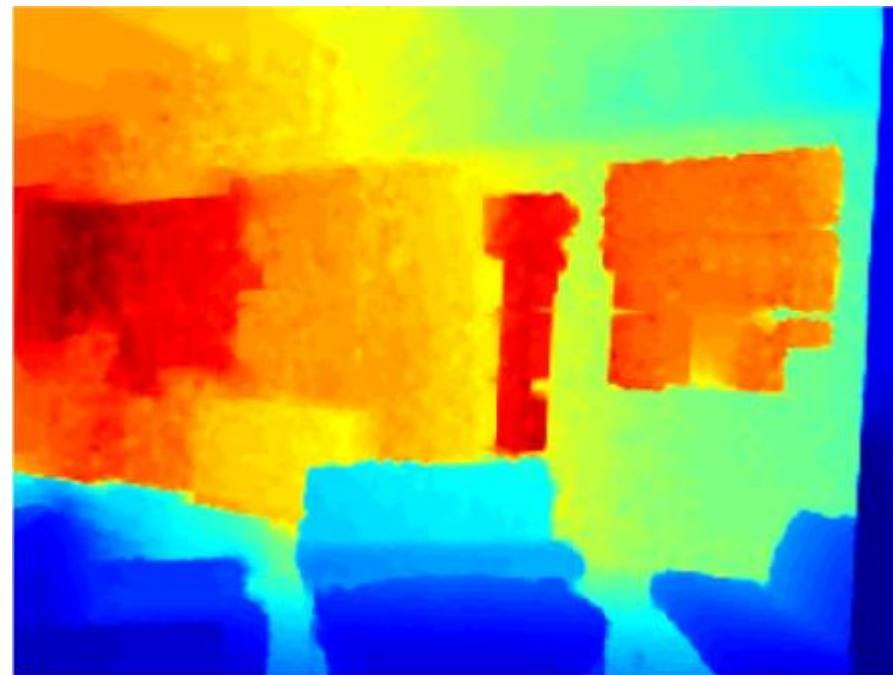


# Monocular Depth Estimation Based on Multi-Scale Graph Convolution Networks

Wooseok Oh

Seoul National University



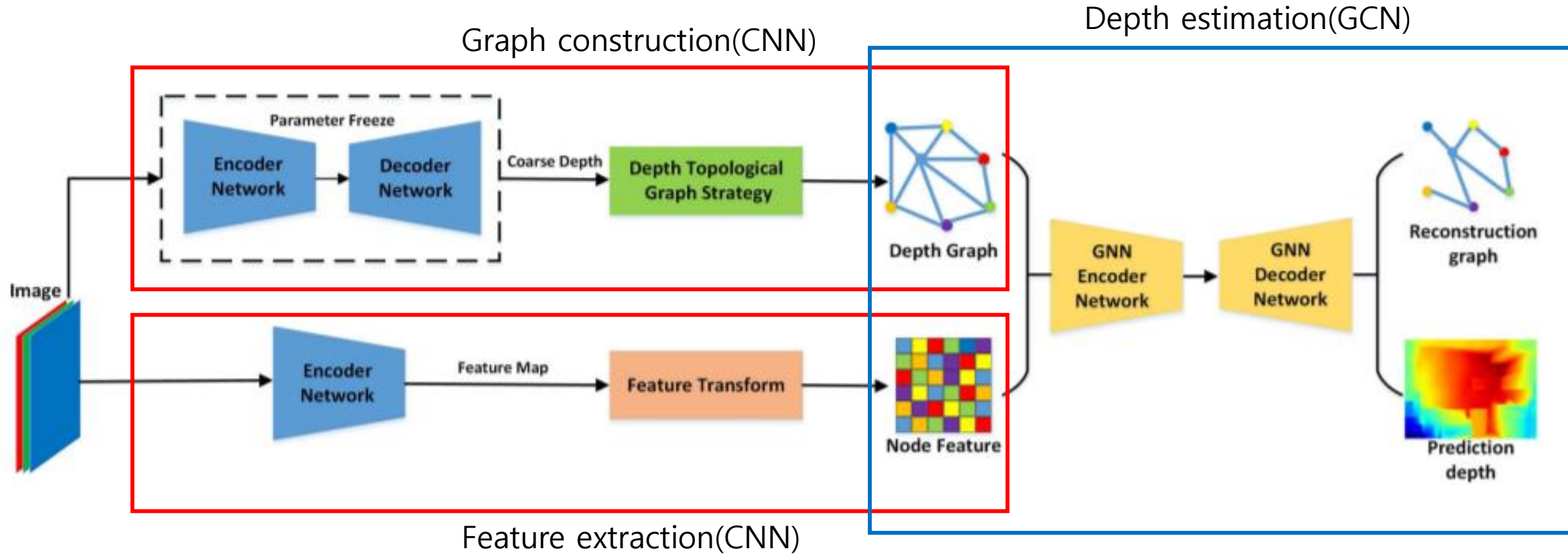
## Motivation to use GCN

- CNN can only extract the feature of receptive field of filter
- Depth value is related to
  - The neighboring pixel values → CNN
  - Other pixels of the same depth value → GCN

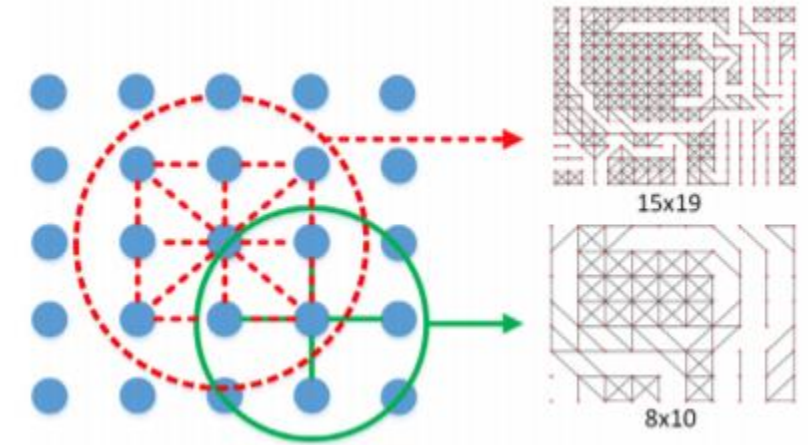
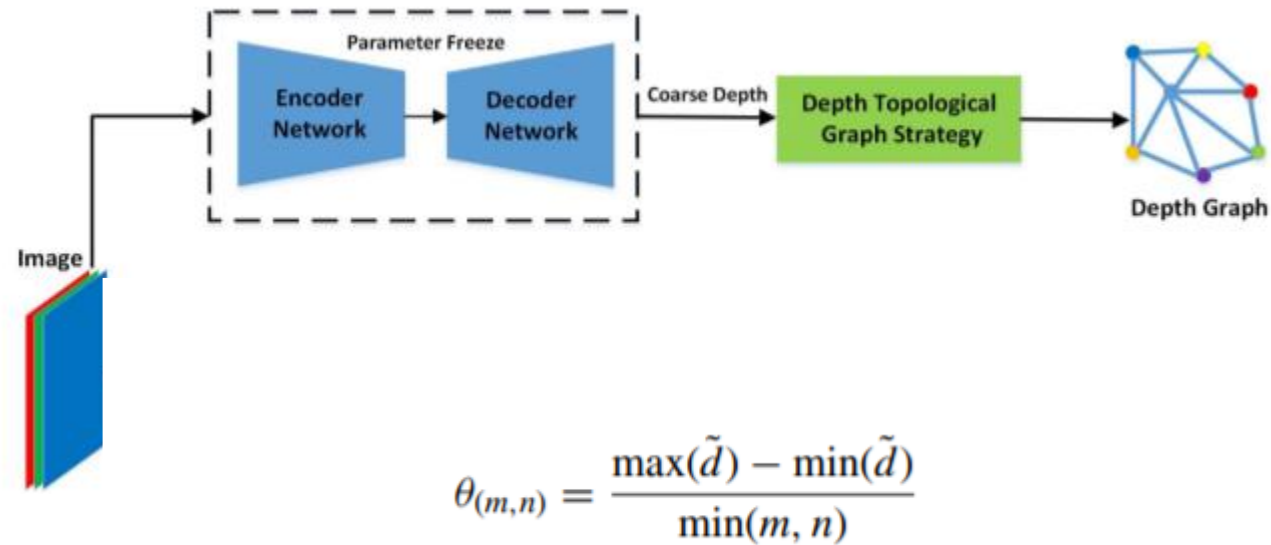


- Neighbor & not related
- Not neighbor & related
- Neighbor & related

# Network Architecture



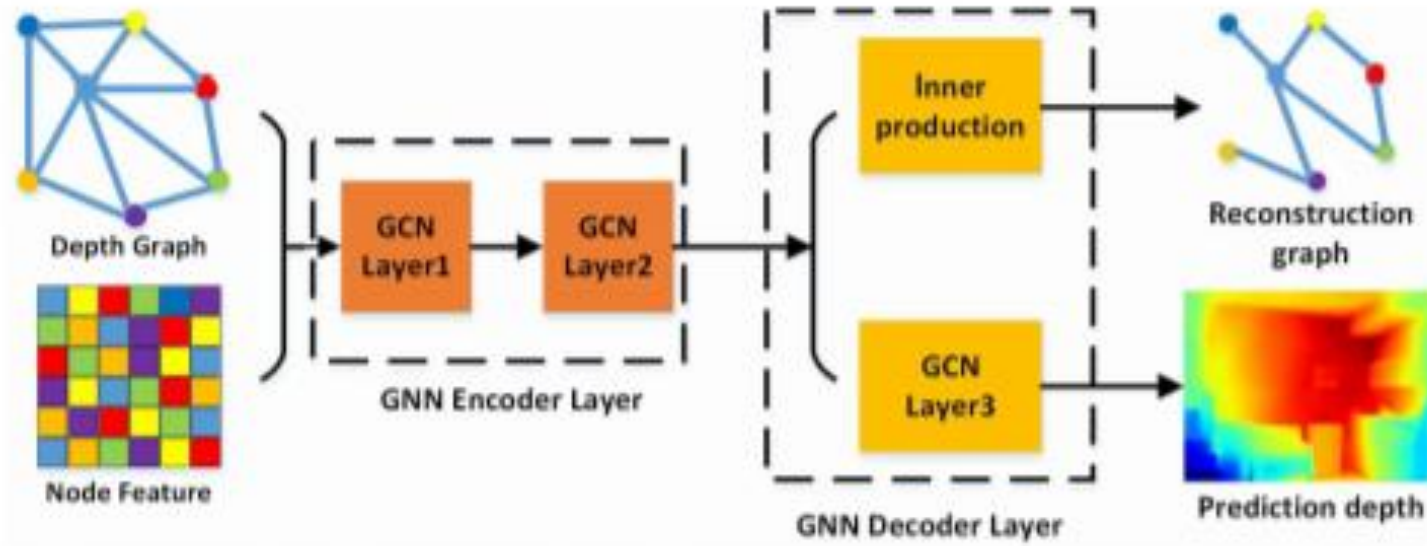
# Graph Construction



Graph construction using threshold  $\theta$

where  $\tilde{d}$  is the coarse-grained depth map.  $m$  is the rows number.  $n$  is the columns number.

# GCN for depth estimation



$$\tilde{A} = A + I \quad (1)$$

$$H^{l+1} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^l W^l) \quad (2)$$

$$\hat{A} = \text{sigmoid}(H^{l+1} H^{l+1}^T) \quad (3)$$

$A$  : adjacency matrix of graph

$\tilde{D}$  : degree matrix of  $\tilde{A}$  (all weights are same)

$H^l$ :  $l$ 'th layer's node feature. (N x D matrix)

( $H^0$  = feature extracted by CNN)

$W^l$ : trainable weight matrix

$\hat{A}$  : reconstruction graph's adjacency matrix

# Multiscale GCN

