

高機能なマンガコンテンツのための
画像解析手法に関する研究

Research on Image Analysis Methods
for Advanced Comic Content

2014年7月

石井 大祐

高機能なマンガコンテンツのための
画像解析手法に関する研究

Research on Image Analysis Methods
for Advanced Comic Content

2014年7月

早稲田大学大学院国際情報通信研究科

国際情報通信学専攻

オーディオビジュアル情報処理研究II

石井 大祐

目次

表紙		A
目次		i
第 1 章	序論	1
1.1	研究の背景と目的	1
1.2	本論文の構成	3
第 2 章	マンガ画像の特性及び高機能なマンガコンテンツ生成に関する課題	5
2.1	まえがき	5
2.2	マンガの基本的な構成および特徴	5
2.3	マンガデータ生成に関する課題	6
2.3.1	マンガデータ生成に関する背景	6
2.3.2	マンガデータ生成の具体例	7
	マンガメタデータを用いたサービス（自動スクロールシステム）	7
	マンガメタデータを用いたサービス（構成要素に基づくサーチ用クエリの取得）	7
	マンガ制作支援システム（キャラクター作画支援）	7
	マンガ制作支援システム（ページ構成設計支援）	8
2.4	むすび	9
第 3 章	マンガ画像からのコマ検出	11
3.1	まえがき	11
3.2	文書画像のレイアウト解析	12
3.2.1	文書画像解析	12
3.2.2	一般化 Hough 変換	13
3.3	コマと視線遷移情報	13
3.4	再帰的 2 分割処理によるコマ分割手法	15
3.5	コマ分割処理の改良	16
3.5.1	コマ分割処理の改良について	16
3.5.2	帯状分割線検出	17
3.5.3	分割線上輝度こう配分布に着目した分割線選択	19

3.5.4	コマ内外識別	21
3.5.5	分割線検出アルゴリズムの改良による高速化	21
	高速化の概要	21
	高速高精度コマ分割処理手順 (FAFSM)	21
3.5.6	コマ位置精度の向上	23
	コマ位置精度向上手法の概要	23
	余白領域の検出	23
	コマ角位置候補検出処理と、コマ角位置の修正	24
3.6	コマ分割実験	25
3.6.1	画像解像度変換と分割パラメータによるコマ分割処理への影響	25
	コマ分割処理の評価法	25
	実験条件	27
	コマ分割速度及び分割精度に関する評価	28
3.6.2	コマ余白除去実験	29
	実験条件と評価法	29
	コマの余白除去に関する評価	31
3.7	むすび	33
第4章	マンガ画像からのキャラクター情報抽出	35
4.1	まえがき	35
4.2	人物検出手法	36
4.2.1	物体検出のアプローチ	36
4.2.2	人物検出処理と画像特徴量	36
	SIFT 特徴量	36
	Haar-Like 特徴量	37
	HOG 特徴量	38
	アニメキャラクター検出	39
4.3	マンガ画像からのキャラクター検出手法	39
4.3.1	キャラクター検出のアプローチ	39
4.3.2	瞳、顔及びキャラクター識別器とその組み合わせ	40
4.4	キャラクター検出実験	41
4.4.1	実験条件	41
4.4.2	キャラクター識別器に関する実験	42
	瞳、キャラクター及び特定キャラクター検出について	42
4.4.3	考察	44
4.5	キャラクターの顔の向き及びデフォルメを考慮した識別処理	46
4.5.1	マンガ画像におけるキャラクターの特徴	46
4.5.2	判別処理アルゴリズム	47
4.5.3	キャラクター識別実験	47

	実験の概要	47
	キャラクターの表情等を考慮しない場合	47
	キャラクターの顔の向き及びデフォルメを考慮した場合	48
4.6	むすび	49
第 5 章	2 値線画像向け画像特徴量 HBOP とマンガへの適用	63
5.1	まえがき	63
5.2	従来の画像特徴量と 2 値線画像に対する問題点	64
5.2.1	一般物体検出に用いられる画像特徴量の概要	64
5.2.2	Haar-like 特徴量	65
5.2.3	HOG 特徴量	65
5.2.4	EHOG 特徴量	67
5.2.5	Edgelet 特徴量	67
5.2.6	複数局所領域の関連性を用いる手法	69
5.3	HBOP 特徴量	70
5.3.1	HBOP 特徴量の概要	70
5.3.2	HBOP 特徴量の計算手順	70
5.3.3	検出処理時の計算量削減	74
	2 値化について	74
5.4	HBOP 特徴量による図形識別	75
5.4.1	HBOP 評価実験の概要	75
5.4.2	単純 2 値細線画像に対する識別実験	75
5.4.3	実験に関する設定及び実験環境	75
	HOG 特徴量及び HBOP 特徴量の学習特性	75
	未知画像識別実験	79
5.4.4	複雑な画像からの円検出実験	79
5.5	HBOP 特徴量によるマンガ画像識別	83
5.5.1	実験に関する設定及び実験環境	83
5.5.2	実験結果	83
5.6	むすび	86
第 6 章	結論	89
6.1	研究成果の概要	89
	第 3 章「マンガ画像からのコマ検出」	89
	第 4 章「マンガ画像からのキャラクター情報抽出」	90
	第 5 章「2 値線画像向け画像特徴量 HBOP とマンガへの適用」	90
	まとめ	91
6.2	本研究成果の応用用途に対する効果と考察	91
	マンガメタデータを用いたサービス（自動スクロールシステム）	91

	マンガメタデータを用いたサービス（構成要素に基づくサーチ用クエリの取得）	92
	マンガ制作支援システム（キャラクター作画支援）	92
	マンガ制作支援システム（ページ構成設計支援）	92
6.3	総括	93
	謝辞	95
	参考文献	97
	図目次	103
	図目次	107
	表目次	109
	表目次	110
	研究業績	111

第 1 章

序論

1.1 研究の背景と目的

マンガは日本のみならず世界的にもよく知られている代表的なコンテンツである。近年、マンガは日本政府によって、クールジャパン戦略の一部として扱われており、2009年に知的財産戦略本部コンテンツ・日本ブランド専門調査会により日本ブランド戦略 [1]、2011年にクールジャパン推進に関するアクションプラン [2] が示された。このように、マンガを含むコンテンツの国外発信に向けた方策が実施されており、マンガに対する社会的関心は大変高い。

近年の電子書籍市場において、マンガは有力なコンテンツであり、電子書籍市場全体の成長において大きな役目を果たしている [3-6]。電子書籍では、従来の紙媒体の書籍と異なり、コンテンツに対して様々な機能を持たせることが可能である。2013年には International Digital Publishing Forum (IDPF) により EPUB3.0 [7] が策定された。従来主流であった携帯電話のみならず、タブレットや専用端末などの電子書籍閲覧可能な端末数が増加している。電子書籍に対する社会的状況は整いつつあり、今後その普及が更に加速すると予想される。

マンガの電子書籍化において、単に紙媒体のマンガを電子化するだけでなく、様々な付加的機能の追加や新たな表現の導入が検討されている。効果音や画像の一部が振動するなどの視覚効果が付与された作品 [8,9] や、一部がアニメーション化された作品 [10] などがある。既存のマンガに対しても、その構成要素に基づく情報を加えることで利便性を向上させることが可能である。携帯電話向けサービスでは、コマごとに表示する方法やページ内部の読み順にそってスクロールする機能 [11] が付与されている。また、他メディアやコンテンツ内の他箇所へのリンク機能等は、マンガや文字ベースの電子書籍を問わず有用な機能であると考えられる。高機能なマンガコンテンツに対する需要は今後更に高まるものと考えられる。

本論文の目的は、マンガ画像からのメタデータ自動抽出を可能とする、マンガ画像解析処理を実現することである。マンガの構造は複雑であり、その構成要素も多い。マンガのインタラクティブ性を向上させるためには、マンガ内部の詳細な情報が必要となる。しかしその取得及び記述作業には多大なコストが必要であり、現実的に利用することが困難な要素も多い。そのため、マンガのアーカイブやコンテンツとしての利便性向上に利用可能な、マンガ画像の意味的要素を自動で取得可能、若しくは情報付与作業において作業者に補助的機能をもたらす解析手法を実現することは有意である。

マンガを取り扱う研究として、これまでに表現、芸術的側面から検討がなされている一方、画像

1.1 研究の背景と目的

工学分野における検討はまだ一部にとどまっている。表現及び芸術的側面からの検討として有名なものとして、夏目房之助による多数の書籍がある [12–15]。マンガから得られるメタデータの取り扱いについては、三原らのグループにより制作支援、情報管理、マンガの構造記述等複数の目的に対する検討 [16–23] が行われている。マンガ関連研究を網羅的に取り扱った研究として、松下らの調査 [24, 25] がある。マンガの制作支援技術として、2004 年度 IPA 未踏ユースに採択された POM [26] がある。POM はマンガの構図やコマ割りの推薦等を行うソフトウェアである。マンガコンテンツの作成支援に関しては、近年、主にコンテンツプロバイダ向けに、オペレータによる入力補助を行うソフトウェア [27–29] が開発されている。しかし、これらコンテンツ作成ソフトウェアは発展途上であり、原則的にオペレータによる多数の入力が必要である。デジタル化に伴うマンガの機能及び表現の進化のためには、今後更に画像工学分野における支援が重要であると考えられる。

これまでに、文書画像の解析を目的としたレイアウト解析処理に関する検討が行われている。ここで文書画像とは主に新聞や論文等の、明確に分離された文字と画像からなる画像である。文書画像では、基本的に行や文字の位置関係及びその角度に一定の規則が存在する。この規則を利用して、文字を構成する画素の並行投影結果やモルフォロジ処理結果を利用することで、要素ごとの領域情報を得ることができる [30–32]。また、対象画像から任意の形状を検出する、一般化 Hough 変換なども提案されている [33, 34]。一般化 Hough 変換では任意の検出対象図形テンプレートを用意し、パラメータ空間上での評価により検出を行う。これらの手法は、領域が十分に分離されている画像や、検出対象の形状が推測可能な場合に有効である。しかしマンガの構造は一般的な文書画像と比較して複雑であり、且つ異なる配置ルールに基づくものである。したがって、従来の文書画像解析に基づく解析手法では、マンガのレイアウトを原理的に解析できない。

マンガ画像の処理に特化した手法として、田中らによるコマ分割処理 [35] の研究がある。コマの枠線に沿った点では、コマの枠線に起因するエッジが存在するという特徴がある。そのため、コマの枠線に沿って引かれた直線上では、多くの点で高い濃度こう配値が得られる。この特徴を利用し、再帰的な画像の 2 分割処理によるコマ分割を行う。画像の 2 分割処理には、画像の 2 辺を結ぶ直線の中で、その直線上の濃度こう配計値が最大となる直線を用いる。このような処理を再帰的に実行することで、各コマへの分割と読み順の取得を行う。コマ分割の従来手法では分割線を決定する際に、濃度こう配値のみを用いるため、ロバスト性及び分割精度が低いという問題がある。更に、再帰処理の各過程において、画像上の 2 辺を結ぶ多数の直線に対して濃度こう配計値を求めるため、計算に係る処理コストが大きい。コマ分割処理の性能向上のためには、分割線の決定における信頼性と速度の改善が必要である。

画像解析分野では、画像特徴量と統計的機械学習を用いた一般物体認識が検討されてきた。顔検出及び人物検出を主な目的として、Haar-like や HOG をはじめとする様々な画像特徴量 [36–48] が提案されている。一般物体認識では、これら画像特徴量を Adaboost [49–51] や SVM [36, 51–54] などの統計的機械学習手法により学習する。電子化されたマンガは、基本的に白地の紙の上に黒色のインク及びトーンを用いて描かれるものであり、本来白黒 2 値のデータである。実際には、低解像度化や入力時のローパスフィルタの影響により一部中間調が出てくる場合がある。しかしその影響は基本的な特性を変えるものではない。したがって、マンガの画像の特性は基本的に全

て急峻なエッジによるものである。一方画像解析手法は多値画像を想定し設計されていることから、実画像に対して有効な解析手法をそのまま適用してもうまく解析が行えない。マンガ画像の要素を解析するために、マンガに対応した手法を検討する必要がある。

本論文では、マンガ画像からの具体的なメタデータ抽出処理として、コマおよびキャラクター情報の抽出を実現するためのマンガ画像解析手法について論ずる。また、マンガの主要構成である線の特徴を捉えるため、2 値線画像解析に利用可能な新たな画像特徴量として、2 値細線の方向ヒストグラムを利用した、Histograms of Binary Orientation Pattern (HBOP) 特徴量について論ずる。

1.2 本論文の構成

以下では本論文の構成について概要と共に示す。

第 1 章は本章であり、本論文の研究背景と目的及び本論文の構成について示している。

第 2 章「マンガ画像の特性及び高機能なマンガコンテンツ生成に関する課題」では、本論文の処理対象であるマンガの基本的な構成及び特徴と、高機能マンガコンテンツ生成支援の実現に関して論じている。マンガは、主に白色の紙の上に線画と黒色のドットからなるトーンを用いて描かれる。高機能なマンガコンテンツを生成するためには、目的に応じた多数の内部要素の情報が必要となる。そこで、マンガの構成要素と画像処理における特徴及び高機能マンガコンテンツ生成のために想定される具体的なアプリケーションについて示している。

第 3 章「マンガ画像からのコマ検出」では、マンガ画像からの自動コマ分割処理について論じている。マンガにおいてコマは、シーン記述の一単位として大変重要な意味を持つ。これまでに、コマ解析手法として、輝度こう配に基づく分割線検出と得られた分割線からなる再帰的 2 分割処理及びその関連手法が提案されている [35, 55]。従来の分割線検出手法では、画像の 2 辺を結ぶ直線を構成する画素上の輝度勾配値を取得する。一方コマの枠線方向は作者により任意に引かれるものであるため、画像に対して水平、垂直に限らないアナログ的な角度となる。任意角度に追従するためには、分割線検出時に基準とする直線の角度の刻み幅を十分に小さくする必要がある。したがって、検出精度を確保するためには、探索する分割線の数が増大となる。また、従来の分割処理では、分割線の決定が輝度こう配のみに依存するため、本来分割不可能な部位にて分割処理が行われるという問題がある。また、2 分割処理の過程で、その処理特性からコマ周囲の余白領域が残る。

本論文では、帯状領域を利用した分割線検出と分割線上の輝度こう配方向分布等を利用した分割線の検査処理を用いたコマ分割処理を行う。本手法では、微小な角度変化に対するロバスト性の向上と分割線の探索数の削減を行い、且つ他のコマ上を通過する分割線を抑制する。また、コマの角位置を取得することで余白領域の除去を行う。解析処理時間については、処理対象画像の解像度を下げることで短縮可能である。一方で、分割精度は低解像度化による情報欠落の影響により低下する。コマ分割実験により、低解像度化による分割精度への影響も含めた、コマ分割の処理時間、処理精度並びにメタデータ付与コストの削減並びに余白領域の除去について示している。

第 4 章「マンガ画像からのキャラクター情報抽出」ではマンガからのキャラクター検出及び識別手法について論じている。キャラクターはマンガの物語を構成する上で必須の情報であり、マ

1.2 本論文の構成

マンガの要素メタデータの中で大変重要である。マンガはその基本構造を線により描かれるため、従来一般的に人間の顔検出で利用されてきた Haar-like 特徴量を用いた識別処理は適さない。そこで、人体のエッジを捉えることで輪郭の識別に有効性を持つ HOG 特徴量を利用し、マンガのキャラクター検出と、その識別処理への利用を試みた。キャラクター検出では、瞳を学習した識別器、顔を学習した識別器をカスケード型に接続することで、単一の要素のみを学習した場合よりも検出適合率が高くなることが確認された。最終的に、瞳、顔の検出結果に対しキャラクターを学習した識別器による識別処理を行うキャラクター検出処理を構成し、その結果を示している。マンガは実際の間人々と異なり、シーンにより非常に大きな形状の変化を持って描かれる。また、作者の特性により、キャラクターの向きに応じてある点を境に輪郭形状が突然変化する場合がある。そこで、キャラクターの状態を正面、横、背面、デフォルメの4系統に分離するアプローチをとり、キャラクターの状況ごとに別クラスを構成した場合の識別性能に対する影響を示している。

第5章「2値線画像向け画像特徴量 HBOP とマンガへの適用」では2値線画像向け画像特徴量 HBOP について論じている。マンガは多数の線によって描かれており、この線の特徴を捉えることで、マンガの解析性能を向上させることができる可能性がある。人間の描いた意味のある線の一つとして文字がある。光学文字認識等では線の特徴を得るために2値細線化を利用している。一方で、第4章で使用した HOG 特徴量は、細線化された2値画像から得られる方向は最大4方向のみとなる。また、直線自体の方向を直接記述することが不可能である。そこで、一般物体認識のアプローチを、2値細線画像解析に応用可能とするために、2値細線分方向ヒストグラムを捉える Histograms of Binary Orientation Pattern (HBOP) 特徴量を提案する。HBOP 特徴量では注目点の8近傍画素において12パターンのオペレータを規定し、これを用いて8方向のヒストグラムを生成する。2値細線で描かれた円と矩形の学習及び識別実験では、HBOP 特徴量は HOG 特徴量と比較して高い識別性能が得られることを示している。また、本研究の処理対象であるマンガ画像に対しても HOG 特徴量との識別性能を比較し、その結果を示している。

第6章は本論文のまとめである。

第 2 章

マンガ画像の特性及び高機能なマンガコンテンツ生成に関する課題

2.1 まえがき

本章では、マンガ解析に係るマンガの特性と高機能マンガコンテンツ生成の具体的課題について述べる。本論文で処理対象とするマンガは人間の想像により描かれ、一般的な自然画像と異なる画像的特徴を持つ。初めに、本論文で取り扱うマンガ画像の基本的な構成と特性についてまとめる。次いで、マンガメタデータ利用に関する具体的検討課題として、高機能なマンガコンテンツの生成に係る課題について述べる。

2.2 マンガの基本的な構成および特徴

本論文で取り扱うマンガ画像の特徴について述べる。マンガは、主に白色の紙の上に線画と黒色のドットからなるトーンを用いて描かれる。トーンとは、多数の黒点によって構成される濃淡変化を表現する画材である。全体の基本的な構造は線画で描かれ、その上にトーンや塗りつぶし等の手法により陰影の表現が重ねられる。したがって、マンガは白黒 2 値により構成される画像である。デジタル化の際にローパスフィルタを用いることで、トーン部分にはグラデーションを検出しようもの、他の箇所では画像の輝度変化は通常画像における、最大のエッジに相当する急峻なものである。このように急峻なエッジにより構成されるマンガは、一般的な自然画像と非常に異なる画像的特性を持つ画像である。

マンガでは、コマと呼ばれる枠の中に、キャラクターや風景などを表現する絵の部分と、吹き出し、セリフ及び効果音などの文字部分が複雑に配置され構成される。セリフ部分では、テキストが、吹き出しとよばれる風船状の枠の中に配置される。マンガを読み進めるむためのページ内の順序はコマによって定義される。ただし、表現のために、コマからのキャラクターや文字要素、吹き出し等のはみ出しを許容するため、そのレイアウトは非常に自由度が高い。マンガ画像の例を図 2.1 に示す。マンガのレイアウトにおいて最も特徴的な点として、コマからの要素のはみ出しがある。要素のはみ出しにより、本来描かれるべきコマの枠線が隠れてしまい、枠線の一部が描かれていない場合がある。更に、集中線などに代表される、コマの枠線と大きさ及び太さが似ている線分が画像内に多数存在する。要素の複雑な重なりはマンガのレイアウト解析を困難にさせる要因となっている。

2.3 マンガデータ生成に関する課題



図 2.1 マンガ画像例 [56]

2.3 マンガデータ生成に関する課題

2.3.1 マンガデータ生成に関する背景

近年のデジタル技術の進歩に伴い、電子書籍化されたマンガの配信が進んでいる [3-6]。電子書籍では、従来の紙に印刷された書籍では不可能であった様々な機能の付与が可能である。電子書籍化されたマンガにおいても、マンガの持つ様々な要素を利用した機能の付与が考えられる。携帯電話向けマンガ配信サービス [11] では、小型端末に対してマンガを提供するために、スキャンされたマンガに対してコマ情報を付与している。Web ベースにおける新たなマンガの表現として、インタラクティブ性の導入についても提案されている [8-10]。デジタルマンガでは、GIF アニメーションのようにコマの内部が動くマンガ表現も可能である。マンガ形態の多様化に伴い、マンガの画像そのものに加え、付加機能のためのメタデータやプレゼンテーション用のアクション作成等、作品制作に係る作業量が増加する。これら付加機能は主にマンガの要素情報に基づくものである。したがって、マンガの要素情報が自動的に取得可能となることで、その作成に関するコストの削減や、従来困難であった新機能の実現が可能となる。高機能マンガコンテンツ生成においてこれを補助するアプリケーションは有用である。

2.3.2 マンガデータ生成の具体例

本論文では、今後のマンガとその関連サービスにおけるデータ生成に対する機能を想定する。以下では、想定する関連サービスの概要と、データ生成における画像処理による支援の効果について述べる。

マンガメタデータを用いたサービス（自動スクロールシステム）

表示領域が限られる場合においてもマンガの閲覧を快適にする機能として、表示領域をスクロールする手法がある [11]。携帯電話は電子書籍発展初期からの主要コンテンツ配信先の一つである [3-6]。一部の大型スマートフォンを除き携帯電話を始めとする小型電子端末では画像表示面積が物理的に小さい。そのため、マンガ画像をそのまま表示した場合に文字やコンテンツの内容も小さいものとなる。そこでページ上をスクロールしつつ表示する手法が取られる。この表示領域のスクロールにはコマの位置と順序からなるコマ情報が必要である。現状のサービスでは、このコマ情報をコンテンツプロバイダが用意している。ここで、コマ情報の入力オペレータの操作によるものである。マンガのコマに関する情報の自動取得を可能とすることで、コマ情報付与にかかるコストが削減できる。

マンガメタデータを用いたサービス（構成要素に基づくサーチ用クエリの取得）

コンテンツ配信サービスを構築する上で検索機能は重要である。マンガにおけるサーチ用クエリとしてキャラクター情報、コマ情報、セリフ情報、擬音情報等が想定される。また、サーチ要のクエリは内部要素を齟齬にリンクするためのキーワード情報としても利用可能である。扱うコンテンツが小説等の文書であれば、その内部要素は文字のみであるため、内部要素をクエリとした検索機能は容易に実現可能である。一方でマンガの場合には、文字だけでなく複雑かつ大量の内部要素が存在する。情報取得対象としての要素量は、その処理を人力で行う場合におけるコストそのものである。多数の要素に関する情報を利用したい場合には、その取得コストは高くなる。また、現状の文書画像処理技術では、マンガからの文字の検出自体が困難であるばかりか、内部要素の情報を得ることができない。マンガの構成要素に関する自動的なメタデータ取得を可能とすることにより、マンガ構成要素に基づくクエリを利用したサーチが可能となる。

マンガ制作支援システム（キャラクター作画支援）

マンガはストーリーを構成する多数の絵画表現やセリフを用いて構成される。マンガはその制作に多くの時間と高度な技術が必要であり、その制作に係るハードルが高い。また、ストーリー性のあるマンガを作成するためには、多数のページが必要である。マンガの作成コストは高く、また、絵を描くための技術が必須である。よって、マンガの制作自体についても課題があるといえる。一方、マンガ同様に、制作のために専門的な技術が必要とされる音楽制作では、ボーカロイドと呼ばれる、主にボーカルに関する要素の作成を補完するソフトウェアである初音ミク [57] の登場により、広範なレベルのユーザによる創作活動が大変活発化された。これは、ソフトウェアによる部分的なサポートが、ユーザの創作活動に対して効果的に作用した結果であると考えることができる。マンガにおいても技術、作業時間に対して有効なソフトウェア的補助により、

2.3 マンガデータ生成に関する課題

幅広いユーザ層による分野全体の創作活動活性化につながると考えられる。

デジタル上での作画機能に重点をおいたマンガ製作用ソフトウェアがある。このようなソフトウェアの代表として、ComicStudio [58] やその後継版である CLIP STUDIO PAINT [59] がある。ComicStudio は総合マンガ制作ソフトウェアであり、ペンタブレットから得られる入力線をデジタルデータ上で滑らかに変換することができる。基本的に紙ベースの作画をコンピュータ上でも行えるようにすることを目的としたものである。制作支援機能として、背景の作成や、アタリを取りやすくする機能がある。近年、これらのソフトウェアはデジタルマンガ制作において一般的に利用されている。一方で、清書の作画についてはほぼすべてマニュアルで行う必要がある。したがって、技能の低いユーザーにとっては要求される技術レベルが高い。

3次元ベースのキャラクターを用いたマンガ作成支援ソフトウェアがある。こみ Po! [60] は、3次元データとして作成されたキャラクターを動かし、漫画上に配置するためのソフトウェアである。また、前述の CLIP STUDIO PAINT においても、3次元モデルの利用が可能である。更に、3次元キャラクターのポーズ入力を補助するための装置 [61] もある。しかし、これらの機能を利用するユーザは一部にとどまっている。現状の3次元形状からレンダリングされるキャラクターでは、マンガ特有の表現等を十分に行えるものではないと考えられる。

マンガ特有の表現を再現可能とするためには、マンガ独自の表現を含む形で作画されたキャラクターに基づく2次元画像ベースのアプローチが有効であると考えられる。また、作品作成の上では、一度生成された画像に対する微調整等の再編集が行えることが望ましい。線で描かれるマンガの表現は重要なものであり、2次元画像ベースのキャラクター生成技術が必要となる。そこで、あらかじめキャラクターごとの特徴を学習しておき、入力されたポーズ情報や表情タグに応じて出力画像を得る形式のマンガ作成支援技術を想定する。キャラクターの生成にはキャラクターごとの構成要素を適宜変形、配置し出力を得る方法が考えられる。キャラクターはその画像的構成要素及び位置関係による構造によって特徴づけられている。構成要素及び構造情報は、入力画像からキャラクター要素（瞳、口等）その位置関係を検出することで得られる。本機能の実現には、2次元で描かれたキャラクターの各要素の検出及び位置等のパラメータ化が必要である。2次元画像による表現であれば、画像生成後にマンガ特有の表現を盛り込むための再編集も行きやすい。キャラクター要素のパラメータ化のためには、キャラクターごとに、複数のバリエーションで取得されたキャラクターの顔を構成する各パーツの位置及びサイズ等の詳細なデータが必要である。2次元画像ベースのキャラクター生成であれば、マンガ特有の表現を十分に行いつつ、技術レベルにかかわらずマンガの作成コストを低減できる可能性がある。

マンガ制作支援システム（ページ構成設計支援）

マンガの表現に関する検討 [12] では、コマの位置やキャラクターの位置による表現について言及されている。そこで、マンガの構成を検討する際に利用可能な、読者の視線移動と注目時間を予測し表示する機能も検討される。読み手側行動の図示は、視覚及び心理的効果を想定したマンガの構成に結びつくものである。読み手の視線フロー表示機能を実現することで、作品のクオリティ向上および表現に寄与できるものと考えられる。

2.4 むすび

本論文の目的は、従来自動取得が困難であったマンガのメタ情報を取得可能することである。マンガの構成要素に関する情報を利用することで、マンガ配信サービスあるいは新たなマンガコンテンツ制作に対する支援が可能となる。本論文では、上記アプリケーション実現のため、マンガの構成要素に関する情報取得を可能とする画像処理手法について検討する。セリフに関する検出処理は関連研究の一部で検討されている [62–64]。本論文では、コマに関する情報、キャラクターに関する情報の解析手法について検討を行う。

マンガ画像は一般的な画像処理対象である自然画像とは異なり、グラデーションや、色に関する情報が得られない。マンガは急峻なエッジにより構成されており、その画像的特性は自然画像とは大きく異なるものである。そのため、一般的に利用される文書画像解析技術や、顔検出に用いられる特徴量では解析が困難である。急峻なエッジのみからなる画像を解析することは、画像解析処理の課題の一つである。

マンガは人間の想像を基に描かれる芸術表現の一つである。マンガの構成要素を識別する処理は、人間の想像で描かれたマンガ上の意味的要素の識別を行うものである。人間の描いた意味的要素の識別処理として文字認識がある。文字は形状や交差の規則がある程度定められたものであり、ある程度の規則化が可能である。一方で、マンガは表現の自由度が高い。例えば人の顔を表現する際に使用される線の数は不定である。そのため、文字解析のような交点のルール化等は困難であり、解析の難易度は高い。人間が描いた文字以外の意味的要素を識別することは、画像解析分野における大きな挑戦である。

以降では本章で述べた課題に対する具体的なアプローチを述べる。本章で例示した個別のアプリケーションに関する考察については 6 章にて示す。

第 3 章

マンガ画像からのコマ検出

3.1 まえがき

マンガにおいてコマは、シーン記述の一単位として大変重要な意味を持つ。ページ内における読み順の定義、シーンの区切り、視線誘導等はこのコマによって与えられるものである。また、画像処理を行う上でも、他のオブジェクトに関する処理対象位置についての情報源となりうる。また、携帯電話向けマンガ配信サービスではコマの情報を基にスクロールがなされる。コマは、マンガの画像上における物理的な位置関係と意味上における物語のシーン推移を示す重要な要素である。

これまでに、文書画像の解析を目的としたレイアウト解析処理に関する検討が行われている。ここで文書画像とは主に新聞や論文等の、明確に分離された文字と画像からなる画像である。文書画像では、基本的に行や文字の位置関係及びその角度に一定の規則が存在する。この規則を利用して、文字を構成する画素の並行投影結果やモルフォロジ処理結果から要素ごとの領域情報を得る手法が提案されている [30–32]。また、対象画像から任意の形状を検出する、一般化 Hough 変換なども提案されている [33, 34]。一般化 Hough 変換では任意の検出対象図形テンプレートを用意し、パラメータ空間上での評価により検出を行う。これらの手法は、領域が十分に分離されている画像や、検出対象の形状が推測可能な場合に有効である。

マンガ画像では、各オブジェクトが明確に分離されておらず、また、コマの形状は矩形に限られない。したがって、先に述べたレイアウト解析手法では正しく認識が行えない。また、一般化 Hough 変換に対しては、マンガ画像のコマ形状は高い自由度を持つため、テンプレートを用意できないという問題点がある。このように、マンガ画像はレイアウトの自由度が高いため、既存の文書画像解析手法は適用不可能である。マンガのレイアウト解析を行うためには、マンガの画像的特性に合わせた解析手法が必要である。

マンガ画像の処理に特化した手法として、田中らによるコマ分割処理 [35] の研究がある。コマの枠線に沿った点では、コマの枠線に起因するエッジが存在するという特徴がある。そのため、コマの枠線に沿って引かれた直線上では、多くの点で高い濃度こう配値が得られる。田中らの手法では、この特徴を利用し、再帰的な画像の 2 分割処理によるコマ分割手法を提案している。画像の 2 分割処理には、画像の 2 辺を結ぶ直線の中で、その直線上の濃度こう配合計値が最も大きい直線を用いる。このような処理を再帰的に実行することで、各コマへの分割と読み順の取得を行う。しかし、分割線を決定する際に、濃度こう配値のみを用いるため、ロバスト性及び分割精

3.2 文書画像のレイアウト解析

度が低いという問題がある。更に、再帰処理の各過程において、画像上の2辺を結ぶ多数の直線に対して濃度こう配計値を計算するため、処理に時間がかかる。コマ分割処理の性能向上のためには、分割線の決定における信頼性と速度の改善が必要である。

本章では、マンガ画像からの高精度かつ高速なコマ分割処理を提案する。提案する2分割の改良アルゴリズムでは、2分割手法の分割線検出に対して、分割線候補の検出にロバスト性をもたせ、分割線の濃度こう配方向の分布による分割線適合検査を導入した。更に、アルゴリズムの効率化による高速化、余白部の除去アルゴリズム等の追加を行うことで、高精度かつ高速なコマ分割処理を実現する。提案手法と従来手法を用いて、実際にマンガのコマ分割実験を行いその精度と効果について示す。

以下、3.2節では文書画像のレイアウト解析と、マンガ画像に対する問題点について述べる。続く3.3節ではマンガ画像の構成要素であるコマの形状、マンガにおける視線遷移情報、3.2節では関連手法である文書画像の領域分割手法及び図形解析手法のマンガ画像に対する適用性について述べる。3.4節では、再帰的2分割処理によるコマ分割処理の従来手法について述べる。3.5節では、再帰的2分割処理の改良手法について述べる。そして3.6節では、コマ分割実験により、低解像度化による分割精度への影響も含めた、コマ分割の処理時間、処理精度並びにメタデータ付与コストの削減と、余白領域の除去について示す。最後に3.7節にて本章のまとめを示す。

3.2 文書画像のレイアウト解析

3.2.1 文書画像解析

文書画像はフォント等により構成される文字領域とグラフや写真等の図画領域から構成されている。レイアウト解析では文書中の領域を文書ブロックや図ブロック等に分解することを目的としている。文書画像では、マンガと異なり、ドキュメント内に含まれる文字要素と画像要素は明確に分離されている。また、文字は整列されており、行間や段落間には均等幅の空白領域が存在する。辻らにより提案された投影法 [30] では、画像の垂直方向と水平方向に対して、それぞれ黒画素の出現頻度を取得することで、行間の位置や文書ブロックの境界を得る。Wongらにより提案された Run Length Smearing Algorithm (RLSA) [31] では、設定したしきい値以下の長さの白画素により分離される黒画素を連結することで、複数の文字からなる文書ブロックを取得する。黄瀬らはボロノイダイアグラムと近隣グラフにより文字ブロックを取得している [32]。従来の文書画像解析はオブジェクト間が分離されているという前提の下設計されたものである。

マンガにおけるコマの配置は文書画像と異なる特徴をもつ。このため、文書画像解析処理をそのまま適用しても、コマの検出が困難である。例えば、マンガ画像に対し投影を行う場合、コマの形状が矩形に限られないため、上下、左右等固定方向のみの投影処理ではコマ間の空白部分をとらえられない。そして、コマの内部でも空白部分が多く存在することから、投影結果として、有意なヒストグラムが得られない可能性が高い。また、コマからの要素のはみ出し等によるコマ同士の連結、コマ枠線の欠落等の存在により、連結成分を用いた解析手法もコマ分割には適さない。更に、コマ内部にもコマの枠とは関係のない直線成分が多数存在する。このため、一般的な直線分の検出処理では、マンガ特有の特徴から影響を受け、領域の検出が行えない。従来の文書

画像解析処理では、マンガのコマ位置の検出が不可能である。

3.2.2 一般化 Hough 変換

任意形状を検出する手法として一般化 Hough 変換がある [33, 34]。一般化 Hough 変換では、Hough 変換 [65] の拡張により、任意のテンプレート形状の検出を可能としている。ただし、テンプレート形状はあらかじめ既知である必要がある。一般化 Hough 変換では、Hough 変換と同様に、パラメータ空間上に投票し評価を行うことで任意形状を検出する。一般化 Hough 変換は、検出対象の形状が既知である場合高い精度で検出処理が可能である。

マンガ画像のコマの形状は矩形等に限定されず高い自由度を持つものである。コマは矩形であっても、その縦横比は可変である。また、矩形に限られない任意の 4 角形である場合も存在する。一方で、一般化 Hough 変換では、あらかじめ検出対象の形状が既知でない場合には、テンプレートを用意できない。マンガからのコマ検出を行うためには、形状の変化に柔軟な手法を考える必要がある。

3.3 コマと視線遷移情報

マンガにおけるコマの重要な役割は、ページ内部における視線遷移情報を与えることである。マンガ画像のページ内部における視線遷移情報はコマの配置によって規定されている。コマの形状と読み順の関係について、隣接するコマ同士の相対的な枠線角度から順序づけを行う手法が提案されている [66]。山田らの手法では、多角形のコマを考慮したコマ間の順序づけを行うことが可能であるが、今回処理対象としたコマの場合には、以下に示す、より単純なルールにより順序を示すことができる。日本で一般的な左開きのマンガでは、ページ内のコマは、右上のコマから順に、右から左かつ上から下の順で読み進める。読み進める過程において、同一段に複数コマが存在する場合にはその段の右から左の順に読みすすめる。同一段に存在するすべてのコマが終わるとひとつ下の段へ移動する。以上を左下のコマに到達するまで繰り返す。コマの中には、作者の意図により、複数のコマが連結される場合や、読み順が明確でない場合がある。したがって一部の場合を除き、基本的にはコマの位置関係を基準として、ページ内部におけるマンガのストーリーを構成する順番の取得が可能である。

コマの形状には様々なバリエーションが存在する。マンガのコマ配置の例を図 3.1 に示す。図中の数字はコマの順序である。ここで示した一例は以下の特徴を持つコマである。

- (A) コマはある程度の大きさをもつ
- (B) コマは矩形に限定されない線分若しくは画像の縁からなる 4 辺に囲まれた領域である
- (C) コマとコマの間には空白部が存在する

一方で、マンガの種類により上記特徴の枠に収まらない、より自由なレイアウトを持つマンガも存在する。具体的には、四角形以外の形状のコマを持つ場合や、コマ同士が重なって配置される場合、コマ間の区切りが線のみで空白を持たない場合などがある。これらのレイアウトは主として少女マンガ等で多く見られる。しかし、そのような自由なレイアウトを持つマンガであっても、

3.3 コマと視線遷移情報

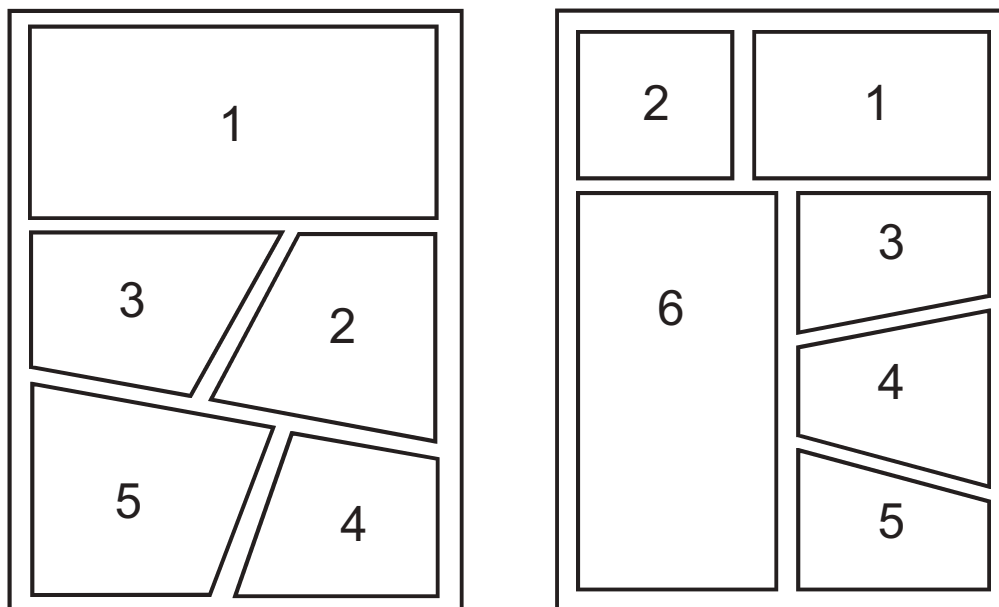


図 3.1 コマの例
自発表 [67] から引用

上記の (A) から (C) の特徴を持つコマもまた多数含まれる。したがって、これはストーリーをもつマンガにおいて、最も普遍的なコマであるとみなすことができる。作者が意図的に読者に対して自由度をもたせる表現を行う場合もある。よって、複雑なレイアウトを持つコマの中には、その順序及び境界線を定義出来ないものも存在する。このようなコマでは読み手側に読むための自由度があり、その順序情報付与にはオペレータが個別に対応する必要がある。読み順及びコマ領域の分割の補助においては、普遍的なコマの分割を可能とすることで、目的の多くを達成可能であると考えられる。本章で取り扱うコマの条件を、次のように規定する。

- コマの形状は矩形に限られない四角形によって構成される
- コマの境界を示す枠線が存在する
- コマ間には少なからず余白が存在する
- コマからのキャラクター、吹き出し、オノマトペ等の要素のはみ出しを許容する
- コマ枠線の欠落を許容する

このようなコマに対して、コマ同士の位置の関係性からコマ順序を得る手法が提案されている。山田ら [66] は、多角形のコマを考慮したコマ間の順序づけについて提案している。一般的なコマ割りにおける順序はより単純化されたルールに基づいて求めることができる。今回処理対象としたコマの場合には、以下の 4 点から順序を求めることができる。

- 各ページでは、右上のコマから順に、右から左かつ上から下の順で読み進める
- 読み進める過程において、同一段に複数コマが存在する場合にはその段の右から左の順に読みすすめる
- 同一段に存在するすべてのコマが終わるとひとつ下の段へ移動する

- 左下のコマに到達するとそのページは終了となる

3.4 再帰的 2 分割処理によるコマ分割手法

マンガの特徴に対応したコマ検出手法として、田中らにより再帰的な 2 分割によるコマ分割手法が提案されている [35]。2 分割の繰り返しによるコマ分割処理のステップを図 3.2 に示す。2 分割処理に用いる分割線を検出するために、画像の 2 辺を結ぶ直線を用いる。本論文ではこの線を検出線とよぶ。分割処理の手順を以下に示す。

1. 画像内の各画素について、Sobel フィルタにより輝度勾配値を求める
2. 1 で得られた輝度こう配値に対し、ガウス関数を用いて画像中心部に重み付けを行う
3. 検出線の始点及び角度を変化させつつ各検出線上の重み付き輝度こう配値を計算する
4. 3 にて得られた最大の重み付き輝度こう配値をもつ線により画像の 2 分割を行う
5. 分割後の画像に対し 1 から 4 の処理をそれぞれ繰り返し適用する

分割線の決定に用いる輝度こう配値は、コマの枠線のみならず、コマ内部の絵や文字成分からも得られる。そのため、コマの枠線以外を捉えた分割線が決定される場合があり、結果として分割精度が低下するという問題がある。

その後の研究 [55] において処理の高精度化とその順序付けに関する検討がなされた。分割線決定に際して、コマ間の白画素の連続領域長比率 WPR と T 型隣接に関する重み付けの概念が導入された。WPR は検出線に沿った検出線近隣における連続する白画素の最大長と、検出線の長さとの比率であり、以下の式で示される。

$$WPR = \text{最大白画素連続長 [pixel]} / \text{検出線の長さ [pixel]} \quad (3.1)$$

T 型隣接は、コマ間の白画素領域の形状が T 字型となるコマの並びである。また、コマ内部の影響を低減するために、Douglas-Peucker アルゴリズムにより四角形の領域が得られた場合にはその内部を塗りつぶす処理を導入している。縦方向と横方向に分割線候補をとり、その交点周辺の小領域について、分割線検出処理を再度実施し、交点における優先方向が決定される。この優先方向となる向きの分割線を正式に分割線とすることで、誤った分割線によるコマの切断を防ぐ。また、2 分割処理本処理の対象としているコマに対する、順序付けルールについて、2 分木構造を構成することで、順序付け可能であることを示している。ここで提案された 2 分木ルールは前節にて述べたコマの順序ルールを 2 分割処理の処理過程と結びつけたものであり、処理時にそのまま 2 分木構造を構成可能である。

この再帰的な 2 分割によるコマ分割処理では、分割線の候補を決定するために検出線上の輝度こう配値を利用している。このため、コマの枠線が少しでも検出線とずれる場合には、本来コマの枠線から得られるべき輝度こう配値が取得できない。そのため、検出線のパラメータ変更ステップを十分に小さくとる必要がある。一方で、パラメータ変更ステップが小さい場合には、分割処理過程において調査する検出線数が多くなる。コマ分割処理に係る計算量は検出線の量に大きく依存する。したがって、十分な精度を保つためには高い処理コストが必要である。また、四角形内部の塗りつぶしが行えないコマも多く、コマ内部の影響が残る場合がある。

3.5 コマ分割処理の改良

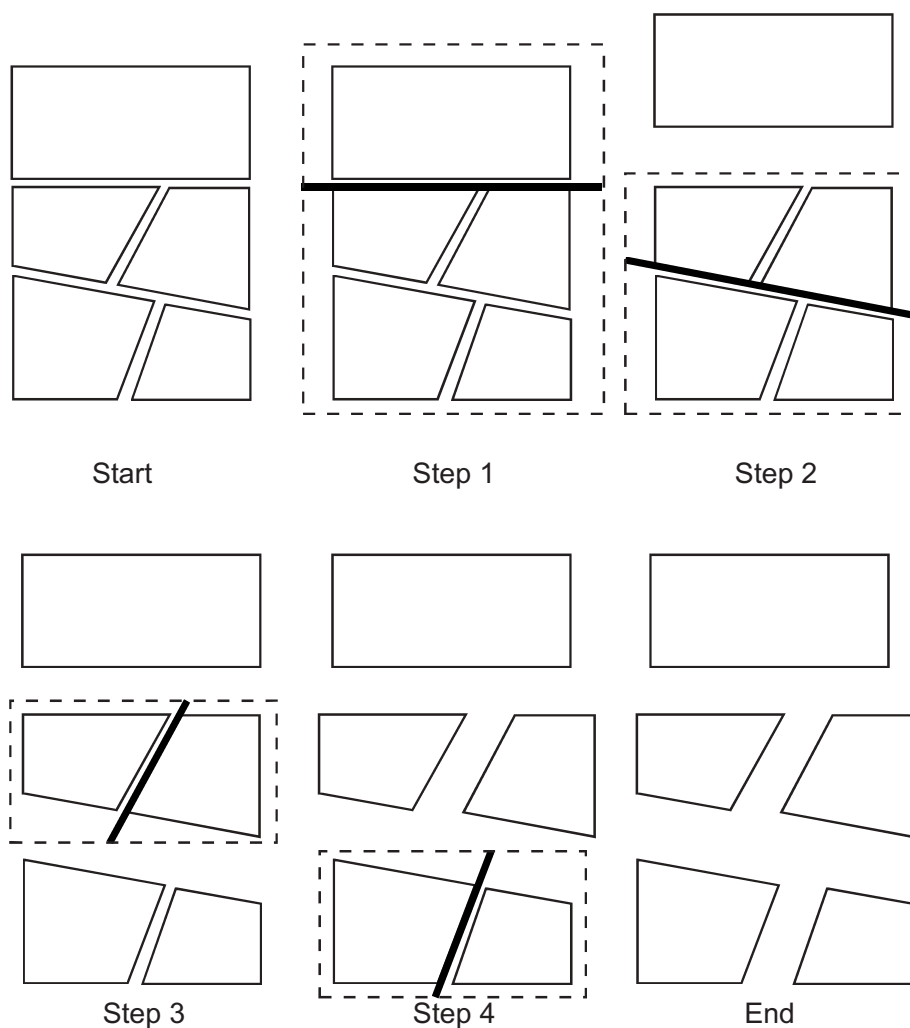


図 3.2 コマ分割のステップ
自発表 [67] から引用

3.5 コマ分割処理の改良

3.5.1 コマ分割処理の改良について

本節では再帰的 2 分割処理に基づくコマ分割処理の改良手法について述べる。はじめに、コマ分割処理の高精度化のために、帯状分割線検出、分割線上の輝度勾配分布に着目した分割線選択、コマ内外識別処理を導入する。次に、高速化として、分割線検出アルゴリズムの改良を行う。ここで、2 分割の繰り返しによるコマ分割処理では、画像内の走査を多数行うため、画像解像度に応じて処理時間が変化する。処理量を削減するためには低解像度化が有効であるが、同時に分割処理の精度に影響がある。そこで、低解像度化による分割精度への影響について明らかにする。2 分割処理では、アルゴリズムの特性により、コマの周囲に余白部分が残る。この余白部分を除去するため、直線検出及び角検出の組み合わせによる、余白除去アルゴリズムについて述べる。提案手法はコマ分割処理の高速化及び高精度化と取得されるコマの位置精度を改善するものである。

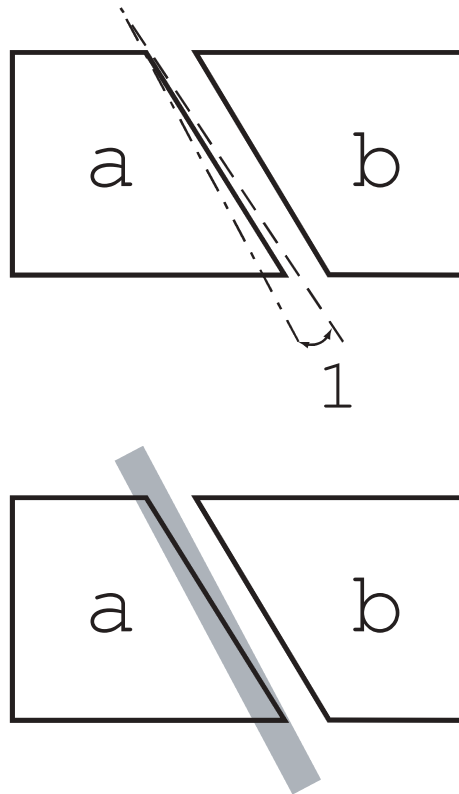


図 3.3 直線及び帯状領域による分割線検出処理
自発表 [68] から引用

3.5.2 帯状分割線検出

提案手法では、1画素以上の幅をもつ帯状領域を利用して分割線の決定を行う。従来手法では、分割線の検出に幅1画素の検出線を用いていた。始点と角度を変化させつつ多数の画像内の2辺を結ぶ直線について、その直線上の輝度勾配合計値をもとに分割線候補の検出を行う。従来の検出方法では、検出角度とコマの枠線角度に違いがあると分割線が検出できない場合がある。これを防ぐため、多数のパラメータを用いて検出線を生成し調査を行う必要がある。これまでの検出線に幅を持たせることで、ロバスト性を向上させた、帯状分割線検出処理を行う。帯状分割線検出により、角度の微小な違いによる分割線の検出漏れを防ぎつつ、操作回数が膨大となることを避けることができる。

直線による分割線検出の概念及び帯状分割線検出の概念を図 3.3 に示す。以降では、従来の検出線を基本検出線と呼ぶこととする。帯状分割線検出では、基本検出線に対して複数画素からなる幅を持たせる。このなかで、輝度こう配値の取得を行う。基本検出線の角度は、角度の直接指定若しくは始点と終点の組み合わせにより決定される。

帯状分割線検出処理では、基本検出線に対して幅を持たせるために、新たに追加する画素の画素値を求めることが必要である。ここで、画像内の直線成分を発見するという目的から、基本検出線の角度は一般的に多数となる。一方で、検出対象となる画素は正方形のグリッドである。通常、直線に対して垂直方向に画素を追加する場合には、直線に対する垂直方向に対して画素数分

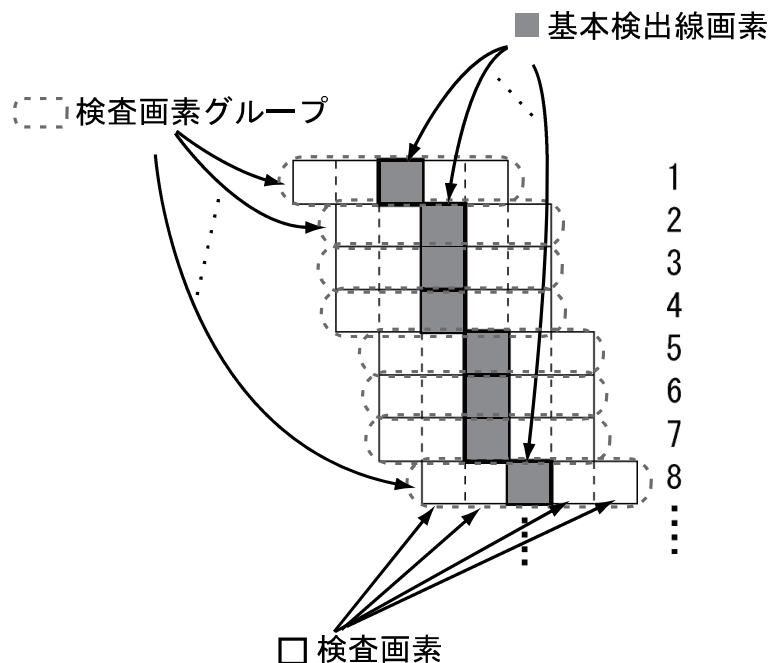


図 3.4 基本検出線に対する追加画素
自発表 [68] から一部を変更した

移動した位置の座標を計算し、その周囲 4 画素からの、距離に応じた重み付き平均値を求める必要がある。多数の直線上の輝度こう配平均値を基に分割線を検出する処理において、重み付き平均値を求める場合、多大な計算コストが必要となる。

提案する帯状分割線検出処理では、計算コスト削減のため、追加する画素については垂直及び水平に拡張するものとする。具体的な追加画素の方向は、以下に基づいて選択する。

$$\text{追加する参照画素の方向} = \begin{cases} \text{Vertical} & 0^\circ \leq |\theta| < 45^\circ \\ \text{Horizontal} & 45^\circ \leq |\theta| \end{cases} \quad (3.2)$$

ここで、 θ は元となる基本検出線の角度である。追加画素の例を図 3.4 に示す。ここでは基本検出線の角度 $|\theta|$ は 45° 以上であるため、水平方向に画素を追加することで帯を構成する。

基本検出線の 1 画素に対して、対応する追加画素（検査画素）を加えることで検査画素グループを構成する。検査画素グループ数は、基本検出線の画素数と同一である。検出線上の重み付き輝度こう配平均値は、検出線を構成する各画素における輝度こう配の合計値を基に算出する。各検査画素グループにおける、最大となる輝度こう配絶対値を持つ画素を代表画素とする。帯を構成する各代表画素を用いて、分割線の評価値を計算する。

一般的にこのような直線検出を行う場合には、Hough 変換の利用が考えられる。しかし、マンガ画像においては、画像内に多数の直線成分が存在するなど、本来検出したい線以外の要素による影響が強くなる。このため、検出する候補数に制限を設けた場合などに、一部の位置の線のみを分割線候補として検出し、正解となるべき分割線が分割線候補の中に含まれないなどの状況が発生する。一方で、候補数の制限をゆるくし過ぎると無数の線が検出されるなど、パラメータの制御が困難である。一方で、始点と終点による基本検出線のパラメータ指定であれば、画面内で均等に分割線候補の検出を行える。提案手法では、コマ内部の直線成分の影響とパラメータ制

御の観点から、分割線候補の検出には Hough 変換は使用しないこととした。

3.5.3 分割線上輝度こう配分布に着目した分割線選択

従来のコマ分割処理では、コマの上を通過するような、本来切断してはならない箇所が分割線として検出される場合がある。これは、コマの内部にある直線の影響が強い場合あるいは一部はコマの枠線に沿うがその直線の延長線上で他のコマに侵入している場合等に発生する。コマ内部の直線に沿う場合には、コマの枠線と同様に検出線上の輝度こう配平均値は高くなる。よって、輝度こう配値のみに基づく従来の分割線決定方式では、このような場合に分割線の誤検出が生じる。2分割処理という特性上、他のコマに分割線が重なる場合には、その後のコマ分割がうまく行かないばかりか、得られる読み順まで正しい結果から外れることになる。検出線が他のコマに侵入する場合には、分割線の候補から除外されるべきである。

本手法では輝度こう配方向の分布を分割線決定時の規範として利用する。コマの内部には、登場人物は背景などの絵や、セリフの文字、オノマトペ等様々な要素が配置されている。コマの枠線と同様の直線成分も多数存在するが、一般的にコマの内部においては、様々な方向の輝度こう配が存在していると考えられる。検出線がコマの内部に侵入する場合、コマの内部に侵入している部分においては、その方向分布の分散は大きくなる。図 3.5 に検出線の位置、図 3.6 に、図 3.5 においてコマ内部に侵入している検出線 (a) 上における検出線に対する方向成分の分布を、図 3.7 に、図 3.5 においてコマ内部に侵入しておらず、コマの枠線に沿った検出線 (b) 上における検出線に対する方向成分の分布をそれぞれ示す。図 3.5 に示された検出線は、実際にマンガ画像に対して分割線の検出処理を行った際に、上位の候補として得られた直線のパラメータである。図 3.5(a) の直線では、20pixel から 220pixel 近辺の間で検出線がコマの内部に侵入しており、図 3.6 において、当該領域にて方向成分の分散が大きいことがわかる。一方で、図 3.5(b) の直線では、検出線がコマの内部に侵入しておらず、図 3.7 においても、全体的に方向成分の分散が小さいことがわかる。

本手法では、分割線の輝度こう配に基づく識別処理を導入する。輝度こう配に基づく識別処理は以下の手順からなる。

1. 対象検査帯を n 個のセグメントに等分し、それぞれのセグメント内で帯に対する各代表点の輝度こう配の角度を調べる。
2. セグメントごとに、1 で得た代表点について、その角度が $(90 \pm \delta)$ 度の範囲に収まる代表点の割合を調べる。
3. 2 で得られた割合が $1 - \rho$ 以下の場合にはそのセグメントを不適合領域とする。
4. 3 で得られた適合するセグメントの数が $n - m$ 以上である場合には、該当する検出帯に対する輝度こう配方向検査を合格とする。

ここで、 n, m, σ は任意とし、 $n > m + 2$ とする。また、コマの周囲に大きな余白が存在する場合がある。余白部分では輝度こう配値は 0 となり、代表点の輝度こう配方向が検出できない。これに対応するため、検査帯の最初と最後の領域では、代表点の除外処理を行う。コマの枠線と接した箇所では、輝度こう配値が大きくなる。そこで、検出線の始点および終点から、それぞれ順に

3.5 コマ分割処理の改良

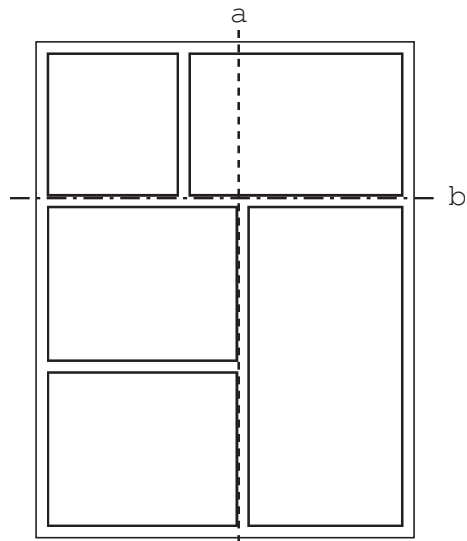


図 3.5 検出線の例
自発表 [68] から引用

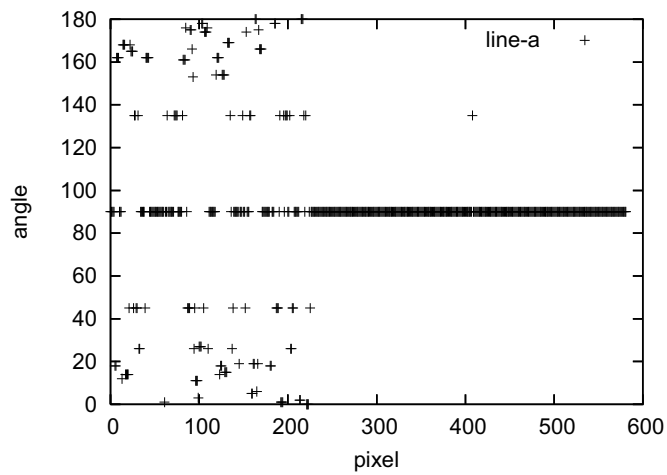


図 3.6 検出線上の輝度こう配方向分布 (コマへの侵入有り)
自発表 [68] から引用

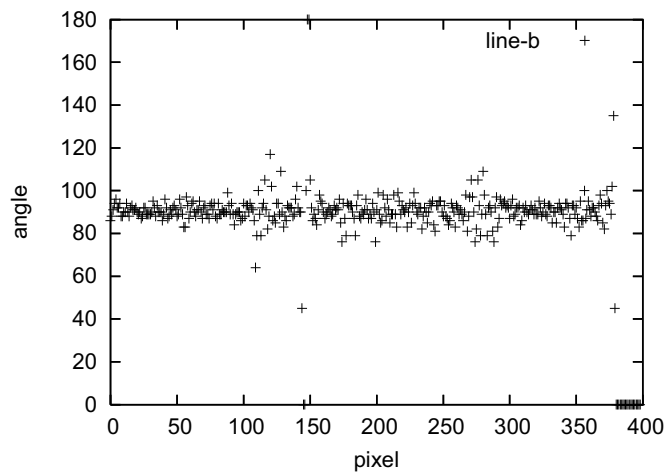


図 3.7 検出線上の輝度こう配方向分布 (コマへの侵入無し)
自発表 [68] から引用

大きな輝度こう配値を確認できる箇所までの画素数を確認し、これを除外する。余白が大きい場合には、その画素数が1セグメントを超える場合がある。この場合には、当該セグメントを処理から除外し、 $n = n - 1, m = m - 1$ とする。これらの処理により、余白領域の影響を避けつつ、輝度こう配方向の分散が大きい箇所を含む直線を分割線の候補から除外する。

3.5.4 コマ内外識別

コマ内部の輝度こう配方向を利用した検査に加え、コマの枠線との交差を考慮した判定処理を導入する。コマの内外識別では、検出線の両側に存在するコマの枠線に関する状況を考慮する。コマの枠線に沿った検出線の場合には、検出線の片側は常に空白領域が続く。一方で、検出線がコマの枠線と交差する場合には、検出線の両側に枠線画素が存在することになる。このような検出線を分割線とした場合には、誤分割が生じるため、これを分割線候補から除外する。

コマ内外識別の具体的な処理は以下のとおりである。分割線候補を選定する際に、検出線両側について最初に出現する黒画素の位置を調べる。検出線の両側において、この黒画素が等位置に存在する場合には、基本検出線がコマの枠線と交差しているとみなし、分割線の候補から除外する。

3.5.5 分割線検出アルゴリズムの改良による高速化

高速化の概要

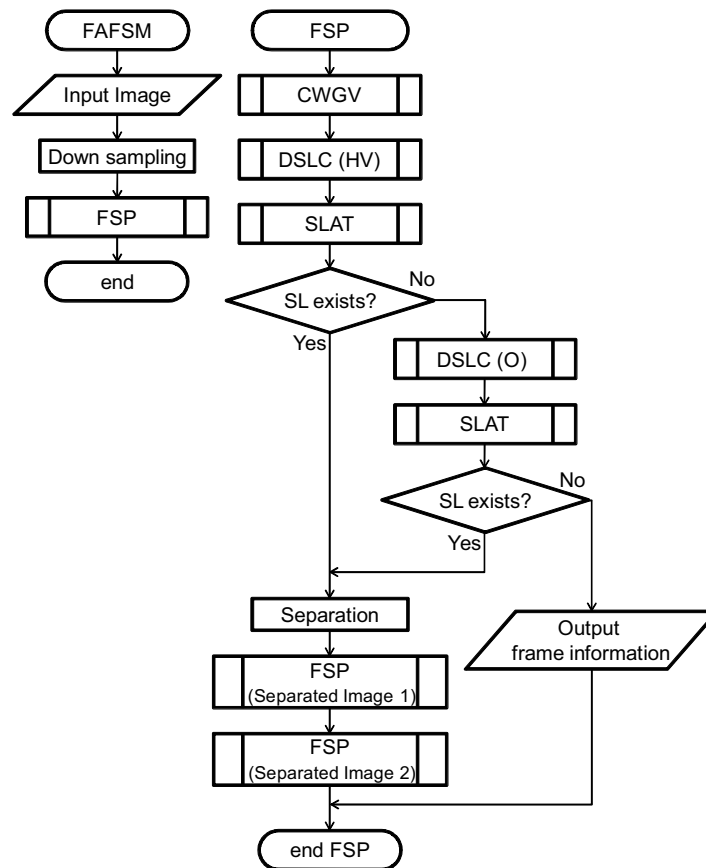
コマ分割処理における処理時間の短縮手法について検討を行う。先に述べた、帯状分割線検出、分割線上輝度こう配分布に着目した分割線選択及びコマ内外識別を導入したコマ分割処理を高精度コマ分割処理:Accurate Frame Separation Method とし、以下 AFSM と記す。AFSM では、多数の基本検出線が必要であること及び、各検出線に伴う帯状領域内の輝度こう配値の計算等の追加処理により計算コストが高い。

これまでの調査より、マンガにおけるコマの枠線は、水平若しくは垂直方向であるものが多いという傾向がある。また、画像の端領域ではコマの境界は多くの場合において存在しない。これらの特徴を利用することで、検出処理の計算コストを低下させることが可能である。更に AFSM では、帯状の検出領域を利用することで、近似パラメータにおける基本検出線では、相互に重複領域を持つことになる。この重複領域があることにより、パラメータの変更きざみを大きくすることができる。ここでは、上記特徴を考慮した AFSM の処理効率改善手法である、高速高精度コマ分割処理:Fast Accurate Frame Separation Method (FAFSM) を提案する。

高速高精度コマ分割処理手順 (FAFSM)

■FAFSM の処理手順 高速高精度コマ分割処理 (FAFSM) の処理フローを図 3.8 に示す。FSP は画像を引数とする分割処理のためのプロシージャ、SL は分割線を示す。FAFSM では入力された画像に対してまず画像解像度変換により低解像度化を行う。そして低解像度化された画像を FSP に渡す。FSP では、分割線候補検出順序の最適化及び冗長な分割線検索の削減に基づいた手順により、分割線の検出及び画像の分割処理を実行する。ここで DSLC(HV) は垂直水平方向のみに限定した分割線検出処理である。DSLC(O) は DSLC(HV) 以外の方向に対する分割線検出処理

3.5 コマ分割処理の改良



DSL (HV) : DSL for horizontal and vertical direction.
 DSL (O) : DSL for oblique direction.

図 3.8 高速高精度コマ分割手法 (FAFSM) の処理フロー
 自発表 [69] から引用

であり、FAFSM では DSL(HV) において分割線が検出されなかった場合のみ実行される。本アルゴリズムにより多くの場合で分割処理時間が低減される。

■冗長な分割線検索の削減 マンガのコマはある程度の大きさを持つため、画像の四隅近傍に複数のコマが存在できない。つまり、画像の四隅近傍には分割線が存在しない。そこで画像の四隅近傍を始点とする分割線候補の検出処理の削減を行う。ここでは、分割線候補の検出処理にでは画像の 4 隅から l [pixel] の範囲を始点座標から削減する。図 3.9 及び 3.11 に始点、終点の設置領域を太線で表示する。図 3.9 は通常の状態、図 3.11 は冗長な分割線検索の削減を行う場合の始点、終点の設置領域を示す。

更に、検出線のパラメータ変更ステップを大きくする。本手法では、検出線に幅をもたせたため、実際のコマの枠線と検出線の角度差に対するロバスト性が高くなる。そこで、検査帯の変化角度を適切に変更することで、処理精度を維持したまま処理の高速化を図る。ここでは、検出線角度の変化量を画像サイズに応じて変更することとし、以下の式によりその値 ξ を決定する。

$$\xi = A \left(\frac{1}{\omega + 1} \right) + B \quad (3.3)$$

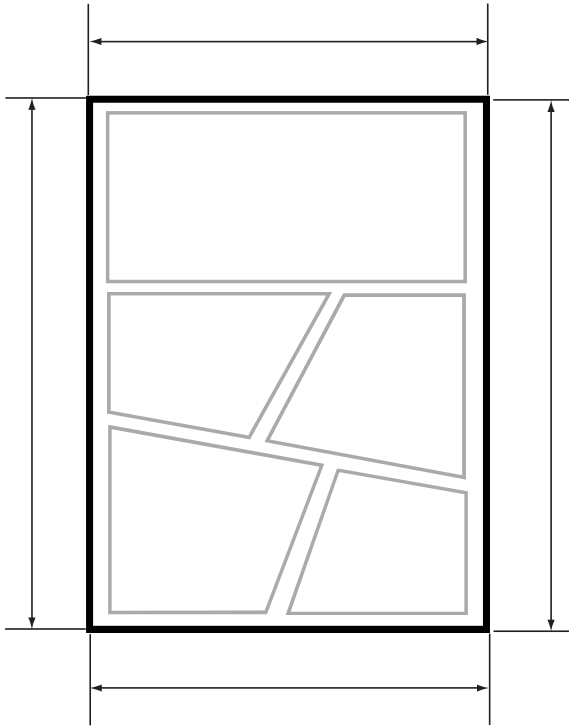


図 3.9 コマ分割処理の始点及び終点位置 (標準)
自発表 [67] から一部を変更した

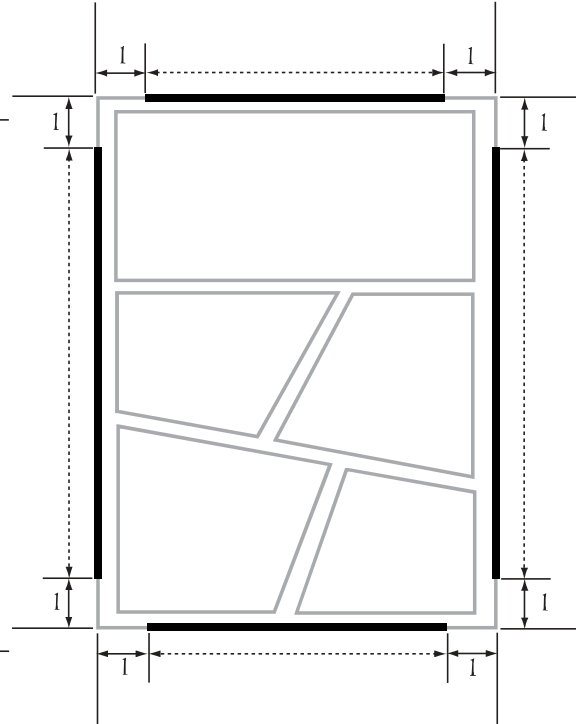


図 3.10 コマ分割処理の始点及び終点位置 (冗
長領域除去後)
自発表 [67] から一部を変更した

3.5.6 コマ位置精度の向上

コマ位置精度向上手法の概要

マンガのコマ周囲には、コマ間を隔てる余白部分が存在する。コマのより正確な位置を得るためには、この余白部分を認識、除去する必要がある。一方で、2分割を繰り返し適用する手法ではコマ周囲に存在する余白部分の除去ができない。コマ周囲の余白部分にはコマからの要素のはみ出しがある場合も多い。コマの位置は不定であり、断ち切りと呼ばれる画像の端部に接続する場合もある。コマからはみ出し及び断ち切りにより、単純な画像周囲からの連結領域を得るだけでは、正しく余白領域を得られない。

コマ内部に存在する様々なマンガの要素によるエッジは、直線検出処理、角検出処理に影響を与える。本処理ではコマ内部の影響を排除し、コマ間の空白領域を検出する。得られたコマの概形に対して直線検出と角検出を組み合わせ、コマの角候補を取得する。コマの角位置候補取得の処理手順を図 3.11 に示す。ここでは、マンガのコマ内部からの画像解析に対する影響を除去するための前処理を導入する。内部要素の影響を除去した画像に対する直線検出と角検出処理の結果と、上記コマ検出結果との組み合わせから、得られたコマの角位置を補正する。

余白領域の検出

マンガのコマ内部の影響を除外し、コマの概形を得るために、コマ間の空白部分の取得を行う。コマ周囲の余白部分にはコマからの要素のはみ出しがある場合も多く、コマの位置は不定であり、

3.5 コマ分割処理の改良

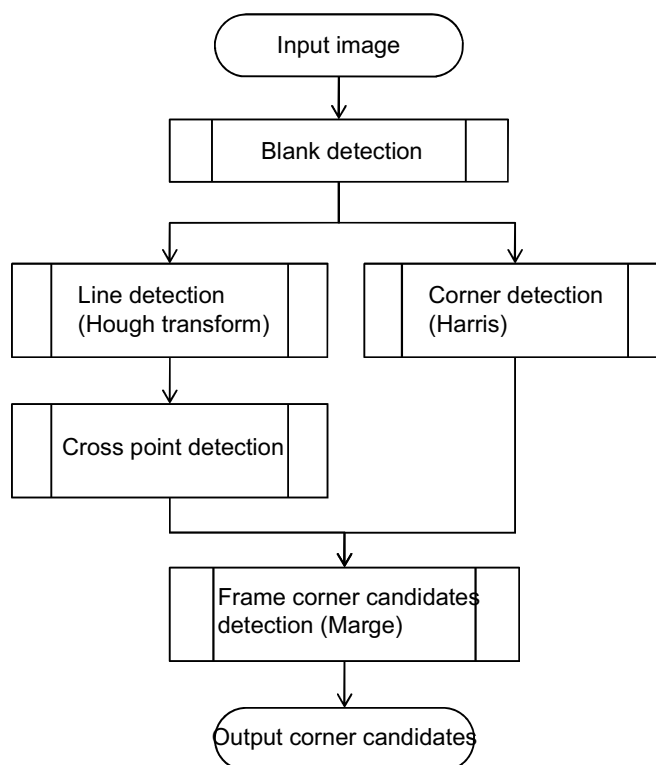


図 3.11 コマの角位置候補取得処理
自発表 [70] を基に一部変更を加えた

画像の端部に接続する場合もある。したがって、画像のどの部分から連結領域の抽出を開始すればよいか不明である。そこで、本余白領域の検出では、画像の外周部分複数箇所から連結領域の取得を行う。複数箇所から始まる連結領域の取得結果を重畳することで、連結領域が途中で途切れている場合や、コマが外周部に接続している場合でもコマ間の空白領域の取得を可能とする。

提案手法では、画像の外周部分に等間隔に、任意数の連結領域取得のための開始点を配置する。各開始点から連結する均等色領域を取得し、投票を行う。しきい値以上の投票値が得られた領域をコマ間の空白領域とする。

図 3.12 に開始点配置箇所の例を、図 3.13 に空白領域の検出結果を示す。ここで、開始点の数を 20 とした。投票数が多い領域を余白領域、投票数の少ない領域をコンテンツ領域として、空白領域を白、コンテンツ領域を黒色で表示した。

コマ角位置候補検出処理と、コマ角位置の修正

余白領域の検出により得られた画像に対して、直線検出処理及び角検出処理を行い、角位置候補の検出を行う。ここでは、直線検出に Hough 変換 [65]、角検出に Harris のコーナー検出処理 [71] を用いることとした。直線検出結果に対して、その検出直線同士の交点を求める。ここで、交点を保持するのは、交差する直線間の角度が 30 度以上となる組み合わせのみとする。得られた交点と、Harris のコーナー検出結果の AND をとる。このようにして得られた結果をコマ角位置候補とする。図 3.14 に余白領域の検出結果に対して Hough 変換による直線検出結果を、3.15 にその交点を示す。また、図 3.16 に同じく余白領域の検出結果に対するコーナー検出結果を、図 3.17

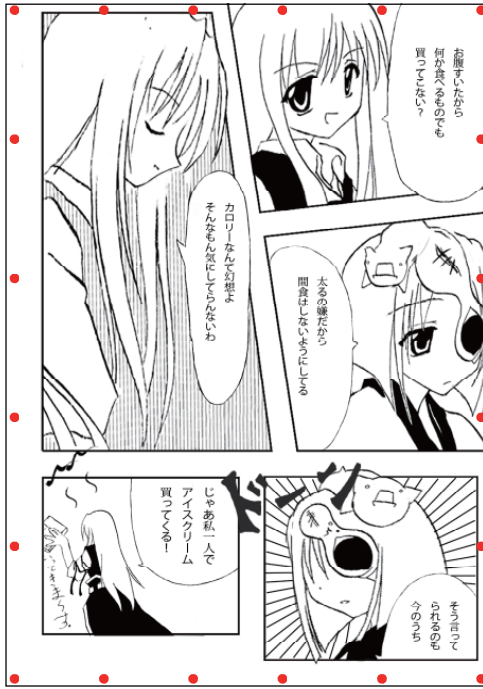


図 3.12 連結領域取得の開始点
自発表 [70] から引用

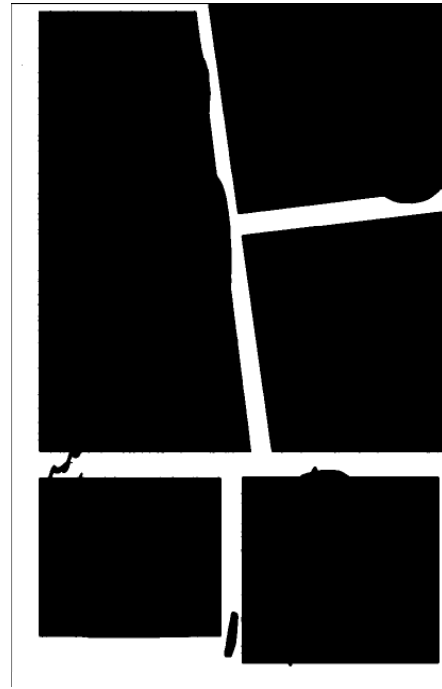


図 3.13 空白領域の検出結果
自発表 [70] から引用

に両処理により得られたコマ角位置候補を示す．コマ角位置の修正では，あらかじめ実施したコマ分割結果の各コマにおける角座標について，最寄りのコマ角位置候補検出処理結果への移動を行う．

3.6 コマ分割実験

本節では，これまでに述べてきたコマ分割処理によるコマ分割実験を行う．低解像度化による分割精度への影響も含めた，コマ分割の処理時間，処理精度並びにメタデータ付与コストの削減と余白領域の除去結果について示す．

3.6.1 画像解像度変換と分割パラメータによるコマ分割処理への影響

コマ分割処理の評価法

■**コマ編集距離** コマの情報及び視線遷移情報となる順序の精度を評価するために，コマ分割に対して編集距離の概念を導入する．編集距離とはある文字列 A に対して，文字の挿入，削除及び置換の操作を行い，他の文字列 B と同一とするために要する操作の数である．マンガにおけるコマを文字と同様にとらえることで編集距離の概念をコマ分割結果に対して適用する．

以下ではコマ編集距離を Comic frame editing distance(CFED) と呼ぶ．ここで，コマは画像の一部領域であり，画像上の全てのコマはある順序で並べられるものとする．CFED 取得時における操作を以下に定義する．

- 「コマの挿入」任意のコマ，あるいは画像上の任意の範囲を新たなコマとして定義し，任意

3.6 コマ分割実験

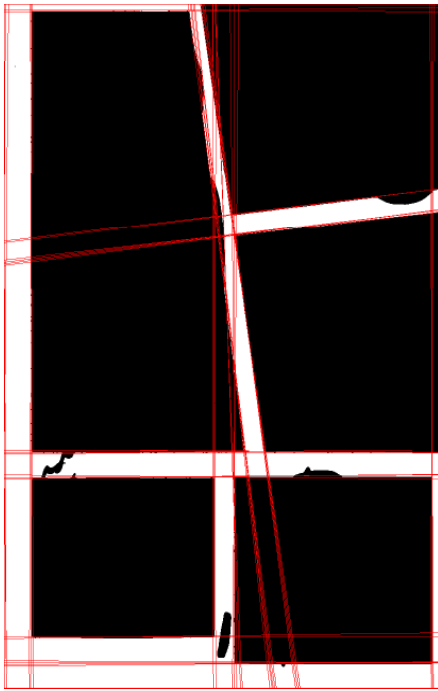


図 3.14 Hough 変換結果
自発表 [70] から引用

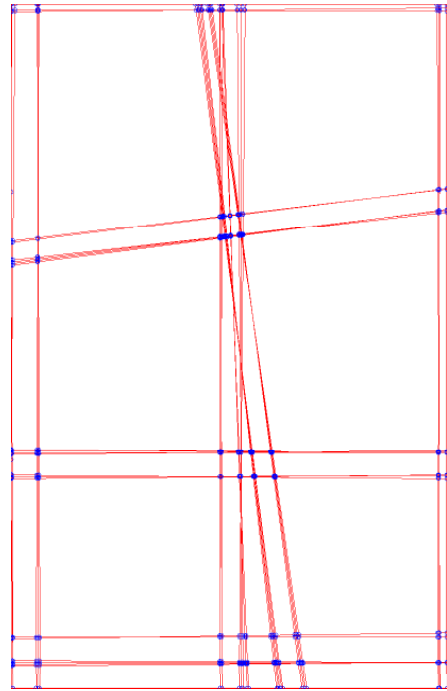


図 3.15 Hough 変換により得られた直線の交点
自発表 [70] から引用



図 3.16 Harris の角検出アルゴリズムによる角
検出結果
自発表 [70] から引用

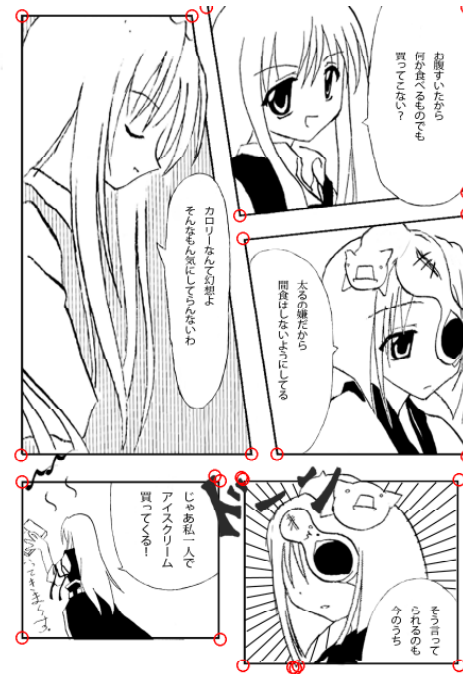


図 3.17 角候補点
自発表 [70] から引用

表 3.1 計算機環境
自発表 [69] から引用

CPU	Intel Core2Quad Q6600 2.4GHz
Memory Size	4GB
OS	Linux

の順番に挿入する

- 「コマの削除」任意のコマを削除する
- 「コマの領域変更」任意のコマの領域を変更する

以下に CFED と操作の例を示す。ここで、あるマンガ画像において、A,B,C,D,E という順番で 5 つのコマが並んでいるとする。

例 1 D のコマを過分割し、A,B,C,D1,D2,E という順番でコマが得られた場合。

操作 1:D2 を削除。結果、A,B,C,D1,E。

操作 2:D1 を D に領域変更。結果、A,B,C,D,E。

よって CFED=2 となる。

例 2 C, D のコマを分割しなかったため、A,B,CD,E という順番でコマが得られた場合。

操作 1:CD を C に領域変更。結果 A,B,C,E。

操作 2:C と E の間に D を挿入。結果 A,B,C,D,E。

よって CFED=2 となる。

例 3 C,D,E を誤った形状 F,G に分割したため、A,B,F,G という順番でコマが得られた場合。

操作 1:F を C に領域変更。結果 A,B,C,G。

操作 2:G を E に領域変更。結果 A,B,C,E。

操作 3:C と E の間に D を挿入。結果 A,B,C,D,E。

よって CFED=3 となる。

CFED を用いることでコマ単位における分割精度を評価する。

■**成功率** CFED は画像中のコマ数が多いほどその値が大きくなる。コマ数が少ない画像では、ほとんど分割が行えていない場合であっても、その値は小さくなる。そこで、ページ単位での成功率についても同時に評価を行う。ここで、マンガ画像 1 ページ内に含まれる全てのコマが規定の順序どおりに全て正しく得られた場合に成功であるとする。上記条件はコマ編集距離が 0 となる場合に満たされる。成功率は全ての対象ページのうちコマ編集距離が 0 となったページの割合を示す。

実験条件

実験には 4 作品合計 110 枚の画像を用いた。実験に用いた計算機環境を表 3.1 に示す。入力画像のサイズ及びページ数を表 3.2 に示す。分割性能の評価では、コマの CFED 及びページ単位成功率を用いた。また、処理速度は画像の入出力、解像度変換、前処理等も含めた総合時間を直接

表 3.2 画像サイズ及び画像枚数
自発表 [69] から引用

	Image Size[pixel×pixel]	Numvers of Pages
Comic-A	740×1200	20
Comic-B	760×1200	30
Comic-C	1024×1536	30
Comic-D	840×1200	30

計測した。

本実験では FAFSM のパラメータはそれぞれ $A = 20, B = -2, E = 10, \rho = 20\%, \delta = 40, n = 7, m = 2$ を用いる。また、今回 FAFSM に対してしきい値 M_{th} を設け、その線上の重み付き平均輝度こう配値が、設定したしきい値以上となるもののみ分割線候補として検出するようにした。 M_{th} の値は実験により最適な値である 60 とした。FAFSM の画像解像度変換には $R \times R$ サイズの平均値フィルタを用いてダウンサンプリングを行う。

比較対象である田中らの手法 [55] を以下では T-method と表記する。T-method では高速化のアプローチがとられていない。そこで FAFSM と同様の処理により画像解像度を縮小することで、実質的な処理時間を削減した。ただし、 $R = 1$ の場合には画像解像度変換は行われず、通常の T-method と同一となる。

コマ分割速度及び分割精度に関する評価

T-method の画像解像度変換によるページ単位の成功率を図 3.18, CFED を図 3.19, 処理時間の変化を図 3.20 に示す。FAFSM の画像解像度変換によるページ単位の成功率を図 3.21, CFED を図 3.22, 処理時間の変化を図 3.23 に示す。T-method では、画像解像度を下げるに従って大幅に成功率が低下している。これに伴い、CFED もコマ数から考えられる最大値に近い値となっている。また、T-method では Comic-A で最大で 85% 程度のページ単位のコマ分割成功率が得られた。一方で、FAFSM では、画像解像度変換による CFED の変化は T-method と比較してゆるく、 $R=2$ までであればページ単位の分割成功率の低下は少ない。また、FAFSM では、Comic-A で最大で 95% 程度のページ単位のコマ分割成功率が得られた。処理手法を問わず、処理時間については、 $R=2$ 以上とすることで大幅に削減されることがわかる。

T-method と FAFSM の処理時間に対するページ単位の成功率を図 3.24 に、CFED を図 3.25 に示す。FAFSM は、T-method と比較して、処理時間に対する識別精度が高いことが確認された。図 3.24 より、FAFSM は 5 秒未満の処理時間で 80% 程度のページ単位正解率が得られている。また、図 3.25 より、T-method と比較して、処理時間は少なく、且つ編集距離が削減できていることがわかる。FAFSM では、従来手法と比較して、コマ情報付与にかかるコストが削減できることが確認された。

次に、FAFSM による、各マンガに対する CFED の分布について示す。縦軸の number は、横軸に示される編集距離となるページ数を表す。全ての作品で、画像解像度変換の係数を大きくすることで、より高い CFED のページ数が増える。Comic-B ではすべての R において、CFED=3,4

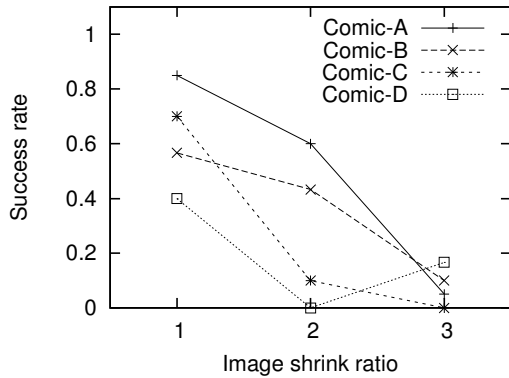


図 3.18 T-method における画像解像度変換による成功率の変化
自発表 [69] から引用

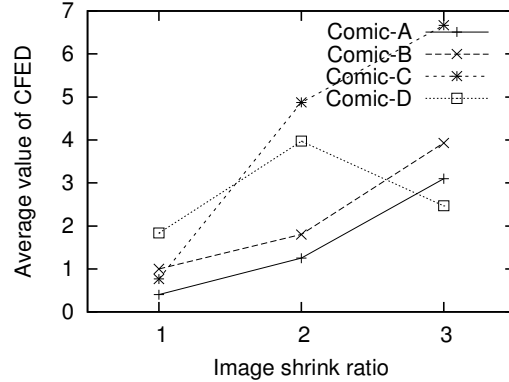


図 3.19 T-method における画像解像度変換による CFED の変化
自発表 [69] から引用

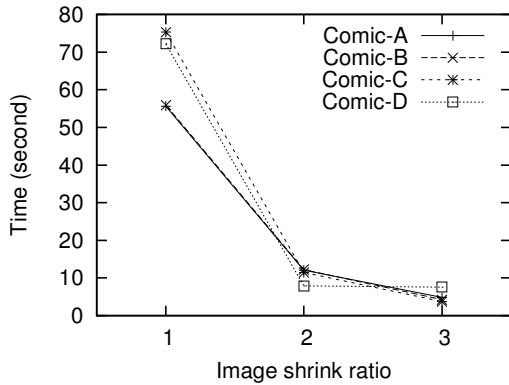


図 3.20 T-method における画像解像度変換による処理時間の変化
自発表 [69] から引用

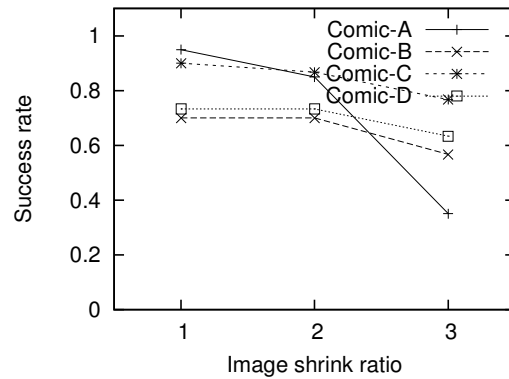


図 3.21 FAFSM における画像解像度変換による成功率の変化
自発表 [69] から引用

の数が多い結果となった。この原因として、Comic-B は他の作品と比較して、コマからの要素のはみ出し、矩形でないコマの数、コマ内部の直線成分が非常に多いことが考えられる。特に、コマ内の直線成分を検出した場合の誤分割はページ全体の CFED を大きく引き上げる可能性がある。したがって、コマ情報の記述を補助するためには、コマ分割処理の誤分割を削減することが重要である。また、コマからはみ出しが多い場合には、平均 CFED を考慮して、コマ分割処理の利用を検討することが必要であると考えられる。

3.6.2 コマ余白除去実験

実験条件と評価法

提案手法により得られた各コマの角位置と、FAFSM により得られた各コマ領域の角座標について比較検討を行う。今回の実験では、得られた角の精度評価に、以下の式に示される Precision と Recall を用いた。これまで検討を行ってきたコマ分割では 2 分割処理を繰り返して、コマ毎に切り離された画像を出力するため、今回は分割された各領域の角座標をそれぞれの手法におけ

3.6 コマ分割実験

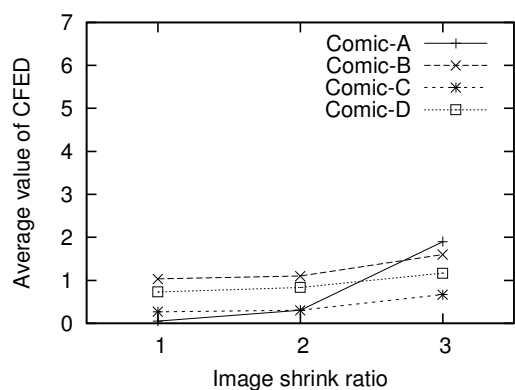


図 3.22 FAFSM における画像解像度変換による CFED の変化
自発表 [69] から引用

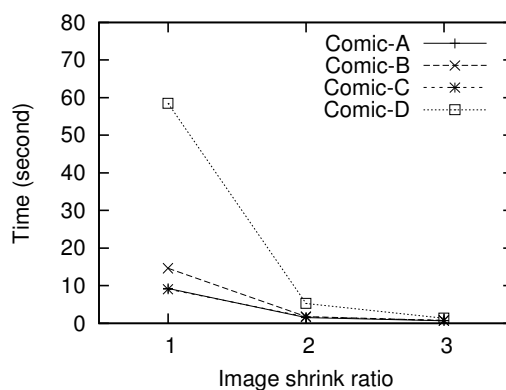


図 3.23 FAFSM における画像解像度変換による処理時間の変化
自発表 [69] から引用

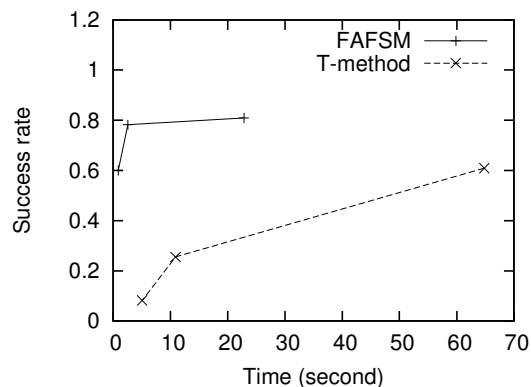


図 3.24 各手法における処理時間と成功率
自発表 [69] から引用

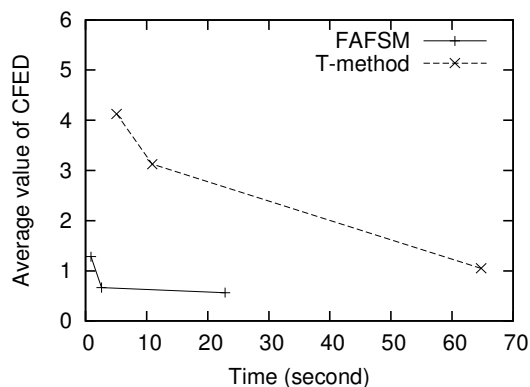


図 3.25 各手法における処理時間と CFED
自発表 [69] から引用

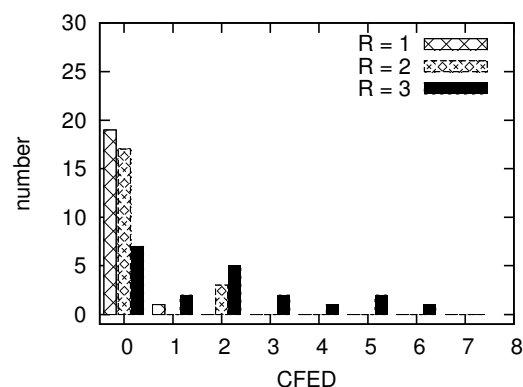


図 3.26 FAFSM における画像解像度変換による CFED の変化 Comic-A
自発表 [69] から引用

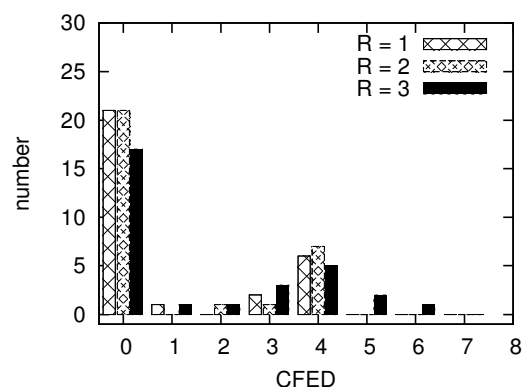


図 3.27 FAFSM における画像解像度変換による CFED の変化 Comic-B
自発表 [69] から引用

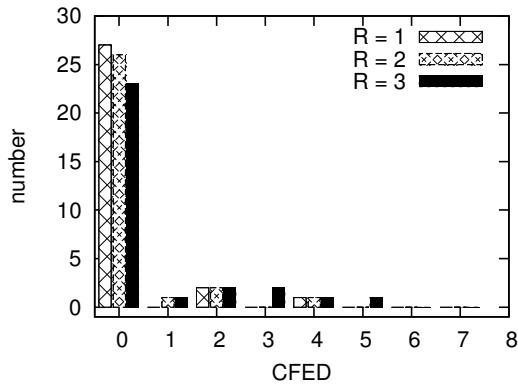


図 3.28 FAFSM における画像解像度変換による CFED の変化 Comic-C
自発表 [69] から引用

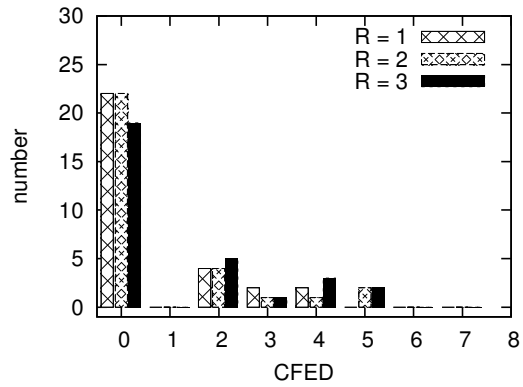


図 3.29 FAFSM における画像解像度変換による CFED の変化 Comic-D
自発表 [69] から引用

る検出されたコマ角位置として比較を行った。実験の対象画像は以下の通りである。画像サイズ及び画像枚数は表 3.2 に示したものと同一である。マンガ B 及び C はコマからはみ出し量の多いシーケンスである。今回入力した画像は基本的なコマの構造により構成されるマンガである。

$$\text{Precision} = \frac{\text{correctcorners} \cap \text{obtainedcorners}}{\text{obtainedcorners}}, \quad (3.4)$$

$$\text{Recall} = \frac{\text{correctcorners} \cap \text{obtainedcorners}}{\text{correctcorners}}. \quad (3.5)$$

コマの余白除去に関する評価

コマ余白除去について実験を行った結果の一例を図 3.30 に示す。図 3.30 において、○は余白除去アルゴリズムにて得られた角候補を、点線は FAFSM により得られた各コマの境界線を、□は FAFSM で得られた角の位置をそれぞれ示す。FAFSM の結果は、分割後も空白部分が残ってしまうため、実際のコマの角位置は取得できない場合が多い。図より、コマ位置精度の向上による結果の点は実際のコマの角位置に近い場所にあることが分かる。得られた角候補とコマ分割結果に対し統合処理を行い、得られた角位置に基づき作図した結果を図 3.31 に示す。統合処理では、コマ分割結果の角から最近傍となる角候補位置への移動を行った。図 3.30 にて得られている FAFSM のコマ分割結果と比較して、統合処理結果では、コマ周囲の余白領域が減少していることが確認できる。

入力画像 110 枚に対して評価を行った結果を表 3.3 及び表 3.4 にそれぞれ示す。ここでは、コマの角座標を正解として評価を行った。空白検出及びコーナー検出のみを行った場合の結果を Propose*、Propose*に対して Hough 変換の交点との AND をとった結果を Propose とする。また、本表中の Conv.I は FAFSM、Conv.T は T-method を示す。Propose*と Propose の比較により、提案手法における Hough 変換の交点による制限を導入した場合には、すでに取得されている候補の Recall の低下なしに Precision を向上することが確認された。FAFSM との比較実験からコマ位置精度向上処理の有効性が確認された。

3.6 コマ分割実験

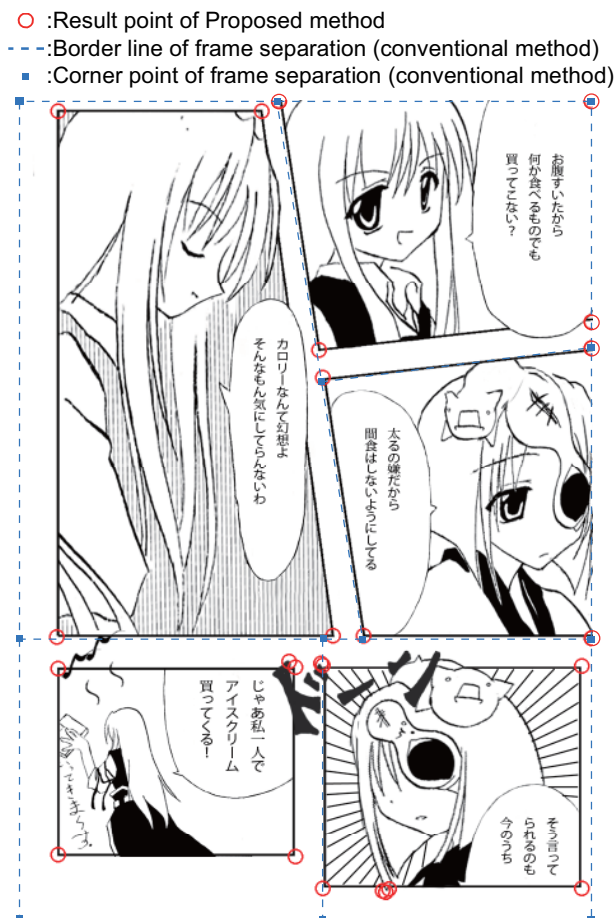


図 3.30 提案手法の結果と、従来のコマ分割手法により分けられた各コマの境界とその角
 自発表 [67] から引用

表 3.3 コマの角位置解析実験における Precision
 自発表 [67] から引用

Works	Propose	Propose*	Conv. I	Conv. T
Comics A	0.936	0.588	0.260	0.258
Comics B	0.778	0.350	0.199	0.214
Comics C	0.708	0.159	0.043	0.057
Comics D	0.937	0.656	0.271	0.212

表 3.4 コマの角位置解析実験における Recall
 自発表 [67] から引用

Works	Propose	Propose*	Conv. I	Conv. T
Comics A	0.936	0.990	0.263	0.245
Comics B	0.958	0.972	0.179	0.196
Comics C	0.908	0.916	0.038	0.040
Comics D	0.976	0.984	0.267	0.202

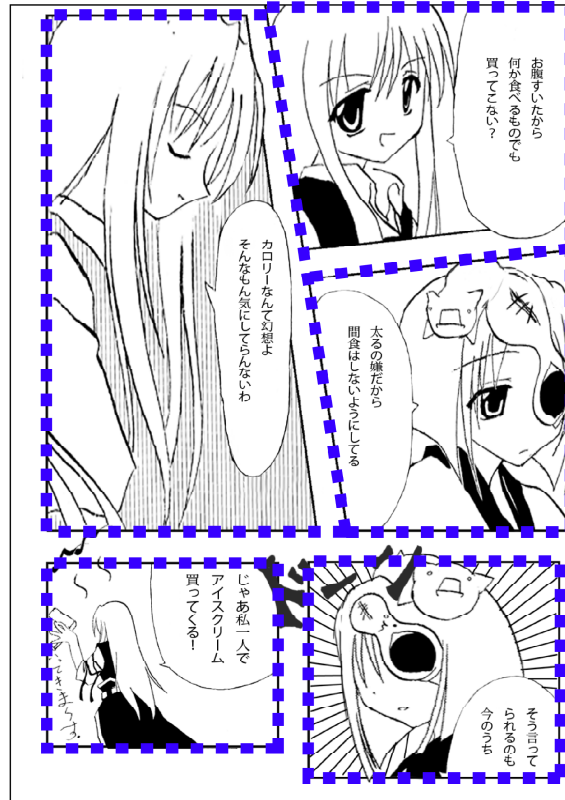


図 3.31 余白除去アルゴリズム適用後のコマ領域
自発表 [70] から引用

3.7 むすび

本章では、コマ分割処理の分割精度及び処理速度向上のためのアルゴリズムを提案した。従来の分割処理では、分割線の決定が輝度こう配のみに依存するため、本来分割不可能な部位にて分割処理が行われるという問題があった。提案手法では分割線自体に幅を持たせることにより、微小な角度変化に対するロバスト性を向上させ、分割線の探索数を減少させた。コマの内部では、様々な絵の成分が存在するためその方向分布も複雑となる。上記特徴を利用し、他のコマ上を通過する分割線が決定されることを防ぐため、分割線上の輝度こう配方向分布等を利用した分割線の検査を提案した。また、分割線検出順序の変更を導入し、従来手法と比較して高精度かつ高速な処理を可能とした。本提案手法は、分割線候補の交差点近辺の直線方向を基に優先分割線を決定するアルゴリズムにより改良された再帰的 2 分割手法と比較しても高い分割精度を得た。

通常のコマ分割処理では、アルゴリズムの特性から、コマ周辺の余白部が残る。マンガでは、コマからの要素のはみだし等により、結果として、分割後のコマ周辺にある白画素が余白であるか否かの判断ができない。そこで、コマの内部塗り潰しアルゴリズムによるコマ内絵領域の除去の後、コーナー検出結果と Hough 変換結果の組み合わせによるコマ角位置検出によるコマ位置精度の向上手法を提案した。提案手法により従来 2 分割処理の繰り返しアルゴリズムでは考慮されない、コマ周辺部の余白除去を可能とした。

第 4 章

マンガ画像からのキャラクター情報抽出

4.1 まえがき

本章では、マンガ画像からのキャラクター情報の抽出を行う。キャラクターはマンガのストーリーを構成する上で大変重要な要素である。マンガにおいて、キャラクターは読者に識別可能なレベルで描き分けがなされている。キャラクター情報の抽出はキャラクター検出とキャラクター識別の 2 段階に分けられる。マンガのメタデータとして利用することを考えた場合には、誰がいるのかという単位でメタデータの取得ができることが望ましい。そのため、キャラクターの検出だけでなく、キャラクターの識別についても検討を行う。

物体検出は任意の対象物と同一の物体を検出する特定物体認識と、検出対象カテゴリに属するものを画像の中から発見若しくは選別する一般物体認識に分けられる。近年、一般物体認識の一分野として顔検出や人物検出が検討及び実用化されてきた。[36–42]。識別対象から得られる画像特徴量を、統計的機械学習等を用いて学習することで、同一のアルゴリズムを用いて学習させた、任意の識別対象に対する識別器を得ることができる。一般物体認識の一処理である人物顔検出は、今日ではほぼすべてのデジタルカメラに搭載されている。このように画像解析技術は今日の技術発展の中で大きな役割を担っている。これまでアプリケーションや、その目的及び検出対象に応じてさまざまな画像特徴量が提案されてきた。一般物体認識に利用される画像特徴量として、Haar-like [39, 40]、HOG [41, 42] などが提案されている。

顔検出に使用される Haar-like 特徴量は、小領域内の平均輝度情報に基づく特徴である。Haar-like 特徴量は人間の瞳や頬の明るさの違いを捉えることができる。一方で、マンガのキャラクターは主に線によって描かれるため、マンガ画像において平均輝度から得られる情報は人間の顔検出の場合と比較して少ない。そのため、平均輝度情報以外の画像特徴を利用する必要がある。

本章では、人間のシルエットを捉えるために使用される HOG 特徴量に着目する。HOG 特徴量では、少領域内の方向成分の分布を捉えることができる。HOG 特徴量を利用し、瞳、顔の学習によりそれぞれ識別器を生成する。更に、キャラクターごとに学習を行ったキャラクター識別器を生成する。最終的に構成される識別器は瞳、顔及びキャラクターの 3 識別器をカスケード型に接続させた、特定キャラクター検出処理を提案する。また、マンガのキャラクターはその向きやシーンにより大変大きな違いをもって描き分けられる場合がある。このような場合に、同一クラスタとして学習処理した場合には、その識別精度を低下させることとなる。そこで、キャラクターの状態ごとに学習を行った場合の識別率の変化についても検討を行った。

4.2 人物検出手法

4.2.1 物体検出のアプローチ

近年物体検出及び認識処理に関する研究として、様々な目的に応じた画像特徴量及び機械学習手法が提案されている。物体検出処理では、検出したい対象である対象領域の画像から、特定の画像特徴量を計算する。同一物体の検出において、検出対象が同一のものであっても、その撮像条件等により画像上では若干の変化が生じる場合がある。このような影響を低減し、更にスケール不変性、回転不変性等まで考慮された画像特徴量等が提案されている。画像特徴量は、様々な画像的変動の影響を最小限に抑え、識別に有効な情報を取り出すために使用される。特定物体認識では、画像特徴量を用い、特定の部位に関する特徴量の一致を図る。これにより別画像からでも検出対象と同一の特徴を持つ箇所を検出が可能である。一般的に剛体の検出には特定物体認識が使用される。一般物体認識では、検出対象カテゴリごとに画像特徴量を計算し、これを統計的機械学習等により学習する。このようにして構成された識別器により、注目する画像領域内における特徴量が検出対象カテゴリに属するか否かを判別する。マンガのキャラクターのように、検出対象に形状変化のある場合には、一般物体認識のアプローチが有効である。

画像からの検出では、複数サイズの検出ウィンドウを用いた画像内をラスタスキャンが行われる。各ウィンドウ内における画像特徴量を計算し、これが検出対象となる画像特徴量と一致するか、若しくは特定カテゴリの特徴量空間に属するかを判定し、検出結果を得る。ラスタスキャンによる検出では、特定カテゴリの検出の場合などに、同一の検出対象を捉えた、位置、スケールのバリエーションによる多数の検出結果が得られる。結果に対して、MeanShift クラスタリング [72] 等による統合処理を追加することで、目的に応じた有意な検出結果を得ることも可能である。

4.2.2 人物検出処理と画像特徴量

画像解析の分野では、近年の機械学習の発展とともに目的に応じた多様な画像特徴量が提案されている。特定物体検出に用いられる画像特徴量として SIFT [73, 74] がある。SIFT では、スケール及び回転不変なキーポイントを記述可能であり、他方向から撮像された同一物体の検出等に利用されている。一般物体検出に用いられる画像特徴量として Haar-like [39] や HOG [41, 42] 等がある。Haar-like 特徴量は局所領域間の平均輝度差を求めるもので、人物顔検出等において高速かつ高精度な検出が可能である。HOG 特徴量は局所的なエッジ方向の分布に着目した特徴量であり、セルと呼ばれる小領域の位置情報と共に、その中に含まれるエッジ成分を取得できるため、人物検出や車検出等に有効な特徴量である。

SIFT 特徴量

スケール変化及び回転にロバストな画像特徴量として Scale-Invariant Feature Transform (SIFT) 特徴量が提案されている。近年特定物体認識や、画像の照合などに広く用いられている。SIFT は、Difference of Gaussian (DoG) を用いてスケール不変なキーポイントを検出し、SIFT descriptor により回転不変な特徴量を記述する手法である。そのため SIFT 特徴量は、回転・ス

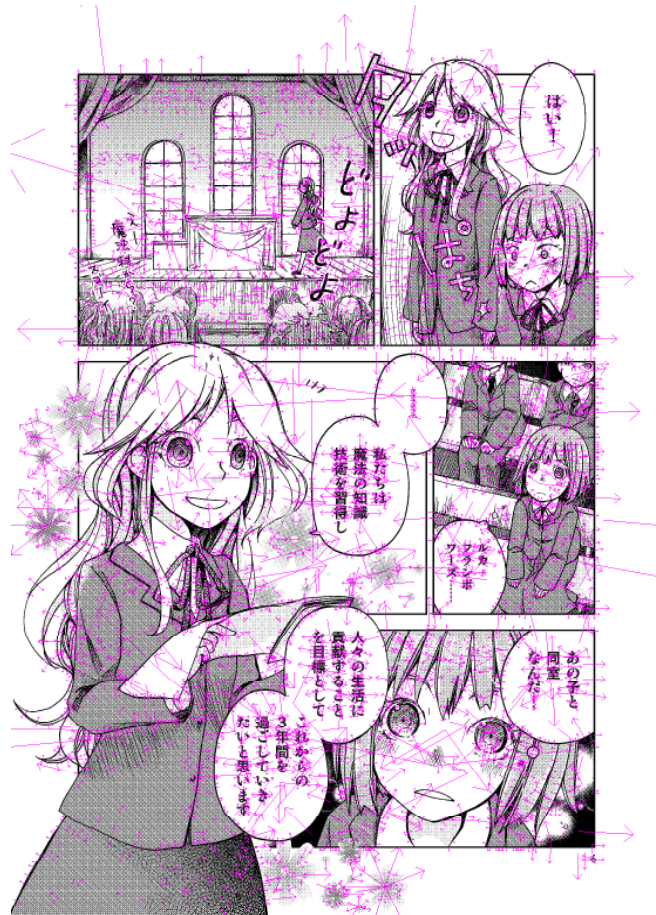


図 4.1 マンガ画像上の SIFT 特徴量
自発表 [75] から引用

ケール変化にロバストな特徴量となっている。

一方でマンガ画像では、線画、トーン部分双方において、SIFT のキーポイントになりうるエッジ成分が無数に存在する。このため、画像中のあらゆる箇所に SIFT のキーポイントが検出されてしまう。マンガ画像上で SIFT 特徴量取得した結果例を図 4.1 に示す。無数のキーポイントが検出されていることが確認できる。SIFT 特徴量は全く同一の画像や部位を照合する目的には使用できる。対象物体に変動がある場合には、キーポイント自体も変化することになるため、SIFT 特徴量は変化のある特定のキャラクター等を見つけ出す用途には向かない。

Haar-Like 特徴量

Viola と Jones によって提案された Haar-like 特徴量を利用した顔画像検出 [39, 40] は、顔画像検出分野において一般的に使用される手法である。Haar-like 特徴量は画像内局所領域の輝度の分布特徴を捉える。特に人物画像では、人物の瞳部分は暗く、周囲の肌部分では明るい輝度が存在する。Haar-like 特徴により人間の顔に特有な輝度分布をとらえることができる。Haar-like 特徴量は 2 つの局所領域に含まれる合計輝度の差から特徴量を求める。この局所領域の形状は検出対象に合わせ多数のバリエーションをもつ。顔検出では瞳、鼻、口などに対してそれぞれ異なる形状の Haar-like 特徴量を用いた検出処理を行い、その結果をカスケード接続することで、高い精度

4.2 人物検出手法

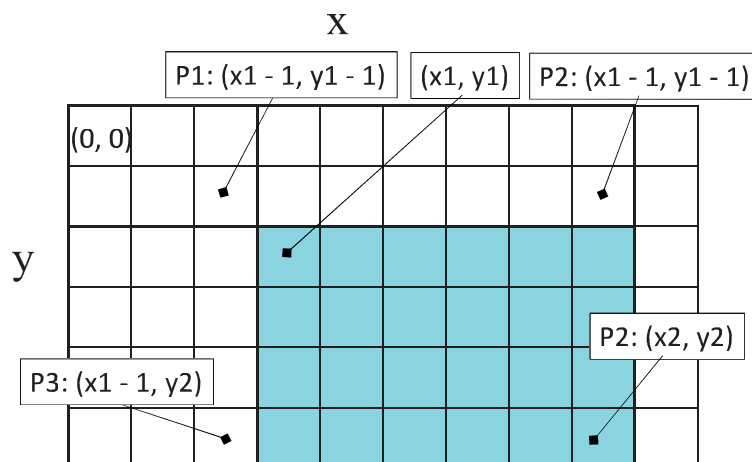


図 4.2 IntegralImage 参照画素
自発表 [76] から引用

で人物の顔検出が可能である。

Haar-like 特徴量の処理過程では、多数の局所領域内の合計輝度を算出する必要があるが、Integral Image を利用することで計算の高速化が図られている。Integral Image の各画素の値は、画像の起点 (0,0) から当該画素 (x, y) を対角点とする矩形領域内に含まれる輝度の合計値として算出される。画像の左上を起点とした場合、画像の右下に向けて Integral Image の値は徐々に大きくなる。矩形領域内の輝度合計値の計算には、局所領域内に含まれる画素数分の参照が必要であったが、Integral Image を用いることで、4 点の参照及び計算のみで実現可能となる。図 4.2 に示される (x1, y1) と (x2, y2) を対角とする矩形領域内の輝度合計値を求める場合、Integral Image 上の P1 から P4 で示された 4 点から $P4 - P2 - P3 + P1$ を計算することでその値が得られる。Integral Image を利用することで多数の矩形領域内の輝度合計値を得る場合に計算コストを削減可能である。

マンガ画像は主に線画により主要成分が描かれている。また、もともとほぼ 2 値のデータであるため、Haar-like 特徴量を計算すると、線画の書き込み量やドットの頻度が得られるが、線の方角等に関する情報は得られない。キャラクターの識別課程において、顔の平均輝度からなる特徴はキャラクター間では似た状態である可能性が高く、解析に十分な情報が得られないと考えられる。

HOG 特徴量

HOG [41] は画像内の輝度勾配方向の分布に着目した特徴量である。人物輪郭などのエッジを捉えることができ、学習により、同一の検出対象がバリエーションを持つ場合にも対応可能である。

HOG 特徴量の計算では、各画素について、上下方向の輝度差 V_d と左右方向の輝度差 H_d から下記の式により勾配ベクトルの大きさ m とその向き θ を計算する。

$$m(x, y) = \sqrt{V_d(x, y)^2 + H_d(x, y)^2} \quad (4.1)$$

$$\theta(x, y) = \tan^{-1} \frac{V_d(x, y)}{H_d(x, y)} \quad (4.2)$$

このようにして求めた係数について、画像を複数の小さなセルに分け、そのセル内においてヒストグラム化する。ヒストグラム化にあたり、方向成分 θ については、 $[0^\circ, 180^\circ)$ の範囲に射影し、その中で任意の N 方向に量子化する。この方向に基づいて、 $m(x, y)$ の値を投票しヒストグラムを生成する。また、複数セルをまとめてブロックとし、ブロック単位にて正規化を行う。HOG 特徴量の次元数は下記の式により求める。

$$\begin{aligned} HOG(Dimension) = & CellNumber \\ & \times DirectionNumber(N) \\ & \times BlockNumber \end{aligned} \quad (4.3)$$

ここで、画像内において、一セルずつ順番にブロックを移動し正規化を行う。総ブロック数は、画像のセル数が縦 7、横 5、ブロック内に含まれるセル数が縦 3 セル、横 3 セルであった場合に 12 となる。このとき、ブロック内のセル数は 9 であり、方向量子化数 N を 9 とすると、HOG 特徴量の次元数は 972 となる。

HOG 特徴量においても、画像をラスタスキャンしつつ対象物体を探索する場合には、セル内のヒストグラム計算を繰り返し実施することとなり、膨大な計算コストが生じる。この計算コスト削減のために、Haar-like 特徴量と同様に、方向成分ごとに Integral Image を作成する手法 [42] が提案されており、現実的な処理時間で画像内ラスタスキャンによる人物検出処理等が可能である。

アニメキャラクター検出

マンガと似たコンテンツとしてアニメがある。アニメのキャラクターも、マンガと基を一つにする場合もあり、その構造は非常に近い。最も大きな違いとして、現状のアニメーションは基本的にカラーで彩色されており、マンガには存在しない色情報及び濃淡情報を利用することができる。高山らはアニメキャラクターの識別手法を提案している [77, 78]。ここで提案される識別処理では、顔を構成する目やアゴの輪郭形状を特徴として、識別時の規範として利用している。アニメキャラクター検出では、処理の多くを上記濃淡情報及び色情報に依存しており、マンガ画像に適用することが出来ない。

4.3 マンガ画像からのキャラクター検出手法

4.3.1 キャラクター検出のアプローチ

一般的にマンガでは、作者の想像に依存して描かれるため変動が大きい。キャラクターの存在有無の特定については、顔全体の安定した検出は困難である場合がある。一方で、瞳はマンガの中において、他の箇所と比較して特徴的な描かれ方をする箇所であり、瞳が検出できれば、そこにキャラクターがいる可能性が高いと判断できる。顔全体の検出処理及び瞳部分の検出処理を組み合わせることで、キャラクターの存在有無を捉えることが考えられる。次に、キャラクターの

4.3 マンガ画像からのキャラクター検出手法

顔位置の特定については、顔を学習した検出器を用い、スライドウィンドウを用いた検索を行う方法が考えられる。また、瞳の検出結果が得られている場合には、瞳周辺に対してのみ顔の検出処理を行うことも可能である。キャラクターの識別については、検出対象キャラクターごとに当該の検出対象キャラクターと検出対象者以外のキャラクター画像を学習する。得られた対象キャラクター識別器を用いて検出対象者であるか否かの識別を行う。キャラクター検出には、複数の検出器を組み合わせるアプローチをとる。

HOG 特徴量は、対象領域を複数のセルに分割し、各領域内で方向ヒストグラムを計算する。得られた方向ヒストグラムをまとめて多次元ベクトルを構成し特徴量としている。これにより、画像内部の物体形状を捉えるのに適した画像特徴量である。HOG 特徴量はシルエットによる人体検出などに利用されている。マンガは多数の線の組み合わせを駆使することでキャラクターを表現している。したがって、これはキャラクターにおけるシルエットの形状を表現していることと近い。方向性の情報を捉えることが可能な HOG 特徴量は、マンガ画像の解析にも有効であると考えられる。そこで本章ではこの HOG 特徴量と機械学習の組み合わせによりキャラクターの検出および識別を行う。

4.3.2 瞳、顔及びキャラクター識別器とその組み合わせ

マンガにおけるキャラクター検出に際し、以下の3種類の識別器を検討する。ただし、すべての学習用正例及び負例の画像は、あらかじめ手作業にて切り出されたものを使用する。

瞳識別器: 瞳部分の画像を正例、瞳以外の部分の画像を負例として学習を行ったもの。

顔識別器: 顔部分の画像を正例、顔以外の部分の画像を負例として学習を行ったもの。

キャラクター識別器: 検出対象キャラクターの顔を正例、検出対象以外のキャラクターの顔を負例として学習を行ったもの。

予備調査により、各識別器は多くの場合弱い識別器となり、多数の FalsePositive を結果として出力することが確認された。本論文ではこれらの組み合わせによる瞳検出、顔検出、キャラクター識別処理について提案する。各ステップを考慮した上記識別器の組み合わせについては以下のパターンが考えられる。

Eye:

画像からスライドウィンドウと瞳識別器を用いて瞳の検出を行う

Face:

画像からスライドウィンドウと顔識別器を用いて顔の検出を行う

Character:

画像からスライドウィンドウとキャラクター識別器を用いて得られた領域が、特定キャラクターの顔であるか識別を行う

Eye-Face(E-F):

Eye を実施し瞳の候補領域を取得後、その領域の周囲に対して顔識別器を用いて顔検出を行う

Eye-Character(E-C):

Eye を実施し瞳の候補領域を取得後、その領域に対してキャラクター識別器を用いて特定キャラクターの顔であるか識別を行う

Face-Character(F-C):

Face を実施し、顔の候補領域を取得後、その領域に対してキャラクター識別器を用いて特定キャラクターの顔であるか識別を行う

Eye-Face-Character(E-F-C):

Eye-Face を実施し、顔の候補領域を取得後、その領域に対してキャラクター識別器を用いて特定キャラクターの顔であるか識別を行う

上記識別に用いる特徴量として HOG 特徴量を利用する。Eye, Face, Character では、入力されたマンガ画像に対してスライドウィンドウを用いて瞳識別器、顔識別器、キャラクター識別器を用いて検出処理を行う。Eye-Face では、瞳検出結果の周囲に対して、複数サイズの領域で新たにウィンドウを定義し、このウィンドウに対して顔識別器による判別処理を行う。顔と認められる範囲があれば、当該ウィンドウを顔と判定する。Eye-Character(E-C), Face-Character(F-C) はそれぞれ瞳検出、顔検出の結果に対して顔識別処理を行い、検出対象のキャラクターか否かの判定を行う。Eye-Face-Character(E-F-C) では Eye-Face の結果として顔であると判定されたウィンドウに対して顔識別処理を実施する。また、マンガ画像における顔の変化度合いを考慮した場合に、瞳の方がその変化パターン総数が少ないと予想されるため、顔の検出ができない場合でもキャラクターの存在を検知するために、瞳の有無から得られる情報を優先させることが有効であると考えられる。顔及び特定キャラクター検出における瞳検出と顔検出の順番については、先に瞳検出を行うこととした。

4.4 キャラクター検出実験

評価実験では、まずキャラクター識別器について、あらかじめ手作業により切り出された顔画像に対する識別実験結果について示す。続いて、提案手法として述べた各識別の組み合わせについて、実際にマンガ画像に対して実験を行った結果を示す。また、上記組み合わせに加え、一部を Haar-like 特徴量を用いて置き換え、HOG 特徴量と Haar-like 特徴量の比較実験結果も併せて示す。

4.4.1 実験条件

SVM の実装には Chang らによる LIBSVM [54] を用いた。HOG 特徴量は方向の正規化数を 9、ブロックサイズを 3×3 セル、セルのサイズは HOG 特徴量計算対象の画像サイズに対して適応的に決定することとし、対象画像を縦方向横方向それぞれ 5 分割するように設定した。従って

4.4 キャラクター検出実験

表 4.1 顔識別の学習と実験に用いた画像数
自発表 [75] から引用

Data Set	Known		Unknown	
Character	Positive	Negative	Positive	Negative
A	93	32	92	86
A2	36	32	92	86
B	26	45	19	142

HOG 特徴量の次元数は $9(\text{ブロック内に含まれるセルの数}) \times 9(\text{正規化回数}) \times 9(\text{方向}) = 729(\text{次元})$ となる。比較実験に用いた Haar-like 特徴量については、OpenCV [79] の実装を使用した。また、Haar-like 特徴量の一般的な利用状況を考慮し、学習には Adaboost を利用した OpenCV の実装である traincascade を用いた。予備実験より学習用画像のサイズは、瞳、顔検出用は 16×16 、キャラクター識別用は 32×32 画素とした。

4.4.2 キャラクター識別器に関する実験

1 作品に登場する主要な 2 人のキャラクター A,B について、本人を正例、本人以外のキャラクターを負例とし、予め切り出された顔画像群に対して識別実験を行った。学習に使用した画像枚数を表 4.1 に、結果を表 4.2 にそれぞれ示す。A2 はキャラクター A に対する学習用正例画像の切り出しもとなる画像数を制限し、学習枚数を減らしたものである。学習に使用した画像に対する識別結果を Known、学習に使用していない未知の画像に対する識別結果を Unknown にそれぞれ示す。ここで、学習及び識別対象の顔画像については、様々な方向を含む顔画像群を使用した。評価には Precision と Recall 及びその平均値である F 値を利用した。ここで

$$Precision = (\text{検出された正解数}) / (\text{全検出結果数}) \quad (4.4)$$

$$Recall = (\text{検出された正解数}) / (\text{全正解数}) \quad (4.5)$$

$$F \text{ 値} = (2 * Precision * Recall) / (Precision + Recall) \quad (4.6)$$

とした。また、検出された正解数は検出対象キャラクターの画像に対して Positive と判定した数、前検出結果数は、本人・他人を問わず Positive と判定された数、全正解数は実験対象画像群全体に含まれる、検出対象キャラクター画像の総数である。各キャラクターの学習に使用したキャラクター A 及び B の正例画像の一部を、図 4.3 及び図 4.4 にそれぞれ示す。

瞳、キャラクター及び特定キャラクター検出について

各識別器の組み合わせによる検出実験を行いその結果を示す。ここでは、識別器の組み合わせとして 4.3.2 に示した HOG 特徴量を用いた組み合わせに対して、下記の Haar-like 特徴量を用いた検出、識別処理を追加し比較を行う。

Eye(HaE):

画像からスライドウィンドウと Haar-like 特徴による瞳識別器を用いて瞳の検出を行う

表 4.2 顔識別成功率
自発表 [75] から引用

Method	Character	Known		
		Precision	Recall	F
HOG	A	0.94	1.00	0.97
HOG	A2	0.95	1.00	0.97
HOG	B	1.00	1.00	1.00
Haar-like	A	0.76	1.00	0.87
Haar-like	A2	0.61	0.97	0.75
Haar-like	B	0.75	0.92	0.83

Method	Character	Unknown		
		Precision	Recall	F
HOG	A	0.57	0.99	0.72
HOG	A2	0.58	0.92	0.71
HOG	B	0.55	0.58	0.56
Haar-like	A	0.56	0.85	0.68
Haar-like	A2	0.52	0.39	0.45
Haar-like	B	0.28	0.53	0.36

Face(HaF):

画像からスライドウィンドウと Haar-like 特徴による顔識別器を用いて顔の検出を行う

Face-Character(HaF-C):

(HaF) を実施し、顔の候補領域を取得後、その領域に対して HOG 特徴によるキャラクター識別器を用いて特定キャラクターの顔であるか識別を行う

Face-Character(HaF-HaC):

(HaF) を実施し、顔の候補領域を取得後、その領域に対して Haar-like 特徴によるキャラクター識別器を用いて特定キャラクターの顔であるか識別を行う

Eye-Face-Character(E-F-HaC):

(E-F) を実施し、顔の候補領域を HOG 特徴により取得後、その領域に対して Haar-like 特徴によるキャラクター識別器を用いて特定のキャラクターの顔であるか識別を行う

キャラクター (A,B) に対する識別器の学習に用いた画像及び画像数は前節で示したものと同一である。実験で用いた入力画像の解像度は 1342×1877 pixel, スライドウィンドウのサイズは 50×50 , 100×100 , 200×200 , 400×400 pixel とし、スライドウィンドウのスライド幅は、ウィンドウサイズの $1/5$ とした。瞳検出後に実行される顔識別処理のウィンドウサイズは検出された瞳のウィンドウサイズの 3 倍, 4 倍, 5 倍とした。各ウィンドウによって得られた領域はすべて 100×100 画素に正規化を行ない HOG 特徴量の計算を行うものとする。スライドウィンドウの

4.4 キャラクター検出実験

サイズ及び正規化解像度等は入力画像のサイズに応じて適宜変更する必要があるが、予備実験から、今回の入力画像に対して良好な結果の得られた値を使用した。今回は、合計 27 枚の画像に対して実験を実施した。

瞳検出器の学習に用いた正例画像の一部を図 4.5 に、顔検出器の学習に用いた正例画像の一部を図 4.6 に、双方に用いた負例画像の一部を図 4.7 にそれぞれ示す。

上記各組み合わせに対する識別結果として、学習に利用した画像を含む既知の画像に対する識別結果を表 4.3 に、学習に利用した画像を含まない未知の画像に対する識別結果を表 4.4 にそれぞれ示す。今回検出対象画像上における全ての検出対象を記述したデータが用意出来なかった。このため Recall についてはこれを求めず、Precision のみの評価とした。Precision の評価基準として、Eye に対しては得られた領域が主に瞳部分を含む場合、Face 及び E-F については得られた領域が主に顔部分を含む場合を Correct とした。また、特定キャラクターの検出については、対応する各キャラクターの顔を含む場合のみを Correct として評価を行った。上記に含まれない全ての検出結果は Incorrect としてカウントし、Precision は

$$Precision = Correct / (Correct + Incorrect) \quad (4.7)$$

とした。

次に、学習に利用した既知画像に対する検出結果例を図 4.9 から図 4.13 に、未知画像に対する検出結果例を図 4.14 から図 4.18 にそれぞれ示す。図中にて赤色でプロットされた領域は瞳の検出結果を、青色にてプロットされた領域は顔の検出結果を、緑色にてプロットされた領域はキャラクターの検出結果をそれぞれ示す。括弧の中の A 及び A2 と B は 4.4.2 節にて示したキャラクター A, B に対応する。また、上記画像例に対するキャラクター A, B の正解位置を図 4.8 に示す。ここで、赤枠で示された箇所がキャラクター A、青枠がキャラクター B の位置を表す。

4.4.3 考察

キャラクター識別器に関する実験について考察を述べる。識別結果より、Known に対する識別結果は高く、学習自体は十分に収束していると考えられる。その上で、Unknown に対する識別結果において、評価尺度である F 値が HOG の場合には 0.56 から 0.72 であり、未知の画像に対する識別も実施可能であると考えられる。また、Haar-like 特徴を利用した場合も、キャラクター A に対しての F 値が 0.68 であり、うまく特徴を捉えられる場合もあると考えられる。今回学習数による検出結果への影響は軽微であった。ただし、他の予備実験ではこの学習枚数内での変動が大きい場合には、検出結果に対して悪影響を与えることが確認された。学習用データセットの作成においては、普遍的なシーンから画像を集めることが有効であると考えられる。

瞳、キャラクター及び特定キャラクター検出実験についての考察を行う。表 4.4 より、未知画像に対する特定キャラクターの検出を行う場合には、HOG を用いた E-F-C のパターンで最も高い 0.93 の識別成功率が得られた。キャラクターの顔同士を学習した Character では、得られている識別器がキャラクター領域とそれ以外を見分けるにあたって非常に弱い。図 4.9, 4.14 の (3),(4) からも、ほぼ全ての領域で True の判定を行なっていることが確認される。2 種以上の識別器を用いた場合、E-F は顔をそのまま検出した結果である Face よりも良い結果が得られた。一方で、E 若しくは F と組み合わせた場合の特定キャラクターの Precision は 0.24 から 0.33 と E-F-C と比

表 4.3 既知画像に対する検出結果
自発表 [75] から引用

Method	Correct	Incorrect	Total	Precision
Eye	109	406	515	0.211
Face	536	551	1087	0.493
Character(A)	6935	40574	47509	0.146
Character(B)	1104	13335	14439	0.076
E-F	34	5	39	0.872
E-C(A)	149	682	831	0.179
E-C(B)	64	86	150	0.427
F-C(A)	222	498	720	0.308
F-C(B)	117	204	321	0.364
E-F-C(A)	4	9	13	0.307
E-F-C(B)	23	1	24	0.958
HaE	43	280	323	0.133
HaF	89	203	292	0.305
HaF-C(A)	111	475	586	0.189
HaF-C(B)	10	11	21	0.476
HaF-HaC(A)	29	156	185	0.157
HaF-HaC(B)	22	163	185	0.119
E-F-HaC(A)	4	35	39	0.103
E-F-HaC(B)	27	11	38	0.711

べ低い値である。E-F-Cの結果はいずれの場合においても他の組み合わせに比べて良好であり、これは、E-Fの結果を利用することで、より高い精度で得られた顔候補領域に対する特定キャラクターの識別処理が実施できたためと考えられる。本実験結果から、検出処理をカスケード化することは有効であるといえる。

HOGを利用した場合と Haar-likeを利用した場合の比較では、既知画像に対する一部と、未知画像に対して直接キャラクターの顔を得ようとする場合にのみ Haar-like がやや良好な結果である。しかし、実験結果より今回対象としている瞳、キャラクター、特定キャラクターの検出に対しては、HOGを用いた場合のほうが有効であることが確認された。

未知画像と既知画像に関しては、手動で切り出した顔画像を用いた識別実験においては顕著な差が見られた。実際の画像に対する検出を伴う実験においては、その差はほぼ無いと判断される。これは、スライドウィンドウによる走査が処理の根本にあるため、この移動幅により、実際に判別対象となっている画像と、学習に使用した画像が一致しないことが原因として挙げられる。

画像上の評価としては、図 4.11 の (10,11) にて、それぞれ検出対象のキャラクター A,B を取得できており、一方で検出対象外のキャラクターは検出結果から除外されていることが確認される。未知画像に対する実験結果において、キャラクター A を対象とした検出処理では、図 4.16 の (11) では検出対象外のキャラクターを検出結果から除外することに成功している。

表 4.4 未知画像に対する検出結果
自発表 [75] から引用

Method	Correct	Incorrect	Total	Precision
Eye	157	690	847	0.185
Face	935	928	1863	0.502
Character(A)	12560	66778	79338	0.158
Character(B)	1368	21707	23075	0.059
E-F	44	23	67	0.657
E-C(A)	365	1154	1519	0.240
E-C(B)	42	136	178	0.236
F-C(A)	420	851	1271	0.330
F-C(B)	151	366	517	0.292
E-F-C(A)	22	25	47	0.468
E-F-C(B)	15	1	16	0.938
HaE	70	482	552	0.127
HaF	152	349	501	0.303
HaF-C(A)	239	747	986	0.242
HaF-C(B)	15	23	38	0.395
HaF-HaC(A)	59	216	275	0.215
HaF-HaC(B)	33	242	275	0.120
E-F-HaC(A)	24	43	67	0.358
E-F-HaC(B)	19	44	63	0.302

4.5 キャラクターの顔の向き及びデフォルメを考慮した識別処理

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、キャラクターの顔の表現が様々に変化する。作中のシーンでは、セリフ等との配置の兼ね合いから、顔に対してオクルージョンが発生する場合もある。顔が正面向きと横向きの場合とで大きな形状の違いがあり、加えて、手描きの絵による表現であるため、方向の遷移における中間状態の取得ができない。しかし各キャラクターは、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。ここでは、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行う。

4.5.1 マンガ画像におけるキャラクターの特徴

マンガ内に登場するキャラクターの画像例を図 4.19 に示す。また、マンガにおけるキャラクター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

表 4.5 学習と実験に用いた画像数
自発表 [80] から引用

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

(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、たとえ同一キャラクターの顔であっても、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、同一キャラクター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。

4.5.2 判別処理アルゴリズム

上記特徴を総合的に考慮し、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。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 4.20 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る。Deformed では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一方で登場回数は少ない。

4.5.3 キャラクター識別実験

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。

キャラクター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 としてそれ以外のキャラクターを用いて学習を行った。学習及び実験に用いた画像枚数を 4.5 に、結果を 4.6 に示す。All は顔画像すべてのパターンを含む。FrontOnly は学習用及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。

表 4.6 識別成功率
自発表 [80] から引用

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 4.7 学習と実験に用いた画像数
自発表 [80] から引用

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

表 4.8 各パターンにおける識別成功率
自発表 [80] から引用

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

キャラクターの顔の向き及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。学習及び実験に用いた画像枚数を 4.7 に、結果を 4.8 に示す。実験では、本人の当該パターンのみを Positive、他人の顔及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。Front 以外の各パターンの学習画像数が少ない理由として、マンガにおいてその登場回数が少ないことが考えられる。

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 の場合には本人及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、顔のバ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バ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。

4.6 むすび

本章ではマンガ上のメタデータとして、マンガのキャラクターの情報を得るための、マンガ画像からの自動キャラクター検出及びその識別手法について検討を行った。一般にマンガのキャラクターには、表現によるキャラクターの表情変化、オクルージョン、描画サイズによる表現の変化等があるため、顔をそのまま学習、検出することは困難である。また、顔の識別に際して、学習に利用する画像の選別方法によりその結果が大きく変動するものである。

マンガ画像上のキャラクター検出という課題に対する具体的な処理として、一般物体認識のアプローチを利用した。画像特徴量と機械学習によりマンガのキャラクターを学習し、その検出とキャラクター間の識別を行った。マンガはその基本構造が線により描かれるため、従来一般的に人間の顔検出で利用されてきた Haar-like 特徴量を用いた識別処理は適さない。そこで、人体のエッジを捉えることで輪郭の識別に有効性を持つ HOG 特徴量を利用し、マンガのキャラクター検出と、その識別処理への利用について検討を行った。

実験により、HOG 特徴量と SVM を組み合わせた識別器を用いることで、マンガのキャラクター検出及びキャラクター間の識別に対しての一定の識別能力があることを確認した。また、キャラクター検出において、瞳を学習した識別器、顔を学習した識別器をカスケード型に接続することで、単一の要素のみを学習した場合よりも検出適合率が高くなることが確認された。また、上記に結果に対しキャラクターを学習した識別器による識別処理を加えることで、キャラクターごとの検出処理について最大 0.93 の適合率が得られた。

マンガでは実際の人間と異なり、シーンにより非常に大きな形状の変化を持って描かれる。また、作者の特性により、キャラクターの向きに応じてある点を境に輪郭形状が突然変化する場合がある。このため、キャラクターの状況ごとに別クラスを構成することで、識別性能が向上する。そこで、キャラクターの状態を正面、横、背面、デフォルメの 4 系統に分離することで、クラス内識別性能の向上を確認した。

4.6 むすび

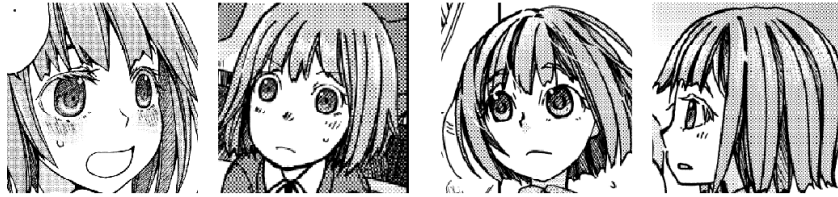


図 4.3 キャラクター A の正例画像
自発表 [75] から引用

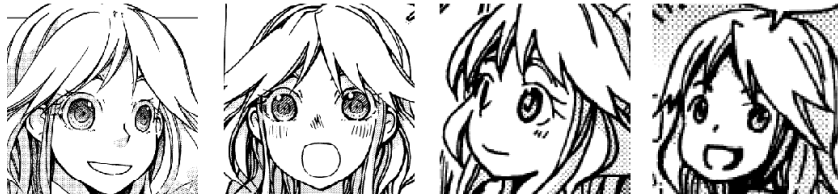


図 4.4 キャラクター B の正例画像
自発表 [75] から引用



図 4.5 瞳の正例画像
自発表 [75] から引用



図 4.6 顔の学習用正例画像
自発表 [75] から引用



図 4.7 瞳及び顔の学習における負例画像
自発表 [75] から引用

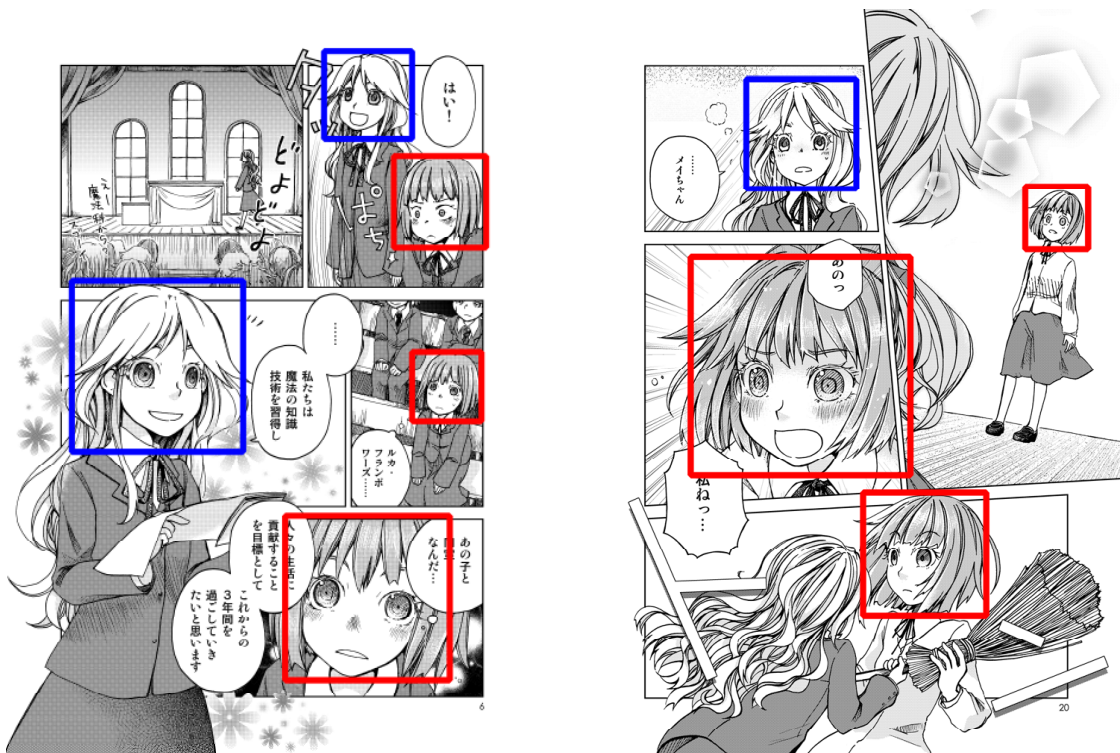
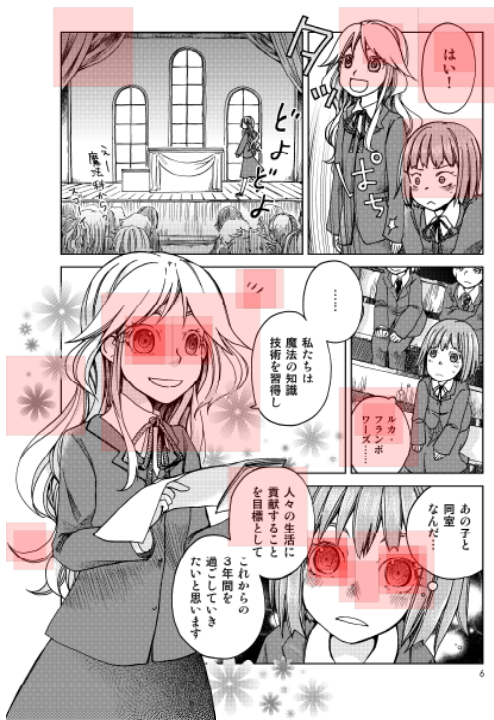
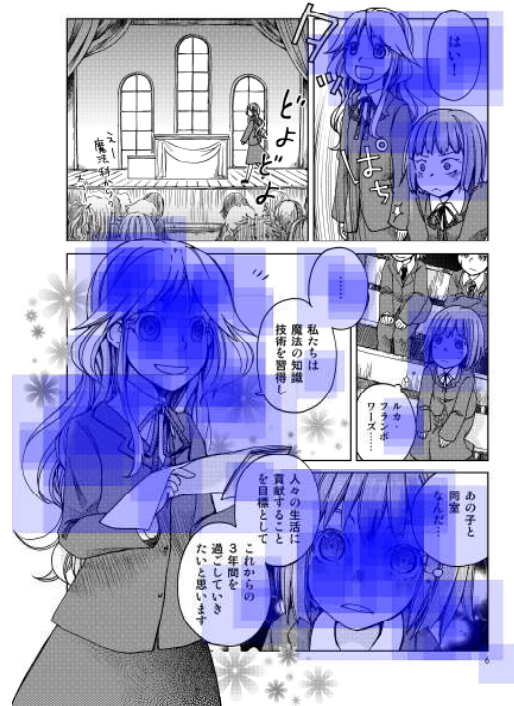


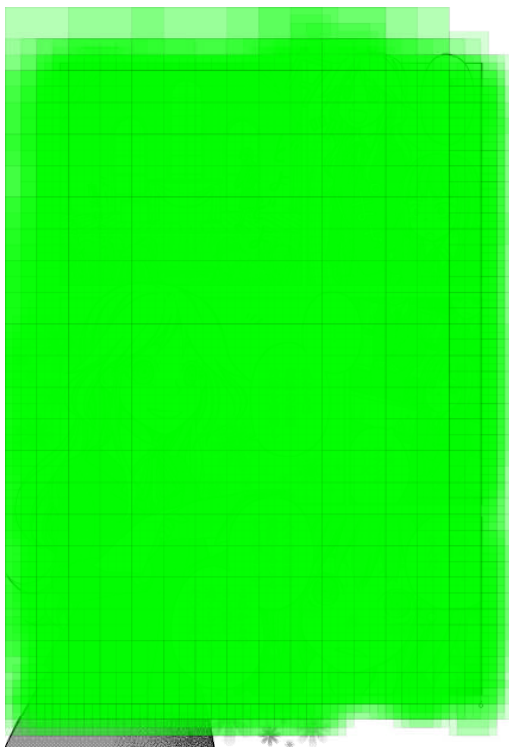
図 4.8 キャラクター A,B の正解領域例
自発表 [75] から引用



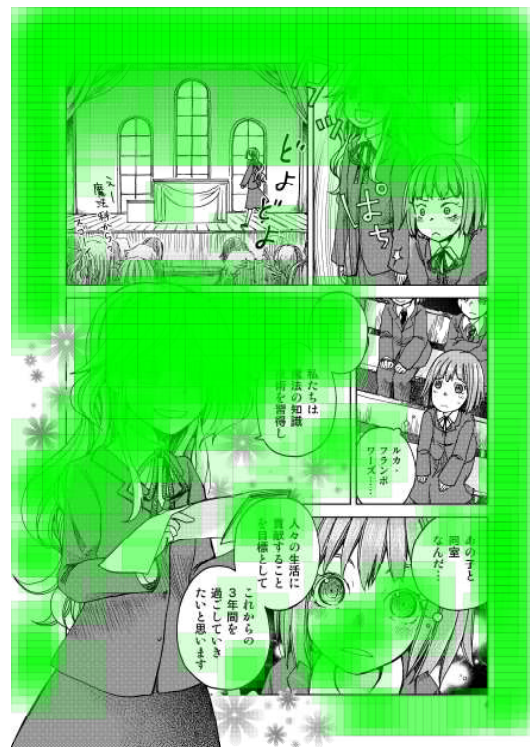
(1)Eye



(2)Face



(3)Character(A)



(4)Character(B)

図 4.9 既知画像上の瞳，顔検出及び顔識別結果 (1)

自発表 [75] から引用



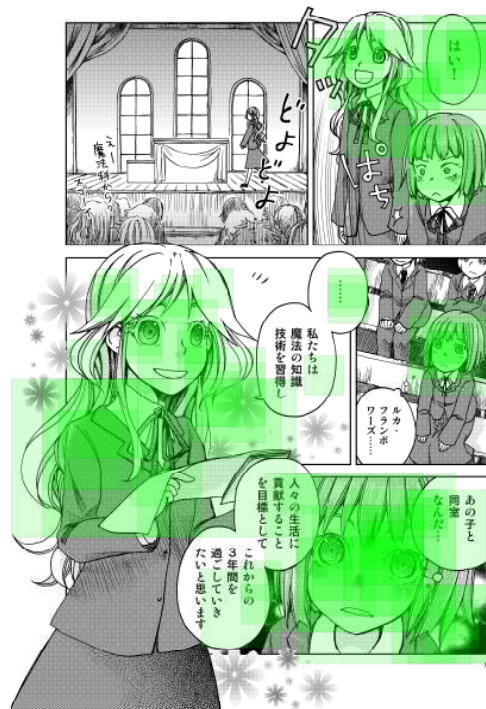
(5)E-F



(6)E-C(A)



(7)E-C(B)



(8)F-C(A)

図 4.10 既知画像上の瞳、顔検出及び顔識別結果 (2)
自発表 [75] から引用



(9)F-C(B)



(10)E-F-C(A)



(11)E-F-C(B)



(12)HaE

図 4.11 既知画像上の瞳，顔検出及び顔識別結果 (3)
自発表 [75] から引用



(13)HaF



(14)HaF-C(A)



(15)HaF-C(B)

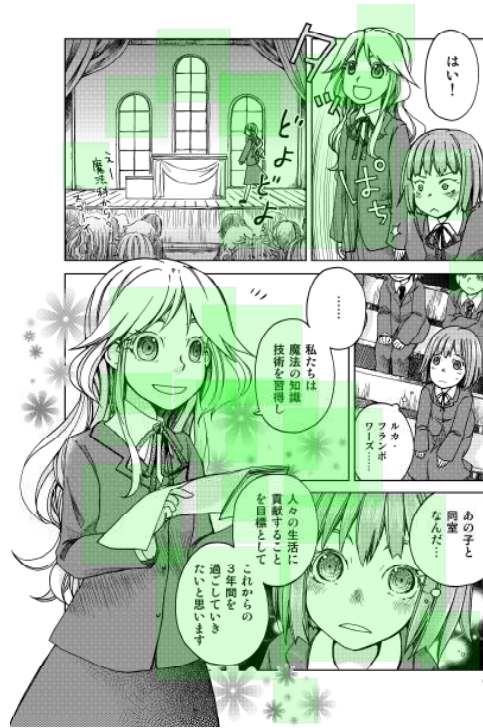


(16)HaF-HaC(A)

図 4.12 既知画像上の瞳，顔検出及び顔識別結果 (4)
自発表 [75] から引用



(17)HaF-HaC(B)

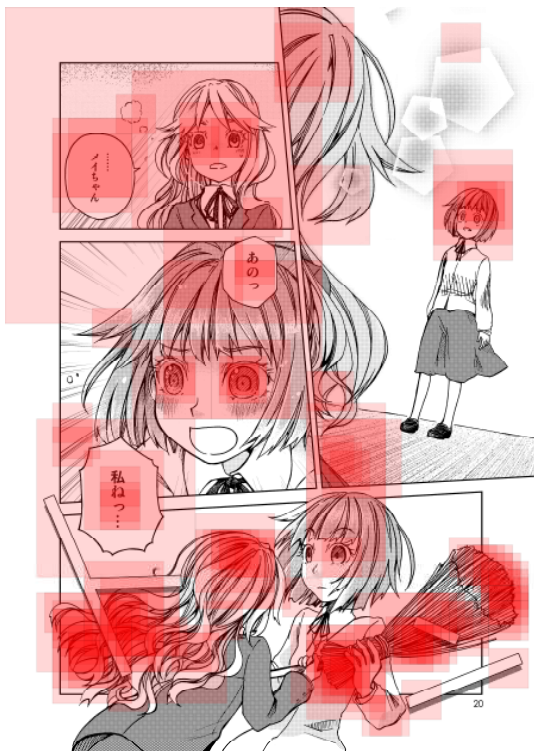


(18)E-F-HaC(A)

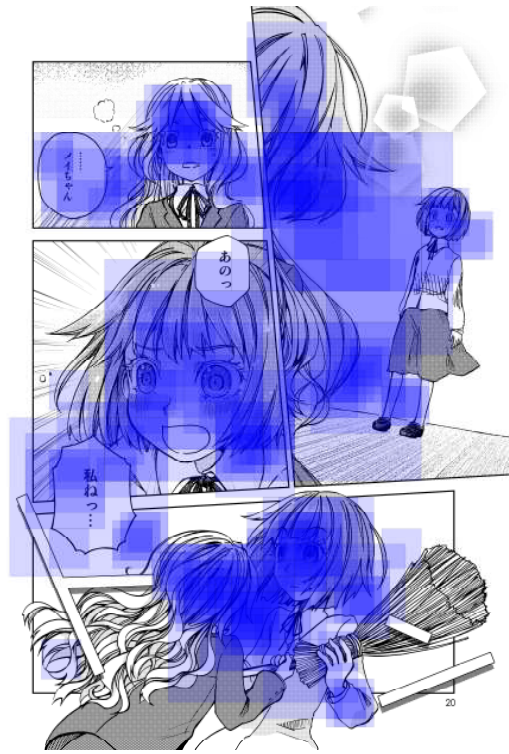


(19)E-F-HaC(B)

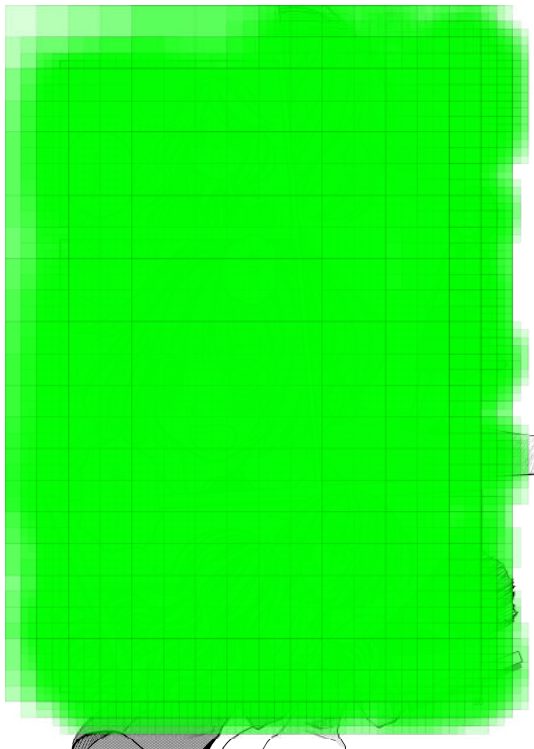
図 4.13 既知画像上の瞳、顔検出及び顔識別結果 (5)
自発表 [75] から引用



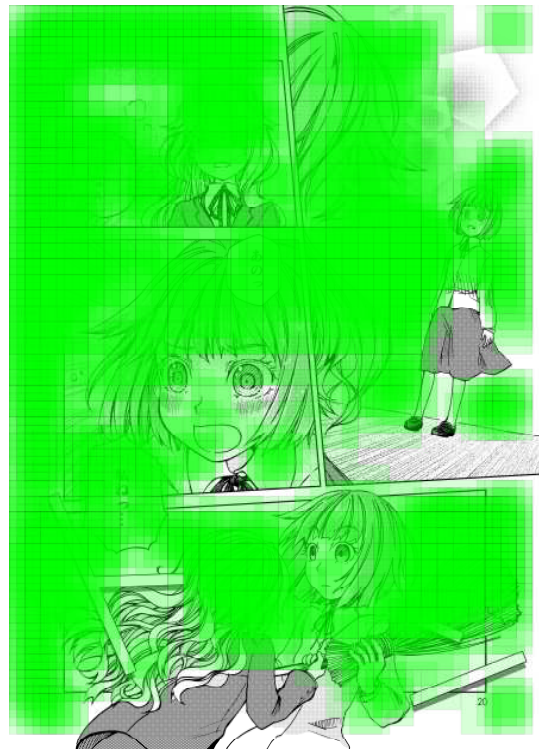
(1)Eye



(2)Face



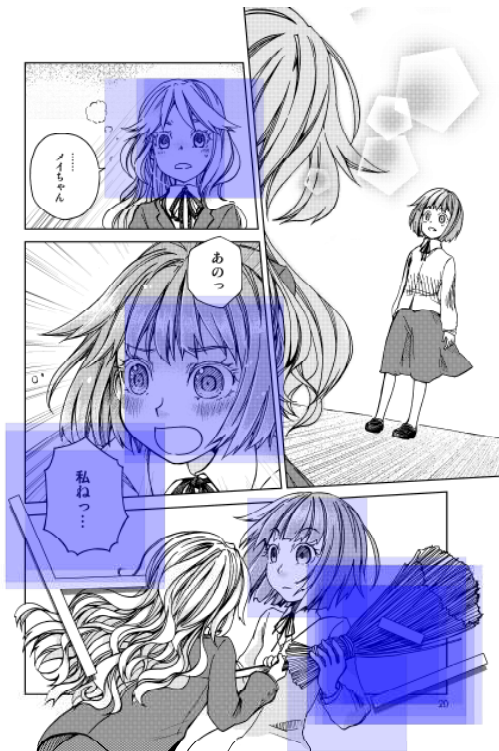
(3)Character(A)



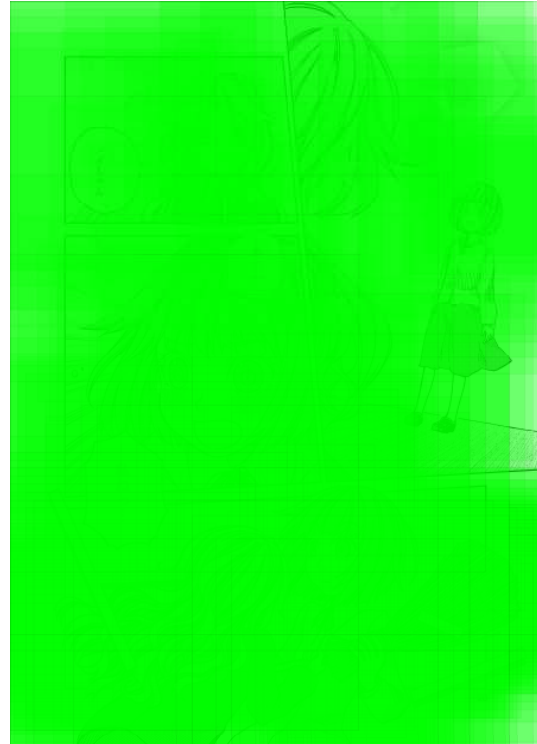
(4)Character(B)

図 4.14 未知画像上の瞳, 顔検出及び顔識別結果 (1)
自発表 [75] から引用

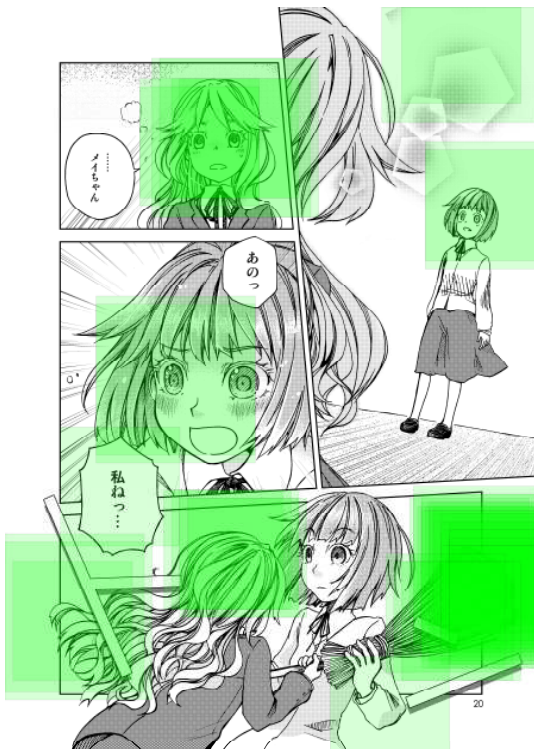
4.6 むすび



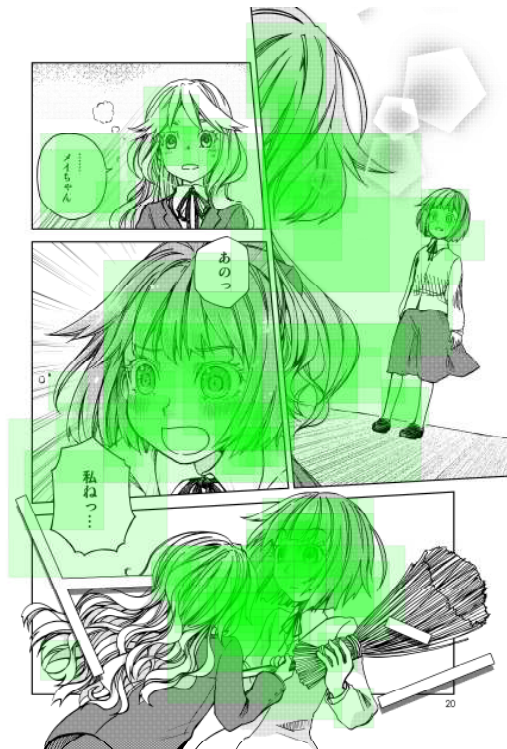
(5)E-F



(6)E-C(A)



(7)E-C(B)



(8)F-C(A)

図 4.15 未知画像上の瞳，顔検出及び顔識別結果 (2)
自発表 [75] から引用

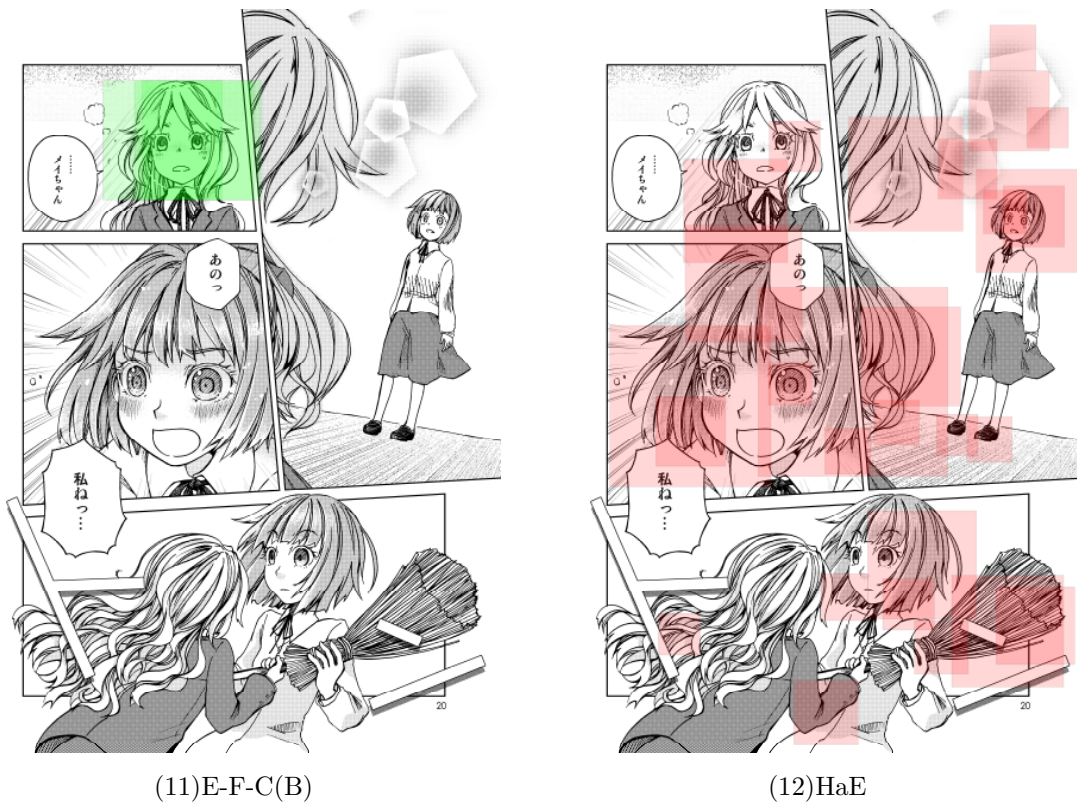
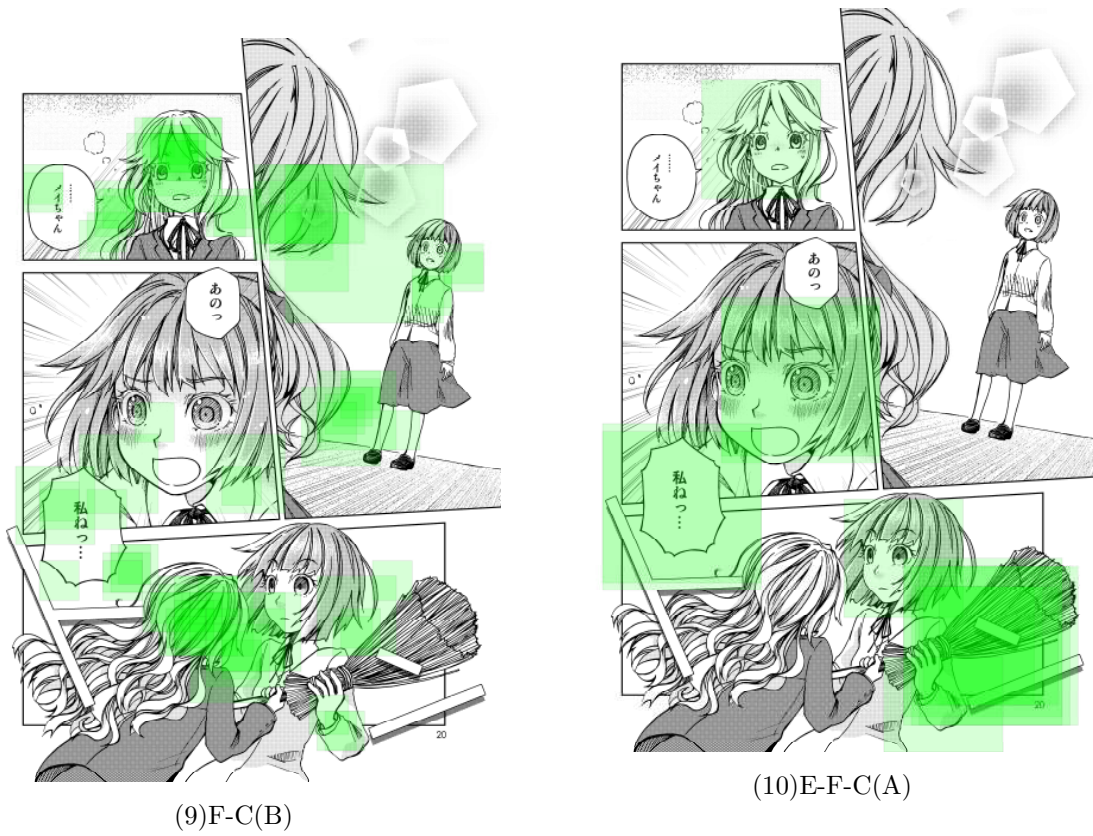
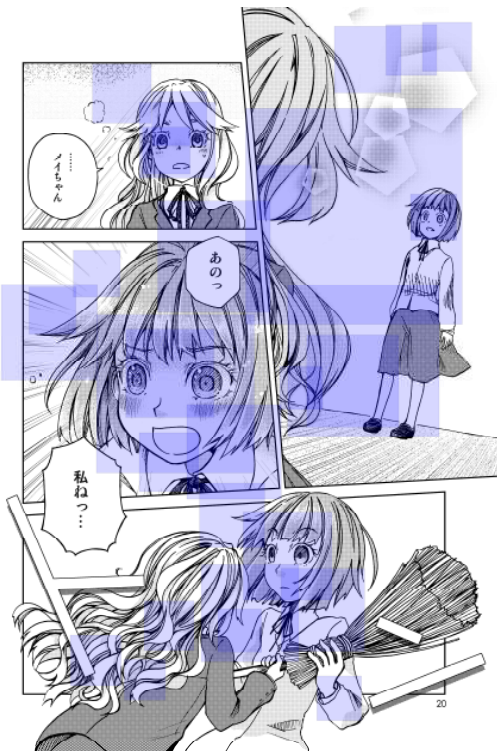
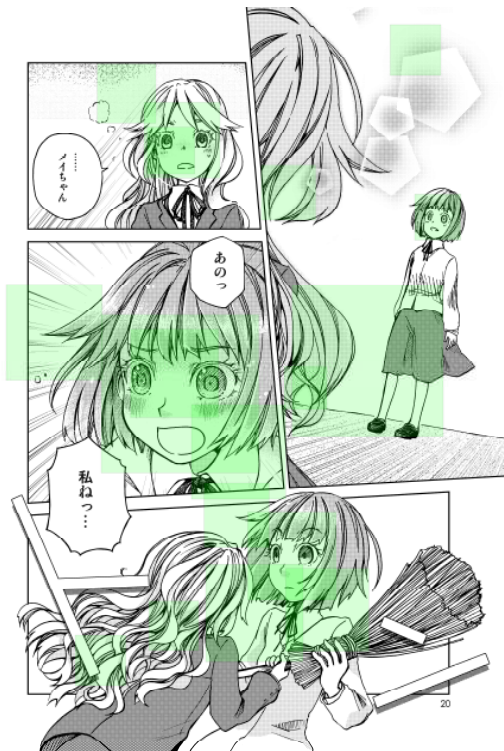


図 4.16 未知画像上の瞳，顔検出及び顔識別結果 (3)
自発表 [75] から引用

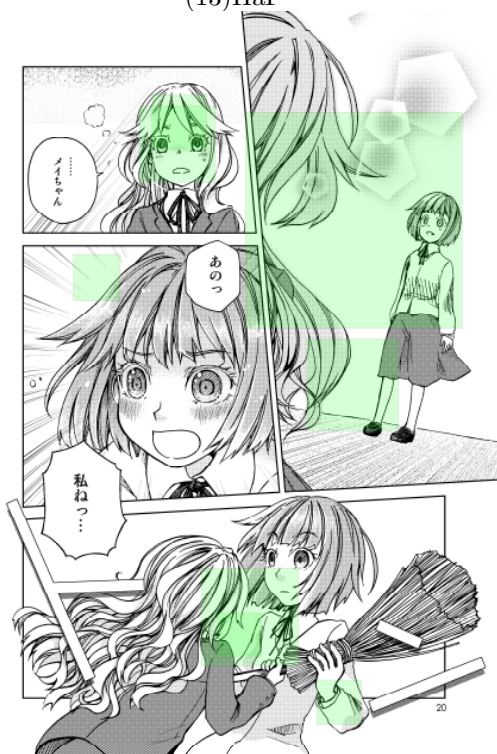
4.6 むすび



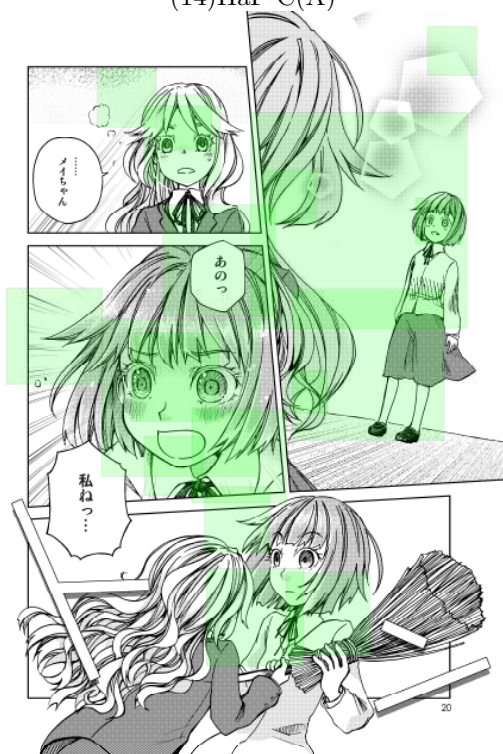
(13)HaF



(14)HaF-C(A)

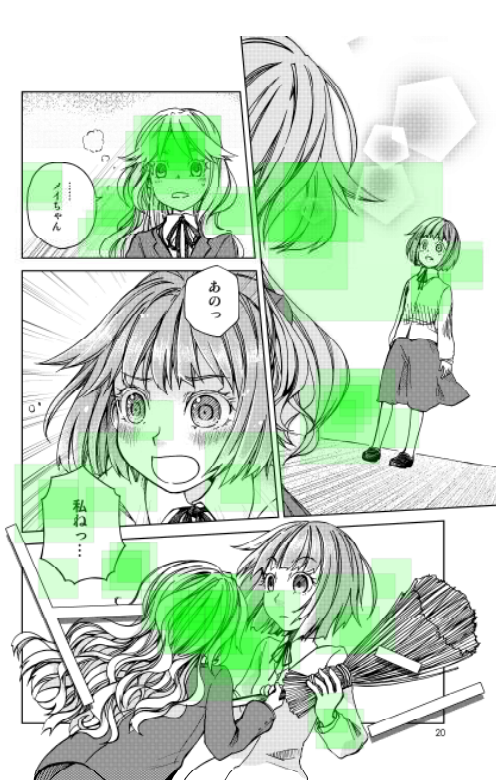


(15)HaF-C(B)

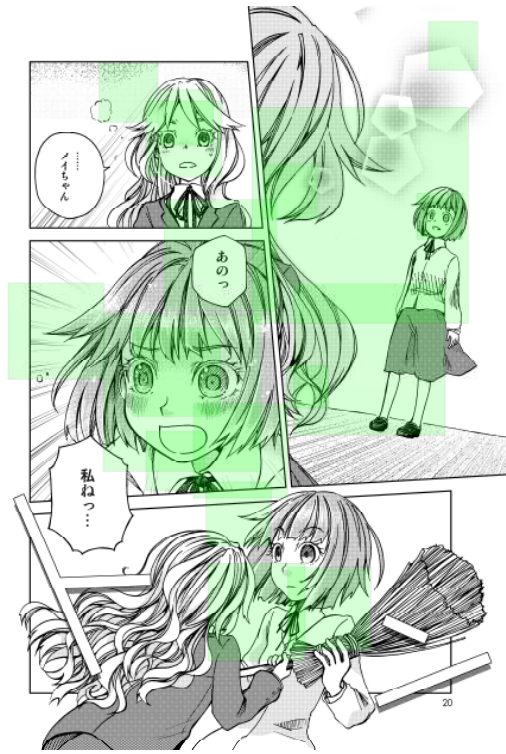


(16)HaF-HaC(A)

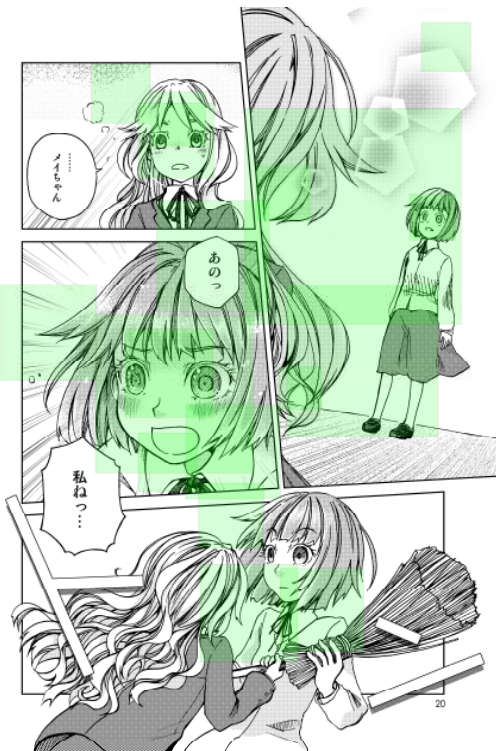
図 4.17 未知画像上の瞳，顔検出及び顔識別結果 (4)
自発表 [75] から引用



(17)HaF-HaC(B)



(18)E-F-HaC(A)



(19)E-F-HaC(B)

図 4.18 未知画像上の瞳，顔検出及び顔識別結果 (5)
自発表 [75] から引用

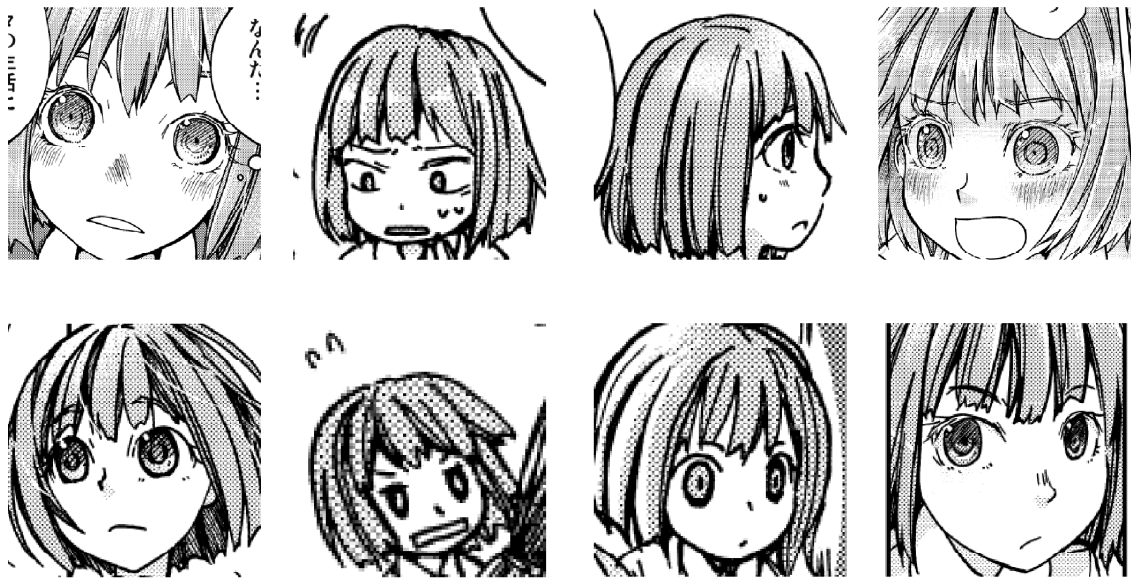


図 4.19 キャラクターの顔の変動
自発表 [80] から引用

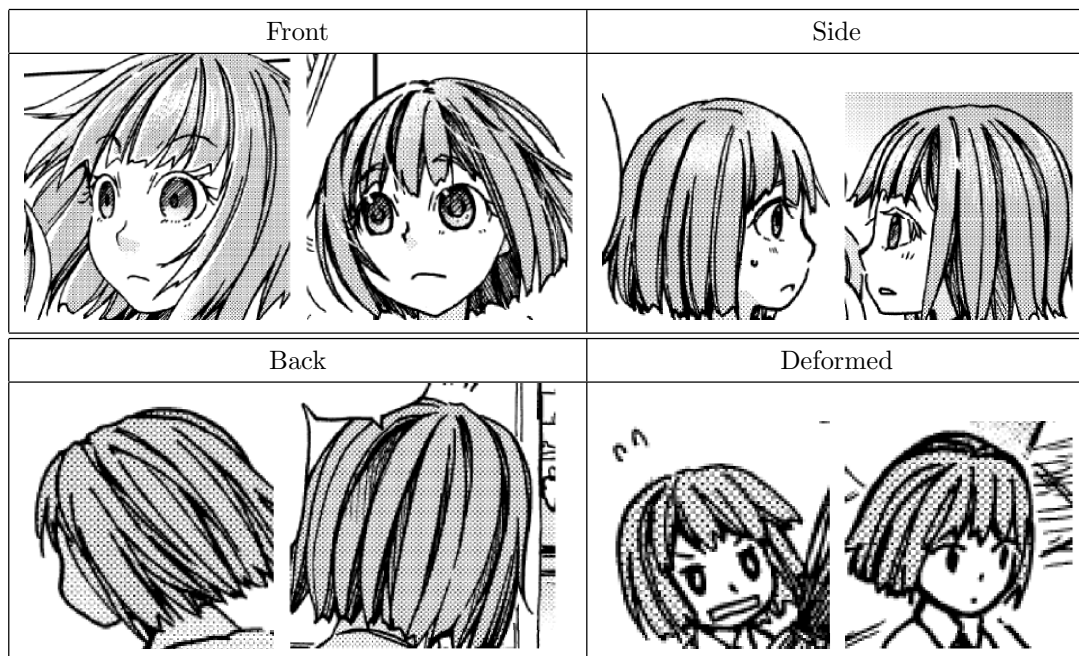


図 4.20 キャラクターの顔パターン分類
自発表 [80] から引用

第 5 章

2 値線画像向け画像特徴量 HBOP とマンガへの適用

5.1 まえがき

画像の表現方法の一つとして線画像がある。本論文で解析対象としているマンガはこの線画像の一形態である。他の線画像例として、画像からのエッジ検出処理結果、文字、イラスト等がある。エッジ検出処理は、特徴抽出処理の一つとして、古くから画像解析分野において利用されている。複雑な構造を持つ線画像によって表現された内容を理解することは、一般的に困難である。従来、画像解析の分野においては、細線化 [81–83] により得られる 2 値細線画像は光学文字認識 [84, 85] 等に用いられてきた。例えば、文字認識を行う際に、フォントの種類や文字を形成する線の幅により、同一の文字でも、異なる形状として認識されてしまう。そこで、2 値画像を細線化することで、構造を捉えやすくし、結果として文字の識別を実現している。このように、線を用いて描かれる形状では、それが細線化されていない画像であっても、細線化処理により 2 値線画像に変換することで、形状の解析が実現可能となる場合がある。

近年検討されてきた物体検出及び物体認識処理は、主にグレースケール若しくはカラー画像といった多値画像の特徴をとらえる様に設計されている。例えばカメラ等で顔検出を行うためにしばしば利用される Haar-like 特徴量 [39, 40] は対象領域内の平均輝度値を利用する。また、監視カメラ等での人物検出に利用される Histograms of Oriented Gradient (HOG) 特徴量 [41] は、縦方向の輝度差分と横方向の輝度差分の比から算出した、方向成分のヒストグラムを利用する。これらの特徴量は輝度レベルの差を用いた画像特徴量であり、輝度レベルが 2 段階しか存在しない 2 値画像、あるいは細線化された 2 値画像である、2 値細線画像では特徴をとらえることが困難である。このため、2 値細線画像に特化した画像特徴量を導入することで、画像解析性能の向上を図る。

本章では、2 値画像を、各点ごとにオブジェクトの有無の情報のみを持つ画像であり、一般的に白黒の 2 値によって構成される全ての画像とする。また、2 値線画像を、2 値画像の構成要素のうちそれが主に線によって描かれているものとし、2 値線画像のうち、その線の太さが 1 画素のみで構成されるものを 2 値細線画像と表記する。上記の広い意味での 2 値画像から 2 値細線画像への変換は、細線化処理、モルフォロジフィルタリング、輪郭抽出、エッジ検出処理等を行うことで可能となる。また、2 値画像に限らないカラー画像、濃淡画像等からも同様にエッジ検出及び

5.2 従来の画像特徴量と 2 値線画像に対する問題点

細線化処理により 2 値細線画像を取得できる。2 値化処理としては大津の 2 値化 [86, 87] をはじめとする各種 2 値化処理が提案されている。これだけでなく、画像からのエッジ検出から細線化までを通して行う Canny フィルタ [88] もある。

ここでは、マンガ画像解析を主目的として 2 値細線画像解析性能の向上を試みる。マンガ画像は多数の線により描かれる白黒の画像である。線によって描かれる形状は、フォントの文字認識と同様に、細線化処理により 2 値線画像とすることで、解析処理に有効であると考えられる。特に、本論文の解析対象であるマンガでは、コマの枠、キャラクター、文字等、ほぼすべての構成要素が線や塗りつぶしによって描かれており、その線数は多く複雑である。マンガに描かれている内容を解析するためには、構造や線の形状から情報を得る必要があり、これは、処理の面からは、フォントの解析と同様の性質をもつものであると考えられる。人物認識や人物認証を、マンガ画像でも実施可能とするためには、線画像の構造を認識することが有効であると考えられる。また、対象カテゴリ内の変動を許容し、形状変化に対応する特徴記述が必要である。

一般的に濃淡画像における線は太さを持つ。濃淡画像を 2 値化した場合に得られる太さを持つ線に対し、処理目的に応じて 2 値細線画像を取得する方法を選択する。例えば線の太さも含めた線の形状自体が重要である場合にはエッジ検出処理後に細線化を行うことで、その線の輪郭形状を 2 値細線画像を取得可能である。一方で、線の太さは関係なく、その線によって示される形状が重要である場合には、線そのものに対して細線化処理を行うか、線がベクターによって得られているものであれば、ベクターの情報から幅 1 画素の線を出力することで、2 値細線画像を取得可能である。

マンガは線としての特徴を持つ一方で、画像の複雑さという面では自然画像に近い。認識対象が文字の場合、細線化後の構造は小文字のアルファベットであれば 26 種類であり、また文字の形状は単純であるため、分岐点の数も限られる。従って、これまで、フォント解析等に代表される 2 値細線画像の解析は、ある程度単純化された構造解析といえる。一方で自然画像に近い特徴を持つ場合、分岐点数等は一定せず、細線化された形状のバリエーションは多い。このような対象を解析する為には、自然画像解析と同様に画像特徴量と、統計的機械学習等を用いたアプローチが有効であると考えられる。

そこで本章では、これまでの検討により、従来の画像特徴量の中で、2 値線画像に対して一定の識別性能を持つ HOG 特徴量を参考に、2 値細線画像の特徴を捉えることができ、一般的な機械学習手法と組み合わせることで識別器を生成可能な 2 値画像向け画像特徴量 Histograms of Binary Orientation Pattern (HBOP) について示し、実験により有効性を示す。

5.2 従来の画像特徴量と 2 値線画像に対する問題点

5.2.1 一般物体検出に用いられる画像特徴量の概要

第 4 章では、一般物体検出に用いられる画像特徴量として Haar-like, HOG について述べた。これに加え、HOG 特徴量の改良版である EHOG [43] や、エッジの出現位置と形状を組み合わせた Edgelet [44] などの画像特徴量がある。HOG 特徴量の派生版である EHOG 特徴量は、HOG 特徴量に対して方向成分の単純化及びセルの除去により特徴量の次元数を削減している。Edgelet

特徴量はエッジの出現位置とその形状から検出対象の特定を行う。Edgelet 特徴を用いることで、人物検出において、人間の両肩に起因するエッジ等を捉えることが可能である。また、画像内複数箇所の特徴量について、その共起性や、出現位置を組み合わせることで、より高精度な検出を行う手法も提案されている。Joint-HOG [46], CoHOG [45], Deformable Parts Model [47] では、検出領域内において、複数箇所の小領域に対し HOG 特徴量を用いた検出結果を組み合わせ、再度学習を行うことにより、従来よりも高い精度をもつ強識別器が構成される。Shapelet 特徴量は複数の小領域内におけるエッジ方向の共起性をとらえるものである。以下ではこれら一般物体検出向け画像特徴量の 2 値細線画像及び本論文の処理ターゲットであるマンガに対する適応性について述べる。

5.2.2 Haar-like 特徴量

Haar-like 特徴量は画像内局所領域の輝度の分布特徴を捉える画像特徴量である。Haar-like 特徴量の概要については第 4 章で示した。Haar-like 特徴量を用いた顔検出処理は高速かつ頑健であるため、非常に多く利用されている。その一方で、2 値細線画像を対象とした処理では、Haar-like 特徴量の計算方法から、局所領域内に存在する線の密度についてのみ検知可能である。しかし円形か四角形などの構造及び線の方向等については検知不可能である。したがって、2 値細線画像の特徴を十分に捉えることができない。

5.2.3 HOG 特徴量

HOG 特徴量は、方向成分の分布を捉えるため、現在一般に使用されている画像特徴量の中では線画像に対しても高い識別性能がある画像特徴量である。HOG 特徴量の概要については第 4 章で示した。HOG 特徴量では、セル単位でヒストグラムが作成されるため、若干の形状変化が許容される。一方で、2 値線画像に対しては、方向及び勾配強度の計算方法に問題がある。以下では HOG 特徴量の 2 値線画像に対する問題点を示す。

HOG 特徴量では先に示した通り、注目画素の上下左右の輝度差からエッジ強度とその方向が算出される。図 5.1 に HOG 特徴量計算時の参照画素を示す。図 5.1 において、丸で示された画素は縦方向、十字で示された画素は横方向の参照画素である。(A) - (D) に示される様に 45 度方向に線を構成する画素が存在する合や、垂直、水平に直線画素が並ぶ場合に、式 (1), (2) により勾配強度を求めると V_d , H_d 共に差分を検出できないため、線を構成する画素上では特徴量が得られない。また、(E), (F) の場合には左右方向の差分は存在するが、上下方向の差分を取得できないため、(E), (F) の線の状況を区別できず、更に直線自体の方向とは異なる 90 度としてエッジ方向を観測する。(G), (H) は (E), (F) とは逆であるが、角度の射影により同じく 90 度として角度が得られる。

直線上の画素では多くの場合上記のように特徴量が得られないため、基本的には図 5.2 に示す様に線と隣接する点において得られる値から間接的に直線を表現することとなる。ただし、この場合においても、観測点が 4 点でかつその値が 1 か 0 のみに限られる。HOG 特徴量では方向を、 $[0^\circ, 180^\circ)$ の区間に射影し、その中で任意の方向数の bin を生成する。よってここで得られる方向は 0 度 (C, D), 45 度 (A, B), 90 度 (G, H), 135 度 (E, F) の 4 種類のみとなる。また、2 線が

5.2 従来の画像特徴量と2値線画像に対する問題点

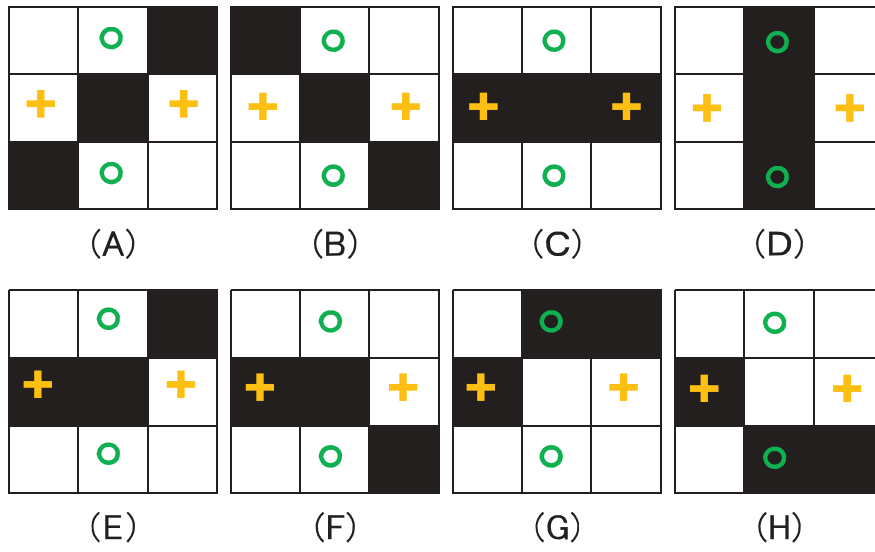


図 5.1 HOG 特徴量計算時における参照画素 (1)
自発表 [76] を基に一部変更を加えた

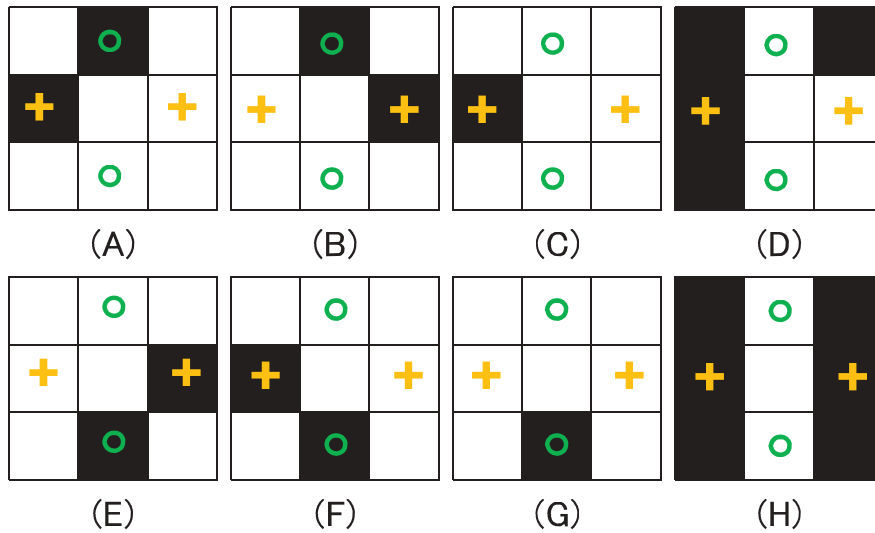


図 5.2 HOG 特徴量計算時における参照画素 (2)
自発表 [76] を基に一部変更を加えた

平行して存在する場合 (G, H) には参照点間の輝度差が 0 となるため、やはり方向の測定が不可能となる。

HOG の方向算出方法は通常の濃淡画像では有効な数値を得られる。しかし 2 値細線画像に対しては直線を構成する画素の上では方向成分を取得できない点、周囲の画素からの影響により本来出現しない方向成分が現れる点が問題となる。2 値細線化されたマンガ画像への適用という観点から評価した場合、マンガ画像では、多数の線により絵が描かれるため、人体検出と比較すると、多数の方向にエッジ成分が出現する。また、対象により描かれる線分のパターンが変化すると考えられるため、対象領域内に存在する線の方向分布に特徴が現れる。HOG では複数方向の分布を表現可能であるが、直線方向検出処理時に、隣接する線や孤立点からの影響を受ける。この影響を除外するためには、線を構成する画素上にて線の方向が記述可能である必要がある。

5.2.4 EHOg 特徴量

HOG 特徴量の派生である Extended Histograms of Oriented Gradient (EHOg) 特徴量 [43] では、主要方向成分エッジ量の相対表現と、ブロックにセルを定義しないことにより、特徴量の次元数が削減される。HOG 特徴量ではセルごとにヒストグラム化した方向成分を、ブロック単位で正規化し、これをすべて HOG の特徴量の各ベクトルに割り当てていた。EHOg では、以下の式により各ブロックにつき $F_b(D)$ を求め、これを特徴量とする。

$$F_b(D) = \sigma_{i \in D} g_b(i) / Z_b \quad (5.1)$$

Z_b は 9 方向全てのヒストグラム合計値である。また、D はブロック内で着目する主方向を表す。ここで、D は 9 方向に量子化された方向成分から、連続した 1 から 3 方向の組み合わせによる 27 パターンの中から選択される。このとき、D の含有する角度は 20° から 60° となる。また、ブロック内にセルを定義しないため、EHOg 特徴量の次元数は総ブロック数となる。したがって特定ブロック内に特定方向のエッジが統計的に得られる場合には、HOG 特徴量同等の検出精度を保ちつつ、大幅な次元数の削減が可能である。

EHOg の主要方向算出アルゴリズムは、対象領域内にシンプルなエッジが存在し、その角度が一定方向に集まるという前提が成り立つ場合に有効とされる。しかしマンガ画像から得られる 2 値細線画像では、出現する線の方向は多岐にわたり上記条件を満たさない。

5.2.5 Edgelet 特徴量

Edgelet [44] は line, arc として直線と弧形状、そしてこれを対象に配置した symmetric-pair などの形状をあらかじめ定義する。そして、定義された形状上の、重み付き評価値を基に学習を行う。検出時には、学習された形状に対して高い相関を示すエッジ成分を探し、その類似度を測る。このため、人物検出等において、人間の両肩や人間の頭部等に起因するエッジをその位置と共に得ることができる。この特徴から、Edgelet 特徴量は人物検出処理など、位置と共に特徴的なエッジが計測されやすい状況であれば良好な識別性能を持つ。Edgelet では、line, arc 等の定義された形状を構成する画素上において、それぞれ重み付き評価値を計算する際に、 3×3 画素の Sobel オペレータ [89] を使用する。評価値の計算では、 $[0^\circ, 180^\circ)$ の区間に射影された角度成分を 6 方

5.2 従来の画像特徴量と2値線画像に対する問題点

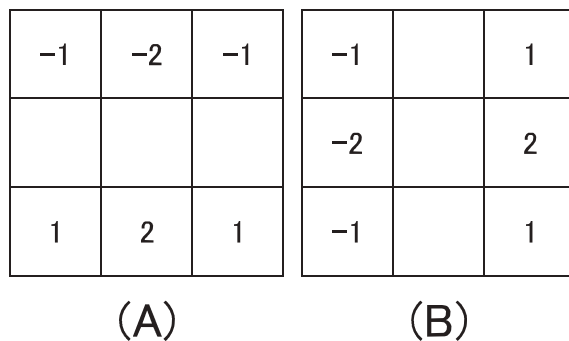


図 5.3 Sobel オペレータ [89]

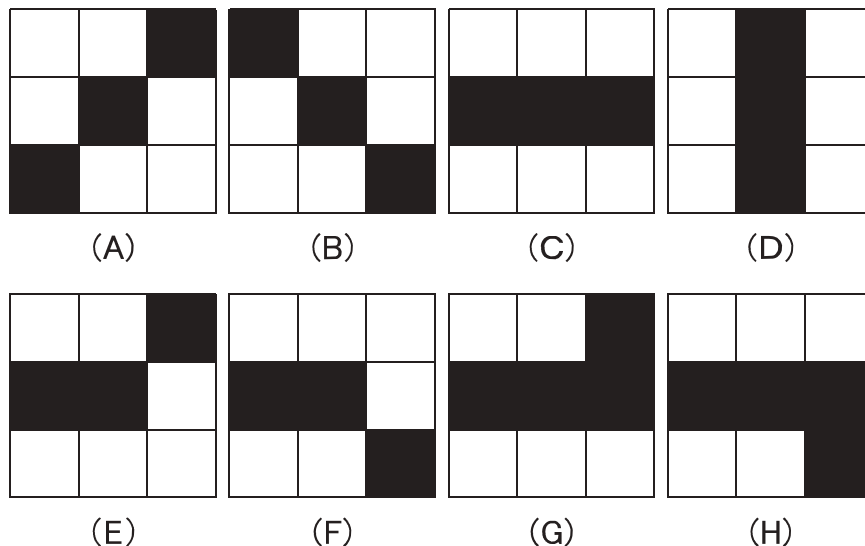


図 5.4 Sobel によるエッジの計算 (1)

向に量子化し、そのインデックス差に応じて重みづけを行う。

Sobel オペレータは2方向の線形フィルタからなり、周囲画素値に対して図4に示される重みを掛けて計算される。Sobel フィルタの参照画素と重みを図5.3に示す。HOGの場合と同様に、2値細線画像から取得される方向を検討する。図5.4(A)-(D)の配置となる場合には縦方向及び横方向の差分が0となり、直線方向が算出できない。直線の構成画素上では図5.4(E), (F)とこれを回転させた8パターンの場合にのみ角度が得られ、 $45^\circ, 135^\circ$ となる。一方、図5.4(G), (H)の場合には、得られる角度はそれぞれ $135^\circ, 45^\circ$ となる。(E),(F)と(G),(H)では、連結数により得られる角度が逆転してしまう。線に隣接する画素では基本的にSobelフィルタにより方向成分を得ることができる。図5.5において、(A)及び(E)では 45° 、(B)及び(F)では 135° 、(C)は 0° 、(G)は 90° となる。ただし、図5.5(D)の様に、線の近隣に他の画素が入る場合にはその影響によ

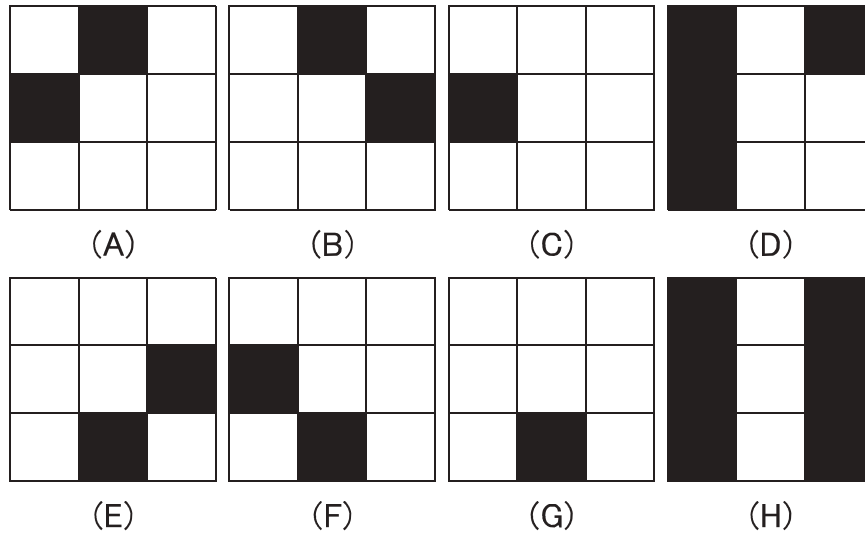


図 5.5 Sobel によるエッジの計算 (2)

り角度が変化する. この場合 (D) で得られる角度は 18° となる. 図 5.5(H) では HOG 特徴量計算時と同様, 2 直線により差分が打ち消されるため直線を検出できない.

このように, 2 値細線画像に対する Sobel フィルタによる角度の取得では, 直線を構成する画素上からは十分に安定した方向成分を得ることが出来ない. 一方で, 隣接画素においては, Sobel オペレータでは HOG 特徴量計算時と比較して参照画素数が増えるため, 線の近隣にノイズや他の直線が存在する場合など, 実際の線とは関係のない方向成分が算出されやすい. そのためマンガのように, 様々な方向の線分が多数存在する場合には, 高い相関を持つ Edgelet 特徴量が対象領域内に無数に観測されるため, 検出精度が著しく低下する.

5.2.6 複数局所領域の関連性を用いる手法

Joint-HOG [46], CoHOG [45], Deformable Parts Model [47] などに代表される共起特徴量では, 検出領域内において, 複数箇所の小領域に対して, 従来の特徴量による検出結果を取得し, これを学習することで, 単一の特徴量を用いるよりも高い精度をもつ強識別器を構成するものである. 上記 3 手法では, 先に示した HOG 特徴量による弱識別機を利用している. Shapelet は複数の小領域内におけるエッジ方向の共起性に着目した特徴量である. Shapelet ではその低レベル特徴として, $[-1,0,-1]$ のカーネルを用いて $0^\circ, 45^\circ, 90^\circ, 135^\circ$ の 4 方向について差分値を取得する. 一般的に用いられる HOG 特徴量と比較して, 方向の分割数が少ないため人物検出精度は低い. また, 2 値細線画像においても方向取得性が低い.

線画像はエッジ抽出処理後の画像とほぼ同等のものである. 一般的に使用されている画像特徴量の方向成分取得アルゴリズムは, 線そのものの方向を記述する為に設計されていない. 特に 2 値細線画像からの特徴量計算では, 内部で利用する低レベル特徴量の出力に依存するものであり,

5.3 HBOP 特徴量

これまでに述べた 2 値画像及びマンガ画像に対する問題点が基本的にそのまま残る。一方で、共起性を用いるという戦略は利用する特徴量に左右されない。したがって、共起特徴量の手法において、その内部で利用する特徴量を、2 値細線画像に有効な手法に置き換えることで、全体としてより高精度な強識別器の構成が可能であると考えられる。

5.3 HBOP 特徴量

5.3.1 HBOP 特徴量の概要

2 値線画像解析処理の高精度化を実現するため、2 値細線画像に特化した特徴量である HBOP について述べる。2 値細線化されたマンガ画像では、一般的な自然画像と比較して多数の線及びテクスチャが存在する。したがって、近隣画素からの影響をなくすため、直線を構成する画素上で、直線方向を記述できることが望ましい。HBOP 特徴量では HOG 特徴量を基に、2 値細線を構成する画素上で直線方向の記述が可能となるようにアルゴリズムを設計した。本節では、2 値細線画像用の画像特徴量である、HBOP 特徴量 [76] の詳細を示す。

従来の画像特徴量の中で、HOG 特徴量は直線近傍画素を用いることで間接的に線分方向を記述できるため、4 に示されるように、その基本成分が線により構成されるマンガ画像の解析においても一定の解析性能を持つ。ただし、HOG 特徴量において得られる方向成分のヒストグラムは線そのものの方向分布を表現するものではない。

提案する HBOP は 2 値の線分の取り扱いに特化したものであり、注目点の 8 近傍に存在する画素のパターンを集約し、領域ごとにその出現量をヒストグラム化する。ヒストグラム化するパターンを線分に出現するものに限ることで、2 値線分を取り扱うために冗長な情報を削減できる。

5.3.2 HBOP 特徴量の計算手順

HBOP 特徴量の計算手順について述べる。HBOP 特徴量の計算手順を図 5.6 に示す。HBOP 特徴量では入力画像を任意数のセルに分割し、セルごとに出現する方向パターンのヒストグラムを計算する。ここで、入力画像については 2 値の線画像とする。HBOP 特徴量では、HOG 特徴量同様に、画像を任意数のセルに分割し、それぞれのセル内にて方向パターンのヒストグラムを作成する。線画素は白黒双方の場合が考えられるが、ここでは、白を背景、黒を線を構成する前景として説明する。学習時には入力画像全体から一つの HBOP 特徴量の計算を行う。処理は図 5.6 に示される 4 ステップに分かれる。画像内のラスタスキャンを行い、そこで得られる 8 近傍画素パターンにより方向パターンへの投票を行う。次に、各セルにおいて出現する方向パターンのヒストグラムを計算する。そして正規化を行うことで HBOP 特徴量を得る。各ステップの詳細を以下に述べる。

ステップ 1 では、画像内に出現する 8 近傍画素パターンを計測する。ここで利用する 8 近傍画素パターンと方向パターンを図 5.7 に示す。このパターンは、大津らによって提唱され、その後米山らによって拡張された局所自己相関パターン [90–92] の一部である。まず、HBOP 特徴量では注目する画素自身が 2 値の線上である場合のみヒストグラムへの追加処理を行う。注目点自身が線上である場合の 8 近傍画素について、HBOP ではこの 12 種類のみを扱うこととする。実際の

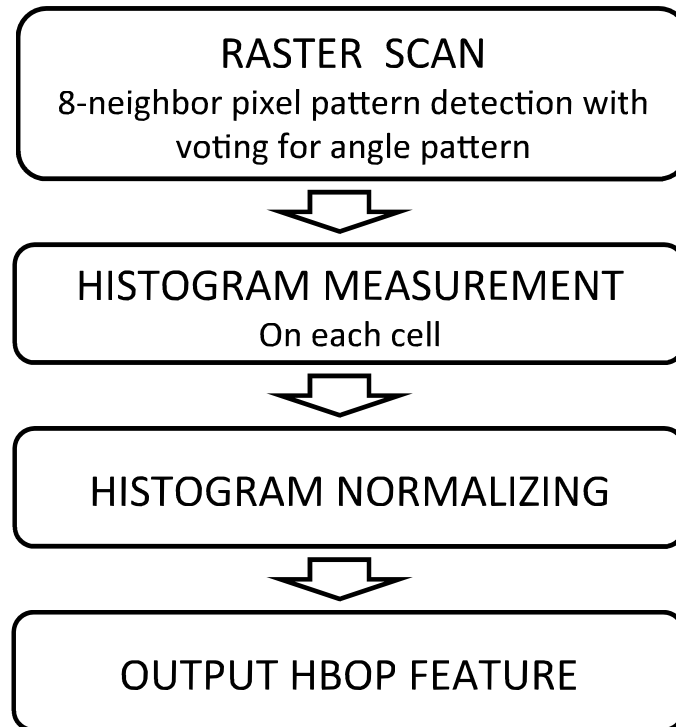


図 5.6 HBOP 特徴量の計算手順
(自発表 [76] から引用)

処理では、図 6 に黒色で示される 3 画素について線上か否かの判定を行う。3 画素とも線上であればそのパターンの属する方向カテゴリに対して投票を行う。図 6 に示された Angle は投票される方向カテゴリを表す。したがってステップ 1 が終了した段階で画像内に出現する各方向パターンの出現数が取得される。

ステップ 2 では、セル内ヒストグラムの計算を行う。HBOP 特徴量では画像全体を任意数のセルに分割し、それぞれのセルにて各方向パターンの合計出現回数をヒストグラム化する。

ステップ 3 では、得られたセル内の全方向パターンの合計により、ステップ 2 で得られたヒストグラムを割ることで正規化する。HOG 特徴量では照度変化に対応するため、複数セルから構成されるブロックによる正規化が行われていた。2 値線画像では、画像内に現れる線はそのままセル内の全方向パターンの合計値として得られ、かつすべての点の強度は一定であるため、ブロック単位での正規化の必要性はない。

ステップ 4 では、ステップ 3 までで得られた特徴量をブロックごとに各方向成分を順番に並べたものを HBOP 特徴量とし、これを出力する。ここで出力される HBOP 特徴量の次元数は次の式により求められる。

$$HBOPDimension = CellNumber \times AnglePatternNumber \quad (5.2)$$

セル数が 5×5 、方向パターン数が 8 の場合、HBOP 特徴量の次元数は 200 となる。

一般的に線の終端や角部分などでは図 5.7 に示す HBOP 特徴量で扱わないパターンが出現する。出現例を図 5.8 及び 5.9 に示す。図 5.8(A) 及び図 5.9(A) に示されるパターンは、図 5.7 には出現しない。一方でこれらのパターンは直線上に多数出現することが考えられる。ただし、

5.3 HBOP 特徴量

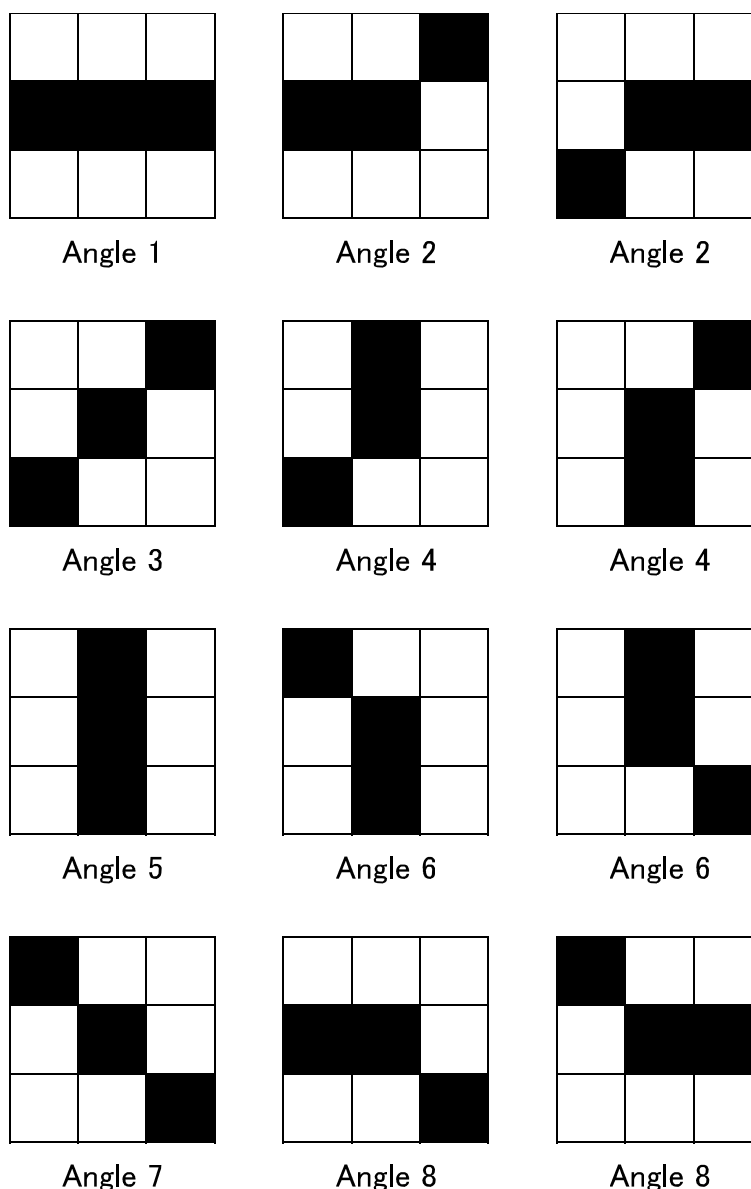


図 5.7 HBOP 特徴における近傍画素のパターン
(自発表 [76] から引用)

HBOP 特徴量計算時には、線を構成する隣接画素において得られる HBOP パターンにより、間接的に線形状を表記できる。具体的には、図 5.8 では (B), (C), 図 5.9 では (B)-(E) に示されるパターンにより、線の形状が表現される。図中の三角などの形状はそれぞれのパターンにおいて参照する画素を示す。また、このような重ね合わせによる記述を行うことで、4 連結、8 連結ともに共通のパターンで記述可能である。HBOP では、HOG 特徴量等における直線の間接表現と比較して、直線を構成する画素上にてパターンを取得するため、近隣に直線がある場合でもその影響を受けない。HBOP で用いるパターン数を増やし、全ての 8 近傍パターンを導入することも可能であるが、特徴量の次元数が増加するというデメリットがある。一方で、上記の理由から、図 5.7 に示すパターンを用いることで、2 値細線形状を表現することは十分に可能である。したがって、

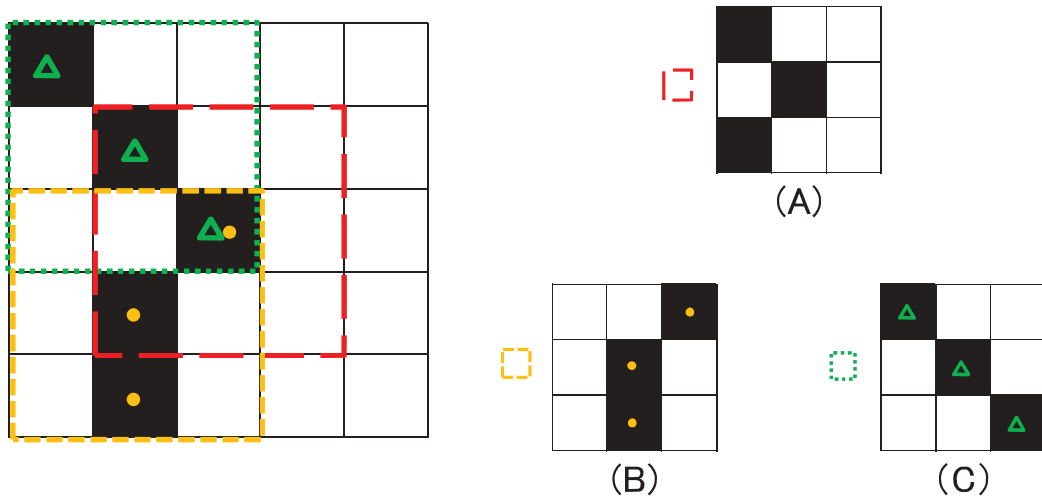


図 5.8 HBOP 特徴における線画素の表現 (1)
(自発表 [76] から引用)

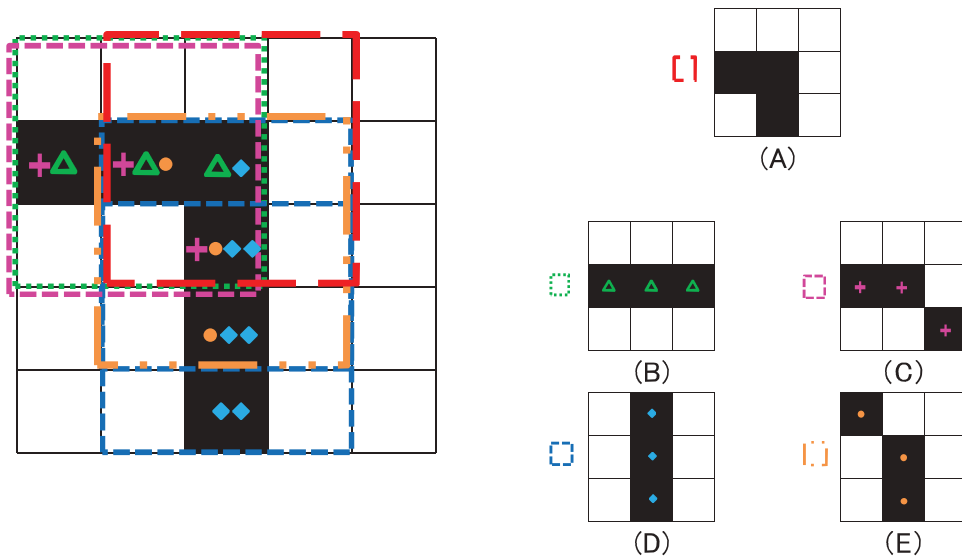


図 5.9 HBOP 特徴における線画素の表現 (2)
(自発表 [76] から引用)

5.3 HBOP 特徴量

本提案手法においては、図 5.7 に示したパターンのみを用いることとした。

5.3.3 検出処理時の計算量削減

HBOP 特徴量を用いた画像解析処理でも、他の画像特徴量と同様に、画像内から検出対象の検出処理を行う場合には、検出ウィンドウを用いたラスタスキャンを行う必要がある。ラスタスキャンにより得られる領域数は多数となり、都度、領域内の方向ヒストグラムを作成すると、その計算量は膨大となる。先に他の処理を行い候補領域の検出を行う場合等、カスケードに対象の検出処理を行う場合でも、その候補数が多い場合には上記同様の問題が生じる。

HBOP 特徴量では、対象領域内の方向パターンごとにその合計出現回数をヒストグラム化する。これは Haar-like 特徴量の領域内合計輝度計算とほぼ同様の処理である。そこで、HBOP 特徴量の計算時に、Integral Image のアルゴリズムと同じように、あらかじめ各方向パターンごとに出現回数をカウントした Integral Image を生成しておくことで、検出処理時の計算量を削減できる。

方向パターンの Integral Image の各画素の値は、は始点画素より x, y を 1 ずつ変化させラスタスキャンを行い、以下の式を用いて計算する。

$$\begin{aligned} Integral_a(x, y) = & P_a(x, y) + I_a(x, y - 1) \\ & + I_a(x - 1, y) - I_a(x - 1, y - 1) \end{aligned} \quad (5.3)$$

$$I_a(x, y) = \begin{cases} 0 & (x = 0 \wedge y = 0) \\ P_a(x, y) + I_a(0, y - 1) & (x = 0 \wedge y \geq 1) \\ P_a(x, y) + I_a(x - 1, y) & (x \geq 1 \wedge y = 0) \\ Integral_a(x, y) & (otherwise) \end{cases} \quad (5.4)$$

ここで、 I_a は計算結果の Integral Image の a 番目の方向パターンの値を示す。 P_a は方向パターンの計算によって得られた a 番目の方向パターンの実測値である。 I_a の計算を行う際に、計算に必要な I_a の値が得られている必要がある。以上の計算により、各方向パターンごとに Integral Image 化された方向パターン画像を取得する。

2 値化について

HBOP 特徴量では 2 値の線画像をその処理対象とする。入力画像がもともと 2 値線画像であればそのまま特徴量計算が可能である。また、一般の多値画像においても、一般的な前処理により 2 値線画像化することで、同様に取り扱うことができる。ただし、得られる画像は前処理に依存するため、用途により有効な 2 値線画像化を行う必要がある。

ここで、HBOP 特徴量におけるパターンのヒストグラム化は画素ごとに行われる。注目する線を構成する画素の近傍画素に存在する全てのパターンに対して投票を行う。このため、4 連結、8 連結どちらの細線化処理でも扱うことが可能である。ただし、前処理として用いる細線化手法により得られる線の位置は一般的に異なるため、学習時と検出時には同様の手法で細線化処理を行う必要がある。

一方で、マンガ画像に対する予備実験では正規化画像サイズが 150×150 [pixel] かつセル数が 5×5 の場合に良好な動作を示した。この条件では 1 セルは 30×30 [pixel] となる。このセルサ

イズは通常の人物検出で用いるものと比較して大きい。セルのサイズが大きく、かつ学習、検出で用いる細線化処理が同一という条件の下では、細線化処理アルゴリズムに起因する線の若干の差異については、セル単位ヒストグラム化により十分に吸収可能であると考えられる。

5.4 HBOP 特徴量による図形識別

5.4.1 HBOP 評価実験の概要

本章では、HBOP 特徴量の 2 値細線画像に対する特性を示すため、HOG 特徴量との比較実験を行う。単純な 2 値線画像に対する学習、識別及び検出実験を行い、HBOP 特徴量の特性を明らかにする。次に、本論文の処理ターゲットであるマンガ画像を用い、学習、識別実験を行い、その性能を示す。

5.4.2 単純 2 値細線画像に対する識別実験

5.4.3 実験に関する設定及び実験環境

HOG 特徴量と本提案方式である HBOP の 2 値線画像に対する特性を調査するため、各特長量において次元数ごとの学習の収束状態、学習時間、識別性能について調査を行った。学習用入力画像はすべて 50×50 pixel の 2 値線画像を用いた。学習用の正例として円を、不例として矩形を用いた。また、円については中心座標及び直径を、矩形については始点、幅及び高さをランダムなパラメータによって決定し作成した画像を使用する。正例及び不例にはそれぞれ 700 枚の画像を使用した。学習用画像の例を図 5.10 及び図 5.11 にそれぞれ示す。

実験において、HBOP はセル数を 3×3 , 5×5 , 7×7 と変化させた。各 HBOP 特長量の次元数はそれぞれ 72 次元, 200 次元, 392 次元である。以下ではそれぞれの次元数に基づき HBOP72, HBOP200, HBOP392 と表記する。HOG 特徴量はすべて 5×5 セル、角度の量子化ステップは 9 とし、正規化ブロックのサイズをそれぞれ 1×1 , 2×2 , 3×3 とした。各 HOG 特徴量の次元数はそれぞれ 225 次元, 576 次元, 729 次元である。以下それぞれの次元数に基づき HOG225, HOG576, HOG729 と表記する。

各特徴量の学習には Real-Adaboost アルゴリズムを使用した。Real-Adaboost の実装には中部大学藤吉研究室にて公開されているプログラム [93] を基に、提案手法に合うように一部変更したものを使用した。実験に使用した環境を表 5.1 に示す。また、実験のため、学習時間の計測時には並列処理を無効とした。

HOG 特徴量及び HBOP 特徴量の学習特性

前節にて示した学習用画像を用いた学習結果について調査を行った。実際に学習を行った際の収束状態について図 6 に示す。今回使用した画像はランダムに生成した円と四角形であるため、学習対象としては単純な部類である。このため収束は早くすべての特長量において、20 回程度までの学習で学習に用いた正例に対する True Positive が 1.0 となった。本結果をもって、今回利用する特徴量の各パラメータにおいて学習は問題なく実施されていることが確認できる。

学習に要した時間の学習回数ごとの変化を図 5.13 に示す。学習回数が極端に少ない時点におい

5.4 HBOP 特徴量による図形識別

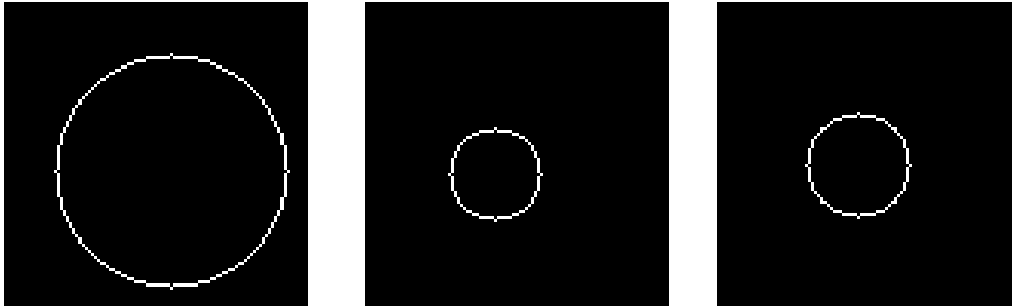


図 5.10 学習用正例画像
自発表 [76] から引用

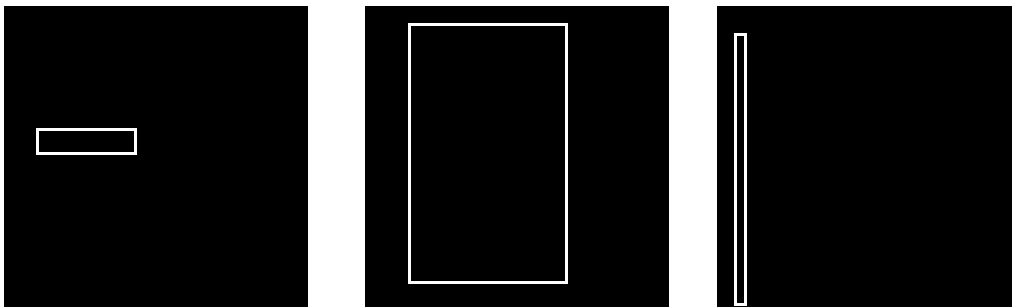


図 5.11 学習用負例画像
自発表 [76] から引用

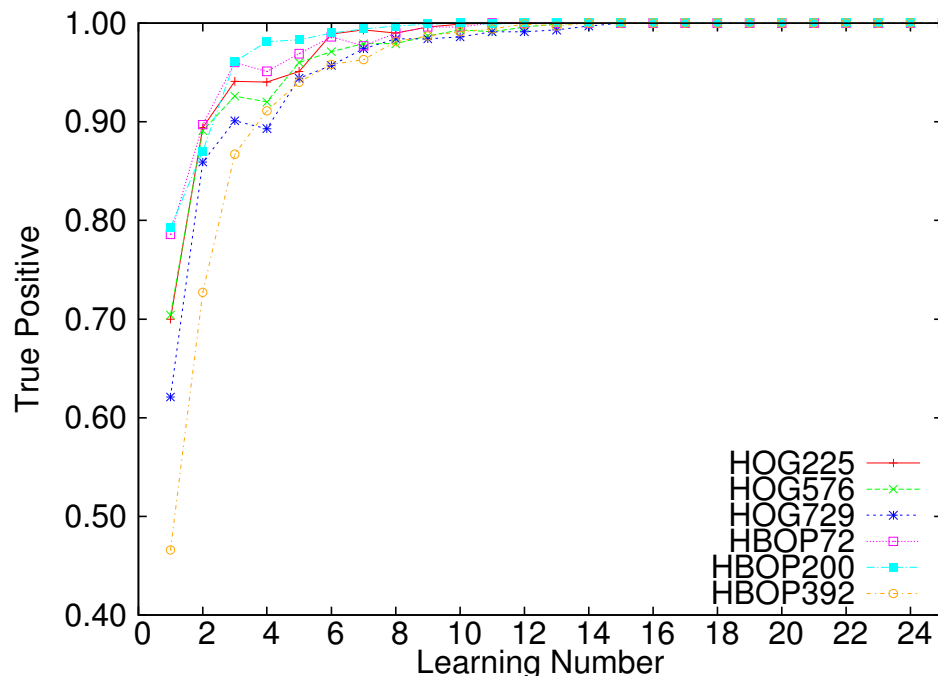


図 5.12 学習の収束状態
自発表 [76] から引用

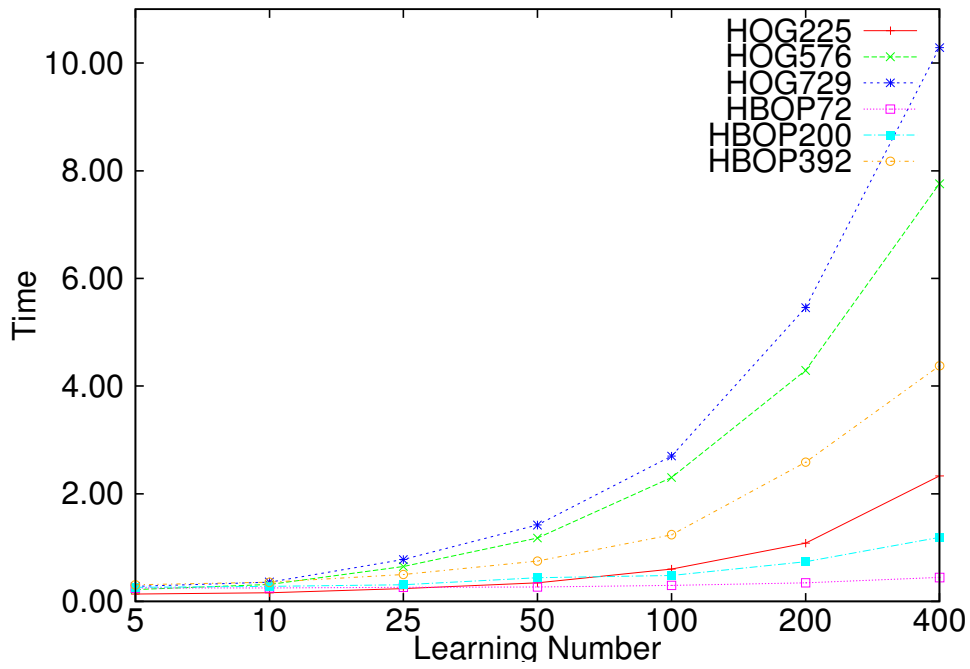


図 5.13 学習処理時間
自発表 [76] から引用

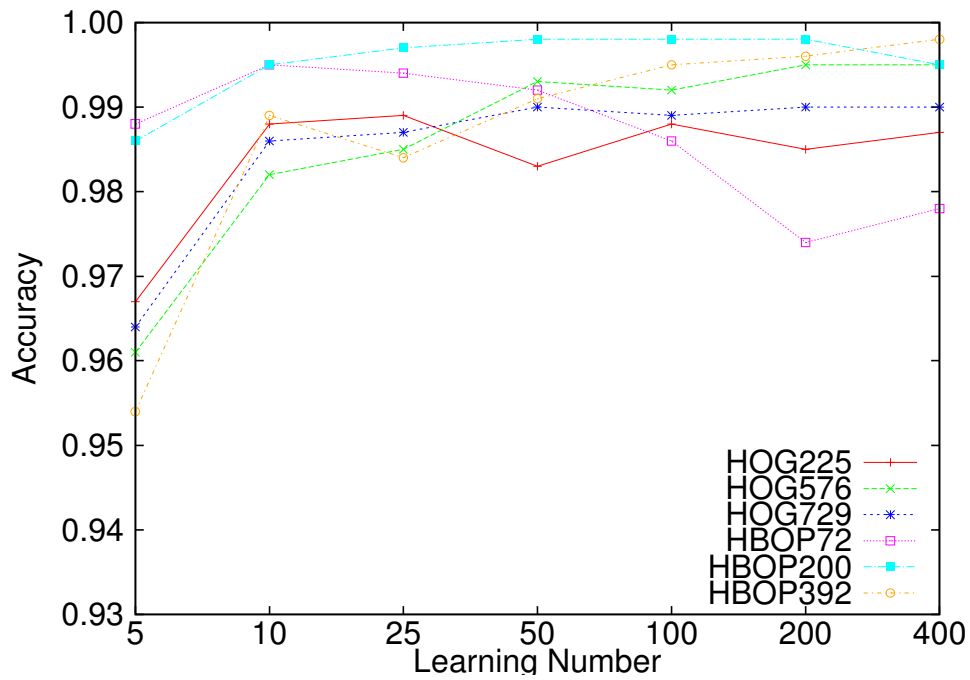


図 5.14 未知入力画像に対する平均識別精度
自発表 [76] から引用

5.4 HBOP 特徴量による図形識別

表 5.1 実験環境
自発表 [76] より一部引用

OS	Ubuntu13.10
CPU	Corei7 3770K
Memory	32GB
OpenCV	Version2.4.5
G++	Version4.8.1

表 5.2 未知入力画像（正例）に対する識別精度
自発表 [76] から引用

Iteration	HOG			HBOP		
	225	576	729	72	200	392
5	0.974	0.952	0.950	0.980	0.982	0.944
10	0.990	0.984	0.984	0.996	0.992	0.984
25	0.990	0.992	0.982	0.992	0.994	0.980
50	0.982	0.994	0.990	0.990	0.996	0.986
100	0.984	0.996	0.992	0.988	0.996	0.990
200	0.982	0.998	0.990	0.974	0.996	0.992
400	0.984	1.000	0.990	0.982	0.994	0.996

表 5.3 未知入力画像（負例）に対する識別精度
自発表 [76] から引用

Iteration	HOG			HBOP		
	225	576	729	72	200	392
5	0.960	0.970	0.978	0.996	0.990	0.964
10	0.986	0.980	0.988	0.994	0.998	0.994
25	0.988	0.978	0.992	0.996	1.000	0.988
50	0.984	0.992	0.990	0.994	1.000	0.996
100	0.992	0.988	0.986	0.984	1.000	1.000
200	0.988	0.992	0.990	0.974	1.000	1.000
400	0.990	0.990	0.990	0.974	0.996	1.000

ては、画像からの特徴量計算に関するオーバーヘッドがある。25 回以降では基本的に特徴量の次元数に応じて学習にかかる計算時間が増加していることが確認できる。

HBOP 特徴量は HOG 特徴量と比較して、ブロック単位での正規化が無いこと、方向成分の量子化パターンが限られることから、基本的に次元数が低い。このため、同程度のセル数を用いた場合には計算量の削減が見込める。学習回数が 400 回の時点においても HBOP72 及び HBOP200 については約 1 秒までの時間で学習されており、複雑なモデルを学習する際など学習の収束に多くの繰り返し学習が必要となる場合に、計算量において利点があると考えられる。

未知画像識別実験

上記学習結果を用いて、未知画像に対する識別実験を行った。本実験で用いる未知画像は学習用画像と同様の手順により生成し、円及び矩形に対する識別成功率を調査した。ここで生成したテスト用画像は、学習用画像と生成時のランダム値が異なる画像である。テスト用画像は正例 500 枚、負例 500 枚の合計 1000 枚を使用した。各特徴量における正例、負例に対する学習回数ごとの識別正解率を表 5.2 及び表 5.3 に、その平均値をグラフ化したものを図 5.14 にそれぞれ示す。実験によりすべての手法で 95 % を超える正解率を持つことが確認された。この中で、HBOP200 の学習回数 50, 100, 200 及び HBOP392 の学習回数 400 の際に最も高い平均識別精度 99.8% を達成した。また、HBOP200 においては、すべての測定点において HOG よりも高い識別成功率を得た。注意点として、すべての手法において、学習が収束していない 5 回程度の学習時点であっても未知画像に対する識別成功率が 95 % を超えている。この主な要因は本テストで学習・識別対象とした図形が比較的単純である点及び今回のテスト用画像として学習用画像同様の手順で生成された画像を使用している点によるものであると考えられる。

HBOP200 とセルの分割数が同一でブロックにおける正規化を行わない HOG225 を比較すると、HOG225 では学習が進んでも結果が安定せず、その識別精度も HBOP200 より低い値となっている。このことから、HBOP で設定した画像方向パターンへの設計は、本論文の目的である 2 値細線画像を対象とした画像特徴を得るという目的において、HOG で用いる特徴よりも高精度かつ高効率であるといえる。

5.4.4 複雑な画像からの円検出実験

より一般的かつ複雑な形状検出への適応性を調べるため、複数の円と矩形が混在する画像上での円検出実験を行う。実験に用いた画像例を図 9 に示す。実験には同様に作成した 50 枚の画像を使用した。

ここでは、先の実験において識別精度の高かった HOG576 の学習回数 400 の場合と、HBOP200 の学習回数 200 及び HBOP392 の学習回数 400 について検出結果をまとめる。

対象画像に対してスライドウィンドウによるラスタスキャンを用いて検出処理を行う。また、多数の近傍領域に対する検出結果を整理するため、Mean Shift Clustering による統合処理を行う。検出ウィンドウは正方形を用い、サイズは 25×25 画素から 125×125 画素まで 25 画素刻みで変化させた。本検出実験の評価は、円形の検出が主目的であるため、検出された領域内にある円の 75% 以上が含まれていれば成功、これ以外を失敗として評価した。ここでは、円を検出するという目的から、上記成功条件を満たしてさえいれば、検出領域内に複数の円が含まれる場合若しくは検出結果に矩形が含まれている場合も正解としている。

検出成功例と検出失敗例を図 5.16 及び図 5.17 に示す。実際に検出を行った例を図 5.18–5.20 に、本実験結果の総検出数及び検出成功率を表 5.4 に示す。ここで、識別精度は全検出結果中の正解数の割合とした。今回の実験では、検出対象の円には多数の重なりが存在し、また、画像中からの検出アルゴリズム中に統合処理が入ることから、再現率については評価を行わず、正解率についてのみ評価を行った。

5.4 HBOP 特徴量による図形識別

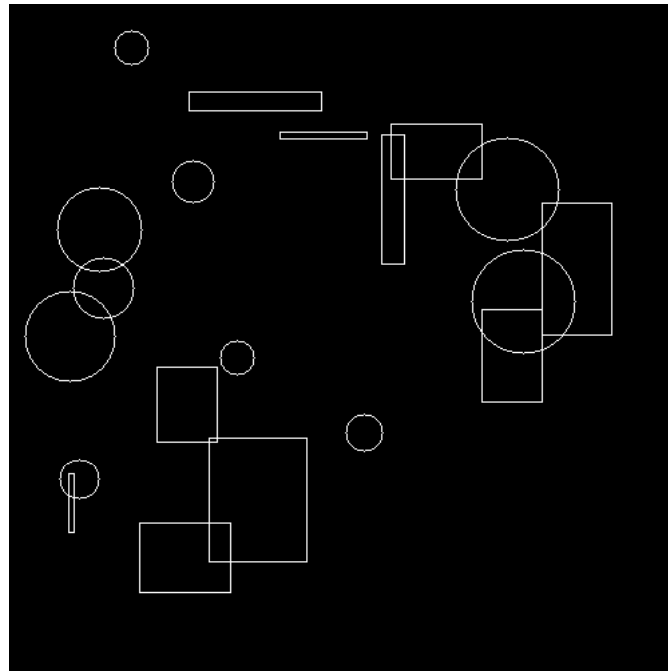


図 5.15 未知入力画像に対する平均識別精度複数の円と矩形が混在する実験画像
自発表 [76] から引用

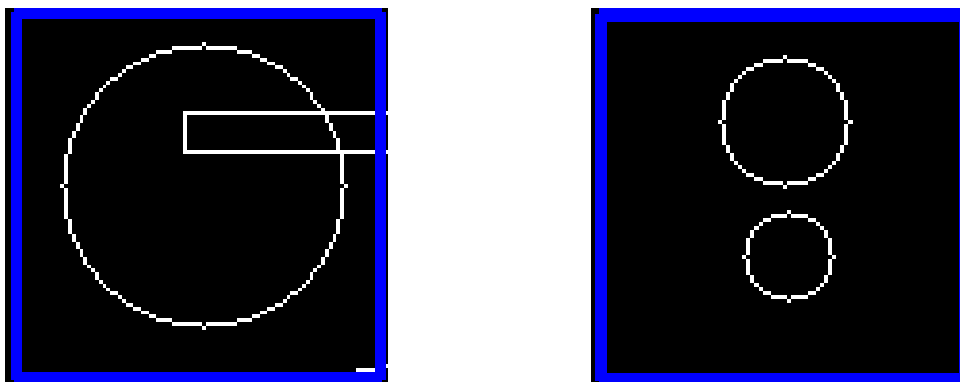


図 5.16 学習用正例画像複雑画像上の円検出における正解画像例
自発表 [76] から引用

表 5.4 に示された実験結果より、HBOP では HOG と比較して True Positive が高くまた False Positive も少ない結果となった。特に False Positive の減少により、全体としての識別成功率では 8% から 10% の向上が見られた。画像では、HOG で誤検出している矩形が重なっている部分において HBOP では誤検出が減少している。以上より、複雑な状況においても HOG 特徴量と比較して、本提案手法である HBOP 特徴量の線画像解析に対する優位性が確認できる。

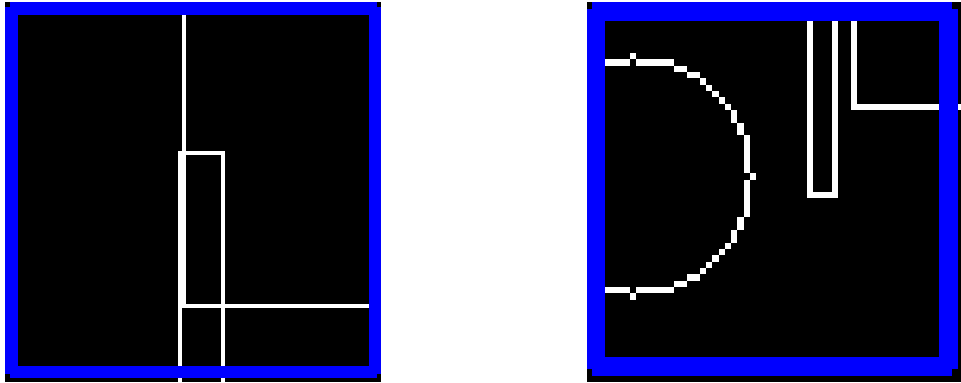


図 5.17 学習用正例画像複雑画像上の円検出における不正解画像例
自発表 [76] から引用

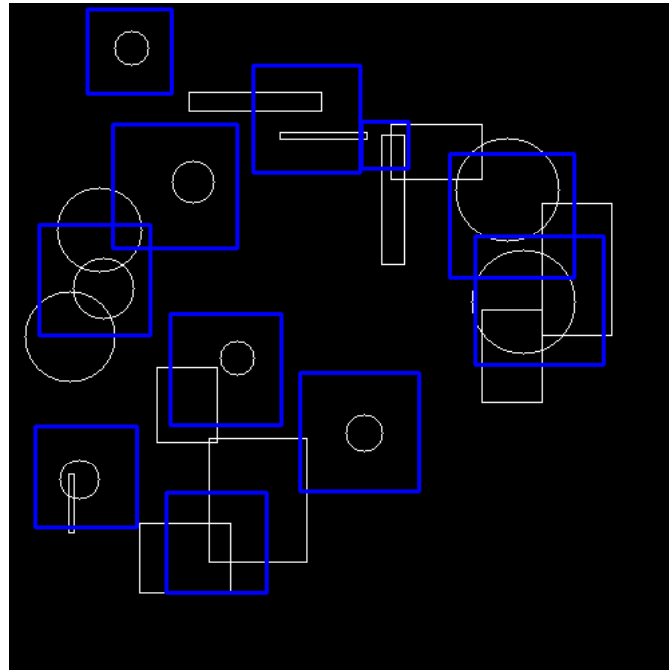


図 5.18 円検出結果 (HOG729)

表 5.4 複雑な画像における円検出実験結果
数値は自発表 [76] から引用

Method	True Positive	False Positive	Precision
HOG576-400	340	44	0.885
HBOP200-200	349	12	0.967
HBOP392-400	345	4	0.989

5.4HBOP 特徴量による図形識別

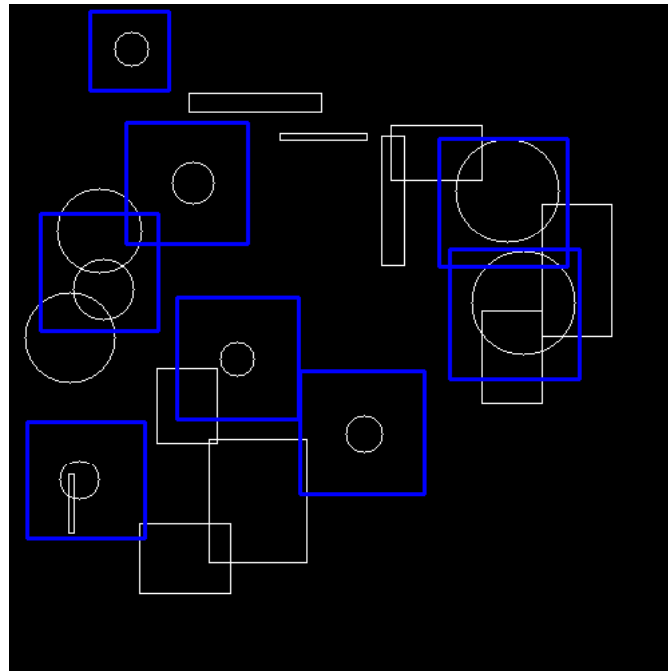


図 5.19 円検出結果 (HBOP200)
自発表 [76] から引用

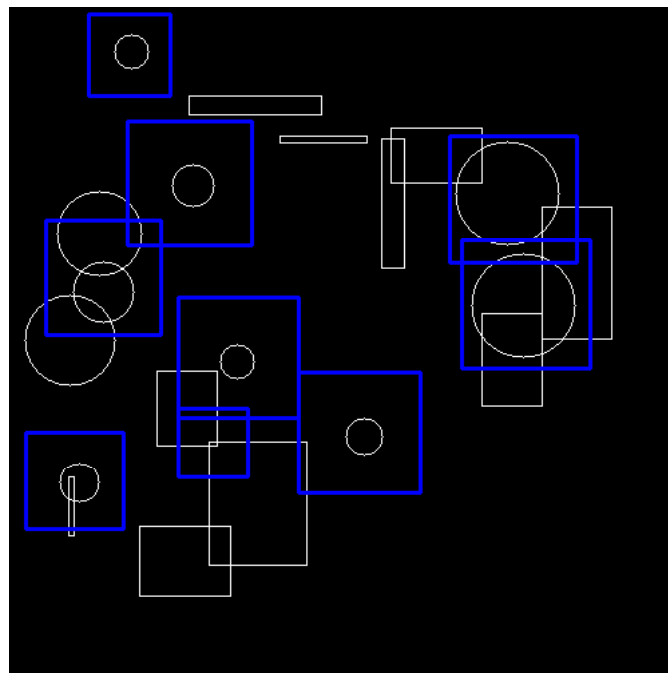


図 5.20 円検出結果 (HBOP392)

表 5.5 学習と実験に用いた画像数
 自発表 [76, 94] から引用

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

5.5 HBOP 特徴量によるマンガ画像識別

5.5.1 実験に関する設定及び実験環境

本論文において処理の主対象とするマンガ画像に対して、識別実験を行う。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、登場人物の形状の中で、特に特徴的に描かれる傾向を持つ瞳部分を学習、識別する。

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。ここで、今回使用したマンガ画像は完全な 2 値画像ではなく、高解像度グレースケール画像としてスキャンされたものを 360dpi に変換したものである。

HOG 特徴量と HBOP 特徴量は、若干の位置ずれに関しては許容可能なものの、シフト不変性は持たない。このため、実験には、画像内の同一部位から切り出し領域をさまざまに変更し取得した画像を使用した。学習用の画像には、瞳部分を切り出した画像を正例、瞳以外の領域を切り出した画像を不例とした。学習用画像数は表 5.5 に示したとおりである。識別実験用の未知画像には正例 250 枚、不例 250 枚を用いた。

HBOP, HOG ともに学習回数は 200 回とし、学習時のセル数は 5×5 、HOG 特徴のブロック数は 3×3 とした。また、HOG 特徴量の量子化方向は 9 とし、特徴量の次元数は HBOP が 200、HOG が 729 である。また、正規化画像サイズは 50×50 , 100×100 , 150×150 , 200×200 の 4 段階とする。入力画像からの 2 値細線画像の生成には、2 値化後にモルフォロジ処理による細線化を用いた。今回の HBOP 特徴量の実装では、線を白、背景を黒と定義しているため、マンガ画像からの 2 値細線化画像生成時に、画像の色反転を行っている。HOG 特徴量では、細線化されていない画像からも特徴を得ることが可能であるため、細線化、2 値化 (BW)、前処理無し (Normal) について実験を行った。実験に用いた正例画像を図 5.21–図 5.23 に、負例画像を図 5.24–図 5.26 にそれぞれ示す。

5.5.2 実験結果

実験結果を表 5.6 に示す。HOG 特徴量と比較して HBOP 特徴量では、2 値細線化された画像において True Positive の値が全パラメータにおいて向上している。True Negative の値は一部下回っているが、その低下率は True Positive の上昇分と比較して小さいものである。HOG 特徴量では、通常の入力画像において識別を行った場合と比較して、2 値画像及び 2 値細線画像を入力した場合の True Positive, True Negative が共に低下している。これは、HOG 特徴の計算過程において 2 値細線画像からの方向及び勾配値取得の問題点によるものと考えられる。HOG 特徴量



図 5.21 実験画像:正例 (BW)
自発表 [76,94] から引用



図 5.22 実験画像:正例 (BWThin)
自発表 [76,94] から引用



図 5.23 実験画像:正例 (Canny)
自発表 [76,94] から引用

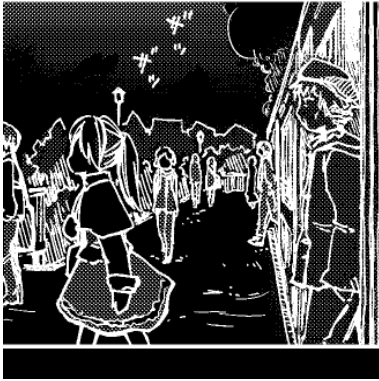


図 5.24 実験画像:負例 (BW)
自発表 [76,94] から引用



図 5.25 実験画像:負例 (BWThin)
自発表 [76,94] から引用

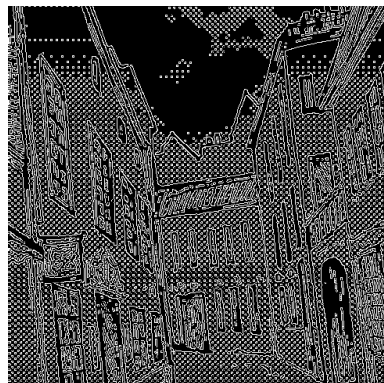
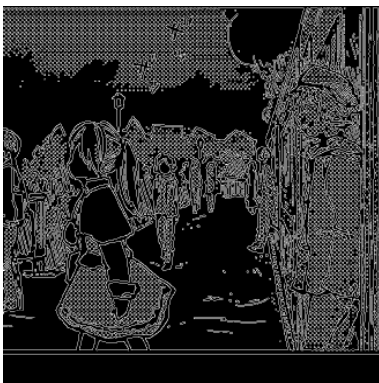


図 5.26 実験画像:負例 (Canny)
自発表 [76,94] から引用

表 5.6 識別結果
自発表 [76] から引用

Method	Resolution	True Positive	True Negative
HBOP	50	0.765	0.908
	100	0.880	0.912
	150	0.972	0.880
	200	0.964	0.805
HOG	50	0.729	0.829
	100	0.841	0.813
	150	0.805	0.884
	200	0.833	0.865
HOG(BW)	50	0.896	0.896
	100	0.861	0.912
	150	0.869	0.869
	200	0.861	0.888
HOG(Normal)	50	0.880	0.912
	100	0.928	0.932
	150	0.892	0.936
	200	0.900	0.936

では前処理なしで正規化サイズが 100 の場合に、最も優れた識別性能を得られている。HBOP における正規化サイズ 150 の場合には、これと比較して、True Negative において 0.052 の低下を伴うものの、True Positive において 0.044 の向上が見られた。一般的に、Joint 特徴量などのカスケード型の識別機を構成する場合には、True Positive の値が高い弱識別器が利用できることが望ましい。この点で HBOP 特徴量では HOG 特徴量と比較して良好な結果を得られたといえる。

5.6 むすび

本章では、マンガ画像解析に対する新たなアプローチとして 2 値細線化による構造解析を試みた。これまでマンガ画像解析で使用してきた HOG 特徴量は、細線化された 2 値画像からでは最大 4 方向の方向成分を得るが、直線自体の方向を直接記述するものではなく、HOG で作成するヒストグラムの一部には投票されないという問題点があった。そこで、2 値線画像解析の為に画像特徴量として、2 値線分の方向ヒストグラムを利用した、Histograms of Binary Orientation Pattern (HBOP) 特徴量を提案した。HBOP 特徴量では、注目点の 8 近傍画素において 12 パターンのオペレータを規定し、8 方向のヒストグラムを生成するものである。

円と矩形の学習及び識別において、HBOP 特徴量では、HOG 特徴量と比較して少ない次元数でより高い検出精度が得られた。また、学習回数を変化させて実施した識別実験においては、200 次元の HBOP 特徴量は HOG 特徴量の識別結果をすべての点で上回る結果が得られた。複数の円と矩形を配置した複雑状況下の円検出実験においては、HBOP 特徴量を使用した場合 HOG 特徴量での実験結果と比較して正解数を減らすことなく、検出結果の成功率が 8% から 10% 高い

精度で検出可能なことを確認した。以上から、HBOP の 2 値線画像特性について、従来の HOG より優れていることが確認された。また、本研究の処理対象であるマンガ画像に対しては、細線化なしの HOG 特徴量による識別結果と比較して 0.052 の True Negative の低下を伴うものの、0.044 高い True Positive を得ることが確認された。

今後の検討課題として Joint 特徴量やその他カスケード型アルゴリズム等における、HOG 特徴量の代替としての HBOP 特徴量の導入が挙げられる。Joint 特徴量等では、人物検出等において HOG 特徴量をそのまま利用した場合と比較して、良好な検出精度を実現している。2 値細線画像に対する処理においては、利用する低レベル特徴として HBOP を用いることで、より高精度な解析処理が実現可能であると考えられる。

第 6 章

結論

6.1 研究成果の概要

本論文では、マンガ画像からのメタデータ抽出を目的として、コマ解析、キャラクター解析からなるマンガ画像解析処理について検討を行った。また、マンガを構成する線画の特徴を捉えるため、2 値線画解析に利用可能な新たな画像特徴量として、2 値線分の方向ヒストグラムを利用した、Histograms of Binary Orientation Pattern (HBOP) 特徴量を提案した。提案手法の概要及び得られた研究成果を以下にまとめる。

第 3 章「マンガ画像からのコマ検出」

第 3 章「マンガ画像からのコマ検出」では、高速高精度コマ分割手法により、コマ分割処理精度の向上と処理時間の削減を行った。コマ分割処理として、これまでに輝度こう配に基づく分割線検出と、得られた分割線からなる再帰的 2 分割処理が提案されていた。従来の分割線検出手法では、画像の 2 辺を結ぶ直線を構成する画素上の輝度勾配値を取得し、この情報を基に画像の 2 分割処理を行う。一方で、コマの枠線は作者により任意に引かれるものであるため、その向きは画像に対して水平、垂直に限られない。任意角度のコマ枠線を捉えるためには、分割線検出において、基準とする直線の、角度変化量を十分に小さくする必要がある。したがって、検出精度を確保するためには、探索する分割線の数が増大となる。また、分割線の決定が輝度こう配のみに依存することから、本来分割不可能な部位にて分割処理が行われるという問題があった。そこで、高速高精度コマ分割手法では、分割線自体に幅を持たせることにより、微小な角度変化に対するロバスト性を向上させ、分割線の探索数を減少させた。コマの内部では、様々な絵の成分が存在するため輝度こう配の方向分布が複雑となる。一方で、コマの枠線はコマ内部の線と比較して単純な形状である。そのため、ノイズの影響を除き、枠線に沿った部位では輝度こう配の方向分布は一様となる。そこで、他のコマ上を通過する分割線が決定されることを防ぐため、分割線上の輝度こう配方向分布等を利用した分割線の検査を導入した。さらに、分割線検出順序の変更を導入し、従来手法と比較して高精度かつ高速な処理を可能とした。実験により、高速高精度コマ分割手法では、分割線候補の交差点近辺の直線方向を基に優先分割線を決定するアルゴリズムにより改良された再帰的 2 分割手法と比較して、処理時間及び処理精度の双方で優位性があることを確認した。また、編集距離の概念を用いた評価によりメタデータ付与に対しても優位性があることを確認した。

6.1 研究成果の概要

通常のコマ分割処理では、アルゴリズムの特性からコマ周辺の余白部を残したまま分割が行われることとなる。しかしマンガでは、コマからの要素のはみ出し等が存在するため、コマ分割処理結果画像のみからでは、周辺にある白画素が余白であるか否かの判断ができない。そこで、コマの内部塗り潰しアルゴリズムによるコマ内の絵領域の除去画像に対するコーナー検出結果及び Hough 変換結果の組み合わせと、コマ分割処理結果とを統合する、コマ角位置検出処理を提案した。これにより従来の 2 分割処理の繰り返しアルゴリズムでは考慮されない、コマ周辺部の余白除去を可能とした。

第 4 章「マンガ画像からのキャラクター情報抽出」

第 4 章「マンガ画像からのキャラクター情報抽出」では、マンガのキャラクター検出とキャラクター間の識別を行った。検出及び識別処理において使用した識別器は、一般物体認識のアプローチを利用してマンガのキャラクターを学習することで構築したものである。マンガはその基本構造が線により描かれるため、従来一般的に人間の顔検出で利用されてきた Haar-like 特徴量を用いた識別処理は適さない。そこで、人体のエッジを捉えることで輪郭の識別に有効性を持つ HOG 特徴量を利用し、マンガのキャラクター検出と、その識別処理への利用について検討を行った。実験により、HOG 特徴量と SVM を組み合わせた識別器を用いることで、マンガのキャラクター検出及びキャラクター間の識別に対しての一定の識別能力があることを確認した。また、人物検出において、瞳を学習した識別器と顔を学習した識別器をカスケード型に接続することで、単一の要素のみを学習した場合よりも検出適合率が高くなることが確認された。更に、上記結果に対し、キャラクターの顔を学習した識別器による識別処理を加えることでキャラクターごとの検出処理を構成し、最大 0.93 の適合率を得た。マンガのキャラクターは実際の人間と異なり、シーンにより非常に大きな形状の変化を持って描かれる場合や、作者の特性により、キャラクターの向きに応じてある点を境に輪郭形状が突然変化する場合がある。このため、キャラクターの状況ごとに別クラスを構成することで、識別性能を向上させられる可能性がある。実験により、キャラクターの状態を正面、横、背面、デフォルメの 4 系統に分離し学習することで、正面及び背面のクラス内識別性能の向上を確認した。

第 5 章「2 値線画像向け画像特徴量 HBOP とマンガへの適用」

第 5 章「2 値線画像向け画像特徴量 HBOP とマンガへの適用」では、2 値線画像解析のための画像特徴量として、2 値細線の方向成分をとらえる HBOP 特徴量を提案した。先の検討で使用した HOG 特徴量は、細線化された 2 値画像から最大 4 方向の方向成分を得るが、直線自体の方向を直接記述するものではなく、HOG で作成するヒストグラムの一部には投票されないという問題点がある。そこで、HBOP 特徴量では、注目点の 8 近傍画素において 12 パターンのオペレータを規定し 8 方向のヒストグラムを生成することで、2 値細線画像における線の方向を直接取得可能とした。

円と矩形の学習及び識別において、HOG 特徴量と比較して少ない次元数でより高い検出精度が得られた。また、学習回数を変化させて実施した識別実験においては、200 次元の HBOP 特徴量は HOG 特徴量の識別結果をすべての点で上回る結果が得られた。複数の円と矩形を配置した複雑状況下の円検出実験においては、HBOP 特徴量を使用した場合 HOG 特徴量での実験結果と

比較して正解数を減らすことなく、検出結果の成功率が 8% から 10% 高い精度で検出可能なことを確認した。以上から、HBOP 特徴量の 2 値線画像特性について、従来の HOG 特徴量より優れていることが確認された。

また、実際のマンガ画像に対して 2 値細線化処理を行った画像を用いて、HBOP 特徴量と HOG 特徴量の比較実験を実施した。本実験においても、HBOP 特徴量では HOG 特徴量を上回る識別精度が得られた。実験結果より、2 値の細線画像解析において、HBOP 特徴量は低次元かつ高精度であり、HOG 特徴量に対し優位性があることが確認された。また、細線化なしの HOG 特徴量による識別結果と比較した場合にも 0.052 の True Negative の低下を伴うものの、0.044 高い True Positive を得ることが確認された。この結果は、カスケード型に多数の識別処理を重ねる場合に有効である。

まとめ

本論文で提案する各手法により、従来困難であったマンガ画像の内容理解について、コマ、キャラクターの検出及び、キャラクターの識別に関して、一定の成果が得られた。更に、HBOP 特徴量により、2 値細線化された画像に対する識別性能全般を向上させることが可能である。HBOP に関する今後の検討課題として Joint 特徴量やその他カスケード型アルゴリズム等における、HOG 特徴量の代替としての HBOP 特徴量の導入が挙げられる。Joint 特徴量等では、人物検出等において HOG 特徴量をそのまま利用した場合と比較して、良好な検出精度を実現している。2 値細線画像に対する処理においては、利用する低レベル特徴として HBOP を用いることで、より高精度な解析処理が実現可能であると考えられる。

6.2 本研究成果の応用用途に対する効果と考察

本論文で得られた研究成果について、第 2 章にて示した応用用途・アプリケーションに関する達成度、検討事項及び残された課題についてまとめる。

マンガメタデータを用いたサービス（自動スクロールシステム）

自動スクロールシステムの為に必要なコマの位置及びその形状に関する情報は、本論文におけるコマ分割処理あるいはその後の研究成果 [27] を利用する事で取得可能である。また、その精度も基本的なコマ割りに対して最大 95% と非常に高い。現状得られている研究成果を利用する事で、自動スクロールシステムを構築するために必要な作業コストを大幅に低下させる事が可能である。

一部の作品では明確に定義可能な読み順が存在しない場合がある。その例として、キャラクターの大きなはみ出し等が存在する場合、複数コマが複雑につながっている場合、複数コマの境界が明確に定義されない場合、表現として複数コマの読み順が同一若しくは未定義の場合等がある。このような場合には、読み順には読み手に依存した自由度が存在する。したがって、スクロールサービスを行う上で設定される読み順については、オペレータの判断で決定する必要がある。また、単純なスクロールではなく、キャラクターの切り出しを伴うより高機能な表示には、オペレータの介在が必要である。ただし、これらの場合においても、基本的なコマ形状を取得可能で

6.2 本研究成果の応用用途に対する効果と考察

あれば、その作業コストを十分に減ずることができる。したがって現時点で本機能に要求される精度及び速度は十分に達成されていると考えられる。

マンガメタデータを用いたサービス（構成要素に基づくサーチ用クエリの取得）

マンガを構成する要素に対する検索機能を実現する為には、マンガ内の要素を理解する必要がある。マンガの内容を十分に記述するためには、“どのキャラクターが”、“いつ”、“どこで”、“何をしているか”という情報を取得する必要がある。この情報を構成する要素としてキャラクター、セリフ等が考えられる。

本論文では、一般物体認識のアプローチを応用することでキャラクターの検出及び識別が可能であることを示した。また、セリフの検出については関連研究成果 [62–64] において一定の成果が得られている。これらを統合する事で、キャラクター名と同一シーンのセリフ情報の取得が可能となる。ただし、キャラクター検出及びその識別精度に関してはまだ十分ではない。ページ内に存在するキャラクターと関連するセリフの組み合わせ手法についても、今後詳細な検討が必要である。例えば、セリフの中身についても特徴を解析する事で、人物との関連性が得られる可能性がある。ここではキャラクターとセリフについて検討したが、他に擬音や背景などのシーン情報もクエリとして考えられる。擬音は文字をベースとするもののその形状はアレンジされ、複雑な背景上に配置されている。これは認証システムでも利用される画像と同様に画像解析が困難である。本分野に対する解析精度及び対象については、今後さらなる検討が必要である。

マンガ制作支援システム（キャラクター作画支援）

2次元画像ベースのキャラクター作成支援を行う為には、キャラクター特徴の解析・学習機能並びに自然な線画の変換機能が必要である。このうち、線画の変換については、モーフィング等の処理が利用できるものと考えられるが、その詳細なアルゴリズムについては今後十分に検討する必要がある。本論文で得られたキャラクターの解析手法では、キャラクターの識別が可能である。また、キャラクター検出の過程で瞳の検出を行っている。本研究では、キャラクターの検出時に複数要素をカスケード型に組み合わせる事で、より高い検出精度を得られる事が確認された。したがって、特定キャラクターの顔として検出された領域に対して瞳検出を行う事で、瞳位置の検出が可能であると考えられる。現状では、顔要素としては瞳の検出について検討したのみであるが、口、鼻、あご等も同様に学習及び検出を行う事で、キャラクターの特徴が得られると考えられる。本機能の実現には今後、線画の変換機能と合わせた、高精度な顔要素検出手法の検討が必要である。

マンガ制作支援システム（ページ構成設計支援）

ページの構成を行う上で、読者の視線移動は作家にとって重要な情報となり得る。読者の視線追跡に関する簡易実験では、読者の着目点がセリフとキャラクターに集まる傾向が見られた。基本的な視線の移動順序に関する情報はコマ情報を用いる事で得られる。本研究成果により人物位置及びコマ情報については取得可能となった。セリフの検出については関連研究成果 [62–64] において一定の成果が得られている。これらを統合することで、手書きの原稿をスキャンした画像からの各要素の検出が可能となる。適切な検出結果が得られない箇所については一部に修正情報

を加える必要があるが、その作業量は大幅に削減される。以上より、本研究結果において支援を行う為に必要な情報は十分に取得可能であると考えられる。更なる精度の向上については今後の検討が必要である。

6.3 総括

マンガは日本を代表するコンテンツである。近年のデジタル化により、マンガのグローバル化が更に進むものと考えられる。また、デジタルデバイスの機能を活用した、新たなマンガ表現も増加すると考えられる。一般的に、マンガの作成には非常に高度な技能と作成時間が必要である。そしてグローバル化においては翻訳など様々な作業が必要となる。音楽制作の分野では、素人を始めとする音楽制作者にとってボーカルの確保が困難であった。この問題に対し、デジタル技術のサポートによりこの問題が一定レベル緩和された。その結果、従来では作品の作成が出来なかった幅広いクリエイター層による作品の発表がなされた。クリエイター層の増加は、当該分野における文化的発展に大いに寄与するものである。マンガにおいても同様に、デジタル技術による補助が有効であると考えられる。マンガコンテンツの作成に有効な画像解析及び画像処理技術は、マンガ分野の文化的発展に大いに寄与できる可能性がある。

本論文の解析対象であるマンガは、人間の想像により描かれる架空の画像である。これまでに、人間による意味的画像表現の理解として文字解析が行われてきた。マンガは文字と比較してその表現の自由度が高く解析対象としての難易度が高い。人間の思考や表現を理解することは、コンピュータサイエンスの大いなる目標の一つである。人間による意味的表現を物体認識のアプローチから識別可能であること示した点は意義がある。

本論文ではマンガのコマ解析、キャラクターの検出及び識別と 2 値細線解析用画像特徴量 HBOP の提案を行った。その結果、高機能マンガコンテンツ生成を支援する上で有用なコマ情報及びキャラクター情報を取得可能とした。これらの機能は高機能マンガコンテンツ生成に対して、直接あるいは機能実現の重要な要素として寄与するものである。HBOP 特徴量は従来の HOG 特徴量と比較して 2 値線画像に対する高い識別性能が得られた。今後、現時点で想定していない機能やメタデータを利用したマンガが提案される可能性がある。柔軟な解析技術を実現するためには、用途に応じた高精度な画像処理技術が必要であり、HBOP 特徴量を利用することで、2 値線画像全般に対する処理精度を向上させることができると考えられる。具体的なメタデータに応じた解析及び編集機能については今後更に検討が必要である。また、マンガ作成支援に関してはキャラクターの生成以外にも多数の機能が考えられる。画像処理の分野においてマンガ研究は現在いい明期といえる段階であり、今後さらなる発展が期待される。

謝辞

本研究を遂行するにあたり、今日に至るまで大変多くの方々の御指導とご協力を賜りました。お力添えを頂いた方々に感謝の辞を述べさせていただきます。

早稲田大学 大学院国際情報通信研究科 渡辺 裕教授には、本研究の機会を与えてくださり、また貴重な時間を割いて研究内容及び研究者としての方向性を示して頂いただけでなく、今日に至るまで懇切丁寧なる御指導、御助言を賜りました。ここに、心より御礼申し上げます。

早稲田大学大学院 国際情報通信研究科 亀山 涉教授には、修士課程の頃より、研究の方向性をはじめ多数の御指導並びに御助言を賜りました。ここに、心より御礼申し上げます。

早稲田大学 大学院国際情報通信研究科 坂井 滋和教授、河合 隆史教授には、本論文をまとめるにあたり大変貴重な御示唆を賜りました。ここに、心より御礼申し上げます。

研究生活において、大変多くの御意見、ご提案を頂いた渡辺研究室の皆様にご礼申し上げます。特に、研究の方向性をはじめ、日頃から研究における心構え、研究者としての考え方、研究におけるツールの使い方に至るまで、御指導、御助言を頂きました早稲田大学 国際情報通信研究センター 客員研究員 石川 孝明氏に心より御礼申し上げます。また、本研究のきっかけを与えてくださり、研究の左右もわからぬ頃より研究の進め方から文章の書き方に至るまで丁寧に御指導いただきました株式会社 KDDI 研究所 河村 圭博士に心より御礼申し上げます。本論文第3章の一部は河村氏との共同研究の成果によるものである。

研究者としての御意見を頂くのみならず、教育者としての立場並びに機会を与えて頂きました津田塾大学 学芸学部 情報科学科 小箱 亮之教授に心より御礼申し上げます。

日頃共に職務を行うだけでなく、研究活動を進める上で大変よい刺激を与えて頂きました早稲田大学国際情報通信研究科助教・助手の皆様にご心より感謝致します。特に、日頃より研究に対する御意見、御助言を頂きました菅沼 睦先生、Sriprasertsuk Pao 先生に心より御礼申し上げます。また、Ong Kok Meng 氏に心より御礼申し上げます。氏との約束は本論文を完成させる上で大変大きな支えとなりました。

助手の職務だけでなく、日頃から大変お世話になりました早稲田大学国際情報通信研究科の先生方並びに事務室の皆様にご心から感謝致します。

私が大学院に進学するきっかけ並びに初めて研究に携わる機会を与えて下さりました、東京理科大学 応用物理学科 荒木修 准教授に心より御礼申し上げます。

本論文で使用しているほぼ全てのマンガ画像は木野陽様、三原鉄也様 <http://www.etheric-f.com/> にご提供頂いた作品からのものである。ご自身の作品を、本論文を始めとする研究における実験への利用および論文内での結果の表示に対して快く許可いただき、また画像データをご提供いただいた、木野陽様、三原鉄也様に深く感謝致します。また、三原氏にはマンガ製作者の側か

らの大変貴重な知識，御意見を頂きました。重ねて御礼申し上げます。

私が長きに渡る学生生活並びに研究生活を送ることができたのは，家族の支えによるものです。家族並びに，大変辛抱強く応援をしてくれた妻に心より感謝致します。

参考文献

- [1] 知的財産戦略本部コンテンツ・日本ブランド専門調査会, “日本ブランド戦略”. http://www.kantei.go.jp/jp/singi/titeki2/houkoku/090310_nihonbland.pdf, 参照 Feb. 23, 2014.
- [2] 内閣府知的財産戦略本部, “クールジャパン推進に関するアクション,” 2012. <http://www.kantei.go.jp/jp/singi/titeki2/kettei/cjap.pdf>, 参照 Feb. 23, 2014.
- [3] 高木利弘, 電子書籍ビジネス調査報告書 2009, 株式会社インプレス R&D (編), 株式会社インプレス R&D, 2009.
- [4] 高木利弘, 電子コミックビジネス調査報告書 2009, 株式会社インプレス R&D (編), 株式会社インプレス R&D, 2009.
- [5] インターネットメディア総合研究所, 電子コミックビジネス調査報告書 2012, インターネットメディア総合研究所 (編), インターネットメディア総合研究所, 2012.
- [6] インターネットメディア総合研究所, 電子書籍ビジネス調査報告書 2013, インターネットメディア総合研究所 (編), インターネットメディア総合研究所, 2013.
- [7] IDPF, “Epub 3–international digital publishing forum,” 2013. <http://idpf.org/epub/30>, 参照 Feb.23, 2014.
- [8] Y. Kayashima, “Manga2.0”. <http://manga2dot0.boiledscope.com/>, 参照 Mar. 6, 2014.
- [9] Y. Kayashima, “hacktothebrain”. <http://hacktothebrain.jp/>, 参照 Mar. 6, 2014.
- [10] 株式会社パピレス, “電子貸本 renta! — 次世代ブック - 電子書籍 (コミック) をお得にレンタル!”. http://renta.papy.co.jp/renta/sc/frm/page/index_n.htm, 参照 Mar. 6, 2014.
- [11] NTT ソルマーレ, “電子書籍ストア コミックシーモア ntt ソルマーレ”. http://www.nttsolmare.com/service/comic_cmoa/, 参照 Mar. 6, 2014.
- [12] 夏目房之介, 竹熊健太郎, ほか, マンガの読み方, 石井慎二 (編), 宝島社, 1995.
- [13] 夏目房之介, マンガはなぜ面白いのか:その表現と文法, 日本放送出版協会, 1997.
- [14] 夏目房之介, マンガ学への挑戦, NTT 出版, 2004.
- [15] 夏目房之介, 竹内オサム, マンガ学入門, 夏目房之介, 竹内オサム (編), ミネルヴァ書房, 2009.
- [16] 三原鉄也, 杉本重雄, “デジタル環境を指向したマンガの制作プロセスのモデル化とそれに基づく制作支援,” 情報処理学会研究報告. 情報学基礎研究会報告 FI, 第 2009–FI–96 巻, pp.1–8, Nov. 2009.
- [17] 三原鉄也, 永森光晴, 杉本重雄, “デジタルマンガにおけるストーリー構造とビジュアル構

- 造を表すメタデータモデル,” 情報処理学会研究報告 IFAT, 第4巻, pp.1-8, Nov. 2011.
- [18] 落合香織, 永森光晴, 杉本重雄, “デジタルマンガの構造化フレームワークの開発-マンガの構成要素を指定する path 式の提案-,” 情報処理学会第74回全国大会講演論文集, pp.593-595, March 2012.
- [19] T. Mihara, M. Nagamori, and S. Sugimoto, “A metadata-centric approach to a production and browsing platform of manga,” 14th International Conference on Asia-Pacific Digital Libraries, pp.87-96, Nov. 2012.
- [20] 落合香織, 三原鉄也, 永森光晴, 杉本重雄, “マンガ path 式を利用したソーシャル web 上におけるデジタルマンガのアノテーション共有,” 第11回情報科学技術フォーラム講演論文集, pp.327-330, Sept. 2012.
- [21] 岩間勇介, 三原鉄也, 永森光晴, 杉本重雄, “Linked open data を利用したマンガへのアクセス支援-メタデータによるマンガ情報の可視化,” 情報処理学会講演論文集, pp.633-635, March 2013.
- [22] 荻原 彰, 三原鉄也, 永森光晴, 杉本重雄, “マンガのデザインメタデータを利用した制作過程の可視化による制作支援,” 第75回情報処理学会全国大会, pp.629-631, March 2013.
- [23] カブリン, 三原鉄也, 永森光晴, 杉本重雄, “Dbpedia を利用したマンガの書誌データからの work の同定,” 第44回デジタル図書館ワークショップ, no.44, pp.11-19, March 2013.
- [24] 松下光範, “コミック工学の可能性,” 第2回 ARG WEB インテリジェンスとインタラクション研究会, pp.63-68, May 2013.
- [25] 松下光範, “コミック工学: マンガを計算可能にする試み,” 日本知能情報ファジィ学会 ファジィシステムシンポジウム講演論文集, 第29巻, pp.875-878, Sept. 2013.
- [26] 小林由佳, 石若裕子, “漫画設計支援システム pom,” コンピュータソフトウェア, vol.25, no.1, pp.82-88, 2008.
- [27] 野中俊一郎, 沢野哲也, 羽田典久, “コミックスキャン画像からの自動コマ検出を可能とする画像処理技術「gt-scan」の開発,” 富士フィルム研究報告, no.57, pp.46-49, 2012.
- [28] 野中俊一郎, 寺横 素, 沢野哲也, 羽田典久, “電子コミック向け画像処理技術の開発とその応用,” 富士フィルム研究報告, no.58, pp.42-47, 2013.
- [29] Inc. CELSYS, “総合電子書籍ソリューション — 株式会社セルシス”. <http://www.bsreader.net/>, 参照 Mar. 6, 2014.
- [30] 辻 三郎, “スプリット検出法による文書画像構造解析,” 電子情報通信学会論文誌 (DII), vol.J81DII, no.4, pp.726-734, April 1991.
- [31] K.Y. Wong, R.G. Casey, and F.M. Wahl, “Document analysis systems,” IBM J. Res. Dev., vol.26, no.6, pp.647-656, Nov. 1982.
- [32] K. Kise, A. Sato, and M. Iwata, “Segmentation of page images using the area voronoi diagram,” Computer Vision and Image Understanding, vol.70, no.3, pp.370-382, 1998.
- [33] D.H. Ballard, “Generalizing the hough transform to detect arbitrary shapes,” Pattern Recognition, vol.13, pp.111-122, 1981.

- [34] 木村彰男, 渡辺孝志, “高速一般化ハフ変換 相似変換不変な任意図形検出法,” 電子情報通信学会論文誌 (DII), vol.J81-D-2, no.4, pp.726–734, April 1998.
- [35] 田中孝昌, 東海林健二, 外山史, “マンガ画像のコマ割り構造解析,” 電子情報通信学会総合大会, no.D-12-89, p.221, March 2006.
- [36] 藤吉弘亘, 山下隆義, 岡田和典, 前田英作, ノジクヴァンソン, 石川尋代, ドウ ソルビエフ ランソフ, -CVIM チュートリアルシリーズ-コンピュータビジョン最先端ガイド 2, 八木康史, 斎藤英雄 (編), アドコム・メディア株式会社, 2010.
- [37] 山下隆義, 藤吉弘亘, “特定物体認識に有効な特徴量,” 情報処理学会研究報告, 第 CVIM 2008(115) 巻, pp.221–236, Nov. 2008.
- [38] 山内悠嗣, 山下隆義, 藤吉弘亘, “画像からの統計的学習手法に基づく人検出,” 電子情報通信学会論文誌 D, vol.J96-D, no.9, pp.2017–2040, sept. 2013.
- [39] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” Computer Vision and Pattern Recognition 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, Vol.1, pp.I-511–I-518, Dec. 2001.
- [40] Viola P. and Jones M., “Robust real-time face detection,” International Journal of Computer Vision, vol.57, no.2, pp.137–154, 2004.
- [41] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), pp.886–893, 2005.
- [42] F. Porikli, “Integral histogram: A fast way to extract histograms in cartesian spaces,” Conference on Computer Vision and Pattern Recognition (CVPR), vol.1, pp.829–836, June 2005.
- [43] C. Hou, H. Ai, and S. Lao, “Multiview pedestrian detection based on vector boosting,” 8th Asian Conference on Computer Vision, vol.Part I, pp.210–219, Nov. 2007.
- [44] B. Wu and R. Navatia, “Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors,” International Journal of Computer Vision, vol.75(2), pp.247–266, Nov. 2007.
- [45] T. Watanabe, S. Ito, and K. Yokoi, “Co-occurrence histograms of oriented gradients for pedestrian detection,” IPSJ Transactions on Computer Vision and Applications, vol.2, pp.39–47, March 2010.
- [46] 三井, 山内悠嗣, 藤吉弘亘, “Joint hog 特徴量を用いた 2 段階 adaboost による人検出,” 画像センシングシンポジウム (SSII), pp.IN1–06, June 2008.
- [47] P.F. Felzenszwalb, R.B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part based models,” IEEE Trans. Pattern Anal. Mach. Intell., vol.32, no.9, pp.1627–1645, Sept. 2010.
- [48] P. Sabzmeydani and G. Mori, “Detecting pedestrians by learning shapelet features,” Computer Vision And Pattern Recognition 2007 CVPR’07 IEEE Conference on, pp.1–8, 2007.

- [49] Y. Freund and R.E. Schapire, “A decision–theoretic generalization of on–line learning and an application to boosting,” Computational Learning Theory Second European Conference EuroCOLT, pp.23–37, March 1995.
- [50] R.E. Schapire and Y. Singer, “Improved boosting algorithms using confidence–rated predictions,” Machine Learning, vol.37, no.3, pp.297–336, 1999.
- [51] 倉爪 亮, 石川 博, 加藤丈和, 佐藤 淳, 三田雄志, –CVIM チューリアルシリーズ–コンピュータビジョン最先端ガイド 1, 八木康史, 斎藤英雄 (編), アドコム・メディア株式会社, 2008.
- [52] B.E. Boser, I.M. Guyon, and V.N. Vapnik, “A training algorithm for optimal margin classifiers,” in Proceedings of the 5th annual workshop on Computational learning theory (COLT’92), pp.144–152, 1992.
- [53] N. Cristianini and B. Scholkopf, “Support vector machines and kernel method – the new generation of learning machines,” AI Magazine, vol.23, no.3, pp.31–41, 2002.
- [54] C.C. Chang and C.J. Lin, “Libsvm: a library for support vector machines,” ACM Transactions on Intelligent Systems and Technology (TIST), vol.2, no.27, pp.27:1–27:27, 2011.
- [55] T. Tanaka, F.T. Kenji Shoji, and J. Miyamichi, “Layout analysis of tree-structured scene frames in comic images,” Proc. IJCAI-07, pp.2885–2890, Jan. 2007.
- [56] 木野陽, ベリーベリークリームショコラふたつのベリー, 自己出版, 2010.
- [57] I. Crypton Future Media, “Vocaloid2 初音ミク (hatsune miku) — クリプトン”. <http://www.crypton.co.jp/mp/pages/prod/vocaloid/cv01.jsp>, 参照 Mar. 6, 2014.
- [58] Inc. CELSYS, “マンガ制作ソフト comicstudio | comicstudio.net”. <http://www.comicstudio.net/>, 参照 Mar. 6, 2014.
- [59] Inc. CELSYS, “イラスト、マンガ制作ソフト clip studio paint — clip studio.net”. <http://www.clipstudio.net/>, 参照 Mar. 6, 2014.
- [60] コミ Po 製作委員会有限責任事業組合 ComiPoLLP, “組み立てるようにマンガをつくる ～漫画作成ソフト コミ po! (コミポ)”. <http://www.comipo.com/>, 参照 Mar. 6, 2014.
- [61] Inc. CELSYS, “Qumarion | 人型入力デバイス”. <http://www.clip-studio.com/quma/>, 参照 Mar. 6, 2014.
- [62] 河村 圭, 石井大祐, 渡辺 裕, “線画の混在する画像におけるテキスト領域抽出の改善手法,” FIT2009 第 8 回情報科学技術フォーラム, no.H-027, pp.155–156, Sep. 2009.
- [63] 河村 圭, 石井大祐, 渡辺 裕, “コミック画像の文字位置特定に関する検討,” 電子情報通信学会総合大会, no.D-12-52, p.163, Mar. 2010.
- [64] K. Hoashi, C. Ono, D. Ishii, and H. Watanabe, “Automatic preview generation of comic episodes for digitized comic search,” MM’11 Proc. of the 19th ACM international conference on Multimedia, pp.1489–1492, 2011.
- [65] P.V.C. Hough, “Method and means for recognizing complex patterns,” 1962.
- [66] 山田雅之, 鈴木茂樹, ラフマツブディアルト, 遠藤守, 宮崎慎也, “携帯電話を利用したコミックの閲覧システムとその評価,” 芸術科学会論文誌, vol.3, no.2, pp.149–158, June

- 2004.
- [67] 石井大祐, 河村 圭, 帆足啓一郎, 滝嶋康弘, 渡辺 裕, “コミック画像におけるコマの角検出に関する一検討,” 情報処理学会オーディオビジュアル複合情報処理研究会研究報告, Vol.2010-AVM69, pp.1-6, 2010.
- [68] 石井大祐, 河村 圭, 渡辺 裕, “分割線選択によるコミックのコマ分割処理に関する検討,” FIT2006 第5回情報科学技術フォーラム, no.J-033, pp.263-264, Sep. 2006.
- [69] 石井大祐, 河村 圭, 渡辺 裕, “コミック画像のコマ分割処理における制御パラメータに関する検討,” 電子情報通信学会パターン認識・メディア理解研究会 PRMU2009-34 (IE2009-43, MI2009-34), pp.187-192, May 2009.
- [70] D. Ishii and H. Watanabe, “A study on frame position detection of digitized comics images,” Workshop Picture Coding Symposium, WP317, pp.124-125, 2010.
- [71] C. Harris and M. Stephens, “A combined corner and edge detector,” Proc. of Fourth Alvey Vision Conference, pp.147-151, 1988.
- [72] D. Comaniciu and P. Meer, “Mean shift: a robust approach toward feature space analysis,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol.24, issue 5, pp.603-619, May 2002.
- [73] D. Lowe, “Distinctive image features from scale-invariant keypoints,” International Journal of Computer Vision, vol.60, no.2, pp.91-110, 2004.
- [74] D.G. Lowe, “Object recognition from local scale invariant features,” IEEE International Conference on Computer Vision (ICCV), pp.1150-1157, 1999.
- [75] 石井大祐, 渡辺 裕, “マンガからの自動人物検出と識別に関する一検討,” 画像電子学会論文誌, vol.42, no.4, pp.457-465, 2013.
- [76] 石井大祐, 渡辺 裕, “2値線画像解析のための2値方向パターンヒストグラム特徴量HBOP,” 早稲田大学 GITI 紀要, pp.-, 2014年発行予定.
- [77] K. Takayama, H. Johan, and T. Nishita, “Face detection and face recognition of cartoon characters using feature extraction,” IIEEJ 3rd Image Electronics and Visual Computing Workshop, no.1C-5, pp.1-5, 2012.
- [78] 高山耕平, ヘンリージョハン, 西田友是, “特徴抽出によるアニメキャラクターの顔認識とキャラクター検索,” Visual Computing グラフィクスと CAD 合同シンポジウム 2012 予稿集, no.3, pp.1-6, 2012.
- [79] OpenCV.org, “Opencv”. <http://opencv.org/>, 参照 Feb. 23, 2014.
- [80] 石井大祐, 山崎太一, 渡辺 裕, “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,” 情報処理学会オーディオビジュアル複合情報処理研究会研究報告, Vol.2012-AVM80, No.1, pp.1-4, Feb. 2013.
- [81] L. Lam, S.-W. Lee, and C.Y. Suen, “Thinning method method -a comprehensive survey,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol.14, no.9, pp.869-885, Sept. 1992. Thinning Survey.
- [82] F.W.M. Stentiford and R.G. Mortimer, “Some new heuristics for thinning binary hand-

- printed character for ocr,” *Systems, Man and Cybernetics, IEEE Transactions on*, vol.SMC-13, no.1, pp.81–84, Feb. 1983. Thinning for OCR.
- [83] P. Tarabek, “A robust parallel thinning algorithm for pattern recognition,” *Applied Computational Intelligence and Informatics (SACI)*, 2012 7th IEEE International Symposium on, pp.75–79, May 2012. Thinning for OCR.
- [84] S. Mori, C.Y. Suen, and K. Yamamoto, “Historical review of ocr research and development,” *Proceedings of the IEEE*, vol.80, no.7, pp.1029–1058, July 1992. OCR Survey.
- [85] L. Lam and C.Y. Suen, “An evaluation of parallel thinning algorithms for character recognition,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol.17, no.9, pp.914–919, 1995. Thinning for OCR.
- [86] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man and Cybernetics*, vol.SMC-9, no.1, pp.62–66, 1979.
- [87] 大津展之, “判別および最小 2 乗基準に基づく自動しきい値選定法,” *電子通信学会論文誌*, vol.J63-D, no.4, pp.349–356, 1980.
- [88] C. John, “A computational approach to edge detection,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol.PAMI-8, pp.679–698, Nov. 1986.
- [89] L.S. Davis, “A survey of edge detection techniques,” *Computer Graphics and Image Processing*, vol.4, pp.248–270, 1975.
- [90] N.Otsu and T. Kurita, “A new scheme for practical flexible and intelligent vision systems,” *Proc. IAPR Workshop on Computer Vision (MVA1988)*, pp.431–435, 1988.
- [91] 大津展之, 島田俊之, 森俊二, “N 次自己相関マスクによる図形の特徴抽出,” *電子通信学会技術報告*, 第 78 巻, pp.81–90, 1978.
- [92] 米山 昇吾 平野敬, 岡田康裕, “図面画像内シンボル抽出方式の検討,” *電子情報通信学会総合大会*, no.D-11-126, p.126, March 2006.
- [93] 中部大学藤吉研究室, “Cv-tutorial”. <http://www.vision.cs.chubu.ac.jp/CVTutorial/>, 参照 Feb. 23, 2014.
- [94] 石井大祐, 張 傑, 石上 諒, 渡辺 裕, “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,” *情報処理学会オーディオビジュアル複合情報処理研究会研究報告*, 第 2014-AVM84 巻, pp.1–5, 2014.

図目次

2.1	マンガ画像例 [56]	6
3.1	コマの例	
	自発表 [67] から引用	14
3.2	コマ分割のステップ	
	自発表 [67] から引用	16
3.3	直線及び帯状領域による分割線検出処理	
	自発表 [68] から引用	17
3.4	基本検出線に対する追加画素	
	自発表 [68] から一部を変更した	18
3.5	検出線の例	
	自発表 [68] から引用	20
3.6	検出線上の輝度こう配方向分布（コマへの侵入有り）	
	自発表 [68] から引用	20
3.7	検出線上の輝度こう配方向分布（コマへの侵入無し）	
	自発表 [68] から引用	20
3.8	高速高精度コマ分割手法 (FAFSM) の処理フロー	
	自発表 [69] から引用	22
3.9	コマ分割処理の始点及び終点位置（標準）	
	自発表 [67] から一部を変更した	23
3.10	コマ分割処理の始点及び終点位置（冗長領域除去後）	
	自発表 [67] から一部を変更した	23
3.11	コマの角位置候補取得処理	
	自発表 [70] を基に一部変更を加えた	24
3.12	連結領域取得の開始点	
	自発表 [70] から引用	25
3.13	空白領域の検出結果	
	自発表 [70] から引用	25
3.14	Hough 変換結果	
	自発表 [70] から引用	26

図目次

3.15	Hough 変換により得られた直線の交点 自発表 [70] から引用	26
3.16	Harris の角検出アルゴリズムによる角検出結果 自発表 [70] から引用	26
3.17	角候補点 自発表 [70] から引用	26
3.18	T-method における画像解像度変換による成功率の変化 自発表 [69] から引用	29
3.19	T-method における画像解像度変換による CFED の変化 自発表 [69] から引用	29
3.20	T-method における画像解像度変換による処理時間の変化 自発表 [69] から引用	29
3.21	FAFSM における画像解像度変換による成功率の変化 自発表 [69] から引用	29
3.22	FAFSM における画像解像度変換による CFED の変化 自発表 [69] から引用	30
3.23	FAFSM における画像解像度変換による処理時間の変化 自発表 [69] から引用	30
3.24	各手法における処理時間と成功率 自発表 [69] から引用	30
3.25	各手法における処理時間と CFED 自発表 [69] から引用	30
3.26	FAFSM における画像解像度変換による CFED の変化 Comic-A 自発表 [69] から引用	30
3.27	FAFSM における画像解像度変換による CFED の変化 Comic-B 自発表 [69] から引用	30
3.28	FAFSM における画像解像度変換による CFED の変化 Comic-C 自発表 [69] から引用	31
3.29	FAFSM における画像解像度変換による CFED の変化 Comic-D 自発表 [69] から引用	31
3.30	提案手法の結果と、従来のコマ分割手法により分けられた各コマの境界とその角 自発表 [67] から引用	32
3.31	余白除去アルゴリズム適用後のコマ領域 自発表 [70] から引用	33
4.1	マンガ画像上の SIFT 特徴量 自発表 [75] から引用	37
4.2	IntegralImage 参照画素 自発表 [76] から引用	38

4.3	キャラクター A の正例画像 自発表 [75] から引用	50
4.4	キャラクター B の正例画像 自発表 [75] から引用	50
4.5	瞳の正例画像 自発表 [75] から引用	50
4.6	顔の学習用正例画像 自発表 [75] から引用	50
4.7	瞳及び顔の学習における負例画像 自発表 [75] から引用	50
4.8	キャラクター A,B の正解領域例 自発表 [75] から引用	51
4.9	既知画像上の瞳, 顔検出及び顔識別結果 (1) 自発表 [75] から引用	52
4.10	既知画像上の瞳, 顔検出及び顔識別結果 (2) 自発表 [75] から引用	53
4.11	既知画像上の瞳, 顔検出及び顔識別結果 (3) 自発表 [75] から引用	54
4.12	既知画像上の瞳, 顔検出及び顔識別結果 (4) 自発表 [75] から引用	55
4.13	既知画像上の瞳, 顔検出及び顔識別結果 (5) 自発表 [75] から引用	56
4.14	未知画像上の瞳, 顔検出及び顔識別結果 (1) 自発表 [75] から引用	57
4.15	未知画像上の瞳, 顔検出及び顔識別結果 (2) 自発表 [75] から引用	58
4.16	未知画像上の瞳, 顔検出及び顔識別結果 (3) 自発表 [75] から引用	59
4.17	未知画像上の瞳, 顔検出及び顔識別結果 (4) 自発表 [75] から引用	60
4.18	未知画像上の瞳, 顔検出及び顔識別結果 (5) 自発表 [75] から引用	61
4.19	キャラクターの顔の変動 自発表 [80] から引用	62
4.20	キャラクターの顔パターン分類 自発表 [80] から引用	62
5.1	HOG 特徴量計算時における参照画素 (1) 自発表 [76] を基に一部変更を加えた	66

図目次

5.2	HOG 特徴量計算時における参照画素 (2) 自発表 [76] を基に一部変更を加えた	66
5.3	Sobel オペレータ [89]	68
5.4	Sobel によるエッジの計算 (1)	68
5.5	Sobel によるエッジの計算 (2)	69
5.6	HBOP 特徴量の計算手順 (自発表 [76] から引用)	71
5.7	HBOP 特徴における近傍画素のパターン (自発表 [76] から引用)	72
5.8	HBOP 特徴における線画素の表現 (1) (自発表 [76] から引用)	73
5.9	HBOP 特徴における線画素の表現 (2) (自発表 [76] から引用)	73
5.10	学習用正例画像 自発表 [76] から引用	76
5.11	学習用負例画像 自発表 [76] から引用	76
5.12	学習の収束状態 自発表 [76] から引用	76
5.13	学習処理時間 自発表 [76] から引用	77
5.14	未知入力画像に対する平均識別精度 自発表 [76] から引用	77
5.15	未知入力画像に対する平均識別精度複数の円と矩形が混在する実験画像 自発表 [76] から引用	80
5.16	学習用正例画像複雑画像上の円検出における正解画像例 自発表 [76] から引用	80
5.17	学習用正例画像複雑画像上の円検出における不正解画像例 自発表 [76] から引用	81
5.18	円検出結果 (HOG729)	81
5.19	円検出結果 (HBOP200) 自発表 [76] から引用	82
5.20	円検出結果 (HBOP392)	82
5.21	実験画像:正例 (BW) 自発表 [76,94] から引用	84
5.22	実験画像:正例 (BWThin) 自発表 [76,94] から引用	84

5.23	実験画像:正例 (Canny)	
	自発表 [76, 94] から引用	84
5.24	実験画像:負例 (BW)	
	自発表 [76, 94] から引用	85
5.25	実験画像:負例 (BWThin)	
	自発表 [76, 94] から引用	85
5.26	実験画像:負例 (Canny)	
	自発表 [76, 94] から引用	85

表目次

3.1	計算機環境	
	自発表 [69] から引用	27
3.2	画像サイズ及び画像枚数	
	自発表 [69] から引用	28
3.3	コマの角位置解析実験における Precision	
	自発表 [67] から引用	32
3.4	コマの角位置解析実験における Recall	
	自発表 [67] から引用	32
4.1	顔識別の学習と実験に用いた画像数	
	自発表 [75] から引用	42
4.2	顔識別成功率	
	自発表 [75] から引用	43
4.3	既知画像に対する検出結果	
	自発表 [75] から引用	45
4.4	未知画像に対する検出結果	
	自発表 [75] から引用	46
4.5	学習と実験に用いた画像数	
	自発表 [80] から引用	47
4.6	識別成功率	
	自発表 [80] から引用	48
4.7	学習と実験に用いた画像数	
	自発表 [80] から引用	48
4.8	各パターンにおける識別成功率	
	自発表 [80] から引用	48
5.1	実験環境	
	自発表 [76] より一部引用	78
5.2	未知入力画像（正例）に対する識別精度	
	自発表 [76] から引用	78

表目次

5.3	未知入力画像（負例）に対する識別精度 自発表 [76] から引用	78
5.4	複雑な画像における円検出実験結果 数値は自発表 [76] から引用	81
5.5	学習と実験に用いた画像数 自発表 [76, 94] から引用	83
5.6	識別結果 自発表 [76] から引用	86

研究業績

論文誌論文	<p>○石井, 渡辺, “2 値線画像解析のための 2 値方向パターンヒストグラム特徴量 HBOP,” 早稲田大学 GITI 紀要 (投稿中).</p> <p>○石井, 渡辺, “マンガからの自動人物検出と識別に関する一検討,” 画像電子学会論文誌, Vol.42, No.4, pp.457-465, July 2013.</p> <p>K. Kawamura, D. Ishii, H. Watanabe, “Automatic Scale Detection for Contour F . . . Vol. E94-D, pp.1998-2005, Oct. 2011.</p> <p>河村, 石井, 渡辺, “全変動最小化の高速計算手法,” 電子情報通信学会論文誌, Vol.J93-D, No.3, pp.326-335, Mar. 2010.</p> <p>○ [レター] 石井, 河村, 渡辺, “コミックのコマ分割処理に関する一検討,” 電子情報通信学会論文誌 D, Vol. J90-D, No.7, pp.1667-1670, July 2007.</p>
査読付き国際会議	<p>○ D. Ishii and H. Watanabe, “Multi Size Eye Detection on Digitized Comic Image,” IEVC2012, Nov. 2012.</p> <p>K. Hoashi, C. Ono, D. Ishii, H. Watanabe, “Automatic Preview Generation of Comic Episodes for Digitized Comic Search,” MM’ 11 Proc. of the 19th ACM international conference on Multimedia, pp. 1489-1492, Nov. 2011.</p> <p>○ D. Ishii and H. Watanabe, “A Study on Frame Position Detection of Digitized Comics Images,” Workshop Picture Coding Symposium, WP3-17, pp.124-125, Dec. 2010.</p> <p>K. Kawamura, D. Ishii, H. Watanabe, “Automatic Scale Detection Based on Difference of Curvature,” 2009 International Workshop on Advanced Image Technology (IWAIT 2009), Jan. 2009.</p>

国内研究会	<p>石井, 張, 石上, 渡辺, “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,” 情報処理学会 AVM 研究会研究報告, Vol.2013-AVM84, No.5, pp.1-5, Feb. 2014.</p> <p>石井, 山崎, 渡辺, “漫画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,” 情報処理学会 AVM 研究会研究報告, Vol.2013-AVM80, No.1, pp.1-4, Feb. 2013.</p> <p>辻, 飯倉, 小館, 石井, 下村, “登場人物の瞳および顔要素の配置特徴に基づく漫画作品類似度評価,” 情報処理学会 AVM 研究会研究報告, Vol.2013-AVM80, No.2, pp.1-5, Feb. 2013.</p> <p>山崎, 石井, 渡辺, “フーリエ記述子と色情報を用いた野菜識別手法,” 情報処理学会 AVM 研究会研究報告, Vol.2012-AVM79, No.27, pp.1-4, Dec. 2012.</p> <p>石井, 渡辺, “マンガ画像からの自動キャラクター位置検出に関する検討,” 情報処理学会オーディオビジュアル複合情報処理研究会 IPSJ-AVM, Vol.2012-AVM76, No. 2, pp. 1-5, March, 2012.</p> <p>石井, 河村, 帆足, 滝嶋, 渡辺, “コミック画像におけるコマの角検出に関する一検討,” 情報処理学会オーディオビジュアル複合情報処理研究会 AVM, AVM-69, pp.1-6, July 2010.</p> <p>石井, 河村, 渡辺, “コミック画像のコマ分割処理における制御パラメータに関する検討,” 電子情報通信学会パターン認識・メディア理解研究会 PRMU2009-34 (IE2009-43, MI2009-34), pp.187-192, May 2009.</p> <p>河村, 石井, 渡辺, “曲率空間尺度を用いたベクタ変換の評価手法に関する検討,” 情報処理学会 AVM 研究会, 2007AVM60-05, Mar. 2008.</p> <p>河村, 石井, 渡辺, “疑似 TV ノルムの数値計算とその領域分割への応用に関する検討,” 情報処理学会 AVM 研究会, 2007AVM59-10(CS2007-39 IE2007-122 pp.57-62), Dec. 2007.</p> <p>河村, 石井, 渡辺, “人工的な多値画像のベクタ変換を考慮した領域分割に関する検討,” 情報処理学会 AVM 研究会, 2007AVM58-08, Sep. 2007.</p>
-------	--

国内大会, シンポジウム等	<p>河村, 石井, 渡辺, “ベクター表現に適した少数画素精度の改良境界線検出に関する検討,” 情報処理学会 AVM 研究会, 2007AVM56-11, Mar. 2007.</p> <p>河村, 山本, 石井, 渡辺, “ベクター表現による文字や網点を含む画像符号化システム,” 情報処理学会 AVM 研究会, 2006AVM54-06, Sep. 2006.</p> <p>石井, 山崎, 渡辺, “漫画上のキャラクター識別に関する一検討,” 情報処理学会全国大会, 3D-2, March 2013.</p> <p>石井, 渡辺, “マンガ画像における瞳検出の解像度依存性に関する検討,” 画像電子学会年次大会, R2-4, June, 2012.</p> <p>河村, 石井, 渡辺, “コミック画像の文字位置特定に関する検討,” 電子情報通信学会総合大会, D-12-52, Mar. 2010.</p> <p>河村, 石井, 渡辺, “領域分割に基づくベクタ表現を用いた画像符号化方式の検討,” 画像符号化シンポジウム PCSJ P-3.09, Oct. 2009.</p> <p>河村, 石井, 渡辺, “線画の混在する画像におけるテキスト領域抽出の改善手法,” FIT2009, H-027, pp.155-156, Sep. 2009.</p> <p>河村, 石井, 渡辺, “グラデーションメッシュ最適化手法の検討,” 電子情報通信学会総合大会, D-11-36, Mar. 2009.</p> <p>石井, 河村, 渡辺, “アンチエイリアシングを考慮したベクタ表現による濃淡画像符号化に関する一検討,” 画像符号化シンポジウム PCSJ2008, P-2.04, Oct. 2008.</p> <p>河村, 石井, 渡辺, “グラデーションメッシュによる画像符号化の基礎検討,” 画像符号化シンポジウム PCSJ2008, P-5.04, Oct. 2008.</p> <p>河村, 石井, 渡辺, “スケール不変な曲線特徴に関する検討,” FIT2008 H-025, Sep. 2008.</p> <p>河村, 石井, 渡辺, “全変動ノルムの離散化手法に関する検討,” 電子情報通信学会総合大会, D-11-25, Mar. 2008.</p>
---------------	---

	<p>石井, 河村, 渡辺, “ベクタ表現を用いた画像のエッジ再現手法に関する一検討,” 情報処理学会全国大会, 3Y-1, Mar. 2008.</p> <p>河村, 石井, 渡辺, “直感的な曲線操作を実現するベクタ変換の検討,” 情報処理学会全国大会, 4B-6, Mar. 2008.</p> <p>石井, 河村, 渡辺, “動画像の時空間ベクタ表現に関する一検討,” 画像符号化シンポジウム PCSJ2007, P-1.03, Nov. 2007.</p> <p>河村, 石井, 渡辺, “人工的な多値画像のベクター変換,” 画像符号化シンポジウム PCSJ2007, P-2.21, Nov. 2007.</p> <p>河村, 石井, 渡辺, “ベクター表現に適した小数画素精度の改良境界線検出に関する検討,” FIT2007 第6回情報科学技術フォーラム, I-045, Sep. 2007.</p> <p>河村, 山本, 石井, 渡辺, “電子透かしを埋め込んだベクター画像の客観評価手法に関する検討,” 電子情報通信学会総合大会, D-11-42, Mar. 2007.</p> <p>石井, 河村, 渡辺, “コミック画像における高速コマ分割処理の改良に関する検討,” 電子情報通信学会総合大会, D-12-158, Mar. 2007.</p> <p>石井, 河村, 渡辺, “コミック画像におけるコマ分割処理の高速化に関する検討,” 情報処理学会全国大会, 1P-2, Mar. 2007.</p> <p>河村, 山本, 石井, 渡辺, “ベクター表現による文字や網点を含む画像符号化システム,” 情報処理学会全国大会, 3B-1, Mar. 2007.</p> <p>河村, 山本, 石井, 渡辺, “ベクター表現によるマンガ符号化システム,” 画像符号化シンポジウム PCSJ2006, P-2.06, Nov. 2006.</p> <p>石井, 河村, 渡辺, “コミックのコマ分割処理に関する一検討,” 画像符号化シンポジウム PCSJ2006, P-2.06, Nov. 2006.</p> <p>石井, 河村, 渡辺, “分割線選択によるコミックのコマ分割処理に関する検討,” FIT2006 第5回情報科学技術フォーラム, J-033, Sep. 2006.</p>
--	--