The World of Fast Moving Objects

Denys Rozumnyi[1,3] Jan Kotera[2] Filip Šroubek[2] Lukáš Novotný[1] Jiří Matas[1]

[1]CMP, Czech Technical University in Prague [2]UTIA, Czech Academy of Sciences [3]SGN, Tampere University of Technology

IEEE 2017 Conference on Computer Vision and Pattern Recognition

CVPR July 21-26 HONOLULU 2017

Introduction

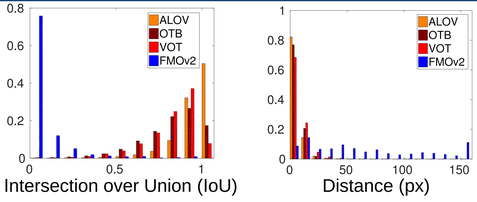
- **Definition (Fast Moving Object, FMO)**: object is fast moving if its trajectory projected on the image plane is larger than its size
  Note: FMO is a property of the relative motion of object w.r.t. the camera
- Image formation model: $I = (1 - P * M)B + P * F$ ... convolution

I – **Image** | **Alpha matte** | B – **Background** | P – **Path** | F – **FMO**

- Difference image: $\Delta = I - B = P * F - (P * M)B$
- Multiple FMOs: $\Delta = \sum P^i * F^i - (\sum P^i * M^i)B$
- **Goal:** given I, recover P^i (and F^i)

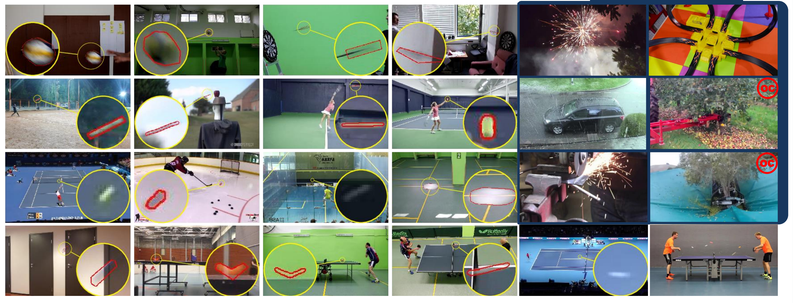
Analysis of Standard Tracking Sets: VOT [1], OTB [2], ALOV [3]

→ No FMOs present
→ Different tracking "worlds"



[1] Kristan et al. "The visual object tracking VOT 2015 challenge results", ICCV 2015
[2] Wu et al. "Online object tracking: A benchmark", CPVR 2013
[3] Smeulders et al. "Visual tracking: An experimental survey", PAMI 2014
[4] http://cmp.felk.cvut.cz/fmo/

FMO Annotated Public Dataset [4] and Examples

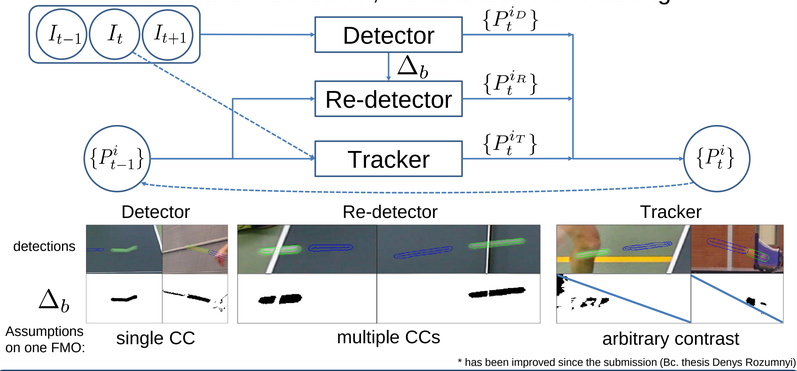


FMO Localization Algorithms

- Detection algorithm*

1. Input images → background (estimated by median)
2. $\Delta_b = |I - B| > \theta$ (threshold set adaptively)
3. Stroky connected components
4. Not a lateral motion (checked by KLT tracker)
5. Not a shadow → FMOs (check of gradient fields)

- Full FMO localization: Detection, Re-detection and Tracking



* has been improved since the submission (Bc. thesis Denys Rozumnyi)

Evaluation

- Criteria: precision, recall, F-score
- SOTA trackers perform poorly

F-score (in %) of the proposed method and SOTA

Sequence name	ASMS	DSST	MEEM	SRDCF	STRUCK	FMO
Average	17	1	1	1	3	32
volleyball	80	0	50	0	10	40
volleyball passing	12	6	95	88	8	0
darts	3	0	6	0	0	0
darts window	0	0	0	0	0	0
softball	0	0	0	0	0	39
archery	5	5	5	5	0	3
tennis serve side	7	0	0	0	6	56
tennis serve back	5	0	0	0	3	41
tennis court	0	0	3	3	0	41
hockey	0	0	0	0	0	7
squash	0	0	0	0	0	21
frisbee	65	0	6	6	0	75
blue ball	30	0	0	0	25	0
ping pong tampere	0	0	0	0	0	67
ping pong side	1	0	0	0	0	46
ping pong top	0	0	0	0	1	74

- Example detections:

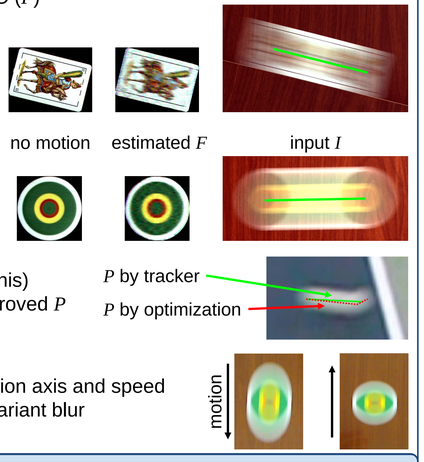
in a sequence | in a single frame

Appearance Reconstruction

- Minimization with priors on FMO appearance (F) and path (P)

$$\min_{P,F} \|I - (1 - P * M)B - P * F\|_1 + \text{prior}(F) + \text{prior}(P)$$

- **More complex than blind deconvolution**:
  - mixing with background (B)
  - blur (P) larger than the size of the FMO (F)
- F arbitrary (~1min)
  - P from tracker assumed correct
  - prior(F): standard total variation
- F simplified (~1s)
  - estimation of F and improved P
  - prior(P): polyline
  - prior(F): circular symmetry (demo)
- prior(F): single-color sphere (table tennis)
- estimation of F (color, radius) and improved P
- Beyond convolution
  - rotating FMO → grid search over rotation axis and speed
  - rolling shutter → complicated space-variant blur



Applications

- Temporal super-resolution
  25 fps vs 250 fps
- Angular velocity
- Object counting
- Projected translational velocity vs radar