Computers in Human Behavior 27 (2011) 1553–1564

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Estimation of word emotions based on part of speech and positional information

Kazuyuki Matsumoto *, Fuji Ren

Faculty of Engineering, University of Tokushima, 2-1 Minamijosanjima, Tokushima 770-8506, Japan

ARTICLE INFO

Article history:
Available online 16 November 2010

Keywords:
Emotion recognition
Emotion corpus
Sensibility

ABSTRACT

Recently, studies on emotion recognition technology have been conducted in the fields of natural language processing, speech signal processing, image data processing, and brain wave analysis, with the goal of letting the computer understand ambiguous information such as emotion or sensibility. This paper statistically studies the features of Japanese and English emotional expressions based on an emotion annotated parallel corpus and proposes a method to estimate emotion of the emotional expressions in the sentence. The proposed method identifies the words or phrases with emotion, which we call emotional expressions, and estimates the emotion category of the emotional expressions by focusing on the three kinds of features: part of speech of emotional expression, position of emotional expression, and part of speech of the previous/next morpheme of the target emotional expression.

1. Introduction

With the rapid diffusion of the Internet, the opportunities to access text information written in foreign languages have increased. Text-based communications such as e-mail and SNS using personal computer or mobile phone are filtering into our daily lives. These situations have enabled us to obtain public reports or opinions on the issues in the world in real time on the Web. However, precisely understanding what is described on the Web is still very difficult when the languages used are unfamiliar to us. We understand the details from the information provided in our native languages, although such information is limited in quantity and diversity. Translation and summarization of information might fail to convey the contents of the original information as it is. Against this background, there are increasingly more demands to understand the information on the Web in the original written language.

One of the research studies undertaken to answer this demand focuses on "opinion mining" and studying how to extract the information such as "estimation," "demand," "emotion," "impression," "division," etc. from text and judging their attributes (Turney, 2002). For opinion analysis, extracting such emotional expressions is inevitable. The extracted emotional expressions will be useful in understanding the writer's intention and emotion. Various emotional expressions are used in our daily conversations. As shown in Table 1, due to the varieties of expressive forms and the emotional levels that each expression describes, it is difficult for us to use the appropriate emotional expressions in the right situations.

In this paper, we analyze the features of emotional expressions using the Japanese–English parallel corpus' annotated emotional information for the purpose of automatically understanding and judging the category of emotional expressions in Japanese and English. Then we propose a method to judge the emotion category expressed with the emotional expression by using the extracted features. Section 2 introduces related works. In Section 3, our proposed method is described. Section 4 evaluates our method using closed and open data. Finally, Section 5 offers a conclusion and indicates possible future work.

2. Related works

There are already methods that attempt to extract emotional expressions or evaluation expressions from a large amount of text data. For example, Nakayama, Eguchi, and Kando (2004) extracted emotional expressions considering the modification relation of the seed words.

Sato and Nakagawa (2006) tried to extract emotional expressions from Amazon.com customer reviews using expression patterns based on the "PlefixSpan" method. Takamura, Inui, and Okumura (2005) expressed the impression of a word with the positive or negative emotion polarity value and applied the value to extract evaluation expressions. Expressions were evaluated to indicate the expression describing whether the object is likable or not. Okanohara and Tsujii (2007) classified evaluation documents written in Japanese using the Support Vector Machine (SVM) and Naive Bayes classifier. They used polarity emphatic word such as "very" emphasizing "good" and classified the documents into three types: positive, negative and neutral. In a research study investigating positive/negative impression of words, Mera, Ichimura, Aizawa, and Yamashita (2002) proposed a method to calculate emotion based on a word's favorability value. They used formulas to calculate

* Corresponding author. Tel.: +81 88 656 9684; fax: +81 88 656 6575.
E-mail address: matumoto@is.tokushima-u.ac.jp (K. Matsumoto).

0747-5632/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chb.2010.10.028

Table 1
Example of emotional expressions.

Type	Example	Emotion
Word	*pun-pun*	Anger
	hisou	Sorrow
Idiom	*hara ga tatsu* (feel aggravated)	Anger
	ware wo wasureru (be all over the place)	Crazy
4-character idiom	*daikyo shisshoku*	Surprise
	gishin annki	Anxiety

emotion for each case frame and judged whether the sentence was positive or negative based on the calculation.

These research studies focused on extracting emotional expressions or estimating emotion from sentences but not estimating the emotions that the emotional words presented. To extract emotional expressions and estimate the emotion categories that the expressions describe it is necessary to judge the emotion attribute value of the words. The emotion attribute value shows the relative emotional strength of a word in a sentence (Matsumoto, Minato, Ren, & Kuroiwa, 2005; Matsumoto, Mishina, Ren, & Kuroiwa, 2007). Our research group has been engaged in developing a technique to recognize human emotion using a computer (Ren, 2009). We have analyzed emotions expressed by words or sentences (Minato, Matsumoto, Tsuchiya, Kuroiwa, & Ren, 2007; Minato, Matsumoto, Ren, & Kuroiwa, 2007) and estimated the emotion of sentences (Matsumoto, Minato, Tsuchiya, & Ren, 2008a; Matsumoto, Minato, Tsuchiya, & Ren, 2008b; Matsumoto, Ren, Kuroiwa, & Tsuchiya, 2007; Minato, Bracewell, & Kuroiwa, 2008; Minato, Matsumoto, Ren, Kuroiwa, & Tsuchiya, 2008).

This paper focuses on the words used for emotional expressions and aims to estimate the emotion attribute value of such words. If we can judge which words express which emotions, a sentence including emotional expressions can be obtained. This can be applied not only to estimate the emotion in a sentence but to extract the important sentences from a document because sentences expressing emotions tend to reflect the opinion of the writer and are considered to be important sentences. WordNet-Affect is a dictionary that lists and classifies words that express emotion. This dictionary classifies words belonging to emotional concepts as hyperonyms, hyponyms, synonyms, antonyms, etc. The labels for emotional concepts are: COGNITIVE STATE, PHYSICAL STATE, HEDONIC SIGNAL, EMOTION-ELICTING SITUATION, EMOTIONAL RESPONSE, BEHAVIOUR, ATTITUDE and SENSATION.

In this study, we used the originally constructed corpus without using such a dictionary.

3. Proposed method

We propose a method to estimate the emotion attribute of emotional expressions based on the features extracted from a Japanese–English bilingual emotion corpus. Fig. 1 shows the flow of extracting and judging the emotion attribute of the emotional expressions in the input sentence.

At first, we explain method 1 to extract an emotional expression and estimate the emotion attribute of the emotional expression.

In step1, an input sentence is analyzed by a morphological analysis system and split into units of morpheme.

In step2, based on the emotional polarity correspondence table (Takamura et al., 2005), a word that is a content word and has an absolute value of emotional polarity over the threshold value is judged to be an emotional expression.

In step3, the emotion category expressed by the word is judged based on the features of the word judged to be an emotional expression and the information about its neighboring words.

Specifically, the following features extracted from the Japanese–English bilingual emotion corpus are used in judging the kind of the emotion.

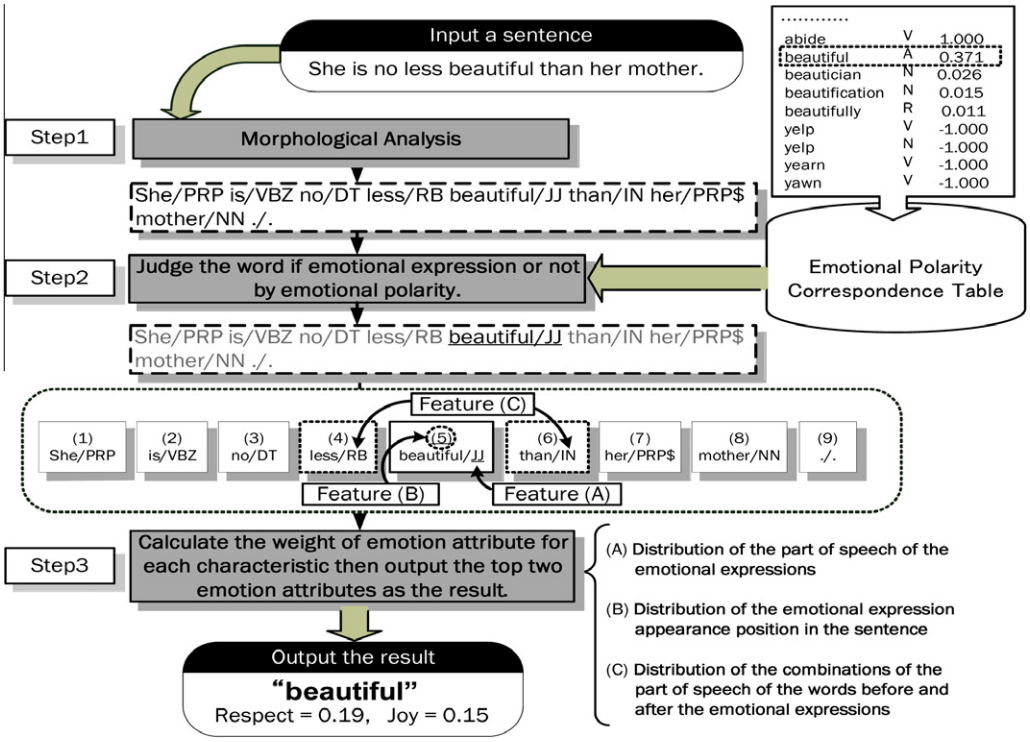

Fig. 1. Flow of the proposed method.