

評価表現に着目した映画レビューからの評価情報抽出

紀本 雅大^{1,a)} 伊藤 淳子² 宗森 純²

概要: 映画レビューサイトでは、映画に対する評論、批評、感想や考察が映画レビュー文章として投稿され、公開されている。公開された映画レビューは映画の内容や評価の情報を探すために利用されることがあるが、得られる情報と要する時間はそのレビュー文章に強く依存している。本研究の目的は、映画レビュー文章から映画の内容や評価の情報を抽出し、利用者に提示することである。物事を評価する語（評価語）と評価語によって評価されている対象（評価属性）の組で表される文章を評価表現とし、評価語の判定には評価表現辞書を用いる。実験では抽出された評価表現を利用者に提示し、特定の情報について知ることができたか、映画の面白さを予想する際に参考になったかをアンケートにより評価した。評価の結果、評価表現を抽出し提示することは映画の評価を知る際に有用であり、映画の面白さを予想する際に参考にされていることが示された。

1. はじめに

インターネットの普及にともない、一般人の情報発信の機会が増加している。その中でも、物事に対する評論、批評や感想が、ブログ、マイクロブログ、レビューサイトなどのような、情報発信が可能なウェブサイトにおいて、レビュー文章として発信される。発信されたレビューはそれぞれのウェブサイトで公開され、他のサイト利用者もこれらのレビューを閲覧できるようになる。

映画.com^{*1}、Yahoo!映画^{*2} やみんなのシネマレビュー^{*3} に代表される映画レビューサイトでは、映画に対する評論、批評、感想や考察が投稿され、公開されている。映画レビューの利用目的としては、これから見る映画を選択するための情報収集、これまでに見た映画の評価の共有や閲覧、見ていない映画の内容を把握するための情報収集が挙げられる。

このように映画レビューは映画の内容や評価の情報を探すために利用されることがある。ここで、評価とは「○○が良かった」「○○が悪かった」のような映画に対する個々の感想である。また、内容とは「○○が○○した」のように映画の登場人物の行動の説明などを指す。

しかし、これらの映画レビューから得られる情報の品質と、目を通すのに要する時間はレビュー文章に強く依存し

ている。例えば、参考になる情報が含まれないレビューや、攻撃的な文章、読みにくい文章で書かれたレビューの場合、かかった時間に対し得られる情報が少ない可能性が高い。利用者は複数投稿されたレビューを読むことで有用な情報を探そうとするが、レビューの数が多い場合、利用者の負担が大きくなると考えられる。

そこで本研究では、映画レビュー文章から映画の内容や評価といった情報を抽出し、利用者に提示する手法を提案する。映画レビュー中に含まれる内容や評価の情報を、膨大なレビュー文章を読まずとも得られるようにすることを目的とする。

2. 関連研究

2.1 レビュー文章に関する研究

映画や小説などのストーリーの要素を持った作品に対するレビューには、作品の評価や感想以外に、作品の内容やあらすじについての文章が含まれている。岩井らは、レビュー文章の文書構造と文脈情報を利用して、作品の詳細や展開に言及している箇所を検出し非表示にすることで、利用者が作品を見る前にレビュー文章からストーリーの内容について知ることを防止している [1]。これらの研究ではレビュー文章に含まれる作品の内容を検出していたが、評価については扱っていない。本研究では作品の評価についても抽出する。

2.2 評価表現に関する研究

物事を評価する表現（評価表現）を文章中から抽出する研究は多く行われている。評価表現は、何らかの評価がな

¹ 和歌山大学大学院システム工学研究科

² 和歌山大学システム工学部

a) s195064@center.wakayama-u.ac.jp

*1 <http://eiga.com/> (2018/12/25 確認)

*2 <https://movies.yahoo.co.jp/> (2018/12/25 確認)

*3 <https://www.jtnews.jp/> (2018/12/25 確認)

されている対象の語（評価属性）とその評価をする語（評価語）の組で表すことができる [2].

小林らは、用言を中心に収集した評価表現に対して、人手で評価極性情報を付与し、日本語評価極性辞書（用言編）*4 を作成した [3]. また、東山らは評価極性を持つ名詞・複合名詞に対して、人手で評価極性情報を付与し、日本語評価極性辞書（名詞編）*4 を作成した [4]. これらの評価極性辞書には、評価語とその極性が登録されており、評価語の検出とその評価語が使用されている評価表現の極性判定に用いることができる。本研究では、映画レビュー文章から評価語を検出する際に、これらの評価極性辞書を用いる。

谷本らは、インターネット上の商品レビュー文章から評価表現を抽出し可視化し、膨大なレビュー文章の内容把握を助けるシステムを提案した [5][6]. レビュー文章において、評価属性と評価語は修飾・被修飾の関係にあるという考えのもと、日本語係り受け解析によって評価語と評価属性を得ている。これに対し映画レビュー文章には、評価に関する文章以外に、作品の内容やあらすじについての文章が含まれていることがある。そのため、係り受けの関係であっても評価表現でない場合が考えられる。映画レビュー文章を対象とする本研究では、評価極性辞書によって評価語が含まれる単位文を検出し、その文章を対象に日本語係り受け解析を行う。

2.3 特徴語による文章からの情報抽出

文章から情報を抽出する手法の一つとして、TF-IDF 指標による特徴語抽出があげられる。TF-IDF とは、自然言語処理の分野で利用される文章中の単語に対する重みを示す指標である。TF（単語の出現頻度）と IDF（逆文書頻度）に基づいて算出される。

映画レビューにおいては、ある映画のレビュー文章中に多く出現する単語は、その映画の特徴を表す単語であるといえ、TF の値が大きくなる。IDF は、多くの映画タイトルにまたがってレビュー文章に出現する単語の重要度を下げ、特定の映画タイトルのレビュー文章のみに出現する語の重要度を上げる役割がある。TF-IDF 指標によって抽出された特徴語は文章の特徴を表す語である。映画レビューにおいてはその映画の内容に関連した語が特徴語となる。

しかし、TF-IDF 指標による特徴語抽出では、「この音楽は良い」という文と「この音楽は悪い」という反対の評価を表す文であっても、「音楽」という語が特徴語として出現する場合は考えられる [7]. このため、映画レビューにおいては、多くの利用者に音楽が良くないと評価された映画であっても、特徴語が「音楽」となり音楽が目玉されている映画として抽出される可能性がある。

ネガ（経験）	ひきつる
ネガ（経験）	ひっかかる
ネガ（経験）	ひもじい
...	
ポジ（評価）	欠点はない
ポジ（評価）	潔い
ポジ（評価）	血の通う人だ

図 1 日本語評価極性辞書（用言編）の例

Fig. 1 Examples of the Japanese Sentiment Dictionary (Volume of Verbs and Adjectives).

そこで本研究では、レビュー中の「良い」「悪い」などの肯定的または否定的な評価表現に着目して、情報抽出を行う。

3. 評価表現によるレビューからの評価抽出

3.1 概要

レビュー文章に含まれる評価表現に着目し、評価の情報を抽出する手法が、これまでに谷本らによって提案されている [5][6]. 本研究ではこれらの手法を元に映画レビューからの情報抽出に利用する。

評価表現とは、ある物事に対してそれを評価する語（評価語）と評価される語（評価属性）によって作られる文章である。評価語とは、文章中で物事を評価する語のことであり、3.2 節で述べる日本語評価極性辞書に登録されている名詞・動詞が評価語となる。評価属性とは、評価語が評価している対象となる語のことで、CaboCha*5 を用いた日本語係り受け解析によって得られる、評価語にかかっている文節に含まれる名詞が評価属性語となる。本研究では、映画レビューのうち評価表現が含まれる文章を対象に、この評価語と評価属性の組を抽出する。

映画レビュー利用者は、映画レビューから抽出された評価語と評価属性の組（以降、評価表現）を閲覧する。これにより、映画レビュー文章全体を読まずとも映画の評価と内容について把握できるようにすることができると考える。

3.2 日本語評価極性辞書

2.2 節でも述べた通り、日本語評価極性辞書 *4 は用言編 [3] と名詞編 [4] にわかれており、それぞれに評価語とその極性が記されている。

日本語評価極性辞書（用言編）の例を図 1 に示す。図 1 に示すように、評価極性辞書（用言編）には、評価に用いられる単語とその極性が記されている。単語にはネガ（否定的）、ポジ（肯定的）の極性と、経験（客観的）、評価（主観的）の分類が付与されており、計 5278 語が登録されて

*4 東北大学 乾・岡崎研究室 / Author(s): Inui-Okazaki Laboratory, Tohoku University

*5 <https://taku910.github.io/cabocha/> (2018/12/25 確認)

神智	p	～がある・高まる (存在・性質)
神通力	p	～がある・高まる (存在・性質)
神田	e	～がある・高まる (存在・性質)

図 2 日本語評価極性辞書 (名詞編) の例

Fig. 2 Examples of the Japanese Sentiment Dictionary (Volume of Nouns).

いる。

次に、日本語評価極性辞書 (名詞編) の例を図 2 に示す。図 2 に示すように、評価極性辞書 (名詞編) には、評価に用いられる単語とその極性が記されている。単語には n (否定的), e (中立), p (肯定的) の極性が付与され、計 13315 語が登録されている。また、単語の用法によって「～がある・高まる (存在・性質)」のように分類がなされている。

本研究においては、これらの評価極性辞書を、評価語の検出とその評価語が使用されている評価表現の極性判定に利用する。

3.3 評価情報抽出の手順

映画レビュー文章から評価表現を抽出する手順について述べる。抽出手順の概要を図 3 に示す。図 3 の映画レビュー文章は、映画レビューサイトから収集したものである。映画レビュー文章は映画タイトルごとにわかれており、映画レビューサイトにおける 1 つの投稿に含まれる本文の文章が、改行を取り除かれて 1 行ごとに記されているテキストファイルである。手順の詳細を以下に示す。

手順① 映画レビュー文章を単位文に分割する。単位文とは、「、」、「。」、「、」、「.」、「!」、「?」の記号、2 つ以上連続したスペース記号、改行コードを終端として文章を分割したものである。映画レビュー文章は単位文の集合で構成されるものとし、以降の処理は単位文ごとに行う。

手順② MeCab を用いて単位文を単語ごとに分割する。手順③ で日本語評価極性辞書を利用するため、この時、単語は原形に変換する。単位文は単語の集合で構成されるものとし、単位文とその構成単語を対応させて記録する。

手順③ 日本語評価極性辞書を用いて、評価語と、その評価語を含む単位文を抽出する。手順② で記録した単語集合の中に、日本語評価極性辞書に登録されている単語があるかを探し、登録されていた場合は、その単語を含む単位文を評価語を含む文章 (評価文) として抽出する。この時、日本語評価極性辞書と一致した単語が評価語である。評価語の検出の際には極性の情報についても付与する。また、評価文中で、「良くない (良い—ない)」、「好きではない (好きだ—ない)」のように「～ない」の形で否定されている評価語については、

否定情報として、極性を反転する処理を加える。

手順④ CaboCha を用いて評価文の日本語係り受け解析を行い、評価属性候補となる文節を抽出する。評価属性と評価語は評価文中で係り受けの関係にあるため、日本語係り受け解析によって、評価語を含む文節に係っている文節を評価属性候補文節として抽出する。

手順⑤ 手順④ で得た評価属性候補文節から評価属性となる単語を抽出する。評価属性は名詞であり、MeCab で利用されている IPA 品詞体系 [8] において、「名詞—一般」、「名詞—固有名詞」、「名詞—サ変接続」のいずれかである。従って、MeCab を用いた日本語形態素解析で、評価属性候補文節からいずれかの品詞に該当する単語を抽出し、その単語を評価属性とする。

以上の手順によって評価語、評価属性の組とその極性を抽出する。手順② 以降の処理は単位文ごとに行うため、1 つの映画レビュー文章から複数の評価表現が抽出される場合がある。抽出された評価表現は映画タイトルごとに評価属性で集約してまとめる。

4. 評価表現の有用性の調査

4.1 概要

3.3 節で述べた手順によって、映画レビュー文章から評価表現の抽出を行う。事前に収集した映画レビュー文章から評価表現を抽出し、出力された単語の傾向を考察する。また、従来の特徴語抽出手法として TF-IDF 手法により出力された単語との比較を行う。抽出結果に含まれる、評価に関連した語や内容に関連した語の出現数から、提案手法による情報抽出の有用性を確かめる。

4.2 調査環境

調査で使用したプログラムは Perl で作成した。事前に行った映画レビュー収集は Windows10 環境でプログラムを実行した。また、評価表現抽出のためのプログラムは CentOS7 環境で構築したサーバ上で実行した。

4.3 対象データ

調査では、事前に収集した映画レビューのうち、映画レビューサイトの「みんなのシネマレビュー」に投稿された映画レビュー文章を対象とする。2017 年 07 月 12 日から 2017 年 07 月 16 日の期間にウェブサイト上で公開されていたレビュー文章を収集し、映画タイトルごとにまとめた。調査では収集した映画レビューのうち、レビュー投稿数の多い映画上位 5 件を評価情報抽出の対象とする。表 1 に対象データの詳細について示す。調査では、表 1 の 5 つの映画について評価情報抽出を行う。

4.4 調査方法

4.3 節で述べたデータを対象に、提案手法により評価表現

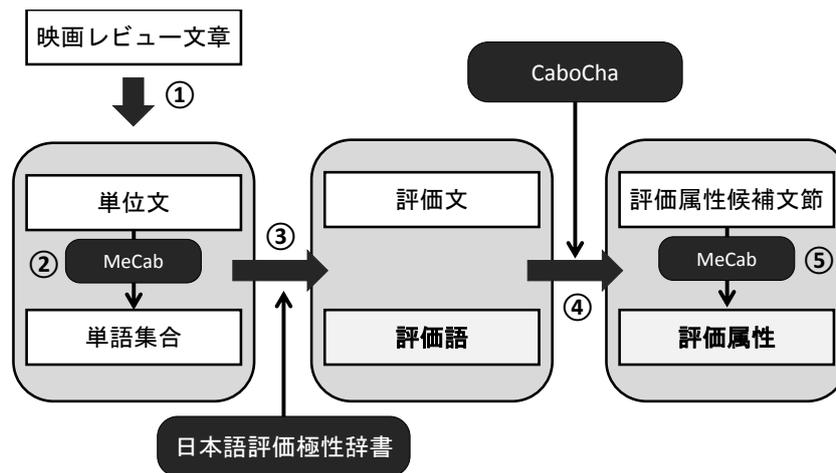


図 3 抽出手順の概要

Fig. 3 Overview of the extraction method

表 1 抽出対象データの詳細

Table 1 The detail of the target data

	レビュー件数 (件)
映画 a	1226
映画 b	879
映画 c	834
映画 d	827
映画 e	810

表 2 抽出された評価表現の数

Table 2 Number of evaluation expressions

	評価表現	
	件数 (件)	評価属性の種類数 (件)
映画 a	741	316
映画 b	766	345
映画 c	586	312
映画 d	459	240
映画 e	483	227

の抽出を行う。抽出結果の評価として、出力された単語の傾向と、従来手法によって出力された単語との比較を行う。従来手法には TF-IDF 指標を使用する。この時、TF-IDF 値を計算する単語は、提案手法の評価属性にあわせて、「名詞—一般」、「名詞—固有名詞」、「名詞—サ変接続」のいずれかとする。

システムによって画面上に抽出結果である単語が従来手法によって出力された単語のいずれかが表示される。提示された単語を、何らかの評価がされている項目を指す単語（評価関連語）、映画の内容に関連している語（内容関連語）、どちらでも無い単語のいずれかに協力者が分類する。協力者は大学生、大学院生 3 名である。2 名以上の分類が一致した場合、評価属性にその分類のラベルを付与する。提案手法により出力された評価属性の出現頻度による上位 20 語と、既存手法である TF-IDF 指標により出力された特徴語の TF-IDF 値上位 20 語の中での、評価関連語および内容関連語の出現数を検証する。

4.5 調査結果

映画タイトルごとの抽出件数を表 2 および表 3 に示す。ここで、TF-IDF 値はすべての単語に対して計算されるため、表 2 における特徴語の抽出件数は、レビュー文章に含まれている「名詞—一般」、「名詞—固有名詞」、「名詞—サ変接続」の単語の総数と同値である。また、評価表現は

表 3 抽出された特徴語の数

Table 3 Number of feature words

	特徴語	
	件数 (件)	種類数 (件)
映画 a	16390	3330
映画 b	13722	3345
映画 c	13155	3161
映画 d	9744	2572
映画 e	9482	2516

評価語と評価属性の組で抽出されるため、評価表現の件数は、評価語の件数と同値である。

表 2 より、いずれの映画においても評価表現が抽出されていることがわかる。評価表現および特徴語の抽出件数は、表 1 に示すように映画レビューの規模が大きいかほど多くなる傾向が見られる。しかし、表 1 において、映画 a のレビュー件数は映画 b よりも多かったが、表 2 での評価表現の出現数は、件数、種類数どちらも映画 b の方が多い。映画レビューの文体や書き方は様々であり、映画のジャンルや投稿者の興味の内容などの要因によって、評価表現が多く含まれる映画レビュー、あまり含まれない映画レビューの存在が考えられる。

抽出結果の例として、映画 a から提案手法によって抽出された評価属性の出現頻度上位 7 単語と、その評価語をま

表 4 映画 a の抽出結果例

Table 4 Examples of the extraction result

評価属性	出現数 (件)	評価語 (括弧内の数字はその評価語の出現数 (件))
映画	68	面白い (13), 良い (10), 素晴らしい (8), 好き (5), いい (4), 嫌い (4), 同様 (2), 大切 (2), 他
ラスト	22	素晴らしい (3), 良い (3), 爽快 (3), 清々しい (3), いい (3), 面白い (1), すばらしい (1), 好き (1), 他
希望	19	素晴らしい (5), 危険 (4), 大事 (4), 大切 (3), 確か (1), 賢い (1), 自由 (1)
シーン	17	爽快 (4), 良い (2), 素晴らしい (2), いい (2), 美しい (2), 大好き (2), 素敵 (1), 不自然 (1), 印象深い (1)
自分	17	良い (3), 好き (3), 残念 (1), 辛い (1), 憎い (1), 不幸 (1), 良くない (1), 面白くない (1), 恨めしい (1), 他
ストーリー	17	面白い (5), 良い (3), 重厚 (1), 安っぽい (1), いい (1), 簡素 (1), 丁寧 (1), 良くない (1), 素晴らしい (1), 他
後味	16	良い (10), いい (4), 悪くない (1), 悪い (1)

とめてを表 4 に示す。評価語の括弧内の数字はその評価語の出現数 (件) である。

表 4 に示すように、映画 a では、「ラスト」、「ストーリー」といった映画レビュー内で何らかの評価を受けていると考えられる単語が評価属性として出力されている。またそれらの評価属性に対しては、「素晴らしい」、「良い」、「面白くない」のような評価を与える単語が評価語として出力されている。このように、評価表現による評価情報抽出では、映画の評価されている属性 (評価属性) が、その評価を与える語 (評価語) とあわせて抽出された。

4.6 調査結果の評価

協力者 3 名により、映画 a, b, c, d, e の 5 つの映画から抽出された単語の内容関連語、評価関連語の分類を行った。対象の単語は提案手法による評価属性の出現頻度による上位 20 語と、既存手法である TF-IDF 指標による特徴語の TF-IDF 値上位 20 語である。

それぞれに分類された単語数を表 5 に示す。表 5 より、いずれの映画タイトルにおいても、評価属性における評価関連語の出現数は、特徴語における出現数よりも多い。ウィルコクソンの順位和検定の結果、これらの二群には有意差が見られた (有意水準 $p = 0.05$)。一方で、内容関連

表 5 内容関連語、評価関連語の分類結果

Table 5 Classification result

		内容関連語	評価関連語
評価属性	映画 a	3	10
	映画 b	6	9
	映画 c	2	12
	映画 d	3	9
	映画 e	1	14
特徴語	映画 a	7	6
	映画 b	8	4
	映画 c	7	5
	映画 d	6	6
	映画 e	5	9

語では、評価属性における出現数が、特徴語における出現数よりも少なく、ウィルコクソンの順位和検定の結果、これらの二群には有意差が見られた (有意水準 $p = 0.05$)。

以上のことから、提案手法による評価情報の抽出は、従来の手法に比べて評価関連語を多く抽出しており、映画レビューから評価の情報を抽出する際に有用であるといえる。

5. 抽出結果の評価実験

5.1 概要

本章では 4 章で抽出した評価表現を利用者に提示するシステムを用いて、映画レビュー中に含まれる内容や評価の情報を利用者が得ることができるか評価を行う。提案手法により抽出した評価表現、既存手法により抽出した特徴語、映画レビューの原文の一部をそれぞれウェブブラウザ上のシステムで被験者に提示する。映画の内容や評価について知ることができたか、映画を選択する際の参考にしたかをアンケートにより調査する。

5.2 実験環境

実験の被験者は大学生・大学院生の男女 18 名である。実験期間は 2018 年 1 月 5 日から 2018 年 1 月 17 日の 12 日間であり、実験期間の初日に被験者に対して、メールで実験内容を伝える。被験者は、実験期間中に各自のコンピュータからウェブブラウザを利用して実験システムにアクセスする。システム利用後、被験者はウェブブラウザで表示されたアンケートに回答する。

5.3 実験用システム

実験で利用するシステムは、4 章で抽出した映画レビュー中の評価表現と、特徴語、レビュー文章をそれぞれウェブブラウザ上で表示する映画評価閲覧システムである。

システムの構成を図 4 に示す。

図 4 に示すように、映画評価閲覧システムはトップページと映画ごとに関連する情報を提示する専用ページで構成されている。

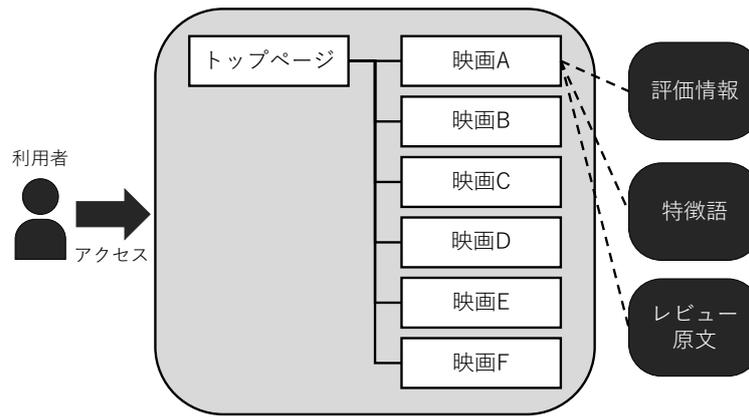


図 4 映画評価閲覧システムの構成
 Fig. 4 Configuration of the system

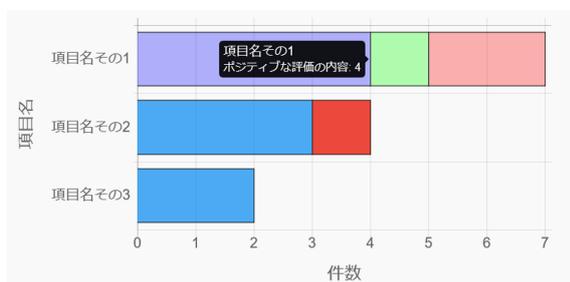


図 5 評価情報の提示例
 Fig. 5 Example of the evaluation expression

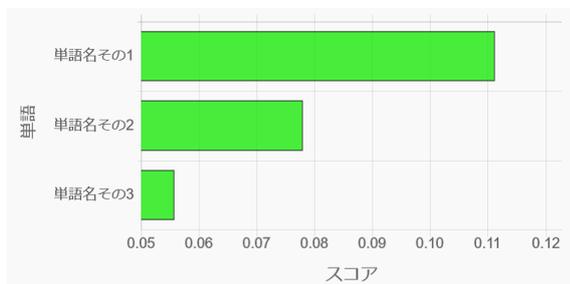


図 6 特徴語の提示例
 Fig. 6 Example of the feature words

平山らの研究より、抽出結果である評価情報のグラフ表示は利用者がその結果を理解するのに有効であると考えられている [9]。本研究においても抽出された評価表現をグラフ表示し、利用者に提示する。

評価表現の提示例を図 5 に示す。評価表現は、図 5 のように、出現数を棒の長さで表す横棒グラフとして提示する。縦には評価属性が出現頻度の上位から 15 語並び、マウスオーバーによってそれぞれの評価語をツールチップで提示する。青色は評価語がポジティブな評価、赤色はネガティブな評価、緑色は中立的な評価であることを表している。

特徴語の提示例を図 6 に示す。特徴語は、図 6 のように、TF-IDF 値をスコアとして棒の長さで表す横棒グラフとして提示する。縦には特徴語が TF-IDF 値の上位から 15 語

No.	レビュー
1	映画のレビュー文章その1
2	映画のレビュー文章その2
3	映画のレビュー文章その3

図 7 レビュー原文の提示例
 Fig. 7 Example of the review text

表 6 抽出対象データの詳細
 Table 6 The detail of the target data

	レビュー件数 (件)
映画 A	794
映画 B	627
映画 C	622
映画 D	620
映画 E	609
映画 F	601

並ぶ。また、評価極性の混同を避けるため、図 5 で中立的な評価として使用した緑色を使用する。

続けて、レビュー原文の提示例を図 7 に示す。レビューの原文は、図 7 のように、リストの形式で提示する。レビューの表示件数は 15 件で、収集した映画レビューからランダムに選択したものを表示する。

5.4 対象データ

実験に使用する映画レビューを決定するにあたり、事前調査として、映画レビューの投稿件数が 600 件以上である映画 22 本のタイトルを列挙し、映画の視聴経験をアンケートで質問した。本実験では、視聴経験による映画への印象の変化を低減させるために、被験者全員から視聴経験が無いと回答された映画 6 タイトルを実験の対象とする。対象データの詳細を表 6 に示す。

5.5 実験方法

5.4 節で述べた映画 6 タイトルを対象に、5.3 節の映画評

表 7 表示形式の組み合わせ
Table 7 Combination of the display format

	被験者グループ		
	グループ A	グループ B	グループ C
映画 A	評価表現	特徴語	レビュー原文
映画 B	レビュー原文	特徴語	評価表現
映画 C	評価表現	レビュー原文	特徴語
映画 D	特徴語	評価表現	レビュー原文
映画 E	特徴語	レビュー原文	評価表現
映画 F	レビュー原文	評価表現	特徴語

価閲覧システムを用いて、映画レビュー中に含まれる内容や評価の情報を利用者が得ることができるか評価を行う。

被験者 18 名を 6 名ずつの 3 グループに分け、それぞれをグループ A, B, C とする。また、実験の対象である映画 6 タイトルに対し、評価表現、特徴語、レビュー原文の 3 種類の表示形式で、専用ページを計 18 ページ用意する。これらのページを、グループごとに異なる組み合わせになるように、1 つの表示形式につき 2 タイトルずつ割り当てる。被験者のグループと表示形式の組み合わせを表 7 に示す。

実験では、被験者に対して「あなたは面白そうな映画を探すために、いくつかの映画の評価や感想を知りたいと思っています。」というタスクを与える。被験者は実験システムにアクセスし、システムを利用しながらアンケートに回答する。アンケートでは映画 1 タイトルごとに、「評価やレビューの数（以降、評価数）」、「映画の内容（以降、内容）」、「映画の中で肯定的または否定的に評価されている点（以降、評価点）」、「肯定的または否定的な意見の傾向（以降、傾向）」の 4 つの項目について、それぞれ知ることができたかを質問する。被験者は 1 を「あてはまらない」、5 を「あてはまる」とした 5 段階で回答する。知ることができたという回答が多い場合、その映画タイトルで使用されていた表示形式は、項目が指す情報を知るのに有用である表示形式であるといえる。また、映画ごとにその映画を面白そうと思ったかについて質問する。被験者はその回答の根拠として、4 つの項目を参考にしたかどうかを、それぞれ「あてはまらない」—「あてはまる」の 5 段階で回答する。面白さの判断の参考にしたという回答が多い場合、その項目は映画の評価を決める際に重視されている項目であり、その表示形式で評価情報を提示した際に、注目された情報であるといえる。

アンケートの回答は表示形式ごとにまとめ、5 段階評価の中央値を求める。評価表現、特徴語、レビュー原文の 3 つの表示形式を比較し、提案手法が映画レビューからの評価情報抽出に有用であるか検証する。

5.6 結果

提案手法により抽出した評価表現、既存手法により抽出した特徴語、映画レビューの原文の一部をそれぞれウェブ

表 8 アンケート結果（中央値）
Table 8 Questionnaire results

表示形式	知ることができたか			
	評価数	内容	評価点	傾向
評価表現	4	3.5	5	4
特徴語	2	4	1.5	2
レビュー原文	4	4	4	4

表示形式	参考にしたか			
	評価数	内容	評価点	傾向
評価表現	4	4	4	4
特徴語	2	4	1.5	1
レビュー原文	3	4	4	4

ブラウザ上のシステムで被験者に提示し、映画の内容や評価について知ることができたか、映画を選択する際の参考にしたかをアンケートで質問した。質問の回答ごとに中央値を求めた結果を表 8 に示す。

表 8 の、「評価数」、「評価点」、「傾向」の 3 つの項目に着目する。各項目において、それらの項目のことを知ることができたかという質問に対し、高い評価を回答する被験者が、特徴語に比べ評価表現において多くなっていることがわかる。ウィルコクソンの順位和検定の結果、これらの二群にはいずれも有意差が見られた（有意水準 $p = 0.01$ ）。また、「内容」については、評価表現と特徴語の間に大きな差は見られず、ウィルコクソンの順位和検定の結果、二群の間に有意な差は無かった。

次に評価表現とレビュー原文について述べる。「評価数」の項目において、評価表現とレビュー原文の間に有意な差は見られなかった。「評価点」、「傾向」の 2 つの項目においては、評価表現の値がレビュー原文よりも大きく、ウィルコクソンの順位和検定の結果、二群の間に有意差が見られた（有意水準 $p = 0.05$ ）。一方で、「内容」については、評価表現よりもレビュー原文の方が知ることができたという回答が多く、ウィルコクソンの順位和検定の結果、これらの二群には有意差が見られた（有意水準 $p = 0.01$ ）。

また、映画の面白さを判断する際に「評価数」、「評価点」、「傾向」の 3 つの項目のことを参考にしたという回答が、特徴語に比べ評価表現において多くなっていることがわかる。ウィルコクソンの順位和検定の結果、これらの二群にはいずれも有意差が見られた（有意水準 $p = 0.01$ ）。一方で、「内容」については、いずれの表示形式の場合でも中央値が 4 と同値であり、これらの間に有意差は見られなかった。

5.7 考察

表 8 の結果より、「評価数」、「評価点」、「傾向」の 3 つの項目において、評価表現と特徴語の間に差が見られたことから、評価表現によって抽出された評価情報の提示は、特徴語を提示するよりも、これら 3 つの項目について情報を

与えることができたといえる。これらの項目は映画の評価に関係しており、提案手法による評価情報抽出が映画の評価を知る際に、有用であることを示している。一方で、“内容”については、評価表現と特徴語で同程度の評価であったことから、映画の内容については、提案手法でも従来手法と同程度の内容を知ることができるといえる。

また、“評価数”の項目において、評価表現とレビュー原文の間に有意な差は見られず、“評価点”、“傾向”の2つの項目においては、評価表現の値がレビュー原文の値よりも大きいという結果が得られた。この結果から、評価表現による評価情報の提示は、レビュー原文をそのまま読むのと同程度、またはそれ以上に評価の情報を利用者に伝えることができるといえる。“内容”については、評価表現よりもレビュー原文の方が知ることができたという回答が多い。これは、評価表現による情報抽出には含まれない、映画の内容を示す情報が映画レビュー文章に存在していることを示している。

続けて、“評価数”、“評価点”、“傾向”の3つの項目において、それらの項目のことを参考にしたという回答が、特徴語に比べ評価表現において多くなっていた。このことから、評価表現を提示した場合、被験者は評価の情報に強く注目しており、映画の面白さの判断にそれらの情報を利用しているといえる。

以上のことから、評価表現による情報の提示は、“評価数”、“評価点”、“傾向”などの映画の評価を知る際に、有用であるといえる。映画の評価に関する情報は、それらの情報を提示した際に、映画の面白さを判断するために注目されており、映画の評価情報を抽出することは、これらの判断を助けたと考えられる。これらの情報は映画レビュー原文と同程度の評価を受けており、このことから、評価表現に着目して評価情報を抽出することで、利用者は膨大なレビューを読まずとも映画の評価について知ることができたといえる。

6. おわりに

本研究では、映画レビュー文章から映画の内容や評価に関する情報を抽出し、利用者に提示する手法を提案した。情報の抽出には、映画レビュー文章中に含まれる、物事を評価する表現（評価表現）を手がかりとした。

収集した映画レビュー文章から、提案した手法によって評価情報を抽出する調査を行ったところ、映画レビュー内で何らかの評価を受けていると考えられる単語が評価属性として出力された。また、それらの評価属性に対して評価を与える単語が評価語として出力された。出力された単語を分類した結果、提案手法による評価情報の抽出は、従来の手法に比べて評価に関連した単語を多く抽出しており、提案手法は映画レビューから評価の情報を抽出する際に有用であるといえる。

続けて、抽出された評価情報を閲覧するシステムを作成し、映画レビュー中に含まれる内容や評価の情報を利用者が得ることができるか評価した。その結果、評価表現によって抽出された評価情報の提示は、映画の評価を知る際に有用であることが示された。また、被験者は映画の面白さを判断するために評価の情報を参考にしており、映画の評価情報を抽出することは、これらの判断を助けたと考えられる。

本研究では、映画の評価を映画に対する個々の感想、映画の内容を映画の登場人物の行動の説明などを指す情報として扱った。しかし、映画レビュー文章には映画の特定の内容を指して評価するレビューが含まれている。

評価情報の抽出結果には、「映画」や「作品」といった、特定の内容を指定せず映画そのものを指す評価属性や、映画全体を通しての構成要素を指す評価属性と、映画の一部分の内容を反映した評価属性が混在している。本研究ではこれらの評価属性をすべて同一のものとして扱ったが、映画の内容が反映された評価表現を抽出するには、評価属性が指す対象に応じての分類が必要であると考えられる。

今後の課題として、これら評価属性の指す対象の整理と、評価属性の分類、そして、映画の内容が反映された情報を抽出できるような評価表現の抽出手法の改善があげられる。

謝辞 本研究の一部は、JSPS 科研費基盤研究(C)(JP16K00371)の助成を受けたものである。

参考文献

- [1] 岩井秀成, 土方嘉徳, 西田正吾:“レビューの文脈一貫性を用いたあらすじ文判定手法”, 情報処理学会論文誌 データベース, Vol.7, No.2, pp.11-23(2014).
- [2] 乾孝司, 奥村学:“テキストを対象とした評価情報の分析に関する研究動向”, 自然言語処理, Vol.13, No.3, pp.201-241(2006).
- [3] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一:“意見抽出のための評価表現の収集”, 自然言語処理, Vol.12, No.3, pp.203-222(2005).
- [4] 東山昌彦, 乾健太郎, 松本裕治:“述語の選択選好性に着目した名詞評価極性の獲得”, 言語処理学会第14回年次大会論文集, pp.584-587(2008).
- [5] 谷本融紀, 太田学:“評価属性を考慮した評判情報の可視化”, 情報処理学会研究報告, Vol.2010-DBS-151, No.12, pp.1-8(2010).
- [6] 谷本融紀, 太田学:“特定評価属性の関連属性自動抽出による評価表現辞書の生成”, 情報処理学会研究報告, Vol.2012-DBS-155, No.12, pp.1-9(2010).
- [7] 林貴宏, 尾内理紀夫:“Web上のレビューを利用した映画推薦システム”, 人工知能学会論文誌, Vol.30, No.1, pp.102-111(2015).
- [8] 浅原正幸, 松本裕治:“ipadic version 2.7.0 ユーザーズマニュアル”, 奈良先端科学技術大学院大学 情報科学研究科 自然言語処理学講座, 平成15年11月発行(2003).
- [9] 平山拓央, 湯本高行, 新居学, 佐藤邦弘:“語の共起と極性に基づく商品レビュー閲覧支援システム”, 情報処理学会研究報告, Vol.2012-DBS-155, No.3, pp.1-9(2012).