

カラー画像からの文字パターン抽出法

4L-1

仙田修司 美濃導彦 池田克夫

京都大学工学部

1 はじめに

カラー画像中に書かれた文字に対して文字認識を行うとき、文字認識処理では文字の色が黒、背景の色が白であると仮定しているため、そのままでは適用できない。そのため、カラー画像から文字領域と背景領域からなる文字パターンを抽出する手法が必要となる。

これまでに、カラー画像からの文字パターン抽出法として、文字の周囲のエッジを用いる手法 [1] や、15 枚の色・明度分解画像を用いる手法 [2] が報告されているが、これらの手法は主に文字と背景とのコントラストによるものであり、色情報を積極的に利用しているとはいえない。本稿では、画像内の単一文字列中の文字は同一色で書かれているという性質に着目し、カラー画像の色分類によって文字パターンを抽出する手法を提案する。本手法は、画像中に多く存在する主要な色を文字色候補とし、これを用いて、1 枚のカラー画像を複数の濃淡画像に変換するものである。

濃淡画像を出力する理由は、小さい文字パターンの認識には濃淡情報が有効となるからである。従来の文書画像処理では、濃淡画像を最初に2値化し、以降の処理は2値化された画像を対象に行っていた。しかし、文字認識処理における文字パターンの大きさの正規化のように、画像の拡大処理が必要となる場合には、濃淡情報を利用することによって高画質な補間が可能となる。本稿では、実際に文字認識率を比較することにより、文字認識における濃淡情報の効果を示す。

2 色分類によるカラー画像の濃淡化

カラー画像の色分類には、均等色空間においてクラスタを検出すればよい [3]。一般的なクラスタ検出手法として、K 平均アルゴリズム [4] などを挙げることができるが、こうした手法を画素を単位として適用するのは処理量の点から好ましくない。処理量削減のために、文献 [3] では、色分布の1次元ヒストグラムを分析することによりクラスタ検出を行っているが、ヒストグラムの分析には例外的な処理が多く、安定してクラ



図 1: 色分類によって得られた濃淡画像の例

スタを検出することが困難である。

本稿での色分類の目的は文字の色をクラスタとして検出することであり、文字を構成する画素は局所的に連結している。そこで、本稿では、局所的な色の同一性に基づく連結領域を生成し、これを初期的なクラスタとすることで処理量の削減を計る。ここで、ある点 p とその8近傍内の1点 q との色差 (RGB空間におけるユークリッド距離の2乗を用いている) が、ある閾値 α よりも小さければ p と q は連結していると定義する。この初期クラスタの生成処理は、画像の平滑化の効果もあるが、後のクラスタ併合の処理量削減が主要な目的であるので、 α は大きくなりすぎないように適当に決めればよい。

次に、クラスタの併合によって主要な色の抽出を行う。まず、全クラスタを画素数の多い順に並び換えたものを順に C_1, C_2, \dots, C_n とする。また、 $d(C_i, C_j)$ は、2つのクラスタ C_i と C_j の代表色(それぞれのクラスタ内の平均の色)の色差(ここでは CIE-L*a*b*表色系におけるユークリッド距離の2乗)を表す。 $i = 1, 2, \dots, n$ の順に、 $\min_{j < i} d(C_i, C_j)$ がある閾値 β よりも小さければ、 C_i を C_j に併合する。 $i = n$ まで行った後、クラスタの併合が起こらなくなるまで上記の手順を繰り返す。画素数によってクラスタを並び換えるのは、クラ

A Method of Extraction of Character Patterns from a Color Image

Shuji SENDA, Michihiko MINOH, Katsuo IKEDA
Faculty of Engineering, Kyoto University

スタの併合による代表色の変化を最小限にとどめ、安定したクラスタを得るためである。

上記のクラスタ併合アルゴリズムは、クラスタ数が単調に減少する点でK平均アルゴリズムよりは単純であるが、色差が最小のクラスタと併合していく点で単純クラスタリング [4] より複雑である。また、K平均アルゴリズムでは、あらかじめ最終クラスタ数 K を決定する必要があるが、文字色の抽出の目的には K を決めない方が都合がよい。

最終的に n 個のクラスタが得られた場合、それぞれのクラスタの代表色を用いて n 枚の濃淡画像を生成する。これらの濃淡画像における各画素の明るさは、対応する原画像上の画素の色と、代表色との色差（ここでは CIE- $L^*a^*b^*$ 表色系におけるユークリッド距離）として定義する。代表色として文字の色が選ばれている場合は、文字領域が黒、背景領域が白の濃淡画像が得られる。代表色として文字色が選ばれているのか、背景色が選ばれているのかの判断は、得られた濃淡画像（もしくはそれをさらに2値化した画像）上で文字らしさの幾何学的な評価（例えば [2]）を行えばよい。

3 色分類の実験および考察

図1は、大きさ 294×224 の 24bit/pixel のフルカラー画像（背景が白、上に赤い文字、下に黒い文字がある）と、それに前章の手法を適用して得られた3枚の濃淡画像である。初期クラスタの生成では、全部で 66150 画素の原画像に対して、5834 個の連結領域が生成された ($\alpha = 100$ とした)。このことから、局所的な処理による初期クラスタ生成によって、全画素を初期クラスタとする場合に比べてクラスタ数が大幅に減少していることが分かる。また、クラスタ併合では、3回の繰り返しで $n = 5834 \rightarrow 7 \rightarrow 3$ となって、最終クラスタ数 3 で収束した ($\beta = 1200$ とした)。この結果、ほぼ正確な色分類が行われているということが出来る。

Sparc Station IPX による実行時間は、初期クラスタの生成が約 2.4 秒、クラスタ併合が約 20 秒であった。

4 濃淡情報を用いた微細文字の認識

文字認識のように高い解像度を必要とする処理を行う際には、原画像中の濃淡情報を用いることによって精度が向上することが予想される。そこで、2値画像と濃淡画像を用いた場合の文字認識率の比較実験を行った (図2)。

JIS 第一水準中の漢字 2965 文字すべてについて、大

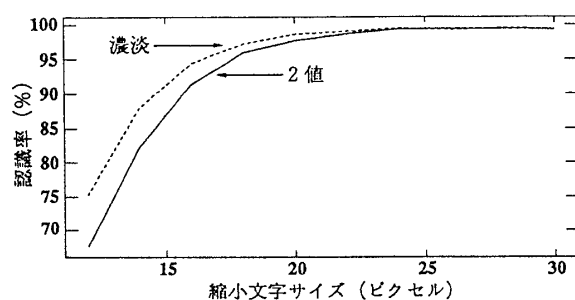


図 2: 2 値と濃淡の文字認識率の比較

きさ約 130×130 の 2 値画像を縮小して濃淡画像を得る。そして、これらを 2 値化してから拡大した場合と、濃淡のまま cubic convolution [4] により拡大してから 2 値化した場合の認識率の比較実験を行った。認識手法には、文献 [5] のものを用いた。その結果、およそ 25×25 以下の大きさの文字では、濃淡画像を用いた拡大の方が最大 7% 程度、認識率が良かった。しかし、濃淡画像の拡大手法として線形補間を用いた場合の認識率は、あらかじめ 2 値化してから拡大した場合とほとんど変わらなかったことから、認識率の向上は拡大手法に依存していると言える。このことから、よりよい拡大手法 (例えば [6]) を用いることで、さらに認識率の向上が期待できる。

5 おわりに

カラー画像を対象として、色分類による文字ボタン抽出法を提案した。本手法の特徴は、局所および大域的な色情報を用いていること、文字ボタンが濃淡画像として得られることである。また、微細な文字の認識に濃淡情報を利用することを提案し、実験によってその有用性を確かめた。

参考文献

- [1] 八木, 美濃, 池田: “文字写真混在画像からの文字エッジ点候補抽出法,” 第 41 回情処全大, 2-210, 1990.
- [2] 松尾, 梅田: “濃淡及び色情報による情景画像からの文字列抽出,” 信学技報, PRU92-121, 1992.
- [3] 富永: “カラー画像の色分類と分割,” 情処論, vol.31, no.11, pp. 1589-1598, 1990.
- [4] 長尾: “画像認識論,” コロナ社, 1983.
- [5] 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, vol.J74-D-II, no.3, pp. 330-339, 1991.
- [6] 新堀, 高木: “空間的高周波成分の外挿による高画質画像拡大,” 情処技報, IE92-38, 1992.