SUSTech

UNIVERSITY OF BIRMINGHAM

JUNE 18-22, 2023

CVPR

VANCOUVER, CANADA

Resource-Efficient RGBD Aerial Tracking

Jinyu Yang[1,2,†], Shang Gao[1,†], Zhe Li[1,†], Feng Zheng[1,3], Ales Leonardis[2]

1 Southern University of Science and Technology

2 University of Birmingham

3 Peng Cheng Laboratory

Overview

- **Background:** current drone-based research scenes focus on urban environments from a birds-eye view.
- **Motivation:** adding depth information can more effectively deal with target and background interference, while RGBD aerial tracking is still unexplored.
- **Contribution 1:** we explore RGBD aerial tracking in an overhead space, which can greatly enlarge the development of drone-based visual perception.
- **Contribution 2:** we propose a large-scale benchmark for RGBD aerial tracking, containing 1,000 drone-captured RGBD videos with dense annotations.
- **Contribution 3:** we propose an efficient RGBD tracker named EMT. Our tracker runs at over 100 fps on GPU, and 25 fps on the edge platform, benefiting from its efficient multimodal fusion and feature matching.

https://github.com/yjybuaa/RGBDAerialTracking

Motivation

- Drones have visibility and illumination limitations in color view. While RGBD aerial tracking is effective to tackle such failures.

- UAV tracking datasets record videos in urban environments from a birds-eye view, while overhead space (2 - 5 meters above the ground) is unexplored.

- New task brings challenges: complex real-world circumstances, limited onboard computational resources, real-time practical applications, and so on…

Therefore, datasets and solutions for RGBD aerial tracking are necessary.

Contributions

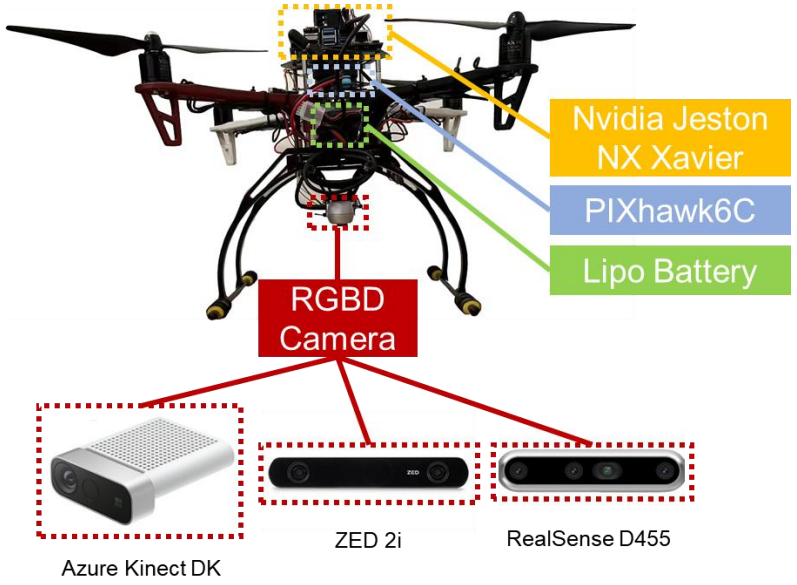
- **New Problem:** We propose a new task of RGBD aerial tracking for newly defined overhead space (2m - 5m).

- **New Benchmark:** We construct a large-scale high-diversity benchmark for RGBD aerial tracking. The advantage is that more categories can be considered than existing aerial tracking datasets.

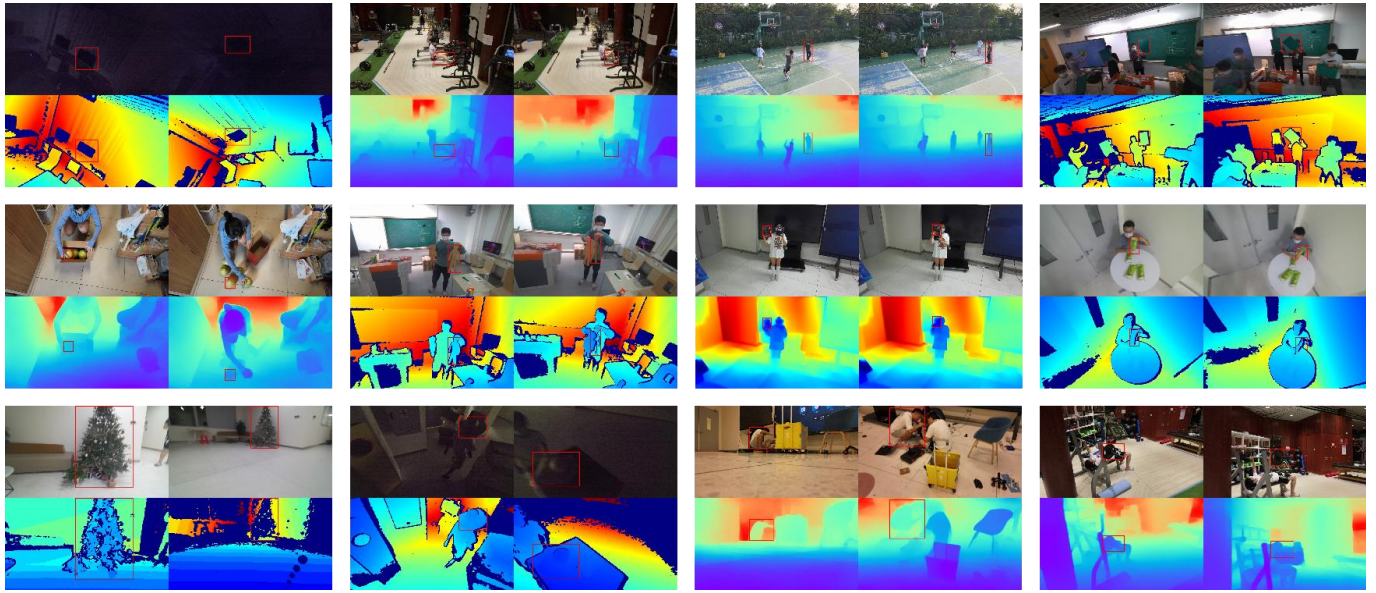
- **New Baseline:** An efficient tracking baseline is proposed for RGBD aerial tracking, which is the first real-time tracker for efficient on-board multimodal tracking.

D²Cube Dataset

- Data Collection:



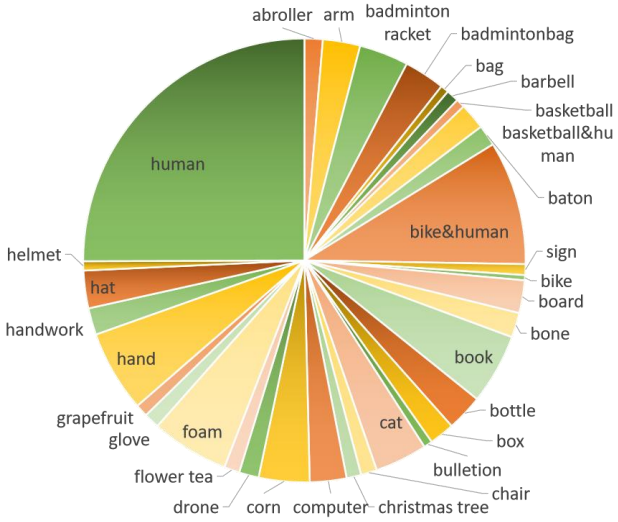
(a) Flight platform



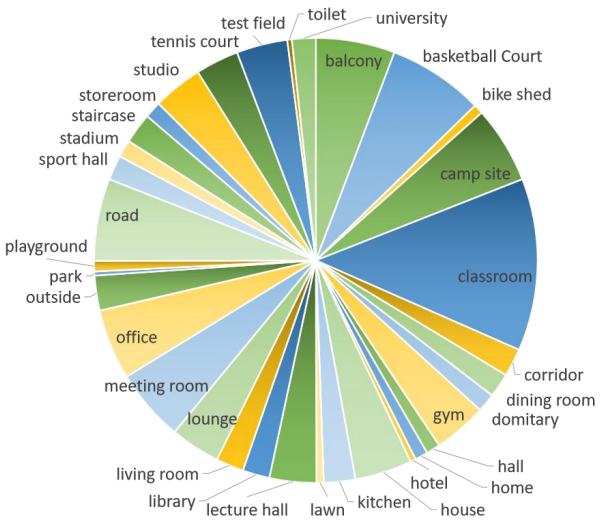
(b) Example video sequences

D²Cube Dataset

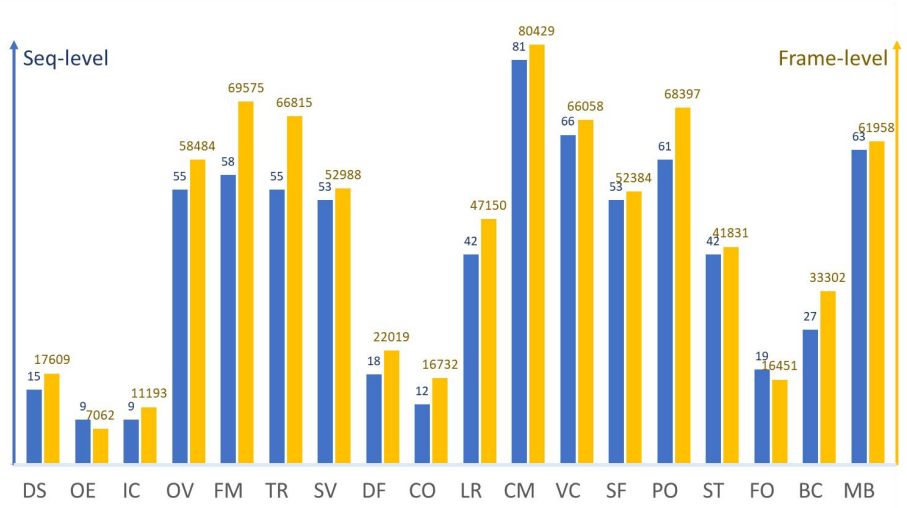
- Data Statistics:



(a) Data distribution of objects

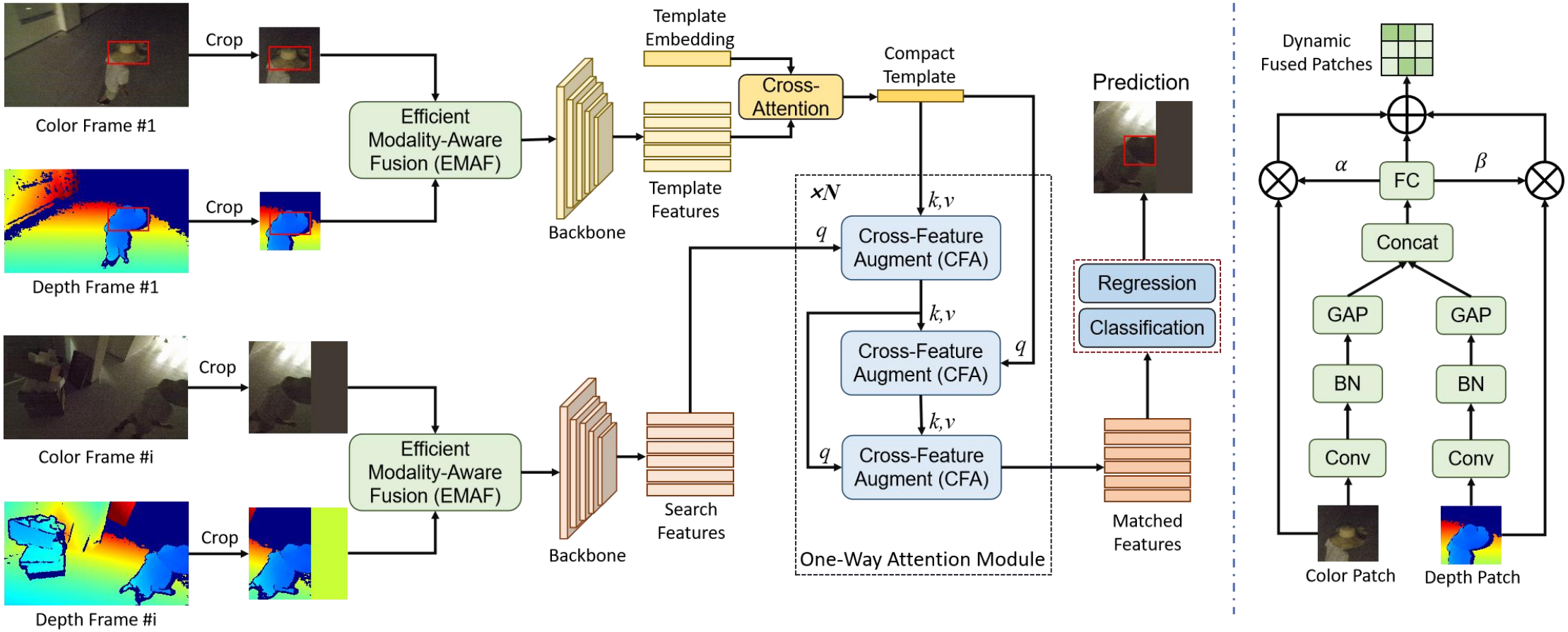


(b) Data distribution of scenarios



(c) Data distribution of attributes

Efficient Multimodal Tracker

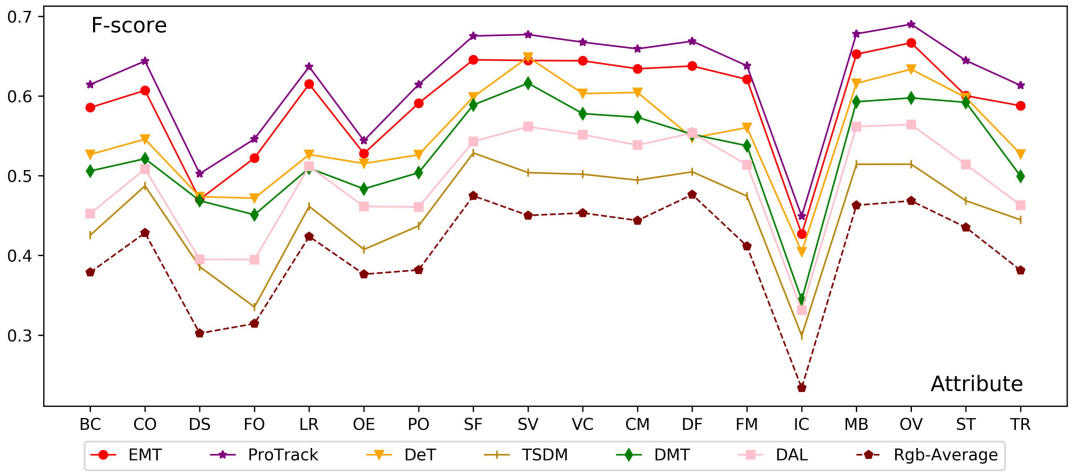


Experiments

- Experimental comparison:

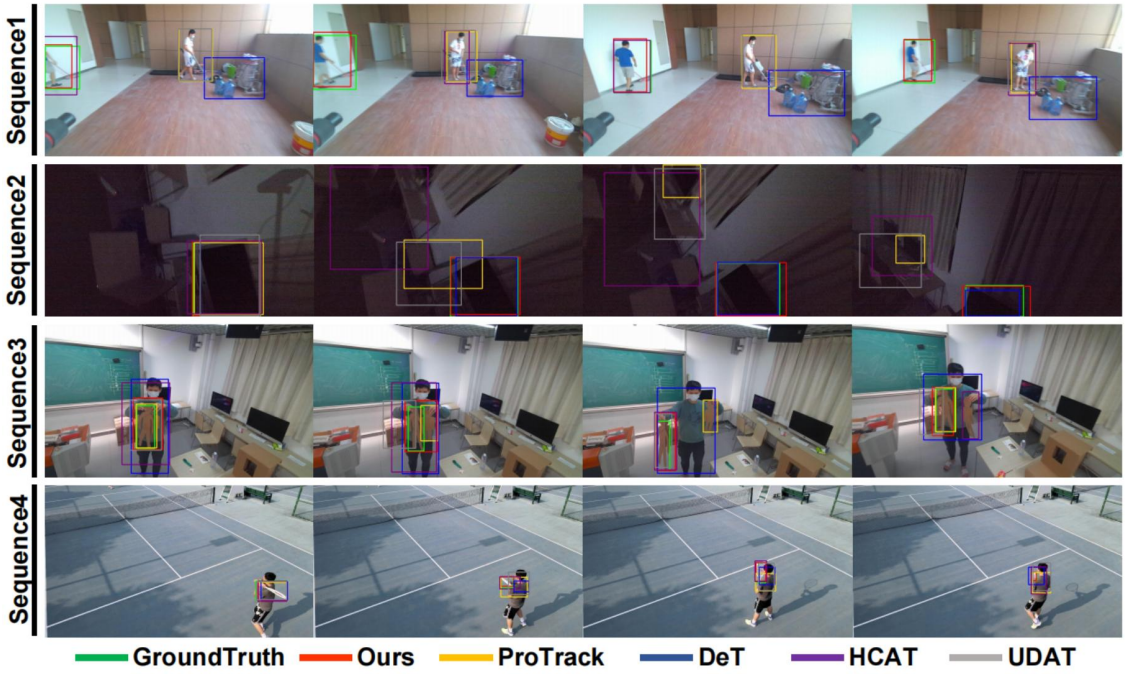
Method	Pr	Re	F-score	Speed
LightTrack [33]	0.500	0.531	0.515	119.5
HiFT [2]	0.404	0.430	0.417	66.9
TCTrack [4]	0.416	0.448	0.432	78.1
SiamAPN [9]	0.413	0.441	0.427	140.2
SiamAPN++ [3]	0.411	0.436	0.423	114.9
DaSiamRPN [44]	0.392	0.415	0.403	200.6
HCAT [5]	0.544	0.578	0.561	148.2
UDT+ [31]	0.387	0.412	0.399	50.4
SiamRPN++ [23]	0.459	0.488	0.473	83.3
UDAT-CAR [39]	0.462	0.492	0.476	33.9
EMT	0.653	0.609	0.630	120.3

Method	DAL [28]	TSDM [42]	DeT [34]	DMT [13]	ProTrack [37]	EMT
Pr	0.529	0.521	0.608	0.584	0.669	0.653
Re	0.565	0.492	0.587	0.569	0.644	0.609
F-score	0.547	0.506	0.597	0.576	0.656	0.630
MACs	15.78G	74.08G	30.57G	40.44G	82.58G	3.43G
Params	19.60M	114.59M	34.63M	38.97M	159.61M	10.05M
Speed	21.3	18.2	26.8	25.5	5.4	120.3

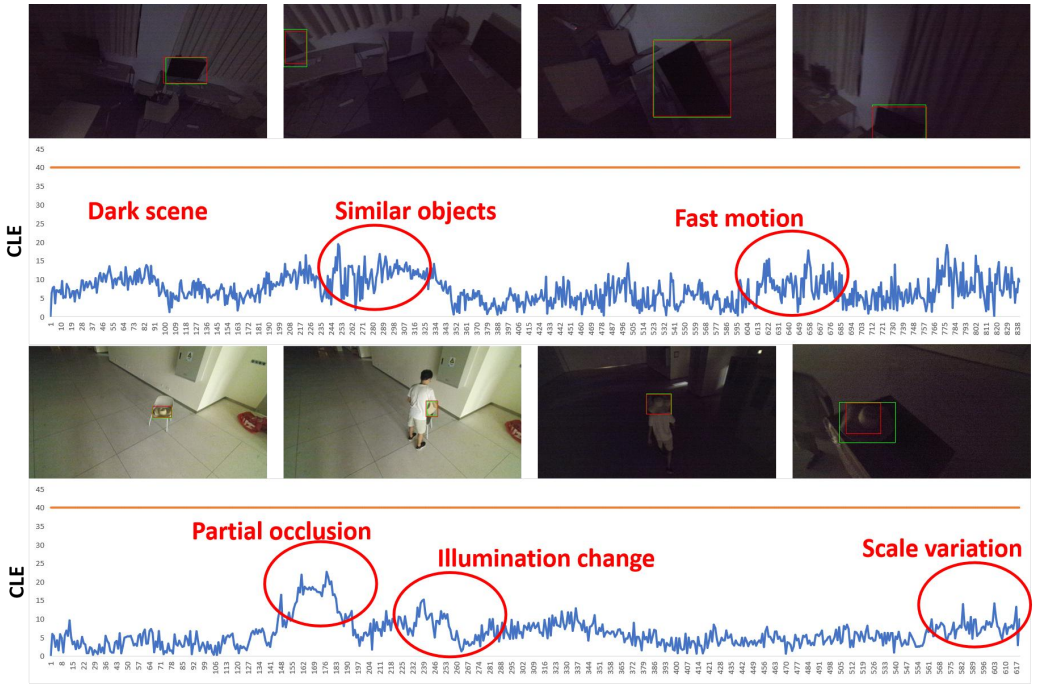


Experiments

- Qualitative results:



- On-board tests:



SUSTech

UNIVERSITY OF BIRMINGHAM

JUNE 18-22, 2023

CVPR

VANCOUVER, CANADA

Thanks for Watching!

jinyu.yang96@outlook.com

GitHub

WeChat