Interactive Videos

Plausible Video Editing using Sparse Structure Points

Chia-Sheng Chang Hung-Kuo Chu Niloy J. Mitra

國立清華大學
NATIONAL TSING HUA UNIVERSITY

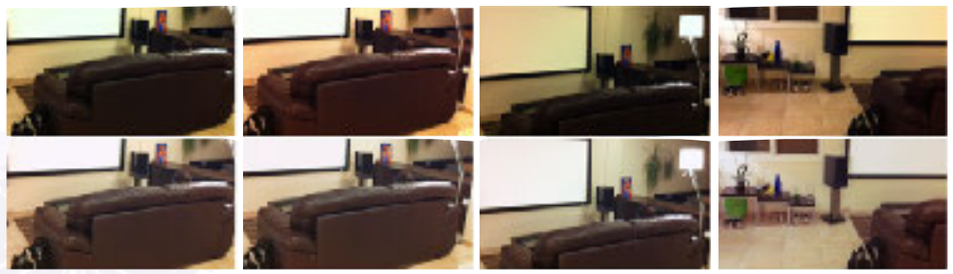
UCL

Video Cameras

Video Research

[Liu et al. 2014]

[Eilertsen et al. 2015]



[Farbman and Lischinski 2011]

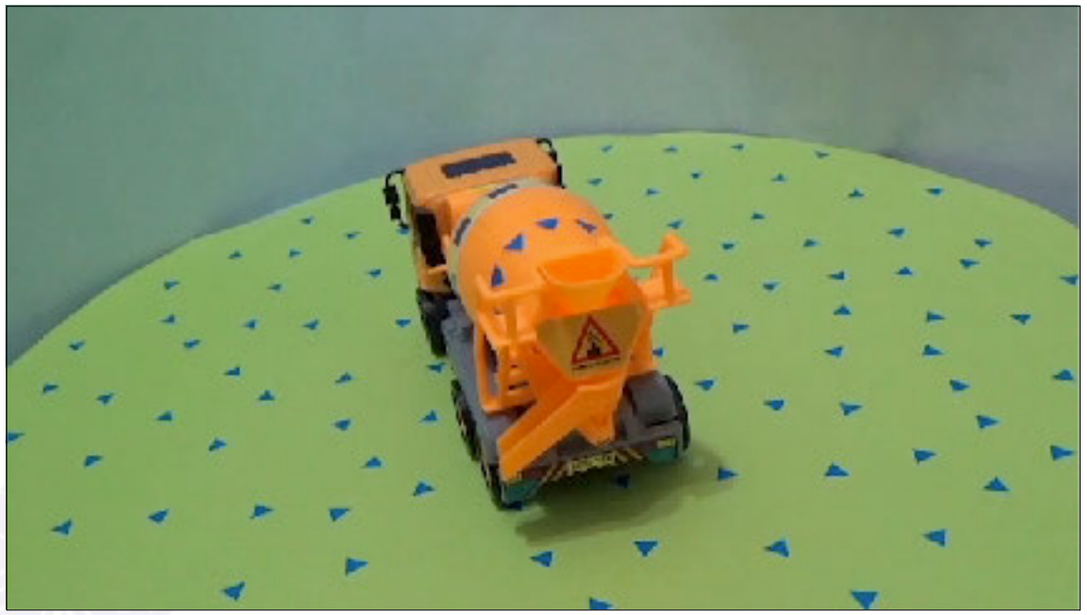
[Cho et al. 2012]

[Bai et al. 2009]

[Fan et al. 2015]

Motivation

- Input RGB videos



Motivation

- Our results

- 3D object-level video edits

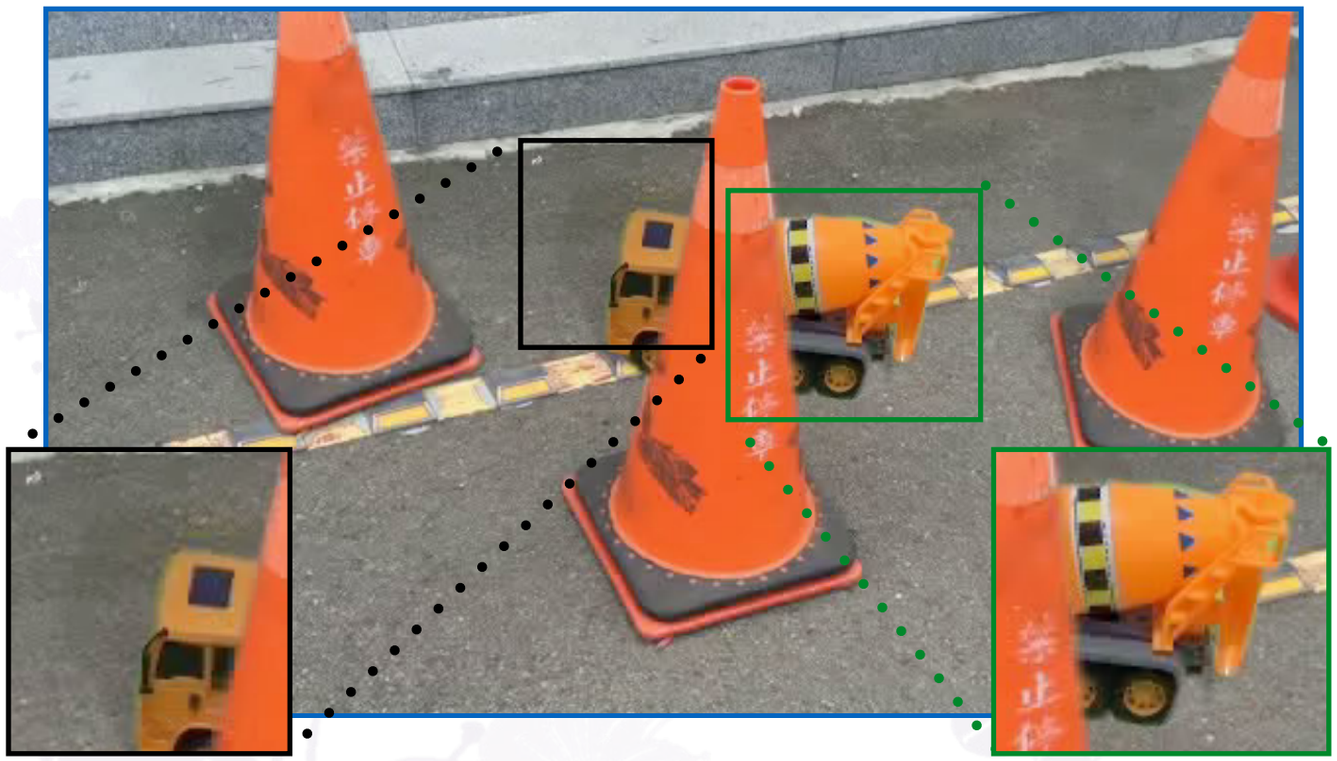
Object duplication	Object transfer	Keyframe animation

Key Issues

- The lack of the underlying 3D model of the scene

Key Issues

- The lack of the underlying 3D model of the scene



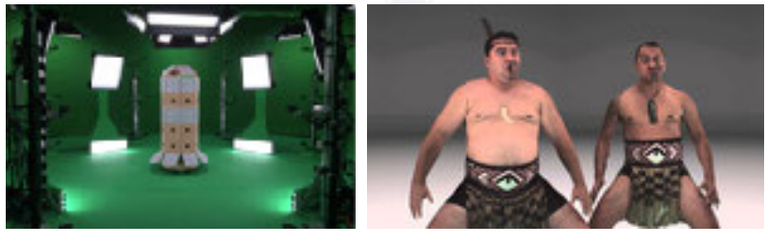
Ensure correct perspective

Handle occlusion effects

Update shadows

State-of-the-art Solutions

- A setup studio with specialized capturing equipments

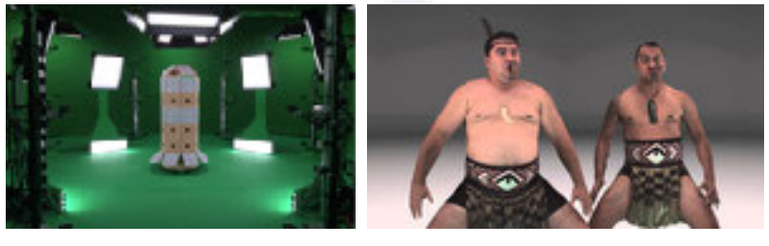


[Collet et al. 2015]

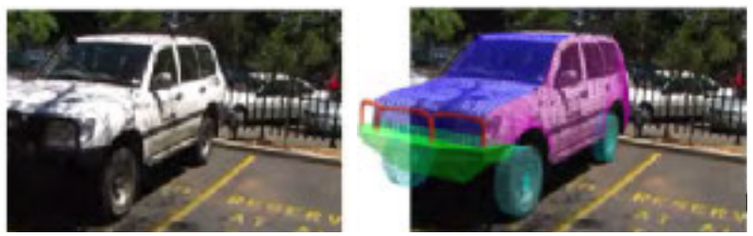
- Setup costs and extensive rigging

State-of-the-art Solutions

- A setup studio with specialized capturing equipments
  - Setup costs and extensive rigging
- Reconstruct 3D models of the scene from videos
  - Complex geometries
  - Significant post-processing



[Collet et al. 2015]



[Hengel et al. 2007]

[123D Catch]

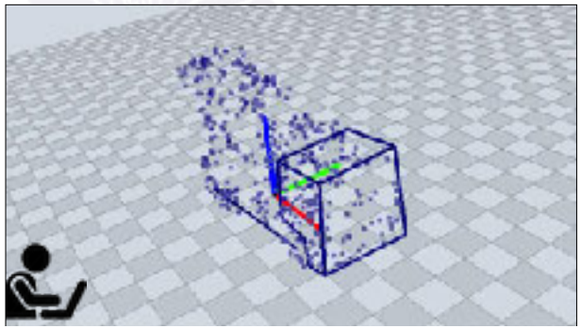
[Vi3Dim]

Our Key Idea

Input video

Our Key Idea

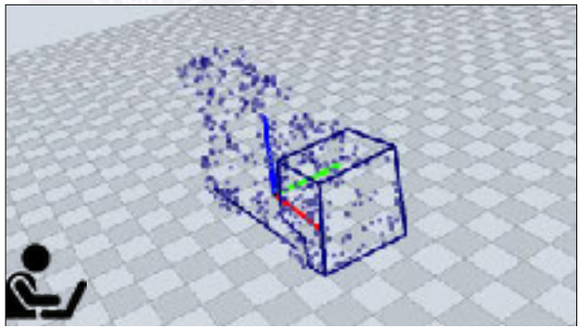
Input video



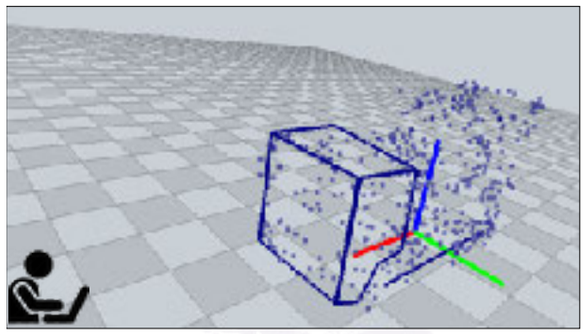
Sparse structure points (**SSP**)

Our Key Idea

Input video

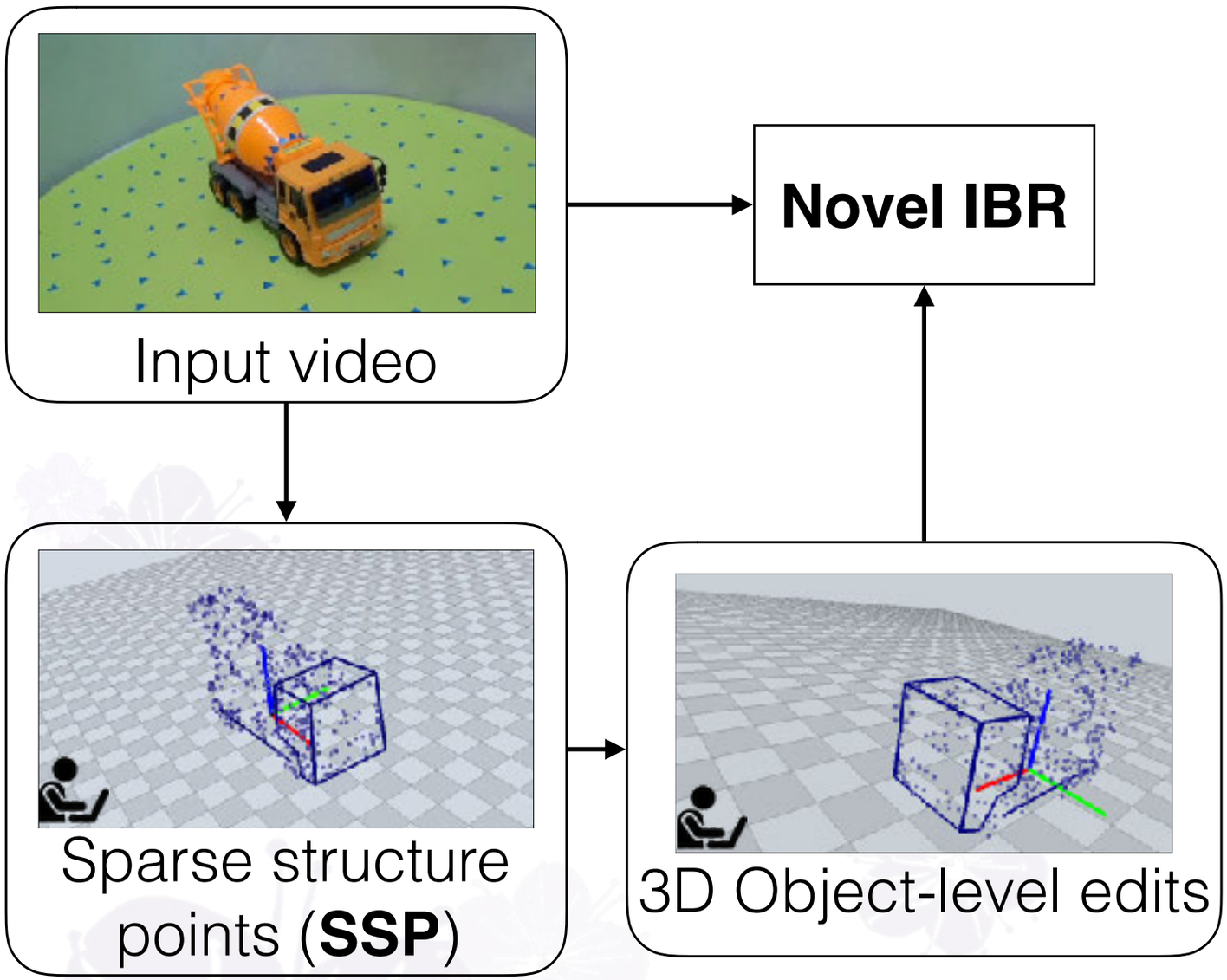


Sparse structure points (**SSP**)



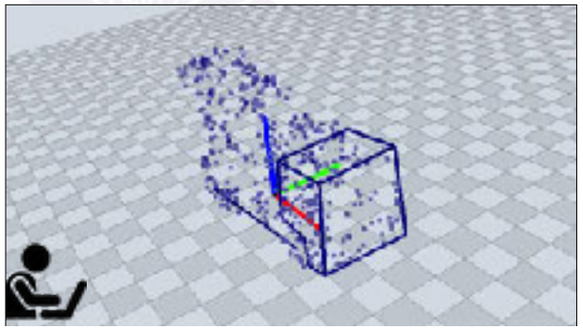
3D Object-level edits

Our Key Idea


Input video
Novel IBR
Sparse structure points (**SSP**)
3D Object-level edits

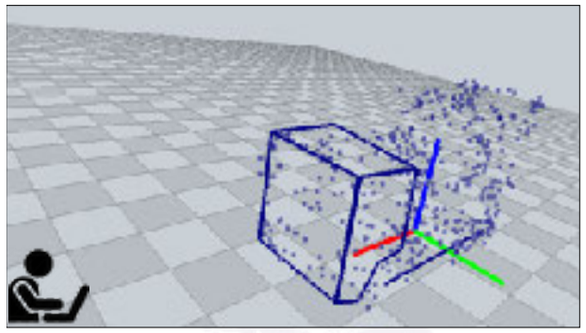
Our Key Idea

Input video



Sparse structure points (**SSP**)

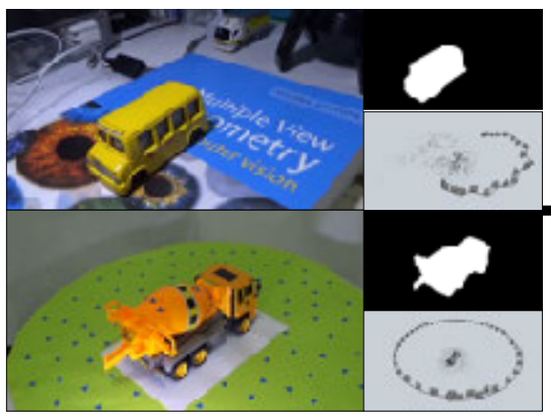
Novel IBR



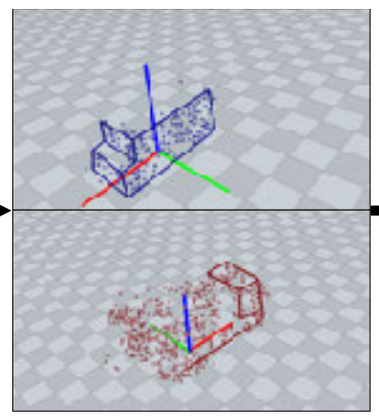
3D Object-level edits

Result

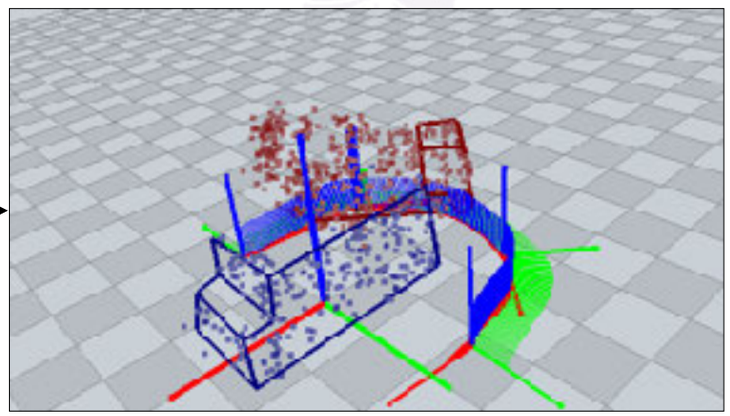
System Overview



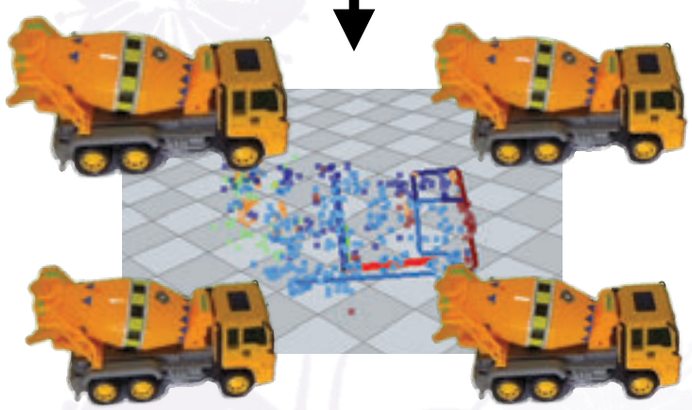
Preprocessing



Scene modeling



Object-level manipulation



Object frame retrieval

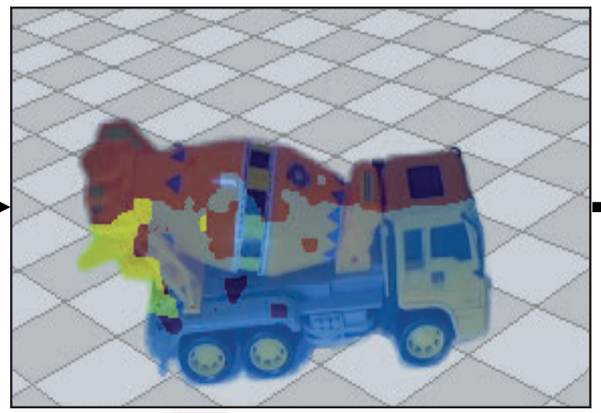
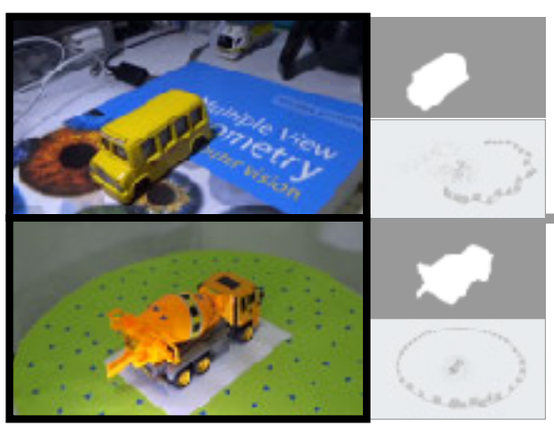


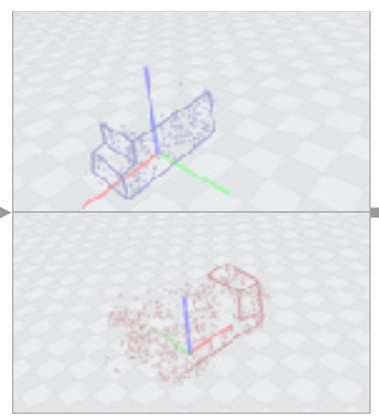
Image warping and stitching

Layer composition

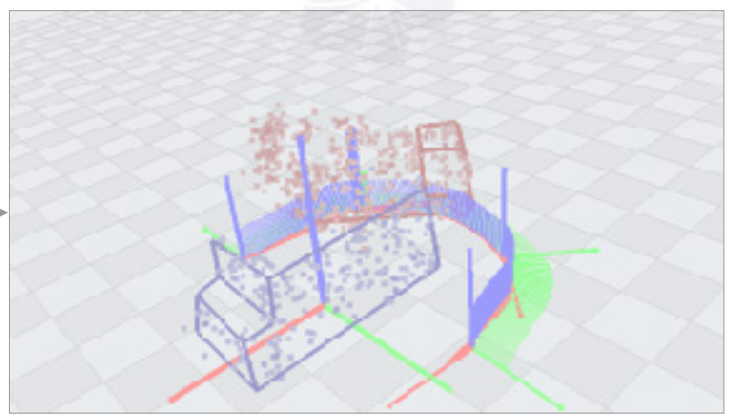
System Overview



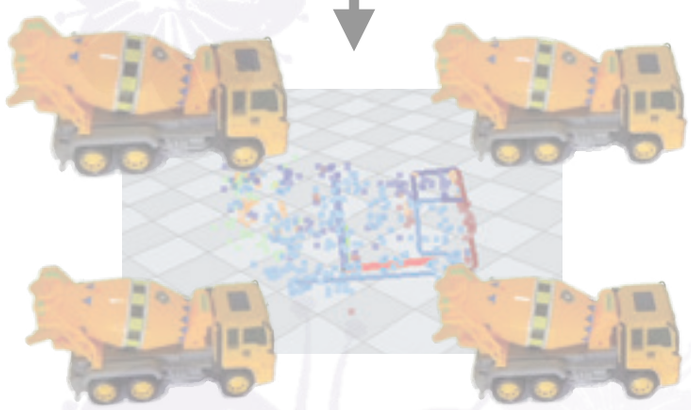
Preprocessing



Scene modeling



Object-level manipulation

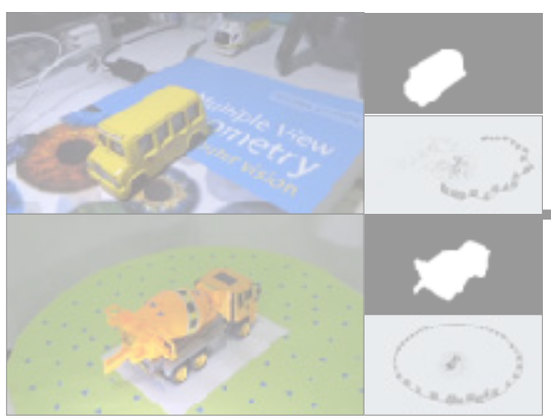


Object frame retrieval

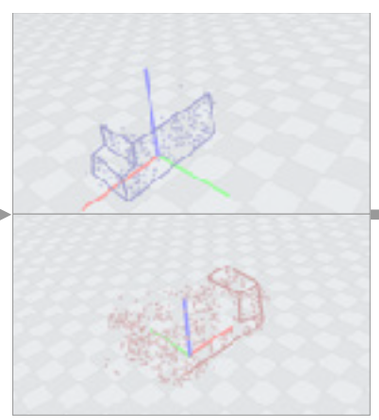
Image warping and stitching

Layer composition

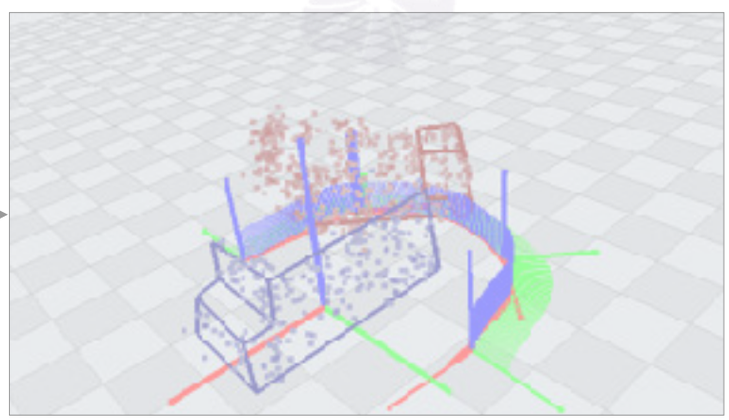
System Overview



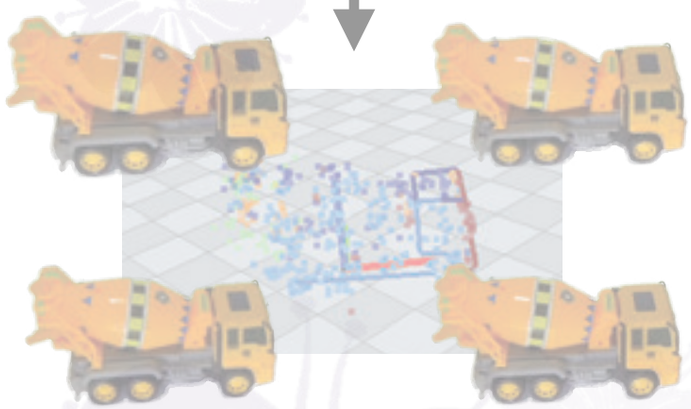
Preprocessing



Scene modeling



Object-level manipulation

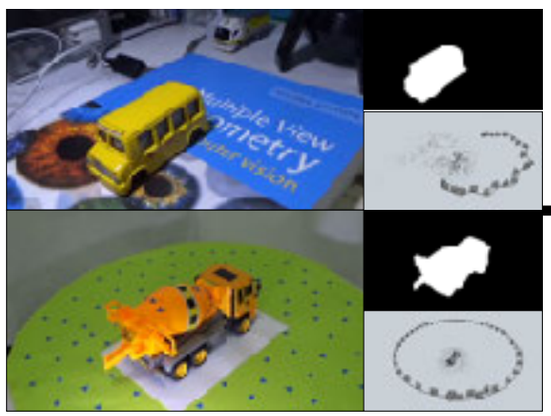


Object frame retrieval

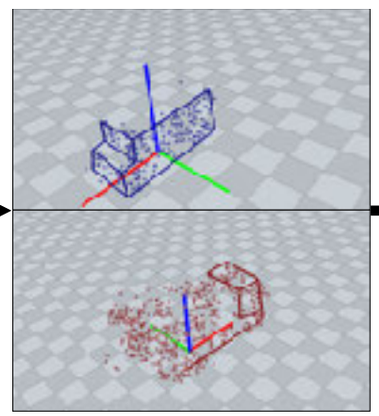
Image warping and stitching

Layer composition

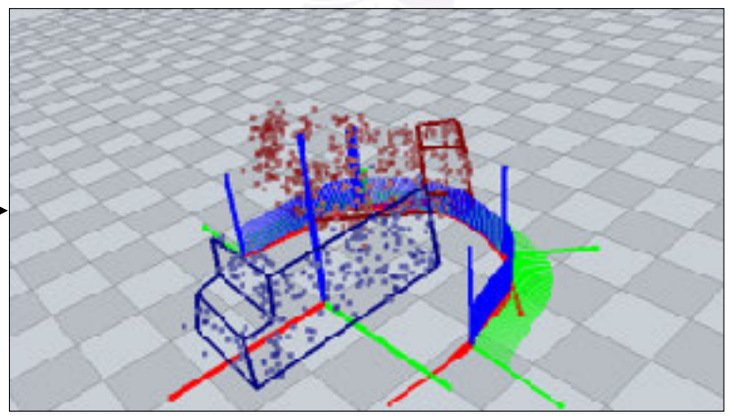
System Overview



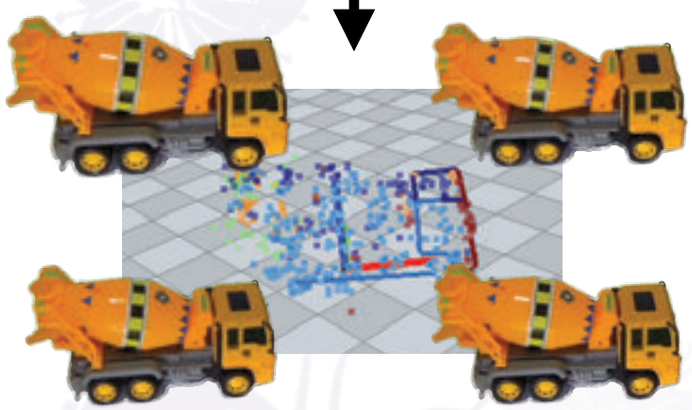
Preprocessing



Scene modeling



Object-level manipulation



Object frame retrieval

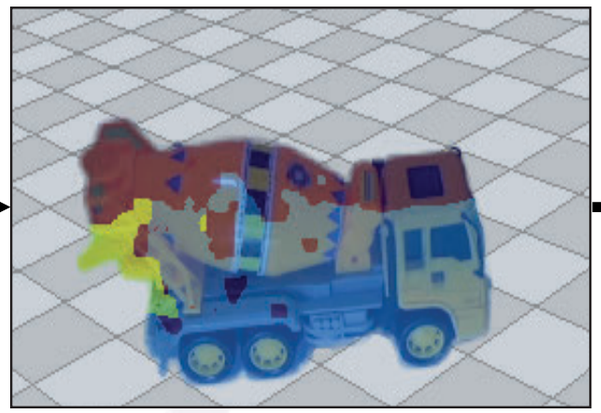
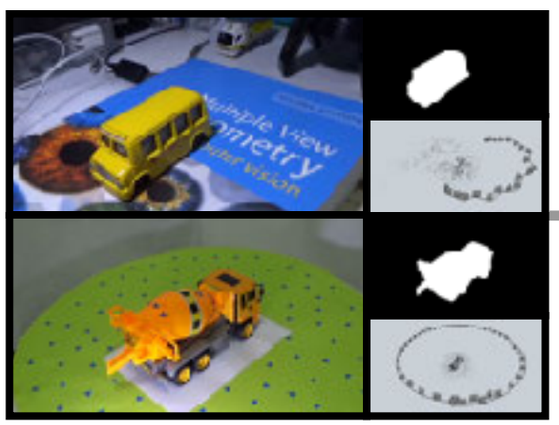


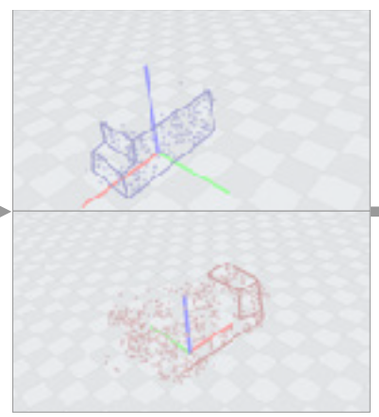
Image warping and stitching

Layer composition

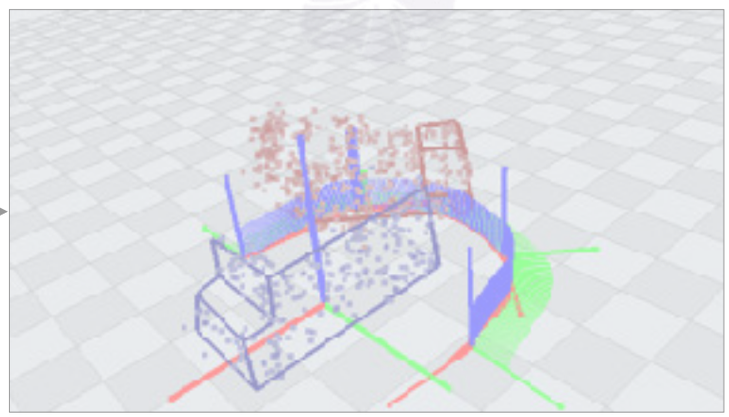
Preprocessing



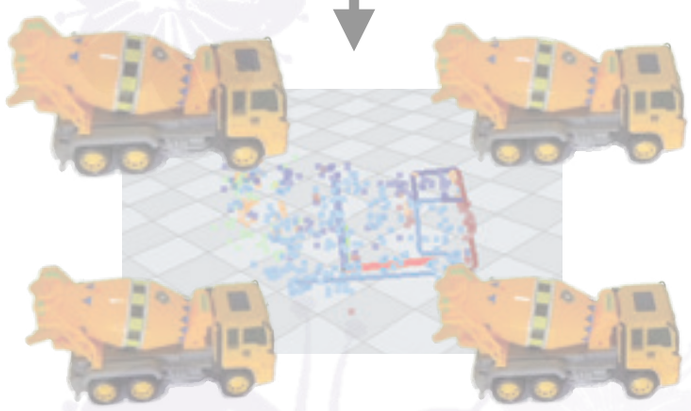
Preprocessing



Scene modeling



Object-level manipulation

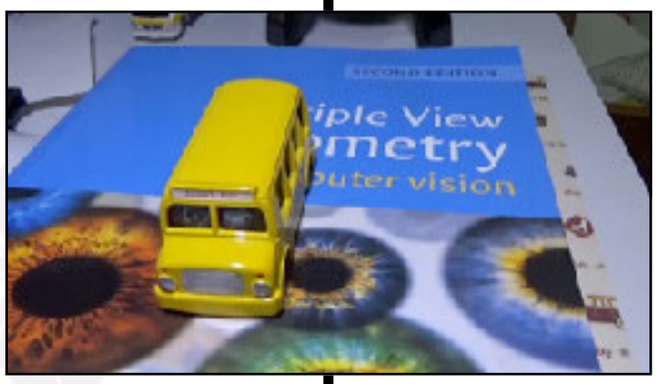


Object frame retrieval

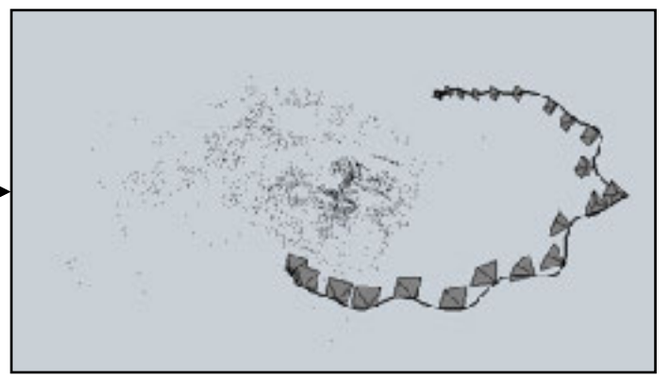
Image warping and stitching

Layer composition

Preprocessing

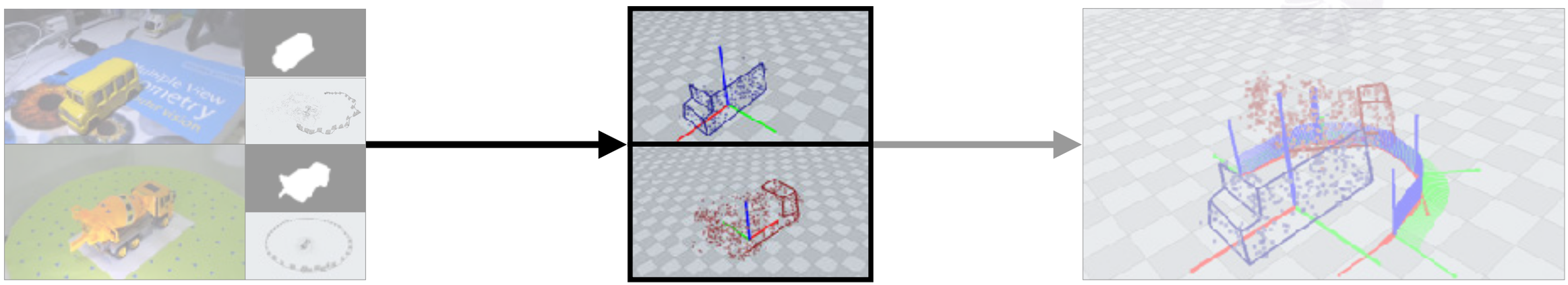


Video snapcut [Bai et al.]
Global matting [He et al.]



Voodoo camera tracker
[VISCODA GmbH]

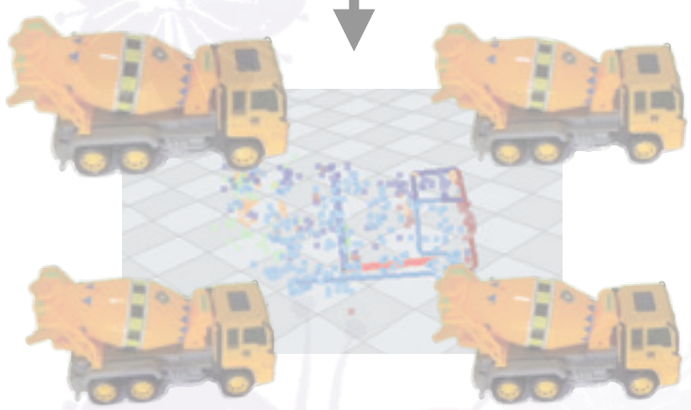
Scene Modeling



Preprocessing

Scene modeling

Object-level manipulation

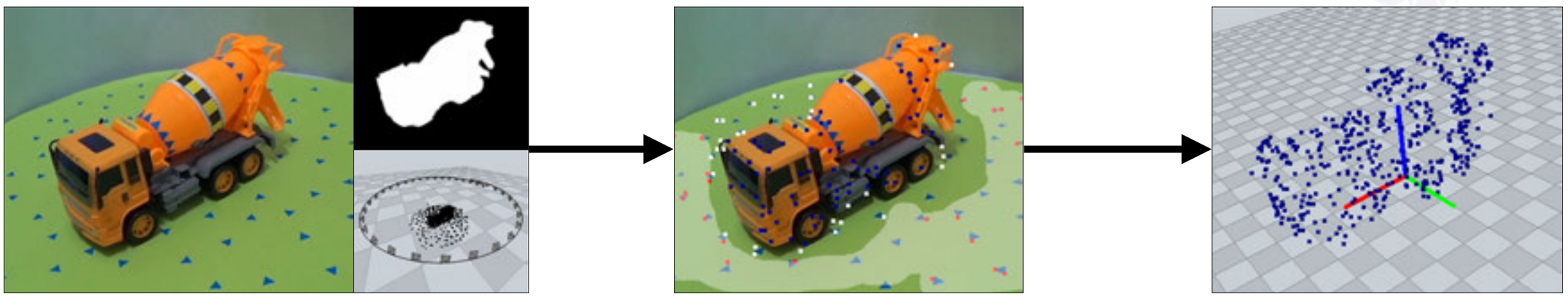


Object frame retrieval

Image warping and stitching

Layer composition

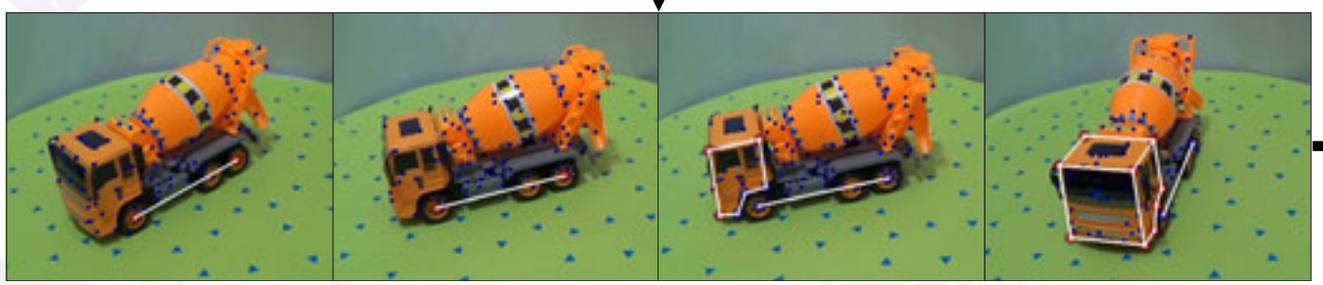
Scene Modeling



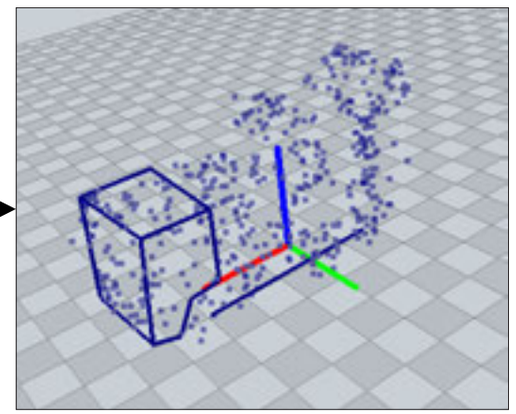
Preprocessed video

Label ground

Sparse Structure Points (**SSP**)

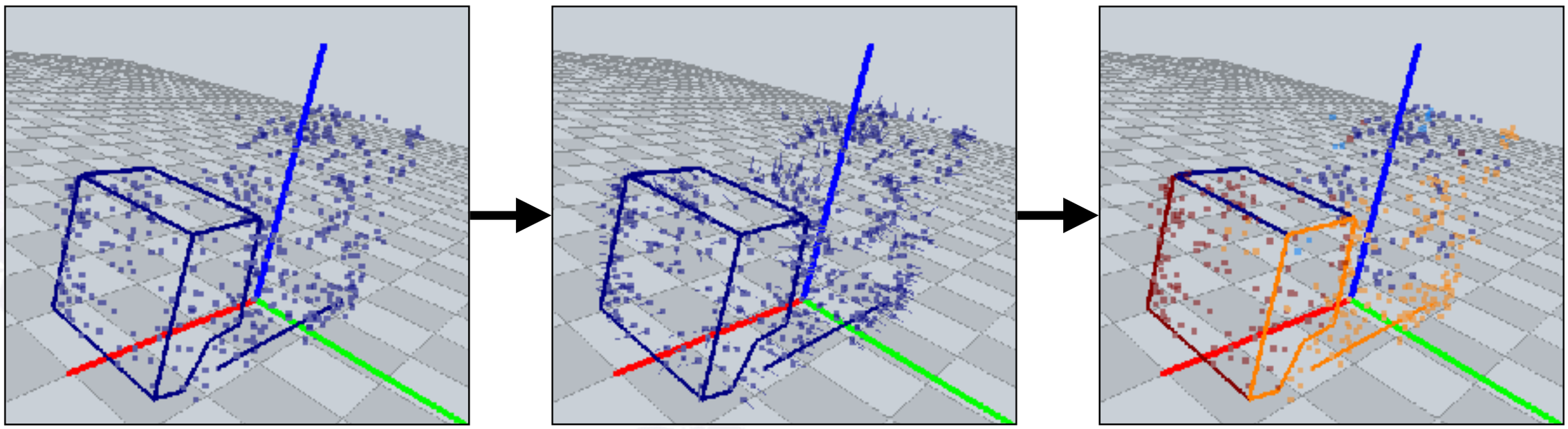


Label polygons and edges
[Hengel et al. 2007]



Polygons and edges

Proxy Geometry Decomposition

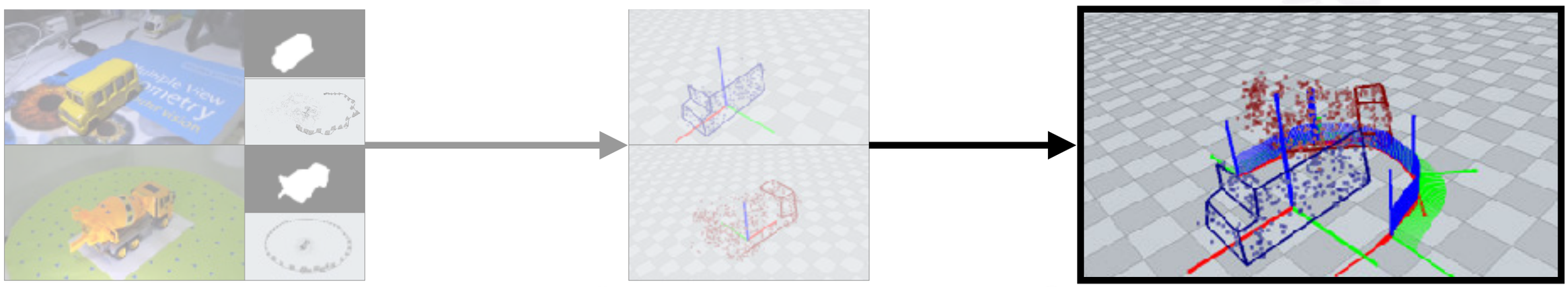


Proxy Geometry

Normal estimation
[Point Cloud Library]

Non-parametric
mean shift clustering
[Comaniciu et al. 2002]

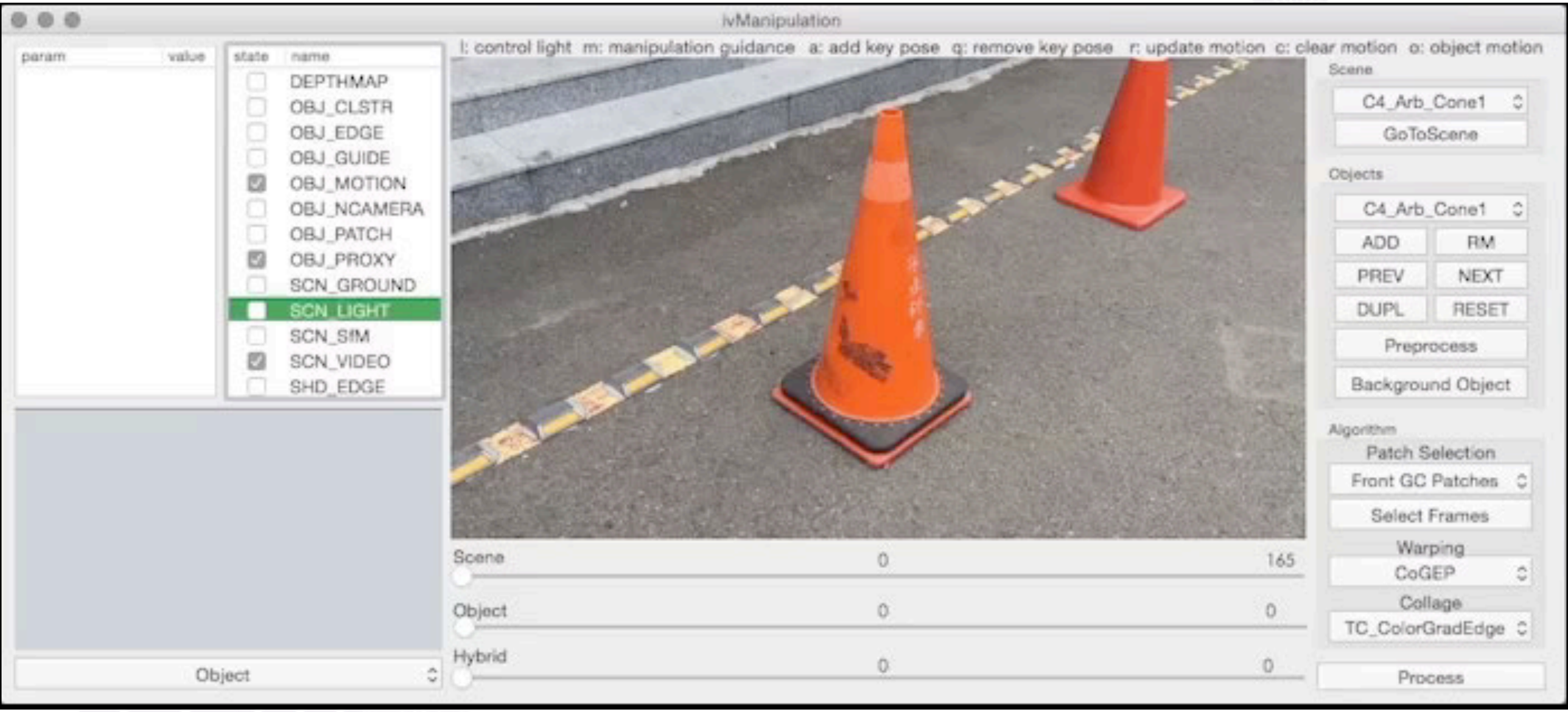
Object-level Manipulation



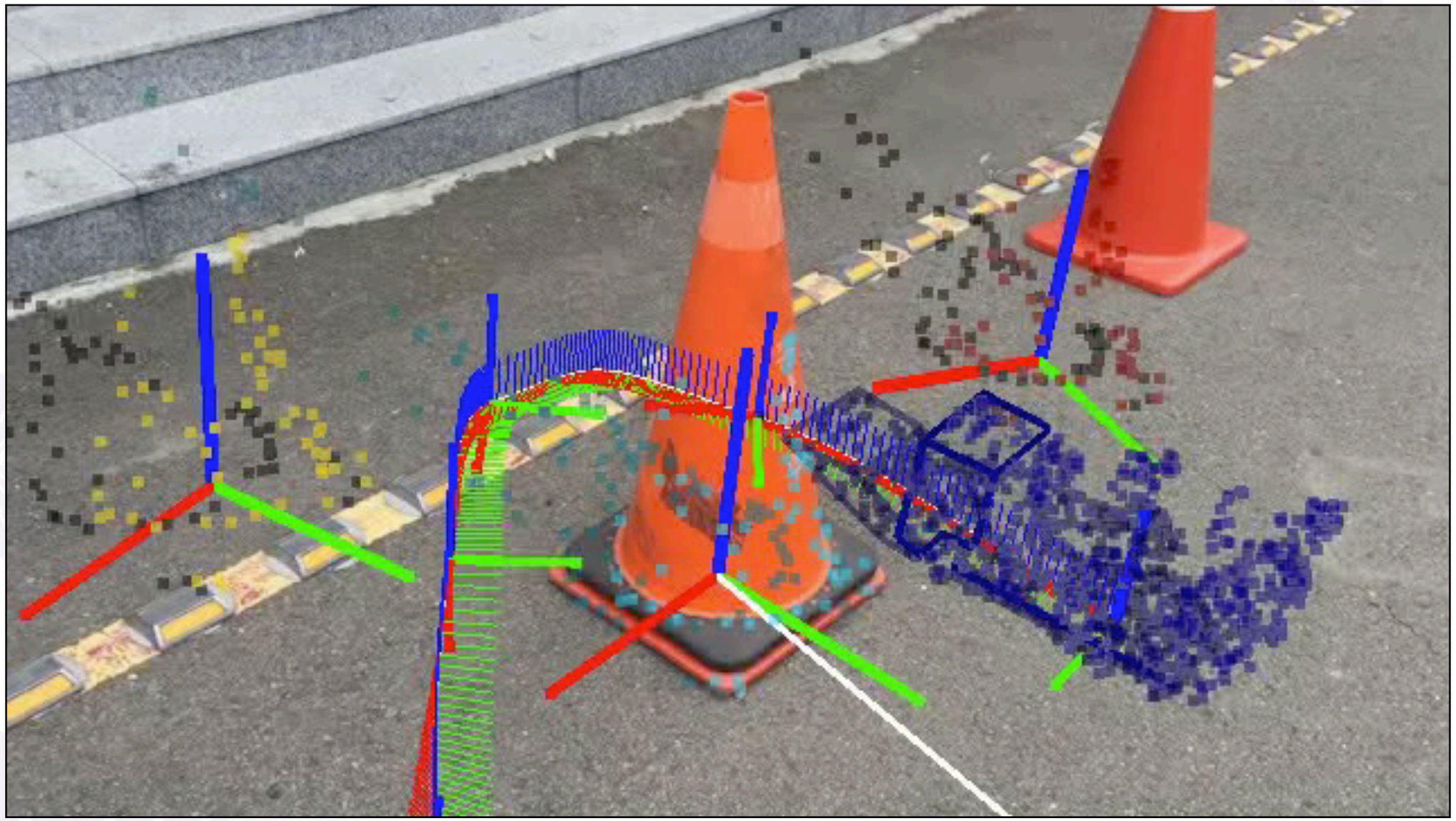
Object-level Manipulation

- Object level 3D manipulations
    - Apply 3D transformations
    - Specify keyframe animation of objects
    - Duplicate objects
    - Transfer object across videos
    - Specify the point light source

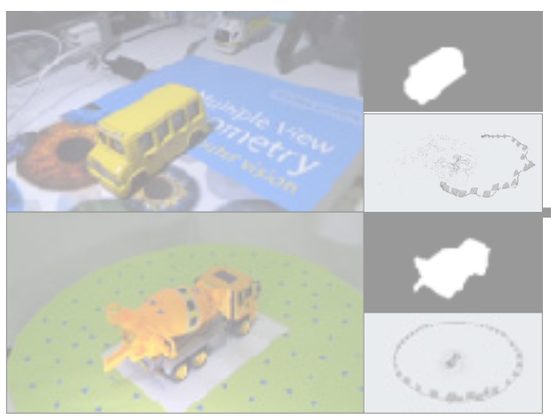
Object-level Manipulation



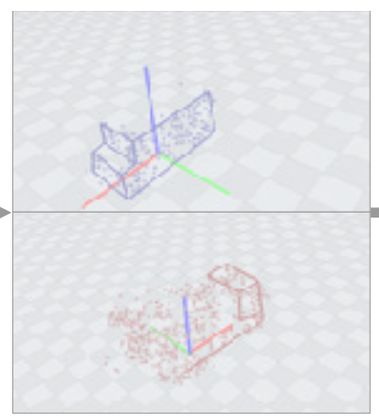
Object-level Manipulation



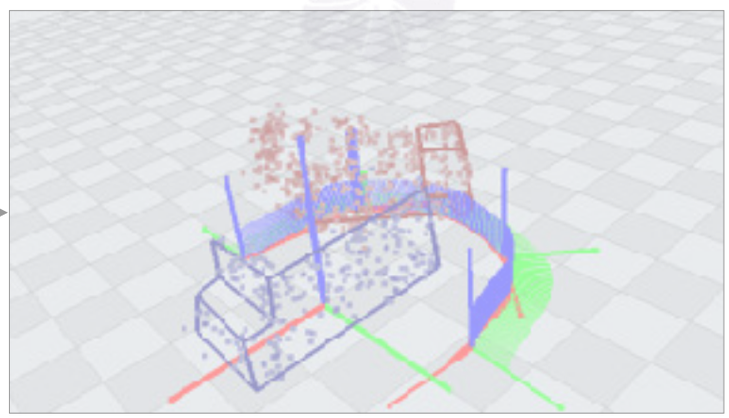
Object Frame Retrieval



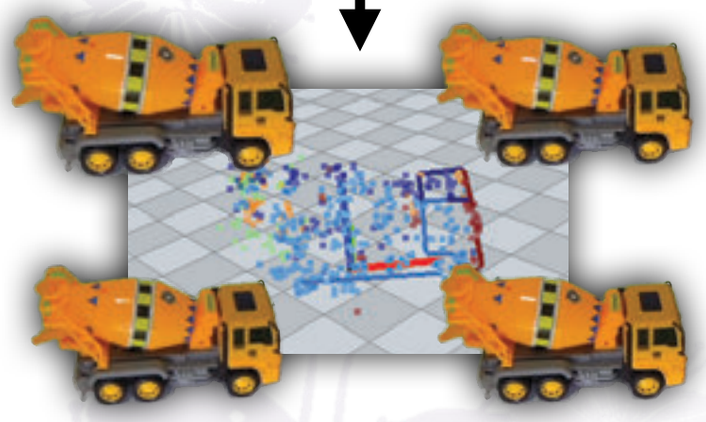
Preprocessing



Scene modeling



Object-level manipulation



Object frame retrieval

Image warping and stitching

Layer composition

Object Frame Retrieval Objectives

- **Frame similarity**
    - The camera view in the retrieved object frame should be close to that in target frame
- **Spatiotemporal smoothness**
    - Adjacent object frames should be selected for spatiotemporal neighbors

Energy terms

- **Frame similarity term**

$$E_{fs}(F) = \sum_{k=1}^{n_t} \sum_{i=1}^{n_x} D_{\text{cam}}(\mathbf{C}_k^t, \mathbf{C}_{\mathcal{F}_{i,k}}^o, \mathbf{P}_i)$$

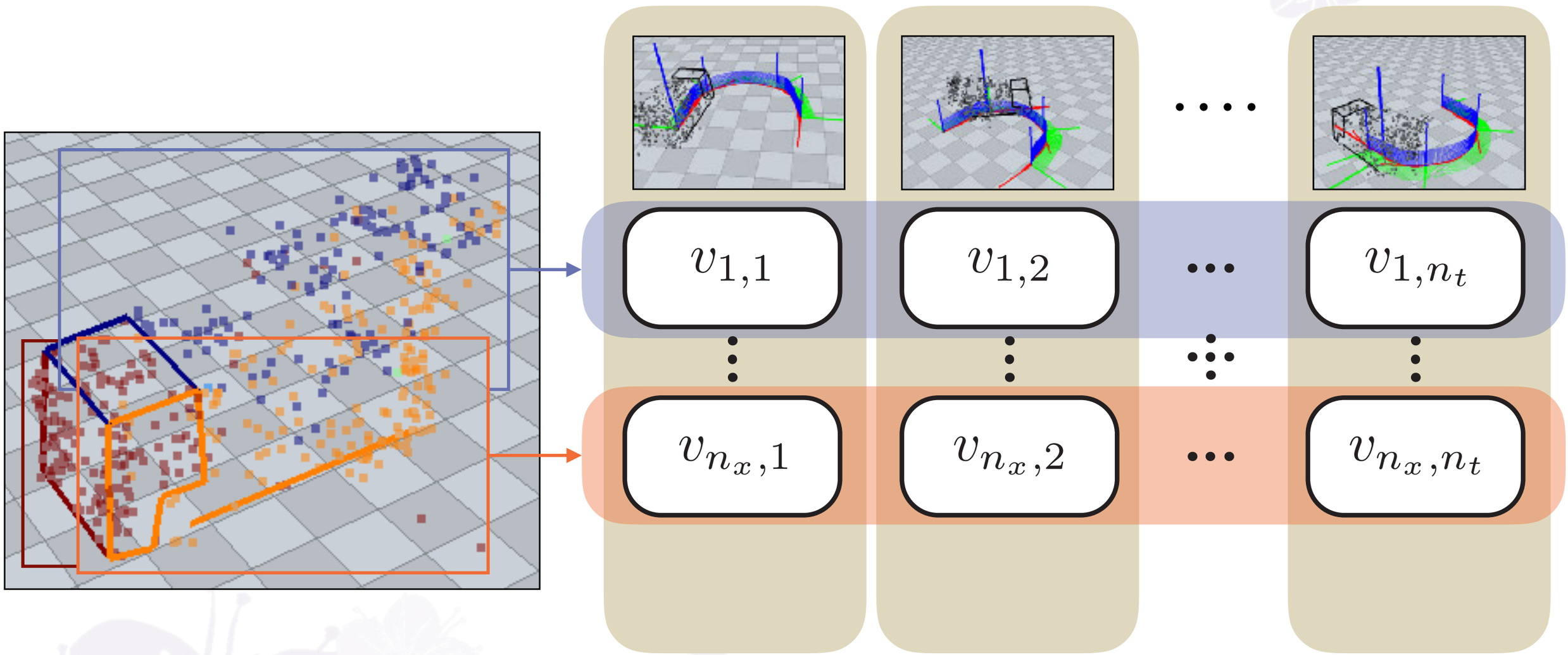
- **Spatial smoothness term**

$$E_{ss}(F) = \sum_{k=1}^{n_t} \sum_{i=1}^{n_x} D_{\text{ang}}(\mathbf{C}_{\mathcal{F}_{n_x+1,k}}^o, \mathbf{C}_{\mathcal{F}_{i,k}}^o, 0)\delta(\mathcal{F}_{n_x+1,k}, \mathcal{F}_{i,k})$$

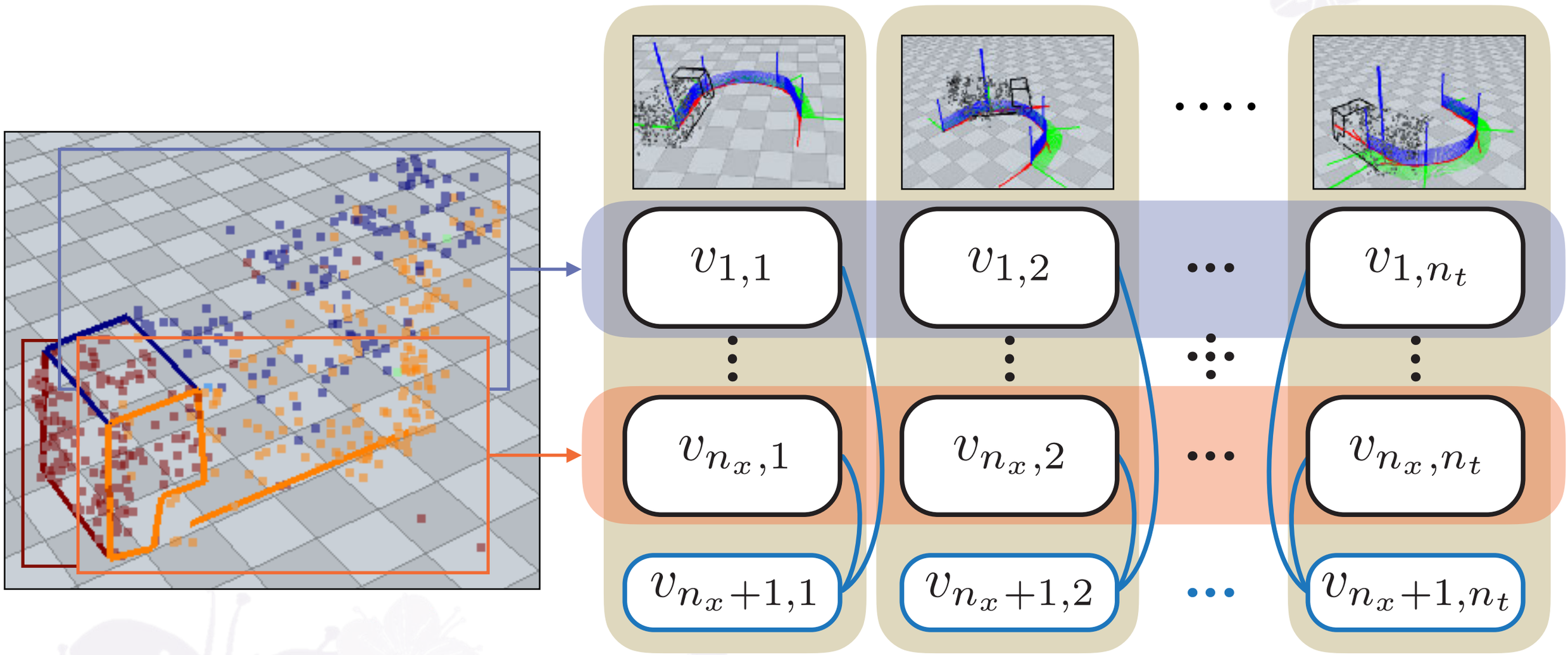
- **Temporal smoothness term**

$$E_{ts}(F) = \sum_{k=1}^{n_t-1} \sum_{i=1}^{n_x} D_{\text{ang}}(\mathbf{C}_{\mathcal{F}_{i,k}}^o, \mathbf{C}_{\mathcal{F}_{i,k+1}}^o, 0)\delta(\mathcal{F}_{i,k}, \mathcal{F}_{i,k+1})$$

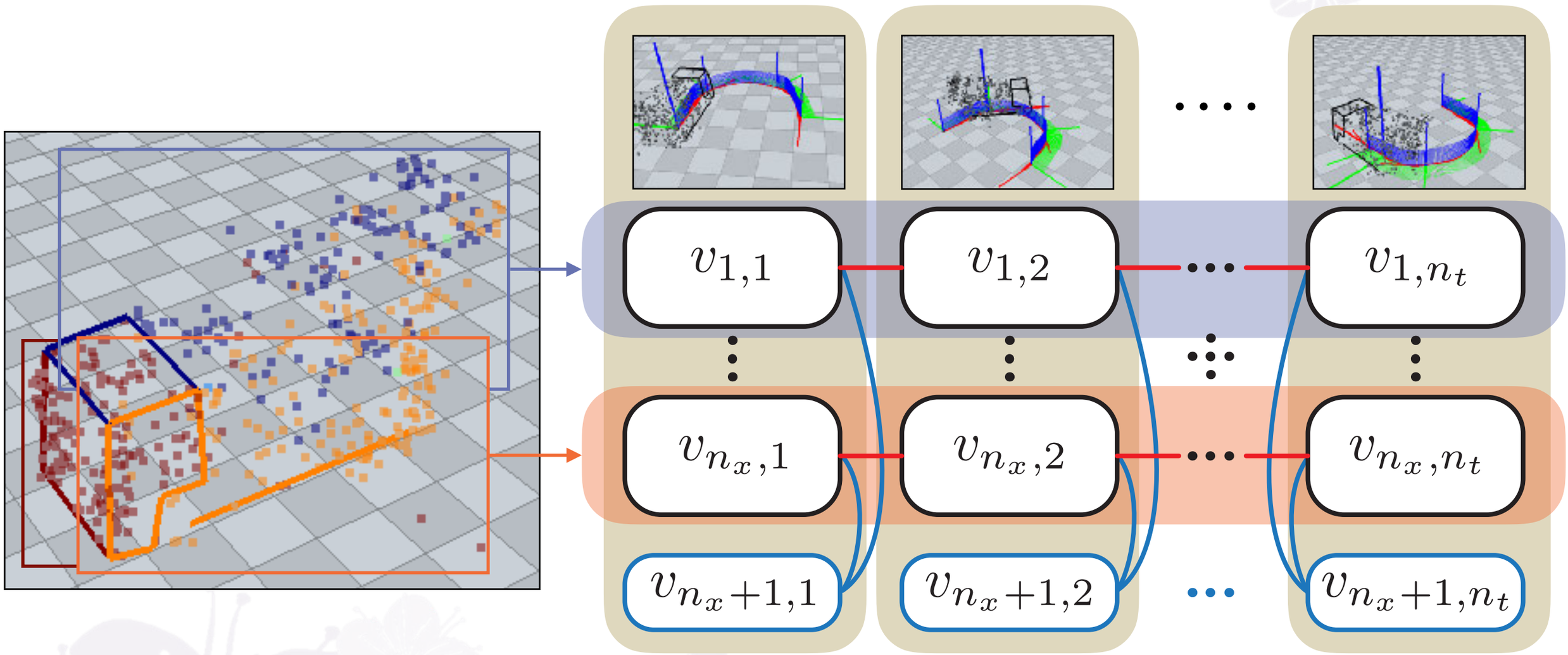
Multiple Label MRF



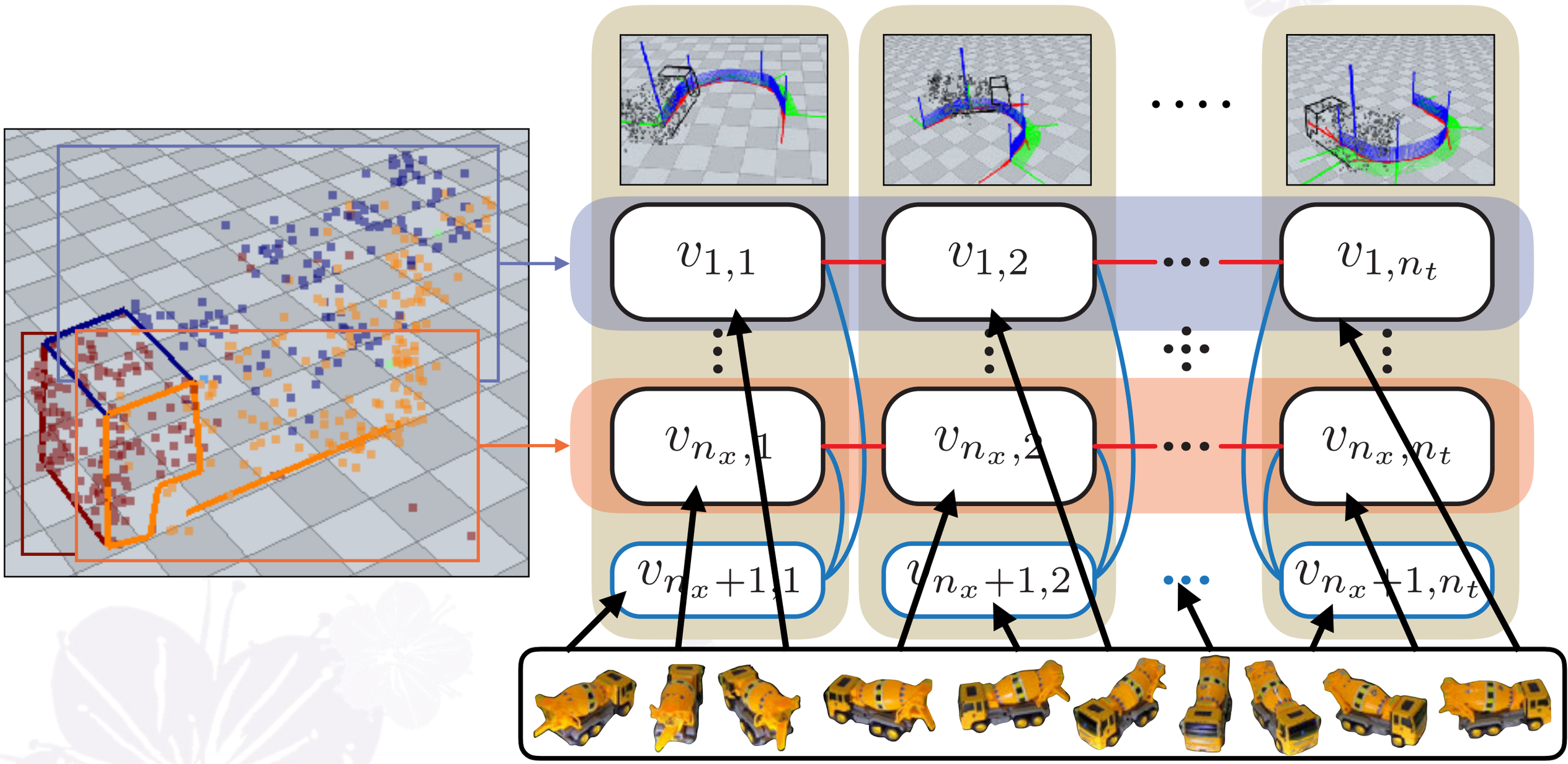
Spatial Edges



Temporal Edges



Multiple Label MRF



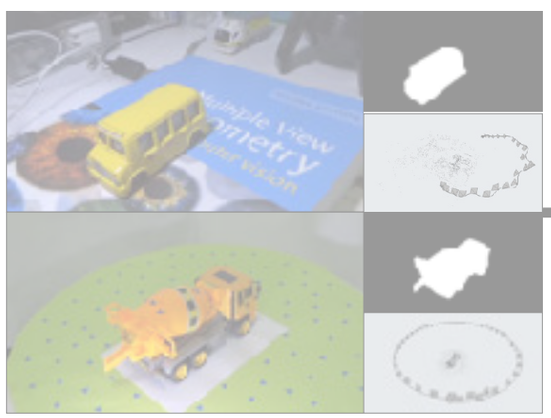
Optimization

- Total energy

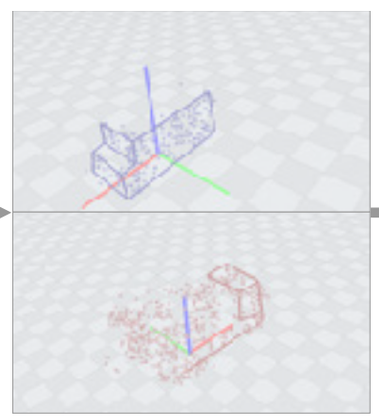
$$F^* = \underset{F}{\operatorname{argmin}}[E_{fs}(F) + \lambda_s(E_{ss}(F) + E_{ts}(F))]$$

- Multi-label graph cut algorithm [Boykov et al. 2001]

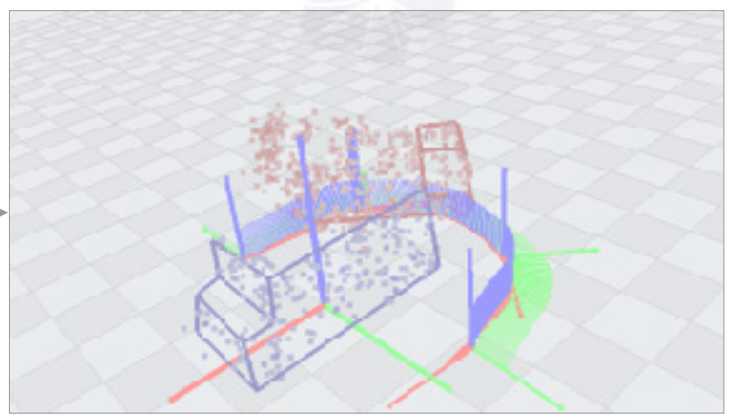
Image Warping and Stitching



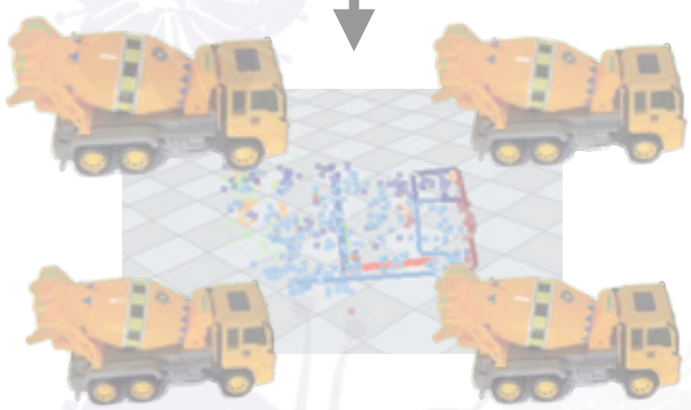
Preprocessing



Scene modeling



Object-level manipulation



Object frame retrieval

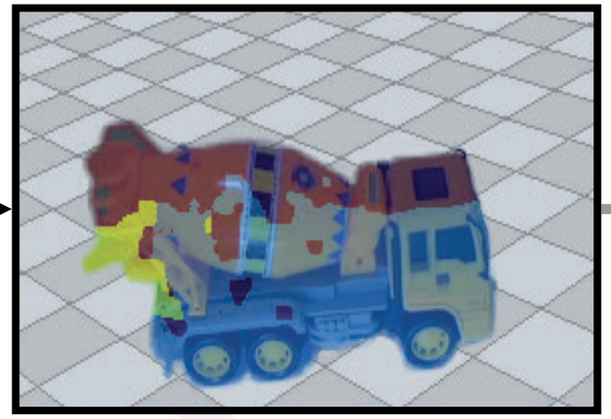
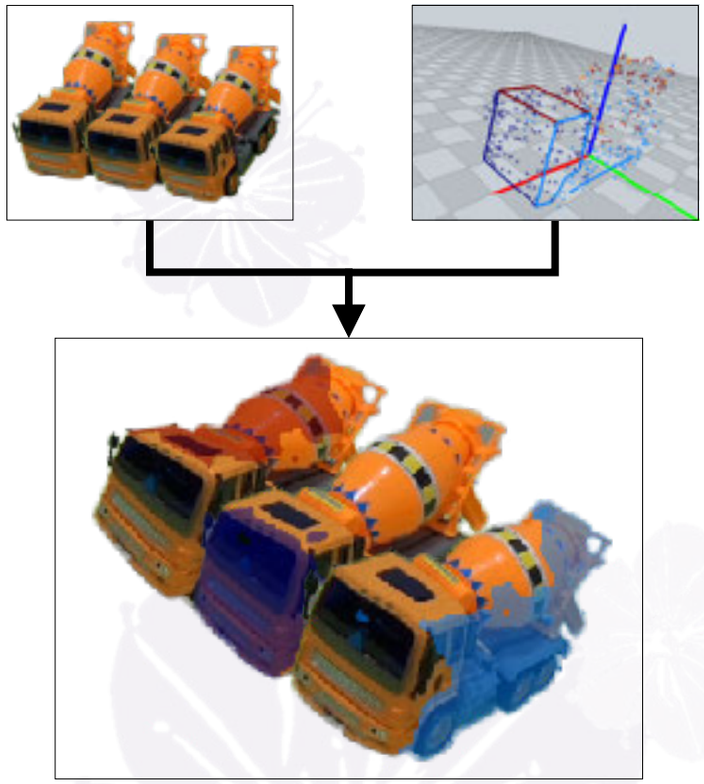


Image warping and stitching

Layer composition

Structure Preserving Image Warping

- We propose a novel structure-preserving image warping that augments the existing system.



[Liu et al. 2009]

Point alignment

- Align the feature points

Edge preservation

- Preserve the labeled edges

[Igarashi et al. 2005]

Similarity transform

- Avoid patch distortion

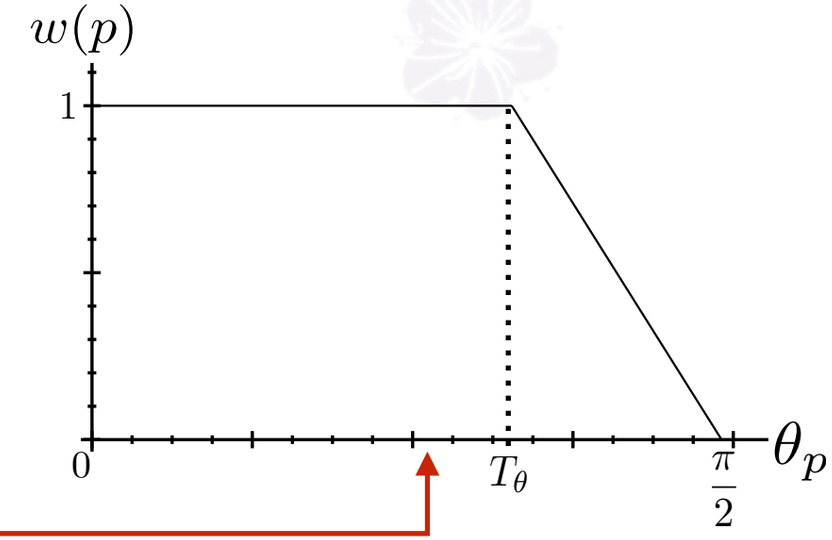
Structure preservation

- Combine the warped patches

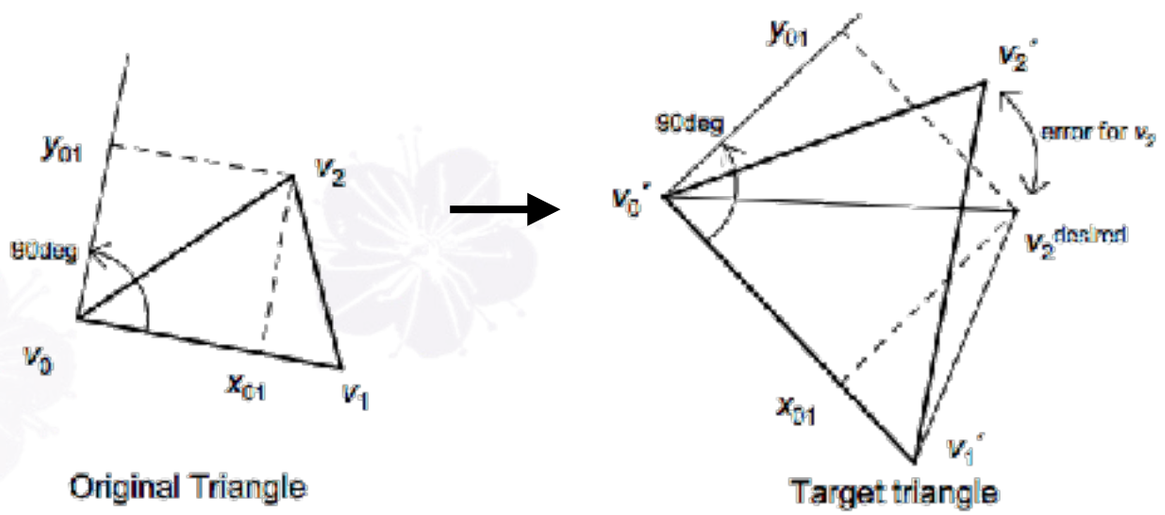
Point Alignment / Similarity Transform

- **Point alignment** [Liu et al. 2009]

$$E_{pa}(V) = \frac{1}{|\tilde{\mathbf{P}}_{\mathcal{F}^o}|} \sum_{p_i \in \tilde{\mathbf{P}}_{\mathcal{F}^o}} \omega(p_i) \|w_i^T V_i - \bar{p}_i^t\|^2$$



- **Local similarity transformation** constraint [Igarashi et al. 2005]



Edge Preservation Term

- 2D Edge Equation e

$$l(x, y, \alpha, b) : \sin(\alpha)x - \cos(\alpha)y + b = 0$$

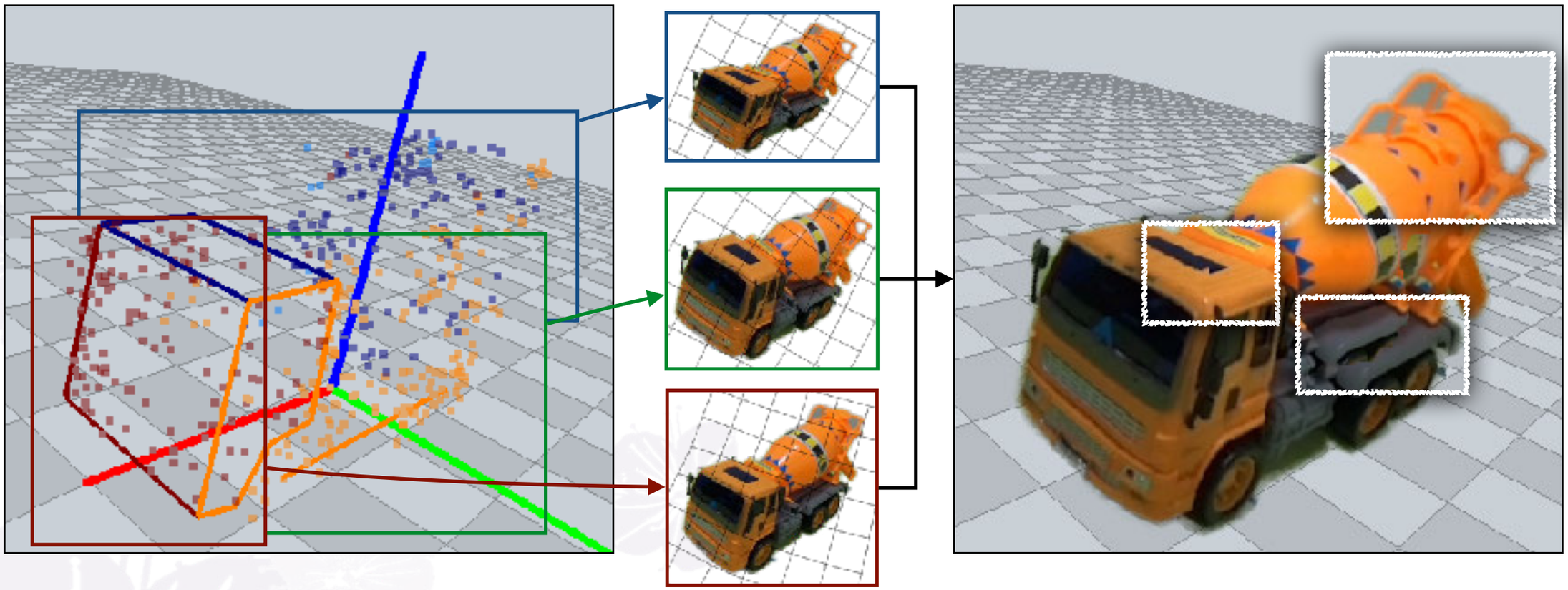
- We uniformly sample a set of 2D points $\mathbf{P}_e$ on each edge e

- **Edge preservation term**

$$E_{ep}(V) = \frac{1}{|\tilde{\mathbf{E}}_{\mathcal{F}^o}|} \sum_{e \in \tilde{\mathbf{E}}_{\mathcal{F}^o}} \sum_{p \in \mathbf{P}_e} l(S_x(p), S_y(p), \alpha_e, b_e)^2$$

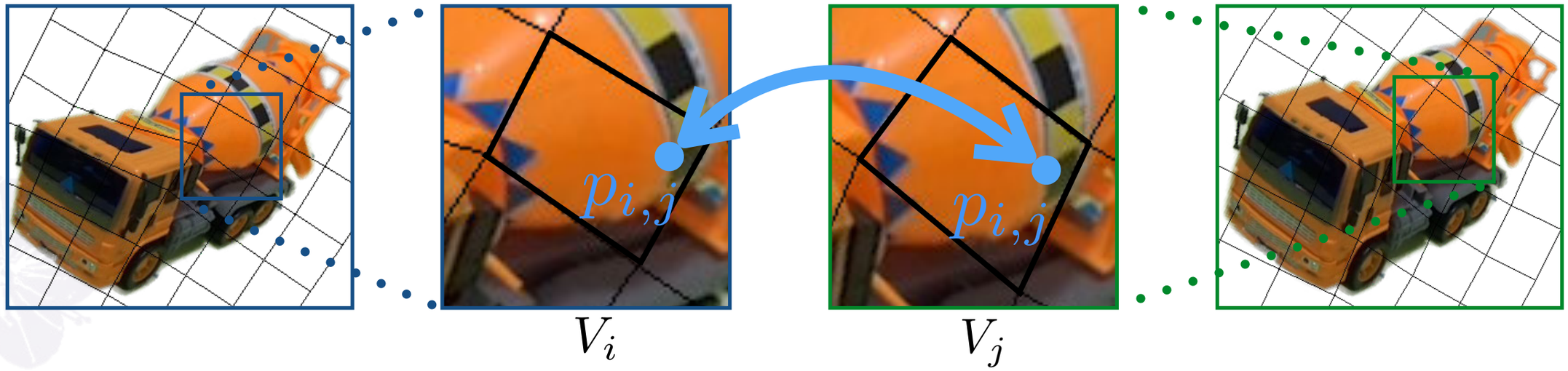
Warping Misalignment

- The misalignment among individually warped images



Structure Preservation Term

- Correlate the individual warps by the shared points

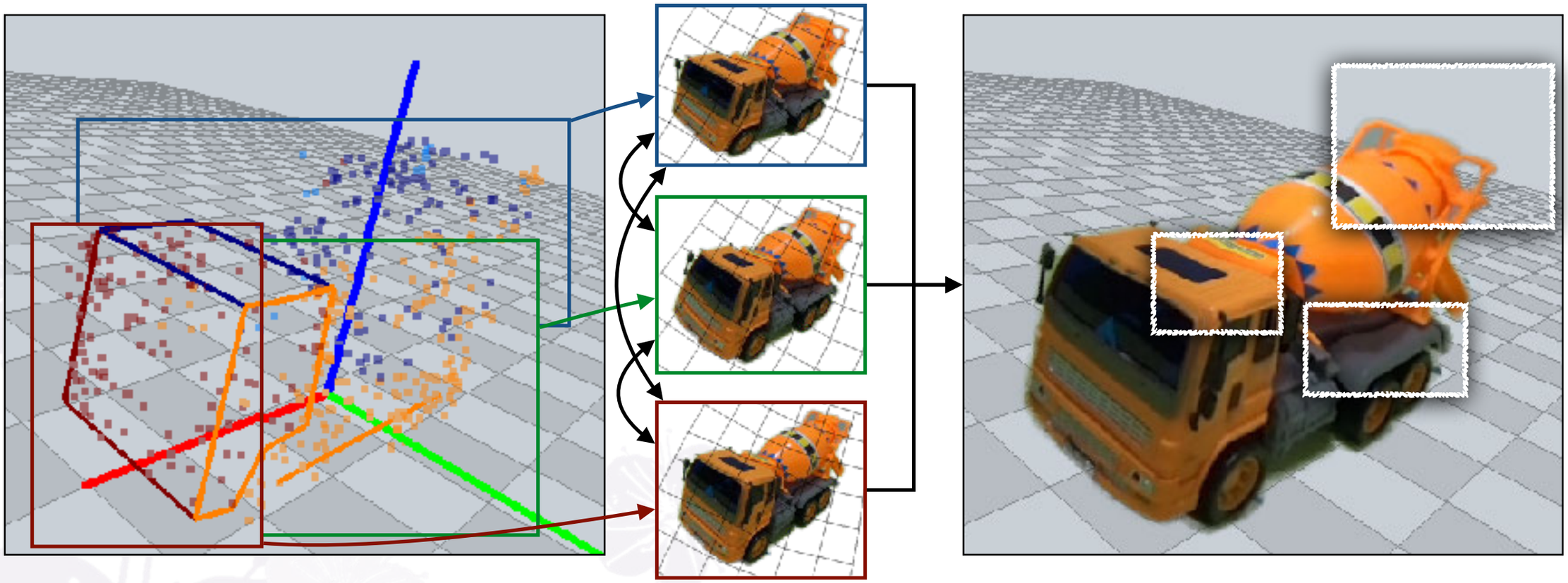


- **Structure preservation term**

$$E_{sp} = \frac{1}{|\mathbf{P}_s|} \sum_{p_{i,j} \in \mathbf{P}_s} \|w_i^T V_i - w_j^T V_j\|^2$$

Structure Preservation

- Align the shared points among the independently warped images



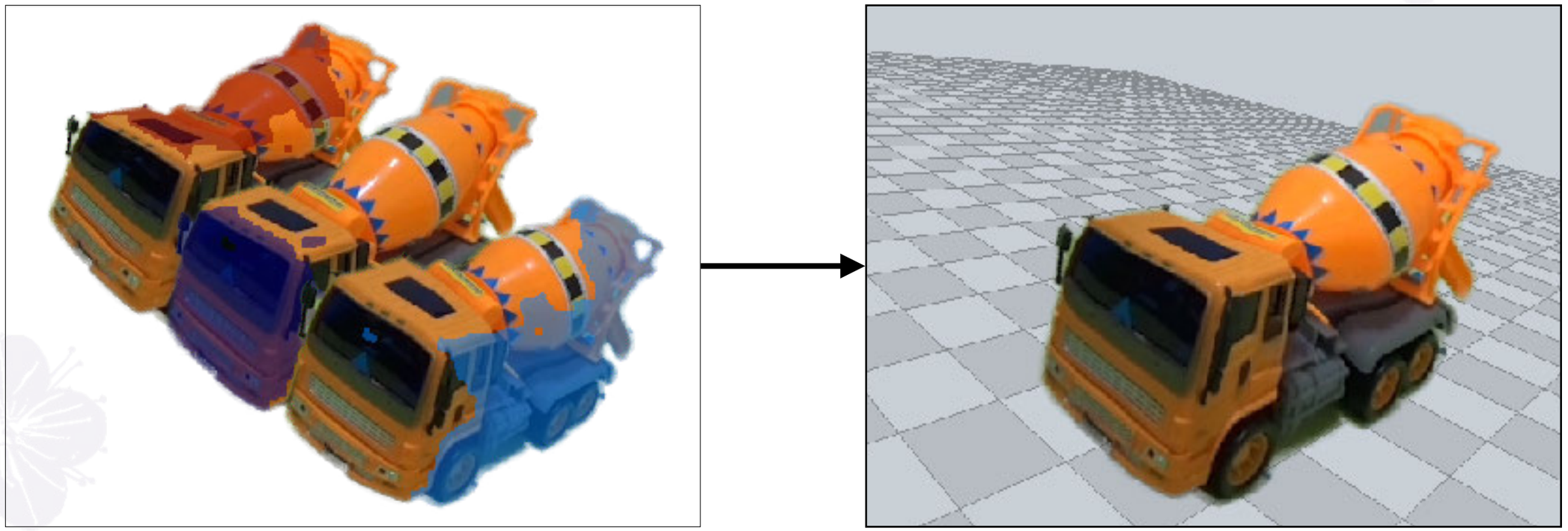
Energy Optimization

- The total energy becomes

$$\underset{\{V_1,\ldots,V_{n_x}\}}{\operatorname{argmin}} \; [\lambda_{sp}E_{sp} + \sum_{i=1}^{n_x} \beta_i(\lambda_{pa}E_{pa}(V_i) + \lambda_{st}E_{st}(V_i) + \lambda_{ep}E_{ep}(V_i))]$$

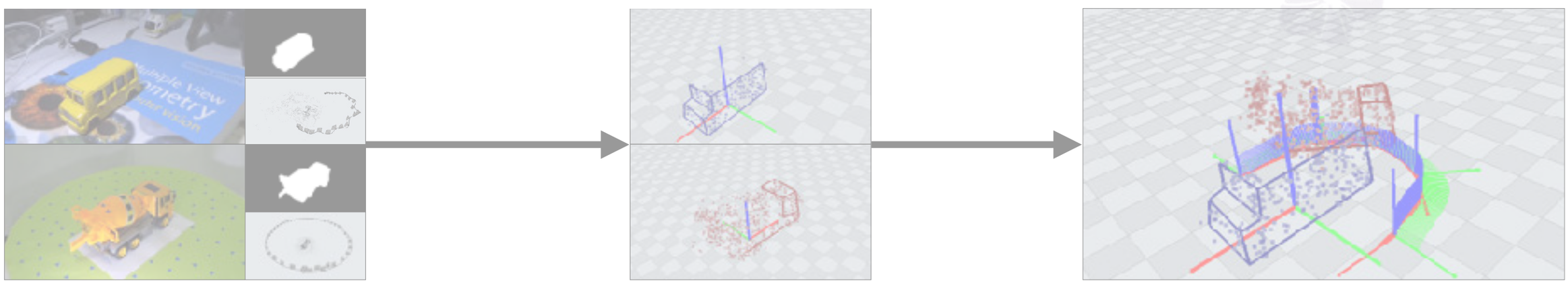
which can be solved efficiently using standard sparse linear system solver [Eigen Library].

Spatiotemporally Coherent Image Stitching



- To collage the warped patches, we follow the same formulation in [Agarwala et al. 2004], and consider the temporal smoothness.

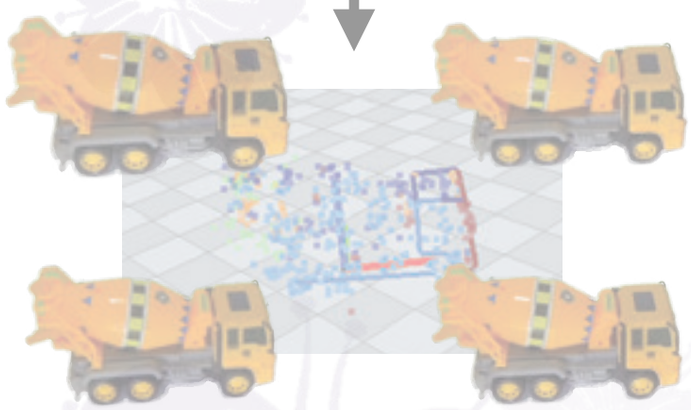
Layer Composition



Preprocessing

Scene modeling

Object-level manipulation



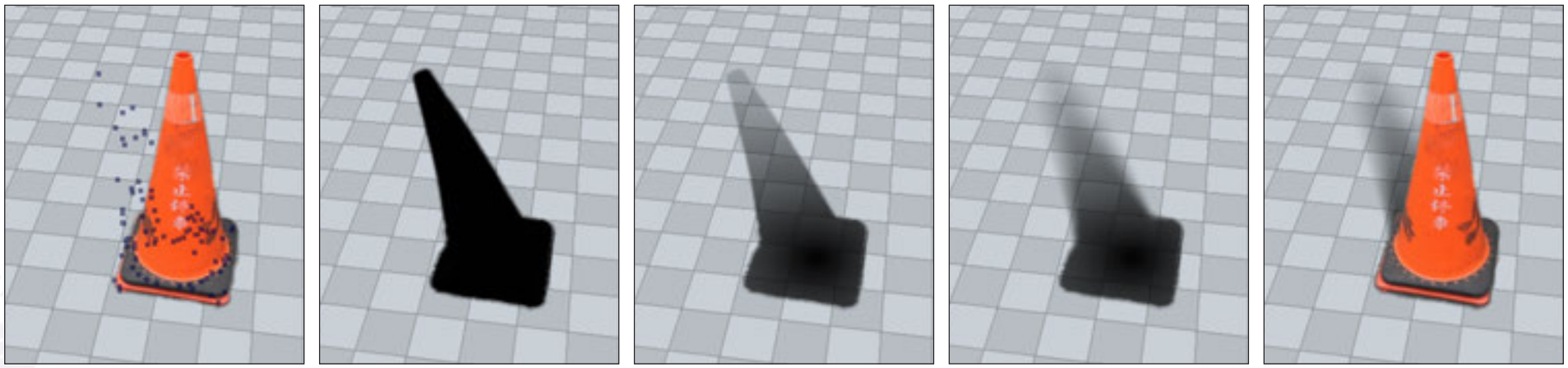
Object frame retrieval

Image warping and stitching

Layer composition

Layer Composition

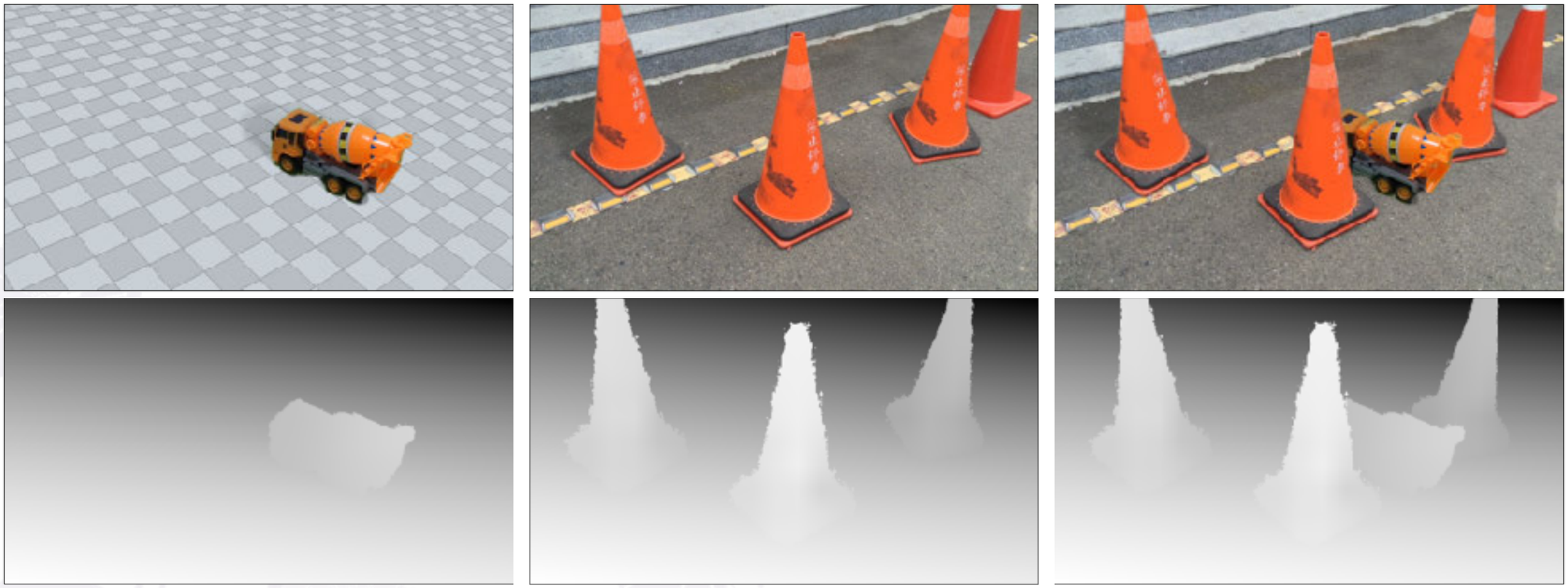
- Shadow map synthesis



- Depth interpolation [Kopf et al. 2014]

Layer Composition

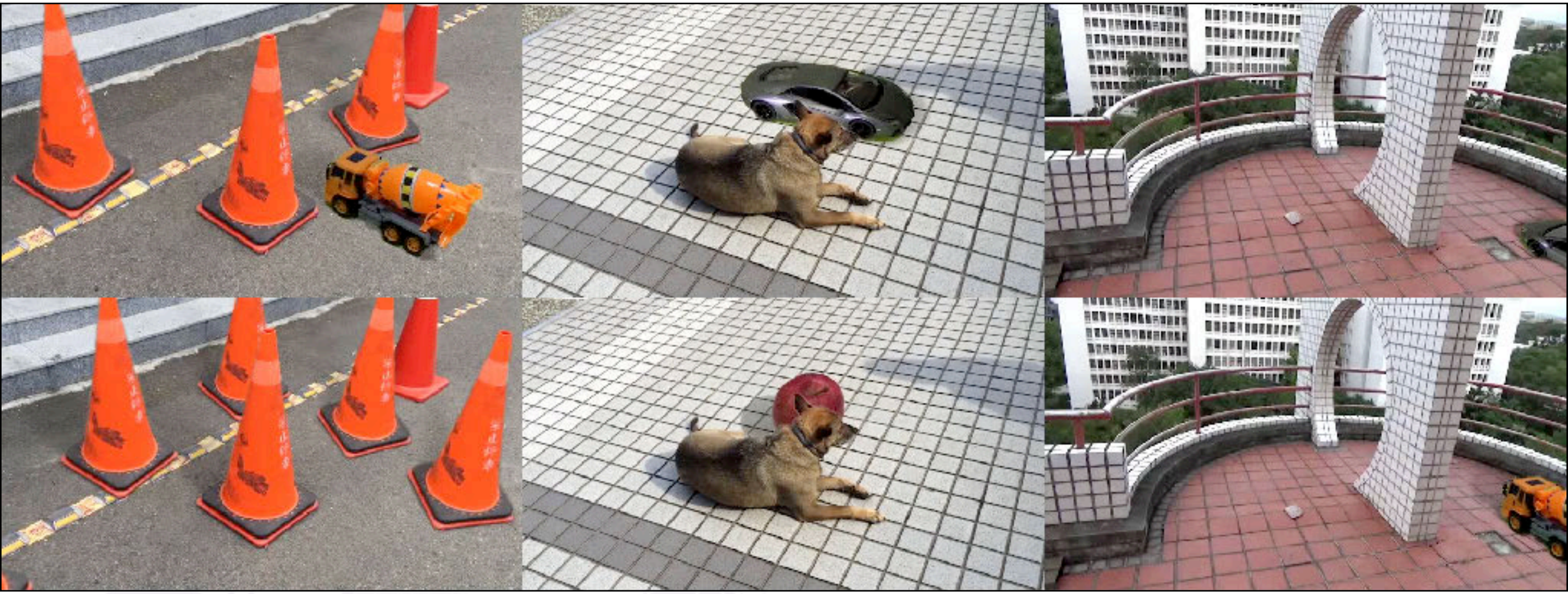
- Layer Composition



Object Scene Result

Results

Results



Object Transfer

Input

Result

Keyframe Animation

Input

Result

Object Duplication

Input

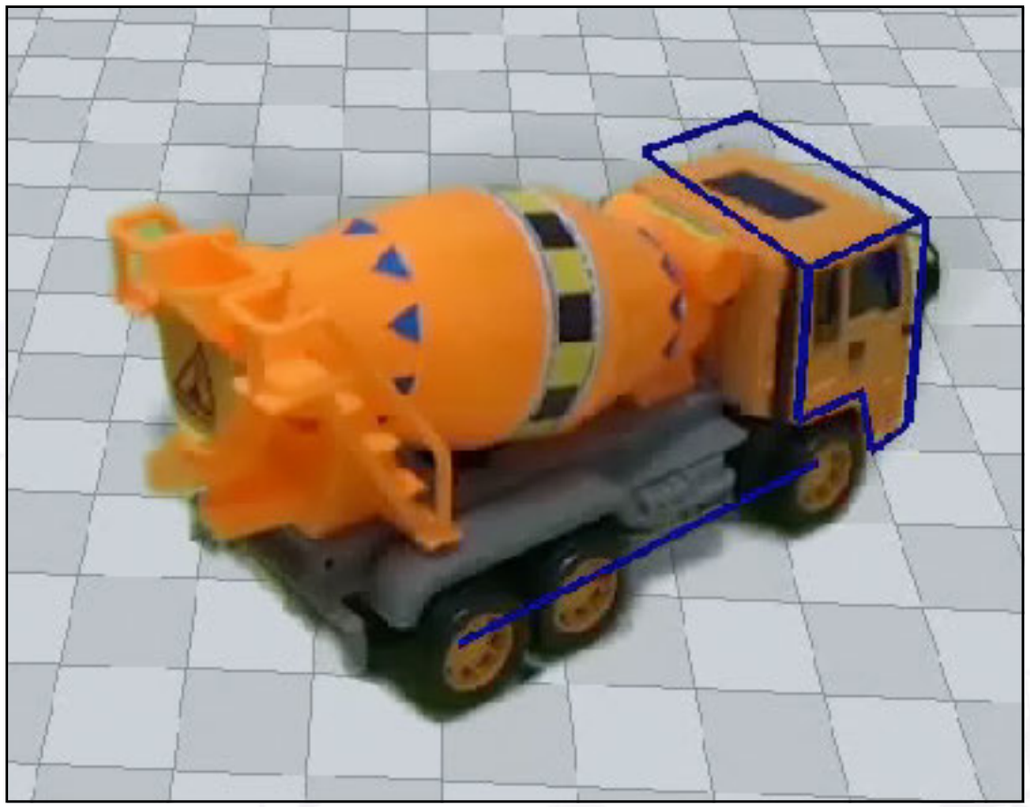
Result

Evaluation

Evaluation

- Evaluate image warp
    - Comparison with a baseline warping approach
    - Validate the effectiveness of edge and structure preservation terms
- Stress test in terms of changing novel camera view
- Comparison with the 3D reconstruction methods

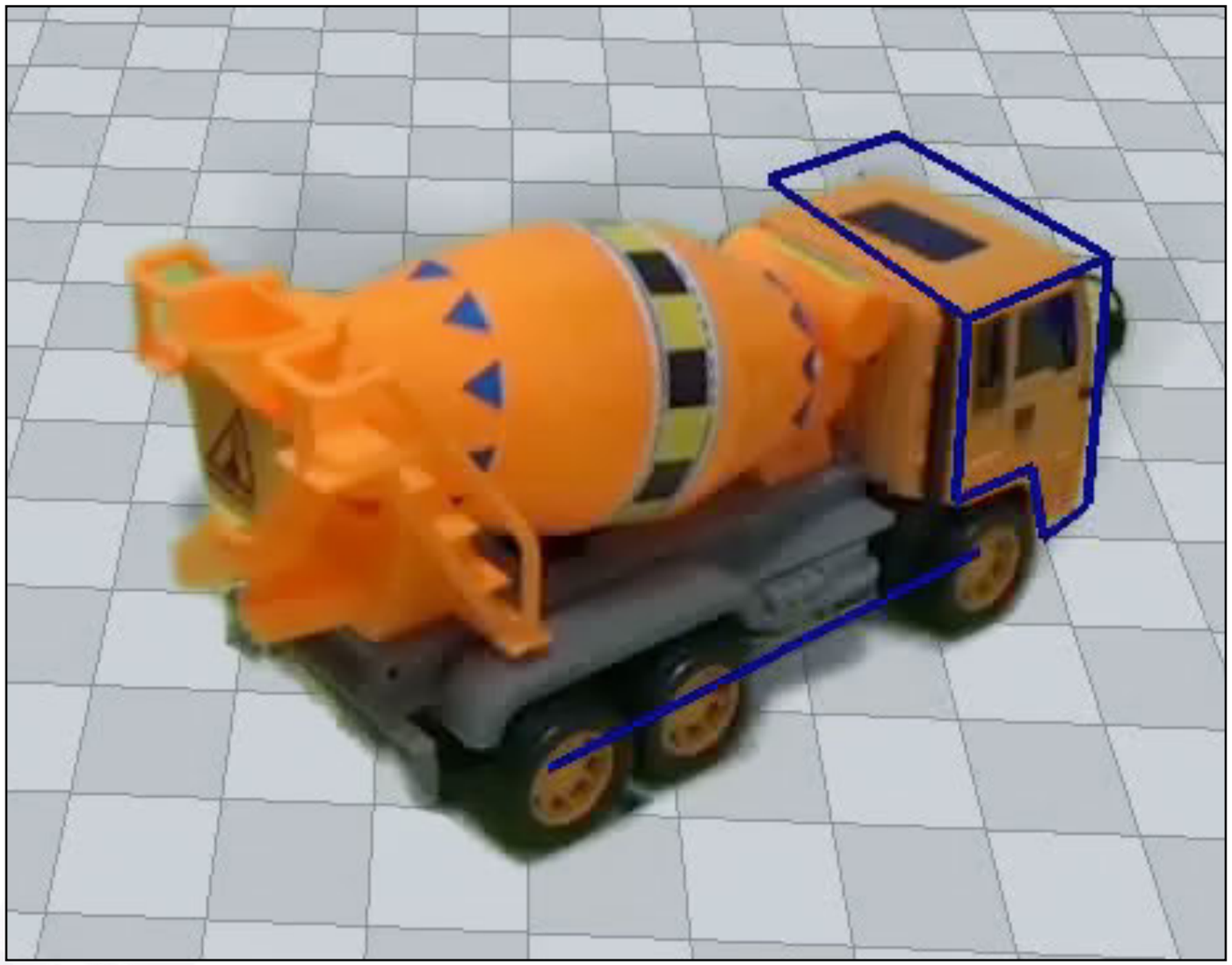
Comparison With Global Warp



Globally warp the nearest object frame

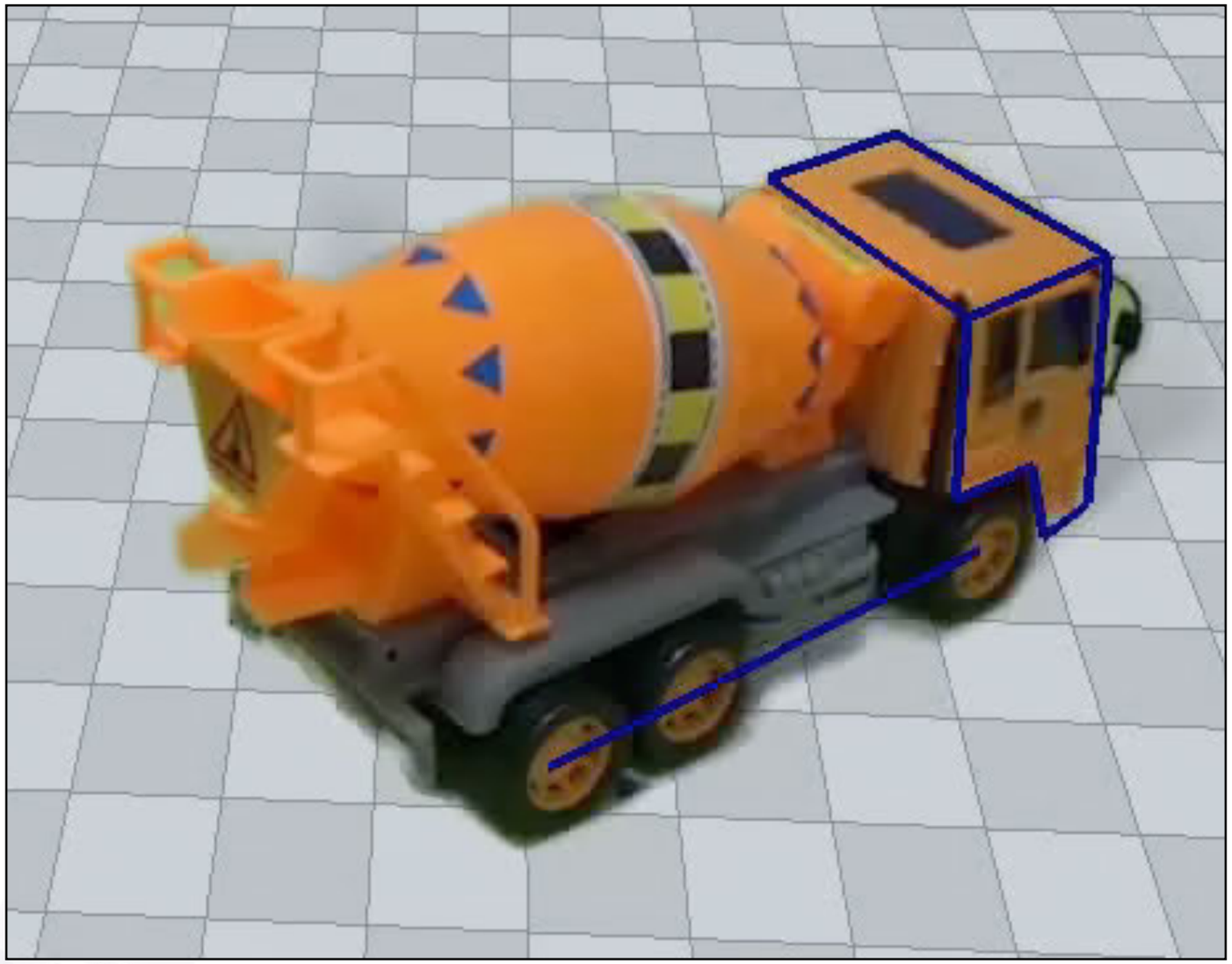
Our result

Comparison With Global Warp



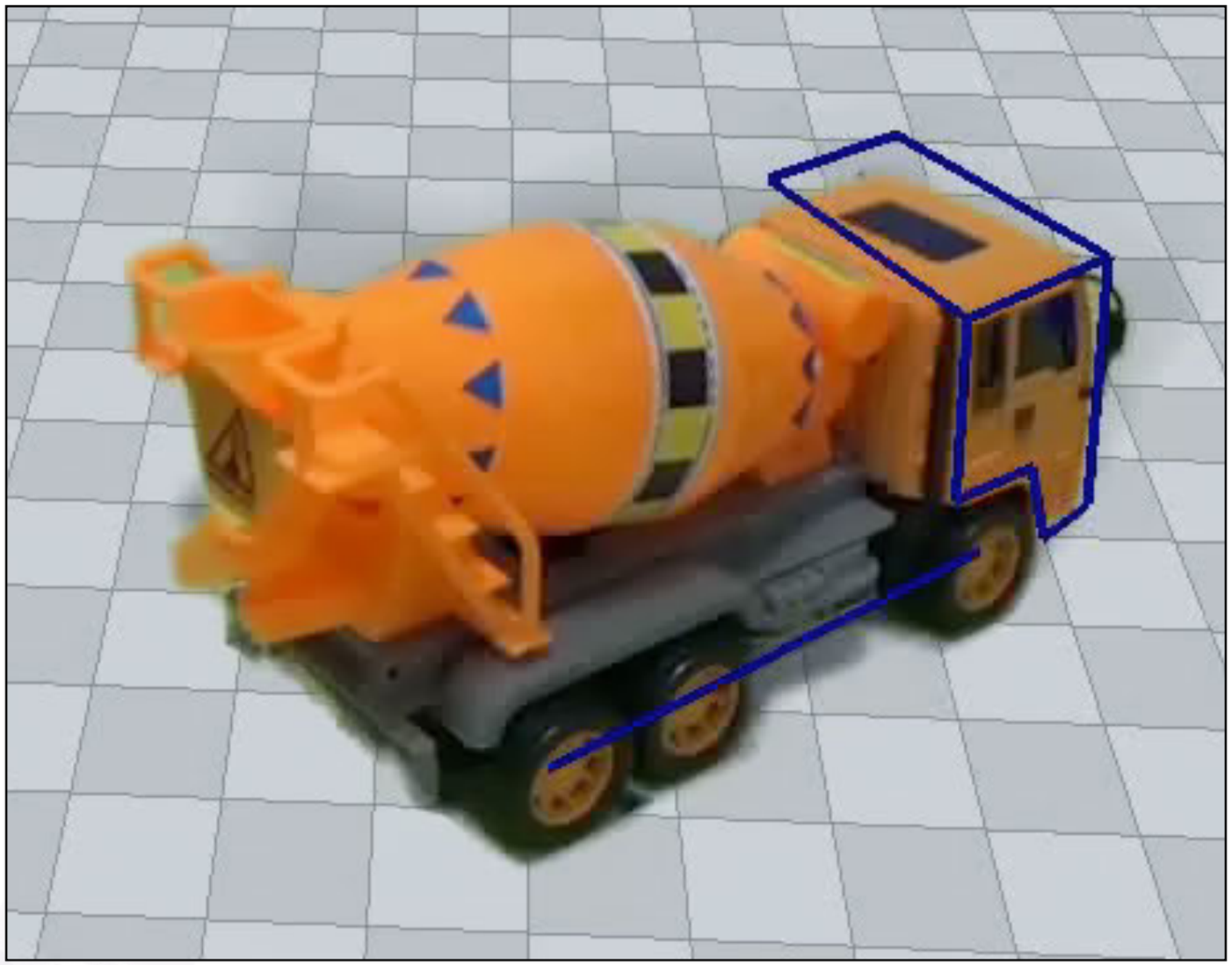
Globally warp the nearest object frame

Comparison With Global Warp



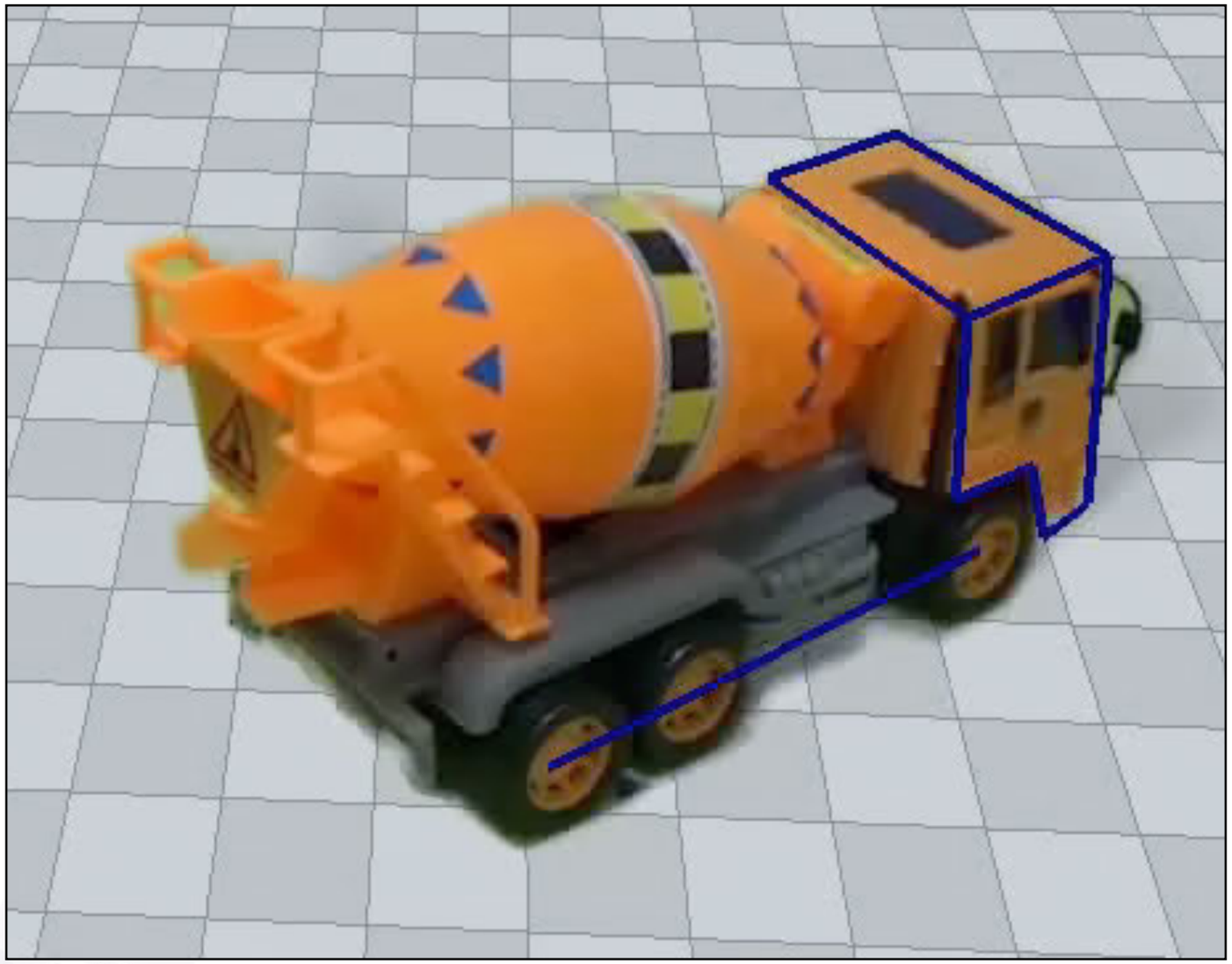
Our result

Comparison With Global Warp



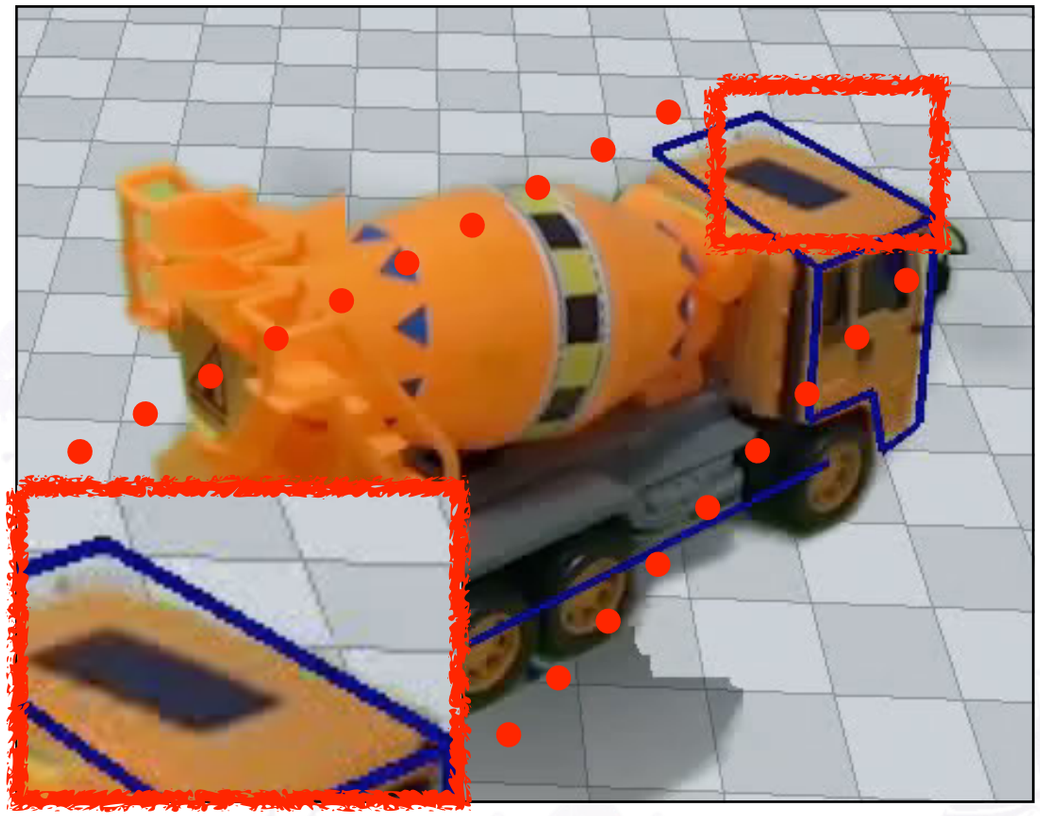
Globally warp the nearest object frame

Comparison With Global Warp

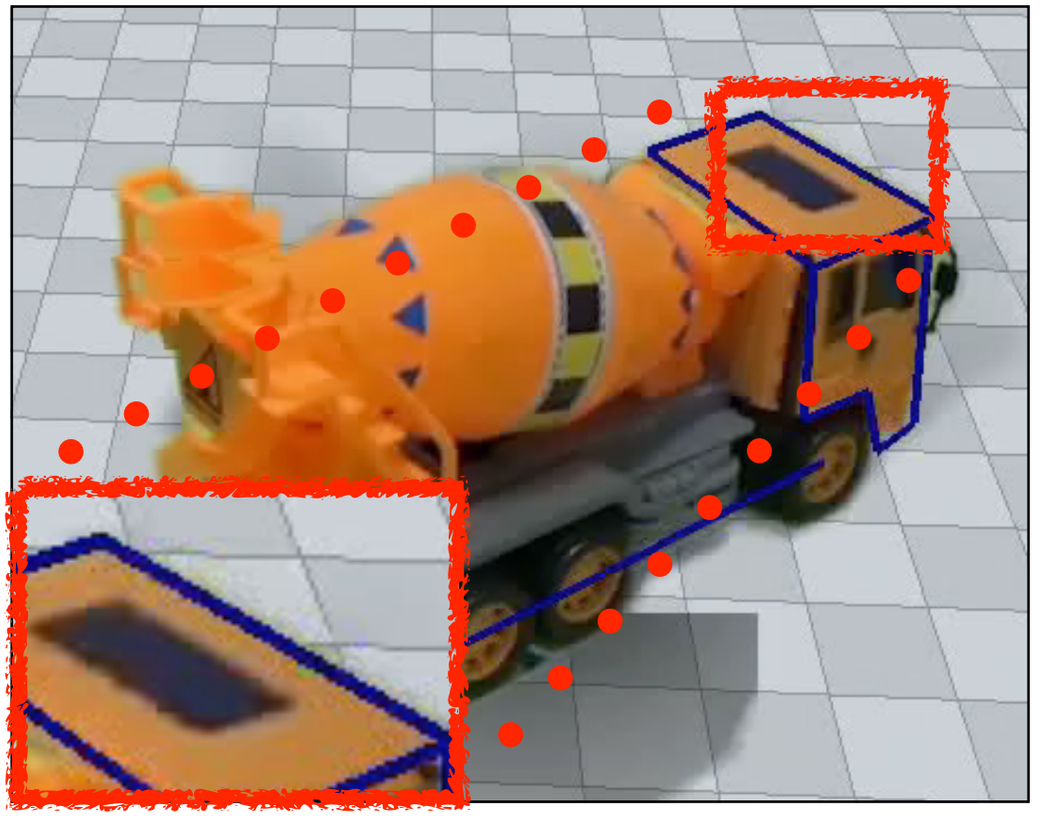


Our result

Effectiveness of Edge Preservation Term

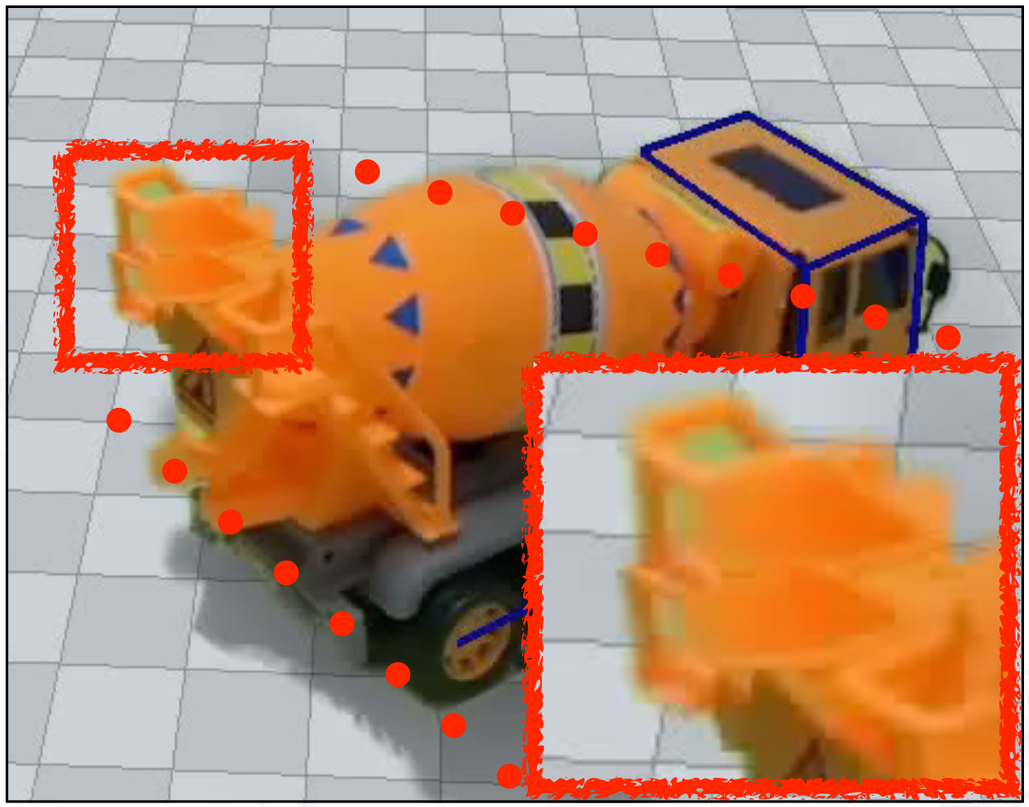


Without edge preservation term

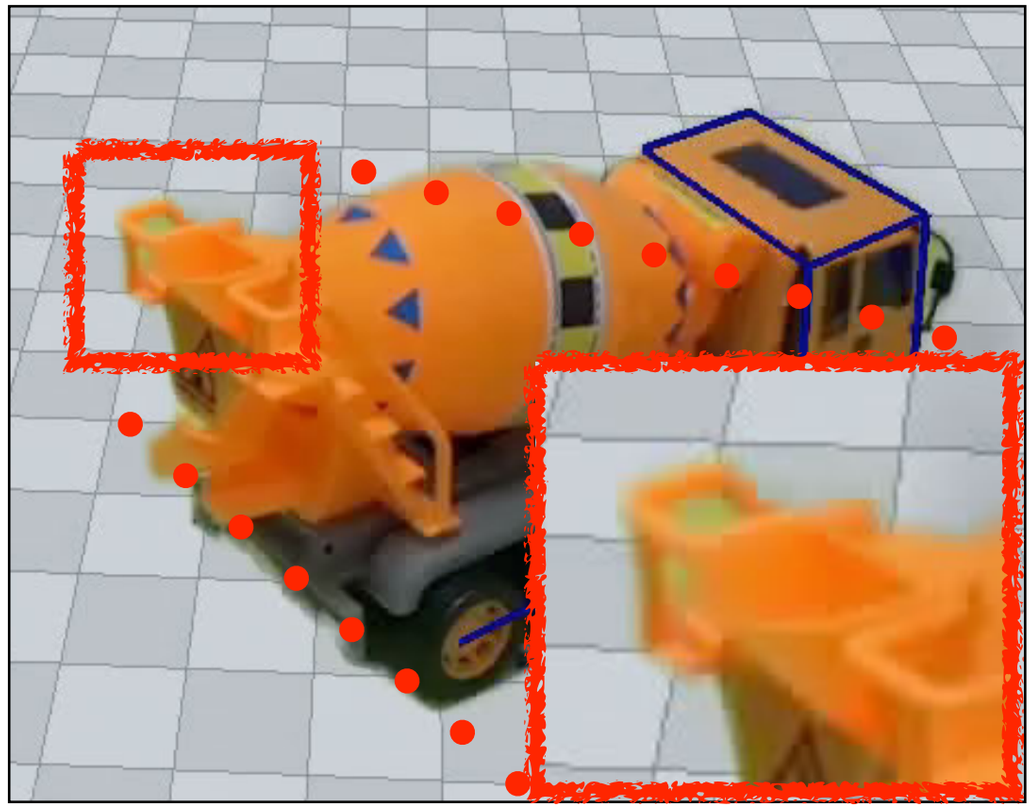


With edge preservation term

Effectiveness of Structure Preservation Term

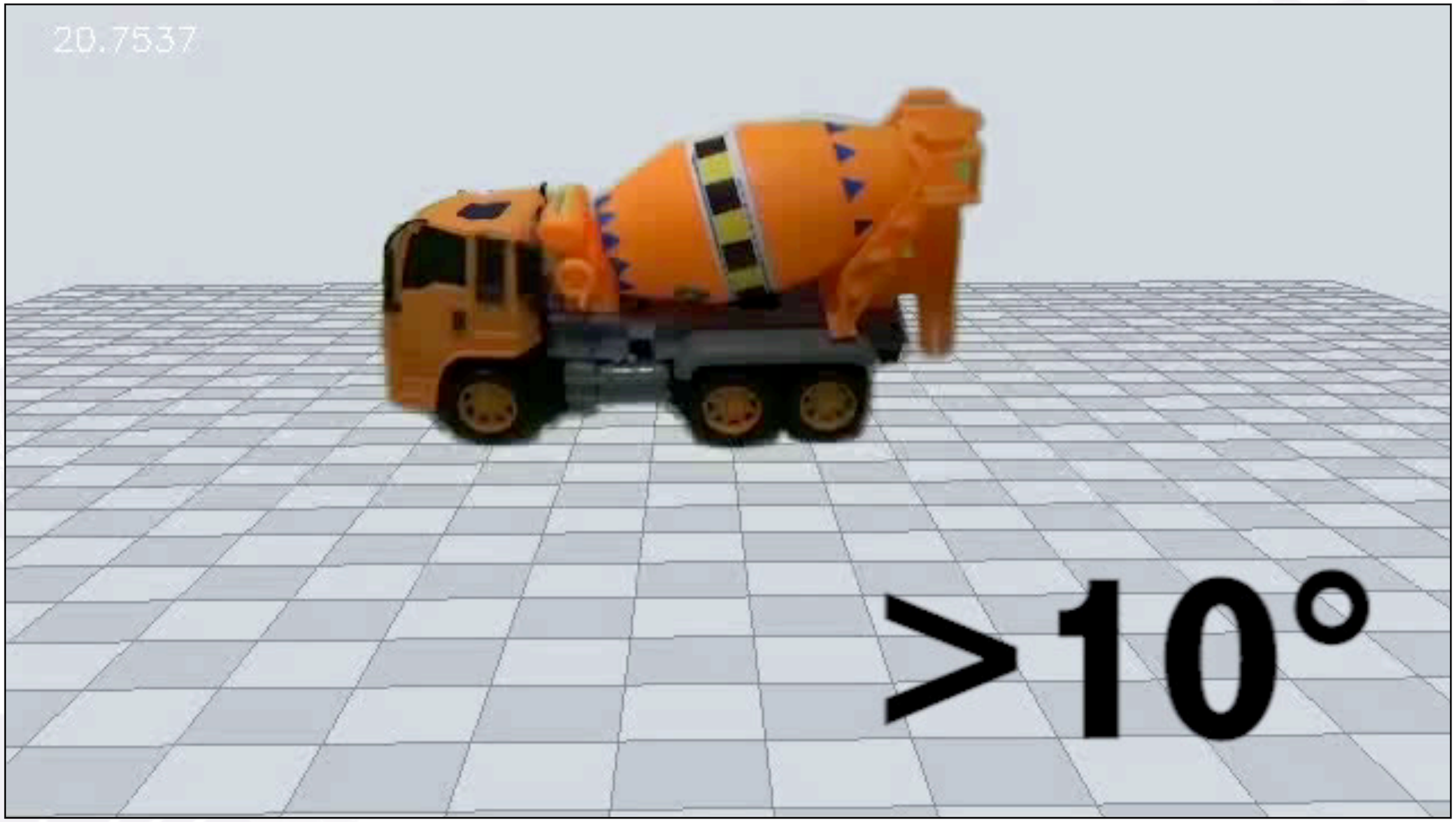


Without structure preservation term



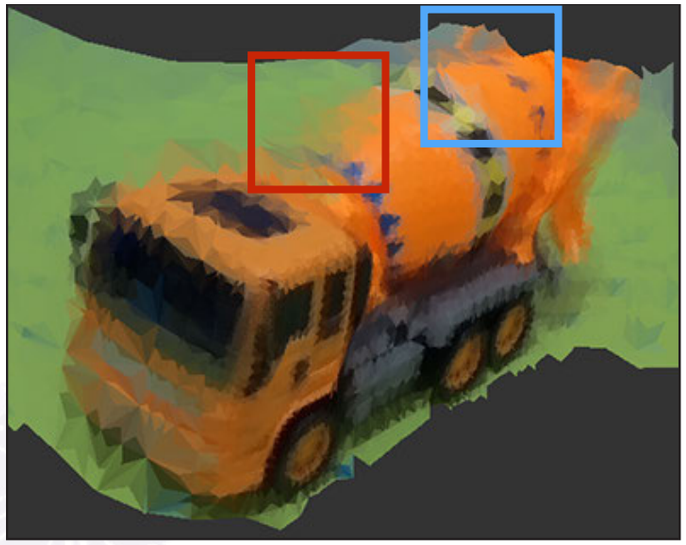
With structure preservation term

Stress Test

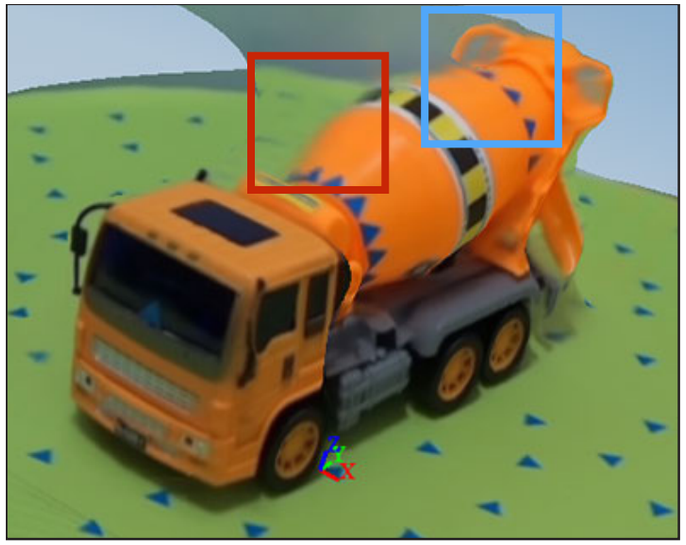


Comparison with 3D Reconstruction Methods

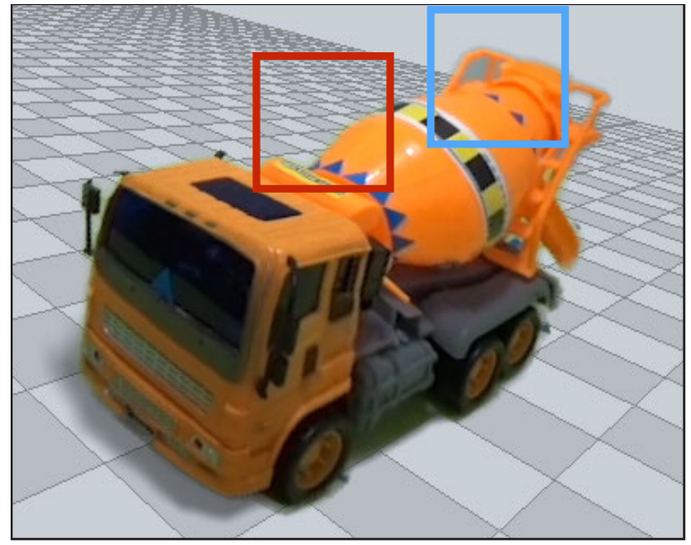
Vi3Dim

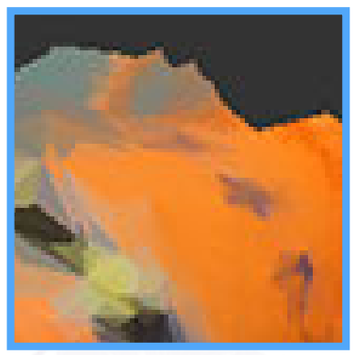


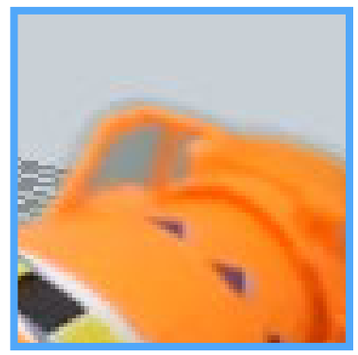
123D Catch



Our Result

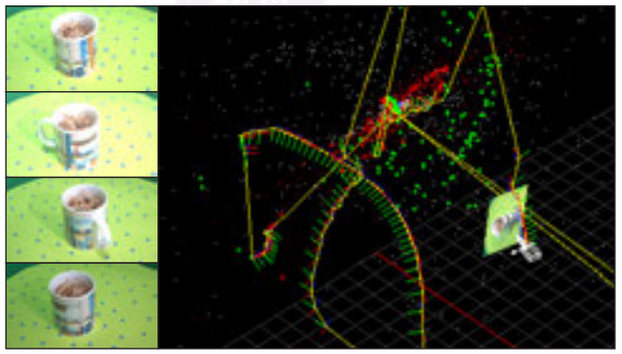


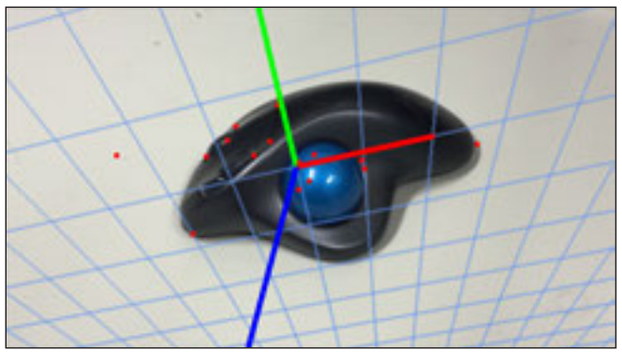


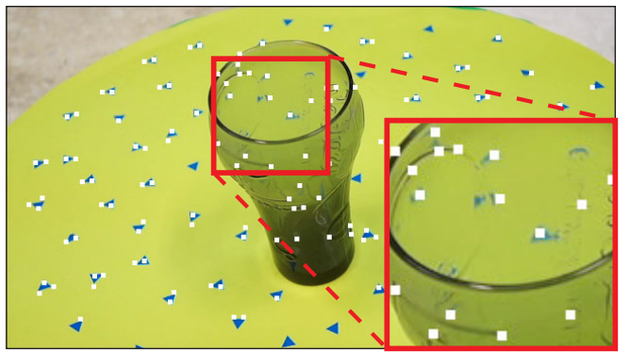


Limitations

- The limitations of SfM algorithm
  - Severe temporal lighting changes
  - Textureless scenes
  - Transparent foreground objects







Conclusion

- A video editing system that enables object level edits of videos without explicitly reconstructing the 3D geometries of the scenes

- A novel image-based rendering algorithm
  - The use of the sparse structure points as manipulation proxy
  - Object frame retrieval
  - Structure preserving image warping
  - Spatiotemporal stitching

Future Work

- Improve the quality of composition by incorporating sophisticated shadow creation, illumination adjustment, and appearance harmonization

- Adaptively devote processing power based on model saliency

- Make advanced version of SSP to support thin tubelike primitives and videos with dynamic foreground

Acknowledgements

- Thank the anonymous reviewers for their invaluable comments

- The project was supported in part by
  - Ministry of Science and Technology of Taiwan
  - ERC Starting Grant SmartGeometry (StG-2013-335373)

Supplementary Materials

- Project webpage

- Videos/codes (non-commercial usage)

- More results and paper videos

Thank You !

Q & A