Zhiyao Duan

University of Rochester
720 Computer Studies Building
Rochester, NY 14627

(585) 275-5302
zhiyao.duan@rochester.edu
hajim.rochester.edu/ece/sites/zduan/

(Last updated on March 8, 2026)

INTERESTS

Computer Audition, Music Information Retrieval, Audio-Visual Learning, Machine Learning, AR/VR

CURRENT APPOINTMENTS

University of Rochester – Rochester, NY, USA
Professor, Department of Electrical and Computer Engineering (primary) — Jul. 2025 – present
Department of Computer Science (secondary)
Goergen Institute for Data Science (affiliated)

Mango Future – Shenzhen, China & Rochester, NY, USA
Co-Founder & Chief Scientist — Dec. 2016 – present

EDUCATION

Northwestern University - Evanston, IL, USA — August 2013
Ph.D., Department of Electrical Engineering and Computer Science
Thesis: *Computational Music Audio Scene Analysis*
Advisor: Bryan Pardo

Tsinghua University - Beijing, China — July 2008
Master of Science, Department of Automation
Thesis: *Research on Polyphonic Music Pitch Estimation*
Advisor: Changshui Zhang

Tsinghua University - Beijing, China — July 2004
Bachelor of Science, Department of Automation
Thesis: *Constructing an Assistant Training System for Long Jump*
Advisor: Changshui Zhang

PROFESSIONAL EXPERIENCE

Kwai Inc. – Bellevue, WA, USA
Principal Research Scientist, Kwai Seattle AI Lab — Jun. 2020 – Jan. 2022
- Led a research team working on multimodal interaction and creation

University of Rochester – Rochester, NY, USA
Associate Professor, Department of Electrical and Computer Engineering (primary) — Jul. 2020 – Jun. 2025
Department of Computer Science (secondary) — Jul. 2020 – Jun. 2025
Goergen Institute for Data Science (affiliated) — Jul. 2020 – Jun. 2025

Assistant Professor, Department of Electrical and Computer Engineering (primary) — Jul. 2013 – Jun. 2020
Department of Computer Science (secondary) — Jun. 2016 – Jun. 2020
Goergen Institute for Data Science (affiliated) — Jun. 2016 – Jun. 2020

Ohio State University - Columbus, OH, USA — Feb. 2013 – Mar. 2013

Visiting Researcher, Department of Computer Science and Engineering
- Investigated applications of deep learning in speech and audio signal processing

Advisor: DeLiang Wang

Northwestern University - Evanston, IL, USA Sep. 2008 – Jun. 2013
Research Assistant, Department of Electrical Engineering and Computer Science
- Developed machine learning algorithms towards audio information retrieval applications, e.g., multi-pitch estimation and tracking of music and speech, audio-score alignment, and source separation

Advisor: Bryan Pardo

Adobe Systems - San Francisco, CA, USA Jun. 2011 - Dec. 2011
Research Intern, Advanced Technology Labs (ATL)
- Invented an online machine learning algorithm for real-time semi-supervised source separation, with an application on real-time speech enhancement in non-stationary noise environments

Advisors: Gautham J. Mysore and Paris Smaragdis

Microsoft Research Asia - Beijing, China Jul. 2007 - Apr. 2008
Research Intern, Speech Group
- Designed algorithms for music tagging and tonality classification for automatic music recommendation

Advisor: Lie Lu

Stanford University - Stanford, CA, USA Apr. 2007 - Jun. 2007
Visiting Researcher, Center for Computer Research in Music and Acoustics (CCRMA)
- Implemented and compared audio signal processing algorithms for extracting guitar excitation signals

Advisor: Julius O. Smith III

Tsinghua University - Beijing, China Sep. 2005 - Mar. 2007
Research Assistant, State Key Laboratory of Intelligent Technology and Systems
- Developed machine learning algorithms towards audio information retrieval applications, e.g., multi-pitch estimation and source separation

Advisor: Changshui Zhang

NTP CO., LTD - Shenzhen, Guangdong, China Jul. 2003 - Aug. 2003
Software and Hardware Developer
- Developed and tested a motor control system

RESEARCH FUNDING

Large Music Language Models for Symbolic Music Analysis and Generation 01/01/2026-12/31/2026
University of Rochester SoundSpace Transdisciplinary Institute pilot research program ($58,250)
PI: Zhiyao Duan, Co-PI: David Temperley

Adaptive AI-Guided Immersive Concert Experiences 01/01/2026-12/31/2026
University of Rochester SoundSpace Transdisciplinary Institute pilot research program ($50,000)
PI: Yukang Yan, Co-PI: Zhiyao Duan

Enhancing Guitar Practice and Songwriting with Automatic Guitar Tablature Transcription
New York State Center of Excellence in Data Science ($30,000) 05/01/2026-04/30/2027
PI: Zhiyao Duan

Zhiyao Duan

University of Rochester
720 Computer Studies Building
Rochester, NY 14627

(585) 275-5302
zhiyao.duan@rochester.edu
hajim.rochester.edu/ece/sites/zduan/

Streaming Voice Transformation for Voice Privacy Protection 10/01/2025-09/30/2026
NAIRR Pilot (25,000 GPU-hours)
PI: Zhiyao Duan

Speaker Static Trait Disentanglement for Speaker Anonymization 09/01/2025-08/31/2026
NAIRR Pilot (12,000 H100 GPU-hours)
PI: Zhiyao Duan

REU Site: Computational Methods for Understanding Music, Media and Minds 04/01/2025-03/31/2028
National Science Foundation ($463,718)
PI: Ajay Anand, Co-PI: Zhiyao Duan

AudioPuzzle: A Masked Generative Generalist Foundation Model for Audio Intelligence 03/01/2025-02/28/2026
NAIRR Pilot (5,000 GPU-hours)
PI: Zhiyao Duan

Toward Speaker-Specific Voice Spoofing Countermeasures 01/01/2024-12/31/2025
New York State Center of Excellence in Data Science ($60,000)
PI: Zhiyao Duan

Attributable Watermarking and Deepfake Detection for Responsible Audiobox 12/01/2024-11/30/2025
Meta Audiobox Responsible Generation Grant Program ($50,000 + Audiobox licence)
PI: Zhiyao Duan

R.A.S.P. - Real-time Anonymization & Speech Protection 09/23/2024-09/22/2027
IARPA Anonymous Real-Time Speech (ARTS) Program (UR share $1.25M)
PI: Tor Finseth (Honeywell), Co-PIs: Zhiyao Duan (UR), Ricardo Gutierrez-Osuna (Texas A&M), John Hansen (UT-Dallas)

Active Control of Room Reverberation 10/01/2024-09/30/2025
University of Rochester Goergen Institute for Data Science seed funding program ($30,000)
PI: Zhiyao Duan, Co-PIs: Michael Heilemann, Mark Bocko

Safeguarding Generative AI by Audio-Visual Deepfake Detection 07/01/2024 – 06/30/2025
The National Artificial Intelligence Research Resource (NAIRR) Pilot (10,000 GPU hours)
PI: Zhiyao Duan

Development of a Machine Learning Algorithm to Identify Probability of Acute Coronary Syndrome using Reported Symptoms and Acoustic Features of 911 Callers Reporting Non-traumatic Chest Pain 07/01/2024 – 06/30/2025
University of Rochester Center for Integrated Research Computing (CIRC) ($51,948)
PI: Dillon J Dzikowicz, Co-PI: Zhiyao Duan

Toward Noise Resilient Voice Spoofing Countermeasures 01/01/2024 – 12/31/2024
New York State Center of Excellence in Data Science ($60,000)
PI: Zhiyao Duan

Unrestricted Gift 2023 – 2024

Adobe Inc. ($70,000)
PI: Zhiyao Duan

Training Audio-Visual Foundation Models to Capture Fine-Grained Dependencies 2023 – 2024
Microsoft Azure Accelerate Foundation Models Research ($20,000 Azure computing credit)
PI: Zhiyao Duan

Language Guided Audio Source Separation 11/01/2023 – 10/31/2024
University of Rochester Goergen Institute for Data Science seed funding program ($20,000)
PI: Zhiyao Duan

Developing and Deploying Spoofing Aware Speaker Verification Systems 01/01/2023 – 12/31/2023
New York State Center of Excellence in Data Science ($59,989)
PI: Zhiyao Duan

Personalized Immersive Spatial Audio with Physics Informed Neural Field 11/01/2022 – 10/31/2023
University of Rochester Goergen Institute for Data Science seed funding program ($20,000)
PIs: Zhiyao Duan and Mark Bocko

Toward an Ecosystem of Artificial Intelligence-Powered Music Production (TEAMuP) 10/01/2022 – 09/30/2026
National Science Foundation – FW-HTF-R ($1,800,000)
PI: Raffaella Borasi, Co-PIs: Zhiyao Duan, Jonathan Herington, Rachel Roberts, Bryan Pardo

Unrestricted Gift 2021-2022
Adobe Inc. ($41,500)
PI: Zhiyao Duan

Automatic Rendering of Augmented Effect in Immersive Concerts 10/01/2021 – 09/30/2022
University of Rochester Goergen Institute for Data Science seed funding program
PIs: Zhiyao Duan ($10,854), Matthew Brown ($5000), and Raffaella Borasi ($3000)

Next Generation Large Scale Robust Automatic Speaker Verification 10/01/2021 – 09/30/2022
New York State Center of Excellence in Data Science ($59,666)
PI: Zhiyao Duan

Exploring Creative Design at the Human-Technology Frontier through the Emerging "Artist-Technologist" Occupation 09/01/2020-08/31/2021
National Science Foundation – FW-HTF-P ($149,675)
PI: Raffaella Borasi, Co-PIs: Mark Bocko, James Doser, Zhiyao Duan, Joe Testani

Interdisciplinary Graduate Training in the Science, Technology, and Applications of Augmented and Virtual Reality 09/01/2019 – 08/31/2024
National Science Foundation – NRT-HDR ($1,560,000)
PI: Mujdat Cetin, Co-PIs: Jannick Rolland, Michele Rucci, Zhen Bai
Senior Personnel: Zhiyao Duan, Ross Maddox, Andrew White, Chenliang Xu, and Yuhao Zhu

REU Site: Computational Methods for Understanding Music, Media, and Minds 03/01/2020 – 02/28/2023
National Science Foundation – REU Site ($405,000)
PI: Ajay Anand, Co-PI: Zhiyao Duan

End-to-End Speaker Verification 09/01/2019 – 08/31/2021
Voice Biometric Group – Unrestricted Gift ($69,531)

PI: Zhiyao Duan

Visually Informed Music Generation 01/01/2020 – 12/31/2020
Kwai Inc. – Unrestricted Gift ($25,000)
PI: Zhiyao Duan

Music Understanding and Generation 10/15/2019 – 10/14/2020
ByteDance – Unrestricted Gift ($50,000)
PI: Zhiyao Duan

CAREER: Human-Computer Collaborative Music Making 06/01/2019 – 05/31/2024
National Science Foundation – CISE IIS-CHS core program ($499,219)
PI: Zhiyao Duan

BIGDATA: F: Audio-Visual Scene Understanding 09/01/2017 – 08/31/2021
National Science Foundation – Big Data Science & Engineering
PI: Chenliang Xu ($349,999), Co-PI: Zhiyao Duan ($300,000)

Real-Time Synthesis of a Virtual Talking Face from Acoustic Speech 07/01/2017 – 06/30/2018
University of Rochester AR/VR Pilot Funding ($50,000)
PIs: Ross Maddox, Zhiyao Duan, and Chenliang Xu

Adding High-Quality Spatial Audio to 3D-VR-360 Recordings for Live Streaming and Building a VR Video Database 07/01/2017 – 06/30/2018
University of Rochester AR/VR Pilot Funding ($69,800)
PIs: Zhiyao Duan, Ming-Lun Lee, and Matthew Brown

Development and Evaluation of an Evidence-Based Mobile Health Caregiver Intervention for FASD
National Institute of Health ($1,504,884) 07/01/2017 – 05/31/2022
PIs: Christie Petrenko and Cristiano Tapparello; Co-Is: Heather Olson, Wendi Heinzelman, and Zhiyao Duan

Algorithms for Query by Example of Audio Databases 09/01/2016 – 08/31/2019
National Science Foundation – CISE III core program
PI: Zhiyao Duan ($299,775), Co-PI: Bryan Pardo ($199,996)

Predicting Adverse Events from Cardiac Signals using Deep Neural Networks 08/22/2016 – 08/21/2017
University of Rochester Goergen Institute for Data Science Collaborative Pilot Award Program in Health Analytics
PI: Mina Attin ($26,995), Co-PI: Zhiyao Duan ($19,701)

TEACHING

Tutorials

[6] Invited Lectures at the Winter School of Speech and Audio Processing (WiSSAP) Dec. 2023
IIT Kanpur, India

- Introduction to Music Information Retrieval
- Music Pitch Analysis
- Music Rhythm Analysis
- Interactive Music Systems

[5] Computer-Assisted Music-Making Systems: Taxonomy, Review, and Live Coding Nov. 2023
International Society for Music Information Retrieval conference (ISMIR), Milan, Italy. Co-presented with Philippe Pasquier and Christos Benetatos.

[4] A Brief History of Interactive Music Systems Aug. 2022
CAAI International Conference on Artificial Intelligence (CICAI) under Theme 3 – Art and Artificial Intelligence

[3] Audio Scene Understanding Jun. 2021
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Tutorial on "Audio-Visual Scene Understanding", organized by Di Hu, Lele Chen, Yapeng Tian, Zhiyao Duan, and Chenliang Xu

[2] Audiovisual Music Processing Nov. 2019
International Society for Music Information Retrieval conference (ISMIR), Delft, The Netherlands. Co-presented with Slim Essid, Bochen Li, and Sanjeel Parekh.

[1] Automatic Music Transcription Oct. 2015
International Society for Music Information Retrieval conference (ISMIR), Malaga, Spain. Co-presented with Emmanouil Benetos.

Courses Designed

[6] ECE 408: The Art of Machine Learning, Undergrad/Grad Level Spring 2023-2026
Instructor, University of Rochester, Rochester, NY, USA

[5] Music and Math, Pre-College Level Summer 2016, 2017, 2019
Instructor, Upward Bound Program, University of Rochester, Rochester, NY, USA

[4] ECE 477: Computer Audition, Grad Level Fall 2014-2015, 2017-2019, 2022-2025
Instructor, University of Rochester, Rochester, NY, USA

[3] Y0250421: Computer Audition, Grad Level Summer, 2015
Instructor, Tsinghua University, Beijing, China

[2] ECE 272/472: Audio Signal Processing, Undergrad/Grad Level Spring 2014-2020
Instructor, University of Rochester, Rochester, NY, USA

[1] ECE 492: Computer Audition and Its Applications in Music, Grad Level Fall 2013
Instructor, University of Rochester, Rochester, NY, USA

Courses Involved

[9] ECE 411: Special Topics in Augmented and Virtual Reality Spring 2024
Co-Instructor, University of Rochester, Rochester, NY, USA

[8] ECE 410: Introduction to Augmented and Virtual Reality: Audio for AR/VR Fall 2022-2024
Co-Instructor, University of Rochester, Rochester, NY, USA

[7] CSC 294: AR/VR Interaction Design Fall 2019
Guest Lecturer, University of Rochester, Rochester, NY, USA

- Designed and gave a lecture on Computer Audition and Music Interaction

Zhiyao Duan

University of Rochester
720 Computer Studies Building
Rochester, NY 14627

(585) 275-5302
zhiyao.duan@rochester.edu
hajim.rochester.edu/ece/sites/zduan/

[6] CSC 249/449: Machine Vision — Spring 2018
Guest Lecturer, University of Rochester, Rochester, NY, USA
- Designed and gave a lecture on Multi-Modal Music Scene Understanding

[5] CSC 412: Human Computer Interaction — Fall 2013
Guest Lecturer, University of Rochester, Rochester, NY, USA
- Designed and gave a lecture on Music Interaction

[4] EECS 349: Machine Learning — Fall 2010, 2011, 2012
Teaching Assistant and Guest Lecturer, Northwestern University, Evanston, IL, USA
- Designed and gave lectures on Ensemble Learning, Memory-based Learning, Gaussian Mixture Models, and Expectation-Maximization;
- Designed homework problems on the above topics and decision trees
- Held office hours; graded homework, exams and final projects

[3] Introduction to Artificial Intelligence — Fall 2007
Teaching Assistant, Tsinghua University, Beijing, China
- Held office hours, graded homework and final projects

[2] Object-Oriented Computer Programming (Visual C++) — Fall 2007
Lab Instructor, Tsinghua University, Beijing China
- Led weekly lab sessions
- Mentored students on final projects; graded homework and final projects

[1] Fundamentals of Computer Programming (C++) — Spring 2006
- Led weekly lab sessions
- Mentored students on final projects; graded homework and final projects

Visiting Scholar Hosting

- Fei Jiang (visiting PhD student from Beijing Institute of Technology, China, Sep. 2019 – Feb. 2021)
- Nan Jiang (visiting PhD student from Tsinghua University, China, September 2019 – March 2020)
- Jianyu Fan (visiting PhD student from Simon Fraser University, Canada, March 2020)
- Rui Lu (visiting PhD student from Tsinghua University, China, October 2017 – August 2018)
- Hongjuan Zhang (Shanghai University, China, December 2016 – December 2017)
- Jun Zhou (Southwest University, China, September 2014 – August 2015)

Doctoral Thesis Supervising

Current
- Venkatakrishnan Vaidyanathapuram Krishnan, co-supervised by Prof. Sam Norman-Haignere
- Ryan Bowering
- Baotong Tian
- Stella Wong
- Xingjian Du
- Huiran Yu
- Meiying (Melissa) Chen
- Frank Cwitkowitz (expected December 2026)
- Mojtaba Heydari (expected December 2026)

- You (Neil) Zhang (expected May 2026)
- Christos Benetatos (expected December 2026)

Graduated

- Yujia Yan, *Structured Analysis and Generation in Music, Audio, and Beyond*, August 2025
- Ge Zhu, *The Generative Perspective for Diverse Audio Applications,* June 2025
- Bochen Li, *Multi-Modal Analysis for Music Performances,* August 2020 (2021 University of Rochester Outstanding PhD Dissertation Award)
- Yichi Zhang, *Sound Search by Vocal Imitation*, December 2019
- Sefik Emre Eskimez, *Robust Techniques for Generating Talking Faces from Speech*, August 2019, co-supervised by Prof. Wendi Heinzelman
- Andrea Cogliati, *Toward a Human-Centric Automatic Piano Music Transcription System*, December 2017

Doctoral Thesis Reading

- Kavya Ranjan Saxena (EE, Indian Institute of Technology Kanpur, November 2025)
- Menglu Li (Electrical, Computer, and Biomedical Engineering, Toronto Metropolitan University, July 2025)
- Hang Hua (CS, May 2025)
- Ben Maman (CS, Tel Aviv University, November 2024)
- David Skrill (Neuroscience, November 2024)
- Yixiao Zhang (EECS, Queen Mary University of London, November 2024)
- Paul Mitchell (BME, October 2024)
- Shadi Sartipi (ECE, August 2024)
- Tong Shan (Biomedical Engineering, May 2024)
- Ho-Hsiang Wu (Music Technology, NYU, October 2023)
- Philip Mehrgardt (CS, University of Sydney, June 2023)
- Jeffrey Rowan (ECE, June 2023)
- Songyang Zhang (CS, May 2023)
- Yapeng Tian (CS, June 2022)
- Lele Chen (CS, May 2022)
- Zhenyuan Yang (CS, August 2021)
- Sahar Hashemgeloogerdi (ECE, November 2019)
- Jianbo Yuan (CS, September 2019)
- Sarah Smith (ECE, May 2019)
- Priyanga Gunarathne (Simon Business School, May 2018)
- Xiaochang Peng (CS, May 2018)
- Chen Wang (ECE, December 2017)
- Ahmed Elliethy (ECE, February 2017)
- Dave Anderson (ECE, January 2017)
- Gang Ren (ECE, November 2015)
- He Ba (ECE, February 2015)
- Na Yang (ECE, March 2015)

Master's Student Advising

Zijian Gu, CS, U Rochester – Misophonia Trigger Detection	Spring 2026
Fei-Yueh Chen, ECE & Linguistics, U Rochester – Symbolic Music Generation	Fall 2025-Spring 2026
Jiajun Wu, ECE, U Rochester – Music Generation for Movies	Fall 2025-Spring 2026
Brynn Lee, Data Science, U Rochester – Speech Anti-spoofing	Fall 2024-Spring 2025
Zehua Li, TEAM, U Rochester – Expressive Text-to-Speech Synthesis	Spring 2023-Fall 2024

Zhiyao Duan

University of Rochester
720 Computer Studies Building
Rochester, NY 14627

(585) 275-5302
zhiyao.duan@rochester.edu
hajim.rochester.edu/ece/sites/zduan/

Siwen Ding, Columbia University – Speaker Anti-Spoofing	06/22-03/23
Zewen Song, ECE, U Rochester – Online Music Structural Segmentation	10/21-05/22
Qiaoyu Yang, ECE, U Rochester – Pitch Relations Between Singing and Speaking of Cantonese Songs	09/21-05/23
Xinhui Chen, CS, U Rochester – Anti-spoofing Speaker Verification	01/21-08/21
Anton Selitskiy, ECE, U Rochester – Audio Bandwidth Extension	09/20-08/22
Mingqing Yun, ECE, U Rochester – Multi-talker Localization in Reverberant Environments	01/18-05/19
Isaac Mosebrook, ECE, U Rochester – Parkinson's Disease Diagnosis from Voice	09/18-05/19
Hilary Mogul, ECE, U Rochester – Beamforming with Ambisonics	Fall 2018
Yufei Zhang, ECE, U Rochester – Audio-Visual Pitch Estimation	Summer 2018
Yiming Zhao, ECE, U Rochester – Automatic Lyrics Display System Design	06/17-05/18
Zhuohuang Zhang, ECE, U Rochester – Speaker Diarization	09/16-05/17
Iris Yuping Ren, ECE, U Rochester – Singing Tutoring System	09/15-12/17
Jonathan Downing, ECE, U Rochester – Thesis on Joint Separation and Dereverberation	01/16-08/16
Jay Biernat, ECE, U Rochester – Music Visualization	Summer 2016
Marko Stamenovic, ECE, U Rochester – Music Recommendation	Summer 2016
Xinzhao Liu, ECE, U Rochester – Thesis on Audio-Visual Music Performance Analysis	05/15-05/16
Hanqing Wen, ECE, U Rochester – Music Onset Detection	06/15-05/16
Zejin Li, ECE, U Rochester – Audio-Visual Guitar Transcription	Summer 2015
Yuhui Chen, ECE, U Rochester – Audio-Visual Guitar Transcription	Summer 2015
Shumin Xu, ECE, U Rochester – Music Universe	09/14-05/15
David Heid, ECE, U Rochester – Vibrato Analysis	01/15-05/15
Andrew Trahan, ECE, U Rochester – Thesis on Drum Kit Transcription	Spring 2014
Jonathan Springer, EECS, Northwestern University – Bird Species Recognition	Fall 2012
Jesse Bowman, EECS, Northwestern University – Real-time Multi-pitch Estimation for Guitars	07/2018-06/2011

Undergraduate Research Advising

Marcos Ibáñez Matles, ECE & CS	Spring 2026
Mai Pham, Data Science & Business Information System, U Rochester	Spring 2026
Nigel Nguyen CS, U Rochester	Spring 2026
Eliss Haddad, CS & Business, U Rochester	Spring 2026
Cynthia Lu, CS, U Rochester	Spring 2026
Derek Llanes, CS & Math, U Rochester	Spring 2026
Elek Yuhas, AME, U Rochester	Spring 2026
Natalie Smith, AME, U Rochester	Spring 2026
Richard Zhu, AME & Business, U Rochester	Spring 2026
Yicheng Wang, CS, U Rochester	Spring 2026
Yiru Zhou, Math, NYU Shanghai	Summer 2024-Fall 2024
Jiajun Wu, AME, U Rochester	Fall 2024-Spring 2025
Paul Berggren, ECE, U Rochester	Summer 2024-Spring 2025
Kyungbok Lee, CS, U Rochester	Fall 2023-Spring 2024
Yutong Wen, AME, U Rochester	Spring 2023-Spring 2024
Enting Zhou, CS, U Rochester	Spring 2023
Yongyi Zang, ECE, U Rochester	Summer 2021-Spring 2023
Jordan Darefsky, CS, U Rochester	Fall 2021-Spring 2024
Chanha Kim, Pomona College, NSF REU Site on Music, Media and Mind	Summer 2021
Farrah Pierre-Louis, Simmons University, NSF REU Site on Music, Media and Mind	Summer 2021
Panzhen Wu, Math & Music	Spring 2021
Qiaoyu Yang, CS, U Rochester	Spring 2021
Mingrui Yuan, EE, Tsinghua University, Visiting Student	Summer 2019

Hangyu Li, EE, Beihang University, Visiting Student	Summer 2019
Yinghao Ma, Math, Peking University, Visiting Student	Summer 2019
Daniel Dopp, Kentucky University, NSF REU Site on Music, Media and Mind	Summer 2019
Nick Creel, Marlboro College, NSF REU Site on Music, Media and Mind	Summer 2019
Yiting Zhang, ECE, University of Rochester, Xerox Engineering Fellowship	Summer 2018 – Spring 2019
Peizhe Gao, CS, Beihang University, Visiting Student	Summer 2018
Junyi Fan, China University of Geosciences, Visiting Student	Spring and Summer 2018
Zhihan Zhou, Math, Zhejiang University, Visiting Student	Summer 2017
Andrew Smith, CS, University of Central Florida, NSF REU Site on Music, Media and Mind	Summer 2017
Arlen Fan, ECE, University of Rochester, NSF REU Site on Music, Media and Mind	Summer 2017
Ayumi Yuasa, ECE, University of Rochester, NSF REU Student	Summer 2017
Ryan Bhular, ECE University of Rochester, NSF REU Student	Summer 2017
Meixiao Han, ECE, University of Rochester, Undergraduate Researcher	Summer 2017
Yukun Chen, ECE, University of Rochester, Undergraduate Researcher	Summer and Fall 2016
Karan Vombatkere, ECE, University of Rochester, Xerox Engineering Fellowship	Summer 2016
Steven Belitzky, ECE, University of Rochester, Discover Grant Researcher	Summer 2016
Almas Abdibayev, CS, Nazarbayev University, Visiting Student	Summer 2015
Ibrahim Akbar, ECE, University of Rochester, Undergraduate Researcher	Spring 2015
Haowen Pan, ECE, University of Rochester, Xerox Engineering Fellowship	Summer 2014
Prem Seetharaman, EECS, Northwestern University	Winter 2012
Jiawei Lyu, Automation, Tsinghua University	Spring 2008

PUBLICATIONS

(†: undergraduate students)

Book Chapters

[3] You Zhang, Fei Jiang, Ge Zhu, Xinhui Chen, and **Zhiyao Duan**, "Generalizing voice presentation attack detection to unseen synthetic attacks and channel variation," in Marcel, S., Fierrez, J., Evans, N. (eds). *Handbook of Biometric Anti-Spoofing: Presentation Attack Detection and Vulnerability Assessment*. Springer, Singapore, 2023.

[2] Bryan Pardo, Antoine Liutkus, **Zhiyao Duan**, Gaël Richard, "Applying source separation to music," in *Audio Source Separation and Speech Enhancement*, eds. E. Vincent, T. Virtanen, S. Gannot. Wiley, pp. 347-376, 2018.

[1] Bryan Pardo, Zafar Rafii, and **Zhiyao Duan**, "Audio source separation in a musical context," in *Springer Handbook of Systematic Musicology*, ed. Rolf Bader. Springer-Verlag Berlin Heidelberg, pp. 285-298, 2018.

Journal Publications

[38] Ge Zhu, Yutong Wen, and **Zhiyao Duan**, "Audio generation through score-based generative modeling: design principles and implementation," accepted by *Foundations and Trends in Signal Processing*, 2026.

[37] Raquel Norel, Jennifer Gewandter, Zhengwu Zhang, Anika Tahsin, Chadi G Abdallah, John Markman, **Zhiyao Duan**, Guillermo Cecchi, and Paul Geha, "Turning patients' open-ended narratives of chronic pain into quantitative measures: Natural language processing study," *JMIR Human Factors*, vol. 12, 2025.

[36] Ge Zhu, Jordan Darefsky†, and **Zhiyao Duan**, "Cacophony: An improved contrastive audio-text model," accepted by *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 2024.

[35] Mojtaba Heydari and **Zhiyao Duan**, "BeatNet+: Real-time rhythm analysis for diverse music audio," accepted by *Transactions of the International Society for Music Information Retrieval*, 2024.

[34] Yujian Yan and **Zhiyao Duan**, "Measure by measure: Measure-based automatic music composition with modern staff notation," accepted by *Transactions of the International Society for Music Information Retrieval*, 2024.

Zhiyao Duan

University of Rochester
720 Computer Studies Building
Rochester, NY 14627

(585) 275-5302
zhiyao.duan@rochester.edu
hajim.rochester.edu/ece/sites/zduan/

[33] Ge Zhu, Juan-Pablo Careres, **Zhiyao Duan**, and Nicholas J. Bryan, "MusicHiFi: Fast high-fidelity stereo vocoding," *IEEE Signal Processing Letters*, 2024.

[32] **Zhiyao Duan***, Peter van Kranenburg*, Juhan Nam*, and Preeti Rao*, "Editorial for TISMIR Special Collection: Cultural Diversity in MIR Research," *Transactions of the International Society for Music Information Retrieval, Special Collection on Cultural Diversity in MIR Research, 2024. (* authors in alphabetical order)*

[31] Yongyi Zang*†, Christodoulos Benetatos*, and **Zhiyao Duan**, "Euterpe: A web framework for music interaction and creation," accepted by *Journal of Audio Engineering Society*, 2023. (* equal contribution)

[30] Tong Shan, Casper E. Wenner, Chenliang Xu, **Zhiyao Duan**, and Ross K. Maddox, "Speech-in-noise comprehension is improved when viewing a deep-neural-network-generated talking face," *Trends in Hearing*, vol. 26, pp. 1-10, 2022.

[29] Ge Zhu, Jordan Darefsky†, Fei Jiang, Anton Selitskiy, and **Zhiyao Duan**, "Music source separation with generative flow," *IEEE Signal Processing Letters*, vol. 29, pp. 2288-2292, 2022.

[28] Christodoulos Benetatos and **Zhiyao Duan**, "Draw and listen! A sketch-based system for music inpainting," *Transactions of the International Society for Music Information Retrieval*, 2022.

[27] 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.

[26] Sefik Emre Eskimez, You Zhang, and Zhiyao Duan, "Speech driven talking face generation from a single image and an emotion condition," *IEEE Transactions on Multimedia*, vol. 24, pp. 3480-3490, 2021.

[25] You Zhang, Fei Jiang, and Zhiyao Duan, "One-class learning towards synthetic voice spoofing detection," *IEEE Signal Processing Letters*, vol. 28, pp. 937-941, 2021.

[24] Fei Jiang and **Zhiyao Duan**, "Speaker attractor network: generalizing speech separation to unseen numbers of sources," *IEEE Signal Processing Letters*, vol 27, pp. 1859-1863, 2020.

[23] Sefik Emre Eskimez, Ross Maddox, Chenliang Xu, and **Zhiyao Duan**, "Noise-resilient training method for face landmark generation from speech," *IEEE/ACM Transactions on Audio Speech and Language Processing*, vol. 28, pp. 27-38, 2019.

[22] Bochen Li, Karthik Denish, 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.

[21] Rui Lu, **Zhiyao Duan**, and Changshui Zhang, "Audio-visual deep clustering for speech separation," *IEEE/ACM Transactions on Audio Speech and Language Processing*, vol. 27, no. 11, pp. 1697-1712, 2019.

[20] Sefik Emre Eskimez, Kazuhito Koishida, and **Zhiyao Duan**, "Adversarial training for speech super-resolution," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 2, pp. 347-358, 2019.

[19] **Zhiyao Duan***, Slim Essid*, Cynthia C. S. Liem*, Gaël Richard*, "Audio-visual analysis of music performances," *IEEE Signal Processing Magazine,* vol. 36, no. 1, pp. 63-73, 2019. (* authors in alphabetical order)

[18] Emmanouil Benetos*, Simon Dixon*, **Zhiyao Duan***, and Sebastian Ewert*, "Automatic music transcription: an overview," *IEEE Signal Processing Magazine,* vol. 36, no. 1, pp. 20-30, 2019. (* authors in alphabetical order)

[17] Yichi Zhang, Bryan Pardo, and **Zhiyao Duan**, Siamese style convolutional neural networks for sound search by vocal imitation, accepted by *IEEE/ACM Transactions on Audio Speech and Language Processing,* vol. 27, no. 2, pp. 429-441, 2019.

[16] Bochen Li*, Xinzhao Liu*, Karthik Dinesh, **Zhiyao Duan**, and Gaurav Sharma, "Creating a multi-track classical music performance dataset for multi-modal music analysis: challenges, insights, and applications," *IEEE Transactions on Multimedia*, vol., 21, no. 2, pp. 522-535, 2019. (* equal contribution)

[15] Rui Lu, **Zhiyao Duan**, and Changshui Zhang, "Listen and look: audio-visual matching assisted speech source separation", *IEEE Signal Processing Letters*, vol. 25, no. 9, pp. 1315-1319, 2018.

[14] Sefik Emre Eskimez, Peter Soufleris, **Zhiyao Duan**, and Wendi Heinzelman, "Front-end speech enhancement for commercial speaker verification systems," *Speech Communication*, vol. 99, no. pp. 101-113, 2018.

[13] Shiwei Yu, Hongjuan Zhang, and **Zhiyao Duan**, "Singing voice separation by low-rank and sparse spectrogram decomposition with pre-learned dictionaries," *Journal of the Audio Engineering Society*, vol. 65, no. 5, pp. 377-388, 2017.

[12] Andrea Cogliati, **Zhiyao Duan**, and Brendt Wohlberg, "Piano transcription with convolutional sparse lateral inhibition," *IEEE Signal Processing Letters*, vol. 24, no. 4, pp. 392-396, 2017.

[11] David Temperley, Iris Ren, and **Zhiyao Duan**, "Mediant mixture and 'blue notes' in rock: An exploratory study," *Music Theory Online*, vol. 23, no. 1, 2017.

[10] Na Yang, Jianbo Yuan, Yun Zhou, Ilker Demirkol, **Zhiyao Duan**, Wendi Heinzelman, and Melissa Sturge-Apple, "Enhanced multiclass SVM with thresholding fusion for speech-based emotion classification," *International Journal of Speech Technology*, vol. 20, no. 1, pp. 27-41, 2017. DOI: 10.1007/s10772-016-9364-2.

[9] Bochen Li and **Zhiyao Duan**, "An approach to score following for piano performances with the sustained effect," *IEEE/ACM Trans. Audio Speech Language Process.*, vol. 24, no. 12, pp. 2425-2438, 2016.

[8] Andrea Cogliati, **Zhiyao Duan**, and Brendt Wohlberg, "Context-dependent piano music transcription with convolutional sparse coding," *IEEE/ACM Trans. Audio Speech Language Process.*, vol. 24, no. 12, pp. 2218-2230, 2016.

[7] Yichi Zhang and **Zhiyao Duan**, "Supervised and unsupervised sound retrieval by vocal imitation," *Journal of Audio Engineering Society*, vol. 64, no. 7/8, pp. 533-543, 2016.

[6] Francisco J. Rodriguez-Serrano, **Zhiyao Duan**, Pedro Vera-Candeas, Bryan Pardo, and Julio J. Carabias-Orti, "Online score-informed source separation with adaptive instrument models," *Journal of New Music Research*, vol., 44, no. 2, pp., 83-96, 2015. DOI: 10.1080/09298215.2014.989174.

[5] Zafar Rafii, **Zhiyao Duan**, and Bryan Pardo, "Combining rhythm-based and pitch-based methods for background and melody separation," *IEEE Trans. Audio Speech Language Process.*, vol. 22, no. 12, pp. 1884-1893, 2014.

[4] **Zhiyao Duan**, Jinyu Han, and Bryan Pardo, "Multi-pitch streaming of harmonic sound mixtures," *IEEE Trans. Audio Speech Language Process.*, vol. 22, no. 1, pp. 138-150, 2014.

[3] **Zhiyao Duan** and Bryan Pardo, "Soundprism: an online system for score-informed source separation of music audio," *IEEE Journal of Selected Topics in Signal Processing*., vol. 5, no. 6, pp. 1205-1215, 2011.

[2] **Zhiyao Duan**, Bryan Pardo, and Changshui Zhang, "Multiple fundamental frequency estimation by modeling spectral peaks and non-peak regions," *IEEE Trans. Audio Speech Language Process*., vol. 18, no. 8, pp. 2121-2133, 2010.

[1] **Zhiyao Duan**, Yungang Zhang, Changshui Zhang, and Zhenwei Shi, "Unsupervised single-channel music source separation by average harmonic structure modeling," *IEEE Trans. Audio Speech Language Process*., vol. 16, no. 4, pp. 766-778, 2008.

Peer-reviewed Conference Publications

[93] Frank Cwitkowitz, Christodoulos Benetatos, Qixin Deng, Huiran Yu, Nathan Pruyne‡, Patrick O'Reilly, Hugo Flores Garcia, **Zhiyao Duan**, Bryan Pardo, "HARP 3.0: Generalizing I/O and API support for machine learning in digital audio workstations," in *NeurIPS 2025 Workshop on AI for Music*, 2025.

[92] Yu Zhang, Baotong Tian, and **Zhiyao Duan**, "Conan: A chunkwise online network for zero-shot adaptive voice

conversion," in *Proc. IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2025.

[91] You Zhang, Andrew Francl, Ruohan Gao, Paul Calamia, **Zhiyao Duan**, and Ishwarya Ananthabhotla, "Towards Perception-Informed Latent HRTF Representations," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2025. (best student paper award)

[90] Frank Cwitkowitz and **Zhiyao Duan**, "Investigating an overfitting and degeneration phenomenon in self-supervised multi-pitch estimation," in *Proc. International Society for Music Information Retrieval (ISMIR)*, 2025.

[89] Kyungbok Lee†, You Zhang, and **Zhiyao Duan**, "Audio visual segmentation through text embeddings," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2025.

[88] You Zhang*, Baotong Tian*, Lin Zhang, and **Zhiyao Duan**, "PartialEdit: Identifying partial deepfakes in the era of neural speech editing," in *Proc. Interspeech*, 2025. (* equal contribution)

[87] Geoffroy Peeters, Zafar Rafii, Magdalena Fuentes, **Zhiyao Duan**, Emmanouil Benetos, Juhan Nam, and Yuki Mitsufuji, "Twenty-five years of MIR research: Achievements, practices, evaluations, and future challenges," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025.

[86] You Zhang, Yongyi Zang, Jiatong Shi, Ryuichi Yamamoto, Tomoki Toda, and **Zhiyao Duan**, "SVDD 2024: The inaugural singing voice deepfake detection challenge," in *Proc. IEEE Spoken Language Technology Workshop (SLT)*, 2024.

[85] Kyungbok Lee†, You Zhang, and **Zhiyao Duan**, "A multi-stream fusion approach with one-class learning for audio-visual deepfake detection," in *Proc. IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2024.

[84] Samuele Cornell, Jordan Darefsky†, **Zhiyao Duan**, and Shinji Watanabe, "Generating data with text-to-speech and large-language models for conversational speech recognition," in Interspeech 2024 Workshop on SynData4GenAI, 2024.

[83] Yujia Yan and **Zhiyao Duan**, "Scoring intervals using non-hierarchical transformer for automatic piano transcription," in *Proc. International Society for Music Information Retrieval Conference (ISMIR)*, 2024. (best paper runner-up)

[82] Huiran Yu and **Zhiyao Duan**, "Note-level transcription of choral music," in *Proc. International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[81] Yongyi Zang, Jiatong Shi, You Zhang, Ryuichi Yamamoto, Jionghao Han, Yuxun Tang, Shengyuan Xu, Wenxiao Zhao, Jing Guo, Tomoki Toda, and **Zhiyao Duan**, "CtrSVDD: A benchmark dataset and baseline analysis for controlled singing voice deepfake detection," in *Proc. Interspeech*, 2024.

[80] Zehua Li, Meiying Chen, Yi Zhong, and **Zhiyao Duan**, "GTR-Voice: Articulatory phonetics informed controllable expressive speech synthesis," in *Proc. Interspeech*, 2024.

[79] Yongyi Zang†, Yi Zhong, Frank Cwitkowitz, and **Zhiyao Duan**, "SynthTab: Leveraging synthesized data for guitar tablature transcription," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024.

[78] Yongyi Zang†, You Zhang, Mojtaba Heydari, and **Zhiyao Duan**, "SingFake: Singing voice deepfake detection," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024.

[77] Enting Zhou†, You Zhang, and **Zhiyao Duan**, "Learning arousal-valence representation from categorical emotion labels of speech," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024.

[76] Ge Zhu, Yutong Wen†, Marc-André Carbonneau, and **Zhiyao Duan**, "EDMSound: Spectrogram based diffusion models for efficient and high-quality audio synthesis," in *NeurIPS 2023 Workshop on Machine Learning for Audio*, 2023.

[75] Hugo Flores Garcia, Christodoulos Benetatos, Patrick O'Reilly, Aldo Aguilar, **Zhiyao Duan**, and Bryan Pardo, "HARP: Bringing deep learning to the DAW with hosted, asynchronous, remote processing," in *NeurIPS 2023 Workshop on Machine Learning for Creativity and Design (ML4CD)*, 2023.

[74] Yutong Wu†, You Zhang, and **Zhiyao Duan**, "Mitigating cross-database differences for learning unified HRTF representation," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2023.

[73] Qiaoyu Yang, Frank Cwitkowitz, and **Zhiyao Duan**, "Harmonic analysis with neural semi-CRF," in *Proc. International Society for Music Information Retrieval (ISMIR)*, 2023.

[72] Meiying Chen and **Zhiyao Duan**, "ControlVC: Zero-shot voice conversion with time-varying controls on pitch and speed," in *Proc. Interspeech*, 2023.

[71] Yongyi Zang†, You Zhang, and **Zhiyao Duan**, "Phase perturbation improves channel robustness for speech spoofing countermeasures," in *Proc. Interspeech*, 2023.

[70] Ge Zhu, Yujia Yan, Juan-Pablo Caceres, and **Zhiyao Duan**, "Transcription free filler word detection with neural semi-CRFs," in *Proc. IEEE Internationsl Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023.

[69] Siwen Ding, You Zhang, and **Zhiyao Duan**, "SAMO: speaker attractor multi-center one-class learning for voice anti-spoofing," in *Proc. IEEE Internationsl Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023.

[68] You Zhang, Yuxiang Wang, and **Zhiyao Duan**, "HRTF field: unifying measured HRTF magnitude representation with neural fields," in *Proc. IEEE Internationsl Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023. (top 3% of all accepted papers)

[67] Mojtaba Heydari, Ju-Chiang Wang, and **Zhiyao Duan**, "SingNet: a real-time singing voice beat and downbeat tracking system," in *Proc. IEEE Internationsl Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023.

[66] Abudukelimu Wuerkaixi, Kunda Yan, You Zhang, **Zhiyao Duan**, and Changshui Zhang, "DyViSE: Dynamic vision-guided speaker embedding for audio-visual speaker diarization," accepted by *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2022.

[65] Abudukelimu Wuerkaixi, You Zhang, **Zhiyao Duan**, and Changshui Zhang, "Rethinking audio-visual synchronization for active speaker detection," accepted by IEEE International Workshop on Machine Learning for Signal Processing (MLSP), 2022.

[64] You Zhang, Ge Zhu, and **Zhiyao Duan**, "A probabilistic fusion framework for spoofing aware speaker verification," in *Proc. The Speaker and Language Recognition Workshop (Odyssey)*, 2022, pp. 77-84.

[63] Mojtaba Heydari and **Zhiyao Duan**, "Singing beat tracking with self-supervised front-end and linear transformers," in *Proc. International Society for Music Information Retrieval (ISMIR)*, 2022, pp. 617-624.

[62] Frank Cwitkowitz, Jonathan Driedger and **Zhiyao Duan**, "A data-driven methodology for considering feasibility and pairwise likelihood in deep learning based guitar tablature transcription systems," in *Proc. The Sound and Music Computing Conference (SMC)*, 2022, pp. 131-138.

[61] Frank Cwitkowitz, Mojtaba Heydari and **Zhiyao Duan**, "Learning sparse analytic filters for piano transcription," in *Proc. The Sound and Music Computing Conference (SMC)*, 2022, pp. 209-216.

[60] Mojtaba Heydari, Matthew McCallum, Andreas Ehmann, and **Zhiyao Duan**, "A novel 1d state space for efficient music rhythmic analysis," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 421-425. (video/poster presentation)

[59] Ge Zhu, Frank Cwitkowitz, and **Zhiyao Duan**, "A study of the robustness of raw waveform based speaker embeddings under mismatched conditions," in *Proc. IEEE International Conference on Acoustics, Speech and*

Signal Processing (ICASSP), 2022, pp. 7657-7661. (video/poster presentation)

[58] Rui Lu, Baigong Zheng, Jiarui Hai, Fei Tao, **Zhiyao Duan**, and Ji Liu, "Progressive teacher-student training framework for music tagging," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 3129-3133. (video/poster presentation)

[57] Yujia Yan, Frank Cwitkowitz, and **Zhiyao Duan**, "Skipping the frame-level: Event-based piano transcription with neural semi-CRFs," in *Proc. The Thirty-fifth Conference on Neural Information Processing Systems (NeurIPS)*, 2021. (video/poster presentation)

[56] Xinhui Chen*, You Zhang*, Ge Zhu*, and **Zhiyao Duan**, "UR channel-robust synthetic speech detection system for ASVspoof 2021," in *Proc. 2021 Edition of the Automatic Speaker Verification and Spoofing Countermeasures Challenge Workshop (ASVspoof)*, 2021, pp. 75-82. (* equal contribution) (oral presentation)

[55] Mojtaba Heydari, Frank Cwitkowitz, and **Zhiyao Duan**, "BeatNet: A real-time music integrated beat and downbeat tracker," in *Proc. International Society for Music Information Retrieval Conference (ISMIR)*, 2021, pp. 270-277. (video/poster presentation)

[54] Abudukelimu Wuerkaixi†, Christodoulos Benetatos, **Zhiyao Duan**, and Changshui Zhang, "CollageNet: Fusing arbitrary melody and accompaniment into a coherent song," in *Proc. International Society for Music Information Retrieval Conference (ISMIR)*, 2021, pp. 786-793. (video/poster presentation)

[53] You Zhang, Ge Zhu, Fei Jiang, and **Zhiyao Duan**, "An empirical study on channel effects for synthetic voice spoofing countermeasure systems," in *Proc. Interspeech*, 2021, pp. 4309-4313, 2021. (video/poster presentation)

[52] Ge Zhu, Fei Jiang, and **Zhiyao Duan**, "Y-vector: Multiscale waveform encoder for speaker embedding," in *Proc. Interspeech*, 2021, pp. 96-100. (video/poster presentation)

[51] Mojtaba Heydari and **Zhiyao Duan**, "Don't look back: An online beat tracking method using RNN and enhanced particle filtering," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 236-240. (video/poster presentation)

[50] Yuxiang Wang, You Zhang, **Zhiyao Duan**, and Mark Bocko, "Global HRTF personalization using anthropometric measures," in *Audio Engineering Society 150th Convention*, 2021.

[49] Nan Jiang, Sheng Jin†, **Zhiyao Duan**, and Changshui Zhang, "When counterpoint meets Chinese folk melodies," in *Proc. The Thirty-fourth Conference on Neural Information Processing Systems (NeurIPS)*, 2020. (video/poster presentation)

[48] Christodoulos Benetatos, Joseph VanderStel, and **Zhiyao Duan**, "BachDuet: A deep learning system for human-machine counterpoint improvisation," in *Proc. International Conference on New Interfaces for Musical Expression (NIME)*, 2020, pp. 635-640. (video/oral presentation)

[47] Sefik Emre Eskimez, Ross Maddox, Chenliang Xu, and **Zhiyao Duan**, "End-to-end generation of talking faces from noisy speech," in *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1948-1952. (video/poster presentation)

[46] Yichi Zhang, Junbo Hu, Yiting Zhang†, Bryan Pardo, and **Zhiyao Duan**, "Vroom!: A search engine for sounds by vocal imitation queries," in *Proc. ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR)*, 2020, pp. 23-32. (oral presentation)

[45] Nan Jiang, Sheng Jin†, **Zhiyao Duan**, and Changshui Zhang, "RL-Duet: Online music accompaniment generation using deep reinforcement learning," in *Proc. AAAI*, 2020, pp. 710-718. (oral/poster presentation)

[44] Lele Chen, Ross K. Maddox, **Zhiyao Duan**, and Chenliang Xu, "Hierarchical cross-modal talking face generation with dynamic pixel-wise loss," in *Proc. CVPR*, 2019.

[43] Bongjun Kim, Madhav Ghei, Bryan Pardo, and **Zhiyao Duan**, "Vocal Imitation Set: a dataset of vocally imitated sound events using the AudioSet ontology," in *Proc. of the Detection and Classification of Acoustic Scenes*

and Events Workshop (DCASE), 2018, pp. 148-152.

[42] Yapeng Tian, Jing Shi†, Bochen Li, **Zhiyao Duan**, and Chenliang Xu, "Audio-visual event localization in unconstrained videos," accepted by *European Conference on Computer Vision (ECCV)*, 2018, pp. 247-263.

[41] Lele Chen, Zhiheng Li†, Ross Maddox, **Zhiyao Duan**, and Chenliang Xu, "Lip movements generation at a glance," accepted by *European Conference on Computer Vision (ECCV)*, 2018, pp. 520-535.

[40] Bochen Li, Akira Maezawa, and **Zhiyao Duan**, "Skeleton plays piano: online generation of pianist body movements from MIDI performance," accepted by *International Society for Music Information Retrieval Conference (ISMIR)*, 2018, pp. 218-224 (oral/poster presentation).

[39] Yujia Yan, Ethan Lustig, Joseph Vaderstel, and **Zhiyao Duan**, "Part-invariant model for music generation and harmonization," accepted by *International Society for Music Information Retrieval Conference (ISMIR)*, 2018, pp. 204-210. (oral/poster presentation)

[38] Sefik Emre Eskimez, Ross K. Maddox, Chenliang Xu, and **Zhiyao Duan**, "Generating talking face landmarks from speech," in Proc. *International Conference on Latent Variable Analysis and Signal Separation (LVA/ICA)*, 2018. (poster presentation)

[37] Zhihan Zhou†, Yichi Zhang, and **Zhiyao Duan**, "Joint speaker diarization and recognition using convolutional and recurrent neural networks," accepted by *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 2496-2500. (poster presentation)

[36] Xueyang Wang, Ryan Stables, Bochen Li, and **Zhiyao Duan**, "Score-aligned polyphonic microtiming estimation," accepted by *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 361-365. (poster presentation)

[35] Sefik Emre Eskimez, **Zhiyao Duan**, and Wendi Heinzelman, "Unsupervised learning approach to feature analysis for automatic speech emotion recognition," accepted by *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 5099-5103. (poster presentation)

[34] Yichi Zhang and **Zhiyao Duan**, "Visualization and interpretation of Siamese style convolutional neural networks for sound search by vocal imitation," accepted *by IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018. (oral presentation)

[33] Rui Lu, **Zhiyao Duan**, and Changshui Zhang, "Multi-scale recurrent neural network for sound event detection," accepted by *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 131-135. (oral presentation)

[32] Lele Chen, Sudhanshu Srivastava, **Zhiyao Duan**, and Chenliang Xu, "Deep cross-modal audio-visual generation," accepted by *ACM Multimedia Thematic Workshops*, 2017. (poster presentation)

[31] Yichi Zhang and **Zhiyao Duan**, "IMINET: convolutional semi-Siamese networks for sound search by vocal imitation," accepted by IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2017. (poster presentation)

[30] Rui Lu, **Zhiyao Duan**, and Changshui Zhang, "Metric learning based data augmentation for environmental sound classification," accepted by IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2017. (oral presentation)

[29] Bochen Li, Karthik Dinesh, Gaurav Sharma, and **Zhiyao Duan**, "Video-based vibrato detection and analysis for polyphonic string music," accepted by International Society for Music Information Retrieval Conference (ISMIR), 2017. (oral presentation) (best paper nomination)

[28] Andrea Cogliati and **Zhiyao Duan**, "A metric for music notation transcription accuracy," accepted by International Society for Music Information Retrieval Conference (ISMIR), 2017. (poster presentation)

[27] Bochen Li, Chenliang Xu, and **Zhiyao Duan**, "Audio-visual source association for string ensembles through

multi-modal vibrato analysis," in *Proc. 14th Sound and Computing Conference (SMC)*, 2017. (oral presentation) (best paper award)

[26] 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," accepted by *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017. (oral presentation)

[25] Karthik Dinesh*, Bochen Li*, Xinzhao Liu, **Zhiyao Duan**, and Gaurav Sharma, "Visually informed multi-pitch analysis of string ensembles," accepted by *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017. (* equal contribution) (poster presentation)

[24] Rui Lu, Kailun Wu, **Zhiyao Duan**, and Changshui Zhang, "Deep ranking: triplet MatchNet for music metric learning," accepted by *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017. (oral presentation)

[23] Sefik Emre Eskimez, Melissa Sturge-Appley, **Zhiyao Duan**, and Wendi Heinzelman, "WISE: web-based interactive speech emotion classification," accepted by 4th Workshop on Sentiment Analysis where AI meets Psychology (SAAIP), 2016. (oral presentation)

[22] Andrea Cogliati, David Temperley, and **Zhiyao Duan**, "Transcribing human piano performances into music notation," in *Proc. International Society for Music Information Retrieval Conference (ISMIR)*, 2016. (poster presentation)

[21] Sefik Emre Eskimez, Kenneth Imade†, Na Yang, Melissa Sturge-Apple, **Zhiyao Duan**, and Wendi Heinzelman, "Emotion classification: How does an automated system compare to naive human coders?," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016. (oral presentation)

[20] Yichi Zhang and **Zhiyao Duan**, "IMISOUND: An unsupervised system for sound query by vocal imitation," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016. (oral presentation)

[19] Andrea Cogliati, **Zhiyao Duan**, Brendt Wohlberg, "Piano music transcription with fast convolutional sparse coding," in *Proc. IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2015. (poster presentation)

[18] Yichi Zhang and **Zhiyao Duan**, "Retrieving sounds by vocal imitation recognition," in *Proc. IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2015. (poster presentation)

[17] Jun Zhou, Shuo Chen, and **Zhiyao Duan**, "Rotational reset strategy for online semi-supervised NMF-based speech enhancement for long recordings," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2015. (poster presentation)

[16] Bochen Li and **Zhiyao Duan**, "Score following for piano performances with sustain-pedal effects," in *Proc. International Society for Music Information Retrieval Conference (ISMIR)*, 2015, pp. 469-475. (poster presentation)

[15] Andrea Cogliati and **Zhiyao Duan**, "Piano music transcription modeling note temporal evolution," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 429-433. (poster presentation)

[14] **Zhiyao Duan** and David Temperley, "Note-level music transcription by maximum likelihood sampling," in *Proc. International Society for Music Information Retrieval Conference (ISMIR)*, 2014, pp. 181-186. (oral presentation)

[13] **Zhiyao Duan**, Bryan Pardo, Laurent Daudet, "A novel cepstral representation for timbre modeling of sound sources in polyphonic mixtures," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 7495-7499. (poster presentation)

[12] Jonathan Springer, **Zhiyao Duan** and Bryan Pardo, "Approaches to multiple concurrent species bird song recognition," in the *2nd International Workshop on Machine Listening in Multisource Environments (CHIME)*, 2013. (poster presentation)

[11] **Zhiyao Duan**, Gautham Mysore and Paris Smaragdis, "Speech enhancement by online non-negative spectrogram decomposition in non-stationary noise environments," in *Proc. Interspeech*, 2012, Portland, Oregon. (oral presentation)

[10] **Zhiyao Duan**, Gautham Mysore and Paris Smaragdis, "Online PLCA for real-time semi-supervised source separation," in *Proc. International Conference on Latent Variable Analysis and Signal Separation (LVA/ICA)*, LNCS 7191, pp. 34-41, 2012. (oral presentation)

[9] **Zhiyao Duan** and Bryan Pardo, "Aligning semi-improvised music audio with its lead sheet," in *Proc. International Society for Music Information Retrieval Conference (ISMIR)*, 2011, pp. 513-518. (poster presentation)

[8] **Zhiyao Duan** and Bryan Pardo, "A state space model for online polyphonic audio-score alignment," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 197-200. (poster presentation)

[7] **Zhiyao Duan**, Jinyu Han and Bryan Pardo, "Song-level multi-pitch tracking by heavily constrained clustering," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* 2010, pp. 57-60. (oral presentation)

[6] **Zhiyao Duan**, Jinyu Han, and Bryan Pardo, "Harmonically informed multi-pitch tracking," in *Proc. International Society for Music Information Retrieval Conference (ISMIR)*, 2009, pp. 333-338. (oral presentation)

[5] **Zhiyao Duan**, Lie Lu, and Changshui Zhang, **"**Collective annotation of music from multiple semantic categories," in *Proc. International Conference on Music Information Retrieval (ISMIR)*, 2008, pp. 237-242. (poster presentation)

[4] **Zhiyao Duan**, Lie Lu, and Changshui Zhang, "Audio tonality mode classification without tonic annotations," in *Proc. International Conference on Multimedia & Expo (ICME)*, 2008, pp. 1361-1364. (poster presentation)

[3] **Zhiyao Duan**, Changshui Zhang, "A maximum likelihood approach to multiple fundamental frequency estimation from the amplitude spectrum peaks," in *Music, Brain and Cognition (MBC) workshop in the Twenty-first Annual Conference on Neural Information Processing Systems (NIPS)*, 2007. (spotlight and poster presentation)

[2] **Zhiyao Duan**, Dan Zhang, Changshui Zhang, and Zhenwei Shi, "Multi-pitch estimation based on partial event and support transfer," in *Proc. International Conference on Multimedia & Expo (ICME)*, 2007, pp.216-219. (poster presentation)

[1] Nelson Lee, **Zhiyao Duan**, and Julius O. Smith, "Excitation signal extraction for guitar tones," in *Proc. International Computer Music Conference (ICMC),* 2007, pp. 450-457.

Patents

[2] Andrea Cogliati, **Zhiyao Duan**, and Brendt Wohlberg, "Context-dependent piano music transcription with convolutional sparse coding," U.S. Patent 9779706, issued in September 2017.

[1] Gautham J. Mysore, Paris Smaragdis, and **Zhiyao Duan**, "Online Source Separation," U.S. Patent US 2013/0121506A1.

INVITED TALKS

[31] Music Learning in the Era of AI

Zhiyao Duan

University of Rochester
720 Computer Studies Building
Rochester, NY 14627

(585) 275-5302
zhiyao.duan@rochester.edu
hajim.rochester.edu/ece/sites/zduan/

Pittsford Community Center — Jul. 2025

[30] *Making Music with AI*
ICME 2025 Workshop on AI for Music Keynote Speech — Jun. 2025
Electroacoustic Music Studios (EMuSE) at Eastman School of Music — Oct. 2025

[28] *Frontiers of Speech Synthesis: Controllability, Expressiveness and Natural Conversations*
Speech and Audio in the North East (SANE) Workshop — Oct. 2024

[27] *AI Powered Interactive Music Making*
AAAI 2025 Workshop on AI for Music — Mar. 2025
Xi'an Jiao Tong University – Xi'an, China — Dec. 2024
Central Conservatory of Music(CCoM) – Beijing, China — Dec. 2024
The 11th China Conference on Sound and Music Technology — Oct. 2024
Boston AI Music Meetup — Oct. 2024

[26] *Some Thoughts about Music Education for Robots*
Dagstuhl Seminar 24302 – Dagstuhl, Germany — Jul. 2024

[25] *A Brief History of Interactive Music Systems* — Aug. 2022
CAAI International Conference on Artificial Intelligence - Art and Artificial Intelligence – Online

[24] *Draw & Listen: A Sketch Based System for Music Inpainting*
Central Conservatory of Music (CCoM) – Online — Mar. 2022

[23] *When Counterpoint Meets Chinese Folk Melodies*
University of Washington NeuroAI Seminar – Online — Jan. 2021

[22] *Towards Human-Computer Collaborative Music Making*
Michigan State University, ECE Department – East Lancing, MI — Sep. 2023
AI Club of Sutherland High School – Pittsford, NY — Mar. 2022
The Summit on Music Intelligence (SOMI), Central Conservatory of Music (CCoM) – Online — Oct. 2021
Invited Keynote Talk, Midwest Music and Audio Day – Bloomington, NY — Jun. 2019

[21] *Computer Audition and Its Potential Application in Digital Health*
International Symposium for Digital Health – Hong Kong — Jun. 2019

[20] *Audio Information Research Lab Overview*
North East Music Information Special Interest Group (NEMISIG) – Ithaca, NY — Feb. 2024
North East Music Information Special Interest Group (NEMISIG) – Newark, NJ — Jun. 2022
North East Music Information Special Interest Group (NEMISIG) – Brooklyn, NY — Feb. 2019
North East Music Information Special Interest Group (NEMISIG) – Rochester, NY — Feb. 2017
North East Music Information Special Interest Group (NEMISIG) – Philadelphia, PA — Feb. 2015
North East Music Information Special Interest Group (NEMISIG) – Ithaca, NY — Feb. 2014

[19] *Audio-Visual Analysis of Music Performances: State and Beyond*
Music and Audio Research Laboratory (MARL), New York University – Brooklyn, NY — Feb. 2019

[18] *Some Thoughts about Singing*

Dagstuhl Seminar 19052 – Dagstuhl, Germany — Jan. 2019

[17] *Deep Learning for Health and Welling Being Diagnosis*
World University Network Meeting on Digital Health – Sydney, Australia — Dec. 2018

[16] *Toward Machine Musicianship*
Upstate New York Sound Meetup – Ithaca, NY — Aug. 2018
Birmingham City University – Birmingham, UK — Aug. 2018

[15] *Multimodal Music Scene Analysis*
Tencent AI Lab – Seattle, WA — May 2018
State University of New York at Fredonia – Fredonia, NY — Mar. 2018
SUSTC, Dept. of Computer Science – Shenzhen, China — May 2017
Fudan University, School of Computer Science – Shanghai, China — May 2017

[14] *Teaching Machines to Listen*
Upstate New York Sound Meetup – Rochester, NY — Aug. 2017
USTC, School of Computer Science and Technology – Hefei, China — May 2017

[13] *Transcribing Piano Music in the Time Domain into Music Notation*
Joint Meeting of the Acoust. Society of America and Acoust. Society of Japan – Honolulu, HI — Dec. 2016

[12] *Towards Complete Music Notation Transcription of Piano*
Western New York Image and Signal Processing Workshop (WNYISPW) – Rochester, NY — Nov. 2016

[11] *The Machine Musicianship: Automatic Music Transcription*
Beihang University, Image Processing Center – Beijing, China — Nov. 2016

[10] *Enriching Sound Interactions through Computer Audition*
Shanghai Jiao Tong University, Dept. of Computer Science and Engineering – Shanghai, China — May 2017
Peking University, Advanced Data & Signal Processing Laboratory – Shenzhen, China — May 2017
Indiana University Bloomington, Department of Computer Science – Bloomington, IN — Sep. 2016

[9] *Retrieving Sounds through Vocal Imitation*
The 3rd Rochester Interdisciplinary Audio Engineering Symposium (RIAES) – Rochester, NY — Aug. 2016
University of Rochester Goergen Institute for Data Science Symposium - Rochester, NY — June 2018

[8] *Computational Music Scene Analysis*
RIT, Center for Applied and Computational Mathematics – Rochester, NY — Mar. 2016
Shanghai University, Department of Mathematics – Shanghai, China — Mar. 2016

[7] *Computational Music Audio Scene Analysis*
Auditory Attention and Scene Analysis workshop and summer school – Delmenhorst, Germany — Jul. 2014

[6] *Note-Level Music Transcription by Maximum Likelihood Sampling*
International Audio Labs Erlangen – Erlangen, Germany — Jul. 2014
1st Rochester Interdisciplinary Audio Engineering Symposium (RIAES) – Rochester, NY — Jun. 2014

[5] *Combining Data-driven and Knowledge-driven Models for Automatic Music Transcription*

Zhiyao Duan

University of Rochester
720 Computer Studies Building
Rochester, NY 14627

(585) 275-5302
zhiyao.duan@rochester.edu
hajim.rochester.edu/ece/sites/zduan/

2nd Midwest Music Information Retrieval Gathering (MMIRG) – Evanston, IL — Jun. 2014

[4] *Transcribing the Pitch Content of Polyphonic Music Audio*
IEEE Signal Processing Society Rochester Chapter IEEE Day Seminar – Rochester, NY — Oct. 2013

[3] *Computer Audition: Analyzing Complex Auditory Scenes*
University of Rochester, Department of Electrical and Computer Engineering – Rochester, NY — Apr. 2013
The Ohio State University, Department of Computer Science and Engineering – Columbus, OH — Mar. 2013
Northwestern University, Department of EECS – Evanston, IL — Jan. 2013

[2] *Music Audio Scene Analysis Informed by a Score*
Ohio State University, Department of Computer Science and Engineering – Columbus, OH — May 2012
Northwestern University, Department of EECS – Evanston, IL — May 2012

[1] *An Approach to Multi-Pitch Tracking of Polyphonic Music*
Dolby Laboratories – Beijing, China — Dec. 2011
Tsinghua University, Department of Automation – Beijing, China — Dec. 2011
Peking University, Institute of Computer Science and Technology – Beijing, China — Dec. 2011
Stanford University, Center for Computer Research in Music and Acoustics – Stanford, CA — Aug.2011

HONORS AND AWARDS

Best Student Paper Award at WASPAA 2025 — Oct. 2025
Best Paper Nomination at ISMIR 2024 — Oct. 2024
Outstanding Reviewer of ICASSP 2023 — Aug. 2023
Top 3% of All Accepted Papers of ICASSP 2023 — April 2023
NSF CAREER Award — Mar. 2019
Best Paper Nomination at ISMIR 2017 — Oct. 2017
Best Paper Award at the 2017 Sound and Music Computing (SMC) Conference — Jul. 2017
Terminal Year Fellowship in Northwestern University — 2012-13
Chinese Government Award for Outstanding Self-Financed Students Abroad — Jun. 2011
Walter P. Murphy Fellowship in Northwestern University — 2008-09
Second-Class Scholarship for Academic Excellent Students of Tsinghua University — 2002-03
Third-Class Scholarship for Academic Excellent Students of Tsinghua University — 2001-02
Third-Class Scholarship for Academic Excellent Students of Tsinghua University — 2000-01

Machine Learning Summer School at Purdue University Scholarship — Jun. 2011
Student Travel Grant for International Society for Music Information Retrieval conference (ISMIR) — 2008, 2010
Excellent Intern in Microsoft Research Asia (MSRA) — Apr. 2008
Champion and Best Control Scheme Prize, Tsinghua University Electronic Design Competition — Dec. 2002

ACADEMIC SERVICE

University of Rochester
Department of Electrical and Computer Engineering Faculty Search Committee Chair — 2024-2025
Steering Committee of the Faculty Council of the College of Arts, Sciences and Engineering — 2024-2026
AME Major Advisor for the Class of 2026 — 2022-2026
Faculty Council of the College of Arts, Sciences and Engineering — 2022-2026
AME Major Advisor for the Class of 2022 — 2018-2022

Zhiyao Duan

University of Rochester
720 Computer Studies Building
Rochester, NY 14627

(585) 275-5302
zhiyao.duan@rochester.edu
hajim.rochester.edu/ece/sites/zduan/

Decanal Review Committee for Hajim School of Engineering and Applied Sciences 2020
ECE Department Graduate Admissions Committee 2014-2018
AME Major Advisor for the Class of 2018 2014-2018
Hajim School Outstanding PhD Dissertation Award Committee 2014-2018
Faculty Search Committee of the Department of ECE 2015-2017
Steering Committee of the Faculty Council of the College of Arts, Sciences and Engineering 2015-2017
Robert L. And Mary L. Sproull University Fellowships Committee 2017

Professional Society

Technical Committee Member, IEEE Signal Processing Society Audio and Acoustic Signal Processing 2025-2027
President, International Society for Music Information Retrieval 2024-2025
President-Elect, International Society for Music Information Retrieval 2022-2023

Journal Editor

Senior Area Editor, IEEE Signal Processing Letters 2024-2026
Associate Editor, IEEE Open Journal of Signal Processing 2022-2023
Topic Editor, Frontiers in Signal Processing 2022
Research Topic on Deep-Learning Based Music Source Separation and Transcription
Guest Editor, Transactions of the International Society for Music Information Retrieval 2021
Special Collection on Cultural Diversity in MIR

Conference Chairing

Scientific Program Co-Chair – International Society for Music Information Retrieval (ISMIR) Conference 2026
Session Chair – Interspeech Conference 2024
Scientific Program Co-Chair – International Society for Music Information Retrieval (ISMIR) Conference 2021
Chair – North East Music Informatics Special Interest Group (NEMISIG) Workshop 2017
Publications Chair - International Society for Music Information Retrieval (ISMIR) Conference 2017
Session Chair - International Society for Music Information Retrieval (ISMIR) Conference 2015

Journal Reviewer

IEEE Transactions on Audio, Speech and Language Processing, IEEE Transactions on Image Processing, IEEE Transactions on Human Machine Systems, IEEE Transactions on Information Forensics and Security, IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Multimedia, IEEE Journal of Selected Topics in Signal Processing, IEEE Multimedia, IEEE Signal Processing Magazine, IEEE Signal Processing Letters, ACM Transactions on Intelligent Systems and Technology, ACM Transactions on Multimedia Computing Communications and Applications, EURASIP Journal on Audio, Speech, and Music Processing, EURASIP Journal on Advances in Signal Processing, Elsevier Computer Science Review, Elsevier Computer Communications, Elsevier Journal on Computer Methods and Programs in Biomedicine, Elsevier Speech Communication, Journal of New Music Research, Music Perception, Neural Processing Letters, Journal of Audio Engineering Society, Frontiers in Psychology, Artificial Intelligence Review, Bulletin of Engineering Informatics, Springer Nature Multimedia Systems

Conference Program Committee / Reviewer

AAAI, ACM Multimedia, AES, Audio Mostly, CSMT, DAFx, EUSIPCO, ICASSP, ICLR, ICME, ICML, Interspeech, ISMIR, NeurIPS, WASPAA, WNYISPW