

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 14 日現在

機関番号：32689

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330137

研究課題名(和文) マンガ検索のための自動要約生成に関する研究

研究課題名(英文) Research on Automatic Digest Generation for Comic Retrieval

研究代表者

渡辺 裕 (Watanabe, Hiroshi)

早稲田大学・理工学術院・教授

研究者番号：10329154

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：マンガの自動要約を実現するためには、マンガに含まれるメタデータの抽出が必要である。このうち最も重要なものは登場キャラクターである。本研究では、キャラクター検出の高精度化を進めた。その結果、HOG特徴量を改良したDeformable Part Model (DPM) が有効であることを示した。また、事前知識がない状態からクラスタリングにより主要キャラクターを同定する手法について検討した。これにより、正例と負例をDPMに与えることが可能となった。さらに、ロールコミュニティモデルというキャラクターの登場場所・回数を検査する手法により、特定ページの重要度が決定でき、自動要約が実現できることがわかった。

研究成果の概要(英文)：It is important to extract metadata contained in comics for realizing automatic digest generation system. The most important metadata is comic characters. In this research, we focused on the improvement of comic character extraction. Through experiments, we showed that Deformable Part Model (DPM), which is based on HOG features, is effective for extraction. We also studied on main character identification without prior knowledge by clustering. By this approach, positive and negative samples can be fed into DPM. Further, "Role Community Model", which is the method to identify the weight of each page by counting locations of character's appearance and their relations, can be used to realize automatic comic digest.

研究分野：マルチメディア

キーワード：マンガ 顔画像検出 機械学習 自動要約 キャラクター抽出 HOG特徴量 Deformable Part Model
ロールコミュニティ

1. 研究開始当初の背景

(1) マンガは内閣官房知的財産戦略本部においても、世界的に競争力のある重要なコンテンツであると認識されている[1] [2]。また、アニメの原作供給源としてもマンガは重要なコンテンツであり、より電子書籍の普及と共に、海外に対してもより一層の発信が望まれている[3]。現在の電子書籍としてのマンガは、その多くが非常に簡単な試読ページしか持たず内容理解が容易ではない。しかも多くの電子書籍やマルチメディアデータは、大規模データベースに包含されつつある。そのため、ビッグデータ時代にはユーザは所望のマンガに到達することが困難な状況にある。マンガの試読インタフェースやアニメの視聴インタフェースは、出版社やポータルサイトのシステム毎に異なっており、日本国内はもとより世界中で異なっている。したがって、システムに依存せずにマンガのストーリー把握を可能とするような、エピソード要約や登場キャラクターの提示は、ユーザにとって検索のための重要な情報となる。マンガに含まれる台詞が英語や中国語に変換されて世界へ発信される場合にも、エピソード要約や登場キャラクター提示はそのまま運用できる利点を有すると考えられる。

(2) マンガからメタデータを抽出するための解析技術は、携帯電話での閲覧を目的としたコマ検出を中心に進められてきた。コマ検出及びコマ順序決定処理に加えて、吹き出し検出による台詞の抽出により、半自動的ではあるが表示面積の狭い携帯電話に対応させることができつつある。さらに、「若手研究(B)平成24年度-平成25年度 マンガ画像解析及びキャラクター検出手法に関する検討、研究代表者：石井大祐」においては、マンガに含まれるキャラクター検出を目的として、瞳形状の検出を試みている[4]。しかし現状ではキャラクター検出までには到達していない。一方、自然画像やアニメなどの多譜調画像に対しては顔画像検出技術が成熟期に達しており、多くの研究成果が得られている。また、アニメ用顔検出プログラム Anime Face が Web 上で公開されている。しかし、2値画像であるマンガに対しては認識精度が必ずしも高くはないという問題点がある。

2. 研究の目的

マンガ及びアニメは日本のデジタルコンテンツ産業の中でも世界的に競争力のあるコンテンツである。本研究では、大規模マンガデータに対してユーザ主導型でのマンガ検索を可能にするため、登場キャラクターの自動配列及びエピソードの自動要約の実現を目的とする。大規模マンガ・アニメデータから特定のエピソードを検索可能とするためには、エピソードとしてマンガ・アニメを集約するための技術が必要となる。特にマン

ガに対しては、エピソード生成の基盤となるマンガキャラクターの顔画像の検出と、登場キャラクターの顔画像リストの自動生成が重要となる。得られた登場キャラクターの顔画像の出現頻度、サイズ、順序情報などによって、キャラクター顔画像リストを判断基準に加えたエピソードの自動要約を実現する。

3. 研究の方法

(1) マンガのキャラクター顔画像に対して有効に動作するハール変換型(Harr-like)の特徴量について検討する。従来のハール変換型(Harr-like)の特徴量は、目、鼻、口の平均輝度パターンの変化を基にしており、Viola&Jones(2001)[5]の研究以降、デジタルカメラ、監視カメラなどのアプリケーションで幅広く使われている。ハール変換型(Harr-like)の特徴量を、線画主体である2値画像の場合にも対応できる特徴量に修正する。

このためには、マンガのキャラクター顔画像を数多く集積したデータセットが必要となる。従来はWebクローリングといった巡回型の処理によって、顔画像切り出しが可能であったが、現状では著作権遵守から困難な状況にある。また後述する教師付き学習には、正解データだけでなく、非対象データの蓄積も必要となる。一般画像については、数多くの研究用画像データセットが、大学や研究機関から提供されている。しかし、マンガコンテンツに関しては研究用データセットを半自動で作成せざるを得ず、時間と人手がかかる作業になると考えられる。

次に、教師付き学習を用いる識別手法SVM(Support Vector Machine)を適用し、顔画像検出を行い、その有効性を確認する。識別データセットのうちの正解例と不正解例を蓄積し、データセットの更新を行う。

(2) 局所的なハール変換型特徴量をマンガ特有の他の特徴量と組み合わせ、共起確率特徴量として AdaBoost などの高度な識別器を多段に組み合わせることで、検出精度の向上を図る。複数の局所の特徴量の関連性、共起性を捉えた特徴量を用いる。具体的には、自然画像に対して用いられている Haar-like 特徴量に共起性を考慮した Joint Haar-like 特徴量の組合せ[6]や Haar-like 特徴量とピクセル差分特徴量の特徴を持ち合わせたスパース特徴量を組み合わせ[7]について検討する。また局所の特徴量を AdaBoost で識別した後に生成した次段の特徴量に対して、さらに共起する特徴量を AdaBoost で識別する手法の適用を検討課題とする。これらの高度な識別器は演算量が膨大になる危険性がある。そのため、画像データセットには高速化のために積分画像に変換されている必要がある。得られた識別器の検出精度の向上を図り、さらにデータセットの更新を行う。

次に、エピソードの自動要約の対象とするマ

マンガに対して顔画像検出を行い、特定の正解画像及びそれに類似する他の正解画像を検索することにより、キャラクター画像の分類を行う。分類されたキャラクターは登場人物と捉えることができ、そのグループに属する正解画像は、特定のキャラクターの顔表情の変形とみなすことができる。これらを判断するための識別方法の検討を行う。前段の顔画像検出器は、顔画像であるかどうかを真偽として判断するため、得られた正解顔画像のグルーピングには顔検出時の識別器を用いることができない。ここではグループ化処理のために、正解画像間の類似度を与える特徴量の検討が課題となる。類似度判定により、顔画像はいくつかの登場キャラクターにグルーピングされる。

さらに、登場キャラクターリストの自動生成を行い、顔画像リストを判断基準に加えたマンガエピソードの自動要約を生成し、その有効性を示す。登場キャラクターの顔画像の出現頻度、サイズ、登場順序などの情報を基に、重要性を示す尺度を求める。得られた重要性尺度によって、登場キャラクターのリストを自動生成する。従来のマンガエピソードの自動要約手法で用いられている特徴量（コマ位置、コマサイズ、台詞）などに加えて、コマに対応するキャラクター情報を加えることで、エピソードの自動要約の精度を高める。得られた登場キャラクターリスト及び自動要約されたエピソードを検索のための情報として提示することにより、大規模マンガデータに対するコンテンツアクセスがより容易になると考えられる。最後に、主観評価実験を行い、提案手法の有効性を示す。

4. 研究成果

(1) マンガエピソードの自動要約生成には、登場キャラクターリストの自動生成が必須である。キャラクターの抽出には、その顔画像抽出が最も有効であると考えられる。キャラクター抽出を目的とした特徴量として、対象をハール変換型(Haar-like)のみに限定せず、勾配ヒストグラム(HOG)を基盤としたDeformable Part Modelなどの最新の手法も線画像に適用してその効果を確認した。また判別にはサポートベクターマシン(SVM)を用いることとした。キャラクター抽出には、マンガのキャラクター顔画像を数多く集積したデータセットが必要となる。そこで研究に協力的なマンガ作家の好意により、いくつかのマンガをデータセットとして提供してもらい顔画像(正例)切出し及び不正解画像(負例)切出しを行い、データベース化に向けた作業に入った。2013年度の研究成果として以下の2点があげられる。

勾配ヒストグラム(HOG)特徴量を用いてキャラクター抽出を試みる場合、正例と負例画像の識別の際にキャラクターが含まれる背景画像も抽出性能に影響を与えることが

分かった。背景がセリフ入りの吹き出し画像、オノマトペが含まれる画像、それ以外に3分類した場合、オノマトペが背景に存在する場合には、キャラクター抽出への影響は小さいことが分かった。

勾配ヒストグラム(HOG)を基盤としたDeformable Part Modelは、顔画像中のいくつかの部分の相対的位置関係を利用するため、ハール変換型(Haar-like)特徴量よりもマンガキャラクターの顔画像抽出精度が高いことが分かった(図1)。

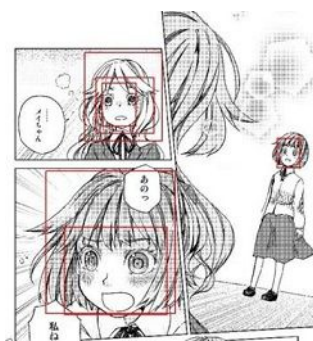


図1 DPMによるキャラクター顔検出例*

(2) 2014年度は、マンガキャラクターの顔検出を中心に各種メタデータの抽出を目標とした。研究方法としては、まず予備実験によりマンガのキャラクター抽出を目的とした。顔画像抽出のための特徴量としては、HOGを基盤としたDeformable Part Modelが最も効果的であることが分かった。マンガキャラクターは、前髪、目、口などの位置がマンガ毎やキャラクター毎に大きく変化する。そこで、Deformable Part Modelをマンガキャラクター抽出へ適用し、その検出特性の改良について検討した。

2014年度に得られた研究成果は以下の3点である。

Deformable Part Modelをキャラクターの顔画像抽出に適用して検出実験を行った。その結果、単純なHOG特徴量とSupport Vector Machine(SVM)識別器を用いる場合に比べて、Deformable Part Modelによる検出時のF-measureは、既知画像に対して17.7%改善され96.3%となり、未知画像に対しては56.8%改善され92.2%となった(表1、2、図1、2)。

表1 従来のHOGとDPMのキャラクター顔画像検出精度の比較実験結果(既知画像)

	適合率	再現率	F-measure
DPM	92.9%	100%	96.3%
HOG	75.0%	82.6%	78.6%

表2 従来のHOGとDPMのキャラクター顔画像検出精度の比較実験結果(未知画像)

	適合率	再現率	F-measure
DPM	85.5%	100%	92.2%
HOG	32.7%	38.7%	35.4%

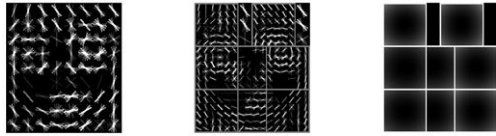


図2 DPMにおけるルートフィルタ HOG 特徴量 (左)、パートフィルタ HOG 特徴量 (中)、パートフィルタ配置 (右) の例

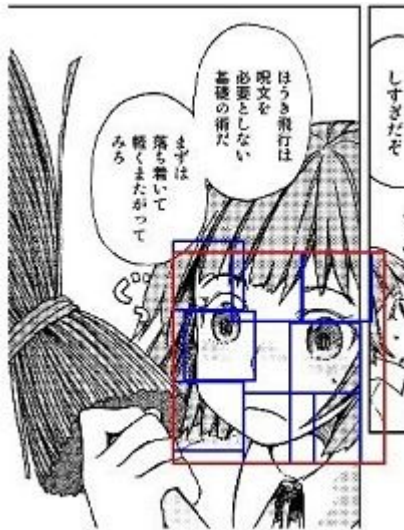


図3 パートフィルタ配置の例*
(*:木野陽:ベリーベリークリームショコラ
ふたつのベリー(2010))

Deformable Part Model の高精度化を目指して、局所バイナリパターン (Local Binary Pattern, LBP) 特徴量を組み合わせた場合に、検出率はほぼ変化せずパートフィルタ位置が変化する結果が得られた。

HOG+SVM 識別器やDPMによる識別器をマンガキャラクターに適用する場合、正例として複数のキャラクターを混合して学習させた後に検出を行う場合と、キャラクター毎に個別に学習させた場合を比較した場合、混合学習よりも個別学習によって検出精度が向上することが分かった。

(3)2015年度に得られた研究成果は以下の3点である。

キャラクター抽出の高精度化(RCNNによる高精度化と横顔対応):
マンガ画像からの登場キャラクターを検出については、HOG 特徴量や Deformable Part Model (DPM) などの一般物体検出に用いられる手法がマンガ画像に対しても有効であることが示されている。一方、現在の一般物体検出の分野において、畳み込みニューラルネットワーク(CNN)を用いた手法が既存手法を上回る検出精度を示していることが示されている。そこで CNN で画像特徴量学習し、物体検出を行なう手法である Regions with

Convolutional Neural Network Features (R-CNN) のマンガ画像への適用について検討を行なった。その結果 R-CNN は DPM と同様に、高い精度でキャラクターの検出が可能であることが分かった。一方で、キャラクターの横顔を含んだ顔検出に関する研究は行われていない。本研究では DPM および R-CNN を用いたマンガキャラクターのマルチビュー顔検出について、複数の横顔モデルと準備する手法の検討を行なった。その結果、Selective Search を利用した RCNN が最も検出率に優れることがわかった。

コミック自動要約(ロールコミュニティの応用):

コミック各巻でのストーリーの本筋を見出して、要約することによって、トレーラーを作成する手法について検討した。提案手法では、映画要約のために提案されたロールコミュニティモデルを利用した。このモデルでは、同じ時間帯、場所での一場面をシーンと定義する。各シーンの登場人物(ロール)の集合をロールコミュニティと呼ぶ。同じロールを要素として持つロールコミュニティをユニークロールコミュニティとしてまとめる。ユニークロールコミュニティの重みを計算することにより、重要度を判定する。与えた重みに比例させて対応するシーンにページ数を割り振り、トレーラーを作成した。主観評価実験の結果、コミックにおいても映画要約と同様の手法で、登場人物の関係性とその頻出度によってストーリーの本筋を見出すことができた。また従来の「試し読み」よりも少ないページ数で理解しやすいトレーラー作成が可能であることがわかった。

主要コミックキャラクターの自動抽出 (Bag-of-Visual Words のクラスタリング):

コミックからキャラクター情報を取得する際には機械学習を前提とする場合がほとんどである。しかし機械学習の際には、正例と負例が必要であり、事前知識として主人公がすでに同定されていなければならない、という問題があった。そこで本研究では、事前知識なしで主要キャラクターを同定する手法について検討した。主要キャラクターとは「物語に大きく関わるキャラクター」と定義すれば、物語において全ての登場キャラクターの中でも主要キャラクターは最も登場回数が多いと考えられる。一般的な顔画像抽出器を用いて多くの候補を抽出した後に、Bag-of-Visual-Words を用いてベクトル化した候補を k-means 法でクラスタリングすることにより、最大クラスターに属する顔画像を主要キャラクターとして抽出する手法を提案した。スコア付与による客観評価実験により、主要キャラクターが効率良く抽出できることがわかった。

図4に研究成果の配置を示す。論文 ①は発表論文、①~⑩は学会発表に対応している。

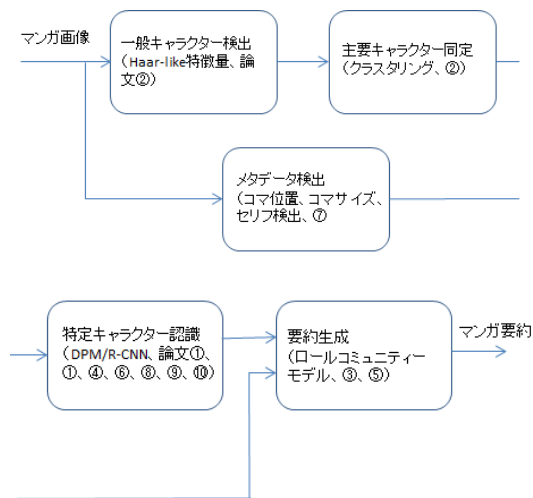


図4 研究成果の配置

<引用文献>

[1] 日本ブランド戦略～ソフトパワー産業を成長の原動力に～内閣官房知的財産戦略本部コンテンツ・日本ブランド専門調査会、平成21年3月10日
Web資料：
http://www.kantei.go.jp/jp/singi/titeki2/houkoku/090310_nihonbland.pdf (最終確認2016/6/14)

[2] 「neo anime」産業のビジネスモデルに関する調査研究報告書、株式会社日本総合研究所、平成20年12月
Web資料：
http://www.meti.go.jp/policy/mono_info_service/contents/downloadfiles/anime.pdf (最終確認2016/6/14)

[3] 技術戦略マップ2012 (コンテンツ分野) 経済産業省商務情報政策局サービス政策課メディアコンテンツ課、平成24年4月
Web資料：
http://www.meti.go.jp/policy/mono_info_service/contents/downloadfiles/120307-2.pdf (最終確認2016/6/14)

[4] 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AVM研究会研究報告，Vol.2012-AVM76，No.2，pp.1-5，Feb. 2012

[5] P. Viola and M. Jones: "Rapid Object Detection Using a Boosted Cascade of Simple Features," IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), ISSN: 1063-6919, Vol. 1, pp. 511-518, December 2001

[6] T. Mita, T. Kaneko, B. Stenger, O. Hori: "Discriminative Feature co-occurrence Selection for Object Detection," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.30, No.7, pp.1257-1269, 2008

[7] C. Huang, H. Ai, T. Yamashita, S. Lao, M. Kawade: "Incremental Learning of

Boosted Face Detector," In Proc. IEEE International Conference of Computer Vision, pp.1-7, 2007

5. 主な発表論文等

〔雑誌論文〕(計2件)

Hideaki Yanagisawa, Daisuke Ishii, Hiroshi Watanabe, Face detection for comic images with deformable part model, The 4th International Workshop on Image Electronics and Visual Computing 2014 (IEVC2014), 査読有, 4A-1, 2014

<http://www.kodama.com.hiroshima-u.ac.jp/iieej/IEVC2014/>

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、査読有、Vol.42、No.4、2013 pp.457-465

https://www.jstage.jst.go.jp/browse/iieej/42/4/_contents/-char/ja/

〔学会発表〕(計10件)

柳澤秀彰、渡辺裕、マンガキャラクターのマルチビュー顔検出に関する検討、電子情報通信学会総合大会、D-11-12、2016年3月15日18日、九州大学、福岡県福岡市

長尾一輝、渡辺裕、コミックにおける主要キャラクター同定の検討、電子情報通信学会総合大会、D-21-3、2016年3月15日18日、九州大学、福岡県福岡市

野口多紀、渡辺裕、電子コミックトレーラーの作成手法の検討、電子情報通信学会総合大会、D-21-4、2016年3月15日18日、九州大学、福岡県福岡市

柳澤秀彰、渡辺裕、R-CNNを用いたマンガキャラクター検出に関する一検討、映像メディア処理シンポジウム、I-4-12、pp.1-2、2015年11月18日20日、ラフォーレ修善寺、静岡県伊豆市

藤岡恭平、渡辺裕、マンガ要約手法に関する一考察、電子情報通信学会ソサイティ大会、A20-11、2015年9月8日11日、東北大学、宮城県仙台市

柳澤秀彰、石井大祐、渡辺裕、マンガの複数キャラクターに対する顔検出率について、電子情報通信学会総合大会、D-12-31、2015年3月10日13日、立命館大学、滋賀県草津市

石井大祐、柳澤秀彰、三原鉄也、永森光晴、渡辺裕、マンガの構成要素に基づく自動シーン分割処理に関する一検討、情報処理学会AVM研究会研究報告、Vol.2014-AVM87、No.15、pp.1-4、2014年12月4-5日、大阪大学、大阪府大阪市

柳澤秀彰、石井大祐、陳明、渡辺裕、マンガ画像からの顔検出におけるパーツ特徴量の一検討、映像情報メディア学会年次大会、17-9、2014年9月2日9月2日、大阪大学、大阪府吹田市

陳明、柳澤秀彰、張傑、石井大祐、渡辺裕、マンガにおけるHOG+AdaBoostによる顔画像

検出の性能評価、映像情報メディア学会年次大会、17-4、2014年8月31日 9月2日、大阪大学、大阪府吹田市

Mei Chen, Hideaki Yanagisawa, Daisuke Ishii, Hiroshi Watanabe, A Note on Face Detection of Comic Image with Different Background, 電子情報通信学会総合大会、BS-1-47、2014年3月18-21日、新潟大学、新潟県新潟市

〔図書〕(なし)

〔産業財産権〕

出願状況(なし)

取得状況(なし)

〔その他〕

マンガデータベース、MANGARESEARCH.ORG、<http://mangaresearch.org/>
渡辺研究室論文リスト
<http://www.ams.giti.waseda.ac.jp/publication-new.htm>

6. 研究組織

(1) 研究代表者

渡辺 裕 (WATANABE, Hiroshi)

早稲田大学・理工学術院・教授

研究者番号：10329154

(2) 研究分担者

石井大祐 (ISHI, Daisuke)

早稲田大学・理工学術院・助手

研究者番号：40581525

(平成27年2月退職)

(3) 連携研究者

なし