

2値線画特徴量を用いたマンガ画像解析に関する一検討

石井 大祐¹ 張 傑¹ 石上 諒² 渡辺 裕¹

概要: マンガは2値線画からなる複雑な模様によって描かれている。従来、画像のエッジをとらえる画像特徴量として、HOGが使用されており、マンガにも一定の解析結果をもたらすことがわかってきている。その一方で、2値の線画を扱ううえで、HOG特徴量では形状をとらえるために冗長な次元がある。また、2値の斜線は縦横2方向の合成によって構成されることから、方向の取得にも難点がある。本稿では我々の提案する2値線画画像用に作成した特徴量 Histograms of Binary Oriented Pattern(HBOP)を用いてマンガ画像の解析を行う。マンガ画像解析において従来のHOG特徴量とHBOPを用いた場合の比較結果を示す。

キーワード: 画像解析, マンガ, 画像特徴量, 2値画像, HOG

A Study on Comic Image Analysis using Binary Line Image Feature

ISHII DAISUKE¹ ZHANG JIE¹ ISHIGAMI RYO² WATANABE HIROSHI¹

Abstract: Comic image was drawn into complex pattern using binary line. In conventional research, HOG it can treat image edge was utilized for to analyze Comic image as image feature. On the other hand HOG feature has redundant dimension for to treat binary line image. Moreover, diagonal line on binary image was consisted by combination of vertical and horizontal edge. Thus it is hard to get line angle. Histograms of Binary Oriented Pattern(HBOP) was proposed for to analyze binary line image by us. In this research, We analyze comic image using HBOP. Then, comparison results between HOG with HBOP on Comic image analysis was derived by examination.

Keywords: Image Analysis, Comics, Image Feature, Binary Image, HOG

1. はじめに

近年、マンガは日本国内のみならず、国際的に大変高く評価されているコンテンツとなっている。マンガがさらに国際的に発展する上で欠かせないのが、そのデジタル化と配信技術である。コンテンツ配信においては、コンテンツの内容を示す様々なメタデータを利用することで、その利便性を高め、コンテンツ価値の向上を図ることができる。

マンガのメタデータは、書籍一般に付与される書誌情報のみならず、マンガのストーリーや、登場人物等内容に関

するものまで多岐にわたる。また、マンガの場合には、絵によってその内容が描かれる。したがって、通常の文書コンテンツには必要のない位置情報等も加わり、メタデータにより記述され得る内容は豊富である。また、記述可能なコンテンツが多いことにより、そのメタデータ付与にかかるコストも膨大となる。文書であればOCRなどによるテキスト情報の抽出を行う、あるいはテキスト情報があらかじめ取得可能である。一方で、マンガは絵によって様々な要素が記述され、かつその配置も複雑なものであるため、解析に必要な技術レベルは飛躍的に高くなる。

自動的にマンガのメタデータを取得することができれば、より短時間でかつ高級なコンテンツを提供することが可能となる。このマンガメタデータを得るために、自動で内容を解析可能な技術が必要となる。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

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

² 基幹理工学部情報理工学科
Department of Computer Science and Engineering,
WASEDA University

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する。

これまで、様々な人物検出手法 [1-3] が提案されている。人物検出等においては、画像特徴量として、濃淡情報や、多値からなるエッジ等を捉えるものが利用されている。一方で、マンガは多くの場合白地の上に、黒色のインクもしくは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。これに対して、我々のこれまでの研究において、エッジを捉える画像特徴量である Histograms of Oriented Features(HOG) 特徴量 [3] を利用することで、ある程度の解析が行えることが確認されている。

HOG 特徴量は多値画像向けに設計された特徴量であり、そのエッジ記述性能は高い。しかし一方で、完全2値の画像を扱う上では、そのエッジ方向を記述するために冗長であり、また、特徴量の次元数も高い。2値画像向けに設計を行った画像特徴量を利用することで、マンガ画像解析の精度を向上できる可能性が考えられる。本稿では我々の提案している2値画像向け画像特徴量 Histograms of Binary Orientation Pattern(HBOP) を利用したマンガ画像解析について検討を行う。

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、これまでの検討でマンガに対して有効性を示した HOG 特徴量について述べる。3章では、我々の提案している HBOP 特徴量について述べる。4章では、実際にマンガを用いた識別実験を行い、その結果を示す。5章は本稿のまとめである。

2. マンガ画像の特徴と HOG 特徴量

2.1 マンガ画像

マンガは、コマと呼ばれるフレームの連続により、時系列を表現し、その中に絵画および文字、絵文字等を組み合わせ、ストーリーを構成した読み物である。絵画部分は登場人物から背景、視覚効果がともに同一領域に描かれるため、その構造は非常に複雑である。

マンガ画像は主に、白地に黒のインクと細かいドットが印刷されたトーンにより描かれる。また、近年は作画においてデジタル化される場合もあり、ペン型タブレット等による入力からベクタ化された線を取得し、これとデジタルトーン等を組み合わせて描かれる。紙、デジタルどちらにおいても、マンガ画像を構成する要素は白地と黒線および黒ドットの組み合わせが基本となる。この絵に文字等が加わり、最終的な出力となる。

線画の特徴を解析するために、線の形状を利用することが考えられる。線の形状は画像のエッジ等を取得することで得られる。また、一方で、2値画像に対する線形状解析手法の一つとして、モルフォロジフィルタ等の組みあわせにより作られる骨格成分抽出がある。細線化された骨格成

分は、形状記述の位置手法であり、文字解析等で利用されており、線により描かれる意味のある形状を解析する上で有効性がある。

2.2 HOG 特徴量

人物の検出や一般物体認識等を目的として、様々な画像特徴量が検討されてきた。ここでは、その中で従来なされた検討から、マンガ画像に対して一定の有効性を示した Histograms of Oriented Gradient (HOG) 特徴量 [3] 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。

我々はこれまでに、画像特徴量としてこの HOG 特徴量を利用し、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 [4, 5]。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 [6]。これらの研究成果により、HOG 特徴量のマンガ画像の解析への有効性が示されている。

HOG 特徴量は、画像上のエッジ成分に着目した手法である。HOG 特徴量の計算手順は次のとおりである。画像を任意数のセルに分割する。セル内における各画素について下記のルールに基づき方向を取得する。注目画素の上下左右方向の輝度差分から、その方向を計算する。セル内について、各画素の方向を量子化された方向成分の bin に投票を行うことで、画像内方向ヒストグラムを生成する。上記のようにして得られた複数セルを組み合わせブロックを構成し、ブロックごとに正規化を行う。HOG 特徴量の次元数 D は次の式によって得られる。

$$D = A \times C \times B \quad (1)$$

ここで、 A は量子化方向数、 C はブロックに含まれるセル数、 B は特徴量抽出画像に対するブロック数である。

HOG は物体の形状を捉えるのに向いた特徴量である。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り、このエッジを捉えることが可能な HOG 特徴量は、マンガ画像に対しても一定の適応性を持つ。一方で、2値画像のエッジを取り扱う上では、方向の計算方法の影響により、得られる方向数が限られる。また、少ない方向成分に対して冗長な次元数を持つこととなる。

3. 2値画像向け画像特徴量 HBOP

ここでは、我々が、2値線画像解析のために提案した画像特徴量である Histograms of Binary Orientation Pattern(HBOP) について述べる。HBOP は、名前の通り、HOG 特徴量の概念を継承し、2値線画像の取り扱いに特化した画像特徴量である。

HBOP では、HOG 特徴量でヒストグラム化を行っていた方向成分について、2値の細線から方向を取得可能であるように変更を行った。濃淡画像の場合と比較して平均輝

度等の影響が少ないことから、ブロック単位による正規化計算を省略した。これにより、2値線画像の特徴を捉えつつ、HOG 特徴量よりも次元数を削減している。

HBOP では、近傍線画素を図 1 に示す 12 パターンに集約する。いずれも中心画素が線を構成する画素である場合に、その周囲 8 画素内に現れるパターンによって、方向成分とし、各 bin へ投票を行う。投票の際に、出現する近傍画素の一部パターンについては集約を行う。これは、8 近傍によって表されるパターンのうち、水平垂直および ± 45 度以外については、同じ角度であっても、位置により 2 パターンに派生するためである。セル単位でヒストグラムを生成することで、画像特徴量を構成する。HBOP の次元数 D_{HBOP} は

$$D_{HBOP} = A_{HBOP} \times C \quad (2)$$

となる。ここで、 A_{HBOP} は HBOP 特徴量の方向数、HOG 特徴量と同様にセル数を示す。ただし、 A_{HBOP} はパターンを集約化した後の値である。我々のこれまでの検討により、本 HBOP 特徴量は、細線で描かれた単純図形の識別において、HOG 特徴量よりも良好な結果を得ることが確認されている。

4. マンガ画像解析実験

4.1 実験の概要

マンガ画像において、登場人物の抽出は重要な課題である。ここでは、登場人物の中で、多くの場合、画像のほか部分と比較して特徴的な描かれ方を有する瞳部分をターゲットとした。HOG 特徴量と HBOP 特徴量を用いて、マンガ画像の学習を行い、瞳とそれ以外の画像の識別率について実験により示す。

4.2 実験条件

実験には、360dpi で描かれたマンガ画像から、人手により切り出した画像を用いる。HOG 特徴量と HBOP 特徴量はともにシフト不変性を持たないことから、検出対象としている瞳部分について、画像内の同一部位から、切り出し領域を様々に変更し、取得したものを利用する。

学習用画像数は表 1 に示したとおりである。識別用の未知画像には正例 250、負例 250 を用意した。ただし、前述の通り、これらの中には切り出し範囲の異なる同一部位の画像が含まれる。学習には [7] にて公開されている RealAdaboost を用いた。また HOG 特徴量の計算も同様に [7] の実装を本実験に合うように調整し、使用した。学習回数は 200 回とし、学習時のセル数は 5×5 、HOG 特徴量のブロック数は 3×3 とした。また、HOG 特徴量の方向量子化数を 9、HBOP 特徴量の方向パターン数を 8 とした。HOG 特徴量、HBOP 特徴量の次元数はそれぞれ、 $D = 729$ 次元、 $D_{HBOP} = 200$ 次元である。

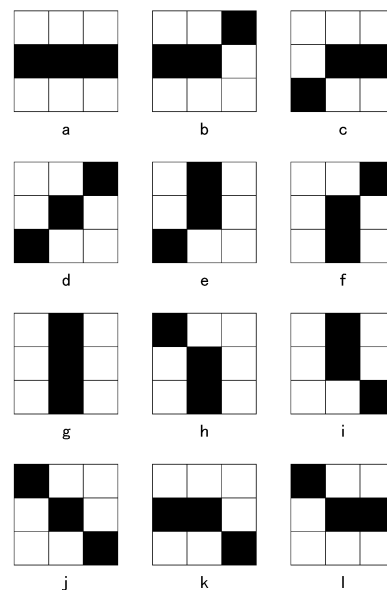


図 1 HBOP 特徴量の近傍画素パターン
Fig. 1 Neighbor pixel pattern of HBOP

表 1 学習と実験に用いた画像数

Data Set	Learn		Test	
	Positive	Negative	Positive	Negative
Image Number	368	499	250	250

4.3 マンガ画像識別実験

上記条件のもと、2手法において、画像特徴量計算時の正規化画像サイズおよび前処理を変更しつつ、識別実験を行った。

実験結果を表 2 に示す。正規化画像サイズは 50×50 , 100×100 , 150×150 , 200×200 の 4 段階とした。前処理は、BW が 2 値化を行ったもの、BWThin が 2 値化後に細線化を行ったもの、Canny はエッジ検出と細線化を同時に行う Canny フィルタ [8] を適用したものである。BW と BWThin については、HBOP 特徴量の線を白、背景を黒と定義しているため、画像の色反転を行っている。実験に用いた正例画像を図 2-図 4 に、負例画像を図 5-図 7 にそれぞれ示す。

実験結果より、HBOP 特徴量において正規化画像サイズを 150、前処理として 2 値化および細線化を行った場合最も高い平均識別成功率を示した。また、全体の中でも TruePositive が高い結果である。

表 2 において、TruePositive が高く、TrueNegative が低い場合、識別された結果の殆どを正であるとしているため、画像内探索を行った場合、得られる候補領域は非常におおくなってしまふ。一方で、TruePositive が低く、TrueNegative が高い場合には、識別器の基準が厳しく、本来正である部分に対してもあまり検出処理を行うことが出来ない結果であると考えられる。

一般的に、このような画像処理においては、識別器を力



図 2 実験画像:正例 (BW)

Fig. 2 Examination Image: Positive Sample(BW)

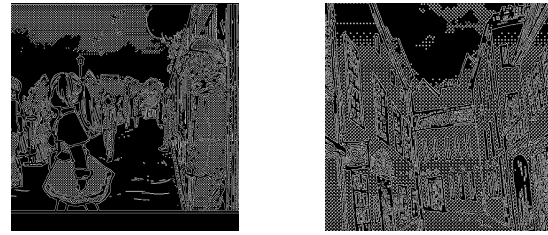


図 7 実験画像:負例 (Canny)

Fig. 7 Examination Image: Negative Sample(Canny)



図 3 実験画像:正例 (BWThin)

Fig. 3 Examination Image: Positive Sample(BWThin)



図 4 実験画像:正例 (Canny)

Fig. 4 Examination Image: Positive Sample(Canny)



図 5 実験画像:負例 (BW)

Fig. 5 Examination Image: Negative Sample(BW)

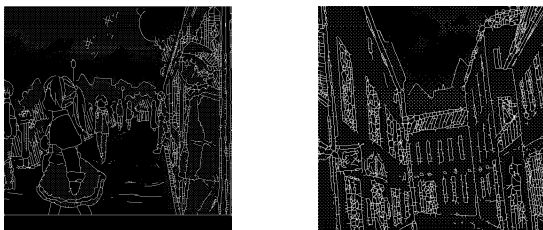


図 6 実験画像:負例 (BWThin)

Fig. 6 Examination Image: Negative Sample(BWThin)

スケードかすることで、多角的に識別を行い、結果の精度向上を図る場合がある。この観点からも、先の結果は、TruePositive の値が高いため、カスケード処理を構築するためにも利用可能であると考えられる。

5. おわりに

本稿では、マンガ解析の精度を向上させるために、マンガの解析に対して、2 値画像向け画像特徴量 HBOP の導入について検討し、これまで有効とされていた HOG 特徴量との比較を行った。あらかじめ切り出された瞳画像と背景画像に対して、HOG 特徴量 HBOP 特徴量双方で学習を行い、未知画像に対する識別実験を行った結果、画像に対して 2 値細線化を行った画像に対して、HBOP 特徴量を用いた場合に最も良好な平均識別率を得た。本パラメータによる解析結果は、TruePositive が高く、今後他の手法とのカスケード化にも利用可能であると考えられる。

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を許可されているものである。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: Rapid object detection using a boosted cascade of simple features, Proc. of Computer Vision and Pattern Recognition 2001, Vol.1, pp.I-511-I-518, Dec. 2001
- [2] P. Viola and M. Jones: Robust Real-Time Face Detection, Trans. IJCV, 57(2), pp.137-154, 2004.
- [3] Dalal.N and Triggs. B: Histograms of Oriented Gradients for Human Detection, IEEE CVPR, pp.886-893, 2005.
- [4] 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.
- [5] Daisuke Ishii and Hiroshi Watanabe: Multi Size Eye Detection on Digitized Comic Image, IEEEJ 3rd Image Electronics and Visual Computing Workshop (IEVC 2012), pp. 1P-4, 2012
- [6] 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.
- [7] 中部大学 藤吉研究室: "CV Tutorial," FLAB : Computer Vision for Visual Surveillance and Mobile Robotics, <http://www.vision.cs.chubu.ac.jp/CVTutorial/>, 参照 Dec. 2013
- [8] Canny John:A Computational Approach to Edge Detection, Pattern Analysis and Machine Intelligence, IEEE Transactions on, pp. 679-698, Nov. 1986
- [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.

表 2 識別結果
Table 2 Recognition Result

Method	Resolution	Preprocess	TruePositive	TrueNegative	Average
HBOP	50 × 50	BW	0.853	0.849	0.851
		BWThin	0.765	0.908	0.837
		Canny	0.410	0.924	0.667
HBOP	100 × 100	BW	0.964	0.450	0.707
		BWThin	0.880	0.912	0.896
		Canny	0.299	0.928	0.614
HBOP	150 × 150	BW	0.992	0.291	0.641
		BWThin	0.972	0.880	0.926
		Canny	0.088	0.960	0.524
HBOP	200 × 200	BW	1.000	0.247	0.624
		BWThin	0.964	0.805	0.884
		Canny	0.032	0.936	0.484
HOG	50 × 50	BW	0.896	0.896	0.896
		BWThin	0.729	0.829	0.779
		Canny	0.590	0.884	0.737
HOG	100 × 100	BW	0.861	0.912	0.886
		BWThin	0.841	0.813	0.827
		Canny	0.430	0.916	0.673
HOG	150 × 150	BW	0.869	0.869	0.869
		BWThin	0.805	0.884	0.845
		Canny	0.558	0.912	0.735
HOG	200 × 200	BW	0.861	0.888	0.875
		BWThin	0.833	0.865	0.849
		Canny	0.741	0.869	0.805