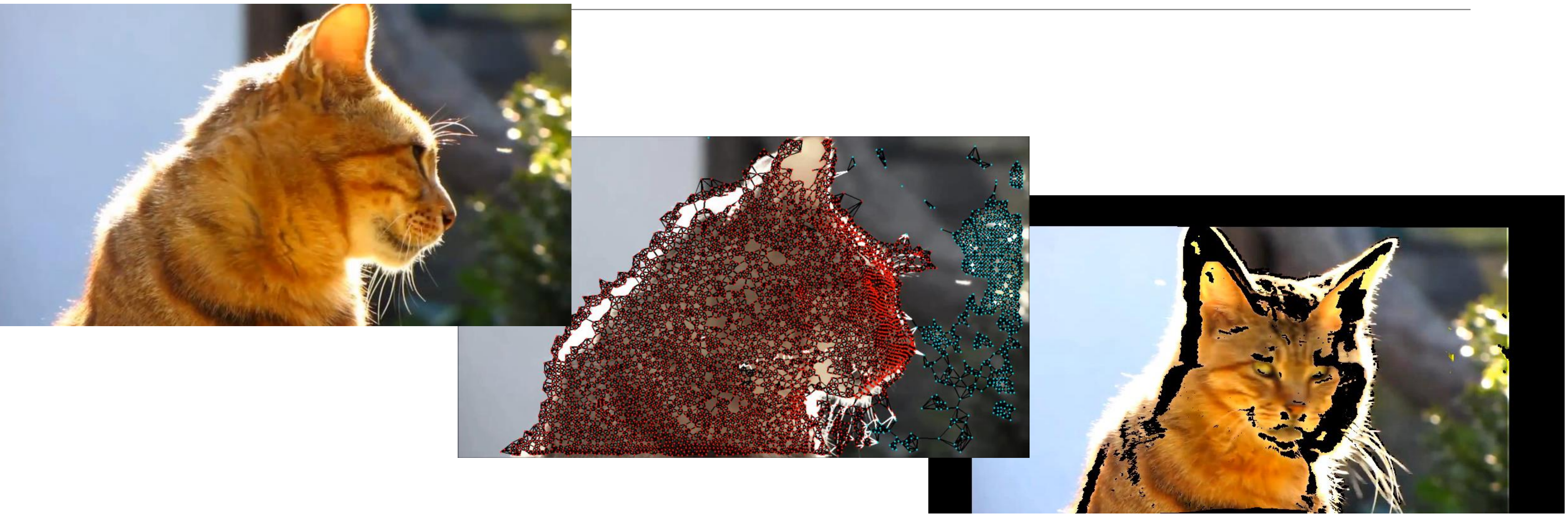
Video Pop-up: Monocular 3D Reconstruction of Dynamic Scenes

CHRIS RUSSELL,† RUI YU,† AND LOURDES AGAPITO

† JOINT FIRST AUTHORSHIP

Aim



Raw video → Segment objects + Reconstruction

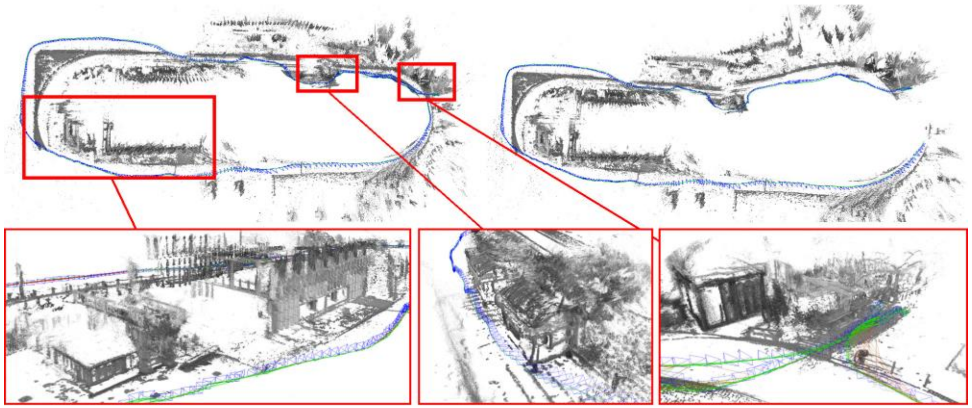
Completely Unsupervised

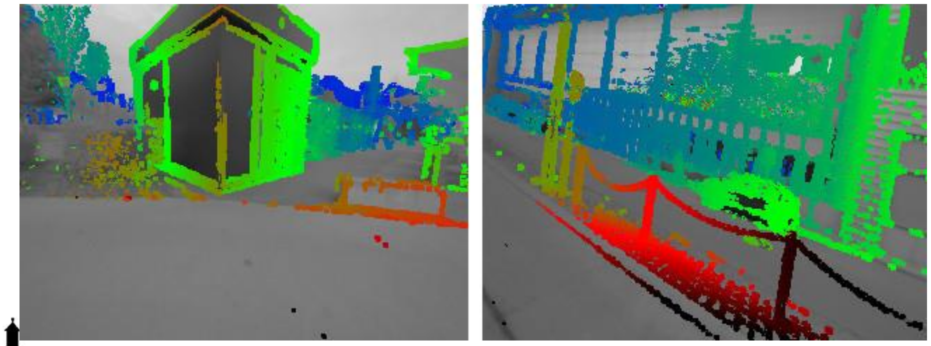
Our approach

Rigid Parts

Piecewise rigid Objects

Rigid reconstruction as gold standard





Fast, reliable, low powered

e.g. LSD-SLAM

Recovers background rather than object of interest in most videos

Non-rigid Structure from Motion

Single object

Presegmented

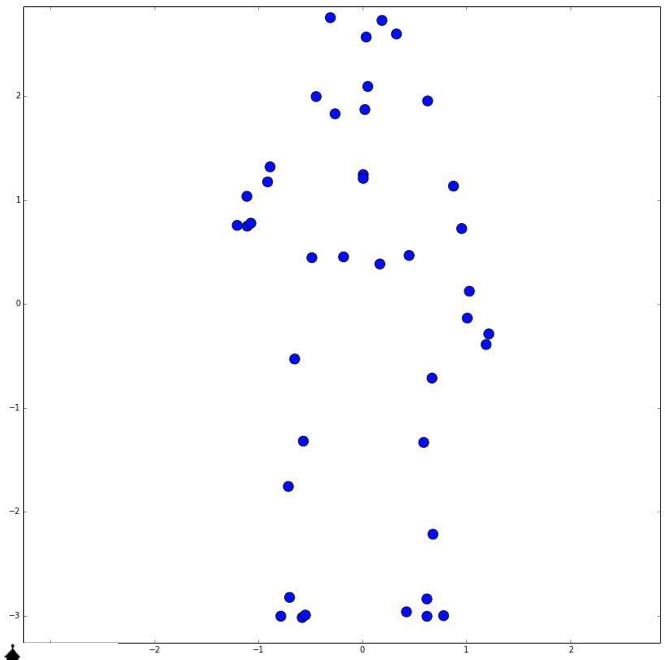
Additional training data for best results

What about the whole world?

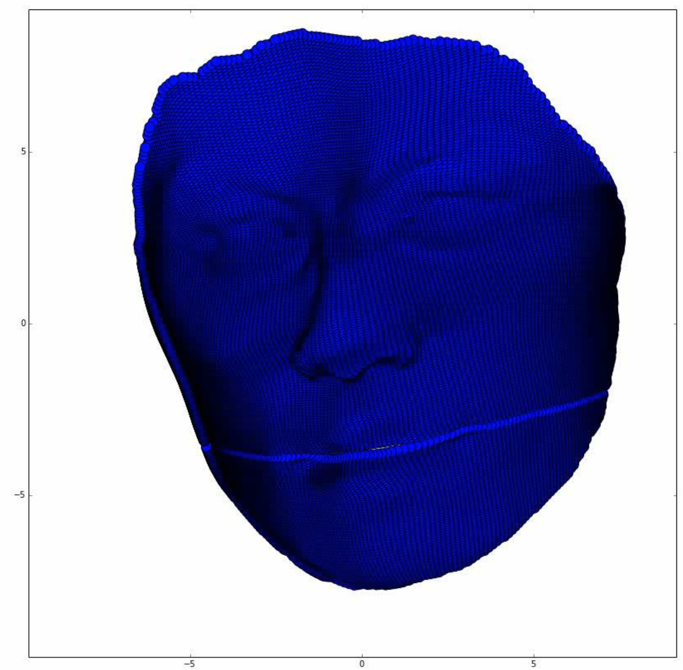
Suwajanakorn et al. ECCV'12

Generic Non-rigid – Synthetic Inputs

DAI ET AL.



GARG ET AL.

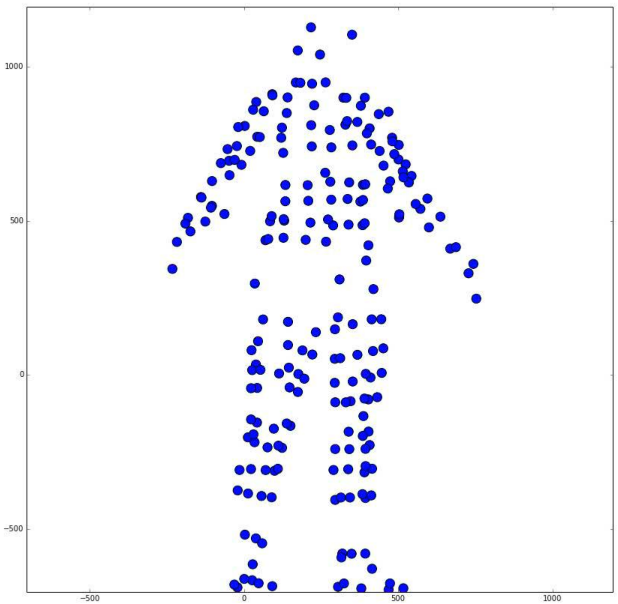


Reconstruction without rotation

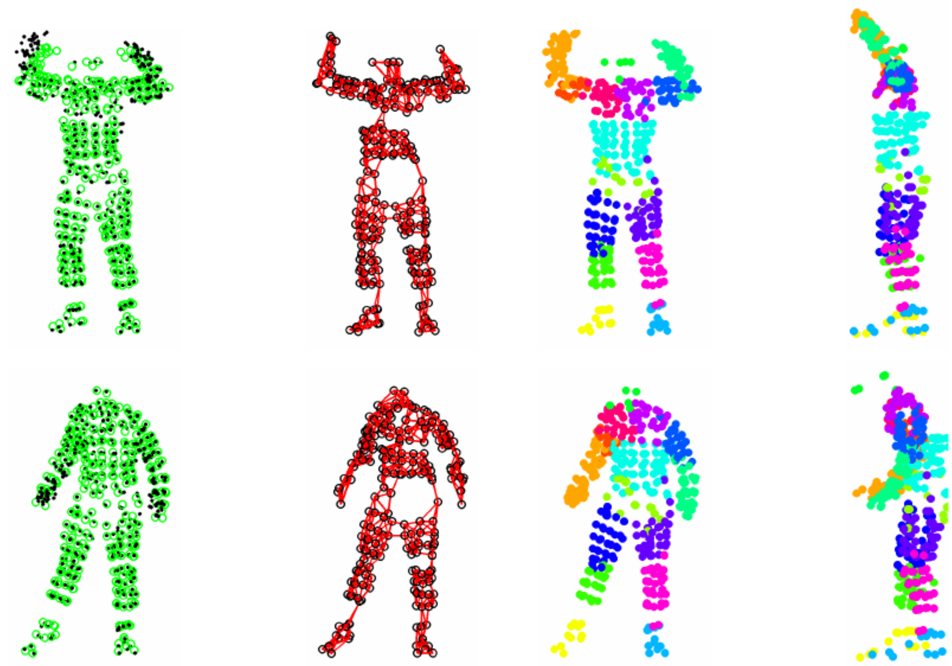
SPLIT OBJECT INTO SIMPLE OVERLAPPING PARTS

RUSSELL ET AL. CVPR'11

FAYAD ET AL. ICCV '11

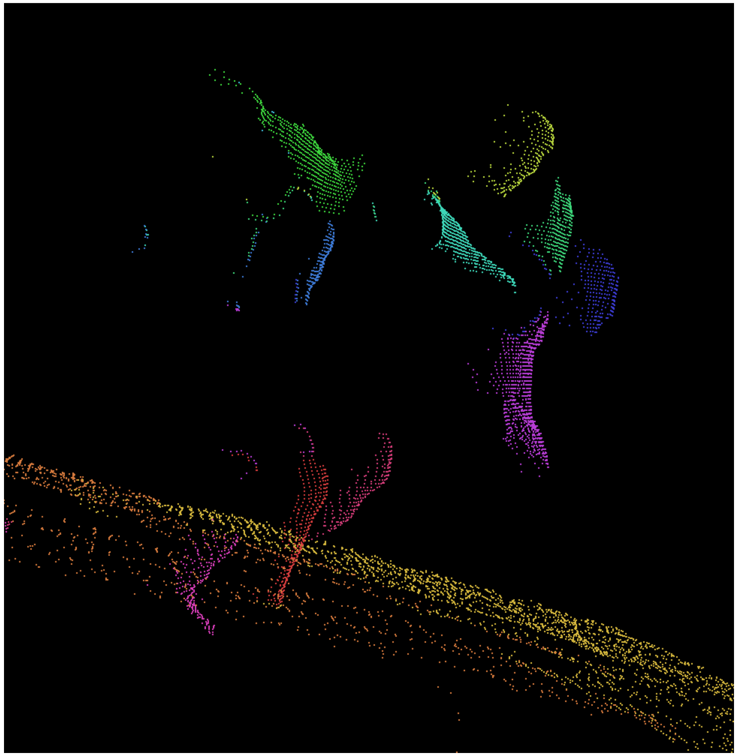


(*a*) Reconstruction vs. GT (*b*) Neighbourhood (*c*) Segmentation

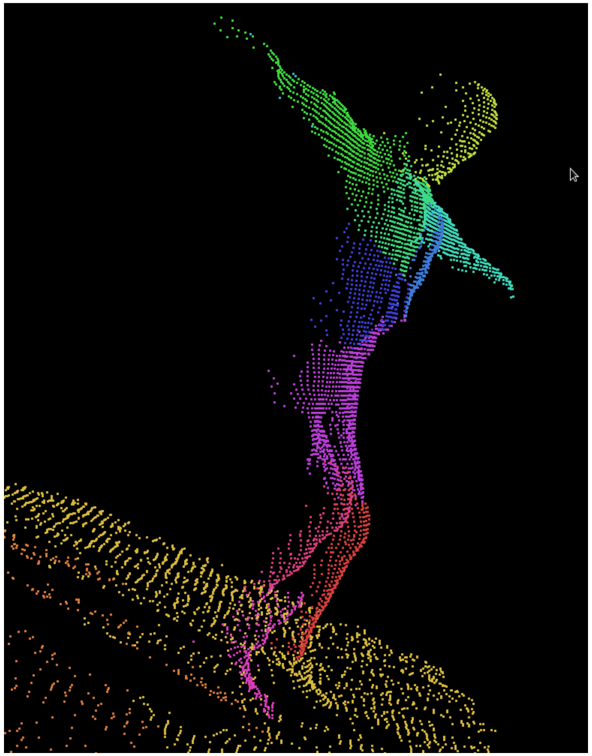


Part Stitching using overlap

WITHOUT OVERLAP

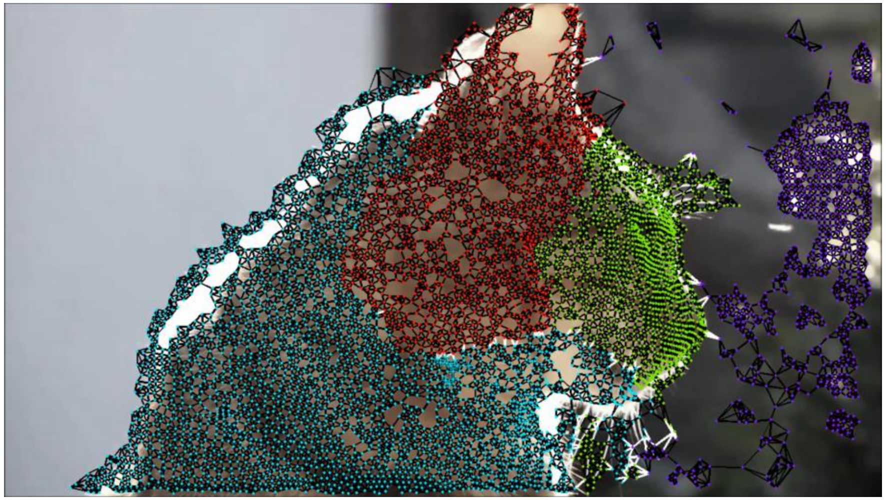


WITH OVERLAP

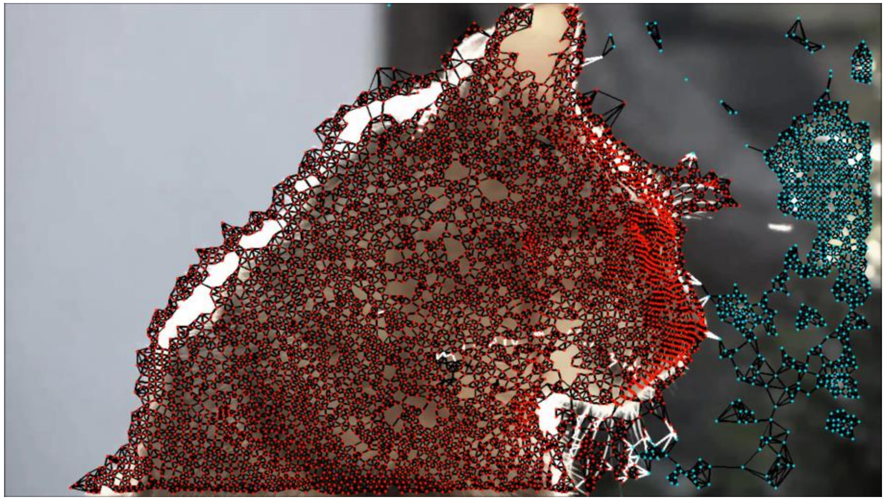


Multi-object reconstruction

Parts from different objects must not overlap



Parts

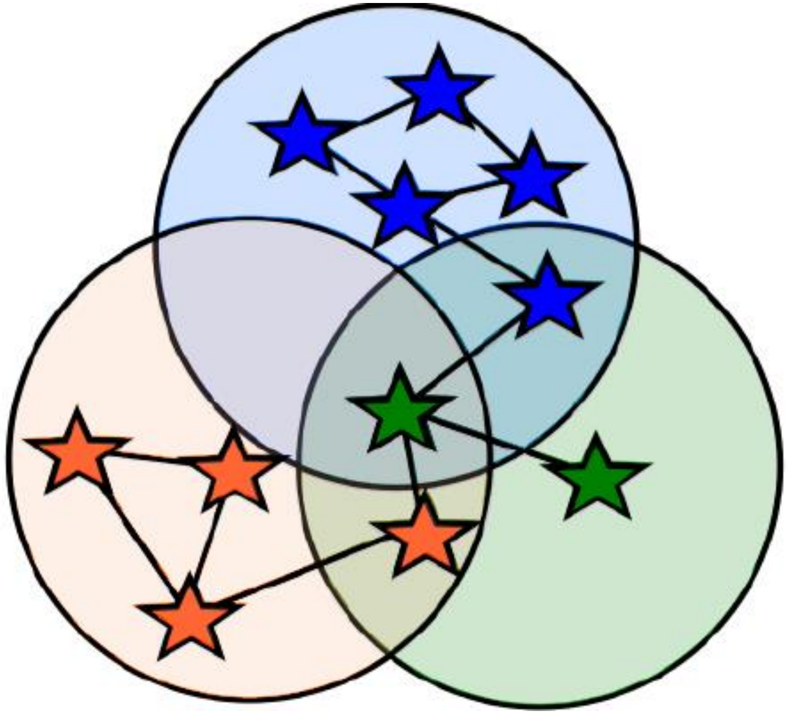


Objects

The formulation

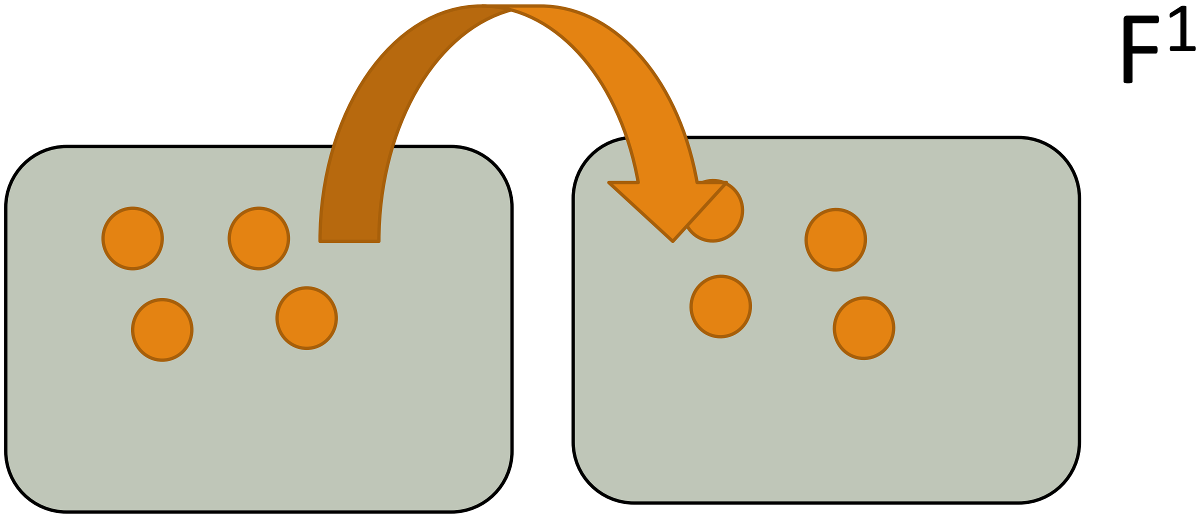
$$C(\mathbf{x}) = \sum_{m \in i \in \mathcal{T}} U_i(m)$$

Part assignment cost



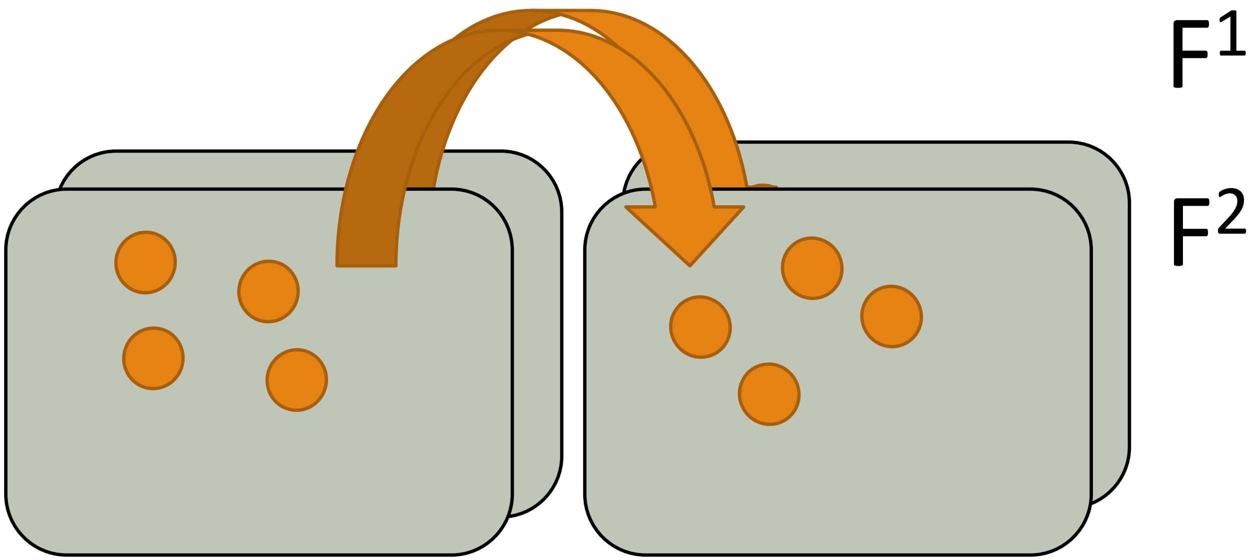
Subject to overlap constraints

Unary Motion Costs



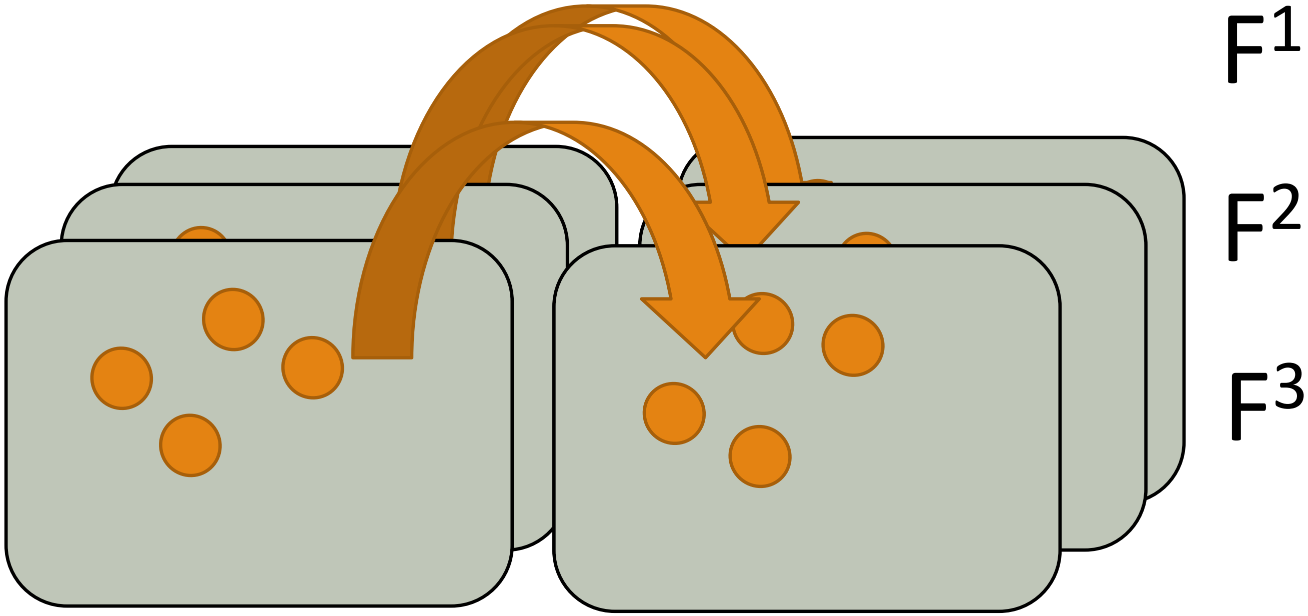
$$U_i(m) = \sum_f \gamma^{-1} ({u_i^{f+1}}^T F_m^f u_i^f)^2$$

Unary Motion Costs



$$U_i(m) = \sum_f \gamma^{-1} (u_i^{f+1^T} F_m^f u_i^f)^2$$

Unary Motion Costs



$$U_i(m) = \sum_f \gamma^{-1} ({u_i^{f+1}}^T F_m^f u_i^f)^2$$

The formulation

$$C(\mathbf{x}) = \sum_{m \in i \in \mathcal{T}} U_i(m)$$

Part assignment cost

Subject to overlap constraints

The formulation

$$C(\mathbf{x}) = \sum_{m \in i \in \mathcal{T}} U_i(m)$$

Part assignment cost

$$+ \sum_{(i,j) \in \mathcal{G}} d_{i,j} \Delta(j \notin N_i)$$

Edge drop cost

Subject to overlap constraints

The formulation

$$C(\mathbf{x}) = \sum_{m \in i \in \mathcal{T}} U_i(m)$$

Part assignment cost

$$+ \sum_{(i,j) \in \mathcal{G}} d_{i,j} \Delta(j \notin N_i)$$

Edge drop cost

$$+ \sum_{m,n \in \mathcal{P}} \Delta(\exists i \in I_m \wedge i \in n)$$

Overlap Sparsity prior

Subject to overlap constraints

The formulation

$$C(\mathbf{x}) = \sum_{m \in i \in \mathcal{T}} U_i(m)$$

Part assignment cost

$$+ \sum_{(i,j) \in \mathcal{G}} d_{i,j} \Delta(j \notin N_i)$$

Edge drop cost

$$+ \sum_{m,n \in \mathcal{P}} \Delta(\exists i \in I_m \wedge i \in n)$$

Overlap Sparsity prior

$$+\mathrm{MDL}(\mathbf{x})$$

Part Sparsity prior

Subject to overlap constraints

UCL

The formulation

$$C(\mathbf{x}) = \sum_{m \in i \in \mathcal{T}} U_i(m)$$

$$+ \sum_{(i,j) \in \mathcal{G}} d_{i,j} \Delta(j \notin N_i)$$

$$+ \sum_{m,n \in \mathcal{P}} \Delta(\exists i \in I_m \wedge i \in n)$$

$$+\mathrm{MDL}(\mathbf{x})$$

Subject to overlap constraints

Part assignment cost

Edge drop cost

Overlap Sparsity prior

Part Sparsity prior

New

Difficult to optimise

Novel alpha expansion formulation

UCL

Without Graph Breaking

Parts

Objects

Graphs leak on object boundaries

Tracks on the edges of objects appear for
a few frames and drift

With Graph Breaking

Parts

Objects

Drop edges that cause a high-cost overlap

Sparsity prior to discard small regions of overlap

The formulation

$$C(\mathbf{x}) = \sum_{m \in i \in \mathcal{T}} U_i(m)$$

Part assignment cost

$$+ \sum_{(i,j) \in \mathcal{G}} d_{i,j} \Delta(j \notin N_i)$$

Edge drop cost

$$+ \sum_{m,n \in \mathcal{P}} \Delta(\exists i \in I_m \wedge i \in n)$$

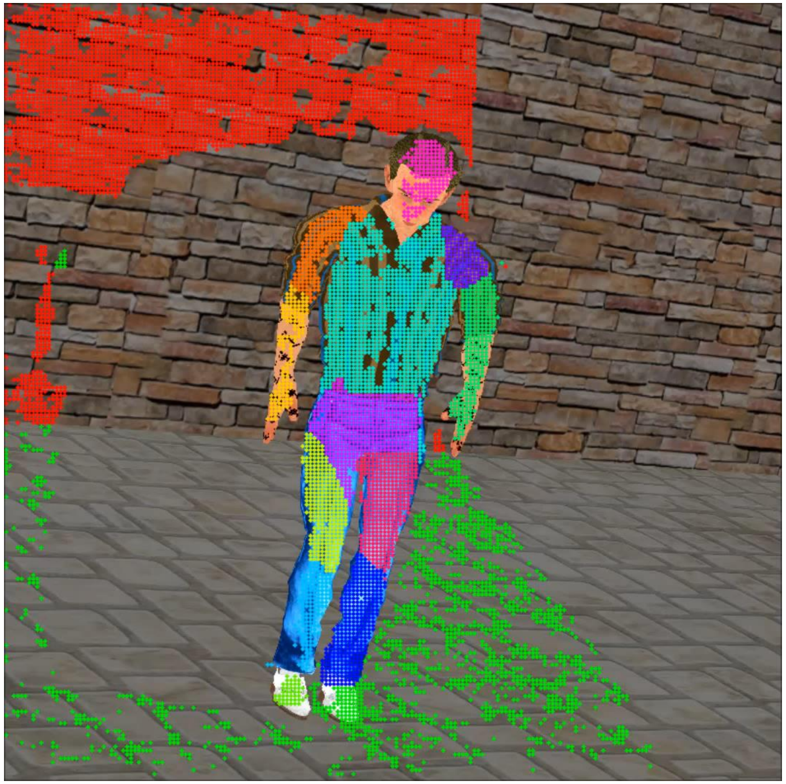
Overlap Sparsity

$$+\mathrm{MDL}(\mathbf{x})$$

Part Sparsity

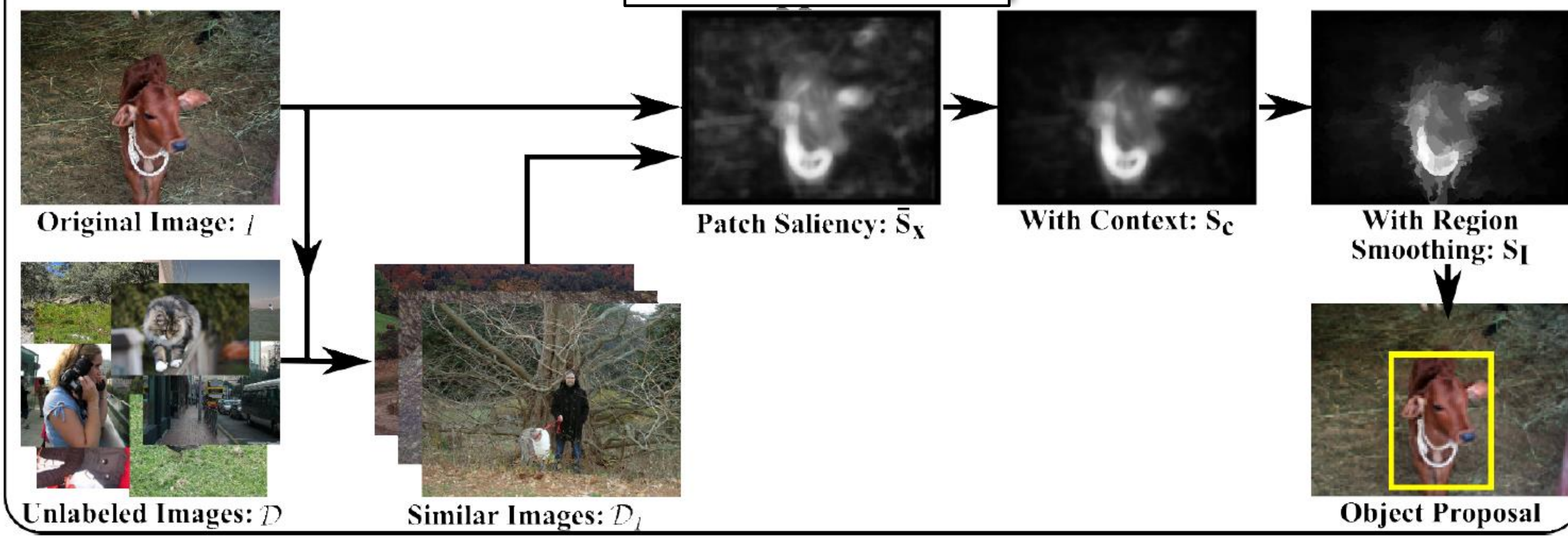
Subject to overlap constraints

Separating objects with weak motion cues



Using Appearance models in Parts

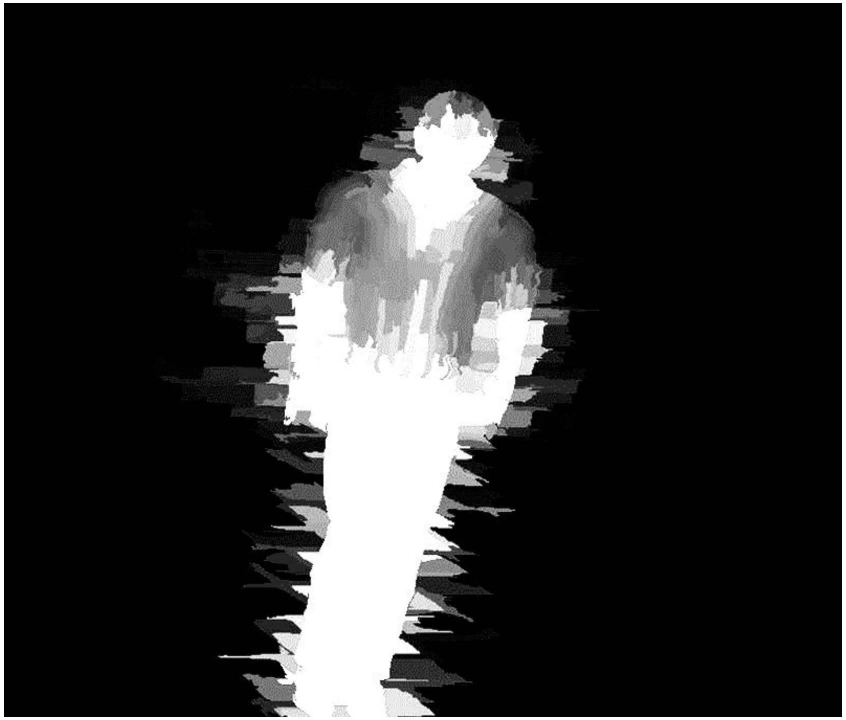
Siva '13



Approach finds uncommon regions that don’t occur in similar images

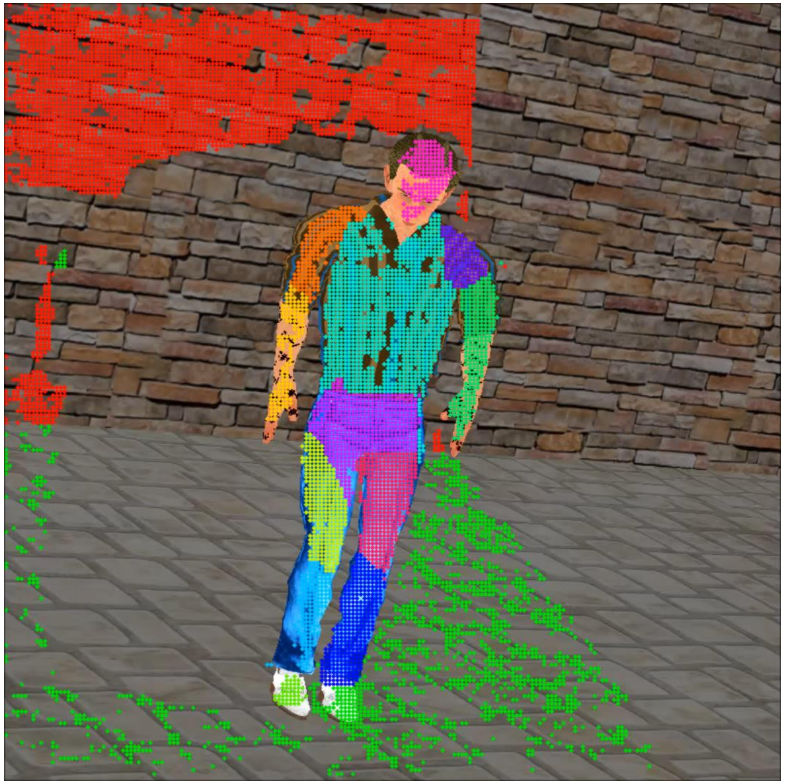
Separating objects with weak motion cues

Siva et al CVPR'13

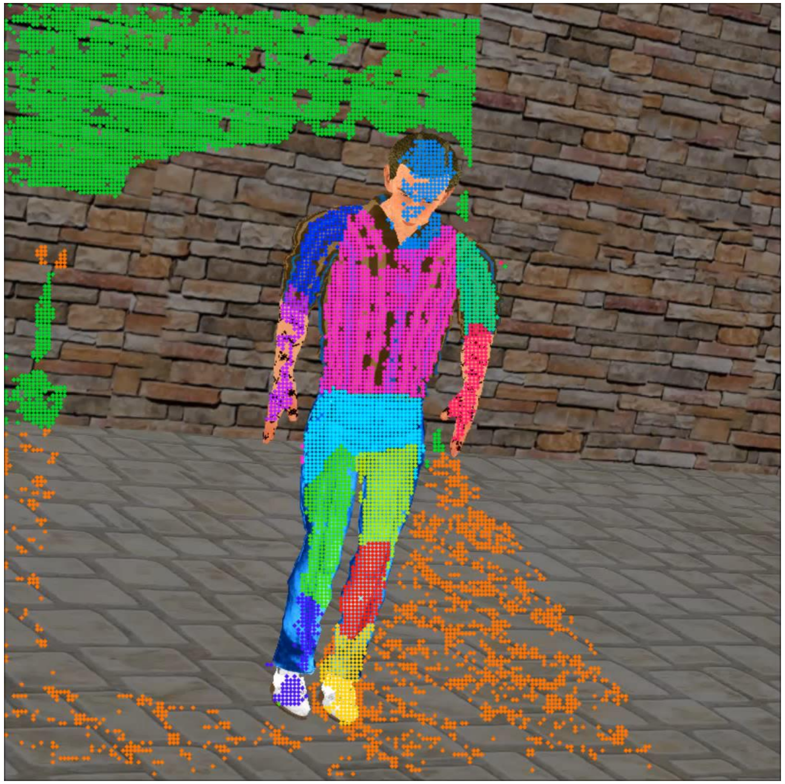


Separating objects with weak motion cues

Siva et al CVPR'13



Without Siva'13



With Siva'13

Reconstruction

Parts → Metric Upgrade → Densify (Blur in xyRGB)

Conclusion

Works on real footage taken from youtube