

SYSU-CMU Joint Institute of Engineering, SYSU, East campus, Guangzhou

SYSU-CMU Shunde International Joint Research Institute, Shunde

liming46@mail.sysu.edu.cn

mingli1@cmu.edu

<http://jie.sysu.edu.cn/~mli/>

Assistant Professor

Sep 2013 ~ present

Sun Yat-Sen University Carnegie Mellon University Joint Institute of Engineering, Sun Yat-sen University, China

Sun Yat-Sen University Carnegie Mellon University Shunde International Joint Research Institute, China

Associate Professor

School of electronics and information technology, Sun Yat-Sen University, China

2016 ~ present

School of mobile information engineering, Sun Yat-Sen University, China

2013 ~ 2015

Adjunct Professor

2015 ~ present

ECE department of Carnegie Mellon University, Pittsburgh, USA

EDUCATION&ACADEMIA

Carnegie Mellon University

Pittsburgh, USA

Visiting Professor in ECE department

Sep 2013 ~ July 2014

http://www.ece.cmu.edu/directory/department/faculty/L/Ming_Li_8171.html

University of Southern California

Los Angeles, USA

Ph.D. in Electrical Engineering

2008-2013

- [Signal Analysis and Interpretation Laboratory](#)
- Advisor: [Shrikanth Narayanan](#)
- Provost Fellowship, Research Assistant

Institute of Acoustics, Chinese Academy of Sciences

Beijing, China

Master of Engineering in Signal and Information Processing

2005-2008

- [ThinkIT Speech Laboratory](#)
- Advisor: [Yonghong Yan](#)
- Research Assistant

Nanjing University

Nanjing, China

Bachelor of Science in Telecommunications engineering

2001-2005

- Graduate with the highest honor
- Rank: top1 within 50 students in communication engineering major

INDUSTRY

Qualcomm Corporate R&D

San Diego, USA

Summer Intern

2011 summer

My current research interests:

My research interests lie in the areas of multimodal signal processing, speech and language processing, machine learning, statistical modeling, affective computing, structure health monitoring. My work is to enhance the robustness and efficiency of the multimodal human state recognition tasks which cover a broad range of applications, notably in security, healthcare and user assistance. Multimodal human state recognition can be considered as a task encompassing term for identifying or verifying various kinds of human centered state labels from multimodal signals, both overtly expressed and covertly present. We can also use signal processing and machine learning techniques for structure health monitoring.

Human state recognition:

1. **Speech signal processing:** Speech recognition, speaker verification, spoken language identification, speaker diarization, speaker age and gender identification, emotion recognition
2. **Multimodal biometrics:** Audio-visual joint biometrics, emerging behavior biometrics (ECG biometrics), multimodal fusion
3. **Body sensing, processing and modeling methods in metabolic health monitoring:** Multimodal physical activity recognition, energy efficient sensing and modeling
4. **Multimodal affective computing:** multimodal emotion recognition, computational paralinguistics
5. **Human behavior signal processing:** gathering, analyzing and modeling multimodal human behavior signals, both overtly and covertly expressed (speech/language/audio/visual/physiological signal analysis and understanding), for children autism
6. **Speech production system based biometrics:** articulation, morphology, inversion

Structure health monitoring:

1. Robust waveform number analysis
2. Data-driven matched field processing
3. Acoustic-Ultrasonic Localization for damages in plates and pipes

My previous research interests:

Audio watermarking: Robust frequency domain audio watermarking, content adaptive audio watermarking in wavelet domain

Computational acoustics scene analysis: co-channel speech separation

Teaching:

Teaching J1799d Design Implementation of Speech Recognition Systems (Fall14, Fall15, Spring15, Fall16)

Teaching J1782 Pattern Recognition (Spring17)

Funding:

- PI of NSFC project “Speaker recognition and language recognition using articulatory information in the speech production system and mid-level discriminative tokenization”, 2015.1~2017.12
- PI of NSFC-Guangdong project “Robust short duration text dependent speaker verification”, 2015.1~2017.12
- PI of the Fundamental Research Funds for the Central Universities project, “Speech information retrieval for targeting Chinese dialects and foreign languages”, 2015.7~2017.6
- PI of JRI project “Multimodal human behavior modeling through computational sensing and analysis for young children with autism spectrum disorders”, 2013.9~2016.9
- PI of project “text dependent speaker verification”, funded by BAIDU, 2015.6-2015.12.
- PI of project “speaker verification and language identification and voice conversion technology”, funded by Jinlin Tech, 2016.1-2016.8.
- PI of project “Multimodal human behavior signal analysis for autism kids”, funded by IBM, 2016.1-2016.12.
- co-PI of project, “Whole body PET/MRI imaging system”, funded by National Key Research and Development Program, 2016-2020.
- co-PI of JRI innovation team project “Structure health monitoring for plates and pipes using ultrasound techniques”, 2014,1~2016.12
- Co-PI of project “Algorithm and Hardware Co-Design for Ultra Low-Power Data Processing of Electrocardiogram (ECG) Biometrics” funded by CMU-SYSU Collaborative Innovation Research Center at Carnegie Mellon University, 2014-2015
- Co-PI of project “Affect Analysis in Human-Human Interactions” funded by CMU-SYSU Collaborative Innovation Research Center at Carnegie Mellon University, 2015-2016

PUBLICATIONS

Journal papers:

- 1) Wenbo Zhao, **Ming Li (*)**, Joel B. Harley, Yuanwei Jin, Jos e M.F. Moura, Jimmy Zhu, "Reconstruction of Lamb wave dispersion curves by sparse representation and continuity constraints", submitted to Journal of the Acoustical Society of America.
- 2) Wenbo Liu, **Ming Li (*)**, Li Yi, "Identifying Children with Autism Spectrum Disorder Based on Their Face Processing Abnormality: A Machine Learning framework", Autism research, 2016.
- 3) Yinliang Xu, Zaiyue Yang, Wei Gu, **Ming Li**, and Zicong Deng, "Robust Real-Time Distributed Optimal Control Based Energy Management in a Smart Grid", IEEE TRANSACTIONS ON SMART GRID, 2015.
- 4) Donna Spruijt-Metz, Cheng K.F. Wen, Gillian O'Reilly, **Ming Li**, Sangwon Lee, Adar Emken, Urbashi Mitra, Murali Annavaram, Gisele Ragusa, Shrikanth Narayanan. "Innovations in the Use of Interactive Technology to Support Weight Management", Current Obesity Reports. 2015
- 5) **Ming Li(*)**, Jangwon Kim, Adam Lammert, Prasanta Kumar Ghosh, Vikram Ramanarayanan, Shrikanth Narayanan, "Speaker verification based on the fusion of speech acoustics and inverted articulatory signals", Computer Speech & Language, 2015.
- 6) **Ming Li (*)**and Wenbo Liu, "Generalized I-vector Representation with Phonetic Tokenizations and Tandem Features for both Text Independent and Text Dependent Speaker Verification", Journal of Signal Processing Systems, 2015.
- 7) **Ming Li(*)**, and Shrikanth Narayanan. "Simplified Supervised I-vector Modeling and Sparse Representation with application to Robust Language Recognition", Computer Speech & Language, 2014.
- 8) **Ming Li(*)**, Kyu J. Hanb and Shrikanth Narayanan, "Automatic Speaker Age and Gender Recognition using acoustic and prosodic level information fusion", Computer speech and language, 2013, vol 27.
- 9) Daniel Bone, **Ming Li**, Matthew P. Black and Shrikanth S. Narayanan. "Intoxicated Speech Detection: A Fusion Framework with Speaker-Normalized Hierarchical Functionals and GMM Supervectors." Computer Speech & Language, 2013.
- 10) Jangwon Kim, Naveen Kumar, Andreas Tsiartas, **Ming Li**, Shrikanth S. Narayanan, "Automatic intelligibility classification of sentence-level pathological speech", to appear in Computer Speech & Language.
- 11) U. Mitra, A. Emken, S. Lee, **M. Li**, V. Rozgic, G. Thatte, H. Vathsangam, D. Zois, M. Annavaram, S. Narayanan, D. Spruijt-Metz, and G. Sukhatme, "KNOWME: a Case Study in Wireless Body Area Sensor Network Design", IEEE Communications Magazine 2012 50:5(116-125).
- 12) Gautam Thatte, **Ming Li**, Sangwon Lee, Adar Emken, Shri Narayanan, Urbashi Mitra, Donna Spruijt-Metz and Murali Annavaram, "KNOWME: An Energy-Efficient and Multimodal Body Area Sensing System for Physical Activity Monitoring," ACM Transactions in Embedded Computing Systems, 2012.
- 13) Adar Emken, **Ming Li**, Gautam Thatte, Sangwon Lee, Murali Annavaram, Urbashi Mitra, Shrikanth Narayanan, Donna Spruijt-Metz, "Recognition of Physical Activities in Overweight Hispanic Youth using KNOWME Networks", Journal of Physical Activity and Health, 9:3(432-441) 2012.
- 14) Gautam Thatte, **Ming Li**, Sangwon Lee, Adar Emken, Murali Annavaram, Shri Narayanan, Donna Spruijt-Metz, Urbashi Mitra, "Optimal Time-Resource Allocation for Energy-Efficient Physical Activity Detection", IEEE Transaction on Signal Processing, vol 59, issue 4, April, 2011.
- 15) **Ming Li(*)**, Viktor Rozgic, Gautam Thatte, Sangwon Lee, Adar Emken, Murali Annavaram, Urbashi Mitra, Donna Spruijt-Metz and Shrikanth Narayanan, "Multimodal Physical Activity Recognition by Fusing Temporal and Cepstral Information," IEEE Transactions on Neural Systems & Rehabilitation Engineering, vol 18, issue 4, August, 2010.
- 16) Hongbin Suo, **Ming Li**, Ping Lu, Yonghong Yan, "Automatic language identification with discriminative language characterization based on SVM", IEICE transaction on Information and Systems, 2008.
- 17) Hongbin Suo, **Ming Li**, Ping Lu, Yonghong Yan, "Using SVM as back-end classifier for language identification", EURASIP Journal on Audio, Speech, and Music Processing, 2008.
- 18) Jianping Zhang, **Ming Li**, Hongbin Suo, Lin Yang, Qiang Fu and Yonghong Yan, "Long span prosodic features for speaker recognition", ACTAACOUSTICA, 2010. (In chinese)

Conference papers:

- 1) Tianyan Zhou, Weicheng Cai, Huadi Zheng, Luting Wang, Xiaoyan Chen, Xiaobing Zou, Shilei Zhang, **Ming Li**, "A Pitch, Energy and Phoneme Duration Features based Speaker Diarization System for Autism Kids' Real-Life Audio

- Data", ISCSLP 2016.
- 2) Danwei Cai, Weicheng Cai, **Ming Li**, "Locality Sensitive Discriminant Analysis for Speaker Recognition", APSIPA ASC 2016.
 - 3) Gaoyuan He, Jinkun Chen, Xuebo Liu, and **Ming Li**, "The SYSU System for CCPR 2016 Multimodal Emotion Recognition Challenge", CCPR 2016.
 - 4) Huadi Zheng, Weicheng Cai, Tianyan Zhou, Shilei Zhang, **Ming Li**, "Text-Independent Voice Conversion Using Deep Neural Network Based Phonetic Level Features", ICPR 2016.
 - 5) Wei Fang, Jianwen Zhang, Dilin Wang, Zheng Chen, **Ming Li**, Entity Disambiguation by Knowledge and Text Jointly Embedding, CONLL 2016.
 - 6) Yandong Wen, Weiyang Liu, Meng Yang, **Ming Li**, "Efficient Misalignment-robust Face Recognition via Locality-constrained Representation", ICIP 2016.
 - 7) Zhun Chen, Wenbo Zhao, Yuanwei Jin, **Ming Li**, Jimmy Zhu, "A Fast Tracking Algorithm for Estimating Ultrasonic Signal Time of Flight in Drilled Shafts Using Active Shape Models", IEEE International Ultrasonics Symposium 2016.
 - 8) Zhiding Yu, Weiyang Liu, Wenbo Liu, Yingzhen Yang, **Ming Li** and Vijayakumar Bhagavatula, "On Order-Constrained Transitive Distance Clustering", AACL, 2016.
 - 9) Shushan Chen, Yiming Zhou and **Ming Li**, "Automatic English Pronunciation Evaluation System", APSIPA ASC 2015.
 - 10) Shitao Weng, Shushan Chen, Lei Yu, Xuewei Wu, Weicheng Cai, Zhi Liu, Yiming Zhou and **Ming Li**, "THE SYSU SYSTEM FOR THE INTERSPEECH 2015 AUTOMATIC SPEAKER VERIFICATION SPOOFING AND COUNTERMEASURES CHALLENGE", APSIPA ASC 2015.
 - 11) Wenbo Liu, Zhiding Yu, Li Yi, Bhiksha Raj, **Ming Li**, "Efficient Autism Spectrum Disorder Diagnosis with Eye Movement: A Machine Learning Framework", ACII 2015.
 - 12) Wenbo Liu, Zhiding Yu, Bhiksha Raj and **Ming Li**, "Locality Constrained Transitive Distance Clustering on Speech Data", INTERSPEECH 2015.
 - 13) Qingyang Hong, Lin Li, **Ming Li**, Ling Huang, Lihong Wan and Jun Zhang, "Modified-prior PLDA and Score Calibration for Duration Mismatch Compensation in Speaker Recognition System", INTERSPEECH 2015.
 - 14) Weicheng Cai, **Ming Li**, Lin Li and Qingyang Hong, "Duration Dependent Covariance Regularization in PLDA Modeling for Speaker Verification", INTERSPEECH 2015.
 - 15) Yingxue Wang, Shenghui Zhao, Wenbo liu, **Ming Li**, Jingming Kuang, "Speech bandwidth extension based on Deep Neural Networks", INTERSPEECH 2015.
 - 16) **Ming Li**, "speaker verification with the mixture of Gaussian factor analysis based representation", ICASSP 2015.
 - 17) Wenbo Liu, Zhiding Yu and **Ming Li**, "An Iterative Framework for Unsupervised Learning in the PLDA based Speaker Verification", ISCSLP 2014.
 - 18) **Ming Li**, Wenbo Liu, "Speaker verification and spoken language identification using a generalized i-vector framework with phonetic tokenization and tandem features", INTERSPEECH 2014.
 - 19) **Ming Li**, "Automatic recognition of speaker physical load using posterior probability based features from acoustic and phonetic tokens", INTERSPEECH 2014.
 - 20) Prashanth Gurunath Shivakumar, **Ming Li**, Vedant Dhandhanian and Shrikanth S. Narayanan, "SIMPLIFIED AND SUPERVISED I-VECTOR MODELING FOR SPEAKER AGE REGRESSION", ICASSP 2014.
 - 21) **Ming Li**, Xin Li, "verification based ECG biometrics with cardiac irregular conditions using heartbeat level and segment level information fusion", ICASSP 2014.
 - 22) **Ming Li**, Andreas Tsiartas, Maarten Van Segbroeck and Shrikanth S. Narayanan, "SPEAKER VERIFICATION USING SIMPLIFIED AND SUPERVISED I-VECTOR MODELING", ICASSP 2013.
 - 23) **Ming Li**, Adam Lammert, Jangwon Kim, Prasanta Ghosh and Shrikanth Narayanan, "Automatic Classification of Palatal and Pharyngeal Wall Morphology Patterns from Speech Acoustics and Inverted Articulatory Signals", Workshop on Speech Production in Automatic Speech Recognition, 2013.
 - 24) **Ming Li**, Jangwon Kim, Prasanta Kumar Ghosh, Vikram Ramanarayanan and Shrikanth Narayanan, "Speaker verification based on fusion of acoustic and articulatory information", Interspeech 2013.
 - 25) Andreas Tsiartas, Theodora Chaspari, Nossos Katsamanis, Prasanta Ghosh, **Ming Li**, Maarten Van Segbroeck,

- Alexandros Potamianos, Shrikanth Narayanan, "Multi-band long-term signal variability features for robust voice activity detection", Interspeech 2013.
- 26) Kyu Jeong Han, Sriram Ganapathy, **Ming Li**, Mohamed K. Omar and Shrikanth S. Narayanan, "TRAP Language Identification System for RATS Phase II Evaluation", Interspeech 2013.
 - 27) Daniel Bone, Theodora Chaspari, Kartik Audhkhasi, James Gibson, Andreas Tsiartas, Maarten Van Segbroeck, **Ming Li**, Sungbok Lee and Shrikanth S. Narayanan, "Classifying Language-Related Developmental Disorders from Speech Cues: the Promise and the Potential Confounds", Interspeech 2013.
 - 28) **Ming Li**, Charley Lu, Anne Wang, Shrikanth Narayanan, "Speaker Verification using Lasso based Sparse Total Variability Supervector and Probabilistic Linear Discriminant Analysis", presented at NIST Speaker Recognition Workshop, Atlanta, 2011. published in Proceedings of APSIPA Annual Summit and Conference, Hollywood, CA, 2012
 - 29) **Ming Li**, Angeliki Metallinou, Daniel Bone, Shrikanth Narayanan, "Speaker states recognition using latent factor analysis based Eigenchannel factor vector modeling", ICASSP 2012.
 - 30) Jangwon Kim, Naveen Kumar, Andreas Tsiartas, **Ming Li** and Shrikanth S. Narayanan, "Intelligibility classification of pathological speech using fusion of multiple high level descriptors", Interspeech 2012.
 - 31) Kartik Audhkhasi, Angeliki Metallinou, **Ming Li** and Shrikanth S. Narayanan, "Speaker Personality Classification Using Systems Based on Acoustic-Lexical Cues and an Optimal Tree-Structured Bayesian Network", Interspeech 2012.
 - 32) **Ming Li**, Xiang Zhang, Yonghong Yan and Shrikanth Narayanan, "Speaker Verification using Sparse Representations on Total Variability I-Vectors", Interspeech Florence, Italy, 2011.
 - 33) **Ming Li**, Shrikanth Narayanan, "Robust talking face video verification using joint factor analysis and sparse representation on GMM mean shifted supervectors", ICASSP, Prague, Czech Republic, 2011.
 - 34) Samuel Kim, **Ming Li**, Sangwon Lee, Urbashi Mitra, Adar Emken, Donna Spruijt-Metz, Murali Annavaram, Shrikanth Narayanan, "Modeling high-level descriptions of real-life physical activities using latent topic modeling of multimodal sensor signals", IEEE conference of engineering in medicine and biology society (EMBC), Boston, 2011.
 - 35) Daniel Bone, Matthew P. Black, **Ming Li**, Angeliki Metallinou, Sungbok Lee and Shrikanth Narayanan, "Intoxicated Speech Detection by Fusion of Speaker Normalized Hierarchical Features and GMM Supervectors", Interspeech, Florence, Italy, 2011.
 - 36) **Ming Li**, Shrikanth Narayanan, "ECG Biometrics by Fusing Temporal and Cepstral Information", 20th conference of the International Association for Pattern Recognition, ICPR 2010, Turkey.
 - 37) **Ming Li**, Chi-Sang Jung and Kyu Jeong Han, "Combining Five Acoustic Level methods for Automatic Speaker Age and Gender Recognition", Interspeech, 2010.
 - 38) Gautam Thatte, Viktor Rozgic, **Ming Li**, Sabyasachi Ghosh, Urbashi Mitra, Shri Narayanan, Murali Annavaram, Donna Spruijt-Metz, "Optimal Time-Resource Allocation for Activity-Detection via Multimodal Sensing," in Proceedings of the Fourth International Conference on Body Area Networks (BodyNets), Los Angeles, CA, 2009.
 - 39) Gautam Thatte, Viktor Rozgic, **Ming Li**, Sabyasachi Ghosh, Urbashi Mitra, Shri Narayanan, Murali Annavaram and Donna Spruijt-Metz, "Optimal Allocation of Time-Resources for Multihypothesis Activity-Level Detection," in Proceedings of the 5th IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS), Marina Del Rey, CA (June 2009). (Best paper award!)
 - 40) Sangwon Lee, Murali Annavaram, Gautam Thatte, Vikor Rozgic, **Ming Li**, Urbashi Mitra, Shri Narayanan and Donna Spruijt-Metz, "Sensing for Obesity: KNOWME Implementation and Lessons for an Architect," in Proceedings of the Workshop on Biomedicine in Computing: Systems, Architectures, and Circuits (BiC2009), Austin, TX (June 2009).
 - 41) Gautam Thatte, **Ming Li**, Adar Emken, Urbashi Mitra, Shri Narayanan, Murali Annavaram and Donna Spruijt-Metz, "Energy-Efficient Multihypothesis Activity-Detection for Health-Monitoring Applications," the 31st Annual International Conference of the IEE Engineering in Medicine and Biology Society (EMBC), Minneapolis, 2009.
 - 42) **Ming Li**, Chuan Cao, Di Wang, Ping Lu, Qiang Fu, and Yonghong Yan, "Cochannel speech separation using multi-pitch estimation and model based voiced sequential grouping", proceeding of International Conference on Spoken Language Processing, INTERSPEECH 2008.
 - 43) **Ming Li**, Hongbin Suo, Xiao Wu, Ping Lu, Yonghong Yan, "Spoken Language Identification Using Score Vector Modeling and Support Vector Machine", INTERSPEECH 2007.
 - 44) **Ming Li**, Yun Lei, Xiang Zhang, Jian Liu, Yonghong Yan, "authentication and quality monitoring based audio

watermark for analog AM shortwave broadcasting”, proceeding of IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIH-MSP 2007.

45) Hongbin Suo, **Ming Li**, Tantan Liu, Ping Lu, Yonghong Yan, “The Design of Backend Classifiers in PPRLM System for Language Identification”, proceeding of International Conference on Natural Computation, ICNC 2007.

46) Hongbin Suo, **Ming Li**, Ping Lu, Yonghong Yan, “Language identification based on parallel PRLM system”, proceeding of Chinese national conference on network security, 2007 (in Chinese).

47) **Ming Li**, Yun Lei, Jian Liu, Yonghong Yan, "A Novel Audio Watermarking in Wavelet Domain", proceeding of IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIH-MSP 2006.

Awards:

2016 IBM Faculty Award

Works co-authored with his colleagues have won awards at Body Computing Slam Contest 2009, IEEE DCOSS 2009, ISCSLP 2014 best paper, Interspeech2011-Speaker State Challenge and Interspeech2012-Speaker Trait Challenge. He was also featured in the Nov 2013 issue of the Science magazine.

Technique committee or Scientific Review Committee member

ICASSP2014, 2015, 2016, INTERSPEECH 2014, 2015, 2016, ISCSLP 2012, 2014, Odyssey 2014,2016

Area Chair

INTERNSPEECH 2016

Reviewer

IEEE transactions on audio, speech and language processing

IEEE transactions on information forensics and security

IEEE computational Intelligence Magazine

IEEE transactions on Emerging Topics in Computing

IEEE transaction on affective computing

IEEE Transactions on Pattern Analysis and Machine Intelligence

IEEE Journal of Biomedical and Health Informatics

IEEE signal processing letter

Speech communication (Elsevier)

EURASIP journal on audio speech and music processing

Journal of Systems and Software (Elsevier)

Health and Technology

Integrated Computer-Aided Engineering

International Journal of Nanomedicine