Academic Year of 2018, Submitted on February 1st, 2019

Graduate School of Science and Engineering, Aoyama Gakuin University

Title: Basic Study about Guidance Timing Optimization of On-Vehicle Spoken Dialogue System

Student Name: Ko UCHIDA
ID Number: 35617132
Degree: Master of Engineering
Course: Intelligence and Information
Thesis Advisor: Guillaume LOPEZ

Abstract

In recent years, spoken dialogue system has been put to practical use in various ways thanks to voice recognition technologies. These systems are particularly useful for car navigation systems because they are believed to reduce hazards due to distracted driving. On the other hand, it is necessary taking a risk about driving while talking into consideration. For this reason, it is better to suspend the system when the driver feel stress. Nevertheless, little is known about drive scenes that driver feel stress.

In this research, to optimize guidance timing of on-vehicle spoken dialogue system, we identify drive scenes driver can't talk. Besides this, we classify these scenes and dialogue situations using driving signals and biological signals.

As examples of dialogue situations, only driving and driving while speaking are mentioned. Even though under the same drive scene, it is assumed that the latter takes more load than the former. Based on this, we classify dialogue situations too. In the experiment, subjects drive a course that consists of scenes with different loads using driving simulator (DS). We measured driving log, change of oxygenated hemoglobin concentration on brain blood flow, and CVRR as driving signal and biological signal. To measure the change of oxygenated hemoglobin concentration, we used a Near-infrared spectroscopy system (NIRS). CVRR is a kind of heart rate interval index.

Based on the number of utterance and the utterance ratio of the transit time of each drive scenes, we found that drive scenes about associated with lane change make the driver feel stress. Thus, we regard these scenes as stressed scenes. In addition, we found that drive scenes that change road width make the driver feel a little stress, and regard these scenes as semi-stressed scenes.

From the result of machine learning using Support vector machine (SVM), we confirmed that NIRS data can classify the stress scenes, semi-stressed scenes, and the other scenes with more 90% accuracy and DS data can classify these with more 80% accuracy. However CVRR data can't classify these, the combination of CVRR data and other data can classify these with high accuracy. In addition, we confirmed that NIRS data can classify the conditions of the driver (only driving/driving while taking) with 95% accuracy and DS data can classify these with more 75%.

In the future, it is necessary to develop the spoken dialogue system using these findings and investigation the convenience of the system.