

CONTACT
INFORMATION

8145 Math Sciences Building
Department of Statistics
University of California, Los Angeles
Los Angeles, CA 90095, USA

Phone: (310) 254-0735
E-mail: zzsi@stat.ucla.edu
WWW: www.stat.ucla.edu/~zzsi

RESEARCH
INTERESTS

Statistical modeling, machine learning, computer vision, information retrieval, optimization.

EDUCATION

University of California, Los Angeles, Los Angeles, California, USA

Ph.D., Statistics, July 2011

Tsinghua University, Beijing, China P.R.

Bachelor, Computer Science and Technology, July 2006

Hong Kong University of Science and Technology, Hong Kong

Undergraduate exchange student, Computer Science, September 2004 - December 2004

PUBLICATIONS

Zhangzhang Si and Song-Chun Zhu. **Learning Hybrid Image Templates (HIT) by Information Projection**. *IEEE Transactions on Pattern Recognition and Machine Intelligence* (in press).

Zhangzhang Si and Song-Chun Zhu. **Unsupervised Learning of AND-OR Templates for Object Recognition**. *ICCV workshop: Stochastic Image Grammars*, 2011.

Zhangzhang Si, Mingtao Pei and Song-Chun Zhu. **Unsupervised Learning of Event AND-OR Grammar and Semantics from Video**. *ICCV*, 2011.

Zhangzhang Si and Ying Nian Wu. **Wavelet Coding, Active Basis, and Shape Script — A Tour in the Sparse Land**. *ACM SIGMM International Conference on Multimedia Information Retrieval*, 2010.

Zhangzhang Si, Haifeng Gong, Ying Nian Wu and Song-Chun Zhu. **Learning active basis models by EM-type algorithms**. *Statistical Science* 2010, Vol. 25, No. 4, 458–475. .

Song-Chun Zhu, Kent Shi and Zhangzhang Si. **Learning Explicit and Implicit Visual Manifolds by Information Projection**. *Pattern Recognition Letter* 2010, Vol. 31, No. 2, 198–235.

Ying Nian Wu, Zhangzhang Si, Haifeng Gong and Song-Chun Zhu **Learning active basis model for object detection and recognition**. *International Journal of Computer Vision* 2010, Vol. 90, No. 2, 198–235.

Zhangzhang Si, Haifeng Gong, Ying Nian Wu and Song-Chun Zhu. **Learning mixed image templates for object recognition**. *IEEE Conference on Vision and Pattern Recognition*, June 2009.

Yingnian Wu, Zhangzhang Si, Chuck Fleming and Song-chun Zhu. **Deformable template as active basis**. *11th International Conference on Computer Vision*, 2007 (Honorable mention for Marr's prize)

Wujie Zheng, Jianmin Li, Zhangzhang Si, Fuzong Lin and Bo Zhang. **Using high-level features for video retrieval**. *International Conference on Image and Video Retrieval*, 2006.

| | | |
|-------------------|---|--------------------|
| HONORS AND AWARDS | Dissertation Year Fellowship, by <i>UCLA</i> | 2010 - 2011 |
| | Best Teaching Assistant - Honorable mention, by <i>UCLA Stat. Dept.</i> | 2010 |
| | University Fellowship, by <i>UCLA</i> | 2006 - 2009 |
| | Chancellor's Fellowship, by <i>UCLA</i> | summer 2007 - 2008 |
| | Marr's Prize – Honorary mention, by <i>ICCV</i> | 2007 |
| | GE Scholarship for Outstanding Student, by <i>General Electric</i> | 2005 |
| | Winner of China Undergraduate Mathematical Contest in Modeling | 2004 |
| | First Prize Scholarship, by <i>Nortel Networks</i> | 2004 |
| | First Prize Scholarship, by <i>OOCL, Hong Kong</i> | 2003 |
| | Second Prize Scholarship for Freshmen, by <i>Tsinghua University</i> | 2002 |

| | | |
|---------------|---------------------------|--|
| PRESENTATIONS | March 2010, Conf. MIR | Wavelet Coding, Active Basis, and Shape Script |
| | Sep. 2009, Tsinghua Univ. | Active Basis for Modeling, Learning, and Recognizing Deformable Template |
| | May 2008, UCLA | Active Basis for Modeling, Learning and Recognizing Deformable Templates |
| | Oct. 2007, ICCV | Deformable template as active basis |

RESEARCH EXPERIENCE **Learning hierarchical deformable templates for object detection** 2009 - current
Project page: <http://www.stat.ucla.edu/~zzsi/hab.html>

In this project, I work on the hierarchical active basis model, a generalized form of the active basis model towards compositional hierarchy and a stochastic image grammar. In this model, the whole object is decomposed into edge groups (i.e. parts) that can move locally relative to the object, and each edge group is in turn decomposed into Gabor elements that can move locally, relative to the center of the shapelet. This model can be automatically learned from aligned (registered) or non-aligned images, and can be used to detect articulated visual objects in challenging public datasets, with an accuracy among the best of current systems.

Discriminative adjustment of active basis model 2009 - current
Project page: http://www.stat.ucla.edu/~zzsi/abpage/abexp3/head_shoulder_adjustment.htm

This project combines the generative active basis model with discriminative optimization and improves its detection accuracy of visual objects. We find the generative model and discriminative training are two sides of a coin, and can be explained in a simple coherent formulation.

Learning image templates composed of structure and texture features 2008 - 2010
Project page: <http://www.stat.ucla.edu/~zzsi/hit.html>

This project extends the active basis model into hybrid image templates (HIT), which integrate heterogeneous features (sketch, texture, flatness and color) in a coherent information theoretic framework. For classification, the proposed HIT model has a sparser representation than many state-of-art methods and demonstrates good performances on par with state-of-art methods on commonly used benchmarks, especially when the training sample size is small.

Shape script as recursive active basis 2010
Project page: <http://www.stat.ucla.edu/~zzsi/ABEllipsoid.html>

This is an experiment to verify a hypothesis that natural visual objects can be efficiently represented by a small number of simple shape motifs (such as ellipsoids), that compose a deformable shape script model. The shape script model is a hierarchical deformable model. Inside the shape script model, shape motifs can locally perturb their locations, orientations, scales and aspect ratios locally relative to the whole object. Shape motifs are represented as deformable active basis models composed of Gabor wavelet elements; Gabor wavelet elements can also perturb their locations and orientations locally relative to the shape motif. The experiment result on images of articulated pelicans confirms the hypothesis and illustrates the power of the proposed model.

Active basis model as deformable templates

2007 - 2010

Project page: <http://www.stat.ucla.edu/~ywu/AB/ActiveBasisMarkII.html>

I work on the *active basis model* as deformable templates for image modeling. This work received honorable mention of Marr's prize in 2007, a highest award in computer vision. In this generative model, a deformable template is in the form of an active basis, which consists of a small number of Gabor wavelet elements at selected locations and orientations. These elements are allowed to slightly perturb their locations and orientations before they are linearly combined to generate the observed image. The active basis model, in particular, the locations and the orientations of the basis elements, can be learned from training images by the shared sketch algorithm. The recognition of the deformable template from an image can be accomplished by a computational architecture that alternates the sum maps and the max maps.

TEACHING EXPERIENCE

2009 Fall TA for Stat 10: Intro. to Statistical Reasoning
2009 Winter TA for Stat 13: Intro. Statistical Methods for Life and Health Sciences
2008 Fall TA for Stat 101A: Intro. Design and Analysis of Experiment

WORKING EXPERIENCE

Postdoctoral fellow at *Statistics Department of UCLA*. Lead and participate in research projects on object recognition and image parsing. My specialty is on supervised and unsupervised learning of hierarchical deformable models for highly articulated visual objects.

Software Engineer Intern at *Google Inc.* Using object recognition and machine learning techniques to improve the quality of image search.

PROFESSIONAL SERVICES

Reviewer for

- Computer Vision and Image Understanding
- International Conference on Computer Vision
- Neural Information Processing Systems
- IEEE Conference on Computer Vision and Pattern Recognition

COMPUTER SKILLS C++/C, Matlab, R, Java, Python.