

コミック画像の文字位置特定に関する検討

A Study on Text Localization in Comic Images

河村圭 石井大祐 渡辺裕
Kei Kawamura Daisuke Ishii Hiroshi Watanabe

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.

1 はじめに

画像からテキスト情報を抽出する技術により、テキスト検索の対象が印刷文書や書籍などへと広がっている。本稿では線画と文字が混在するコミック画像を対象として、文字位置を特定する (text localization) 手法を提案する。コミック画像には様々なサイズの文字や、線画や背景が混在しているという課題がある。

2 従来手法の問題点

これまで多くの文字位置特定手法が提案されてきた。文書画像を対象とする手法では、画像中の要素の大部分が文字であると仮定して最適化している場合が多い。その結果、線画の混在するコミック画像において、線画を全て文字として位置特定をする。特に、文字と線画をまとめて一つの文字として位置特定をした場合、文字認識処理の結果をフィードバックしにくいという問題がある。

線画の混在する画像として地図画像を対象とする文字位置特定の手法が志久らにより提案されている [1]。同一ポイント数の定型文字を切り出すことを目的に、テンプレートの種類を1種類に削減可能であること示した。背景との接触にも強いという特徴がある。テンプレートサイズのロバスト性が不十分で、誤検出が極めて多いという問題がある。

3 提案手法

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 [2]。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。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。画像内においては様々な文字サイズが存在するが、文字列内では同一サイズになるという傾向がある。そこで、文字を結合する前後にそれぞれ文字サイズを付与し、一度文字サイズが決定された矩形は、その後矩形の統合には利用しない。

得られた文字群から文字領域を生成する手法を検討する。文字領域を階層的クラスタリングの一つである単連

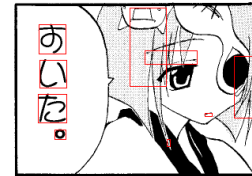


図1 文字位置特定の結果

結法によりまとめる。単連結法では、文字領域を統合するごとに距離の再計算を行う必要がない、結合後の距離は0とする。なお、クラスタリング終了はしきい値以下の距離がなくなった時とする。

4 実験・考察

実際のコミック画像に提案手法を適用した結果の一部を図1と、文字領域生成のクラスタリング終了時における文字間距離グラフの行列表現を表1示す。なお、行列中の値は距離をしきい値で正規化してある。

図より、文字以外の線画を文字として切出す場合も存在しているが、文字と線画が統合されることはない。また、表より1以下の距離は0となり、1, 2, 3, 5と4, 6がそれぞれひとまとまりの文字領域となる。

5 おわりに

本稿では、線画や文字が混在するコミック画像を対象として、様々な文字サイズに対応した文字位置特定手法と文字同士を統合する手法を提案した。実験により、提案手法では線画と統合されることなく文字位置を特定し、文字同士の統合が実現可能であることを示した。

参考文献

- [1] 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.
- [2] 河村ら, “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,” FIT2009, H-027, Sep. 2009.

表1 文字間距離グラフのクラスタリング結果

	1	2	3	4	5	6	7	8	9
1a	0.0	0.0	0.0	2.4	0.0	3.0	3.9	4.4	5.0
2a	0.0	0.0	0.0	2.5	0.0	3.0	3.4	4.1	4.9
3a	0.0	0.0	0.0	2.9	0.0	3.3	3.1	4.0	4.9
4b	2.4	2.5	2.9	0.0	3.1	0.0	2.5	2.3	2.6
5a	0.0	0.0	0.0	3.1	0.0	3.4	2.7	3.8	4.8
6b	3.0	3.0	3.3	0.0	3.4	0.0	2.2	1.7	2.0
7	3.9	3.4	3.1	2.5	2.7	2.2	0.0	1.2	2.4
8	4.4	4.1	4.0	2.3	3.8	1.7	1.2	0.0	1.2
9	5.0	4.9	4.9	2.6	4.8	2.0	2.4	1.2	0.0