

顔文字を対象とする印象と感情の相互変換手法の提案

～ 読み手が感じる「印象」と読み手が推測する書き手の「感情」を対象として ～

On a Method for Mutually Converting Impressions and Emotions Expressed by Emoticons

— Focusing on Readers' Impressions and Writers' Emotions Inferred by Readers —

熊本 忠彦[†]

Tadahiko Kumamoto[†]

1. まえがき

近年、顔文字を対象とする研究が活性化している。人工知能学会誌では「顔文字の科学—Web 上の非言語表現・行動に関する新研究分野の誕生」という特集 [1] が生まれ、顔文字を対象とした様々な研究に関する概観 [2] が示されるとともに、工学的・心理学的側面から見た顔文字の分析に関する研究動向 [3] や顔文字の推薦技術に関する研究動向 [4] など紹介されている。

顔文字に関する研究を感性的側面で分類すると、感情に関する研究と印象に関する研究に分けられる。感情に関する研究は、対象となる感情の種別に応じて、さらに顔文字の送り手（書き手）の感情 [5][6][7][8][9]、顔文字の受け手（読み手）が推測した書き手の感情 [10]、読み手の感情の変化 [11] に分けることができる。同様に、印象に関する研究も、対象となる印象の種別に応じて、書き手に対する読み手の印象 [12] とメッセージ内容に対する読み手の印象 [13] に分けることができる。このように、様々な角度から顔文字の感性的側面に関する研究が行われているが、そのほとんどの研究は印象か感情のどちらかを対象としており、この 2 種類の感性を同時に扱うような研究は少なく、両者の依存関係を定式化しようという研究は見当たらない。

そこで本論文では、読み手が感じる顔文字の印象と顔文字を見て読み手が推測する書き手の感情を定量化し、どのような関係があるかを分析するとともに、両者の依存関係を相互に定式化する。なお、本研究では、Twitter を対象としており、必ずしも面識があるとは限らないフォロワー他への顔文字の使用を前提としている。

ここで、本研究に対するアプローチを示す。まず、予備的なアンケート調査により、一定以上の頻度で閲覧、投稿している 20 代～50 代の Twitter ユーザ（ヘビー閲覧ユーザとヘビー投稿ユーザ）を選出するとともに、ヘビー投稿ユーザが Twitter 上で普段用いている顔文

字を収集し、出現頻度上位の顔文字を抽出する。次に、この出現頻度上位の顔文字を用いて、さらに 2 種類のアンケート調査を行う。一つはそれぞれの顔文字を見てどのような印象を感じるかを問うアンケート調査であり、もう一つはそれぞれの顔文字に書き手のどのような感情が込められていると感じるかを問うアンケート調査である。いずれのアンケート調査もヘビー閲覧ユーザを対象とし、インターネット調査会社を介して Web 上で実施する。以上のアンケート調査の結果を整理することで得られる顔文字印象データと顔文字感情データに対し、因子分析やクラスタ分析を行うことで、印象どうし、感情どうし、印象と感情の間での関係を明らかにするとともに、重回帰分析を行うことで、印象と感情の依存関係を相互に定式化する。

なお、本論文で対象とする感情は、国語学者の中村が提唱した 10 種類の基本感情「喜、好、安、哀、厭、怖、怒、恥、昂、驚」[14] であり、回答者には評価しやすいように、「喜び、好き、安心、悲しい、嫌い、怖れ、怒り、恥ずかしい、高揚、驚き」の 10 個を感情に関する評価項目として提示する。一方、本論文で対象とする印象は、著者が提案している 8 種類の基本印象「攻撃的/不愉快、ネガティブ、感じの良い、楽しい/愉快、ポジティブ、ほのぼの、鬱陶しい、怖い」[15] であり、この 8 個を印象に関する評価項目として提示する。

以下、2. で関連研究について述べ、本研究の新規性を明らかにする。3. で顔文字印象データと顔文字感情データの収集について述べ、4. で様々な多変量解析を行うことで、印象どうし、感情どうし、印象と感情の間での関係を明らかにするとともに、印象と感情の依存関係を相互に定式化する。5. で提案手法の精度を評価し、その有効性を検証する。最後に 6. で本論文のまとめと今後の課題について述べる。

2. 関連研究

近年、顔文字の感情的側面に関して様々な角度から研究が行われている。本章では、顔文字の印象や感情に関する先行研究について述べ、本研究の新規性を明

[†] 千葉工業大学 情報科学部 情報ネットワーク学科
Chiba Institute of Technology,
2-17-1, Tsudanuma, Narashino, 275-0016 Chiba, Japan

らかにする。

書き手の感情を対象とする研究としては、まず、顔文字の推薦に関する研究 [5][6][7][8] が挙げられる。顔文字推薦では、書き手が表現したいと思っている感情や表情などに応じて顔文字を推薦するため、書き手の感情を対象としていることが多いが、その種類は様々である。例えば、Suzuki ら [5] が対象とした感情は、Plutchik の感情の輪 [16] としてモデル化されている 8 種類の基本感情「joy, trust, fear, surprise, sadness, disgust, anger, anticipation」とその強度に対して定義された 29 個の感情語に、メールから抽出した 12 個の感情語を加えたものとなっている。江村ら [6] は、中村の 10 種類の基本感情 [14] をベースに拡張した 27 種類の感情カテゴリを対象としているが、顔文字の推薦にあたっては、この感情カテゴリに加え、コミュニケーションカテゴリ（感謝、挨拶、意気込みなど 9 種類）と動作カテゴリ（睡眠の 1 種類）も考慮している。卜部ら [7] は、中村の 10 種類の基本感情 [14] をそのまま用いている。吉田ら [8] は、Plutchik の 8 種類の基本感情 [16] を用いているが、基本感情を推定するための感情語を中村の感情表現辞典 [14] から選出している。一方、顔文字辞書の構築に関する研究 [9] でも書き手の感情を対象としていることが多い。例えば、Ptaszynski ら [9] は、顔文字の構成要素に基づいて顔文字を抽出し、その顔文字が表現している感情を特定する CAO システムを提案しているが、対象となっている感情は中村の 10 種類の基本感情 [14] であり、様々な顔文字を Web 上で収集する際に、Web ページ作成者によって付与されている各顔文字の感情クラス名も取得し、感情表現辞典 [14] と照合することで、顔文字が表現している感情を特定している。

読み手が推測した書き手の感情を対象とする研究としては、心理学的側面からの研究 [10] が挙げられる。例えば、廣瀬ら [10] は、携帯メールでメッセージを受け取るとき、顔文字もしくは絵文字の付与によって受信者への感情の伝わり具合がどのように変化するかを顔文字と絵文字で比較している。このとき対象となった感情は「喜び、悲しみ、怒り、不安」の 4 項目であった。

読み手の感情の変化を対象とする研究に関しても、心理学的側面からの研究 [11] が多い。例えば、荒川ら [11] は、怒りを感じている相手から謝罪のための携帯メールが送られてきたときに、顔文字が付与されているか否かによってメールを受け取った人の感情がどう異なるかを調べ、検討している。このとき対象となった感情は「怒り、親しみ、礼儀正しさ、反省の度合い」の 4 項目であった。

表 1: 顔文字の使用頻度に関する調査結果

		(1)	(2)	(3)	(4)	(5)	小計
男性	20代	30	33	14	11	23	111
	30代	26	30	20	21	16	113
	40代	28	40	15	12	16	111
	50代	18	29	15	25	24	111
女性	20代	50	35	10	9	7	111
	30代	38	46	9	12	7	112
	40代	26	39	20	18	8	111
	50代	32	33	22	15	9	111
	小計	248	285	125	123	110	891

(注) (1) よく使う, (2) たまに使う, (3) あまり使わない, (4) ほとんど使わない, (5) 全く使わない

読み手の印象に関する研究は、書き手に対する印象を対象としたもの [12] とメッセージ内容に対する印象を対象としたもの [13] に分けられる。例えば、竹原ら [12] は、顔文字の有無によって書き手であるメール送信者の印象がどのように異なるかを調べている。竹原が対象とした印象は、外交性・友好性・誠実性・衝動性（神経質性）・開放性・専門性・真実性に関わる 16 個の質問項目 [17] から構成されていた。一方、田口 [13] は、顔文字を付与したり、文頭や文末にあるひらがなを小文字にしたりすることで、メール文（お礼文、挨拶文、依頼文）の印象がどう変わるかを印象評定に基づいて調べている。このとき対象となった印象は、田口が独自に作成した「積極的な-消極的な、責任感の強い-無責任な、親しみやすい-親しみにくい、真面目な-不真面目な、親切な-不親切な、知的な-知的でない、外向的な-内向的な」などの 14 個の形容詞対であった。

読み手の感情の変化とメッセージに対する読み手の印象の両方を扱った研究 [18][19] もわずかながらある。例えば、荒川らは、文献 [18] において、メールに顔文字を付与することで、受信者の感情がどのように変化するかや、受信者の感情が変化するとき、受信者が顔文字から受ける印象と受信者の感情の変化量との間に何らかの関係があるのかとといったことを調べている。このとき対象となった印象は「感情理解度（送信者が受信者の感情をどの程度理解していると思うか）、親しみやすさ、礼儀正しさ、誠実さ、好ましさ」の 5 項目であり、対象となった感情は「悲しみ、怒り、不安、喜び」の 4 項目であった。荒川らは、各印象項目に付けられた得点を合計することで、「良い印象」という総合指標を作り、それぞれの感情の変化量との間にどのような相関関係があるかを調べた。その結果、印象の良さ喜びとの間には正の相関があり、怒りとの間には負の相関があることや、悲しみ・不安の間には相

表 2: 出現頻度 5 以上の顔文字

順位	頻度	顔文字	累積数	累積割合	順位	頻度	顔文字	累積数	累積割合
1	100	(^^)	100	6.2%	28	7	(^ω^)	640	39.9%
2	62	^^	162	10.1%	28	7	\ (^^) /	647	40.4%
3	61	m(_ _)m	223	13.9%	28	7	? (' ω ') ?	654	40.8%
4	47	^_^	270	16.8%	28	7	(≧▽≦)	661	41.2%
5	35	(>_<)	305	19.0%	28	7	(^ ^) ヽ	668	41.7%
6	33	(^_^)	338	21.1%	28	7	(^O^)	675	42.1%
7	27	\ (^o^) /	365	22.8%	28	7	!(^^)!	682	42.5%
8	24	(*^^*)	389	24.3%	35	6	^o^	688	42.9%
9	22	(^o^)	411	25.6%	35	6	><	694	43.3%
10	21	(^-^-)	432	26.9%	35	6	(' _ ')	700	43.7%
11	18	(^ω^)	450	28.1%	35	6	(o^o)	706	44.0%
11	18	(^^)/	468	29.2%	35	6	(^-^-)/	712	44.4%
13	16	(^O^)	484	30.2%	35	6	(T^T)	718	44.8%
14	15	(^^	499	31.1%	35	6	(-_-)	724	45.2%
14	15	(T_T)	514	32.1%	35	6	(*^ω`*)	730	45.5%
16	14	^^^	528	32.9%	35	6	(*^-^*)	736	45.9%
17	12	(' ω ')	540	33.7%	44	5	m(_ _)m	741	46.2%
17	12	(*^_~*)	552	34.4%	44	5	(o^v^o)	746	46.5%
19	11	(' ω `)	563	35.1%	44	5	(^ v ^)	751	46.8%
20	10	(^^♪)	573	35.7%	44	5	(^v^)	756	47.2%
20	10	(*^?^*)	583	36.4%	44	5	(; ω ;)	761	47.5%
20	10	(*_*^)	593	37.0%	44	5	(; Π)	766	47.8%
23	8	(^▽^)	601	37.5%	44	5	()	771	48.1%
23	8	(` ω `)	609	38.0%	44	5	(^^)	776	48.4%
23	8	(^v^)	617	38.5%	44	5	(^^ゞ	781	48.7%
23	8	(^ω^)	625	39.0%	44	5	(*^艸^)	786	49.0%
23	8	(^.^)	633	39.5%	44	5	(#^.#)	791	49.3%

関がないことを示しているが、印象と感情の依存関係を定式化しようというものではない。また、文献 [19] では、メールの顔文字の部分に半角のゴシック体を用いた場合と全角の明朝体を用いた場合、あるいは顔文字の顔に当たる部分に改行を挿入した場合とで、メール内容に対する読み手の印象と読み手の感情がどのように影響されるかを調べている。このとき、印象に関しては、「ていねいな-らんぼうな、明るい-暗い、親しみのある-よそよそしい、あたたかい-つめたい、楽しい-つまらない、わかりやすい-わかりにくい、良い-悪い」の7項目を、感情に関しては、「悲しみ、不安、怒り」の3項目を採用しているが、印象と感情のそれぞれに関して調べただけであり、両者の関係については触れていない。

以上で述べたように、顔文字の印象や感情に関して様々な角度から研究が行われているが、そのほとんどの研究は印象か感情のどちらかを対象としており、この2種類の感性を同時に扱うような研究は少ない。特に、本研究のように、8種類の印象項目と10種類の感情項目の間の依存関係を定式化しようという研究は見当たらない。

3. 顔文字データの収集

本章では、まず、Twitterを一定以上の頻度で利用しているヘビー投稿ユーザとヘビー閲覧ユーザを抽出し、ヘビー投稿ユーザが普段使用している顔文字を収集する。次に、収集した顔文字の中から出現頻度の高い顔文字を抽出し、ヘビー閲覧ユーザに印象評価もしくは感情評価を行ってもらうことで、顔文字の印象に関するデータと読み手が推測した書き手の感情に関するデータを取得する。

3.1 Twitterヘビーユーザの抽出

本節では、アンケート調査(予備調査A)を行い、一定以上の頻度でつぶやいているヘビー投稿ユーザと一定以上の頻度で閲覧しているヘビー閲覧ユーザを抽出した。

まず、20代~50代のインターネットユーザを対象とする予備調査Aを行い、男性10,808人、女性11,002人からなる21,810人分の回答データを収集した。具体的には、設問Q1でTwitterアカウントの有無を尋ね、「はい」、「いいえ」、「ツイッターが何かわからない」の3択に対し「はい」と回答した人には、設問Q2とQ3においてTwitterの閲覧頻度と投稿頻度を尋ねた。閲覧

頻度に関しては「ほぼ毎日 (1日に4回以上)」、「ほぼ毎日 (1日に2回~3回)」、「ほぼ毎日 (1日に1回)」、「週4~5日」、「週2~3日」、「週1日」、「たまに閲覧している」、「ほとんど閲覧していない」の8つの中から、投稿頻度に関しては「ほぼ毎日 (1日に4回以上)」、「ほぼ毎日 (1日に2回~3回)」、「ほぼ毎日 (1日に1回)」、「週4~5日」、「週2~3日」、「週1日」、「たまにつぶやいている」、「ほとんどつぶやいていない」の8つの中からそれぞれ最も近いものを1つ選んでもらった。

最終的には、Twitter アカウントを持っていると答えた 10,672 人のうち、「ほとんど閲覧していない」と回答した 1,750 人を除く 8,922 人 (83.6%) をヘビー閲覧ユーザとして抽出するとともに、「ほとんどつぶやいていない」と回答した 5,062 人を除く 5,610 人 (52.6%) をヘビー投稿ユーザとして抽出した。

3.2 顔文字の収集

本節では、ヘビー投稿ユーザ 891 人が参加するアンケート調査 (予備調査 B) を行い、Twitter 上で用いられている顔文字を収集した。

まず、設問 Q1 で Twitter 利用時における顔文字の使用頻度を尋ね、「よく使う、たまに使う、あまり使わない、ほとんど使わない、全く使わない」のうち「全く使わない」と回答した 110 人を除く 781 人に対し、設問 Q2 で普段使っている顔文字を入力してもらった。このとき、回答欄は 5 つあり、1~5 個の顔文字が入力可能であった。結果、このアンケート調査により、1,635 個の顔文字を収集することができた。ここで、顔文字の使用頻度に関する調査結果を男女別・年齢別でまとめたものを表 1 に示す。なお、表 1 から、「よく使う、たまに使う」と回答した人は 533 人おり、割と使用頻度の高い人が全体の 59.8% を占めていることがわかる。

次に、この 1,635 個の顔文字に対し各顔文字の出現頻度を求めた。結果の一部を表 2 に示す。但し、表 2 には、出現頻度 5 以上の顔文字 (54 個) が順位・出現頻度・累積数・累積割合とともに示されている。

本論文では、この出現頻度上位の顔文字 54 個を次節以降のアンケート調査で提示刺激として用いることにする。

3.3 顔文字印象データの収集

本節では、ヘビー閲覧ユーザ 530 人が参加するアンケート調査 (本調査 1) を行い、顔文字を見た読み手がどのような印象を感じたかに関するデータ (顔文字印象データ) を取得した。

まず、3.2 節で選出された 54 個の顔文字を均等に分け、18 個の顔文字からなるデータセットを 3 つ作成す

表 3: 基本統計量による分析結果

(a) 基本印象					
	μ	σ	最大値	最小値	幅
楽しい/愉快	2.92	0.81	3.99	1.37	2.62
ポジティブ	2.76	0.71	3.80	1.40	2.40
感じの良い	2.82	0.63	3.71	1.49	2.23
ほのぼの	2.85	0.60	3.55	1.40	2.15
ネガティブ	1.99	0.64	3.45	1.38	2.07
鬱陶しい	1.88	0.29	2.49	1.55	0.94
攻撃的/不愉快	1.70	0.23	2.21	1.40	0.81
怖い	1.57	0.20	2.08	1.33	0.75

(b) 基本感情					
	μ	σ	最大値	最小値	幅
喜び	2.87	1.02	4.34	1.21	3.14
悲しい	1.84	0.83	4.31	1.21	3.10
好き	2.69	0.86	3.84	1.23	2.61
高揚	2.39	0.66	3.71	1.29	2.42
安心	2.39	0.64	3.32	1.24	2.08
恥ずかしい	2.08	0.43	3.28	1.44	1.85
驚き	1.87	0.33	3.06	1.36	1.70
怖れ	1.60	0.42	2.69	1.21	1.48
嫌い	1.62	0.43	2.67	1.21	1.46
怒り	1.52	0.31	2.44	1.21	1.23

るとともに、各データセット内の 18 個の顔文字を 176~178 人の回答者に見てもらい、それぞれの顔文字から感じる印象を評価してもらった。このとき、各回答者には、提示された 8 個の印象「攻撃的/不愉快、ネガティブ、感じの良い、楽しい/愉快、ポジティブ、ほのぼの、鬱陶しい、怖い」[15] のそれぞれに対し、「印象を強く感じる (5 点)、印象を割と感じる (4 点)、印象をやや感じる (3 点)、印象をあまり感じない (2 点)、印象を全然感じない (1 点)」の中から最も近いものを 1 つ選んでもらった。なお、予備調査 B に参加したユーザはあらかじめ回答者から除外され、またデータセット間で回答者が重複しないよう調整が行われた。

次に、以上の結果得られた「54 個の顔文字 × 8 種類の印象 × 5 段階評価値」からなるデータに対し、それぞれの顔文字の各印象において 5 段階評価値の平均を求め、顔文字印象データとした。このとき、評価結果に偏りがある回答者のデータは削除した。

3.4 顔文字感情データの収集

本節では、ヘビー閲覧ユーザ 528 人が参加するアンケート調査 (本調査 2) を行い、顔文字を見た読み手が書き手の感情をどのように推測したかに関するデータ (顔文字感情データ) を取得した。

まず、前節で作成された各データセット内の 18 個の顔文字を 174~177 人の回答者に見てもらい、それぞれの顔文字に書き手のどのような感情が込められている

表 4: 因子分析の結果 (因子数 5, プロマックス回転)

	因子				
	1	2	3	4	5
好き	1.01	0.00	0.00	-0.01	0.10
高揚	1.00	-0.07	0.06	0.31	-0.03
喜び	0.98	-0.05	0.04	0.07	-0.04
安心	0.95	0.12	-0.08	-0.21	0.06
楽しい/愉快	0.92	-0.06	-0.03	0.04	-0.04
感じの良い	0.84	-0.18	0.04	-0.07	-0.03
ポジティブ	0.83	-0.11	-0.06	0.05	-0.10
ほのぼの	0.81	-0.02	-0.13	-0.15	0.14
攻撃的/不愉快	-0.11	1.03	-0.17	0.03	-0.02
怒り	-0.06	0.84	0.15	-0.07	-0.13
怖い	-0.04	0.81	0.04	0.05	0.05
鬱陶しい	-0.34	0.56	0.01	0.11	0.12
嫌い	-0.08	0.52	0.49	-0.03	-0.05
悲しい	-0.15	0.02	0.87	-0.10	0.02
ネガティブ	-0.34	0.10	0.61	-0.02	0.02
怖れ	-0.35	0.09	0.53	0.08	0.08
驚き	0.06	0.05	-0.08	0.78	0.03
恥ずかしい	0.03	-0.06	0.04	0.04	0.75

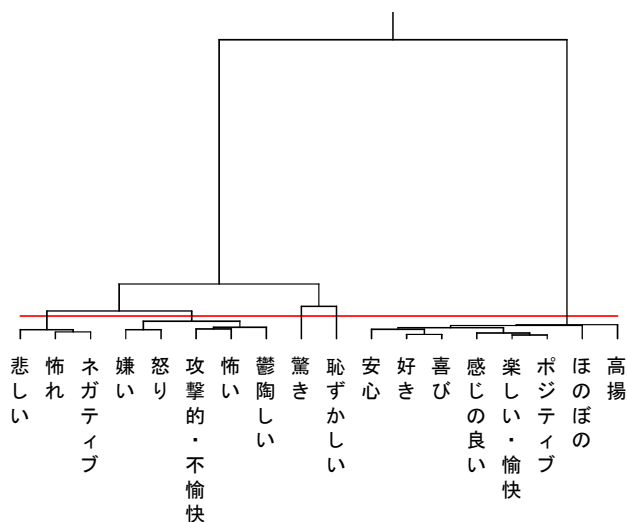


図 1: クラスタ分析の結果 (クラスタ数 5)

かを評価してもらった。このとき、各回答者には、提示された 10 個の感情「喜び、好き、安心、悲しい、嫌い、怖れ、怒り、恥ずかしい、高揚、驚き」のそれぞれに対し、「感情が強く込められている (5 点)、感情が割と込められている (4 点)、感情がやや込められている (3 点)、感情があまり込められていない (2 点)、感情が全然込められていない (1 点)」の中から最も近いものを 1 つ選んでもらった。なお、予備調査 B と本調査 1 に参加したユーザはあらかじめ回答者から除外され、またデータセット間で回答者が重複しないよう調整が行われた。

次に、以上の結果得られた「54 個の顔文字 × 10 種類

表 5: 重回帰分析の結果 (自由度修正済み決定係数)

印象 (目的変数)	R^2	感情 (目的変数)	R^2
攻撃的/不愉快	0.86	喜び	0.97
ネガティブ	0.98	好き	0.95
感じの良い	0.94	安心	0.93
楽しい/愉快	0.97	悲しい	0.95
ポジティブ	0.96	嫌い	0.92
ほのぼの	0.88	怖れ	0.93
鬱陶しい	0.84	怒り	0.88
怖い	0.70	恥ずかしい	0.18
		高揚	0.91
		驚き	0.22

の感情 × 5 段階評価値」からなるデータに対し、それぞれの顔文字の各感情において 5 段階評価値の平均を求め、顔文字感情データとした。このとき、評価結果に偏りがある回答者のデータは削除した。

4. 顔文字データの分析

本章では、前章で得た顔文字印象データと顔文字感情データに対し、基本統計量による分析や因子分析、クラスタ分析を行うことで、印象どうし、感情どうし、印象と感情の間での関係を明らかにするとともに、重回帰分析を行うことで、印象と感情の依存関係を相互に定式化する。

4.1 基本統計量による分析

本節では、顔文字印象データと顔文字感情データの概要を知るために、それぞれの基本統計量 (平均 μ , 標準偏差 σ , 最大値, 最小値, 幅) を求めた。結果を表 3 に示す。表 3 において、幅は最大値と最小値の差であり、幅の降順で各印象、各感情の基本統計量を示している。

表 3 によれば、ポジティブな印象や感情の幅は大きく、ネガティブな印象や感情の幅は小さくなっており、約 3 倍の差が生じていることがわかる。3.2 節で顔文字を収集する際に、普段使っている顔文字を入力してもらった結果、ネガティブなものがあまり多く現れなかったためと考えられる。この点については今後の課題としたい。

4.2 因子分析

本節では、顔文字印象データと顔文字感情データを統合したデータに対し、探索的に因子分析 (主因子法、プロマックス回転) を行った。その結果、因子数 5 のとき、最も良い結果が得られた。このときの各印象・各感情の因子負荷量を表 4 に示す。

表 4 によれば、第 1 因子にはポジティブな印象と感情が属しており、第 2 因子と第 3 因子にはネガティブ

表 6: 偏回帰係数 (目的変数: 印象, 説明変数: 感情)

	攻撃的/不愉快	ネガティブ	感じの良い	楽しい/愉快	ポジティブ	ほのぼの	鬱陶しい	怖い
喜び				0.668	0.208			
好き		-0.126	0.572					
安心					0.183	0.675	-0.157	
悲しい		0.348						
嫌い		0.140	-0.260				0.370	0.397
怖れ		0.477		-0.317	-0.411	-0.378		
怒り	0.575							
恥ずかしい			-0.117		-0.091	0.137	0.073	
高揚	-0.060				0.332			
驚き	0.111		-0.108				0.078	
定数項	0.760	0.694	2.152	1.511	1.785	1.556	1.360	0.928

表 7: 偏回帰係数 (目的変数: 感情, 説明変数: 印象)

	喜び	好き	安心	悲しい	嫌い	怖れ	怒り	恥ずかしい	高揚	驚き
攻撃的/不愉快				-0.580	0.716		1.703	-0.964		
ネガティブ				1.719	0.538	0.633	0.198			
感じの良い			0.106							-0.691
楽しい/愉快	1.518	1.027	0.212						0.549	
ポジティブ								-0.364	0.806	0.804
ほのぼの	-0.415		0.555	0.356	0.131			0.494	-0.395	
鬱陶しい			-0.966				-0.412	1.255	0.728	0.749
怖い			1.180				-0.532			
定数項	-0.380	-0.308	-0.144	-1.607	-1.036	0.346	-0.153	0.953	-1.681	0.197

な印象と感情が属している。第2因子と第3因子を比べてみると、第2因子は外向的であり、第3因子は内向的であると言える。第4因子には「驚き」が属しており、第5因子には「恥ずかしい」が属している。ポジティブかネガティブかという単なる二値分類ではなく、ネガティブが2つのグループに分かれている点や「驚き」と「恥ずかしい」がそれぞれ別のグループを形成している点が興味深い。この点については今後もさらに追及していきたい。

4.3 クラスタ分析

本節では、顔文字印象データと顔文字感情データを統合したデータに対し、クラスタ分析（階層的クラスタリング、ワード法）を行った。結果を図1に示す。

図1を見てみると、クラスタ数が5のとき、表4に示した因子構造と同じ内容になっているのがわかる。因子分析は対象となったデータの背後にある共通因子を明らかにするものであるのに対し、クラスタ分析は相関の強いものどうしをグループ化するものである。結果、同じグループが形成されたことから、グループ構成の妥当性が伺える。

4.4 重回帰分析

本節では、それぞれの顔文字印象データを目的変数、すべての顔文字感情データを説明変数とする重回帰分

析とそれぞれの顔文字感情データを目的変数、すべての顔文字印象データを説明変数とする重回帰分析を行った。それぞれの重回帰分析において算出された自由度修正済み決定係数を表5に示す。

表5によれば、自由度修正済み決定係数は、「恥ずかしい」と「驚き」を除けば、いずれも0.5以上となっており、得られた重回帰式の当てはまり具合が良いことを示している。ここで、それぞれの重回帰分析において得られた各印象・各感情に対する偏回帰係数をそれぞれ表6と表7にまとめる。

5. 評価実験

本章では、重回帰式による感情値から印象値の推測ならびに印象値から感情値の推測がどの程度うまくいっているのかを調べるために、まず学習データに対する精度を評価し、次に学習データを用いて5分割交差検定を行うことにより未知データに対する精度を評価する。

5.1 学習データに対する精度評価

本節では、前章で求めた偏回帰係数（表6、表7参照）を用いてそれぞれの印象値と感情値を算出し、学習データに対する精度を評価した。このとき、評価指標としては平均二乗誤差（Root-Mean-Square-Error: RMSE）を採用した。結果を表8に示す。

表 8: 学習データに対する平均二乗誤差 (RMSE)

印象	RMSE	感情	RMSE
攻撃的/不愉快	0.08	喜び	0.17
ネガティブ	0.10	好き	0.19
感じの良い	0.15	安心	0.16
楽しい/愉快	0.14	悲しい	0.17
ポジティブ	0.13	嫌い	0.12
ほのぼの	0.20	怖れ	0.11
鬱陶しい	0.11	怒り	0.10
怖い	0.11	恥ずかしい	0.37
		高揚	0.19
		驚き	0.28
平均	0.13	平均	0.19

表 8 によれば, 各印象における RMSE は 0.08~0.20 の範囲であり, その平均は 0.13 であった. 一方, 各感情における RMSE は 0.10~0.37 の範囲であり, その平均は 0.19 であった. 印象値や感情値が取りうる値が 1.0~5.0 ということを考え, この幅 (= 5.0 - 1.0 = 4.0) で正規化してみると, 印象に対する平均誤差 (=RMSE の平均/4.0) は 3.2%となり, 感情に対する平均誤差は 4.7%となることからわかる. いずれも 10%未満であり, 十分高い精度と言える. ここで, RMSE が大きかった「恥ずかしい」と「驚き」を除外して感情の RMSE に対する平均と平均誤差を求めてみると, それぞれ 0.15 と 3.8%になり, 印象との差はだいぶ縮まる. 「恥ずかしい」も「驚き」も, その自由度修正済み決定係数は 0.5 よりかなり小さく, 重回帰式の当てはまり具合が良くないことを示していたが, その原因の一つとして, 「恥ずかしい」や「驚き」と相関のある印象が基本印象の中になかったという点が挙げられる. 「恥ずかしい, 驚き」と 8 個の印象との間で相関係数の絶対値が最も大きかったのは「鬱陶しい」であり, いずれも 0.33 であった.

5.2 未知データに対する精度評価

本節では, 学習データに対し 5 分割交差検定を行うことで, 未知データに対する精度を評価した. 結果を表 9 にまとめる. なお, 表 9 には各未知データセットに対する RMSE とその平均 (平均 RMSE) を示す.

表 9 によれば, 平均 RMSE が最も大きかったのは, 印象に対しては「ほのぼの」の 0.24 であり, 感情に対しては「恥ずかしい」の 0.48 であった. 前節で行ったように, 幅 (= 4.0) で正規化してみると, 「ほのぼの」の平均誤差 (=RMSE の平均/4.0) は 6.1%であり, 「恥ずかしい」の平均誤差は 12.0%であった. 平均誤差が 10.0%を超えたのは「恥ずかしい」のみであり, 他の印象や感情は 2.7% (「ネガティブ, 怖れ」) ~8.8% (「驚き」) の範囲であった. したがって, 「恥ずかしい」以外

表 9: 5 分割交差検定の結果 (RMSE)

(a) 目的変数: 印象, 説明変数: 感情						
印象	未知データセット					平均 RMSE
	1	2	3	4	5	
攻撃的/不愉快	0.12	0.11	0.14	0.10	0.13	0.12
ネガティブ	0.08	0.10	0.14	0.09	0.13	0.11
感じの良い	0.24	0.23	0.26	0.16	0.17	0.21
楽しい/愉快	0.19	0.26	0.14	0.13	0.12	0.17
ポジティブ	0.19	0.26	0.18	0.16	0.14	0.19
ほのぼの	0.26	0.26	0.30	0.21	0.18	0.24
鬱陶しい	0.18	0.10	0.23	0.14	0.16	0.16
怖い	0.18	0.07	0.20	0.12	0.10	0.13

(b) 目的変数: 感情, 説明変数: 印象						
感情	未知データセット					平均 RMSE
	1	2	3	4	5	
喜び	0.19	0.24	0.14	0.15	0.19	0.18
好き	0.23	0.25	0.22	0.15	0.20	0.21
安心	0.19	0.22	0.19	0.12	0.12	0.17
悲しい	0.13	0.22	0.21	0.14	0.24	0.19
嫌い	0.13	0.08	0.22	0.11	0.18	0.14
怖れ	0.08	0.13	0.16	0.06	0.12	0.11
怒り	0.06	0.10	0.24	0.12	0.15	0.13
恥ずかしい	0.44	0.62	0.37	0.64	0.33	0.48
高揚	0.19	0.34	0.40	0.21	0.20	0.27
驚き	0.40	0.34	0.36	0.41	0.25	0.35

に対しては, 十分高い精度が得られていると言える.

6. まとめ

本論文では, 読み手 (顔文字を受け取った人) が感じた顔文字の印象と読み手が顔文字を見て推測した書き手 (顔文字を送った人) の感情を定量化し, 印象どうし, 感情どうし, 印象と感情の間にどのような関係があるかを多変量解析手法を用いて分析するとともに, 両者の依存関係を相互に定式化した. 具体的には, 一定以上の頻度で投稿している Twitter ユーザが普段用いている顔文字を収集し, 出現頻度上位の顔文字を抽出するとともに, 各顔文字を見てどのような印象を感じるかを問うアンケート調査と各顔文字に書き手のどのような感情が込められていると思うかを問うアンケート調査を行うことで, 顔文字印象データと顔文字感情データを取得した. これらのデータに対し, 因子分析やクラスタ分析を行うことで, 印象や感情を 5 つのグループに分けられることを示した. すなわち, 第 1 グループには「好き, 高揚, 喜び」のようなポジティブな印象や感情が分類され, 第 2 グループと第 3 グループにはネガティブな印象や感情が分類された. 但し, 第 2 グループには「攻撃的/不愉快, 怒り」のような外向的な印象や感情が分類されたのに対し, 第 3 グループには「悲しい, ネガティブ」のような内向的な印象や感

情が分類された。一方、第4グループには「驚き」が分類され、第5グループには「恥ずかしい」が分類された。以上のことから、印象や感情を分類する際に、ポジティブかネガティブかという単なる二値分類では不十分であることが示唆された。また、それぞれの顔文字印象データを目的変数、すべての顔文字感情データを説明変数とする重回帰分析とそれぞれの顔文字感情データを目的変数、すべての顔文字印象データを説明変数とする重回帰分析を行うことで、印象と感情の相互の依存関係を重回帰式という形で定式化した。

提案した手続きにより作成される重回帰式を用いて、未知データに対する精度評価(5分割交差検定)を行った結果、「恥ずかしい」以外の平均RMSEは0.11~0.35であり、十分高い精度で推測できることが示された。一方、「恥ずかしい」の平均RMSEは0.48であり、精度の改善が必要とされる。今後の課題としたい。また、印象と感情の依存関係を相互に定式化することで、(1)顔文字印象データがあれば顔文字感情データを、顔文字感情データがあれば顔文字印象データを高精度に自動生成できるようになった。特に、顔文字感情データもしくはそれに類するデータは、顔文字推薦に関する研究や顔文字辞書構築に関する研究において作成されることが多く、本研究での再利用は十分可能と考えられる。今後の課題としたい。(2)顔文字推薦や顔文字辞書構築に際して、書き手の感情だけでなく、読み手がどのような印象を受けるかや読み手が書き手の感情をどう推測するかを考慮できるようになった。厳密には、書き手の感情と読み手が推測する書き手の感情の関係を明らかにし、定式化する必要があるが、これについては現在研究を進めているところである。また、今回得られた重回帰式の構成を精査することで、重回帰式が表す印象と感情の依存関係の妥当性についてより深く考察する必要がある。今後の課題としたい。

謝辞

本研究の一部は、JSPS 科研費 17K00430 ならびに福田将治奨学寄付金による研究助成の成果であり、ここに記して感謝の意を表す。

参考文献

- [1] 奥村紀之, 尾崎知伸, 特集「顔文字の科学—Web上の非言語表現・行動に関する新研究分野の誕生—」にあたって, 人工知能学会誌, Vol. 32, No. 3, pp. 331-332, 2017.
- [2] プタシンスキ ミハウ, 奥村紀之, ジェブカ ラファウ, 顔文字の現象および研究の概観, 人工知能学会誌, Vol. 32, No. 3, pp. 333-341, 2017.
- [3] 奥村紀之, プタシンスキ ミハウ, ジェブカ ラファウ, 工学的・心理学的側面から見た顔文字の分析に関する研究動向, 人工知能学会誌, Vol. 32, No. 3, pp. 342-349, 2017.
- [4] ト部有記, プタシンスキ ミハウ, ジェブカ ラファウ, 顔文字推薦技術の研究動向, 人工知能学会誌, Vol. 32, No. 3, pp. 356-363, 2017.
- [5] Nobuo Suzuki, and Kazuhiko Tsuda, Express Emoticons Choice Method for Smooth Communication of e-Business, KES2006, Part II, LNAI 4252, pp. 296-302, 2006.
- [6] 江村優花, 関洋平, マイクロブログにおける感情・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, 自然言語処理, Vol. 19, No. 5, pp. 401-418, 2012.
- [7] ト部有記, ジェブカ ラファウ, 荒木健治, 顔文字の表す感情を用いた顔文字推薦システムの構築, 言語処理学会第19回年次大会発表論文集, pp. 648-651, 2013.
- [8] 吉田綾奈, 邱起仁, 樋山淳雄, 顔文字推薦のための感情を付与した顔文字データベースの構築, 情処研報(エンタテインメントコンピューティング), Vol. 2014-EC-31, No. 35, pp. 1-6, 2014.
- [9] Michal Ptaszynski, Jacek Maciejewski, Pawel Dybala, Rafal Rzepka, and Kenji Araki, CAO: A Fully Automatic Emoticon Analysis System Based on Theory of Kinesics, IEEE Transactions on Affective Computing, Vol. 1, No. 1, pp. 46-59, 2010.
- [10] 廣瀬信之, 牛島悠介, 森周司, 携帯電話メールによる感情の伝達に顔文字と絵文字が及ぼす影響, 感情心理学研究, Vol. 22, No. 1, pp. 20-27, 2014.
- [11] 荒川歩, 鈴木直人, 謝罪文に付与された顔文字が受け手の感情に与える効果, 対人社会心理学研究, Vol. 4, pp. 128-133, 2004.
- [12] 竹原卓真, 栗林克匡, 様々なエモティコンを付加した電子メールが受信者の印象形成に及ぼす効果, 日本感性工学会研究論文集, Vol. 6, No. 4, pp. 83-90, 2006.
- [13] 田口雅徳, 顔文字の付与および文頭・文末の小文字化がメール文の印象に与える影響, 情報学研究(獨協大学情報学研究所), Vol. 3, pp. 105-111, 2014.
- [14] 中村明, 感情表現辞典, 東京堂出版, 東京, 1993.
- [15] 熊本忠彦, Twitterの利用状況とツイートの印象に関する分析, 情処研報(自然言語処理), Vol. 2016-NL-227, No. 16, pp. 1-13, 2016.
- [16] Robert Plutchik, A General Psychoevolutionary Theory of Emotion, In *Emotion: Theory, Research, and Experience*, Vol. 1, No. 3, pp. 3-33, 1980.
- [17] 金官圭, CMC (Computer mediated communication)における印象形成に関する探索的検討, 社会心理学研究, Vol. 14, pp. 123-132, 1999.
- [18] 荒川歩, 竹原卓真, 鈴木直人, 顔文字付きメールが受信者の感情緩和に及ぼす影響, 感情心理学研究, Vol. 13, No. 1, pp. 22-29, 2006.
- [19] 荒川歩, 河野直子, 顔文字の表示形態および中途での改行がメールの印象評定および受信者の感情に与える影響, 感情心理学研究, Vol. 15, No. 2, pp. 107-114, 2008.