MIAO ZHANG

MASTER STUDENT, STANFORD UNIVERSITY

650-441-4465
miaoz18@stanford.edu
Palo Alto, California, USA

EDUCATION

Stanford University, CCRMA (Center for Computer Research in Music and Acoustics)
Music Science and Technology **(Present GPA: 3.9/4)** *Sep' 18 — Now*

- Courses: Programming Abstractions; Computer Organization & Systems; Applied Machine Learning. Audio Signal Processing; Spatial Sound; Symbolic Music Information;

Beijing University of Posts and Telecommunications, School of Electrical Engineering
Bachelor of Engineering, Electronic Engineering **(GPA: 3.85/4)** *Sep' 14 — July '18*

- Courses: C Programming and Assembly; C++ Programming; Data Structures; Digital Circuit Analysis Advanced Mathematics; Linear Algebra; Probability Theory; Digital Signal Processing

WORK EXPERIENCE & PROJECTS

CSLT (Center for Speech and Language Technology), Tsinghua University *May '17 - June '18*
Machine Learning Intern *Supervisor : Prof. Dong Wang*

I proposed machine speaker recognition (SRE) tasks based on human trivial events (e.g. cough, laugh, sniff), which are highly valuable in particular circumstances such as forensic examination, as they are less subjected to intentional change, so can be used to discover the genuine speaker from disguised speech.

- Iteration 1:
  - Developed an iOS app "smart recording" using Objective-C. Collected and constructed trivial events database (Github: ViVi_SRE), and released the data for public use.
- Iteration 2:
  - Built deep feature learning structure (Covolutional Time-delayed Nerual Network) with Python scripts based on Kaldi Toolkit. The model uses MFCC feature, trained with Fisher English database, tested with ViVi_SRE.
  - Test results analyzed by visualizing tool t-SNE show the rich speaker information within trivial events: The EER can reach 10%-14%, despite their extremely short durations (0.2-1.0 seconds), which greatly surpass the baseline i-vector system. All experiments are ran in Linux environment.
- Iteration 3:
  - Developed a website for human test purpose using HTML and PHP, and managed dataset with SQL.

Interactive Audio-visual Computer Games: Dancing Canon *October '18*
Instructor : Prof. Ge Wang

- Emphasized on interactive systems, aesthetics, and product design. Users can listen to canon harmonies by mouse or keyboard input.
- This interactive game was built on Unity using C# programming language.

Interactive Audio-visual Music Sequencer UI Design: Jump in Jungle *December '18*

- Boards with 9 different heights. And users can control jumping to play notes by keyboard input.
- This interface was built on OpenGl and Unity using C++ programming language.

PUBLICATIONS

Miao Zhang, Yixiang Chen, Lantian Li and Dong Wang; "Speaker Recognition with Cough, Laugh and "Wei"". APSIPAP ASC 2017.

Miao Zhang, Xiaofei Kang, Yanqing Wang, Lantian Li, Zhiyuan Tang, Haisheng Dai, Dong Wang*; "Human and machine speaker recognition based on short trivial events". ICASSP | IEEE 2018.

AWARDS & ACHIEVEMENTS

YOFC Enterprise Scholarship, 2016 (Rank, **2/90**)
First Class Scholarship, 2017(Rank, **3/90**)
Honorable award in The Mathematical Contest in Modeling, 2016

TECHNICAL SKILLS

Programming languages: C, C++, Python, Java, Html, Css, Matlab, Php, Assembly Language
Software & Tools: Qt, Git, Linux, kaldi, Latex