Large-displacement 3D Object Tracking with Hybrid Non-local Optimization

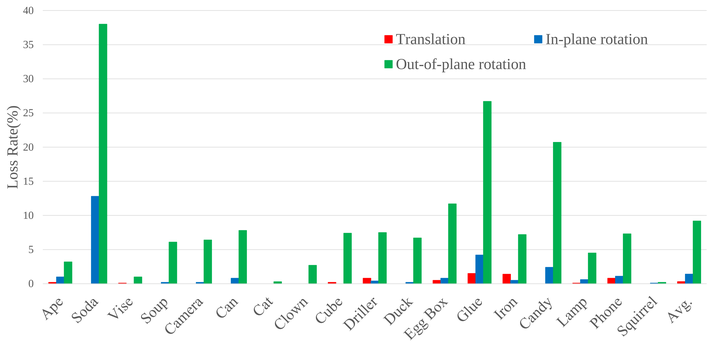
Xuhui Tian, Xinran Lin, Fan Zhong, Xueying Qin
Shandong University, China

ECCV TEL AVIV 2022

Introduction

Motivation:

- Most previous optimization-based 3D object tracking methods only search for the local minima of cost function and thus are sensitive to large inter-frame displacements.
- We decompose the rotation as in-plane rotation and out-of-plane rotation and find that most tracking failures in previous methods are caused by out-of-plane rotations.

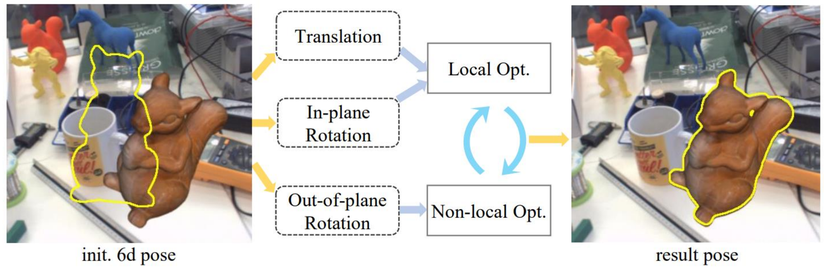


The statistics of RBGT tracking failures

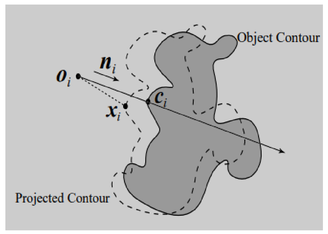
Contribution:

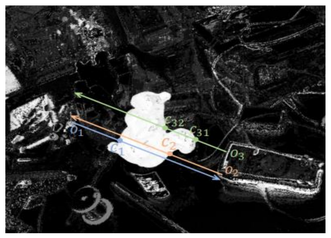
- We proposed a hybrid non-local tracking method to handle large displacements and can run in real-time with only CPU.
- An improved contour-based local tracking method with long pre-computed search lines and multiple candidate correspondences is proposed.
- We proposed an efficient non-local optimization only for the out-of-plane rotation

Method



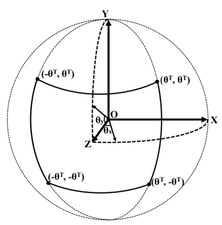
- A fast local pose optimization method that is more adaptive to frame displacements are proposed, with long search lines that can be precomputed for acceleration.

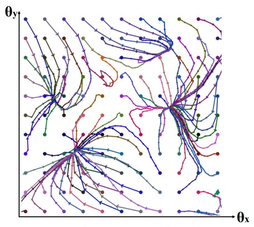




Left: the contour-based tracking model. Right: the exemplar search lines and correspondences

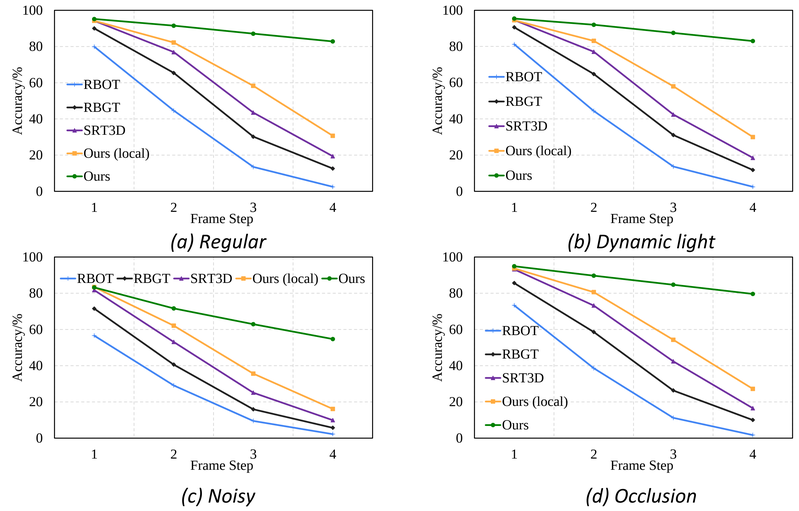
- Non-local search is applied only for out-of-plane rotation, which only requires to do sampling and conduct grid search in a 2D space. We also propose grid pre-termination, path pre-termination and near-to-far search for further acceleration.





Left: The parameterized 2D sampling space of out-of-plane rotation. Right: Converge paths of grid search in the 2D space.

Experiments



Comparison for different frame steps on RBOT. Frame step means the interval between 2 tracked frames. Previous methods' accuracies decrease rapidly with larger displacements while ours can still maintain high accuracies.

	naive	+GP	+PP	+GP&PP	+GP&PP&N2F	local(*Ours*$^-$)
UpdateItrs	1085	718	823	578	468	59
Time	47.7ms	34.0ms	37.9ms	27.6ms	22.3ms	9.7ms
Accuracy	85.1	83.5	84.6	82.7	81.7	30.7

The ablation studies to the acceleration strategies. Using them all can reduce more than half of the time with only about 3% sacrifice in accuracy compared with naïve grid search.

(a) Initial pose (b) SRT3D (c) Ours(local) (d) Ours

Visual examples and comparisons of large displacements (frame step = 8). Our non-local method can still successfully track the object while SRT3D fails.