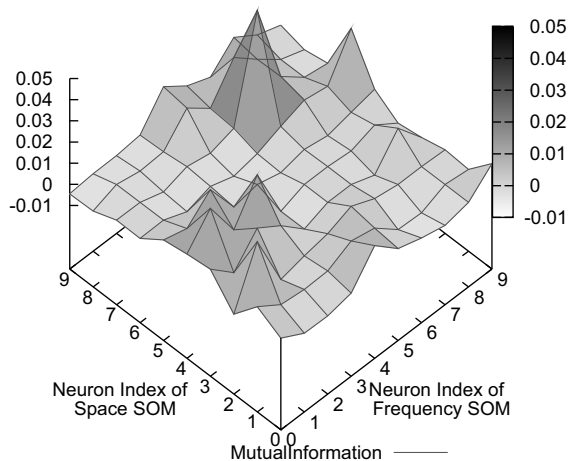


Fig. 10: 相互情報量

素子が低く二極化しているが、類似度係数の場合にはそのような特徴は見られず、各素子が様々な値を取るようにになっている。また将来的に自己組織化に反映させることも考えて、次のように各素子に付与させる類似度係数、相互情報量を定義する。

$$\begin{aligned}
 I s_i &\equiv \max_j I_{ij} \\
 I f_j &\equiv \max_i I_{ij} \\
 C s_i &\equiv \max_j C_{ij} \\
 C f_j &\equiv \max_i C_{ij}
 \end{aligned}$$

(13)

$I s_i$  は空間領域 SOM の  $i$  番目の素子に付与される相互情報量,  $I f_i$  は周波数領域 SOM の  $j$  番目の素子に付与さ

れる相互情報量,  $C s_i$  は空間領域 SOM の  $i$  番目の素子に付与される類似度係数,  $C f_j$  は周波数領域 SOM の  $j$  番目の素子に付与される類似度係数である。これらの素子に付与される相互情報量と各類似度係数を比較した結果は Fig.11 のようになり、正の相関が見られる。これは類似度係数が相互情報量に近い評価能力を持っていて、類似度係数による情報統合も可能であることを示している。

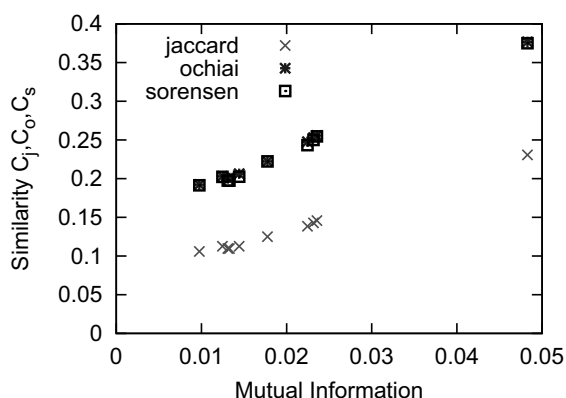


Fig. 11: 相互情報量と各類似度係数の相関

## 5 むすび

われわれは、地雷可視化システムにおける空間領域テクスチャ特徴量と周波数領域テクスチャ特徴量という2つの異なるモードの情報を統合しながら、自己組織的に地雷の概念を構築してゆく方法を提案した。具体的には、相互情報量に応じて、自己組織化マップのダイナミクスの中の近傍領域の範囲を変調するものである。実際に模擬地雷を使用した実験を行い、提案方式によってCSOMにおいて地雷部分のデータが効果的に集合し、地雷の概念形成および空間領域テクスチャと周波数領域テクスチャの情報統合が行われることを示した。また計算コスト低減のため、類似度係数による情報統合と相互情報量による情報統合の比較実験を行い、その可能性を示した。

## 参考文献

- 1) D. Roy and A. Pentland, "Learning words from sights and sounds: a computational model," *Cognitive Science*, vol. 26, no. 1, pp. 113/146(2002.)
- 2) D. Roy and N. Mukherjee, "Towards situated speech understanding: Visual context priming of language models," *Computer Speech and Language*, vol. 19, no. 2, pp. 227/248(2005)
- 3) T. Kohonen, S. Kaski, and H. Lappalainen, "Self-organized information of various invariant-feature in the adaptive subspace SOM," *Neural Computation*, vol. 9, pp. 1321/1344(1997).
- 4) T. Kohonen, "Generalization of the self-organizing map," in *Proceeding of International Joint Conference on Neural Networks*, pp. 457/462(1993).
- 5) K. Kitahara and A. Hirose, "A concept generation method based on mutual information quantity among multiple self-organizing maps," in *Proceedings of the International Conference on Neural Inform. Processing (ICONIP) 2009 Bangkok*, pp. 263/272(2009).
- 6) Y. Nakano and A. Hirose, "Improvement of plastic landmine visualization performance by use

of ring-som and frequency-domain local correlation," *IEICE Transactions on Electronics*, vol. E92-C, no. 1, pp. 102/108 (2009).

# 位相感受型ニューラルネットワークを用いたミリ波イメージングシステム

○小野島昇吾 廣瀬明 (東京大学)

## Adaptive millimeter-wave imaging system based on phase-sensitive neural networks

\* S.Onojima and A.Hirose (The University of Tokyo)

**Abstract**— We construct an adaptive millimeter-wave security imaging system that can visualize objects concealed under jackets or bags of passengers walking at Shinkansen and general railway gates. We aim to solve the following two problems in the conventional millimeter-wave systems: (1) measurement time too long to observe walking persons and (2) privacy invasion caused by the see-through nature of millimeter-wave. We develop an array antenna integrated with envelope-phase detecting (EPD) front-end circuit. We also employ our complex-valued self-organizing map (CSOM) that realizes automatic visualization of objects even for a low spatial-resolution data free of the privacy problem.

**Key Words:** complex-valued neural network, coherent active imaging, envelop phase detecting (EPD)

### 1 はじめに

近年、ミリ波と呼ばれる電磁波が私たちの身の回りで役立てられるようになってきている。その中で特に注目されてきているのがミリ波イメージング技術であり<sup>1)</sup>、空港などで乗客が危険物を保持していないかを探知するスキャンシステムへの応用に利用されている。

従来のスキャンシステムでは、測定に時間がかかるため、乗客は静止しないと測定できない、ミリ波は衣服を透過するため撮像画像のプライバシーがあるといった二つの問題があった。

そこで我々は、これらの問題に対し以下の手法を提案した。一つ目は、受信側にアレイアンテナを用いて乗客がスキャンを通過する間に測定を行い測定時間の短縮を目指すというものである。二つ目は、撮像画像の解像度を低くしても物体区分ができるよう複素ニューラルネットワークを利用するというものである。<sup>2)</sup>

本研究ではこれらの手法を用いたイメージングを行う図1のようなスキャンシステムの構築を目指す。

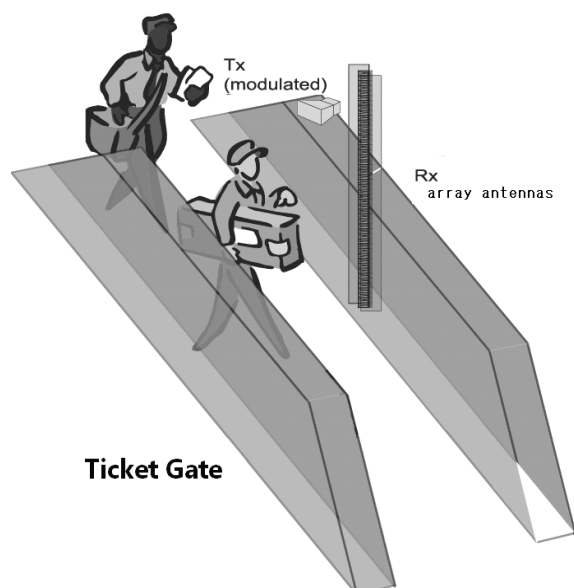


図1 構築を目指すスキャンシステム

### 2 スキャンシステムの構成

現時点で製作できたシステムの構成を図2に示す。キャリア周波数として34GHzのミリ波を利用し、検波においては1GHz程度の周波数帯を利用する。そのため進行方向への空間解像度は1GHzの波長である30cm程度となる。測定の流れとしてはまず34GHzと1GHz程度の電磁波をミキシングした後1つのホーンアンテナから送信する。次に被写体に反射し戻ってきた反射波を8つのアンテナで受信する。このアンテナは我々が設計・製作したもので、受信後すぐに全波整流し Envelop Phase Detection (EPD)<sup>3)</sup>を行うことができる回路が組み込まれている。その後アンテナで受信された電磁波はIQ demodulatorを用いてホモダイン検波を行うことで、振幅だけでなく位相のデータも取得する。それらの振幅と位相のデータはA/D変換を通してPersonal Computer (PC)に渡され処理される。

以上の測定までの流れではLabVIEWプログラミング環境を利用して、Signal Generator (SG)やA/D変換からのデータ取得を行っている。続く処理部分においてもLabVIEWプログラミング環境を用いてニューラルネットワークの一種である複素Self Organizing Map (CSOM)のアルゴリズムをプログラミングした。CSOMの処理では、まず取得データから局所領域に対する実空間と周波数に関する特徴量を算出する。その後それらの特徴量を用いてSOMの自己組織化を行い、適応的に実空間に対して区分を行う。これらの測定から処理の一連の流れはLabVIEWのプログラムによってオンラインで行われる。

### 3 フロントエンド

#### 3.1 アンテナの特性測定

システム構築のための予備実験として、製作したアンテナの特性を測定した。その特性とは、受信アンテナとIQ demodulatorを通して測定される位相と受信

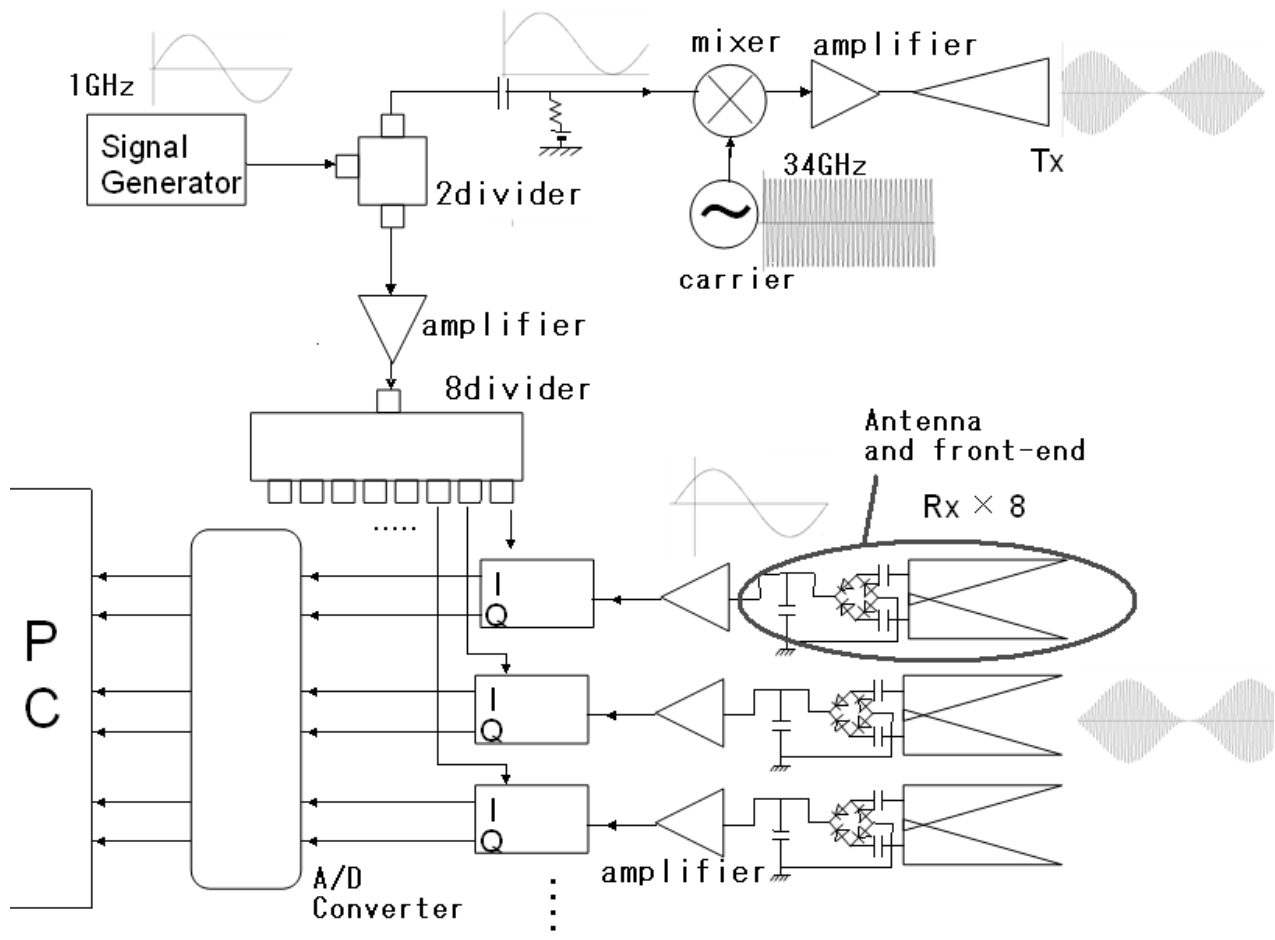


図2 システム構成図

強度のデータが距離に対してどのように変化していくかというものである。

図3に実験で用いた装置を示す。送信アンテナ(Tx)では34GHz帯に利用できるホーンアンテナを使用している。送信アンテナから受信アンテナ(Rx)の距離を2~6cm間で0.1cmずつ離していった。その際に位相と受信強度のデータを取得した。

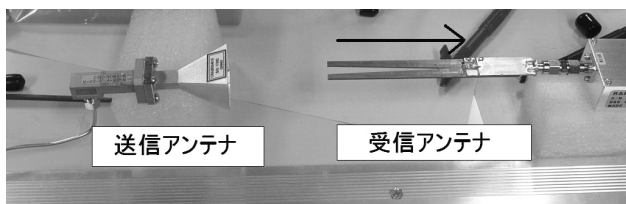


図3 アンテナの特性測定の実験装置

### 3.2 測定結果

図4,5に測定結果を示す。受信強度に関しては2~4cm間は離れるほど減衰していったが、4~6cm間は停滞していた。また、位相に関してはほぼ一致していたが、4~6cm間は傾きが大きくなっていった。このように4cm以降に変化が起きた原因はテーブルからの反射波をアンテナが受信してしまったためだと考えられる。

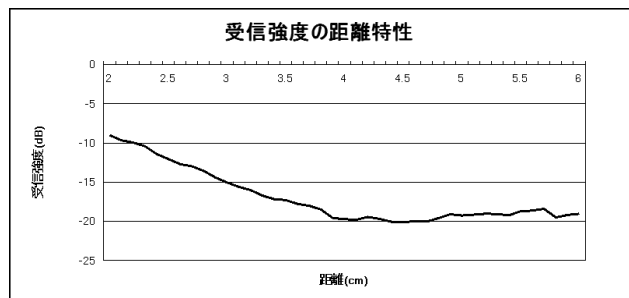


図4 距離に対する受信強度の変化

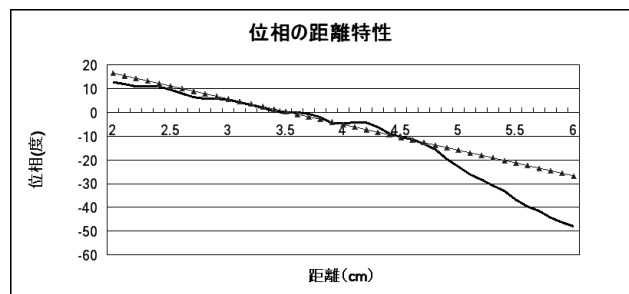


図5 距離に対する位相の変化



## 4 CSOM による適応区分

### 4.1 単体アンテナでの測定及び CSOM による処理

アレイアンテナの構築のための予備実験として単体のアンテナを利用して測定を行った。フロントエンドは図6のように送信用のホーンアンテナとわれわれが作成した受信用のアンテナが隣り合って固定されている。測定対象としては図7のようなアルミ箔で包まれた高さの異なる2つの箱と水の入ったペットボトルである。対象物の下には電波吸収剤を敷いてある。XYステージを利用してフロントエンドを50cm四方に走査し測定を行った。測定点の実空間解像度は1cmとし2500点のデータを取得し、各点に対しSGの周波数を860MHzから32MHz間隔で3つの周波数を利用してデータを測定した。測定後、CSOMによる適応区分を行った。

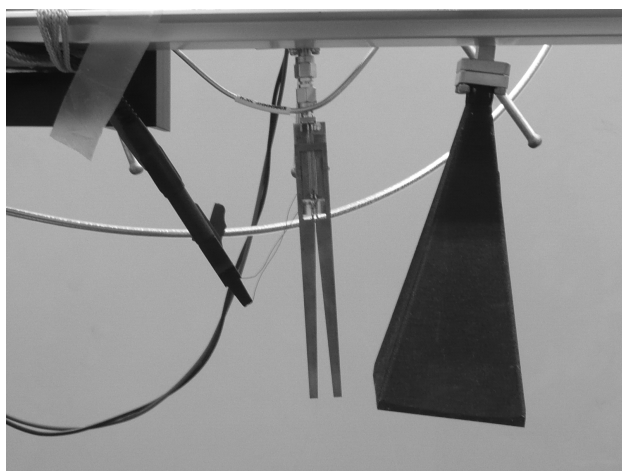


図6 フロントエンド

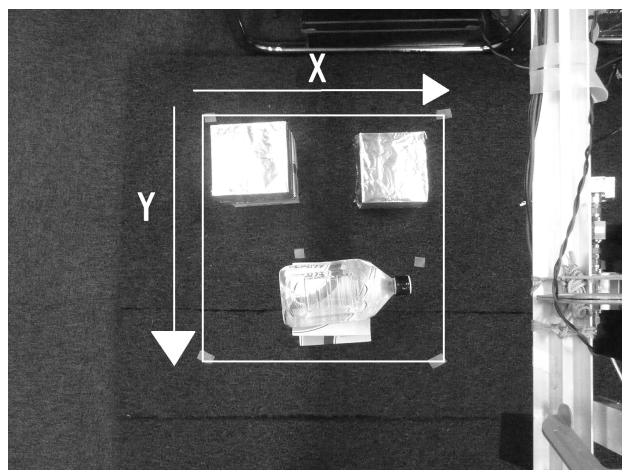


図7 測定対象

### 4.2 CSOM による適応区分の結果

取得した振幅と位相のデータのうち周波数が840MHzのものを図8,9に示す。2つの図のx軸及びy軸は図7で表示されているx方向とy方向と同じ向きである。目盛りの間隔は1cmである。図7と比較する

と、アルミ箔の箱やペットボトルの置かれている位置では振幅と位相ともに周囲に比べて変化があるように見える。これらの振幅と位相のデータから実空間の特徴量と周波数空間の特徴量を計算しCSOMで処理して区分化した結果が図10である。図の位置や向きは図7のものと同じである。図10と図7を比較してみると箱やペットボトルの位置だけ周辺とは異なるよう区分されており、図8,9に比べピークが抑えられて区分化することができている。

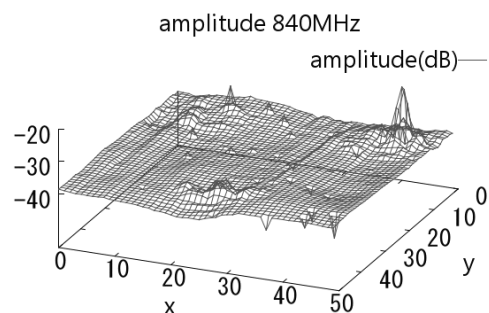


図8 840MHzで測定された振幅の生データ

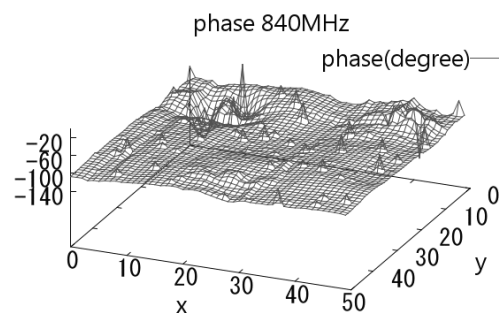


図9 840MHzで測定された位相の生データ

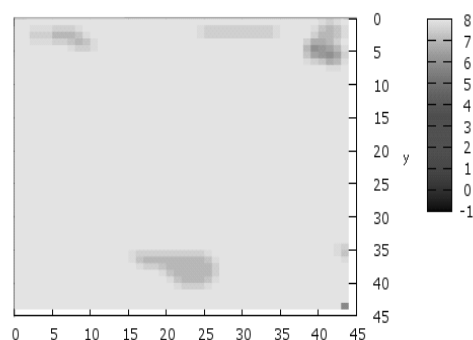


図10 CSOMによって区分化された判断画像

### 4.3 処理速度の比較

CSOMの処理速度を以前別の研究で開発されたMATLABによる処理のものと比較した。測定する時間は特徴量抽出にかかる時間とSOMによる学習および適応区分にかかる時間の2つである。今回は単にCSOMの

処理時間を計測することが目的であるため、以前の研究で用いられていたデータを利用した。

処理における計算時間の結果を図 11 に示す。特徴量抽出では LabVIEW の方が時間がかかってしまったが、CSOM での処理は速かった。これは、MATLAB が行列計算に優れているため特徴抽出の際並列処理により CSOM における自己組織化の計算が速く行われたためだと考えられる。

	LabView	MATLAB
特徴量抽出(ms)	3044.7	663.7
SOM(ms)	2291.6	14125.0

図 11 処理速度の比較

## 5 まとめ

我々は現在、送信アンテナから振幅変調されたミリ波を放射し受信のアレイアンテナで EPD を利用した検波を行い、取得した位相と受信強度のデータから CSOM を利用し画像の区分化をオンラインで行うスキミングシステムの構築を行っている。今回はその予備実験となるアンテナの特性測定、CSOM の処理速度の実験、単体アンテナを利用した測定から CSOM による適応区分の実験を行った。

## 参考文献

- 1) A. Dallinger, S. Bertl, and J. Detlefsen: Coherent millimeter-wave imaging for security applications, Radar Conference, 2007. EuRAD 2007. European, 28/31 (2007)
- 2) Takashi Aoyagi, Damri Radenamad, Yukimasa Nakano, and Akira Hirose: Complex-valued self-organizing map clustering using complex inner product in active millimeter-wave imaging, Neural Networks (IJCNN), The 2010 International Joint Conference on, 1/6 (2010)
- 3) A. Hirose, S. Hamada and R. Yamaki: Envelope phase detection for millimetre-wave active imaging, Electronics letters, **45**-6, 331/332 (2009)

# 複素行列因子分解と音響経路推定に基づく自動採譜

○池内亮太 池田和司 (奈良先端科学技術大学院大学)

## An Automatic Music Transcription Based on Complex Nonnegative Matrix Factorization and Sound Path Estimation

\*R. Ikeuchi and K. Ikeda (Nara Institute of Science and Technology)

**Abstract**— An automatic music transcription method is proposed that assumes a generative model for music sounds. The model takes into account the translation of spectrum for an instrument, and the sound path from the instrument to a microphone. The fundamental frequency (note), the spectrum of the instrument (basis pattern) and the sound path are estimated simultaneously using an extended complex nonnegative matrix factorization. The effectiveness of the proposed method is confirmed by computer simulations.

**Key Words:** Music transcription, Nonnegative matrix factorization

### 1 はじめに

音楽情報処理において重要な機能の一つが自動採譜である<sup>1, 2, 3, 4, 5, 6, 7</sup>。

楽器音は音声に比べ、一般にスペクトルの多様性は少ない。しかし、倍音が豊富に含まれていること、複数楽器の複数音階が含まれていることなどから、自動採譜はいまだに十分な精度を持つアルゴリズムが構築されていない。

音楽情報処理が音声情報処理と異なるもう一つの点は、音源からマイクまでの音響経路が存在することである。音声認識では多くの場合、話者はマイクの近くにおり、その経路の違いは無視できる。しかし自動採譜においては、どのような音響条件下で録音がなされたかはその性能に大きな影響を与える。実際、異なるCDに収録された音楽で学習された判別器は同一のCDの場合よりも性能が低いことが実験的に示されている<sup>8, 9</sup>。しかしこれまでの採譜システムでは、この点はほとんど考慮されてこなかった。

本研究では上記の点を考慮し、通常モデルの後に音響経路を設定した生成モデルに基づく自動採譜法を提案する。さらに、楽器音スペクトルに対しシフト不変性を仮定する。すなわち、同じ楽器の各音階のスペクトルは、あるパターンをシフトしたものだとする (Fig. 1)。

この時、与えられたスペクトログラムからモデルパラメータを推定する問題は、行列因子分解 (matrix factorization, MF) として定式化される。音響信号の振幅は一般に非負であるため、非負 MF (non-negative MF, NMF)<sup>10</sup> が用いられることが多いが、実際には複素信号であることから、Kameoka et al. による複素 NMFの方が適切である<sup>11</sup>。本研究でも複素 NMF を拡張して利用する。

### 2 楽器音の生成モデル

本論文で扱うモデルは、楽譜に相当するアクティビティ、楽器音スペクトルに相当する基底パターン、楽器から録音マイクまでの音響経路で構成される。時刻  $t = 1, \dots, T$  の周波数  $f = 1, \dots, F$  の振幅および位相は、複素数  $Y_{f,t}$  によって

$$Y_{f,t} \approx X_{f,t} = r_f \sum_k B_{f,k} A_{k,t} \exp(j\phi_{k,f,t}) \quad (1)$$

と表される。ここで  $A_{k,t}$ ,  $B_{f,k}$  および  $r_f$  はそれぞれアクティビティ、基底パターン、音響経路の周波数応答である。それぞれの詳細は以下の通りである。

#### 2.1 アクティビティ

本研究では西洋音楽の記法を対象とし、音楽は音階を示す音符で構成されていると考える。アクティビティは行列  $A = \{A_{k,t}\}$  で表し、 $A_{k,t}$  はある楽器の特定の音階  $k$  の時刻  $t$  での強さを示す。一般に  $A$  はスパース表現を持つ、すなわちほとんどの  $k$  および  $t$  について  $A_{k,t} = 0$  が成立することに注意する。

#### 2.2 基底パターン

基底パターン行列  $B = \{B_{f,k}\}$  の各行ベクトル  $B_k$  はある楽器の特定の音階のスペクトルを表す。以下では楽器が一つの場合を考えるが、これは複数楽器の場合にも容易に拡張できる。

同じ楽器については異なる音階のスペクトルは似た形をしている。このスペクトルの類似性は、これまでの研究でも利用されている<sup>12, 13, 14</sup>。Eggert et al. では、同一楽器の基底パターン  $B_k$  はあるパターンを平行移動して得られると仮定する。すなわち、

$$B_k = S^{k-1}b \quad (2)$$

である。ここで  $S$  はシフト行列

$$S = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} \quad (3)$$

であり、 $b$  は未知の定ベクトルである<sup>12</sup>。ここで、最左下成分を 1 にすると、理論的解析が容易になる<sup>9</sup>。Raczynski et al. は基底行列の初期値に、基本周波数の部分以外をゼロに設定したものを用いた<sup>13</sup>。Ochiai et al. はスペクトルの平行移動ではなくなめらかな変化を仮定し、隣接した音階ではスペクトルが類似しているというペナルティ関数を導入した<sup>14</sup>。本研究では、最も単純なモデルである Eggert et al. の方法を用いる。

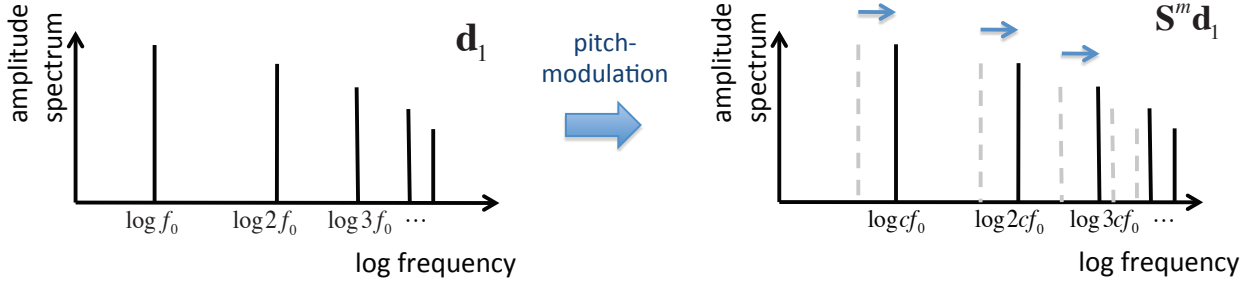


Fig. 1: Spectra of different tones are almost translation of each other.

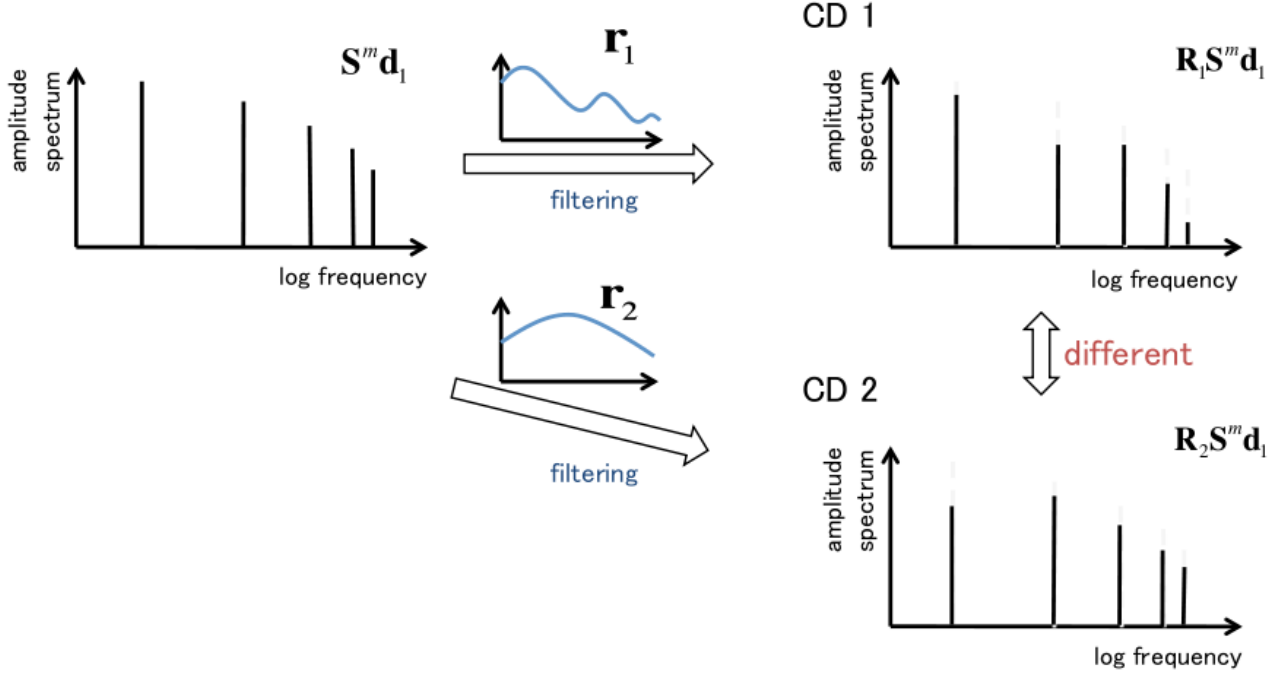


Fig. 2: Schematic view of the proposed model. A sound path is settled from an instrument to a microphone.

### 2.3 音響経路

本研究のモデルの特徴は、楽器とマイクの中に音響経路の存在を仮定する点である (Fig. 2). この経路は音楽が録音された環境に依存し、その周波数応答を  $r = \{r_f\}$  で表す.

## 3 複素行列因子分解

(1) のようなモデルは非負行列因子分解 (NMF) を用いて解くことができ<sup>10)</sup>, 音楽情報処理への応用も数多くの例がある<sup>12, 13, 14)</sup>. しかし従来の NMF は音信号の位相を考慮しておらず, そのまま用いるのは適切ではない. すなわち複素数については  $|a + b| \neq |a| + |b|$  である点に注意する必要がある. Kameoka et al. は時間に依存する位相スペクトル  $\phi_{k,f,t}$  を各項  $B_{f,k} A_{k,t}$  に導入し, この問題を克服する複素 NMF を提案した<sup>11)</sup>. 本研究では (1) のように音響経路  $r$  があるので, 項を  $r_f B_{f,k} A_{k,t}$  に置き換えて複素 NMF を用いる.

複素 NMF ではコスト関数

$$\begin{aligned} \Psi(\theta) &= \sum_{f,t} |Y_{f,t} - X_{f,t}|^2 + 2\lambda U & (4) \\ &= \sum_{f,t} \left| Y_{f,t} - r_f \sum_k B_{f,k} A_{k,t} \exp(j\phi_{k,f,t}) \right|^2 + 2\lambda U, & (5) \end{aligned}$$

を拘束条件

$$\sum_f r_f^2 = 1, \quad \sum_f b_f^2 = 1, \quad (6)$$

の元で最小化する. ここで  $\lambda$  は二つの項の重み付けをするパラメータ,  $\theta = \{r, b, A, \phi\}$  であり,

$$U = \sum_{k,t} |A_{k,t}|^p \quad (0 < p < 2) \quad (7)$$

は行列  $A$  にスパース性を導入するための正規化項である. 拘束条件はスケールの不定性を除去する役割を果たす.

なお,  $B$  のシフト不変性により  $r$  および  $b$  には一般に不定性は生じない. これは直感的には,  $r$  と  $b$  の自

自由度が  $2F$  であるのに対し、そのままの行列  $\text{diag}(r)B$  では自由度が  $FT$  であることによる。

$\Psi(\theta)$  を効率的に最小化する反復法は、既存の方法<sup>3, 11)</sup>と同様に補助関数  $\Psi^+(\theta, \bar{\theta})$  を用いて導出できる。 $\Psi^+(\theta, \bar{\theta})$  が

$$\Psi(\theta) = \min_{\bar{\theta}} \Psi^+(\theta, \bar{\theta}) \quad (8)$$

を満たす時、 $\Psi(\theta)$  は

$$\bar{\theta} \leftarrow \arg \min_{\bar{\theta}} \Psi^+(\theta, \bar{\theta}), \quad \theta \leftarrow \arg \min_{\theta} \Psi^+(\theta, \bar{\theta}). \quad (9)$$

という更新に対して非増加関数であることが知られている。

本研究では、補助関数として Jensen の不等式に基づく

$$\begin{aligned} \Psi^+(\theta, \bar{\theta}) &\equiv \sum_{k,f,t} \frac{1}{\beta_{k,f,t}} \left| \bar{Y}_{k,f,t} \right. \\ &\quad \left. - r_f \sum_{f'} (S^{k-1})_{f,f'} b_{f'} A_{k,t} \exp(j\phi_{k,f,t}) \right|^2 + U^+ \quad (10) \end{aligned}$$

$$\begin{aligned} &\geq \sum_{f,t} \left| Y_{f,t} \right. \\ &\quad \left. - r_f \sum_k \sum_{f'} (S^{k-1})_{f,f'} b_{f'} A_{k,t} \exp(j\phi_{k,f,t}) \right|^2 + U \quad (11) \end{aligned}$$

$$= \Psi(\theta), \quad (12)$$

を用いる。ここで

$$\begin{aligned} U^+ &= \sum_{k,t} \left( \frac{p|\bar{A}_{k,t}|^{p-2}}{2} A_{k,t}^2 + |\bar{A}_{k,t}|^p - \frac{p|\bar{A}_{k,t}|^p}{2} \right) \\ &\geq U \quad (13) \end{aligned}$$

および

$$\sum_k \bar{Y}_{k,f,t} = Y_{f,t}, \quad \sum_k \beta_{k,f,t} = 1 \quad \beta_{k,f,t} > 0 \quad (14)$$

である。ここで、(11) の等号成立は

$$\begin{aligned} \bar{Y}_{k,f,t} &= r_f \sum_{f'} (S^{k-1})_{f,f'} b_{f'} A_{k,t} \exp(j\phi_{k,f,t}) \\ &\quad + \beta_{f,t,k} \left( Y_{f,t} - \sum_k r_f \sum_{f'} (S^{k-1})_{f,f'} b_{f'} A_{k,t} \exp(j\phi_{k,f,t}) \right) \quad (15) \end{aligned}$$

および  $\bar{A}_{k,t} = A_{k,t}$  の時であることに注意する。

$r, b, A$ , および  $\phi$  の更新式は、以下のように与えられる。なお、拘束条件 (6) は  $r$  および  $b$  の大きさを更新後に正規化することで実現する。最小値の必要条件式

$$\frac{\partial \Psi^+}{\partial r_f} = 0, \quad \frac{\partial \Psi^+}{\partial b_l} = 0, \quad \frac{\partial \Psi^+}{\partial A_{k,t}} = 0, \quad (16)$$

を解いて、

$$r_f = \frac{\sum_{k,t} \frac{A_{k,t}}{\beta_{k,f,t}} \sum_{f'} S_{f,f'}^{k-1} b_{f'} \Re \left[ \bar{Y}_{k,f,t}^* \exp(j\phi_{k,f,t}) \right]}{\sum_{k,t} \frac{A_{k,t}^2}{\beta_{k,f,t}} \left( \sum_{f'} S_{f,f'}^{k-1} b_{f'} \right)^2}, \quad (17)$$

$$\begin{aligned} b_l &= \frac{\sum_{k,f,t} \frac{r_f A_{k,t} S_{f,l}^{k-1}}{\beta_{k,f,t}} \Re \left[ \bar{Y}_{k,f,t}^* \exp(j\phi_{k,f,t}) \right]}{\sum_{k,f,t} \frac{r_f^2 A_{k,t}^2 (S_{f,l}^{k-1})^2}{\beta_{k,f,t}}} \\ &\quad - \frac{\sum_{k,f,t} \frac{r_f A_{k,t}^2 S_{f,l}^{k-1}}{\beta_{k,f,t}} \left( \sum_{f'} S_{f,f'}^{k-1} b_{f'} - S_{f,l}^{k-1} b_l \right)}{\sum_{k,f,t} \frac{r_f^2 A_{k,t}^2 (S_{f,l}^{k-1})^2}{\beta_{k,f,t}}}, \quad (18) \end{aligned}$$

$$A_{k,t} = \frac{\sum_f \frac{r_f}{\beta_{k,f,t}} \sum_{f'} S_{f,f'}^{k-1} b_{f'} \Re \left[ \bar{Y}_{k,f,t}^* \exp(j\phi_{k,f,t}) \right]}{\sum_f \frac{r_f^2}{\beta_{k,f,t}} \left( \sum_{f'} S_{f,f'}^{k-1} b_{f'} \right)^2 + \lambda p |\bar{A}_{k,t}|^{p-2}}, \quad (19)$$

が得られる。ここで、

$$S_{f,l}^{k-1} = \begin{cases} 1 & \text{for } k = l - f + 1 \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

により (18) は

$$\begin{aligned} b_l &= \frac{\sum_{f,t} \frac{r_f A_{k',t}}{\beta_{k',f,t}} \Re \left[ \bar{Y}_{f,t}^* \exp(j\phi_{k',f,t}) \right]}{\sum_{f,t} \frac{r_f^2 A_{k',t}^2}{\beta_{k',f,t}}} \\ &\quad - \frac{\sum_{f,t} \frac{r_f A_{k',t}^2}{\beta_{k',f,t}} \left( \sum_{f'} S_{f,f'} A_{k',t} b_{f'} - S_{f,l,k'} b_l \right)}{\sum_{f,t} \frac{r_f^2 A_{k',t}^2}{\beta_{k',f,t}}}, \quad (21) \end{aligned}$$

と簡単化される。ただし

$$k' = l - f + 1, \quad 1 \leq k' \leq K \quad (22)$$

である。

$\phi_{k,f,t}$  の更新式は  $\Psi^+$  を  $\phi$  に依存する項とそれ以外に分ける、すなわち

$$\begin{aligned} \Psi^+ &= c - 2 \sum_{k,f,t} \frac{r_f A_{k,t}}{\beta_{k,f,t}} \sum_{f'} S_{f,f'}^{k-1} b_{f'} \\ &\quad \times |\bar{Y}_{k,f,t}| \cos(C_{k,f,t} - \phi_{k,f,t}) \quad (23) \end{aligned}$$

とすることで得られる。ここで  $C_{k,f,t}$  は

$$\cos C_{k,f,t} = \frac{\Re \left[ \bar{Y}_{k,f,t} \right]}{|\bar{Y}_{k,f,t}|}, \quad \sin C_{k,f,t} = \frac{\Im \left[ \bar{Y}_{k,f,t} \right]}{|\bar{Y}_{k,f,t}|} \quad (24)$$

で与えられ、 $c$  は  $\phi_{k,f,t}$  に依存しない定数である。(23) により、 $\Psi^+$  は  $\phi_{k,f,t} = C_{k,f,t}$  の時に最小値をとることがわかる。

(14) の自由パラメータ  $\beta$  は Kameoka et al.<sup>11)</sup> に従い、

$$\beta_{k,f,t} = \frac{A_{k,t} \sum_{f'} S_{f'}^k b_{f'}}{\sum_n A_{n,t} \sum_{f'} S_{f'}^k b_{f'}} \quad (25)$$

と定める。

## 4 実験と結果

提案手法の有効性を確かめるため、以下のような実験を行った。自動採譜問題は不良設定問題なので、ここではアクティビティ、基底パターンおよび音響経路を既知として合成した音信号を用いている (Fig. 3)。

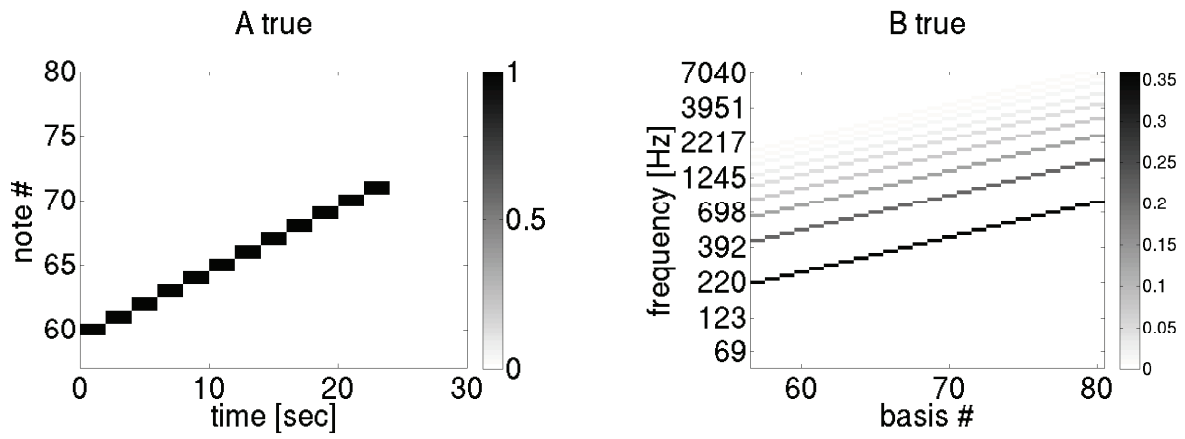
Fig. 4 は  $\lambda = 0.1$  とした時の提案手法の結果を示している。提案アルゴリズムが、アクティビティと基底パターンをほぼ正しく推定できていることがわかる。なお、音信号が存在しない周波数帯域では、音響経路の周波数応答はゼロと推定されている。

## 5 結論

本論文では、楽器音スペクトルのシフト不変性と、楽器からマイクへの音響経路を考慮した自動採譜法を提案した。複素 NMF を拡張した手法により、アクティビティと楽器音の基底パターン、音響経路が同時に推定できることが、簡単な人工データにより示された。今後、本手法を実データに応用し、その優位性を示していきたい。

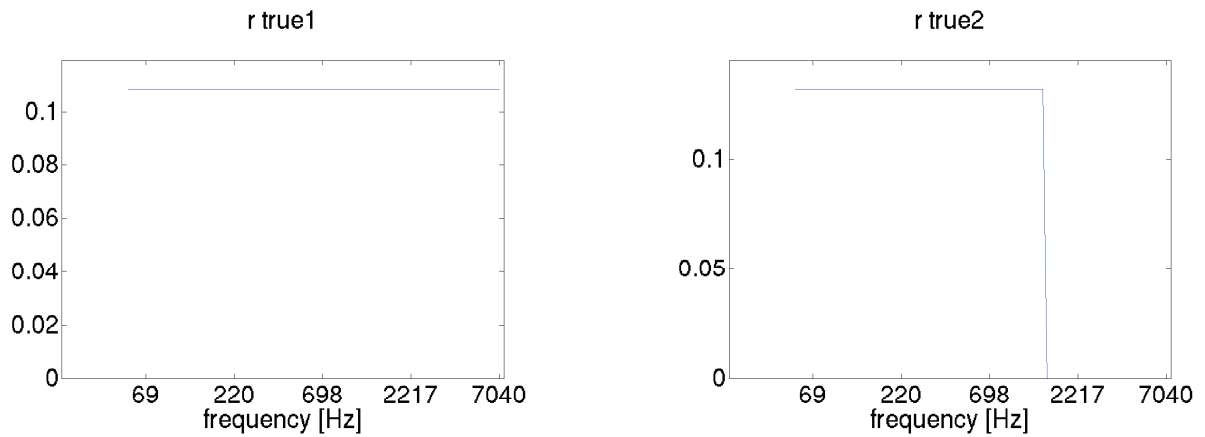
## 参考文献

- 1) Klapuri, A., Davy, M. (eds.): Signal Processing Methods for Music Transcription, Springer (2006)
- 2) Lee, D.D., Seung, H.S.: Learning the Parts of Objects by Non-Negative Matrix Factorization, Nature, **401**, 788/791 (1999)
- 3) Lee, D.D., Seung, H.S.: Algorithms for Nonnegative Matrix Factorization, NIPS, 556/562 (2000)
- 4) Plumbley, M., Abdallah, S., Bello, J., Davies, M., Monti, G., Sandler, M.: Automatic Music Transcription and Audio Source Separation, Cybernetics and Systems, **33-6**, 603/627 (2002)
- 5) Blumensath, T., Davies, M.: Sparse and Shift-Invariant Representations of Music, IEEE Trans. ASLP **14-1**, 50/57 (2006)
- 6) Ihara, M., Maeda, S., Ishii, S.: Estimation of the Source-Filter Model Using Temporal Dynamics, Proc. IJCNN, 3098/3103 (2007)
- 7) Ihara, M., Maeda, S., Ishii, S.: Instrument Identification in Monophonic Music Using Spectral Information, Proc. ISSPIT, 607/611 (2007)
- 8) Ihara, M.: Statistical Approach to the Single-Channel Sound Source Extraction, Ph.D. Thesis, Nara Institute of Science and Technology (2010)
- 9) Ihara, M., Maeda, S., Ikeda, K., Ishii, S.: Low-Dimensional Feature Representation for Monophonic Music Instrument Identification, to appear.
- 10) Cichocki, A., Zdunek, R., Phan, A., Amari, S.: Non-negative Matrix and Tensor Factorizations, John Wiley & Sons Ltd., (2009)
- 11) Kameoka, H., Ono, N., Kashino, K., Sagayama, S.: Complex NMF: A New Sparse Representation for Acoustic Signals, Proc. ICASSP, 3437/3440 (2009)
- 12) Eggert, J., Wersing, H., Koerner, E.: Transformation-Invariant Representation and NMF, Proc. IJCNN (2004)
- 13) Raczynski, S.A., Ono, N., Sagayama, S.: Multipitch Analysis with Harmonic Nonnegative Matrix Approximation, Proc. ISMIR, 381 (2007)
- 14) Ochiai, K., Nakano, M., Ono, N., Sagayama, S.: Parallel Nonnegative Matrix Factorization of High-Time-Resolution and High-Frequency-Resolution Spectrograms for Multipitch Analysis of Music Signals, Proc. ASJ Spring Conf., 705/723 (2011)



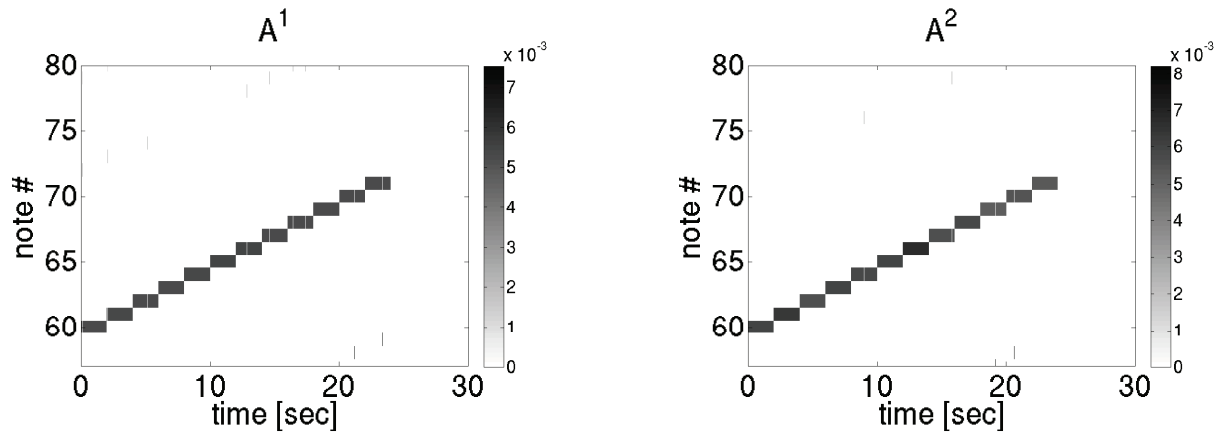
(a) Note activations. Each column shows the note id at a time.

(b) Basis patterns. Each column shows the spectrum of a note.

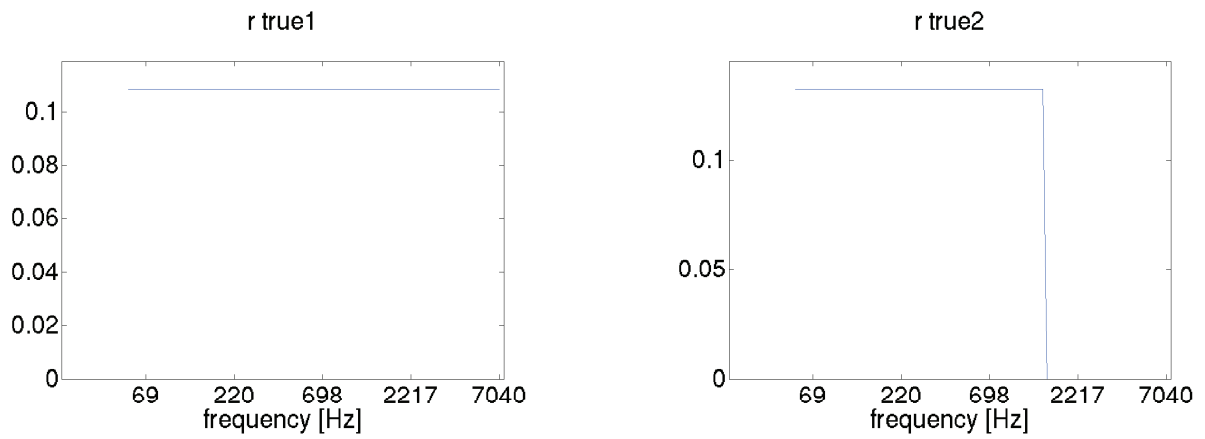


(c) Frequency responses of sound paths for cases 1 and 2.

Fig. 3: True note activations, basis patterns and frequency responses (two cases) of sound paths used in the experiments.



(a) Note activations for cases 1 and 2.



(b) Frequency responses of sound paths for cases 1 and 2.

Fig. 4: Estimated note activations and frequency responses of sound paths for sparseness parameter  $\lambda = 0.1$ .



# 定数項を用いた複素連想記憶

○北原倫理 小林正樹 (山梨大学)

## Complex-valued Associative Memory with Strong Bias Terms

\*M. Kitahara and M. Kobayashi (University of Yamanashi)

**Abstract**— Complex-valued Associative Memory (CAM) is an extension model of Hopfield Associative Memory, and has complex-valued neurons and connection weights. So CAM can treat multi-state patterns, it is often used for operating gray-scale patterns. When CAM takes  $K$  states, it stores  $K-1$  spurious patterns, called rotated patterns, for every learning pattern. The rotated patterns cause decrease in noise robustness. In this paper, we'll propose a model which is CAM with bias terms. By using the bias terms, we can expect to improve noise robustness of CAM. We confirmed the effectiveness by computer simulations.

**Key Words:** Associative memory, Complex-valued neural networks

### 1 はじめに

複素連想記憶 (Complex-valued Associative Memory: CAM) は高度なニューラルネットワークのモデルの1つとして研究されている (Jankowskiら<sup>1)</sup>). CAMは多値パターンを学習することができ, しばしば濃淡画像の処理などに使われる (Aoki<sup>3)</sup>). CAMは学習パターンだけでなく, 回転パターンと呼ばれる偽記憶の代表的なパターンも記憶する (Kuroe<sup>2)</sup>). CAMが $K$ 状態をとるとき, CAMは1つの学習パターンにつき $K-1$ 個の回転パターンを記憶するため雑音耐性の低下を招く. 本論文では, CAMの雑音耐性を向上させるために定数項を用いる. 定数項はCAMが回転パターンを記憶することを防ぐ. 定数項を単純に用いた場合, 定数項の効果は複素ニューロンの離散化により弱められてしまう. そこで, 我々は擬似逆行列 (Aoki<sup>3)</sup>) を用いた学習と定数項を用いて複素ニューロンの離散化を克服した. この擬似逆行列を用いた学習を直交学習と呼ぶ.

定数項を制御するために制御パラメータ $t$ を用いた. コンピュータシミュレーションにより,  $t$ がある値以下の場合 $t$ が大きくなるほど効果が大きくなり,  $t$ がある値以上の場合 $t$ の効果は一定となることが分かった. 更に実験により学習パターン数 $P$ , ニューロン数 $N$ が $P/N \leq 0.1$ を満たす場合, 定数項は雑音耐性を向上させることが分かった.  $P$ が小さく $K$ が大きい場合, 特に定数項の効果は大きい.  $P/N < 0.1$ はアプリケーションにとって十分に大きな値である. 実際, DeguchiとIshii<sup>4,5)</sup>, HattoriとHagiwara<sup>6)</sup>, Osanaら<sup>7)</sup>が提案した連想記憶は $P/N < 0.1$ を満たしている. このことから, 定数項は有用であることが分かる.

### 2 複素連想記憶

複素連想記憶 (Complex-valued Associative Memory: CAM) について簡単に説明する.

まず, 複素ニューロンを定義する. 離散化数が $K$ のとき, 複素ニューロンの状態は複素単位円周上の $K$ 等分点 $s_k (k = 0, 1, \dots, K-1)$ のいずれかを値にとる.

$$s_k = \exp(2k\theta_K\sqrt{-1}) \quad (k = 0, 1, \dots, K-1) \quad (1)$$

$$\theta_K = \pi/K \quad (2)$$

この等分点を離散化点と呼ぶ. 複素ニューロンへの入

力値は, 活性化関数 $f(\cdot)$ によって $\{s_k\}$ の中で最もユークリッド距離が近い値に離散化される.

次に, CAMについて説明する. CAMのニューロン数を $N$ とする.  $i$ 番目のニューロンから $j$ 番目のニューロンへの結合荷重を $w_{ji}(i, j = 1, 2, \dots, N)$ とすると,  $w_{ji} = \overline{w_{ij}}$ が成り立つ. ここで $\overline{w_{ij}}$ は $w_{ij}$ の複素共役を意味する.  $j$ 番目のニューロンの定数項を $\gamma_j$ ,  $i$ 番目のニューロンの状態を $z_i$ とすると,  $j$ 番目のニューロンへの入力 $I_j$ は

$$I_j = \sum_{\substack{i=1 \\ i \neq j}}^N w_{ji}z_i + \gamma_j \quad (3)$$

と定義される.

最後にCAMの学習について説明する. 学習パターンを $\mathbf{x} = (x_1, x_2, \dots, x_N)^T$ とする. ここで学習パターンは $\{s_k\}$ のいずれかを値にとり,  $T$ は転置を意味する. CAMに $\mathbf{x}$ を入力し,  $x_j = f(I_j) (\forall j)$ が成り立つ場合を安定状態と呼ぶ. 学習は, 全ての学習パターンについて安定状態が成り立つように結合荷重と定数項を設定することに相当する.  $\mathbf{x}$ にノイズを乗せたパターン $\mathbf{x}'$ をCAMに入力すると, CAMはノイズを除去して元の $\mathbf{x}$ を想起しようとする. このノイズ除去能力は雑音耐性と呼ばれ, 連想記憶の性能を示す重要な指標の1つとなっている.

### 3 直交学習

定数項の値が全て $0 (\gamma_j = 0 \quad \forall j)$ の場合における直交学習について説明する. 学習パターン数を $P$ ,  $p$ 番目の学習パターンを $\mathbf{x}^{(p)} = (x_1^{(p)}, x_2^{(p)}, \dots, x_N^{(p)})^T (p = 1, 2, \dots, P)$ とすると, 学習パターン行列 $\mathbf{X}$ は $\mathbf{X} = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(P)})$ と表現できる.  $\mathbf{X}$ は $N \times P$ 行列である. このとき,  $N \times N$ 行列 $\mathbf{W}$ を以下のように定義する.

$$\mathbf{W} = \mathbf{X}(\mathbf{X}^*\mathbf{X})^{-1}\mathbf{X}^* \quad (4)$$

\*は共役転置を意味する.  $(\mathbf{X}^*\mathbf{X})^{-1}\mathbf{X}^*$ は $\mathbf{X}$ の擬似逆行列である.  $\mathbf{W}$ は $\mathbf{W}^* = \mathbf{W}$ の関係が成り立っているため,  $\mathbf{W}$ の $(j, i)$ 要素 $w_{ji}$ と $(i, j)$ 要素 $w_{ij}$ は $w_{ji} = \overline{w_{ij}}$ を満たす. また,  $w_{jj}$ は実数値である.  $j \neq i$ のとき $w_{ji}$ を $i$ 番目のニューロンから $j$ 番目のニューロンへの結合

荷重とみなす。ニューロンは自身への結合を持たないので、対角成分  $w_{jj}$  をニューロンの状態の更新に使用することはない。 $\mathbf{W}$  は  $\mathbf{W}\mathbf{X} = \mathbf{X}$  を必ず満たすため、

$$x_j^{(p)} = \sum_{i=1}^N w_{ji}x_i^{(p)} \quad \forall p, j \quad (5)$$

が成り立つ。学習パターン  $\mathbf{x}^{(p)}$  を CAM に入力するとき  $j$  番目のニューロンへの入力  $I_j$  は

$$I_j = \sum_{\substack{i=1 \\ i \neq j}}^N w_{ji}x_i^{(p)} = (1 - w_{jj})x_j^{(p)} \quad \forall j \quad (6)$$

となるから、 $w_{jj} < 1 \quad \forall j$  であるとき全学習パターン ( $\mathbf{x}^{(p)} \quad \forall p$ ) は安定状態を示す。

$s_k \mathbf{x}^{(p)}$  ( $k = 1, 2, \dots, K-1$ ) を  $\mathbf{x}^{(p)}$  の回転パターンと呼ぶ。 $s_k \mathbf{x}^{(p)}$  を CAM に入力すると

$$I_j = \sum_{\substack{i=1 \\ i \neq j}}^N w_{ji}(s_k x_i^{(p)}) \quad (7)$$

$$= s_k \left( \sum_{\substack{i=1 \\ i \neq j}}^N w_{ji}x_i^{(p)} \right) \quad (8)$$

$$= s_k x_j^{(p)} \quad (9)$$

が成り立つから、学習パターン  $\mathbf{x}^{(p)}$  が安定状態を示すとき回転パターン  $s_k \mathbf{x}^{(p)}$  もまた安定状態を示す。この現象を回転不変性と呼ぶ (Zemel ら<sup>11</sup>)。離散化数  $K$  が大きくなるにつれて、2つの問題が深刻化する。一方は回転パターンの増加に伴う偽記憶の増加である。CAM は1つの学習パターンにつき  $K-1$  個の回転パターンを記憶する。他方は学習パターンと回転パターンとの位相差が小さくなる問題である。 $K$  が大きくなるにつれて離散化点  $s_1$  と  $s_{K-1}$  は1に近づくように回転パターン  $s_1 \mathbf{x}^{(p)}$  と  $s_{K-1} \mathbf{x}^{(p)}$  は  $\mathbf{x}^{(p)}$  に近づく。この2つの問題は CAM の雑音耐性の低下の大きな原因となっている。

#### 4 定数項を用いた直交学習

本論文の提案モデルである、定数項を用いた CAM の直交学習について説明する。定数項は回転パターンが安定状態に陥らないように作用する。定数項の位相を設定するために、実数値の制御パラメータ  $t$  を用いる。これに伴い、学習パターン  $\mathbf{x}^{(p)}$  を  $\tilde{\mathbf{x}}^{(p)} = (\mathbf{x}^{(p)T}, t)^T$  に、学習パターン行列  $\mathbf{X}$  を  $(N+1) \times P$  行列の  $\tilde{\mathbf{X}} = (\tilde{\mathbf{x}}^{(1)}, \tilde{\mathbf{x}}^{(2)}, \dots, \tilde{\mathbf{x}}^{(P)})$  に、結合荷重  $\mathbf{W}$  を  $(N+1) \times (N+1)$  行列の  $\tilde{\mathbf{W}}$

$$\tilde{\mathbf{W}} = \tilde{\mathbf{X}}(\tilde{\mathbf{X}}^* \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^* \quad (10)$$

に拡張する。 $\tilde{\mathbf{W}}$  は  $\tilde{\mathbf{W}} = \tilde{\mathbf{W}}^*$  と  $\tilde{\mathbf{W}} \tilde{\mathbf{X}} = \tilde{\mathbf{X}}$  を満たす。 $j \leq N$  のとき、以下の式が成り立つ。

$$\sum_{1 \leq i \leq N} w_{ji}x_i^{(p)} + t w_{j(N+1)} = x_j^{(p)} \quad (11)$$

$\tilde{\mathbf{W}}$  の要素  $w_{ji}$  ( $1 \leq i, j \leq N, i \neq j$ ) を結合荷重とみなす。 $\tilde{\mathbf{W}} = \tilde{\mathbf{W}}^*$  であるから  $w_{ij} = \overline{w_{ji}}$  を満たす。また、 $t w_{j(N+1)}$  を  $j$  番目のニューロンの定数項  $\gamma_j$  とみなす。

$p$  番目の学習パターン  $\mathbf{x}^{(p)}$  を CAM に入力するとき、 $j$  番目のニューロンへの入力  $I_j$  は以下のように表現できる。

$$I_j = \sum_{\substack{1 \leq i \leq N \\ i \neq j}} w_{ji}x_i^{(p)} + \gamma_j \quad (12)$$

$$= (1 - w_{jj})x_j^{(p)} \quad (13)$$

$\tilde{\mathbf{W}} = \tilde{\mathbf{W}}^*$  より  $w_{jj}$  は実数値である。 $w_{jj} < 1$  ( $1 \leq j \leq N$ ) のとき、全ての学習パターン  $\mathbf{x}^{(p)}$  ( $\forall p$ ) が安定状態をとる。 $\mathbf{x}^{(p)}$  が安定状態を示し、その回転パターンの1つ  $s_k \mathbf{x}^{(p)}$  を CAM に入力するとき、 $j$  番目のニューロンへの入力  $I_j$  は以下のように表現できる。

$$I_j = s_k \sum_{\substack{1 \leq i \leq N \\ i \neq j}} w_{ji}x_i^{(p)} + \gamma_j \quad (14)$$

$$= s_k \left( \sum_{\substack{1 \leq i \leq N \\ i \neq j}} w_{ji}x_i^{(p)} + \overline{s_k} \gamma_j \right) \quad (15)$$

$$= s_k \left( \sum_{1 \leq i \leq N} w_{ji}x_i^{(p)} + \gamma_j - w_{jj}x_j^{(p)} - \gamma_j + \overline{s_k} \gamma_j \right) \quad (16)$$

$$= s_k (x_j^{(p)} + (\overline{s_k} - 1)\gamma_j - w_{jj}x_j^{(p)}) \quad (17)$$

$$= s_k x_j^{(p)} (1 - w_{jj} + (\overline{s_k} - 1)\gamma_j \overline{x_j^{(p)}}) \quad (18)$$

いくつかの  $j$  ( $1 \leq j \leq N$ ) において

$$|\arg(1 - w_{jj} + (\overline{s_k} - 1)\gamma_j \overline{x_j^{(p)}})| > \theta_K \quad (19)$$

$$\theta_K = \pi/K \quad (20)$$

が成り立つとき、 $s_k \mathbf{x}^{(p)}$  は安定状態をとらない。ここで、 $\arg(c)$  は複素数  $c$  の位相を意味している。 $\mathbf{x}^{(p)}$  は安定状態を示すから  $1 - w_{jj}$  は正の実数値をとる。定数項  $\gamma_j$  が負の実数値をとり、かつ  $\gamma_j$  の絶対値が  $1 - w_{jj}$  の絶対値に対してはるかに大きい場合、条件 (19) を満たさない。よって、 $\gamma_j$  が大きな値をとるように制御パラメータ  $t$  を設定する必要がある。

制御パラメータ  $t = 0$  のとき  $\tilde{\mathbf{X}}^* \tilde{\mathbf{X}} = \mathbf{X}^* \mathbf{X}$  が成り立つから、 $1 \leq j, i \leq N$  における  $\mathbf{W}$  の要素と  $\tilde{\mathbf{W}}$  の要素は等しくなる。よって、 $t = 0$  のときには提案モデルである定数項を用いた CAM の直交学習は、定数項を用いない従来の CAM の直交学習と同じになる。提案モデルは従来のモデルの拡張であり、制御パラメータ  $t$  は定数項を制御する役割を果たす。 $t = 1$  のとき

#### 5 制御パラメータと定数項

制御パラメータ  $t$  が大きくなるにつれて定数項  $\gamma_j$  が大きくなることが予想されるが、これは自明ではない。制御パラメータと定数項の関係をコンピュータシミュ

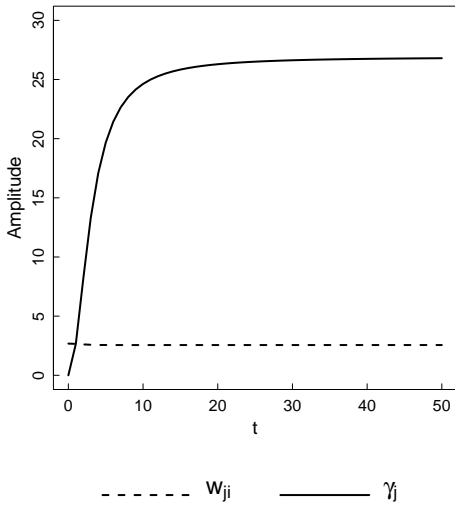


Fig. 1: Results of average amplitudes of  $w_{ji}$  and  $\gamma_j$ .

レーションによって調べた。パラメータはニューロン数  $N = 100$ , 離散化数  $K = 8$ , 学習パターン数  $P = 10$ , 制御パラメータ  $t = 0 \sim 50$  である。学習パターンの各要素は  $\{s_k\}$  のいずれかを値にとり、ランダムに生成した。試行回数は  $t$  毎に 1000 回とした。

Fig. 1 は  $i$  番目のニューロンから  $j$  番目のニューロンへの結合荷重  $w_{ji}$  の振幅の平均値と  $t$  との関係,  $j$  番目のニューロンの定数項  $\gamma_j$  の振幅の平均値と  $t$  との関係を表している。X 軸は制御パラメータ  $t$ , Y 軸は結合荷重  $w_{ji}$  と定数項  $\gamma_j$  の振幅の平均値である。シミュレーション結果から結合荷重と制御パラメータは関係が無いことが分かる。一方定数項と制御パラメータの関係は, 制御パラメータがある値に達するまでは定数項の値は増加, 制御パラメータがある値に達した後は定数項はほぼ一定の値を示した。定数項  $\gamma_j$  の振幅の平均値と結合荷重  $w_{ji}$  の振幅の平均値との割合  $r$  を考える。本シミュレーションは  $t$  が小さいときには  $r$  は急激に増加し,  $t$  がほぼ一定以上に大きくなると  $r$  は一定の値をとることを示している。  $r$  は十分に大きいことが求められる。

## 6 雑音耐性

CAM の雑音耐性における制御パラメータ  $t$  の効果をコンピュータシミュレーションによって調べた。パラメータはニューロン数  $N = 100$ , 離散化数  $K = 10, 20, 30$ , 学習パターン数  $P = 10, 20, 30$ , 制御パラメータ  $t = 0, 1, 10$ , ノイズ数  $L = 1 \sim 80$  である。組み合わせ  $(K, P, t)$  毎に学習パターンを生成し CAM に学習させる。学習パターンの各要素は  $\{s_k\}$  のいずれかをランダムにとる。これを 100 回繰り返す。学習パターンを 1 回生成する毎に以下を繰り返す。

- $L = 1 \sim 80$  毎に以下をを 100 回繰り返す。
  1.  $P$  個の学習パターンの中からランダムに 1 の学習パターン  $D$  を選択する。
  2.  $D$  の要素の  $L$  箇所をランダムに選び,  $\{s_k\}$  のいずれかの値に置き換えてノイズパターンを生成する。

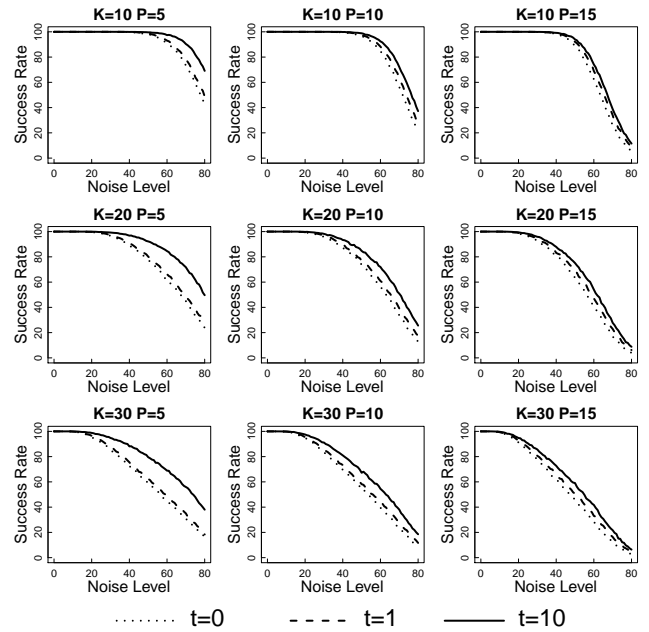


Fig. 2: Results of noise robustness.

3. ノイズパターンを CAM に入力する。CAM がノイズをのせる前のパターンを想起できれば成功とする。

$t = 0$  のときは定数項が無い。  $t = 1$  のときは 1 個の 1 を常に出力するニューロンが存在する。 Fig. 2 では,  $t < 10$  のとき定数項の振幅は急激に増加し,  $t > 10$  ではほぼ一定の値をとっている。よって,  $t \geq 10$  のとき定数項と  $t$  は関係が無い。

Fig. 2 は雑音耐性の実験結果を表す。X 軸はノイズ数  $L$ , Y 軸はノイズ除去の成功率である。実験から CAM の雑音耐性について以下の結果が得られた。

1. 全ての条件において,  $t$  が大きい方が雑音耐性が強い。
2.  $t = 1$  において, わずかな雑音耐性の向上が見られる。
3.  $t = 1, 10$  において, 両者の雑音耐性の向上の差は,  $P$  が小さくなるほど,  $K$  が大きくなるほど, 大きくなる。

我々の実験において, 特に  $P/N \leq 0.1$  の場合に定数項は強い効果を発揮した。この定数項の効果は Table 1 にあげられる  $P/N \leq 0.1$  を満たすアプリケーションで有効である。

## 7 むすび

本論文では, 定数項を用いた CAM の直交学習を提案した。定数項を用いない場合, 離散化数  $K$  の CAM は回転不変性により 1 つの学習パターンにつき  $K - 1$  個の回転パターンを記憶するため, 雑音耐性の低下につながる。これに対して, 提案モデルは定数項が回転パターンの記憶を防ぐため, 雑音耐性の向上が期待できる。ただし, 定数項を用いた場合でも, 結合荷重に対して定数項の値があまりに小さい場合には定数項の効果は期待できない。この定数項の値を制御するため

Applications	$P/N$
Deguchi and Ishii <sup>4)</sup>	0.04
Deguchi and Ishii <sup>5)</sup>	0.075
Hattori and Hagiwara <sup>6)</sup>	0.04
Osana et. al. <sup>7)</sup>	0.011
Yano and Osana <sup>8)</sup>	0.009
Nakada and Osana <sup>9)</sup>	0.008
Kitahara et. al. <sup>10)</sup>	0.008

Table 1:  $P/N$  of some applications

に、我々は制御パラメータ  $t$  を定義した。実験により、定数項の振幅の平均値と結合荷重の振幅の平均値との割合  $r$  について以下のことが分かった。

1.  $t$  が小さいとき、 $t$  が大きくなるにつれて  $r$  も大きくなる。
2.  $t$  が大きいとき、 $r$  はほぼ一定の値となるため  $r$  との関係がない。

更に、雑音耐性の実験により提案モデルが雑音耐性を向上させることが分かった。また、学習パターン数  $P$  が小さいほど、離散化数  $K$  が大きいほど雑音耐性の向上することを示した。

## 参考文献

- 1) S. Jankowski, A. Lozowski and J. M. Zurada: Complex-valued multistate neural associative memory, *IEEE Trans. Neural Networks*, vol.7, 1491/1496 (1996)
- 2) Y. Kuroe: A Model of complex-valued associative memories and its dynamics, *Complex-Valued Neural Networks, Theories and Applications*, Ed. A.Hirose, World Scientific, 57/79 (2003)
- 3) H. Aoki, M. R. Azimi-Sadjadi and Y. Kosugi: Image association using a complex-valued associative memory model, *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol.E83-A, 1824/1832 (2000)
- 4) T. Deguchi and N. Ishii: Control of associative dynamics by matching features in chaotic neural network, *Systems and Computers in Japan*, vol.27, no.5, 47/54 (1996)
- 5) T. Deguchi and N. Ishii: Search for cyclic patterns by a chaotic neural network, *Systems and Computers in Japan*, vol.29, no.11, 1/8 (1998)
- 6) M. Hattori and M. Hagiwara: Multimodule associative memory for many-to-many associations, *Neurocomputing*, vol.19, 99/119 (1998)
- 7) Y. Osana, M. Hattori and M. Hagiwara: Chaotic bidirectional associative memory, *Proceedings of IEEE International Conference on Neural Networks*, 816/821 (1996)
- 8) Y. Yano and Y. Osana: Chaotic complex-valued bidirectional associative memory, *Proceedings of IEEE and INNS International Joint Conference on Neural Networks*, 433/438 (2009)
- 9) M. Nakada and Y. Osana: Chaotic complex-valued associative memory, *2007 International Symposium on Nonlinear Theory and its Applications NOLTA '07*, 493/496 (2007)
- 10) M. Kitahara, M. Kobayashi and M. Hattori: Chaotic rotor associative memory, *2009 International Symposium on Nonlinear Theory and its Applications NOLTA '09*, 399/402 (2009)
- 11) R. S. Zemel, C. K. I. Williams and M. C. Mozer: Lending direction to neural networks, *Neural Networks*, 503/512 (1995)

# 時間的に変化する不応性のパラメータを有する カオス複素多方向連想メモリ

吉田明生 ○長名優子 (東京工科大学)

## Chaotic Complex-valued Multidirectional Associative Memory with Variable Scaling Factor

A. Akio and \*Y. Osana (Tokyo University of Technology)

**Abstract**— In this paper, we propose a Chaotic Complex-valued Multidirectional Associative Memory with variable scale factor which can realize one-to-many associations of  $M$ -tuple multi-valued patterns. The proposed model is based on the Multidirectional Associative Memory, and is composed of complex-valued neurons and chaotic complex-valued neurons. In the proposed model, associations of multi-valued patterns are realized by using complex-valued neurons, and one-to-many associations are realized by using chaotic complex-valued neurons. Moreover, in the proposed model, the appropriate parameters of chaotic complex-valued neurons can be determined easily than in the original Chaotic Complex-valued Multidirectional Associative Memory. We carried out a series of computer experiments and confirmed that the proposed model has superior one-to-many association ability than that of the conventional model.

**Key Words:** One-to-many association, Chaotic complex-valued multidirectional associative memory, Scaling factor of refractoriness

### 1 はじめに

近年、生物の脳や神経系に見られるような柔軟な情報処理を行う手法として、ニューラルネットワークの研究が盛んに行われている。そのなかでも人間の連想記憶機能を模倣する連想記憶に関する研究が盛んに行われており、多くの連想メモリが提案されている。しかし、これらの連想メモリは2値パターンを扱うことを前提としており、多値パターンを扱うことはできない。

また、多値パターンを扱うことができるモデルとして複素ニューロンモデルが提案されている<sup>1)</sup>。複素ニューロンモデルはニューロンの入出力や内部状態を複素数値にしたモデルで、複素数値を多値に割り当てることで、多値パターンを扱うことができる<sup>2)-5)</sup>。複素ニューロンモデルから構成される複素連想メモリ<sup>2)</sup>では多値パターンの連想が扱えるだけでなく、耐ノイズ特性が向上することが報告されている。

一方で柔軟な情報処理を行う手法の一つとしてカオスが注目されている。カオスは決定論的な時間発展をする非線形なシステムに生じる長期予測不可能な現象で、生体の脳や神経系でも観測されており、脳における記憶や学習に重要な役割を果たしていると考えられている<sup>6)</sup>。ニューラルネットワークの研究においても、実際の神経細胞においてみられる時空間加算、不応性、連続値出力などを考慮することでカオスを導入したカオスニューロンモデルが提案されている<sup>7)</sup>。また、カオスニューロンモデルから構成される連想メモリでは、記憶したパターンを動的に想起できることが知られている<sup>7)8)</sup>。カオス双方向連想メモリ<sup>9)</sup>では、双方向連想メモリ<sup>10)</sup>の一部にカオスニューロンモデルを導入することで、1対多の連想を実現している。

複素ニューロンモデル<sup>11)</sup>とカオスニューロンモデル<sup>7)</sup>を組み合わせたカオス複素ニューロンモデル<sup>12)</sup>も提案されている。カオス複素ニューロンモデルから構成されるカオス複素連想メモリ<sup>12)</sup>では、多値パター

ンの動的な想起が実現されている。また、カオス複素双方向メモリ<sup>13)14)</sup>では、双方向連想メモリ<sup>10)</sup>に複素ニューロンモデルとカオス複素ニューロンモデルを導入することで、多値パターンの1対多の連想を実現している。

さらに、カオス複素多方向連想メモリ<sup>15)</sup>では、多方向連想メモリ<sup>16)</sup>に複素ニューロンモデルとカオス複素ニューロンモデルを導入することで3つ組データの多値パターンの1対多の連想を実現している。文献<sup>17)</sup>ではカオス複素多方向連想メモリの1対多の想起能力について調べ、カオス複素ニューロンのパラメータ、特に不応性のスケーリングファクタによって1対多の想起能力が大きく変わってくることを示されている。

本報告では、カオス複素多方向連想メモリの不応性のパラメータを時間的に変化させることで1対多の想起能力を向上する手法を提案する。

### 2 カオス複素ニューロンモデル

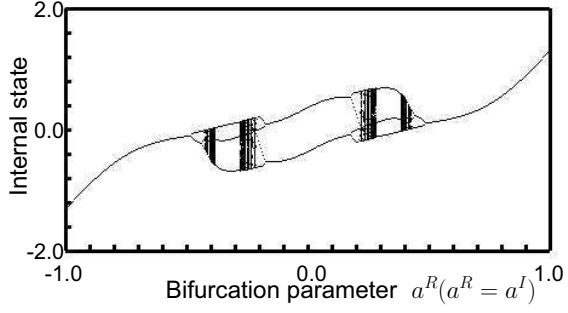
ここでは、提案モデルで用いるカオス複素ニューロンモデル<sup>12)</sup>について説明する。カオス複素ニューロンモデルは、カオスニューロンモデル<sup>7)</sup>を内部状態や出力として複素数値を扱えるように拡張したモデルであり、カオス複素ニューロンから構成されるカオス複素連想メモリでは、記憶した複数の多値パターンを動的に想起できることが知られている<sup>12)</sup>。

カオス複素ニューロンのダイナミクスは以下のように表される。

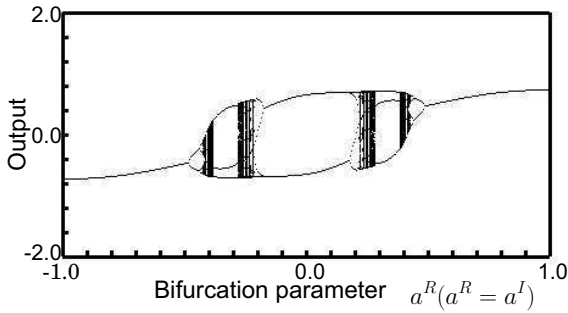
$$x(t+1) = f \left( A(t) - \alpha \sum_{d=0}^t k^d x(t-d) - \theta \right) \quad (1)$$

(ただし、 $A(t), x(t), \theta \in \mathbb{C}$   $k, \alpha \in \mathbb{R}$ )

ここで、 $x(t)$ は時刻 $t$ におけるニューロンの出力、 $A(t)$ は時刻 $t$ における外部入力、 $\alpha$ は不応性に関するスケーリングファクタ、 $k$ は時間減衰定数( $0 \leq k < 1$ )、 $\theta$ は



(a) Internal State (real part)



(b) Output (real part)

Fig. 1: Bifurcation diagram of chaotic complex-valued neuron model.

ニューロンの閾値である。また、 $f(\cdot)$  は出力関数であり、以下のように与えられる。

$$f(u) = \frac{\eta u}{\eta - 1.0 + |u|} \quad (\text{ただし、}\eta \in \mathbb{R}) \quad (2)$$

ここで、 $\eta$  は  $\eta > 1$  を満たす定数である。

式 (1) は、外部入力  $A(t)$  を一定 ( $A(t) = A$ ) であるとする以下のように変形することができる。

$$\begin{aligned} x(t+1) &= f(u(t+1)) \\ &= f(ku(t) - \alpha f(u(t)) + (A - \theta)(1 - k)) \\ &= f(ku(t) - \alpha f(u(t)) + a) \end{aligned} \quad (3)$$

ここで、 $u(t)$  は時刻  $t$  におけるニューロンの内部状態、 $a = (A - \theta)(1 - k)$  は分岐パラメータである。 $\alpha = 1.0$ 、 $k = 0.8$  として分岐パラメータを変化させたときのカオス複素ニューロンの内部状態と出力の実部の分岐図を図 1 に示す。なお、図 1 では出力関数のパラメータ  $\eta$  は 1.1 としている。また、分岐パラメータ  $a$  は複素数値であるが、図 1 では実部と虚部に同じ値 ( $a^R = a^I$ ) を用いている。

### 3 時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリ

ここでは、提案する時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリ (Chaotic

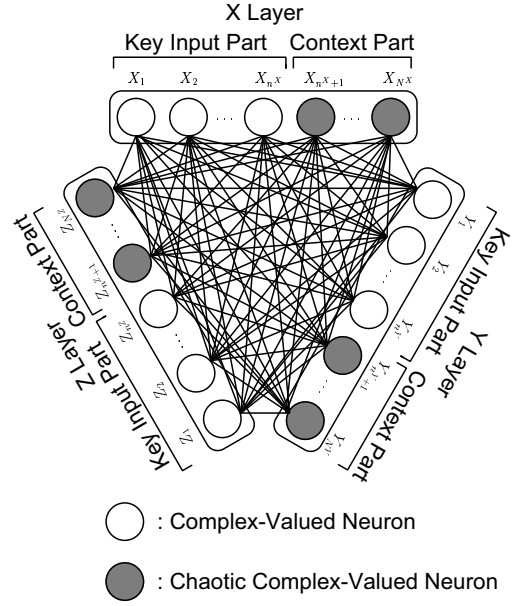


Fig. 2: Structure of proposed CCMAM-VS.

Complex-valued Multidirectional Associative Memory with Variable Scaling Factor : CCMAM-VS) について説明する。このモデルはカオス複素多方向連想メモリ<sup>15)</sup>に基づいたモデルである。

#### 3.1 構造

時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリは、多方向連想メモリ<sup>16)</sup>と同様に 3 つ以上の層から構成されており、図 2 の例では X 層、Y 層、Z 層の 3 つの層から構成されている。各層は、図 2 に示すようにキー入力部 (Key Input Part) と文脈情報部 (Context Part) の 2 つに分けられている。キー入力部は複素ニューロンモデル<sup>11)</sup>から、文脈情報部はカオス複素ニューロンモデル<sup>12)</sup>から構成されている。

#### 3.2 学習過程

関連学習では 1 対多の関係にある学習組を記憶することはできない。カオス双方向連想メモリ<sup>9)</sup>では、1 対多の関係にある学習組に文脈情報を付加することで擬似的に 1 対 1 の関係に変換し、1 対多の関係にある学習組の記憶を実現している。カオス複素多方向連想メモリ<sup>15)</sup>や時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリでも、同様の方法を用いて学習を行う。

$$\{(X_1, Y_1, Z_1), (X_1, Y_2, Z_2), (X_2, Y_3, Z_3)\} \quad (4)$$

のような 1 対多の関係にある学習組を 3 つの層から構成される時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリに学習させる場合、各学習組に異なる文脈情報パターンを付加することで以下のような 1 対 1 の学習組を擬似的に得ることができる。

$$\begin{aligned} &\{(X_1.C_{X1}, Y_1.C_{Y1}, Z_1.C_{Z1}), \\ &(X_1.C_{X2}, Y_2.C_{Y2}, Z_2.C_{Z2}), \end{aligned}$$



$$\{(\mathbf{X}_2, \mathbf{C}_{X3}, \mathbf{Y}_3, \mathbf{C}_{Y3}, \mathbf{Z}_3, \mathbf{C}_{Z3})\} \quad (5)$$

3つの層から構成される時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリでは、文脈情報を付加したパターンを用いて各層の間の重みを直交学習により以下のように決定する。

$$\mathbf{w}^{YX} = \mathbf{Y}(\mathbf{X}^* \mathbf{X})^{-1} \mathbf{X}^* \quad (6)$$

$$\mathbf{w}^{XY} = \mathbf{X}(\mathbf{Y}^* \mathbf{Y})^{-1} \mathbf{Y}^* \quad (7)$$

$$\mathbf{w}^{ZX} = \mathbf{Z}(\mathbf{X}^* \mathbf{X})^{-1} \mathbf{X}^* \quad (8)$$

$$\mathbf{w}^{XZ} = \mathbf{X}(\mathbf{Z}^* \mathbf{Z})^{-1} \mathbf{Z}^* \quad (9)$$

$$\mathbf{w}^{ZY} = \mathbf{Z}(\mathbf{Y}^* \mathbf{Y})^{-1} \mathbf{Y}^* \quad (10)$$

$$\mathbf{w}^{YZ} = \mathbf{Y}(\mathbf{Z}^* \mathbf{Z})^{-1} \mathbf{Z}^* \quad (11)$$

ここで、\*は共役転置、 $-1$ は逆行列を表す。また、 $\mathbf{X}$ 、 $\mathbf{Y}$ 、 $\mathbf{Z}$ は学習パターン行列であり、それぞれ

$$\mathbf{X} = \{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p)}, \dots, \mathbf{X}^{(P)}\} \quad (12)$$

$$\mathbf{Y} = \{\mathbf{Y}^{(1)}, \dots, \mathbf{Y}^{(p)}, \dots, \mathbf{Y}^{(P)}\} \quad (13)$$

$$\mathbf{Z} = \{\mathbf{Z}^{(1)}, \dots, \mathbf{Z}^{(p)}, \dots, \mathbf{Z}^{(P)}\} \quad (14)$$

で与えられる。ここで、 $\mathbf{X}^{(p)}$ 、 $\mathbf{Y}^{(p)}$ 、 $\mathbf{Z}^{(p)}$ はX層、Y層、Z層の $p$ 番目の学習パターン、 $P$ は学習パターンの数を表す。

なお、 $M$ 個の層から構成される時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリでは、 $x$ 層から $y$ 層への重み、 $y$ 層から $x$ 層への重みはそれぞれ

$$\mathbf{w}^{yx} = \mathbf{X}_x (\mathbf{X}_y^* \mathbf{X}_y)^{-1} \mathbf{X}_y^* \quad (15)$$

$$\mathbf{w}^{xy} = \mathbf{X}_y (\mathbf{X}_x^* \mathbf{X}_x)^{-1} \mathbf{X}_x^* \quad (16)$$

で与えられる。ここで、 $\mathbf{X}_x$ 、 $\mathbf{X}_y$ はそれぞれ $x$ 層、 $y$ 層に記憶させる学習パターン行列であり

$$\mathbf{X}_x = \{\mathbf{X}_x^{(1)}, \dots, \mathbf{X}_x^{(p)}, \dots, \mathbf{X}_x^{(P)}\} \quad (17)$$

$$\mathbf{X}_y = \{\mathbf{X}_y^{(1)}, \dots, \mathbf{X}_y^{(p)}, \dots, \mathbf{X}_y^{(P)}\} \quad (18)$$

で与えられる。

### 3.3 想起過程

時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリでは、文脈情報を付加した学習組を用いて学習を行うが、想起時には文脈情報が未知であると考えるのが自然であるため、キー入力のみからの想起を行う。

式(5)の学習組では、例えば $\mathbf{X}_1$ が時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリに与えられる入力となる。時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリにおいて、 $\mathbf{X}_1$ が入力として与えられると、文脈情報に相当する部分のカオス複素ニューロンの状態がカオスによって遷移することにより、

$$\begin{aligned} (\mathbf{X}_1, \mathbf{0}, ?, ?) &\rightarrow \dots \rightarrow (\mathbf{X}_1, \mathbf{C}_{X1}, \mathbf{Y}_1, \mathbf{Z}_1) \rightarrow \dots \\ &\rightarrow (\mathbf{X}_1, \mathbf{C}_{X2}, \mathbf{Y}_2, \mathbf{Z}_2) \rightarrow \dots \end{aligned} \quad (19)$$

のように1対多の想起が行われることが期待できる。

3つの層から構成される時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリにおいてX層に入力が与えられた場合、想起は以下のように行われる。

#### Step 1: X層への入力

X層にパターンを入力する。

#### Step 2: X層からY層、Z層への伝搬

X層にパターンが入力された場合、Y層とZ層のキー入力部へと情報が伝搬される。Y層のキー入力部のニューロン $k$ の出力 $x_k^Y(t)$ は、

$$x_k^Y(t) = f \left( \sum_{j=1}^{N^X} w_{kj}^{YX} x_j^X(t) \right) \quad (20)$$

で与えられる。ここで、 $N^X$ はX層のニューロン数、 $w_{kj}^{YX}$ はX層のニューロン $j$ からY層のニューロン $k$ への重み、 $x_j^X(t)$ は時刻 $t$ におけるX層のニューロン $j$ の出力を表す。また、 $f(\cdot)$ は式(2)で与えられる出力関数である。

Z層のキー入力部のニューロン $l$ の出力 $x_l^Z(t)$ は、

$$x_l^Z(t) = f \left( \sum_{j=1}^{N^X} w_{lj}^{ZX} x_j^X(t) \right) \quad (21)$$

で与えられる。ここで、 $w_{lj}^{ZX}$ はX層のニューロン $j$ からZ層のニューロン $l$ への重みを表す。

#### Step 3: Y層、Z層からX層への伝搬

X層のキー入力部のニューロン $j$ の出力 $x_j^X(t+1)$ は、

$$x_j^X(t+1) = f \left( \sum_{k=1}^{n^Y} w_{jk}^{XY} x_k^Y(t) + \sum_{l=1}^{n^Z} w_{jl}^{XZ} x_l^Z(t) + v A_j \right) \quad (22)$$

で与えられる。ここで、 $n^Y$ はY層のキー入力部のニューロン数、 $n^Z$ はZ層のキー入力部のニューロン数、 $w_{jk}^{XY}$ はY層のニューロン $k$ からX層のニューロン $j$ への重み、 $w_{jl}^{XZ}$ はZ層のニューロン $l$ からX層のニューロン $j$ への重み、 $v$ は外部入力に対する重み、 $A_j$ はX層のニューロン $j$ への外部入力の大きさを表す。なお、外部入力として与えられるパターンについては3.4で詳しく述べる。

X層の文脈情報部のニューロン $j$ の出力 $x_j^X(t+1)$ は、

$$\begin{aligned} x_j^X(t+1) &= f \left( \sum_{k=1}^{n^Y} w_{jk}^{XY} \sum_{d=0}^t k_m^d x_k^Y(t-d) \right. \\ &\quad \left. + \sum_{l=1}^{n^Z} w_{jl}^{XZ} \sum_{d=0}^t k_m^d x_l^Z(t-d) \right. \\ &\quad \left. - \alpha(t) \sum_{d=0}^t k_r^d x_j^X(t-d) \right) \end{aligned} \quad (23)$$

で与えられる。ここで、 $k_m, k_r$  は時間減衰定数である。式 (23) において  $\alpha(t)$  は時刻  $t$  における不応性のスケールリングファクタであるが、このモデルではこの値を時間的に変化させることで 1 対多の想起能力を向上させている。なお、 $\alpha(t)$  の変化のさせ方としては様々な方法が考えられるが本研究では

$$\alpha(t) = a + b \sin(c \cdot t) \quad (24)$$

としている。ここで、 $a, b, c$  は係数である。

#### Step 4 : 繰り返し

Step 2, 3 を繰り返す。

$M$  個の層から構成される時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリにおいて  $x$  層に入力が与えられた場合、Step 2 において  $y$  層 ( $y \neq x$ ) のキー入力部のニューロン  $k$  の出力  $x_k^y(t)$  は

$$x_k^y(t) = f \left( \sum_{j=1}^{N^x} w_{kj}^{yx} x_j^x(t) \right) \quad (25)$$

で与えられる。ここで、 $N^x$  は  $x$  層のニューロン数、 $w_{kj}^{yx}$  は  $x$  層のニューロン数  $j$  から  $y$  層のニューロン  $k$  への重み、 $x_j^x(t)$  は時刻  $t$  における  $x$  層のニューロン  $j$  の出力を表す。また、Step 3 において、 $x$  層のキー入力部のニューロン  $j$  の出力  $x_j^x(t+1)$  は

$$x_j^x(t+1) = f \left( \sum_{y \neq x}^M \left( \sum_{k=1}^{n^y} w_{jk}^{xy} x_k^y(t) \right) + v A_j \right) \quad (26)$$

で与えられる。ここで、 $n^y$  は  $y$  層のキー入力部のニューロン数、 $w_{jk}^{xy}$  は  $y$  層のニューロン  $k$  から  $x$  層のニューロン  $j$  への重みを表す。また、 $x$  層の文脈情報部のニューロン  $j$  の出力  $x_j^x(t+1)$  は

$$x_j^x(t+1) = f \left( \sum_{y \neq x}^M \left( \sum_{k=1}^{n^y} w_{jk}^{xy} \sum_{d=0}^t k_m^d x_k^y(t-d) \right) - \alpha(t) \sum_{d=0}^t k_r^d x_j^x(t-d) \right) \quad (27)$$

ここで、 $x_j^x(t+1)$  は時刻  $t+1$  における  $x$  層のニューロン  $j$  の出力、 $M$  は層の数、 $n^y$  は  $y$  層のキー入力部のニューロン数、 $w_{jk}^{xy}$  は  $y$  層のニューロン  $k$  から  $x$  層のニューロン  $j$  への重み、 $k_m, k_r$  は時間減衰定数、 $\alpha(t)$  は時刻  $t$  における不応性のスケールリングファクタである。また、 $f(\cdot)$  は式 (2) で与えられる出力関数である。

### 3.4 外部入力

時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリでは、パターンが入力された層の文脈情報部の状態がカオスによって変化するとその影響は他の層へと伝搬し、さらにはパターンが入力された層の全体に及ぶ。そのため、記憶パターンの数が多くなると想起の過程においてキー入力部のパターンが他のキーパターンに遷移してしまう可能性が高くなる。カオス複素双方向連想メモリ<sup>13)14)</sup>では、他の

キーパターンへの遷移を防ぐ方法として、パターンが入力された層のキー入力部のニューロンに外部入力  $A_j$  を与え続けるという方法を用いている。時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリでも、同様にパターンが入力された層に対して外部入力を与え続ける方法を用いる。

$x$  層に入力が与えられた場合、初期入力にノイズが含まれていないことが明らかであれば、外部入力  $A_j = x_j^x(0)$  としても構わない。しかしながら、実際には入力パターンにノイズが含まれている可能性を考慮する必要があるため、ネットワークの状態が最初に安定したとき、つまり

$$t_{in} = \min \left\{ t \left| \sum_{j=1}^{n^x} (\hat{x}_j^x(t) - \hat{x}_j^x(t-1)) = 0 \right. \right\} \quad (28)$$

で与えられる時刻  $t_{in}$  における  $x$  層のキー入力部の量子化した出力の値  $\hat{x}_j^x(t_{in})$  を正しいキー入力のパターンであるとみなして、外部入力として用いる。ここで、 $n^x$  は  $x$  層のキー入力部のニューロン数を表す。また、 $\hat{x}_j^x(t)$  は時刻  $t$  における  $x$  層のニューロン  $j$  の量子化した出力であり、学習パターンが  $S$  値パターンのとき、

$$\hat{x}_j^x(t) = \omega^{\hat{s}} \quad (29)$$

で与えられる。ただし、

$$\hat{s} = \underset{s}{\operatorname{argmin}} (\omega^s - x_j^x)^* (\omega^s - x_j^x) \quad (30)$$

(ただし、 $s = 1, 2, \dots, S-1$ )

である。なお、 $\omega$  は以下のように表される。

$$\omega = \exp(i2\pi/S) \quad (31)$$

ここで、 $i$  は虚数単位を表す。

## 4 計算機実験

ここでは、提案する時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリの動作を確認し、有効性を示すために計算機実験を行った。

### 4.1 1対多の想起結果

ここでは、4層の時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリにおいて多値パターンの 1 対多の想起が行えることを確認する実験を行った。

図3のような 1 対多の関係にある 4 値パターンの学習組を 4 層のキー入力部のニューロン数が 400 (=20×20)、文脈情報部のニューロン数が 100 のネットワークに学習させ、実験を行った。図4に提案モデルの想起結果(方向余弦)を示す。 $t=0$  において *pattern A* が第1層のキー入力部に与えられると、*Pattern Set 1, 2* が想起されている。 $t=30$  において *pattern E* が第2層のキー入力部に与えられると、*Pattern Set 2, 3* が、 $t=60$  において *pattern I* が第3層のキー入力部に与えられると、*Pattern Set 3, 4* が想起されている。 $t=80$  において *pattern J* が第4層のキー入力部に与えられると、*Pattern 1, 4* が想起されている。

以上より、提案モデルにおいて 4 値パターンの 1 対多の想起が行えることが分かる。様々な条件のもとで実験を行い、提案モデルにおいて多値パターンの 1 対多の想起が行えることを確認した。



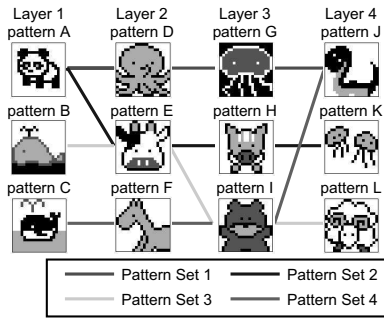
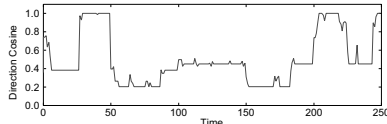
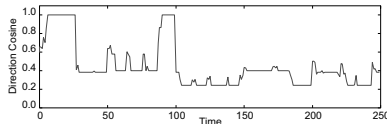


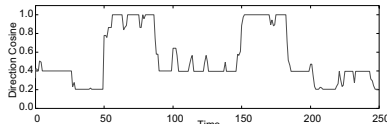
Fig. 3: 学習パターンの例 ( $M = 4, S = 4$ ).



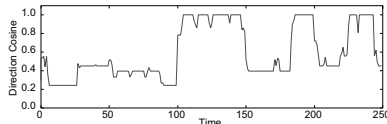
(a) Pattern Set 1



(b) Pattern Set 2



(c) Pattern Set 3



(d) Pattern Set 4

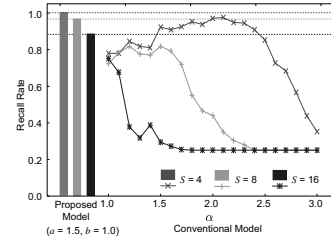
Fig. 4: 想起結果 (方向余弦).

#### 4.2 1対多の想起能力の比較

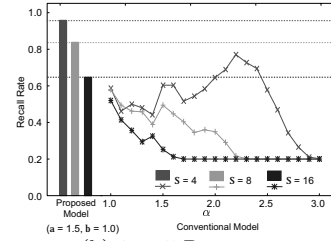
ここでは、提案モデルと従来のカオス複素多方向連想メモリ<sup>15)</sup>の1対多の想起能力を比較する実験を4, 8, 16値のランダムパターンを用いて行った。この実験では、1対 $N$  ( $N=4\sim 6$ )の関係にある $N$ 個のパターンを学習させたネットワークにおいて、共通項を初期入力として与えた。図5~8はそれぞれのモデルにおける想起率の100回の試行の平均を表している。これらの結果より、提案モデルの想起率は従来のモデルよりも高いことが分かる。

#### 4.3 提案モデルの1対多の想起能力

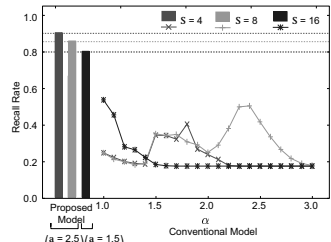
ここでは、提案モデルの1対多の想起能力を調べる実験を4, 8, 16値のランダムパターンを用いて行った。この実験では、1対 $N$  ( $N=4\sim 8$ )の関係にある $N$ 個のパターンを学習させたネットワークに共通項を初期入力として与えた。図9~12は想起率の100回の試行の平均を表している。図9と10より、3層と4層のネットワークでは、1対4のパターンに対しては $a=1.5, b=1.0$ のときに1対多の想起能力が高くなっていることが分



(a) 1-to-4 Patterns

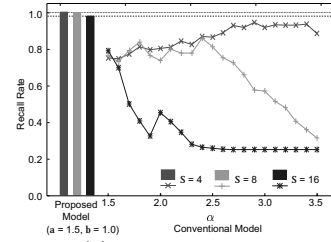


(b) 1-to-5 Patterns

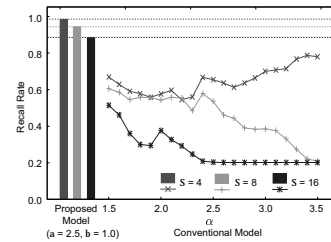


(c) 1-to-6 Patterns

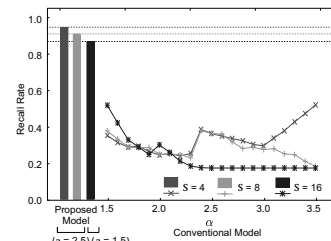
Fig. 5: One-to-Many Association Ability (3-layered Network).



(a) 1-to-4 Patterns

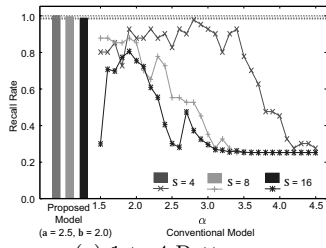


(b) 1-to-5 Patterns

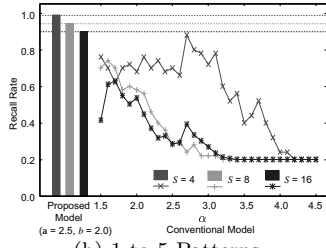


(c) 1-to-6 Patterns

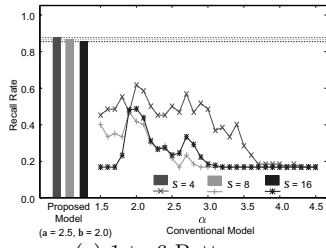
Fig. 6: One-to-Many Association Ability (4-layered Network).



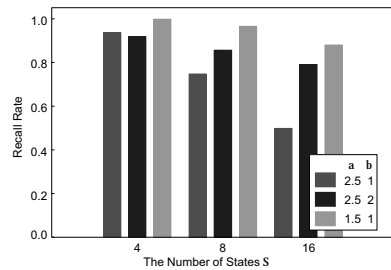
(a) 1-to-4 Patterns



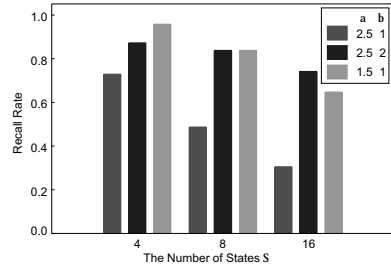
(b) 1-to-5 Patterns



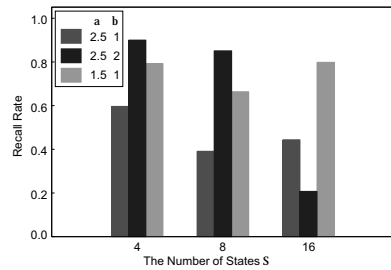
(c) 1-to-6 Patterns



(a) 1-to-4 Patterns

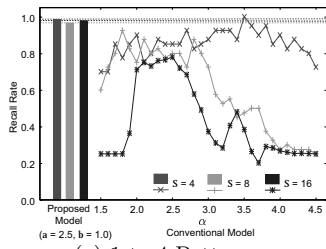


(b) 1-to-5 Patterns

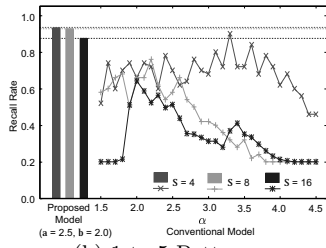


(c) 1-to-6 Patterns

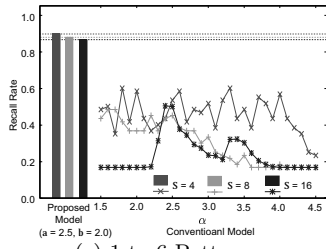
Fig. 7: One-to-Many Association Ability (5-layered Network).



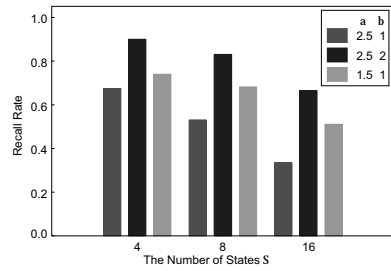
(a) 1-to-4 Patterns



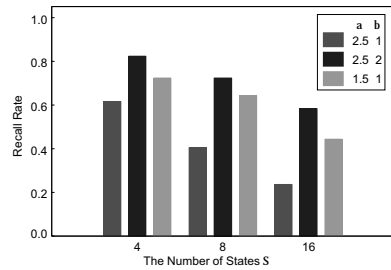
(b) 1-to-5 Patterns



(c) 1-to-6 Patterns



(d) 1-to-7 Patterns



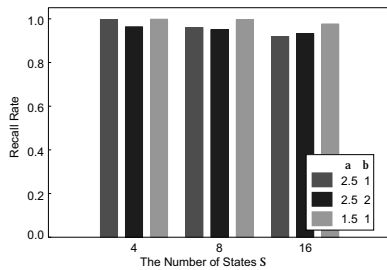
(e) 1-to-8 Patterns

Fig. 8: One-to-Many Association Ability (6-layered Network).

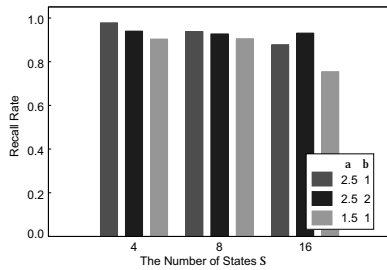
Fig. 9: One-to-Many Association Ability (3-layered Network) (2).

かる。また、1対 $N$ の $N$ の値が大きい場合には、 $a$ と $b$ のパラメータが大きい方が、高い想起率が得られている。

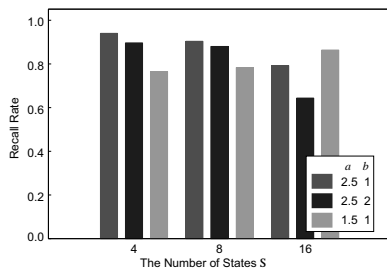
また、図 11 より、5層のネットワークでは1対多の想起能力は $a=2.5, b=2.0$ のときに高くなっていることが分かる。また、図 12 より6層のネットワークでは $a=2.5, b=1.0$ のときに高い想起率が得られていることが分かる。



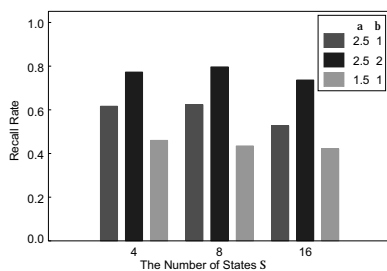
(a) 1-to-4 Patterns



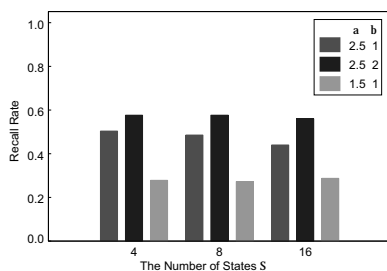
(b) 1-to-5 Patterns



(c) 1-to-6 Patterns



(d) 1-to-7 Patterns



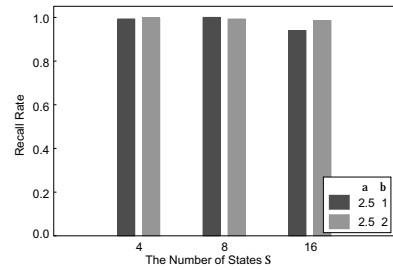
(e) 1-to-8 Patterns

Fig. 10: One-to-Many Association Ability (4-layered Network) (2).

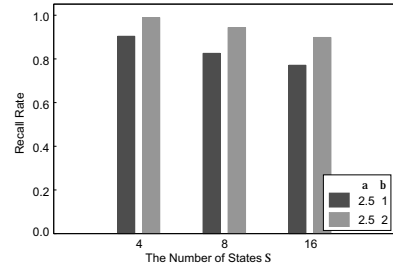
## 5 おわりに

本論文では、時間的に変化する不応性のパラメータを有するカオス複素多方向連想メモリを提案した。提案モデルの動作を確認し、有効性を示すために計算機実験を行い、以下のような性質があることを確認した。

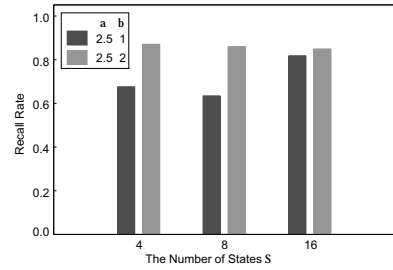
- (1) 多値パターンの1対多の学習・想起が行える。
- (2) 従来のカオス複素多方向連想メモリに比べ、1対



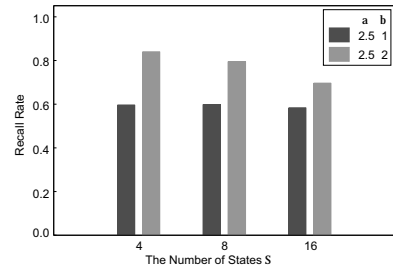
(a) 1-to-4 Patterns



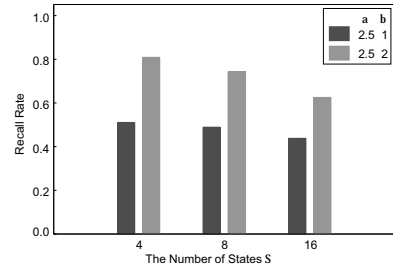
(b) 1-to-5 Patterns



(c) 1-to-6 Patterns



(d) 1-to-7 Patterns



(e) 1-to-8 Patterns

Fig. 11: One-to-Many Association Ability (5-layered Network) (2).

多の想起の想起能力が改善されている。

- (3) 従来のカオス複素多方向連想メモリでは不応性のスケーリングファクタ  $\alpha$  の値により、大きく想起能力が変化してしまい、適切なパラメータを設定することが困難であったが、提案モデルでは、不応性のスケーリングファクタ  $\alpha$  の値を時間的に変化させることで高い想起率が得られるようなパラメータの設定を比較的簡単に行うことができる。

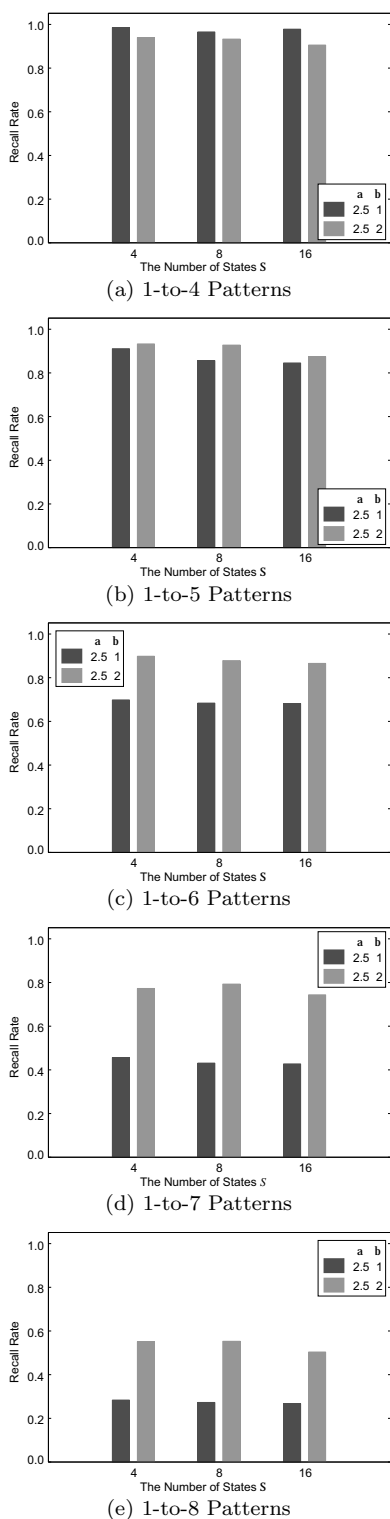


Fig. 12: One-to-Many Association Ability (6-layered Network) (2).

## 参考文献

- 1) 根本 幾, 河野 知志: “神経回路網の複素数値化の試み,” 電子情報通信学会論文誌 D-II, Vol.J74, No.9, 1282/1288 (1991)
- 2) 青木 宏之, 小林 幸夫: “ペナルティ項を有する複素連想記憶モデルの性質,” 電子情報通信学会論文誌 A, Vol.J81, No.11, 1538/1546 (1998)
- 3) 青木 宏之: “複素数体上に拡張したホップフィールド連想記憶の平衡状態の解析,” 電子情報通信学会論文誌 A,

Vol.J78, No.9, 1238/1241 (1995)

- 4) S. Jankowski, A. Lozowski and J. M. Zurada : “Complex-valued multistate neural associative memory,” IEEE Transactions on Neural Networks, Vol.7, No.6, 1491/1496 (1996)
- 5) 黒江 康明: “複素連想記憶のモデルとそのダイナミクス,” 電子情報通信学会論文誌, Vol.87, No.6, 454/459 (2004)
- 6) 合原 一幸編: カオス -カオス理論の基礎と応用-, サイエンス社 (1990)
- 7) K. Aihara, T. Takabe and M. Toyoda : “Chaotic neural networks,” Physics Letter A, Vol.144, No.6 & 7, 333/340 (1990)
- 8) Y. Osana and M. Hagiwara : “Separation of superimposed pattern and many-to-many associations by chaotic neural networks,” Proceedings of IEEE and INNS International Joint Conference on Neural Networks, Anchorage, Vol.1, 514/519 (1998)
- 9) 長名優子, 服部元信, 萩原 将文: “カオス双方向連想メモリ,” 電気学会論文誌 C, Vol.116, No.7, 741/747 (1996)
- 10) B. Kosko : “Bidirectional associative memories,” IEEE Transactions on Systems Man and Cybernetics, Vol.18, No.1, 49/60 (1988)
- 11) 橋本 直樹, 黒江 康明, 森 武宏: “複素ニューラルネットワークに対するエネルギー関数について,” 電子情報通信学会技術報告, No.NC98-152, 121/128 (1999)
- 12) M. Nakada and Y. Osana : “Chaotic complex-valued associative memory,” Proceedings of International Symposium on Nonlinear Theory and its Applications, Vancouver (2007)
- 13) Y. Yano and Y. Osana : “Chaotic complex-valued bidirectional associative memory,” Proceedings of IEEE and INNS International Joint Conference on Neural Networks, Atlanta (2009)
- 14) Y. Yano and Y. Osana : “Chaotic complex-valued bidirectional associative memory - one-to-many association ability -,” Proceedings of International Symposium on Nonlinear Theory and its Applications, Sapporo (2009)
- 15) Y. Shimizu and Y. Osana : “Chaotic complex-valued multidirectional associative memory,” Proceedings of IASTED Artificial Intelligence and Applications, Innsbruck (2010)
- 16) M. Hagiwara : “Multidirectional associative memory,” Proceedings of IEEE and INNS International Joint Conference on Neural Networks, Washington D.C., Vol.1, 3/6 (1990)
- 17) A. Yoshida and Y. Osana : “One-to-many association ability of chaotic complex-valued multidirectional associative memory,” Proceedings of International Symposium on Nonlinear Theory and its Applications, Kobe (2011)
- 18) Y. Osana : “Recall and separation ability of chaotic associative memory with variable scaling factor,” Proceedings of IEEE and INNS International Joint Conference on Neural Networks, Hawaii (2002)

# 複素多層パーセプトロンの探索空間と探索法

○鈴木 真矢 中野 良平 (中部大学)

## Search Space and Search Method for Complex-Valued Multilayer Perceptron

\* S. Suzumura and R. Nakano (Chubu University)

**Abstract**— A complex-valued MLP (multilayer perceptron) is expected to have greater potential than a real-valued MLP. The search space of complex-valued MLP seems to be in quite extraordinary shape full of crevasses having huge condition numbers for Hessians. In such extraordinary search space any ordinary method will not work. We propose a new search method for complex-valued MLP to find an excellent solution in such abnormal search space. Our experiments using 8-bit parity problem and an unbounded function evaluate the proposed method and the role of an activation function.

**Key Words:** Complex-valued MLP, Search space, Search method, Hessian matrix, Condition number

### 1 はじめに

多層パーセプトロン (MLP) はニューラルネットの中で最も知られるモデルである。特に、入出力や重みがすべて実数である実 MLP が広く知られるが、入出力や重みが複素数である複素 MLP は実 MLP には無い魅力と可能性を秘めている。例えば、複素 MLP における BP 学習では、実部要因と虚部要因の補完によって、実 MLP に比べて学習が速く進むという報告がある<sup>12)</sup>。また、複素数を扱う分野では、複素 MLP が自然に適用可能であり、実 MLP のように実部と虚部を分割して扱う必要がない<sup>9)</sup>。

複素 MLP の活性化関数は複素 MLP の能力や特性に重要な影響を持つと思われる、多くの関数形が提案されてきた<sup>5)</sup>。よく知られるのは、Kim ら<sup>4)</sup>と Leung ら<sup>6)</sup>が独立に提案した単純シグモイド型、そして、新田ら<sup>11)</sup>や Benvenuto ら<sup>1)</sup>が独立に提案した個別シグモイド型である。前者は、有界でなく、特異点が存在するため学習が収束しにくいとの報告がある。後者は、有界であり、収束性に関する難点が解消されたとの報告がある。

一方、実であれ複素であれ、MLP の探索空間がどのような形状をしているかに関する研究は計算機を駆使して実証的に進めるのが有効と考えるが、これまで十分に解明されたとはいえない。最急降下法やニュートン法などは、スキーマのゲレンデを多次元に拡張したような空間では有効に働くが、MLP の探索空間はそのようなバランス良い形状ではなく、異常に大きい条件数で特徴付けられる切り立ったクレバス帯に満ちていると思われる<sup>8)</sup>。そうした異常な探索空間では通常の探索法は機能しないので、異常形状に即した新探索法の提案が望まれる。

本稿では、複素 MLP の探索空間の形状を計算機実験を通して解明し、それを意識した新探索法を提案する。提案法は、通常の最急降下法に、クレバス帯を進むための探索法および可約性写像を利用した探索法を付加したものである。なお、活性化関数

として、Kim・Leung の活性化関数を若干修正したものをを用いる。新探索法の有効性を評価するため、8bit パリティ問題と有界でない関数を用いた計算機実験を行う。また、他の活性化関数も用いて、活性化関数の違いを比較評価する。

### 2 複素多層パーセプトロンモデル

本稿が対象とする複素3層パーセプトロンを Fig.1 に示す。重みや入出力はすべて複素数である。隠れ

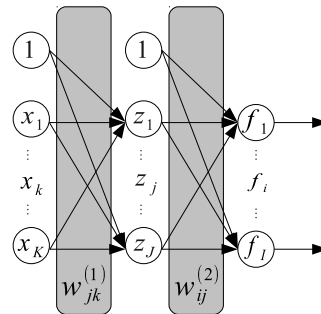


Fig. 1: 複素3層パーセプトロン

ユニットの出力  $z_j^\mu$  と出力ユニットの出力  $f_i^\mu$  の計算式を以下に示す。ただし、 $x_0 = z_0 = 1$  とし、右肩の  $\mu$  はサンプル番号を表す。

$$z_j^\mu = g(h_j^\mu), \quad h_j^\mu = \sum_{k=0}^I w_{jk}^{(1)} x_k^\mu \quad (1)$$

$$f_i^\mu = \sum_{j=0}^J w_{ij}^{(2)} z_j^\mu \quad (2)$$

$g(h)$  は活性化関数を表す。後述する提案法の  $g(h)$  は以下とする。なお、 $h = x + iy$ ,  $i = \sqrt{-1}$  とする。

$$g(h) = \frac{1}{1 + i + e^{-h}} \quad (3)$$
$$= \frac{e^{-x} \cos y + 1 + i(e^{-x} \sin y - 1)}{e^{-2x} + 2\sqrt{2}e^{-x} \sin(y + \frac{3\pi}{4}) + 2}$$

この活性化関数には以下の特徴がある。

- A : 特異点が存在する.
- B : 周期関数を含む.

これら特徴から次の恩恵が得られる. 特性 A より,  $f(x) = 1/x$  のように, 極のある関数が近似できる. 特性 B より, 周期関数が近似できる. 例えば,  $n$ -bit パリティ問題の近似である.  $n$ -bit 偶数パリティ問題は, 例えば以下のように近似できる.

$$f(\mathbf{x}) = \sin^2 \left( \frac{1}{2} \pi \sum_{k=1}^N x_k \right), \quad \text{where } x_k \in \{0, 1\} \quad (4)$$

周期性がない活性化関数では,  $(n/4)$  周期分の関数値が必要となり,  $n$  が巨大になると, 周期性のない活性化関数を用いた MLP では近似困難となる.

なお, 活性化関数  $g(x, y) = \cos x + i \sin y$  を用いれば, 任意の周期関数が近似できる. この活性化関数は,  $x, y$  に関する偏微分が常に滑らかであり, 探索空間の形状が比較的線形であると考えられるため, 周期関数の近似に適している.

### 3 複素多層パーセプトロンの探索空間

複素 MLP の探索空間について述べる. 探索空間の形状を知る有力な手段としてヘッセ行列がある. 複素 MLP では, 変数 (入出力, 重み) は複素数であるが, 目的関数は実関数である. 工学的には, そうした設定は少なく, Wirtinger 微分<sup>2)</sup> が古くから確立している. そこで, Wirtinger 微分を用いてヘッセ行列を求める. 記法には,  $r$  表現,  $c$  表現 (R 形式),  $c$  表現 (C 形式),  $z$  表現があるが,  $c$  表現 (C 形式) のヘッセ行列  $\mathbf{H}_c^C$  は以下となる. ただし,  $\mathbf{w}$  は重みベクトルで,  $\mathbf{c}^T = (\mathbf{w}^T \bar{\mathbf{w}}^T)$  とする.

$$\mathbf{H}_c^C = \frac{\partial^2 E}{\partial \bar{\mathbf{c}} \partial \mathbf{c}^T} \quad (5)$$

$\mathbf{H}_c^C$  はエルミート行列で, 固有値は実数である. 計算機実験では, 探索空間の形状を知るため, 固有値の他に, 条件数 (condition number) も求める.

計算機実験で後述するが, 複素 MLP では, 実 MLP<sup>8)</sup> と同様, ヘッセ行列の条件数が巨大になる場合が非常に多く, 探索空間は極めて悪条件の領域に満ちていると思われる. ヘッセ行列の条件数が巨大な空間は楕状の形状をしている. ここでは, 条件数が  $10^3 \sim 10^6$  の切り立った空間を楕状帯,  $10^6 \sim 10^{20}$  の極度に切り立った空間をクレバス帯と呼ぶ. 楕状帯もクレバス帯もその底の傾斜は非常に微小である. クレバス帯は極度に切り立っているため, 最急降下法では, 双方の壁の間を往復するのみで, 底に沿って進めず, やがて探索幅が極端に小さくなって停止する. クレバス帯のイメージ図を Fig.2 に示す.

### 4 新探索法

新探索法は, 通常最急降下法に, クレバス帯探索法と可約性写像探索法を付加したものである.

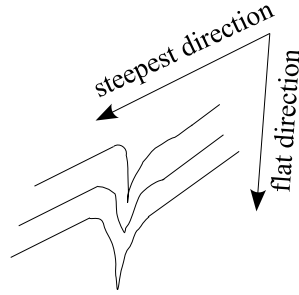


Fig. 2: クレバス帯のイメージ図

#### 4.1 最急降下法

複素 MLP における BP 学習は, 実 MLP に比べて学習が数倍速く進むという報告がある<sup>12)</sup>. また, 最急降下法は簡単に実装できるため広く利用される. 最急降下法の重み更新式は以下である. なお,  $\mathbf{w}$  は重みベクトル,  $E$  は目的関数,  $\eta$  は探索幅である.

$$\mathbf{w} \leftarrow \mathbf{w} - \eta \frac{\partial E}{\partial \bar{\mathbf{w}}} \quad (6)$$

以下, 複素 3 層パーセプトロンにおける最急降下法の重み更新式を具体的に求める. 目的関数は以下のスカラー実関数とする. ただし,  $y_i^\mu$  は出力ユニット  $i$  のサンプル  $\mu$  における教師信号である.

$$E = \sum_{\mu} \sum_i \delta_i^\mu \bar{\delta}_i^\mu, \quad \delta_i^\mu \equiv f_i^\mu - y_i^\mu \quad (7)$$

目的関数の勾配は以下ようになる.

$$\frac{\partial E}{\partial w_{jk}^{(1)}} = \sum_{\mu} \sum_i \delta_i^\mu \overline{w_{ij}^{(2)} g'(h_j^\mu) x_k^\mu} \quad (8)$$

$$\frac{\partial E}{\partial w_{ij}^{(2)}} = \sum_{\mu} \delta_i^\mu \bar{z}_j^\mu \quad (9)$$

このとき, 重み更新式は以下となる. なお,  $\eta$  は固定せず, 直線探索などで求める最適探索幅とする.

$$w_{jk}^{(1)} \leftarrow w_{jk}^{(1)} - \eta \frac{\partial E}{\partial w_{jk}^{(1)}} \quad (10)$$

$$w_{ij}^{(2)} \leftarrow w_{ij}^{(2)} - \eta \frac{\partial E}{\partial w_{ij}^{(2)}} \quad (11)$$

#### 4.2 クレバス帯探索法

クレバス帯や楕状帯の底に沿って探索することを考える. その勾配は極微小であるが, それに沿って進めば, 急激に降下する地点が見つかると思われる. 後述する計算機実験において, 目的関数  $E$  がずっと変わらない時期を過ぎると, 急激に降下する場合があることがそれを物語っている. ここでは, クレバ

ス帯を通過する一方法として、ランダム探索を利用する。その処理の流れを以下に示す。

#### クレバス帯探索法

- 1: 隠れユニット数を  $J$ , 探索開始時の誤差と重みベクトルを  $\hat{E}(J)$ ,  $\hat{\mathbf{w}}(J)$  とする。
- 2: **for**  $n = 1, 2, \dots, N_{max}$  **do**
- 3: 重みを更新する ( $\hat{w}_{jk}, \hat{w}_{ij} \in \hat{\mathbf{w}}(J)$ ).  
 $w'_{jk} \leftarrow \hat{w}_{jk} + \phi(n) \text{rand}()$   
 $w'_{ij} \leftarrow \hat{w}_{ij}$
- 4:  $\mathbf{w}'$  から再学習する (最急降下法)。
- 5: 学習後の誤差と重みベクトルを  $E'$ ,  $\mathbf{w}''$  とする。
- 6: **if**  $E' < \hat{E}(J)$  **then**
- 7:  $\hat{E}(J) \leftarrow E'$
- 8:  $\hat{\mathbf{w}}(J) \leftarrow \mathbf{w}''$
- 9: **end if**
- 10: **end for**

ここで、 $\phi(n)$  は、カウンタ  $n$  の増加に伴って減少する以下の関数である。

$$\phi(n) = \frac{N_{max} - n}{N_{max}} \quad (12)$$

なお、 $\text{rand}()$  は範囲  $[-1, 1]$  の一様乱数とし、 $N_{max}$  は近傍の最大探索回数である。また、本稿が対象とする複素 3 層パーセプトロンでは、出力層に活性化関数を設けておらず、隠れ層から出力層への入力線形であるため、ランダムに摂動する重みは、入力層と隠れ層間の重み  $\{w_{jk}\}$  のみとした。

#### 4.3 可約性写像探索法

実 MLP における重みの一意性と冗長性<sup>7)</sup>に対応して、複素 MLP でも重みの一意性と冗長性が知られる<sup>10)</sup>。そして、実 MLP の重みの冗長性を利用して、隠れユニット数が  $J-1$  の MLP から、それと等価な働きをする隠れユニット数  $J$  の MLP への写像 (可約性写像と呼ぶ) を考えることができる<sup>3)</sup>。実 MLP では、 $\alpha$ ,  $\beta$ ,  $\gamma$  の 3 種の可約性写像が知られる<sup>13)</sup>。複素 MLP の可約性写像も、複素 MLP の冗長性を利用して同様に考えることができる。

最適重み  $\hat{\mathbf{u}}$  を持つ複素 MLP( $J-1$ ) と等価な働きをする複素 MLP( $J$ ) を考える。MLP( $J$ ) の重みを  $\mathbf{w}$  とする。  $\hat{\mathbf{u}}$  から  $\mathbf{w}$  への可約性写像の一つに、 $w_{i1}^{(2)} = 0$ , ( $i = 1, \dots, I$ ) とする写像  $\alpha$  がある。重み  $\hat{\mathbf{u}}$  に写像  $\alpha$  を適用して得られる領域は以下となる。

$$\{\mathbf{w} \mid w_{i0}^{(2)} = \hat{u}_{i0}^{(2)}, w_{i1}^{(2)} = 0, \\ w_{ij}^{(2)} = \hat{u}_{ij}^{(2)}, w_{jk}^{(1)} = \hat{u}_{jk}^{(1)}, j = 2, \dots, J\}$$

写像  $\alpha$  では、 $w_{i1}^{(2)} = 0$  なので  $w_{1k}^{(1)}$  は自由である。幾つかある可約性写像の中で写像  $\alpha$  を採用したのは、重み  $\{w_{jk}\}$  の摂動を可能とするためである。

さて、特異領域は、全勾配が 0 で、ヘッセ行列の固有値の中に 0 になるものが存在する領域として定

義できる。MLP の探索空間の中には特異領域が存在するが、特異領域は可約性写像によって作られる。勾配型の探索は特異領域内では停滞するため、それを回避する工夫もなされている。しかし、一方では、特異領域を積極的に利用して良質の解を求める探索法<sup>8)</sup>も提案されており、可約性写像や特異領域は良質の解を得るための有力な出発点であると考えられる。ここでは、上記の可約性写像を利用して、隠れユニット数を一つ増して良質の解を求める探索法を提案する。その処理の流れを以下に示す。

#### 可約性写像探索法

- 1: 隠れユニット数を  $J$ , 探索開始時の誤差と重みベクトルを  $\hat{E}(J+1)$ ,  $\hat{\mathbf{w}}(J)$  とする。
- 2: 重み  $\hat{\mathbf{w}}(J)$  を  $\mathbf{w}(J+1)$  に写像する (可約性写像  $\alpha$ )。このとき追加する隠れユニット番号を  $j$  とし、 $\{w_{jk}\}$  は未定のままとする。
- 3: **for**  $m = 1, 2, \dots, M_{max}$  **do**
- 4: 写像時未定の重み  $w_{jk} \in \mathbf{w}(J+1)$  を  $[-1, 1]$  の範囲でランダムに初期化する。
- 5:  $w_{jk}$  初期化済みの  $\mathbf{w}(J+1)$  から再学習する (最急降下法)。
- 6: 学習後の誤差と重みベクトルを  $E'$ ,  $\mathbf{w}'$  とする。
- 7: **if**  $E' < \hat{E}(J+1)$  **then**
- 8:  $\hat{E}(J+1) \leftarrow E'$
- 9:  $\hat{\mathbf{w}}(J+1) \leftarrow \mathbf{w}'$
- 10: **end if**
- 11: **end for**

#### 4.4 新探索法の全体の処理

提案法は、最急降下法、クレバス帯探索法、可約性写像探索法から構成される。ここでは、学習法として最急降下法を用いたが、準ニュートン法など別の学習法を用いることも考えられる。提案法の全体の処理の流れを以下に示す。ここで、 $J_{max}$  は、利用者が指定する最大隠れユニット数である。

#### 新探索法の全体の処理

- 1:  $J \leftarrow 1$ .
- 2: 最急降下法を実行する。
- 3: **while** true **do**
- 4: クレバス帯探索法を実行する。
- 5: **if**  $J < J_{max}$  **then**
- 6: 可約性写像探索法を実行する。
- 7: **else**
- 8: 停止する。
- 9: **end if**
- 10:  $J \leftarrow J + 1$ .
- 11: **end while**

提案法を用いれば、隠れユニット数  $J = 1, 2, \dots, J_{max}$  毎の最小誤差  $\hat{E}(J)$  と最適重み  $\hat{\mathbf{w}}(J)$  が求まる。

## 5 計算機実験

計算機実験によって、提案法の有効性および活性化関数の違いを評価する。比較する活性化関数は以下の3種とした。ただし、 $h = x + iy$  とする。

$$\begin{aligned} g_1(h) &= \frac{1}{1+i+e^{-h}} \\ g_2(h) &= \cos x - i \sin y \\ g_3(h) &= \frac{1}{1+e^{-x}} + i \frac{1}{1+e^{-y}} \end{aligned}$$

### 5.1 8-bit パリティ問題の学習実験

一般に、 $n$ -bit パリティ問題は、ハミング距離が1だけ異なる入力信号が別クラスに属するため、実MLPでは識別が極めて困難な問題である。8-bit 偶数パリティ問題のデータを Table 1 に示す。

Table 1: 8-bit 偶数パリティ問題のデータ

入力信号	教師信号
00000000	0
00000001	1
00000010	1
⋮	⋮
11111101	1
11111110	1
11111111	0

複素MLPに入力する際は、表中の入力信号は入力変数群の実部に与え、虚部は0とする。複素MLPの構成は、入力ユニット数が8、出力ユニット数が1である。

#### (1) 最急降下法の実験

探索空間の形状調査、および提案法との性能比較のために、最急降下法のみを用いた実験を行った。最初に、実験条件を Table 2 の条件 1a とし、乱数のシード番号を変えて30回試行した。3種の活性化

Table 2: 実験条件 (8-bit パリティ問題, 最急降下法)

項目	条件 1a	条件 1b
学習サンプル数 ( $N$ )	256	256
隠れユニット数 ( $J$ )	1	3
初期重み	$[-5, 5]$	$[-5, 5]$

関数を用いた学習過程における誤差の変化を Fig.3 から Fig.5 に示す。また、全30回の試行のうち、学習後の誤差が0.1以下となった回数を Table 3 に示す。図や表から、活性化関数が異なると、探索結果に大きな違いが生じることがわかる。なお、最急降下法+ $g_1$  では、収束する時期が試行毎に大きく異なっている。

Table 3: 学習成功数 (8-bit パリティ問題, 最急降下法,  $J=1$ , 30 試行)

活性化関数	$g_1$	$g_2$	$g_3$
学習成功数	17回	11回	0回

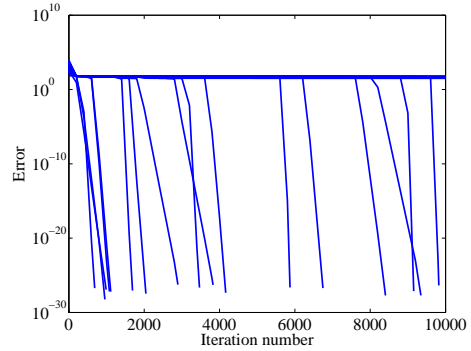


Fig. 3: 誤差の変化 (8-bit パリティ問題, 最急降下法+ $g_1$ ,  $J=1$ , 30 試行)

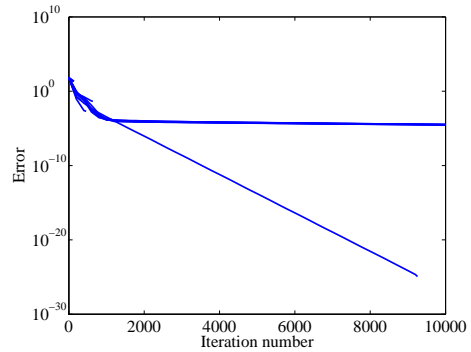


Fig. 4: 誤差の変化 (8-bit パリティ問題, 最急降下法+ $g_2$ ,  $J=1$ , 30 試行)

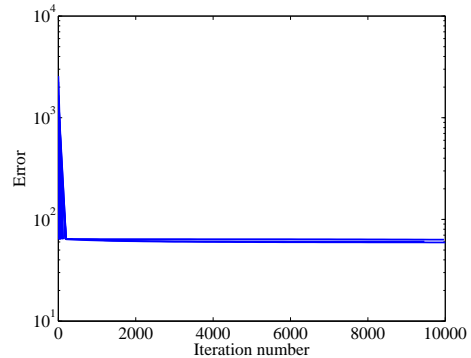


Fig. 5: 誤差の変化 (8-bit パリティ問題, 最急降下法+ $g_3$ ,  $J=1$ , 30 試行)

次に、活性化関数を  $g_1$  とした学習過程において、乱数シード番号1から5における誤差の変化、および条件数を Fig.6 と Fig.7 に示す。さらに、ヘッセ行列の固有値の変化の一例を Fig.8 に示す。これらの図を見ると、学習の初期は、誤差の減少がほとんどなく、条件数は  $10^5$  前後である。その後突然、誤差が急激に減少し、それと同期して条件数が急激に増大し、 $10^{16} \sim 10^{18}$  に至る。各固有値の変化の図を見ると、条件数の急激な増大は、最小固有値が急激に0に近づいたことが原因であることがわかる。このとき、ヘッセ行列は縮退する。誤差を最小にする正解は切り立ったクレバスの底に存在する。

最後に、実験条件を Table 2 の条件 1b、つまり、



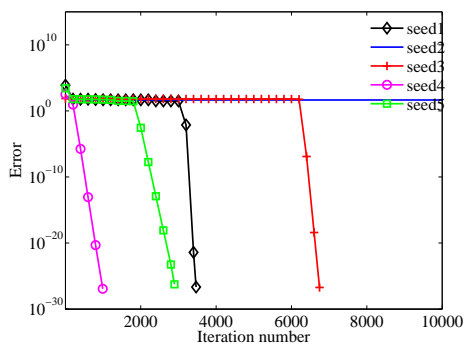


Fig. 6: 誤差の変化 (8-bit パリティ問題, 最急降下法+ $g_1$ ,  $J=1$ , 5 試行)

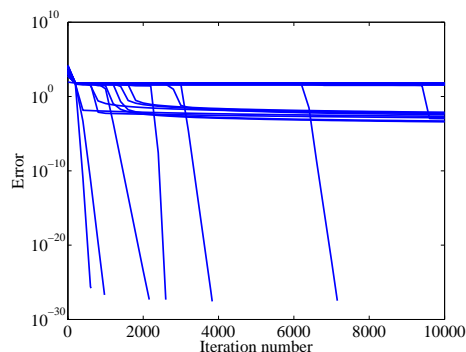


Fig. 9: 誤差の変化 (8-bit パリティ問題, 最急降下法+ $g_1$ ,  $J=3$ , 30 試行)

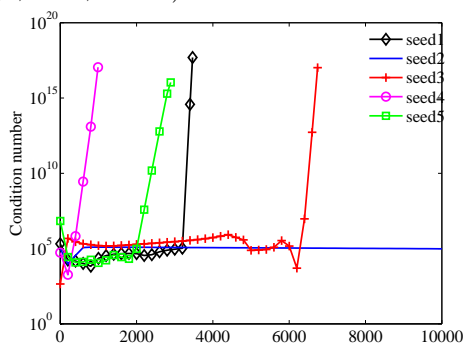


Fig. 7: 条件数の変化 (8-bit パリティ問題, 最急降下法+ $g_1$ ,  $J=1$ , 5 試行)

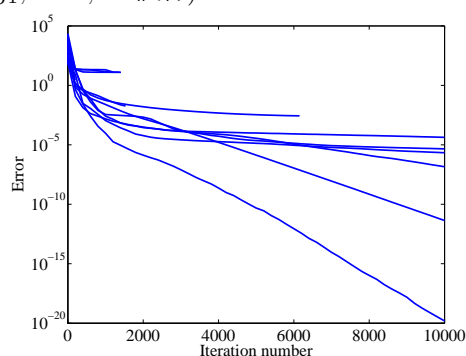


Fig. 10: 誤差の変化 (8-bit パリティ問題, 最急降下法+ $g_2$ ,  $J=3$ , 30 試行)

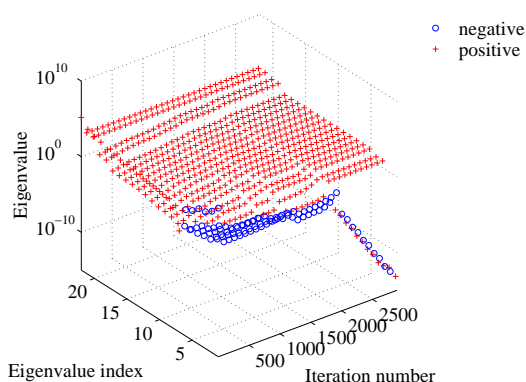


Fig. 8: 固有値の変化 (8-bit パリティ問題, 最急降下法+ $g_1$ ,  $J=1$ , シード No.5)

隠れユニット数を 3 として, 同様な実験を行った. 3 種の活性化関数を用いた学習過程における誤差の変化を Fig.9 から Fig.11 に示す. 全 30 回の試行のうち, 学習後の誤差が 0.1 以下となった回数を Table 4 に示す. 隠れユニット数を 1 から 3 へ増やしても成功数は増えず, むしろ減少した. 8-bit パリティ問題に対しては, 複素 MLP は隠れユニットが 1 個あれば十分であり, 余分な隠れユニットは探索に貢献しない結果となった.

## (2) 提案法の実験

提案法を 8-bit 偶数パリティ問題に適用した. 実験条件は, Table 5 の条件 2a とし, 乱数のシード

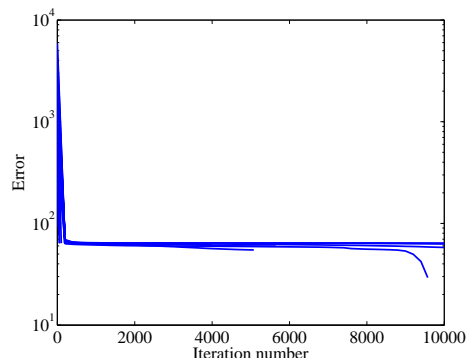


Fig. 11: 誤差の変化 (8-bit パリティ問題, 最急降下法+ $g_3$ ,  $J=3$ , 30 試行)

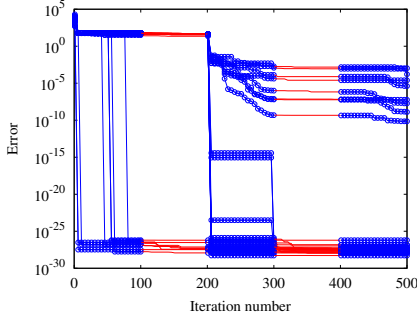
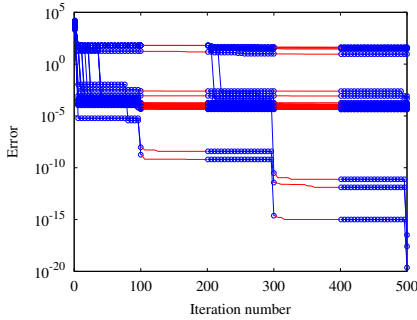
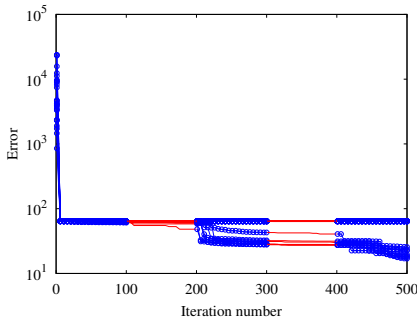
Table 4: 学習成功数 (8-bit パリティ問題, 最急降下法,  $J=3$ , 30 試行)

活性化関数	$g_1$	$g_2$	$g_3$
学習成功数	15 回	8 回	0 回

番号を変えて 30 回試行した. 3 種の活性化関数を用いた学習過程における誤差の変化を Fig.12 から Fig.14 に示す. 各図において, iteration number の 0 から 100 まで, 200 から 300 まで, 400 から 500 までがクレバス探索で, 100 から 200 までと 300 から 400 までが可約性写像探索である. 全 30 回の試行のうち, 学習後の誤差が 0.1 以下となった回数を Table 6 に示す. 提案法においても, 活性化関数

Table 5: 実験条件 (8-bit パリティ問題, 提案法)

項目	条件 2a	条件 2b
学習サンプル数 ( $N$ )	256	64
最大隠れユニット数 ( $J_{max}$ )	3	1
最大探索回数 ( $N_{max}$ )	100	100
最大探索回数 ( $M_{max}$ )	100	100
初期重み	$[-5, 5]$	$[-5, 5]$

Fig. 12: 誤差の変化 (8-bit パリティ問題, 提案法 + $g_1$ ,  $J_{max}=3$ , 30 試行)Fig. 13: 誤差の変化 (8-bit パリティ問題, 提案法 + $g_2$ ,  $J_{max}=3$ , 30 試行)Fig. 14: 誤差の変化 (8-bit パリティ問題, 提案法 + $g_3$ ,  $J_{max}=3$ , 30 試行)

が異なると探索結果に大きな違いが生じる。また、同じ活性化関数  $g_1$  を用いた 30 試行のうち、最急降下法の成功は 15~17 回であるが、提案法では全部成功した。ただ、早期の収束もあれば、遅い収束もある。

### (3) 汎化能力の評価

8-bit 偶数パリティ問題を用いて、提案法+ $g_1$  の汎化能力を評価する。実験条件は Table 5 の条件 2b

Table 6: 学習成功数 (8-bit パリティ問題, 提案法,  $J_{max}=3$ , 30 試行)

活性化関数	$g_1$	$g_2$	$g_3$
学習成功数	30 回	25 回	0 回

とする。全 256 サンプルから無作為抽出した 64 サンプルを学習データとし、残り 192 サンプルを未学習データとする。学習後は、未学習の 192 サンプルを含む全 256 サンプルに対する誤差を評価する。実験の結果、全 256 サンプルに対する誤差  $E$  は  $4.7353 \times 10^{-27}$  であった。また、学習後の重みを Table 7 に示す。

Table 7: 学習後の重み (8-bit パリティ問題, 提案法+ $g_1$ ,  $J_{max}=1$ ,  $N=64$ )

入力-隠れ層の重み $w_{jk}$ $j=1, k=0,1,\dots,8$	隠れ-出力層の重み $w_{ij}$ $i=1, j=0,1$
1.10216 + 2.33163i	2.62821 - 0.05229i
0.00000 - 3.14159i	-2.06599 - 1.95563i
0.00000 - 3.14159i	-
0.00000 + 3.14159i	-
0.00000 - 3.14159i	-
0.00000 - 3.14159i	-
0.00000 + 3.14159i	-
0.00000 - 3.14159i	-
0.00000 - 3.14159i	-

実験の結果、未学習の 192 サンプルを含め、全 256 サンプルに対して教師信号を正しく表現できた。提案法+ $g_1$  は周期関数の近似において高い汎化能力を持つと考えられる。

## 5.2 有界でない関数の学習実験

次に、以下のような有界でない関数を用いて提案法の有効性を評価する。

$$f(x) = 2x + i \frac{1}{10x} \quad (13)$$

ここで、 $x$  は実数変数であり、 $f(x)$  は  $x$  を変数とする複素関数  $f(x) : \mathbb{R} \rightarrow \mathbb{C}$  である。 $x \in [-1, 1]$  のとき、式 (13) は変化量が 4 の関数  $2x$  と変化量が無限大の関数  $\frac{1}{10x}$  の和であり、その近似は相当に困難である。また、 $x = -1, -0.99, -0.98, \dots, 1$  のとき、 $\frac{1}{10x}$  は、高次の導関数が 0 に収束せず、ルンゲ現象によって多項式補間できない関数である。

今、 $x$  を区間  $[-1, 1]$  において 0.01 間隔にサンプリングし、 $f(x^\mu)$  を求める。 $f(x^\mu)$  を教師信号として、 $\{x^\mu, f(x^\mu)\}$  の組を複素 MLP の学習データとする。ただし、 $x=0$  は除外する。

以下、複素関数  $f(x)$  と複素 MLP の出力結果を可視化して対比し、提案法の学習性能を評価する。提案法+活性化関数  $g_1$  による学習結果を、真の関数値とともに Fig.15 に示す。また、 $x$  を区間  $[-2, 2]$  において  $10^{-3}$  間隔に刻んだサンプルの多くは未知サンプルである。それらの出力結果を、真の関数値とともに Fig.16 に示す。

学習結果の Fig.15 から、提案法+ $g_1$  は、有界でない関数、即ち、特異点(極)を持つ関数の近似に適していると考えられる。また、Fig.16 を見ると、

Table 8: 実験条件 (非有界関数, 提案法)

学習サンプル数 ( $N$ )	200
最大隠れユニット数 ( $J_{max}$ )	10
最大探索回数 ( $N_{max}$ )	100
最大探索回数 ( $M_{max}$ )	100
初期重み	$[-5, 5]$

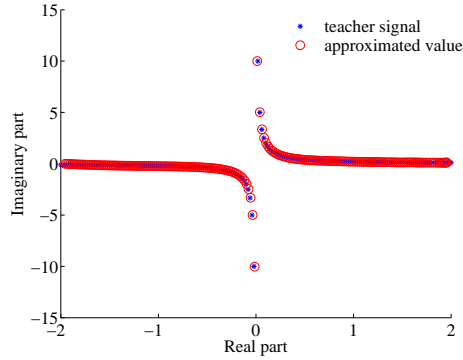


Fig. 15: 学習データの出力と真の出力 (非有界関数, 提案法+ $g_1$ )

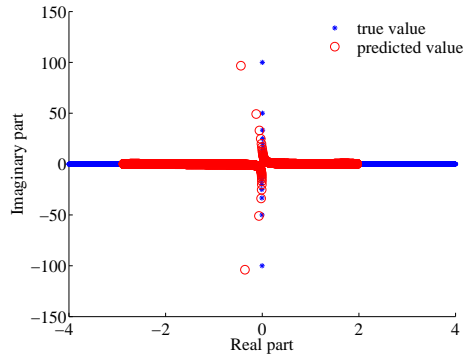


Fig. 16: 未知データの出力と真の出力 (非有界関数, 提案法+ $g_1$ )

未知サンプルに対しては, 虚部の出力は真の関数を正確に表現できているが, 実部の出力は  $x = 0$  の近傍でややずれが生じている.

## 6 おわりに

複素 MLP の探索空間と新探索法を論じ, 実験で評価した. 今回の実験では以下の傾向が見られた. 複素 MLP の探索空間はヘッセ行列の条件数が  $10^5 \sim 10^{17}$  となる切り立ったクレバス帯に満ちている. 複素 MLP の活性化関数は, 極を含んでいても, 周期性を持てば, 実 MLP には無い強力な近似能力を持つ. 同じ探索法を用いても活性化関数が異なると探索能力に大きな違いが生じる. 今後は提案法を他の問題に適用し改良を進める予定である.

## 参考文献

- 1) N. Benvenuto and F. Piazza. On the complex backpropagation algorithm. *IEEE Trans. Signal Processing*, 40(4):967/969 (1992).
- 2) K.K. Delgado. The complex gradient operator and the CR-calculus. *ECE275A-Lecture Supplement* Fall (2006).

- 3) K. Fukumizu and S. Amari. Local minima and plateaus in heirarchical structure of multilayer perceptrons. *Neural Networks*, 13(3):317/327 (2000).
- 4) M.S. Kim and C.C. Guest. Modification of backpropagation networks for complex-valued signal processing in frequency domain. In *Proc. IJCNN'90 vol.3*, 27/31 (1990).
- 5) Y. Kuroe, M. Yoshida, and T. Mori. On activation functions for complex-valued neural networks -existence of energy functions-. In *Proc. ICANN/ICONIP 1993*, 985/992 (2003).
- 6) H. Leung and S. Haykin. The complex backpropagation algorithm. *IEEE Trans. Signal Processing*, 39(9):2101/2104 (1991).
- 7) H. J. Sussmann. Uniqueness of the weights for minimal feedforward nets with a given input-output map. *Neural Networks*, 5(4):589/593 (1992).
- 8) 佐藤聖也, 大脇高之, 中野良平. 多層パーセプトロンの特異領域を利用した探索法. 信学技法 N2010-81, 85/90 (2010).
- 9) 新田徹. 複素ニューラルネットワーク. 電子情報通信学会誌, 83(8):612/615 (2000).
- 10) 新田徹. 複素ニューラルネットワークにおける一意性定理とパラメータの冗長性. 電子情報通信学会論文誌, J85-D-II(5):796/804 (2002).
- 11) 新田徹, 古谷立美. 複素バックプロパゲーション学習. 情報処理学会論文誌, 32(10):1319/1329 (1991).
- 12) 新田徹, 古谷立美. 複素バックプロパゲーション学習アルゴリズムの学習特性. 情報処理学会論文誌, 34(1):29/38 (1993).
- 13) 中野良平. ニューラル情報処理の基礎数理. 数理工学社 (2005).



# 複素ネットワークインバージョンによる逆問題解法と正則化

○中村恭介 小川毅彦 (拓殖大学)

## Solution of Inverse Problems by Complex-Valued Network Inversion with Regularization

\* K. Nakamura and T. Ogawa (Takushoku University)

**Abstract**— Complex-valued network inversion is a solution for inverse problems using a complex-valued multilayer neural network. In complex-valued network inversion, the complex input is estimated from the provided complex output using a trained network. In general, there is a problem attributable to the ill-posedness on the inverse problems. Regularization imposes specific conditions on an ill-posed inverse problem to convert it into a well-posed problem. In this study, the effect of the regularization for complex-valued network inversion is examined by the ill-posed inverse problems. We examine the inverse estimation of the complex mapping and the distributed generation.

**Key Words:** complex-valued network inversion, inverse problems, regularization.

### 1 はじめに

科学・工学分野において、観測結果から原因を推定する逆問題解法は重要である<sup>1)</sup>。多層型ネットワークを用いて逆問題を解く方法としてネットワークインバージョンが提案されている<sup>2)</sup>。従来のネットワークインバージョンの解法は実数領域の入出力と荷重を持つ多層型ネットワークの解法であるが、複素ニューラルネットワークを用いた逆問題解法として複素ネットワークインバージョンが提案された<sup>3)</sup>。

逆問題には不良設定性の問題があり、解の存在・一意性・安定性が保証されない<sup>4)</sup>。不良設定性の解決のために、正則化法によって拘束条件を与える方法が用いられる<sup>5)</sup>。複素ネットワークインバージョンにおいても不良設定性の問題は重要であり、正則化法を用いた解決法が考えられる。

本研究では複素ネットワークインバージョンに正則化法を導入する際の効果をシミュレーションによって確かめる。問題としては、複素写像の逆推定問題と分散電源逆推定問題を考え、一意性に関する不良設定性問題を対象とする。

### 2 逆問題とニューラルネットワーク

順問題が原因から結果を導くものであるのに対し、逆問題は結果から原因を求める、あるいは出力から入力を求めるものである。Fig. 1に示すように、一般に逆問題には不良設定性の問題があり、解の存在・一意・安定性が保証されない。この解決のために、不良設定問題に拘束条件を与え、良設定問題化する方法として正則化法がしばしば用いられる。

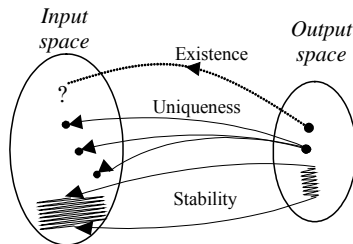


Fig.1 Concept of inverse problem and ill-posedness.

### 2.1 複素ネットワークインバージョン

通常が多層型ニューラルネットワークは、入力から出力への方で学習を行い、学習によって得られた順方向の関係をj用いて入力から出力を推定しているため、順問題の解法であると言える。これに対し、学習によって得られた順方向関係を逆にj用いて出力から入力を推定することによって、多層型ニューラルネットワークを用いて逆問題を解く方法がネットワークインバージョンである。ネットワークインバージョンを複素多層型ネットワークに拡張したものが複素ネットワークインバージョンである。

複素ネットワークインバージョンでは、通常のネットワークインバージョンの入力修正の原理を複素領域に拡張し、学習済みの複素多層型ニューラルネットワークを用いて、複素出力から複素入力を推定する。実際には、学習済みの複素多層型ニューラルネットワークに対して複素ランダム入力を与え、その複素出力と目標複素出力の誤差を逆伝搬し、複素入力の修正を繰返し行うことによって、複素出力から複素入力を推定する。

学習時は、まず複素学習入力  $x = x_R + ix_I$  と複素学習出力  $y = y_R + iy_I$  をj用いて、複素出力誤差  $E = E_R + iE_I$  を計算する。そして、複素荷重  $w = w_R + iw_I$  の実部と虚部をそれぞれ、

$$\begin{aligned} w_R(n+1) &= w_R(n) - \varepsilon_t \left( \frac{\partial E_R}{\partial w_R} + \frac{\partial E_I}{\partial w_R} \right) \\ w_I(n+1) &= w_I(n) - \varepsilon_t \left( \frac{\partial E_I}{\partial w_I} - \frac{\partial E_R}{\partial w_I} \right) \end{aligned} \quad (1)$$

の式に基づき更新する。  $\varepsilon_t$  は学習係数を表す。この更新手順を繰返すことによって、順方向の関係を学習することができる。これは通常の複素バックプロパゲーション法<sup>6)</sup>に基づく学習手順である。

逆推定時は、学習時に得られた関係を固定したまま、複素入力  $x = x_R + ix_I$  に仮の入力を与え、出力  $y = y_R + iy_I$  から複素出力誤差  $E = E_R + iE_I$  を計算する。そして、複素入力の実部と虚部をそれぞれ、

$$\begin{aligned} x_R(n+1) &= x_R(n) - \varepsilon_e \left( \frac{\partial E_R}{\partial x_R} + \frac{\partial E_I}{\partial x_R} \right) \\ x_I(n+1) &= x_I(n) - \varepsilon_e \left( \frac{\partial E_I}{\partial x_I} - \frac{\partial E_R}{\partial x_I} \right) \end{aligned} \quad (2)$$

の式に基づき更新する。  $\varepsilon_e$  は入力更新の係数を意味する。 ネットワークインバージョンによる入力の繰り返し更新の概念を Fig. 2 に示す。 この更新手順を繰り返すことによって、複素入力出力の値を満たすように修正されていく。 結果として、学習によって得られた複素荷重分布を使用して、複素出力から複素入力を繰り返し法によって逆推定することができる。

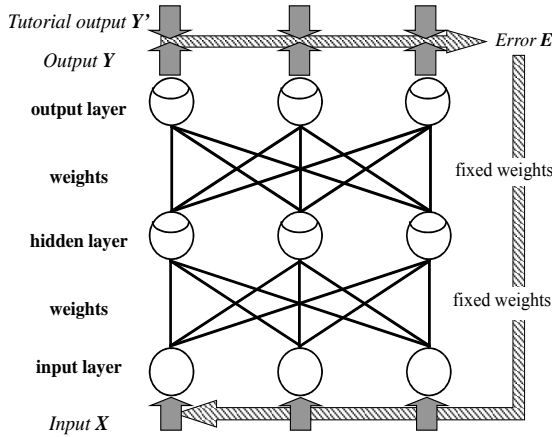


Fig. 2 Iterative correction of input from provided output in network inversion.

## 2.2 複素ネットワークインバージョンと正則化

複素ネットワークインバージョンにおいても、不良設定性の問題は生じる。複素ネットワークインバージョンに対しても、正則化法の考え方を適用することが可能である。正則化法では、ある条件を与えるための汎関数を繰り返し法の誤差関数に追加する。例えば、正則化汎関数  $K(x)$  を追加した誤差関数  $E$  を

$$E = \frac{1}{2} \sum_r (y'_r - y_r)^2 + \lambda \sum_k K(x_k) \quad (3)$$

として最小化することで、拘束条件を満たす誤差最小解を探すことを行う。ここで  $K(x_k)$  と  $\lambda$  はそれぞれ正則化汎関数と正則化係数を表す。

この方法はネットワークインバージョンおよび複素ネットワークインバージョンにも適用できる。拘束条件を与えるために、ネットワークインバージョンでの逆推定時にネットワーク出力誤差に加えて正則化汎関数を最小化する。エネルギー関数を、出力誤差関数に正則化項を加え、実部と虚部それぞれ独立に

$$\begin{aligned} E_R &= \frac{1}{2} \sum_r (y'_{rR} - y_{rR})^2 + \lambda \sum_k K(x_{kR}) \\ E_I &= \frac{1}{2} \sum_r (y'_{rI} - y_{rI})^2 + \lambda \sum_k K(x_{kI}) \end{aligned} \quad (4)$$

と定義する。ここで  $y'_r, K, x_k$  はそれぞれネットワーク

出力、正則化汎関数、および入力を表す。パラメータ  $\lambda$  は正則化係数を表す。この正則化項の目的は、入力  $x$  が満たすべきで、誤差項と同時に最小化されるべき拘束条件を表現する。この正則化項を用いることで、出力誤差最小かつ拘束条件満足の許容解を求めることが期待される。正則化項は入力に関する拘束条件なので、例えばノルム最小解や、複数解間の誤差最小解などを、正則化項によって表現することで選ぶことができる。本研究では、正則化法を一意性に関する不良設定問題を緩和する方法として導入する。

## 3 シミュレーション

本研究では、不良設定性をもつ逆問題として、複素写像の逆推定問題と、分散電源の逆推定問題を取り上げる。以下で問題の概要と不良設定性について説明し、さらに、正則化法を導入した複素ネットワークインバージョンによる逆推定シミュレーションを行い、正則化法による不良設定性緩和の効果を示す。

### 3.1 複素写像逆推定問題

最初に、複素領域にわたる簡単な逆問題の例として、複素写像の逆推定問題を取り上げる。この問題ではまず、複素ニューラルネットワークによって複素平面上の点間に与えられた写像が学習される。そして、複素平面上に任意に与えられる点の逆写像を、学習時に得られた順方向関係を逆に用いて推定する<sup>7)</sup>。

本研究では、複素数の入力データに対して、実部が負の場合はその虚数部の絶対値をとった複素数の出力データに写像するような複素写像を考える。ただし、実部が正の場合は虚部も正のデータのみを考える。すなわち、

$$\begin{cases} y_R + iy_I = x_R + i|x_I| & \text{for } x_R < 0 \\ y_R + iy_I = x_R + ix_I & \text{for } x_R, x_I > 0 \end{cases} \quad (5)$$

という入出力データ間の対応を学習し、与えられた出力データに対応する入力データの逆推定を試みる。この写像では、複素数平面の第1象限に与えられた出力に対する入力は一意に決まるが、第2象限に与えられた出力に対する入力は2通り存在し、さらに第3・第4象限に与えられた出力に対する入力は存在しないという、一部不良設定逆問題である。

Table 1 に示すパラメータをもつネットワークを用いて、式(5)の関係を学習させた。一意性に関する不良設定性緩和の効果を確認するために、学習を行ったネットワークに対して、第1・第2象限に分布する出力を与えて逆推定を行った。Fig. 3 は、正方向への拘束条件を加えた正則化による逆推定結果、Fig. 4 は負方向への拘束条件による逆推定結果である。これらによると、解が一意に決まる第1象限においてはどちらも同様な入力が推定されているが、解が2通り考えられる第2象限出力に対しては、正則化法によってそれぞれ第2象限および第3象限の入力が選択的に推定されていることがわかる。



Table 1 Network parameters for the simulation of inverse complex mapping problem.

Number of input neurons	1
Number of hidden neurons	10
Number of output neurons	1
Training rate $\varepsilon_t$	0.01
Input correcting rate $\varepsilon_e$	0.01
Max. number of training epoch	10000
Max. number of estimating epoch	10000
Training error to be attained	0.001
Estimation error to be attained	0.001

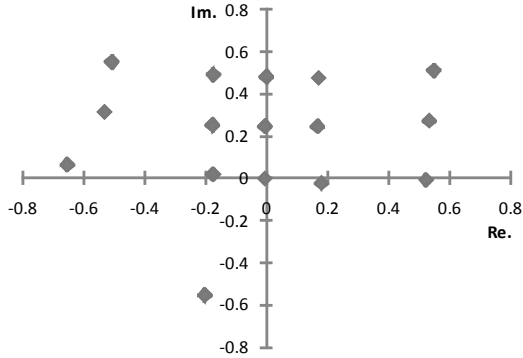


Fig.3 Estimated result of inverse mapping with regularization toward positive direction of imaginary axis.

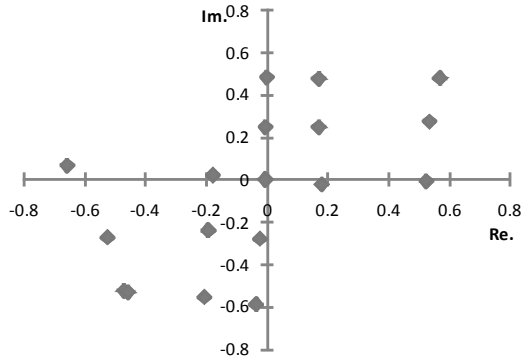


Fig.4 Estimated result of inverse mapping with regularization toward negative direction of imaginary axis.

### 3.2 分散電源逆推定問題

需要地近傍に分散配置される小規模な電源が分散電源であり<sup>8)</sup>, 分散電源パラメータを推定する問題は重要である. この問題は複数の測定点における複素電位・電流データから分散電源の複素電圧データを推定する逆問題である<sup>9)10)</sup>.

本研究では Fig. 5 に示す 2 つの分散電源からなる回路を考え, 出力電位から分散電源のパラメータを逆推定する問題を考える. この回路において各インピーダンスの値を  $Z1 = Z2 = Z3 = 1 + i [\Omega]$  とし, 電源のパラメータを変更することによって学習データを得る. 電源のパラメータとしては電圧を  $0 \sim 200[V]$ , 位相を  $-180 \sim 180[度]$  に変化させた. 本モデルでは,  $VG1$  と  $VG2$  のパラメータについて解の一意性に関する不良設定問題となる. 以下では  $VG1$  および  $VG2$  のパラメータに関して正則化法によって制約条件を与える場合の, 不良設定性による推定結果への影響をシミュレーション

によって確かめる.

Table 2 に示すパラメータをもつネットワークを用いて,  $VG1$ ,  $VG2$ ,  $VG3$  および  $IZ3$  の関係を学習させた.  $VG1$  および  $VG2$  に関する一意性に関する不良設定性緩和の効果を確認するために,  $VG2$  に拘束条件を与えて逆推定を行った. Fig. 6 は正則化による拘束条件を与えなかった場合の推定結果, Fig. 7 は  $VG2$  に関して拘束条件を与えた正則化による逆推定結果である. これらによると, 正則化による拘束条件を与えない場合は, 解の推定が正しく行われないのに対し,  $VG2$  に拘束条件を与えることによって,  $VG1$  が正しく推定されていることがわかる.

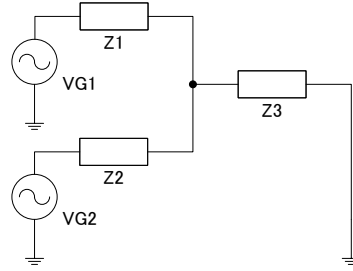
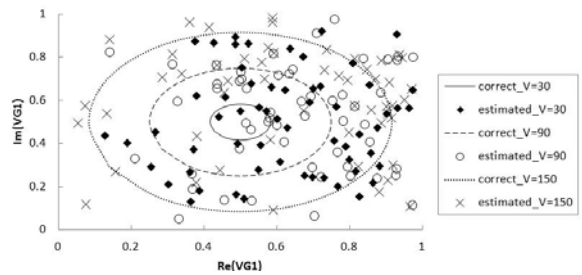


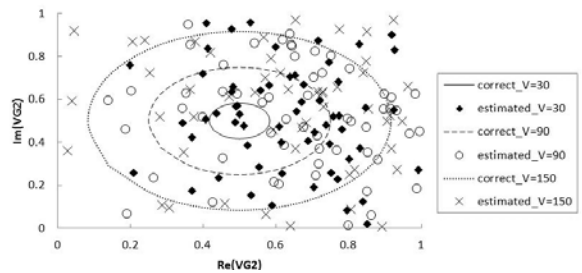
Fig. 5 Distributed generation circuit model for simulation, which has two power sources and an output.

Table 2. Network parameters for the simulation of inverse estimation of distributed generation

Number of input neurons	2
Number of hidden neurons	10
Number of output neurons	2
Training rate $\varepsilon_t$	0.0001
Input correcting rate $\varepsilon_e$	0.0001
Max. number of training epoch	10000
Max. number of estimating epoch	10000
Training error to be attained	0.0001
Estimation error to be attained	0.0001



(a)



(b)

Fig. 6. Estimated results of the power supplies (a)  $VG1$  and (b)  $VG2$  without regularization.

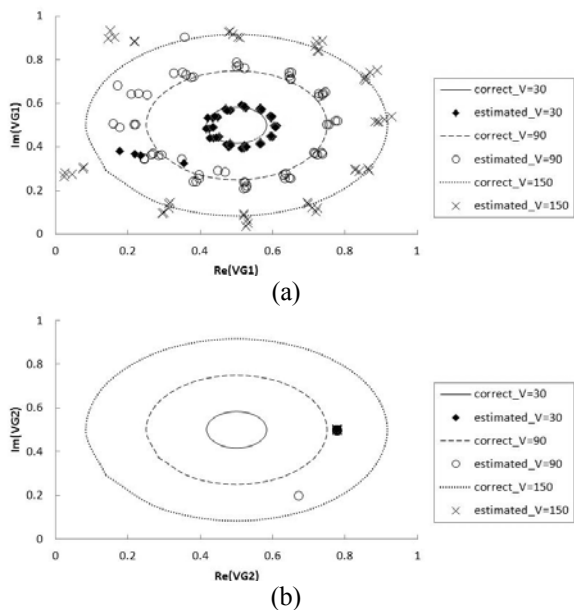


Fig. 7. Estimated results of the power supplies (a) VG1 and (b) VG2 with regularization for VG2.

#### 4 おわりに

本研究では複素領域に拡張した複素ネットワークインバージョンに正則化法を導入することで一意性に関する不良設定性の緩和を試みた。具体的には、複素写像逆推定問題と分散電源逆推定問題を用いて、正則化法を導入して不良設定性の解決をシミュレーションによって確認した。

今後の課題として、存在性や安定性に関する不良設定性緩和のための正則化法の検討や、推定精度の向上のための方策の検討、分散電源問題においては電源を増やした場合の推定の検討が挙げられる。

#### 謝辞

本研究の一部は、平成 23 年度科学研究費補助金（課題番号 21700260）の助成を受けて行われた。

#### 参考文献

- 1) C. W. Grötsch: Inverse problems in the mathematical sciences, Informatica International (1993)
- 2) A. Linden and J. Kindermann: Inversion of multilayer nets, Proc. Int. Joint Conf. on Neural Networks, 425/430 (1989)
- 3) 小川, 金田: 複素逆問題の解法のための複素ネットワークインバージョン, 信学論, **J88-D-II-9** 1954/1962 (2005)
- 4) A. N. Tikhonov and V. Y. Arsenin: Solutions of ill-posed problems, Winston & Sons (1977).
- 5) Y. P. Petrov and V. S. Sizikov: Well-posed, ill-posed, and intermediate problems with application, Koninklijke Brill NV (2005)
- 6) T. Nitta: An extension of the backpropagation algorithm to complex numbers, Neural Networks, **10-8**, 1392/1415 (1997)
- 7) S. Fukami, T. Ogawa and H. Kanada: Regularization for complex-valued network inversion, Proc. SICE Annual Conf., 1237/1242 (2008)
- 8) G. T. Heydt: The next generation of power distribution systems, IEEE Trans. on Smart Grid, **1-3**, 225/235 (2010)
- 9) N. V. Korovkin, V. L. Chechurin and M. Hayakawa: Inverse problems in electric circuits and electromagnetics, Springer (2007)
- 10) T. Ogawa, S. Fukami and H. Kanada: Inverse estimation of distributed generation by complex-valued network inversion, Proc. SICE Annual Conf., 247/250 (2010)



# 量子ビット遺伝的アルゴリズムの基本性能評価

○村本憲幸 磯川悌次郎 松井伸之 (兵庫県立大学)

## Qubit-Inspired Genetic Algorithm and Its Performance

\* N. Muramoto, T. Isokawa and N. Matsui (University of Hyogo)

**Abstract**—Qubit-inspired Genetic Algorithm (QGA) is an extension of genetic algorithm in which quantum mechanics and its representations are introduced. It has been shown that this scheme has better performances than the classical ones in several problems, but this has not yet been applied to more practical problems. In this paper, a chromosome in QGA is concretized as a series of quantum-bit (qubit) described by its complex-valued representation, and phase-rotation gates are embedded into the selection process over generations. We show the diversity of individuals in a population for this QGA by solving  $N$ - $K$  landscape problem, Knapsack Problem and Maximum Search. We also show the performance of QGA by applying the positioning problem in 2D figures.

**Key Words:** Qubit-inspired, Genetic Algorithm, Complex-valued representation, Knapsack problem, Maximum Search,  $N$ - $K$  landscape problem, Positioning problem

## 1 まえがき

Fogel の進化的プログラミング(Evolutionary Programming)<sup>1)</sup>, Rechenberg & Schwefel らの進化戦略(Evolution Strategy)<sup>2)</sup>そして Holland の遺伝的アルゴリズム(Genetic Algorithm:GA)<sup>3)</sup>が 1960 年代後半から提唱されて以来, これらは組合せ最適化問題や関数最適化などに対する優れた計算手法であるがゆえに, 実用拡大をめざして進化計算の性能向上法が様々に展開され活発に研究されてきた. 近年では, 量子アルゴリズム研究<sup>4)</sup>の活性化に伴って量子計算概念を取り入れた量子進化的計算も試みられている. 進化計算は生物進化の特徴を模した計算過程であり各種の組合せ最適化問題などへの強力な解法手法であるが, 量子アルゴリズムもまた, Shor の素因数分解解法や Grover のデータベース検索解法から明らかのように, 組合せ問題を高速に解く有望なアルゴリズム<sup>4)</sup>であり, 従来の性能を向上する手法として有望視されつつある. しかしながら, Shor と Grover の量子アルゴリズム以外の有望な量子アルゴリズムは現状では知られておらず, 量子計算手法の有効性を示す研究が国内外で切望されている. 本研究では, このような状況も鑑みて量子進化的計算の基本性能を調べようとしたものである.

量子進化的アルゴリズム(Quantum-inspired Evolutionary Algorithm:QEA)は1996年にNarayananらにより提案され<sup>5)</sup>, その後Hanらは染色体を量子ビットで表現し, ナップサック問題を用いてQEAが従来の遺伝的アルゴリズムよりも収束速度が速く, より良い解が求まることを示した<sup>6)</sup>. このようにQEAがGAに比べて収束速度などに優れていることは数値実験的に示されてきている. 我々も複素数表示に基づく量子ビット表現を導入した量子ビット遺伝的アルゴリズム(Qubit-inspired Genetic Algorithm: QGA)を構築し, その性能を評価し画像処理応用を試みてきた<sup>7)</sup>.

本報告では, ナップサック問題, 最大値探索問題および $N$ - $K$ ランドスケープ問題をベンチマーク例としてこのQGAの基本性能をさらに精査した結果を報告するとともに, 2次元図形の位置決定問題を通して, QGAの実応用への可能性を探る.

## 2 量子ビットと量子状態

本節では, 量子計算の基本的な概念の説明を行う. はじめに, 量子状態の説明を行い, 続いてQGAにおいて次世代の個体集団の決定に用いる1ビット回転ゲートの説明を行う<sup>4)</sup>.

量子計算では, 量子ビットと呼ばれる情報のビットに対応した概念が用いられる. 量子ビットには $|0\rangle, |1\rangle$ の2量子状態が存在し, これらは古典的ビットの0,1に対応する. 量子ビットでは古典的ビットとは異なり $|0\rangle, |1\rangle$ の重ね合わせ状態が存在し, 任意の状態 $|\phi\rangle$ は次式のように表現される.

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1 \quad (1)$$

ここで,  $\alpha$ と $\beta$ は確率振幅と呼ばれる複素数であり, 式(1)は量子状態 $|\phi\rangle$ が確率 $|\alpha|^2$ で状態 $|0\rangle$ へ遷移し, 確率 $|\beta|^2$ で状態 $|1\rangle$ へ遷移することを示している.

式(1)に基づく量子状態 $|\phi\rangle$ の幾何学的表現は次式で示される Bloch 球表示<sup>4)</sup>によって表すことができる.

$$|\phi\rangle = \cos \theta |0\rangle + e^{i\varphi} \sin \theta |1\rangle \quad (2)$$

本報告では, 1ビット回転量子ゲートと2ビット制御 NOT 量子ゲート機能を表現しうる最小の量子状態表現として, 式(2)において $e^{i\varphi} = 1$ とし,  $|0\rangle$ を実軸に,  $|1\rangle$ を虚軸に対応させた次式の複素数表現:

$$|\phi\rangle = \cos \theta + i \cdot \sin \theta = e^{i\theta} \quad (3)$$

を用いる<sup>7)</sup>. ただし,  $i = \sqrt{-1}$ は虚数単位である.

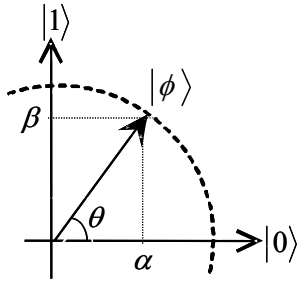


Fig. 1 Complex-valued representation of Quantum State

実部  $\cos \theta$  は  $|0\rangle$  の確率振幅  $\alpha$ , 虚部  $\sin \theta$  は  $|1\rangle$  の確率振幅  $\beta$  に対応する. 式(3)の概念図を Fig. 1 に示す.

1 ビット回転量子ゲートは 1 量子ビットの状態を  $\theta$  だけ回転させる機能を持つゲートであるが, 式(3)の表現によりこれを位相のシフトとして簡単に表現できる. これは 3 節でみるように, QGA では次世代の個体集団を決定する重要な役割を担っている.

### 3 量子ビットを導入した遺伝的アルゴリズム

本節では, 量子ビットを導入した遺伝的アルゴリズム (Qubit-inspired Genetic Algorithm: QGA) の染色体構造およびアルゴリズムについて説明を行う. はじめに, QGA の染色体構造について説明する. 続いて, QGA の具体的なアルゴリズムについて説明する. なお, 本報告の QGA は, 染色体を量子状態の位相として表現している点と, 「個体集団の更新」の処理が文献(6)とは異なる. 染色体を量子状態の位相として表現することにより計算量が少なくなり, 「個体集団の更新」を改良することにより適応度の上昇速度が速くなる.

#### 3.1 QGA の染色体構造

QGA の染色体は確率振幅で表される. 第  $t$  世代において,  $n$  個体で構成される個体集団  $Q(t)$  は次式で表現される.

$$Q(t) = \{q_1^t, q_2^t, \dots, q_n^t\} \quad (4)$$

ここで,  $q_j^t$  は  $j$  番目の個体の染色体を表し, 次式で与えられる.

$$q_j^t = [\theta_1^t | \theta_2^t | \dots | \theta_m^t] \quad (5)$$

なお, 式(5)の  $m$  は染色体の長さを表す. これにより,  $|0\rangle$  の確率振幅  $\alpha = \cos \theta_i^t$ ,  $|1\rangle$  の確率振幅  $\beta = \sin \theta_i^t$  が与えられる.

#### 3.2 QGA のアルゴリズム

QGA は「初期個体集団の生成」, 「個体集団の評価」, 「更新」, 「交叉」, 「突然変異」により実行され, 全体的な処理の流れは GA と同じである. しかし, QGA では染色体が量子ビットで表現されており, また, 「更新」で 1 ビット回転ゲートを用いている. Fig. 2 に QGA のアルゴリズムを示し, 以降の節にて各処理の詳細を説明する.

##### 3.2.1 初期個体集団の生成

QGA の初期個体集団  $Q(0)$  の生成では, 量子状態が  $|0\rangle$  または  $|1\rangle$  にそれぞれ 50% の確率で遷移するように確率振幅を設定する. よって, 式(5)の各遺伝子  $\theta_i^t$  は次式のように初期化される.

$$\theta_i^t = \frac{\pi}{4} \quad (i = 1, 2, \dots, m) \quad (6)$$

```

begin
  t ← 0
  initialize Q(t)
  while (t < Max_Generation) do
    begin
      make P(t) by observing Q(t) states
      evaluate P(t)
      store the best solution b among P(t)
      update Q(t)
      crossover Q(t)
      mutation Q(t)
      t ← t + 1
    end
  end
end

```

Fig. 2 Procedure of QGA

##### 3.2.2 個体の評価

個体の評価では, まず確率振幅で表される  $Q(t)$  を量子力学の確率解釈に基づいて観測し, その結果  $\{0, 1\}$  の 2 値数列からなる  $P(t)$  を得る. この  $P(t)$  に対して適応度関数を適用することにより各個体の適応度を算出し評価を行う. なお,  $P(t)$  は式(7), (8)のように表される.

$$P(t) = \{p_1^t, p_2^t, \dots, p_n^t\} \quad (7)$$

$$p_j^t = \{x_1, x_2, \dots, x_m\}, \quad x_i \in \{0, 1\} \quad (8)$$

量子力学の確率解釈により, 量子状態  $|\phi\rangle$  は確率  $|\alpha|^2 = |\cos \theta_i^t|^2$  で  $|0\rangle$  に遷移し, 確率  $|\beta|^2 = |\sin \theta_i^t|^2$  で  $|1\rangle$  に遷移するので, これを式(5)に適用することにより式(8)が得られる. また,  $P(t)$  の全要素の中で最も適応度の高いものを現時点での最良解  $b$  と定義する.

##### 3.2.3 更新: 個体の選択

従来の GA では, ルーレット選択やトーナメント選択を用いることにより, 適応度の高い個体を次世代に残している. 一方 QGA では, 適応度の高い個体を残すために 1 ビット回転ゲートを用いる. 第  $t+1$  世代の個体集団  $Q(t+1)$  は次式に従い決定する.

$$\theta_i^{t+1} = \theta_i^t + \Delta \theta_i \quad (9)$$

ここで, 式(9)の  $\Delta \theta_i$  は式(10), 式(11)ならびに Table 1 に従って決定する.

$$\Delta \theta_i = s(\theta_i) \delta_i \quad (10)$$

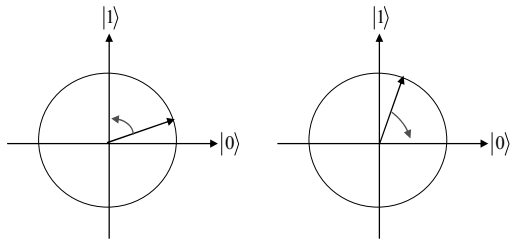
$$\delta_i = 0.01\pi \times (1 - k \times \frac{t}{Max\_Generation + 1}) \quad (11)$$

Table 1 The lookup table for  $s(\theta_i)$

$x_i$	$b_i$	$f(p) < f(b)$	$s(\theta_i)$			
			$\theta_i$ in the first or third quadrant	$\theta_i$ in the second or fourth quadrant	$\theta_i = 0, \pi$	$\theta_i = \pi/2, 3\pi/2$
0	0	true	0	0	0	0
0	0	false	0	0	0	0
0	1	true	1	-1	0	$\pm 1$
0	1	false	-1	1	0	0
1	0	true	0	+1	$\pm 1$	0
1	0	false	-1	0	0	0
1	1	true	0	0	0	0
1	1	false	0	0	0	0

ここで、 $k$ は0~1の実数、Table 1における $f(\cdot)$ は適応度関数である。

Table 1の $s(\theta_i)$ は、3.2.2において定義した現時点での最良解 $b$ と各個体の観測結果 $p$ を比較して決定している。Fig. 3(a)のように $x_i=0, b_i=1$ 、量子状態が第1象限に存在する場合、 $|1\rangle$ に近づくように位相を反時計回りに回転する。このとき、 $s(\theta_i)=1$ とする。また、Fig. 3(b)のように $x_i=1, b_i=0$ 、量子状態が第1象限に存在する場合、 $|0\rangle$ に近づくように位相を時計回りに回転する。このとき、 $s(\theta_i)=-1$ とする。この操作によって、各個体の染色体は現時点で最も高い適応度を持つ個体の染色体に近づいていく。

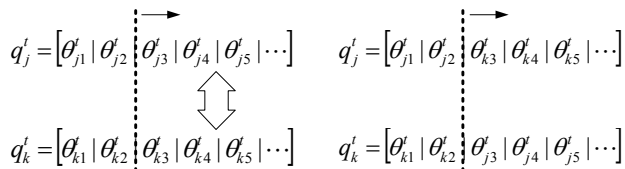


(a)  $x_i=0, b_i=1,$   $\theta_i$  in first quadrant (b)  $x_i=1, b_i=0,$   $\theta_i$  in first quadrant  
Fig. 3 Determination of  $s(\theta_i)$

GAとQGAでは次世代の個体を決定するために適用する操作は異なるが、両者ともに現世代において適応度の高い個体を次世代に残すという点では共通している。

### 3.2.4 交叉

交叉とは、異なる個体間で染色体の一部を入れ替える操作である。Fig. 4に示すように本報告のQGAにおいては一点交叉を採用する。Fig. 4(a)と(b)を比較すると交叉点以降の遺伝子列が入れ替わっていることがわかる。



(a) Before Crossover (b) After Crossover  
Fig. 4 Crossover operation in QGA

### 3.2.5 突然変異

突然変異とは、染色体の一部を変化させる操作のことである。QGAの突然変異は、 $|0\rangle$ と $|1\rangle$ の確率振幅を入れ替える操作とする。本論文では、確率振幅を位相 $\theta$ として表現しているの、突然変異の様子はFig. 5のようになる。

$$q_j^i = [\theta_{j_1} | \theta_{j_2} | \theta_{j_3} | \theta_{j_4} | \theta_{j_5} | \dots] \quad q_j^i = [\theta_{j_1} | \theta_{j_2} | 90 - \theta_{j_3} | \theta_{j_4} | \theta_{j_5} | \dots]$$

(a) Before Mutation (b) After Mutation  
Fig. 5 Mutation operation in QGA

## 4 実験結果

本節では、QGAの性能評価を行う。まず、4.1、4.2節においてQGAの基本的な性能を $N$ - $K$ ランドスケープ問題およびナップサック問題により評価する。続いて、4.3節では環境の変化に対するロバスト性を調べるため関数の最大値探索問題、4.4節では応用例として、2次元図形の位置決定を行う。また、GAとの性能比較を行うことによりQGAの有用性を示す。

### 4.1 $N$ - $K$ ランドスケープ問題

$N$ - $K$ ランドスケープ問題における $M$ は遺伝子の長さを示しており、各遺伝子は1か0の値をとる。その中のある遺伝子はそれ以外の $K$ 個の遺伝子と関係して適応度が決まる。エピスタシスと呼ばれるこの影響のされ方は、遺伝子がより全体として複合的に作用して適応度を決定する指標となっており、 $N$ - $K$ ランドスケープ問題はエピスタシスのある場合の進化過程のモデルとして研究されてきた<sup>8)</sup>。本実験では、 $M$ 個の遺伝子の中で着目する遺伝子とそれに影響を及ぼす $K$ 個の遺伝子に対し、 $M$ 個の遺伝子の値の総和を適応度として、この適応度が最大となる個体を求めている。

#### 4.1.1 実験方法および実験条件

$N$ - $K$ ランドスケープ問題では個体数50、最大世代数1000とし、 $N=100$ に対して $K=2, 6, 10, 14$ の問題について実験を行う。なお、各パターンにおける値は0または1であり、いずれの問題も最大適応度は100となる。

#### 4.1.2 実験結果

Table 2に $N$ - $K$ ランドスケープ問題の実験結果を示す。Table 2のfitnessは1000世代経過後の適応度、min, max, averageは10回の試行における最小値、最大値、平均値を示している。また、各試行における最良個体の適応度の平均と世代数の関係をFig. 6に示す。さらに、個体集団の多様性の有無を調べるために、Fig. 7には個体集団において染色体情報が異なる個体の数(個体の種類数)の変化を示す。Table 2において $K$ が小さい場合の最終的な適応度はGAの方が高いが、 $K$ が大きくなるとQGAの方が高い解が求まっていることがわかる。

また、Fig. 6, 7からGAでは早い段階(数十世代)で多様性が失われ、適応度がほとんど上がらなくなっているが、QGAの方は世代を経ても多様性が維持され、緩やかではあるが適応度が増え続けていることがわかる。

Table 2 Fitness for  $N$ - $K$  landscape problem

n-k	fitness					
	GA			QGA		
	min	max	average	min	max	average
100-2	94	99	96.9	89	95	91.9
100-6	87	95	91	89	98	94.4
100-10	80	86	83.6	83	87	85.3
100-14	76	84	79.4	79	85	83.5

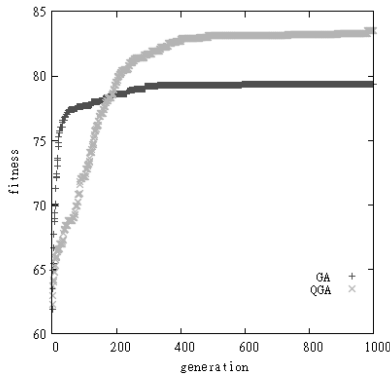


Fig. 6 Profit values over generation for  $N$ - $K$  landscape problem ( $N=100, K=14$ )

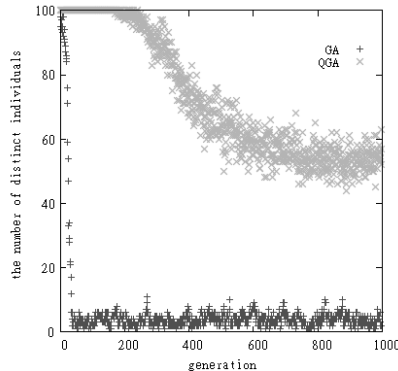


Fig. 7 Transition of the number of distinct individuals over generation for  $N$ - $K$  landscape problem ( $N=100, K=14$ )

## 4.2 ナップサック問題

ナップサック問題とは、 $L_w$ の重さまで荷物を入れることができるナップサックと、重さと価値の決まった  $N$ 個の荷物が存在するとき、ナップサックの重量制限以下で価値の総和が最大となる荷物の組み合わせを求める問題である。

### 4.2.1 実験方法および実験条件

ナップサック問題では、GA, QGAともに個体数100, 最大世代数1000, 交叉率0.9, 突然変異率0.01とし、初期個体を変化させ10回試行し、適応度と個体の多様性が世代を経るにつれてどのように変化していくかを調べる。本実験では、ナップサックの重量制限は全荷物の総重量の1/2とする。また、適応度はナップサックに詰め込むことができた荷物の価値の総和とする。ただし、ナップサックの重量制限を超えた個体の適応度は0とする。

### 4.2.2 実験結果

Table 3 にナップサック問題の実験結果を示す。Table3の  $N$ はナップサック問題の荷物の種類数を表しており、profitは1000世代経過後の適応度、min, max, averageは10回の試行における最小値、最大値、平均値を示している。また、Fig. 8に適応度の推移のグラフを示し、Fig. 9に多様性の推移のグラフを示す。Table 3より、QGAの方がGAよりも適応度(価値)の高い解

Table 3 Profit values for KP(10 trials)

number of items (N)	profit					
	GA			QGA		
	min	max	average	min	max	average
100	33775	35205	34606	38865	38958	38926
200	63195	71114	67031	79526	79725	79614
300	93291	105347	98768	121521	122960	122270
400	121978	132236	126400	155771	158088	156589
500	146609	163110	155831	193882	197631	196043

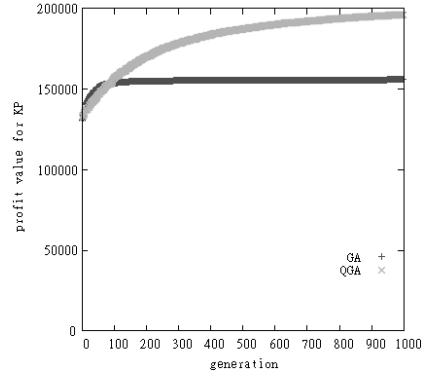


Fig. 8 Profit value over generation for KP with  $N=500$

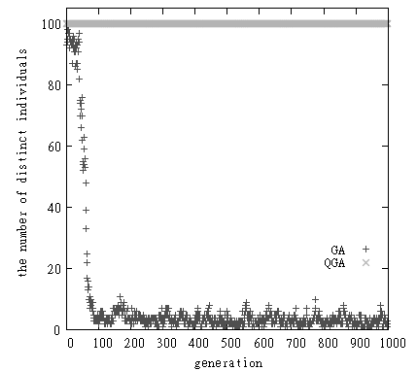


Fig. 9 Transition of the number of distinct individuals over generation for KP with  $N=500$

が求まっていることがわかる。Fig. 8からはQGAの方が適応度の上昇速度が速いことがわかる。また、Fig. 9から、GAでは世代を経るに従って個体の種類数が減少していく、すなわち、個体の多様性が失われていくが、QGAでは1000世代経過後も個体の種類数は0世代目とほぼ変わっておらず、世代を経ても個体の多様性を維持できていることがわかる。

## 4.3 関数の最大値探索問題

本節では、世代交代の途中で環境が変化した場合のQGAの性能について調べるために最大値探索問題にQGAを適用し、その性能を調べる。

### 4.3.1 実験方法及び実験条件

最大値探索問題では、GA, QGAともに個体数20, 最大世代数400とし、第100世代目までは式(12)の関数  $f_1(x)$  (Fig. 10), 第101世代目以降は式(13)の関数  $f_2(x)$  (Fig. 11)の最大値を求める。 $f_1(x)$ は  $x=0.82$ のとき最大、 $f_2(x)$ は  $x=2.07$ のとき最大となる。この問題では、位置  $x$  における関数  $f_1(x)$  および  $f_2(x)$  の値

を適応度として適応度の算出を行う。

$$f_1(x) = \cos x - \frac{1}{2} \sin 3x + \sin 10x + \cos 15x + 3 \quad (12)$$

$$f_2(x) = \cos 3x - \frac{1}{2} \sin 3x + \sin 10x + \cos 15x + 3 \quad (13)$$

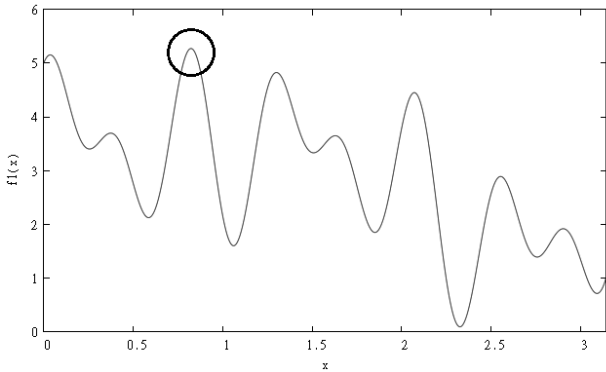


Fig. 10 Target function  $f_1(x)$  in 0~100 generations

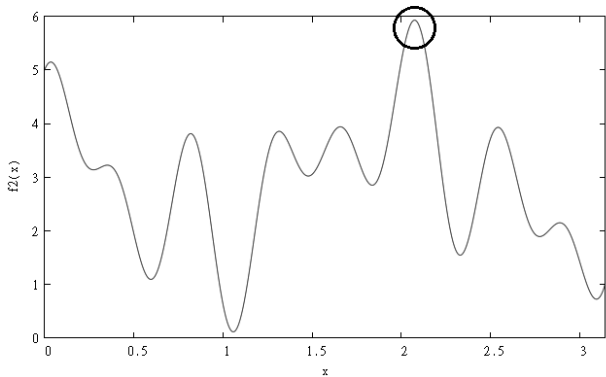


Fig. 11 Target function  $f_2(x)$  in 101~400 generations

### 4.3.2 実験結果

Fig. 12 に適応度( $f_1(x)$ ,  $f_2(x)$ の値)の変化のグラフ, Fig. 13 に多様性の変化のグラフを示す. Fig. 12 より, GA により求めた最大値は 5.07, QGA により求めた最大値は 5.27 であり,  $f_1(x)$  の真の最大値は 5.27 であるので QGA は最大値を正確に得ている. また, 101 世代目以降に注目すると, GA では  $f_2(x)$  の最大値が求まっていないのに対して, QGA では  $f_2(x)$  の最大値が求まっていることがわかる. また, Fig. 13 から, GA では  $f_1(x)$  の最大値が存在する  $x=0.8$  に近づくように個体の淘汰が行われ, 個体の多様性が失われていくことがわかる. このため, 101 世代目以降において,  $f_1(x)$  と最大値の位置が異なる  $f_2(x)$  の最大値が求まらなかったと考えられる. また,  $f_1(x)$  の  $x=0.8$  付近が極値となっていたため, GA では極値を最大値と誤認したと考えられる. 一方, QGA では  $f_1(x)$  の最大値が存在する  $x=0.8$  に近づくように世代交代が行われているが, GA とは異なり個体の多様性が維持できていたために,  $f_1(x)$ ,  $f_2(x)$  両方の最大値が求められたと考えられる.

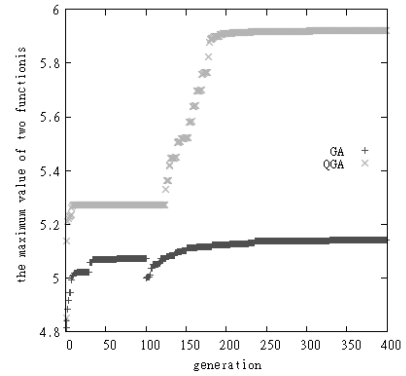


Fig. 12 Fitness value over generation for Maximum Search Problem

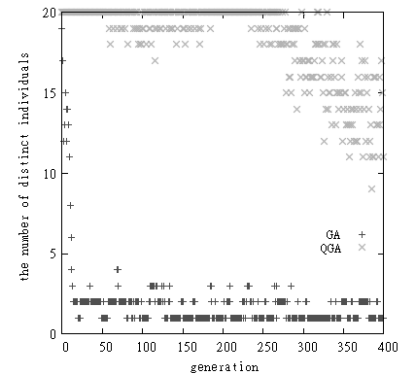


Fig. 13 Transition of the number of distinct individuals over generation for Maximum Search Problem

### 4.4 2次元図形の位置決定

本節では, 実用的な応用例として2次元図形の位置決定を行う<sup>9)</sup>.

2次元図形の位置決定では, 位置( $x, y$ ), 大きさ( $m$ ), 画像に対して垂直な軸に関する回転角度( $angle$ )が未知であるモデル図形を原画像から検出することを目的とする. モデル図形は次式のように  $n$  個の点で表されているものとする.

$$P = \{p_1(x_1, y_1), p_2(x_2, y_2), \dots, p_n(x_n, y_n)\} \quad (14)$$

#### 4.4.1 実験方法および実験条件

各個体の染色体はFig. 14のように求める4つのパラメータ  $x, y, m, angle$  で構成される.

この問題の適応度は, モデル図形の重心を中心として  $angle$  だけ回転させてから  $m$  倍に拡大し, 図形の重心を原図形の座標( $x, y$ )に移動させたときの一致率であり,  $n_{total}$  を一致した画素の総数であるとする, 適応度は次式で表される.

$$\frac{01101101}{x} \quad \frac{10010100}{y} \quad \frac{10101}{m} \quad \frac{10110101}{angle}$$

Fig. 14 Chromosome in the problem of positioning 2D figures

$$fitness = \frac{n_{total}}{n} \quad (14)$$

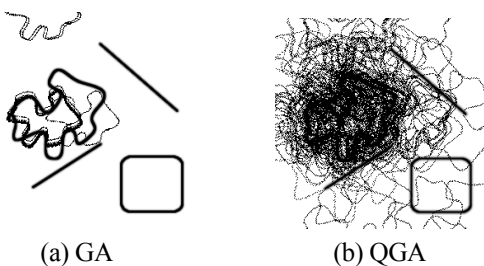
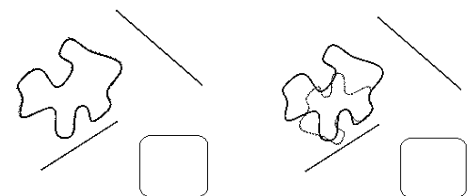
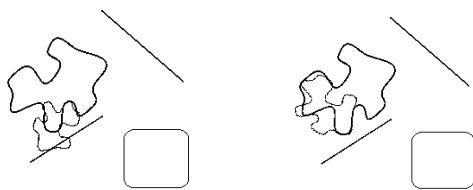
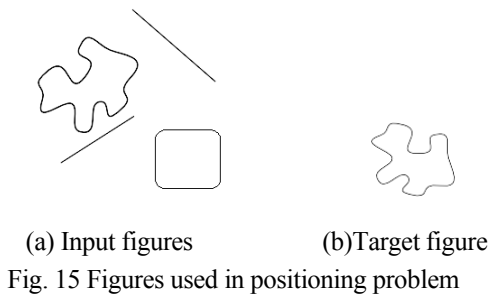
式(15)よりモデル図形が完全に一致した場合、適応度の最大値は1となる。

本実験では個体数100, 最大世代数500, 染色体長29とする。

#### 4.4.2 実験結果

原画像(Fig. 15(a))からモデル図形(Fig. 15(b))を抽出する。適応度85%以上で検出成功とすると、試行回数10回のうちGAの0回に対してQGAは6回検出することに成功した。この場合のGAおよびQGAの最終的な検出結果例をFig. 16および17に示す。

GAでは、Fig. 16に示した例のように、モデル図形を正しく検出することができなく、50%以上の一致率を得ることができなかった。一方、QGAでは、Fig. 17



左図に示した例のように検出に成功している。また、Fig. 18に示した200世代における探索の分布例からわかるように、GAではほぼ同じ所(a)を探索しており、多様性が失われているが、QGAではGAに比べ広範囲(b)を探索しているため、多様性が失われていないことがわかる。

## 5 結論

本論文では、 $N$ - $K$  ランドスケープ問題、ナップサック問題、最大値探索問題により QGA の基本的な性能を調査した。まず、基本問題例として取り扱った  $N$ - $K$  ランドスケープ問題、ナップサック問題により QGA が GA よりも高い適応度を得ることができるということがわかった。次に、環境が変化した場合のロバスト性を調べるために行った最大値探索問題では、最大値を求める関数を世代交代の途中で変更した場合、GA では変更後の関数の最大値を求めることができなかったが、QGA では両方の関数の最大値を求めることができた。最後に 応用問題として取り扱った 2 次元図形の位置決定問題では GA では検出できなかったモデル図形を QGA では求めることができた。これにより、QGA の実用的な問題における有用性を示すことができた。これらの数値実験を通して、QGA の性能の良さが個体の多様性の高い維持に起因していることも示唆できたが、QGA が GA を真に凌駕しうる適用を見出すことおよび精査は今後の課題である。

### 謝辞

本研究の一部は科研費(基盤研究(C) 23500286)の助成を受けたものである。

### 参考文献

- 1) Fogel L.J. et al., "Artificial Intelligence through Simulated Evolution," John Wiley & Sons (1966)
- 2) Bäck, T., Hoffmeister, F. and Schwefel, H.-P., "A Survey of Evolution Strategies," Proc. 4<sup>th</sup> Int. Conf. on Genetic Algorithms, 2/9 (1991)
- 3) Holland, J.H., "Adaptation in natural and artificial systems," University of Michigan Press (1975)
- 4) Michal A. Nielsen and Isaac Chuang, "Quantum Computation and Quantum Information", Cambridge University Press (2000)
- 5) Narayanan, A. and Moore, M., "Quantum-inspired genetic algorithms," Proc. IEEE Int. Conf. Evolutionary Computation, 61/66 (1996)
- 6) Han, K.H. and Kim, J.H., "Genetic quantum algorithm and its application to combinatorial optimization problem," Proc. IEEE Congr. on Evolutionary Computation, 1354/1360 (2000)
- 7) Takata, T., Isokawa, T. and Matsui, N., "Performance Analysis of Quantum-Inspired Evolutionary Algorithm," Journal of Advanced Computational Intelligence and Intelligent Informatics, 15-8 (2011, to be published)
- 8) 伊庭斉志, "遺伝的アルゴリズムと進化のメカニズム", 13/21, 岩波書店 (2002)
- 9) 長尾智晴, "進化的画像処理", 48/57, 昭晃堂 (2002)