A Multilayer-Based Framework for Online Background Subtraction with Freely Moving Cameras

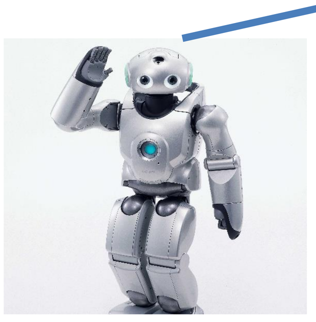
Yizhe Zhu Ahmed Elgammal
Rutgers University, Department of Computer Science
yizhe.zhu@rutgers.edu, elgammal@cs.rutgers.edu

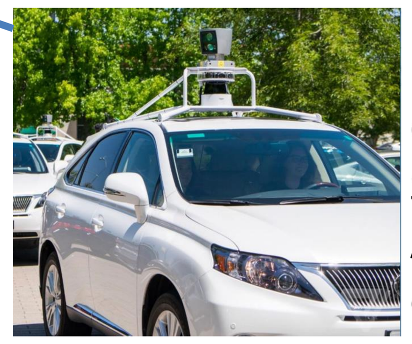
Paper & code here!

Venice, Italy
October 22-29, 2017

Motivation

- An increasing amount of videos are captured from **moving platforms**





Robots,
Cell phones,
Self-driving cars,
Ariel drones,
etc.

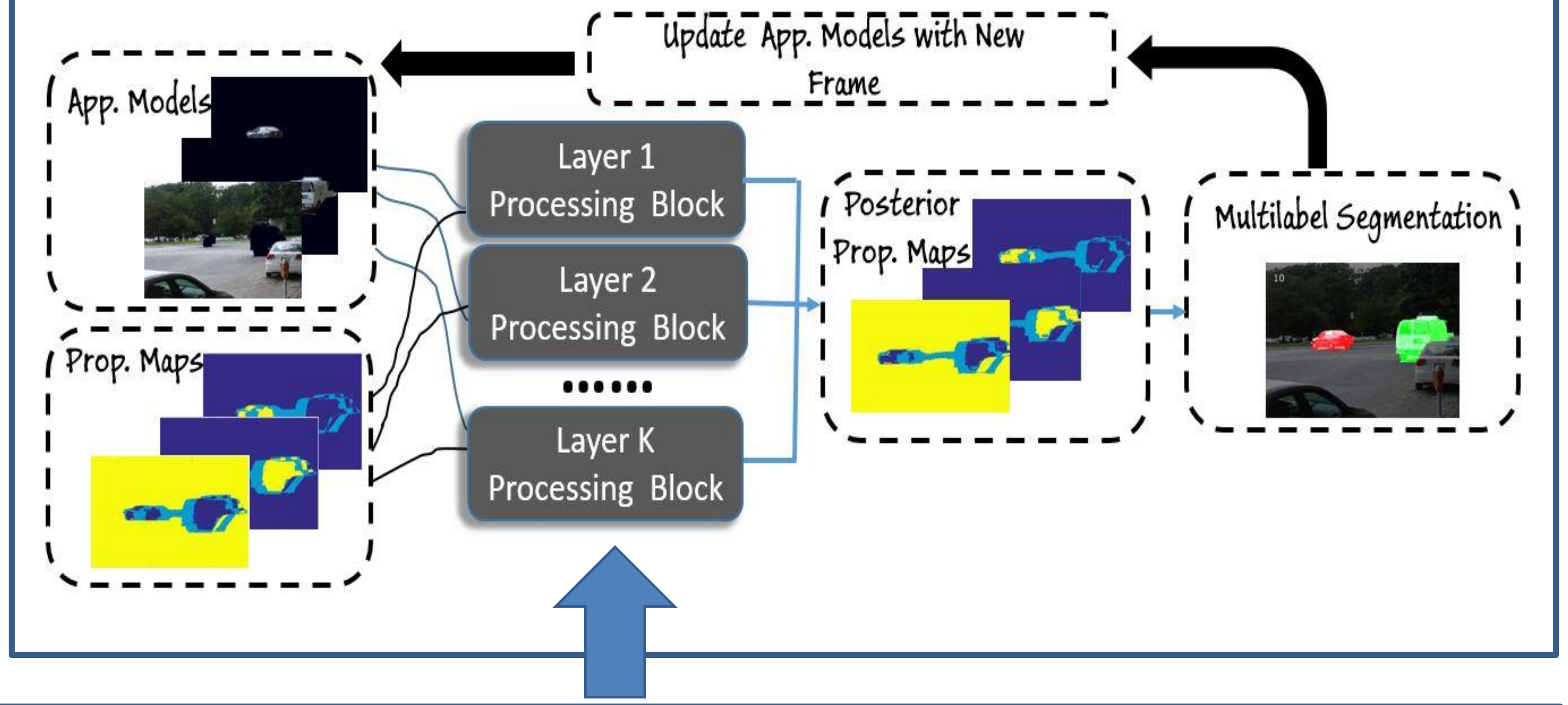
- Videos need to be analyzed as they being streamed in real time.
- Most general Background Subtraction methods are restricted to binary segmentation.

Contributions

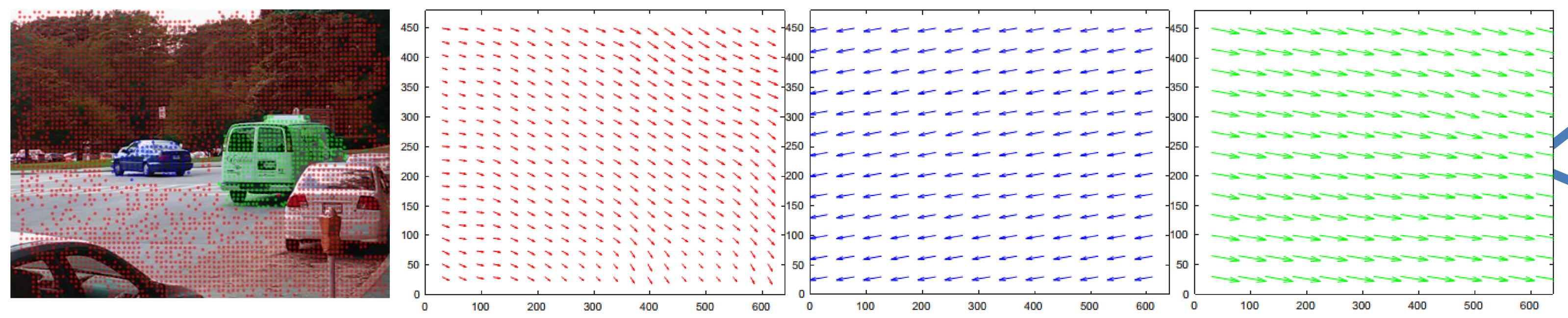
- We formulate Background Subtraction as an online multi-label segmentation problem by modeling multiple foreground objects in different layers.
- We design an processing block to handle the information in each layer separately and simultaneously.

Our Proposed Framework

- The accumulated appearance models and probability maps of each layer are fed to "processing blocks".
- With the collection of probability maps produced by each "processing block", we use Multi-label Graphcut to carry out the segmentation.
- Add the new color feature (RGB) in the current frame to update the corresponding appearance model.



Processing Block



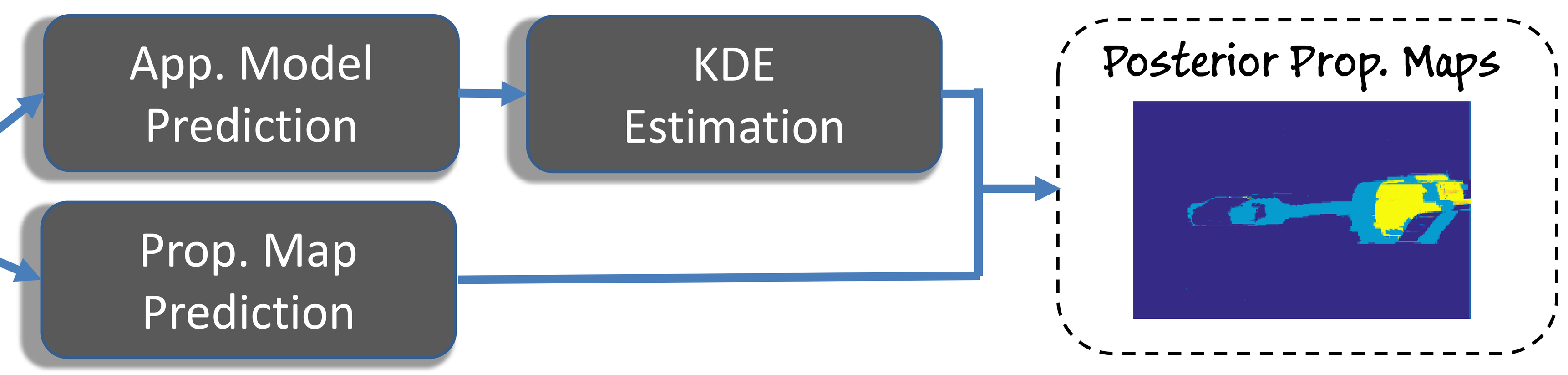
Trajectory labeling:

- Recursive normalized cuts to get the initial labels of trajectories.
- Label propagation to infer labels in the following frames.

Motion estimation:

- In each layer, Motion estimation is performed based on Gaussian Belief Propagation with motion of trajectories as the evidence.

$$P(\mathcal{M}_t^k|\mathcal{P}_t^k) \propto \prod_{(i,j)\in\varepsilon} \Psi(m_{k,t}^i, m_{k,t}^j) \prod_{i\in S_{k,t}} \Phi(m_{k,t}^i)$$



- Shift the appearance model and the prior probability from the previous one with the estimated motion model.
- Estimate Posterior map

$$P_{post}(l_t^i|I_t^i,\mathcal{P}_t) \propto P(I_t^i|l_t^i)P_{prior}(l_t^i|\mathcal{P}_t)$$

Likelihood estimated by KDE with the appearance model $\frac{1}{N}\sum_{f=1}^{N} K_G(I_t^i - I_f^i)$

Prior propagated from last frame

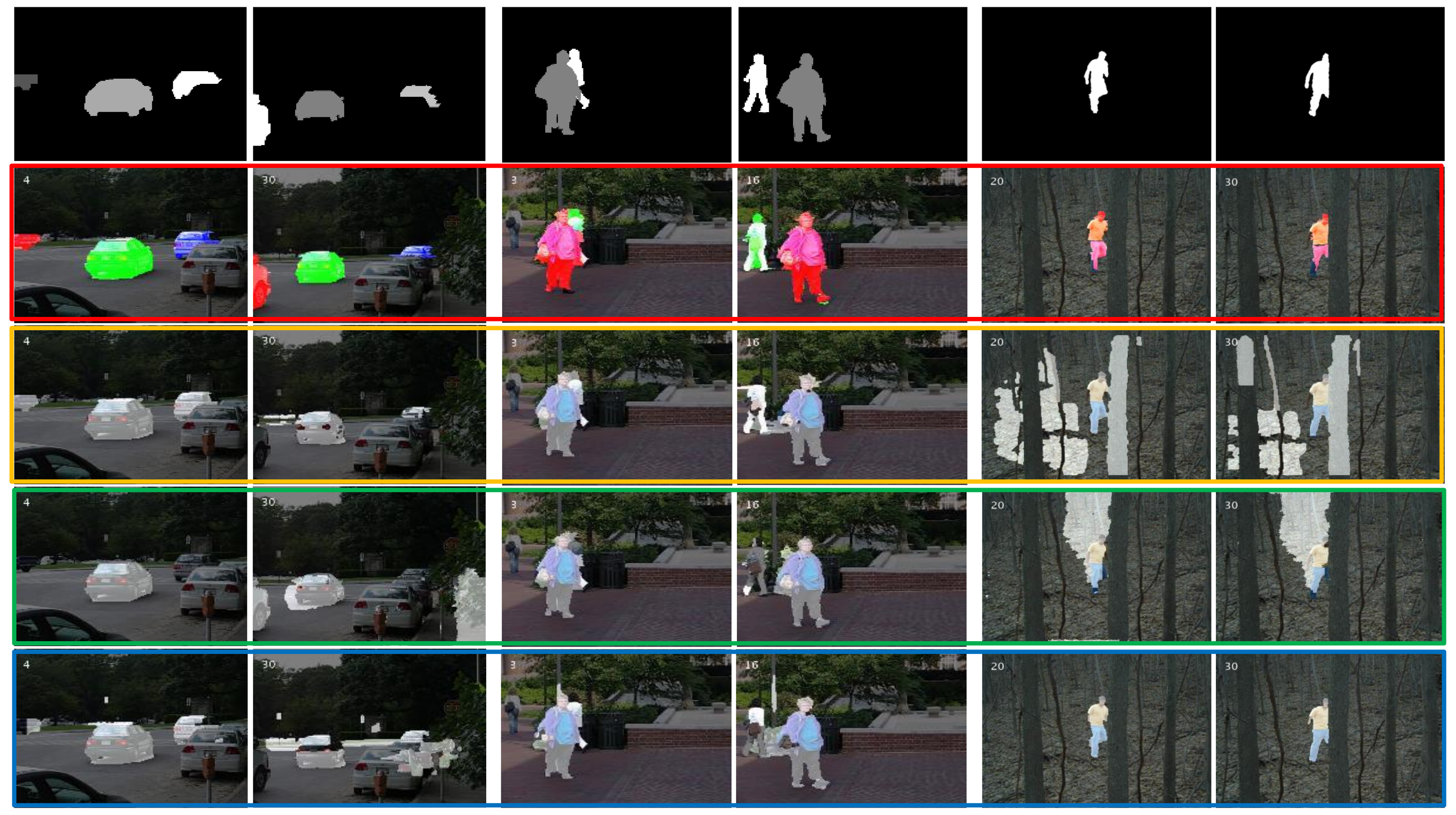
Results Compared with State-of-the-art

- Two-label Background Subtraction

Table 1: Two-label background subtraction performance comparison on the videos with different numbers of moving objects.

	MLBS	GBSSP[1]	FOF[2]	OMCBS[3]	GBS[4]	BSFMC[5]
1	**84.05**	80.70	59.20	54.22	47.82	31.20
2	**91.14**	81.51	83.19	71.48	80.30	62.75
≥ 3	**74.99**	65.05	66.26	57.02	55.72	45.05

Figure 1: Qualitative comparisons. Ours , GBS , OMCBS , GBSSP

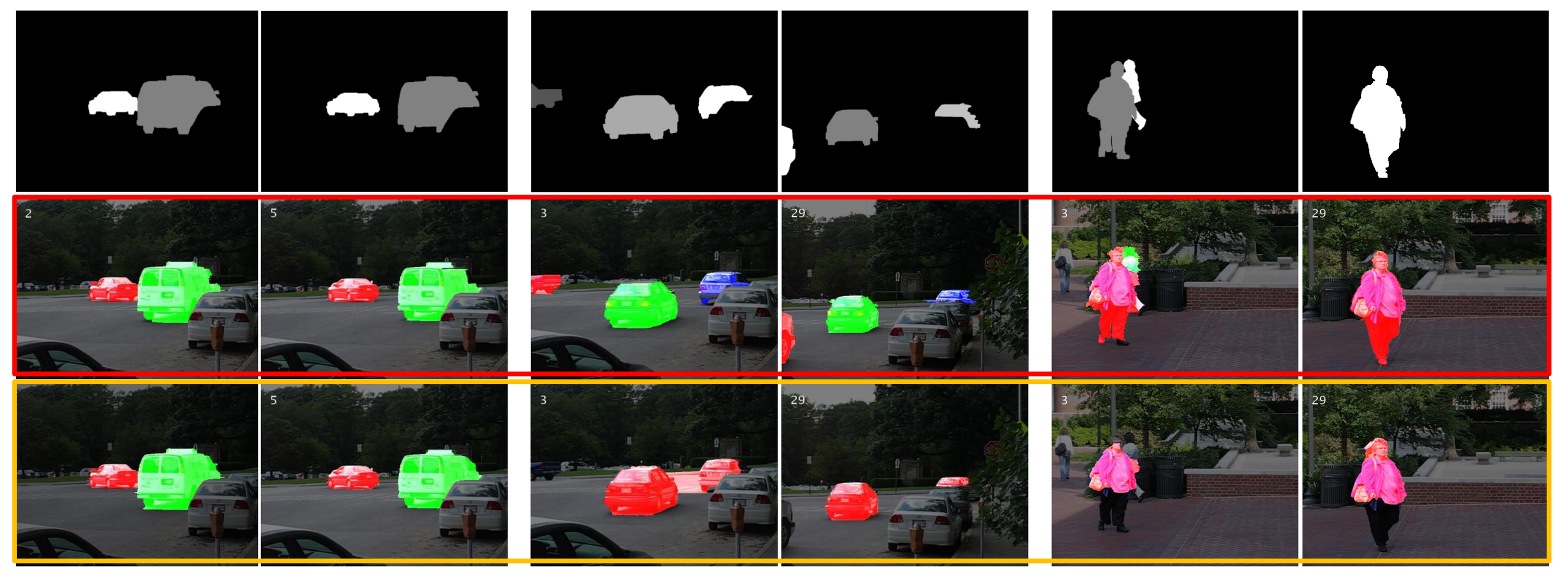


- Multi-label Background Subtraction

Table2: Performance comparison of Multi-foreground segmentation on Hopkins Dataset.

		Precision	Recall	F-Score
First 10 Frames	ours	89.79	**83.00**	**85.06**
	SLT	**90.70**	78.94	83.41
	Baseline	61.23	62.32	61.35
All Frames	ours	88.60	**85.05**	**86.16**
	SLT	**89.51**	78.24	82.00
	Baseline	63.63	66.22	64.35

Figure 2: The Multilabel segmentation performance comparison with SLT[6]



[1] Lim *et al*. Generalized background subtraction using superpixels with label integrated motion estimation. ECCV'14
[2] Narayana *et al*. Coherent motion segmentation in moving camera videos using optical flow orientations. ICCV'13.
[3] Elqursh *et al*. Online moving camera background subtraction ECCV'12
[4] Kwak *et al*. Generalized background subtraction based on hybrid inference by belief propagation and bayesian filtering. ICCV'11
[5] Sheikh *et al*. Background subtraction for freely moving cameras. ICCV'09
[6] Ochs *et al*. Segmentation of moving objects by long term video analysis. TPAMI'14