GUANLONG ZHAO

+1 (979) 213-2315 | guanlongzhao@gmail.com | github.com/guanlongzhao | linkedin.com/in/guanlongzhao

Research Areas

Speech synthesis, voice conversion, accent conversion, acoustic modeling for speech recognition

Education

Ph.D. in Computer Science, Texas A&M University, advised by Dr. Ricardo Gutierrez-Osuna — May 2020 (expected)

B.S. in Physics & Minor in Computer Science, University of Science and Technology of China — June 2015

Technologies

Python, C/C++, Shell, SQL, LaTeX, HTML/CSS, Matlab, PyTorch, Kaldi ASR, Caffe

Experience

Software Engineering Intern (C++ & Python) — ***Google, Geo Machine Perception Team*** — May – Aug 2019

- Built an unsupervised semantic segmentation model for large (*multi-TB*) Google Street View 3D Lidar point-cloud data using a combination of the DeepLab model, 2D-to-3D projection, and dense CRF
- Constructed and deployed a Lidar data labeling tool into production to collect ground-truth semantic annotations
- Obtained *18%* relative improvements in segmentation accuracy compared with the previous internal system

Software Engineering Intern (C++ & Python) — ***Google, Speech Team*** — June – Aug 2018

- Implemented a GMM forced-aligner that can generate graphemic forced-alignment for low-resource languages
- Established an end-to-end pipeline to train alignment-based graphemic acoustic models for several Indic languages
- Improved the load-balancing strategy of the production acoustic-model-refreshing infrastructure (still in use)

Research Assistant (Python & Matlab) — ***Texas A&M University*** — Sept 2015 – May 2019

- Developed accent conversion algorithms to reduce the foreign accents in non-native speech using sequence-to-sequence models (*PyTorch*) and DNN-based acoustic modeling (*Kaldi ASR*)
- Led the *L2-ARCTIC* project (psi.engr.tamu.edu/l2-arctic-corpus) that open-sourced the first accent-diverse non-native English corpus for speech synthesis

Awards

Graduate Student Presentation Grant, Texas A&M University — 2017

Outstanding Graduate Award, University of Science and Technology of China — 2015

Outstanding Undergraduate Scholarship, University of Science and Technology of China — 2011 – 2014

Selected Publications

Zhao, G., and R. Gutierrez-Osuna. "Using Phonetic Posteriorgram Based Frame Pairing for Segmental Accent Conversion." *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 27.10 (2019): 1649-1660.

Zhao, G., et al. "Foreign Accent Conversion by Synthesizing Speech from Phonetic Posteriorgrams." *Interspeech*. 2019.

Liu, Y., **G. Zhao**, et al. "Image Dehazing: Improved Techniques." *Deep Learning through Sparse and Low-Rank Modeling*. Academic Press, 2019. 251-262.

Zhao, G., et al. "L2-ARCTIC: A Non-native English Speech Corpus." *Interspeech.* 2018.

Ding, S., **G. Zhao**, et al. "Improving Sparse Representations in Exemplar-Based Voice Conversion with a Phoneme-Selective Objective Function." *Interspeech*. 2018.

Liu, Y., **G. Zhao**, et al. "Improved techniques for learning to dehaze and beyond: A collective study." arXiv preprint arXiv:1807.00202 (2018).

Zhao, G., et al. "Accent conversion using phonetic posteriorgrams." *ICASSP*. IEEE, 2018.

Liberatore, C., **G. Zhao**, and R. Gutierrez-Osuna. "Voice Conversion through Residual Warping in a Sparse, Anchor-Based Representation of Speech." *ICASSP*. IEEE, 2018.

Zhao, G., and R. Gutierrez-Osuna. "Exemplar selection methods in voice conversion." *ICASSP*. IEEE, 2017.