

車載用安全運転支援装置のためのパルスニューロンモデルによる音源接近検出及び音源種類識別システム

岩佐 要^{†a)} 藤角 岳史[†] マウリシオ クグレ[†] 黒柳 奨[†]
 岩田 彰[†] 段野 幹男^{††} 宮治 正廣^{†††}

A Detection System of Approaching and Recognition Sound Source by Pulsed Neuron Model for On-Vehicle Safety Driving Support Device

Kaname IWASA^{†a)}, Takeshi FUJISUMI[†], Mauricio KUGLER[†],
 Susumu KUROYANAGI[†], Akira IWATA[†], Mikio DANNO^{††}, and Masahiro MIYAJI^{†††}

あらまし 現在，自動車の普及とともに交通事故が増加しており安全運転支援装置の開発が望まれている．その中でも音を利用した接近車両検出手法は障害物の多い交差点や死角に入る二輪車の検出など対象物を視覚的に検出できない環境において有利である．本論文では，車両の接近を検出する手段として音を利用し，ハードウェア上に実装したパルスニューロンモデルを用いて新たに音源接近検出及び音源種類の識別を行うシステムを提案する．実験の結果，本システムが音源の接近検知と音の種類の識別を同時に行い，かつ入力音に対してリアルタイムに処理を行うことができ，従来手法と比べて小さな回路規模で実装できることが確認できた．

キーワード 運転支援システム，パルスニューロンモデル，ハードウェア実装，音源接近検出，音源種類の識別

1. ま え が き

近年，自動車は社会に広く普及し必要不可欠な存在となっている．その半面，交通事故の発生も年々増加しており [1]，わき見や不注意などの運転者の認知不足が原因となる事故が多発している．このような背景から，周辺環境の情報を把握し運転者に対して警告するような運転支援システムの開発が望まれている [2]．本研究では特に，交通事故の要因として発生頻度の高い，自車が交差点などで停止または徐行していたときに，死角から二輪車が接近することで起こる巻き込み事故や緊急車両が接近していることに気づかず衝突事故を起こすといった状況を想定する．運転者が周辺環境の

情報を得る方法としては目視などによる視覚的手段と警報音を聞くことによる聴覚的手段があり，これらを利用して接近車両の検知や交通状況を調べるといった運転支援システムの研究・開発が行われている [3] ~ [7]．

視覚的手段に関しては，カメラを車両に設置し画像処理によって接近車両などを検知する手法や [3], [4]，センサを用いて周辺環境を検知する手法 [5] などが数多く提案されている．しかし視覚的手段では障害物の多い街路地や夜間，交差点などで二輪車が死角に入るときなど，車両を直接確認しにくい環境に対応することが難しい．また，画像処理による方法では車両の検出などを高精度に行うことができる反面，演算コストや装置規模の点で高価になることが多い．これに対して聴覚的手段を用いた場合，上記のような対象物を視覚的に検出することのできない環境であっても，音情報は得られるため音源の接近を検出するには非常に有効であるといえる．加えて画像処理装置に比べて装置規模が小さくコストの面からも有利である．このような利点から，音を利用した接近車両検知システムについての研究も各種行われている [6], [7]．しかしこれら

[†] 名古屋工業大学，名古屋市
 Nagoya Institute of Technology, Gokiso-cho, Showa-ku,
 Nagoya-shi, 466-8555 Japan

^{††} (株)トヨタ IT 開発センター，東京都
 Toyota InfoTechnology Center, Co., Ltd., 6-6-20 Akasaka,
 Minato-ku, Tokyo, 107-0052 Japan

^{†††} トヨタ自動車株式会社，豊田市
 Toyota Motor Corporation, 1 Toyota-cho, Toyota-shi, 471-
 8572 Japan

a) E-mail: kaname@mars.elcom.nitech.ac.jp

は車内に計算機を用意して識別処理を行っており、装置が大きかりなものになっている。

一方、我々はこれまでに FPGA (Field Programmable Gate Array) を用いてパルスニューロンモデル (Pulsed Neuron 以下 PN) モデル [8] をハードウェア上に実装し競合学習ネットワークを構成した聴覚情報処理システムを提案している [9]。本システムは高性能な計算機を用いなくても入力音に対してリアルタイムに音源方向の定位と音源種類の識別処理を同時に行うことができ、1 個の FPGA (3 cm×3 cm, 約 48,000 [ALUTs]) 上にすべての回路を実装することができるため小さな装置でシステム全体を実現することが可能である。

以上のことから本論文では、FPGA 上に実装された PN モデルにより入力音から音源の接近検知とその音源種類を識別するシステムを提案する。車両の接近を検知する方法として FPGA で実装された PN モデルを組み合わせることで、新たに入力信号と一定時間前の入力信号の間の音圧差を各周波数成分ごとに検出するネットワークを構成し、音源の接近を判断する。PN モデルを用いることで、基本素子である PN モデルの構成に変更は加えなくても、その組合せを変更するだけで接近の検出を行うことができる。また基本素子が同じであるため検出回路や識別回路などを一つのハードウェアに搭載することが容易であり、装置の規模を小さくすることができる。更に検出された周波数成分ごとの音圧差の情報を音圧差検出部と同じく FPGA 上で構成された競合学習ネットワーク [9] に与えることで音源の接近及び離脱の検出と同時に音源の種類を識別する。入力音の一定時刻間の音圧差を用いて車両の接近を検出する方法は星野の提案する手法 [7] でも用いられているが、本研究では PN モデルで構成された音圧差検出ネットワークにより音圧差情報を検出する点が異なる。次章より、提案するシステムとその実装方法について説明した後、実験より提案システムが入力された音信号から音源の接近検出及びその種類の識別を行うことができることを示す。また従来手法と比較して回路規模の面において有利であることを示す。

2. 提案システム

本論文で提案するシステムについて概略図を図 1 に示す。本システムは大きく分けて周波数-パルス変換部、音圧差検出部、音圧差情報量子化部、音源種類及び接近識別部からなる。音圧差検出部、音圧差情報量

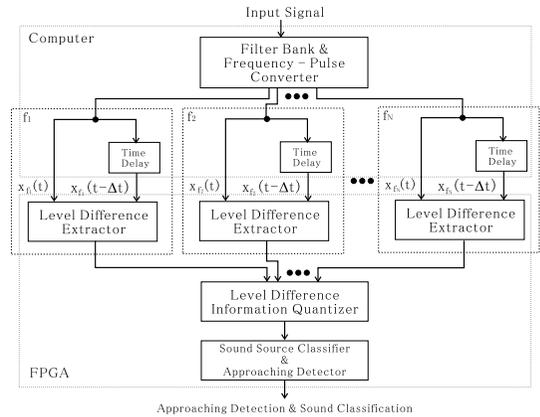


図 1 提案システムの概略

Fig. 1 The abstract of proposed system.

子化部、音源種類及び接近識別部については PN モデルで構成されており、FPGA 上に実装されている。なお本論文においては周波数-パルス変換部及び時間遅れを与える部分については計算機で処理される。

入力された音信号は計算機上の周波数-パルス変換部で複数の周波数成分に分割された後パルス情報へと変換される。変換されたパルス情報には並行して時間遅れが与えられ、音圧差検出部に入力される。音圧差検出部で検出された音圧差情報を音圧差情報量子化部である競合学習ネットワークに入力することで量子化を行う。最後に量子化された情報から教師あり学習を行った競合学習ネットワークにて音源の種類を識別する。最後に得られた接近検出結果及び音源種類の識別結果は再び計算機へ返され、計算機上で出力結果を表示する。

2.1 パルスニューロンモデル

音信号のような時系列データに関して情報処理を行うために PN モデルが提案されている [8], [10]。PN モデルと通常のニューロンモデルとの違いは、PN モデルが入出力に神経パルスを想定し 0 か 1 の値をとるパルス信号を用いている点である。このため各素子における演算が単純であり、大規模なネットワークを少ないハードウェア資源で構成することができる。また入力の情報を時定数で減衰する膜電位として保持しており、音信号のような時系列信号の扱いに適している。本論文で用いた PN モデルの詳細については、紙面の都合上、文献 [10], [11] を参照されたい。

PN モデルの構造を図 2 に示す。PN モデルは他のニューロンと結合する n 本のシナプス、各シナプスに

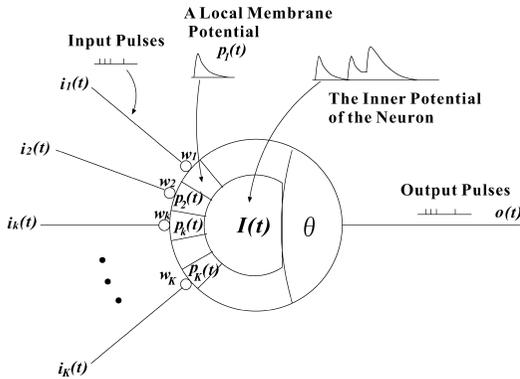


図 2 パルスニューロンモデル
Fig. 2 A pulsed neuron model.

おける局所膜電位 $p_k(t)$, 各局所膜電位の総和である内部電位 $I(t)$, 出力を他のニューロンに伝達する軸索から構成される. PN モデルにおいては, まず各シナプス k において入力パルス $i_k(t)$ が入力されると, それぞれのシナプスに対応した結合重み w_k の値だけ局所膜電位 $p_k(t)$ が増加し, その後時間の経過とともに時定数 τ_k で指数関数的に減衰する (式 (1)). 次に PN モデルは単位時間ごとに局所膜電位 $p_k(t)$ の総和である内部電位 $I(t)$ を計算する. この内部電位 $I(t)$ がしきい値 θ を超えると発火, 時刻 t において出力 $o(t)$ として 1 を出力する. またある PN モデルが一度発火した後, 一定時間 RP の間は内部電位がしきい値を超えた場合はこの PN モデルは発火できない. この時間を不応期と呼ぶ.

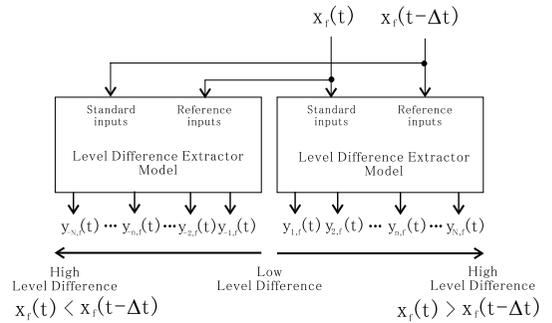
$$p_k(t) = w_k i_k(t) + p_k(t - 1) \exp\left(-\frac{dt}{\tau}\right) \quad (1)$$

2.2 周波数-パルス変換部

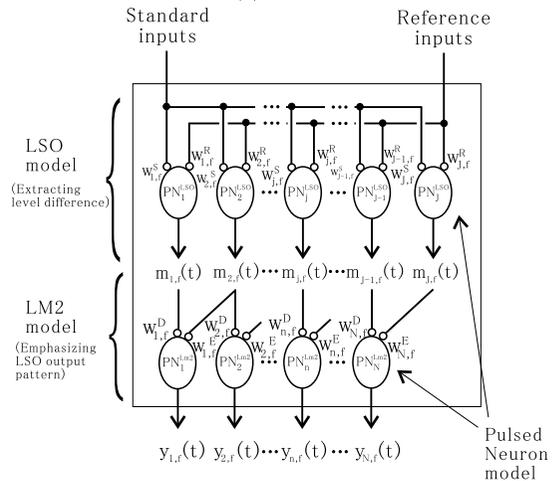
まず音信号は周波数-パルス変換部に入力される. 周波数-パルス変換部では, 左右の信号ごとにまず周波数情報に変換された後, Band Pass Filter 群によって各中心周波数の特性をもった周波数帯域に分割される. なお各フィルタはデジタルフィルタである. 各フィルタの周波数特性は中心周波数を対数尺度において 300 Hz~3 kHz までの間で 11 分割し, それぞれが等間隔になるように決定してある. 次に各周波数帯の情報ごとにパルス列に変換する. パルス列への変換は, 周波数変換後の振幅値が大きいほど確率的に多くのパルス列が含まれるようにする.

2.3 音圧差検出部

音源の接近を検出する手法としては, 黒柳らの提案



(a) 全体図



(b) 音圧差検出モデル

図 3 音圧差検出ネットワーク

Fig. 3 The network of extracting time difference.

した音圧差検出手法 [10] を用いる. 黒柳らの手法では音源定位を行うために左右二つの信号の間の音圧差を検出しているが, 本システムでは, 入力される一つの音信号のうちある周波数成分 f に関するパルス列 $x_f(t)$ と, 同じパルス列に一定の遅れ時間を与えたパルス列 $x_f(t - \Delta t)$ を音圧差検出部に与えて音圧差情報を検出する. 検出の結果, 前の時間における音圧より現在の時間における音圧の方が大きい場合, 対象となる音源が接近していると判断する.

音圧差検出部では各周波数ごとにパルス変換部で変換されたパルス列を図 3 (a) に示す音圧差検出ネットワーク [10] に入力する. 音圧差検出ネットワークは, 図 3 (b) に示すように PN モデルで構成された LSO モデルと LM2 モデルからなり, 入力された二つのパルス情報から音圧差を検出する. LSO モデルにおける各 PN モデル PN_j^{LSO} の重み $w_{j,f}^S, w_{j,f}^R$ は式 (2) で与

えられ、添字 n が大きいニューロンほど重みの値が負の方向に大きくなる．また入力パルスの頻度は振幅値に比例しているため振幅値の差が大きく入力パルスの頻度の差が大きくなるほど添字 n が大きいニューロンまで発火するように設定されている．

LM2 モデルにおける各 PN モデル PN_n^{LM2} の重み $w_{n,f}^D, w_{n,f}^E$ は式 (3) で与えられ、側抑制の効果により LSO モデルにおける出力パルスパターンの特徴を強調する働きがあり、音圧差検出部の出力 $y_n(t)$ より入力信号の音圧差を判断することができる．その他音圧差検出部の詳細については、紙面の都合上、文献 [10] を参照されたい．

$$w_{j,f}^S = 1.0, w_{j,f}^R = \begin{cases} -10^{\frac{j}{\gamma}} & 0 < j < b \\ -10^{\frac{j-i}{\epsilon}} & j \geq -b \end{cases} \quad (2)$$

$$w_{n,f}^D = 1.0, w_{n,f}^E = -1.0 \quad (3)$$

2.4 音圧差情報量子化部

音圧差情報量子化部には PN モデルによって構成された競合学習ネットワーク CONP (COmpetitive learning Network using Pulsed neurons) [11] を用いる．音圧差情報量子化部の CONP では教師なし学習を行う．この教師なし学習では、ニューロンが発火した場合に各入力 k のもつ膜電位 $p_k(t)$ の値と学習係数 α を積算した値を重みに加算した後、重みベクトルのノルムが 1 になるよう正規化して更新する．これにより各ニューロンの重みは入力ベクトルに従った値となる．そして各ニューロンは各入力ベクトルに対応することになり、ニューロンの出力からベクトル量子化を行うことができる．

CONP の学習では、入力ベクトルに対して最も類似した重みをもつニューロン（勝者ニューロン）を決定し、勝者ニューロンのみが発火して学習するようにしなければならない．しかし出力値が実数値ではなく 0 または 1 のパルス信号であり各素子が独立非同期に動作している PN モデルでは勝者を決定することが困難である．そこで、外部にある制御ニューロン (Control Neuron) が出力の状態を判断し各競合学習ニューロン (Competitive Learning Neuron 以下 CL ニューロン) のもつ内部電位を一律に調整し、CL ニューロンが一つだけ発火するような状態にするという制御を行っている [11]．制御ニューロンとしては、CL ニューロンが一つも発火していないことを検出する無発火検出 (No Firing Detection) ニューロンと、CL ニューロンが複

数個発火していることを検出する複数発火検出 (Multi Firing Detection) ニューロンを用意する．そしてしきい値 θ の他に制御しきい値 θ_{gate} を設定し、入力パルス頻度 $p_{in_k}(t)$ の総和が制御しきい値以下であれば制御ニューロンからのしきい値の調整は行わない．各 CL ニューロンにおいては内部電位から入力パルス頻度を一定の割合 β で引くことで、内部電位の急激な変化を抑え、CL ニューロンの内部電位の調整における追従性を向上させている．その他 CONP の詳細については、紙面の都合上、文献 [9], [11] を参照されたい．

2.5 音源種類及び接近識別部

音源種類及び接近識別部では音圧差情報量子化部でベクトル量子化されたパターンが入力され、入力パターンからどの音であるかを識別すると同時に、その音源が接近しているか離脱しているかということも識別する．音源種類及び接近識別部も CONP で構成されている．出力部として競合学習ネットワークを用いた理由として、背景雑音や自車のエンジン音などの情報に対して頑健にする他に、適用する環境で学習させることによって識別するためのパラメータである各ニューロンの重みの状態をシステムが自動的に環境に合わせることができ、利用者が手動で設定するという不便さをなくすという目的がある．

学習に関しては、識別すべき音の種類を明示的に示す必要があるため黒柳らの提案した PN モデルにおける教師あり学習法 [12] に基づき計算機上で各 PN モデルの重みを更新する．具体的には、まず学習のために出力に関する膜電位 $p_{out}(t)$ と教師パルスに関する膜電位 $p_{sup}(t)$ を用意しておく．各ニューロンが発火し、かつ教師パルスが与えられた場合に式 (4) に従い重みが更新される．この教師あり学習により各ニューロンの重みはそれぞれ識別すべき音の入力ベクトルに近づき、識別すべきでない入力ベクトルから遠ざかる．これによりニューロンの出力と識別音を対応させることができる．

$$w_k(t+1) = w_k(t) + \alpha \cdot p_{in_k}(t) \cdot \{p_{sup}(t) - p_{out}(t)\} \quad (4)$$

ただし制御ニューロンからのしきい値制御により、必ずいずれかのニューロンが出力結果を返すことになってしまうため、後述の実験においては学習した目的音以外の音が入力された場合のために“その他”と識別するための素子を用意した．

3. FPGA への実装

3.1 パルスニューロンモデルの実装

FPGA 上で電子回路として実現した PN モデルのブロック図を図 4 に示す [9]. 従来のニューロンモデルにおいては入力信号がアナログ値であるため, “入力 × 重み” という計算を行うために積算器を必要とする. しかし PN モデルでは入力が 0 または 1 のパルス信号であるため, “入力 × 重み” の計算が積算ではなく, 入力信号が入力されたときに重みの値だけ出力する AND 回路で実現することができる. AND 回路より出力されたそれぞれの値は加算器を用いて加算され, 比較器にてしきい値と比較して出力が決定される. 時定数による減衰は, ビットシフトと補数演算で実現しておりもとの内部電位の何分の一かを減算するという演算を行う. 本来減衰は各膜電位ごとに行う必要があるが, 本実験では各膜電位の時定数を同じ値としたため膜電位の総和である内部電位に対し減衰を行うことで等価な回路とした.

なお今回の実装にあたっては, PN モデルのもつ学習機構は FPGA 上には実装されていない. このため, 計算機上であらかじめ学習して生成された重みを固定小数点表現に変換して回路に与えることで代用し, 認識動作のみを行うこととしている.

3.2 音圧差検出部の実装

PN モデルによる音圧差検出部のうち, LSO モデルに関しては FPGA 上では図 5 のように構成した. 本来図 4 のように各ニューロンごとに各入力に対する重みである AND 素子を用意することが適切であるが, 図 3(b) に示す重み $w_{j,f}^S$ は式 (2) に示すように 1.0 で一律でありかつ入力もすべてのニューロンに同時に与

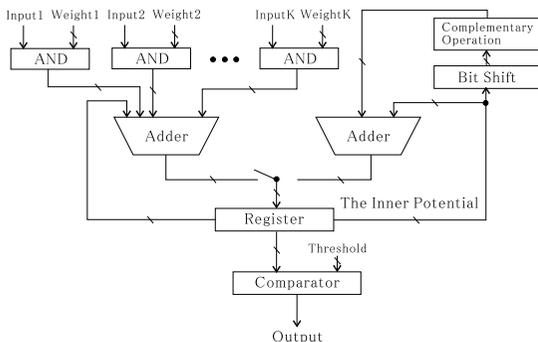


図 4 FPGA におけるパルスニューロンモデル
Fig. 4 A pulsed neuron model on FPGA.

えられるため一つの AND 回路を用いるだけで等価となる.

LM2 モデルに関してはそれぞれ $w_{n,f}^D, w_{n,f}^E$ である二つの重みの AND 回路をもつ PN モデルを複数個並べることによって構成した.

3.3 競合学習ネットワークの実装

PN モデルによる競合学習ネットワークを図 6 に示す [9]. 図 4 の FPGA による PN モデルを基本とし, 競合学習ニューロン (以下 CL ニューロン) と制御ニューロンを配置する. 入力パルスの頻度の計算につ

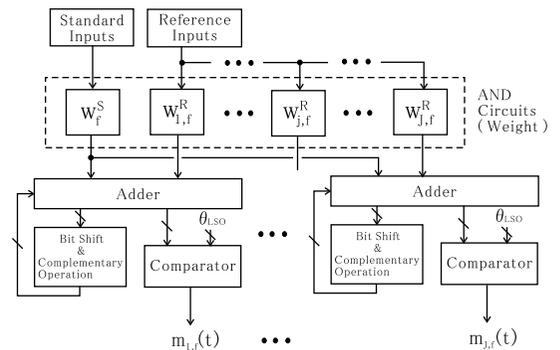


図 5 FPGA における LSO モデル
Fig. 5 The detection of approaching network on FPGA.

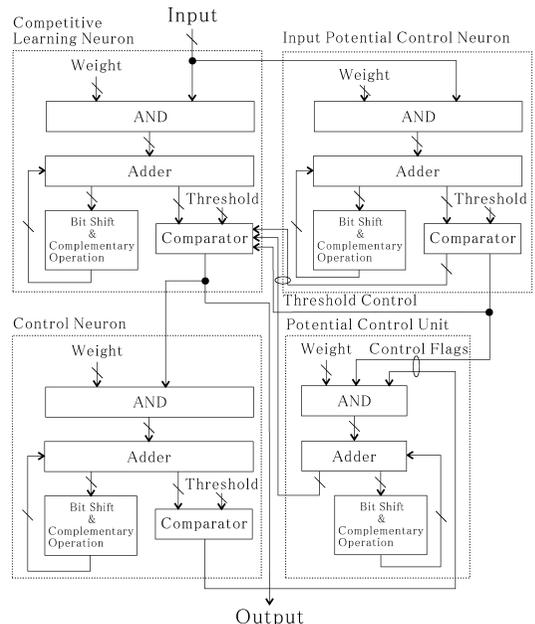


図 6 FPGA における競合学習ネットワーク
Fig. 6 The competitive learning network on FPGA.

いては、すべての CL ニューロンにおいて時定数及び重みは同じであるため外部に一つ、入力パルス頻度を計算するためのニューロン (Input Potential Control Neuron: 以下 IP ニューロン) を用意することで等価な回路とした。

そして単一発火のための制御方法は、内部電位を調整することはしきい値を調整することと等価であるため、各 CL ニューロンの内部電位に対して直接加減算するのではなく外部からしきい値を与えて発火の判定を行うことで実現した。すなわち入力パルスの頻度から内部電位の減少させるには IP ニューロンから CL ニューロンに対してしきい値 $\theta_{in}(t)$ を与えることで等価な回路とした。同じく制御ニューロンからのしきい値の制御に関しても外部にしきい値制御ユニット (Threshold Control Unit) を無発火検出用と複数発火検出用に一つずつ用意してしきい値 $\theta_{nfd}(t)$ と $\theta_{mfd}(t)$ の 2 種類を各 CL ニューロンに与えることで等価な回路とした。

4. 実験

4.1 実験条件

前章までに説明したシステムを用いて音源接近検出及びその種類の識別実験を行う。

本実験では自車が交差点などで停止または徐行していたときに、死角から二輪車や緊急車両が接近している状況を想定する。このため学習用の音としては、図 7 のように自車が停止中に音源が時速 30 km/h で 2 秒間で接近しその後 2 秒間で離脱すると仮定して音圧を変化させた音を計算機上で作成した。具体的には、最も接近したときに 1 m 離れた所に位置しこのときの音圧を S_{base} として、各時間における音圧 $S(t)$ を距離 $d(t)$ の二乗に反比例させる。音源の種類としては“パトカーのサイレン音”、“救急車のサイレン音”、“スクーターのエンジン音”をの 3 種類と、“その他”として出力させるために“路上の背景雑音”を用意した。このため識別する数としては 3 種類の音の接近と

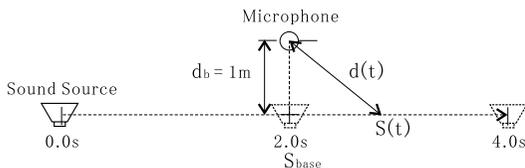


図 7 学習実験に用いた音源の移動

Fig. 7 Sound source movement on experiments.

離脱、その他の合計 7 種類となる。また本システムの実装に用いた計算機及び FPGA の仕様を表 1 に示す。

はじめにこれらの音を用いて表 2 に示すパラメータを用いて計算機上で各 CONP の学習を行った。このうち LSO ニューロン, CL ニューロンの時定数に関しては時間的な音圧差を検出する場合には 0.1 [s] 程度の分解能があれば十分であると考えて設定し、各ニューロン数に関しては FPGA に実装することが可能な最大の個数であることから表の値を設定した。その後、学習の結果得られた重みや、表 2 に示す競合学習パ

表 1 使用する機器の仕様

Table 1 The specifications of using instruments.

計算機		FPGA	
CPU	Pentium4 3.8 GHz	開発ボード	Altera 社 StratixII EPS60
メモリ	2 GB	総回路数	48,352 [ALUTs]
OS	WindowsXP	計算機との通信規格	USB 2.0
コンパイラ	Microsoft Visual Studio C++ .NET	論理合成ソフトウェア	Altera 社 QuartusII 7.0

表 2 実験におけるパラメータ

Table 2 The parameters of each unit on experiment.

入力音	
サンプリング周波数	16 [kHz]
量子化ビット数	8 [bit]
周波数-パルス変換部	
用いる周波数帯域数	11 [Channel]
音圧差検出部	
時間遅れ Δt	0.4 [s]
LSO ニューロン総数 (2N)	42 [個]
b	6
LSO ニューロン間隔	1 [個]
LSO ニューロン重み係数 γ / ϵ	60 / 60
LM2 ニューロン総数 (2L)	20 [個]
しきい値 $\theta_{LSO} / \theta_{LM2}$	0.001 / 0.001
時定数 τ_{LSO} / τ_{LM2}	0.1 [s] / 35 [μ s]
不応期 RP_{LSO} / RP_{LM2}	1.0 / 1.0 [ms]
競合学習ニューロン	
音圧差情報量子化部 CL ニューロン数	23 [個]
音源種類及び接近識別部 CL ニューロン数	7 [個]
しきい値 θ	1.0×10^{-4}
制御しきい値 θ_{gate}	50.0
入力パルス頻度に対する割合 β	3.125×10^{-2}
膜電位減衰の時定数 τ_p	0.1 [s]
不応期 RP_{CL}	1.0 [ms]
無発火検出ニューロン / 複数発火検出ニューロン	
時定数 τ_{no} / τ_{multi}	0.5 / 1.0 [ms]
しきい値 $\theta_{no} / \theta_{multi}$	$-1.0 \times 10^{-3} / 2.0$
CL ニューロンへの結合重み	20.0 / -20.0
競合学習パラメータ	
学習係数 α	6.0×10^{-7}
学習回数	1000
学習半径の収束回数	200

表 3 FPGA 上での各回路の回路数
Table 3 The number of each circuit on FPGA.

部位	回路数 [ALUTs]
FPGA 入出力部	83
音圧差検出部	25,498
音圧差パターン量子化部	16,170
音源種類及び接近識別部	1,349
合計	43,100

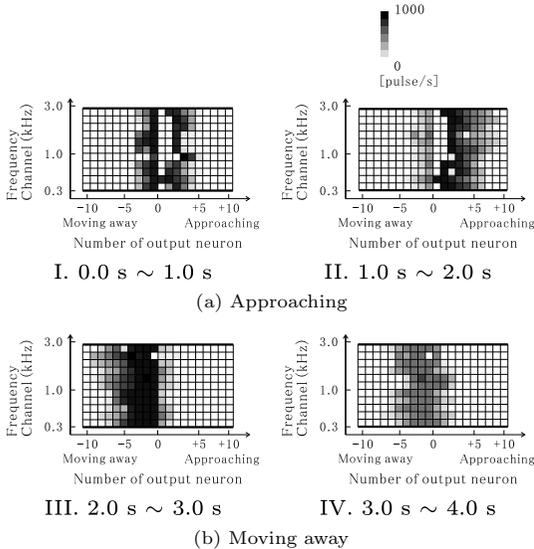


図 8 各時間ごとにおける音圧差検出部の出力結果
Fig. 8 Level difference extractor output of the “scooter” dataset.

ラメータ以外のパラメータの値に従い FPGA 上に各部を実装した．実装の際に必要な各部の回路数を表 3 に挙げる．合計が 43,100 [ALUTs] であるため使用した開発ボードにすべての回路を搭載することが可能である．

4.2 音圧差検出実験

音源の接近検出を行うには音圧差検出部において現在の音信号と一定時間前の同じ音信号との音圧差を検出する．学習音の一つである“スクーターのエンジン音”に対する各時間間隔ごとの音圧差検出部の出力を図 8 に示す．なお音圧差を検出する際に与える時間遅れ Δt は予備実験の結果から十分な音圧差が得られる最小時間として 0.4 [s] に設定してある．図の横軸は音圧差検出部の出力ニューロン番号 n ，縦軸に周波数帯域を示す．図の横軸の数値が大きいほど現時刻の入力 $x_f(t)$ が遅れ時間差を与えた入力 $x_f(t - \Delta t)$ より大きな音圧差であることを示す．また横軸のうち + 側は $x_f(t)$ が $x_f(t - \Delta t)$ より大きいことを表し，音源

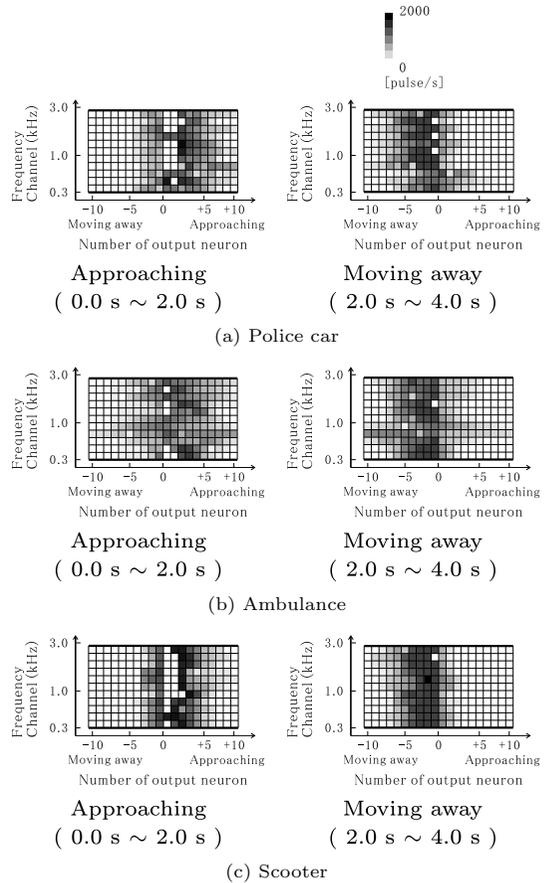


図 9 各入力音ごとにおける音圧差検出部の出力結果
Fig. 9 Comparing the output of level difference information for each dataset.

が接近している (Approaching) ことを示す．逆に横軸の - 側は $x_f(t)$ が $x_f(t - \Delta t)$ より小さいことを表し，音源が離脱している (Moving away) ことを示す．

結果より，音源の接近時 (0.0~2.0s) には横軸 + 側のニューロンが多く発火している．更に音源がマイクにより近い 1.0~2.0s のときの方が横軸の数値が大きいニューロンが発火し，より接近していることを示している．逆に離脱時には - 側のニューロンが多く発火しており，音源がマイクに近い 2.0~3.0s のときの方が数値が小さいニューロンが発火している．以上のことから，音源の移動に伴い音圧差検出部が正しく音圧差を検出できているといえる．

次に接近時及び離脱時における各学習音ごとの音圧差検出部の出力を図 9 に示す．例えば“救急車のサイレン音”であれば約 800 Hz の周波数帯域において横

表 4 音源種類と接近識別部の認識結果

Table 4 The results of sound recognition and approaching detection.
(A = Approaching, M = Moving away)

Input Sound		Recognition Rate [%]							
		police car		ambulance		scooter		other	
		A	M	A	M	A	M		
police	A	<u>99.9</u>	0.0	0.0	0.1	0.0	0.0	0.0	0.0
	M	0.7	<u>99.3</u>	0.0	0.0	0.0	0.0	0.0	0.0
ambulance	A	3.2	0.0	<u>93.4</u>	0.0	3.2	0.0	0.0	0.0
	M	0.0	0.0	1.0	<u>99.0</u>	0.0	0.0	0.0	0.0
scooter	A	0.6	0.0	0.0	0.0	<u>99.6</u>	0.0	0.0	0.0
	M	0.0	0.0	0.0	0.0	1.8	<u>98.2</u>	0.0	0.0



図 10 録音マイクの取付位置

Fig. 10 The position of recording microphone.

軸の数値の大きなニューロンが多く発火しているように、それぞれの音に関して大きな音圧差を示す周波数帯域が異なることが分かる。そのため、音圧差検出部の出力パターンから音源の種類を推定することができると予測できる。

4.3 音源種類及び接近識別実験

音源種類の識別を行うにはまず音圧差検出部の出力を音圧差情報量子化部へ入力しベクトル量子化を行う。次に量子化されたパターンを音源種類及び接近識別部に与えて音源の種類識別と接近の検出を行う。表 4 に学習音を入力した場合の FPGA 上における音源種類及び接近識別部の各ニューロンの出力頻度の割合を示す。下線は割合が最大となる出力結果を示す。

結果から学習用に用いた各入力音に対して適切なニューロンが最も高い頻度で出力しており、各音の接近・離脱を正しく学習し識別できているといえる。

次に本システムが実環境で利用できるかどうかを検証するために、実際に得られる音に関して同様に接近検出及び音源種類の識別実験を行った。ただし緊急車両に関しては音を収録することが困難であるため、本論文では自転車に対して二輪車が接近する状況のみに限定して実験を行った。

実験環境としては、晴天時のアスファルト上で行い、図 10 のように自転車の両側のルーフに SHURE 社製無

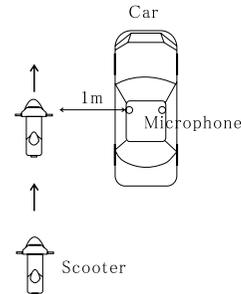


図 11 路上音の録音状況

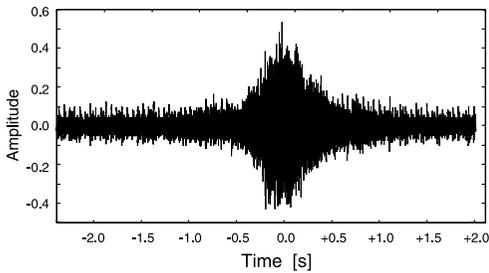
Fig. 11 Recording situation on the road.

指向性マイク “WL-93” を取り付ける。そして図 11 のようにアイドリング状態の自転車に対して左後方からスクーターが接近し通過するまでの音を録音する。

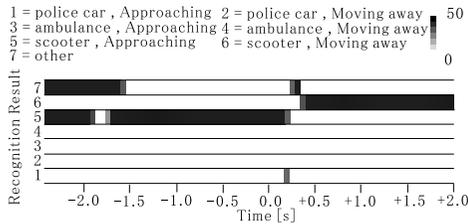
録音した音に関して提案手法を用いて識別実験を行った結果の一例を図 12 (b) に示す。図の横軸は時間であり音源が真横にきた時間を 0.0 [s] とする。縦軸は振幅と識別結果であり識別結果の濃淡は各ニューロンの出力パルス頻度を示し黒いほど多くのパルスを出していることを表す。接近と離脱の識別が変わる時間が実際の音より遅れているが、これは音圧差検出部では一定時間前の入力と比較しているためである。

また学習データとして計算機上で仮想的に設定した音を用いたが、実験結果のように実環境で得られる音に関して接近検出と識別ができている。これは識別部において勝者ニューロンのみが発火するように制御しているため、環境により入力音が変化しても、正答に対応する学習音と入力音との間の類似度が他の学習音よりも高ければ、正答に対応するニューロンのみが発火し正しく識別できるためと考えられる。このため提案手法は動作環境がある程度変化しても重みやパラメータを再設定することなく対応することができる。

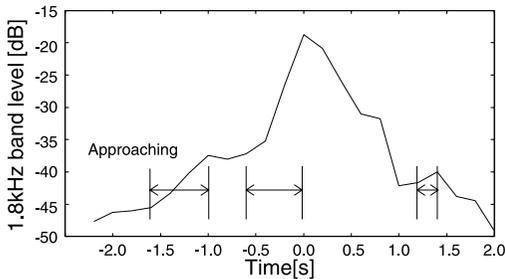
続いて比較実験として、まず接近検出については大塚らの手法 [6] や星野の提案しているレベルしきい値またはレベル上昇度から判断し検知する手法 [7] のように音源のもつ特定の周波数信号のレベルが過去の時間より上昇することから接近と判断する手法について同様の実験を行った。スクーターの特徴周波数として 1.8 kHz を選択しこの周波数帯域におけるレベルの時間変化を図 12 (c) に示す。更に背景雑音と区別するためにしきい値を -45 dB とし、しきい値以上でありかつレベル値が過去の値より上昇している区間も図に示す。この区間では音源が接近していると判断する。結



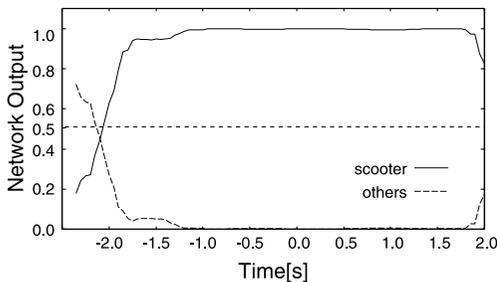
(a) 入力音



(b) 提案手法における出力結果



(c) レベル判断による手法の結果



(d) 周波数スペクトル識別による手法の結果

図 12 路上音における出力結果
Fig. 12 The results on the road.

果では -1.6 [s] より接近を検知できている．しかしスクーターが通りすぎた $+1.2$ [s] から $+1.4$ [s] においても接近と検出してしまうことになる．これは 1.8 kHzの周波数帯域のみのレベルの時間変化を参照しているためであると考えられる．

次に音源の検出と識別を行う手法として、星野の提

案しているニューラルネットワークによる走行音の周波数スペクトルから検知する手法 [7] についても実験を行った．実験条件としては文献 [7] を参考に、本手法と同じく 3 種類の音源の識別も同時に行うために入力層のニューロン数を $I = 60$ 、中間層、出力層のニューロン数を 3 倍である $M = 15$ 、 $O = 3$ へと増やした 3 層の階層型ニューラルネットワークを用いると仮定し、文献 [7] と同じく出力値を評価しやすくするために過去 0.5 [s] の間の出力を平均してスムージング処理を行い最終的な出力値とした．このときのスクーターを示すニューロンの出力値とその他のニューロンの出力値の合計を図 12 (d) に示す．結果より [7] の手法では出力値が 0.5 以上である -2.0 [s] から“スクーター”として音源の接近を検出できる．ただし手法の都合上、出力値からは音源の接近・離脱及び停止を直接区別することはできず、これらを区別するにはレベル値による方法と併用したり出力値の時間的変化を見る必要がある．

これらの手法と比較して提案手法では、 -1.5 [s] より以前では“その他”を示すニューロンも交互に発火しており“スクーターの接近”として確実に識別できる時刻が従来手法より遅い．この理由は、提案手法では音源の接近検出及び種類の識別を同時に行うために周波数帯域ごとの音圧の変化を用いているが、スクーターが遠方にあるときには音の音圧変化が小さく、接近として識別ができないためである．しかし比較した従来手法においても接近及び離脱を検出するためには時間的な音圧の変化を計算する必要があり、この場合提案手法と同様に検出可能距離が短くなることが考えられる．また、人間が物体を認知してから対応するまでの時間は通常 1.0 [s] であるといわれており [2]、提案手法の実験結果でも 1.5 [s] 前より接近の検出ができていたため提案手法は検知システムとしては問題ないといえる．

以上の比較実験から、提案手法は従来手法では個別に行っていた音源の接近及び離脱の検出と音源種類の識別を同時に行うことができ、また接近検知システムとして十分な識別精度をもつことが確認できた．

4.4 処理時間及び回路数の比較

交通安全装置の開発において重要な点として、出力を得るまでの応答時間がある．本システムにおける応答時間を検証するため、入力音として音長が 4.00 [s] のデータを用い本システムを動作させた場合の、FPGA における各部の単独の処理時間及び全動作を行った場

表 5 各部の処理時間の比較
Table 5 The time cost of each unit.

部位	提案手法		従来手法
	計算機 [s]	FPGA [s]	
入力音長	4.00	4.00	4.00
音圧差検出部	11.20	0.40	
音圧差情報量子化部	8.96	0.38	
音源種類及び接近識別部	0.42	0.37	
全動作	20.58	0.40	0.71

合の処理時間を表 5 に示す．また比較のために，計算機シミュレーションにおける本手法の処理時間と，従来手法として星野の提案しているニューラルネットワークによる走行音の周波数スペクトルから接近を検知する手法 [7] についても処理時間を示す．なお従来手法については先の実験と同じく入力層，中間層，出力層のニューロン数を $I = 60$ ， $M = 15$ ， $O = 3$ と仮定した 3 層の階層型ニューラルネットワークを用いた．また計算機シミュレーションで利用した環境は表 1 のとおりである．ただし各手法ともに FFT による周波数変換などの識別部に入力するデータを生成するまでの時間は考えない．

表 5 より，計算機を用いて処理を行った場合，提案手法は従来手法と比べて識別のための複数の手順があり順次計算する必要があるため計算時間が長くなる．しかし提案手法は FPGA 上に実装することにより並列に計算することができるため大幅な速度の短縮ができていくことが分かる．よって FPGA 上に実装することにより提案手法でも従来手法と同じく入力音に対して十分短い時間で出力が得られ，リアルタイムに動作を行うことができる．

従来手法もハードウェア上に実装することで並列に計算が可能となり高速な計算が可能であると予測される．そこでそれぞれの手法における周波数変換などを含むすべての処理に必要な演算から，ハードウェア上にシステムとして実装した場合に必要な回路数の概算を表 6 に示す．従来手法では FFT の窓長さ N を 256 とし計算に必要な係数はあらかじめ用意されているものとする．このため積算の回数は正規化を含めて $N \log_2 N + N \rightarrow 2,304$ だけ必要とし，加算は $1 + N \rightarrow 257$ だけ必要となる．ニューラルネットワーク部分では各ニューロン数が入力層 $I = 60$ ，中間層 $M = 15$ ，出力層 $O = 3$ であるため積算が $I \cdot M + M \cdot O \rightarrow 955$ だけ行われ，しきい値による比較も加算と同等の演算とすると総和と比較により加算の総数は $2M + 2O \rightarrow 36$ となる．これに対して提

表 6 各手法における並列化した際に用いる演算回路の個数

Table 6 The circuit number of pallareled operation in each method.

	提案システム	従来システム
	周波数-パルス変換部	FFT + 正規化
積算	308	2,304
加算	132	257
	PN モデル部分	ニューラルネットワーク部分
積算	0	945
加算	7,598	36
必要回路数	196,060 [ALUTs]	768,203 [ALUTs]

案手法では，使用する周波数帯域の数を $F = 11$ ，音圧差情報量子化部での量子化数を $Q = 23$ ，音源種類及び接近識別部での識別数を $C = 7$ とすると，周波数-パルス変換部で積算回数が $12F \rightarrow 132$ ，加算回数が $28F \rightarrow 308$ となる．続いて PN モデルにて実装された部分に関しては積算を必要とせず，加算が音圧差検出部で $174F \rightarrow 1,914$ ，音圧差情報量子化部で $(Q + 1)(20F + 4) + 2(Q + 8) \rightarrow 5,438$ ，音源種類及び接近識別部で $(C + 1)(Q + 4) + 2(C + 8) \rightarrow 246$ だけの加算が必要となる．

表 6 より提案システムの方が全回路の個数は多くなる．しかし提案システムでは従来システムと比較して積算の個数が少ないことが分かる．16 bit 固定小数点同士の二つの数について加算と積算を行う回路を，本研究で実装に用いた開発ソフトウェアである Altera 社の QuartusII を使用して論理合成した場合，加算を行う場合に必要な回路数が 16 [ALUTs] になるのに対して，積算の場合には 235 [ALUTs] の回路が必要となる．よって表 6 に示すように実装を行った場合，必要な回路数の概算は提案システムは 196,000 [ALUTs] 程度，従来システムは 768,000 [ALUTs] 程度となり最終的な回路規模では提案システムの方が小さくできる．なお本実験では文献 [7] を参考に従来手法の中間層のニューロン数を 15 個と仮定したが，学習の状況によってはどちらの手法においても更に多くのニューロン数を必要とすることがあり得る．この場合従来手法では積算の回数が増えるため，提案手法と従来手法の回路数の差は更に広がり，提案手法の方が有利になる．

以上の実験から，提案システムは従来システムでは個別に行っていた音源の接近と種類の識別を同時に行うことが可能であり，かつ従来システムより少ない回路数でハードウェアに実装でき，入力信号に対してリアルタイムに結果を出力できることが確認できた．

5. む す び

本論文では、自転車に対して接近する二輪車や緊急車両を音を用いて検知する方法として、FPGA 上に実装されたパルスニューロンモデルを用いて新たに音圧差検出モデル、競合学習ネットワークを構成し、入力される音情報から周波数ごとに音源の時間的な音圧差の情報を抽出することで音源の接近の検出と接近する音源の種類を識別するシステムを提案した。実験の結果、提案システムが従来システムでは個別に行っていた音源の接近・離脱の検出と接近する音源種類の識別を同時に行うことができ、かつ従来システムより少ない回路数で実現できることを示した。

音による車両の接近検出手法は障害物の多い街路地や交差点、夜間などの環境において有利であるが周囲の雑音が非常に大きい場合などには不利である。結果として画像処理のような視覚的手段と音を用いた聴覚的手段を併用することが最適であると考えられる。視覚的・聴覚的手段を併用する場合にはそれぞれのシステムの装置規模ができるだけ小さく安価であることが望ましい。実験結果で示したように、本システムは従来手法より回路規模が小さく併用する際にも有効な方法であるといえる。また基本素子である PN モデルの構成は先に提案した音源定位と音源種類の識別を行う回路 [9] と同じであり、識別手法ごとに個別の回路を製作する必要がなく一つの装置にまとめることが容易であるため、最終的な装置の規模を小さくできる。加えて併用時には各識別結果を相互に反映させることが容易であり、性能の向上も期待できる。

本研究に関する今後の課題として、本研究では計算機を使用した周波数-パルス変換部もハードウェア上に実装し、実際に搭載して実験することが挙げられる。特に、実際の環境では背景雑音が非常に大きな場合や接近する車両が複数ある場合など音環境が大きく変化するため、新たに複数音源を分離する機構などを適用する必要があると考えられる。

謝辞 本研究の一部は財団法人堀情報科学振興財団、文部科学省科学研究費・若手研究 (B) による支援を受けた。記して感謝致します。

文 献

- [1] 警察庁, 国土交通省, “交通安全マップ”
<http://www.kotsu-anzen.jp/index.html>
- [2] 佐藤 武, 自動車工学全書 16 自動車の安全, 山海堂, 1980.
- [3] 松本美智子, 須藤 智, 小沢慎治, “交差点通行時におけ

る危険警報のための接近車両検出” 電学論 (C), vol.122, no.3, pp.512-522, 2002.

- [4] G. Surendra, M. Osama, F.K.M. Robert, and P.P. Nikolaos, “Detection and classification of vehicles,” IEEE Trans. ITS, vol.3, no.1, pp.37-47, 2002.
- [5] W. Chieh-Chi, C. Thorpe, and S. Thrun, “Online simultaneous localization and mapping with detection and tracking of moving objects: Theory and results from a ground vehicle in crowded urban areas,” Proc. ICRA 2003, pp.842-849, 2003.
- [6] 大塚紳一郎, 原 寛徳, 小沢慎治, “車載マイクロフォンによる緊急車両の存在と方向検知システム” 電学論 (D), vol.124, no.4, pp.388-395, 2004.
- [7] 星野博之, “ドライバ支援のための走行音による接近車両検知システム” 音響誌, vol.62, no.3, pp.265-274, 2006.
- [8] W. Maass and C.M. Bishop, Pulsed Neural Networks, MIT Press, 1998.
- [9] 岩佐 要, 黒柳 奨, 岩田 彰, “FPGA を用いたパルスニューロンモデルによる音源定位及び音源種類の識別システムの実装” 信学論 (D), vol.J90-D, no.11, pp.3079-3090, Nov. 2007.
- [10] 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.
- [11] 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ネットワーク” 信学論 (D-II), vol.J87-D-II, no.7, pp.1496-1504, July 2004.
- [12] 黒柳 奨, 岩田 彰, “パルスニューロンモデルのための教師あり学習則” 信学技報, NC97-151, 1998.

(平成 19 年 7 月 31 日受付, 10 月 12 日再受付)



岩佐 要 (学生員)

平 15 岐阜高専専攻科・電子システム工学専攻了。平 17 名工大大学院工学研究科情報工学専攻博士前期課程了。現在, 名工大大学院工学研究科情報工学専攻博士後期課程に在学中。ニューラルネットワーク, 時系列情報処理に関する研究に従事。



藤角 岳史

平 15 名工大・電気情報工学卒。現在, 同大学院工学研究科情報工学専攻修士課程に在学中。



マウリシオ クグレ

平 13 Electrical Engineering, Federal Center of Technological Education, Brazil 卒。平 15 MSc in Biomedical Engineering, Federal Center of Technological Education, Brazil 了。平 19 名工大大学院工学研究科情報工学専攻博士後期課程了。同年名工大大学院・助教, 現在に至る。マシンラーニング, 大規模分類問題, 生医学信号処理, パルスニューロンモデルのハードウェア化に関する研究に従事。博士(工学)。IEEE 会員。



宮治 正廣

昭 46 名大・工・応用物理卒。昭 48 同大大学院修士課程了。昭 48 トヨタ自動車(株)入社, 技術開発部門に所属し, 現在に至る。車両の安全システム, マルチメディア車載機器, ボデー機能部品の企画・開発に従事。自動車技術会会員。



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, 平 18 同大学院・助教授, 平 19 同大学院・准教授, 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セ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。



段野 幹男

昭 52 早大・商卒。昭 55 インドボンベイ大学経済学部修士課程了。平 13(株)トヨタ IT 開発センターシニアリサーチャー, ビジネス企画部に所属し, 技術調査を担当。ITS ジャパン, 自動車技術会会員。