

# 情景画像からの文字抽出における 色分解画像の複合を用いた候補画像生成と特徴量補強に関する研究

大原 宏太<sup>†</sup> 小嶋 和徳<sup>†</sup> 伊藤 慶明<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 技術を利用した文書画像中の文字認識は実用段階にある。しかし、情景画像からの文字認識を行う場合、画像中に含まれる複雑な背景パターンがノイズとなり文字位置の特定が困難となる為、文字認識が難しい。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新たに候補画像を生成するほか、特徴量を補強することで認識率の向上を図る。

## 2. 従来手法

芦田らの手法<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する。そして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。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、クラスタリングを行った際にクラスタ数の設定や環境光による影響によって文字領域の画素が集中しない場合が見られる。また、本研究ではこれまで環境光による色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<sup>2)</sup>が、小領域では背景との誤認識がみられ誤認識をやや見受けられた。これらを受けて、本稿で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や環境光の影響によっては、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 による学習及び認識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、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 つの画像中に集めることで、図 2 のように元の文字形状を維持し、それを用いて特徴量を取得できるようにする。



図 2 色分解画像の組み合わせ例

### 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、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。

A Research of generating candidate by combination of color separated image and consideration of feature 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.

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つのブロックの面積比が 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小方形の高さの 6 割以上( $l \geq 0.6c$ )

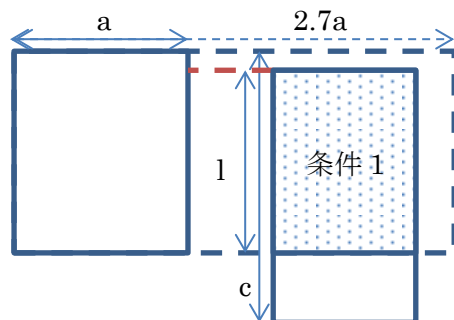


図3 結合条件

### 3.3. ストローク特徴による特徴量補強

文字としての特徴として、直線や曲線の集合体であることが考えられる。そこで、似た太さの線の長さを特徴量として抽出する。

画像中を縦横に走査し、線幅のヒストグラムを取得する。このヒストグラムから最も頻度の高かった幅の頻度値を特徴量とした。

## 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、 $\gamma = 6.0$ 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 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 中の文字が含まれる 300 枚を利用し、学習に 150 枚、評価に 150 枚としたものを用いる。

### 4.3. 実験結果・考察

実験結果を表 1 に示す。色分解画像の複合のみを行った場合には再現率で 8% 上昇した。これは色分解画像において文字形状を維持することができ、良好な特徴抽出につながった為であると考えられる。一方で適合率が 5% 低下しているのは、候補数が増大し、誤認識が増えた為である。

表 1 色分解画像複合・領域結合の実験結果

|       | 再現率[%] | 適合率[%] | F 尺度 |
|-------|--------|--------|------|
| 岩渕    | 59     | 68     | 63   |
| 複合    | 67     | 63     | 64   |
| 結合    | 55     | 69     | 61   |
| 複合・結合 | 71     | 65     | 67   |
| 複・結・特 | 69     | 69     | 69   |

領域結合のみを行った場合には適合率の 1% の向上があるものの、再現率で 4% の低下がみられた。環境光の影響が強くなってしまいうまく文字形状を抽出できていないと考えられる。これら 2 つの手法を組み合わせた場合においては再現率で 12%、適合率で 3% の低下がみられた。領域が大きくなったものの、色分解画像の複合を併せたことによって環境光の影響を抑えることができたと考えられる。ストローク特徴を加えて実験した場合、特徴を加えなかった場合と比べて適合率が 4% 向上し、F 尺度では 2% の向上がみられた。これは、ストローク特徴によって文字領域と背景領域の差分化が行えていると考えられる。再現率が 2% 低下してしまっているのは、縦横の走査のみで曲線への考慮を行っていないのではないかと考えている。

## 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ほか、ストローク特徴を用いて文字領域判別の精度向上を図った。

## 参考文献

- 1) 芦田和毅, 永井弘樹, 岡本正行, 宮尾秀俊, 山本博章, “情景画像からの文字抽出”, 電子情報通信学会論文誌 D- II, Vol.JJ88-D- II, No9(2005), pp.1817-1824.
- 2) 岩渕慎也, 大原宏太, 小嶋和徳, 伊藤慶明, 石亀昌明, “看板画像からの文字領域判別に関する特徴量の検討”, 平成 22 年度 電気関係学会東北支部連合大会, p.114.