Meng YU

yumeng.ssym@gmail.com • Tel: 949-648-8557

EMPLOYMENT

Tencent AI Lab

Principal Researcher • Bellevue, WA Sep 2016 – Present

- Lead research and engineering on front-end speech and audio projects
- Far field speech processing and enhancement system supporting various types of microphone arrays
- Data driven R&D for single-channel and multi-channel speech enhancement, separation, echo suppression, VAD and localization
- Joint training & optimization of front-end speech enhancement and recognition
- Acoustic scene detection, audio classification, spatial audio generation

Audience Inc.

Staff Algorithm Engineer • Mountain View, CA Mar 2013 – Aug 2016

- R&D for voice processing and enhancement in various acoustic environments
- Data driven algorithms for voice processing
- Multi-channel acoustic echo cancellation and noise suppression algorithms
- Target speech enhancement for automatic speech recognition assist

Cisco Systems Inc.

Software Engineer • Milpitas, CA Jun 2012 – Mar 2013

- R&D for automatically generating speaker metadata for an ever growing set of videos with minimal user input
- Noise and reverberation robustness for speaker segmentation and recognition

EDUCATION:

University of California, Irvine • Irvine, CA Sep 2007 – Jun 2012

Ph.D. in Applied Mathematics: blind source separation and speech signal processing

Peking University • Beijing China Sep 2003 – Jun 2007

B.S. in Applied Mathematics

SKILLS

- Programming languages: C, C++, Java, Python, Matlab
- Language: Chinese (native), English (fluent)

ACADEMIC REVIEW

- IEEE Transactions on Speech Audio and Language Processing
- IEEE Transactions on Signal Processing
- EURASIP Journal on Audio, Speech, and Music Processing
- Speech Communication
- Journal of Computer Science and Technology
- Communications in Mathematical Sciences
- Mathematical Reviews
- ACM Multimedia
- International Speech Communication Association (Interspeech)

- IEEE workshop on application of signal processing to audio and acoustics
- IEEE Digital Signal Processing Workshop

PUBLICATIONS

Peer-reviewed Journal Publications

- [5] Audio-Visual Speech Separation and Dereverberation with a Two-Stage Multimodal Network, K.Tan, Y. Xu, S. Zhang, M. Yu, and D. Yu, IEEE Journal of Selected Topics in Signal Processing, 2020
- [4] DurIAN: Duration Informed Attention Network for Multimodal Synthesis, C. Yu, H. Lu, N. Hu, M. Yu, C. Weng, K. Xu, P. Liu, D. Tuo, S. Kang, G. Lei, D. Su, and D. Yu, arXiv: 1909.01700
- [3] Multi-channel l1 regularized convex speech enhancement model and fast computation by the split Bregman method, M. Yu, W. Ma, J. Xin and S. Osher, IEEE Tran. on Audio, Speech and Language Proc., 20(2), pp. 661-675, 2012
- [2] A convex model and l1 minimization for musical noise reduction in blind source separation, M. Yu, W. Ma, J. Xin and S. Osher, Communications in Mathematical Sciences, 10(1), pp 223- 238, 2012.
- [1] Stochastic approximation and a nonlocally weighted soft-constrained recursive algorithm for blind separation of reverberant speech mixtures, M. Yu and J. Xin, vo. 28, no. 4, Discrete and Continuous Dynamical Systems, 2010.

Peer-reviewed International Conference Papers

- [36] End-to-End Multi-Look Keyword Spotting, M. Yu, X. Ji, B. Wu, D. Su, D. Yu, submitted to Interspeech 2020
- [35] Neural Spatio-Temporal Beamformer for Target Speech Separation, Y Xu, M Yu, SX Zhang, L Chen, C Weng, J Liu, D Yu, arXiv preprint arXiv:2005.03889
- [34] Speech signal processing model training method, electronic device and storage medium, L. Chen, M Yu, M Luo, D Su, US Patent App. 16/655,548
- [33] Integration of multi-look beamformers for multi-channel keyword spotting, X. Ji, M. Yu, J. Chen, J. Zheng, D. Su, D. Yu, ICASSP, 2020
- [32] Speaker-aware target speaker enhancement by jointly learning with speaker embedding extraction, X. Ji, M. Yu, C. Zhang, D. Su, T. Yu, X. Liu, D. Yu, ICASSP, 2020
- [31] Enhancing end-to-end multi-channel speech separation via spatial feature learning, R. Gu, SX Zhang, L. Chen, Y. Xu, M. Yu, D. Su, Y. Zou, D. Yu, ICASSP, 2020
- [30] Far-field location guided target speech extraction using end-to-end speech recognition objectives, A. Subramanian, C. Weng, M. Yu, SX Zhang, Y. Xu, S. Watanabe, D. Yu, ICASSP, 2020
- [29] Overlapped speech recognition from a jointly learned multi-channel neural speech extraction and representation, B Wu, M Yu, L Chen, C Weng, D Su, D Yu, arXiv preprint arXiv:1910.13825
- [28] A Unified Framework for Speech Separation, F Bahmaninezhad, SX Zhang, Y Xu, M Yu, JHL Hansen, D Yu, arXiv preprint arXiv:1912.07814
- [27] End-to-End Multi-Channel Speech Separation, R. Gu, J. Wu, S. Zhang, L. Chen, Y. Xu, M. Yu, D. Su, Y. Zou, D. Yu, arXiv:1905.06286, 2019
- [26] Time Domain Audio Visual Speech Separation, J. Wu, Y. Xu, S. Zhang, L. Chen, M. Yu, L. Xie, D. Yu, ASRU, 2019
- [25] Syllable-Dependent Discriminative Learning for Smal Footprint Text-Dependent Speaker Verification. J. Peng, Y. Zou, N. Li, D. Tuo, D. Su, M. Yu, C. Zhang, D. Yu, ASRU, 2019
- [23] Improving Speech Enhancement with Phonetic Embedding Features, B. Wu, M. Yu, L. Chen, M. Jin, D. Su, D. Yu, ASRU, 2019
- [23] Jointly Adversarial Enhancement Training for Robust End-to-End Speech Recognition, B. Liu, S. Nie, S. Liang, W. Liu, M. Yu, L. Chen, S. Peng and C. Li, Interspeech, 2019
- [22] Direction-aware Speaker Beam for Multi-channel Speaker Extraction, G. Li, S. Liang, S. Nie, W. Liu, M. Yu, L. Chen, S. Peng and C. Li, Interspeech, 2019

- [21] Neural Spatial Filter: Target Speaker Speech Separation Assisted with Directional Information, R. Gu, L. Chen, S. Zhang, J. Zheng, Y. Xu, M. Yu, D. Su, Y. Zou, D. Yu, Interspeech, 2019
- [20] A comprehensive study of speech separation: spectrogram vs waveform separation, F. Bahmaninezhad, J. Wu, R. Gu, S. Zhang, Y. Xu, M. Yu, and D. Yu, Interspeech, 2019
- [19] Improved Speaker-Dependent Separation for CHiME-5 Challenge, J. Wu, Y. Xu, S. Zhang, L. Chen, M. Yu, L. Xie, D. Yu, Interspeech, 2019
- [18] Boundary Discriminative Large Margin Cosine Loss for Text-independent Speaker Verification, R. Li, N. Li, D. Tuo, M. Yu, D. Su, and D. Yu, ICASSP, 2019
- [17] Joint Training of Complex Ratio Mask Based Beamformer and Acoustic Model for Noise Robust Asr, Y. Xu, C. Weng, L. Hui, J. Liu, M. Yu, D. Su and D. Yu, ICASSP, 2019
- [16] Seq2Seq Attentional Siamese Neural Networks for Text-dependent Speaker Verification, Y. Zhang, M. Yu, N. Li, C. Yu, J. Cui and D. Yu, ICASSP, 2019
- [15] Multi-band PIT and Model Integration for Improved Multi-channel Speech Separation, L. Chen, M. Yu, Y. Qian, D. Su, and D. Yu, ICASSP, 2019
- [14] Deep extractor network for target speaker recovery from single channel speech mixtures, J. Wang, J. Chen, D. Su, L. Chen, M. Yu, Y. Qian, D. Yu, Interspeech, 2018
- [13] Text-Dependent Speech Enhancement for Small-Footprint Robust Keyword Detection, M. Yu, X. Ji, Y. Gao, L. Chen, J. Chen, J. Zheng, D. Su, and D. Yu, Interspeech, 2018
- [12] Permutation Invariant Training of Generative Adversarial Network for Monaural Speech Separation, L. Chen, M. Yu, Y. Qian, D. Su, and D. Yu, Interspeech, 2018
- [11] Method and System for Reducing Interference and Noise in Speech Signals, M. Yu, J. Hershey, Mitsubishi Electric Research Lab, Grant Date: 06-02-2015, US Patent Number: 9048942
- [10] Speech dereverberation by constrained and regularized multi-channel spectral decomposition: evaluated on REVERB challenge, M. Yu, F. K. Soong, REVERB Challenge Workshop 2014
- [09] Constrained multichannel speech dereverberation, M. Yu, F. K. Soong, Interspeech 2012
- [08] Exploring off-time nature for speech enhancement, M. Yu, J. Xin, Interspeech 2012
- [07] A triple-microphone real-time speech enhancement algorithm based on approximate array analytical solutions, M. Yu, R. Ritch, J. Xin, Interspeech 2012
- [06] Exploring feature collocation for semantic concept identication, M. Yu, pp. 94 - 95, ICCM, 2012
- [05] Identication of semantic categories: a sparse representation approach, S. Zhang, M. Yu, M. Lee and J. Xin, CogSci 2011
- [04] Convexity and fast speech extraction by split Bregman method, M. Yu, W. Ma, J. Xin and S. Osher, pp. 398 - 401, Interspeech 2010.
- [03] Reducing musical noise in blind source separation by time-domain sparse lters and split Bregman method, M. Yu, W. Ma, J. Xin and S. Osher, pp. 402 - 405, Interspeech 2010.
- [02] A nonlocally weighted soft-constrained natural gradient algorithm and blind separation of strong reverberant speech mixtures, M. Yu, J. Xin, Y. Qi, H. Yang and F-G Zeng, pp. 346-350, 43rd Asilomar Conference on Signals, Systems, and Computers 2009.
- [01] A nonlocally weighted soft-constrained natural gradient algorithm for blind separation of reverberation speech, J. Xin, M. Yu, Y. Qi, H. Yang and F-G Zeng, pp. 81 - 84, WASPAA, 2009.