Cheng-chieh Yeh

National Taiwan University Speech Processing Lab

r06942067@ntu.edu.tw acetylSv +886988601707

Education

National Taiwan University (NTU), MS in Data Science Group of Communication Engineering
GPA: 4.30/4.30 *2017 - 2019*
Advisors: Prof. Lin-shan Lee and Prof. Hung-yi Lee
Thesis: Multi-target Rhythm-flexible Voice Conversion without Parallel Data (Link)

National Taiwan University (NTU), BS in Computer Science
GPA (Last 60) : 3.73/4.30 *2013 - 2017*

Selected Course (A+, all graduate level):
Introduction to Digital Speech Processing, Natural Language Processing, Machine Learning Foundations, Machine Learning, Machine Learning and Having it Deep and Structured, Applied Deep Learning

Research Experiences

Non-parallel Voice Conversion
NTU Speech Processing Lab *2017 - present*
- Dealing with many-to-many speaker pairs conversion with only one model trained, as compared to conventional approaches with many one-to-one speaker pair models, (Publication [1], Award [c])
- Considering the flexibility of rhythmic patterns in voice conversion between speakers with very different speaking rates, (Publication [2])
- One-shot unsupervised voice conversion by separating speaker and content into two embeddings using Instance Normalization (IN) and Adaptive Instance Normalization (adaIN), (Publication [4])

Text-to-Speech
NTU Speech Processing Lab *2017 - present*
- Github open-sourced project with re-implementation of Google's end-to-end TTS system (GST Tacotron)
- Proposed an end-to-end TTS for low-resource languages by transfer learning, including a learnable Phonetic Transformation Network to map the input representations between different languages, (Publication [3])
- Trying to gain more controllability over unsupervised emotion and style modeling in TTS

Audio2Text Abstractive Machine Summarization
NTU Speech Processing Lab Undergraduate Research Project *2015 - 2017*
- Built a system which takes audio news stories and generates the headlines, including data collection, and proposing an abstractive machine summarization approach with reduced performance gap caused by ASR errors (Award [d])

Working Experiences

Research Internship
Siri Understanding, Apple Inc. *July 2019 - October 2019*
- Speaker modeling in end-to-end neural speech synthesis, seeking better unseen speaker adaptation results
- Try out different strategies on improving the stability and scalability of end-to-end systems

Teaching Experiences

Teaching Assistant *2015 - 2017*
- Digital Speech Processing Undergraduate Research Project (Instructor: Lin-shan, Lee), Machine Learning (Instructor: Hung-yi, Lee), Machine Learning and Having it Deep and Structured (Instructor: Hung-yi, Lee), Applied Deep Learning (Instructor: Yun-Nung, Chen)

Working Group Leader for a National Grand Challenge "Talk to AI" *June 2017 - September 2017*
- A Chinese spoken language understanding competition held by the Ministry of Science and Technology of Taiwan
- Leading a working group of graduate students offering technical support to the nation-wide Kaggle competition for undergraduate and high school students, including collecting data, sharing fundamental techniques and tools, preparing baselines for grading, etc

Publications

[1] **Multi-target Voice Conversion without Parallel Data by Adversarially Learning Disentangled Audio Representations**, J.-C. Chou, **C.-C. Yeh**, H.-Y. Lee, L.-S. Lee, *Interspeech 2018 (selected as one of the 12* ***BEST STUDENT PAPER FINALIST****, award [c])* (Paper link)

[2] **Rhythm-Flexible Voice Conversion without Parallel Data Using Cycle-GAN over Phoneme Posteriorgram Sequences**, **C.-C. Yeh**, P.-C. Hsu, J.-C. Chou, H.-Y. Lee, L.-S. Lee, *SLT 2018* (Paper link)

[3] **End-to-end Text-to-speech for Low-resource Languages by Cross-Lingual Transfer Learning**, Tao Tu, Yuan-Jui Chen, **C.-C. Yeh**, H.-Y. Lee, *Interspeech 2019* (Paper link)

[4] **One-shot Voice Conversion by Separating Speaker and Content Representations with Instance Normalization**, J.-C. Chou, **C.-C. Yeh**, H.-Y. Lee, *Interspeech 2019* (Paper link)

Awards and Grants

[a] **2019 The Phi Tau Phi Scholastic Honorable member**

[b] **2018 National Taiwan University Advanced Speech Technologies Scholarship**

- Grants ($16,000 USD) for students who excel in speech related technologies and researches

[c] **2018 Interspeech Best Student Paper Finalist: Non-parallel Voice Conversion (Publication [1])** (Paper link)

[d] **2017 NTU CSIE Undergraduate Research Competition: Audio2Text Summarization** (Poster link)

- Departmental **FIRST PRIZE**, also Appier second prize, Viscovery third prize, Leopard Mobile first prize