

## 書棚画像の背表紙の文字情報を用いた書籍境界検出

小林 駿也<sup>†</sup>明星大学情報学部情報学科<sup>†</sup>丸山一貴<sup>‡</sup>明星大学情報学部情報学科<sup>‡</sup>

## 1 はじめに

図書館や書店では、書籍の場所や在庫の有無をデータベース上で管理している。このデータを利用した書籍検索システムによって利用者が特定の書籍を探す際、書籍が収納されている書棚まではたどり着くことができる。しかし、利用者が戻す位置を誤る場合があるため、書棚中での書籍の位置は特定できない。この問題は、書棚を撮影した画像から書籍を検索することができれば解決できる。書棚を撮影した画像から書籍単位に分割する手法が研究されている [1][2][3]。しかし、一度にすべての書籍を正確に分割することは難しい。例えば、双書のような似た背表紙を持つ書籍が並んでいる際に分割の失敗が起きやすい。そこで、双書において背表紙中に類似した文字列が、類似した座標に存在することに注目した。

本研究の目的は、光学文字認識 (Optical Character Recognition, 以下、「OCR」という。) を用いて文字の画像中の座標を含めた文字情報を取得し、レーベンシュタイン距離を用いて境界を検出する事である。

## 2 関連研究

石川ら [1] は文字が縦方向に連続して並んでいる領域を背表紙文字領域として扱った。本研究とは文字の座標を用いている点で一致している。しかし、本研究は文字列の類似性を判断する際に画像中の文字列座標を使用している点異なる。

Yang ら [3] はテキスト/非テキスト CNN モデルを適用して文字列の顕著性マップを生成して、書籍タイトルの場所を抽出する手法を使用した。生成した顕著性マップ上で、顕著性の値が大きい部分をタイトルとして扱い分割を行った。この研究は背表紙の文字列の顕著性に注目しており、本研究では OCR の処理結果である文字情報を利用している。顕著性を使用することで背表紙に文字がある限りその顕著性を取得することができるが、薄い書籍が並んでいる場合と背表紙文字が 2 列の書籍の識別が難しいという性質がある。それに対し、OCR の

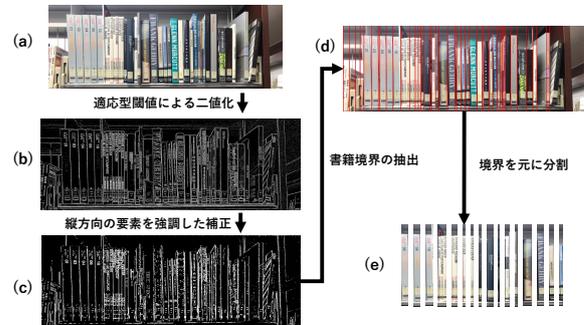


図1 書棚画像から分割までの処理の流れ

処理結果を使用する場合は文字列を取得することができれば書籍の厚さに影響されないが、背表紙の文字列が OCR によって取得できない書籍の識別が難しいという性質がある。

茨田ら [2] は書棚画像に対して適応型閾値を使用した二値化処理を行うことで書籍境界を抽出した。この研究は書籍同士の境界に色の差が生まれることを利用している。本研究では適応型閾値による二値化を使用し、二値化した画像に補正処理を加えて分割を行った上で、OCR の処理結果による再分割を行う。

## 3 提案手法

適応型閾値による二値化を行った結果に、補正処理を加えて書籍単位に分割する。分割後の書籍画像に対して OCR を使用し、その結果からさらに分割が可能であると判断した画像を再度分割する。システムには Java にて OpenCV を使用して作成した。本研究で使用する OCR は、Google が提供している Cloud Vision API である。この OCR では、単語単位の文字列と画像中の座標を取得できる。本研究では、入力画像は以下の条件を満たすことを前提としている。

- 書棚正面から撮影した画像であること
- 画像内の書籍が垂直に並んでいること
- 書棚画像が一段ごとの状態に切り出されていること

## 3.1 書棚画像からの分割

書棚画像から分割までの処理の流れを図1に示す。

入力した書棚画像に対して適応型閾値による二値化処理を行う。二値化処理を行った結果が図1(b)である。適応型閾値を用いた二値化処理の実装には OpenCV のメソッドを利用した。次に二値化画像に対して境界抽出

Book Spine Segmentation Based on Text Information by OCR

<sup>†</sup> Shunya Kobayashi, Department of Information Science, School of Information Science, Meisei University<sup>‡</sup> Kazutaka Maruyama, Department of Information Science, School of Information Science, Meisei University

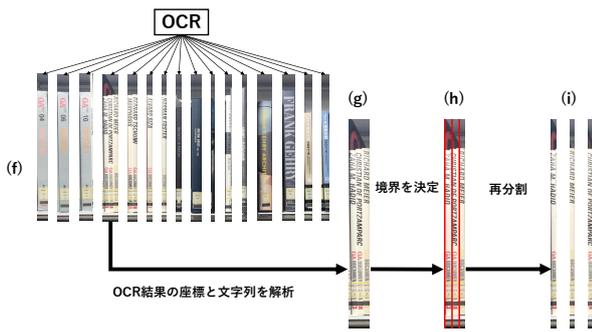


図2 再分割の処理の流れ

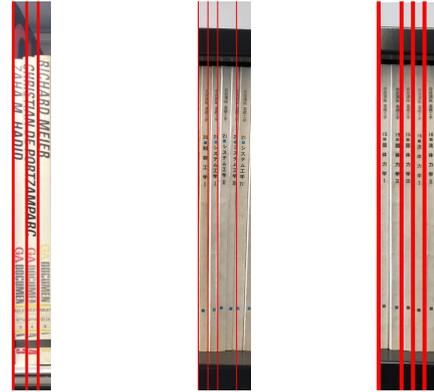


図3 再分割対象の例

の補助を目的とした補正を行う。補正にはノイズの除去と膨張処理，収縮処理を使用する。ノイズ除去は縦方向に伸びていない要素をノイズとして除去する。この処理を行うことにより縦方向の要素が強く残り，境界抽出の精度が上がる。その後，膨張処理を二回行った後に収縮処理を一回行う。これは境界要素の隙間を塞ぐことのほか，ノイズ除去によって掠れた境界線を修復することが目的である。補正した画像が図1(c)である。次に境界抽出を行う。二値化画像内で境界要素が縦方向へ十分に長く伸びている x 座標を境界座標とみなす。境界として認識した座標を赤い線で示した画像が図1(d)である。抽出した境界を元に書棚画像から書籍画像へ分割する。分割された書籍画像が図1(e)である。

### 3.2 再分割

再分割の処理の流れを図2に示す。書棚画像からの分割によって得られたそれぞれの書籍画像にOCRを使用する。図2(f)の状態に分割された書籍画像にOCRを使用して得られた文字列のレーベンシュタイン距離を，各書籍画像ごとに計測する。書籍画像中に，レーベンシュタイン距離と y 座標が近い，対の文字列が存在する画像を再分割対象の書籍画像とする。分割対象画像内の前述の条件に合致する文字列同士の x 座標の差を書籍一冊の厚さとして扱い，再分割を行う。しかし，ある文字列の対となる文字列が二つ以上あった場合は，右隣りにある文字列を隣接した書籍の背表紙として書籍の厚さを算出する。

## 4 実験

実験では，書棚画像から誤って1冊として分割された書籍画像を再分割した際の，正しく書籍ごとに分割することができた割合を算出した。正しい分割は，再分割された書籍画像を目視で確認し，タイトル文字列が欠けることなく1冊単位に分割できた書籍をカウントした。再分割する際の分割する座標を赤線で示した画像の例を図3に示す。本実験では12枚の書籍画像を再分割対象とした。

再分割した12枚の画像中で正しく分割できた書籍の

割合は，47冊中42冊で89.3%となった。5冊の書籍の分割に失敗した要因は，製本する際に片側に背表紙文字列がずれた書籍であったことと，OCRによって正しい書籍の背表紙文字を得ることができなかったためである。

## 5 まとめ

本研究ではOCRの結果を利用して複数書籍が含まれる書籍画像から，書籍ごとの境界を検出する手法を提案した。提案手法により，適応型閾値を用いた二値化処理を使用した手法で正しく分割できない場合があった双書の分割をすることができた。

本研究では入力画像の条件を指定していたが，二値化の前に回転等の補正を行う必要がある。[1],[3]の手法を組み合わせることで，背表紙で使用されている色が黒を基調として書籍が並んでいて，それらの書籍が同じ文字列を保有していなかった場合に再分割することができない場合や，背表紙文字の取得が難しい場合の改善が見込まれる。

## 参考文献

- [1] 石川幸博, 大倉充, 塩野充, 橋本禮治, 書棚画像からの書籍の背文字領域の抽出について. 岡山理科大学紀要. A, 自然科学, Vol.32, pp.163-171, 1996.
- [2] 茨田将史, 渡辺大地, 柿本正憲, 漫画本整理のための背表紙認識の研究, 情報処理学会研究報告, IPSJ SIG Technical Report, Vol.2015-CG-161 No.23, Vol.2015-CVIM-199, No.23, 2015.
- [3] X.Yang, D.He, W.Huang, Z.Zhou, A.Ororbis, D.Kifer, C.L.Giles, Smart library: Identifying books in a library using richly supervised deep scene text reading, CoRR, 2016, <<http://arxiv.org/abs/1611.07385>> (参照 2020-01-08).