Yuxuan Wu

@ Yuxuan.Wu@mbzuai.ac.ae | | Irislucent | BowOfAtlas

Education

Mohamed bin Zayed University of Artificial Intelligence — Abu Dhabi, UAE
Ph.D in Machine Learning; ***GPA: N/A*** — *Aug 2023 – Now*

Carnegie Mellon University — Pittsburgh, USA
M.Sc in Music and Technology; ***GPA: 3.91/4.00*** — *Aug 2021 – Aug 2023*

Shanghai Jiao Tong University — Shanghai, China
B.Eng in Information Engineering; ***GPA: 3.55/4.30*** — *Sep 2017 – Jun 2021*

Selected Research Experience

Carnegie Mellon University — Pittsburgh, USA
Research Assistant — *Aug 2021 – Present*

Motif-based Music Representation Learning — *Feb 2023 – Aug 2023*

- Designed and implemented algorithms for better encoding of music motifs. Experiments demonstrated the effectiveness of the proposed method compared to baselines. Also, an intuitive visualization of music structure based on motif representations is conducted.
- Designed and collected data for model training and evaluation.

Jazz Improvisation Tutoring with Generative Models — *Aug 2022 – April 2023*

- Designed and implemented the deconstruction functions for jazz master MIDI data, which degenerate data in the musical sense to enable self-supervised learning.
- Participated in designing and conducting user behavior experiments.

Music Timbre Transfer with Flexible Timbre Control — *Feb 2022 – May 2022*

- Used a concisely structured Autoencoder-based model to achieve music timbre transfer among multiple instruments with a single model. The Autoencoder functions on CQT features and the waveform is generated by a Diffwave vocoder trained on music. The model performance is comparable to state-of-the-art one-to-one timbre transfer method based on GAN on certain transferring pairs.

Lab of Multimedia Intelligence, Shanghai Jiao Tong University — Shanghai, China
Undergraduate Research Assistant — *Jul 2020 – June 2021*

Client Software for Speech Synthesis and Splicing

- Built a cross-platform client software for speech synthesis and splicing. It contains an interface to call Chinese text-to-speech models on remote servers, and tools for voice activity detection, voice splicing and WSOLA-based time stretching and pitch shifting. This software has been put into actual use in telemarketing by a listed company.

Lab of Audio and Music Technology, Fudan University — Shanghai, China
Undergraduate Research Assistant — *Aug 2019 – June 2021*

Vocal Register Recognition with Deep Learning

- Recorded and labelled the dataset for voice register recognition. Used a 1D-CNN model with Mel-spectrograms as the inputs, and reached an accuracy of 99.15% on the self-constructed dataset, which surpassed the performances of previous feature engineering methods.

Working Experience

Machine Learning Department, MBZUAI — Abu Dhabi, UAE
Teaching Assistant — *Jan 2024 – May 2024*

- Teaching Assistant of ML711: *Intermediate Music AI*

School of Computer Science, Carnegie Mellon University — Pittsburgh, USA
Teaching Assistant — *Aug 2022 – Now 2022*

- Teaching Assistant of 11755/18797: *Machine Learning for Signal Processing*
- Teaching Assistant of 15322/15622: *Introduction to Computer Music*

GEC Academy Beijing, China

Teaching Fellow *Jul 2023 – Mar 2024*

* Teaching Fellow of *Introduction to Computer Music*

Speech, Audio and Music Intelligence, TikTok, ByteDance Inc. Santa Clara, USA

Research Intern *May 2022 – Aug 2022*

* Designed and implemented an algorithm to generate music accompaniment texture on given chord progressions, according to the midi template library. The algorithm supports the generation of non-chord tones and can be applied to a variety of music situations.
* Implemented the support for bus effects in the in-house music rendering system.
* Built a versatile off-line software for quickly testing the in-house music rendering system.

AI Lab Speech & Audio, ByteDance Inc. Shanghai, China

Musician Intern *Jul 2020 – Jul 2021*

* Designed and implemented a multi-track music texture generator: The program generates stylized multi-track MIDI according to users' input of chord progressions and styles using Markov models. Music created under its assistance have been played for thousands of times in TikTok.
* Improved the performance of AI melody generation models: Applied hierarchical generation structure to the Conditional-VAE melody generation model with prior music knowledge. Reduced the manual participation rate in AI composing by over 20%.
* Built multiple stylized music arrangement templates. The templates use Reaper API scripts to call the music generation library and render music audio.
* Conducted studio monitoring for text-to-speech data collection.

Publications & Patents

Shuqi Dai, Siqi Chen, **Yuxuan Wu**, Ruxin Diao, Roy Huang, Roger B. Dannenberg: "SingStyle111: A Multilingual Singing Dataset with Style Transfer", ISMIR 2023

Yuxuan Wu, Yifan He, Xinlu Liu, Yi Wang, Roger B. Dannenberg: "TransPlayer: Timbre Style Transfer with Flexible Timbre Control", ICASSP 2023 (Oral)

Ruibin Yuan, **Yuxuan Wu**, Jacob Li, Jaxter Kim: "DeID-VC: Speaker De-identification via Zero-shot Pseudo Voice Conversion", INTERSPEECH 2022

Xia Liang, Yuan Wan, **Yuxuan Wu**, Bilei Zhu, Zejun Ma: "A Method for Music Texture Generation Based on Markov Chain", China Patent (Pending), 2022

Awards & Achievements

AI Song Contest 2022: Our team *3+i* won the **3rd Place** with our song "A to I"

Music Exhibition at CSMT 2020: Xinyue Zhu and I won the **2nd prize** with our Max immersive experience sound system "Project Ming"

The 3rd Think Youth Digital Creation Competition: 3rd prize with the song "Feather"

The 22nd SJTU Annual Singer Contest: 3rd place

SJTU Merit Student: 2018