

卒業論文概要書

Summary of Bachelor's Thesis

Date of submission: (2/3/2015)

専攻名 (専門分野) Department	情報理工	氏名 Name	長尾一輝	指導 教員 Advisor	渡辺 裕 印 Seal
研究指導名 Research guidance	オーディオビジュアル情報処理研究	学籍番号 Student ID number	CD 1W120371-8		
研究題目 Title	Bag-of-Visual Words を用いたコミックにおける主要キャラクター同定 Identification of Major Character in Comics using Bag-of-Visual Words				

1. まえがき

近年、日本の電子書籍市場において電子コミックの需要が高まっている。この需要の高まりは今後も続くと推測できるため、コミックの電子化が次々と起こると考えられる。また、電子コミックは紙媒体と異なりメタデータを付与することができる。メタデータには、コミックのタイトルや著者名、キャラクター情報などがある。このメタデータを利用することで新しいインデクシングが可能になる。そこで、本論文では主要キャラクターを事前知識なしで自動的に同定する手法を提案する。

2. コミックにおける主要キャラクター同定

本論文では、汎用識別器[1]を利用しキャラクター顔画像をコミック画像から抽出することで、SURF 特徴[2]と Bag-of-Visual Words[3]手法を用いて主要キャラクターを事前知識なしで自動的に同定する手法を提案する。コミックにおいて主要キャラクターとは、「ストーリーに大きく関わり、登場回数の多いキャラクター」であると仮定した。従って、各キャラクター顔画像を正確にクラス分類できると仮定すると、主要キャラクターを中心としたクラスには他の非主要キャラクターに比べて、より多くの顔画像が集まるはずである。そこで、本論文ではこのクラスを取得し、クラス所属度 1 位のキャラクターを抽出することで主要キャラクターがどの登場キャラクターなのかを同定する。

3. 研究成果

本研究では、使用するコミックは計 10 冊であり、それぞれのコミックにおける主要キャラクターを予め仮定し、そのキャラクター顔画像が抽出されたかどうかで提案手法の評価を行った。使用したキャラクター顔画像は、100*100 サイズと 150*150 サイズの 2 種類とし、各サイズにおける提案手法の精度を確認した。実験の結果、150*150 サイズの場合には、88%の精度で主要キャラクターを同定できた。従って、使用したコミックに関しては提案手法で主要キャラクターを同定可能であることを確認できた。また、提案手法では使用するキャラクター顔画像サイズを大きくすることで主要キャラクター同定精度の向上が可能であった。これより、提案手法でコミックにおける主要キャラクターの同定が可能であることを示せた。

4. まとめ

本研究では、汎用識別器と SURF 特徴、Bag-of-Visual Words 手法を使いコミックにおける主要キャラクター画像の同定を行った。実験の結果、使用したコミックにおいては提案手法で主要キャラクターの同定が可能であることを示せた。しかし、現状では提案手法において設定するパラメタ値の最適化が行えていない。従って、最適なパラメタ値を検討することで更なる精度向上が臨めると推測できる。

参考文献

- [1] Animeface, https://github.com/nagadomi/lbpcascade_animeface, 参照 Jan.16, 2016.
- [2] H.Bay, A.Ess, T.Tuytelaars, L.V.Gool, Speed-Up Robust Features (SURF), Computer Vision and Image Understanding 110(2008), pp.346-359, 2008.
- [3] G.F.Pineda, H.Koga, T.Watanabe, Scalable Object Discovery : A Hash-Based Approach to Clustering Co-occurring Visual Words, IEICE Transactions, Vol.E94-D, No.10, pp.2024-2035, 2011.

2015 年度 卒業論文

Bag-of-Visual Words を用いた
コミックにおける主要キャラクター同定
Identification of Major Character in Comics
Using Bag-of-Visual Words

指導教員 渡辺 裕 教授

早稲田大学 基幹理工学部

情報理工学科

1W120371-8

長尾一輝

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 主要キャラクターについて	3
2.1 まえがき	3
2.2 主要キャラクターの定義	3
2.3 本研究における主要キャラクターの分類	4
2.4 主要キャラクター登場回数比較実験	6
2.4.1 実験内容	6
2.4.2 実験結果	7
2.5 条件「事前知識なし」の主要キャラクター同定について	11
第3章 キャラクター顔認識手法	12
3.1 まえがき	12
3.2 キャラクター顔認識既存手法	12
3.2.1 従来のキャラクター顔画像認識について	12
3.2.2 LBP (Local Binary Pattern) 画像特徴	13
3.2.3 AnimeFace 識別器	15
3.3 Speeded Up Robust Features (SURF)	15
3.3.1 SURF の概要	15
3.3.2 SURF における特徴点計算	16
第4章 Bag-of-Visual Words	18
4.1 まえがき	18
4.2 局所特徴による画像認識	18
4.3 Bag-of-Visual Words について	19
4.4 Bag-of-Visual Words を用いたキャラクター顔ヒストグラム	21

第5章 コミックにおける主要キャラクターの同定手法	22
5.2 提案手法について	22
5.2.1 提案手法概要	22
5.2.2 提案手法手順	23
5.3 主要キャラクター同定実験	27
5.3.1 実験内容	27
5.3.2 評価方法	28
5.3.3 実験結果	32
5.4 考察	33
第6章 結論	35
6.1 総括	35
6.2 課題	35
謝辞	36
参考文献	37
図一覧	39
表一覧	40

第1章

序論

1.1 研究の背景

近年，スマートフォンやタブレット等の小さなネットワーク接続可能な移動体通信端末が普及している．通信端末の中でも，スマートフォンは2009年ごろより本格的に世界的に普及し，スマートフォンの販売台数は2009年から2011年にかけて，2.7倍と大きく拡大した[1]．この変化に伴い，日常生活の隙間時間にこれらの端末を使用して過ごす機会が増えた．その結果，人々の生活や市場が大きく変化した．特にこの変化に影響を受けたものは，これまで紙媒体を用いたサービスを中心に展開していた書籍市場である．

端末を使用することで紙媒体をデータとして端末に保存し，各々のタイミングで自由に読むことが可能になる．その為，紙媒体を使用する機会が減り，紙媒体での書籍市場は縮小し，電子書籍市場が大きく拡大した．

特に日本の市場においてこの変化が著しく見られ，紙媒体での書籍市場規模は，2008年から2013年にかけて19.7%減少した[2]．一方，電子書籍市場は2012年ごろから急速に拡大し，2019年ごろには市場規模は3400億円に到達すると予測されている[3]．また日本の電子書籍市場の特徴として，電子コミックが電子書籍売上の多くを占める特徴が挙げられる．例えば，2015年1月に講談社は，コミック雑誌を全て電子書籍化し，紙と電子版の同時配信を発表した[4]．

この例の様に，電子書籍需要の高まりと共に紙媒体であったコミックの電子化が次々と起こると考えられる．今後は電子コミックが日本の電子書籍コンテンツの中心になるとみられる．

本論文では，コミック電子化における作業効率の向上や新しいインデクシングを可能にすることを旨とし，メタデータである主要キャラクター情報を自動的に抽出する手法を提案する．

1.2 研究の目的

電子書籍におけるコミックでは、紙媒体と異なりメタデータを付与することで、コミックにおける利用者の要求に対して、様々な手段を用いて解決可能になる。例えば、コミックのタイトルや著者名、ページ数をコミックデータと共に付与することで、効率的な探索を可能とする研究が行われている[5]。他にもコミックにおけるキャラクター情報やジャンル、タイプを付与させることで、利用者や管理者が、コミックを読む前から多少の理解を得られる。

そこで本研究では、コミックにおける主要キャラクター情報に注目し、自動的に主要キャラクター情報を抽出する。この研究成果により、利用価値の高いメタデータである主要キャラクター情報が抽出可能なため、コミック電子化の流れにおいて、自動要約やキャラクター検索等の新しいコミックサービスの創出に利用可能になると考えている。

また、本論文では事前知識である対象コミックにおける登場キャラクター情報を必要としない手法を提案する。これは、コミック電子化の際に手動過程があれば効率性が失われるためである。今後のコミック電子化に提案手法を推進するためには、キャラクター情報などの事前知識なしで自動的に主要キャラクターの同定を実行する必要がある。

1.3 論文の構成

本論文の構成と概要は以下のとおりである。

第1章は本章であり、本研究の背景と目的、構成について述べる。

第2章では、本研究における主要キャラクターの定義を行い、条件である対象コミックにおける登場キャラクター情報の事前知識なしで研究することについて述べる。

第3章では、コミックにおけるキャラクター認識の既存手法と本研究で使用した SURF 特徴について述べる。

題4章では、本研究で使用する Bag-of-Visual Words について述べる。

題5章では、提案手法と実験結果について述べる。

題6章では、本研究における結論と今後の課題について述べる。

第2章

主要キャラクターについて

2.1 まえがき

本章では、本主要キャラクターの定義を行い、本研究における主要キャラクターについての概要および本研究の条件である対象コミックにおける登場キャラクター情報の事前知識なしで研究することについて述べる。

2.2 主要キャラクターの定義

主要キャラクターとは、小説や演劇等の中心人物のことであり、特に雑誌や単行本として連載されるコミックや絵本、アニメーション、映画等の様々な分野において、対象のストーリーにおいて欠かせない登場人物となる場合が多い。特にコミックにおいては、主要キャラクターの紹介ページとして巻頭にまとめることで、読者がストーリー展開を容易く理解可能にする工夫がされているものもある。紹介ページの例を図1に示す。

本研究では、登場するキャラクターの中でもストーリーの進行上欠かせない役割を持つ登場人物を主要キャラクターと定義している。例えば、コミックにおけるキャラクター紹介ページにまとめられた登場キャラクターは本研究でも主要キャラクターと考える。他にも、対象コミックの公式ホームページ等でキャラクターの紹介がある場合は、作者や出版社がそのキャラクターをストーリー進行上重要と判断したキャラクターである。そこで、本研究はこれらの登場キャラクターも主要キャラクターと考える。また、主要キャラクターはストーリーの進行上重要な役割を持つ場合が多いため、コミックにおけるアクション、恋愛、サスペンス等のあらゆるジャンルにおいて、主要キャラクターは多くのコマで登場すると考えられる。

ここで述べるコマとは、コミックにおいて何かしらのアクションを登場キャラクターが

起こす場面や、何も起こらずとも登場キャラクターが描かれている場面など全てのコマがある。従って、主要キャラクターと定義可能である登場キャラクターは、ストーリーに関わらない別の登場キャラクターに比べて登場回数が多くなると推測できる。

以上の点より本研究における主要キャラクターは「ストーリーに大きく関わり、登場回数の多いキャラクター」と定義した。



図1 キャラクター紹介ページ例[6][7]

2.3 本研究における主要キャラクターの分類

コミックにおいて全ての登場キャラクターは、登場回数や他キャラクターとの関係性を基に、ストーリー進行に対してどの程度重要な役割を担うかを評価することで、階層構造で表現可能と推測した。例えば、主人公とは小説や脚本などの中心人物であり、ストーリーにおいて全ての事象の中心軸として描かれることが多いキャラクターである。また主人公は全ての登場キャラクターの中でも、最も他キャラクターとの関係性が読者にわかりやすく描かれている。コミックの場合、他主要キャラクターと同じシーン内に描かれ、登場回数が最も多くなると考えられる。そこで、主人公は登場回数が最も多いキャラクターと仮定する。

以上の点から、主人公は階層構造の頂点に属するキャラクターとして分類可能と考えられる。同様にして、一階層毎に登場回数や他キャラクターとの関係性を考慮すると、下の

階層に進むにつれて徐々にストーリー進行との関係性が低いキャラクター群になると考えられる。

階層構造は、上から (1) 主人公群, (2) ストーリーとの関係性が高く主人公と共に描かれるキャラクター群, あるいは主人公と相対する敵主力キャラクター群, (3) 主人公の所属する集団に属するキャラクター群あるいは主人公と相対する敵キャラクター群, (4) 主人公あるいはストーリーとの関係性が無い一時的に描かれるキャラクター群, (5) ただ一つのコマのみに描かれるストーリー進行上価値の無いキャラクター群, 計 5 種類の階層構造になると仮定した。コミックにおける登場キャラクター階層図を図 2 に示す。

この階層構造の中で「ストーリーに大きく関わり, 登場回数が多いキャラクター」である主要なキャラクターは, 第 1 階層と第 2 階層のキャラクターであると考えられる。第 3 階層以下のキャラクターの場合, 本筋であるストーリーに対して大きく影響を与えることがなく, 一時的に登場回数が増える場合もあるが, 対象の巻全体では登場回数が他キャラクターに比べ, 多くはならないと推測される。従って, 第 1 階層と第 2 階層に属するキャラクターを同定することが可能であれば, 対象コミックの主要キャラクターの同定が可能になると考えられる。

また本研究では, コミックにおいて主要キャラクターは「ストーリーに大きく関わり, 登場回数が多いキャラクター」と定義している。そのため, あるタイトルのコミックシリーズに注目した際に, 1 巻と 2 巻でストーリーの方向性が違う場合には, 同じコミックのシリーズにおいて巻毎に主要キャラクターは異なる場合があると推測している。

従って本研究における主要キャラクターの同定とは, あるコミックシリーズ全体に対しての主要キャラクターではなく, 対象コミックにおいて主要なキャラクターを同定する研究となる。

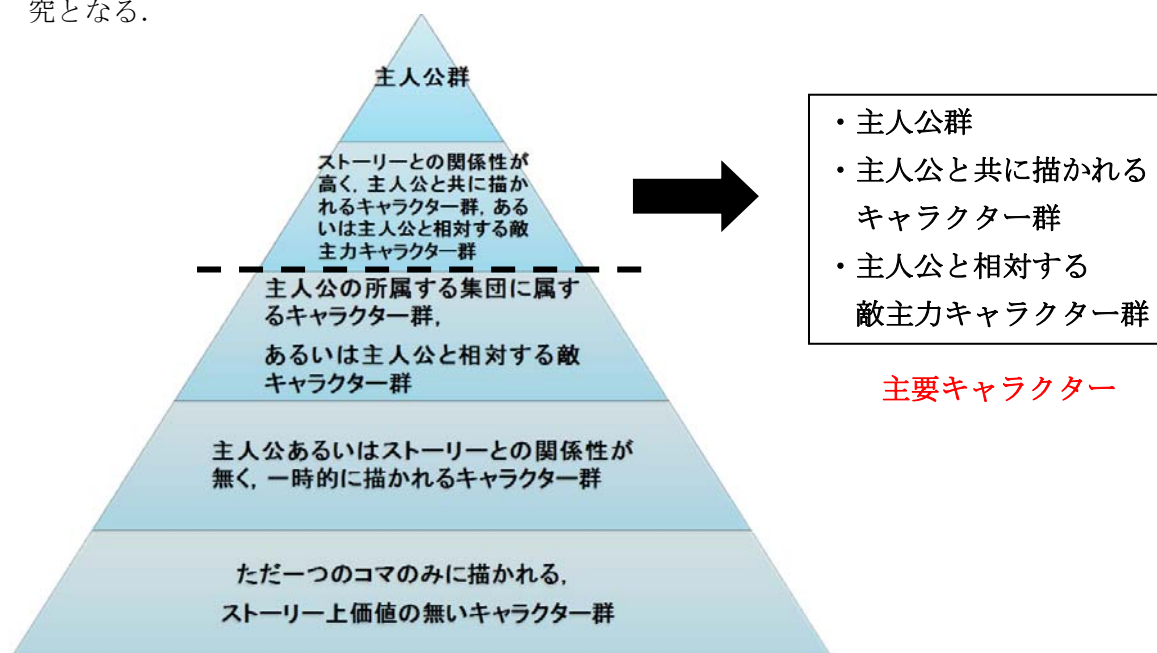


図 2 登場キャラクター分類

2.4 主要キャラクター登場回数比較実験

2.4.1 実験内容

本研究では、主要キャラクターを「ストーリーに大きく関わり、登場回数が多いキャラクター」と定義している。そこで、主要キャラクターを予め特定可能なコミックを使用し、手動で登場回数を計測し、登場回数が多いキャラクターが主要キャラクターであるか検証する。

使用するコミックは、学園、コメディ、バトルアクション、恋愛の4種類のジャンルに分類されたコミックであり、主要キャラクターを予め特定することが可能なものを選択する。使用したコミックを表1に示す。また手動で登場回数を計測する際、主観的な判断で明確にキャラクターの名称が特定可能な登場キャラクターのみを計測した。キャラクターが後ろ向きで描かれる場合や、ストーリーを読まなければ名称を明確に特定することが困難なキャラクターは対象外とした。

使用コミックにおける対象外のキャラクター例を図3に示す。上のコマが対象外のキャラクターであり、下のコマが対象として計測するキャラクターである。

表1 使用コミック一覧

ジャンル	使用コミック名	登場主要キャラクター数
学園	アオハライド[8] 1巻	5
コメディ	干物妹うまるちゃん[9] 1巻	8
バトル	緋弾のARIA AA[10] 1巻	10
恋愛	ラブひな[11] 1巻	7



図3 対象外のキャラクター例[11]

2.4.2 実験結果

本研究で仮定した階層構造に従い分類した結果を表 2-5, 計測実験の結果を図 4-7 に示す.

予め主要キャラクターとした登場キャラクターを手動で計測した結果, 非主要キャラクターに比べて登場回数が増える傾向があることが確認できた. また, 第 1 層と第 2 層のキャラクターの登場回数を比較した場合, 第 1 層である主人公群が最も多く登場することがわかった. この傾向はジャンルに関係無く, 殆どのコミックにおいて見られる傾向であると推測できる.

従って, 前述した主要キャラクターの定義である「ストーリーに大きく関わり, 登場回数が多いキャラクター」に基づき研究を進めることで主要キャラクターの同定が可能と判断した.

表 2 階層毎に分類したキャラクター名称と登場回数 1

コミック名	アオハライド 1巻					
キャラクター名称	吉岡	馬淵	田中	榎田	村尾	非主要キャラクター
所属階層	1	2	2	2	2	3層~5層
登場回数	286	75	59	39	7	91

表 3 階層毎に分類したキャラクター名称と登場回数 2

コミック名	干物妹うまるちゃん 1巻						
キャラクター名称	うまる (変形)	タイハイ	うまる	海老名	猛	シルフィン フォード	アレックス
所属階層	1	1	1	2	2	2	2
登場回数	216	184	105	39	13	6	6

切絵	非主要キャラクター
2	3層~5層
5	49

表 4 階層毎に分類したキャラクター名称と登場回数 3

コミック名	緋弾のエリア AA 1巻								
キャラクター名称	間宮	神埼	佐々木	火野	風魔	間宮 (妹)	峰	レキ	星伽
所属階層	1	1	2	2	2	2	2	2	2
登場回数	271	130	123	25	6	6	1	1	1

遠山	非主要キャラクター
2	3層～5層
1	41

表 5 階層毎に分類したキャラクター名称と登場回数 4

コミック名	ラブひな 1巻							
キャラクター名称	浦島	鳴瀬川	前原	紺野	青山	カオラ	浦島 (叔母)	非主要 キャラクター
所属階層	1	1	2	2	2	2	2	3層～5層
登場回数	294	206	70	69	63	41	14	22

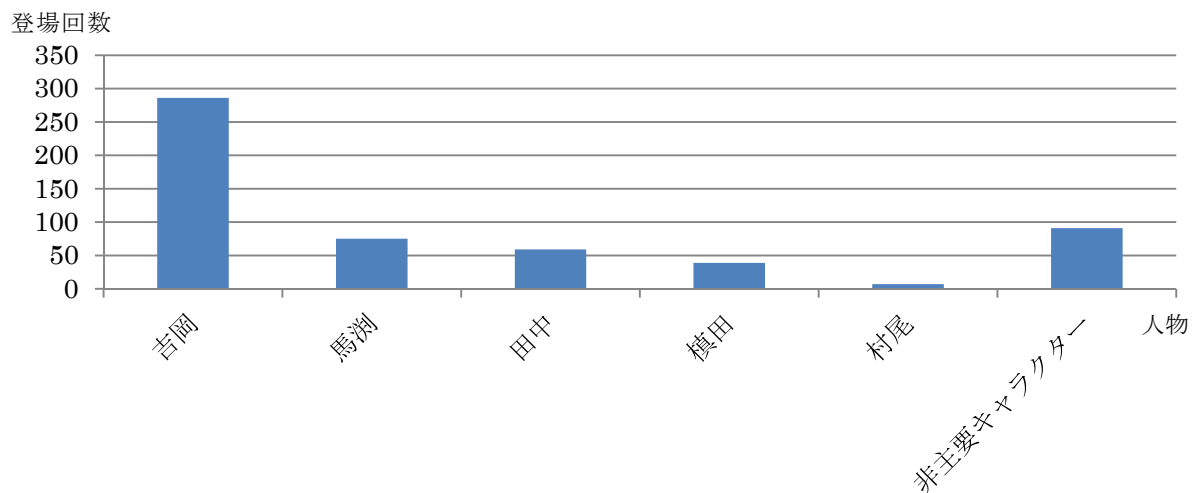


図 4 アオハライドにおける登場回数比較

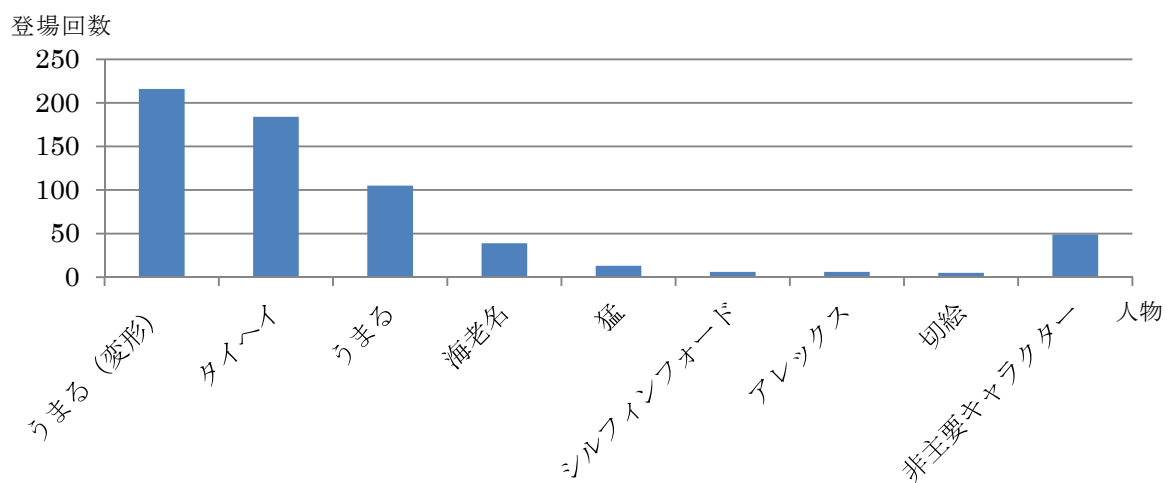


図 5 干物妹うまるちゃんにおける登場回数比較

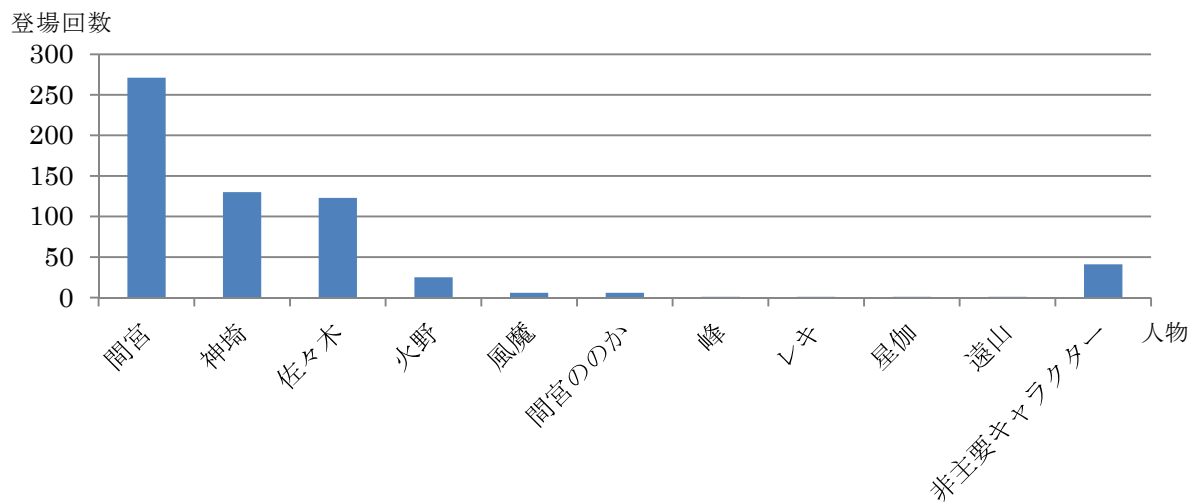


図 6 緋弾のエリア AA における登場回数比較

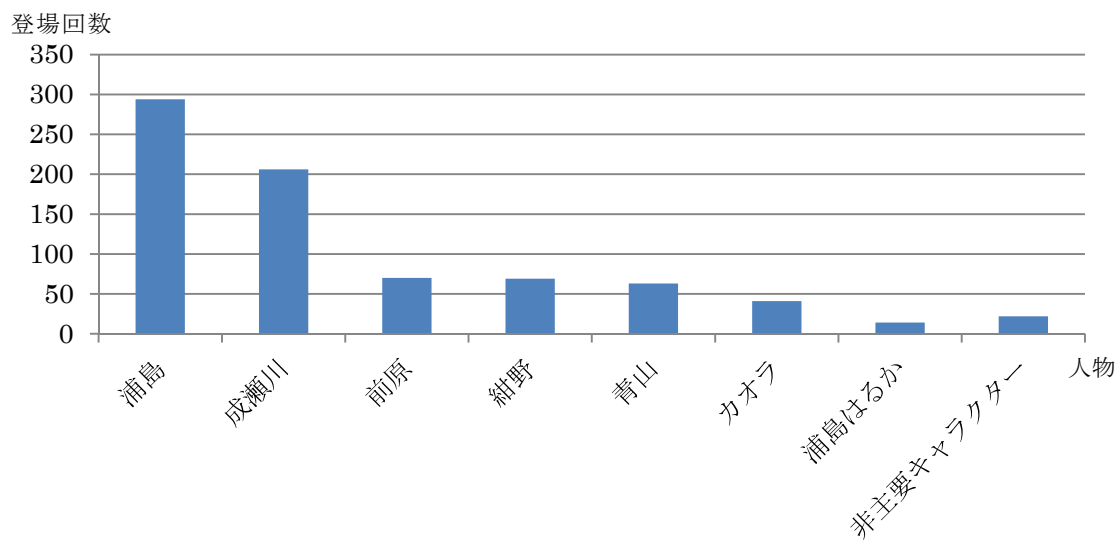


図 7 ラブひなにおける登場回数比較

2.5 条件「事前知識なし」の主要キャラクター同定について

従来、コミックからキャラクター情報の抽出を行なう場合には機械学習を前提としていた。例えば、認識対象の形状や色、特徴量などを事前に学習し、学習データを予め作成することで、入力された対象の分類や認識を機械が判断し出力する。この手法は、物体認識をはじめ、画像認識や音声認識などの様々な分野で用いられている。近年では機械学習の新たな手法であるディープラーニングに関する研究が進み、精度の向上が見られることから注目を集めている[12][13]。

これらの機械学習では、学習データを作成する際に、対象の画像が含まれる正例画像と、含まれない負例画像を必要とし、事前知識を前提とした学習となる場合が多い。しかし、事前知識を取得するためには人の介入が必要であり、処理時間の短縮が困難である。コミックの電子化においても、メタデータを付加させるためには人間が対象コミックを読むことが必要であり、ストーリー理解に時間を割くことが不可欠となる。

本研究では、利用価値の高いメタデータである主要キャラクター情報を抽出し、コミック電子化の流れにおいて、自動要約やキャラクター検索等の新しいコミックサービスの創出に利用ことを目的としている。

従って、本研究では主要キャラクターの同定において、主要キャラクターがどのキャラクターであるかを人が学習するプロセスを省き、自動的に同定可能とする手法を検討する。そのため、予め人間が学習を行わない状態を想定し、主要キャラクターの同定プロセスでは、「事前知識なし」の条件を設け実験を行った。

第3章

コミックにおける 主要キャラクターの同定手法

3.1 まえがき

この章では、まず既存のキャラクター顔画像認識手法と画像特徴について述べ、本研究で使用する SURF 特徴について述べる。

3.2 キャラクター顔認識手法

3.2.1 従来のキャラクター顔認識について

コミックにおけるキャラクター認識は、既存手法を用いて様々な特徴を抽出し実現している。現状では、主に人間の顔識別で用いる手法が採用されている。例えば、Haar-like 特徴を用いて Adaboost で学習させ、カスケード分類器を作成し、対象画像から顔の認識を行う手法が挙げられる[14][15]。

Haar-like 特徴とは、矩形領域の輝度値から平均値を算出し、その差を特徴とするものであり、Adaboost は弱小識別器を直列に繋ぎ、一つの頑強な分類器を形成させる手法である。これらは主に人間の顔認識に用いられる手法であるが、コミックにおけるキャラクターの顔認識にも対応させることが可能である。しかし、キャラクターの顔描写は作者の趣向により大きく異なるため、人間顔の形状など決められた形状が存在しない。例えば、キャラクターが線画で描かれる場合や、髪型が奇抜な形状で描かれる場合などが挙げられる。他にも登場キャラクターが人型で無い場合もあり、動物顔のキャラクターも存在する。図 8 に非一様な特徴を持つ登場キャラクターと、作者の趣向で顔の表現が変化し異なる見た目となったキャラクターの例を示す。これらのキャラクターは人間の特徴と大きく異なり、

一様な特徴を持たず、人間の顔特徴を基に顔判別を行う既存手法では安定した精度を維持することが困難である。また、図8において赤枠はHaar-like識別器で認識可能な顔である。しかし、等しいキャラクター顔画像でも顔形状が崩れた右図の場合では認識が困難である。これより、等しい識別環境でもキャラクター顔画像に対する識別精度が維持できない場合があることがわかる。従って、コミックにおいて推奨されるキャラクター顔認識手法が定められていない現状がある。

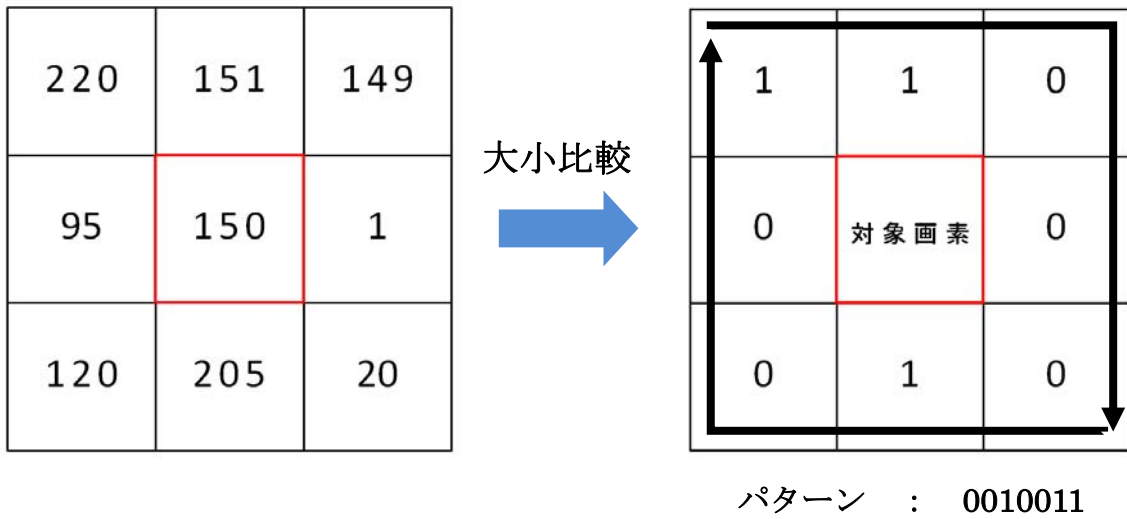


図8 非一様な特徴を持つ登場キャラクター例[16]

3.2.2 LBP (Local Binary Pattern) 画像特徴

LBPとは、通常 3×3 画素領域からバイナリーパターンを算出することで画像の局所特徴量を抽出する手法である[17]。まず、画素領域中央と、その近傍領域である8画素の差をとり画素の正負を比較する。正であれば1、負であれば0を割り当てることで、近傍領域を8ビットの2進数で表記可能になる。そして、この2進数を進数変換することで、局所特徴を10進数表記で表すことが可能になる。この際、最大8ビットの2進数になる為、0から256の範囲で局所特徴を表すことが出来る。従って、LBPは256階調であり、正規化することで256階調グレースケール画像となる。図9にLBPの算出例を示す。

またLBPは、局所特徴を2進数で表現する為、同様の図柄を持つ場合には輝度値による影響を受けない。これは、正負の比較プロセスにおいて輝度が変化した場合に、ほぼ全体に対して同様の輝度影響が起きる為、特に2値の間で関係が変化することは無く、大小関係を保つからである。従って、ほぼ同じ特徴で表記することができ、LBP画像特徴は照明変化に頑強な画像特徴となる。さらに、対象画像における位置情報を局所特徴に付与していないため、画素位置の変化でバイナリーパターンに変化が生じない。従ってLBP画像特徴は位置不変性を備えている画像特徴である。



1	2	4
128	対象画素	8
64	32	16

$$0010011 \rightarrow 0 + 0 + 32 + 0 + 0 + 2 + 1 = 35$$

対象画素は, 35

対象画素 1 の局所特徴 + 対象画素 2 の局所特徴 + ... = 画像全体の特徵

図9 LBPの算出例

3.2.3 AnimeFace 識別器[18]

AnimeFace とは、学習済みのアニメキャラクター顔に特化した汎用識別器である。コミックにおけるキャラクター顔は、人間の顔と大きく異なるだけではなく、著者により描画の仕方が異なる。そのため、定まった特徴を持たない。

そこで、様々なパターンに対して対応可能にするために、複数のキャラクター顔を学習させることで汎用な識別器を作り、キャラクター顔認識を可能にした。

使用する画像はランダムなキャラクター顔画像であり、予め画像を顔領域と顔以外の領域に分け、それぞれ正画像と負画像とした。これらを各 20000 枚用意し、特徴を抽出する。使用されている画像特徴は LBP である。

人間が主観で明確にキャラクター顔領域を判断可能である領域ならば、おおよそキャラクター顔を判別可能である。しかし、コミックにおいてキャラクターは、非一様な形状であるロボットや動物などがキャラクターとして登場する場合も有る。この場合は精度が維持出来ず、顔を認識出来ないコミックもあることが確認されている。

本研究は、対象コミックにおいて事前知識なしで主要キャラクターを同定する手法を検討している。従って、対象コミックに特化することが無い汎用な識別器である AnimeFace を用いて、対象コミックについて事前知識を持たずにキャラクター顔画像を抽出させた。

3.3 Speeded Up Robust Features (SURF)

3.3.1 SURF の概要

SURF とは特定物体認識手法の一つであり、画像のスケール変化や照明変化等に頑強な画像特徴である。特定物体認識では、対象である特定物体の状態に影響されず、変化に柔軟に対応する必要があった。そこで、SIFT (Scale Invariant Feature Transform) が提案され、変化に頑強な特徴を抽出可能になった[19]。

SIFT は大きく分け二つのプロセスが行われる。まず特徴点の検出を行う。スケール変化に頑強に対応する為に DoG (Difference of Gaussians) 画像を作成し、重要な特徴点とスケールを選択する。次に特徴記述を行う。回転や照明変化に頑強に対応する為に、精度の高い輝度方向を算出し正規化する。しかし、SIFT は計算コストが増加する場合もあり、瞬時に処理を行いたい場合に困難な特徴であった。そこで、SIFT の特性であるスケール変化や回転、照明変化の頑強さを維持し、高速に処理可能な SURF が提案されている[20]。

3.3.2 SURF における特徴点計算[20]

Harbert らによると処理コストが高くなりやすい傾向を持つ Hessian 行列において, SURF では Integral Image[21]と box filters を使用することで計算処理速度を向上させている.

Hessian 行列は, 2 次微分演算と Gaussian 関数の標準偏差で式(3.1)として表すことができる. これは, エッジ成分を表す式であり, 判別式(3.2)よりそのエッジの種類を算出することが出来る.

$$\mathcal{H}(\mathbf{x}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) \\ L_{xy}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) \end{bmatrix} \quad (3.1)$$

$$\det(\mathcal{H}(\mathbf{x}, \sigma)) = L_{xx}(\mathbf{x}, \sigma) L_{yy}(\mathbf{x}, \sigma) - (L_{xy}(\mathbf{x}, \sigma))^2 \quad (3.2)$$

L は Gaussian 関数において 2 次微分演算で方向要素を画像に畳み込む応答値となる. 各応答値を式(3.3), 式(3.4), 式(3.5)として表す.

$$L_{xx}(\mathbf{x}, \sigma) = I_x \frac{\partial^2}{\partial x^2} g(\sigma) \quad (3.3)$$

$$L_{yy}(\mathbf{x}, \sigma) = I_x \frac{\partial^2}{\partial y^2} g(\sigma) \quad (3.4)$$

$$L_{xy}(\mathbf{x}, \sigma) = I_x \frac{\partial^2}{\partial x \partial y} g(\sigma) \quad (3.5)$$

しかし, 式(3.2)において 2 次微分演算は計算コストが高くなり, エッジの種類を算出する際に時間がかかる. そこで, Integral Image と box filters を利用することで計算コストを抑える.

まず box filters を利用し式(3.3), 式(3.4), 式(3.5)を D_{xx} , D_{yy} , D_{xy} として近似する. 近似した D_{xx} , D_{yy} , D_{xy} を用いて, 判別式(3.2)は近似判別式(3.6)と表記できる.

$$\det(\mathcal{H}_{\text{approx}}) = D_{xx} D_{yy} - (w D_{xy})^2 \quad (3.6)$$

例えば, $\sigma = 1.2$ の場合について box filters を図 10, 図 11 に示す. box filters は 9×9 サイズである. 式(3.6)におけるパラメタ w は, 標準偏差と box filters のサイズを基に, 式(3.7)より $w = 0.9$ と算出できる. この w は近似により発生した誤差を抑える為の Harbert らに

より算出されたパラメタである.

図 10 は L_{yy} と L_{xy} を xy 次元で表現した図であり, 図 11 は近似された D_{yy} と D_{xy} を同じく表現した図である.

$$w = \frac{|L_{xy(1.2)}|_F |D_{yy(9)}|_F}{|L_{yy(1.2)}|_F |D_{xy(9)}|_F} = 0.912... \approx 0.9 \quad (3.7)$$

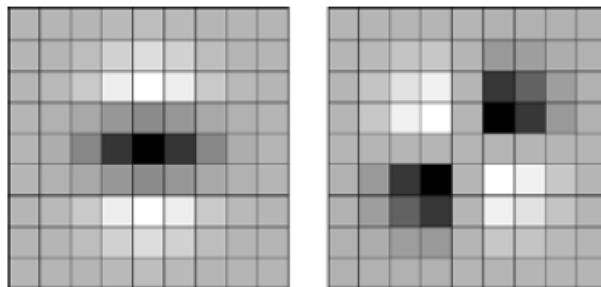


図 10 box filters 例 1 [20]

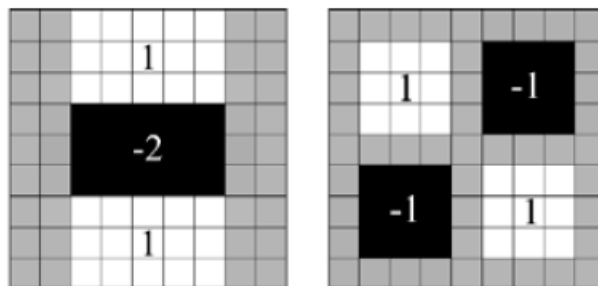


図 11 box filters 例 2 [20]

この box filters により計算コストの高い処理を簡易に近似して行うことが出来るため, Integral Image が適応可能になる. 次に, 予め用意した画像の積分演算結果を基に矩形領域の輝度和を算出する. この処理により輝度和の演算を行うことができ, 各要素に対して高速に近似 Hessian 行列を算出することが可能になる. これらの処理をすることで SURF は, SIFT に比べ計算コストを抑え, 特徴点を選択可能な新たな特徴として注目された.

第 4 章

Bag-of-Visual Words

4.1 まえがき

本章では、Bag-of-Visual Words の手順について述べる。また、キャラクター顔をヒストグラム化することについて述べる。

4.2 局所特徴による画像認識

従来の画像認識で代表的な手法では、対象の全体に対して特徴量を検出し、そのパターンを認識することで画像識別や機械学習を行う。しかし、同じ対象物体が与えられた際に、スケール変化の影響が大きい場合や形状が異なる場合がある。他にも、3次元空間上に対象物体がある画像において、物体の前後に別の物体が置かれ一部隠れてしまい、その画像において対象の認識が困難になる場合がある。これらの例の様に全体に対して特徴を検出させる手法だけでは、画像認識において限界があった。

そこで、全体に対して特徴を算出するのではなく、物体のパーツに注目することでその対象を認識する手法が提案された。例えば、Part-based がこれに相当する手法である。この手法は、Support Vector Machine にも使用可能であるため、近年では多くの画像認識手法で用いられている[22][23]。

また局所部分に注目する手法では、対象画像にある特定の物体のみに対して認識を試みるものである。近年では対象画像内にある複数の物体の関係から確率的に空間全体に対して関連付けを行い、認識可能であることが分かっている。これらはシーン認識と呼ばれ、近年では様々な手法と共に使われる代表的なものとなっている[24]。

4.3 Bag-of-Visual Words について

Bag-of-Visual Words (BoVW) とは、画像認識に局所特徴量を用いる手法の一つである [25]。BoVW は Bag-of-Keypoints (BoK) や Bag-of-Features (BoF) と呼ばれることもある。BoVW は、言語処理分野においてテキストデータから単語をベクトルにすることで文章の特徴を抽出する Bag-of-Words (BoW) が基となる手法である [22]。

BoW では単語をベクトルに対応させ、文章中における単語出現頻度を対象文章の特徴としている。これに対して BoVW は、単語を局所特徴に置き換え、画像を局所特徴の集合として処理をする。そして局所特徴をベクトルに対応させることで、そのベクトルの頻度からヒストグラムを作成し、対象の特徴とすることで認識を行う。

従来の認識手法では、局所特徴の次元数を類似度算出の際に用いる為、計算コストが増加するが多かった。一方、BoVW は画像の特徴を Visual Words の頻度から作られたヒストグラムで表現することが出来る。この処理により画像特徴はヒストグラムのみとなる為、大幅な次元数の削減が可能となった。

BoVW の概要図を図 12 に示す。BoVW は以下の手順で行われる。

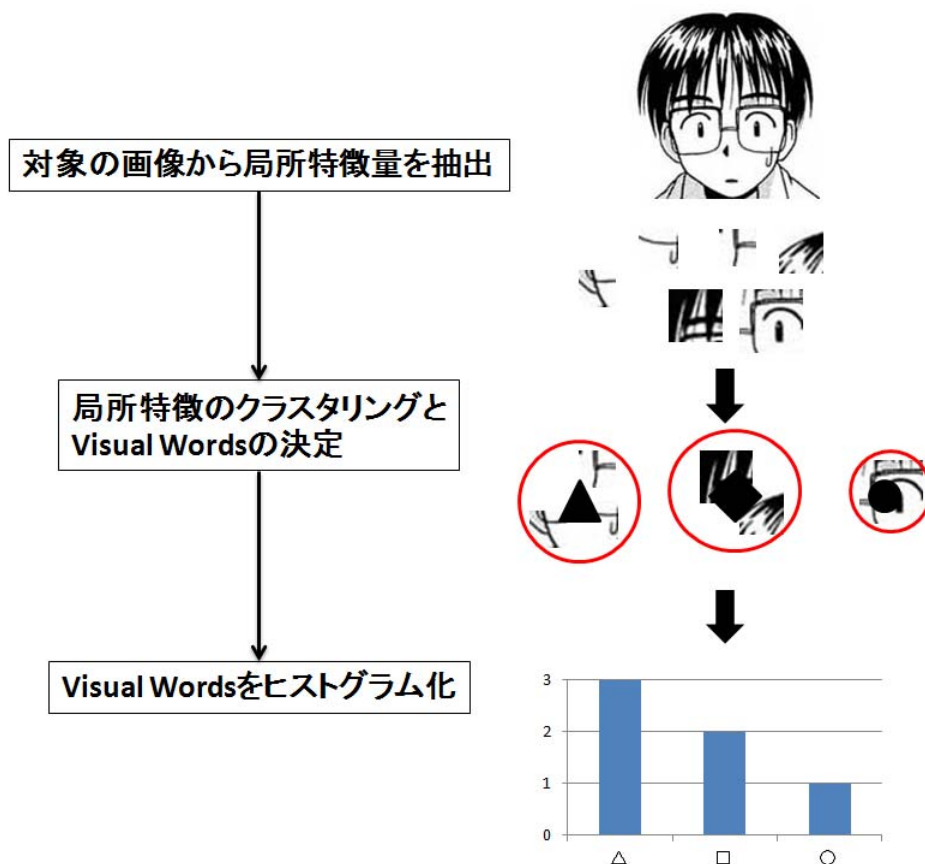


図 12 BoVW 概要図[11]

手順 (1) 対象の画像から局所特徴を抽出

認識対象である画像から局所特徴を抽出する。BoVW が扱う認識問題は、全体特徴を用いて対象物体の認識が困難なものであり、画像の部分毎に局所特徴を抽出する必要がある。そこで、BoVW は局所特徴として画像のスケール変化や照明変化等に強い SIFT や SURF を局所特徴として使用する。

本研究では、対象画像のスケール変化に対して、計算コストを低く抑えることが可能な SURF を選択している。キャラクター顔画像から SURF を抽出した例を図 13 に示す。



図 13 コミックにおけるキャラクター顔画像 SURF 特徴例[16]

手順 (2) 局所特徴のクラスタリングと Visual Words の決定

抽出した局所特徴をクラスタリングし Visual Words を決定する。対象から局所特徴を抽出した後、これらを K-means 法によりクラスタリングする[26]。K-means とは、予め K の値を決定することで、与えられた要素を K 個のクラスに分類をする手法である。

まず要素からランダムに K 個選択する。その要素と各要素群との距離を算出し、K 個のクラスに分類していく。このクラス毎の中心を次の要素選択の際に使用する。この手順をクラス結果に変化が生じなくなるまで繰り返すことで K-means 法によるクラスタリングを実現している。

最後に K 個にクラスタリングされた局所特徴群それぞれの中心であるセントロイドを選択する。これが Visual Words である。局所特徴の中でも最も類似している特徴同士を一つの特徴として扱う為、次元数を削減可能になる。

手順 (3) Visual Words をヒストグラム化

局所特徴から算出した Visual Words からヒストグラムを作成する。対象画像中の特徴と Visual Words の距離を計算し、最も近い Visual Words 同士を同じ集合として集める。これを全ての Visual Words に対して行うことで、各 Visual Words 毎の頻度が得られる。この頻度をヒストグラム化することで、与えられた画像群全体に対して、それぞれの画像を特徴的なヒストグラムとして表現できる。このヒストグラムが対象画像の特徴となるため、事前知識が無い場合においても分類が可能である。

4.4 Bag-of-Visual Words を用いたキャラクター顔ヒストグラム

本研究では、事前知識なしで主要キャラクターの同定手法を行う。BoVW は画像認識に使用できるため、近年では類似画像検索への利用などが提案されている[26]。そこで、コミックにおけるキャラクター顔画像も BoVW で分類可能であると考えられる。

特定物体認識と同様に、キャラクター顔もそれぞれ異なる形状をしている。特にコミックにおいてキャラクターは、髪型、顔立ちなどに強く特徴を持たせることで描き分けている。従って、これらの特徴を Visual Words として選択可能ならば、キャラクター毎にヒストグラムが変化するため、自動分類が可能であると考えられる。

本研究では、はじめに汎用な識別器である AnimeFace を利用している。そのため、人型のキャラクターの顔画像のみが抽出されることになる。従って、本研究では非一様な特徴を持つキャラクターを扱わないこととして、人型のキャラクターにおける自動分類を BoVW で行う手法を検討する。

第 5 章

コミックにおける 主要キャラクターの同定手法

5.1 まえがき

本章では、まず提案手法の概要と提案手法手順、評価方法について述べる。次に提案手法で主要キャラクターの同定実験を行い、実験結果とその考察を述べる。

5.2 提案手法について

5.2.1 提案手法概要

本稿ではコミックにおける主要キャラクターの同定に、SURF 特徴と BoVW を用いる。まず、電子コミックデータからキャラクター顔画像を抽出し、その顔画像から SURF 特徴を抽出する。次に、BoVW を使用することで、キャラクター毎にヒストグラムを作成する。最後に作成したヒストグラムを K-means 法で分類する。

本研究において、主要キャラクターは登場回数が多いと仮定している。従って、各キャラクターを正確にヒストグラムで表せると想定した場合、クラスタリング結果において主要キャラクターを含むクラスは保有画像が多くなる。そこで、ヒストグラムを K-means 法でクラスタリングし、キャラクター分類を行うことで、保有画像の多いクラスを判定し、主要キャラクタークラスを取得する。このクラスにおける所属度 1 位のキャラクター顔画像が、主要キャラクターと考えられる。

提案手法では、これらのプロセスを経て主要キャラクターと考えられるキャラクター顔画像を自動で抽出する。以上が提案する主要キャラクター同定手法の概要であり、提案手

法による主要キャラクター同定の妥当性を検討する．提案手法の概要図を図 14 に示す．

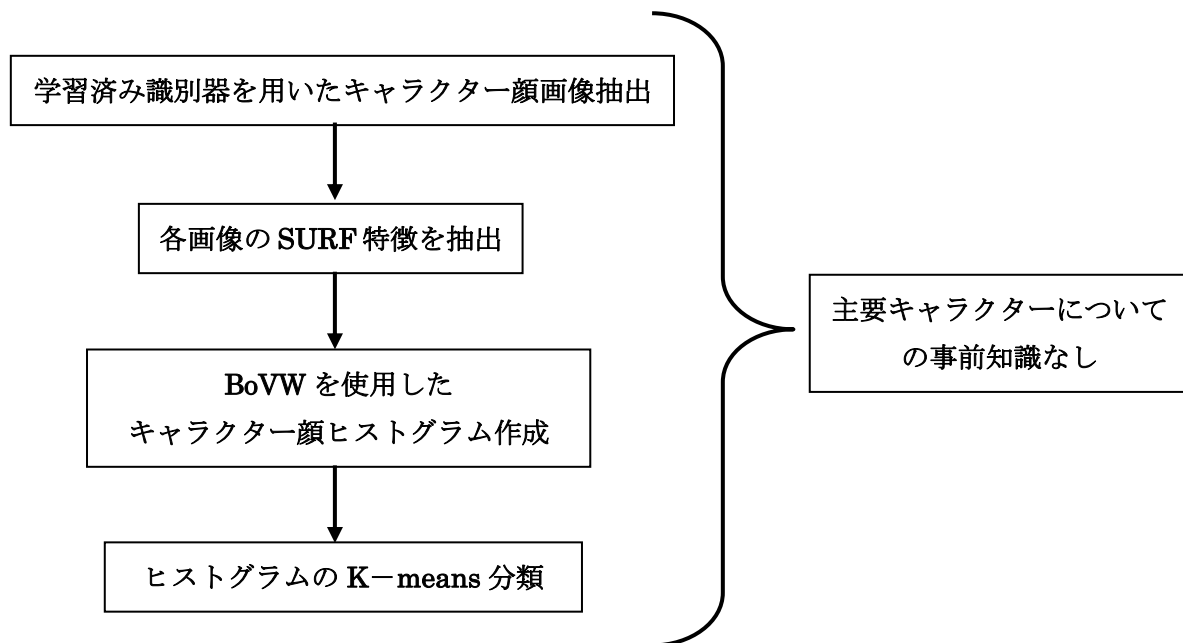


図 14 提案手法概要図

5.2.2 提案手法手順

手順 (1) 学習済み汎用識別器を用いたキャラクター顔画像抽出

電子コミック画像からキャラクター顔画像を検出し，その領域を抽出する．対象となる画像は全てモノクロ画像とする為，予め全てのコミックデータをモノクロ変換している．使用する電子コミック画像は JPEG 画像であり，画像のサイズは幅が約 1100～1200 ピクセル，高さが約 1700 ピクセルである．本研究において，これらの画像サイズは定めていない．理由としては，このプロセスはコミック画像からキャラクター領域を抽出するものであり，コミックデータをリサイズする必要が無いためである．

キャラクター領域を抽出するために，学習済みの汎用識別器である AnimeFace を使用する．この識別器により抽出することが可能なコミックにおける登場キャラクター顔画像を使用して実験を行う．この識別器は，ロボットや動物などの非一様な特徴を持つキャラクターに対して検出が出来ない．そこで，本研究では人型のキャラクターが主要キャラクターとなるコミックについてのみで実験することとし，提案手法の妥当性を確認する．この識別器を使用する際のパラメタを表 6 に示す．

表 6 パラメタ設定

パラメタ要項	設定
抽出画像スケール比	縦横比 1 : 1
最小近傍矩形数	3
最小顔検出サイズ	50 * 50 ピクセル

・抽出画像スケール比

コミックにおける人型キャラクター顔は、基本的に正方形型の枠内に収まるように描かれていることが多い。本研究では、キャラクター顔画像において重要な顔パーツは、髪型の中でも主に前髪、顔形、目、鼻、口と考えている。従来の人間顔認識研究においても、同様に髪型や顔形、顔正面パーツを重要顔特徴と考えている提案が多い。そこで、コミックにおけるキャラクター顔画像においても、これらのパーツが重要であると考え、これらの顔特徴が取得可能な画像スケール比にした。

・最小近傍矩形数

識別器でキャラクター領域を検出する際、検出矩形領域を変化させつつ複数回実行している。その為、一つのキャラクター顔領域に対して複数個の顔領域が検出される。そこで検出矩形数の最小個数を定め、より識別器が出力する顔領域が正確な結果となるようにパラメタの設定を行う必要がある。

コミックにおいてキャラクター顔は、大きく精密に描かれるパターンと、小さいコマに描かれている為に、細部まで顔パーツが描画されていないパターンが存在する。識別器を使用し顔領域を抽出する場合には、全てのパターンに対応可能なパラメタを設定することが理想である。

しかし、本稿で提案する手法において、効率的な実験結果が得られる顔画像数が未知数である。従って、最小近傍矩形数を顔サイズに関係なく、精度を維持しつつも多くの顔領域を抽出することが可能であると考えられる 3 を設定し実験を行った。

・最小顔検出サイズ

検出されるキャラクター顔画像は様々なサイズがある。試験的に実験で使用する識別器を用いてキャラクター顔画像を検出した際、対象コミックにおける画像サイズはどのコミックにおいても約 $50 * 50$ ピクセル～ $500 * 500$ ピクセルとなった。本研究では、可能な限り多くのキャラクター顔画像を抽出し、主要キャラクター判定処理に用いる。そこで、大抵のコミックにおいて人間が明確に認識可能であると推測される $50 * 50$ ピクセルを試験的に最小顔検出サイズとして設定した。

手順 (2) キャラクター顔画像の事前処理

コミックにおいて登場キャラクターは様々な画像サイズとなる。

抽出したキャラクター顔画像のサイズにより実験結果が変動することを防ぐため、全ての抽出されたキャラクター顔画像のリサイズ処理を事前に行う。識別器で抽出されるキャラクター画像サイズは、約 $50 * 50$ ピクセル～ $500 * 500$ ピクセルである。そこで、小さいサイズの画像解像度が極端に下がらないと考えられる大ききでリサイズ処理を行う。本研究では、試験的に $100 * 100$ ピクセルと $150 * 150$ ピクセルでリサイズ処理を行い、提案手法においてリサイズした画像の大ききが結果に影響を与えるかを検証した。

手順 (3) SURF 特徴の抽出

リサイズしたキャラクター顔画像から SURF 特徴の抽出を行う。SURF 特徴の記述において勾配方向と強度が算出される。特徴記述には基本的に 64 次元の特徴を用いて表現する。また、精度の向上を目的として向きと強さの条件を加えた 128 次元の特徴も用意されている。

そこで、本研究ではより精度の向上が望める SURF-128 を特徴として選択し、キャラクター顔画像から抽出した。

手順 (4) キャラクター顔画像のヒストグラム化

BoVW を使用し、抽出した SURF-128 の頻度からヒストグラムを作成する。まず Visual Words を決定する。この際、パラメタとして Visual Words の数 (次元数) を設定している。しかし、コミックにおけるキャラクター顔画像の持つ SURF-128 特徴を使用した Visual Words において、最も精度良くキャラクター顔画像が分類される表現が可能な次元数は現状では未知数である。従って本研究では試験的に 500 次元の Visual Words を決定しヒストグラムの作成をした。作成されたキャラクターのヒストグラム例を図 15、図 16 に示す。

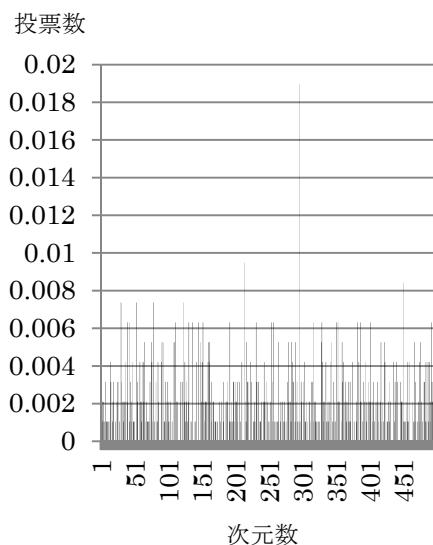


図 15 キャラクターA

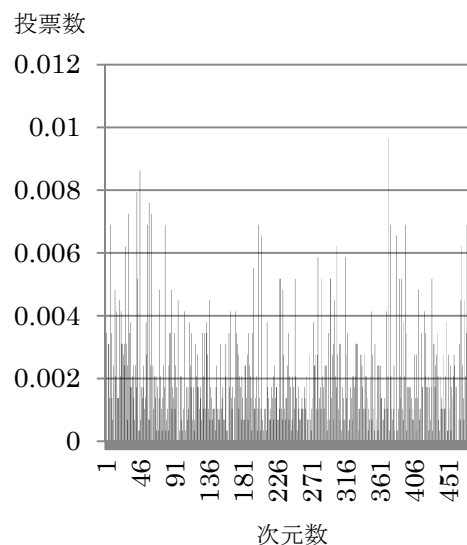


図 16 キャラクターB

手順 (5) K-means 法によるクラスタリング

作成したキャラクター顔画像毎のヒストグラムを K-means 法により分類する。作成したヒストグラムでは、500 次元で対象画像における各 Visual Words の頻度が表現されている。このクラスタリング処理には、クラスタリングツール bayon を使用した[28]。bayon はヒストグラムを K-means 法でクラスタリング可能な公開ツールである。

本研究において、主要キャラクターは登場回数が多いと定義している。従って、各キャラクターを正確にヒストグラムで表せると想定した場合、クラスタリング結果において主要キャラクターを含むクラスは保有画像が多くなる。そこで、クラス数 K をパラメタとして設定し、 K 個のクラスにおいて最も保有画像の多いクラスにおける所属度 1 位のキャラクター顔画像を抽出することで主要キャラクターが同定可能になる。

しかし、正確に主要キャラクターを分類可能なクラス数が現状では未知数である。従って、試験的に 30 クラスでの分類を試みた。ある程度大きな値である 30 を設定した理由は、識別器でキャラクター画像を抽出する際、キャラクター顔領域を含まないノイズデータも同時に抽出しているため、少ないクラス数に設定すると、目的の画像が含まれていないノイズを中心としたクラスが作成される可能性が高くなる。

5.3 コミックにおける主要キャラクター同定実験

5.3.1 実験内容

コミックにおける主要キャラクターに対して，提案手法により同定を試みた．使用するコミックを表 7 に示す．また提案手法のステップ毎に設定したパラメータを表 8 に示す．使用するコミックは A～J までの計 10 巻であり，予め主要キャラクターを特定してから実験を行った．またこの実験において，キャラクター顔画像のリサイズ処理により結果が影響を受けるか検証実験を行った．リサイズ処理は，試験的に 100*100 ピクセルと 150*150 ピクセルとした．それぞれについて提案手法を行い結果の比較を行う．

提案手法により主要キャラクターを同定した後，評価方法に則り検証を行い，提案手法の妥当性を確認する．

表 7 使用コミック一覧表

ジャンル	コミック名称	コミック No
恋愛	アオハライド 1 巻	A
〃	アオハライド 2 巻[29]	B
〃	失恋ショコラティエ 1 巻[30]	C
〃	失恋ショコラティエ 4 巻[31]	D
バトル	東京喰種 1 巻[32]	E
〃	東京喰種 2 巻[33]	F
〃	緋弾のアリア AA 1 巻	G
〃	緋弾のアリア AA 2 巻[34]	H
コメディ	干物妹うまるちゃん 1 巻	I
〃	干物妹うまるちゃん 2 巻[35]	J

表 8 パラメータ設定

キャラクター顔画像サイズ	100*100 ピクセル, 150*150 ピクセル
局所特徴量	SURF-128
Visual Words 次元数	500 次元
クラスタリング手法	K-means 法
クラス数	30 クラス(K = 30)

5.3.2 評価方法

提案手法において、主要キャラクターの自動同定には事前知識を用いていない。従って、全てのコミックにおいて主要キャラクター情報が事前知識として得られていない。そこで、それぞれのコミックにおける主要キャラクターを実験前に同定し、提案手法による実験結果と比較し妥当性を確認する。

実験で使用するコミックについて、予め公式ホームページや内容を確認した実験協力者 5 人に、本研究で仮定した登場キャラクター分類方法に基づき登場キャラクターの分類を依頼した。これらを基に登場キャラクターを分類した。作成した登場キャラクター階層表をそれぞれ表 9～表 16 に示す。

提案手法で主要キャラクターを同定した後、登場キャラクター階層図と比較することで妥当と考えられる結果が得られるかどうか、提案手法の妥当性を検証する。コミックにおいて登場キャラクターは 5 種類の階層構造をとると仮定した。そこで、階層毎にポイントをつけることで対象コミックにおける提案手法の妥当性を評価する。ポイントは、第 1 階層より順に第 5 階層までを 4, 3, 2, 1, 0 とした。評価表を表 17 に示す。

提案手法で主要キャラクターが同定可能ならば、実験で使用する全てのコミックにおいて 4 あるいは 3 ポイントの評価になると推測される。このポイントの平均を算出することで提案手法の妥当性を検証する。

表 9 アオハライド 1 巻, 2 巻

キャラクター名	階層
吉岡	1
馬淵	2
田中	2
槇田	2
村尾	2
非主要キャラクター	3～5

表 10 失恋ショコラティエ 1 巻

キャラクター名	階層
小動	1
吉岡	1
オリヴィエ	2
井上	2
非主要キャラクター	3～5

表 11 失恋ショコラティエ 4巻

キャラクター名	階層
小動	1
吉岡	2
関谷	2
井上	2
オリヴィエ	2
加藤	2
小動(妹)	2
六道	2
非主要キャラクター	3～5

表 12 東京喰種 1巻

キャラクター名	階層
金木	1
リゼ	2
霧島	2
芳村	2
永近	2
西尾	2
非主要キャラクター	3～5

表 13 東京喰種 2巻

キャラクター名	階層
金木	1
霧島	1
芳村	2
笛口	2
四方	2
亜門	2
真戸	2
非主要キャラクター	3～5

表 14 緋弾のエリア AA 1 卷

キャラクター名	階層
間宮	1
神埼	1
佐々木	2
火野	2
風魔	2
間宮ののか	2
峰	2
レキ	2
星伽	2
遠山	2
非主要キャラクター	3~5

表 15 緋弾のエリア AA 2 卷

キャラクター名	階層
間宮	1
神埼	1
佐々木	2
火野	2
島	2
風魔	2
間宮ののか	2
峰	2
レキ	2
星伽	2
遠山	2
非主要キャラクター	3~5

表 16 干物妹うまるちゃん 1, 2 巻

キャラクター名	階層
うまる(変形)	1
タイヘイ	1
うまる	1
海老名	2
猛	2
シルフィンフォード	2
アレックス	2
切絵	2
非主要キャラクター	3~5

表 17 評価表

階層	分類キャラクター名	主要 キャラクタ ー	ポイント
1	主人公群	○	4
2	ストーリーとの関係性が高く 主人公と共に描かれるキャラクター群, あるいは主人公と相対する敵主力キャラクター群	○	3
3	主人公の所属する集団に属するキャラクター群, あ るいは主人公と相対するキャラクター群	—	2
4	主人公あるいはストーリーとの関連性が無く, 一時的に描かれるキャラクター群	—	1
5	ただ一つのコマのみに描かれる, ストーリー上価値の無いキャラクター群	—	0

5.3.3 実験結果

提案手法における実験結果を表 18, 表 19 に示す. 全ての結果が正解の場合, 結果は 4 ポイントとなる. 従って, 100*100 ピクセルの場合は正解率 63%, 150*150 ピクセルの場合は正解率 88%の結果が得られたことがわかる. また, リサイズ処理による結果の影響を見るために 100*100 ピクセルと 150*150 ピクセルの結果を比較する. 大きくリサイズ処理を行うことで 25%の正解率向上がみられた. 一部, 0 ポイントとなり同定が出来ないコミックがあった. この時, 主要キャラクターとして抽出された画像は, キャラクター顔画像ではなく, 何かしらの物体が描画されたノイズデータであった. ノイズデータの例を図 17 に示す.

表 18 100*100 ピクセルの場合における主要キャラクター同定結果

	コミック名										平均ポイント
	A	B	C	D	E	F	G	H	I	J	
1 位同定ポイント	0	3	3	3	0	4	3	3	4	4	2.7
2 位同定ポイント	0	0	3	3	0	2	4	3	4	4	2.3
各コミックにおける平均同定ポイント	0	1.5	3	3	0	3	3.5	3	4	4	2.5

表 19 150*150 ピクセルの場合における主要キャラクター同定結果

	コミック名										平均ポイント
	A	B	C	D	E	F	G	H	I	J	
1 位同定ポイント	4	4	4	3	4	3	4	4	4	4	3.8
2 位同定ポイント	4	3	4	4	3	3	4	4	0	3	3.2
各コミックにおける平均同定ポイント	4	3.5	4	3.5	3.5	3	4	4	2	3.5	3.5



図 17 ノイズデータの例

5.4 考察

実験結果より, 使用コミックにおいて $150 * 150$ ピクセルのリサイズ処理の場合には 88% と高い正解率となることが確認できた. 従って, 今回使用したコミックにおいては, 提案手法を用いることで主要キャラクターの同定が可能となった. しかし, 一部の結果において, 所属度 1 位と 2 位の結果が等しいキャラクターを抽出していた. そのため, 今回の実験では等しいキャラクターが検出された場合に, 所属度 3 位となるキャラクターを結果として用いた. この結果は, 提案手法において自動同定の不完全さを表している. 現状では提案手法のみを用いて主要キャラクターを同定することは困難である.

主要キャラクターの自動同定に提案手法を使用する為には, パラメタや今回用いた手法などの改善を行い, 正解率の更なる向上を目指す必要がある.

また本実験において, 考えられる改善余地が 2 点ある.

1 点目は, 提案手法に最適なパラメタが明確に定められていないことである. パラメタの最適化のために, 更に多くのコミックを使いデータを集める必要があると考えている. パラメタの最適化が進めば, 明確にキャラクターの特定が行われ, ノイズデータが減少する可能性がある. また, 今回使用したジャンル以外の場合では提案手法が全く機能しない可能性もある. 従って, より多くのジャンルに属するコミックを使用し提案手法の性能確認実行する必要がある.

2 点目は, 主要キャラクターの定義やキャラクターの分類における調査が十分ではないことである. 本実験で主要キャラクターとして抽出された結果が, 明確に正解であるかの妥当性を検証する必要がある. この点についても, 更に多くのコミックを実験対象とすることで, 明確な根拠を得る必要がある. また提案手法において, リサイズ処理におけるサイズ設定が実験結果に大きく影響を与えることが分かった. リサイズ処理においてある程度大きなサイズを設定することにより, キャラクター顔の特徴を抽出しやすくなる傾向があ

る.

本研究の成果より, **SURF** 特徴はコミックにおけるキャラクター顔に対しても有効であると考えられる. 本研究では, **BoVW** 手法と **SURF** を用いてコミックにおけるキャラクターの同定を行ったが, **SURF** 特徴を別の手法と組み合わせることで, 更に精度の良い結果が得られる可能性もある.

第6章

結論

6.1 総括

本論文では SURF 特徴と BoVW 手法を用いて、コミックにおける主要キャラクターの同定を行った。その結果、実験で使用したコミックに関して 150×150 ピクセルのキャラクター顔画像を抽出することで 88% の正解率で主要キャラクター同定が可能になったことがわかった。この結果から、提案手法を用いることで事前知識である対象コミックにおける登場キャラクター情報を必要としない主要キャラクターの同定が可能であることを確認できた。

6.2 課題

現状では本論文で提案する手法において、パラメタの最適値が選択出来ていない観点から主要キャラクター同定の結果が不完全である。従って、より多くのコミックを用いて実験をする必要がある。今後の課題として以下の点が挙げられる。

- ・コミックにおけるキャラクター顔識別器の作成

コミックからキャラクター顔検出を行う際、公開された汎用な識別器では非一様なキャラクターなどに対応していない。そこで、大量のコミックよりキャラクター画像を学習し全てのパターンに対応する識別器を作成する必要がある。

- ・SURF と BoVW の最適なパラメタ設定の決定

提案手法におけるパラメタの設定が最適ではない。設定を変更した実験データを集め、明確な根拠のあるパラメタの最適値を調査する必要がある。

・主要キャラクターの定義について

本研究において、主要キャラクターを主観的な判断で定義した。しかし、登場キャラクターが多く、全てのキャラクターの登場回数に差がないコミックの場合には、明確に登場回数が増減しないため、主要キャラクターの判断が困難である。従って、何らかの手法により主要キャラクターの定義を明確に定める必要がある。

謝辞

本研究にあたり素晴らしい研究環境を与えて下さり、また貴重な時間を割いて丁寧にご指導していただいた渡辺裕教授に感謝致します。

並びに、テーマの決定から始まり、既存手法におけるアルゴリズムの理解への助言、日々研究の進め方についてご意見を頂き、本研究にご助力頂きました渡辺研究室の皆様へ感謝致します。

本研究を行うにあたり、コミック画像の提供および論文への掲載を許可して頂いた東京大学相澤・山崎研究室の皆様へ感謝致します [36]。

また、実験を行うにあたり、プログラムの使用を許可して頂いた皆様へ感謝致します。

最後に、本研究を行うにあたり、ご助力頂いた全ての皆様へ深く感謝致します。

参考文献

- [1] "平成 24 年情報通信白書", 総務省, <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h24/html/nc122110.html>, 参照 Jan.16, 2016.
- [2] 特別企画: "出版関連業者の経営動向調査", 帝国データバンク, Jul.23, 2015.
- [3] インプレス総合研究所: "電子書籍ビジネス調査報告書 2015", 株式会社インプレス, Jul.30, 2015.
- [4] 講談社: "講談社からのお知らせ", <http://www.rbbtoday.com/article/2015/01/05/127027.html>, Jun.5, 2015.
- [5] 安田つくし, 三原鉄也, 永森光晴, 杉本重雄, コミックマーケットのジャンルに関するメタデータを利用した同人誌探索支援ツールの提案, 情報処理学会第 76 回全国大会講演論文集 2014(1), 541-542, 2014.
- [6] 南澤久佳, "魔法使い養成専門マジックスター学院 トリプルスター 1", Manag109, 2015.
- [7] 菅野博士, "ばらいそロード", Manag109, 1993.
- [8] 咲坂伊緒, "アオハライド 1", マーガレットコミックス, 2011.
- [9] サンカクヘッド, "干物妹うまるちゃん 1", ヤングジャンプコミックス, 2013.
- [10] 赤松中学, "緋弾のアリア AA (1)", ヤングガンガンコミックス, 2011.
- [11] 赤松 健, "ラブひな 1 巻", Manag109, 1999.
- [12] Q.V.Le, Builing high-level features using large scale unsupervised learning, ICASSP, 2013 IEEE International Conference, pp.8595-8598, 2013.
- [13] G.Hinton, L.Deng, D.Yu, G.Dahl, A.Mohamed, N.Jaitly, A.Senior, V.Vanhoucke, P.Nguyen, T.Sainath and B.Kingsbury, Deep Neural Networks for Acoustic Modeling in Speech Recognition, IEEE Signal Processing Society, pp.82 -97, 2012.
- [14] C.P. Papageorgiou, M.Oren, T.Poggio, A General Framework for Object Detection, Computer Vision, 1998.Sixth International Conference, pp.555-562 ,1998.
- [15] 山田剛士, 呉梅元, 和田俊和, GA と Adaboost を用いた顔検出, 電気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 106(469), pp.43-48, 2007.
- [16] 赤松健, 魔法先生ネギま! (14), 少年マガジン KC, 2006.
- [17] T.Ojala, M.Pietikänen, D.Harwood, A comparative study of texture measures with classification based on feature distributions, Pattern Recognition Society, Vol.29, No.1, pp.51-59, 1996.
- [18] AnimeFace, https://github.com/nagadomi/lbpcascade_animeface, 参照 Jan.16, 2016.

- [19] D.G.Lowe, Distinctive Image Features from Scale-Invariant Keypoints, Accepted for publication in the International Journal of Computer Vision November 2004, Vol.60, pp.91-110, 2004.
- [20] H.Bay, A.Ess, T.Tuytelaars, L.V.Gool, Speed-Up Robust Features (SURF), Computer Vision and Image Understanding 110(2008), pp.346-359, 2008.
- [21] P.Viola, M.Jones, Rapid Object Detection using a Boosted Cascade of Simple Features, Accepted conference on computer vision and pattern recognition 2001, pp.I 511-I 518, 2001.
- [22] 柳井啓司, 一般物体認識の現状と今後, 情報処理学会論文誌 CVIM48, pp.1-24,2007.
- [23] G.Wang, Y.Zhang, L.Fei-Fei, Using dependent regions for object categorization in a generative framework, Computer vision and pattern recognition, IEEE Computer Society Conference, Vol.2, pp.1597-1604, 2006.
- [24] K.Barnard, P.Duygulu, R.Guru, P.Gabbur, D.Forsyth, The Effects of Segmentation and Feature Choice in a Translation Model of Object Recognition, Computer vision and pattern recognition, Vol.2, pp.675-682, 2003.
- [25] G.F.Pineda, H.Koga, T.Watanabe, Scalable Object Discovery : A Hash-Based Approach to Clustering Co-occurring Visual Words, IEICE Transactions, Vol.E94-D, No.10, pp.2024-2035, 2011.
- [26] A.K.Jain, Data clustering : 50 years beyond K-means, Pattern Recognition Letters Vol.31, pp.651-666, 2010.
- [27] 獅子君, 古賀久志, 顕著特徴領域を利用した BoVW ベース類似画像検索の改善方式の検討, IEICE Technical Report PRMU2013-110, 2014-CVIM-190(30), pp.1-6, 2014.
- [28] クラスタリングツール bayon, <https://github.com/fujimizu/bayon>, 参照 Jan.16, 2016.
- [29] 咲坂伊緒, "アオハライド 2", マーガレットコミックス, 2011.
- [30] 水城せとな, "失恋ショコラティエ 1", 小学館, 2009.
- [31] 水城せとな, "失恋ショコラティエ 4", 小学館, 2011.
- [32] 石田スイ, "東京喰種 1", ヤングジャンプコミックス, 2012.
- [33] 石田スイ, "東京喰種 2", ヤングジャンプコミックス, 2012.
- [34] 赤松中学, "緋弾のエリア AA (2)", ヤングガンガンコミックス, 2011.
- [35] サンカクヘッド, "干物妹うまるちゃん 2", ヤングジャンプコミックス, 2013.
- [36] Y.Matsui, K.Ito, Y. Aramaki, T.Yamasaki, K. Aizawa, Sketch-based Manga Retrieval using Manga109 Dataset, arXiv:1510.04389, 2015.

図一覧

図 1	キャラクター紹介ページ例	4
図 2	登場キャラクター分類	5
図 3	対象外のキャラクター例	6
図 4	アオハライドにおける登場回数比較	9
図 5	干物妹うまるちゃんにおける登場回数比較	9
図 6	緋弾のアリア AA における登場回数比較	10
図 7	ラブひなにおける登場回数比較	10
図 8	非一様な特徴を持つ登場キャラクター例	13
図 9	LBP の算出例	14
図 10	box filters 例 1 [20]	17
図 11	box filters 例 2 [20]	17
図 12	BoVW 概要図	19
図 13	コミックにおけるキャラクター顔画像 SURF 特徴例[16]	20
図 14	提案手法概要図	23
図 15	キャラクターA	26
図 16	キャラクターB	26
図 17	ノイズデータ例	33

表一覧

表 1	使用コミック一覧	6
表 2	階層毎に分類したキャラクター名称と登場回数 1	7
表 3	階層毎に分類したキャラクター名称と登場回数 2	7
表 4	階層毎に分類したキャラクター名称と登場回数 3	8
表 5	階層毎に分類したキャラクター名称と登場回数 4	8
表 6	パラメタ設定	24
表 7	使用コミック一覧表	27
表 8	パラメタ設定	27
表 9	アオハライド 1 巻, 2 巻	28
表 10	失恋ショコラティエ 1 巻	28
表 11	失恋ショコラティエ 4 巻	29
表 12	東京喰種 1 巻	29
表 13	東京喰種 2 巻	29
表 14	緋弾のARIA AA 1 巻	30
表 15	緋弾のARIA AA 2 巻	30
表 16	干物妹うまるちゃん 1, 2 巻	31
表 17	評価表	31
表 18	100*100 ピクセルの場合における主要キャラクター同定結果	32
表 19	150*150 ピクセルの場合における主要キャラクター同定結果	32

研究業績

題目	発表年月	発表	連名者
コミックにおける 主要キャラクター同定の検討	2016年3月 15日(火)	電子情報通信学会 2016年 総合大会(講演分野:D-21 マ ルチメディア情報ハイデ ィング・エンリッチメント)	長尾一輝, 渡辺裕