

Ground-truth segmentation comparison:

The segmentation obtained from the proposed approach and TcMVS [41] is compared against ground-truth and the evaluation of the segmentation is shown in Figure S1 depicting the improvement using the proposed method. Region in red represents region missing from ground-truth and green represents region not present in ground-truth.

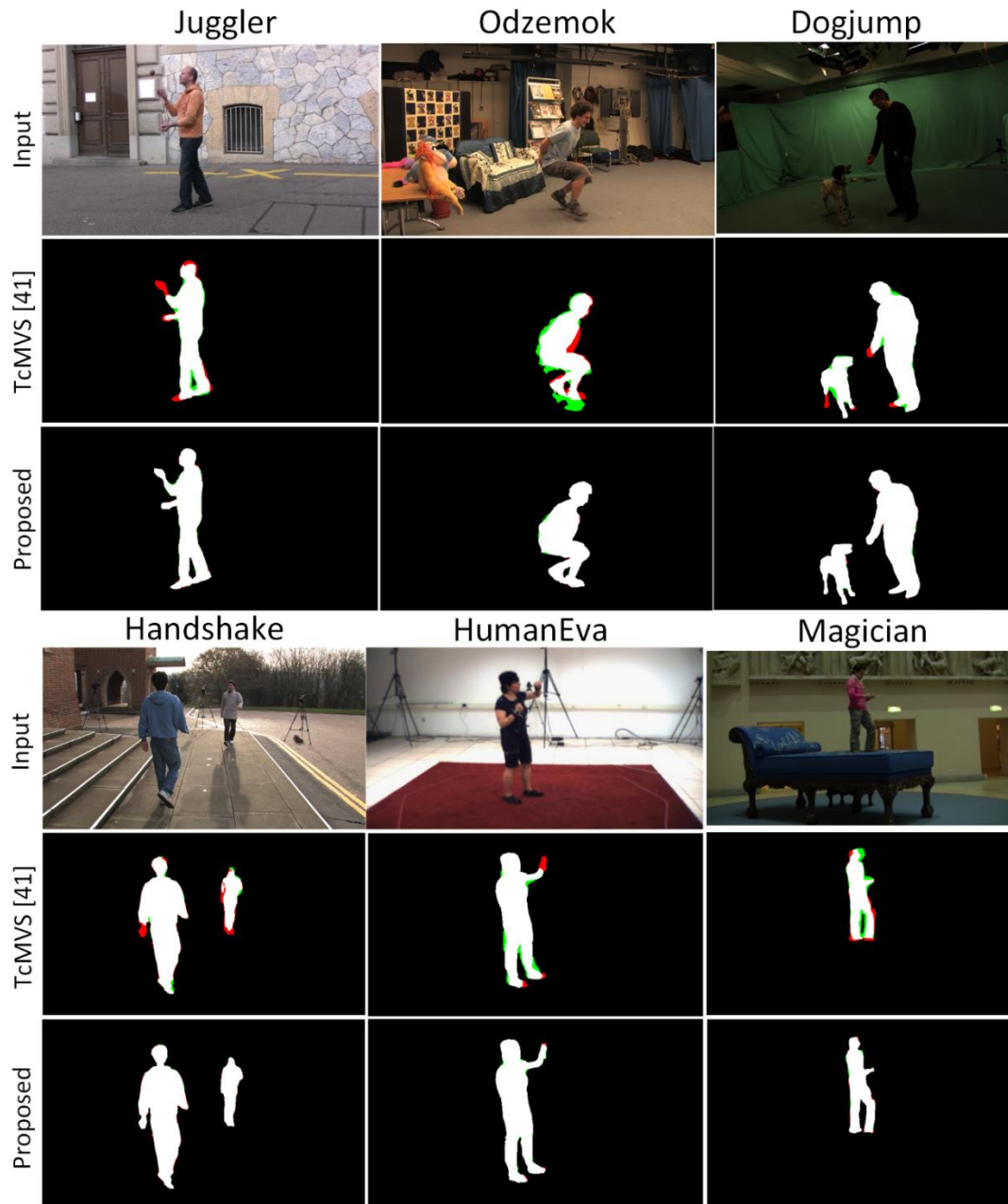


Figure S1. Ground-truth segmentation comparison with [41] on multi-view datasets

Segmentation comparison with Kundu[33]:

Kundu presents results for single view video of street scenes. Comparison with the proposed approach is presented in Fig. S2 and S3. This shows that our method gives better segmentation boundaries but filters out thin structures.

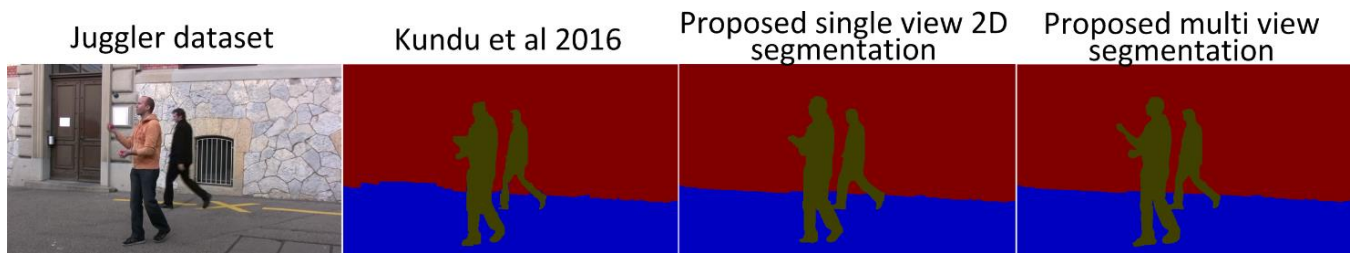


Figure S2. Segmentation comparison with [33] on multi-view dataset

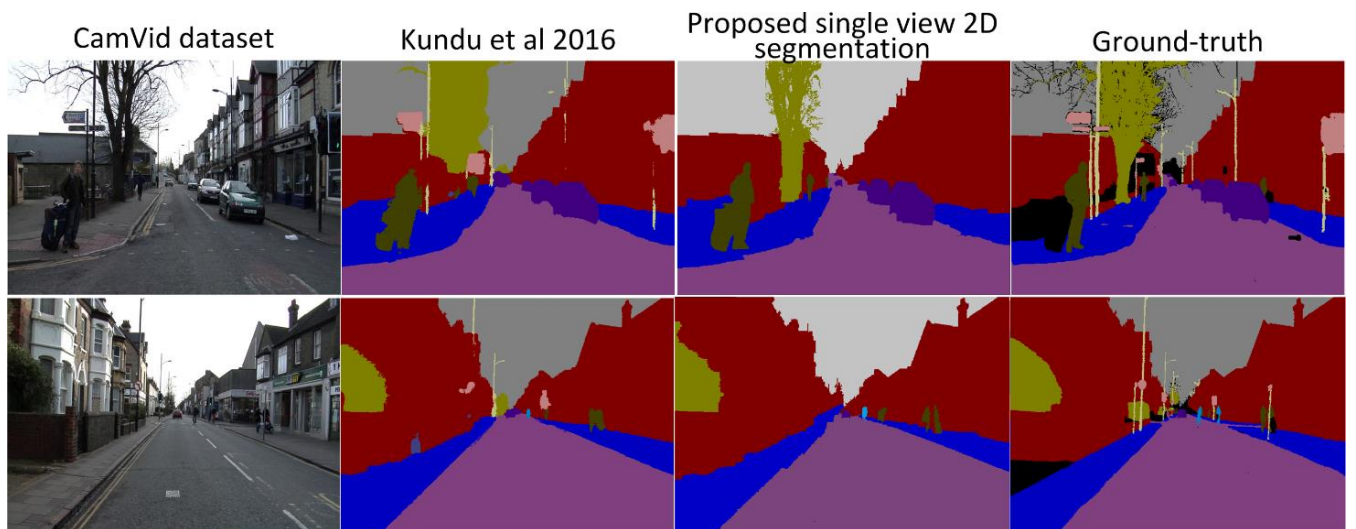


Figure S3. Segmentation comparison with [33] on street sequence

Computation complexity:

Results are insensitive to parameter setting for all indoor and outdoor scenes. Table S1 shows the parameters used, with constant contrast cost $\lambda_{ca} = \lambda_{cl} = 0.5$ and smoothness cost $\lambda_s^T = 0.6$ and $\lambda_s^S = 0.4$.

Parameters/ Datasets	$E_{\text{multi}}()$			$E_{\text{single}}()$	
	λ_d	λ_{sm}	λ_s	λ_a and λ_{sem}	λ_c
Odzemok	1.0	0.7	0.1	0.5	5
Breakdance	1.0	0.7	0.1	0.5	5
DogJump	1.0	0.7	0.1	0.5	5
HumanEva	1.0	0.7	0.1	0.5	5
Juggler	1.0	0.9	0.2	0.6	7.5
Handshake	1.0	0.9	0.2	0.6	7.5
Magician	1.0	0.9	0.2	0.6	7.5

Table S1. Parameters for datasets