Topics

Prediction of Low Recognition Rate Words

Ryuta Terashima, Human Interface Lab.

1. Introduction

In recent years, an interface using speech recognition has been adopted in in-vehicle information equipment suchi as a car navigation system for the drivers convenience and improved safety. Since the recognition rate of such a system significantly affects the task achievement time, it is important to detect words of low recognition rate in the development phase and take necessary measures for assuring the system quality. However, the car navigation system is required to recognize some several hundred thousand words for the names of places. Therefore, it is practically difficult to evaluate the recognition performance by collecting speech data for all words. In this connection, we have developed a method to detect words of low recognition without using actual pronunciation.

2. Prediction method of word recognition rate

2.1 Word pair and word recognition rates

We will discuss a vocabulary set W consisting of N number of words. Supposing that the minimum value of the word pair recognition rate for one word (minimum word pair recognition rate), that is, on the word pair consisting of words w_i and w_j, the word pair recognition rate of w_i is $r_{wi,\,j}$, and the upper limit of the word recognition rate for w_i is considered to be $\min(r_{wi,1'}\ r_{wi,2'}\ \ldots,\ r_{wi,\,N})$. Therefore, the present method will predict the difficulty of word recognition from the minimum word pair recognition rate.

2.2 Prediction of word pair recognition rate using the distribution of log likelihood difference

Generally, the goodness of fit (log likelihood) of a word to voice can be calculated using the sum of the log likelihood of respective phoneme segmentations. In the recognition of a word pair, if a value obtained by deducting the log likelihood of an incorrect word from that of the correct word is a positive number, it means that the word has been correctly recognized. Therefore, the word pair recognition rate can be predicted by obtaining the distribution of the log likelihood difference and using the area ratio of the correct region in the distribution of the log likelihood. The present method will predict the distribution of the word pair log likelihood difference using the following method (**Fig. 1**):

(1) Obtain the distribution of the phoneme log likelihood difference in advance for the combination of all the phonemes.

(2) Decompose words into phonemes. Obtain the distribution of the word log likelihood difference by summing the log likelihood difference distribution of mutually different phonemes in the word.

3. Word pair recognition test and results

We conducted a prediction test of a word pair recognition rate by using the distribution of the phoneme log likelihood difference obtained from the experimental conditions given in **Table 1**. **Fig. 2** shows the results of the test with 350 word pairs extracted at random from the NTT word database. The correlation coefficient between the predicted and the measured recognition rates is 0.87, and it is understood that the present method can predict the word pair recognition rate.

4. Conclusion

We have discussed a method for predicting the word pair recognition performance on the basis of the distribution of the log likelihood difference. We have verified that the present method is applicable to detecting words of low recognition rate. In the future, we intend to study the method for further improving the prediction accuracy.

Reference

1) Tanaka, K., et al. : "Estimation of a Degree of Speech Recognition Difficulty for Word Sets", J. Acoust. Soc. Jpn(E), 19-5(1998), 339

(Report received on Feb. 1, 2001)

Table 1 Experimental condition.

Decoder	HVite(HTK)
Acoustic Model	monophone, 4mixture/state(IPA'97)
Utterance	NTT Japanese Words Database (Command set and Japanese city names)
Acoustic Analysis	Sampled at 16kHz with 16bit 12_{th} order MFCC + ΔMFCC + ΔLog Pow

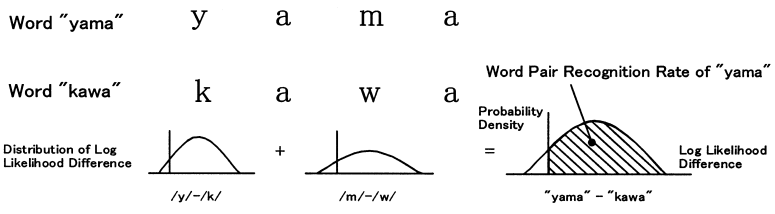

Fig. 1 Prediction of word pair recognition rate.

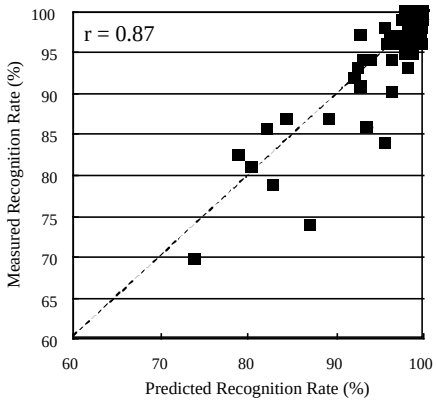

Fig. 2 Prediction result.

Copyright (C) 2001 Toyota Central R&D Labs., Inc.