

# 顔文字の表す感情を用いた顔文字推薦システムの構築

ト部有記, ジェプカ ラファウ, 荒木健治

北海道大学大学院情報科学研究科

{y\_urabe, kabura, araki}@media.eng.hokudai.ac.jp

## 1. はじめに

近年, 電子メール, Twitter や Facebook 等の SNS (ソーシャルネットワークサービス) において文字ベースのコミュニケーションでの顔文字の使用が一般化している. 今日存在する顔文字辞書に含まれる顔文字の数は, 現在も増加の一途をたどっている. 例えば, iPhone の顔文字辞書アプリケーション「girl's 顔文字辞典 5000」[1]では5,000個, 「顔文字辞書 The Facemark Dictionary」[2]では58,130個の顔文字が登録されている. このため, これらの中からユーザが文章で伝えたい感情が含まれた顔文字を選択するのは困難である.

本稿では, ユーザが入力した文章に適切な顔文字の選択支援を目的とし, 顔文字に含まれる感情を被験者から採ったアンケート調査によって数値化し, その数値を用いてユーザの入力文から推定された感情に一致する感情を持つ顔文字を推薦するシステムを提案する.

## 2. 関連研究

顔文字を用いた研究に関しては, 川上[3]が, 顔文字が表す感情と強調に関する顔文字データベースの作成を行った. 彼の調査は, 1つの顔文字に複数の感情を表す場合があることが示した. Ptaszynski ら[4]が構築した顔文字解析システム CAO は顔文字の抽出と感情解析を行い 90%以上の精度をあげている. Ptaszynski らが作成した顔文字データベースは, オンライン上の顔文字辞書のラベルを感情表現解析システム ML-Ask[6]を用いて分析し, 各ラベルに含まれている顔文字を10個の感情「喜・哀・怒・怖・安・恥・驚・好・昂・厭」に分類することにより構築されている. 江村ら[5]は, 顔文字の使い方を研究し, 顔文字は感情だけでなくコミュニケーションや動作も表すことを明らかにして顔文字推薦システムの構築を行った. 江村らが作成した顔文字データベースは, twitter のつぶやきからコーパスを作成し, アンケート調査によって各顔文字がどの感情に当てはまるかを分類することにより構築されている.

本研究は, 川上の手法を参考にして顔文字がどの感情に当てはまるかではなく, 顔文字が各感情をどの程度表現しているのかをアンケート調査により数値化し, 顔文字と得られた数値を用いて顔文字データベースを作成する.

## 3. 顔文字データベースの作成

本手法では, 大きく分けて感情推定と顔文字推薦の2種類の処理から成り立っている. 感情推定には Ptaszynski らが構築した感情表現解析システム ML-Ask[6]を用いた. ML-Ask は, ML-Ask で解析する10個の感情と本研究で作成する顔文字データベースで用いる感情を合致させるために用いた.

顔文字データベースは iPhone に用いられているモバイルオペレーションシステム iOS 5.0 に登録されている顔文字 140 個中 59 個を用いた. 59 個の顔文字の選択は, “( )” の外に記号がないものという基準を用いて行った. 例えば, “(\*^\_^\*)” は採択したが, “(・ω・)／” は除外した. “( )” の中には, 基本的に左から「目」, 「口」, 「目」と, 顔の表情を構成している. 本稿では, 顔文字データベースの作成を行う最初の段階として, “( )” の中身に使用される記号の違いだけに注目し, 被験者がそれぞれの顔文字に対して感じる感情を調査した. 選択した 59 個の顔文字をアンケート調査で被験者にそれぞれ 10 個の感情「喜・哀・怒・怖・安・恥・驚・好・昂・厭」(感情表現辞典[7]より抜粋)のどれに合致するかの評価をしていただいた.

### 3. 1. アンケート調査

アンケート調査は 2012 年 7 月 1 日から 7 月 30 日に大学及び大学院に所属する学生 60 名(男性 22 名, 女性 38 名)に対して実施した. 60 名の被験者のうち, 33 名が文系, 27 名が理系の学生であった.

59 個の顔文字がそれぞれどの程度の感情を表しているか 5 件法での評価を求めた. 5 件法とは, 対象物を評価するために, 評価するための複数の短文を用いて回答するという評価尺度法で用いられる方法の一つである[3].

顔文字の評価に用いられた感情は, 感情表現辞典に定義されている「喜・哀・怒・怖・安・恥・驚・好・昂・厭」の 10 個である. また, 同時に被験者がどのような顔文字を使いたいと思うのかについての調査を行った. 『これから 59 個の顔文字に対してそれぞれの感情を 5 段階 (1. 全く当てはまらない, 2. 当てはまらない, 3. どちらでもない, 4. 当てはまる, 5. とても当てはまる) で評価してください. また, 最後に「この顔文字を使ってみたいか。」という質問に対して「はい」又は「いいね」のどちらかに丸を付け

てください。』という教示と評価例（図1）を示し、各感情を説明する単語を6つずつ例示した（表1）。また、例示する単語を6つずつとしたのは、感情表現辞書より各感情を広く表現する単語を選択する際に最低6つの単語が必要であると考えたためである。

この際、59 個の顔文字はランダムな順序で並べた。

(ë\_ë)

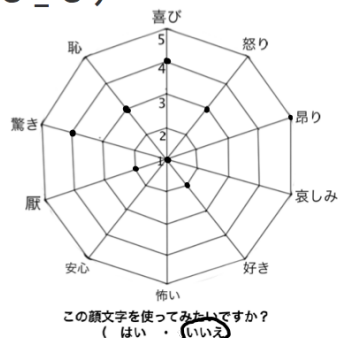


図1：顔文字の評価例

表1：各感情を表す単語例

各感情について表す単語を例示したので参考にしてください。

喜：嬉しい, 満悦, 楽しい, 面白い, 快樂, 幸福感  
怒：立腹, 激怒, 憤り, 敵意, むかつく, 叱責  
昂：焦る, そわそわ, 自棄, 感激, 感服, 昂揚  
哀：悲哀, 愁い, 嘆く, 悲涙, 不憫, 侘しい  
好：友情, 愛する, 恋しい, 惚れる, 敬愛, 懐かしむ  
怖：不気味, 恐怖, 怯える, 臆病, 不安, 心細い  
安：安堵, 和やか, 平静, のんびり, 安楽, ほっと  
厭：悔しい, 不快, 疎む, 憎悪, 気怠い, 鬱  
驚：吃驚, 動揺, 慌てる, 呆然, 意外, 思いもよらず  
恥：照れる, はにかむ, 恥辱, 冷や汗, 肩身が狭い, 紅潮

### 3. 2. アンケート結果と考察

59 個の顔文字に対する、各感情評定値及び感情の相関係数、顔文字の相関係数を算出した。各感情評定値は顔文字推薦の際、ユーザの入力文を ML-Ask で感情推定した結果より顔文字データベースの顔文字の推薦順序を並び替える際に用いる。評定値の一部を表2に示す。感情評定値は、数値が5に近い程顔文字が当該の感情に当てはまり、1に近い程、当該の感情に当てはまらない。例えば、表2の“(^O^)”は「喜」という数値が4.63と一番高く、「厭」という数値が1.38と一番低くなっている。従って、この顔文字は「喜」という感情に近く、「厭」という感情から遠いことが明らかとなった。

感情の相関係数において、互いに0.80以上の高い正の相関を示した感情は、喜と好、喜と安、好と安、哀と怖、怖と厭である。互いに-0.80以下の高い負の相

関を示した感情は、厭と喜、厭と安、厭と好、怖と安、怖と好、怖と喜である。驚、恥、昂、怒からは正負共々高い相関はなかった。

表2：喜の得点を降順に並べた顔文字の感情評定値

顔文字	喜び	怒り	昂り	哀しみ	好き	怖い	安心	厭	驚き	恥
(^^O^^)	4.85	1.22	1.95	1.25	4.24	1.18	3.55	1.22	1.75	1.80
(*^^*)	4.75	1.42	1.90	1.43	3.98	1.30	3.57	1.30	1.48	1.98
(^^O^^)	4.70	1.22	1.65	1.22	3.93	1.20	3.48	1.24	1.47	1.85
(^▽^)	4.65	1.37	1.73	1.28	3.64	1.32	3.75	1.27	1.32	1.33
(^O^)	4.63	1.47	1.83	1.48	3.61	1.33	3.42	1.38	1.75	1.42
(#^#)	4.55	1.28	2.02	1.27	3.93	1.25	3.20	1.23	1.53	2.32
(^O^)	4.55	1.30	1.95	1.33	3.68	1.23	3.40	1.28	1.88	1.48
(^▽^)	4.53	1.28	1.63	1.25	3.68	1.23	3.55	1.28	1.48	1.43
(O^^O)	4.52	1.23	1.85	1.22	4.22	1.20	3.73	1.25	1.32	2.02
(^^)	4.48	1.32	1.78	1.28	3.68	1.28	3.50	1.35	1.42	1.37
(*^*)	4.47	1.40	1.63	1.37	4.12	1.25	3.58	1.33	1.40	1.98
(^O^)	4.43	1.20	1.60	1.25	3.59	1.25	3.73	1.28	1.62	1.40
(=^X^=)	4.40	1.22	1.73	1.22	3.69	1.22	3.05	1.23	1.42	1.67
(^◇^)	4.33	1.28	1.73	1.23	3.29	1.20	3.47	1.27	1.53	1.38

顔文字の相関係数の一部を表3に示す。顔文字の相関係数の結果から、顔文字同士に高い正の相関がある場合、括弧内の記号が一部同じであることが明らかとなった。例えば、表3の“(^O^)”と“(^▽^)”の違いは口の部分が、前者が“O”，後者が“▽”であり、相関係数が0.996と高い数値が確認された。従って、括弧内にほとんど同じ記号を用いた顔文字は人が顔文字より受け取る感情はほぼ同じであることが得られた数値より確認された。逆に、高い負の相関がある場合、括弧内の記号が全く異なることが明らかとなった。表3の顔文字の相関係数の結果より、“(^O^)”と“(一一;)”は、“( )”内の記号が全く異なり、これらの顔文字の相関係数は-0.856と高い負の相関を示した。従って、括弧内の記号が全く異なる場合は、人が顔文字より受け取る感情が異なることが確認された。

表3：顔文字の相関係数（一部）

	(^O^)
(^O^)	1.000000
(^▽^)	0.995785
(^O^)	0.995760
(^.)	0.995750
(~::~)	-0.855559
(一一;)	-0.842154

アンケート結果を全体的に見ると、ある感情に対して括弧内に使用される記号に共通性があることが考えられる。例えば、表2の結果より、「喜」に対して多く用いられる記号は、目の部分は「^, ^, ^」, 口の部分は「O, O, ▽」であることが確認された。

## 4. 顔文字推薦システムの構築

3章で作成した顔文字データベースを用いた顔文字推薦システムの全体の流れについて説明する。

- (1) ユーザが文章を入力。
- (2) 入力文章の感情を ML-Ask を用いて推定。
- (3) 感情推定結果を用いて顔文字データベースから顔文字を 5 件ずつ取り出して推薦候補として表示。推薦候補を 5 件としたのは、5 件が最も選択しやすい数であると考え、今回の実験では仮に 5 件としたが、今後最適な値を実験により明らかにする必要がある。
- (4) ユーザが顔文字を選択。表示された 5 件の推薦候補の中に適切な顔文字がない場合は「次の顔文字候補」を選択し、次の候補 5 件を表示することを可能とした。
- (5) ユーザの入力文と顔文字が同時に出力。

## 5. 評価実験

### 5. 1. 実験方法

実験は 2012 年 10 月 22 日から 10 月 31 日の間に大学及び大学院に所属する学生 20 名（男性 8 名、女性 12 名）を被験者として実施した。20 名の被験者のうち文系と理系の学生はそれぞれ 10 名ずつであった。

SD(Semantic Differential)法[8]を用いて、従来手法と本研究で実装した顔文字推薦システムを提案手法として双方の比較評価実験を行った。SD 法とは、「遅い-速い」などの対立する形容詞の対を用いて、対象物の感情的なイメージを、5 段階あるいは 7 段階で評価する方法である。従来手法は、iPhone(注1)で用いられている顔文字の順番の並べ方である。iPhone では、顔文字を選択すると次の選択時にその顔文字が登録されている顔文字の中で一番上位となる方法を用いている。システムに入力する文章は、実験にかかる時間を考慮して両手法とも 15 文ずつとした。15 文は、10 個の感情から感情表現辞典より感情語を 5 個ずつ用意し、10 個の感情のうちランダムに 3 個の感情に含まれている合計 15 個の感情語を各文に 1 つずつ使用するように指示をした。感情語の選び方は、感情表現辞典の各感情に登録されている 1 つの感情だけを定義された感情語のうち Google(注2)のキーワード検索でのヒット件数が高い感情語の上位 5 個である。1 つの感情を表す単語だけを用いたのは、ユーザが 1 つの感情を表す文章を入力した時にどの顔文字を選択するのかを確認するためである。

(注1) : iPhone は Apple Inc. の登録商標である。

(注2) : <https://www.google.com/>

表 4 : SD 法で用いられた単語対

SD 法で用いた 23 個の単語対
・面白さ つまらない-面白い 楽しくない-楽しい 不愉快な-愉快的な 親しみにくい-親しみやすい 興味がない-興味がある 感じの悪い-感じの良い 気持ちの悪い-気持ちのよい ありがちな-新鮮な 劣っている-優れている 役に立たない-役に立つ 一般的な-特殊な
・使いやすさ 遅い-速い のろい-すばやい 重い-軽い 選びやすすくない-選びやすい 不便な-便利な 不自由な-自由な 頼りない-頼もしい わがままな-思いやりのある
・簡単さ 複雑な-単純な 難しい-分かりやすい 不親切な-親切的な 不正確な-正確な

SD 法で用いた単語対は 23 個あり、「面白さ」「使いやすさ」「簡単かどうか」という点を考慮し、井上正明らの「日本における SD 法による研究分野とその形容詞対構成の概観」[9]に記載されている単語対を参考に選択した(表 4)。

また、被験者が顔文字を選択する際にかかった時間も計測した。

実験の手順について説明する。

- (1) 参加者が 1 つ目の手法を利用する。

従来手法もしくは提案手法のどちらかを先に使うことで表れる違いを確認するため最初に使用する手法をランダムに指定した。

- (2) 23 個の単語対を用いて 1 つ目の手法を評価する。
- (3) (1) と同様に 2 つ目の手法を利用する。
- (4) (2) と同様に 2 つ目の手法を評価する。

### 5. 2. 実験結果と考察

従来手法と提案手法それぞれの SD 法より得られた数値(表 5)を比較した。提案手法が従来手法よりも評価が特に高かったのは、「選びやすい-選びにくい」が 1.05 点、「正確な-不正確な」が 0.55 点、「特殊な-一般的な」が 0.55 点であった。反対に、従来手法が提案手法よりも評価が特に高かったのは、「単純な-複雑な」が 0.40 点、「親しみやすい-親しみにくい」が 0.30 点であった。

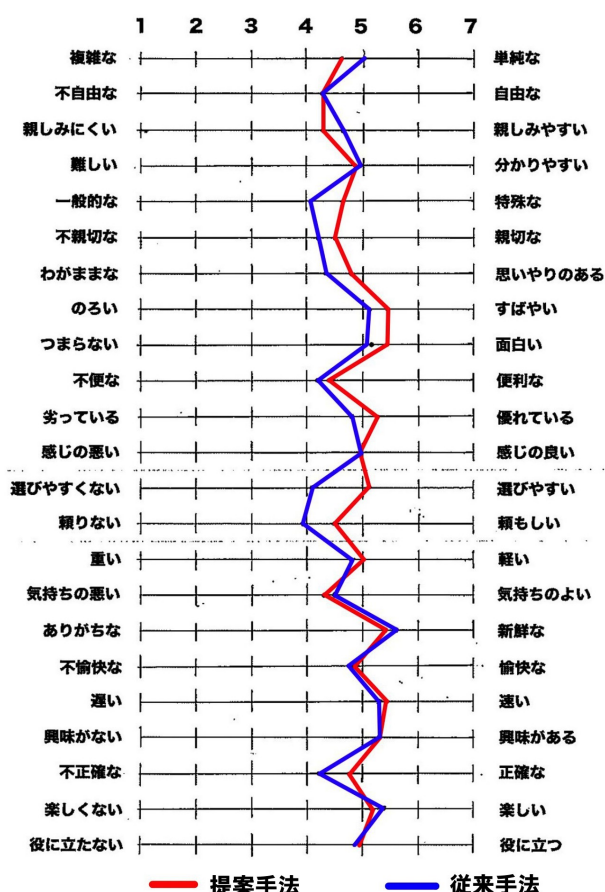
また、被験者が顔文字の選択にかかった時間は、感情によって変化が見られた。「怒り」と「厭」の感情語を含む文章に適切な顔文字を選択する時間は提案手法が従来手法よりもおよそ 3.50 秒速いことが確認された。しかし、「怖い」と「昂り」に対しては提案手法が従来手法よりもおよそ 2.00 秒遅いことが確認された。顔文字データベースに登録されている顔文字において、「怒り」や「厭」という感情に対して高い数値を持つ顔文字は多いが、「怖い」や「昂り」という感情に対して高い数値をもつ顔文字が少ないからであると考えられる。したがって、顔文字データベースに登録している顔文字が増加すれば、それぞれの感情に対して高い数値を持つ顔文字が得られるの



でこの問題は解決可能であると考えられる。また、「怒り」「厭」「怖い」「昂り」以外の残りの6つの感情の感情語を含んだ文章に適切な顔文字を選択した際には、提案システムと従来システムの両方にかかる時間に大きな違いはなかった。

全体の結果から従来システムと提案システムを比較した。全体の結果から従来システムと提案システムの比較を行った。従来システムは、顔文字を一度選択するとその顔文字が顔文字データベースの中で上位に上がる仕組みとなっているので、特定の顔文字だけを使うユーザにとっては従来手法のほうが使い易いとが考えられる。一方、本研究で提案した手法は、相手に感情をより正確に伝えることを重要とするユーザにとって使い易いと考えられる。これは、提案システムが受け手が顔文字に対して感じる感情を数値化することにより作成した顔文字データベースを用いているためである。

表5：SD法に基づいたアンケート調査の結果



## 6. まとめと今後の課題

顔文字それぞれが持つ感情の値をアンケートで調査し、その値を用いた顔文字推薦システムを構築した。本手法は、入力文章より推定された感情に適切な感情を持つ顔文字を推薦することを可能とした。提案手法と従来手法それぞれをSD法で得られた結果から、提案手法は従来手法よりも「選びやすい」選

びにくい」が1.05点、「正確な-不正確な」が0.55点、「特殊な-一般的な」が0.55点高いことが確認された。反対に、従来手法は提案手法よりも「単純な-複雑な」が0.40点、「親しみやすい-親しみにくい」が0.30点高いことが明らかとなった。以上の結果から、提案手法は相手に感情を正確に伝えたいユーザに対して選び易い手法であることが考えられる。

後の課題としては、顔文字データベースのデータを増加することで該当する感情により適切な顔文字を選択する可能性を増加させることが挙げられる。また、従来手法と提案手法を併用することにより、各ユーザに適応した顔文字推薦システムを構築することが可能となると考えられる。今回の実験で用いたインタフェースはMac OS Xに付属しているUnix系端末エミュレータのターミナルであったが、本実験システムをアプリケーション化することによりスマートフォンでの利用を可能とし、ユーザの反応を確認する予定である。

## 参考文献

- [1] girl's 顔文字辞典 5000, PEDIGREE (ペディグリー), <https://play.google.com/store/apps/details?id=kao.moji>.
- [2] 顔文字辞書 The Facemark Dictionary, <http://www.kaomoji.sakura.ne.jp/>.
- [3] 川上正造: 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, 2008-01-31.
- [4] Michal Ptaszynski, Jacek Maciejewski, Pawel Dybala, Rafal Rzepka, Kenji Araki: CA0: A Fully Automatic Emoticon Analysis System Based on Theory of Kinesics, IEEE Transaction on Affective Computing, Vol. 1, No. 1 pp. 46-59, 2010-1.
- [5] 江村優花, 関洋平: テキストに表れる感情, 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. 情報学基礎研究会報告 2012-IFAT-106(1), pp. 1-7, 2012-03-19.
- [6] Michal Ptaszynski, Pawel Dybala, Wenhan Shi, Rafal Rzepka, Kenji Araki: A System for Affect Analysis of Utterances in Japanese Supported with Web Mining, 日本知能情報ファジイ学会誌, Vol. 21, No. 2, pp. 194-213, 2009.
- [7] 中村明: 感情表現辞典, 東京堂出版, 1993-5.
- [8] 大石康彦, 金濱聖子, 比屋根哲, 田口春孝: 森林空間が人に与えるイメージと気分の比較: POMS およびSD法を用いた森林環境評価, 日本林學會誌 85(1), pp. 70-77, 2003-02-16.
- [9] 井上正明, 小林利宣: 日本におけるSD法による研究分野とその形容詞対尺度構成の概観, 教育心理学研究 33(3), pp. 253-260, 1985-09-30.