

SEMINAR ANNOUNCEMENT

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
COLLEGE OF DESIGN AND ENGINEERING
Website: <https://cde.nus.edu.sg/ece>

Area: Signal Analysis & Machine Intelligence

Host: Assistant Prof Mike Shou Zheng

TOPIC	:	Audio-Visual Speaker Diarization And Social Analysis
SPEAKER	:	Mr. Xu Zhongcong Graduate Student, ECE Dept, NUS
DATE	:	Friday, 22 July 2022
TIME	:	4.00PM to 4.30PM
WEBINAR	:	Join Zoom Meeting https://nus-sg.zoom.us/j/81507059776?pwd=Uk9NbJBEWXNFOEhkdGRqR3cvNWILUT09 Meeting ID: 815 0705 9776 Passcode: 067223

ABSTRACT

Audio-visual speaker diarization aims at detecting “who spoke when” using both auditory and visual signals. Existing audio-visual diarization datasets are mainly focused on indoor environments like meeting rooms or news studios, which are quite different from in-the-wild videos in many scenarios such as movies, documentaries, and audience sitcoms. To develop diarization methods for these videos, we create the AVA Audio-Visual Diarization (AVD) dataset, which extends the AVA movie dataset with speaker diarization annotations. We also design an audio-visual relation network which introduces a simple yet effective modality mask to capture discriminative information based on face visibility. Experiments show that our method not only can outperform state-of-the-art methods but is more robust as varying the ratio of off-screen speakers. In addition to third person view, we also collect a new egocentric dataset (EGO4D) for audio-visual diarization and social analysis, experiments on baseline frameworks shows challenging nature of the egocentric perspective.

BIOGRAPHY

Mr. Zhongcong Xu is a PhD student at Showlab, ECE Dept, NUS, supervised by Prof. Mike Shou Zheng. His research interests are multi-modal discriminative and generative models for human faces.

<https://cde.nus.edu.sg/ece/highlights/events/>