TRACKING LEARNING DETECTION

DEMO

OBJECTIVES

Our goal is **long-term, real-time tracking** of arbitrary objects. The object is defined by a region of interest in a single frame. The video sequence is unconstrained, the object might significantly change appearance, get partially or fully occluded or move in and out of the field of view.

MOTIVATION

Long-term tracking of arbitrary objects is a the core problem in many computer vision applications: surveillance, object auto-focus, SLAM, games, HCI, video annotation.

CHALLENGES

Real-time performance, partial and full occlusions, illumination changes, large displacements, background clutter, similar objects, low video quality.

THE APPROACH

Decomposition of the long-term tracking task into three components: tracking, learning and detection (TLD). Each of these components deals with different aspect of the problem, the components are running in parallel and are combined in a synergetic manner to suppress their drawbacks.

AVAILABLE ONLINE

The demo application, sequences with ground truth.
http://info.ee.surrey.ac.uk/Personal/Z.Kalal/

FUTURE WORK

Document the code and make it publically available. Automatic initialization, test different tracker and detector, eliminate planarity assumption, explicitly handle out-of-plane rotation, track multiple targets, learn shape.

UNIVERSITY OF SURREY

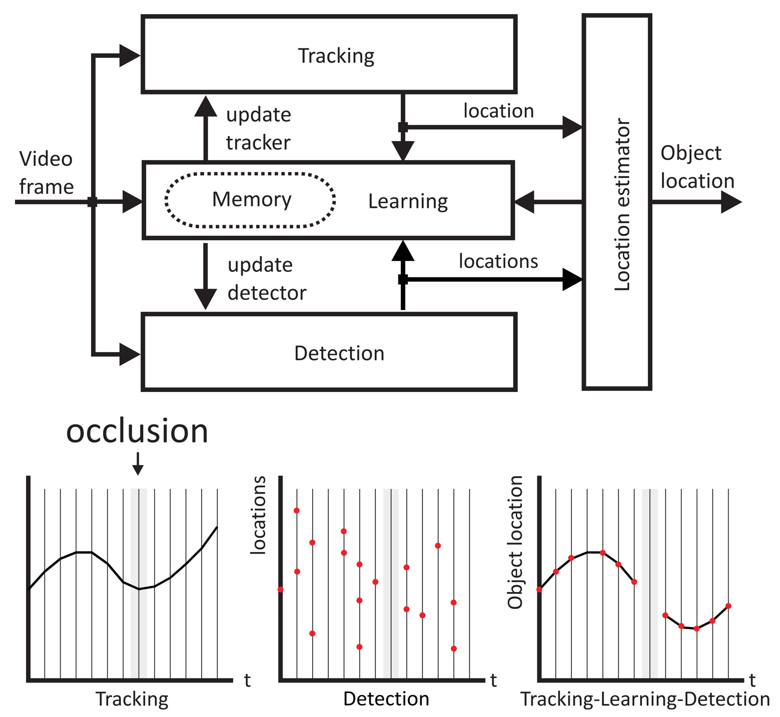
Zdenek **Kalal**, Krystian **Mikolajczyk**, Jiri **Matas**
zdenek.kalal@gmail.com

TLD

A framework addressing long-term tracking. TLD trains a detector of an object after initialization from a single patch and its warps.

The tracker and the detector are running in parallel and both contribute to estimated location of the object. "Not visible" is possible output.

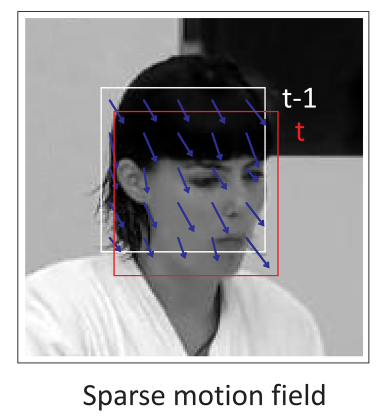
Updates of the tracker and the detector depends on the learning module described below.



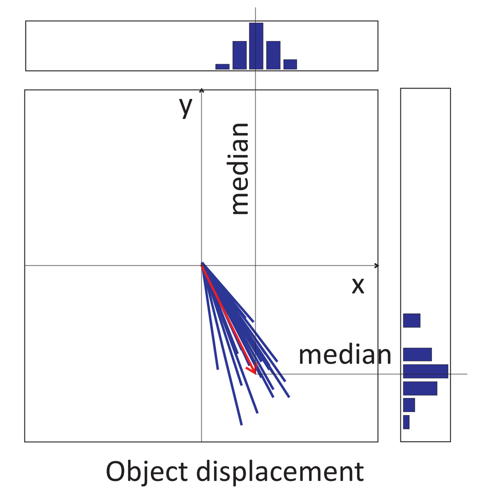
TRACKING

Median-shift tracker - tracker of a rectangle, based on the Lucas-Kanade tracker, robust to partial occlusions. Estimates translation and scale.

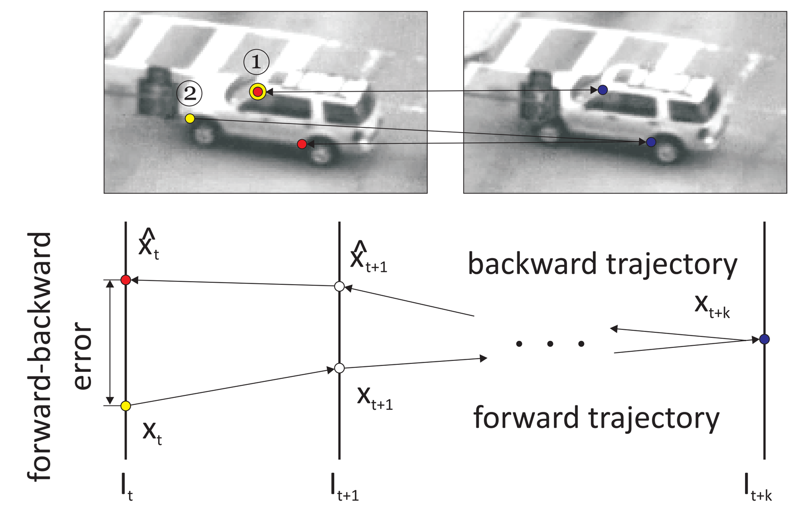
Tracker validation - detector is updated as long as the trajectory is forward-backward consistent.



Sparse motion field

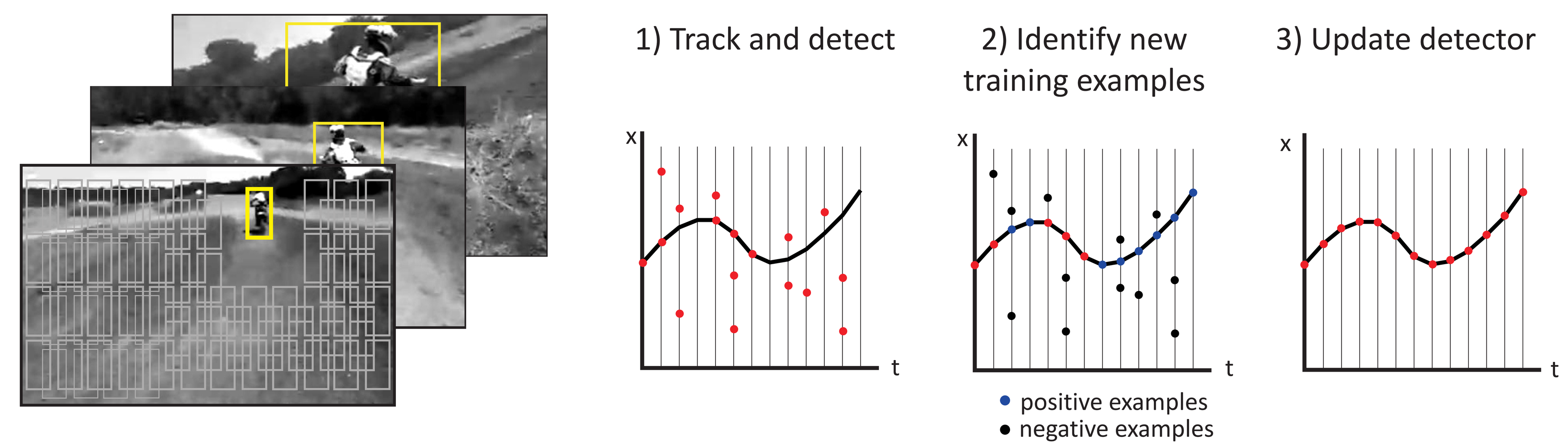


Object displacement



LEARNING

The learning is implemented withing the P-N Learning framework. Object is tracked by a tracker. Patches close to the trajectory update the detector with positive label (P-consaints). The object is detected by the detector, non-maximaly confident detections update the detector with negative label (N-constraints). Both constraints make errors, the learning stability is achieved by their mutual compensation.



DETECTION

1st stage filter:
Randomized forest, 2bitBP features

2nd stage classifier:
1-NN, 10x10 patch, NCC
Confidence = $d^- / (d^- + d^+)$