# XUEJIAN RONG

| | |
|---|---|
| *email* | xrong@fb.com |
| *phone* | (650)418-1381 |
| *address* | Reality Labs, Meta, 1101 Dexter Ave N, Seattle, WA 98109 |

## EDUCATION

*2013-2019*  **The City University of New York (CUNY), USA**

**Ph.D. in Electrical Engineering**
**Research Focus**: Computer Vision & Machine Learning, specifically on Semantic Scene Understanding
**Committee**: Yingli Tian, Ioannis Stamos, Jizhong Xiao, Zhigang Zhu, and Denis Demandolx
**Thesis Title**: Deep Features for Context-Aware Scene Text Image Enhancement and Interpretation

*2009-2013*  **Nanjing University of Aeronautics and Astronautics (NUAA), China**

**B.E. in Control Science and Engineering**     Outstanding Undergraduate Thesis Award

## EXPERIENCES

*2/18/2020-*
*Now*
**Applied Research Scientist, Reality Labs, Meta.** *Seattle, WA*

Working in the Computational Photography Group of Meta Reality Labs on projects related to 3D Photos/Videos and Photorealistic Metaverse, including monocular and stereo video depth estimation, novel view synthesis with neural rendering, and 3D Videos prototypes on wearable headsets at the frontier of AR/VR.

*5/06/2019-*
*8/23/2019*
**Research Intern, Facebook Research.** *Seattle, WA*

Worked in the Computational Photography group on projects related to deep burst denoising. Our proposed method is the first unified end-to-end train-able deep burst denoising framework which enables scalable joint spatial-temporal modeling through burst frames at no extra cost, and potential low-latency burst denoising on edge devices.

*5/29/2018-*
*12/14/2018*
**Research Intern, Siemens Corporate Research.** *Princeton, NJ*

Worked on the project to incrementally generate complete and consistent 2D or 3D scenes with learned scene priors, while real observations of an actual scene can be incorporated, and unobserved parts of the scene can be hallucinated. In the limit of observing real data at each point, our method converges to solving the SLAM problem. In the limit of never observing real data, it samples entirely imagined scenes from the prior distribution. Applications include autonomous driving and few-shot learning.

*9/01/2013-*
*4/30/2018*
**Research Assistant, Media Lab, CUNY.** *New York, NY*

Designed novel deep networks to model the relationship between scene text instances and context concepts in surrounding environments, which results in better image captioning and visual question answering performance. Applications include scene understanding and active agent exploration.

Designed new feed-forward network based inference algorithms for scene text detection, retrieval, and recognition in the wild, in the presence of image degradations like blur, distortion, noise, cluttered background, etc.

## SELECTED PUBLICATIONS

**Xuejian Rong**, Jia-Bin Huang, Ayush Saraf, Changil Kim, and Johannes Kopf. *Boosting View Synthesis with Residual Transfer* . IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**), 2022.

Johannes Kopf, **Xuejian Rong**, and Jia-Bin Huang. *Robust Consistent Video Depth Estimation*. IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**), 2021. (*Oral Presentation*)

**Xuejian Rong**, Denis Demandolx, Kevin Matzen, Priyam Chatterjee, and Yingli Tian. *Burst Denoising via Temporally Shifted Wavelet Transforms*. European Conference on Computer Vision (**ECCV**), 2020.

**Xuejian Rong**, Chucai Yi, and Yingli Tian. *Unambiguous Text Localization, Retrieval, and Recognition for Cluttered Scenes*. IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**).

Benjamin Planche, **Xuejian Rong**, Ziyan Wu, Srikrishna Karanam, Harald Kosch, YingLi Tian et al. *Incremental Scene Synthesis*. Thirty-third Conference on Neural Information Processing Systems (**NeurIPS**), 2019.

Haiyan Wang, **Xuejian Rong**, Liang Yang, and Yingli Tian. *Towards Weakly Supervised Semantic Segmentation in 3D Graph-Structured Point Clouds of Wild Scenes*. British Machine Vision Conference (**BMVC**), 2019. (*Oral Presentation*)

**Xuejian Rong**, Chucai Yi, and Yingli Tian. *Unambiguously Indicated Characterness for Referring Scene Text Segmentation*. IEEE Transactions on Image Processing (**TIP**).

**Xuejian Rong**, Chucai Yi, and Yingli Tian. *Unambiguous Text Localization and Retrieval for Cluttered Scenes*. IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**), 2017. (*Spotlight Oral Presentation*)

Yang Xian, **Xuejian Rong**, Xiaodong Yang, and Yingli Tian. *Evaluation of Low-Level Features for Real-World Surveillance Event Detection*. IEEE Transactions on Circuits and Systems for Video Technology (**TCSVT**), 2017.

**Xuejian Rong**, Bing Li, Aris Arditi, and Yingli Tian. *Guided Text Spotting for Assistive Blind Navigation in Unfamiliar Environments*. International Symposium on Visual Computing (**ISVC**), 2016. (*Oral Presentation*)

**Xuejian Rong** and Yingli Tian. *Adaptive Shrinkage Cascades for Blind Image Deconvolution*. IEEE International Conference on Digital Signal Processing (**DSP**), 2016. (*Oral Presentation*)

Yuancheng Ye, **Xuejian Rong**, Xiaodong Yang, and Yingli Tian. *Region Trajectories for Video Semantic Concept Detection*. ACM International Conference on Multimedia Retrieval (**ICMR**), 2016.

J. Pablo Munoz, Bing Li, **Xuejian Rong**, Jizhong Xiao, Yingli Tian, Aris Arditi. *Demo: Assisting Visually Impaired People Navigate Indoors*. The 25th International Joint Conference on Artificial Intelligence (**IJCAI**), 2016.

**Xuejian Rong**, Chucai Yi, Xiaodong Yang, and Yingli Tian. *Scene Text Recognition in Multiple Frames based on Text Tracking*. IEEE International Conference on Multimedia and Expo (**ICME**), 2014.

## SKILLS

| | |
|---|---|
| *Programming Languages* | C, C++, Python, Matlab, and Shell with practical experiences |
| *Tools* | OpenCV, PyTorch, TensorFlow, Caffe, LaTeX, GNU Linux |

## PROFESSIONAL ACTIVITIES

| | |
|---|---|
| *Conference Reviewer* | CVPR, ICCV, ECCV, NeurIPS, ICLR, AAAI, WACV, BMVC, ACCV |
| *Journal Reviewer* | TPAMI, TIP, TMM, TCSVT, JVCI |