

# 方向特徴を用いた抽出対象の文字部分にのみ注目した 認識手法に関する考察

平野 真至† 山本 和彦† 加藤 邦人†

†岐阜大学工学部 〒501-1193 岐阜県岐阜市柳戸 1-1

E-mail: hirano@yam.info.gifu-u.ac.jp, {kkato, yamamoto}@info.gifu-u.ac.jp

あらまし 文字認識において、一般的な認識手法では、本来の文字にはないパターンが画像中の探索対象の文字領域内に存在すると、認識に失敗する可能性がある。そこで、本研究では方向特徴を用いて抽出対象の文字部分にのみ注目して認識を行う手法を考察する。そして、文字部分のみに注目することで、認識に無関係なパターンの影響を受けずに文字を認識することが可能であることを示し、代表的な認識手法であるテンプレートマッチングとの比較実験を行った。

キーワード MAP マッチング, 文字認識, 方向特徴, 画像処理

## 1. はじめに

一般的な認識システムでは、多くの場合、文字候補領域をセグメンテーションしてから認識を行う<sup>[1][2]</sup>。一般的な書籍や、文字が見やすいようにデザインされた印刷物のように、文字領域の背景が単純なもの、あるいは文字の位置があらかじめ既知なものにおいては、セグメンテーションは容易である。しかし、雑誌の表紙やポスター、商品のパッケージなど、文字が絵や柄を背景に描かれている印刷物の場合、背景と文字を正しく区別することは困難である。

また、テンプレートマッチング法では、認識対象の文字に別の文字や線などのパターンが重なっている場合、それらのパターンの影響を受けるため、対象となる文字を正しく認識するのは難しい。

これらの問題に対応可能な方法として、MAP(Multi-Angled Parallel)マッチングという、セグメンテーションを行わずに目的の文字を認識するアルゴリズムが提案されている<sup>[3][4]</sup>。本研究では、この MAP マッチングの手法を用いて、方向特徴を用いて抽出対象の文字部分にのみ注目した文字認識手法について考察した。また、文字部分のセグメンテーションが困難な事例として、認識対象の文字に別の線が隣接している場合における本手法の有効性について実験を行った。

## 2. 文字認識手法について

パターンマッチングの手法として一般的に利用されるテンプレートマッチングは、文字の辞書画像を、認識対象である入力画像と照らし合わせて、類似度を計算する。しかし、図1の画像のように認識対象の文字に別の線などが重なっているとき、または入力画像中の認識対象の文字の背景が複雑であるとき、文字とは別の特徴が文字領域に存在することで、類似度が下がってしまい、正しく認識できない場合がある。

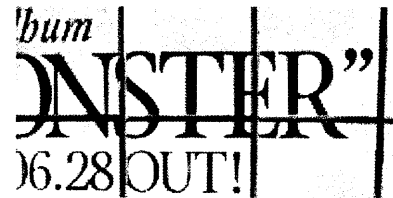


図1 文字領域に別のパターンが重なっている画像例

また、一般的に文字認識は、入力画像中の文字候補領域のセグメンテーションを行ってから認識を行う。しかし、文字領域のセグメンテーションを行うときは、図1のように、基本的に文字と思われる部分の周りに文字とは別の線などが重なると、文字領域のセグメンテーションに失敗する場合がある。

この問題に対応可能な方法として、MAP マッチングという認識手法が提案された。本研究では、MAP マッチングのアルゴリズムを用いて、認識対象の画像中の文字の部分にのみ注目し、セグメンテーションせずに認識を行う方法を考察する。

## 3. 使用する認識手法

### 3.1. 特徴抽出

MAP マッチングでは、先述した問題を解決するために、辞書画像の方向特徴面によるエッジ特徴から選択されたエッジ部分を、入力画像の方向特徴面によるエッジ特徴画像と照合する。

ここでは、エッジ抽出オペレータは Prewitt オペレータを用いた。45度刻み8方向分のエッジを抽出し、エッジの方向ごとにエッジ画像を作成する。図2に方向  $d$  の定義と8方向分の Prewitt オペレータの係数を、図3に特徴抽出例を示す。

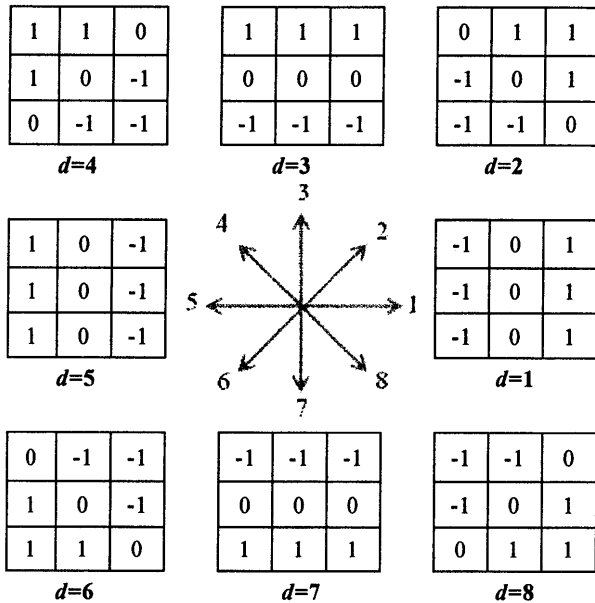


図 2 Prewitt オペレータ (d は方向を数値で示す)

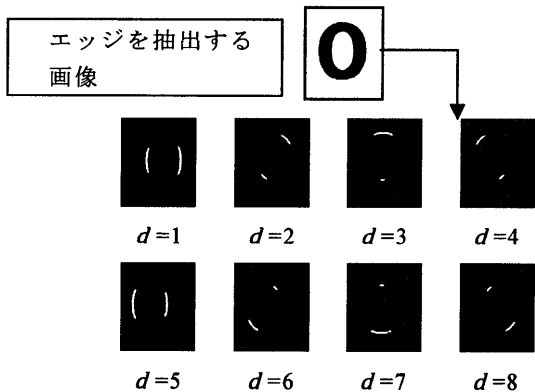


図 3 特徴抽出例

### 3.2. 評価点列の作成

辞書画像から抽出された 8 枚のエッジ画像について、画像の画素を走査し、画素値 (エッジ特徴量) がある値  $t_1$  以上であれば、その画素を探索文字の照合に利用する評価点  $R_m$  とする。ただし、 $1 \leq t_1 \leq 255$  の範囲で固定された値とし、 $m$  は 1 を始まりとする評価点の番号である。 $R_m$  は式(1)のように、エッジ特徴の方向  $d_m$ 、その画素の座標  $(u_m, v_m)$ 、エッジ特徴量の  $e_m$  からなる。 $w$  を辞書画像の横幅、 $h$  を辞書画像の縦幅とすると、 $0 \leq u_m \leq w$ 、 $0 \leq v_m \leq h$  である。また、 $M$  を評価点の総数とする。

$$R_m = \{(d_m, u_m, v_m, e_m) \mid m=1, 2, \dots, M\} \quad (1)$$

### 3.3. 投票処理

次に、評価点と入力画像中のエッジ特徴を照合し、特徴量がある程度一致しているときに、カウンタ面という入力画像と同じサイズの二次元配列の対応する座

標に投票し、加算していく。

入力画像から抽出した方向  $d_m$  のエッジ画像の座標  $(x+u_m, y+v_m)$  の特徴量を  $U(x+u_m, y+v_m, d_m)$  として、 $R_m$  のエッジ特徴量  $e_m$  と  $U(x+u_m, y+v_m, d_m)$  との差分の絶対値  $S_m(x, y)$  を取る。ここで、 $x$  の範囲は、0 以上、入力画像の横幅から辞書画像の横幅を引いた値未満であり、 $y$  の範囲も同様に 0 以上、入力画像の縦幅から辞書画像の縦幅を引いた値未満である。 $x, y$  の範囲が投票を行う範囲である。

差分値  $S_m(x, y)$  についての式は(2)となる。

$$S_m(x, y) = |U(x + u_m, y + v_m, d_m) - e_m| \quad (2)$$

$S_m(x, y)$  の値の範囲は 0 以上 255 以下であり、 $S_m(x, y)$  が小さいほど、2 点の特徴量の値が近いことを意味する。 $S_m(x, y)$  が小さいほど投票する値が大きくなるよう、投票する値  $S'_m(x, y)$  は次の式(3)で求める。また、 $S_m(x, y)$  が  $t_2$  より小さければ投票を行うよう、投票の閾値  $t_2$  を設定する。

$$S'_m(x, y) = \begin{cases} 255 - S_m(x, y) & (S_m(x, y) \leq t_2) \\ 0 & (S_m(x, y) > t_2) \end{cases} \quad (3)$$

そして、カウンタ面の座標  $(x, y)$  に、 $S'_m(x, y)$  の値を加算する。この一連の処理が投票である。

図 4 に投票の方法を示す。入力画像中の画素から、その真下にあるカウンタ面の座標へと投票を行う。一つの評価点  $R_m$  について、カウンタ面全体に投票が行われたら、次の評価点である  $R_{m+1}$  についての投票を行う。このとき、カウンタ面を  $x$  方向に  $(u_{m+1} - u_m)$ 、 $y$  方向に  $(v_{m+1} - v_m)$  移動する。評価点の位置にしたがってカウンタ面を平行移動させることで、カウンタ面の座標に対する入力画像の座標との相対位置は、 $R_m$  について投票するときと変わらず入力画像の真下のままである。ゆえに、 $R_{m+1}$  についての投票を行うと、 $R_m$  について投票したときと同じ座標に加算される。

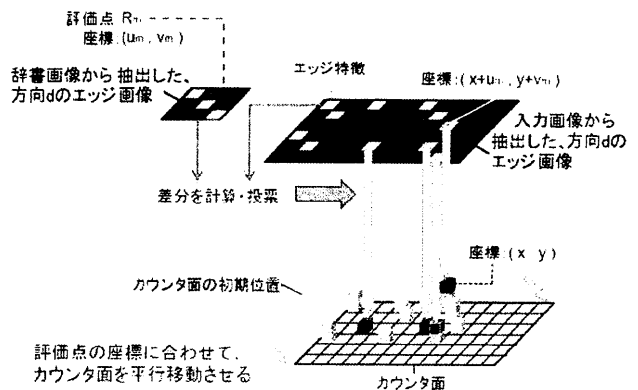


図 4 評価点  $R_m$  に対する投票

最終的にカウンタ面の座標 $(x,y)$ に投票される値の総計  $C(x,y)$ を、式(4)で表す。

$$C_m(x,y) = \sum_{m=1}^M S'_m(x,y) \quad (4)$$

$S_m(x,y)$ の最大値が 255 であるので、 $C(x,y)$ の最大値は  $M \times 255$  となる。

すべての投票の終了後、 $C(x,y)$ からパーセンテージで表される投票率  $P(x,y)$ は以下の式(5)で計算される。 $P(x,y)$ が高いほど、その領域が探索文字に近いことを表す。

$$P(x,y) = \frac{C(x,y)}{M \times 255} \times 100 \quad (5)$$

#### 4. 実験

##### 4.1. 実験概要

MAP マッチングの手法を用いて、実際にイメージキャナやデジタルカメラによって取得した画像から、目的のパターンを認識する実験を行った。

実験には図 5、6 の 2 種類の画像を使用し、それぞれ以下の画像を用いた。データ 1 については入力画像中の「E」の文字が認識対象であり、データ 2 については画像中の 3 つの「郵便局」の記号が認識対象である。また、実験の流れを図 7 に示す。



辞書画像(36×53pixels)      入力画像(193×97pixels)  
図 5 データ 1



辞書画像(38×38pixels)      入力画像(484×252pixels)  
図 6 データ 2

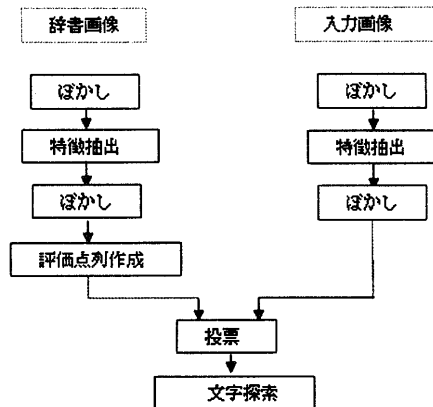


図 7 フローチャート

評価点列を作成する際の閾値  $t_1$  は、認識対象の形の微小変動に対処するためにぼかした特徴から、投票に適切な特徴量を得るために 40 とした。また、投票する際の閾値  $t_2$  は、それぞれの場合において最も対象文字部分に重点的に投票が行われるように閾値を選んだ結果、データ 1 については 50、データ 2 については 40 が最適であったため、 $t_2$  はそれぞれこの値とした。

##### 4.2. 認識結果

###### 4.2.1. データ 1 の結果

データ 1 の例を用い、カウンタ面へ投票を行った様子を、3 次元グラフで図 8 に表す。底面に垂直な方向のパラメータは、底面上の点 $(x,y)$ を中心とした領域の投票率  $P(x,y)$ である。底面はカウンタ面を表し、それが入力画像中ではどこにあたるかを明確にするために、入力画像を描画している。右下図中の中央部にある底面と平行な平面は、認識対象の文字であるかを判定する閾値を表す。すべての評価点に対して投票を行い、閾値より高いならば、その部分を認識対象の文字部分として抽出する。図 8 の右下の図から、「E」の部分の投票率  $P(x,y)$ が平面より高いため、この部分を文字「E」として認識できたことが確認できる。

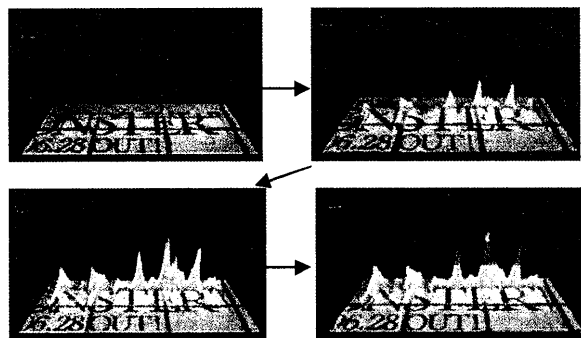


図 8 投票結果 (データ 1)

また、本手法との比較のために、テンプレートマッチ

ングを用いて認識を行い、結果を比較した。図9に実験の流れを、認識結果を表1に示す。

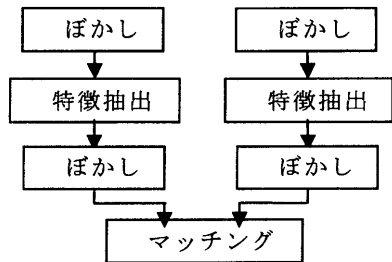


図9 テンプレートマッチングのフローチャート

表1 比較結果 (データ1)

入力画像の文字	提案手法 (投票率)(%)	テンプレート法 (類似度)(%)
① E	82.84	64.02
② R	72.10	63.02
①-②	10.73	1.00

表1から、本手法では「E」と「R」の投票率の差がテンプレートマッチングでの「E」と「R」の類似度の差よりも大きい。テンプレートマッチングでは差が1%であることから、認識対象の「E」が辞書画像に比べて変形しているような画像で認識を行う場合、「R」を「E」と誤認識する可能性がある。しかし、MAPマッチングでは差が10.73%であるゆえに、多少文字が変形しても本手法の方が正確な結果を得られやすいと言える。

4.2.2. データ2の結果

図10に示すとおり、郵便局の記号部分にそれぞれ①～③の番号を付けた。また、④と示した部分は、郵便局の記号部分以外で最も投票率の高かった部分を表す。次に、投票結果を図11に示す。さらに、画像中の3つの郵便局の記号部分の投票率を表2に示す。



図10 記号部分の番号 (データ2)

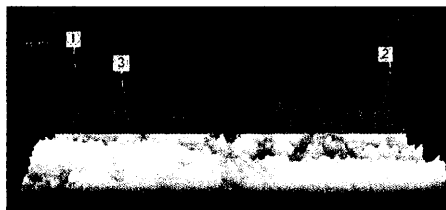


図11 投票結果 (データ2)

表2 認識結果 (データ2)

位置	記号	投票率(%)
①	郵便局	93.34
②	郵便局	85.84
③	郵便局	91.32
④	?	43.69

図11の3次元グラフおよび表2から、記号部分とそれ以外の部分の投票率の差が十分にあるため、画像中の3つの記号をすべて認識することができた。したがって、この認識手法が有効であることを確認した。

5. まとめ

本稿では、文字領域のセグメンテーションの困難な場合や、認識対象の文字に別のパターンが重なっている場合でも文字認識が可能なMAPマッチングの認識手法について考察を行った。また、セグメンテーションを行わずに、文字以外のパターンが文字に重なる場合でも認識が可能であることを確認した。そして、テンプレートマッチングと比較し、本手法が有効であることを確認した。

本手法により、認識対象の文字に線などが重なっている状況でも使用可能なアプリケーションの構築が可能と考えられる。今後の課題として、認識対象の文字の形の違いに対する対応や、類似するパターンとの識別などが挙げられる。

文 献

[1] 岩塚, 山本, 加藤: “文字認識を用いた盲導犬システムの提案”, 電子情報通信学会総大会, p.237, (2003)

[2] Norie Sasayama, Kazuhiko Yamamoto, and Kunihito Kato, “ A Proposal of Mobile Support System for Blind Person with Character Recognition ”, Proc. of 13<sup>th</sup> FCV, pp.192-195,(2007)

[3] 山田, 山本, 細川, 中村: “ MAP(Multi-Angled Parallel) 演算法と MAP 整合法による地形図の特徴抽出と認識 ”, 信学論(D-II), Vol.J-74-D-II, No.4, pp. 718-726 (1991)

[4] 澤, 山本, 加藤: “ MAP マッチングによる複雑背景からの文字探索手法 ”, 電子情報通信学会総大会, p.248,(2002)