Philip Anastassiou

www.philipanastassiou.com | philip.anastassiou@pm.me |

Education

Columbia University — New York, NY
M.S. Computer Science (Machine Learning Track) — Dec. 2022

- Coursework included: artificial intelligence, machine learning, advanced topics in neural networks and deep learning, advanced topics in spoken language processing and speech recognition, natural language processing, computer vision, databases, and analysis of algorithms.
- Conducted research advised by Professor Peter Belhumeur on developing self-supervised optical character recognition model whose training procedure is guided by separate language model via student-teacher knowledge distillation paradigm without use of ground-truth labels.

B.A. Drama and Theatre Arts — May 2018

Experience

TikTok/ByteDance — San Jose, CA
Speech Synthesis Engineer — Dec. 2022 – Present
Software Engineer Intern — May 2022 – Nov. 2022

- Conduct research on Data-Speech team (previously Speech, Audio, and Music Intelligence) on state-of-the-art neural speech synthesis and audio processing algorithms for various TikTok and ByteDance products to empower content creation and consumption.

Research

My research interests include deep generative modeling, self-supervised representation and transfer learning, knowledge distillation, and digital signal processing. I am especially interested in neural end-to-end learning for audio and natural language processing tasks.

Publications

Philip Anastassiou*, Zhenyu Tang*, Kainan Peng, Dongya Jia, Jiaxin Li, Ming Tu, Yuping Wang, Yuxuan Wang, Mingbo Ma (*equal contribution), *"VoiceShop: A Unified Speech-to-Speech Framework for Identity-Preserving Zero-Shot Voice Editing,"* ArXiv, 2024. [Paper, Demo]

Patents

Philip Anastassiou, Zhenyu Tang, Jiaxin Li, Kainan Peng, Dongya Jia, Qiao Tian, Mingbo Ma, Yuping Wang, Yuxuan Wang, *"Identity-Preserving Zero-Shot Many-to-Many Accent and Speech Style Conversion via Bottleneck-to-Bottleneck and Diffusion-based Modeling,"* pending CN patent application filed by ByteDance Ltd., 2024.

Presentations

Nokia Bell Labs — Murray Hill, NJ
Research Presentation ("Squashed" Software) — Feb. 2019

- Developed software for Nokia Bell Labs Experiments in Art and Technology program in collaboration with Columbia Computer Music Center to produce musically desirable digital artifacts in audio signals with lossy LAME MP3 encoder.

Lamont-Doherty Earth Observatory — Palisades, NY
Research Presentation ("Novel Song" Software) — May 2018

- Developed software to convert text of fictional novels into music with VADER Python package for sentiment analysis to interpolate emotional valence scores into harmonic relationships in RTcmix, developed at Columbia Computer Music Center and presented at annual *Research as Art* exhibition at Lamont-Doherty Earth Observatory.

Projects

VAE-GAN for Speech-to-Speech Style Transfer — Dec. 2021

- Implemented proposed variational autoencoder-generative adversarial network (VAE-GAN) architecture with domain-specific decoders for non-autoregressive speech-to-speech style transfer based on AlBadawy, et al. (2020) and Bonnici, et al. (2021). [Code]

Deep Convolutional Spectral Autoencoder for Neural Audio Synthesis — May 2021

- Implemented deep convolutional autoencoder in Python using TensorFlow for audio synthesis of musical notes based on Engel, et al. (2017), trained on subset of Google Magenta's *NSynth* corpus.

Skills

Languages: Python, Java, C, MATLAB, SQL, Unix shell scripting
Software: Git, PyTorch, PyTorch Lightning, TensorFlow, Keras, Pandas, SciPy, NumPy, Matplotlib, Scikit-Learn, FFmpeg, Librosa, Kaldi, ESPNet, Praat, NLTK, Conda, LaTeX