Master Thesis Defense Shweta Mahajan

Date: Monday, 16.12.2019
Time: 16:30
Room: A032

Title: Realtime Deep Learning for Multispectral Human Detection

Abstract:
Multispectral human detection is the detection of humans using colour and thermal images i.e. sensor fusion. The complementary information provided by the two make the detector more robust to environmental factors. This work aims to leverage sensor fusion to detect humans from an Unmanned Aerial Vehicle (UAV) mounted with a visible light and thermal camera. With this in view, we define a multispectral dataset called the Fraunhofer dataset consisting of two sequences captured with our hardware setup.

Fraunhofer dataset's small size and incompatibility with a deep learning-based detector make it less than ideal for evaluation. Instead, we use the KAIST Multispectral pedestrian detection dataset to train and evaluate our proposed detectors. From state-of-the-art object detection networks, Faster R-CNN with Resnet101 and SSD with MobilenetV2 are chosen as candidate architectures. Additionally, we evaluate weighted addition, LAB and stacking (RGB-T) image fusion techniques to achieve sensor fusion. Since no fusion technique performs significantly better than others, stacking is chosen based on the simplicity of computations. We compare the RGB-T input based Faster R-CNN Resnet101 and SSD MobilenetV2 with ACF+T+THOG and other state-of-the-art detectors Halfway Fusion, Fusion RPN + BDT, and MSDS-RCNN. MSDS-RCNN performs best with a miss rate of 18.50% but at the cost of computational complexity which is ~7x compared to Faster R-CNN Resnet101 and ~46x compared to SSD MobilenetV2. Faster R-CNN Resnet101 and SSD MobilenetV2 achieve miss rates of 38.19% and 41.28% which are both better than the baseline ACF+T+THOG. Better accuracy of Faster R-CNN compared to SSD is offset by their performance in terms of speed. At ~81 FPS, SSD is ~8x faster than Faster R-CNN on an NVIDIA GeForce GTX 1080 Ti GPU.

Finally, we curate the Fraunhofer dataset by applying image processing techniques to the colour and thermal images, obtaining synchronised, aligned images of the same size which are compatible with proposed detectors. To deploy the SSD MobilenetV2 on the UAV, the network is 8-bit quantised as required by Google's Coral USB Accelerator; a USB drive-sized portable processor. On the Fraunhofer test set, the final SSD MobilnetV2 model achieves a miss rate of 89.04% and mAP of 16.78% while running at realtime inference speeds of ~14 FPS on the Coral TPU.