A Method for Classifying Emotion of Text Based on Emotional Dictionaries for Emotional Reading

Futoshi Sugimoto and Yoneyama Masahide

***Abstract*— Representing emotional expressions in text-to-speech synthesis is an interesting subject. The ultimate purpose of our research is to develop an automatic reading system which reads text aloud such as novels with emotion. Our strategy for constructing the system is that we classify the emotion of a text in perspective based on the distribution of emotional words, and classify the emotion of a sentence based on the emotion of nouns, adjectives or verbs composed in the sentence instead of understanding the meaning of the text, and synthesize speech using partially optimum prosodic parameters.**

I. Introduction

In recent years, it has become possible to make a computer read expository text naturally with synthetic speech because of the technical progress of speech synthesis. There still, however, remain problems to be solved in text-to-speech when we treat text, which includes emotional description and express the emotion in synthetic speech. These problems consist of the control of prosodic parameters to express emotion and the classification of emotion in a text. The emotion expressed in a speech can be controlled by operating three prosodic parameters: tempo, pitch and power [1]. Although the relation between these parameters and emotion is qualitatively clear, we attempt to evaluate the quantitative characteristics of the parameters on the five kinds of emotions, which are joy, sorrow, anger, surprise, and neutral emotion, through the analysis of the speeches which some subjects read novels with emotion [2].

We separate the classification of emotion into two processes. In one process, we partition a text into discourse units by using the Hearst method, and determine the emotion of each discourse unit based on the frequency of emotional words distributed throughout the unit. In the other process, the sentences where the emotion is raised are found, and the emotion of each sentence is classified by each method suitable for that type of sentence. In this paper, we focus on the emotional classification of text.

II. Emortional Reading System

Although we focus on the classification of emotion in a text in this paper, we explain the outline of the system that we are challenging as the ultimate goal and our reading policy.

A. Outline of the System

The system that we propose in this paper is organized into two parts, emotional classification and speech synthesis. The classification of emotion is separated into two processes. In one process, a text is partitioned into discourse units by using a method that is often used in the segmentation of text, and the emotion of each discourse unit is simply classified based on the frequency of emotional words distributed throughout the unit. In the other process, the emotion of each sentence is classified based on the type of the sentence and the emotional

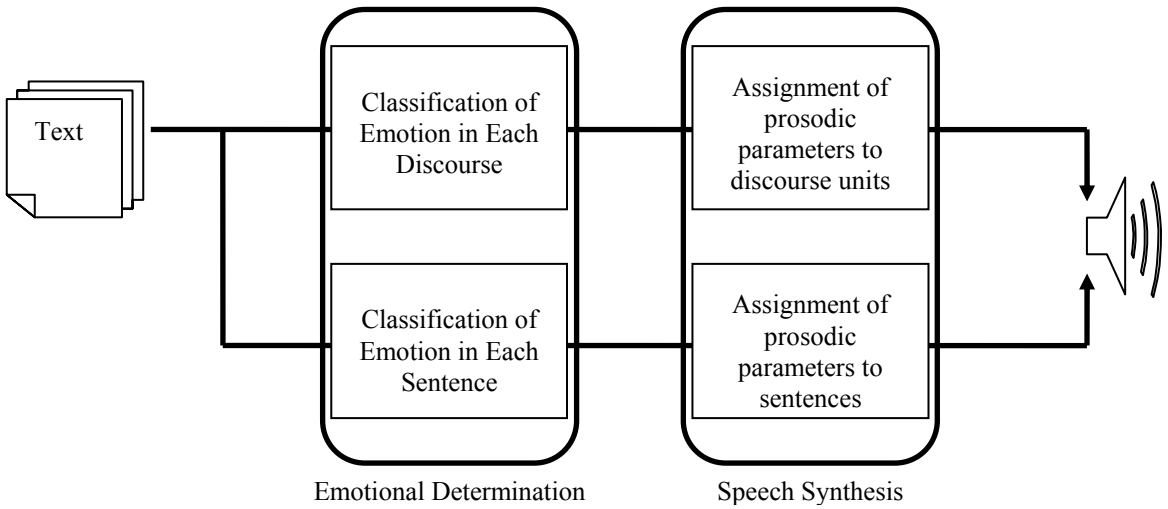

Fig. 1. Outline of the Emotional Reading System

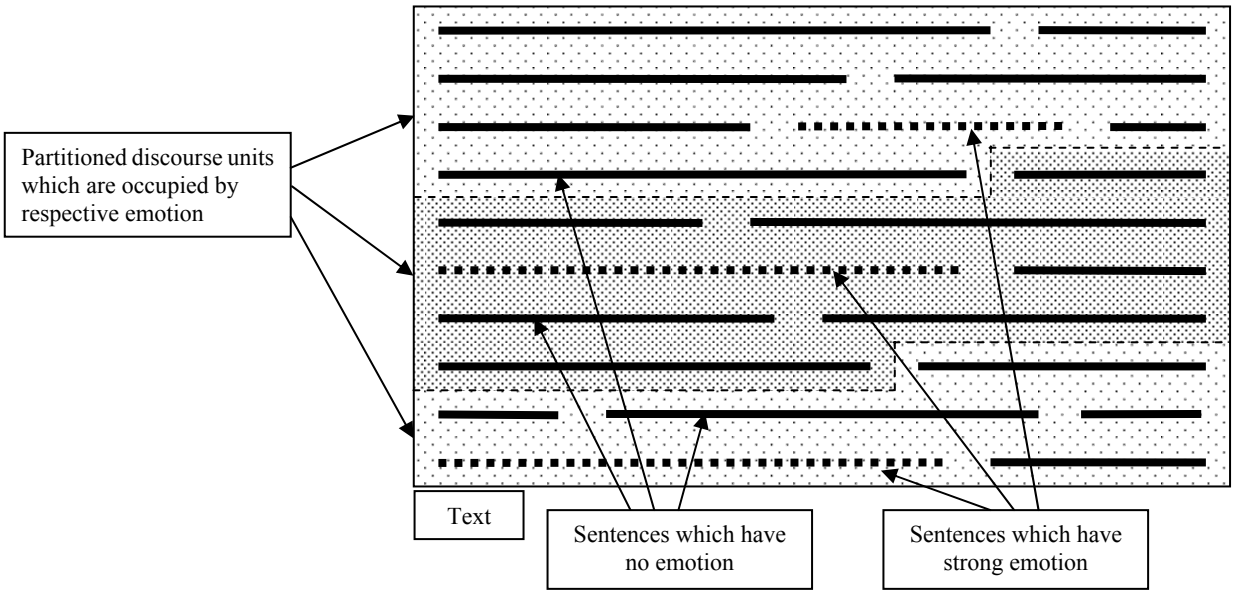

Fig. 2. Partitioning of Text

words appeared in the sentence.

B. Reading Policy

Figure 2 illustrates the partitioning of the text. Each discourse unit is occupied respective emotion. Each discourse unit is occupied respective emotion. We do not assign the extreme values of prosodic parameters, pitch, power and tempo, for reading a whole discourse unit, because the purpose in this process is to express the atmosphere of each discourse. Meanwhile, the sentence which has stronger emotion within the discourse unit is read with the prosodic parameters controlled strongly, and then the raised emotion can be expressed.

III. Emotional Classification Of Discourse Units

In the classification of emotion in perspective, we partition a text into discourse units, and decide the emotion that is expressed in each discourse unit.

A. Partitioning of text

In partitioning text into discourse units, we applied to text tiling algorithm [3]. A part of text is divided by putting two blocks on the text, and the similar level between the two blocks is calculated by the equation (1).

$$sim(b_1, b_2) = \frac{\sum_t w_{t,b_1} w_{t,b_2}}{\sqrt{\sum_t w_{t,b_1}{}^2 \sum_t w_{t,b_2}{}^2}} \quad (1)$$

Where t can be described as the words that occur in the text and wtbi is the weight that word t has within the context in block i. This weight is generally calculated as the word's frequency within the block. Finally, dividing points are decided at the points where the similar level is lower than a criterion.

Although word t is generally assigned to one word, we assigned it to a group of words that has the same meaning referring to a thesaurus [4]. This extension from words to synonyms can make the similarity in the equation (1) have more significant connection.

B. Emotional Classification

In order to decide the emotion that is expressed in each discourse unit, we tag all nouns, adjectives and adverbs appeared in the discourse unit with pleasantness, unpleasantness or neutral using an emotional words dictionary, and decide the emotion of each unit based on the frequency of pleasantness or unpleasantness.

The emotional words dictionary that we used in this paper was compiled through an experiment using 20 subjects whose ages vary from 20s to 60s. Each subject was required to classify nouns, adjectives, verbs and adverbs into five emotions mentioned above. The words used in this experiment are contained in IPAL [5]. The emotional strength of each word is calculated by the following equation:

$$\frac{\text{number of subjects who classified the word into the emotion}}{20} \quad (2)$$

Furthermore, we divided the words into pleasantness, unpleasantness or neutral based on the strength of emotion that each word has. The words whose emotion is joy were classified into pleasantness, while the words whose emotion is sorrow and anger were classified into unpleasantness. The

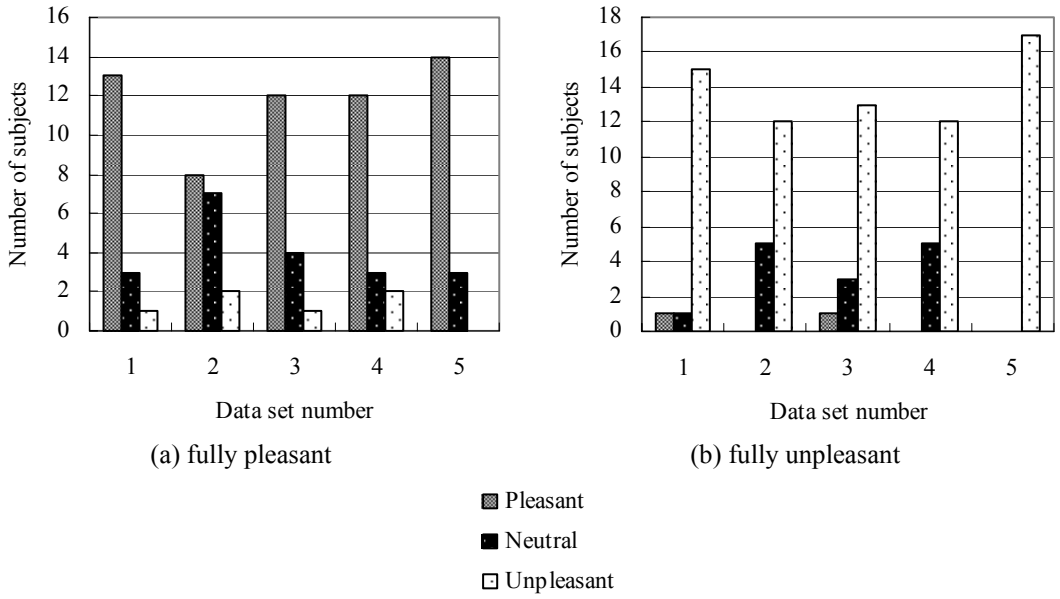

Fig. 3. Results of Verification about "fully pleasant" and "fully unpleasant"

words whose emotion is surprise were classified into pleasantness or unpleasantness based on the meaning of each word.

C. Verification

We verified the suitability of the method mentioned above through an experiment. The material used in the experiment is textbooks for lower classes of elementary school. We prepared four types of discourse units, which are "fully pleasant", "fully unpleasant", "almost pleasant", and "almost unpleasant". In "fully pleasant" type of unit, all the words are pleasant; in "almost pleasant", almost the words are pleasant. We made 5 data sets for each type of unit. 17 subjects were asked to read all data sets and classify each discourse unit into pleasant, unpleasant, or neutral.

Results of verification about "fully pleasant" and "fully unpleasant" are shown in Figure 3. The majority of the subjects in case of "fully pleasant" classified into pleasant or neutral, while in case of "fully unpleasant" unpleasant or neutral. The result shows that the method adopted in this paper is rather effective.

IV. Emotional Classification Of Sentence

In classification of local emotion, a simple sentence is classified into one of the five emotions; joy, sorrow, anger, surprise, and neutral emotion based on the grammatical structure and the emotional words.

A. Grammatical Structure of Simple Sentence

The grammatical structure of a simple sentence is obtained through morphological analysis and dependency analysis. The morphological analysis is a technique to divide a sentence into morphemes that are primitives of a sentence, and allocate a part of speech to each morpheme. On the other hand, the dependency analysis is to obtain the dependency among clauses. The sentence pattern can be judged based on these two analyses. We use three types of sentence, adjective sentence, noun sentence and verb sentence.

After detecting the sentence pattern, the sentence is classified into one of the five emotions mentioned above using an emotional elements dictionary.

B. Emotional Elements Dictionary

Japanese generally has tree types of sentences; adjective sentence, noun sentence, and verb sentence. The rules that determine emotion of a sentence are different from each type of the sentence. The emotional elements dictionary contains the combinations of the words as well as the words that play an important role in classifying emotion of a sentence. We compiled this dictionary with the words that were described in Section 3.2 and the combination that will be described below.

C. Emotion of Sentence

Adjective sentence and noun sentence correspond to an English sentence, S+V+C. In the adjective sentence C is an adjective, while in the noun sentence a noun. Verb sentence corresponds to S+V or S+V+O. The emotion of adjective/noun sentences is influenced by the emotion of an adjective or of a noun assigned in the complement as in the following examples. In these cases, the adjective and the noun become the emotional elements. The following

examples are adjective sentence. As these examples show, the adjective in the complement controls the emotion of sentence.

(1) この本はおもしろい。This book is interesting.
emotion of sentence = Joy 1.0
S : This (Neutral) book = Neutral
V : is
C : interesting = Joy 1.0

(2) 彼女は幸せだ。She is happy.
emotion of sentence = Joy 1.0
S : She = Neutral
V : is
C : happy = Joy 1.0

The emotion of verb sentences is determined by a combination of a noun of the subject and a verb of the predicate as in the following examples. In this case, the combination becomes the emotional element.

(1) 美しい婦人が嘆き悲しんでいる。A beautiful lady weeps.
emotion of sentence = Sorrow 0.9
S : A beautiful (Joy 0.7) lady = Friendly
V : weeps = Sorrow 0.9

(2) 憎い敵が負ける。A hateful enemy loses.
emotion of sentence = Joy 0.7
S : A hateful (Anger 0.9) enemy = Hostile
V : loses = Sorrow 0.8

In the example (1), as a beautiful (Joy 0.7) lady is friendly to a reader, the sentence inherits the same emotion from the verb of predicate. On the other hand, in the example (2), as a hateful (Anger 0.9) enemy is hostile to a reader, the emotion of the sentence changes from the emotion of the verb of the predicate. As these examples show, the emotion of verb sentence changes from the emotion of verb depending on the characteristics of the subject. The emotional elements dictionary also contains these combinations.

V. Conclusions

The system proposed in this paper can classify the emotion of a text in perspective based on the distribution of emotional words and classify the emotion of a sentence based on the emotion of nouns, adjectives or verbs composed in the sentence, instead of understanding the meaning of the text. Although Artificial Intelligence technology has been developed, it is still impossible to understand the meaning of text. Therefore, we proposed this method which is based on the dictionary. Our dictionaries do not cover enough fields, so we should increase the number of words and the combinations, and then we can expand the number using a thesaurus as well.

REFERENCES

[1] Iida, A., Higuchi, F., Campbell, N., Yasumura, M. "A Corpus-based Speech Synthesis System with Emotion", *Speech Communication*, Vol. 40, No.1-2, 1999, 161-187.

[2] Sugimoto, F., Yazu, K., Murakami, M. and Yoneyama, M. "A Method to Classify Emotional Expressions of Text and Synthesize Speech", *Proceedings of the IEEE-EURASIP International Symposium on Control, Communications, and Signal Processing*, 2004, 611-614.

[3] Hearst, M. "A. Multi-Paragraph Segmentation of Expository Text", *Proceedings of the 32nd Annual Meeting of Association for Computational Linguistics*, 1994, 9-16.

[4] National Language Laboratory in Japan, *Bunrui Goi Hyou*, Dainihon Tosyo, 2002.

[5] Information-technology Promotion Agency, "IPA Lexicon of the Japanese Language for computers", *http://www.ipa.go.jp/STC/NIHONGO/IPAL/ipal.html*, 1990.