

Ryuichi Yamamoto

Voice Team
LINE Corporation
Yotsuya Office/Yotsuya Tower 23rd FL.
1-6-1 Yotsuya, Shinjuku-ku, Tokyo

Last updated: January, 2022
Publications: [Google Scholar](#)
E-mail: zryuichi@gmail.com
Website: r9y9.github.io

EDUCATION

2011 – 2013 **M.Eng**, Graduate School of Engineering
Nagoya Institute of Technology, Nagoya, Japan
Supervisor: Prof. Tadashi Kitamura

2007 – 2011 **B.Eng**, Department of Computer Science
Nagoya Institute of Technology, Nagoya, Japan
Supervisor: Prof. Tadashi Kitamura

PROFESSIONAL EXPERIENCE

2021 – present **Senior Research Scientist**
Voice Team
LINE Corporation, Tokyo, Japan

2018 – 2020 **Research Engineer**
Voice Team
LINE Corporation, Tokyo, Japan

2018 – 2019 **Research Engineer**
Clova Voice
NAVER Corporation, Seongnam, Gyeonggi-do, Korea

2013 – 2017 **Software Engineer**
Computer Vision Team
teamLab Inc., Tokyo, Japan

RESEARCH AREAS

- Statistical Speech Synthesis, Machine Learning, Voice Conversion
- Music Signal Processing, Music Information Retrieval

PROGRAMMING SKILLS

- Experienced in Linux/Windows programming based on C/C++, Python, Bash, Emacs, Git.
- Experienced in speech processing toolkit (SPTK, HTK, HTS, Merlin, ESPnet)
- Experienced in deep learning framework (PyTorch, Keras)

LANGUAGES

| | |
|----------|--------------|
| Japanese | Native |
| English | Intermediate |

MEMBERSHIPS


- The Institute of Electrical and Electronics Engineers, Inc. (IEEE), Member
- The Acoustical Society of Japan (ASJ), Member

AWARDS

| | |
|------|--|
| 2021 | IEEE Signal Processing Society (SPS) Japan Young Author Best Paper Award |
| 2013 | Best Presentation Award in the Acoustic Society of Japan (ASJ) |
| 2012 | Best Presentation Award in the Acoustic Society of Japan (ASJ), Tokai |

PUBLICATIONS

BOOKS

- 2021 **Ryuichi Yamamoto**, Shinnosuke Takamichi, “Text-to-speech with Python,” Impress.
- Website: book.impress.co.jp/books/1120101073
 - Code:  [r9y9/ttslearn](https://github.com/r9y9/ttslearn)

JOURNALS

- 2013 Eita Nakamura, Haruto Takeda, **Ryuichi Yamamoto**, Yasuyuki Saito, Shinji Sako, Shigeki Sagayama, “Score Following Handling Performances with Arbitrary Repeats and Skips and Automatic Accompaniment,” *Journal of Information Processing Society of Japan*, Vol. 54, No. 4, pp. 1338-1349, 2013.

CONFERENCE PROCEEDINGS (PEER-REVIEWED)

- 2021 Min-Jae Hwang, **Ryuichi Yamamoto**, Eunwoo Song, Jae-Min Kim, “High-Fidelity Parallel WaveGAN with Multi-Band Harmonic-Plus-Noise Model,” *Proc. Interspeech*, pp. 2227–2231, 2021.
- Website: min-jae.github.io/interspeech2021/
- Kosuke Futamata, Byeongseon Park, **Ryuichi Yamamoto**, Kentaro Tachibana, “Phrase Break Prediction with Bidirectional Encoder Representations in Japanese Text-to-Speech Synthesis,” *Proc. Interspeech*, pp. 3126–3130, 2021.
- Website: matasuke.github.io/demos/pbp_bert
- Ryuichi Yamamoto**, Eunwoo Song, Min-Jae Hwang, Jae-Min Kim “Parallel Waveform Synthesis Based on Generative Adversarial Networks with Voicing-Aware Conditional Discriminators,” *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6039–6043, 2021.
- Website: r9y9.github.io/demos/projects/icassp2021/

Min-Jae Hwang, **Ryuichi Yamamoto**, Eunwoo Song, Jae-Min Kim, “TTS-by-TTS: TTS-Driven Data Augmentation for Fast and High-Quality Speech Synthesis,” *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6598–6602, 2021.

- Website: min-jae.github.io/icassp2021/

Eunwoo Song, **Ryuichi Yamamoto**, Min-Jae Hwang, Jin-Seob Kim, Ohsung Kwon, Jae-Min Kim, “Improved Parallel Wavegan Vocoder with Perceptually Weighted Spectrogram Loss,” *Proc. Spoken Language Technology Workshop (SLT)*, pp. 470–476, 2021.

- Website: sewplay.github.io/demos/wavegan-pwsl/

2020

Eunwoo Song, Min-Jae Hwang, **Ryuichi Yamamoto**, Jin-Seob Kim, Ohsung Kwon, Jae-Min Kim, “Neural Text-to-Speech with a Modeling-by-Generation Excitation Vocoder,” *Proc. Interspeech*, pp. 3570–3574, 2020.

- Website: sewplay.github.io/demos/mbg_excitnet/

Katsuki Inoue, Sunao Hara, Masanobu Abe, Tomoki Hayashi, **Ryuichi Yamamoto**, Shinji Watanabe, “Semi-Supervised Speaker Adaptation for End-to-End Speech Synthesis with Pretrained Models,” *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7634–7638 2020.

Min-Jae Hwang, Eunwoo Song, **Ryuichi Yamamoto**, Frank Soong, Hong-Goo Kang, “Improving LPCNET-Based Text-to-Speech with Linear Prediction-Structured Mixture Density Network,” *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7219–7223, 2020.

Ryuichi Yamamoto, Eunwoo Song, Jae-Min Kim, “Parallel WaveGAN: A Fast Waveform Generation Model Based on Generative Adversarial Networks with Multi-Resolution Spectrogram,” *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6199–6203, 2020.

- Website: r9y9.github.io/demos/projects/icassp2020/

Tomoki Hayashi, **Ryuichi Yamamoto**, Katsuki Inoue, Takenori Yoshimura, Shinji Watanabe, Tomoki Toda, Kazuya Takeda, Yu Zhang, Xu Tan, “ESPnet-TTS: Unified, Reproducible, and Integratable Open Source End-to-End Text-to-Speech Toolkit,” *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7654–7658, 2020.

- Website: espnet.github.io/icassp2020-tts/

2019

Ryuichi Yamamoto, Eunwoo Song, Jae-Min Kim, “Probability Density Distillation with Generative Adversarial Networks for High-Quality Parallel Waveform Generation,” *Proc. Interspeech*, pp. 699–703, 2019.

- Website: r9y9.github.io/demos/projects/interspeech2019/

Shigeki Karita, Nanxin Chen, Tomoki Hayashi, Takaaki Hori, Hirofumi Inaguma, Ziyang Jiang, Masao Someki, Nelson Enrique Yalta Soplin, **Ryuichi Yamamoto**, Xiaofei Wang, Shinji Watanabe, Takenori Yoshimura, Wangyou Zhang, “A Comparative Study on Transformer vs RNN in Speech Applications,” *Proc. Automatic Speech Recognition and Understanding Workshop (ASRU)*, pp. 449–456, 2019.

- 2014 Shinji Sako, **Ryuichi Yamamoto**, Tadashi Kitamura, “Ryry: A Real-Time Score-Following Automatic Accompaniment Playback System Capable of Real Performances with Errors, Repeats and Jumps,” *Proc. International Conference on Active Media Technology (ICAMT)*, pp. 134–145, 2014.
- 2013 **Ryuichi Yamamoto**, Shinji Sako, Tadashi Kitamura, “Robust On-line Algorithm For Real-time Audio-to-score Alignment Based on A Delayed Decision and Anticipation Framework,” *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 191–195, 2013.
- Ryuichi Yamamoto**, Shinji Sako, Tadashi Kitamura, “Accurate and Low Computational Audio-to-score Alignment Using Segmental CRF with An Explicit Continuous Tempo Model,” *Proc. of Communications and Signal Processing (NCSP)*, pp. 345–348, 2013.







CONFERENCE PROCEEDINGS (NON PEER-REVIEWED)


- 2021 Ryo Terashima, **Ryuichi Yamamoto**, Kentaro Tachibana, “An Investigation of Data Augmentation Using CycleGAN Voice Conversion for Text-to-Speech Synthesis,” *The Acoustic Society of Japan (ASJ)*, pp. xxx–xxx, 2021 (in Japanese).
 • Website: ryojerky.github.io/demo/
- 2013 **Ryuichi Yamamoto**, Shinji Sako, Tadashi Kitamura, “Ryry: Automatic Accompaniment System Capable of Polyphonic Instruments,” *Proc. Interaction*, 2013 (in Japanese).
- Ryuichi Yamamoto**, Shinji Sako, Tadashi Kitamura, “Score Following Based on a Combined Model of Score Position and Tempo and Application to Audio-based Automatic Accompaniment,” *The Acoustic Society of Japan (ASJ)*, pp. 1065–1066, 2013 (in Japanese).
- 2012 **Ryuichi Yamamoto**, Shinji Sako, Tadashi Kitamura, “Real-time Audio to Score Alignment Using Segmental Conditional Random Fields and Linear Dynamical System,” *Proc of The Music Information Retrieval Evaluation eXchange (MIREX)*, 2012.
- Ryuichi Yamamoto**, Shinji Sako, Tadashi Kitamura, “Audio to Score Alignment Using Semi-Markov Conditional Random Fields,” *The Acoustic Society of Japan (ASJ)*, pp. 935–936, 2012 (in Japanese).
- Ryuichi Yamamoto**, Eita Nakamura, Yasuyuki Saito, Shinji Sako, Shigeki Sagayama, “Eurydice: Automatic Accompaniment System with Jumping Capability,” *Proc. Information Processing Society of Japan (IPSJ)*, MUS-96(18), pp. 1–10, 2012 (in Japanese).
- Ryuichi Yamamoto**, Shinji Sako, Tadashi Kitamura, “Real-time Audio to Score Alignment using Hidden Semi-Markov Model and Linear Dynamical System,” *Proc. Information Processing Society of Japan (IPSJ)*, MUS-96(13), pp. 1–6, 2012 (in Japanese).
- Eita Nakamura, **Ryuichi Yamamoto**, Shinji Sako, Yasuyuki Saito, Shigeki Sagayama, “Modeling ornaments in polyphonic MIDI score following and its application to automatic accompaniment,” *Proc. The Acoustic Society of Japan (ASJ)*, pp. 929–930, 2012 (in Japanese).
- Eita Nakamura, **Ryuichi Yamamoto**, Shinji Sako, Yasuyuki Saito, Shigeki Sagayama, “Modeling Performance Indeterminacies for Polyphonic Midi Score Following and Its Application to Automatic Accompaniment,” *Proc. Information Processing Society of Japan (IPSJ)*, MUS-96(14), pp. 1–6, 2012 (in Japanese).

- 2011 **Ryuichi Yamamoto**, Shinji Sako, Tadashi Kitamura, “Cooperative Automatic Accompaniment System Using Predictive Models of Expression in Music Performance Based on CRFs,” *Proc. Information Processing Society of Japan (IPSJ)*, MUS-91(11), pp. 1–6, 2011 (in Japanese).


SELECTED SOFTWARE

LIBRARIES

- 2021 – present **ttslearn**
Library for the book “Text-to-speech with Python”
- Role: Creator and core developer
 - Code:  [r9y9/ttslearn](https://github.com/r9y9/ttslearn)
 - Website: r9y9.github.io/ttslearn/
- 2020 – present **nnsvs**
Neural network-based singing voice synthesis library for research
- Role: Creator and core developer
 - Code:  [r9y9/nnsvs](https://github.com/r9y9/nnsvs)
- 2017 – present **nnmnkwii**
Library to build speech synthesis systems designed for easy and fast prototyping
- Role: Creator and core developer
 - Code:  [r9y9/nnmnkwii](https://github.com/r9y9/nnmnkwii)
 - Website: r9y9.github.io/nnmnkwii/latest/
- 2015 – present **pysptk**
A python wrapper for Speech Signal Processing Toolkit (SPTK).
- Role: Creator and core developer
 - Code:  [r9y9/pysptk](https://github.com/r9y9/pysptk)
 - Website: pysptk.readthedocs.io/
- 2015 – present **pyworld**
A Python wrapper for the high-quality vocoder “World”
- Role: Core contributor and maintainer
 - Code:  [JeremyCCHsu/Python-Wrapper-for-World-Vocoder](https://github.com/JeremyCCHsu/Python-Wrapper-for-World-Vocoder)
- 2014 – 2020 **WORLD.jl**
A lightweight Julia wrapper for WORLD - a high-quality speech analysis, modification and synthesis system
- Role: Creator and core developer
 - Code:  [r9y9/WORLD.jl](https://github.com/r9y9/WORLD.jl)
 - Website: r9y9.github.io/world.jl/latest/
- 2014 – 2017 **MelGeneralizedCepstrums.jl**
Mel-Generalized Cepstrum analysis
- Role: Creator and core developer
 - Code:  [r9y9/MelGeneralizedCepstrums.jl](https://github.com/r9y9/MelGeneralizedCepstrums.jl)

- 2015 – 2019 **librosa**
 Python library for audio and music analysis.
- Role: Contributor
 - Code:  [librosa/librosa](https://github.com/librosa/librosa)
 - Website: librosa.org/



RESEARCH PROJECTS

- 2019 – present **ESPnet**
 End-to-End Speech Processing Toolkit
- Role: Discussions and reviews for text-to-speech features
 - Code:  [espnet/espnet](https://github.com/espnet/espnet)
 - Website: espnet.github.io/espnet/
- 2017 – 2021 **wavenet__vocoder**
 WaveNet vocoder: neural network based waveform generation models
- Role: Creator and core developer
 - Code:  [r9y9/wavenet__vocoder](https://github.com/r9y9/wavenet__vocoder)
 - Website: r9y9.github.io/wavenet__vocoder/
- 2017 – 2020 **deepvoice3__pytorch**
 PyTorch implementation of convolutional neural networks-based text-to-speech synthesis models
- Role: Creator and core developer
 - Code:  [r9y9/deepvoice3__pytorch](https://github.com/r9y9/deepvoice3__pytorch)
 - Website: r9y9.github.io/deepvoice3__pytorch/
- 2017 – 2020 **gantts**
 PyTorch implementation of GAN-based text-to-speech synthesis and voice conversion
- Role: Creator and core developer
 - Code:  [r9y9/gantts](https://github.com/r9y9/gantts)
- 2017 – 2019 **tacotron__pytorch**
 PyTorch implementation of Tacotron speech synthesis model
- Role: Creator and core developer
 - Code:  [r9y9/tacotron__pytorch](https://github.com/r9y9/tacotron__pytorch)
- 2014 – 2016 **VoiceConversion.jl**
 Statistical Voice Conversion in Julia
- Role: Creator and core developer
 - Code:  [r9y9/VoiceConversion.jl](https://github.com/r9y9/VoiceConversion.jl)
- 2014 – 2016 **RobustPCA.jl**
 Robust Principal Component Analysis in Julia
- Role: Creator and core developer
 - Code:  [r9y9/RobustPCA.jl](https://github.com/r9y9/RobustPCA.jl)

SUPERVISOR FOR STUDENTS

- 2021.09 – present Reo Yoneyama (Nagoya University)
- 2021.03 – 2021.09 Takaaki Saeki (The University of Tokyo)

PRESENTATIONS

- 2021 Tomohiro Tanaka, **Ryuichi Yamamoto**, “Report on Participation in Interspeech2021,” SIG Technical Reports, Dec 2021, Online.
- 2020 **Ryuichi Yamamoto**, “Parallel WaveGAN: Fast and High-Quality GPU Text-to-Speech,” Conference on Computer Science for Enterprise (CCSE), Dec 2020, Online.
- Ryuichi Yamamoto**, “Parallel WaveGAN: Fast and High-Quality GPU Text-to-Speech,” Main Session in LINE DEVELOPER DAY, Nov 2020, Online.
- Recording:  youtube.com/watch?v=knzT7M6qsl0
- Togami Masahito, Yusuke Kida, **Ryuichi Yamamoto**, Keisuke Imoto, “Current progress on speech technologies and its future prospects,” Panel Discussion in LINE DEVELOPER DAY, Nov 2020, Online.
- Recording:  youtube.com/watch?v=iSPBCot6n7g
- Tomoki Hayashi, **Ryuichi Yamamoto**, Katsuki Inoue, Takenori Yoshimura, Kazuya Takemura, Tomoki Toda, Shinji Watanabe, “ESPnet-TTS: A toolkit to accelerate research on end-to-end speech synthesis,” Special session of The Acoustic Society of Japan (ASJ), Mar 2020, Online.
- **Invited talk**
 - Website: kan-bayashi.github.io/asj-espnet2-tutorial/
- 2018 **Ryuichi Yamamoto**, “WaveNet: A Generative Model for Raw Audio: What I Learned from Developing An Open-Source Implementation,” Invited Talk in National Institute of Information and Communications Technology (NICT), Feb 2018, Kyoto.
- **Invited talk**
- Ryuichi Yamamoto**, “An Attempt to Reproduce WaveNet-based Text-to-Speech Synthesis,” MACHINE LEARNING Meetup KANSAI, Jun 2018, Kyoto.
- 2016 **Ryuichi Yamamoto**, “The Julia C++ Interface,” JuliaTokyo #6, Sep 2016, Tokyo.
- 2015 **Ryuichi Yamamoto**, “Speech Signal Processing in Julia,” JuliaTokyo #3, Apr 2015, Tokyo.
- 2014 **Ryuichi Yamamoto**, “BinDeps.jl,” JuliaTokyo #2, Sep 2014, Tokyo.