A Study of Effectiveness of Speech Enhancement for Cognitive Load Classification in Noisy Conditions

Phu Ngoc Le[1,2,3], Eliathamby Ambikairajah[1,2], Tharmarajah Thiruvaran[1,4], Tuan Thanh Nguyen[3]

[1] School of Electrical Engineering and Telecommunications,
The University of New South Wales, UNSW Sydney, NSW 2052, Australia
[2] ATP Research Laboratory, National ICT Australia (NICTA), Eveleigh 2015, Australia
[3]Department of Electrical and Electronic Engineering, Ho Chi Minh City University of Technology, Vietnam
[4] Department of Electrical and Electronic Engineering, University of Jaffna, Sri Lanka

E-mail: phule@unsw.edu.au, ambi@unsw.edu.au, thiruvaran@jfn.ac.lk, nttuan@hcmut.edu.vn

***Abstract*— In the last decade, speech-features have been effectively utilized for estimating cognitive load level in ideal conditions where recorded speech is clean. However, in more realistic conditions, the recorded speech data is corrupted by noise. Hence, the employment of speech enhancement is essential to reduce the noise. In this paper, the effectiveness of three speech enhancement algorithms proposed in our previous studies are compared based on performance and processing time and the most suitable method is utilized to denoise the input noisy speech before feeding it to a cognitive load classification system in order to improve its performance. The results of this study indicate that the use of speech enhancement can reduce 3.0% of average relative error rate for the system under the effect of various noisy conditions.**

Keywords—cognitive load; noisy conditions; speech enhancement; discrete cosine transform.

I. INTRODUCTION

Cognitive load is defined as the amount mental workload imposed on cognitive system of users while they are performing a task. Since human working memory is limited, their task performance will decrease if they are overloaded or underloaded. It is therefore essential to measure their cognitive load (CL) level and accordingly tailor the amount of workload assigning on them [1]. Among the methods have been proposed for CL level measurement, speech-based method has been recognized as an effective one due to the non-intrusive and inexpensive properties of speech signal [2-3].

In most of the previous studies in this research area, speech was recorded in laboratory environments where noise does not exist and hence speech is clean [2-3]. However, this is not the case for practical scenarios where speech is recorded at working environments e.g. office, airport, inside vehicles, etc. where the contamination of noise to input speech of the system is inevitable. It has been shown that noise can cause a mismatch between the feature distributions of speech in training and the testing phases which consequently decreases the classification accuracy and make the system less applicable for industrial implementation. Therefore it is essential to develop techniques that can reduce the contamination of noise for cognitive load classification. One possible scheme to address this issue is to use speech enhancement techniques to preprocess noisy speech before feeding it into the system. It was reported that the incorporation of speech enhancement can yield higher accuracy for a speech recognition system under the corruption of noise [4].

Numerous researchers have been attracted in development speech enhancement algorithm and a large number of techniques have been proposed for speech enhancement including spectral subtraction [5], Wiener filtering [6], statistical-models [7], and Kalman filtering [8]. Among the methods have been developed, Kalman filtering-based methods have been reported to yield less musical tone and low distortion for output speech [8]. Furthermore, empirical mode decomposition (EMD) has been indicated to be effective for denoising speech due to its ability to separate speech and noise into different components [9]. In addition, speech enhancement methods using discrete cosine transform (DCT) have been shown to be efficient due to the energy compaction nature of DCT [10].

However, the demand for speech enhancement is still high that requires the development of even more effective enhancement methods. In this line, three new speech enhancement algorithms were proposed in our previous research, namely non-uniform subband Kalman filtering [11], empirical mode decomposition based speech enhancement method [12], improved soft thresholding on Discrete Cosine Transform [13]. Speech enhancement methods can be very effective in cognitive load classification under noisy conditions. However, the application of cognitive load classification requires not only improved performance but also least processing time. Thus in this paper the above speech enhancement methods are compared not only based on performance but also based on their processing time in order to select a suitable method for cognitive load classification. Then the selected method is applied for improving the robustness of a cognitive load classification system in noisy conditions.

978-1-4673-8374-5/15/$31.00 ©2015 IEEE