

# ZHENYU TANG

Santa Clara, CA 95051 ♦ zhy@cs.umd.edu ♦ <https://github.com/RoyJames>

## EDUCATION

### **University of Maryland - College Park**

PhD in Computer Science, supervised by **Dinesh Manocha**

Received Dean's Fellowship in 2018 and 2019, GPA: 3.7/4.0

College Park, MD, USA

2018 – 2022

### **Zhejiang University** (Chu Kochen Honors College)

Bachelor in Engineering (**with Honor**), Opto-Electronic Science and Engineering

Cumulative GPA: 3.73/4.0, Major GPA: 3.83/4.0 (top 5%)

Hangzhou, Zhejiang, China

2013 – 2017

**Research interests:** large speech-language models, speech/music synthesis & analysis, generative AI

## PROFESSIONAL EXPERIENCE

### **Meta (formerly Facebook)**

*Senior Research Scientist in GenAI*

Menlo Park, CA

Apr. 2022 – Present

### **TikTok (ByteDance)**

*Research Scientist in Speech, Audio & Music Intelligence*

San Jose, CA

June 2022 – Apr. 2024

- Develop algorithms for large audio/speech dataset curation, annotation, and quality control
- Research in large speech models for conversational/expressive text-to-speech applications
- Research in generative AI models (e.g., Diffusion, Flow, GAN) for speech/audio, topics including speech/singing voice conversion, generative speech enhancement, timbre+accent editing

### **Amazon Lab126**

*Research Intern in Audio Data Engineering*

San Jose, CA

July 2021 – Oct. 2021

- Propose novel methods for generating high-quality acoustic training data

### **Meta Reality Labs**

*Research Intern in Audio Team*

Redmond, WA

May 2020 – Oct. 2020

- Developed efficient algorithms for acoustic simulation with custom microphone arrays
- Built pipeline for training speech models with synthetic spatial audio data (one co-authored publication)

### **Adobe Systems**

*Creative Intelligence Lab Intern in Audio Group*

Seattle, WA

May 2019 – Sept. 2019

- Devised novel user-friendly methods for synthesizing realistic virtual sound in augmented reality
- Integrated our learning-based acoustic analyzer as a service in Adobe's Sensei cloud AI framework

## ACADEMIC SERVICES

I served as a meta-reviewer in the Senior Program Committee (SPC) for the **AAAI 2023&2024 Conference on Artificial Intelligence**. I also served as a reviewer for

- International Symposium on Mixed and Augmented Reality (ISMAR) Journal track 2022, 2023, 2024
- ACM Multimedia 2023
- Eurographics 2022
- SIGGRAPH 2020, SIGGRAPH Asia 2021
- The IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR) 2020 and 2021
- Transactions on Audio, Speech and Language Processing (TASLP) 2021
- Acta Acustica Journal 2020, 2021
- IEEE Transactions on Multimedia 2020
- Journal of the Acoustical Society of America (JASA) 2020

## **TECHNICAL STRENGTHS**

**Programming:** C/C++, Python, Matlab, R, bash scripting

**Software and Tools:** Docker/Singularity, Pytorch, TensorFlow, Blender, pybind11

## **SELECTED PUBLICATIONS**

***MESH2IR: Neural Acoustic Impulse Response Generator for Complex 3D Scenes***

Anton Ratnarajah, **Zhenyu Tang**, Rohith Aralikatti, Dinesh Manocha  
*Proceedings of the 2022 ACM on Multimedia Conference*. ACM, 2022

***GWA: A Large High-Quality Acoustic Dataset for Audio Processing***

**Zhenyu Tang**, Rohith Aralikatti, Anton Ratnarajah, Dinesh Manocha  
*SIGGRAPH 2022 Conference Proceedings*

***FAST-RIR: Fast Neural Diffuse Room Impulse Response Generator***

Anton Ratnarajah, Shi-Xiong Zhang, Meng Yu, **Zhenyu Tang**, Dinesh Manocha, Dong Yu  
*International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2022*

***IR-GAN: Room Impulse Response Generator for Speech Augmentation***

Anton Ratnarajah, **Zhenyu Tang**, Dinesh Manocha  
*INTERSPEECH 2021*

***Learning Acoustic Scattering Fields for Dynamic Interactive Sound Propagation***

**Zhenyu Tang**, Hsien-Yu Meng, Dinesh Manocha  
*IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR) 2021*

***Low-frequency Compensated Synthetic Impulse Responses for Improved Far-field Speech Recognition***

**Zhenyu Tang**, Hsien-Yu Meng, Dinesh Manocha  
*International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2020*

***Improving Reverberant Speech Training using Diffuse Acoustic Simulation***

**Zhenyu Tang**, Lianwu Chen, Bo Wu, Dong Yu, Dinesh Manocha  
*International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2020*

***Scene-Aware Audio Rendering via Deep Acoustic Analysis***

**Zhenyu Tang**, Nicholas J. Bryan, Dingzeyu Li, Timothy R. Langlois, Dinesh Manocha  
*IEEE VR 2020 Journal, Transactions on Visualization and Computer Graphics (TVCG)*

***Regression and Classification for Direction-of-Arrival Estimation with Convolutional Recurrent Neural Networks***

**Zhenyu Tang**, John D. Kanu, Kevin Hogan, Dinesh Manocha  
*INTERSPEECH 2019*

Full publication list at <https://scholar.google.com/citations?user=gPGVGtKAAAAJ&hl=en&oi=ao>