**Problem Statement**
- Issue with depth map from RGB-D sensor: noise, quantization effects, coarse resolution.

**Contribution**
- Jointly solve depth super resolution and camera motion estimation.
  - Simple data acquisition setup.
  - No lighting calibration.
  - A variational formulation.
  - End-to-end frame work.

**Variational Formulation**

```
\min_{\{R(t), t, \phi\}} \sum_{i} \phi_i (R(t), R_i, \phi_i, t) + \tau \|D_2 - D_1\|^2_F
```

- **PS term** is evaluated using robust estimator \(\phi_i\), with \(\phi_i(R(t), R_i, \phi_i, t)\).
- \(\tau\) : re-projection operator.

**Numerical Resolution**

- Rotation \(R\) and translation \(t\) are parametrized by the corresponding 6-dim twist coords \(\xi\), i.e. \(R = R(\xi), t = t(\xi)\).
- Separate normals into linear part by \(a_i[z]\) and nonlinear part \(d_i[z]\) with \(a_i[z] n_i[z], d_i[z] n_i[z]\).
- Solve nonlinear nonconvex problem by alternating reweighted least square.

**Background**
- Using a first-order spherical harmonic approximation for image formation model [1].

**Experimental Results**

**Synthetic Datasets**

<table>
<thead>
<tr>
<th>Ground Truth</th>
<th>Input Depth (z_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic</td>
<td>MAE: 27.42</td>
</tr>
<tr>
<td>[2] Ours</td>
<td>25.86</td>
</tr>
<tr>
<td>[4] Ours</td>
<td>20.49</td>
</tr>
<tr>
<td>[5] Ours</td>
<td>12.73</td>
</tr>
<tr>
<td>[6] Ours</td>
<td>12.54</td>
</tr>
</tbody>
</table>

**Real World Datasets**

<table>
<thead>
<tr>
<th>Ground Truth</th>
<th>Input Depth (z_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanford Bunny</td>
<td>MAE: 32.11</td>
</tr>
<tr>
<td>Joyful Yell</td>
<td>21.42</td>
</tr>
</tbody>
</table>

**Reference**

[7] Lu Sang, Bjorn Haefner, Daniel Cremers. Department of Informatics, Technical University of Munich

**Image**

- Multi-View Photometric Stereo
- An LED light source is attached to a RGB-D sensor.
- Camera and light source move simultaneously.
- Input depth refined normals refined depth
- Simple data acquisition setup.
- Jointly solve depth super resolution and camera motion estimation.
- PS term is evaluated using robust estimator \(\phi_i\), with \(\phi_i(R(t), R_i, \phi_i, t)\).
- PS term is evaluated using robust estimator \(\phi_i\), with \(\phi_i(R(t), R_i, \phi_i, t)\).

**Contributions**
- Simple data acquisition setup.
- A variational formulation.
- End-to-end frame work.