FMODetect: Robust Detection of Fast Moving Objects

Denys Rozumnyi[1,4], Jiří Matas[4], Filip Šroubek[2], Marc Pollefeys[1,3], Martin R. Oswald[1]
[1]Department of Computer Science, ETH Zurich [3]Microsoft Mixed Reality and AI Zurich Lab
[2]UTIA, Czech Academy of Sciences [4]Visual Recognition Group, Czech Technical University in Prague

2021 ICCV OCTOBER 11-17 VIRTUAL

Introduction:

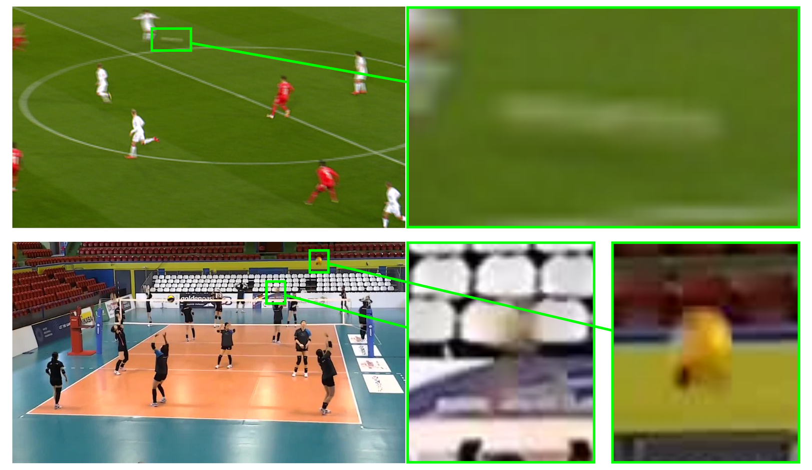
Definition: Fast Moving Objects (FMOs) – objects that move over a distance larger than their size within the camera exposure time.

Inputs:
- image *I* with potentially several objects moving fast and appearing blurred,
- background *B* without the object.

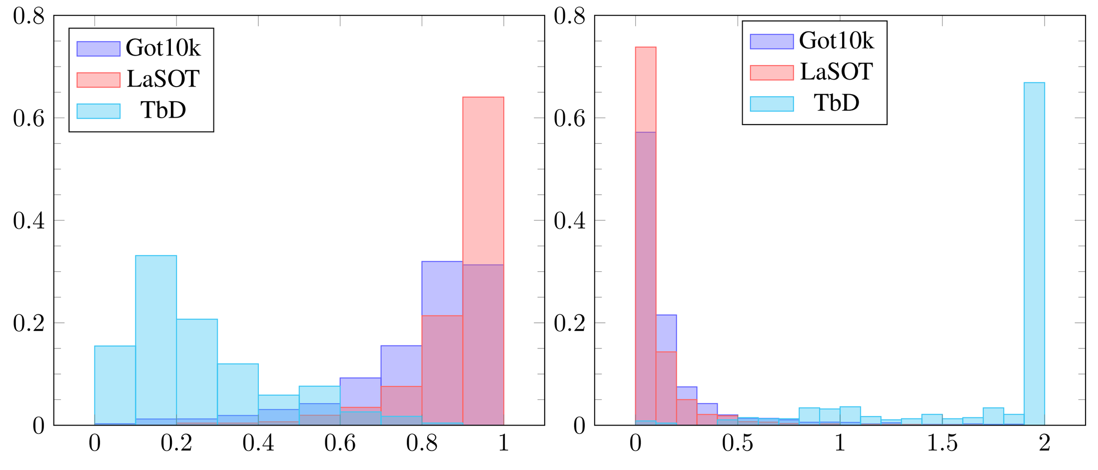
Task: detect all FMOs, estimate their trajectories (*H*), deblur their appearance (*F*,*M*). **Model**: $I = H * F + (1 - H * M)\,B$

Contribution: the first learned approach for FMO detection and deblurring.

Detection examples:



The problem of detecting and tracking FMOs has been unnoticed by the research community, and such objects are not present in standard tracking datasets:

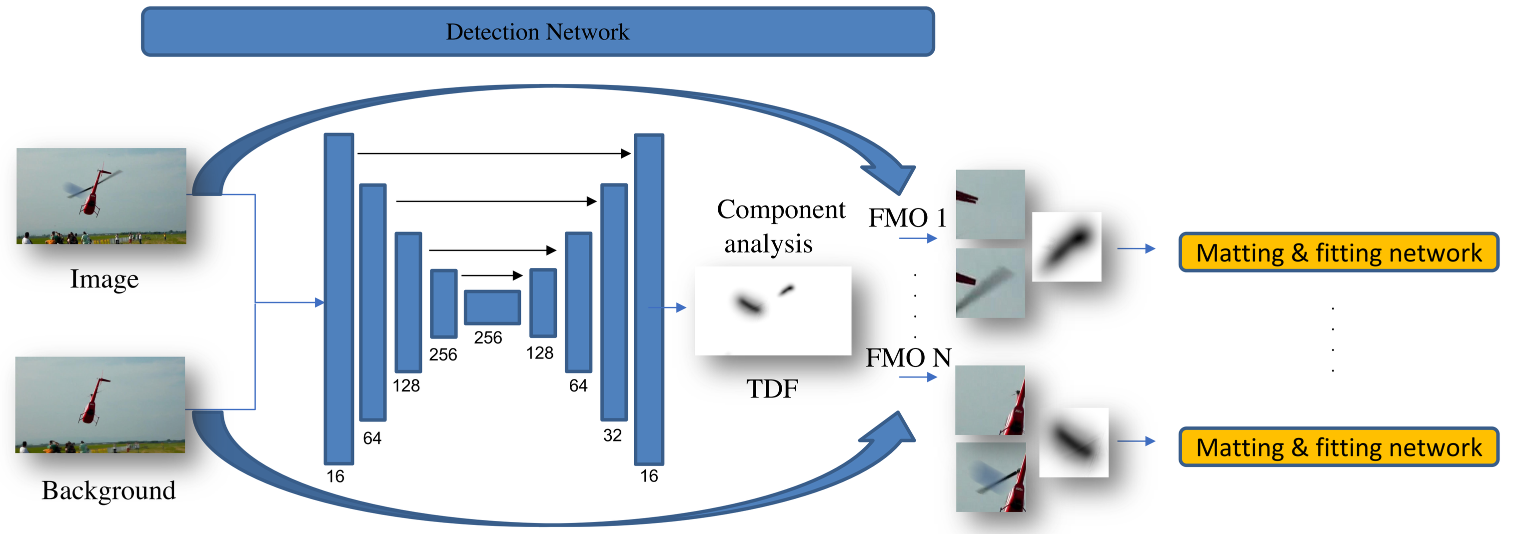


(a) IoU(GT_t,GT_{t+1}) (b) Normalized speed

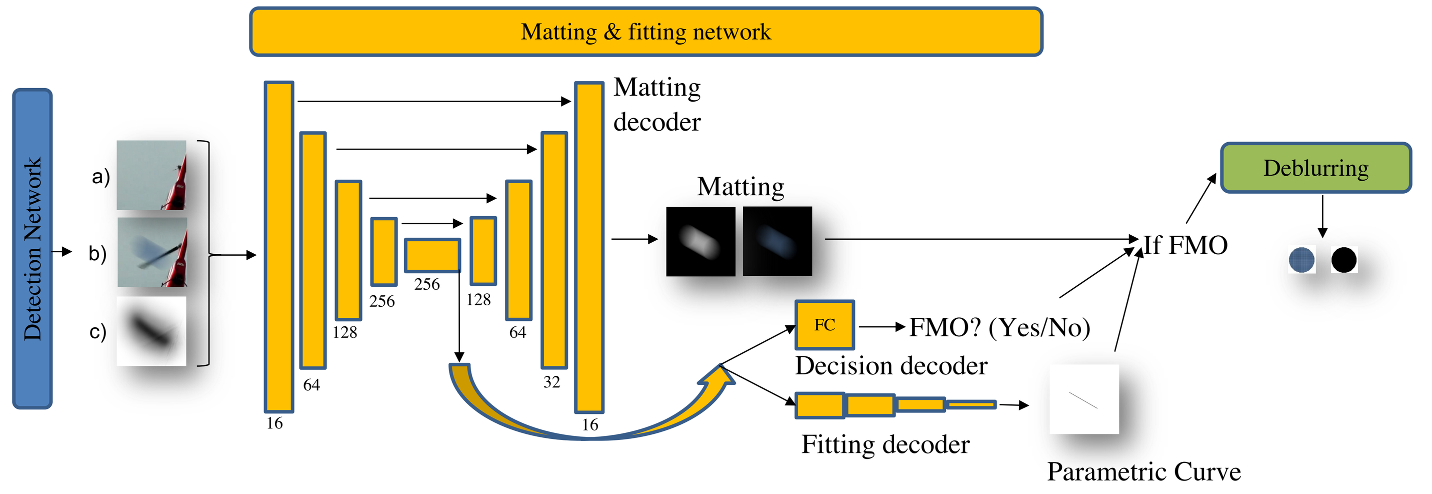
FMO retrieval on YouTube videos by FMODetect in real-time:

Method:

Step one (detection): the network detects all FMOs in the scene using a truncated distance field (TDF) to object trajectory.



Step two (matting): solve the matting problem that separates the background from the input image.

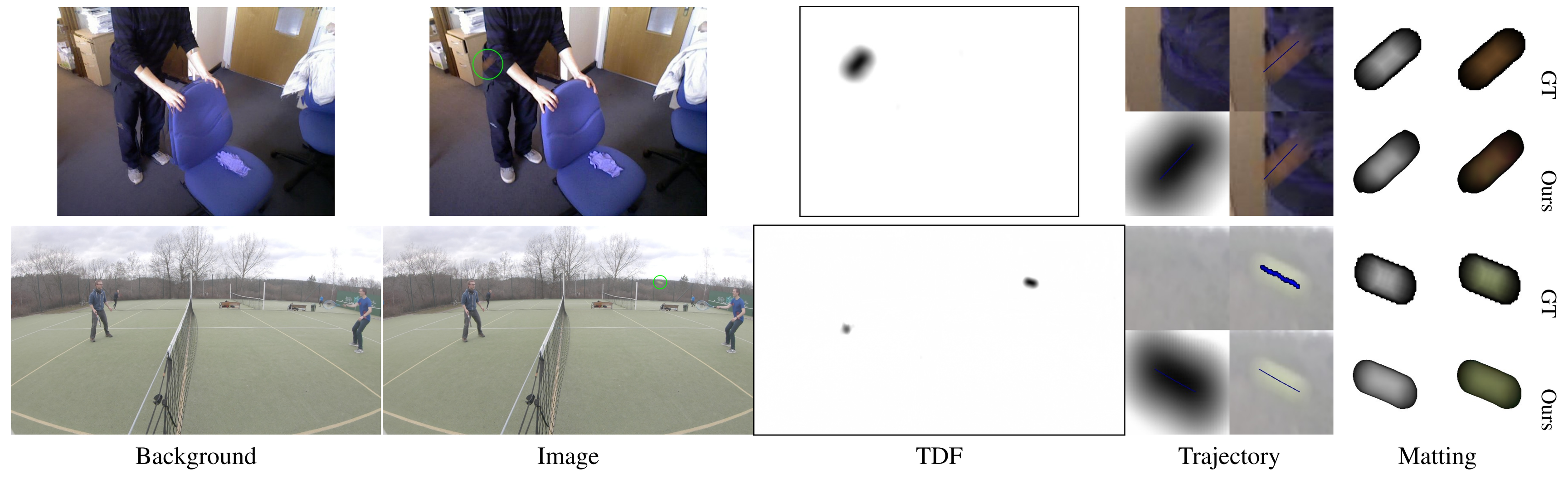


- Trained on a new synthetic dataset.
- Generalizes well to unseen and more difficult real-world data.

Step three (deblurring): sharp appearance (*F*,*M*) estimation. We propose a novel energy minimization-based deblurring.

$$\min_{F,M,H} \frac{1}{2}\left(\|H * F - \hat{H}_F\|_2^2 + \|H * M - \hat{H}_M\|_2^2\right) + \alpha_F \|\nabla F\|_1 + \alpha_M \|\nabla M\|_1$$

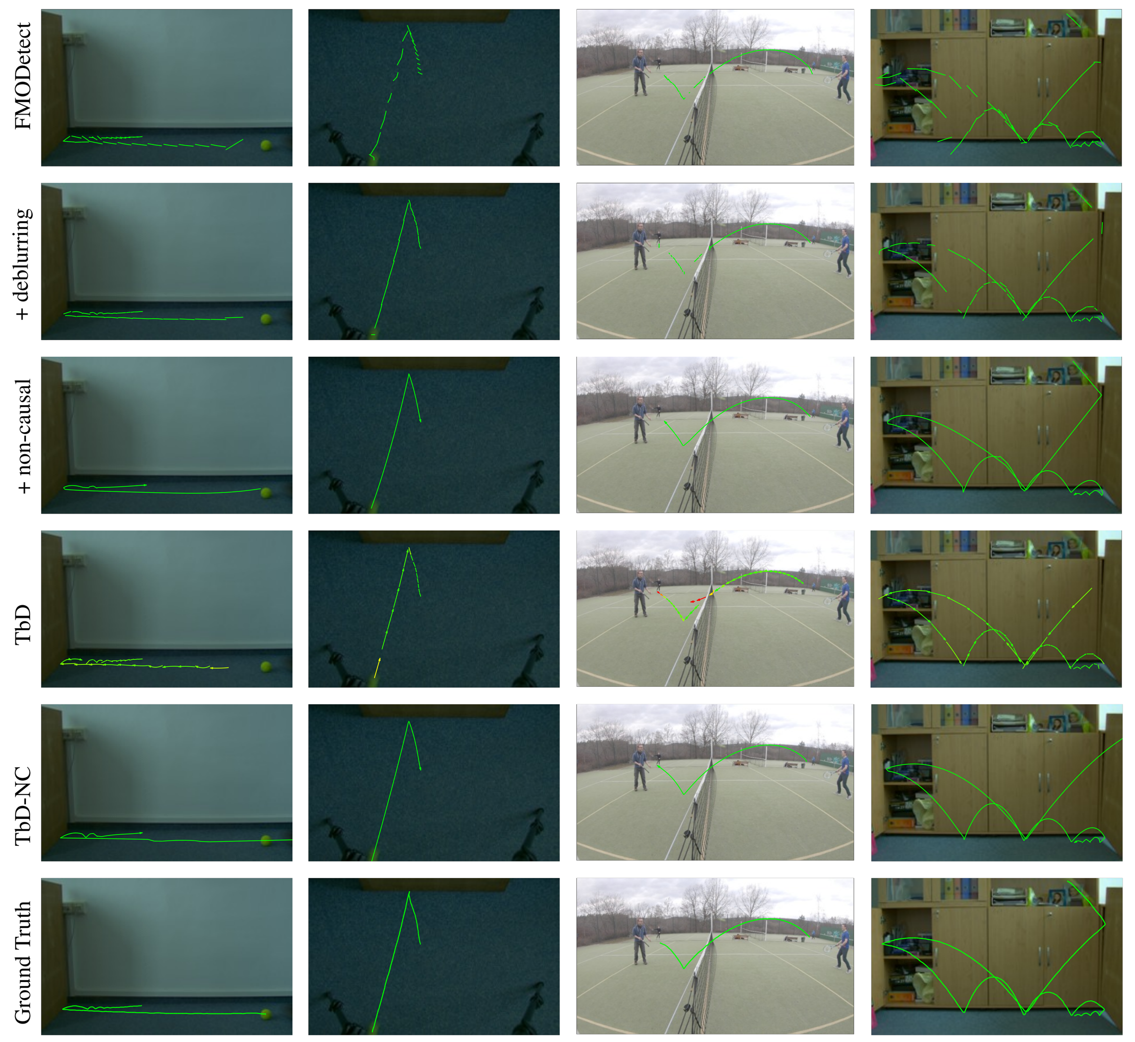
FMODetect steps one and two:



- Trajectories are overlaid on the input image.
- False positives in TDF are usually rejected by the decision decoder.

Experiments:

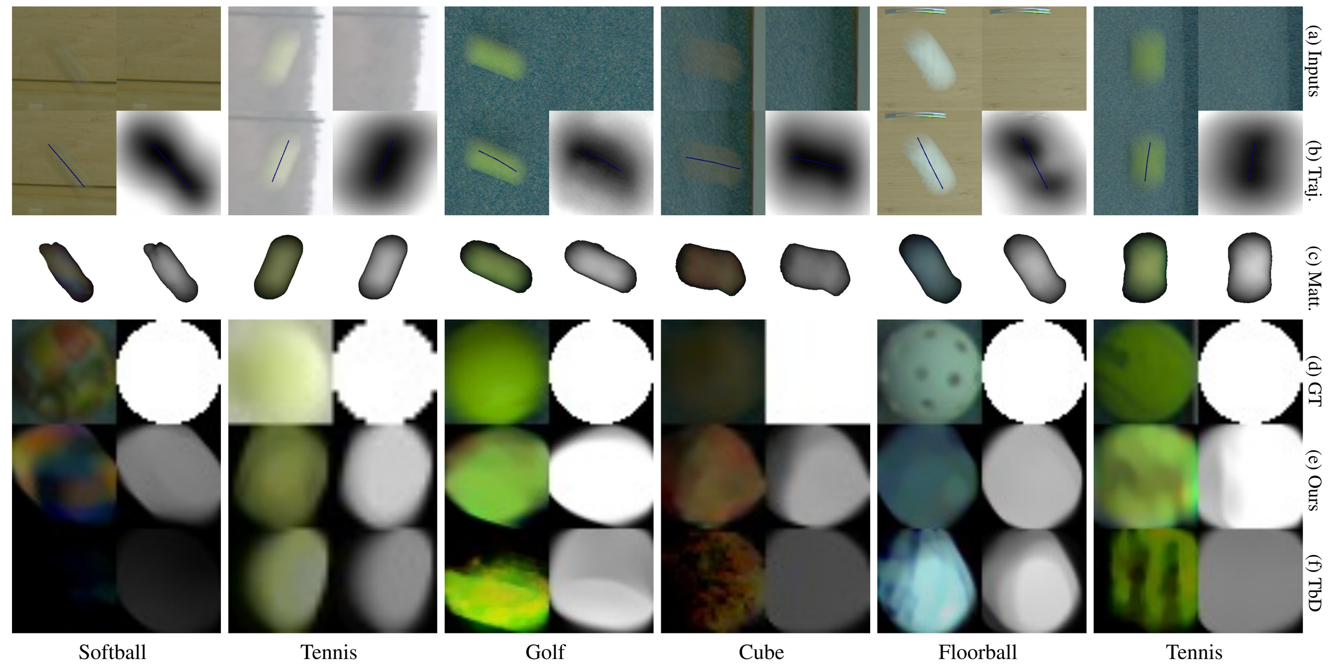
Trajectory estimation, compared to previous methods such as [TbD] and [TbD-NC] and the ground truth from the high-speed camera:



The proposed detection and trajectory estimation method is much faster (runs in real-time) and outperforms the previous methods:

	Causal processing				Non-causal processing	
	FMO	TbD	(a) FMODetect	(b) + deblurring	(c) + NC	TbD-NC
Recall↑	0.56	0.96	**0.97**	**0.97**	0.99	0.99
TIoU↑	0.352	0.713	0.519	**0.715**	0.781	0.779
Runtime [1/s]↑	1 fps	0.2 fps	**20 fps**	0.4 fps	N/A	N/A

In terms of deblurring quality, the reconstructed object is sharper:



Conclusion:

- The first **learning-based**, **real-time,** approach for FMO detection.
- Compared to the previous methods, FMODetect is **simpler**, does not require extensive **parameter tuning**, and works with the **same settings** for a wide range of scenarios.
- Code is on GitHub: https://github.com/rozumden/FMODetect

References:
[TbD] Kotera et al. Intra-frame Object Tracking by Deblatting, ICCV VOT 2019
[TbD-NC] Rozumnyi et al. Non-Causal Tracking by Deblatting, GCPR 2019
[FMO] Rozumnyi et al. The World of Fast Moving Objects, CVPR 2017