

マンガ制作支援に向けたメタデータモデル

伊藤 博典[†] 浅野 泰仁[†]

[†] 京都大学情報学研究科 〒606-8501 京都府京都市左京区吉田本町 36-1

E-mail: [†]h-ito@db.soc.i.kyoto-u.ac.jp, ^{††}asano@i.kyoto-u.ac.jp

あらまし 本研究では、マンガ構図推薦による制作支援を目標とし、マンガ画像のアノテーションに必要な要素を定義したメタデータモデルを提案する。マンガは画像とテキストの組み合わせによって個々の場面、ストーリーを表現するマルチモーダルな表現物であり、一般にテキストのみの表現よりも読者に強い印象を与えることができる。そのため、教育や広告に利用したいと考える人は多い。一方、その表現には専門知識や暗黙的な技術・知識が用いられており、一般にそうした人々が効果的にこの表現手法を利用するのは難しい。そこで、最終的に脚本からのマンガ構図の自動生成によるマンガ制作支援を目指し、ここではその基盤となる、脚本による言語情報と画像情報の観点から必要要素を定義したメタデータモデルを提案する。

キーワード コミック工学, メタデータモデル, アノテーション, XML

1. はじめに

近年、マンガは子供が物語を楽しむだけのものにとどまらず、幅広い世代に様々な形で楽しまれている。マンガはその表現手法としてのメリットから学習マンガとして教育に用いられている他、広告にも使用されるなど、エンターテインメントの領域以外にも幅広い分野で表現手法の一つとして利用されている。

この表現手法としてのマンガのメリットとしては、次のようなものがある。(1) テキストのみの表現よりも目に止まりやすく、記憶に残りやすい。(2) ストーリーに乗せて情報を伝えるため、感情移入しやすい。(3) テキストのみでは伝わりにくいことを分かりやすく表現できる。

そのため、一般の人々においても教育や広告に利用したいと考える人は多く、近年はコミポ[1]といった絵を書かなくとも、3Dモデルを利用してマンガを描くことができるソフトウェアなども市販されている。しかし、マンガの表現には専門知識や暗黙的な技術・知識が用いられており、一般にそうした人々が効果的にこの表現手法を利用するのは難しい。

このマンガの表現における専門知識や暗黙的な技術・知識の一例としては、場面や状況に応じたパースや人物・吹き出しなどの配置、サイジングなどがある。これらを踏まえずにマンガを制作すると、どのような状況下に登場人物があるのかといったことや何を伝えたいのかといったことが伝わりにくいものとなってしまう、先に挙げたマンガという表現手法によるメリットを活用することができていない。

また、マンガのメタデータモデルに関する研究はいくつか存在するが、表現する場面や状況に関して脚本のような言語的表現とマンガにおける画像的表現とを結びつけられるものは存在しておらず、これら双方の要素を考慮したモデルによって、構図推薦や脚本生成といった様々な用途への応用が考えられる。

そこで本研究では、効果的に内容を伝えるために、脚本のような言語的表現とマンガにおける構図を結びつけて取り出す上で必要な要素を定義するとともに、この脚本から得られる言語

情報とマンガから得られる画像情報を表現するためのメタデータモデルを提案する。

本研究において提案するメタデータモデルは、最終的に脚本からのマンガ構図の自動生成によるマンガ制作支援におけるマンガのアノテーションデータセット作成の基盤として利用することを想定し、アノテーションデータ作成の上で重要となる時間的なコストを踏まえた上で、脚本と画像情報の観点から必要要素を定義している。

本研究において提案するメタデータモデルにおいては、構図を表現する要素として、コマを基準単位として、コマ内における以下の10個の要素を定義した。

(1) 場面タイプ, (2) アングル, (3) 対象距離, (4) 背景, (5) オノマトペ, (6) 台詞, (7) キャラクタ, (8) アイテム, (9) 効果, (10) 領域

また脚本に関する要素としては、場面を基準単位として、場面ごとに以下の7個の要素を定義した。

(1) 場所, (2) 対象コマ, (3) 動作主, (4) 感情, (5) 動作, (6) 動作対象, (7) 発話内容。

提案するメタデータモデルにおいては、以上の各要素ごとにマンガ画像にアノテーションを付与する上で必要な情報を更に細分化して、定義しており、これらの詳細については第3章において説明する。

評価実験は、2種類のマンガから複数のコマを取り出し、このメタデータモデルを用いて、それらのコマ画像にアノテーションを事前に行い、このアノテーション情報のみを被験者に与え、それをもとにコミポを用いてコマを作成してもらい、別の被験者に元の画像と比較して再現できているかどうかを評価してもらうことで行った。

本論文の構成は次のとおりである。まず、第2章では関連研究について述べる。第3章では、提案するメタデータモデルの詳細について説明する。第4章では評価実験について説明し、第5章ではまとめと今後の課題について述べる。

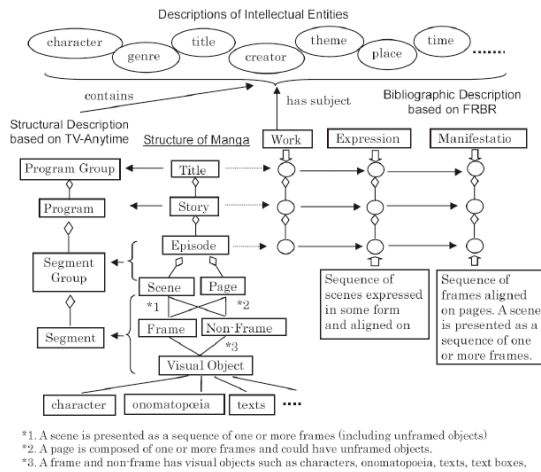


図 1 Morozumi らのマンガメタデータフレームワーク

2. 関連研究

本研究に関連する先行研究がいくつか存在する。それについて以下に述べる。

日本のマンガのメタデータモデルに関する研究としては、Morozumi らの研究 [2] や Fujimoto らの研究 [6] が挙げられる。

Morozumi ら [2] は、ネットワーク環境におけるマンガの発見やアクセス、再利用のためのメタデータフレームワークを提案している。このフレームワークでは、出版された具体物としてのマンガに関する書誌記述、マンガにおける表現やストーリーに関する構造記述、マンガに登場する主要人物や動作主体、コンセプトに関するオントロジー記述の 3 つの観点から議論し、それぞれ「FRBR [3]」、「TV-Anytime [4]」、「知的実体を示す辞書項目 [5]」を基礎として拡張したモデルを提案している。図 1 にこのメタデータモデルを示す。

Morozumi らの手法においては、デジタルマンガを主な対象として、先の 3 つの観点から個別の表現に依らず、内容に関する要素をマンガの構造に従って抽象化された実体として記述するためのメタデータフレームワークを定義している。そのため、書誌情報やキャラクタ、動作、話数、シリーズといった観点からインターネット上におけるマンガ内外における検索やプロによるマンガ制作における再利用に向けたモデルとなっているが、各情報を横断的に結びつけるアノテーションの困難性やそのコストの大きさ、また内容要素に従い抽象化されているため、データとして機械学習への応用が難しい点が問題である。

本研究では、マンガの制作について詳しくないユーザを対象として、脚本からマンガ構図を自動生成することで、マンガの制作支援を行うことを目標としており、内容要素だけでなく、それを効果的に伝えるために必要となるアングルや効果などの表現要素にも着目したメタデータモデルを提案する。また提案するメタデータモデルにおいては、話数やシリーズといった大きな粒度のストーリー構造ではなく、マンガ制作の基となる脚本から会話や動作を用いて、マンガにおける内容や表現を記述する点においても、本研究は Morozumi らの研究と異なる。

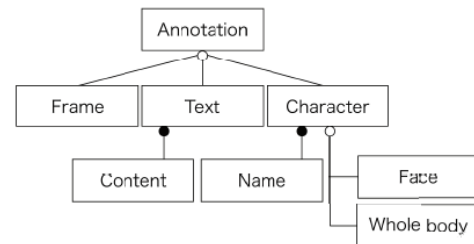


図 2 Fujimoto らのメタデータモデル

Fujimoto ら [6] は、マンガ画像における登場人物や文字の認識、テキストと画像からなるマルチモーダルな表現に対応した検索に関する技術を発展させることを目標として、学術利用可能な大規模で高品質なオープンマンガデータセットを構築するとともに、アノテーションに用いるメタデータモデルについて述べている。

Fujimoto らのメタデータモデル [6] は、長方形による囲いを基準として、コマ、登場人物、テキストの 3 つの観点から定義しており、各要素の存在する画像中の空間座標、登場人物の名前や顔、全体像、台詞の内容のみから成る簡潔なモデルである。図 2 にこのメタデータモデルを示す。

このモデルは、キャラクタや吹き出しといった観点からの画像的特徴抽出やその画像認識など機械学習に向けたモデルであり、そのアノテーションのコストも小さく、データセット作成に向けたものとなっているが、マンガの表現方法における意味内容についての観点がないため、ストーリーの内容やそれを表現する人物や台詞以外の要素が不足している点が、本研究の最終的な目的を達成する上で必要となる構図と脚本の意味内容の関連についての学習において問題となる。

本研究で提案するメタデータモデルでは、マンガにおける表現としての画像情報と脚本における言語資源との関連をあわせて表現するとともに、アイテムや背景などの場面状況を認識するために必要な要素を記述要素として新たに定義している。

また日本のマンガではないが、コミックやグラフィックノベルの構造に関するメタデータ記述規則についての研究としては、Comic Book Markup Language(CBML) [7] が挙げられる。

これはデジタル環境での人文学資料のマークアップのためのガイドラインである Text Encoding Initiative (TEI) [8] を基礎にしたコミックとグラフィックノベルのマークアップのための記述規則である。CBML は画像データとしてデジタル化されたマンガのマークアップを目的として、マンガの表現構造を記述するための語彙を定義し、XML の構造を定義するためのスキーマ言語 RelaxNG で記述されたスキーマを公開している。特に画像中のテキストについて、セリフやキャプションといったテキストの種類、コマとの関係、式などの詳細な定義を行う記述が用意されており、本研究の構図を表現するメタデータモデルの定義において、粒度を決定する際に参照した。

3. 提案手法

まずはじめに、マンガ画像の構図を表現するメタデータモデ

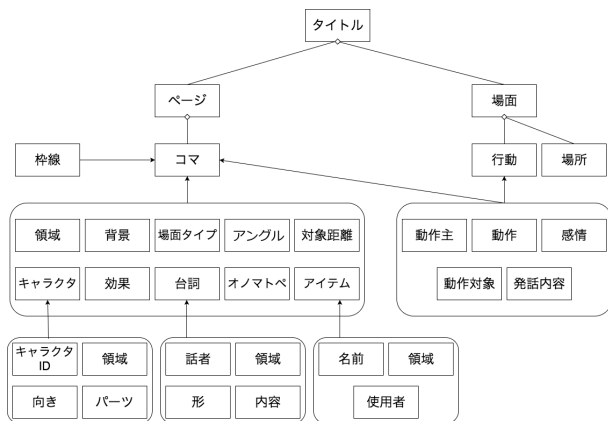


図 3 提案メタデータモデルの全体図

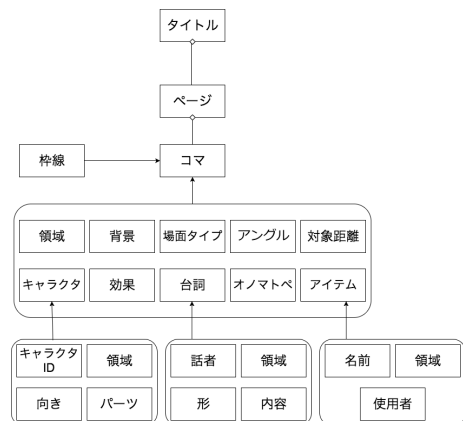


図 4 構図に関するメタデータモデル

ルとその脚本を表現するメタデータモデルの2つを含めた、本論文で提案するメタデータモデルの全体像を図3に示す。

以下、3.1節においてマンガ画像の構図を表現するメタデータモデルの構造と各部の詳細について説明し、3.2節においてマンガ画像の表す内容に応じた脚本を表現するメタデータモデルの構造と各部の詳細について説明を行う。

3.1 構図メタデータモデル

本節では、構図に関するメタデータモデルの構造について述べた後に、モデルにおける各部の詳細ならびにそのガイドラインについて説明する。

3.1.1 構図メタデータモデルの構造

図4に構図に関するメタデータモデルの構造を示す。この図のように、あるマンガにおいて巻数を含めたその「タイトル」を構造の最上位の実体として、「ページ」はそのマンガの各ページを示す実体であり、「コマ」は各ページ内の個別のコマの実体である。

「コマ」は、そのコマの画像中の始点座標とサイズを示す「領域」、その枠線の有無を示す「枠線」、コマ内においてフォーカスされているものを示す「場面タイプ」、コマの描写角度を示す「アングル」、コマ内で描写される中心物に対する距離感を示す「距離」、コマの背景を示す「背景」、コマ内に描かれる登場人物を示す「キャラクター」、人物の感情や雰囲気、動きの表現を付与するために描かれる記号などの種類を示す「効果」、コマ内の吹き出し・発話やト書きを示す「台詞」、「効果」と同様の表現を付与するための擬音語や擬態語を示す「オノマトペ」、コマ内でキャラクターの動作に関わるものを示す「アイテム」の11個の実体によって具体化される。

3.1.2 構図メタデータモデルの詳細とガイドライン

以下において、提案モデルにおいて定義した各部の詳細とそのガイドラインについて説明する。本提案モデルの定義において、先行研究から新たに定義している実体は、主に「コミ Po! マンガ入門」[9]を参考としている。

● コマ

コマに描写されている舞台とその順序は、脚本と構図を結びつける上で基本的な情報であり、一意に割り当てるIDによってこれを表現する。この各コマがもつ情報を先に挙げた11個の

実体により表現する。

● 領域

この実体では、コマの属するページ内における始点座標とそのサイズに関する情報をもつ。ここではコマを、単一のシーンを描写する画像領域として定義する。基本的には枠線を基準とするが、枠線が存在しない場合もコマとして定義される場合がある。このような場合、Fujimotoらの定義に従い、仮想的に枠線を作成し、それをコマの領域として定義する。

● 枠線

この実体では、コマに枠線が存在するかどうかの情報をもつ。この要素は、先の領域の実体の説明において述べた枠線が存在しないコマに対応するためのものである。先行研究においては実体として定義されていないが、構図において枠線の有無は認識の上でも大きく関わる重要な要素であるため、本研究では新たに定義している。

● 背景

この実体では、そのコマ内で描かれている場所の情報をもつ。ここでは前後のコマの場所情報に関わりなく、コマ内に場所の要素が描かれているかのみから判断して、場所の情報を記述する。Morozumiらの研究では、作品を通しての舞台としての場所に関する記述はあるが、各コマごとに場所に関して記述する実体はない。ストーリー上において背景として描かれる場所は、まともりはあれどコマごとに変化するものであり、この場所としての背景の要素は構図に大きく関わる要素であるため、本研究においては各コマごとに実体として定義している。

● 場面タイプ

この実体では、コマ内で表現されている場面を「一般」、「人物」、「風景」の三種類のいずれかに分類する。ここでは、コマ内でキャラクターに焦点を当てて、アップで描いているものを「人物」として、背景に焦点を当てて、引きで描いているものを「風景」として、それ以外を「一般」として分類する。

この要素は本研究において新たに定義したものであり、各コマにおいてキャラクターに焦点を当てているか、風景に焦点を当てているかは、ストーリーと密接に関連する要素であり、脚本と構図を結びつける上で重要な要素であるため、新たに追加している。

- アングル

この実体では、コマ内の背景やキャラクタがどのアングルから描かれているかを示す。ここでは、コマの枠をカメラの画面として、対象物の平面上の角度をアングルとして定義し、この平面の向きをもとに画角変化表の粒度で表現する。

この要素は、構図を考える上で欠かせない概念であり、各コマの場面を効果的に見せる上でも重要なものである。アングルを考慮することで、キャラクタの立ち位置の切り返しなどによる不自然なコマの流れを防ぎ、ストーリーに沿った自然なカメラ位置の流れを演出することができる。

- 対象距離

この実体では、コマ内において描写されているものが「アップ」、「ロング」のどちらで描かれているかの2つに分類する。ここでは、キャラクタの顔や物の一部分を近くから見た構図を「アップ」として分類し、キャラクタや物を遠くから見た引いて描かれた構図を「ロング」として分類する。例えば「うれしい」「楽しい」という表現の場合、キャラクタをアップにすると効果的に見せる事ができ、キャラクタの心理描写を伝えるにはアップとロングショットを使い分けることが効果的であるため、本研究において新たにメタデータモデルに追加している。

- キャラクタ

この実体は、コマ内の登場人物に関する情報を「キャラクタID」、「領域」、「向き」、「パーツ」の4つの実体により具体化する。「キャラクタID」は、あるタイトルのマンガ全体を通して主に登場するキャラクタに対して一意に割り当てたIDによって表現し、主なキャラクタ以外はモブとして表現する。「領域」は、コマ内のキャラクタの位置を始点座標とサイズで表現し、「向き」では、アングルに対するそのキャラクタの体の向きを、正面、後、右、左で表現する。「パーツ」ではそのコマ内に収まっているキャラクタの部位を顔、首、胸、右腕、左腕、胴、右足、左足、全身、上半身からそれぞれ選択して記述する。

ここで「キャラクタID」と「領域」については、マンガでストーリーを表現する上で中核をなす要素であるため、先行研究においても定義されている。本研究では、ストーリーを分かりやすく伝え、キャラクタの心理を効果的に伝えるには、アングルや対象距離が重要であるという観点に立ち、ここからアングルや対象距離に関わりのある要素として、「向き」と「パーツ」もキャラクタを表現する実体として定義している。

- 効果

この実体では、コマ内で用いられている効果線・効果背景の種類についての情報をもつ。ここで効果線・効果背景とは、主にキャラクタの心理状態や動きなどを表現するために用いられる線や点からなる表現を指し、例えば、集中線やスピード線、タレ線などがある。特に効果背景では、場所背景がなく、効果のみが描かれている背景を指す。効果については、Morozumiらのモデルにおいても定義されており、キャラクタの心理状態や動きを表現する上で重要な要素として定義している。

- 台詞

この実体は、コマ内の吹き出しに関する情報を「話者」、「領域」、「形」、「内容」の4つの実体により具体化する。「話者」はその

吹き出しの内容を話す人物を先のキャラクタIDで表現し、「領域」はコマ内での吹き出しの位置を始点座標とサイズで表現する。「形」は吹き出しの形は一般的に用いられる12種類の中から選択し、「内容」ではその吹き出し内の文字を記述する。台詞については、キャラクタ同様にストーリーの進行において欠かすことのできないものとして先行研究においても定義されている。

- オノマトペ

この実体では、コマ内に存在するすべてのオノマトペをその表記によって表現する。オノマトペは擬音語や擬態語をデザインして描いているものであり、場面状況を体感として視覚的に伝えるための重要な演出の一つであり、Morozumiらの研究においても定義されている。

- アイテム

この実体では、コマ内において登場人物と接するもしくは使用する道具の情報について「名前」、「領域」、「使用者」の3つの実体により具体化する。「名前」はその道具の名前を示し、「領域」はコマ内においてその道具の存在する位置を始点座標とサイズで示し、「使用者」はその道具に接するもしくはそれを使用する人物を先のキャラクタIDで表現する。Morozumiらの研究において、キャラクタの持ち物などの小道具・小物はストーリーで重要な意味を持つことがあり、こういった小道具についても記述されるべきとしており、本研究においても同様に定義している。

3.2 脚本メタデータモデル

本節では、脚本部に関するメタデータモデルの構造について述べた後に、モデルにおける各部の詳細ならびにそのガイドラインについて説明する。

3.2.1 脚本メタデータモデルの構造

図5に脚本に関するメタデータモデルの構造を示す。この構造の最上位の実体は、構図メタデータモデルと同様に、あるマンガにおいて巻数を含めたその「タイトル」であるが、「場面」はそのマンガにおける場所の切り替わりにより定義される実体である。この実体は、場面内で行われる各行動を示す「行動」とその行動が行われる場所を示す「場所」の2つの実体によって具体化される。

「行動」は、場面中において動作を行った主体について示す「動作主」、場面中において行われた行動を示す「動作」、その動作主の感情を示す「感情」、動作が効果を及ぼす対象を示す「動作対象」、動作が発話やナレーションであった場合にその内容を示す「発話内容」の5つの実体によって具体化され、これらの5つの実体で具体化された「行動」ごとにコマの情報と結びつけられる。

3.2.2 脚本メタデータモデルの詳細とガイドライン

マンガ制作における脚本の作成方法は、各作者に応じて様々であり、一般的な定義が存在するわけではないが、ここでは「マンガ原作の書き方」[10]の定義を基にしたマンガの脚本の表現を行うメタデータモデルを提案する。

以下において、「マンガ原作の書き方」を踏まえた上で定義したモデルの各部の詳細とそのガイドラインについて説明する。本研究においては、ストーリーを表現する上では、「どこで誰がど

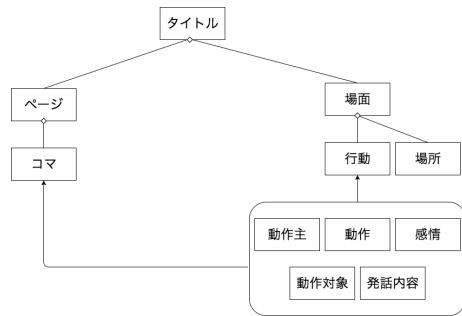


図 5 脚本メタデータモデル

んな感情で何にどんなことをするか」といったことが基本であるという観点から以下の実体によって脚本を定義している。

- 場面

場面の定義については様々なものがあるが、本研究においては、マンガ内で場所の切り替わるコマを場面の切り替わりポイントとして定義し、同じ場所で行われる一連の行動を一つ場面とする。

- 場所

この実体では、一連の動作が行われている場所について記述する。

- 行動

この実体は、アイテムの動作とキャラクタの動きの詳細に関する情報を以下の5つの実体により具体化される。

- 動作主

この実体では、ある動作や動きの主体に関する情報を持ち、キャラクタ ID もしくはアイテム名により表現される。

- 動作

この実体では、動作主が行う動作について動詞で表現する。ここで使用される動詞は、日本語 WordNet [11] の定義を利用する。

- 感情

この実体では、動作主の感情を Ekman ら [12] の提唱する「怒り」、「嫌悪」、「恐れ」、「幸福感」、「悲しみ」、「驚き」、「面白さ」、「恥」などを含む 17 種類の感情の分類に、「中立」を加えた 18 種類を用いて表現する。

- 動作対象

この実体では、動作が作用する対象について表現する。ここでは、対象とそれが係る助詞まで含めて記述し、例えば「A に話す」といった内容の場合、「A に」の部分を動作対象として記述する。

- 発話内容

この実体では、動作が発話や思考、ナレーションに類するものであった場合に、その内容を記述する。

4. 評価実験

本研究において提案するメタデータモデルが、マンガを作成するにあたりその構図を十分に表現できているかどうかについて評価するためにユーザ実験を行った。

4.1 実験概要

実験は、はじめに「坊っちゃん」[13] と「銀河鉄道の夜」[14]

```
<frame>
<ID>4</ID>
<Area></Area>
<Framing>
<SceneType>人物</SceneType>
<Angle>正面</Angle>
<Range>アップ</Range>
<Back>
<Onomatope>
<Effect>困惑_02, 怒り騒音_001</Effect>
<Characters>
<Character>
<ID>A</ID>
<Area>(710,0), 480x880</Area>
<Part>上半身</Part>
<Direction>左向き</Direction>
<Character>
<ID>B</ID>
<Area>(190,0), 500x880</Area>
<Part>上半身</Part>
<Direction>右向き</Direction>
<Character>
<Characters>
<Balloons>
<ID>1</ID>
<Area>(1070, 290), 250x350</Area>
<Content>ずいぶん遅かったな！</Content>
<Type>01-2</Type>
<Speaker>A</Speaker>
```

図 6 コマ ID4 の構図側のアノテーションの一部

```
<Scene>
<ID>1</ID>
<Place>職員室</Place>
<Sequences>
<Sequence>
<ID>1</ID>
<Frame_ID>4</Frame_ID>
<Agent>A</Agent>
<Emotion>怒り</Emotion>
<Action>話す</Action>
<Target>Bに</Target>
<Content>ずいぶん遅かったな！</Content>
<Sequence>
<ID>2</ID>
<Frame_ID>4</Frame_ID>
<Agent>B</Agent>
<Emotion>怒り</Emotion>
<Action>話す</Action>
<Target>Aに</Target>
<Content>ああ君のせいでな/罰金を出したまえ</Content>
<Sequence>
</Sequences>
</Scene>
```

図 7 コマ ID4 の脚本側のアノテーションの一部

の 2 種類のマンガからそれぞれ 5 コマを取り出し、この計 10 コマに対して、提案メタデータモデルをもとにアノテーションを行い、このアノテーションのみをもとに 5 名の被験者（マンガ制作経験者 2 名、未経験者 3 名）にコミ Po! [1] を用いて、各コマを再現してもらった。今回使用したアノテーションの一例を図 6 と 7 に示す。次にこれを別の被験者 3 名により、アノテーション元のコマと比較し、構図上の類似度を 3 段階で評価してもらった。また、合わせてアノテーションをもとにマンガのコマを作成してもらった被験者それぞれに対して、本アノテーションを作成に用いる上で困難だった点などに関するインタビューを行った。

4.2 実験結果

表 1 と表 2 に評価結果を示す。また、制作者 ID1 と制作者 ID3 の評価値の高いコマと低いコマの比較を図 8 と図 9 に示し、コマ ID6 とコマ ID8 の評価値の高いコマと低いコマの比較を図 10 と図 11 に示す。ここで、制作者 ID3、ID4 の制作者がマンガ制作経験者である。そして評価者には、制作者に関する情報については知らせずに、以下の基準で評価してもらった。

- 評価値 3: 元のコマと制作されたコマの各要素がほぼ一致している
- 評価値 2: 制作されたコマは元のコマの各要素と一部異なっているが、不自然でない範囲（実用に耐えるレベル）
- 評価値 1: 制作されたコマが元のコマの各要素と大きく異

表 1 制作者ごとの評価値平均

制作者 ID	1	2	3	4	5	全体平均
評価値平均	1.73	1.83	2.03	1.73	1.97	1.86

表 2 コマごとの評価値平均

コマ ID	1	2	3	4	5
評価値平均	1.93	2.07	2.13	1.80	1.73
コマ ID	6	7	8	9	10
評価値平均	2.13	1.67	1.60	1.87	1.80



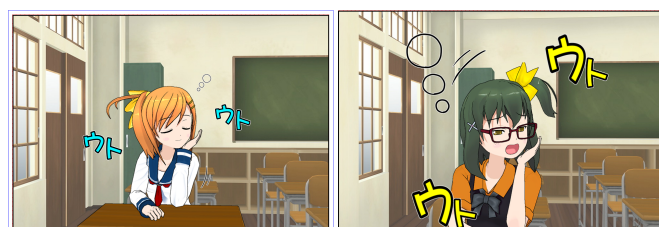
[1] 評価値の高かったコマ (ID1) [2] 評価値の低かったコマ (ID10)

図 8 制作者 ID1 が制作した評価値の高いコマと低いコマの比較



[1] 評価値の高かったコマ (ID1) [2] 評価値の低かったコマ (ID4)

図 9 制作者 ID3 が制作した評価値の高いコマと低いコマの比較



[1] 評価値の高かったコマ [2] 評価値の低かったコマ

図 10 コマ ID6 において評価値の高いコマと低いコマの比較



[1] 評価値の高かったコマ [2] 評価値の低かったコマ

図 11 コマ ID8 において評価値の高いコマと低いコマの比較

なっている、もしくは不自然なものになっている

今回の評価実験の結果として、実用に耐えるラインである評価値 2 に評価値平均がわずかに及ばなかった。この原因としては、主に以下の 3 つの問題が考えられる。

- 1). 提案メタデータモデルにおける問題
- 2). 制作者に対する説明の問題
- 3). 評価における問題

一つ目に挙げた提案メタデータモデルにおける問題として、効果やオノマトベに関するアノテーションが不足している点がある。今回提案するモデルではアノテーションコスト削減のため、オノマトベや効果については種類や表記のみの記述に留め、その位置やサイズについては記述していない。

そのため、アノテーションの読み手である制作者にその置き方やサイズを委ねているが、その結果として図 8[2] のように不自然な効果の配置になったり、図 10 のように制作者によってオノマトベと効果の配置やサイジングが大きく異なり、コマの与える印象が大きく変わり、不自然に感じるものも見受けられた。

ここから、オノマトベや効果の比重の大きいコマ ID1 や ID6, ID10 などの評価値が低くなったと考えられる。また、オノマトベや効果の配置とサイジングによってコマが不自然なものになっているケースは、特にマンガ制作未経験者らが制作したコマにおいてよく見受けられた。本メタデータモデルの最終的な応用先としてマンガ制作未経験者に対するマンガ制作支援があり、また今回の評価の傾向から構図の上でオノマトベや効果の配置が大きな比重を占めていることが見て取れたことから、オノマトベや効果の配置やサイズに関するアノテーションの記述について再考する必要があると考える。

二つ目に挙げた制作者に対する説明の問題として、メタデータモデル自体に対するいくつかの点における説明不足とマンガの表現手法である効果背景についての説明が不足していた点が挙げられる。

メタデータモデルにおいて説明が不足していた点の一つは、アングルとキャラクタの向きに関する説明である。本提案モデルにおいては、アングルに対するベクトルからキャラクタの向きを定義しているが、この説明が不足していたため、キャラクタの体の向きが不自然なものがあり、評価が下がる一因となったと考える。

また、もう一つ説明が不足していた点としては、構図アノテーションと脚本アノテーションの優先度に関する説明である。今回、メタデータモデルの表現力を評価するにあたって構図アノテーションを優先してもらった必要があったが、この説明が不足していたため、脚本の場面描写をもとにアレンジを加えていたことも評価が下がる一因となっており、これが制作者 ID3, ID4 のマンガ制作経験者間の評価値平均において差が生じた理由であると考えられる。

また、表現手法としての効果背景についての説明不足と先のアノテーションにおける優先順位に関する説明が不足していたこともあり、効果背景を用いずに、脚本側のアノテーションを優先し、場所背景を用いていたものがあっても評価値平均を下げた要因であると考えられる。この傾向は、特に特にマンガ制作未経験者らが制作したコマにおいてよく見受けられ、ここから効果背景を用いていたコマ ID2 や ID4 の評価値が低くなったと考えられる。

そして 3 つ目にあげた評価における問題としては、比較する

上での着目する点に関する説明が行き届いていなかった点があり、先の2つ目に挙げた説明不足に関する問題も含め、再度実験を行う必要があると考える。

また、制作者に対するインタビューから、先に挙げたオノマトベや効果の配置、アングルに対するキャラクタの向きの観点以外に、吹き出しにおけるフォントタイプの指定や、脚本側のアノテーションにおいて動作や場所以外に相談シーンや喧嘩シーンといった場面の特徴をより明確に示すものを追加することで、より元のコマに近いものを再現しやすいアノテーションを作成することができると思う。

5. まとめと今後の課題

本論文では、マンガ表現において意味内容を伝えるための画像情報と意味内容としての言語情報をともに表現することを目的とし、マンガのコマにおける各要素とそのコマに対する脚本的要素を定義したマンガメタデータモデルの提案を行った。

この提案メタデータモデルをもとに、複数のコマのアノテーションを作成し、このアノテーションのみをもとに作成してもらったマンガのコマとアノテーション元のコマと比較してもらうことで評価実験を行った。この評価結果から、実用に耐えるラインに表現力がわずかに及ばなかった。この結果から、効果やオノマトベに対するアノテーションなどの観点から、アノテーションコストも踏まえて取捨選択する必要があるが、メタデータモデルの改善を行う必要があり、また実験における不備もいくつか存在したため、改善モデルを用いた再実験を行う必要があると考える。

今後の課題としては、アノテーションデータ作成の上で重要となる時間的なコストを踏まえた上で、効果やオノマトベ、フォントなどに対するアノテーションの改善を行っていく必要がある他、クラウドソーシングなどを用いて、大規模なアノテーションデータセットを作成するために、複雑なアノテーションタスクをいかに分解して、素人でもできるようにするかを考えていく必要がある。

今後は本論文で提案したモデルを改善するとともに、このモデルを元にしたデータセットを作成することで、画像からアノテーションを生成する機械学習の可能性やマンガ制作支援のための構図アノテーション自動生成の可能性についての検討も行う必要がある。

謝 辞

本研究は JSPS 科研費 15K00423 および栢森情報科学振興財団の助成を受けたものです。ここで心より感謝申し上げます。

文 献

- [1] “コミ Po!” <https://www.comipo.com/>
- [2] Ayako Morozumi et al. “Metadata Framework for Manga: A Multi-paradigm Metadata Discription Framework for Digital Comics.” Proceedings of DC-2009, pp. 61–70, 2009.
- [3] “Functional Requirements for Bibliographic Records Final Report.” International Federation of Library Associations

- and Institutions, 1998. <http://www.ifla.org/VII/s13/frbr/frbr.pdf>
- [4] TV-Anytime Forum. “Welcome to the TV-Anytime website.” <http://www.tv-anytime.org/>
- [5] 両角彩子, 永森光晴, 杉本重雄. ストーリーの知的内容を表すメタデータ記述項目の提案:Wikipedia 上のマンガ・小説作品記事を対象として. 情報処理学会研究報告, Vol. 2008-FI-92, No. 105, pp. 1-14, 2008.
- [6] Azuma Fujimoto et al. “Manga109 Dataset and Creation of Metadata.” Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding(MAMPU'16), Article No. 2, 2016.
- [7] J.A. Walsh. “Comic book markup language: An introduction and rationale.” Digital Humanities Quarterly, Vol.6, No.1, 2012
- [8] TEI Consortium. “TEI: Text Encoding Initiative”, <http://www.tei-c.org/index.xml>
- [9] 田中 圭一, 竹熊 健太郎. コミ Po! マンガ入門. 太田出版. 2011.
- [10] 大石 賢一. マンガ原作の書き方. 言視社. 2012.
- [11] 国立研究開発法人情報通信研究機構 (NICT). “日本語 WordNet”, <http://compling.hss.ntu.edu.sg/wnja/index.ja.html>
- [12] P. エクマン, W. フリーセン. 表情分析入門: 表情に隠された意味をさぐる第7版. 工藤力訳, 誠信書房, 2000.
- [13] 夏目 漱石, Team バンミカス. 坊っちゃん. イースト・プレス, 2016.
- [14] 宮沢 賢治, パラエティ・アートワークス. 銀河鉄道の夜. イースト・プレス, 2007.