Ryuichi Yamamoto

Speech and Acoustic AI
LY Corporation
Kioi Tower
1-3 Kioicho, Chiyoda-ku, Tokyo, 102-8282, Japan

Last updated: October, 2023
Publications: Google Scholar
E-mail: zryuichi@gmail.com
Website: r9y9.github.io

EDUCATION

2022 – present Ph.D course, Graduate School of Informatics
Nagoya University, Nagoya, Japan
Superviser: Prof. Tomoki Toda

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

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

PROFESSIONAL EXPERIENCE

2023 – present **Senior Research Scientist**
Speech and Acoustic AI
LY Corporation, Tokyo, Japan

2021 – 2023 **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, Singing Voice Synthesis, Singing Voice Conversion
- Music Signal Processsing, Music Information Retrieval

PROGRAMMING SKILS

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

LANGUAGES

Japanese	Native
English	Intermediate

MEMBERSHIPS

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

AWARDS

2022	Student Presentation Award in Graduate School of Informatics, Nagoya University
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 (in Japanese).
- Website: book.impress.co.jp/books/1120101073
- Code: 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 (in Japanese).

CONFERENCE PROCEEDINGS (PEER-REVIEWED)

2023 **Ryuichi Yamamoto**, Reo Yoneyama, Lester Phillip Violeta, Wen-Chin Huang, Tomoki Toda, "A Comparative Study of Voice Conversion Models with Large-Scale Speech and Singing Data: The T13 Systems for the Singing Voice Conversion Challenge 2023," *Proc. Workshop on Automatic Speech Recognition and Understanding (ASRU)*, pp. xxxx–xxxx, 2023.
- Website: anonymous7n.github.io/asru2023/

Robin Scheibler, Takuya Hasumi, Yusuke Fujita, Tatsuya Komatsu, **Ryuichi Yamamoto**, Kentaro Tachibana, "Foley Sound Synthesis with a Class-conditioned Latent Diffusion Model," *Proc. Workshop on Detection and Classification of Acoustic Scenes and Events (DCASE)*, pp. xxxx–xxxx, 2023.

Ryuichi Yamamoto, Reo Yoneyama, Tomoki Toda, "NNSVS: A Neural Network-Based Singing Voice Synthesis Toolkit," *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. xxxx–xxxx, 2023.
- Website: r9y9.github.io/projects/nnsvs/
- Code: nnsvs/nnsvs

Reo Yoneyama, **Ryuichi Yamamoto**, Tomoki Toda, "Non-parallel High-Quality Audio Super Resolution with Domain Adaptation and Resampling CycleGANs," *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. xxxx–xxxx, 2023.
- Website: chomeyama.github.io/DualCycleGAN-Demo/
- Code: chomeyama/DualCycleGAN

Yuma Shirahata, **Ryuichi Yamamoto**, Eunwoo Song, Ryo Terashima, Jae-Min Kim, Kentaro Tachibana, "Period VITS: Variational Inference With Explicit Pitch Modeling For End-to-End Emotional Speech Synthesis," *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. xxxx–xxxx, 2023.
- Website: yshira116.github.io/period_vits_demo/

Masaya Kawamura, Yuma Shirahata, **Ryuichi Yamamoto** Kentaro Tachibana, "Lightweight and High-Fidelity End-to-End Text-to-Speech with Multi-Band Generation and Inverse Short-Time Fourier Transform," *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. xxxx–xxxx, 2023.
- Website: masayakawamura.github.io/Demo_MB-iSTFT-VITS/
- Code: MasayaKawamura/MB-iSTFT-VITS

2022 Byeongseon Park, **Ryuichi Yamamoto**, Kentaro Tachibana, and Min-Jae Hwang, "A Unified Accent Estimation Method Based on Multi-Task Learning for Japanese Text-to-Speech," *Proc. Interspeech*, pp. 1931–1935, 2022.
- Website: 6gsn.github.io/demos/mtl_accent/

Ryo Terashima, **Ryuichi Yamamoto**, Eunwoo Song, Yuma Shirahata, Hyun-Wook Yoon, Jae-Min Kim, Kentaro Tachibana, "Cross-Speaker Emotion Transfer for Low-Resource Text-to-Speech Using Non-Parallel Voice Conversion with Pitch-Shift Data Augmentation," *Proc. Interspeech*, pp. 3018–3022, 2022.
- Website: ryojerky.github.io/demo_vc-tts-ps/

Eunwoo Song, **Ryuichi Yamamoto**, Ohsung Kwon, Chan-Ho Song, Min-Jae Hwang, Suhyeon Oh, Hyun-Wook Yoon, Jin-Seob Kim, Jae-Min Kim, "TTS-by-TTS 2: Data-selective Augmentation for Neural Speech Synthesis Using Ranking Support Vector Machine with Variational Autoencoder," *Proc. Interspeech*, pp. 1941–1945, 2022.
- Website: sewplay.github.io/demos/txt2/

Takaaki Saeki, Kentaro Tachibana, **Ryuichi Yamamoto**, "DRSpeech: Degradation-Robust Text-to-Speech Synthesis with Frame-Level and Utterance-Level Acoustic Representation Learning," *Proc. Interspeech*, pp. 793–797, 2022.
- Website: takaaki-saeki.github.io/drspeech_demo/

Hyunwook Yoon, Ohsung Kwon, Hoyeon Lee, **Ryuichi Yamamoto**, Eunwoo Song, Jae-Min Kim, and Min-Jae Hwang, "Language Model-Based Emotion Prediction Methods for Emotional Speech Synthesis Systems," *Proc. Interspeech*, pp. 4596–4600, 2022.
- Website: christophyoon.github.io/lmemotiontts/

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, Ziyan 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)

2023 **Ryuichi Yamamoto**, Reo Yoneyama, Tomoki Toda, "NNSVS: A Neural Network-Based Singing Voice Synthesis Toolkit," *The Acoustic Society of Japan (ASJ) Autumn*, pp. xxx–xxx, 2023 (in Japanese).

Robin Scheibler, Takuya Hasumi, Yusuke Fujita, Tatsuya Komatsu, **Ryuichi Yamamoto**, Kentaro Tachibana, "Foley Sound Synthesis with a Class-Conditioned Latent Diffusion Model and FAD-Based Post-filtering," *The Acoustic Society of Japan (ASJ) Autumn*, pp. xxx–xxx, 2023 (in Japanese).

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

2020 – present **nnsvs**
Neural network-based singing voice synthesis library for research
- Role: Creator and core developer
- Code: nnsvs/nnsvs
- Website: nnsvs.github.io/

2021 – present **ttslearn**
Library for the book "Text-to-speech with Python"
- Role: Creator and core developer
- Code: r9y9/ttslearn
- Website: r9y9.github.io/ttslearn/

2017 – present **nnmnkwii**
Library to build speech synthesis systems designed for easy and fast prototyping
- Role: Creator and core developer
- Code: 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
- 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

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
- Website: r9y9.github.io/world.jl/latest/

2015 – 2019 **librosa**
Python library for audio and music analysis.
- Role: Contributor
- Code: librosa/librosa
- Website: librosa.org/

2014 – 2017 **MelGeneralizedCepstrums.jl**
Mel-Generalized Cepstrum analysis
- Role: Creator and core developer
- Code: r9y9/MelGeneralizedCepstrums.jl

RESEARCH PROJECTS

2019 – 2021 **ESPnet**
End-to-End Speech Processing Toolkit
- Role: Discussions and reviews for text-to-speech features
- Code: 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
- 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
- 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

2017 – 2019 **tacotron_pytorch**
PyTorch implementation of Tacotron speech synthesis model
- Role: Creator and core developer
- Code: r9y9/tacotron_pytorch

SUPERVISOR FOR STUDENTS

2023.08 – 2023.09 Reo Shimizu (Tohoku University)

2021.09 – 2023.01 Reo Yoneyama (Nagoya University)

2021.03 – 2022.04 Takaaki Saeki (The University of Tokyo)

LECTURES

2022 AI and Business
Lecture on text-to-speech methods and applications
Graduate School of Medicine, Juntendo University, Nov 2022.

2022 Pattern Recognition III
Lectore on research and development for TTS in industry
Graduate School of Engineering, Nagoya Institute of Technology, Jan 2022, Online.

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 Sesson 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.