

登場人物の行為に着目したシーンの言語化に関する一検討: 少女漫画を対象として

井上 須美^{†,a} 朴 炳宣[‡] 安尾 萌[‡] 松下 光範^{†,b}

[†] 関西大学 総合情報学部 [‡] 関西大学大学院 総合情報学研究科

a) k776512@kansai-u.ac.jp b) mat@res.kutc.kansai-u.ac.jp

概要 本研究の目的は、「告白」や「嫉妬」のようにコミック内の特定の内容を含むシーンを検索可能にすることである。このような検索を可能にするには、予めコミックの内容を解析してシーンに分割し、各シーンを言語表現に対応付けて検索可能にする必要がある。その実現にあたっては、シーンの分割や言語化、およびそれらの要約など、様々な処理が必要になる。本稿ではその実現のための基礎検討として、シーンの言語化に関する調査を行う。具体的には、時間的に連続するコマを単一シーンとみなして分割し、各シーンに登場する人物の行為を「誰が」「誰に対して」「何をする」の記述項目に沿って実験協力者に言語化してもらう。得られた言語表現のバリエーションや特徴の観察を通じて、シーン検索の実現可能性について検討する。

キーワード シーン検索, シーン分割, 言語化

1 はじめに

電子書籍の市場規模は年々拡大している。その主たる要因は電子コミックの増加であり、市場の約8割を占めている[4]。コンテンツの増加に伴って、それを利用するユーザの興味も多様化している。そのため、多様なコミックの中から自らの嗜好や興味に合ったコンテンツを発見する検索手法が求められている。

従来の紙媒体コミックの場合、内容を検索して特定の目的や関心に沿ったコミックを検索することは難しく、表紙やタイトル・著者などの書誌情報を手がかりとした作品検索の方法が主流であった。そのため、特定のシーンに対する検索要求(例:「主人公が告白するところが見たい」や「ライバルに嫉妬している場面を探したい」)やキャラクターの属性に対する検索要求(例:「ツンデレ女子高生の出てくるマンガが読みたい」や「一途な男の子の話が読みたい」)に基づいて作品を探すことは容易ではない。コミックの内容情報に基づく検索が実現することは、従来の書誌情報に基づく検索では不可能だった情報要求に応えることが可能になる。

本研究ではコミックのシーンに着目し、特定のシーンに基づくコミック作品検索システムの実現を目指す。そのためには、検索の元となる、シーンに対するメタ情報データ(例:「ヒロインが意中の相手に告白」や「主人公がライバルに嫉妬」)が必要である。しかし、現在コミック作品の各シーンを表す言語データセットの作成手法は確立されておらず、特定のシーンに基づく作品検索システムは実現困難である。

本稿では、コミックのシーンに対する言語データセッ

ト作成のための、シーンの言語化に関する調査を行い、特定のシーンに基づく作品検索を可能にすることを目指す。言語化における条件や、制約による表記のばらつきを調査し、シーンに基づいたコミック作品検索の実現可能性について検討する。

2 関連研究

2.1 シーン検索に関する研究

島脇らは、サッカーの試合の中継映像において、ゴールシーン等の特定のシーンを検索可能にするシステム構築のための、シーンの自動解析手法を提案した[6]。スポーツのテレビ中継は次々とショットが切り替わるため、映像から特定のショットを画像として切り出すことが可能である。切り出した画像を解析することによってショット(シーン)の切り替わりを解析する。提案手法では、中継映像から特定のカメラ映像を切り出すために、16階調の色ヒストグラムを用いて画像中のグラウンドと選手のユニフォームの色を自動登録し、ショットの切り替わりを認識した。コートを囲むラインの見え方から中央にあるセンターカメラショットであるかどうかを検証し、直線の交点の世界座標と画像座標との対応関係からカメラ姿勢を推定した。評価実験の結果、93.6%という高い値でセンターカメラからの映像であると正しく認識された。

またテキストを用いたシーンの解析に関する研究として、三浦らのドラマのシーン検索に関する研究がある[5]。三浦らは、ドラマ映像のシナリオを解析し、シーンに対してテキスト情報を用いたインデクスを作成することにより、ユーザが所望するシーンを検索可能にするシステムを試作した。ドラマの脚本における、シナリオを

構成する「シーン名」「ト書き」「セリフとその人物名」の3つの要素を自然言語処理によって解析し、「シナリオ解析システム」という、人物等の行動を構造的に記述したデータベースを作成した。具体的には、「ト書き」からは人物の存在や行動記述といったシーンを記述する情報をインデクス情報として抽出し、「セリフ」からは、感性語を抽出し、決定木法によりシーンの雰囲気や感性インデクス情報として抽出した。評価実験において、検索被験者の主観的評価による正解とシステムが結果として出す正解の適合率は高く、検索システムの有用性を示した。適合しなかった場合においては、機械学習が導き出す一般性と被験者の主観の相違により精度が低くなっていたと筆者は考察した。今後の課題として、システム使用者の主観に沿うようなカスタマイズ性が必要であると述べた。

2.2 本研究の位置付け

今回対象とするコミックは絵とテキストによる複合的なコンテンツであるため、先行研究で取り上げた映像やテキスト情報とは違い、計算機による物語の解析は難しいと考えられる。映像のように画像情報である絵に着目することは、作者によって絵の作風が大きく異なり、描かれた人物や物体の認識が難しい点と、コマによってストーリーが生成されることから、計算機を用いた絵による内容解析は難しい。コマ内に記載されるテキスト情報においても、小説やドラマの台本とは違い登場人物の言動やナレーションによって内容の全てを表しているわけではなく、絵のみで状況を表すコマが存在する。よってテキストによる機械的な内容解析も難しい。これらの理由から、計算機によるコミックのシーン解析は難しく、検索のための言語化は困難である。そこで本研究では、人手によるシーンの言語化手法の検討を行う。各シーンにおいて、登場人物の行為に着目し物語の内容を把握することによって、説明可能な形で言語化を行う。これにより、シーン検索に用いる言語データセットの作成を試みる。

3 提案手法

特定のシーンを自然言語で検索可能にするために、各シーンをメタ情報となる言語データに対応づける必要がある。本章では、まず作品の物語をシーン単位に分割し、各シーンの言語化を行うための手法について記述する。

3.1 作品のシーン分割

要約の前段階として、作品の物語を人手で複数のシーンに分割する。久行らの研究を参考に [8]、同一の時間の連続するコマを単一のシーンと定義する。久行らとの違いは、同一の場所を定義から除外した点である。コミックでは背景の描写がなされていない場合があり、

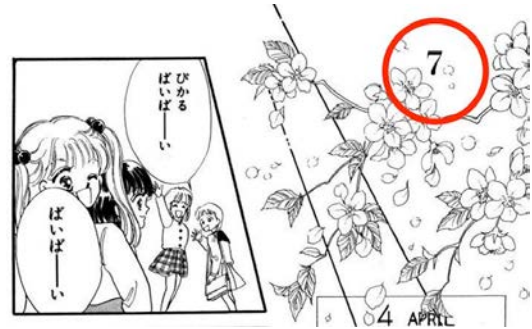


図1 シーン転換箇所の例 (作品:「びびる☆元気です!」¹⁾)

場所の判別が付かないコマが多く存在するためである。具体的なシーンの転換箇所として、3つのタイプのコマを設定した。久行らの研究 [7] で提示された、1) ナレーションがある、2) セリフがない、3) 主要キャラクターがないの3タイプのコマに加えて、“時間の連続するコマを単一のシーン”として定義したことから、4) 昼夜が変化したと見られるコマについてもシーンの転換箇所とした(例:図1)。これら4タイプのコマをシーンの転換箇所として、シーン分割を行った。

3.2 シーンの要約

次に、記述項目に沿って登場人物の行為の要約を行い、各シーンの言語化を行った。

要約の手順は、作品を読み、各シーンに対して、どの登場人物が何を行なったシーンであるかを人手で記述する。行為の表記は、“物語は人の行為によって進むものである”という、プロットの物語に対する定義に基づいている [3]。

記述のルールとして、「誰が」「誰に対して」「何のイベント(シチュエーション)の時に」「何をする」の4項目で構成される文章を、1シーンあたり1文で記述してもらう。4項目のうち、「誰に対して」「何のイベント(シチュエーション)の時に」の2項目に関しては、当てはまる情報が作品に記載されていなければ表記無しでも可とした。

「誰が」「誰に対して」の項目は、それぞれ行為の主體と対象者を表す登場人物名が該当する。この項目に関して、作品の始まりのシーンなど、登場人物の名前が判明していないシーンに関しては、作品を読んだ時点で判明している情報(「主人公」「ライバル」等)を用いて表記を行うこととした。「何をする」の項目は、登場人物の行為を表す動詞が該当する。この項目に当てはめる動詞は1つのみとした。要約にあたっては、データに一般性を持たせるために複数人で行い、その中からシーンを表す要約文を決定することとした。

¹©Shoko Kuriki, Shogakukan Ltd.

表1 データセット作成に用いた作品

タイトル	作者	出版社
ぴかる☆元気です!	栗城祥子	小学館
どきどき DO!れみ	うえだ未知	小学館
だって・思春期	草薙竜樹	実業之日本社
ケーキはいかが!	栗原まもる	A-WAGON
あすか BOY	栗城祥子	ビーグリー
信じる者は救われる!	片山こずえ	オフィス漫
チャーミーぷっぷ ダビデにお願い!	高瀬直子	オフィス漫

3.3 シーンを表す言語データの決定

最後に、要約文同士を比較し、他の要約文と類似度が最も高い要約文をシーンを表す言語データとして一意に決定した。

まず、各要約文の形態素解析を行う。形態素解析には MeCab(バージョン 0.996)[1] を用いた。次に、各要約文に含まれる単語を、Word2Vec[2] を用いて分散表現に置き換える。各文における単語の特徴ベクトル平均を算出し、それを文自身の特徴ベクトルとする。そして、(1)式を用いて各文章間の \cos 類似度を算出する。 \cos 類似度は二つの n 次元のベクトル間の距離を測る際に用いられ、値が1に近いほど二つのデータが似通っていることを表す。そのため、本手法での \cos 類似度は、文章間の意味的な類似度を表す。

$$\cos(\vec{q}, \vec{d}) = \frac{\vec{q} \cdot \vec{d}}{|\vec{q}| |\vec{d}|} \quad (1)$$

各文において、他の文章との \cos 類似度を合計して比較し、最も値の高かった文章を「シーンを表す言語データ」として決定する。

4 データセットの作成

4.1 人手によるシーンの言語化

データセット作成にあたって対象とした作品は、1990年代に出版された少女漫画うち、恋愛が物語の中心テーマである学園系の作品を選出した。表1に選出した作品の一覧を示す。7作品全てにおいて、完結までの全巻を対象とし、1作品につき5人にシーンの言語化を行ってもらった。言語化の対象となるデータの作成にあたって、あらかじめ3.1節で示したシーン分割を行い、言語化の対象となるコミックデータを作成した。

言語化データセット作成にあたって、まず参加者に対する説明を行った。参加者は、18歳-22歳の大学生男女23名である。説明後、コミックデータを読み進めていき、3.2節で示した手法に倣ってテキストファイルにシーンの記述を行ってもらった。データセット作成時に

表2 「ぴかる☆元気です!」第7シーンに対する要約文

A	ひかるが智と大地は兄弟だと知る
B	ひかるがサッカーボールから小学生をかばって怪我をする
C	ひかるが男の子と再会する。
D	ひかるが智と再会する
E	ひかるが智と小学校で再会する。

は他の参加者とのコミュニケーションは取らず、個別で行ってもらった。

4.2 シーンに対する一意的な言語化データの決定

参加者によるデータセット作成後、3.2節の提案手法に倣って、各シーンにつき5つの要約文の中からシーンを表す言語化データを一意に決定した。

例として、「ぴかる☆元気です!」第7シーン(p.22-31)を取り上げる。対象シーンに対し、参加者A-Eの5人が言語化を行った。結果は表2の通りである。

これら5つの要約文を、3.3節で示した手法に基づいて分散表現に置き換え、 \cos 類似度を算出した結果、E:「ひかるが智と小学校で再会する。」が他の要約文との類似度が一番高いため、第7シーンを表す言語化データとして決定した。同じ処理を7作品317シーンに対し行い、シーンの言語化データセットを作成した。

5 考察

本稿では、少女漫画を題材とし、コミックの各シーンに対する言語化の実現可能性の検討を行った。提案手法を用いたデータセットの作成を複数人で行った結果、シーンによって要約文の表記にばらつきが見受けられた。一例として、「ぴかる☆元気です!」における要約文を下記の表に示している。表3は第10シーン(pp.40-42)、表4は第8シーン(pp.32-34)に対する要約文の一覧である。また、表5、表6はそれぞれ第10シーン、第8シーンに対する要約文(表3、表4参照)間で算出した \cos 類似度とその合計を示している。

第10シーンにおける、参加者AとD、BとEの要約

表 3 表記のばらつきが少ない例

A	ひかるが智に教科書を貸す
B	ひかるが智に国語の教科書を貸す
C	ひかるが智くんに教科書を貸す
D	ひかるが智に教科書を貸す
E	ひかるが智に国語の教科書を貸す

表 4 表記のばらつきが多い例

A	ひかるが女の子に憎まれる
B	ひかるが女子から陰で智くんに助けられたことを言われる
C	ひかるが大地に無視される
D	大地がひかるに対して不愛想になる
E	大地がひかるのあいさつを無視する

文の cos 類似度はそれぞれ 1 を示した (表 5). cos 類似度は 1 に近いほど二つのデータが似通っていることを示すため, 1 という値は, 2 つの文が全く同じ文章であることを示す. 参加者 A と D が「ひかるが智に教科書を貸す」, B と E が「ひかるが智に国語の教科書を貸す」と全くの同文を記述していたため, 正確に同文であると結果が出力された. また, 表 5 は, 全ての要約文間の類似度が 0.90 台と非常に高い値を示している. これは, 「誰が」の項目に「ひかるが」, 「誰に対して」は「智」, 「智くん」, 「何をする」は「教科書を貸す」, 「国語の教科書を貸す」と, 文章間で異なる人物や行為についての記述はなく, 表記揺れ程度の差異に留まったためだと考えられる.

対して第 8 シーン (表 4) では, 各要約文間の cos 類似度は 0.75-0.87 と, 第 10 シーンと比べると低い値を示した (表 6). これは, 記述項目「何をする」において, A:憎まれる, B:言われる, C:無視される, D:無愛想になる, E:無視する, のように, 参加者間で表記にばらつきが生じたためであると考えられる.

6 おわりに

本稿では, 特定のシーンに基づくコミック作品の検索を目指した, シーンの言語化手法の提案を行った. 具体的には, 登場人物の行為に着目し, 複数人で各シーンの要約を行い, 複数の要約文の中からシーンを表す文を一意的に決定し言語化データセットの作成を行った.

今後の課題として, 提案したシーンの言語化手法の普遍性を高めるために作品数の追加が必要である. 複数ジャンルの作品でシーンの言語化を行い, 提案した手法における記述の制約でシーンを表す要約文が作成できるかを検証する.

表 5 要約文間の cos 類似度と合計

	A	B	C	D	E	計
A		0.96	0.97	1.00	0.96	3.89
B	0.96		0.93	0.96	1.00	3.85
C	0.97	0.93		0.97	0.93	3.81
D	1.00	0.96	0.97		0.96	3.89
E	0.96	1.00	0.93	0.96		3.85

表 6 要約文間の cos 類似度と合計

	A	B	C	D	E	計
A		0.78	0.75	0.77	0.76	3.07
B	0.78		0.81	0.76	0.78	3.14
C	0.75	0.81		0.82	0.87	3.27
D	0.77	0.76	0.82		0.81	3.17
E	0.76	0.78	0.87	0.81		3.23

参考文献

- [1] Kudo, T., Yamamoto, K. and Matsumoto, Y.: Applying Conditional Random Fields to Japanese Morphological Analysis, *Proc. the 2004 Conference on Empirical Methods in Natural Language Processing*, pp. 230–237 (2004).
- [2] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. and Dean, J.: Distributed Representations of Words and Phrases and Their Compositionality, *Advances in Neural Information Processing Systems*, pp. 3111–3119 (2013).
- [3] Propp, V. R.: 昔話の形態学 (訳: 北岡 誠司, 福田 美智代), 水声社 (1987).
- [4] インプレス総合研究所: 電子書籍ビジネス調査報告書, インプレス (2019).
- [5] 三浦健仁, 中川裕志: シナリオを用いたドラマのシーン検索システム, *情報処理学会研究報告*, Vol. 1998, No. 48, pp. 55–62 (1998).
- [6] 島脇巧, 三浦純, 白井良明: シーン検索システムのための長時間サッカー中継映像の解析, *情報処理学会研究報告*, Vol. 2004, No. 40, pp. 125–132 (2004).
- [7] 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*, 第 1 分冊, pp. 631–632 (2015).
- [8] 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, *人工知能学会全国大会論文集*, 2J4OS08a3 (2016).