

漫画本整理のための背表紙認識の研究

茨田将史^{†1} 渡辺大地^{†2} 柿本正憲^{†2}

概要: 近年、漫画本の販売数は上昇傾向にある。図書館で書籍整理を行うロボットの開発が盛んに行われてきたが、図書館等の限定した環境下でしか行えず、背表紙に認識用のマーカやバーコードを貼る等の手間がかかる。また漫画本の背表紙は決まったフォントがなく、OCR等での認識は難しい。本研究ではSIFT等の画像認識を用いて漫画本の背表紙を認識する。これにより認識用のマーカやバーコードを後付する必要をなくし、書籍整理機への応用を念頭に置いた認識を行うことでユーザの手間を削減した。

キーワード: 漫画本, 背表紙, 画像認識, SIFT, 書籍整理

Research of the spine cover recognition for comic book organizing

MASASHI BARADA^{†1} TAICHI WATANABE^{†2} MASANORI KAKIMOTO^{†2}

Abstract: In recent years, comic sales of book is on the rise. Development of robots that perform the books organizing at the library has been actively. It can be performed in limited the environment of libraries, etc. recognition paste a marker and bar code for troublesome the spine cover. Because there is no definite font in the spine cover of the comic book it is difficult to recognize in the OCR and the like. This study recognize the spine cover of a comic book using the image recognition such as SIFT. This eliminates the need for retrofitting a marker or bar code for recognition and reduced the labor of the user by performing the recognition considering the application of book sorters.

Keywords: comic book, spine cover, image recognition, SIFT, book organizing

1. はじめに

近年、漫画本の販売数は上昇傾向にある。オリコンの調査[1]では書籍全体の推測売上部数は11億7775万7千部(前年比97.3%)となり減少傾向にある。その中で、漫画本の推測売上部数は5億48万2千部(前年比101.3%)となり2年連続で上昇している。またネットオフが18~78歳の会員995名を対象に行った家の中にどれだけの書籍があるのかアンケート調査[2]を行ったところ、「書籍」は平均356冊、「コミック」は平均298冊だった。このことから個人でも大量に漫画本を所持していることがわかる。このような大量の書籍を整理するにはかなりの手間がかかる。

このような問題に対して、主に図書館の書籍整理について研究が行われている。今中ら[3]は書籍を取り出す際に人が行う動作を参考に、3本指を有するロボットアームを用いることで任意の把持力を実現することに成功した。

石田ら[4]はグリッパ付きアーム、カメラ、Laser Range Finderを搭載した移動車を用いて、傾いた書籍や倒れた書籍を把持することを可能にした。また、取り出し動作と収納動作を実現することにより書籍の並べ替えを行うことに成功した。

しかし、これらの研究は書籍の背表紙に認識用の自作マ

一カと自作バーコードを貼り付けるため、個人での運用を考えると手間がかかりすぎることが難点である。また、漫画本には決まったフォントがなく、OCR等での認識は難しい。

そこで、本研究ではSIFT特徴量とHSVヒストグラムを組み合わせた画像認識を用いて漫画本の背表紙の認識を可能にした。これにより、認識用のマーカやバーコードを背表紙に後付する必要をなくし、個人向けの書籍整理機への応用を念頭に置いた場合のユーザの手間を削減した。

2. 提案手法

2.1 前提条件

本研究では主に漫画本の整理について取り扱っていく。漫画本を選んだ最大の理由としては、日本で売られている書籍の半分近くが漫画本だという点が挙げられる。他にも、ページ数にバラツキが少ない、同じタイトルのものが何冊も出版される、背表紙の形式が同じ(タイトル・巻数・著者名の順に並ぶ)、バーコードの位置がほぼ一定などの一貫性があり、画像認識による整理に向いていることがあげられる。

前提条件として本棚の本は真っ直ぐに近い状態で立っているものとする。本棚の本はタイトルごとに巻数順でソートされており、1巻2巻4巻のように巻数の欠損があってもよいものとする。本棚内に入っている書籍は、一度はバーコードリーダーに通してデータベース化を行っているものとする。

^{†1} 東京工科大学大学院 バイオ・情報メディア研究科
メディアサイエンス専攻

Graduate School of Bionics, Computer and Media Sciences,
Tokyo University of Technology.

^{†2} 東京工科大学 メディア学部

School of Media Science, Tokyo University of Technology.

実際にどのような状況で使われるか想像しやすいように書籍整理の想定を示す。本棚は整理されている本が入る本棚部分と整理したい本を入れる書籍整理部分に分かれている。以下の図1が書籍整理の概要図である。

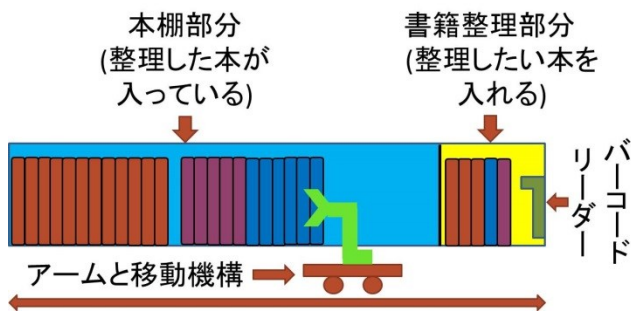


図1. 書籍整理の概要図

書籍の整理は書籍整理部分の本を本棚部分の同じタイトルの本の隣に移動し整理することを想定している。

2.2 書籍整理の流れの想定

最初に利用者が本棚から本を取り出し、読み終わったら書籍整理部分に本を入れる。

次に、書籍整理部分に備え付けてあるバーコードリーダーが書籍のバーコードを読み取り、ローカルのデータベース内に登録されている本か判断し、登録されていなければ本のバーコード情報と背表紙画像をデータベースへ登録し、登録されていればデータベースからバーコードの出版者情報を元に同じ著者の本があるか検索する。

データベースの情報から同じ著者の本が見つかった場合、同タイトルの並びが本棚の何処にあるのか画像認識を用いて探し、同タイトルの本がない場合は一番右に収納する。

画像認識で得た情報を元にアームが移動し、書籍整理部分から本を1冊取り出して本棚部分に移動する。以上を書籍整理部分の本がなくなるまで繰り返す。

本論文では、本の背表紙から本棚の何処に本を入れればいいのかを検索する部分を実装していく。

2.3 本棚の本の分割

元となる本棚の画像が以下の図2である。画像サイズは横482ピクセル縦221ピクセルである。



図2. 本棚の画像

書籍の認識手順は以下のようにした。図2の本棚部分の画像に対し、二値化を行う。同じような背景色の本の場合、図3のような決まった閾値で二値化を行うと本の境界線が

曖昧になりやすい。



図3. 二値化画像(閾値128)

そこで、二値化の閾値にはJainの適応型閾値[5]を用いた。適応型閾値とは各ピクセルの周囲の領域から重み付けの平均を計算し、閾値を決定することである。この手法は照明や反射等による輝度勾配等に強い特徴がある。この手法を使う利点は図4のように同じタイトルの本でも境界線が明確に出ることである。



図4. 適応型閾値による二値化

二値化画像から境界線の検出を行う。境界線の検出にはSuzukiらの境界抽出アルゴリズム[6]を利用する。図5に境界線を抽出した画像を示す。白くなっている部分が境界線である。

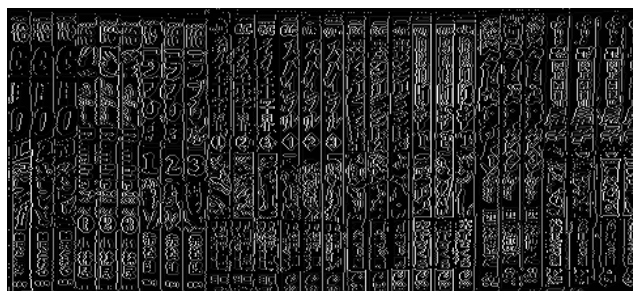


図5. 本棚の境界線を抽出した画像

境界線を元に縦方向に分割することで本1冊1冊を認識する。横方向に1ピクセルずつ縦方向にいくつ境界線の要素があるのかを計測し、一定間隔で要素が多い場所をピックアップする。間隔は写真撮影時にカメラと本棚の距離が一定なので、本1冊の横幅の大体の数値が想定できる。その本1冊想定値よりも少し狭い範囲を間隔とすることで分割線の数を増やし取りこぼしを軽減する。図6に分割線を引いた本棚の画像を示す。ピンク色の線が本の境界線を表す。



図 6. 分割線を引いた本棚画像

このままだと線が多すぎるため、先ほど検出した境界線間の距離の平均値を取得し、平均値の半分以下の値の場合、分割線を除外する。これにより多重に引かれている線を取り除くことができる。

図 7 に余分な分割線を除外し、本棚部分で本 1 冊 1 冊を認識した画像を示す。ピンクの縦線が本の境界線を示している。本 1 冊の幅は 12 ピクセルから 23 ピクセルの間となった。



図 7. 本棚の本を 1 冊単位に分割

2.4 SIFT 特徴量を用いた比較

SIFT を用いた理由は漫画本のタイトルや著者名の部分は同じフォントや背景であることが多く、SIFT による同一タイトル検出が期待できるためである。

書籍整理部分の本 1 冊に対し、本棚部分の本すべてと 1 冊ずつ総当たりで Lowe の SIFT 特徴量[7]を用い比較し、もっとも一致している可能性が高い本を特定する。比較には OpenCV の BruteForceMatcher を用いた。BruteForceMatcher は 1 つの特徴点に対して比較対象のすべての特徴点と総当たりで比較を行い、もっともユークリッド距離の小さい特徴点を見つけ出す手法である。特徴点の 1 つ 1 つのユークリッド距離の小さいものから順番に 20 個集め、ユークリッド距離を合計する。合計のユークリッド距離が小さいほど一致している可能性が高いことになる。

画像の範囲を狭め比較対象をタイトルのみにした場合、特徴点の検出数が著しく減少するという問題が生じた。そのため、SIFT の比較結果と HSV ヒストグラムの比較結果を乗算により組み合わせることで精度を上げる手法に切り替えた。

2.5 HSV モデルを用いた比較

SIFT 特徴量のみを用いた場合、同じタイトルの巻数違いでも似ていないと識別してしまうことがある。これは本の

背表紙が日焼けして特徴が変化してしまう等が原因として考えられる。そこで、H(色相)S(彩度)V(明度)モデルのヒストグラムを用い、色に対して比較を行うことでさらなる精度向上を図った。書籍整理部分の本 1 冊に対し、本棚部分の本すべてを総当たりで比較した。HSV のそれぞれの要素を 256 段階に分け、正規化を行った上で比較する。

比較にはバタチャリア距離を用いた。バタチャリア距離は 0.0 から 1.0 の間で表し、完全一致が 0.0、完全不一致が 1.0 となる。バタチャリア距離は HSV それぞれの要素の比較に用いた。これら 3 つのバタチャリア距離を合計し、SIFT の結果と乗算することでハイブリッドな計算結果を求めた。この乗算結果の値が低いほど類似性が高いと判断するものとした。

今回は漫画本の性質を利用して精度向上を図る。漫画本の巻数違いはタイトルの位置や色が同じことが多い。そこで、本を上下半分に分割し、上のタイトル部分だけを比較することにした。よりタイトルの要素を高めるために、横幅についてはそれぞれの本の幅の中央に 1 ピクセル、左右に 3 ピクセルの合計 7 ピクセルの範囲で画像を切り取り、比較を行った。画像サイズは横 7 ピクセル縦 110 ピクセルとなった。

3. 実験結果

3.1 SIFT 特徴量を用いた比較

「スパイラル」3 巻を注目画像とし、本棚のすべての本それぞれを比較画像とし、注目画像と各比較画像との SIFT 特徴量を用いた比較結果を図 8 にまとめた。「マッチング数」は元画像とどれだけマッチングした特徴点があるか、「ユークリッド距離」はどれだけ元の画像と似ているかをそれぞれ表す。ユークリッド距離が小さい方がより注目画像に近い。今回は一致率の高い特徴点 20 点のマッチングのユークリッド距離の合計を比較した。

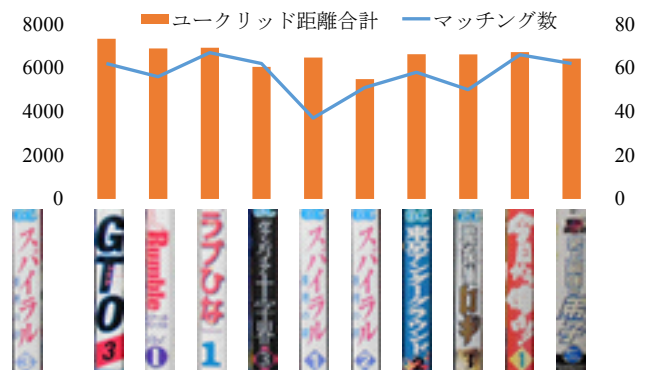


図 8. スパイラル”3 巻 (左下端) との比較

比較の結果、同タイトルの巻数違いである”スパイラル”2 巻のユークリッド距離の合計が一番小さいという結果になった。しかし”スパイラル”1 巻は 5 番目という結果になった。これは本の日焼けによる色の変色が原因だと考えられる。また 2 番目に位置する”ヴァンパイア十字界”は原作者

と著者が注目画像と同じであるため合計ユークリッド距離が小さいと考えられる。

上記と同じ条件でそれぞれのタイトルの1巻を注目画像とした実験を行った。結果は”東京アンダーグラウンド”と”魔探偵ロキ”と”今日から俺は!!”以外の本はすべて同タイトルの巻数違いの本との合計距離が小さく、似ていると識別できた。

“東京アンダーグラウンド”については巻数で背景色が違うため、SIFT 特徴量では違う特徴が検出されやすくなっているのではないかと推測する。

“魔探偵ロキ”は2巻が下から2番目の値となった。これはタイトルの文字の色や著者名の背景色、キャラクター絵等の違いがあり、似ている特徴が出にくいのではないかと推測する。

“今日から俺は!!”の3巻は2巻のユークリッド距離の合計と比べると三分の二程度でもっとも良い値になった。2巻は全体の半分よりも下位であった。原因としては、本の一番上にあるキャラクターの絵が違うことや2巻が著しく日焼けしていることが考えられる。

注目画像と同じタイトルの比較画像のユークリッド距離合計がもっとも小さい場合、認識したと定義する。SIFT 特徴量を用いた比較では最終的に全体の本の認識率は60個中37個で61.6%となった。少なくとも同じタイトルの本を1冊でも認識できる確率は28個中23個で82.1%となった。同じタイトルの本をすべて認識できる確率は28個中12個で42.8%であった。本1冊1冊の比較なのでタイトル以外の背景絵等の特徴を感知してしまい、その分認識率が低くなったと考えられる。

3.2 HSV モデルを用いた比較

色を考慮して精度を上げるために、2.5 で述べた手法に基づく実験を行った。ここでは、”スパイラル”3巻を注目画像とし、すべての本それぞれを比較画像とし、注目画像と各比較画像との HSV ヒストグラムによるそれぞれの要素の比較結果を得た。それらの結果のうち10件の比較画像の結果を図9にまとめた。結果の数値はHSVそれぞれのヒストグラムについて計算したバタチャリア距離の加算合計である。



図 9. “スパイラル”3巻との HSV モデルを用いた比較

本の上半分であるタイトル部分を比較したことにより、背景色が”スパイラル”と同じ白色のものが比較的良いスコアをマークした。今回の比較は注目画像と同じタイトルの巻数違いの2冊が1番目と2番目に似ていると認識された。

上記と同じ条件で9件のタイトルのそれぞれ1巻を注目画像として実験を行った。結果は”東京アンダーグラウンド”と”ラブひな”以外のタイトルについてすべて同タイトルの巻数違いの比較画像でバタチャリア距離が小さい値となり、似ていると識別できた。

“東京アンダーグラウンド”については巻数違いの2冊は「4番目と13番目」の順位となり、明らかに劣った結果となった。これは”東京アンダーグラウンド”の1巻だけ他の巻との背景色が異なるデザインであることが大きな原因だと考えられる。

”ラブひな”の注目画像については”スクールランブル”の比較画像と似ているという結果が出た。これは本の背景色が同じ白で、色相や明度が類似したヒストグラム分布を示したことが原因だと考えられる。

注目画像と同じタイトルの比較画像が、同じタイトル以外の比較画像と比べてときに値がもっとも低い場合、認識したと定義する。HSV ヒストグラムを用いた比較では最終的に全体の本の認識率は60個中41個で68.3%となった。少なくとも同じタイトルの本を1冊でも認識できる確率は28個中22個で78.5%となった。同じタイトルの本をすべて認識できる確率は28個中16個で57.1%となった。こちらはタイトル部分だけの比較だったので、SIFT 特徴量による比較よりも高い認識率となった。

3.3 SIFT と HSV ヒストグラムのハイブリッド比較

“スパイラル”第3巻を注目画像として本棚の本すべてとの比較を SIFT 特徴量と HSV ヒストグラムによるハイブリッド化を行った。それらの結果のうち、10件を図10にまとめた。

SIFT 特徴量から得たユークリッド距離の合計と HSV ヒストグラムから得たバタチャリア距離の合計を乗算することでハイブリッド化した認識結果を得た。



図 10. “スパイラル”3巻とのハイブリッド比較結果
 複数の比較結果を組み合わせることにより、注目画像と

同タイトルの別の巻数の比較画像が他のタイトルの比較画像よりも類似性が高い結果となった。これにより、同タイトルの本が本棚の何処にあるのかということ画像から認識できるようになった。このことから、SIFT 特徴量だけでは対応できなかった日焼けした本への対応ができたといえる。

それぞれのタイトルの1巻を注目画像として、本棚の本すべてとの比較を、SIFT 特徴量と HSV ヒストグラムによるハイブリット化手法によって行った。結果は”東京アンダーグラウンド”と”魔探偵ロキ”の本以外の本はすべて同タイトルの巻数違いの比較画像が高い類似性となり、似ていると識別できた。これにより SIFT や HSV ヒストグラム単体では認識できなかった本が認識できるようになった。

しかし、”東京アンダーグラウンド”や”魔探偵ロキ”といった元々の背景色や文字の色が著しく違うものは認識できないという結果になった。

各タイトルを各巻数別に注目画像としたとき、同タイトルの比較画像でもっとも類似性が高いものの順位を最高順位とし、もっとも類似性の低いものの順位を最低順位として以下の表1にまとめた。”史上最強の弟子”以外は各タイトル3巻ずつあるため、最高順位が1位かつ最低順位が2位となるのが最良の結果である。”史上最強の弟子”は4巻あるため、最高順位が1位かつ最低順位が3位となるのが最良の結果である。

表 1. 同タイトルの認識順位

タイトルおよび巻号		最高順位	最低順位
GTO	1	1	2
	2	1	2
	3	1	3
スクールランブル	1	1	2
	2	1	2
	3	1	2
ラブひな	1	1	2
	2	1	2
	3	1	2
ヴァンパイア十字界	1	1	2
	2	1	2
	3	1	2
スパイラル	1	1	2
	2	1	2
	3	1	2
東京アンダーグラウンド	1	3	4
	2	1	12
	3	1	15
魔探偵ロキ	1	2	10
	2	1	8

今日から俺は!!	3	4	6
	1	1	2
	2	1	2
史上最強の弟子	3	1	2
	1	1	3
	2	1	6
	3	1	3
	4	1	3

注目画像と同じタイトルの比較画像が、違うタイトルの比較画像のいずれよりも類似性が高い場合、認識したと定義する。ハイブリット比較では最終的に全体の本の認識率は60個中49個で81.6%となった。同じタイトルの本を1冊でも認識できる確率は28個中25個で89.2%となった。同じタイトルの本をすべて認識できる確率は28個中20個で71.4%となった。

SIFT や HSV ヒストグラムに比べて全体の本の認識率は20~30%高くなった。同じタイトルの本を1冊でも認識できる確率は約10%上がった。同じタイトルをすべて認識できる確率は15~30%高くなった。

同じ作品シリーズの中で背景色やタイトル色が例外的に異なるような巻号を取り除いておくと全体の認識率は95.8%、1冊でも認識できる確率は100%、同じタイトルをすべて認識できる確率は92%となった。このことから、背景やタイトルが他の巻と例外的に異なる背表紙を認識できるようになれば100%に近い確率で漫画本を認識できようである。

4. まとめと今後の課題

本研究では発行部数の多い漫画本に着目した。SIFT 特徴量と HSV ヒストグラムのハイブリット認識を行った結果、同タイトルの別の本が本棚の何処にあるのかということ画像から認識できる確率が、SIFT や HSV に比べて10%上昇した。これにより、認識用のマーカやバーコードを背表紙に後付する必要をなくし、書籍整理機への応用を念頭に置いた場合のユーザの手間の削減が期待できる。

課題点としては、背表紙の背景色や背景画像が同じタイトルの巻ごとに著しく違う画像にも対応できるように手法を改良していく必要がある。

また、本の巻数の認識には至っていないため、バーコードの情報などを用いて巻数を認識する手法を確立する必要がある。さらに、ロボットアームとの連携などを考えて行く必要がある。

参考文献

- 1) ORICON STYLE : 【オリコン年間】『ONE PIECE』7年連続首位『進撃の巨人』は2位
<http://www.oricon.co.jp/news/2045222/>, 2014-12-01, 参照 2015-10-12.
- 2) ネットオフ株式会社 : 家にある、“読んでいない書籍”の数は平均131冊!

<http://www.netoff.co.jp/corp/pdf/20131125.pdf>, 2013-11-25, 参照
2015-10-12.

3) 今中辰弥, 杉山正治, 吉川恒夫 : 3 本指ハンドを有するロボットアームによる書籍の整理, ロボティクス・メカトロニクス講演会講演路概要集 2010, ” 1A2-A29(1)” - ” 1A2-A29(4)” , 2010.

4) 石田弘樹, 杉山正治, 吉川恒夫 : 移動マニピュレータを用いた書籍整理システム, ロボティクス・メカトロニクス講演会講演路概要集 2011, ” 1A1-D09(1)” - ” 1A1-D09(4)” , 2011.

5) A. Jain : Fundamentals of Digital Image Processing. Prentice-Hall, Englewood Cliffs, N.J. , 1989.

6) S.Suzuki, K.Abe. : Topological structural analysis of digital binary image by border following. Computer Vision, Graphics, and Image Processing, Vol. 30, No. 1, pp. 32-46, 1985.

7) David G. Lowe : Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision 2004, Vol. 60, Issue. 2, pp. 91-110, 2004.