Soo-Whan Chung

RESEARCH SCIENTIST – CLOVA SPEECH, NAVER CLOUD

NAVER 1784, 95, Jeongjail-ro, Bundang-gu, Seongnam-si, Gyeonggi Province, South Korea

✉ soowhan.chung@gmail.com | 🏠 soowhanchung.github.io | 💼 soowhan-chung | 💻 kevin6293 | 🎓 Google Scholar

"We will do something wonderful!"

Education

Doctor of Philosophy   *Mar. 2016 - Feb. 2021*

DEPARTMENT OF ELECTRICAL & ELECTRONIC ENGINEERING, YONSEI UNIVERSITY   *Seoul, S.Korea*

DIGITAL SIGNAL PROCESSING & ARTIFICIAL INTELLIGENCE (DSP&AI) LAB. (Supervisor. HONG-GOO KANG)

Bachelor of Engineering   *Mar. 2012 - Feb. 2016*

DEPARTMENT OF ELECTRICAL & ELECTRONIC ENGINEERING, YONSEI UNIVERSITY   *Seoul, S.Korea*

FULL SCHOLARSHIP – THE NATIONAL SCHOLARSHIP FOR SCIENCE AND ENGINEERING (KOSAF)

Interests

Speech signal processing

SPEECH ENHANCEMENT/SEPARATION/RESTORATION/GENERATION, SPEAKER RECOGNITION, SPEECH RECOGNITION, KEYWORD SPOTTING

Multi-modal Learning

AUDIO-VISUAL, AUDIO-TEXT, CROSS-MODAL LEARNING

Career

NAVER Cloud   *S. Korea*

RESEARCH SCIENTIST @ SPEECH (PREV. NAVER CORP.)   *Nov. 2020 - Present*

- Machine / Deep Learning
- Speech Signal Processing
- Multi-modal Learning

NVIDIA   *USA*

DEEP LEARNING INSTITUTE CERTIFIED INSTRUCTOR

- Qualify for NVIDIA DLI University Ambassadorship
- Instruct deep learning courses using DLI program toolkit provided by NVIDIA corporation

NAVER   *S. Korea*

INTERNSHIP RESEARCH SCIENTIST   *Jul. 2018 - Dec. 2018*

- Collaborate research projects with NAVER Clova Speech team
- Conduct researches on audio-visual signal processing and speech signal processing

Publication

Dissertation

Self-supervised Cross-modal Representation for Visual-Guided Speech Separation   *2021*

ISNI : 0000 0004 9330 4678 [ENG/KOR]   *Ph.D.*

GRADUATE SCHOOL, YONSEI UNIVERSITY (Supervisor. HONG-GOO KANG)

Journal

Perfect Match: Self-Supervised Embeddings for Cross-modal Retrieval   *2020*

Soo-Whan Chung, JOON SON CHUNG, HONG-GOO KANG

IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, VOL. 14, NO. 3, PP.568–576, MAR., 2020

A study on speech disentanglement framework based on adversarial learning for speaker recognition   *2020*

YOOHWAN KWON, **Soo-Whan Chung**, HONG-GOO KANG

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF KOREA, VOL. 39, NO. 5, PP.447-453, SEP., 2020

Generic uniform search grid generation algorithm for far-field source localization *2018*

JeeSok Lee, **Soo-Whan Chung**, Min-Seok Choi, Hong-Goo Kang

The Journal of the Acoustical Society of America, 143 (1), el37-el42, Jan., 2018

Conference

HD-DEMUCS: General Speech Restoration with Heterogeneous Decoders *2023*

Doyeon Kim, **Soo-Whan Chung**, Hyewon Han, Youna Ji, Hong-Goo Kang

INTERSPEECH

MF-PAM: Accurate Pitch Estimation through Periodicity Analysis and Multi-level Feature Fusion *2023*

Woo-Jin Chung, Doyeon Kim, **Soo-Whan Chung**, Hong-Goo Kang

INTERSPEECH

Imaginary Voice: Face-styled Diffusion Model for Text-to-Speech *2023*

Jiyoung Lee, Joon Son Chung, **Soo-Whan Chung**

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

MoLE : Mixture of Language Experts for Multi-Lingual Automatic Speech Recognition *2023*

Yoohwan Kwon, **Soo-Whan Chung**

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

An Empirical Study on Speech Restoration Guided by Self-supervised Speech Representation *2023*

Jaeuk Byun, Youna Ji, **Soo-Whan Chung**, Soyeon Choe, Min-Seok Choi

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Diffusion-based Generative Speech Source Separation *2023*

Robin Scheibler, Youna Ji, **Soo-Whan Chung**, Jaeuk Byun, Soyeon Choe, Min-Seok Choi

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Learning Audio-Text Agreement for Open-vocabulary Keyword Spotting *2022*

Hyeon-Kyeong Shin, Hyewon Han, Doyeon Kim, **Soo-Whan Chung**, Hong-Goo Kang

INTERSPEECH – 🏆 Best Student Paper Award Finalist

SASV 2022: The First Spoofing-Aware Speaker Verification Challenge *2022*

Jee-weon Jung, Hemlata Tak, Hye-jin Shim, Hee-Soo Heo, Bong-Jin Lee, **Soo-Whan Chung**, Ha-Jin Yu, Nicholas Evans, Tomi Kinnunen

INTERSPEECH

Baseline Systems for the First Spoofing-Aware Speaker Verification Challenge: Score and Embedding Fusion *2022*

Hye-jin Shim, Hemlata Tak, Xuechen Liu, Hee-Soo Heo, Jee-weon Jung, Joon Son Chung, **Soo-Whan Chung**, Ha-Jin Yu, Bong-Jin Lee, Massimiliano Todisco, Héctor Delgado, Kong Aik Lee, Md Sahidullah, Tomi Kinnunen and Nicholas Evans

Odyssey

Phase Continuity: Learning Derivatives of Phase Spectrum from Speech Enhancement *2022*

Doyeon Kim, Hyewon Han, Hyeon-Kyeong Shin, **Soo-Whan Chung**, Hong-Goo Kang

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Looking into Your Speech: Learning Cross-modal Affinity for Audio-visual Speech Separation *2021*

Jiyoung Lee*, **Soo-Whan Chung***, Sunok Kim, Hong-Goo Kang, Kwanghoon Sohn

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)

Look Who's Talking: Active Speaker Detection in the Wild *2021*

You Jin Kim, Hee-Soo Heo, Soyeon Choe, **Soo-Whan Chung**, Yoohwan Kwon, Bong-Jin Lee, Youngki Kwon, Joon Son Chung

INTERSPEECH

End-to-end Lip Synchronisation based on Pattern Classification *2021*

You Jin Kim, Hee-Soo Heo, **Soo-Whan Chung**, Bong-Jin Lee

IEEE Spoken Language Technology Workshop (SLT)

FaceFilter: Audio-visual speech separation using still images *2020*

Soo-Whan Chung, Soyeon Choe, Joon Son Chung, Hong-Goo Kang

INTERSPEECH – 🏆 Best Student Paper Award

Seeing voices and hearing voices: Learning discriminative embeddings using cross-modal self-supervision *2020*

Soo-Whan Chung, Hong-Goo Kang, Joon Son Chung

INTERSPEECH

MIRNet: Learning multiple identities representations in overlapped speech *2020*

Hyewon Han, **Soo-Whan Chung**, Hong-Goo Kang

INTERSPEECH

Intra-class variation reduction of speaker representation in disentanglement framework *2020*

Yoohwan Kwon, **Soo-Whan Chung**, Hong-Goo Kang

INTERSPEECH

Perfect match: Improved Cross-modal Embeddings for Audio-visual Synchronisation *2019*

Soo-Whan Chung, Joon Son Chung, Hong-Goo Kang

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Gradient-based Active Learning Query Strategy for End-to-end Speech Recognition *2019*

Yuan Yang, **Soo-Whan Chung**, Hong-Goo Kang

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

A study on search grid points for data-driven 3-D beamsteering *2017*

JeeSok Lee, **Soo-Whan Chung**, Min-Seok Choi, Hong-Goo Kang

Hands-free Speech Communications and Microphone Arrays (HSCMA)

Domestic Conference

임베딩 매트릭스를 기반으로 한 비정상적 잡음 제거 알고리즘의 분석과 딥러닝 음질개선 방법들과의 성능비교 *2018*

Soyeon Choe, **Soo-Whan Chung**, Hong-Goo Kang

한국음향학회 추계발표대회

비학습 데이터 적응화 기법을 이용한 딥러닝 기반 한국어 음성 인식 기술 *2018*

Yuan Yang, **Soo-Whan Chung**, Hong-Goo Kang

한국음향학회 추계발표대회

음성 인식 기반의 방송미디어 디바이스 제어 및 편집 시스템 구현 *2017*

Sangshin Oh, **Soo-Whan Chung**, Hong-Goo Kang

대한전자공학회 추계학술대회

Research Projects

Naver / Naver Cloud

Keyword Detection

Main role: Scientific Research *Nov. 2020 - Present*

- Implement open-vocabulary keyword detection system
- Design acoustic models with small model size and little latency

Speech Enhancement

Main role: Scientific Research *Apr. 2021 - Present*

- Improve speech quality of recorded speech signals
- Design the speech enhancement model for the noise reduction and the de-reverberation

Yonsei University

Deep learning-based Audio-Visual Speech Separation Algorithm using Multi-modal Information *w. NAVER Corp.*

Main tasks: Audio-visual speech enhancement / Speech separation / Speaker recognition *Jan. 2019 - Oct. 2019*

- Implement an audio-visual speech separation system using face video streams
- Separate target speech using given visual information represented on facial appearances
- Extract speech signals in a recursive approach for the unknown number of speakers

Implementation of Speech Enhancement and Speech Recognition for CHiME Challenge *w. NAVER Corp.*

Main tasks: Speech separation / Speech enhancement / Speech recognition *Dec. 2017 - Jun. 2018*

- Imlement speech recognition system robust on various environments (noise, interference, multi-talker, reverberation)
- Separate overlapped speech signals using deep learning-based methods
- Recognize speech signals based on Kaldi ASR platform

Assessment of Speaker Presentation Skills using Deep learning-based Speech Signal Analysis

w. Graduate School of Education, Yonsei Univ.

MAIN TASKS: SPEECH ANALYSIS / SPEECH ASSESSMENT

Dec. 2017 - Jun. 2018

- Construct audio-visual database including speech presentation, debate and score
- Analyze acoustic characteristics on observed speech signals
- Estimate the assessment score for each speech signal using deep learning framework

Controlling Broadcasting Media Device and Editing System using Personalized Speech/Gesture Recognition

w. Institute for Information and Communications Technology Promotion (IITP)

MAIN TASKS: HUMAN-COMPUTER INTERACTION / WORD-LEVEL SPEECH RECOGNITION / GESTURE RECOGNITION

Apr. 2017 - Dec. 2017

- Design a simple software to edit media content using speech and gesture inputs
- Recognize pre-enrolled speech commands achieved from real-time microphone
- Control the editor using gesture poses robust on environments

Deep learning-based Multi-channel Speech Enhancement Algorithm for High-Quality Speech Recognition

w. Naver Corp.

MAIN TASKS: BEAMFORMING / SPEECH ENHANCEMENT / SOURCE LOCALIZATION

Jul. 2016 - Jun. 2017

- Analyze spatial/spectral/temporal characteristics of signals and indicate the direction of the input target speech
- Implement a deep learning-based beamformer using spatio-temporal features related to the time-delay of arrival
- Reject undesirable noise and reconstruct speech signals on the time-frequency domain

Skills

Programming	C/C++, Python, MATLAB
Deep learning	PyTorch, Tensorflow, MATLAB
Languages	Korean, English