

Fanyi Xiao

CONTACT INFORMATION	2078 Academic Surge University of California, Davis Davis, CA 95616, USA	WWW: fanyix.cs.ucdavis.edu Mobile: 949-491-2293 E-mail: fyxiao@ucdavis.edu
RESEARCH INTERESTS	I am interested in computer vision and machine learning. Particularly I am fascinated by the opportunities of learning from videos for all types of visual tasks related to understanding and creativity (e.g., detection, segmentation, generation, etc.).	
EDUCATION	University of California Davis, Computer Science Dept. , Davis, CA, USA Ph.D. student, Computer Science <ul style="list-style-type: none">• Advisor: Prof. Yong Jae Lee Carnegie Mellon University, Robotics Institute , Pittsburgh, PA, USA M.S., Robotics, 2014 <ul style="list-style-type: none">• Advisors: Prof. Martial Hebert and Prof. Yaser Sheikh• Thesis: <i>Model Recommendation for Large Scale Exemplar-based Object Detection</i> Central South University, Computer Science Department , Changsha, China B.S., Computer Science, 2012 <ul style="list-style-type: none">• Thesis: <i>Facial Expression Analysis with Active Appearance Model</i>	
PUBLICATIONS	<ul style="list-style-type: none">[1] Fanyi Xiao and Yong Jae Lee. Video object detection with an aligned spatial-temporal memory. In <i>European Conference on Computer Vision (ECCV)</i>, 2018.[2] Wenjian Hu, Krishna Kumar Singh*, Fanyi Xiao*, Jinyoung Han, Chen-Nee Chuah, and Yong Jae Lee (* equal contribution). Who will share my image? predicting the content diffusion path in online social networks. In <i>ACM International Conference on Web Search and Data Mining (WSDM)</i>, 2018.[3] Fanyi Xiao, Leonid Sigal, and Yong Jae Lee. Weakly-supervised visual grounding of phrases with linguistic structures. In <i>Computer Vision and Pattern Recognition (CVPR)</i>, 2017.[4] Fanyi Xiao and Yong Jae Lee. Track and segment: An iterative unsupervised approach for video object proposals. In <i>Computer Vision and Pattern Recognition (CVPR)</i>, 2016. (Spotlight presentation).[5] Krishna Singh, Fanyi Xiao, and Yong Jae Lee. Track and transfer: Watching videos to simulate strong human supervision for weakly-supervised object detection. In <i>Computer Vision and Pattern Recognition (CVPR)</i>, 2016.[6] Fanyi Xiao and Yong Jae Lee. Discovering the spatial extent of relative attributes. In <i>International Conference on Computer Vision (ICCV)</i>, 2015. (Oral presentation).[7] Fanyi Xiao and Martial Hebert. Efficient model evaluation with bilinear separation model. In <i>Winter Conference on Applications of Computer Vision (WACV)</i>, 2015.[8] Fanyi Xiao, Martial Hebert, Yaser Sheikh, Yair Movshovitz-Attias, Mei Chen, and Denver Dash. Runtime model recommendation for exemplar-based object detection. Technical report, Robotics Institute, Carnegie Mellon University, 2014.	

- [9] Zhiding Yu, Chunjing Xu, Deyu Meng, Fanyi Xiao, Wenbo Liu, and Jianzhuang Liu. Transitive distance clustering with k-means duality. In *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [10] Iljoo Baek, Taylor Stine, Denver Dash, Fanyi Xiao, Yaser Ajmal Sheikh, Yair Movshovitz-Attias, Mei Chen, Martial Hebert, and Takeo Kanade. Physical querying with multi-modal sensing. In *Winter Conference on Applications of Computer Vision (WACV)*, 2014.

AWARDS

- Best Graduate Researcher Award, CS Dept. of UC Davis, 2018
- Azure Research Award, Microsoft, 2017
- Graduate Fellowship, UC Davis, 2015
- AWS Research Grant (\$10000), Amazon Web Services, Inc., 2015
- Graduate Research Assistantship, CMU, 2013-2014
- Excellent Undergraduate Thesis, CSU, 2012
- Top Grade Scholarship (University-wide highest honor, 0.8%), CSU, 2010
- Sunward Scholarship (0.4%), Sunward Corporation, 2010
- National Scholarship (1%), Ministry of Education of China, 2009
- 1st Grade Scholarship (6%), CSU, 2009

EXPERIENCE

NVIDIA Research, Santa Clara, CA

Summer Intern

July 2017 - Oct 2017

- Work on action detection in videos. Proposed an iterative method for detecting action tubes. [Under submission]

Disney Research, Pittsburgh, PA

Summer Intern

June 2016 - Sept 2016

- Work on weakly supervised vision-language alignment (more specifically, producing segmentation masks for free-form language inputs) by exploiting linguistic structure. [CVPR 2017]

University of California Davis, Davis, CA

Graduate Student Researcher

Sept 2014 - Present

- Propose an algorithm to learn disentangled representations for the identity and pose of objects, with which we can then generate novel images. [Under submission]
- Design of a video object detection algorithm equipped with a *spatial-temporal memory*, which achieves the state-of-the-art performance on ImageNet VID dataset. [ECCV 2018]
- Study how images get propagated in a social network, with the proposed *Diffusion-LSTM*. [WSDM 2018]
- Design of a video object proposal algorithm, which takes a video as input, to output proposals that are either object-like or having distinct motion against the background. [CVPR 2016]
- Proposed a weakly supervised object detection algorithm that transfers *tracked object tubes from video*, i.e., exploiting the extra motion information, to improve object localization. Achieved state-of-the-art performance on PASCAL VOC 2007/2010 in the weakly supervised setting. [CVPR 2016]
- Development of a *visual attribute mining* algorithm which takes the dataset with image-level relative attribute annotation as the input to automatically discover visual elements corresponding to semantic attributes. [ICCV 2015]

Carnegie Mellon University, Pittsburgh, PA

Graduate Research Assistant

Sept 2012 - June 2014

- Proposed the *Bilinear Separation Model* to approximate the exemplar models with low-rank estimators which are learnt by optimizing a task-specific max-margin formulation. [WACV 2015]
- Development of a framework which uses *collaborative filtering* to recommend object detection models for testing images during runtime to avoid exhaustive search, thus scale up the exemplar-based object detection. [CMU TR 2014]

Intel Science and Technology Center, Pittsburgh, PA

Student Researcher

Sept 2012 - Aug 2013

- Development of a multi-modal sensing retailing assistant named “Marvin”. Lead developer of the visual recognition module. [WACV 2014]

Central South University, Changsha, China

Undergraduate Senior Thesis

Sept 2011 - June 2012

- Implementation and analysis of a facial expression classification system based on the Active Appearance Model (AAM) representation.

SKILLS

- Programming: Python, C/C++, Lua, MATLAB, Java
- Misc: PyTorch, Torch7, Caffe, LINUX, L^AT_EX

RELATED
GRADUATE
COURSES

- CMU: Computer Vision / Machine Learning / Convex Optimization
Math Fundamentals for Robotics / Learning-based Methods in Vision
Mechanics of Manipulation
- UC Davis: Visual Recognition

SERVICE

- Reviewer, Computer Vision and Pattern Recognition (CVPR), 2018, 2019
- Reviewer, Asian Conference on Computer Vision (ACCV), 2018
- Reviewer, Winter Conference on Applications of Computer Vision (WACV), 2015-2018