

カラー文書画像処理

長谷博行†, 米田政明†, チン・スエン‡

† 富山大学工学部知能情報工学科

〒930-8555 富山市五福 3190

‡ CENPARMI, コンコーディア大学

1455 de Maisonneuve Blvd. West, Montreal, Quebec, Canada H3G 1M8

E-mail: {hase,yoneda}@pat.iis.toyama-u.ac.jp, suen@cenparmi.concordia.ca

あらまし カラー文書画像から文字列を抽出・認識することを目的として処理する際に生じるさまざまな問題点と対策について述べる．まず，スキャナを用いてカラー画像を入力する場合に，解像度が画像処理に及ぼす影響について述べる．次に，色を利用した領域分割と文字列抽出の方法について述べる．最後にカラー文書画像中によく見られる変形文字列等の再配置と認識方法について述べ，いくつかの実験結果を示す．

キーワード : カラー文書画像, オフセット印刷, 文字列抽出, 文字認識

On a Processing of Color Document Images

Hiroyuki Hase †, Masaaki Yoneda †, Ching Y. Suen ‡

† Dept. of Intellectual Information Systems Engineering, Toyama University

3190 Gofuku, Toyama-shi, 930-8555 Japan

‡ Centre for Pattern Recognition and Machine Intelligence, Concordia University

1455 de Maisonneuve Blvd. West, Montreal, Quebec, Canada H3G 1M8

Abstract We discuss some issues on color document image processing for layout analysis and character recognition. First of all, we mention the problem that high resolution affects image segmentation using color information. Next, our color segmentation and character string extraction method are considered in comparison with the conventional techniques. Then we present a realignment model of rotated and/or distorted characters that we can often see on color documents. Lastly, we show some experimental results.

Keywords : color document image, offset printing, character string extraction, character recognition

1. はじめに

我々の身の回りにはカラー文書が溢れている。文書をカラー印刷する利点は言うまでもなく人の目を引くことと、忠実に色再現した画像が提供できることである。フルカラーメディアから文字列を抽出する研究は、メディアにより分類することができる。ひとつは情景中の文字列やテレビ画面中のフリップの抽出と認識[1,2]、二つ目は Web ページにおける画像中の文字列抽出[3]、そして最後にカラー印刷されたメディアからの文字列抽出・認識の研究[4]である。前二者は画像ソースがフルカラーであるのに対し、最後の印刷メディアの場合、ソース(印刷のインク)が限定された4色であるという点において異なる。本報告ではカラー印刷されたメディア(文書)に焦点を絞って議論する。

カラー文書を読み取る研究はこれまでにいくつか報告されている[例えば5,6]。この研究における主な課題は、画像入力法、文字列抽出法、文字認識法である。画像入力法では、文字領域を背景から明瞭に区別するために画像の濃度むらを如何に除去するか[7]、またカラー印刷の方式に起因して生じるノイズの除去や4色のインクで印刷されている文書から如何に見た目の色で入力するかを追求することである[8]。文字列抽出法では、カラー文書は文字色が不明であり、白黒文書のように必ずしも縦または横方向に文字が並んでいるとは限らないので文書中から如何に文字列(候補)を抽出するかが問題となる[5,9]。文字認識では、前述のように縦または横方向に並んでいない文字列を如何に認識するか[10,11]、またカラー文書によく見られる特殊なフォントに如何に対処するかが課題となっている。カラー画像処理の研究は大きく分けてこれら3つの要素からなっており、各要素において最良の方法を選択することが望ましい。

本報告では、筆者らがこれまでに行ってきたカラー文書画像処理方式の概要と今後の課題について述べる。

2. 画像入力

2.1 カラー印刷方式

フルカラー文書は4つのメッシュスクリーンにそれぞれブラック、マゼンタ、イエロー、シアンの4色のインクをのせ、順に印刷される。その際、スクリーンのメッシュの向きはそれぞれ決められた角度

に設定されている。印刷方法には大きく分けてオフセット印刷とグラビア印刷がある。オフセット印刷は高品質なカラー印刷法であり、スクリーンメッシュの大きさと密度が調整される。グラビア印刷はメッシュの密度のみが調整され高速印刷が可能であるが、見た目に粗く週刊誌のグラビア写真などの大量印刷に利用されている。本報告ではオフセット印刷されたカラー文書に限定して議論する。

2.2 高解像度カラー文書画像

通常のオフセットカラー印刷は175lpi(line per inch)の解像度を持つが、これは350dpiに相当する[12]。故に単一色(図1(a))を光学的に拡大すると(b)のような印刷パターンが現われ、これを白黒文書で通常用いられている400dpiの高解像度でスキャナから入力すると見た目には単一色となるが、それを1画素の色が見えるまで拡大すると図1(c)のようになる。

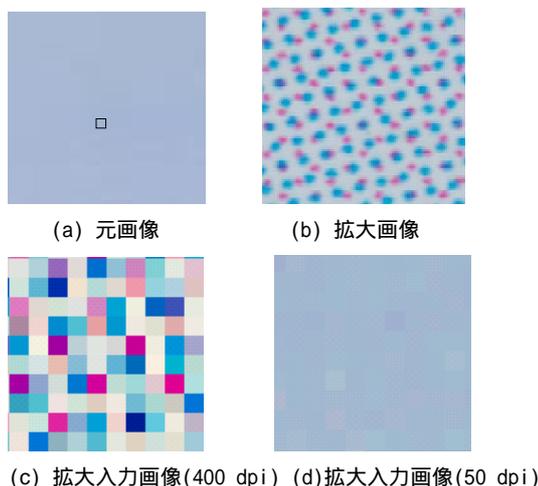
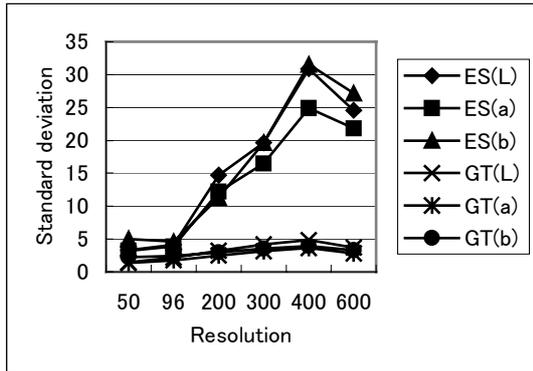
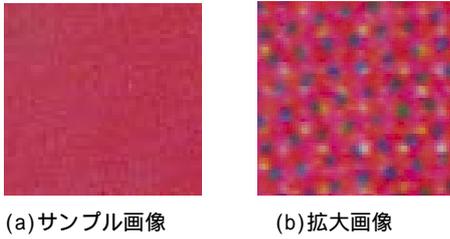


図1 異なった解像度で入力した画像の比較

ところが図1(a)の画像を50dpiの解像度で入力し、(c)と同様に拡大表示すると(d)のように各画素が同色に近くなる。なお、(c)と(d)では1画素の大きさは同じにしてある。見た目に同一色の印刷物(図2(a))を、解像度を変えて入力し各成分の標準偏差を見てみると、図2(c)のようなグラフとなる。図には6つのグラフが表されているが、それは $L^*a^*b^*$ 均等色空間を用いていることと[13]、2種類の光学解像度の異なるスキャナを用いているためである。図中、1600dpiの光学解像度をもつスキャナからのグラフをESと表し、600dpiのスキャナからのグラフをGTと表している。最近の市販スキャナの光学解像度は高く設定されており、高解像度で入力するとES

に近いばらつきが見られる。故に、画像処理する時、高解像度のカラー画像には何らかの処理を施さなければならないことがわかる。



(c) 解像度と標準偏差

図 2 解像度の違いが画像に及ぼす影響

そこで、すぐ考えられるのは平均化であるが、その効果を図 3 に示す。図 3 では見た目に単一色で異なる 4 色の L^* 成分について測定した結果が示してある。図では横軸の値 i に対して平均化対象領域は $(2i+1) \times (2i+1)$ である。結果から $i=1$ 、すなわち 3×3 の領域で平均化すると標準偏差が $1/4$ 以下になることがわかる。しかし、これだけではまだ不十分であり平均化はボケを伴うので筆者らは選択的局所色平均化法 (SLCA と略す) を考案した [8]。そのアルゴリズムを次に示す。

まず、元画像を f とし、 3×3 の領域で平均化さ

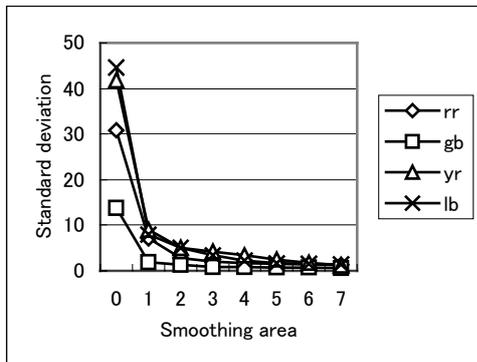


図 3 平均化の効果(ES)

れた画像を f_s とする。次に、 $t=0$ とし f_s を更新画像 $f^{(t)}$ にコピーする。そのとき、SLCA アルゴリズムは次のようである。

[Step 1] $f^{(t)}$ において注目画素を中心に $N \times N (= M)$ 局所領域を考える。それらの画素をグレイレベルに変換した後、1次元配列 $h(k)$ ($k=1, \dots, M$) に格納する。その際、座標も記憶しておく。1次元配列中の注目画素の要素番号を k^* とする。

[Step 2] グレイレベルの値で並べ替えを行う。

[Step 3] \bar{k} ($\bar{k}=1, \dots, M-1$) で配列要素を 2 分し、それぞれの領域 $\{1, \dots, \bar{k}\}$ と $\{\bar{k}+1, \dots, M\}$ において平均 μ_1, μ_2 と分散 σ_1^2, σ_2^2 を計算する。

[Step 4] 2つの領域の分離度を次式により計算する。

$$\rho(\bar{k}) = \frac{\mu_1 - \mu_2}{\sqrt{(k\sigma_1^2 + (M-k)\sigma_2^2) / M}}$$

その時、 $k_{\max} = \{\bar{k} \mid \max \rho(\bar{k})\}$ を得る。

[Step 5] もし、次の条件の内ひとつ以上を満たせば $f^{(t)}$ の $N \times N$ 局所領域の平均を $f^{(t+1)}$ の注目画素の値とし、Step 7 へ。そうでなければ Step 6 へ進む。

$$\rho_{\max} < \alpha$$

$$k^* = 1 \text{ and } k_{\max} = 1$$

$$k^* = M \text{ and } k_{\max} = M - 1,$$

なお、 α は事前に定められた定数である。

[Step 6] もし、 $k^* \leq k_{\max}$ であれば、 k ($k=1, \dots, k_{\max}$) が記憶している $f^{(t)}$ の画素の平均を $f^{(t+1)}$ の注目画素の値とする。そうでなければ、 k ($k=k_{\max}+1, \dots, M$) が記憶している $f^{(t)}$ の画素の平均を $f^{(t+1)}$ の注目画素の値とする。

[Step 7] 全ての画素に Step 1-6 を適用する。

[Step 8] もし、 $f^{(t+1)}$ が十分であれば停止する。それでなければ、 $t=t+1$ とし Step 1-8 を繰り返す。

- 終 -

このアルゴリズムを適用した例を図 4 に示す。(a) の元画像の一部を拡大したものが (b) であり、 3×3 の平均化を 1 度行った画像が (c) である。さらに

SLCA アルゴリズムを 1 回適用した画像が(d)であり、3 回適用した画像が(e)である。図より SLCA アルゴリズムを適用する回数を重ねるごとに平均化され、かつ境界領域が明瞭になっていく様子がわかる。



(a)元画像 (b)拡大画像 (c)平均化画像

(d)SLCA(1 回) (e)SLCA(3 回)

図 4 局所色平均化アルゴリズムの効果

また、いくつかのサンプルにおいて単一色と思われる文字領域と背景領域をそれぞれ指定し、本アルゴリズムを適用して得られた標準偏差を平均した結果を表 1 に示す。図 2 (c)と対比すると、SLCA を 3 回適用した結果においては約 50dpi で入力した画像レベルにまで均一化されていることがわかる。

表 1 SLCA の効果

	f	f_s	SLCA (1 回)	SLCA (3 回)
文字領域	22.36	5.90	4.08	3.44
背景領域	31.10	6.01	4.24	3.59

2.3 代表色選択

カラー画像の領域分割方法の一般的手法は代表色選択である。これは 1670 万色以上の色を数少ない色に置き換えることである。この方法は古くから種々の方法が考えられている[14,15]。PA 法、MCA 法[16]は一般画像から代表色を選択する方法であり、優位な色が代表色となる傾向がある。また、文字列抽出を目的とした色空間分割法も提案されている[17]。我々は $L^*a^*b^*$ 色空間を一辺 w の多数の立方体 (小領域) に分割し、局所色平均化された画像に対し画素の値をその色空間に写像してヒストグラムを

求めた。代表色は周囲の 26 近傍の度数に比べて大きな値を示す小領域の重心とした。この方法は計算量が少なく、クラスタリング等の他の方法と比較しても、性能的に遜色ないことが示されている[18]。また、不均一な背景から低品質文字パターンを効果的に抽出する手法として文献[19]が提案されている。

3. 文字列抽出

3.1 課題

通常白黒文書の場合には文字と非文字の違いは明白であり、文字列も横書きあるいは縦書きであるため抽出も容易である。カラー文書では文字列にフレーム、影、立体陰影などの飾りが施されていることがあり、文字列本体を抽出することは容易ではない。さらに、文字列は単一色のほかにグラデーションで描かれているものも時々見られる。色を利用する文字列抽出手法のほとんどの場合は文字列が見た目に単一色で描かれていることが前提になっている。そのため、グラデーション文字には対応できないが、アルゴリズムは簡単になる。逆に色を利用しない手法としてモルフォロジー演算により同じ線幅をもつ領域を抽出する手法が提案されている[20]。この方法は背景に制限を設けていないため、汎用的な手法であると考えられる。折衷的な手法として[4]があるが、分割した等色領域をクラスタリングすることによりグラデーションに対処している。しかし、カラー文書中の大部分の文字列は単一色で描かれており、これまでに提案されている手法は全て色を積極的に用いている[6,21]。筆者らが提案した手法[5]においても、ひとまとまりの文字列は見た目に単一色で描かれていると仮定している。

3.2 提案手法

文字列抽出法はこれまで数多く提案されている。我々も段階的弛緩法を提案した[22]。この方法自体はボトムアップ手法であり、他手法との能力比較において特に優れているとは言えない。しかし、ここでの目的は文字列候補を抽出することであり、筆者らの方法ではカラー文書画像から抽出される文字列候補の 90%以上は非文字列である。例えば、文字飾りや文字の穴などが連結したものも候補になり得る。そこで問題となるのは、如何に真の文字列をこれらの中から選択するかということである。筆者らは文献[5]でひとつの方法を提案している。

文字列としての評価はこれまで文字認識で得られ

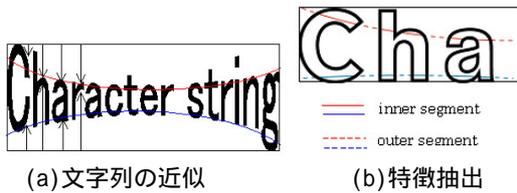
る評価値を用いることが多いが、ほとんど非文字列が候補となるカラー文書画像ではあまり効果がない。すなわち、文字飾りや陰影などは文字らしい形状をしており、また文字自体のフォントも特殊なことが多いため、文字認識の評価値では誤りやすいことが実験で明らかとなっている。そこでひとつのフィルタとして、文字列の並びを図5(a)に示すように2次元関数で近似し、図5(b)に示すように特徴を近似曲線が文字内と文字外を通る長さで与える。その特徴を使って、文字列らしさの評価を次式で定義した。

$$\eta = \frac{1}{1 + (\mathbf{x} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})}$$

ここに、特徴 $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6)^t$ であり、 x_1, x_2 は2つの近似曲線の文字外領域長さの標準偏差、 x_3, x_4 は2つの近似曲線の文字内領域長さの標準偏差、 x_5 は文字列矩形に対する黒画素数の比、 x_6 は線幅の標準偏差とした。

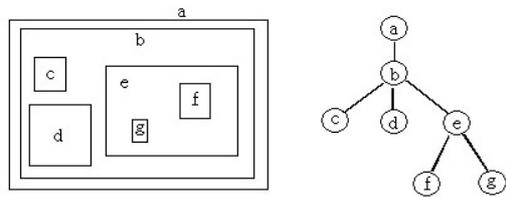
実験の結果、真の文字列のは大きな値をとり、比較的良好な結果が得られた。この方法の利点は文字列の並びにほぼ制限はなく、文字認識法は別として、取り扱える文字列の形状の自由度が高いことにある。さらに、文字の並び情報は文字認識の際にも利用できる。

次に文字列候補から真の文字列を選択する手

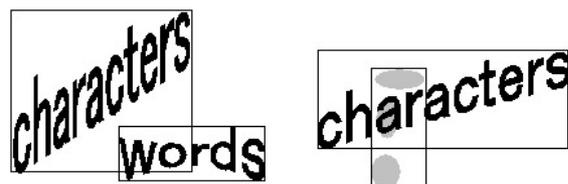


(a) 文字列の近似

(b) 特徴抽出



(c) 包含関係



(a) Case 1

(b) Case 2

(d) 重なり関係

図5 文字列抽出

法であるが、図5(c)(d)に示すように、隣接候補との包含関係と重なり関係を用いることにより解決した。すなわち、包含関係において、文字列である条件はお互いに空間的に離れている場合であり、かつ大きな評価値をもつ候補を文字列と考える。重なり関係では、図のCase1はどちらも文字列の可能性があるが、Case2ではどちらかが文字列ではない。この判定を文字列評価値 η の比較により行った。この文字列選択法は直線的でない文字列や特殊フォント、そして飾りのある文字列などにも対応できる点で効果的である。

4. 文字認識

4.1 課題

前述のように、カラー文書には時折傾斜した文字列や回転した文字列さらに歪んだ文字列などが見られる。これらの文字認識問題を扱った報告は筆者らのもの[11]のほか[10]などがある。両方法においては立場が異なっており、筆者らはカラー文書で比較的良好に見られる歪に注目しているのに対し、後者ではロボットの視覚を想定し、歪のない文字列が視点の位置が変わることにより歪んで見える文字形状に注目している。どちらの提案においても数学的モデルで歪補正の解決を試みている。しかし、対応できる歪には手法上の制限があり、またそれぞれにおいて副作用も生じている。今後、広範囲に対応できるモデルの提案が望まれる。

以下では、筆者らが提案する手法について述べる。

4.2 文字並びの補正

筆者らが対象とする変形文字列や文字歪を図6に示す。

図において(a)(b)は通常の文字列(タイプ0)、(c)

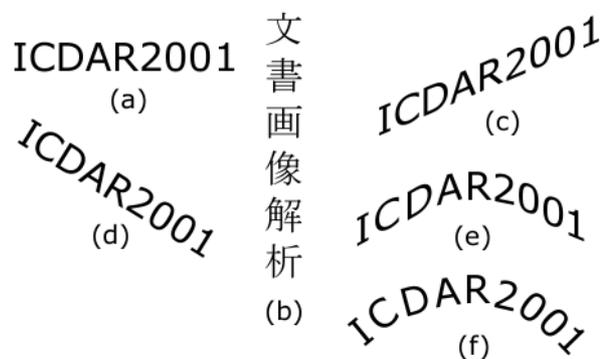


図6 対象文字列タイプ

は垂直方向にせん断歪がある文字列(タイプ 1), (d) は回転文字列(タイプ 2), (e)は文字位置によりせん断歪の大きさが異なる文字列(タイプ 3), (f)は曲線上に配置された文字列(タイプ 4)である. タイプ検出と文字形状補正は次のアルゴリズムで行った. 概略を示す.

[Step 1] 文字列並びに対し図 7 のように境界画素座標から 2 次関数近似を行う.

[Step 2] 近似関数により文字列が直線的であるか, 曲線的であるか判定する.

[Step 3] 2 つの射影パターンを作成する. ひとつは近似関数の接線に垂直方向に射影を行う. これを $h_n(s)$ とする. もうひとつは垂直方向に射影を行う. これを $h_v(s)$ とする (図 7).

[Step 4] それぞれの射影パターンにおいて $\eta = \sum_s (h(s) - h(s-1))^2$ を計算し, それらを η_n , η_v とする.

[Step 5] もし, 文字並びが水平あるいは垂直方向であればタイプ 0, 傾斜し直線的かつ $\eta_v > \eta_n$ であればタイプ 1, また $\eta_n > \eta_v$ であればタイプ 2 とする. また, 文字並びが曲線的であり $\eta_v > \eta_n$ であればタイプ 3 であり, $\eta_n > \eta_v$ であればタイプ 4 とする.

[Step 6] タイプの決定によりタイプ 1 と 3 の場合は文字の歪を補正し[23], 文字並びを再配置する. タイプ 2 とタイプ 4 の近似曲線に垂直な向きの場合には法線方向の角度を検出して文字を回転し再配置する. 隣接文字との間隔は元の文字間隔を利用する.

- 終 -

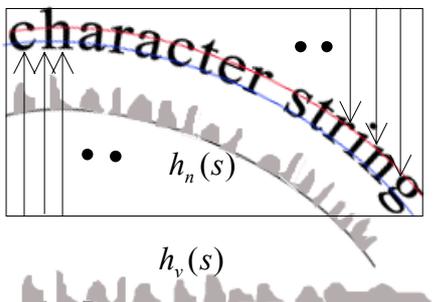


図 7 文字列近似と文字の向き判定

4.3 文字認識

本研究で用いた文字認識アルゴリズムは, 孫, 大町らが提案した方法である[24,25]. これは 64×64 の文字画像を細線化して 7×7 の小領域に分割し, 各

小領域で 4 方向に量子化した方向線素成分を計数し $196 (= 7 \times 7 \times 4)$ 次元特徴とする手法である. この方法では 3000 以上の英数字・記号・日本文字が対象であり, 通常の日本語フォントであれば 99% 以上の認識性能を示すことが確認されている.

5. 実験

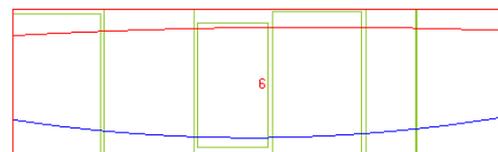
実験には実際の画像を用いた. このような歪文字列のデータ収集は難しく実験に用いた数はまだ少ない. また, モデルの制限により吸収できない歪みをもつ文字列は対象外としているので, 処理例は参考として示した.

図 8 に各タイプの処理例を示す. ここでは, 表示上の都合で文字列抽出が既に完了していると仮定するが, 前述の代表色選択, 文字列抽出アルゴリズムを適用してこれらの文字列が得られている. 各処理例では, 2 値化された画像を(a)に, 2 次関数近似の結果を(b)に, タイプ判定後文字歪を補正し文字を水平に並べ替えた画像を(c)に, 文字認識結果を(d)に示した.

まず, 例(A)はタイプ 0 の例であり, この場合は文字を並べ替える必要がないために(c)の画像がない. 文字認識は全て 1 位候補の文字を表示しているが, この例では正しく認識されている.(B)はタイプ 1 の例である. 近似曲線が直線的であり, 正しくタイプが判別されているが, 文字認識において隣接した 2 つの英文字が 1 つの日本文字に誤る現象が見られる.

手紙の書き方

(a) Original image (Type 0)



(b) Approximation by functions

手紙の書き方

(d) Recognition result

(A)タイプ 0 の処理例

No annual fee

(a) Original image(Type 1)



(b) Approximation by functions

No annual fee

(c) Realigned characters

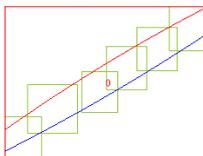
No a n n u a l 化 e

(d) Recognition result

(B)タイプ1の処理例

AWARDS

(a) Original image(Type 2)



(b) Approximation by functions

AWARDS

(c) Realigned characters

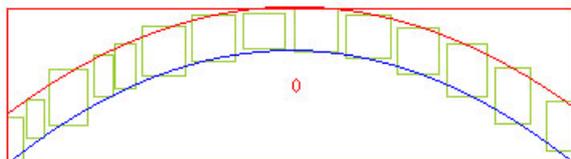
A W A R D S

(d) Recognition result

(C)タイプ2の処理例

20世紀最後の歳末大売出し

(a) Original Image(Type 3)



(b) Approximation by functions

20世紀最後の歳末大売出し

(c) Realigned characters

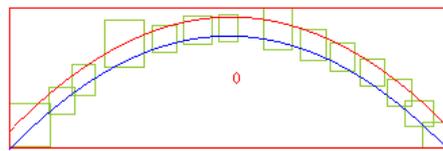
2。世紀最後の歳末大売出し

(d) Recognition result

(D)タイプ3の処理例

For Your Success

(a) Original image(Type 4)



(b) Approximation by functions

For Your Success

(c) Realigned characters

de or Y our S u c c e s s

(d) Recognition result

(E)タイプ4の処理例

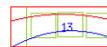
図8 各タイプの処理例

(C)はタイプ2と正しく判断されているが、文字フォントの影響で誤認識が生じている。(D)はタイプ3の例であるが、長い文字列では良い近似曲線が得られる傾向にある。文字の再配置では、多少のゆらぎが生じているがあまり文字認識には影響がないようである。(E)はタイプ4の例である。この例では良好な曲線近似に見えるが文字を再配置すると先頭文字の角度補正が不十分であることがわかる。そのため認識誤りが生じている。

次に、副作用が生じた例を示す。この手法は2次関数による近似がタイプを決めるため、短い文字列はタイプ判定を誤りやすい傾向にあると言える。図9(1)は左端の文字の影響でタイプ4と誤った例である。誤り例2は曲線近似に使う下側の文字境界座標の影響でタイプ3に誤った例である。今後は曲線近似のための境界画素の効果的選択により近似精度を上げる予定である。

Java

(a)



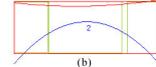
(b)

Java

(c)

BYTE

(a)



(b)

BYTÉ

(c)

(1)誤り例1

(2)誤り例2

図9 タイプ判定誤りの例

6. まとめ

筆者らのこれまでの研究を通してカラー文書画像処理の現状と問題点について述べた。筆者らが対象としたカラー文書は雑誌表紙や新聞折込チラシなどであるが、これらは写真など複雑な背景上に飾りのある文字列が多く、比較的難しい対象であると考えられる。カラー文書画像処理は、画像入力法、文字列抽出法、文字認識法の3つに分けることができるが、画像入力法と文字列抽出法に関しては、処理時間の問題は残っているが実用的な手法が提案されていると思われる。しかし、文字認識法では文字列の補正のために想定した文字列タイプはまだ不十分であり、今後さらなる検討が必要である。従来手法のように、早急な実用化を目指すには対象を限定するのむひとつの方法であろう。

また、今後の開かれた研究開発のためにはカラー文書画像のデータベースが必要となるであろう。しかし、カラー文書の情報量は膨大であり、A4サイズの文書を400dpiで入力すると、4670×3300画素となる。新聞折込チラシの場合ではA3サイズなどもあり、入力自体が困難となる。故に筆者らは、認識対象と考えられる文字列を含む部分画像のデータベースを現在構築している。

文 献

- [1] K.Matsuo, K.Ueda and M.Umeda: "Extraction of Character string region on signboard from scene image using adaptive threshold method," IEICE, Vol. J80-D, No.6, pp.1617-1626 (1997).
- [2] A. Jain and B. Yu, "Automatic text location in images and video frames," Pattern Recognition, Vol.31, No.12, pp.2055-2076 (1998).
- [3] Y. Yang and H. Zhang: "HTML page analysis based on video cues," ICDAR01, pp.859-864 (2001).
- [4] 上羽, 武田, 岡田: "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, PRU-94-28, pp.9-16 (1994).
- [5] H.Hase, T.Shinokawa, M.Yoneda and C. Y. Suen: "Character string extraction from color documents," Pattern Recognition, Vol.34, No.7, pp.1349-1365 (2001).
- [6] Y. Zhong, K. Karu and A. Jain, "Locating text in complex color images," Pattern Recognition 28, pp.1523-1535 (1995).
- [7] 後藤, 阿曾: "様々な画像に適用できる文字パターン抽出手法について - サーベイおよび構成法 -," 信学技報, PRMU99-234, pp.23-30 (2000).
- [8] 長谷, 米田, 酒井, 丸山, 加藤: "高解像度入力したカラー文書画像に現れる網点ノイズの効果的軽減手法," 映像情報メディア学会誌, Vol.54, No.7, pp.1107-1111 (2000).
- [9] T.Perroud, K. Sobottka, and H. Bunke: "Text extraction from color documents - clustering approaches in three and four dimensions," ICDAR01, pp.937-941 (2001).
- [10] 萩原, 渡辺, 河岡: "歪を有する活字文字列における形状補正方法," 信学技報, PRMU2001-190, pp.101-108 (2001).
- [11] H.Hase, M.Yoneda, J.Kato and C.Y.Suen: "Alignment of Free Layout Color Texts for Character Recognition," ICDAR2001, pp.932-936 (2001).
- [12] 一ノ瀬: "スキャナの最新技術," 情報処理, Vol.39, No.8, pp.769-775 (1998).
- [13] J.Tajima: "Uniform color scale applications to computer graphics," CVGIP21, pp.305-325 (1983).
- [14] E.Riseman and M.Arbib: "Segmentation of static scenes," Computer Graphics & Image Processing, Vol.6, pp.221-276 (1977).
- [15] 富永: "カラー画像の色分類と分割," 情報処理, Vol.31, No.11, pp.1589-1596 (1990).
- [16] P.Heckbert: "Color image quantization for frame buffer display," Computer Graphics, Vol.16, No.3, pp.297-307 (1982).
- [17] 池田, 田島: "カラー画像の限定色表現を利用したパターン切り出し方法," 春季信学全大, D-416, p.7-168 (1990).
- [18] 長谷, 米田, 酒井, 丸山: "カラー文書画像中の文字領域抽出を目的とした色分割代表色選択法," 信学論, Vol. J83-D-II, No.5, pp.1294-1304 (2000).
- [19] 後藤, 平山, 阿曾: "局所多値しきい値処理による濃淡文書画像からの文字パターンの抽出," 信学論 (D-II), Vol. J82-D-II, No.11, pp.2188-2192, pp.17-24 (1999).
- [20] 顧, 田中, 金子, R.M.Haralick: "表紙画像からの文字領域抽出方式," 信学論 D-II, Vol. J80-D-II, No.10, pp.2696-2704 (1997).
- [21] T.Perroud, K.Sobottka and H.Bunke: "Text extraction from color documents - clustering approaches in three and four dimensions -," Proc. of 6th ICDAR2001, pp.937-941 (2001).
- [22] 篠川, 長谷, 米田: "段階的弛緩法を用いた文字列抽出," 画像電子学会 Vol.29, No.5, pp.471-479 (2000).
- [23] A.Rosenfeld and A.C.Kak: "Digital picture processing," Academic press, (1978).
- [24] 大町, 阿曾: "品質を考慮した印刷文字の効率的认识法," 信学技報, PRMU92-38, pp.25-32 (1992).
- [25] 孫, 田原, 阿曾, 木村: "方向線素特徴量を用いた高精度文字認識," 信学論 D-II, Vol. J74-D-II, No.3, pp.330-339 (1991).