Im2depth: Scalable Exemplar based Depth Transfer

Mohammad Haris Baig\textsuperscript{1}, Vignesh Jagadesh\textsuperscript{2}, Robinson Piramuthu\textsuperscript{2}, Wei Di\textsuperscript{2}, Anurag Bhardwaj\textsuperscript{2}, Neel Sundaresan\textsuperscript{2}
**Im2depth: The Problem**

**Goal:** Given a **Single Image** of an **Indoor Scene**, we want to perform **Dense Depth Estimation**

![Image](image_url)

- **True Depth Map**
- **Estimated Depth Map**

**Our Method**
**Im2depth: First Impressions**

**Depth Estimation**
- Shape
- Absolute Depth

<table>
<thead>
<tr>
<th>Image</th>
<th>Ground Truth</th>
<th>Im2Depth</th>
<th>BU/Google</th>
<th>NPS</th>
<th>Global Prior</th>
<th>Make3D</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Image" /></td>
<td><img src="groundtruth1" alt="Ground Truth" /></td>
<td><img src="im2depth1" alt="Im2Depth" /></td>
<td><img src="bu-google1" alt="BU/Google" /></td>
<td><img src="nps1" alt="NPS" /></td>
<td><img src="global-prior1" alt="Global Prior" /></td>
<td><img src="make3d1" alt="Make3D" /></td>
</tr>
<tr>
<td><img src="image2" alt="Image" /></td>
<td><img src="groundtruth2" alt="Ground Truth" /></td>
<td><img src="im2depth2" alt="Im2Depth" /></td>
<td><img src="bu-google2" alt="BU/Google" /></td>
<td><img src="nps2" alt="NPS" /></td>
<td><img src="global-prior2" alt="Global Prior" /></td>
<td><img src="make3d2" alt="Make3D" /></td>
</tr>
<tr>
<td><img src="image3" alt="Image" /></td>
<td><img src="groundtruth3" alt="Ground Truth" /></td>
<td><img src="im2depth3" alt="Im2Depth" /></td>
<td><img src="bu-google3" alt="BU/Google" /></td>
<td><img src="nps3" alt="NPS" /></td>
<td><img src="global-prior3" alt="Global Prior" /></td>
<td><img src="make3d3" alt="Make3D" /></td>
</tr>
</tbody>
</table>
Observation: New Large Scale RGBD Datasets can be used to form data dependent depth transfer approaches.

Observation: Depth Maps share a lot of similarities in Global Structure.

Visualizing Depths via an Image Space

Visualizing Depths via a Depth Space
**Im2depth: RGB & Depth Spaces**

**Depth Space**

Exploiting Global Structure

\[
\begin{array}{c}
\text{Exemplar Basis} \\
\text{Depth}
\end{array}
= \begin{array}{c}
\text{Depth Space Basis} \\
0 \\
0 \\
0.9
\end{array}
\]

Visualization of Depth Maps via Depth Space Descriptor

**Image Space**

**Observation:** We want an image representation that encodes Global Properties of the Image

Global Image Feature

**Observation:** Information from the Neighborhood can be used

Rich, Compact Image Space Descriptor

Visualization of Depth Maps via Image Space Descriptor
Im2depth: Putting it all Together

Test Image

1) Image Space Descriptor

2) Transformation

Depth Space Descriptor

Exemplar Basis

3) Super Pixels

0.6 0 0.1

4) Coarse Depth

Refined Depth

Ground Truth

Exemplar Basis

3)
Im2depth: Learning a Mapping

Use an RGBD dataset, to learn Transformation between Image Space and Depth Space

Images

Image Space

Learn a Transformation

Supervised Learning
Linear Regression

Depths

Depth Space

0.2 0 0.6
0 0 0.8
0.1 0.4 0
0 0.7 0
0.8 0 0
**Im2depth: Evaluation**

**Dataset**

NYUV2 (RGBD Indoor Scene Dataset)
- 27 Scene Categories
- 464 Distinct Scenes
- 1449 RGBD Examples
  - 1200 Training
  - 249 Testing

**Quantitative Results**

<table>
<thead>
<tr>
<th>Methods</th>
<th>L1 Err (m)</th>
<th>NCC Metric</th>
<th>Relative Err</th>
<th>RMS Err</th>
<th>Log10 Err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>0.6901</td>
<td>0.68529</td>
<td>0.2941</td>
<td>0.85205</td>
<td>0.11241</td>
</tr>
<tr>
<td>BG [10]</td>
<td>0.81695</td>
<td>0.61387</td>
<td>0.33981</td>
<td>1.0092</td>
<td>0.13172</td>
</tr>
<tr>
<td>NPS [9]</td>
<td>0.82514</td>
<td>0.6079</td>
<td>0.35482</td>
<td>1.0133</td>
<td>0.13301</td>
</tr>
<tr>
<td>GLB</td>
<td>0.90222</td>
<td>0.58058</td>
<td>0.38923</td>
<td>1.0988</td>
<td>0.14415</td>
</tr>
<tr>
<td>MK3D [14]</td>
<td>3.4781</td>
<td>0.40379</td>
<td>1.6228</td>
<td>3.9521</td>
<td>0.36463</td>
</tr>
</tbody>
</table>

Im2Depth outperforms state of the art methods on all 5 metrics.
Im2depth: Experiments

Image Space

Objective: Find a Global Image Descriptor that yields the most informative Image Space

Conclusion: Classemes is the most informative descriptor

Depth Space

Objective: Finding good size of exemplar dictionaries

Conclusion: Use bigger sized Visual space Dictionaries and smaller Depth Dictionaries.
Im2depth: Conclusions

Dense Depth Estimation Method for Indoor Scenes

- Extremely Fast
- Highly Scalable
- Simple closed form Training
- Closed form Inference
- Well Suited for Mobile (Light-Weight)