

## 聴覚情報処理のための競合学習パルスニューラルネットワーク

### CONP-II

黒柳 奨<sup>†a)</sup> 岩田 彰<sup>†</sup>

A Competitive Learning Pulsed Neural Network for Auditory Information Processing CONP-II

Susumu KUROYANAGI<sup>†a)</sup> and Akira IWATA<sup>†</sup>

あらまし 我々はパルスニューラルネットワークを用いた聴覚情報処理システムによる、周辺環境把握システムの構築に関する研究を行っており、これまでに入力信号の時間的な変化に追従して勝者ニューロンを決定できる競合学習ネットワーク CONP を提案している。実環境に CONP を適用することを考えた場合、CONP は入力パターンベクトルをそのまま量子化するため、複数の音源から音が発せられた場合にはその混在パターンを学習してしまうという問題点がある。そこで、本論文では CONP の改良を行い、残響成分等の影響を受けずに複数音源から発せられた信号の分離学習が可能な新しい競合学習ネットワーク CONP-II を提案する。計算機シミュレーションの結果、音の立上りをずらした複数の音を混合した信号に対して CONP-II を用いた場合、それぞれの音を分離してベクトル量子化が可能であることが確認できた。

キーワード パルスニューラルネットワーク、競合学習、ベクトル量子化、聴覚情報処理、混合音分離

### 1. ま え が き

我々は危険物体（接近する自動車など）の回避や、目標捕そく（子供の位置確認、携帯電話の呼出し）などの行動を行うために音による周辺環境把握を行っており、この機能を工学的に応用することは聴覚障害者の補助、自律移動ロボットの環境把握など様々な用途に非常に有益である。このような生体の神経情報処理を工学的に応用する方法として、ニューラルネットワークモデルを用いる研究が盛んに行われている。その中でも聴覚情報処理のような時系列信号の処理に適したモデルとしてパルスニューロン（Pulsed Neuron；以下 PN）モデルが注目されている[1]。PN モデルは入出力信号としてパルス列を用いるニューロンモデルであり、一般に用いられているマカロックピッツ（McCulloch Pitts）タイプのニューロンモデル[2]と比較した場合、モデル内にダイナミックスをもつのが

大きな特徴となっている。また、各 PN モデルは簡易な演算のみから構成されているため、デジタル回路として容易に、かつ少ないリソースで実装が可能であり[3]、更に各 PN モデルは独立非同期に動作可能なため、ニューラルネットワークのハードウェア実装手法として適している。我々は、この PN モデルによって構成されるニューラルネットワーク、パルスニューラルネットワーク（Pulsed Neural Network；以下 PNN）を用いた聴覚情報処理システムによる周辺環境把握システムの構築に関する研究を行っている[4]、[5]。本システムでは初段に設置された音-パルス変換器以降の処理を、すべて PN モデル群で処理することを目指しており、大量の PN モデルを FPGA、LSI 等の集積回路に実装し、これらを並列動作させることで非常に高速なニューラルネットワークを低コストで実現できると考える。本システムに関して我々はこれまでに、音-パルス列変換モデル並びに水平面の音源定位モデル[5]、入力信号のスペクトルなどの特徴量パターンを教師あり学習で獲得するモデル[6]などを実現しており、また競合学習により入力信号の特徴パターンを教師なし学習で獲得するモデル[7]（以降、本

<sup>†</sup> 名古屋工業大学大学院情報工学専攻，名古屋市  
Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya-shi, 466-8555 Japan  
a) E-mail: bw@nitech.ac.jp

モデルを Competitive learning Neural network using Pulsed neurons; 略称 CONP とする) を提案している. CONP は Kohonen による競合学習モデル並びに自己組織化マップ (Self-Organizing Maps; 以下 SOM)[8] を PNN に適用したものであり, 入力パターンが時間的に変化する場合にも, パターンの変化に追従して勝者ニューロンを求めることが可能である.

音による周辺環境把握の基本機能は音源定位と音源認識である. 我々の提案システムにおいて, 音源定位については, 二つの入力信号間の時間差情報を抽出した発火パターンを CONP により低次元に位相情報を保持したまま写像することで, 自動的に入力信号の方位をマッピングできることを確認しており [9], また音源認識についても, CONP により稼働環境における代表的な周波数スペクトルパターンを統計的に学習できることが期待できる.

今, 実環境に CONP を適用することを考えた場合, 入力音として複数の音源からの音が混在していることを前提とする必要がある. また, 一般に音には残響成分または持続成分が存在し, CONP は SOM と同様に入力パターンベクトルをそのまま量子化するため, 複数の音源から発せられた各音に残響などがある場合にはその混在パターンを学習してしまうという問題点がある. そこで, 本論文では競合学習ネットワーク CONP の改良を行い, 残響成分等の影響を受けずに複数音源から発せられた音の分離学習が可能な新しい競合学習ネットワークの構築を行う. なお, 聴覚情報処理で用いられている情報としては, 入力信号の瞬時的スペクトルパターンのような静的な情報だけではなく, スペクトルパターンの時間変化のような時系列情報も非常に重要である. 我々の提案するシステムでは, まず入力信号の極小時間ごとの特徴量を抽出し, 次にこれら量子化されたパターンの時間変化を処理する方法を検討しており [7], CONP はこのうちのベクトル量子化部分に相当するネットワークである.

以降, まず 2. において CONP について説明を行う. 次に 3. において混合音分離学習手法について検討し, 新しい競合学習ネットワーク CONP-II を提案する. 最後に 4. において計算機上のシミュレーションにより提案手法の有効性を検証する.

## 2. パルスニューロンモデルのための競合学習ネットワーク CONP

本章では, 本論文で提案するネットワークモデルの

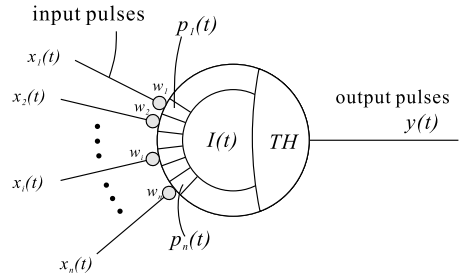


図1 パルスニューロンモデル  
Fig. 1 A pulsed neuron model.

もとなる PN モデル及び我々の提案する競合学習ネットワーク CONP について述べる.

### 2.1 パルスニューロンモデル

PN モデルの模式図を図 1 に示す. この図において,  $i$  番目の入力チャネルからパルス  $x_i(t) = 1$  が到着すると, その部分の局所膜電位  $p_i(t)$  が結合重み  $w_i$  分上昇し, その後時定数  $\tau$  で静止電位まで減衰する. PN モデルの内部電位  $I(t)$  はその時刻の各局所膜電位の総和として表される. ニューロンはこの内部電位がしきい値  $TH$  を超えたとき発火 (出力パルスを発生) する. ただし神経細胞には発火に関する不応期  $RP$  が存在するため, 本モデルにおいてもある発火から  $RP$  の間は内部電位がしきい値を超えた場合でも発火はしない. 本 PN モデルの動作を明確にするため, 式により以下のように定義する.

今システムをサンプリング周波数  $F_s$  の離散時間系とし,  $\Delta t = 1/F_s$  とする. ここで引数として時刻  $t$ , 減衰時定数  $\tau$ , 結合重み  $w$ , 時刻  $t$  における入力信号  $x(t)$  の四つをもつ関数  $F$  を導入し,

$$\begin{aligned} F(0, \tau, w, x(0)) &= 0 \\ F(t, \tau, w, x(t)) &= F(t - \Delta t, \tau, w, x(t - \Delta t)) \cdot e^{-\frac{\Delta t}{\tau}} + w \cdot x(t) \end{aligned} \quad (1)$$

とする.

このとき, 時刻  $t$  における PN モデルの内部電位  $I(t)$  は局所膜電位  $p_i(t)$  の総和として次式のように記述できる.

$$I(t) = \sum_{i=1}^n p_i(t) \quad (2)$$

$$= \sum_{i=1}^n F(t, \tau, w_i, x_i(t)) \quad (3)$$

ここで、 $\tau$  は  $p_i(t)$  の減衰時定数である．PN モデルの不应期を  $RP$ ， $ET(t)$  を時刻  $t$  における前回発火からの経過時間とし、 $ET(0) > RP$  とすると、PN モデルの出力値  $y(t)$  は以下のアルゴリズムにより計算される．

```

if  $I(t) \geq TH$  and  $ET(t) > RP$ 
  then  $y(t) = 1, ET(t) = 0$ 
  else  $y(t) = 0, ET(t) = E(t - \Delta t) + \Delta t$ 
    
```

パラメータ  $\tau, \mathbf{w} = (w_1, w_2, \dots, w_n)$ ， $TH$  は各ニューロンにより可変の値であり、この組合せにより PN モデルの動作は決定される．

PN モデルは各ニューロン間のデータ転送路が単線であるためハードウェア上に複数のニューロンモデルを実装する際に非常に有利である．また、伝達する信号が頻度変調されたパルスのストリームであるため、各ニューロンは基本的に非同期独立に演算が可能であり、多数のニューロンを実装した場合にも厳密な同期処理を行う必要はないという特徴をもつ．

### 2.2 競合学習ネットワーク CONP

CONP は Kohonen の競合学習則 [8] を PN モデルへ適用したものである．CONP は入力信号の時間変化に追従して勝者ニューロンが自動的に変化するという特徴をもつ．以下にその動作アルゴリズムについて述べるが、紙面の都合上詳細については文献 [7] を参照されたい．

CONP では各ニューロンにおける入力ベクトルとの類似度の評価値としてニューロンの内部電位  $I(t)$  を用いる．

競合学習における勝者ニューロンを決定するためには一般に最も評価値の高いニューロンを検索するが、CONP においては各ニューロンのしきい値を同時に同じ量だけ変化させることにより、最も高い評価値をもつニューロンすなわち勝者ニューロンのみが発火するようにする．図 2 に CONP の模式図を示す．CONP では各ニューロンのしきい値を制御するために競合学習ニューロンが一つも発火していないときに発火する「無発火検出 (No-Firing Detection: 以下 NFD) ニューロン」と、競合学習ニューロンが二つ以上発火しているときに発火する「複数発火検出 (Multi-Firing Detection: 以下 MFD) ニューロン」の二つを用いる．これら状態検出ニューロンは図 1 で示した PN モデルと同じものである．CONP はこれら状態検出ニューロンの出力パルスに応じて各競合学習ニューロンのし

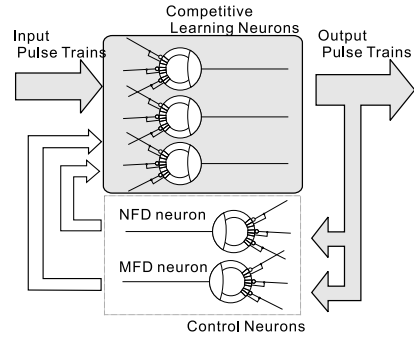


図 2 競合学習ニューラルネットワーク CONP  
Fig. 2 The competitive learning neural network CONP.

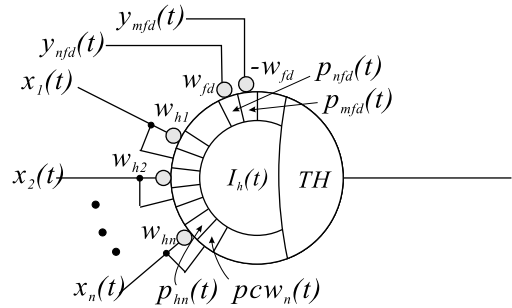


図 3 競合学習用パルスニューロンモデル  
Fig. 3 The pulsed neuron model for competitive learning.

きい値を一律に変化させることで唯一のニューロンが発火する状況を保持する．なお、CONP は状態検出ニューロンによるフィードバック制御を行っているため、時間的な入力パターンの変化に対する追従性には限界があるが、文献 [7] における実験結果より、PN モデルの不应期  $RP$  の 10 倍程度の時間分解能を有することが確認されており、実用上は十分な性能をもつと考えられる．

CONP ではネットワークにおいて発火した競合学習ニューロンを勝者ニューロンと考えるため、各ニューロンが発火した場合に学習を行う．学習すべき入力パターンの表現法としては、結合重みを 1 に固定したシナプスにおける局所膜電位  $pcw(t)$  を用いる．図 3 に図 1 に対して学習に必要な要素を加えた競合学習用パルスニューロンモデルを示す．ここで入力パルス列数は  $n$  であり、このニューロンは  $M$  個ある競合学習ニューロンの  $h$  番目とする．この図において、時刻  $t$  における NFD, MFD ニューロンの出力をそれぞれ  $y_{nfd}(t)$ ， $y_{mfd}(t)$ 、競合学習ニューロンの NFD, MFD ニューロンに対する結合重みを  $w_{fda}$ ， $-w_{fda}$  ( $w_{fda} > 0$ ) とし、

時刻  $t$  における  $h$  番目の競合学習ニューロンの内部電位  $I_h(t)$  は前述の関数  $F$  を用いて次式のように記述できる。

$$I_h(t) = \sum_{i=1}^n p_{hi}(t) + p_{nfd}(t) + p_{mfd}(t) \quad (4)$$

$$p_{hi}(t) = F(t, \tau, w_{hi}, x_i(t)) \quad (5)$$

$$p_{nfd}(t) = F(t, \tau_{fd}, w_{fd}, y_{nfd}(t)) \quad (6)$$

$$p_{mfd}(t) = F(t, \tau_{fd}, -w_{fd}, y_{mfd}(t)) \quad (7)$$

なお、CONP においては  $p_{nfd}$ 、 $p_{mfd}$  を発火しきい値の動的変化量として扱うことで制御を行うため、減衰時定数  $\tau_{fd}$  は時定数  $\tau$  に対して十分大きいものと仮定する。時刻  $u$  における勝者ニューロンの結合重み  $w_{win,i}$  の更新は、学習係数を  $\alpha$  としたとき次式で表すことができる。

$$w_{win,i}(u) = w_{win,i}(u - \Delta t) + \alpha \cdot pcw_i(u) \quad (8)$$

$$(1 \leq i \leq n, 0 \leq \alpha)$$

$$pcw_i(u) = F(u, \tau, 1, x_i(u)) \quad (9)$$

各更新の後、結合重みベクトル  $w(u)$  はノルムが 1 となるよう正規化される。なお、ノルムを 1 に正規化するためには一般に二乗計算と除算が必要であり、これらは PN モデルにおける演算として高負荷であり望ましいものではない。これについては、FPGA 等のハードウェア上への CONP の実装する際に、近似的な正規化手法を用いたり、または同時に学習が行われるニューロン数がたかだか 1 であることからハードウェア上に正規化演算ユニットを一つだけ準備するなどの対応により適応可能であると考えられる。これについては CONP のハードウェア上への実装を含め今後の検討課題としたい。

入力パルス列によって発生する内部電位の総量が大きく変動する場合、この変動量を吸収するためにしきい値の変化が生じることになり、しきい値の変化が入力ベクトルの方向変化に追従できない場合がある。そこで、CONP では内部電位  $I(t)$  に対して、 $pcw$  の総和を一定の比率  $\beta_{pcw}$  ( $0 \leq \beta_{pcw} \leq 1$ ) であらかじめ差引くことで、入力信号のノルム変動に対する内部電位の変化を抑制している。これにより式 (4) の  $I_h(t)$  は以下のように修正される。

$$I_h(t) = \sum_{i=1}^n p_{hi}(t) + p_{nfd}(t) + p_{mfd}(t) - S_{pcw}(t) \quad (10)$$

$$S_{pcw}(t) = \beta_{pcw} \sum_{i=1}^n pcw_i(t) \quad (11)$$

以上のアルゴリズムにより PNN において Kohonen の競合学習を実現することが可能であり、これにより入力信号に含まれるスペクトルパターンなどが時々刻々と変化する場合でも、これを統計的に学習、ベクトル量子化することが可能である。なお、これまで述べてきた結合重みの更新を勝者ニューロンの近傍のニューロンに対しても適用することにより、SOM アルゴリズムは容易に実現可能である。4. での実験においても SOM アルゴリズムを用いている。この実験では末端のないリング状に競合学習ニューロンが結合しているとし、 $h$  番目のニューロンと  $h-1$  番目及び  $h+1$  番目のニューロン、そして 1 番目と  $M$  番目のニューロンが隣接関係をもつ。このとき学習半径は学習回数により  $M$  より linear に減少し、 $EN_0$  回の学習で 0 になるものとした。

### 3. 複数音源混合信号の分離学習のための競合学習ネットワーク

#### 3.1 混合音の分離手法

実環境に CONP を適用することを考えた場合、単一の音源から発せられる音のみが存在する状況は極めてまれであり、入力音として複数の音源からの音が混在していることを前提とする必要がある。複数の音源が存在する場合であっても、各音の発生時間が時間軸上で異なっていれば CONP を用いることでそれらの音を独立してベクトル量子化することが可能であるが、一般に音には残響成分または持続成分が存在する。CONP は SOM と同様に、入力パターンベクトルをそのまま量子化するため、複数の音源から発せられた各音に残響などがある場合には、ある時刻に発生した音のパターンと、それ以前に発生した音の残響などの混在パターンを学習してしまうという問題点がある。このように複数の音源から発せられた音を分離する手法としては、独立成分解析を用いる方法 [10], [11] がある。この手法は  $N$  個のセンサで受信した信号を行列計算により別の  $N$  個の信号に変換することを基本としており、変換後の各信号の独立性がなるべく大きくなるような変換行列を学習により求めることで分離を実現する。この方法は分離後の信号をそのまま取り出すことが可能であるという特徴があるが、 $N$  個の信号を分離するためには  $N$  個の混合信号が必要であり、

分離される信号の数はセンサの個数によって限定されるという欠点がある。

本論文では、分離する入力信号を聴覚系にて受信される音信号に限定し、分離した出力をベクトル量子化によりパターン認識された結果とすることにより、CONPの改良版ネットワークによって混合音を分離しながら競合学習を行う方法を提案する。

なお、SOMによって複数の音を分離する手法としては、Gautamaら[12]により、二つのSOMネットワークをフィードバック結合したネットワークを用いる手法が提案されている。この方法では、混合音より作成した信号をそれぞれのSOMネットワークに入力し、このネットワークの出力結果より、量子化された入力信号を復元する。そしてこの復元信号をもう一方のSOMネットワークの入力から引き算することにより、ネットワークの動作後にはそれぞれのSOMネットワークに別々の入力信号に対応する出力が現れるというものである。しかし、この手法ではあらかじめ分離したい信号をそれぞれ独立して学習しておく必要があり、新規の混合信号を分離しながらベクトル量子化することは想定していない。

今、音を立上り成分と残響成分に分け、かつ音による環境把握システムの実現を考えた場合、入力音の周波数スペクトルパターンが一定の場合は、その残響については除去をしても差し支えない。また、スペクトルパターンが途中で変化する場合、変化した瞬間に新たなスペクトルパターンの立上りが起こるため、やはり残響成分については除去をしてもその影響は少ないと考えることができる。そこで、本論文では競合学習ネットワークCONPの改良を行い、残響成分などの影響を除去して複数音源から発せられた混合音の分離学習が可能な新しい競合学習ネットワークの構築を行う。

この残響成分を除去して競合学習を行うという手法を言い換えると、入力音信号の各周波数成分の立上り時刻のほぼ同じものをグルーピングし、これを一つの信号として競合学習を行うということができる。この手法の問題点として、ほぼ同時刻に複数の音源から音が発生する場合には、音の立上り成分もほぼ同時刻に発生するため分離が困難であることが想定できるが、人間の聴覚系においても、単一センサからの同時発生音を別の音として分離することは困難であり、これらは一つの音源から発生した別種の音として認識されると考えてよいであろう。

我々の提案している競合学習ネットワークCONPは各時刻の入力ベクトルを競合学習によりベクトル量子化することが可能であり、入力信号の変化に追従して自動的に勝者ニューロンを変化させることが可能である。よってこのCONPの入力に対して、各周波数成分の立上りから特定期間のみに特異的に反応するアクティブフィルタを適用すれば、ほぼ同時刻に立ち上がった周波数成分のみを次々にベクトル量子化、すなわちグルーピングすることが可能となる。つまり、

(1) 各周波数成分の立上りに反応して、立上りからある特定の期間だけゲートを開くアクティブフィルタ

(2) CONPの競合学習によるグルーピング効果を組み合わせることで複数音源からの混合信号を分離し、競合学習することが可能である。本手法で用いるPNモデルは内部にダイナミクスをもつため、特別な機構を新たに導入することなく動的なフィルタリングを実装可能である。

### 3.2 混合音分離競合学習ネットワークCONP-II

以上の議論に基づき、混合音を分離し競合学習するパルスニューラルネットワークCONP-IIを提案する。図4に提案手法の模式図を示す。この図においては入力信号は $n$ チャンネルの時系列パルス信号である。CONP-IIは各入力チャンネルの立上りを検出するONニューロンをCONPに追加した構造をもち、各競合学習ニューロンには $n$ チャンネルの入力パルス信号と、ONニューロンの出力パルス信号の合計 $2n$ チャンネルが入力される。競合学習ニューロンの数は $M$ 個とする。

ONニューロンは時間 $T_d$ の遅れ素子を用いることによりPNモデルで実現可能である。今、 $i$ 番目の入力チャンネル $x_i$ に対するONニューロンの時刻 $t$ の内

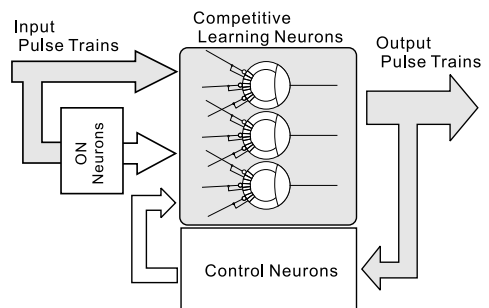


図4 競合学習ニューラルネットワークCONP-II  
Fig. 4 The competitive learning neural network CONP-II.

部電位を  $I_i^{ON}(t)$  とすると

$$I_i^{ON}(t) = F(t, \tau^{ON}, w^{ON}, x_i(t)) - F(t - T_d, \tau^{ON}, w^{ON}, x_i(t - T_d)) \quad (12)$$

と記述できる．ここで， $\tau^{ON}$ ， $w^{ON}$  は ON ニューロンの減衰時定数及び結合重みで，いずれのチャンネルにおいても同じ値を用いる．ON ニューロンはこの  $I_i^{ON}$  によりしきい値  $TH^{ON}$  に従い出力  $y_i^{ON}(t)$  を生成する．

競合学習ニューロンの入力信号において上述のアクティブフィルタを実現するためには，ON ニューロンが発火した時刻より一定時間だけ非零となるようなフィルタ値を作成し，この値と入力パルス列による局所膜電位をかけ合わせたものを各競合学習ニューロンの局所膜電位とすればよい．また，このフィルタ値は ON ニューロンの出力パルスを入力とする局所膜電位で実現可能である．今，COMP-II における  $h$  番目 ( $h = 1 \dots M$ ) の競合学習ニューロンの  $i$  番目 ( $i = 1 \dots n$ ) のチャンネルにおけるフィルタ値を  $AF_{hi}^*(t)$  とすると

$$AF_{hi}^*(t) = F(t, \tau^{AF}, 1, y_i^{ON}(t)) \quad (13)$$

と記述できる． $\tau^{AF}$  はフィルタ値の減衰時定数である．

よって競合ニューロンにおける結合重みを  $w'_{hi}$ ，局所膜電位の減衰時定数を  $\tau$ ，入力信号を  $x_i(t)$  とすると求める局所膜電位  $p_{hi}^*(t)$  は

$$p_{hi}^*(t) = AF_{hi}^*(t) \cdot F(t, \tau, w'_{hi}, x_i(t)) \quad (14)$$

と記述できる．しかし，この演算のためには PN モデル内において乗算を実装する必要があり，PN モデルの演算コストが大幅に増加してしまう．そこで，結合重み  $w'_{hi}$  が固定値であることに着目し，この値を ON ニューロンの出力値により動的に変化させることで上記膜電位の演算を実装することとした．ON ニューロンの出力値  $y_i^{ON}(t)$  を入力とし，結合重み =  $w'_{hi}$ ，時定数 =  $\tau^{AF}$  で変化する局所膜電位を改めて  $AF_{hi}(t)$  とする．

$$AF_{hi}(t) = F(t, \tau^{AF}, w'_{hi}, y_i^{ON}(t)) \quad (15)$$

この  $AF_{hi}(t)$  を競合学習ニューロンの  $i$  番目の入力に対する結合重みとして用いることで，所望する局所膜電位を得ることができる．よって時刻  $t$  における  $h$  番目の CONP2 における競合学習ニューロンの内部電位

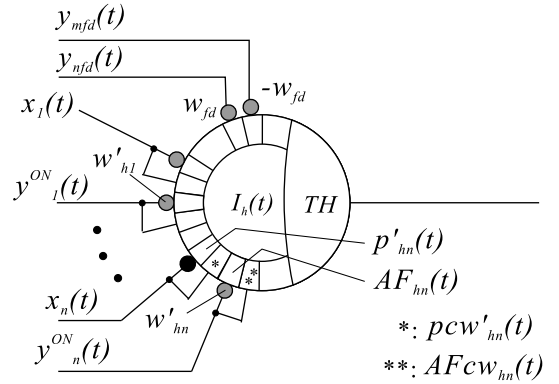


図5 CONP-II における競合学習パルスニューロンモデル  
Fig. 5 The competitive learning pulsed neuron model for CONP-II.

$I_h^{CONP2}(t)$  は次式のように記述できる．

$$I_h^{CONP2}(t) = \sum_{i=1}^n p'_{hi}(t) + p_{nfd}(t) + p_{mfd}(t) - S_{pcw}(t) \quad (16)$$

$$p'_{hi}(t) = F(t, \tau, AF_{hi}(t), x_i(t)) \quad (17)$$

図5に提案手法を実装した CONP-II 用の競合学習ニューロンの模式図を示す．この図において，局所膜電位  $AF_{hn}(t)$  の値が局所膜電位  $p'_{hn}(t)$  の結合重みの値（図中黒点で示す）となる．

NFD，MFD の各状態検出ニューロンも CONP と同じものがそのまま流用可能であり，しきい値の制御方式も CONP と同様である．学習に関して CONP と同様に発火した競合学習ニューロンを勝者ニューロンと考え，学習が行われる．ただし，学習に用いる局所膜電位は式 (8) の  $pcw_i(u)$  の代わりに下記の  $pcw'_i(u)$  を用いる．

$$pcw'_i(u) = F(u, \tau, AFcw_i(t), x_i(u)) \quad (18)$$

$$AFcw_i(u) = F(u, \tau^{AF}, 1, y_i^{ON}(u)) \quad (19)$$

以上の手法により，ほぼ同時刻に立ち上がった周波数成分のみを次々にベクトル量子化する競合学習ネットワークが実現できる．この CONP-II のための競合学習ニューロンを従来の競合学習ニューロンと比較した場合，内部で保持すべき局所膜電位の数が倍に増えてはいるが，それ以外の構造は同じものであり，従来の PN モデルを用いて実装が可能であることが分かる．また，CONP-II で新たに導入された ON ニューロンも同様に従来の PN モデルで実装が可能であり，

CONP-II は CONP 同様, PN モデルから構成されたネットワークモデルとして実現可能なことが分かる .

#### 4. 計算機シミュレーションによる提案手法の検証

本章では計算機上のシミュレーションにより CONP-II の動作を検証する . なお, ニューロンへの入力信号が十分に小さい場合において発火しきい値が下がりすぎることを防ぐため, CONP においては  $pcw_i(u)$ , CONP-II においては  $pcw'_i(u)$  の総和があるしきい値  $TH_g$  よりも小さくなった場合には, 内部ポテンシャルにかかわらず発火を停止し, また制御ニューロンからの入力をキャンセルするようにした .

##### 4.1 実験 1 : 人工データによる学習実験

まず, 人工的に作成したデータを用いて, CONP-II の動作確認並びに CONP との動作比較を行った . 図 6 (a) に実験に用いた入力データを示す . この図において横軸は時間, 縦軸は入力チャンネルを示す . 各マスは 10 ms 当りのパルス頻度を濃淡で表しており, 色が濃いほどパルス頻度が高いことを示す . 本データは計 11 チャンネルのパルス列であり, 4 種類の入力信号を想定している . 1 と 2, 3 と 4, 5 と 6 チャンネルはそれぞれ 1000/s の発火頻度で同時に立ち上がった後に次第に頻度が減少する入力を意味し, 7 から 11 チャンネルは 1 から 6 チャンネルとは独立に発生する瞬時入力を想定したものである . これらの四つの信号はそれぞれ異なるタイミングで発生しており, 本実験においてはそれぞれを独立した信号として分離して学習することを期待する .

このデータを CONP 及び CONP-II に入力して SOM アルゴリズムにより競合学習を行った . システムのサンプリング周波数は 48 kHz とし, 競合学習ニューロンはそれぞれ 6 個用意した . 各ニューロンは終端が存在しないようにループ状にした次元上に配置

しているとする . 各ニューロンの結合重みの初期値は乱数により与えられ, それぞれ 2000 回の学習を行った . なお, 学習には SOM アルゴリズムを用いているため, 結果としてすべての競合学習ニューロンは学習の初期に入力パターンの平均値を学習しており, 本実験においては Kohonen 学習則を用いたときのような学習結果の初期値依存性は解消されている . ただし, 結合重みの初期値によって学習後の結合重みの順番が位相関係を保持したままループ的にシフトをすることが考えられるが, これは実験結果の評価には影響を与えない . 実験に用いたネットワークの各種パラメータは表 1 のとおりである . 表中の  $w^{NFD}$ ,  $TH^{NFD}$ ,  $\tau^{NFD}$ ,  $w^{MFD}$ ,  $TH^{MFD}$ ,  $\tau^{MFD}$  はそれぞれ NFD ニューロン, MFD ニューロンの結合重み, 発火しきい値, 膜電位減衰時定数を表す . CONP 及び CONP-II で異なる値については, I/II の順で併記してある .

学習後のネットワークに対して学習データを改めて入力したときの CONP, CONP-II それぞれのネットワークの出力を図 6 (b), (c) に示す . これらの図においては縦軸は競合学習ニューロンを表す . また, 図 7 に学習後の CONP 及び CONP-II のそれぞれの競合学習ニューロンの結合重みを示す . 以上の実験結果よ

表 1 実験 1 におけるニューロンの各種パラメータ  
Table 1 The parameters in the experiment 1.

$RP$	1 msec	ON ニューロン	
$n$	11	$T_d$	3 ms
競合学習ニューロン		$w^{ON}$	1.0
$TH$	0.0	$TH^{ON}$	1.2
$\alpha$	$1 \times 10^{-5}$	$\tau^{ON}$	3 ms
$\tau$	10 ms	NFD ニューロン	
$\tau^{AF}$	10 ms	$w^{NFD}$	-1.0
学習回数	2000	$TH^{NFD}$	-0.01
$EN_0$	200	$\tau^{NFD}$	1 ms
$w_{fd}$ I/II	0.2/4.0	MFD ニューロン	
$\beta_{pcw}$ I/II	0.30/0.33	$w^{MFD}$	1.0
$\tau_{fd}$	100 ms	$TH^{MFD}$	2.0
$TH_g$ I/II	3.0/2.0	$\tau^{MFD}$	1 ms

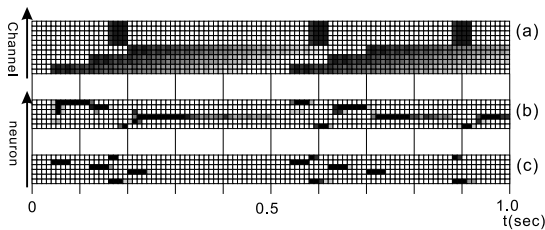


図 6 実験 1 における学習結果  
Fig. 6 Learning results of the experiment 1.

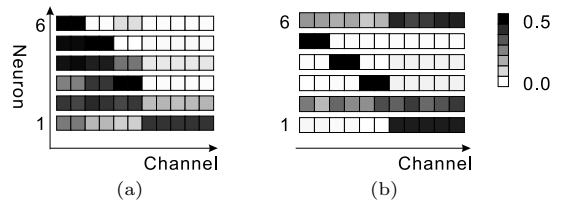


図 7 各ニューロンの学習後の結合重み  
Fig. 7 The connection weights of each neuron after learning process.

り、まず CONP-II では図 6(c) より新しいパターンが入力されるたびに勝者ニューロンが変化しており、10ms 程度の時間解像度においては、入力パターンの変化に追従してベクトル量子化が行われていることが分かる。また図 7(b) より第 1, 3, 4, 5 ニューロンの結合重みにより 4 種類の入力信号を分離して学習していることが分かる。なお入力信号のパターン数に比べて用意した競合ニューロンが多かったため、第 2 ニューロンは一度も発火しないデッドノードとなっており、第 6 ニューロンは第 1 ニューロンとほぼ同じパターンを学習している。CONP-II では各信号の立上りに反応するアクティブフィルタを内蔵しているため、入力信号が持続的なものであっても信号の立上りからある一定時間が経過するとニューロンへの入力信号が消滅してしまう。このため各ニューロンの発火の持続時間は最大でも約 50ms となっている。

次に CONP においては図 6(b) より新しいパターンが入力されるたびに勝者ニューロンは変化しており、ベクトル量子化が行われたことを示しているが、図 7(a) より各時刻における全 11 チャンネルの入力パターンを学習したために第 2, 3, 4, 5 ニューロンにおいて複数信号の混合パターンを学習しており、4 種類の入力信号に対して 6 パターンの結合重みが生成されていることが分かる。つまり CONP においては複数の信号が同時刻に存在する場合にはこれをもとの信号とは別のパターンとして学習していることが確認された。

以上の結果より、CONP-II により新たに提示された信号をそれ以前の入力から分離して学習可能であることが明らかとなり、複数信号の混合信号に対する CONP-II の有効性が確認できた。

#### 4.2 実験 2：混合音による学習実験

次にシンセサイザにより生成された混合器楽音を用いて音信号に対する CONP-II の評価実験を行った。音の作成には Cakewalk 社の Audio Workstation Software SONAR4 を使用し、Software Synthesiser TTS-1 により生成された 3 音程のフルートの音（サンプリング周波数 48kHz, 16bit 量子化）とバックグラウンドホワイトノイズの計四つの音を混合した。三つのフルート音はそれぞれ C4, E5, G6 の音程で、同音量である。ホワイトノイズは各フルート音に対して -24dB の音量で混合した [13]。図 8 に各音の発音タイミングを示す。この図において、横軸は時間を示し、一つのますが 50ms を表す。C4, E5, G6 のそれ



図 8 実験 2 における入力信号  
Fig. 8 The input signal of the experiment 2.

ぞれの音について発音した時間（音が持続している時間）を黒帯で示した。ホワイトノイズについては終始発音し、混合した。本実験ではこの音を CONP 及び CONP-II にて学習し、CONP-II においては、各音の立上り成分以外を除去して学習することにより、発音タイミングの違いから混合音を分離して学習することを期待する。

なお、CONP-II においては複数の音程がほぼ同時に発音した場合には、これらの混合音を一つの音として認識することを想定している。このような音はいわゆる和音であり、各音程の音とは別の新種類の音として識別することが妥当であろう。ただし、別種類の音源からの音がほぼ同時に発音した場合には、それぞれの音の包絡線（音圧の時間変化）情報が異なることにより別々の立上りが検出され、独立した二つの音として学習されることは妥当である。また、今回は単一センサによるモノラル音源を想定しているが、複数のセンサを用いた場合には同時発音した音であっても、その音源の定位方向の違いを用いることで音を分離学習することが可能であろう。

本学習実験では上述の四つの混合音信号をパルス列に変換してネットワークへの入力信号とした。以下、その変換方法を簡潔に述べる。入力信号はまず駒木根らの Lyon 型帯域フィルタ群 [14], [15] により周波数成分に分解される。なお、今回は周波数方向の分解能を向上させるために各フィルタの中心周波数を Bark 軸上の 1/3 から 20 まで、1/3 [Bark] ごとに 60 チャンネル用意し、このうちの第 6（中心周波数約 0.20 kHz）から第 55 チャンネル（中心周波数約 4.8 kHz）の 50 チャンネルを使用した。各周波数成分は文献 [4] で述べた非線形変換によりエンベロープの検出を行い、生成パルス間隔はこのエンベロープの大きさに反比例するとして、パルス列を生成した。また、発火頻度による振幅情報の表現能力を十分にたせるため、1 チャンネル当り 10 本のパルス列を用いた。なお各パルス列は CONP, CONP-II のニューロンと同様の不応期  $RP$  をもつとし、最高発火頻度が 1000/s に制限されてい



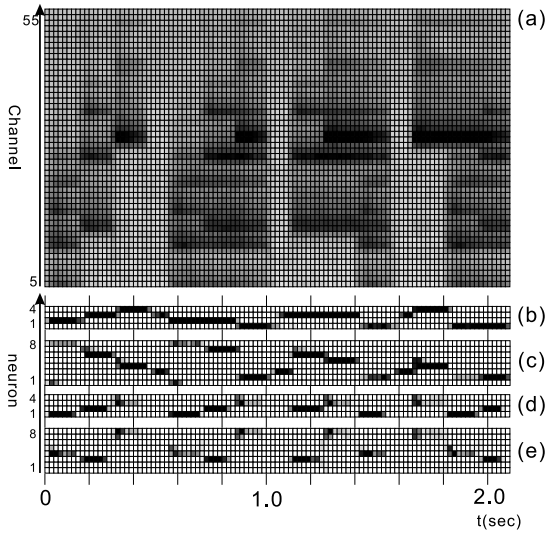


図 9 実験 2 における学習結果  
Fig. 9 The learning results of the experiment 2.

る．以上の方法で生成された入力パルス信号を図 9 (a) に示す．この図において各マスは 20 ms 当りの発火頻度を示し，発火頻度が高いほど濃い色で塗りつぶされている．最大発火頻度は 200/マスである．この図から三つのフルート音それぞれが別の倍音構成をもっており，混合音はそれら倍音の混合から構成されていることが分かる．

CONP, CONP-II を用いて，上述のデータを学習し，学習後のネットワークに対してデータを改めて入力したときのネットワークの出力を図 9 (b), (c), (d), (e) に示す．図 9 (b), (c) は CONP, (d), (e) は CONP-II を用いて学習し，競合学習ニューロンの数を (b), (d) は 4, (c), (e) は 8 とした．実験 1 と同様にシステムのサンプリング周波数は 48 kHz とし，各競合ニューロンは終端が存在しないようにループ状に配置した．これらの図において各マスは図 9 (a) と同様に 20 ms 当りの競合学習ニューロンの発火頻度を表し，最大発火頻度は 20/マスである．実験に用いたネットワークの各種パラメータは表 2 のとおりである．この表において CONP 及び CONP-II で異なる値，若しくは競合学習ニューロンの数を変えた場合に異なる値については，I/II または I-4 (CONP でニューロン数 4 の場合)/I-8/II のように併記した．NFD, MFD ニューロンについては実験 1 と同じパラメータを用いているため省略した．また，図 10 に図 9 (a) より抜粋した入力音 C4, E5, G6 に対応する入力信号の発

表 2 実験 2 におけるニューロンの各種パラメータ  
Table 2 The parameters in the experiment 2.

$RP$	1 ms	競合学習ニューロン	
$n$	50	$\beta_{pcw}$ (I/II)	0.15/0.17
競合学習ニューロン		$\tau_{fd}$	100 ms
$TH$	0.0	$TH_g$ (I/II)	2.0/4.0
$\alpha$ (I-4/I-8/II)	2/5/1 †	ON ニューロン	
$\tau$	20 ms	$T_d$	10 ms
$\tau^{AF}$	20 ms	$w^{ON}$	1.0
学習回数	2000	$TH^{ON}$	1.3
$EN_0$	300	$\tau^{ON}$	10 ms
$w_{fd}$ (I-4/I-8/II)	0.4/0.1/8.0	† × 10 <sup>-6</sup>	

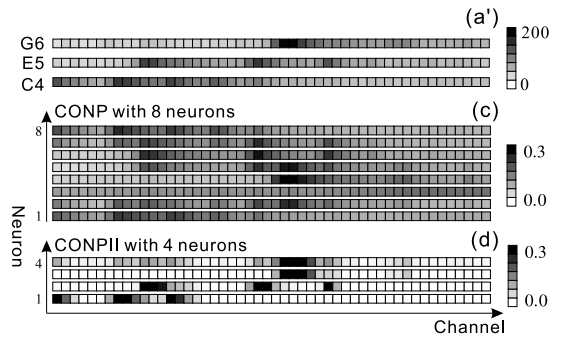


図 10 各ニューロンの学習後の結合重み  
Fig. 10 The connection weights of each neuron after learning process.

火パターン (図 10 (a')) と図 9 (c) 及び (d) で用いたネットワークにより得られた学習後の各競合ニューロンの結合重みを示す．縦軸はニューロン番号を表し，図 9 (c), (d) の縦軸に相当する．横軸は周波数チャンネルを表し，図 9 (a) の縦軸に相当する．

これら実験結果より，CONP の実験結果では図 9 (b), (c) とともに競合ニューロンの中に発火タイミングが同期しているニューロンがないことから，それぞれすべての競合ニューロンを使って入力信号をベクトル量子化しており，入力として用いた独立音 3 種以上のパターンをベクトル量子化していることが分かる．また，図 10 (c) においても，例えば第 4, 6 ニューロンはそれぞれ G6, E5 の単独音を学習しているが，その間にある第 5 ニューロンはこの 2 音の混合音を学習していることが確認でき，上述のとおり CONP は混合前の単独音ではなく，各時刻における全周波数チャンネルの入力パターンをベクトル量子化していることが分かる．

これに対して，CONP-II の実験結果については，まず図 9 (e) では入力パターンのバリエーションに対して競合ニューロンの数が多すぎるためにデッドニュー

ロンが存在しており(第1, 2, 5ニューロン), かつ同一のパターンを複数のニューロンが学習する様子が確認できる(第7及び第8ニューロン). そして図9(d)では四つのうち二つのニューロンが同一のパターンを学習している(第3及び第4ニューロン)ため, 結局3種類のパターンを学習していることが分かる. ここで, 図10(d)より, 第1, 2, 3ニューロンはそれぞれC4, E5, G6の構成周波数に相当するパターンを学習しており, 混合音を構成する三つの音をそれぞれ学習していることが確認できる. なお, 図10(d)において第4ニューロンはG6のパターンを基本として若干C4のパターンを付加したパターンを学習している. これは競合学習ニューロンをループ上に配置したために, 第1ニューロンと第3ニューロンの両方のパターンを学習したと考えることができる.

以上の実験結果から, CONP, CONP-II いずれにおいても競合学習により入力信号をベクトル量子化できているが, CONPでは複数の音が同時に存在する場合にはその混合音をベクトル量子化しているのに対し, CONP-IIでは各音の立上りの違いを用いてそれぞれの構成音を分離してベクトル量子化していることが確認できた. これより, CONP-IIを用いることで立上りのそろった周波数成分のみを一つのパターンとみなしてベクトル量子化が可能であり, それぞれ独立して発生した混合音を分離して競合学習可能であることが明らかとなった.

## 5. む す び

本論文では, 複数の音源からの音が混在する状況においてそれぞれの音の残響成分を除去することで各音を分離して競合学習することが可能な, PNNによる新しい競合学習ネットワーク CONP-IIを提案した. 計算機シミュレーションの結果より, 音の立上りをずらした複数の音を混合した信号に対して CONP-IIを用いることでそれぞれの音を分離してベクトル量子化が可能であることが確認できた.

CONP-IIは各周波数成分の立上りの同期性を用いてそれぞれの音をグルーピングすることが可能であり, PNモデルのもつダイナミクスを有効に利用することで, 本機能を実現している. またこのような機能を実装する場合, 通常はアクティブフィルタと入力信号の乗算演算が必要となりモデルの計算コストが増大するが, 本論文ではPNモデルにおける結合重みの値をアクティブフィルタ値とする手法を提案し, 乗算を用

いることなく本機能を実現できた.

1. でも述べたように PNN を用いた聴覚情報処理システムの実現は, デジタル回路への実装手法として最適なものと考えられ, 本手法の有効性は実際にハードウェア上に実装した際の計算コストについても検討する必要がある. 今後は CONP-II 並びに提案システムを FPGA 等のデバイス上に実装し, その動作に関して検証を行っていく予定である. また, 聴覚情報処理で識別すべき情報は瞬時的な特徴量パターンのみではなく, これらの時系列パターンであり, CONP-II だけではこのような特徴量をベクトル量子化することはできない. 今後は, この時系列パターンのベクトル量子化問題について検討を行い, 周辺環境把握のための聴覚情報処理システムの構築を行っていききたい.

謝辞 本研究の一部は文部科学省科学研究費補助金(若手研究(B))による. また, 本論文で用いた Lyon 型フィルタの設計については当時名古屋工業大学電気情報工学科在籍の森本寛君(現株式会社デンソー)に御協力を頂いた. 記して感謝致します.

## 文 献

- [1] W. Maass and C.M. Bishop, Pulsed Neural Networks, MIT Press, 1998.
- [2] W.H. Pitts and W.S. McCulloch, "A logical calculus of the ideas immanent in nervous activity," Bull. Math. Biophysics, vol.5, pp.115-133, 1943.
- [3] 田中愛久, 黒柳 奨, 岩田 彰, "FPGAのためのニューラルネットワークのハードウェア化手法," 信学技報, NC2000-179, 2001.
- [4] S. Kuroyanagi and A. Iwata, "Auditory pulse neural network model to extract the inter-aural time and level difference for sound localization," IEICE Trans. Inf. & Syst., vol.E77-D, no.4, pp.466-474, April 1994.
- [5] 黒柳 奨, 岩田 彰, "音源方向定位聴覚神経系モデルによる ITD, ILD の脳内マッピングの実現," 信学論 (D-II), vol.J79-D-II, no.2, pp.267-276, Feb. 1996.
- [6] 坂口晋也, 黒柳 奨, 岩田 彰, "環境把握のための音源識別システム," 信学技報, NC99-70, 1999.
- [7] 黒柳 奨, 岩田 彰, "聴覚情報処理システムのためのパルスニューロンモデルを用いた競合学習ニューラルネットワーク," 信学論 (D-II), vol.J87-D-II, no.7, pp.1496-1504, July 2004.
- [8] T. Kohonen, Self-Organizing Maps, Springer-Verlag, 1995.
- [9] 黒柳 奨, 平田浩一, 岩田 彰, "パルスニューラルネットワークのための競合学習手法," 信学技報, NC2001-210, 2002.
- [10] A.J. Bell and T.J. Sejnowski, "An information maximization approach to blind separation and blind deconvolution," Neural Comput., vol.7, pp.1129-1159, 1995.

- [11] 甘利俊一, “独立成分解析: 多変量データの新しい分解法 ; 情報処理, vol.41, no.10, pp.1090–1095, 2000.
- [12] T. Gautama and M.M.V. Hulle, “Separation of acoustic signals using self-organizing neural networks,” Proc. IEEE Neural Network for Signal Processing Workshop 1999, pp.324–332, 1999.
- [13] <http://mars.elcom.nitech.ac.jp/~bw/>より WAV ファイルを取得可能 .
- [14] 駒木根隆士, 平原達也, “蝸牛の周波数分析機能を模擬するフィルターバンクの一構成法 ;” 信学技報, SP87-45, 1987.
- [15] 早川昭二, 梶田将司, 板倉文忠, “蝸牛フィルタを用いた話者の個人性の分析・認識 ;” 信学技報, EA92-19, 1992.  
(平成 17 年 4 月 11 日受付, 12 月 27 日再受付)



黒柳 奨 (正員)

平 3 名工大・電気情報卒 . 平 5 同大大学院博士前期課程了 . 平 8 同大学院博士後期課程了 . 同年名工大・電気情報・助手 . 平 15 同大大学院・助手, 現在に至る . ニューラルネットワーク, 聴覚情報処理に関する研究に従事 . 博士 (工学) . 日本音響学会, 日本神経回路学会, 日本生体医工学会会員 .



岩田 彰 (正員)

昭 48 名大・工・電気卒 . 昭 50 同大大学院修士課程了 . 同年名工大・情報・助手 . 昭 57 年 4 月より昭 58 年 10 月まで, ドイツ連邦共和国ゲーゼン大学医学部医用情報研究所客員研究員 . 昭 59 名工大・情報・助教授 . 平 5 同大・電気情報・教授 . 平 14 同大・副学長, 平 16 同大大学院・教授, 現在に至る . ニューラルネットワーク, 情報セキュリティに関する研究に従事 . 工博 . 平 5 年度本会論文賞受賞, 平 10 年度情報処理学会 Best Author 賞受賞 . 情報処理学会, 日本生体医工学会, 日本心電図学会, 日本神経回路学会, 日本医療情報学会各会員, IEEE Senior Member .