Takaaki Saeki

saefrospace@gmail.com | takaaki-saeki.github.io | linkedin.com/in/takaaki-saeki

Research Interests

Speech synthesis: Text-to-speech synthesis, voice conversion
Spoken language processing: Speech recognition, speech translation, multimodal speech processing
Other: Signal processing, natural language processing, machine learning

Education

The University of Tokyo Tokyo, Japan
Ph.D. Degree in Graduate School of Information Science and Technology *Apr. 2021 – Present*
- Supervisor: Prof. Hiroshi Saruwatari

The University of Tokyo Tokyo, Japan
Master's Degree in Graduate School of Information Science and Technology *Apr. 2019 – Mar. 2021*
- Thesis Title: "Real-Time, Full-Band, High-Quality Neural Voice Conversion with Sub-Band Modeling and Data-Driven Phase Estimation"
- Supervisor: Prof. Hiroshi Saruwatari

The University of Tokyo Tokyo, Japan
Bachelor's Degree in Aerospace Engineering *Apr. 2015 – Mar. 2019*
- Supervisor: Prof. Kimiya Komurasaki

Work Experience

The University of Tokyo Tokyo, Japan
Research Assistant *May 2020 – Present*
- Researching on text-to-speech synthesis and voice conversion.
- Advisor: S. Takamichi & H. Saruwatari

LINE Corporation Tokyo, Japan
Part-Time Researcher *Mar. 2021 – Present*
- Designing research proposals and implementing new text-to-speech synthesis methods.
- Advisor: K. Tahchibana & R. Yamamoto

Preferred Networks, Inc. Tokyo, Japan
Research Intern *Aug. 2021 – Sep. 2021*
- Researched on application of machine learning to music creation
- Advisor: D. Higurashi, B. Vogel, and J. Naradowsky

NEC Data Science Research Laboratories Kanagawa, Japan
Research Intern *Aug. 2019 – Sep. 2019*
- Worked on speech enhancement.
- Advisor: O. Hoshuyama

The University of Tokyo Tokyo, Japan
Teaching Assistant *Apr. 2019 – July 2019*
- Lectured how to develop audio-based interaction system in "Project Practice" course.

Recruit Holdings Co., Ltd. Tokyo, Japan
Data Engineering Intern & Part-Time Engineer *Feb. 2019 – June 2019*
- Worked on data analysis and developed recommendation engine.
- Lectured a training course of data analysis for newcomers.
- Advisor: N. Osugi

Delight Inc. Tokyo, Japan
Part-Time Engineeer *May 2017 – Apr. 2019*
- Developed simulation software of fluid dynamics.
- Read research papers and implemented algorithms of thermo-fluid dynamics.

Technical Skills

Languages: Python, C/C++, Shell, LaTeX, SQL (BigQuery), HTML/CSS
Developer Tools: Git, Docker, Amazon Web Service, Google Cloud Platform, VS Code
Libraries: NumPy, pandas, PyTorch, Tensorflow, Chainer

Language skills

Japanese (native) and English (fluent)

Honors and Awards

Telecom System Technology Award for Students | *The Telecommunication Advancement Foundation* *Mar. 2022*
Best Student Presentation Award | *Acoustical Society of Japan* *Oct. 2021*
Research Fellowship for Young Scientists (DC2) (2,400,000 JPY/year) *Apr. 2022 – Mar. 2024*
Tobitate Study Abroad Initiatives (160,000 JPY/month) *July 2021*
TOYOTA/Dwango AI Scholarship (1,200,000 JPY/year) *Apr. 2021 – Mar. 2022*
IEICE ISS Best Student Poster Award | *IEICE Speech Committee 2021* *Mar. 2021*
Award for Excellence (2nd place) | *Natural Language Processing Hackathon at Recruit Co., Ltd.* *Feb. 2019*

Publications (Journal Paper)

1. **Takaaki Saeki**, Yuki Saito, Shinnosuke Takamichi, and Hiroshi Saruwatari,
"Real-time Full-Band Voice Conversion with Sub-Band Modeling and Data-Driven Phase Estimation of Spectral Differentials,"
IEICE Transactions on Information and Systems, vol. E104-D, no.7, pp. 1002–1016, Jul. 2021.

2. **Takaaki Saeki**, Shinnosuke Takamichi, and Hiroshi Saruwatari,
"Incremental Text-to-Speech Synthesis Using Pseudo Lookahead with Large Pretrained Language Model,"
IEEE Signal Processing Letters, vol. 28, pp. 857–861, Apr. 2021.
(The 37th Telecom System Technology Award for Students from TAF)
(To be presented at *ICASSP2022*)

Publications (Peer-Reviewed Conference Paper)

1. **Takaaki Saeki**, Shinnosuke Takamichi, and Hiroshi Saruwatari,
"Low-latency Incremental Text-to-Speech Synthesis with Distilled Context Prediction Network,"
in *Proc. ASRU*, Dec. 2021.

2. **Takaaki Saeki**, Yuki Saito, Shinnosuke Takamichi, and Hiroshi Saruwatari,
"Real-Time, Full-Band, Online DNN-Based Voice Conversion System Using a Single CPU,"
in *Proc. INTERSPEECH*, pp. 1021–1022, Shanghai, China, Oct. 2020.

3. Naoki Kimura, Zixiong Su, and **Takaaki Saeki**,
"End-to-End Deep Learning Speech Recognition Model for Silent Speech Challenge,"
in *Proc. INTERSPEECH*, pp. 1025–1026, Shanghai, China, Oct. 2020..

4. **Takaaki Saeki**, Yuki Saito, Shinnosuke Takamichi, and Hiroshi Saruwatari,
"Lifter Training and Sub-Band Modeling for Computationally Efficient and High-Quality Voice Conversion Using Spectral Differentials,"
in *Proc. ICASSP*, pp. 7784–7788, Barcelona, Spain, May 2020.

Publications (Preprint)

1. **Takaaki Saeki**, Shinnosuke Takamichi, Tomohiko Nakamura, Naoko Tanji, and Hiroshi Saruwatari, "SelfRemaster: Self-Supervised Speech Restoration with Analysis-by-Synthesis Approach Using Channel Modeling,"
in *arXiv preprint arXiv:2203.12937*, Mar. 2022.

2. Yuta Matsunaga, **Takaaki Saeki**, Shinnosuke Takamichi, and Hiroshi Saruwatari, "Personalized Filled-Pause Generation with Group-Wise Prediction Models,"
in *arXiv preprint arXiv:2203.09961*, Mar. 2022.

3. Shinnosuke Takamichi, Ludwig Kürzinger, **Takaaki Saeki**, Sayaka Shiota, and Shinji Watanabe, "JTubeSpeech: Corpus of Japanese Speech Collected from YouTube for Speech Recognition and Speaker Verification,"
in *arXiv preprint arXiv:2112.09323*, Dec. 2021.

4. Tomoki Hayashi, Ryuichi Yamamoto, Takenori Yoshimura, Peter Wu, Jiatong Shi, **Takaaki Saeki**, Yooncheol Ju, Yusuke Yasuda, Shinnosuke Takamichi, and Shinji Watanabe, "ESPnet2-TTS: Extending the Edge of TTS Research,"
in *arXiv preprint arXiv:2110.07840*, Oct. 2021.

Exhibitions

Sainokuni Buisiness Arena 2021 ONLINE
"Research on Stress-Free, Real-Time, and Full-Band Voice Conversion Based on Perceptual Model" *Jan. 2021*

CEATEC 2020 ONLINE
"Research on Stress-Free, Real-Time, and Full-Band Voice Conversion Based on Perceptual Model" *Oct. 2020*

Patents (Selected)

Shinnosuke Takamichi, Yuki Saito, **Takaaki Saeki**, and Hiroshi Saruwatari, "Voice Conversion System, Method, and Program,"
International Patent Application No. PCT/JP2021/4367. (Submitted)

Shinnosuke Takamichi, Yuki Saito, **Takaaki Saeki**, and Hiroshi Saruwatari, "Voice Conversion System, Method, and Program,"
International Patent Application No. PCT/JP2020/031122. (Submitted)

(Last update: 2022/03/28)