

携帯メールにおける顔文字自動付与手法

鈴木信雄 津田和彦

筑波大学大学院 ビジネス科学研究科

1. はじめに

近年、携帯電話は日常的なツールとなり、個人の情報通信ツールとして広く普及するに至っている。このような携帯電話は、インターネットに接続され、音声通話に加えてメールによるコミュニケーションに多く利用されている。通常、メールでは、文字を使ったコミュニケーションのために、言葉によるコミュニケーションでは発話速度や音の高低で表現される感情が伝えづらいという欠点が生じる。

この欠点を補うため、顔文字が多用されている^[1]。本論文では、携帯から入力された文章を収集し分類した結果から、使用された顔文字の感情情報を分析すると共に、携帯メール本文の持つ感情を予測し、自動的に文末に顔文字を付与する手法について述べる。

2. 感情の分類

まず、顔文字がどのような感情を表すかを定めるためには、人間の感情そのものの分類を定義する必要がある。これまで数多くの感情分類法が提案されているが、ここでは、以下の理由により、多様な感情を表現できるブラッチクの8つの基本感情に基づいた分類法を使用する^[2]。

- ・感情の強度を表現している
- ・感情の対極を表現している
- ・感情に伴う行動が定義されており、行動と感情の関連が示されている
- ・基本感情の混合が定義されており、多様な感情を説明できる

表1にブラッチクの感情と感情語の関係を示す。今回は、これらの感情語と次項で述べる追加の感情語を用い、顔文字の感情を表現した。

3. 感情を表す顔文字の定義

前項にて示した感情語に対応する顔文字を定義する。使用する顔文字は、Microsoft IME 2003で定義されている172個の顔文字の中から、複数人のアンケートにより各感情語に相当する代表的な顔文字をピックアップした。

ここで、単純にIMEとブラッチクの感情語を対

表1. ブラッチクの感情分類

基本感情	(強い) 強度 (弱い)
受容	愛、好意、信頼、寛容、受容
恐れ	戦慄、怯え、恐れ、懸念、慎重
驚き	仰天、驚き、困惑、動揺、疑念
悲しみ	悲嘆、憂い、悲しみ、落胆、悲哀
嫌悪	憎悪、嫌悪、反感、嫌い、厭気
怒り	憤怒、激怒、怒り、憤然、敵意
予期	予期、期待、警戒、好奇心、気配り
喜び	得意、喜び、満足、愉快、平穩

応させると、高い頻度で使用される顔文字や感情語が脱落してしまうことが懸念される。そのため、収集したデータから携帯メールで良く使われるが、ブラッチクの感情語には出てこない感情語を抜き出して追加した。また、逆に、ブラッチクの感情語にはあるが、IMEで不足しているものも追加した。追加の感情語を表2に示す。

表2. 追加の感情語

焦り	普通の別れ
呆れ	悲しい別れ
謝罪	依頼
眠い	恐れ
疲労	愛
挨拶	受容

以上のように顔文字と感情語のペアをピックアップした。その一部を表3に示す。これらを「顔文字辞書」としてシステム内に保持する。

表3. 顔文字辞書(一部)

顔文字	感情語
(^_^)/	挨拶
(^_^)	焦り
(>_<)	嫌悪
(^m)	嬉しさ
(@_@)	驚き
(T_T)	悲しみ
(^_^)	期待
(...?)	疑念
m(_)_m	謝罪、依頼
(.o.)	動揺
<(^')>	得意

4. 感情判別のための形態素解析

携帯メール文にどのような感情が現れているか

Automatic face-characters generation method for cellular phone E-mail
SUZUKI Nobuo, TSUDA Kazuhiko
Graduate School of Business Sciences, University of Tsukuba

を判別するために、本手法では形態素解析を利用する。形態素解析を用いることにより、全ての文字列をサーチする手法に比べて高精度に処理できることが期待できる。形態素解析には、茶筌^[2]を使用する。

茶筌にて定義されている形態素の中から、感情を表現する品詞を選定した。これを「感情品詞」と呼ぶ。表4に感情品詞の一覧を示す。

表4. 感情品詞一覧

感情品詞	単語例	茶筌辞書の単語数
名詞 サ変接続	悪化 ひと安心	12,041
名詞 形容動詞語幹	安易 だめ	3,313
名詞 ナイ形容詞語幹	申し訳 仕方	42
形容詞 形容詞・イ段	哀しい 楽しい	654
形容詞 不変化型	かっこいい きもちいい	8
合計単語数		16,058

次に、実際に収集した携帯電話から入力された2,218個の文から、各感情品詞に対する単語を抽出した。これらの単語がどのような感情を表現しているかをアンケートにより決定し、表5のような「感情辞書」を構築した。

表5. 感情辞書（一部）

品詞	感情語	感情
名詞-サ変接続	お願い	依頼
	うっかり	動揺
名詞-形容動詞語幹	不安	困惑
	不気味	嫌悪
名詞-ナイ形容詞語幹	仕方	悲哀
	申し訳	謝罪
形容詞・イ段	よろしく	挨拶
	あいくるしい	愛
形容詞 不変化型	かっこいい	得意

5. 顔文字自動付与手法

前項までで述べた顔文字辞書と感情辞書を用い、携帯メール文に対して、感情を表現する顔文字を付与する手法について述べる。

本手法は、以下の手順により実行される。

1文を入力。

形態素解析により、文の最後の感情品詞を取得する。（携帯メール文は、文の最後に顔文字が付くことが多い。本手法でも文末に顔文字を付与する。）

感情品詞と実際の単語をキーとして、感情辞書から感情語を取得する。

その感情語に対する顔文字を顔文字辞書から取得する。

取得した感情品詞が「名詞-ナイ形容詞語幹」である場合、直後に「助動詞 特殊・ナイ」（「～ない」など）があるかをチェックする。あれば、で取得した感情と反対の感情を示す顔文字とする。（プラチックの感情分類法により反対の感情を求めることが可能）

入力された1文の文末の句読点を、上記の顔文字に変換して出力する。

6. 評価実験

携帯電話から入力された顔文字付きの65文に対して、実際の顔文字と本手法で取得された顔文字とを比較した。結果を表6に示す。

表6. 評価結果

	文の数	割合
正解	57	87.7%
不正解	8	12.3%
合計	65	100.0%

評価結果の中で不正解になった原因は次のとおりである。

接続助詞の解釈不可

「けどね」などは、当該1文のみの情報では意味が曖昧で、人間でも顔文字が無いと感情の判別は難しい。例：「楽しかったけどね(^-^)」

範囲外の形態素

例えば、「もし心配なら猫の本とか見て調べてみて下さい V(^-^)^V」では、本手法では「心配」に対する顔文字を選択したが、実際は「調べてみて下さい」（動詞 五段・ラ行特殊）に対する顔文字となっている。これは、対象とする形態素を拡大することで対応可能である。

7. まとめと今後の課題

評価実験の結果、87.7%の正解率となり、本手法の有効性を確認できた。今後は、あいまいな表現などに対応するため文脈による感情の理解について検討を進める予定である。

参考文献

- [1]中村,池田,乾,小谷:対話システムにおける顔文字の学習,情報処理学会 NLP154-24, 2003/3/7
- [2]福井:感情の心理学,川島書店,1990
- [3]松本:形態素解析システム「茶筌」,情報処理 Vol.41 No.11,2000/11