

The Articulated Scene Model: Model-less Priors for Robot Object Learning?

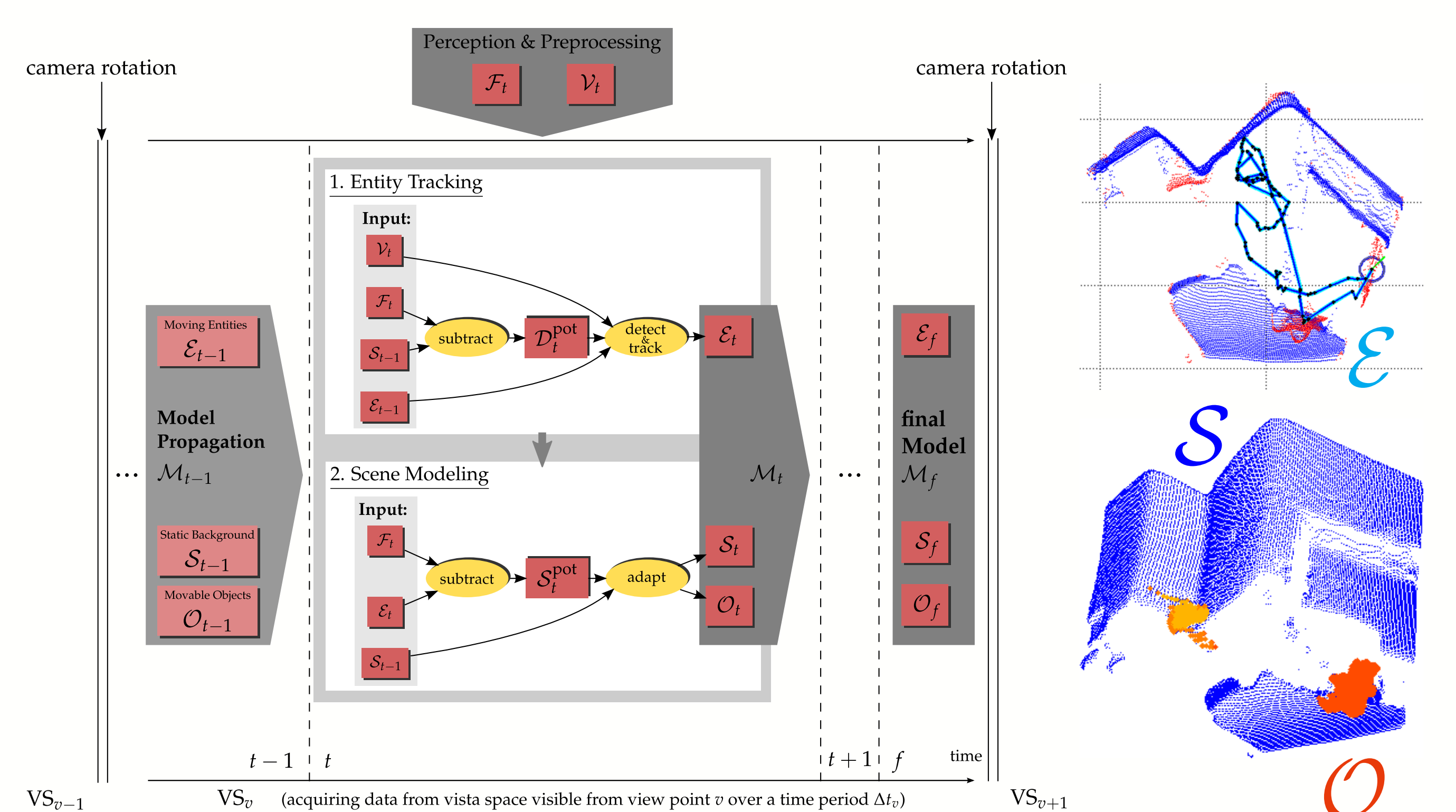
Motivation

How can articulated scene parts be extracted from short and disturbed observations?

- analyze spatial changes
- change := variation of structure
- change detection := comparison of current perception with memorized representation

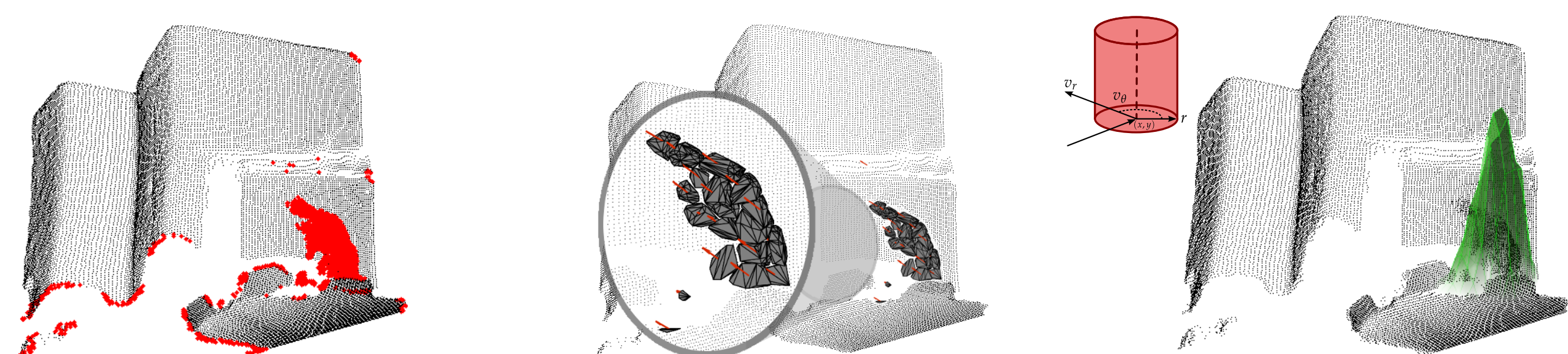


System overview



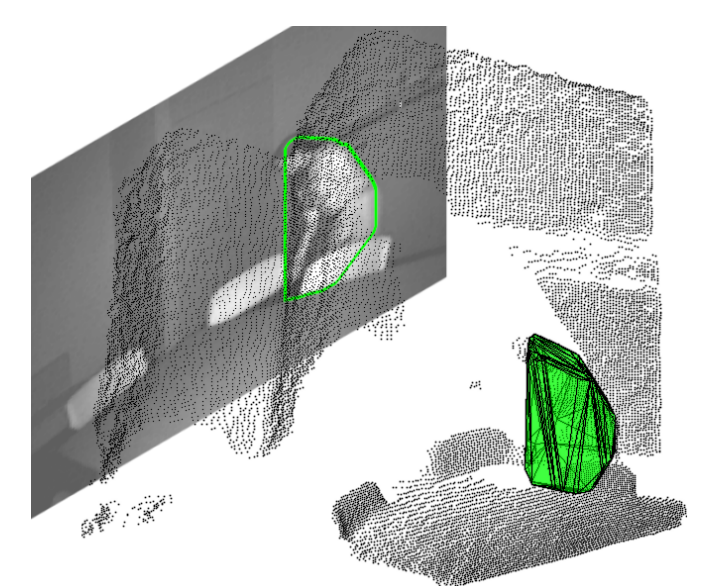
Tracking

- potentially dynamic points ($\mathcal{F}_t - \mathcal{S}_{t-1}$)
- motion-attributed clusters based on spatial proximity and velocity homogeneity
- entity hypotheses rated by value in density approximation ρ ($\rho(e_k) = K_r(e_k) \sum_{l \in e_k} K_d(l, e_k) K_v(l, e_k)$)



Modeling

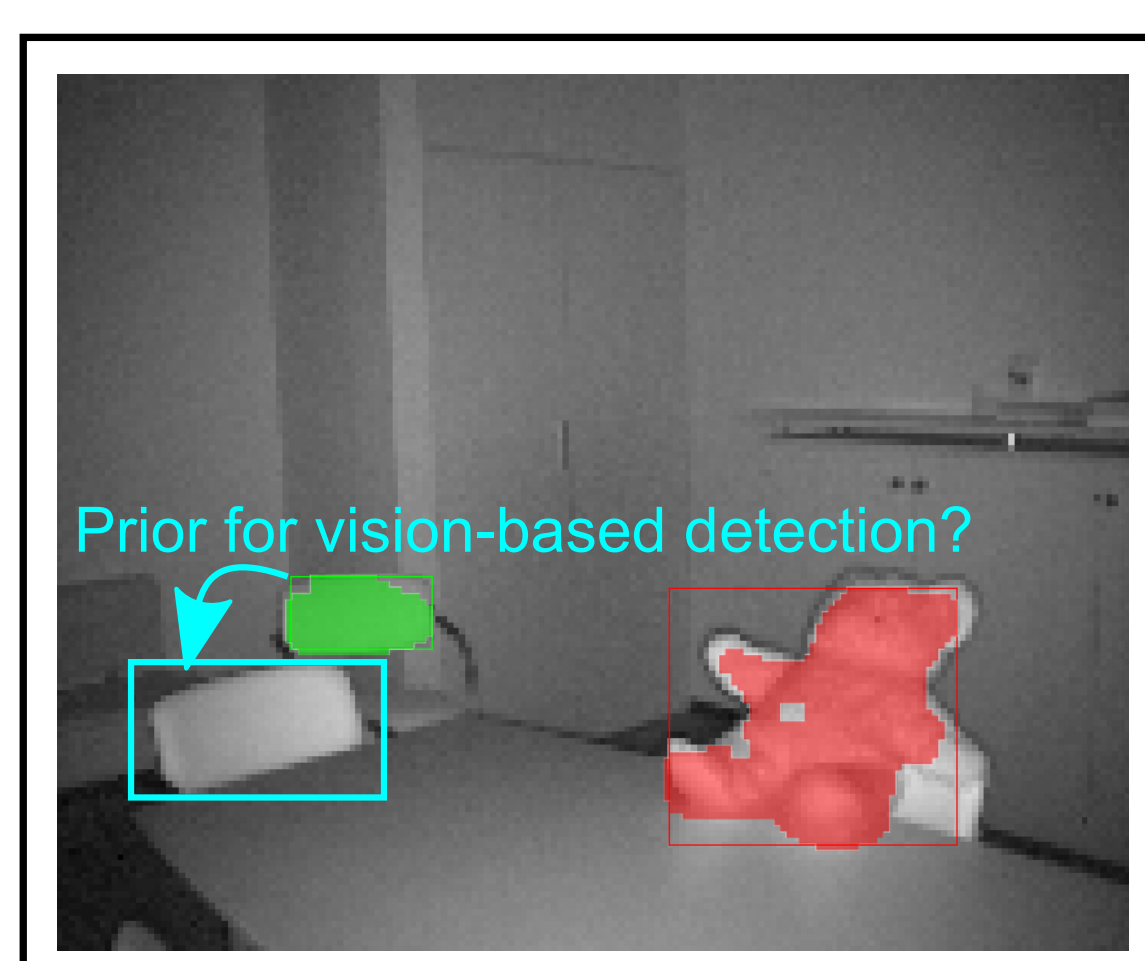
- potentially static points ($\mathcal{F}_t - \mathcal{E}_t$)
- static background := *farthest* static depth measurement along a ray
- movable objects := *model-less* detection from depth measurements emerging in front of a known background



Applications

Object segmentation:

- articulated scene model provides example of a movable object (e.g. chair)
- learn object appearance (SIFT features, color distribution, ...)
- detect other objects of the same type which could possibly be moved due to the given example (\rightarrow attention areas for robot)



Object articulation:

- history of object extractions provides a model-less object trajectory
- train candidate kinematic models using object tracks (e.g. rigid model for drawer)

