

Chenliang Xu

Prepared on December 11, 2018

CONTACT	3005 Wegmans Hall 250 Hutchison Road University of Rochester Rochester, NY 14627	PHONE: (585) 275-5426 EMAIL: chenliang.xu@rochester.edu WEB: www.cs.rochester.edu/~cxu22
CURRENT APPOINTMENT	Assistant Professor Department of Computer Science University of Rochester	ROCHESTER, NY 9/2016 - PRESENT
RESEARCH INTERESTS	Computer vision and its relations to natural language, robotics, and data science with a focus on problems in video understanding such as video segmentation, activity recognition, and multi-modal vision-and-x modeling.	
RESEARCH AREA	Computer Vision Deep Learning Multimodal Modeling	
EDUCATION	University of Michigan Ph.D. in Computer Science and Engineering ADVISOR: Jason J. Corso DISSERTATION TITLE: “Scale-Adaptive Video Understanding”	ANN ARBOR, MI 8/2016
	State University of New York at Buffalo M.S. in Computer Science and Engineering	BUFFALO, NY 2/2012
	Nanjing University of Aeronautics and Astronautics B.S. in Information and Computing Science	NANJING, CHINA 6/2010
DISTINCTIONS	National Science Foundation BIGDATA Award <i>Audio-Visual Scene Understanding</i>	2017
	Best Paper Award at 14th Sound and Music Computing Conference	2017
	University of Rochester AR/VR Pilot Grant Award <i>Real-Time Synthesis of a Virtual Talking Face from Acoustic Speech</i>	2017
	Rackham Conference Travel Grant – University of Michigan	2015, 2016
	Travel Award of IEEE ICCV 2013	2013
	Best Open Source Code Third Prize at IEEE CVPR 2012	2012
	Best Demo Prize at the 2nd Greater New York Multimedia and Vision Meeting	2012
	First Tier Jiangsu Province Award – Mathematics Modeling Competition in China	2007
	Outstanding Student Scholarship – Nanjing Univ. of Aero. and Astro.	2007, 2008, and 2009
PUBLICATIONS	PEER-REVIEWED JOURNAL ARTICLES, CONFERENCE AND WORKSHOP PROCEEDINGS ¹ <i>Qualifiers added where known: Acceptance Rate (AR), Impact Factor (IF), h5-Index (h5) provided by Google Scholar.</i>	
	2018	
	P24. Y. Tian, J. Shi, B. Li, Z. Duan, and C. Xu. Audio-visual event localization in unconstrained videos. In <i>Proc. of European Conference on Computer Vision</i> , pages 252–268, 2018. AR: 31.8%, h5: 104 (ECCV)	

¹CVPR, ICCV and ECCV are premier conferences in computer vision. For each, typical number of submissions is several thousand and the overall acceptance rate is under 30%. CVPR is the highest rated publication venue for computer vision and eighth-highest across all engineering and computer science, according to Google Scholar metrics.

- P23. L. Chen, Z. Li, R. K. Maddox, Z. Duan, and C. Xu. Lip movements generation at a glance. In *Proc. of European Conference on Computer Vision*, pages 538–553, 2018.
AR: 31.8%; h5: 104 (ECCV)
- P22. L. Chen, E. Eskimez, Z. Li, Z. Duan, C. Xu, and R. K. Maddox. Toward a visual assistive listening device: Real-time synthesis of a virtual talking face from acoustic speech using deep neural networks. *The Journal of the Acoustical Society of America*, 143(3):1813–1813, 2018.
IF: 1.8 (JASA)
- P21. L. Ding and C. Xu. Weakly-supervised action segmentation with iterative soft boundary assignment. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 6508–6516, 2018.
AR: 28%; h5: 188 (CVPR)
- P20. M. Chakraborty, C. Xu, and A. D. White. Encoding and selecting coarse-grain mapping operators with hierarchical graphs. *The Journal of Chemical Physics*, 149(13):134106, 2018.
IF: 2.8 (JCP)
- P19. S. E. Eskimez, R. K. Maddox, C. Xu, and Z. Duan. Generating talking face landmarks from speech. In *Proc. of International Conference on Latent Variable Analysis and Signal Separation*, pages 372–381, 2018.
(LVA-ICA)
- P18. T. Chen, C. Xu, and J. Luo. Improving text-based person search by spatial matching and adaptive threshold. In *Proc. of IEEE Winter Conference on Applications of Computer Vision*, pages 1879–1887, 2018.
AR: 37%; h5: 38 (WACV)
- P17. L. Zhou, C. Xu, and J. J. Corso. Towards automatic learning of procedures from web instructional videos. In *Proc. of AAAI Conference on Artificial Intelligence*, pages 7590–7598, 2018.
AR: 24%; h5: 59 (AAAI)
- P16. L. Chen, Y. Wu, A. M. DSouza, A. Z. Abidin, A. Wismüller, and C. Xu. Mri tumor segmentation with densely connected 3d cnn. In *Proc. of SPIE Medical Imaging 2018: Image Processing*, volume 105741, page 105741F, 2018.
Oral Presentation (SPIE)

2017

- P15. B. Li, C. Xu, and Z. Duan. Audio-visual source association for string ensembles through multi-modal vibrato analysis. In *Proc. of Sound and Music Computing Conference*, pages 159–166, 2017.
Best Paper Award (SMC)
- P14. L. Chen, S. Srivastava, Z. Duan, and C. Xu. Deep cross-modal audio-visual generation. In *Proc. of the on Thematic Workshops of ACM Multimedia*, pages 349–357, 2017. (ACMMMW)
- P13. L. Zhou, C. Xu, P. Koch, and J. J. Corso. Watch what you just said: Image captioning with text-conditional attention. In *Proc. of the on Thematic Workshops of ACM Multimedia*, pages 305–313, 2017. (ACMMMW)
- P12. A. Newswanger and C. Xu. One-shot video object segmentation with iterative online fine-tuning. In *The 2017 DAVIS Challenge on Video Object Segmentation - CVPR Workshops*, pages 1–5, 2017.
h5: 47 (CVPRW)
- P11. Y. Yan, C. Xu, D. Cai, and J. J. Corso. Weakly supervised actor-action segmentation via robust multi-task ranking. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1022–1031, 2017.
AR: 28%; h5: 188 (CVPR)
- P10. T. Han, H. Yao, C. Xu, X. Sun, Y. Zhang, and J. J. Corso. Dancelets mining for video recommendation based on dance styles. *IEEE Transactions on Multimedia*, 19(4):712–724, 2017.
IF: 4.0; h5: 63 (TMM)

2016 and Before

- P9. C. Xu and J. J. Corso. Actor-action semantic segmentation with grouping process models. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 3083–3092, 2016.

- P8. **C. Xu** and J. J. Corso. Libsvx: A supervoxel library and benchmark for early video processing. *International Journal of Computer Vision*, 119(3):272–290, 2016. IF: 11.5; h5: 65 (IJCV)
- P7. **C. Xu**, S.-H. Hsieh, C. Xiong, and J. J. Corso. Can humans fly? action understanding with multiple classes of actors. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 2264–2273, 2015. AR: 28%; h5: 188 (CVPR)
- P6. **C. Xu**, R. F. Doell, S. J. Hanson, C. Hanson, and J. J. Corso. A study of actor and action semantic retention in video supervoxel segmentation. *International Journal of Semantic Computing*, 7(4):353–375, 2013. Selected as a Best Paper from ICSC. h5: 12 (IJSC)
- P5. **C. Xu**, S. Whitt, and J. J. Corso. Flattening supervoxel hierarchies by the uniform entropy slice. In *Proc. of the IEEE International Conference on Computer Vision*, pages 2240–2247, 2013. AR: 28%; h5: 124 (ICCV)
- P4. **C. Xu**, R. F. Doell, S. J. Hanson, C. Hanson, and J. J. Corso. Are actor and action semantics retained in video supervoxel segmentation? In *Proc. of IEEE International Conference on Semantic Computing*, pages 286–293, 2013. AR: 30%; h5: 15 (ICSC)
- P3. P. Das, **C. Xu**, R. F. Doell, and J. J. Corso. A thousand frames in just a few words: Lingual description of videos through latent topics and sparse object stitching. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 2634–2641, 2013. AR: 26%; h5: 188 (CVPR)
- P2. **C. Xu**, C. Xiong, and J. J. Corso. Streaming hierarchical video segmentation. In *Proc. of European Conference on Computer Vision*, pages 626–639, 2012. Oral Presentation AR: 2.8%; h5: 104 (ECCV)
- P1. **C. Xu** and J. J. Corso. Evaluation of super-voxel methods for early video processing. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1202–1209, 2012. AR: 24%; h5: 188 (CVPR)

THESIS, TECHNICAL REPORTS AND SUBMISSIONS IN REVIEW

- T11. Y. Tian, Y. Zhang, Y. Fu, and **C. Xu**. Tdan: Temporally deformable alignment network for video super-resolution. Technical report, arXiv:1812.02898, 2018. (In Review)
- T10. Y. Tian, C. Guan, J. Goodman, M. Moore, and **C. Xu**. An attempt towards interpretable audio-visual video captioning. Technical report, arXiv:1812.02872, 2018. (In Review)
- T9. S. Wang, W. Zhao, Z. Kou, and **C. Xu**. How to make a blt sandwich? learning to reason towards understanding web instructional videos. Technical report, arXiv:1812.00344, 2018. (In Review)
- T8. W. Zhao, S. Wang, Z. Xie, J. Shi, and **C. Xu**. Gan-em: Gan based em learning framework. Technical report, arXiv:1812.00335, 2019. (In Review)
- T7. J. Bi, T. Xiao, Q. Sun, and **C. Xu**. Navigation by imitation in a pedestrian-rich environment. Technical report, arXiv:1811.00506, 2019. (In Review)
- T6. Y. Yan, **C. Xu**, D. Cai, and J. J. Corso. A weakly supervised multi-task ranking framework for actor-action semantic segmentation. In Review at *International Journal of Computer Vision*, 2018. (In Review)
- T5. L. Ding and **C. Xu**. Tricornet: A hybrid temporal convolutional and recurrent network for video action segmentation. Technical report, arXiv:1705.07818, 2017. (arXiv Report)
- T4. **C. Xu**, C. Xiong, and J. J. Corso. Action understanding with multiple classes of actors. Technical report, arXiv:1704.08723, 2017. (arXiv Report)

- T3. **C. Xu.** *Scale-adaptive video understanding.* PhD thesis, University of Michigan, 2016.
(Ph.D. Thesis)
- T2. S. Oh, A. Perera, I. Kim, M. Pandey, K. Cannons, H. Hajimirsadeghi, A. Vahdat, G. Mori, B. Miller, S. McCloskey, Y.-C. Cheng, Z. Huang, C.-H. Lee, **C. Xu**, R. Kumar, W. Chen, J. J. Corso, L. Fei-Fei, D. Koller, V. Ramanathan, K. Tang, A. Joulin, and A. Alahi. *Trecvid 2013 genie: Multimedia event detection and recounting.* In *NIST TRECVID Workshop*, 2013.
(NIST Report)
- T1. A. Perera, S. Oh, M. Pandey, T. Ma, A. Hoogs, A. Vahdat, K. Cannons, H. Hajimirsadeghi, G. Mori, S. McCloskey, B. Miller, S. Venkatesha, P. Davalos, P. Das, **C. Xu**, J. J. Corso, R. Srihari, I. Kim, Y.-C. Cheng, Z. Huang, C.-H. Lee, K. Tang, L. Fei-Fei, and D. Koller. *Trecvid 2012 genie: Multimedia event detection and recounting.* In *NIST TRECVID Workshop*, 2012.
(NIST Report)

FUNDING

Total Funding: \$1.8 million (\$1.2 million as PI)

Funding is sorted by start date (recent first).

- F8. **PI:** *Synthesizing Fashion Clothe* \$55,000
SOURCE: Viscovery 9/2018-8/2019
OBJECTIVE: The main objective is to design novel models to synthesize fashion clothe by different design attributes.
- F7. **PI:** *RI: Small: Learning Dynamics and Evolution towards Cognitive Understanding of Videos* \$449,990
SOURCE: Robust Intelligence, NSF 9/2018-8/2021
COLLABORATORS: Jiebo Luo (Co-PI, CS)
OBJECTIVE: The main objective is to build computational models to study higher-order inference in understanding web instructional videos.
- F6. **Co-PI:** *CDS&E: D3SC: Applying Video Segmentation to Coarse-grain Mapping Operators in Molecular Simulations* \$488,605
SOURCE: Division of Chemistry, NSF 8/2018-7/2021
COLLABORATORS: Andrew White (PI, CE)
OBJECTIVE: The main objective is to apply advances in computer vision, e.g., video segmentation and action recognition, to improve models of multiscale systems in chemistry.
- F5. **PI:** *Deep Reinforcement Learning for Instructional Video Grounding* \$60,000
SOURCE: Tencent AI Lab Rhino-Bird Program 6/2018-5/2019
OBJECTIVE: The main objective is to develop a novel deep reinforcement learning framework to perform higher-order inference in understanding web instructional videos.
- F4. **PI:** *BIGDATA: F: Audio-Visual Scene Understanding* \$650,000
SOURCE: BIGDATA Program, NSF 9/2017-8/2021
COLLABORATORS: Zhiyao Duan (Co-PI, ECE)
OBJECTIVE: The main objective is to develop algorithms to achieve a human-like audio-visual bimodal scene understanding that overcomes the limitations in single-modality analysis through big data analysis of Internet Videos.
- F3. **Co-PI:** *Real-Time Synthesis of a Virtual Talking Face from Acoustic Speech* \$50,000
SOURCE: University of Rochester AR/VR Pilot Program 8/2017-7/2018
COLLABORATORS: Ross K. Maddox (PI, BOE), Zhiyao Duan (Co-PI, ECE)
OBJECTIVE: Develop a real-time talking face by analyzing the correlations between auditory and visual signals to assist hearing-impaired people.
- F2. **PI:** *Video Understanding Beyond Human Activities* GPU Donation
SOURCE: NVIDIA GPU Grant Program 8/2017
OBJECTIVE: The main objective is to achieve a joint recognition of actors and their actions in videos that overcomes the limitations of separated or cascaded approaches.
- F1. **PI:** *Real-Time Streaming Video Object Detection* \$30,000
SOURCE: Markable, Inc. 8/2017-1/2018
OBJECTIVE: Develop an online real-time video object detection system that handles various video challenges, such as motion blur, video defocus, object occlusion and rare pose.

* IEEE CVPR Workshop *Brave New Ideas For Motion and Spatio-Temporal Representations* 2017

PROGRAM COMMITTEE

* IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016, 2017, 2018, 2019
* The AAAI Conference on Artificial Intelligence (AAAI) 2017, 2018, 2019
* European Conference on Computer Vision (ECCV) 2018
* IEEE International Conference on Computer Vision (ICCV) 2015, 2017
* British Machine Vision Conference (BMVC) 2015, 2016, 2017, 2018
* International Conference on Semantic Computing (ICSC) 2017
* Neural Information Processing Systems (NIPS) 2016
* Asian Conference on Computer Vision (ACCV) 2016, 2018
* International Symposium CompIMAGE'16 2016
* IEEE/ISPRS 3rd Joint Workshop on Multi-Sensor Fusion for Dynamic Scene Understanding 2015
* ISPRS Geospatial Week – Image Sequence Analysis 2015
* XXIIIrd ISPRS Congress 2016
* Indian Conference on Computer Vision, Graphics and Image Processing 2014

JOURNAL REVIEWER

* IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI)
* IEEE Transactions on Image Processing (IEEE TIP)
* IEEE Transactions on Circuits and Systems for Video Technology (IEEE TCSVT)
* IEEE Transactions on Multimedia (IEEE TMM)
* International Journal of Computer Vision (IJCV)
* Pattern Recognition (PR)
* Image and Vision Computing (IVC)
* Computer Vision and Image Understanding (CVIU)
* Signal Processing: Image Communication
* Machine Vision and Applications
* IET Computer Vision
* OSA Biomedical Optics Express
* Electronic Letters on Computer Vision and Image Analysis
* IPSJ Transactions on Computer Vision and Applications

UNIVERSITY SERVICES

PhD Admission Committee (CS, University of Rochester) 2016-2017, 2017-2018, 2018-2019
Lab Committee (CS, University of Rochester) 2018-2019
Website Committee (CS, University of Rochester) 2017-2018
MS Admission Committee (GIDS, University of Rochester) 2017-2018

INVITED TALKS

When Computer Vision Meets Audition: ²
From Cross-Modal Generation to Audio-Visual Scene Understanding
- Western New York Image and Signal Processing Workshop 10/2018
- Data Science Summer Colloquium Series, University of Rochester 6/2018
- Rochester Institute of Technology 11/2017

Towards the What, Who and Where of Video Understanding
- The DAVIS Challenge on Video Object Segmentation, IEEE CVPR Workshop 7/2017

Actor-Action Semantic Segmentation with Grouping Process Models
- Midwest Vision Workshop 4/2016

Scale-Adaptive Video Understanding
- School of Computing, University of Utah 3/2016
- Department of Computer Science, Stevens Institute of Technology 3/2016
- Department of Computer Science, University of Rochester 3/2016
- Department of Computer Science & Engineering, Michigan State University 2/2016

Pulling Information from Scales
- SPEECS Seminar, University of Michigan 1/2016

Action Understanding with Multiple Classes of Actors
- AI Lab Mini-Symposium, University of Michigan 5/2015

Hierarchical Video Segmentation: Methods, Perception and Application

²I had to turn down two invited platform talks at ECCV 2018 workshops due to visa issue.

- SPEECS Seminar, University of Michigan 9/2014
- LIBSVX and Video Segmentation Evaluation
- IEEE CVPR 2014 Tutorial on Video Segmentation 6/2014
- Are Actor and Action Semantics Retained in Video Supervoxel Segmentation?
- IEEE International Conference on Semantic Computing 9/2013

TEACHING

NSF REU Site: Computational Methods for Understanding Music, Media, and Minds
University of Rochester, Summer 2017, Summer 2018
 Teach a multi-day deep learning workshop to NSF REU site students.

Pre-College Data Science Program
Hajim Engineering, University of Rochester, Summer 2018
 Teach a deep learning tutorial to a group of 17 high school students (including from South Korea, China, Spain and Haiti).

CSC 249/449 Machine Vision
University of Rochester, Spring 2018
 Instructor for the cross-listed undergraduate- and graduate-level course that covers both basics and advances in modern computer vision.

CSC 298/578 Deep Learning and Graphical Models
University of Rochester, Spring 2017
 Instructor for the cross-listed undergraduate- and graduate-level course that covers both essentials in deep learning and probabilistic graphical models. It is a new CS course at UR.

CSC 577 Advanced Topics in Computer Vision
University of Rochester, Fall 2016, Fall 2017, Fall 2018
 Instructor for the graduate-level seminar course that covers special topics in computer vision.

EECS 598 Foundations of Computer Vision
University of Michigan, Fall 2015
 Graduate student instructor for this course. Lecture tutorials and reviews, hold office hours, and grade student presentations.

Beyond the M.S. Mentoring Program
University of Michigan, Fall 2015
 Encourage and provide guidance for master’s students to continue on for a Ph.D.

STUDENT ADVISING

- PH.D. ADVISEES
- D5. Zhiheng Li (CS, University of Rochester) FALL 2018 - PRESENT
 - D4. Lele Chen (CS, University of Rochester) FALL 2018 - PRESENT
 - D3. Jie Chen (CS, University of Rochester; co-supervised with Jiebo Luo) FALL 2017 - PRESENT
 - D2. Jing Shi (CS, University of Rochester) FALL 2017 - PRESENT
 - D1. Yapeng Tian (CS, University of Rochester) FALL 2017 - PRESENT

- PH.D. COMMITTEE
- S9. Nabil Hossain (CS, University of Rochester)
 - S8. Kushal Kafle (Rochester Institute of Technology)
 - S7. Taylan Sen (CS, University of Rochester)
 - S6. Haichuan Yang (CS, University of Rochester)
 - S5. Haofu Liao (CS, University of Rochester)
 - S4. Shupeng Gui (CS, University of Rochester)
 - S3. Feng Yang (CS, University of Rochester)
 - S2. Jianbo Yuan (CS, University of Rochester)
 - S1. Chris Bates (BCS, University of Rochester)

- MS STUDENTS
- M12. Yujian Wu (DS, University of Rochester) FALL 2018 - PRESENT
 - M11. Justin Tian (CS, University of Rochester) FALL 2018 - PRESENT
 - M10. Jing Bi (ECE, University of Rochester) SUMMER 2018 - PRESENT

M9. Hao Huang (CS, University of Rochester)	SUMMER 2018 - PRESENT
M8. Wei Zhang (CS, University of Rochester)	SUMMER 2018 - PRESENT
M7. Wentian Zhao (CS, University of Rochester)	SUMMER 2018 - PRESENT
M6. Shaojie Wang (CS, University of Rochester)	SUMMER 2018 - PRESENT
M5. Lele Chen (CS, University of Rochester)	SPRING 2017 - SPRING 2018
M4. Li Ding (DS, University of Rochester)	SUMMER 2017
M3. Sudhanshu Srivastava (CS, University of Rochester)	SPRING, SUMMER 2017
M2. Mingyang Zhou (ECE, University of Michigan)	SPRING 2016
M1. Yao Li (CSE, SUNY at Buffalo)	SPRING 2013

UNDERGRADUATES

U16. Chenxiao Guan (Xerox Fellow, CE, University of Rochester)	SUMMER 2018 - PRESENT
U15. Tianyou Xiao (CS, University of Rochester)	SUMMER 2018 - PRESENT
U14. Qiuyue Sun (CS, University of Rochester)	SUMMER 2018
U13. Justin Goodman (REU Student from University of Maryland)	SUMMER 2018
U12. Marc Moore (REU Student from Mississippi State University)	SUMMER 2018
U11. Amos Newswanger (CS, University of Rochester)	SUMMER 2017, FALL 2017
U10. Max Torop (DS, University of Rochester)	SUMMER 2017
U9. Zhiheng Li (Visiting Student from Wuhan University)	SUMMER 2017
U8. Wei Zhao (Visiting Student from USTC)	SUMMER 2017
U7. Wes Smith (REU Student from University of Edinburgh)	SUMMER 2017
U6. Moses Bug (REU Student from Brandeis University)	SUMMER 2017
U5. Shengqi Suizhu (CS, University of Rochester)	FALL 2016
U4. Austin Schaffer (CS, SUNY at Buffalo)	SPRING 2014
U3. Libing Wu (CS, SUNY at Buffalo)	SPRING 2014
U2. Spencer Whitt (CS, SUNY at Buffalo)	SUMMER 2013
U1. Tyler Ganter (CS, SUNY at Buffalo)	SUMMER 2013

SOFTWARE & DATA SETS

AVE 2018
 Audio-Visual Event (AVE) dataset is a large video dataset that consists of 4143 10-second videos with both audio and video tracks for 28 audio-visual events and their temporal boundary annotations. It is the largest dataset for sound event detection.
<http://www.cs.rochester.edu/~cxu22/d/ave/>

YouCook2 2018
 YouCook2 is the largest task-oriented, instructional video dataset in the vision community. It contains 2000 long untrimmed videos from 89 cooking recipes; on average, each distinct recipe has 22 videos. The procedure steps for each video are annotated with temporal boundaries and described by imperative English sentences.
<http://youcook2.eecs.umich.edu/>

A2D 2015
 A dataset and benchmark for action recognition and segmentation with multiple classes of actors. It considers seven actor classes (adult, baby, dog, etc.) and eight action classes (climb, crawl, eat, etc.). It contains 3782 videos with at least 99 instances per valid actor-action tuple.
<http://www.cs.rochester.edu/~cxu22/a2d/>

Video2Text.net 2013
 A website and web-service for automatic conversion of videos to natural language sentences based on the video content. This website showcases our work in the vision+language domain.
<http://www.video2text.net>

YouCook 2013
 Dataset of third-person cooking videos categorized into six styles of cooking and selected from open-source web videos of different kitchens and complexity levels. It contains object and action bounding boxes as well as multiple natural language descriptions of each video.
<http://www.cs.rochester.edu/~cxu22/youcook/>

LIBSVX

2012, 2013

Supervoxel library: a set of methods for early video processing by computing supervoxel segmentations as well as a quantitative benchmark for fair comparisons of those segmentations.

<http://www.supervoxels.com>

– *Winner Best Demo Prize at 2nd Greater New York Multimedia and Vision Meeting.* 6/2012

– *Winner Best Open Source Code 3rd Prize at IEEE CVPR 2012.* 6/2012

PAST POSITIONS

Research Assistant

ANN ARBOR, MI

Electrical Engineering and Computer Science

9/2014 - 8/2016

University of Michigan

ADVISOR: Jason J. Corso

PRIMARY FOCUS: Investigating unified models of video representation integrating probabilistic methods, machine learning, and dynamic adaptive graphs to advance our ability to solve the video understanding problem.

Research Intern

CUPERTINO, CA

NEC Lab American, INC.

5/2014 - 8/2014

MENTOR: Manmohan Chandraker

PROJECT: Video labeling and 3D localization for monocular road scene understanding.

Research Assistant

BUFFALO, NY

Computer Science and Engineering

8/2011 - 5/2014

State University of New York at Buffalo

ADVISOR: Jason J. Corso

PRIMARY FOCUS: Improving the representation and indexing of objects and events in large-scale video analysis by efficiently encoding the low-level perceptual entities in the video and grounding them with rich high-level semantics.