Hua Yu

Language Technology Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

(412) 268-5479 (o)
(412) 422-2762 (h)
hyu@cs.cmu.edu
http://www.cs.cmu.edu/∼hyu

Objective   A challenging position in speech recognition, statistical modeling/learning, natural language processing and related fields.

Research Interests

- Large vocabulary conversational speech recognition, handwriting recognition, or sequential modeling in general

- Statistical modeling, machine learning, pattern recognition

Education

Doctor of Philosophy in Language Technologies   August 2004
Thesis: Recognizing Sloppy Speech
Carnegie Mellon University, Pittsburgh, PA
Recipient of Graduate Research Fellowship, Language Technologies Institute, 1996-2004

Master of Science in Language Technologies   May 1998
Carnegie Mellon University, Pittsburgh, PA

Master of Science and Engineering in Computer Science   June 1996
Tsinghua University, Beijing, China
Recipient of Motorola Scholarship, 1994

Bachelor of Science and Engineering in Computer Science   June 1994
Tsinghua University, Beijing, China
Recipient of first class prize for excellent students, 1989-1994

Research Experiences

- **Research Assistant**   Sept. 1996 – present
  School of Computer Science, Carnegie Mellon University   Pittsburgh, PA

  I have extensive experiences in building LVCSR systems, which involves developing and maintaining the JANUS speech recognition system of around 200K lines of C code. I have led the efforts in developing:

  - the ISL Switchboard system, which achieves state-of-the-art performance (23.5% word error rate) on the DARPA RT-03 evaluation;

  - the Broadcast News transcription system, which is also used for a live dictation demo;

  - automatic meeting transcription system.

  My thesis topic is to improve the recognition of sloppy speech. To this end, I have explored several novel approaches, including single tree clustering, Gaussian transition modeling and thumbnail features.

  I have also been involved in the following research projects:

  - face recognition: We developed a new, direct LDA algorithm for classification of high dimensional data, such as face images.

  - automatic segmentation and clustering of broadcast news/meeting data.

- grapheme-to-phoneme mapping: We developed a program that consults DECTalk, MITalk, as well as static dictionaries to answer pronunciation queries from the network.
  - voice-driven web browser: Sphinx-II is used to recognize hyperlinks as well as a number of navigation commands.
  - automatic clustering of text documents.

- **Research Assistant** 1994-1996
  Speech Lab, Tsinghua University Beijing, China

  I worked on automatic language identification and speaker-independent Chinese syllable/phrase recognition. I also volunteered as a system/network administrator in the lab.

- **Various Side Projects**

  - Designing an intelligent controller for brushless DC motor with a single-chip controller;
  - Tracking down a new virus and developing an anti-virus program;
  - Developing a postal service window system;
  - Many others.

Other Experiences

- Teaching two lectures on HMMs for CMU CS11-751: Speech Recognition and Understanding

- Reviewer for Pattern Recognition and ICMI'2003

- Consultant for the Spoken Language Technology group, Sony Inc. May 2000. My job is to assist them in developing an LVCSR system.

- Teaching Assistant for CMU CS15-229: Multimedia Signal Processing, by Prof. R. Thibadeau, Prof. R. Dannenberg and Prof. R. Reddy, 1999

- LTI admission committee member, 1998

Programming Skills

Proficient in C, Perl, TclTk, Linux/Unix System Administration, LaTeX, TCP/IP
Experienced in C++, Matlab, VisualBasic, 80x86 Assembly, PASCAL, Lisp, etc.

Publications (Speech Related)

1. H. Yu. **Phase Space Representation of Speech — Revisiting the delta and double-delta features**. ICSLP, Jeju Island, Korea, 2004

2. H. Soltau, H. Yu, F. Metze, C. Fuegen, Q. Jin and S. Jou. **The ISL Transcription System for Conversational Telephony Speech**. ICASSP, Montreal, 2004

3. H. Yu and A. Waibel. **Integrating Thumbnail Features for Speech Recognition Using Conditional Exponential Models**. ICASSP, Montreal, 2004

4. H. Yu and T. Schultz. **Enhanced Tree Clustering with Single Pronunciation Dictionary for Conversational Speech Recognition**. Eurospeech, Geneva, 2003

5. H. Yu and T. Schultz. **Implicit Trajectory Modeling through Gaussian Transition Models for Speech Recognition**. HLT-NAACL, Edmonton, 2003

6. H. Yu and A. Waibel. **Flexible Parameter Tying for Conversational Speech Recognition**. ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition, Tokyo, 2003

7. H. Soltau, H. Yu, F. Metze, C. Fuegen, Q. Jin and S. Jou. **The ISL RT-03 Conversational Telephone Speech Recognition System**. Rich Transcription Workshop, Boston, MA, 2003

8. S. Burger, V. MacLaren and H. Yu. **The ISL Meeting Corpus: The Impact of Meeting Type on Speech Style**. ICSLP, Denver, 2002

9. H. Soltau, H. Yu, F. Metze, C. Fuegen, Y. Pan and S. Jou. **ISL Meeting Recognition**. Rich Transcription Workshop, Vienna, VA, 2002

10. C. Hori, S. Furui, R. Malkin, H. Yu and A. Waibel. **Automatic Speech Summarization Applied to English Broadcast News Speech**. ICASSP, Orlando, 2002

11. A. Waibel, M. Bett, F. Metze, K. Ries, T. Schaaf, T. Schultz, H. Soltau, H. Yu and K. Zechner. **Advances in Automatic Meeting Record Creation and Access**. ICASSP, Salt Lake City, 2001

12. A. Waibel, H. Yu, M. Westphal, H. Soltau, T. Schultz, T. Schaaf, Y. Pan, F. Metze and M. Bett. **Advances in Meeting Recognition**. HLT, San Diego, 2001

13. H. Yu and A. Waibel. **Streamlining the Front-End of a Speech Recognizer**. ICSLP, Beijing, 2000

14. H. Yu, T. Tomokiyo, Z. Wang and A. Waibel. **New Developments in Automatic Meeting Transcription**. ICSLP, Beijing, 2000

15. R. Gross, M. Bett, H. Yu, X. Zhu, Y. Pan, J. Yang and A. Waibel. **Towards a Multimodal Meeting Record**. ICME, New York, 2000

16. H. Yu, M. Finke and A. Waibel. **Progress in Automatic Meeting Transcription**. Eurospeech, 1999

17. H. Yu, C. Clark, R. Malkin and A. Waibel. **Experiments in Automatic Meeting Transcription using JRTk**. ICASSP, Seattle, USA, May 1998

18. D. Fang, H. Yu and S. Li. **Speech Recognition based on Normal Distribution Hypothesis**. Intl. Conf. on Chinese Computing, Singapore, 1994

19. H. Yu, S. Li, S. Qing and D. Fang. **Speaker-independent Isolated Word/Phrase Recognition — a Statistical Approach**. National Conf. on Human-Machine Communication, Chongqing, Oct. 1994

Publications (Non-Speech Areas)

1. H. Yu and J. Yang. **A Direct LDA Algorithm for High-Dimensional Data — with Application to Face Recognition**. Pattern Recognition 34(10), 2001, pp. 2067-2070

2. J. Yang, H. Yu and W. Kunz. **An Efficient LDA Algorithm for Face Recognition**. ICARCV, Singapore, 2000

3. H. Yu. **Automatically Determining Number of Clusters**. Information Retrieval course (CMU CS11-741) final report, 1998

4. H. Yu and Z. Wang, **A Survey on Anonymous Digital Cash Systems**. Security and Cryptography course(CMU CS15-827) final report, 1997

Selected Presentations

- **Phase Space Representation of Speech — Revisiting the delta and double-delta features** Sphinx Speech Group, Carnegie Mellon University, Pittsburgh, PA, July, 2003

- **Implicit Pronunciation Modeling for Conversational Speech Recognition**. Joint Speech Seminar, Carnegie Mellon University, Pittsburgh, PA, May 2, 2003

- **Development of the Broadcast News System**. Interactive Systems Labs, Carnegie Mellon University, Pittsburgh, PA, May, 1999

- **Experiments in Automatic Meeting Transcription using JRTk**. ICASSP, Seattle, USA, May, 1998

Personal Native speaker of Chinese, fluent in English. Citizenship: China. US permanent resident.

References

Alex Waibel
ahw@cs.cmu.edu
412-2687676
Professor
Interactive Systems Labs
Carnegie Mellon Univ.

Tanja Schultz
tanja@cs.cmu.edu
412-2684278
Research Scientist
Interactive Systems Labs
Carnegie Mellon Univ.

Richard Stern
rms@cs.cmu.edu
412-2682535
Professor
ECE Dept.
Carnegie Mellon Univ.

Jie Yang
yang+@cs.cmu.edu
412-2688020
Research Scientist
HCII
Carnegie Mellon Univ.

Hagen Soltau
hsoltau@us.ibm.com
914-9452778
Research Staff Member
IBM T.J. Watson Research Lab