# Bochen Li

https://bochenli.github.io
Google Scholar Profile

+1(917)-969-5428
bochen1106@gmail.com


## Interests

- Music information retrieval, interactive music design, multi-modal scene analysis

## Education

- **University of Rochester** — Rochester, NY, USA
  PhD, Electrical and Computer Engineering (ECE) — *October 2020*
  Master of Science, Electrical and Computer Engineering (ECE) — *May 2016*
- **University of Science and Technology of China** — Hefei, China
  Bachelor of Science, Electronic Engineering and Information Science — *June 2014*

## Research/Industrial Experience

- **TikTok** - *Intelligent Creation Audio* — Mountain View, CA, USA
  *Tech Lead* — *December 2019 - Present*
  - Main developer and tech lead for music creation algorithms to support Ripple, an AI-powered music creation App released in 2023.
  - Developing multi-modal music effects in TikTok to promote content consumption among 1+ billion users.

- **University of Rochester** - *Audio Information Research Lab* — Rochester, NY, USA
  *Research Assistant* — *September 2014 - August 2020*
  - **PhD Thesis: Multi-Modal Analysis for Music Performance**:
    - Created two audiovisual music datasets, URMP and URSing, and addressed association problem of multiple music modalities.
    - Applied multi-modal analysis for traditional music information retrieval tasks and new frontiers of emerging research topics, e.g., visually-informed multi-pitch analysis, source separation, and cross-modal localization/retrieval/generation.

- **Bytedance AI Lab** - *Speech, Audio, and Music Intelligence Team* — Palo Alto, CA, USA
  *Research Intern* — *February 2019 - May 2019*
  - Proposed the audio-visual singing separation algorithm.

- **Spotify** - *Music Intelligence Team* — New York, NY, USA
  *Research Intern* — *June 2018 - August 2018*
  - Music-query by video. Developed a two-stream network to learn the cross-modal distance between music and unconstrained videos via latent emotion space, which includes audio/video emotion tagging branches and cross-modal distance learning framework. The model recommends a music/playlist given user-uploaded video clip.

- **Yamaha** - *Music AI Team* — Hamamatsu, Shizuoka, Japan
  *Research Intern* — *October 2017 - December 2017*
  - Visual performance generation. Developed a system to learn the music context of the given MIDI data (music score) and generate expressive whole-body visual performance as pianist skeleton key points, using convolutional and recurrent neural networks.

- **Knowles Intelligent Audio** - *Speech Interface Team* — Mountain View, CA, USA
  *Intern* — *May 2017 - August 2017*
  - Performed the keyword spotting process including data augmentation, training, parameter tuning, and testing.
  - Developed the framework for talker ID recognition based on the Gaussian mixture models (GMM).

## Teaching Experience

- **Academic Tutorials**
  - "Audio-visual Music Processing" given at the *ISMIR2019* conference. *Fall 2018*
- **Guest Lectures**
  - "Intro to Music Information Retrieval and Industrial Applications", *University of Wisconsin - Stout.* *Fall 2020*
  - "Audio-visual analysis for music performance", *University of Rochester.* *Fall 2018*
  - "Machine learning for audio signal processing", *University of Rochester.* *Spring 2018*
- **Teaching Assistant**
  - "Audio Signal Processing", *University of Rochester* *Spring 2018*
  - "Music and Math", for pre-college students from the *Upward Bound Program.* *Summer 2016*
  - "Circuits and Signals", *University of Rochester* *Spring 2015*
  - "Intro to C/C++ Programming", *University of Rochester* *Fall 2014*

## Publications

- Qiuqiang Kong, **Bochen Li**, Jitong Chen, and Yuxuan Wang, "GiantMIDI-Piano: a large-scale MIDI dataset for classical piano music," *Transactions of the International Society for Music Information Retrieval*, vol. 5, no. 1, pp.87-98, 2022.
- **Bochen Li**, Yuxuan Wang, and Zhiyao Duan, "Audiovisual singing voice separation", *Transactions of the International Society for Music Information Retrieval*, vol. 4, no. 1, pp.195-209, 2021.
- Qiuqiang Kong, **Bochen Li**, Xuchen Song, Yuan Wan, and Yuxuan Wang, "High-resolution piano transcription with pedals by regressing onsets and offsets times," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp.3707-3717, 2021.
- **Bochen Li**, Karthik Dinesh, Chenliang Xu, Gaurav Sharma, and Zhiyao Duan, "Online audio-visual source association for chamber music performances," *Transactions of the International Society for Music Information Retrieval*, vol. 2, no. 1, pp.29-42, 2019.
- **Bochen Li** and Aparna Kumar, "Query by video: cross-modal music retrieval," in *Proc. International Society for Music Information Retrieval (ISMIR)*, 2019.
- **Bochen Li***, Xinzhao Liu*, Karthik Dinesh, Zhiyao Duan, and Gaurav Sharma, "Creating a musical performance dataset for multimodal music analysis: challenges, insights, and applications," *IEEE Transactions on Multimedia*, vol. 21, no. 2, pp. 522-535, 2019. *(* Equal contribution)*
- **Bochen Li**, Akira Maezawa, and Zhiyao Duan, "Skeleton plays piano: online generation of pianist body movements from MIDI performance," in *Proc. International Society for Music Information Retrieval (ISMIR)*, 2018.
- **Bochen Li** and Akira Maezawa, "MIDI2Pose: online keyboard performance motion generation from performance data," in *Proc. Information Processing Society of Japan*, 2018.
- Yapeng Tian, Jing Shi, **Bochen Li**, Zhiyao Duan, and Chenliang Xu, "Audio-visual event localization in unconstrained videos," in *Proc. European Conference on Computer Vision (ECCV)*, 2018.
- Xueyang Wang, Ryan Stables, **Bochen Li**, and Zhiyao Duan, "Score-aligned polyphonic microtiming estimation," in *Proc. International Conference on Audio, Speech and Signal Processing (ICASSP)*, 2018.
- **Bochen Li**, Karthik Dinesh, Gaurav Sharma, and Zhiyao Duan, "Video-based vibrato detection and analysis for polyphonic string music," in *Proc. International Society for Music Information Retrieval (ISMIR)*, 2017. (**Best Paper Nomination**)
- **Bochen Li**, Chenliang Xu, and Zhiyao Duan, "Audio-visual source association for string ensemble videos through multi-modal vibrato analysis," in *Proc. Sound and Music Computing Conference*, 2017. (**Best Paper Award**)

- **Bochen Li**, Karthik Dinesh, Zhiyao Duan, and Gaurav Sharma, "See and listen: score-informed association of sound tracks to players in chamber music performance videos," in *Proc. International Conference on Audio Speech and Signal Processing (ICASSP)*, 2017.
- Karthik Dinesh*, **Bochen Li***, Xinzhao Liu, Zhiyao Duan, and Gaurav Sharma, "Visually informed multi-pitch analysis of string ensembles," in *Proc. International Conference on Audio, Speech and Signal Processing (ICASSP)*, 2017. *(* Equal contribution)*
- **Bochen Li** and Zhiyao Duan, "An approach to score following for piano performances with sustained effect," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 12, 2016.
- **Bochen Li** and Zhiyao Duan, "Score following for piano performances with sustain-pedal effects," in *Proc. International Society for Music Information Retrieval (ISMIR)*, 2015.

## PATENTS

- **Bochen Li**, Vibert Thio, Haonan Chen, Xuefan Hu, and Jitong Chen, "Approach to automatic music remix based on style templates," Publication of *US20230360619A1*, November 2023.
- Vibert Thio, **Bochen Li**, Haonan Chen, and Jitong Chen, "Automatic and interactive mashup system," Publication of *US20230360618A1*, November 2023.
- **Bochen Li**, Andrew Shaw, and Jitong Chen, "Converting audio samples to full song arrangements," Publication of *WO2023214937A1*, November 2023.
- **Bochen Li**, Rodrigo Castellon, Daiyu Zhang, and Jitong Chen, "Beatboxing transcription," Publication of *US 20230282188A1*, September 2023.
- Zhihao Ouyang, **Bochen Li**, Daiyu Zhang, "Automatic and fast generation of music audio content for videos," Publication of *US11763849B1*, September 2023.
- **Bochen Li**, Daiyu Zhang, Shawn Chan, and Jitong Chen, "Interactive movement audio engine," Publication of *US20230197040A1*, June 2023.
- Shuai Yuan, **Bochen Li**, Qiuhong Xu, Na Zhao, Zhengyi Fang, Peidao Li, and Shengli Wang, "Method and device for determining audio frequency, electronic equipment and storage medium," Publication of *CN115831080A*, March 2023.
- **Bochen Li** and Aparna Kumar, "Systems, methods & computer program products for associating media content having different modalities," Publication of *US20200394213A1*, December 2020.
- Akira Maezawa and **Bochen Li**, "Information processing method," Publication of *US20200365126A1*. November 2020.

## Awards

- **Outstanding PhD Dissertation Award**, *University of Rochester* *June, 2021*
- **Best Paper Nomination**, *18th International Society for Music Information Retrieval (ISMIR)* *October, 2017*
- **Best Paper Award**, *14th Sound and Music Computing Conference (SMC)* *July, 2017*

## Academic Service

- Committee members
  - Music program chair, *International Society for Music Information Retrieval*, 2021
  - Technical chair, *North East Music Information Special Interest Group*, 2017
- Reviewer for journals
  - *Journal of New Music Research*

- *The Journal of the Acoustical Society of America*
- *Digital Audio Processing: A Review Journal*
- *EURASIP Journal on Audio Speech and Music Processing*
- *IEEE Transactions on Multimedia*
- *Transactions of the International Society for Music Information Retrieval*
- *IEEE/ACM Transactions on Audio, Speech and Language Processing*
- *IEEE Transactions on Affective Computing*
- *IEEE Access*
- *The Journal of Supercomputing*
- *Journal of Scientific Programming*

- Reviewer for conferences
  - *ACM Symposium on Applied Perception (SAP)*
  - *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*
  - *International Society for Music Information Retrieval (ISMIR)*
  - *ACM International Conference on Multimedia* (Program Committee Member)
  - *IEEE Technical Committee on Multimedia Computing*
  - *IEEE International Symposium on Multimedia*

- Reviewer for book chapters
  - *Audio Source Separation and Speech Enhancement*