

マンガ固有の特徴を利用した マンガ登場人物識別に関する一検討

石井 大祐¹ 山崎 太一¹ 渡辺 裕¹

概要: 近年マンガは電子書籍市場における重要なコンテンツである。一方でマンガ画像の解析は現状あまり進んでいない。我々はマンガ画像における自動メタデータ生成技術の実現を目指している。本稿では、マンガの登場人物の顔画像に対する識別処理を行う手法を提案する。HOG と SVM による判別処理に対する学習と判別の為のマンガの特徴に基づく顔のパターン分類を示した。実験により、パターン分けを利用しない場合で本人に対して 80%、他人に対して 40%程度の識別率が得られ、パターン分けを用いた場合には、正面向きの場合および後ろ向きの場合に、より良好な結果が得られる事を確認した。

A Study on Character Recognition Utilizing Specific Features of Comics

ISHII DAISUKE¹ YAMAZAKI TAICHI¹ WATANABE HIROSHI¹

Abstract: Recently Japanese comics is important contents in electronic book market. On the other hand comic analysis is now developing. We aim to create auto metadata extraction method for comics. In this paper we propose face recognition process. We propose face pattern classification relate to specific features of comics for prediction process by HOG and SVM. We obtained 80% true positive ratio and 40% true negative ratio by examination without face pattern classification. Then we obtained more good result with face pattern classification.

1. はじめに

近年デジタル書籍閲覧に利用可能な端末が各社から発売されるなど、デジタル出版の発展が顕著である。また、その利用者数及びコンテンツ数は飛躍的に増大している。特に、日本においては、デジタル書籍コンテンツの中でマンガの占める割合は大変大きく、マンガの内容取得やアクセシビリティを向上させることは有意である。一方で、マンガ画像の解析技術については未だ発展途上である。

我々はメタデータの自動付与を目的としたマンガの解析手法を確立することを目的としている。マンガのメタデータとしては、マンガの読み順やシーンの移り変わりを示すコマ、コマの内部に描かれる登場人物及びテキストからなるセリフなどが挙げられる。また、登場人物はマンガの構

成上重要である。

マンガは多くの場合白地の上に黒色のインクやトーンを用いて描かれるため、これをデジタル化すると、白黒 2 値のデータとなる。低解像度化を行うことで一部中間調が出てくるが、全体としてはほぼ白黒で描かれている画像であると捉えることができる。これが、マンガ画像の解析が困難である要因の一つとなっている。

これまで、人物検出 [1] や、一般物体認識などにも利用可能な様々な画像特徴量 [2, 3] 及び、その統計的機械学習手法 [4] 等が検討されてきた。特に、人物の顔はその濃淡や色、エッジ等をうまく捉えることでその検出や認識等が行われてきた。一方で、マンガは前述の通り、人物画像とは異なる 2 値の線や点によって描かれており、一般的な人物検出等の処理がうまく働かない。

マンガに対する研究では内容要約を目的とした手法 [5] や、Histograms of Oriented Gradient(HOG) 特徴量 [3] と Support Vector Machine(SVM) [4] を使ったアプ

¹ 早稲田大学 大学院国際情報通信研究科
Graduate School of Global Information and Telecommunication Studies, WASEDA University, 1011 Nishi-Tomida, Honjo-shi, Saitama 367-0035, Japan



図 1 マンガ画像例 [9]
Fig. 1 A sample of
manga image [9].



図 2 マンガ画像上の
SIFT 特徴量
Fig. 2 SIFT Feature
on Comic Image

ローチ [6, 7] が検討されている。しかし、マンガの登場人物の識別は未だ実現されていない。

マンガの登場人物識における問題点として、マンガにおける表現として表情及び方向による顔形状のドラスティックな変化がある。今回は登場人物の識別として、HOG と SVM を利用した識別を土台として、表情変化および顔の向きの変化という特徴による影響と、これを考慮した顔識別について検討を行った。

本稿の構成は以下のとおりである。2 章ではマンガの画像的特徴と、一般的な画像解析手法及びマンガ画像解析の現状について述べる。3 章では、今回我々が着目したマンガの特徴と、検討するマンガの登場人物識別手法について述べる。4 章では、実際にマンガ上の登場人物の識別を行い、その結果を示す。5 章は本稿のまとめである。

2. マンガ画像解析

2.1 マンガについて

マンガは、主に線とトーンを用いて描かれる絵と、文字によって構成される。マンガのシーンはコマと呼ばれる枠により分けられ、コマの並びによりページ内の読み順が決められる。各コマの中には今回の識別対象である登場人物、背景などの絵、セリフと吹き出しなどの要素が描かれる。また、しばしば表現の為に、コマからの要素のみ出しが存在する。マンガ画像の例を図 1 に示す。

2.2 一般的な画像解析技術

人物の検出や一般物体認識等を目的として、これまで様々な画像特徴量と、これを利用した画像解析手法が検討されてきた。ここでは、その中で代表的な手法について、その概要を示し、マンガ画像への利用について述べる。

顔検出に対する研究として、Viola と Jones は、高速かつ頑健な顔検出手法を提案した [1]。この顔検出手法では、画像特徴量として、様々な形状の Haar-like 特徴量を用い

ることで、主に人の目をはじめとする顔の各部位の濃淡から、人物の顔領域を検出する。マンガ画像では、基本的に濃淡情報ではなく、線によって描かれる。作品により、特性はやや異なるが、主に線画による表現のマンガ画像からは、Haar-like 特徴量を用いても有意な情報を取り出すことが困難である。

Scale-Invariant Feature Transform (SIFT) 特徴量 [2] は、スケール変化や方向の変化に対して頑健なキーポイントとその特徴量を同時に求める手法である。Difference of Gaussian (DoG) から求めた各スケール間における変動の局値の位置を求めるため、スケールに対して頑健にキーポイントの検出が行える。これにより、特定物体認識や、画像の対応点検出および特定物体の追跡などに利用可能である。SIFT をマンガ画像に対して適用した結果の一部を図 2 に示す。このように、マンガ上には様々なエッジの成分が無数に存在するため、SIFT によりキーポイントと識別されてしまう点の数が多く、これを制御し有意な情報のみ取り出すことは困難である。

Histograms of Oriented Gradient (HOG) 特徴量 [3] は、画像上のエッジ成分に着目した手法であり、HOG は物体の形状を捉えるのに向いた特徴量であるため、人物のシルエットによる人検出などに利用可能である。マンガ画像のキャラクターは線画により描かれているため、エッジより形状が形成されている状態である。そのため、概形を捉えやすい HOG 特徴量は、マンガ画像の解析に有用であると考えられる。

2.3 マンガ画像解析

マンガ画像解析について、これまでにいくつかの研究がなされている。また、前節の最後に述べた HOG 特徴量を利用したマンガ画像の解析手法が検討されている。

帆足ら [5] は、マンガのシーン要約を目的とした研究の中で、自動でコマの分割とセリフ部分の文字および吹き出し部分の検出を行う手法を提案している。この中で、コマの分割時には、分割線上のエッジの方向性を利用している。エッジの方向性は、1 次元化された HOG 特徴量とほぼ等価であると言える。

我々は、これまでに、HOG 特徴量と Support Vector Machine (SVM) [4] を利用した、マンガ画像上における登場人物の瞳検出について検討を行った [6]。マンガ画像上の登場人物は顔のパリエーションや、オクルージョンによる影響を多々受ける。この影響により、顔全体を学習し検出を行うよりも、瞳部分のみを検出する方が対象領域の検出再現率が高くなるという結果が得られている。

新井らは同じく HOG 特徴量と SVM からマンガ画像の登場人物の顔検出を試みている [7]。新井らの検討ではどの程度とれているとされる。この結果はどういうものであると捉えることができるマンガは作品ごとに絵の傾向は異

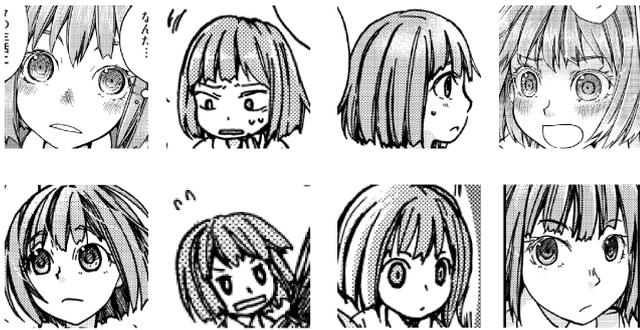


図 3 登場人物の顔の変動

Fig. 3 Variety of face on character.

なり、また、作品内においても顔のバリエーションが豊富である。このため、顔検出を統一された検出器で行うには限界があると考えられる。

3. マンガ画像の特徴を利用した登場人物識別

各登場人物は、多くの場合、人間の読み手には判別可能となるように描かれている。本研究では、顔部分の切り出しが行えているという条件の下、切り出された顔画像の識別を目的とする。

マンガ画像の主要成分は線画で描かれるため、方向性の特徴を利用することが可能であると考えられる。また、顔の構造的特徴等は Haar-like 特徴量では取得が困難である。このことからマンガの人物識別にも HOG 特徴量を利用することが考えられる。そこで我々は、顔画像に対する基本的な識別処理として HOG 特徴量と SVM による判別処理を利用する。これに加え、今回はマンガ表現の特徴を加味した判別処理について検討を行う。

3.1 マンガ画像における登場人物の特徴

マンガではストーリーを、セリフ等の文字だけでなく絵を用いて表現することから、登場人物の顔の表現が様々に変化する。マンガ内に登場する登場人物の画像例を図 3 に示す。また、マンガにおける登場人物の描かれ方に関する特徴を以下に示す。

- (1) 顔の向きによる離散的かつ劇的な変化
- (2) シーンによる表情の変化
- (3) キャラクターのデフォルメ

(1) の特徴より、顔の方向ごとに得られる特徴量はまったく異なるものとなる。このため、たとえ同一人物の顔であっても、一つの判別器にすべてのパターンをまとめて学習することは、誤判定を招く要因となりうる。(2),(3) の特徴は、同一人物でかつ方向の顔であっても、画像的に離れた表現が存在し、かつ (1) の特徴同様に、表情や表現の間のつながりについては原理的に取得不可能である。

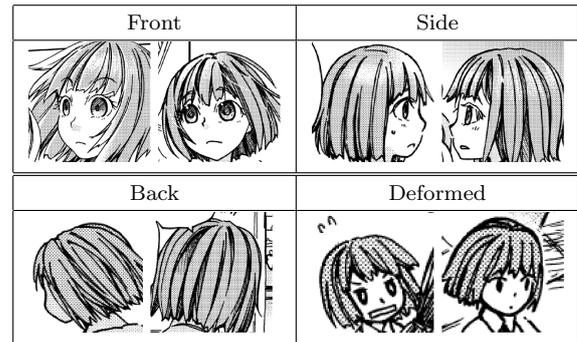


図 4 登場人物の顔パターン分類

Fig. 4 Claccification of face pattern.

3.2 判別処理アルゴリズム

上記特徴を総合的に考慮し、我々は、顔画像をいくつかのパターンに分類し、各パターンごとに学習を行う識別処理を検討した。

前節の特徴を考慮し、今回は 4 パターンへの分割を行った。各パターンの画像例を図 4 に示す。Front は正面向きの画像である。顔の表情変化は正面向きの場合だけでも多数取得可能であるため、これを含む最も大きな要素数を含む分類となる。Side は横向きの顔であり、顔の表情変化は正面向きについて多い。右向きと左向きがあるが、今回は反転処理により左向きに統一して学習を行った。Back は後ろ向きであり顔の表情変化は無い。Deformed はマンガ特有の、簡略化された表現であり、画像上における大きさや描かれる際の線数、描き方の特徴などが異なる。表情の種類も豊富であるが、一方で登場回数は少ない。

4. 登場人物識別実験

4.1 実験の概要

初めに、顔画像のパターン分けを行わない場合の識別性能について実験を行い、次いで各パターンに対して実験を実施した。共通の条件を以下に示す。実験には 1 作品 26 ページのマンガ画像から切り出した顔画像群を利用した。ここで、学習には HOG 特徴量を用いるが、HOG 特徴量はシフト不変性が無いため、同一の顔画像に対して位置をずらした複数の切り出しを行っている。学習数、判別数はこの複数の切り出し画像をすべてカウントしたものである。

4.2 登場人物の表情等を考慮しない場合

顔画像の複数パターンを含む場合の識別性能について実験を行った。Positive として主人公を、Negative としてそれ以外の人物を用いて学習を行った。学習および実験に用いた画像枚数を表 1 に、結果を表 2 に示す。All は顔画像すべてのパターンを含む。FrontOnly は学習用および実験用の画像集合から Side 及び Back を取り除いたものである。

実験結果より判別器を作成する際に、顔画像のパターン

表 1 学習と実験に用いた画像数

Table 1 Image number utilized for learn and examination.

Data Set	Learn		Unknown	
	Positive	Negative	Positive	Negative
All	50	50	135	68
Front Only	50	39	40	30

表 2 識別成功率

Table 2 Recognition Rate.

Data Set	Learn		Unknown	
	Positive	Negative	Positive	Negative
All	1.00	0.98	0.79	0.44
Front Only	1.00	0.85	0.88	0.47

表 3 学習と実験に用いた画像数

Table 3 Image number utilized for learn and examination.

Data Set	Learn		Unknown	
	Positive	Negative	Positive	Negative
Front	60	39	71	133
Side	19	52	9	223
Back	6	14	5	278
Deformed	9	20	8	266

を分けない場合には、顔の正面のみの場合と比べて、学習対象の分散が大きいため、その識別結果が悪化したものと考えられる。

4.3 登場人物の顔の向きおよびデフォルメを考慮した場合

顔画像の複数パターンを含む場合の識別性能について実験を行った。学習および実験に用いた画像枚数を表 3 に、結果を表 4 に示す。実験では、本人の当該パターンのみを Positive、他人の顔および本人の当該パターン以外を Negative として学習及び判別を実施した。ここで、Front 以外の各パターンの学習画像数が少ない理由は、マンガにおいてその登場回数が少ないためである。

実験結果より、Front については、全パターンを学習した場合よりも高い検出精度が得られた。また、その他のパターンにおいても他人に対する識別成功率は高く、とりわけ Back の場合には本人および他人の両方に対して高い識別精度が得られた。Side と Deformed では本人に対する識別成功率が低い結果となっている。双方の学習に用いた画像群に対する識別結果は良好である。Side と Deformed については、顔のパリエーションが多く、一方で登場回数はそれほど多くないため識別対象に対する学習のパリエーションが少ないことが誤判別の原因となっていると考えられる。Back も学習画像数が少ないが、Back の場合には識別対象もあまり変化しないため高い識別率となっている。

5. おわりに

本稿では、マンガの登場人物の顔画像に対する識別手

表 4 各パターンにおける識別成功率

Table 4 Recognition rate of each pattern.

Data Set	Learn		Unknown	
	Positive	Negative	Positive	Negative
Front	1.00	0.97	0.80	0.60
Side	1.00	1.00	0.44	0.96
Back	1.00	1.00	1.00	0.83
Deformed	1.00	0.95	0.38	0.80

法の検討を行った。HOG と SVM を用いた判別機により、あらかじめ切り出された顔画像に対して、そのまま学習を行った場合と、マンガの特徴を用いパターン分けを行うアプローチを用い、双方を用いて実際に識別実権を実施した。実験により、パターン分けをしない場合で本人に対して 80%、他人に対して 40%程度の識別率が得られ、パターン分けをした場合には、正面向きの場合および後ろ向きの場合に、上記より良好な結果が得られた。一方で、横向きとデフォルメされた場合にはまだ識別率が低く、学習枚数などに改善が必要であると考えられる。全体の識別率向上の為に、各パターンで十分な学習枚数が得られる場合には、パターン分けを行った方が良好な識別結果を得られると考えられる。

謝辞 本稿にて例示したマンガ画像はすべて木野陽様 <http://www.etheric-f.com/> より学術目的の為に使用許可をいただいたものである。マンガ画像の提供及び原稿への掲載を許可いただいた木野陽様に深く感謝する。本研究は JSPS 科研費 12018464 の助成を受けたものである。

参考文献

- [1] P. Viola and M. Jones: Robust Real-Time Face Detection, Trans. IJCV, 57(2), pp.137-154, 2004.
- [2] D. G. Lowe: Object Recognition from Local Scale - Invariant Features, International Conference on Computer Vision, pp.1150-1157, 1999.
- [3] Dalal.N and Triggs. B: Histograms of Oriented Gradients for Human Detection, IEEE CVPR, pp.886-893, 2005.
- [4] C. Cortes and V. N. Vapnik: Support-vector networks, Machine Learning, vol.20, pp.273-297, 1995.
- [5] K. Hoashi, C. Ono, D. Ishii, H. Watanabe: Automatic Preview generation of comic episodes for digitized comic search, MM'11 Proceedings of the 19th ACM international conference on Multimedia, pp.1489-1492, Nov. 2011.
- [6] 石井大祐, 渡辺裕: マンガからの自動キャラクター位置検出に関する検討, 情報処理学会 AVM 研究会研究報告, Vol.2012-AVM76, No.2, pp.1-5, Feb. 2012.
- [7] 新井俊宏, 松井勇佑, 相澤清晴: 漫画画像からの顔検出, 2012 電子情報通信学会総合大会, d-12-67, pp.161, Mar. 2012.
- [8] C. C. Chang and C. J. Lin: LIBSVM: a Library for support vector machines, ACM Transactions on Intelligent Systems and Technology, 2:27:1 - 27:27, 2011.
- [9] 木野陽: ベリーベリークリームショコラ ふたつのベリー, 2010.