Xiangxi Shi

Mobile: +65 9448 8095| Email: xxshi@ntu.edu.sg

EDUCATION & WORK EXPERIENCE

Nanyang Technological University, Singapore *Aug.2017-present*

Project Officer of Rapid-Rich Object Search (ROSE) Lab

University of Science and Technology of China, Hefei, China *Sept.2013-June.2017*

Bachelor of Engineering in Automation | GPA: 3.10/4

PAPERS & WORKSHOP

Xiangxi Shi, J. Cai, S. Joty, J. Gui. Watch It Twice: Video Captioning with a Refocused Video Encoder

Recently accepted by the *27th ACM International Conference on Multimedia (ACMMM19)*

- State-of-the-art among the current proposed methods
- Introduce a reinforcement learning based keyframe selection method to pick out the better key frame of a video to represent it
- Introduce a novel bi-directional video encoder based on the selected keyframe
- Train the selection model without labeled data by the weakly supervised reward calculated from generated captions

Xiangxi Shi, J. Cai, S. Joty, J. Gui. Video Captioning with Boundary-Aware Hierarchical Language Decoding and Joint Video Prediction

In Proceedings for *Neural Computing*

- Introduce a binary gate into the low-level GRU language decoder to detect the language boundaries and generate captions at phases level with a hierarchical language decoder
- Introduce the video and language reconstruction to learn the better representation for both sides

Bastan M, **Shi X**, Gu J, et al. NTU ROSE Lab at TRECVID 2018: Ad-hoc Video Search and Video to Text[J]. 2018.

- Re-implemented the CST-captioning model and enhanced it with multiple additional data resources, including static frame features, motion features and audio features
- Achieved the 3rd place in caption generation task and 5th place at retrieval task in TRECVID supported by NIST

Shi X, Kang K, Cao Y. An iterative method for optical flow estimation with motion blur[C]//2016 Visual Communications and Image Processing (VCIP). IEEE, 2016: 1-4. Present a method for estimating the optical flow of image sequences while considering the blur effect

- Performed two steps until convergence after an initial optical flow, 1) the blur kernel is estimated using the information from optical flow; 2) the optical flow is estimated considering the blur kernel
- Achieved Average of Endpoint Error (AEE) of 0.79795

OTHER WORKS

Few-Shot Recognition for Indian Food, ROSE Lab, NTU *Nov.2018 -present*

- Implement few-shot recognition to realize food recognition on multiple datasets with limited data
- Improve the few-shot learning network with a distance prediction network
- Achieve accuracy of 71.28% for base classes, 74.56% for novel classes and 60.44% for all classes, better than the initial CVPR2018 paper claimed

Dispersion Detection Algorithm in Anomaly Detection Project, ROSE Lab, NTU *Aug.2017- Oct.2018*

- Implemented an algorithm to detect the dispersion event in videos as a clue of the video anomaly detection
- Implemented a threshold-based dispersion detection based on the dense of crossover points of different humans' tracks

Fire rescue training agent, summer research in University of Newcastle, Australia *Aug. 2016-Oct. 2016*

- Built a VR system for fire rescuing training, including a VR environment, intelligent agents and hardware using Unity
- Implemented a VR environment for test using Unity and C#
- Search the escape route using greedy algorithm

Automatic Navigation of Four-rotor UAV, research training program in USTC *Jun. 2015-Oct. 2015*

- Implemented the computer vision system for an UAV to avoid the carriers and fly safely during the trip based on
- Achieve 3D ground plane region and scene depth estimation based on monocular image.
- Apply fusion of image defocus, image saturation and dark channel prior to estimate the relative depth map of scene.
- Highest score for National Undergraduate Training Programs for Innovation and Entrepreneurship

PROGRAMMING & SKILLS

C++(***NOIP Fujian 2010 First Prize***), Python, Matlab, PyTorch, OpenCV, Vim, Unity3D, VirtualBox, Unix/Linux, Git