Depth Prediction and RGBD Images for Recognition

Yihui He, Metehan Ozten

yihuihe@foxmail.com, m_ozten@umail.ucsb.edu

May 25, 2016
Related work and motivation

Yihui He, Metehan Ozten
Depth Prediction and RGBD Images for Recognition
overview

our project: depth estimation & Classification on RGBD images

implement previous work

Go further

(2) Build a RGBD CIFAR10 based on indoor depth knowledge

(3) Compare RGBD and RGB
\[ label = f(RGBD) \]
\[ label = f(RGB) \]
first part: implement previous work

infer depth from RGB image
At training time, we combine two objective function\(^1\)

1. regress to ground truth depth image (Kinect, PrimeSense)
\[ \sum_p (y_p - \hat{y}_p)^2, \ p \text{ stands for pixel.} \]

2. Similarity between superpixels.
\[ R_{pq} = \sum_{k=1}^{K} \beta_k S_{pq}^{(k)} \]
\(\beta\) is trainable weight. \(S\) is similarity function.

Architecture: Deep convolutional Neural Field

Image Segments → Superpixel Neighbor Graph → Unitary Network → Similarity Function

Joint Loss: \[ \sum (d_{true} - d_{set})^2 + \lambda \sum (y_{true} - y_{set})^2 \]
Infer depth from RGB image: Supervised part

using traditional CNN.
Compare performance with original paper

<table>
<thead>
<tr>
<th>Method</th>
<th>Error (lower is better)</th>
<th>Accuracy (higher is better)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rel</td>
<td>log10</td>
</tr>
<tr>
<td>Our implementation</td>
<td>0.252</td>
<td>0.103</td>
</tr>
<tr>
<td>Original paper</td>
<td>0.230</td>
<td><strong>0.095</strong></td>
</tr>
</tbody>
</table>

**Table:** Sanity check (**Bold** is better)
Classification on RGBD images
build RGBD CIFAR dataset

32x32x3

400x400x1

400x400x3

32x32x4

Yihui He, Metehan Ozten

Depth Prediction and RGBD Images for Recognition
Yihui He, Metehan Ozten
Depth Prediction and RGBD Images for Recognition
architecture

32x32x4

airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck
R vs G vs B vs D: training time

Yihui He, Metehan Ozten

Depth Prediction and RGBD Images for Recognition
R vs G vs B vs D: testing time

Epoch

Validation accuracy

Yihui He, Metehan Ozten

Depth Prediction and RGBD Images for Recognition
results

Num Iters vs. Accuracy

Accuracy

Num Iters

- RGBD-Train
- RGBD-Val
- RGBD-Test
- RGB-Train
- RGB-Val
- RGB-Test
our contribution

1. reproduce previous work on depth estimation
2. create the first RGBD CIFAR10 dataset
3. define a new metric for depth prediction problem
4. prove that depth channel has a better feature representation
5. show that training on RGBD images can somehow improve accuracy
questions?²

²code, references, report and slides can be access here: https://github.com/yihui-he/Depth-estimation-with-neural-network