

語の出現頻度と語と語の相関に着目した 感情語辞書の自動生成の検討

柳瀬 恵里

西崎 博光

山梨大学大学院医学工学総合教育部

{eririn,nisizaki}@alps-lab.org

1 はじめに

近年、音声認識技術は発達しているが、基本的に話されている言葉を文字化するのみで、話者の感情まで判別することはできない。例えば、コールセンターの対話記録を参照したい場合、ある語が発話されている場面を取り出すことはできるが、“怒ったお客様”との対応の部分を取り出すことは難しい。従って、音声認識とともに話者の感情を判別し、利用することができれば、音声を利用することの有用性がより高まると考えられる。

話者の感情を判別する研究は様々なところで試みられており、これらは音響的特徴を用いて判別する例が圧倒的に多い [1]。どのような語が話されているかを調べるような言語的な特徴を用いた研究例は少ない中、例えば Shuller らは [2] 音声認識結果で得られた単語を元に SVM (Support Vector Machine) を用いた機械判別手法を提案している。言語情報を用いる場合、音声認識が正しく行えなければ、言語的特徴の利用が困難であること、また、特に日本語の場合は感情判別のためのコーパスが少なく、どのようなフレーズや語句を感情語として設定したら良いのかが難しいことなどが挙げられる。単語に感情極性を付与した辞書 [3][4] が提案されているが、評価軸はポジティブまたはネガティブの2値のみという現状である。そこで本研究では、感情判別に利用できる日本語の単語・フレーズ (本稿ではこれらを感情語と記す) を Twitter¹ から収集し、感情判別のための辞書の自動構築を目指す。また、近年では通常の音声認識よりも高い精度で単語を抽出することができる音声中の検索語検出 (Spoken Term Detection: STD) 技術の研究が行われており [5]、これを利用することで感情判別の性能が改善できるのではないかと考えた。

本研究では感情の分類方法として、8つの基本感情 (喜, 信頼, 怖, 驚, 悲, 嫌, 怒, 予期) と、これら2

つを組み合わせた複合感情で形成される, Plutchik の感情分類 [6] を採用する。まず, Twitter のツイートを集集し, 感情判別のための語句やフレーズを抽出することで, 各感情に対応する感情語を設定した “感情語辞書” を自動生成する。そして, 入力された発話に対し STD 技術を用い, どのような感情語が含まれているのかを検出し, 感情を判別する。

本稿では, 各感情に対する感情語を設定するという目標に重点を置き, 全く同じツイートデータから2種類の方法で感情語辞書を自動生成し, それぞれの辞書を用いた感情の判別結果を比較した。それぞれの辞書は, 語の頻度やエントロピーの情報を用いて構築したもの (辞書1), IBM社の提供する ContentAnalytics [7] を用い, 語と語の相関に着目し構築したもの (辞書2) の2つである。実験の結果, 辞書1では36.3%, 辞書2ではそれを上回る42.2%の判別精度が得られた。

2 感情語辞書の生成

感情語辞書を生成するため, Twitter のツイートを収集し, 感情判別のための語句やフレーズを抽出する。

2.1 感情語設定のためのデータ収集・選定

Twitter のツイート収集のための感情キーワードを設定するため, 感情表現辞典 [8] を用い, 遠藤ら [9] の手法を参考に, 次の条件に該当する単語・表現を候補として拾い出した。

- 書き言葉で使われるような表現, 自分や他人の状態を解説するような表現を含まない,
- 「ない」「ぬ」などの否定語を含まない,
- 「が」「の」「を」などの格助詞を含まない。

さらに得られた単語・表現を, 形容詞, 名詞-形容動詞語幹, 名詞-サ変接続, 副詞, 動詞-自立, 名詞-一般の順で優先的に選択し, 各感情カテゴリにつき10個程度, 合計124個まで絞り込んだ。各感情カテゴリにおける感情キーワードの例を表1に示す。

¹<http://twitter.com/>

表 1: 各カテゴリにおける感情キーワードの例

感情カテゴリ	検索語の例
喜	喜ばしい, 嬉しい など
信頼	和やか, 安心 など
怖	気味悪い, 不安 など
驚	びっくり, ショック など
悲	悲しい, 後悔 など
嫌	切ない, 不快 など
怒	腹立たしい, むかつく など
予期	期待, そわそわ など

この感情キーワードを検索語とし, Twitter API² を利用することで, Twitter に投稿されたツイートを集集³した。続いて, 収集したツイートデータをもとに, 次の2種類の感情語辞書をそれぞれ述べる手順により生成した。

2.2 (辞書 1) 頻度とエントロピーに基づく辞書

2.2.1 クリーニング処理とフィルタリング処理

収集した大量のツイートは, 1 ツイートが複数の文で構成されていたり, 本研究では不要である記号や顔文字を大量に含んでいた⁴りする等, 扱いづらいデータである。そこで, ツイート中の不要な記号類を削除し, 文末を示す記号で1文ごとに区切る⁵等の整形を行ったのち, 感情キーワードが含まれる文のみを抽出, 各感情カテゴリと対応付ける。しかし, 先の処理で取り出した「感情表現キーワードが含まれる文」は, 確実にその感情であるとは言えず, 有効なデータとはならない。そのため, 長谷川ら [10][11] の手法を参考に, 次のフィルタリング条件を設定し, 合致する文章を除外した。

- 感情キーワードと仮定, 引用, 否定, 命令, 逆説の表現が伴うものを除外,
- 感情キーワードが自立語を修飾するものを除外。

2.2.2 形態素・形態素の組み合わせの抽出

各感情カテゴリに対応する文集を, 形態素解析エンジン Mecab⁶ を用いて形態素に分割する。この際, より感情判別に有用なフレーズが得られる可能性がある。接続する2つ組の形態素も合わせて抽出しておく。続いて, あるカテゴリで頻繁に出現する語を調べるため, 感情カテゴリ $C = \{c_1, c_2, \dots, c_n\} (n = 8)$ のうち,

²<https://dev.twitter.com>

³収集期間: 2013年10月12日~2014年8月21日

⁴本研究では音声ターゲットとしているため扱わない。

⁵約7,898万文が得られた。

⁶Yet Another Part-of-Speech and Morphological Analyzer, <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

各カテゴリ c_i での形態素 w の出現回数 $N_{c_i}(w)$ を数える。また, あるカテゴリに偏って出現している形態素を調べるため, 出現回数が100回を超える形態素 w のうち, 各感情カテゴリ c_i における出現確率 $P(w|c_i)$ とエントロピー H_w を, それぞれ式(1)と式(2)により求める。接続する形態素の組み合わせに関しては, あるカテゴリ内で50個以上出現するペアに対して, 2つ組の形態素どちらかが該当カテゴリでの出現確率0.25以上という条件を満たしたものを抽出する。

$$P(w|c_i) = \frac{N_{c_i}(w)}{\sum_{i=1}^n N_{c_i}(w)} \quad (1)$$

$$H_w = - \sum_{i=1}^n \{P(w|c_i) \log_2 P(w|c_i)\} \quad (2)$$

出現確率 $P(w|c_i)$ とエントロピー H_w を利用することで, それぞれの語と各感情カテゴリとの対応付けを行う。各感情カテゴリにおける出現確率に大きな差が出なくなることが確認された2.7未満のエントロピーを持つ形態素を, 出現確率が高い感情カテゴリへ最大2つまで紐付ける。さらに, 出現確率が0.25以下のもの, 語の長さが1文字であるものは, 感情の判別には有用でないとし, 紐付けの対象としない。

以上の条件を満たした形態素について, 式(3)を用いて感情カテゴリ毎に形態素がどれくらいその感情判定に寄与するかを示すスコア $S(w|c_i)$ を計算する。ここで, 形態素の組み合わせに対するスコアは, 両形態素の出現回数, 出現確率の平均を利用して計算することとした。

$$S(w|c_i) = \log_{10} N_{c_i}(w) \times P(w|c_i) \quad (3)$$

抽出した語, 対応する感情カテゴリ, 計算したスコアの3つを一組とし, 感情語辞書に登録する。本節で示した辞書1には, 合計78,505語が登録されている。

2.3 (辞書 2) 語と語の相関に基づく辞書

IBM社が提供するIBM Content Analytics with Enterprise Search Ver 3.0 (以下ICAとする) と呼ばれる大規模テキストマイニングツールを利用する[7]。ICAを用いて, 本研究で設定した感情キーワードと関連の強い語を得ることで, より有用な感情語が得られることが期待される。

収集したツイートの整形は, 重複ツイートの削除, 文末を示す記号で1文ごとに区切るの2点のみであり, これらのツイートに対応する感情カテゴリを対しICA内に登録する。続いて, 初めに設定した感情キーワード124語それぞれと相関の高い語を, 名詞, 動詞, 形容詞, 副詞の4種類の品詞に分け, それぞれ上位から最大500語抽出する。この相関値[12]は, あ

る語 w の全体の出現数に対し、各感情キーワード k と共起する割合がどの程度かを表す値であり、単純な頻度情報ではない。ここで得られた語 w に関連する感情カテゴリ c_i 、相関値 $I_k(w)$ 、出現頻度 $F_k(w)$ を用い、辞書 1 と同様に語と感情カテゴリとを対応付け、式 (4) にて計算した s を 8 カテゴリで 1 で正規化することでその語のスコアとして登録した。

$$s(w|c_i) = \sum_{k \in c_i} \sum_w \{\log_{10} F_k(w) \times I_k(w)\} \quad (4)$$

辞書の形式は上記辞書 1 と同じであり、本節で示した辞書 2 には、合計 8,248 語が登録されている。

3 感情の判別手法

本節では、感情語辞書に登録された語を用いて感情判別を行う手法を述べる。

入力は辞書の生成に利用していないツイート、またはこれを筆者が読み上げた音声であり、出力は Plutchik の 8 つの感情カテゴリまたは感情なしの合計 9 つのカテゴリから、2 つまでとする。この際、正反対の感情同士や、感情なしのカテゴリと他のカテゴリの組み合わせはないものとする。

各入力発話からの感情語の検出には次の 2 種類の方法を用いる。

- テキストで与える入力発話について、感情語辞書の単語を単純な文字列検索で抽出、
- 音声で与える入力発話に対して、感情語辞書の語をクエリとし、名取ら [5] の PTN (Phoneme Transitional Network: PTN) を利用した STD 手法⁷で抽出。

式 (5) に示すように、各入力発話について検出された感情語辞書の語 w のスコア $S(w|c_i)$ を、感情カテゴリ毎に足し込み、最大のスコアとなった感情カテゴリ \hat{c} を決定する。

$$\hat{c} = \arg \max_{c_i \in C} \left\{ \sum_w S(w|c_i) \right\} \quad (5)$$

今回は、発話中に感情語が 1 つも含まれていなかった場合のみ、感情なしが出力されることとした。また、スコアが 2 番目に大きい値となった感情カテゴリについて、1 位とのスコア差が 10.0 以内であれば、このカテゴリを含め上位 2 つのカテゴリを出力するものとする。

4 評価実験

各テストデータに対し、提案手法を用いて得られた出力と、人間が分類した結果を比較する評価を行った。

⁷10 種類の音声認識システムの出力結果を利用した。使用した音響モデル、言語モデルは名取らの STD 手法と同様である。

表 2: 感情判別の正解率 [%]

	辞書 1	辞書 2
テキストツイート	42.2	45.7
音声ツイート (STD 利用)	36.3	42.2

4.1 評価実験条件

4.1.1 テストデータ

感情語辞書を作成するために利用したツイートを含まない 240 ツイートを用意し、各ツイートについて、6 名の被験者に Plutchik の 8 種類の感情および感情なしの合計 9 個のカテゴリから、最大 2 つまでの感情を選択して頂いた。次の条件を満たすようにテストデータを選別し、各ツイートについて人間の感覚で分類した結果を得るとともに、223 ツイートを有効なテストデータとして採用した。

- 過半数の回答が得られているカテゴリかつ、評価の票数が上位 2 位までの感情カテゴリを人間の分類結果として採用、
- 2 位のカテゴリが過半数の回答を得られていなければ、上位 1 位のカテゴリのみを採用、
- 3 つ以上のカテゴリが同じ得票数、どのカテゴリも過半数の回答を得られていないものなどは、評価対象から除外。

4.1.2 感情判別正解基準

ツイート音声の感情判別評価では、4.1.1 節で説明した人間による感情判別結果に対して、提案手法による判別結果がどの程度一致しているのかの正解率を評価尺度とする。今回、人間による感情判別でも 1 つのツイートにつき複数の感情が付与されることが多く見られたことから、複数 (2 つまで) の感情ラベルを持つツイートに対しては、提案手法による判別結果がどちらか一方の感情ラベルと一致すれば正解とした。また、判別においては最大 2 つの感情ラベルを出力すると説明しているが、このうちどちらかの感情ラベルが人間の判定した感情ラベルと一致すれば正解と判定する。

4.2 評価実験結果と考察

2 節で生成したそれぞれの辞書を用いて、テキストツイート、音声ツイートに対して感情判別実験を行った。判別結果を表 2 に示す。辞書 1 の語数 78,505 語を大幅に下回る 8,248 語を持つ辞書 2 の方が、テキスト、音声ツイートのどちらの形式のデータに対してもより高い正解率が得られている。

また音声ツイートから STD によって感情語を抽出すると、STD の誤検出により本来音声中に存在しない感情語を多く検出してしまい、結果的に不要なスコアがつけられてしまっている。このスコアの悪影響が、

テキストツイートに比べて正解率が低下した原因であると考えられる。これを防止するため、辞書の内容に関わらず、短い感情語での湧き出し誤りという STD の弱点に対応する必要があると考えている。

4.2.1 (辞書 1) 頻度とエントロピーに基づく辞書

本辞書の語とスコアは、語が出現する頻度に大きく依存するため、「新しい」が“悲”，「いいね」が“怒”等、そぐわないカテゴリに対応付けられている感情語が見受けられた。形態素の組み合わせも同様であるため、「た-事」「て-泣い」など、感情語らしくないものが辞書には多いことが確認された。しかしその中で、「いちばん-感動」「ぜったい-楽しい」などの長い組み合わせでスコアがつき、判別に成功している例も稀に見られている。長い語であれば STD の誤検出の悪影響が小さくなるため、このような語が多く抽出できるような工夫をすることが望ましい。

4.2.2 (辞書 2) 語と語の相関に基づく辞書

辞書 1 ではそぐわないカテゴリに対応付けられている感情語が原因で不正解となったテストデータについて、辞書 2 では正解できた例も存在している。例えば、「おかしいだろ」と含まれる発話では、辞書 1 では“信頼”スコアがつけられていたが、辞書 2 では“嫌”が強く、「いつまでもこんなことで」と含まれる発話では“悲”カテゴリへの関連が強かった等が挙げられる。しかし、辞書 2 では辞書 1 と異なり、1 文字の語を登録しないという制限を設けなかったため、より多くの短い感情語が辞書に登録されてしまっているという欠点が挙げられる。語と語の相関をうまく取り入れたまま、辞書 1 のようにより長い単位で感情語を抽出する、語と語の組み合わせでスコアを付与するなどの対策が必要そうである。

5 おわりに

本稿では、言語的特徴を用いて話者の感情を判別することを目標とし、Twitter のツイートを収集することで感情判別のための語句やフレーズを抽出し、感情語辞書を自動生成する 2 種類の方法を提案した。生成した 2 種類の感情語辞書を用いて、感情判別とその評価を行い、音声ツイートに対して辞書 1 では 36.3%、辞書 2 では 42.2% の判別精度が得ることができた。ツイートの集合から、頻度やエントロピーの情報だけではなく、語と語の相関にも着目して感情語辞書を生成することでより効果が高まる可能性がある。今後は、辞書へ登録する語やスコア付与の方法の改善、STD による湧き出し誤りへの対策により正解率の向上を目指す。

謝辞

本研究の実施に際して、IBM 社の“IBM Content Analytics with Enterprise Search Ver 3.0”を利用して頂きました。本ソフトウェアの基本的な使い方、活用法等をご教授くださいました IBM 社東京基礎研究所の那須川哲哉様に感謝いたします。

参考文献

- [1] 門谷信愛希他. 音声に含まれる感情の判別に関する検討. 信学技報 LC2000-34, SP2000-82, pp. 43-48, 2000.
- [2] B. Schuller et al. Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine - belief network architecture. In Proc. of ICASSP2004, Vol.I, pp. 577-580, 2004.
- [3] 小林のぞみ他. 意見抽出のための評価表現の収集. 自然言語処理 Vol.12 No.3, pp. 203-222, 2005.
- [4] 東山昌彦他. 述語の選択選好性に着目した名詞評価極性の獲得. 言語処理学会第 14 回年次大会論文集, pp. 584-587, 2008.
- [5] S. Natori et al. Spoken term detection using phoneme transition network from multiple speech recognizers' outputs. *Journal of Information Processing, Vol.21, No.2*, pp. 176-185, 2013.
- [6] The nature of emotions by plutchik. (online), <<http://www.fractal.org/Bewustzijns-Besturings-Model/Nature-of-emotions.htm>>, (2013-06-10).
- [7] W-D. Zhu et al. In *IBM Watson Content Analytics Discovering Actionable Insight from Your Content*. IBM Redbooks, 2014.
- [8] 中村明. 感情表現辞典. 東京堂出版, 1993.
- [9] 遠藤大介他. 係り受けを利用した感情生起表現の抽出. 言語処理学会第 12 回年次大会, pp. 947-950, 2006.
- [10] 長谷川貴之他. オンライン上の対話における聞き手の感情予測. *DEIM Forum 2013 D9-1*, 2013.
- [11] 長谷川貴之他. 聞き手の感情を喚起する発話の分類と生成. 信学技報, NLC2012-29, pp. 25-30, 2012.
- [12] 那須川哲哉他. 機器の不具合を記述した日本語と英語のコーパスにおけるオノマトペ. 言語処理学会第 16 回年次大会, pp. 154-157, 2010.