

部分領域の探索による孤立文字認識 Isolated Character Recognition by Search of Partial Regions

根岸 和也[†] 岩村 雅一[†]
Kazuya Negishi Masakazu Iwamura

大町 真一郎[†] 阿曾 弘具[†]
Shinichiro Omachi Hirotomoto Aso

1. まえがき

切り出された活字文字に対する認識精度は実用上十分であるが、それを支える文字切り出し技術には制約が多い。従来の切り出し手法は文字列(行)を抽出した後、個々の文字に分割する。従って、文字列を成さない「孤立した文字」を切り出すことは容易でない。また、数式認識においては、数式構造の複雑さから文字列の抽出は難しく、従来法の適用は困難である[1]。さらに、接触文字と呼ばれる、2文字以上が接触した文字の認識も問題となっている。

そこで、我々はアクティブ探索法[2]を利用して「孤立した文字」の切り出しと認識を同時に行う手法を提案した[3]。この手法は従来法では認識が難しかった接触文字に対しても適用でき、参照する文字と認識したい文字が同一フォントでは高い認識精度が得られたが、フォントが異なる場合を想定しておらず、そのような場合には必ずしも高い認識精度が得られなかった。

そこで本論文では、文献[3]で提案した手法を、フォントが異なる場合でも認識出来るように改良する。具体的には、参照画像をいくつかの領域に分割したのち、各部分領域の特徴を個別に検索する。検索した特徴の配置が特定の文字を表していると認められれば、そこに文字があるとみなし、文字の位置と種類が同時に決定する。このとき、文字のフォントの違いによる字形の違いを領域の位置ずれとして許容する。

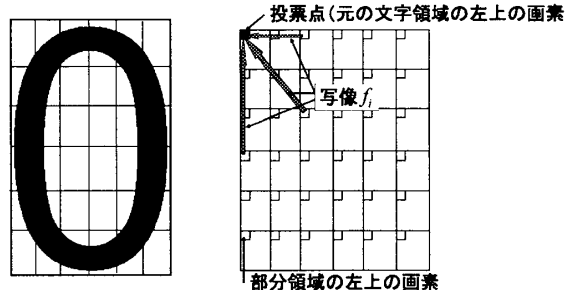
2. 提案手法

2.1 参照画像の部分領域の探索

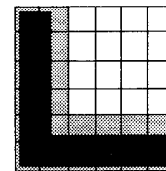
文字フォントの画像を参照画像、認識したい文字を含む画像を入力画像と呼び、入力画像中から参照画像を探索する。

本手法では、まず図1(a)のように1つの参照画像を D 個の部分領域に分割する。図1(a)は本論文の実験で用いた $D=36$ の様子を表している。次に、部分領域ごとに探索を行い、特徴が類似する領域を見つける。アクティブ探索法では色ヒストグラムを特徴量とするが、本手法では文字画像のエッジを Sobel Filter で抽出し、抽出したエッジの方向に応じたヒストグラムを特徴量として用いる。ここで、一定強度以下のエッジについてはエッジを持たない画素と判断し、これについてもヒストグラムを作成する。従って、ヒストグラムを作成する際のエッジの方向の刻み幅を θ とすると、エッジを持たない画素を考慮して、部分領域の特徴量の次元数は $\frac{360}{\theta}+1$ となる。本論文の実験では $\theta=11.25$ とした。従って、特徴量の次元数は17次元となる。

類似度が閾値以上の領域には文字が存在する可能性がある。その領域に今探索中の文字が存在するかどうかを



(a) 部分領域

(b) 投票で用いる写像 f_i 

(c) ホワイト領域 (図中の白色の領域)

図1: 部分領域と投票点への写像

判断するために投票を行う。本論文で用いた投票の詳細については次節で述べる。

参照画像を部分領域に分割した時、図1(c)の白色の領域のように、部分領域がエッジを持たない画素のみで構成される場合がある。このような領域を『ホワイト領域』と呼ぶことにする。参照画像に含まれるホワイト領域数を N_w とする。ホワイト領域を参照画像として探索を行うと、類似度1でマッチする領域が多数見つかる。ホワイト領域は文字の位置の特定に意味を成さないので、探索には用いない。

2.2 投票方法

各部分領域に類似した領域が、探索中の文字の部分領域であるかどうかを判断するために投票を行う。

まず、投票する位置について述べる。 i 番目の部分領域の左上の画素から文字領域の左上の画素への写像を f_i とする(図1(b)参照)。写像 f_i を用いる事で各部分領域で類似度が大きかったという情報は、分割前の文字画像の左上の画素(投票点)の近辺に集められる。

次に投票点付近での投票方法を具体的に述べる。 i 番目の部分領域での投票値を $v_i(x, y)$ とし、予め全ての (x, y) について $v_i(x, y) = 0$ と初期化しておく。類似度が閾値以上の画素 (X, Y) について、以下の処理を行う。

1. $f_i(X, Y)$ を中心とする半径10画素の円に含まれる

[†] 東北大学 大学院工学研究科

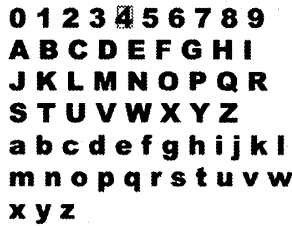


図 2: match(参照画像に 4 を使用)

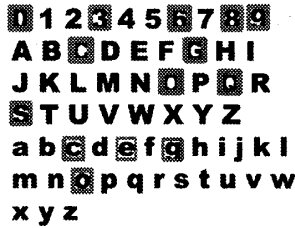


図 3: match(参照画像に C を使用)

領域 $(\sqrt{(x - f_i(X))^2 + (y - f_i(Y))^2} < 10)$ を G とする。

2. 領域 G 上で

$$g(x, y; f_i(X), f_i(Y)) = \exp\left(-\frac{(x - f_i(X))^2 + (y - f_i(Y))^2}{2 \cdot 2\sigma^2}\right) \quad (1)$$

を定義する。実験では $\sigma^2 = 10$ とした。

3. 領域 G 上の画素について、投票値 $v_i(x, y)$ を

$$v_i(x, y) = \max(g(x, y; f_i(X), f_i(Y)), v_i(x, y)) \quad (2)$$

で更新する。なお、 $v_i(x, y)$ は $v_i(x, y) \leq 1$ を満たす。

全ての部分領域での探索が終了した後、全領域を用いた投票値 $V(x, y)$ 以下の式により求める。

$$V(x, y) = \sum_i v_i(x, y) \quad (3)$$

$V(x, y)$ は、座標 (x, y) が探索中の文字画像の左上の点である可能性の高さを表す。

3. 認識実験

孤立した文字を含む画像に対する認識実験を行った。用いた参照画像は、フォント Arial、認識対象は ArialBlack の英数字 62 文字である。投票数が $V(x, y) > (D - N_w) \times 0.7$ を満たす箇所は文字が存在するとして矩形で出力する。

実験の評価方法について述べる。本手法はエッジの方向に応じたヒストグラムを用いているので、類似のヒストグラムを持つ文字同士は誤認識が起りやすい。そこで、提案手法の評価をあらかじめ決めておいた次の match と miss の基準で考える。match は探索している参照画像と同じ文字を切り出した場合で、図 2 のように

		match	miss
従来法	割合	80.7%	19.3%
	個数	50	12
提案手法	割合	95.1%	4.9%
	個数	59	3

表 1: 参照画像 62 文字の認識結果

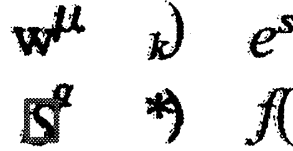


図 4: 接触文字に対する認識結果

参照画像に『4』を用いた時の結果で、切り出しおよび認識が正しく行われている場合と、図 3 のように参照画像と同じ文字を切り出しはしているが他の文字も多数切り出しをしている場合がある。miss は参照画像の文字を切り出す事が出来なかった場合を表す。参照画像 62 文字の認識結果を表 1 にまとめた。

画像中から参照画像と同じ文字を切り出せなかった文字はわずかに 3 文字だけであった。反対に同じ文字を切り出したものは 95% を越え非常に良い結果であった。他の文字も多数切り出した文字もあるものの、探索中の文字を含んでいるため、後処理を行う事で正しい認識が行えると考えられる。これにより提案手法はフォントが異なった場合でも非常に有効であることが分かる。

また、図 4 は参照画像に Arial の『S』を用い、接触文字に対する認識実験を行った時の結果である。このように、本手法は接触文字に対しても有効に切り出しと認識を行うことができる。

4. まとめ

本論文では、エッジの方向に応じたヒストグラムを用い、切り出しと認識を同時に行う手法を提案した。提案手法は参照画像を部分領域に分割し、部分領域ごとにマッチングを行い、最後に情報を統合する事でフォントの異なる文字に対しても有効に切り出しと認識ができ、実験により手法の有効性を示した。今後の課題として、今回用いた閾値はすべての文字に共通の値を用いたために精度の低下を招いてしまったので、個々の文字に適した閾値を動的に設定する必要があると考えられる。

参考文献

- [1] 内田誠一, 野村明弘, 鈴木昌和, “数学文書データベースの解析,” 信学技報, PRMU2003-234, 2004.
- [2] 村瀬 洋, V. V. Vinod, “局所色情報を用いた高速物体探索-アクティブ探索法-, ” 信学論 D-II, Vol.J81-D-II, No.9, pp.2035-2042, 1998.
- [3] 根岸和也, 岩村雅一, 大町真一郎, 阿曾弘具, “切り出しを同時に行う孤立文字認識法,” 平成 16 年度電気関係学会東北支部連合大会, 2004.