

THE TRANSACTIONS OF THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS D-II

cic | **電子情報通信学会**
D-II | **論文誌** VOL.J80-D-II NO.10 OCTOBER
情報・システムⅡ—情報処理 ————— 1997

情報・システムソサイエティ

社団法人 **電子情報通信学会**

THE INFORMATION AND SYSTEMS SOCIETY

THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS

本方式は、大きく領域分割部と文字領域抽出部からなる。領域分割部では、数学モルフォロジーを用いて入力画像である濃淡画像の領域幅を求め、得られた領域幅に基づきいくつかの部分画像を生成する。文字線幅と文字サイズには相関が見られるので1枚の部分画像には似通った大きさの文字が存在する。領域幅検出はモルフォロジー演算を用いて記述している。文字領域抽出部は、更に、特徴強調・文字候補領域抽出・整形の三つの処理からなる。特徴強調では、部分画像に対し、雑音の除去、文字領域の強調を行う。次に、この画像のヒストグラムを計算し、そのピーク値を利用して文字領域の抽出を行う。最後に、抽出された文字候補領域に対して条件付き dilation (意味は 3. に述べる) を利用して整形を経て文字領域が得られる。2. では、モルフォロジーについて簡単に述べる。3. では、本提案の方法を詳細に述べる。4. では、本方式の有効性の検証を行うために雑誌の黒白およびカラー表紙画像を用いた抽出実験を記述する。結果および考察は 5., 6. では、まとめを記述する。

2. 構造化要素について

2.1 モルフォロジー

モルフォロジー [10] は、2 値画像を対象とするものと濃淡画像を対象とするものに分けることができる。その基本的な演算は dilation, erosion, opening, closing の四つであり、それぞれの演算記号を、2 値画像に対しては $\oplus, \ominus, \circ, \bullet$ で示し、濃淡画像に対しては添字 g を付け $\oplus_g, \ominus_g, \circ_g, \bullet_g$ で示す。例えば 2 値画像の dilation であれば $A \oplus B$ と記し、 A を被演算画像 (図形) B を構造化要素と呼ぶ。dilation はミンコフスキー和とも呼ばれ、二つの集合 A, B の任意の要素のベクトル和を求める演算である。より直観的には、集合 B の原点を集合 A のすべての要素点上へ移動させたときの B の軌跡ということになる。erosion は、ミンコフスキー差とも呼ばれ、集合 B の原点を x に移動したとき集合 B が集合 A に完全に包含される x を求める演算である。より直観的には、集合 A の原点を集合 B のすべての要素上に移動したときの集合 A の積集合と言うことができる。opening は、同一の構造化要素を用いた erosion, dilation の連続的な演算であり

$$A \circ B = (A \ominus B) \oplus B \quad (1)$$

である。同様に closing は、

$$A \bullet B = (A \oplus B) \ominus B \quad (2)$$

である。

2.2 構造化要素の選定

使用する構造化要素 (Structure Element) は次の 2 種類である。

- (1) 円盤 (Disk)
- (2) 正方形 (Square)

この 2 種類の構造化要素は、次のとおり定義される。

- 半径 i の円盤型の構造化要素: $r_i B_{disk}$
- 辺の長さ $2i$ の正方形の構造化要素: $r_i B_{square}$

この 2 種の構造化要素のうちで本方式では正方形を用いる一つの演算を除いて残りすべてのモルフォロジー演算において円盤を用いた。円盤を用いた理由は方向に依存しない対称性を必要としたためである。また、正方形を用いた箇所は試行錯誤による選択である。

この 2 種類の構造化要素の構成法について説明する。ある単純な構造化要素 B 、例えば、2 次元空間 E^2 中の円盤型 (Disk)、正方形 (Square) および三角型 (Triangle)、あるいは、3 次元空間 E^3 の球体 (Sphere) および立方体 (Cube) を考える。サイズ r_i のとき、その構造化要素の集合 X_i は、次のように与えられる。

$$X_i = r_i B \quad (3)$$

もし、 B が有界凸状であるという条件を満足すれば、整数 r_i に対して式 (3) は、

$$X_i = B \oplus B \oplus \dots \oplus B \quad (r_i \text{ times}) \quad (4)$$

となる。

本研究では、式 (3), (4) に基づく、一連の構造化要素の集合 $r_i B_{square}$ と $r_i B_{disk}$ を作成する。 $r_i B_{square}$ は、単純性を持ち、式 (3) より循環的に作成される。すなわち

$$r_i B_{square} = r_1 B_{square} \oplus r_1 B_{square} \oplus \dots \oplus r_1 B_{square} \quad (r_i \text{ times}) \quad (5)$$

$r_i B_{disk}$ では、単純な 3×3 元要素からの構成はできないため、元要素 $r_1 B_{square}$ と $r_1 B_{rhombus}$ (斜方体) の両方を用いなければならない。

$$r_i B_{disk} = r_1 B_{rhombus} \oplus r_1 B_{square} \oplus \dots \oplus r_1 B_{square} \quad (r_i \text{ times}) \\ (i = 2n : r_1 B_{square}; i = 2n - 1 : r_1 B_{rhombus}, n = 1, 2, 3, \dots) \quad (6)$$

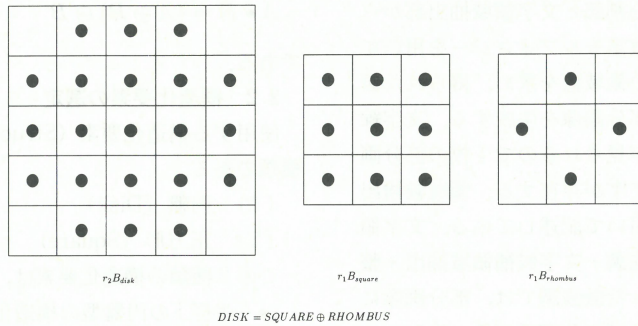


図1 構造化要素 r_2B_{disk} , r_1B_{square} と $r_1B_{rhombus}$
 Fig.1 Structure element r_2B_{disk} , r_1B_{square} and $r_1B_{rhombus}$.

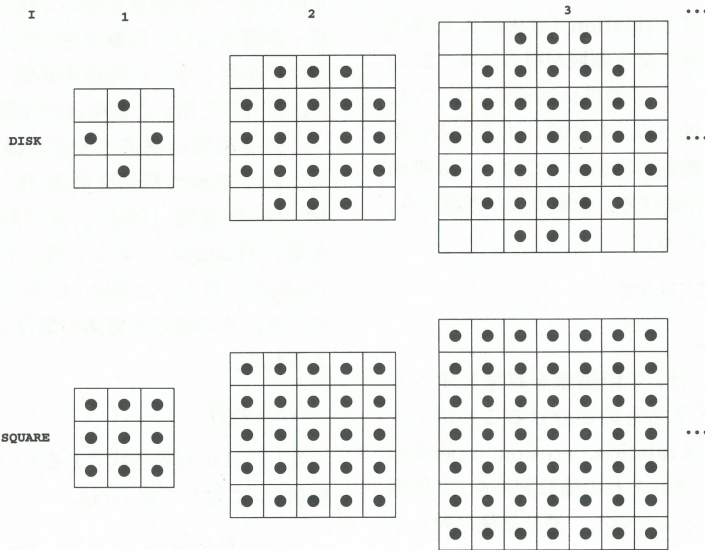


図2 構造化要素 B の例とその相似拡大集合 $r_i B$ ($i=1,2,3,\dots$)
 Fig.2 Example of figures used as structure element B, and their n-multiplied figures $r_i B$. ($i=1,2,3,\dots$)

このような構成方式の例を図1に示した。上記の2種類の構造化要素集合を図2に示す。

3. 文字領域抽出方式 [11]

3.1 領域分割部

領域分割処理は、次のモルフォロジカルアルゴリズムにより行う。その原画像 X_0 は濃淡画像であり、部分画像は2値画像に変換される。

$$X_i = |(X_0 - X_0 \circ_g r_i B_{disk}) - (X_0 - X_0 \circ_g r_{i-1} B_{disk})|_B - X'_{i-1}$$

$$X'_i = \bigcup_{0 < j \leq i} X_j, \quad X'_0 = \emptyset \quad (7)$$

あるいは

$$X_i = |(X_0 \circ_g r_i B_{disk} - X_0) - (X_0 \circ_g r_{i-1} B_{disk} - X_0)|_B - X'_{i-1}$$

$$X'_i = \bigcup_{0 < j \leq i} X_j, \quad X'_0 = \emptyset \quad (8)$$

ここで、 $i = 1, \dots, i_{max}$

X_0 濃淡画像の原画像、

X'_i 分割された2値化画像の集合、

$| \cdot |_B$ 特定のしきい値(濃淡画像では $B=2$, カラーでは $B=31$)で2値化処理

式(7), 式(8)を入力画像に対してそれぞれ独立に適用する。式(7)の X_i には、 $r_i B_{disk}$ の直径に等し

い幅をもつ、周囲より浮き出した領域が順次求まる。従ってその論理和である X'_i には $r_i B_{disk}$ 以下の幅をもつ周囲より浮き出した（大きい階調値をもつ）領域が累積される。式 (8) も同様であるがここでは周囲よりも沈み込んだ（小さい階調値をもつ）領域が求まることになる。なお、この両式はそれぞれ i_{max} 回反復されるがこの i_{max} は、抽出できる最大の文字線幅 ($2i_{max} + 1$) を指定することになるため最大の文字線幅が抽出されるように選定する必要がある。なお、上式に示した 2 値化しきい値については実験的に決定したものである。

3.2 文字領域抽出部

3.2.1 特徴強調

部分画像は、文字領域と雑音が共存している。そこで、本研究では、まず、文字領域の特徴を強調するため次のアルゴリズムで特徴強調処理を行う。

$$E_i = \begin{cases} ((X'_i \circ r_{i-1} B_{disk}) \\ \oplus r_{i+1} B_{disk}) \\ \circ r_{2i} B_{disk}) \times X_0 & (i \leq 10) \\ (X'_i - X'_{10}) \times X_0 & (i > 10) \end{cases} \quad (9)$$

ここで、 \times なる演算は同サイズ画像間の算術積を表す。 X'_i は $r_i B_{disk}$ の直径以下の幅をもつ領域が累積されている。式 (9) ではひとまわり小さなディスク ($r_{i-1} B_{disk}$) と $r_i B_{disk}$ とを用いて領域の欠けの修復、孤立領域の除去を行った後、原画像との算術積により濃淡画像に変換している。なお、領域幅の大きい範囲 ($i > 10$) では原画像との算術積のみとした。次に、ヒストグラムを用いて文字領域の階調値の選択を行う。

3.2.2 領域抽出

特徴が強調された部分画像 E_i の中に、文字領域の階調値がピークをもつ。階調ヒストグラムを求めて、その曲線中のすべてのピーク値の平均値より大きいピーク値を検出することにより文字領域のもっている階調を得ることができる。異なる階調をもっている文字列が存在するためにピークは複数存在し、その中で平均値より大きいすべてのピーク値を用いて 2 値化を行っている。なお、2 値化は原理的にはヒストグラムに対する opening によるピーク検出に基づいた自動化が可能であるが現段階では半自動である。ヒストグラム手法により得られた階調によって、部分画像 E_i に 2 値化処理を施して 2 値画像 H_i を作成する。この画像にはまだ若干のノイズが存在し、ノイズ処理のため、次の整形処理を行う。

3.2.3 整形

条件付き dilation を用いた次のアルゴリズムにより整形処理を実施する。

$$\begin{aligned} R_{i0} &= H_i \circ r_{i-1} B_{square} \\ R_{in} &= (R_{i(n-1)} \oplus r_5 B_{disk}) \cap |X_0|_B \\ &\text{if } R_{ik} = R_{i(k-1)} \text{ then stop} \end{aligned} \quad (10)$$

ここで、 $|X_0|_B$ は原画像の 2 値化画像であり、その 2 値化には E_i から H_i を生成したときと同じしきい値（複数）を用いている。この整形処理は、まず H_i から小さめの構造化要素 ($r_{i-1} B_{square}$) を用いた文字領域の核を取り出す。この核を条件付きの dilation により順次太めることにより文字領域を得る。その際条件付き dilation の制約条件として $|X_0|_B$ を用いた。これにより最終的な文字領域抽結果が得られる。結果を表示するために画像 R_{ik} を 1 枚の画像に統合する。

$$X_r = \bigcup_{1 \leq i \leq i_{max}} R_{ik} \quad (11)$$

4. 実験結果

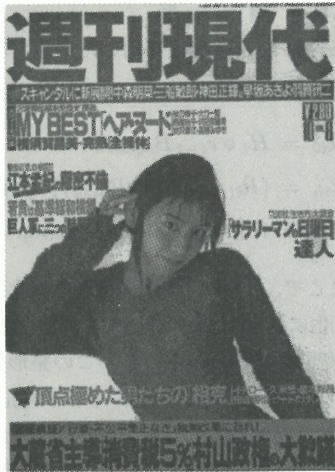
4.1 実験対象画像

表紙画像の 1 カテゴリーと考えられる雑誌の表紙画像には、比較的平坦な背景部と文字部からなるものから、絵と文字が相当入り込んだものまでさまざまなバリエーションが存在し、表紙画像として十分な一般性があると考えられる。そこで、実験対象としては、雑誌の表紙画像を主として用いた。提案方式の適用性を検証するため濃淡画像とカラー画像の両方に対して実験を行った。

4.2 実験

実験対象画像である雑誌の表紙を画像として取り込む。パソコンに接続されたスキャナによってこれらの画像データを読み取り、EWS 上にインストールし画像データベースを構築した。A4 の画像は 100dpi, 1170 × 848 の画像として蓄積される。その代表例を図 3 に示す。

実験の手順は、3. に詳述したとおりであるが、領域分割部については、式 (7), (8) を用いた。その復帰の制約条件を $i_{max} = 40$ とした。これは、実際の対象画像を統計的に調査し、最大文字線幅が 81 画素 ($r_{40} B_{disk}$ の直径) より小さいと考えられたからである。図 3 の (b) に対して、部分画像 X'_4, X'_{10} と X'_{30} は図 4 のようになる。但し図 4 は、式 (8) で得られたものである。



(a) 週刊現代



(b) CarEx



(c) 少年マガジン (濃淡)



(d) 少年マガジン (カラー)

図3 原画像の例
Fig.3 Examples of source images.

文字領域抽出部により、抽出された部分画像 H_{30} と整形した部分画像 $R_{30,27}$ の一例として図5に示す。ここで、(a)は、図4の(c)に対して文字領域抽出した結果であり、(b)は、それに続いて整形処理を施した結果である。図の中の白抜き部分は分割あるいは抽出された文字領域である。

カラー画像の場合については、まず原画像をRGB画像に変換し、それぞれ三つの濃淡画像に対して提案された本手法によって実験を行った。文字領域抽出部では、カラー画像の特徴に対応して本方式の一部の改

良を行った。式(9)の代わりに、次の式を導入した。

$$E_{ki} = \begin{cases} (((X'_{ki} \circ r_{i-1} B_{disk}) \\ \oplus r_{i-1} B_{disk}) \\ \text{or } r_{2i} B_{disk}) & (i \leq 10) \\ (X'_{ki} - X'_{k10}) & (i > 10) \end{cases} \quad (12)$$

$$H_{ki} = \bigcap_{m \in c \cap k^c} |E_{ki} \times X_{m0}|_B$$

$$X_0 \Rightarrow X_{c0} = \{X_{R0}, X_{G0}, X_{B0}\}$$

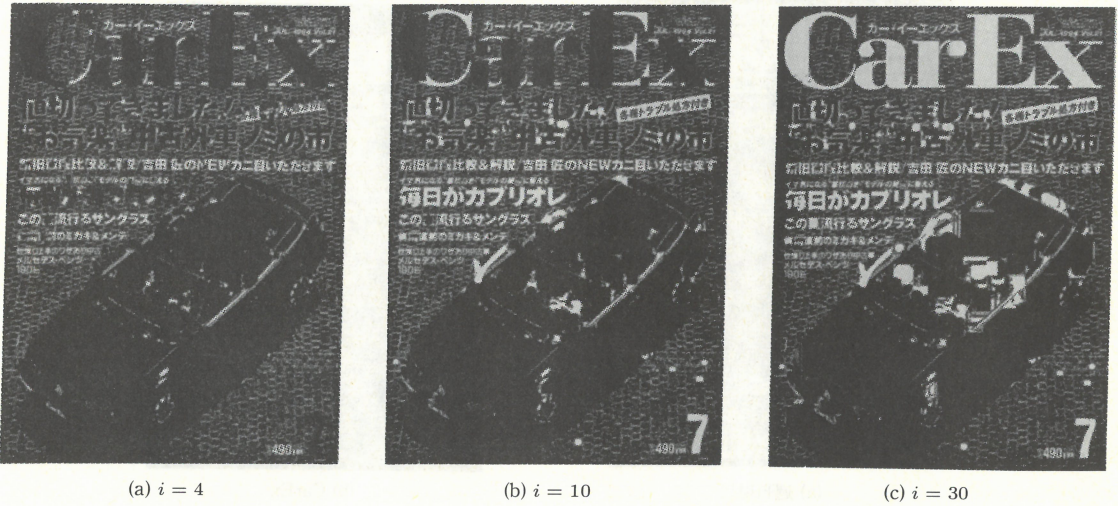


図4 分割された部分画像の例
Fig. 4 Examples of discomposed subimages.



図5 抽出および整形された部分画像の例
Fig. 5 Examples of extracted and refined subimages.

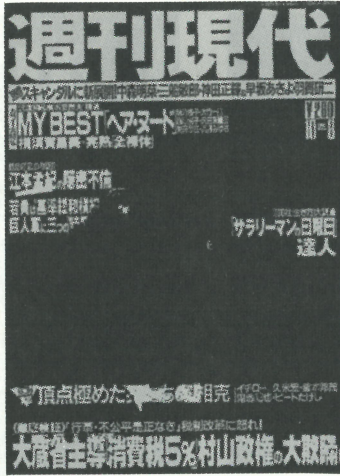
$$c = R, G, B, \quad k \in c$$

- $$\left\{ \begin{array}{l} X_{c0} : R, G, B \text{ 原画像の集合;} \\ k^c : k \text{ の補集合;} \\ |B| : \text{文字領域の階調により 2 値化処理.} \end{array} \right.$$

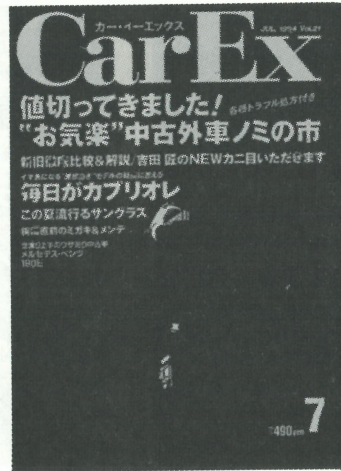
R, G, Bをそれぞれ独立に処理し最終結果の論理和をとることにより1枚の結果画像を得る。

5. 結果および考察

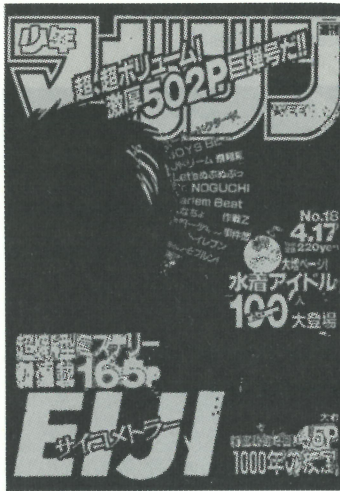
雑誌の表紙に着目して、濃淡画像とカラー画像の2種類の対象画像に対する文字領域抽出実験を行った。濃淡画像に対する実験については、3種類の異なる情景をもつ対象画像を利用した。それは絵と文字が相当複雑に入り込んだもの「少年マガジン(濃淡)」(図3(c))、文書画像に似た比較的平たんな背景部と文



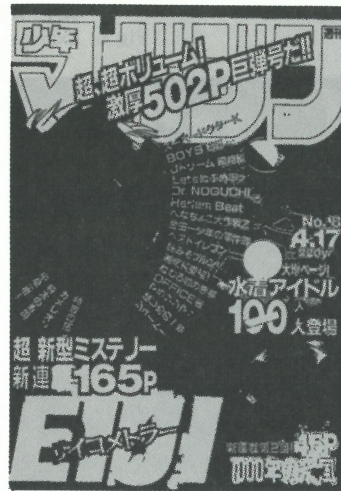
(a) 週刊現代



(b) CarEx



(c) 少年マガジン (濃淡)



(d) 少年マガジン (カラー)

図 6 結果画像の例

Fig. 6 Examples of result images.

表 1 文字領域抽出結果

Table 1 The results of characters extraction.

原画像	種類	総文字数	小さすぎ	コントラスト不足	抽出成功文字数	成功率	誤抽出領域数	誤抽出率
週刊現代	濃淡	320	20	0	300	94%	11	3%
CarEx	濃淡	230	50	0	180	78%	5	2%
少年マガジン	濃淡	303	29	62	212	70%	13	4%
少年マガジン	カラー	303	29		274	90%	12	4%
少年ジャンプ	カラー	267	10		257	96%	9	3%
スピリッツ	カラー	243	8		235	97%	8	3%
スパリオール	カラー	231	15		216	94%	4	2%

字部から構成されているもの「週刊現代」(図 3(a)), およびこの二つのちょうど中間要素をもつ画像「CarEx」

(図 3(b)) である。実験結果を図 6 および表 1 に示す。「週刊現代」のような単純な背景部をもつものでは

94%の抽出率となり、中間的な「CarEx」では78%となった。この両者とも抽出ができなかった文字は後記する「小さすぎ」る文字であり、これを除外すると100%の抽出に成功したことになる。複雑な背景をもつ「少年マガジン(濃淡)」では70%の抽出率となり、「小さすぎ」る文字を除外すると77%の抽出率となる。「少年マガジン(濃淡)」の場合、原画像では色コントラストは十分にあるが、濃淡画像に変換したときに、十分にコントラストが得られない部分が生じ、抽出不能となった文字が62文字であった。なお、このコントラスト不足の判定は目視により行った。このコントラスト不足を解消するには、減色処理をうまく行うことでも対処し得ると考えられるが、ここでは、カラー画像を用いて実験を行ってみた。その結果、「小さすぎ」る文字は依然抽出されないものの、コントラスト不足の文字は1文字も見られず、平均95%の抽出率を得ることができた。その一例として図3(d)に示した「少年マガジン(カラー)」は図6(d)のような結果となった。また、すべての画像に共通する「小さすぎ」る文字は、文字サイズが最小の構造化要素と同じか更に小さいものである。この場合、原理的に抽出は不可能になってしまう。これは画像解像度を上げれば抽出可能であると推定される。この特定の領域の文字を除けば100%の文字が正しく抽出されていることになる。

抽出された文字の周辺の若干の雑音と、文字の切れや欠けの修正などがあり今後の課題として残されている。雑音については、文字の特徴をもち、文字候補領域から除去されなかった“誤抽出領域”と考えられる。これを定量的に評価することは困難であるが、本研究では、“誤抽出領域数”として総文字数に対する割合を“誤抽出率”と定義した。この考えに従って表1の右2欄に実験結果を示した。

また、本方式は雑誌の表紙画像に特化して開発を進めたものではなく、十分に汎用性を考慮しており、情景画像一般に適用可能であると考えられる。更に、多くの他のカテゴリーの情景画像を用いた文字領域抽出実験を進めたいと考えている。

6. む す び

本論文では、モルフォロジーを用いた表紙画像からの文字領域抽出手法を提案した。表紙画像中の文字領域は、“背景部より浮き出している、あるいは沈み込んだ”細長い平たんなテクスチャをもった領域という特徴をもつ。この特徴に着目し「細長い領域」を検出す

ることに基づく文字領域方式を提案した。その一例として雑誌の表紙画像を対象として文字領域抽出実験を行った。その結果、複雑な構造をもつ画像にも適用可能でありかなり良好な抽出結果を得ることができた。

今後の課題としては、複雑な画像の一部で抽出が不可能となったものの改善を図る必要がある。また、文字領域の周りに残された雑音部分の除去、雑誌の表紙以外の他の情景画像への適用を考えたい。

謝辞 この研究の一部は、電気通信普及財団の研究助成により行われた。

文 献

- [1] 上羽 葵, 武田哲也, 岡田至弘, “等色線処理によるカラー画像からの文字領域の抽出”, 信学技報, PRU94-28, Sept. 1994.
- [2] 仙田修司, 美濃専彦, 池田克夫, “文字列の単色性に着目したカラー画像の文字パターン抽出法”, 信学技報, PRU94-29, Sept. 1994.
- [3] R.M. Haralick, S.R. Sternberg, and X. Zhang, “Image analysis using mathematical morphology,” IEEE Trans. Pattern Anal. & Machine Intell., vol.9, no.4, pp.532-550, July 1987.
- [4] I. Pitas and A.N. Venetsanopoulos, “Mathematical shape decomposition,” IEEE Trans. Pattern Anal. & Machine Intell., vol.12, no.1, pp.38-45, Jan. 1990.
- [5] X. Zhang and R.M. Haralick, “Morphological structure element decomposition,” Computer Vision, Graphics, Image Processing, vol.35, pp.370-382, Sept. 1986.
- [6] H. Park and R.T. Chin, “Decomposition of arbitrarily shaped morphological structure elements,” IEEE Trans. Pattern Anal. & Machine Intell., vol.17, no.1, pp.2-15, Jan. 1995.
- [7] 出口光一郎, 篠沢恭子, 森下 巖, “Morphologyによる図形分解における構造化要素の評価法,” 計測自動制御学会論文集, vol.30, no.2, pp.240-242, 1994.
- [8] S. Liang and M. Ahmadi, “A morphological approach to text string extraction from regular periodic overlapping text/background images,” CVGIP, vol.56, no.5, pp.402-413, Sept. 1994.
- [9] M. Okamoto and H. Hayashi, “Character extraction from headlines with background patterns by using shrinking/expanding methods,” IEICE Technical Report PRU90-151, pp.47-54, 1991.
- [10] R.M. Haralick and L.G. Shapiro, “Computer and Robot Vision,” Addison-Wesley Publishing, 1992.
- [11] 顧 力翔, 田中直樹, R.M. Haralick, “情景画像からの文字領域抽出に関する研究”, 信学技報, PRU95-221, March 1996.

(平成8年8月9日受付, 9年2月27日再受付)



顧 力 翔 (学生員)

昭 64 中国龍商院・工卒。平 8 神戸商船大・工・情報修士課程了。現在、豊橋技術科学大学電子情報工学博士課程に在学中。パターン認識・画像処理に関する研究に従事。



田中 直樹 (正員)

昭 56 阪大・工・通信卒。昭 61 同大学院博士後期課程了。工博。同年近畿大助手。平 2 神戸商船大助教授。平 4~5 ワシントン大客員助教授。パターン認識・画像処理などの研究に従事。



金子 豊久 (正員)

昭 37 東大・工卒。昭 39 同大学院修士課程了。昭 45 年 1 月米国プリンストン大学工学部電気工学科博士課程了・Ph.D 修得。同年 2 月米国 IBM ワトソン研究所に入社。画像処理の研究に従事。昭 57 日本アイ・ビー・エム東京基礎研究所に移り、東京サイエンティフィックセンター長など歴任後、平 5 年 4 月より豊橋技術科学大学情報工学系・教授。専門は画像処理、コンピュータグラフィックス、仮想現実感など。



R.M. Haralick

1964 カンザス大・数学卒。1966 同大電気卒。1967 同大学院修士課程了。1969 同大学院博士課程了。Ph.D. 取得。同年同大の教官となる。1975 同大教授。1979 バージニア工科大学教授。1984~1986 マシンビジョン国際研究所副所長。現在ワシントン大学教授。コンピュータビジョンの分野に多数の貢献をもつ。IEEE Fellow.