

同時事後確率推定ニューラルネットを用いた双腕動作識別法

島 圭介*・平松 侑樹**
芝 軒 太 郎**・辻 敏 夫**

Discrimination of Dual-arm Motions Using a Joint Posterior Probability Neural Network

Keisuke SHIMA*, Yuki HIRAMATSU**,
Taro SHIBANOKI** and Toshio TSUJI**

This paper proposes a novel dual-arm motion discrimination method combining of the posterior probabilities estimated independently for left and right arm motions. In the proposed method, first, only the posterior probability of each single-arm motion is estimated through learning from measured biological signals using recurrent probabilistic neural networks. The posterior probabilities output from the recurrent probabilistic neural networks are then combined based on motion dependency between arms, making it possible to calculate a joint posterior probability of dual-arm motions. With this method, all the dual-arm motions consisting of each single-arm motion can be discriminated through learning of single-arm motions only. In the experiments, the proposed method was applied to discrimination of 15 dual-arm motions which consist of three right-arm motions, three left-left arm motions and nine combined dual-arm motions. The results showed that the proposed method could achieve high discrimination performance though learning of three motions for each arm only (average discrimination rates: $97.49 \pm 2.37\%$). In addition, the possibility of applying the proposed method for a human interface was confirmed through operation experiments for the glovebox system.

Key Words: neural network, gesture discrimination, dual-arm motion, acceleration signal

1. はじめに

身振りや手振りなどのジェスチャは非言語コミュニケーションとしての意思伝達が可能であり^{1),2)},このようなジェスチャを自動識別することができれば意思伝達のツールや機器制御インタフェースの入力手段として利用できる可能性がある。しかしながらジェスチャで表現できるものは感情や単語, アルファベットなど多岐にわたることから, その対応関係を自動的に識別するためにはジェスチャの正確な計測と識別が必要となる。

従来, 各種センサを用いてジェスチャ時の身体運動に伴う生体信号を計測し, 自動識別を行なう研究が数多く行なわれてきた^{3)~10)}。ジェスチャには手や腕を利用して作る形状そのものを情報量として用いた静的なパターンと, 手や腕の時系列の動きを情報量として用いる動的なパターンがある。静的なパ

ターンを対象とするものには, 岡本ら³⁾のシェイプセンサと確率ニューラルネット (Probability Neural Network, 以下 PNN と略記) を用いた 31 種の指形状識別や, Bowden ら⁴⁾の動画像と隠れマルコフモデル (Hidden Markov Model, 以下 HMM と略記) を用いた 26 種のジェスチャ識別などがある。

一方, 動的なパターンのジェスチャ識別に関する試みもさまざま行なわれており, たとえば Sawada ら⁸⁾は, 計測した加速度信号 (acceleration signal, 以下, ACC 信号と略記) から腕の速度や移動距離・動作角度などを抽出し, DP (Dynamic Programming) マッチング法を用いて 10 種類のジェスチャの識別を実現した。また Vogler ら⁹⁾はカメラを用いて腕の 3 次元位置の変位や速度, 動作角度などから HMM やパターンマッチングを用いることで 53 種のジェスチャを識別している。ただし, これらは識別対象が単腕の動作のみである。ジェスチャの多くは双腕の動きが複雑に組み合わせられて表現されるため, 自動識別には単腕動作のみならず双腕動作を精度よく識別できる必要がある。

これに対し西野ら¹⁰⁾はデータグローブを用いて 3 次元空間内における手の位置や向き, 指関節の曲げ角度など多数の特徴を抽出し, リカレント型の NN を用いて双腕を用いた動的な 18 種のジェスチャ識別を実現した。ただし, この手法では識別精度向上のために大量の学習データを必要とし, 学習に時間がかかってしまう問題がある。また単一の NN を用

* 横浜国立大学大学院工学研究院
横浜市保土ヶ谷区常盤台 79-5

** 広島大学大学院工学研究院 東広島市鏡山 1-4-1

* Graduate School of Engineering, Yokohama National University, 79-5 Tokiwadai, Hodogaya-ku, Yokohama

** Graduate School of Engineering, Hiroshima University, 1-4-1 Kagamiyama, Higashihiroshima

(Received September 1, 2011)

(Revised February 12, 2013)

いているため、手の形状や向きが類似した動作など識別対象とする動作数が増加すると識別率が低下してしまう可能性がある。各腕の動きがさまざまに組み合わせられて実現される数多くの双腕動作を精度良く識別するためには、これらの点を解決できる新しい方法論が必要である。

本論文では双腕動作識別のための新しい PNN を提案し、これを用いた双腕動作識別法を構築する。提案する PNN は双腕動作識別の際に各腕の動作を個別に推定し、あらかじめ学習させた各腕の動作に対する事後確率を算出する。そして人間が両腕を動かす際、点対称や線対称以外の方向に同時に動かすことが困難であること¹¹⁾を考慮し、各腕の従属性に基づいた確率修正を行なうことで、双腕動作の同時事後確率を算出する。これにより、PNN に各腕の動作のみを学習させれば双腕動作が識別可能となるため、動作学習に要する時間や手間を増加させることなく、識別対象動作を増加できる。

以下、2. で提案する PNN について説明し、3. では提案 PNN を用いた双腕動作識別法について述べる。そして4. で提案法によるジェスチャ識別実験および機器制御実験の結果を述べ、提案 PNN の有効性を示す。

2. 同時事後確率推定ニューラルネットワーク

双腕動作は各腕の動きが複雑に組み合わせられて実現される。本論文では各腕の動作を個別に推定し、双腕動作の同時事後確率を推定可能な PNN を提案する。以下、提案する PNN を同時事後確率推定ニューラルネットワーク (Joint posterior Probability Neural Network: J-PNN) と呼び、そのネットワーク構造および学習則をそれぞれ説明する。

2.1 ネットワーク構造

Fig. 1 にネットワークの構成を示す。J-PNN は時系列パターン識別部と確率結合部から構成される。時系列パターン識

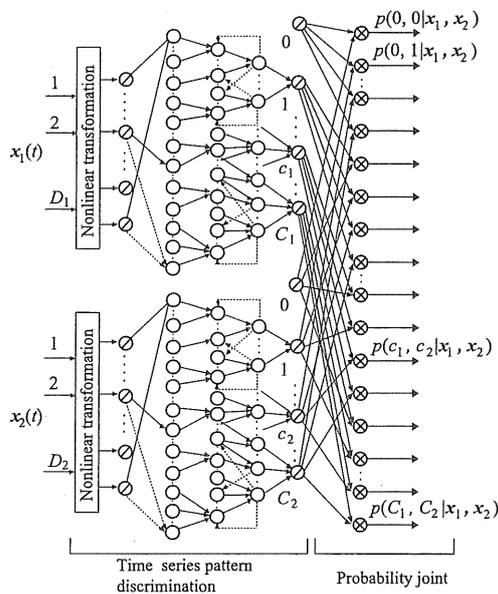


Fig. 1 Structure of the joint posterior probability neural network

別部は2つの時系列ベクトルを個別に入力し、各入力ベクトル $x_\alpha(t)$ ($\alpha \in \{1, 2\}$) の各識別対象クラス c_α ($c_\alpha = 1, \dots, C_\alpha$) に対する事後確率を算出する。確率結合部では入力間の従属性に基づいて事後確率を修正し、同時事後確率を算出する。提案ネットワークを用いることで2つの事象を組み合わせる表現される全クラスに対する同時事後確率を算出することができる。たとえば左右の腕の動作を個別に学習・推定することで、その同時事後確率から双腕動作の推定・識別が可能である。以下、各部の詳細について説明する。

(1) 時系列パターン識別部

時系列パターン識別部は混合正規分布モデル (Gaussian Mixture Model, 以下 GMM と略記) と HMM を内包したりカレント構造を有する PNN である Recurrent Log-Linearized Gaussian Mixture Network¹²⁾ (以下、R-LLGMN と略記) に基づいて構成する。これにより時系列データが従う統計分布を学習的に推定する。まず2つのベクトル $x_\alpha(t) \in \mathbb{R}^{D_\alpha}$ を入力する。 $\alpha \in \{1, 2\}$, $t = 1, 2, \dots, T$ であり、 T は時系列長である。つぎに $x_\alpha(t)$ それぞれに対し、GMM を構成する各コンポーネントに対応する正規分布を線形演算で表現するため次式に示す変数変換を行なう。

$$X_\alpha(t) = [1, x_\alpha(t)^T, x_{\alpha,1}(t)^2, x_{\alpha,1}(t)x_{\alpha,2}(t), \dots, x_{\alpha,1}(t)x_{D_\alpha}(t), x_{\alpha,2}(t)^2, x_{\alpha,2}(t)x_{\alpha,3}(t), \dots, x_{\alpha,2}(t)x_{D_\alpha}(t), \dots, x_{\alpha,D_\alpha}(t)^2]^T \quad (1)$$

各入力ベクトルに対し識別を行なうネットワークの第1層は変換された入力ベクトル $X_\alpha(t) \in \mathbb{R}^{U_\alpha}$ の次元数 $U_\alpha = 1 + D_\alpha(D_\alpha + 3)/2$ 個のユニットから構成されている。各ユニットの入出力関数には恒等関数を用いており、第1層の入力 $X_{\alpha,u_\alpha}(t)$ ($u_\alpha = 1, \dots, U_\alpha$) をそのまま⁽¹⁾ $O_{\alpha,u_\alpha}(t)$ として出力する。

つぎに第2層のユニット $\{c_\alpha, k_\alpha, k'_\alpha, m_\alpha\}$ ($c_\alpha = 1, \dots, C_\alpha$; $k_\alpha, k'_\alpha = 1, \dots, K_{\alpha,c}$; $m_\alpha = 1, \dots, M_{\alpha,c,k}$) は、第1層のユニットの出力に重み係数⁽²⁾ $w_{\alpha,k'_\alpha,k_\alpha,m_\alpha,u_\alpha}^{c_\alpha}$ が掛け合わされたものを入力として受け取る。第2層の入出力関係は次式で表現される。

$${}^{(2)}I_{\alpha,k'_\alpha,k_\alpha,m_\alpha}^{c_\alpha}(t) = \sum_{h_\alpha=1}^{U_\alpha} {}^{(1)}O_{\alpha,u_\alpha}(t) {}^{(2)}w_{\alpha,k'_\alpha,k_\alpha,m_\alpha,u_\alpha}^{c_\alpha} \quad (2)$$

$${}^{(2)}O_{\alpha,k'_\alpha,k_\alpha,m_\alpha}^{c_\alpha}(t) = \exp\left({}^{(2)}I_{\alpha,k'_\alpha,k_\alpha,m_\alpha}^{c_\alpha}(t)\right) \quad (3)$$

ここで C_α は識別対象となるクラス数、 $K_{\alpha,c}$ は HMM の状態数、 $M_{\alpha,c,k}$ はクラス c_α , 状態 k_α に対応する GMM のコンポーネント数を表す。

第3層のユニット $\{c_\alpha, k_\alpha, k'_\alpha\}$ への入力は、第2層のユニット $\{c_\alpha, k_\alpha, k'_\alpha, m_\alpha\}$ ($m_\alpha = 1, \dots, M_{\alpha,c,k}$) の出力を統合した⁽³⁾ $I_{\alpha,k'_\alpha,k_\alpha}^{c_\alpha}(t)$ であり、入力に1時刻前の第4層の出力を乗じた値が第3層の出力⁽³⁾ $O_{\alpha,k'_\alpha,k_\alpha}^{c_\alpha}(t)$ となる。

$${}^{(3)}I_{\alpha, k'_\alpha, k_\alpha}^{c_\alpha}(t) = \sum_{m_\alpha=1}^{M_{\alpha, c_\alpha, k_\alpha}} {}^{(2)}O_{\alpha, k'_\alpha, k_\alpha, m_\alpha}^{c_\alpha}(t) \quad (4)$$

$${}^{(3)}O_{\alpha, k'_\alpha, k_\alpha}^{c_\alpha}(t) = {}^{(4)}O_{\alpha, k'_\alpha}^{c_\alpha}(t-1) {}^{(3)}I_{\alpha, k'_\alpha, k_\alpha}^{c_\alpha}(t) \quad (5)$$

ただし、初期状態は ${}^{(4)}O_{\alpha, k'_\alpha}^{c_\alpha}(0) = 1.0$ とする。

さらに第4層の入出力 ${}^{(4)}I_{\alpha, k'_\alpha}^{c_\alpha}(t)$, ${}^{(4)}O_{\alpha, k'_\alpha}^{c_\alpha}(t)$ は、

$${}^{(4)}I_{\alpha, k'_\alpha}^{c_\alpha}(t) = \sum_{k'_\alpha=1}^{K_{\alpha, c}} {}^{(3)}O_{\alpha, k'_\alpha, k_\alpha}^{c_\alpha}(t) \quad (6)$$

$${}^{(4)}O_{\alpha, k'_\alpha}^{c_\alpha}(t) = \frac{{}^{(4)}I_{\alpha, k'_\alpha}^{c_\alpha}(t)}{\sum_{c'_\alpha=1}^{C_\alpha} \sum_{k'_\alpha=1}^{K_{\alpha, c'_\alpha}} {}^{(4)}I_{\alpha, k'_\alpha}^{c'_\alpha}(t)} \quad (7)$$

で与えられる。

最後に、第5層のユニット c は第4層の $K_{\alpha, c}$ ユニット $\{c_\alpha, k_\alpha\}$ ($k_\alpha = 1, \dots, K_{\alpha, c}$) の出力を統合したものである。入出力関係は、

$${}^{(5)}I^c(t) = \sum_{k_\alpha=1}^{K_{\alpha, c}} {}^{(4)}O_{\alpha, k_\alpha}^{c_\alpha}(t) \quad (8)$$

$${}^{(5)}O_\alpha^c(t) = {}^{(5)}I^c(t) \quad (9)$$

となる。この出力 ${}^{(5)}O_\alpha^c(t)$ がクラス c_α に対する事後確率 $p(c_\alpha | x_\alpha)$ ($\alpha \in \{1, 2\}$; $c_\alpha = 1, \dots, C_\alpha$) に相当する。

以上のように時系列パターン識別部は第3層と第4層の間にリカレント結合を有することで、1時刻前の出力を考慮することが可能である。これにより、各パターンの時系列的な繋がりを利用可能とし、各クラス間の重複が大きいパターン分布に対しても精度よい学習が可能となる。また第1層と第2層の間の重み係数を学習的に調節することで、各クラスに対する事後確率をGMMに基づいて統計的にモデル化することができる。

(2) 確率結合部

確率結合部では、時系列パターン識別部の出力である事後確率 $p(c_\alpha | x_\alpha)$ ($\alpha \in \{1, 2\}$; $c_\alpha = 1, \dots, C_\alpha$) を用いて、同時事後確率 $p(c_1, c_2 | x_1, x_2)$ を次式より算出する。

$$p(c_1, c_2 | x_1, x_2) = \prod_{\alpha \in \{1, 2\}} \frac{w_{\alpha, c_\alpha} p(c_\alpha | x_\alpha)}{\sum_{c_\alpha=1}^{C_\alpha} w_{\alpha, c_\alpha} p(c_\alpha | x_\alpha)} \quad (10)$$

ここで w_{α, c_α} は確率修正係数であり、次式で定義される事後確率のあいまいさであるエントロピー H_α を用いて表現される。

$$w_{\alpha, c_\alpha} = \begin{cases} 1 - H_\alpha & (H_\alpha > H_{th} | (c_1, c_2) \in \Xi) \\ 1 & (\text{otherwise}) \end{cases} \quad (11)$$

$$H_\alpha = -\frac{1}{\log C_\alpha} \sum_{c_\alpha=0}^{C_\alpha} p(c_\alpha | x_\alpha) \log p(c_\alpha | x_\alpha) \quad (12)$$

Ξ は2つのクラスの組 (c_1, c_2) を要素とし、識別対象クラスのうち一方のクラスの後事確率がもう一方のクラスの後事確率に影響を与える可能性がある組（以下、従属クラスと呼ぶ）を要素として含む集合で、この集合に含まれる従属クラ

スは事前に与えられているものと仮定する。すなわち、(11)式は Ξ に含まれる従属クラスのうち、各事後確率があいまいな場合にエントロピーを用いて事後確率を抑制することを意味する。(10), (11)式を用いることで、各クラスに従属性に基づいた同時確率の推定を行なう。なお、 $p(0 | x_\alpha)$ は次式で定義する余事象確率である。

$$p(0 | x_\alpha) = \begin{cases} 1 & (x_\alpha \in \{\phi\}) \\ 0 & (x_\alpha \notin \{\phi\}) \end{cases} \quad (13)$$

一般的にPNNは入力がない場合（未入力状態、 $x_\alpha \in \{\phi\}$ ）、事後確率を算出しない。そこで余事象確率 $p(0 | x_\alpha)$ は入力ベクトル x_α の状態に従って1または0をとるものとした。これによりたとえば一方の入力ベクトルがない場合、ネットワークの余事象確率が1となり、もう1つの入力ベクトルに対する事後確率がそのまま出力される。なお、J-PNNの識別対象クラス数 J は $J = (C_1 + 1)(C_2 + 1) - 1$ となる。

2.2 ネットワークの学習則

PNNの教師あり学習では、通常教師信号とネットワーク出力の誤差を最小化するように重みを修正する。J-PNNでは時系列パターン識別部を構成する各ネットワークの学習を行なうことでネットワーク全体の学習が可能である。

まず各ネットワークへの入力ベクトル $x_\alpha(t)^{(n)}$ ($n = 1, 2, \dots, N_\alpha$; $t = 1, 2, \dots, T_\alpha$ であり、 N_α は学習データサンプル数、 T_α は時系列長である) に対して、教師信号ベクトル $P_\alpha^{(n)} = [P_1^{(n)}, \dots, P_{c_\alpha}^{(n)}, \dots, P_{C_\alpha}^{(n)}]^T$ ($\alpha \in \{1, 2\}$) がそれぞれ与えられた場合を考える。 $x_\alpha(t)^{(n)}$ がクラス c_α に属するとき $P_{c_\alpha}^{(n)} = 1$ とし、それ以外は $P_{\hat{c}_\alpha}^{(n)} = 0$ ($\hat{c}_\alpha \neq c_\alpha$) とする。学習用データに対するネットワークの評価関数 E_α は、

$$E_\alpha = -\sum_{n=1}^{N_\alpha} \sum_{c_\alpha=1}^{C_\alpha} P_{c_\alpha}^{(n)} \log {}^{(5)}O_\alpha^c(T_\alpha)^{(n)} \quad (14)$$

と定義し、これを最小化するかわり対数尤度を最大化するように学習を行なう。ただし、 ${}^{(5)}O_\alpha^c(T_\alpha)^{(n)}$ は入力ベクトル $x_\alpha(t)^{(n)}$ に対する時刻 T_α での出力を意味している。

重みの更新には通時誤差逆伝播学習法 (BPTT)¹³⁾を用い、さらに学習を決められた時間内で収束させるためZak¹⁴⁾が提案したターミナルアトラクタを学習則に導入し、収束時間の上限値を設定可能とした。

3. 双腕動作識別法

Fig. 2にJ-PNNを用いた双腕動作識別法の構成を示す。提案法は双腕動作時に発生する各腕のACC信号から動作ごとの特徴を抽出し、J-PNNを用いて動作を識別する。

まず、被験者の各腕に3軸ACCセンサを1つずつ装着し、動作時のACC信号を計測する。計測したACC信号をA/D変換したものを $a_m^v(t)$ ($m \in \{r, l\}$, $v \in \{x, y, z\}$) (l は左手、 r は右手、また x, y, z はそれぞれACC信号の x, y, z 軸成分を表わす添字) とする。ここで身体運動に伴うACC信号は比較的低周波であるため、カットオフ周波数 f_c [Hz] のローパ

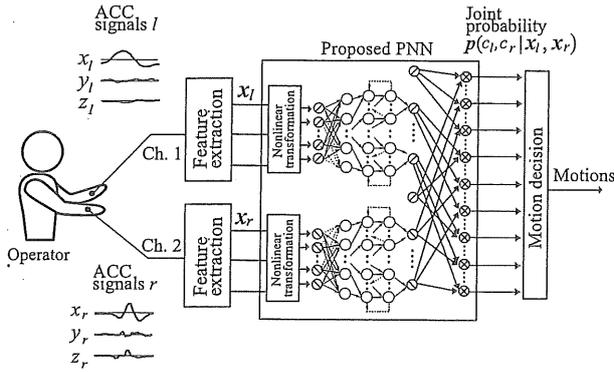


Fig. 2 Overview of the discrimination method for dual-arm motions

スフィルタを用いて低周波成分の信号 $g_m^v(t)$ を抽出する¹⁵⁾. さらに $g_m^v(t)$ の最大値がそれぞれ1となるように次式を用いて正規化する.

$$\hat{g}_m^v(t) = \frac{g_m^v(t)}{\max(g_m^v(t))} \quad (15)$$

ここで $\max(g_m^v(t))$ は事前に計測した $g_m^v(t)$ の最大値である. そして, 正規化した各腕の ACC 信号ベクトル $\hat{g}_m(t) = [g_m^x(t), g_m^y(t), g_m^z(t)]^T$ の大きさを次式で算出し, 動作発生 の判定を行なう.

$$\hat{g}_m(t) = \sqrt{\hat{g}_m^x(t)^2 + \hat{g}_m^y(t)^2 + \hat{g}_m^z(t)^2} \quad (16)$$

Fig. 3 に動作発生 の判定例を示す. ここでは, $\hat{g}_m(t)$ が Th_{start} 以上になる時刻を動作開始時刻 t_{start} とし, $\hat{g}_m(t)$ が Th_{end} [s] 連続して Th_{end} 以下になったときの最初の時刻を動作終了時刻 t_{end} とする. そして, 動作区間 $[t_{start}, t_{end}]$ における $\hat{g}_m(t)$ の最初の極大値から最後の極大値までを識別区間 $[t_{start'}, t_{end'}]$ とし, この識別区間の $\hat{g}_m(t) \in \mathcal{R}^3$ を識別用データとして抽出する.

つぎに, 動作ごとに識別用データの時系列長が S となるようリサンプリングを施し, 識別区間 $[t_{start'}, t_{end'}]$ における $\hat{g}_m(t)$ の最大値が1となるように正規化したものを特徴パターンベクトル $x_m(t)$ として J-PNN の入力に用いる. J-PNN は $x_m(t)$ と $x_{m'}(t)$ ($m' \in \{r, l\} \neq m$) を入力として識別を行なう. た

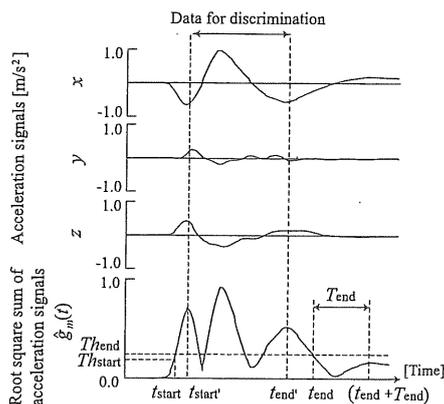


Fig. 3 An example of gesture detection results

だし各腕の動作には時間差が生じる可能性があるため, PNN に $x_m(t)$ が入力されてから T [s] 以内に $x_{m'}(t)$ が入力されなければ $x_{m'}(t) = \{\phi\}$ とする. ここで $c_m = 1, 2, \dots, C_m$ であり, C_m は各腕の識別対象クラス数である.

また前述のように, 人間が両腕を同時に動かす際に点対称や線対称方向以外へ動かすことが困難である点¹¹⁾を考慮し, 双腕動作においては両腕の対称運動が従属クラスであると仮定し, 本論文では, 各腕の動作 $c'_m = \operatorname{argmax}_{c_m} p(c_m | x_m)$ ($m \in \{r, l\}$) が $c'_r = c'_l$ となる組み合わせを集合 \mathcal{E} の要素とした. そして J-PNN の出力である同時事後確率 $p(c_r, c_l | x_r, x_l)$ が最大となるクラスを被験者が行なった動作として識別する.

4. 実 験

4.1 双腕動作識別実験

J-PNN の有効性を検証するため, 双腕を用いたジェスチャの識別実験を行なった. ACC 信号計測には3軸 ACC センサ (MicroStone 社製 MA3-04Ac) と無線モーションレコーダ (MicroStone 社製 MVP-RF) を用い, センサを被験者の両手首 (Ch. 1: 右手首, Ch. 2: 左手首) に医療用テープで固定した. 計測はサンプリング周波数 $f_s = 200$ Hz で行ない, ジェスチャ識別のための各パラメータは, $f_c = 10.0$ Hz, $Th_{start} = 0.2$, $Th_{end} = 0.25$, $T_{end} = 0.15$ s, $T = 0.4$ s, $S = 30$ とした. なお, 動作発生 の判定に関するパラメータは予備実験より試行錯誤的に決定した.

被験者は健康な男子大学生5名であり, 識別対象動作を各腕3動作の組み合わせにより実現できる全15動作 (Fig. 4 参照) として各動作を50試行行なわせた. 各被験者は本システムの操作経験はなく, 対象とする動作や実験順序の確認のため実験前に30分程度の練習時間を設けた. 実験ではまず, 各50試行のデータを無作為に学習データ10試行と未学習の識別精度検証用データ (検証データ) 40試行に分けた. J-PNN の時系列パターン識別部を構成する R-LLGMN は高い学習能力を有しており, 1試行の学習データを用いた場合でもある程度識別ができることが示されている¹⁶⁾. そのため本論文では, 10試行のデータから各腕基本となる3動作のデータを無作為に1試行ずつ選出したものを学習データとして用いた. なお, 提案法では各腕の動作を個別に推定し, 双

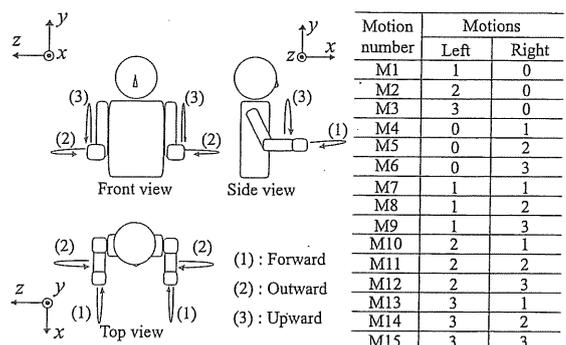


Fig. 4 One-armed motions of each hands

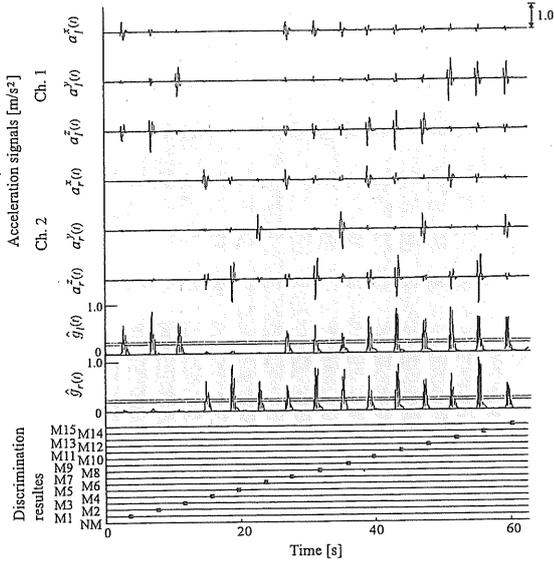


Fig. 5 An example of ACC signals and discrimination results

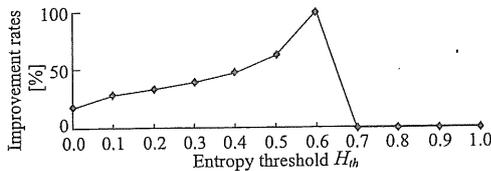


Fig. 6 Ratio of improvement in misdiscrimination data for each H_{th} (Subject A)

腕動作時の従属性を考慮して組み合わせることで双腕動作の識別を可能としているため、各腕3動作 (M1~M6) のみを学習すればよいことに注意されたい。ここで、予備実験より学習データの識別率は100%であり、高い精度で識別が行なえていることをあらかじめ確認している。

(1) 実験結果

Fig. 5 に識別対象動作 (M1~M15) を順に行なったときの識別結果を示す。図は上から、被験者の各腕より計測したACC信号の各軸成分 $a_m^n(t)$ とACC信号ベクトルの大きさ $\hat{g}_m(t)$ 、識別結果を表わしている。図より動作ごとに特徴の異なるACC信号が計測されており、J-PNNに入力することで正しく識別を行なえていることがわかる。また学習を行っていない動作 (M7~M15) に対しても識別が行なえている。これはM7~M15が学習動作の組み合わせで表現される動作であり、J-PNNによる同時事後確率の算出が正しく行なえたことを示している。

つぎに双腕動作識別のための適切なエントロピー閾値を設定するため、エントロピー閾値 H_{th} が識別精度に与える影響を調べた。ここでは、双腕動作の独立性を仮定したときに誤識別する試行数 Λ に対して、(10)式によって正しく識別される試行数 Ψ の比率 $\frac{\Psi}{\Lambda}$ (識別改善率) を算出した。

Fig. 6 に Sub. A の識別改善率を示す。図は H_{th} を0.0から1.0まで0.1刻みで変化させた場合における結果改善率を示している。図では H_{th} が0.6を超えたところで改善率が0

Table 1 Discrimination results of each motion (Subject A)

	M1	M2	M3	M4	M5	M6	M7	M8	M9
DR	99.55	100	99.94	100	100	100	99.80	99.77	97.15
SD	3.51	0	0.39	0	0	0	1.56	1.13	6.70
	M10	M11	M12	M13	M14	M15	I	II	III
DR	99.68	100	99.24	99.78	99.86	99.59	99.91	99.43	99.62
SD	2.86	0	1.80	1.28	0.58	1.08	0.18	0.88	0.72

DR: Discrimination rate [%] SD: Standard deviation [%]

I: Average discrimination rate of single-arm motions (M1~M6)

II: Average discrimination rate of dual-arm motions (M7~M15)

III: Average discrimination rate of all motions

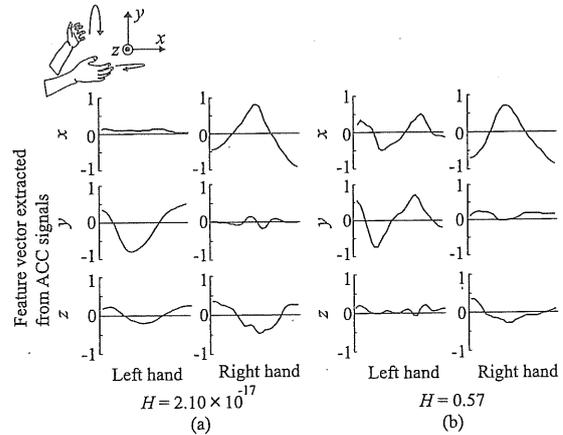


Fig. 7 Examples of ACC signals in a dual-arm motion (M13)

となっている。これは H_{th} を大きくしていくと修正対象となる試行数が少なくなり、確率修正によって結果が改善する試行の数が減少してしまうためである。 $H_{th} = 0.6$ において識別改善率が最大となることから、本論文では $H_{th} = 0.6$ として実験を行なうこととした。

Table 1 に $H_{th} = 0.6$ として識別を行なった場合の各動作に対する平均識別率を示す。ここでは、学習データによる識別結果への影響を低減するため、あらかじめ計測した50試行のデータを (i) 学習用データのセット (10試行分) と識別精度検証用データのセット (40試行分) に分割し、(ii) 学習用データセットの中から各腕の基本3動作のデータを1試行ずつ無作為に選出したものを学習データ1試行として定義した。そして (i) のデータセットの分割を10回、(ii) の学習データ1試行の選出を20回行ない、計200回の識別を行なった際の平均値を算出した。結果より平均識別率は $99.62 \pm 0.72\%$ となり、動作を精度よく識別できていることが確認できる。また双腕動作 (M7~M15) の平均識別率は $99.43 \pm 0.88\%$ であることから、学習を行っていない双腕動作に対しても高い精度で識別できることがわかる。

また、識別対象動作中、最も識別率が低い動作はM13 (左手をy軸方向に振る、右手をx軸方向に振る) であった。Fig. 7 は被験者にM13を行なうよう指示した際に得られた特徴ベクトルの例である。(a)の特徴ベクトルは腕を振った軸方向 (左手:y軸, 右手:x軸) の振幅が大きいことがわかる。一方、(b)は(a)と同じ動作を行なっているにもかかわらず、左

Table 2 Relationships between the number of correct and incorrect classification for the data belonging to the set of dependent classes Ξ with high entropy

Subject	A	B	C	D	E	Ave. \pm S.D.
The number of trials in dependent classes Ξ	58	107	110	140	93	101.6 \pm 29.8
The number of incorrect classification using Eq. (17)	46	100	92	106	83	85.4 \pm 23.7
The number of correct classification using Eq. (10)	46	100	83	100	76	81.0 \pm 22.2

手の y 軸だけでなく x 軸の特徴が大きくなっている。これは左手と右手の動きが対称ではないため同時に動かすことが困難となり、左手が右手の動き (x 軸方向へ振る動作) につられて動いてしまったと考えられる。ここで、双腕動作時の各腕の動きが互いに独立であると仮定すると、(10) 式の重み係数 $w_{\alpha, c_{\alpha}}$ はすべて 1 となり、以下のように簡略化できる。

$$p'(c_1, c_2 | x_1, x_2) = \prod_{\alpha \in \{1, 2\}} p(c_{\alpha} | x_{\alpha}) \quad (17)$$

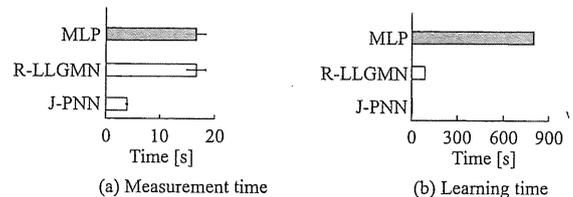
両腕を x 軸方向に振る動作は集合 Ξ に属する従属クラスとなるため、これらのデータに対して (17) 式を用いて識別結果を算出すると、(b) では M7 へ誤識別されてしまうのに対し、(10)、(11) 式を用いた J-PNN では正しく識別が可能である。これは J-PNN が各腕の従属性に基づいて確率修正を行なうため、各腕が対称でない動きに対しても正しく識別可能であることを示した結果である。同様に全動作において集合 Ξ に属し、かつエントロピーが高い試行のうち、(17) 式を用いた際に誤識別する数と、(10) 式を用いて正しく識別できる数をまとめたものを **Table 2** に示す。誤識別する 85.4 \pm 23.7 試行のうち、(10)、(11) 式を用いて 81.0 \pm 22.2 試行を正しく識別できた。このとき、従属性試行のうち (17) 式および (10) 式を用いて正しく識別できる試行数の割合を算出したところ、それぞれ 15.7 \pm 7.2%、80.2 \pm 8.3% となり、有意水準 0.1% で有意差が得られた。以上の結果から、双腕動作を用いたジェスチャ識別には各腕の従属性を考慮した識別が有効であることが明らかになった。ただし、確率修正を行わずに識別を行なった場合の全試行に対する識別率は 97.01 \pm 3.60% となり、従属性試行以外に対しては十分に高い精度で識別可能であることがわかる。

(2) 従来法との比較

J-PNN の識別能力を検証するため、一般的にパターン識別に用いられる NN である多層パーセプトロン (Multi Layered Perceptron, 以下、MLP と略記)、および Tsuji らによって提案された R-LLGMN¹²⁾ を用いて、全 15 動作を学習、識別した。使用した MLP は入力層、隠れ層 10、出力層の計 12 層構成とし、入力層のユニット数は 180 (3 次元 \times 時系列データ長 S: 30 \times 2)、隠れ層のユニット数は各層 200、出力層は識別動作数に合わせて 15 とした。入出力関数には対数シグモイド関数を利用した。学習則には誤差逆伝播法を用い、誤差閾値を 0.01、学習率を 0.01 とした。R-LLGMN は入力層 6、出力層 15 となり、その他のパラメータは J-PNN と同じとし

Table 3 Average discrimination rates of MLP, R-LLGMN¹²⁾ and J-PNN

Subject	A	B	C
MLP	60.68 \pm 29.39	56.34 \pm 20.39	62.04 \pm 15.85
R-LLGMN	93.90 \pm 6.61	94.33 \pm 5.34	87.97 \pm 9.13
J-PNN	99.62 \pm 0.72	99.02 \pm 2.93	97.58 \pm 3.82
Subject	D	E	Average
MLP	60.12 \pm 26.97	60.41 \pm 26.66	62.03 \pm 5.59
R-LLGMN	85.96 \pm 17.68	88.17 \pm 13.30	90.07 \pm 3.80
J-PNN	93.54 \pm 3.92	97.65 \pm 3.44	97.49 \pm 2.37

Ave. \pm S.D. [%]**Fig. 8** ACC measurement time and learning time of each neural network for all motions

た。PC は Intel Core (TM) i7 (クロック周波数: 3.3 GHz, メモリ: 3.0 GB) を使用した。

Table 3 に各被験者の識別結果の詳細を示す。表から、J-PNN が最も識別率が高く、MLP が最も低いことがわかる。MLP は動作数が増加したことによって学習が難しくなり、汎化性が低下している。これに対し、R-LLGMN および J-PNN は時系列信号の統計分布を GMM や HMM に従って推定することにより、識別率が向上したと考えられる。各手法の平均識別率はそれぞれ MLP が 62.03 \pm 5.59%、R-LLGMN が 90.07 \pm 3.80%、J-PNN が 97.49 \pm 2.37% であった。また、MLP および R-LLGMN を用いて各腕 3 動作ずつを学習させて全 15 動作の識別を行なったところ、平均識別率はそれぞれ 50.6 \pm 17.4%、97.01 \pm 3.60% となった。以上の結果より、双腕動作を構成する各片腕動作を個別に学習し、双腕動作時の従属性に従って組み合わせることで精度よい識別が実現できることを明らかにした。

Fig. 8 には各手法の学習にかかる時間 (学習データ取得に要した時間、および PNN の学習時間) の全被験者平均を示している。MLP および R-LLGMN は識別対象動作 (15 動作 \times 1 試行) をすべて学習する必要があるのに対し、J-PNN では各腕の基本動作のみを学習すればよいため、学習データ取得に要する時間を短縮できることがわかる (**Fig. 8 (a)**)。学習データ取得に要する時間は MLP および R-LLGMN が 16.60 \pm 1.77 s、提案法が 3.88 \pm 0.23 s であった。ここで、今回識別対象とした動作は左右対称の 3 動作を基本動作としており、両腕を「前 (x 軸) 方向に振る」、「上 (y 軸) 方向に振る」、「外 (z 軸) 方向に振る」という 3 動作を実施した際の各腕の計測データを学習に用いればよく、学習データ取得時間を低減可能である。また R-LLGMN や J-PNN は学習則にターミナルアトラクタを導入しており、収束時間の上限値を設定できるため学習時間も短い。さらに J-PNN では各腕 3 動作のみの学習でよいことから学習

時間が大幅に低下できている (Fig. 8 (b)). 学習時間は MLP が 804.30 ± 6.13 s, R-LLGMN が 87.48 ± 0.20 s, J-PNN が 1.18 ± 0.16 s であった. 以上の結果より, J-PNN を用いることで学習の時間や時間を大幅に削減できることを示した.

4.2 機器制御実験

提案法を応用して機器制御実験を行なった. 実験では提案法を用いてジェスチャを識別し, その結果をロボットなどの機器制御に用いる. これにより操作者の意図どおりに機器を制御する.

本論文では特殊環境下における作業状況を想定し, グローブボックス (glovebox) を用いた作業中に glovebox 内の環境制御とロボットマニピュレータの制御を行なう¹⁷⁾. glovebox はボックス内部の湿度や気圧を操作し, 内部気体を調整することで特殊な環境を作り出すことができる装置であり, ボックスに取り付けられたグローブを介することで作業者の安全を確保しつつ作業を行なうことが可能である. ただしグローブにより使用者の動作が制限され, 作業効率が低下する問題がある. そこで glovebox 内にロボットマニピュレータを搭載し, 提案法を用いて制御することで作業効率の改善を試みる. 以下, 構築した glovebox 制御システムと実験結果について述べる.

(1) glovebox 制御システム

glovebox は外部のコントローラ (タッチパネル) を用いて操作できる. glovebox の環境制御では, 識別動作をあらかじめボックス内の減圧や湿度の調整など, 制御項目の「選択」および「決定」の動作として割り当てておく. そして, 各制御命令の選択と決定を繰り返すことで, glovebox へ制御命令を送信しボックス内の環境制御を行なうことができる. ロボットマニピュレータは, プログラム制御方式とマニュアル制御方式の2つの制御方式を用いて制御できる. プログラム制御では, あらかじめ使用頻度の高い作業や反復的な作業を複数プログラムしておく. そして環境制御と同様に, プログラムの選択, 決定を行なうことでプログラムどおりの作業をマニピュレータに実行させる. マニュアル制御は, 識別動作に対しエンドエフェクタの X 軸, Y 軸, Z 軸への「移動」を割り当てる. また同じ動作をもう一度行なうことで停止させることができる.

(2) 操作例

glovebox (INKP-HM30, 美和製作所製) にロボットマニピュレータ (MELFA RV-1A, 三菱電機製) を搭載して実験を行なった. Fig. 9 に実験に用いた glovebox 制御システムのハードウェア構成を示す. 被験者は男子大学生 1 名 (21 歳) とし, ACC センサを両腕の尺側手根伸筋の皮膚表面に 1 つずつ取り付けた. 識別対象動作は各腕を「1: 前方向に振る」, 「2: 外側方向に振る」という 2 動作を組み合わせて表現可能な計 8 種類の双腕動作とした. そして, それぞれの動作を制御命令の選択, 決定およびマニピュレータの各軸方向への移動に割り当てた. 被験者には (A) ボックス内の蛍光灯を点灯し, (B) プログラム制御を用いて物体を指定位置へ移動後,

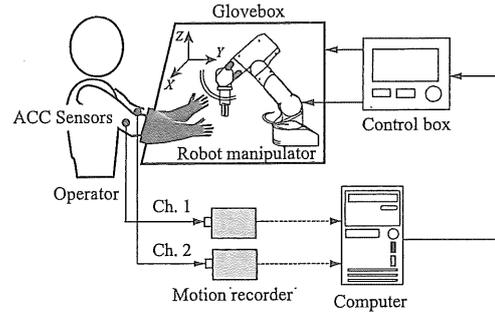


Fig. 9 Hardware components of glovebox control system

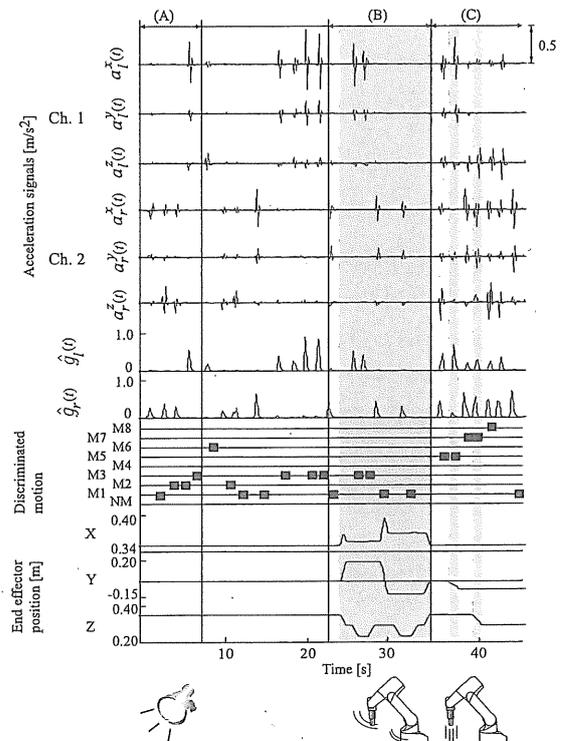


Fig. 10 An example of operation for glovebox control system

(C) マニュアル制御によるマニピュレータのエンドエフェクタ位置の微調整を順に行なうように指示した.

Fig. 10 に実験結果の一例を示す. 波形は上から, 計測された各センサの 3 軸 ACC 信号 $a_m^n(t)$, 各センサの ACC 信号ベクトルの合成波形 $\hat{g}_m(t)$, 識別結果, エンドエフェクタの座標を表わしており, 陰影部はマニピュレータの動作区間を表わしている. 被験者はまず蛍光灯をつけるため制御項目 (メニュー) の選択と決定を行ない, その後同様に制御項目の選択を繰り返し, 23s 付近でプログラム制御を行ないマニピュレータを移動させている. その後マニュアル制御に切り替え, 36s, 38s 付近で位置や高さを調整している. 結果より, glovebox 内の環境制御やロボットマニピュレータの位置調整など, 被験者の意図した制御が行なえていることが確認できる. また Table 3 からわかるように, MLP を用いた場合の識別率は低く, 動作識別が困難であった.

以上のように, 提案法を用いて双腕動作の識別を行なうこ

とで, glovebox システムなどの作業支援が可能である. 本論文では glovebox システムを対象としたが, 提案法はその他のさまざまなヒューマン・インタフェース^{8), 10), 18)}へも応用可能である.

5. まとめと今後の課題

本論文では, 時系列パターン識別部と確率結合部から構成される同時事後確率推定ニューラルネット (J-PNN) を提案し, 双腕動作識別に応用した. J-PNN では各腕の少ない動作学習のみで, それらを組み合わせた双腕動作が識別できるため, 学習時間・時間の軽減が可能である. また, 各腕間の従属性に基づく確率修正を行なうことで精度よいジェスチャ識別を実現できる. 実験では, 左右の腕の動きを組み合わせた全 15 動作を 97.49% の精度で識別できることを確認した. また機器制御実験により, 双腕動作を用いて被験者の意図した機器制御が実現できることを示した.

今後は被験者および識別対象動作を増加させ, 提案法の有効性を検証していく. また, エントロピー閾値 H_{th} などのパラメータを被験者ごとに適切に設定する方法について検討し, さらなる識別精度の向上を図る.

参考文献

- 1) 高田英一: 手話の森を歩く一言語としての手話 その秘密をさぐる, 全日本ろうあ連盟出版局 (2003)
- 2) 厚生労働省: 平成 18 年身体障害児・者実態調査結果, <http://www.mhlw.go.jp/index.shtml> (2006)
- 3) 岡本, 村上, 柴, 辻: 木構造を有する階層型確率ニューラルネットの提案と指形状識別への応用, 計測自動制御学会論文集, **43-2**, 128/134 (2007)
- 4) R. Bowden and M. Sarhadi: A non-linear model of shape and motion for tracking finger spelt American sign language, *Image and Vision Computing*, **20-9-10**, 597/607 (2002)
- 5) 坂口, 岡林, 金森, 井口: 運動覚センサと筋電位センサを併用したジェスチャ識別, 電子情報関連学会誌, **81-2**, 2385/2393 (1998)
- 6) B. Gonzalo, R. Daniel, T. Gerhard and T. Gracian: Real time gesture recognition using Continuous Time Recurrent Neural Networks, *ICST*, 978-963-06-2193-9 (2007)
- 7) 高橋, 関, 小島, 岡: ジェスチャー動画像のスポッティング認識, 信学論 D-II, **J77D-2-8**, 1552/1561 (1994)
- 8) H. Sawada and S. Hashimoto: Gesture Recognition Using an Acceleration Sensor and Its Application to Musical Performance Control, *Electronics and Communications in Japan, Part 3*, **80-5**, 9/17 (1997)
- 9) C. Vogler and D. Metaxas: Adapting Hidden Markov Models for American Sign Language Recognition by using Three-dimensional Computer Vision Methods, *IEEE International Conference on Systems, Man and Cybernetics*, 156/161 (1997)
- 10) 西野, 凍田, 宇都宮: オンライン学習機能を備えた対話型両手ジェスチャインタフェース, 信学論 D-II, **J81D-2-5**, 897/905 (1998)
- 11) 鎌倉矩子: 手のかたち手のうごき, 医歯薬出版 (1989)
- 12) T. Tsuji, N. Bu, O. Fukuda and M. Kaneko: A Recurrent Log-Linearized Gaussian Mixture Network, *IEEE Trans. on Neural Networks*, **14-2**, 304/316 (2003)
- 13) P.J. Werbos: Backpropagation through time: What it does and how to do it, *Proc. IEEE*, **78-10**, 1550/1560 (1990)
- 14) M. Zak: Terminal Attractors for Addressable Memory in Neural Networks, *Physics Letters A*, **133-1-2**, 18/22 (1988)
- 15) 辻, 吉久, 島: 加速度センサを利用した義手型マニピュレータの筋音制御, *日本ロボット学会誌*, **25-6**, 874/880 (2007)
- 16) 辻, 島, 村上: 筋シナジーに基づく複合動作のパターン識別, *日本ロボット学会誌*, **28-5**, 606/613 (2010)
- 17) 山田, 石井, 島, 辻, 山本, 和田: ロボットマニピュレータを搭載した EMG 制御型グローブボックスの開発, 第 8 回計測自動制御学会システムインテグレーション部門公演論文集, 389/390 (2007)
- 18) K. Shima, M. Okamoto, N. Bu and T. Tsuji: Novel Human Interface for Game Control Using Voluntarily Generated Biological Signals, *Journal of Robotics and Mechatronics*, **18-5**, 626/633 (2006)

[著者紹介]

島 圭介 (正会員)



2005 年広島大学工学部第二類卒業. 2007 年同大学大学院工学研究科博士課程前期修了. 2009 年同博士課程修了. 2007~2008 年日本学術振興会特別研究員 (DC1), 2009~2012 年同特別研究員 (PD). 2012 年より横浜国立大学大学院工学研究科助教, 現在に至る. 博士 (工学). 生体信号解析, ニューラルネット, ヒューマンインタフェースなどの研究に従事. 日本機械学会, IEEE などの会員.

平松 侑樹



2010 年広島大学工学部第二類卒業. 2012 年同大学大学院工学研究科博士課程前期修了. 在学中は生体信号解析, ニューラルネットに関する研究に従事.

芝 軒 太郎



2008 年徳島大学工学部知能情報工学科卒業. 2010 年広島大学大学院工学研究科博士課程前期修了. 2012 年同博士課程後期修了. 日本学術振興会特別研究員 (DC2) を経て, 現在, 同特別研究員 (PD). 博士 (工学). 生体信号解析, 次元削減, マン・マシン・インタフェースなどの研究に従事. IEEE などの会員.

辻 敏夫 (正会員)



1985 年広島大学大学院工学研究科博士課程前期修了. 同大学工学部助手, 同助教授を経て, 2002 年より同大学大学院工学研究科教授, 現在に至る. 工学博士. 人間とロボットの運動制御, 生体信号解析, ニューラルネット, ヒューマン・マシンシステムなどの研究に従事. IEEE, 日本人間工学会, バイオメカニズム学会, 日本機械学会, 電気学会などの会員.