**Deep Depth From Focus**

Caner Hazirbas, Sebastian G. Soyer, Maximilian C. Staab, Laura Leal-Taixé and Daniel Cremers

**Introduction**
- Depth can be recovered from a focal stack
- Pixel sharpness determines where the pixel is focused in the stack, and hence the depth of pixel

**Contributions**
- We propose an end-to-end trained network for **depth from focus**
- We introduce a large indoor dataset with 720 light-fields and co-registered ground truth depth maps
- We compare several state-of-the-art methods for DFF and analyse several network variations

**DDFF 12-Scene Dataset**
- 720 light-fields with groundtruth
- 9x9 undistorted sub-apertures
- 383x552 image resolution
- Indoor, real-world challenges

**Mobile Depth From Focus Dataset**
- 202 focal stacks captured with an Android smartphone
- Registered depth maps captured by the smartphone
- Publicly available

**Experiments**
- Quantitative results on the **DDFF 12-Scene** benchmark

<table>
<thead>
<tr>
<th>Method</th>
<th>MSE</th>
<th>RMS</th>
<th>Bumpness</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSPNet</td>
<td>$9.4 \times 10^{-4}$</td>
<td>0.03</td>
<td>0.55</td>
</tr>
<tr>
<td>Lytro</td>
<td>$2.1 \times 10^{-4}$</td>
<td>0.04</td>
<td>1.02</td>
</tr>
<tr>
<td>PSP-LF</td>
<td>$2.7 \times 10^{-4}$</td>
<td>0.05</td>
<td>0.54</td>
</tr>
<tr>
<td>DFLF</td>
<td>$4.8 \times 10^{-4}$</td>
<td>0.06</td>
<td>0.65</td>
</tr>
<tr>
<td>VDFF</td>
<td>$7.3 \times 10^{-4}$</td>
<td>0.08</td>
<td>0.79</td>
</tr>
<tr>
<td>DDFF-CC3</td>
<td>$9.7 \times 10^{-4}$</td>
<td>0.03</td>
<td>0.59</td>
</tr>
</tbody>
</table>

**Several variations of the encoder-decoder architecture**

**DDFF vs. state-of-the-art DFF and depth from light-field methods**

**Results on the **mDFF** dataset**