

マンガ上のキャラクター識別に関する一検討

石井 大祐[†] 山崎 太一[†] 渡辺 裕[†]

早稲田大学大学院 国際情報通信研究科[†]

1. まえがき

近年、マンガは、日本国内ばかりでなく、海外からの評価も高いコンテンツである。また、電子書籍市場におけるマンガの占める割合も大変大きい[1]。その一方で、マンガ画像の解析技術については、未だ発展途上である。マンガは、通常の写真等と異なる画像特徴を持ち、また、作家によりその描かれ方が異なるため、一般的な画像特徴量を用いた顔検出手法[2]では有用な情報を取り出す事が困難である。マンガにおいて、キャラクターはストーリーを構成する上で非常に重要な役割を担う。今回、我々は、マンガキャラクターの顔が入力された場合に、そのキャラクターの識別を行う手法について検討を行った。

2. マングの画像特徴と画像解析手法

これまで、顔画像に対する検出、認識について研究が行われてきた。現在の顔検出の主流である Viola と Jones による手法[2]では、Haar-like 特徴量を利用した解析が行われている。人の顔画像には、顔の部位ごとにその輝度が異なるため、これを解析の手掛かりとして利用している。一方で、マンガの場合には、顔は線によって描かれ、肌部分は実質空白である場合がほとんどであり、平均輝度をとらえる Haar-like 特徴量では、マンガの顔特徴を捉えにくいと考えられる。

3. マングキャラクターの顔識別法

本稿では、あらかじめ切り出された顔領域に対して、これをキャラクターごとに学習し、識別を実施する方法を提案する。マンガのキャラクターには、グラデーション等による濃淡情報は少ない。このため、我々は、エッジの方向性をとらえる Histograms of Oriented Gradient (HOG) 特徴量[3]を用いる。

マンガ画像上に描かれるキャラクターの顔サイズはさまざまである。キャラクターの瞳検出に関する我々の研究[4]より、HOG 特徴量計算時に解像度変換を実施しても、識別結果に対する

影響が少ないことが分かっている。そこで、本研究では、HOG 特徴量を同一条件で計算するために、切り出された顔画像に対して、同一解像度への解像度変換を実施したのち、HOG 特徴量の計算を行うこととする。

キャラクター識別では、本研究の目的から、当該キャラクターの登場後数十枚程度を正例、当該キャラクター以外の顔画像数十枚を負例として学習を行う。少ない学習枚数で判別結果を得るために、学習には Chang らによる LIBSVM[5]を用いた。

4. キャラクター顔識別実験

実際に、漫画に登場するキャラクターの顔部分を切り出し、識別実験を行った。はじめに、2 作品における顔画像の識別を行い、次に、その中の 1 キャラクターに対して、学習枚数による識別結果への影響を調査した。

本実験で利用するすべての顔領域は、マンガ画像より手動で切り出したものである。学習はキャラクターごとに、検出対象キャラクターの顔画像を正例、当該作品における検出対象キャラクター以外の顔画像を負例として行った。各登場人物は、その登場回数が異なるため、学習のための枚数および判定実験用の未知画像枚数は各々異なる。学習及び実験に使用した画像枚数を表 1 および表 2 に、学習に使用した画像例を図 1 にそれぞれ示す。切り出された顔画像はその画素数が異なるが、今回は、HOG 特徴量計算時にすべて 50 x 50 画素に解像度変換を行っている。

実際にキャラクター識別を実施した結果を表 3 及び表 4 にそれぞれ示す。ここで、Learning Positive 学習に使用した本人の画像に対する識別の正解率を、Learning Negative は学習に使用した本人以外の画像に対する識別正解率を示し、それぞれ既知画像に対する True Positive に相当する。および True Negative の割合に相当する。Test Positive は、学習に使用していない本人の画像に対する識別正解率を、Test Negative は学習に使用していない本人以外の画像に対する識別正解率を示し、それぞれ未知画像に対する True Positive, True Negative の割合に相当する。

検出結果例の画像を図 2 に示す。表より、学習に使用した既知の画像ではほぼすべての結果で

表 1. 学習及び実験画像枚数 (作品 1)

Table1. Number of Learning and Test Image (Work1)

Character	Learning	Learning	Test	Test
	Positive	Negative	Positive	Negative
1-A	35	39	55	32
1-B	21	41	24	76

表 2. 学習枚数及び実験枚数 (作品 2)

Table2. Number of Learning and Test Image (Work2)

Character	Learning	Learning	Test	Test
	Positive	Negative	Positive	Negative
2-A	30	40	42	227
2-B	30	40	29	240
2-C	30	40	18	251
2-D	20	40	11	268
2-E	20	40	7	272

表 3. 顔識別結果正解率 (作品 1)

Table3. Success Ratio of Face Recognition (Work1)

Character	Learning	Learning	Test	Test
	Positive	Negative	Positive	Negative
1-A	1.00	0.95	0.84	0.60
1-B	1.00	1.00	0.63	0.99

表 4. 顔識別結果正解率 (作品 2)

Table4. Success Ratio of Face Recognition (Work2)

Character	Learning	Learning	Test	Test
	Positive	Negative	Positive	Negative
2-A	1.00	1.00	0.81	0.82
2-B	1.00	1.00	0.59	0.99
2-C	1.00	0.98	0.67	0.94
2-D	0.90	1.00	0.91	0.90
2-E	1.00	1.00	1.00	0.99

100%の検出率を得られている。一方未知の画像に対しては、キャラクターにより正解率にばらつきが得られた。これはキャラクターの表情パターンに偏りがあるため、この影響で判別可否が分かれたものと考えられる。

次に、学習枚数による識別結果への影響を調査するため、作品 1 のキャラクター 1-A を利用して、学習に使用した正例の数を 20 枚から 50 枚まで変化させて実験を実施した。結果を表 5 に示す。学習用の正例が増えるにしたがって、True Positive の値が増えているが、同時に False Positive の値も増える結果となった。学習用の正例を増やすと、学習時に多数の表情パターンを内包できる分、本人の表情のばらつきに対する許容性が増すが、同時に他人の顔に対する許容性も増加したため、False Positive も増えたと考えられる。

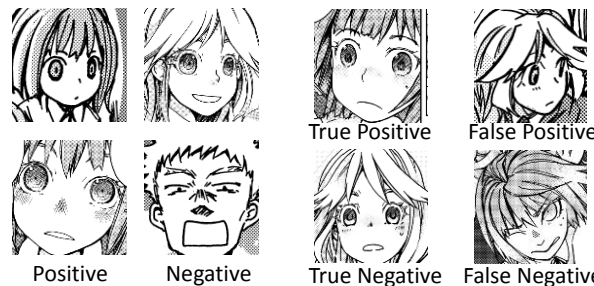


図 1 学習画像

Fig.1 Learning Image

図 2 識別結果

Fig.2 Recognition Results

表 5. 学習枚数と顔識別結果正解率 (1-A)

Table5. Learning Number Relative Success Ratio of Face Recognition (1-A)

Learning Positive Number	Learning Positive	Learning Negative	Test Positive	Test Negative
	20	1.00	1.00	0.69
30	1.00	0.95	0.72	0.67
40	1.00	0.92	0.82	0.57
50	1.00	0.85	0.88	0.47

5. むすび

本稿では、画像特徴量の一つである HOG と学習手法 SVM を用い、マンガのキャラクター識別を行った。実際の漫画に対するキャラクター識別実験により、未知の本人画像に対して 59% から 100%、未知の他人画像に対して 56% から 99% の識別成功率を得ることが確認された。

6. 謝辞

本稿にて例示したマンガ画像 [5] は、木野陽様 <http://www.etheric-f.com/> より学術目的の為に使用許可を頂いた物である。マンガ画像の提供およ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す。

参考文献

- [1] 高木ら, "電子書籍ビジネス調査報告書," 三橋昭和, インプレス R&D, 東京, 2009.
- [2] P. Viola and M. Jones, "Robust Real-Time Face Detection", Trans. IJCV, 57(2), pp. 137-154, 2004.
- [3] Dalal, N, Triggs, B, "Histograms of Oriented Gradients for Human Detection," IEEE CVPR, pp.886-893, 2005.
- [4] D. Ishii, T. Yamazaki and H. Watanabe, "Multi Size Eye Detection on Digitized Comic Image," IEEEJ 3rd Image Electronics and Visual Computing Workshop (IEVC 2012), Kuching, Malaysia, 1P-4, Nov. 2012.
- [5] 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.
- [5] 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.