

書字運動モデルと遺伝的アルゴリズムによる変形文字の生成

春日 裕之[†] 住田 桂一[†]
大川 慶[†] 和田 安弘[†]

文字認識システムを構築するためには多種多様な文字を含んだデータベースが必要不可欠であるが、データベースの構築には非常に多くの時間とコストがかかる。本論文では、多種多様な文字を含んだデータベース構成の手段として、Wada & Kawato の書字運動モデルと遺伝的アルゴリズムを応用した文字の変形生成方法を提案した。変動エントロピーによって、生成した文字パターンの多様性を確認したところ、手書き文字データベースである ETL と同程度の多様性を持っていることを確認した。次に、本手法によって生成した文字パターンによって構成した辞書の有効性を検証するために、提案手法により作成した文字による辞書が実際の手書き文字による辞書データベースの代替となることの確認を行った。確認には、Glucksman の特徴抽出法により抽出した特徴を使用し、複合類似度法、線形識別法、k-nearest neighbor 法による 3 つの識別法を使用したシステムを各々構築し、認識実験を行った。実際の手書き文字データベースである ETL6, ETL9 の文字と本手法によって生成した文字データベースを使って構成した各々のパターン認識システムの性能を比較した結果、ほぼ同程度の認識率を得ることができ、提案手法の有効性を確認できた。

A Deformed Character Generation by a Computational Handwriting Model and a Genetic Algorithm

HIROYUKI KASUGA,[†] KEITI SUMIDA,[†] KEI OHKAWA[†]
and YASUHIRO WADA[†]

In pattern recognition, a large number of diversiform characters is necessary to train/test a handwritten character recognition system. However, it is not easy to collect a large number of natural samples. In this paper, we propose a method for generating a large variety of characters by Wada & Kawato's handwriting model and a genetic algorithm. The following two points were examined to evaluate the character patterns generated by the proposed approach. (1) Is there variety in the character patterns generated by the method compared with the character patterns in an existing handwritten database (Japanese character database ETL)? (2) Is there availability in composing a character recognition system that uses handwritten characters generated by the method as training data? The variety in the generated characters by the proposed method was found to be greater than or almost equivalent to the variety in the human handwritten characters in the ETL database. The effectiveness of the pattern generation by the method was evaluated by comparing the recognition accuracy rates of the following systems; (1) Multiple similarity method, (2) k-nearest neighbor method, and (3) Linear discriminant method. Two systems were composed for each method, that is, one used the ETL database as training data and the other used the generated data as training data. The recognition accuracy results by the test data for each method are almost the same and are very high. We confirmed that the proposed method can be applied as an alternate method to collect a large number of diversiform character samples instead of collecting a large number of natural samples.

1. はじめに

現在までに手書き文字認識システムとして、ニューラルネットワークを応用した学習システムや、ある特徴に対する統計処理などによる認識システム^{19),21),23)}

が数多く提案されている。このような認識システムを構築するためには、通常多くの文字データが必要となるが、多種多様なデータを集めるには時間とコストがかかる。そのため、大量の手書き文字のデータの収集作業を、変形文字を人工的に生成することで代替し、文字認識システムにおける学習パターンのトレーニングセットに用いる研究の重要性が指摘⁸⁾されており、これまでもこれに関する研究が報告されている。

[†] 長岡技術科学大学
Nagaoka University of Technology

標準文字に不規則な摂動を与えることによって変形文字を生成する方法^{16),17)}, 文字の特性値の定義から, その主成分分析に基づいて文字を生成する方法¹⁸⁾, 書字過程を確率モデルによって表現して文字の生成を行う方法¹³⁾, 少数のユーザ手書き文字をもとに手書き変形傾向を文字に依存しないようにルール化して, 文字を生成する方法¹⁵⁾, パターンを非線形変換した際に得られる大域的個人性パラメータと局所個人性を表現する線形結合パラメータによって文字生成を行うもの¹⁴⁾, 手書き文字を方向線素の連結によって表現し, それに遺伝的アルゴリズムを適用して文字を生成する方法⁹⁾, 文字の直線や弧をパラメータを持つ関数によって表現して, そのパラメータをランダムに変化させることによって文字を生成する方法³⁾, 文字構造情報を含むオンライン文字筆跡データとサンプル文字パターンから得た方向線素の分布と線分の傾斜角度を統計的に求めて, 逸脱したパターンの生成を抑制するようにした方法²²⁾, などが提案されている。

これらの手法はいずれも多様な手書き文字生成に有効であることが報告されているが, 多くは2次元平面の画像文字パターンを何らかの方法によって変形しようとする手法であり, オフラインデータの画像パターンの変形操作に着目したものであった。また, ほとんどの手法は, 基本的に1つの文字パターンからの変形であって, 複数の文字の特徴を組み合わせて変形文字を生成するものではなかった。

我々が提案する方法は, パターン生成のヒトの運動モデルを基本としている。このモデルは, 特徴(経路点)を入力することによって, 文字パターンの時系列軌道を生成することができるものである。この運動モデルは, ヒトの運動制御の分野において研究・提案されたものであるが, パターンの滑らかさを考慮した最適化問題を解くことによって, ヒトの生成する書字運動パターンをよく再現することが指摘されている。このときパターン上の点から何点かを適当に選んで, その選択された点を運動パターンを生成するモデルに入力することが必要である。このパターン上から選択された点を経路点と呼ぶ。経路点は, いわばパターンの特徴量と考えられるものである。これらの運動パターンを生成するモデルおよび適切な経路点を抽出するモデルについては, Wada & Kawato^{11),12),24)}によって提案されている。

本論文において, 我々が提案する方法は, より多様な文字パターンを生成するために, 複数の特徴を持つ文字パターンの部分を組み合わせて, 1つの文字パターンの特徴を構成しようとするものである。また, 特徴

に摂動を与えて文字を生成した方が, より多様な文字を生成できる可能性が高いと考えられる。そこで, 最初に, 比較的少人数から採取した文字パターンから各文字パターンごとの経路点組を抽出する。次に, それらの経路点に対して遺伝的アルゴリズムを適用することで, 採取した経路点組とは異なる新しい経路点位置情報(空間情報)と経路点通過時刻情報(時間情報)を持った経路点の組を作り出す。そして, これらを経路点情報として書字運動モデルに与え, 文字パターンの生成を行うことによって, 計測したパターンとは異なるパターンを新たに生成する方法である。つまり, 複数の人の個人的特徴の部分を組み合わせて, 1つの文字パターンを構成する方法であり, 比較的少数の採取文字パターンから, より多様な文字パターンの生成が期待される。

本論文では, 以下の点にしばって, 提案手法の有効性を確認する。提案手法で生成した変形文字を認識辞書として用いた文字認識システムとヒトの手書きデータベース(ETL9B, ETL6)の文字を認識辞書とした文字認識システムの認識率を比較することにより, 提案手法による変形文字によって構成されたデータベースの認識辞書としての有効性を検証する。つまり, パターン認識システムを構築する際に収集するデータとして, 実際にサンプルするのではなく, 本論文で提案した手法によって生成して, 辞書データとして利用することを想定して, 本手法で生成した文字データを辞書データベースとして利用する有効性に関する比較実験を行う。

以下, 2章では本論文で用いる文字生成モデルの基本となる Wada & Kawato の書字運動モデルを概説し, 3章で遺伝的アルゴリズムを用いた変形文字の生成手法を説明する。次に本手法で生成した文字の多様性と有効性を確認するため, 4章では, 文字の多様性を変動エントロピーを用いて評価し, ETLデータベースと比較する。また, 5章で ETL のデータを辞書として使用した場合と提案手法により生成した文字を使用した場合の認識システムの認識精度を比較することにより, 変形文字の有効性について検討する。

2. 書字運動モデル

2.1 経路点推定モデル

手書き文字のような複雑な運動軌道の経路点を抽出するために, 我々は以前に計測軌道から経路点を抽出し, 抽出した経路点から計測軌道の近似軌道を作るモデル¹²⁾を提案している。

複雑な動作の経路点の獲得は次の最適化問題を解く

ことにより得ることができる。\$X_{data}\$ は与えられた複雑な運動とし、\$X_{reconst}\$ はモデルによって再現される軌道とする。経由点推定問題は式 (1) を満たす最小個数の経由点 \$n\$ 個の組を見出すことである。

$$|X_{data} - X_{reconst}| < \varepsilon \quad (1)$$

すなわち、経由点推定アルゴリズムは誤差が閾値 \$\varepsilon\$ 以下になるような経由点組のうち最小個数の経由点組を探す方法である。なお、推定される経由点の位置は、定性的には軌道の曲率の大きな点に推定される可能性が高く、経由点の数は、\$\varepsilon\$ が同じであれば、ほぼ同じ数の経由点が推定される。

2.2 軌道生成モデル

いわゆるヒト運動の軌道は躍度最小規範²⁾ や指令トルク変化最小規範^{6),10)} のような最適化原理に基づいたモデルによって近似的に生成可能であることが指摘されている。後者の方が、ヒトの運動軌道をよりよく再現することが指摘されており、前者は後者の近似モデルと考えることができる。しかしながら、前者は、解析的に軌道を表す関数(スプライン関数)を求めることができるが、一方、後者は、非線形最適化問題を解く必要があり、解析的な解(軌道)を得ることは困難であり、数値的解を得ることも比較的難しい問題である。したがって、本論文では、比較的容易に軌道を求めることができる躍度最小規範(2)をパターン生成モデルとして使用する。

$$\frac{1}{2} \int_0^T \left\{ \left(\frac{d^3x}{dt^3} \right)^2 + \left(\frac{d^3y}{dt^3} \right)^2 + \left(\frac{d^3z}{dt^3} \right)^2 \right\} dt \rightarrow Min \quad (2)$$

ここで、\$x, y, z\$ は軌道の座標を表し、\$T\$ は運動時間を表す。この最適化問題(2)の最適軌道は、5次のスプライン関数によって表すことができ、このスプライン関数は、始点、経由点、終点の位置および各点の通過時刻を与えることによって一意に決定される。

3. 遺伝的アルゴリズムによる文字の生成手法

本章では、最初に、遺伝的アルゴリズムを応用するために必要となる文字の評価のための適応度関数を提案し、次に、これを用いた遺伝的アルゴリズムによる文字生成手法について述べる。

3.1 適応度関数

躍度最小化規範に基づいた書字軌道生成のための目的関数は経由点の通過を制約条件とした式(3)の最適化問題として表現することができる。この目的関数は躍度最小規範に従った軌道生成において、与えられた経由点を通することを境界条件とした軌道生成の

ための評価関数に対応する。

$$\int_0^T \left\{ \left(\frac{d^3x}{dt^3} \right)^2 + \left(\frac{d^3y}{dt^3} \right)^2 + \left(\frac{d^3z}{dt^3} \right)^2 \right\} dt + \Sigma \{ (x(t_{via}) - x_{via}^*)^2 + (y(t_{via}) - y_{via}^*)^2 + (z(t_{via}) - z_{via}^*)^2 \} \rightarrow Min \quad (3)$$

ここで、\$x, y, z\$ は各時刻での軌道の座標を示し、\$x_{via}^*, y_{via}^*, z_{via}^*\$ は経由点の座標、\$x(t_{via}), y(t_{via}), z(t_{via})\$ は経由点通過時刻 \$t_{via}\$ での軌道の座標を示す。

式(3)の第1項は運動における躍度が滑らかなようになるように軌道を生成するための項であり、第2項は \$x_{via}^*, y_{via}^*, z_{via}^*\$ をある文字パターンの経由点とすれば、通過すべき経由点を制約するための項である。したがって、式(3)は、ある経由点を通する滑らかな軌道(文字パターン)の評価値を示していることになる。逆に、軌道 \$(x, y, z)\$ が与えられたとき(以後、この軌道を評価パターンと呼ぶ)、その軌道に対する式(3)を計算すると、評価パターン \$(x, y, z)\$ と \$(x_{via}^*, y_{via}^*, z_{via}^*)\$ を通過する躍度最小軌道 \$(x^*, y^*, z^*)\$ (以後、標準パターンと呼ぶ)との一種の距離が定義される。本論文では、標準パターン \$(x^*, y^*, z^*)\$ と、評価パターン \$(x, y, z)\$ の式(3)の差を、標準パターンに対する評価パターンの変動を評価する値と考える。本論文では、標準パターンの \$W\$ の値を1に正規化する意味で、式(4)で示すような指数関数表現とした。また、具体的に式(4)を適用する際には、標準パターンと評価パターンの経由点の個数の差が問題となる。本研究では、付録A.1に示したような方法によって、標準パターンと評価パターンの経由点の対応をとるようにしている。

$$W = \exp \left\{ - \frac{\{C_J(x, y, z) - C_J(x^*, y^*, z^*)\}^2}{\alpha^2} \right\} \cdot \exp \left\{ - \frac{1}{\beta^2} \sum_{via=1}^{N_{via}} (x(t_{via}) - x_{via}^*)^2 + (y(t_{via}) - y_{via}^*)^2 + (z(t_{via}) - z_{via}^*)^2 \right\} \quad (4)$$

ここで \$C_J\$ は式(3)における第1項を示し、\$x, y, z\$ は評価パターンの座標、\$x^*, y^*, z^*\$ は標準パターンの座標を表し、\$T\$ は \$T=1\$ となるように運動時間を正規化する。\$\alpha, \beta\$ は適当な定数である。

我々は様々な文字に対して心理実験(付録A.2)を行い、式(4)の評価の妥当性を検討した。1つの文字に対して24のパターンを用意し、11人の被験者による実験を行った。評価モデル(式(4))とヒトによる評価との相関係数は、約0.8と非常に高い値を示した。

これより、式 (4) はヒトの選好関係を比較的良好に定量化できると考えられる。よって、この評価式を適応度関数とした遺伝的アルゴリズムによる多様な文字生成のための方法を提案する。

3.2 遺伝的アルゴリズムによる文字の生成

Wada & Kawato の書字運動モデルにおいて、ある文字から抽出した経路点のうち、少なくとも 1 つの経路点の座標や時間を摂動させると、元の文字と異なった形状の文字を生成することが可能である。つまり、経路点情報を適当に変化させることで、元の文字の変形が可能となる。しかし、ランダムに経路点情報を変化させると、ヒトが文字と可読できないような文字を多数発生させることになる。

また、1 つの文字の経路点を変化させるよりも、複数のヒトの書いた文字を組合せて、その経路点情報を変化させた方がより多様な文字を生成できる可能性が高いと考えられる。そこで、本論文では複数の経路点組を用意して、遺伝的アルゴリズムを適用することで文字の経路点を交叉・摂動させ、様々な経路点の組合せを発生させ、多様な形状の文字の生成を行う。手順としては、1) 交叉に使用する 2 つの文字の経路点間の 1 点交差を行い、2) 経路点を突然変異により摂動させる、3) 経路点情報から書字運動モデルによって軌道を生成し、4) 式 (4) を使用して文字の評価を行い、選択・淘汰を行う。これらの処理の詳細を以下に述べる。

3.2.1 遺伝子座情報

遺伝的アルゴリズムで用いる遺伝子座の情報は、経路点の空間情報 x_{via} , y_{via} , z_{via} と通過時間 t_{via} である。ただし、経路点の空間情報については式 (5) を用いて経路点間の距離の総和でパターンの大きさを正規化する。

$$\begin{aligned} \tilde{x}_{via}^i &= \hat{x}_{via}^{i-1} + \frac{x_{via}^i - x_{via}^{i-1}}{L} \\ \tilde{y}_{via}^i &= \hat{y}_{via}^{i-1} + \frac{y_{via}^i - y_{via}^{i-1}}{L} \\ \tilde{z}_{via}^i &= \hat{z}_{via}^{i-1} + \frac{z_{via}^i - z_{via}^{i-1}}{L} \end{aligned} \quad (5)$$

$$L = \sum_{i=1}^{N+1} \sqrt{(x_{via}^i - x_{via}^{i-1})^2 + (y_{via}^i - y_{via}^{i-1})^2 + (z_{via}^i - z_{via}^{i-1})^2}$$

ここで、 L は経路点間の距離の総和、 x_{via}^i , y_{via}^i , z_{via}^i は正規化前の経路点座標、 \tilde{x}_{via}^i , \tilde{y}_{via}^i , \tilde{z}_{via}^i は正規化後の経路点の座標である。ただし、経路点数は N 個であり、 $i = 0$ は始点を、 $i = N + 1$ は終点を示す。また、経路点の通過時刻は式 (6) を用いて全体の運動

時間で正規化を行う。

$$\tilde{t}_{via}^i = \frac{t_{via}^i - t_{via}^0}{t_{via}^{N+1} - t_{via}^0} \quad (6)$$

ここで、 t_{via}^i は正規化前の経路点の通過時刻、 \tilde{t}_{via}^i は正規化後の経路点の通過時刻である。

生成した文字パターンの位置ずれを修正するために、各経路点の x 座標、 y 座標の原点が各経路点座標の平均で与えられる重心となるように式 (7) を用いて移動する。

$$\begin{aligned} \bar{x}_{via}^i &= \tilde{x}_{via}^i + \frac{\sum_{i=0}^{N+1} \tilde{x}_{via}^i}{N+2} \\ \bar{y}_{via}^i &= \tilde{y}_{via}^i + \frac{\sum_{i=0}^{N+1} \tilde{y}_{via}^i}{N+2} \end{aligned} \quad (7)$$

ここで、 \bar{x}_{via}^i , \bar{y}_{via}^i は経路点の重心を基準とした経路点座標を表す。

遺伝子情報は位置の情報として、正規化後の経路点の x , y 座標の相対座標値、 z 座標の絶対座標値、時間の情報として、正規化後の経路点の相対的な運動時間からなっている。 z 軸を絶対座標としたのは、 z 軸は「i」などの 3 次元運動を行う文字の 1 画目から 2 画目への移動時の z 軸の高さは人によって違うため z 軸が最も高い点で交叉の処理が行われた場合 1 画目の書字面より 2 画目の書字面が高い座標になることがあるためである。

3.2.2 交叉

親を 0 世代目として、 n 世代目の交叉の対象とする個体は $0 \sim n - 1$ 世代目に生成されたパターンの中から乱数により 2 つ選択する。選択された 2 つの個体は Wada & Kawato の書字運動モデルにより、経路点を抽出して遺伝子情報に変換し、2 つのパターンの遺伝子情報の対応 (付録 A.1) をとり、対応をとった経路点の中から乱数によって 1 点を選択し、1 点交叉を行う。

3.2.3 突然変異

はじめに用意した初期個体の遺伝子情報のパターンの組合せだけでは、パターンの変動に限界がある。そこで、より大きくパターンを変動させるために交叉によって新たに生成した遺伝子情報に対して突然変異を与える。確率 p_m で突然変異を起こす遺伝子を決定し、その遺伝子の中から確率 p_v で突然変異を起こす遺伝子座を選択する。選択された遺伝子座の x_{via} , y_{via} , z_{via} , t_{via} に係数 ε_x , ε_y , ε_z , ε_t を乗じて新たな遺伝子情報を生成する。係数 ε_x , ε_y , ε_z , ε_t は正の乱

数である。

3.2.4 軌道の生成

交叉・突然変異の処理を終了した新たな遺伝子情報を経由点情報に変換し、Wada & Kawato 書字運動モデルを使用して経由点情報から書字軌道を生成する。

3.2.5 選択・淘汰

生成した文字の中には文字として認識できないような文字が生成される可能性があり、それを取り除くために式 (4) で生成した文字に対する評価値を計算し、評価値が閾値を超えない文字は人間に判別することができない文字として淘汰する。

この手順を繰り返し、多様な文字の生成を行うが、標準パターンに基づいた最適な軌道を得ることが目的ではないため、エリート戦略などの手法によって、評価値に従って個体を優先的に選択するような操作は行わず、ただ、ヒトの判別が困難な文字の淘汰だけを行って、多様な文字の生成を可能とする。

4. 文字の多様性の評価

提案手法で生成したデータベースの文字の変動を定量的に評価し、実際の手書き文字のデータベースと比較し、本手法で生成したデータベースの変動量が十分なものであることを確認するために、以下の式 (8) で表される変動エントロピー⁵⁾を用いた。

$$h = - \sum_{i=1}^I \sum_{j=1}^J \frac{n(i,j)}{N} \log \frac{n(i,j)}{m} \quad (8)$$

ここで、 i, j はパターンの座標を表し、各パターンのサイズは $I * J$ である。座標 (i, j) での黒画素数を $n(i, j) \leq m$ とした。また、 m は重ね合わせたパターンの枚数、 N は重ね合わせたパターンの黒点の総和である。

初期個体に使用する個数を 2 つから 7 つと変えて、各世代の交叉回数 50 回で 10 世代まで文字の生成を行い、生成された文字の各世代ごとの変動エントロピーを計算した。その結果を図 1 に示す。図は『ほ』について行った結果である。早い世代の文字は文字群の変動量が小さいが、世代が進むにつれ文字群の変動量が大きくなっていることが確認できる。今回の実験では 4 個以上の初期個体では、変動エントロピーの変化の相違は大きくなかったが、2 個の初期個体よりも 3 個以上の初期個体を使った方が、より多様な文字を生成できることが確認でき、複数の文字を遺伝的アルゴリズムを使って交叉処理を行う有効性が確認できる。本実験では、4 個以上の個体数での変動エントロピーの差はあまり確認できなかった。初期個体自体の形状の

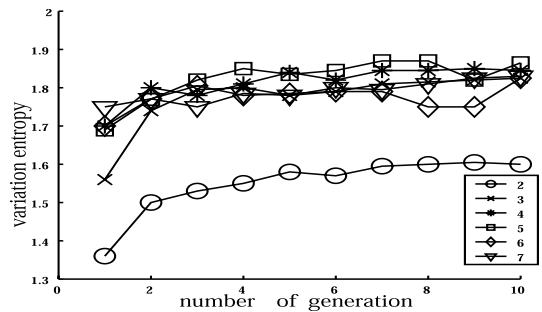


図 1 初期個体数による変動エントロピーの変化

Fig. 1 Change of variation entropy according to number of initial genes.

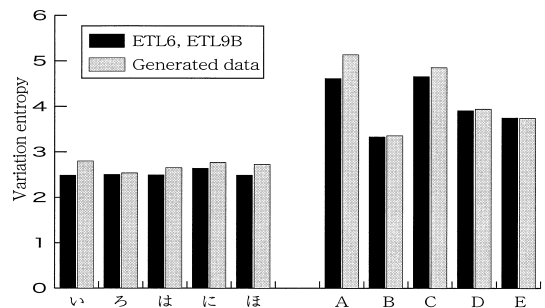


図 2 ETL と生成文字の変動エントロピーの比較

Fig. 2 Comparison between the ETL database and generated characters.

影響や選択淘汰のしきい値の問題などが考えられ、今後の検討課題である。

初期個体 4 つを用いて本手法によって生成したアルファベットと仮名文字の変動エントロピーと ETL6, ETL9B にあるアルファベットと仮名文字の変動エントロピーを計算したものを図 2 に示す。アルファベットについては、生成した約 1,000 個の文字から 100 文字を無作為に 3 セット抽出して、その変動エントロピーを平均したものと、ETL6 にあるアルファベットの約 1,300 個から 100 文字を無作為に 3 セット抽出して変動エントロピーを平均したものを示している。仮名文字については、約 1,000 個の生成文字から 200 文字を無作為に 3 セット抽出して変動エントロピーを平均したものと、ETL9B にある仮名文字の約 200 個の変動エントロピーを示している。図 2 の結果より仮名文字、アルファベットの文字において ETL9B, ETL6 と同程度の変動量を持った文字群を生成することが確認できる。

5. 文字認識システム構築の有用性

5.1 文字の生成

本手法を用いて、図 3, 4, 5 に示す 4 つの初期個

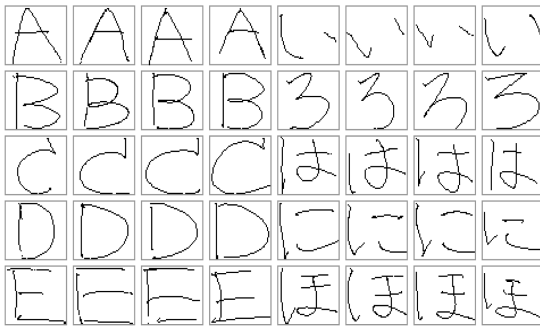


図 3 初期個体

Fig. 3 Initial patterns.

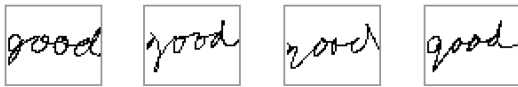


図 4 初期個体 : good

Fig. 4 Initial patterns: good.

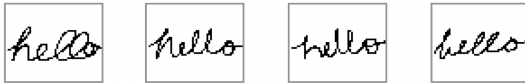


図 5 初期個体 : hello

Fig. 5 Initial patterns: hello.

体から生成したアルファベット, アルファベットの続け字, 仮名文字の一例を図 6, 7, 8, 9, 10, 11 に示す. また, 図 6~11 を生成したときに淘汰された文字の一例を図 12, 13, 14 に示す. 図 6~11 の文字の生成においては, 図 3~5 に示す一番左の初期個体を標準個体として文字の生成を行った. 図 6~11 には初期個体にはない異なったパターンが含まれていることが確認できる.

5.2 認識システムの構築

5.2.1 認識方法

本論文では提案手法の有効性を検証するための認識実験を行った. 認識対象字種は仮名文字「い」「ろ」「は」「に」「ほ」、アルファベット「A」「B」「C」「D」「E」で, 仮名文字, アルファベットの各々に分けた認識実験を行った. 文字生成のための初期個体として与えたデータは, 60Hz でサンプリングした 4 人の文字データを使った. 実験では手書き文字データベース ETL9B, ETL6 に含まれる文字の中から, 各文字から乱数により抜き出した 100 個の文字をテストデータ(以後, ETL テスト)とした. また, ETL テストに使用した文字とは重ならないように各文字から乱数により抜き出した 100 個の文字の辞書データ(以後, ETL 辞書)と提案手法により生成した文字の辞書データ(以後, GA 辞書)を各々使用した認識システムを

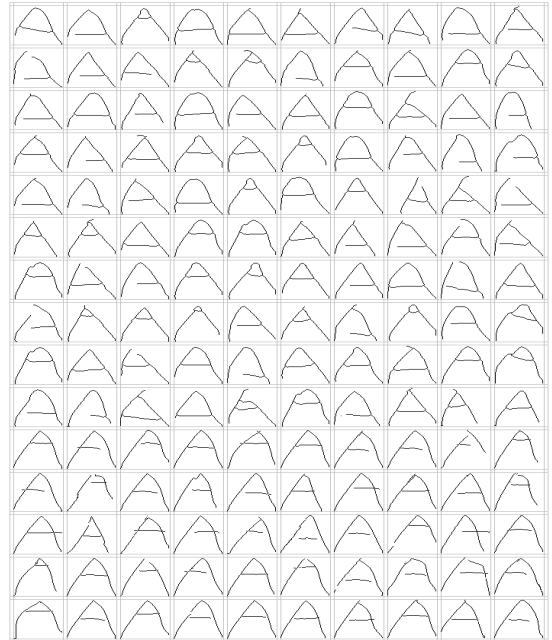


図 6 生成文字 : A

Fig. 6 Generated characters: A.

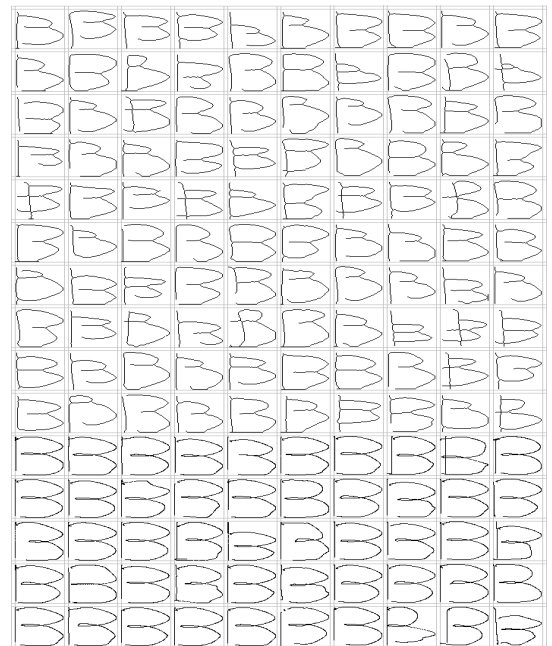


図 7 生成文字 : B

Fig. 7 Generated characters: B.

構築した. GA 辞書は, 初期個体と表 1 に示すパラメータを使用して作成し, アルファベットは約 1,000 個, 仮名文字は約 20,000 個の文字を生成し, その中から乱数を使用して 100 個の文字を抜き出したものを

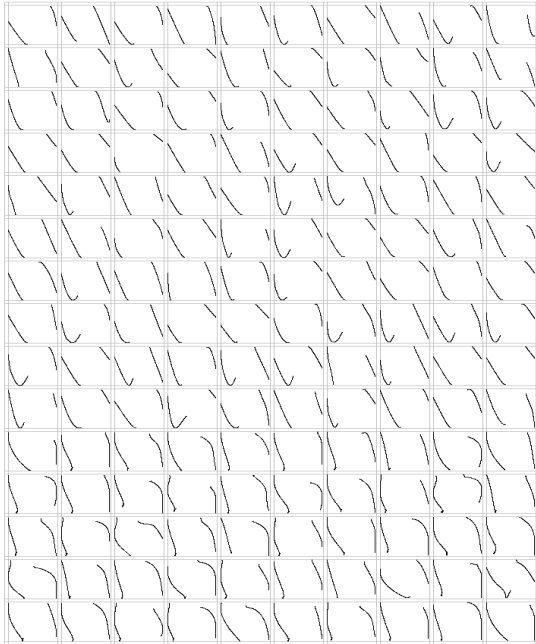


図 8 生成文字：い

Fig. 8 Generated characters: Japanese character い.

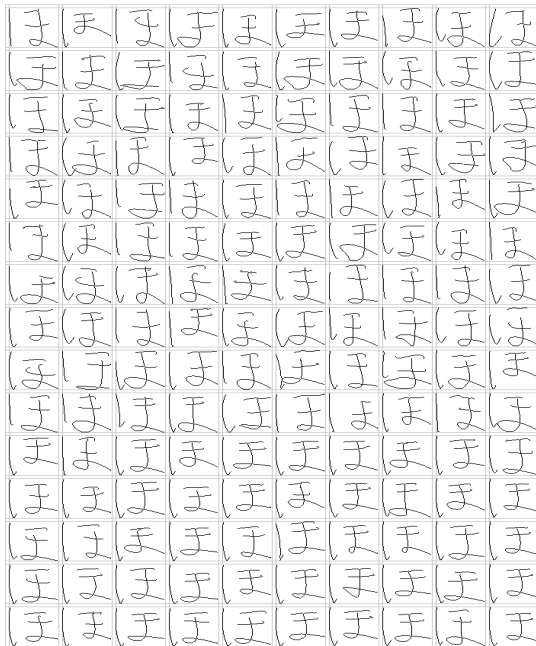


図 9 生成文字：ほ

Fig. 9 Generated characters: Japanese character ほ.

GA 辞書とした。また、選択淘汰に用いた式 (4) の閾値は、各文字の平均で、 5×10^{-2} 、 α は 1.6×10^{13} 、 β は、0.32 である。ただし、本論文では、閾値設定は、生成された軌道を確認して、主観的な判断で、図 12~



図 10 生成文字：good

Fig. 10 Generated script: good.



図 11 生成文字：hello

Fig. 11 Generated script: hello.

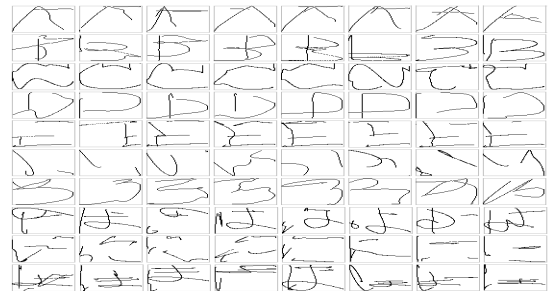


図 12 淘汰文字：A, B, C, D, E, い, ろ, は, に, ほ

Fig. 12 Selected characters.



図 13 淘汰文字：good

Fig. 13 Selected scripts: good.



図 14 淘汰文字：hello

Fig. 14 Selected scripts: hello.

14 で示したような文字が淘汰されるように閾値を決定した。

ETL データについては、含まれているノイズを除

表 1 文字の生成パラメータ

Table 1 Parameter values for character formation.

| | アルファベット | 仮名文字 |
|---------------|---------|--------|
| 世代数(世代) | 10 | 200 |
| 交叉回数(回数) | 50 | 50 |
| 突然変異が起こる確率(%) | 10 | 70 |
| 遺伝子座の変異確率(%) | 20 | 20 |
| 突然変異時の摂動率(%) | 60~140 | 90~110 |

去し、文字の細線化を行い、縦横比を損なわないように 64*63 のサイズに正規化を行った。GA データについても同様の正規化を行った。

認識実験では、識別用特徴として Glucksman の特徴抽出法⁴⁾による 81 次元のベクトルを使用した。識別方法としては複合類似度法²⁰⁾、線形識別法⁷⁾、k-nearest neighbor 識別¹⁾、の 3 つの識別方法を使用した。線形識別法では、各文字ごとに 1 つの識別関数を構成した。つまり、「い」「ほ」「A」「E」の仮名文字、アルファベットに対応する各々 5 つの識別関数を構成した。最終的な認識結果は、各 5 つの識別関数で最も高い出力をした文字を認識結果とした。k-nearest neighbor 識別では、 $k = 10$ として実験した。

5.3 実験結果

ETL 辞書、GA 辞書の 2 つの辞書を使用して、複合類似度法、線形識別法、k-nearest neighbor 識別を使用した識別法により認識実験を行った結果、表 2, 3, 4, 5 のようになった。認識結果はすべて同一のテストデータである ETL テストに対する結果であり、GA 辞書については、生成した文字データの中から乱数によって 100 文字を 3 セット抽出して各々の認識システムを構成し、各々に対してテストを行い、その平均正認識率を表に示した。また、認識実験結果において用いた認識のためのパラメータ値は、ETL データによる辞書と、GA 生成文字データによる辞書の両者の場合について同じ値を用いて実験を行っている。なお、本実験では、リジェクトは発生しないようにしており、必ずどれかのカテゴリに分類されるものとしたので、リジェクト率は 0% である。したがって、誤認識率は、100% - 正認識率で与えられる。

ETL 辞書によって構成したアルファベット、仮名文字に対する認識システムの正認識率(誤認識率)の平均は、各々 100%(0.0%), 94.7%(5.3%) であった。また、GA 辞書によって構成した認識システムの正認識率(誤認識率)の平均は、各々 100%(0.0%), 94.5%(5.5%) であった。本論文で提案した手法によって生成した文字で構成した認識システムは、ETL で構成した認識システムとほぼ同程度の認識精度が示されているとい

表 2 アルファベット: ETL を辞書として使用した場合の正認識率

Table 2 Alphabet recognition rate: ETL dictionary.

| | A | B | C | D | E |
|---------------------|-----|-----|-----|-----|-----|
| Multiple similarity | 100 | 100 | 100 | 100 | 100 |
| Linear discriminant | 100 | 100 | 100 | 100 | 100 |
| k-nearest neighbor | 100 | 100 | 100 | 100 | 100 |

表 3 アルファベット: GA を辞書として使用した場合の正認識率

Table 3 Alphabet recognition rate: GA dictionary.

| | A | B | C | D | E |
|---------------------|-----|-----|-----|-----|-----|
| Multiple similarity | 100 | 100 | 100 | 100 | 100 |
| Linear discriminant | 100 | 100 | 100 | 100 | 100 |
| k-nearest neighbor | 100 | 100 | 100 | 100 | 100 |

表 4 仮名文字: ETL を辞書として使用した場合の正認識率

Table 4 Kana recognition rate: ETL dictionary.

| | い | ろ | は | に | ほ |
|---------------------|----|-----|----|-----|----|
| Multiple similarity | 93 | 100 | 90 | 100 | 90 |
| Linear discriminant | 95 | 100 | 90 | 98 | 91 |
| k-nearest neighbor | 89 | 100 | 90 | 99 | 96 |

表 5 仮名文字: GA を辞書として使用した場合の正認識率

Table 5 Kana recognition rate: GA dictionary.

| | い | ろ | は | に | ほ |
|---------------------|----|-----|----|-----|----|
| Multiple similarity | 90 | 100 | 89 | 100 | 95 |
| Linear discriminant | 91 | 100 | 90 | 100 | 92 |
| k-nearest neighbor | 90 | 100 | 87 | 100 | 94 |

え、本手法の有用性が確認できた。

また、比較のために、初期個体だけを辞書とした認識システムを構成してみたところ、アルファベットに対しては、正認識率: 96.2%、仮名文字に対しては、正認識率: 72.2%、であった。これより、初期個体だけでは、ETL 辞書や GA 辞書と同レベルの認識精度を持つ認識システムを構築することができないことが分かる。本手法によって生成した文字を使った GA 辞書は、初期個体だけではカバーできないような文字パターンの認識を可能にしていると考えられる。しかも、実際に採取した文字パターン数は 4 つで、ほぼ ETL データベースの 100 パターン分で構成する認識システムと同程度の認識精度を持つ認識システムが構築できることを示しており、本手法の有効性が理解できる。また、特に、パターンの変動が大きいと推定される仮名文字に対しては、認識率は飛躍的に改善されており、多様なパターンが必要な認識システムの構築において、本手法が有効であることが分かる。

次に我々は、生成文字 100 文字以上の文字数によって辞書を構成した場合のかな文字の認識率の向上を確認した。実験は、100 文字の場合と同様に行い、200 文字、400 文字の辞書 3 セットを使って認識実験を

行った．認識率の平均は，200 文字：95.7%，400 文字：94.8%であった．文字数を増やすことによる認識率の向上が確認でき，本手法の認識システム構築への応用の有効性が確認できる．しかしながら，400 文字に文字数を増やした場合は，200 文字に比べて若干の認識率の低下があった．これは，特徴量に代表される認識アルゴリズムの問題，遺伝的アルゴリズムによって生成する際の淘汰のパラメータの問題などが 1 つの要因の可能性と考えられ，これらの検討は今後の課題である．

6. 結 言

本論文では，Wada & Kawato の書字運動モデルによって抽出した経路点を摂動させることにより，多様な形状の文字の生成を行った．経路点の摂動の方法として遺伝的アルゴリズムを使用するため，文字生成のための適応度関数を定義し，心理実験により確認を行った．定義した適応度関数を使用して，文字を生成し，生成した文字を人の手書き文字と比較して十分な変動量であることを確認するために変動エントロピーを用いて，ETL9B，ETL6 のデータと生成文字を比較し，同程度の変動量であることを確認した．また，本手法により生成した変形文字と ETL9B，ETL6 の文字データを使用し，複数の認識実験を行った．その結果，ETL9B，ETL6 の文字データを辞書として用いたときと生成文字の認識結果はほぼ等しく，文字変形モデルの有効性を確認することができた．このことより，人の手書き文字のデータベースの代替えとして活用できることの可能性が確認できた．

本論文で提案した手法は，ヒトの時系列の運動軌道を取得して，その軌道の経路点を推定して，それを摂動させて，文字パターンの生成を行うものであり，従来提案されているオフラインの画像データに対する変形処理を施す手法とは異なり，応用の広い手法である．たとえば，オフラインのデータベースの構築だけでなく，書字時間の異なるオンラインのデータベースの構築などにも応用が可能である．また，文字パターンだけでなく，今後，ヒューマンインタフェースにおいて重要性が増してくると考えられる手話や身振りなどのパターンに対しても適用が可能であり，広い応用範囲を持つと考えられる．

謝辞 本研究を行うにあたり，手書き文字データベースを提供していただいた産業技術総合研究所の皆様に感謝いたします．

参 考 文 献

- 1) Cover, T.M. and Hart, P.E.: Nearest neighbor pattern classification, *IEEE Trans. Inf. Theory*, Vol.IT-13, No.1, pp.21-27 (1967).
- 2) Flash, T. and Hogan, N.: The coordination of arm movements: An experimentally confirmed mathematical model, *Journal of Neuroscience*, Vol.5, pp.1688-1703 (1985).
- 3) Ghosh, D. and Shivaprasad, A.: An analytic approach for generation of artificial hand-printed character database from given generative models, *Pattern Recognition*, Vol.32, pp.907-920 (1999).
- 4) Glucksman, H.A.: Classification of mixed-font alphabets by characteristic loci, *IEEE Computer Conf.*, pp.138-141 (1967).
- 5) 長谷博行, 米田政明, 酒井 充, 吉田順作: 変動エントロピーによる文字変動の解析, 電子情報通信学会論文誌, Vol.J71-D, No.6, pp.1048-1056 (1988).
- 6) Nakano, E., Imamizu, H., Osu, R., Uno, Y., Gomi, H., Yoshioka, T. and Kawato, M.: Quantitative examinations of internal representations for arm trajectory planning: minimum commanded torque change model, *Journal of Neurophysiology*, Vol.81, No.5, pp.2140-2155 (1999).
- 7) Nilsson, N.J.: *Learning Machines*, McGraw-Hill, New York (1965).
- 8) 小川英光: パターン認識・理解の新たな展開, 信学会 (1994).
- 9) 但馬文昭, 張 軍: 遺伝的アルゴリズムによる個性と変動を伴った手書き風文字の生成, *Et98-89*, 信学技報, pp.125-132 (1998).
- 10) Uno, Y., Kawato, M. and Suzuki, R.: Formation and control of optimal trajectory in human multijoint arm movement - minimum torque change model, *Biological Cybernetics*, Vol.61, pp.89-101 (1989).
- 11) Wada, Y. and Kawato, M.: A neural network model for arm trajectory formation using forward and inverse dynamics models, *Neural Networks*, Vol.6, No.7, pp.919-932 (1993).
- 12) Wada, Y. and Kawato, M.: Theory for cursive handwriting based on the minimization principle, *Biological Cybernetics*, Vol.73, No.1, pp.3-13 (1995).
- 13) ブーワット・アッタチュー, 近藤正三: 手書き過程のモデル化とその評価, *Pru86-15*, 信学技法, pp.39-48 (1986).
- 14) 安本 護, 池田尚志, 堀井 洋: 大域的個人性と局所的個人性に基づく手書き風フォントの生成, 電子情報通信学会論文誌 D-II, Vol.J80-D-

- II, No.11, pp.2930-2939 (1997).
- 15) 喜多伸之：個人筆記モデルを用いた手書き文字認識の検討，画像の認識・理解シンポジウム MIRU'92, pp.247-253 (1992).
 - 16) 石井健一郎：変形文字パターン発生法とその応用，電子情報通信学会論文誌，Vol.J66-D, No.11, pp.1270-1277 (1983).
 - 17) 石井健一郎：変形文字を用いた識別辞書の設計，電子情報通信学会論文誌 D-II, Vol.J72-D-II, No.5, pp.669-577 (1989).
 - 18) 大澤一爽：文字の科学，法政大学出版会，pp.129-146 (1985).
 - 19) 中嶋 孝，若林哲史，木村文隆，三宅康二：混合識別関数による類似文字認識の高精度化，電子情報通信学会論文誌 D-II, Vol.J83-D-II, No.2, pp.623-633 (2000).
 - 20) 飯島泰蔵：パターン認識，コロナ社 (1973).
 - 21) 木村義政，若原 徹，伴野 明：統計的手法と神経回路の結合による高精度多字種認識，電子情報通信学会論文誌 D-II, Vol.J83-D-II, No.10, pp.1986-1994 (2000).
 - 22) 鈴木雅人，加藤 寧，根元義章，市村 洋：オンライン文字筆跡情報を用いた手書き文字パターンの自動生成方法，電子情報通信学会論文誌 D-II, Vol.J84-D-II, No.11, pp.2353-2361 (2001).
 - 23) 鈴木雅人，加藤 寧，根本義章，市村 洋：2次混合マハラノビス関数を用いた類似文字識別手法，電子情報通信学会論文誌 D-II, Vol.J84-D-II, No.4, pp.659-667 (2001).
 - 24) 和田安弘，小池康晴，川人光男：連続書字運動の計算論的モデル，電子情報通信学会論文誌 D-II, Vol.J76-D-II, No.11, pp.2400-2410 (1993).

付 録

A.1 経路点の対応

標準パターンと評価パターンの経路点の数が異なるため，経路点のそれぞれの始点から順に対応をとると定性的に全然異なる位置の経路点に対応することになる．そこで，互いの経路点の対応をとる必要がある．本論文では，式 (9) によって経路点の対応をとり，2つの文字の対応する適切な経路点を選び評価を行う．

$$p_{ij} = \exp \left\{ -\frac{(x_{via,j} - x_{via,i}^*)^2}{\sigma_x^2} - \frac{(y_{via,j} - y_{via,i}^*)^2}{\sigma_y^2} - \frac{(z_{via,j} - z_{via,i}^*)^2}{\sigma_z^2} - \frac{(t_{via,j} - t_{via,i}^*)^2}{\sigma_t^2} \right\}$$

$$\hat{p}_{ij} \leftarrow \max_j p_{ij} \tag{9}$$

ここで， $x_{via,i}^*, y_{via,i}^*, z_{via,i}^*$ は標準個体の経路点 i

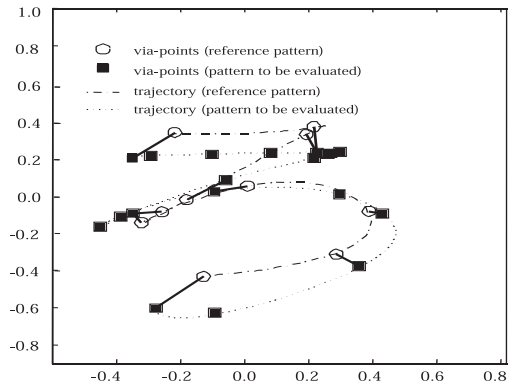


図 15 対応する経路点
Fig. 15 Corresponding via-point between reference pattern and pattern to be evaluated.



図 16 主観的評価実験に用いたパターンセットの例
Fig. 16 Examples for test patterns.

の座標， $t_{via,i}^*$ は標準個体の経路点 i の通過時刻， $x_{via,j}, y_{via,j}, z_{via,j}$ は評価パターンの経路点 j の座標， $t_{via,j}$ は評価パターンの経路点 j の通過時刻， $\sigma_x, \sigma_y, \sigma_z$ は経路点間距離の平均 $\times \frac{1}{3}$ ， σ_t は標準個体の経路点間時間の平均 $\times \frac{1}{6}$ とした．式 (9) で，評価パターンの経路点から見た標準パターンの経路点の中で最も定性的に対応がとれている標準パターンの経路点との対応をとることが可能になる．図 15 に「る」の場合についての例を示す．標準パターンの経路点数が 10 個，評価パターンの経路点数が 18 個のケースであるが，上記の対応法によって，定性的に妥当な経路点間の対応がとれることが分かる．

A.2 評価関数の妥当性確認のための心理実験

本節では，評価式 (4) による文字変動の評価が妥当であることを被験者による心理実験によって確認を行った結果を示す．ここでは，人の主観的な評価と評価式 (4) による評価の相関により確認を行った．まず，ある文字パターンの標準経路点の位置と通過時刻を撰動させてパターンを 24 パターン作成した．図 16 に，作成したパターンセットの例を示す．人の主観的な評

表 6 人の評価と評価式による評価の相関係数

Table 6 Correlation coefficient between model evaluation and human evaluation.

| | 実験 I | 実験 II | 実験 III | 実験 IV |
|---|-------|-------|--------|-------|
| い | 0.724 | 0.714 | 0.714 | 0.701 |
| ろ | 0.815 | 0.806 | 0.812 | 0.796 |
| は | 0.788 | 0.776 | 0.776 | 0.777 |
| に | 0.828 | 0.846 | 0.852 | 0.854 |
| ほ | 0.768 | 0.811 | 0.812 | 0.809 |
| へ | 0.851 | 0.834 | 0.839 | 0.861 |
| と | 0.816 | 0.822 | 0.822 | 0.858 |
| a | 0.774 | — | — | — |
| b | 0.881 | — | — | — |
| c | 0.913 | — | — | — |

価では、このパターンセットを印刷してカードにし、以下の4つの方法によってカードを被験者に並べてもらい、パターンセットに順位付けを行った。

実験 I: 24枚のカードすべてを被験者に渡して、被験者がきれいだと思う順に並べる。

実験 II: カードを2枚ずつ見せ、きれいだと思う方を残し順位をつける。

実験 III: 実験 IIの終了後に、その並びを修正する。

実験 IV: 24枚のカードと標準パターンを見せ、標準パターンに近いと思う順に並べる。

実験 I~実験 IIIは、被験者がそれぞれ記憶として持っている標準パターンとの比較であると考えられる。しかし、被験者が記憶に持っている標準パターンが各個人で異なる可能性も考えられるため、標準パターンを統一する意味で、実験 IVを行った。評価実験に使用した文字は、仮名文字「い」「ろ」「は」「に」「ほ」「へ」「と」の7文字とアルファベット筆記体「a」「b」「c」の3文字とした。評価式(4)による評価では、パターンセットの評価値を計算し、評価値が大きい順に順位付けを行った。表6に各文字の実験ごとの評価式と主観的評価の相関係数を示す。各文字で高い相関が得られていることが確認でき、この結果、評価式(4)は比較的よく人間の文字の選好を定量化するものと考えられる。

(平成14年6月25日受付)

(平成15年7月3日採録)



春日 裕之(正会員)

平成14年長岡技術科学大学工学部電子機器工学課程卒業。現在(株)NSコンピュータサービス勤務。



住田 桂一

平成14年長岡技術科学大学大学院電気・電子システム工学専攻修士課程修了。現在(株)松下通信ITS勤務。



大川 慶

平成12年長岡技術科学大学大学院電気・電子システム工学専攻修士課程修了。現在(株)ハイマックス勤務。



和田 安弘

昭和55年東京工業大学工学部制御工学科卒業。昭和57年同大学大学院総合理工学研究科システム科学専攻修士課程修了。同年川崎製鉄(株)入社。平成元年ATR視聴覚機構研究所に出向。平成4年ATR人間情報通信研究所に出向。平成6年博士(工学)。平成9年より長岡技術科学大学助教授。運動制御および神経回路網の研究に従事。神経回路学会理事。電子情報通信学会、神経回路学会、計測自動制御学会等各会員。