Michael I Mandel

mim@mr-pc.org · 347-881-6165 · http://mr-pc.org

Current position

Research Scientist, Facebook Reality Labs

Associate Professor, Computer & Information Science, Brooklyn College (CUNY). Faculty in the Computer Science and Linguistics PhD Programs at the CUNY Graduate Center.

Research interests

Signal processing and deep learning for surface electromyography
Deep learning for speech, audio, and music
Noise robust automatic speech recognition and speech enhancement
Very high quality speech enhancement via speech synthesis
Psychoacoustics of speech perception in noise

Education

2010 Feb — PhD with distinction in Electrical Engineering, Columbia University
- Dissertation: "Binaural Model-Based Source Separation and Localization"
- Committee: Daniel Ellis (advisor), Barbara Shinn-Cunningham, Shih-Fu Chang, Richard Stern, Xiaodong Wang

2008 May — MPhil in Electrical Engineering, Columbia University

2006 Feb — MS in Electrical Engineering, Columbia University, GPA: 4.1/4.0

2004 Jun — BS in Computer Science and Engineering, MIT, GPA: 4.9/5.0

Academic positions

2018 – present — **Brooklyn College, CUNY**, Computer & Info. Science, Associate Professor

2015 – 2018 — **Brooklyn College, CUNY**, Computer & Info. Science, Assistant Professor

2016 – present — **Graduate Center, CUNY**, Linguistics PhD Program, Assistant Professor

2015 – present — **Graduate Center, CUNY**, Computer Science PhD Program, Assistant Professor

2015 Jul–Aug — **Jelinek Speech and Language Technologies Workshop**, Far-field Speech team, Senior member

2012 – 2015 — **The Ohio State University**, Computer Science & Eng., Research Scientist

2014 May–Jun — **Télécom ParisTech**, Signal & Image Processing, Visiting professor, AAO Group

2009 – 2010 — **Université de Montréal**, Département d'informatique et de recherche opérationnelle, Postdoctoral researcher, LISA Lab

2004 – 2009 — **Columbia University**, Electrical Engineering, Research Assistant, LabROSA

2008 May–Jun — **Boston University**, Cog. & Neur. Sys., Visiting scholar, Shinn-Cunningham Lab

2003 – 2004 — **MIT**, CS/AI Lab, Undergraduate RA for Prof Bill Freeman

2002 – 2004 — **MIT**, MediaLab, Undergraduate RA for Prof Barry Vercoe

Work experience

2020 – present	**Facebook Reality Labs**, New York, NY, Research Scientist
2018 – 2019	**CTRL-Labs**, New York, NY, Visiting Scientist
2017 Jun–Aug	**HearUNow, Inc**, Woodside, CA, Consultant
2010 – 2012	**Audience, Inc**, Mountain View, CA, Algorithm developer
2009 – 2010	**Musically Intelligent Machines LLC**, New York, NY, Founder, CEO
2007 Jun–Aug	**Google, Inc.**, New York, NY, Software Engineering Intern, Google News
2006 Mar–Sep	**Owl Multimedia**, New York, NY, Co-founder, Dir. Technology
2004 Jun–Aug	**Bose Corporation**, Framingham, MA, Research intern, uMusic™ project

Funding

National Science Foundation Award OPP-1839185, September 2018 – August 2023. " Collaborative Research: Navigating the New Arctic (NNA): Soundscape ecology to assess environmental and anthropogenic controls on wildlife behavior." PI: Michael Mandel. $535,581.

National Science Foundation Award IIS-1750383, June 2018 – May 2023. "CAREER: Integrating perceptual models of auditory importance into deep learning-based noise-robust speech recognition." PI: Michael Mandel. $497,162.

National Science Foundation REU Supplement to Award IIS-1618061. July 2017 – December 2017. $8,000.

Alfred P Sloan Foundation CUNY Junior Faculty Research Award for Science and Engineering (JFRASE). April 2017 – March 2018. $50,000.

National Science Foundation Award IIS-1618061, June 2016 – May 2019. "RI: Small: Concatenative Resynthesis for Very High Quality Speech Enhancement." PI: Michael Mandel. $449,958.

PSC-CUNY Research Award, Trad-B Project #69638-00 47, July 2016, "A game for identifying important speech cues." PI: Michael Mandel. $5,931.

Google Research Award, February 2016, "Incorporating a speech model into multichannel spatial clustering." PI: Michael Mandel. $50,430.

National Endowment for the Humanities Award HD-228966-15, May 2015 – October 2016. "Automatic Music Performance Analysis and Comparison Toolkit (AMPACT): An empirical exploration of expressive musical performance." PI: Johanna Devaney. Co-PI: Michael Mandel. $59,843

National Science Foundation Award IIS-1409431, June 2014 – May 2017. "RI: Medium: Deep Neural Networks for Robust Speech Recognition through Integrated Acoustic Modeling and Separation." PI: Eric Fosler-Lussier, Co-PIs: Michael Mandel and DeLiang Wang. $798,082.

Telecom ParisTech, February 2014, "Learning to recognize sounds for the separation of musical mixtures." PI: Michael Mandel. $5,641.

Google Research Award, August 2013, "Learning to recognize sounds for separation." PI: Michael Mandel. $49,308.

Awards

Top 10% of reviewers NIPS 2018, awarded free registration

Third prize for project "Auditory Bubbles Game", New York City Media Lab Summit, Demo Expo, 2017, $500

Outstanding undergraduate research mentor, Ohio State University, 2013

Postdoctoral research fellowship, Le Fonds québécois de la recherche sur la nature et les technologies, Merit Scholarship Program for Foreign Students 2009–2010, $35,000

Dissertation with distinction, top 10% of Columbia dissertations

Presidential Fellowship, Columbia University School of Engineering and Applied Sciences, 2004–2009, $116,700 plus tuition:

- Sep 2004 – Aug 2005: $30,000 + 2 semesters' tuition
- Sep 2005 – May 2006: $22,500 + 2 semesters' tuition
- Jan 2007 – May 2007: $14,600 + 1 semester's tuition
- Sep 2007 – Aug 2008: $35,000 + 2 semesters' tuition
- Jan 2009 – May 2009: $14,600 + 1 semester's tuition

Second place, Columbia Venture Competition 2009, Columbia University School of Enchineering and Applied Sciences, $7,000

First place, Music Information Retrieval Evaluation eXchange 2008 Audio Artist and Classical Composer Identification task. Tied for first place in Audio Tag Classification task.

First place, Music Information Retrieval Evaluation eXchange 2005 Audio Artist Identification.

Honorable mention, NSF Graduate Research Fellowship Program, 2004.

Top 5% of 180 students in 6.003: Signals and Systems, May 2002.

Emerson Music Scholarship to study saxophone with Jeff Harrington at the Berklee School of Music, 2001–2002 and 2002–2003, $1,200 total.

Publications

Books, Chapters, Theses

M. Mandel, J. Salamon, and D. P. Ellis, eds. *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2019 Workshop (DCASE2019)*. New York University, NY, USA, 2019.

M. I. Mandel, S. Araki, and T. Nakatani. "Multichannel clustering and classification approaches". In: *Audio Source Separation and Speech Enhancement*. Ed. by E. Vincent, T. Virtanen, and S. Gannot. Wiley, 2018. Chap. 12.

M. I. Mandel and J. P. Barker. "Multichannel spatial clustering using model-based source separation". In: *New Era for Robust Speech Recognition: Exploiting, Deep Learning*. Ed. by S. Watanabe, M. Delcroix, F. Metze, and J. R. Hershey. Springer, 2017. Chap. 3.

X. Xiao, S. Watanabe, H. Erdogan, M. Mandel, L. Lu, J. R. Hershey, M. L. Seltzer, G. Chen, Y. Zhang, and D. Yu. "Discriminative beamforming with phase-aware neural networks for speech enhancement and recognition". In: *New Era for Robust Speech Recognition: Exploiting, Deep Learning*. Ed. by S. Watanabe, M. Delcroix, F. Metze, and J. R. Hershey. Springer, 2017. Chap. 4.

J. Devaney, M. I. Mandel, D. Turnbull, and G. Tzanetakis, eds. *Proceedings of the 17th International Society for Music Information Retrieval Conference (ISMIR)*. New York, 2016.

M. I. Mandel. "Binaural Model-Based Source Separation and Localization". PhD thesis. Columbia University, 2010.

T. Bertin-Mahieux, D. Eck, and M. I. Mandel. "Automatic Tagging of Audio: The State-of-the-Art". In: *Machine Audition: Principles, Algorithms and Systems*. Ed. by W. Wang. IGI Publishing, 2010. Chap. 14, pp. 334–352.

Journal

V. A. Trinh and M. I. Mandel. "Directly comparing the listening strategies of humans and machines". In: *IEEE Transactions on Audio, Speech, and Language Processing* 29 (2021), pp. 312–323.

M. I. Mandel, V. Grover, M. Zhao, J. Choi, and V. Shafer. "The Bubble-Noise Technique for Speech Perception Research". In: *Perspectives of the ASHA Special Interest Groups* 4.6 (2019), pp. 1653–1666.

M. I. Mandel, S. E. Yoho, and E. W. Healy. "Measuring time-frequency importance functions of speech with bubble noise". In: *Journal of the Acoustical Society of America* 140 (4 2016), pp. 2542–2553.

H. Larochelle, M. I. Mandel, R. Pascanu, and Y. Bengio. "Learning Algorithms for the Classification Restricted Boltzmann Machine". In: *Journal of Machine Learning Research* 13 (2012), pp. 643–669.

J. Devaney, M. I. Mandel, D. P. W. Ellis, and I. Fujinaga. "Automatically extracting performance data from recordings of trained singers". In: *Psychomusicology: Music, Mind & Brain* 21.1—2 (2012), pp. 108–136.

M. I. Mandel, R. Pascanu, D. Eck, Y. Bengio, L. M. Aiello, R. Schifanella, and F. Menczer. "Contextual tag inference". In: *ACM Transactions on Multimedia Computing, Communications and Applications* 7S.1 (2011), 32:1–32:18.

R. Weiss, M. I. Mandel, and D. P. W. Ellis. "Combining localization cues and source model constraints for binaural source separation". In: *Speech Communication* 53.5 (2011), pp. 606–621.

M. I. Mandel, S. Bressler, B. Shinn-Cunningham, and D. P. W. Ellis. "Evaluating source separation algorithms with reverberant speech". In: *IEEE Transactions on Audio, Speech, and Language Processing* 18.7 (2010), pp. 1872–1883.

M. I. Mandel, R. J. Weiss, and D. P. W. Ellis. "Model-based expectation maximization source separation and localization". In: *IEEE Transactions on Audio, Speech, and Language Processing* 18.2 (2010), pp. 382–394.

M. I. Mandel and D. P. W. Ellis. "A Web-Based Game for Collecting Music Metadata". In: *Journal of New Music Research* 37.2 (2008), pp. 151–165.

T. S. Huang, C. K. Dagli, S. Rajaram, E. Y. Chang, M. I. Mandel, G. E. Poliner, and D. P. W. Ellis. "Active Learning for Interactive Multimedia Retrieval". In: *Proceedings of the IEEE* 96.4 (2008), pp. 648–667.

M. I. Mandel, G. E. Poliner, and D. P. W. Ellis. "Support vector machine active learning for music retrieval". In: *Multimedia systems* 12.1 (2006), pp. 1–11.

Conference

"Towards Large Scale Ecoacoustic Monitoring With Small Amounts of Labeled Data". In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. To appear. 2021.

Z. Ni, Y. Xu, M. Yu, B. Wu, S. Zhang, D. Yu, and M. I. Mandel. “WPD++: an improved neural beamformer for simultaneous speech separation and dereverberation”. In: *IEEE Workshop on Spoken Language Technologies*. 2020.

H. S. Kavaki and M. I. Mandel. “Identifying Important Time-frequency Locations in Continuous Speech Utterances”. In: *Proceedings of Interspeech*. 2020, pp. 1639–1643.

V. A. Trinh and M. I. Mandel. “Large Scale Evaluation of Importance Maps in Automatic Speech Recognition”. In: *Proceedings of Interspeech*. 2020, pp. 1166–1170.

H. Ghaly and M. I. Mandel. “Using Prosody to Improve Dependency Parsing”. In: *Speech prosody*. 2020.

E. B. Çoban, D. Pir, R. So, and M. I. Mandel. “Transfer learning from YouTube soundtracks to tag arctic ecoacoustic recordings”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2020, pp. 726–730.

S. Maiti and M. I. Mandel. “Speaker independence of neural vocoders and their effect on parametric resynthesis speech enhancement”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2020, pp. 206–210. arXiv: 1911.06266.

Z. Ni and M. I. Mandel. “Mask-dependent Phase Estimation for Monaural Speaker Separation”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2020. arXiv: 1911.02746.

S. Maiti and M. I. Mandel. “Parametric Resynthesis with Neural Vocoders”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2019, pp. 303–307. arXiv: 1906.06762.

S. Maiti and M. I. Mandel. “Speech denoising by parametric resynthesis”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2019, pp. 6995–6999.

S. Maiti, J. Ching, and M. I. Mandel. “Large vocabulary concatenative resynthesis”. In: *Proceedings of Interspeech*. 2018, pp. 1190–1194.

V. A. Trinh, B. McFee, and M. I. Mandel. “Bubble cooperative networks for identifying important speech cues”. In: *Proceedings of Interspeech*. 2018, pp. 1616–1620.

A. R. Syed, V. A. Trinh, and M. I. Mandel. “Concatenative resynthesis with improved training signals for speech enhancement”. In: *Proceedings of Interspeech*. 2018, pp. 1195–1199.

S. Maiti and M. I. Mandel. “Concatenative resynthesis using twin networks”. In: *Proceedings of Interspeech*. 2017, pp. 3647–3651.

A. Syed, A. Rosenberg, and M. I. Mandel. “Active Learning for Low-Resource Speech Recognition: Impact of Selection Size and Language Modeling Data”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2017.

J. Devaney and M. I. Mandel. “An evaluation of score-informed methods for estimating fundamental frequency and power from polyphonic audio”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2017.

M. I. Mandel and J. P. Barker. “Multichannel spatial clustering for robust far-field automatic speech recognition in mismatched conditions”. In: *Proceedings of Interspeech*. 2016, pp. 1991–1995.

M. I. Mandel. “Directly comparing the listening strategies of humans and machines”. In: *Proceedings of Interspeech*. 2016, pp. 660–664.

H. Erdogan, J. Hershey, S. Watanabe, M. I. Mandel, and J. L. Roux. “Improved MVDR beamforming using single-channel mask prediction networks”. In: *Proceedings of Interspeech*. 2016, pp. 1981–1985.

X. Xiao, S. Watanabe, H. Erdogan, L. Lu, J. Hershey, M. L. Seltzer, G. Chen, Y. Zhang, M. Mandel, and D. Yu. “Deep Beamforming Networks for Multi-Channel Speech Recognition”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 2016, pp. 5745–5749.

D. Bagchi, M. I. Mandel, Z. Wang, Y. He, A. Plummer, and E. Fosler-Lussier. “Combining spectral feature mapping and multi-channel model-based source separation for noise-robust automatic speech recognition”. In: *Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding*. 2015, pp. 496–503.

M. I. Mandel and Y. S. Cho. “Audio super-resolution using concatenative resynthesis”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2015.

S. S. Tirumala and M. I. Mandel. “Exciting estimated clean spectra for speech resynthesis”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2015.

M. I. Mandel and N. Roman. “Enforcing consistency in spectral masks using Markov random fields”. In: *Proceedings of EUSIPCO*. 2015, pp. 2028–2032.

M. I. Mandel, Y.-S. Cho, and Y. Wang. “Learning a concatenative resynthesis system for noise suppression”. In: *Proceedings of the IEEE GlobalSIP conference*. 2014.

M. I. Mandel, S. E. Yoho, and E. W. Healy. “Generalizing time-frequency importance functions across noises, talkers, and phonemes”. In: *Proceedings of Interspeech*. 2014.

M. I. Mandel and A. Narayanan. “Analysis-by-synthesis feature estimation for robust automatic speech recognition using spectral masks”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2014.

A. Nandi, L. Jiang, and M. I. Mandel. “Gestural Query Specification”. In: *Proceedings of the International Conference on Very Large Data Bases*. Vol. 7. 2014.

M. I. Mandel. “Learning an intelligibility map of individual utterances”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2013.

N. Roman and M. Mandel. “Classification based binaural dereverberation”. In: *Proceedings of Interspeech*. 2013.

J. Devaney, M. I. Mandel, and I. Fujinaga. “A Study of Intonation in Three-Part Singing using the Automatic Music Performance Analysis and Comparison Toolkit (AMPACT)”. In: *Proceedings of the International Society for Music Information Retrieval conference*. Porto, 2012.

J. Devaney, M. I. Mandel, and I. Fujinaga. “Characterizing Singing Voice Fundamental Frequency Trajectories”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2011, pp. 73–76.

M. I. Mandel, D. Eck, and Y. Bengio. “Learning tags that vary within a song”. In: *Proceedings of the International Society for Music Information Retrieval conference*. 2010, pp. 399–404.

J. Bergstra, M. I. Mandel, and D. Eck. “Scalable genre and tag prediction with spectral covariance”. In: *Proceedings of the International Society for Music Information Retrieval conference*. 2010, pp. 507–512.

E. Law, K. West, M. I. Mandel, M. Bay, and J. S. Downie. “Evaluation of algorithms using games: the case of music annotation”. In: *Proceedings of the International Society for Music Information Retrieval conference*. 2009, pp. 387–392.

M. I. Mandel and D. P. W. Ellis. “The ideal interaural parameter mask: a bound on binaural separation systems”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2009, pp. 85–88.

J. Devaney, M. I. Mandel, and D. P. W. Ellis. “Improving MIDI-audio alignment with acoustic features”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2009, pp. 45–48.

R. J. Weiss, M. I. Mandel, and D. P. W. Ellis. “Source separation based on binaural cues and source model constraints”. In: *Proceedings of Interspeech*. Brisbane, Australia, 2008, pp. 419–422.

M. I. Mandel and D. P. W. Ellis. “Multiple-instance learning for music information retrieval”. In: *Proceedings of the International Society for Music Information Retrieval conference*. 2008, pp. 577–582.

D. P. W. Ellis, C. V. Cotton, and M. I. Mandel. “Cross-correlation of beat-synchronous representations for music similarity”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. Las Vegas, NV, 2008, pp. 57–60.

M. I. Mandel and D. P. W. Ellis. “EM localization and separation using interaural level and phase cues”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 2007, pp. 275–278.

M. I. Mandel and D. P. W. Ellis. “A web-based game for collecting music metadata”. In: *Proceedings of the International Society for Music Information Retrieval conference*. Ed. by S. Dixon, D. Bainbridge, and R. Typke. Vienna, Austria, 2007, pp. 365–366.

M. I. Mandel, D. P. W. Ellis, and T. Jebara. “An EM Algorithm for Localizing Multiple Sound Sources in Reverberant Environments”. In: *Advances in Neural Information Processing Systems*. Ed. by B. Schölkopf, J. Platt, and T. Hoffman. Cambridge, MA: MIT Press, 2007, pp. 953–960.

M. I. Mandel and D. P. W. Ellis. “Song-Level Features and Support Vector Machines for Music Classification”. In: *Proceedings of the International Society for Music Information Retrieval conference*. Ed. by J. D. Reiss and G. A. Wiggins. London, UK, 2005, pp. 594–599.

E. B. Sudderth, M. I. Mandel, W. T. Freeman, and A. S. Willsky. “Distributed Occlusion Reasoning for Tracking with Nonparametric Belief Propagation”. In: *Advances in Neural Information Processing Systems*. Ed. by L. K. Saul, Y. Weiss, and L. Bottou. Cambridge, MA: MIT Press, 2005, pp. 1369–1376.

Other

T. Cai, M. I. Mandel, and D. He. “Music Autotagging as Captioning”. In: *First Workshop on NLP for Music and Audio*. Montreal, QC, 2020.

S. Watanabe, M. I. Mandel, J. Barker, and E. Vincent. *CHiME-6 Challenge: Tackling Multispeaker Speech Recognition for Unsegmented Recordings*. 2020. arXiv: 2004.09249.

Z. Ni and M. I. Mandel. “Onssen: an open-source speech separation and enhancement library”. In: 2020, pp. 7269–7273. arXiv: 1911.00982.

L. Mandel, M. I. Mandel, and C. Streb. “Soundscape Ecology: How listening to the environment can shape design and planning”. In: *American Society for Landscape Architects Conference on Landscape Architecture*. San Diego, CA, 2019.

V. Grover, M. I. Mandel, V. Shafer, Y. Syed, and A. Twine. “Understanding Acoustic Cues Non-Native Speakers Use for Identifying English /v/-/w/ Using Bubble Noise Method”. In: *ASHA Convention*. 2018.

H. Ghaly and M. I. Mandel. “Analyzing Human and Machine Performance In Resolving Ambiguous Spoken Sentences”. In: *1st Workshop on Speech-Centric Natural Language Processing (SCNLP)*. 2017, pp. 18–26.

J. Choi and M. I. Mandel. “Perception of Korean fricatives and affricates in ’bubble’ noise by native and nonnative speakers”. In: *International Circle of Korean Linguistics*. 2017.

M. I. Mandel and N. Roman. “Integrating Markov random fields and model-based expectation maximization source separation and localization”. In: *Acoustical Society of America Spring Meeting*. Pittsburgh, PA, 2015.

M. I. Mandel, S. E. Yoho, and E. W. Healy. “Listener consistency in identifying speech mixed with particular “bubble” noise instances”. In: *Acoustical Society of America Spring Meeting*. Pittsburgh, PA, 2015.

M. I. Mandel and S. H. Chon. “Using auditory bubbles to determine spectro-temporal cues of timbre”. In: *Cognitively Based Music Informatics Research (CogMIR)*. Toronto, ON, 2014.

A. Nandi and M. I. Mandel. “The Interactive Join: Recognizing Gestures for Database Queries”. In: *CHI Works-In-Progress*. 2013.

M. Mandel, R. Pascanu, H. Larochelle, and Y. Bengio. *Autotagging music with conditional restricted Boltzmann machines*. 2011. arXiv: 1103.2832.

M. I. Mandel and D. P. W. Ellis. “A Probability Model for Interaural Phase Difference”. In: *ISCA Workshop on Statistical and Perceptual Audio Processing SAPA*. Pittsburgh, PA, 2006, pp. 1–6.

E. B. Sudderth, M. I. Mandel, W. T. Freeman, and A. S. Willsky. “Visual Hand Tracking Using Nonparametric Belief Propagation”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2004, pp. 189–197.

Teaching

CUNY Graduate Center, Computer Science Program

2019 Spring 83060: Speech and Audio Understanding, Instructor, 10 PhD students

2016 Fall 83060: Speech and Audio Understanding, Instructor, 9 PhD students

CUNY Graduate Center, Linguistics Program

2020 Spring 83800: Methods in Computational Linguistics II, Instructor, 17 masters students

2017 Fall 78100: Methods in Computational Linguistics I, Instructor, 12 masters students

Brooklyn College, Department of Computer and Information Science

2020 Spring 3620: Computer Graphics, Instructor, 40 undergraduate students

2019 Spring 3620: Computer Graphics, Instructor, 40 undergraduate students

2018 Fall 7610X: Multimedia databases, Instructor, 13 masters students

2018 Spring 1600: Intro. to Multimedia Computing, Instructor, 40 undergraduate students

2017 Fall 7610X: Multimedia databases, Instructor, 20 masters students

2017 Spring 1600: Intro. to Multimedia Computing, Instructor, 35 undergraduate students

2016 Spring 7610X: Multimedia databases, Instructor, 12 masters students

2016 Spring 1600: Intro. to Multimedia Computing, Instructor, 35 undergraduate students

2015 Fall 1600: Intro. to Multimedia Computing, Instructor, 32 undergraduate students

The Ohio State University, Department of Computer Science and Engineering

2014 Fall 5226: Neural networks, Instructor, 25 masters students

6539: Speech & language reading group, Co-instructor, 16 students

2014 Spring 6539: Speech & language reading group, Co-instructor, 10 students

2013 Fall 6539: Speech & language reading group, Co-instructor, 11 students

2013 July Machine learning, Sennheiser Technology & Innovation Center, Instructor

- 5-day course for 8 Sennheiser employees
- Designed course, created materials, presented lectures and labs

Columbia University, Department of Electrical Engineering

2009 Spring 6820: Speech & audio processing & recognition, Co-lecturer, 7 students

2008 Fall 4810: Digital Signal Processing, Teaching Assistant, 60 masters students

2008 Summer 6820: Speech & audio processing & recognition, Manager, 5 PhD students

2008 Spring 6820: Speech & audio processing & recognition, Co-lecturer, 9 PhD students

Theses supervised

PhD 2020 Jan Hussein Ghaly, "Computational approaches to the syntax-prosody interface: using prosody to improve parsing." CUNY Graduate Center, Linguistics

2018 Feb Min Ma, "Adaptation and augmentation: Towards better rescoring strategies for automatic speech recognition and spoken term detection." CUNY Graduate Center, Computer Science

Masters	2019 Apr	Mengxuan Zhao, "The Perception of Mandarin Tones in "Bubble" Noise by Native and L2 Listeners." CUNY Graduate Center, MA in Computational Linguistics
	2018 Jun	Daniel Chait, "Extract and synthesize: Percussive instruments." Brooklyn College, MA in Computer Science
	2018 Apr	Jiyoung Choi, "Speech perception in 'bubble' noise: Korean fricatives and affricates by native and non-native Korean listeners." CUNY Graduate Center, MA in Computational Linguistics
Bachelors	2019 May	Shelby Ahmed, "Human Listening Cues and Machine Listening." Brooklyn College, Computer and Information Science Honors Thesis.
	2018 May	Oleksandr Loyko, "deep beamforming network for multi-channel speech recognition." Brooklyn College, Computer and Information Science Honors Thesis
	2013 May	Jordan Hawkins, "Automating Music Production with Music Information Retrieval." Ohio State University, Electrical and Computer Engineering, Honors Research Thesis

Students supervised

PhD	2019–2020	Enis Berk Çoban, Hussein Ghaly, Felix Grezes, Zhaoheng Ni, Soumi Maiti, Hassan Salami, Ali Raza Syed, Trinh Viet Anh
	2018–2019	Hussein Ghaly, Felix Grezes, Zhaoheng Ni, Soumi Maiti, Hassan Salami, Ali Raza Syed, Trinh Viet Anh
	2017–2018	Hussein Ghaly, Felix Grezes, Min Ma, Zhaoheng Ni, Soumi Maiti, Ali Raza Syed, Trinh Viet Anh
	2016–2017	Felix Grezes, Min Ma, Zhaoheng Ni, Soumi Maiti, Ali Raza Syed, Trinh Viet Anh
	2014–2015	Young Suk Cho
	2013–2014	Young Suk Cho
MA	2018–2019	Mengxuan Zhao
	2017–2018	Daniel Chait, Jiyoung Choi, Mengxuan Zhao
	2016–2017	Daniel Chait, Jiyoung Choi
	2015–2016	Sreyas Srimath Tirumala
BS	2019–2020	Michael Aleksa, Vadzim Arlou, Amara Auguste, Abdelkader Draou, Minlu Jiang, Alex R Jones, Isaiah Khan, Chao Liang, Xin Liu, Christian Sarcona, Yang Tao, Elsie Urena, Xiaojun Wu
	2018–2019	Shelby Ahmed, Ethan Boiangu, Lauren Burgess, Joey Ching, Hakeem Gayle, Abdullah Gulfam, Xin Liu, Nikita Reshetov, Christian Sarcona, Harry Shomer
	2017–2018	Shelby Ahmed, Eugene Chen, Joey Ching, Xiaowen Huang, Dzmitry Kasinets, Oleksandr Loyko, Max Ohsawa, Christian Sarcona, Max Shteyman, Muhammad Tahir Vali
	2016–2017	Alex Aquino, Eugene Chen, Heriberto Cortes, Renee Esses, Klanti Islam
	2014–2015	Thomas Lyons, Benjamin Oberhaus, Rachel Nelson
	2013–2014	Santosh Kantharaj, Kyle MacNicholas, Austin Mackey, Erik Ringman
	2012–2013	Jordan Hawkins

Invited talks

"High fidelity modeling of speech with neural synthesizers"

2019 Sep 27	Betaworks Render: Hearing Voices

"Navigating the new arctic with ecoacoustics"

2019 Sep 15 Bloomberg Data 4 Good Exchange

"Very high quality speech enhancement via speech synthesis"

2019 Sep 19 CUNY Graduate Center Computer Science Back to School Event
2019 Aug 16 Betaworks Deep Dive
2019 Jul 11 CUNY Summer Undergraduate Research Program
2019 Jun 27 Google Sound Understanding Group Seminar
2019 Jun 23 NYC MediaLab Synthetic Media working group
2019 Jun 11 NYC MediaLab Machines+Media working group

"Building noise robust machine listeners with data and inspiration from humans"

2019 Mar 07 Brooklyn College Biology Seminar
2019 Jan 22 Henan Institute of Science and Technology
2018 Nov 08 CUNY City Tech
2018 Apr 23 Brooklyn College visiting lecture, Computer Assisted Composition
2018 Apr 13 Ohio State University
2018 Apr 03 Northeastern University
2018 Mar 05 New York University

"Noise robust speech processing using strong and weak models"

2017 Sep 07 JP Morgan Chase

"Auditory bubbles: Estimating time frequency importance functions of speech and music"

2016 Dec 16 Spotify
2016 Nov 11 New York University

"Multichannel spatial clustering at the 2015 Jelinek Workshop "

2015 Oct 22 Speech and Audio in the Northeast (SANE) Workshop

"Analysis-by-synthesis for source separation and speech recognition"

2015 Sep 09 Columbia University

"Noise robustness in Automatic Speech Recognition"

2015 Jun 29 Jelinek Speech and Language Technologies Summer School

"Machine learning and optimization in speech analysis-by-synthesis systems"

2015 Apr 01 Telecom ParisTech

"Auditory bubbles: Estimating time frequency importance functions"

2015 Mar 20 CCRMA Hearing Seminar
2014 Jul 11 McMaster University
2014 Jun 25 École Normal Supérieure

"Analysis-by-synthesis for speech recognition and source separation"

2015 Mar 18 Google

"Rich models of digital media: Driving analysis from human perception"

2015 Mar 16 York University

"Detailed models for understanding speech in noise"

2015 Jan 28 University of Illinois, Urbana-Champaign,

2014 Jun 13 Sheffield University
2014 Feb 12 Toyota Technological Institute, Chicago
2014 Feb 07 Mitsubishi Electric Research Labs

"Strong models for understanding sounds in mixtures"
2014 May 02 Queen Mary University London

"Context-dependent models for understanding speech in noise"
2014 Jan 15 Dartmouth Computer Science Colloquium

"Extracting descriptive tags from audio using restricted Boltzmann machines"
2013 Nov 15 CIRMMT Workshop on symbolic music processing, semantic audio, and music information retrieval

"Model-based source separation in reverberant mixtures"
2013 Apr 30 Telecom ParisTech

"Model based source separation"
2012 Oct 06 First Samsung Intenational Symposium on Hearing Aids

"Training automatic music taggers"
2012 Jun 29 CCRMA Music Information Retrieval Workshop 2012
2010 Apr 19 Google

"Evaluating reverberant source separation"
2012 May 11 CCRMA Hearing Seminar

"Automatically describing music"
2009 Nov 25 New York University

"Binaural Model-based Source Separation and Localization"
2009 Oct 16 Drexel University

"MajorMiner: Automatically describing music"
2008 Dec 15 Last.fm
2008 Nov 05 Dorkbot NYC
2008 Nov 04 McGill University, Music Technology Student Colloquium music"

"Model-based EM source separation and localization in reverberant mixtures."
2008 Dec 15 Cambridge University
2008 Dec 10 Sheffield University
2008 Jun 13 Boston University Hearing Research Seminar
2008 Feb 18 Université de Montréal

"EM localization and separation using interaural level and phase cues."
2007 Nov 16 New York University
2007 Oct 09 Université de Montréal

Service

Dept. 2018–2021 BC CIS Appointments Committee
2017–Pres. BC CIS Undergraduate Outcomes Assessment Committee (Chair)
2012–2015 Organizer for the OSU CSE AI Seminar

	2008	Co-founder of the Columbia Electrical Engineering Signal and Information Processing Seminar Series (EESIP SS), 2008 organizer
College	2018–2021	Faculty Council representative, School of Natural & Behavioral Sciences
	2018–Pres.	Undergraduate Admissions Committee, Chair
	2017–2018	Undergraduate Admissions Committee
Univ.	2019–Pres.	Deputy Executive Officer, Computer Science PhD program, CUNY Graduate Center
	2016–Pres.	CUNY Interdisciplinary Research Grant review committee
Field	2017–2020	Associate Editor, Journal of the Acoustical Society of America, Speech Communication area
	2020	Co-Organizer, CHiME-6 Challenge and CHiME-2020 Workshop on Speech Processing in Everyday Environments
	2019	Program Chair, Workshop on Detection and Classification of Acoustic Scenes and Events (DCASE)
	2019	Organizer, North Eastern Music Information Special Interest Group (NEMISIG) Workshop
	2019	ICASSP Area Chair, Audio and Speech Source Separation
	2017	Co-organizer, Speech and Audio in the NorthEast (SANE) Workshop
	2017	Invited participant, IRCAM Workshop on reverse correlation for high-level audio cognition, Paris
	2016–2022	Member, IEEE Technical Committee on Audio and Acoustic Signal Processing
	2016	Publications chair, International Society of Music Information Retrieval Conference
	2014	Lead Organizer and moderator for the panel "The Future of Audio Multimedia" with Gerald Friedland at ACM Multimedia, panelists Dan Ellis, Gerald Friedland, Youngmoo Kim, Josh McDermott, and Paris Smaragdis.
	2011	Publicity chair for the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics
	2008	Co-organizer of the Montreal Music and Machine Learning workshop at the Université de Montréal
	2008	Tutorial and panel chair, ISMIR
	2008	Co-organizer of the Audio Tag Classification task, Music Information Retrieval Evaluation eXchange (MIREX)
	2007	Co-founder of the North Eastern Music Information Special Interest Group (NEMISIG), 2007 co-organizer

Other contributions

Program Committees	2020	Associate for the Advancement of Artificial Intelligence (AAAI) Senior Program Committee member
	2019	International Society for Music Information Retrieval Conference (ISMIR)
	2018	International Society for Music Information Retrieval Conference (ISMIR)
	2018	Machine Learning in Speech and Language Processing Workshop (MLSLP)
	2018	International Conference on Machine Learning (ICML)
	2018	Association for the Advancement of Artificial Intelligence (AAAI)
	2017	International Society for Music Information Retrieval Conference (ISMIR)
	2017	International Conference on Machine Learning (ICML)
	2016	Annual Conference of the International Speech Communication Association (Interspeech)
	2016	Speech Processing in Everyday Environments Workshop at Interspeech

2016 International Society for Music Information Retrieval Conference (ISMIR)
2016 International Conference on Machine Learning (ICML)
2015 IEEE Workshop on Applications of Signal Processing to Audio & Acoustics (WASPAA)

Journal Reviews
- IEEE Transactions in Audio Speech and Language Processing, 2007–19
- IEEE Transactions on Multimedia, 2010–18
- IEEE Transactions on Signal Processing, 2013–18
- ACM Transactions on Knowledge and Data Engineering, 2013
- Computer Speech & Language, 2016–18
- Journal of the Acoustical Society of America Express Letters, 2013–2019
- Journal of the Acoustical Society of America, 2012–2019
- Speech Communication, 2012–19
- EURASIP Journal on Audio, Speech, and Music Processing, 2012–13, 2017
- IEEE Signal Processing Letters, 2010–17

Conference Reviews
- Intl. Conference on Learning Representations (ICLR), 2013–19
- IEEE Intl. Conference on Audio Speech and Signal Processing (ICASSP), 2006–19
- Neural Information Processing Systems (NeurIPS), 2017–19
- Intl. Conference on Machine Learning (ICML), 2013–19
- International Conference on Artificial Intelligence and Statistics (AISTATS), 2017–19
- Annual Conference of the Intl. Speech Communication Association (INTERSPEECH), 2014–2015, 2019
- Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT) 2019
- Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA), 2017
- Intl. Society of Music Information Retrieval Conference, 2006–15
- Intl. Conference on Very Large Data Bases (VLDB) 2013
- IEEE Intl. Conference on Emerging Signal Processing Applications, 2011

Associations
- IEEE Student member 2007–2009, Member 2010–present
- ACM Member 2013–present
- Acoustical Society of America, associate member 2015–present
- Society for Music Theory, joint member, 2015–present

Brooklyn, July 20, 2021