ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Tracking the Invisible: Learning Where the Object Might be

Helmut Grabner[1], Jiri Matas[2], Luc Van Gool[1,3], Philippe Cattin[4]

[1]ETH-Zurich, [2]Czech Technical University, Prague, [3]K.U. Leuven, [4]University of Basel

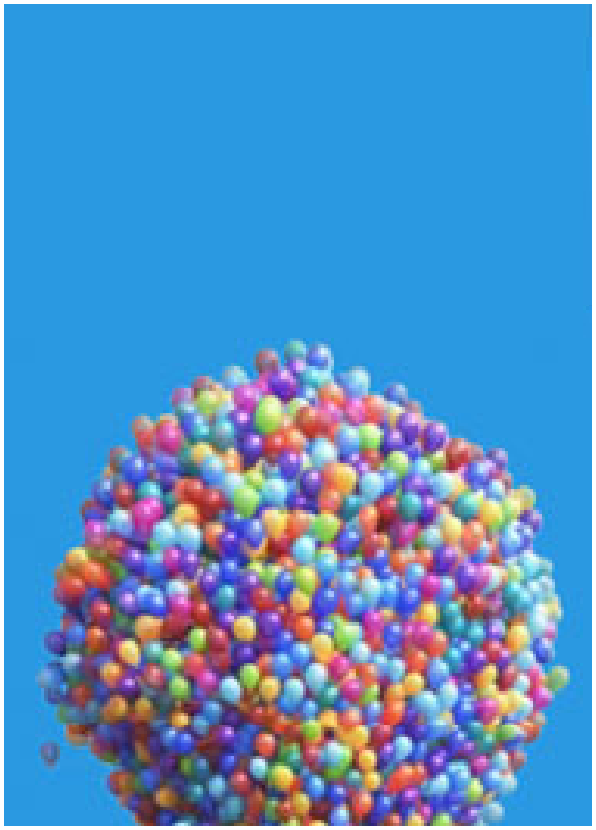
I'm Carl – Track me…

Tracking Carl

Goal: Estimate the Position of an Object

Track the Object…

… with occlusions



… outside the image

Many temporary, but potential very strong links exists between a tracked object and parts of the images.

We discover the dynamic elements – called SUPPORTERS – that predict the position of the target.

[L. Cerman, J. Matas, J., V. Hlavac, **Sputnik Tracker,** *SCIA*, 2009]

[M. Yang, Y. Wu, G. Hua. **Context-Aware Visual Tracking** *PAMI*, 2009]

SUPPORTERS…

- … came with different strength.
- … change over time.

SUPPORTERS…

- … came with different strength.
- … change over time.

SUPPORTERS help Tracking of…

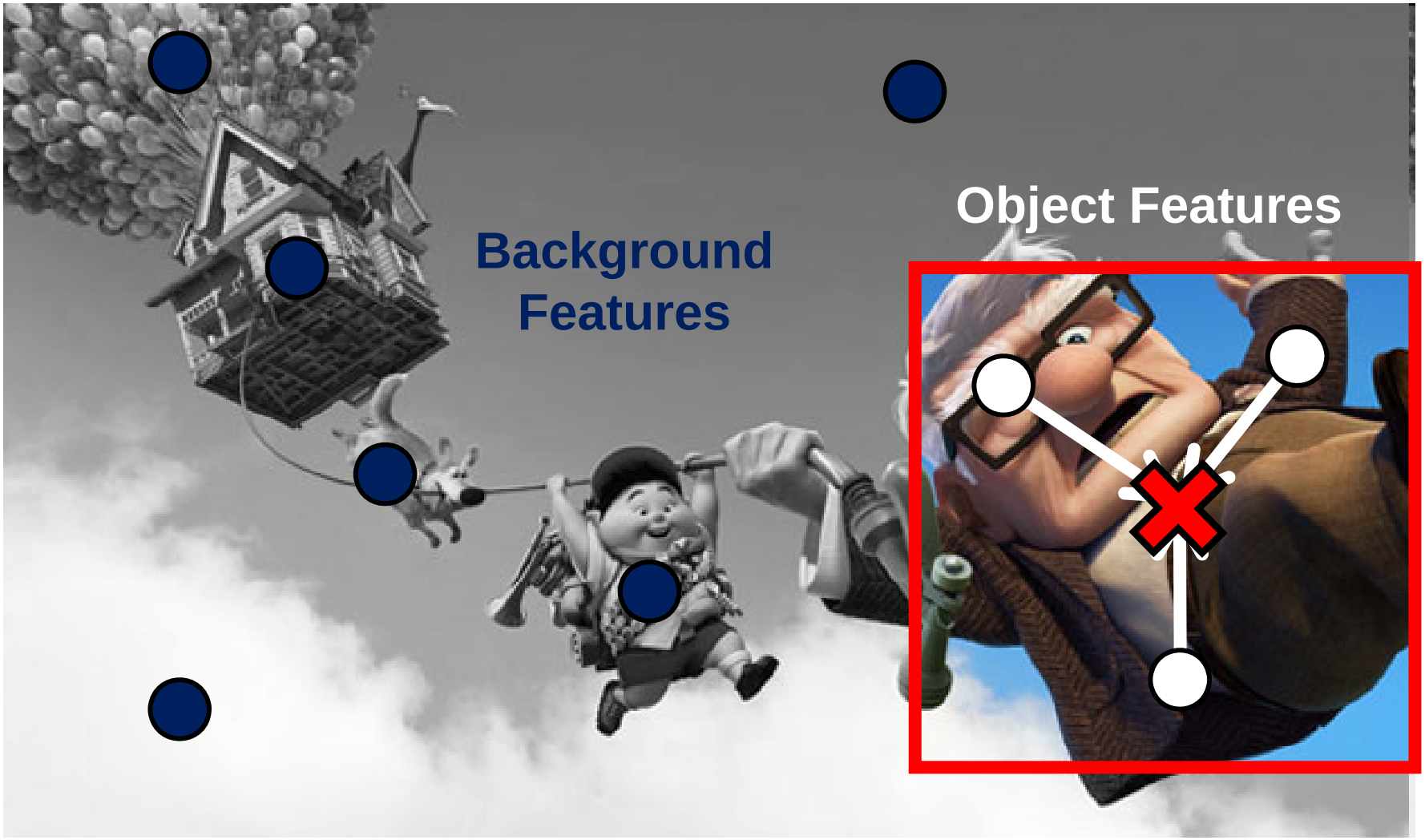
- … objects which change there appearance very quickly.
- … occluded objects or object outside the image.
- … small and/or low textured objects or even “virtual points”.

Local Image Features as Supporters

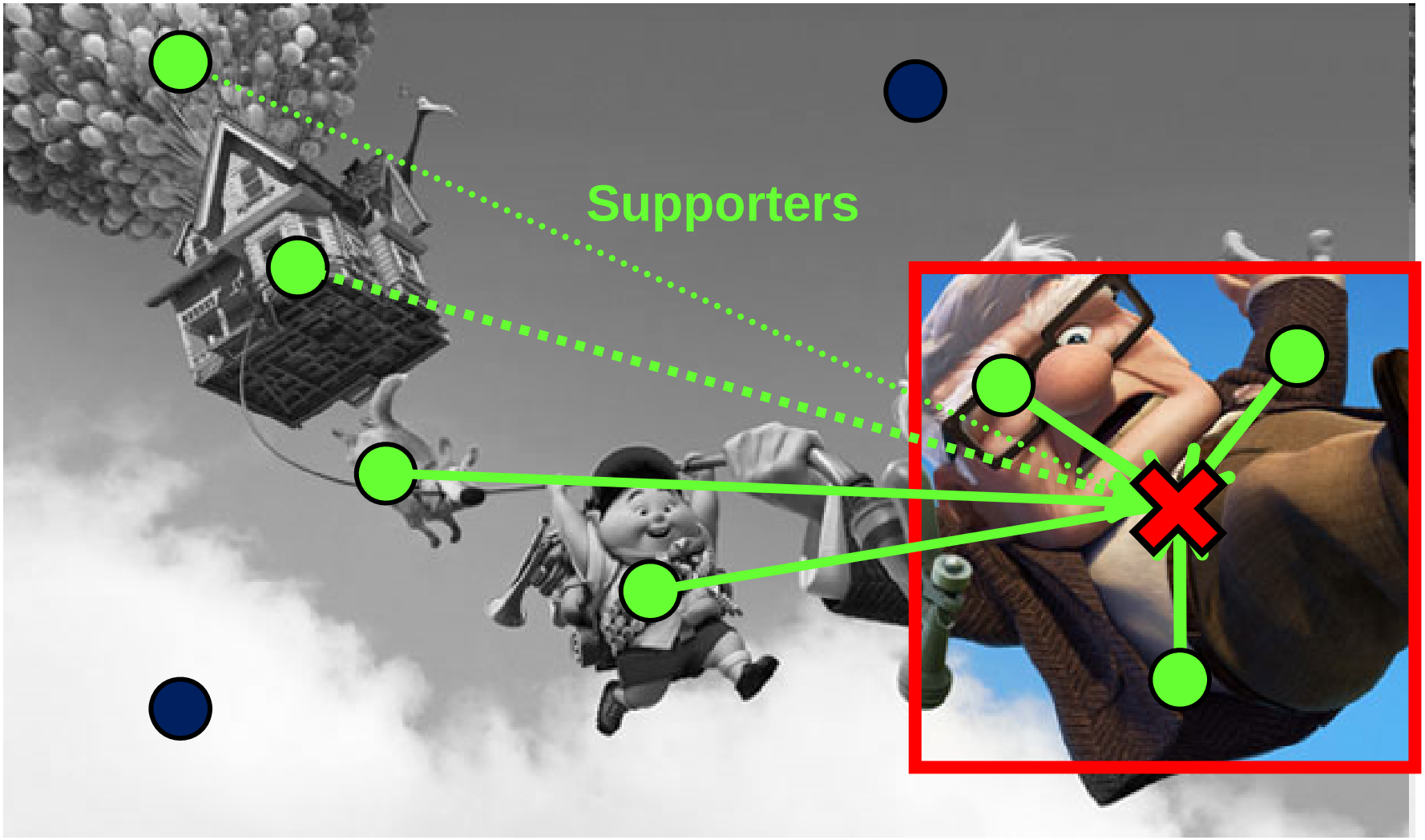
Local Image Features as Supporters

Local Image Features as Supporters

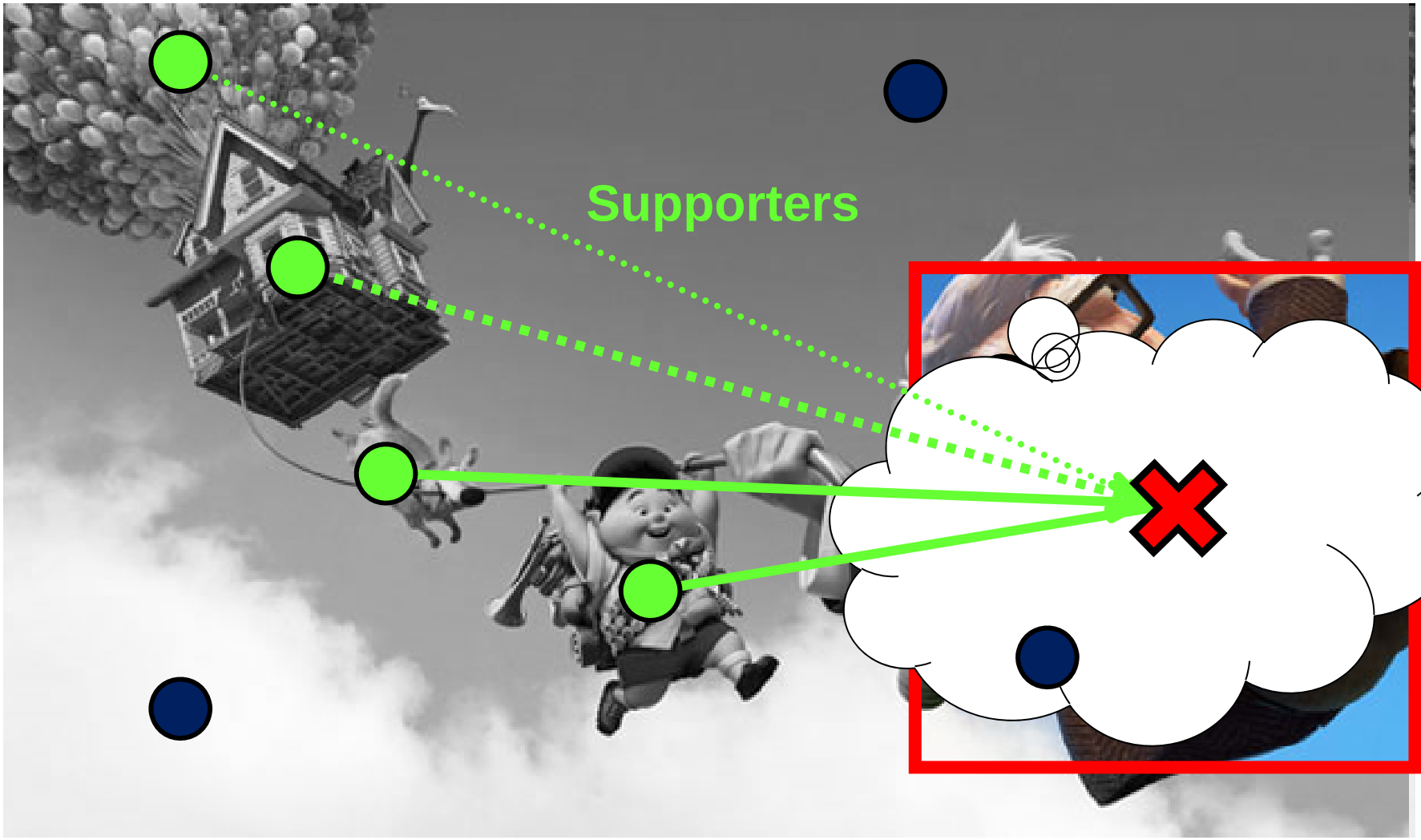
Local Image Features as Supporters



Local Image Features as Supporters



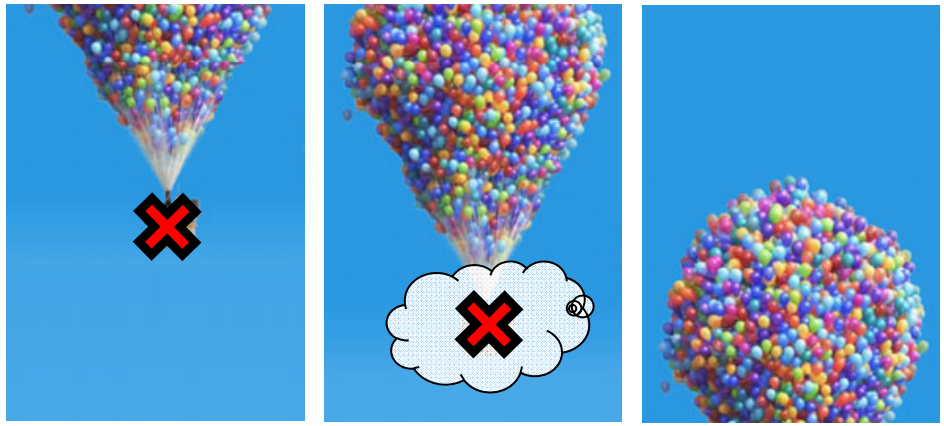
Local Image Features as Supporters



SUPPORTERS are features which contribute to the prediction of the target position.

They at least temporarily move in a way which is statistically related to the motion of the object.

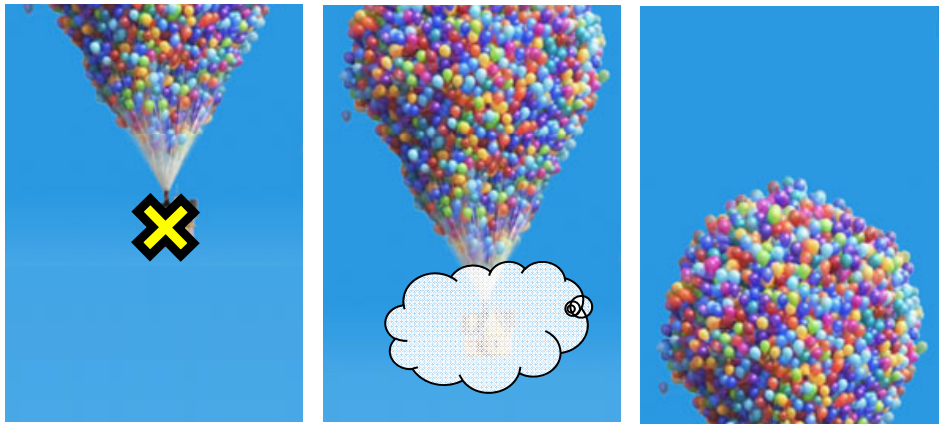
Discovering the Supporters



Predicted postion

Reliable measurments

Model & learning

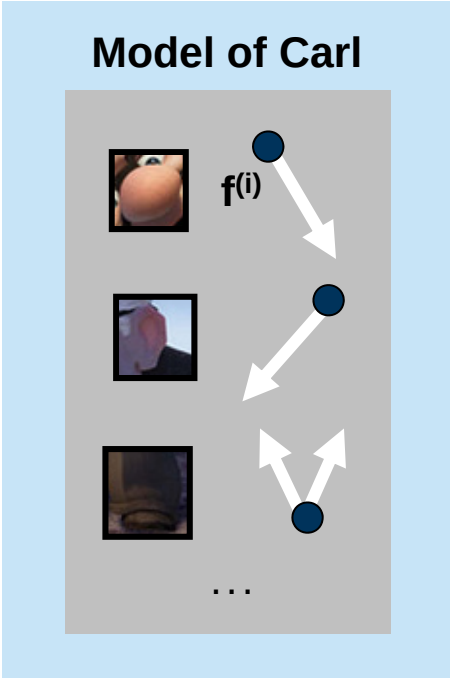


Model

$$P(\mathbf{x}|I)$$

Implicit Shape Model

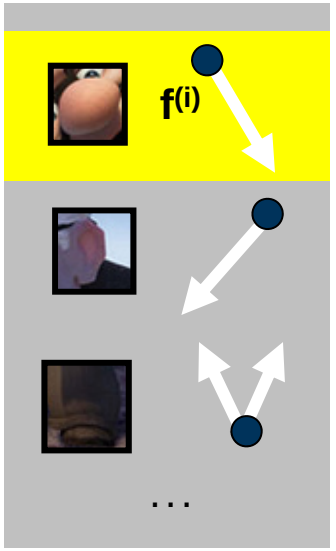
$$P(\mathbf{x}|I) \propto S = \sum_{\mathbf{f}\in\mathcal{F}} P(\mathbf{x}|\mathbf{f})P(\mathbf{f}|I)$$



[B. Leibe, A. Leonardis, B. Schiele, **Robust Object Detection with Interleaved Categorization and Segmentation,** *IJCV*, 2007]

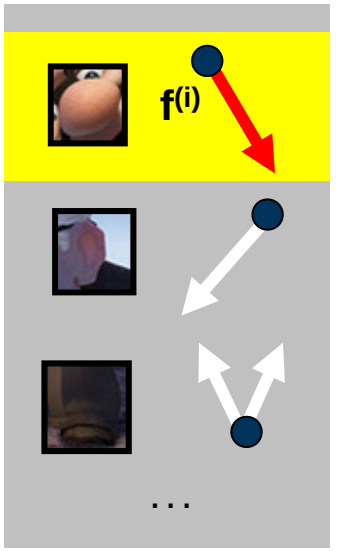
Implicit Shape Model - Features

$$P(\mathbf{x}|I) \propto S = \sum_{\mathbf{f}\in\mathcal{F}} P(\mathbf{x}|\mathbf{f})P(\mathbf{f}|I)$$



Implicit Shape Model – Object Displacement

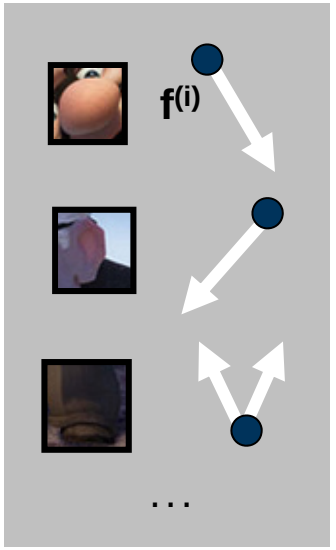
$$P(\mathbf{x}|I) \propto S = \sum_{\mathbf{f} \in \mathcal{F}} P(\mathbf{x}|\mathbf{f}) P(\mathbf{f}|I)$$



Implicit Shape Model

$$P(\mathbf{x}|I) \propto S = \sum_{\mathbf{f}\in\mathcal{F}} P(\mathbf{x}|\mathbf{f})P(\mathbf{f}|I)$$

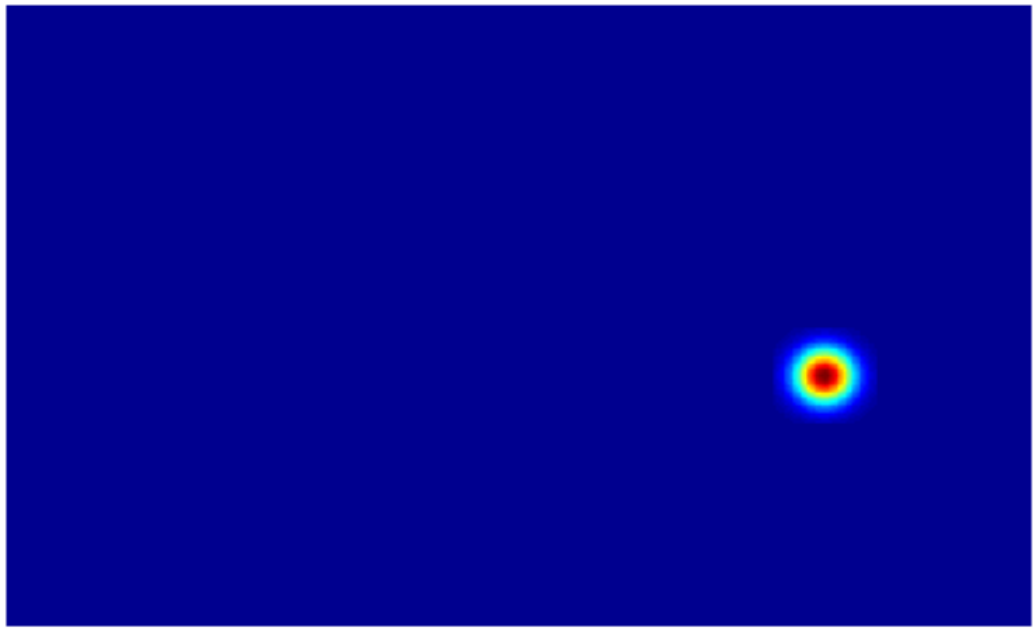
Model of Carl



Implicit Shape Model

$$P(\mathbf{x}|I) \propto S = \sum_{\mathbf{f}\in\mathcal{F}} P(\mathbf{x}|\mathbf{f})P(\mathbf{f}|I)$$

Voting Space



Implicit Shape Model

$$\hat{\mathbf{x}} = \max \sum_{\mathbf{f} \in \mathcal{F}} P(\mathbf{x}|\mathbf{f}) P(\mathbf{f}|I)$$

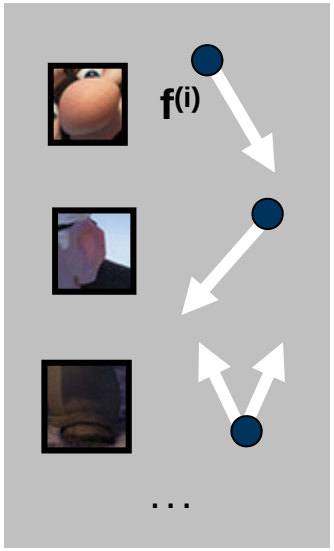
„Model of Carl“ – Object Detection

Large dataset of labeled images

OFF-line training

Prototypes
Object Displacment

Model of Carl

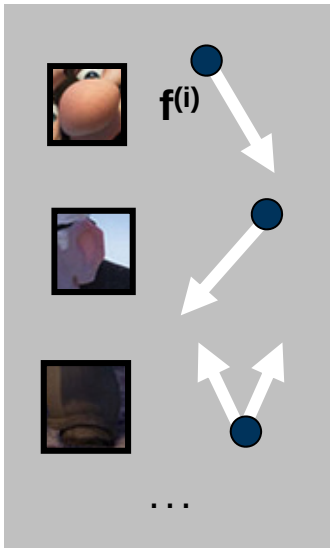


Model of Supporters

Update Prototypes

Update Object Displacment

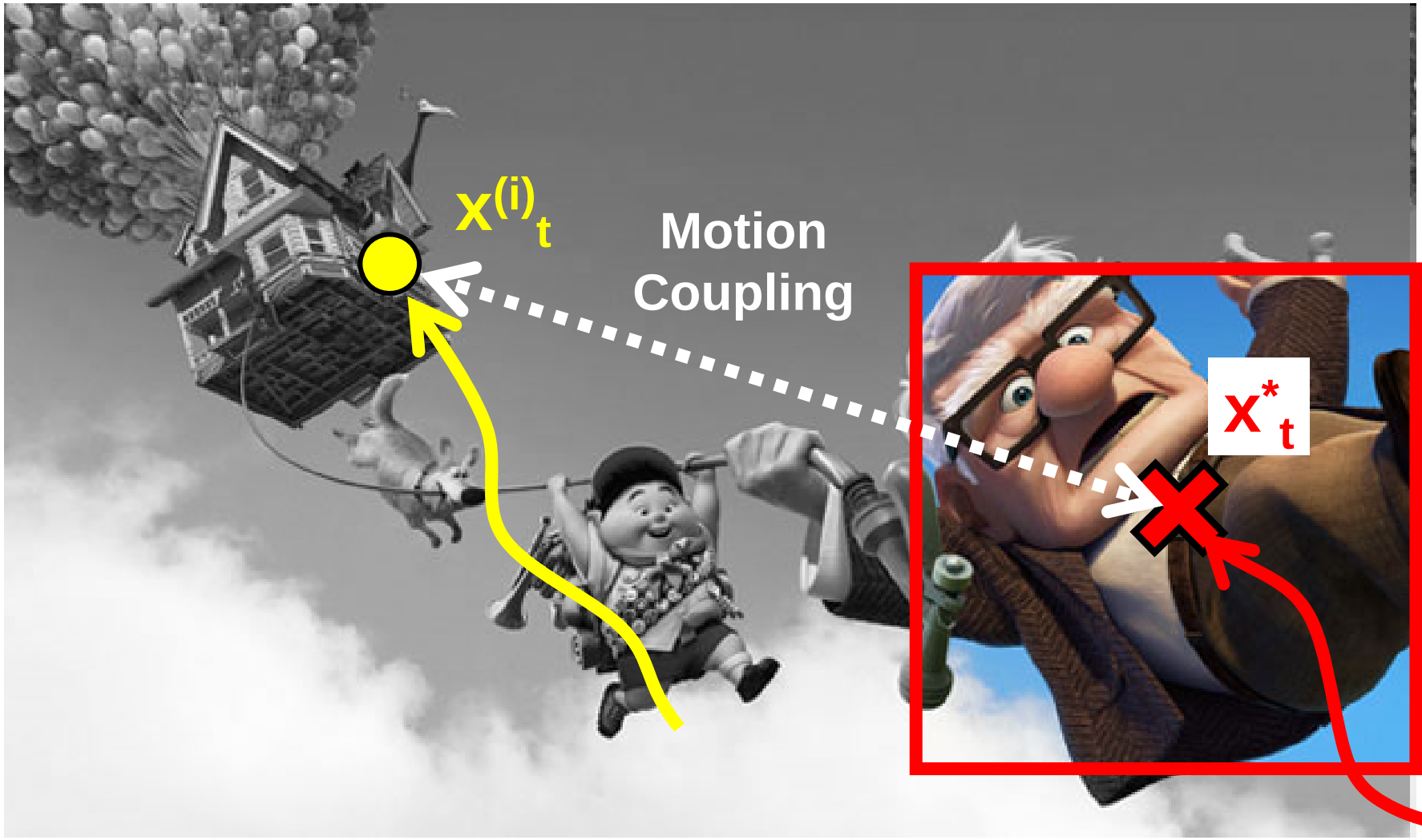
Model at time t



ON-line training

Model at time t+1

Reliable Information & Motion Coupling



Supporter

Strong Supporter

Strong motion coupling

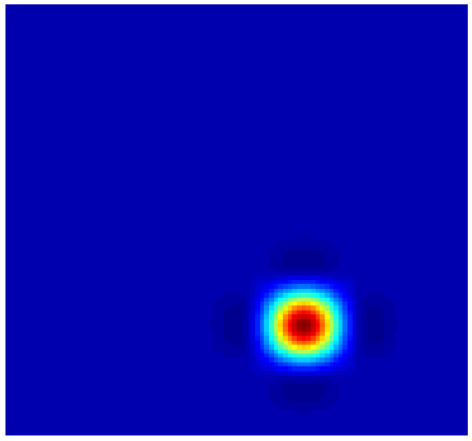
Peaky vote

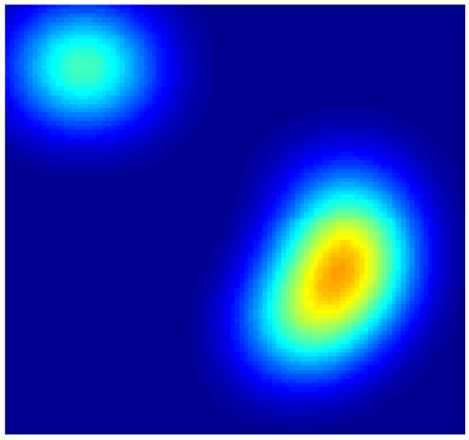
Weak Supporter

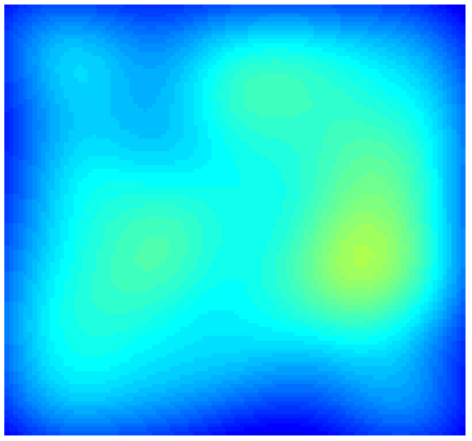
Weak motion coupling

Blurred vote

Almost Unrelated Features





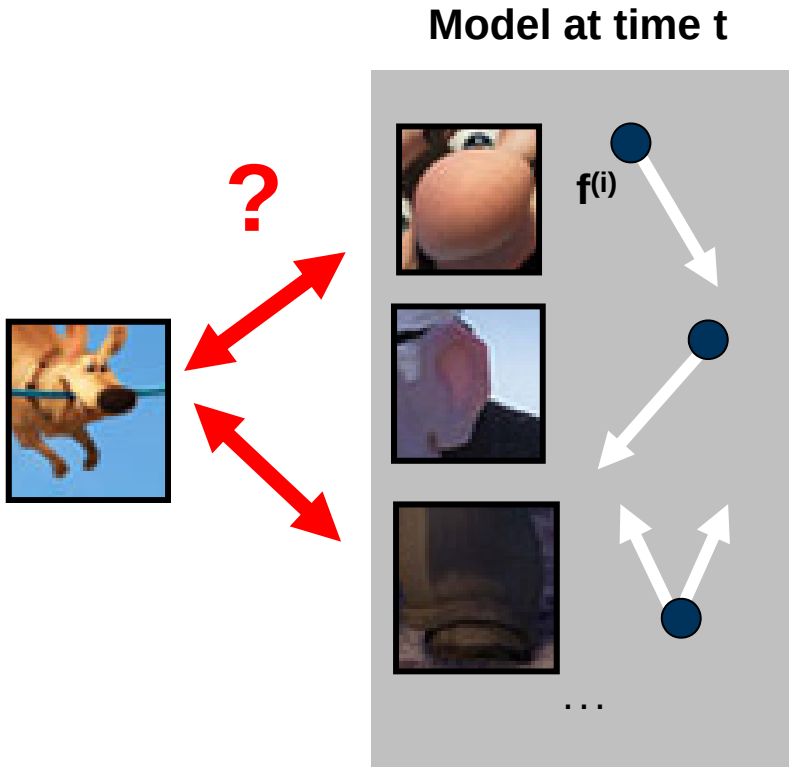


Implementation / Approximations

$$P(\mathbf{x}|I) \propto S = \sum_{\mathbf{f}\in\mathcal{F}} P(\mathbf{x}|\mathbf{f})P(\mathbf{f}|I)$$

- **Database of Supporters**

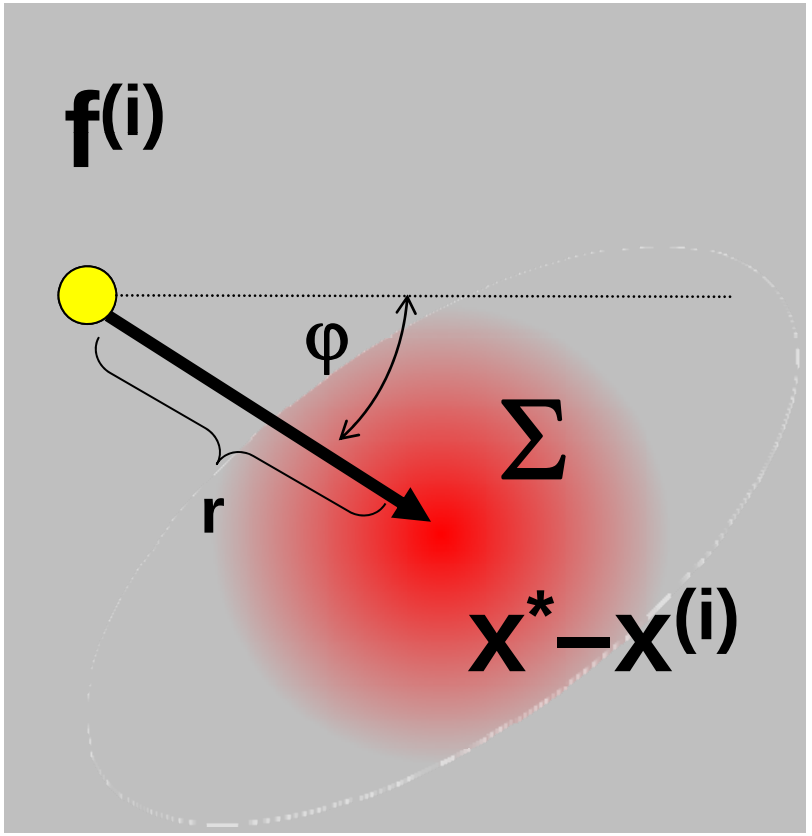
- Harris points
- Feature point matching using a „light“ SIFT descibtor
- KLT tracking



Implementation / Approximations

$$P(\mathbf{x}|I) \propto S = \sum_{\mathbf{f} \in \mathcal{F}} P(\mathbf{x}|\mathbf{f}) P(\mathbf{f}|I)$$

- Relative object position with respect to the feature
- Single Gaussian
- On-line update using exponential forgetting

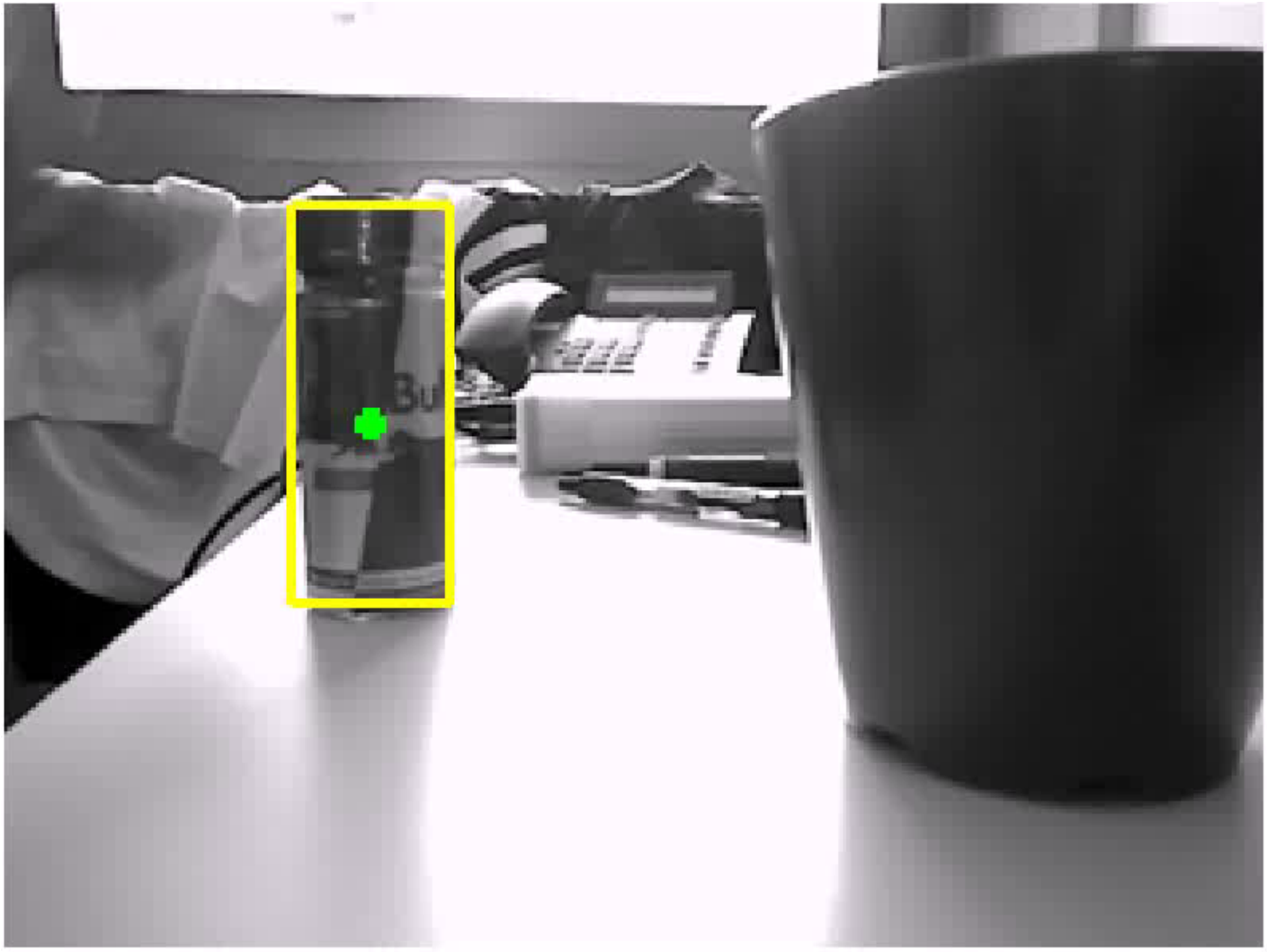


Experimental Results: *ETH-Cup* Sequenze

ETH-Cup: Humans

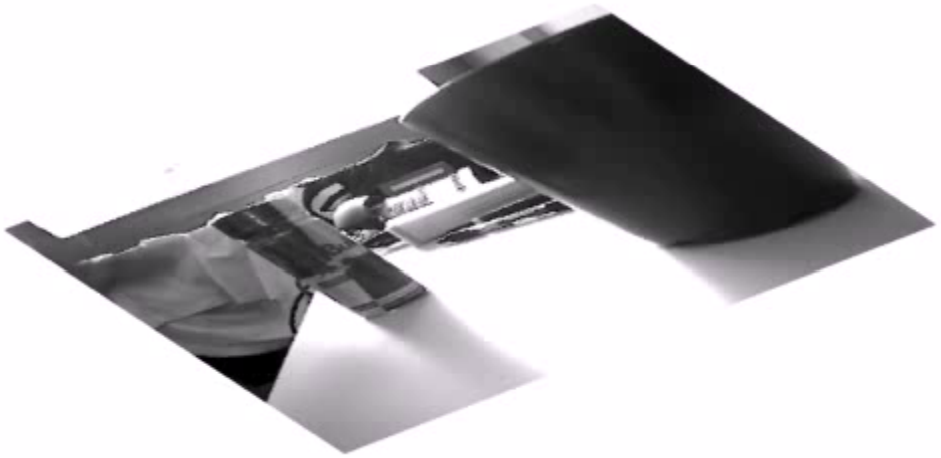
www.vision.ee.ethz.ch/boostingTrackers

ETH-Cup: Of the Web Tracker

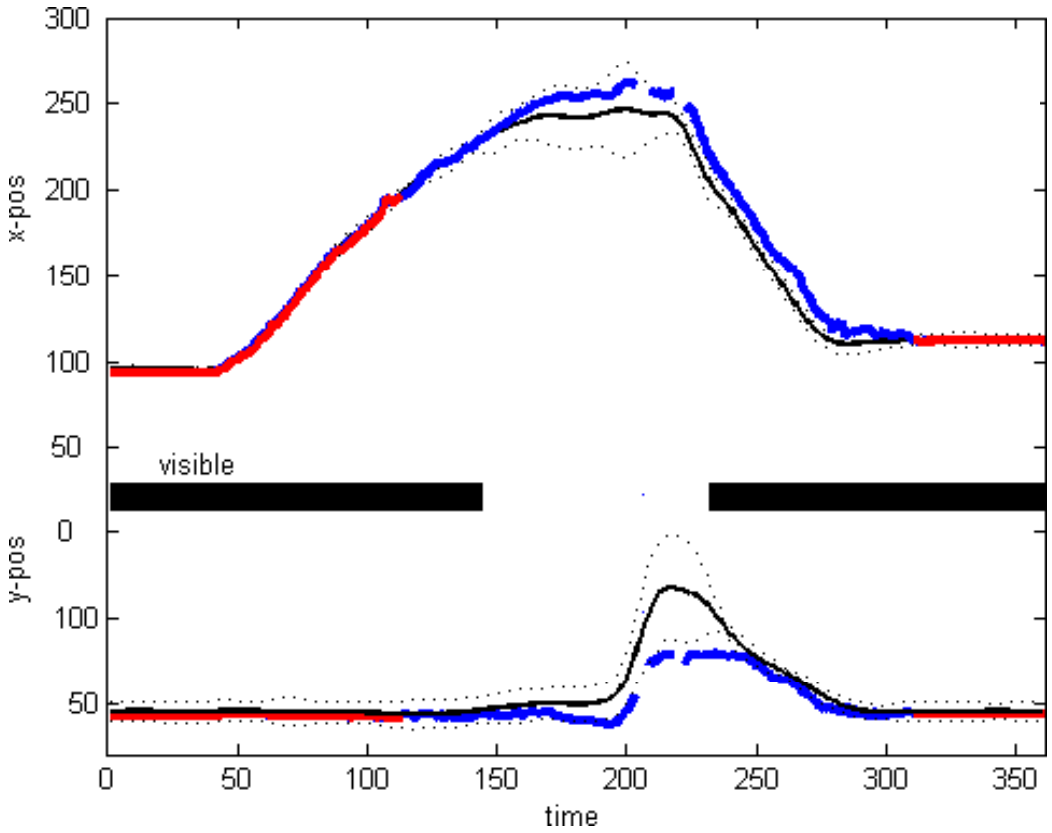


[S. Stalder, H. Grabner, L. Van Gool, **Beyond Semi-Supervised Tracking**, OLCV, 2009]

ETH-Cup: Our Result – Voting Space



ETH-Cup: Improving Object Tracking

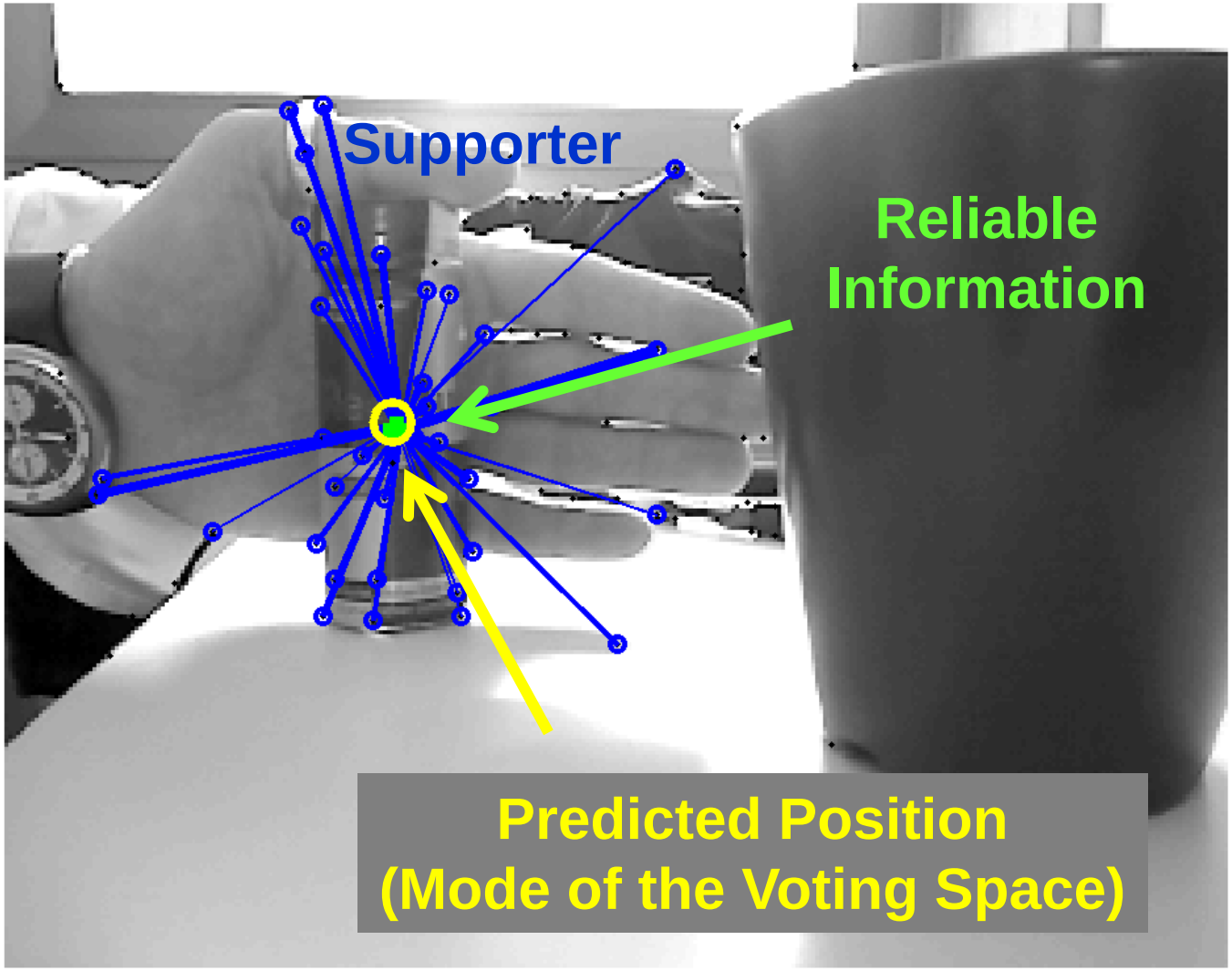


10 Humans
($\mu \pm 2\,\sigma$)

	Rec.	Prec.
Tracker [Stalder et al, OLCV'09]	45 %	100 %
+ supporter	89 %	97 %

Please note, a simple interpolation would not work.

ETH-Cup: Supportes



ETH-Cup: Supporters

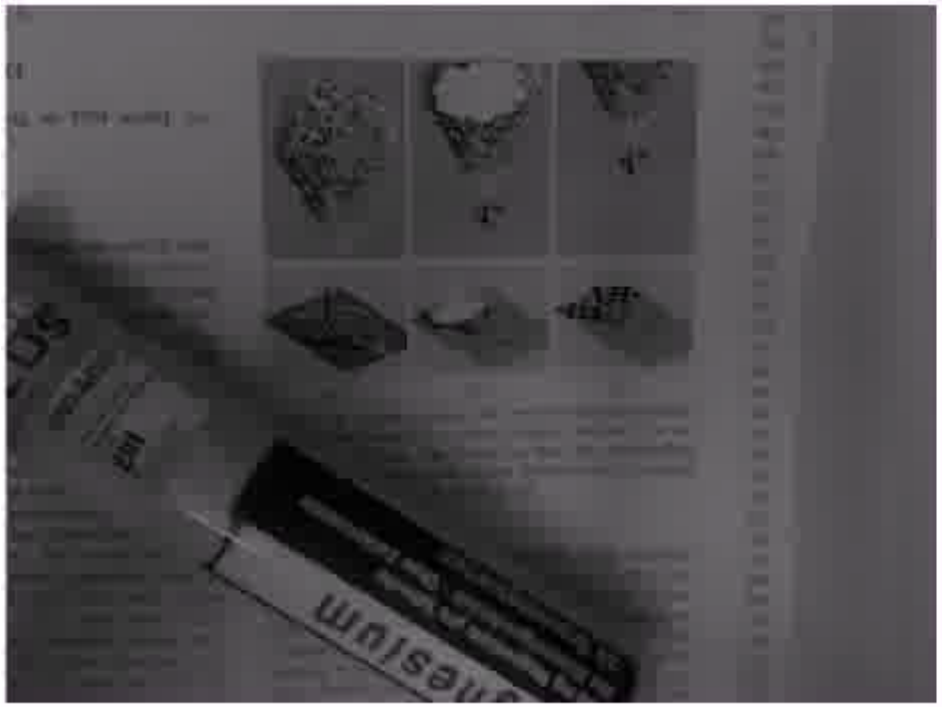
Beyond the Image

Voting Space

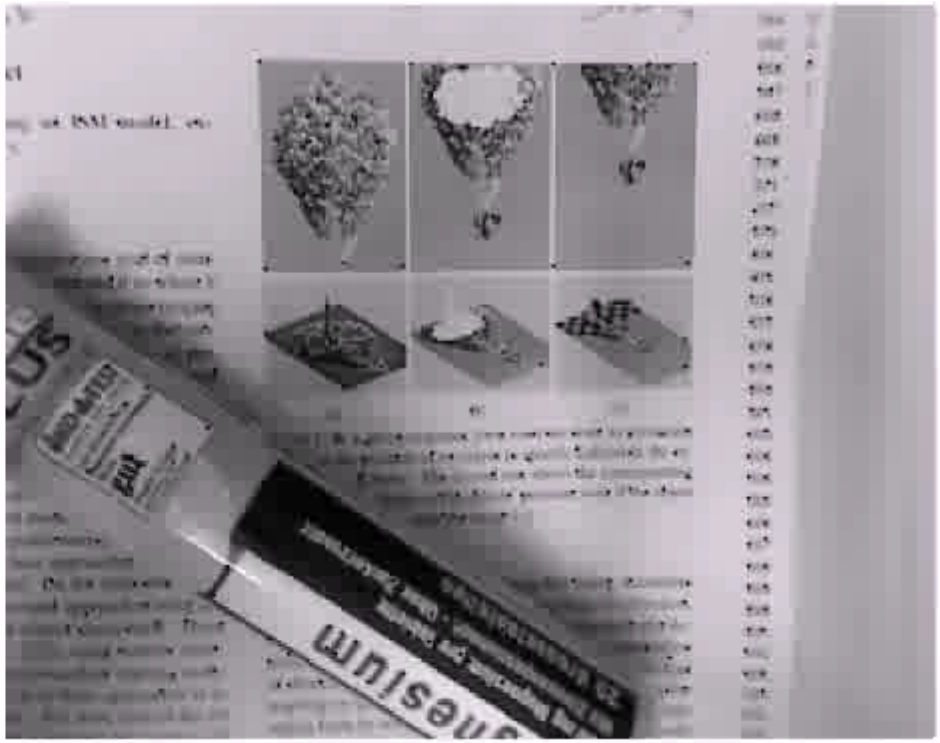
Supporters

Changing Supporter

Voting Space

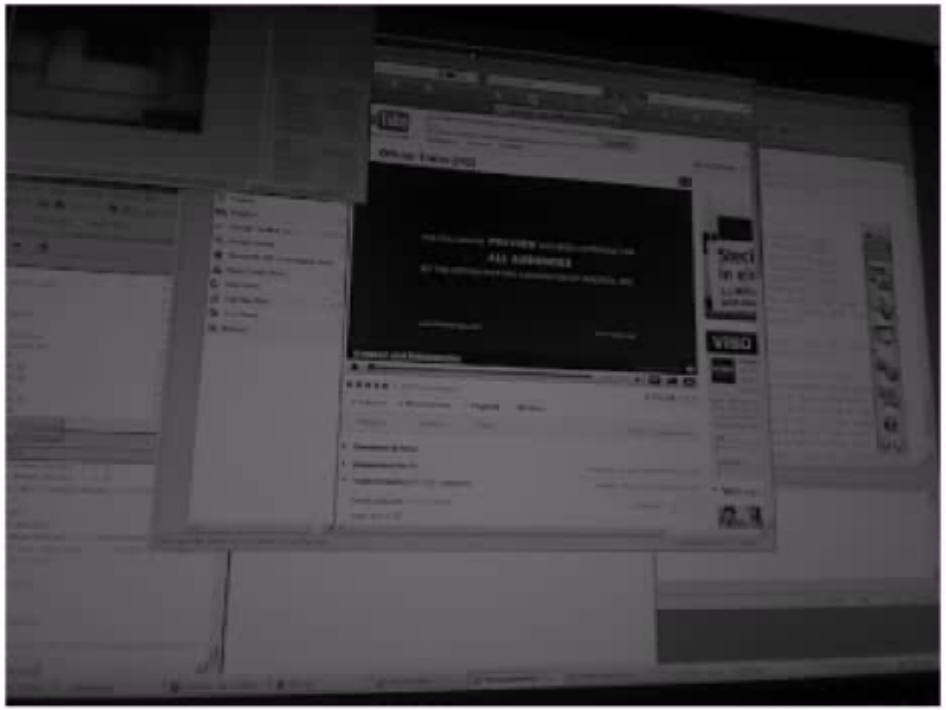


Supporters

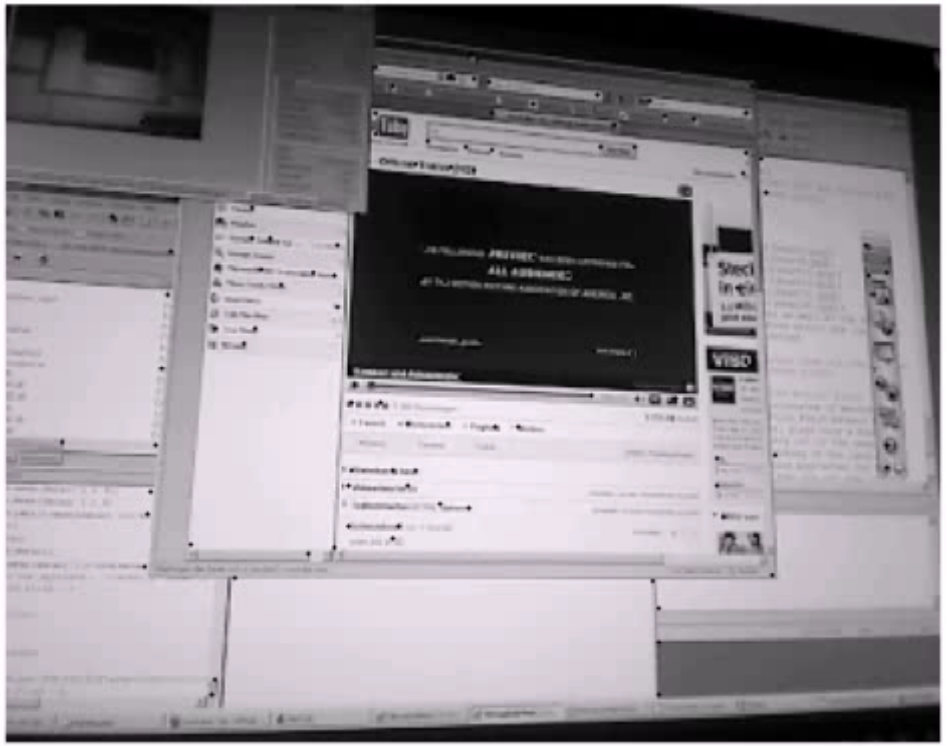


Appearance Change

Voting Space



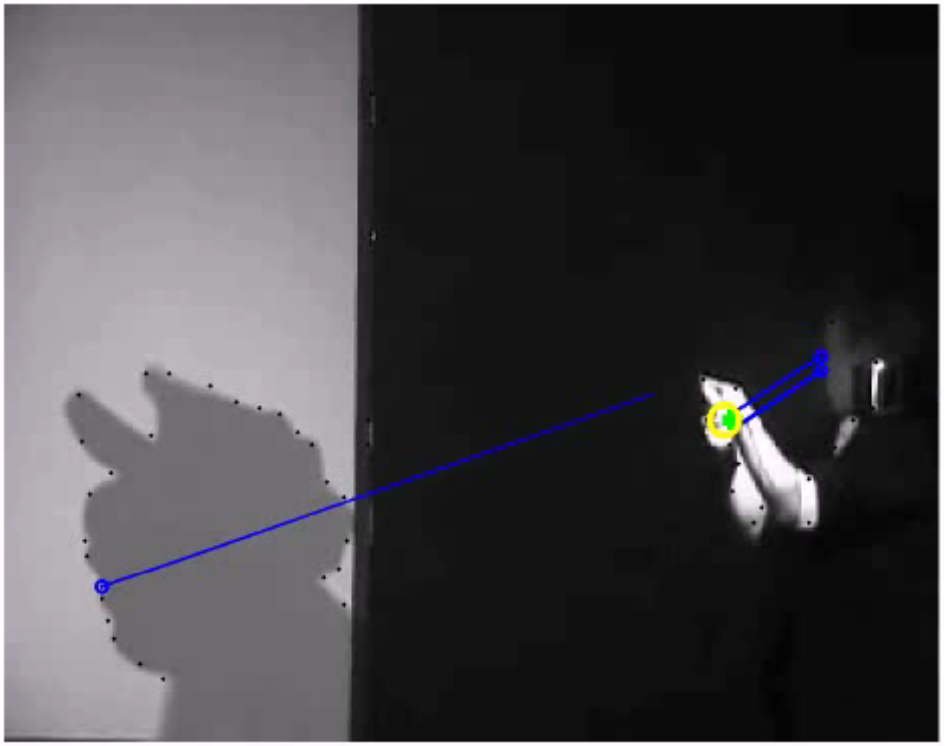
Supporters



Coupled Motion

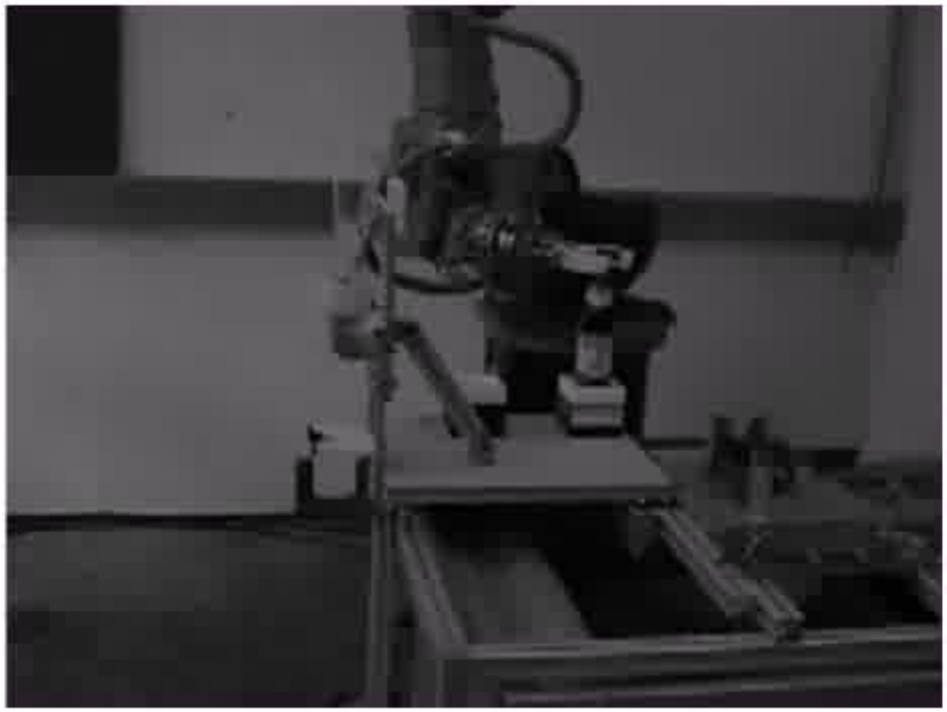
Voting Space

Supporters

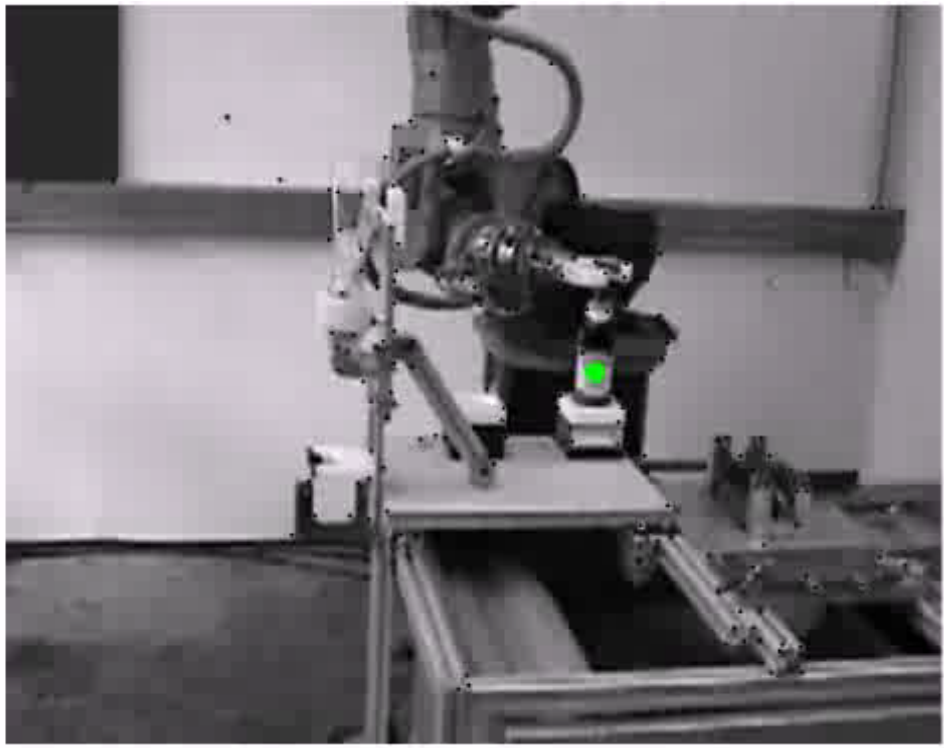


Changing Supporters

Voting Space



Supporters



Obviously, there are failure cases…

…. and magician knows that.

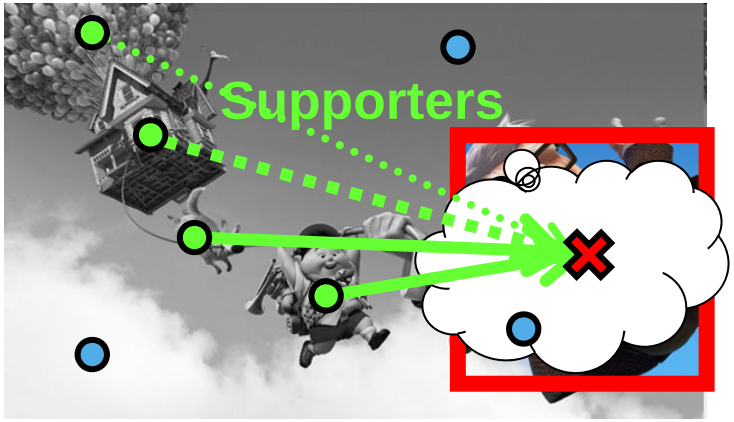
Voting Space

Supporters

Conclusions

Many dynamic relationships between parts of image, i.e., SUPPORTERS, and the object exists.

SUPPORTERS enhance tracking.



Uncertainty of motion prediction is used to maintain the set of SUPPORTERS.

Prediction is robustly combined by voting.




Look for SUPPORTERS!

They are common and take many forms.