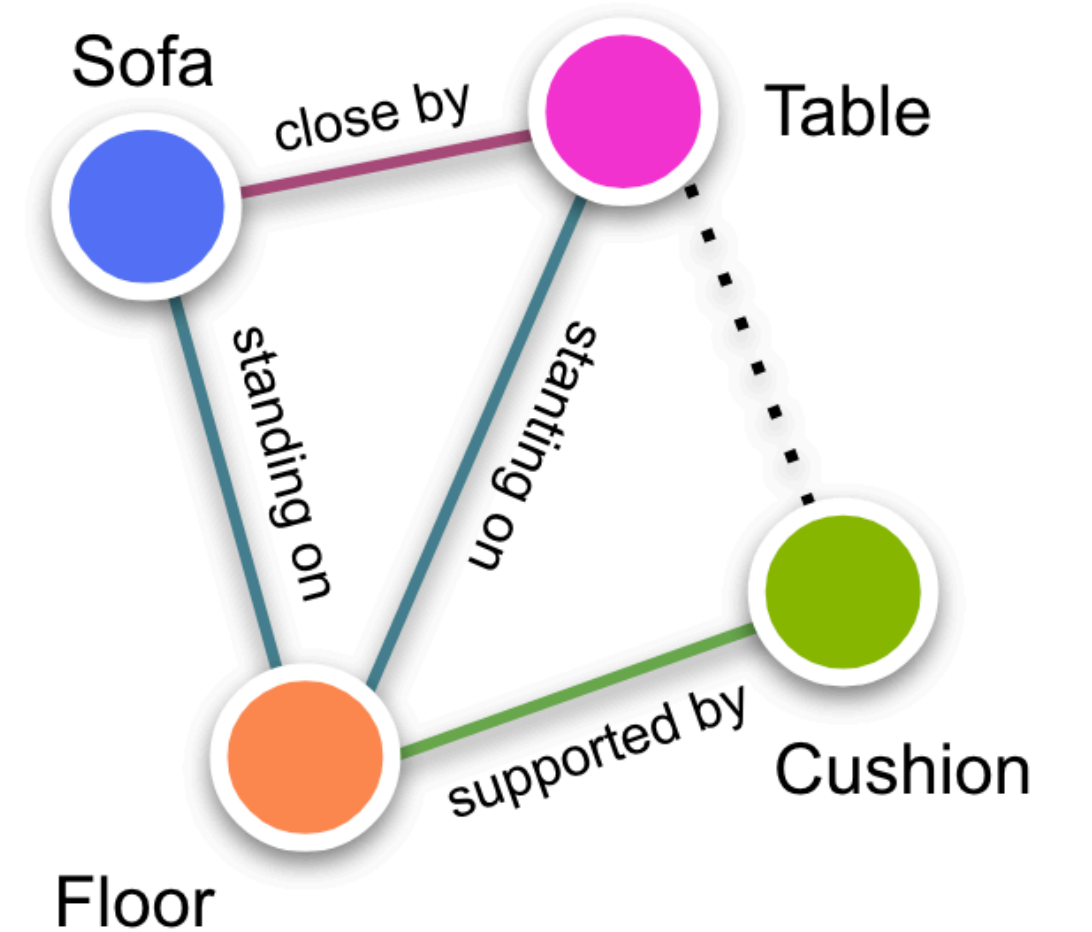
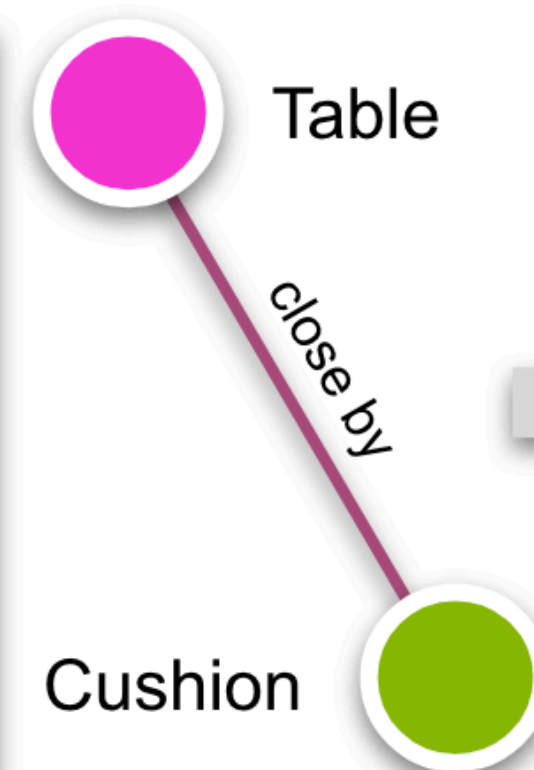
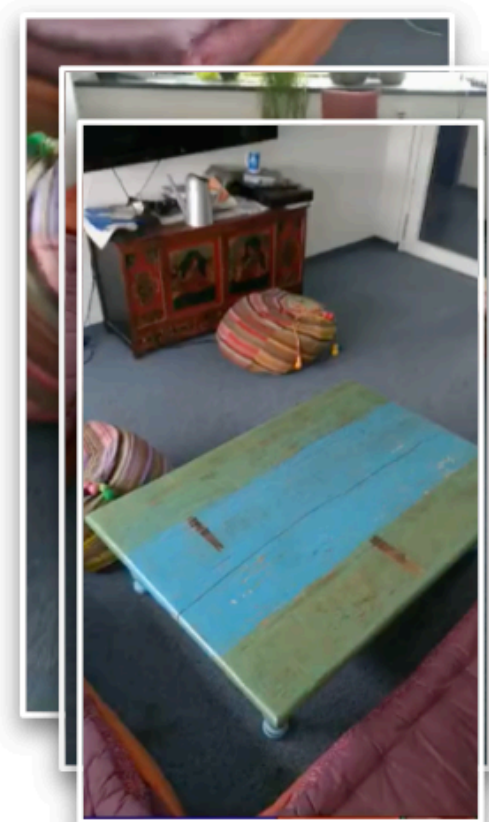


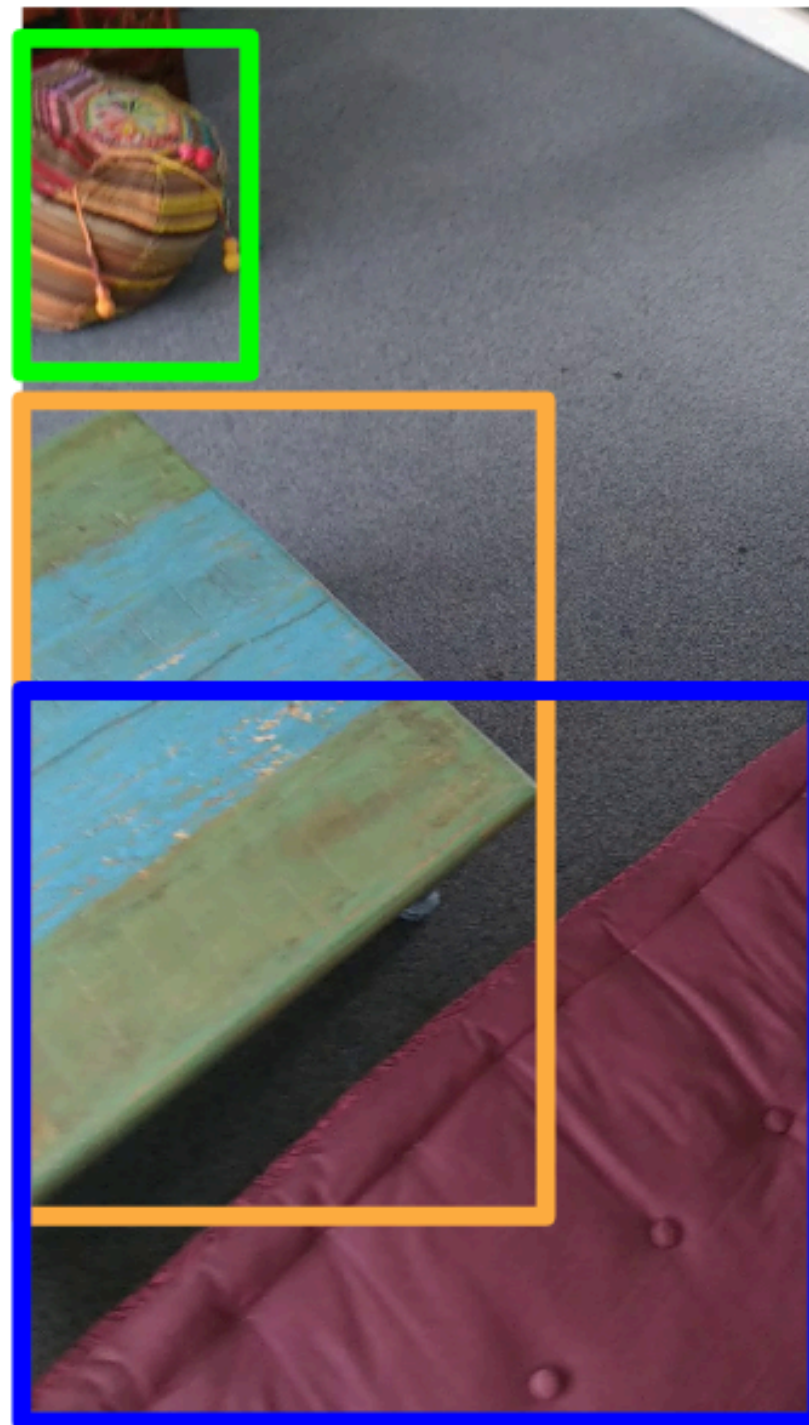
Incremental 3D Semantic Scene Graph Prediction from RGB Sequences

Shun-Cheng Wu¹ Keisuke Tateno² Nassir Navab¹ Federico Tombari^{1,2}

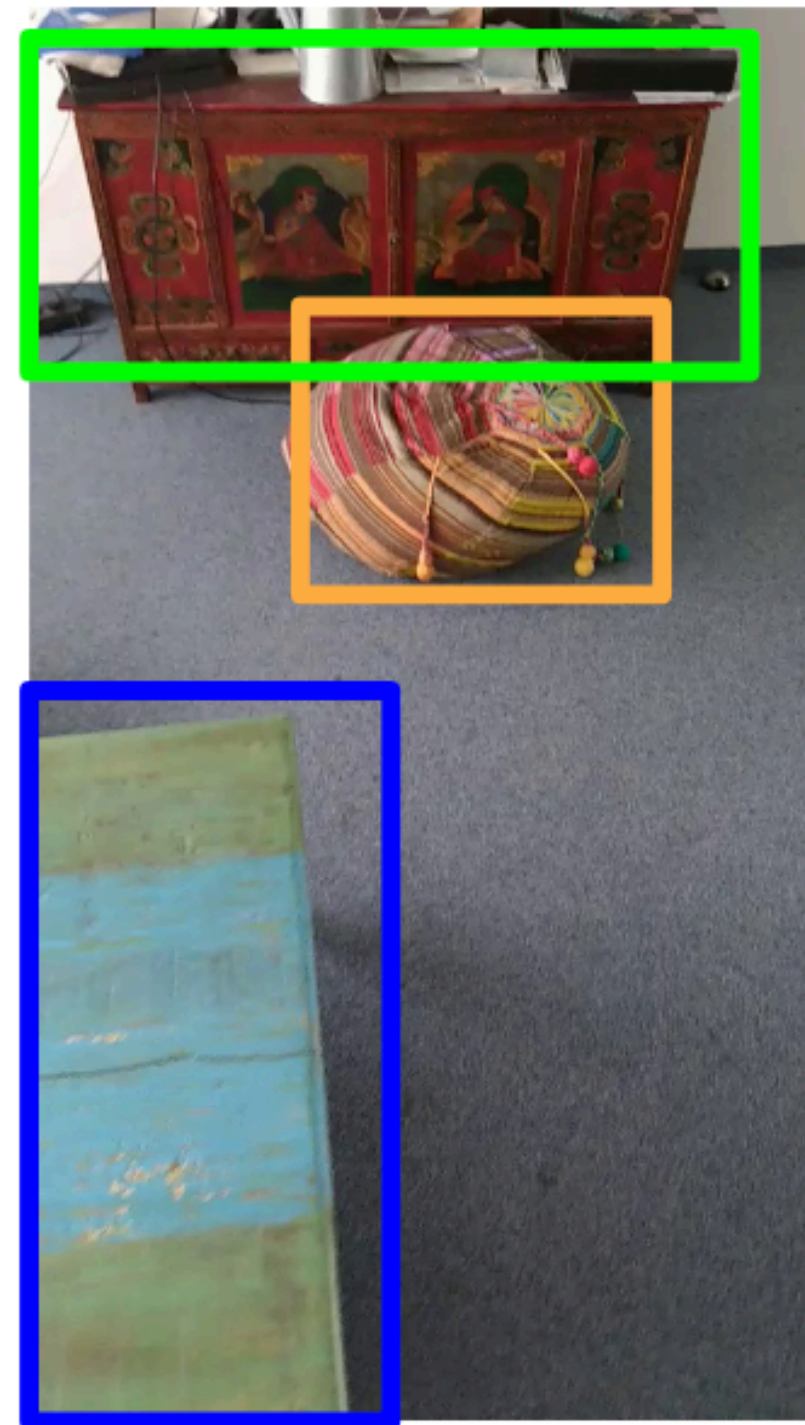
¹Technical University of Munich, ²Google



View-based Semantic Scene Graphs Generation (SSG)



frame-000006



frame-000018

Relationship estimation is limited by the field of view

Lack of temporal consistency

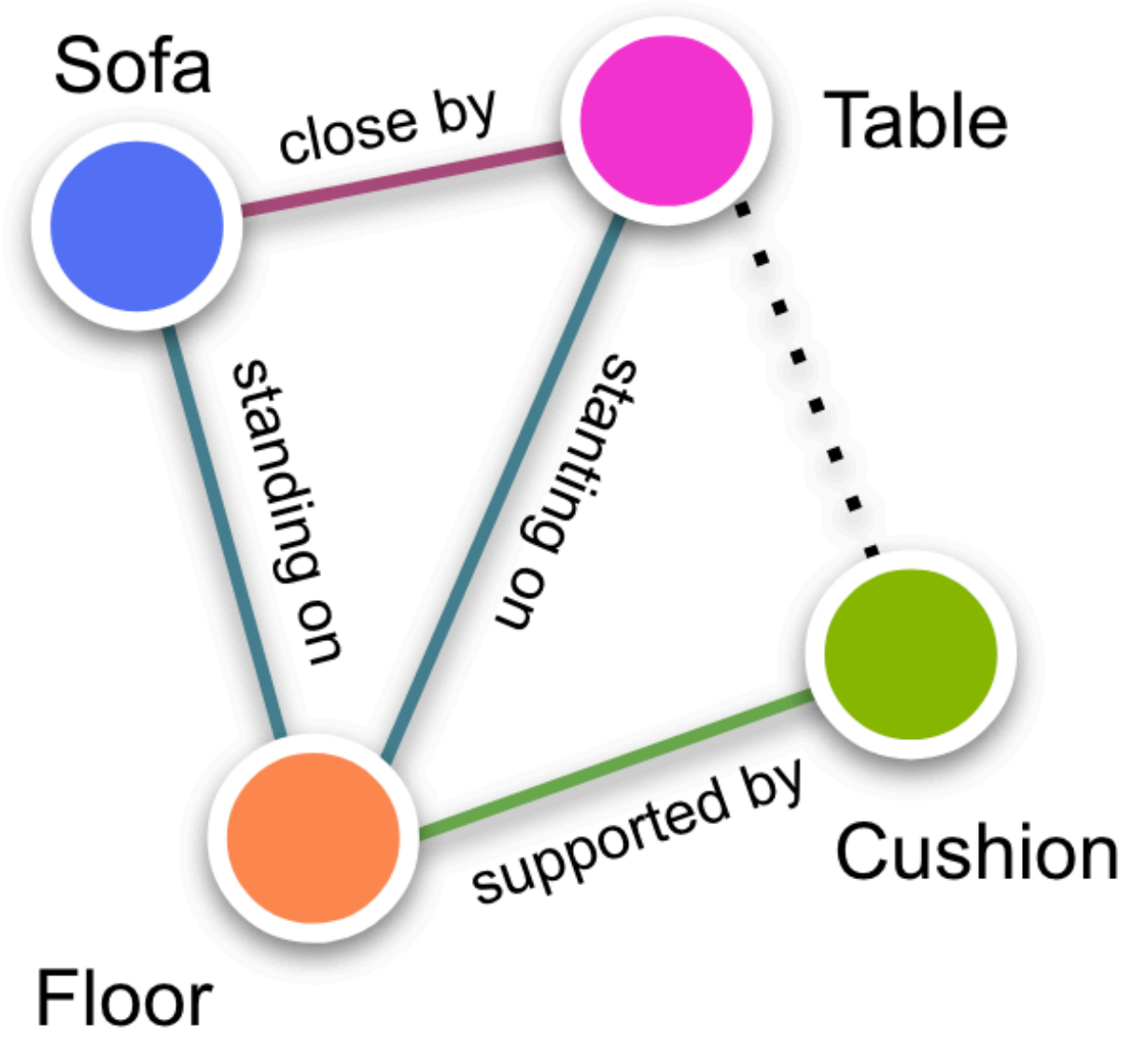
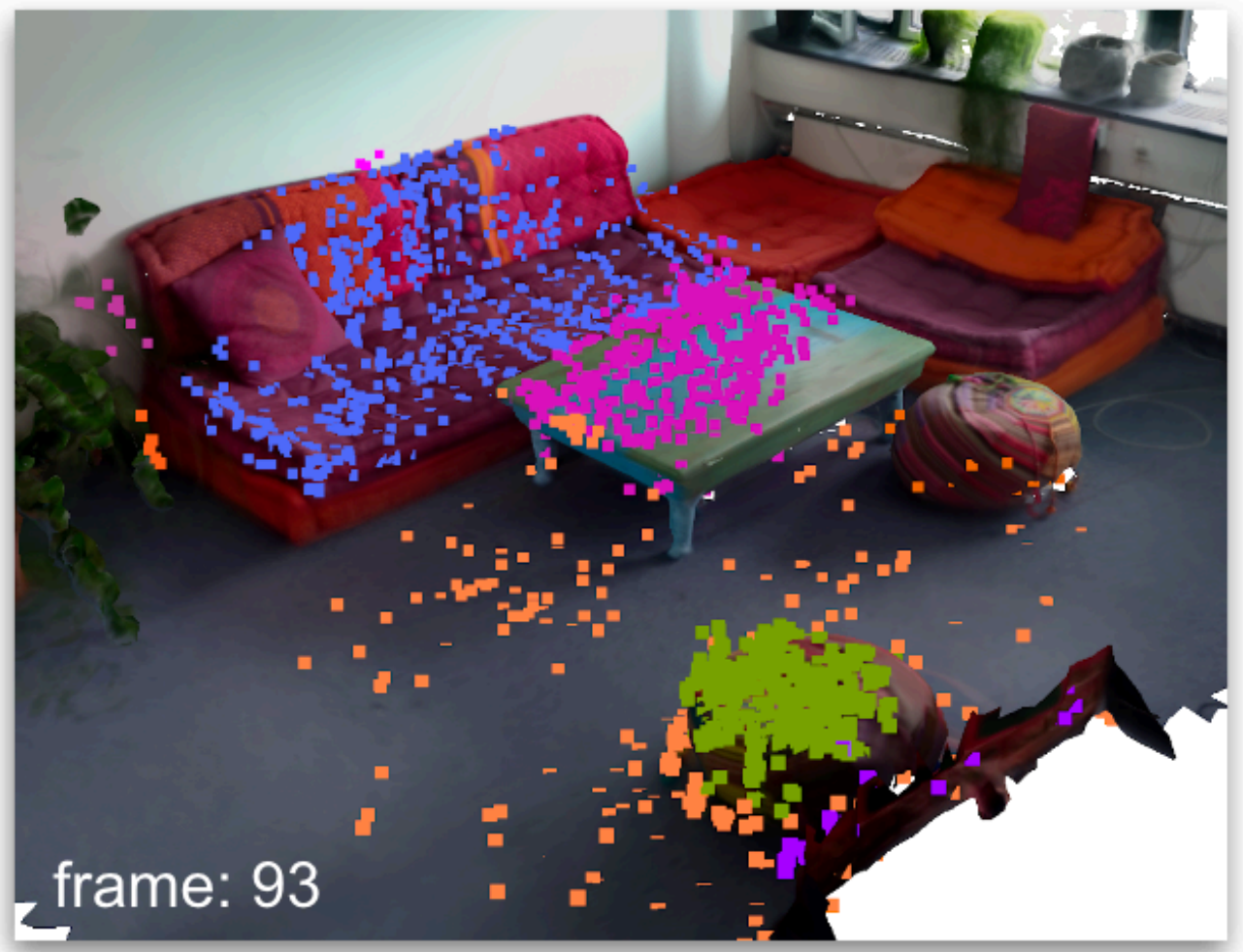
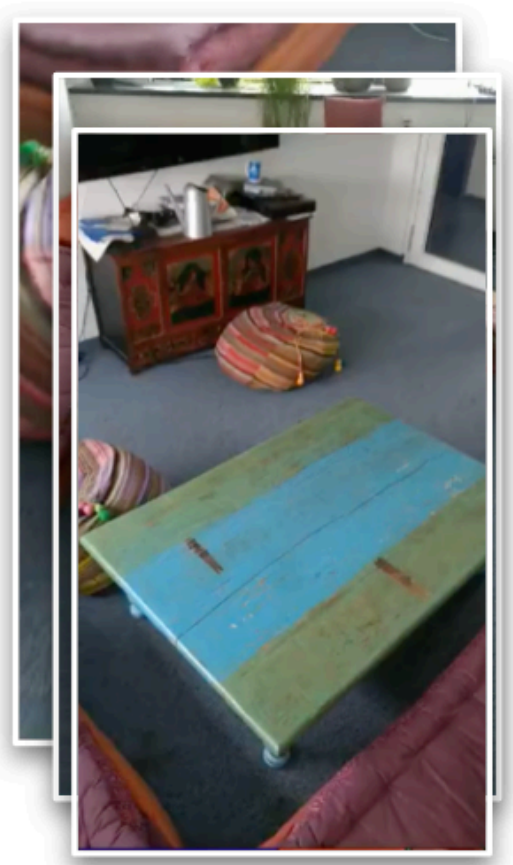
3D-based Semantic SSG



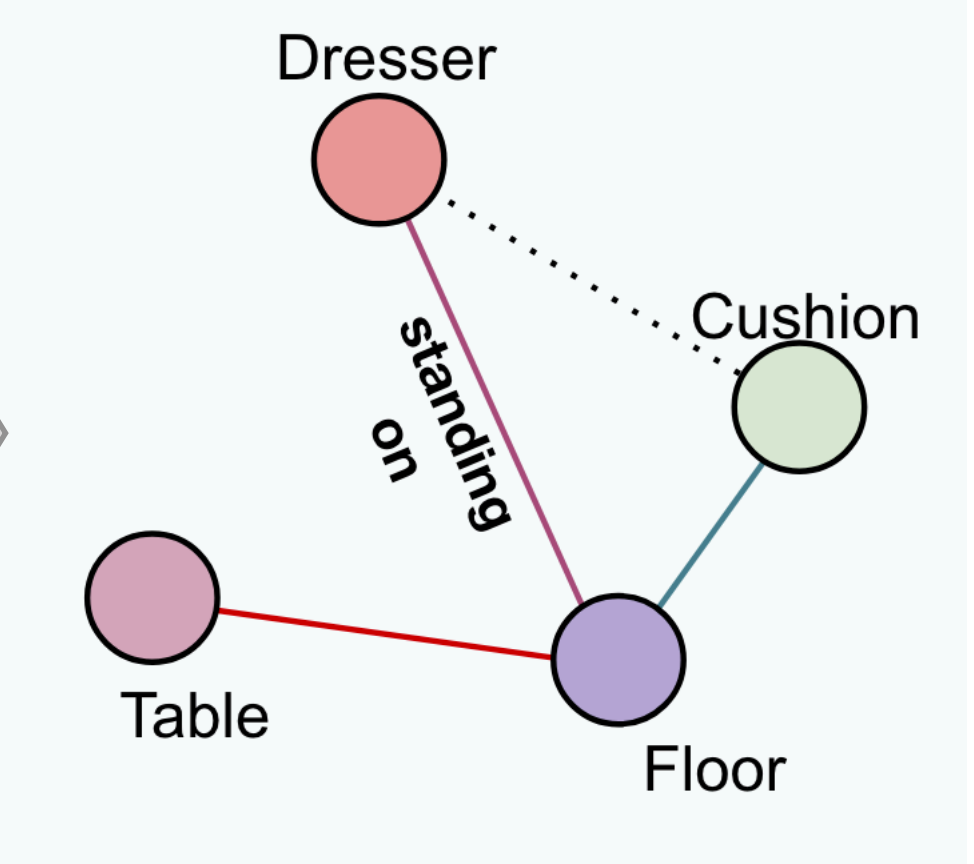
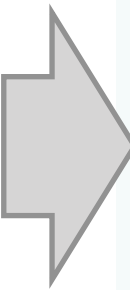
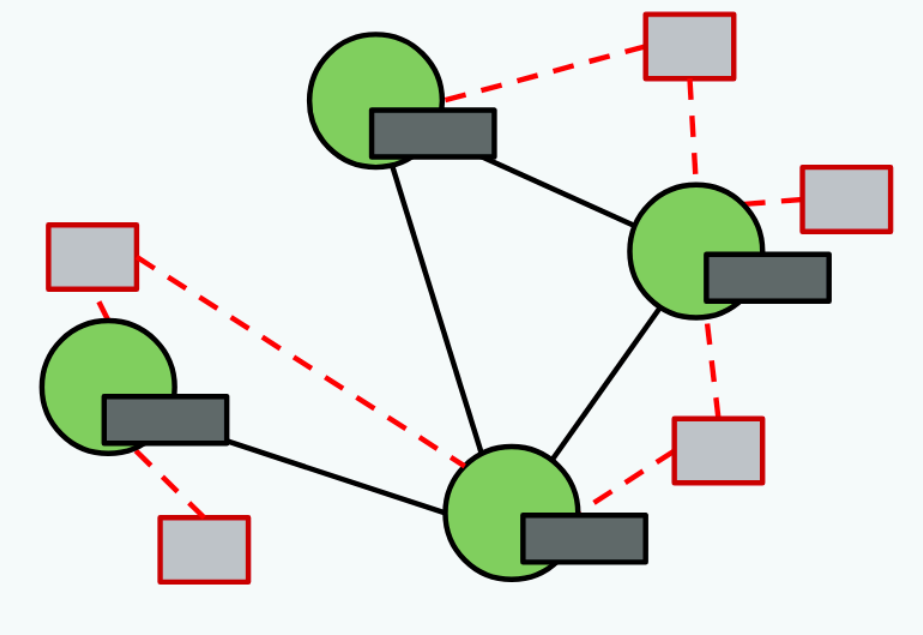
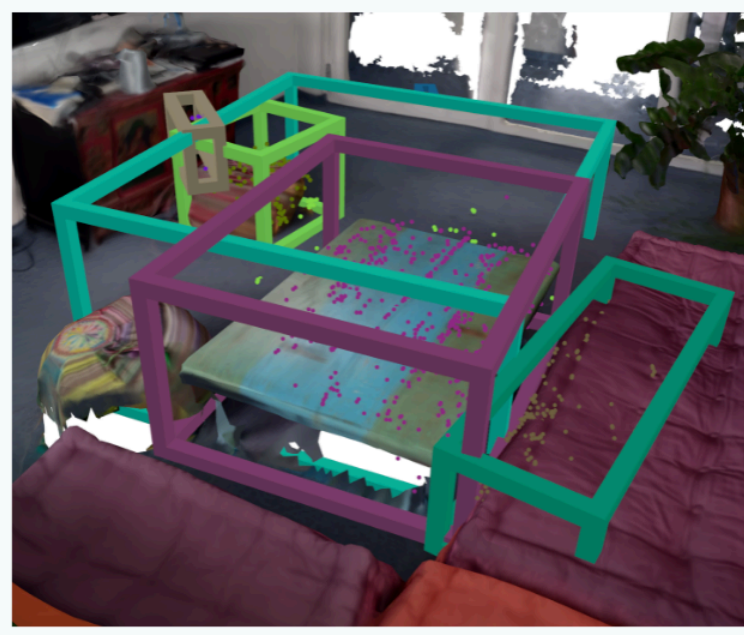
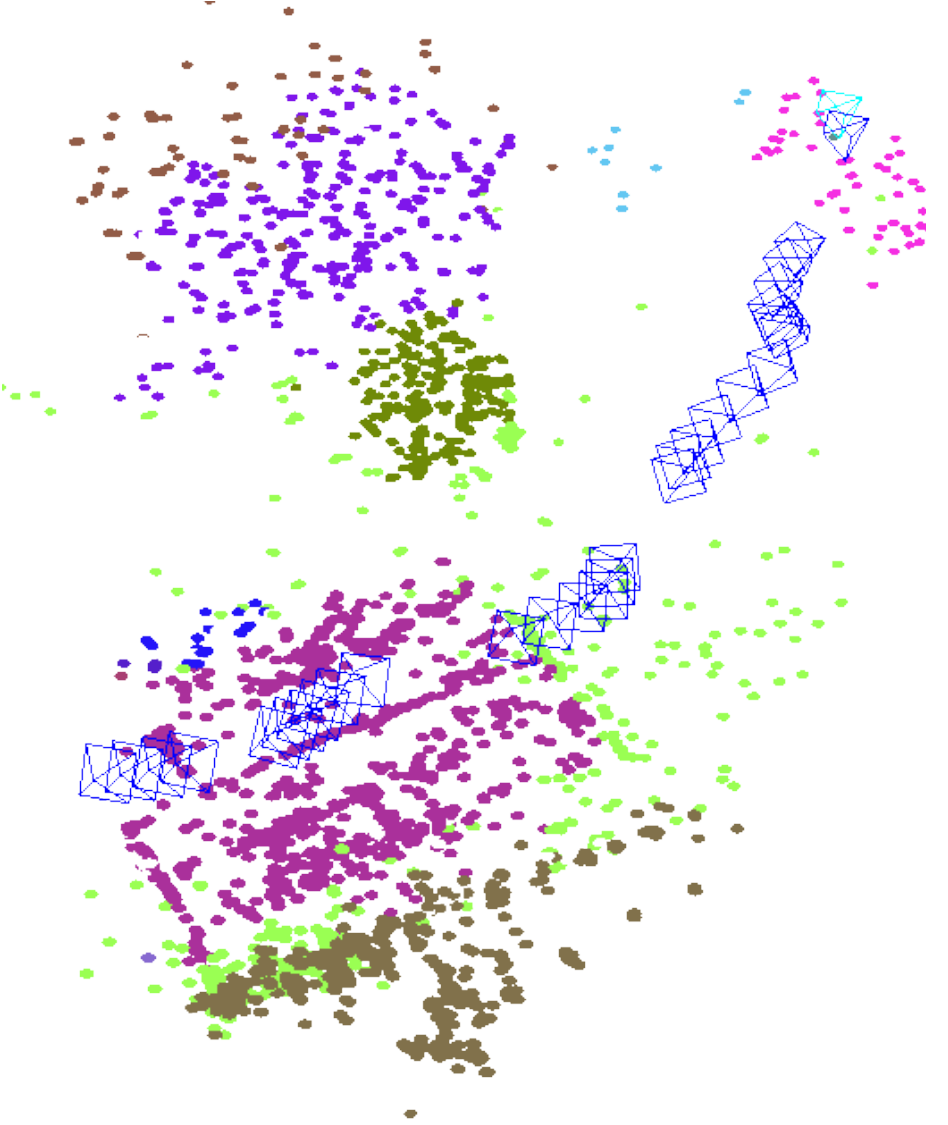
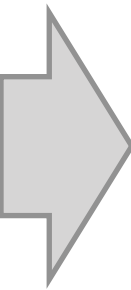
Dense Geometry is not always available

Feature learning on unrealistic texture

Our method



Our method



● : Entity □ : Keyframe ▭ : Properties
— : Neighbour graph edge - - - : Visibility graph edge

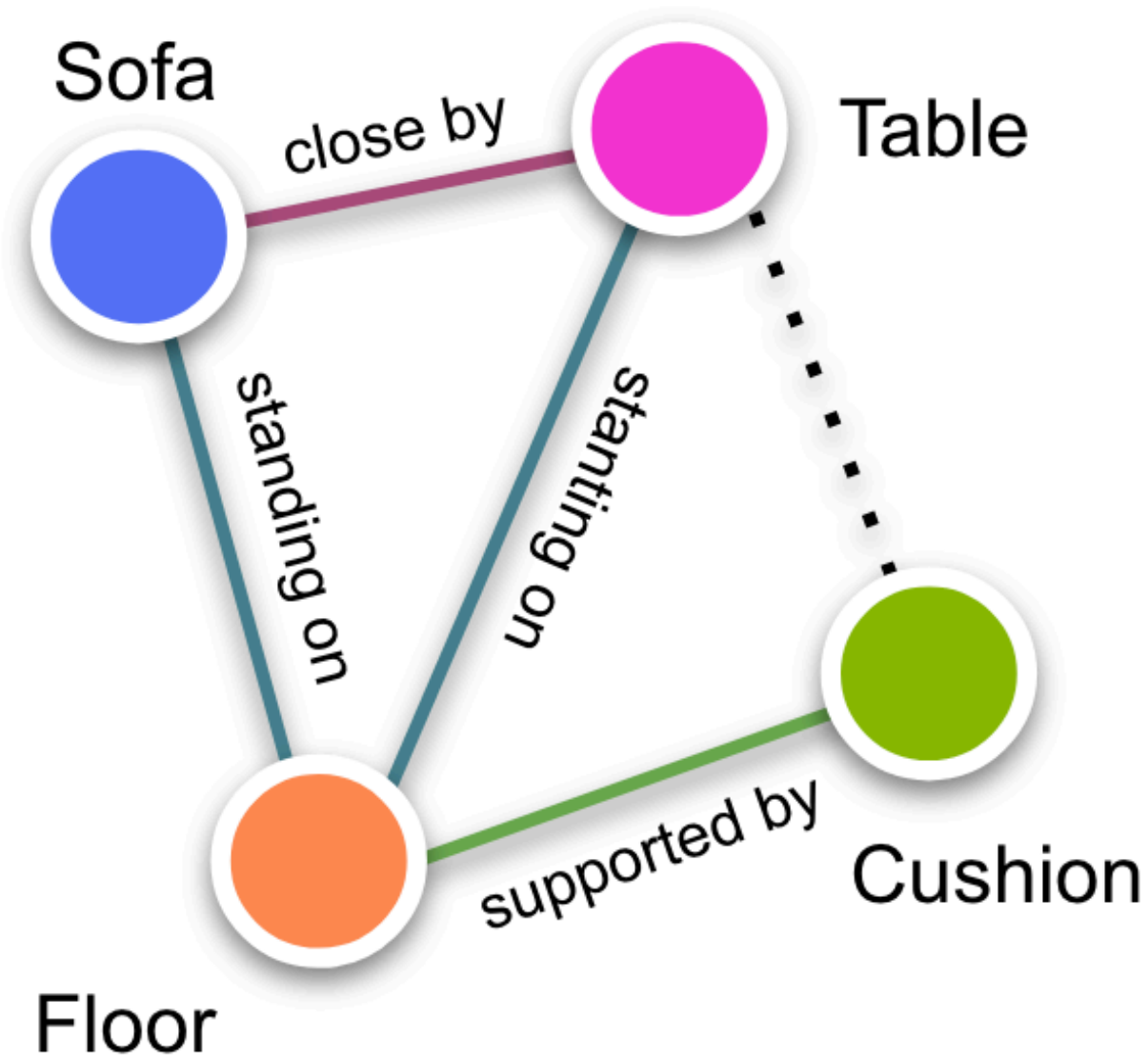
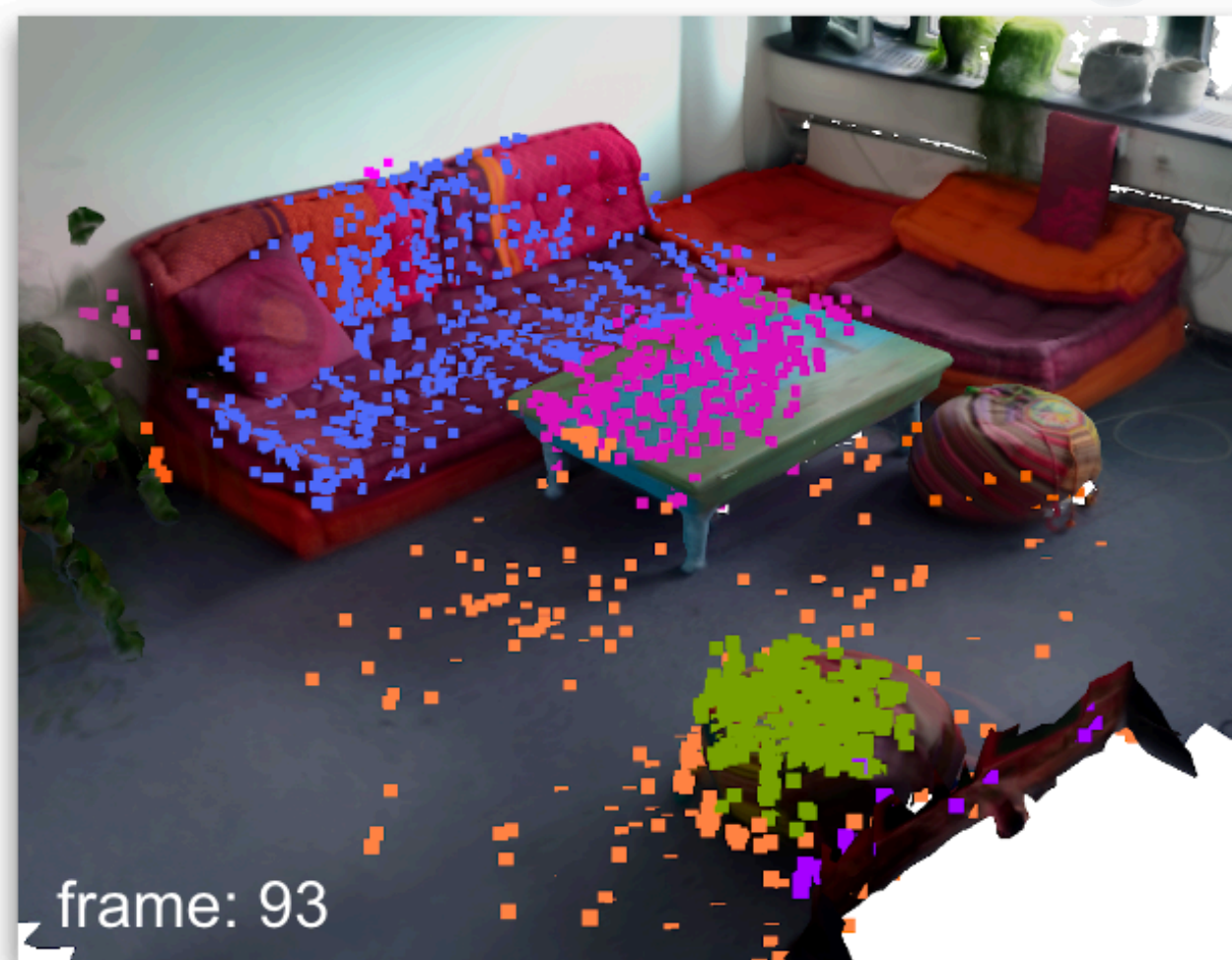
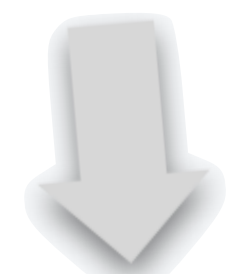
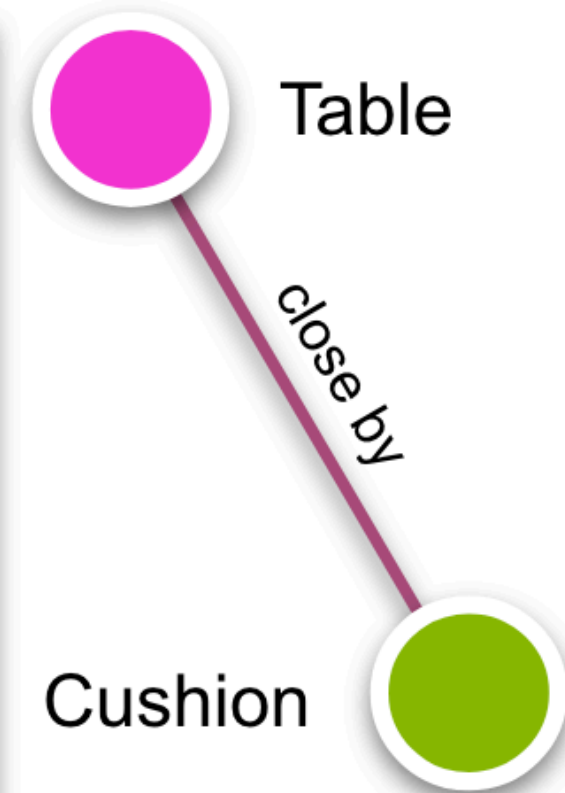
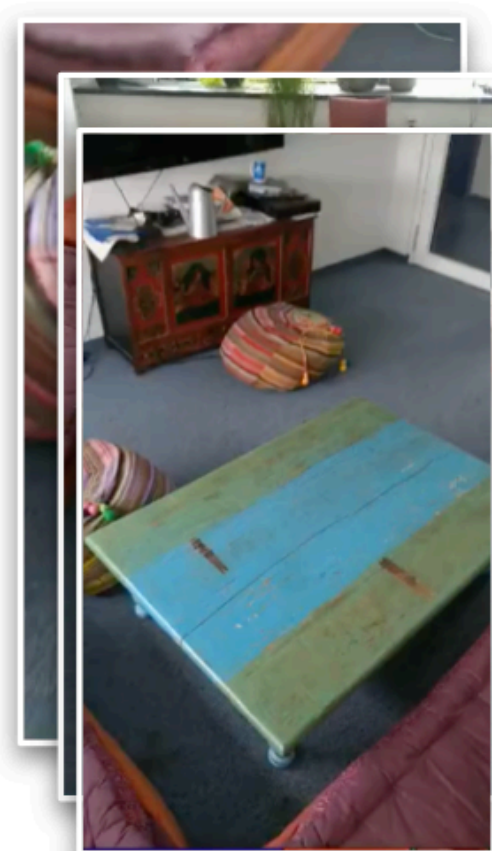
Different input types

Method	Recall(%)			mRecall(%)		
	Rel.	Obj.	Pred.	Obj.	Pred.	
GT	IMP [66]	49.8	70.1	94.3	53.0	38.1
	VGfM [15]	49.3	69.4	94.8	57.5	44.6
	3DSSG [61]	34.6	58.0	95.2	46.8	58.7
	SGFN [64]	41.8	63.8	94.3	57.7	65.5
	Ours	66.1	81.2	95.6	77.4	71.5
Dense	IMP [66]	25.8	51.8	90.4	30.0	23.0
	VGfM [15]	28.3	53.3	90.7	31.6	24.4
	3DSSG [61]	17.5	41.4	88.2	31.9	26.6
	SGFN [64]	31.4	56.7	89.6	38.3	30.5
	Ours	34.1	58.1	89.9	43.0	33.3
Sparse	IMP [66]	7.9	27.5	90.7	20.6	14.0
	VGfM [15]	8.2	26.9	90.8	17.6	15.4
	3DSSG [61]	0.9	9.7	87.9	5.9	15.1
	SGFN [64]	1.7	12.6	88.9	8.3	14.4
	Ours	9.9	29.5	90.4	23.5	16.5
	Ours (i)	10.7	30.2	90.4	24.5	15.9

With fewer labels & predicates

Method	Recall(%)			mRecall(%)	
	Rel.	Obj.	Pred.	Obj.	Pred.
IMP [66]	44.5	35.9	9.0	18.7	4.9
VGfM [15]	44.5	37.9	14.7	17.9	6.5
3DSSG [61]	46.8	29.6	68.8	11.7	25.5
SGFN [64]	45.2	29.4	42.8	11.8	13.5
Ours	52.7	56.7	50.4	27.2	23.9

With fine-grained labels



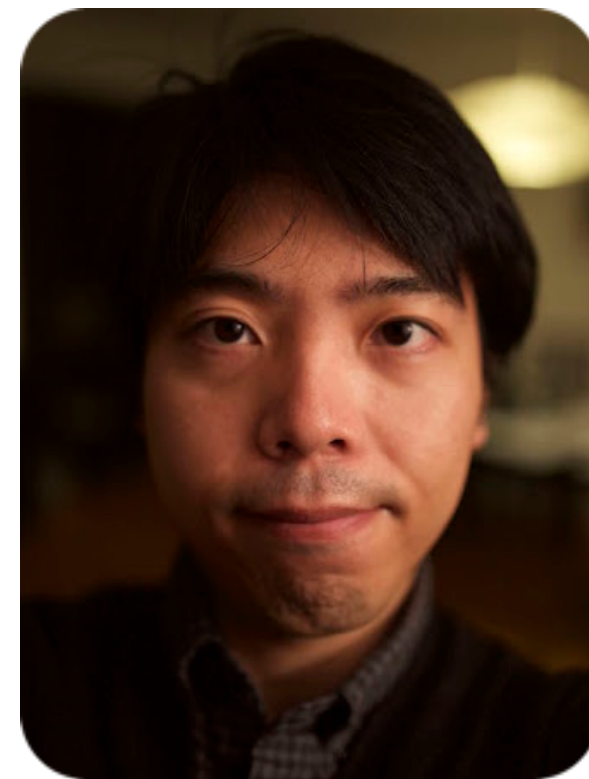
TUE-PM-088

ID: 5050

Thanks for watching.
Check out our poster and paper!



Shun-Cheng Wu



Keisuke Tateno

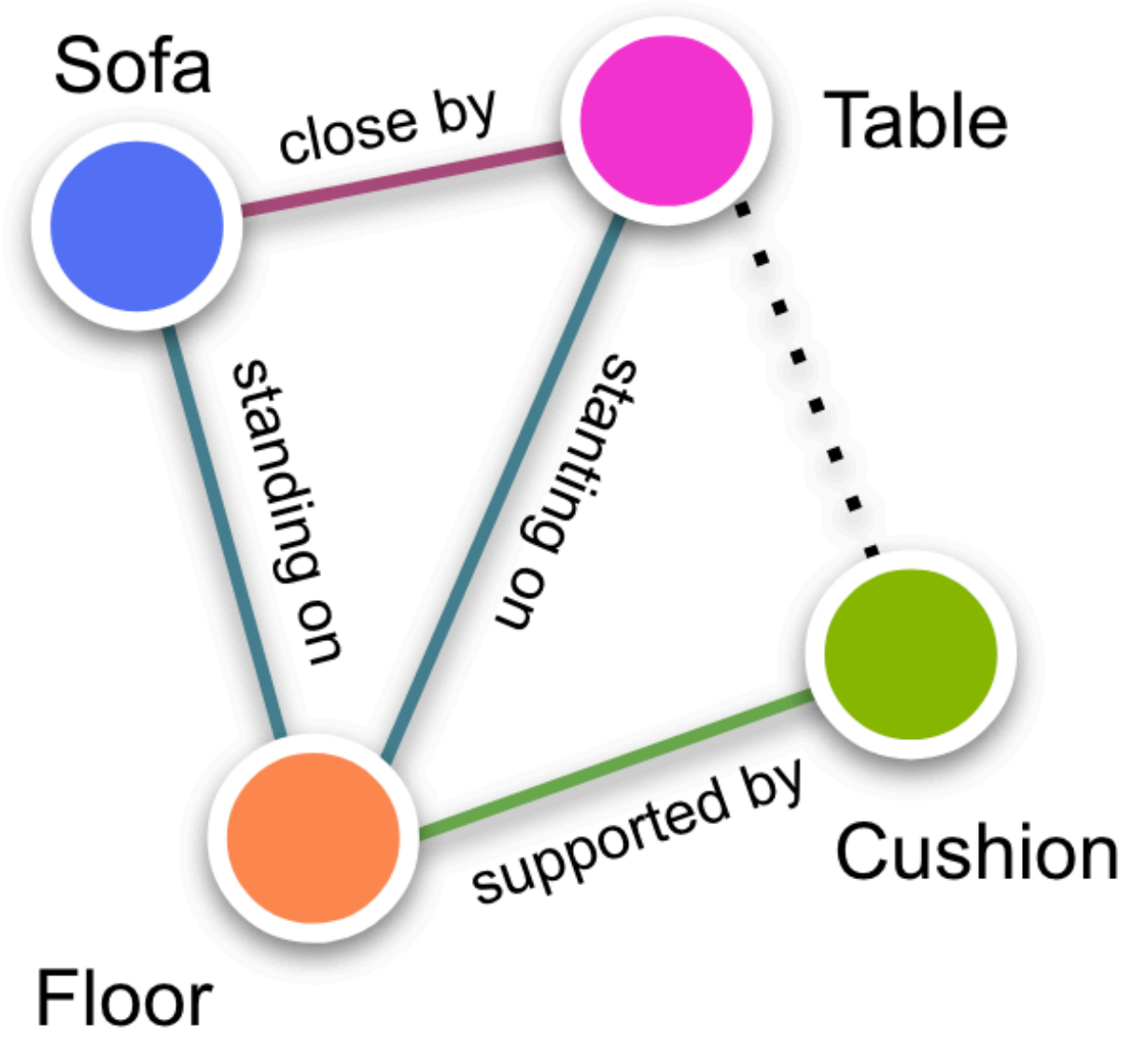
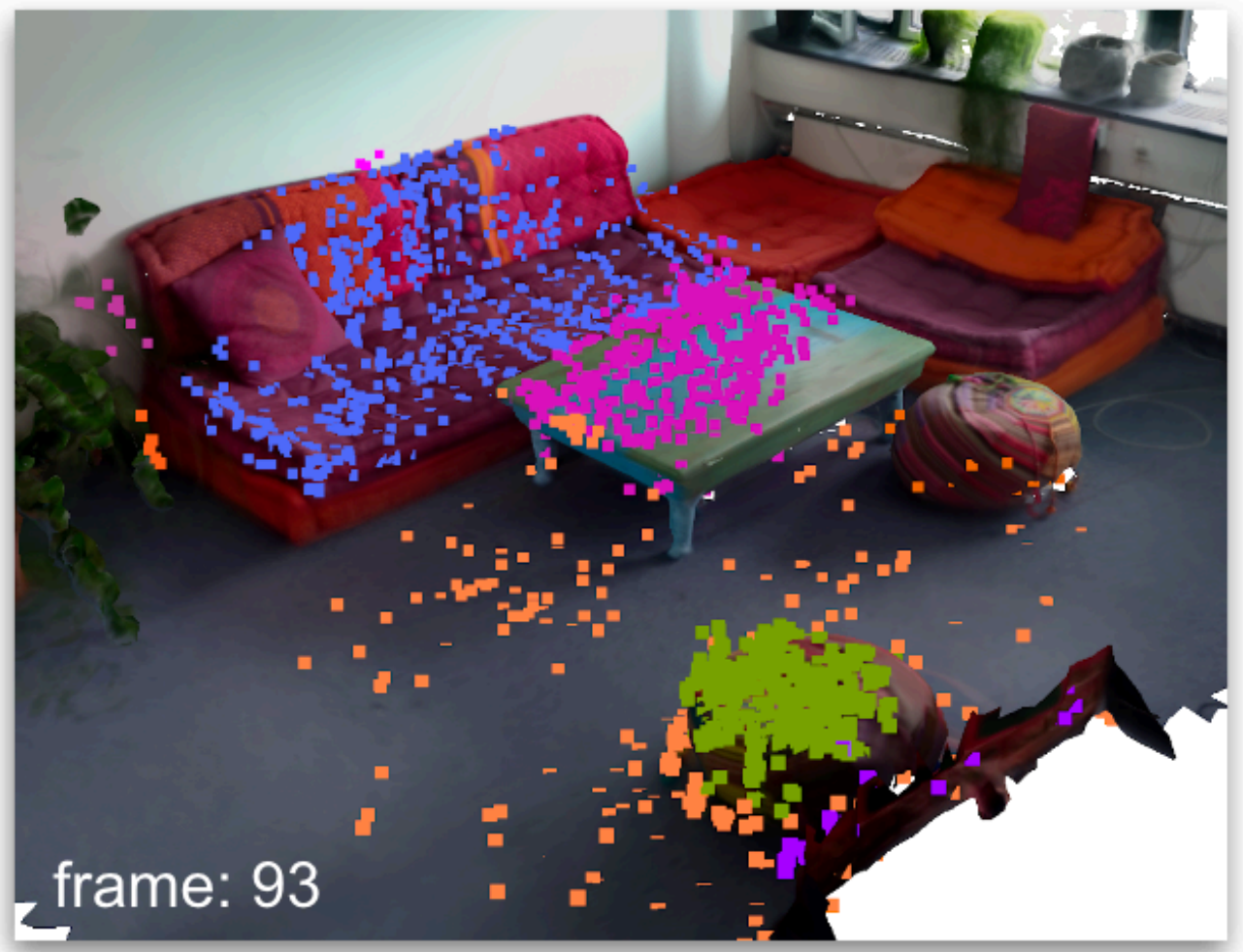
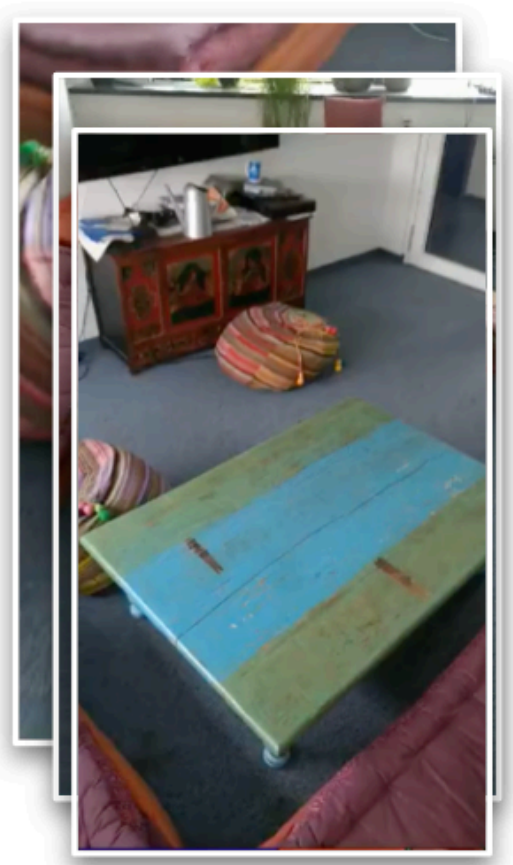


Nassir Navab

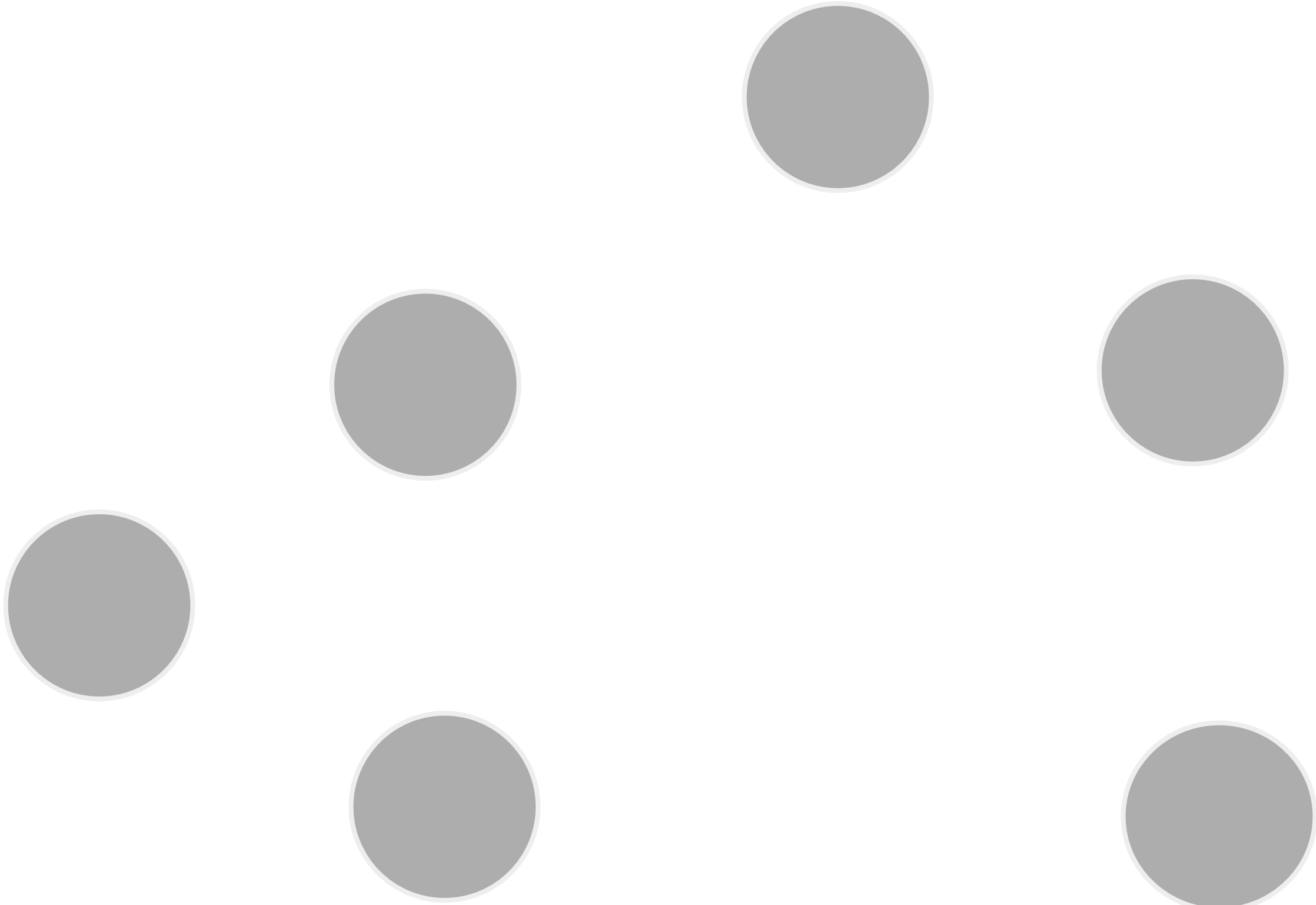


Federico Tombari

Our method

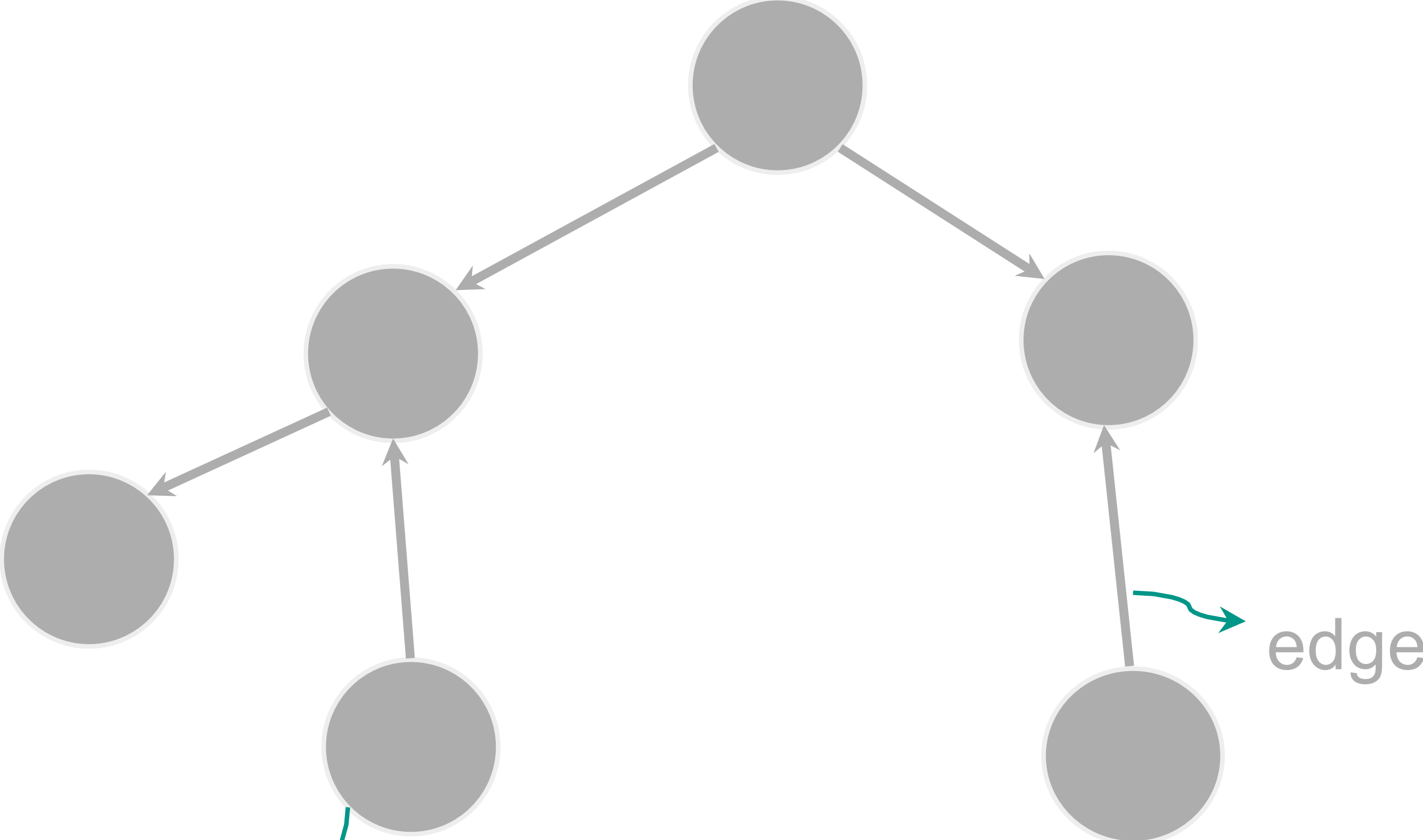


Semantic Scene Graphs



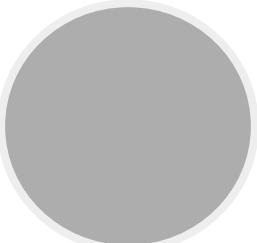
 Node

Semantic Scene Graphs



node

edge

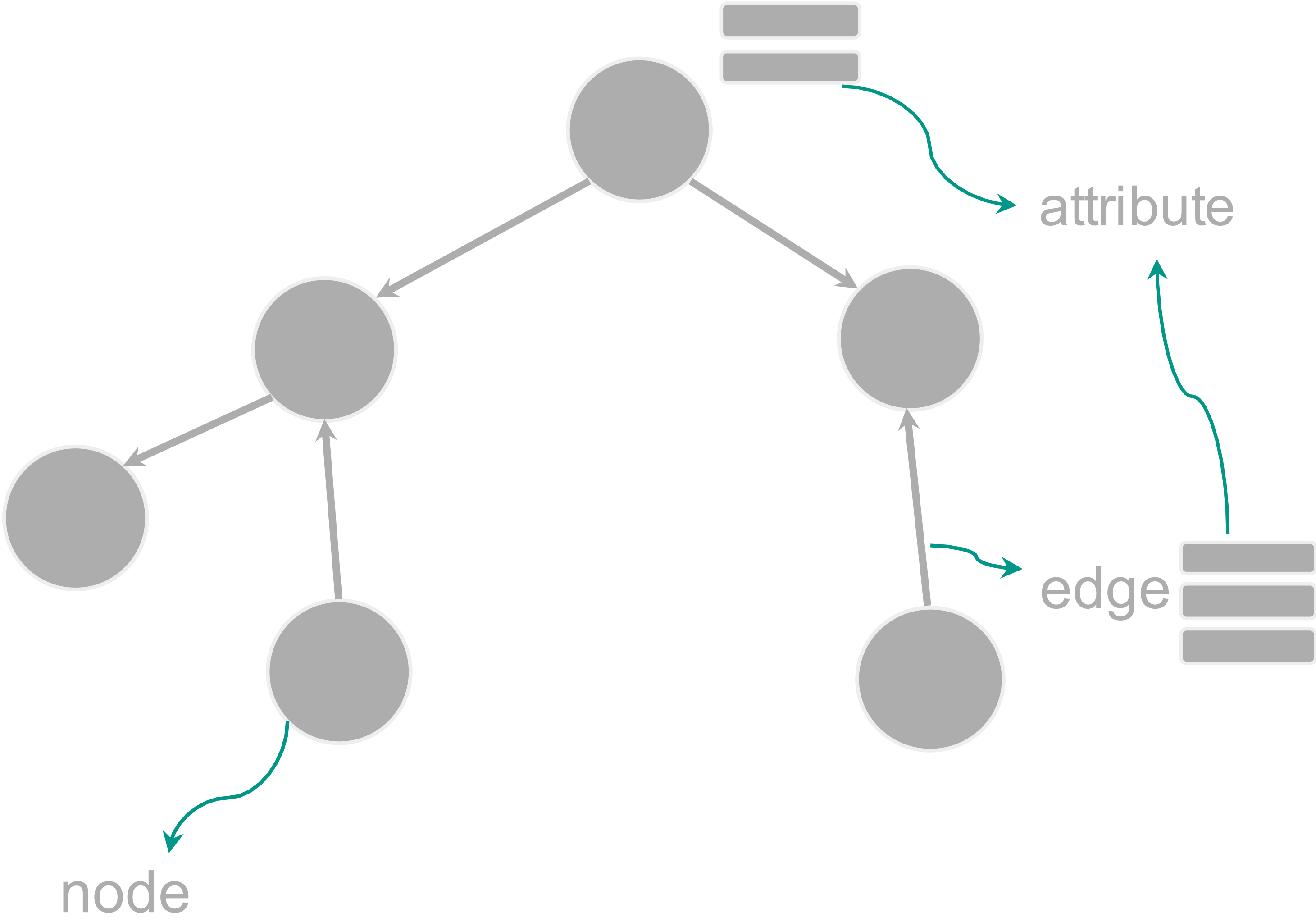


Node



Edge

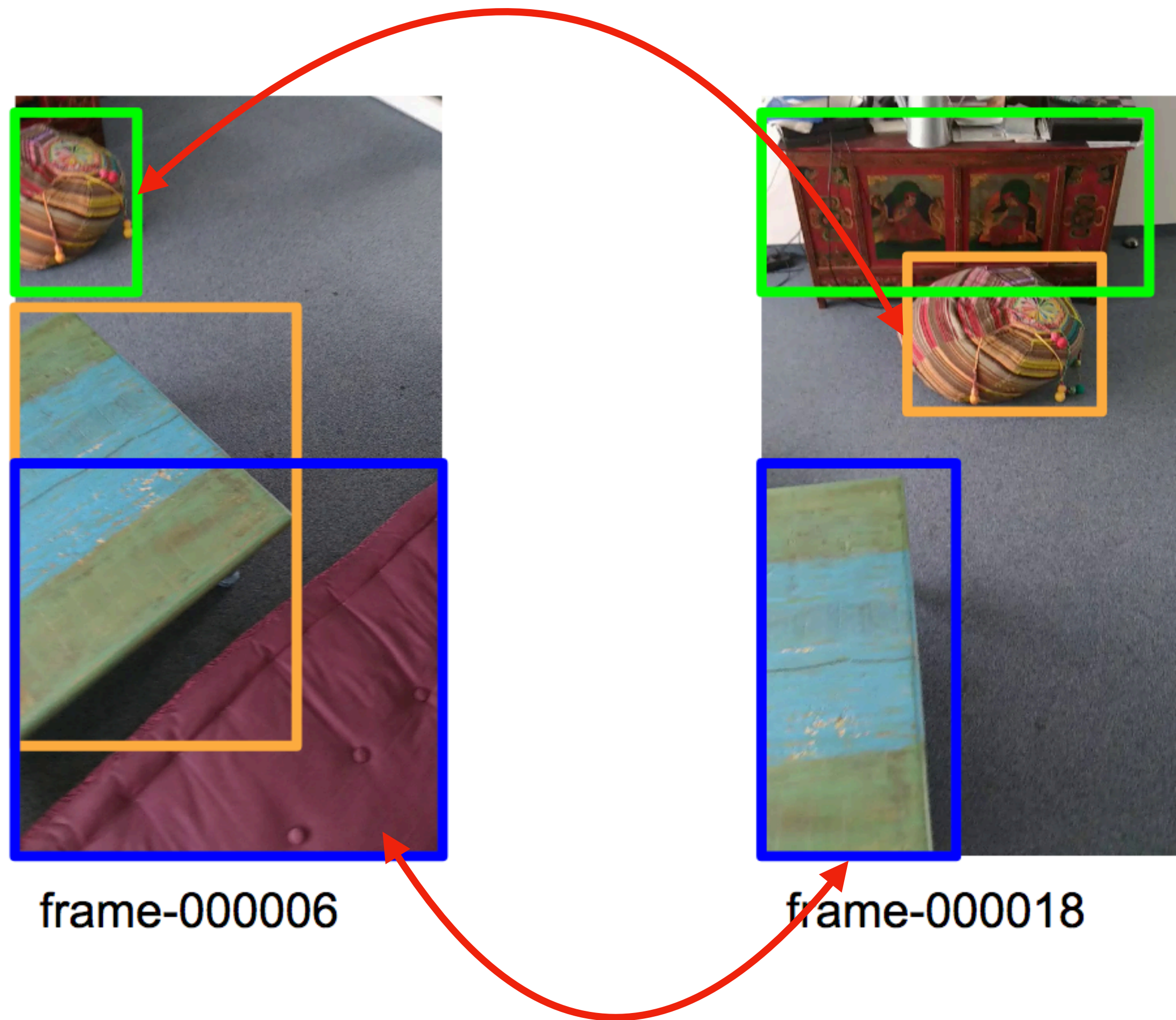
Semantic Scene Graphs



Node Edge Attribute

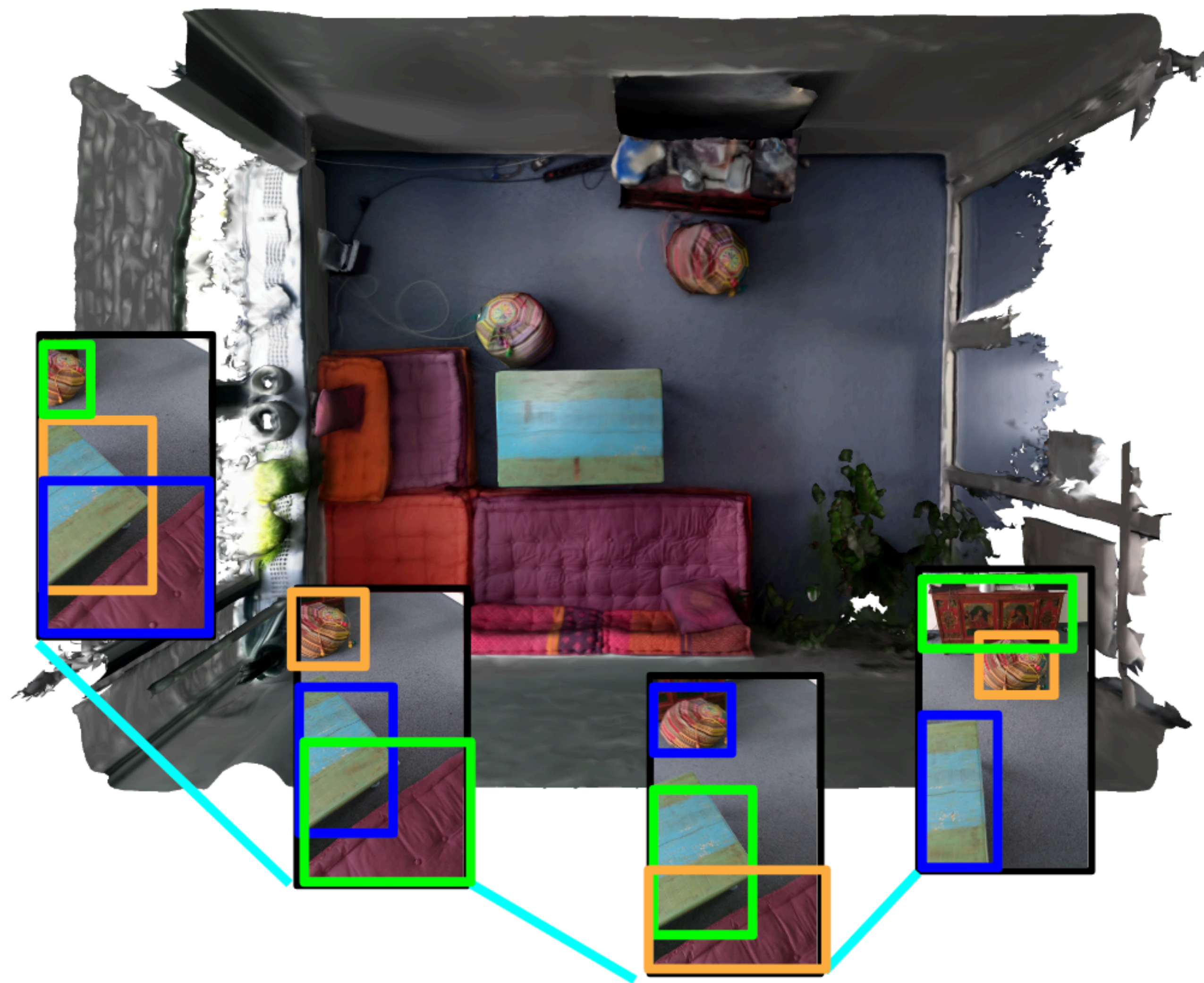
View-based Semantic Scene Graphs Generation (SSG)

Estimated scene graphs lacking temporal consistency!



The relationship between table and sofa cannot be estimated!

View-based SSG + Temporal consistency



Relationship estimation is limited by the field of view

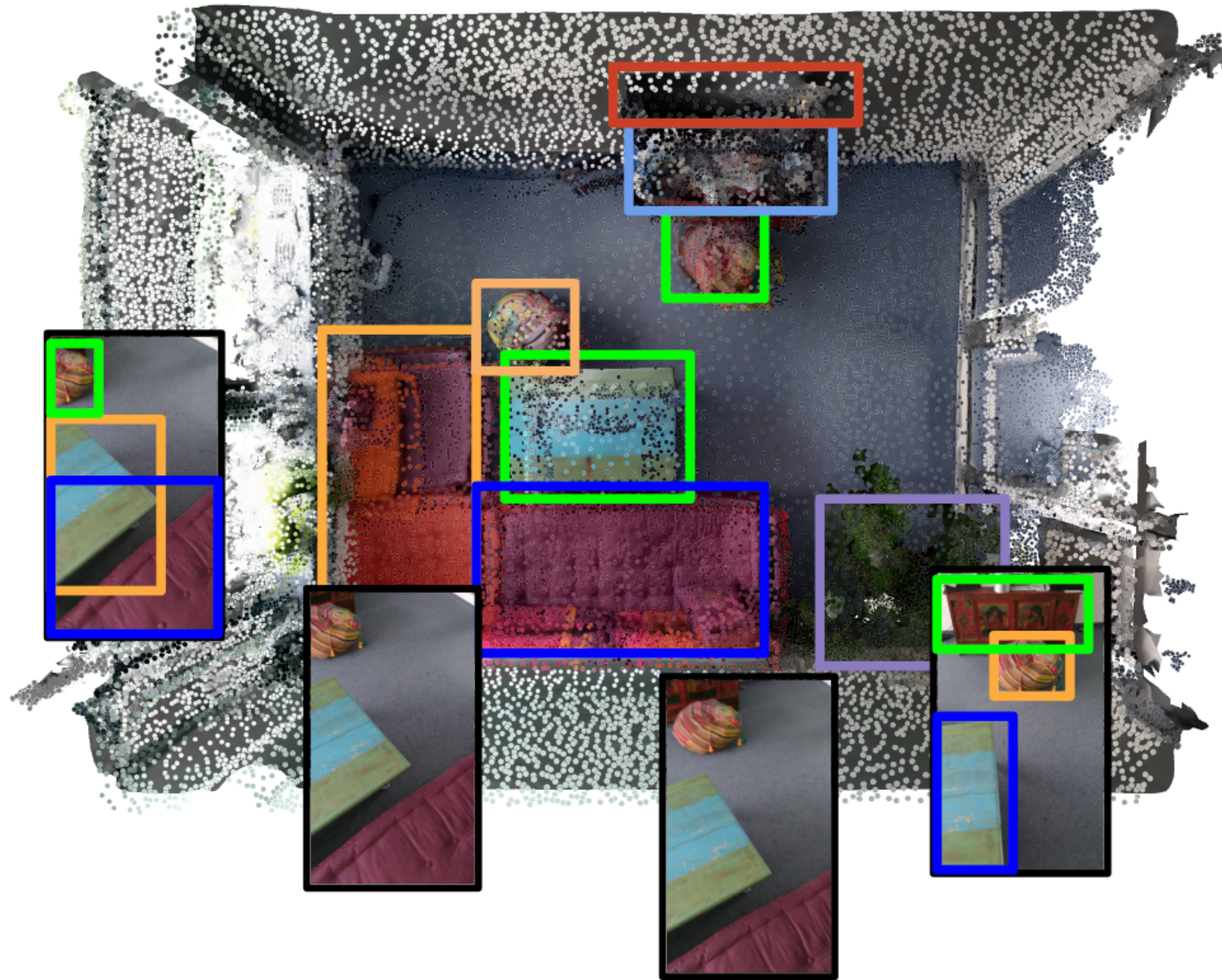
3D-based Semantic SSG



Dense Geometry is not always available

Feature learning on unrealistic texture

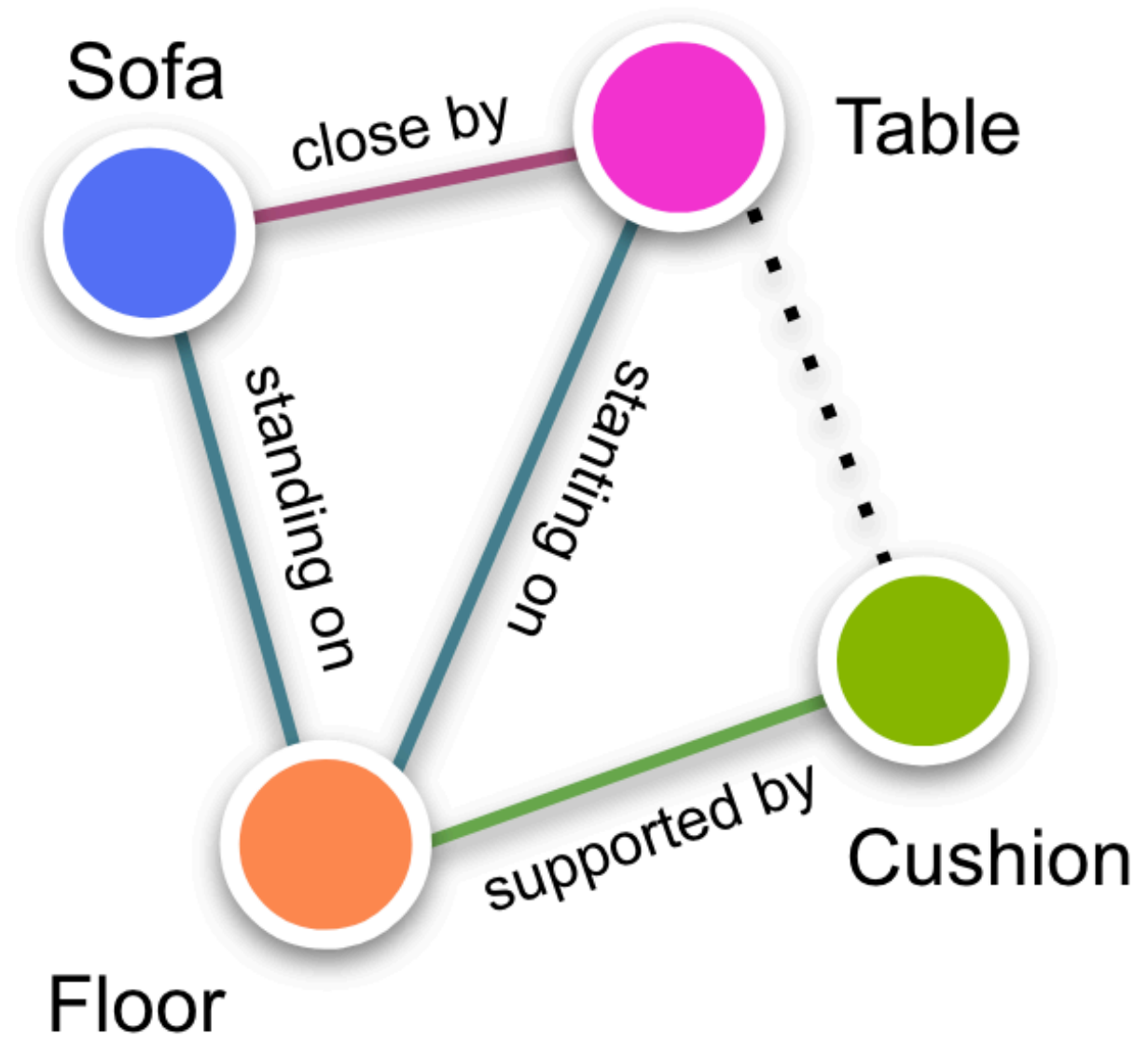
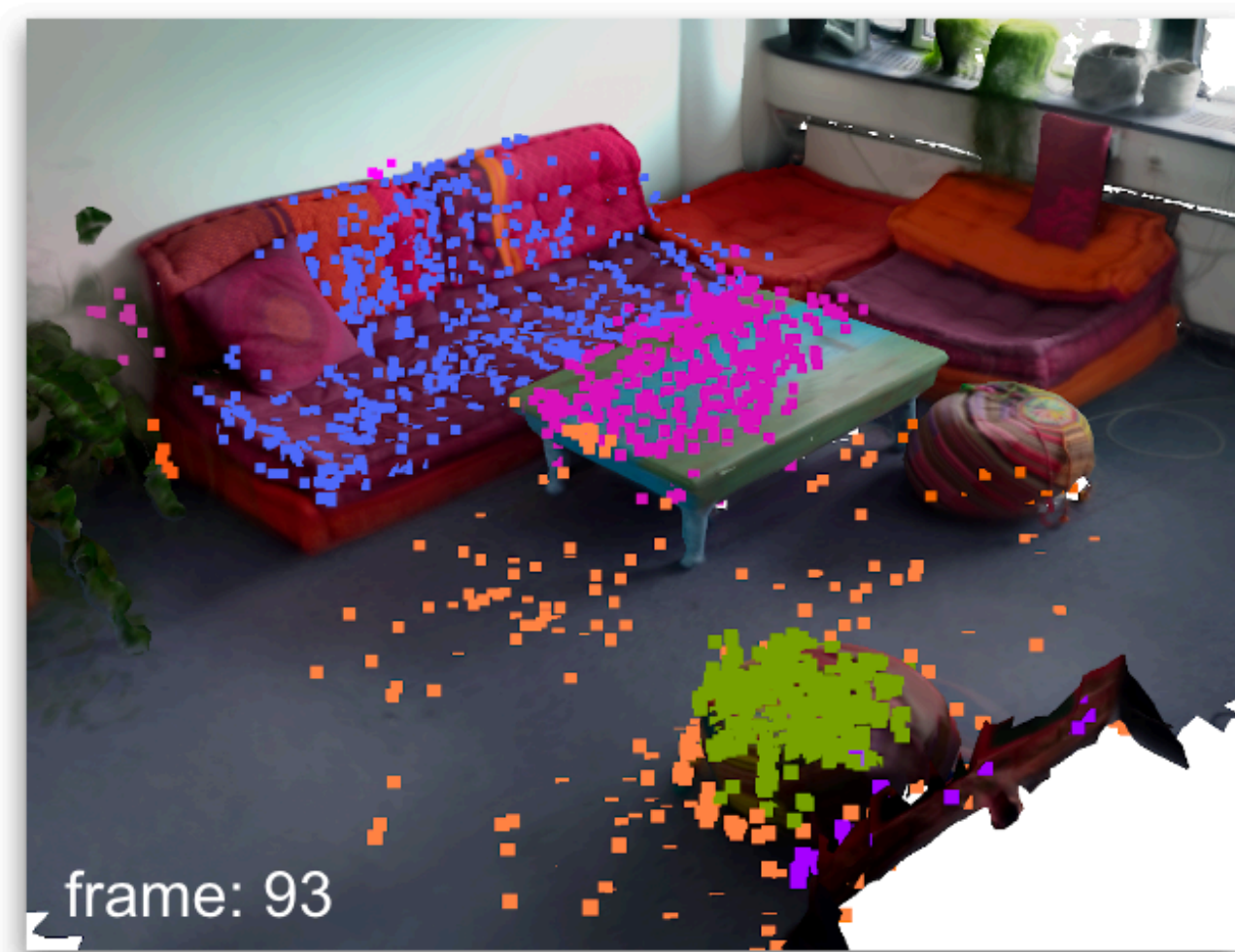
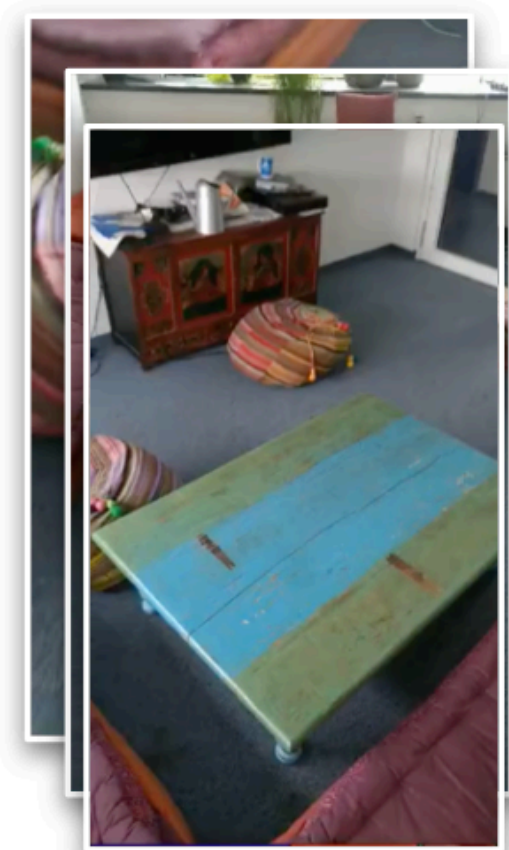
Direct estimation from 3D geometry



Node feature from 2D

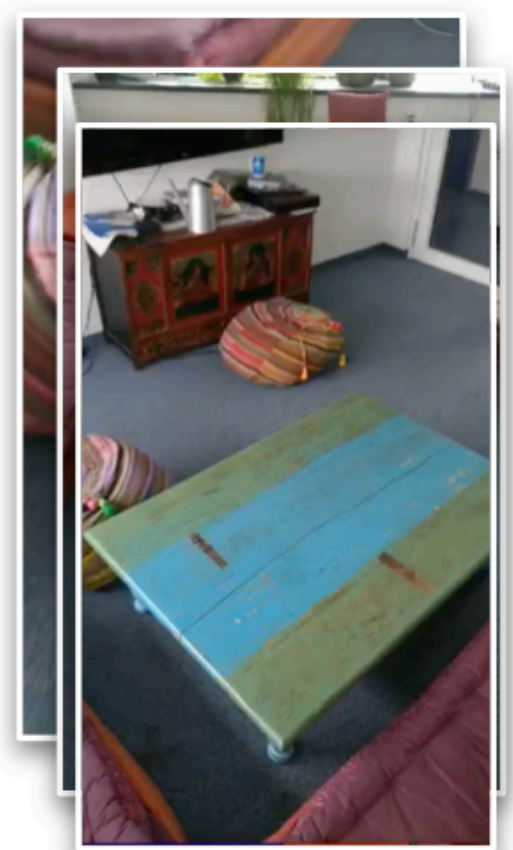
Edge feature from 3D

Our idea

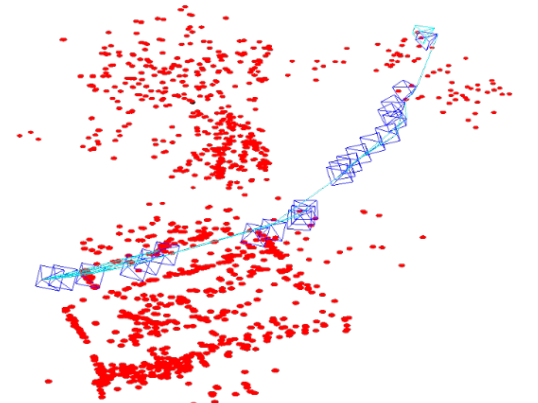


Incremental Entity Estimation

RGB Sequence



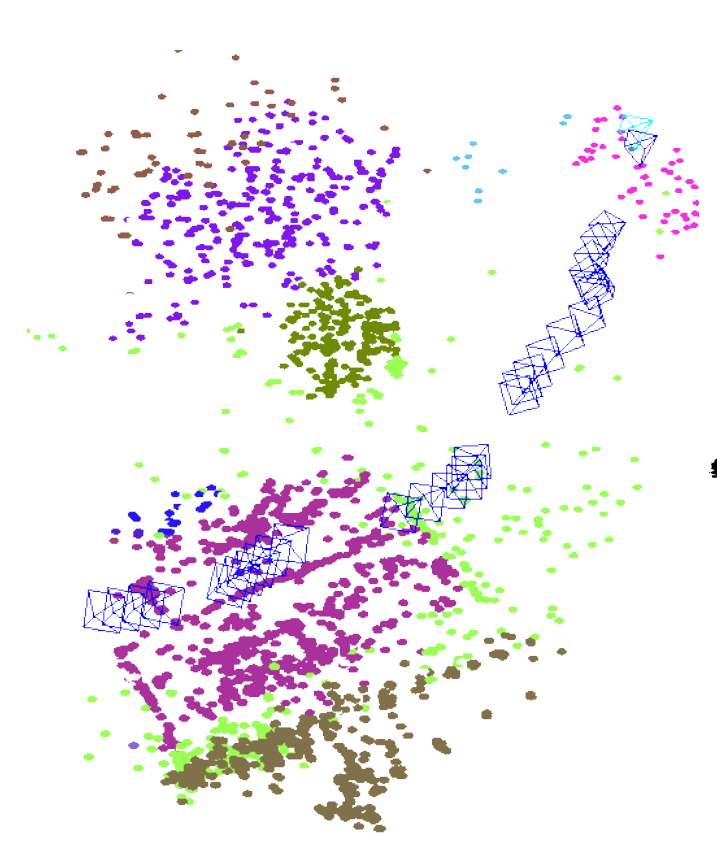
SLAM



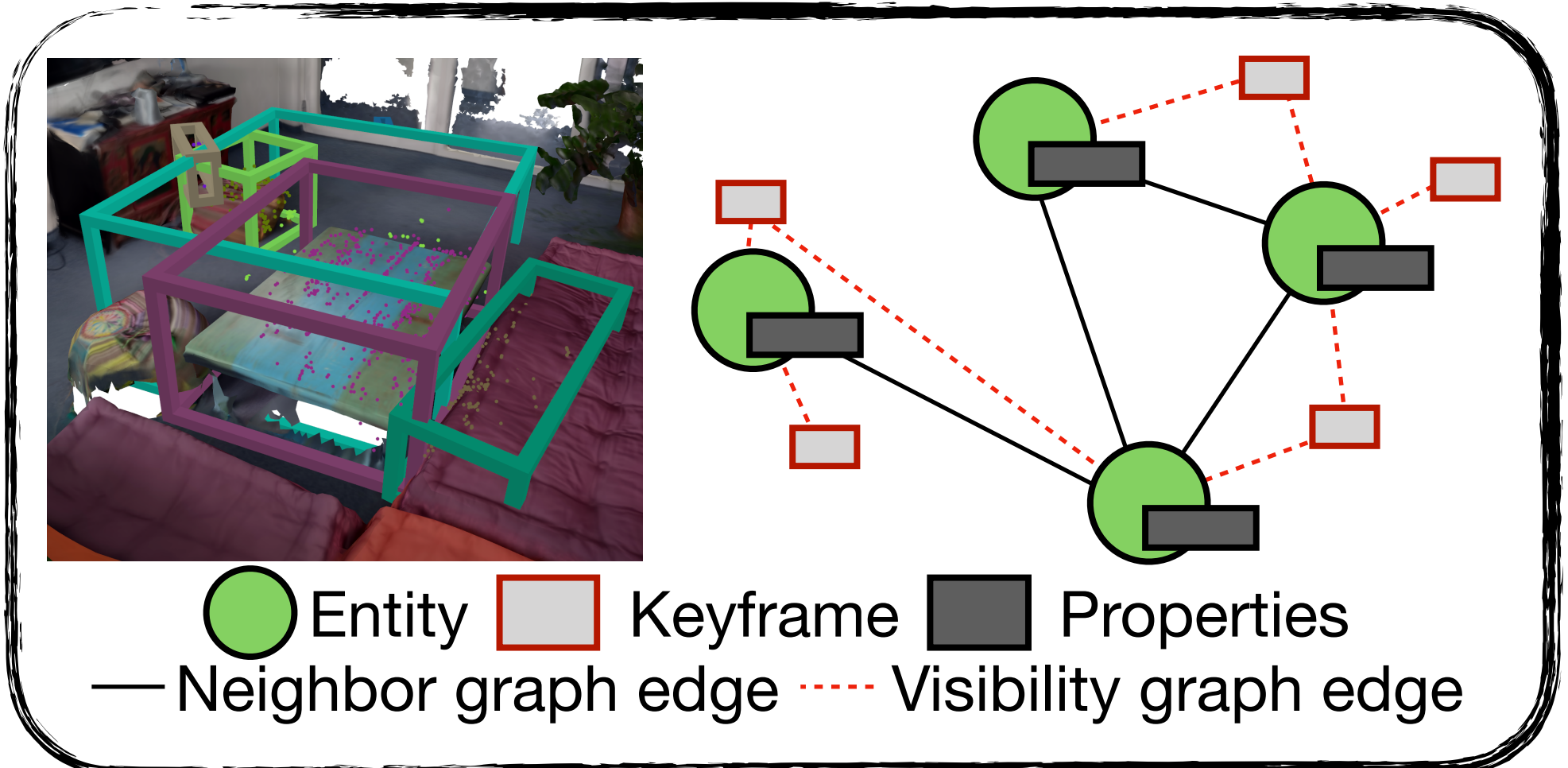
2D Entity Prediction



3D Entity Association

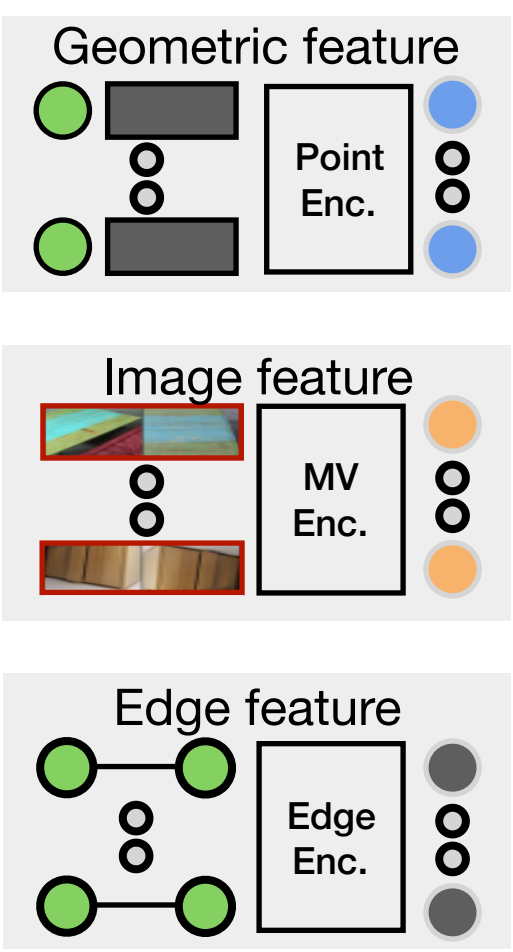


Graph Extraction

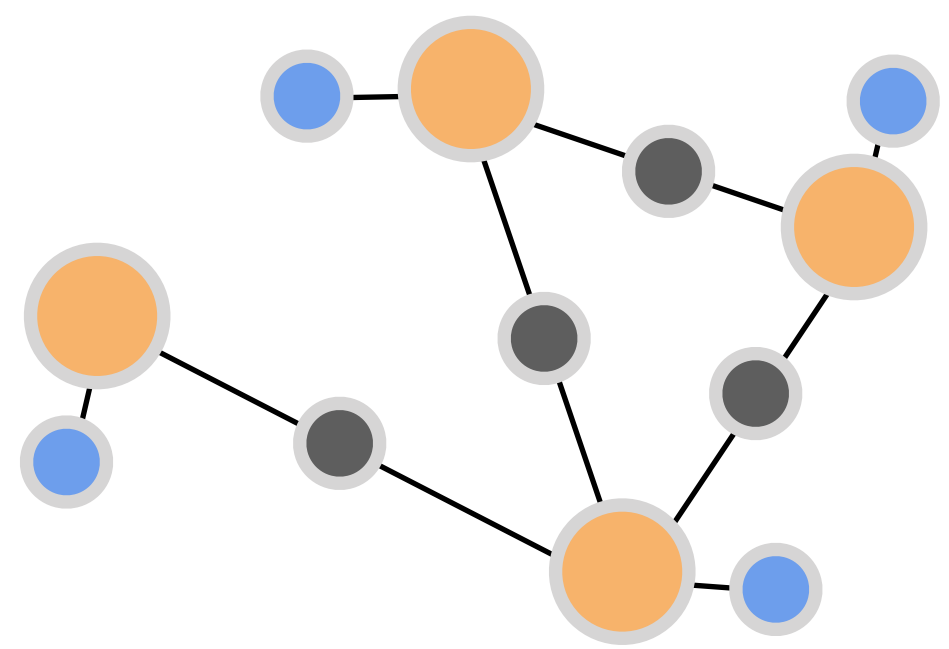


Semantic Scene Graph Prediction

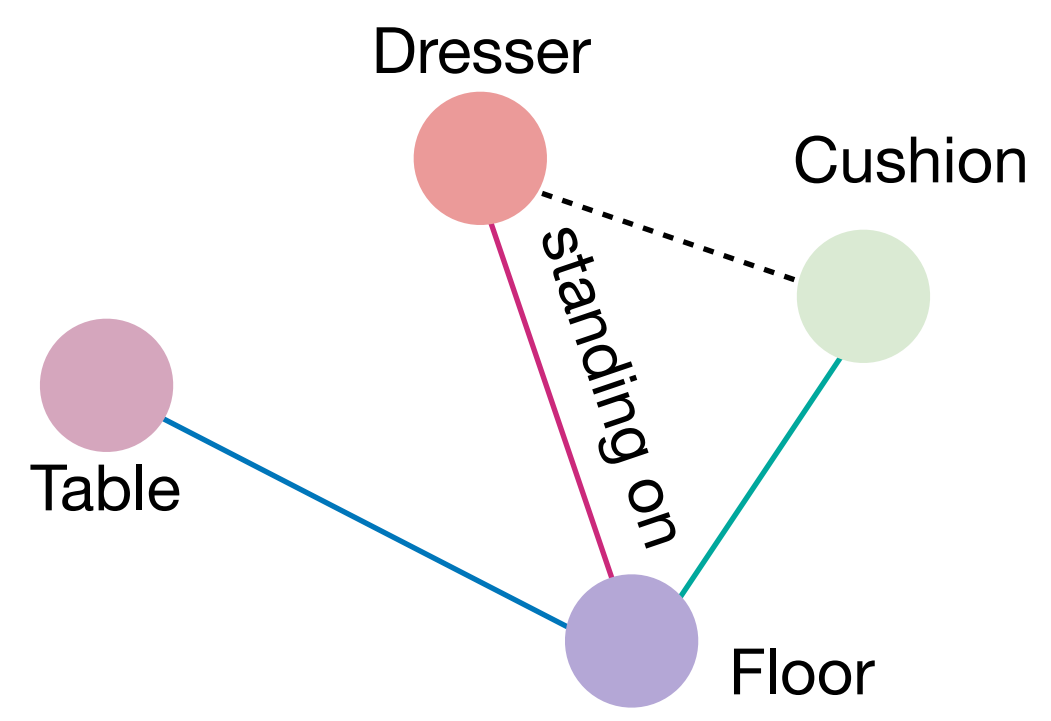
Feature Encoding



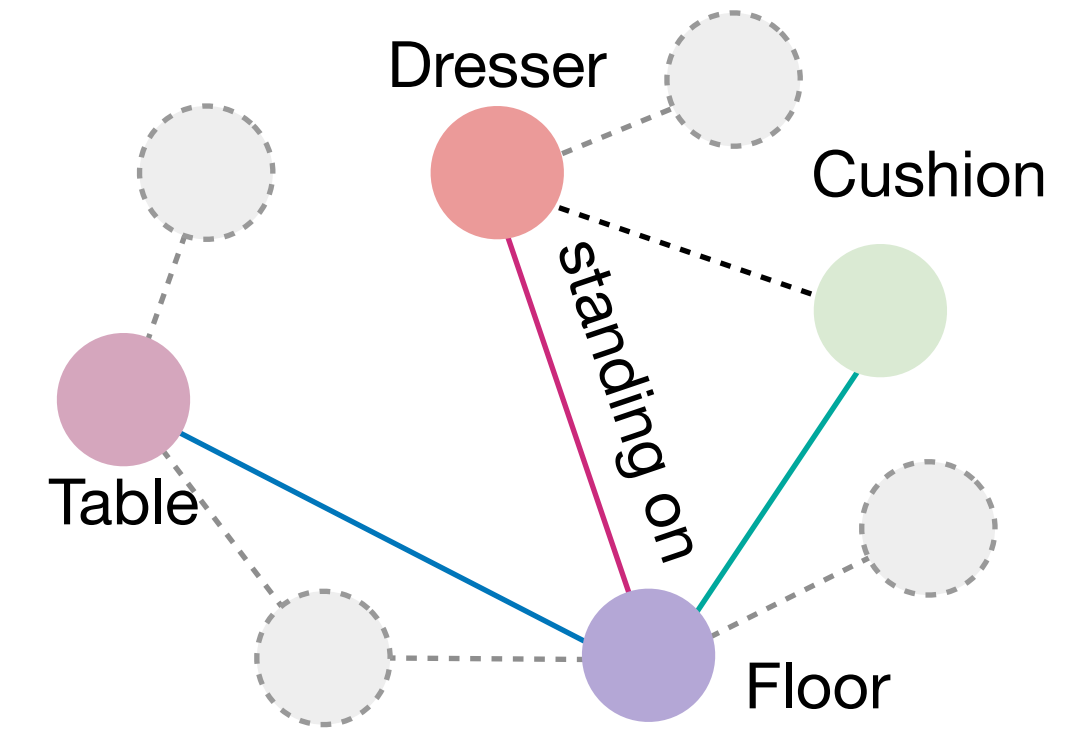
Message Passing



SSG Prediction



Fusion

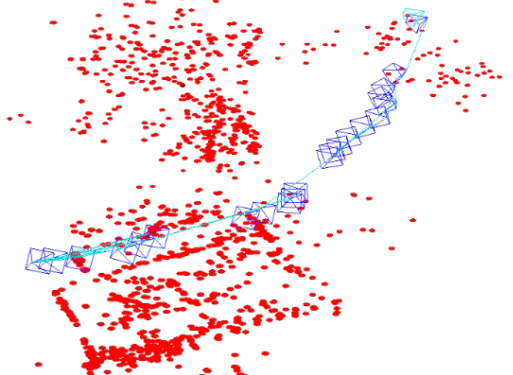


Incremental Entity Estimation

RGB Sequence



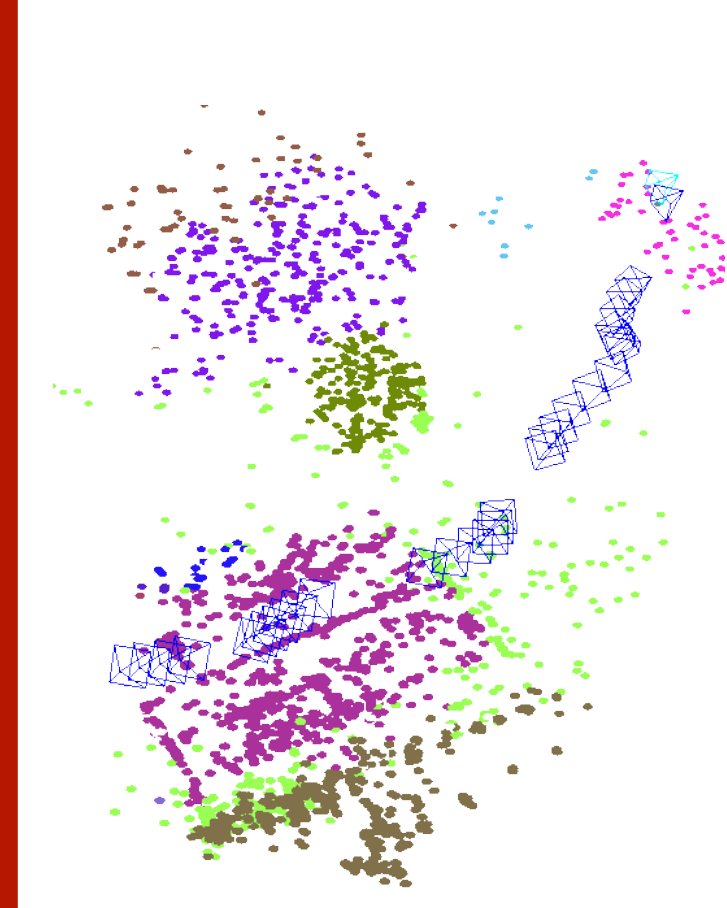
SLAM



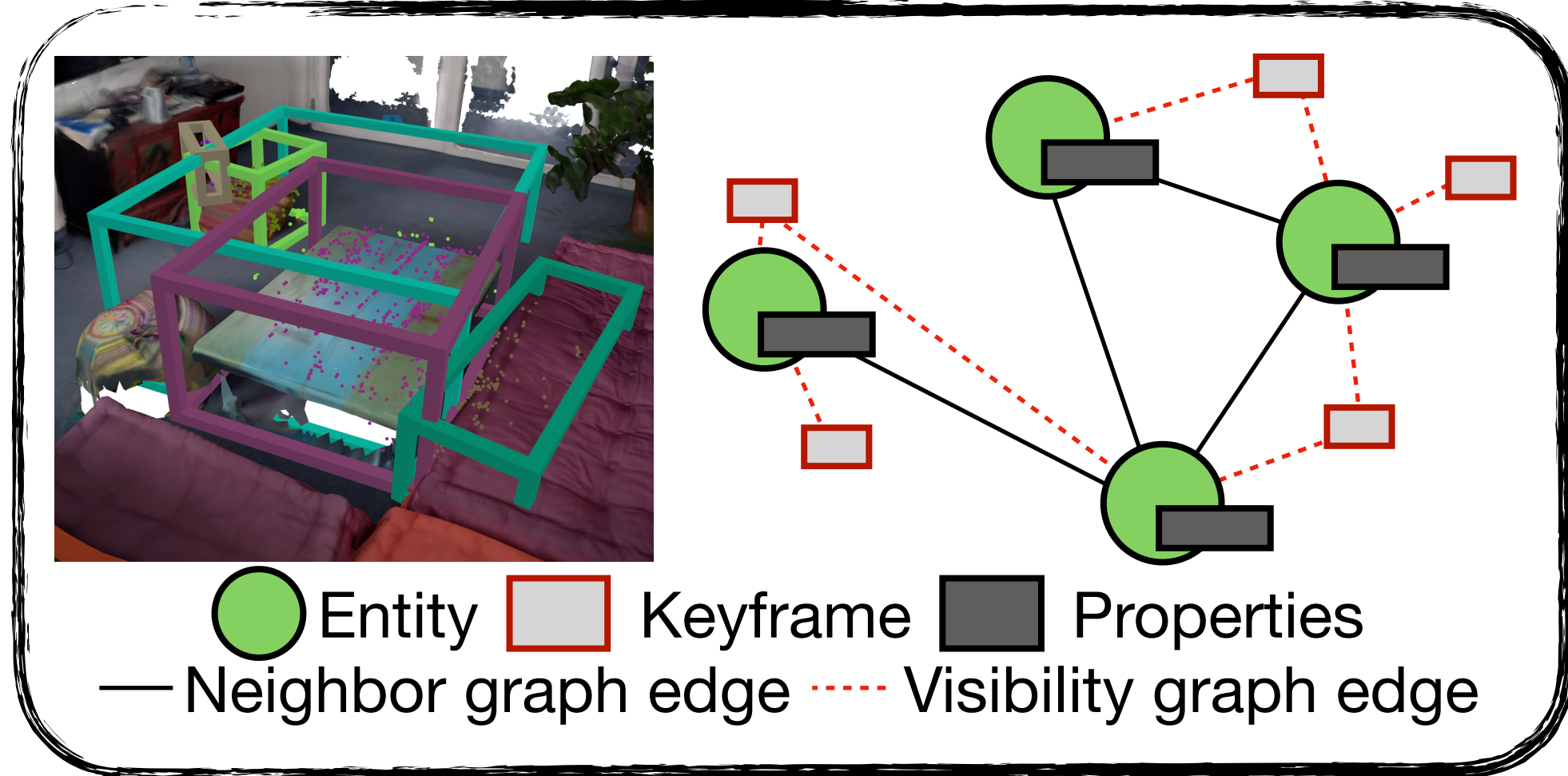
2D Entity Prediction



3D Entity Association

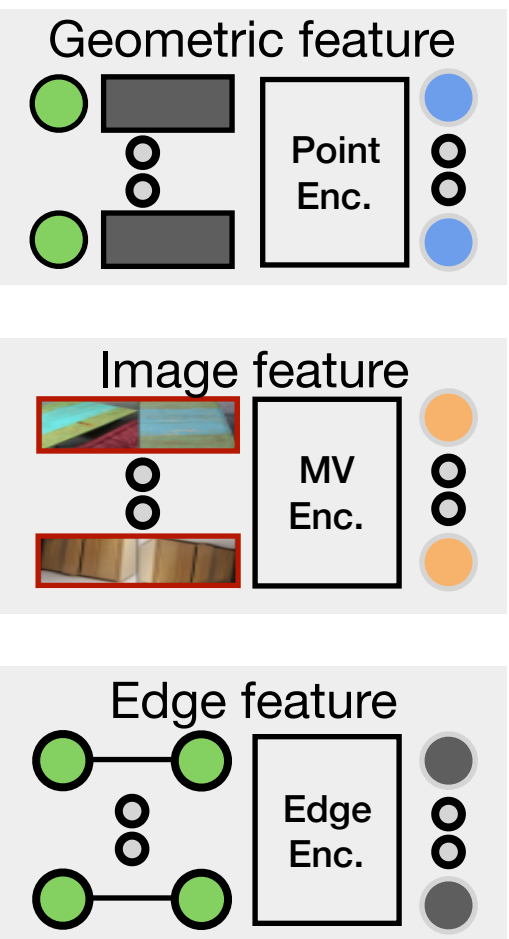


Graph Extraction

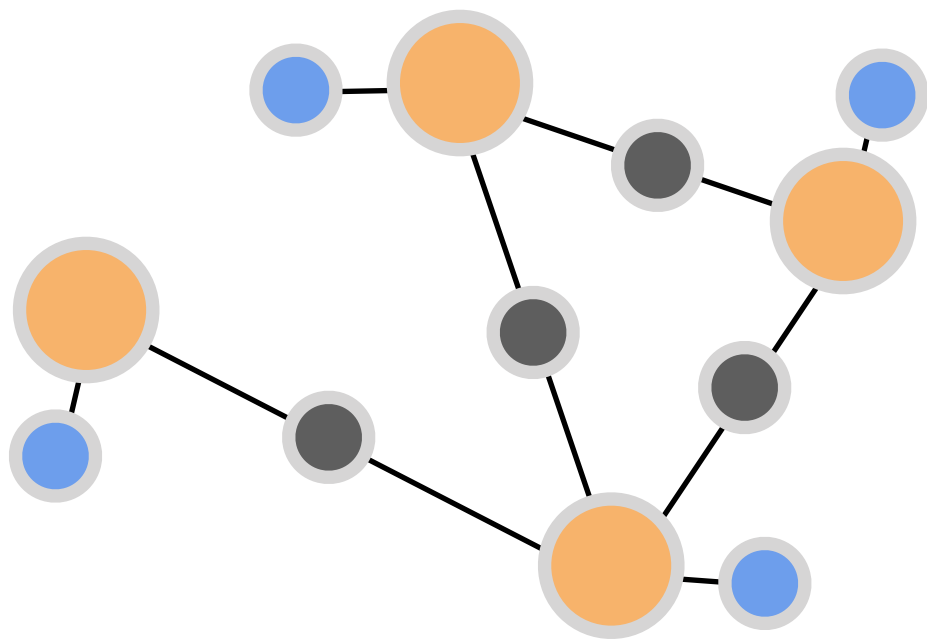


Semantic Scene Graph Prediction

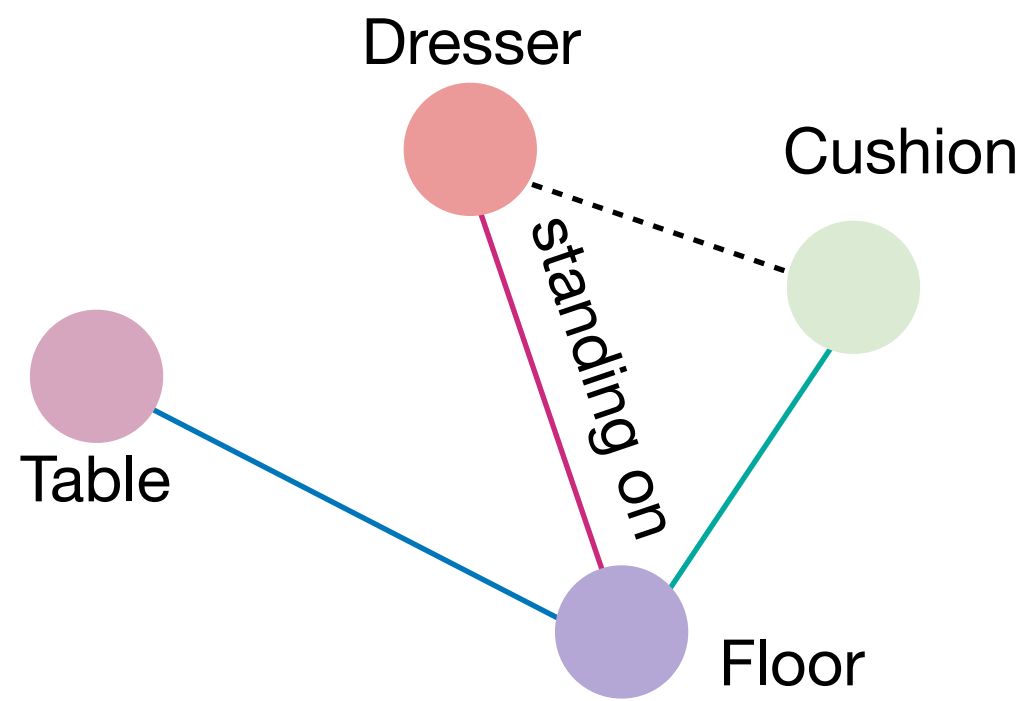
Feature Encoding



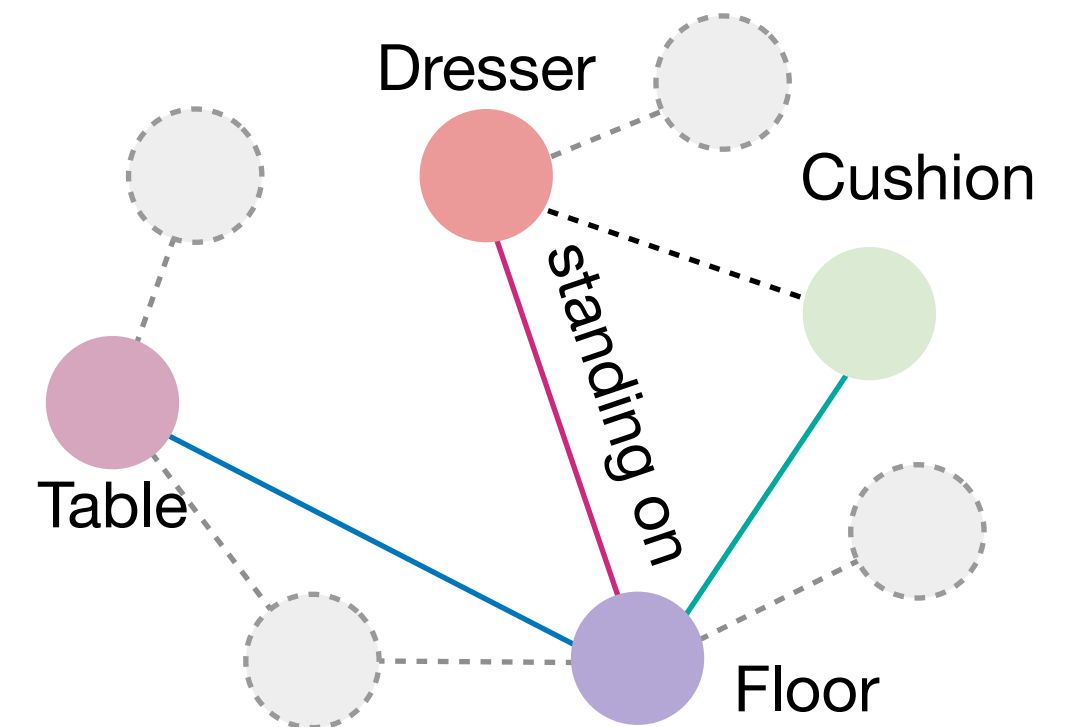
Message Passing

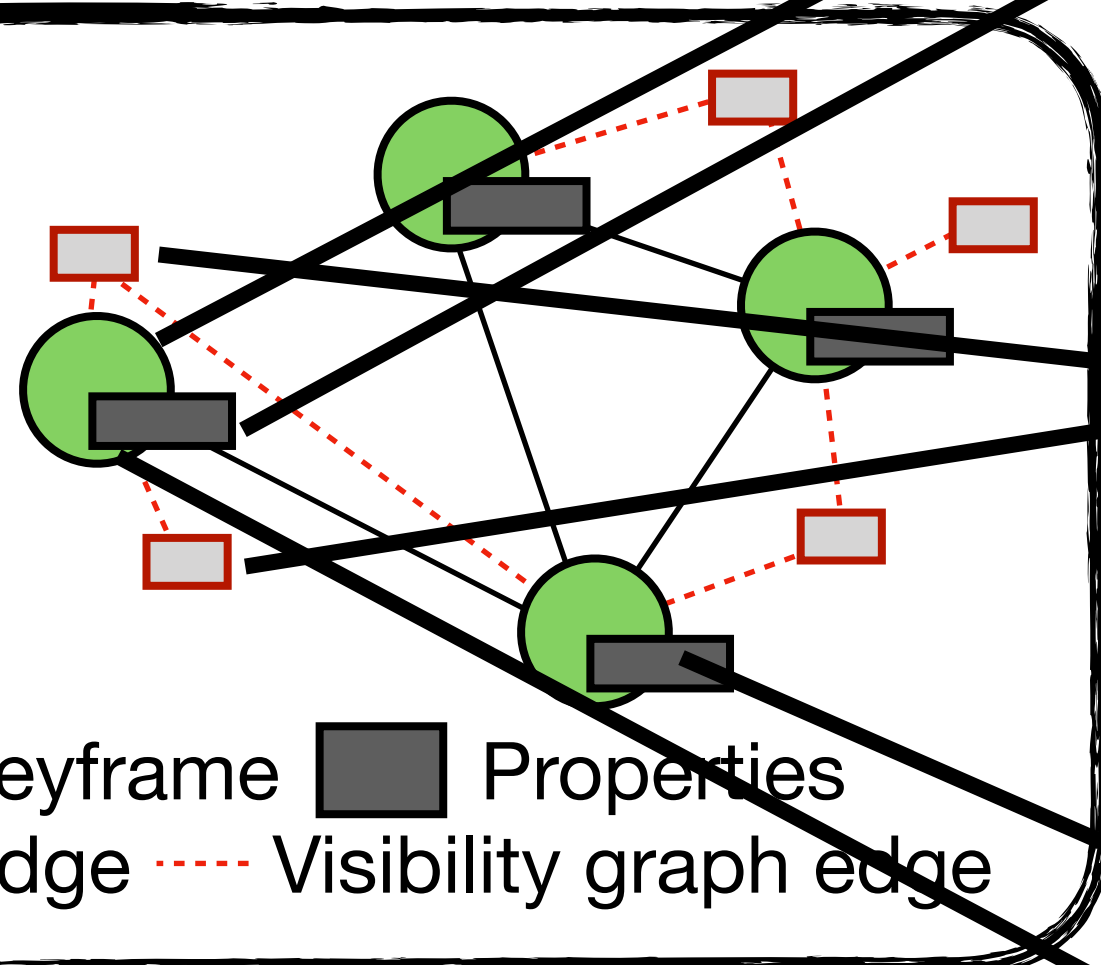
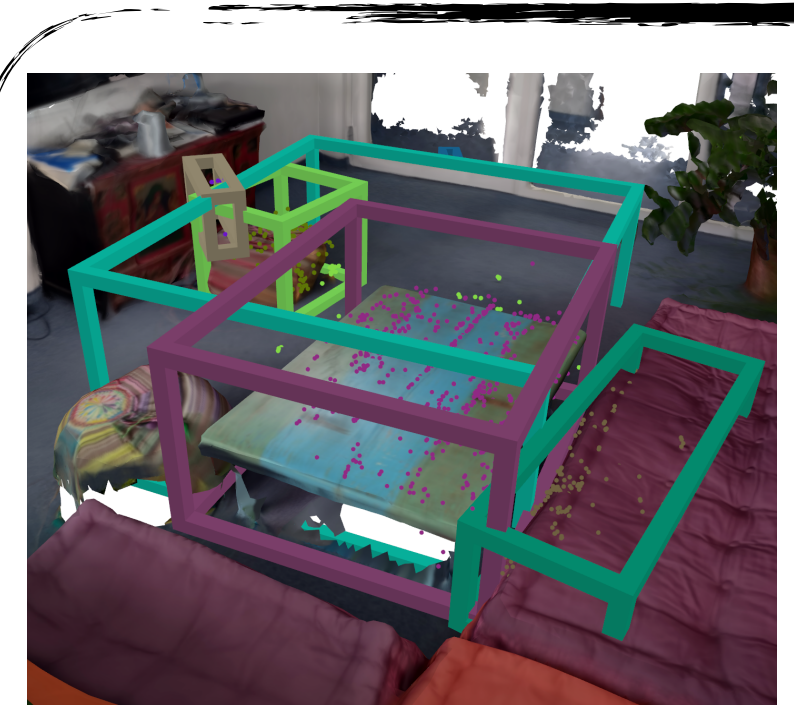


SSG Prediction

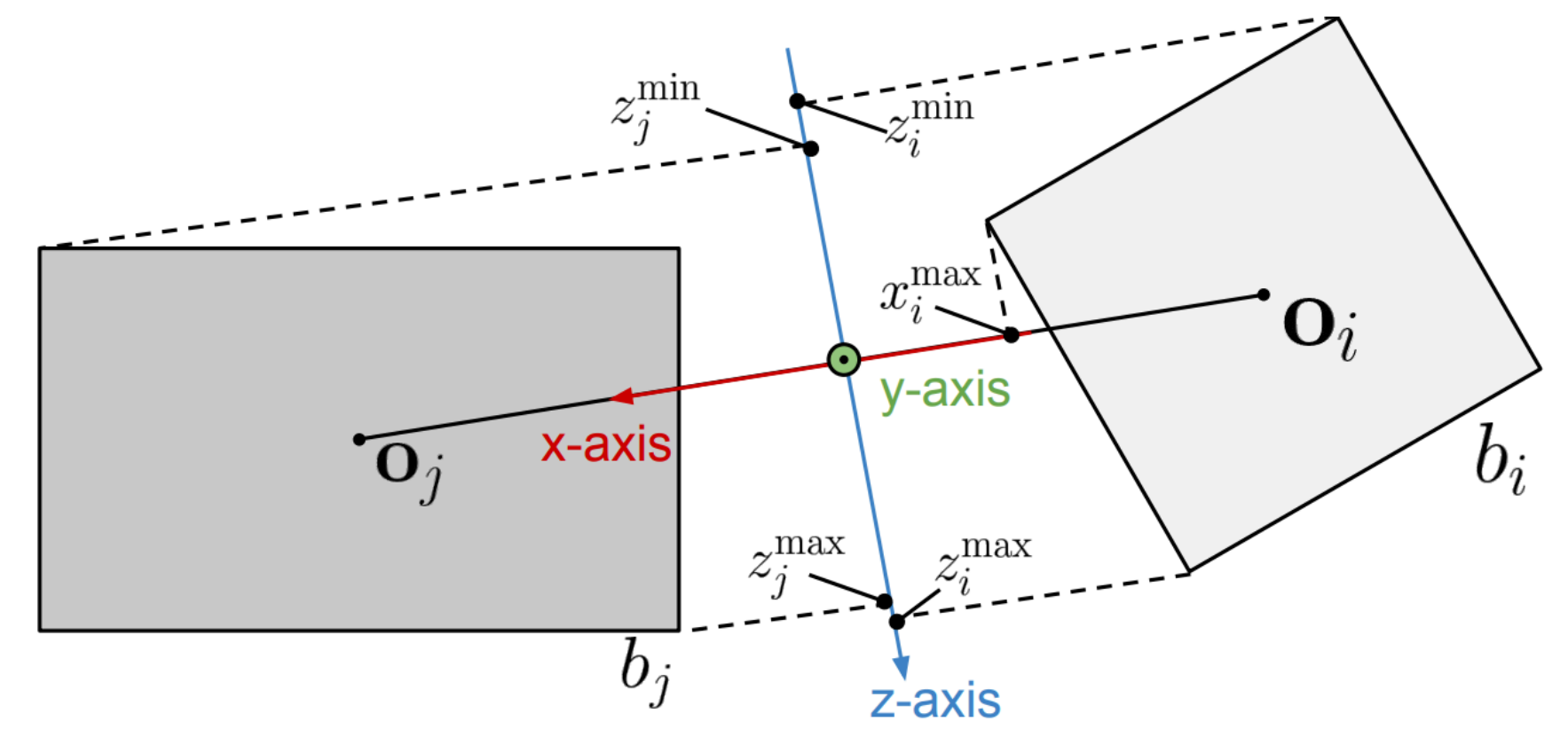
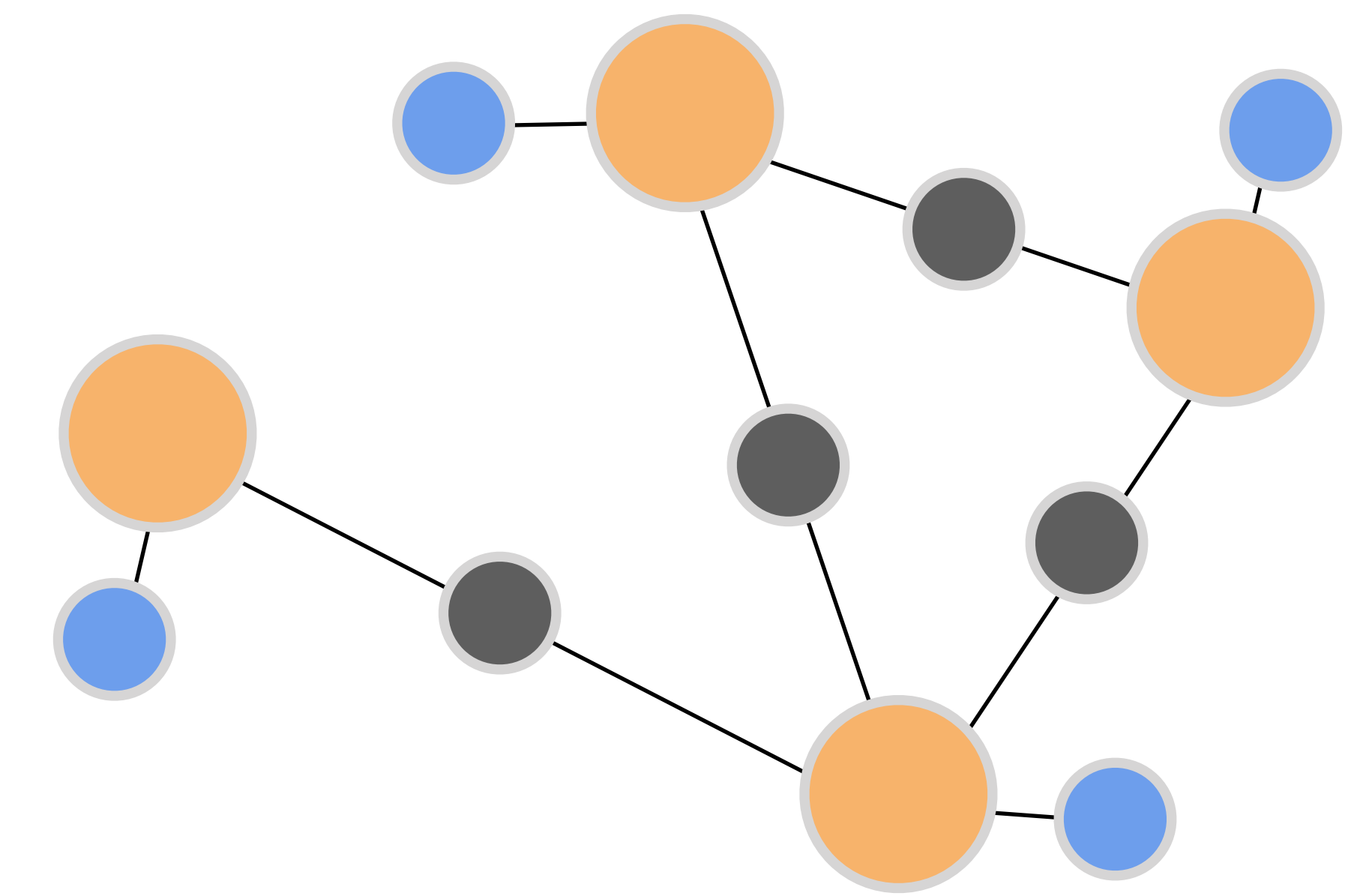
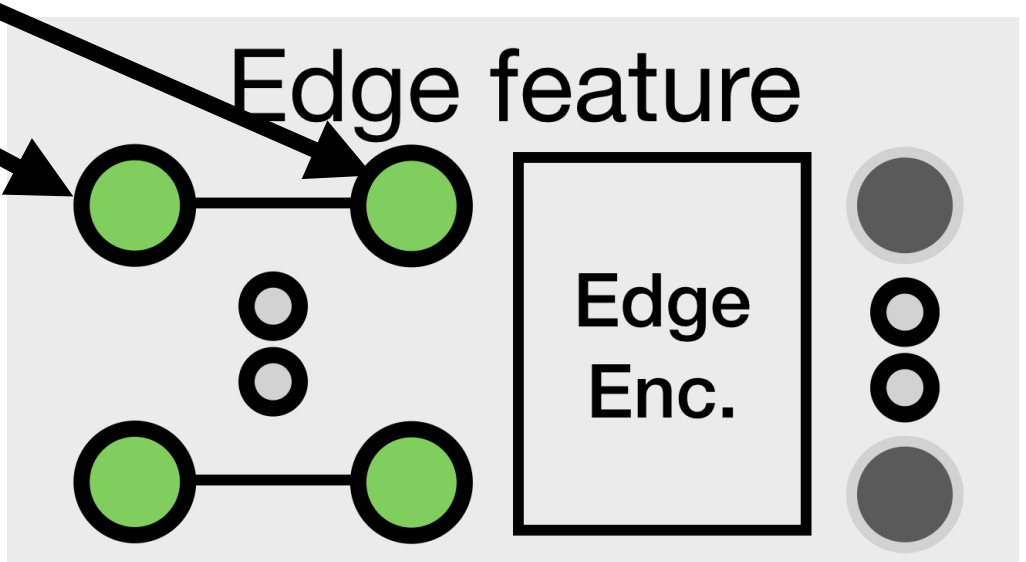
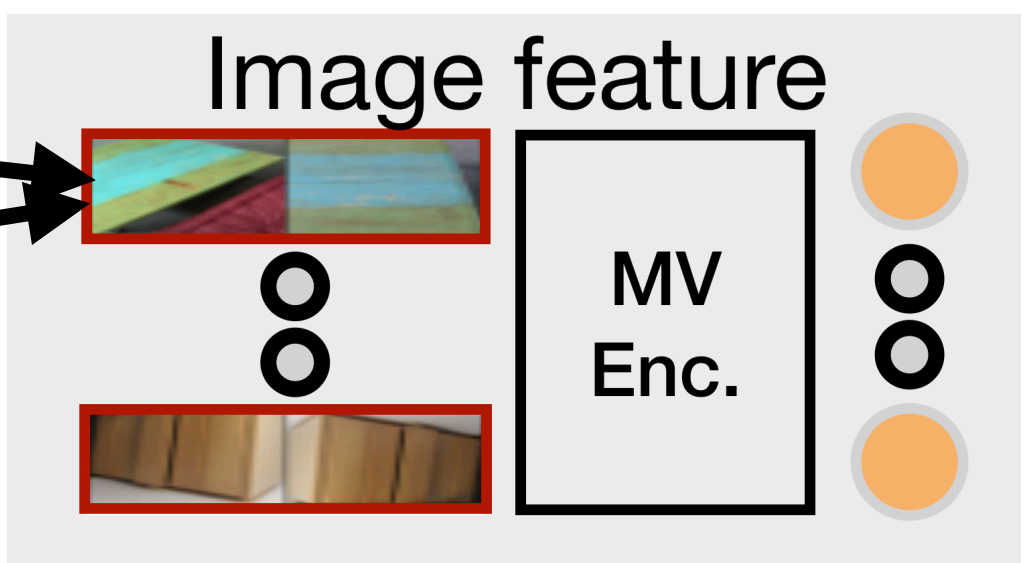
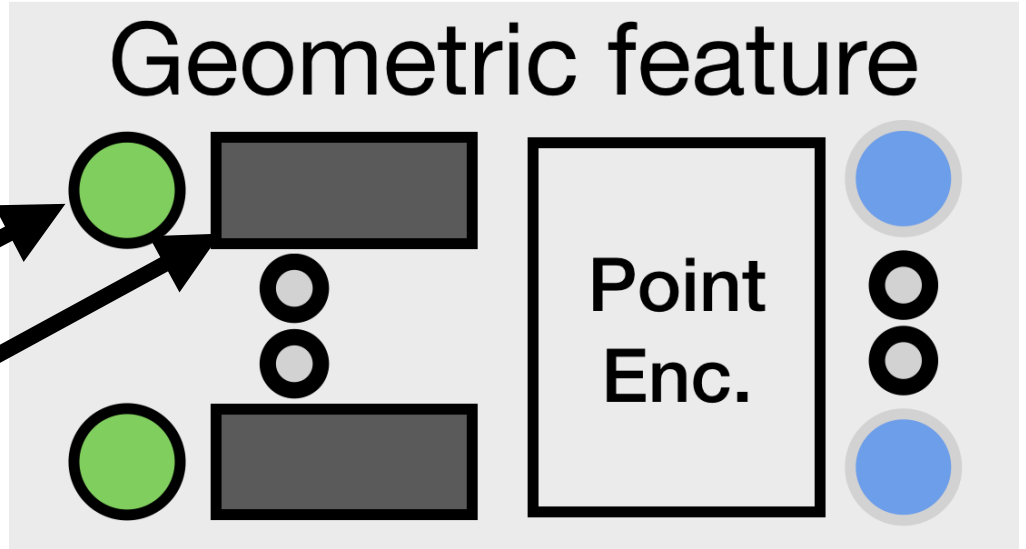


Fusion



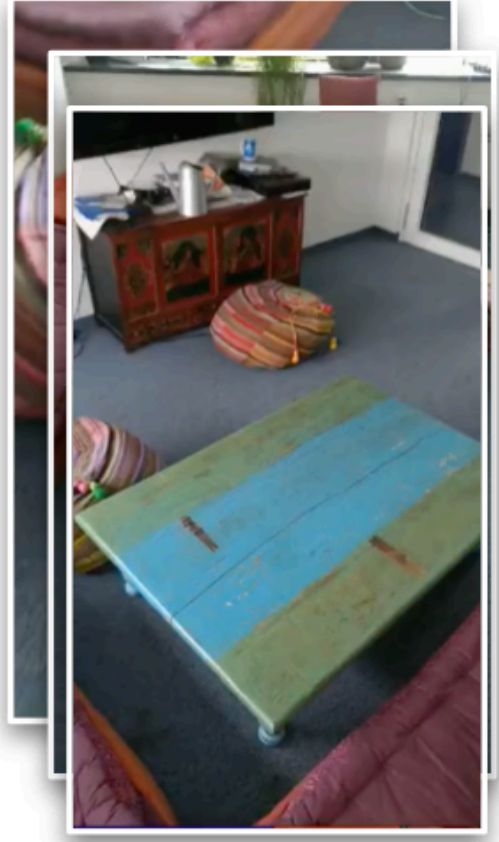


● Entity □ Keyframe ■ Properties
— Neighbor graph edge - - - Visibility graph edge

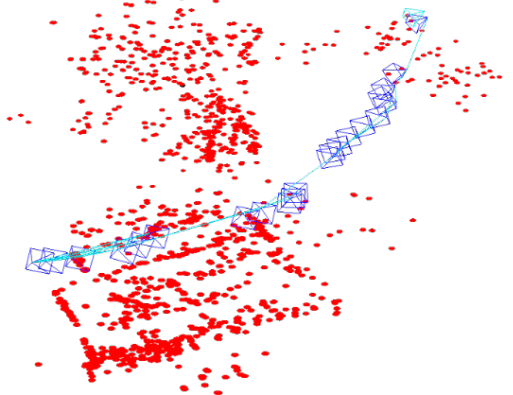


Incremental Entity Estimation

RGB Sequence



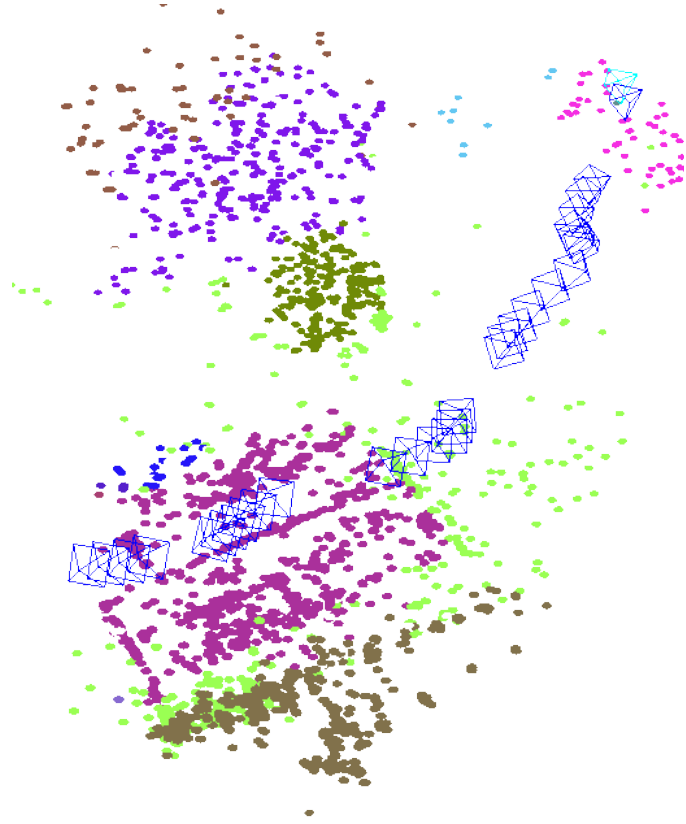
SLAM



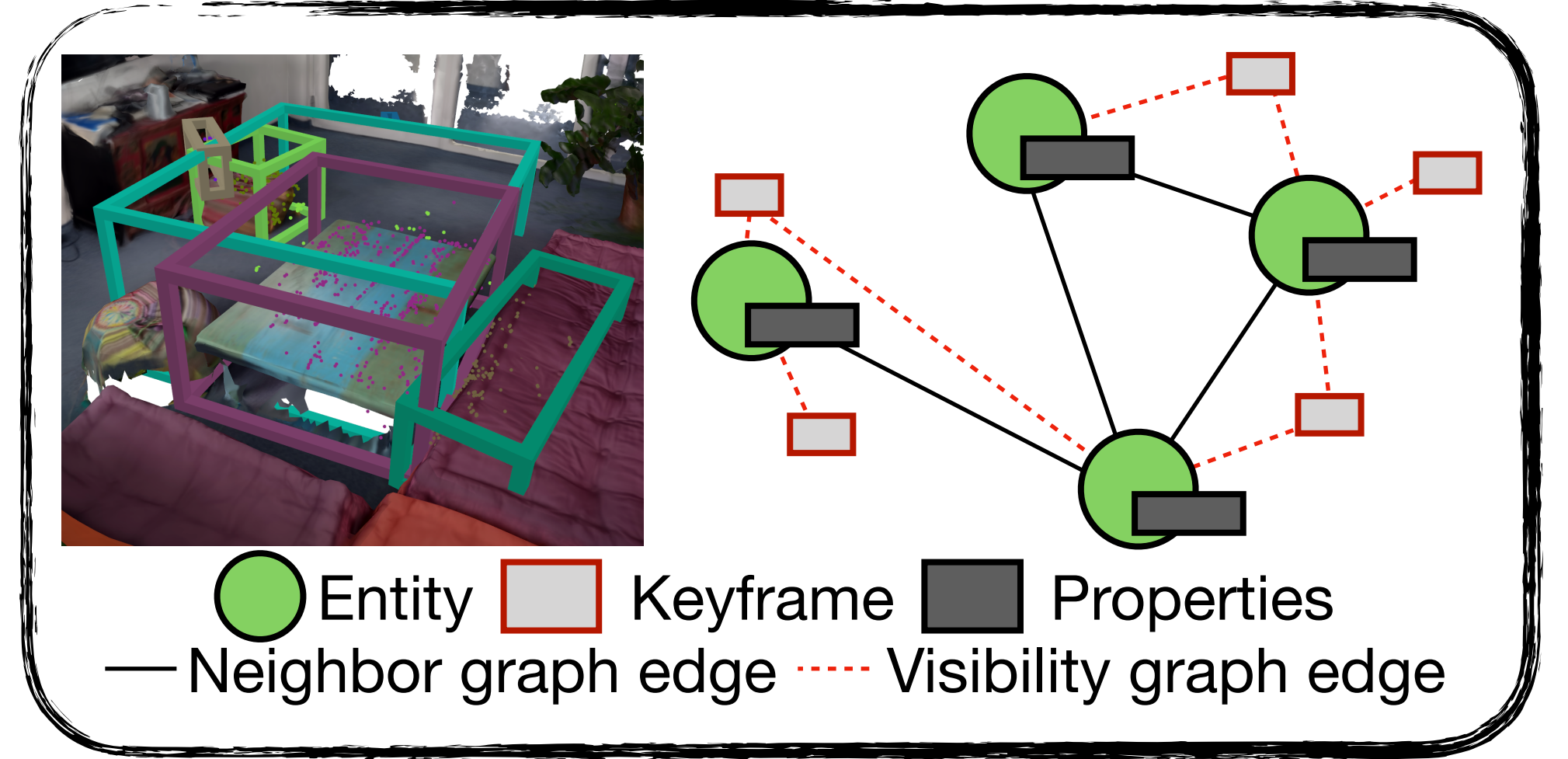
2D Entity Prediction



3D Entity Association

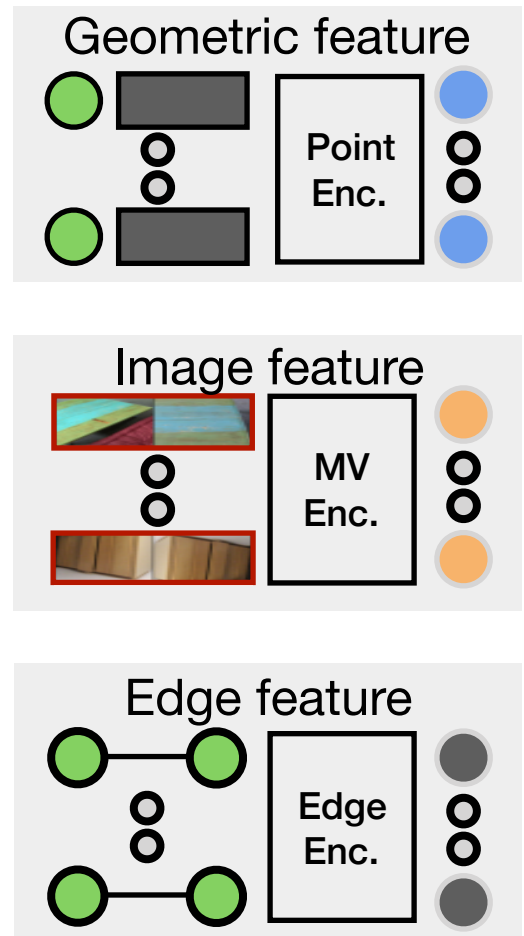


Graph Extraction

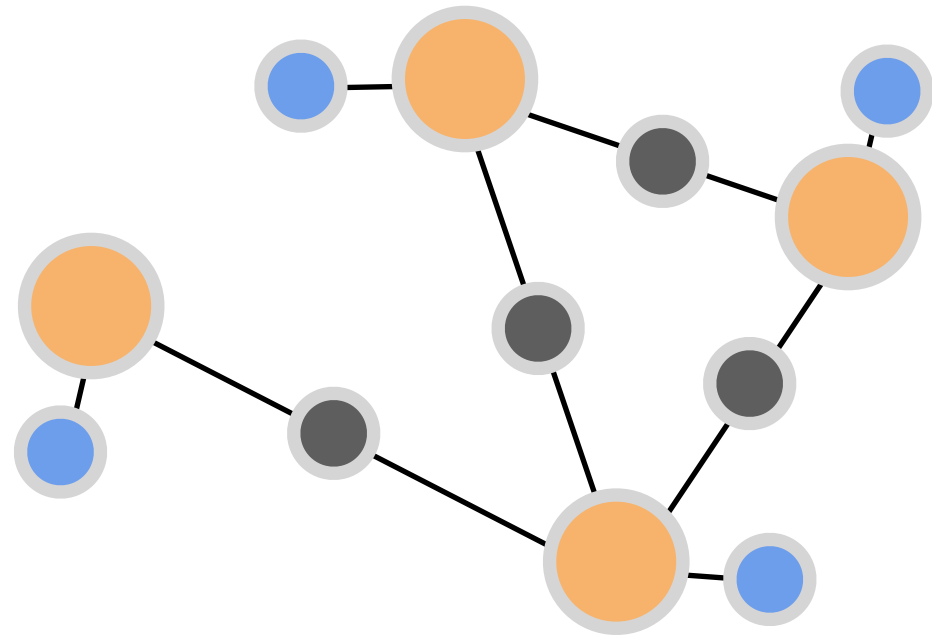


Semantic Scene Graph Prediction

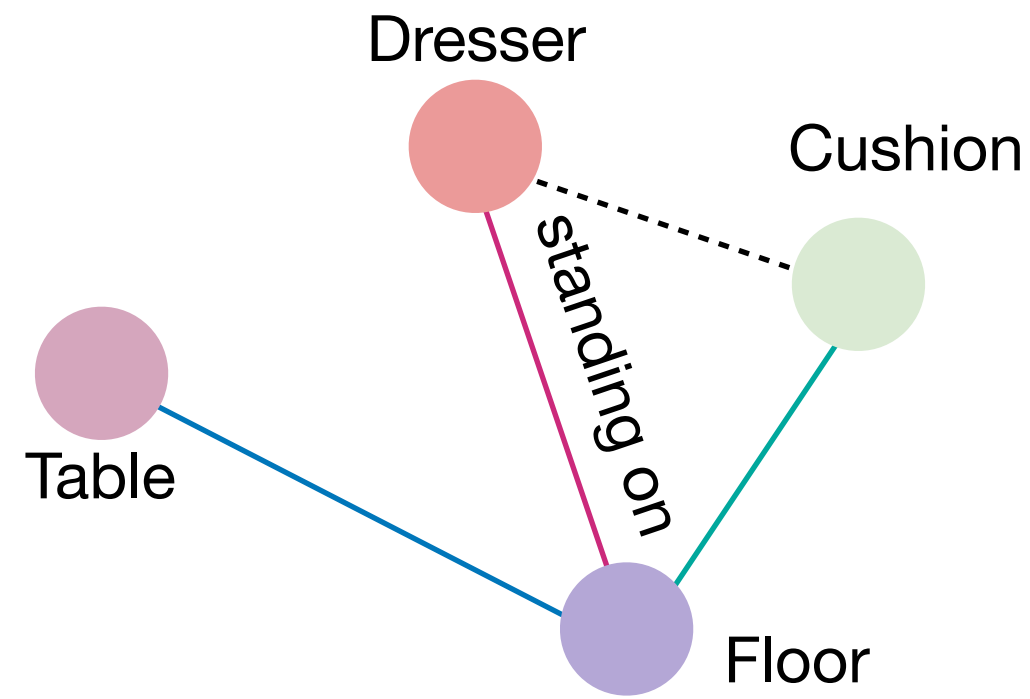
Feature Encoding



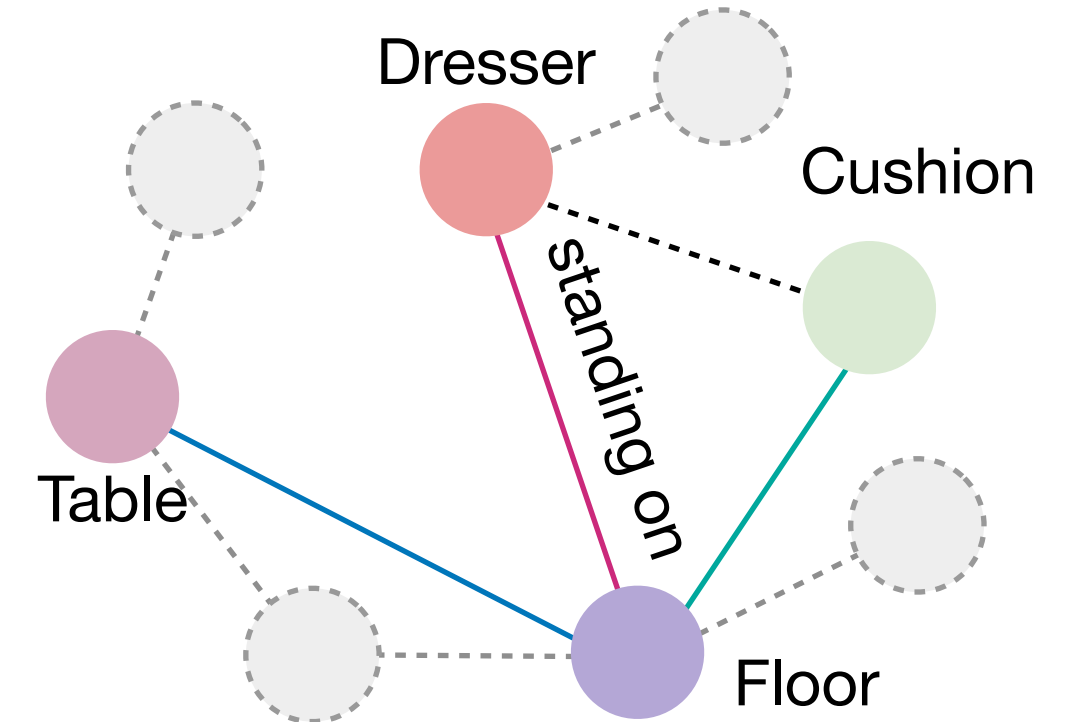
Message Passing

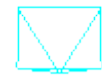


SSG Prediction



Fusion





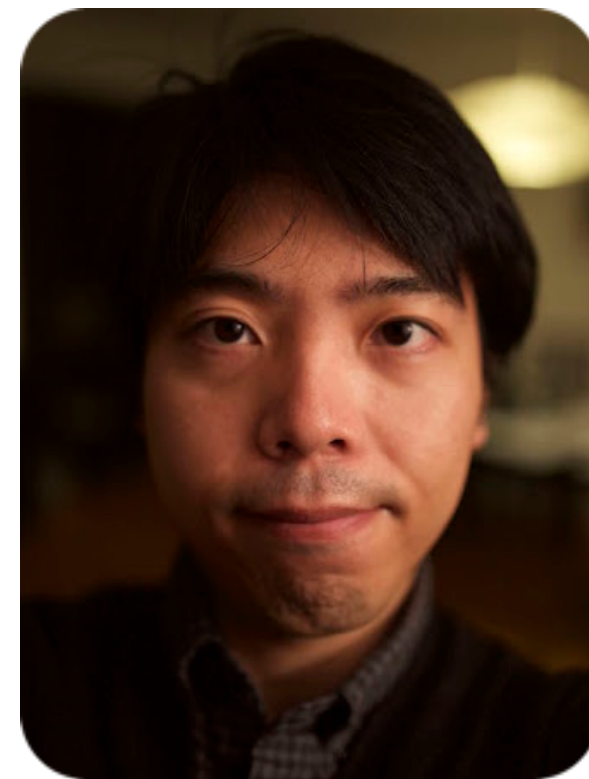
TUE-PM-088

ID: 5050

Thanks for watching.
Check out our poster and paper!



Shun-Cheng Wu



Keisuke Tateno



Nassir Navab



Federico Tombari