Eunwoo Song

Last updated: May 2022

CONTACT INFORMATION

Email gregorio.song@gmail.com

Address 6, Buljeong-ro, Bundang-gu, Gyeonggi-do, 13561, Korea

Phone +82-10-3191-9108

EXPERIENCE

Seongnam, Korea

Mar 2017 - present

Naver Corp.

Senior Research Scientist

HDTS Team Lead, Clova

Topic: Speech synthesis

- Research and development of hybrid speech synthesis system, combining deep learning and unit-selection TTS models. Implementing cloud-based real-time TTS products for Deep learning-based acoustic parameter estimation
  - Naver AI news anchor (Korean celeb voice, Sep 2021, May 2020)
  - Gatebox (Japanese character voice, Oct 2019).
  - Line Car Navi (Japanese navigation, Sep 2019),
  - Naver Maps (Korean navigation, Sep 2019),
  - Naver AI speaker (Korean cel eb voice, Nov 2018),
- Research and development of neural excitation vocoder, incorporating linear prediction filter to WaveNet architecture for quality improvement.
- Research and development of Parallel WaveGAN vocoder, incorporating generative adversarial network to non-autoregressive WaveNet generator.
- Research and development of end-to-end expressive speech synthesis system, leveraging variance autoencoder-based emotion embedding methods.
- Implementing and evaluating state-of-the-art speech synthesis models, such as Tacotron, Transformer, FastSpeech, WaveNet, WaveRNN, WaveFlow. Experimenting on these models by architectural and feature-level modifications.
- Implementing and evaluating parametric vocoders for speech synthesis back-end, such as ITFTE, WORLD, STRAIGHT, Glottal Vocoder, HNM, MBE, MELP. Experimenting on these vocoders by architectural modifications for TTS.

San Diego, CA

Aug 2016 – Nov 2016

Qualcomm Technologies Inc.

Intern for Multimedia Group

Mentor: Dr. Deep Sen

Topic: Spatial audio

- Fixed-point implementation of MPEG-H 3D Audio Decoder

Beijing, China

Apr 2016 – Jun 2016

Sep 2015 – Feb 2016

Microsoft Research Asia

Student Consultant for Speech Group

Mentor: Dr. Frank Soong

Topic: Speech synthesis

- Deep learning-based TTS system using ITFTE vocoder

EDUCATION

Seoul, Korea

Sep 2010 – Feb 2019

Yonsei University

Combined M.S. and Ph.D., Electrical and Electronic Engineering

- Dissertation: Improved time-frequency trajectory excitation vocoder for deep learning-based statistical parametric speech synthesis system
- Advisor: Prof. Hong-Goo Kang

Seoul, Korea

Mar 2006 – Aug 2010

Yonsei University

B.S., Electrical and Electronic Engineering

TALKS

1. "Voice synthesis and applications", KAIST and SNU (2022)
2. "Introduction to text-to-speech", Naver (2021)
3. "Deep learning-based text-to-speech", Yonsei Univ. and Korea Univ. (2021)
4. "Clova AI: Text-to-speech technology", Yonsei Univ. (2020)
5. "Parallel WaveGAN", Naver (2020)
6. "WaveNet". Naver (2018)
7. "Clova voice: From unit-selection to deep learning-based TTS", ASK Conference (2018)
8. "Speaker-adaptive text-to-speech", Naver AI Colloquium (2018)

PUBLICATIONS

1. M.-J, Hwang, R. Yamamoto, **E. Song**, J.-M. Kim, "High-fidelity Parallel WaveGAN with multi-band harmonic-plus-noise model," Proc. INTERSPEECH, 2021, pp. 2227-2231.
2. H.-K. Nguyen, K. Jeong, S. Um, M.-J. Hwang, **E. Song**, H.-G. Kang, "LiteTTS: A decoder-free lightweight text-to-wave synthesis based on generative adversarial networks," Proc. INTERSPEECH, 2021. pp. 3595-3599.
3. R. Yamamoto, **E. Song**, M.-J. Hwang, J.-M. Kim, "Parallel waveform synthesis based on generative adversarial networks with voicing-aware conditional discriminators," Proc. ICASSP, 2021, pp. 6039-6043.
4. M.-J. Hwang, R. Yamamoto, **E. Song**, J.-M. Kim," TTS-by-TTS: TTS-driven data augmentation for fast and high-quality speech synthesis," Proc. ICASSP, 2021, pp. 6598-6602.
5. **E. Song**, R. Yamamoto, M.-J. Hwang, J.-S. Kim, O. Kwon, J.-M. Kim, "Improved Parallel WaveGAN with perceptually weighted spectrogram loss," Proc. SLT, 2021, pp. 470-476.
6. M.-J. Hwang, F. K. Soong, **E. Song**, X. Wang, H. Kang, H.-G. Kang, "LP-WaveNet: Linear prediction-based WaveNet speech synthesis," Proc. APSIPA, 2020, pp. 810-814.

7. S. Oh, H. Lim, K. Byun, M.-J. Hwang, **E. Song**, H.-G. Kang, "ExcitGlow: Improving a WaveGlow-based neural vocoder with linear prediction analysis," Proc. APSIPA, 2020, pp. 831-836.
8. **E. Song**, M.-J. Hwang, R. Yamamoto, J.-S. Kim, O. Kwon, J.-M. Kim, "Neural text-to-speech with a modeling-by-generation excitation vocoder," Proc. INTERSPEECH, 2020, pp. 3570-3574.
9. **E. Song**, J.-S. Kim, K. Byun, H.-G. Kang, "Speaker-adaptive neural vocoders for parametric speech synthesis systems," Proc. MMSP, 2020, pp. 1-5.
10. R. Yamamoto, **E. Song**, J.-M. Kim, "Parallel WaveGAN: A fast waveform generation model based on generative adversarial networks with multi-resolution spectrogram," Proc. ICASSP, 2020, pp. 6194-6198.
11. M.-J. Hwang, **E. Song**, R. Yamamoto, F. K. Soong, H.-G. Kang, "Improving LPCNet-based text-to-speech with linear predictions-structured mixture density network," Proc. ICASSP, 2020, pp. 7214-7218.
12. R. Yamamoto, **E. Song**, J.-M. Kim, "Probability density distillation with generative adversarial networks for high-quality parallel waveform generation," Proc. INTERSPEECH, 2019, pp. 699-703.
13. **E. Song**, K. Byun, H.-G. Kang, "ExcitNet vocoder: A neural excitation model for parametric speech synthesis systems," Proc. EUSIPCO, 2019, pp. 1179-1183.
14. K. Byun, **E. Song**, J. Kim, J.-M. Kim, H.-G. Kang, "Excitation-by-SampleRNN model for text-to-speech," Proc. ITC-CSCC, 2019, pp. 356-359.
15. J. Y. Lee, S. J. Cheon, B. J. Choi, N. S. Kim, **E. Song**, "Acoustic modeling using adversarially trained variational recurrent neural network for speech synthesis," Proc. INTERSPEECH, 2018, pp. 917-921.
16. M.-J. Hwang, **E. Song**, J.-S. Kim, H.-G. Kang, "A unified framework for the generation of glottal signals in deep learning-based parametric speech synthesis systems," Proc. INTERSPEECH, 2018, pp. 912-916.
17. M.-J. Hwang, **E. Song**, H.-G. Kang, "Modeling-by-generation-structured noise compensation algorithm for glottal vocoding speech synthesis system," Proc. ICASSP, 2018, pp. 5669-5673.
18. **E. Song**, F. K. Soong, H.-G. Kang, "Perceptual quality and modeling accuracy of excitation parameters in DLSTM-based speech synthesis systems," Proc. ASRU, 2017, pp. 671–676.
19. **E. Song**, F. K. Soong, H.-G. Kang, "Effective spectral and excitation modeling techniques for LSTM-RNN-based speech synthesis systems," IEEE/ACM Trans. Audio, Speech, and Lang. Process., vol. 25, no. 11, pp. 2152–2161, 2017.
20. **E. Song**, F. K. Soong, H.-G. Kang, "Improved time-frequency trajectory excitation vocoder for DNN-based speech synthesis," Proc. INTERSPEECH, 2016, pp. 874–878.

21. **E. Song**, H.-G. Kang, "Multi-class learning algorithm for deep neural network-based statistical parametric speech synthesis," Proc. EUSIPCO, 2016, pp. 1951–1955.
22. **E. Song**, H.-G. Kang, "Deep neural network-based statistical parametric speech synthesis system using improved time-frequency trajectory excitation model," Proc. INTERSPEECH, 2015, pp. 874–878.
23. K. Byun, **E. Song**, H. Sim, H. Lim, H.-G. Kang, "A constrained two-layer compression technique for ECG waves," Proc. EMBC, 2015, pp. 6130–6133.
24. **E. Song**, Y. S. Joo,, H.-G. Kang, "Improved time-frequency trajectory excitation modeling for a statistical parametric speech synthesis system," Proc. ICASSP, 2015, pp. 4949–4953.
25. **E. Song**, H.-G. Kang, J. Lee, "Fixed-point implementation of MPEG-D unified speech and audio coding decoder," Proc. DSP, 2014, pp. 110–113.
26. **E. Song**, J. Ryu, H.-G. Kang, "Speech enhancement for pathological voice using time-frequency trajectory excitation modeling," Proc. APSIPA, 2013, pp. 1–4.

PREPRINT

1. O. Kwon, **E. Song**, J.-M. Kim, H.-G. Kang, "Effective parameter estimation methods for an ExcitNet model in generative text-to-speech systems," arXiv preprint arXiv:1905. 08486, 2019.

PATENTS

1. KR10-2198598, "Method for generating synthesized speech signal, neural vocoder, and training method thereof," Dec. 2020.
2. KR10-2198597, "Neural vocoder and training method of neural vocoder for constructing speaker-adaptive model," Dec. 2020.

HONORS & AWARDS

1.	Ranked No. 2 in N Innovation Award 2020, Naver Corp.	Dec 2020
2.	The Best Paper Award, APSIPA ASC 2020	Dec 2020
3.	Ranked No. 1 in N Innovation Award 2019, Naver Corp.	Dec 2019
4.	Ranked No. 1 in N Innovation Award 2018, Naver Corp.	Nov 2018
5.	Excellent intern award, Microsoft Research Asia	Jun 2016
6.	Excellent intern award, Microsoft Research Asia	Feb 2016