Curriculum Vitae

Yu Tsao (曹昱)

Assistant Research Fellow
Bio-Acoustic Signal Processing (Bio-ASP) Lab
Research Center for Information Technology Innovation, Academia Sinica
128 Academia Rd., Sec. 2, Nankang District, Taipei 115, Taiwan
TEL: +886-22-2787-2390
yu.tsao@citi.sinica.edu.tw
http://www.citi.sinica.edu.tw/pages/yu.tsao/

RESEARCH INTERESTS

- Speech and Audio Signal Processing
- Hearing Assistive Technology (Design, Practice, and Evaluation)
- Deep Neural Network for Speech and Audio Signal Processing
- Multimedia Signal and Information Processing

EDUCATION

GEORGIA INSTITUTE OF TECHNOLOGY, Atlanta, Georgia
Ph.D. in Electrical and Computer Engineering   Aug. 2003-Dec. 2008
- Major: Digital Signal Processing; Minor: Statistics
- Research Topic: Robust Speech Recognition, under advisor: Dr. Chin-Hui Lee
- Leadership Activities: President, Taiwanese Student Association (TSA): represented the school; assisted students and visiting scholars from Taiwan.

NATIONAL TAIWAN UNIVERSITY, Taipei, Taiwan
Master of Science in Electrical Engineering   Sept. 1999-June 2001
Bachelor of Science in Electrical Engineering   Sept. 1995-June 1999
- Graduate Research Topic: Rapid Speaker Adaptation, under advisor Dr. Lin-Shan Lee
- Leadership Activities: Chairman, Public Relations Department of the Student Association: organized events to cultivate strong relationships among members.

PROFESSIONAL EXPERIENCE

RESEARCH CENTER FOR INFORMATION TECHNOLOGY INNOVATION, ACADEMIA SINICA, Taipei, Taiwan
Assistant Researcher Fellow   Nov. 2011-Present
- Research and develop speech, speaker and language recognition algorithms.
- Establish algorithms for acoustic and language modeling to handle mismatch caused by noises.
- Derive multimedia signal and information processing for audio and speech applications.
- Apply pattern recognition and machine learning theories on audio event detection tasks.

NATIONAL INSTITUTE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY, Kyoto, Japan
Expert Researcher   April 2009-Sept. 2011
- Developed researches to handle non-native accent issues for automatic speech recognition applications.
- Developed digital signal processing based solutions to improve the performance of speech recognizer under real-world adverse conditions.
- Contributed to a development of VoiceTra multilingual speech to speech translation application on IPhone.
- Carried out projects of field tests and dissemination of spoken dialog interface technologies.

GEORGIA INSTITUTE OF TECHNOLOGY, Atlanta, Georgia, USA
Research Assistant   Aug. 2003-March 2009
- Conducted research for the Automatic Speech Attribute Transcription (ASAT) Program.
- Established detection-based automatic speech recognition framework on a smaller vocabulary database.
- Proposed a novel model-based noise robustness approach called Ensemble Speaker and Speaking Environment Modeling (ESSEM); enhanced performance robustness by characterizing noise interference.

TEXAS INSTRUMENTS INCORPORATED, Dallas, Texas, USA
Intern, Speech Technologies Laboratory DSP Solutions R&D Center   Summers 2004, 2005, 2006
- Contributed to a patent-pending project to develop high-performance and low-computational cost algorithms to improve efficiency and accuracy of automatic speech recognition on devices (2006).
- Reduced mismatch caused by speaker variations for mobile speech recognition applications; devised algorithms for the project (2005).
- Enhanced performance under noisy conditions for mobile speech recognition applications (2004).

APPLIED SPEECH TECHNOLOGIES (AST) COMPANY, Taipei, Taiwan
Intern   May 2000-July 2001
- Calculated algorithms that helped enhance the performance and reduce computational cost for a Mandarin-based speech recognition product called Mega Voice Mandarin Recognition System.

NATIONAL TAIWAN UNIVERSITY, Taipei, Taiwan
Research Assistant, Graduate Institute of Communication Engineering   Sept. 1999-June 2001
- Analyzed and developed segmental eigenvoice speaker adaptation approach by studying rapid speaker adaptation; created algorithms for Chinese syllables and tone adjustments for Mandarin recognition.

DEVELOPED TOOLS

A smartphone-based assistive listening system, **SmartHear**, using wireless technologies for individuals with mild-to-moderate hearing loss. **SmartHear** has been developed as a **mobile application program** and is available on **Google Play**: https://play.google.com/store/apps/details?id=van1211.mycode.test2pair&hl=zh_TW (since 2015, there are more than **10,000 downloads**). A demo **video** about **SmartHear** can be viewed at https://www.youtube.com/watch?v=e9HqIj09QJs. **SmartHear** has been repored in the Central News Agency (CNA): http://www.cna.com.tw/postwrite/Detail/188673.aspx#.VtLmmfmqqkr.

HONORS

- Top three team in BigMM 2016 Challenge, 2016.
- Co-advised Student, T.-H. Wen, Received the best paper award nomination, Interspeech 2013.
- Excellent Paper Award, Technologies and Applications of Artificial Intelligence (TAAI), 2012.
- Co-advised Student, H.-T. Hwang, Received the Interspeech 2012 Travel Grant Award, 2012.
- Fellowship from Texas Instruments Leadership University Program (TILU), 2005-2008.

PROFESSIONAL ACTIVITIES

- Special Session Chair, ISCSLP 2016, Tianjin, China (2016/10).
- Workshop Co-organizer, IEEE BigMM 2016, Taiwan (2016/05).
- Co-organizer, Taiwan-Japan Symposium on Psychological, Physiological, and Electro-Acoustics, National Tsing Hua University, Taiwan (2015/10).
- Session Chair, ROCLING 2015, National Central University, Taiwan (2015/10).
- Organizer, SWS 2015, IIS, Academia Sinica, Taiwan (2015/03).
- Special Session Chair, ICASSP 2015, South Brisbane, Queensland, Australia (2015/04).
- Session Chair, ROCLING 2014, National Central University, Taiwan (2014/09).
- Session Chair, ROCLING 2013, National Sun Yat-Sen University, Taiwan (2013/10).
- Special Session Chair, APSIPA 2013, Kaohsiung, Taiwan (2013/10).
- Organizer, NGASR Workshop 2013, CITI, Academia Sinica, Taiwan (2013/09).
- Session Chair, ROCLING 2012, Yuan Ze University, Taiwan (2012/09).
- Committee Members, 2012 IEEE SPS Summer School, National Tsing Hua University, Taiwan (2012/07).
- Session Chair, ISCSLP 2010, National Cheng Kung University, Taiwan (2010/11).

INVITED TALKS

- Deep Learning in Speech, Deep Learning Workshop in Taiwan, Academia Sinica, 2015.
- Intelligibility-oriented speech enhancement using GMAPA and maximum speech intelligibility algorithm, Taiwan-Japan Symposium on Psychological, Physiological, and Electro-Acoustics, 2015.
- Machine learning and signal processing technologies for assistive hearing devices, MediaTek Inc., 2015.
- Machine learning and signal processing technologies for assistive hearing devices, National Taiwan University, 2015.
- Some challenges and recent advances in speech technology, National Taiwan University of Science and Technology, 2014.
- Recent advances in robust speech recognition, speaker and event recognition, and voice conversion, National Tsing Hua University, 2013.
- Some challenges and recent advances in speech technology, National Sun Yat-Sen University, 2013.
- Some challenges and recent advances in speech technology, National Taiwan Normal University, 2013.
- Some challenges and recent advances in speech technology, National Central University, 2013.
- Recent advances in robust speech recognition, speaker and acoustic event recognition, and voice conversion, National Chiao Tung University, 2013.
- Challenges and recent advances in robust speech recognition, speaker and event recognition, and text classification, Nara Institute of Science and Technology, Japan, 2013.
- Some challenges and recent advances in speech technology, The University of Hong Kong, 2012.
- Some challenges and recent advances in speech technology, College of Information Science and Technology, Xinjiang University, 2012.
- Increasing discrimination on the ensemble speaker and speaking environment modeling approach for robust ASR, Tokyo Institute of Technology, Tokyo, Japan, 2009.
- MAP-based online mapping function estimation for ensemble speaker and speaking environment modeling approach, NTT Communication Science Laboratories, Kyoto, Japan, 2009.

PUBLICATIONS

(1) Journal Papers

Published

[1] Y.-H. Lai, F. Chen, X. Lu, **Y. Tsao***, and C.-H. Lee, "A deep denoising autoencoder approach to improving the intelligibility of vocoded speech in cochlear implantation," IEEE Transactions on Biomedical Engineering (SCI), accepted to appear.

[2] S.-H. Fang, W.-H. Chang, **Y. Tsao***, H.-C. Shih, and C. Wang, "Channel state reconstruction using multi-level discrete wavelet transform for improved fingerprinting-based indoor localization," IEEE Sensors Journal, 2016.

[3] S.-H. Fang , H.-H. Liao , Y.-X. Fei , K.-H. Chen , J.- W. Huang , Y.-D. Lu and **Y. Tsao**, "Transportation modes classification using sensors on smartphones," Sensors, 2016.

[4] P. Lin, S.-W. Fu, Y.-H. Lai, and **Y. Tsao***, "Maximum entropy learning for deep belief networks," Entropy (SCI), 2016.

[5] T.-E. Chen, S.-I. Yang, L.-T. Ho, K.-H. Tsai, Y.-H. Chen, Y.-F. Chang, Y.-H. Lai, and **Y. Tsao***, C.-C. Wu, "S1 and S2 heart sound recognition using deep neural networks," IEEE Biomedical Engineering (SCI), 2016.

[6] S.-S. Wang, A. Chern, **Y. Tsao***, J.-w. Hung, X. Lu, Y.-H. Lai, and B. Su "Wavelet speech enhancement based on nonnegative matrix factorization," IEEE SPL (SCI), 2016.

[7] H.-L. S. Wang, I-C. Chen, C.-H. Chiang, Y.-H. Lai, and **Y. Tsao***, "Auditory perception, suprasegmental speech processing, and vocabulary development in Chinese preschoolers," Perceptual and Motor Skills(SCI), 2016.

[8] F. Chen, **Y. Tsao**, Y.-H. Lai, "Modeling speech intelligibility with recovered envelope from temporal fine-structure stimulus," Speech Communication (SCI), 2016.

[9] **Y. Tsao*** and Y.-H. Lai, "Generalized maximum a posteriori spectral amplitude algorithm for speech enhancement," Speech Communication (SCI) , 2016.

[10] S.-H. Fang, C.-H. Wang, **Y. Tsao***, "Compensating for orientation mismatch in robust Wi-Fi localization using histogram equalization," IEEE Transaction on TVT (SCI), 2015.

[11] Y.-C. Lin, Y.-H. Lai, H.-W. Chang, **Y. Tsao***, Y.-p. Chang, and R. Y. Chang, "SmartHear: A smartphone based remote microphone haring assistance system using wireless technologies," IEEE System Journal (SCI). 2015.

[12] J. Li-You, Y.-R. Chien, and **Y. Tsao**, "Rapid converging M-max partial update least mean square algorithms with new variable step-size methods," IEICE Transactions on Communications (SCI). 2015.

[13] Y.-H. Lai, Fe. Chen, and **Y. Tsao***, "Effects of adaptation rate and noise suppression on the intelligibility of compressed-envelope based speech," PLoS One (SCI), 2015.

[14] C.-C. Hsu, K.-M. Cheong, T.-S. Chi, **Y. Tsao**, "Robust voice activity detection algorithm based on feature of frequency modulation of harmonics and its DSP," IEICE Transactions on Information and Systems (SCI), 2015.

[15] **Y. Tsao***, P. Lin, T.-y. Hu, and X. Lu, "Ensemble environment modeling using affine transform groups for robust automatic speech recognition," Speech Communication (SCI), 2015.

[16] **Y. Tsao**, S.-H. Fang, and Y. Shiao, "Acoustic echo cancellation using a vector-space-based adaptive filtering algorithm," IEEE Signal Processing Letter (SCI), 2015.

[17] **Y. Tsao***, T.-y. Hu, S. Sakti, S. Nakamura, and L.-s. Lee, "Variable selection linear regression for robust speech recognition," IEICE Transactions on Information and Systems (SCI) 2014.

[18] **Y. Tsao***, X. Lu, P. Dixon, T.-y. Hu, S. Matsuda, and C. Hori, "Incorporating local information of the acoustic environments to MAP-based feature compensation and acoustic model adaptation," Computer Speech and Language (SCI), 2014.

[19] **Y. Tsao***, S. Matsuda, C. Hori, H. Kashioka, and C.-H. Lee, "A MAP-based online estimation approach to ensemble speaker and speaking environment modeling," IEEE Transaction on ASLP (SCI), 2014.

[20] Y.-H. Lai, **Y. Tsao***, and F. Chen, "A study of adaptive WDRC in hearing aids under noisy conditions," International Journal of Speech & Language Pathology and Audiology, 2013, (invited paper).

[21] **Y. Tsao** and C.-H. Lee, "An ensemble speaker and speaking environment modeling approach to robust speech recognition," IEEE Transactions on ASLP (SCI), 2009.

[22] **Y. Tsao**, S.-M. Lee, and L.-S. Lee, "Segmental eigenvoice with delicate eigenspace for improved speaker adaptation," IEEE Transactions on SAP (SCI), 2005.

Under Revision or Submitted
[23] H.-y. Lee, B.-H. Tseng, T.-H. Wen, and **Y. Tsao**, "Personalizing recurrent neural network based language model by social network," submitted to IEEE Transaction on ASLP (SCI) (Decision: Major Revision).

[24] Y.-T. Tsai, B. Su, **Y. Tsao**, and S.-S. Wang, "Robust beamforming against DoA mismatch using subspace-constrained diagonal loading," submitted to IEEE Snesors (SCI).

[25] S.-W. Hsiao, H.-C. Sun, M.-C. Hsieh, M.-H. Tsai, H.-C. Lin, **Y. Tsao**, and C.-C. Lee, "A computational framework for automatic scoring of pre-service school principal's oral presentation using fusion of audio-visual information," submitted to IEEE Transaction on Affective Computing (SCI).

[26] Y.-H. Lai, **Y. Tsao**, X. Lu , Y.-T. Su, J. K.-C. Chen, M.-J. Lien, S. Y. Chen, Z.-M.Yeh, Y.-X. Chen, L. P.-H. Li, C.-H. Lee, "A deep learning–based noise reduction approach to improve the speech intelligibility for cochlear implantees under challenging listening conditions," submitted to PlosOne (SCI).

(2) Selected International Conference Papers

Acoustic modeling and speech processing
[1] J.-C. Hou, S.-S. Wang, Y.-H. Lai, J.-C. Lin, **Y. Tsao**, H.-W. Chang, and H.-M. Wang, "Audio-visual speech enhancement using deep neural networks," to appear in APSIPA 2016

[2] S.-W. Fu, **Y. Tsao**, X. Lu, "SNR-aware convolutional neural network modeling for speech enhancement," Interspeech 2016.

[3] S.-S. Wang, H.-T. Hwang, Y.-H. Lai, **Y. Tsao**, X. Lu, H.-M. Wang, and B. Su, "Improving denoising auto-encoder based speech enhancement with the speech parameter generation algorithm," APSIPA 2015.

[4] H.-T. Hwang, **Y. Tsao**, H.-M. Wang, Y.-R. Wang, and S.-H. Chen, "A probabilistic interpretation for artificial neural network-based voice conversion," APSIPA 2015.

[5] P. Lin, D.-C. Lyu, Y.-F. Chang, and **Y. Tsao**, "Temporal alignment for deep neural networks" GlobalSIP 2015.

[6] X. Lu, P. Shen, **Y. Tsao**, C. Hori, H. Kawai, "Sparse representation with temporal max-smoothing for acoustic event detection," Interspeech 2015.

[7] P. Lin, D.-C. Lyu, Y.-F. Chang, and **Y. Tsao**, "Speech recognition with temporal neural networks," Interspeech2015.

[8] W.-C. Chen, P.-T. Lai, **Y. Tsao**, and C.-C. Lee, "Multimodal arousal rating using unsupervised fusion technique," ICASSP 2015.

[9] Y.-H. Lai, S.-S. Wang, P.-C. Li, and **Y. Tsao**, "A discriminative post-filter for speech enhancement in hearing aids," ICASSP 2015.

[10] H.-S. Lee, **Y. Tsao**, H.-M. Wang and S.-K. Jen, "Clustering-based I-vector formulation for speaker recognition," Interspeech 2014.

[11] H. Jing, T.-Y. Hu, H.-S. Lee, W.-C. Chen, C.-C. Lee, **Y. Tsao** and H.-M. Wang, "Ensemble of machine learning algorithms for cognitive and physical speaker load detection," Interspeech 2014.

[12] P. Lin, F. Chen, S.-S. Wang, **Y. Tsao**, and Y. H. Lai, "Automatic speech recognition with primarily temporal envelope information," Interspeech 2014.

[13] X, Lu, **Y, Tsao**, S, Matsuda and C, Hori, "Ensemble modeling of denoising autoencoder for speech spectrum restoration," Interspeech 2014,

[14] X. Lu, **Y. Tsao**, S. Matsuda, C. Hori, "Sparse representation based on a bag of spectral exemplars for acoustic event detection," ICASSP 2014.

[15] H.-t. Fan, J.-w. Hung, X. Lu, S.-S. Wang, **Y. Tsao**, "Speech enhancement using segmental nonnegative matrix factorization," ICASSP 2014.

[16] H.-T. Hwang, **Y. Tsao**, H.-M. Wang, Y.-R. Wang and S.-H. Chen, "Alleviating the over-smoothing problem in GMM-based voice conversion with discriminative training," Interspeech 2013.

[17] T.-H. Wen, A. Heidel, H.-y. Lee, **Y. Tsao** and L.-S. Lee, "Recurrent neural network based language model personalization by social network crowdsourcing," Interspeech 2013.

[18] X. Lu, **Y. Tsao**, S. Matsuda and C. Hori, "Speech enhancement based on deep denoising autoencoder," Interspeech 2013.

[19] B. Li, **Y. Tsao** and K. C. Sim, "An investigation of spectral restoration algorithms for deep neural networks based noise robust speech recognition," Interspeech 2013.

[20] S.-S. Wang, **Y. Tsao**, J.-W. Hung, "Filtering on the temporal probability sequence in histogram equalization for robust speech recognition," ICASSP 2013.

[21] Y.-C. Su, **Y. Tsao**, J.-E. Wu, F.-R. Jean, "Speech enhancement using generalized maximum a posteriori spectral amplitude estimator," ICASSP 2013.

Assistive listening device
[1] Y.-T. Liu, R. Y. Chang, **Y. Tsao**, and Y.-p. Chang, "A new frequency lowering technique for Mandarin-speaking hearing aid users," GlobalSIP 2015.

[2] Y.-T. Liu, **Y. Tsao**, and R. Y. Chang, "A deep neural network based approach to Mandarin consonant/vowel separation," ICCE 2015.

[3] Y.-H. Lai, F. Chen, and **Y. Tsao**, "Effect of adaptive envelope compression in simulated electric hearing in reverberation," ISIC 2014.

[4] Y. H. Lai, F. Chen, and **Y. Tsao**, "An adaptive envelope compression strategy for speech processing in cochlear implants," Interspeech 2014,

[5] Y.-H. Lai, Y.-C. Su, **Y. Tsao**, S.-T. Young, "Evaluation of generalized maximum a posteriori spectral amplitude (GMAPA) speech enhancement algorithm in hearing aids," ISCE 2013.

Machine learning
[1] H. Jing, A.-C. Liang, S.-D. Lin, and **Y. Tsao**, "A transfer probabilistic collective factorization model to handle sparse data in collaborative filtering," ICDM 2014 (acceptance rate=9.5%)

[2] H. Jing, **Y. Tsao**, K.-Y. Chen and H.-M. Wang, "Semantic naive Bayes classifier for document classification," IJCNLP 2013.

[3] H.-y. Lee, T.-y. Hu, H. Jing, Y.-F. Chang, **Y. Tsao**, Y.-C. Kao and T.-L. Pao, "Ensemble of machine learning and acoustic segment model techniques for speech emotion and autism spectrum disorders recognition," Interspeech 2013.

[4] H. Jing and **Y. Tsao**, "Sparse maximum entropy deep belief nets," IJCNN 2013.

(3) Patents

[1] S, Matsuda, S. Nakamura, and **Y. Tsao**, "音響モデルの話者適応装置及びそのためのコンピュータプログラム," JP patent, 2013/09.
[2] K. Yao and **Y. Tsao**, "Efficient speech recognition with cluster methods," US patent publication, 2008/12.
[3] **Y. Tsao** and Y.-C. Su, "generalized maximum a posteriori spectral amplitude estimator for speech enhancement," TW patent, 2016/01.
[4] **Y. Tsao**, S.-H. Fang, Y.-R. Chien, and Y. Shiao, "acoustic echo cancellation method and system using prior knowledge of the acoustic conditions," US patent, submitted 2014, result pending.
[5] B. Su, Y.-T. Tsai, **Y. Tsao**, and S.-S. Wang, "System and method for beamforming design against direction-of-arrival mismatch," US patent, submitted 2016, result pending.

TECHNOLOGY TRANSFERS

[1] ASUS, "Speech enhancement for distant automatic speech recognition," signing contract.
[2] Advanced and Wise (A&W) Technology, "Speech enhancement technologies for mobile devices," 2015/11.
[3] iMediPlus Inc., "Bio-acoustic signal enhancement," 2016/01.