

Takaaki Saeki

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

RESEARCH INTERESTS

Spoken language processing: Speech synthesis, voice conversion, speech representation learning

Natural language processing

Deep learning

Signal processing

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

EXPERIENCE

Google Research

Tokyo (Remote), Japan

Student Researcher (Intern)

Apr 2022 – Present

- Researching on spoken language processing at Brain Applied Research Team.
- Advisor: H. Zen

LINE Corporation

Tokyo (Remote), Japan

Part-Time Researcher

Mar. 2021 – Mar. 2022

- Worked on text-to-speech synthesis at Voice Team.
- Advisor: K. Tahchibana and R. Yamamoto

Preferred Networks, Inc.

Tokyo (Remote), Japan

Research Intern

Aug. 2021 – Sep. 2021

- Researched on application of deep learning to music creation at Music Team.
- Advisor: D. Higurashi and B. Vogel

NEC Data Science Research Laboratories

Kanagawa, Japan

Research Intern

Aug. 2019 – Sep. 2019

- Worked on speech enhancement.
- Advisor: O. Hoshuyama

Recruit Holdings Co., Ltd.

Tokyo, Japan

Data Engineering Intern & Part-Time Engineer

Feb. 2019 – June 2019

- Worked on data analysis and developed recommendation engine.
- Advisor: N. Osugi

Delight Inc.

Tokyo, Japan

Part-Time Engineer

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

Best Paper Award | *The Institute of Electronics, Information and Communication Engineers (IEICE)* Apr. 2022

Ranked 1st Place in 10/16 Metrics (22 Teams in Total) | *VoiceMOS Challenge 2022* Mar. 2022

Telecom System Technology Award for Students | *The Telecommunication Advancement Foundation* Mar. 2022

Best Student Presentation Award | *Acoustical Society of Japan* Oct. 2021

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

GRANTS AND SCHOLARSHIPS

Research Fellowship for Young Scientists (DC2) (2,400,000 JPY/year) Apr. 2022 – Mar. 2024

UTokyo-TOYOTA Study Abroad Scholarship in AI Field (220,000 JPY/month) Mar. 2022

Tobitate Study Abroad Initiatives (160,000 JPY/month) July 2021

TOYOTA/Dwango AI Scholarship (1,200,000 JPY/year) Apr. 2021 – Mar. 2022

PUBLICATIONS (JOURNAL PAPER)

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.

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)

Yuta Matsunaga, **Takaaki Saeki**, Shinnosuke Takamichi, and Hiroshi Saruwatari,
“Personalized Filled-Pause Generation with Group-Wise Prediction Models,”
in *Proc. LREC*, June 2022. (Accepted)

Naoki Kimura, Zixiong Su, **Takaaki Saeki**, and Jun Rekimoto,
“SSR7000: A Synchronized Corpus of Ultrasound Tongue Imaging for End-to-End Silent Speech Recognition,”
in *Proc. LREC*, June 2022. (Accepted)

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

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.

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..

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)

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.

Takaaki Saeki, Kentaro Tachibana, Ryuichi Yamamoto,
“DRSpeech: Degradation-Robust Text-to-Speech Synthesis with Frame-Level and Utterance-Level Acoustic
Representation Learning,”
in *arXiv preprint arXiv:2203.15683*, Mar. 2022.

Takaaki Saeki*, Detai Xin*, Wataru Nakata*, Tomoki Koriyama, Shinnosuke Takamichi, and Hiroshi
Saruwatari (* Equal Contribution),
“UTMOS: UTokyo-SaruLab System for VoiceMOS Challenge 2022,”
in *arXiv preprint, 2204.02152*, Mar. 2022.

Yoshifumi Nakano, **Takaaki Saeki**, Shinnosuke Takamichi, Katsuhito Sudoh, and Hiroshi Saruwatari,
“vTTS: Visual-Text to Speech,”
in *arXiv preprint arXiv:2203.14725*, Mar. 2022.

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.

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.

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)

MISCELLANEOUS WORK

Sainokuni Business 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

(Last update: 2022/04/26)