

令和 2 年 6 月 15 日現在

機関番号：32503

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K00430

研究課題名(和文) 個々の書き手・読み手を意識した統合的な感性情報マイニング技術に関する研究

研究課題名(英文) A Study on Integrated Kansei Information Mining Techniques for Individual Writers and Readers

研究代表者

熊本 忠彦 (Kumamoto, Tadahiko)

千葉工業大学・情報科学部・教授

研究者番号：30358890

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：本研究では、ツイートなどの文章から書き手の感情、読み手が感じた印象、読み手が推測した書き手の感情という3種類の感性情報を抽出することを目的とする。

まず、アンケート調査に基づいてツイートとツイートから抽出される感情や印象との関係を分析し、それらの関係を多変量解析手法や機械学習手法を用いて定式化することで、高精度かつロバストにツイートの感情や印象を抽出できる手法を開発した。

なお、本研究では、「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感情と「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」の8印象を対象とする。

研究成果の学術的意義や社会的意義

文章に内在する感性情報を抽出する研究において、書き手の感情、読み手の印象、読み手が推測する書き手の感情の3種類を同時に抽出し、相補的に利用するという研究は、これまでなかった。それぞれの感性情報を相補的に利用することで、より高精度な感性情報抽出(感性情報マイニング)を実現することが期待される。また、個々の書き手や読み手の感性情報をより高精度に抽出できるようになることで、既存の研究(トレンド分析やニュース記事推薦、リアルタイムイベント抽出、ユーザの印象選好の可視化、SNS等でのフォロワー推薦など)に与える影響も大きい。

研究成果の概要(英文)：The purpose of this study is to extract three kinds of affective information from sentences such as tweets: emotions of the writers who wrote the sentences, impressions which readers felt from reading the sentences, and the writers' emotions which the readers estimated by reading the sentences.

First, we conduct questionnaire surveys, and analyze the relationships between tweets and emotions and impressions extracted from the tweets. And then, we formulate these relationships using multivariate analysis methods and machine learning methods, resulting in high accuracy and robustness. We have developed a method that can extract the emotions and impressions of tweets.

Note that, in this study, ten kinds of emotions of "joy, love, relief, sad, dislike, fear, anger, embarrassment, uplifting, and surprise" and eight kinds of impressions of "aggressive/unpleasant, negative, pleasant, pleasant/happy, positive, heartwarming, annoying, and scary" are used.

研究分野：感性情報処理

キーワード：感情 印象 ツイート

## 様式 C-19、F-19-1、Z-19（共通）

### 1. 研究開始当初の背景

Twitter や Facebook といった様々なタイプの SNS (Social Networking Service) が日常生活に溶け込んでいる。中でも、Twitter は、(i) 一度に入力可能な文字数が 140 字と少なく、気軽に閲覧したり投稿したりすることができる、(ii) スマートフォンなどを用いていつでもどこからでも簡単にアクセスでき、リアルタイムな情報の獲得や発信が可能である、(iii) 友人や知人といった面識のある人だけでなく、芸能人や同好の士といった面識のない人でもフォローでき、様々な情報を得ることができるといった点に特徴があり、日本では年齢や性別によらず幅広い世代に支持されている。このような Twitter の特徴を利用した研究は数多くなされており、トレンド分析やニュース記事推薦、リアルタイムイベント抽出といった応用研究だけでなく、ツイートの印象やツイート投稿者の感情などツイートに内在する感性情報を対象とした基礎研究も行われている。

メールや新聞記事、ツイートなどの文章に内在する感性情報を抽出する研究は、申請者らの研究をはじめ、多数存在しているが、書き手の感情、読み手の印象、読み手が推測する書き手の感情の 3 種類を明確に区別し、統合的に抽出しようという研究は、これまでなかった。申請者らは、この 3 種類の感性情報の対応関係を分析し、明らかにすることで、それぞれの感性情報を相補的に利用する、より高精度な感性情報抽出（感性情報マイニング）の実現を目指す。

### 2. 研究の目的

本研究は、ツイートのような短い文章から①その文章を書いた人（書き手）の感情や②その文章を読んだ人（読み手）が感じた印象、ならびに③読み手が推測した書き手の感情を抽出するための統合的な基盤技術（感性情報マイニング技術）を確立することを目的とする。このとき、本研究において抽出の対象とする感情には、国語学者の中村明が提唱している 10 種類の基本感情をアレンジした「喜び、好き、安心、悲しい、嫌い、怖れ、怒り、恥ずかしい、高揚、驚き」を用い、印象には、申請者らが提案している 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」を用いる。なお、文章を読んだことによる読み手の感情の変化を扱うためには、より高次のユーザモデルが必要になると考え、本研究では扱わない。

### 3. 研究の方法

本研究では、ツイートのような短い文章から書き手の感情、読み手が感じた印象、読み手が推測した書き手の感情という 3 種類の感性情報を抽出するための統合的な基盤技術（感性情報マイニング技術）を確立することを目的として、以下の 3 つの研究を行う。

#### ①書き手の感情を抽出するための技術の確立

まず、インターネットユーザを対象とする大規模なアンケート調査 A を行い、Twitter アカウントの有無に加え、Twitter 上での投稿頻度や閲覧頻度を問うことで、一定以上の投稿頻度を有するヘビー投稿ユーザと一定以上の閲覧頻度を有するヘビー閲覧ユーザを抽出する。次に、このうちのヘビー投稿ユーザを対象とする大規模なアンケート調査 B を行い、各アンケート回答者に 1 個以上のツイートを新たに創作し、回答してもらうとともに、それぞれのツイートを創作したときの自分の感情を 10 種類の基本感情を評価項目とする 5 段階尺度で評価してもらうことにより、ツイートと基本感情との対応関係を明らかにするためのデータを得る。なお、各回答者には「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 語を評価項目として提示し、それぞれの評価項目に対して、該当する基本感情を「含む（4 点）」から「含まない（0 点）」の 5 段階で評価してもらう。

このアンケート調査の結果から得られるデータを分析することで、基本感情どうしの共起確率を求めるとともに、ツイート（ツイートを構成する形態素に基づいて生成されるツイートベクトル）と各基本感情（5 段階評価値）との対応関係を得ることができる。このようにして得られるツイートと基本感情との対応関係を多変量解析手法や機械学習手法を用いて定式化することで、平均的な感性を持った書き手の感情を抽出するための手法を開発する。

#### ②読み手が感じた印象と読み手が推測した書き手の感情を抽出するための技術の確立

まず、研究項目①のアンケート調査 A の結果に基づいて抽出されたヘビー閲覧ユーザを対象に大規模なアンケート調査 C を行う。この調査では、各回答者（すなわち読み手）にアンケート調査 B で収集したツイートを読んでもらい、そのツイートからどのような印象を感じたかを 8 種類の基本印象を評価項目とする 5 段階尺度で、またそのツイートの投稿者（すなわち書き手）がどのような感情を抱いていると思うのかを 10 種類の基本感情を評価項目とする 5 段階尺度で評価してもらうことにより、ツイートと基本印象との対応関係とツイートと基本感情との対応関係を得る。なお、基本感情として回答者に提示する評価項目は、研究項目①と同様、「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 項目であり、それぞれの評価項目に対して、該当する基本感情を「含む（4 点）」から「含まない（0 点）」の 5 段階で評価してもらう。一方、基本印象として回答者に提示する評価項目は、「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」の 8 項目であり、該当する基本印象を「感じる（4 点）」から「感じない（0 点）」の 5 段階で評価してもらう。

以上の手順により得られるデータを分析することで、研究項目①と同様、基本感情どうしや基本印象どうし、基本感情と基本印象の各組み合わせに対し、共起確率を求めるとともに、ツイート（ツイートを構成する形態素に基づいて生成されるツイートベクトル）と各基本感情（5段階評価値）や各基本印象（5段階評価値）との対応関係を得ることができる。このようにして得られるデータを、研究項目①で試した多変量解析手法や機械学習手法を用いて定式化することで、平均的な感性を持った読み手が感じた印象を抽出するための手法と平均的な感性を持った読み手が推測した書き手の感情を抽出するための手法を開発する。

### ③個々の書き手・読み手に応じて適応的に感性情報を抽出するための技術の確立

まず、研究項目①および②で開発された手法を用いて算出された書き手の感情、読み手が感じた印象、読み手が推測した書き手の感情と、各ツイートに対する書き手や読み手の5段階評価の結果を比べることで、理論値と実際値のズレを定量的に扱えるようにする。このとき、読み手が感じた印象や読み手が推測した書き手の感情に関しては、研究項目②で取得するデータを用いることができるが、書き手の感情に関しては、個人差を扱えるようなデータがないので、新たなアンケート調査Dを行い、収集する。具体的には、まず、研究項目①で抽出したヘビー投稿ユーザの中から、アンケート調査Bに参加していないヘビー投稿ユーザを抽出し、このヘビー投稿ユーザを対象とするアンケート調査Dを行う。この調査では、研究項目①と同様、各回答者（書き手）に新たに創作したツイート（20個程度）を回答してもらうとともに、それぞれのツイートを創作したときの感情を10種類の基本感情を評価項目とする5段階尺度で評価してもらう。これにより、回答者1人につき約20個のツイートと各ツイートに対する基本感情との対応関係を得ることができる。

次に、書き手と読み手の各々に対し、各人がツイートを付与した実際値（5段階評価値）と提案手法により算出された理論値とのズレ（差分）を、その人の感性モデルと捉え、ズレ方の似ている人どうしをクラスタ分析によりグループ化する。さらに、各ユーザグループに対し、理論値からのズレ方を多変量解析手法や機械学習手法を用いて定式化することで、研究項目①および②で開発された汎用的な手法により算出された感情値や印象値を個々の書き手・個々の読み手に応じて適応的に補正する手法を開発する。

## 4. 研究成果

本研究では、ツイートのような短い文章から書き手の感情、読み手が感じた印象、読み手が推測した書き手の感情という3種類の感性情報を抽出するための統合的な技術を確認することを目的とし、以下の研究を行った。なお、本研究において、対象とする感情は「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種類であり、印象は「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」の8種類である。

平成29年度は、アンケート調査を実施し、ツイートを収集するとともに、書き手の感情、読み手が感じた印象、読み手が推測した書き手の感情に関するデータを取得した。また、これらのデータを分析することで、感情どうし、印象どうし、感情と印象の組み合わせに対する対応関係を明らかにした。

平成30年度は、ツイートと印象や感情との対応関係を定式化するとともに、ツイートの印象値（読み手が感じた各印象の強さを示す値）や感情値（読み手が推測した書き手の各感情の強さを示す値）を算出する手法を開発した。また、顔文字がツイートの印象や感情に与える影響を定式化し、顔文字を含まないツイートを対象として設計されている従来手法と組み合わせることで、顔文字付きツイートの印象値や感情値を算出する手法を開発した。

平成31年度（令和元年度）は、ツイートの印象値や感情値を算出するために公開されている複数の従来手法を組み合わせることで、ロバスト性（入力文章の印象値や感情値を算出するにあたって算出不能になりにくい性質）と精度面（交差検定時の平均Root-Mean-Square Error）においてより優れた手法を開発した。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 熊本忠彦	4. 巻 J102-D
2. 論文標題 顔文字に対する読み手の印象と読み手が推測する顔文字使用者の感情を相互に変換する手法の提案	5. 発行年 2019年
3. 雑誌名 電子情報通信学会論文誌	6. 最初と最後の頁 854-864
掲載論文のDOI（デジタルオブジェクト識別子） 10.14923/transinfj.2019JDT0001	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 蔣承志, 熊本忠彦	4. 巻 19
2. 論文標題 顔文字がツイート感情の推測に及ぼす影響の定式化 ~ ツイート閲覧者が推測するツイート投稿者の感情を対象として ~	5. 発行年 2020年
3. 雑誌名 日本感性工学会論文誌	6. 最初と最後の頁 39-47
掲載論文のDOI（デジタルオブジェクト識別子） 10.5057/jjske.TJSKE-D-19-00020	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 熊本忠彦	4. 巻 12
2. 論文標題 叫喚ツイート抽出手法の拡張	5. 発行年 2019年
3. 雑誌名 情報処理学会論文誌データベース	6. 最初と最後の頁 1-5
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計11件（うち招待講演 0件 / うち国際学会 2件）

1. 発表者名 秋山和寛, 三澤賢祐, 成田和弥, 熊本忠彦, 瀬本明代
2. 発表標題 消費者投稿文の節単位のネガティブ感情分類における手法の比較検討
3. 学会等名 第9回ソーシャルコンピューティングシンポジウム
4. 発表年 2018年

1. 発表者名 熊本忠彦
2. 発表標題 叫喚表現化によるツイート感情の変化 ~ 読み手が推測する書き手の感情を対象として ~
3. 学会等名 電子情報通信学会技術報告
4. 発表年 2018年

1. 発表者名 秋山和寛, 三澤賢祐, 成田和弥, 熊本忠彦, 瀬本明代
2. 発表標題 CRFを用いた複数ドメインの消費者投稿文におけるネガティブ感情分類
3. 学会等名 第11回Webとデータベースに関するフォーラム
4. 発表年 2018年

1. 発表者名 熊本忠彦
2. 発表標題 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 ~
3. 学会等名 第17回情報科学技術フォーラム
4. 発表年 2018年

1. 発表者名 秋山和寛, 熊本忠彦, 瀬本 明代
2. 発表標題 レビューの印象に基づいた評価の高いレビューの分析
3. 学会等名 第11回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2019年

1. 発表者名 蒋承志, 熊本忠彦
2. 発表標題 顔文字がツイート投稿者の感情推測に及ぼす影響の分析
3. 学会等名 第11回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2019年

1. 発表者名 植木舞, 熊本忠彦
2. 発表標題 ツイート印象の時間依存性に関する分析
3. 学会等名 第14回日本感性工学会春季大会
4. 発表年 2019年

1. 発表者名 Akiyama Kazuhiro, Mitsuzawa Kensuke, Kazuya Narita, Kumamoto Tadahiko, Nadamoto Akiyo
2. 発表標題 Clause-level Negative-opinion Analysis for Classifying Reviews on Multiple Domains
3. 学会等名 The 20th International Conference on Information Integration and Web-based Applications & Services (国際学会)
4. 発表年 2018年

1. 発表者名 秋山和寛, 熊本忠彦, 灘本明代
2. 発表標題 ツイートからの多次元感情抽出手法の考察
3. 学会等名 第8回ソーシャルコンピューティングシンポジウム
4. 発表年 2017年

1. 発表者名 秋山和寛, 三澤賢祐, 成田和弥, 熊本忠彦, 瀬本 明代
2. 発表標題 CRFを用いたレビューにおける節単位毎の感情推定
3. 学会等名 第10回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2018年

1. 発表者名 Kazuhiro Akiyama, Tadahiko Kumamoto, and Akiyo Nadamoto
2. 発表標題 Emotion-based Method for Latent Follower Recommendation in Twitter
3. 学会等名 the 19th International Conference on Information Integration and Web-based Applications & Services (国際学会)
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担者	瀬本 明代  (Nadamoto Akiyo)  (30359103)	甲南大学・知能情報学部・教授   (34506)	