P2D: a self-supervised method for depth estimation from polarimetry

Marc Blanchon¹*, Désiré Sidibé², Olivier Morel¹, Ralph Seulin¹, Daniel Braun¹, Fabrice Meriaudeau¹

¹ ERL VIBOT CNRS 6000, ImVia EA 7535, UBFC, France
² IBISC, Univ Evry, France

*fr.marc.blanchon@gmail.com
**Objective**

Polarimetry to Depth (P2D)

- Generic
- Polarization adapted
- Transparency / Specularity sensitive

**Input Polarimetric Image**

**Model**

**Depth**
Previous approaches

Supervised Methods

Eigen et al.\textsuperscript{1}

DORN\textsuperscript{2}

\checkmark Required ground truth depth - No polarimetric dataset
\checkmark Lower performance compared to unsupervised / self-supervised
\xmark Less generalization

Previous approaches

Unsupervised Methods

Godard et al.\(^1\)

- ![Input](image1)
- ![Monodepth2 (M)](image2)
- ![Monodepth2 (S)](image3)
- ![Monodepth2 (MS)](image4)

EPC++\(^2\)

- ![Input](image5)
- ![Monodepth2 (M)](image6)

✅ No required groundtruth depth
✅ Generalization through loss function
❌ Mainly colorimetric approaches
❌ Limitations of the modalities exported in the model - Specularity/transparency insensitive

---


Contextualization

Contextualization

Failure Cases

Method

Why Polarization?

✅ Sensitive to specularity
✅ Particular link between acquired polarization parameters and normal
❌ No colorimetric (strong differentiable features) space
❌ No generic approach leveraging the polarization to normal relationship
Method

What is Polarization?

From a 4x4 grid of micro-polarizer to polarization representative images
Method

Linking $\alpha$ to $\vec{n}$

Legend

- $\Gamma(\vec{E})$: Reprojection of Electric field
- $\alpha$: Angle of Polarization
- $\Delta$: Angular difference

Image plane
Method

A three term minimizable error evaluation

\[ \Lambda = \mu L_r + \lambda L_s + \tau L_p \]
Method

A three term minimizable error evaluation

\[ \Lambda = \mu L_r + \lambda L_s + \tau L_p \]

Penalizes reprojection error between two consecutive frame using perspective geometry statement

---

Method

A three term minimizable error evaluation

\[ \Lambda = \mu L_r + \lambda L_s + \tau L_p \]

Penalizes reprojection error between two consecutive frame using perspective geometry statement

Ensures second order prior smoothness while preserving salient edges

Method

A three term minimizable error evaluation

$$\Lambda = \mu L_r + \lambda L_s + \tau L_p$$

Penalizes reprojection error between two consecutive frame using perspective geometry statement\(^1\)

Ensures second order prior smoothness while preserving salient edges\(^2\)

Penalizes the deviation between acquired polarimetric data and normal deduced from computed disparity

---


Method

The model

Raw Polarimetric Image at t and t-1
Method

The model

Concatenated Polarimetric parameters
Method

The model

$\text{PoseCNN}$

$[R, t]_{t+1}$
The model

\[ \text{ PoseCNN } \]

\[ [R, t]_{t+1}^t \]

\[ \text{ UNet w/ Resnet50 Blocks } \]

\[ D_t \]
Method

* Multi-scale loss similar to Godard et al. approach through consecutive downscaling of the input
Results


P2D - Our approach - On polarimetric stack image

On Grayscale image

Our approach - On polarimetric stack image
## Results

### Benchmark

<table>
<thead>
<tr>
<th>Type</th>
<th>Network</th>
<th>Abs_Rel</th>
<th>Sq.Rel</th>
<th>RMSE</th>
<th>RMSE_log</th>
<th>$\delta &gt; 1.25$</th>
<th>$\delta &gt; 1.25^2$</th>
<th>$\delta &gt; 1.25^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>$G_{RGB}$</td>
<td>0.471</td>
<td>10.809</td>
<td>25.161</td>
<td>0.680</td>
<td>0.485</td>
<td>0.707</td>
<td>0.804</td>
</tr>
<tr>
<td></td>
<td>$G_I$</td>
<td>0.482</td>
<td>9.144</td>
<td>22.332</td>
<td>0.617</td>
<td>0.431</td>
<td>0.695</td>
<td>0.838</td>
</tr>
<tr>
<td></td>
<td>P2D</td>
<td><strong>0.322</strong></td>
<td><strong>4.504</strong></td>
<td><strong>20.651</strong></td>
<td><strong>0.484</strong></td>
<td><strong>0.537</strong></td>
<td><strong>0.801</strong></td>
<td><strong>0.896</strong></td>
</tr>
<tr>
<td>Cropped</td>
<td>$G_{RGB}$</td>
<td>0.533</td>
<td>14.050</td>
<td>29.312</td>
<td>0.780</td>
<td>0.449</td>
<td>0.658</td>
<td>0.771</td>
</tr>
<tr>
<td></td>
<td>$G_I$</td>
<td>0.415</td>
<td>11.247</td>
<td>25.899</td>
<td>0.678</td>
<td>0.467</td>
<td>0.729</td>
<td>0.850</td>
</tr>
<tr>
<td></td>
<td>P2D</td>
<td><strong>0.245</strong></td>
<td><strong>5.650</strong></td>
<td><strong>24.009</strong></td>
<td><strong>0.531</strong></td>
<td><strong>0.604</strong></td>
<td><strong>0.825</strong></td>
<td><strong>0.910</strong></td>
</tr>
<tr>
<td>Specular</td>
<td>$G_{RGB}$</td>
<td>0.341</td>
<td>8.249</td>
<td>7.236</td>
<td>0.306</td>
<td>0.666</td>
<td>0.808</td>
<td>0.896</td>
</tr>
<tr>
<td></td>
<td>$G_I$</td>
<td>0.208</td>
<td>2.248</td>
<td>5.491</td>
<td>0.233</td>
<td>0.639</td>
<td>0.877</td>
<td>0.952</td>
</tr>
<tr>
<td></td>
<td>P2D</td>
<td><strong>0.147</strong></td>
<td><strong>1.583</strong></td>
<td><strong>4.898</strong></td>
<td><strong>0.166</strong></td>
<td><strong>0.796</strong></td>
<td><strong>0.921</strong></td>
<td><strong>0.973</strong></td>
</tr>
</tbody>
</table>

### Sky Reconstruction

<table>
<thead>
<tr>
<th>Network</th>
<th>$G_{RGB}$</th>
<th>$G_I$</th>
<th>P2D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_s$</td>
<td>0.055</td>
<td>0.388</td>
<td><strong>0.532</strong></td>
</tr>
</tbody>
</table>
In Brief

A Preliminary Approach to Depth Estimation via a Polarimetric Monocular

✅ Extracting usable characteristics from polarimetry and infuse them in a model
  - Stacking strategy for Deep Learning usable image representation
  - Generalization through loss function
  - Abstraction of Fresnel equations (Light-to-surface relationships)
In Brief

A Preliminary Approach to Depth Estimation via a Polarimetric Monocular

✅ Extracting usable characteristics from polarimetry and infuse them in a model
- Stacking strategy for Deep Learning usable image representation
- Generalization through loss function
- Abstraction of Fresnel equations (Light-to-surface relationships)

✅ Create a network able to infer depth from monocular polarimetry
In Brief

A Preliminary Approach to Depth Estimation via a Polarimetric Monocular

- Extracting usable characteristics from polarimetry and infuse them in a model
  - Stacking strategy for Deep Learning usable image representation
  - Generalization through loss function
  - Abstraction of Fresnel equations (Light-to-surface relationships)

- Create a network able to infer depth from monocular polarimetry

- A better case-specific depth estimation and sky reconstruction
In Brief

A Preliminary Approach to Depth Estimation via a Polarimetric Monocular

✅ Extracting usable characteristics from polarimetry and infuse them in a model
  - Stacking strategy for Deep Learning usable image representation
  - Generalization through loss function
  - Abstraction of Fresnel equations (Light-to-surface relationships)

✅ Create a network able to infer depth from monocular polarimetry

✅ A better case-specific depth estimation and sky reconstruction

❌ Very few datasets and not enough groundtruth

❌ Some inconsistencies on the surfaces close to the sensor
What's next?

Using other source of information to reduce the complexity and improving accuracy
  ↪ Multi-modal / multi-space fusion (RGB, segmentation,..)
  ↪ Combining the benefits
  ↪ Improving accuracy and reducing the loss complexity
What’s next?

➡ Using other source of information to reduce the complexity and improving accuracy
   ↪ Multi-modal / multi-space fusion (RGB, segmentation,..)
   ↪ Combining the benefits
   ↪ Improving accuracy and reducing the loss complexity

➡ More and more usage of polarization as the sensor price decreases
Thank you
See you at PS T3.3

fr.marc.blanchon@gmail.com