

# Xinshuo Weng

Mobile: +1-412-320-3565

Personal Website: [www.xinshuoweng.com](http://www.xinshuoweng.com)

Email: [xinshuo.weng@gmail.com](mailto:xinshuo.weng@gmail.com)

Linkedin: [www.linkedin.com/in/xinshuoweng](http://www.linkedin.com/in/xinshuoweng)

Github: [www.github.com/xinshuoweng](http://www.github.com/xinshuoweng)

## RESEARCH INTERESTS

---

### Computer Vision

- Visual Recognition: Image Segmentation, 2D Object Detection and Tracking, High-Resolution Recognition
- 2D/3D Keypoint Detection: Facial Landmark Detection, Human Pose Estimation, Hand Pose Estimation
- 2.5D Vision: Depth Estimation, Surface Normal Estimation, Ground Plane Normal and Horizon Line Estimation
- 3D Vision: 3D Object Detection, Tracking and Forecasting, Point Cloud Generation, Registration and Forecasting
- Video Analysis: Video Action Recognition, Visual Lipreading

### Machine Learning for Vision

- Deep Learning: Equivariance Modeling
- Self/Unsupervised Learning: Supervision via Consistency

## EDUCATION

---

**Carnegie Mellon University**, Ph.D. in Robotics Institute, School of Computer Science Aug.2018 - May.2023

**Carnegie Mellon University**, M.S. in Computer Vision, Computer Science, **GPA: 4.06/4.3** Aug.2016 - Dec.2017

**The Johns Hopkins University**, Summer Research Program in Computer Science Jun.2016 - Aug.2016

**Wuhan University**, China, B.S. in Electrical Engineering, **GPA(Major): 3.8(3.9)/4.0, Ranking: 2** Sep.2012 - Jun.2016

**University College Dublin**, Ireland, Exchange Program in Computer Science, **GPA: 4.1/4.2** Jan.2016 - May.2016

## WORK EXPERIENCE

---

**Research Engineer, Oculus Research Pittsburgh (Facebook Reality Lab)** Feb.2018 - Aug.2018

- *Super-Resolution Dense Landmark Detection for Hands and Body*

**Research Collaborator, Oculus Research Pittsburgh (Facebook Reality Lab)** Jan.2017 - Dec.2017

- *Unsupervised Cycle Lucas-Kanade Network for Landmark Tracking*
  - Proposed the cycle loss, minimizing the reprojection error between forward-backward tracking and initialization
  - Implemented deep inverse compositional Lucas-Kanade network and achieved unsupervised training using cycle loss
- *Supervision-by-Registration: An Unsupervised Approach to Improve the Precision of Facial Landmark Detectors*
  - Proposed a registration loss using optical flow as a free supervision to learn a coherent landmark detector in video
- *Semi-Automatic Multiview Keypoint Annotation Tool*
  - Provided 100X speed-up and coherent landmarks across viewpoints and frames using optical flow and epipolar constraint

**Research Intern, Facebook** May.2017 - Aug.2017

- *ChopNet: Hierarchical Chopping Network for Super-Resolution Facial Landmark Alignment*
  - Achieved average L2 distance error less than 0.8mm over 224 landmarks on images with resolution of  $5120 \times 3840$

**Research Intern, The Johns Hopkins University** Jun.2016 - Aug.2016

- *Unsupervised Learning of Object Semantic Parts from Internal States of CNNs by Population Encoding*
  - Proposed to learn semantic object parts in CNNs by clustering activations from the intermediate layers using K-means++
  - Achieved object detection built upon learned semantic object parts and obtained the 83% mAP on PASCAL3D+

- *Automatic Omnidirectional Parking Assist System for Autonomous Driving*
  - Implemented image stitching pipeline with omnidirectional camera calibration, homography estimation, and blending
  - Achieved lane detection using convolutional neural networks (CNNs), and built an embedded system for wheel control

Research Intern, Shanghai Baolong Automotive Corporation

Jun.2015 - Oct.2015

- *Real-Time 3D Panoramic Vehicle System with Pedestrian Detection*
  - Implemented 3D dense reconstruction with omnidirectional cameras and achieved pedestrian detection using HOG+SVM
  - Adapted Real-Time Object Detection (YOLO) network for pedestrian detection and deployed the system to the product

PUBLICATIONS

---

- **X. Weng**, K. Kitani, “Learning Spatio-Temporal Features with Two-Stream Deep 3D CNNs for Lipreading”, arXiv:1905.02540, 2019.
- **X. Weng**, K. Kitani, “Monocular 3D Object Detection with Pseudo-LiDAR Point Cloud”, arXiv:1903.09847, 2019.
- **X. Weng**, “On the Importance of Video Action Recognition for Visual Lipreading”, arXiv:1903.09616, 2019.
- A. Manglik, **X. Weng**, E. Ohn-Bar, K. Kitani, “Future Near-Collision Prediction from Monocular Video: Feasibility, Dataset, and Challenges”, arXiv:1903.09102, 2019.
- **X. Weng**, W. Han, “CyLKs: Unsupervised Cycle Lucas-Kanade Network for Landmark Tracking”, arXiv:1811.11325, 2018.
- Y. Man, **X. Weng**, X. Li, K. Kitani, “GroundNet: Monocular Ground Plane Estimation with Geometric Consistency”, arXiv:1811.07222, 2018.
- X. Dong, S. Yu, **X. Weng**, S. Wei, Y. Yang, Y. Sheikh, “Supervision-by-Registration: An Unsupervised Approach to Improve the Precision of Facial Landmark Detectors”, Computer Vision and Pattern Recognition (**CVPR**), 2018.
- **X. Weng**, S. Wu, F. Beainy, K. Kitani, “Rotational Rectification Network: Enabling Pedestrian Detection for Mobile Vision”, IEEE Winter Conf. on Applications of Computer Vision (**WACV**), 2018.
- N. Lee, **X. Weng**, V. Boddeti, Y. Zhang, F. Beainy, K. Kitani, T. Kanade, “Visual Compiler: Synthesizing a Pedestrian Pose Estimator from a Single Image”, arXiv:1612.05234, 2016.

AWARDS AND HONORS

---

- Outstanding Graduate Award, Wuhan University, 2016.
- Wuhan University Scholarship (4%), 2013, 2015, 2016.
- CSC (China Scholarship Council) Scholarship (1%), 2015.
- Yang Gui Scholarship (4%), Wuhan University, 2015.
- Undergraduate Research Fellowship, Wuhan University, 2014, 2015.
- China National Scholarship (1%), 2014.

TEACHING

---

- Teaching Assistant (TA) with Martial Hebert, **Geometry-Based Methods in Computer Vision (16-822)**, CMU, Fall 2018
- Teaching Assistant (TA) with Kris Kitani, **Computer Vision (16-385)**, CMU, Fall 2019

PROFESSIONAL SERVICE

---

- Conference Reviewer: **CVPR, ACCV, ICCV**
- Journal Reviewer: **TCSVT**

SKILLS

---

- Software: C/C++, OpenCV, Python, Torch/PyTorch/Lua, Matlab, Tensorflow, Caffe,  $\LaTeX$
- Language: Mandarin Chinese (Native), English