Unsupervised Learning for Representing Visual Concepts

Michael Chang
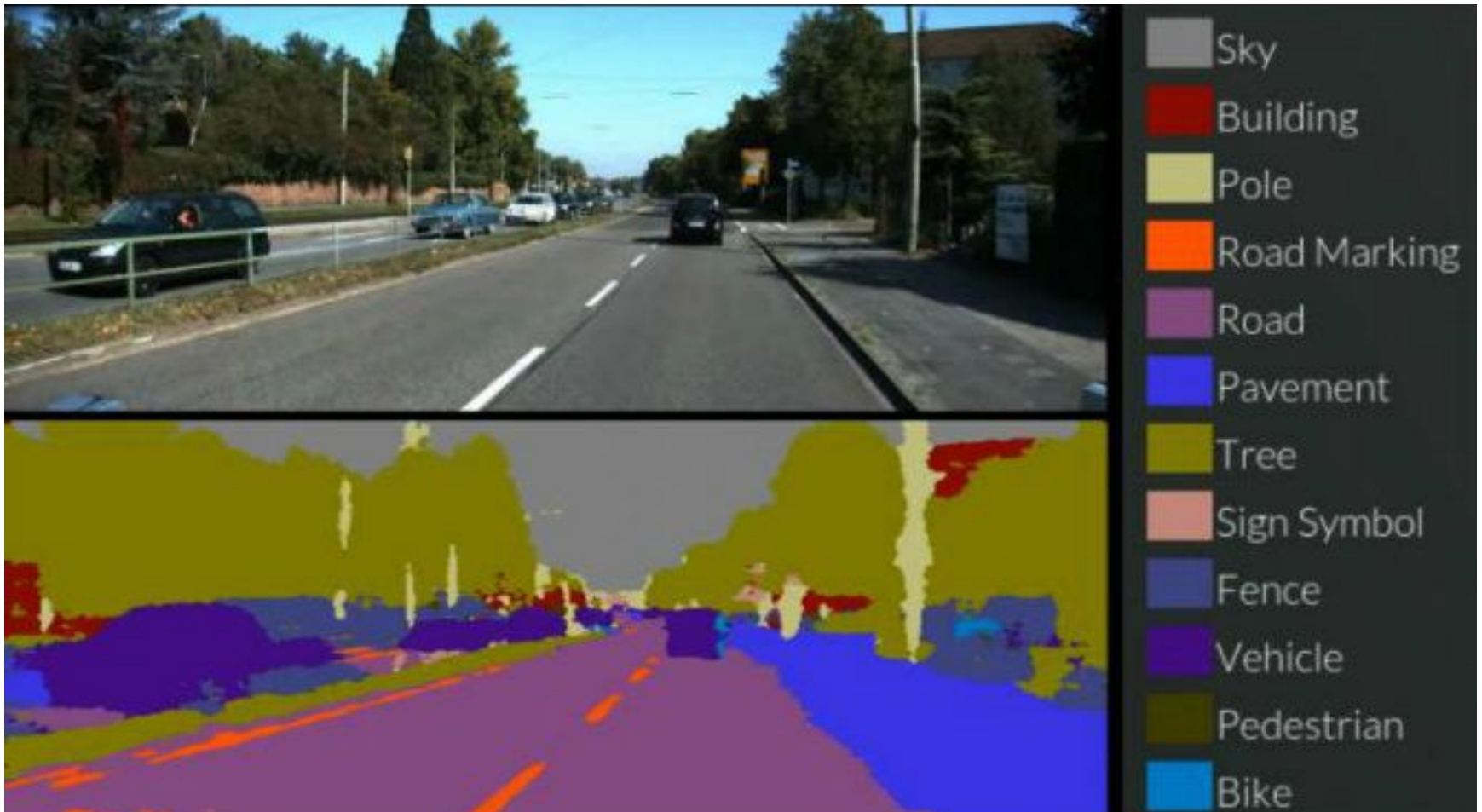
http://arxiv.org/pdf/1602.06822v1.pdf

[William Whitney, Michael Chang, Tejas Kulkarni, Joshua Tenenbaum]

Automatic Image Captioning

"A group of young people playing a game of frisbee."
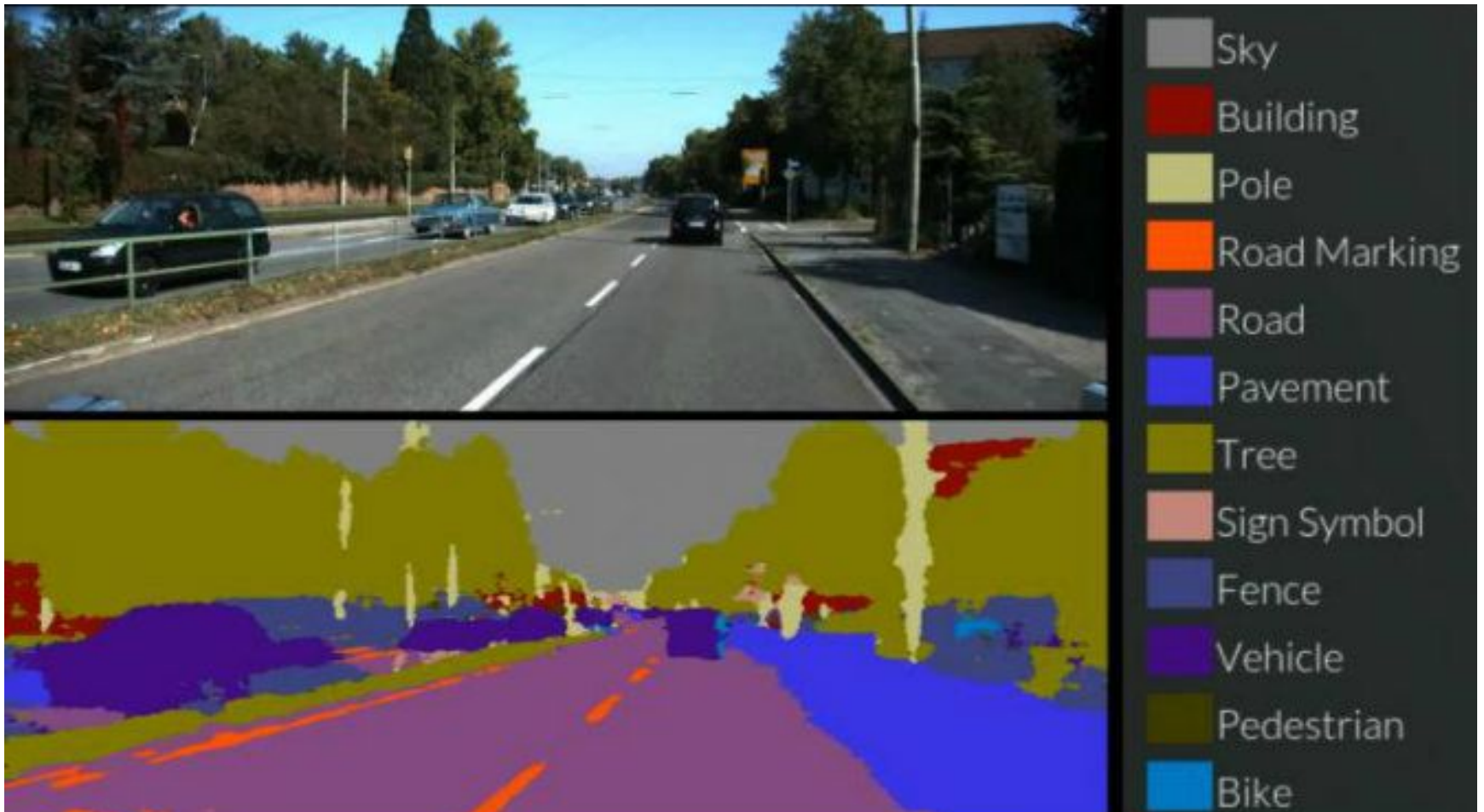
Scene Segmentation

[Segnet]

So. Much. Data.

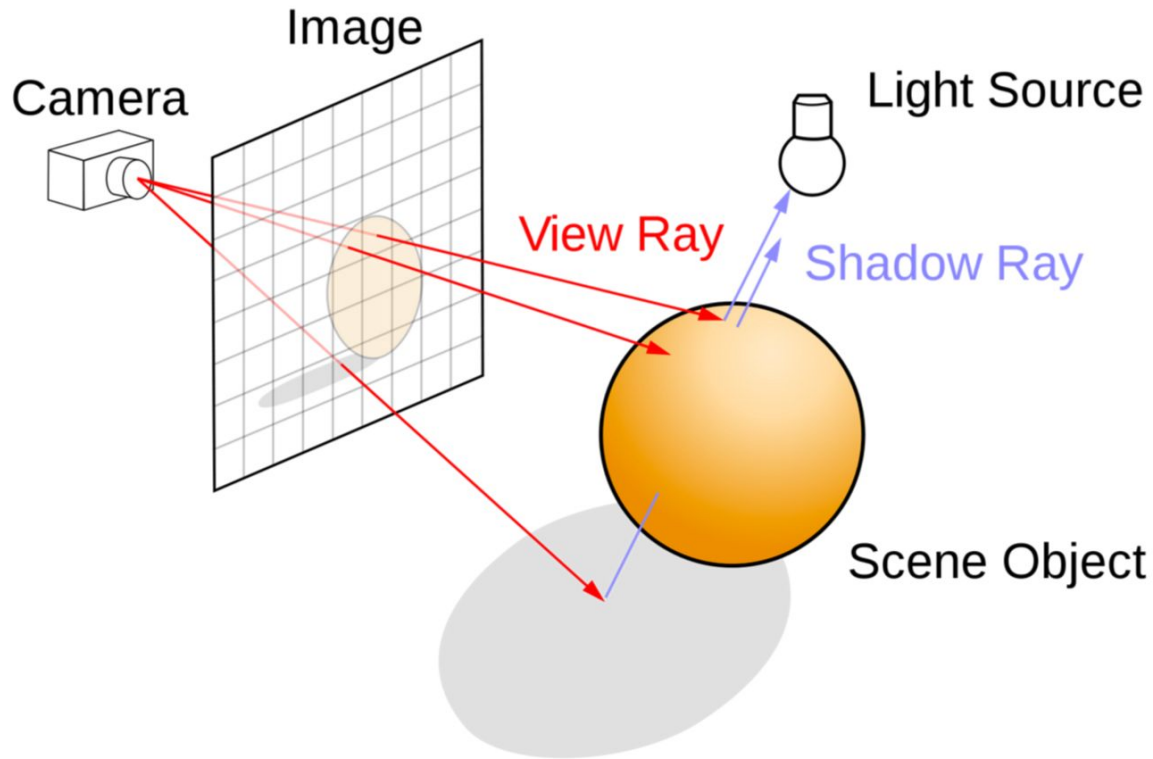# Hours of Manual Labor

[Segnet]

Task-Specific
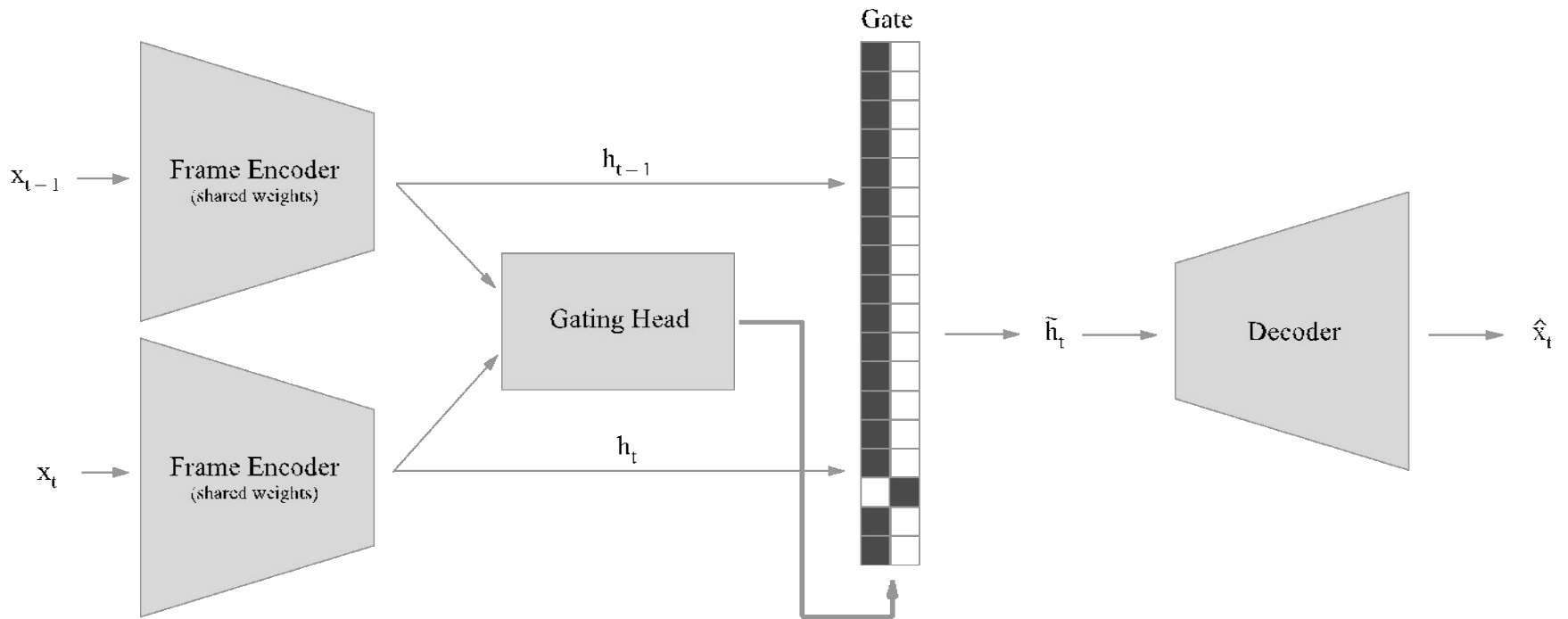

"A group of young people playing a game of frisbee."

Computer Vision as Inverse Graphics

Disentangled representation: objects, parts, pose, lighting

Reflects causal structure in the world

[Whitney]

Model

[William Whitney, Michael Chang, Tejas Kulkarni, Joshua Tenenbaum]
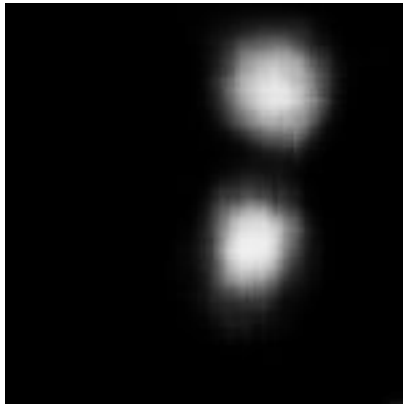
Learning Binary Decisions

- Hard binary decisions are hard to learn
  - no gradients
- Continuation method for smoothing between soft weighting and hard decision

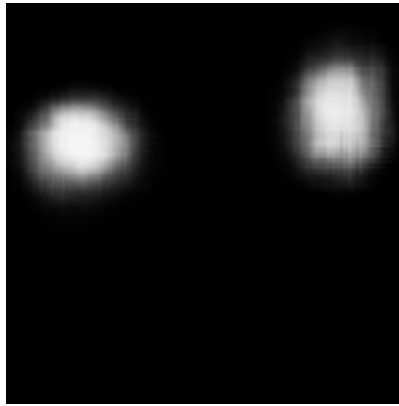$$w_i' = \frac{\left(w_i + \mathcal{N}(0, \sigma^2)\right)^{\gamma}}{\sum_j w_j^{\gamma}}$$
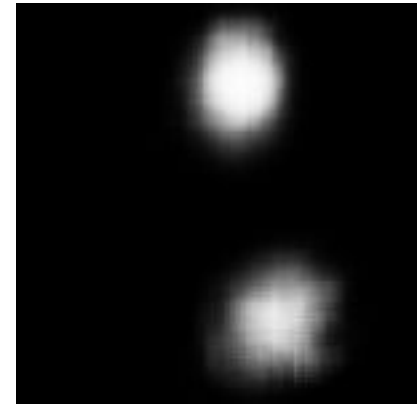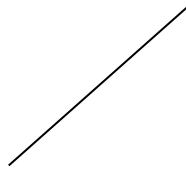
Unsupervised Learning for Representing Visual Concepts

[William Whitney, Michael Chang, Tejas Kulkarni, Joshua Tenenbaum]

Unsupervised Learning for Representing Visual Concepts

52

1

17

[William Whitney, Michael Chang, Tejas Kulkarni, Joshua Tenenbaum]

Unsupervised Learning for Representing Visual Concepts

[William Whitney, Michael Chang, Tejas Kulkarni, Joshua Tenenbaum]