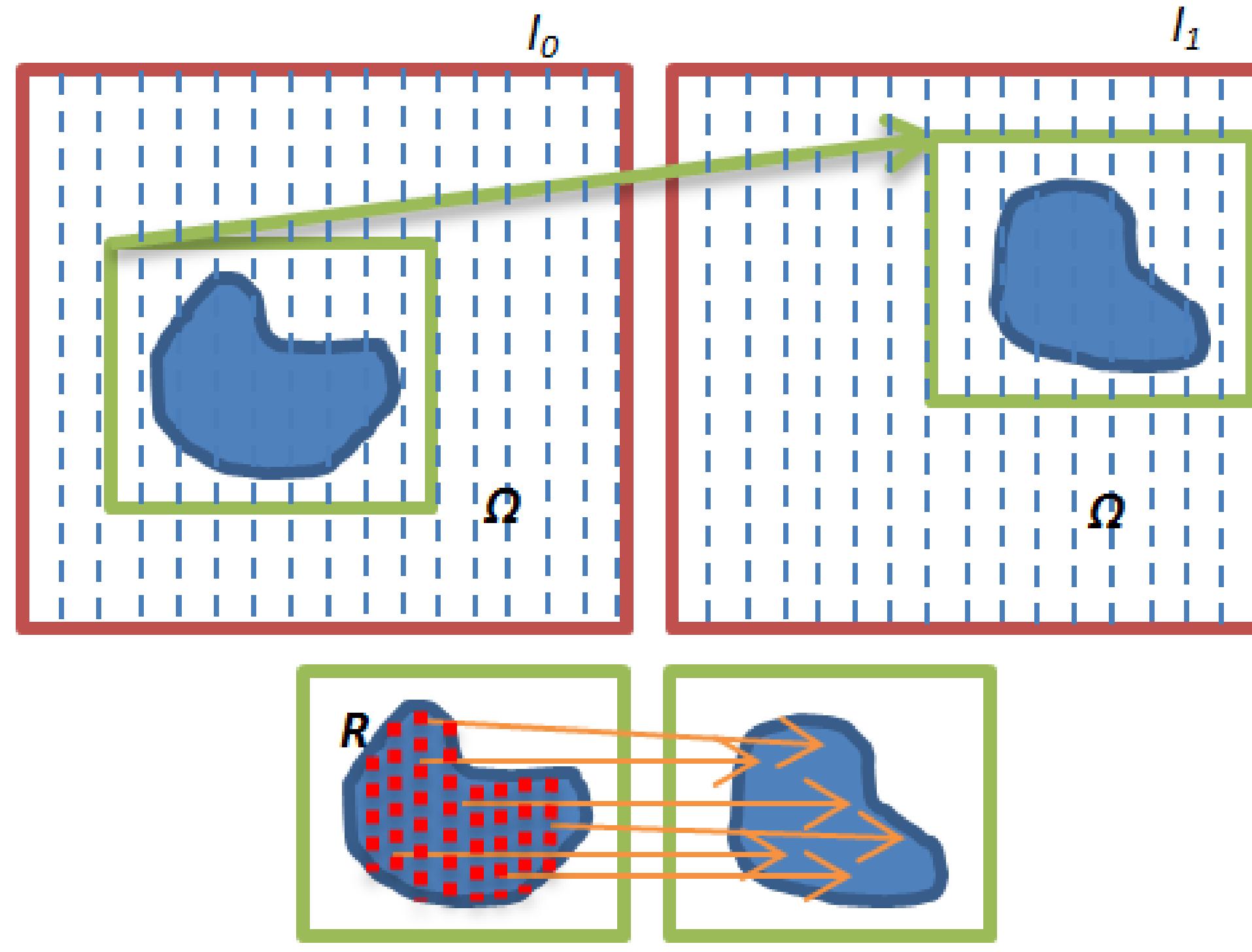


# OBJECT FLOW: A PER-OBJECT DENSE MOTION DESCRIPTOR

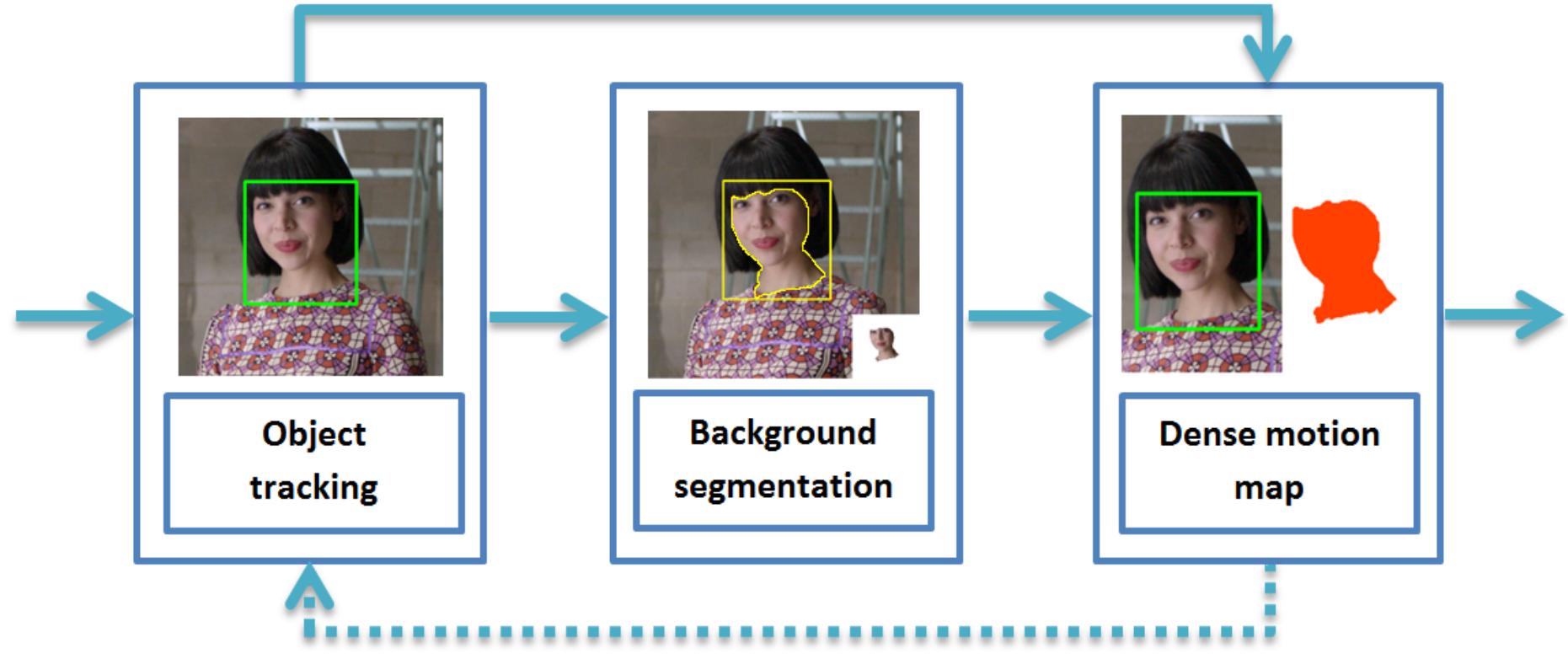
JUAN M. PEREZ RUA, TOMAS CRIVELLI AND PATRICK PEREZ

## 1. DEFINITION

Given a video sequence  $I_t, t : 0..N - 1$ , and an initial bounding box loosely indicating the position of an object in  $I_0$ . Let  $\mathcal{R} \in \Omega$  be the region corresponding to the support of the object in the bi-dimensional grid  $\Omega$ . Then, the **object flow**,  $\mathcal{O}(x)$ , is defined as  $\mathcal{O}(x) = d_{0,t}(x), \forall x \in \mathcal{R}$ . With  $d_{0,t}(x)$  a displacement vector between the frame 0 and the frame  $t$  for the pixel  $x$ .



## 2. PIPELINE

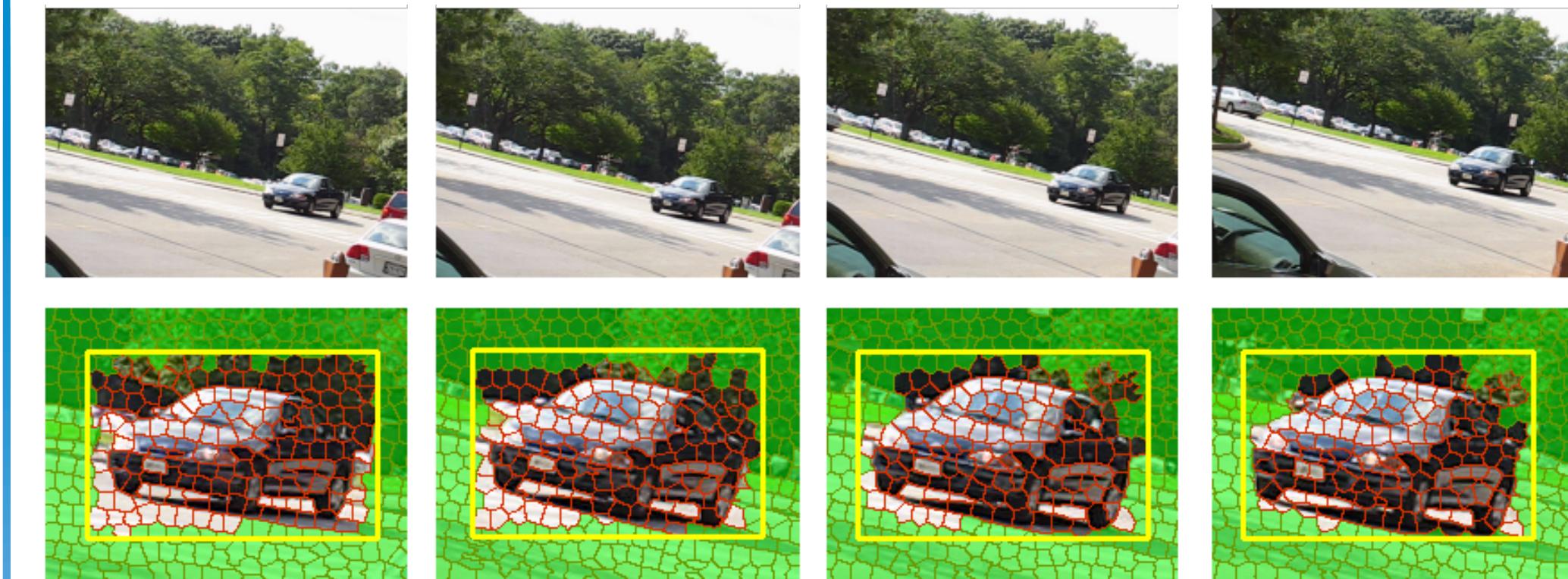


**Tracking.** Tracking-by-detection methods are reliable and can be improved by context information [4].

**Segmentation.** Video dynamics and tracker window can be used to estimate object boundaries.

**Flow estimation.** Strong smoothness assumption is a valid prior within the estimated object boundaries [2], and can be used to complete the global motion estimation given by the object tracker.

## 3. BACKGROUND TRACKING



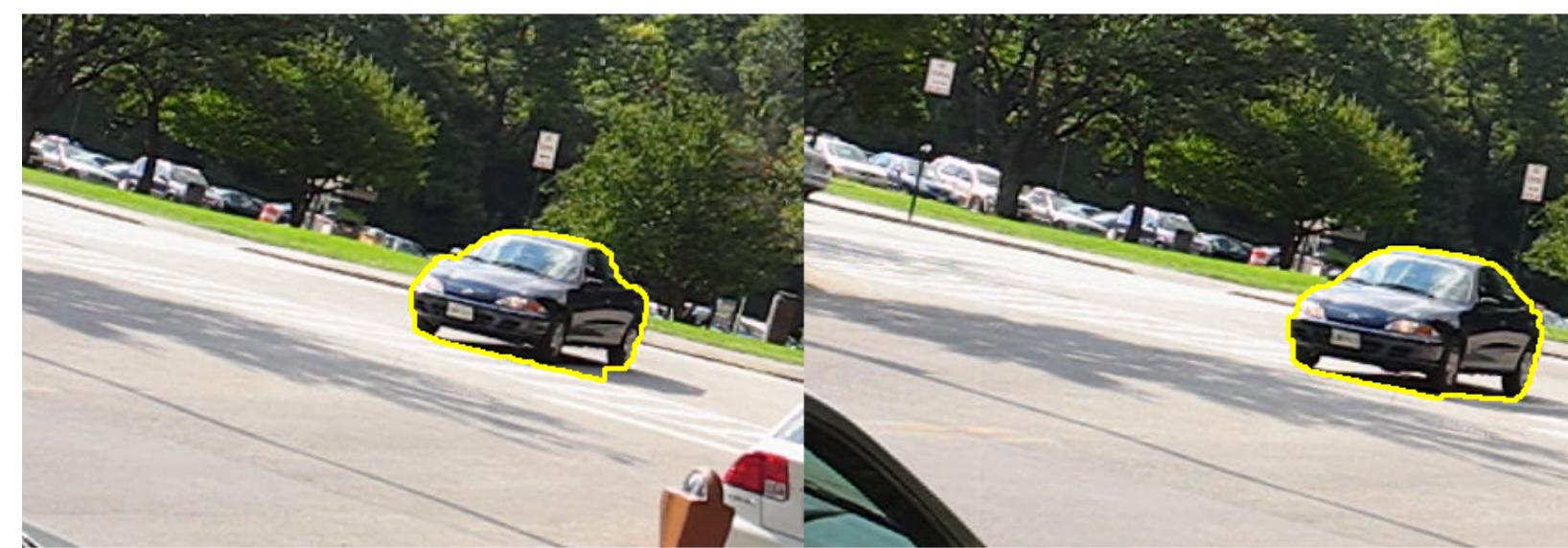
Obvious background superpixels (Outside tracking window) are labeled and tracked with the **superpixel flow**. The superpixel flow is an energy based matching method for superpixels:

$$E(l) = \sum_{p \in \Gamma} D_p(l_p; I_0, I_1) + \sum_{(p,q):q \in \mathcal{N}_r} S_{p,q}(l_p, l_q),$$

with,  $D_p(l_p; I_0, I_1) = \rho(\mathbf{h}(p), \mathbf{h}(p'))$ ,

and,  $S_{p,q}(l_p, l_q) = \lambda(p) \sqrt{\frac{|u_{p_c} - u_{q_c}|}{\|p_c - q_c\|}} + \frac{|v_{p_c} - v_{q_c}|}{\|p_c - q_c\|}$ .

The mask is finally refined with [3].

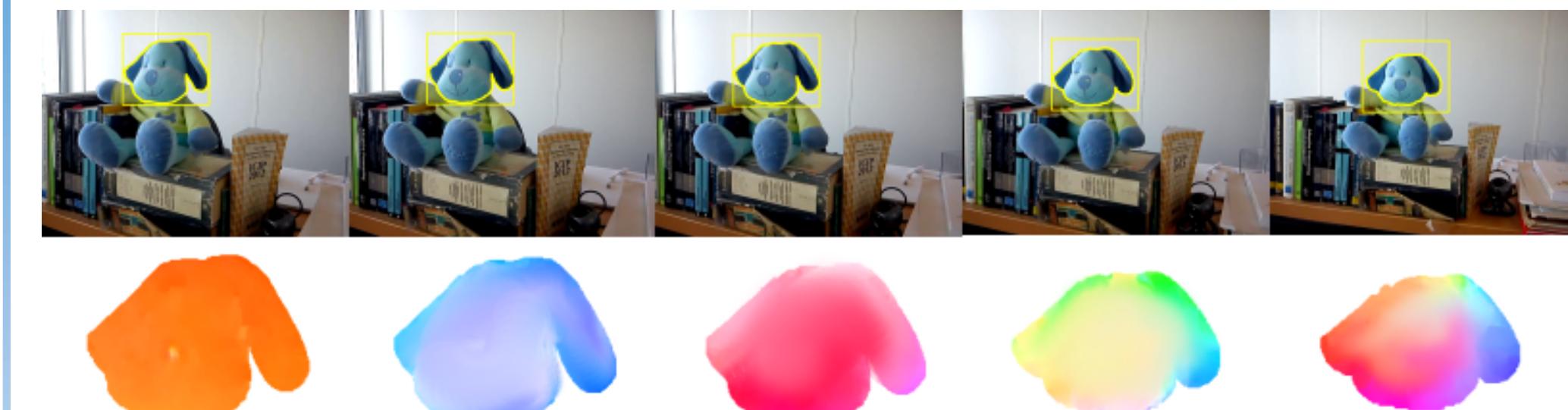


## 4. MOTION ESTIMATION

**SimpleFlow** method [1] is modified to use segmentation masks ( $S_0$  and  $S_1$ ) as smoothness boundaries.

The flow vector  $(u, v)$  has to explain motion in the pixel  $(x_0, y_0)$  and also in its neighborhood  $\mathcal{N}_0$ ,

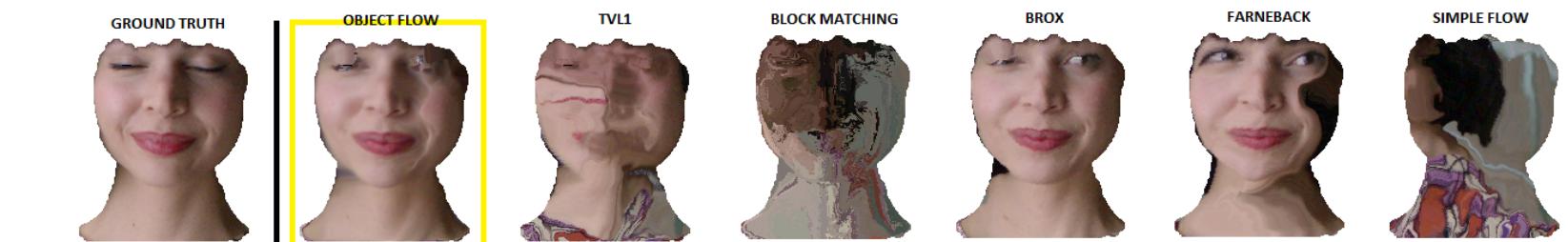
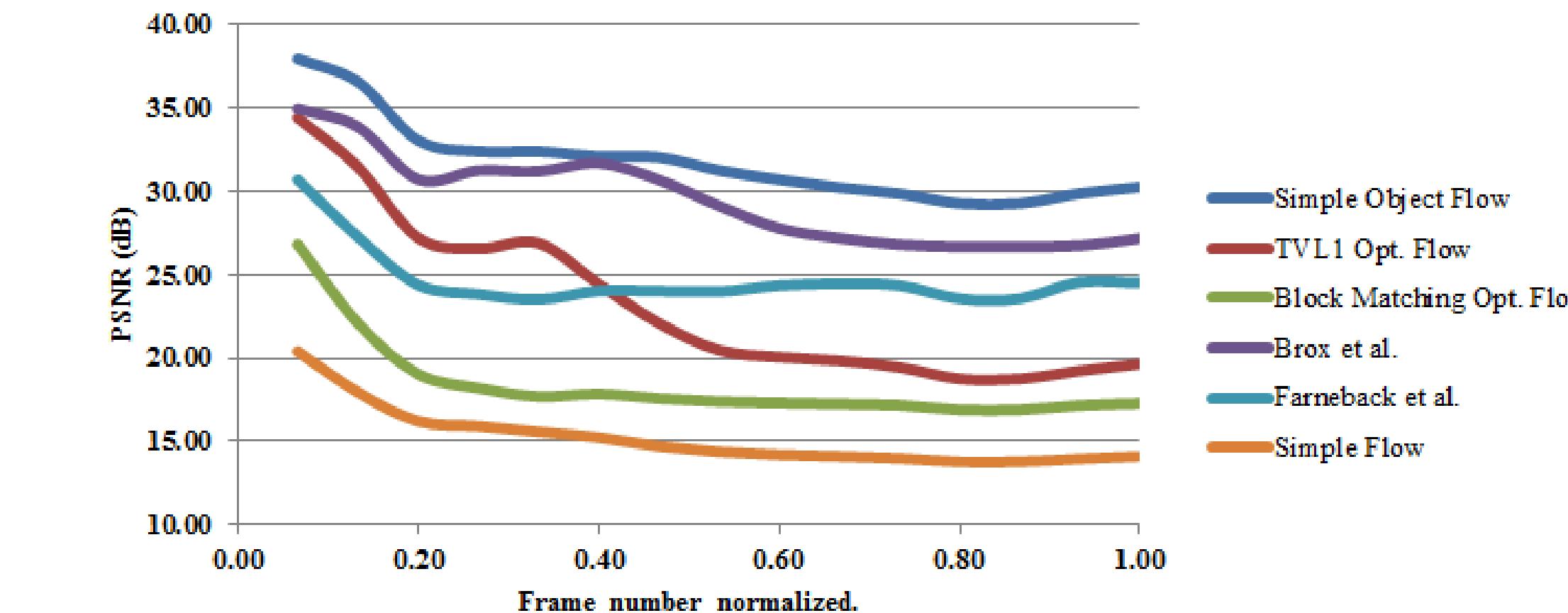
$$E(x_0, y_0, u, v) = \sum_{(x,y) \in \mathcal{N}_0} w_d w_c \|I_0(x, y) - I_1(x+u, y+v)\|^2, \forall (x, y \in S_0), (x+u, y+v \in S_1).$$



The energy is computed with a cross-bilateral filter, and the motion  $(u, v)$  is found as the one that minimizes it in its local region. Motion is more deeply regularized with a final Gaussian filter applied only inside these boundaries.

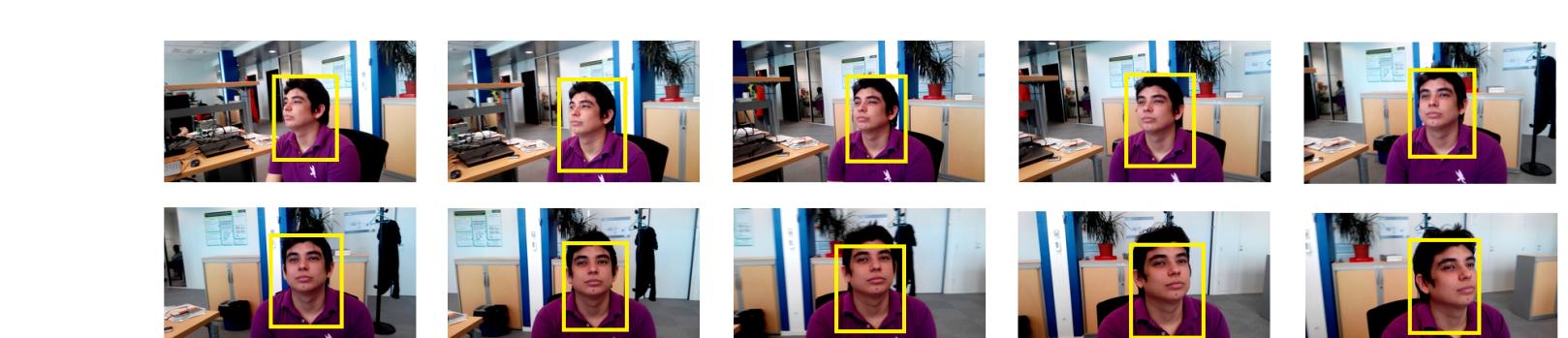
## 5. RESULTS

The object flows captures the motion details much better than globally computed optical flow methods.



## 6. APPLICATIONS

Video editing and Object-SfM can use the object flow for precise and dense results.



## 7. CONTRIBUTIONS

Our main contributions are

1. A method to compute superpixel matchings (The superpixel flow).
2. A segmentation method for objects in video (Background regions tracking).
3. A framework to combine object trackers and optical flow methods.
4. A method to extend the Simple Flow method to use input segmentation masks.
5. A better sampling method for tracking-by-detection methods.

## 8. BIBLIOGRAPHY

- [1] M. Tao, J. Bai, P. Kohli, and S. Paris. SimpleFlow: A Non-iterative, Sublinear Optical Flow Algorithm, *Eurographics '12*
- [2] S. Baker, D. Scharstein, J.P. Lewis, S. Roth, M.J. Black and R. Szeliski. A Database and Evaluation Methodology for Optical Flow, *IJCV '13*
- [3] Y. Boykov, M.-P. Jolly. Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D images, *ICCV '13*
- [4] Y. Wu, J. Lim and M.-H. Yang. Online object tracking: A benchmark, *CVPR '13*