

# 実環境文字認識のためのオートマトン制御付き最適境界検出法

酒井 恵<sup>†</sup> 内田 誠一<sup>†</sup> 岩村 雅一<sup>††</sup> 大町真一郎<sup>†††</sup> 黄瀬 浩一<sup>††</sup>

<sup>†</sup>九州大学大学院システム情報科学研究院 〒812-8581 福岡市東区箱崎6-10-1

<sup>††</sup>大阪府立大学大学院工学研究科 〒599-8531 大阪府堺市学園町1-1

<sup>†††</sup>東北大学大学院工学研究科 〒980-8579 仙台市青葉区荒巻字青葉6-6-05

E-mail: <sup>†</sup>{sakai, uchida}@human.is.kyushu-u.ac.jp, <sup>††</sup>{masa, kise}@cs.osakafu-u.ac.jp,

<sup>†††</sup>machi@ecei.tohoku.ac.jp

あらまし カメラを用いた実環境中の文字パターンの高精度認識を目標として、複比パターンと呼ばれる縞模様状のカテゴリ情報を文字パターンそのものに埋め込む方法が検討されている。この方法では、各縞の幅から計算される複比の値が文字カテゴリを表している。従ってカテゴリ情報抽出のためには、各縞の境界を検出する必要がある。本論文では、動的計画法と有限状態オートマトン (FSA) を組み合わせたロバストな境界検出法を提案する。評価実験の結果、本手法により安定して境界を検出できることがわかった。さらに検出された境界から計算した複比の値を用いて文字認識を行なったところ、非常に高い認識率が得られた。

キーワード 実環境文字認識, 境界検出, セグメンテーション, 有限状態オートマトン, 動的計画法, 情報埋め込み

## FSA-Based Optimal Segmentation for Camera-Based Character Recognition

Megumi SAKAI<sup>†</sup>, Seiichi UCHIDA<sup>†</sup>, Masakazu IWAMURA<sup>††</sup>, Shinichiro OMACHI<sup>†††</sup>, and Koichi KISE<sup>††</sup>

<sup>†</sup> Faculty of Information Science and Electrical Engineering, Kyushu University  
Hakozaki 6-10-1, Higashi-ku, Fukuoka-shi, 812-8581 Japan

<sup>††</sup> Graduate School of Engineering, Osaka Prefecture University Sakai-shi, Osaka, 599-8531 Japan

<sup>†††</sup> Graduate School of Engineering, Tohoku University, Sendai-shi, 980-8579 Japan

E-mail: <sup>†</sup>{sakai, uchida}@human.is.kyushu-u.ac.jp, <sup>††</sup>{masa, kise}@cs.osakafu-u.ac.jp,

<sup>†††</sup>machi@ecei.tohoku.ac.jp

**Abstract** For accurate camera-based character recognition, a method where category information is embedded onto every character image as a striped pattern has been investigated. The cross ratio calculated from the widths of stripes represents the category of the character. In this paper, an algorithm to detect the boundaries of the stripes for measuring the widths is proposed. The algorithm is realized by a combination of a dynamic programming-based optimal segmentation framework and a finite state automaton (FSA) which represents the characteristics of the striped pattern. Experimental results showed the ability of the proposed method on not only accurate boundary detection but also accurate recovery of the embedded cross ratio.

**Key words** camera-based character recognition, segmentation, finite state automaton, dynamic programming,

### 1. ま え が き

安価で高性能なデジタルカメラが普及してきたことを受け、カメラによる実環境中の文字認識が研究されている [1], [2]。スキャナに比べ、カメラの自由度は魅力的である。加えて、文字を手がかりとした画像検索、海外における看板翻訳、ITS における標識認識など、新たなニーズも生み出しうる。反面、カメラによる文字認識には様々な問題があることも事実である。例

えば、射影変換歪み、低解像度、非均質照明、ぼけ・手ぶれ、といった問題がある。これらの解決には、従来の文字認識手法の延長線では限界があると考えられる。

これに対し、文字パターンそのものをカメラで読みやすいように工夫しておくという新たな方法が検討されている [3]。この方法では、複比パターンと呼ばれる縞模様状のパターン (図 1(a)) を文字パターンに埋め込む。複比パターンの各縞の幅から計算される複比は射影変換に影響されないという性質に着

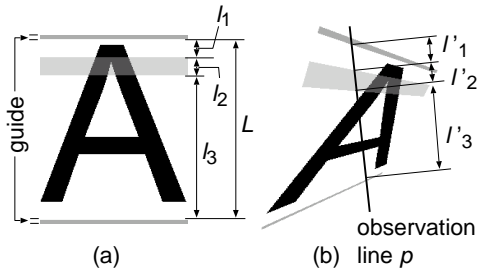


図 1 (a) 複比パターンを埋め込んだ文字画像および (b) 射影変換例

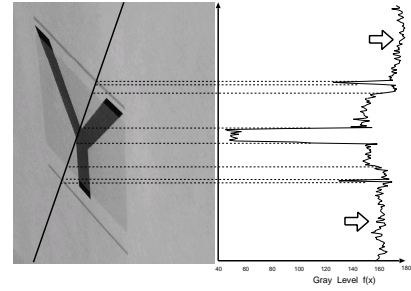


図 2 実環境中の濃淡値変動

目し、各文字に対しそのカテゴリに応じた複比パターンを割り当てておく。その結果、撮影角度に依らない安定したカテゴリ情報の抽出すなわち文字認識が実現する。

カメラで撮影した文字画像から実際に複比を計算するためには、複比パターンの縞の境界を検出する必要がある。ただしこの境界検出には 2 つの問題がある。第 1 は、濃淡値変動の問題である。すなわち、対象となる複比パターンは文字パターンとともに実環境においてカメラで撮影されて入力されるため、それらの濃淡値は照明等の影響により大きく変動し、境界検出を困難にする。第 2 は、検出される境界には複比計算に必要なもの(複比パターンの縞の境界)だけでなく不要なもの(文字線と背景の境界)が混在する点である。従って、単純なセグメンテーション処理だけでなく、複比計算に必要な境界だけを自動的に選出する必要がある。文献 [3] では、これら境界検出に係る問題については将来の課題とされ、検討されていなかった。

本論文では、第 1 の問題に対処すべく、境界検出を一種の最適化問題として定式化し、動的計画法 (Dynamic Programming, 以下 DP) に基づいた解法を提案する。さらに、第 1 および第 2 の両方の問題に対処すべく、複比パターンに関する事前知識の導入を図る。具体的には事前知識を有限状態オートマトン (Finite State Automaton, 以下 FSA) として表現し、DP アルゴリズムに組み込む。こうして得られたアルゴリズムを用いて、実際にカメラで撮影された文字パターンから埋め込まれた複比を抽出した結果、ならびにその複比により文字認識を行なった結果についても論ずる。

## 2. カメラによる文字認識のためのカテゴリ情報埋め込み [3]

本章では、文献 [3] で提案されている、複比パターンによるカテゴリ情報埋め込み、ならびに埋め込まれたカテゴリ情報抽出の基本方針について述べる。

### 2.1 複比パターンによる情報埋め込み

複比パターンとは、5 本の平行縞を 1 セットとした模様である。図 1 (a) は複比パターンを文字画像 “A” に重畳して印刷した例である。5 本の縞のうち、最初と最後の縞はガイドと呼ばれ、複比パターンの存在領域を表す。ガイドは背景領域に配置され、文字線とは交差しないものとする。残る 3 本の縞はカテゴリ情報を表す。具体的には幅  $l_1, l_2, l_3$  から計算される複比

$$r = \frac{(l_1 + l_2)(l_2 + l_3)}{l_2(l_1 + l_2 + l_3)} \quad (1)$$

により、文字カテゴリを表現する。カテゴリは離散的であるため、複比についても  $K$  通りの値  $r_k$  ( $k = 1, 2, \dots, K$ ) に量子化する。複比パターンからガイドを除いた部分の幅を  $L$  とする。すなわち  $L = l_1 + l_2 + l_3$  とする。各縞の幅の設定、すなわち  $r_k$  の設定法については様々な方式が考えられるが、本論文では  $L$  ならびに  $l_1$  を固定とし、 $l_2, l_3$  を可変とする。従って  $l_2 + l_3 = L - l_1$  を満たす  $l_2, l_3$  の組を  $K$  通り設定することで  $r_k$  を定めることになる。

各縞の濃淡値設定も様々考えられる。本論文では、図 1 (a) のように設定する。5 本の縞のうちガイドと中央の縞は有色とする。すなわち、その縞領域の濃淡値は元々の文字線・背景の濃淡値とは異なる。残る 2 つの縞は無色とする。

### 2.2 複比抽出および認識の基本方針

複比  $r_k$  を埋め込んだ文字画像から複比を抽出するための基本方針は以下の通りである。まず上下 2 つのガイドおよび文字を縦断するような直線を引く(図 1(b) 中の  $p$ )。以下この直線を検出線と呼ぶ。次に検出線上のガイドにはさまれた区間について、次章以降で述べる境界検出法により縞の境界を求め、それから検出線上での縞の幅(同図  $l'_1, l'_2, l'_3$ )を求める。これら縞の幅  $l'_1, l'_2, l'_3$  を  $l_1, l_2, l_3$  の代わりに用いて、(1) により複比を求める。複比は射影不変なので、縞の幅に誤差がなければ、どのように検出線を引いても、またどのような角度から撮影しても、埋め込んだものと同じ複比  $r_k$  すなわちカテゴリ情報が求まる。

実際には縞の境界の誤検出やデジタル画像特有の離散化誤差等により  $l'_1, l'_2, l'_3$  には誤差が入るため、複比にも誤差が入る。そこで、 $P$  本の検出線を引き、それぞれで求めた  $P$  個の複比の値から、多数決により真値を推定することにする。具体的にはまず  $P$  本それぞれの検出線について上述の方法で複比  $r$  を求め、各々  $r$  に最も近い  $r_k$  を選択する。この選択処理を全  $P$  本の検出線について行った後、最も多く選択された  $r_k$  をその文字パターンに埋め込まれた複比として採用する。

簡単のため本論文では、すべてのカテゴリに異なった複比  $r_k$  が割り当てられていると仮定する。この場合、 $K$  はカテゴリ数に一致し、上記で採用された  $r_k$  のインデクス  $k$  をそのまま認識カテゴリとすることができる。

## 3. 境界検出

### 3.1 DP に基づく最適境界検出アルゴリズム

前述のように、文字カテゴリの決定には、各検出線において濃淡値を手掛かりに複比パターンの縞の境界を求める必要があ

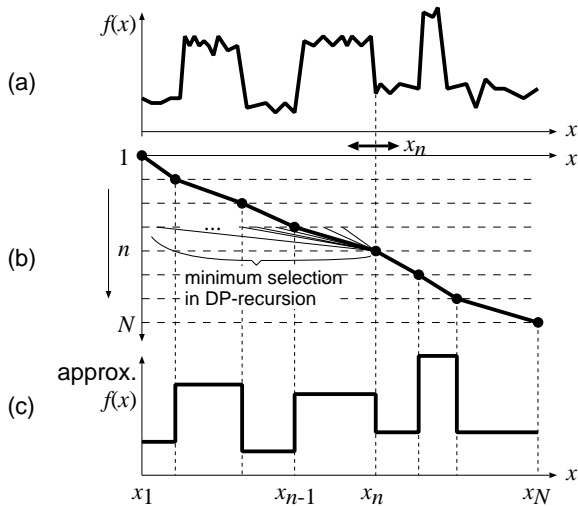


図3 (a) 検出線上の濃淡値  $f(x)$ . (b) DP による境界探索過程. (c) 階段関数近似された濃淡値および最適境界  $x_n$ .

る. この境界を単純なしきい値処理で検出するのは困難である. これは, 実際に観測される濃淡値 (すなわち撮影された文字画像の濃淡値) が照明等の影響により変動するためである. 図2はカメラで撮影された文字画像について, 1つの検出線上の濃淡値変化をプロットしたものである. 白抜き矢印で示された部分は, 背景であり元々は同じ濃淡値を持つ箇所であるが, 大きく異なった値となっている. また元々階段状に変化しているはずの濃淡値が, 細かく変動していることもわかる.

そこで本節では, 濃淡値変動に対してロバストとなるよう設計された境界検出アルゴリズムについて述べる. 具体的には境界検出を最適化問題として定式化し, DP アルゴリズムで解く.

検出線  $p$  上の位置  $x$  での濃淡値を  $f(x)$  とする. 図3(a)はその模式図である. 本手法では  $f(x)$  を同図(c)のように  $N$  ステップからなる階段関数で最適近似し, その各段の両端点として境界を検出する. ここで  $N$  は検出線  $p$  上の (両端点を含めた) 境界数である. 境界数  $N$  の決定法については, 3.3にて後述する. なお, この境界には, 複比計算に必要な複比パターンの縞の境界だけでなく, 文字線と背景の境界も含まれている点に注意されたい. これらの区別については3.2で論ずる.

関数  $f(x)$  の階段近似すなわち境界検出は, 次の目的関数  $J = J(x_1, \dots, x_n, \dots, x_N)$  の最小化問題として定式化できる.

$$J = \sum_{n=2}^N \int_{x_{n-1}}^{x_n} \|f(x) - \bar{f}_{x_{n-1}, x_n}\| dx \quad (2)$$

この最小化問題は, 区間  $[x_{n-1}, x_n]$  の観測値  $f(x)$  ( $x \in [x_{n-1}, x_n]$ ) と平均値  $\bar{f}_{x_{n-1}, x_n}$  が二乗誤差の意味で最も近くなるように区間境界を定める問題となっている<sup>(注1)</sup>. なお,  $x_1, x_N$  はそれぞれ検出線の両端座標に固定する.

目的関数  $J$  の最小化問題は,  $x_n$  を  $n = 2$  から  $N - 1$  まで順次決定していくという多段決定過程として考えることができる.

(注1): ある区間を一定値で近似する場合, その一定値として平均値  $\bar{f}_{x_{n-1}, x_n}$  を用いるのが最良である. このため, 式(2)では平均値  $\bar{f}_{x_{n-1}, x_n}$  からのずれを誤差基準としている.

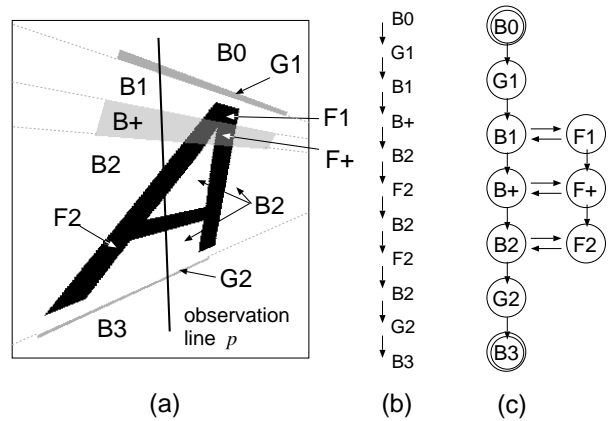


図4 (a) 各領域への状態の割り当て. (b) 検出線  $p$  上の状態変化. (c) すべての状態変化を受理する FSA.

一般に多段決定過程上での最適決定問題は DP により効率的に解けることが知られている. 具体的には以下の DP 漸化式

$$g_n(x) = \min_{x' < x} \left[ g_{n-1}(x') + \int_{x'}^x \|f(\tau) - \bar{f}_{x', x}\| d\tau \right] \quad (3)$$

を,  $n$  を 2 から  $N$  まで増やしなが, すべての  $x \in [x_1, x_N]$  において計算することで (図3(b)),  $J$  を最小化する最適境界  $x_n$  を求めることができる. ただし初期条件は

$$g_1(x) = \begin{cases} 0 & \text{if } x = x_1 \\ \infty & \text{otherwise} \end{cases}$$

とする. 最適性の原理より  $g_n(x)$  は第  $n$  境界  $x_n$  を位置  $x$  に設定する場合の最小累積誤差になる. 従って,  $g_N(x_N)$  が  $J$  の最小値となる. また (3) 計算時に最小値を与えた  $x'$  を  $x_N$  から遡る処理 (バックトラック処理) により, 最適境界が  $x_N \rightarrow \dots \rightarrow x_n \rightarrow x_{n-1} \rightarrow \dots \rightarrow x_1$  の順に求まる.

DP を使った境界検出すなわちセグメンテーションは, 本手法以外にも様々な問題において提案されている. 例えば, 連続点列の多角形近似 [4], 運動軌跡分割 [5] が挙げられる. また文字列認識の代表的な手法である recognition-based segmentation [6] においても, DP によるセグメンテーションが利用されている.

### 3.2 DP アルゴリズムへの FSA の組み込み

#### 3.2.1 FSA 組み込みの目的

前節の DP アルゴリズムにより最適境界を求めることができる. しかし, 得られた境界には, 複比計算に不要な境界 (文字線と背景の境界) が含まれているために, それら以外の境界 (複比パターンの縞の境界) だけを自動的に選出する必要がある. また, 濃淡値変動の影響により依然として境界検出に失敗する場合も多い. 例えば文字線と背景の境界など濃淡値が大きく変化する箇所に余分な境界が挿入されることがある. 以下では, これら 2つの問題を解決する方法として, 文字画像の事前知識を表現する FSA をアルゴリズムに組み込む方法について述べる.

#### 3.2.2 濃淡値変化規則の FSA 表現

検出線上の濃淡値には, 文字線, 背景および複比パターンの濃淡値に由来した一定の規則性がある. このことを見るために,

元々の文字画像を図 4 (a) のように 10 領域に分けて考える．領域 B0, B1, B2, B3 は背景領域である．いずれも元々は同じ濃淡値を持つ領域であるが，複比パターンを境界として 4 領域に区別している．また B+ は背景領域に複比パターンの中央の縞が重畳された領域である．同様に F1, F2 は文字線領域であり，F+ はそこに複比パターンの中央の縞が重畳された領域である．残る G1, G2 は上下のガイドの領域である．ここで，例えば同図 (a) の検出線  $p$  を考えると，(b) のように領域を遷移することになる．他の検出線では当然異なる領域を遷移する．しかし，どのように検出線を引いても，同図 (c) のような遷移規則に従うことがわかる．以下では，領域を状態と見なし，(c) を検出線上の濃淡値変化を表現する FSA と考える．

### 3.2.3 FSA を組み込んだ DP アルゴリズム

境界検出の高精度化ならびに複比計算に必要な境界と不要な境界の区別の自動化を目的として，3.1 の DP アルゴリズムに前節の FSA すなわち文字画像に関する事前知識を組み込む．このため，状態集合  $\Omega = \{B0, B1, B2, B3, B+, F1, F2, F+, G1, G2\}$  と  $s \in \Omega$  に先行可能な状態の集合  $\Phi(s)$  を定義する．例えば， $\Phi(B+) = \{B1, F+\}$  となる．この  $\Phi(s)$  により FSA を組み込んだ DP アルゴリズムにおける DP 漸化式は以下ようになる．

$$g_n(x, s) = \min_{\substack{x' < x \\ s' \in \Phi(s)}} \left[ g_{n-1}(x', s') + \int_x^{x'} \|f(\tau) - \bar{f}_{x',x}\| d\tau + q(s, s', x, x') \right] \quad (4)$$

このように， $g_n(x)$  は状態  $s$  を含めたものに拡張され， $g_n(x, s)$  となる．FSA 規則は，最小値選択における先行状態制約  $s' \in \Phi(s)$  と次節で述べる評価項  $q(s, s', x, x')$  に反映される．この 2 点がそのまま漸化式 (4) と (3) の違いとなっている．

DP 漸化式 (4) を  $n$  を 2 から  $N$  まで増やしながらか，すべての  $x \in [x_1, x_N]$  と  $s \in \Omega$  について計算する．ただし初期条件は

$$g_1(x, s) = \begin{cases} 0 & \text{if } x = x_1 \text{ and } s = B0 \\ \infty & \text{otherwise} \end{cases}$$

となる．最小誤差  $\min J$  は  $g_N(x_N, B3)$  で得られる．

DP に基づく境界検出アルゴリズムに FSA を組み込むという考え方は，前出の recognition-based segmentation にも見られる．そこでは単語辞書表現に FSA が利用されている．また連続音声認識 [7] や書棚画像解析 [8] でも類似の検討がなされている．ただし，FSA と DP を用いて最適階段関数近似を行い，境界を検出する方法は，他には見られないものである．特に 3.2.5 に述べるモデル評価項は本手法独特の考え方と思われる．

### 3.2.4 複比の抽出

DP 漸化式終了後， $(x_N, B3)$  からバックトラックすることで，境界系列と状態遷移系列が求まる．その状態遷移系列のうち， $B1 \Rightarrow F1$  と  $B2 \Rightarrow F2$  の計 4 種類の遷移に係る境界については文字線と背景間の境界として無視する．残る遷移に係る境界を用いて  $l'_1, l'_2, l'_3$  を求め，複比を計算することで，その検出線が表現している複比  $r_k$  を抽出できる．

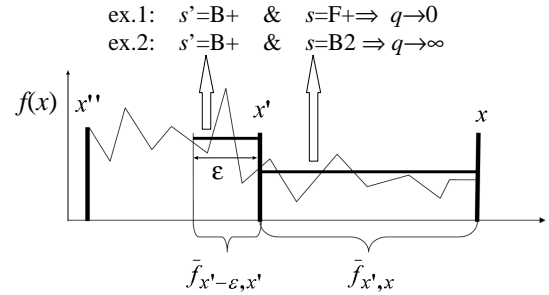


図 5 モデル評価項  $q$  の評価領域と評価例

### 3.2.5 FSA モデルに基づく評価の実際

モデル評価項  $q(s, s', x, x')$  は，前状態  $s'$  と現状態  $s$  に関する事前知識を用いて，境界検出結果を制約するものである．例えば，状態 B+ から状態 F+ に遷移する際は，濃淡値が明るくなるべきであり，もし暗くなるならばその遷移を排除すればよい．

より詳細には，図 5 に示すように，現区間  $[x', x]$  の平均値  $\bar{f}_{x',x}$  と前区間後半部  $[x' - \epsilon, x']$  の平均値  $\bar{f}_{x'-\epsilon, x'}$  の大小関係を見て，それが 2 状態  $s', s$  の本来の濃淡値の大小関係に合致するなら  $q = 0$ ，合致しないなら  $q = \infty$  とする．ここで  $\epsilon$  は正定数である<sup>(注2)</sup>．先の  $s' = B+$ ， $s = F+$  の例では， $\bar{f}_{x'-\epsilon, x'}$  より  $\bar{f}_{x',x}$  のほうが値が小さい場合 (図 5 の状況) にのみ  $q$  は 0 とし，そうでなければ  $\infty$  とする．このように上の FSA を導入した DP アルゴリズムでは，文字画像に関する事前知識に基づき平均濃淡値の変化パターンを規定している．検出される境界もそれに依って制約を受けることになり，従って境界の誤検出を低減することができる．

モデル評価項  $q$  では前区間の平均値を  $\bar{f}_{x'-\epsilon, x'}$  としているが，本来ならば  $\bar{f}_{x'', x'}$  ( $x''$  は前区間の始点) を用いるほうが自然である．ところがこの場合， $q$  は  $q(s, s', x, x', x'')$  となり，従って  $x$  (すなわち  $x_n$ ) の決定は  $x''$  (すなわち  $x_{n-2}$ ) に影響されることになる．よって  $x_n$  の決定過程が 2 次マルコフ過程となってしまう，DP では扱いにくい問題となる<sup>(注3)</sup>．この点を考慮し，今回のモデル評価項では  $\bar{f}_{x'', x'}$  の代わりに  $\bar{f}_{x'-\epsilon, x'}$  を近似値として用いて，問題を単純マルコフ過程 ( $x$  の決定は直前の決定  $x'$  にのみ依存) で扱えるようにしている．

モデル評価項に関するもう一つの工夫は，評価の中に濃淡値の絶対値を入れず，大小関係だけを入れている点である．例えば「F1, F2 ならば濃淡値は 0 に近く，F+ ならば 128 に近い」といった知識は敢えて使用していない．これは全体的な濃淡値変化により，元々の濃淡値と実際の濃淡値に乖離が生じてしまう場合への配慮である．

### 3.3 境界数 $N$ の決定

検出線上の境界数  $N$  は事前には判らない．これは認識対象となっている文字が未知であり，従って文字線と背景の境界数も未知となるためである．よって境界検出に際しては，境界数

(注2)：ガイドのように狭い区間にも対応するべく，実際には  $\epsilon$  の値を前状態  $s'$  に応じて適宜切り替えている．このため厳密には  $\epsilon(s')$  と表記すべきであるが，ここでは簡単のため単に  $\epsilon$  とした．

(注3)：過去 2 回分の決定 ( $x, x'$ ) を一括して扱うことで 2 次マルコフ過程を単純マルコフ過程に帰着させることも可能だが，計算量が急激に増加する．



図 6 複比パターンを埋め込んだ文字画像

$N$  の最適推定も併せて行なう必要がある。

本手法では、 $N$  を最小値 6 から最大値 12 まで変えながら、それぞれの  $N$  で境界検出を行い、最終的なコスト  $g_N(x_N)$  が最小となる  $N$  を境界数として決定するという、比較的単純な方針を採った。ここで重要なのは、 $N$  を増やす度に DP 計算を最初からやりなおす必要はない点である。具体的には、境界数を  $N$  と仮定したときの結果が得られていたとすると、(4) を用いて  $g_N(x)$  から  $g_{N+1}(x)$  を求めさえすれば、境界数を  $N+1$  としたときの結果が求まったことになる。

#### 4. 実験

本章では前章で述べた境界検出法の評価実験を行なった結果について述べる。結果の評価には、複比抽出精度に加え、2.2 の方針で求めた認識精度も用いた。

##### 4.1 実験諸元

プリンタを用いて白紙に複比パターンを重畳した文字画像を印刷した。フォントは Arial とし、英語大文字 26 カテゴリを用いた。文字画像を図 6 に示す。各カテゴリに異なる複比を埋め込むという方針により、複比パターンの種類  $K$  は 26 となる。原理的にはどの複比パターンをどのカテゴリに埋め込んでもよいが、ここでは複比パターンの境界と文字線・背景の境界が重ならないように配慮した。これは FSA で想定していない濃淡値変化 (例えば  $B+ \rightarrow F2$ ) の発生を極力避けるためである。

この文字画像をカメラにより撮影した。撮影方向はカメラと紙面を正対させたものを 1 方向、カメラと紙面を正対させていないもの 8 方向の計 9 方向である (図 7)。従って本実験では  $26 \times 9 = 234$  画像を使用した。撮影環境を室内とし、カーテンを締め切って蛍光灯の光を照明とした。画像サイズは最小で  $300 \times 225$ 、最大で  $493 \times 370$ 、平均  $381 \times 291$  となった。なおカメラ撮影時はカラー画像であるが、これを濃淡画像 (グレースケール) に変換したものを処理対象とした。

こうして撮影された画像を上下 2 方向から走査することで、まずガイドを検出した。この検出の際に限り、背景が元々白地であることを利用したしきい値処理を用いた。次に、上下ガイドをそれぞれ 10 等分割し、9 つの分割点を求めた。この上下それぞれの 9 分割点を互いにすべて結ぶことで、各文字画像に  $P = 81$  本の検出線を定めた。

##### 4.2 境界検出結果

まず、境界検出の結果を図 8 を例として考察する。同図は FSA の有無による境界検出精度の違いも表している。すなわち

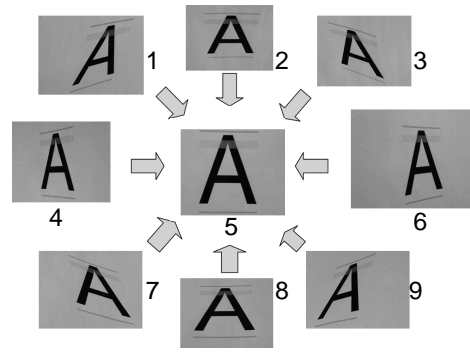


図 7 画像の撮影方向

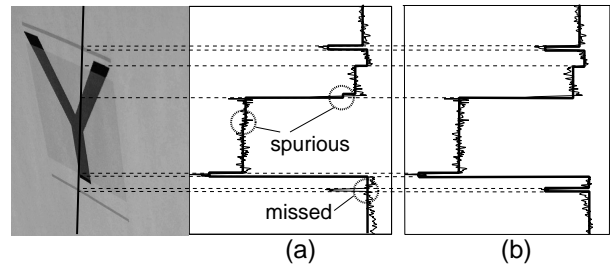


図 8 境界検出結果の例。(a) FSA なし。(b) FSA あり。

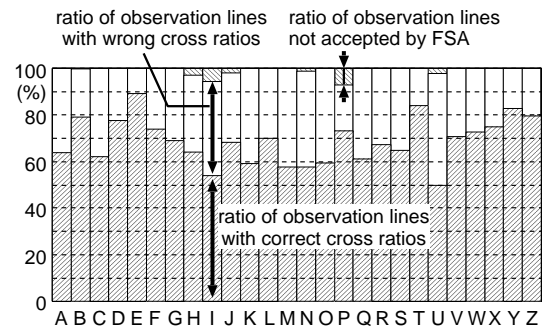


図 9 文字種別複比抽出成功率

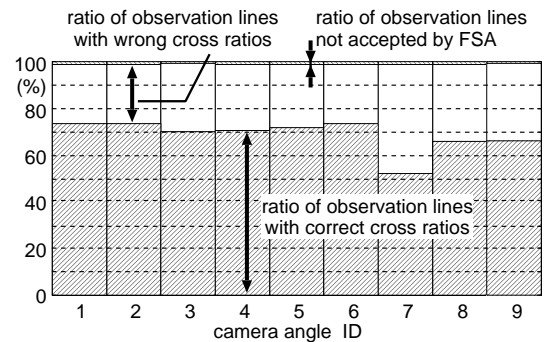


図 10 撮影方向別複比抽出成功率。横軸の撮影角度 ID は図 7 中の数字に対応。

同図 (a) は FSA を組み込まず、(2) の最小化問題を単純に DP で解いた場合の結果であり、(b) は FSA を組み込んだ場合の結果である。FSA を組み込まない場合、不要な境界が 2 つ発生し、逆に下部ガイドに関する境界が 2 つとも検出されていない。一方、FSA を組み込んだ場合、すべての境界が正しく検出されている。この改善は、同図 (a) の誤った階段近似 (5 回連続で平均濃淡値が減少) が FSA で受理されず、結果から排除された

ためと考えられる．このように FSA により濃淡値変化の事前知識を導入することで，境界検出精度を向上できている．

#### 4.3 複比抽出成功率

次に，複比抽出成功率により本手法の精度を定量的に評価する．この複比抽出成功率とは，全検出線のうち正しい複比  $r_k$  を与えた検出線の割合である．図 9 および図 10 は，それぞれ文字種別および撮影角度別の複比抽出成功率である．詳細には，図 9，図 10 はそれぞれ  $81 \times 9$ ， $81 \times 26$  本のうち正しい複比を与えた検出線の割合が成功率となる．いずれの図からも半数以上の検出線において正しい複比が抽出されていることがわかる．複比が正しいということは，その検出線上のすべての境界検出に成功していることを意味している．従って，検出線ではなく，各境界を単位とした境界検出率（それぞれの境界が検出できているか否か）を考えると，50%よりも遥かに高い値になっていると考えられ，本境界検出手法は十分な精度を持つものと考えられる．また，図 10 からわかるように，成功率は撮影角度にほとんど影響されない．これは複比の射影不変性を実証している．

複比抽出の失敗要因には次の 3 つがある．

- 第一は，境界数  $N$  の誤推定である．本実験で用いた全検出線  $81 \times 234$  本のうち，正しい  $N$  が推定されたのは，80.9%であった．すなわち，2 割の検出線では境界が過剰若しくは過少検出されており，その影響があったと言える．

- 第 2 は，境界位置を間違っても FSA に受理される場合が存在するためである．すなわち，FSA によるチェックを通過してしまう場合である．例えば，境界位置の検出誤差が比較的小さな場合が相当する．

- 第 3 は，濃淡値変化  $f(x)$  が FSA に従わない場合である．文字線・背景境界と複比パターン境界が一致してしまった場所は，濃淡値変化が FSA で想定されているものにならず，境界が求まらない．ただし，4.1 で述べたように，複比パターン埋め込みの時点で，こうした一致はなるべく起こらないように配慮したため，結果的にあまり大きな問題にはなっていない．実際，図 9，10 には，FSA 受理のために複比抽出に失敗した検出線の割合も表されているが，非常に少ないことがわかる．

#### 4.4 文字認識率

2.2 で述べたように，本手法では全 81 本の検出線で得られる複比の多数決で認識カテゴリ  $k$  を決定する．こうして得られた文字の認識率は  $99.6\% = 233/234$  であった．すなわち誤った認識結果はわずかに 1 例であった．前節で述べたように，半数以上の検出線で正しい複比が求まったことは，ほとんどの文字画像で多数決により正しい複比が選出されることを意味しており，本結果を裏付けている．

失敗した 1 例は，“O” ( $k = 14$ ) の “N” ( $k = 13$ ) への誤認識であった．81 本の検出線のうち，27 本で  $r_{13}$ ，26 本で  $r_{14}$  が求まり，多数決の結果  $r_{13}$  すなわち “N” が選ばれていた．

### 5. む す び

カメラによる文字認識のために，縞模様状のパターン — 複比パターン — を文字に埋め込む手法が検討されている．縞の幅から計算される複比の値が，射影変換に依らず抽出可能なカ

テゴリ情報を表す．ただし実際に複比を抽出するためには，濃淡値を手掛かりとして縞の境界を検出する必要がある．そこで本論文では，動的計画法 (DP) に基づく最適化アルゴリズムとして境界検出法を設計した．また，高精度化と複比計算に必要な境界と不要な境界の自動識別を目的として，複比パターンの濃淡値変化規則を表現した有限状態オートマトン (FSA) を導出し，さらに DP アルゴリズムへの組み込みを提案した．評価実験の結果，照明等の影響による濃淡値変動があっても，本手法により安定して境界を検出できることがわかった．さらに検出された境界から複比を計算し，それをういて文字認識を行なったところ，9 種類の角度から撮影された英語大文字 234 文字画像のうち 233 文字を正しく認識できることを確認した．

今後の課題としては以下が挙げられる．

- 複比パターンのデザインの再検討：本論文のような比較的目的易い複比パターンではなく，影付き文字の影の部分を利用したカテゴリ情報埋め込み [9] も考えられる．また，文字列を認識対象とするためのデザインも重要である．

- 形状による文字認識の併用：本論文では文字種と複比パターンを 1 対 1 としていたが，カテゴリ数が多くなると多対 1 の関係になり得る．こうした場合，文字形状を援用してカテゴリを唯一に定める必要がある [3]．またその際，どのカテゴリにどの複比を埋め込むかが問題になる．この点についても文献 [10] の結果を参考にしながら検討する予定である．

謝辞 本研究の一部は，科学研究費補助金（若手研究 (B) 17700198），H17 年度公益信託マイクロソフト知的財産研究助成基金研究助成，マツダ財団第 21 回マツダ研究助成によった．

#### 文 献

- [1] 黄瀬浩一，大町真一郎，内田誠一，岩村雅一，“カメラを用いた文字認識・文書画像の解析の現状と課題，” 信学技報，PRMU2004-246，2005.
- [2] D. Doermann, et al., “Progress in camera-based document image analysis,” Proc. ICDAR, pp. 606-616, 2003.
- [3] 内田誠一，岩村雅一，大町真一郎，黄瀬浩一，“カメラによる文字認識のための付加情報埋め込みに関する検討，” 信学論，vol. J-89D, no. 2, pp. 344-352, 2006.
- [4] J.-C. Perez and E. Vidal, “Optimum polygonal approximation of digitized curves,” Pattern Recognition Letters, vol. 15, pp. 743-750, 1994.
- [5] R. Mann, A. D. Jepson, and T. El-Maraghi, “Trajectory segmentation using dynamic programming,” Proc. ICPR, vol. 1 of 4, pp. 331-334, 2002.
- [6] R. G. Casey and E. Lecolinet, “A survey of methods and strategies in character segmentation,” IEEE Trans. Pat. Anal. Mach. Intell., vol. 18, no. 7, pp. 690-706, 1996.
- [7] H. Sakoe, “A generalized two-level DP-matching algorithm for continuous speech recognition,” Trans. IEICE, vol. E-65, no. 11, pp. 649-656, 1982.
- [8] 平 英二，高山誠悟，内田誠一，迫江博昭，“モデル当てはめによる書棚画像解析，” 信学論，vol. J87-DII, no. 2, pp. 565-573, 2004.
- [9] 大町真一郎，岩村雅一，内田誠一，黄瀬浩一，“高精度実環境文字認識のための面積比による付加情報埋め込み，” 信学技報，PRMU2005-213, 2006.
- [10] 岩村雅一，内田誠一，大町真一郎，黄瀬浩一，“情報付加による認識率 100%の実現 —人にも機械にも理解可能な情報伝達のために—，” 画像の認識理解シンポジウム (MIRU2005), IS2-85, 2005.