

Benjamin (Zhenyu) Yao

Homepage: www.stat.ucla.edu/~zyyao

Email: zyyao@ucla.edu

Phone number: 424-270-4437

Educations

- Post-doc Researcher in Computer Vision** June. 2012 - present
Center for Vision, Cognition, Learning, and Art, University of California Los Angeles (UCLA).
- Ph.D. in Statistics** June. 2012
Department of Statistics, University of California Los Angeles (UCLA).
- M.S. in Electrical Engineering** July. 2007.
Graduate University of Chinese Academy of Sciences (GUCAS), Beijing, China.
- B.S. in Electrical Engineering** July. 2003.
University of Science and Technology of China (USTC), Hefei, China.

Publications

- **Benjamin Z. Yao**, Z. Liu, S.-C. Zhu, “*Animated Pose Templates for Modeling and Detecting Human Actions*”, IEEE Trans. of Pattern Analysis and Machine Intelligence, (under review), 2012.
- Y. Lu, **B. Yao**, Y. Wang and S.-C. Zhu, “*Reconfigurable Templates for Robust Vehicle Detection and Classification*”, Workshop on Application of Computer Vision (WACV), Colorado, 2012.
- Z. Si, M. Pei, **B. Yao**, and S.-C. Zhu, “*Unsupervised Learning of Event And-Or Grammar and Semantics from Video*”, Int'l Conf. on Computer Vision (ICCV), Barcelona, Spain, 2011.
- J. Zhang, W. Hu, **B. Yao**, Y. Wang and S.-C. Zhu, “*Inferring Social Roles in Long Timespan Video Sequence*”, Int'l Workshop on Video Event Categorization, Tagging and Retrieval for Real World Applications, Spain, 2011.
- **B. Yao**, X. Yang, L. Lin, M. Lee and S.-C. Zhu, “*I2T: Image Parsing to Text Description*”, Proceedings of IEEE, 2010.
- **B. Yao**, S.-C. Zhu, “*Learning Deformable Action Templates from Cluttered Videos*”, Int'l Conf. on Computer Vision (ICCV), Tokyo, Japan, 2009.
- **B. Yao**, X. Yang and T. F. Wu, “*Image Parsing with Stochastic Grammar: The Lotus Hill Dataset and Inference Scheme*”, SIG-09: First International Workshop on Stochastic Image Grammars, CVPR'09 workshop.
- J. Porway, **B. Yao**, S.-C. Zhu, “*Learning compositional models for object categories from small sample sets*”, Book Chapter in *Object Categorization: Computer and Human Vision Perspectives*, Cambridge University Press. 2009.
- **B. Yao**, Liang Wang and S.-C. Zhu, “*Learning a Scene Contextual Model for Tracking and Abnormality Detection*”, Proc. 3rd Int'l Workshop on Semantic Learning and Applications in Multimedia, Anchorage, Alaska, June, 2008
- **Z. Yao**, X. Yang and S.-C. Zhu, “*Introduction to a large scale general purpose ground-truth dataset: methodology, annotation tool, and benchmarks*”, EMMCVPR, Springer LNCS 4679, Ezhou, China, Aug 2007.

Research experiences

Researcher at the statistics department, UCLA. (Sept. 2008 – present)

- **V2T: Video parsing, event understanding and text description**
 - Develop discriminatively trained deformable templates model for detecting and tracking objects from video.
 - Build a real-time system based on CUDA library and GPU hardware for detecting and tracking objects.
 - Propose a video event understanding algorithm based on trajectories and interactions between multiple objects.
- **Action recognition from video**
 - Model human actions as shape and optical-flow templates that can be learned from a set of aligned frames.
 - Develop a Dynamic Space-Time Warping (DSTW) algorithm as well as a HMM-type algorithm for detecting human actions from videos as well as aligning frames with templates.
 - Propose a Semi-supervised learning algorithm that learns action templates from weakly supervised videos by iterating the learning and inference steps as stated above.

Intern at the Microsoft Research, Redmond (July. 2010 – Sept. 2010)

- **Action detection from cluttered video with discriminatively trained model**
 - Build an animated action template model for detecting actions from cluttered videos and develop a learning algorithm to train model parameters based on the Latent SVM algorithm.
 - Develop a dynamic programming based inference algorithm for fast action detection from video.

Intern at the Lotus Hill Institute of Computer Vision (LHI), Ezhou, China. (Jan. 2006 – July. 2007)

- **Video surveillance project**
 - Develop an object tracking algorithm based on GMM background modeling and mean-shift tracking algorithm.
 - Co-develop the software infrastructure of the LHI video surveillance system.
 - Co-develop an object classification algorithm for recognizing the tracked objects using HoG features and SVM.
- **LHI image dataset project** (see website <http://www.imageparsing.com/>)
 - Design the data structure of the LHI dataset – a large-scale image database with ground-truth annotation.
 - Setup and maintain several public image datasets and benchmarks.

Programming skills

- Experienced with C/C++ programming for computer vision algorithms.
- Experienced with OpenCV library and Matlab programming.
- Familiar with CUDA library and GPU programming.
- Other programming language: Python, R (statistical programming).
- Familiar with programming on both Windows and Linux operation systems.

Presentations

- Presentation at the kickoff meeting of the MURI project: Knowledge representation, reasoning and learning for understanding scenes and events. UCLA, Sept. 2010
- Presentation at the review of an ONR grants collaborated by the UCLA vision group and Object Video Inc.
- Presentation at the 1st Int'l Workshop on Stochastic Image Grammars (SIG-09), CVPR'09
- Presentation at the 3rd Int'l Workshop on Semantic Learning and Applications in Multimedia (SLAM-08).
- Presentation at the EMMCVPR 2007.

Awards

- Chancellors Scholarship for Outstanding Students, UCLA, 2007-2008
- Excellent Paper Award for Research Plan for College Student, USTC, 2003
- Third-grade Outstanding Student Scholarship, USTC, 2002
- Third-grade Outstanding Student Scholarship, USTC, 1999
- Second-grade Outstanding Freshman Scholarship, USTC, 1998

Hobbies

Photography, tennis, guitar and bicycling.