# Learning Occlusion with Likelihoods for Visual Tracking

**Suha Kwak  Woonhyun Nam  Bohyung Han  Joon Hee Han**

POSTECH — POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

**Department of Computer Science and Engineering, POSTECH, Korea**

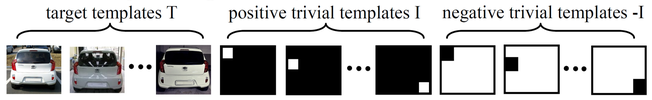
ICCV 2011

## Summary

- ***Problem:*** occlusion detection and handling for visual tracking
- ***Contribution:*** explicit occlusion detection by a single classifier, which is learnt based on likelihoods
  - Simple and active occlusion detection
  - Not perfect, but improving tracking performance significantly
  - Universal classifier for occlusion detection

## $L_1$ Minimization Tracking

- ***Particle filter + sparse representation*** **[1]**
  - Dictionary = target templates + trivial templates




  - Observation based on sparse representation

$$\boldsymbol{y} = [\mathbf{T}\ \mathbf{I}\ -\mathbf{I}]\begin{bmatrix}\mathbf{a}\\ \mathbf{e}_+\\ \mathbf{e}_-\end{bmatrix} = \mathbf{Bc}, \text{ subject to } \mathbf{c} \geq \mathbf{0}.$$

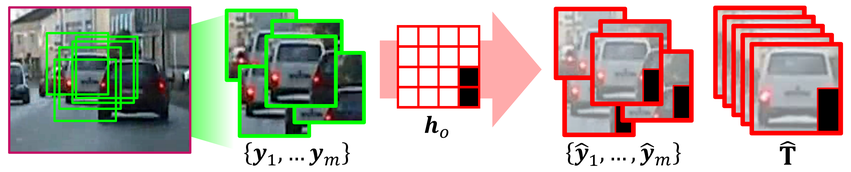
$$\mathbf{c}_i = \arg\min_{\mathbf{c}} \|\mathbf{Bc} - \boldsymbol{y}\|_2^2 + \tau\|\mathbf{c}\|_1$$

  - Observation likelihoods of samples

$$\ell(\boldsymbol{y}_i) = \exp\left(-\frac{\lambda\|\boldsymbol{y}_i - \mathbf{T}\boldsymbol{a}_i\|^2}{\dim \boldsymbol{y}_i}\right)$$

## Tracking with Occlusion Reasoning

- ***Tracking with an occlusion mask***



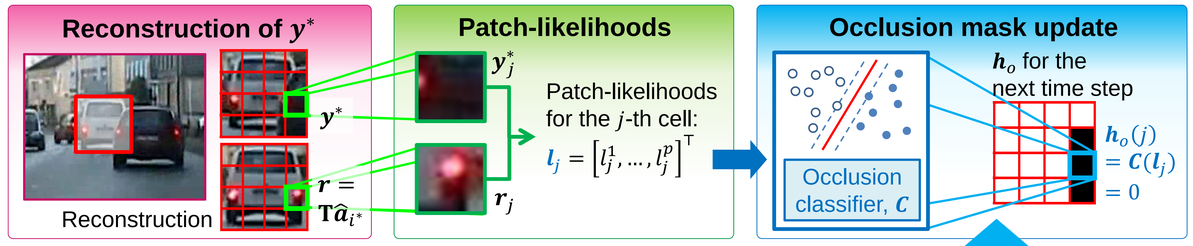

- Observation likelihoods only by unoccluded parts:

$$\ell(\hat{\boldsymbol{y}}_i) = \exp\left(-\frac{\lambda\|\hat{\boldsymbol{y}}_i - \hat{\mathbf{T}}\hat{\boldsymbol{a}}_i\|^2}{\dim \hat{\boldsymbol{y}}_i}\right)$$

- Tracking results:

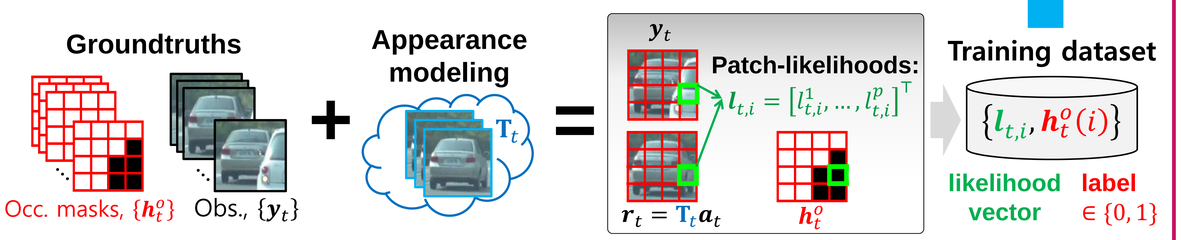
$$i^* = \arg\max_i \ell(\hat{\boldsymbol{y}}_i), \quad \boldsymbol{y}^* = \boldsymbol{y}_{i^*}$$

- ***Occlusion reasoning by classification***

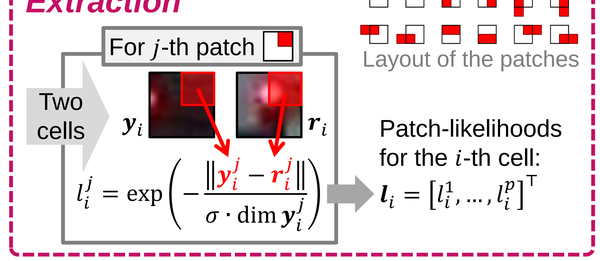



1: Init. $\mathbf{T}$, $\mathbf{h}_s$ (degeneracy mask), $\mathbf{h}_o$ (occlusion mask)
2: $\mathbf{h} \leftarrow \mathbf{h}_s \vee \mathbf{h}_o$, $\gamma \leftarrow \sum_{i=1}^{q} \mathbf{h}(i)/q$.
3: **while** the target is in the scene **do**
4:  **if** $\gamma < \gamma_1$ **then**
5:  Apply $\mathbf{h}$ to $\{\mathbf{y}_1, \dots, \mathbf{y}_m\} \rightarrow \{\hat{\mathbf{y}}_1, \dots, \hat{\mathbf{y}}_m\}$.
6:  Apply $\mathbf{h}$ to $\mathbf{T} \rightarrow \hat{\mathbf{T}} = [\hat{\mathbf{t}}_1, \dots, \hat{\mathbf{t}}_n]$.
7:  $[\mathbf{y}^*, \hat{\mathbf{c}}] \leftarrow L_1$_Track$(\{\hat{\mathbf{y}}_1, \dots, \hat{\mathbf{y}}_m\}, \hat{\mathbf{T}})$.
8:  Compute $\{\mathbf{l}_1, \dots, \mathbf{l}_q\}$ based on $\mathbf{y}^*$ and $\mathbf{T}\hat{\mathbf{a}}_{i^*}$.
9:  $\mathbf{h}_o(i) \leftarrow C(\mathbf{l}_i),\ \forall i = 1, \dots, q$.
10:  The occlusion ratio $\gamma_o \leftarrow \sum_{i=1}^{q} \mathbf{h}_o(i)/q$.
11:  **if** $\|\mathbf{y}^* - \mathbf{t}_{\text{argmax}_i \hat{\mathbf{a}}_i}\| > \xi$ and $\gamma_o < \gamma_2$ **then**
12:  $\mathbf{T} \leftarrow$ Template_Update$(\mathbf{y}^*, \mathbf{T})$.
13:  Compute $\mathbf{h}_s$ from $\mathbf{y}^*$.
14:  **end if**
15:  **else**
16:  Predict the target state $\mathbf{x}_o$ with motion prior.
17:  $[\mathbf{y}_o, \mathbf{c}_o] \leftarrow L_1$_Track$(\{\mathbf{y}_o\}, \mathbf{T})$.
18:  Compute $\{\mathbf{l}_1, \dots, \mathbf{l}_q\}$ based on $\mathbf{y}_o$ and $\mathbf{Ta}_o$.
19:  $\mathbf{h}_o(i) \leftarrow C(\mathbf{l}_i),\ \forall i = 1, \dots, q$.
20:  **end if**
21:  $\mathbf{h} \leftarrow \mathbf{h}_s \vee \mathbf{h}_o$, $\gamma \leftarrow \sum_{i=1}^{q} \mathbf{h}(i)/q$.
22: **end while**

## Learning Occlusion with Likelihoods

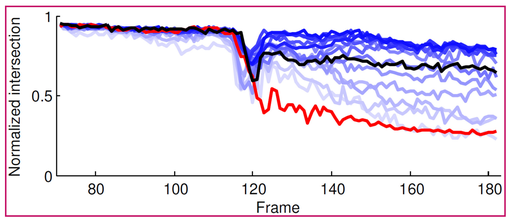


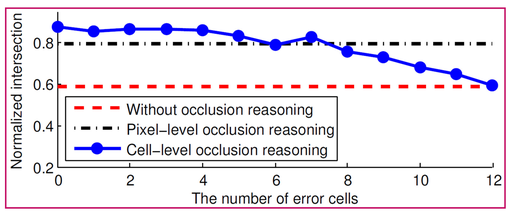

### *Patch-likelihoods Extraction*




## Experiments

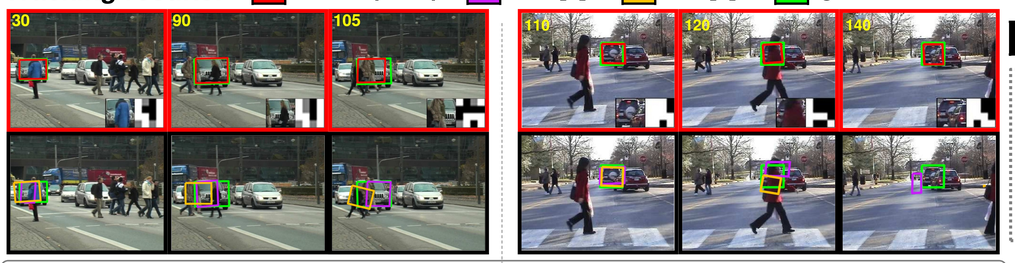
- ***Simulation***



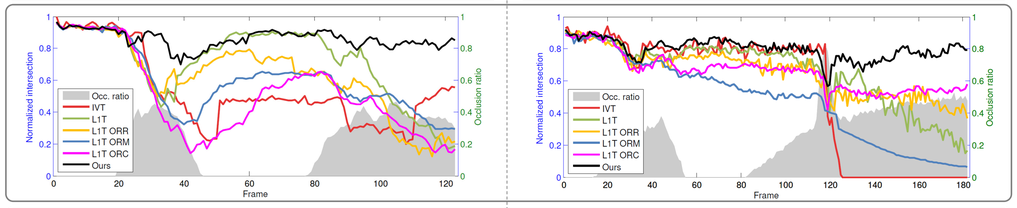

- $L_1$ minimization tracking + groundtruth occlusion masks with simulated errors
- Decent occlusion reasoning improved tracking performance significantly.

- ***Tracking results:*** L1TOR(Ours) L1T [1] IVT [2] groundtruth

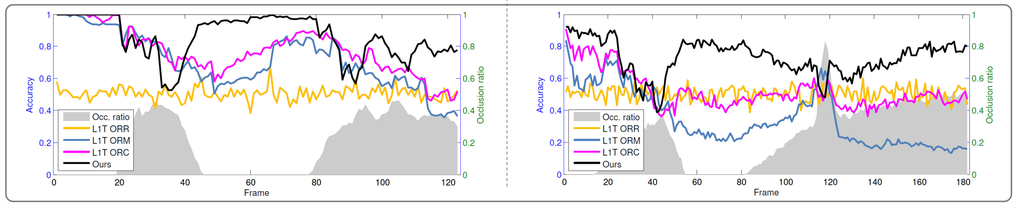



occlusion degeneracy

[1] Mei and Ling, *Robust visual tracking using $L_1$ minimization.* In ICCV'09.
[2] Ross et al., *Adaptive probabilistic visual tracking with incremental subspace update.* In ECCV'04.

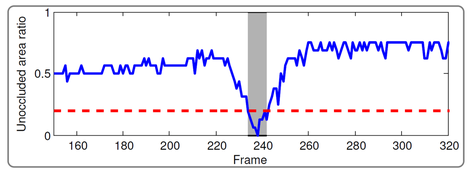



**The quantitative comparisons of 6 tracking algorithms over time**




**The quantitative comparisons of 4 occlusion reasoning algorithms over time:**
- ORR – random guessing
- ORM – one cell likelihood
- ORC – more conservative ORM

- ***Tracking with full occlusion***


- ***Our occlusion reasoning + IVT*** **[2]**