

Li Li

AI Lab, Cyberagent, Inc.,
22F Shibuya Scramble Square, 2-24-12 Shibuya, Shibuya-ku, Tokyo, 150-6121, Japan

☎ +81 50-5530-9270 | ✉ lili-0805@ieee.org | 🌐 <https://lili-0805.github.io/>

RESEARCH INTERESTS

Audio signal processing, source separation, machine learning, speech analysis

EDUCATION

- MAR. 2021 Ph.D. degree in Engineering, Department of Computer Science, Graduate School of Systems and Information Engineering, University of Tsukuba, Japan.
Topic: Determined multichannel source separation based on signal independence
Advisors: Prof. Shoji Makino and Dr. Hirokazu Kameoka
- MAR. 2018 M.S. degree in Engineering, Department of Computer Science, Graduate School of Systems and Information Engineering, University of Tsukuba, Japan.
Topic: Monaural speech enhancement with non-negative matrix factorization
Advisors: Prof. Shoji Makino and Dr. Hirokazu Kameoka
- JULY 2014 B.E. degree in Engineering, Department of Information Management and Engineering, Shanghai University of Finance and Economics, China.

RESEARCH & WORK EXPERIENCES

- JAN. 2023-CURRENT Research scientist at AI Lab, Cyberagent, Inc., Japan.
- APR. 2021-CURRENT Postdoctoral researcher at Toda Lab, Nagoya University, Japan.
- APR. 2021-DEC. 2022 Adjunct researcher at NTT communication Science Laboratories, NTT Corporation, Japan.
- JUN. 2021-DEC. 2022 Researcher at TARVO Inc., Japan.
- APR. 2018-MAR. 2021 Research fellow (DC1) of Japan Society of Promotion of Science (JSPS), Japan.
- APR. 2018-MAR. 2021 Research intern at NTT Communication Science Laboratories, NTT Corporation, Japan.
- JULY 2019-OCT. 2019 Internship at Applied Sciences Lab, Microsoft Corporation, USA.
Topic: Geometrically constrained independent vector analysis for speech enhancement
- JULY 2018-AUG. 2018 Research student at Toda Lab, Nagoya University, Japan.
Topic: Singing voice modeling for singing style control
- OCT. 2014-MAR. 2016 Research student at Kameoka Lab and Saruwatari Lab, University of Tokyo, Japan.

AWARDS

- MAR. 2022 The 51th Awaya Prize Young Researcher Award, The Acoustical Society of Japan (ASJ).
- MAR. 2021 President's Award of University of Tsukuba.
- DEC. 2020 IEEE Signal Processing Society Japan Student Conference Paper Award.
- DEC. 2019 The Best Student Presentation Award, IEICE Electroacoustics Symposium.
- NOV. 2018 IEEE Signal Processing Society Tokyo Joint Chapter Student Award.
- MAR. 2018 Chair Award of the Department of Computer Science, Graduate School of Systems and Information Engineering, University of Tsukuba.
- SEP. 2016 Best Student Presentation Award, The Acoustical Society of Japan (ASJ).

FUNDINGS

- MAY. 2019 Grants for Researchers Attending International Conferences, The Telecommunications Advancement Foundation.
- 2018-2021 Grant-in-Aid for JSPS Fellows, Japan Society of the Promotion of Science (JSPS).
- AUG. 2017 Travel Support for Overseas Dispatch of Graduate Students, University of Tsukuba.
- MAR. 2017 Travel Support for Overseas Dispatch of Graduate Students, University of Tsukuba.

LANGUAGES

Chinese (native), English (fluent), Japanese (fluent)

COMPUTER SKILLS

Programming: Matlab, Python

PUBLICATIONS

JOURNAL PAPERS

- [1] **L. Li**, H. Kameoka, and S. Makino, “FastMVAE2: On improving and accelerating the fast variational autoencoder-based source separation algorithm for determined mixtures,” *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 31, pp. 96–110, Oct. 2022.
- [2] **L. Li**, H. Kameoka, S. Inoue, and S. Makino, “FastMVAE: A fast optimization algorithm for the multichannel variational autoencoder method,” *IEEE Access*, vol. 8, pp. 228740–228753, Dec. 2020.
- [3] **L. Li**, H. Kameoka, and S. Makino, “Majorization-minimization algorithm for discriminative non-negative matrix factorization,” *IEEE Access*, vol. 8, pp. 227399–227408, Dec. 2020.
- [4] M. Qin, **L. Li**, and S. Makino, “Deep complex-valued neural network-based triple-path mask and steering vector estimation for multichannel target speech separation,” *Journal of Signal Processing*, vol. 27, no. 4, pp. 87–91, July, 2023.
- [5] R. Takahashi, **L. Li**, S. Makino, and T. Yamada, “VMInNet: Interpolation of virtual microphones in optimal latent space explored by auto encoder,” *Journal of Signal Processing*, vol. 25, no. 6, pp. 245–250, Nov. 2021.
- [6] N. Murashima, H. Kameoka, **L. Li**, S. Shogo, and S. Makino, “Single-channel multispeaker separation with variational autoencoder spectrogram model,” *Journal of Signal Processing*, vol. 25, no. 4, pp. 145–149, Jul. 2021.
- [7] S. Seki, H. Kameoka, **L. Li**, T. Toda, and K. Takeda, “Underdetermined Source Separation Based on Generalized Multichannel Variational Autoencoder,” *IEEE Access*, vol. 7, No. 1, pp. 168104–168115, Nov. 2019.
- [8] H. Kameoka, **L. Li**, S. Inoue, and S. Makino, “Supervised determined source separation with multichannel variational autoencoder,” *Neural Computation*, vol. 31, no. 9, pp. 1–24, Sep. 2019.
- [9] H. Kameoka, T. Higuchi, M. Tanaka, and **L. Li**, “Non-negative matrix factorization with basis clustering using cepstral distance regularization,” *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 26, no. 6, pp. 1025–1036, Jun. 2018.

SELECTED PEER-REVIEWED CONFERENCE PAPERS

- [1] **L. Li**, H. Kameoka, and S. Seki, “HBP: An efficient block permutation solver using Hungarian algorithm and spectrogram inpainting for multichannel audio source separation,” in *Proc. 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2022)*, pp. 516–520, Singapore, May 2022.
- [2] **L. Li**, K. Koishida, and S. Makino, “Online directional speech enhancement using geometrically constrained independent vector analysis,” in *Proc. The 21th Annual Conference of the International Speech Communication Association (Interspeech2020)*, pp. 61–65, Shanghai, Oct. 2020.

- [3] **L. Li**, H. Kameoka, and S. Makino, “Determined audio source separation with multichannel star generative adversarial network,” in Proc. *The 30th IEEE International Workshop on Machine Learning for Signal Processing (MLSP2020)*, Espoo, Sep. 2020.
- [4] **L. Li** and K. Koishida, “Geometrically constrained independent vector analysis for directional speech enhancement,” in Proc. *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2020)*, pp. 846–850, Barcelona, May 2020.
- [5] **L. Li**, T. Toda, K. Morikawa, K. Kobayashi, and S. Makino, “Improving singing aid system for laryngectomees with statistical voice conversion and VAE-SPACE,” in Proc. *20th International Society for Music Information Retrieval Conference (ISMIR2019)*, pp. 784–790, Delft, Nov. 2019.
- [6] **L. Li**, K. Yamaoka, Y. Koshino, M. Matsumoto, and S. Makino, “Voice activity detection under high levels of noise using gated convolutional neural networks,” in Proc. *International Congress on Acoustics (ICA2019)*, pp. 2862–2869, Aachen, Sep. 2019.
- [7] **L. Li**, H. Kameoka, and S. Makino, “Fast MVAE: Joint separation and classification of mixed sources based on multichannel variational autoencoder with auxiliary classifier,” in Proc. *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2019)*, pp. 546–550, Brighton, May 2019.
- [8] **L. Li** and H. Kameoka, “Deep clustering with gated convolutional networks,” in Proc. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2018)*, pp. 16–20, Calgary, April 2018.
- [9] **L. Li**, H. Kameoka, and S. Makino, “Mel-generalized cepstral regularization for discriminative non-negative matrix factorization,” in Proc. *The 27th IEEE International Workshop on Machine Learning for Signal Processing (MLSP2017)*, Tokyo, Sep. 2017.
- [10] **L. Li**, H. Kameoka, T. Toda, and S. Makino, “Speech enhancement using non-negative spectrogram models with mel-generalized cepstral regularization,” in Proc. *The 18th Annual Conference of the International Speech Communication Association (Interspeech2017)*, pp. 1998–2002, Stockholm, Aug. 2017.
- [11] **L. Li**, H. Kameoka, and S. Makino, “Discriminative non-negative matrix factorization with majorization-minimization,” in Proc. *The 5th Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA2017)*, pp. 141–145, San Francisco, Mar. 2017.
- [12] **L. Li**, H. Kameoka, T. Higuchi, and H. Saruwatari, “Semi-supervised joint enhancement of spectral and cepstral sequences of noisy speech,” in Proc. *The 17th Annual Conference of the International Speech Communication Association (Interspeech2016)*, pp. 3753–3757, San Francisco, Sep. 2016.