

# 局所的ポテンシャルと最小点追跡を用いた 手書き漢字文字列の切り出し

姉 川 正 紀\* ・ 鶴 原 光 輝\*\*  
中 村 彰\*\*\* ・ 黒 田 英 夫\*  
宮 原 末 治\*

## Handwritten Chinese Character Segmentation using Local Potential Threshold and Minimum Potential Search

by

Masanori ANEGAWA\*, Kouki TSURUHARA\*\*, Akira NAKAMURA\*\*\*

Hideo KURODA\* and Sueharu MIYAHARA\*

Human can segment a character from handwritten Chinese character text line having many overlaps or joins. Therefore, a fairly high degree of accuracy in the character segmentation may be attained by imitating the human visual processing. In the cognitive science field, it is proposed that the method called "field of induction on the retina" is similar to human subjective image processing. But, there are some problems (overlaps or joins characters) in character segmentation method using field of induction on the retina.

To solve these problems, this paper proposes the local potential threshold method to imitate human movement of view point and the minimum potential search method to segment joined characters.

### 1. まえがき

文字列からの文字の切り出し精度は、その後の認識処理の精度を大きく左右する重要な要素である。しかし、手書き漢字文字列は、印刷漢字文字列に比べ重なりや接触等が多く、文字切り出しは困難である。そこで現在、試行錯誤的に文字を切り出し、その後の認識処理や知識処理（構成要素の組み合わせの妥当性をチェックする等）の結果を文字切り出し処理にフィードバックさせ、適切な切り出しを決定する手法が主流である。しかし、この様な手法では、処理負担が大きくなるため、認識処理や知識処理をおこなわず、高精度に手書き漢字文字列から文字を切り出す手法が期待されている。

これに対し人間は、重なりや接触等のある文字列から、高精度に文字を切り出す事が可能である。そこで、人間の視覚処理系を模倣し、文字切り出しに適用する事で、高精度な文字切り出しが実現出来るものと思われる。認知科学の分野において、人間の主観的（感覚的）な処理結果と比較的よく一致する“視覚の誘導場”が提案されている<sup>1)</sup>。従来、視覚の誘導場を利用した手書き漢字文字列からの文字切り出し手法が提案されているが<sup>2)</sup>、文字の重なりや接触等に弱いといった問題点がある。

本論文では、これらの問題を解決するために、人間の視点移動を模倣した“局所的ポテンシャル閾値法”と接触文字を高精度に分割可能な“誘導場最小点追跡

平成10年11月27日受理

\*情報システム工学科 (Department of Computer and Information Sciences)

\*\*日通工(株) (Nitsuko Co., Ltd.)

\*\*\*熊本工業大学電子工学科 (Department of Electronic Engineering, Kumamoto Institute of Technology)

法”を新たに提案する。手書き漢字文字列からの文字切り出しに適用した結果、本手法は、従来提案されている視覚の誘導場を利用した手法に比べ、約20%文字切り出し精度が向上した。

## 2. 視覚の誘導場

DOG(Difference Of Gaussian)フィルタやLoG(Laplacian Of Gaussian)フィルタは、人間の初期視覚系の処理結果に一致する事が知られている<sup>3)</sup>。しかし、その処理結果は、必ずしも人間の主観的(感覚的)な処理結果と一致しない。これに対し、認知科学の分野において、人間の主観的(感覚的)な処理結果と比較的良く一致する“視覚の誘導場”が提案されている。

視覚の誘導場とは、認知科学の分野において、生理学、心理学現象を基に定義されたポテンシャル場であり、式(1)で定義される。図1に示す様に、視覚の誘導場のポテンシャル場は、従来の画像処理的な手法に比べ、人間の主観的(感覚的)な処理に比較的良好に一致する。

従来、視覚の誘導場を利用した文字切り出し手法では、ポテンシャル場の円形度及び階層構造を利用して、この手法では、文字が適切に切り出されるために、個々の文字を囲む閉曲線が存在しなければならない、文字同士が接触や接近しすぎではならない等の問題点が存在する。

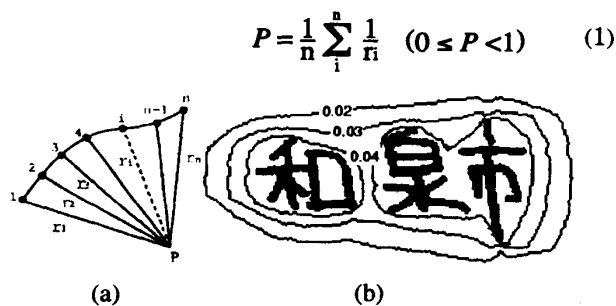


図1 視覚の誘導場

## 3. 局所的ポテンシャル閾値法

従来の手法では、文字列全体に対して、視覚の誘導場のポテンシャル場を決定していたため、適切なポテンシャル場が得られない事があった。

これらの問題を解決するために、本論文では、人間の視点移動を模倣した“局所的ポテンシャル閾値法”を提案する。

### 3.1 文字列周辺領域

本論文では、文字列周辺の空白領域と文字を切り出すポテンシャル場の関係について、予備実験をおこな

った。まず、図2(a)に示す様に、文字列の平均的な文字幅を $W$ 、文字列から一定距離離れたポテンシャル値を $P_0, P_1$ とする。 $P_0, P_1$ のいずれかポテンシャル値が高い方を基準とし、このポテンシャル値を囲む等ポテンシャル閉曲線の面積を $S_0$ とする。また、人間が文字の切り出しに最適と判断したポテンシャル値を $S_a$ とする(図2(b)参照)。なお、これらの実験は、個々の文字の特徴を抽出するために、数文字程度の短い文字列でおこなった。

この結果、図3に示すように、式(2)で定義される文字列周辺の空白領域と文字を切り出すポテンシャル値の比 $S_{ab}$ は、ほぼ0.3~0.5の範囲に分布している事が判明した。この相関関係を利用する事によって、ほぼ適切な文字の切り出しが可能である。

$$S_{ab} = \frac{S_a}{S_b} \quad (2)$$

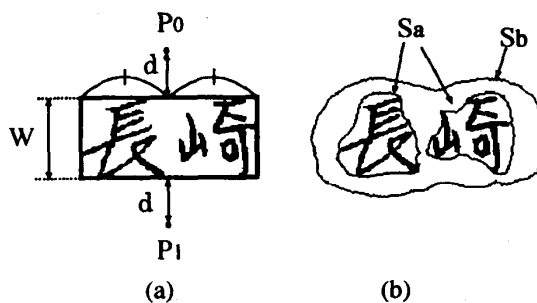


図2 文字列の周辺領域

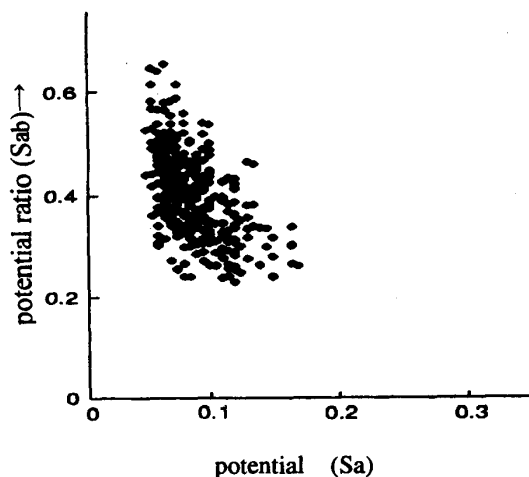


図3  $S_{ab}$ とポテンシャル( $S_a$ )の関係

### 3.2 局所マスクの設定

実際の文字列は、3.1でおこなった予備実験に比べ文字列が長く、適切なポテンシャル場を得る事が困難である。

そこで、本論文では、図4示す様に、複数の長さの局所マスクを設定する。それぞれの局所マスク内で、3.1で述べた手法を用い、文字切り出し候補となる閉

曲線を抽出する。その結果、図4に示すように、局所マスクの長さによって、異なった形状の文字切り出し候補の閉曲線が抽出される。

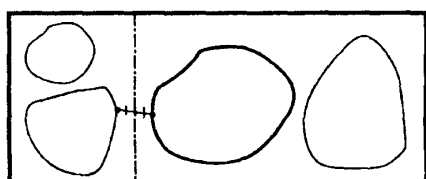


mask1(M1)      mask2(M2)      mask3(M3)

図4 局所マスクとポテンシャル( $S_a$ )の関係

### 3. 3 文字分離点の決定

個々の局所マスク内で、3.2で抽出された文字切り出し候補の閉曲線の位置関係を調べる。閉曲線が上下方向に大きく重なっている場合、これらの閉曲線は同一の文字から構成されているものとする。また、上下方向の重なりが少ない場合(本実験では、予備実験より10%未満とする)、他の閉曲線との中点を文字分離候補点とする(図5参照)。この様にして得られた複数の文字分離候補点の中で、最も座標の小さい(左側にある)点を文字分離点とする。



↑ segment point

図5 文字分離点の決定

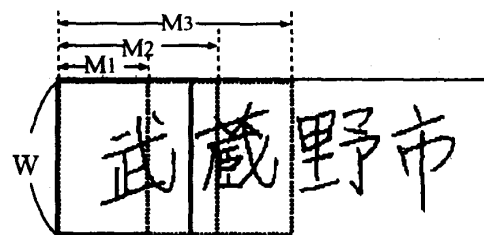
### 3. 4 局所マスクの再設定

3.3の処理で、文字分離点を発見した場合、その分離点を新たな局所マスクの開始点として、文字列の終端まで、上記の処理を繰り返す。また、文字分離点が発見できなかった場合、最も局所マスクの長さが短いマスクの終点を新たな局所マスクの開始点として、同様の処理をおこなう(図6参照)。

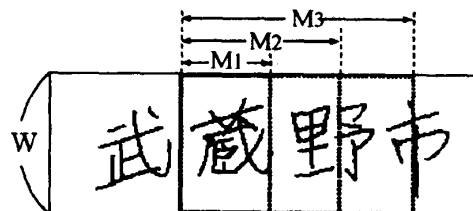
この動作は、人間が文字列を読む時に視点を移動する動作に相当している。

### 3. 5 局所ポテンシャル値の線形補間

3.4で抽出された文字分離点の座標とポテンシャル値を基に文字切り出しをおこなう。しかし、文字切り出しのポテンシャル値が不連続に変化する場合、文字切り出しに悪影響を及ぼす可能性がある。そこで、図7に示すように各文字切り出しのポテンシャル値を線形補間し、適切な文字切り出しを可能にする。



↑ segment point



↑ new mask start point

図6 局所マスクの再設定

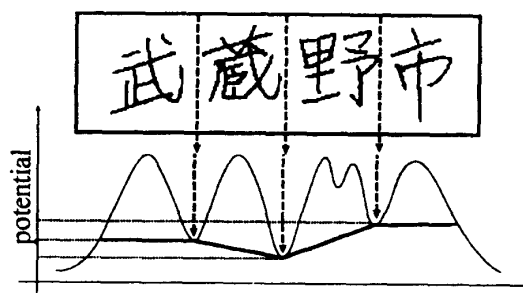


図7 局所ポテンシャルの線形補間

## 4. 誘導場最小点追跡法

局所的ポテンシャル閾値法によって、従来手法に比べ、重なりが多い文字列からも適切な文字切り出しが可能になる。しかし、文字同士が接触している場合、文字の切り出しが適切におこなえない可能性がある。

そこで、本論文では、接触した文字を切り出すための“誘導場最小点追跡法”を提案する。

### 4. 1 探索領域の設定

局所的ポテンシャル閾値法によって、切り出された文字を囲む外接矩形の面積を $S_c$ 、文字の縦と横の長さの比を $R_c$ とする。 $S_c$ 、 $R_c$ の値が式(3)の条件を満たした場合、文字同士が接触しているものと推測し、文字の分割処理をおこなう。

文字の垂直中心線から左右に式(4)の範囲を分割線探索領域とする。式(4)に示す様に、探索領域は、文字の横方向の長さ に比例して決定される。

$$S_c < W^2 \text{ and } R_c > C_0 \quad (C_0 = \text{constant}) \quad (3)$$

$$D_l = C_1 \cdot W \cdot R_c \quad (C_1 = \text{constant}) \quad (4)$$

#### 4. 2 分割線の探索

探索領域の水平中心線(DI)の各点を探索開始点とする。各探索開始点から最もポテンシャル値が低い点を通過する様に、上下方向に探索をおこなう。この時、各探索線と原画像の交差回数の最も少ない探索線を文字の分離線とする(図8参照)。

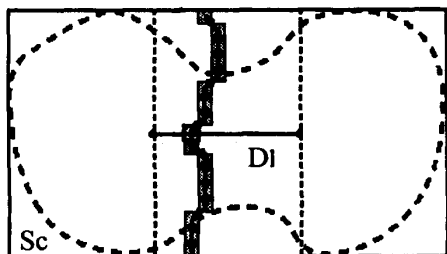


図8 誘導場最小点追跡

#### 5. 実験結果

本手法を手書き漢字文字列(56文字列)に適用し、実験をおこなった結果を表1に示す。表1より、局所的ポテンシャル閾値法は、従来手法に比べ、約10%切り出し率が向上した。また、誘導場最小点追跡法を併用する事により、全体で約20%の切り出し率が向上した。

図9に本手法の具体的な文字切り出し結果を示す。外接矩形は、局所的ポテンシャル閾値法による切り出し、外接矩形中の曲線は、誘導場最小点追跡法による切り出しを示している。接触文字に対しても、誘導場最小点追跡法により、適切な切り出しが可能である事が分かる。

表1 文字切り出し率

手法	切り出し率(%)
従来手法	70.9
局所的ポテンシャル閾値法	79.9
局所的ポテンシャル閾値法 +誘導場最小点追跡法	89.1

#### 6. まとめ

本論文では、人間の視点移動を模倣した、“局所的ポテンシャル閾値法”と接触文字を高精度に分割可能な“誘導場最小点追跡法”を提案した。これらの手法を実際の手書き漢字文字列に適用した結果、従来手法に比べ約20%の切り出し率が向上した。

文字切り出しの失敗例を調査した結果、そのほとんどが漢字以外の文字であった(図10参照)。現在、漢字以外の文字切り出しについて検討中である。

塩田町 大字谷所 甲

(a) 原画像

塩田町 大字谷所 甲

塩田町 大字谷所 甲

(b) 局所的ポテンシャル閾値法

塩田町 大字谷所 甲

(c) 誘導場最小点追跡法

図9 文字切り出し例

14-11 シャト-幸 303

図10 文字切り出し失敗例

#### 参考文献

- 1) Zensho YOKOSE: "A Study of Character Patterns Based Upon the Theory of Psychological Potential Field", Japanese Psychological Research, Vol.12, No.1, pp.18-25, 1970
- 2) Michihiro NAGASHI: "A Consideration on segmentation of Handwritten Characters using Field of Induction on the Retina", Technical Report of IEICE, Japan, PRU92-120, pp.17-24, Jan.1993
- 3) Marr, D.: "VISION - A Computational Investigation into the Human Representation and Processing of Visual Information", 産業図書, 1982