

特集 確率とパルス信号を用いたニューラルネットワークとそのハードウェア化*

An Architecture of Digital Neuro-hardware Using Probabilistically-coded Pulse Neurons

川島 毅 上杉 浩 田中裕章 秋田成行 石黒章夫 大熊 繁

Takeshi KAWASHIMA Hiroshi UESUGI Hiroyuki TANAKA Shigeyuki AKITA Akio ISHIGURO Shigeru OKUMA

In this paper, we present an architecture of neuro-hardware that can be realized on by far a small-scaled circuit compared to the conventional approach. In order to reduce the scale of the circuits, the architecture employs a new method of computing the membrane potential and the sigmoid function by encapsulating the probability properties into relative delay between two pulses. Proposed architecture enables to integrate more than one hundred of neurons on a latest FPGA chip, which means approximately thirteen-fold miniaturization compared to conventional architecture.

Key Words : Neural network, Pulse signals, Probabilistic operation, Digital circuit, Small-scaled realization

1. はじめに

ニューラルネットワークは生物の脳をヒントに考案されたシステムであり、人間が行っているようなあいまいで複雑な情報処理を可能にする仕組みとして期待されている¹⁾。その演算単位はニューロンと呼ばれる細胞であり、これが多数組合わさってネットワークを構成している。代表的なニューロンのモデルとしてシグモイド関数モデルが提案されており²⁾、このモデルを用いてFig. 1のようにニューラルネットワークをデジタル回路で実現しようとする研究が行われている³⁾。ところが、このモデルには乗算や非線型関数が含まれているために、そのままの形で実現しようとすると乗算回路や非線形関数を記憶するメモリ回路が必要になる。そのために回路が複雑になってしまい、多くのニューロンを集積化することは難しい。

また、生物の脳の中に乗算回路やメモリ回路があって演算をしているとは考えにくく、実はもっと単純な仕組みを使って動作しているのではないかと考えられる。脳生理学分野の研究から、ニューロンはFig. 2のように電気的なパルス信号を使って信号を伝達していることが分かってきた⁴⁾。もっともパルスと言っても明瞭な信号ではなく、脳波を観察してみると分かるようにランダムな信号成分が多く含まれている。

我々は、このパルス信号とそのランダムさに着目した結果、ニューロンを簡単な演算で表現できる新しいモデルを開発した。開発したモデルでは主に乱数を用いて演算を行うため、従来に比べて約1/13の回路規模でニューロンを実現することができる。さらに、二

ューロン間の信号伝達もランダムさを許容できるため、ノイズを含んだセンサ信号などをそのままニューラルネットワークに入力して処理することができる。

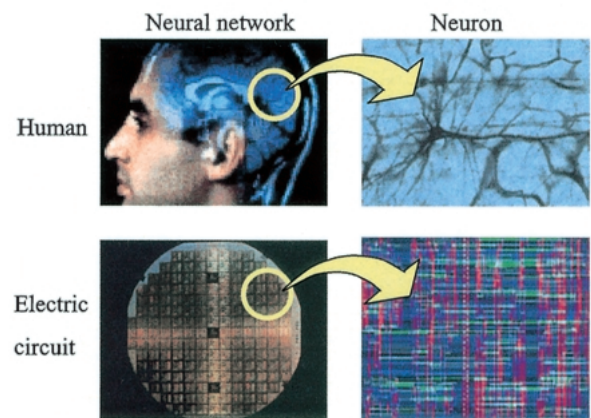


Fig. 1 Biological and artificial neural network

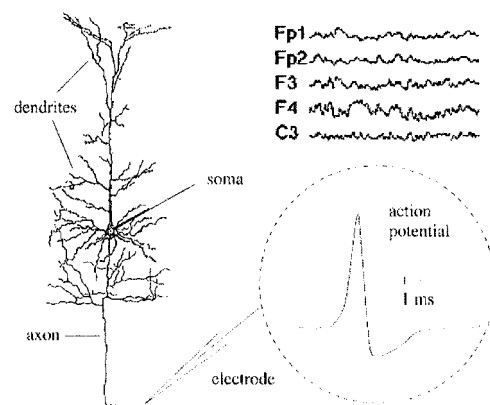


Fig. 2 A pulse signal on a neuron and the electroen-cephalogram of the brain.

* IEEE の了解を得て、下記を和訳し一部加筆して転載

© 2000 IEEE. Reprinted with permission, from "International Conference on Industrial Electronics, Control and Instrumentation -21st Century Technologies and Industrial Opportunities-" ; Oct. 2000, held in Nagoya, Japan, "An Architecture of Small-scaled Neuro-hardware Using Probabilistically-coded Pulse Neurons"

2. 提案するニューロンモデル

シグモイド関数モデルにおいて，ニューロンは*i*個の入力 a_i と1個の出力 y を持ち，次式のように表される．

$$y=f(x) \tag{1}$$

$$x = \sum_i w_i \cdot a_i \tag{2}$$

ここで， w_i は結合係数である(1)式にシグモイド関数：

$$f(t) = \frac{1}{1+\exp(-t)} \tag{3}$$

が含まれ，(2)式に乗算が含まれている．

次に提案モデルについて説明する．提案モデルは，ニューロンの入力と出力にパルス信号を用いることと，そのパルス信号を確率的に発生させることが特徴である．Fig. 3に示すように，入力 A_i と出力 Y はパルス信号(0または1)で表される．ここで，入力として新たに基準パルス入力 T を追加して， T には一定間隔のパルス信号を入力しておく．

$$T=\{T^1, T^2, \dots, T^m\} \tag{4}$$

A_i に inputs するパルスは，(4)式のパルス T^k のそれぞれに対応させて，

$$A_i=\{A_i^1, A_i^2, \dots, A_i^m\} \tag{5}$$

とする．そして，対応するパルス T^k と A_i^k の時間差：

$$\{a_i^1, a_i^2, \dots, a_i^m\} \tag{6}$$

を考え，次の性質を持つように A_i^k パルスを与える．すなわち，(6)式の a_i^k はそれぞれランダムでよいが，全体として現れる確率が平均値 a_i の正規分布となるように A_i^k を与える．ここで正規分布とは，Fig. 4に示すように平均値に近いほど現れる確率が高く，外れるほど確率が低くなるような確率の分布のことである．そして，この平均値 a_i がニューロンへの入力値，すなわち(2)式の a_i を表す．このように，個のパルス列の組を使って一つの入力値 a_i を表すことにする．

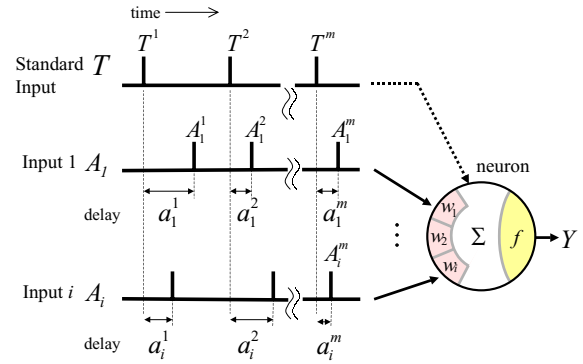


Fig. 3 Input pulse train into the proposed neuron model.

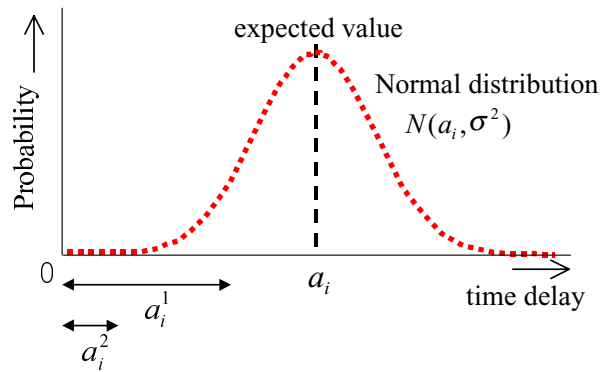


Fig. 4 Normal distribution of time delay.

以上が提案するモデルである．少々複雑なので次のような例を考えてみよう．Fig. 5は，ある赤外線距離センサの出力を繰り返し計測した結果である．センサの出力には誤差が含まれているために，計測距離が同じであっても出力にはばらつきが生じる．しかし計測を何度も繰り返し行えば，その出力は真値を平均値とする正規分布になる．計測を一定時間ごとに行って，得られるセンサ出力をパルス信号で表せばFig. 3のような入力を得られる．

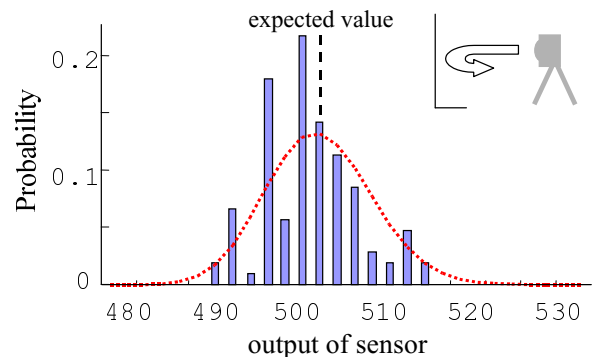


Fig. 5 An example of the normal distribution.

3. ニューロンの演算原理

提案するニューロンモデルを用いると、乗算とシグモイド演算とを、乱数を使った等価的演算へ置き換えることができる。以下では、この等価的演算の原理について述べる。

3.1 乗算

ニューロンの内部では、まず(2)式の乗算 $w_i a_i$ を計算する。通常は乗算回路を用いるのであるが、提案モデルでは乱数を用いて乗算を行う。簡単のために、Fig. 6の「コイン投げ」を考えてみよう。コインを1回投げたときに表が出る確率は0.5であるから、100回コインを投げたとき表が出る回数は平均50回である。これは $0.5 \times 100 = 50$ の乗算を行っていることと等価である。

そこでこの原理を拡張して、表が出る確率が w_i/w_M (w_M :規格化定数)のコインを使ってコイン投げを a_i 回行えば、 $w_i \cdot a_i$ を確率的に求めることができる。数学的には、乗算結果は平均値が $w_i a_i/w_M$ 分散が $(w_i/w_M)^2 a_i^2$ の正規分布となる。この原理を回路で実現するには、時間の長さ a_i をシステムクロック数で表し、クロックの発生ごとに乱数と w_i の大きさを比較することでコイン投げ乗算を等価的に実現できる。

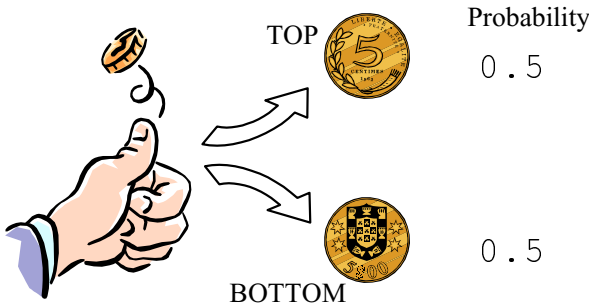


Fig. 6 Coin tossing

3.2 シグモイド演算

引き続き、ニューロン内部では(3)式のシグモイド演算を計算する。(3)式は、Fig. 7に示すように単調算を計算する。(3)式は、Fig. 7に示すように単調増加関数ではあるものの傾きが一定でないため、回路で実現するにはメモリに記憶させておく必要があった。提案するモデルでは、乱数を使うことで、メモリなしでシグモイド演算を計算することができる。

Fig. 7に示すように、シグモイド関数は正規分布を積分した関数 $F(x)$ ：

$$F(x) = \int_{-\infty}^x N(0, \sigma^2) dt \quad (7)$$

とよく一致している。そこで、この積分関数を計算することを考える。例えば、正規分布において $x < 0$ まで積分した値Fig7の(斜線部分)は0.5であるが、これは $F(0)=0.5$ を意味している。このようにシグモイド演算を、入力 x までの正規分布 $N(x, \sigma^2)$ の積分として求めることができる。

この原理を回路で実現するには、平均値が x の正規分布 $N(x, \sigma^2)$ の確率に応じて発生した乱数 x_1, x_2, \dots, x_m (これを正規乱数と呼ぶ)を使って、次の関数： $\hat{F}(x_k)$

$$\hat{F}(x_k) = \frac{1}{m} \sum_{k=1}^m \text{sign}(x_k) \quad (8)$$

を計算すればよい。 $\text{sign}(\cdot)$ は符号関数である。(8)式は、乱数の正負だけを見てシグモイド演算が計算できることを示している。従って回路的には、正規乱数を発生する回路と、その乱数の正負を調べる比較回路、正となる回数を数えるカウンタ回路で構成できる。

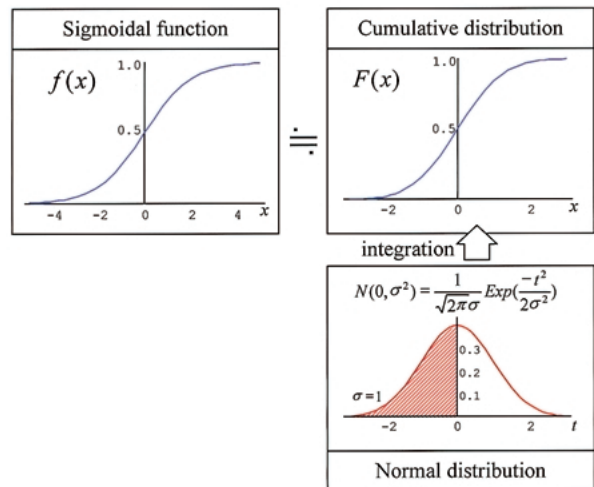


Fig. 7 Approximation of a sigmoid function by the integration of normal distribution.

3.3 正規分布乱数の発生

先ほどのシグモイド演算を行うためには、平均値が x になるような正規乱数が必要である。そのような正規乱数は、一様乱数を足し合わせることで簡単に発生させることができる。

今度は、サイコロを例に考えてみよう。1個のサイコロを振ったとき出る目は、Fig. 8に示すようにどれも $1/6$ の確率で現れる。これを一様乱数という。では次に、サイコロを3個振ってその和を調べてみると、結果は一様ではなく、10と11になる確率が最も高く、

3 や18になる確率は低い．これは確率論では中心極限定理として知られている性質である⁶⁾．この例から分かるように，一様乱数を加算するだけで正規乱数を作り出すことができる．

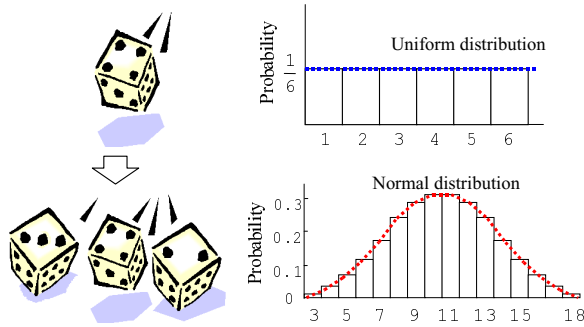


Fig. 8 Tumbling dice.

4. ニューロンの回路規模

以上の演算原理を回路ブロックで表した図がFig 9である．1個のニューロンは，乗算ブロック，加算ブロック，シグモイド演算ブロック f ，および，正規乱数発生ブロックからなる．提案モデルでは，主に一様乱数回路や，比較回路を用いて構成できる．これらの回路は，乗算回路の1/20以下の規模で実現できるため，ニューロンの小規模化に対して有効である．

回路規模を比較するため，提案モデルと従来手法とを電子回路（FPGA; Field Programmable Gate Array）で実現した際の回路規模をシミュレーションにより見積もった．従来手法は，ニューロンへの入力と結合係数との乗算を行う i 個の乗算回路，乗算結果を加算する加算回路，およびシグモイド関数を記憶するメモリを用いて構成した．

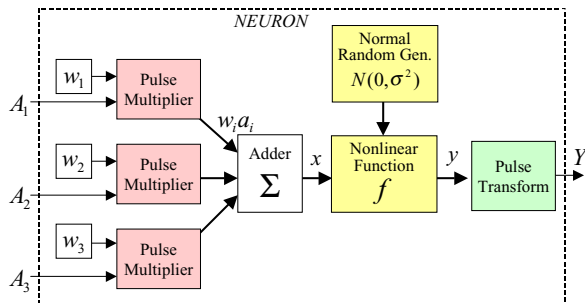


Fig. 9 Circuit configuration of proposed neuron model.

見積もりは，両方の回路をRTL(Resister Transfer Level)記述した後，論理合成・マッピングして行った．結果をFig. 10に示す．従来手法は，ニューロンの機

能をそのままハードウェア化するという意味で最もシンプルな実現方法であるが，回路が大規模化するという問題点がある．

例えば，8ビット精度で入力数 $i=10$ のニューロンを1個構成するのに，10個の8ビット固定整数型並列乗算回路，8ビット入出力を有するメモリ，および加算回路を用いた場合，約6.3kLCs（Logic Cells；FPGAの機能単位で，1LC=10～20Gatesに相当する）の回路が必要となり，最新のFPGAを用いても9ニューロン程度しか集積することができない．また大規模化の原因を分析すると，乗算回路とシグモイド関数用メモリが全体の回路量の95%を占めていることが分かる．これに対して，提案モデルにおける1ニューロン当たりの回路規模は約480LCsであった．このことから提案モデルでは回路規模を約1/13へ小規模化可能であることが見てとれる．

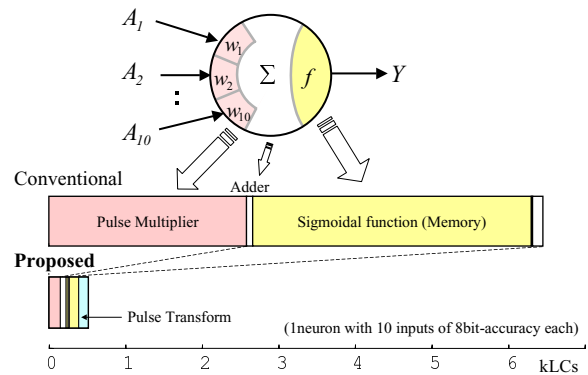


Fig. 10 Quantitative comparison of the resultant hardware-scale between proposed and conventional model.

5. 動作検証

最後に，提案モデルの動作を確認するために，EX-OR関数を構成してシミュレーション上で検証を行った．EX-OR関数は2入力(a, b)出力へ関数であり， $(0,0) 0, (0,1) 1, (1,0) 1, (1,1) 0$ となる滑らかな非線形関数である．この関数はニューラルネットワークの検証によく用いられている．6個のニューロンを用いて，Fig. 11に示すニューラルネットワークを構成した．

結合係数は文献⁶⁾にしたがって $w_{11}=3.7, w_{12}=3.7, w_{13}=1.5, w_{21}=2.9, w_{22}=2.9, w_{23}=-4.6, w_{31}=4.5, w_{32}=-5.2, w_{33}=-2.0$ とした．

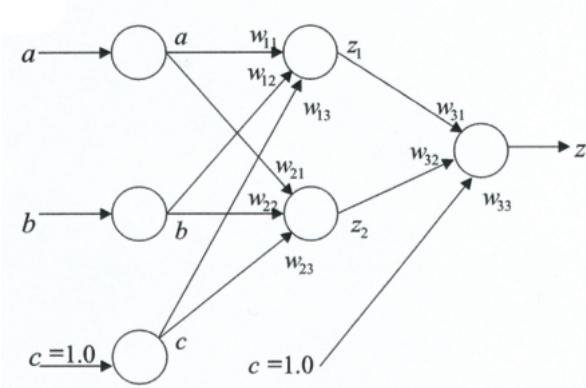


Fig. 11 The neural network for EX-OR problem.

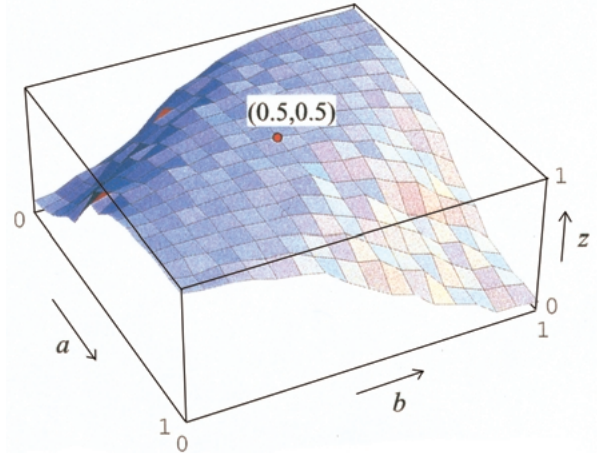


Fig. 12 An output of proposed neural network for EX-OR problem.

ニューラルネットワークの出力を Fig. 12 に示す。Fig. 12の結果は従来モデル，すなわちシグモイド関数を用いたニューラルネットワーク出力とほぼ一致している(誤差 ± 0.09 以内)。このことから，提案モデルは従来モデルと同様の機能を持っていることが分かる。ただし Fig. 12 から分かるように，ネットワークの出力には凹凸があり滑らかに変化していない。これは提案モデルが乱数を用いた確率演算を行っているために，出力にばらつきが生じるためである。ばらつきを抑制するためには，入力値を表すパルス数(式(4)における m)を増加させて，正規分布をより正確に再現すればよい。今回の検証では $m=1024$ とした。

逆に提案モデルは，正確な演算はできない代わりに，ばらつきを持った入力を取扱うことができることが特徴である。例えば，Fig. 13は平均入力が $a=0.5, b=0.5$ (Fig. 12の赤点に対応する)における EX-OR ネットワークの出力 z を示したものである。横軸に入力回数(式(5)の $k=1, 2, \dots$)を示し，縦軸にはその瞬間の入力値(式(6)の時間差 a_i^*)を示している。

ニューラルネットワークには，その瞬間ごとのばらついた入力を与えているにもかかわらず，ネットワークの出力は安定して得られていることが分かる。

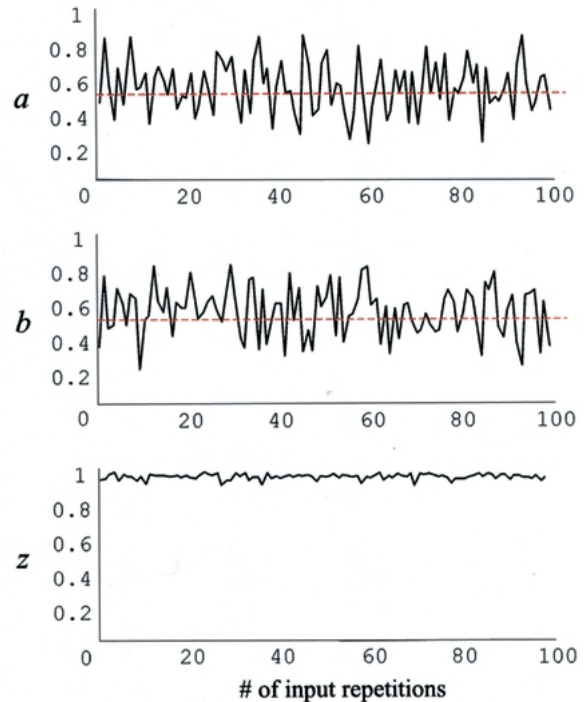


Fig. 13 Repetitive output for noisy input signals.

6. まとめ

本稿では，小規模デジタル回路でニューロハードウェアを構成するためのニューロンモデルを提案した。提案モデルは，パルス信号を伝達し乱数を用いて演算を行うため，従来に比べて $1/13$ の回路規模でニューロンを実現することができる。また，ニューロン間の信号伝達もランダムさを許容できるため，ノイズを含んだセンサ信号などをそのままニューラルネットワークに入力して処理できるという特徴を備えている。

今後は、提案モデルに学習機能を付与した上で、実際にハードウェア化を行って自律移動ロボット等へ搭載する研究を進めている⁷⁾。

なお、本研究は(財)名古屋産業科学研究所「創発型ソフトコンピュータ開発プロジェクト」の協力を得て行われた。

参考文献

- 1) 甘利 俊一: “ニューラルネットの新展開”, サイエンス社 (1993)
- 2) 中野 肇 他: “入門と実習ニューロコンピュータ”, 技術評論社 (1989)
- 3) M.Yasunaga, N.Masuda, M.Asai, M.Yamada, A.Masaki, and Y.Hirai: “A wefer scale integration neural network utilizing completely digital circuits”, proceedings of IJCNN ' 89, Vol.2, pp.213-217 (1989)
- 4) Wolfgang Maass, Christofer M.Bishop: “Pulsed Neural Networks”, The MIT Press (1999)
- 5) 伊藤 清: “確率論I,II”, 岩波書店 (1976)
- 6) J.E.Dayhoff, 桂井 浩 訳: “ニューラルネットワークアーキテクチャ入門”, 森北出版 (1992)
- 7) P.Eggenberger, A.Ishiguro, S.Tokura, T.Kondo, T.Kawashima and T.Aoki: “Toward Seamless Transfer from Simulated to Real Worlds: A Dynamically- Rearranging Neural Network Approach”, Proc. of the 8th European Workshop on Learning Robot (EWLR-8), pp.4-13 (1999)

< 著 者 >



川島 毅
(かわしま たけし)
基礎研究所 第1研究
創発型ソフトコンピュータの開発
に従事。



上杉 浩
(うえすぎ ひろし)
基礎研究所 第1研究
画像処理回路, 創発型ソフトコン
ピュータの開発に従事。



田中 裕章
(たなか ひろあき)
基礎研究所 第1研究
自動車用LSI, ASIC機能回路の研
究開発に従事。



秋田 成行
(あきた しげゆき)
基礎研究所 第1研究
自動車用LSI, ASIC機能回路, 半
導体センサの研究開発に従事。



石黒 章夫
(いしぐろ あきお)
名古屋大学大学院 工学研究科 計
算理工学専攻 助教授
創発システム, ロボット工学の研
究に従事。



大熊 繁
(おおくま しげる)
名古屋大学大学院 工学研究科
電気工学専攻 教授
工学博士
電気工学に関する教育・研究, 主
として, 創発型ソフトコンピュ
ータの開発, ロボティクス, パワー
エレクトロニクスの研究に従事。