Learning to Find Occlusion Regions

Ahmad Humayun Oisin Mac Aodha Gabriel J. Brostow
University College London
http://visual.cs.ucl.ac.uk/pubs/learningOcclusion/

Presented at 24th **IEEE Conference on Computer Vision and Pattern Recognition**. Colorado Springs, USA. June 20-25, 2011

Contents

1 Quantitative Results **2**

2 Qualitative Results on Stein and Hebert [39] **3**

3 Qualitative Results on Lobaton *et al.* [27] **and Sigal *et al.*** [$PAMI2004$] **5**

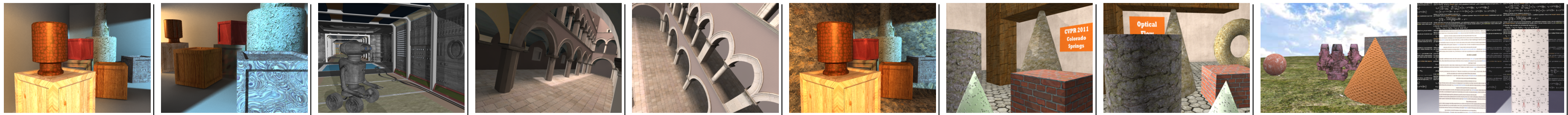
4 Full vs. Lean Results **7**

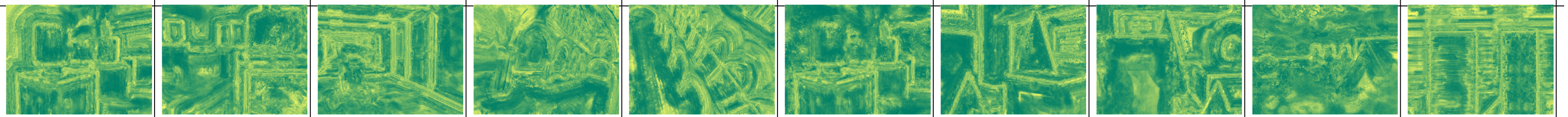
5 Angle vs. Texture **8**

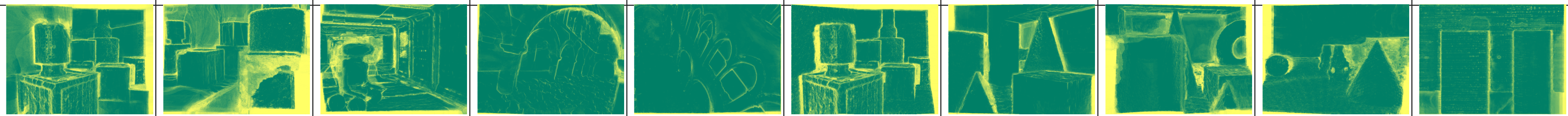
6 Superpixels **9**

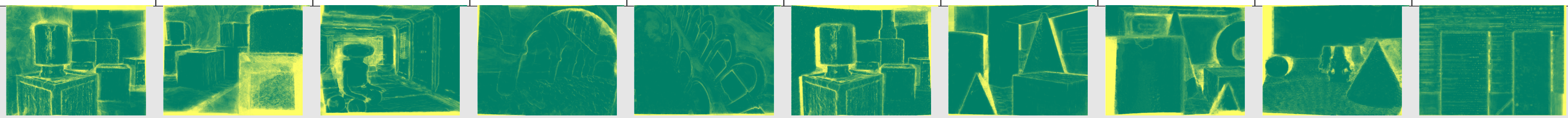
7 Qualitative Results on Ayvaci *et al.* [3] **and Lobaton *et al.*** [27] **10**

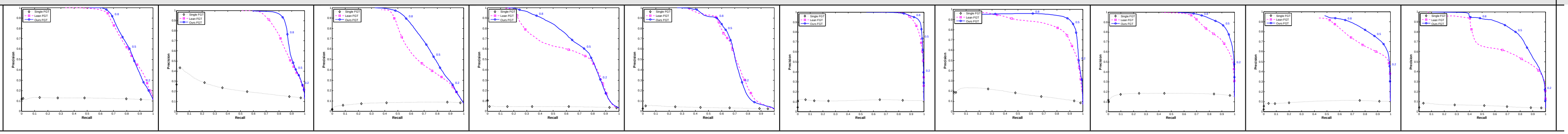
1 Quantitative Results











	Crates1	Crates2	Robot	Sponza1	Sponza2	Crates1txtr	Brickbox1t1	Brickbox2of	Mayan1	Text1
Single FGT	0.564	0.658	0.536	0.633	0.640	0.511	0.766	0.604	0.508	0.626
Lean FGT	**0.950**	0.961	0.922	0.940	**0.929**	0.996	0.986	0.981	0.976	0.985
Ours FGT	0.942	**0.969**	**0.936**	**0.947**	0.928	**0.997**	**0.992**	**0.991**	**0.986**	**0.991**
Precision-Recall										

Table 1: Supplementary results for Table 1 in the paper (Leave-one-out scores). All results are presented on the Full Ground Truth FGT data. The **first row** shows the posteriors from the single image features; **second row** gives posteriors of our lean classifier; **third row** gives posteriors using the full set of features; and **fourth** row gives precision-recall curves on FGT for all three methods.

2 Qualitative Results on Stein and Hebert [39]

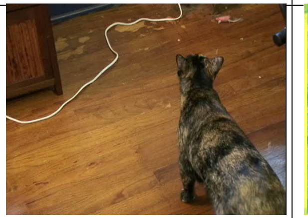
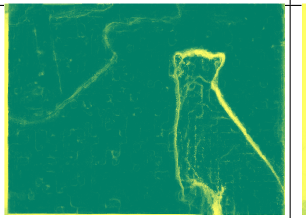
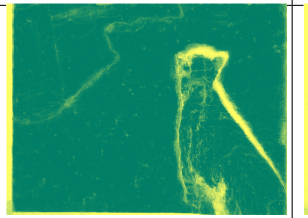
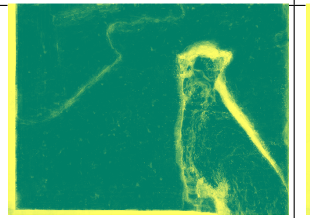
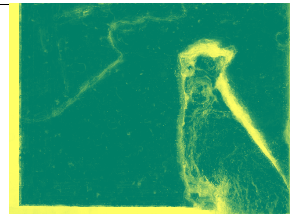
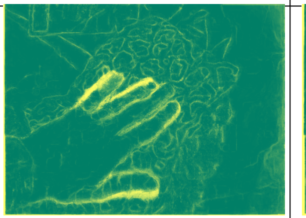
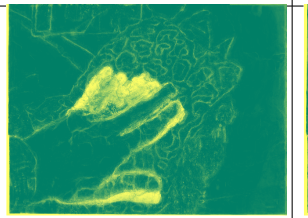
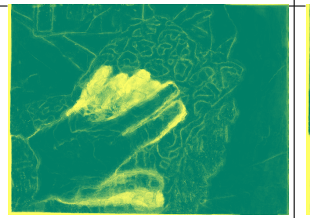
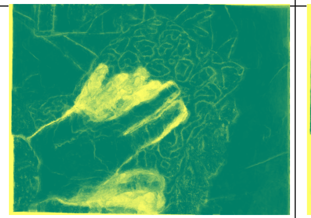
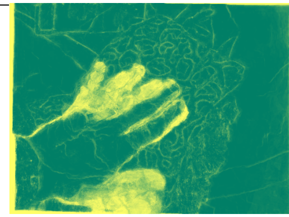
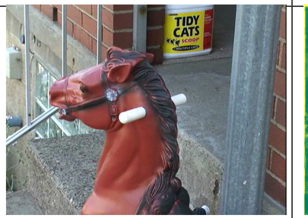
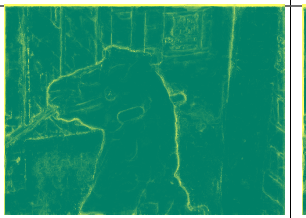
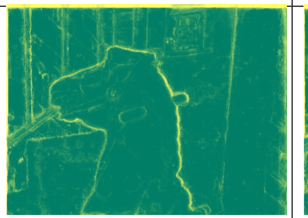
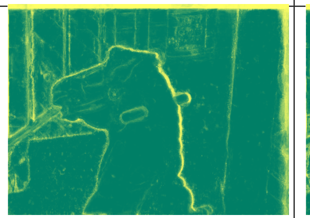
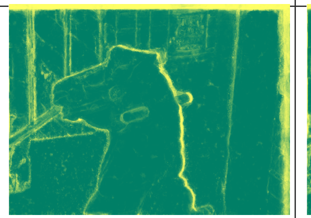
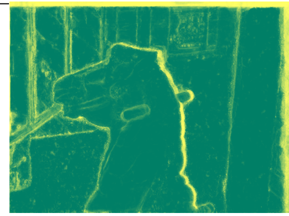
	I_1	$I_1 \to I_2$	$I_1 \to I_3$	$I_1 \to I_4$	$I_1 \to I_5$	$I_1 \to I_6$
zoe1						
hand3						
rocking horse						

Table 2: Each row shows occlusion posteriors from from different sequences from Stein and Hebert [39]. These are supplementary qualitative results to those given in Table 3 in the paper (qualitative results). The **first column** shows the first frame in each sequence. The **second column** shows occlusion posteriors between the first and second frames. Each **successive column** shows the posteriors as a result of increasing the frame gap.

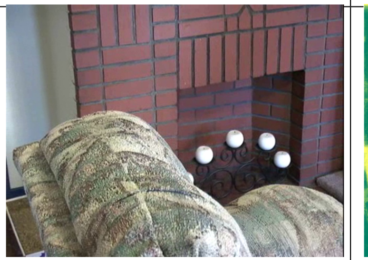
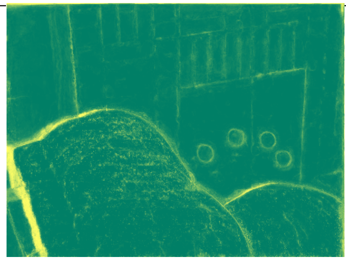
	I_1	$I_1 \to I_4$
couch corner		

Table 3: Scene with little occlusion with respect to frame gap.

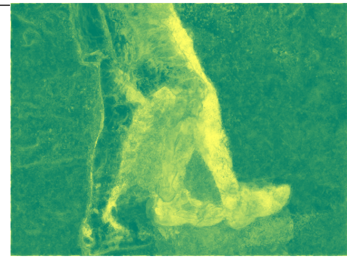
	I_1	$I_1 \to I_4$
walking legs		

Table 4: This scene features very large motion for a frame gap of 4 and as a result the posterior exhibits artefacts.

3 Qualitative Results on Lobaton *et al.* [27] and Sigal *et al.* $[PAMI2004]$

	I_1	$I_1 \rightarrow I_4$
Sequence 07		

Table 5: Sigal *et al.* $[PAMI2004]$ sequence 07 frames 72 and 75.

	I_1	$I_1 \rightarrow I_4$
Sequence 09		

Table 6: Sigal *et al.* $[PAMI2004]$ sequence 09 frames 36 and 39.

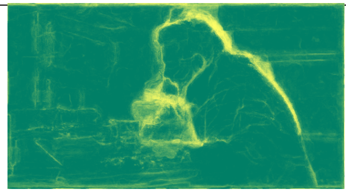
	I_1	$I_1 \rightarrow I_4$
Sequence 13		

Table 7: Sigal *et al.* $[PAMI2004]$ sequence 13 frames 325 and 328.

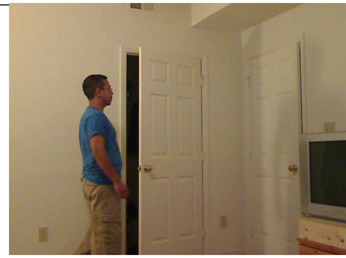
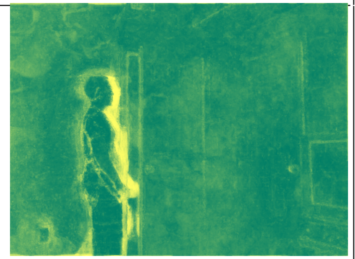
	I_1	$I_1 \rightarrow I_4$
Person walking		

Table 8: Lobaton *et al.* [27] walking sequence frames 4 and 7.

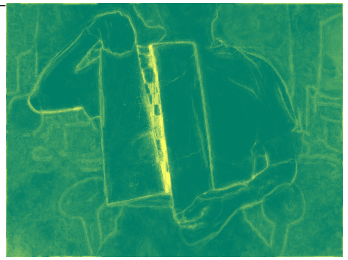
	I_1	$I_1 \rightarrow I_4$
Macbeth board		

Table 9: Lobaton *et al.* [27] macbeth board sequence frames 4 and 7.

4 Full vs. Lean Results

As discussed in the paper we provide a lean version for our algorithm which uses a subset of the full feature vector. It consists of a $d = 122$ dimensional feature vector f_i,as opposed to $d = 227$ for the full version, which is computed for each pixel, using the flow algorithms set $K = \{1, 2\}$ ([44, 45]) and two scale-space choices $S_1 = \{1, 4\}$ and $S_2 = \{1, 10\}$:

$$f_i = \left\{ f_{\mathrm{ED}}\left(\mathbf{x}, S_2\right),\ f_{\mathrm{PC,}}\left(\mathbf{x}, S_1\right),\ f_{\mathrm{TG}}\left(\mathbf{x}, S_2\right), f^{n}_{\mathrm{AV},K}\left(\mathbf{x}, S_1\right),\ f^{n}_{\mathrm{LV},K}\left(\mathbf{x}, S_1\right),\ f^{n}_{\mathrm{CS},K}\left(\mathbf{x}, S_1\right),\ f_{\mathrm{RC},K}\left(\mathbf{x}, S_2\right),\ f_{\mathrm{RA},K}\left(\mathbf{x}, S_2\right),\ f_{\mathrm{FA}}\left(\mathbf{x}, S_2\right),\ f_{\mathrm{FN}}\left(\mathbf{x}, S_2\right) \right\} \tag{1}$$

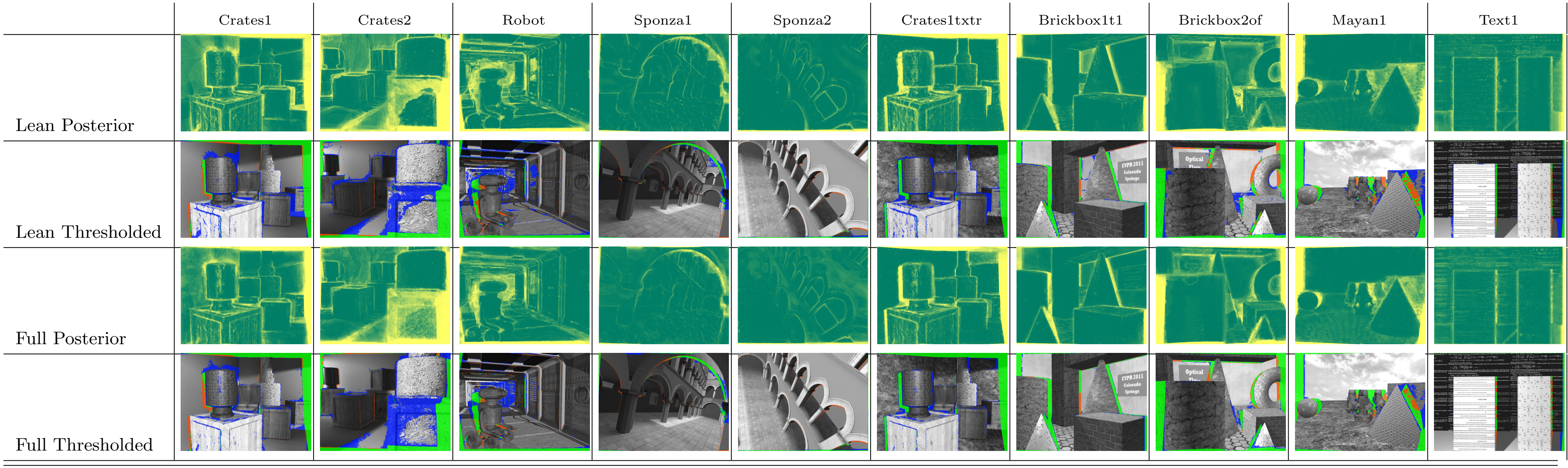


Table 10: Full vs. lean comparison on leave-one-out tests for the FGT. The **first** and **third rows** give the posteriors for the lean and full version respectively, whereas the **second** and **fourth rows** show the corresponding thresholded results. The threshold was set at 0.5. The thresholded images show the true-positives in green overlay, false-positives in blue, and false-negatives in reddish orange.

5 Angle vs. Texture

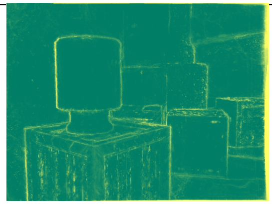
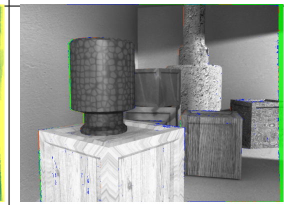
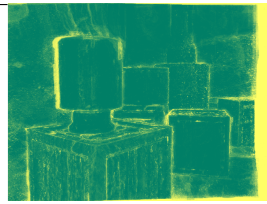
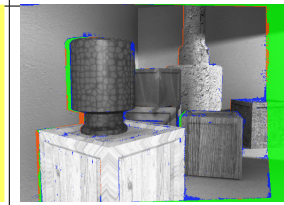
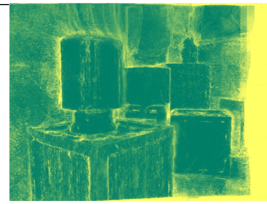
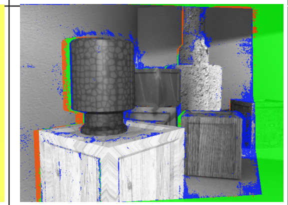
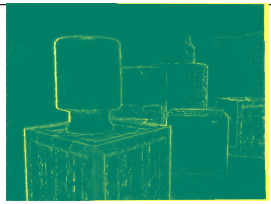
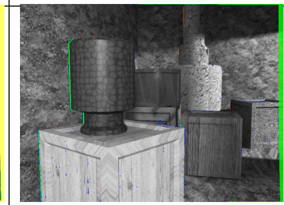
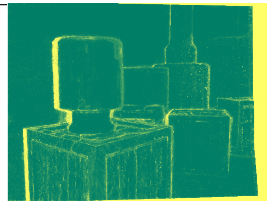
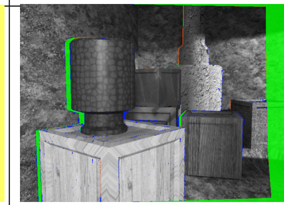
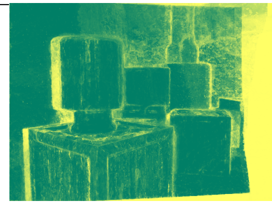
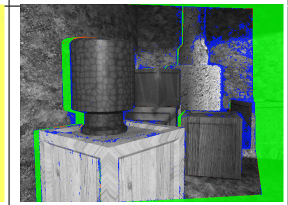
	Posterior 1°	Thresholded 1°	Posterior 4°	Thresholded 4°	Posterior 7°	Thresholded 1°
						
						

Table 11: Supplementary results for Figure 4 in the paper. The two rows represent the same scene with different background texture. For the second frame the camera was rotated about the nearest corner of the wooden box in the foreground along the x axis for 1°, 4° and 7°. The posteriors and thresholded results are given for all 6 pairs. The threshold, as before, was set at 0.5.

6 Superpixels

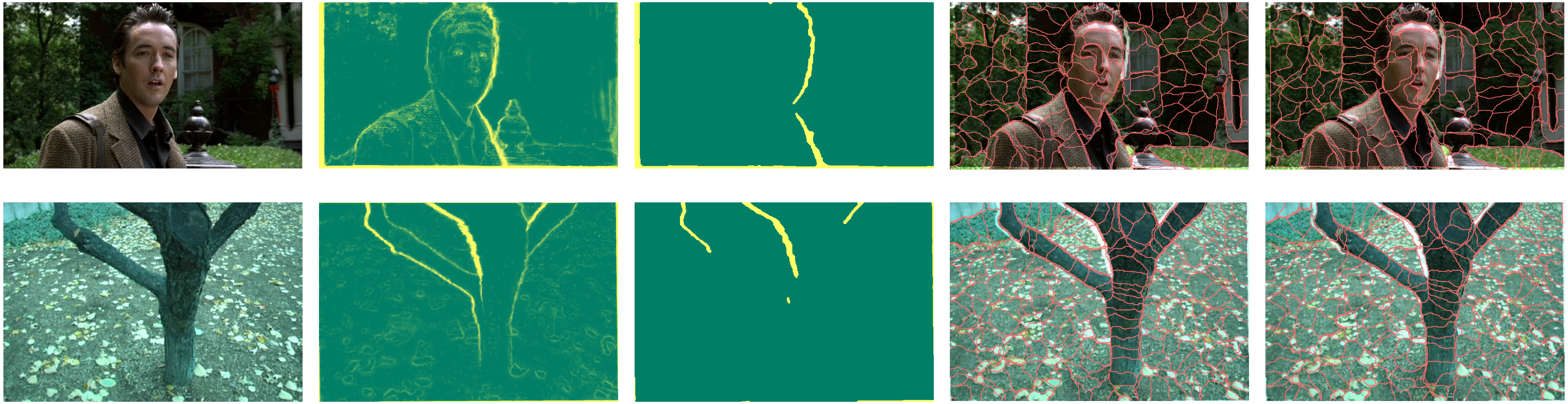


Figure 1: Occlusion-aware superpixel results for two additional sequences from [27, 35]. As Figure 5 in the paper, for each row: the first images is the first image of the pair, the second is the posterior, the third is the regularised posterior from graph cuts, the fourth is the result of image based over-segmentation [32] and the fifth is with the addition of our occlusion term.

7 Qualitative Results on Ayvaci *et al.* [3] and Lobaton *et al.* [27]

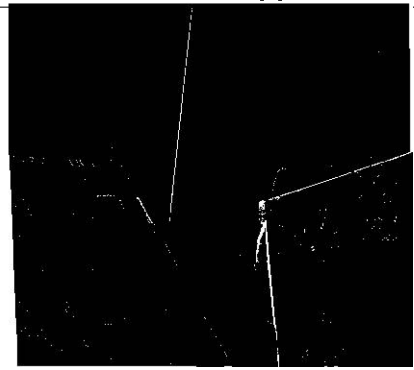
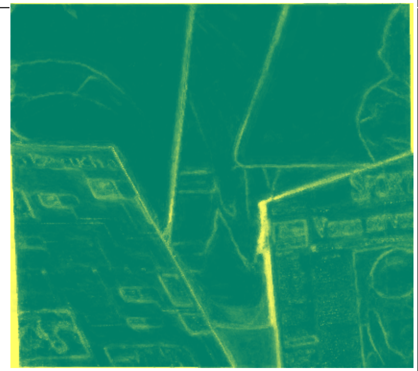
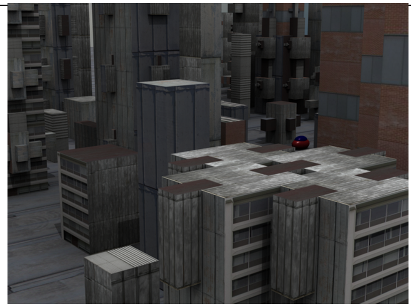
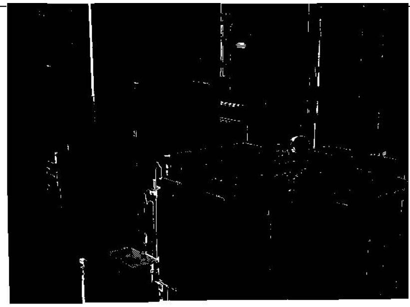
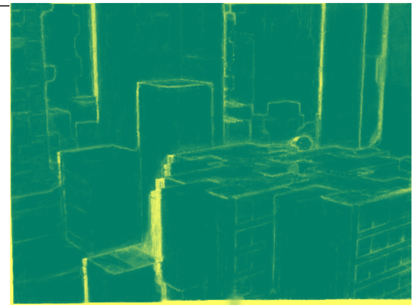
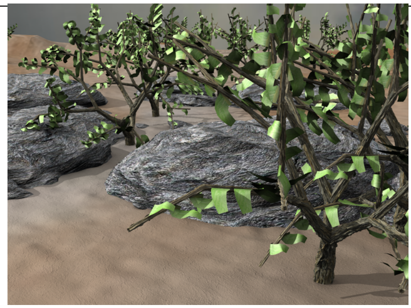
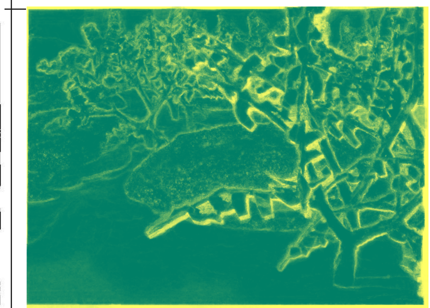
	I_1	Ayvaci *et al.* [3] e_1	Our Posterior
venus [6]			
urban2 [6]			
grove3 [6]			

Table 12: Shows comparative results against Ayvaci *et al.* [3].

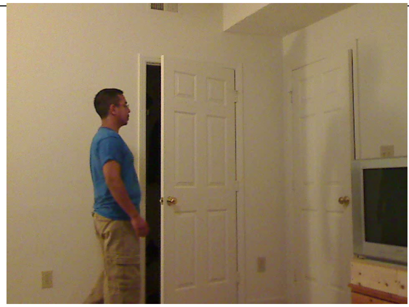
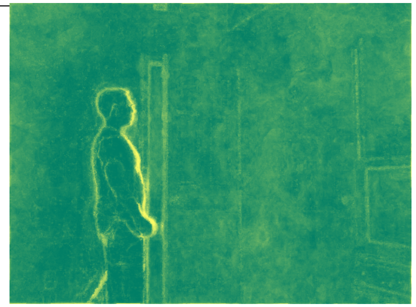
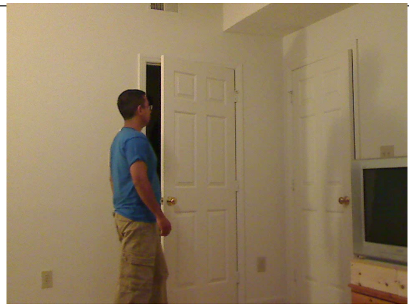
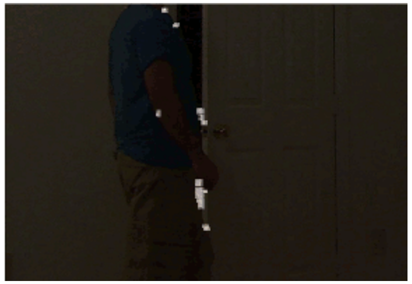
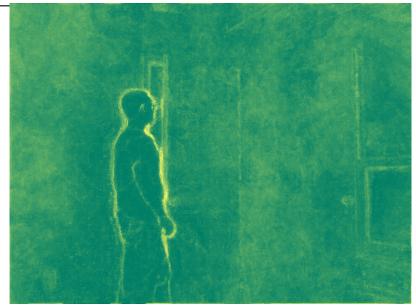
	I_1	Lobaton *et al.* [27] occlusion inset	Our Posterior
frame 0000			
frame 0012			

Table 13: Shows comparative results against Lobaton *et al.* [27].