Microsoft Research - Inria
JOINT CENTRE

INVENTEURS DU MONDE NUMÉRIQUE

THOTH

UNIVERSITÉ
Grenoble
Alpes

LABORATOIRE
JEAN KUNTZMANN
MATHÉMATIQUES APPLIQUÉES - INFORMATIQUE

Towards Robust Visual Object Tracking: Proposal Selection & Occlusion Reasoning

Yang HUA

PhD Defense, June 10 2016

Rapporteur	Dr. Patrick PÉREZ
Rapporteur	Pr. Deva RAMANAN
Examinateur	Pr. Jiří MATAS
Examinateur	Dr. Florent PERRONNIN
Directeur de these	Dr. Cordelia SCHMID
Co-encadrant de these	Dr. Karteek ALAHARI

Motivation

- Emerging technologies which require robust object tracking
  - Microsoft HoloLens

Image courtesy of Microsoft (https://www.microsoft.com/microsoft-hololens/en-us)

Motivation

- Emerging technologies which require robust object tracking
  - Microsoft HoloLens

Video courtesy of Microsoft (https://www.microsoft.com/microsoft-hololens/en-us)

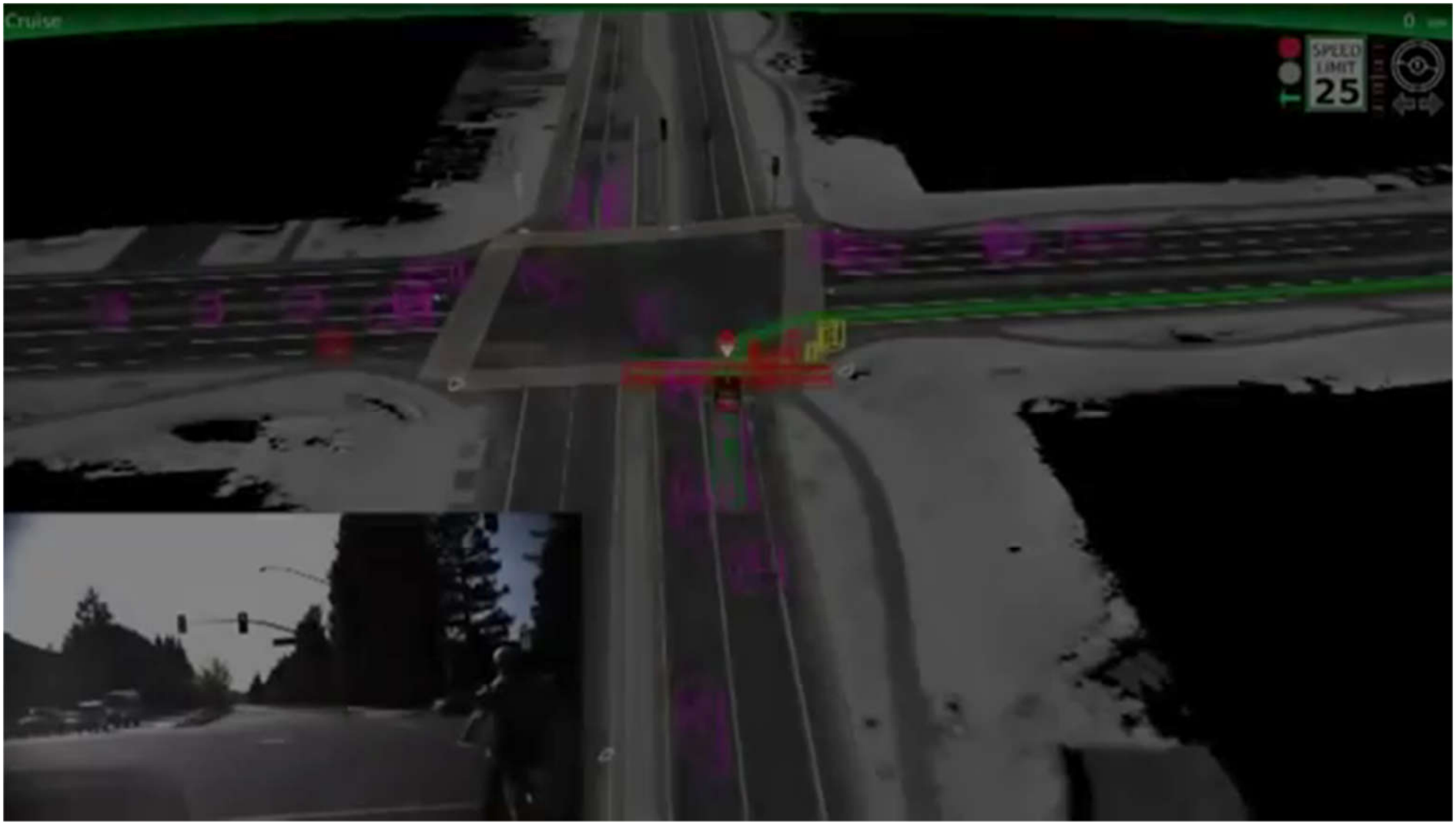
Motivation

- Emerging technologies which require robust object tracking
  - Google Self-Driving Car

Image courtesy of Google (https://www.google.com/selfdrivingcar/)

Motivation

- Emerging technologies which require robust object tracking
  - Google Self-Driving Car



Video courtesy of Google (https://www.google.com/selfdrivingcar/)

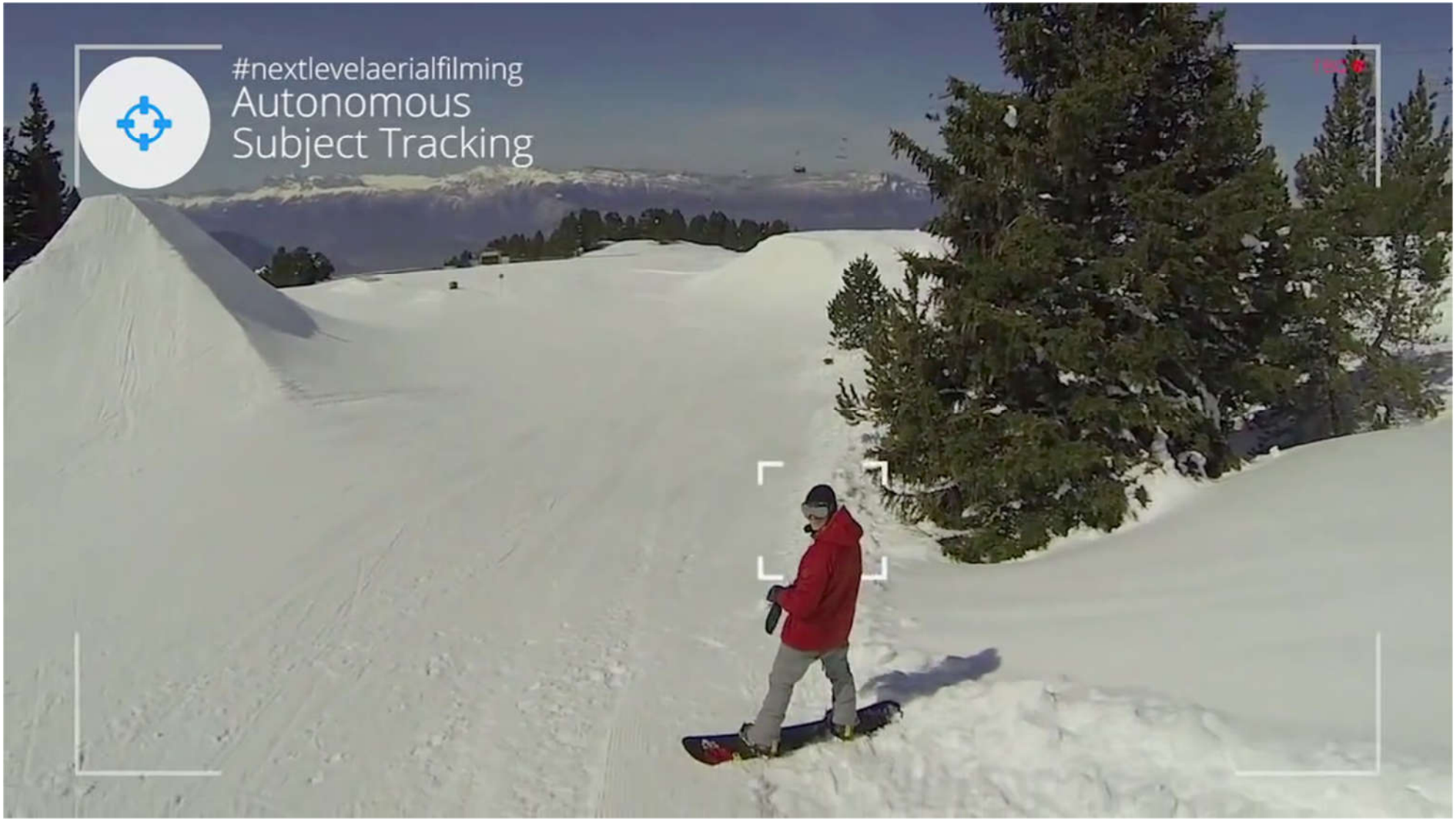
Motivation

- Emerging technologies which require robust object tracking
  - HEXO+ Drone

Image courtesy of HEXO+ (https://hexoplus.com/)

Motivation

- Emerging technologies which require robust object tracking
  - HEXO+ Drone



Video courtesy of HEXO+ (https://hexoplus.com/)

Visual Object Tracking

- "… the problem of estimating the **trajectory** of an object in the image plane **as it moves around a scene**" [Yilmaz'06]
- Setup: initial object position is defined at the beginning of a video

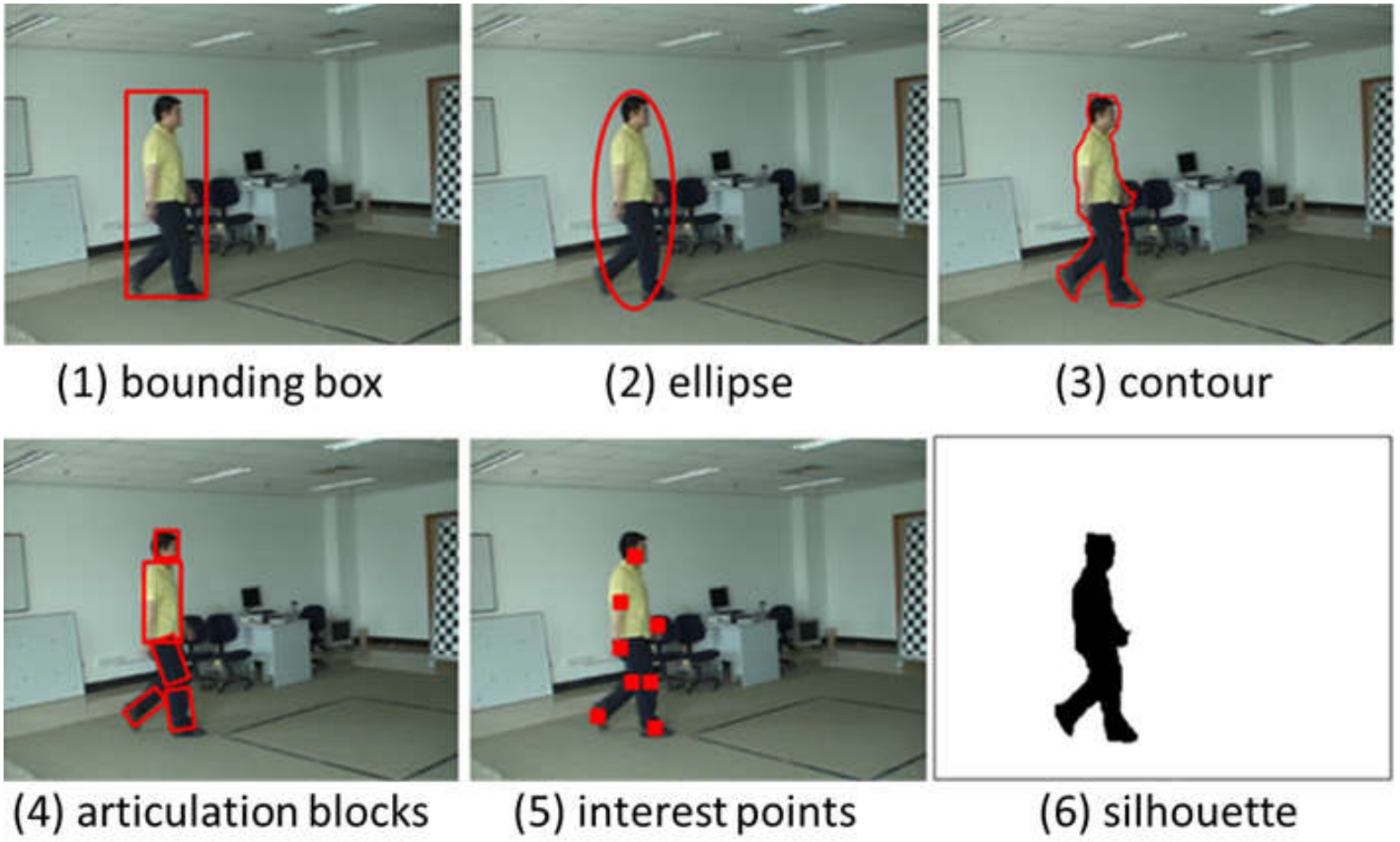


Image courtesy of [Li'13]

Visual Object Tracking

- "… the problem of estimating the **trajectory** of an object in the image plane **as it moves around a scene**" [Yilmaz'06]
- Setup: initial object position is defined at the beginning of a video

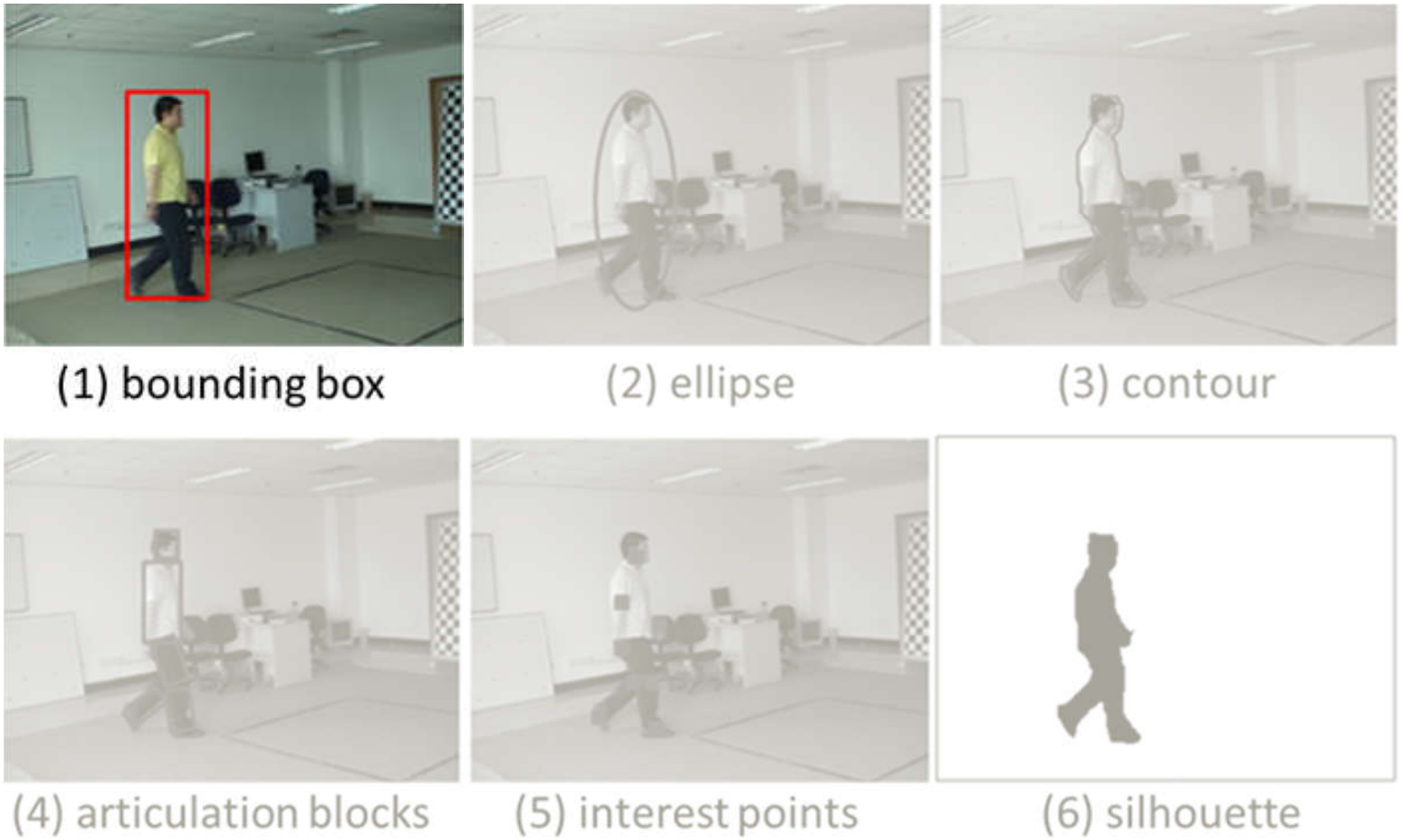


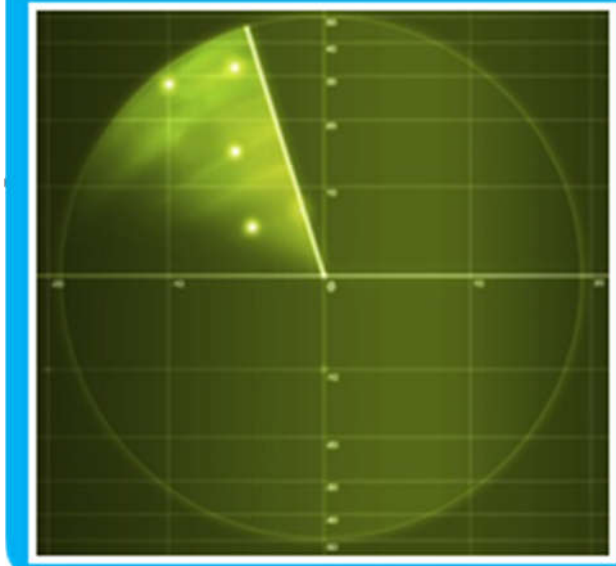
Image courtesy of [Li'13]

Visual Object Tracking: A brief history

1950s

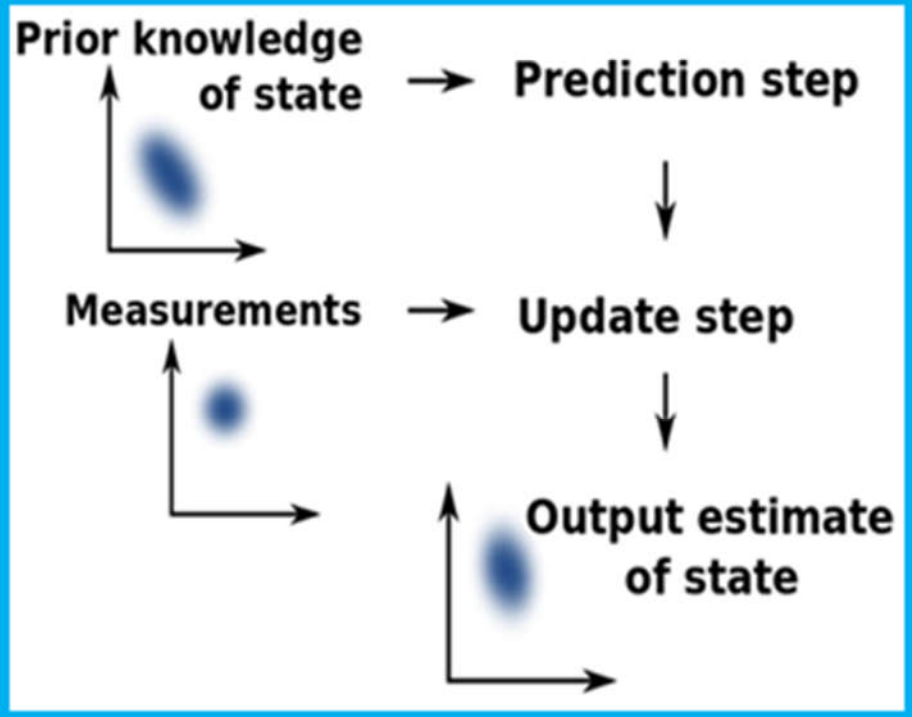
RADAR

[Wax, 1955]



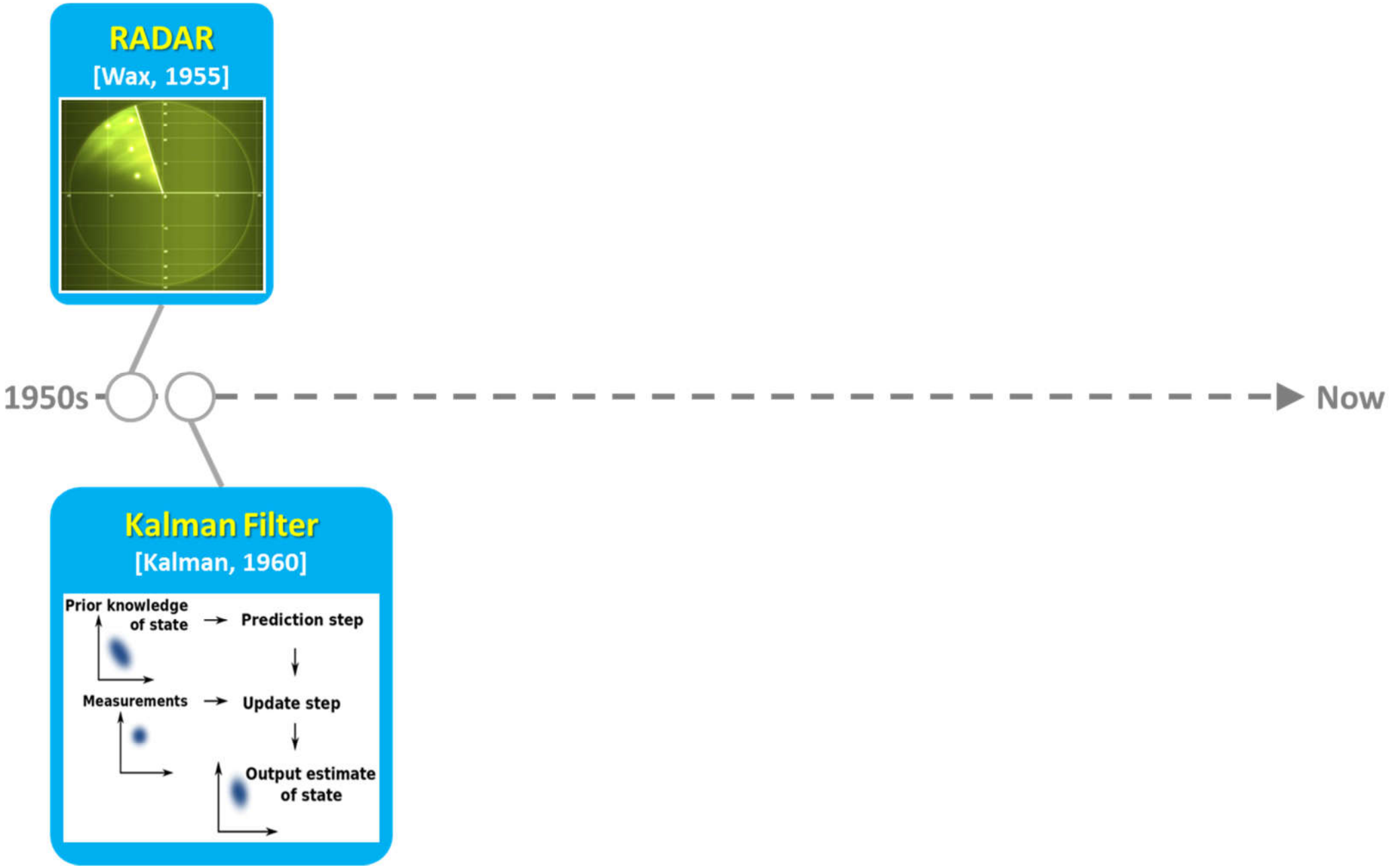
Kalman Filter

[Kalman, 1960]



Now

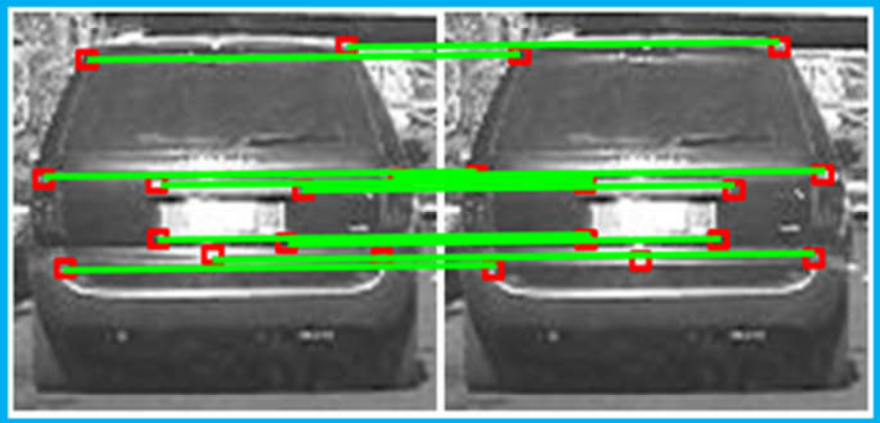
Visual Object Tracking: A brief history



Visual Object Tracking: A brief history

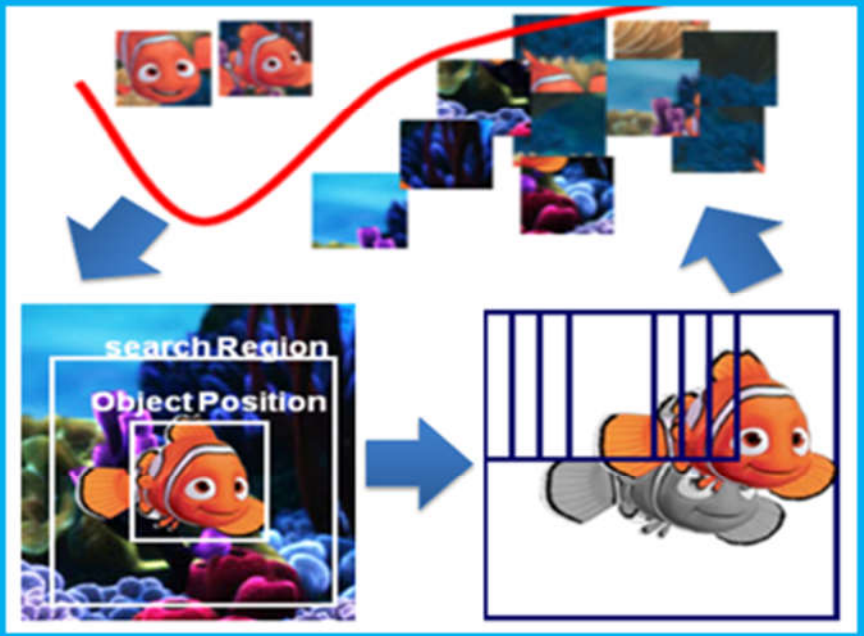
Tracking-By-Detection Template matching

[Lucas and Kanade, 1981, Tomasi and Kanade, 1991]



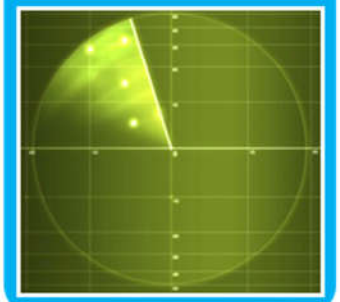
Tracking-By-Detection Discriminative modeling

[Avidan, 2004, Grabner et al., 2006, Hare et al., 2011]



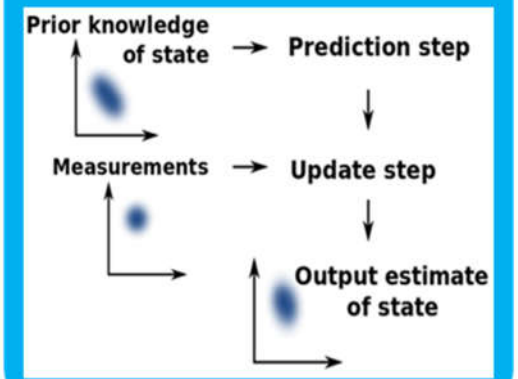
Visual Object Tracking: A brief history

RADAR
[Wax, 1955]

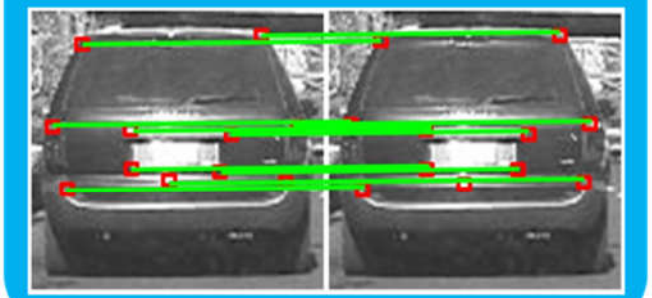


1950s ---------- Now

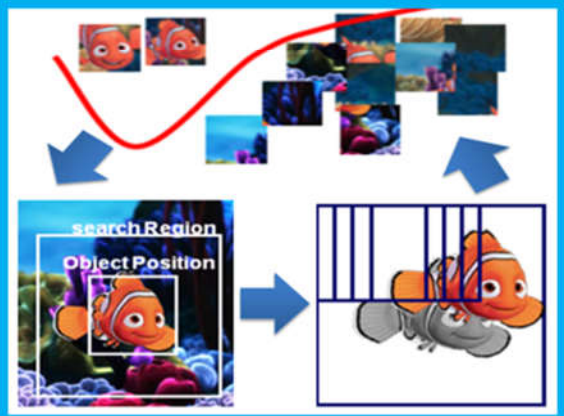
Kalman Filter
[Kalman, 1960]



Tracking-By-Detection
Template matching
[Lucas and Kanade, 1981, Tomasi and Kanade, 1991]



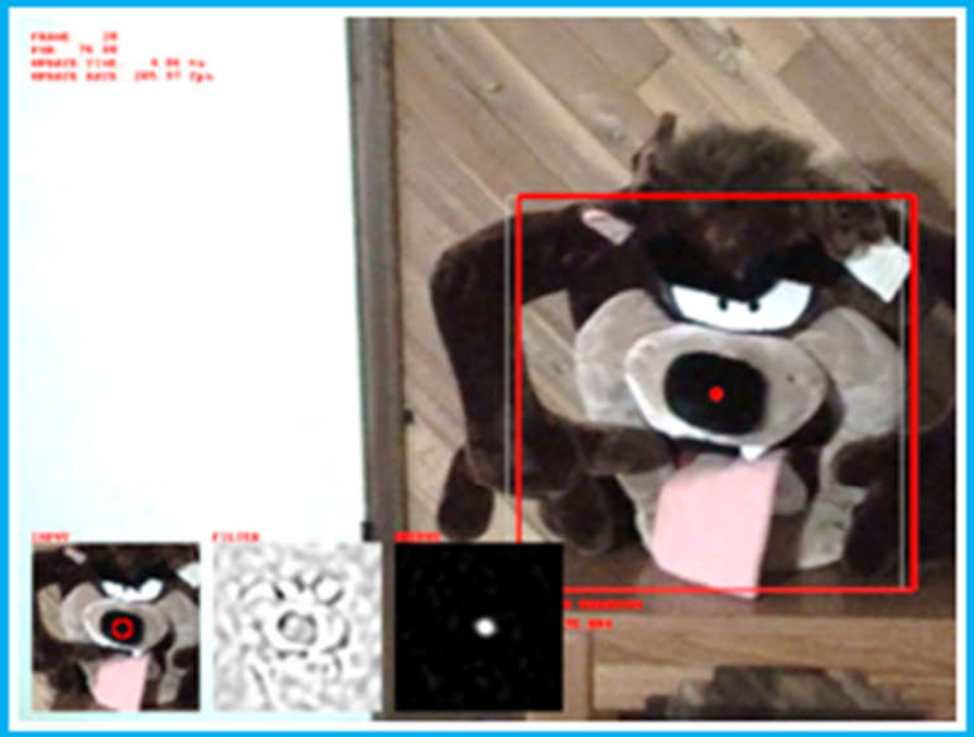
Tracking-By-Detection
Discriminative modeling
[Avidan, 2004, Grabner et al., 2006, Hare et al., 2011]



Visual Object Tracking: A brief history

Correlation Filters

[Bolme et al., 2010,
Henriques et al., 2015]



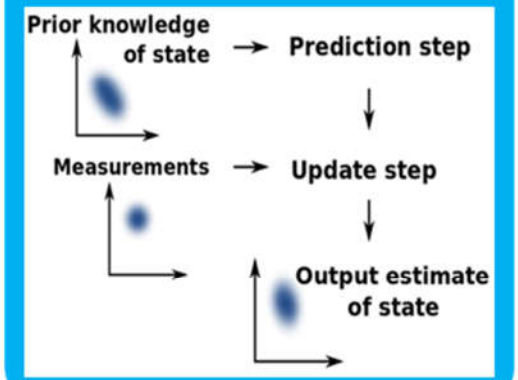
Visual Object Tracking: A brief history

RADAR
[Wax, 1955]

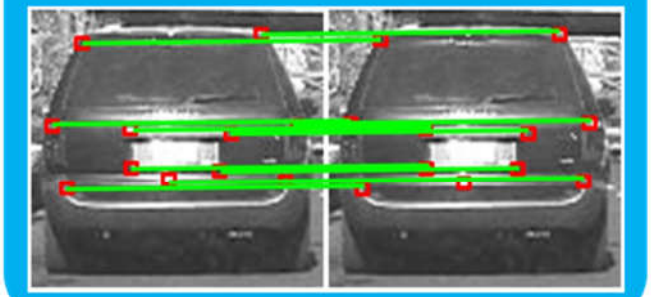
Correlation Filters
[Bolme et al., 2010, Henriques et al., 2015]

1950s ——— Now

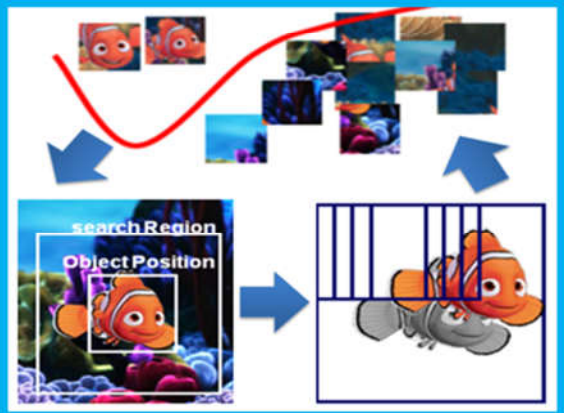
Kalman Filter
[Kalman, 1960]



Tracking-By-Detection Template matching
[Lucas and Kanade, 1981, Tomasi and Kanade, 1991]



Tracking-By-Detection Discriminative modeling
[Avidan, 2004, Grabner et al., 2006, Hare et al., 2011]



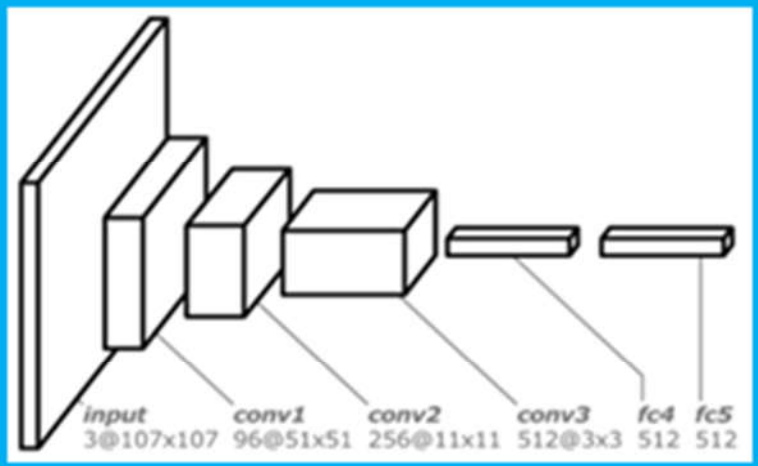
Visual Object Tracking: A brief history

Tracking Dataset & Challenge

[Wu et al., 2013, Kristan et al., 2013]

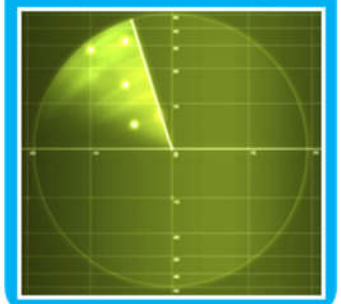
Deep Learning for Tracking

[Nam & Han, 2016]



Visual Object Tracking: A brief history

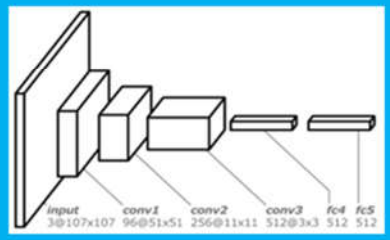
RADAR
[Wax, 1955]



Correlation Filters
[Bolme et al., 2010, Henriques et al., 2015]

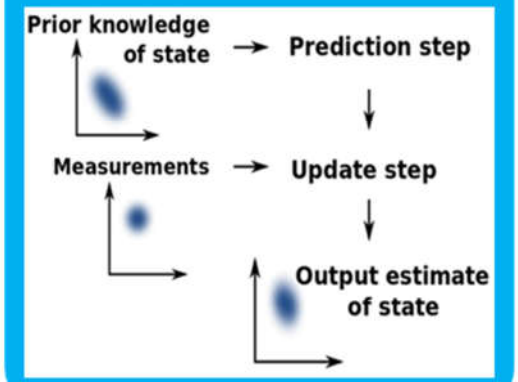
Tracking Dataset & Challenge
[Wu et al., 2013, Kristan et al., 2013]

Deep Learning for Tracking
[Nam & Han, 2016]

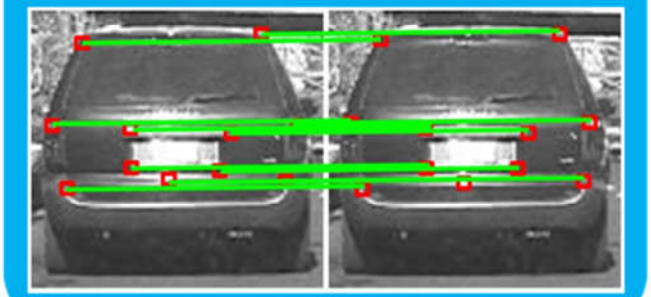


1950s ▶ Now

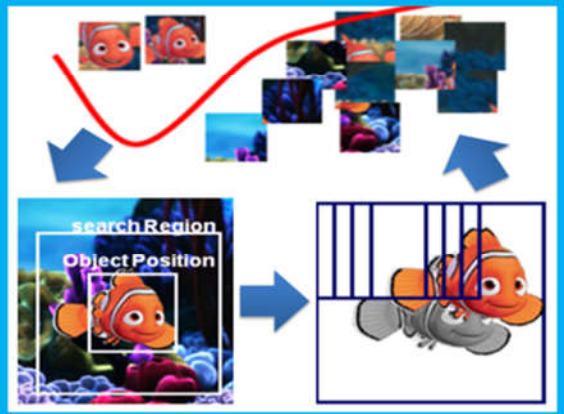
Kalman Filter
[Kalman, 1960]



Tracking-By-Detection Template matching
[Lucas and Kanade, 1981, Tomasi and Kanade, 1991]



Tracking-By-Detection Discriminative modeling
[Avidan, 2004, Grabner et al., 2006, Hare et al., 2011]



Visual Object Tracking: Applications

- Plays a fundamental role for high-level computer vision tasks
  - Video segmentation [Li'13]

Image courtesy of [Li'13]

Visual Object Tracking: Applications

- Plays a fundamental role for high-level computer vision tasks
  - Action localization [Weinzaepfel'15]

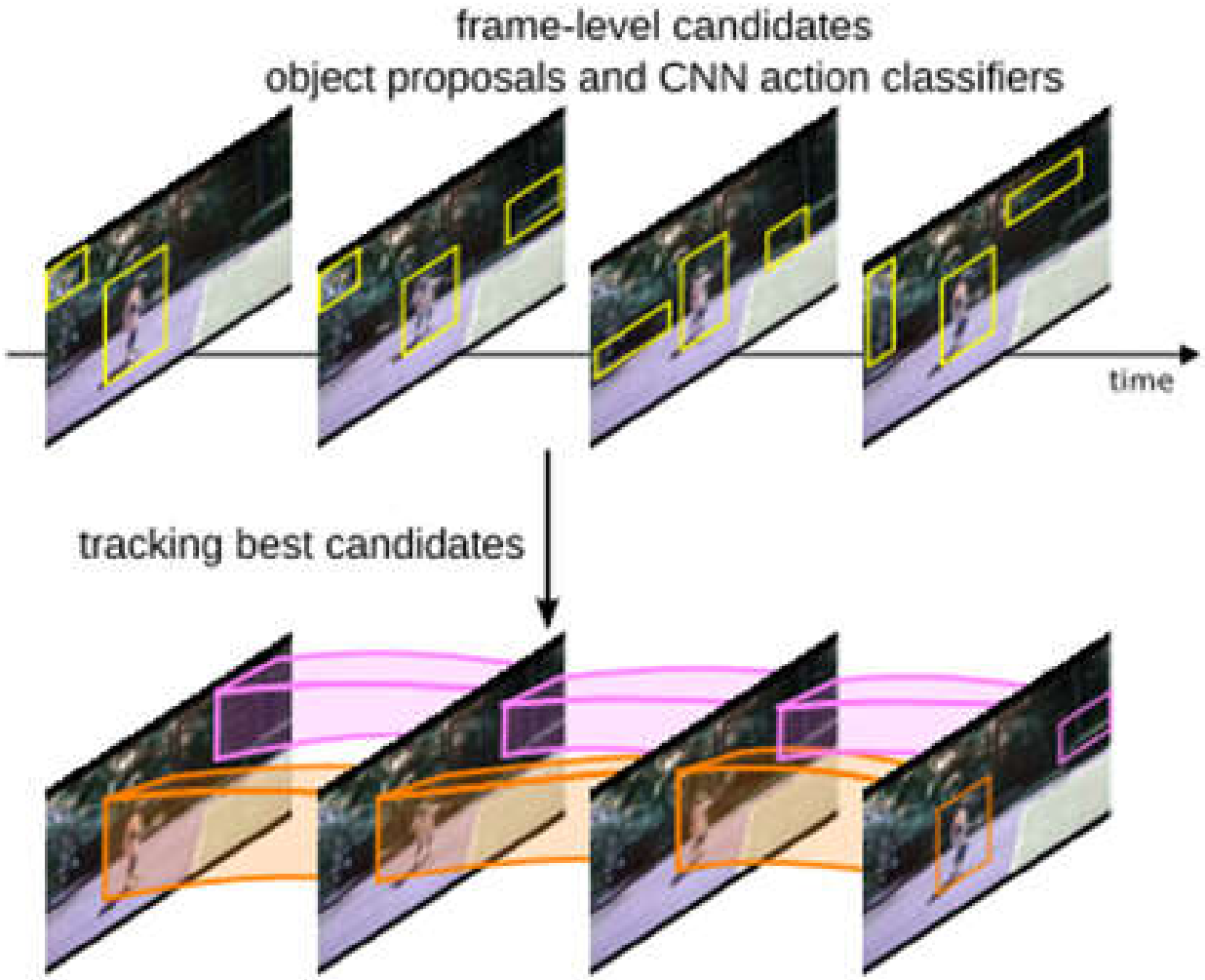
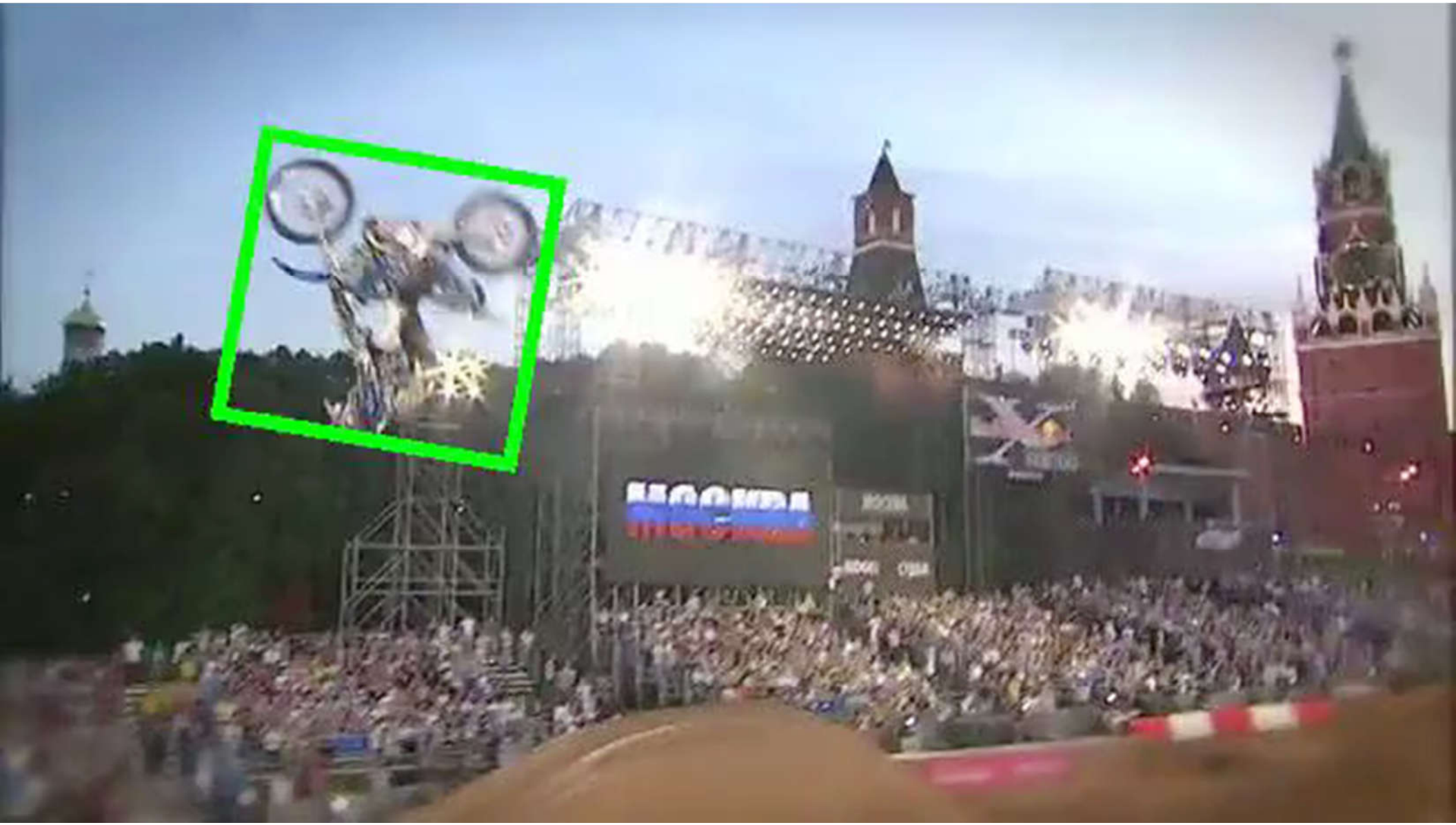
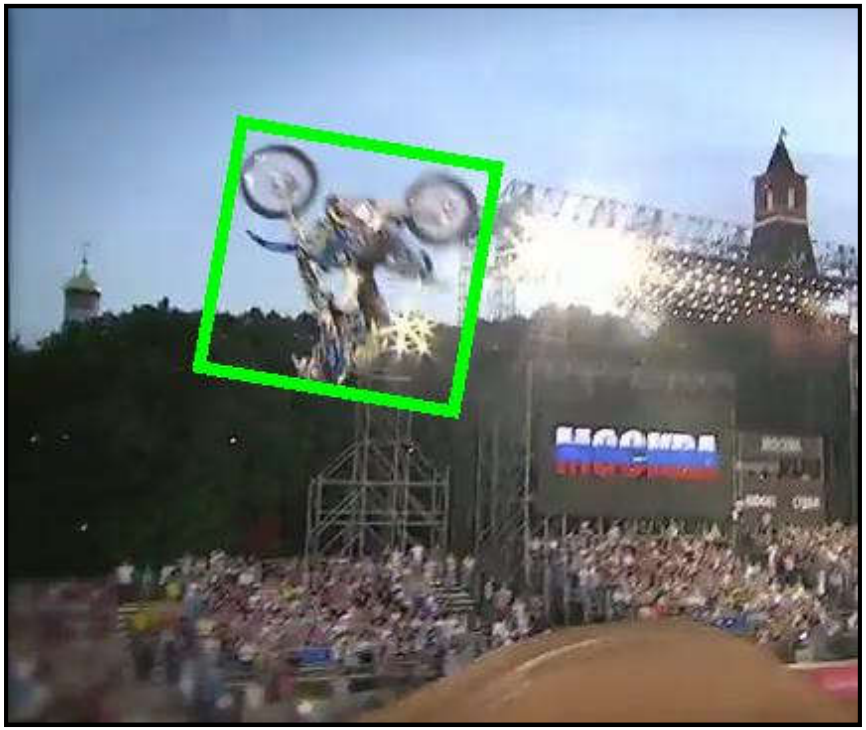


Image courtesy of [Weinzaepfel'15]

Challenge: Significant Transformations



Challenge: Significant Transformations







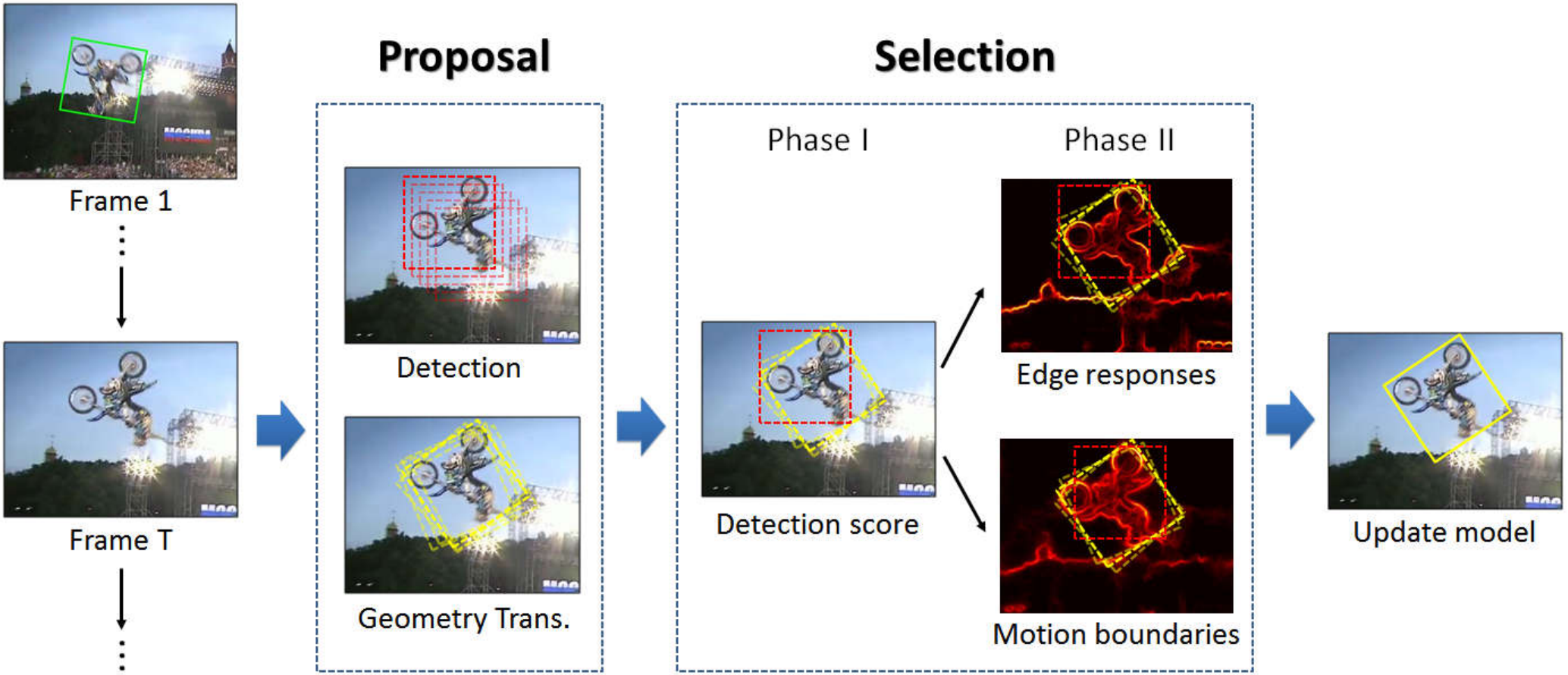



Challenge: Occlusion or Leaving/Re-entering the Field-of-View

Challenge: Occlusion or Leaving/Re-entering the Field-of-View

Contributions of this thesis

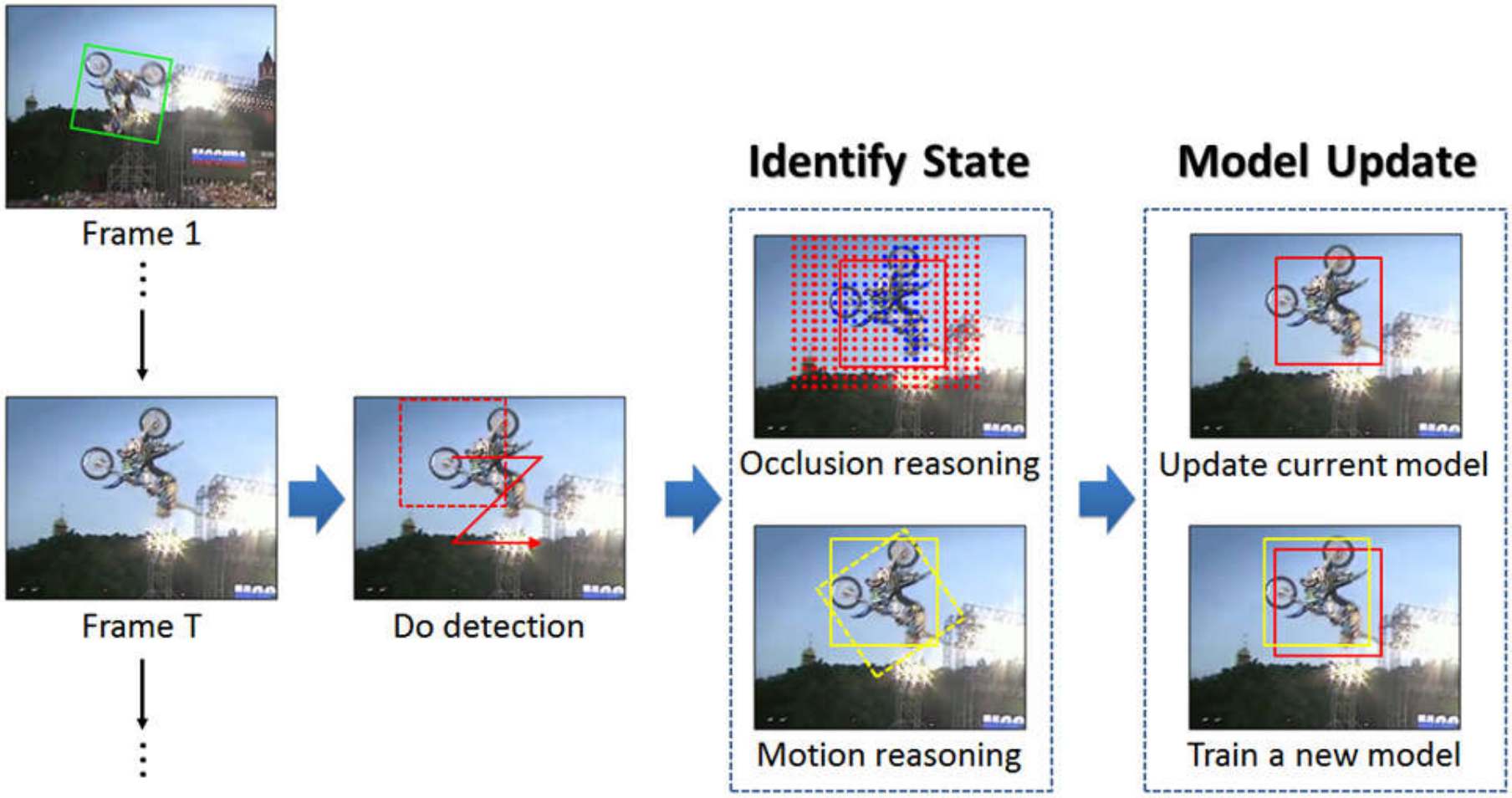
- Proposal and selection framework for short-term tracking



- Publication: Y. Hua, K. Alahari, and C. Schmid. Online object tracking with proposal selection. In ICCV, 2015
- Award: “Winning tracker” in the VOT-TIR2015 challenge

Contributions of this thesis

- Robust model update in the context of long-term tracking



- Publication: Y. Hua, K. Alahari, and C. Schmid. Occlusion and motion reasoning for long-term tracking. In ECCV, 2014

Outline

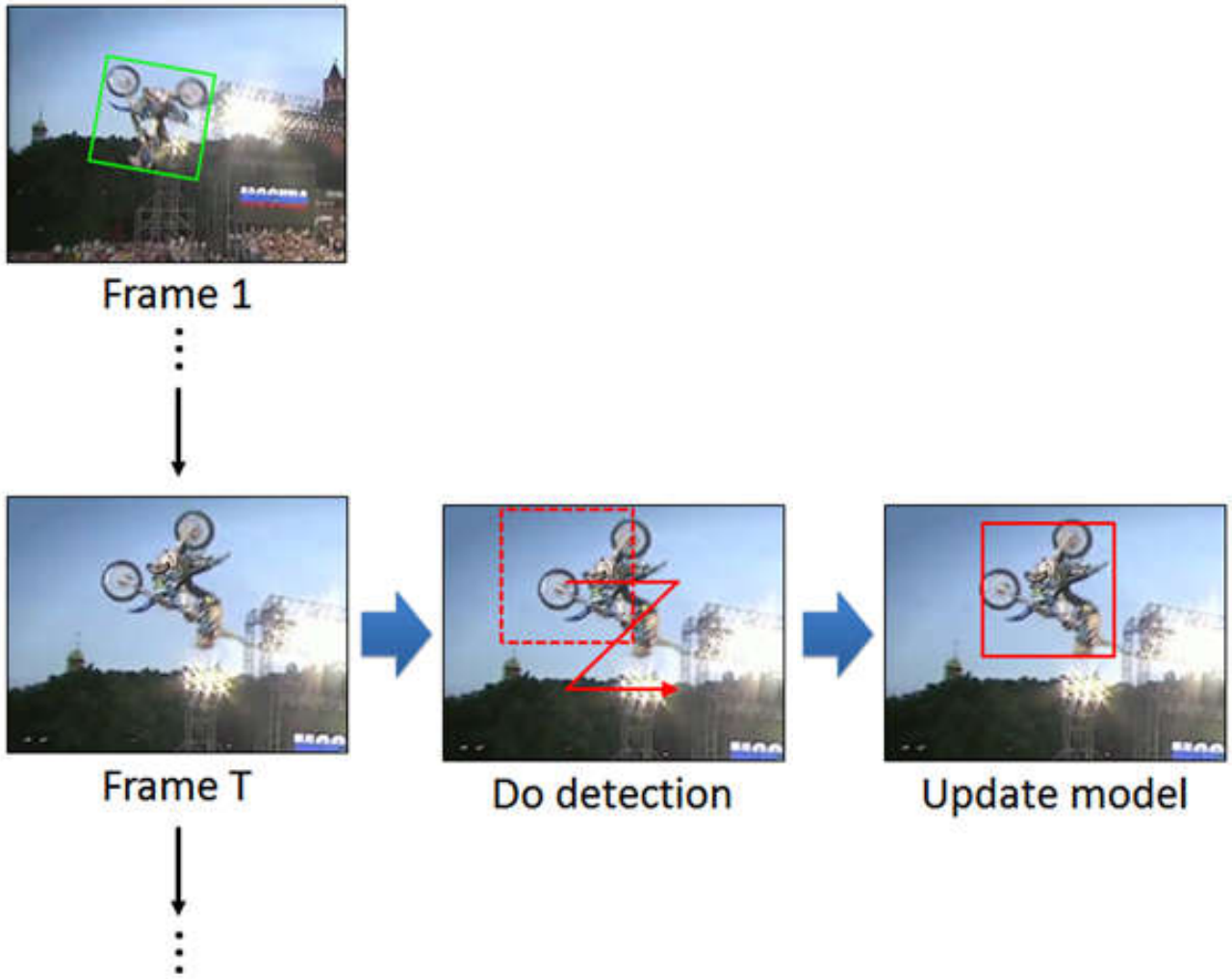
- ***Online Object Tracking with Proposal Selection***

- Occlusion and Motion Reasoning for Long-term Tracking

- Conclusion and Future Work

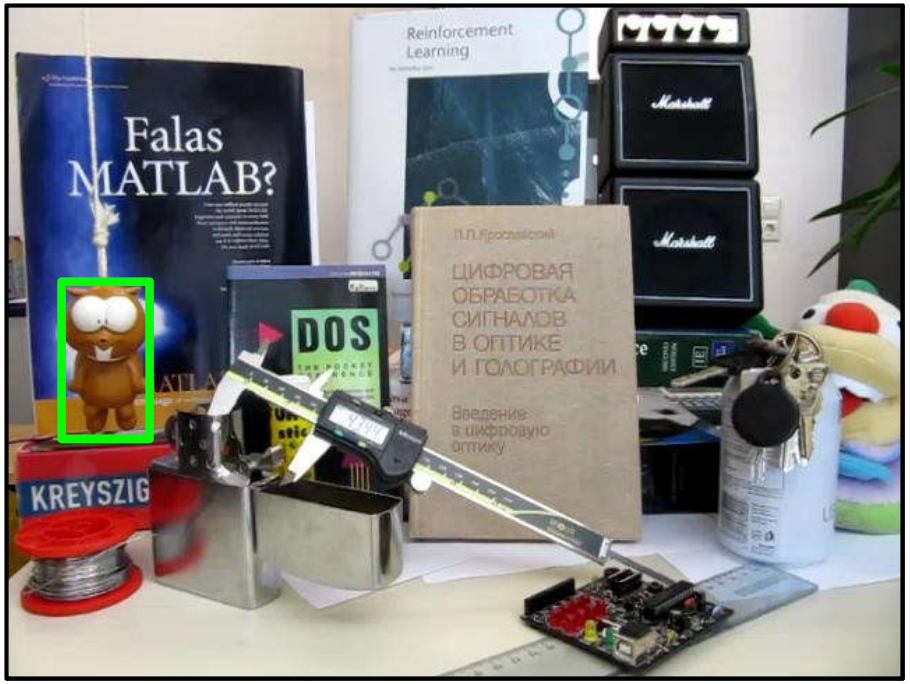
Tracking-By-Detection

- A successful approach on diverse benchmarks [Wu'13, Kristan'13/'14]
  - Structured Output Tracking with Kernels (Struck) [Hare'11]
  - Pixel based LUT Tracker (PLT) [Heng'12]
  - Discriminative Scale Space Tracker (DSST) [Danelljan'14]

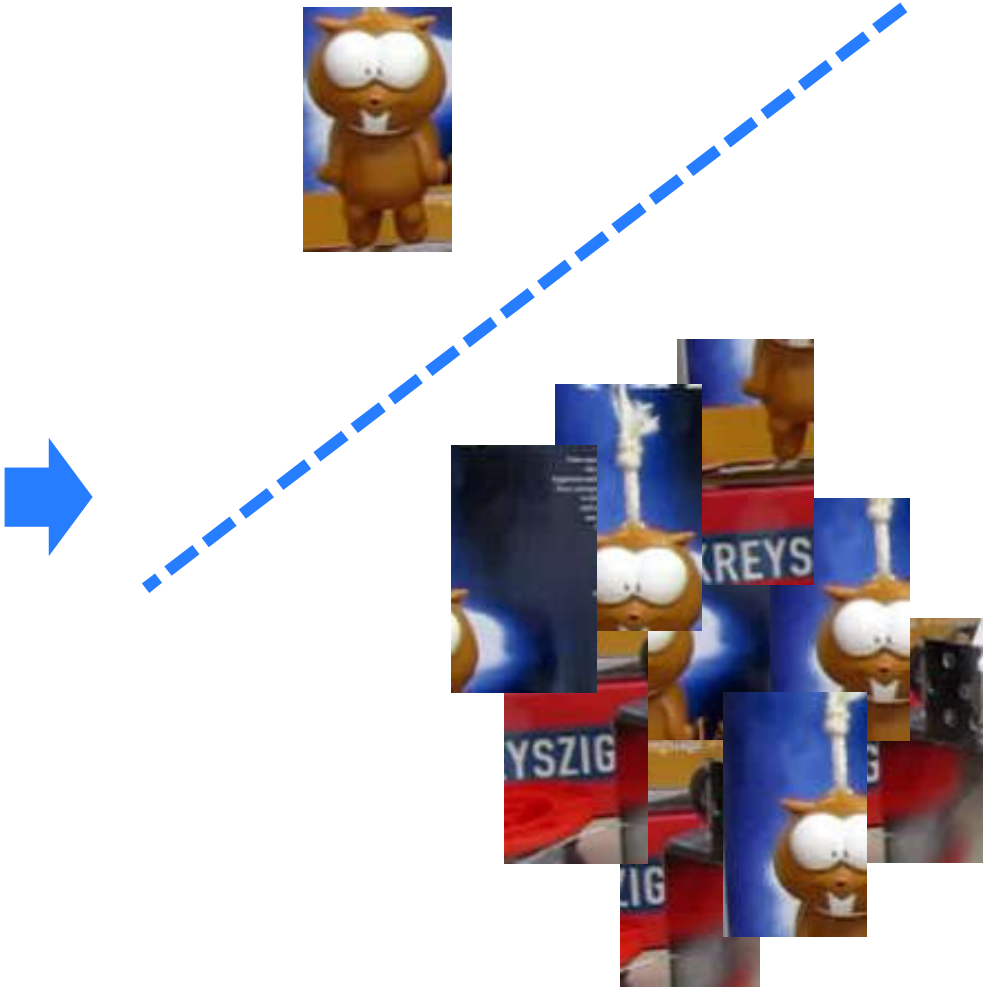


Tracking-By-Detection

- Two key points
  - Discriminative learning

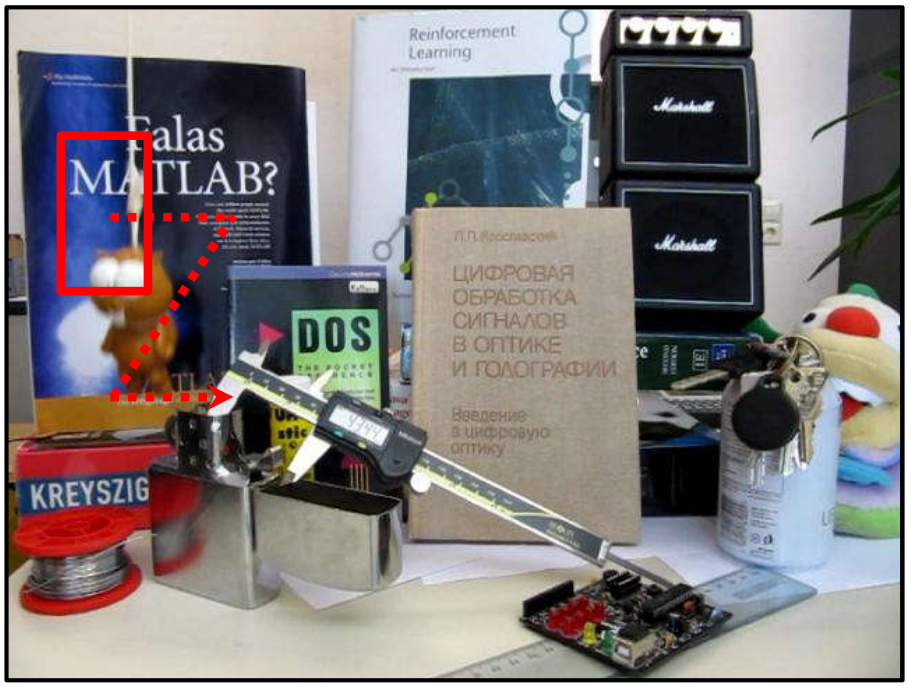


Frame 1



Tracking-By-Detection

- Two key points
  - Discriminative learning



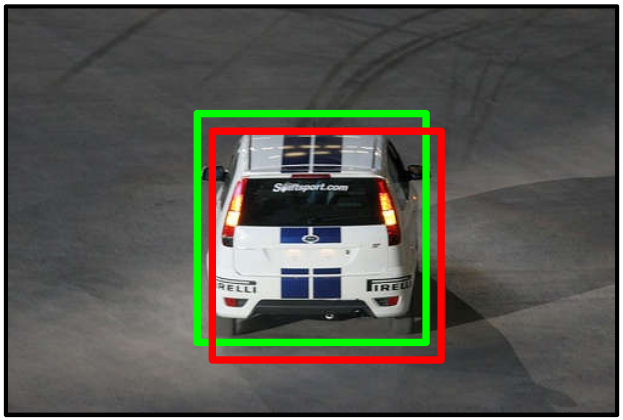
Frame 382

Similarity

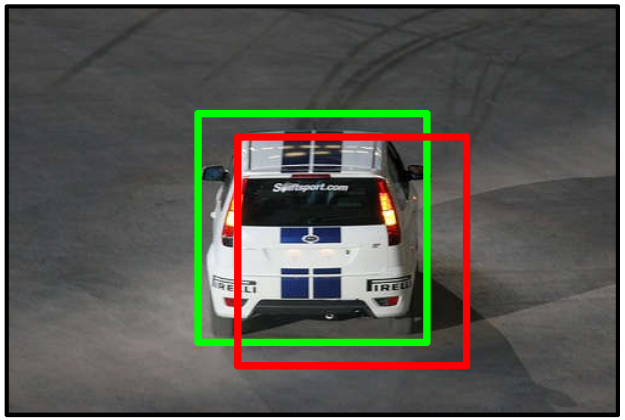
Dissimilarity

Tracking-By-Detection

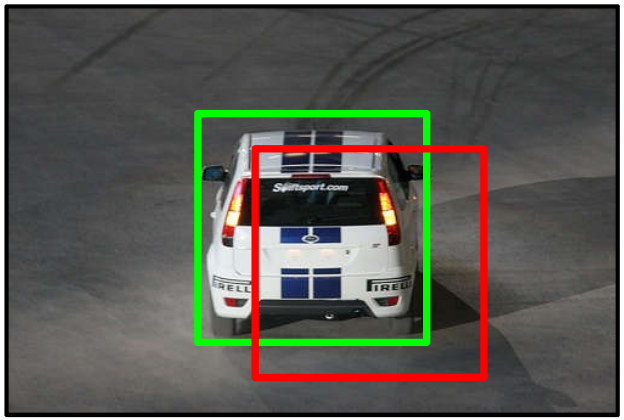
- Two key points
  - Discriminative learning
  - Pixel-accurate localization



IoU = 0.9



IoU = 0.7

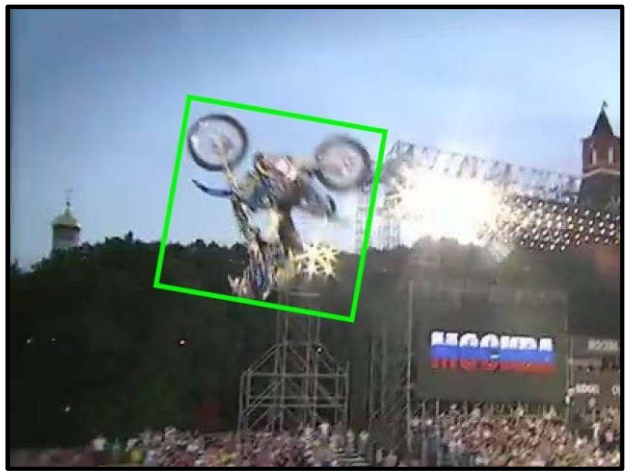


IoU = 0.5

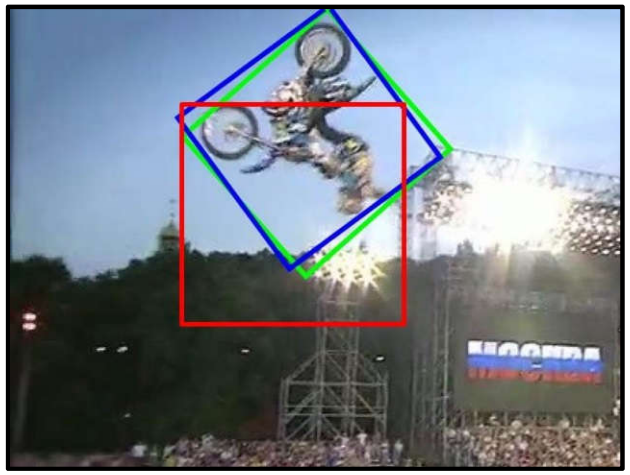
Tracking-By-Detection

- Limitations of existing methods
  - Can not handle challenging conditions where an object undergoes transformations, e.g., severe rotation
  - Select tracking results based on detection score only

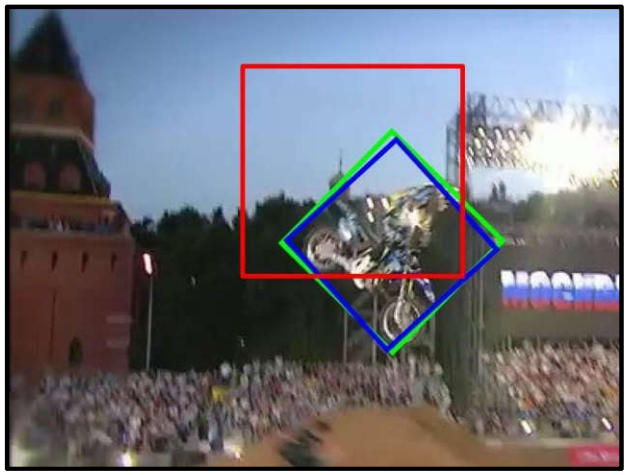
DSST Ours Groundtruth



Frame 1

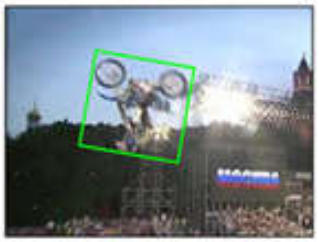


Frame 10



Frame 30

Our approach: Proposal Selection Tracking



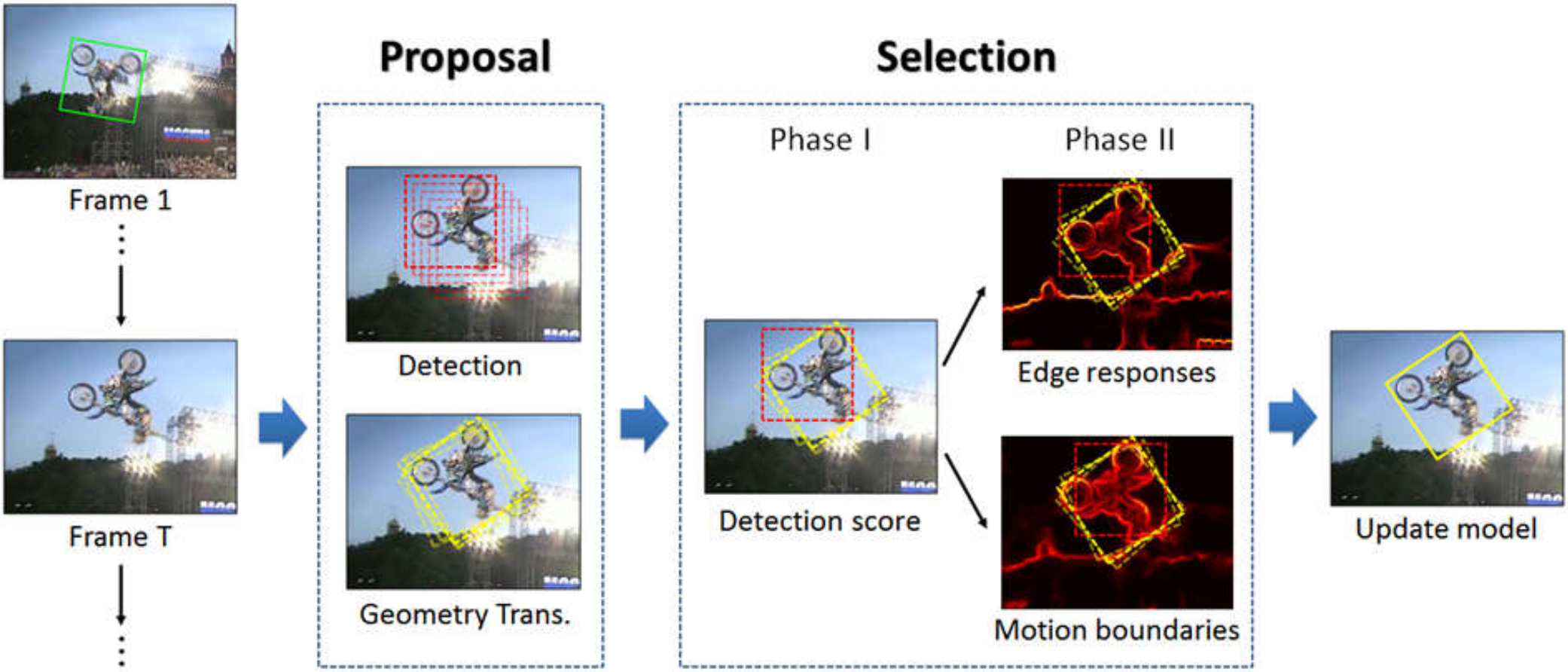
Frame 1

Frame T

Do detection

Update model

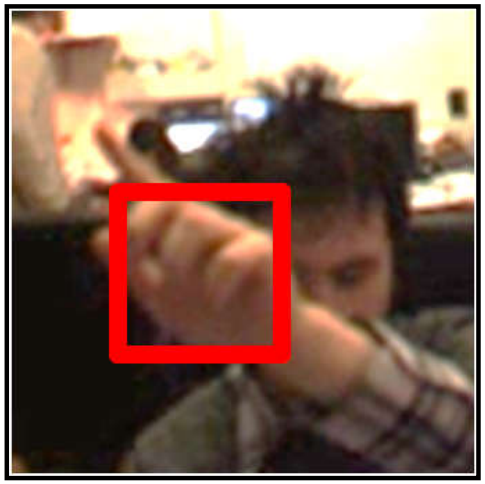
Our approach: Proposal Selection Tracking



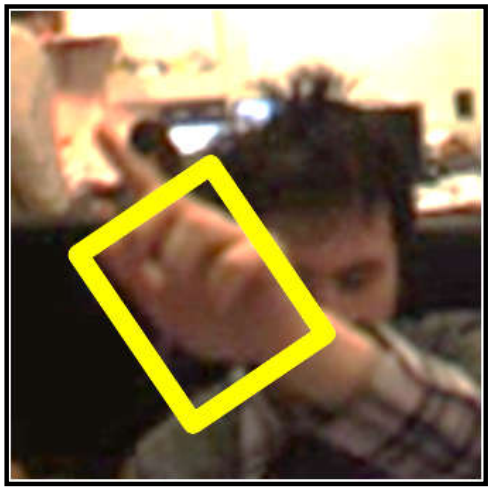
Proposals

- Detection proposals
- Geometry proposals
  - Compute frame-to-frame pixel correspondences with optical flow [Brox'11]
  - Estimate similarity transformations with Hough transform voting [Lowe'04]

Ground truth



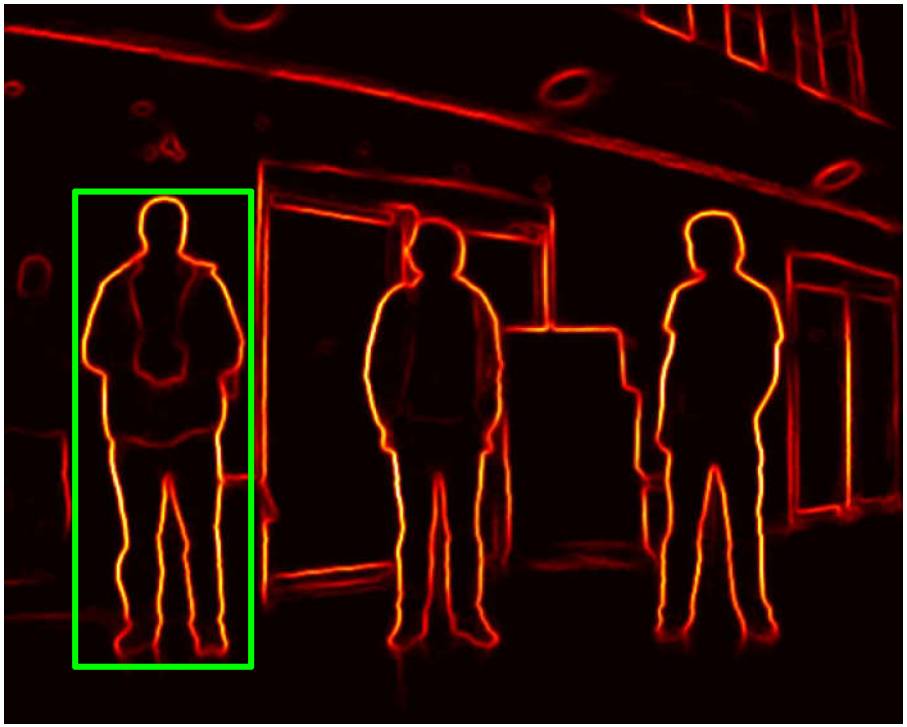
Detection proposal



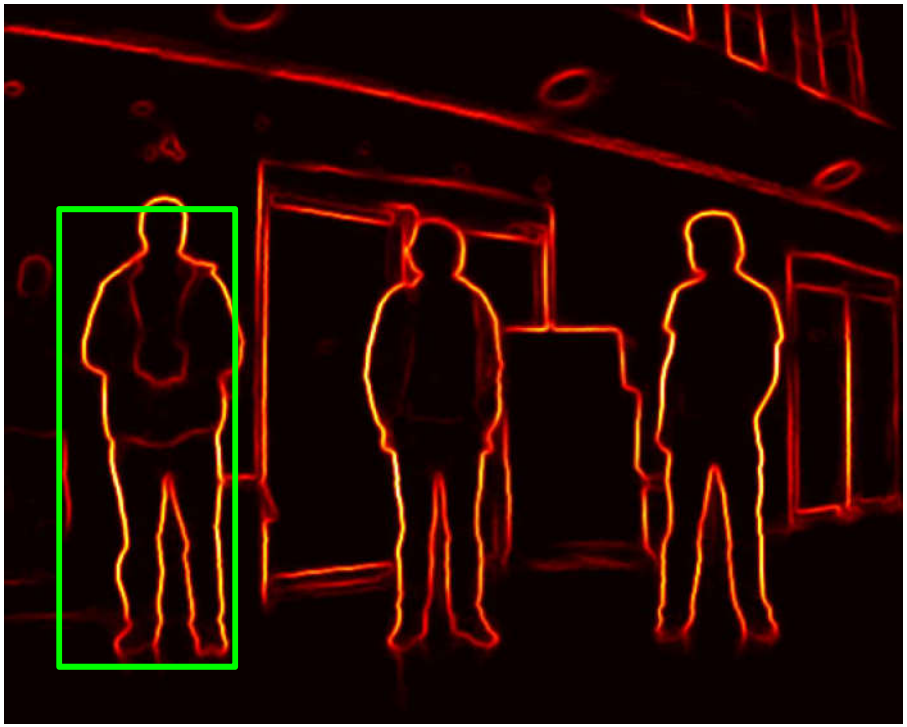
Geometry proposal

Selection

- Multiple cues for selection
  - Detection scores
  - Edgebox score [Zitnick'14], originally from edge response [Dollár'13]



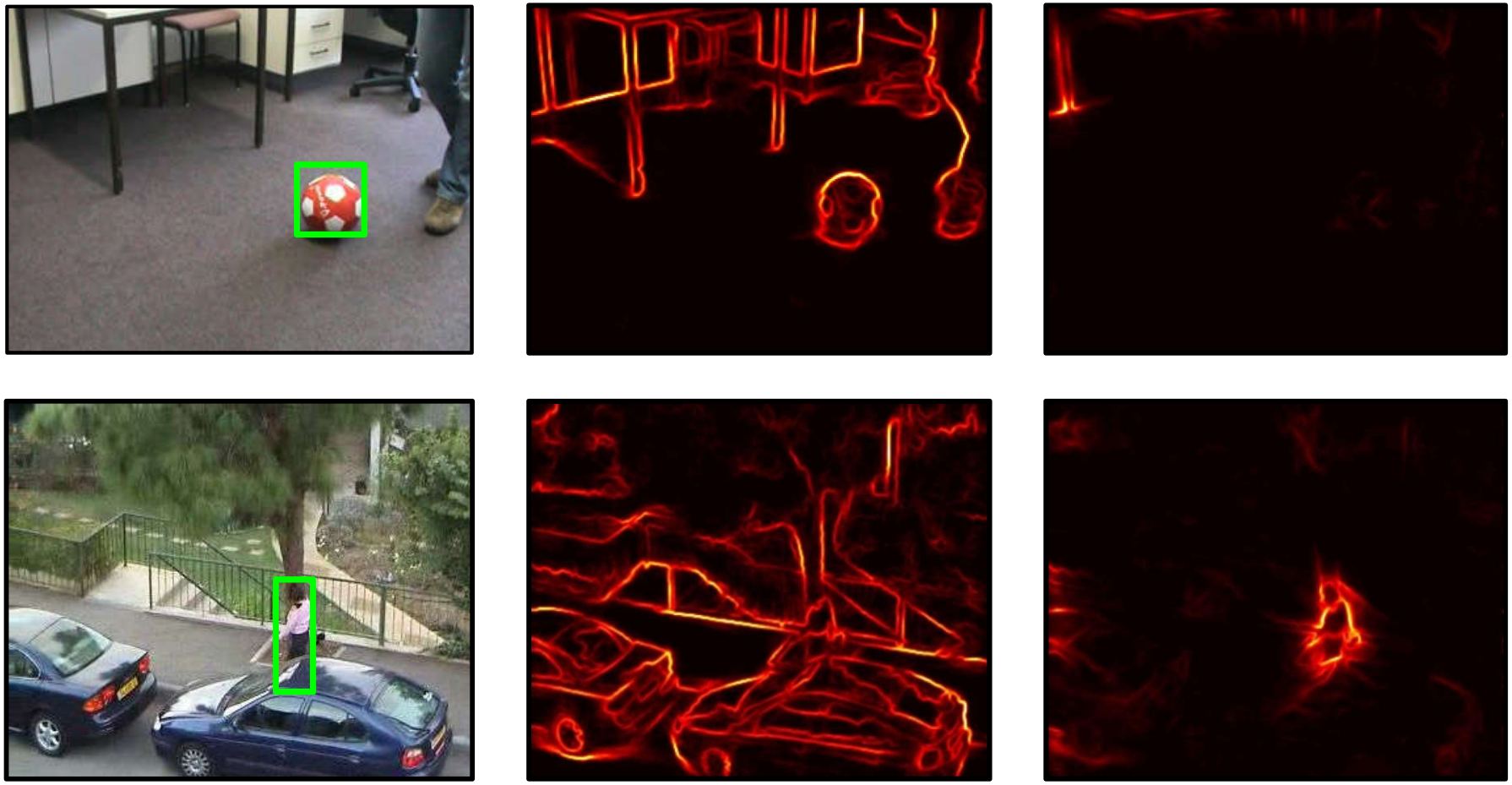
High edgebox score



Low edgebox score

Selection

- Multiple cues for selection
  - Edgebox scores from edge responses and motion boundaries [Weinzaepfel'15] are **complementary**

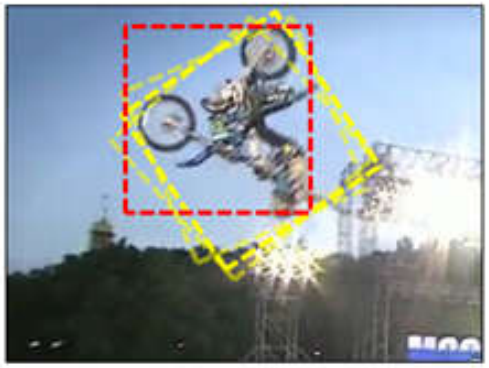


Edge Responses Motion boundaries

Selection

- How to combine multiple cues?
  - Propose a two-phase strategy

Phase I



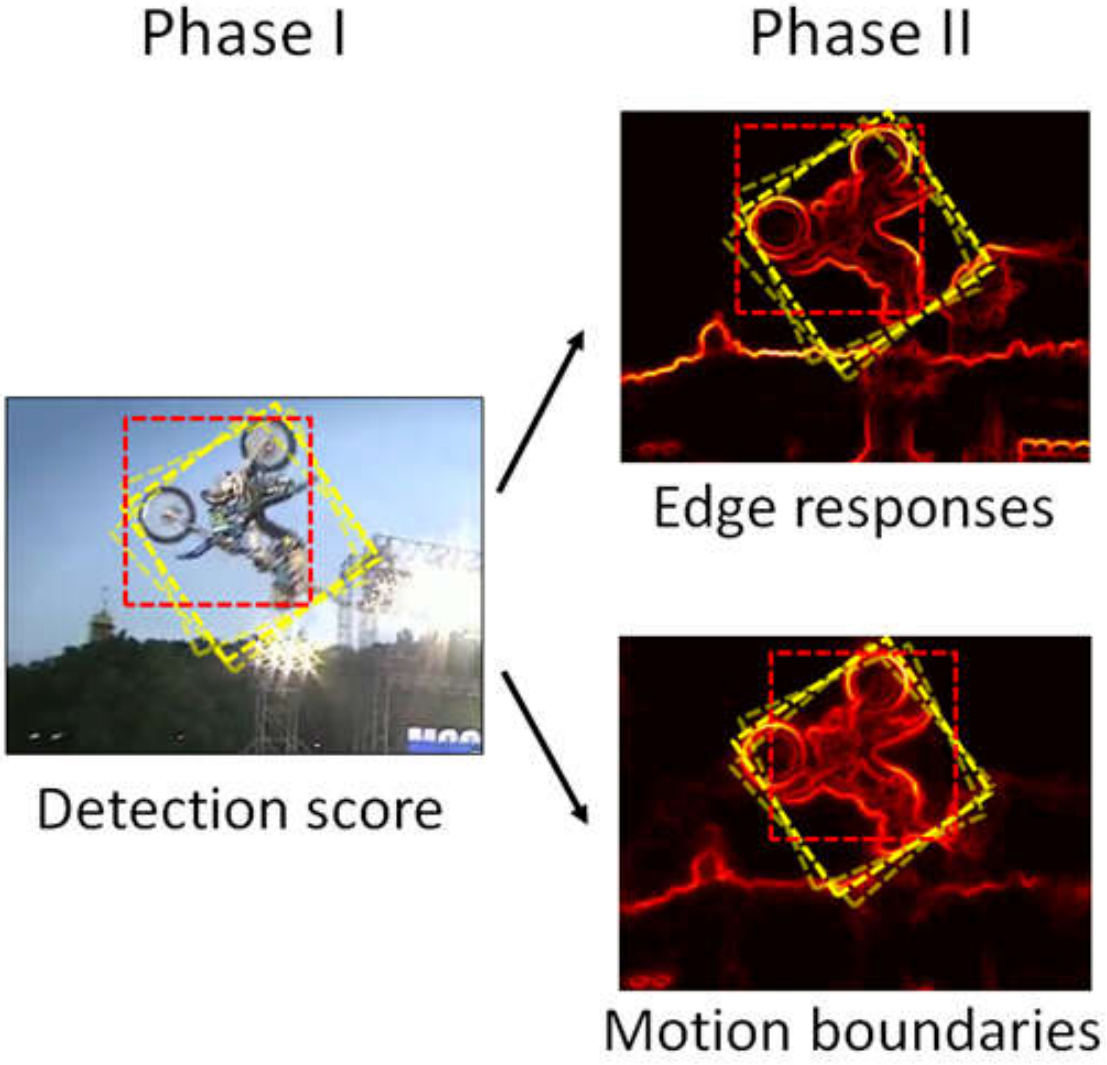
Detection score

- Phase I
  In frame t, select candidates whose detection scores (DS_i^t) are statistically similar to the best detection score (DS_{best}^t), i.e.,

$$\frac{DS_{best}^t - DS_i^t}{DS_{best}^t} < 1\%$$

Selection

- How to combine multiple cues?
  - Propose a two-phase strategy

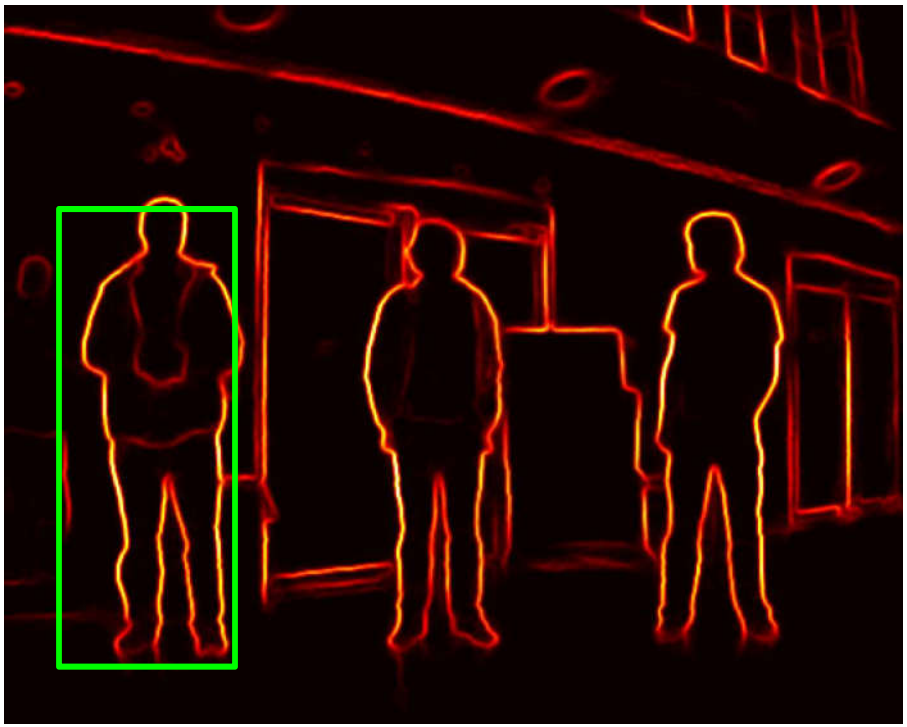


- Phase II
  Determine result with the best normalized edgebox scores from edge responses and motion boundaries

$$ES^t_{norm} = \frac{ES^t - \mu_{[t-5,t-1]}}{\sigma_{[t-5,\,t-1]}}$$

Selection

- How to combine multiple cues?
  - Need for the two-phase strategy

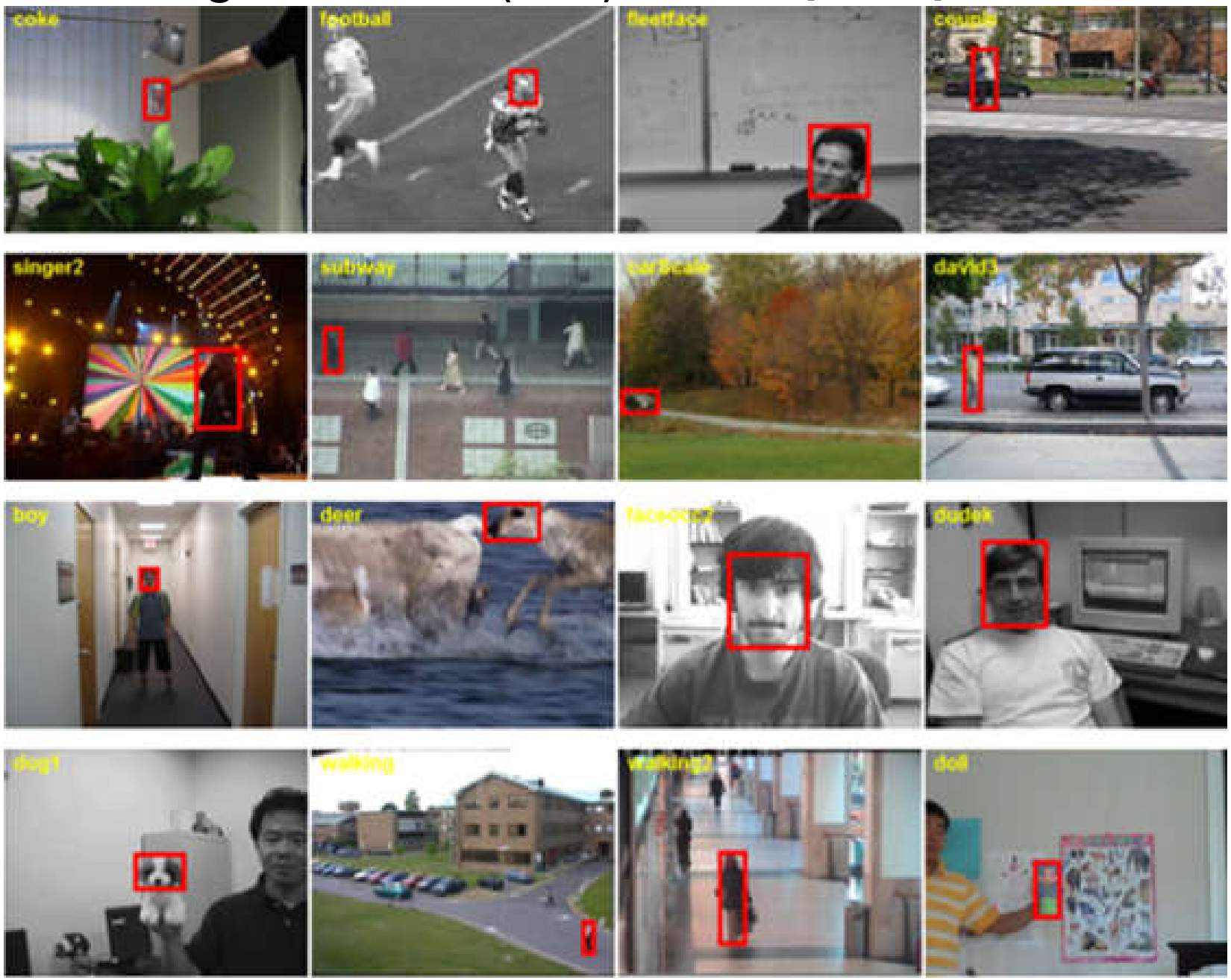


Low edgebox score, but
high detection score

High edgebox score, but
low detection score

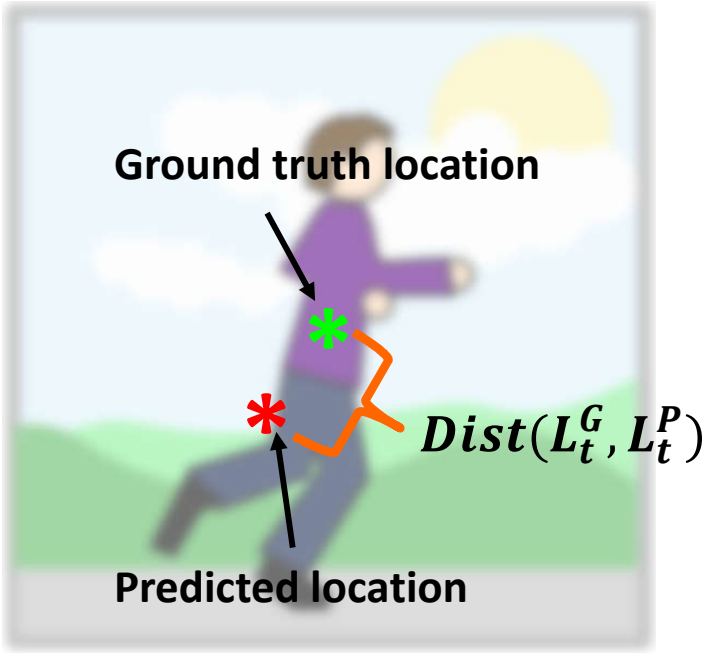
Experimental Results

- Online Tracking Benchmark (OTB) Dataset [Wu'13]



Experimental Results

- Online Tracking Benchmark (OTB) Evaluation Metrics
  - Precision Plot: Percentage of frames with location error less than threshold
  - Success Plot: Percentage of frames with overlap larger than threshold



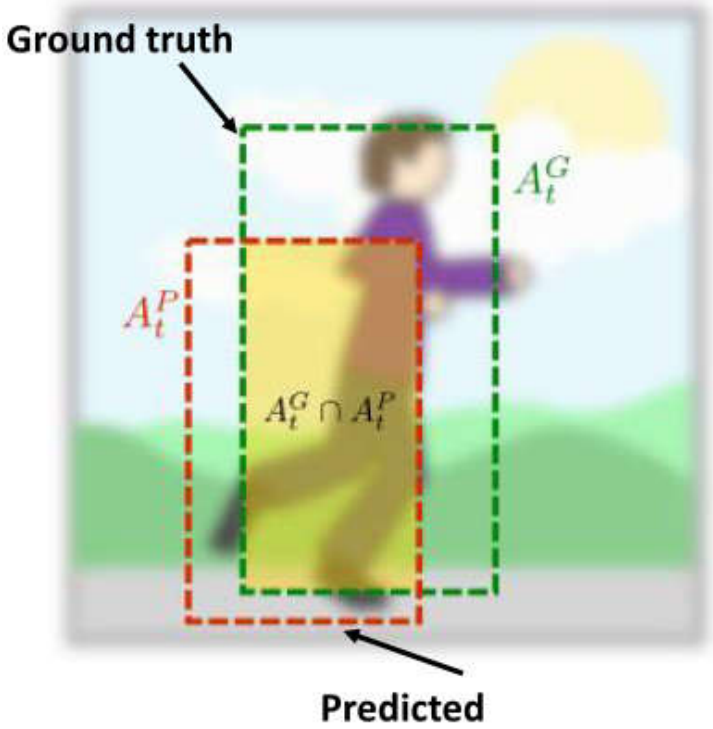
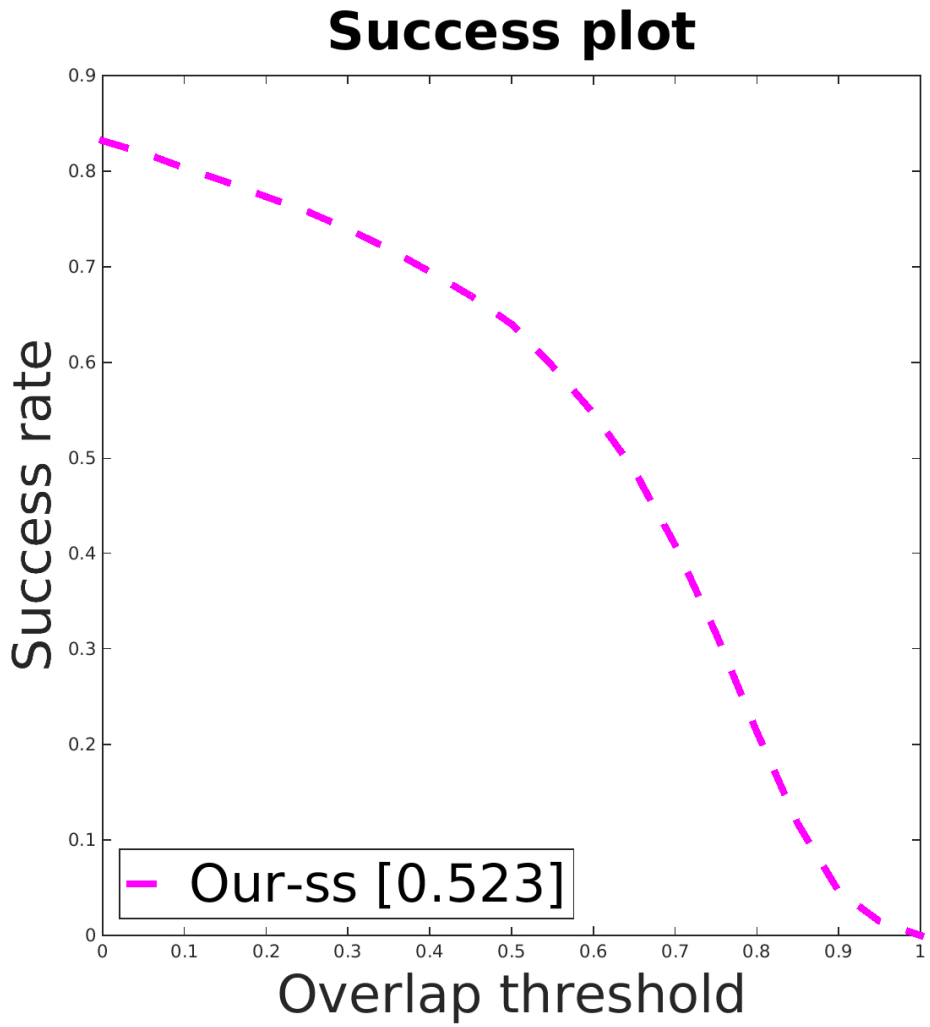


Image courtesy of [Kristan'14]

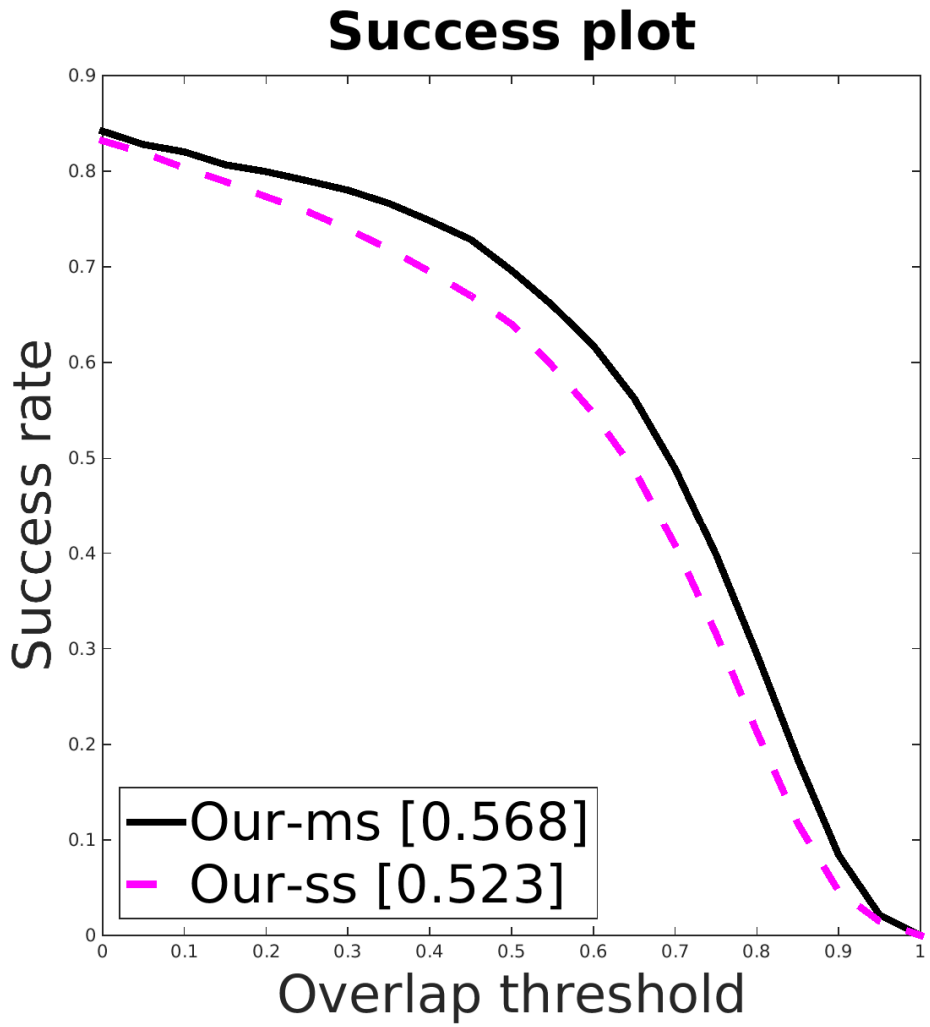
Experimental Results

- Online Tracking Benchmark (OTB) Results: Our variants
  - Our-ss
    - Proposal: detection (single scale)
    - Selection: detection score



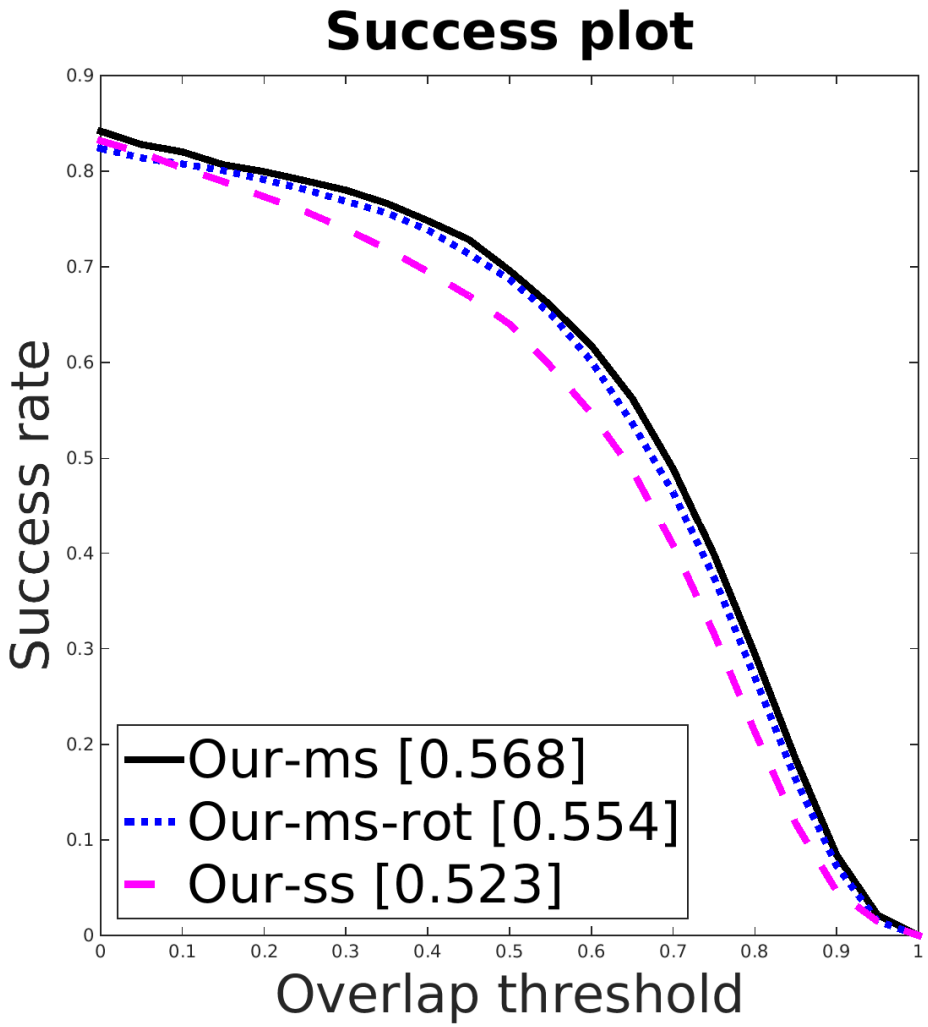
Experimental Results

- Online Tracking Benchmark (OTB) Results: Our variants
  - Our-ms
    - Proposal: detection (multiple scales)
    - Selection: detection score



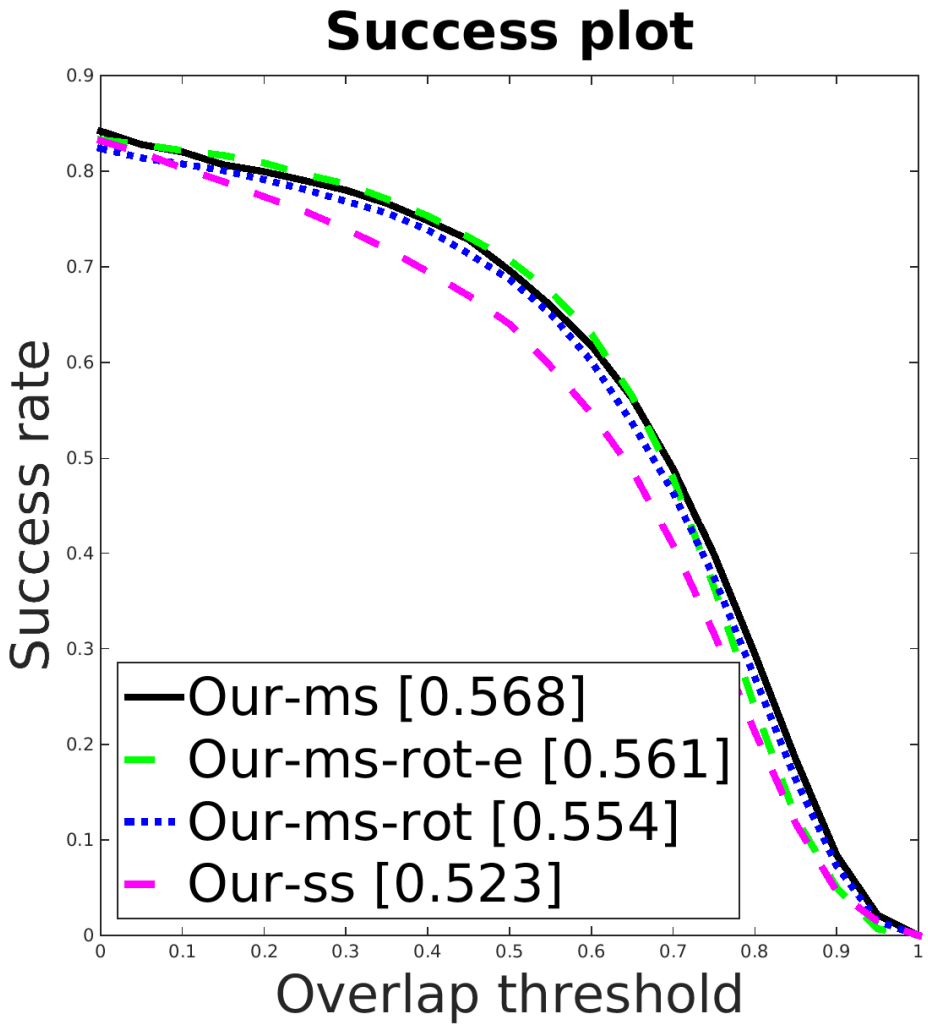
Experimental Results

- Online Tracking Benchmark (OTB) Results: Our variants
  - Our-ms-rot
    - Proposal: detection (multiple scales) + geometry
    - Selection: detection score



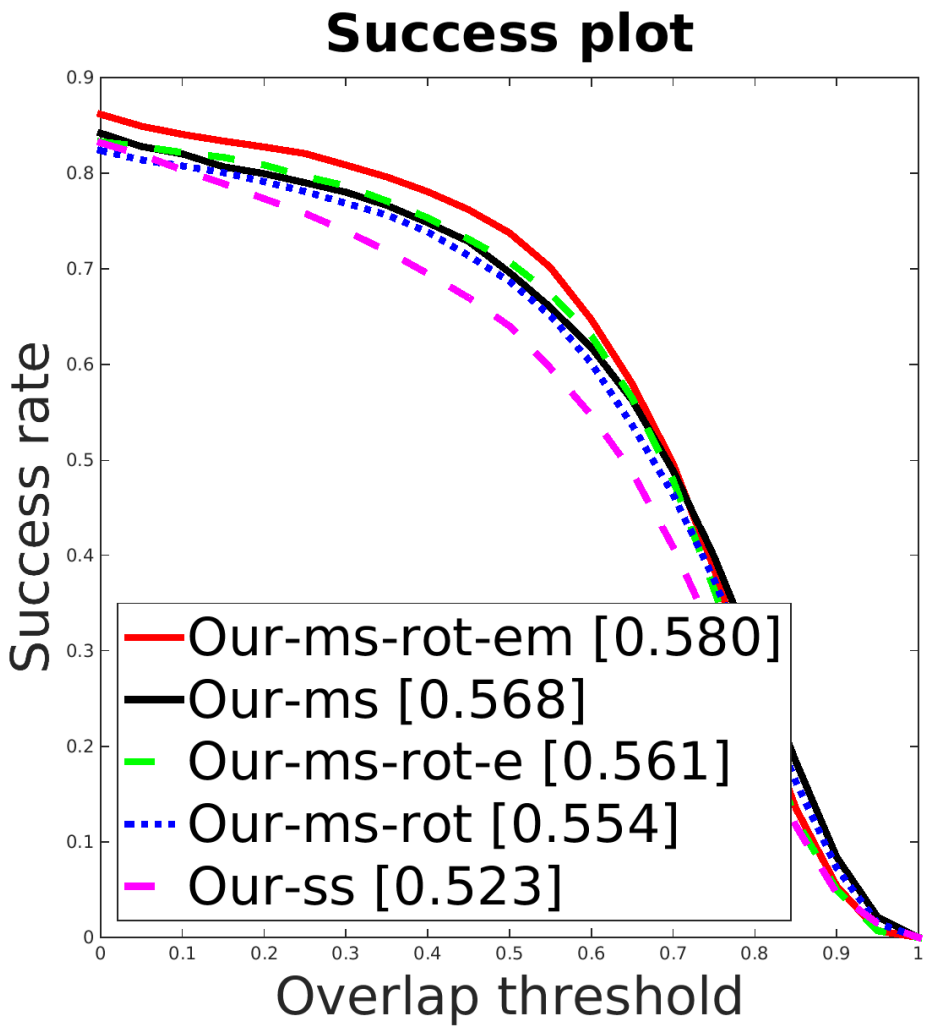
Experimental Results

- Online Tracking Benchmark (OTB) Results: Our variants
  - Our-ms-rot-e
    - Proposal: detection (multiple scales) + geometry
    - Selection: detection + edgebox (edge response) scores



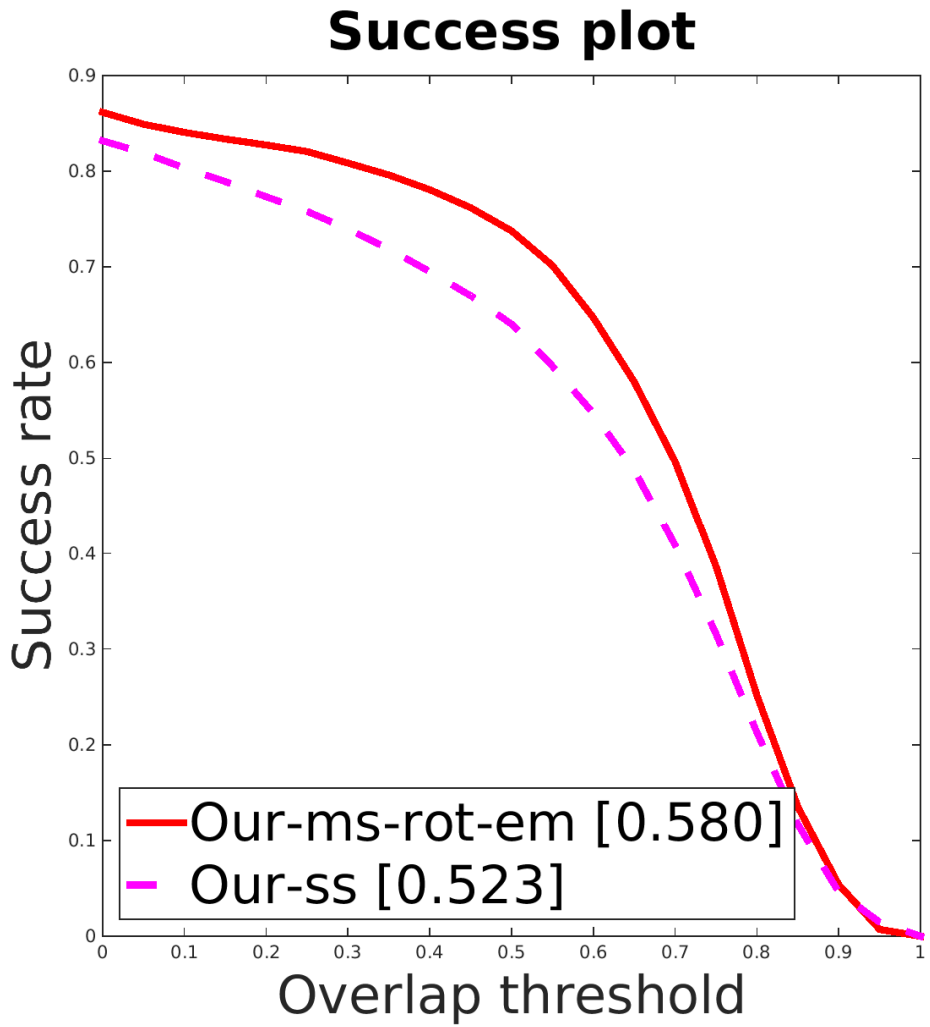
Experimental Results

- Online Tracking Benchmark (OTB) Results: Our variants
  - Our-ms-rot-em
    - Proposal: detection (multiple scales) + geometry
    - Selection: detection + edgebox (edge & motion boundary response) scores



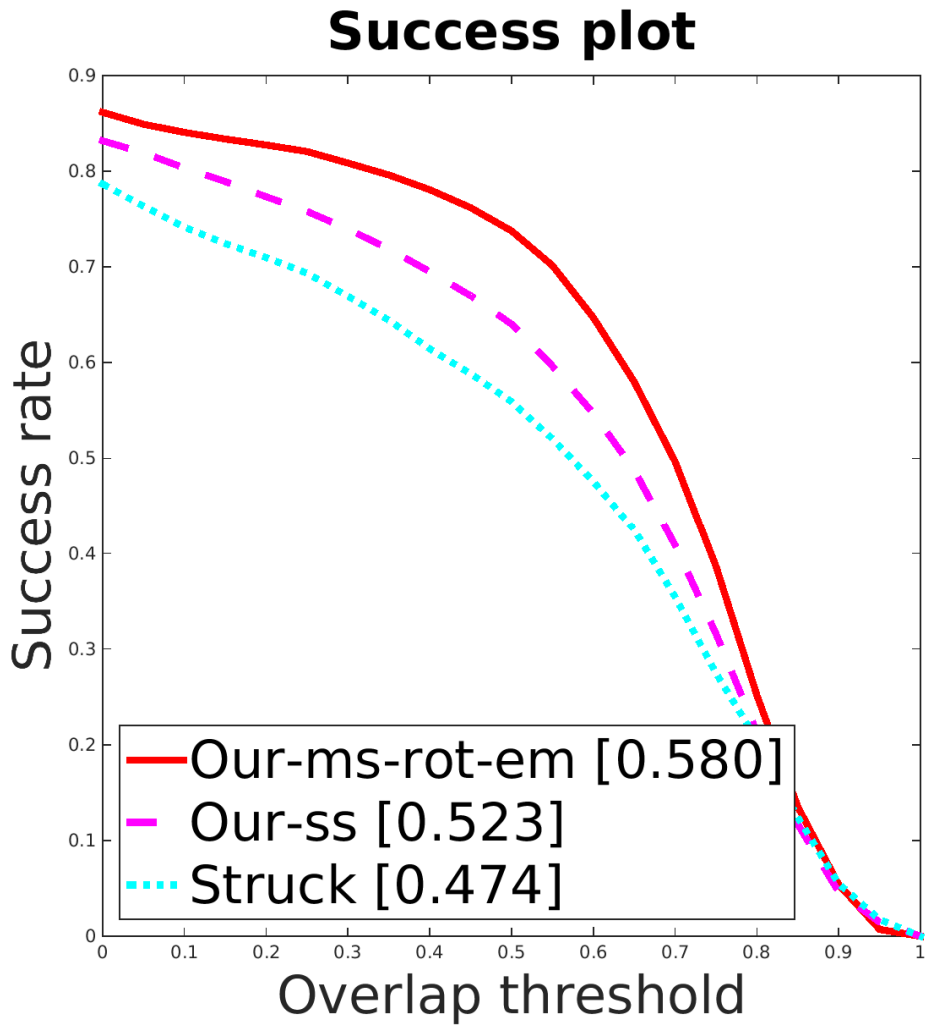
Experimental Results

- Online Tracking Benchmark (OTB) Results: Ours vs. others



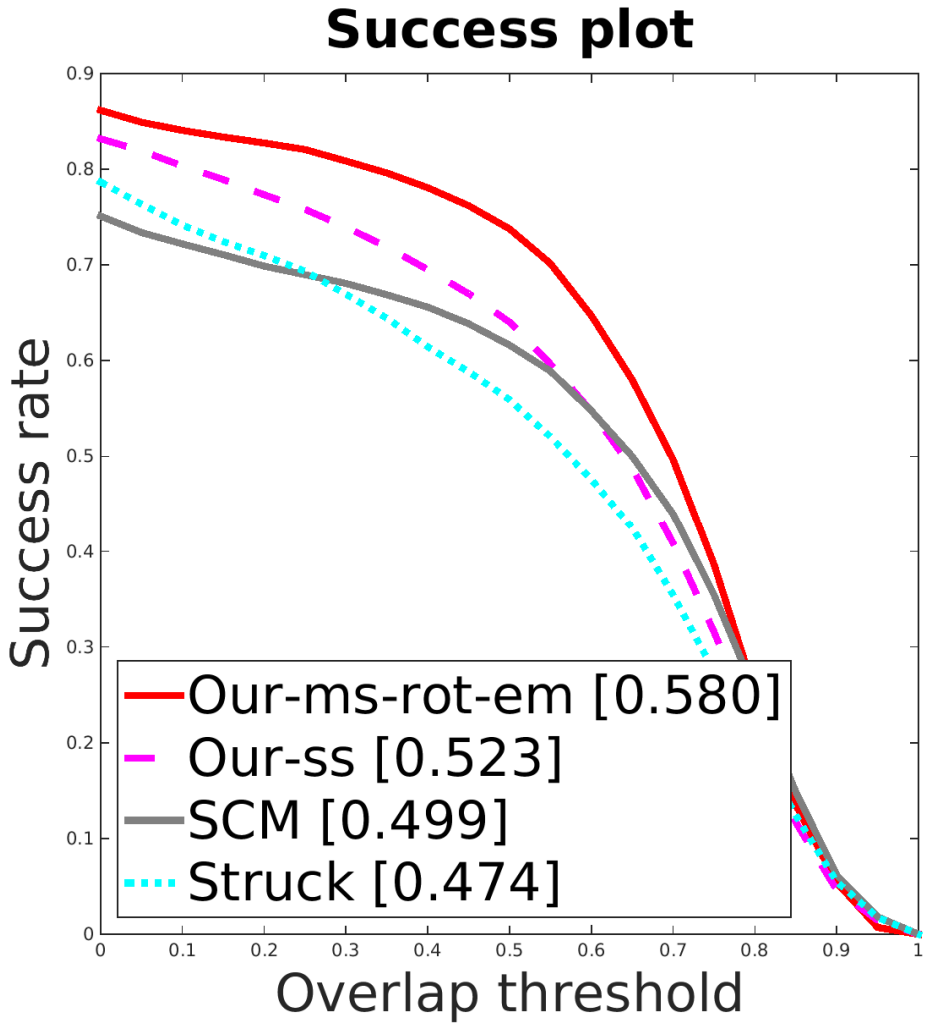
Experimental Results

- Online Tracking Benchmark (OTB) Results: Ours vs. others



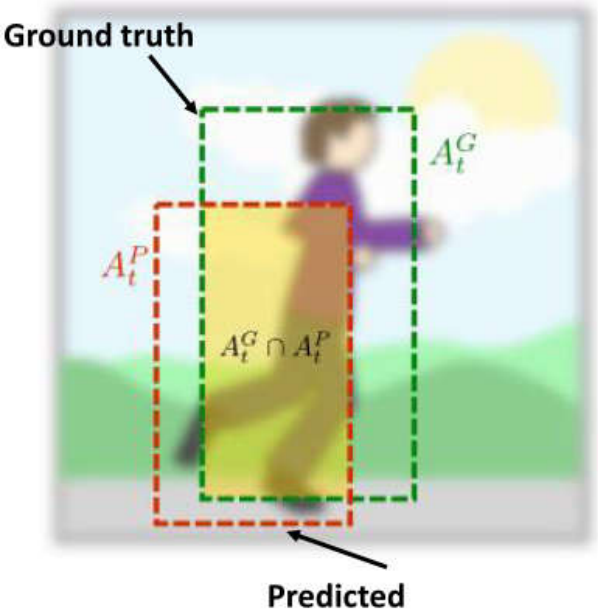
Experimental Results

- Online Tracking Benchmark (OTB) Results: Ours vs. others



Experimental Results

- Participated in VOT 2015 competitions with a **simplified framework**
- Visual Object Tracking (VOT) Challenge Evaluation Metrics
  - Accuracy: Average overlap during successful tracking
  - Robustness: Number of times a tracker drifts off the target



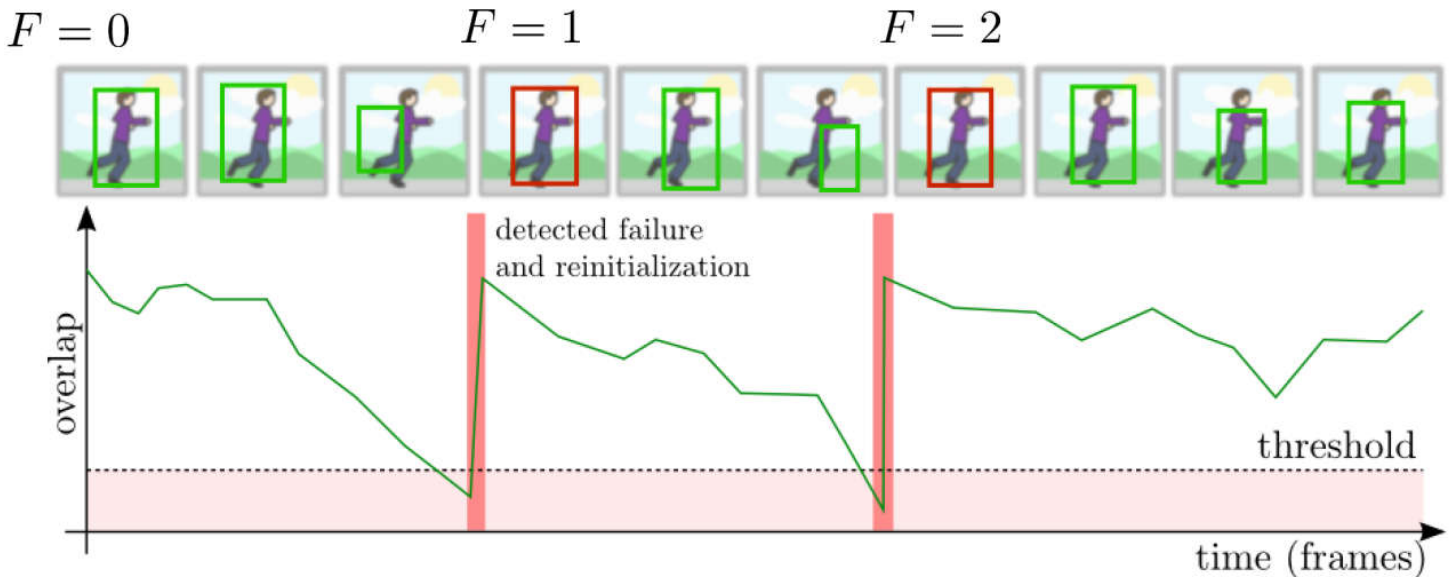
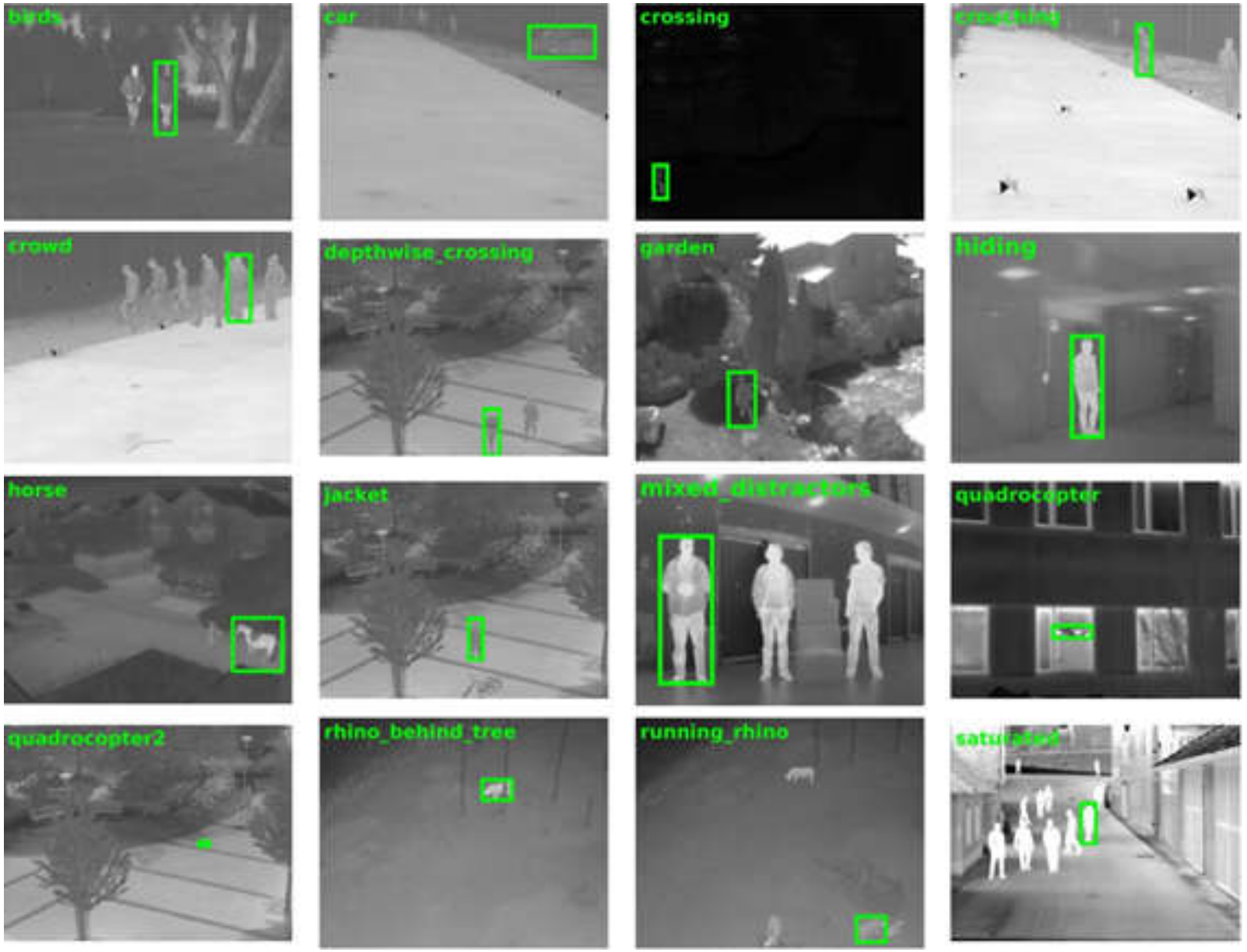


Image courtesy of [Kristan'14]

Experimental Results

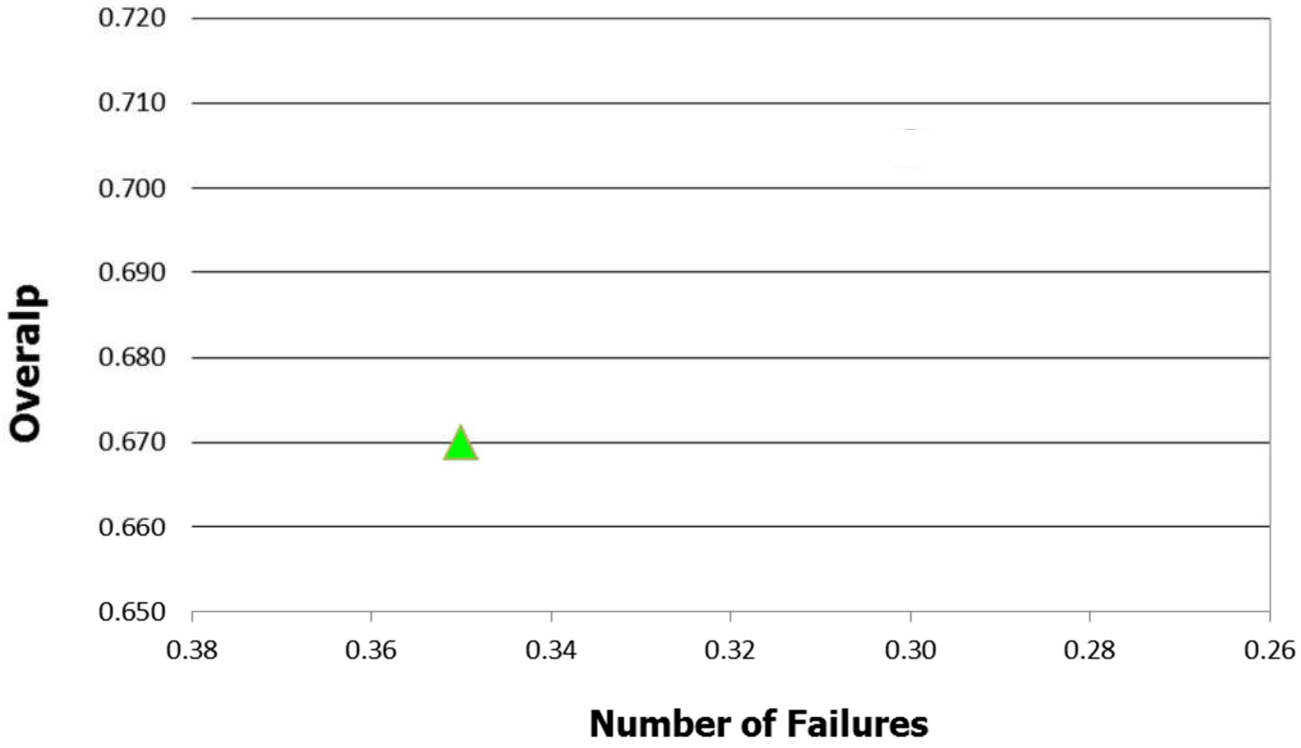
- Thermal Infrared Visual Object Tracking (VOT-TIR) 2015 Challenge Dataset (20 seq.) [Felsberg'15]



Experimental Results

- Thermal Infrared Visual Object Tracking (VOT-TIR) 2015 Challenge Results

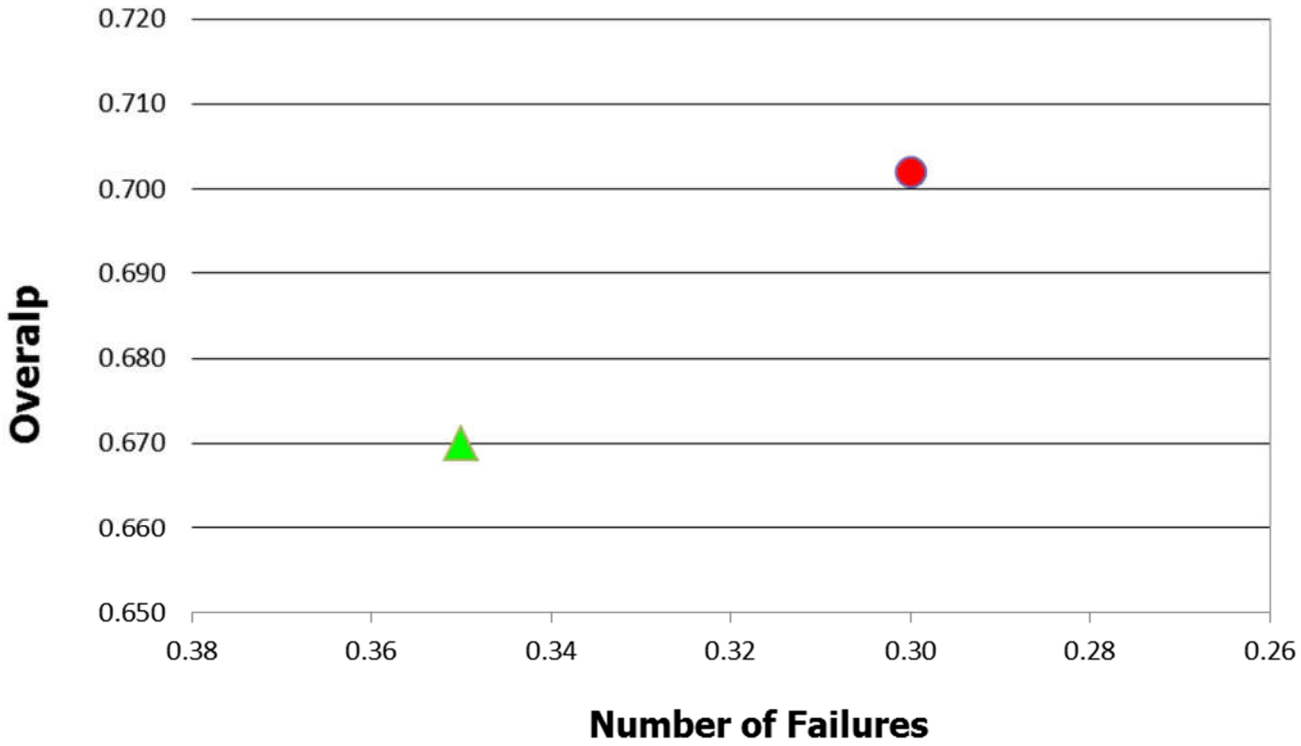
▲ Selection: detection score only	
Acc. (Overlap)	Rob. (#Failures)
0.670	0.35



Experimental Results

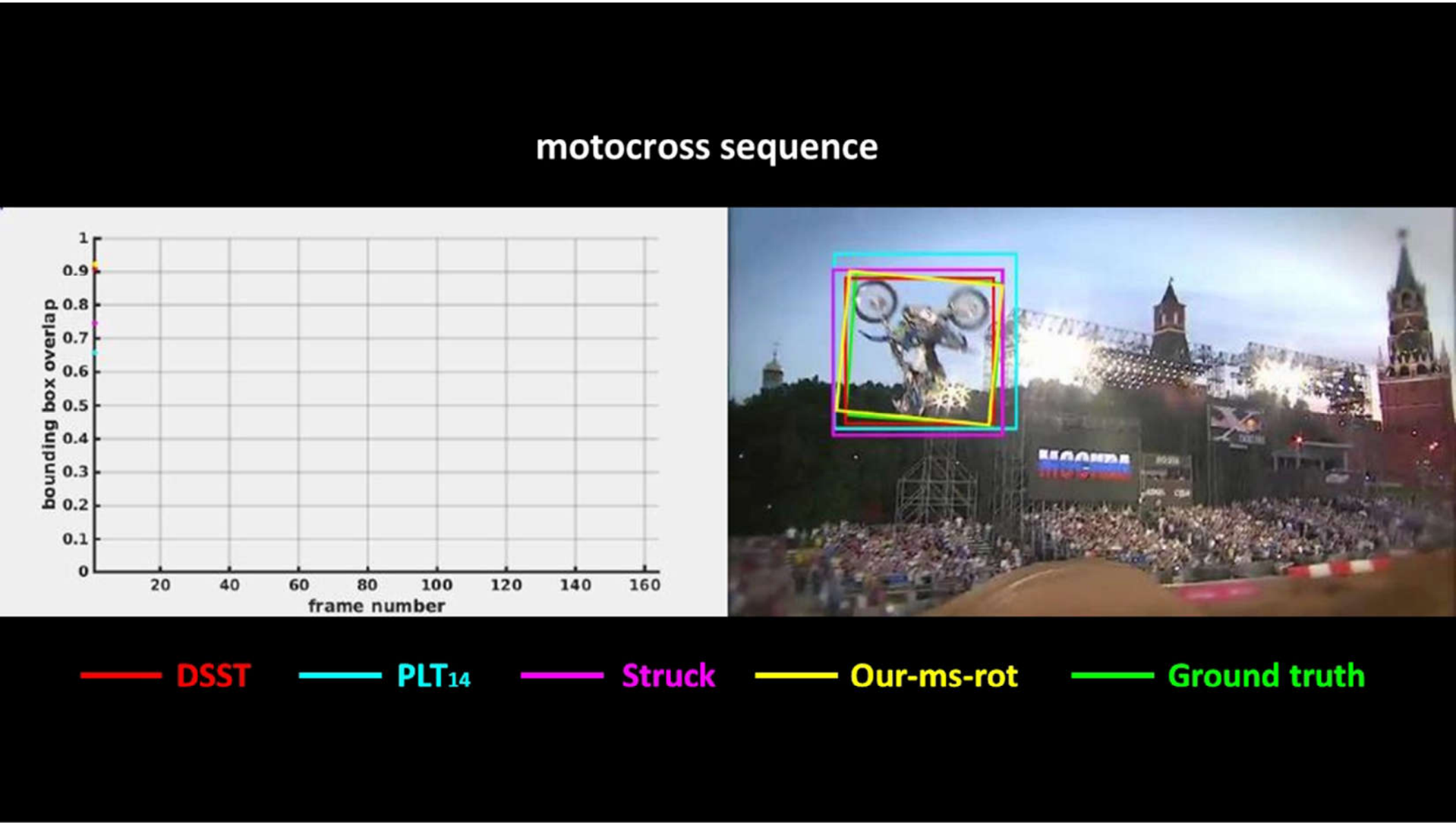
- Thermal Infrared Visual Object Tracking (VOT-TIR) 2015 Challenge Results

▲ Selection: detection score only		● Selection: detection + edgebox score	
Acc. (Overlap)	Rob. (#Failures)	Acc. (Overlap)	Rob. (#Failures)
0.670	0.35	**0.702**	**0.30**



Experimental Results

- Video demo



Summary: Proposal Selection Tracking

- Proposed a generalized discriminative tracking-by-detection framework for short-term tracking
  - New geometry proposals
  - A novel selection scheme based on multiple cues
- Achieved state-of-the-art performance on challenging datasets
- Participated in recent challenges
  - "Winning tracker" from 24 trackers in the VOT-TIR2015
  - Ranked sixth among 62 trackers in the VOT2015

- Publication
  - Y. Hua, K. Alahari, and C. Schmid. Online object tracking with proposal selection. In ICCV, 2015
- Source code released at
  - http://thoth.inrialpes.fr/research/pstracker/

Outline

- Online Object Tracking with Proposal Selection
- ***Occlusion and Motion Reasoning for Long-term Tracking***
- Conclusion and Future Work

Background

- Many tracking methods suffer from the **template update problem** [Matthews'04]
  - To update, or not to update?

Update in every frame

Frame 1 Frame 245 Frame 341 Frame 468

No update

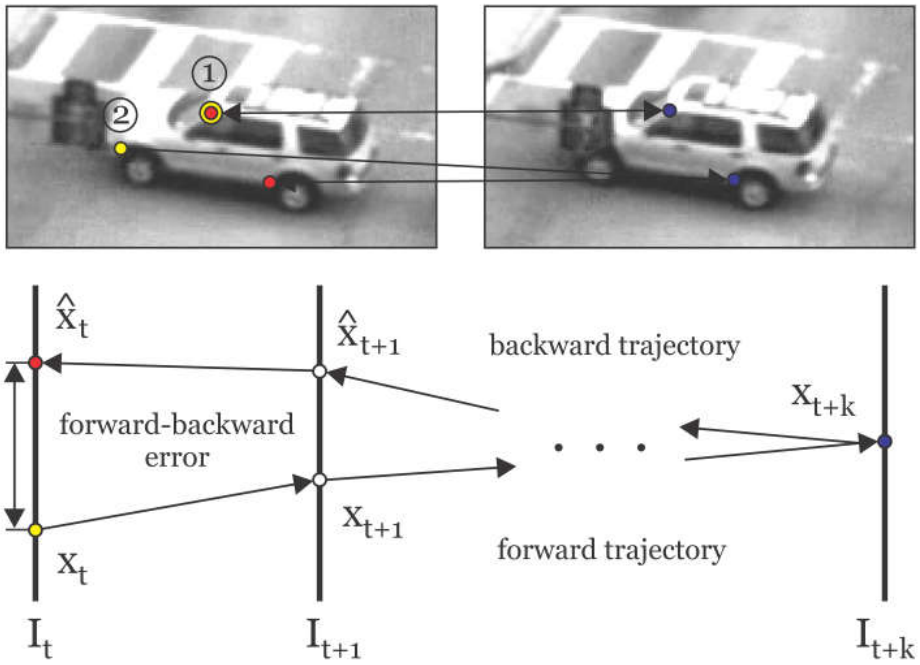
Related Work

- Long-term tracking
  - Investigated in “Tracking-Learning-Detection” [Kalal’12]
  - “… where the object may become **occluded**, significantly change scale, and **leave/re-enter the field-of-view**” [Supancic III’13]

Image courtesy of [Kalal’12]

Related Work

- Key step for long-term tracking: **When to update the model**
  - Forward-backward tracking for checking errors [Kalal'10/12]
  - Self-paced learning for collecting relevant samples [Supancic III'13]



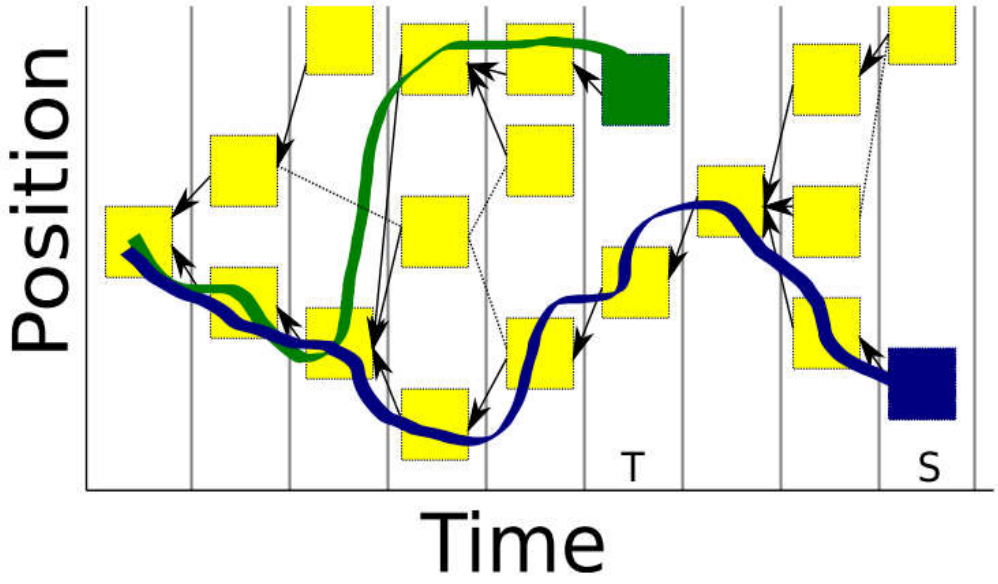
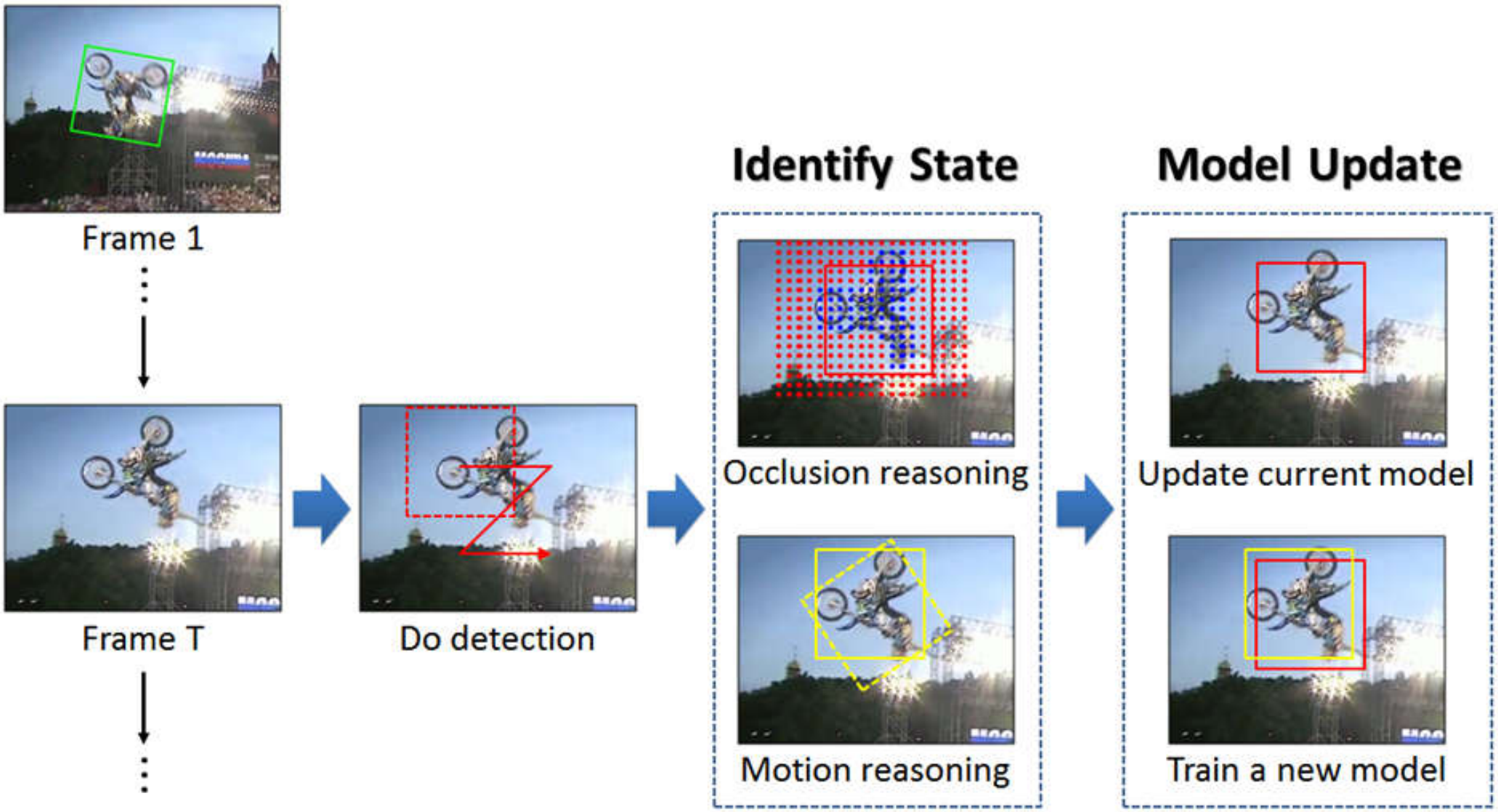


Image courtesy of [Kalal'10, Supancic III'13]

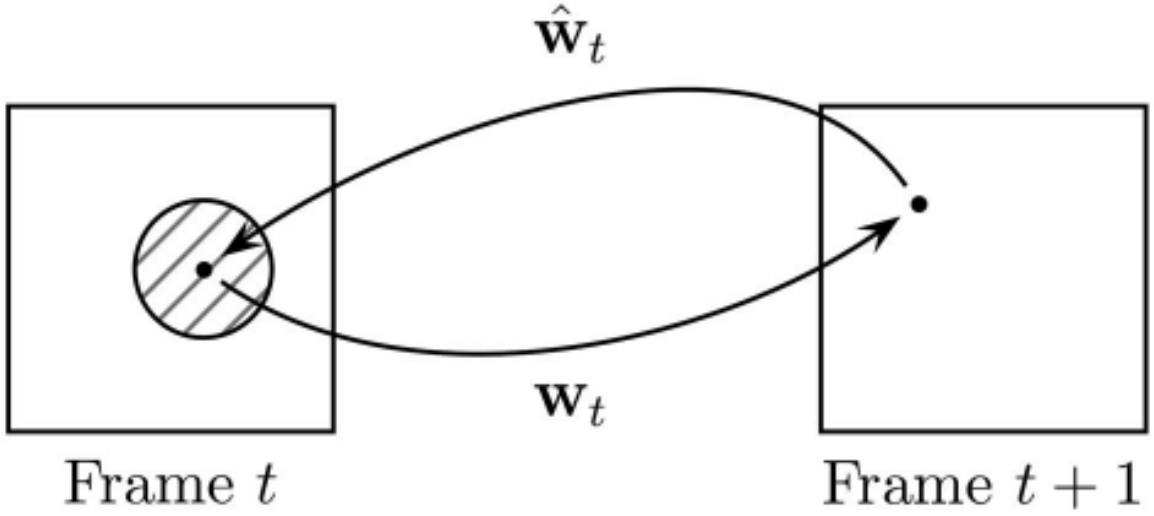
Our approach: Occlusion and Motion Reasoning

Our approach: Occlusion and Motion Reasoning



Motion Cues

- Optical flow
- Long-term tracks [Ochs'14]
  - Built on dense optical flow method [Brox'11]
  - Verification step: consistency of forward and backward flow



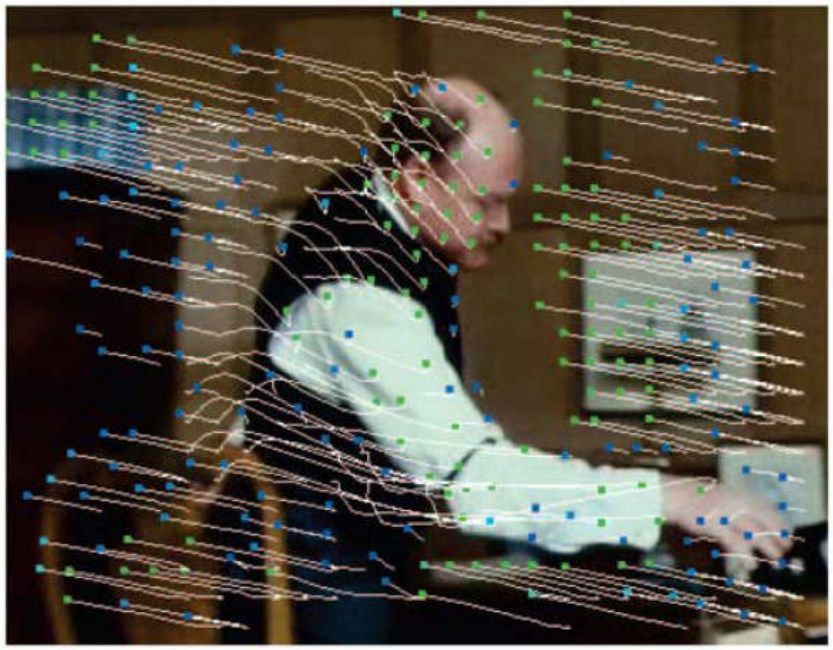
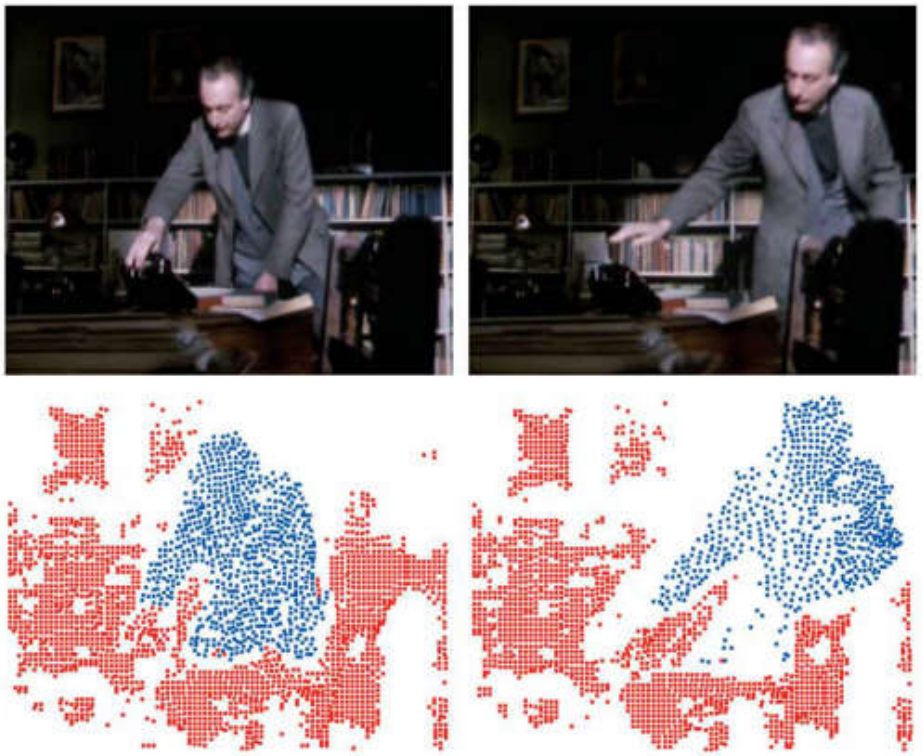


Image courtesy of [Ochs'14]

Motion Cues

- Application
  - Segmentation of Moving Objects by Long Term Video Analysis [Ochs'14]
  - Spatio-temporal Video Segmentation with Long-range Motion Cues [Lezama'11]



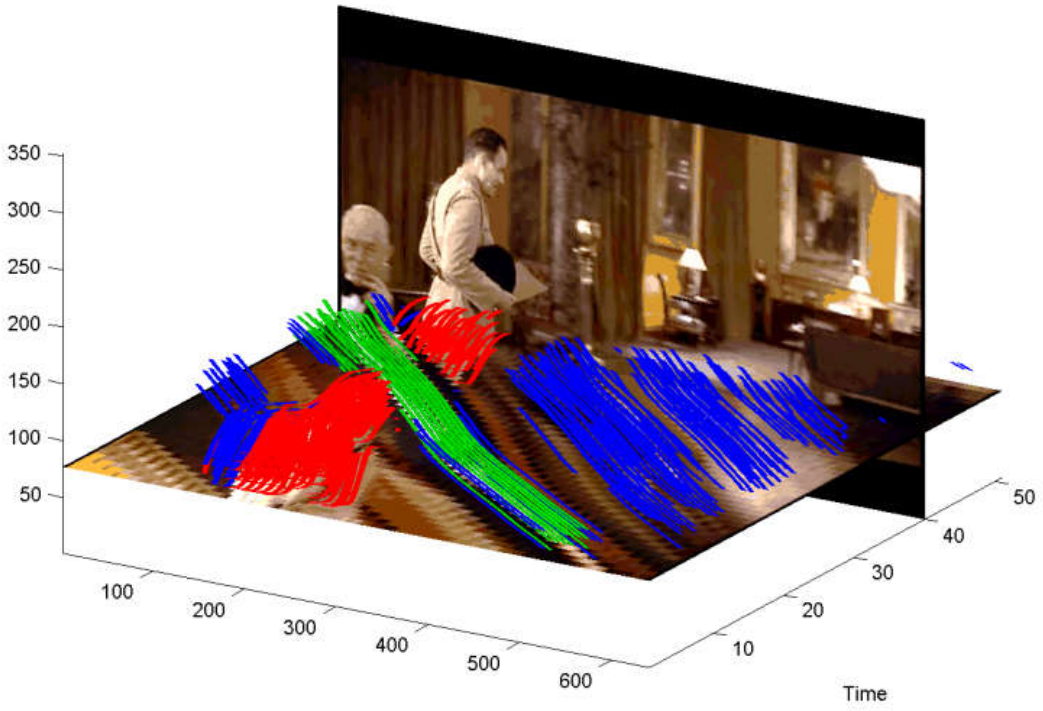


Image courtesy of [Ochs'14, Lezama'11]

Occlusion Reasoning

- Goal: Label long-term tracks as **foreground** or **background**
- Formulation: Energy minimization problem

$$E(x) = \sum_{i=1}^{n} \phi_i(x_i) + \lambda \sum_{(i,j)\in\varepsilon} \phi_{ij}(x_i, x_j)$$

$$where\ \phi_i(x_i = 1) = \begin{cases} 1 - \dfrac{1}{1 + \exp(\alpha * d_t + \beta)} & if\ track_i \in box_t \\ 0.5 & otherwise \end{cases}$$

detection score of box_t

Occlusion Reasoning

- Goal: Label long-term tracks as **foreground** or **background**
- Formulation: Energy minimization problem

$$E(x) = \sum_{i=1}^{n} \phi_i(x_i) + \lambda \sum_{(i,j)\in\varepsilon} \phi_{ij}(x_i, x_j)$$

$$where\ \phi_{ij}(x_i, x_j) = \begin{cases} \exp\left(-\lambda_d d(i,j)\right) & if\ x_i \neq x_j \\ 0 & otherwise \end{cases}$$

pointwise distance between tracks

Occlusion Reasoning

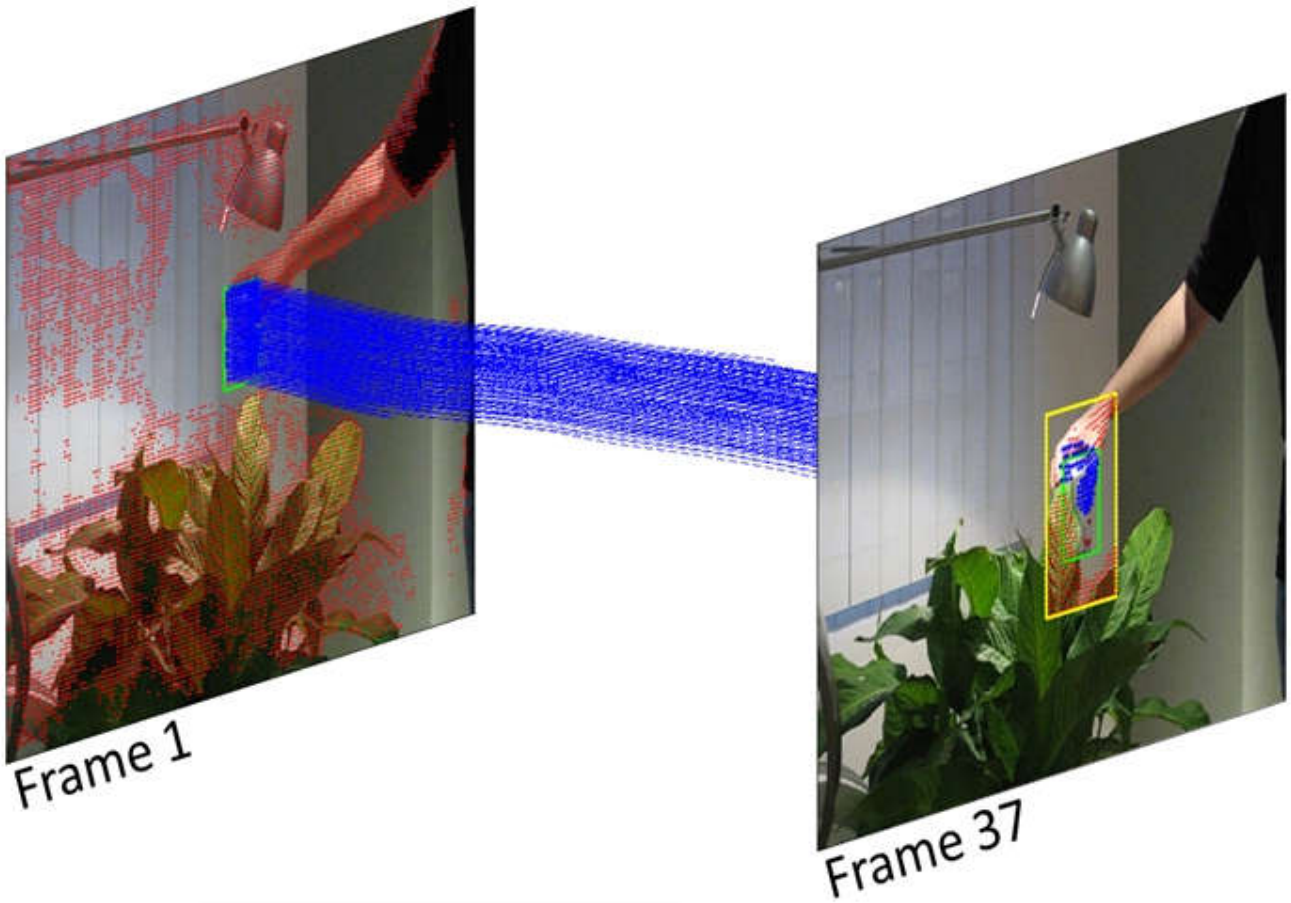
- Goal: Label long-term tracks as **foreground** or **background**
- Formulation: Energy minimization problem

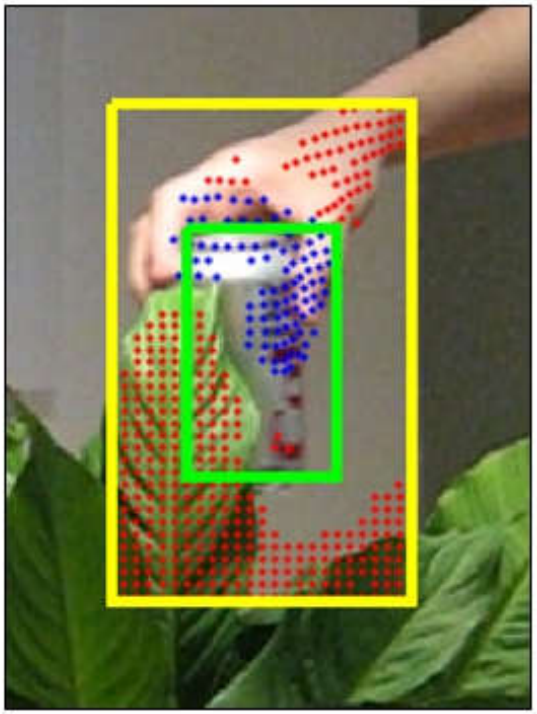
$$E(x) = \sum_{i=1}^{n} \phi_i(x_i) + \lambda \sum_{(i,j)\in\varepsilon} \phi_{ij}(x_i, x_j)$$

- Solver: Graph cuts algorithm [Kolmogorov'04]

Occlusion Reasoning

- Occlusion state estimation via long-term tracks

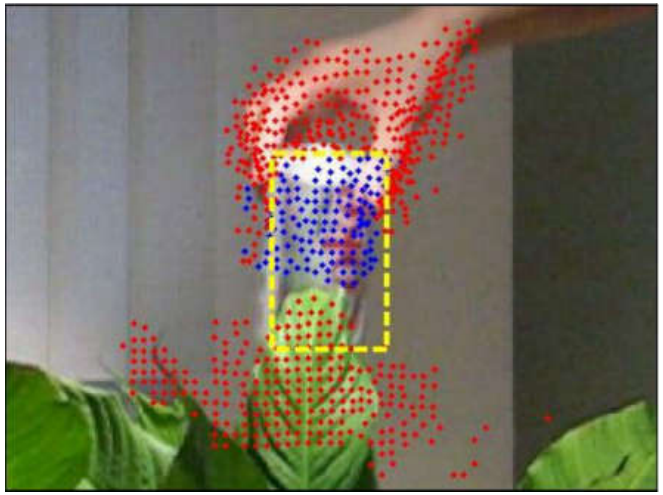




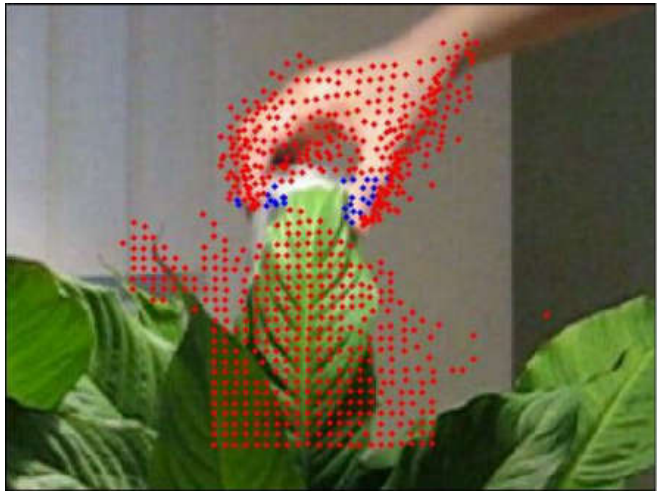
Track labels

Occlusion Reasoning

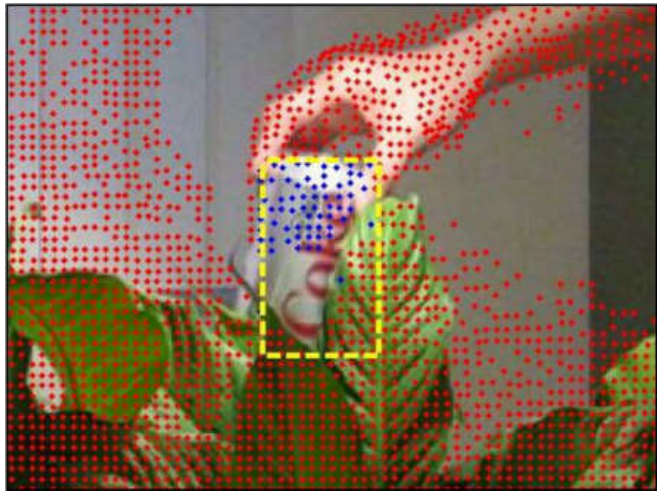
- Selectively update model according to the state of the object



Frame 251: Partial occlusion
Continue to track and no model updating



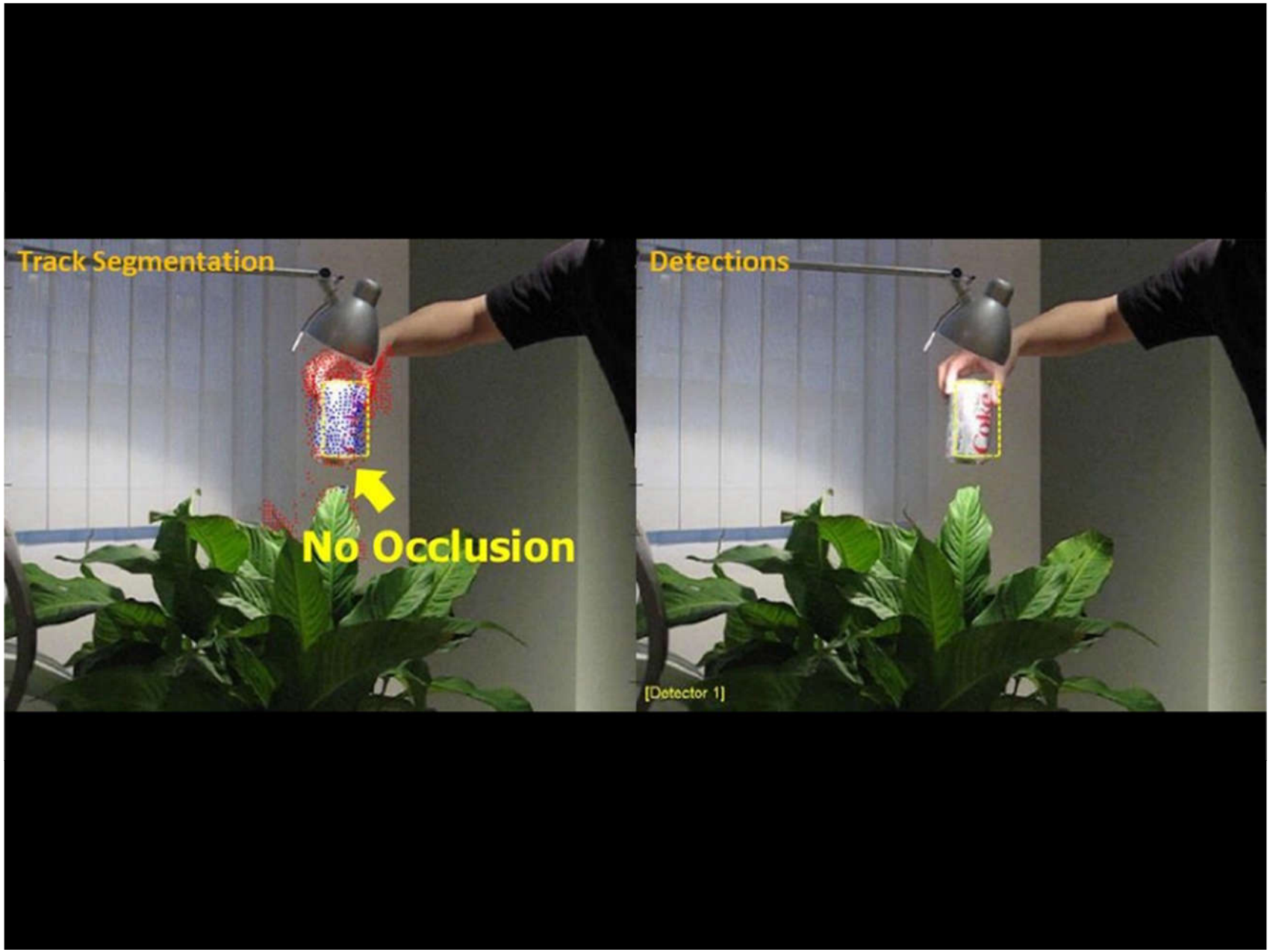
Frame 254: Full occlusion
Stop tracking and no model updating



Frame 269: Object reappears
Recover from occlusion by scanning detector globally

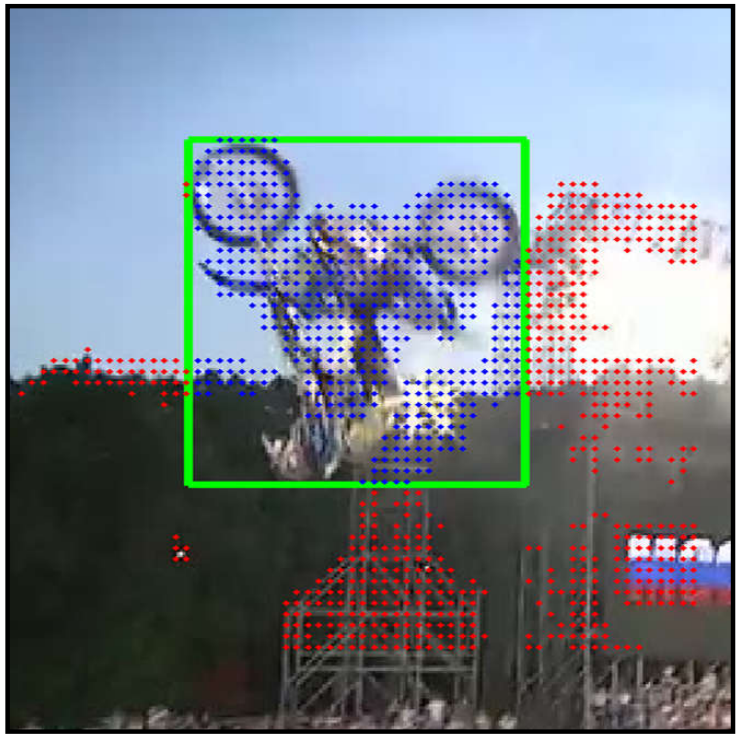
Occlusion Reasoning

- Video demo

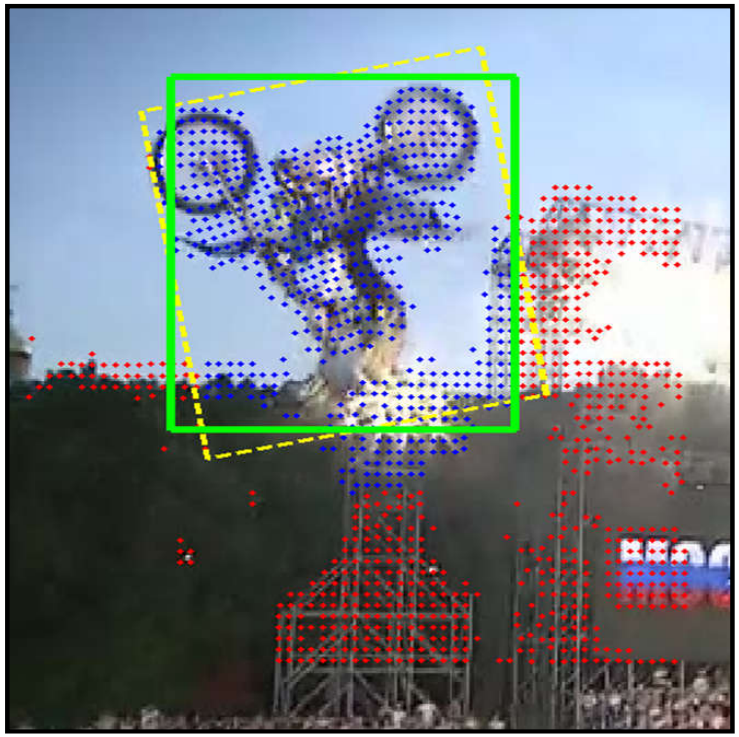


Motion Reasoning

- Estimate accumulated similarity transformations over frames
- Add a new detector if a significant change occurs



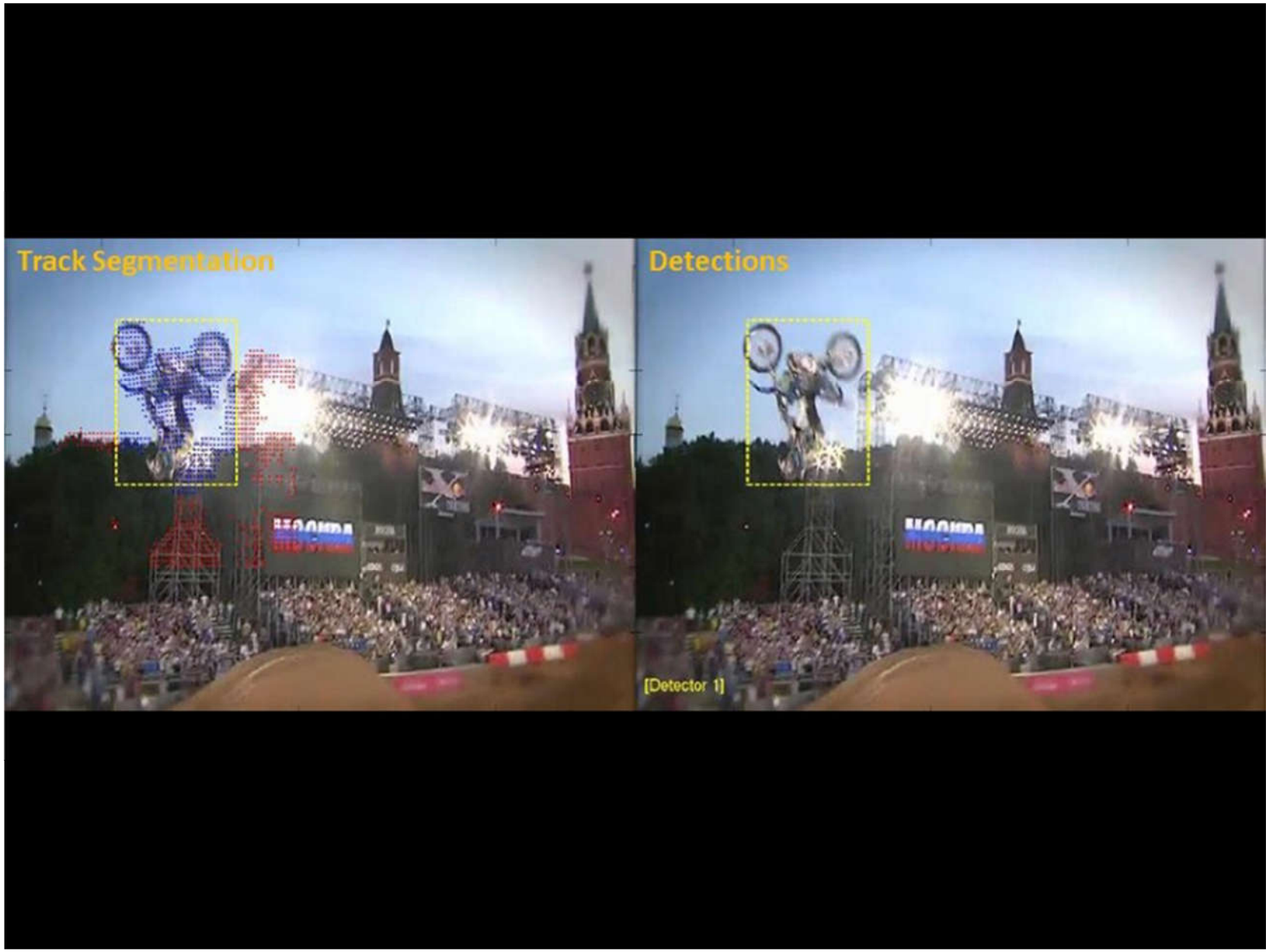
Frame 1



Frame 4

Motion Reasoning

- Video demo



Experimental Results

- Evaluation metrics [Kalal'12]
  - $F_1\ score = 2\ * precision * recall\ /\ (precision + recall)$
  - The threshold of overlap is set to 0.5

- Methods for comparison
  - Struck: Structured Output Tracking with Kernels (Struck) [Hare'11]
  - Tracking-Learning-Detection (TLD) [Kalal'12]
  - Self-paced learning for long-term tracking (SPLTT) [Supancic III'13]

Experimental Results

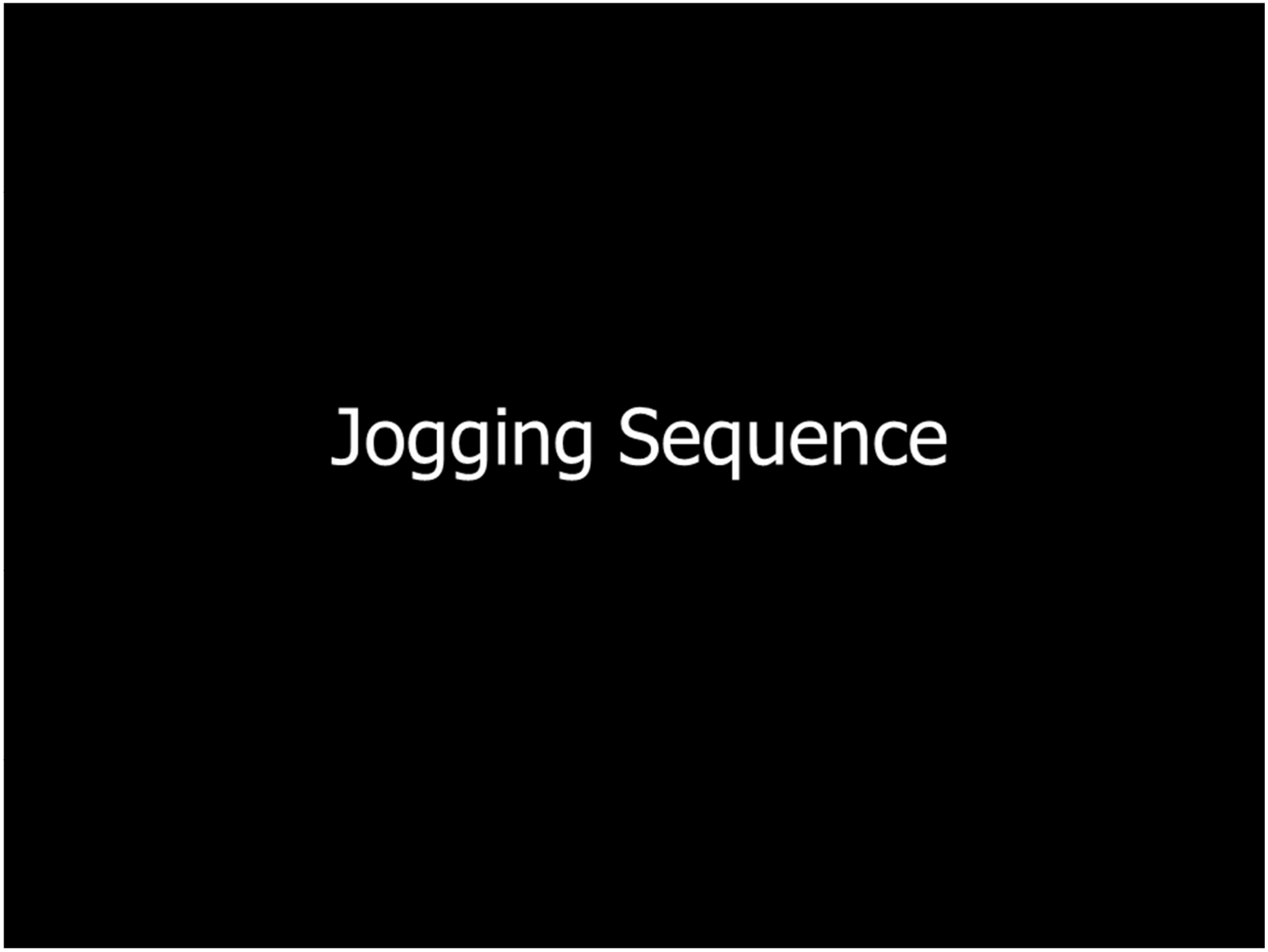
- Online Tracking Benchmark (OTB) Dataset (50 seq.) [Wu'13]
  - Overall results on 50 sequences comparing with other methods

	Struck	TLD	SPLTT	Ours
F_1 score	0.565	0.513	0.661	0.657

  - The results comparable to our ICCV 2015 tracker
    - i.e., Proposal Selection Tracker – Single Scale (PST-ss)
    - However, significantly better on sequences with occlusion

Experimental Results

- Online Tracking Benchmark (OTB) Dataset (50 seq.) [Wu'13]
  - Video demo

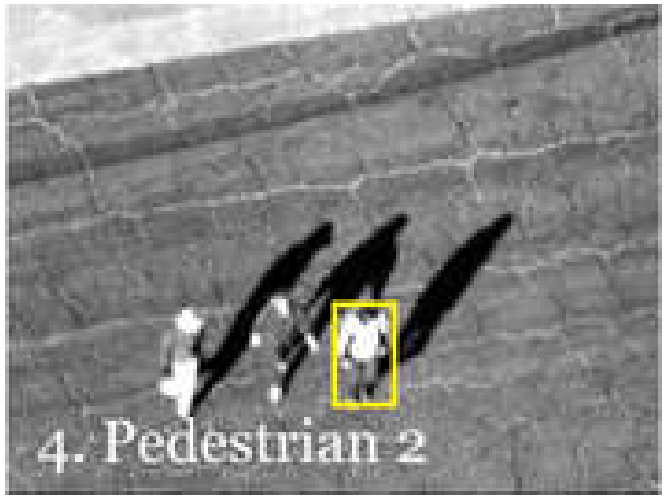


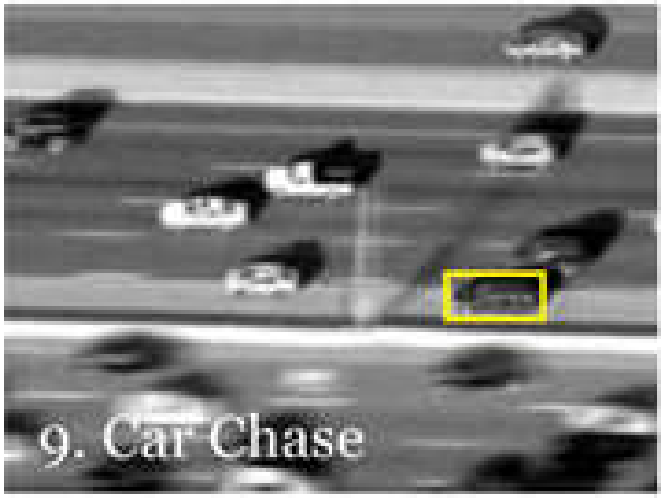
Experimental Results

- Online Tracking Benchmark (OTB) Dataset [Wu'13]
  - Ambiguity of ground truth annotation

Experimental Results

- Tracking-Learning-Detection (TLD) Dataset (3 seq.) [Kalal'12]



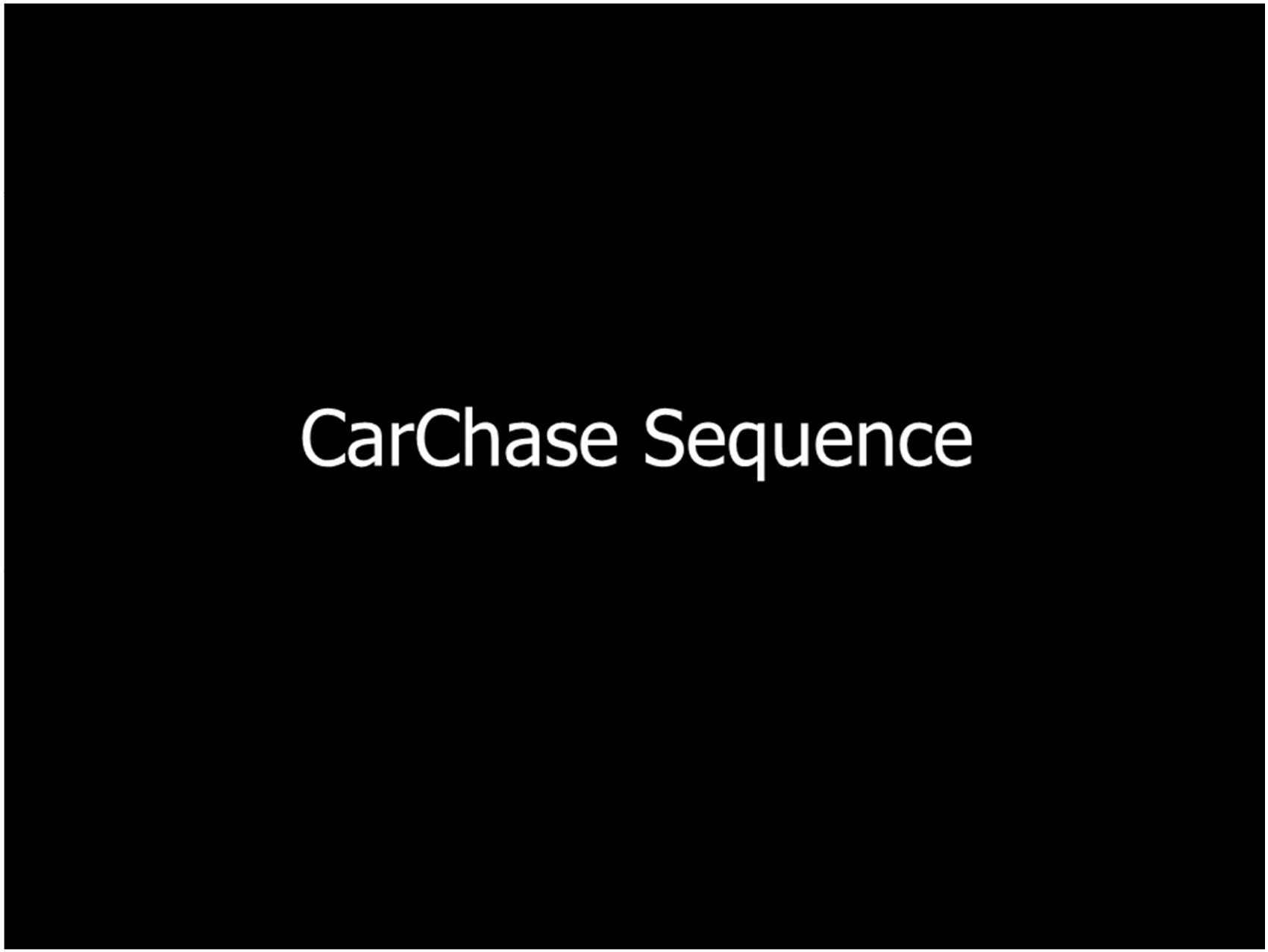


- Results on three long-term sequences

	Struck	TLD	SPLTT	Ours
Pedestrian 2	0.175	0.500	0.950	**0.979**
Pedestrian 3	0.353	0.886	0.989	**1.000**
Car Chase	0.036	**0.340**	0.290	0.312

Experimental Results

- Video demo



Summary: Occlusion and Motion Reasoning

- Proposed a principled way to identify the state of the object based on motion cues
  - Identify occlusion state via long-term track segmentation
  - Estimate change-in-viewpoint with geometric transformations
- Addressed model update problem for long-term tracking
  - Selectively update/create the object model based on the state of the object

- Publication
  - Y. Hua, K. Alahari, and C. Schmid. Occlusion and motion reasoning for long-term tracking. In ECCV, 2014

Outline

- Online Object Tracking with Proposal Selection
- Occlusion and Motion Reasoning for Long-term Tracking
- ***Conclusion and Future Work***

Conclusion

- Online Object Tracking with Proposal Selection
  - Proposed a generalized discriminative tracking-by-detection framework for short-term tracking
  - Achieved state-of-the-art performance on challenging datasets
- Occlusion and Motion Reasoning for Long-term Tracking
  - Proposed a principled way to identify the state of the object using motion cues
  - Addressed the model update problem for long-term tracking

- Publications & Award
  - Y. Hua, K. Alahari, and C. Schmid. Occlusion and motion reasoning for long-term tracking. In ECCV, 2014
  - Y. Hua, K. Alahari, and C. Schmid. Online object tracking with proposal selection. In ICCV, 2015
  - “Winning tracker” in the VOT-TIR2015 challenge

Future Work: Proposal Selection Tracking

- Handle deformable objects
  - Propose more candidates from
    - General object proposals [Zhu'15, Ren'15]
    - Deformable object trackers [Godec'11, Liu '15]

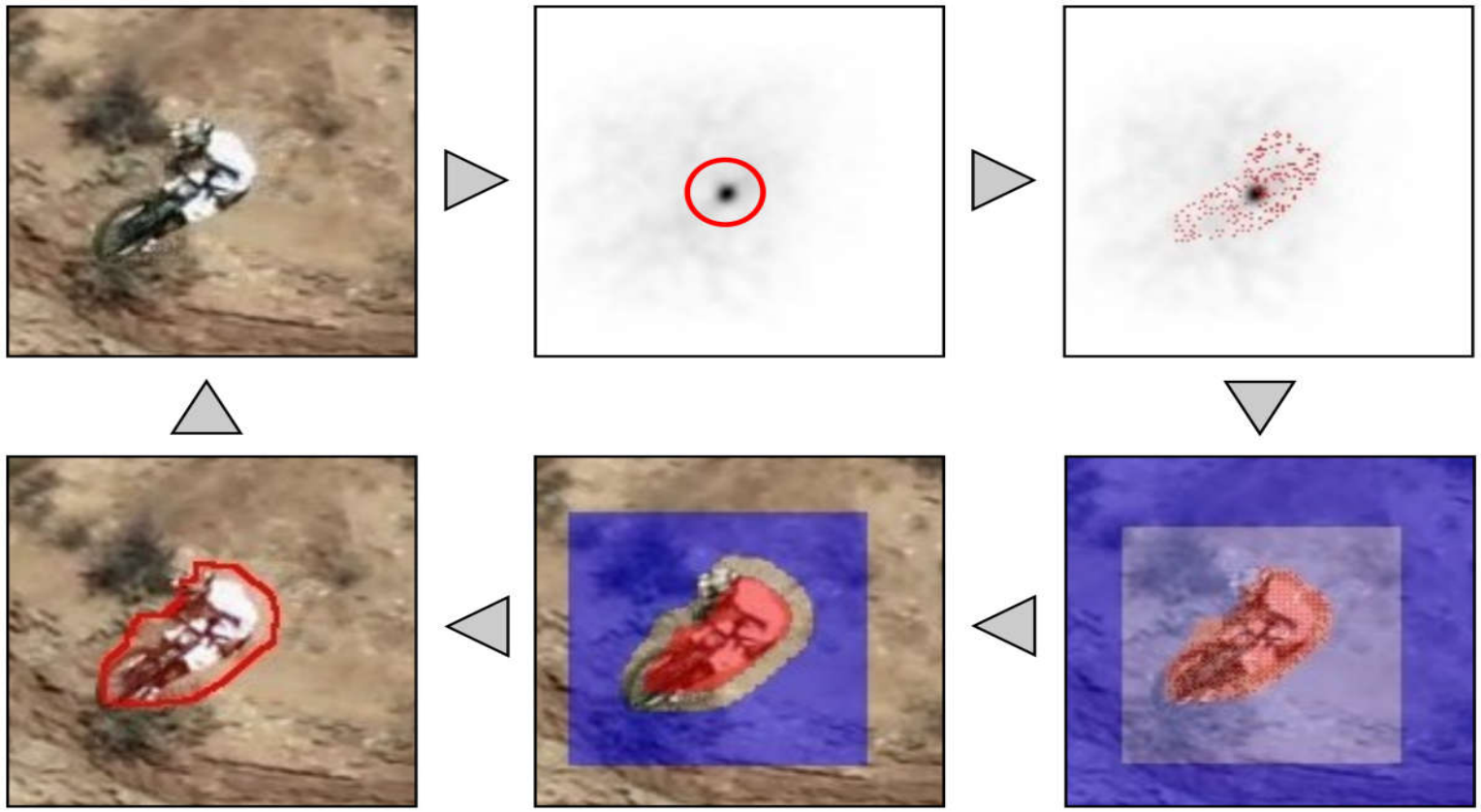


Image courtesy of [Godec'11]

Future Work: Proposal Selection Tracking

- Handle deformable objects
  - Propose more candidates from
    - General object proposals [Zhu'15, Ren'15]
    - Deformable object trackers [Godec'11, Liu '15]
    - Object segmentation [Li'13, Wen'15]

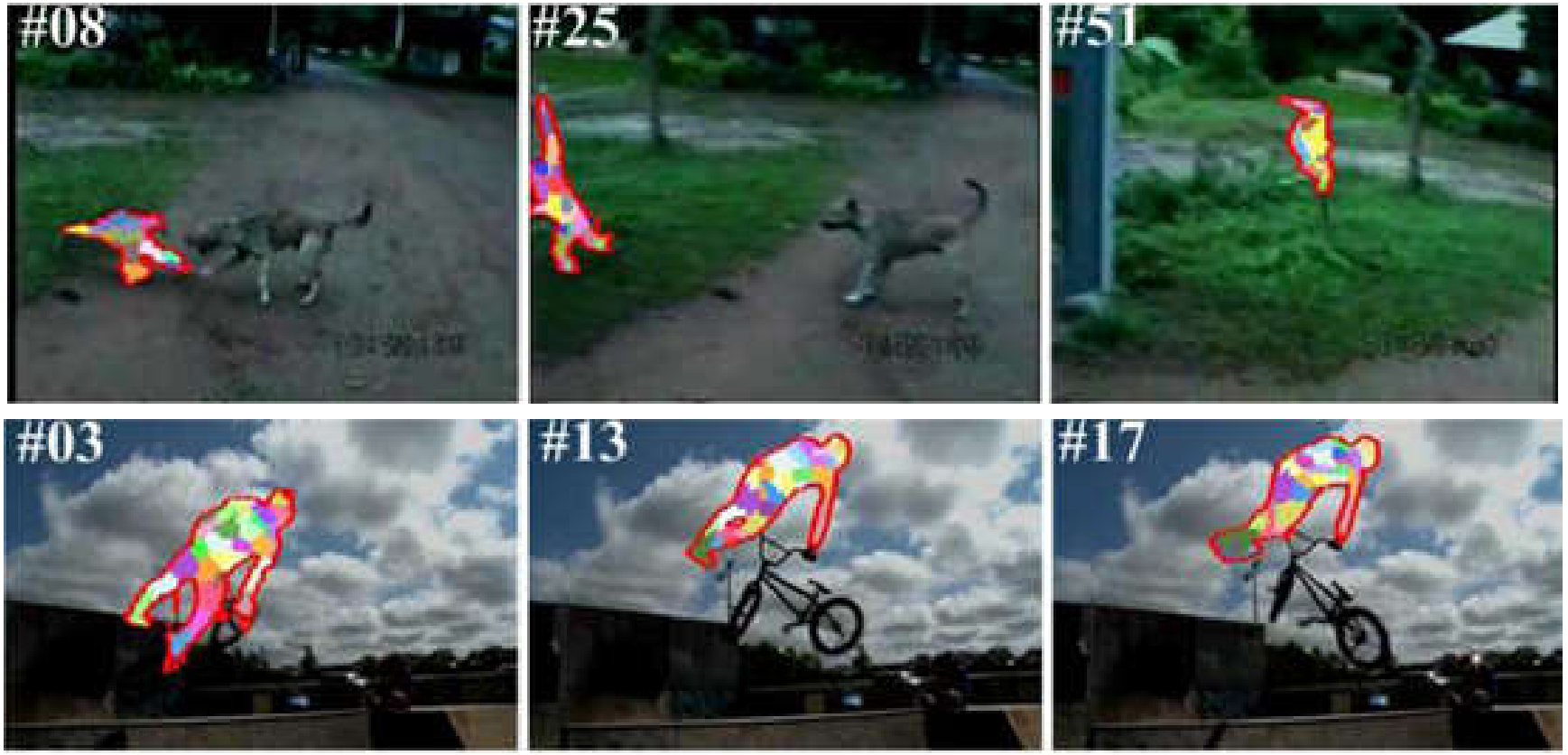


Image courtesy of [Wen'15]

Future Work: Proposal Selection Tracking

- Handle deformable objects
  - Select best candidates based on
    - Matching [Cho'14, Revaud'15]
    - Multi-hypothesis trajectory analysis [Lee'15]

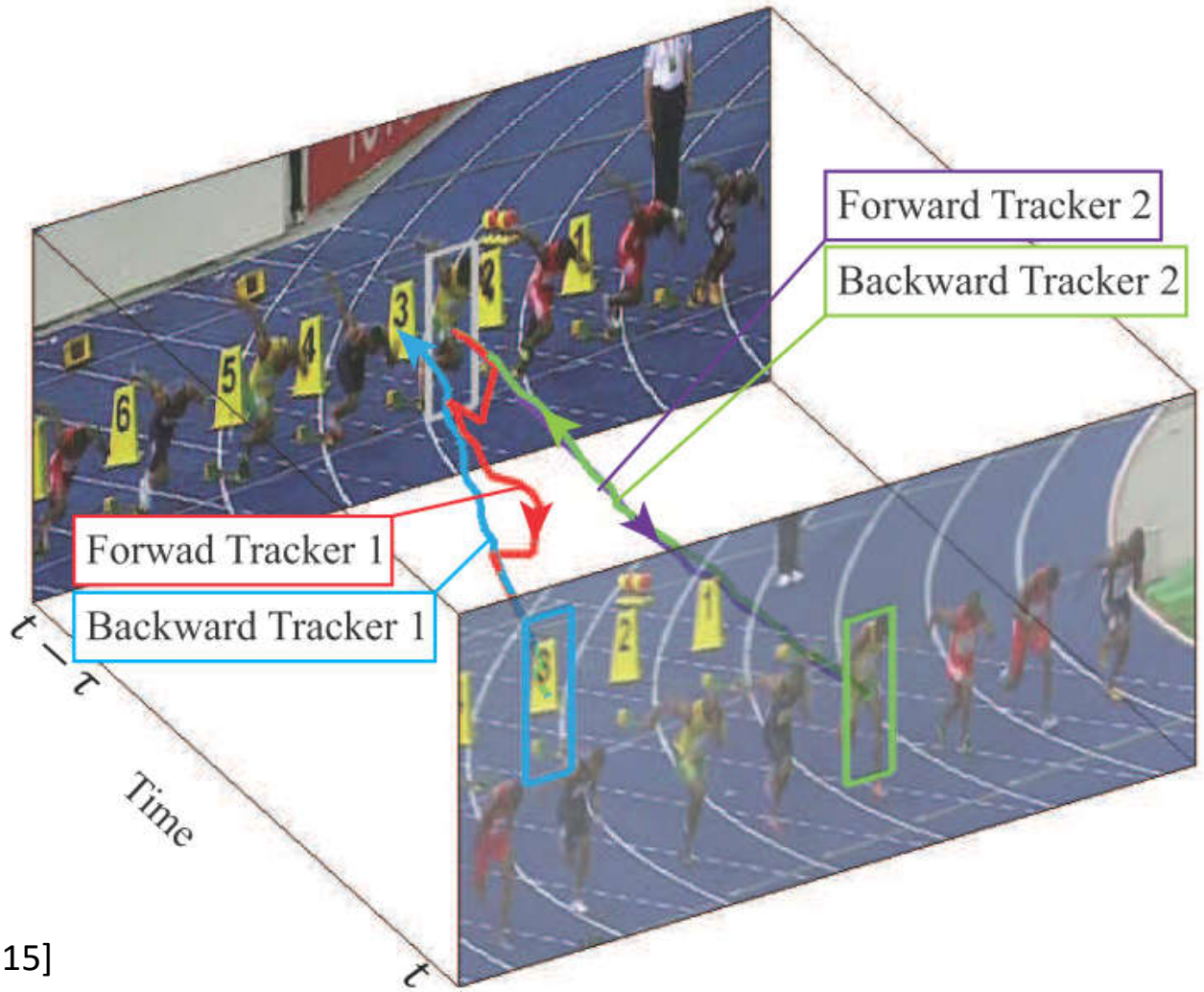


Image courtesy of [Lee'15]

Future Work: Occlusion and Motion Reasoning

- Utilize deep learning for motion representation
  - Learn motion patterns from video data
  - Learn to identify the state of the object
  - Harvest training data without violating model-free setting

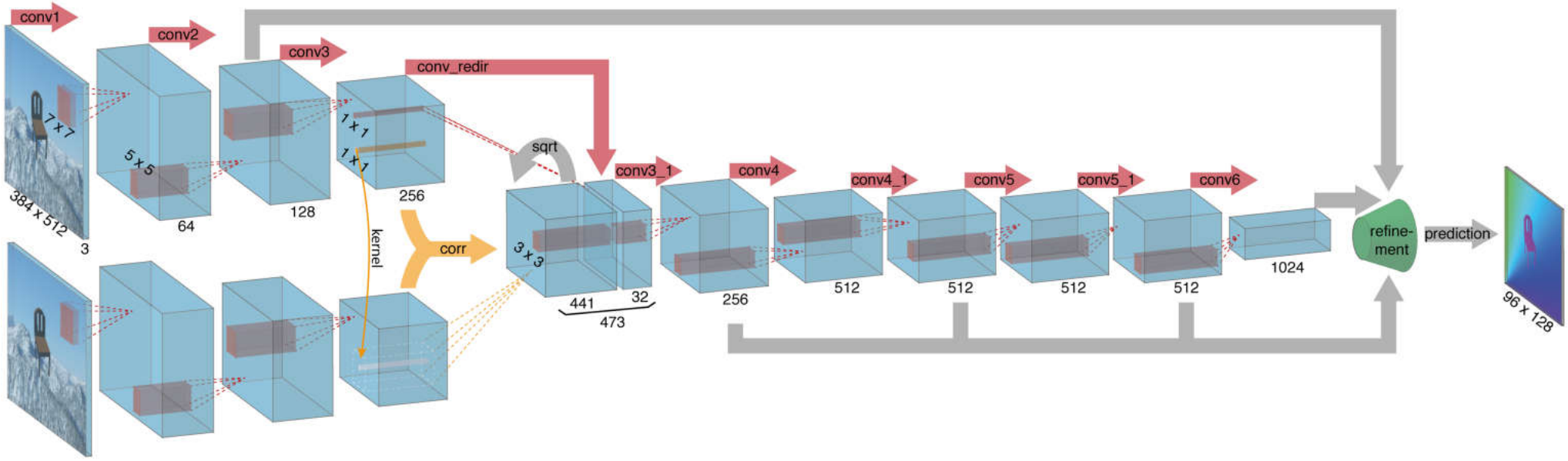


Image courtesy of [Dosovitskiy'15]

Thank You

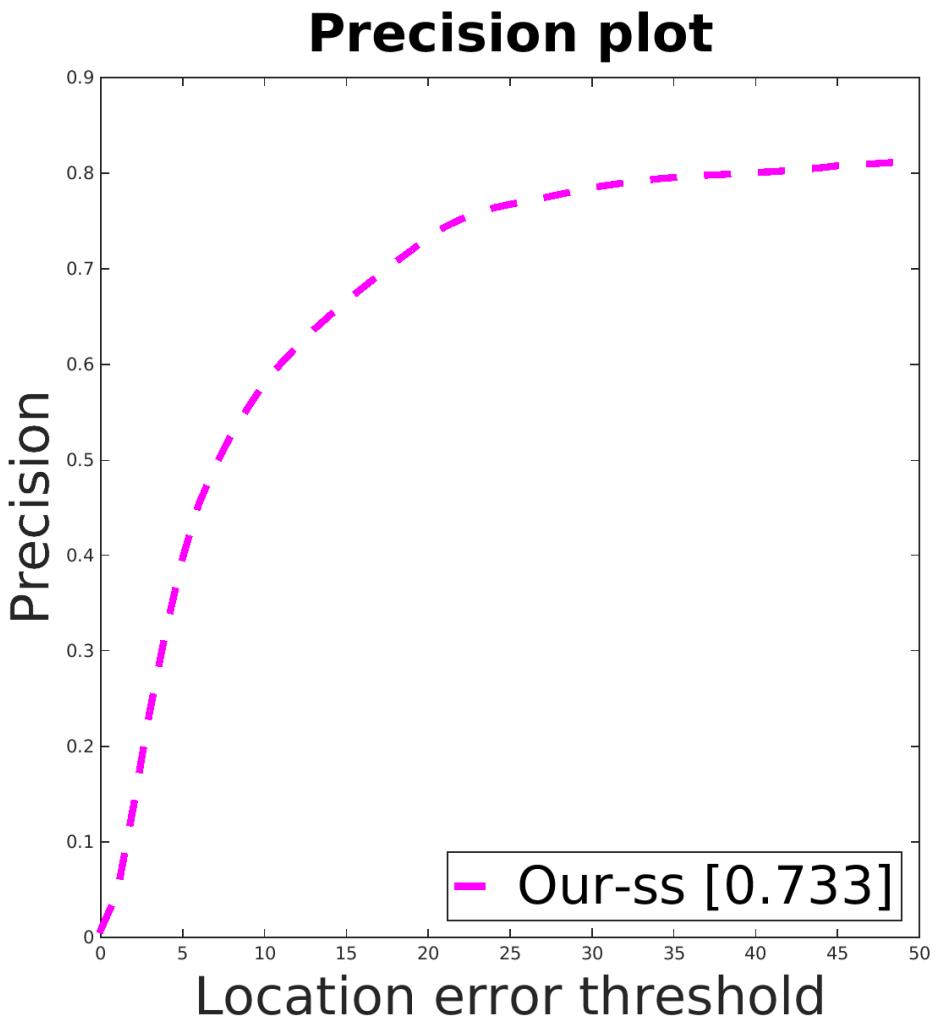
This PhD thesis is supported in part by the MSR-Inria joint project, Google Faculty Research Award, and the ERC advanced grant ALLEGRO

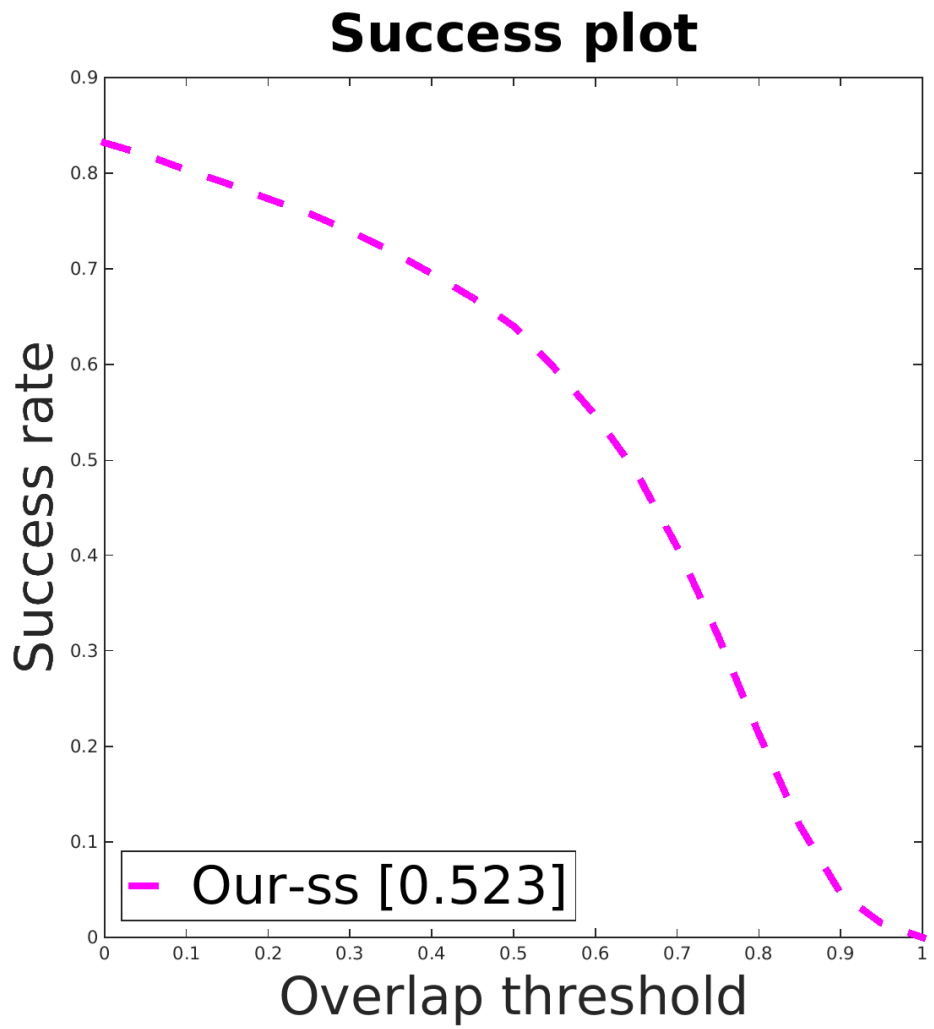
Appendix

- Experimental results: OTB (Precision & Success Plot)
- Experimental results: VOT2014
- Framework of simplified PST
- Experimental results: VOT2015
- Rank table of VOT2015 and VOT-TIR2015
- A brief review of tracking problems

Experimental Results

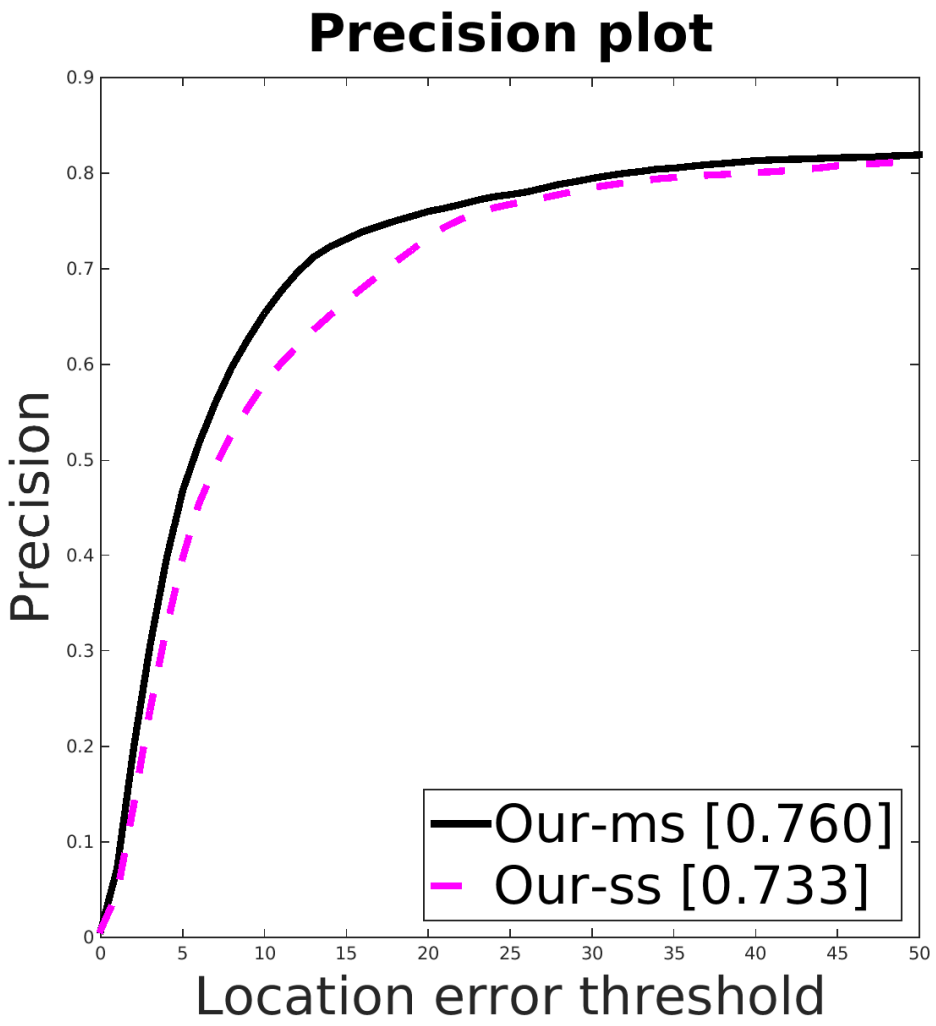
- Online Tracking Benchmark (OTB) Results: Our variants
  - Our-ss
    - Proposal: detection (single scale)
    - Selection: detection score

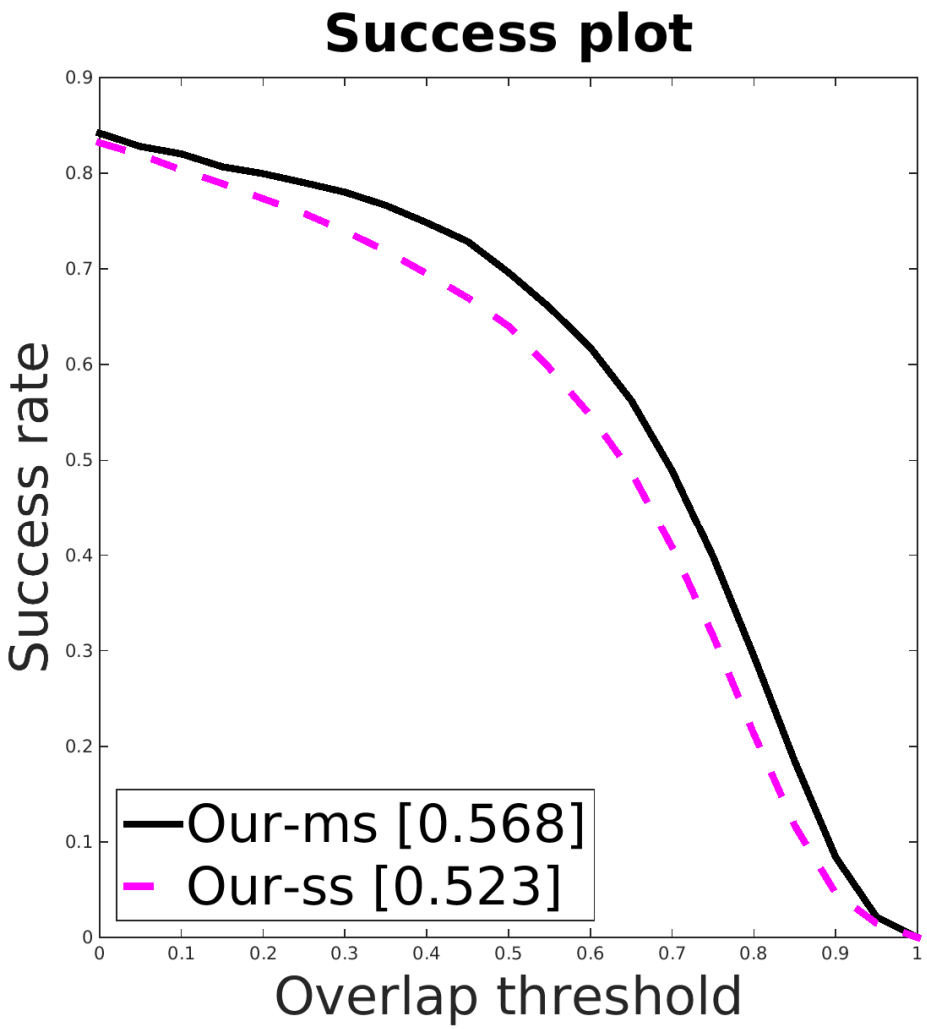




Experimental Results

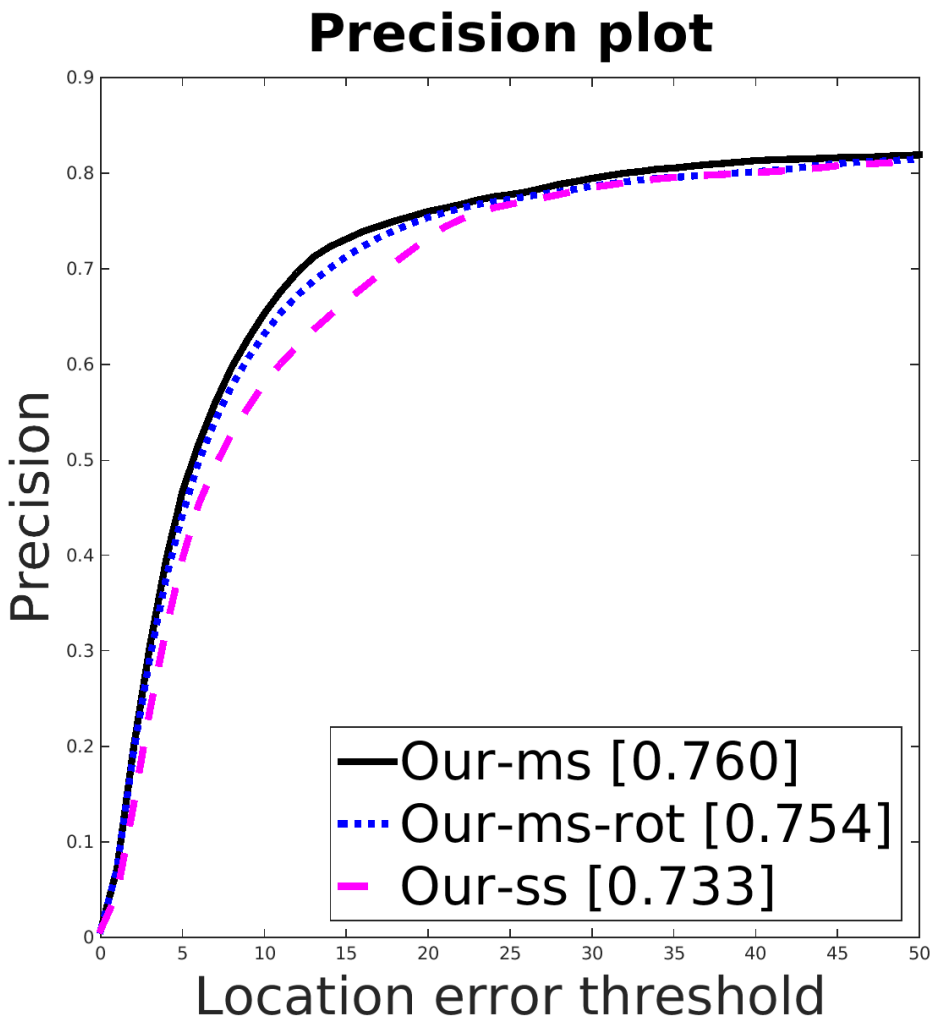
- Online Tracking Benchmark (OTB) Results: Our variants
  - Our-ms
    - Proposal: detection (multiple scales)
    - Selection: detection score

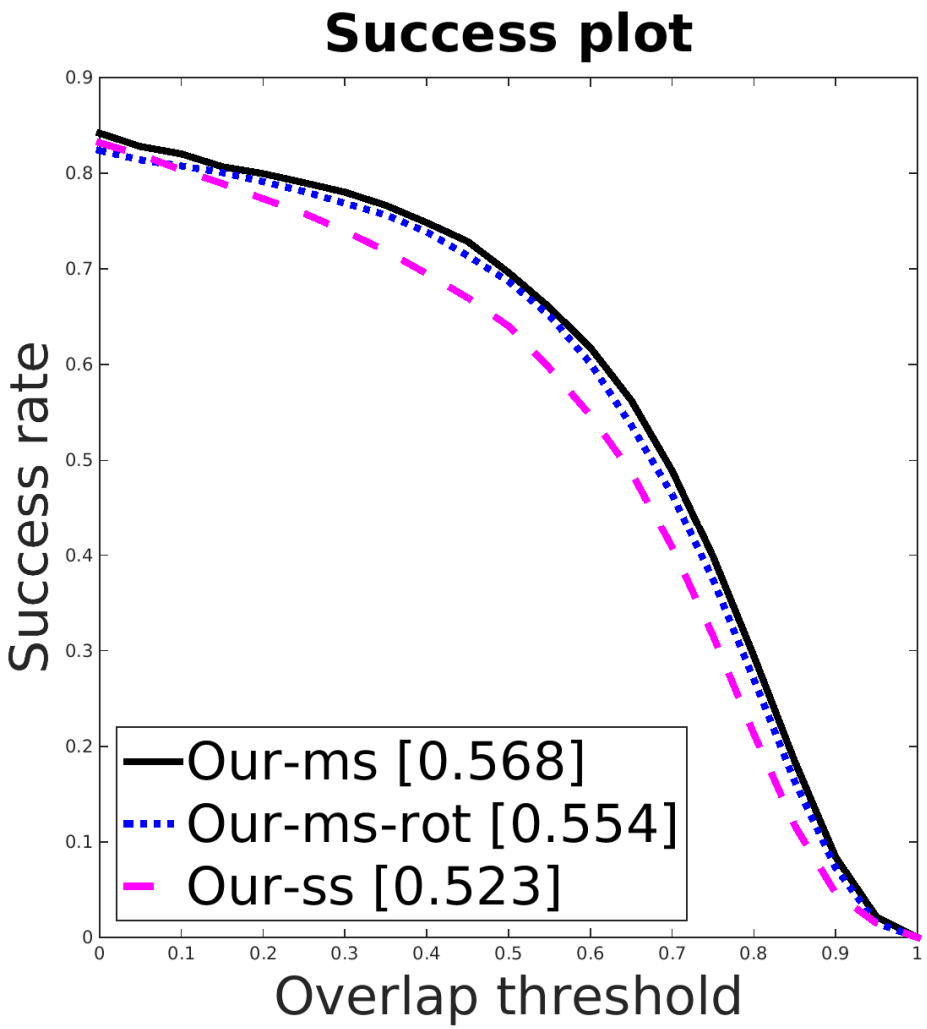




Experimental Results

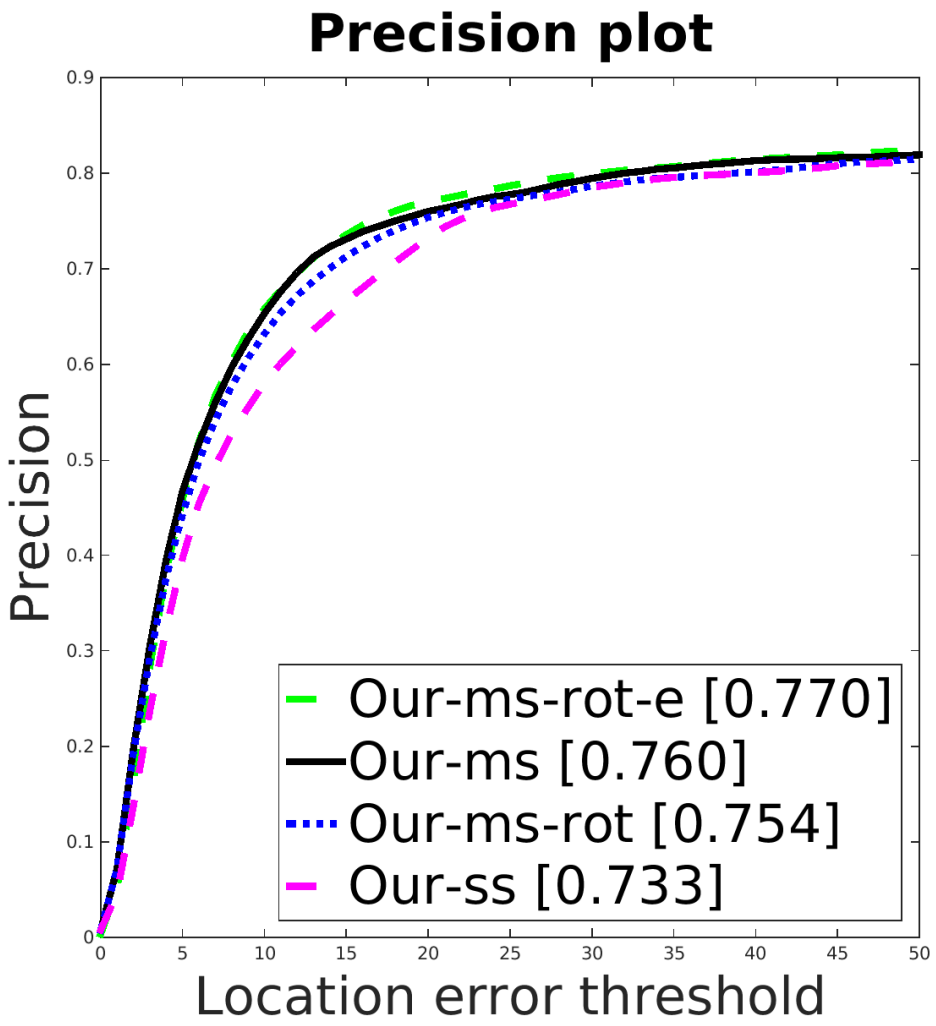
- Online Tracking Benchmark (OTB) Results: Our variants
  - Our-ms-rot
    - Proposal: detection (multiple scales) + geometry
    - Selection: detection score

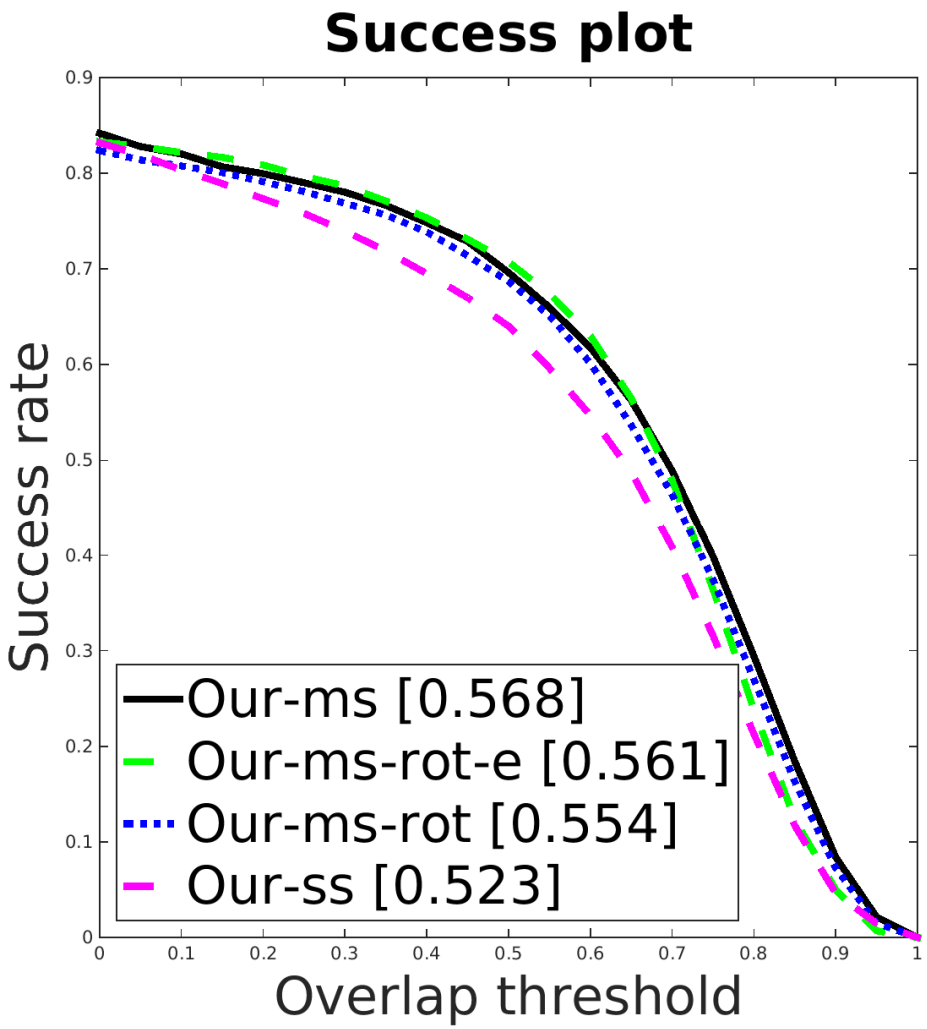




Experimental Results

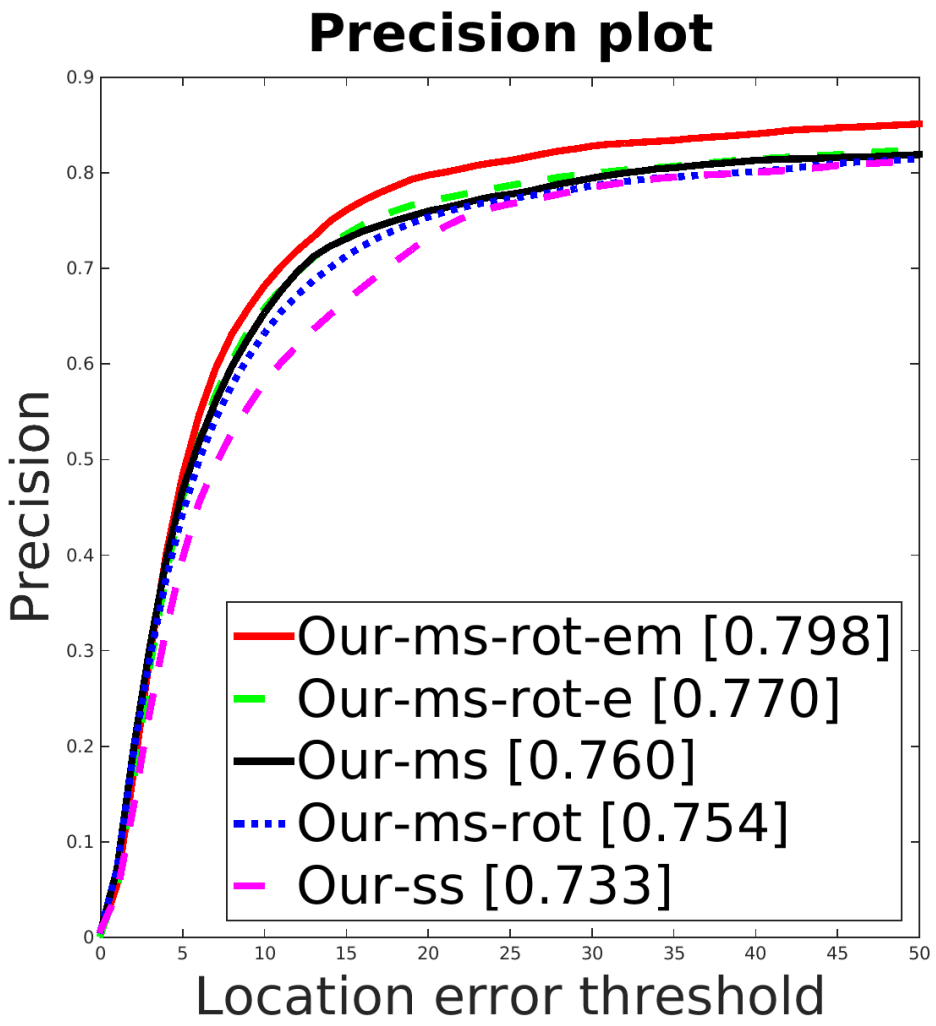
- Online Tracking Benchmark (OTB) Results: Our variants
  - Our-ms-rot-e
    - Proposal: detection (multiple scales) + geometry
    - Selection: detection + edgebox (edge response) scores

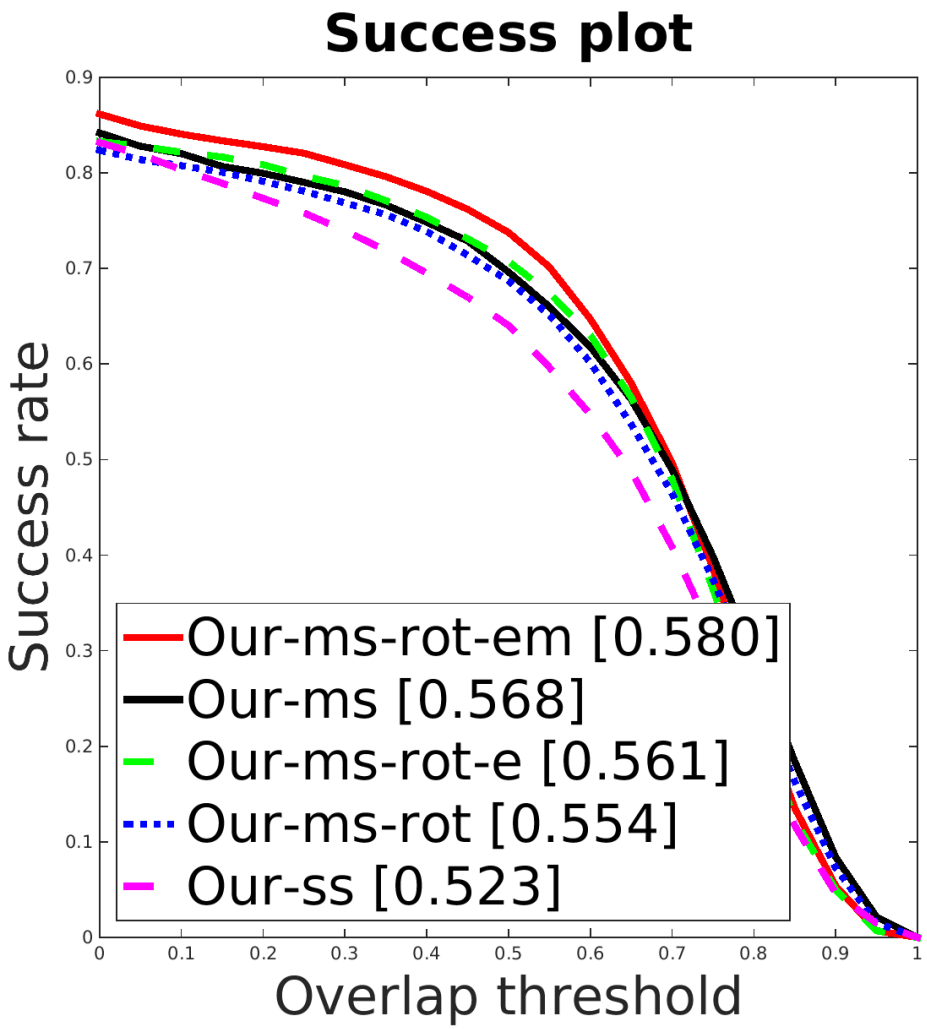




Experimental Results

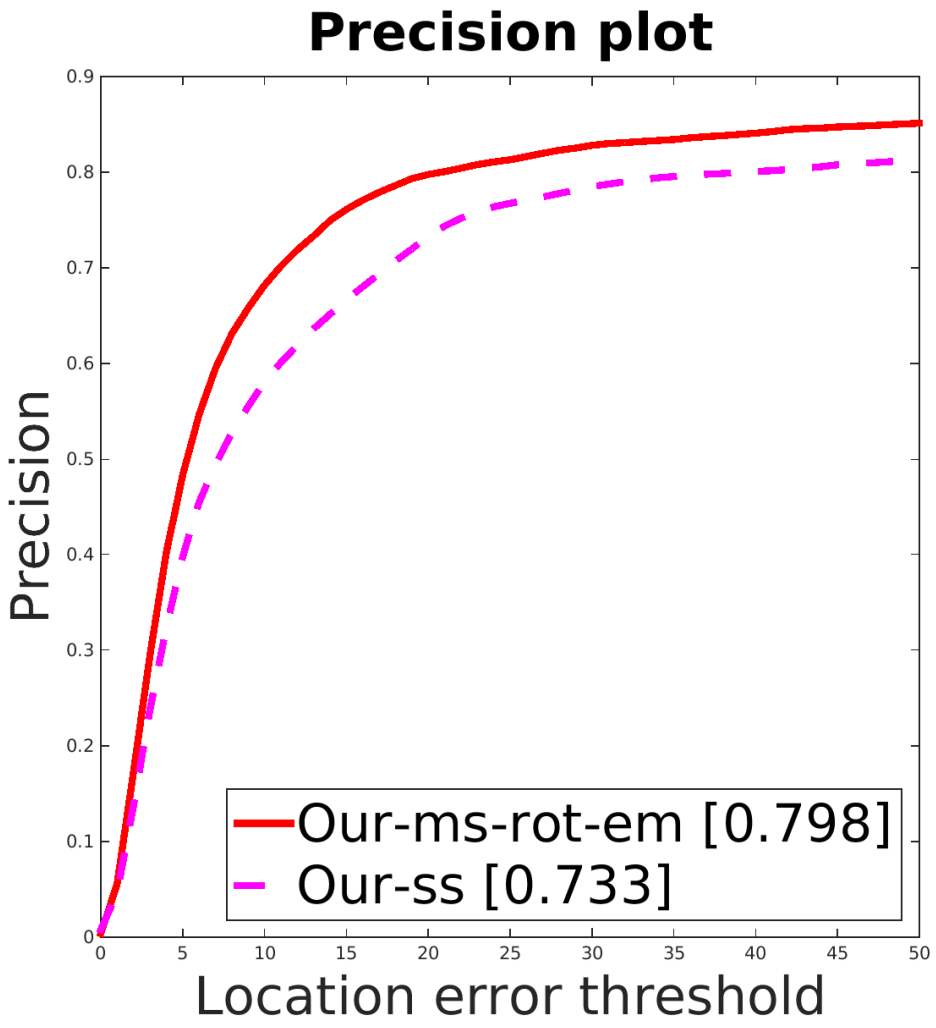
- Online Tracking Benchmark (OTB) Results: Our variants
  - Our-ms-rot-em
    - Proposal: detection (multiple scales) + geometry
    - Selection: detection + edgebox (edge & motion boundary response)

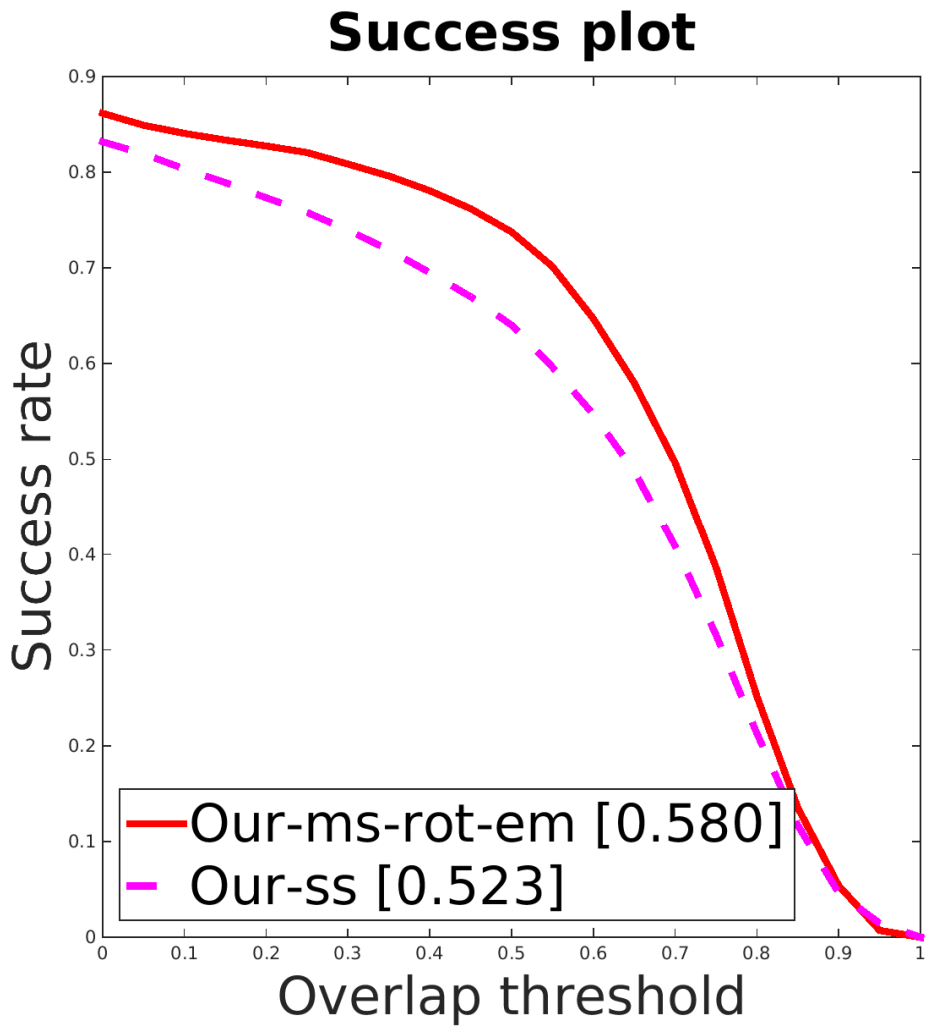




Experimental Results

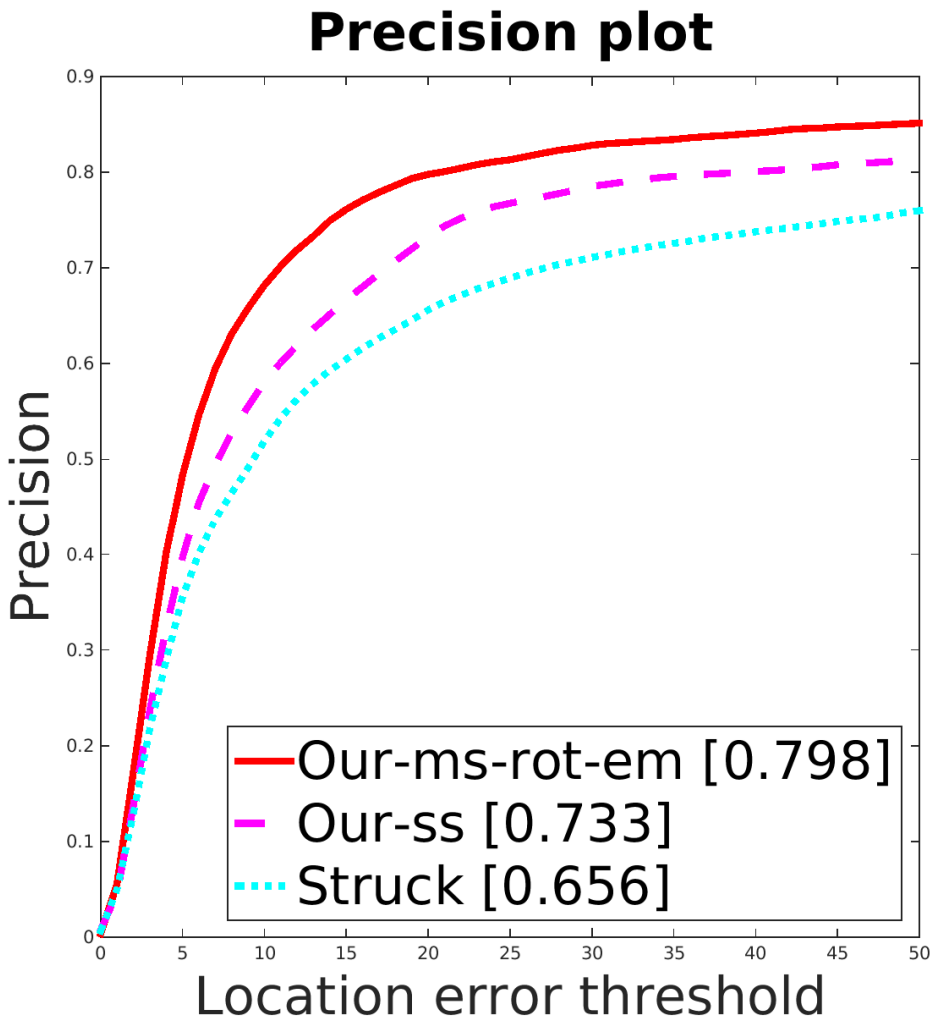
- Online Tracking Benchmark (OTB) Results: Ours vs. others

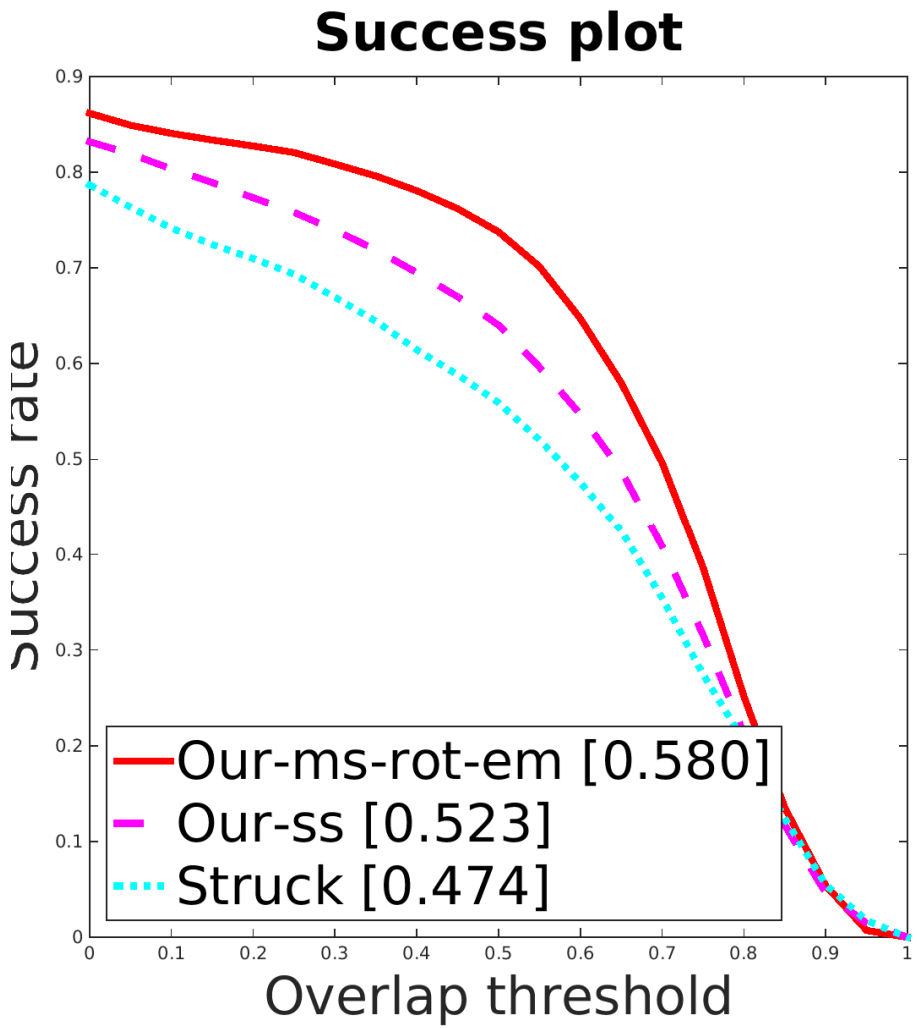




Experimental Results

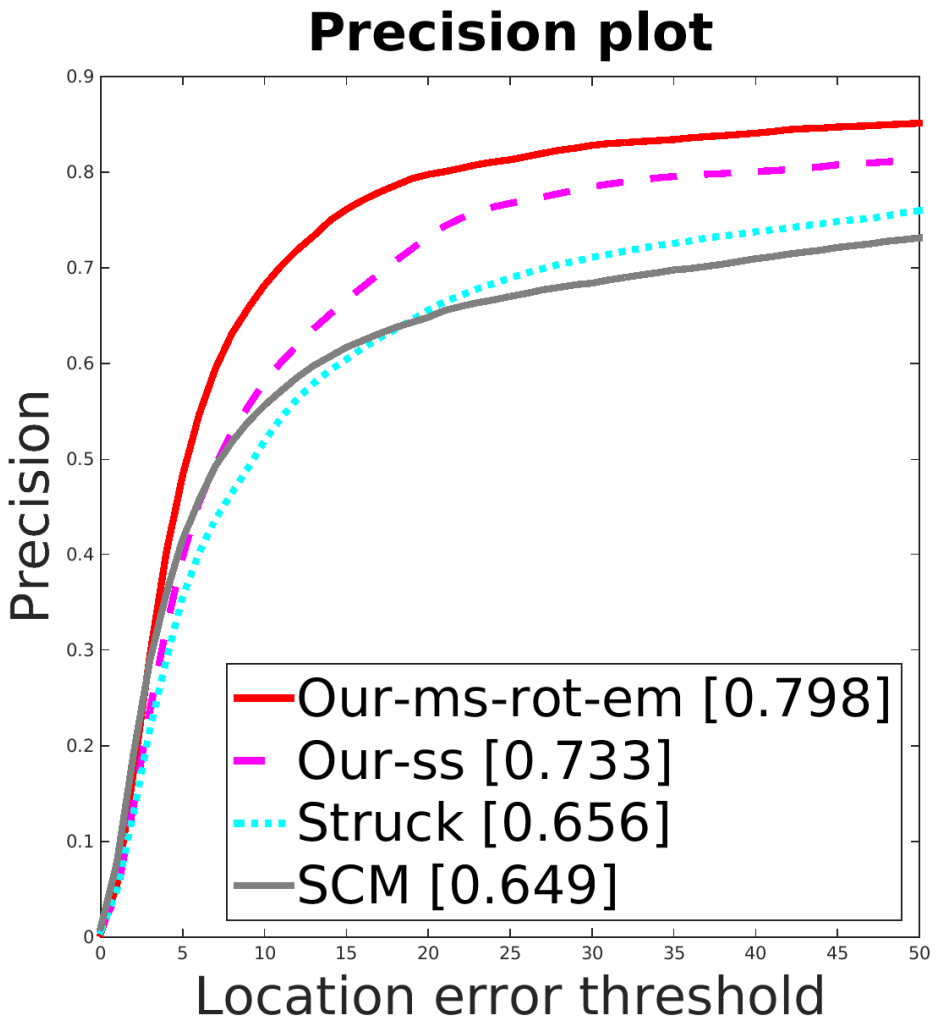
- Online Tracking Benchmark (OTB) Results: Ours vs. others

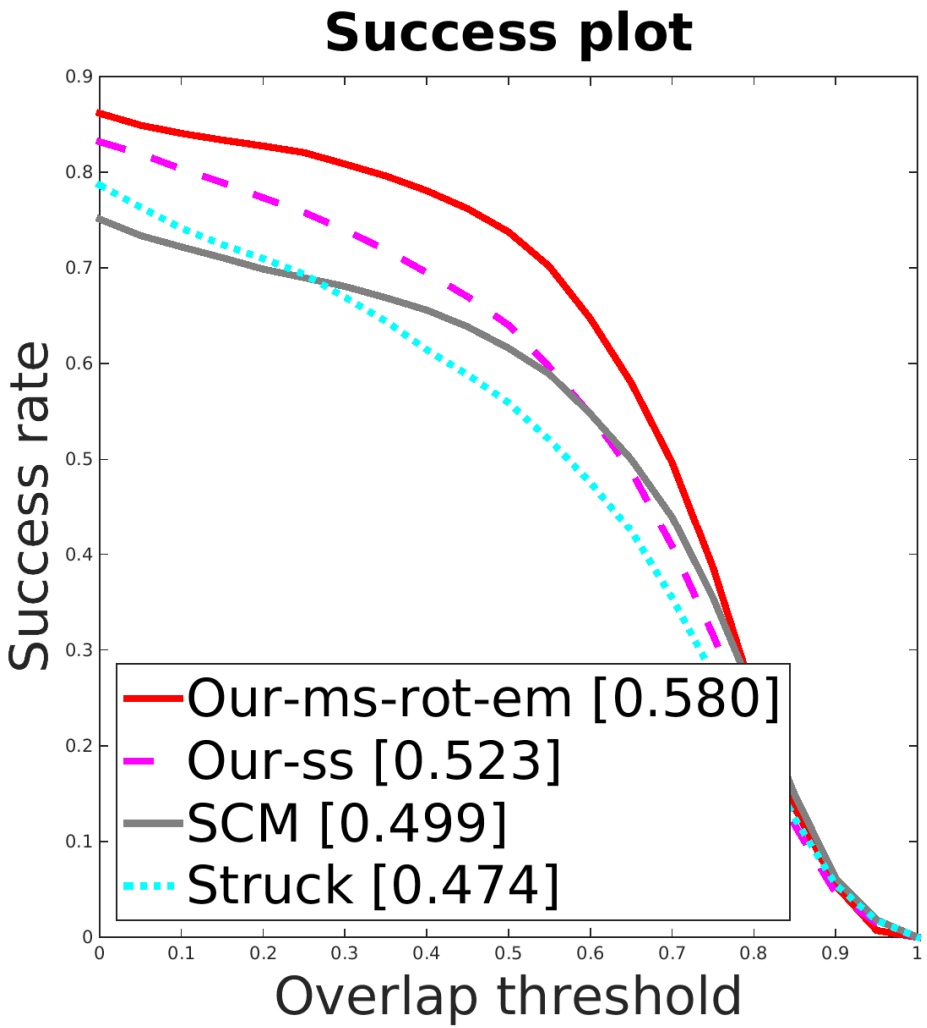




Experimental Results

- Online Tracking Benchmark (OTB) Results: Ours vs. others



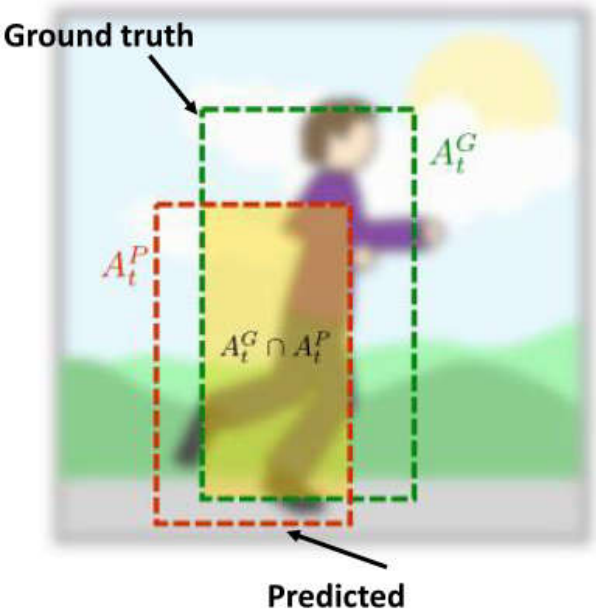


Experimental Results

- Visual Object Tracking (VOT) 2014 Challenge Dataset (25 seq.)
  [Kristan'14]

Experimental Results

- Visual Object Tracking (VOT) 2014 Challenge Evaluation Metrics
  - Accuracy: Average overlap during successful tracking
  - Robustness: Number of times a tracker drifts off the target



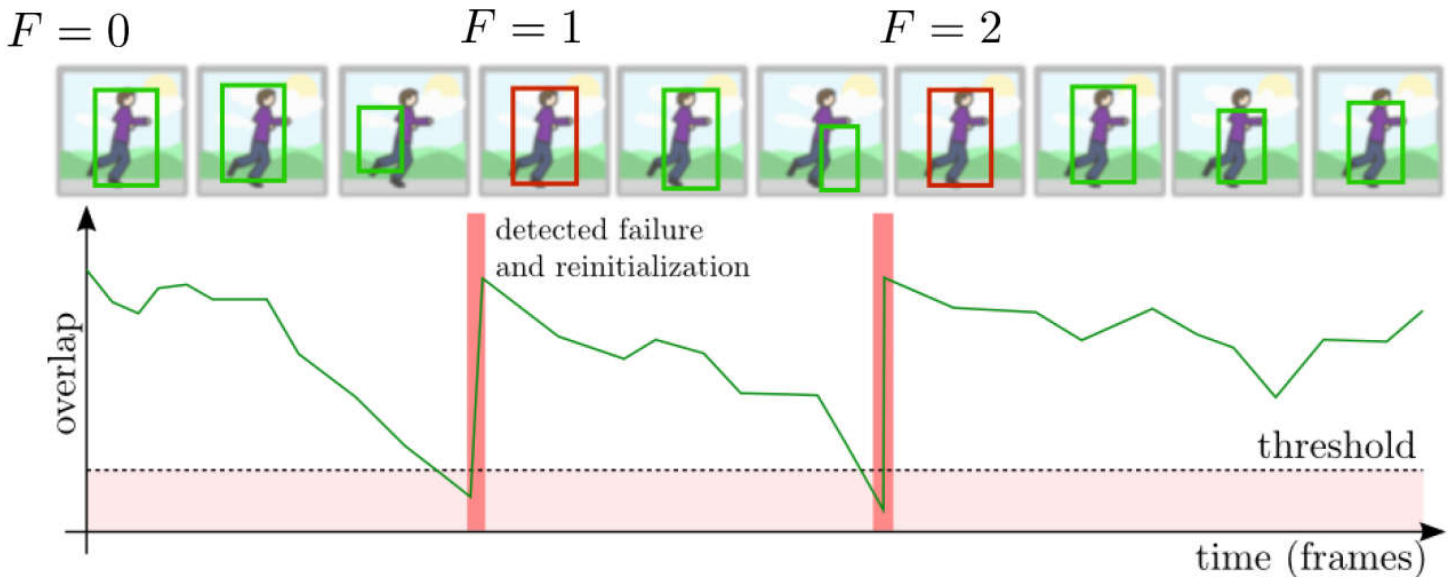
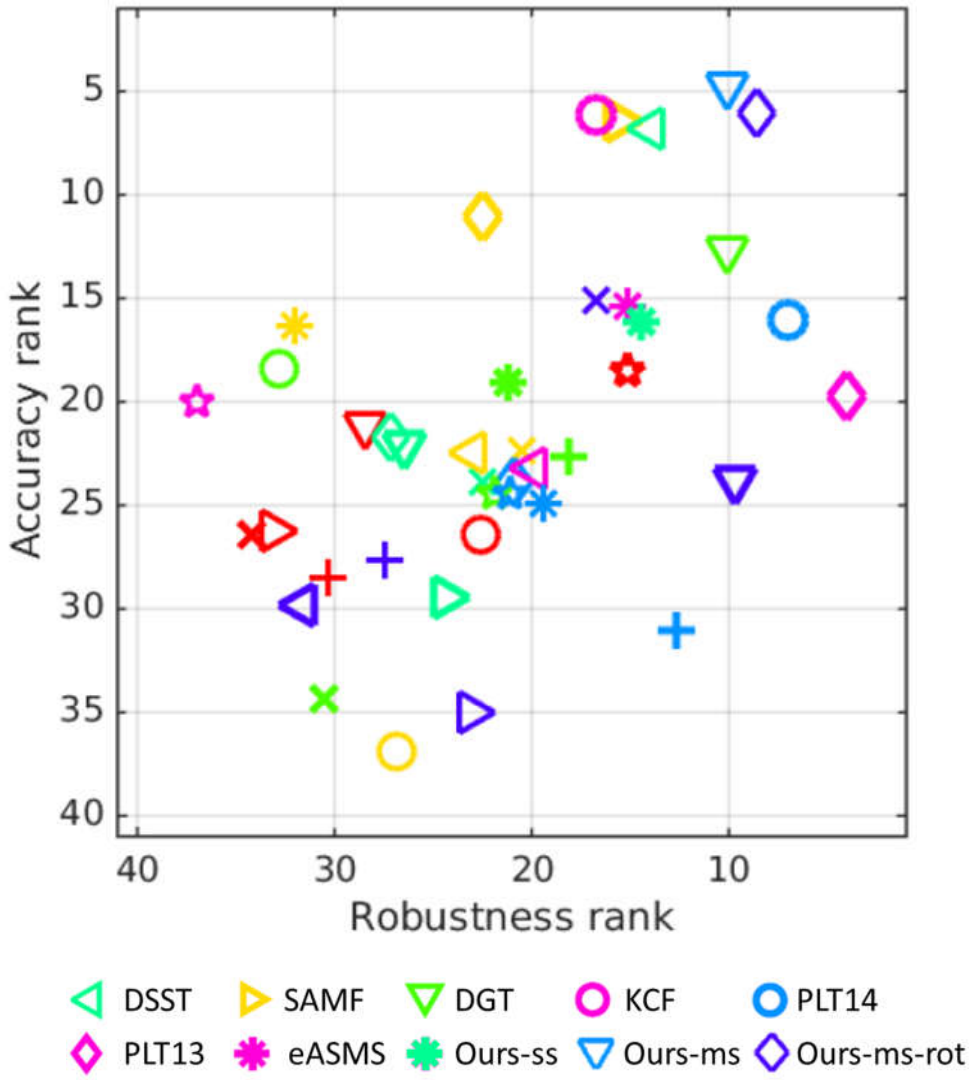


Image courtesy of [Kristan'14]

Experimental Results

- Visual Object Tracking (VOT) 2014 Challenge Results

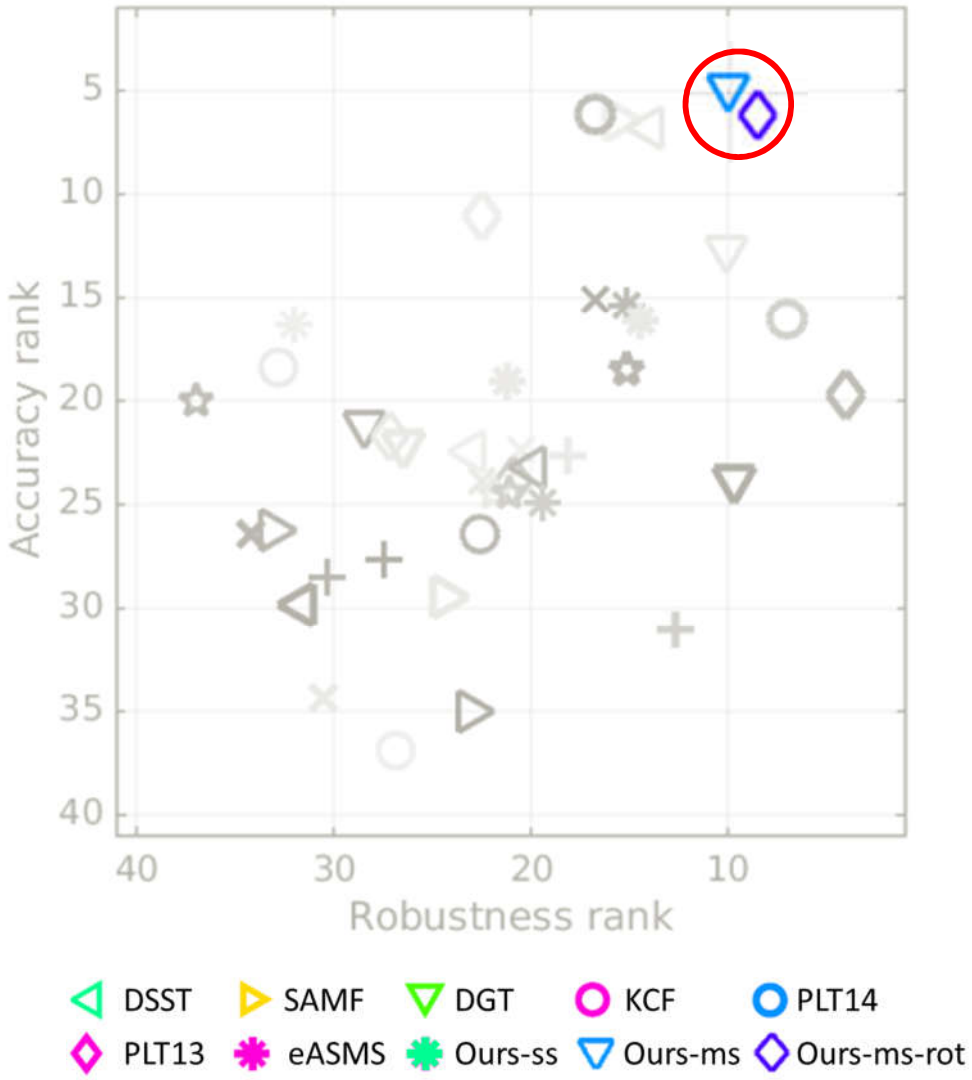
Method	Accuracy	Robust.	Average
Our-ms-rot	**6.07**	**8.58**	**7.33**
Our-ms	**4.73**	10.13	**7.43**
DSST	6.78	13.99	**10.39**
SAMF	6.46	15.65	11.06
DGT	12.67	10.13	11.40
KCF	**6.16**	16.71	11.44
PLT14	16.04	**6.98**	11.51
PLT13	19.74	**4.00**	11.87
eASMS	15.37	15.10	15.24
Our-ss	16.11	14.47	15.29



Experimental Results

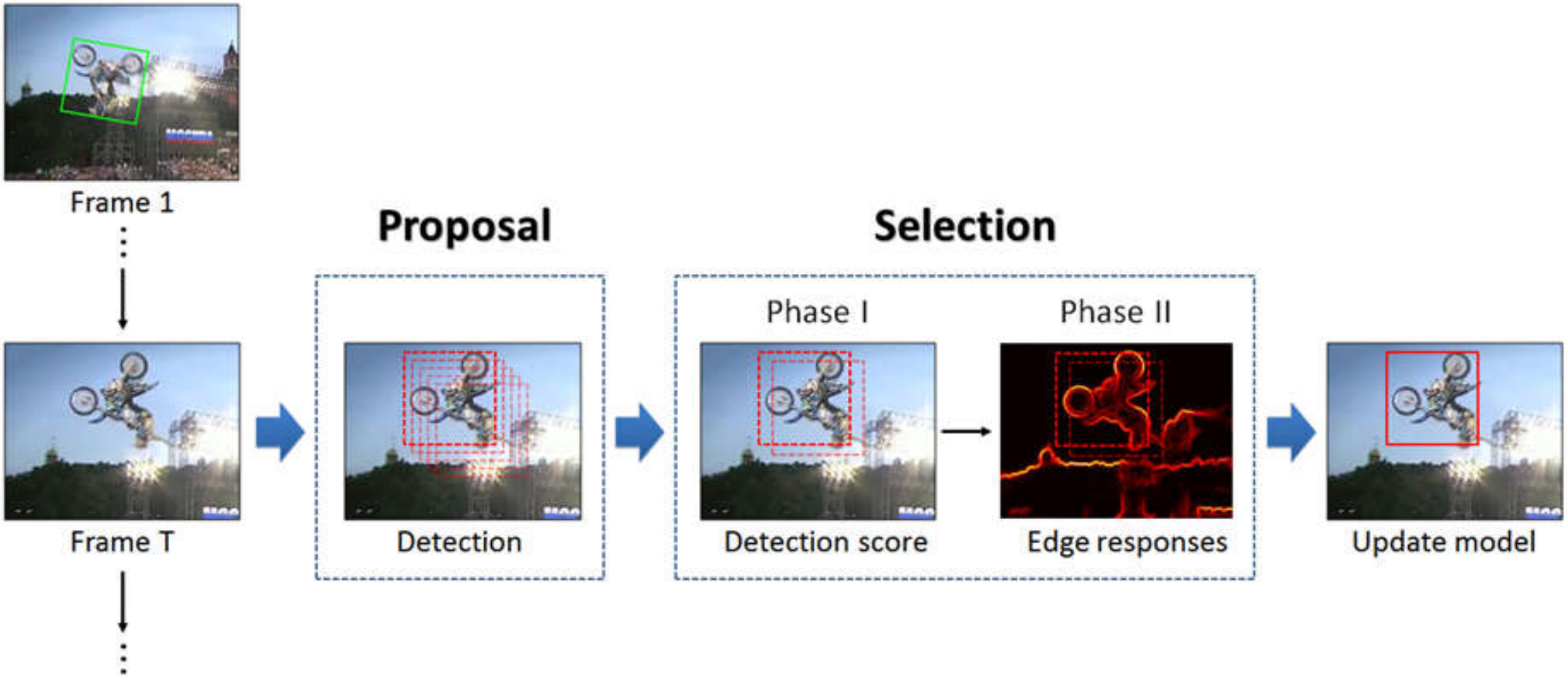
- Visual Object Tracking (VOT) 2014 Challenge Results

Method	Accuracy	Robust.	Average
Our-ms-rot	**6.07**	**8.58**	**7.33**
Our-ms	**4.73**	10.13	**7.43**
DSST	6.78	13.99	**10.39**
SAMF	6.46	15.65	11.06
DGT	12.67	10.13	11.40
KCF	**6.16**	16.71	11.44
PLT14	16.04	**6.98**	11.51
PLT13	19.74	**4.00**	11.87
eASMS	15.37	15.10	15.24
Our-ss	16.11	14.47	15.29



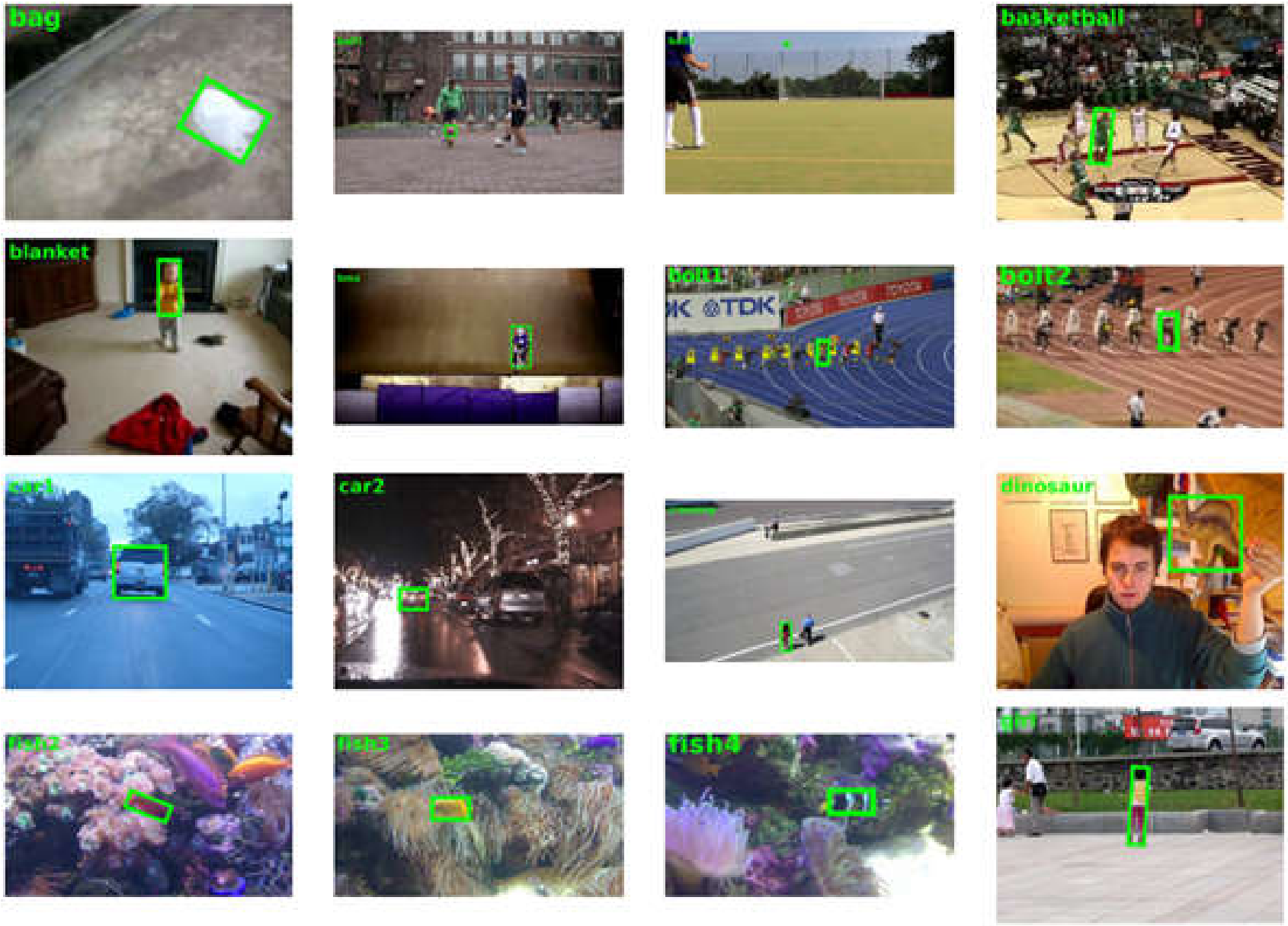
Experimental Results

- Participated in VOT competitions with a simplified Proposal Selection Tracking framework



Experimental Results

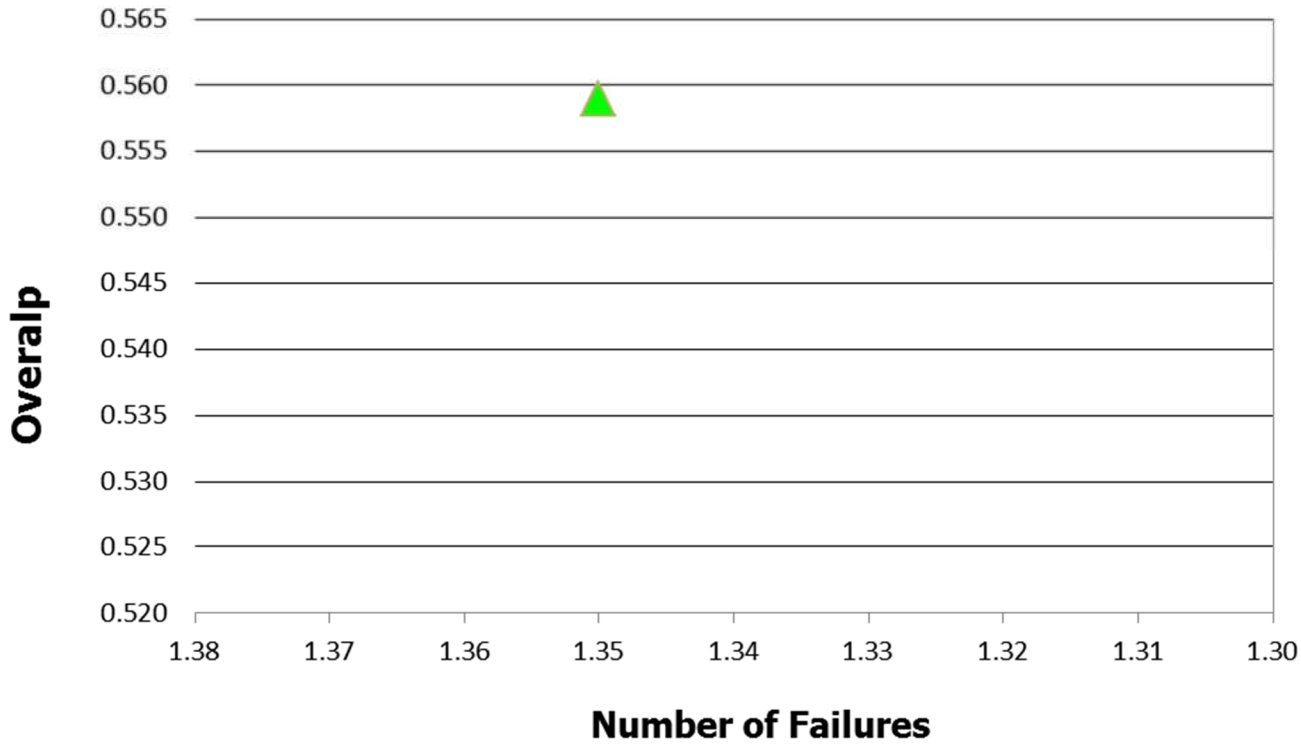
- Visual Object Tracking (VOT) 2015 Challenge Dataset (60 seq.) [Kristan'15]



Experimental Results

- Visual Object Tracking (VOT) 2015 Challenge Results

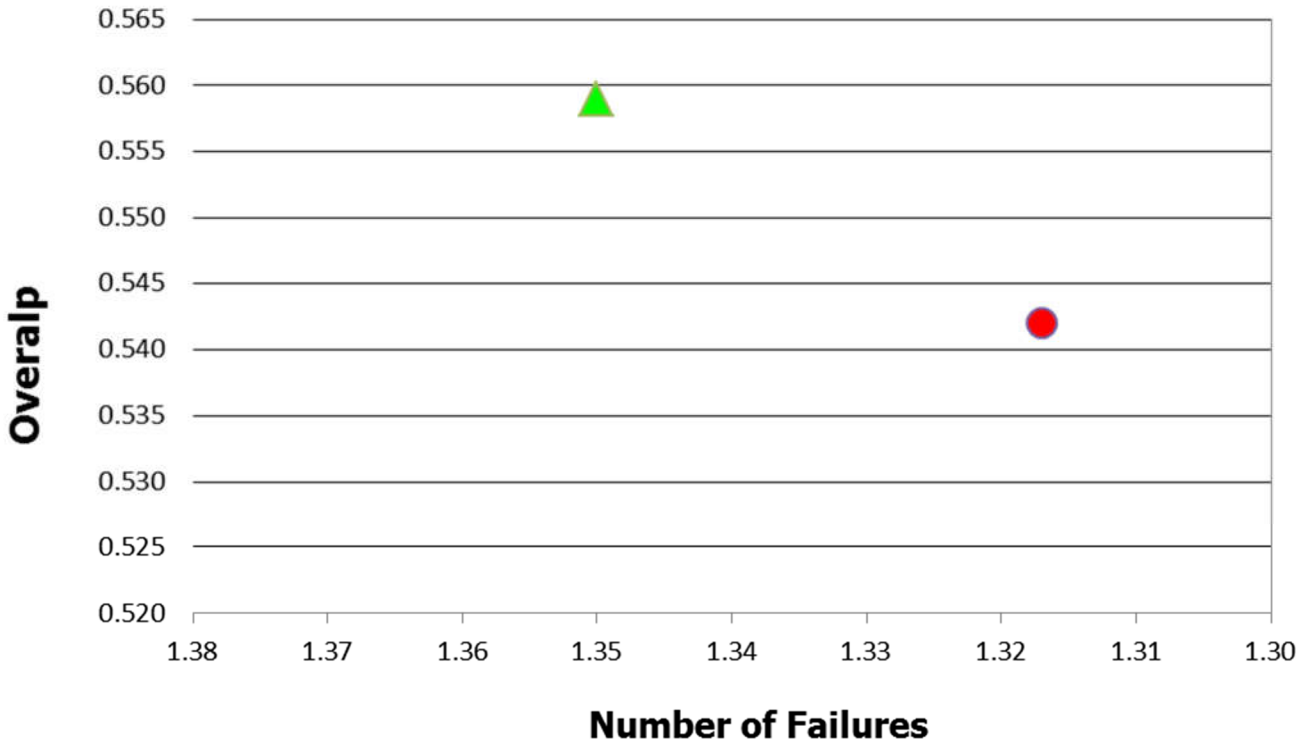
▲ Selection: detection score only	
Acc. (Overlap)	Rob. (#Failures)
0.559	1.35



Experimental Results

- Visual Object Tracking (VOT) 2015 Challenge Results

▲ Selection: detection score only		● Selection: detection + edgebox score	
Acc. (Overlap)	Rob. (#Failures)	Acc. (Overlap)	Rob. (#Failures)
0.559	1.35	0.542	**1.32**



Experimental Results

- Rank table of VOT2015 Challenge Results

Method	Accuracy	Robust.	Φ
MDNet*	**0.60**	**0.69**	**0.38**
DeepSRDCF*	0.56	**1.05**	**0.32**
EBT	0.47	**1.02**	**0.31**
SRDCF*	0.56	1.24	0.29
LDP*	0.52	1.84	0.28
sPST*	0.55	1.48	0.28
SC-EBT	0.55	1.86	0.25
NSAMF*	0.53	1.29	0.25
Struck*	0.47	1.61	0.25
RAJSSC	**0.57**	1.63	0.24

Φ: Expected average overlap

Experimental Results

- Rank table of VOT-TIR2015 Challenge Results

Method	Accuracy	Robust.	Φ
SRDCFir*	0.65	0.58	0.70
sPST*	0.66	2.18	0.64
MCCT*	0.67	3.34	0.55
EBT	0.50	3.50	0.43
CCFP*	0.63	8.55	0.36
ABCD*	0.63	5.81	0.34
Struck*	0.58	8.48	0.30

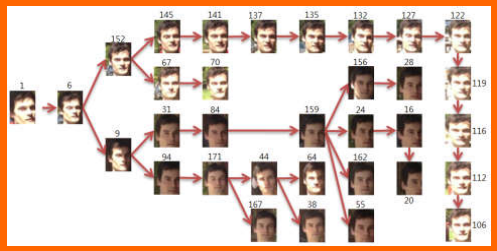
Φ: Expected average overlap

Visual Object Tracking: A brief history (problems)

Point Tracking
[Sethi and Jain, 1987, Salari and Sethi, 1990]

Multiple Cameras Tracking
[Stauffer and Tieu, 2003]

Offline/Non-causal Tracking
[Black and Jepson, 1998, Hong and Han, 2014]



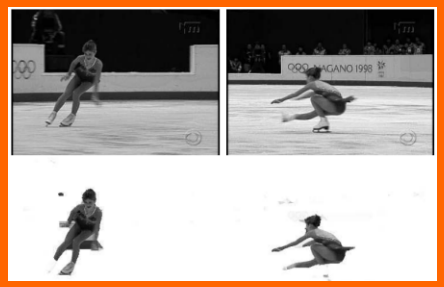
RGBD Tracking
[Song and Xiao, 2013]

1950s ▶ Now

Model-Specific Tracking
[Wren et al., 1997]

Multiple objects Tracking
[Bose et al., 2007]

Non-rigid Object Tracking
[Ren and Malik, 2007]



Long-term Tracking
[Kalal, Mikolajczyk and Matas, 2012]