

丁少锦

shjd@tamu.edu, +1-(979)-412-5190

LinkedIn: <https://www.linkedin.com/in/shaojin-ding-5b9092162/>

研究方向

- 语音合成, 语音说话人转换, 语音识别, 物体识别, 人脸识别

核心技能

编程语言: **Python, MATLAB**, C/C++, Ruby, HTML, Javascript

工具包: Pytorch, Kaldi, Caffe, TensorFlow, Django

教育经历

德州农工大学, 大学城, 德克萨斯, 美国

计算机专业工程博士 (Ph.D.) 在读

2016 - 2021 (预计毕业时间)

西安交通大学, 西安, 陕西, 中国

电子信息工程专业学士学位 (B.S.)

2011 - 2015

西安交大少年班

2009 - 2011

科研项目

语音说话人转换和非本土英语口语转换 (美国国家科学基金委 **NSF** 支持) 2016 年秋季 - 现在

- 搭建了若干种不同的语音说话人转换系统。这些系统能够在保持语音内容不变的情况下将说话人转换为指定的目标说话人。
- 研究方法主要基于稀疏编码 (Sparse Coding) 和深度神经网络 (Deep Neural Network)。 (方法细节见学术论文部分)
- 在 ARCTIC 数据集上达到超过 3.5 Mean Opinion Score 的音频质量。
- 将语音说话人转换技术应用在非本土英语口语转换, 实现了将非本土口音的英语语音转换为对应的本土口音英语语音。

Attentive But Diverse 行人重识别

2018 年秋季 - 现在

- 虽然局部注意力机制在行人重识别中取得了很好的效果, 但是这种机制很有可能过拟合局部噪音。所以我们需要更多样化的特征去构建一个更加紧凑的更具有表现力的特征空间。
- 为了解决这些问题, 我们遵循 “attentive but diverse” 的原则设计神经网络。我们在较低的卷积层加入 self-channel attention 和 orthogonality constraint, 并且在较高的卷积层加入 Dual attention 模块。
- 在 Market 1501 数据集上的初步结果实现了 ~92.5% top-1 accuracy 和 ~81.0% mAP (基于 Densenet121 网络)。

非限制环境下人脸验证

2014 年秋天 - 2016 年春天

- 提出了一个新的 CNN 网络层用于提取从整体到局部的面部特征——包括整体面部的特征以及眼睛，鼻子和嘴部的局部特征。
- 将提出的 CNN 层集成在 Google 的 FaceNet 的框架中去提取面部特征，然后使用 Joint-Bayesian 人脸验证模型实现人脸验证。
- 这个算法在第三届智慧城市视频分析竞赛的人脸识别与验证项目中取得了第三名。

触屏手机输入法输入校正模型

2014 年 7 月 - 2014 年 9 月

- 设计并实现了一个非监督在线学习的算法用于学习特定人的输入习惯。这个算法自适应于特定的使用者，并且能够有效减少输入错误。
- 这个算法被集成到了微软必应手机输入法中。

工作经历

科研助理

2016 年秋季 - 现在

PSI 实验室，计算机科学与工程系，德州农工大学

导师：Ricardo Gutierrez-Osuna 博士

- 进行语音说话人转换，非本土口音转换，以及语音识别方向的研究。

科研助理

2014 年秋季 - 2016 年春季

人工智能和机器人研究所，西安交通大学

导师：王进军博士

- 进行人脸检测与验证方向的研究。

实习生

2014 年 7 月 - 2014 年 9 月

无线网络组，微软亚洲研究院，北京，中国

导师：沈国斌

- 进行触屏手机输入法校正模型的研究。

学术论文

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*.