

早稲田大学大学院国際情報通信研究科

博士論文概要

論文題目

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

申請者

石井	大祐
Daisuke	Ishii

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

2014年 3月

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、2011年にクールジャパン推進に関するアクションプランが示された。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。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や、一部がアニメーション化された作品などがある。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。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。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．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 や SVM 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．

本論文の構成について概要と共に示す．

第 1 章では本論文の研究背景，目的及び概要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。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勾配のみ依存するため、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勾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。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値細線分の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。