List of Publications

Yi-Chiao Wu

March 2021

Ph.D. Thesis

[1] Yi-Chiao Wu, "Incorporating prior knowledge on speech production mechanism into neural speech waveform generation," in the graduate school of informatics (artificial intelligent group) of Nagoya University, Feb. 2021.

Journal Papers

[1] Y.-C. Wu, P. L. Tobing, K. Kobayashi, T. Hayashi, and T. Toda, "Non-parallel voice conversion system with WaveNet vocoder and collapsed speech suppression," in IEEE Access, vol. 8, pp. 62094–62106, Apr. 2020.

[2] Y.-C. Wu, T. Hayashi, P. L. Tobing, K. Kobayashi, and T. Toda, "Quasi-Periodic WaveNet: an autoregressive raw waveform generative model with pitch-dependent dilated convolution neural network," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 29, pp. 1134-1148, 2021.

[3] Y.-C. Wu, T. Hayashi, T. Okamoto, H. Kawai, and T. Toda, "Quasi-Periodic Parallel WaveGAN: a non-autoregressive raw waveform generative model with pitch-dependent dilated convolution neural network," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 29, pp. 792-806, 2021.

[4] H-T. Hwang, Y.-C. Wu, Y.-H. Peng, C.-C. Hsu, Y. Tsao, H.-M. Wang, Y.-R. Wang, and S.-H. Chen, "Voice conversion based on locally linear embedding," in Journal of Information Science and Engineering, vol. 34, pp. 1469–1491, 2018.

[5] H-T. Hwang, Y.-C. Wu, S.-S. Wang, C.-C. Hsu, Y. Tsao, H.-M. Wang, Y.-R. Wang, and S.-H. Chen, "Locally linear embedding based post-filtering for speech enhancement," in Journal of Information Science and Engineering, vol. 34, pp. 1493–1516, 2018.

[6] P. L. Tobing, Y.-C. Wu, T. Hayashi, K. Kobayashi, and T. Toda, "Voice conversion eith cycleRNN-based spectral mapping and finely tuned WaveNet vocoder," in IEEE Access, vol. 7, pp. 171114–171125, Apr. 2019.

[7] 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. Le 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, Z.-H. Ling, "ASVspoof 2019: a large-scale public database of synthetic, converted

and replayed speech," in Computer Speech and Language, Vol. 64, Article 101114, 25 pages, Nov. 2020.

[8] P. L. Tobing, Y.-C. Wu, K. Kobayashi, T. Hayashi, and T. Toda, "An evaluation of voice conversion with neural network spectral mapping models and WaveNet vocoder," in APSIPA Transactions on Signal and Information Processing, vol. 9, e26, pp. 1-14, Nov. 2020.

[9] W. -C. Huang and T. Hayashi and Y. -C. Wu and H. Kameoka and T. Toda, "Pretraining Techniques for Sequence-to-Sequence Voice Conversion," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 29, pp. 745-755, 2021.

International Conferences

[1] Y.-C. Wu, H-T. Hwang, C.-C. Hsu, Y. Tsao, and H.-M. Wang, "Locally linear embedding for exemplar-based spectral conversion," Proc. INTERSPEECH, pp. 1652–165, Sept. 2016.

[2] Y.-C. Wu, H.-T. Hwang, S.-S. Wang, C.-C. Hsu, Y.-H. Lai, Y. Tsao, and H.-M. Wang, "A locally linear embedding based postfiltering approach for speech enhancement," Proc. ICCASP, pp. 5555–5559, Mar. 2017.

[3] Y.-C. Wu, H.-T. Hwang, S.-S. Wang, C.-C. Hsu, Y. Tsao, and H.-M. Wang, "A postfiltering approach based on locally linear embedding difference compensation for speech enhancement." Proc. INTERSPEECH, pp. 1953–1957, Aug. 2017.

[4] Y.-C. Wu, P. L. Tobing, T. Hayashi, K. Kobayashi, and T. Toda, "The NU non-parallel voice conversion system for the voice conversion challenge 2018," Proc. Speaker Odyssey, pp. 211–218, Les Sables d'Olonne, France, Jun. 2018.

[5] Y.-C. Wu, K. Kobayashi, P. L. Tobing, T. Hayashi, and T. Toda, "Collapsed speech segment detection and suppression for WaveNet vocoder," Proc. INTERSPEECH, pp. 1988–1992, Hyderabad, India, Sep. 2018.

[6] Y.-C. Wu, T. Hayashi, P. L. Tobing, K. Kobayashi, and T. Toda, "Quasi-Periodic WaveNet vocoder: a pitch dependent dilated convolution model for parametric speech generation," Proc. INTERSPEECH, pp. 196–200, Graz, Austria, Sep. 2019.

[7] Y.-C. Wu, P. L. Tobing, T. Hayashi, K. Kobayashi, and T. Toda, "Statistical voice conversion with quasi-periodic WaveNet vocoder," Proc. SSW10, pp. 63–68, Vienna, Austria, Sep. 2019.

[8] Y.-C. Wu, T. Hayashi, T. Okamoto, H. Kawai, and T. Toda, "Quasi-Periodic Parallel WaveGAN vocoder: a non-autoregressive pitch-dependent dilated convolution model for parametric speech generation," Proc. INTERSPEECH, Full virtual, Oct. 2020.

[9] Y.-C. Wu, P. L. Tobing, K. Yasuhara, N. Matsunaga, Y. Ohtani, and T. Toda, "A cyclical post-filtering approach to mismatch refinement of neural vocoder for text-to-speech systems," Proc. INTERSPEECH, Full virtual, Oct. 2020.

[10] C.-C. Hsu, H-T. Hwang, Y.-C. Wu, Y. Tsao, and H.-M. Wang, "Dictionary update for NMF-based voice conversion using an encoder-decoder network," Proc. ISCSLP, pp. 1–5, 2016.

[11] C.-C. Hsu, H-T. Hwang, Y.-C. Wu, Y. Tsao, and H.-M. Wang, "Voice conversion from non-parallel corpora using variational auto-encoder," Proc. APSIPA, pp. 1–6, 2016.

[12] C.-C. Hsu, H-T. Hwang, Y.-C. Wu, Y. Tsao, and H.-M. Wang, "Voice conversion from unaligned corpora using variational autoencoding wasserstein generative adversarial networks," Proc. INTERSPEECH, pp. 3364–3368, Aug. 2017.

[13] Y.-H. Peng, C.-C. Hsu, Y.-C. Wu, H-T. Hwang, Y.-W. Liu, Y. Tsao, and H.-M. Wang, "Fast locally linear embedding algorithm for exemplar-based voice conversion," Proc. APSIPA, pp. 591–595, 2017.

[14] P. L. Tobing, Y.-C. Wu, T. Hayashi, K. Kobayashi, and T. Toda, "NU voice conversion system for the voice conversion challenge 2018," Proc. Speaker Odyssey, pp. 219–226, Les Sables d'Olonne, France, Jun. 2018.

[15] Y.-H. Peng, H-T. Hwang, Y.-C. Wu, Y. Tsao, and H.-M. Wang, "Exemplar-based spectral detail compensation for voice conversion," Proc. INTERSPEECH, pp. 486–490, Hyderabad, India, Sep. 2018.

[16] P. L. Tobing, T. Hayashi, Y.-C. Wu, K. Kobayashi, and T. Toda, "An evaluation of deep spectral mappings and WaveNet vocoder for voice conversion," Proc. SLT, pp. 297–303, Athens, Greece, Dec. 2018.

[17] P. L. Tobing, Y.-C. Wu, T. Hayashi, K. Kobayashi, and T. Toda, "Voice conversion with cyclic recurrent neural network and fine-tuned WaveNet vocoder," Proc. ICASSP, pp. 6815–6819, Brighton, UK, May 2019.

[18] W.-C. Huang, Y.-C. Wu, H.-T. Hwang, P. L. Tobing,T. Hayashi, K. Kobayashi, T. Toda, Y. Tsao, and H.-M. Wang, "Refined WaveNet vocoder for variational autoencoder based voice conversion," Proc. EUSIPCO, pp. 1–5, A Coruna, Spain, Sep. 2019.

[19] W.-C. Huang, Y.-C. Wu, C.-C. Lo , P. L. 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," Proc. INTERSPEECH, pp. 709–713, Graz, Austria, Sep. 2019.

[20] P. L. Tobing, Y.-C. Wu, T. Hayashi, K. Kobayashi, and T. Toda, "Non-parallel voice conversion with cyclic variational autoencoder," Proc. INTERSPEECH, pp. 674–678, Graz, Austria, Sep. 2019.

[21] W.-C. Huang, Y.-C. Wu, K. Kobayashi, Y.-H. Peng, H.-T. Hwang, P.L. Tobing, Y. Tsao, H.-M. Wang, and T. Toda, "Generalization of spectrum differential based direct waveform modification for voice conversion," Proc. SSW10, pp. 57–62, Vienna, Austria, Sep. 2019.

[22] P.L. Tobing, Y.-C. Wu, T. Hayashi, K. Kobayashi, and T. Toda, "Efficient shallow WaveNet vocoder using multiple samples output based on Laplacian distribution and linear prediction," Proc. ICASSP, Full virtual, pp. 7204–7208, May 2020.

[23] P.L. Tobing, T. Hayashi, Y.-C. Wu, K. Kobayashi, and T. Toda, "Cyclic spectral modeling for unsupervised unit discovery into voice conversion with excitation and waveform modeling," Proc. INTERSPEECH, Full virtual, Oct. 2020.

[24] 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," Proc. INTERSPEECH, Full virtual, Oct. 2020.

[25] P.L. Tobing, Y.-C. Wu, and T. Toda, "Baseline system of voice conversion challenge 2020 with cyclic variational autoencoder and parallel WaveGAN," Proc. Joint Workshop for the Blizzard Challenge and Voice Conversion Challenge 2020, Full virtual, Oct. 2020.

[26] 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," Proc. Joint Workshop for the Blizzard Challenge and Voice Conversion Challenge 2020, Full virtual, Oct. 2020.