Prem Seetharaman

Contact and Info
954-558-3546
56seeth@gmail.com
pseeth.github.io
github.com/pseeth

Research interests

Lowering the barrier to entry for creative work in video and audio by leveraging and developing state-of-the-art generative modeling techniques. My research interests are primarily in computer audition, with a focus on generative audio modeling (e.g. speech synthesis, music generation), audio representation learning, audio compression, and audio-to-audio mapping (e.g. speech/music separation and enhancement).

Education

Northwestern University, Evanston, IL
09/2019 - PhD, Computer Science

Northwestern University, Evanston, IL
BS, Computer Science, Music Composition - 2013

Professional employment

Descript, Inc.
Sr. Research Scientist: 2020 - Present

- Led Studio Sound - state-of-the-art speech enhancement.
  https://www.descript.com/studio-sound
- Core contributor to Overdub - voice cloning for speech editing.
  https://www.descript.com/overdub
- Developed the Descript Audio Codec - high-fidelity audio compression, which powers internal audio language models for editing speech and generating music.
  https://github.com/descriptinc/descript-audio-codec

Northwestern University, Evanston, IL
Postdoctoral Scholar: 2019 - 2020

Northwestern University, Evanston, IL
Doctoral Student in Interactive Audio Lab: 2013 - 2019

Mitsubishi Electric Research Labs, Cambridge, MA
Research Intern: 2018

Adobe Research, San Francisco, CA
Research Intern: 2017-2018

Gracenote, Emeryville, CA
Research Intern: 2016

Publications

Thesis

[1] **Seetharaman, Prem**. "Bootstrapping the Learning Process for Computer Audition". PhD thesis. Northwestern University, 2019.

Patents

[1] Prem Seetharaman, Gautham J Mysore, and Bryan A Pardo. *Sound Quality Prediction and Interface to Facilitate High-Quality Voice Recordings*. US Patent App. 16/296,122. Sept. 2020.

[2] Markus K Cremer, Zafar Rafii, Robert Coover, and Prem Seetharaman. *Automated Cover Song Identification*. US Patent App. 15/698,557. July 2018.

[3] Zafar Rafii and Prem Seetharaman. *Audio Identification Based on Data Structure*. US Patent App. 15/698,532. Mar. 2018.

Journal Articles

[1] Vivian Tang, Prem Seetharaman, Kevin Chao, Bryan A Pardo, and Suzan Van Der Lee. "Automating the detection of dynamically triggered earthquakes via a deep metric learning algorithm". In: *Seismological Research Letters* 91.2A (2020), pp. 901–912.

[2] Bryan Pardo, Mark Cartwright, Prem Seetharaman, and Bongjun Kim. "Learning to build natural audio production interfaces". In: *Arts*. Vol. 8. 3. MDPI. 2019, p. 110.

[3] Eric J Humphrey, Sravana Reddy, Prem Seetharaman, Aparna Kumar, Rachel M Bittner, Andrew Demetriou, Sankalp Gulati, Andreas Jansson, Tristan Jehan, Bernhard Lehner, et al. "An introduction to signal processing for singing-voice analysis: High notes in the effort to automate the understanding of vocals in music". In: *IEEE Signal Processing Magazine* 36.1 (2018), pp. 82–94.

[4] Prem Seetharaman and Bryan Pardo. "Audealize: Crowdsourced Audio Production Tools". In: *Journal of the Audio Engineering Society* 64.9 (2016), pp. 683–695.

Papers

[1] Hugo Flores Garcia, Prem Seetharaman, Rithesh Kumar, and Bryan Pardo. "VampNet: Music Generation via Masked Acoustic Token Modeling". In: *Proceedings of the 24th International Society for Music Information Retrieval Conference*. 2023.

[2] Rithesh Kumar, Prem Seetharaman, Alejandro Luebs, Ishaan Kumar, and Kundan Kumar. "High-Fidelity Audio Compression with Improved RVQGAN". In: *arXiv preprint arXiv:2306.06546* (2023).

[3] Noah Schaffer, Boaz Cogan, Ethan Manilow, Max Morrison, Prem Seetharaman, and Bryan Pardo. "Music Separation Enhancement with Generative Modeling". In: (2022).

[4] Ho Hsiang Wu, Magdalena Fuentes, Prem Seetharaman, and Juan Pablo Bello. "How to Listen? Rethinking Visual Sound Localization". In: *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. Vol. 2022. 2022, pp. 876–880.

[5] Ho-Hsiang Wu, Prem Seetharaman, Kundan Kumar, and Juan Pablo Bello. "Wav2clip: Learning robust audio representations from clip". In: *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2022, pp. 4563–4567.

[6] Ethan Manilow, Patrick O'Reilly, Prem Seetharaman, and Bryan Pardo. "Source separation by steering pretrained music models". In: *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2022, pp. 126–130.

[7] Max Morrison, Rithesh Kumar, Kundan Kumar, Prem Seetharaman, Aaron Courville, and Yoshua Bengio. "Chunked autoregressive GAN for conditional waveform synthesis". In: *International Conference on Learning Representations*. 2021.

[8] Nicolas Turpault, Romain Serizel, Scott Wisdom, Hakan Erdogan, John R Hershey, Eduardo Fonseca, Prem Seetharaman, and Justin Salamon. "Sound event detection and separation: a benchmark on desed synthetic soundscapes". In: *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2021, pp. 840–844.

[9] Andreas Bugler, Bryan Pardo, and Prem Seetharaman. "A study of transfer learning in music source separation". In: *arXiv preprint arXiv:2010.12650* (2020).

[10] Prem Seetharaman, Gordon Wichern, Bryan Pardo, and Jonathan Le Roux. "AutoClip: Adaptive Gradient Clipping for Source Separation Networks". In: *2020 IEEE 30th International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE. 2020, pp. 1–6.

[11] Omkar Ranadive, Grant Gasser, David Terpay, and Prem Seetharaman. "Otoworld: Towards learning to separate by learning to move". In: *ICML 2020 Workshop on Self-supervision in Audio and Speech*. 2020.

[12] Scott Wisdom, Hakan Erdogan, Daniel PW Ellis, Romain Serizel, Nicolas Turpault, Eduardo Fonseca, Justin Salamon, Prem Seetharaman, and John R Hershey. "What's all the fuss about free universal

sound separation data?" In: *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2021, pp. 186–190.

[13] Nicolas Turpault, Scott Wisdom, Hakan Erdogan, John R Hershey, Romain Serizel, Eduardo Fonseca, Prem Seetharaman, and Justin Salamon. "Improving Sound Event Detection In Domestic Environments Using Sound Separation". In: *DCASE Workshop-Detection and Classification of Acoustic Scenes and Events 2020*. 2020.

[14] Alisa Liu, Alexander Fang, Gaëtan Hadjeres, Prem Seetharaman, and Bryan Pardo. "Incorporating Music Knowledge in Continual Dataset Augmentation for Music Generation". In: *ML4MD Workshop at ICML 2020*. 2020.

[15] Alexander Fang, Alisa Liu, Prem Seetharaman, and Bryan Pardo. "Bach or mock? a grading function for chorales in the style of JS Bach". In: *ML4MD Workshop at ICML 2020*. 2020.

[16] Prem Seetharaman, Gordon Wichern, Jonathan Le Roux, and Bryan Pardo. "Bootstrapping unsupervised deep music separation from primitive auditory grouping principles". In: *ICML 2020 Workshop on Self-supervision in Audio and Speech*. 2020.

[17] Ethan Manilow, Prem Seetharaman, and Bryan Pardo. "Simultaneous separation and transcription of mixtures with multiple polyphonic and percussive instruments". In: *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2020, pp. 771–775.

[18] Fatemeh Pishdadian, Bongjun Kim, Prem Seetharaman, and Bryan Pardo. "Classifying Non-Speech Vocals: Deep vs Signal Processing Representations". In: 2019.

[19] Ethan Manilow, Gordon Wichern, Prem Seetharaman, and Jonathan Le Roux. "Cutting Music Source Separation Some Slakh: A Dataset to Study the Impact of Training Data Quality and Quantity". In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2019.

[20] Prem Seetharaman. "Bootstrapping the Learning Process for Computer Audition". PhD thesis. Northwestern University, 2019.

[21] Alisa Liu, Prem Seetharaman, and Bryan Pardo. "Model Selection for Deep Audio Source Separation via Clustering Analysis". In: 2020.

[22] Prem Seetharaman, Gautham Mysore, Bryan Pardo, Paris Smaragdis, and Celso Gomes. "VoiceAssist: Guiding Users to High-Quality Voice Recordings". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2019.

[23] Prem Seetharaman, Gordon Wichern, Shrikant Venkataramani, and Jonathan Le Roux. "Class-Conditional Embeddings for Music Source Separation". In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019*. 2018.

[24] Prem Seetharaman, Gordon Wichern, Jonathan Le Roux, and Bryan Pardo. "Bootstrapping Single-Channel Source Separation via Unsupervised Spatial Clustering on Stereo Mixtures". In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019*. 2018.

[25] Ethan Manilow, Prem Seetharaman, Fatemah Pishdadian, and Bryan Pardo. *NUSSL: the northwestern university source separation library*. 2018.

[26] Julia Wilkins, Prem Seetharaman, Alison Wahl, and Bryan Pardo. "VocalSet: A Singing Voice Dataset". In: *Proceedings of the 19th International Society for Music Information Retrieval Conference*. 2018.

[27] Prem Seetharaman, Gautham J Mysore, Paris Smaragdis, and Bryan Pardo. "Blind Estimation of the Speech Transmission Index for Speech Quality Prediction". In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018*. 2018.

[28] Prem Seetharaman, Fatemeh Pishdadian, and Bryan Pardo. "Music/Voice Separation Using the 2D Fourier Transform". In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2017.

[29] Ethan Manilow, Prem Seetharaman, Fatemeh Pishdadian, and Bryan Pardo. "Predicting Algorithm Efficacy for Adaptive Multi-Cue Source Separation". In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2017.

[30] Prem Seetharaman and Zafar Rafii. "Cover Song Identification with 2D Fourier Transform Sequences". In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017*. 2017.

[31] Taylor Zheng, Prem Seetharaman, and Bryan Pardo. "SocialFX: Studying a Crowdsourced Folksonomy of Audio Effects Terms". In: *Proceedings of the 2016 ACM on Multimedia Conference*. ACM. 2016, pp. 182–186.

[32] Prem Seetharaman and Bryan Pardo. "Simultaneous Separation and Segmentation in Layered Music". In: *Proceedings of the 17th International Society for Music Information Retrieval Conference*. 2016, pp. 495–502.
[33] Prem Seetharaman and Bryan Pardo. "Crowdsourcing a Reverberation Descriptor Map". In: *Proceedings of the 22nd ACM international conference on Multimedia*. ACM. 2014, pp. 587–596.
[34] Prem Seetharaman and Stephen P Tarzia. "The Hand Clap as an Impulse Source for Measuring Room Acoustics". In: *Audio Engineering Society Convention 132*. Audio Engineering Society. 2012.

Extended Abstracts

[1] Vivian Tang, Prem Seetharaman, Kevin Chao, Bryan A Pardo, and Suzan van der Lee. "Siamese networks for triggered earthquakes detection". In: *AGU Fall Meeting Abstracts*. Vol. 2018. 2018, S11E–0428.
[2] Kevin Chao, Prem Seetharaman, Vivian Tang, Bryan A Pardo, and Suzan Van der Lee. "Automatic classification of triggered tectonic tremor with deep learning". In: *AGU Fall Meeting Abstracts*. Vol. 2018. 2018, T33E–0463.
[3] Prem Seetharaman and Bryan Pardo. "Reverbalize: A Crowdsourced Reverberation Controller". In: *Proceedings of the 22nd ACM international conference on Multimedia*. ACM. 2014, pp. 739–740.
[4] Michael Donovan, Prem Seetharaman, and Bryan Pardo. "A Web Audio Node for the Fast Creation of Natural Language Interfaces for Audio Production". In: (2017).

References

Bryan Pardo
Full Professor
Department of Computer Science, Northwestern University
Evanston, IL
847.491.7184
pardo@northwestern.edu

Ken Forbus
Walter P. Murphy Professor of Computer Science
Department of Computer Science, Northwestern University
847.491.7699
forbus@northwestern.edu

Paris Smaragdis
Associate Professor
Department of Computer Science, University of Illinois at Urbana Champaign
217.265.6893
paris@illinois.edu

Sara Sood
Chookaszian Family Teaching Professor, Northwestern University
847.491.5708
sara@northwestern.edu

Gautham Mysore
Principal Scientist, Adobe Research
Adjunct Professor, Stanford University
650.353.1142
gmysore@adobe.com

Jonathan Le Roux
Senior Principal Research Scientist, Mitsubishi Electric Research Laboratories (MERL)
617.621.7547
leroux@merl.com