Object Tracking

Samreen Anjum & Danna Gurari

The University of Texas at Austin

Fall 2019

https://www.ischool.utexas.edu/~dannag/Courses/CrowdsourcingForCV/CourseContent.html

Review

- Last week:
  - Video classification and localization applications
  - Evaluating video classification and localization
  - Crowdsourcing video classification and localization

- Assignments (Class Website & Canvas)
  - Project proposal due yesterday
  - Project outline due in two weeks

- Questions?

Today's Topics

- Object tracking applications
- Evaluating object tracking methods
- Crowdsourcing object tracking strategies
- Object tracking datasets & challenges
- Lab: video annotation & creating figures/tables in latex

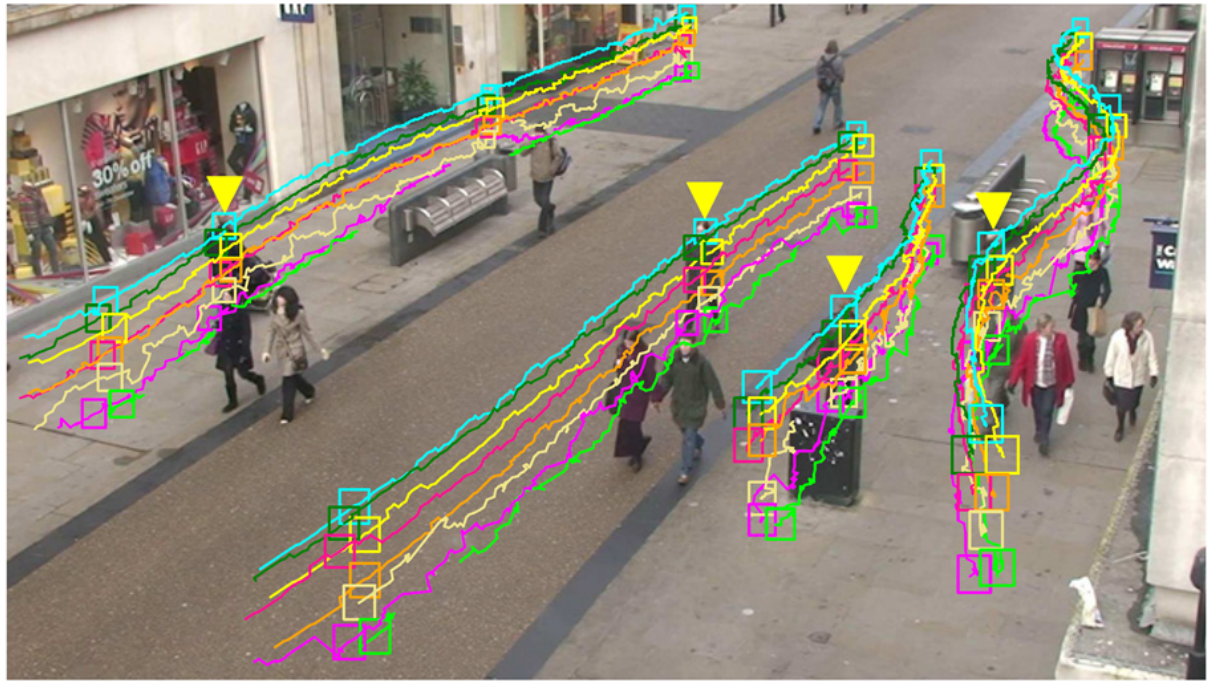
Today's Topics

- Object tracking applications
- Evaluating object tracking methods
- Crowdsourcing object tracking strategies
- Object tracking datasets & challenges
- Lab: video annotation & creating figures/tables in latex

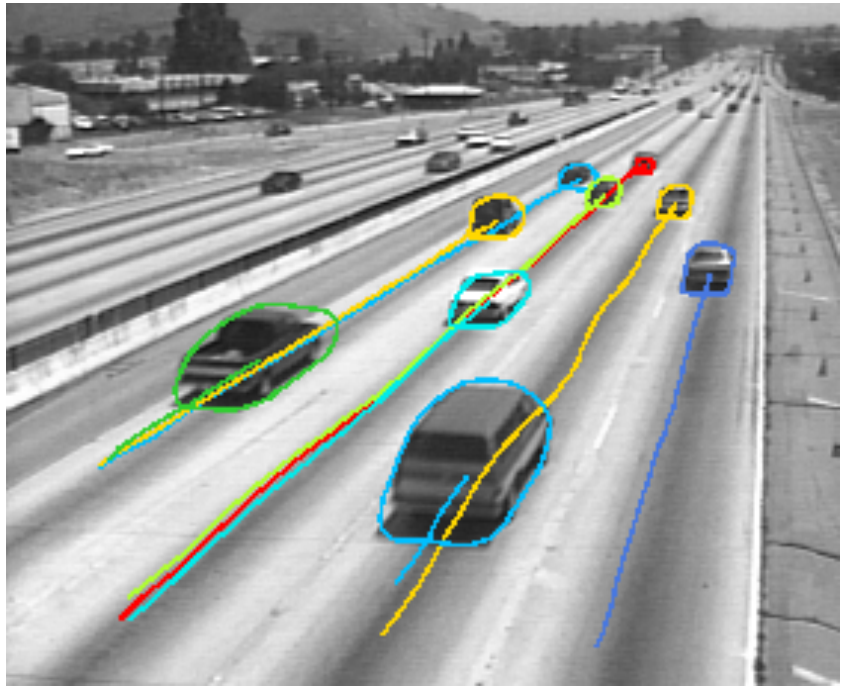
Definition

- Each object in a video is localized (object detection) and followed as it moves over time (tracking)

Application: Surveillance



People Tracking

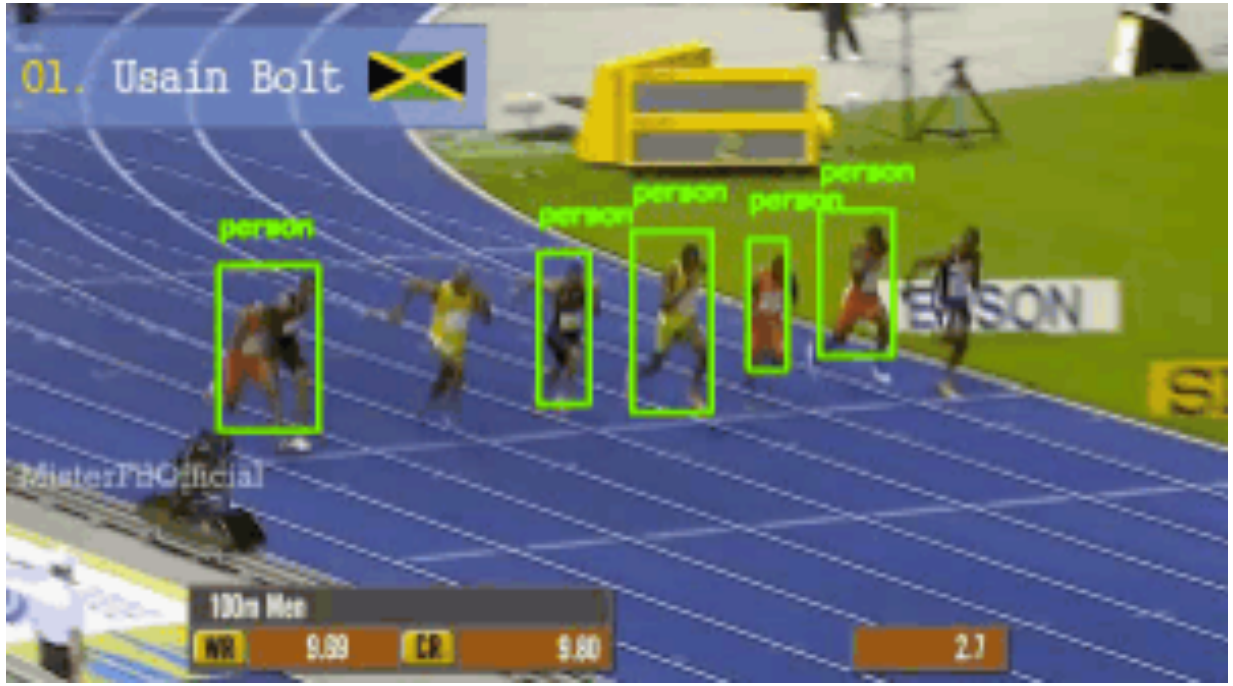


Traffic Control

http://www.vision.caltech.edu/koller/MOU-83.html

https://heartbeat.fritz.ai/the-5-computer-vision-techniques-that-will-change-how-you-see-the-world-1ee19334354b

Application: Sports Analysis



https://www.pyimagesearch.com/2018/10/29/multi-object-tracking-with-dlib/

https://www.pyimagesearch.com/2018/08/06/tracking-multiple-objects-with-opencv/

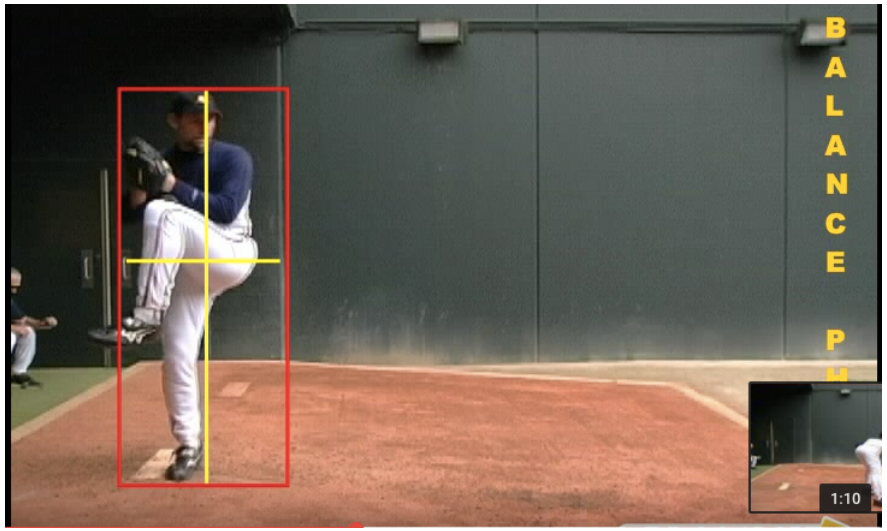
Application: Performance Analytics

Calculate Bat speed from video!

NEW! Track Bowling Ball Path!

Works great for putting!

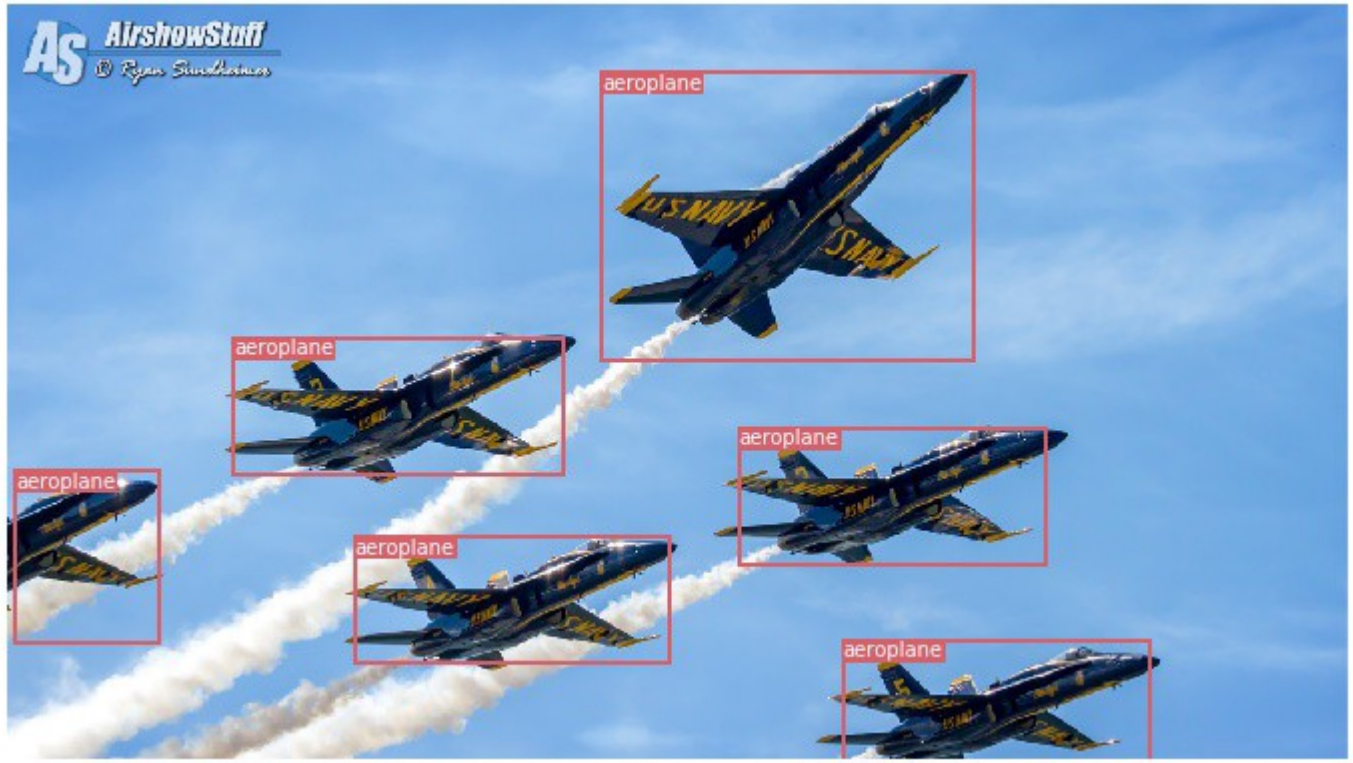
http://www.motionprosoftware.com/



Watch: Dartfish Analysis

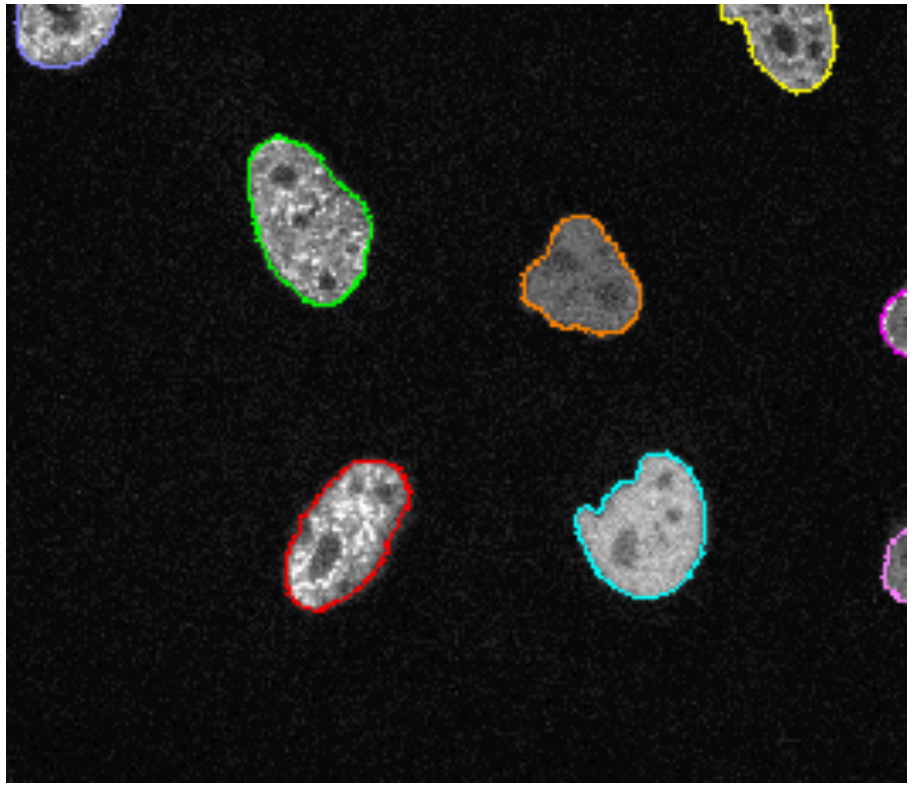
Watch: Pitch Analysis

Application: Military Defense



https://towardsdatascience.com/object-detection-and-tracking-in-pytorch-b3cf1a696a98

Application: Cell Tracking & Mitosis Detection

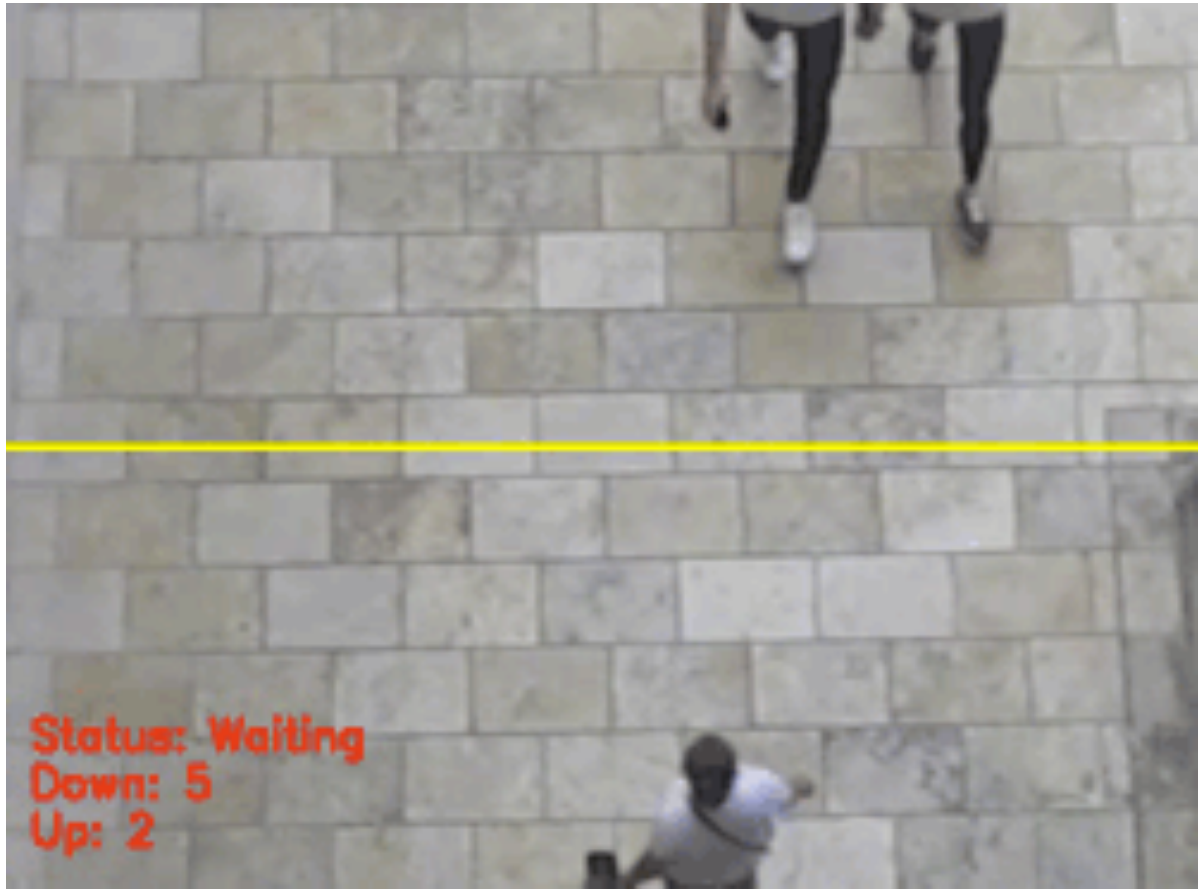


http://celltrackingchallenge.net/

Application: Self-driving Cars

https://www.pyimagesearch.com/2018/08/06/tracking-multiple-objects-with-opencv/

Application: People Counting



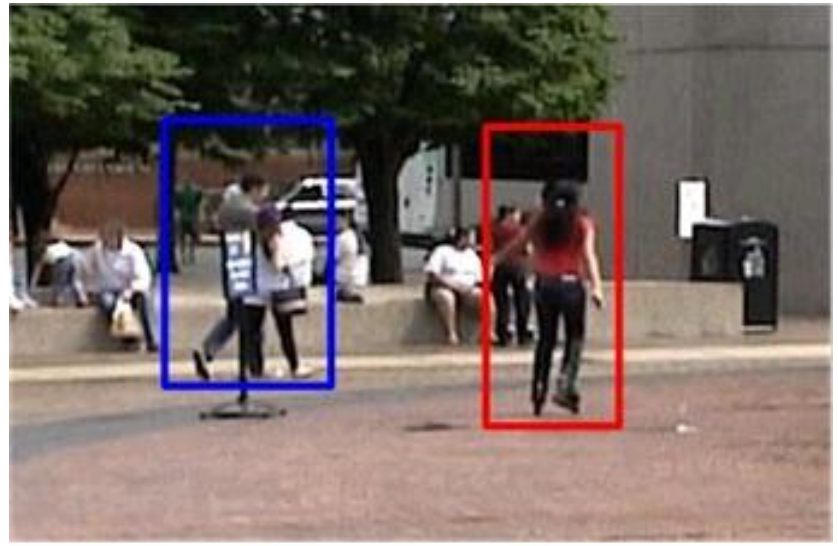
https://www.pyimagesearch.com/2018/08/13/opencv-people-counter/

Application: Human Computer Interaction

Roboceptionist

https://alapkshirsagar.github.io/travel/namerica/usa/wv/47.jpg

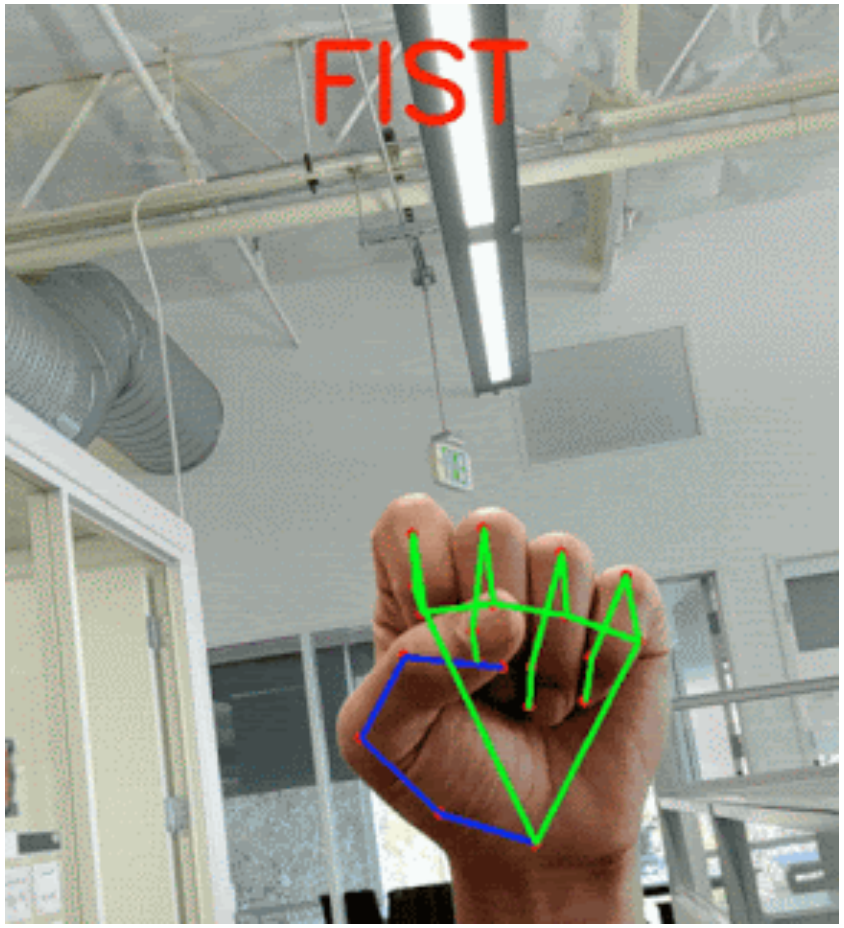
Application: Activity Recognition



https://www.bu.edu/ids/research-projects/action-recognition/

https://www.crcv.ucf.edu/projects/realistic_action_recognition/

Application: Sign Language Recognition



https://techcrunch.com/2019/08/19/this-hand-tracking-algorithm-could-lead-to-sign-language-recognition/

Application: Marketing

https://www.nytimes.com/2013/07/15/business/attention-shopper-stores-are-tracking-your-cell.html

Application: Augmented Reality

https://www.geekwire.com/2017/augmented-reality-shopping-phone-patent-hints-amazons-aspirations/

https://virtualrealitypop.com/object-recognition-in-augmented-reality-8f7f17127a7a

Applications

What other applications can you think of where object tracking could be useful?

Applications: Many More

- Robotics
- Space and satellite
- Industrial production
- Ecological monitoring
- Healthcare
- Anomaly detection

Today's Topics

- Object tracking applications
- Evaluating object tracking methods
- Crowdsourcing object tracking strategies
- Object tracking datasets & challenges
- Lab: video annotation & creating figures/tables in latex

Evaluation Metrics - I

Precision

Distance between the centers of bounding boxes for each frame

A = Ground Truth

B = Predicted Track

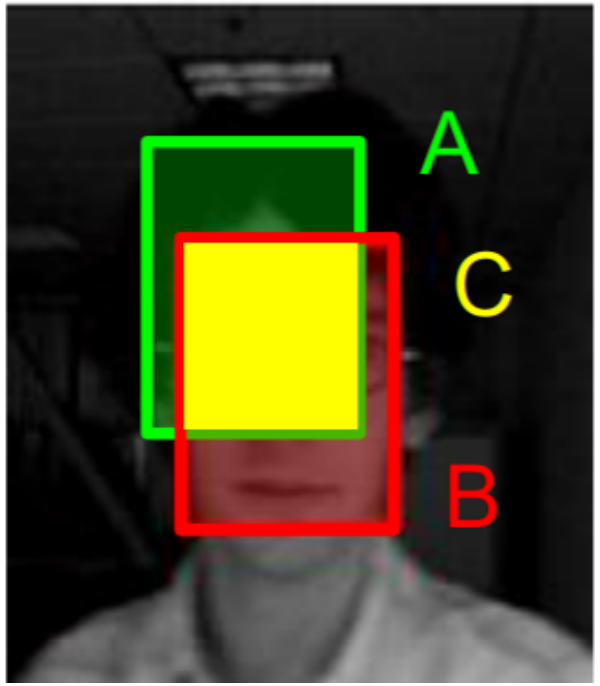
$$p = \|c_A - c_B\|$$

Slide credit: https://ags.cs.uni-kl.de/fileadmin/inf_ags/opt-ss15/OPT_SS2015_lec11.pdf

Evaluation Metrics - II

Overlap

Intersection over Union (IoU) of the bounding boxes for each frame



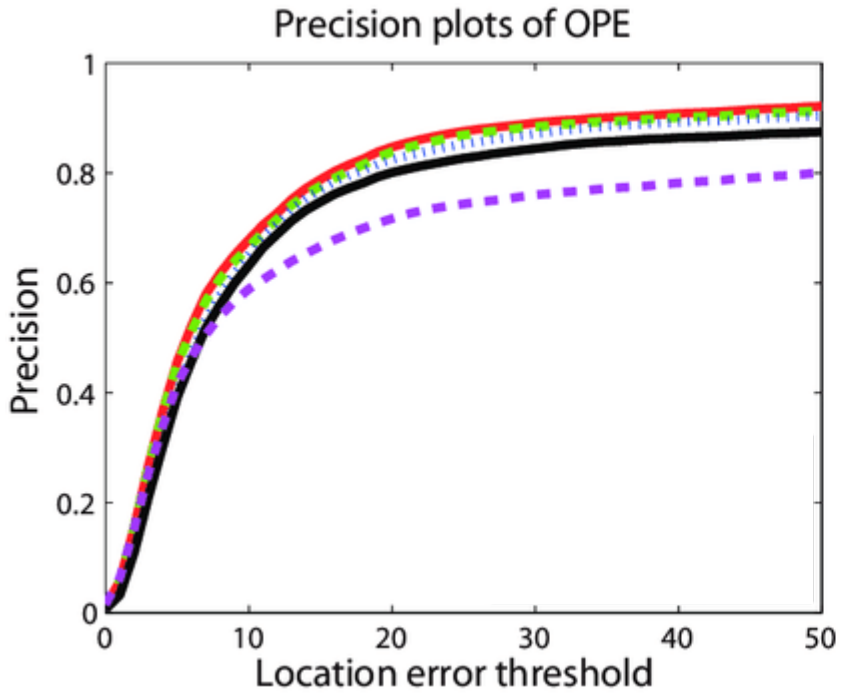
A = Ground Truth
B = Predicted Track
C = Intersection

$$o = \frac{C}{A + B - C}$$

Slide credit: https://ags.cs.uni-kl.de/fileadmin/inf_ags/opt-ss15/OPT_SS2015_lec11.pdf

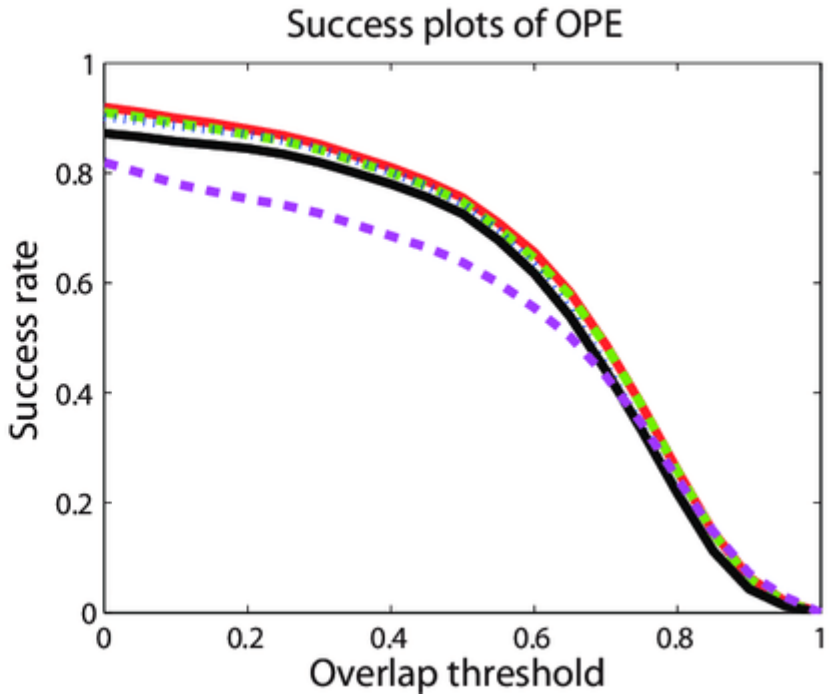
Visualization of Evaluation Metrics - I

- **Precision plot**: Percentage of frames with predicted location within a given threshold distance of ground truth (for e.g., 20 pixels)



Visualization of Evaluation Metrics - II

- **Overlap plot**: Percentage of frames where the overlap between the bounding boxes is larger than a given threshold (for e.g., 0.5)



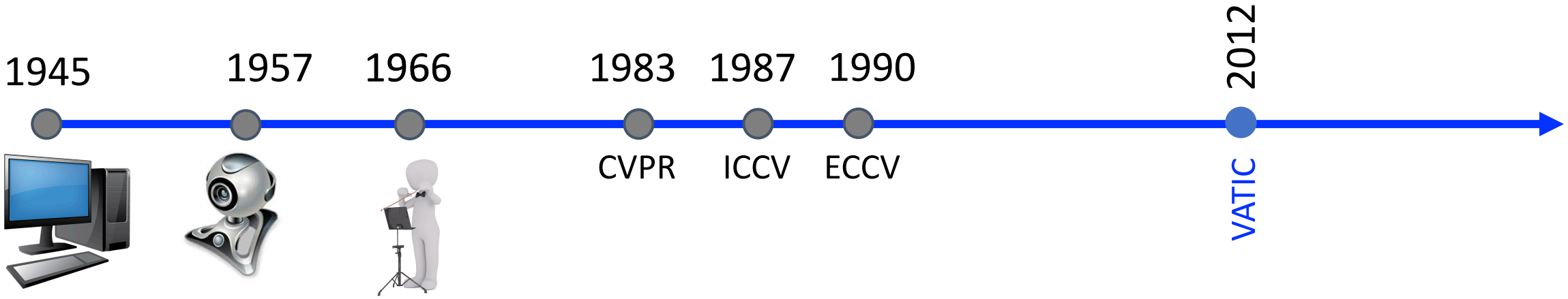
Evaluation References

- Gao, C., Shi, H., Yu, J. G., & Sang, N. (2016). Enhancement of elda tracker based on cnn features and adaptive model update. *Sensors*, *16*(4), 545
- Wu, Y., Lim, J., & Yang, M. H. (2013). Online object tracking: A benchmark. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2411-2418)

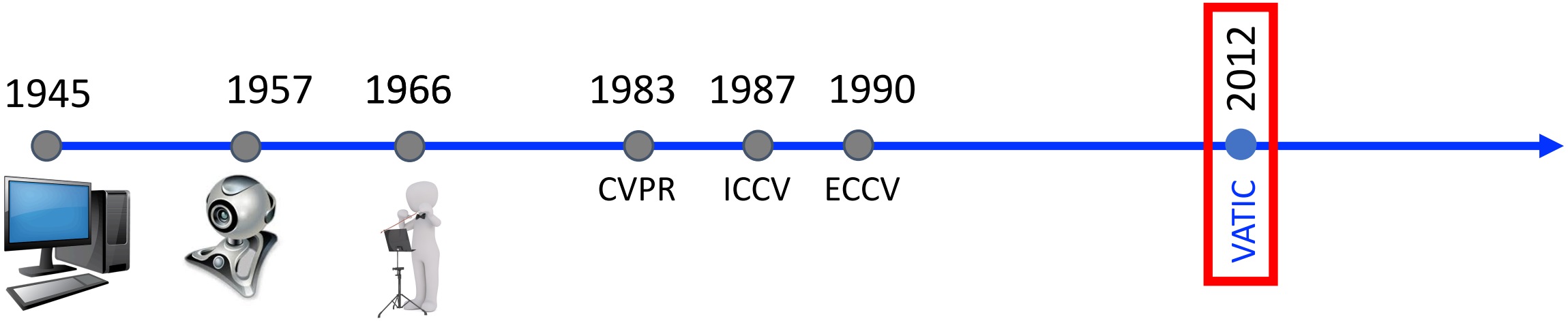
Today's Topics

- Object tracking applications
- Evaluating object tracking methods
- Crowdsourcing object tracking strategies
- Object tracking datasets & challenges
- Lab: video annotation & creating figures/tables in latex

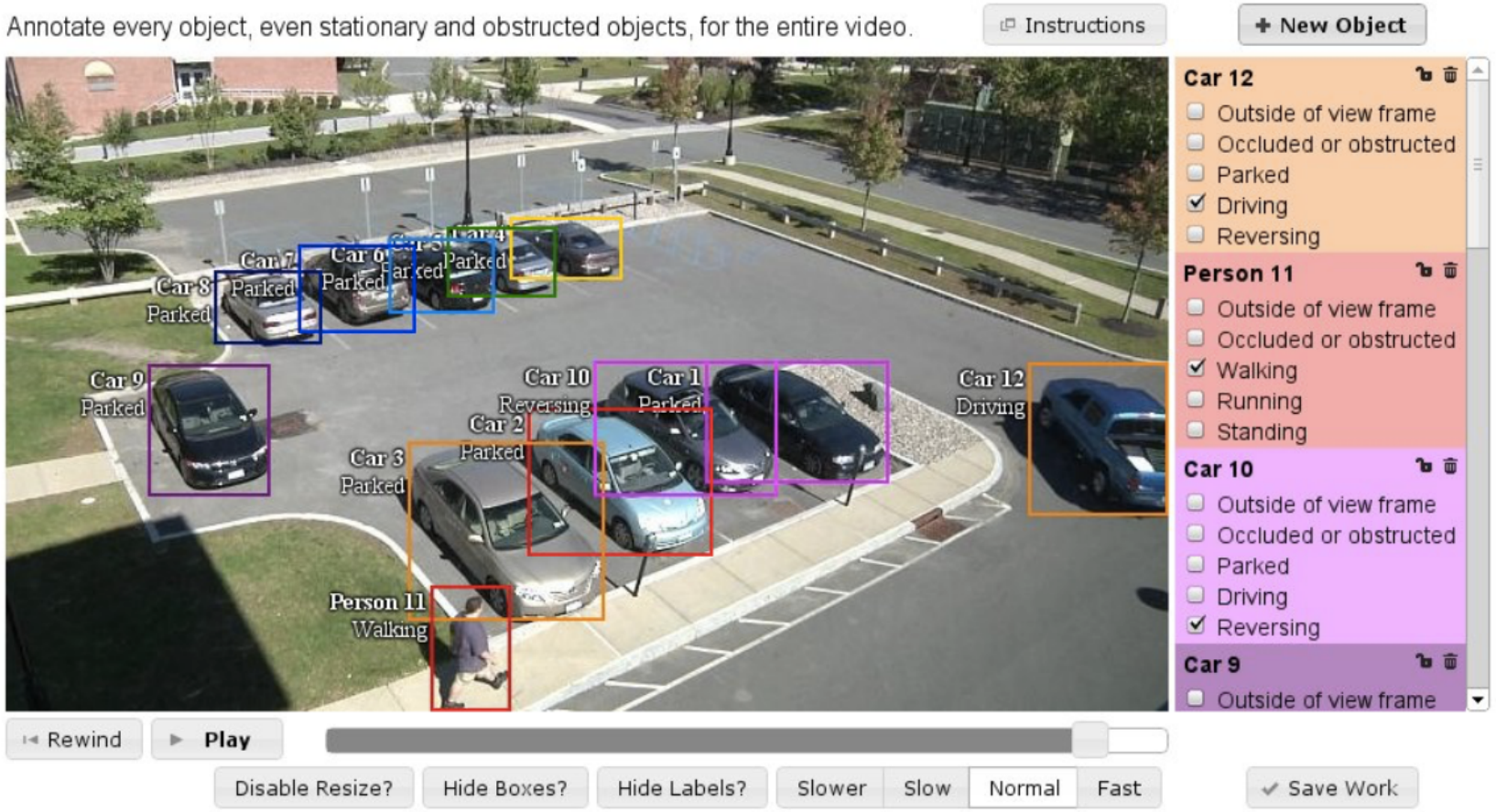
Object Tracking Approaches



Object Tracking Approaches: VATIC



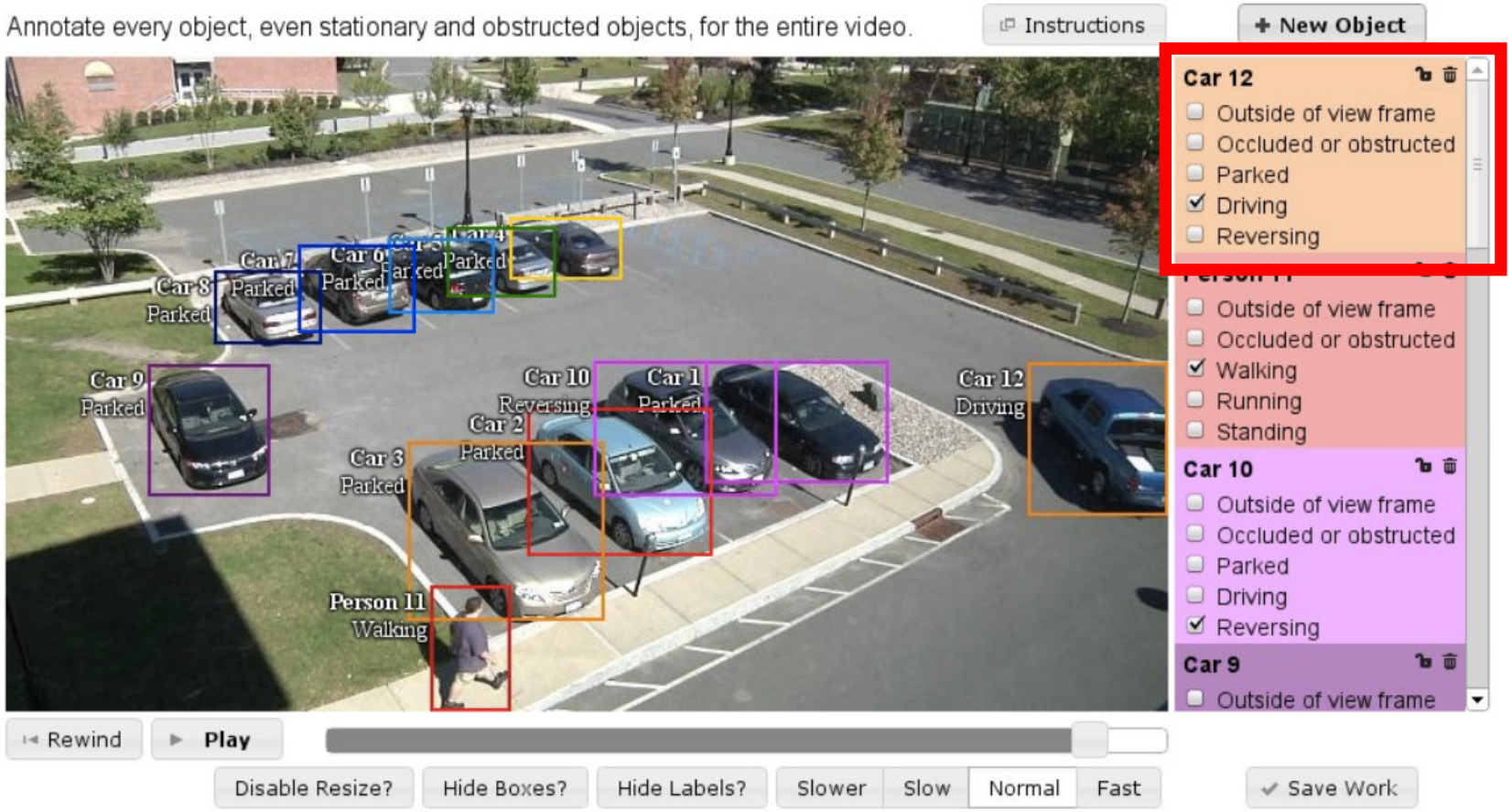
Object Tracking Approaches: VATIC



Demo: https://www.youtube.com/watch?v=ljI5pAowACc

Carl Vondrick, Donald Patterson, and Deva Ramanan. Efficiently Scaling Up Crowdsourced Video Annotation: A Set of Best Practices for High Quality, Economical Video Labeling. IJCV 2012.

Object Tracking Approaches: VATIC



Metadata about each object: e.g., activity, attributes, etc.

Carl Vondrick, Donald Patterson, and Deva Ramanan. Efficiently Scaling Up Crowdsourced Video Annotation: A Set of Best Practices for High Quality, Economical Video Labeling. IJCV 2012.

VATIC Design

- Goal: user studies to establish best-practices
- Data: 3 videos

1. Scripted: large objects, linear motion

2. Basketball: small object occluded frequently, nonlinear motion

3. VIRAT: slow moving/stationary cars, linear motion

Carl Vondrick, Donald Patterson, and Deva Ramanan. Efficiently Scaling Up Crowdsourced Video Annotation: A Set of Best Practices for High Quality, Economical Video Labeling. IJCV 2012.

VATIC Studies: Task Design

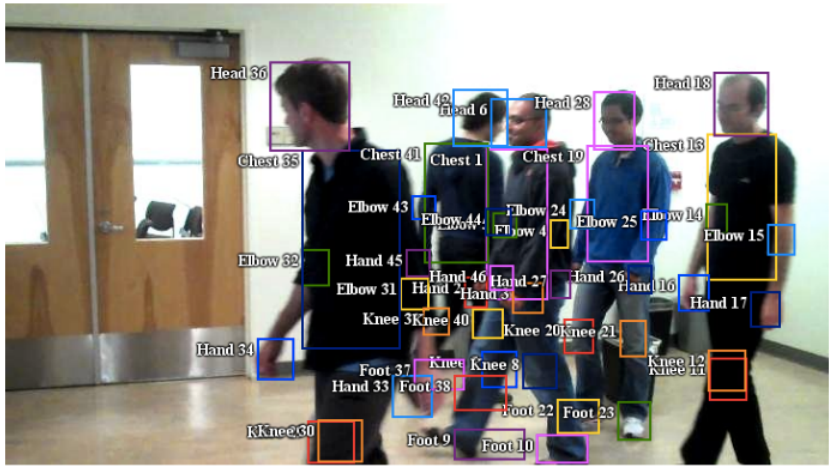
	Scripted				Basketball				VIRAT			
Subject	User	Fixed	Ratio	Saved	User	Fixed	Ratio	Saved	User	Fixed	Ratio	Saved
A	599	**463**	0.77	136	1,457	**1,323**	0.91	134	**220**	244	1.11	-24
B	653	**247**	0.38	406	4,555	**2,275**	0.50	2,280	**176**	**178**	1.03	-2
C	476	**275**	0.58	201	1,216	**830**	0.68	386	338	**215**	0.64	123
D	772	**432**	0.56	340	1,505	**1,497**	0.99	8	489	**302**	0.62	187
*E	605	**371**	0.61	234	**935**	1501	1.61	-566	269	**231**	0.85	38
*F	654	**472**	0.72	182	**1,672**	1,858	1.11	-186	372	**326**	0.87	46
*G	235	**193**	0.82	42	**591**	696	1.18	-105	165	**120**	0.73	45
*H	**312**	331	1.06	-19	**656**	748	1.14	-92	172	**164**	0.95	8
Mean	538	**348**	0.66	190	1,573	**1,341**	0.96	232	275	**223**	0.83	53

Fixed interval key frames significantly faster than user defined key frames!

Carl Vondrick, Donald Patterson, and Deva Ramanan. Efficiently Scaling Up Crowdsourced Video Annotation: A Set of Best Practices for High Quality, Economical Video Labeling. IJCV 2012.

VATIC Studies: Task Design

Subject	Human Joints						Basketball						Preferred	Order
	One		Group		All		One		Group		All			
A	**1639**	1.00	1763	1.07	-	-	476	1.00	358	0.75	**253**	0.53	One	OGA
B	**2071**	1.00	2100	1.01	2440	1.17	**1382**	1.00	1618	1.17	2403	1.73	One	AOG
C	**1867**	1.00	2106	1.12	-	-	681	1.00	813	1.19	**599**	0.87	One	GAO
D	**2399**	1.00	-	-	-	-	494	1.00	**448**	0.91	960	1.94	One	AOG
Mean	**1994**	1.00	1989	1.07	2440	1.17	**759**	1.00	809	0.99	1085	1.12		



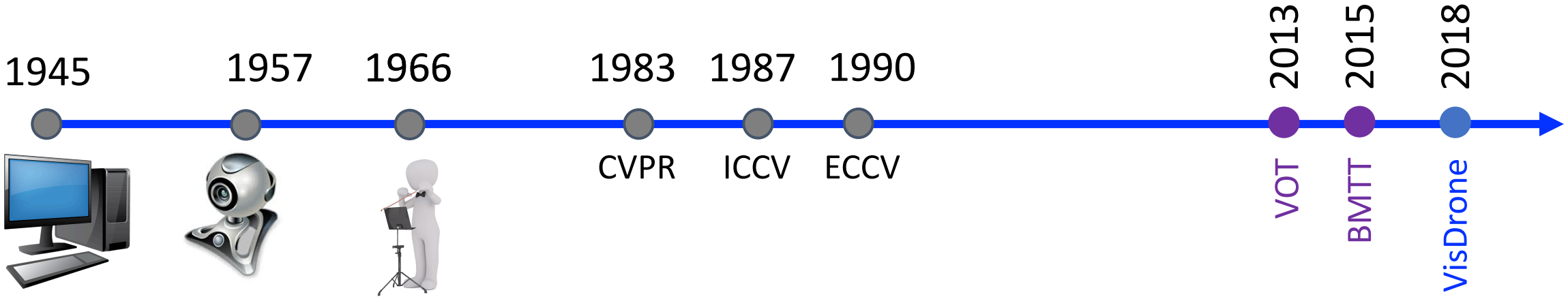
Annotating one object at a time is more efficient and preferred!

Carl Vondrick, Donald Patterson, and Deva Ramanan. Efficiently Scaling Up Crowdsourced Video Annotation: A Set of Best Practices for High Quality, Economical Video Labeling. IJCV 2012.

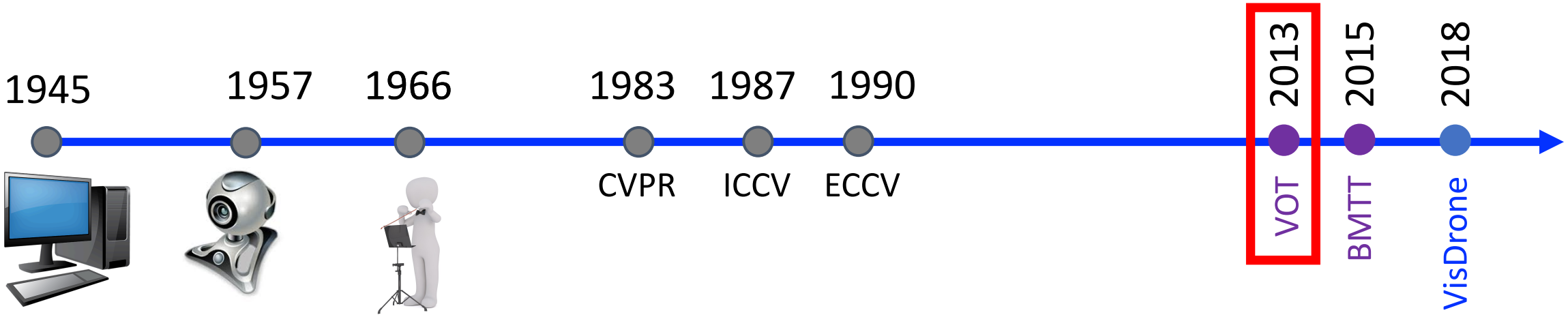
Today's Topics

- Object tracking applications
- Evaluating object tracking methods
- Crowdsourcing object tracking strategies
- Object tracking datasets & challenges
- Lab: video annotation & creating figures/tables in latex

Object Tracking Datasets



Object Tracking Datasets



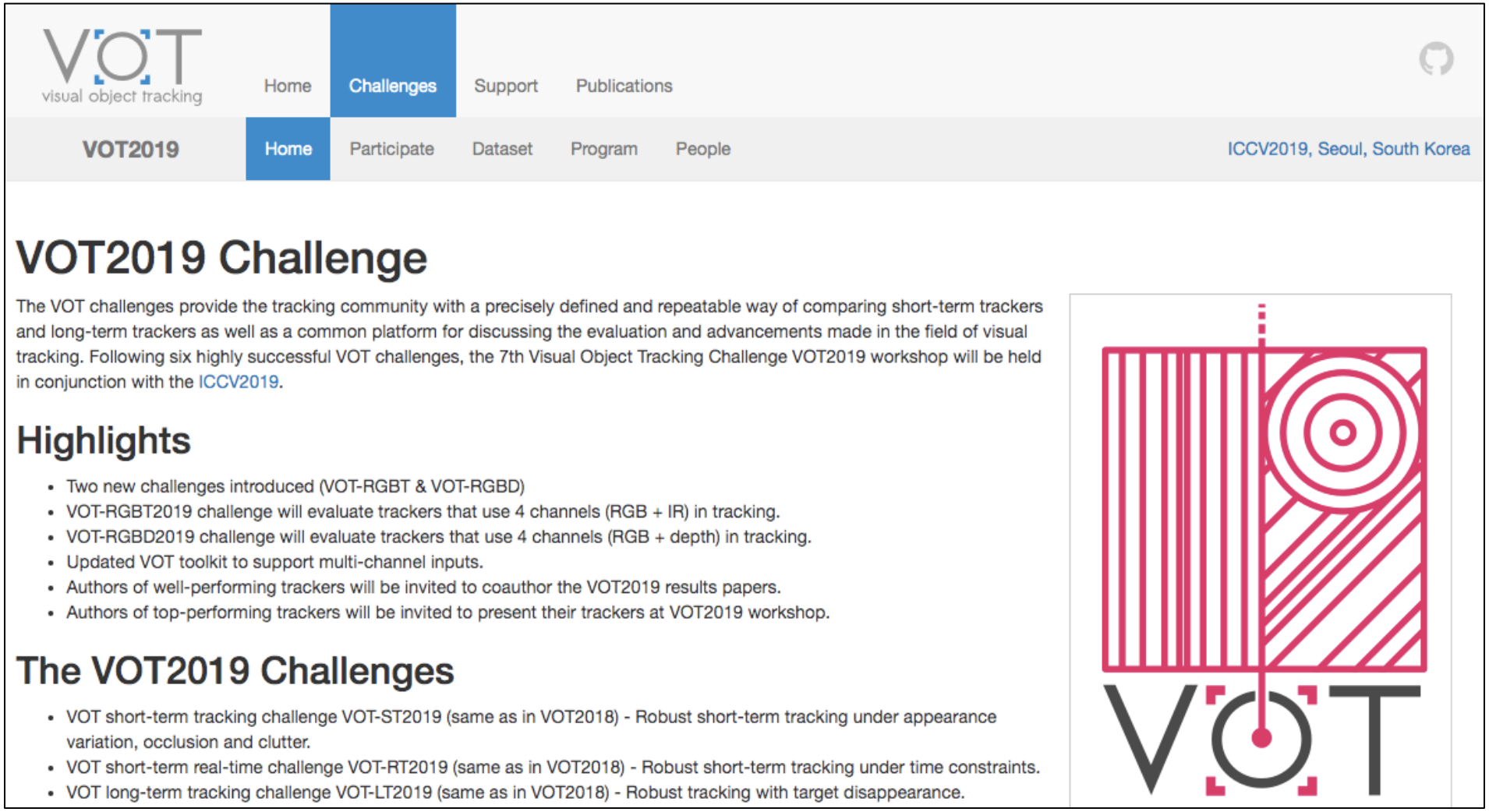
Object Tracking Datasets: VOT

- Collected existing datasets, which resulted in inconsistent annotation methodologies across videos (e.g., different bounding box criteria)

- Authors re-annotated object tracking for videos they deemed to have unsuitable annotations

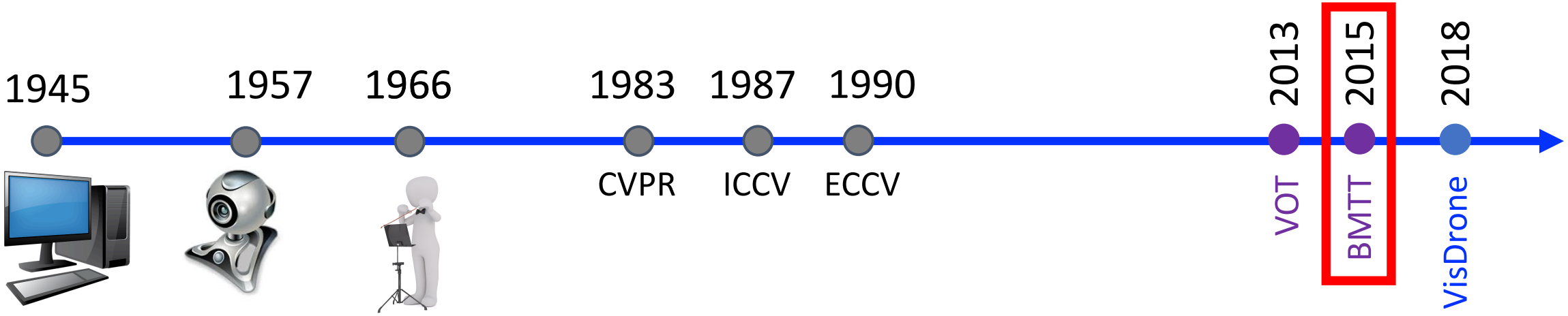
Matej Kristan et al. The Visual Object Tracking VOT2013 Results. ICCV Workshop 2013.

VOT Annual Challenge (7th year now)



http://www.votchallenge.net/vot2019/

Object Tracking Datasets



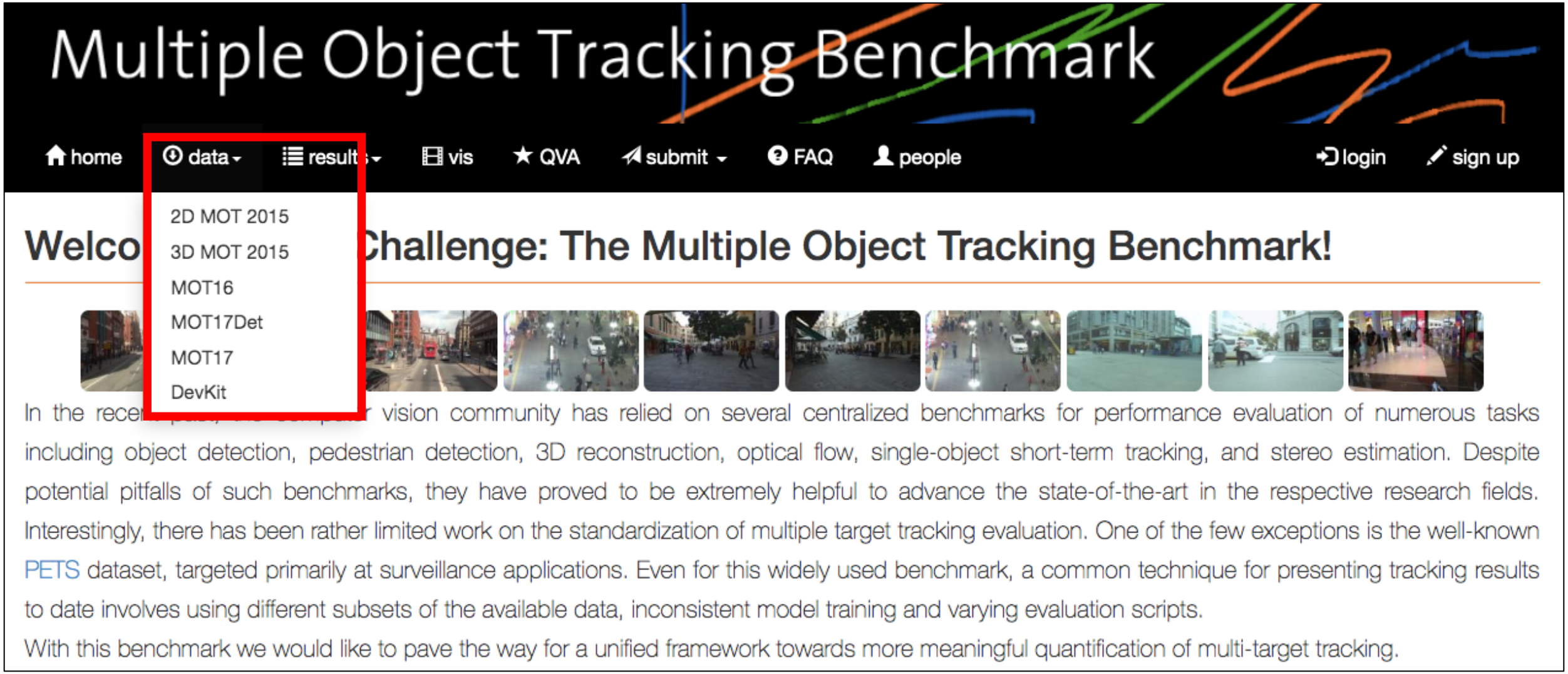
Object Tracking Datasets: BMTT

- Authors collected 22 videos

- Unspecified how object tracking annotations were collected

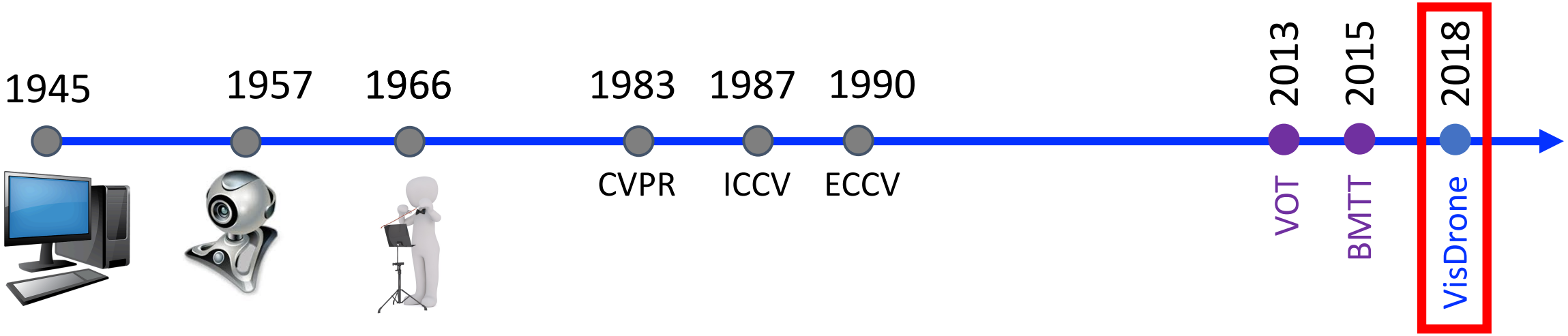
Laura Leal-Taixe, Anton Milan, Ian Reid, Stefan Roth, and Konrad Schindler.
MOTChallenge 2015: Towards a Benchmark for Multi-Target Tracking. arXiv 2015.

BMTT MOT Annual Challenge



https://motchallenge.net/

Object Tracking Datasets



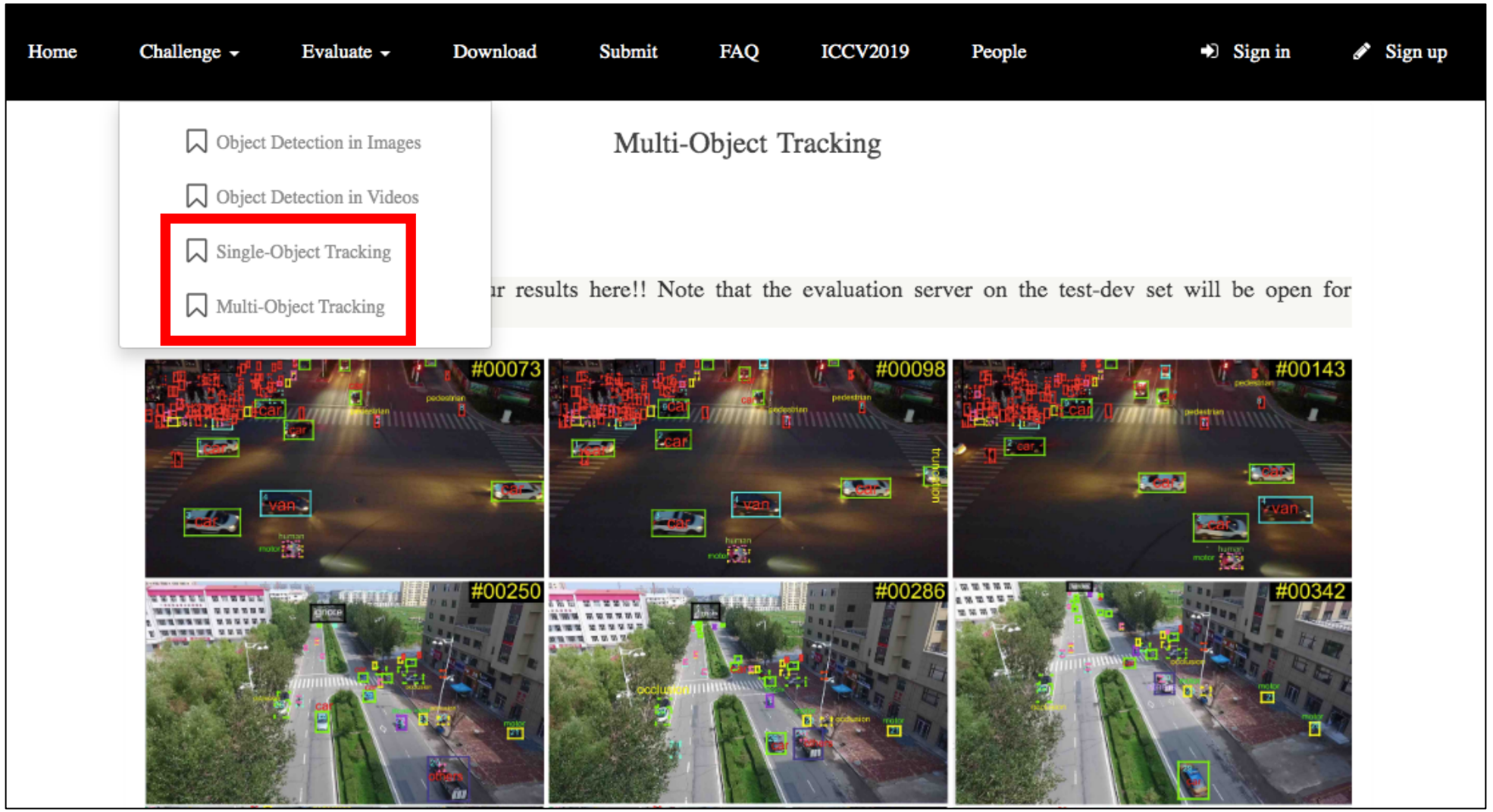
Object Tracking Datasets: VisDrone

- Authors collected 263 video clips (179,264 frames) from drones in Asia

- Annotations created for over 2.5 million object instances, however it is unspecified how these annotations were collected

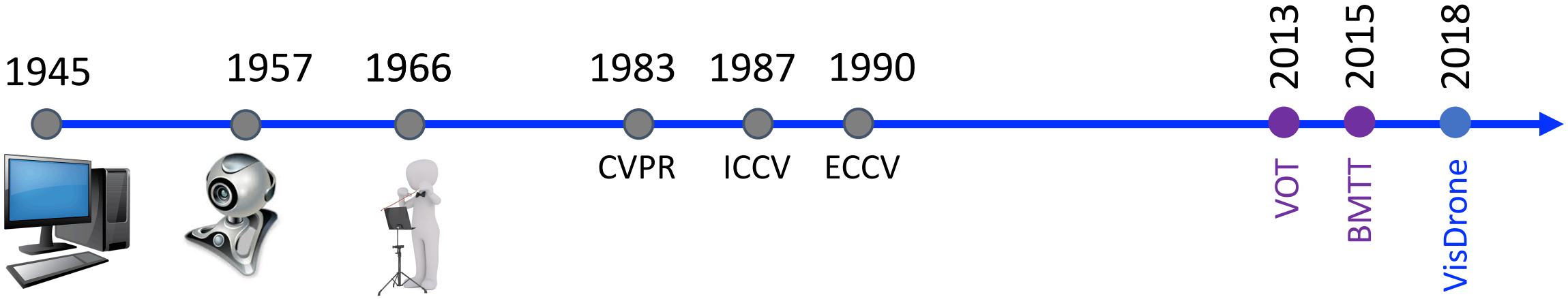
Pengfei Zhu, Longyin Wen, Xiao Bian, Haibing Ling, and Qinghua Hu. Vision Meets Drones: A Challenge. arXiv 2018.

VisDrone Challenge



http://www.aiskyeye.com/views/index

Object Tracking Datasets



Discussion

- Design a crowdsourcing system to collect **high quality** object tracking annotations. Specify how your system design will handle difficult scenarios such as when an object is:
  - Partially visible
  - Occluded
  - Is reflected in reflective surfaces such as mirrors or windows

- What will be the total crowdsourcing task cost if we use your system to annotate 1,000 1-minute videos where you need to track 5 humans/video (assume 30 frames/second)?

Today's Topics

- Object tracking applications
- Evaluating object tracking methods
- Crowdsourcing object tracking strategies
- Object tracking datasets & challenges
- Lab: video annotation & creating figures/tables in latex