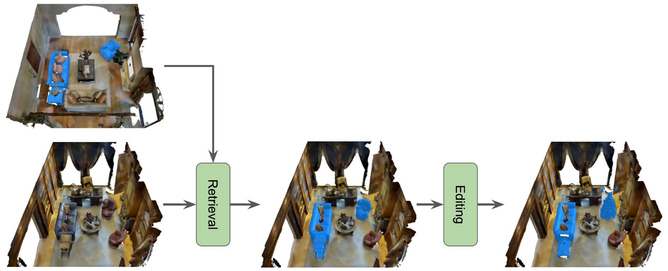
3DSSR: 3D Subscene Retrieval

Reza Asad, Manolis Savva

SFU AI | gruvi {graphics} U {vision} | SFU

Motivation

With increasing number of complex reconstructed 3D scenes, it is important to be able to retrieve 3D subscens of interest.

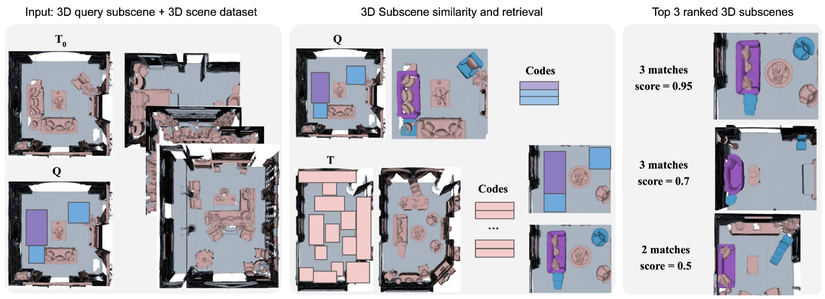


Our goal is to retrieve 3D subscenes with similar geometry and object arrangements.

Our contributions:

- We introduce the task of 3D subscene retrieval (3DSSR), generalizing prior work on 3D object and 3D scene retrieval.
- We tackle 3DSSR using our self-supervised point cloud encoder PointCrop that outperforms supervised and prior self-supervised encoders in subscene retrieval.

Task



Assumption: 3D scenes in the database come with axis-aligned bounding boxes but object categories are not provided.

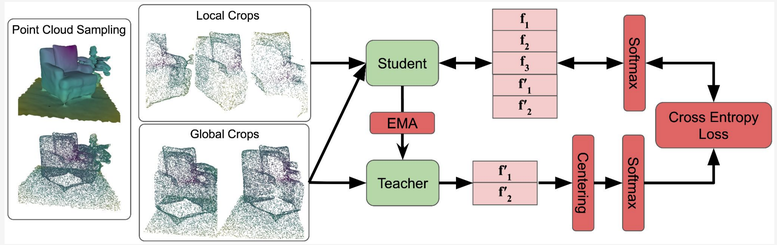
Intrinsic similarity: We compute the cosine similarity of codes derived from a trained encoder.

Extrinsic similarity: We measure the geometric overlap between the query and target boxes in the world coordinate.

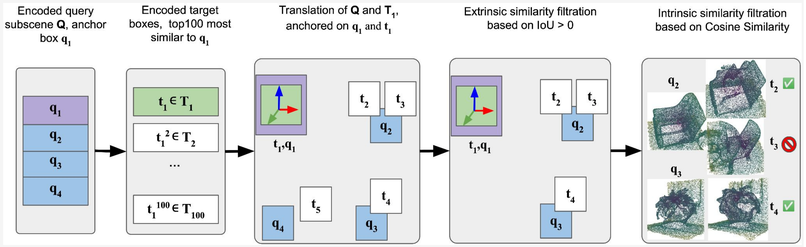
3D Subscene Retrieval Using PointCrop

Intrinsic similarity through self-distillation:

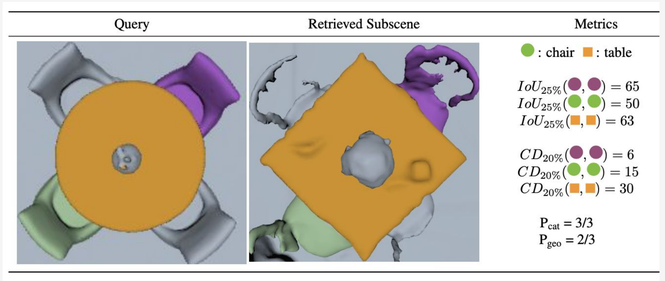
- All crops are fed to the student, while the teacher only takes the global crops.
- The student and the teacher have the same Point Transformer architecture.
- We apply centering on the output of the teacher to avoid collapse.
- The parameters of the teacher are an exponential moving average of the student's parameters (no SGD)



Extrinsic similarity:



Evaluation metric:



Quantitative Results

Method	mAP_{cat}	mAP_{geo}	AUC (mAP_{cat})	AUC (mAP_{geo})
OracleCatRank	19.69	7.99	9.97	4.27
OracleCatRank[+IoU]	**27.73**	9.16	**13.12**	5.08
OracleGKRank	17.22	5.69	8.70	2.76
TransformerRank	16.61	**20.88**	8.87	**10.86**
TransformerRankDisc	16.51	6.20	7.54	2.86
RandomRank	0.68	1.45	0.36	0.64
CSCRank	12.65	16.10	6.98	8.36
PointCropRank	**14.94**	**25.21**	**8.06**	**13.14**

Method	$mAP_{cat\&geo}$	AUC ($mAP_{cat\&geo}$)
OracleCatRank[+IoU]	9.16	5.08
TransformerRank	9.80	**5.49**
CSCRank	7.91	4.45
PointCropRank	**9.91**	**5.49**
CatRankV2[+IoU]	2.54	1.42
TransformerRankV2	**6.71**	**3.64**
CSCRankV2	4.46	2.42
PointCropRankV2	5.87	3.40

Using predicted 3D bounding boxes at test time (Right table bottom group)

Method	AUC[dist+CD]	AUC[angle+CD]	AUC[dist+angle+CD]	AUC[dist+angle+Cat+CD]
OracleCatRank	5.88	5.75	4.88	4.88
OracleCatRank[+IoU]	8.58	8.99	8.07	8.07
OracleGKRank	5.52	5.22	4.03	4.03
TransformerRank	16.14	16.03	14.88	9.46
RandomRank	2.49	1.87	1.60	0.16
CSCRank	14.73	14.71	13.32	8.52
PointCropRank	**22.53**	**23.05**	**20.98**	**11.52**

Softer evaluation metrics for 3DSSR

Method	AUC[dist+CD]	AUC[angle+CD]	AUC[dist+angle+CD]	AUC[dist+angle+Cat+CD]
OracleCatRank[+IoU]	8.68	9.09	8.09	8.09
TransformerRank	15.49	15.40	14.28	8.81
CSCRank	15.26	15.40	13.84	8.76
PointCropRank	**21.01**	**21.37**	**19.37**	**10.48**

Evaluating with a rotation module

Qualitative Results