

Representative Publications

A. Recent Work

1. “Spoken Content Retrieval–Beyond Cascading Speech Recognition with Text Retrieval”, to appear on IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol. 23, No. 9, Sept. 2015, pp. 1389-1420.
(This an overview paper in the journal. The overview papers were started in 2010 in this journal. It has to be on an important problem in the area, having co-authors from different groups to have a balanced view, can be much longer than a regular paper and at most 4 such papers each year. This paper gives an overview of the new concepts and directions of speech information retrieval, co-authored by a scientist in MIT. This paper is not yet published but will appear recently.)
2. “Spoken Document Understanding and Organization”, IEEE Signal Processing Magazine, Special Issue on Speech Technology in Human-machine Communication, Vol. 22, No.5, Sept. 2005, pp.42-60.
(This is one of 9 papers selected by the special issue on speech technology. Usually there is only one such special issue in a few years. This is a review paper presenting the new framework for easier browsing the retrieved speech information, including a prototype system developed at Taiwan University with functionalities first seen in the world at the time of publication.)
3. “Voice-based Information Retrieval - how far are we from the text-based information retrieval?” (invited paper), IEEE Automatic Speech Recognition and Understanding Workshop, Merano, Italy, Dec 2009.
(invited paper for the invited speech in the workshop, overviewing the technology of voice-based information retrieval, comparing it with the currently very successful text-based counterpart, analyzing the differences and presenting possible future directions)
4. “Multi-layered Summarization of Spoken Document Archives by Information Extraction and Semantic Structuring”, Interspeech Conference, Sept. 2006, Pittsburgh, USA, International Speech Communication Association (ISCA).
(full paper presenting the technologies and performance evaluation of “spoken document understanding and organization”; presented in the Special Session of Speech Summarization in the Interspeech Conference as 1 out of 6 papers accepted globally; the presentation including a demonstration system with functionalities not seen elsewhere before)
5. “Voice Access of Global Information for Broadband Wireless: Technologies of Today and Challenges of Tomorrow” (invited paper), Proceedings of the IEEE, Jan. 2001, pp.41-57.
(invited paper in a Special Issue on Broadband Wireless Communications)
6. “Spoken Knowledge Organization by Semantic Structuring and a Prototype Course Lecture System for Personalized Learning”, IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 22, No. 5, May 2014, pp. 883-898.
(proposing new technologies for semantic structuring and organizing the recorded courses using speech technologies for convenient learning selected knowledge by learners, with a courses lecture system developed at National Taiwan University as an example)
7. “Improved Semantic Retrieval of Spoken Content by Document/Query Expansion with Random Walk over Acoustic Similarity Graphs”, IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 22, No. 1, Jan 2014, pp. 80-94.

- (proposing new approaches for semantic retrieval of spoken content by random walk)*
8. “Semantic Analysis and Organization of Spoken Documents Based on Parameters Derived from Latent Topics”, IEEE Transactions on Audio, Speech and Language Processing, Vol. 19, No. 7, Sept 2011, pp. 1875-1889.
(proposing two parameters useful in different tasks of semantic analysis of speech, including key term extraction, title and summary generation and topic classification; results not limited to Chinese although tested with data in Chinese)
 9. “Integrating Recognition and Retrieval with Relevance Feedback for Spoken Term Detection”, IEEE Transactions on Audio, Speech and Language Processing, Vol. 20, No. 7, Sep 2012, pp. 2095-2110.
(proposing a completely new concept for speech information retrieval: integrating recognition and retrieval and optimize them together as a whole rather than directly cascading them; not limited to Chinese although tested with data in Chinese)
 10. “Model-based Unsupervised Spoken Term Detection with Spoken Queries”, IEEE Transactions on Audio, Speech, and Language Processing, Vol. 21, No. 7, Jul 2013, pp. 1330-1342.
(proposing a completely new concept for unsupervised speech information retrieval without recognition: matching based on models rather than on signals; not limited to Chinese although tested with data in Chinese)
 11. “Enhanced Spoken Term Detection Using Support Vector Machines and Weighted Pseudo Examples”, IEEE Transactions on Audio, Speech and Language Processing, Vol. 21, No. 6, Jun 2013, pp. 1272-1284.
(proposing new methods for spoken term detection using support vector machine)
 12. “Interactive Spoken Document Retrieval with Suggested Key Terms Ranked by a Markov Decision Process”, IEEE Transactions on Audio, Speech and Language Processing, Vol. 20, No. 2, Feb 2012, pp. 632-645.
(proposing new approaches for interactive retrieval of spoken documents)
 13. “Performance Analysis for Lattice-Based Speech Indexing Approaches Using Words and Subword Units”, IEEE Transactions on Audio, Speech and Language Processing, Vol. 18, No. 6, Aug 2010, pp. 1562-1574.
(the first complete analysis on a whole set of speech indexing approaches considering retrieval accuracy and computation requirements; results not limited to Chinese although tested with data in Chinese)
 14. “Discriminating Capabilities of Syllable-based Features and Approaches of Utilizing Them for Voice Retrieval of Speech Information in Mandarin Chinese”, IEEE Transactions on Speech and Audio Processing, Vol.10, No.5, July 2002, pp. 303-314.
(first journal paper proposing most discriminative indexing features for Chinese spoken document retrieval)
 15. “A Recursive Dialogue Game for Personalized Computer-Aided Pronunciation Training”, IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol. 23, No. 1, Jan 2015, pp. 127-141.
(proposing new approaches for computer assisted language learning using dialogue games)
 16. “Supervised Detection and Unsupervised Discovery of Pronunciation Error Patterns for Computer-Assisted Language Learning”, IEEE/ACM Transactions on Audio, Speech and

Language Processing, Vol. 23, No. 3, Mar 2015, pp. 564-579.

(proposing new approaches to detect and discover pronunciation error patterns for computer-assisted language learning)

17. “An Improved Framework for Recognizing Highly Imbalanced Bilingual Code-Switched Lectures with Cross-Language Acoustic Modeling and Frame-Level Language Identification”, IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol. 23, No. 7, Jul 2015, pp. 1144-1159.
(first proposing complete framework for recognizing Mandarin-English bilingual code-switched speech)
18. “Improved Features and Models for Detecting Edit Disfluencies in Transcribing Spontaneous Mandarin Speech”, IEEE Transactions on Audio, Speech and Language Processing, Vol. 17, No. 7, Sept 2009, pp. 1263-1278.
(proposing a whole set of features and models for processing the very difficult spontaneous speech in Chinese)
19. “Higher Order Cepstral Moment Normalization for Improved Robust Speech Recognition”, IEEE Transactions on Audio, Speech and Language Processing, Vol. 17, No. 2, Feb 2009, pp.205-220.
(proposing the new family of higher order cepstrum moment normalization techniques for robust speech recognition under noisy environment)
20. “Pronunciation Modeling with Reduced Confusion for Mandarin Chinese Using A Three-stage Framework”, IEEE Transactions on Audio, Speech and Language Processing, Vol.15, No.2, Feb 2007, pp.661-675.
(proposing a new framework to model pronunciation variation in spontaneous Mandarin speech)

B. Pioneering Contributions on Chinese Spoken Language Processing

21. “Voice Dictation of Mandarin Chinese”, Special Section on Signal Processing in Asia, IEEE Signal Processing Magazine, Vol. 14, Nov 4, July 1997, pp. 63-101.
(This is a review paper presenting the problems and approaches in speech recognition for Mandarin Chinese. This is one of the two feature articles of the issue, showing the pioneering leadership of the nominee in the area. The national flags of Taiwan (ROC) and China (PRC) were both shown on the cover of the issue sent to readers and libraries worldwide, with that of Taiwan (ROC) slightly larger, obviously because people in both Taiwan and China speak Chinese, but the author of the paper is in Taiwan.)
22. “Structural Features of Chinese Language: Why Chinese Spoken Language Processing is Special and Where We Are”, Keynote Speech, International Symposium on Chinese Spoken Language Processing, Singapore, Dec 1998, pp. 1-15.
(opening keynote speech in the first International Symposium on Chinese Spoken Language Processing, reviewing the past achievements in different areas and looking forward to the future, primarily focusing on the approaches to handling structural features of Chinese language and considering the network environment.)
23. “Golden Mandarin (I) — A Real-time Mandarin Speech Dictation Machine for Chinese Language with Very Large Vocabulary”, IEEE Transactions on Speech and Audio Processing, Vol. 1, No.2, Apr 1993, pp. 158-179.

- (first journal paper on Mandarin dictation system in the world)*
24. “Complete Recognition of Continuous Mandarin Speech for Chinese Language with Very Large Vocabulary Using Very Limited Training Data”, IEEE Transactions on Speech and Audio Processing, Vol.5, No.2, March 1997, pp. 195-200.
(first journal paper on continuous Mandarin Speech dictation system in the world)
 25. “An Efficient Natural Language Processing System Specially Designed for the Chinese Language”, Computational Linguistics, Vol. 17, No. 4, Dec 1991, pp. 347-374.
(The first journal paper in the world presenting computer analysis of Chinese natural language sentence structures)
 26. “Special Speech Recognition Approaches for the Highly Confusing Mandarin Syllables Based on Hidden Markov Models”, Computer Speech and Language, Vol.5, No.2, Apr 1991, pp.181-201.
(first journal paper on Mandarin syllable recognition using Hidden Markov Models in the world)
 27. “Markov Modeling of Mandarin Chinese for Decoding the Phonetic Sequence into Chinese Characters”, Computer Speech and Language, Vol.5, No.4, Oct 1991, pp.363-377.
(first journal paper on Chinese language modeling in the world)
 28. “The Synthesis Rules in A Chinese Text-to-Speech System”, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-37, No. 9, Sept 1989, pp. 1309-1320.
(The first journal paper in the world presenting Chinese text-to-speech synthesis)

C. Typical Contributions in Communications

29. “An Exact Performance Analysis of the Clipped Diversity Combining Receiver for FH/MFSK Systems Against A Band Multitone Jammer”, IEEE Transactions on Communications, Vol.42, No.2/3/4, Feb/ March/ Apr 1994, pp.700-710.
(typical example for work in spread spectrum)
30. “Multi-h Phase Coded Modulations with Asymmetric Modulation Indices”, IEEE Journal on Selected Areas in Communications, Vol.SAC-7, No.9, Dec 1989, Special Issue on Bandwidth and Power Efficient Coded Modulations, pp.1450-1461.
(typical example for work on modulation techniques)
31. “Minimum Likelihood – A New Concept for Symbol Synchronization”, IEEE Transactions on Communications, Vol. COM-35, No.5, May 1987, pp. 545-549.
(typical example for work on synchronization)
32. “A General Theory for Asynchronous Speech Encryption Techniques”, IEEE Journal on Selected Areas in Communications, Vol. SAC-4, No.2, Mar. 1986, Special Issue on Military Communications, pp.280-287.
(typical example for work on communication security)

D. Contribution in Engineering Education

33. “Taiwan: Meeting the New Challenges” in “Special Issue on Engineering Education - A Global View”, IEEE Communications Magazine, Vol. 30, No. 11, Nov. 1992. pp. 18-26.
(on engineering education, one out of the nine papers invited globally in the Special Issue)