Learning to Segment Moving Objects in Videos - Code Release

Please cite our paper [3] if you are using this code for your research. The code has been tested on Ubuntu 64 machine with MATLAB versions 2013b. You need to install caffe and gop packages included following the instructions in their READMEs.

The code uses the Large Displacement Optical Flow (LDOF) [2] but any other optical flow could be used instead.

The moving objectness detector follows the R-CNN paradigm [6] : region proposals are packed into batches and evaluated through the network. Though, it can be modified for faster evaluation using Fast RCNNs [5]. The dual pathway moving objectness regressor is initialized from image and flow stack weights that each one has been finetuned to the task beforehand, as shown in the finetuning folder. For the moving objectness finetuning the standard train and test splits of VSB100 [4] and FBMS [1] have been used:

- Training:

  - In VSB100: alec-baldwin anteater avalanche big-wheel bowling campanile car-jump chrome deoksugung dominoes drone excavator floorhockey galapagos gray-squirrel guitar hippo-fight horse-riding juggling kia-commercial knot lion lion2 lukla-airport pouring-tea rock-climbing roller-coaster rolling-pin sailing sea-snake sea-turtle sitting-dog snow-shoes soccer space-shuttle swing tarantula tennis trampoline zoo
  - In FBMS bear01 bear02 cars2 cars3 cars6 cars7 cars8 cars9 cats02 cats04 cats05 cats07 ducks01 horses01 horses03 horses06 lion02 marple1 marple3 marple5 marple8 marple10 marple11 marple13 meerkats01 people04 people05 rabbits01 rabbits05

- Testing:

  - In VSB100: airplane angkor-wat animal-chase arctic-kayak ballet baseball beach-volleyball belly-dancing beyonce bicycle-race birds-of-paradise buck buffalos capoeira chameleons fish-underwater fisheye freight-train frozen-lake gokart harley-davidson hockey hockey-goals horse-gate hummingbird humpback jungle-cat kangaroo-fighting kim-yu-na koala monkeys-behind-fence nba-commercial new-york nordic-skiing octopus palm-tree panda panda-cub penguins pepsis-wasps planet-earth-1 planet-earth-2 riverboat rock-climbing rock-climbing2 salsa samba-kids shark-attack sled-dog-race slow-polo snow-leopards snowboarding snowboarding-crashes street-food swimming up-dug up-trailer vw-commercial white-tiger yosemite
  - In FBMS: camel01 cars1 cars4 cars5 cars10 cats01 cats03 cats06 dogs01 dogs02 farm01 giraffes01 goats01 horses02 horses04 horses05 lion01 marple2 marple4 marple6 marple7 marple9 marple12 people03 people1 people2 rabbits02 rabbits03 rabbits04 tennis

Parameter Tuning   The most important parameter is $para.objectness_threshold$: the lower it is, the larger the proposal pool, the better the performance, yet the slower the computation.

demos   The following demos are included:

- demo-STMOP.m: Our full method.

- demo-objectness.m: Our object proposal generation and moving objectness ranking in the frames of the sample video.

- demo-objectness-imdb.m: Our object proposal generation and moving objectness ranking in the train and test frames of VSB100 and FBMS.

- demo-supervoxels.m: Computing hierarchical supervoxels by smoothing superpixel labels in time.

- demo-tr-clustering.m: Motion segmentation via point trajectory normalized cut clustering.

References

[1] T. Brox and J. Malik. Object segmentation by long term analysis of point trajectories. In *ECCV*. 2010.

[2] T. Brox and J. Malik. Large displacement optical flow: Descriptor matching in variational motion estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(3):500–513, Mar. 2011.

[3] K. Fragkiadaki, P. Arbelaez, P. Felsen, and J. Malik. Learning to segment moving objects in videos. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[4] F. Galasso, N. S. Nagaraja, T. J. Cardenas, T. Brox, and B. Schiele. A unified video segmentation benchmark: Annotation, metrics and analysis. In *ICCV*, 2013.

[5] R. B. Girshick. Fast R-CNN. *CoRR*, abs/1504.08083, 2015.

[6] R. B. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *CVPR*, 2014.