

IAF SYMPOSIUM ON INTEGRATED APPLICATIONS (B5)  
Satellite Commercial Applications (3)

Author: Mr. Wenyuan Du  
University of Electronic Science and Technology of China(UESTC), China

Prof. Dong Zhou  
University of Electronic Science and Technology of China(UESTC), China

Dr. Chengjun Guo  
University of Electronic Science and Technology of China(UESTC), China

RESEARCH ON VIDEO MOVING TARGET DETECTION ALGORITHM BASED ON IMPROVED  
TARGET DETECTION NETWORK**Abstract**

With the continuous development of computer vision, some valuable achievements have been made in target detection in images. However, the performance of target detection algorithms given videos in prior works are not quite satisfactory. That is because one video is composed of multiple frames of images which have temporal and spatial correlation. And those correlations are often ignored before. Thus, there are two main problems if we only use the image target detection algorithm to detect the target in videos. One is that it does not make good use of the temporal and spatial correlation between different frames, and the other is that the calculation workload will increase significantly if we calculate the result of each frame with the algorithm. To solve the first problem, this paper replaces the original Region Proposal Network (RPN) with the feature point matching algorithm based on Faster RCNN. Besides, we also introduce the temporal and spatial correlation between frames into our new algorithm. By matching the feature points of adjacent image frames, the algorithm chooses candidate frames with different sizes and proportions for each matched point. The other steps are consistent with the Faster RCNN algorithm. To solve the second question, we make the improvement based on the DFF model. The new model first distinguishes key frames from non-key frames, then improved Faster RCNN is used to directly detect key frames. And for those non-key frames, we use wrap to merge the features obtained by key frames through backbone and use FlowNet to extract the optical flow information of non-key frames. This method is able to speed up the algorithm significantly. In a word, we evaluate our architecture on the highly competitive object recognition benchmark tasks (ImageNet VID dataset). The framework proposed in this paper effectively improves the accuracy and speed of video target detection, providing a foundation for the wilder use of video target detection.