

Pei Yin

School of Interactive Computing
College of Computing
Georgia Institute of Technology
85-5th St., N.W.
Atlanta, GA 30332-0760
+1 (404) 791-0615 Tel
+1 (404) 894-0673 Fax

pyin@cc.gatech.edu
<http://www.cc.gatech.edu/~pyin>

RESEARCH INTERESTS

Machine Learning and Computer Vision

EDUCATION

- 2005-2009 **Ph.D., Computer Science**
GEORGIA INSTITUTE OF TECHNOLOGY – Atlanta, Georgia
Advisor: Dr. Thad Starner, Dr. Irfan Essa, and Dr. James M. Rehg
Thesis: Segmental Discriminative Analysis for American Sign Language Recognition
GPA: 4.0/4.0
- 2002-2005 **M.Sc., Computer Science**
GEORGIA INSTITUTE OF TECHNOLOGY – Atlanta, Georgia
GPA: 4.0/4.0
- 1998-2002 **B.Eng., Computer Science and Technology**
TSINGHUA UNIVERSITY – Beijing, China
GPA: 86/100

SKILLS

PROFESSIONAL: C/C++, C#, MATLAB, LISP, HTML

OPERATING SYSTEMS: MS Windows, Linux/Unix

SOFTWARE: MS Office, L^AT_EX, Adobe Photoshop

LANGUAGES: English (fluent), Chinese (native)

ACADEMIC & TEACHING EXPERIENCE

- 2006,2007,2008 **Georgia Institute of Technology**, Atlanta, Georgia
GRADUATE TEACHING ASSISTANT – College of Computing
CS4641/7641 Machine Learning
CS3600 Introduction to Artificial Intelligence
CS4235 Introduction to Information Security
CS4731/8803 Game AI
- 2002-Present **Georgia Institute of Technology**, Atlanta, Georgia
GRADUATE RESEARCH ASSISTANT – College of Computing
TOPIC: Feature Selection for Sequence Classification
Sequence classification predicts one label for an entire observation sequence. For the sequences that have internal structure, such as American Sign Language, we propose extracting discriminative features in a segmental manner. Experiments on publicly available datasets for American Sign Language recognition, lip reading, speech recognition, and human gait identification show 17-70% reduction of error compared with traditional HMM-based recognizers and their variants. Our technique will be integrated to an interactive sign language game to help deaf children develop their language ability at early age.
IMPLEMENTATION: C++, MATLAB, HTK, BNT, SPHINX
- 2001-2002 **Tsinghua University**, Beijing, China
UNDERGRAD RESEARCH ASSISTANT – Department of Computer Science and Technology
TOPIC: BATCH MODE ACTIVE LEARNING FOR SUPPORT VECTOR MACHINES
Relevance feedback, which creates significant improvements in retrieval accuracy, is costly to obtain. In order to minimize the cost, we propose active learning heuristics based on the framework of the Support Vector Machine classifiers. Comparing with previous works, our heuristics are more effective with batched user feedback.
IMPLEMENTATION: C++, SVM Light

INDUSTRIAL WORK EXPERIENCE

- 2006 **Microsoft Research Cambridge (MSRC)**, Cambridge, England
SUMMER RESEARCH INTERN – MACHINE LEARNING AND PERCEPTION GROUP
TOPIC: BILAYER VIDEO SEGMENTATION
We may want to replace the background in video chat due to privacy reasons. The challenge is to retrieve the depth information from a monocular video stream. We propose a novel method to infer segmentation likelihood from the spatial context of motion. In order to choose the best learning algorithm for this application, we also propose a general taxonomy of tree-based classifiers to facilitate theoretical and experimental comparison. The resulting segmentation algorithm, based on Random Forests and Conditional Random Fields, achieves correct segmentations even in the presence of large background motion with nearly stationary foreground. This algorithm then allows the user to replace the background with a static image or a pre-recorded video.
IMPLEMENTATION: C++, MATLAB

- 2005 **Microsoft Research**, Redmond, the United States
 SUMMER RESEARCH INTERN – COMMUNICATION COLLABORATION SYSTEMS GROUP
 TOPIC: ACTIVE SPEAKER DETECTION FOR DISTRIBUTED MEETING VIDEOS
 One of the most desired features in video conferencing is to provide remote users with a close-up of the current speaker. The challenges for active speaker detection include high reverberation noise for the audio-based (microphone array) algorithms and low resolution for the video-based algorithms. In addition, the detection has to be efficient to implement on a 100-MIPS DSP chip. In this research, we fuse acoustic and visual features at signal level automatically using a decision cascade. The detection needs only two features on average while the error is reduced by more than 40%.
- IMPLEMENTATION: C#
- 2001-2002 **Microsoft Research Asia (MSR-Asia)**, Beijing, China
 LONG-TERM RESEARCH INTERN – MEDIA COMPUTING GROUP
 TOPIC: AUTOMATIC VIDEO CONTENT INDEXING AND RETRIEVAL
 Digital cameras and camcorders are producing multimedia data beyond human’s ability to manage. In this research, we study various multimedia indexing and retrieval technologies, such as shot boundary detection, semantic feature extraction, and video search. Part of the results are reported to TREC-11 Video Track.
- IMPLEMENTATION: C++

SELECTED COURSE PROJECTS

- Sentimental Learning from Computer Hardware Reviews (C++)
- Audio-visual Information Fusion and the “Cocktail Party Effect” (C++, MATLAB)
- Voting the Label: Image Understanding Using Page Rank (MATLAB)

PUBLICATIONS

REFEREED JOURNAL PUBLICATIONS

- P. Yin**, I. Essa, T. Starner, J. M. Rehg, “Segmental Feature Selection for Hidden Markov Models in Sequence Classification,” draft, to be submitted to IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI).
- P. Yin**, A. Criminisi, J. Winn, I. Essa, “Tree-based Classifiers for Bilayer Video Segmentation,” submitted to IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI).
- C. Zhang, **P. Yin**, Y. Rui, R. Cutler, P. Viola, X. Sun, N. Pinto, Z. Zhang, “Boosting-Based Multimodal Speaker Detection for Distributed Meeting Videos”, IEEE Trans. on Multimedia, pp. 1541-1552, Vol. 10, No. 8, 2008.
- S. Zhang, **P. Yin**, S. Hu, “Error propagation in Geometric Transformation,” Journal on Computer Aided Design And Computer Graphics, pp 537-540, Vol. 15, No. 5, 2003. A version of this paper also appears in Proc. CSIAM Geometric Design And Computing, pp 100-105, June 2002 as an invited paper.

REFEREED CONFERENCE PUBLICATIONS

P. Yin, T. Starner, H. Hamilton, I. Essa, James M. Rehg, “Learning Basic Units in American Sign Language Using Discriminative Segmental Feature Selection”, in Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009), Apr. 2009.

P. Yin, I. Essa, T. Starner, J. M. Rehg, “Discriminative Feature Selection for Hidden Markov Models Using Segmental Boosting,” in Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008), Mar. 2008.

P. Yin, A. Criminisi, J. Winn, I. Essa, “Tree-based Classifiers for Bilayer Video Segmentation,” in Proc of IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007), June 2007.

C. Zhang, **P. Yin**, Y. Rui, R. Cutler, P. Viola, “Boosting-Based Multimodal Speaker Detection for Distributed Meetings,” in Proc. of IEEE international workshop on Multimedia and Signal Processing, Oct. 2006, Victoria, BC, Canada.

P. Yin, I. Essa, J. M. Rehg, “Asymmetrically Boosted HMM for Speech Reading,” in Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004), II:pp 755-761, June 2004.

P. Yin, I. Essa, J. M. Rehg, “Boosted Audio-Visual HMM for Speech Reading,” in Proc. of IEEE International Workshop on Analysis on Modeling of Faces and Gestures (AMFG 2003), pp 68-73, Oct. 2003, held in conjunction with ICCV-2003. A version of this paper also appears in Proc. of Asilomar Conference on Signals, Systems, and Computers, pp 2013-2018, Nov. 2003 as an invited paper.

X. Hua, **P. Yin**, H. Zhang, “Efficient Video Text Recognition Using Multiple Frame Integration,” in Proc. of International Conference on Image Processing (ICIP 2002), II:pp 397-400, Sept. 2002.

P. Yin, X. Hua, H. Zhang, “Automatic Time Stamp Extraction System for Home Videos,” in Proc. of IEEE International Symposium on Circuits and Systems, Theme: Circuits and Systems for Ubiquitous Computing (ISCAS 2002), II:pp 73-76, May 2002.

TECHNICAL REPORTS

P. Yin, I. Essa, J. M. Rehg, “Segmental Boosting Algorithm for Time-series Feature Selection,” in Proc. of Snowbird Learning workshop, Mar. 2007.

X. Hua, **P. Yin**, H. Wang, J. Chen, L. Lu, M. Li, and H. Zhang, “MSR-Asia at TREC-11 Video Track,” in Proc. NIST TREC-VID, Nov. 2002.

P. Yin (advisors: B. Zhang, F. Lin), “SVM Active Learning in Text Classification,” Bachelor’s Thesis, Tsinghua University, June 2002.

PATENT

Patent (Applicaiton#20070297682): “Identification of People Using Multiple Types of Input” (Microsoft)

SERVICE**PROFESSIONAL**

Reviewer for IEEE International Conference on Computer Vision (ICCV),
IEEE Computer Vision and Pattern Recognition (CVPR),
IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI),
and IEEE Trans. on Circuits and Systems for Video Technology (TCSVT).

OTHER

Student advisor for freshmen,
Department of Computer Science & Technology, Tsinghua University 2000-2001

English guide at the International Olympiad in Informatics (IOI 2000),
Beijing, China Sept. 2000

HONORS & AWARDS

Dean's Fellowship Award,
College of Computing, Georgia Institute of Technology 2002-2003

Tsinghua-Lenovo, Tsinghua-ISS, Tsinghua-GE Scholarship 1999, 2000, 2001

REFERENCES

Available upon request