WEN-CHIN HUANG

vincenthuang316@gmail.com ⋄ wen.chinhuang@g.sp.m.is.nagoya-u.ac.jp
https://unilight.github.io/

EDUCATION

Nagoya University Apr. 2021 - Mar. 2024 (expected)
Ph.D candidate, Graduate School of Informatics
Advisor: Prof. Tomoki Toda

Nagoya University Apr. 2019 - Mar. 2021
M.S., Graduate School of Informatics
Advisor: Prof. Tomoki Toda
Thesis: Transfer Learning for Sequence-to-Sequence Voice Conversion

National Taiwan University Sep. 2014 - Jun. 2018
B.S. in Computer Science & Information Science
Overall GPA 3.79/4.3; Last 60 4.04/4.3

PROFESSIONAL EXPERIENCES

NTT Communication Science Laboratories, NTT Corporation, Japan
Aug. 2019 - Sep. 2019
Research Intern
Advisor: Dr. Hirokazu Kameoka

Institute of Information Science, Academia Sinica, Taiwan Jul. 2017 - Mar. 2019, Aug. 2020 - Present
Research Assitant
Advisor: Dr. Hsin-Min Wang, Dr. Yu Tsao

ACADEMIC ACTIVITIES

Organizer/Committee Member

1. Organizing Committee, Voice Conversion Challenge 2020

Reviewer

1. IEEE Speech Processing Letters (2020)
2. IEEE Transactions on Audio, Speech and Language Processing (2020)
3. Neural Networks (2020)
4. Joint Workshop for the Blizzard Challenge and Voice Conversion Challenge 2020

HONORS

- **Research Fellowship for Young Scientists (DC1)**, from Japan Society for the Promotion of Science (JSPS), April 2021 - March 2024
- **Scholarship for International Students**, JEES Docomo, April 2019 - March 2021
- **Travel grant**, ISCA and Interspeech 2019
- **Best Student Paper Award**, The 11th International Symposium on Chinese Spoken Language Processing (ISCSLP), 2018

PUBLICATIONS

Journals

1. W.-C. Huang, T. Hayashi, Y. C. Wu, H. Kameoka, and T. Toda, "Pretraining techniques for sequence-to-sequence voice conversion," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 745–755, 2021
2. H. Kameoka, W.-C. Huang, K. Tanaka, T. Kaneko, N. Hojo, and T. Toda, "Many-to-many voice transformer network," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 656–670, 2021
3. W.-C. Huang, H. Luo, H.-T. Hwang, C.-C. Lo, Y.-H. Peng, Y. Tsao, and H.-M. Wang, "Unsupervised Representation Disentanglement Using Cross Domain Features and Adversarial Learning in Variational Autoencoder Based Voice Conversion," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 4, no. 4, pp. 468–479, 2020
4. X. Wang, J. Yamagishi, M. Todisco, H. Delgado, A. Nautsch, N. Evans, M. Sahidullah, V. Vestman, T. Kinnunen, K. A. Lee, L. Juvela, P. Alku, Y.-H. Peng, H.-T. Hwang, Y. Tsao, H.-M. Wang, S. L. Maguer, M. Becker, F. Henderson, R. Clark, Y. Zhang, Q. Wang, Y. Jia, K. Onuma, K. Mushika, T. Kaneda, Y. Jiang, L.-J. Liu, Y.-C. Wu, W.-C. Huang, T. Toda, K. Tanaka, H. Kameoka, I. Steiner, D. Matrouf, J.-F. Bonastre, A. Govender, S. Ronanki, J.-X. Zhang, and Z.-H. Ling, "Asvspoof 2019: a large-scale public database of synthetized, converted and replayed speech," *Computer Speech & Language*, vol. 64, p. 101114, 2020

Peer-reviewed Conferences and Workshops

1. W.-C. Huang, C.-H. Wu, S.-B. Luo, K.-Y. Chen, H.-M. Wang, and T. Toda, "Speech recognition by simply fine-tuning bert," *arXiv preprint arXiv:2102.00291*, 2021, to appear at 2021 IEEE International Conference on Acoustics, Speech and Signal Processing
2. W.-C. Huang, T. Hayashi, Y.-C. Wu, and T. Toda, "Any-to-One Sequence-to-Sequence Voice Conversion using Self-Supervised Discrete Speech Representations," *arXiv preprint arXiv:2010.12231*, 2020, to appear at 2021 IEEE International Conference on Acoustics, Speech and Signal Processing
3. K. Kobayashi, W.-C. Huang, Y.-C. Wu, P. L. Tobing, T. Hayashi, and T. Toda, "CRANK: an Open-Source Software for Nonparallel Voice Conversion based on Vetor-Quantized Variational Autoencoder," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021
4. T. Hayashi, W.-C. Huang, K. Kobayashi, and T. Toda, "Non-autoregressive sequence-to-sequence voice conversion," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021
5. Y.-W. Chen, K.-H. Hung, S.-Y. Chuang, J. Sherman, W.-C. Huang, X. Lu, and Y. Tsao, "Ema2s: An end-to-end multimodal articulatory-to-speech system," *arXiv preprint arXiv:2102.03786*, 2021, to appear at 2021 IEEE International Symposium on Circuits and Systems (ISCAS)
6. Z. Yi, W.-C. Huang, X. Tian, J. Yamagishi, R. K. Das, T. Kinnunen, Z.-H. Ling, and T. Toda, "Voice Conversion Challenge 2020 – Intra-lingual semi-parallel and cross-lingual voice conversion –," in *Proc. Joint Workshop for the Blizzard Challenge and Voice Conversion Challenge 2020*, 2020, pp. 80–98
7. R. K. Das, T. Kinnunen, W.-C. Huang, Z.-H. Ling, J. Yamagishi, Z. Yi, X. Tian, and T. Toda, "Predictions of Subjective Ratings and Spoofing Assessments of Voice Conversion Challenge 2020 Submissions," in *Proc. Joint Workshop for the Blizzard Challenge and Voice Conversion Challenge 2020*, 2020, pp. 99–120
8. W.-C. Huang, T. Hayashi, S. Watanabe, and T. Toda, "The Sequence-to-Sequence Baseline for the Voice Conversion Challenge 2020: Cascading ASR and TTS," in *Proc. Joint Workshop for the Blizzard Challenge and Voice Conversion Challenge 2020*, 2020, pp. 160–164
9. W.-C. Huang, P. L. Tobing, Y.-C. Wu, K. Kobayashi, and T. Toda, "The NU Voice Conversion System for the Voice Conversion Challenge 2020: On the Effectiveness of Sequence-to-sequence

Models and Autoregressive Neural Vocoders," in *Proc. Joint Workshop for the Blizzard Challenge and Voice Conversion Challenge 2020*, 2020, pp. 165–169

10. W.-C. Huang, T. Hayashi, Y.-C. Wu, H. Kameoka, and T. Toda, "Voice Transformer Network: Sequence-to-Sequence Voice Conversion Using Transformer with Text-to-Speech Pretraining," in *Proc. Interspeech*, 2020, pp. 4676–4680
11. W.-C. Huang, Y.-C. Wu, K. Kobayashi, Y.-H. Peng, H.-T. Hwang, P. Lumban Tobing, T. Toda, Y. Tsao, and H.-M. Wang, "Generalization of Spectrum Differential based Direct Waveform Modification for Voice Conversion," in *Proc. 10th ISCA Speech Synthesis Workshop*, 2019, pp. 57–62
12. W.-C. Huang, Y.-C. Wu, C.-C. Lo, P. Lumban Tobing, T. Hayashi, K. Kobayashi, T. Toda, Y. Tsao, and H.-M. Wang, "Investigation of F0 Conditioning and Fully Convolutional Networks in Variational Autoencoder Based Voice Conversion," in *Proc. Interspeech*, 2019, pp. 709–713
13. W.-C. Huang, Y.-C. Wu, H.-T. Hwang, P. Lumban Tobing, T. Hayashi, K. Kobayashi, T. Toda, Y. Tsao, and H.-M. Wang, "Refined WaveNet Vocoder for Variational Autoencoder Based Voice Conversion," in *Proc. 27th European Signal Processing Conference (EUSIPCO)*, Sep 2019
14. W.-C. Huang, H.-T. Hwang, Y.-H. Peng, Y. Tsao, and H.-M. Wang, "Voice conversion based on cross-domain features using variational auto encoders," in *Proc. The 11th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, Nov 2018
15. C.-C. Lo, S.-W. Fu, W.-C. Huang, X. Wang, J. Yamagishi, Y. Tsao, and H.-M. Wang, "MOSNet: Deep Learning based Objective Assessment for Voice Conversion," in *Proc. Interspeech*, 2019, pp. 1541–1545

Domestic Conferences

1. W.-C. Huang, Y.-C. Wu, H.-T. Hwang, P. Lumban Tobing, T. Hayashi, K. Kobayashi, T. Toda, Y. Tsao, and H.-M. Wang, "Reducing mismatch of WaveNet vocoder for variational autoencoder based voice conversion," in *ASJ*, Mar 2019
2. W.-C. Huang, C.-C. Lo, H.-T. Hwang, Y. Tsao, and H.-M. Wang, "Wavenet vocoder and its applications in voice conversion," in *Proc. The 30th ROCLING Conference on Computational Linguistics and Speech Processing (ROCLING)*, Oct 2018

MISCELLANEOUS WORKS

Invited/visiting talks

1. W.-C. Huang "Machine Reading Comprehension with Deep Learning", National Taiwan University of Science and Technology, May. 2018

Open-source software development

1. ESPnet: End-to-end speech processing toolkit.
2. CRANK: Open-source software for nonparallel voice conversion based on vetor-quantized variational autoencoder

SKILLS

Languages	Mandarin: native English: fluent (TOEIC 960, TOEFL 108) Japanese: intermediate (JLPT N2)
Programming languages	Python, Matlab, C++
Deep learning frameworks	PyTorch, TensorFlow, Chainer

Last updated: 2021/02/26