H-027

線画の混在する画像におけるテキスト領域抽出の改善手法
An Improvement of Text Area Detection in Binary Images Containing Line Drawings

河村 圭　　石井 大祐　　渡辺 裕
Kei Kawamura　　Daisuke Ishii　　Hiroshi Watanabe

1. はじめに

テキスト検索やインデキシング技術の向上により電子化テキストの有用性が広く認識され，既存印刷文書（活字）の電子化技術の重要性が増している．とくに画像中に含まれる文字を認識することは必須である．しかしながら，文書画像を対象とした既存の文字認識手法では，線画や文字が混在する画像において線画を文字と誤判定する問題がある．

そこで我々は，従来から研究されている手書き文字列から文字を切り出す方式に着目した．これらの手法では，外接矩形の面積や縦横比などの形状的特徴を文字の判定に用いる．そして，線画と文字が混在する画像から文字領域を抽出する手法を提案してきた．

本稿では，あらかじめ文字サイズを指定することなく，文字切出しを実現する手法を提案する．画像内に複数の文字サイズが混在している場合でも，文字列内では同一サイズになるという傾向を利用する．なお文字サイズとは，文字のポイント数ではなく，文字の外接矩形における長辺の画素数と定義する．また，コンテンツとして需要の高いマンガを主な対象画像とする．

2. 従来手法の問題点

これまで多くの文字切出し手法，文字認識手法が提案されてきた．本稿では文字切出しのみに着目し，文字認識は対象外とする．ところで，文字認識結果や単語照合結果を文字切り出しにフィードバックする方式も検討されている．しかし，形状的な特徴に基づく切り出しがより正確になれば，フィードバック情報の確度が高まり，性能がより一層向上する．このため，文字の形状的特徴のみに基づく文字切り出しの高精度化は，重要な課題である．

文書画像を対象とする文字切出し手法においては，文字領域の抽出に先立ち行を抽出する手法がある．文字と思わしきもの（外接矩形）を文字行に垂直な直線に射影し，その分布密度を用いて行を抽出する．

後藤ら [1] は，罫線やノイズの混入を考慮しない単純な外接矩形が誤抽出の原因であることを示した．そこで，文書画像中の区分直線状の要素を文字行と仮定して抽出することにより，これを解決している．文書構造に関する知識を必要とせず，画像のゆがみにも耐性がある．本手法は必要なしきい値が多く，特に文字サイズに対するロバスト性が十分でない．また，罫線以外の要因に対する考慮が十分でなく，線画が混在すると誤検出が増加する．

線画の混在する画像として地図画像を対象とする文字切出し手法においては，文字枠図形をテンプレートとして用いる手法がある．テンプレートと原画像の適合度（黒画素密度）を抽出基準として文字を切出す．志久らは [2] は，地図画像から同一ポイント数の定型文字を切出すことを目的に，テンプレート数を 1 種類に削減する手法を提案している．背景との接触にも強いという特徴がある．本手法は，複数のポイント数が混在している画像から得られる結果の統合についての検討が不十分である．

3. 提案手法

3.1 文字切出し手法の概要

我々は以前より，文字の形状的特徴を用いる文字切出し手法を提案してきた [3]．まず，我々がこれまで提案してきた文字切出し手法の概要を示す．文字サイズ S を仮定して処理を進める．なお，文字サイズ S は等比数列的に大きくしていく．予備実験より，比率は 1.25 を採用する．

入力画像から，黒画素の連結成分を抽出して外接矩形を得る．矩形の長辺が S より長い場合，その矩形を破棄する．注目する矩形の周辺を探索して，x 軸，y 軸それぞれについて，矩形中心の位置と矩形の面積を射影してヒストグラムを生成する．ヒストグラムを平滑化してから中心付近に存在する山を検出し，文字列中心と文字中心を特定する．なお，探索範囲は文字列の垂直方向へ $S \times 1.3$，水平方向に $S \times 3$ とする．得られた文字列中心と文字中心を用いて，分離している矩形を統合して一文字に相当する矩形を得る．

3.2 文字サイズの自動決定手法

画像内においては様々な文字サイズが存在しうるが，文字列内では同一サイズになるという傾向がある．そこ

*早稲田大学大学院 国際情報通信研究科，
Graduate School of Global Information and Telecommunication Studies, Waseda University.

表 1 Recall rate of charactor segmentation from comic images.

Method	Recall rate [%]
Multi size with detection	86.2
Multi size without detection	19.1
Single size	87.0

で，文字を結合する前後にそれぞれ文字サイズを付与する手法を提案する．

先に結合後における文字サイズ付与を述べる．探索範囲に含まれる矩形のうち，射影後の文字列中心が十分近く，射影後の文字中心が十分離れ，かつ矩形長辺の長さが近い場合に，同一文字列を構成していると判断する．探索範囲の大きさはおおよそ 3 文字分であるから，行頭である場合を考慮して，文字数が 2，または 3 文字の時に正しく文字列が検出されていると判断する．これらの矩形に対して矩形の長辺を文字サイズとして与える．

結合前においては，探索範囲を文字列の垂直方向へ $S \times 0.2$，水平方向に $S \times 3.2$ として，注目する矩形の周辺を探索する．これは文字列のみを検出するためである．範囲内に存在する矩形について，矩形の長辺の中央値を文字サイズとして与える．これにより，文字サイズが決定されると，そのサイズが文字列に伝搬される．

4. 実験・考察

4.1 文字切出しの再現率

対象画像の代表例であるマンガ（B6 サイズ）を 300dpi でスキャンし入力画像とする．対象画像は 16 枚で，含まれる文字数は 1299 文字である．文字サイズを 1 種類のみに限定する場合，文字サイズの自動決定をする場合としない場合の 3 種類について比較を行う．実験結果を表 1 に示す．評価は一文字ごとに行い，ページ再現率（1 ページに含まれるすべての文字のうち，実際に検出された文字数）の平均値を示した．

まず，提案手法なしに文字サイズ S を変えると，再現率が著しく低下する．予備実験より，S が実際の文字サイズの 2 倍以上になった場合に顕著であった．複数の文字をまとめて一文字と判定してしまうことが原因である．

次に，提案手法を導入した場合，文字サイズ S を変えても再現率の低下は 1%以下となる．依然として複数の文字が結合されてしまう場合がある一方で，タイトルページ含まれるような大きな文字サイズにおいても検出が可能となった．また，入力解像度の異なる場合，すなわち異なる文字サイズの画像においても，同様の再現率が得られた．

ところで，一般的な文字切出し手法と比べて，提案手法の再現率は高くない [1]．一般的な文書画像と比べて，マンガのセリフには三点リーダー，感嘆符，疑問符などの記号が多く含まれていることに起因する．これらの記号を考慮した文字切出しは今後の課題である．

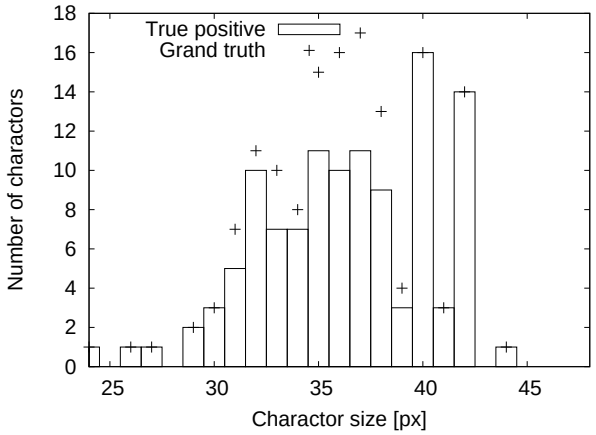

図 1 Charactor size and number of charactors.

4.2 文字サイズの分布

16 枚のうち 1 枚について，文字サイズと含まれる文字数（Grand truth），及び正しく検出された文字数（True positive）を図 1 に示す．なお，文字サイズと文字種の関係としては，32 画素付近がセリフにおけるひらがな，36 画素付近がセリフにおける漢字，40 画素付近が下書きに相当している．この図より，提案している文字切出し手法は様々な文字サイズに対応していることが確認できる．

5. おわりに

本稿では，線画の混在する画像から，様々な文字サイズに対応した文字切り出し手法を提案した．ある文字列に含まれる文字群は同一の文字サイズであるという仮定を用いて，文字サイズを自動的に決定している．実験により，提案手法では文字サイズをあらかじめ指定することなく，文字切出しを実現可能であることを示した．

参考文献

[1] 後藤ら，“文字行の局所的な直線性を利用した頑健・高速な文字行抽出法，” 信学論 D–II，vol. J78–D–II，no. 3，pp. 465–473，Mar. 1995.

[2] 志久ら，“地図からの文字切り出し，” 情処論，vol. 34，No. 2，pp. 273–280，Feb. 1993.

[3] 河村ら，“線画の混在する画像におけるテキスト領域の抽出に関する検討，” 信学総大，D-11-99，Mar. 2006.