The 1st JAIST World Conference (JWC2018)
at JAIST, February 27-28, 2018

Reverberant speech enhancement based on phase and reverberation aware DNN

Longbiao Wang[1,*], Zeyan Oo[2], Seiichi Nakagawa[3], Jianwu Dang[1,4], Masahiro Iwahashi[2]

[1]*Tianjin Key Lab. of Cognitive Computing and Application, Tianjin University, Tianjin, China*
[2] *Nagaoka University of Technology, Nagaoka, Niigata, Japan*
[3] *Toyohashi University of Technology, Toyohashi, Aichi, Japan*
[4] *Japan Advanced Institute of Science and Technology (JAIST), Ishikawa, Japan*
*e-mail:longbiao_wang@tju.edu.cn

Keywords: Speech enhancement, Deep neural network, Phase information, Reverberant environment

Reverberation can greatly decay the performance of many acoustical systems. With the consideration of improving the Deep Neural Network (DNN) performance we propose a method by adding online reverberation information of the reverberant speech signal. Moreover, in many speech enhancement systems, the phase information has been ignored because of its complicated structure. As phase correlates closely to magnitude signal we exploited this relationship to achieve better performance using DNN. Then, phase information is augmented with magnitude information and used as the input for DNN. We denote this method as phase aware DNN. Finally, both phase information and reverberation were added to reverberant speech to achieve better speech enhancement performance, which is denoted as reverberation and phase aware DNN. The proposed method was evaluated using the REVERB (Reverberant Voice Enhancement and Recognition Benchmark) challenge 2014 database. Results were significantly improved results with respect to both reconstructed speech quality (PESQ: Perceptual Evaluation of Speech Quality) and influence of reverberation (SRMR: Speech to Reverberation Modulation Energy Ratio). As compared to the conventional DNN-based approach, this proposed reverberation and phase aware DNN improved PESQ from 2.34 to 2.70 and SRMR from 4.84 to 5.92, indicating that our proposed method could efficiently enhance reverberant speech under various conditions.

Acknowledgement The research is supported partially by the National Natural Science Foundation of China (No. 61771333 and No. U1736219) and JSPS KAKENHI Grant (No. 16K12461 and No. 16K00297).

References

1) R. M. Hegde et al., “Significance of the Modified Group Delay Feature in Speech Recognition,” IEEE Trans. on Audio, Speech, and Language Processing, Vol. 15, No.1, pp. 190-202, 2007.

2) X. Xiao et.al, “Speech dereverberation for enhancement and recognition using dynamic features constrained deep neural networks and feature adaptation,” EURASIP Journal of Advances in Signal Processing, 2016.