Shaojin Ding

shjd@tamu.edu, +1-(979)-412-5190
LinkedIn: https://www.linkedin.com/in/shaojin-ding-5b9092162/

Research Interests

- **Speech synthesis**, **Voice conversion**, **Speech recognition**, Person Re-ID, Face recognition

Key Skills

Programming Languages: **Python**, **MATLAB**, C/C++, Ruby, HTML, Javascript
Toolkits: Pytorch, Kaldi, Caffe, TensorFlow, Django

Education

Texas A&M University, College Station, TX, USA
Ph.D. program in Department of Computer Science and Engineering — 2016 – 2021 (expected)

Xi'an Jiaotong University, Xi'an, Shaanxi, China
B.S. in Electronic and Information Engineering — 2011 – 2015
Special Class for the Gifted Youth — 2009 – 2011

Projects

Voice Conversion and Foreign Accent Conversion (Sponsored by NSF) Fall 2016 – present
Advisor: Dr. Ricardo Gutierrez-Osuna

- Developed various speech synthesis systems that convert the speech from a source speaker to sound as if a target speaker had produced them.
- Focused on sparse representation based methods and deep neural network based methods (see publications for detailed methods).
- Achieved over 3.2 Mean Opinion Score acoustic quality on ARCTIC dataset.
- Applied voice conversion techniques on foreign accent conversion to modify non-native English speech to have native accents.

Attentive But Diverse Person Re-ID Fall 2018 – present
Advisor: Dr. Zhangyang Wang

- Although local attention mechanisms achieved reasonable results on Person Re-ID tasks, they are likely to overfit local noise. As a result, we will need to obtain more diverse features to span a compact and representative feature space.
- To address the problem, we designed the network according to an attentive but diverse paradigm. We proposed a self-channel attention and an orthogonality constraint in lower layers, and a dual attention module in higher layers.
- Preliminary results achieve ~92.5% top-1 accuracy and ~81.0% mAP on Market 1501 dataset (Densenet 121 backbone).

Face Verification in Unconstrained Conditions Fall 2014 – Spring 2016

- Proposed a CNN layer to extract global to local facial features concerning the feature of the whole image and specific areas including eyes, nose and mouth.
- Integrated the proposed layer into Google's FaceNet to extract facial features, and implemented Joint-Bayesian face verification model.
- This algorithm achieved 3rd place in the Face Detection and Verification task in National Smart-City Video Parsing Competition.

Typing correction model for Touchscreen Keyboards Jul. 2014 – Sep. 2014

- Designed an unsupervised online learning algorithm that is able to continuously learn the keyboard typing pattern. The algorithm is adaptable to specific users, and the learned pattern can effectively eliminate typing errors.
- The algorithm has been integrated to Microsoft Bing IME for Android.

Experiences

Research Assistant Fall 2016 – present

PSI Lab, Department of Computer Science and Engineering, TAMU.

Advisor: Dr. Ricardo Gutierrez-Osuna

- Conducted research on speech synthesis, voice conversion, and accent conversion.

Research Assistant Fall 2014 – Spring 2016

Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University.

Advisor: Dr. Jinjun Wang

- Conducted research on face detection and verification.

Intern Jul. 2014 – Sep. 2014

Wireless and network group, Microsoft Research Asia, Beijing, China

Advisor: Lead Researcher Jacky Shen

- Conducted research on Typing correction model for Touchscreen Keyboards.

Publications

Shaojin Ding, Christopher Liberatore and Ricardo Gutierrez-Osuna "Learning Structured Dictionaries for Exemplar-based Voice Conversion," in *Proceedings of INTERSPEECH, 2018.*

Shaojin Ding, Guanlong Zhao, Christopher Liberatore and Ricardo Gutierrez-Osuna "Improving Sparse Representations in Exemplar-Based Voice Conversion with a Phoneme-Selective Objective Function," in *Proceedings of INTERSPEECH, 2018.*

Shaojin Ding, Christopher Liberatore, Guanlong Zhao, Sinem Sonsaat, Evgeny Chukharev-Hudilainen, John Levis and Ricardo Gutierrez-Osuna "Golden Speaker Builder: an interactive online tool for L2 learners to build pronunciation models," *in Proceedings of Pronunciation in Second Language Learning and Teaching, 2017.*