

Scott T. Wisdom, Ph.D.

CONTACT INFORMATION

E-mail: <first name>.thomas.<last name>@gmail.com

Website: stwisdom.github.io

SUMMARY

I am passionate about machine learning, deep learning, statistical models, and signal processing for real-world problems, especially for separation, enhancement, detection, and classification of sound.

EDUCATION

University of Washington, Seattle, Washington, USA

Ph.D., Electrical Engineering, August 2017

- Dissertation Topic: “Improving and Unfolding Statistical Models of Nonstationary Signals”
- Summary: I proposed novel statistical models of complex-valued time-frequency representations of audio signals and created new principled deep and recurrent network architectures by unfolding statistical model inference algorithms. These approaches lead to improved performance on detection of speech and acoustic events, acoustic scene classification, pixel-by-pixel image classification, and audio source separation with single and multiple microphones.
- Committee: Les Atlas (co-advisor), James Pitton (co-advisor), John Hershey, Mari Ostendorf, Mike Seltzer, Marina Meila

University of Washington, Seattle, Washington, USA

M.S. Electrical Engineering, March 2014

- Thesis title: “Improved Statistical Signal Processing of Nonstationary Random Processes Using Time-Warping”
- Summary: I proposed a new method of processing nonstationary signals, where signals are time-warped to make them more stationary. My approach improved detection of sonar signals and enhancement of noisy and reverberated speech.
- Committee: Les Atlas (co-advisor), James Pitton (co-advisor), Rico Malvar

University of Colorado, Boulder, Colorado, USA

B.S. Electrical and Computer Engineering, May, 2010

B.A. English Literature, May, 2010

PROFESSIONAL AND ACADEMIC EXPERIENCE

Google, Cambridge, Massachusetts, USA

Research Scientist

July 2018 - present

- Working as a member of the Sound Understanding team within Google AI Machine Perception, focusing primarily on audio source separation and speech enhancement.

Affectiva, Boston, Massachusetts, USA

Senior Speech Research Scientist

September 2017 - June 2018

- Formulated, implemented, and trained statistical models and deep networks for voice activity detection, multispeaker detection and speaker counting, and noise-robust speech emotion recognition.
- Supervised gathering and labeling of real-world audio data.
- Interfaced with data, product, engineering, marketing, and computer vision teams to define and implement requirements.

Interactive Systems Design Lab, University of Washington, Seattle, Washington, USA

Research Assistant

June, 2012 - September 2017

- Performed novel research in machine learning, deep learning, and statistical signal processing, especially applied to audio. Authored and co-authored over a dozen peer-reviewed papers.
- As most senior graduate student, mentored other graduate students in the lab and helped them develop their academic, technical, and professional skills.
- Maintained the lab Linux server for general-purpose GPU computing.

Jelinek Summer Workshop on Speech and Language, Seattle, Washington, USA

Graduate Student Team Member

June, 2015 - August 2015

- Member of the Far-Field Speech Recognition team, explored new methods to make speech recognition more robust to mismatched training data, especially in the presence of unknown noise and reverberation. Results are now published as a book.
- Created a new type of deep network architecture that directly processes complex-valued, frequency-domain, multi-microphone audio data for improved separation of speech sources.

Mitsubishi Electric Research Labs (MERL), Cambridge, Massachusetts, USA

Research Intern

January, 2015 - March 2015

- Worked with John Hershey, Jonathan Le Roux, and Shinji Watanabe on deep unfolding of generative speech models for multichannel source separation of reverberant and noisy speech.
- Derived and implemented a variety of existing and novel speech separation methods in Matlab.

Microsoft Research, Redmond, Washington, USA

Research Intern

June, 2014 - September 2014

- Worked with Jie Liu and the Sensing and Energy Research Group (SERG) on signal processing for wireless signal localization.
- Came up to to speed quickly on state-of-the-art methods for localization, designed a new method for indoor localization, and ran extensive experiments.

Sensor Systems Lab, University of Washington, Seattle, Washington, USA

Research Assistant

June, 2011 - December, 2012

- Worked with Josh Smith on the Wireless Resonant Energy Link (WREL) project, a wireless power system for moving devices.
- Helped design hardware, wrote embedded firmware in C, and authored a Python interface.
- Designed a robust tracking algorithm (now patented) to maintain maximum wireless energy transfer regardless of receivers position and orientation.

Fluke Networks, Everett, Washington, USA

Software Engineer in Advanced Engineering Program (AEP)

July, 2010 - June, 2012

- As part of AEP, a career acceleration program for high-potential recent graduates, worked on a wide variety of solo and team projects building innovative network test tools for copper, fiber, and wireless media.

PUBLICATIONS

E. Tzinis, S. Wisdom, J.R. Hershey, A. Jansen, D.P.W. Ellis, “Improving Universal Sound Separation Using Sound Classification”, in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Barcelona, Spain, May 2020.

S. Sonning, C. Schultdt, H. Erdogan, S. Wisdom, “Performance Study of a Convolutional Time-Domain Audio Separation Network for Real-Time Speech Denoising”, in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Barcelona, Spain, May 2020.

I. Kavalerov, S. Wisdom, H. Erdogan, B. Patton, K. Wilson, J. Le Roux, and J.R. Hershey, “Universal Sound Separation”, in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, New York, USA, October 2019.

S. Wisdom, J.R. Hershey, K. Wilson, J. Thorpe, M. Chinen, B. Patton, and R.A. Saurous, “Differentiable Consistency Constraints for Improved Deep Speech Enhancement”, *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Brighton, UK, May 2019.

Le Roux, J., S. Wisdom, H. Erdogan, and J.R. Hershey, “SDR-half-baked or well done?”, in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Brighton, UK, May 2019.

S. Wisdom, T. Powers, J. Pitton, and L. Atlas, Deep Recurrent NMF for Speech Separation by Unfolding Iterative Thresholding, in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, New York, USA, October 2017.
(Best student paper award, WASPAA student travel grant awarded)

J. R. Hershey, J. Le Roux, S. Watanabe, S. Wisdom, Z. Chen, Y. Isik, “Novel deep architectures in speech processing,” in *New Era for Robust Speech Recognition: Exploiting Deep Learning*, S. Watanabe, M. Delcroix, F. Metze, J. R. Hershey, Eds. Springer, 2017.

S. Wisdom, T. Powers, J. Pitton, and L. Atlas, Building Recurrent Networks by Unfolding Iterative Thresholding for Sequential Sparse Recovery, in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, New Orleans, Louisiana, USA, March 2017.

S. Wisdom, T. Powers, J.R. Hershey, J. Le Roux, and L. Atlas, “Full-Capacity Unitary Recurrent Neural Networks,” in *Advances in Neural Information Processing Systems (NeurIPS)*, Barcelona, Spain, Dec 2016. **(NeurIPS student travel grant awarded)**

S. Wisdom, T. Powers, J. Pitton, and L. Atlas, “Interpretable Recurrent Neural Networks Using Sequential Sparse Recovery,” in *Proc. NeurIPS 2016 Workshop on Interpretable Machine Learning in Complex Systems*, Barcelona, Spain, December 2016. **(Student travel grant awarded)**

T. Powers, J. Bilmes, S. Wisdom, D. Krout, and L. Atlas, “Constrained Robust Submodular Optimization,” in *Proc. NeurIPS Workshop on Optimization for Machine Learning*, Barcelona, Spain, Dec 2016.

S. Wisdom, L. Atlas, J. Pitton, and G. Okopal “Benefits of Noncircular Statistics for Nonstationary Signals,” in *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov 2016. **(Invited paper)**

S. Wisdom, L. Atlas, and J. Pitton, “On Spectral Noncircularity of Natural Signals,” in *Proc. IEEE Sensor and Multichannel (SAM)*, Rio de Janeiro, Brazil, July 2016. **(Invited paper)**

S. Wisdom, J.R. Hershey, J. Le Roux, and S. Watanabe, “Deep Unfolding for Multichannel Source Separation,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Shanghai, China, March 2016.
(Travel grants awarded by NSF, IEEE SPS, and UW)

G. Okopal, S. Wisdom, and L. Atlas “Speech Processing with the Strong Uncorrelating Transform,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Nov 2015.

S. Wisdom, G. Okopal, L. Atlas, and J. Pitton, “Voice Activity Detection Using Subband Non-circularity,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*

S. Wisdom, J. Pitton, and L. Atlas, “Extending Coherence for Optimal Detection of Nonstationary Harmonic Signals,” in *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov 2014.

S. Wisdom, T. Powers, L. Atlas, and J. Pitton, “Enhancement of Reverberant and Noisy Speech by Extending Its Coherence,” in *Proc. REVERB Challenge Workshop*. Florence, Italy, May 2014.

S. Wisdom, L. Atlas, and J. Pitton, “Extending Coherence Time for Analysis of Modulated Random Processes,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

A.P. Sample, B. Waters, S. Wisdom, and J.R. Smith, "Enabling Seamless Wireless Power Delivery in Dynamic Environments," *Proc. IEEE*, vol. 101, no. 6, pp. 1343-1358, June 2013.

PATENTS A.P. Sample, J.R. Smith, B.H. Waters, S. Wisdom, “Enabling Seamless Wireless Power Delivery in Dynamic Environments.”

TECHNICAL SKILLS

- Programming languages: Python, Matlab, C/C++, bash scripting.
- Other proficiencies: Keras, TensorFlow, Theano, Linux, Git, embedded systems.
- Algorithms: Experience deriving and programming implementations of deep learning, inference algorithms for probabilistic graphical models (exact and variational), speech processing, convex optimization, beamforming, statistical estimation and detection.

ACADEMIC SERVICE

- Member IEEE Audio and Acoustic Signal Processing (AASP) Technical Committee (2019-2022).
- Reviewer for ICASSP, NeurIPS, ICML, MLSP, CHI, IEEE/ACM Transactions on Signal Processing, IEEE/ACM Transactions on Audio, Speech, and Language Processing, IEEE Signal Processing Letters, Journal of the Acoustical Society of America, Computer Speech and Language, Neural Computation, EURASIP Journal on Advances in Signal Processing, Speech Communication, GlobalSIP, UbiComp.

INVITED TALKS		
	Dolby SoHo, New York City, New York	April 2019
	“Advances in Enhancing Noisy Speech, and Beyond”	
	Mitsubishi Electric Research Labs (MERL), Cambridge, Massachusetts	March, 2018
	“Theory and Applications of Sparse Model-Based Recurrent Neural Networks”	
	Google, Cambridge, Massachusetts	November, 2017
	“Unitary Recurrent Neural Networks and Deep Recurrent Nonnegative Matrix Factorization”	
	Pacific Northwest National Laboratory, Seattle, Washington	March, 2017
	“Unfolding Statistical Inference Algorithms into Better Deep Networks”	
	Microsoft Research, Redmond, Washington	April, 2016
	“Improving and Unfolding Statistical Models for Multichannel Audio”	
	Mitsubishi Electric Research Labs (MERL), Cambridge, Massachusetts	January, 2016
	“Unfolding Generative Models for Multichannel Source Separation”	

Amazon, Seattle, Washington	June, 2015
“Noncircularity: A New Feature for Speech Processing”	
Mitsubishi Electric Research Labs (MERL), Cambridge, Massachusetts	September, 2014
“Beyond Locally Stationary and Circular”	
NATO Centre for Maritime Research & Experimentation (CMRE), La Spezia, Italy	May, 2014
“Improved Statistical Signal Processing by Extending Coherence Time”	