

顔文字の役割に着目したツイートの文の感情値抽出手法の提案

山本 湧輝[†] 若井 祐樹^{††} 熊本 忠彦^{†††} 灘本 明代^{††††}

[†] 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

^{††} 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

^{†††} 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

^{††††} 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

E-mail: [†]si071166@center.konan-u.ac.jp, ^{††}m1224008@center.konan-u.ac.jp

^{†††}kumamoto@net.it-chiba.ac.jp ^{††††}nadamoto@konan-u.ac.jp

あらまし Twitter などのマイクロブログサービスはその手軽さから多くのユーザが気軽にツイートを投稿している。そういったマイクロブログサービスには今起こったことや感じたことを投稿している場合が多いので、そのユーザの感情が現れやすい傾向がある。また、ツイートには顔文字という記号を用いて人の表情を表現するようなものが多く用いられている。しかしながら、同じ顔文字でも文との関係によって使われ方が異なる。そこで本論文では、文と顔文字の関係を分類するとともに、その関係を役割と呼び、役割を考慮した感情値抽出手法を提案する。

キーワード 感情抽出, Twitter, 顔文字

1. はじめに

近年、Twitter などのマイクロブログサービスが普及してきている。Twitter とは、140 文字以内のツイートと呼ばれる短文を投稿することができるツールである。短文しか投稿することができないということから、従来の SNS やブログと違い、手軽に投稿することが可能となった。そういったマイクロブログサービスには今起こったことや感じたことを気軽に投稿している場合が多いので、そのユーザの感情が現れやすい傾向がある。しかしながら、テキストでのコミュニケーションでは実際の会話と違い、ジェスチャーや表情といった非言語的コミュニケーションを用いることができないため、正確に自分の考えや感情を伝えることが難しい。そのため、ツイートではテキストで伝えることの難しい情報を伝えるための手段として顔文字が利用されている。顔文字とは「(´▽`)」 「(´ω`)」のような記号を用いて人の表情を表現したものである。顔文字を用いることで、言葉で説明することが困難な感情を表現することができる。したがって、顔文字が文に付与されることによりツイートの感情が変化すると考えられる。例えば、「そういう態度腹立たしい(´O`)」のように「腹立たしい」というきつい言葉を使っているが、顔文字が入ることで文のきつい言い回しが和らげられているように感じられる。このことから、顔文字が文に付与されることにより全体の感情が変化していると考えられる。しかしながら、同じ顔文字でも使い方によっては文に与える影響は変わってくる。例えば、先ほど用いた顔文字「(´O`)」を例に挙げると、「今日はとても楽しい(´O`)」のように楽しいという言葉に顔文字が強めているように感じられる。このように同じ顔文字でも文との関係によって使われ方が異なる。したがって、本論文では文の感情と顔文字の感情との関係によって感情の変化を分類する。また、このような文と顔文字との関係を役割と呼び、その役割を考慮した感情値抽出手法を提案

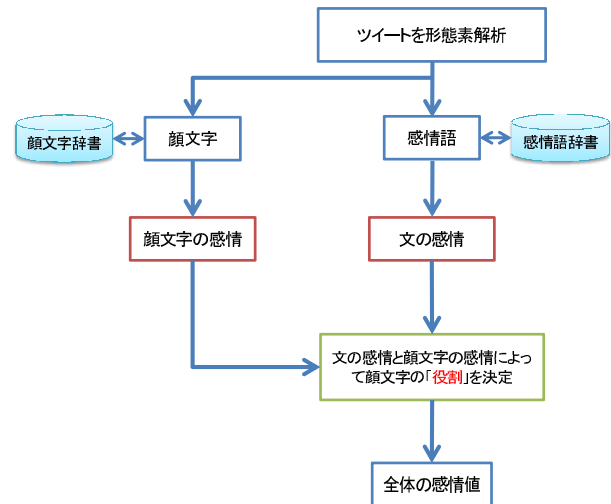


図1 感情抽出の流れ

する。これまで我々はニュースに対するツイートの感情抽出を行った [1]。ここでは、文のみの場合と顔文字を合わせた場合の感情値を比較することで、顔文字の役割を分類した。しかしながら、顔文字の機能を分類し役割を決定したが、それぞれの役割が文に与える影響の大きさは求めていなかった。そこで本論文では、文と顔文字の感情ごとに重みを決定し、その重みを考慮した感情値抽出手法を提案する。以下と図1に役割と重みを考慮した感情抽出手順を示す。

(1) Twitter からツイートを取得する。

(2) 取得したツイートを形態素解析し、単語を抽出する。

(3) 感情語辞書とのマッチングにより各単語の感情値を算出する。

(4) 顔文字を含んでいるツイートでは、文の感情と顔文字の感情から顔文字の役割を決定し、ツイート全体の感情値を算出する。また、顔文字を含んでいない場合は (3) の感情値をツ

イト全体の感情値とする。

以下、第2章で関連研究について、第3章で感情抽出手法について、第4章で顔文字の役割について、第5章で評価実験について、第6章でまとめと今後の課題について述べる。

2. 関連研究

近年、感情表現を抽出する研究が行われている。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のモデル[2]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。

熊本ら[3]は新聞記事を読んだ人々が記事に対して受けた印象をアンケート調査で調べ、分析することにより、6本の感情軸を提案している。各感情軸を反義語関係となる2種類の印象語で構成しており、「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本が提案されている。また、これらの感情軸をベースにした感情値抽出手法[4][5]も提案しており、新聞記事に対して有効であることを示している。しかしながら、本論文ではTwitterに着目し、ツイートから感情を抽出するという点で異なる。

顔文字に対する研究は多数存在する。風間ら[6]は顔文字を自動抽出する手法と、さらにその多様な顔文字を用いてツイートにおける顔文字の使用状況を可視化する顔文字クラウドを提案し、感情の変化を可視化することを提案している。本論文では文と顔文字に着目して感情抽出を行っている点が異なる。顔文字の感情分析に関する研究には顔文字のみと顔文字と文を合わせた感情分析に分類される。文献,[7],[8],[9],[10],[11],[12]は顔文字のみを対象としている。本論文では顔文字と文を合わせた感情分析を行うため、顔文字のみの感情分析とは異なる。顔文字と文を合わせて感情分析を行っている研究とは類似している。しかしながら、中村[13]は顔文字をポジティブな顔文字とネガティブな顔文字に分けて文との関係により感情分析を行っている。本論文では顔文字も多次元的な感情を持っているとして分析している点が異なる。加藤ら[14]は携帯メールにおけるコミュニケーションの中で怒りの感情に着目し顔文字が付与されることによって感情へ及ぼす影響を調べた。本論文では顔文字の及ぼす影響を他の感情についても分析している点が異なる。江村ら[15]は文の感情を推定し、その結果から適切な顔文字の推薦を行っている。本論文では顔文字が文に及ぼす影響を調べている為異なる。篠山ら[16]は顔文字が文に付与されることにより感情がどのように変化するかを調べている。しかしながら、感情の強弱については分析していない為、本論文とは異なる。

3. 感情語辞書と顔文字辞書

3.1 感情軸と感情語辞書

本論文ではまず、感情抽出の対象とする感情軸を決定する。様々な感情軸が提案されているが、本論文では中村明の感情表

表1 感情表現の例

感情軸	感情表現		
喜	喜び	明るい	楽しみ
好	愛	好き	好み
安	安心	溜め息	ほっと
哀	悲しさ	哀しさ	悲しみ
厭	厭	嫌い	不快
怖	不気味	怖い	恐い
怒	怒り	憤り	腹立ち
恥	恥ずかしい	決まりが悪い	あられもない
昂	焦る	苛立つ	気を揉む
驚	驚く	魂消る	驚き入る

表2 感情語辞書の例

感情軸	感情語	感情値	感情語	感情値
喜	楽しみだ	0.91	大切だ	0.54
好	にこやかだ	0.64	アルバム	0.51
安	安心	0.45	心地よい	0.81
哀	悲しい	0.45	憂鬱	0.82
厭	嫌だ	0.82	嫌いだ	0.75
怖	怖い	0.45	恐怖	0.64
怒	怒る	0.87	怒鳴る	0.54
恥	恥ずかしい	0.72	恥じらい	0.87
昂	昂る	0.45	すごい	0.57
驚	驚く	0.71	驚愕	0.54

現辞典[17]で分類されている「喜」、「好」、「安」、「哀」、「厭」、「怖」、「怒」、「恥」、「昂」、「驚」の感情軸を用いる。中村明の感情表現辞典では各感情軸を表現する語を感情表現としている。10軸の感情とそれに含まれている感情表現の例を表1に示す。しかしながら、感情表現辞典に登録されている語は定量化されていないため定量化された辞書を構築する必要がある。そこで、熊本ら[4]が提案している感情辞書構築システムを用いて感情語辞書を構築する。熊本らのシステムは各感情軸を表す語を印象語として、その印象語と文書集合内の共起している語を感情語として登録し、その共起頻度から感情値を算出するものである。

ここで、本論文で構築する感情語辞書を生成する際に用いる文書集合を決定する。本論文ではツイートを対象としているが、ツイートは短い文が多いため表記揺れや文法の乱れが多くあり、共起関係が正しくないと考え、ツイートと同様に口語的ながら、表記揺れや文法の乱れが少ない映画のレビューを文書集合として辞書構築を行った。その結果今回用いる10本の感情軸に対して32,322個の感情語と感情値のペアを感情語辞書に登録した。構築した感情語辞書の例を表2に示す。

次に、構築した感情語辞書を用いてツイートの文の感情を決定する。まず、ツイートに対して形態素解析エンジンJuman^(注1)

(注1) : <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

表 3 実験で用いた顔文字の例

顔文字	回数	割合
(^o^)	2832	0.160
(^^)	935	0.053
(>_<)	917	0.052
('_`)	742	0.042
(^-^)	527	0.030
('▽`)	504	0.029
('ω`)	499	0.028
(;ω;`)	413	0.023
(^-▽`)	389	0.022
(^ω^)	361	0.020

を用いて形態素解析し、ツイート中の形態素と感情語とのマッチングを行う。しかしながら、表記揺れが発生する場合は感情語が正しくマッチングしない。例えば、感情語辞書に「楽しい」という言葉が感情語として登録されているが、ツイートが「今日は楽しもう」だと形態素に分けた際に「楽しもう」になるためにマッチングしない。このような問題を解決するために Juman には表記揺れを緩和する目的で代表表記という基本語彙に付与されている ID のようなものが存在する。ツイート中の形態素を代表表記に統一し感情語辞書との単語のマッチングを行う。そのために、感情語辞書に記載されている語の中で代表表記を有していて、代表表記が未登録なものを辞書に追加した。また、形態素に同形が存在する場合は曖昧な単語なので考慮しないものとする。

また、ツイートに否定語が入っている場合に正しい感情値が抽出することができないという問題が発生する。例えば、「楽しくない」のような語の場合だと Juman による出力結果は形容詞「楽しい」と形容詞性述語接尾辞「ない」に分けられる。本論文では感情語は単語ベースのマッチングを行っているので「楽しい」という形容詞が感情語辞書とマッチしこの文の感情は「喜」になる。しかしながら、否定されているので感情語を打ち消さなければならない。熊本ら [4] は単語の印象を考慮するため、品詞を変換する目的で、Juman の出力結果を変換するルールを用いている。本論文ではその中で、否定語についてのルールを適用することで否定語の判定を行い、否定語を含む場合の感情語を打ち消す。

3.2 顔文字辞書の構築

次に、顔文字辞書の構築手法を述べる。顔文字の役割は顔文字のみで決まっているのではなく、文と顔文字の感情を複合的に考慮することで顔文字の役割が決定する。したがって、感情軸毎に顔文字を追加する必要がある。既存の顔文字辞書では本論文の感情軸ごとに分類されていない為、ユーザ実験を行い顔文字辞書を作成した。

3.2.1 実験条件

顔文字辞書作成のため、Twitter から取得した顔文字付きの 17,647 ツイートの中で出現頻度が高い顔文字 100 件を用いて、10 名の被験者によるユーザ実験を行なった。また、今回実験で用いる顔文字 100 件で収集した顔文字の 92% の割合を占める。実験で用いた顔文字の出現回数と全体の割合の例を表 3 に

表 4 顔文字辞書の例

感情	顔文字		
喜	(^o^)	(^^)	(^-^)
好	(^o^)	('▽`)	(≥▽≤)
安	('▽`)	('ω`)	('_`)
哀	(>_<)	('_`)	(ToT)
厭	('_`)	('_`)	(;ω;`)
怖	(;ω;`)	(((((;Π°))))))	(;Π)
怒	('Π`)	('ω`)	('Π`)
恥	(>_<)	(^-^;)	("ω")
昂	('Π`)	(≥▽≤)	('_ω`)
驚	(°Π°)	(^-▽`)	(°_°)

表 5 パーツごとに割り当てられた感情の例

目	感情
(^^)	喜 好 安
(><)	哀 厭 驚
(. .)	哀 厭 恥
(° °)	驚
(≥≤)	喜 好 昂

示す。実験の方法は、顔文字を見てもらいその顔文字がどの感情軸に当てはまるか最大 3 つの感情から選んでもらい、1 番に選択した感情を 3 点、2 番に選択した感情を 2 点、3 番に選択した感情を 1 点としてアンケート結果を収集した。

3.2.2 実験結果

各感情に割り当てられた顔文字の例を表 4 に示す。本論文では、顔文字は複数の感情を持っている場合もあるとしているため、複数の感情軸に割り当てられている顔文字も存在する。また、「怖」、「怒」、「恥」などの感情軸によっては顔文字の数が少なくなっている。理由としては、Twitter でよく使われている顔文字を実験に使用したため、顔文字のデータに偏りができたのだと考えられる。また、登録されている顔文字が少ないために顔文字辞書の拡張を行った。基本的に顔文字は顔を構成するためのパーツに分かれている為、先ほどの実験で各感情に割り当てられた顔文字に対して顔文字のパーツに着目する。本論文では、基本的な顔文字のパーツのみに焦点を当て分析を行った。今回分析を行ったパーツは「目」、「口」、「眉」、「頬」の 4 パーツとする。その結果、各パーツにはいくつかの感情を表すものがある事がわかった。その例を表 5 に示す。表 5 より、感情を表すパーツは「目」のパーツが多く見られた。これは日本人は人の感情を理解するときに目に着目する傾向がある [18] からと考えられる。このパーツを参考に顔文字辞書に顔文字を手で追加した。その結果、顔文字辞書に登録した顔文字を 500 個まで増やした。

表 6 役割と特徴

役割	特徴の例	ツイートの例
強調	同じ表現を繰り返すもの	きたあああゝ(´▽`)/
	文の意味=顔文字の意味	好きだな(*´▽`*)
	強調記号+顔文字	観る♪(´ε`)
	文の意味≠顔文字の意味	かわいいわあゝ(;▽;)ノ
自嘲	否定形+(笑)	できてない(笑)
弛緩	命令形+やわらかめな顔文字	飲み会来るなよ(´O`)

4. 顔文字の役割

本節では、文と顔文字の関係から顔文字の役割を決定するために実験を行う。また、文と顔文字の感情ごとに重みも決定する。

4.1 予備実験

被験者は10代から50代の男女各10人の合計100人である。実験手順は以下のとおりであり、(1) 顔文字を含んだ100件のツイートとそのツイートから顔文字を除いた100件のツイートの双方を見てもらい、それぞれのツイートに対し当てはまると感じた感情を10段階(1~10)で評価してもらった、(2) 当てはまらなかった感情は0点として扱い、ツイート毎に全被験者分の合計値を求めた、(3) 文の感情と顔文字の感情の組み合わせ毎に、該当するツイートの合計値の平均値を求めた。

4.2 実験結果

顔文字がある場合と顔文字がない場合の結果を比べたところ、顔文字がある場合に感情が強められている時と、感情が弱められている時、感情が変化しない時があることがわかった。

これまでに我々は文のみの場合と顔文字を合わせた場合の感情値を比較することで、顔文字の役割を3つに分類した[1]。以下に3つの役割を示す。

強調 顔文字が文の意味(良い意味、悪い意味含め)をより強めている。

弛緩 顔文字が文の意味を少しでも弱めている、和らげている。

自嘲 顔文字があることで、自分に呆れて笑うさま、自分で自分をつまらぬものとして軽蔑すること・・・といったニュアンスを感じる。

顔文字の役割と特徴を表6に示す。本論文では、自嘲は弛緩と包含関係にあると考え村上ら[19]の提案している「強調」、「弛緩」の役割を用いる。次に、本論文で用いる「強調」、「弛緩」の役割を説明する。文の感情を強めている役割のことを「強調」、文の感情を弱めている役割のことを「弛緩」と呼ぶ。例えば、「喜」の感情の中で大きく感情が強められていた「上原すごい(=^▽^)」のとき、顔文字がない場合に「喜」の感情値の合計が138なのに対して、顔文字が含まれている場合における「喜」の感情値の合計は194と56大きくなっている。このような役割を「強調」と呼ぶ。また、「哀」の中で大きく感情が弱められていた「ホント勉強ばかり嫌になる(^○^)」だと、顔文字がない場合における「哀」の感情値の合計が

表 7 各感情ごとの重み

文 顔文字	喜	好	安	哀	厭	怖	怒	恥	昂	驚
喜	1.34	1.23	1.13	0.00	0.78	0.76	0.45	0.91	1.04	0.68
好	1.26	1.23	1.10	0.12	0.77	0.83	0.39	1.05	1.00	0.65
安	1.16	1.24	1.24	0.15	0.78	0.88	0.40	1.11	0.91	0.76
哀	0.59	0.61	1.03	1.34	1.04	0.93	0.42	0.98	0.83	1.05
厭	0.59	0.61	1.03	1.23	1.10	0.93	0.42	0.98	0.83	1.05
怖	0.35	0.75	0.73	0.56	1.32	1.03	1.63	1.25	1.05	0.57
怒	0.10	0.85	0.75	0.56	1.07	1.02	1.62	1.13	0.98	0.70
恥	1.06	1.41	0.98	0.45	1.01	0.60	0.53	1.41	1.04	0.77
昂	0.93	1.21	0.73	0.49	0.80	0.72	1.32	1.35	1.04	0.82
驚	0.24	0.58	0.75	0.43	0.89	0.81	0.64	0.73	1.01	1.67

102なのに対して、顔文字含まれている場合における「哀」の感情値の合計は60と42小さくなっている。このような役割を「弛緩」と呼ぶ。

次に今回の実験の結果から顔文字毎の役割を決める。しかしながら、同じ役割でも感情の組み合わせによっては感情の変化する量が異なる。例えば、「今親元を離れたばっかだから心配ごと多い(´▽`)」と、「付き合ってたときより別れたあとのが辛い(´\_`)」という文を比べると、両方とも文の感情は「哀」を示し、顔文字の感情は前の文は「喜」で後の文は「驚」を示す。両方の文とも顔文字の役割は「弛緩」を表すが、実験結果より、感情の変化は前の文が156小さくなっている。それに対して、後の文は56小さくなっている。このようなことから、文と顔文字の感情の組み合わせによって感情の変化する量は異なることがわかった。そこで、文と顔文字の全ての組み合わせに対して重みを決定する必要がある。我々は、顔文字は文の感情を補助するものだと考えており、その為顔文字の重みをツイートの感情に掛けることによりツイート全体の感情値を決定する。そこで、重みを求める式を以下に示す。

$$F_i = 1 + \frac{y_i - x_i}{\max |y_i - x_i|} \quad (1)$$

文と顔文字の感情の組み合わせを*i*とし、*y_i*は顔文字を含む文の感情値の平均を示し、*x_i*は顔文字を含まない文の感情値の平均を示す。また、 $\max |y_i - x_i|$ は各文と顔文字の組み合わせの中で絶対値を取った際の変化量の最大のものをさす。

次に式(1)の式を用いて各感情の組み合わせごとに計算した重みを表7に示す。重みの値により顔文字の役割を決定する。重みの値が1より大きいものは、顔文字を含んでいない文の感情値より顔文字を含んでいる文の感情値のほうが大きいため、顔文字が文の感情を強めている。そのような顔文字の役割を「強調」とする。また、重みの値が1より小さいものは顔文字を含んでいない文の感情値より、顔文字を含んでいる感情値の方が小さかったので顔文字が文の感情を弱めている。そのような顔文字の役割を「弛緩」とする。

表7の結果から、感情の変化を見るとポジティブな感情に対しての変化より、ネガティブな感情に対して感情を弱める傾向が強かった。戸梶ら[20]によると情緒表現記号(顔文字)は、イ

表 8 年代別の比較

	喜	好	安	哀	厭	怖	怒	恥	昂	驚
10代	-216	-81	-124	-154	-79	246	-13	-20	19	75
50代	111	7	95	128	182	45	41	62	66	8

ンフォーマルな文書内容において使用することはポジティブな印象を受け手に対して与える可能性が高いとしている。Twitter というインフォーマルな文書において顔文字はポジティブな印象を受けやすい傾向があるとしていることから、ネガティブな感情に対しての「弛緩」において高い値を示す結果になったと考えられる。また表 7 の結果より、「昂」「驚」「恥」の感情に関しては文によって感情の変化の仕方が大きく異なっていた。これは、これらの感情が必ず他の感情とセットになる傾向があるからだと考えられる。例えば、「テンションが上がってきた」と「赤点取ったら怒るからね!」の顔文字を含んでいない 2 つの文を見てみると結果は、前の文が「喜」と「昂」がセットになっているのに対して、後の文は「怒」と「昂」がセットになっている。このように「昂」「驚」「恥」は他の感情とセットになって 2 番手の感情になっている場合が多い、その場合前の感情の影響に引きずられるためこの 3 つの感情に対しては値が正しく取れていなかったのだと考えられる。

また、顔文字の認知度の違いによって感情の変化に違いがないかを判断するために 4.1 節の実験に対して年代別に感情の変化を分析した。実験の被験者は 10 代、20 代、30 代、40 代、50 代の男女各 10 人なのでその中でも特に顔文字を用いてメールやマイクロブログを使用している 10 代とあまり使用していないであろう 50 代を比較してみる。10 代と 50 代の各感情軸ごとの変化を表 8 に示す。

表 8 の結果より、頻繁に顔文字を使う 10 代に関しては顔文字の表情を理解していると考えられる。しかしながら、50 代のあまり顔文字を使わないであろう年代になると、顔文字がついているということだけを理解して判断しているため、顔文字の表情までは理解しないで値を付けているのがわかる。

5. 評価実験

提案手法の有用性を確かめるため、評価実験を行った

5.1 実験方法

Twitter から取得した顔文字付きのツイート 100 件を取得し、各ツイートごとに、被験者 3 人が与えた評価結果と、顔文字を考慮していない場合のシステム、提案手法である顔文字を考慮したシステムがそれぞれに算出した感情値を比較する。

まず、被験者にはツイートを見てもらい当てはまると感じた感情を 10 段階 (10 点, 9 点, …, 1 点) で評価してもらう。評価してもらった感情の合計がある閾値以上の時このツイートは感情を持っているものとした。同様に、感情抽出システムを用いて各ツイートの感情値を求める。算出された感情値が閾値以上の場合に、このツイートは感情を持っているものとした。閾値は各感情軸の平均を元にすべてのツイートに対して感情が 3

表 9 評価実験の比較

	適合率	再現率	F 値
顔文字を考慮しない	24%	37%	29%
顔文字を考慮する	38%	44%	40%

表 10 感情軸別の比較

	顔文字を考慮しない			顔文字を考慮する		
	適合率	再現率	F 値	適合率	再現率	F 値
喜	27%	18%	21%	52%	35%	42%
好	33%	25%	29%	62%	65%	63%
安	14%	35%	20%	18%	53%	26%
怒	6%	67%	11%	50%	33%	40%
昂	37%	47%	42%	36%	38%	37%
哀	30%	33%	31%	73%	33%	46%
厭	17%	29%	22%	26%	29%	28%
恥	27%	43%	33%	27%	43%	33%
怖	22%	68%	34%	33%	53%	41%
驚	22%	50%	31%	33%	44%	38%

個以上にならないように決定した。

5.2 実験結果

顔文字を考慮していない場合と顔文字を考慮している場合の適合率、再現率、F 値を表 9 に示す。実験結果より顔文字を考慮した場合の方が顔文字を考慮していない時よりも適合率、再現率、F 値それぞれ向上し良い結果となった。また、感情軸ごとの適合率、再現率、F 値を表 10 に示す。

感情軸ごとに見ると、顔文字を考慮した方が良い結果になっている感情軸と、あまり変わらない結果となっている軸があった。「喜」、「好」、「安」、「哀」、「厭」、「怖」の 6 軸は良い結果となっていた。理由としては、この 6 軸はポジティブな軸なのか、ネガティブな軸なのかははっきりしているということがある。そのような軸に関しては顔文字の影響を大きく受ける結果になった。また、「昂」、「驚」、「恥」の軸は良くない結果となっていた。その理由は 2 つあると考えられる。

1 つ目は、4.2 節にあるように「昂」「驚」「恥」の感情軸は他の感情とセットになって 2 番手の感情になっている場合が多いので、その場合前の感情の影響に引きずられるためこの 3 つの感情に対しては値が正しく取れていなかったと考えられる。

2 つ目は、「昂」「驚」「恥」の軸はポジティブな軸なのかネガティブな軸なのかの判断がつかないために、それ単体では顔文字の影響を受けないということが考えられる。

評価実験の結果から「喜」、「好」、「安」、「哀」、「厭」、「怖」などの顔文字の影響を受けやすい軸と、「昂」、「驚」、「恥」などの顔文字の影響を受けにくい軸があるということがわかった。

6. まとめと今後の課題

本論文では、顔文字の役割を考慮してツイートの感情値を抽出をするはじめの一歩として、感情語辞書の構築と顔文字辞

書の作成のためのユーザ実験を行った。その結果感情語辞書に 32,322 個の感情語と感情値、顔文字辞書には 500 個の顔文字を登録した。また、文と顔文字の感情から顔文字の役割を決定するためにユーザ実験を行なった。その結果を用い、文と顔文字の感情ごとに役割の重みを計算し分析を行った。そして、実験で求めた役割と重みの精度を測るために評価実験を行なった。

今後の課題として、テキストの感情辞書の拡張において、出現頻度の低い重要な単語の抽出手法の検討、顔文字の影響を考慮した感情軸の検討、最後に、顔文字の役割とその重みを考慮した感情抽出の機能をシステムの中に組み込んでいきたい。

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。

- [15] 江村優花, 関洋平. テキストに現れる感情, コミュニケーション, 動作タイプの推定に基づく顔文字の推薦, 情報処理学会研究報告. DD, [デジタル・ドキュメント] 2012-DD-85(1), 7 pages, 2012.
- [16] 篠山学, 松尾朋子. 顔文字を考慮した対話テキストの感情推定に関する研究, 香川高等専門学校研究紀要 1, pp. 151-153, 2010.
- [17] 中村明. 感情表現辞典. 東京堂出版, 1993.
- [18] <http://lynx.let.hokudai.ac.jp/CSM/interview/09.html>
- [19] 村上浩司, 山田薫, 萩原正人, “顔文字情報と文の評価表現の関連性についての一考察”, 第 17 回言語処理学会発表論文集, pp.1155-1158, 2012.
- [20] 戸梶亜紀彦. コンピューター上でのコミュニケーションに見られる情緒表現に関する研究, 広島県立大学紀要, 1997, pp.125-139.

文 献

- [1] 若井祐樹, 田中美羽, 熊本忠彦, 灘本明代, “顔文字を考慮したニュースに対するツイートの感情抽出手法の提案”, 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, No.D9-5, 2013.
- [2] R.Plutchik. The nature of emotions. American Scientist, Vol.89, pp. 344-355, 2011
- [3] T.Kumamoto. Design of Impression Scales for Assessing Impressions of News Articles, Lecture Notes in Computer Science, LNCS6193, Springer, pp. 285-295, In International Workshop on Social Networks and Social Media Mining on the Web (SNSMW'10), 2010.
- [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, 電子情報通信学会論文誌, Vol.J94-D, No.3, pp. 540-548, 2011.
- [5] 熊本忠彦, 河合由起子, 張建偉. ユーザ印象評価データの分析に基づく印象マイニング手法の設計と評価, 情報処理学会論文誌データベース, Vol. 6, No. 2, pp. 1-15, 2013.
- [6] 風間一洋, 榊剛史, 鳥海不二夫, 篠田孝祐, 栗原聡, 野田五十樹. 顔文字に着目したツイートの感情変化の分析, 第 6 回 Web とデータベースに関するフォーラム, 2013
- [7] 中村純平, 池田剛, 乾伸雄, 小谷善行. 対話システムにおける顔文字の学習, 情報処理学会研究報告. 自然言語処理研究会報告 2003(23), pp. 169-176, 2003.
- [8] M.Ptaszynski. 顔文字処理-取るに足らない表現をコンピュータに理解させるには-, 情報処理 53(3), pp. 204-210, 2012.
- [9] CHO HEERYON, 稲葉利江子, 石田亨, 高崎俊之, 森由美子. 絵文字コミュニケーションにおけるセマンティクス, 情報処理学会研究報告. ICS, [知能と複雑系] 2006(110), pp. 1-8, 2006.
- [10] 川上正浩. 顔文字が表す感情と強調に関するデータベース, 大阪樟蔭女子大学人間科学研究紀要 7, pp. 67-82, 2008.
- [11] 伊藤淳子, 宗森純. 擬人化エージェントを介したチャットにおける顔文字と対話雰囲気との関連性の分析, 情報処理学会研究報告.GN, [グループウェアとネットワークサービス] 2008(31), pp. 127-132, 2008.
- [12] 加藤尚吾, 加藤由樹, 小林まゆ, 柳沢昌義. 電子メールで利用される顔文字から解釈される感情の種類に関する分析, 教育情報研究: 日本教育情報学会学会誌 22(4), pp. 31-39, 2007.
- [13] 中丸茂. 顔文字が文章の信頼度に及ぼす影響, 言語・音声理解と対話処理研究会 37, pp. 173-176, 2003.
- [14] 加藤由樹, 加藤尚吾, 赤堀侃司. 携帯メールを使用したコミュニケーションにおける怒りの感情の喚起に関する調査, 教育情報研究: 日本教育情報学会学会誌 22(2), pp. 35-43, 2006.