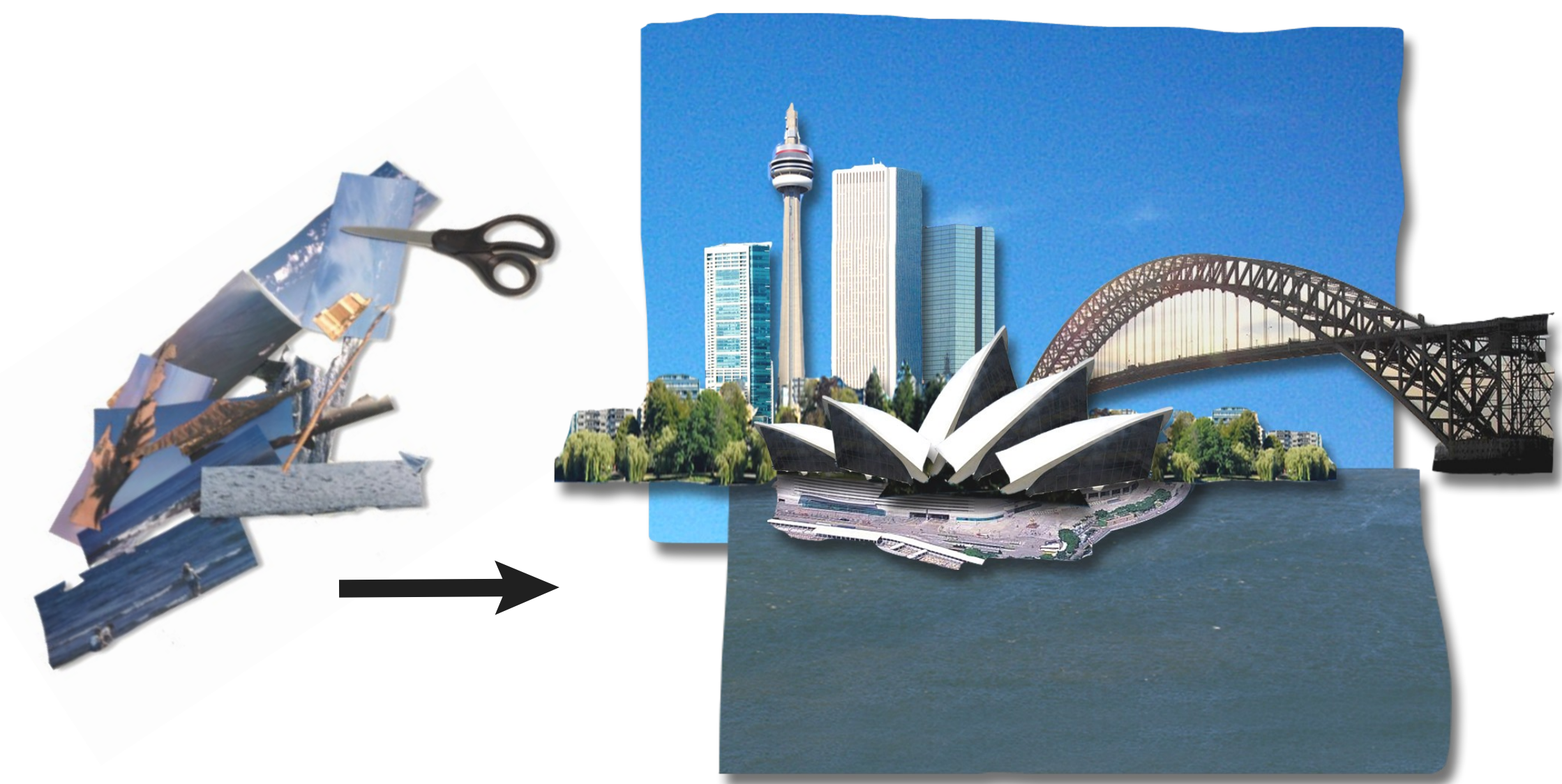


Motivation



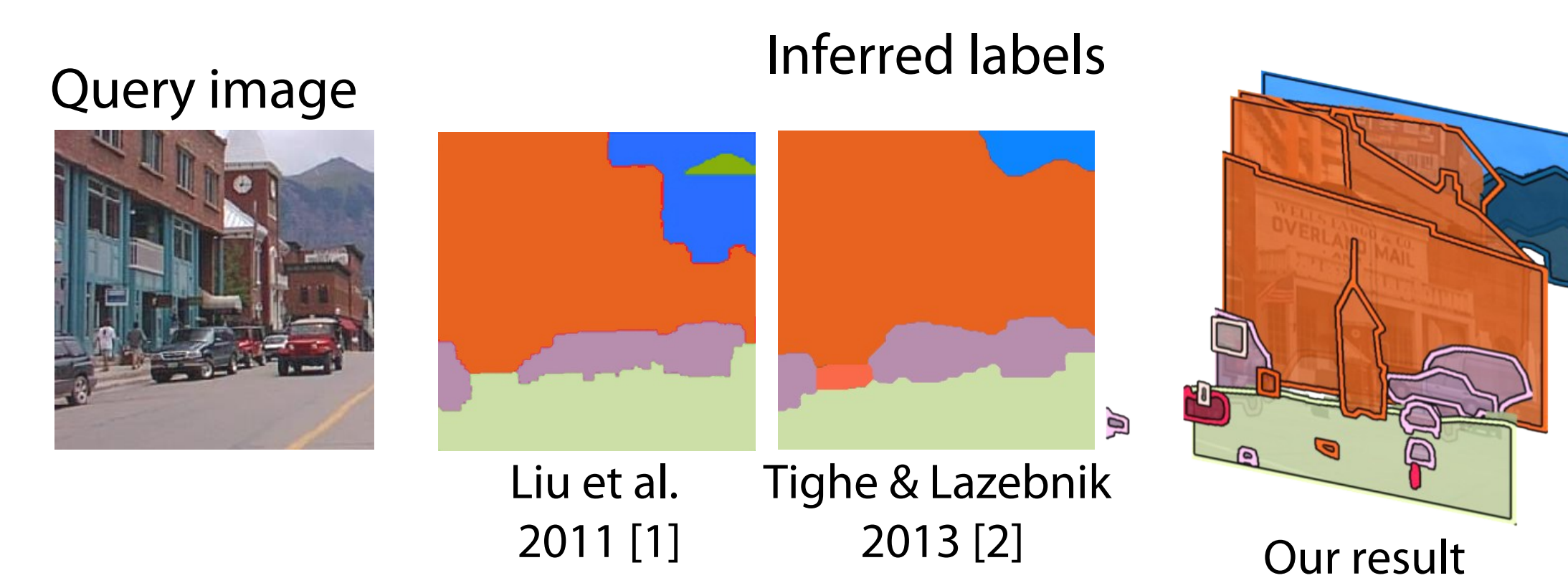
An artist can quickly synthesize a scene by collaging it together from found pieces.

Many current models represent a scene as a 2D grid of pixel labels:

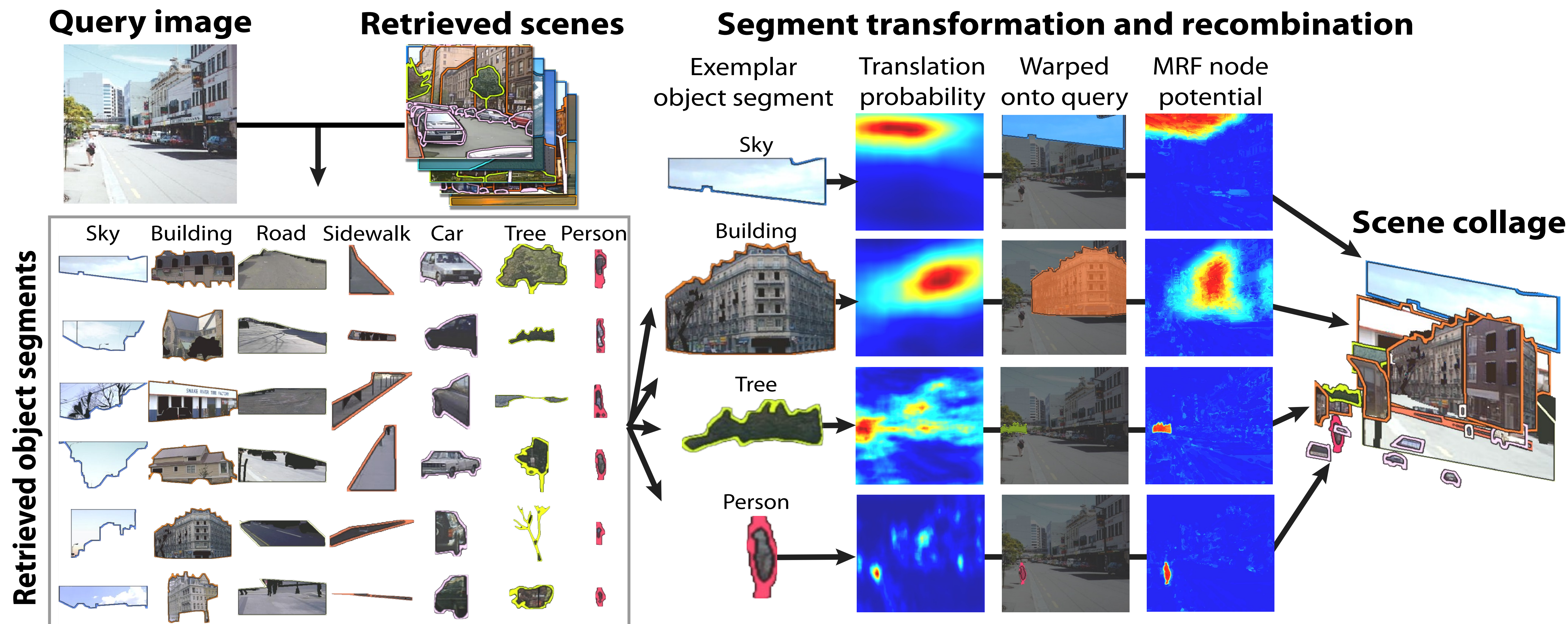
- Lack individuated objects
- No depth information
- Do not model occlusion

To address these issues, we introduce **scene collages**:

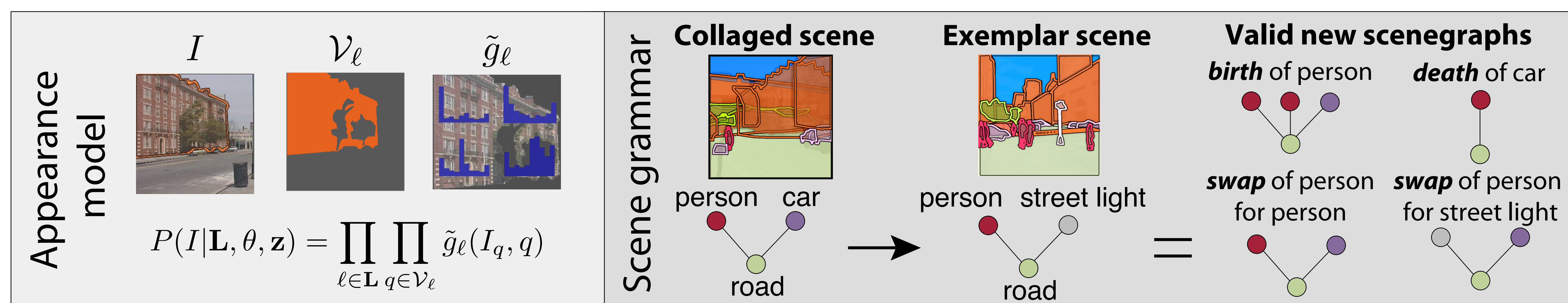
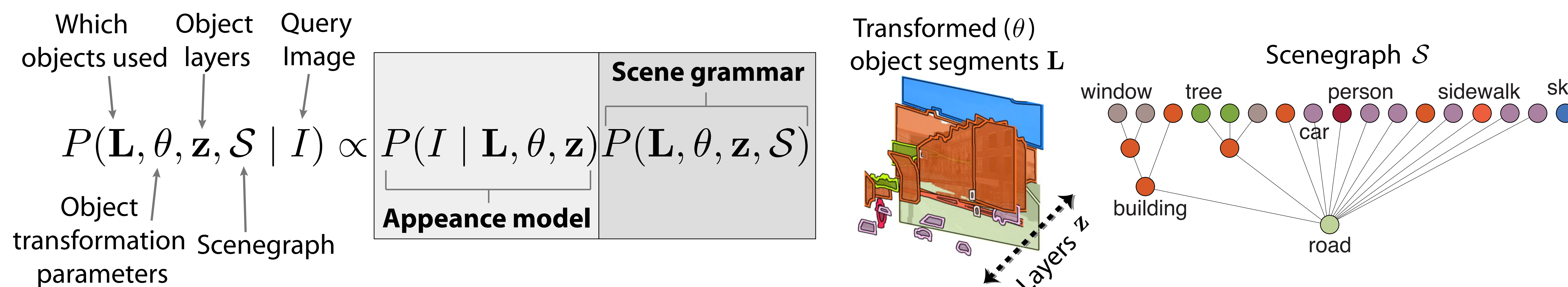
- Individuated **objects**
- **Layer-world** depth
- Represents **occlusion**
- **Generative**
- Can synthesize scenes
- Parse via **Analysis-by-synthesis**



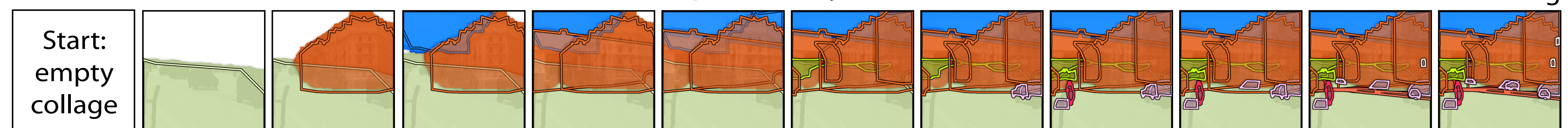
Method



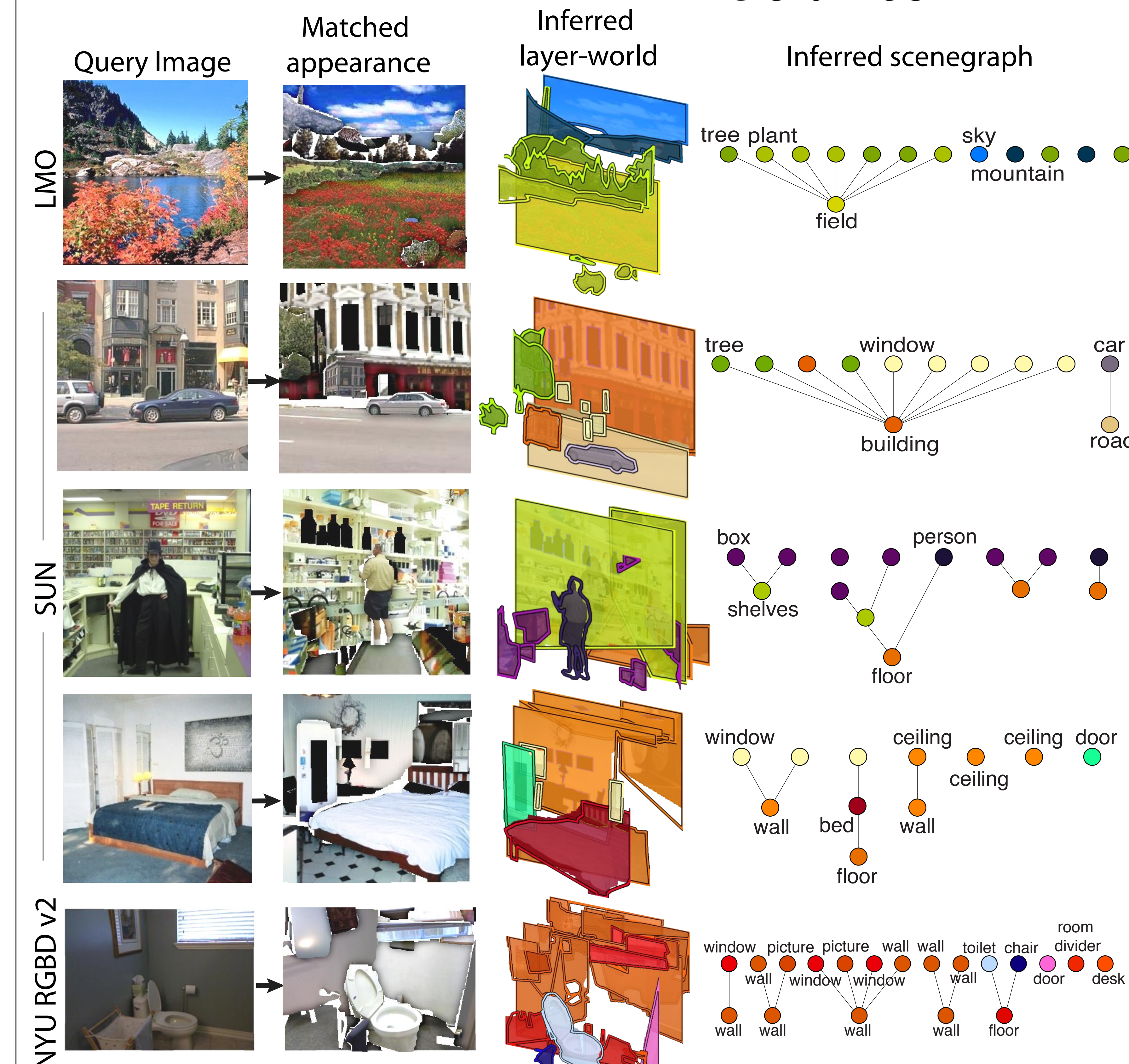
Scene model



Inference – greedy composition



Results



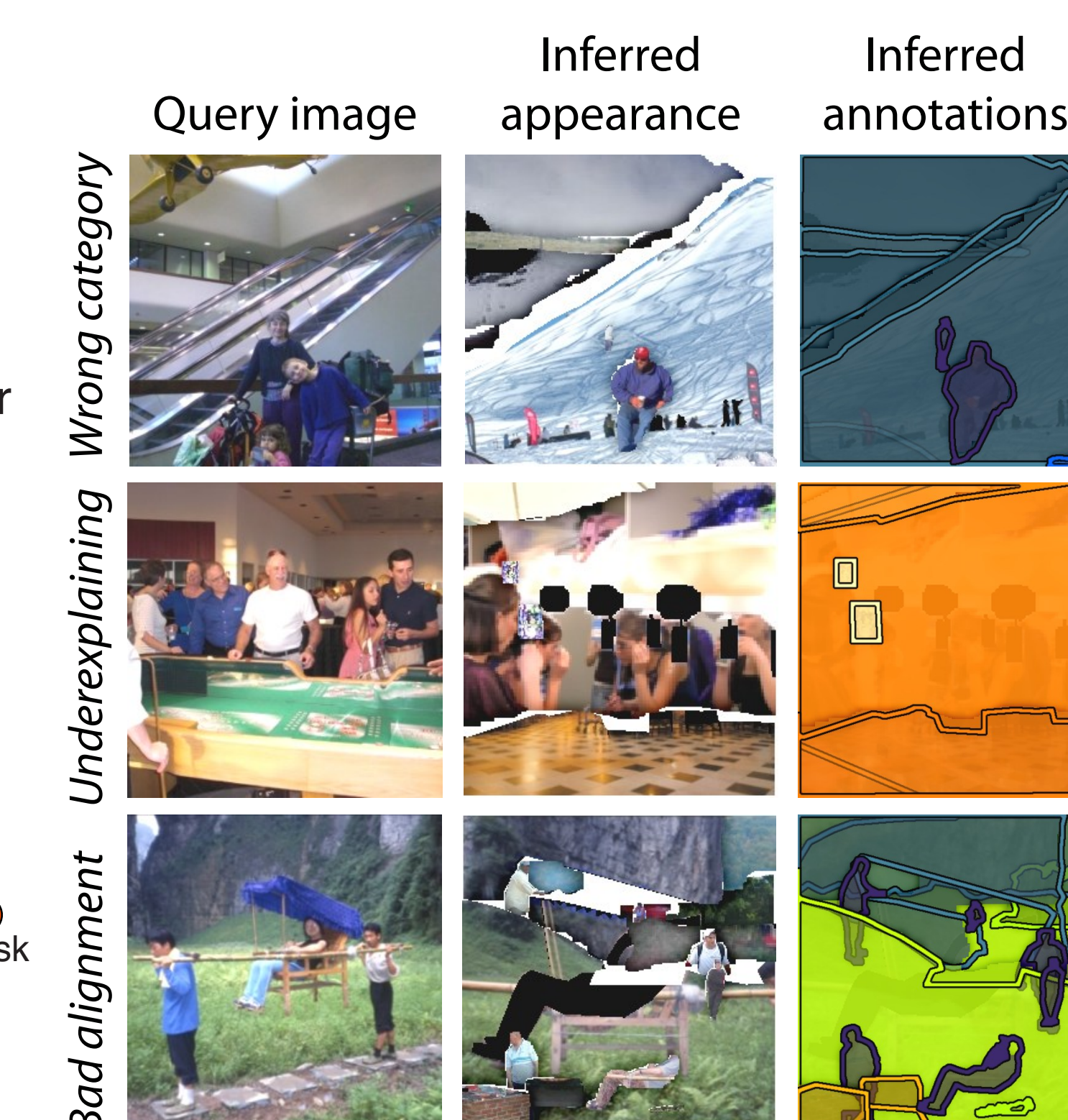
Performance of our algorithm

Dataset	Mean per-pixel accuracy	Mean per-class accuracy
LMO	0.70	0.26
SUN	0.40	0.03
NYU RGBD v2	0.30	0.01

Comparison with other methods on LMO

Algorithm	Mean per-pixel accuracy	Mean per-class accuracy
Liu et al. [1]	76.7	N/A
Tighe and Lazebnik [2]	78.6	39.2
Fabaret et al. [3]	78.5/74.2	29.6/46.0

Common failure cases



Applications

Random scene synthesis

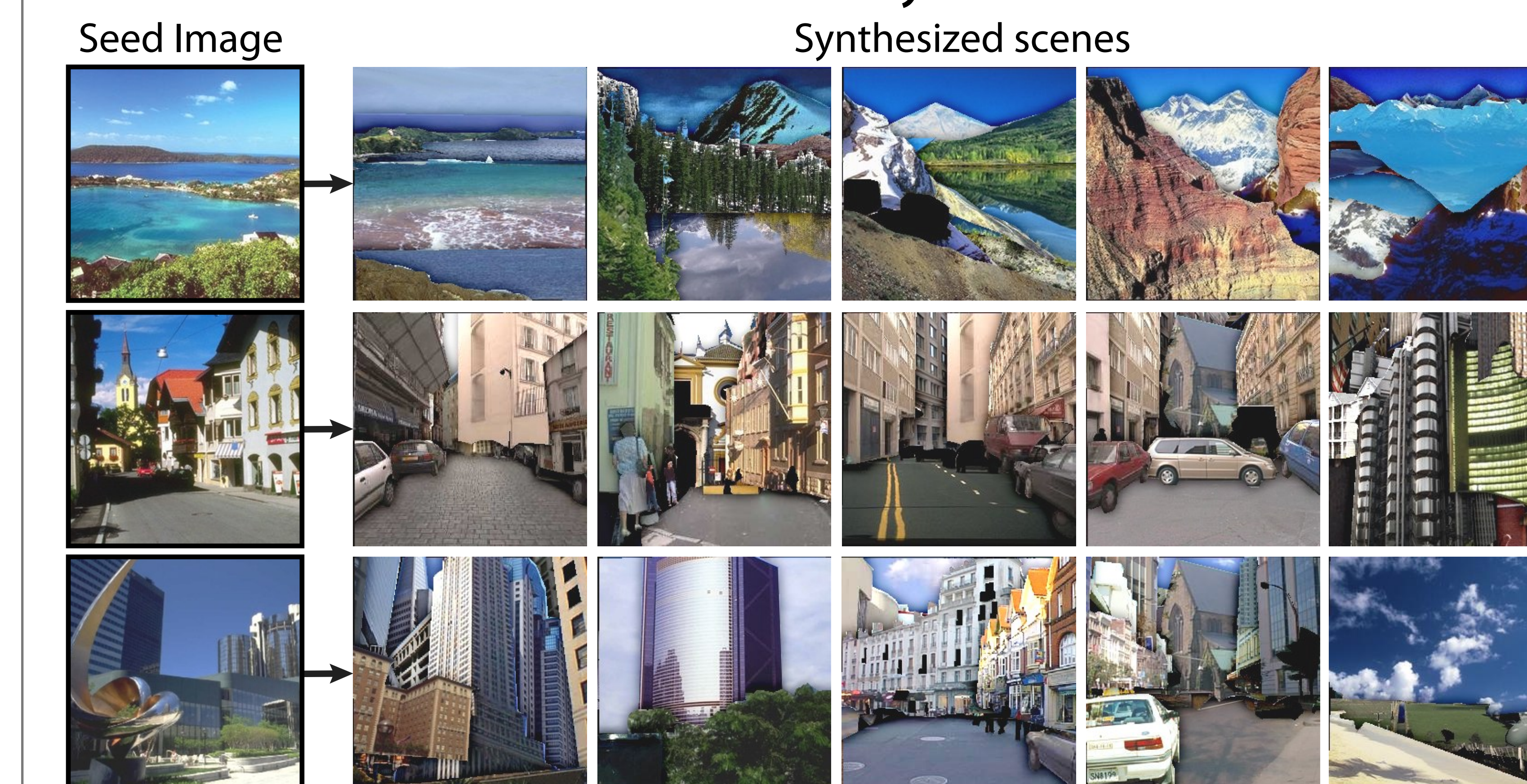


Image editing

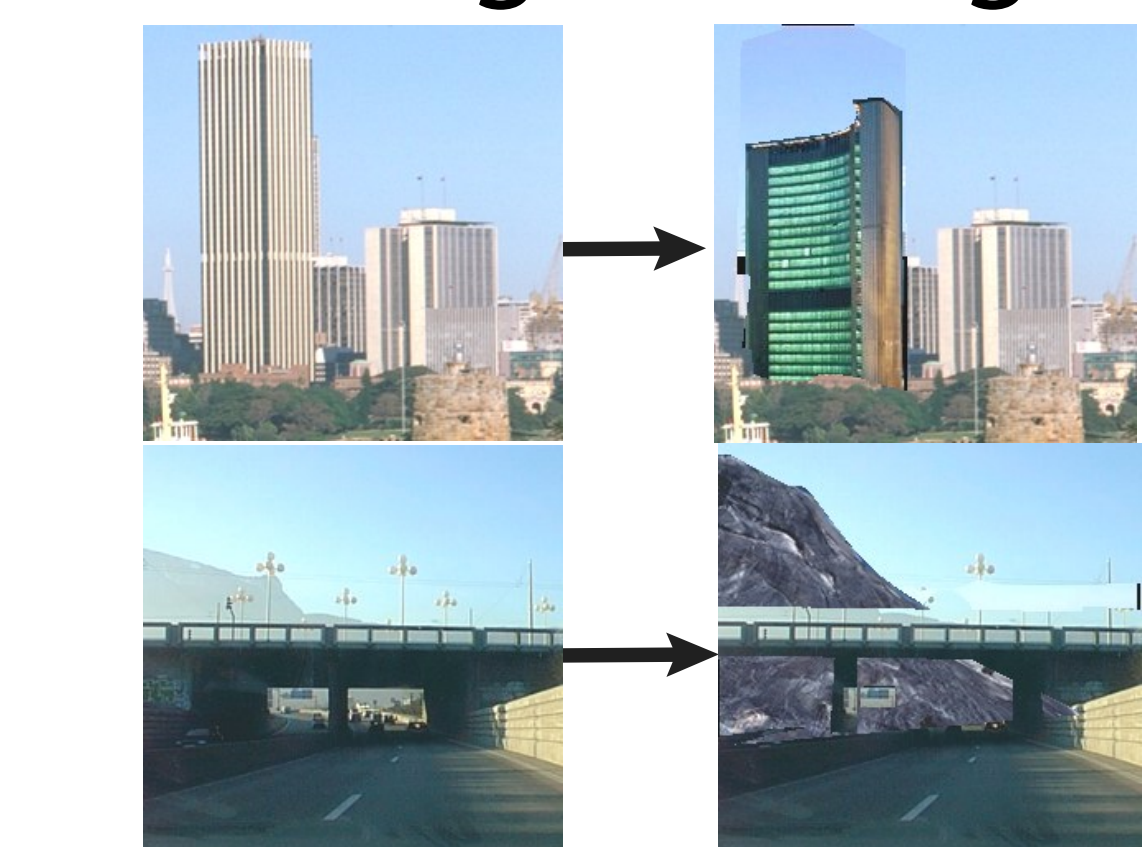
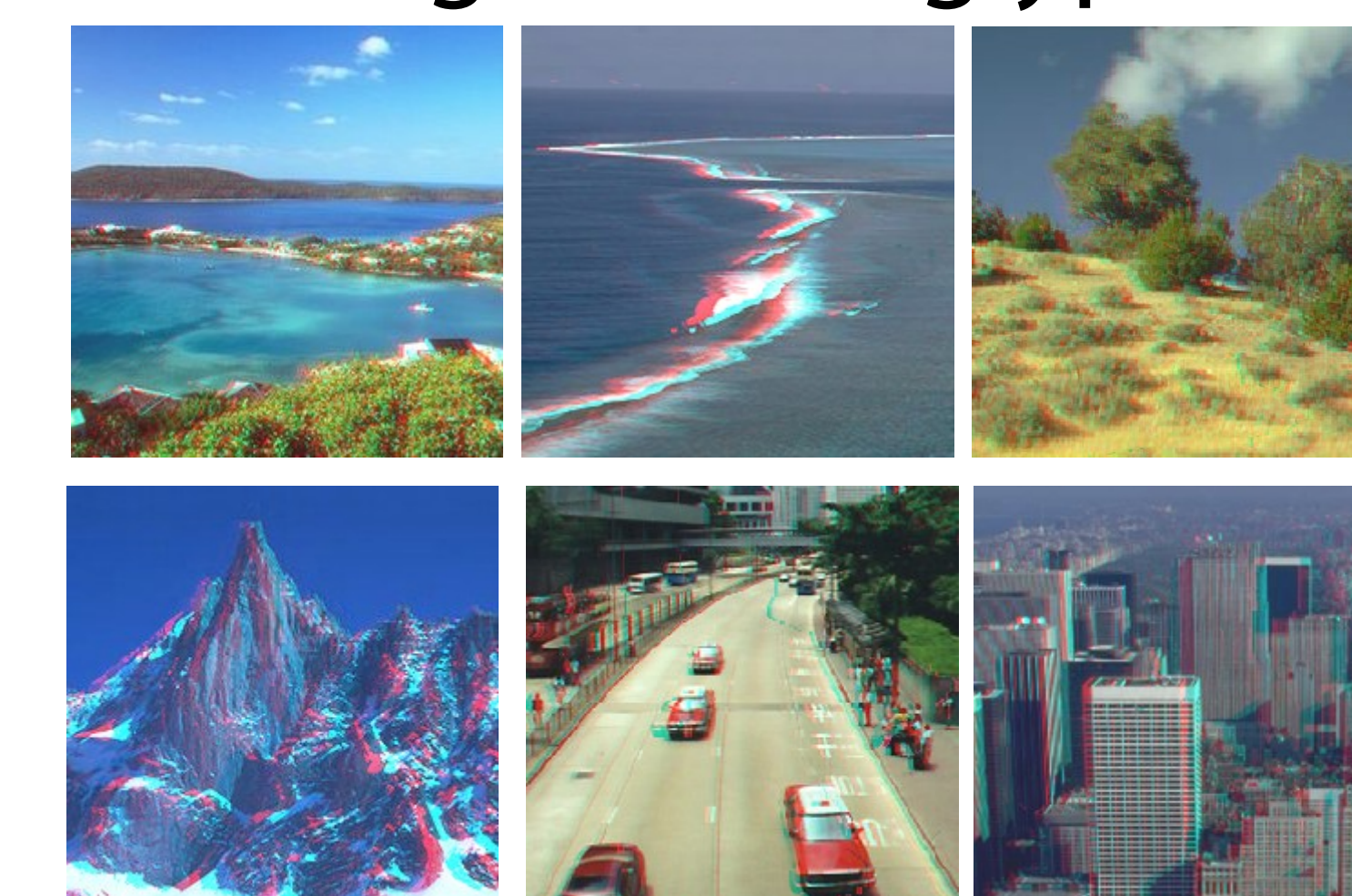


Image-to-anaglyph



[1] C. Liu, J. Yuen, and A. Torralba. Nonparametric Scene Parsing via Label Transfer. PAMI, 2011.

[2] J. Tighe and S. Lazebnik. Finding things: Image parsing with regions and per-exemplar detectors. CVPR, 2013.

[3] C. Farabet, C. Couprie, L. Najman, and Y. LeCun. Scene Parsing with Multi-scale Feature Learning, Purity Trees, and Optimal Covers. ICML, 2012.