REPLY TO REVIEWERS' COMMENTS

Reference No.: TCSVT-01704-2017.R1
Manuscript Title: CNN-Based Multiple Path Search for Action Tube Detection in Videos
Authors: Erick Hendra Putra Alwando, Yie-Tarng Chen and Wen-Hsien Fang

The authors would once again like to express their gratitude to the anonymous reviewers for many thoughtful comments and constructive suggestions to enhance the quality and the readability of this manuscript. Following the suggestion by the 3rd reviewer, this manuscript has been proofread by a native speaker. For easy check-up, the modifications which address the reviewer's concern and the wording changes in the proofreading process are highlighted in blue and red, respectively.

Reply to Reviewer 3's Comments

This paper proposes a two-stream faster R-CNN based approach for action tubes detection in videos. A video localization refinement scheme is proposed to iteratively rectify the detected bounding boxes. Then, a fusion strategy combining appearance, optical flow, and motion saliency is developed to retrieve the detection scores. Finally, a multiple path search algorithm based on dynamic programming is proposed to find multiple paths in a single run. Overall, the revisions clearly addressed the concerns raised by the reviewers. I hereby recommend an acceptation with minor revision. The following issues need to be addressed in the revision.

1. The fonts in some figures, like Fig. 8, Fig. 9 and Fig. 14, are too small to read. Make them legible.

Point taken. We have enlarged the fonts in Figs. 8, 9, and 14 of the revision to make them more clear and easier to read.

2. It's not clear how the computation time was calculated? with GPU or CPU, and setting of experiments.

The computational time is the total time required by the algorithm to obtain all possible paths from all of the video frames in each dataset. Only the steps up to the fusion strategy is run with GPU and the other steps including the proposed MPS algorithm and the temporal trimming algorithm are only with CPU. For clarity, these discussions have been added in the revision. The settings of experimented are also elaborated more in Sec. IV-B of the revision.

3. The English needs to be proofread and polished by a native speaker.

Point taken. We have sent this manuscript to American Journal Experts, suggested by this journal, to have it proofread and polished by a native speaker. The invoice can also be found along with the revision.

4. The following paper in object localization and tracking using classical approach is also related to the work: Real-time obstacle detection and tracking for sense-and-avoid mechanism in UAVs. IEEE Transactions on Intelligent Vehicles, 2018.

Point taken. We have added in the reference suggested, which is related to object localization and tracking using classical approach, in Sec. II of the revision. Some related discussions of this paper (Ref. [41] in the revision) is furnished as well.