

GUANLONG ZHAO

📧 [guanlongzhao.github.io](https://github.com/guanlongzhao) | 📞 +1 (979) 213-2315 | ✉ guanlongzhao@gmail.com | 🔗 [linkedin.com/in/guanlongzhao](https://www.linkedin.com/in/guanlongzhao)

Expertise

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 2020
B.S. in Applied Physics (minor in Computer Science), University of Science and Technology of China 2015

Technologies

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

Work 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

Research Assistant (Python & Matlab) *Texas A&M University* Sept 2015 – May 2020

- 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 Travel Grant, Texas A&M University 2017, 2019
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

G. Zhao and R. Gutierrez-Osuna, "Using phonetic posteriorgram based frame pairing for segmental accent conversion," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 10, pp. 1649–1660, 2019.

S. Ding, G. Zhao, C. Liberatore, and R. Gutierrez-Osuna, "Learning structured sparse representations for voice conversion," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 343–354, 2019.

G. Zhao, S. Ding, and R. Gutierrez-Osuna, "Foreign accent conversion by synthesizing speech from phonetic posteriorgrams," in *Interspeech*, 2019, pp. 2843–2847.

Y. Liu, G. Zhao, B. Gong, Y. Li, R. Raj, N. Goel, S. Kesav, S. Gottimukkala, Z. Wang, W. Ren, and D. Tao, "Chapter 10 — Image dehazing: Improved techniques," in *Deep Learning through Sparse and Low-Rank Modeling*: Elsevier, 2019, pp. 251–262.

G. Zhao *et al.*, "L2-ARCTIC: A non-native English speech corpus," in *Interspeech*, 2018, pp. 2783–2787.

C. Liberatore, G. Zhao, and R. Gutierrez-Osuna, "Voice conversion through residual warping in a sparse, anchor-based representation of speech," in *ICASSP*, 2018, pp. 5284–5288.

G. Zhao, S. Sonsaat, J. Levis, E. Chukharev-Hudilainen, and R. Gutierrez-Osuna, "Accent conversion using phonetic posteriorgrams," in *ICASSP*, 2018, pp. 5314–5318.

G. Zhao and R. Gutierrez-Osuna, "Exemplar selection methods in voice conversion," in *ICASSP*, 2017, pp. 5525–5529.