

Yi-Chiao Wu 吳宜樵

KomaHills 310, Kiyozumi-cho 2-30-5, Chikusa-ku, Nagoya-shi, Aichi Prefecture 464-0034, Japan
+81-80-7891-5085 | yichiao.wu@g.sp.m.is.nagoya-u.ac.jp | <https://bigpon.github.io/>

EDUCATION

- 2017-present **Nagoya University (NU)** **Nagoya, Japan**
- Ph.D. candidate in Graduate school of Informatics (artificial intelligent system group)
 - Research topic: Voice conversion with neural-based speech generation models
 - Advisor: Tomoki Toda
- 2005-2011 **National Chiao Tung University (NCTU)** **Hsinchu, Taiwan**
- M.S. in Communication Engineering (specialized in speaker recognition)
 - B.S. in Communication Engineering
 - Master's thesis: speaker recognition system for intelligent home robot
- 2009-2011 **Industry-university Cooperative Research Project** **Hsinchu, Taiwan**
- Built a human identification system with a microphone array and face and voiceprint recognitions
 - Transferred the speaker recognition system (with windows UI) to COMPAL Electronics

WORK EXPERIENCE

- 2017-present **Graduate School of Informatics, Nagoya University** **Nagoya, Japan**
Research Assistant
- Proposed a pitch-dependent structure for WaveNet to improve the robustness of unseen data
 - Proposed a collapsed speech detection and suppression method for WaveNet vocoder
 - Got an overall performance ranking 2/12 in Voice Conversion Challenge 2018 non-parallel VC task
- 2019 **National Institute of Information and Communications Technology** **Kyoto, Japan**
Summer Intern
- Reduced 30% training time and model size of WaveGlow with the depthwise CNN
- 2015-2017 **Institute of Information Science, Academia Sinica** **Taipei, Taiwan**
Research Assistant
- Got an overall performance ranking 7/17 in Voice Conversion Challenge 2016
 - Combined manifold learning techniques with an exemplar-based speaker voice conversion system
 - Developed an exemplar-based post-filtering speech enhancement technique
 - Integrated exemplar-based post-filtering methods with neural-based speech enhancement systems
- 2013-2015 **Da Vinci Innovation Lab, ASUS** **Taipei, Taiwan**
Software R&D Engineer
- Developed ASUS Zenbo robot's text-independent speaker recognition engine
 - Designed and implemented a speaker recognition mobile application on an Android platform
 - Integrated speaker recognition with a microphone array, noise reduction, and face recognition systems
 - Designed confidence measure algorithms for noisy environment speaker verification
- 2012-2013 **Multimedia BU II, Realtek** **Hsinchu, Taiwan**
System Designer
- Worked with IC component designers to design and verify TV audio systems
 - Ported audio drivers for five mass-production projects (for TOSHIBA, SONY, Skyworth, etc.)
 - Worked with software engineers to pass the Dolby and DTS certification for audio systems

RESEARCH INTEREST

- Speech Generation: voice conversion, speech enhancement, text to speech, and bandwidth expansion

SKILLS

- Machine learning and audio processing tools: PyTorch, TensorFlow, SPTK, WORLD, STRAIGHT
- Programming languages: Python, MATLAB, C/C++, JAVA, UNIX shell script
- Google Scholar Citations: More than 350 citations
- English certification: GRE 323/340; TOEFL 95/120

PUBLICATION AND AWARDS

- 2020
 - NEC 2020 C&C grant for non-japanese researcher
 - P. L. Tobing, **Y. C. Wu** et al., "Efficient shallow WaveNet vocoder using multiple samples output based on laplacian distribution and linear prediction," *Proc. ICASSP*, 2020.
 - W. C. Huang, T. Hayashi, **Y. C. Wu** et al., "Voice Transformer network: sequence-to-sequence voice conversion using Transformer with text-to-speech pretraining", *Submitted to Interspeech*, 2020.
- 2019
 - **Y. C. Wu**, T. Hayashi et al., "Quasi-periodic WaveNet vocoder: a pitch dependent dilated convolution model for parametric speech generation," *Proc. Interspeech*, 2019. (with Interspeech travel grants)
 - **Y. C. Wu**, P. L. Tobing et al., "Statistical voice conversion with quasi-periodic WaveNet vocoder," *Proc. SSW*, 2019.
 - P. L. Tobing, **Y. C. Wu** et al., "Voice conversion with cycleRNN-based spectral mapping and finely tuned WaveNet vocoder," *Proc. IEEE Access*, 2019.
 - P. L. Tobing, **Y. C. Wu** et al., "Voice conversion with cyclic recurrent network and finetuned WaveNet vocoder," *Proc. ICASSP*, 2019.
 - P. L. Tobing, **Y. C. Wu** et al., "Non-parallel voice conversion with cyclic variational autoencoder," *Proc. Interspeech*, 2019.
 - W. C. Huang, **Y. C. Wu** et al., "Investigation of F0 conditioning and fully convolutional network in variational autoencoder based voice conversion," *Proc. Interspeech*, 2019.
 - W. C. Huang, **Y. C. Wu** et al., "Generalization of spectrum differential based direct waveform modification for voice conversion," *Proc. SSW*, 2019.
 - W. C. Huang, **Y. C. Wu** et al., "Refined WaveNet vocoder for variational autoencoder based voice conversion," *Proc. EUSIPCO*, 2019.
- 2018
 - **Y. C. Wu**, K. Kobayashi et al., "Collapsed speech segment detection and suppression for WaveNet vocoder," *Proc. Interspeech*, 2018.
 - **Y. C. Wu**, P. L. Tobing et al., "The NU non-parallel voice conversion system for the voice conversion challenge 2018," *Proc. Odyssey*, 2018.
 - P. L. Tobing, **Y. C. Wu** et al., "NU voice conversion system for the voice conversion challenge 2018," *Proc. Odyssey*, 2018.
 - H. T. Hwang, **Y. C. Wu** et al., "Voice conversion based on locally linear embedding," *Proc. Journal of Information Science and Engineering*, 2018.
 - H. T. Hwang, **Y. C. Wu** et al., "Locally linear embedding based post-filtering for speech enhancement," *Proc. Journal of Information Science and Engineering*, 2018.
 - Y. H. Peng, H. T. Hwang, **Y. C. Wu** et al., "Exemplar-based spectral detail compensation for voice conversion," *Proc. Interspeech*, 2018.
 - P. L. Tobing, T. Hayashi, **Y. C. Wu** et al., "An evaluation of deep spectral mappings and WaveNet vocoder for voice conversion," *Proc. IEEE SLT*, 2018.
- 2017
 - **Y. C. Wu**, H. T. Hwang et al., "A post-filtering approach based on locally linear embedding difference compensation for speech enhancement," *Proc. Interspeech*, 2017.
 - **Y. C. Wu**, H. T. Hwang et al., "A locally linear embedding based postfiltering approach for speech enhancement," *Proc. ICASSP*, 2017.
 - C. C. Hsu, H. T. Hwang, **Y. C. Wu** et al., "Voice conversion from unaligned corpora using variational autoencoding Wasserstein generative adversarial networks," *Proc. Interspeech*, 2017.
 - Y. H. Peng, C. C. Hsu, **Y. C. Wu** et al., "Fast locally linear embedding algorithm for exemplar-based voice conversion," *Proc. APSIPA*, 2017.
- 2016
 - **Y. C. Wu**, H. T. Hwang et al., "Locally linear embedding for exemplar-based spectral conversion," *Proc. Interspeech*, 2016. (oral presentation; special session: voice conversion challenge 2016)
 - C. C. Hsu, H. T. Hwang, **Y. C. Wu** et al., "Dictionary update for NMF-based voice conversion using an encoder-decoder network," *Proc. ISCSLP*, 2016.
 - C. C. Hsu, H. T. Hwang, **Y. C. Wu** et al., "Voice conversion from non-parallel corpora using variational auto-encoders," *Proc. APSIPA*, 2016.
- 2014
 - United States patent (US8879799B2): human identification system