赵冠龙

🌐 guanlongzhao.github.io I 📱 +1 (979) 213-2315 I ✉ guanlongzhao@gmail.com I 🔗 linkedin.com/in/guanlongzhao

专业领域

语音合成，说话人转换，口音转换，语音识别中的声学模型

教育经历

博士，计算机科学	德州农工大学，美国	2020
理学学士，应用物理学（辅修计算机科学）	中国科学技术大学，中国	2015

专业技能

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

工作经历

软件工程实习生（C++ & Python）   谷歌，地图组   2019/5 – 2019/8

- 搭建了可用于大规模 TB 级 3D 激光雷达点云数据的无监督语义分割模型，也即自动给谷歌街景车所采集到的每一个激光雷达点标注类别标签，例如"车辆"，"道路"，"建筑物"等
- 在生产环境中，实现并部署了一个用于手动标注激光雷达点类别标签的交互式标注工具
- 与之前的内部基准系统相比，这一新的语义分割模型将标注准确度提高了 18%

软件工程实习生（C++ & Python）   谷歌，语音组   2018/6 – 2018/8

- 实现了一个基于 GMM 的强制对齐工具（forced-aligner），该工具可生成字素（grapheme）粒度的强制对齐
- 搭建了一套端到端的模型训练流程，可为数种语言（印地语，泰米尔语，孟加拉语）训练基于字素的声学模型
- 改进了生产环境中用于声学模型自动更新的基础架构的负载均衡策略

研究助理（Python & Matlab）   德州农工大学   2015/9 – 2020/5

- 发明了数种利用序列至序列模型（PyTorch）以及神经网络声学模型（Kaldi ASR）来进行口音转换的新算法
- 领导创建并发布了 L2-ARCTIC 语料库（psi.engr.tamu.edu/l2-arctic-corpus），该语料库是当时第一个开源的、包含多种非母语口音的英文语音合成数据集

奖项

研究生科研交流基金	德州农工大学	2017, 2019
优秀毕业生奖	中国科学技术大学	2015
优秀本科生奖学金	中国科学技术大学	2011 – 2014

代表性论文

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.