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 speech and audio.

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 representatinos 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

Affectiva, Boston, Massachusetts, USA

Senior Speech Research Scientist **September 2017 - present**

- Formulating, implementing, and training statistical models and deep networks for voice activity detection, multispeaker detection and speaker counting, noise-robust speech emotion recognition, source separation, and acoustic event classification.
- Supervising gathering and labeling of real-world audio data.
- Interfacing with 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 both novel and existing speech separation methods using 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 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**

- Worked closely with R&D engineering team building innovative network test tools for copper, fiber, and wireless media.
- As part of AEP, a career acceleration program for high-potential recent graduates, worked on a wide variety of solo and team projects, including implementing a 802.11 receiver on an FPGA as part of an ambitious new product prototype and writing a new Linux kernel display driver that fixed a critical path product issue.

Selected Conference Publications

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)

S. Wisdom, T. Powers, J. Pitton, and L. Atlas, Building Recurrent Networks by Unfolding Iterative Thresholding for Sequential Sparse Recovery, *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, New Orleans, Lousiana, 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 (NIPS)*, Barcelona, Spain, Dec 2016. **(NIPS student travel grant awarded)**

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, UW)**

S. Wisdom, G. Okopal, L. Atlas, and J. Pitton, "Voice Activity Detection Using Subband Noncircularity," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Brisbane, Austrailia, April 2015.

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.

Journal & Book Publications

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.

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.

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

Patents

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

Technical Skills

- Programming languages: Python, Matlab, C/C++, bash scripting, Scheme.
- Other proficiencies: Keras, TensorFlow, Theano, Linux, Git version control, 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

Reviewer for NIPS, ICASSP, IEEE/ACM Transactions on Signal Processing, IEEE/ACM Transactions on Audio, Speech, and Language Processing, IEEE Signal Processing Letters, EURASIP Journal on Advances in Signal Processing, Speech Communication, GlobalSIP, UbiComp.

Invited Talks

Google, Cambridge, Massachusetts, USA — November, 2017
"Unitary Recurrent Neural Networks and Deep Recurrent Nonnegative Matrix Factorization"

Pacific Northwest National Laboratory, Seattle Washington, USA — March, 2017
"Unfolding Statistical Inference Algorithms into Better Deep Networks"

Microsoft Research, Redmond, Washington, USA — April, 2016
"Improving and Unfolding Statistical Models for Multichannel Audio"

Amazon, Seattle, Washington, USA — June, 2015
"Noncircularity: A New Feature for Speech Processing"

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