Wang Rui - Research CV

Birth:	30th October 1993
Research:	Spatial audio, Speech signal processing, Speech enhancement/separation, Target speaker extraction, Deep learning
Tools:	Python, Shell, C#, Matlab
Language:	Chinese, Japanese, English
Email:	rui.wang@g.sp.m.is.nagoya-u.ac.jp
Scholar:	https://scholar.google.com/citations?user=N3UBXW8AAAAJ&hl=en&authuser=2
Linkedin:	https://www.linkedin.com/in/rui-wang-aa2b0619b/

Summary

I have several years of research experience in speech signal processing, focusing on spatial hearing and speech enhancement in challenging environments. At JAIST, I worked on monaural 3D sound localization using HRTF features under Prof. Masashi Unoki. For my Ph.D. at Nagoya University with Prof. Tomoki Toda, my main topic was directional target speaker extraction (TSE) in noisy and underdetermined conditions, resulting in publications such as TASLP. **My future goal is to extend statistical signal processing (such as independent/low-rank and spatial covariance modeling) by coupling it with DNN priors and latest LLM-based context, aiming for identifiable, sample-efficient, and real-time/low-latency streaming speech enhancement.**

Work experience

2021.8 - 2021.10 **Summer internship** — *Kyoto, Japan*
National Institute of Information and Communications Technology (NICT), ASTREC
Research on robust speech recognition

2022.3 - 2022.4 **Winter internship** — *Tokyo, Japan*
Nippon Telegraph and Telephone Corporation (NTT), CS lab
Research on robust speech separation

2025.5 - **Research Engineer** — *Shanghai, China*
Midea Group, AI Research Institute
Research on robust multi-task speech interaction system in challenge environments

Education&Research

2012.9 - 2016.6 **BS degree** — *Hangzhou, China*
China Jiliang University
Measurement and Control Technology and Instruments

2016.9 - 2018.8 **Master's course** — *Beijing, China*
National Institute of Metrology, China
Fluid Mechanics (Dropout due to lack of interest)

2018.10 - 2021.3 **Master's degree** — *Ishikawa, Japan*
Japan Advanced Institute of Science and Technology (JAIST)
Akagi & Unoki Laboratory of speech
Computer Science, focus on HRTF-based DOA estimation and spatial hearing

2021.4 - 2025.3 **Doctor's degree** — *Nagoya, Japan*
Nagoya University
Toda Laboratory of speech
Computer Science, focus on target speaker extraction in challenge environments

Publications & Awards

Journal paper 2023	**R. Wang, B. N. Khanh, D. Morikawa, and M. Unoki, "Method of estimating three dimensional direction-of-arrival based on monaural modulation spectrum," Applied Acoustics, 203, 109215, 9 pages, Feb. 2023.**
Journal paper 2024	**R. Wang, L. Li, T. Toda, "Dual-channel target speaker extraction based on conditional variational autoencoder and directional information," IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol. 32, pp. 1968-1979, Mar. 2024.**
Journal paper 2024	**R. Wang, T. Fujimura, and T. Toda, "Target Speaker Extraction under Noisy Underdetermined Conditions Using Conditional Variational Autoencoder, Global Style Token, and Neural Postfilter," APSIPA Transactions on Signal and Information Processing, Vol. 14, No. 1, e2, pp. 1-26, Jan. 2025.**
Conference paper 2021	**N. Li, L. Wang, M. Unoki, S. Li, R. Wang, Me. Ge, J. Dang, "Robust Voice Activity Detection Using a Masked Auditory Encoder Based Convolutional Neural Network," in 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 6828–6832, Jun. 2021.**
Conference paper 2021	**R. Wang, B. N. Khanh, D. Morikawa, and M. Unoki, "Method of Estimating 3D DOA based on Monaural Modulation Spectrum,"In: 2021 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP), pp. 137-140, Mar. 2021.**
Conference paper 2022	**R. Wang, L. Li, and T. Tomoki, "Direction-aware target speaker extraction with a dual-channel system based on conditional variational autoencoders under underdetermined conditions," in Proc. IEEE Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf., 2022, pp. 347–354.**
Conference paper 2023	**R. Wang, T. Toda, "Directional target speaker extraction under noisy underdetermined conditions through conditional variational autoencoder with global style tokens," Proc. IEEE WASPAA, New Paltz, USA, Oct. 2023, pp. 1–5.**
Domestic paper 2021	**R. Wang, B. N. Khanh, D. Morikawa, and M. Unoki, "Method of estimating DOA based on monaural modulation spectrum,"** 日本音学会春季研究表会演文集**, 3-1-21, pp. 321-324, Mar. 2021.**
Domestic paper 2022	**R. Wang, L. Li, T. Toda, "Target speaker extraction based on conditional variational autoencoder and directional information in underdetermined condition", Technical Report of IEICE, Vol. 121, No. 383, EA2021-76, pp. 76-81, Mar. 2022.**
Domestic paper 2022	**R. Wang, Li Li, and T. Toda, "Direction-aware target speaker extraction with conditional variational autoencoders and its sensitivity to direction-of-arrival error, "** 日本音学会春季研究表会演文集**, 2-2-6, pp. 195-196, Sep. 2022.**
Award 2021	**RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP)-Student paper award.**

Award 2021	**Acoustical Society of Japan (ASJ)-Student paper award (Hokuriku branch).**
Award 2022	**Acoustical Society of Japan (ASJ)-Student paper award.**
Award 2023	**IEEE WASPAA 2023 Travel Grants.**
Under Review 2025	**R. Wang et al., "End-to-End Direction-Aware Keyword Spotting with Spatial Priors in Noisy Environments," ICASSP 2026.**