

An Empirical Analysis of Visual Features for Multiple Object Tracking in Urban Scenes



Mehdi MiahPolytechnique Montréal



Justine PepinPolytechnique Montréal

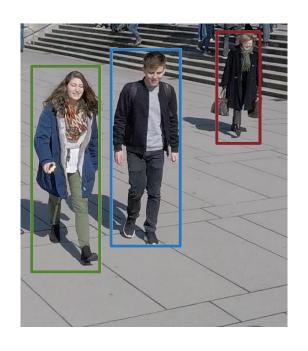


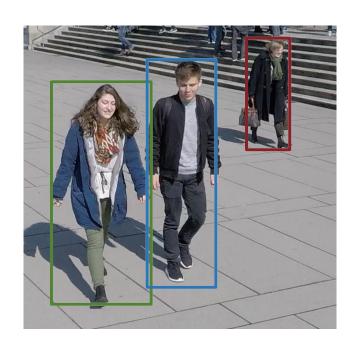
Nicolas SaunierPolytechnique Montréal

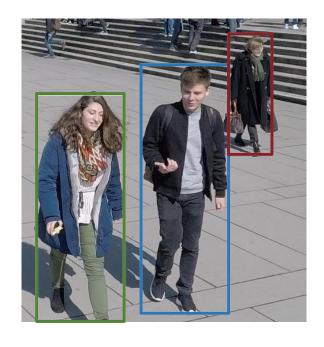


Guillaume-Alexandre Bilodeau Polytechnique Montréal

Tracking-by-detection







Objective: rank several visual features for multiple object tracking focused on urban scenes

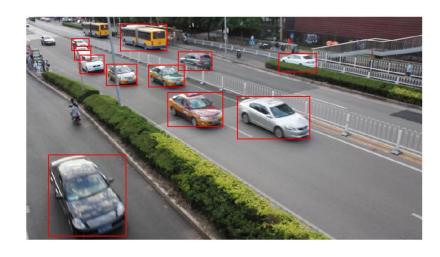
Query object at frame *t* Methodology Visual feature of the query object at frame tReturns the object's Affinity identity whose feature Descriptor is the "most similar" to measure the query's one Descriptor Set of visual features of objects at frame $t+\Delta t$

Candidate objects at frame $t+\Delta t$

Methodology



Query object at frame t



Candidate objects at frame t+1

White Gaussian noise parametrized by sigma

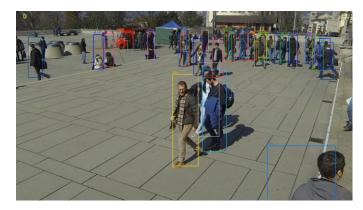
Skip some frames parametrized by step

Methodology

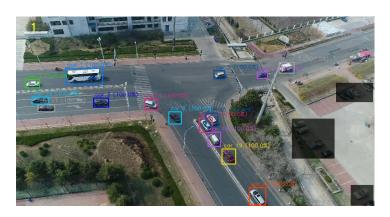
Visual features: **Affinity measures:** ☐ RGB histogram ☐ L1 distance ☐ grayscale histogram ☐ L2 distance histogram of oriented-gradients ☐ Rank-1 counts VGG-19 features Bhattacharyya distance ResNet-18 features cosine similarity DenseNet-101 features | EfficientNet-B0 features ☐ re-identification features (persons or vehicles)

Methodology

Datasets:



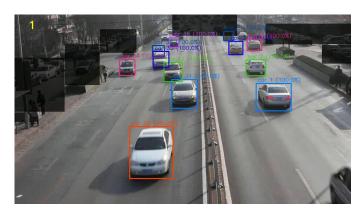
WildTrack



UAVDT



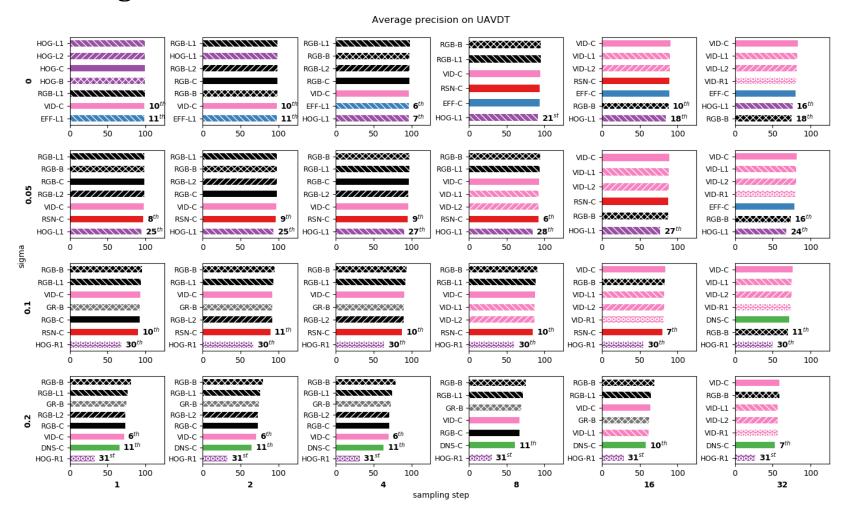
MOT17



DETRAC

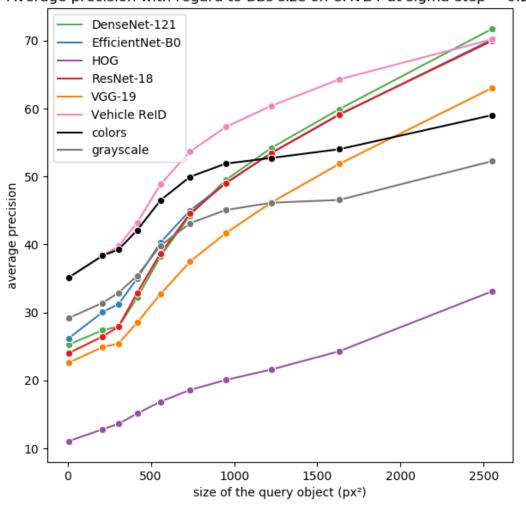
Results: average precision of 35 pairs visual features-affinity, over 24 configurations and 4 datasets

Results on ranking



Results on size effect





An Empirical Analysis of Visual Features for Multiple Object Tracking in Urban Scenes



Mehdi MiahPolytechnique Montréal



Justine PepinPolytechnique Montréal



Nicolas SaunierPolytechnique Montréal



Guillaume-Alexandre Bilodeau
Polytechnique Montréal

More information available on www.mehdimiah.com/visual_features