

# Jing Zheng Program Director

Speech Technology And Research Laboratory SRI International, Menlo Park, CA 94025 USA

Email: [zj@speech.sri.com](mailto:zj@speech.sri.com), Tel: (650) 859-6129 (o) (408) 636-3928 (m) Fax: (650) 859-5984

## Research Interests

- Automatic Speech Recognition
- Machine Translation
- Spoken Language Systems
- Character Recognition
- Machine Learning
- Statistical Modeling
- Natural Language Processing

## Education

Ph.D. E.E., Tsinghua University, Beijing, China, January 1999

Thesis: Spatial-Temporal Unified Model Based on Statistical Structure Description and its Application in On-line handwritten Chinese Character Recognition. (Tsinghua University Excellent Ph.D. Dissertation Award)

B. Engineering., E.E., Tsinghua University, Beijing, China, July 1994

## Professional History

- Program Director, SRI International, Menlo Park, CA (2010-present)
- Senior Research engineer, SRI International, Menlo Park, CA (2006-2010)
- Research Engineer, SRI International, Menlo Park, CA (2001-2006)
- International fellow, SRI International, Menlo Park, CA (1999–2001)
- Research assistant, Tsinghua University, Beijing, China (1994–1999)

## Experience

### SRI International

#### Speech Technology And Research Laboratory (1999-Present)

#### Large-Scale Machine Translation Systems of Text & Audio Documents (2005-2009)

*Funded by DARPA under GALE (Global Autonomous Language Exploitation)*

Led SRI's effort of building high-accuracy large-scale Arabic-to-English and Chinese-to-English machine translation (MT) systems; developed the state-of-the-art hierarchical phrase-based MT, syntax-augmented MT, and string-to-dependency MT.

Co-led SRI consortium's multi-site MT effort involving 7-8 international partners; coordinated the initial proposal development of the MT task.

Led SRI's consecutive participations in NIST open MT evaluations in 2006, 2008 and 2009; ranked the first in MT08 Chinese-to-English constrained training track jointly with MSR and NRC by a combined submission.

Developed an ASR discriminative training technique that combined fMPE, MPFE and MLP approaches in SRI's GALE evaluation system leading to improved accuracy.

### **Real-Time Two-Way Speech-to-Speech Portable Translation Systems (2005-present)**

*Funded by DARPA under TRANSTAC (Spoken Language Communication and Translation System for Tactical Use)*

Served as Co-PI and technical leader. Coordinated research and development efforts among SRI, RWTH Aachen (Germany), USC/ISI, Language Weaver, and University of Washington.

Created and led development of SRI's SRInterp™ real-time cross-platform translation engine, which was used in SRI's IraqComm™ Iraqi Arabic/English speech-to-speech translation system and its clone for several other language pairs, as well as SRI GALE MT evaluation systems.

Improved accuracy and speed performance of SRI's DynaSpeak® real-time continuous speech recognition engine. Augmented functionalities of Dynaspeak® engine to support dynamic vocabulary, fast language model rescoring, online language model adaptation, etc.

Leading effort porting IraqComm™ to Nexus One smart phone platform, by aggressively reducing model size and improving space efficiency of both ASR and MT engines (ongoing).

### **Large Vocabulary Meeting Recognition Systems (2005-2007)**

*Funded by DARPA under CALO (Cognitive Agent that Learn and Organize)*

Developed a new discriminative training technique, fMPE-MAP algorithm, which significantly improved SRI' meeting system's accuracy. Participated in and won NIST meeting evaluations in 2006 and 2007.

### **Large Vocabulary Broadcast News and Conversational Telephone Speech Recognition Systems (2002-2004)**

*Funded by DARPA under EARS (Effective Affordable Reusable Speech to Text)*

Extensively involved in research and development of new algorithms and evaluation systems: Investigated rate-dependent acoustic models; Implemented Heteroscedastic Linear Discriminant Analysis (HLDA) transform; Developed various discriminative Training techniques, including MMI, MPE, and a new algorithm MPFE (Minimum Phone Frame Error) with improved performance; Improved DECIPHER™ decoder and trainer implementation for better efficiency, and support advanced acoustic models.

### **Speaker Recognition System (2004)**

*Funded by NSF/DoD under KDD (Knowledge Discovery and Dissemination)*

Investigated the use of structured MAP adaptation and Heteroscedastic Linear Discriminant Analysis (HLDA) in speaker recognition.

### **Mobile Translation Systems on Handheld Devices (2001-2003)**

*Funded by DARPA under Babylon/CAST*

Implemented various core ASR algorithms for embedded applications, contributed to Phraselator® as a key developer

### **Speech Recognition in Noisy Environment (2001)**

Contributed to the development of SRI's 2001 SPINE evaluation system, which had the highest performance among all participants

### **Large Vocabulary Conversational Speech Recognition (1999-2000)**

Investigated in speaking-rate dependent acoustic modeling; participated in LVCSR2000 NIST evaluation, in which SRI's system ranked 2<sup>nd</sup>.

### **Tsinghua University, Beijing, China**

**Center for Intelligent Image and Document Information Processing, E.E. Dept (1994-1999)**

### **PhD. Candidate and Graduate Student Researcher (1994-1999)**

Single-handedly developed a high-performance online Chinese Character recognition system, which won 1995 National High-Tech R&D Program (863 Program) evaluation, and ranked second in 1998 evaluation; as a group member received the second grade award of 1999 National Scientific and Technical Progress Prize for project "THOCR-97" Synthetic Integrated Chinese Character Recognition System.

### **Professional Activities**

- Program committee, MT Summit XII 2009
- Program committee, ACL 2010, ACL2011
- NSF Panelist, Human Language Technology Area, 2008
- Reviewer, EMNLP 2008
- Reviewer, COLING 2008
- Reviewer, ICPR 2006
- Reviewer, ICASSP 2004, 2005, 2006, 2007, 2008, 2009,2010,2011
- Reviewer, Interspeech 2006, 2007, 2008, 2009,2010
- Reviewer, IEEE Transactions on Audio, Speech and Language Processing. 2002, 2003, 2004, 2006, 2007, 2009,2010
- Reviewer, Journal of Computer Speech and Language. 2005, 2008

- Reviewer, IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002
- Reviewer, Language Resources and Evaluation, 2009
- Member of IEEE and ISCA
- Reviewer, IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002
- Reviewer, Language Resources and Evaluation, 2009
- Member of IEEE and ISCA

## Publications

- J. Zheng, A. Mandal, X. Lei, M. Frandsen, N. F. Ayan, D. Vergyri, W. Wang, M. Akbacak, and K. Precoda. "Implementing SRI's Pashto Speech-to-Speech Translation System on Smartphone," in *Proc. 2010 IEEE Workshop on Spoken Language Technology (SLT'10)*, (Berkeley, CA), December 2010
- W. Wang, G. Tur, J. Zheng, N. F. Ayan, "Automatic Disfluency Removal for Improving Spoken Language Translation," in *Proc. ICASSP2010*, (Dallas, TX), March 2010
- X. Lei, A. Mandal, J. Zheng, "Fast Likelihood Computation Using Hierarchical Gaussian Shortlists", in *ICASSP2010*, (Dallas, TX), March 2010.
- J. Zheng, N. F. Ayan, W. Wang, and D. Burkett, "Using Syntax in Large-Scale Audio Document Translation," in *Proc. Interspeech'2009*, (Brighton, England), September 2009.
- J. Zheng, N. F. Ayan, W. Wang, and D. Burkett, "Using Syntax in Large-Scale Audio Document Translation," in *Proc. Interspeech'2009*, (Brighton, England), September 2009.
- W. Wang, A. Mandal, X. Lei, A. Stolcke, and J. Zheng, "Multifactor Adaption for Mandarin Broadcast News and Conversation Speech Recognition," in *Proc. Interspeech'2009*, (Brighton, England), September 2009.
- M. Yang and J. Zheng, "Toward Smaller, Faster, and Better Hierarchical Phrase-Based SMT," to appear in *Proc. ACL-IJCNLP'2009*, (Singapore), August 2009.
- M. Akbacak, H. Franco, M. Frandsen, S. Hasan, H. Jameel, A. Kathol, S. Khadivi, X. Lei, A. Mandal, S. Mansour, K. Precoda, C. Richey, D. Vergyri, W. Wang, M. , Yang, and J. Zheng, "Recent Advances in SRI's IraqComm™ Iraqi Arabic-English Speech-to-Speech Translation System," in *Proc. ICASSP'2009*, (Taipei), April 2009.
- N. F. Ayan, J. Zheng, and W. Wang, "Improving alignments for better confusion networks for combining machine translation systems," in *Proc. of the 22nd International Conference on Computational Linguistics (COLING'08)*, (Manchester, UK), August 2008
- J. Zheng, W. Wang, and N. F. Ayan, "Development of SRI's translation systems for broadcast news and broadcast conversations," in *Proc. of Interspeech'2008*, (Brisbane, Australia), September 2008.
- D. Vergyri, A. Mandal, W. Wang, A. Stolcke, J. Zheng, M. Graciarena, D. Rybach, C. Gollan, R. Schlüter, K. Kirchhoff, A. Faria, and N. Morgan, "Development of the SRI/Nightingale Arabic ASR system," in *Proc. Interspeech'2008*, (Brisbane, Australia), September 2008.
- Kathol and J. Zheng, "Strategies for building a farsi-english SMT system from

limited resources," in *Proc. Interspeech'2008*, (Brisbane, Australia), September 2008.

- Mandal, D. Vergyri, W. Wang, J. Zheng, A. Stolcke, G. Tur, D. Hakkani-Tur, and N. F. Ayan, "Efficient data selection for machine translation," in *2008 IEEE Workshop on Spoken Language Technology*, (Goa, India), December 2008.
- J. Zheng, O. Cetin, M.-Y. Hwang, X. Lei, A. Stolcke, and N. Morgan, "Combining discriminative feature, transform, and model training for large vocabulary speech recognition," in *Proc. IEEE ICASSP'2007*, (Honolulu, Hawaii), April 2007.
- J. Zheng and A. Stolcke, "fMPE-MAP: Improved discriminative adaptation for modeling new domains," in *Proc. Interspeech'2007*, (Antwerp), August 2007.
- A. Stolcke, K. Boakye, O. Cetin, A. Janin, M. Magimai-Doss, C. Wooters, and J. Zheng, "The SRI-ICSI Spring 2007 meeting and lecture recognition system," in *Proc. NIST 2007 Rich Transcription Workshop*, 2007.
- W. Wang, A. Stolcke, and J. Zheng, "Reranking machine translation hypotheses with structured and web-based language models," in *Proc. IEEE ASRU Workshop*, (Kyoto, Japan), December 2007.
- M. Yang, J. Zheng, and A. Kathol, "A semi-supervised learning approach for morpheme segmentation for an arabic dialect," in *Interspeech'2007*, (Antwerp, Belgium), August 2007
- M.-Y. Hwang, W. Wang, X. Lei, J. Zheng, O. Cetin, and G. Peng, "Advances In Mandarin Broadcast Speech Recognition," in *Interspeech'2007*, (Antwerp, Belgium), August 2007.
- K. Precoda, J. Zheng, D. Vergyri, H. Franco, C. Richey, A. Kathol, and S. Kajarekar, "IraqComm: A next generation translation system," in *Interspeech'2007*, (Antwerp, Belgium), August 2007.
- A. Janin, A. Stolcke, X. Anguera, K. Boakye, Ö. Cetin, J. Frankel, and J. Zheng, "The ICSI-SRI Spring 2006 meeting recognition system," in *Machine Learning for Multimodal Interaction: Third International Workshop, MLMI'2006* (S. Renals, S. Bengio, and J. Fiscus, eds.), Lecture Notes in Computer Science, (Washington, D.C.), Springer, 2007
- A. Stolcke, B. Chen, H. Franco, V. R. R. Gadde, M. Graciarena, M.-Y. Hwang, K. Kirchhoff, N. Morgan, X. Lin, T. Ng, M. Ostendorf, K. Sönmez, A. Venkataraman, D. Vergyri, W. Wang, J. Zheng, and Q. Zhu, "Recent innovations in speech-to-text transcription at SRI-ICSI-UW," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, pp. 1729--1744, September 2006.
- S. Narayanan, P. G. Georgiou, A. Sethy, D. Wang, M. Bulut, S. Sundaram, E. Ettelaie, S. Ananthakrishnan, H. Franco, K. Precoda, D. Vergyri, J. Zheng, W. Wang, R. R. Gadde, V. Abrash, M. Frandsen, and C. Richey, "Speech recognition engineering issues in speech to speech translation system design for low resource languages and domains," in *Proc. IEEE ICASSP*, vol. 5, (Toulouse), May 2006
- A. Stolcke, X. Anguera, K. Boakye, Ö. Cetin, F. Grézl, A. Janin, A. Mandal, B. Peskin, C. Wooters, and J. Zheng, "Further progress in meeting recognition: The ICSI-SRI Spring 2005 speech-to-text evaluation system," in *Machine Learning for Multimodal Interaction Second International Workshop, MLMI 2005* (S. Renals and S. Bengio, eds.), vol. 3869 of Lecture Notes in Computer Science, Springer, 2006.
- J. Zheng and A. Stolcke, "Improved discriminative training using phone lattices," in

*Proc. Eurospeech*, (Lisbon), September 2005.

- S. S. Kajarekar, L. Ferrer, E. Shriberg, K. Sonmez, A. Stolcke, A. Venkataraman, and J. Zheng, "SRI's 2004 NIST speaker recognition evaluation system," in *Proc. IEEE ICASSP*, (Philadelphia), March 2005.
- J. Zheng, H. Franco, and A. Stolcke, "Effective acoustic modeling for rate-of-speech variation in large vocabulary conversational speech recognition," in *Proc. Int'l. Conf. Spoken Language Processing*, (Jeju, Korea), October 2004.
- Venkataraman, A. Stolcke, W. Wang, D. Vergyri, V. R. R. Gadde, and J. Zheng, "An efficient repair procedure for quick transcriptions," in *Proc. Int'l. Conf. Spoken Language Processing*, (Jeju, Korea), October 2004.
- S. Kajarekar, L. Ferrer, K. Sonmez, J. Zheng, E. Shriberg, and A. Stolcke, "Modeling NERFs for speaker recognition," in *Proc. Odyssey 04 Speaker and Language Recognition Workshop*, (Toledo, Spain), June 2004.
- M. Graciarena, H. Franco, J. Zheng, D. Vergyri, and A. Stolcke, "Voicing feature integration in SRI's Decipher LVCSR system," in *Proc. IEEE ICASSP*, (Montreal), May 2004.
- K. Precoda, H. Franco, A. Dost, M. Frandsen, J. Fry, A. Kathol, C. Richey, S. Riehemann, D. Vergyri, J. Zheng, and C. Culy, "Limited-domain speech-to-speech translation between English and Pashto," in *Proc. HLT/NAACL 2004 Demonstrations*, (Boston, MA), January 2004.
- J. Zheng, H. Franco, and A. Stolcke, "Modeling word-level rate-of-speech variation in large vocabulary conversational speech recognition," *Speech Communication*, vol. 41, pp. 273--285, 2003.
- H. Franco, J. Zheng, K. Precoda, F. Cesari, V. Abrash, D. Vergyri, A. Venkataraman, H. Bratt, C. Richey, and A. Sarich, "Development of phrase translation systems for handheld computers: from concept to field," in *Proc. Eurospeech*, (Geneva, Switzerland), 2003.
- Hodjat, H. Franco, H. Bratt, K. Precoda, A. Stolcke, A. Venkataraman, D. Vergyri, and J. Zheng, "Iterative statistical language model generation for use with an agent-oriented natural language interface," in *10th International Conference on Human-Computer Interaction*, (Crete), June 2003.
- V. R. R. Gadde, A. Stolcke, D. Vergyri, J. Zheng, K. Sonmez, and A. Venkataraman, "Building an ASR system for noisy environments: SRI's 2001 SPINE evaluation system," in *Proc. International Conference on Spoken Language Processing*, (Denver, CO), September 2002.
- H. Franco, J. Zheng, J. Butzberger, F. Cesari, M. Frandsen, J. Arnold, V. R. R. Gadde, A. Stolcke, and V. Abrash, "DynaSpeak: SRI's scalable speech recognizer for embedded and mobile systems," in *Proc. Human Language Technology Conference*, (San Diego, CA), March 2002.
- J. Zheng, J. Butzberger, H. Franco, and A. Stolcke, "Improved maximum mutual information estimation training of continuous density HMMs," in *Proc. Eurospeech*, (Aalborg, Denmark), September 2001.
- J. Zheng, H. Franco, F. Weng, A. Sankar, and H. Bratt, "Word-level rate of speech modeling using rate-specific phones and pronunciations," in *Proc. ICASSP*, (Istanbul), June 2000.
- J. Zheng, H. Franco, and A. Stolcke, "Rate-dependent acoustic modeling for large

vocabulary conversational speech recognition," in *Proc. ISCA Tutorial and Research Workshop on Automatic Speech Recognition: Challenges for the new Millennium*, (Paris), September 2000.

- Stolcke, H. Bratt, J. Butzberger, H. Franco, V. R. Rao Gadde, M. Plauch, C. Richey, E. Shriberg, K. Somez, F. Weng, and J. Zheng, "The SRI March 2000 Hub-5 conversational speech transcription system," in *Proc. NIST Speech Transcription Workshop*, (College Park, MD), May 2000.
- J. Gao, X.Q. Ding, & J. Zheng, Image Pattern Recognition Based on Examples -- A Combined Statistical and Structural-Syntactic approach, Proceedings of joint IAPR international workshops on syntactical and structural pattern recognition and statistical pattern recognition, Alicante, Spain, September, 2000
- J. Zheng, X.Q. Ding & Y.S. Wu, Spatio-Temporal Unified Model for On-line Handwritten Chinese Character Recognition, Proceedings of International Conference on Document Analysis and Recognition, Bangalore, India, September
- J. Zheng, X.Q. Ding, & Y.S. Wu, Dynamic classifier combination method based on minimum cost criterion, *Tsinghua Science and Technology*, pp. 1615~1619, Vol. 4, No. 4, 1999
- J. Zheng, X.Q. Ding, & Y.S. Wu, Dynamic Combination of Multi Classifiers Based on Minimum Cost Criterion, *The Chinese Journal of Computers*, Vol. 22, No 2, 1999 (in Chinese)
- J. Zheng, X.Q. Ding, & Y.S. Wu, Dynamic Classifier Combination Method Based on Minimum Cost Criterion and its Application, Proceedings of 2nd International Conference Multimodal Interface, Hong Kong, January, 1999
- J. Zheng, X.Q. Ding, & Y.S. Wu, Combination of Multi Algorithms Solving Coarse Classification Problem for OLCCR, *Pattern Recognition and Artificial Intelligence*. pp. 239~244, Vol. 11, No. 2, 1998 ( In Chinese)
- J. Zheng, X.Q. Ding, & Y.S. Wu, Combinatorial Coarse Classification Method for OLCCR, Proceedings of Asian Conference on Computer Vision, Hong Kong, V2, pp. 145~152, January.1998
- J. Zheng, X.Q. Ding, Y.S. Wu, F.X. Guo & M. Yin, On-line Handwritten Chinese Character Recognition Method Handling both Stroke Concatenation and Stroke Order Variation," *Journal of Tsinghua University (Science and technology edition)*, Vol. 37, No.9, September 1997. (In Chinese)
- J. Zheng, X.Q. Ding, Y.S. Wu, & M. Yin, Study on Combinatorial Pre-classification Approaches for On-line Handwritten Chinese Character Recognition, Proceedings of the 3rd National Academic Conference on Computer Intelligent Interface and Application, pp. 53~58, Zhang Jia Jie, China, August, 1997 (In Chinese)
- J. Zheng, X.Q. Ding, & Y.S. Wu, Recognizing On-line Handwritten Chinese Character via FARG Matching, Proceedings of International Conference on Document Analysis and Recognition, pp. 621~624, Ulm, Germany, August 1997
- J. Zheng, X.Q. Ding, & Y.S. Wu, An On-line Handwritten Chinese Character Recognition Method Based on Fuzzy Attributed Relational Graph Matching," Proceedings of International. Conference. on Computer Processing of Oriental Languages, Hong Kong, pp. 234~237, Hong Kong, April 1997

## Patents

- H. Franco, G. Myer, J. Zheng, F., Cesari, and C. Cowan, Method and apparatus for error correction in speech recognition applications.(US 7756710)
- J. Butzberger, H. Franco, L. Neumeyer, J. Zheng, Method of dynamically altering grammars in a memory efficient speech, US Patent 7324945
- V. Abrash, F. Cesari, H. Franco, C. George, J. Zheng, Method and apparatus for obtaining complete speech signals for speech recognition applications, US Patent 7610199

## Honors

- 2008 SRI Spot Award for outstanding contributions to GALE and TRANSTAC teams.
- 1999 Excellent Doctoral Dissertation, Tsinghua of University
- Second grade award of 1999, National Scientific and Technical Progress Prize, P. R. China, for project “THOCR-97 Synthetic Integrated Chinese Character Recognition System,”
- First grade award of 1998 Scientific and Technical Progress Prize by Ministry of Education, P. R. China, for project “THOCR-97 Synthetic Integrated Chinese Character Recognition System”.

## Personal Information

US citizen.

Languages: English (fluent), Chinese (native)