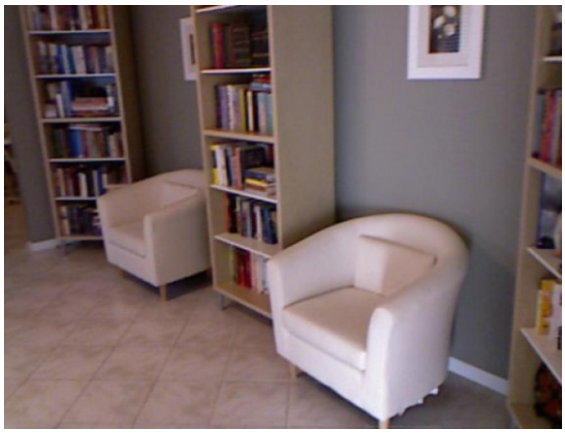
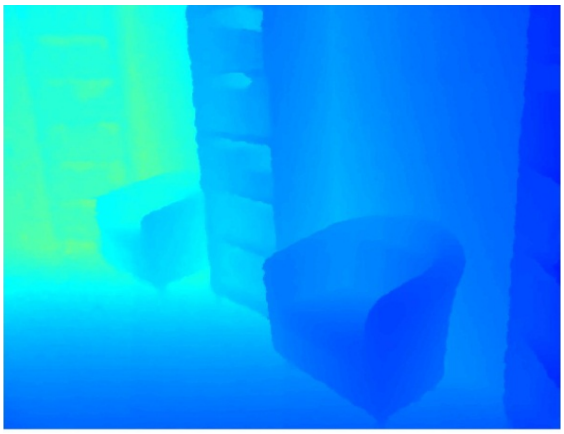
DEPTH FROM SINGLE IMAGE



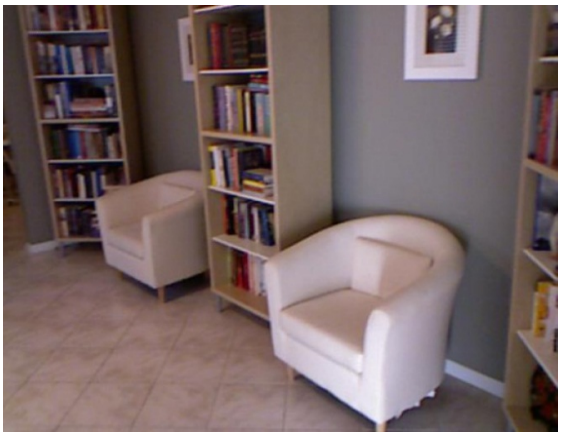
Input

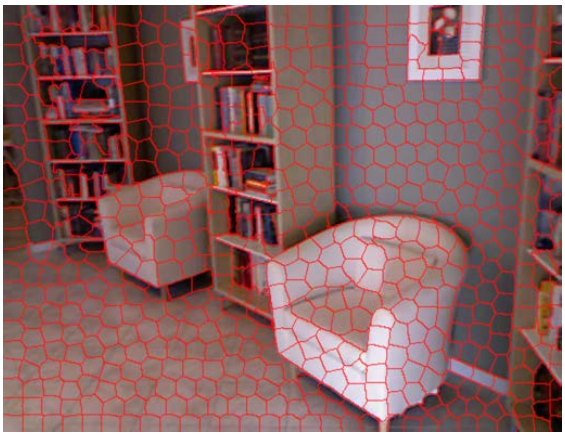


Output

LOCAL PREDICTIONS

Superpixels:

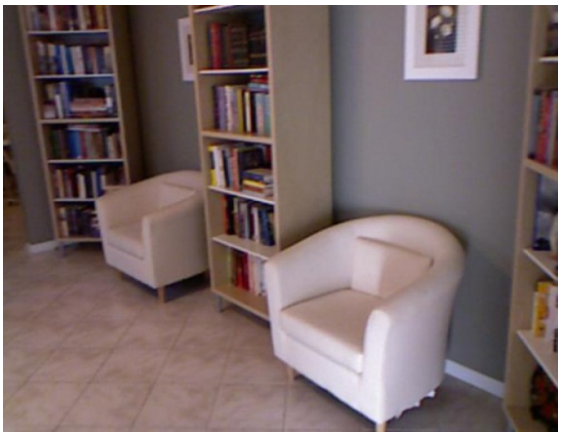


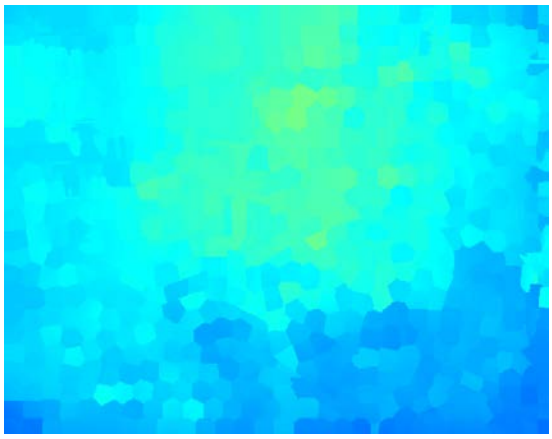


Achanta et al., PAMI'12

LOCAL PREDICTIONS

Train a regressor to predict superpixel depth:





—> Noisy predictions.

Encouraging coherence

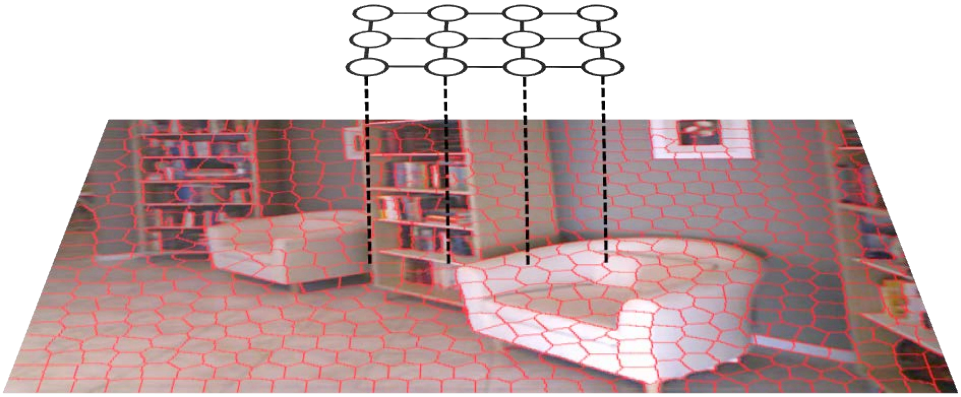
Connect the neighboring superpixels

Encourage their depths to be consistent.

Saxena et al., PAMI'09

MARKOV RANDOM FIELD

Graph with vertices and edges

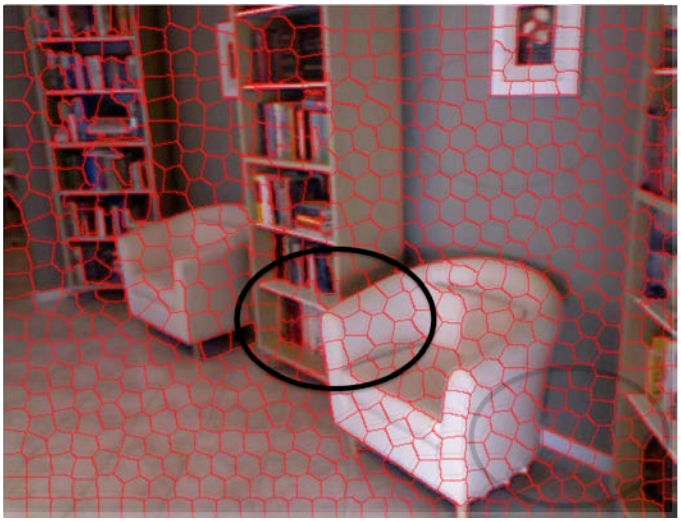


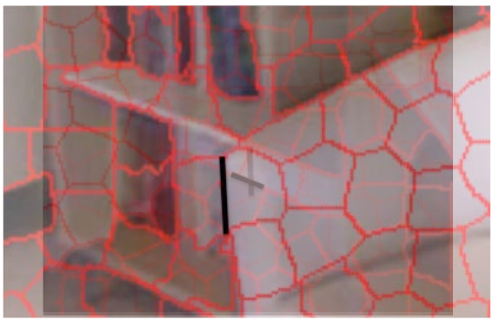
Assign values to the nodes to minimize

$$E(Y) = \sum_{i} \varphi(y_i) + \sum_{(i,j)} \psi(y_i, y_j)$$

unary pairwise

REASONING ABOUT EDGES

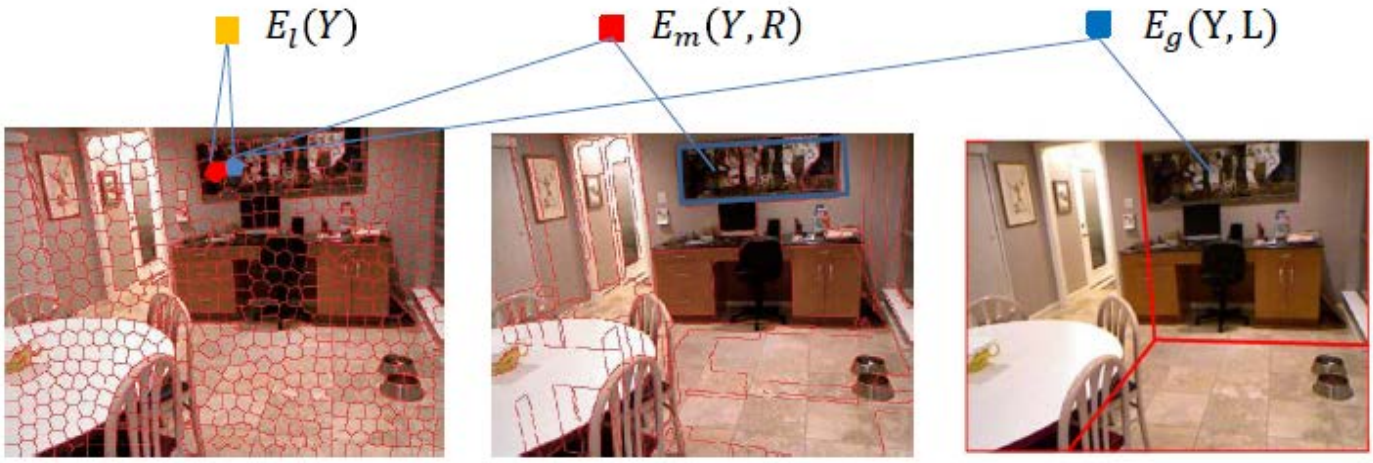




Liu et al., CVPR 2014

HIGHER ORDER TERMS

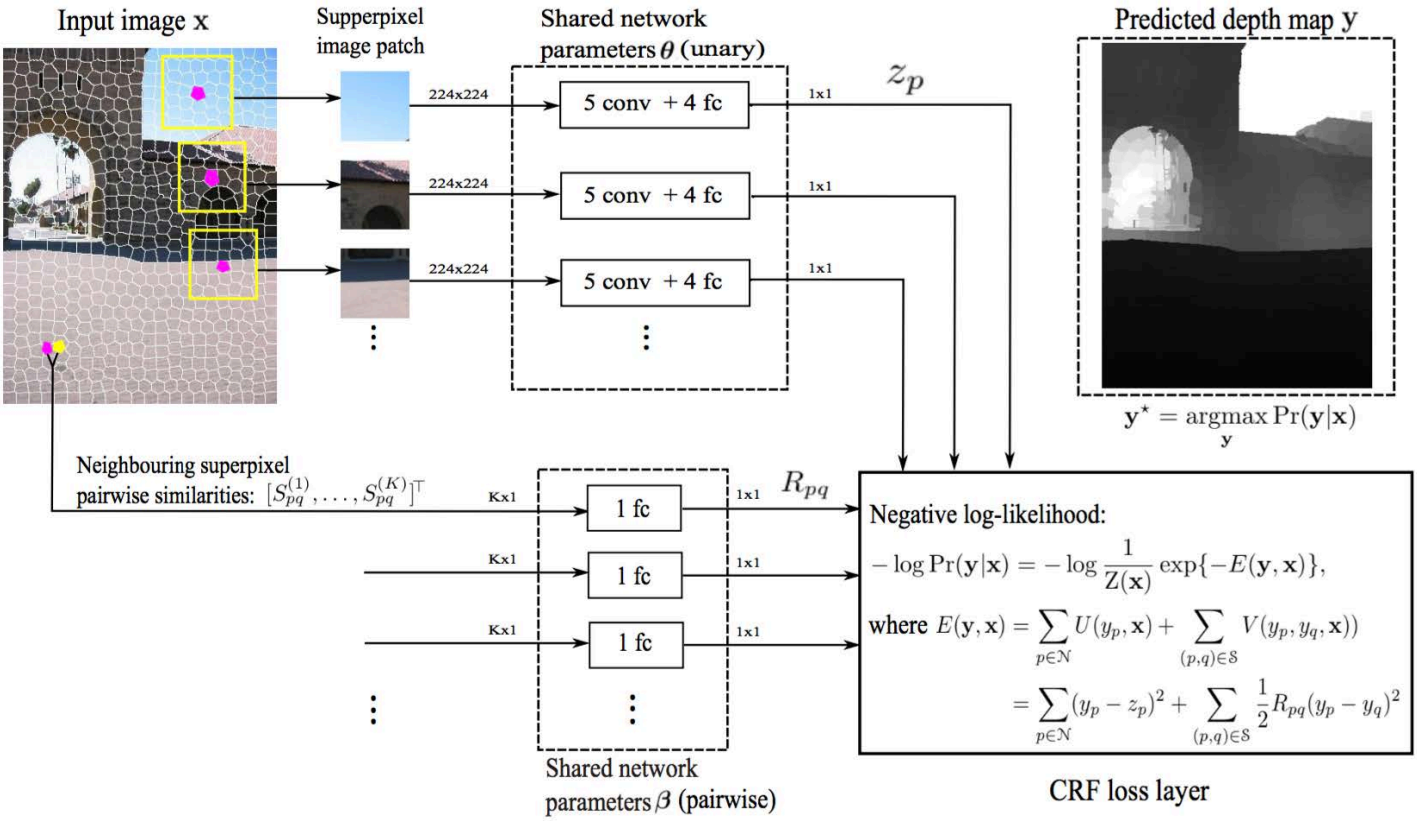
Larger regions can help reason about the scene



Local Mid-level Global

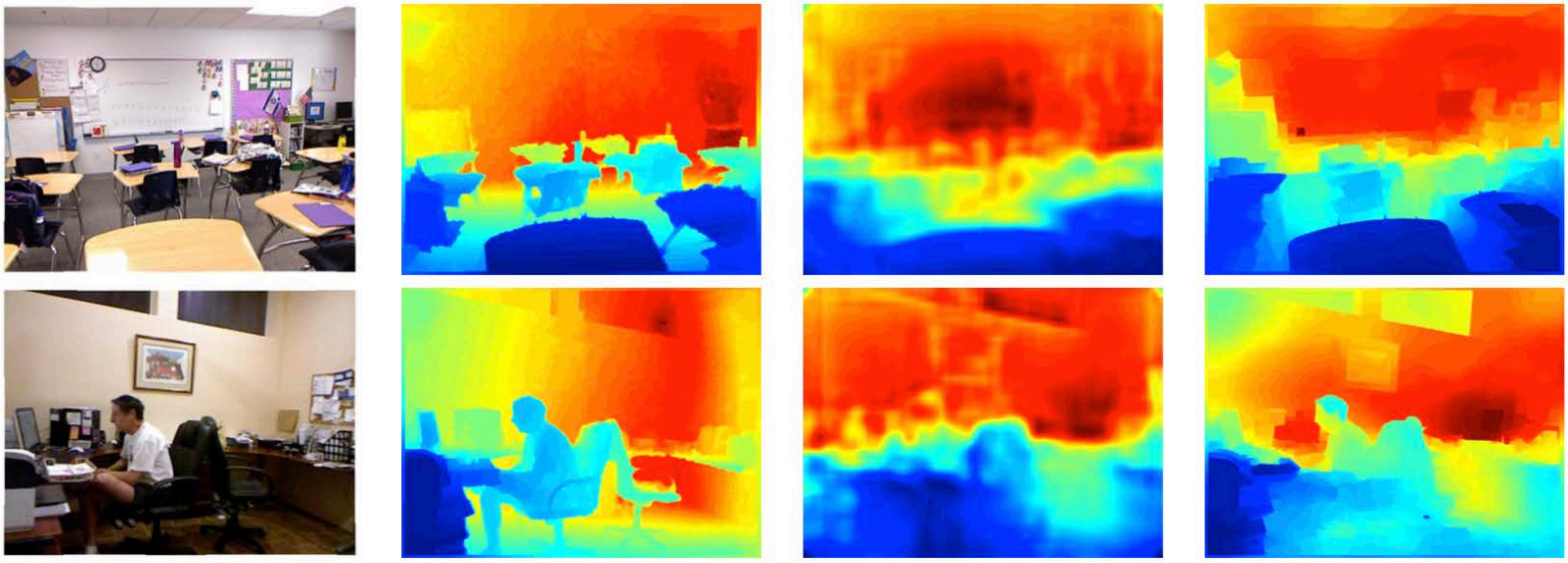
Zhuo et al., CVPR 2015

DEEP LEARNING WITH MRF



Liu et al., PAMI 2016

DEPTH FROM A SINGLE IMAGE

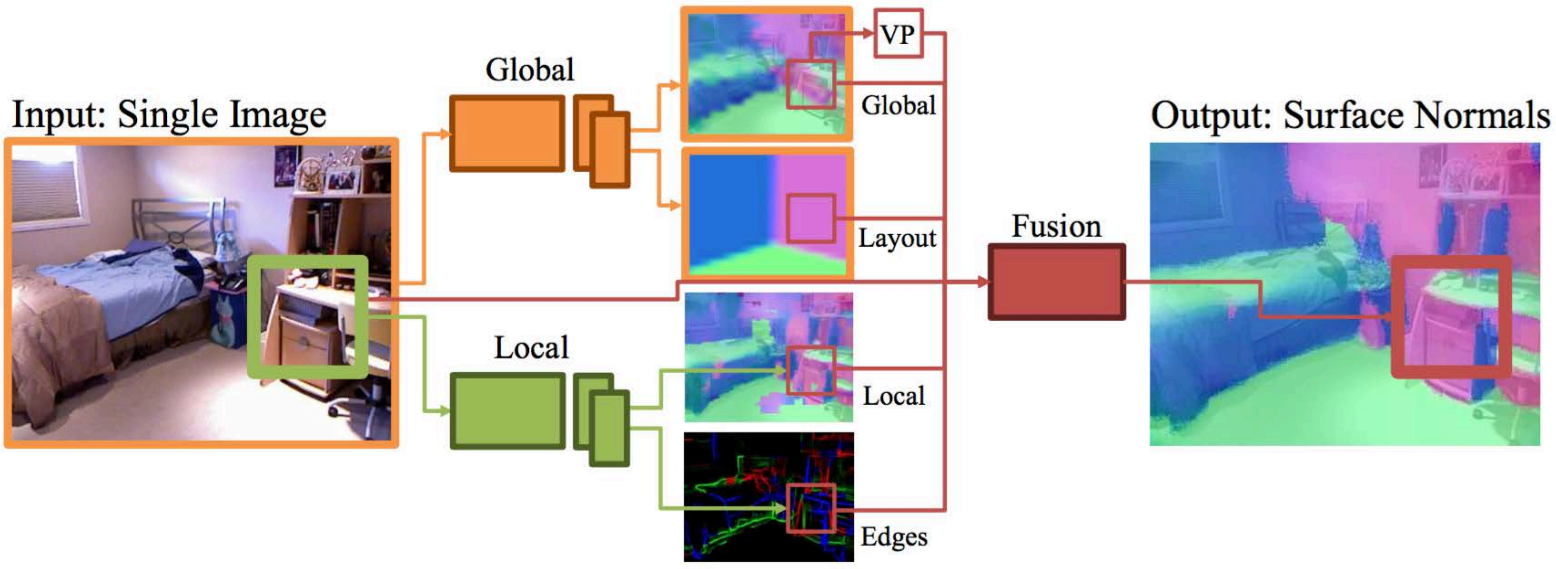


Test image | Ground-truth | Eigen et al. [3] | DCNF-FCSP

Liu et al., PAMI 2016

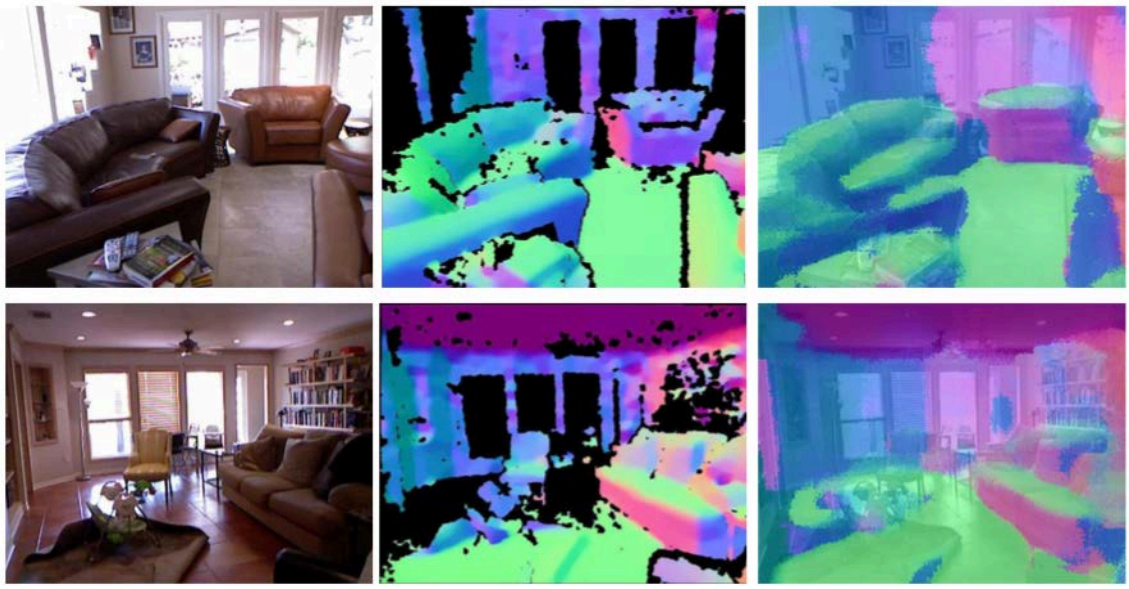
PREDICTING NORMALS

Using deep learning



Wang et al., CVPR 2015

NORMALS FROM A SINGLE IMAGE



Input Ground Truth Output

Wang et al., CVPR 2015

OLD VARIATIONAL METHODS

Minimize:

$$\int\int \left(\left[I(u,v) - Ref(\frac{\delta z}{\delta u}, \frac{\delta z}{\delta v}) \right]^2 + \lambda \left[\left(\frac{\delta^2 z}{\delta u^2} \right)^2 + \left(\frac{\delta^2 z}{\delta u \delta v} \right)^2 + \left(\frac{\delta^2 z}{\delta v^2} \right)^2 \right] \right) dudv$$

Brightness constraint

Smoothness term

or:

$$\int\int \left([I(u,v) - Ref(p,q)]^2 + \lambda \left[\left(\frac{\delta p}{\delta u} \right)^2 + \left(\frac{\delta p}{\delta v} \right)^2 + \left(\frac{\delta q}{\delta u} \right)^2 + \left(\frac{\delta q}{\delta v} \right)^2 \right] + \mu \left[\frac{\delta p}{\delta v} - \frac{\delta q}{\delta u} \right]^2 \right) dudv$$

Integrability constraint

STRENGTHS AND LIMITATIONS

Strengths:

- More general than shape-from-texture.
- Leverages data.

Limitations:

- Requires training data for specific scenes.
- Currently, only limited geometrical reasoning.