

## 情景画像からの文字抽出における 色分解画像組み合わせに関する考察

大原 宏太<sup>†</sup> 小嶋 和徳<sup>†</sup> 伊藤 義明<sup>†</sup> 石亀 昌明<sup>†</sup>  
岩手県立大学大学院ソフトウェア情報学研究科<sup>†</sup>

### 1. はじめに

OCR 技術による文書画像からの文字認識は現在では実用段階にある。しかし、情景画像からの文字認識は OCR 技術を用いても困難である。これは、情景画像が複雑な背景パターンを有しており、文字位置の特定が難しい為である。もし情景画像における文字認識精度が向上すれば様々な分野での活用が期待できる。本稿では、色情報のクラスタリング後に取得できる色分解画像を組み合わせることで、認識率の向上を図る。

### 2. これまでの手法

これまでの手法<sup>1)</sup>では、入力画像に対して Sobel フィルタによりエッジを抽出し、エッジの閉空間を利用して処理領域を抽出し、近接する領域の位置や大きさの情報を利用して処理領域を限定する。次に処理領域ごとに色情報を利用したクラスタリング処理により 2 値画像を生成する。その際に輝度ヒストグラムの頂点数を用いてクラスタ数を設定している。これにより領域中の色数を考慮が可能であると考えたためである。得られた 2 値画像から文字または背景の特徴量を抽出し、SVM で学習・評価することで入力画像に文字領域が存在するか識別を行う。

しかし、輝度ヒストグラムの頂点数を用いてクラスタ数を決定しても、文字領域の画素が集中しない場合が見られる。また、色の変化を考慮して抽出された領域を結合させず、1 文字程度の小領域での処理を行っていたが、小領域では背景との

誤認識をやや見受けられた。これらを受けて、本稿は次に示す手法を提案する。

### 3. 提案手法

#### 3.1. 色分解画像どうしの統合

クラスタリングを行う際にクラスタ数が適切に決定されない場合、文字領域の画素が 1 つのクラスタに集中せず、図 1 に示すような文字形状を崩した 2 値画像を得ることになる。これは SVM による認識及び学習に悪影響を及ぼすことになる。



図 1 クラスタリング結果の不良例

そこで、本稿では取得した色分解画像をそれぞれ組み合わせることで別々のクラスタに分類されてしまった文字領域の画素を 1 つのクラスタに集中させて、元の文字形状を維持した状態で特徴量を取得できるようにする。

#### 3.2. 領域の結合

これまでは光による色変化を考慮して、処理領域を結合せずに識別を行っていたが、小領域である分、背景領域と差が出づらいう事と「i」や「j」などの単調な形状の場合、識別が困難である問題が確認できた。そこで、本稿ではエッジによって得られた処理領域を近接するもの同士で結合して処理を進める。結合の方法については処理領域を限定する為に用いた以下の条件を利用して結合を行う。

1. 一方の処理領域の横幅を 2.7 倍してもう一方の処理領域と重なる領域が存在する
2. 2 つのブロックの比が 1.8 以下
3. 重なる領域の高さが 2 つの処理領域の外接最

A study about combination of color separation image for Letter Extraction from Scene Image

<sup>†</sup> Kouta O'Hara, Kazunori Kojima, Yoshiaki Itoh, Masaaki Ishigame. Iwate Prefectural University Graduate School of Software and Information Science.

小方形の高さの6割以上( $1 \geq 0.6c$ )

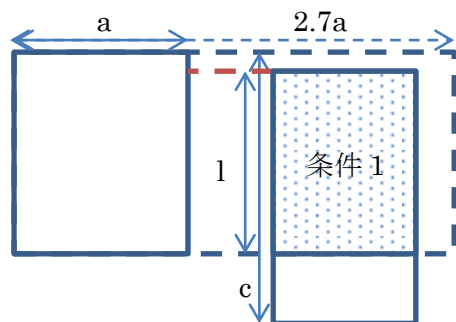


図2 結合条件

#### 4. 実験及び評価

##### 4.1. 評価方法

識別には LIBSVM を用いる。カーネルは RBF を使用し、 $\gamma = 0.6$ とした。本稿では SVM による識別で文字領域とされた領域に含まれる文字数によって評価を行う。

##### 4.2. 実験条件

学習及び評価データとしては Google 画像検索にて「park」「sign」のキーワードで得られた検索結果上位 500 中の 300 枚を利用し、学習データを 150 枚、評価データとして残りの 150 枚を利用して評価実験を行った。

##### 4.3. 実験結果・考察

表1 実験結果

手法	再現率[%]	適合率[%]	F 尺度
前手法 <sup>1)</sup>	59	49	53
本稿	71	31	41

表1が実験結果となる。再現率が7割程度であるのに対して適合率が3割程度と大きく差が開いてしまっている。前手法と比較すると再現率では10%程度の上昇が見られる一方で適合率では下がってしまっている。図3に出力結果の一例を示すが、文字列が良好に抽出できている一方で背景領域が文字領域として抽出されている領域が多く見られる。色分解画像を組み合わせたことによって求める文字領域の画像が得られることにより認識でも特徴量が良好に取得されている一方で、背景領域についても識別候補がより多く生成される為、文字領域と近似する特徴量を持った領

域を誤認識してしまう場合が増え、適合率の低下に繋がったと考えられる。



図3 出力結果例1 (左:本稿 右:前)

また、図4のように上手く連結を行えない場合に単文字として独立してしまった文字領域が識別の結果、背景領域とされてしまうところも存在するため結合の条件を見直し、文字列を良好に形成することで精度の向上が図れると考える。



図4 出力結果例2 (左:本稿 右:前)

#### 5. おわりに

本稿では、クラスタリングによって得られる色分解画像を組み合わせることで文字形状を維持した画像の取得を図った。今後は誤認識された領域に対する処理について考慮する必要があると考えられる。

#### 参考文献

- 1) 大原宏太, 小嶋和徳, 伊藤義明, 石亀昌明, ”看板画像からの文字抽出における処理領域の限定と特徴量の補強に関する研究”, 情報処理学会第73回全国大会, pp.319-320, 2011.
- 2) “LIBSVM”, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/> (最終アクセス 2012/01/10)