

# 映画の実況ツイートにおける時系列毎の感情抽出手法の提案

若井 祐樹<sup>†</sup> 山本 湧輝<sup>††</sup> 熊本 忠彦<sup>†††</sup> 灘本 明代<sup>††</sup>

<sup>†</sup> 甲南大学大学院 自然科学研究科 〒658-8501 兵庫県神戸市東灘区岡本 8-9-1

<sup>††</sup> 甲南大学 知能情報学部 〒658-8501 兵庫県神戸市東灘区岡本 8-9-1

<sup>†††</sup> 千葉工業大学 情報科学部 〒275-0016 千葉県習志野市津田沼 2-17-1

E-mail: <sup>†</sup>m1224008@center.konan-u.ac.jp, <sup>††</sup>si071166@center.konan-u.ac.jp, nadamoto@konan-u.ac.jp,  
<sup>†††</sup>kumamoto@net.it-chiba.ac.jp

**あらまし** 近年, Twitter などのマイクロブログが普及しており, 自分の身近で起こったことを気軽にツイートすることができる. 実際に Twitter を用いて, スポーツやドラマ, 映画などのテレビ番組を視聴しながら実況ツイートを発するユーザが増加している. ユーザ自身がその時視聴しているシーンに対して, 興味を持ったことやどのように感じたのかツイートすることによって, 同じ番組を見ている他ユーザと共有することができる. またツイートには, Twitter 特有表現が多く含まれており, 感情をよりわかりやすく表現するために用いられているものと考えられる. そこで本論文では Twitter 特有表現も考慮に入れた感情の抽出を行う. また, 1つの番組に対して, 盛り上がるシーンや見所があるが, 時系列によってユーザが感じた感情が異なっていると考えられる. そこで本論文では, テレビで放送されている映画に焦点を当て, ツイートの感情を時系列に抽出し, 分析を行う.

**キーワード** Twitter, 感情抽出, 時系列, 特有表現

## 1. はじめに

近年, Twitter などのマイクロブログが普及しており, 一般の人々の様々な声がインターネット上に存在している. Twitter は, 自分の身近で起こったことを気軽に発信 (ツイート) することができる. 実際に, Twitter を用いてスポーツや映画, ドラマ等のテレビ番組を視聴しながら実況ツイートを発するユーザが増加している. 特に映画やドラマ等の実況ツイートは主に, 視聴している番組の出演者やキャラクターのセリフをツイートしたり, 番組に対する自分の意見や感想をツイートしていることが多い. 特に後者の番組に対するツイートの中には, 視聴者の感情が含まれていることが考えられる. つまり, 自分が思っていることや感じたことを発信していることが推測される. そこで本論文では, テレビで放送された映画の実況ツイートに着目し, 映画に対するツイートの感情を時系列毎に抽出し, 感情の変化を分析する. 映画の実況を対象とした理由として, 映画は1作品でストーリーが完結している作品が多く, 起承転結がはっきりしているため, 感情の流れが大きく変化するのではないかと考えられる.

また実際に, 実況ツイートを見てみると, 「きたあああ!」といった, 語句の後ろに同じ母音を繰り返して叫んでいるように表現しているもの (以下 Twitter 特有表現と呼ぶ) も存在する. このような Twitter 特有表現はよりそのツイートの感情を明示的に表していると考えられる. 以前我々は Twitter 特有表現に着目したユーザ実験を行い, 特有表現があることによって, 感情が強調されたり弱められたりしていることがわかった [1]. そこで, 本論文では, Twitter 特有表現を考慮に入れて, 実況ツイートの感情抽出を行う. 具体的には, 新しい感情語辞書の構築を行い, 構築した辞書を用いて, ツイートの感情を定量化す

る. Twitter 特有表現が観測された場合, 感情の種類によって重み付けを行い, ツイートの感情を抽出する. 最後に, 映画に対するツイートを時系列から感情の変化を分析する.

以下, 第2章では関連研究を, 第3章では感情表現語辞書について, 第4章では Twitter 特有表現を考慮したときの感情抽出手法の提案について, 第5章では Twitter 特有表現の重みを考慮したときの評価実験について, 第6章では映画の実況ツイートの時系列ごとの分析について, 最後に第7章ではまとめと今後の課題について述べる.

## 2. 関連研究

現在, 感情表現を抽出する研究が多数行われている. その中で, 感情表現を表す感情モデルが提案されており, 多次元の感情モデルが提案されている. 代表的な感情モデルとして, Plutchik [2] のモデルがある. 人間の感情は「嫌悪⇔信頼」「悲しみ⇔喜び」「驚き⇔予測」「恐れ⇔怒り」の8つの基本となる感情に分類され4次元のベクトルで表されている. また, 中村 [3] は「喜, 好, 安, 哀, 厭, 怖, 怒, 恥, 昂, 驚」の10軸の感情で表しており, 熊本ら [4] は, 「楽しい⇔悲しい」「うれしい⇔怒り」「面白い⇔つまらない」「楽観的⇔悲観的」「のどか⇔緊迫」「驚き⇔ありふれた」の6軸の感情軸で表している. このように様々な感情軸が存在するが, 本研究では, 中村の感情軸を用いる.

高岡ら [5] は中村の提案する10次元の感情軸から6次元の感情軸へ次元削減を行い, この6次元の感情軸を用いて, 名言から感情を表す語 (感情語) を抽出している. そして, これら感情軸と感情語を用いて, ユーザの気分にあった名言を検索するシステムを提案している. 本研究では映画の実況ツイートに着目し, そのツイートから感情を抽出する点が異なる.

表 1 感情表現辞典の感情とその一部の単語例

感情	単語	感情	単語
喜	楽しい, 面白い, 嬉しい, 笑い, 堪能, 至福	怒	怒る, 腹ただしい, 憤怒, 腹立ち, 煮え返る, 仕返し
哀	悲しい, 痛い, 孤独, 寂しい, がっかり, 溜め息	怖	怖い, 不安, 震える, 気味悪い, 臆する, 萎縮
恥	恥ずかしい, もじもじ, 照れる, ハニカム, 赤面, 俯き	好	恋しい, 愛しい, 好き, 恋しい, 愛おしい, 敬愛
厭	暗い, 憎い, 辛い, 不愉快, せせら笑う, 気不味い	昂	奮起, 感動, 動揺, やきもき, 騒ぐ, 熱狂
安	のんびり, すっきり, あっさり, 安らぎ, 弛緩, 懐かしい	驚	ぼんやり, 歓喜, 慌てる, 思いがけない, 面食らう

徳久ら [6] は, ユーザの発話内容から感情を推定する手法を提案している. 発話内容から感情を生起する要因の文を獲得する (この事態の集合を感情生起要因コーパスと呼ぶ). 感情モデルには, (嬉しい, 楽しい, 安心, 恐い, かなしい, 残念, 嫌, 寂しい, 心配, 腹立たしい) と (neutral) の 11 種類の感情を用いている. 堀宮ら [7] は, Twitter の特徴である他者との会話機能を用いて, 人間と人間に対する推測能力に着目し, ユーザへの発言に対する他者の反応であるリプライを利用して感情推定を行う手法を提案している. 堀宮らは, Ekman が定義した (幸福, 驚き, 恐れ, 悲しみ, 怒り, 嫌悪) の基本 6 感情を用いて, ユーザの感情を推定している. 本研究では, 人間と人間との発言から感情を提示するのではなく, 映画の実況ツイートから感情の提示を行う点が異なる. 水岡ら [8] は, 同じ感情表現を持つメッセージは時間的に近傍に出現しやすいと考え, Twitter のログを利用して感情表現を収集している. 感情表現の収集対象をあらかじめ決めておき, 時間的に近いつぶやきに似た感情表現をしているものを判定する. 感情表現には, (かっこいい, かわいい, 泣ける, 笑える) の 4 種類の感情を用いている. 本研究は中村の感情軸から新しい感情表現語辞書を構築し, 感情抽出を行う点が異なる.

浅井ら [9] は, マイクロブログ上で使われる突発的な感情の例である, 「○○きたああ」のような語尾の母音を繰り返す表現をする語を「叫喚フレーズ」と定義し, 特定のイベントにおける叫喚フレーズの発生頻度の推移について調査するとともに, その抽出手法を提案している. 本研究では, 浅井らが定義した叫喚フレーズのうち語句の後ろに母音 (大文字・小文字は区別しない) が 3 つ以上続いている表現のことを Twitter 特有表現と定義し, Twitter 特有表現を考慮したツイートの感情抽出手法を提案する.

山内ら [10] は, テレビ番組関連のツイートを対象とし, ツイートした時間の番組シーンの視聴者の感情を推定している. 感情推定の結果を利用し, シーンにおける視聴者の感情をグラフ化し, 提示を行っている. 中澤ら [11] はテレビ番組に関連するツイート数の変動から重要シーンを自動的に検出し, ツイート内容から主要人物とイベント内容を推定して, その結果をシーンを表すラベルとして付与する手法を提案している. 宮森ら [12] は, 番組実況チャットに基づく視聴者反応を利用したテレビ番組のメタデータ抽出手法を提案している. 視聴者の反応を利用することによって, 番組の盛り上がり場面や, 視聴者の嗜好・興味に沿ったリアクションなどを効率良く抽出できることを示している. 佃ら [13] は, ニコニコ動画を対象に, 視聴者

の動画への反応に基づいた動画の検索や動画の推薦を可能とするシステムを提案している. 動画の各場面におけるコメントの数や内容から, 動画の各登場人物の活躍シーン及び活躍の大きさを推定している. 本研究は, Twitter 特有表現を考慮した感情抽出手法を提案しているため, これら研究とは異なる.

### 3. 感情語辞書

ツイートから感情を抽出し定量化するため, まずはじめに, 語句毎に感情の定量化がされている辞書の構築を行う. 構築の手法として, 熊本ら [4] が提案する感情表現語辞書の構築手法を用いる. 熊本らの手法は, 辞書を構築するための大量のデータに現れる任意の単語とあらかじめ定義してある感情語群との共起関係を調べ, その結果に基づいて感情表現語辞書を構築している. 各辞書の感情軸は, 反義語関係にある 2 つの感情語から構成されている.

熊本らは, 新聞記事データベースから感情表現語辞書を構築している. 感情軸は新聞記事を読んだ人々がその記事からどのような印象を受けたかをアンケート調査により調べた結果に基づき, 「楽しい⇔悲しい」, 「うれしい⇔怒り」, 「のどか⇔緊迫」, 「面白い⇔つまらない」, 「楽観的⇔悲観的」, 「驚き⇔ありふれた」で提案されており, 新聞記事の印象を定量的に捉えるための尺度 (スケール) となっている. 6 本の感情軸のうち「楽しい⇔悲しい」, 「うれしい⇔怒り」, 「のどか⇔緊迫」の 3 つの感情軸を用いて辞書の構築を行っている.

熊本らが構築した感情表現語辞書の感情軸では, ユーザが自由に発信しているツイートの感情を抽出していくには不十分であることが考えられる. また, Twitter では, 様々な感情を持つツイートをすることが多く, 3 つの感情軸で感情を表現することが困難であることが考えられる. さらに, ユーザ自身が自由にツイートしている点で, 感情表現語辞書に含まれていない単語がツイート内に多く含まれていることが問題点である.

そこで本研究では, Yahoo!映画<sup>(注1)</sup>のレビューデータ 74,000 文書内の任意の単語と中村の感情表現辞典から抽出した感情語群との共起関係に基づいて感情語辞書を構築する. 映画のレビューデータを用いた理由として, ツイートより文章がしっかりしている点, 見た映画に対してどのように感じたのか感情を示す語が多いため, 辞書を構築するためのデータとして最適であると考えたためである. 中村の感情表現辞典を選んだ理由は, 感情表現辞典は表 1 のように 10 軸の感情が分類されており, ツイートに含まれる様々な感情に対応できると考えたからであ

(注 1) : Yahoo!映画. <http://movies.yahoo.co.jp/>

表 2 新しく構築した感情表現語辞書の一部

	哀⇔喜	厭⇔好	恥⇔安	怖⇔怒	驚⇔昂
爆笑できる	0.9266	0.0142	0	0.3717	0.500
恋する	0.1267	0.8927	-0.0176	-0.1385	0.0561
嫌らしい	-0.2417	-0.9616	0	0.0424	-0.0938
安全だ	-0.1591	0.0414	0.9713	-0.1843	0.1148
悔しい	0.0284	0.1238	-0.1801	0.9403	0.1148
恐ろしい	-0.2891	-0.2304	0.0375	-0.9662	0.14
熱心だ	-0.0481	0.1551	0.183	0.231	0.9697
意外だ	-0.0081	0.1663	0.207	0.003	-0.9906

る。また、1つの感情語辞書に対する2つの対極の感情の決定に関しては、Plutchikの感情の輪[2]で向かい合っている感情は対極関係であることを参考にして、「哀⇔喜」、「厭⇔好」、「恥⇔安」、「怖⇔怒」、「驚⇔昂」とする。定量化した値(感情値)は、構築した辞書の中の1つの単語に軸毎の感情値が-1.0から1.0の間で付与されており、-1.0に近いほど「哀」などの左側の感情が、1.0に近いほど「喜」などの右側の感情が強いことを示している。新しく構築した辞書を感情語辞書と呼ぶ。感情語辞書の一部を表2に示す。

#### 4. Twitter 特有表現を考慮したツイート感情値算出手法

##### 4.1 ツイートの感情値算出手法の手順

ツイートから感情を抽出するために、構築した感情語辞書を用いて、Twitter特有表現を考慮したツイート毎の感情値算出手法を提案する。算出手順を以下に示す。

(1) 実況ツイートをJuman<sup>(注2)</sup>を用いて形態素解析を行う。

(2) 実況ツイートに含まれる単語の内、構築した感情語辞書に含まれる単語の感情値を辞書から抽出する。

(3) 抽出した単語の感情値の合計を単語数で割った値をそのツイートの感情値とする。

(4) 実況ツイート内にTwitter特有表現が含まれている場合、設定したTwitter特有表現の感情値の重みをツイートの感情値に加算する。

(5) 算出された感情値のうち、最も大きい値の感情軸をそのツイートの感情とする。

ユーザは自由にツイートをするため、様々な表現方法がされ、表記ゆれが起きる場合がある。例えば、「可愛い」という単語では、「かわいい」というように漢字・ひらがなで表現する場合があります。異なってしまうことがある。そこで本研究では、このような表記ゆれに対応するために、形態素解析器Jumanの「代表表記」に着目した。例えば、「楽しい/たのしい」のように、複数の表記方法がある単語に対して、「代表表記」が付与される。そこで、ツイートを形態素解析を行う際、「代表表記」に着目して、感情値の算出を行うことによって、表記ゆれに対応する。

##### 4.2 Twitter 特有表現の重みの設定

Twitter特有表現の重みを求めるためにユーザ実験を行った。データセットは、映画の実況ツイート中から、Twitter特有表現が含まれている任意の50ツイートをを用い、被験者は20代男性11名である。実験手順は以下に示す。

(1) 抽出した50のツイートからTwitter特有表現を削除したツイートをユーザに見せ、感情表現辞典の10軸の感情の中からそのツイートの感情に適した感情軸を最大3軸選んでもらい、選択した感情軸に対して10段階評価を行った。

(2) (1)と同じツイートからTwitter特有表現を削除せず、ツイートそのままをユーザに見せ、(1)と同様の評価をしてもらった。

実験の結果を表3に示す。表3ではTwitter特有表現がある場合とない場合を比較し、Twitter特有表現があることで感情が強くなったと評価されたときの割合を感情(強)で示し、弱くなったときの割合を感情(弱)で示している。

表 3 特有表現によって感情の変化があったときのツイートの割合

感情	感情(強)	感情(弱)
喜	67%	2%
哀	35%	10%
好	80%	2%
厭	33%	14%
安	8%	41%
恥	14%	0%
怒	27%	2%
怖	29%	10%
昂	100%	0%
驚	84%	4%

表3の結果より、ある感情*i*のツイート特有表現の重み $TW_i$ を以下の式で求める。

$$TW_i = (SS_i - SW_i) / \max(SS_i, SW_i) \quad (1)$$

$SS_i$ は、表3の結果の感情*i*の感情(強)の値を示し、 $SW_i$ は、表3の結果の感情*i*の感情(弱)の値を示す。上記の式で求めた重み $TW_i$ は、喜:0.97、哀:0.71、好:0.98、厭:0.58、安:-0.80、恥:1.00、怒:0.93、怖:0.66、昂:1.00、驚:0.95である。

(注2): Juman. <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

表 4 各感情軸の特有表現があるツイートの数と適合率

感情	ツイート数	適合率
喜	3	100.0%
哀	17	0%
好	18	77.8%
厭	2	0%
安	4	0%
恥	16	0%
怒	6	0%
怖	14	0%
昂	18	94.9%
驚	2	50.0%

## 5. Twitter 特有表現の重みを考慮したときの評価実験

我々が Twitter 特有表現を考慮したツイートの感情抽出手法の有用性を示すために 2 種類のデータを用いた評価実験を行った。

### 5.1 視聴率の高い番組での評価実験

多人数が視聴している視聴率の高い番組を対象として、Twitter 特有表現を用いた感情抽出手法の有用性を示す実験を行った。実験で使用したデータは「紅白歌合戦」に関するハッシュタグが付与されている、2013 年 12 月 31 日の 19 時 30 分から 20 時までのツイート 1,032 ツイートのうち、Twitter 特有表現を用いている 103 ツイートを用いた。

#### 実験結果と考察

5 つのそれぞれの感情軸に対して、提案手法を用いた結果を表 4 に示す。結果より、Twitter 特有表現に対する適合率を見ると、「喜」、「好」、「昂」に対して精度が良いことがわかった。精度が良い全ての感情は Twitter 特有表現の感情値の重みを大きく設定している感情であることから、重み付けに有用性があることが考えられる。また、3 つの感情の共通点として、ポジティブな感情で気持ちが高揚しているさまを表現する時に、Twitter 特有表現が使われていることが多いと考えられる。

しかし、「哀」や「恥」、「怖」の 3 つの感情は、提案手法によって抽出されたツイート数が多かったが、精度が悪くなかった。これは対象としているツイートが、「紅白歌合戦」に関するツイートに絞っていることから、ネガティブな感情が含まれる表現が少なかったことが考えられる。今後は、異なったジャンルのテレビ番組に関する実況ツイートを収集し、ジャンルによって感情の変化が異なるのか検証を行っていくことが課題である。

### 5.2 Twitter 特有表現を含む映画の実況ツイートに対する評価実験

実際に、Twitter 特有表現が含まれる映画の実況ツイートの評価実験を行った。被験者は 8 名で、実験データは Twitter 特有表現が含まれるツイートと Twitter 特有表現の部分を削除したツイートの合計 54 ツイートを用いる。実験方法は以下に示す。

表 5 各感情軸の Twitter 特有表現の重みを考慮しない場合の評価

	再現率	適合率	F 値
喜	0.000	0.000	0.000
哀	0.400	0.111	0.174
好	0.417	0.139	0.208
厭	0.308	0.333	0.320
安	0.286	0.067	0.108
恥	0.000	0.000	0.000
怒	0.333	0.111	0.167
怖	0.222	0.063	0.098
昂	0.261	0.231	0.245
驚	0.133	0.091	0.108

表 6 各感情軸の Twitter 特有表現の重みを考慮した場合の評価

	再現率	適合率	F 値
喜	0.172	0.357	0.233
哀	0.636	0.389	0.483
好	0.654	0.472	0.548
厭	0.370	0.833	0.513
安	0.167	0.133	0.148
恥	0.429	0.091	0.150
怒	0.417	0.278	0.333
怖	0.444	0.250	0.320
昂	0.451	0.885	0.597
驚	0.313	0.455	0.370

(1) ユーザは感情表現辞典の 10 軸の感情の中からツイートの感情に適した感情軸を最大 3 つまで選び、選択した感情軸に対して 10 点満点にて評価した。

(2) 提案した感情値算出手法を用いて、Twitter 特有表現の感情値の重みを考慮した場合としない場合のツイートの感情値を算出して感情の出力を行う。なお、算出された値が閾値以上だった場合のみ、そのツイートに関する感情と判別するものとする。

(3) ユーザによる実験結果を正解データとし、感情値算出手法から得られた出力結果と比較し、再現率・適合率・F 値を求め、分析を行う。

#### 実験結果と考察

Twitter 特有表現の重みを考慮した場合と、しなかった場合の結果を表 5、表 6 に示す。表 5、表 6 の結果から、「哀」、「好」、「厭」、「昂」の感情が特に Twitter 特有表現に関する重み付けを行うことによって、重み付けをしない場合より F 値が良いことがわかる。これは、重み付けを文の感情値に加算することによって、感情をより強調する効果が見られたため、良い結果となったと考えられる。しかしながら、「喜」、「怒」、「怖」、「驚」の感情は、重み付けを考慮したほうが F 値が大きくなっているが、F 値は良いとはいえないことがわかる。これは「驚」の場合、「昂」のような対となっている感情が多く判断されたため、F 値が低くなってしまったことが考えられる。重みを -0.80 と負の値に設定してあった「安」の感情に関しても F 値が良くなっていることがわかる。文のみで感情値を算出したときに「安」と判断されていたが、重み付けをすることによって、「安」の感

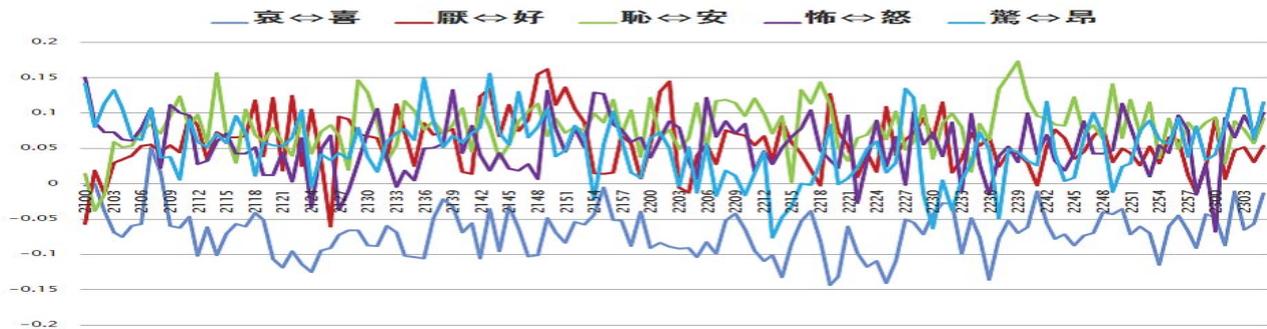


図1 時系列ごとの感情値

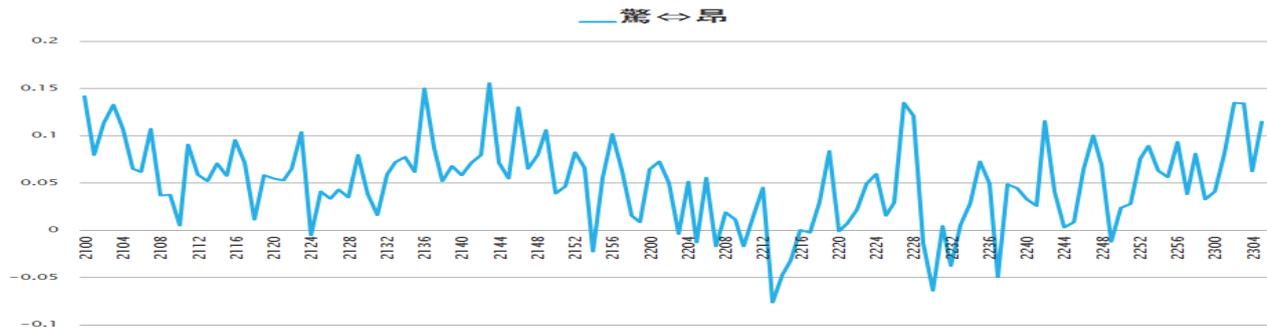


図2 「驚↔昂」の時系列ごとの感情値

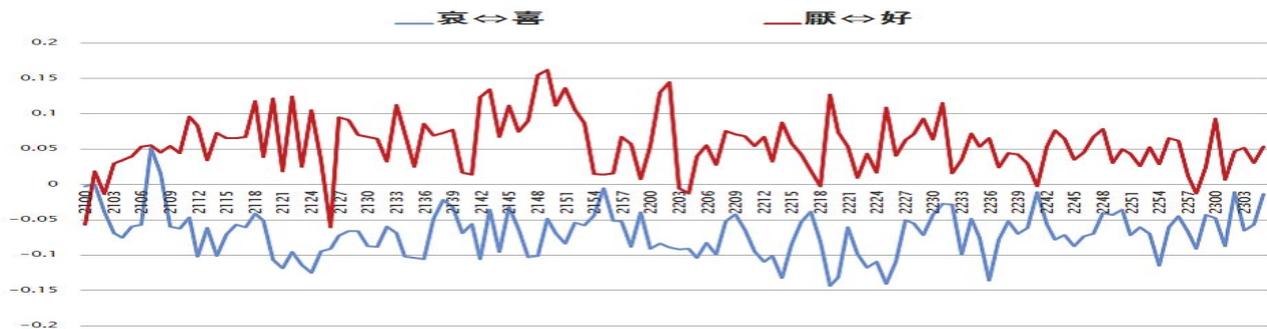


図3 「哀↔喜」, 「厭↔好」の時系列ごとの感情値

情を弱めることができたためであると考えられる。「安」と加えて「恥」の感情はユーザ実験で両者とも、感情であると判断された数が少なかったことも問題点であると考えられる。今後は、さらに実験データを増やして、検証を行っていくことが課題である。

## 6. 映画の実況ツイートの時系列ごとの分析

実際にテレビで放送された、映画の実況ツイートから感情値を抽出する。本研究で使用した映画の実況ツイートは「ONE PIECE FILM Z」に関する実況ツイート 5,549 ツイートとそのツイートされた時間を用いた。本研究では、時系列による感情の変化の分析を行う。

を用いた。ツイートされた時間を用いて、時系列の変化の分析を行う。分析するにあたり、本研究では、時系列ごとのツイートの感情にどのような変化があるのかの分析を行う。

### 6.1 時系列ごとのツイートの感情値の分析

時系列ごとのツイートの感情値から、感情の変化の分析を行う。

図1に、時系列ごとの実況ツイートの感情値を示す。横軸はツイートされた時間で、縦軸は感情値を示す。5つの折れ線グラフは各々の感情軸に関する感情値である。図2に「驚↔昂」を示す。図2の結果より、「驚↔昂」の軸に注目すると、全体的に、「昂」が高いことがわかる。また、個人で映画の見どころだと判断したシーンが「驚↔昂」の感情値が大きく変化していることがわかった。例えば、主人公が初登場した場面や戦闘シーンでは、「昂」が大きく振れており、主人公が敵に敗れてしまった場面では、「驚」が大きく振れていることがわかった。また、このことから、「驚↔昂」の軸は実況ツイートの感情を抽出する際に、適していると考えられる。図3に「哀↔喜」, 「厭↔好」を示す。図3より「厭↔好」の軸では「好」の感情に振れており、「哀↔喜」の軸では「哀」の感情が振れている。これらの結果

を合わせることによって哀しいけど好き、つまり感動したユーザが多いと推測される。「怖⇔怒」、「恥⇔安」の2つの感情軸では、映画を視聴しながら比較したところ、感情軸自体が合っていないことが推測される。これは、対象とした映画が戦闘がある場面が多く、映画を視聴した際に「怖⇔怒」、「恥⇔安」の軸の判断が困難であることが考えられる。解決策として、ジャンルによって感情の重み付けが考えられる。例えば、今回用いた戦闘する場面が多い映画では、「昂」に重みを大きく付けたり、ホラー関連の映画では、「怖」に重みを大きく付けたりすることによって、あるゆるジャンルに適應できることが考えられる。

## 7. まとめと今後の課題

本研究では、ツイートから Twitter 特有表現を考慮した感情を抽出する手法を提案した。ツイートの中から、テレビ番組で放送された映画の実況ツイートを対象にして、時系列毎に感情の変化の分析を行った。ツイートから感情を抽出するために、Yahoo!映画のレビューをテキストデータとし、新しく感情語辞書の構築を行った。また、Twitter 特有表現を考慮するために、各感情に特有表現の感情値の重みを設定し、感情の抽出を行った。Twitter 特有表現の重みの有用性を示すために評価実験を行った。実際に、映画の実況ツイートから我々が提案する感情抽出手法を用いて、感情を抽出し、時系列ごとに見た感情がどのように変化したのか分析を行った。

今後の課題を以下に示す。

- 感情語辞書の拡張

本研究では、10 軸の感情に分類されている中村 [3] の感情表現辞典と、Yahoo!映画の大量のレビューデータを用いて、新しい感情語辞書を構築し、感情の抽出を行った。今後は、感情抽出の精度を高めるために、さらに感情語辞書の拡張を行いたい。

- 重み付けの設定

Twitter 特有表現を考慮に入れるために、重み付けの設定を行った Twitter 特有表現が含まれていることで、ツイートに影響されるようにするため、重みを再度、検討していきたい。

- 感情軸ごとの重みつけ

実況ツイートの対象になっている番組によって、感情が変化することが考えられる。そこで、ジャンルごとに大きくなる感情に重みをつけることによって、番組の特徴が出やすくなるのではないかと考えられる。そのためには、被験者に、あらゆるジャンルがどういった感情を持つのか検証するためのユーザ実験を検討している。

- 他の実況ツールとの比較

本研究では、Twitter を用いて、テレビ番組を視聴しながらツイートする実況ツイートに注目した。しかし、Web 上でテレビ番組に対するコメントをするのは、Twitter だけでなく 2 ちゃんねるなど様々存在する。今後は、Twitter と他の実況ツールとの相違が存在するのか比較をしていきたい。

**謝辞** 本研究の一部は JSPS 科研費 24500134 の助成によるものです。ここに記して謝意を表します。

- [1] 若井祐樹, 熊本忠彦, 灘本明代. “映画に対する実況ツイートの感情抽出手法の提案”, 第 158 回データベースシステム研究会, 2013.
- [2] R.Plutchik. “The nature of emotions”. *American Scientist*, Vol.89, pp. 344–355, 2011.
- [3] 中村明:「感情表現辞典」. 東京堂出版, 1993.
- [4] 熊本忠彦, 河合由起子, 田中克己. “新聞記事を対象とするテキスト印象マイニング手法の設計と評価”, 信学論, Vol. J94-D, No.3, pp. 540–548, 2011.
- [5] 高岡幸一, 灘本明代. “名言のための多次元感情ベクトルの生成”, 第 4 回 Web とデータベースに関するフォーラム (WebDB2011), 9 pages, 2011 年 11 月.
- [6] 徳久良子, 乾健太郎, 松本裕治. “Web から獲得した感情生起要因コーパスに基づく感情推定”, 情報処理学会論文誌, Vol.50, pp. 1365–1374, 2009.
- [7] 堀宮ありさ, 坂野遼平, 佐藤晴彦, 小山聡, 栗原正二, 沼澤政信. “Twitter における発話者のリプライを用いたユーザの感情推定手法”, 第 4 回データ工学と情報マネジメントに関するフォーラム, 2012.
- [8] 水岡良彰, 鈴木優. “マイクロブログを用いた感情表現収集”, 情報科学技術フォーラム, FIT 2011, pp. 291–294, 2011.
- [9] 浅井洋樹, 秋岡明香, 山名早人. “きたあああああああああああああ!!!! 1 1 : マイクロブログを用いた教師なし叫喚フレーズ抽出”, 第 5 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2013), A4-4, 2013.
- [10] 山内崇資, 中野有紀子. “Twitter の感情分析に基づく TV 番組シーン探索システム”, 第 26 回人工知能学会全国大会, IC-R-5-3, 2012.
- [11] 中澤昌美, 帆足啓一郎, 小野智弘, “Twitter によるテレビ番組重要シーン検出及びラベル付与手法”, 第 3 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2011), F5-6, 2011.
- [12] 宮森恒, 中村聡史, 田中克己. “番組実況チャットを利用したテレビ番組のメタデータ自動抽出方式”, 情報処理学会論文誌: データベース (TOD), Vol.46, No.SIG10(TOD28), pp.59–71, 2005.
- [13] 佃洗撰, 中村聡史, 田中克己. “視聴者の反応に基づく動画検索および推薦システムの提案”, 第 19 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2011), 2011.