

提出日 2015年2月6日

Summary of Bachelor's Thesis
2015年3月修了卒業論文概要書

Name 氏名	石橋 延高	ID number 学籍番号	1W080043-9
題目 Title (日本語の場合は英文題目も記入)	Supervisor 指導教員 渡辺 裕		印
日本語 Japanese	マンガ画像における手書きオノマトペ抽出		
英語 English	Extraction of Handwritten Onomatopoeia for Comic Images		

概要 Summary

近年電子コミックの普及が著しく進んでいる。さらに元々海外で人気だった日本のマンガコンテンツが積極的に翻訳され海外でも出版されるようになってきている。マンガの電子化、翻訳化の際に吹き出し等を抽出する技術は研究が進んでいるが、マンガ内に描かれるオノマトペ(擬声語)は背景と重なっているなどの理由で抽出が難しく翻訳化の際の大きな障壁となっている。

そこで、本論文ではマンガ画像内における手書きオノマトペ抽出を行うことで、将来的なマンガの電子化、翻訳化の利便性向上を目指した。

今回はオノマトペが一定線幅の線分で構成される点に着目して、モルフォロジー演算と領域分割手法の二種類の手法を用いて抽出を行った。モルフォロジー演算は対象画像と任意形状の構造要素との集合演算により対象の輪郭が滑らかにすることができ、ノイズ除去や文字列抽出などの分野で用いられる手法である。領域分割手法は画像を特定面積のブロックごとに分割、分散値を計算しその値によって画像変換を行う手法である。

実験の結果、モルフォロジー演算では黒字で書かれたオノマトペに関してはある程度抽出ができるが、最も登場頻度の高い白文字に黒枠で書かれたオノマトペの抽出度合いは著しく低かった。領域分割手法では白文字に黒枠で書かれたオノマトペも抽出がうまくいく場合が多かったが、平均線幅よりも細いオノマトペは抽出されず背景画像とともに除去されてしまい、太いオノマトペは内部や一部が欠損してしまい、抽出度合いは低かった。またいずれの手法でもオノマトペと同程度の面積を持つ領域は除去されずに抽出されてしまった。

注意 1. 提出期限/ Deadline of Submission 2015年2月6日(金)17:00 /February 6, 2015
17:00

基幹理工学部情報理工学科

Bachelor's Thesis
卒業論文

Title

論文題目

マンガ画像における手書きオノマトペ抽出

Extraction of Handwritten Onomatopoeia for Comic Images

Student ID 学籍番号	1W080043-9
Name 氏名	Nobutaka Ishibashi 石橋 延高

Supervisor 指導教員	Hiroshi Watanabe 渡辺 裕	印
--------------------	--------------------------	---

2015年3月15日

目次

第1章 序論	1
1.1 研究の背景	1
1.2 研究の目的	2
1.3 論文の構成	2
第2章 画像からの文字抽出手法	3
2.1 はじめに	3
2.2 領域分離法	3
2.3 合焦位置に基づく文字列抽出	3
2.4 SVMによる文字パターン識別	4
2.5 モルフォロジー演算	4
第3章 オノマトペ抽出	7
3.1 はじめに	7
3.2 マンガ画像における手書きオノマトペの特徴	7
3.3 モルフォロジー演算	8
3.3.1 提案手法	8
3.3.2 実験内容	9
3.3.3 実験結果	10
3.4 領域分割法	12
3.4.1 提案手法	12
3.4.2 実験内容	14
3.4.3 実験結果	14
第4章 考察	17
第5章 結論	19
5.1 総括	19
5.2 今後の課題	20
謝辞	21
参考文献	22
図一覧	24

第 1 章 序論

1.1 研究の背景

マンガは日本を代表する娯楽コンテンツであり、日本の出版社の漫画雑誌や単行本が欧州や北米を中心に海外でも出版されるなど世界的にも知られている代表的なコンテンツである。特に近年では経済産業省が主体となったクールジャパン政策[1]など、日本のコンテンツ産業の中心としてマンガを海外に発信していく傾向が強まっており、マンガに対する社会的関心は高まっている。

近年ではスマートフォンやタブレット端末の普及に伴う電子書籍市場の発展が顕著である。その国内市場規模は 2013 年度で 963 億円であり前年度の 729 億円から 200 億円以上増加している[2]。電子書籍市場においてもマンガは重要なコンテンツであり 2013 年度の売上高は 731 億円と全体の 75.9%を占めており、今後電子書籍市場の発展に伴って電子コミックが持つ役割もより重要になっていくと考えられる。また電子書籍の市場規模は国内では書籍全体の 8%程度なのに対し、アメリカでは約 20%程度を占めており国内よりも書籍の電子化が活発である[3]。このような背景から日本のマンガコンテンツが電子化され、さらに海外の言語に翻訳され配信される機会がこれからより増えていくことが予想される。

マンガ画像を構成する要素としてはコマ割り、台詞や心情などが書かれた吹き出し、背景、人物の絵などが存在するが、それらに加えてオノマトペ(onomatopée)と称される擬音語/擬態語もマンガ画像を構成する大きな要素として挙げられる。吹き出しを同定する技術に関しては、ページ内の文字領域を AdaBoost によって特定し、その領域を元に吹き出し候補を検出、また SVM によって吹き出し形状の分類を行う手法が提案されており、この手法により 86%の吹き出しの同定が行えることが報告されている[4]。これに対してオノマトペは文字を表しているものの、吹き出しに書かれた台詞と違い活字印刷されたものではなく手書きで書かれており、絵と一体化するようにデフォルメされ、さらに背景や人物と重なるように描かれているため、活字印刷で書かれた文字に比べて同定抽出が非常に困難である。そのため、日本のマンガが海外で翻訳される際にはオノマトペは翻訳されずに日本語のままであったり、翻訳言語が添字として併記してあったりするなどがほとんどで、オノマトペ自体を翻訳していることは少ない[5]。コスト的な要因も重なってマンガ画像内のオノマトペはマンガを海外で翻訳出版する際の大きな障壁となっているのが現状である。

1.2 研究の目的

本研究では、マンガ画像における手書きオノマトペの抽出を目的とする。これにより研究背景で述べたマンガを海外で翻訳出版する際の効率化、またマンガを電子化する際にオノマトペに視覚効果を付けるなど、電子コミックならではの付加価値をつける際の利便性の向上にも繋がると考えている。

1.3 論文の構成

本論文の構成は以下のとおりである。

第1章は本章であり、本研究の背景、目的について述べる。

第2章では画像からの文字抽出手法を先行研究を交えて述べる。

第3章でははじめにマンガ画像における手書きオノマトペの特徴を述べ、2種類の提案手法を示し、それぞれの実験内容及び実験結果を述べる。

第4章では実験結果を元に考察を述べる。

第5章では結論として本論文の総括と今後の課題について述べる。

第 2 章 画像からの文字抽出手法

2.1 はじめに

この章ではマンガ画像に限らず、背景画像からの文字抽出において用いられる手法についての解説や用語の定義などを行う。

2.2 領域分離法

この手法は 2 値画像と網点画像が混在する原稿について、それらの領域を精度よく分割する手法として我孫子らによって提案された分割濃度差法をもとに改良を加え拡張した領域分離法である[6]。この手法では分離パラメータ設定に柔軟性を加える事で、分離精度の向上、再生画像の品質向上を狙っている。この手法の概要を以下に示す。

- (1)入力画像全体を 4*4 画素を 1 ブロックとする微小なブロックに分割する
- (2)ブロック内の平均濃度を下記計算式で求める。

$$A = \frac{1}{4 \times 4} \sum_{i=1}^4 \sum_{j=1}^4 L_{ij} \quad (L_{ij}: \text{各画素の読み取り値})$$

- (3)A の値に応じて各ブロックを以下の 3 種類に分類する。

- i. 網点ブロック
- ii. 2 値ブロック
- iii. 文字を含む網点領域

(4)iii に分類されたブロックに対して特別な補正を施し領域拡張を行ったうえで、最終的にこの領域から文字を構成する画素を抽出する。

2.3 合焦位置に基づく文字列抽出

この手法[7]は文字領域を携帯電話のカメラ等を用いて撮影することを想定し、カメラの合焦領域に含まれた文字部分を起点として、隣接する他の文字を探索することで文字列抽出を行う。またこの手法を基に、文字の縦方向への移動や高さが不均一な文字列に対応するための手法も提案されている[8]。抽出アルゴリズムを以下に示す。

- (1)合焦領域の色分布の中心色をクラスタリングで求め、文字と背景の候補色とする。

クラスタリングには Mean-Shift[9]を用いる.

(2)文字と背景の候補色と似た画素を合焦一を基準として探索し, 領域成長により探索領域を広げていくことで単一の文字候補領域を抽出する.

(3)抽出した単一文字候補を基準として水平方向へ新たな文字候補を探索する

(4)候補色ごとに求めた文字列パターン候補を, 形態特徴に基づいて文字列と背景に識別する

2.4 SVM による文字パターン識別

SVM(Support Vector Machine)はパターン認識手法の一つであり, 線形関素子を用いて 2 クラスのパターン識別器を構成する手法である[10]. 背景画像からの文字列抽出にも多く用いられている[11][12]. 岡本らの提案手法[13]ではエッジの強度, 色情報などに基づき画像をブロック化クラスタリングを行い文字列パターン候補を得た後, SVM を用いて文字パターンと背景パターンの識別を行っている. その際の幾何学的特徴量として下記の組み合わせを利用している.

- (1) 走査線ごとの相互相関に基づく特徴量
- (2) 黒ラン長を用いた特徴量
- (3) 輪郭線の極大点と極小点を用いた特徴量
- (4) 黒白変化点を用いた特徴量
- (5) 黒画素の密度を用いた特徴量
- (6) 外接最小矩形の縦横比による特徴量

2.5 モルフォロジー演算

モルフォロジー演算(morphological operation)は, 2 値画像やグレー画像中における対象図形の形状を変化させる処理であり, 与えられた画像に対して構造要素を適用して近傍に関して集合演算を行い, 近傍の形を決定する[14]. 構造要素は任意の形状とサイズを持つ 0 と 1 のみで構成される行列で作られる. 構造要素の例を図 1 に示した. 図で示した 5*5 ピクセルのサイズよりもさらに大きくすることや十字やひし形以外にも任意の形状を指定することができる.

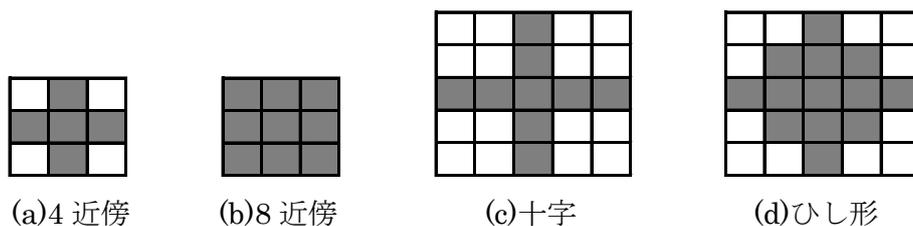
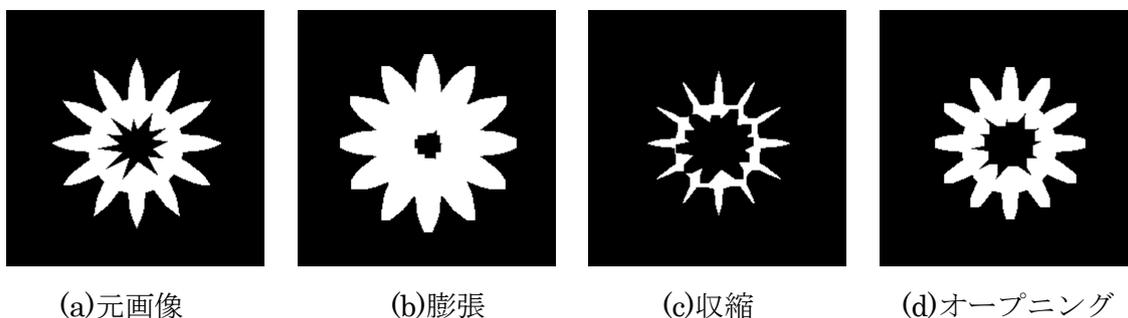


図1 モルフォロジー演算の構造要素例

モルフォロジー演算による処理は膨張(dilation)と収縮(erosion)の2種類に大きく分別される。2値画像の場合、注目画素あるいはその近傍の画素のいずれかが画素値1となる場合に注目画素の画素値を1とする処理を膨張と呼び、注目画素あるいはその近傍の画素のいずれかが画素値0となる場合に注目画素の画素値を0とする処理を収縮と呼ぶ。グレイ画像の場合、注目画素の画素値を、膨張では近傍画素の最大輝度値、収縮では最小輝度値に置き換えることから、それぞれ最大値フィルタ、最小値フィルタと呼ばれる。

膨張・収縮処理を単独で行うことは少なく、それぞれの処理を複数回繰り返すことで実際の処理を行う。オープニング(opening)は、収縮を n 回繰り返したのち、膨張を n 回繰り返す処理である。それに対してクロージング(closing)は膨張を n 回繰り返したのち、収縮を n 回繰り返す処理である。オープニング処理により、小さなとげや単一画素のスパイク雑音のような対象外のものを取り除く効果があり、対象の輪郭が滑らかになる。クロージング処理により、小さい穴やギャップのような対象内のものを埋める効果があり、対象の輪郭が滑らかになる。さらに、膨張した画像から収縮した画像を減算してエッジを検出するグラジエント(gradient)、元画像からオープニングした画像を減算するトップハット(Top-Hat)、クロージングした画像から元画像を減算するブラックハット(Black-Hat)などといった処理も用いられる。2値画像に対してそれぞれのモルフォロジー演算処理を行った画像例を図2に示す。

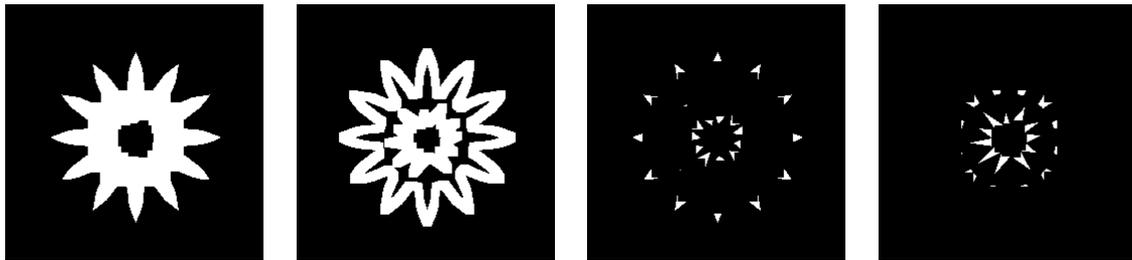


(a)元画像

(b)膨張

(c)収縮

(d)オープニング



(e)クロージング (f)グラジエント (g)トップハット (h)ブラックハット

図2 2値画像に対するモルフォロジー演算の処理例

モルフォロジー演算は、対象物の形状をそのまま扱うことができるので、ノイズ処理、形状抽出処理に有効であり、画像から文字を抽出する手法も多く提案されている。特定の長さの線分に着目して、実文書画像から文字や符号を抽出する手法[15]や、劣化の激しい古地図画像を対象にノイズ除去、エッジ抽出を行い、エッジ情報をブロックごとに集計し文字と非文字ブロックを判別して文字領域を抽出する手法[16]、文字領域の線幅に注目し、平均線幅以下の領域をモルフォロジー演算により排除することで配送伝票画像からの重ね書き文字の抽出を行う手法[17]などがあげられる。

第3章 オノマトペ抽出

3.1 はじめに

この章では、まずマンガ画像における手書きオノマトペの特徴について述べ、実際に抽出を行う際に用いた3つの手法についてその解説を行い、それぞれの手法で行った実験内容、実験結果を順番に説明する。

3.2 マンガ画像における手書きオノマトペの特徴

マンガ作品4作品、単行本77冊分、計14754ページ分のうち、手書きオノマトペが含まれるページを約1100ページを抽出して、マンガ画像における手書きオノマトペにどのような特徴が含まれるかを調査を行った。その結果、以下の要素があげられることがわかった。

(1)同一コマ内において書体は一定

手書きオノマトペはマンガ背景と一体化するように、また演出上の理由から大きくデフォルメされて書かれることが多いが、同一コマ内においては書体、文字の幅、大きさが一定である。

(2)カタカナ文字

ほとんどのオノマトペがカタカナもしくはひらがなで書かれている。割合としてはカタカナのほうが圧倒的に多い。かな文字以外の、漢字やアラビア数字、アルファベットが使われることは稀である。

(3)最前面に配置

マンガ画像の構成要素として、背景、人物、吹き出し、などがあげられるが、オノマトペがそれらと重なって書かれる場合、最前面に位置することがほとんどであり、オノマトペが他の要素に重なり一部が見えなくなっていることは稀である。

(4)文字色と反対色の枠

文字自体が白色の場合は黒色、黒色の場合は白色の枠で囲まれており、文字が目立つような工夫がなされている。白色(黒色)の代わりに色の薄い(濃い)スクリーントーンが用いられることもある。具体例を画像で図3に示す。また、使われる場面としては、ページ全体の輝度が高く黒要素が多い場合は白文字が、逆に白要素が多い場合は黒文字が使われる傾向が多い。全体としては図3(a)のような黒枠に白文字であることが多く、その割合は1100ページ中700ページ以上と約64%を占めている。

以上の点から考慮するとマンガ画像における手書きオノマトペの画像的特徴としては、カタカナで書かれるため直線成分で構成されることが多く、一定幅の線分で囲まれた領域であると特徴づけることができる。



(a)黒枠に白字[18]

(b)白枠に黒字[19]

図3 手書きオノマトペの例

3.3 モルフォロジー演算

3.3.1 提案手法

オノマトペ領域がマンガ画像全体において最前面に存在し、オノマトペ以外の背景や人物、吹き出しなどよりも太字で書かれていることに着目し、オノマトペ以外の細かい要素を除去することを目標として、マンガ画像にモルフォロジー演算を適用する。処理プログラムにはOpenCVのモルフォロジー演算関数を用い、1枚の画像に対して膨張、収縮、オープニング、クロージング、グラジエント、トップハット、ブラックハットの全7種類の処理を行う。その際に用いる構造要素は、 3×3 矩形形状、 9×9 矩形形状、 9×9

十字形状, 9*9 ひし形形状の 4 種類を使用する. 4 種類の構造要素を図 4 に示す. また, オープニング, クロージングともに収縮・膨張の繰り返し回数は 1 回だけに設定する.

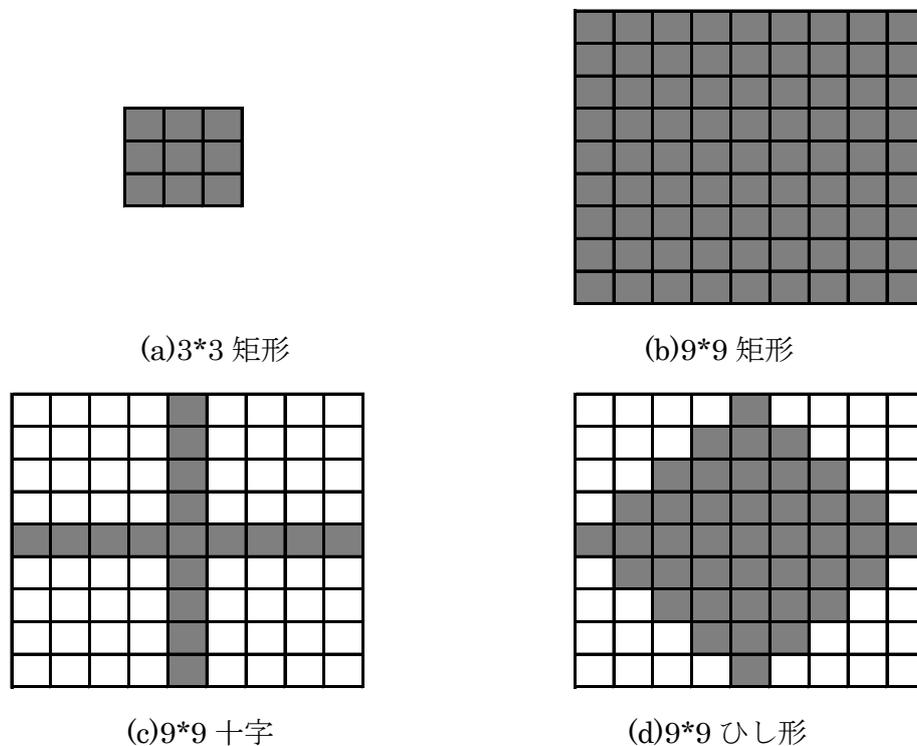


図 4 使用した 4 種類の構造要素

3.3.2 実験内容

使用するマンガ画像はすべて 1 ページのサイズが 644*1033 ピクセルのものを使用した. 特徴調査の際に用いた約 1100 ページ分のオノマトペが含まれる画像のうち, 以下の条件を満たす画像を選んだ.

- (1) 1 ページ内にオノマトペが全て収まっている
- (2) オノマトペが最前面に配置されており隠れている部分がない
- (3) 同一画像内に大きな書体の変化がない

条件を満たさない画像例を図 5 に示した. 画像例のうち(a)は「ド」という文字の両端がページの外に出ている. (b)は「ダンッ」という文字がキャラクターの背面に位置し文字の位置が隠れてしまっている. (c)は「ピーッ」という黒枠に白文字と「ドオッ」という白枠に黒文字の 2 種類の書体の違う文字が含まれている.

これらの 3 つの条件を満たす画像計 200 枚を抽出し, 3.3.1 項で述べたモルフォロジー演算処理を 1 枚の画像に対して 4 種類の構造要素をそれぞれ適用し, 処理後の画像を元画像と比較してオノマトペの抽出度合いを確認する.



図 5 条件を満たさない画像例

3.3.3 実験結果

モルフォロジー演算処理後の画像のオノマトペ抽出度合いは用いた画像によってかなりバラつきがあったが、大別すると以下のように区分された。

- (1) クロージング処理によりオノマトペがはっきりと抽出された
- (2) いずれの処理でもオノマトペが抽出されなかった

(1)に該当する画像はすべてオノマトペ部分が黒字で書かれたものであり、オノマトペ部分以外の細かい線や人物画像などがクロージング処理によって除去され、比較的はっきりと抽出できている。(1)に該当する元画像の例を図 6 に示す。この場合でもキャラクターの髪の毛や背景のベタ塗り部分など、オノマトペ部分と同程度、もしくはそれ以上の面積を持つ黒色部分に関しては除去されずに残ったままになってしまうため、抽出度合いが高かったといえるのは、オノマトペ部分が黒字でかつ背景部分に黒領域が少ない画像のみであった。また、4種類の構造要素をそれぞれ用いたが最も抽出度合いが高かったのはいずれの画像においても 9*9 矩形形状のものであった。

(2)に該当する画像はすべてオノマトペ部分が白字またはスクリーントーンで書かれたものであり、クロージング処理では背景とともにオノマトペ部分も除去されてしまい抽出がまったくされなかった。クロージング以外の処理についても同様に、オノマトペ部分と他の要素を区別する画像的特徴がなく、抽出度合いは著しく低かった。(2)に該当する元画像の例を図 7 に示す。



[23]



[24]

図 6 3.3.3 項(1)に該当する画像例



[25]



[26]

図 7 3.3.3 項(2)に該当する画像例

3.4 領域分割法

3.4.1 提案手法

前節のモルフォロジー演算ではオノマトペの抽出度合いが著しく低かった"黒枠に白字"のオノマトペを抽出することに着目する。この"黒枠に白字"タイプのオノマトペは調査対象のマンガ画像の中でも約 64%を占め、最も登場頻度が高く抽出対象としては最も重要であると考えられる。

この提案手法はオノマトペの線幅が一定であることに着目し、調査対象であるオノマトペの黒枠の平均線幅を求め、一定線幅以下の黒領域と一定線幅以上の黒領域を除去することでオノマトペ領域のみを抽出することを目標としている。画像全体を小さなブロックごとに分割し、ブロックごとの輝度の分散値を計算し、分散値によって輝度変換を行う。今回は 2 パターンの手法を用いたため、それぞれ順番に述べる。

(a)パターン A

まず、画像全体を $n \times n$ ピクセルの小さなブロックに分割する。ブロック内に含まれる $n \times n$ 個の各ピクセルについてその輝度値を 0 から 255 の 256 段階で求める。ブロックごとに輝度の分散値を計算し、分散値が一定値以上の場合はブロック内すべてのピクセルの輝度値を 255 に変換し、ブロック全体を白にする。分散値が一定値以下の場合も同様にブロック内すべてのピクセルの輝度値を変換しブロック全体を白に変更する。これによりある分散値が一定の範囲内であるブロックのみが元画像の輝度値を保つ。

(b)パターン B

まず、一定の輝度値を基準値とし、基準値以上のピクセルは輝度値を 255 に変換、基準値以下のピクセルは輝度値を 0 に変換し、画像全体を 2 値画像に変換する。変換した 2 値画像を $n \times n$ ピクセルの小さな領域に分割し、パターン A と同様に輝度の分散値を計算、一定値以上の場合はブロック全体を白に変換する。その後、改めて画像全体を $m \times m (n < m)$ ピクセルの先程よりも大きなブロックに分割し、ブロックごとに輝度の分散値を計算を行い、一定値以下の場合そのブロック全体を白に変換する。1 段階目の小さなブロックでの変換でオノマトペ部分よりも細かい領域を、2 段階目の変換でオノマトペ部分よりも広い領域の除去を目指している。この 2 段階の変換処理を画像全体すべてのブロックに適用する。

2 パターンの手法の流れを図 8 に示す。

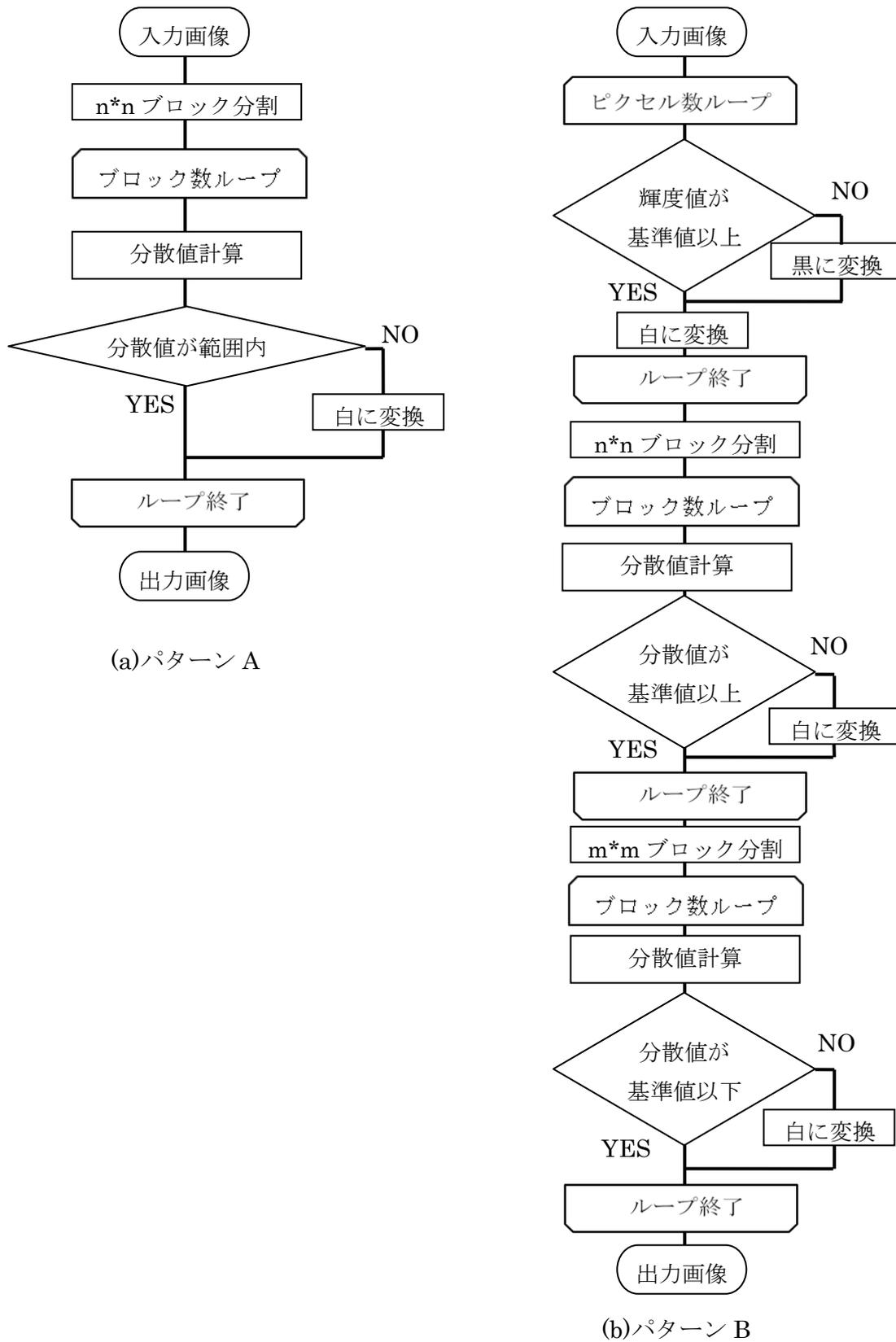


図 8 領域分割手法の流れ

3.4.2 実験内容

2 パターンの手法ともに、前節のモルフォロジー演算と同様に、633*1033 ピクセルのオノマトペが含まれるマンガ画像を入力画像として用いる。どちらの手法も以下の条件を満たす画像を用いる。

- (1) 1 ページ内のオノマトペがすべて黒枠に白字である
- (2) 1 ページ内にオノマトペが全て収まっている
- (3) オノマトペが最前面に配置されており隠れている部分がない

条件を満たす画像の具体例としては図 7 の 2 枚などが挙げられる。条件を満たす画像 200 枚を対象に 2 パターンの手法で領域分割を行った。オノマトペの黒枠の太さと輝度を測定したところ、太さは平均 2~4 ピクセル、輝度は 45~55 に収まるものがほとんどであった。この値を元に、パターン A においては 1 ブロックの大きさを 5*5 ピクセルと設定し、分散値の上限は 10, 100, 500, 1000 の 4 パターン、下限は 1, 5, 10 の 3 パターンをそれぞれ変動させる。パターン B においては輝度値の基準値を 60, ブロックの大きさを 1 段階目を 2*2 ピクセル, 2 段階目を 5*5 ピクセルに設定し、分散値の基準値を 0, 10, 100 の 3 パターンをそれぞれ変動させて処理を行う。1 枚の画像に対して 2 パターンの手法を分散値の基準値を変動させて処理を行い、変換後の画像と元画像を比較してオノマトペの抽出度合いを確認する。

3.4.3 実験結果

2 パターンの手法の実験結果を順番に述べる。

パターン A に関しては、分散値の上限下限の変動によらずオノマトペの抽出度は著しく低かった。分散値の上限を 10, 100 などの低い値に設定した場合、背景部分だけではなく抽出対象のオノマトペ部分も同様に除去されてしまい、面積が広い一部の黒領域を除いて出力画像全体がすべて白くなってしまいうページが 200 枚中ほとんどを占めた。上限値を 500, 1000 などの大きい値に設定した場合は、除去対象の背景部分も抽出対象のオノマトペ部分とともに細切れに残ってしまい、抽出度合いとしては非常に低いものが多くなった。分散値の下限に関しては値を変動させてもほとんど変化がなく、1, 5, 10 の 3 パターンどの値に設定しても、面積の広い黒領域の変化はほぼ一定で、オノマトペの抽出度合いにはほとんど差は出なかった。パターン A においてオノマトペの抽出度合いが比較的高かったのは図 9 に示した 2 枚の画像のように背景の輝度値が低く、全体的に黒領域が目立つページのみにとどまった。



[27]

[28]

図9 パターンAにおいてオノマトペ抽出度合いが高い画像例

パターンBに関しては、200枚のうちほとんどのページにおいてパターンAと比較すると抽出度合いが高かった。特にパターンAで抽出度合いが著しく低かった画像全体の輝度値が高く白領域が多いページに関してもある程度の抽出度合いを確保することができた。基準とした分散値に関しては0, 10, 100の3パターンどの値に設定しても抽出度合いに大きな変化は見られなかったが、100に設定した場合、抽出対象のオノマトペの黒枠の一部が除去される場合があり、0に設定した場合が最も抽出度合いが高かったといえる。抽出度合いが著しく低かった画像例を図10にあげる。今回は分割するブロックの大きさを、オノマトペ黒枠の太さの平均値を参考に2*2ピクセルと5*5ピクセルに設定したため、平均から外れた細かいものに関しては背景ともに除去されてしまい、太いものに関してはオノマトペ黒枠内部や一部が除去されかけてしまうことが多かった。また、いずれの場合においても画像を2値化した際の黒領域と白領域の境界線部分や輪郭線に関しては完全に除去することができず、結果として元画像の輝度値が低い、黒領域が目立つページのオノマトペ抽出度合いが低くなってしまった。



(a) 黒枠が細い[29]

(b) 黒枠が太い[30]

図 10 パターン B においてオノマトペ抽出度合いが低い画像例

第4章 考察

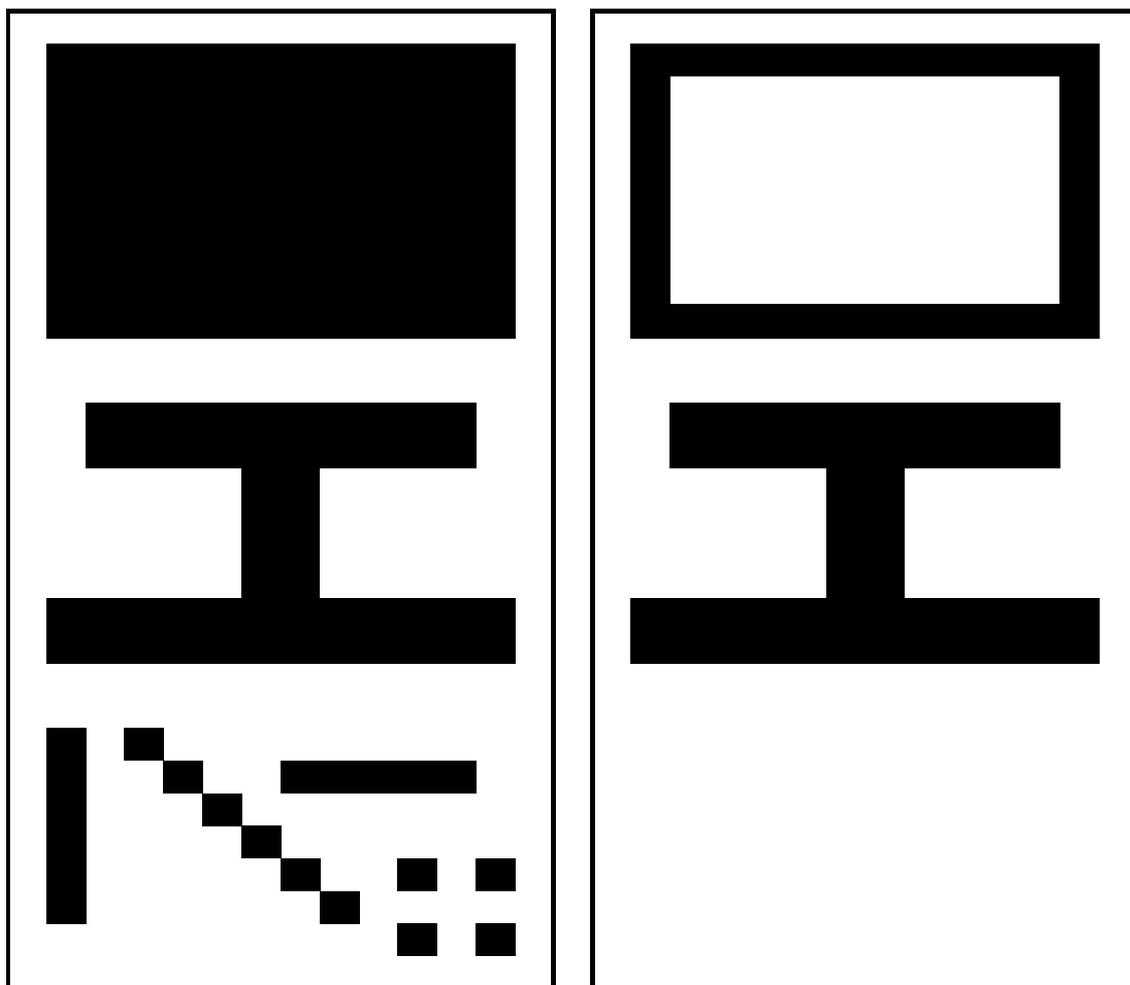
第3章で述べた実験ではモルフォロジー演算と領域分割法の2種類の手法でマンガ画像における手書きオノマトペの抽出を目指した。

モルフォロジー演算においてオノマトペ抽出度合いが高かった画像は画像特徴的にオノマトペと背景の輝度がはっきりと区別されているものにとどまった。第2章で述べたモルフォロジー演算を利用した文字抽出の先行研究では線分の長さであったり、平均線幅の差異などに着目して抽出を行っていたが、マンガ画像内のオノマトペでは文字自体がデフォルメされて書かれた"絵"であり、背景に書かれた"絵"と比較して視覚的に判断できる差異をモルフォロジー演算のみで抽出することは困難であった。また、情景画像からの文字抽出手法[13]では、文字列パターンと背景パターンをあらかじめ学習させたSVMを用いてその識別を行っている。この識別に用いた特徴量は2.4節に述べた組み合わせがあげられるが、マンガ画像内のオノマトペは書体が一定でないため情景画像からの文字抽出と同様の手法では識別が困難であると考えられる。実験で用いた構造要素のうち、最も抽出度合いが高かったのは9*9矩形形状であったが、これは入力画像のサイズに依存するもので、用いる画像サイズ、ひいては抽出対象となるオノマトペのサイズによって適した構造要素のサイズも変化すると考えられる。形状に関しては、今回の実験の場合、抽出対象が垂直方向と平行方向どちらへも依存度が低かったため、十字やひし形よりも矩形形状が最も適していた。

領域分割法では、モルフォロジー演算では抽出が困難であった"黒枠に白字"のオノマトペについて抽出を目指した。パターンAの手法では分散値をどのように変更しても、オノマトペ部分と背景部分の黒領域を明確に分割することはできなかった。分散値が低いブロックは黒領域または白領域が広い領域、分散値が高いブロックはさまざまな輝度で細かい画像が含まれている領域として、それぞれでの大小でオノマトペ部分よりも広い領域と狭い領域を区別しようとしたが、結果として分散値の大小のみではオノマトペとそれ以外の背景部分を区別して抽出することはできなかった。

パターンBの手法では、まず画像を2値化しオノマトペを含む黒領域のみを残し、除去対象の背景部分を大きく減らしている。これにより、後の分散値計算で"一定面積の黒領域"を抽出しやすくしている。さらにブロックごとの輝度分散値計算を2段階に分け、狭い領域を除去してから広い領域を除去することで、指定の面積の領域を確実に抽出することができた。しかし、この手法の問題点は、完全に黒領域の面積のみを基準に判断して画素値の変更を行っているため、面積だけ一致していれば形状に依らずに抽出してしまう点がまずあげられる。さらに、第1段階では小さいブロックで分散値計算を行い輝度変換するため、面積が広い領域に関しては境界線が線分の形状で残り、オノ

マトペ黒枠部分の面積と近似してしまうため、第2段階の大きいブロックによる輝度変換においても除去されずに残ってしまうという点があげられる。これを簡易的に表したものが図11である。抽出対象は中央「エ」と書かれた領域であり、下部の線分と点が第1段階、上部の長方形部分が第2段階で除去する対象である。



(a)元画像

(b)出力画像

図11 パターンBにおける画像変換例

第5章 結論

5.1 総括

第1章では研究の背景及び、研究の目的、論文の構成を述べた。

日本を代表する娯楽コンテンツであるマンガは作者や出版社自身のみならず、政府主体となって海外へ発信していく傾向が強まるなど、社会的な関心が近年非常に高まっている。その中でマンガに含まれる要素の一つである手書きオノマトペがマンガの電子書籍化、さらに海外言語に翻訳して出版される際の障壁となっていることを述べ、マンガ画像における手書きオノマトペを抽出することで、将来的なマンガ翻訳、電子化する際の効率化につなげることを本研究の目的とした。

第2章ではマンガ画像に限らず、画像から文字を抽出する手法を、領域分離法、合焦位置に基づく文字列抽出、SVMによる文字パターン識別、モルフォロジー演算の順で本研究で用いる用語の解説も交えながら述べた。

第3章ではマンガ作品を調査して得られた手書きオノマトペの特徴を述べ、提案手法として「モルフォロジー演算」と「領域分割法」をあげ、それぞれの手法内容、実験内容、実験結果を順番に述べた。

モルフォロジー演算では、オノマトペが黒字で書かれ、背景の輝度が高い画像に関してはクロージング処理によりある程度の抽出度合いを確保することができたが、最も登場頻度の多かった黒枠に白字のオノマトペは抽出がうまくいかなかった。

領域分離法では、黒枠に白字のオノマトペに対象領域を絞り2パターンの手法で抽出を行った。微小ブロック内の輝度分散値の大小によって画像の輝度を変換、除去する手法では予想に反して抽出対象のオノマトペも除去されてしまい、抽出がうまくいかなかった。次に、画像を2値化したのち、微小ブロックとそれより大きいブロックの2段階に分割しそれぞれの輝度分散値の大小によって輝度変換を行う手法では、比較的高い抽出度合いでオノマトペを抽出できた。しかし、平均から大きく外れた太さのオノマトペはうまく抽出できない点や、黒領域と白領域の境界線部分に関して除去がうまくいかない点が問題点として残った。

第4章では実験結果についての考察をおこなった。主にそれぞれの提案手法について、抽出がうまく行かなかった場合の要因について、図を交えながら述べた。

第5章は本章であり、本論文の総括と今後の課題について述べる。

5.2 今後の課題

本論文で提案した手法はいずれもオノマトペとそれ以外の部分の面積に着目し、背景部分の除去及びオノマトペ部分の抽出をおこなった。そのため、特定面積の領域以外を抽出できないことが課題としてあげられる。今回用いた分割領域手法では事前に調査した平均サイズをもとに、基準値となるブロックサイズを決定し、その範囲内に収まる領域のみを抽出している。そのため、事前にオノマトペのサイズが分かっている限り抽出は困難である。今後は画像全体の輝度値などからオノマトペのサイズを割り出し基準値を決定するなどして、オノマトペがいかなるサイズであっても抽出度合いを一定に保てるような手法を検討したい。

また、提案手法では面積が近似であればオノマトペと全く関係のない境界線部分も抽出されてしまう。そのため、オノマトペがカタカナで書かれ、直線成分で構成されることが多い点など、面積だけではなく形状に着目するなど、別手法も用いてより抽出度合いを高めていくことが今後の課題である。

謝辞

本研究の機会及び実験環境を与えて下さり、個人的な都合で研究が遅れてしまった際にも貴重な時間を割いてご指導いただいた渡辺裕教授に心から感謝いたします。

並びに、OpenCV プログラミングの基礎から始まり、日々研究の進め方について助言、ご意見を頂き、研究に行き詰まった際も熱心に指導していただいた石井大祐助手に心から感謝いたします。

日頃のゼミで様々な提案、アドバイスを下さった渡辺研究室の皆様に感謝いたします。

最後に、私をここまで育てて下さり、あらゆる面で多大な心配をおかけしました家族にあらためて深く感謝いたします。

2015年2月6日

参考文献

- [1]経済産業省クールジャパン/クリエイティブ産業, "平成 27 年 1 月「クールジャパン政策について」". http://www.meti.go.jp/policy/mono_info_service/mono/creative/150119CjseisakunituiteJanuary.pdf, 2015/1/23 参照.
- [2]デジタルコンテンツ協会(編), "デジタルコンテンツ白書 2014", 一般財団法人デジタルコンテンツ協会(DCAJ), 2014.
- [3]インプレス総合研究所(編), "電子書籍ビジネス調査報告書 2014", インプレス総合研究所, 2014.
- [4]田中孝昌, 外山史, 宮道壽一, 東海林健二, "マンガ画像の吹き出し検出と分類", 映像情報メディア学会誌, VOL.64, No.12, pp.1933-1939, 2010.
- [5]陳焯, 白水菜々重, 松下光範, "中国人を対象とした日本語コミックにおけるオノマトペの理解に関する調査", 2013 年度人工知能学会, 2013.
- [6]安孫子収, 白川智昭, 北島秀夫, 小川吉彦, "網点上の文字を考慮した文字・網点画像高速分離法", 北海道大学工学部研究報告, VOL.153, pp.51-56, 1990.
- [7]B.E.kim, S.Lee, J.Kim, "Scene text extraction using focus of mobile camera", 10th Inter. Conf. on Document Analysis and Recognition(ICDAR '09), pp166-170, 2009.
- [8]王濟凱, 馬国鋒, 白井啓一郎, 岡本正行, "合焦位置を起点とした文字列抽出の一検討", 電子情報通信学会, VOL.110, pp287-292, 2011.
- [9]D.Comaniciu, P.Meer, "Mean shift:A robust approach toward feature space analysis", IEEE Trans. on Pattern Analysis and Machine Intelligence,VOL.24., o. 5,pp.603-619,2002.
- [10]栗田多喜夫, "サポートベクターマシン入門", <http://home.hiroshima-u.ac.jp/tkuri/ta/lecture/svm.pdf>, 2015/1/26 参照.
- [11]山口拓真, 丸山稔, "階層型識別器を用いた情景画像からの文字抽出手法", 信学論(D-11), VOL.J88-D-11, No.6, pp.1047-1055, 2005.
- [12]芦田和毅, 永井弘樹, 岡本正行, 宮尾秀俊, 山本博章, "情景画像からの文字抽出", 信学論(D-11), VOL.J88-D-11, No.9, pp.1817-1824, 2005.
- [13]若林昌格, 白井啓一郎, 岡本正行, 山本博章"カラー情景画像からの文字抽出高精度化の一検討"電子情報通信学会技術研究報告 PRMU,パターン認識・メディア理解 107 (491),pp149-154,2008.
- [14]白鳥則郎, "画像処理", 共立出版, pp78-80, 2014.
- [15]童衛青, 辰巳昭治, 志水 英二, "モルフォロジー演算による線図形画像から特定の

長さを持つ線図形の抽出", 電気学会論文誌. C, 電子・情報・システム部門誌, VOL.118, pp1536-1537, 1998.

[16]若間俊旭, 芝公仁, 岡田至弘, "モルフォロジー演算を用いた古地図画像からの文字領域抽出法", 人文科学とコンピュータシンポジウム 2009 論文集, pp35-40, 2009.

[17]松尾賢一, 上田勝彦, 梅田三千雄"配送伝票画像からの重ね書き文字領域の分離抽出"電気学会論文誌. C, 電子・情報・システム部門誌, Vol.118, No.12, pp1730-1738, 1998.

[18]井上雄彦, "スラムダンク", 集英社, Vol.1, pp132, 1991.

[19]井上雄彦, "スラムダンク", 集英社, Vol.5, pp108, 1991.

[20]井上雄彦, "スラムダンク", 集英社, Vol.24, pp33, 1995.

[21]井上雄彦, "スラムダンク", 集英社, Vol.25, pp140, 1995.

[22]井上雄彦, "スラムダンク", 集英社, Vol.31, pp21, 1996.

[23]井上雄彦, "スラムダンク", 集英社, Vol.29, pp33, 1996.

[24]森田まさのり, "ROOKIES", 集英社, Vol.24, pp15, 2003.

[25]井上雄彦, "スラムダンク", 集英社, Vol.5, pp141, 1991.

[26]羽海野チカ, "ハチミツとクローバー", 集英社, Vol.1, pp24, 2002.

[27]大場つぐみ, 小畑健, "DEATH NOTE", 集英社, Vol.1, pp38, 2004.

[28]森田まさのり, "ROOKIES", 集英社, Vol.16, pp22, 2001.

[29]大場つぐみ, 小畑健, "DEATH NOTE", 集英社, Vol.12, pp82, 2006.

[30]井上雄彦, "スラムダンク", 集英社, Vol.5, pp57, 1991.

図一覧

図 1	モルフォロジー演算の構造要素例	5
図 2	2 値画像に対するモルフォロジー演算の処理例	5
図 3	手書きオノマトペの例	8
図 4	使用した 4 種類の構造要素	9
図 5	条件を満たさない画像例	10
図 6	3.3.3 項(1)に該当する画像例	11
図 7	3.3.3 項(2)に該当する画像例	11
図 8	領域分割手法の流れ	13
図 9	パターン A においてオノマトペ抽出度合いが高い画像例	15
図 10	パターン B においてオノマトペ抽出度合いが低い画像例	16
図 11	パターン B における画像変換例	18