SeeThrough: Finding Chairs in Heavily Occluded Indoor Scene Images

Moos Hueting
University College London

Pradyumna Reddy
University College London

Ersin Yumer
Adobe Research

Vladimir G.Kim
Adobe Research

Nathan Carr
Adobe Research

Niloy J.Mitra
University College London

Goal: extract 3D scene mock up from single image (focused on chairs and other highly occluded objects)

Context is Important

Global context

Local context

No context

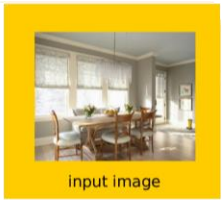
Context is Important

Global context

Local context

No context

Pipeline



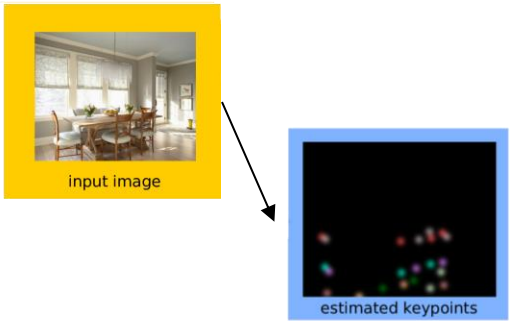
input

stage

intermediate result

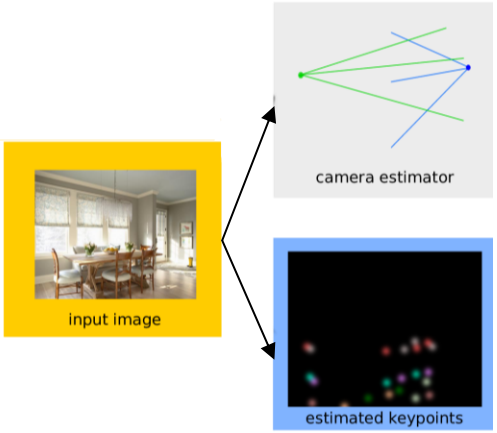
output

Pipeline



input
stage
intermediate result
output

Pipeline



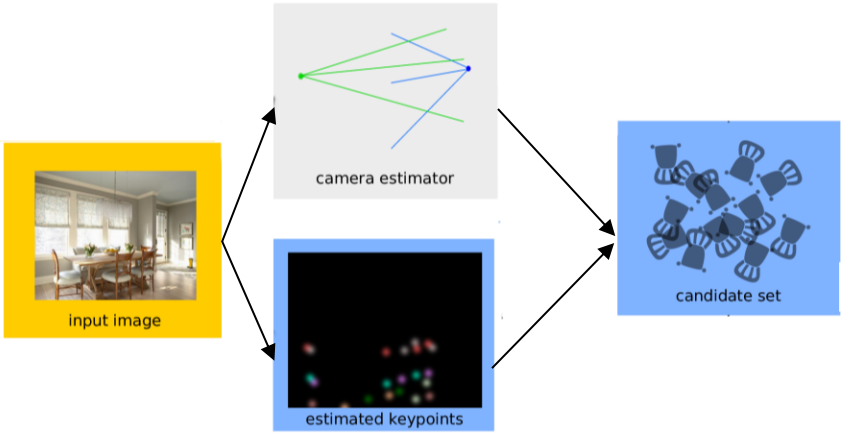
input

stage

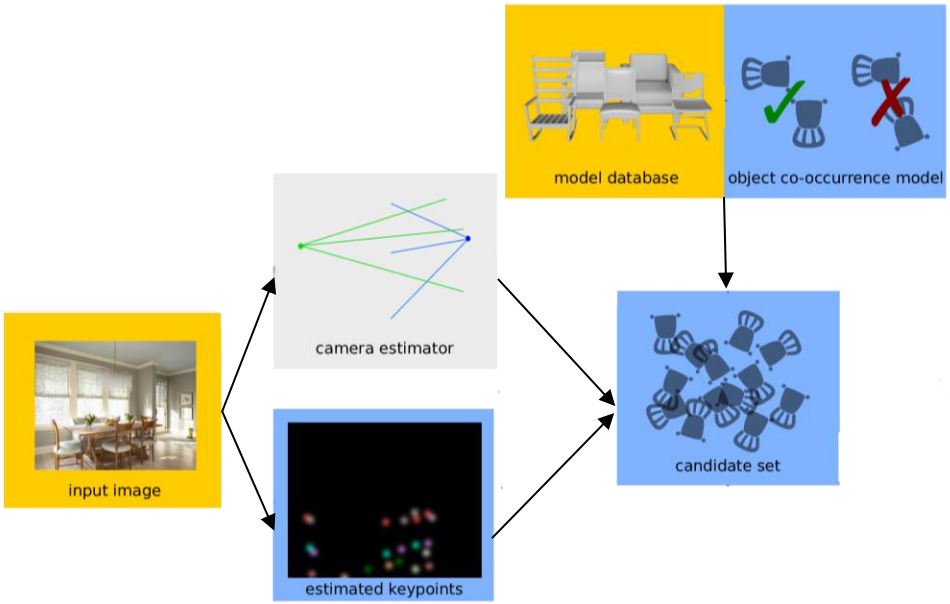
intermediate result

output

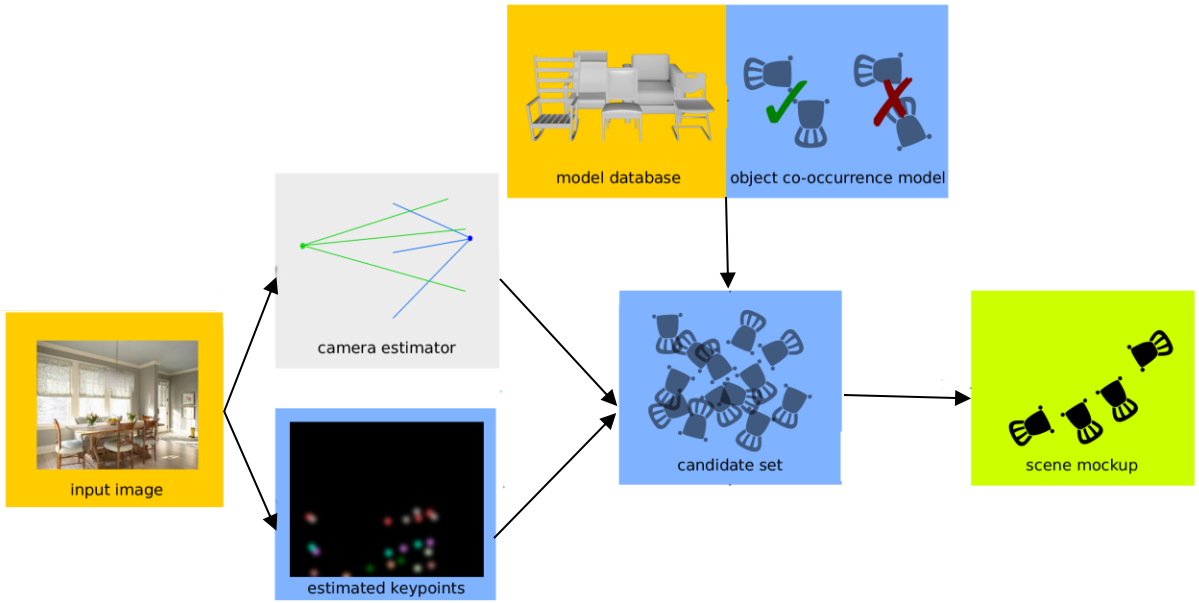
Pipeline


camera estimator
input image
estimated keypoints
candidate set
input
stage
intermediate result
output

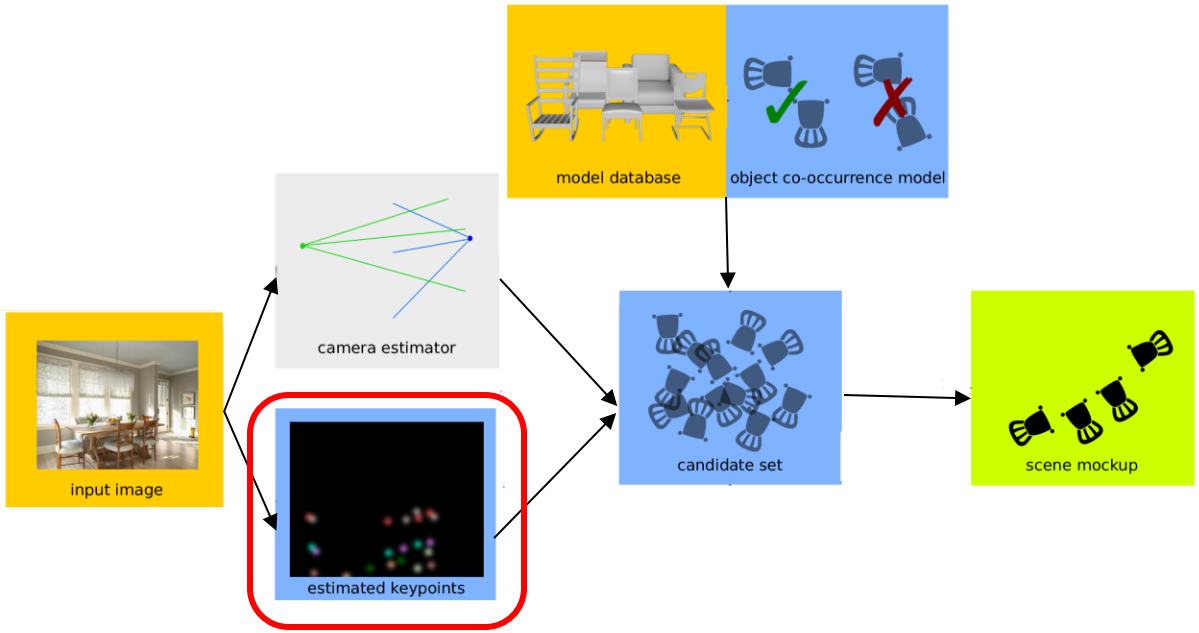
Pipeline


model database
object co-occurrence model
camera estimator
input image
estimated keypoints
candidate set
input
stage
intermediate result
output

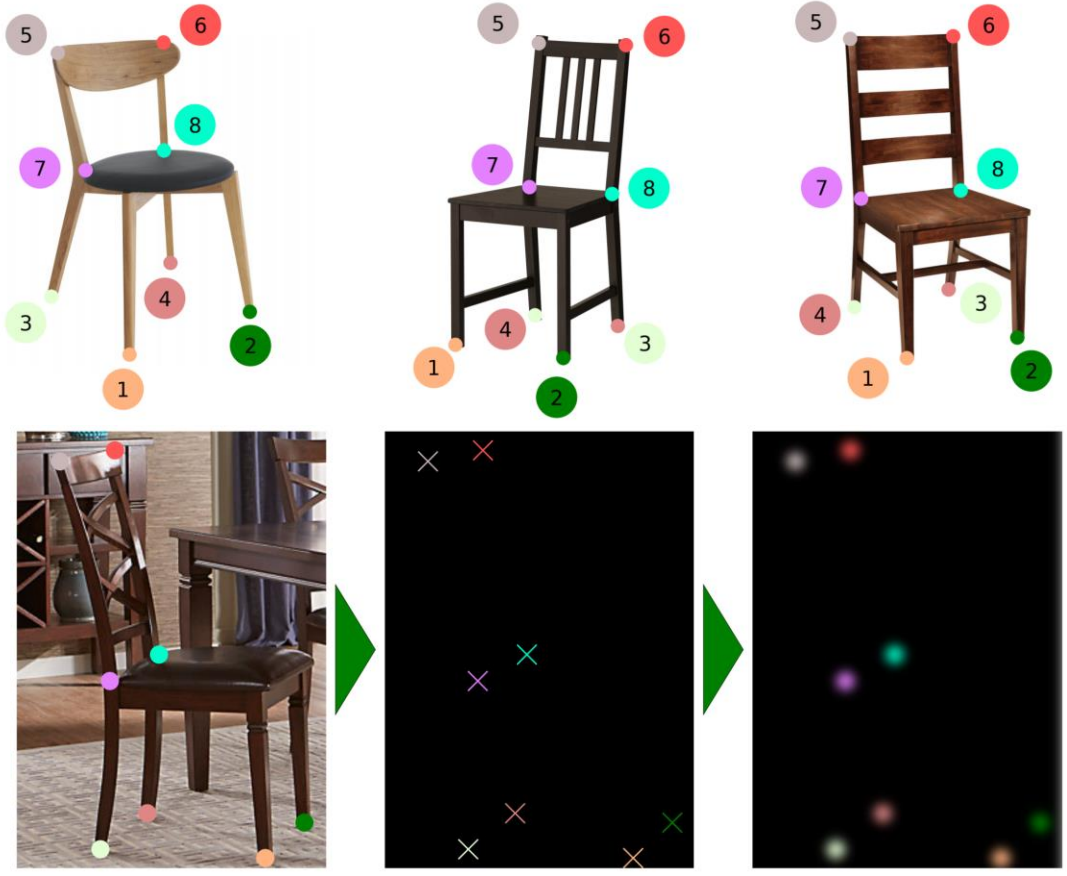
Pipeline


model database
object co-occurrence model
camera estimator
input image
estimated keypoints
candidate set
scene mockup
input
stage
intermediate result
output

Pipeline


model database
object co-occurrence model
camera estimator
input image
estimated keypoints
candidate set
scene mockup
input
stage
intermediate result
output

Keypoint estimation



Keypoint Dataset

Input image

Objectnet3D Ground truth annotation

Selecting Vertices of the overlaid CAD model

cvgl.stanford.edu/projects/objectnet3d/

Keypoint thresholding

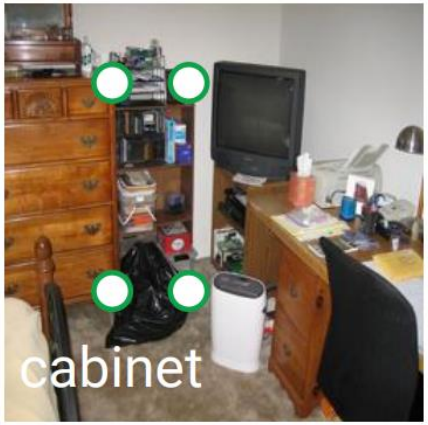
Ground truth keypoints

Network output

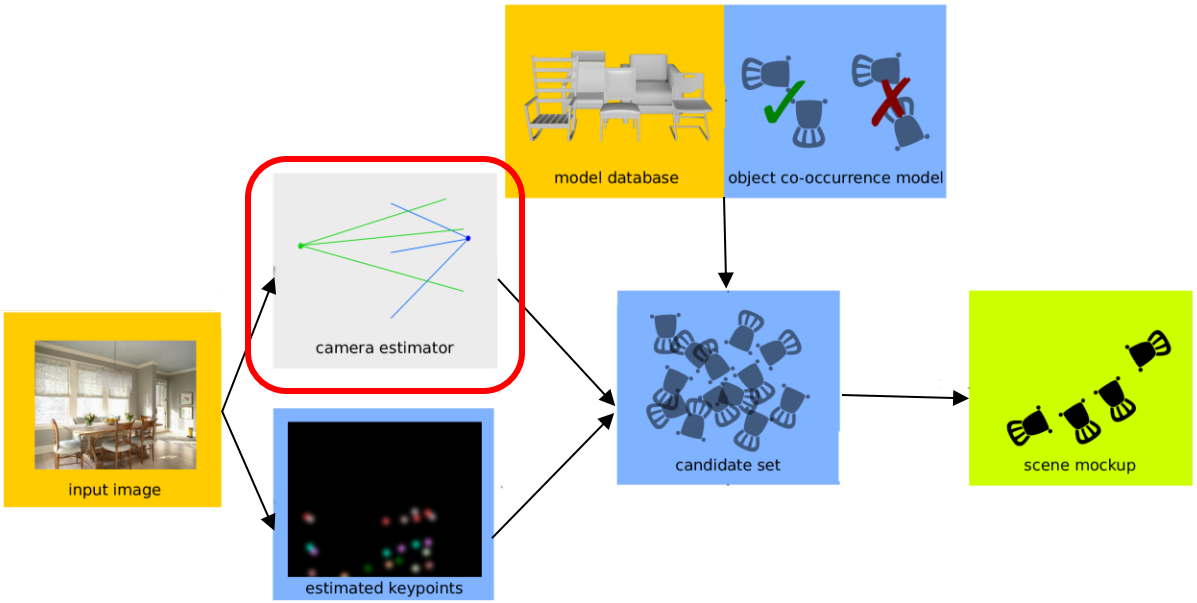
After thresholding

Local maxima

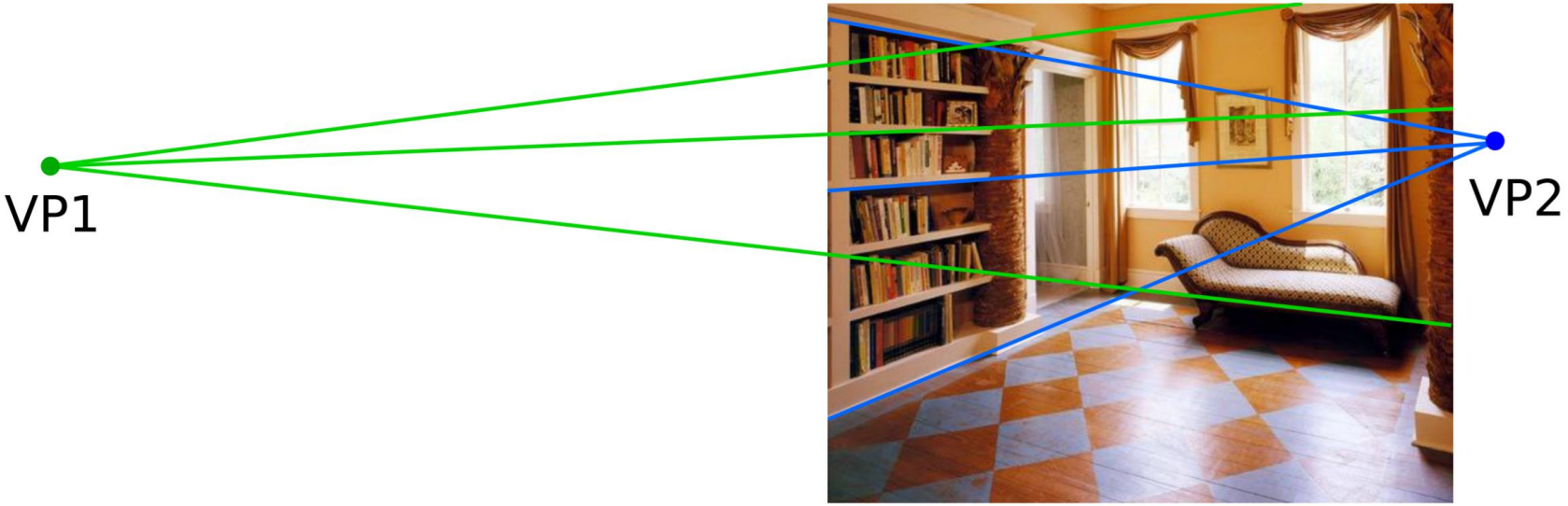
Keypoint thresholding



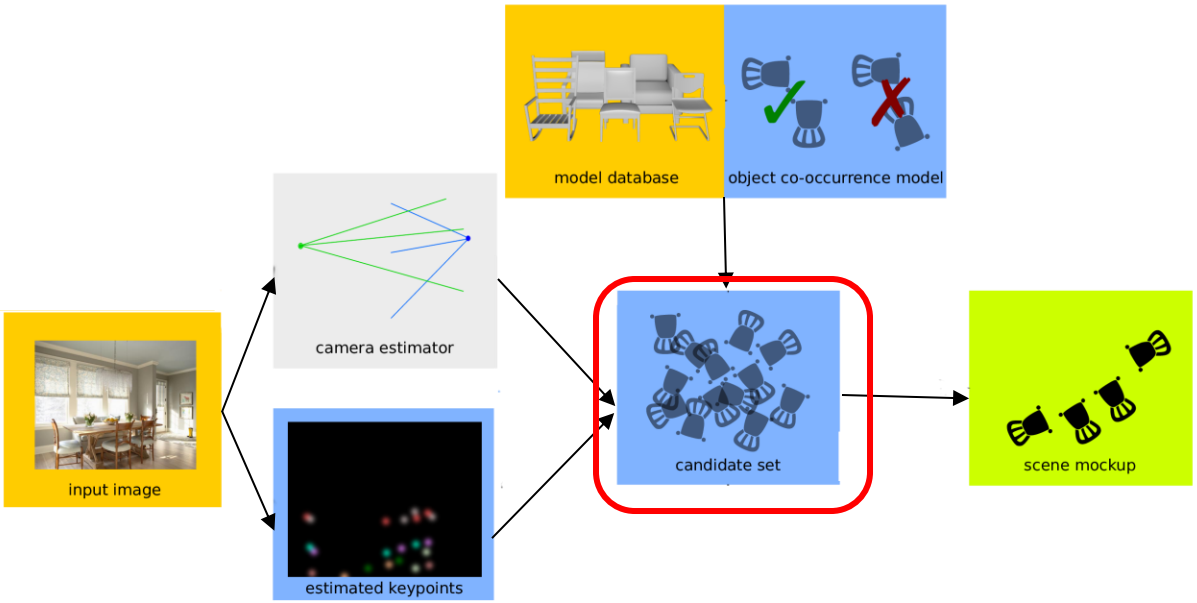
Pipeline


model database
object co-occurrence model
camera estimator
input image
estimated keypoints
candidate set
scene mockup
input
stage
intermediate result
output

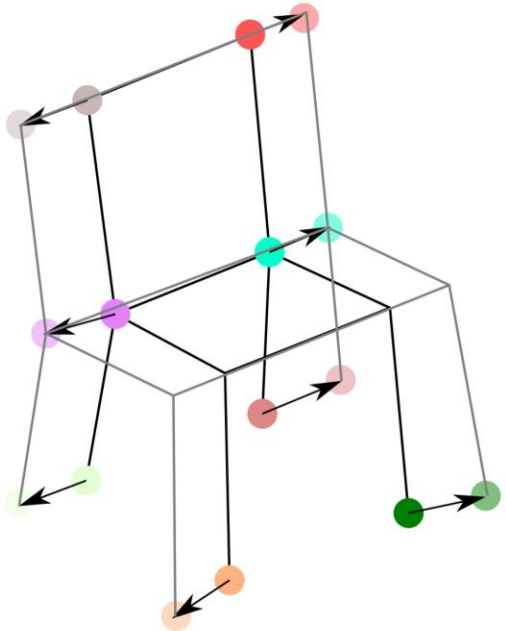
Vanishing point estimation



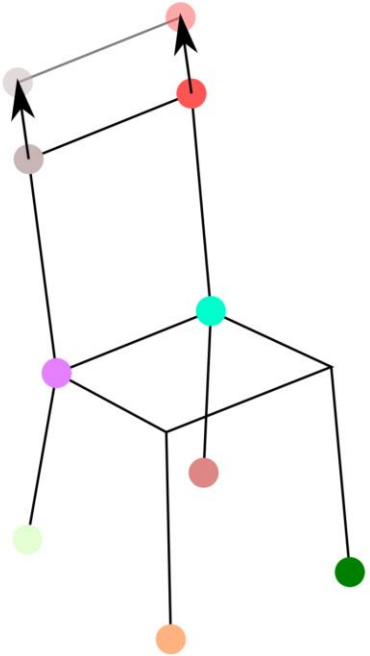
Pipeline


model database
object co-occurrence model
camera estimator
input image
estimated keypoints
candidate set
scene mockup
input
stage
intermediate result
output

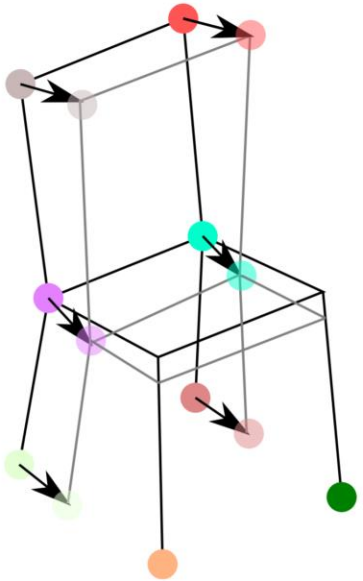
PCA template



PCA dimension 1

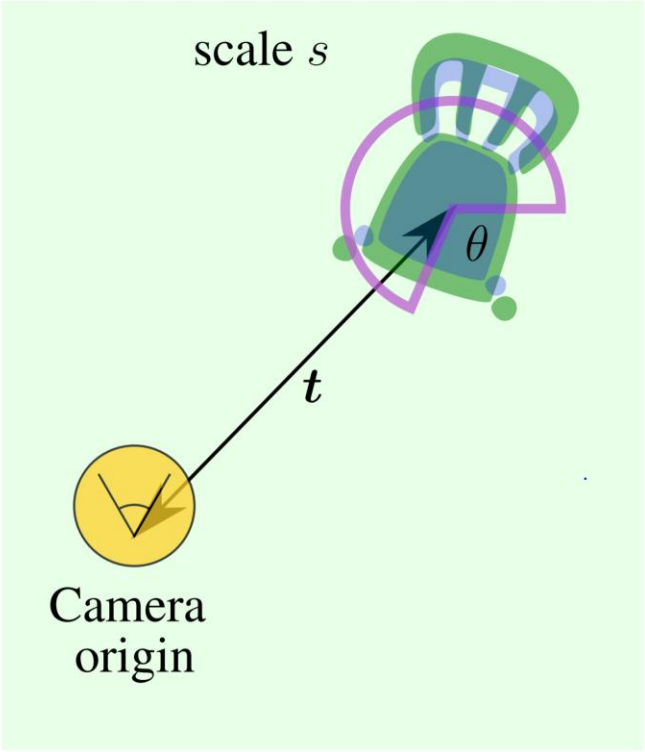


PCA dimension 2

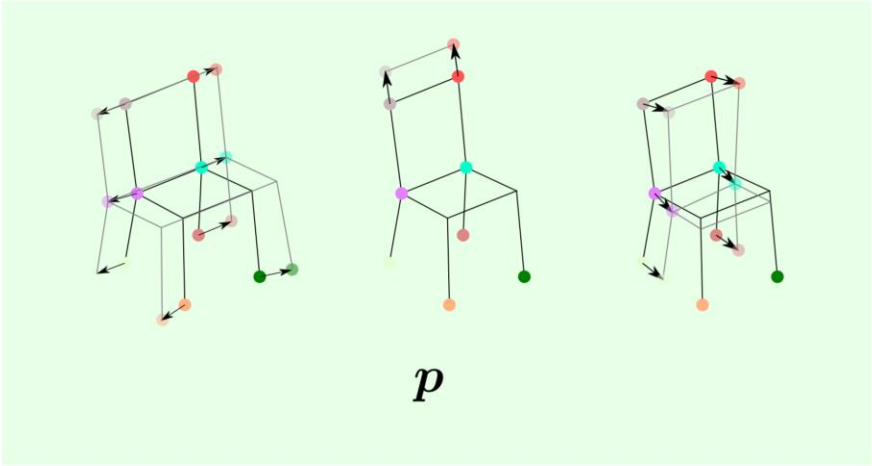


PCA dimension 3

Fit parameters

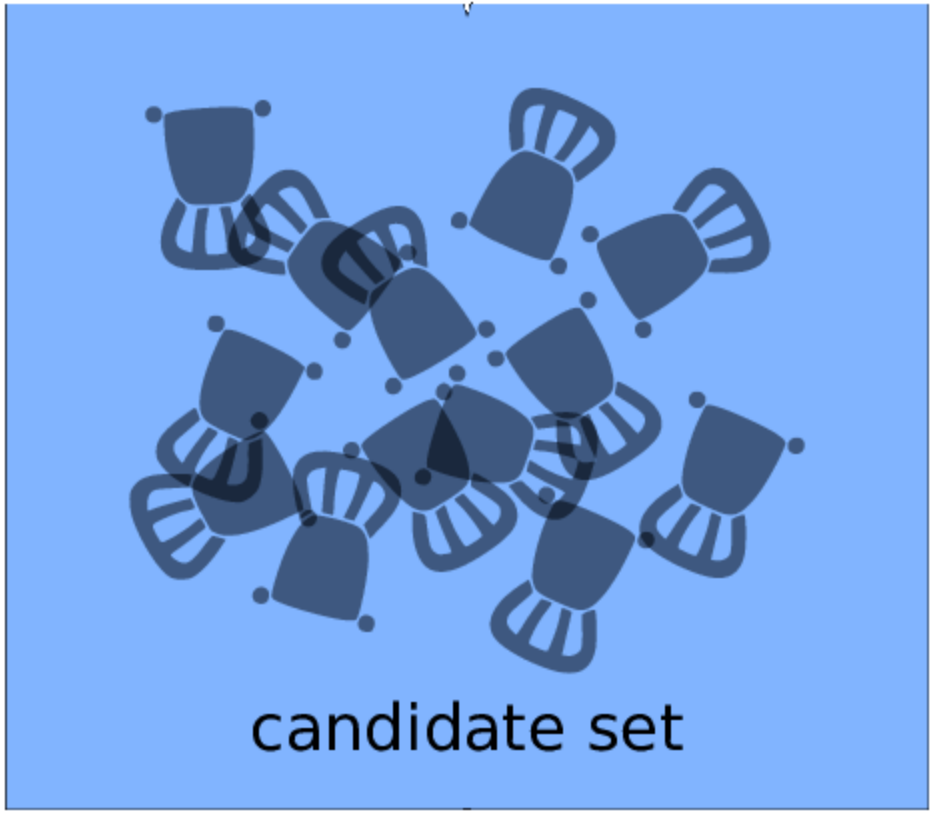


Transform parameters

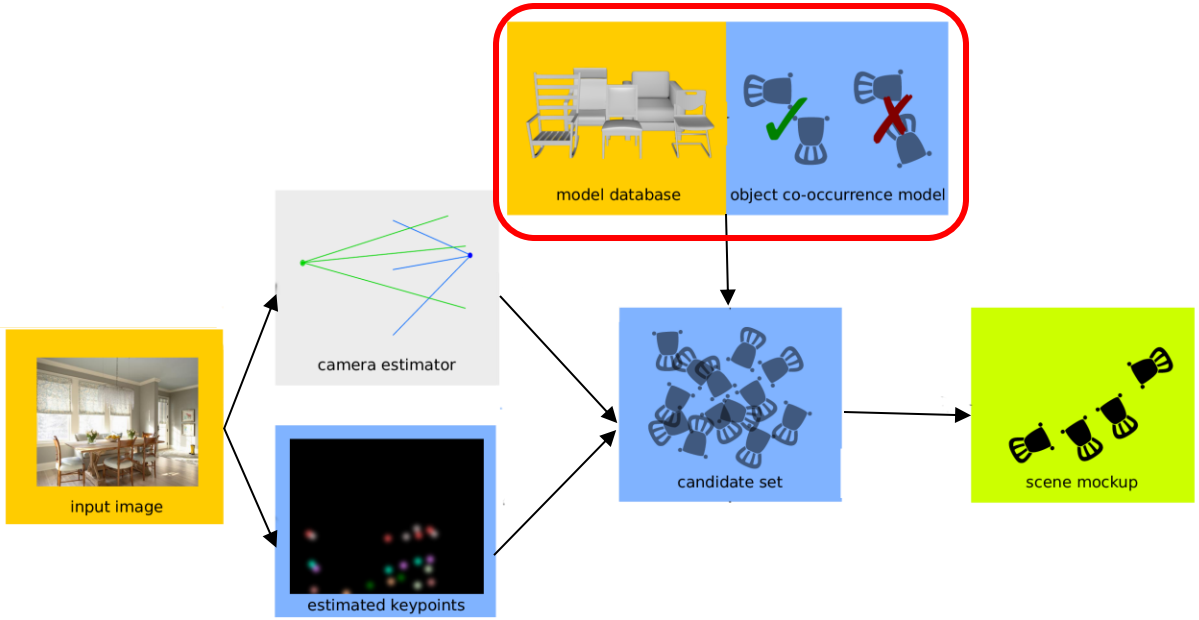


Template parameters

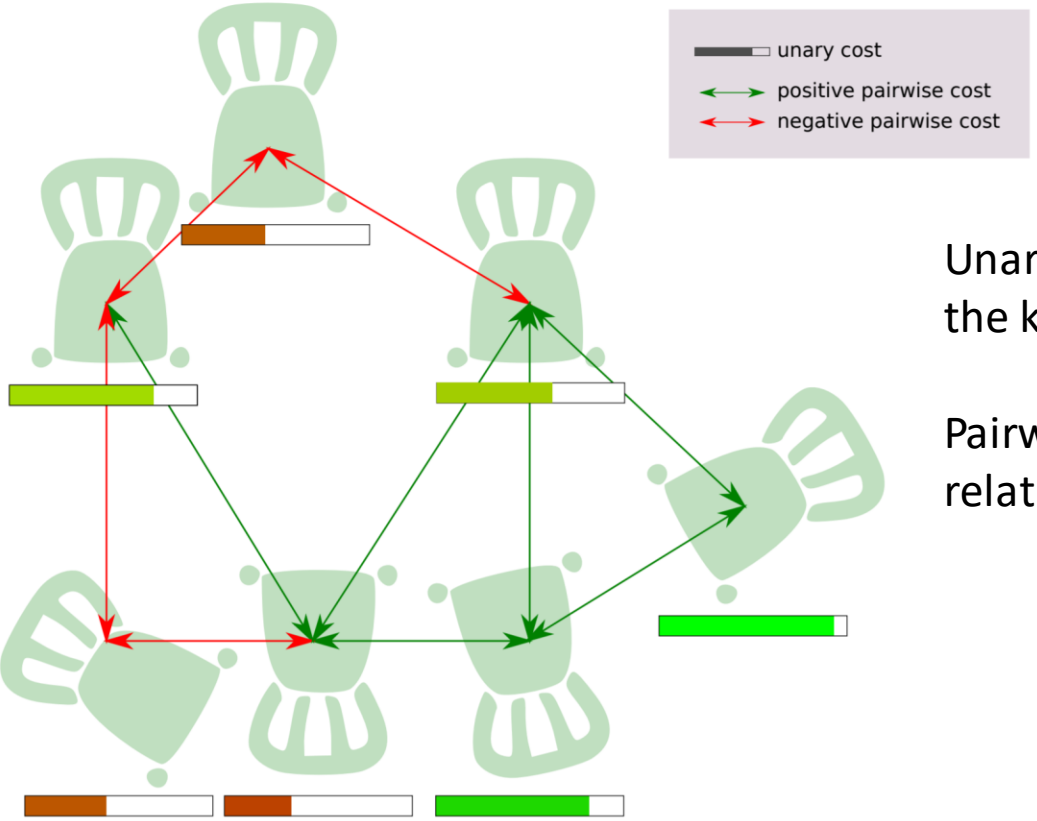
Candidate Set



Pipeline


model database
object co-occurrence model
camera estimator
input image
estimated keypoints
candidate set
scene mockup
input
stage
intermediate result
output

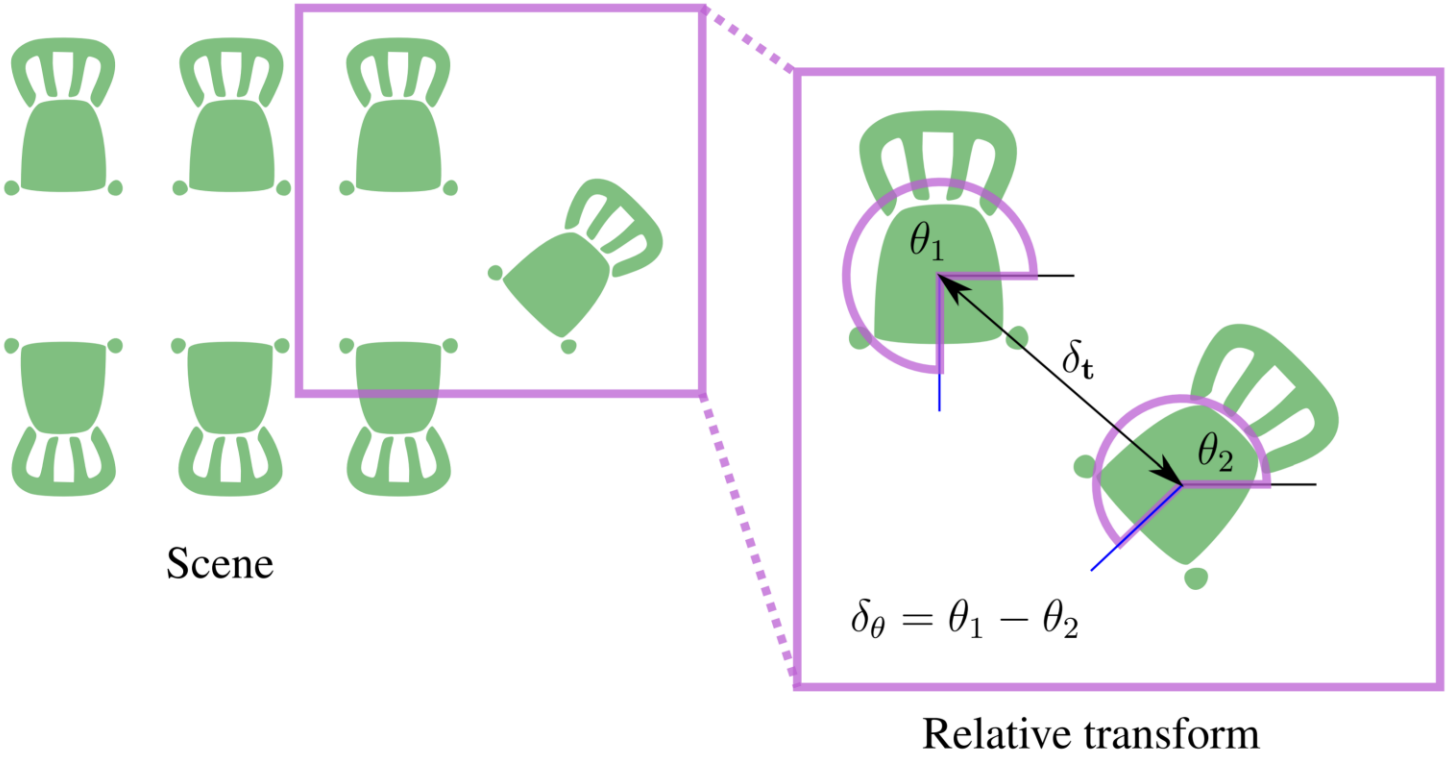
Candidate selection



Unary Costs: measure how well the key points explain the object

Pairwise Costs: Capture relationship between objects

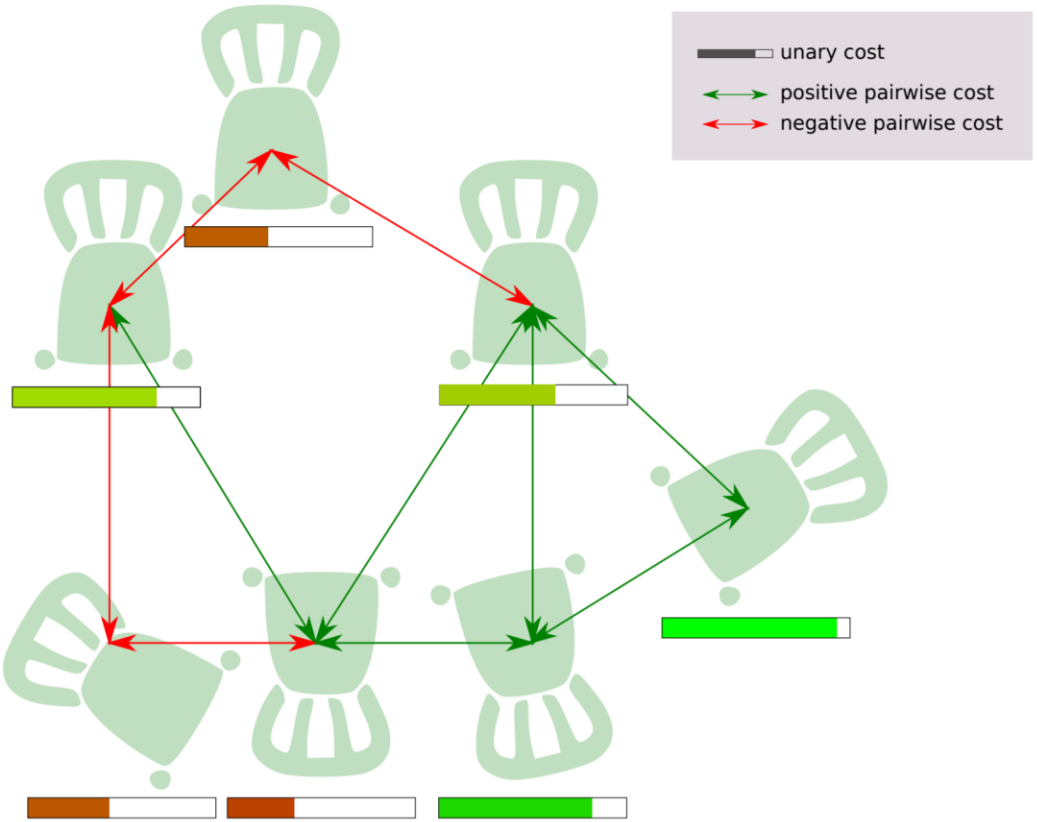
Relative transform



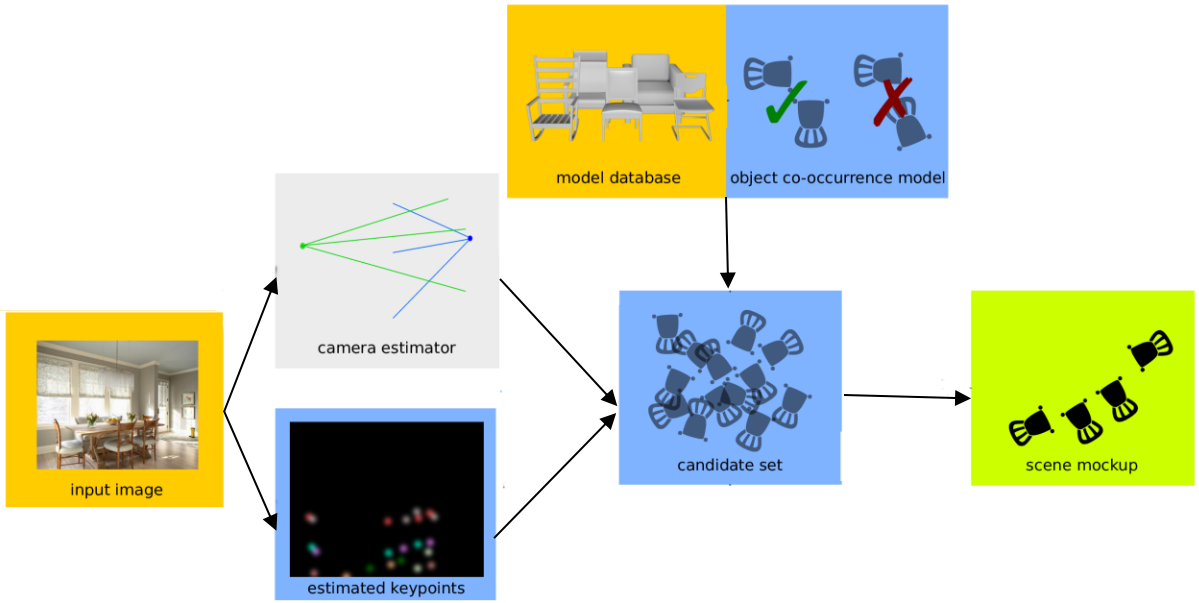
δ_θ Relative Rotation

δ_t Relative Translation

Candidate selection



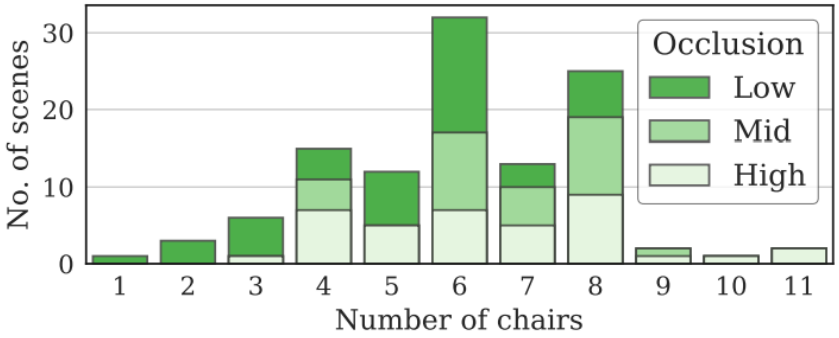
Pipeline

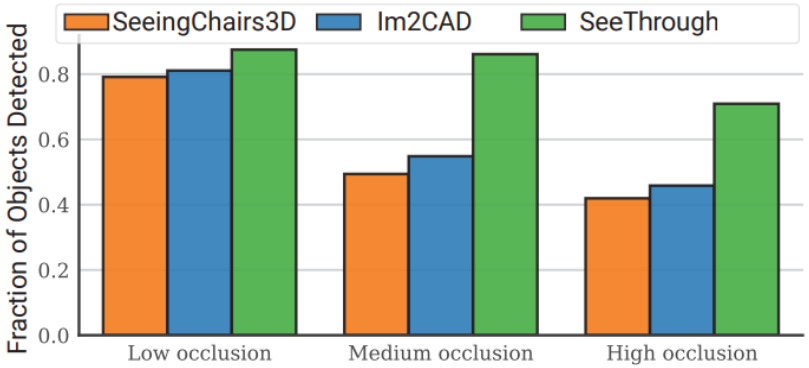

model database
object co-occurrence model
camera estimator
input image
estimated keypoints
candidate set
scene mockup
input
stage
intermediate result
output

Results

http://geometry.cs.ucl.ac.uk/projects/2018/seethrough/

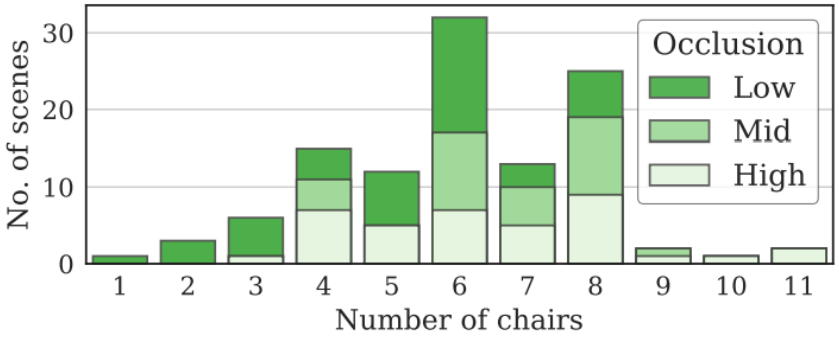
Results and Dataset

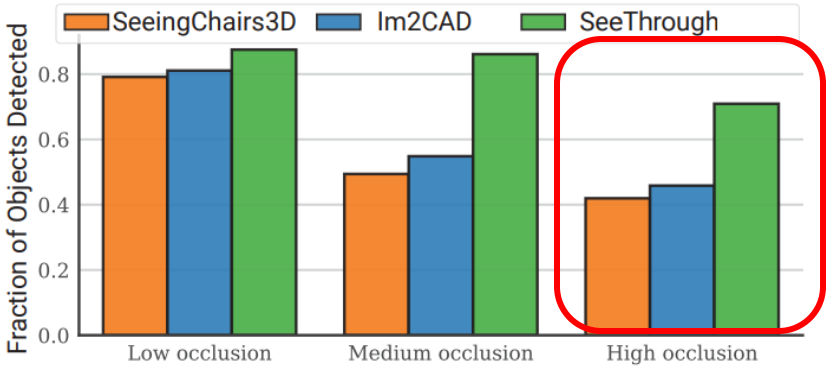




http://geometry.cs.ucl.ac.uk/projects/2018/seethrough/

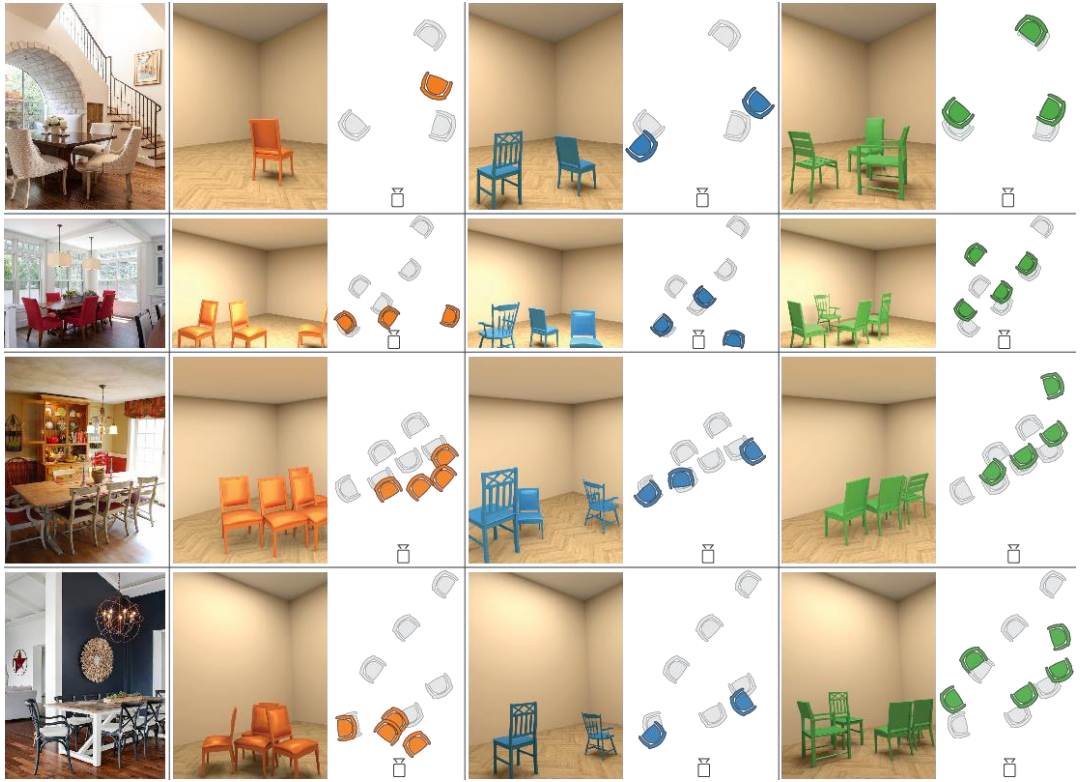
Results and Dataset





http://geometry.cs.ucl.ac.uk/projects/2018/seethrough/

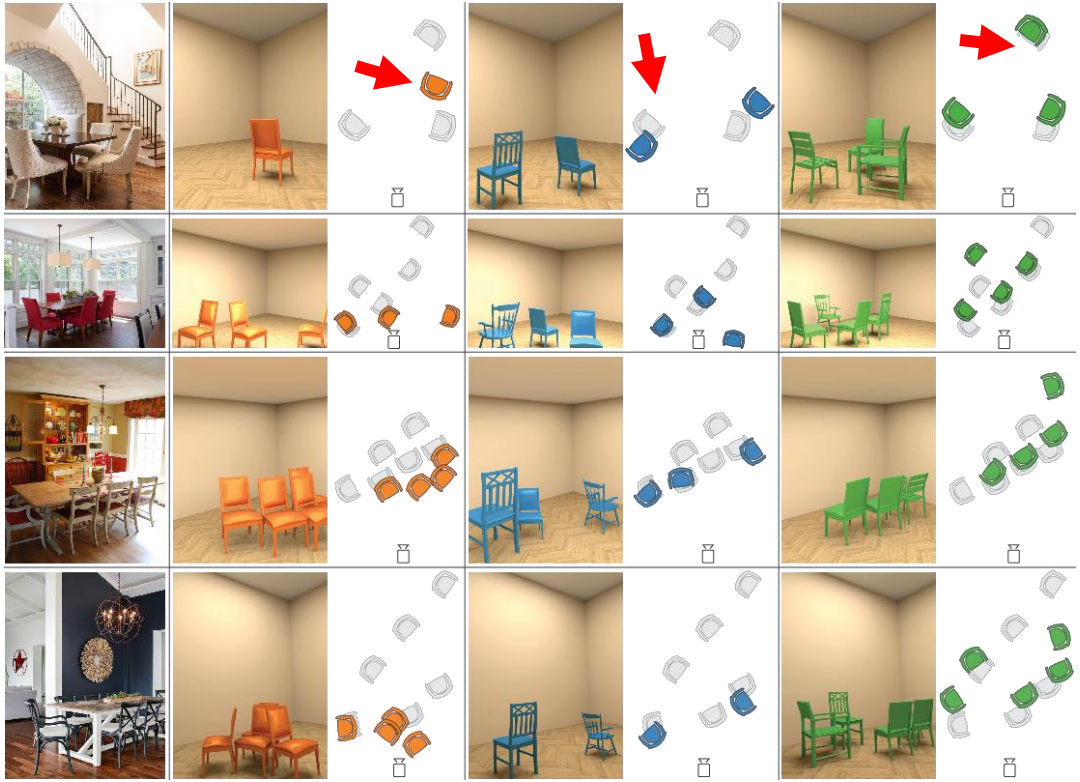
Results



Real World Images | SeeingChairs | Im2CAD | Ours

http://geometry.cs.ucl.ac.uk/projects/2018/seethrough/

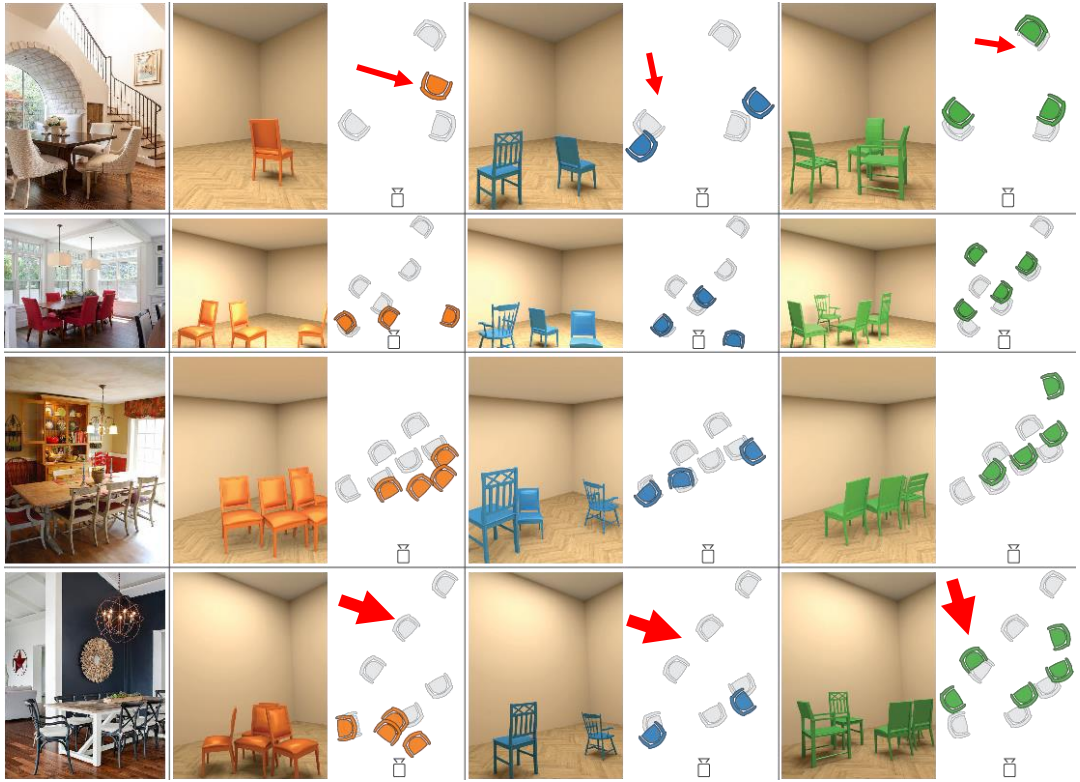
Results



Real World Images | SeeingChairs | Im2CAD | Ours

http://geometry.cs.ucl.ac.uk/projects/2018/seethrough/

Results



Real World Images | SeeingChairs | Im2CAD | Ours

http://geometry.cs.ucl.ac.uk/projects/2018/seethrough/

Performance comparison

	AvgMaxIoU (precision)	AvgMaxIoU (recall)	AvgMaxIoU (F1)	
3D-INN [128] + FasterRCNN [94]	0.316	0.150	0.198	
SeeingChairs [6]	0.195	0.128	0.149	
Ours	**0.386**	**0.250**	**0.293**	
	PercCorrect (precision)	PercCorrect (recall)	PercCorrect (F1)	
3D-INN [128] + FasterRCNN [94]	0.263	0.124	0.165	
SeeingChairs [6]	0.071	0.043	0.052	
Ours	**0.298**	**0.167**	**0.207**	
	PercCorrectFull (precision)	PercCorrectFull (recall)	PercCorrectFull (F1)	
3D-INN [128] + FasterRCNN [94]	0.04	0.015	0.021	
SeeingChairs [6]	0.013	0.007	0.009	
Ours	**0.285**	**0.161**	**0.198**	
	AvgMax2DIoU (precision)	AvgMax2DIoU (recall)	AvgMax2DIoU (F1)	AngleDiff (in degrees)
3D-INN [128] + FasterRCNN [94]	0.526	0.336	0.401	55.8
SeeingChairs [6]	0.372	0.325	0.341	11.4
Ours	**0.628**	**0.470**	**0.525**	**7.3**

Goal: extract 3D scene mock up from single image (focused on chairs and other highly occluded objects)

Main insight: cases with significant occlusion can be improved by using high-level contextual knowledge about how scenes “work”

Main result: resulting scene mock ups significantly better than combinations of state-of-the-art methods which are reliant on object detection algorithms.

Limitations

- First, we plan to extend the evaluation to more classes of objects beyond those considered.
- Second, one can explore higher fidelity models to better recover fine scale features in the recovered models.
- Finally, we would like to explore templates that can express a broader understanding of the multi-object spatial relationships including symmetry and regularity.

Acknowledgement

This work is in part supported by the Microsoft PhD fellowship program, and ERC Starting Grant SmartGeometry (StG-2013-335373). Also, special thanks to Aron Monszpart, James Hennessey, Carlo Innamorati, Paul Guerrero, and other group members for invaluable help at various stages of the project.

Thank You

Code available:
geometry.cs.ucl.ac.uk/projects/2018/seethrough/paper_docs/Code_Data.zip