

4D-7

力学的弛緩整合システムの 自動微分による実装

永崎 健[†] 柳田 正[‡] 中川 正樹[†][†]東京農工大学工学部電子情報工学科 [‡]富士ゼロックス株式会社

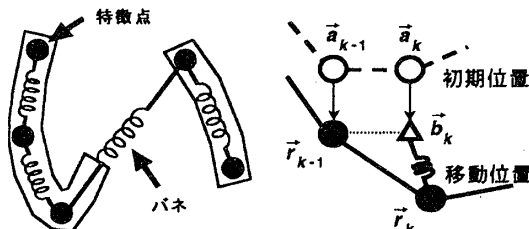
1. はじめに

本稿では、オフライン手書き文字の弛緩整合法について述べる。これまでに、文字パターンを2次元上の点集合とみなし、これを整合する手法がいくつか提案されている[1,2,3]。これらは、パターン整合をパターン間距離関数の最小化問題として定式化している。本稿では、パターン間距離関数をファジー論理から導出し定式化する。また最小化の計算を、自動微分を使った最急降下法により行う。

2. 文字パターンの力学的モデル

弾性ストロークモデルは、手書き文字の力学的モデルである。これは筆順に沿って並ぶ特徴点と、その間を結ぶバネから構成される。概念図を図1に掲げる。

本研究では、標準パターンとして弾性ストロークモデルを採用し、これを变形させることで入力パターンとの整合を行う。標準パターン R の变形度はバネの弾性エネルギー $E_{deform}(R)$ で測る。変形の度合いが低いほど、弾性エネルギーは小さいとする。



a)モデル概念図 b)相対初期位置バネ
図1 弾性ストロークモデル

3. 文字パターン間の類似度評価

標準パターンと入力パターンの類似度はファジー論理に基づいて定義する。入力パターン H を構成する黒点の集合を $\{\vec{h}_1, \vec{h}_2, \dots, \vec{h}_I\}$ 、標準パターン R を構成する黒点の集合を $\{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_J\}$ で表す。このとき文字パターン間の類似度を、次の条件：

- 1) 入力パターン上の点のいずれにも、それに近い標準パターン上の点が存在し、かつ
- 2) 標準パターン上の点のいずれにも、それに近い入力パターン上の点が存在する

が満たされる場合に、より高いと定義する。

ここで2点間の距離の近さを、ファジー論理関数 $f_{ij} = f(h_i, r_j)$ で表すとする。関数 f_{ij} の値は、2点間 (h_i, r_j) が近いほど1に近く、遠いほど0に近いとする。このとき上記の条件は、次のファジー論理式で表現できる。

$$\left[\bigcap_i \{f(h_i, r_j)\} \right] \cap \left[\bigcap_j \{f(h_i, r_j)\} \right] \quad (1)$$

この式の値が1に近いほど、2つのパターン H と R は似ていることになる。

式(1)のような論理式は計算機上での扱いが難しい。そこで論理式をエネルギー形式に変形する。エネルギーの式を次に掲げる。

$$\sum_j \{ \min_i \{ -\log(f_{ij}) \} \} + \sum_i \{ \min_j \{ -\log(f_{ij}) \} \} \quad (2)$$

式(2)は、式(1)中のOR・AND論理をmax・積演算で置き換え、式全体の $-\log$ を取った結果得られたものである。ここで式(2)を文字パターン間エネルギー $E_{pat}(H, R)$ と定義する。文字パターンの類似度の大小は、エネルギーの大小で評価できる。

4. 文字パターンの段階的整合

パターン整合の問題は、標準パターン R について目的関数：

$$E_{total}(H, R) = E_{deform}(R) + E_{pat}(H, R) \quad (3)$$

を最小化する問題として定式化される。本稿ではこの問題を最急降下法によって解く。

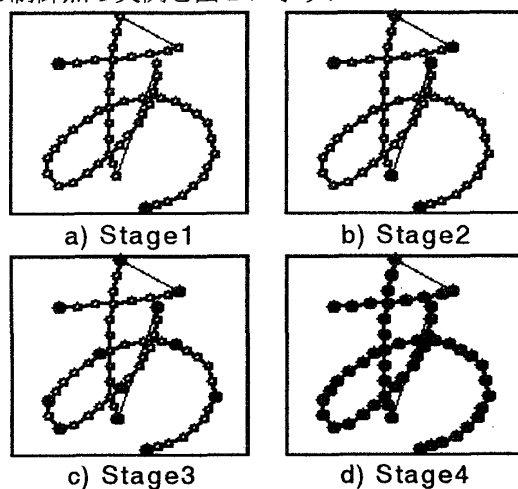
最急降下法の過程は次の漸化式で表せる。

$$\vec{R}^{(k+1)} = \vec{R}^{(k)} - \alpha \cdot \nabla_{\vec{R}} E_{total}(H, R^{(k)}) \quad (4)$$

右肩の数字(k)は最急降下法におけるk回目の反復過程を示す。 $\vec{R}$ は、標準パターン R 中の全特徴点の位置を表すベクトルである。 $\nabla_{\vec{R}} E_{total}$ は、関数 E_{total} をベクトル $\vec{R}$ の各成分で偏微分した勾配ベクトルを意味する。また、 α は移動量を制御する定数である。

しかし、式(4)のように多数の特徴点が同時に動く解法は、極小解に陥りやすい。そこで、最初に文字パターン全体の位置、傾き、大きさを合わせ、次に各ストローク、その次にストローク内の折れ曲がり、最後に全ての特徴点を合わせる、といった段階的整合を行うことにする。段階的整合は、標準パターンの特徴点の一部を制御点として選び、制御点のみを自由に動かし、その他の点は制御点

に合わせて動かすことで実現する。各段階における制御点の実例を図2に示す。



●:制御点 ○:一般特徴点
図2 段階的整合における制御点の例

5. ボトムアップ式自動微分による実装

最急降下法では目標関数の偏微分係数を求める必要がある。従来は関数の定義式から手計算により偏微分方程式を導き、これを実装していた。しかし、研究段階ではエネルギーの定義式をしばしば変更する。また、段階的整合を行う場合には、偏微分対象の変数が動的に変化する。このため偏微分方程式を手計算で解くことは、手間も大きく、過ちを犯しやすい。

そこで、自動微分による実装を行うことにする。これにより目標関数 $E_{total}(H, R)$ の定義をプログラムに記述するだけで、関数値と偏微分係数が同時に求まる。自動微分の実現方法には様々な方法がある[4]。今回は、計算速度的には不利だが、単純で実現しやすいボトムアップ方式を採用した。

ボトムアップ式自動微分では「数」を独自に定義する(AD数)。仮に偏微分対象の変数が n 個あるとすれば、AD数の構造は $AD = \{v, d_1, \dots, d_n\}$ となる。 v は関数の値を保持する場所である。また d_k ($k=1 \dots n$) は、関数を k 番目の変数で偏微分した値を保持する。

自動微分では、このAD数を用いてプログラムを組む。このとき、AD数の定数、変数、基本演算は、次のように定義される。

- (1) 定数: AD数 C は定数値 a を持つ
 $C.v = a, C.d_k = 0$ ($k=1 \dots n$)
- (2) 変数: AD数 X は値 a を持つ i 番目の変数
 $X.v = a, X.d_i = 1, X.d_k = 0$ ($k=1 \dots n, k \neq i$)
- (3) 和演算: AD数の和 $Z=X+Y$
 $Z.v = X.v + Y.v, Z.d_k = X.d_k + Y.d_k$ ($k=1 \dots n$)
- (4) 積演算: AD数の積 $Z=X \times Y$
 $Z.v = X.v \times Y.v$

$$Z.d_k = X.d_k \times Y.v + X.v \times Y.d_k \quad (k=1 \dots n)$$

(5) 初等関数: AD数の初等関数演算 $Z=f(X)$

$$Z.v = f(X.v), Z.d_k = X.d_k \times f'(X.v) \quad (k=1 \dots n)$$

(但し、 f' は f の導関数)

AD数と上記の基本演算を組み合わせることでプログラムを組むと、関数の値と偏微分係数が同時に求まる。C++などにある演算子オーバーロード機能を使えば、プログラムの外観上は、ほとんど関数定義の記述のみで済む。

6. 認識実験

本整合法について認識実験を行った。その結果を表1に掲げる。標準ボタンは各字種1ボタンとし、ETL8B見本ボタンから作成した。入力ボタンとしてETL8Bの平仮名46字種20セットを利用した。入力ボタンに対する非線形正規化、ノイズ除去などは行っていない。標準ボタンのバネとして特徴点間バネと相対初期位置バネを用い、最高認識率98.8%を得た。

表1 相対初期位置バネによる認識率[%]

k1 \ k2	0	0.1	1	10	100
0	97.3	98.8	97.5	96.3	94.8
0.1	97.3	98.4	97.5	96.3	94.7
1	97.5	97.4	97.1	96.1	95.0
10	96.0	96.1	96.0	95.9	94.1
100	95.0	94.9	94.9	94.5	93.6

k1:特徴点間バネ k2:相対初期位置バネ

7. おわりに

本稿では、文字の弾性モデルに基づくオフライン手書き文字ボタンの整合法について述べた。これは弾性ストロークモデルにより標準ボタンを構成し、文字ボタン間の類似度をファジー論理に基づいて計算する。ボタン整合の問題は、類似度から導かれた文字ボタン間エネルギーの最小化問題として定式化される。この最小化問題は最急降下法により弛緩的に解かれる。最急降下法の計算では偏微分係数の計算が必要であるが、本稿ではこれを自動微分により計算させる。これによりプログラム実装が容易になる。

本研究では、電子技術総合研究所の手書き教育漢字データベースETL8Bを利用させていただいた。

参考文献

- [1] Rodney Webster, Masaki Nakawaga: "An On-line/Off-line Compatible Character Recognition Method Based on a Dynamic Model", IEICE Transactions on Information and Systems, vol.E80-D, no.6 (Jun 1997)
- [2] 永崎健, 柳田正, 中川正樹: "弾性ストロークモデルに基づく文字パターン整合の自動微分的実現", PRMU98-141, vol.98, pp.39-46, (Dec 1997)
- [3] 若原徹, 小高和己: "GAT/LATを用いた手書き文字の適応型整形", PRMU, Vol.96, No.598, pp.43-50 (March 1997)
- [4] 久保田光一, 伊理正夫: "アルゴリズムの自動微分と応用", コロナ社, (1998)