

自動本棚整理機のための
漫画書籍タイトル認識手法に関する研究

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

茨田 将史

自動本棚整理機のための
漫画書籍タイトル認識手法に関する研究

指導教員 渡辺 大地

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

茨田 将史

論文の要旨

論文題目	自動本棚整理機のための 漫画書籍タイトル認識手法に関する研究
執筆者氏名	茨田 将史
指導教員	渡辺 大地
キーワード	漫画本、書籍整理、画像認識、SIFT、HSV ヒストグラム

[要旨]

賛近年、漫画本の販売数は上昇傾向にある。オリコンの調査では書籍の半数近くが漫画本である。またネットオフが家の中にどれだけの書籍があるのかアンケート調査を行ったところ、「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。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、これらの従来研究では書籍の背表紙に認識用の自作マー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。

本研究では、画像の特徴に着目し、漫画の背表紙を認識できるのではないかと考えた。近年では画像から特徴を見つけ出し比較する手法として SIFT や SIFT を高速化した SURF、FAST 等の手法が提案されている。本研究では SIFT 特徴量と HSV ヒストグラムを用いて漫画本の背表紙の認識を可能にした。SIFT 特徴量では比較する特徴の高さを考慮することで認識頻度を向上させた。HSV ヒストグラムでは比較対象をタイトルだけにするだけで認識頻度を向上させた。SIFT 特徴量と HSV ヒストグラムの 2 つの幾何平均をとったハイブリッドでは、SIFT ではあまり考慮されていない色の要素を HSV ヒストグラムと掛け合わせ、考慮することで認識頻度を向上させた。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。

同タイトルの漫画本を最低 1 冊認識できる頻度はハイブリッドでは 100% であった。同タイトルの漫画本をすべて認識できる頻度はハイブリッドでは 93% であった。

A b s t r a c t

Title	A study on the comic book title recognition method for automatic bookshelf sorter
Author	Masashi Barada
Advisor	Taichi Watanabe
Key Words	Comic, Book sorter, Image processing, SIFT, HSV histogram

[summary]

In recent years, comic sales of book is on the rise. In Oricon survey nearly half of the book is a comic book. Net off was carried out or questionnaire survey there is a Books of how much inside the house. "Book" is average 356 books. "Comic" is average 298 books. From this, I understand that even an individual possesses a comic book in large quantities. It takes considerable time to organize such a large quantity of books.

For such a problem, a study is conducted about the book organizeing of the library mainly. Referring the human action performed when taking out the Books, was it possible to grip the tilted books or fallen Books by using a robotic arm having three fingers. They also succeeded in carrying out the rearrangement of the Books by realizing the take-out operation and the storing operation. However, pasted your own marker and self-made bar code for recognition on the back cover of the Books to in these previous studies. take a lot of trouble Given that use in the one person.

In this study, focusing on the characteristics of the images were considered to recognize the back cover of cartoons. In late years, SIFT is finds a feature from in image for method of comparing. SURF and FAST was speeding up the SIFT methods. In the present study is to enable the recognition of the back cover of the comic book by using the SIFT feature amount and the HSV histogram. In SIFT feature values is improved the recognition frequency by considering the features of height for comparing. In HSV histogram improved the recognition frequency by only title comparisons. Hybrid was the geometric mean of the SIFT feature value and HSV histogram. Do not consider the color elements of the SIFT. Therefore, improved recognition frequency by combining the HSV histogram. As a result, it eliminates the need to retrofitted with markers and bar code for recognition on the back cover. It was reduce user effort when considering the application of the personal to the Books sorter.

In the hybrid, the probability that the comic book of the same title can be recognized at least one book was 100%. In The hybrid, The probability of recognizing all the comic book of the same title was 93%.

目次

第1章	はじめに	1
1.1	研究背景と目的	2
1.2	論文構成	4
第2章	漫画本認識の提案手法	5
2.1	漫画本について	6
2.2	書籍整理の想定と前提条件	7
2.3	漫画本の背表紙の分割について	9
2.3.1	漫画本の境界検出	9
2.3.2	境界の分割方法	11
2.4	比較方法について	14
2.5	SIFT 特徴量を用いた類似画像検出	15
2.6	HSV ヒストグラムを用いた類似画像検出	16
2.7	ハイブリッドによる類似画像検出	17
2.8	バーコードについて	17
第3章	検証と考察	21
3.1	本棚の分割	22
3.2	SIFT 特徴量を用いた類似画像検出	23
3.3	HSV ヒストグラムを用いた類似画像検出	29
3.4	ハイブリッドによる類似画像検出	33
3.5	各手法の類似画像検出率の比較	37
3.6	漫画本の並べ替えについて	38
第4章	おわりに	40
	謝辞	42
	参考文献	44

目 次

1.1	従来研究の認識法で使用するマーカー	3
2.1	背景色が巻数ごとに違う漫画本	6
2.2	書籍整理の想定図	8
2.3	対象とする本棚画像	9
2.4	本棚の二値化画像 (閾値 128)	10
2.5	適応型閾値を用いた二値化画像	10
2.6	抽出した輪郭線画像	11
2.7	推測した幅	11
2.8	スロットの間隔を狭くする	11
2.9	スロット毎に輪郭要素測定	12
2.10	輪郭要素の多い部分に境界線候補描画	12
2.11	すべてのスロットに境界線候補を描画	12
2.12	余分な境界線候補の除去方法	13
2.13	漫画本 1 冊 1 冊の境界線	14
2.14	評価方法	15
2.15	バーコードの意味	18
2.16	1 巻のバーコード	20
2.17	2 巻のバーコード	20
2.18	3 巻のバーコード	20
3.1	5 冊 5 タイトル 3 冊 1 タイトル	22
3.2	11 冊と 7 冊と 10 冊	22
3.3	同タイトル 28 冊	23
3.4	“スパイラル” 3 巻を注目画像とした SIFT での比較結果	27
3.5	各比較画像を注目画像とした SIFT での比較結果	28
3.6	“スパイラル” 3 巻を注目画像とした HSV での比較結果	31

3.7	各比較画像を注目画像とした HSV での比較結果	32
3.8	“ スパイラル ” 3 巻を注目画像としたハイブリッドでの比較結果	35
3.9	各比較画像を注目画像としたハイブリッドでの比較結果	36
3.10	各手法の類似画像検出率	37

表 目 次

2.1	出版者コードの桁数	19
3.1	各マッチング数での背表紙認識頻度	24
3.2	“スパイラル”3巻を注目画像とした SIFT での比較結果	26
3.3	“スパイラル”3巻を注目画像とした HSV ヒストグラムでの比較結果	30
3.4	“スパイラル”3巻を注目画像としたハイブリッドでの比較結果	34
3.5	出版者コードが共通した漫画本の ISBN	38

第 1 章

はじめに

1.1 研究背景と目的

近年，漫画本の販売数は上昇傾向にある．オリコン [1] の調査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 [2] が 18 歳から 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を行ったところ，「書籍」は平均 356 冊，そのうち「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，これらの従来研究では書籍の背表紙に認識用の自作マーカと自作バーコードを貼り付けている．図 1.1 に従来研究の認識方法を示す．このようなマーカ等を張り付ける作業は，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 等での文字認識は難しい．

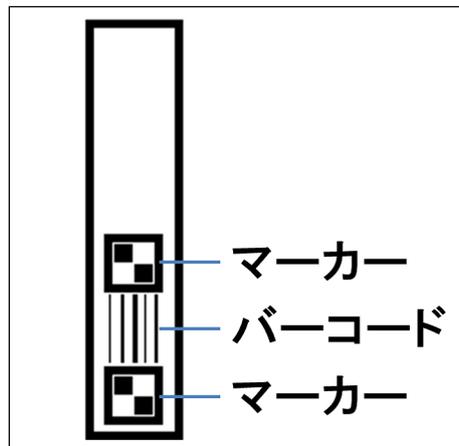


図 1.1 従来研究の認識法で使用するマーカー

本研究では、漫画の背表紙画像の特徴から同じタイトルの漫画本を特定することを目的とした。近年では画像から特徴を見つけ出し比較する手法が確立されている。SIFT[5][6][7] や SIFT を高速化した SURF[8][9] や FAST[10] 等の手法が提案されている。また、SURF をモバイル機器でも扱えるように消費電力の減少や処理の高速化の研究 [11][12][13] も行われている。

SIFT を利用した従来研究として特徴点を利用して画像を張り合わせてパノラマ写真を作る研究 [14] が挙げられる。紙に絵を書き、紙を動かし、SIFT で動きをキャプチャーすることでアニメーションを作成する [15] など様々な分野で応用されている。

現実の物を画像認識で識別する例として、高木ら [16] の SIFT 特徴量を用いた交通道路標識認識が挙げられる。この研究では車載カメラによってとった写真から交通標識を認識することで、ドライバーの交通標識の見落としを防ぐことを目的に行われている。車載カメラで取った画像と 17 種類の交通標識のテンプレート画像とのマッチングを SIFT を用いて行い、88.7% の認識精度となった。このように近年では画像同士を比較し、類似物かどうか判断出来るようになりつつある。

本研究では従来研究のようなマーカーやバーコードを張り付ける作業をなくすために、画像認識を用いて漫画本の背表紙認識を行った。画像認識には、SIFT 特徴量や HSV ヒストグラム、これ

ら 2 つの幾何平均をとったハイブリッドを用いて比較を行った．SIFT 特徴量では比較する特徴の高さを考慮することで認識頻度が向上した．HSV ヒストグラムでは比較対象をタイトルだけにするだけで認識頻度が向上した．ハイブリッドでは SIFT ではあまり考慮していない色の要素を HSV ヒストグラムと掛け合わせ，考慮することで認識頻度が向上した．

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．

1.2 論文構成

本論文の構成は，以下の通りである．第 2 章では，漫画本やバーコードの説明と，漫画本の背表紙の分割法や漫画本同士の比較方法について述べる．第 3 章では，本研究の提案手法をプログラムにて実装し，漫画本の認識の検証について述べる．第 4 章では，本研究における成果と意義をまとめ，今後の展望を述べる．

第 2 章

漫画本認識の提案手法

2.1 節では認識に利用できる漫画本の特徴や問題点について述べる．2.2 節では書籍整理の想定と前提条件について述べる．提案手法は大きく分けて分割と比較の 2 段階で行った．分割については 2.3 節で述べる．比較については 2.4 節で比較方法の説明について述べる．2.5 節，2.6 節，2.7 節ではそれぞれの手法での比較について述べる．

2.1 漫画本について

漫画本には背表紙を画像認識をするにあたって，問題となる点と利用できる特徴がある．漫画本認識における問題点は 3 点ある．第 1 は背表紙に使われている文字のフォントは統一されていない点である．そのため，OCR 等の文字認識で漫画本の背表紙を認識するのは難しい．文字認識以外の方法で背表紙を認識する必要がある．

第 2 はタイトルごとに背表紙の色が様々で二値化等を行う際の閾値の設定などが難しい点である．また，同じタイトルの漫画本でも背景色が異なる場合がある．図 2.1 にその例を示す．これらのことを漫画本の背表紙を 1 冊単位に分割する際に考慮する必要がある．

第 3 は漫画本の背表紙が日焼けするため，同じタイトルの物でも色が著しく変わってしまう場合がある．そのため，漫画本の日焼けなどを考慮した手法を考える必要がある．



図 2.1 背景色が巻数ごとに違う漫画本

以下に利用できる漫画本の特徴を挙げる．

漫画本の横幅

漫画本はページ数にバラツキが少なく，少年漫画だと大体 200 ページ前後になる．これにより，大体の漫画本の幅に予測が付けられるようになる．

同タイトルの背表紙

同じタイトルの漫画本が 1 巻，2 巻，3 巻と連続して何冊も出版されることが挙げられる．多い物だと 100 巻を超えているものもあり，同じタイトルをまとめておく需要が高い．

背表紙の形式

背表紙の形式が似ていることが挙げられる．上から順にタイトル，巻数，著者名となっていることが多い．この特徴を考慮することでタイトル同士の比較などができるようになり，認識精度の向上が期待できる．

バーコードの位置

最後に裏表紙についているバーコードの位置がほぼ一定で情報が読み取り易いことが挙げられる．

これらの点を今回の認識に利用できると考えた．

2.2 書籍整理の想定と前提条件

本研究では本棚とロボットアームを連動させて書籍を整理する想定である．図 2.2 は書籍整理の想定図である．背景が灰色の左側の部分が本棚部分で，タイトルと巻数でソートした漫画本が入っている．背景が薄い黄色の右側の部分が回収箱で，読み終わった漫画本や新しく買ってきた漫画本を入れる場所である．回収箱にはバーコードリーダーが付いており，バーコード情報から本棚内に同じタイトルの本があるかどうか判断し，新しい漫画本に対応する．

アームと移動機構にはカメラと距離センサを取り付け、本棚や回収箱の画像情報を取得し、画像認識を行った。アーム、カメラ、距離センサは横向きだが、本来は本棚側を向いている。

画像認識の結果を元に回収箱から本棚部分へ書籍を移動することで整理を行う。本論文では画像認識を用いて本棚の何処に同じタイトルの漫画本があるのか検索する部分を実装した。

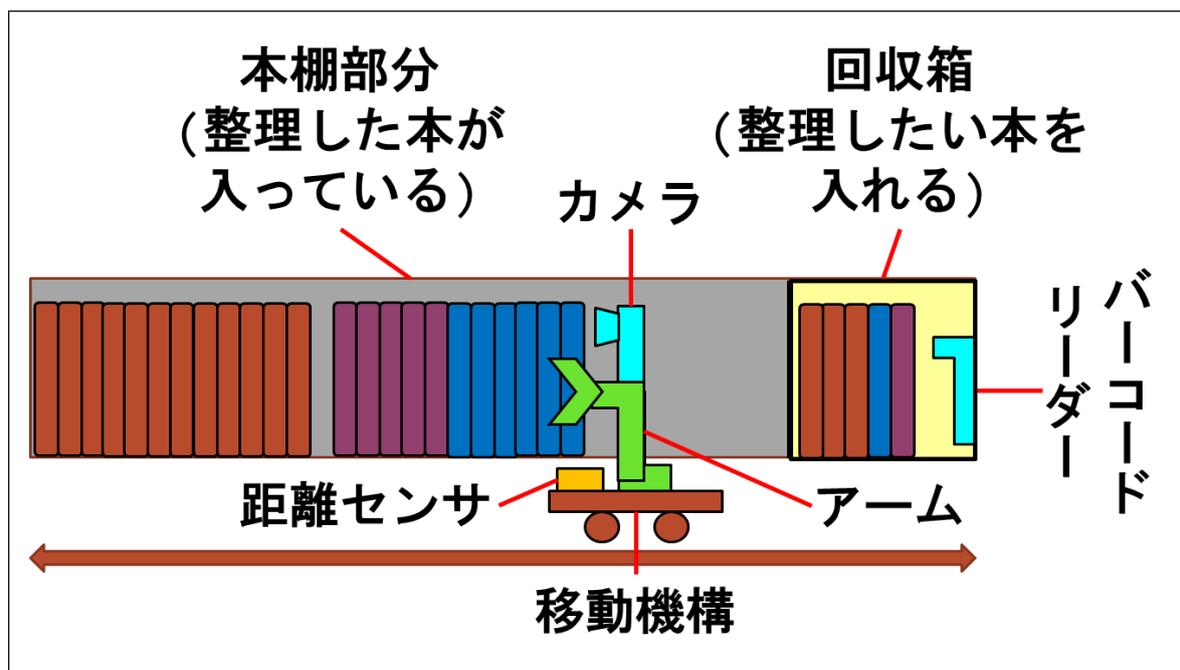


図 2.2 書籍整理の想定図

今回の対象は図 2.3 の本棚画像である。この画像は横 487 ピクセル縦 221 ピクセルの画像である。漫画本は 3 巻あるタイトルが 8 タイトル、4 巻あるタイトルが 1 タイトルの計 28 冊である。本棚画像に様々な処理を加えることで漫画本を 1 冊単位で認識することや、1 冊単位にした漫画本をそれぞれの手法で比較することを行った。



図 2.3 対象とする本棚画像

認識を行う前提条件として、本棚部分と回収箱の漫画本は鉛直に近い状態で立っているものとする。本棚部分の漫画本はタイトルごとに巻数順でソーティングしてあり、1巻2巻4巻のように間に欠損があってもよいものとする。

2.3 漫画本の背表紙の分割について

本棚内にある漫画本の背表紙を1冊単位に分割するためには漫画本の境界を求める必要がある。2.3.1項では漫画本の境界の要素を求めるために様々な画像処理について述べる。2.3.2項では2.3.1項で求めた境界の要素を元に本棚内の漫画本を1冊単位に分割する方法について述べる。

2.3.1 漫画本の境界検出

本棚内にある漫画本の背表紙を1冊単位に分割するために漫画本同士の境界の要素を求める。

境界を求める前処理として二値化を行った。漫画本の背表紙には様々な背景色が使われており、隣合う漫画本の背景色が違うと境界ができやすいが、同じタイトルで同じ背景色の物だと閾値を固定で二値化すると、本来境界があるべき場所に境界が出ない場合が多々ある。図2.4は閾値128で本棚画像を二値化した画像である。同じタイトルの漫画本の境界が消えているのがよく分かる。

適応型閾値で二値化した画像に対して輪郭線の抽出を行った．輪郭線の抽出には鈴木ら [18] の輪郭抽出アルゴリズムを利用した．図 2.6 が輪郭線要素を抽出した画像である．

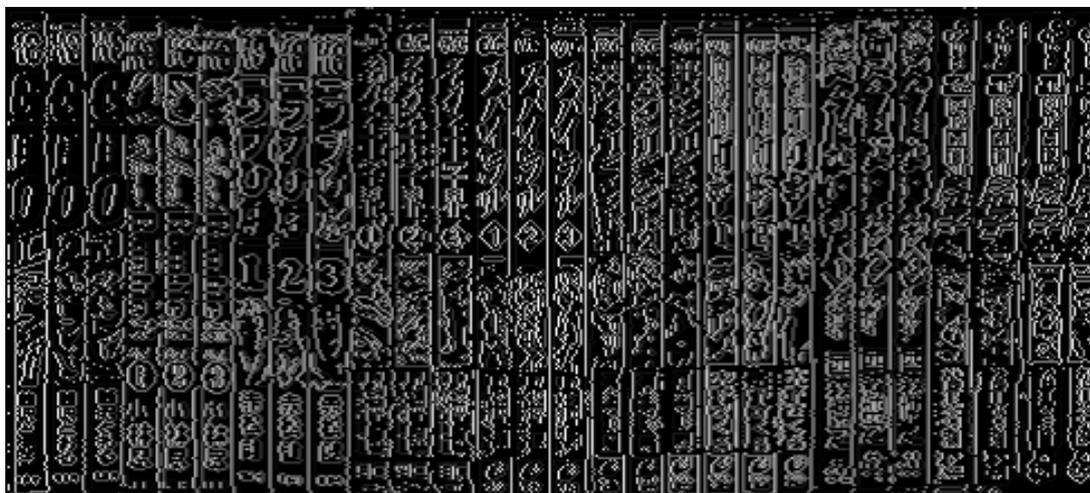


図 2.6 抽出した輪郭線画像

2.3.2 境界の分割方法

境界線の分割は様々な漫画本の特徴を利用することで実現する．2.1 項で説明したように漫画本のページ数はある程度決まっているため，本棚の幅とカメラから写真を撮る位置を考慮することで，本棚に入るおおよその漫画本の数が推測できる．

図 2.7 は推測した漫画本の幅である．このような推測した幅のことを以後スロットと呼ぶ．図 2.8 はスロットの間隔を狭めた図である．これにより，境界線候補の量を増やし，取りこぼしをなくす処理を行った．

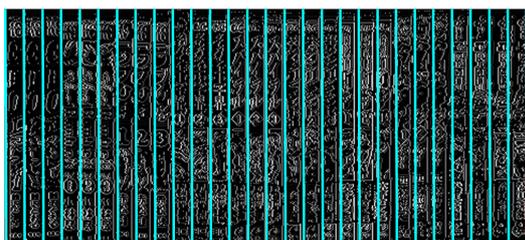


図 2.7 推測した幅

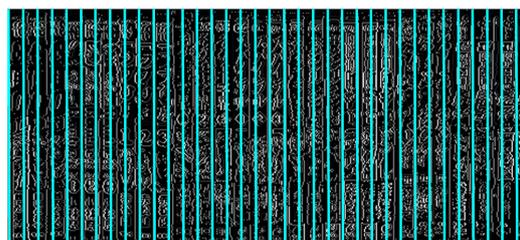


図 2.8 スロットの間隔を狭くする

図 2.9 は 2.3.1 節で検出した輪郭線の要素をスロット毎に計測する方法である．計測は横方向 1 ピクセル毎に縦方向にどれだけ輪郭線の要素があるか測っている．図 2.10 はスロット内で最も輪郭要素が多かった場所に輪郭線候補を 1 本引いた画像である．



図 2.9 スロット毎に輪郭要素測定

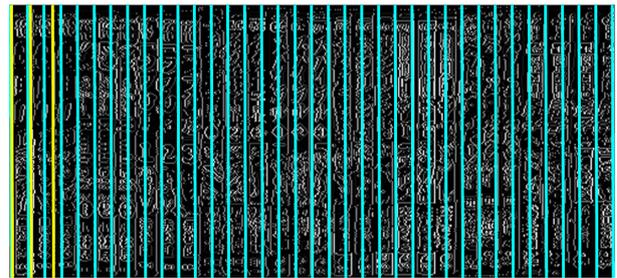


図 2.10 輪郭要素の多い部分に境界線候補描画

すべてのスロットに対し，境界線候補を引いたのが，図 2.11 である．漫画本の境界に線が引けているが，スロットを狭くしたため，無駄な境界線も多い．



図 2.11 すべてのスロットに境界線候補を描画

余分な境界線の除去を行うことで漫画本を 1 冊単位に分割する．先ほど検出した境界線候補間

の距離の平均値を計算し、平均値の半分より小さい値の場合、境界線候補の除外を行った。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た。余分な境界線候補の除去方法は、図 2.12 は除去方法の図解である。

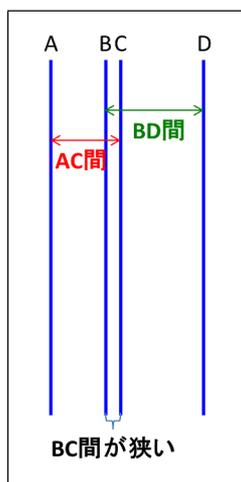


図 2.12 余分な境界線候補の除去方法

連続する 4 本の境界線候補を左から A,B,C,D と置く。境界線候補は y 軸に平行である。 x 座標を A_x, B_x, C_x, D_x とする。境界線候補間の距離の平均値を H とした場合、式 (2.1) で境界線の除去が行われるか判断した。

$$|B_x - C_x| < \frac{H}{2} \quad (2.1)$$

BC 間の距離が平均の半分より小さかった場合、式 (2.2) で除去される境界線を判断した。

$$|A_x - C_x| \leq |B_x - D_x| \quad (2.2)$$

AC 間の距離が BD 間の距離より小さかった場合、境界線候補 B を削除し、BD 間の距離が AC 間の距離より小さかった場合、境界線候補 C を削除する。これにより境界線間の距離が平均的になる。図 2.13 が余分であったすべての境界線候補を除去し、漫画本 1 冊 1 冊の境界線を区切った画像である。



図 2.13 漫画本 1 冊 1 冊の境界線

2.4 比較方法について

本来は回収箱の画像が本棚部分の何処に入れるかを決定するが、今回は回収箱の画像の代わりに本棚内の漫画本の 1 冊を決定し、その漫画本を回収箱の画像の代わりとする。回収箱の画像の代わりに漫画本を注目画像と呼ぶ。本棚内の注目画像以外の画像を比較画像と呼ぶ。本棚内には漫画本が 28 冊あるため、比較画像は 27 冊分の背表紙画像となる。

図 2.14 に比較方法を図解する。注目画像と比較画像 1 枚 1 枚を順番に、それぞれの手法で比較し、最も注目画像に近かった画像が同じタイトルの漫画本かどうかを評価していく。

本研究では SIFT 特徴量と HSV ヒストグラムを合わせたハイブリッドな類似画像検出手法を用いて、類似画像を検出した。2.5 節では SIFT 特徴量を用いた類似画像検出について述べる。2.6 節では HSV ヒストグラムを用いた類似画像検出について述べる。2.7 節では SIFT 特徴量と HSV ヒストグラムを合わせたハイブリッドを用いた類似画像検出について述べる。これにより、類似画像を検出し、同じタイトルの漫画本を収納することで漫画本の整理を行う。

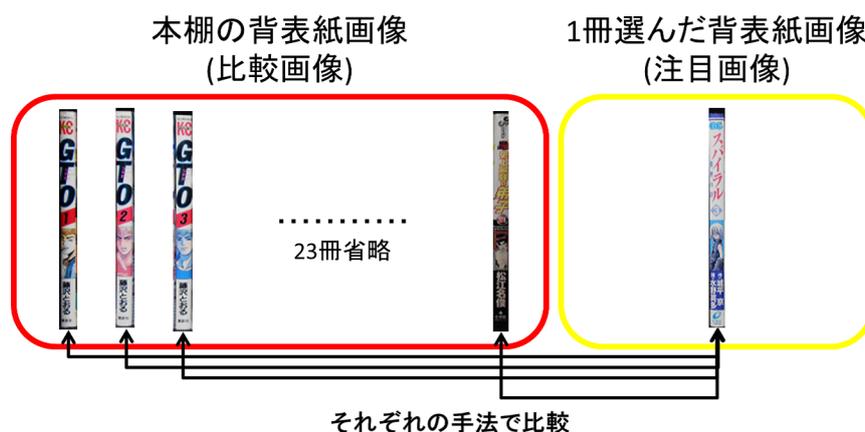


図 2.14 評価方法

2.5 SIFT 特徴量を用いた類似画像検出

Lowe[5] が提案した SIFT 特徴量では特徴を得たい画像に対し強度の違う平滑化処理を行い、差分をとることで画像内で特徴点を抽出する。SIFT の特徴量では周辺領域を 4×4 の 16 に分割し、それぞれのマスに 8 方向の輝度勾配ヒストグラムを作成する。これにより、 $4 \times 4 \times 8$ の 128 次元の特徴量が算出され、回転やスケールの変化に影響されない比較ができる。この 128 次元の特徴を比較することでユークリッド距離として類似度を検出している。ユークリッド距離の値が小さい方が比較元の特徴と類似していることになる。

本研究では、類似画像の検出は以下のように行った。最初に注目画像の SIFT 特徴量を検出する。次に比較画像でも SIFT 特徴量を検出し、注目画像の各特徴点と比較画像の各特徴点と総当たりでユークリッド距離の計算を行った。ユークリッド距離を n 個選び、小さい方から L_1, \dots, L_n とし、平均の値 M を計算する。

$$M = \frac{\sum_{k=1}^n L_k}{n} \quad (2.3)$$

n の最適な個数については、次章で検証を行う。この処理をすべての比較画像に行った。各比較画像の中でユークリッド距離の平均の値が最小だった比較画像をこの手法での正解 (類似画像) と

した。

次に、漫画本の特徴を考慮し、特徴点同士の位置の差で比較結果から除外する手法を述べる。同タイトルの漫画本の場合、タイトル等の高さはほぼ一定になると想定した。

特徴点の位置の差は式 (2.4) で求めた。 E は特徴点同士の差の割合を表す。 y_o は注目画像の特徴点の y の座標を表す。 y_p は比較画像の特徴点の y の座標を表す。 h は漫画本 1 冊の高さを表す。

$$E = \frac{|y_o - y_p|}{h} \quad (2.4)$$

これにより漫画本の背表紙の縦方向の頂点から底辺の距離を 1.0 とした時、 y_o と y_p がどれだけ離れているのかを 0.0 から 1.0 の値で評価した。値が低ければ低い程、特徴点の縦方向の距離が近いということになる。

$$E < G \quad (2.5)$$

G は閾値を表し、検証は 0.2 と 0.05 の値で行った。式 (2.5) が満たされない場合、特徴点が離れすぎているため、同じ漫画本の特徴を参照していないと判断し、比較結果から除外した。

2.6 HSV ヒストグラムを用いた類似画像検出

HSV[19] とはより直感的に色が扱えるようにマンセル表色系に近い、色相 (Hue)、彩度 (Saturation)、明度 (Value) の 3 属性で表される色空間である。色相は色の違いを示す属性である。彩度は色の鮮やかさを示す属性である。明度は各色相の明るさを示す属性である。

類似画像の検出は以下のように行う。注目画像や比較画像の背表紙画像を上下半分に分割し、上のタイトル部分を利用する。タイトルだけの比較を行うことで精度の向上を図っている。

注目画像と比較画像について色相、彩度、明度それぞれのヒストグラムを算出した。ヒストグラムのビン数については 16, 32, 64, 128, 256 で試行した結果、256 の値が最もよかったため 256 とした。

注目画像と比較画像の色相，彩度，明度の各属性同士のバタチャリア距離 [20][21][22] を式 (2.6) で求めた． d はバタチャリア距離を表す． H はそれぞれの画像のヒストグラムの 1 属性を表す． m は属性のビン数を表す． i は属性のビン番号を表す．

$$d(H_1, H_2) = -\log\left(\sum_{i=0}^m \sqrt{H_{1i}H_{2i}}\right) \quad (2.6)$$

正規化を行った，色相，彩度，明度，それぞれの属性に 0.0 から 1.0 の間のバタチャリア距離を算出した．ヒストグラムが完全一致だった場合は 0.0 となり，ヒストグラムが完全不一致だった場合は 1.0 となる．色相，彩度，明度の 3 属性のバタチャリア距離を加算合計し，最小の比較画像をこの手法での正解 (類似画像) とした．

2.7 ハイブリッドによる類似画像検出

ハイブリッドでは 2.5 節で求めた SIFT 特徴量のユークリッド距離を J とし，2.6 節で求めた HSV ヒストグラムのバタチャリア距離を K とし，式 (2.7) にあるように，幾何平均 I をとることで実現した．

$$I = \sqrt{JK} \quad (2.7)$$

幾何平均の結果が最小となる比較画像をこの手法での正解 (類似画像) とした．

2.8 バーコードについて

漫画本には裏表紙の左上の方に元々付いているバーコードがある．漫画本のバーコードは上下 2 段の 13 桁ずつの番号からなる書籍 JAN コード [23] が採用されている．図 2.15 に書籍 JAN コードの意味を記す．



図 2.15 バーコードの意味

下段は書籍の日本図書コードの分類コードと定価を表す価格コードに使われている。

上段は左から 3 桁が書籍用のバーコードであることを表す数字「978」「979」で始まる。次に ISBN と呼ばれる 10 桁があり、この 2 つの要素で構成される。

ISBN には ISBN10 と ISBN13 がある。元々 ISBN は 10 桁であったが、出版点数の増加などの理由により、2 年の移行期間を経て 2007 年 1 月から 13 桁へ拡張された。ただし、ISBN13 は最初に書籍接頭辞という文字が 3 桁追加されただけで、他の 10 桁は ISBN10 と変わらない。ISBN10 の最初の 1 桁は国等の地域の分類、次の数桁が出版者コード、さらに次の数桁が本 1 冊 1 冊に割り振る書名コードとなっている。また、最後の 1 桁は数字の間違いを確認するためのチェックデ

ジットである．出版者コードと書名コードの桁数は国によって違い，日本では出版者コードの最初の 2 桁で決まっている [24]．また，出版者コードと書名コードは 2 つ合わせると 8 桁になる．出版者コードと桁数の関係を表 2.1 にまとめた．

表 2.1 出版者コードの桁数

出版者コードの桁数	最初の 2 桁	出版者コード
2 桁	00 ~ 19	00 ~ 19
3 桁	20 ~ 69	200 ~ 699
4 桁	70 ~ 84	7000 ~ 8499
5 桁	85 ~ 89	85000 ~ 89999
6 桁	90 ~ 94	900000 ~ 949999
7 桁	95 ~ 99	9500000 ~ 9999999

ISBN を利用し，外部のデータベースから本のタイトルや著者名等の様々な情報を得ることができ．しかし，外部データベースは利用料が必要であったり，インターネット接続が必要など利用者に負担がかかる．また，外部のデータベースで並べ替えを行うとなると，本棚内の漫画本すべてのバーコード情報を再度取得する必要がある．本棚内の漫画本の位置情報とバーコード情報を関連付けしたとしても，システム外でユーザにより位置が変更されてしまう可能性がある．そこで，本研究ではローカルなデータベースが作れないか考えた．

回収箱内でバーコードリーダーを使い，ISBN を読み取ることで漫画本の出版者の特定を行う．漫画本の巻数が進むごとに書名コードは増加していた．これを利用し，疑似的な巻数の並べ替えを行う．この出版者コードと書名コードを読み取り，背表紙画像と共にデータベース化を行う．

本棚内の漫画本はデータベース化ができていると仮定した場合，回収箱でバーコードを読み取り，出版者コードを比較することで同じ出版者の漫画本があるかどうか確認できる．同じ出版者の漫画本があった場合には書名コードを利用することで巻数順にソートできるのではないかと考える．対象としたバーコードの書名コードより小さいもののうち最大のものを左側にする．対象

としたバーコードの書名コードより大きいもののうち最小のものを右側にする．これにより，対象としたバーコードに最も近い巻数を求めることができる．

ただし，書名コードは発行日順に割り振られているため，同タイトルの本か判断は出来ない．そのため，データベース内で同出版者コードの最も近い書名コードの背表紙画像とハイブリッドを用いて比較することで，同タイトルの本かどうか特定する．

図 2.16 から図 2.18 はスパイラル各巻のバーコードの上段である．4桁の出版者コードが1巻から3巻で同じことが確認できる．また，4桁の書名コードは巻数が増えるごとに増加している．



図 2.16 1巻のバーコード



図 2.17 2巻のバーコード



図 2.18 3巻のバーコード

第 3 章

検証と考察

本研究では OpenCV2.49[25][26][27] と C++ を用いて実装を行った。

3.1 本棚の分割

今回対象とした本棚の背表紙画像以外にも，今回提案した手法を用いて漫画本の背表紙の分割を行った．図 3.1 と図 3.2 は今回対象とした本棚内の漫画本を入れ替えて分割を行ったものである．図 3.3 はネット上にあった本棚画像に対して分割を行ったものである．



図 3.1 5冊 5タイトル 3冊 1タイトル



図 3.2 11冊と7冊と10冊



図 3.3 同タイトル 28 冊

<http://rurounikensinjoho.blog.so-net.ne.jp/2014-08-08>

どの画像も漫画本の境界部分に境界線を引くことに成功した．また，ネット上にある画像での分割にも成功した．

3.2 SIFT 特徴量を用いた類似画像検出

ここでは，式 (2.3) におけるマッチングの平均個数 n の最適値の検証について述べる．今回は平均をとったマッチング数が 1 個，3 個，5 個，10 個，15 個，20 個での認識冊数の検証を行った．それぞれのマッチング数で注目画像と同じタイトルの漫画本を何冊認識できたかを表 3.1 にまとめた．1 個から 20 個はそれぞれのマッチング数を表す．タイトルと巻数は，注目画像としたタイトルと巻数を表す．0 冊から 3 冊は注目画像と同タイトルの比較画像が他のタイトルの比較画像より，何冊比較結果が良かったかを表す．3 巻あるタイトルの最大認識数は注目画像を除いた 2 冊で，4 巻あるタイトルの最大認識数は注目画像を除いた 3 冊である．

表 3.1 各マッチング数での背表紙認識頻度

タイトルと巻数	1 個	3 個	5 個	10 個	15 個	20 個
GTO1 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
GTO2 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
GTO3 巻	1 冊	2 冊	2 冊	2 冊	2 冊	2 冊
スクールランブル 1 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
スクールランブル 2 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
スクールランブル 3 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
ラブひな 1 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
ラブひな 2 巻	2 冊	2 冊	2 冊	2 冊	2 冊	1 冊
ラブひな 3 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
ヴァンパイア十字界 1 巻	1 冊	2 冊	2 冊	2 冊	2 冊	2 冊
ヴァンパイア十字界 2 巻	1 冊	2 冊	2 冊	2 冊	2 冊	2 冊
ヴァンパイア十字界 3 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
スパイラル 1 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
スパイラル 2 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
スパイラル 3 巻	2 冊	2 冊	2 冊	2 冊	2 冊	1 冊
東京アンダーグラウンド 1 巻	1 冊	1 冊	1 冊	1 冊	1 冊	1 冊
東京アンダーグラウンド 2 巻	1 冊	1 冊	1 冊	1 冊	1 冊	1 冊
東京アンダーグラウンド 3 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
魔探偵口キ 1 巻	2 冊	2 冊	2 冊	1 冊	0 冊	0 冊
魔探偵口キ 2 巻	1 冊	1 冊	1 冊	1 冊	1 冊	1 冊
魔探偵口キ 3 巻	1 冊	1 冊	1 冊	1 冊	0 冊	0 冊
今日から俺は!!1 巻	2 冊	2 冊	2 冊	2 冊	1 冊	1 冊
今日から俺は!!2 巻	2 冊	2 冊	2 冊	2 冊	2 冊	2 冊
今日から俺は!!3 巻	1 冊	2 冊	1 冊	1 冊	1 冊	1 冊
史上最強の弟子 1 巻	3 冊	3 冊	3 冊	2 冊	2 冊	2 冊
史上最強の弟子 2 巻	3 冊	3 冊	3 冊	3 冊	2 冊	2 冊
史上最強の弟子 3 巻	3 冊	3 冊	3 冊	3 冊	3 冊	3 冊
史上最強の弟子 4 巻	3 冊	2 冊	2 冊	2 冊	2 冊	2 冊

マッチング数が多ければ多いほど精度が下がることがわかった。マッチング数が 3 個の場合が、

今回最も高い精度となった。また、マッチング数 1 個では極端に数が少なすぎて誤認識を起していると考えられる。これ以降の検証では最も精度が高かった、マッチング数が 3 個を用いることとする。

“スパイラル” 3 巻を注目画像とし、それぞれの比較画像と比較を行った。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表 3.2 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のかを表す。「ユークリッド距離」は式 (2.3) で求めた値で、どれだけ元の画像と似ているかを表す。高さ考慮とは式 (2.4) と式 (2.5) の結果、特徴点同士の高さの差が大きいマッチングを除外したことを表す。除外した結果、マッチング数が 3 未満になったものはマッチ不足と記す。ユークリッド距離が小さい方がより注目画像に近い。

表 3.2 “スパイラル” 3 巻を注目画像とした SIFT での比較結果

タイトルと巻数	ユークリッド 距離	マッチング 数	ユークリッド 距離 (高さ考慮)	マッチング 数 (高さ考慮)
GTO1 巻	298.7	62	399.3	3
GTO2 巻	352.9	37	443.6	3
GTO3 巻	295.3	70	338.2	4
スクールランブル 1 巻	293.4	58	424.7	3
スクールランブル 2 巻	301.5	46	418.6	6
スクールランブル 3 巻	361.7	54	367.8	7
ラブひな 1 巻	291.4	75	360.8	8
ラブひな 2 巻	331.4	61	388.8	7
ラブひな 3 巻	319.7	45	363.7	6
ヴァンパイア十字界 1 巻	233.6	69	233.6	15
ヴァンパイア十字界 2 巻	307.8	58	370.6	7
ヴァンパイア十字界 3 巻	244.1	50	289.2	8
スパイラル 1 巻	134.9	33	134.9	18
スパイラル 2 巻	109.4	54	109.4	20
スパイラル 3 巻	0.0	57	0.0	57
東京アンダーグラウンド 1 巻	237.7	32	447.0	3
東京アンダーグラウンド 2 巻	252.4	53	287.5	8
東京アンダーグラウンド 3 巻	272.3	46	402.5	4
魔探偵ロキ 1 巻	309.1	23	マッチ不足	0
魔探偵ロキ 2 巻	289.4	34	311.1	7
魔探偵ロキ 3 巻	255.0	43	265.1	9
今日から俺は!!1 巻	287.1	59	392.1	6
今日から俺は!!2 巻	339.8	29	マッチ不足	2
今日から俺は!!3 巻	218.6	52	297.1	6
史上最強の弟子 1 巻	282.1	55	453.4	3
史上最強の弟子 2 巻	281.0	46	404.5	3
史上最強の弟子 3 巻	304.7	36	マッチ不足	1
史上最強の弟子 4 巻	283.8	36	459.0	3

式 (2.4) と式 (2.5) で高さを考慮しなかった SIFT で比較した結果、注目画像と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 1 番小さいという結果になった。” スパイラル ” 1 巻は 2 番目に小さいという結果になり、2 冊とも注目画像に似ていると判断した。背景色がまるで違う ” ヴァンパイア十字界 ” が比較的低いユークリッド距離となった。これは注目画像と原作者と著者が同じため、著者名等の特徴を参照したと考えられる。

式 (2.4) と式 (2.5) で高さを考慮した SIFT で比較した結果を図 3.4 にグラフ化した。グラフの値は高さを考慮したユークリッド距離である。左から順に表 3.2 のタイトルと巻数と同じ順番で並んでいる。注目画像と同じタイトルの結果は赤色になっている。

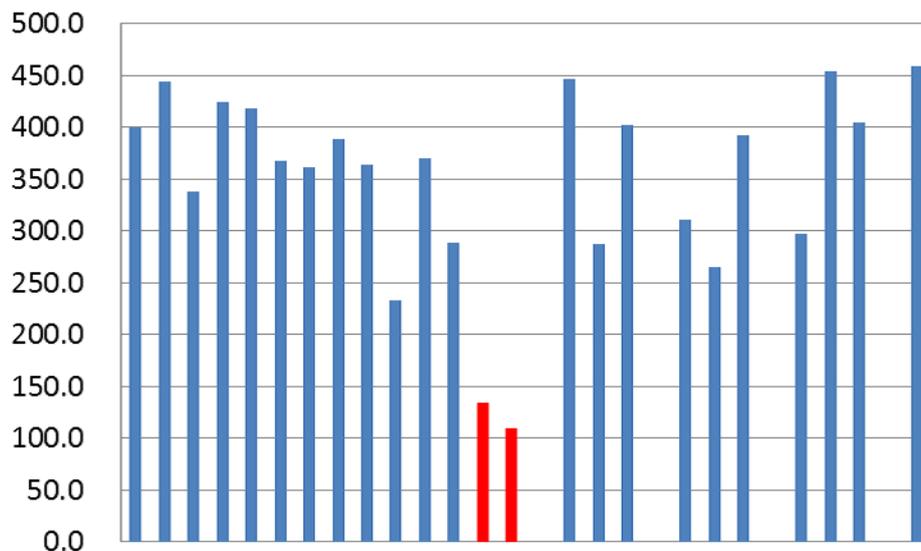


図 3.4 “ スパイラル ” 3 巻を注目画像とした SIFT での比較結果

高さを考慮しない場合と同じく、” スパイラル ” 1 巻 2 巻が最も似ている結果となった。高さを考慮したことでマッチング数が減少し、注目画像と同じタイトル以外の漫画本はユークリッド距離が増加した。これにより、誤認識が軽減できると考える。

上記と同じ条件で本棚内のすべての漫画本を順番に注目画像とし、実験を行った。その結果を図 3.5 にグラフ化した。グラフの値は高さを考慮したユークリッド距離で、注目画像と同じタイ

トルの比較画像の平均値を青色，注目画像と違うタイトルの比較画像の平均値を赤色とした．左から順に表 3.2 のタイトルと巻数と同じ順番で並んでいる．

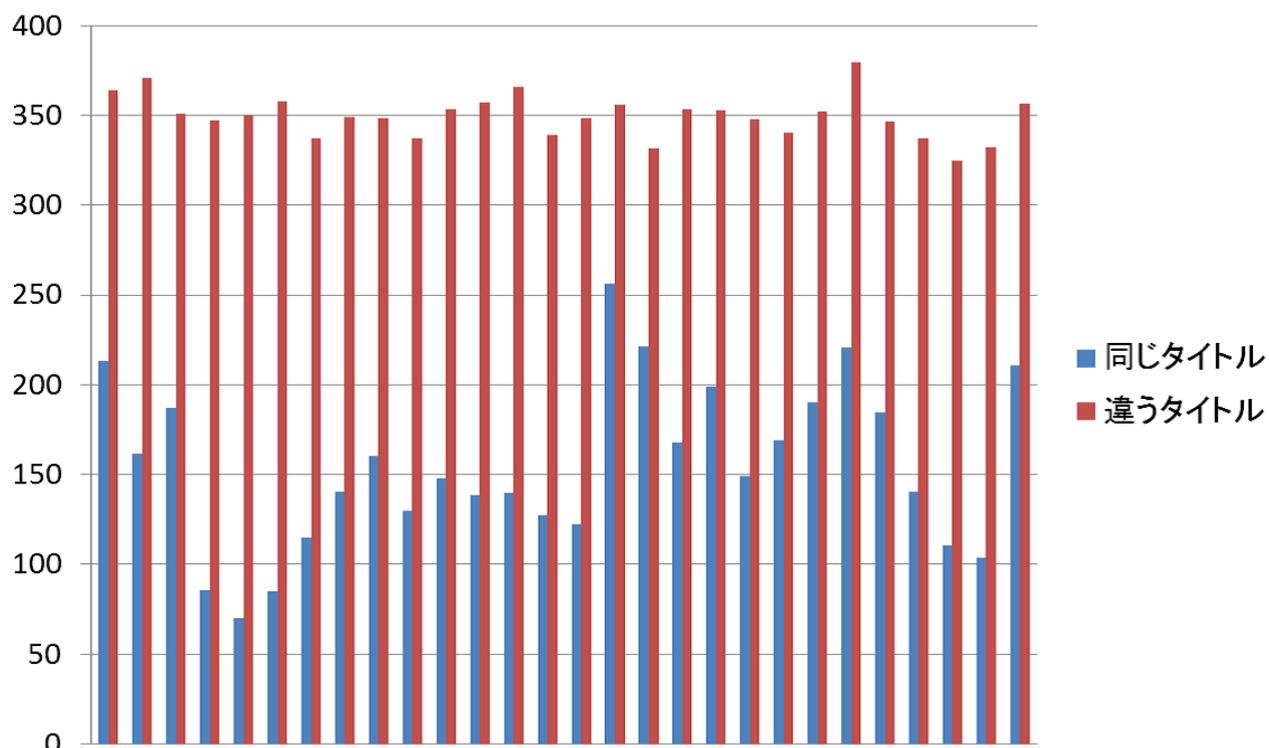


図 3.5 各比較画像を注目画像とした SIFT での比較結果

結果は ”東京アンダーグラウンド” と ”魔探偵ロキ” と ”史上最強の弟子” 以外の漫画本は同タイトルの巻数違いをすべて識別できた．

“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” について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．

“史上最強の弟子” については 1 巻が日焼けをしていることと，タイトルの文字の色が異なるため，似ている特徴が出にくいのではないかと推測する．

高さを考慮しない SIFT で比較した結果は、注目画像と同じタイトルの本が 1 位をとった頻度は 28 冊中 28 冊となり、100% となった。注目画像と同じタイトルの本が比較の結果、上位を独占した頻度は 28 冊中 22 冊となり、79% となった。

高さを考慮した SIFT の閾値が 0.2 の場合、通常の SIFT と認識頻度は同じであった。高さを考慮した SIFT の閾値が 0.05 の場合、注目画像と同じタイトルの本が比較の結果、1 位をとった頻度は 28 冊中 28 冊となり、100% となった。注目画像と同じタイトルの本が比較の結果、上位を独占した頻度は 28 冊中 23 冊となり、82% となった。高さを考慮することにより、上位を独占した頻度で 1 タイトル多く認識できるようになった。

高さを考慮してもあまり効果が得られなかったのは、注目画像と同タイトルの漫画本は元々同じような高さを比較しており、背景色や文字の色に差がないタイトルでは、あまりユークリッド距離の変化は見られなかったためと考えられる。これは背景色や文字の色が違うタイトルでは注目画像と同タイトルの漫画本でもユークリッド距離の変化が起きてしまったため、認識頻度がそこまで向上しなかったと思われる。

3.3 HSV ヒストグラムを用いた類似画像検出

“スパイラル” 3 巻を注目画像とし、それぞれの比較画像と比較を行った。注目画像と各比較画像との HSV ヒストグラムを用いたそれぞれの属性の比較結果を表 3.3 にまとめた。結果の数値は HSV それぞれの属性のヒストグラムについて計算したバタチャリア距離と 3 属性の加算合計である。

表 3.3 “スパイラル” 3巻を注目画像とした HSV ヒストグラムでの比較結果

タイトルと巻数	H	S	V	加算合計
GTO1 巻	0.457	0.533	0.510	1.50
GTO2 巻	0.437	0.495	0.620	1.55
GTO3 巻	0.446	0.459	0.656	1.56
スクールランブル 1 巻	0.462	0.426	0.560	1.45
スクールランブル 2 巻	0.456	0.425	0.583	1.46
スクールランブル 3 巻	0.525	0.468	0.496	1.49
ラブひな 1 巻	0.464	0.388	0.465	1.32
ラブひな 2 巻	0.595	0.407	0.365	1.37
ラブひな 3 巻	0.526	0.494	0.456	1.48
ヴァンパイア十字界 1 巻	0.567	0.495	0.773	1.84
ヴァンパイア十字界 2 巻	0.554	0.416	0.753	1.72
ヴァンパイア十字界 3 巻	0.510	0.415	0.816	1.74
スパイラル 1 巻	0.432	0.361	0.356	1.15
スパイラル 2 巻	0.410	0.353	0.347	1.11
スパイラル 3 巻	0.0	0.0	0.0	0.0
東京アンダーグラウンド 1 巻	0.671	0.406	0.528	1.61
東京アンダーグラウンド 2 巻	0.603	0.575	0.557	1.73
東京アンダーグラウンド 3 巻	0.579	0.482	0.623	1.68
魔探偵ロキ 1 巻	0.555	0.432	0.708	1.70
魔探偵ロキ 2 巻	0.505	0.422	0.664	1.59
魔探偵ロキ 3 巻	0.522	0.424	0.759	1.70
今日から俺は!!1 巻	0.617	0.458	0.686	1.76
今日から俺は!!2 巻	0.591	0.438	0.836	1.86
今日から俺は!!3 巻	0.697	0.475	0.754	1.93
史上最強の弟子 1 巻	0.626	0.523	0.870	2.02
史上最強の弟子 2 巻	0.535	0.507	0.911	1.95
史上最強の弟子 3 巻	0.552	0.490	0.925	1.97
史上最強の弟子 4 巻	0.661	0.466	0.931	2.06

“スパイラル” 3巻を注目画像とした比較結果を図 3.6 にグラフ化した。グラフの値は 3 属性の

バタチャリア距離の加算合計である．左から順に表 3.3 のタイトルと巻数と同じ順番で並んでいる．注目画像と同じタイトルの結果は赤色になっている．

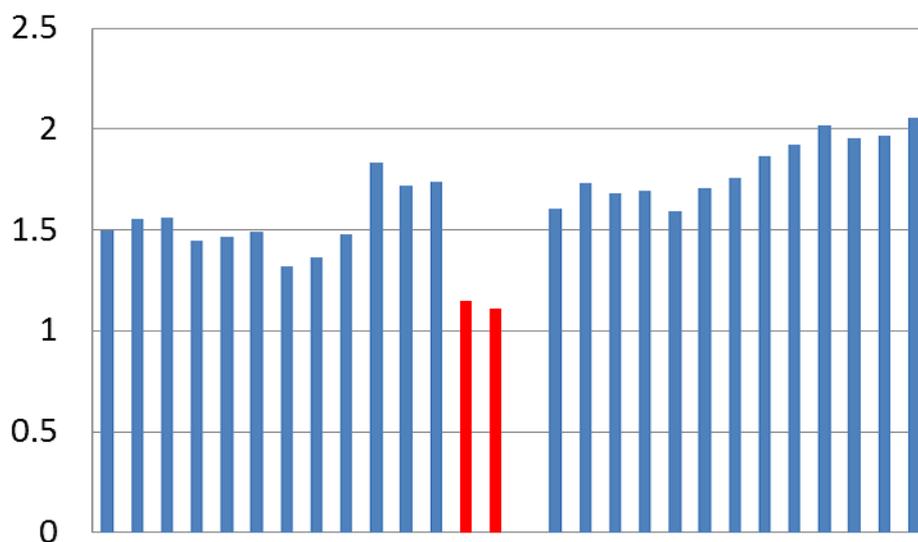


図 3.6 “スパイラル” 3 巻を注目画像とした HSV での比較結果

注目画像と同タイトルの巻数違いである“スパイラル” 2 巻のバタチャリア距離の合計が 1 番小さいという結果になった．“スパイラル” 1 巻は 2 番目に小さいという結果になり，2 冊とも注目画像に似ていると判断された．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低いバタチャリア距離になった．

上記と同じ条件で本棚内のすべての漫画本を順番に注目画像とし，実験を行った．その結果を図 3.7 にグラフ化した．グラフの値は 3 属性のバタチャリア距離の加算合計で注目画像と同じタイトルの比較画像の平均値を青色，注目画像と違うタイトルの比較画像の平均値を赤色とした．左から順に表 3.3 のタイトルと巻数と同じ順番で並んでいる．

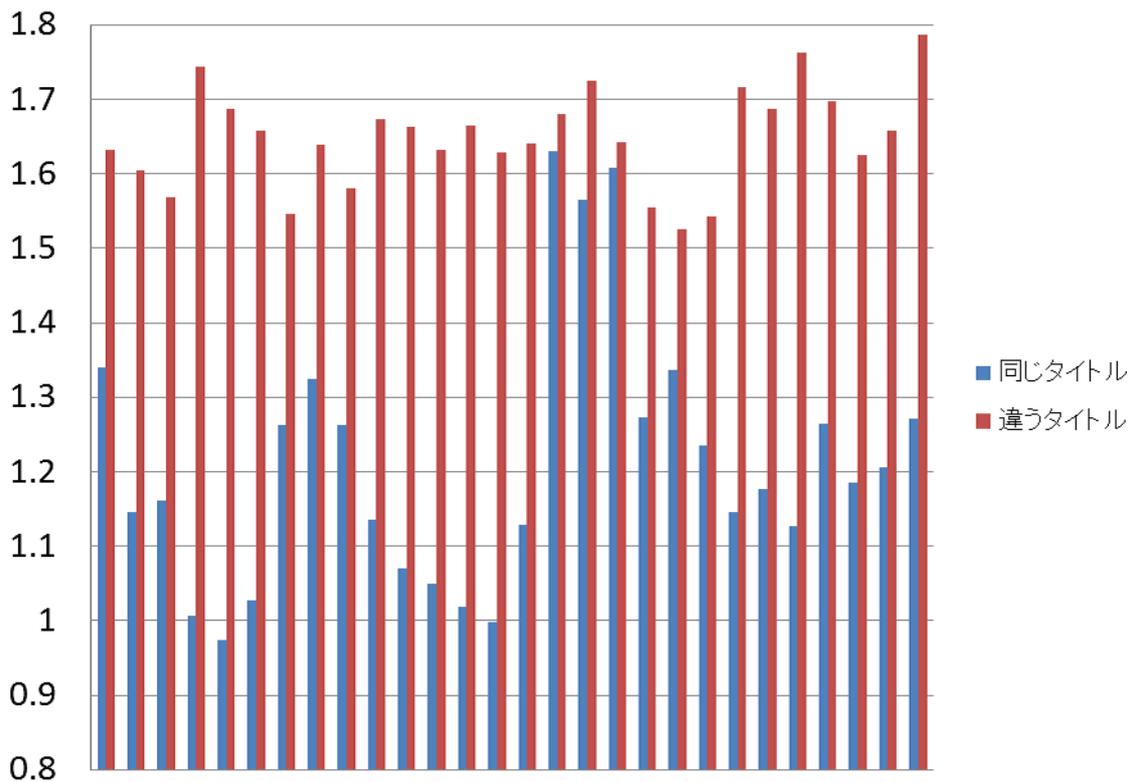


図 3.7 各比較画像を注目画像とした HSV での比較結果

結果は“ラブひな”の1巻2巻，“東京アンダーグラウンド”1巻から3巻，“魔探偵ロキ”1巻は、1冊も同タイトルの本を認識できなかった。他の漫画本は少なくとも1冊は同タイトルを認識している。

“ラブひな”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。

“東京アンダーグラウンド”についてはそれぞれの巻が異なる背景色を持つデザインであることが大きな原因だと考えられる。

“魔探偵ロキ”についてはタイトルの文字の色が異なるデザインが原因だと考えられる。

HSV ヒストグラムによる比較の結果、注目画像と同じタイトルの本が1位をとった頻度は22

冊中 28 冊となり，79% となった．注目画像と同じタイトルの本が比較の結果，上位を独占した頻度は 28 冊中 16 冊となり，57% となった．SIFT と比べるとかなり低い認識頻度となった．背景色の同じ本にはある程度の精度があるが，背景色が違う物が混ざってしまうとかなり認識頻度が落ちる結果になった．

3.4 ハイブリッドによる類似画像検出

“スパイラル” 3 巻を注目画像とし，それぞれの比較画像と比較を行った．注目画像と各比較画像との SIFT(高さ考慮，閾値 0.05) の比較結果，HSV ヒストグラムの比較結果，それら 2 つの幾何平均であるハイブリッドの比較結果を表 3.4 にまとめた．SIFT に関しては閾値で除外した後にマッチング数が 3 未満のものはマッチ不足と記す．

表 3.4 “スパイラル” 3 巻を注目画像としたハイブリッドでの比較結果

タイトルと巻数	バタチャリア 距離	ユークリッド 距離 (高さ考慮)	幾何平均 (ハイブリッド)
GTO1 巻	1.50	399.3	598.9
GTO2 巻	1.55	443.6	688.4
GTO3 巻	1.56	338.2	528.2
スクールランブル 1 巻	1.45	424.7	614.9
スクールランブル 2 巻	1.46	418.6	613.3
スクールランブル 3 巻	1.49	367.8	547.7
ラブひな 1 巻	1.32	360.8	475.1
ラブひな 2 巻	1.37	388.8	531.6
ラブひな 3 巻	1.48	363.7	537.0
ヴァンパイア十字界 1 巻	1.84	233.6	428.8
ヴァンパイア十字界 2 巻	1.72	370.6	638.4
ヴァンパイア十字界 3 巻	1.74	289.2	503.5
スパイラル 1 巻	1.15	134.9	155.0
スパイラル 2 巻	1.11	109.4	121.4
スパイラル 3 巻	0.0	0.0	0.0
東京アンダーグラウンド 1 巻	1.61	447.0	717.8
東京アンダーグラウンド 2 巻	1.73	287.5	498.8
東京アンダーグラウンド 3 巻	1.68	402.5	677.9
魔探偵ロキ 1 巻	1.70	マッチ不足	マッチ不足
魔探偵ロキ 2 巻	1.59	311.1	495.0
魔探偵ロキ 3 巻	1.70	265.1	451.8
今日から俺は!!1 巻	1.76	392.1	690.6
今日から俺は!!2 巻	1.86	マッチ不足	マッチ不足
今日から俺は!!3 巻	1.93	297.1	572.2
史上最強の弟子 1 巻	2.02	453.4	915.2
史上最強の弟子 2 巻	1.95	404.5	790.1
史上最強の弟子 3 巻	1.97	マッチ不足	マッチ不足
史上最強の弟子 4 巻	2.06	459.0	944.8

“スパイラル” 3 巻を注目画像とした比較結果を図 3.8 にグラフ化した。グラフの値は SIFT の

ユークリッド距離と HSV のバタチャリア距離の幾何平均である．左から順に表 3.4 のタイトルと巻数と同じ順番で並んでいる．注目画像と同じタイトルの結果は赤色になっている．

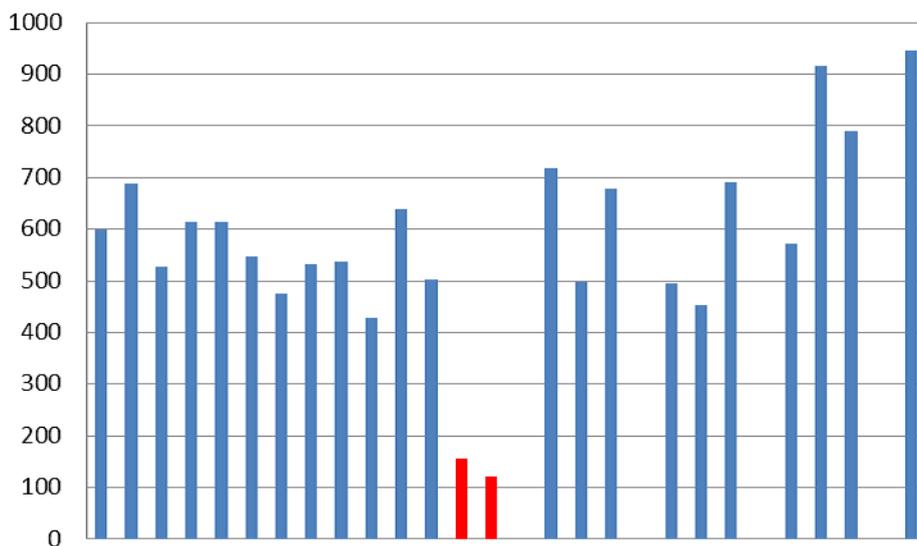


図 3.8 “スパイラル” 3 巻を注目画像としたハイブリッドでの比較結果

注目画像と同タイトルの巻数違いである“スパイラル” 2 巻の幾何平均が 1 番小さいという結果になった．“スパイラル” 1 巻は 2 番目に小さいという結果になり，2 冊とも注目画像に似ていると判断された．SIFT や HSV ヒストグラムに比べて注目画像と違うタイトルの本の幾何平均の値はかなり大きい物となり，注目画像と同タイトルの本との差が開く結果となった．

上記と同じ条件で本棚内のすべての漫画本を順番に注目画像とし，実験を行った．その結果を図 3.9 にグラフ化した．グラフの値は SIFT のユークリッド距離と HSV のバタチャリア距離の幾何平均で，注目画像と同じタイトルの比較画像の平均値を青色，注目画像と違うタイトルの比較画像の平均値を赤色とした．左から順に表 3.4 のタイトルと巻数と同じ順番で並んでいる．

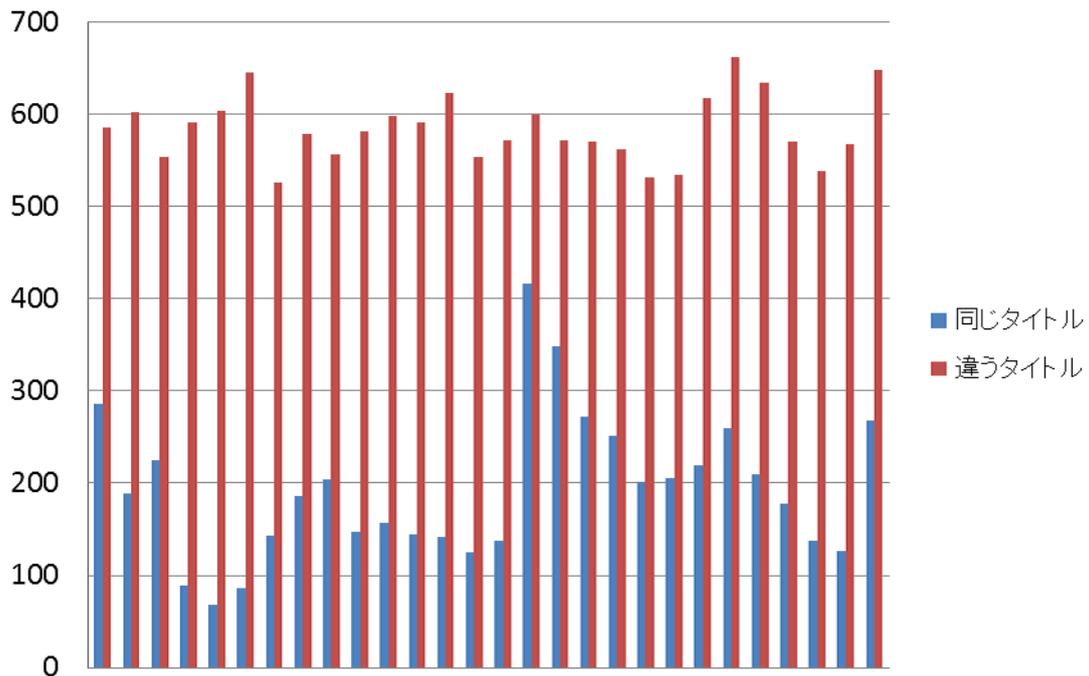


図 3.9 各比較画像を注目画像としたハイブリッドでの比較結果

結果は“東京アンダーグラウンド”1巻2巻以外の漫画本はすべて同タイトルの巻数違いをすべて識別できた。

“東京アンダーグラウンド”については元々 SIFT と HSV では認識出来ていなかったため、幾何平均をとってもあまり効果がなかったと考えられる。

SIFT では認識できなかった“魔探偵ロキ”2巻3巻と“史上最強の弟子”4巻で同タイトルの巻数違いをすべて認識できるようになった。

“魔探偵ロキ”2巻3巻は SIFT と HSV では同タイトルの巻数違いをすべて認識することはできなかったが、認識できなかった巻も比較的上位に位置していた。SIFT と HSV を組み合わせることで誤認識していたタイトルが下位に落ち、注目画像と同じタイトルが上位を維持したと考えられる。

SIFT では認識できていなかった“史上最強の弟子”4巻については、注目画像とは違うタイト

ルの背景色が違う色の漫画本を正解としていたが、HSV の色の要素が組み合わさることで、元々 2 位だった比較画像が 6 位まで下がり、元々 3 位だった注目画像と同タイトルの本が 2 位に上昇した。

ハイブリッドによる比較の結果、注目画像と同じタイトルの本が 1 位をとった頻度は 28 冊中 28 冊となり、100% となった。注目画像と同じタイトルの本が上位を独占した頻度は 28 冊中 26 冊となり、93% となった。

3.5 各手法の類似画像検出率の比較

各手法の類似画像の検出率を図 3.10 にまとめた。SIFT は特に手を加えていない SIFT 特徴量の認識頻度を表す。SIFT(高さ考慮) は高さを考慮した場合の SIFT 特徴量の認識頻度を表す。HSV ヒストグラムは HSV のヒストグラム同士で比較したときの認識頻度を表す。ハイブリッドは SIFT の高さを考慮したものと HSV ヒストグラムで幾何平均をとったときの認識頻度を表す。同タイトルが 1 位とは注目画像と同じタイトルの比較画像が比較の結果、1 位になった頻度を表す。同タイトルが上位独占とは注目画像と同じタイトルの比較画像が比較の結果、3 巻ある漫画本なら 1 位と 2 位、4 巻ある漫画本なら 1 位と 2 位と 3 位になった頻度を表す。

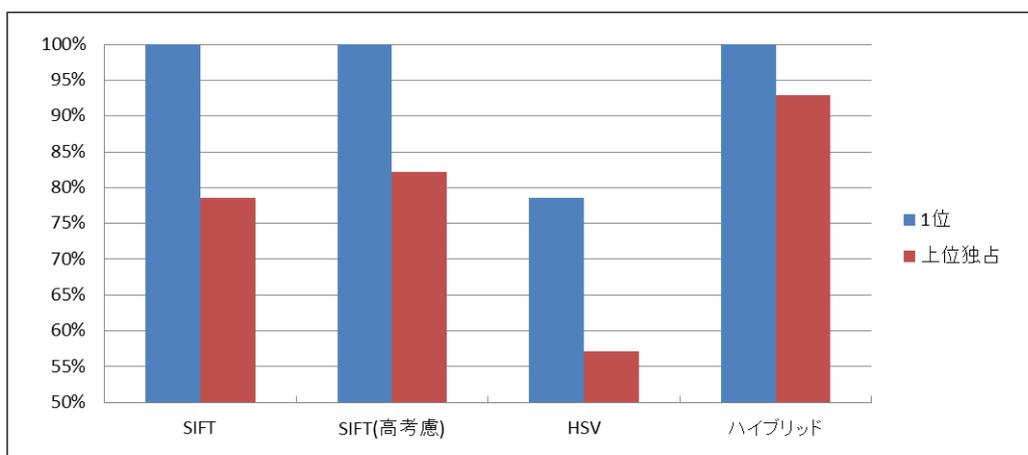


図 3.10 各手法の類似画像検出率

同タイトルが1位になる頻度では，SIFT とハイブリッドでは100%であった．HSV ヒストグラムでは79%であった．

同タイトルが上位を独占する頻度では，通常の SIFT が79%，高さを考慮した SIFT では82%であった．HSV ヒストグラムでは57%であった．ハイブリットでは93%であった．

HSV ヒストグラムは全体的に低い数値になったが，色情報をあまり考慮しない SIFT と組み合わせることで，精度向上に繋がった．

3.6 漫画本の並べ替えについて

バーコードを用いて漫画本の巻数の推測を行った．“スパイラル”2巻が回収箱に入っていると想定した場合，“スパイラル”1巻と3巻の前後関係が判断できるか実験を行った．

表 3.5 出版者コードが共通した漫画本の ISBN

タイトルと巻数	ISBN	出版者コード	書名コード
スパイラル 1 巻	4757501751	7575	0175
スパイラル 2 巻	4757502710	7575	0271
スパイラル 3 巻	4757503700	7575	0370
ヴァンパイア十字界 1 巻	4757511051	7575	1105
ヴァンパイア十字界 2 巻	4757512066	7575	1206
ヴァンパイア十字界 3 巻	475751297X	7575	1297

“スパイラル”2巻の ISBN は4757502710である．ISBN の最初の1桁の“4”は国等の分類のため除外し，次の出版者コードの最初の2桁で出版者コードと書名コードの桁数が決まる．この場合，出版者コードの最初の2桁が“75”なので出版者コードは4桁となる．出版者コードと書名コードは合わせて8桁なので，書名コードは4桁となる．

まずは，出版者コードが完全一致する漫画本を探す．今回のローカルデータベースでは“スパイラル”2巻と出版者コードが完全一致する物は“スパイラル”1巻と3巻，“ヴァンパイア十字

界”1巻,2巻,3巻である。“スパイラル”2巻より書名コードが小さいのは“スパイラル”1巻だけであった。“スパイラル”2巻より書名コードが大きいのは“スパイラル”3巻と“ヴァンパイア十字界”1巻,2巻,3巻であった。“スパイラル”3巻と“ヴァンパイア十字界”1巻,2巻,3巻のうち,最も“スパイラル”2巻の書名コードに近いのは“スパイラル”3巻であった。これにより,“スパイラル”2巻は“スパイラル”1巻と3巻の間に置くのが最適と判断できた。

ISBNを用いて本棚内に同じタイトルの漫画本の有無を判断できるか検証を行った。検証には,式(2.7)で算出したハイブリッドの値を用いる。それぞれ,本棚内に同じタイトルの漫画本がある場合とない場合でどの程度の値になるか調べ,同タイトルを探す判断材料となる閾値を設定できないかと考えた。

その結果,同タイトルの漫画本が本棚内にある場合のハイブリッドの値は,最小値が59,最大値が350,平均値は182となった。同タイトルの漫画本が本棚内がない場合のハイブリッドの値は,最小値が255,最大値が527,平均値は398となった。

28タイトルすべてを識別するために,同タイトルの漫画本が本棚内にある場合の最大値350を閾値とすると,28タイトル中22タイトルで本棚内に同じタイトルの漫画本があると正しく判断できる。28タイトル中6タイトルで本棚内に同じタイトルの漫画本があると誤認識してしまう。同タイトルの漫画本が本棚内がない場合の最小値226を閾値とすると,28タイトル中24タイトルで本棚内に同じタイトルの漫画本があると正しく判断できる。28タイトル中4タイトルで本棚内に同じタイトルの漫画本があると誤認識してしまう。このように,本研究の手法では本棚内の同タイトルの漫画本の有無を完全に判断するのは難しいのが現状である。

第 4 章

おわりに

本研究では漫画本の背表紙を認識するために、SIFT 特徴量や HSV ヒストグラム、これら 2 つの幾何平均をとったハイブリッドを用いて比較を行った。SIFT 特徴量では比較する特徴の高さを考慮することで認識頻度が向上した。HSV ヒストグラムでは比較対象をタイトルだけにするこ
とで認識頻度が向上した。ハイブリッドでは SIFT ではあまり考慮されていない色の要素を HSV
ヒストグラムと掛け合わせ、考慮することで認識頻度が向上した。また、既存のバーコードを用
いることで疑似的に漫画本を巻数順に並べ替えることができた。

今後の課題として、同タイトルで背景色の著しく違う画像への対応が挙げられる。現在の手法
だと、SIFT では特徴の違い、HSV では色の違いが課題となっている。また、回収箱に入れる漫
画本と同タイトルの漫画本が本棚内に入っているかが特定できない。

今後の展望として画像の認識ではタイトルの複雑度等が計測出来れば今の手法と組み合わせ
て更なる認識頻度の向上に繋がりそうである。また、回収箱に入れる漫画本と同タイトルの漫画
本が本棚内に入っているか特定する仕組みを考えなければならない。ロボットに関してもアームの
移動機構や、実際の取出し動作、ロボットへの命令系統など様々な物が挙げられる。

なお、本研究は、情報処理学会 グラフィクスと CAD 研究会 第 161 回研究発表会において“漫
画本整理のための背表紙認識の研究 [28]”として発表した内容を含む。

謝辭

本研究を締めくくるにあたり，ご指導ならびに適切なお助言を下さいました先生方に感謝の意を表します．また，様々な相談に応じて下さった，研究室のメンバーに深く感謝致します．

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/>. 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>. 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. 日本機械学会, ロボティクス・メカトロニクス講演会講演概要集, pp. “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. 日本機械学会, ロボティクス・メカトロニクス講演会講演概要集, pp. “1A1-D09(1)”-“1A1-D09(4)”, 2011.
- [5] D. G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the International Conference on Computer Vision-Volume 2 - Volume 2*, ICCV '99, pp. 1150-, Washington, DC, USA, 1999. IEEE Computer Society.
- [6] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Journal of Computer Vision*, Vol. Vol.60, No.2, pp. 91-110, 2004.
- [7] M. Brown, D. G. Lowe. Automatic panoramic image stitching using invariant features. *Journal of Computer Vision*, Vol. Vol. 74, No.1, pp. 59-73, 2007.
- [8] Herbert Bay, Tinne Tuytelaars, Luc Van Gool. Surf: Speeded up robust features. *European Conference on Computer Vision*, pp. 404-417, 2006.
- [9] Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, Vol. Vol. 110 Issue 3, pp. 346-359, 2008.

- [10] Rosten, Edward and Porter, Reid and Drummond, Tom. Faster and Better: A Machine Learning Approach to Corner Detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 32, No. 1, pp. 105–119, January 2010.
- [11] Xin Yang and Kwang-Ting (Tim) Cheng. Accelerating surf detector on mobile devices. In *Proceedings of the 20th ACM International Conference on Multimedia*, MM '12, pp. 569–578, New York, NY, USA, 2012. ACM.
- [12] Stephen J. Thomas, Bruce A. MacDonald, and Karl A. Stol. Real-time robust image feature description and matching. In *Proceedings of the 10th Asian Conference on Computer Vision - Volume Part II*, ACCV'10, pp. 334–345, Berlin, Heidelberg, 2011. Springer-Verlag.
- [13] Jan Herling and Wolfgang Broll. An adaptive training-free feature tracker for mobile phones. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, VRST '10, pp. 35–42, New York, NY, USA, 2010. ACM.
- [14] Noah Snavely, Steven M. Seitz, and Richard Szeliski. Photo tourism: Exploring photo collections in 3d. *ACM Trans. Graph.*, Vol. 25, No. 3, pp. 835–846, July 2006.
- [15] Connelly Barnes, David E. Jacobs, Jason Sanders, Dan B Goldman, Szymon Rusinkiewicz, Adam Finkelstein, and Maneesh Agrawala. Video puppetry: A performative interface for cutout animation. *ACM Trans. Graph.*, Vol. 27, No. 5, pp. 124:1–124:9, December 2008.
- [16] 高木雅成, 藤吉弘巨. Sift 特徴量を用いた交通道路標識認識. 電気学会論文誌, Vol. Vol.129-C, No.5, pp. 824–831, 2009.
- [17] Jain, Anil K. *Fundamentals of Digital Image Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.

- [18] Satoshi Suzuki, Keiichi Abe. Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, Vol. 30, No. 1, pp. 32–46, 1985.
- [19] デジタル画像処理編集委員会, 画像情報教育振興協会. デジタル画像処理 : digital image processing. 画像情報教育振興協会, 2004.
- [20] A. Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society*, Vol. 35, pp. 99–109, 1943.
- [21] Gary Bradski and Adrian Kaehler and 松田 晃一. 詳解 OpenCV : コンピュータビジョンライブラリを使った画像処理・認識. オライリー・ジャパン, 2009.8 2009.
- [22] しましま. Hellinger 距離 - 機械学習の「朱鷺の杜 Wiki」. <http://ibisforest.org/index.php?Hellinger%E8%B7%9D%E9%9B%A2>. 参照:2016.1.19.
- [23] 上綱 秀治. バーコード - cyberlibrarian. <http://www.asahi-net.or.jp/~ax2s-kmtn/bcodes.html>. 参照:2015.12.20.
- [24] International ISBN Agency. ISBN Ranges _ International ISBN Agency. https://www.isbn-international.org/range_file_generation. 参照:2015.1.19.
- [25] Itseez. OpenCV — OpenCV. <http://opencv.org/>. 参照:2015.12.22.
- [26] 奈良先端科学技術大学院大学 OpenCV プログラミングブック制作チーム. OpenCV プログラミングブック. 毎日コミュニケーションズ, 東京, Japan, 2007.9 2007.
- [27] OpenCV2 プログラミングブック制作チーム. OpenCV2 プログラミングブック : OpenCV2.2/2.3 対応. マイナビ, 2011.12 2011.
- [28] 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. 研究報告コンピュータビジョンとイメージメディア (CVIM) , Vol. 2015-CVIM-199, pp. 1 – 6, 2015.