

モルフォロジー演算と顔検出を用いたマンガ内文字領域検出精度の評価 Evaluation on accuracy of text region detection in manga images using morphological operations and face detection

中川 諒[†] 梅澤 猛[‡] 大澤 範高[‡]
Ryo Nakagawa[†] Takeshi Umezawa[‡] Noritaka Osawa[‡]

1. はじめに

マンガの電子化において、テキスト情報を正確に抽出することができれば、他言語への自動翻訳やテキストによる検索など利便性が高い機能を実現できる。しかし、マンガ作品は、誌面構成の特性(コマ割り、フキダシ)や線画と文字が混在する様態のために、既存の文字認識手法をそのまま適用するだけでは十分な精度が得られない。このため、現状ではマンガ画像からテキスト情報を抽出するには手作業によらざるをえず、テキスト情報を電子化する障害となっている。そこで、本研究では文字認識を自動化することを目的とし、マンガ画像において文字が含まれている領域を検出する際の精度向上を目指す。

2. 関連研究

マンガ画像に対する文字領域検出では、検出領域と正解となる文字領域の対応が一意に定まらない場合が多く、精度の評価が難しい。また、正解の領域以外を検出した場合にはペナルティとして評価する必要がある。Wolfらは、情景画像を対象として、このような複雑な評価が必要になる場合に適切な評価を行うための手法を提案している[1]。マンガ画像に対する文字領域検出においても定量的な評価手法として利用することができる。

田中らは、マンガ画像中のフキダシを検出する過程で、Harr-like 特徴量による AdaBoost を用いて文字領域の検出を行っている[2]。フキダシの検出が目的であるため、文字位置検出はセリフ位置を特定する間接的な手段であり、文字領域検出の精度については検討されていない。また、モノログやオノマトペのようにフキダシ以外の文字領域は検出の対象となっていない。

荒巻らは、文字の連結成分による文字領域検出と領域に対する分類を統合した手法を提案している[3]。幾何特徴を用いて分類した連結成分を統合して領域を構成し、さらに機械学習を使って分類することで、従来手法を大きく上回る精度を達成している。しかし、未検出領域は少ないが、誤検出領域が多く、検出された領域には文字以外の領域も含まれているため、精度の高い文字領域検出に応用するには誤検出を抑えることが課題となる。

データセットとしては、プロの漫画家による作品 109 冊、21,142 ページから作成されたマンガ画像用のデータセット Manga109 [4] を使用する。



図1 文字と判断された連結成分例 ©赤松健

3. 連結成分の分類を用いた文字領域検出手法

連結成分を用いた文字領域検出における誤検出の傾向を調べるために、機械学習によって文字成分と非文字成分に分類する予備実験を行った。分類手順は次の通りである。

- (i). 入力画像を2値化し、黒画素を8近傍で連結する。
- (ii). 連結成分ごとに特徴量を抽出する。特徴量には荒巻ら[3]と同じく、連結成分の面積と周囲長、外接矩形の辺長、凸包の面積、近似楕円の長短径の長さとし、オイラー数の合計9次元ベクトルとする。
- (iii). ランダムフォレストによるモデル構築を行い、連結成分を分類する。

100 ページ分のマンガ画像から作成したすべての連結成分を対象に分類実験を行ったところ、適合率と再現率の調和平均である F 尺度は最大で 0.95 であった。しかし、図 1 に示す画像例のように、コマの背景部分に生じた小さな連結成分およびキャラクターの顔の一部を誤検出する事例が多くみられた。背景中の成分については、ふりがなや濁点などの小さな文字と特徴が似ているためだと考えられる。キャラクターの顔の一部は、周囲に白画素が多い点がフキダシ内の文字と共通していることや、サイズや形状が大きな文字と類似していることが原因であると考えられる。

4. マンガ表現の特徴を考慮した誤検出成分の軽減

本研究では、マンガの表現上の特徴を考慮した 2 つの処理を行うことで、誤検出が多発した要因を文字領域検出結果から除外する。

4.1 クローニング処理

小さな連結成分について、単純に大きさと判断して除外すると、ふりがなや濁点などの小さな文字成分も誤って除外してしまう恐れがある。そこで、モルフォロジー変換によって小さな連結成分と文字成分を区別する。本稿では、

[†] 千葉大学大学院融合理工学府数学情報科学専攻

[‡] 千葉大学大学院工学研究院

最も効果的であったクロージングについて述べる。クロージング処理は、膨張演算の後に同数の収縮演算を行うものであり、小さな穴埋めや切断部分の結合に有効である。文字は可読性を考えた配置のために周囲に他の連結成分が生じる割合が少ないのに対して、誤検出となる小さな連結成分は背景などの黒画素が密集している箇所が多い。したがって、連結成分が密集している箇所では、クロージング処理により周囲の成分同士が結合され、文字領域としての検出結果から除外可能であると期待できる。

4.2 顔検出処理

マンガ表現の特性上、キャラクターの顔に重なるように文字が配置されることは極めて少ない。そこで、顔検出を行うことで、キャラクターの顔部分を文字領域検出の対象から除外する。

4.3 実験手順

誤検出要因を除外するための処理は、それぞれ文字の連結成分を検出する前の 2 値画像に対して適用した。

クロージング処理については、膨張と収縮を 1 回ずつ実施した。顔検出については、HOG 特徴量によってマンガ画像の顔の特徴を SVM で学習させ、得られた検出器を用いた。その後、検出された顔を含む矩形領域内を白画素で塗りつぶすことで、連結成分を用いた文字領域検出の対象から除外した。学習データとしては、顔部分をすべて含むような矩形領域設定したマンガ画像を予め用意し、矩形領域の座標をデータとして与えた。

比較のため、2 つの処理を行わないもの、一方だけを適用したもの、両方を適用したものの計 4 条件で、ジャンルの異なるマンガ 5 作品に対して文字領域の検出を行い、Wolf らの手法[1]を用いて評価した。

4.4 実験結果

評価結果を表 1 に、文字領域検出の画像例を図 2 に示す。また、マンガ作品ごとの評価結果を図 3 に示す。

精度が最も向上したのは、クロージング処理と顔検出処理の両方を行った場合であった。作品によって効果に差があるが、実験に使用したすべての作品において処理を加えることによる適合率の向上が確認できた。なお、最も F 尺度の向上がみられたのは、作品『ラブひな』の 0.05 増加であった。

表 1 各手法における文字領域検出の平均精度

条件	適合率	再現率	F 尺度
処理なし	0.219	0.452	0.294
顔検出	0.226	0.450	0.300
クロージング	0.255	0.444	0.323
顔検出+クロージング	0.263	0.443	0.329

5. 考察

クロージングと顔検出により文字領域の検出精度は改善したが、誤検出箇所は残っており課題が残る結果となった。

顔検出では、複雑な形状の顔を正しく認識できていない例があった。また、前処理なしで正しく認識できていた文字領域の一部が、クロージング処理により認識できなくなる事例もみられた。

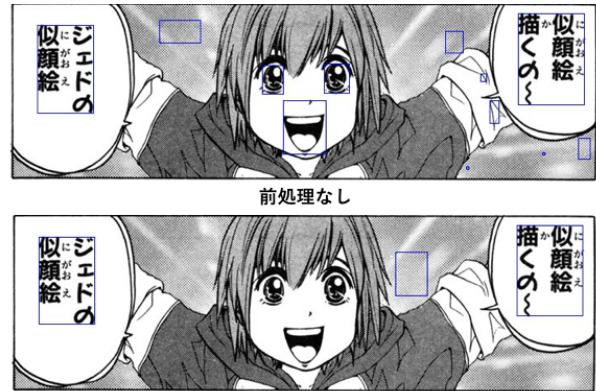


図 2 処理の有無による文字領域検出例 ©出口竜正

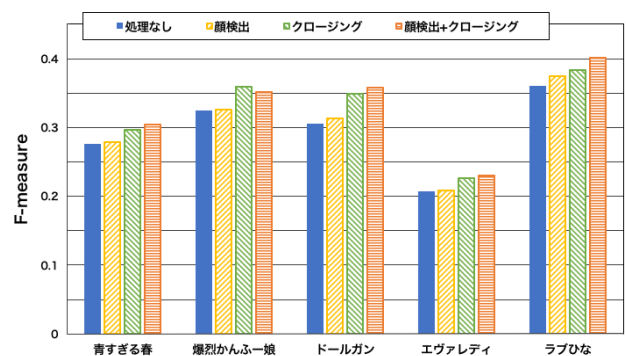


図 3 各マンガの文字領域検出の平均精度

クロージング処理による悪影響に関しては、クロージングを行った後の文字を学習データとして追加することで改善できると考える。顔検出の問題に関しては、顔画像のデータ量が十分でなかったことが一因であると考えられる。

6. おわりに

本稿では、文字の統合を行う前段階の処理によって精度改善を試みたが、その後の処理である文字の統合方法によって文字領域検出の精度に大きな影響があることが実験によりわかっている。そのため今後は文字領域の統合のためのグルーピング法についても検討していく必要がある。また近年画像処理の分野で目覚ましい成果を上げている畳み込みニューラルネットワークによる文字領域の認識やキャラクターの判別を検討する予定である。

参考文献

- [1] Christian Wolf, Jean-Michel Jolion, "Object count/area graphs for the evaluation of object detection and segmentation algorithms", International Journal on Document Analysis and Recognition, Vol.8, No.4, pp280-296 (2006).
- [2] 田中 孝昌, 外山 史, 宮道 壽一, 東海林 健二, "マンガ画像の吹き出し検出と分類", 映像情報メディア学会誌, Vol.64, No.12, pp.1933-1939, (2010).
- [3] 荒巻 裕治, 松井 勇祐, 山崎 俊彦, 相沢 清晴, "漫画における文字領域の検出と認識", 電子情報通信学会技術研究報告, Vol.115, No.495, pp.73-78, (2016).
- [4] Yusuke Matsui, Kota Ito, Yuji Aramaki, Toshihiko Yamasaki, Kiyoharu Aizawa, "Sketch-based Manga Retrieval using Manga109 Dataset", arXiv:1510.04389, pp.1-13, (2015).