王緒翔

中央研究院 資訊科技創新研究中心 博士

https://www.citi.sinica.edu.tw/people/postdoctoral-fellows

Dr. Syu Siang Wang received the Ph.D. degree (2018) in the Graduate Institute of Communication Engineering, National Taiwan University. The topic of his Ph.D. research is on wavelet speech enhancement and feature compression. He won the PhD Thesis Award at ACLCLP. In addition, he gained twice opportunities to be an summer intern in National Institute of Information and Communications Technology, Japan, in Sep. 2015 and Department of Electrical and Electronic Engineering, SUSTC, China in Jun. 2016.

From August 2018 to July 2019, he was the postdoctoral researcher in MOST Joint Research Center for AI Technology and All Vista Healthcare, where he engaged in research on developing algorithm for healthcare applications . Several papers were published based on his research achievements.

Currently, he is the postdoctoral researcher in the Research Center for Information Technology Innovation, Academia Sinica. His research interests include speech and speaker recognition, acoustic modeling, audio-coding, and bio-signal processing.

Single- and Multi-channel Speech Enhancement System

Dr. Syu Siang Wang

Real-world environments are always contain stationary and/or time-varying noises that are received together with speech signals by recording devices. The received noises inevitably degrade the performance of human--human and human--machine interfaces, and this issue has attracted significant attention over the years. To address this issue, an important front-end speech process, namely speech enhancement, which extracts clean components from noisy input, can improve the voice quality and intelligibility of noise-deteriorated clean speech. These speech-enhancement systems can be split into two categories in terms of the physical configurations: single- and multi-channels. For single-channel-based speech enhancement systems, the speech waveform was recorded essentially from an microphone, and then enhanced through the enhancement system, which is derived based on the temporal information of the input. Multiple microphones are used to record the input speech in a multi-channel-based speech enhancement system. The system is designed by simultaneously exploiting the spatial diversity and temporal structures of received signals. In this talk, we present our recent research achievements using machine learning and signal processing on improving speech perception abilities for both configurations.