

情景画像からの文字抽出における高周波情報を用いた背景領域削減 Reduction of Background Area using High Frequency Information in Character Extraction from Scene Images

齋藤 武蔵[†] 小嶋 和徳[†] 陳 国躍[‡] 伊藤 慶明[†]

Takeshi Saito Kazunori Kojima Guoyue Chen Yoshiaki Ito

1. はじめに

現在、OCR 技術による文書画像からの文字認識は、実用化されているほど高精度な認識が可能である。それに対して、情景画像からの文字認識の場合、背景領域がノイズとなるため認識が困難とされている。そこで、情景画像中の文字領域を特定することで OCR での認識を可能にする研究が行われている。芦田らの研究[1]、大原らの研究[2]では、画像中のエッジ情報および色情報から、領域を色ごとに分割した 2 値化画像である色分解画像を生成し、その特徴量から文字・背景領域の識別を行うことで情景画像中の文字を抽出する。また、稲見らの研究[3]では、屋内の画像に限定しているが、エッジ抽出と周波数特徴を用いて情景画像から看板領域を抽出している。これらに対して、本研究ではこれまで高周波情報を用いた閾値処理によりエッジの弱い画像、ノイズの多い画像それぞれに適切な処理を施す手法を提案することで再現率の向上を行ってきた。本稿では、画像のエッジが高周波数成分を持つことに着目し、ウェーブレット変換画像における水平方向の高周波ヒストグラムから閾値を設定し、抽出された文字候補領域から背景らしい領域を削減する手法を提案する。

2. 従来手法

芦田ら[1]や大原ら[2]の研究では、はじめに入力画像のエッジ情報を用いて文字候補となる領域をブロック化する。抽出した文字候補領域から画像の色情報を用いて文字部分を黒とした 2 値化画像を生成する。最後に、得られた色分解画像から特徴量(相互相関、凹凸変化点等 6 種類)を取得し、SVM による文字・背景領域の識別を行う。従来手法では画像によっては、背景から

文字とは異なる多くの領域を文字候補として抽出している画像が存在した。図 1 にその一例を示す。背景領域から余分な文字候補領域が抽出されている。

そこで本研究では、文字候補の高周波情報に着目し、明らかに背景らしい領域を削減する手法を提案する。



図 1 従来手法の文字候補領域ブロック化画像

3. 高周波ヒストグラムによる候補領域削減処理

グレースケール画像に対してウェーブレット変換を行うと、水平、垂直、対角方向の高周波成分画像が得られる。水平方向画像の濃度ヒストグラムの一例を図 2 に示す。

図 2 は、文字領域画像 2 枚、背景領域 2 枚について、濃度値 40~100 の相対度数を表している。図 2 のようなヒストグラムにおいて、相対度数に大きな差異が見られた濃度値 45 以上の時の相対度数 0.005 以上、濃度値 60 以上の時の相対度数 0.003 以上の値が出力される回数をカウントし、そのカウント数が 3 以上になる場合を背景領域として削減する処理を追加した。また、候補領域の幅と高さの和が 40 以下の極端に小さい候補も文字ではないものと考えて領域削減を行った。

[†]岩手県立大学, Iwate Prefectural University

[‡]秋田県立大学, Akita Prefectural University

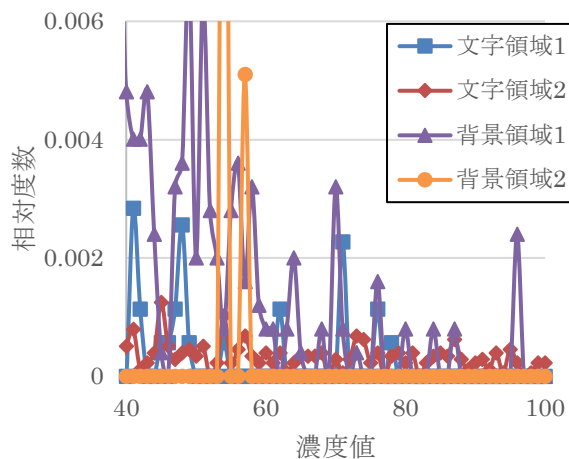


図 2 候補領域別の水平方向高周波濃度ヒストグラム例

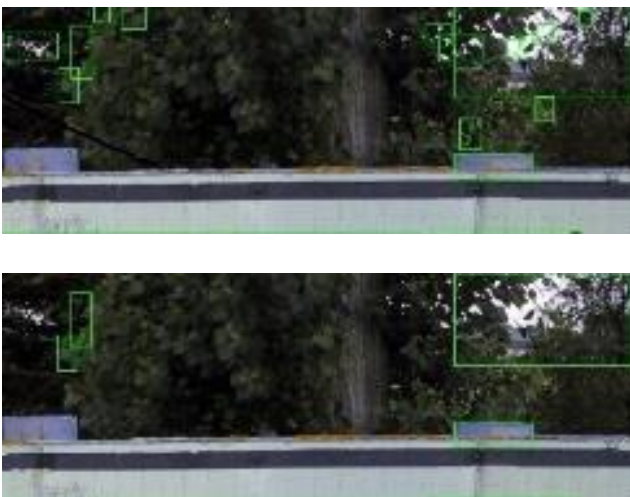


図 3 提案手法による背景領域削減画像例

4. 文字候補領域削減結果

提案手法による文字候補領域の削減結果を表 1 に示す。画像は、ICDAR2003 の画像セットの 200 枚を利用した。また、表 1 の文字抽出率は、

$$\frac{\text{文字候補領域に含まれる文字数}}{\text{画像中の全文字数}}$$

で得られる。

表 1 文字候補領域削減実験結果

	文字抽出率平均	候補領域数平均
従来手法	92%	59.4 個
提案手法	90%	29.5 個

表 1 から、提案手法により高い文字抽出率を維持したまま、平均で約 50%の候補領域を削減することが可

能となった。しかし、従来手法では抽出されていた文字領域部分を削減してしまう場合もあり、文字抽出率が 2%下がった。

5. 考察

提案手法により、画像の高周波情報の利用が文字候補領域からの背景らしい領域の削減に有効であることが確認できた。特に、文字候補数が 100 以上抽出され、そのほとんどが背景領域の画像においては、半分以上を削減することができ、誤抽出の抑制に期待できる成果が得られた。一方で、抽出されていた文字領域が削減されたものも存在したため、より多様な画像に対応した最適な閾値を設定する必要があると考えられる。加えて、削減後に残った候補領域の中に目視で明らかに背景らしい領域もあり、高周波以外の他の画像情報と組み合わせることで、さらに候補領域を絞り込んでいく手法も検討したい。また、従来手法で利用された特徴量に画像の周波数特徴も用いて SVM での学習・識別を行うことで文字候補からの文字領域・背景領域の識別精度も向上すると考えられる。

6. おわりに

提案手法を用いることにより、文字候補領域からの背景領域の削減に対して、高周波ヒストグラムを利用した閾値処理の有効性が確認できた。一方で、文字候補として抽出されている必要がある文字の抽出率が低下するという課題も見られた。今後はより最適な閾値処理の検討、特徴量の追加によりさらなる識別精度向上を図っていく。

参考文献

- [1] 芦田 他: “情景画像からの文字抽出”, 電子情報通信学会論文集 D-II, Vol.J88, No.9, pp.1817-1824, 2005.
- [2] 大原 他: “情景画像中の文字抽出における色分解画像の複合による候補画像生成と特徴量補強に関する研究”, 第 75 回情報処理学会全国大会講演論文集, 2T-3, pp.2-447-448 (2013)
- [3] 稲見他, “エッジ抽出と周波数特徴を応用した情景画像からの看板領域抽出の検討”, 日本知能情報ファジィ学会 合同シンポジウム 2010, 2010.