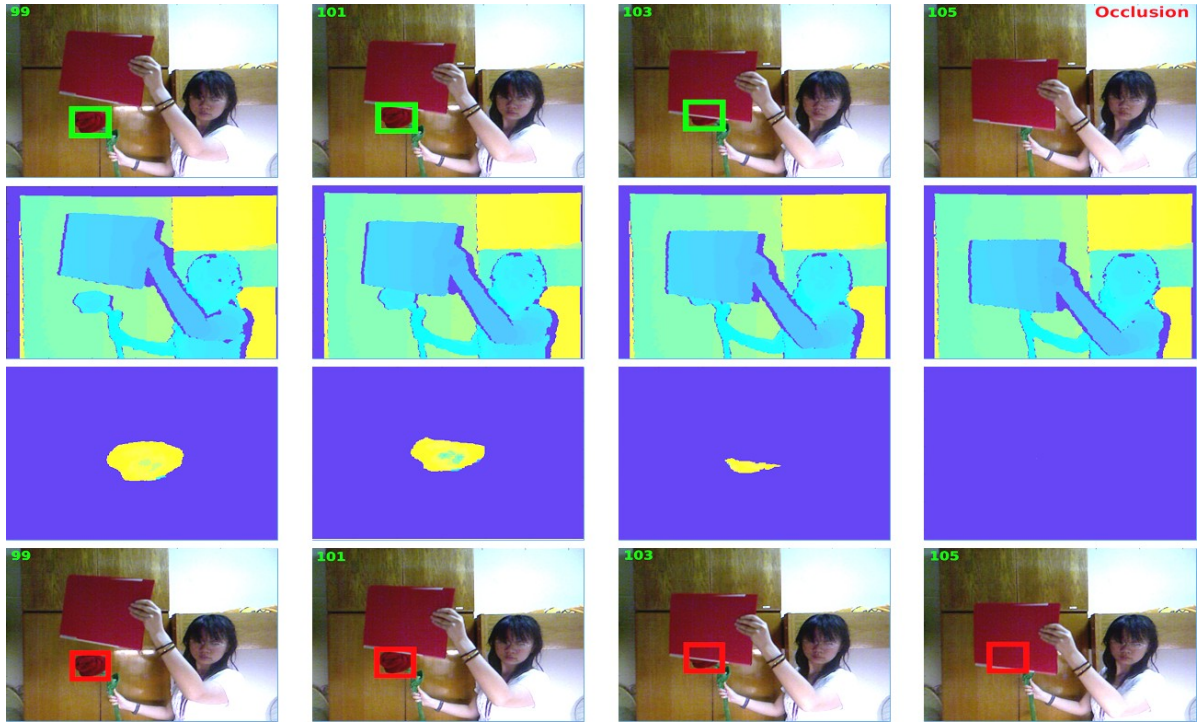
VOT
visual object tracking

The Visual Object Tracking Challenge Results
VOT-RGBD 2020

Matej Kristan, Aleš Leonardis, Jiri Matas, Michael Felsberg, Roman Pflugfelder, *Joni-Kristian Kämäräinen*, Martin Danelljan, Luka Čehovin Zajc, Alan Lukežič, Ondrej Drbohlav, Linbo He, Yushan Zhang, Song Yan, Jinyu Yang, et al.

Why to add depth?

- In robotics people ask »Why to add RGB?«
- Depth simplifies:
  - 3D sensing
  - Occlusion detection
  - Object-to-background separation
- Challenges
  - How to combine RGB and D?
  - OR when to switch between RGB and D?
  - Much less RGBD than RGB tracking datasets



RGBD Datasets

- Existing datasets:
  - PTB [1]: 100 sequences, average sequence length: 214 frames, short disappearances, synchronization problems, indoor only
  - STC [2]: 36 sequences, average sequence length: 255 frames, no target disappearances, limited outdoor scenarios
- Problems:
  - Small number of sequences and/or they are short
  - Depth is overly stable (lack of out-of-plane rotations)
  - Single sensor used (overfitting)

[1] S. Song et al., Tracking Revisited Using RGBD Camera: Unified Benchmark and Baselines, ICCV 2013.
[2] J. Xiao et al., Robust Fusion of Color and Depth Data for RGBD Target Tracking Using Adaptive Range-Invariant Depth Models and Spatio-Temporal Consistency Constraints. IEEE TCyB 2018..

The VOT-RGBD 2020 Dataset (2019)

- 80 sequences, average length 1274 frames
- Frequent and long-lasting target disappearances
  - Average target absence period: 56 frames
- Significant target 3D pose changes
- Axis-aligned bounding box
- Per-frame visual attributes:

  Full Occlusion, Target out-of-frame,
  Partial occlusion, Aspect change, Size change, Fast motion, Similar objects, Out-of-plane rotation,
  Reflective target, Depth change, Deformable target, Dark scene, Unassigned

[3] CDTB: A Color and Depth Visual Object Tracking Dataset and Benchmark (A. Lukezic, U. Kart, J. Käpylä, A. Durmush, J.-K. Kämäräinen, J. Matas and M. Kristan), In Int. Conf. on Computer Vision (ICCV), 2019.

The VOT-RGBD 2020 Dataset

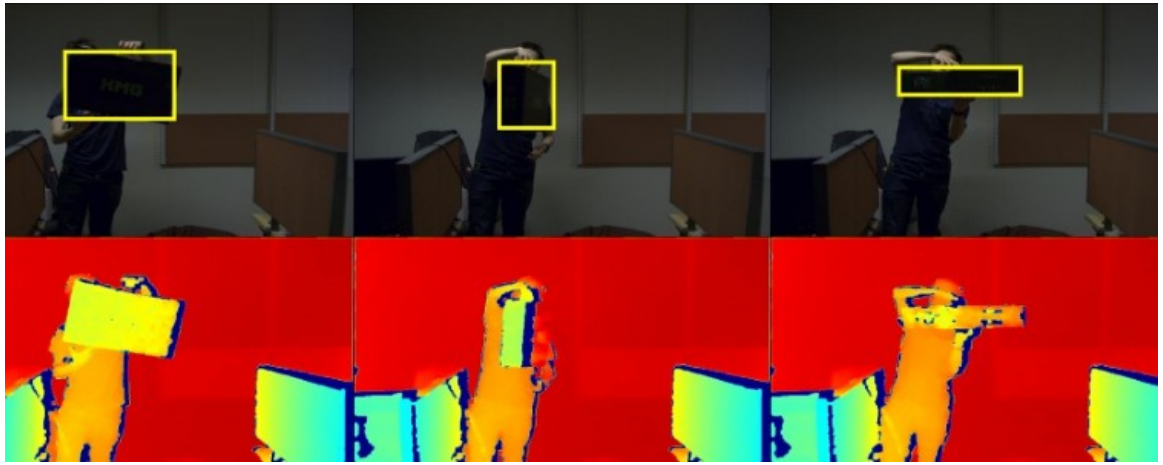
- Diverse sensors:

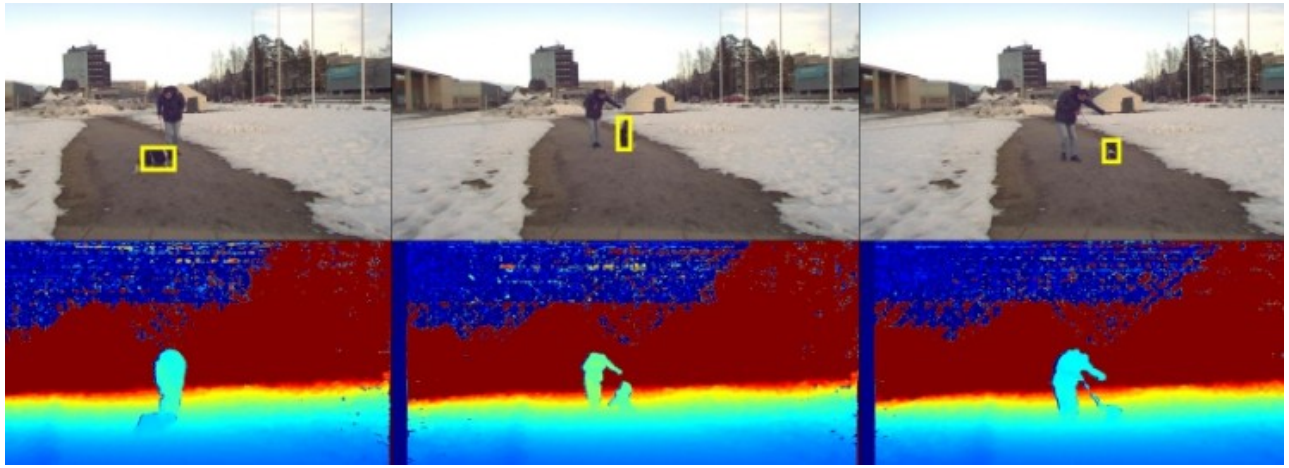
Kinect-v2:

ToF-RGB pair:

Stereo-camera pair:

- Indoor and outdoor sequences, spatially aligned and temporally synchronized (ask for raw data)





VOT-RGBD setup

- Long-term tracking
  - Full/partial occlusions (occlusion handling needed)
  - Target may disappear (re-detection needed)
  - Evaluation uses the VOT-LT protocol
- No sequestered data (thanks to COVID-19)
- Required outputs: bounding box and presence confidence

VOT-RGBD 2020 Performance Evaluation

- No-reset experiment
  - Tracker initialized in the first frame, tracks to the end of the sequence
- Long-term tracking performance measures [1,2]
  - Tracking Precision (Pr): accuracy of predicted bboxes (when predictions given)
  - Tracking Recall (Re): accuracy of predicted bboxes (when target visible)
  - Tracking F-measure: F = (2 * Pr * Re) / (Pr + Re)

[1] A. Lukežič et al., Performance Evaluation Methodology for Long-Term Visual Object Tracking, arxiv: abs/1906.08675.
[2] M. Kristan et al., The sixth Visual Object Tracking VOT2018 challenge results, ECCVW 2018.

VOT-RGBD from 2019 to 2020

2019 – 4 valid entries (1. SiamDW-D, 2. ATCAIS, 3. LTDSEd, and 4. SiamM_Ds)

2020 – 4 valid entries (ATCAIS, CLGS_D, DdiMP, and Siam_LTD)

We added baselines and the best trackers from all VOT categories

VOT-RGBD 2020 Results

Tracker	Pr	Re	F-Score	ST/LT	RGB/RGBD
ATCAIS	0.709②	0.696①	0.702①	LT	RGBD
DDiMP	0.703③	0.689②	0.696②	ST	RGBD
CLGS_D	0.725①	0.664	0.693③	LT	RGBD
SiamDW_D	0.677	0.685③	0.681	LT	RGBD
LTDSEd	0.674	0.643	0.658	LT	RGBD
RLT_DiMP	0.625	0.632	0.629	LT	RGB
LTMU_B	0.680	0.581	0.626	LT	RGB
Megtrack	0.694	0.551	0.614	LT	RGB
RPT	0.601	0.546	0.572	ST	RGB
Siam_LTD	0.626	0.489	0.549	LT	RGBD
OceanPlus	0.577	0.502	0.537	ST	RGB
AlphaRef	0.491	0.547	0.518	ST	RGB

Table 6. List of trackers that participated in the VOT-RGBD2020 challenge along with their performance scores (Pr, Re, F-score) and categorizations (ST/LT, RGB/RGBD). 2020 submissions are ATCAIS, DDiMP, CLGS_D and Siam_LTD. SiamDW_D and LTDSEd are 2019 submissions (SiamDW_D was the winner). RGB trackers are the three top performers of VOT-ST2020 and VOT-LT2020.

VOT-RGBD 2019 Per-Attribute Analysis (F-measure)

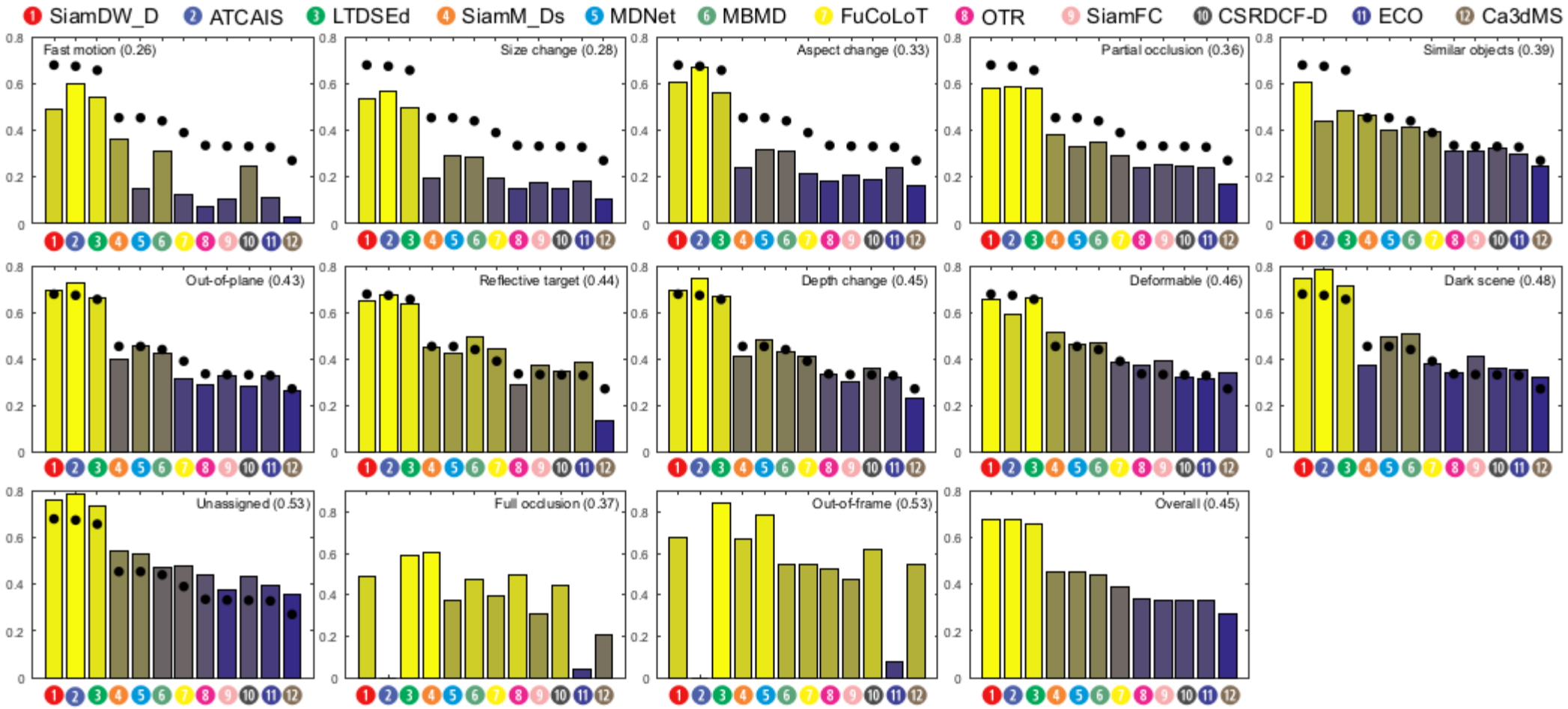


Figure 16. VOT-RGBD2019 challenge: tracking performance w.r.t. visual attributes. The first eleven attributes correspond to scenarios with a visible target (showing F-measure). The overall tracking performance is shown in each graph with black dots. The attributes full occlusion and out of view represent periods when the target is not visible and true negative rate is used to measure the performance.

VOT-RGBD 2020 Per-Attribute Analysis (F-measure)

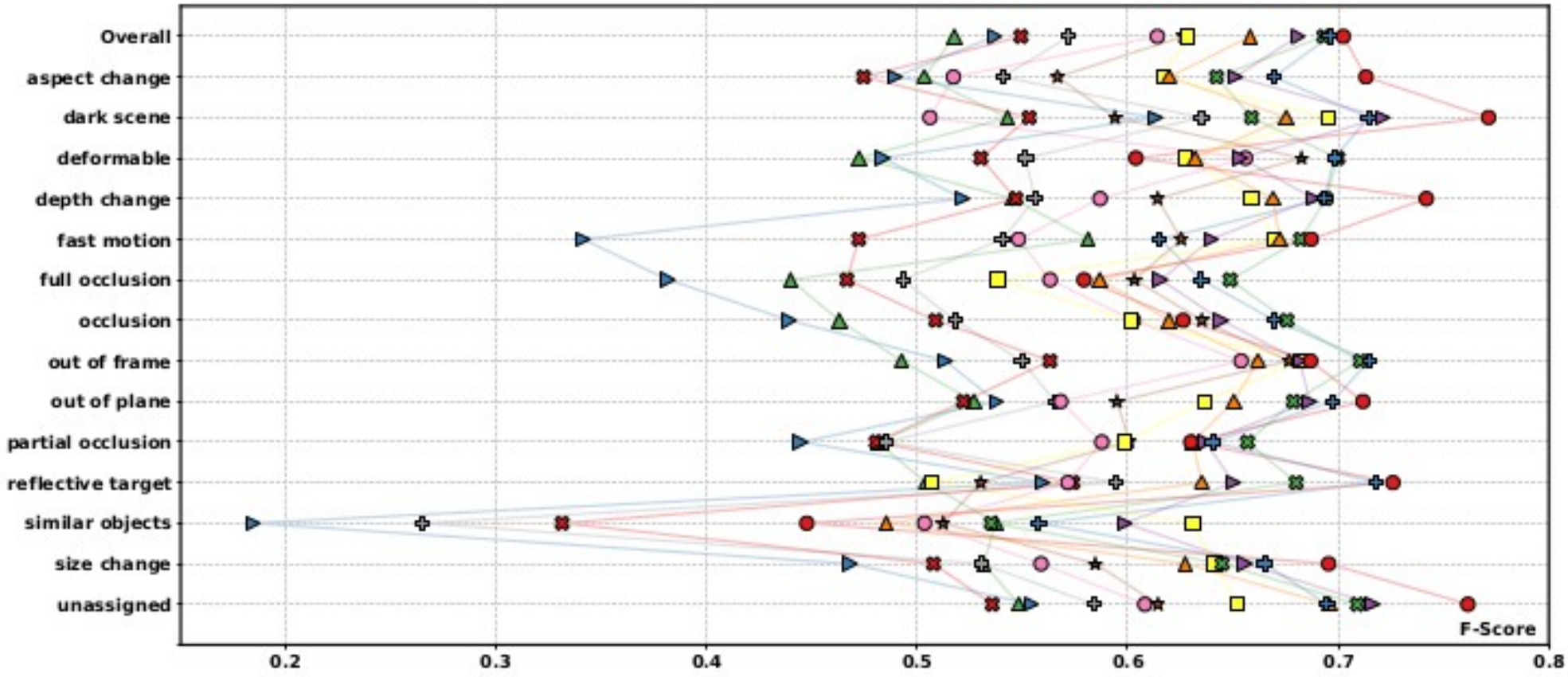
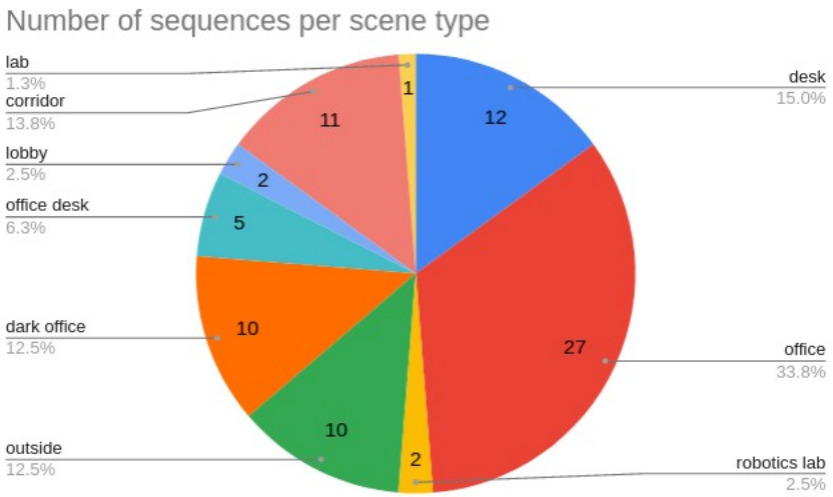
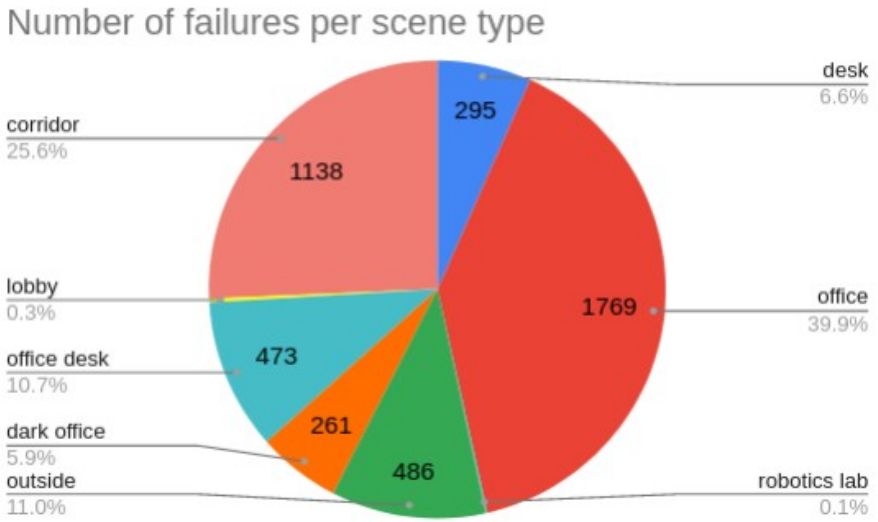
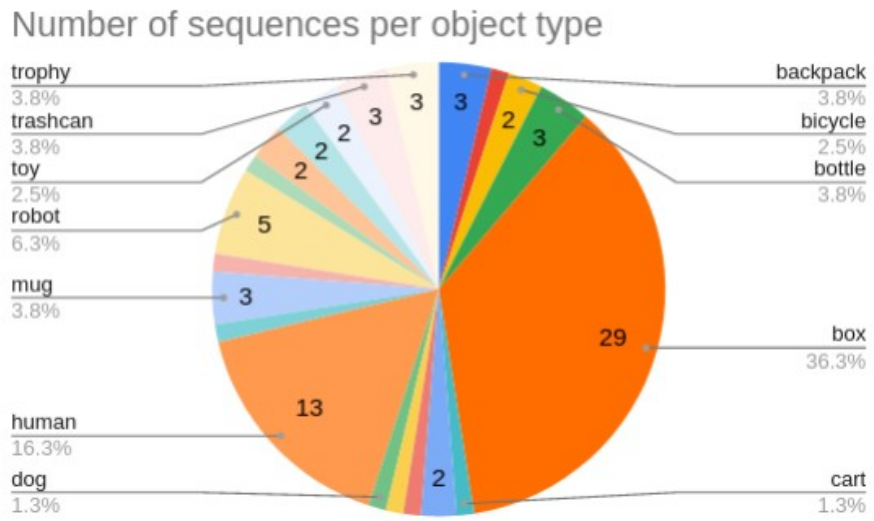


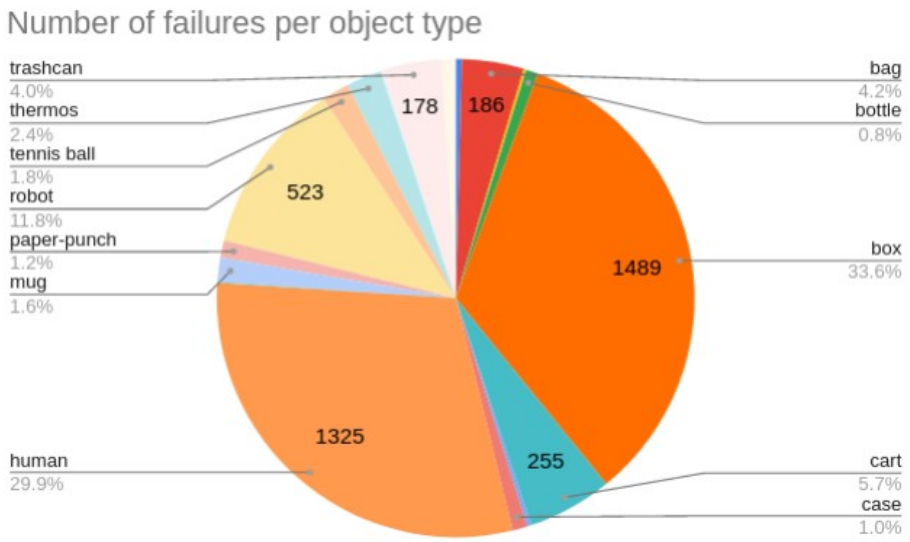
Fig. 14. VOT-RGBD2020 challenge: tracking performance w.r.t. visual attributes.

VOT-RGBD 2020 General Findings









VOT-RGBD 2020 Challenge Summary

Ingredients of the best RGBD trackers:

- Take the best available long-term RGB tracker with all tricks and hacks
- Use depth-based hacks and tweaks for occlusion detection

Open questions:

- Can depth provide discriminative features for tracking?
- Can we learn optimal depth usage from data?
- Depth for 3D tracking?

[1] K. He et al., Deep Residual Learning for Image Recognition, CVPR 2016.
[2] M. Danelljan et al., ATOM: Accurate tracking by overlap maximization, CVPR 2019.
[3] Q. Wang, et al., Fast online object tracking and segmentation: A unifying approach, CVPR 2019.
[4] B. Li et al., High Performance Visual Tracking With Siamese Region Proposal Network, CVPR 2018.

VOT-RGBD 2020 Awards

Winner of the VOT-RGBD 2020 challenge:

ATCAIS by Y. Wang, L. Wang, D. Wang, H. Lu and X. Yang

Tracker	Pr	Re	F-Score	ST/LT	RGB/RGBD
ATCAIS	0.709②	0.696①	0.702①	LT	RGBD
DDiMP	0.703③	0.689②	0.696②	ST	RGBD
CLGS_D	0.725①	0.664	0.693③	LT	RGBD
SiamDW_D	0.677	0.685③	0.681	LT	RGBD
LTDSEd	0.674	0.643	0.658	LT	RGBD
RLT_DiMP	0.625	0.632	0.629	LT	RGB
LTMU_B	0.680	0.581	0.626	LT	RGB
Megtrack	0.694	0.551	0.614	LT	RGB
RPT	0.601	0.546	0.572	ST	RGB
Siam_LTD	0.626	0.489	0.549	LT	RGBD
OceanPlus	0.577	0.502	0.537	ST	RGB
AlphaRef	0.491	0.547	0.518	ST	RGB

Table 6. List of trackers that participated in the VOT-RGBD2020 challenge along with their performance scores (Pr, Re, F-score) and categorizations (ST/LT, RGB/RGBD). 2020 submissions are ATCAIS, DDiMP, CLGS_D and Siam_LTD. SiamDW_D and LTDSEd are 2019 submissions (SiamDW_D was the winner). RGB trackers are the three top performers of VOT-ST2020 and VOT-LT2020.