

機械学習によるシーン抽出のためのマンガのメタデータ提供システム

—Linked Open Data に基づくコマの属性記述と利用—

Manga Metadata Provider to help Scene Extraction by Machine Learning

- Description of Frame Attributes of Manga and its Use based on Linked Open Data -

久行 智恵^{*1} 三原 鉄也^{*2} 永森 光晴^{*2} 杉本 重雄^{*2}
 Chie Hisayuku Tetsuya Mihara Mitsuharu Nagamori Shigeo Sugimoto

^{*1} 筑波大学情報学群情報メディア創成学類

College of Media Arts, Science and Technology University of Tsukuba

^{*2} 筑波大学図書館情報メディア系

Faculty of Library, Information and Media Studies University of Tsukuba

When analyzing contents of Manga, it is necessary to use semantic information related to objects in an image which always cause difficulties when distinguishing them by computers. In this research, we constructed a system for the analysis of Manga by machine learning process that provides such semantic information and corresponding image information as metadata based on Linked Open Data. This system discriminates scene change frame by vectorizing and learning the metadata of the objects and layouts, shapes of frames. It also presents the result together with the Manga image.

1. はじめに

デジタルマンガは年々流通量が増加しており、近年は手軽にデジタルマンガに触れることができるようになった。こうした状況を背景として、デジタルの特性を活かし、計算機の援用によるマンガの利活用に関する研究が広がりつつある。

そのひとつとして、これまでにマンガのダイジェストの自動生成を目的としたシーンの抽出が試みられてきた[久行 15a][久行 15b]。特に久行[久行 16]では、コマの属性とシーン転換箇所の関係を機械学習することによるシーン転換コマの抽出について検討している。しかし、コマの属性やシーン転換箇所のような画像データから直接読み取れないマンガの内容に関わる情報が研究利用しやすい形式で提供され、さらにこれらの情報に対応するマンガ画像も合わせて利用できることは稀である。本研究では、コマの属性を用いたシーン抽出のためのより効率的な機械学習のために、コマの属性やシーン転換箇所をマンガ画像のメタデータとして提供し、機械学習に利用するシステムを構築した。

2. 機械学習によるマンガの内容解析

2.1 機械学習を用いたマンガの内容解析に関する研究

マンガ画像を学習することによる内容解析の研究として、石井[石井 13]とChu[Chu 15]が挙げられる。石井は、画像解析手法のひとつでエッジの方向を捉える特徴量であるHOG(Histograms of Oriented Gradient)と統計的機械学習であるSVM(Support Vector Machine)を用いてマンガ画像から特定のキャラクターを検出する手法を提案している。学習データは学術目的の使用許可を取った1作品のマンガ画像から手作業で切り出したキャラクターの顔や瞳の画像を用いている。また Chu は、LDA(Latent Dirichlet Allocation)に基づくマンガの作風についての評価モデルを定義し、機械学習によってマンガ画像の作風による著者分類や作風分類を試みている。

2.2 マンガのデータセットと研究での利用

マンガ画像を研究に利用するためには著作権処理が必要になる場合が多い。こうした手続きを軽減してマンガの研究を支援するために提供されているマンガ画像のデータセットとして、Manga109[Matsui 15]とeBDtheque[Rigaud 15]が挙げられる。Manga109 では著作者から許諾を得た研究利用可能な日本のマンガ作品 109 冊(約 2 万ページ)の画像データセットを提供しており、そのうち 4 作品(784 ページ)にはページ毎のセリフデータが CSV 形式で提供されている。ただしセリフデータは十分な構造化がされておらず、作品ごとにファイルが、ページ毎に行が分かれており、1 列目にページ番号、2 列目にセリフのテキストが記述されるというデータ構造になっている。eBDtheque では日本、アメリカ、ヨーロッパのマンガ 100 ページ分の画像データが、SVG 形式で記述されたテキストの内容やテキスト、フキダシ、コマの位置情報、フキダシとテキストの対応についての情報と合わせて提供されている。

2.3 マンガの内容解析におけるメタデータ提供の課題

内容解析をはじめとするマンガの研究では、マンガ画像だけでなく、マンガ画像から計算機で判別することが困難なコマやキャラクターといった画像内の細粒度オブジェクトや著者などのマンガの内容に関わる情報も利用できることが望ましいが、このような情報は画像にも増して提供量が少ない。しかし機械学習による内容解析を行うには、大量の学習データが求められることに加えて、利用したい特徴が網羅されたデータを用意する必要がある。Manga109 や eBDtheque では内容に関わる情報が提供されているものの、利用できるデータセットは限られているため、マンガ一般の内容解析を行うには不十分である。タイトルや著者などの書誌情報や大まかなストーリー、キャラクター設定などの情報は Wikipedia や書籍販売サイトといった Web 上の情報資源から得ることができる[三原 15]が、画像内の細粒度オブジェクトに関する情報が Web 上の情報資源として提供されることは稀である。また Web 上で提供される情報資源の多くはそれに対応するマンガ画像と紐付けされておらず、情報と画像を一体にして利用することはできない。

連絡先: 永森光晴, 筑波大学図書館情報メディア系, 茨城県つくば市春日 1-2, nagamori@slis.tsukuba.ac.jp

そこで、マンガの内容の特徴に関連する情報をマンガ画像に対するメタデータにすることで、画像上では点や線の集合で表されるオブジェクトの意味をはじめとしたマンガの内容に関わる情報を内容解析に利用できるようにする。本研究では、コマに含まれる要素やコマの配置、形状といった属性のメタデータを提供し、そのメタデータを利用して機械学習によるシーン転換箇所を推測するシステムの構築を行う。

3. 機械学習によるシーン抽出におけるマンガのメタデータの提供と利用

本研究では次の要求要件を満たすメタデータを提供できるようにする。まず、マンガの内容に関わる細かな情報が内容解析に利用できるようにメタデータにする。これを実現するには、著者などの書誌情報やコマなどの画像内のオブジェクト、シーンなどのストーリーを構成する要素が細かく扱えるようなメタデータモデルを考える必要がある。次に、解析結果がマンガ画像と合わせて確認できるメタデータにする。そのためには、画像とメタデータを同時に提供できる環境を用意した上で、マンガの内容に関わる情報がマンガの画像領域と紐付けされる必要がある。最後に、Linked Open Data(LOD)に基づき広く共有、利用できるメタデータにする。すなわち、Web 上で扱いやすいメタデータの記述仕様である RDF に基づいてメタデータを記述しなければならない。

これらの要求要件を満たすために、本研究では International Image Interoperability Framework(IIIF)を用いたマンガの内容の特徴に関する細かな構造メタデータを提供及び利用するシステムを設計した。まずメタデータのモデルを設計し、マンガの内容情報を細かく構造化し、画像に対するメタデータとして提供できるようにした。次に、提供されたコマの属性やシーン転換箇所のメタデータを利用した、機械学習によるシーン転換箇所の解析を行った。さらに、この解析結果の検証を助けるために、解析結果をアノテーションとしてマンガ画像上に提示した。

4. IIIF に基づくマンガの細粒度構造メタデータ

4.1 IIIF の概要

IIIF[IIIF 17]とは Web 上でデジタル画像の豊かなやり取りを実現するための国際的な枠組みであり、技術仕様として Image API と Presentation API を取り決めている。Image API は構造的な URI を用いることで LOD に基づいた画像データのやり取りを実現するための仕様である。Presentation API は Image API に基づいてアクセスした画像をビューワ上に表示するために必要な画像の構造やレイアウトに関する情報を記述するための仕様である。

この IIIF を用いることで、マンガのメタデータを提供し利用するための環境を容易に実装することができる。まず Image API によってコマやキャラクターのようなオブジェクト情報と画像領域が URI を用いて紐付けることができる。また Presentation API によってマンガ画像を作品毎にバインディングして順番に表示することができる。加えて画像に対するアノテーションを付与することができるため、マンガの内容に関わる情報をアノテーションとしてマンガ画像上に表示することができる。

IIIF を利用するために、本研究では Presentation API が定める画像に関するリソースタイプのモデル(図 1)をマンガのメタデータモデルに組み込む。リソースタイプは 9 つあり、ひとつの画像データセットの構造やレイアウトに関わる情報をリソースタイプに則して分類することで、Presentation API が定める JSON フォーマットどおりに画像をビューワ上に表示するためのメタデータ

を作成することができる。以下にリソースタイプの概要を述べる。Manifest(マニフェスト)はリソースタイプの核であり、ひとつの画像データセット全体を表す。Collection(コレクション)は複数のマニフェストやコレクションをグルーピングしたものである。Canvas(カンバス)はマニフェストを構成する、画像や画像に紐付けられた情報の入れ物であり、Sequence(シーケンス)によってカンバスの順序が定められる。Content(コンテンツ)は画像や画像に紐付けられた情報そのものを指し、Annotation(注釈)によってカンバスと紐付けられる。AnnotationList(注釈リスト)はコンテンツのうち画像以外のリソースを紐付けている注釈をカンバスごとにまとめたものであり、カンバスに紐付けられる。Range(レンジ)は複数のカンバスやレンジをグルーピングしたものであり、Layer(レイヤー)によってレンジと注釈リストを紐付ける。これらのリソースタイプとマンガの内容に関わる情報をマッピングさせることで、ビューワによる画像とメタデータの同時提供を実現する。

4.2 マンガの細粒度構造メタデータ

本研究で提案するマンガのメタデータモデルは図 2 で示すように WorkEntities、StoryEntities、ExpressionEntities、DataEntities の 4 つの実体集合から成り立つ。

WorkEntities、StoryEntities、ExpressionEntities はデジタル環境でマンガを扱うためのメタデータモデル[Morozumi 09][萩原 13]に基づいて設計しており、マンガの内容に関わる情報をそれぞれの実体集合に分類した。WorkEntities はマンガを識別し属性を与えるための書誌情報に関する実体を集めており、マンガ作品名や著者名、作品間の繋がりに関するメタデータを提供することができる。StoryEntities は物語の構成や設定に関する実体を集めており、シーンやエピソード、キャラクターのプロファイルのような設定資料に関するメタデータを提供することができる。ExpressionEntities は、マンガとして認知できる具体的な表現物である画像や画像内の細粒度オブジェクトとその構造に関する実体を集めており、落合[落合 12]のメタデータモデルに基づいて画像内のオブジェクトを細かく構造化したことで、コマやキャラクター、セリフなどに関するメタデータを提供することができる。

DataEntities は 4.1 節で述べたリソースタイプのモデルに基づいて設計しており、ビューワによる画像と合わせたメタデータの提示を実現するためのメタデータを提供する。そのために、他 3 つの実体集合に分類された実体と DataEntities の実体のマッピングを行った。図 3 に「ブラックジャックによろしく」を用いたマッピング例を示す。WorkEntities に属する情報のうち、「ブラックジャックによろしく 1」のような出版物は Manifest に、「ブラックジャックによろしく」のような出版物が属する作品シリーズは

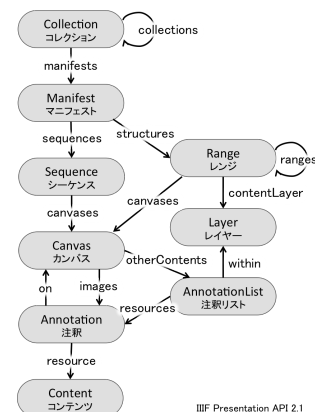


図 1 画像のリソースタイプのモデル

Collection にマッピングできる。StoryEntities に属する情報のうち、「基礎研修編」のような物語中のひと区切りは Collection に、「第 1 話」のような出版物内のひと区切りや「手術シーン」のような場面は Range にマッピングできる。ExpressionEntities に属する情報のうち、ページ番号は Sequence に、「1 ページ目」のような画像リソースの入れ物は Canvas に、さらに、StoryEntities における「手術シーン」というシーン名や「斉藤英二郎」のキャラクター設定資料の内容情報、ExpressionEntities におけるページ画像や画像内のオブジェクトの意味や内容情報は Content にマッピングされ、Annotation でページに対応づけられる。

本研究では 3 作品分(計 678 ページ)の StoryEntities のシーンに関する情報と ExpressionEntities のコマ、キャラクター、セリフに関する情報、これらに対応する DataEntities を RDF の turtle 形式で記述し、システムで約 20 万トリプル RDF メタデータを提供できるようにした。

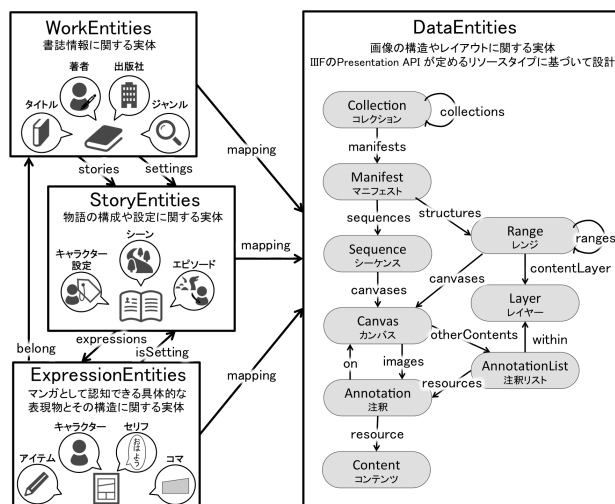


図 2 マンガのメタデータモデルの構造

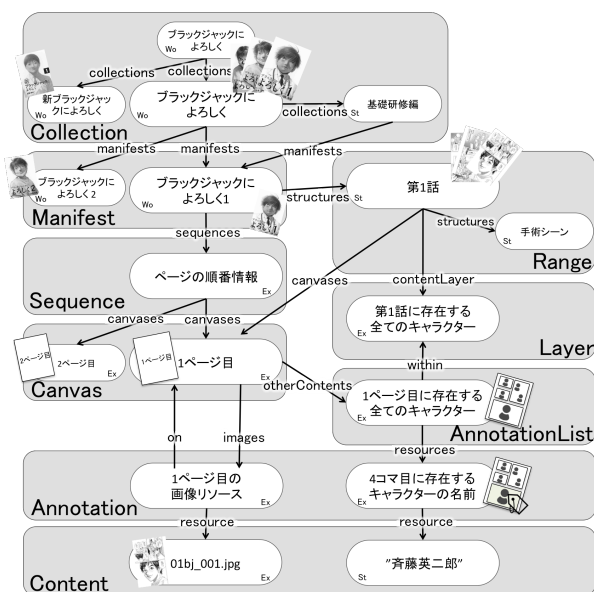


図 3 マッピングの例

5. 機械学習によるコマの属性からのシーン転換箇所抽出

本研究では、コマの属性とシーン転換箇所の関係を SVM で機械学習し、解析結果としてシーン転換箇所を示したメタデータを提供するシステムを設計した。システムの実装には Ruby を用いており、RDF メタデータをストアし検索するための SPARQL サーバには Virtuoso[Virtuoso 10]を用いた。SVM の実行ツールには LIBSVM[Lin 11]を用いた。

図 4 に本システムの構成を示す。本システムは内容解析部とメタデータ提供部の 2 つから成る。内容解析部では、コマの属性に関するメタデータを用いてシーン転換箇所の機械学習を行う。まず学習用データ生成プログラムは、3 作品の中からユーザーが選択した作品の学習に必要なメタデータを SPARQL サーバに問い合わせる。コマの頂点座標、コマ内のキャラクター及びテキストの有無、シーンが転換するコマについてのメタデータを得たプログラムは、これらを特徴量化したデータを LIBSVM に渡し、LIBSVM はシーン転換箇所の解析を行う。このシーン転換箇所推定結果はメタデータ提供部に渡される。メタデータ提供部では、ビューワ上にマンガ画像と合わせて解析結果を表示するためのメタデータ(マニフェストファイル)をユーザーに提供する。マニフェストファイル生成プログラムは、SPARQL サーバから問い合わせた選択作品の画像に関するメタデータ、すなわち 4 章の DataEntities のメタデータを用いてビューワにマンガ画像を表示し、内容解析部から受け取ったシーン転換箇所推定結果をコマに対するアノテーションとして提示するようなマニフェストファイルを生成する。

6. Presentation API を利用したシーン転換箇所推定結果の提示

本システムが提供するマニフェストファイルは IIIF の Presentation API が定める JSON フォーマットに則して生成される。このマニフェストファイルを IIIF 対応ビューワに読み込むことで、シーン転換箇所推定結果がマンガ画像と合わせて閲覧することを可能にした。IIIF 対応ビューワのひとつである Mirador[Mirador 16]に本システムにマニフェストファイルを表示させた様子を図 5 に示す。アノテーションがあるコマ枠が青色で表示され、コマにマウスカーソルを合わせると推定結果がテキストでポップアップされる。シーン転換箇所として正しく推測された場合は「Match Origin & SVM Predict!」、シーン転換箇所だが推測できなかった場合は「Origin Scene Change Frame」、推測結果が誤っている場合は「SVM Predict Frame」が現れる。

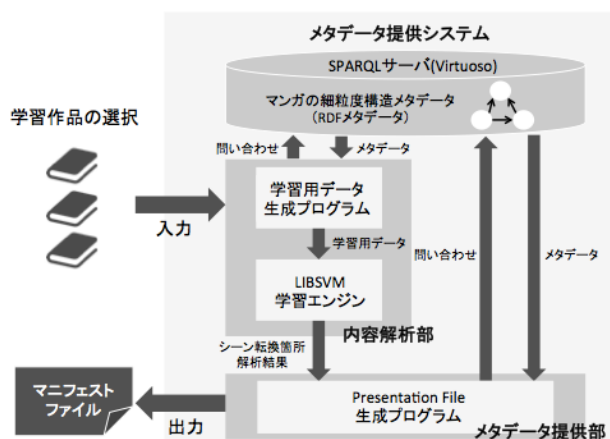


図 4 シーン転換箇所抽出システムの構成図

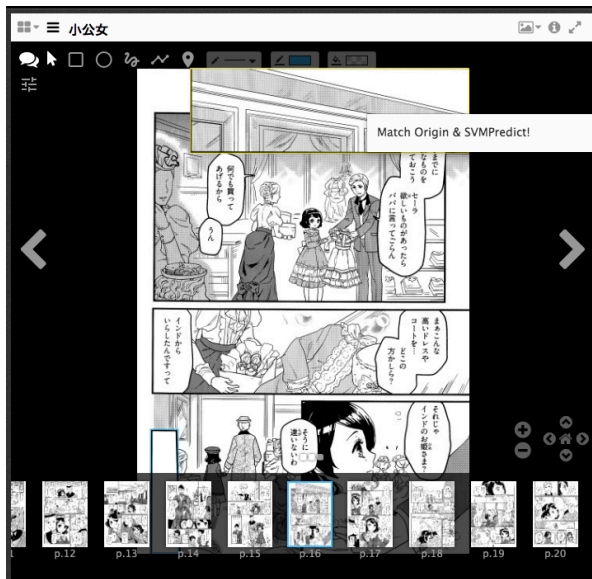


図 5 ビューワによるシーン推測結果の表示

7. おわりに

本研究では、マンガの意味的情報の中でもコマとコマに含まれるオブジェクトに関するメタデータモデルを定義し、それらに基づくコマの属性とシーン転換箇所の関係のメタデータ記述を機械学習に用いるシステムを開発した。これにより、マンガ画像そのものの解析に依らず、マンガの内容に則した機械学習によるシーン抽出を実現した。さらに機械学習による解析結果とマンガ画像の同時提示を実現した。これらによりマンガのシーン抽出の機械学習の検討が効率的に行えるようになったと考えられる。今後は実際の機械学習によるシーン抽出を進めることでその効率化の評価を図る必要がある。

本研究で開発したシステムでは、特徴量の選択や変換の方法が柔軟に選択できず、解析の対象についてもシーン転換箇所だけに留まった。より多彩な機械学習によるマンガの内容解析をするに当たっては、解析対象や特徴量、学習モデルをユーザが柔軟に選択できることは不可欠である。

ビューワについては、解析の対象がシーン転換箇所ではない場合を考慮し、テキストだけでなく、画像や映像などの多様なデータや作品に関連する Web 上の情報資源もアノテーションとして容易に利用できるべきである。

加えて、高精度な機械学習のためには、より多くのマンガ作品が学習に利用できることが望ましい。これには多くのマンガについて細粒度のメタデータが必要であり、その効率的な作成手法が求められる。

参考文献

- [出版科学研究所 16] 出版科学研究所: 特集「紙コミック&電子コミックの最新動向～コミック市場 2015～」, 出版月報 2016 年 2 月号, Vol.58, No.2 (2016).
- [久行 15a] 久行智恵, 三原鉄也, 永森光晴, 杉本重雄: ダイジェスト作成のためのマンガのシーン抽出手法の検討, 情報処理学会 第 77 回全国大会講演論文集, 2015(1), pp. 631-633 (2015).
- [久行 15b] 久行智恵, 三原鉄也, 永森光晴, 杉本重雄: マンガの概念のおよび形態的要素のメタデータに基づくシーン抽出手法の検討, HCG シンポジウム 2015, B-2-2 (2015).

- [久行 16] 久行智恵, 三原鉄也, 永森光晴, 杉本重雄: コマの持つ属性を用いたマンガのシーン自動抽出, 2016 年度 人工知能学会全国大会, 2J4-OS-08a-3 (2016).
- [石井 13] 石井大祐, 渡辺祐: マンガからの自動人物検出と識別に関する一検討, The Journal of the Institute of Image Electronics Engineers of Japan, Vol. 42, No. 4 (2013).
- [Chu 15] Wei-Ta Chu, Chia-Hsiang Yu, Hsin-Han Wang: Optimized Comics-Based Storytelling for Temporal Image Sequences, IEEE Transactions on Multimedia, Vol. 17, No. 2, pp. 201-215 (2015).
- [Matsui 15] Matsui, Y., Ito, K., Aramaki, Y., and Aizawa, T. Y. K.: Sketch-based Manga Retrieval using Manga109 Dataset, CoRR, Vol. abs/1510.04389, pp. 1-13 (2015).
- [Rigaud 15] Christophe Rigaud, Clément Guérin, Dimosthenis Karatzas, Jean-Christophe Burie, Jean-Marc Ogier: Knowledge-driven understanding of images in comic books, International Journal on Document Analysis and Recognition (2015).
- [三原 15] 三原鉄也, 永森光晴, 杉本重雄: マンガメタデータフレームワークに基づくデジタルマンガのアクセスと制作の支援—デジタル環境におけるマンガのメタデータの有効性の考察—, 電子情報通信学会論文誌 A, Vol. J98-A, No.1, pp. 29-40 (2015).
- [ImageAPI 12] IIIF Image API 2.1, <http://iiif.io/api/image/2.1/> (2012).
- [PresentationAPI 12] IIIF Presentation API 2.1, <http://iiif.io/api/presentation/2.1/> (2012).
- [Morozumi 09] A.Morozumi, S.Nomura, M.Nagamori, S.Sugimoto: Metadata Framework for Manga: A Multiparadigm Metadata Description Framework for Digital Comics, Proceedings of DC2009, pp.61-70 (2009).
- [萩原 13] 萩原彰, 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課程の可視化による制作支援, 全国大会講演論文集, 2013(1), pp. 629-631 (2013).
- [落合 12] 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—, 全国大会講演論文集, 2012(1), pp. 593 (2012).
- [Lin 11] Chih-Chung Chang and Chih-Jen Lin: LIBSVM : a library for support vector machines, ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm> (2011).
- [Virtuoso 10] OpenLink Virtuoso - Product Value Proposition Overview, [https://www.openlinksw.com/dataspace/doc/kidehen@openlinksw.com/weblog/kidehen@openlinksw.com/s%20BLOG%20\[127\]/1609](https://www.openlinksw.com/dataspace/doc/kidehen@openlinksw.com/weblog/kidehen@openlinksw.com/s%20BLOG%20[127]/1609) (2010).
- [Mirador 16] Mirador, <http://projectmirador.org/> (2016).