JOHN WALTER McDONOUGH

Carnegie Mellon University
Language Technologies Institute
GHC 5411
5000 Forbes Avenue
Pittsburgh, PA, USA 15213
johnmcd@cs.cmu.edu

5535 Ellsworth Avenue
Pittsburgh, PA, USA 15232
john.mcdonough3141@gmail.com
http://distant-speech-recognition.org
412 952 2881 (mobile)

EXPERIENCE

Visiting Scientist   Carnegie Mellon University
Pittsburgh, PA   January, 2011 to present

- *Professional Activities:*
  - Involved in proposal writing efforts for HP Labs Innovation Research Program and the National Science Foundation.
  - Serving as publications chairman for WASPAA 2011; see **http://www.waspaa.com**.

- *Research:*
  - Continuing research in acoustic array processing and distant speech recognition, with emphasis on voice prompt suppression (VPS), and multi-speaker voice activity detection.
  - Have begun to investigate integration of adaptive beamforming techniques with spherical microphone arrays.
  - Developing techniques to segment and align extended multi-speaker conversations.
  - Working with other researchers at the CMU-LTI to integrate distant speech recognition and natural language processing (NLP) components.

Visiting Scientist   Disney Research, Pittsburgh
Pittsburgh, PA   February, 2010 to January, 2011

- *Research:*
  - Conducted research in all areas related to distant speech recognition, including acoustic array processing, voice activity detection, and voice prompt suppression.
  - Began research on audio-visual speaker tracking and focus of attention recognition.
  - Integrated research components into an online demo.

- *Project Management:*
  - Founded research effort in distant speech recognition.
  - Oversaw two large-scale data collection campaigns.
  - Recruited team of researchers and scientists.

Researcher   Saarland University
Saarbrücken, Germany   February, 2007 to Feburary, 2010

- *Research:*
  - Conducted research on acoustic array processing and weighted finite state transducers for distant speech recognition.

- Developed Python extension modules for manipulation of finite state transducers, beamforming and speaker tracking.

- *Teaching:*

  - Established and held the lecture *Distant Speech Recognition*;
    see **http://www.lsv.uni-saarland.de/e_distant_ws0910.htm**.

  - Established and held the lecture *Weighted Finite-State Transducers in Speech and Natural Language Processing*; see **http://www.lsv.uni-saarland.de/e_finite_ss09.htm**.

- *Project Management:*

  - Served as lead author in successful proposal writing effort for the European Union SCALE (Speech Communication with Adaptive LEarning) project, which supports 12 graduate students and 2 postdocs at 6 academic sites.

  - Organized first plenary meeting of SCALE, the SCALE Winter School 2010.

- *Evaluation Campaigns:*

  - Led the winning team for the Speech Separation Challenge, Part II.

  - Participated in the NIST RT 2009 Evaluation.

Researcher   Universität Karlsruhe
Karlsruhe, Germany   January, 2000 to December, 2006

- *Research:*

  - Conducted research on speaker adaptation, discriminative training, and weighted finite state transducers for automatic speech recognition.

  - Founded and led research effort into the use of microphone arrays for person tracking and distant speech recognition.

  - Developed Python extension modules in C++ (approximately 50,000 lines of source code) for automatic speech recognition, manipulation of finite state transducers, beamforming and source localization.

- *Teaching:*

  - Held the lecture *Verbal Man-Machine Communication* in the German language;
    see **http://isl.ira.uka.de/sprachVorlesung**.

  - Established and taught *Microphone Arrays: Gateway to Hands-Free Automatic Speech Recognition*; see **http://isl.ira.uka.de/ma2005**.

- *Project Management:*

  - Led research in audio technologies and coordinated data collection at the Universität Karlsruhe (UKA) for the European Union integrated project CHIL.

  - Served as site leader for the European Union project NESPOLE on speech-to-speech translation.

- *Evaluation Campaigns:*

  - Led participation of UKA in NIST RT05 meeting evaluation for the acoustic source localization task.

  - Coordinated participation of UKA in the NIST RT06 meeting evaluations for the speech-to-text task.

  - Participated in the CLEAR 2006 evaluation in the single and multiple speaker acoustic tracking tasks, as well as the audio-video person tracking task.

Scientist   BBN Corporation
Cambridge, Massachusetts   January 1993 to August 1997

- Participated in successful Switchboard evaluation efforts in 1995 through 1997.

- Developed theory of and implemented speaker-adapted training.

EDUCATION

Doctor of Philosophy   The Johns Hopkins University
Baltimore, Maryland   September 1997 to April 2000

- Developed theory of speaker adaptation based on all-pass transforms.

- Implemented and maintained public domain extensions to HTK, the hidden Markov model toolkit.

- Participated in the Johns Hopkins Center for Language and Speech Processing Summer Workshop in 1997 and 1998.

PUBLICATIONS

- *Books and Tutorials:*

  - **Robust Automatic Speech Recognition**, B. Raj, R. Singh and T. Virtanen (eds.), Wiley, to appear.
  - **Single and Multi Channel Feature Enhancement for Distant Speech Recognition**, tutorial with F. Faubel and M. Wölfel at *ICASSP*, March, 2010.
  - **Distant Speech Recognition**, with M. Wölfel, Wiley, 2009.
  - **Computers in the Human Interaction Loop**, A. Waibel and R. Stiefelhagen (eds.), Springer, ISBN 1848820534, 2009.
  - **Machine Learning for Multimodal Interaction**, S. Bengio and H. Bourlard (eds.), Springer Lecture Notes in Computer Science, Volume 4892/2008, pgs. 283-294, 2008.

- *Distant Speech Recognition:*

  - "Maximum Kurtosis Beamforming with a Subspace Filter for Distant Speech Recognition," with K. Kumatani, and B. Raj, submitted to *Proc. ASRU*, December, 2011.
  - "An Information Filter for Voice Prompt Suppression," with W. Chu, K. Kumatani, B. Raj and J. F. Lehman, to appear in *Proc. Asilomar*, November, 2011.
  - "On the Combination of Voice Prompt Suppression with Maximum Kurtosis Beamforming," with K. Kumatani and B. Raj, to appear in *Proc. WASPAA*, October, 2011.
  - "Channel Selection based on Multichannel Cross-Correlation Coefficients for Distant Speech Recognition," with K. Kumatani, J. F. Lehman and B. Raj, in *Proc. HSCMA*, May, 2011.
  - "Distant Speech Recognition: No Black Boxes," with M. Wölfel, K. Kumatani, B. Rauch, F. Faubel, and D. Klakow, in *Proc. ITG-Fachtagung*, October, 2008.
  - "Distant Speech Recognition: Bridging the Gaps," with M. Wölfel and E. Stoimenov, in *Proc. HSCMA*, May, 2008.

- *Weighted Finite-State Transducers:*

  - "An Algorithm for the Fast Composition of Weighted Finite-State Transducers," with E. Stoimenov and D. Klakow, *Proc. ASRU*, December 2007.
  - "Memory Efficient Modeling of Polyphone Context with Weighted Finite-State Transducers," with E. Stoimenov, *Proc. Interspeech*, August, 2007.
  - "Modeling Polyphone Context with Weighted Finite-State Transducers," with E. Stoimenov, winner of the *Proc. ICASSP* Best Student Paper Award, May, 2006.

- *Speaker Adaptation:*

  - "Performance Comparisons of All-Pass Transform Adaptation with Maximum Likelihood Linear Regression," with A. Waibel, in *Proc. ICASSP*, 2004.

  - "Speaker Adaptation with All-Pass Transforms," with T. Schaaf and A. Waibel, *Speech Communication Special Issue on Adaptation Methods in Speech Recognition*, January, 2004.

  - "Speaker Compensation with Sine-Log All-Pass Transforms," with F. Metze, H. Soltau and A. Waibel, in *Proc. ICASSP*, 2001.

  - "Speaker Compensation with Sine-Log All-Pass Transforms," Ph.D. Thesis, the Johns Hopkins University, Baltimore, Maryland, April, 2000.

  - "On the Incremental Addition of Regression Classes for Speaker Adaptation," with W. Byrne, in *Proc. ICASSP*, 2000.

  - "Single-Pass Adapted Training with All-Pass Transforms," with W. Byrne, in *Proc. Eurospeech*, 1999.

  - "Speaker Adaptation with All-Pass Transforms," with W. Byrne, in *Proc. ICASSP*, May, 1999.

  - "Speaker Normalization with All-Pass Transforms," with W. Byrne and X. Luo, in *Proc. ICSLP*, Sydney, Australia, May, 1998.

- *Discriminative HMM Training:*

  - "On Maximum Mutual Information Speaker-Adapted Training," with M. Wölfel, and E. Stoimenov, *Computer Speech and Language*, vol. 22, pp. 130-147, 2008.

  - "Maximum Mutual Information Speaker Adapted Training with Semi-Tied Covariance Matrices," with A. Waibel, in *Proc. ICASSP*, May, 2003.

  - "On Maximum Mutual Information Speaker Adapted Training," with T. Schaaf and A. Waibel, in *Proc. ICASSP*, Orlando, Florida, May, 2002.

- *Robust Feature Extraction:*

  - "Particle Filter Based Soft-Mask Estimation for Missing Feature Reconstruction," with F. Faubel, H. Raja, and D. Klakow, in *Proc. IWAENC*, September, 2008.

  - "A Phase-Averaged Model for the Relationship between Noisy Speech, Clean Speech and Noise in the log-Mel Domain," with F. Faubel and D. Klakow, in *Proc. IWAENC*, September, 2008.

  - "Minimum Variance Distortionless Response Spectral Estimation: Review and Refinements," with M. Wölfel, *IEEE Signal Processing Magazine Special Issue on Speech Technology and Systems in Human-Machine Communication*, September, 2005.

  - "Minimum Variance Distortionless Response on a Warped Frequency Scale," with M. Wölfel, in *Proc. Eurospeech*, 2003.

  - "Warping and Scaling of the Minimum Variance Distortionless Response," with M. Wölfel and A. Waibel, in *Proc. ASRU*, 2003.

- *Evaluation System Descriptions:*

  - "Advances in Lecture Recognition: The ISL RT-06S Evaluation System," with C. Fügen, M. Wölfel, S. Ikbal, F. Kraft, K. Laskowski, M. Ostendorf, S. Stüker and K. Kumatani, *Proc. Interspeech*, September, 2006.

  - "Multi-Source Far-Distance Microphone Selection and Combination for Automatic Transcription of Lectures," with M. Wölfel, C. Fügen and S. Ikbal, *Proc. Interspeech*, September, 2006.

  - "Tracking Multiple Speakers with Probabilistic Data Association Filters," with T. Gehrig, in *Proc. CLEAR Evaluation Workshop*, September, 2006.

- *Person Tracking:*

  - "Tracking and Beamforming for Multiple Simultaneous Speakers with Probabilistic Data Association Filters," with T. Gehrig, U. Klee, S. Ikbal, M. Wölfel and Christian Fügen, in *Proc. Interspeech*, September, 2006.

  - "Tracking and Far-Field Speech Recognition for Multiple Simultaneous Speakers," with T. Gehrig, in *Proc. MLMI*, April, 2006.

  - "Kalman Filters for Time Delay of Arrival-Based Source Localization," with U. Klee and T. Gehrig, *EURASIP Journal of Advanced Signal Processing, Special Issue on Multichannel Speech Processing*, January, 2006

  - "Kalman Filters for Audio-Video Source Localization," with T. Gehrig, K. Nickel, H. K. Ekenel, and U. Klee, in *Proc. WASPAA*, October, 2005.

- *Acoustic Array Processing:*

  - "Towards Online Maximum Kurtosis Beamforming," with K. Kumatani and B. Raj, to appear in *Proc. WASPAA*, October, 2011.

  - "Maximum Negentropy Beamforming with Superdirectivity," with K. Kumatani, L. Lu, A. Ghoshal, and D. Klakow, in *Proc. EUSIPCO*, October, 2010.

  - "Maximum Kurtosis Beamforming with the Generalized Sidelobe Canceller," with K. Kumatani, B. Rauch, P. Garner, and W. Li, in *Proc. Interspeech*, September, 2008.

  - "On Hidden Markov Model Maximum Negentropy Beamforming," with B. Rauch, K. Kumatani, F. Faubel, and D. Klakow, September, in *Proc. IWAENC*, 2008.

  - "Adaptive Beamforming with a Maximum Negentropy Criterion," with K. Kumatani, J. McDonough, D. Klakow, P. N. Garner, W. Li, in *Proc. HSCMA*, May, 2008.

  - "Adaptive Beamforming with a Minimum Mutual Information Criterion," with K. Kumatani, T. Gehrig, U. Mayer, E. Stoimenov, and M. Wölfel, *IEEE Trans. Audio, Speech, Language Processing*, November, 2007.

  - "Microphone Array Driven Speech Recognition: Influence of Localization on the Word Error Rate," with M. Wölfel and K. Nickel, in *Proc. MLMI*, March, 2005.

  - "Combining Multi-Source Far Distance Speech Recognition Strategies: Beamforming, Blind Channel and Confusion Network Combination," with M. Wölfel, in *Proc. Interspeech*, September, 2005.

  - "A Cepstral Domain Maximum Likelihood Beamformer for Speech Recognition," with D. Raub and M. Wölfel, in *Proc. Interspeech*, October, 2004.

PERSONAL REFERENCES

1. Dr. Mark Gales (`mjfg@eng.cam.ac.uk`), Reader in Information Engineering, Cambridge University, Cambridge, UK.

2. Dr. Mari Ostendorf (`mo@ee.washington.edu`), Professor of Electrical Engineering, University of Washington, Seattle, Washington, USA.

3. Dr. Bhiksha Raj (`bhiksha@cs.cmu.edu`), Associate Professor of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.

4. Dr. Richard Stern (`rms@cs.cmu.edu`), Professor of Electrical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.