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

Marc Blanchon[1*], Désiré Sidibé[2], Olivier Morel[1], Ralph Seulin[1], Daniel Braun[1], Fabrice Meriaudeau[1]

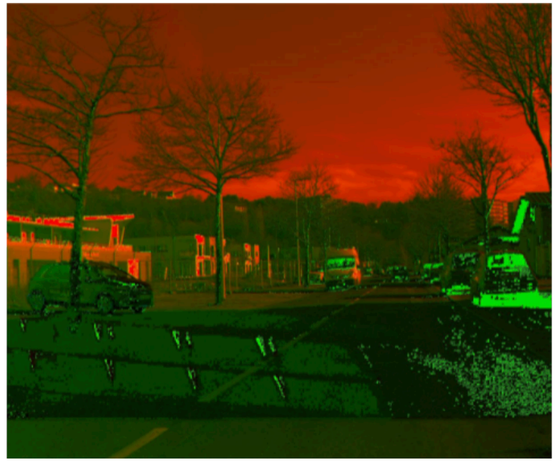
[1] ERL VIBOT CNRS 6000, ImViA EA 7535, UBFC, France
[2] IBISC, Univ Evry, France

*fr.marc.blanchon@gmail.com

25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Objective

Polarimetry to Depth (P2D)



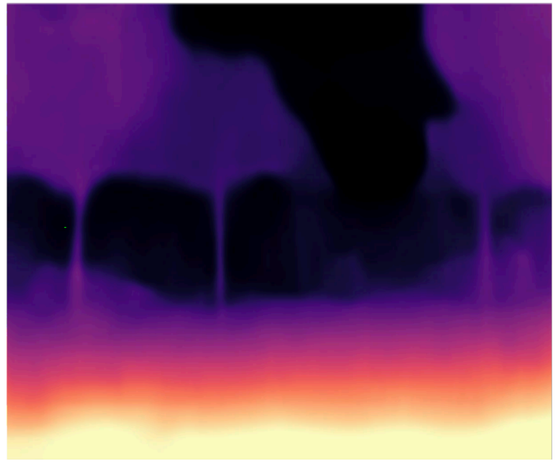
Input Polarimetric Image

- Generic

- Polarization adapted

- Transaprency / Specularity sensitive

Model



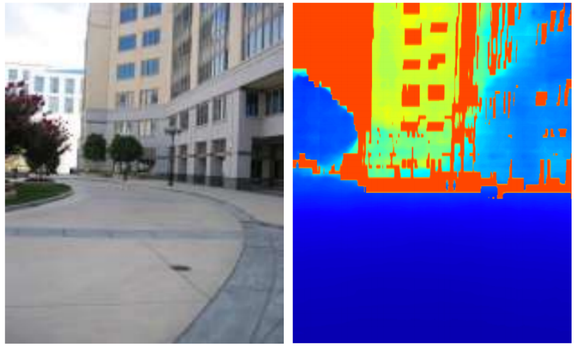
Depth

Previous approaches

Supervised Methods

Eigen et al.[1]

DORN[2]



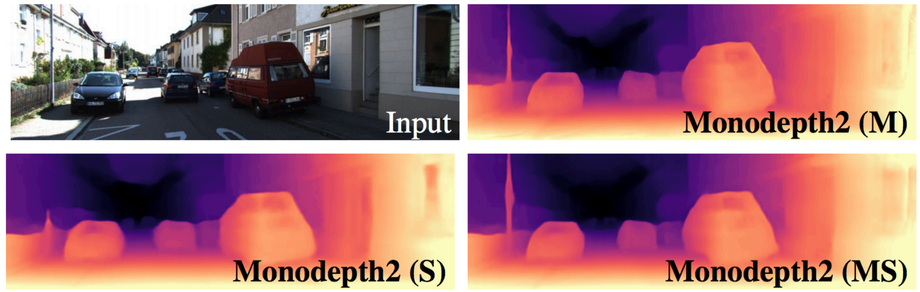
❌ **Required ground truth depth - No polarimetric dataset**
❌ **Lower performance compared to unsupervised / self-supervised**
❌ **Less generalization**

[1] David Eigen, Christian Puhrsch, and Rob Fergus. Depth map prediction from a single image using a multi-scale deep network. In NeurIPS, 2014.
[2] Huan Fu, Mingming Gong, Chaohui Wang, Kayhan Batmanghelich, and Dacheng Tao. Deep ordinal regression network for monocular depth estimation. In CVPR, 2018.

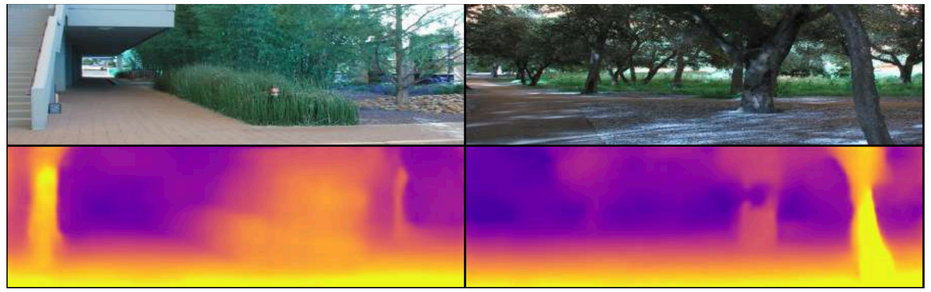
Previous approaches

Unsupervised Methods

Godard et al.[1]



EPC++[2]

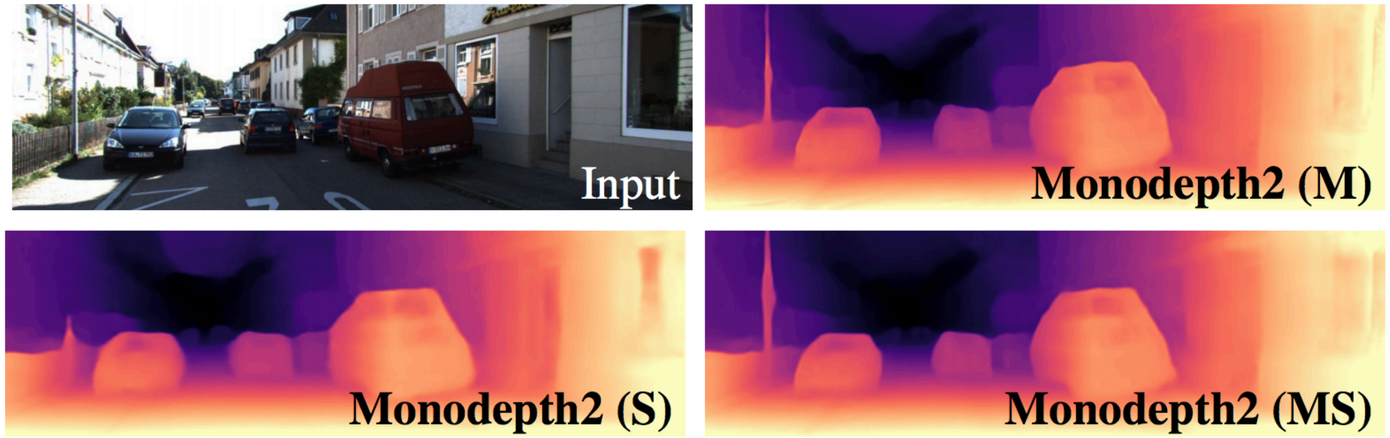


✅ **No required groundtruth depth**
✅ **Generalization through loss function**

❌ **Mainly colorimetric approaches**
❌ **Limitations of the modalities exported in the model - Specularity/transparency insensitive**

[1] Godard, Clément, et al. "Digging into self-supervised monocular depth estimation." *Proceedings of the IEEE international conference on computer vision*. 2019.
[2] Luo, Chenxu, et al. "Every pixel counts++: Joint learning of geometry and motion with 3d holistic understanding." IEEE transactions on pattern analysis and machine intelligence 42.10 (2019): 2624-2641.

Contextualization



Effective Case

Godard, Clément, et al. "Digging into self-supervised monocular depth estimation." *Proceedings of the IEEE international conference on computer vision*. 2019.

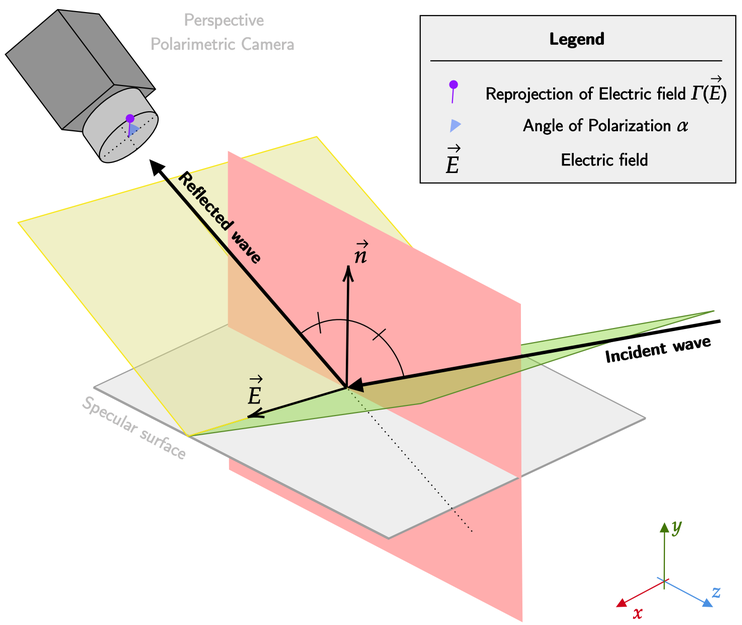
Contextualization

Failure Cases

Godard, Clément, et al. "Digging into self-supervised monocular depth estimation." *Proceedings of the IEEE international conference on computer vision*. 2019.

Method

Why Polarization?

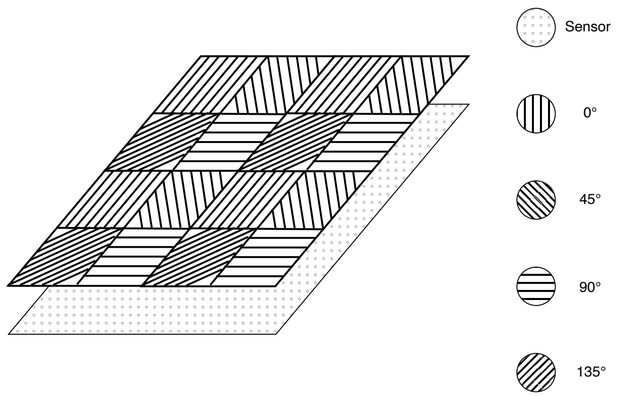


✅ **Sensitive to specularity**
✅ **Particular link between acquired polarization parameters and normal**

❌ **No colorimetric (strong differentiable features) space**
❌ **No generic approach leveraging the polarization to normal relationship**

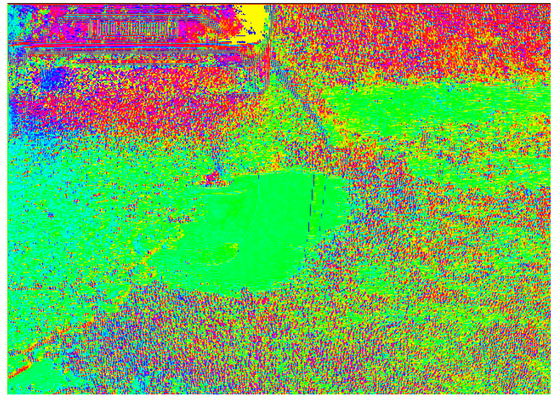
Method

What is Polarization?



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

ι

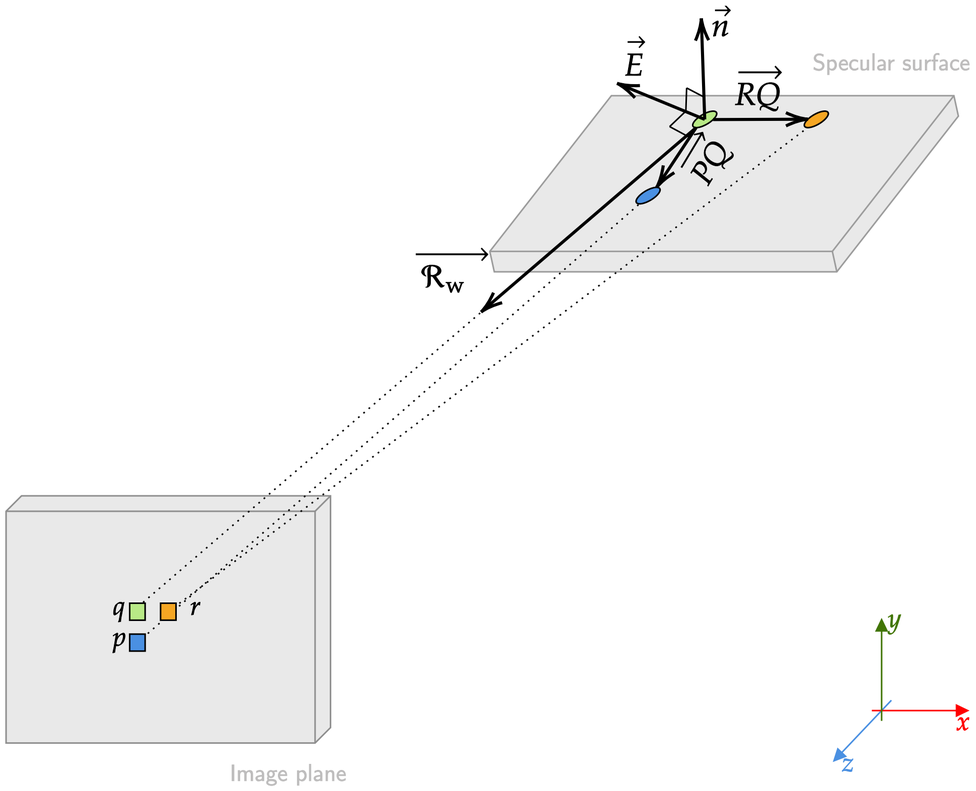


α

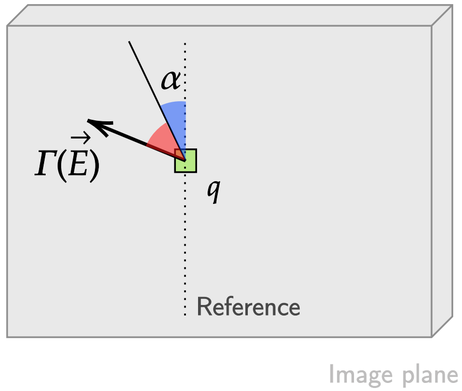
ρ

Method

Linking α to $\vec{n}$



Legend	
$\Gamma(\vec{E})$	Reprojection of Electric field
α	Angle of Polarization
	Angular difference



Method

A three term minimizable error evaluation

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

Reprojection Term

Smoothness Term

Polarimetric Term

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]

[1] Godard, Clément, et al. "Digging into self-supervised monocular depth estimation." *Proceedings of the IEEE international conference on computer vision*. 2019.

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]

[1] Godard, Clément, et al. "Digging into self-supervised monocular depth estimation." *Proceedings of the IEEE international conference on computer vision*. 2019.
[2] Woodford, Oliver, et al. "Global stereo reconstruction under second-order smoothness priors." IEEE transactions on pattern analysis and machine intelligence 31.12 (2009): 2115-2128.

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

[1] Godard, Clément, et al. "Digging into self-supervised monocular depth estimation." *Proceedings of the IEEE international conference on computer vision*. 2019.
[2] Woodford, Oliver, et al. "Global stereo reconstruction under second-order smoothness priors." IEEE transactions on pattern analysis and machine intelligence 31.12 (2009): 2115-2128.

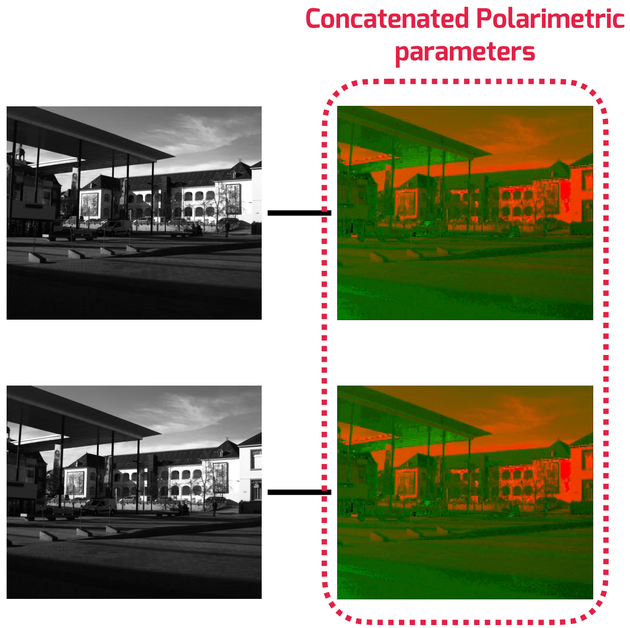
Method

The model

Raw Polarimetric Image
at t and t-1

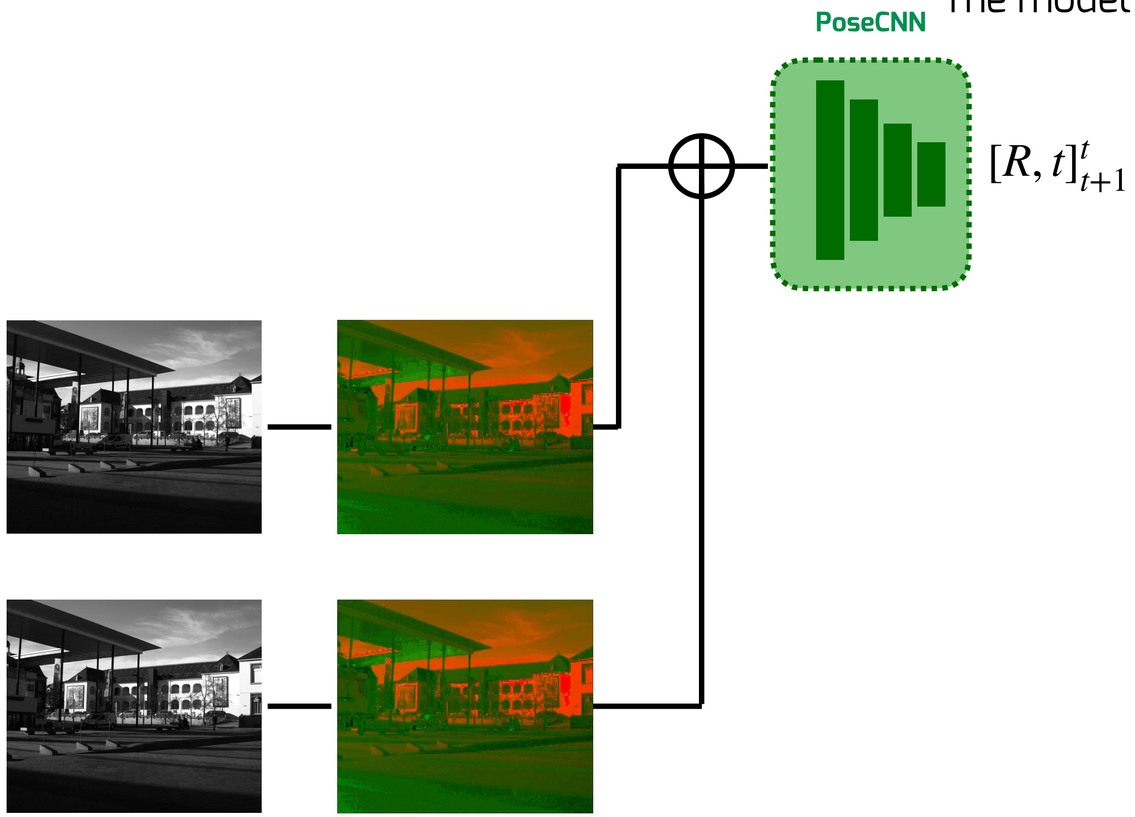
Method

The model



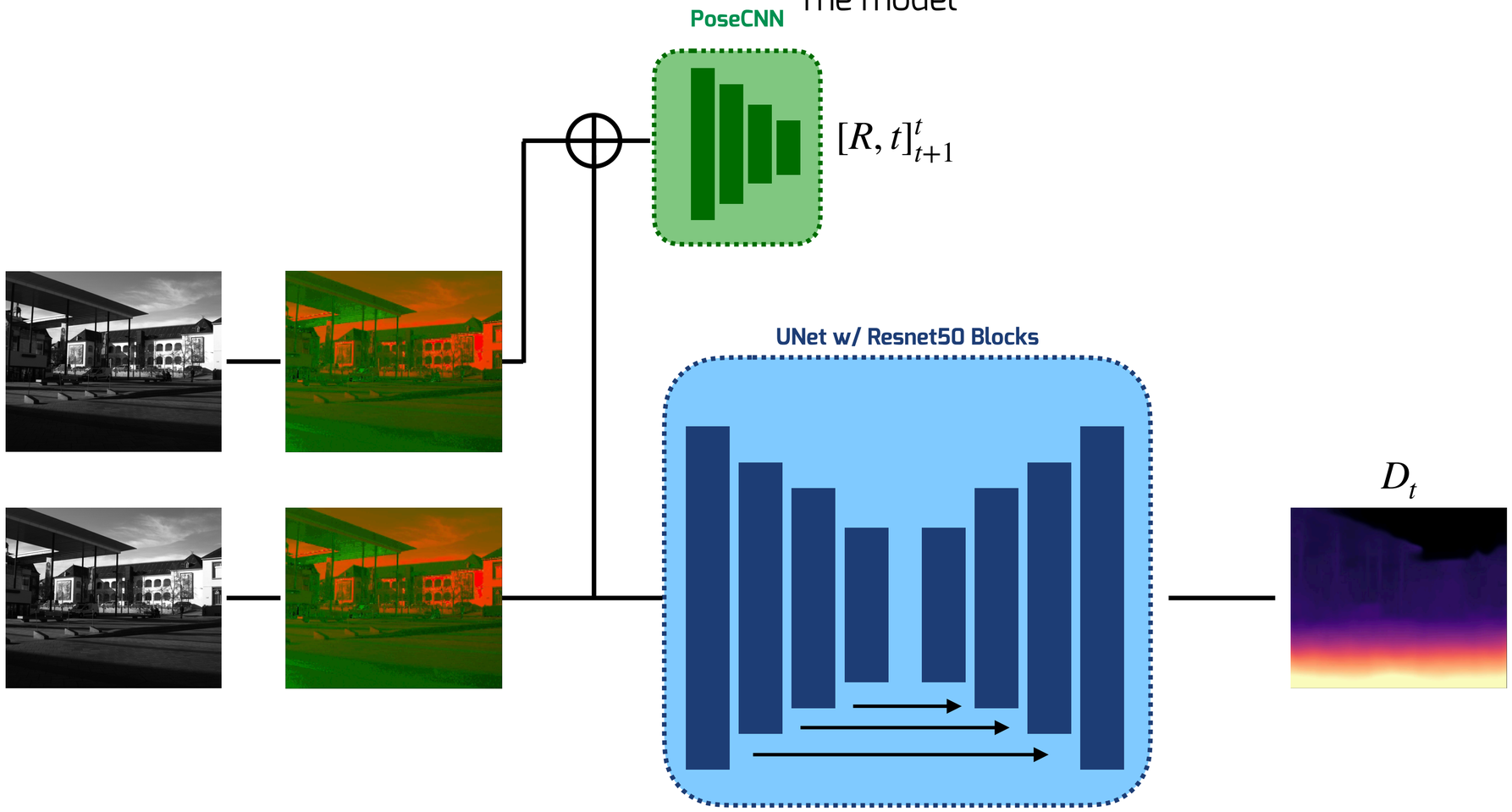
Method

The model



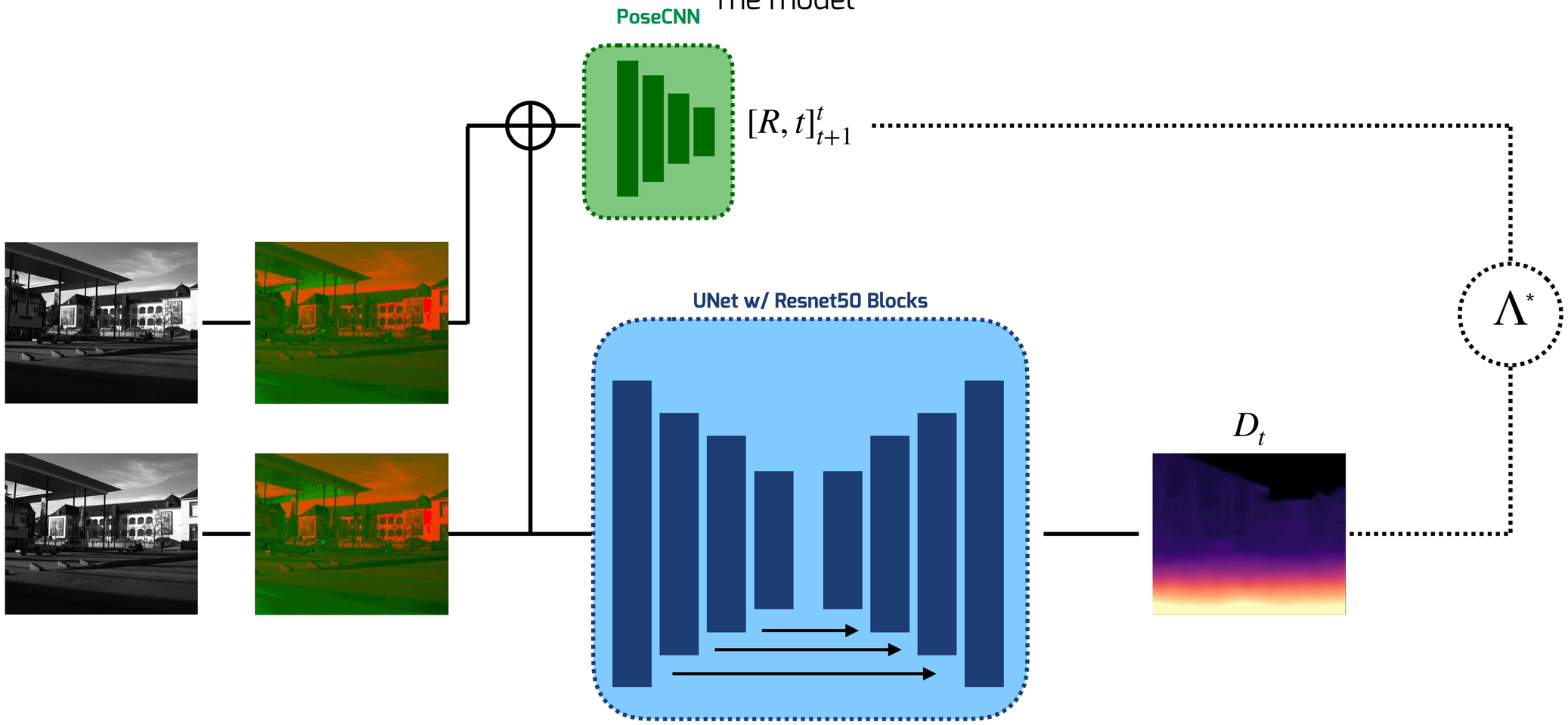
Method

The model



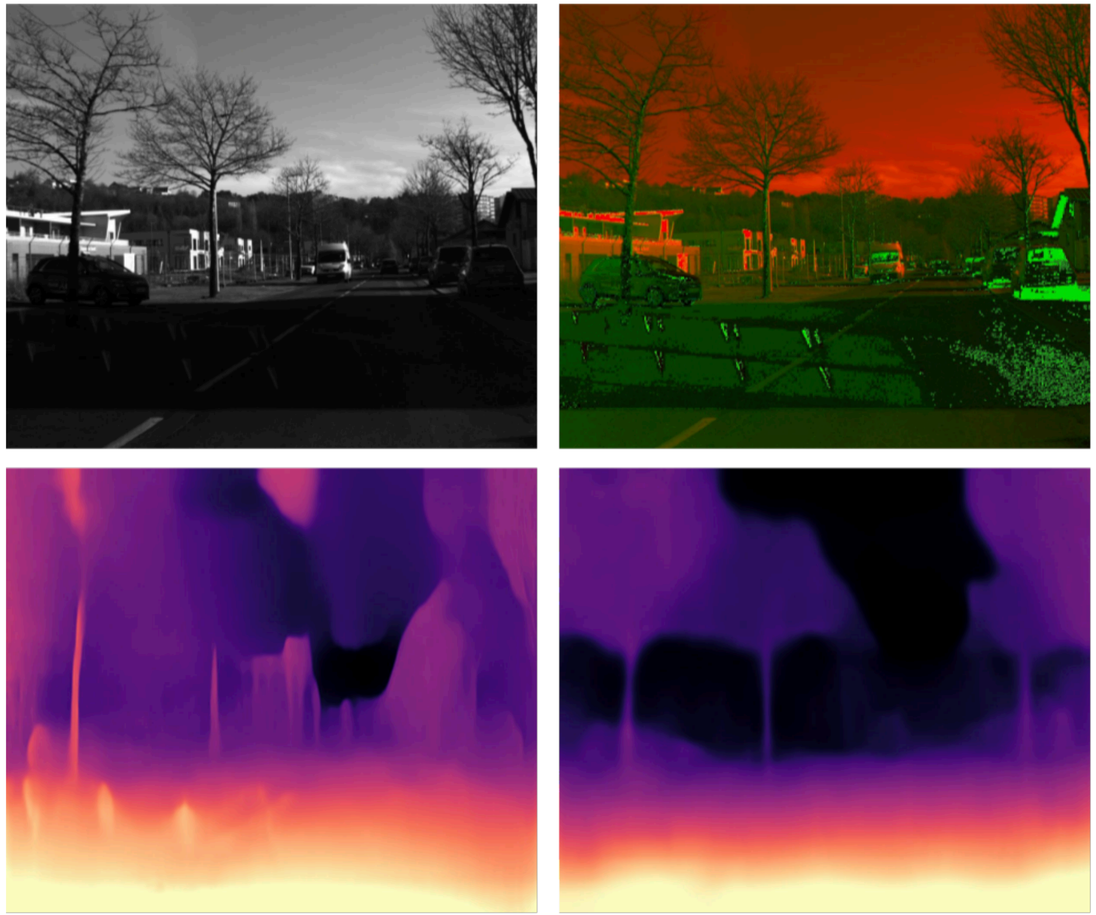
Method

The model



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

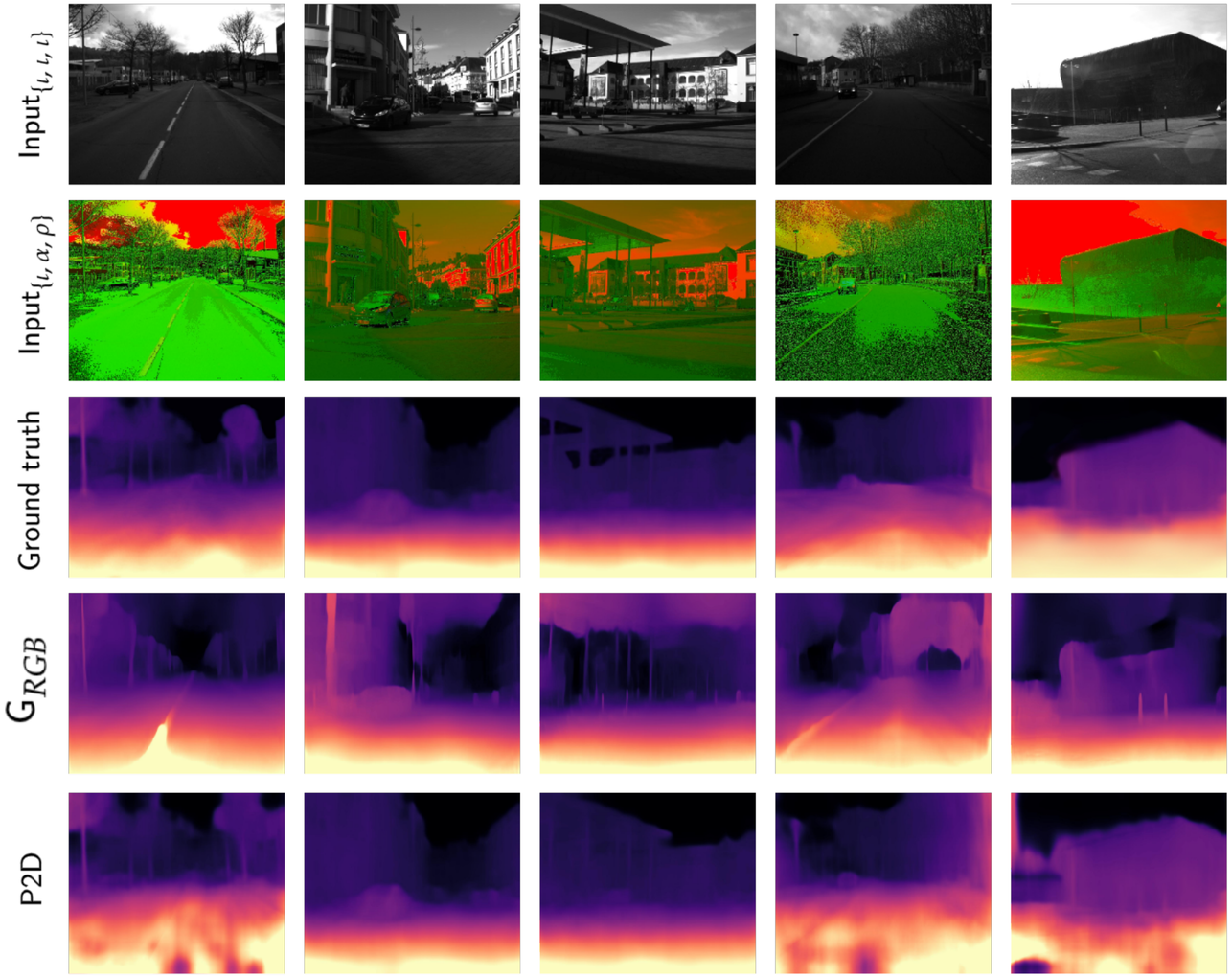
Results



G_{RGB} P2D

G_{RGB} - Godard, Clément, et al. "Digging into self-supervised monocular depth estimation." - *On Grayscale image*
P2D - Our approach - On polarimetric stack image

Results



G_{RGB} - Godard, Clément, et al. "Digging into self-supervised monocular depth estimation." - *On Grayscale image*
P2D - Our approach - On polarimetric stack image

Results

Benchmark

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

Sky Reconstruction

Network	G_{RGB}	G_I	P2D
R_s	0.055	0.388	**0.532**

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

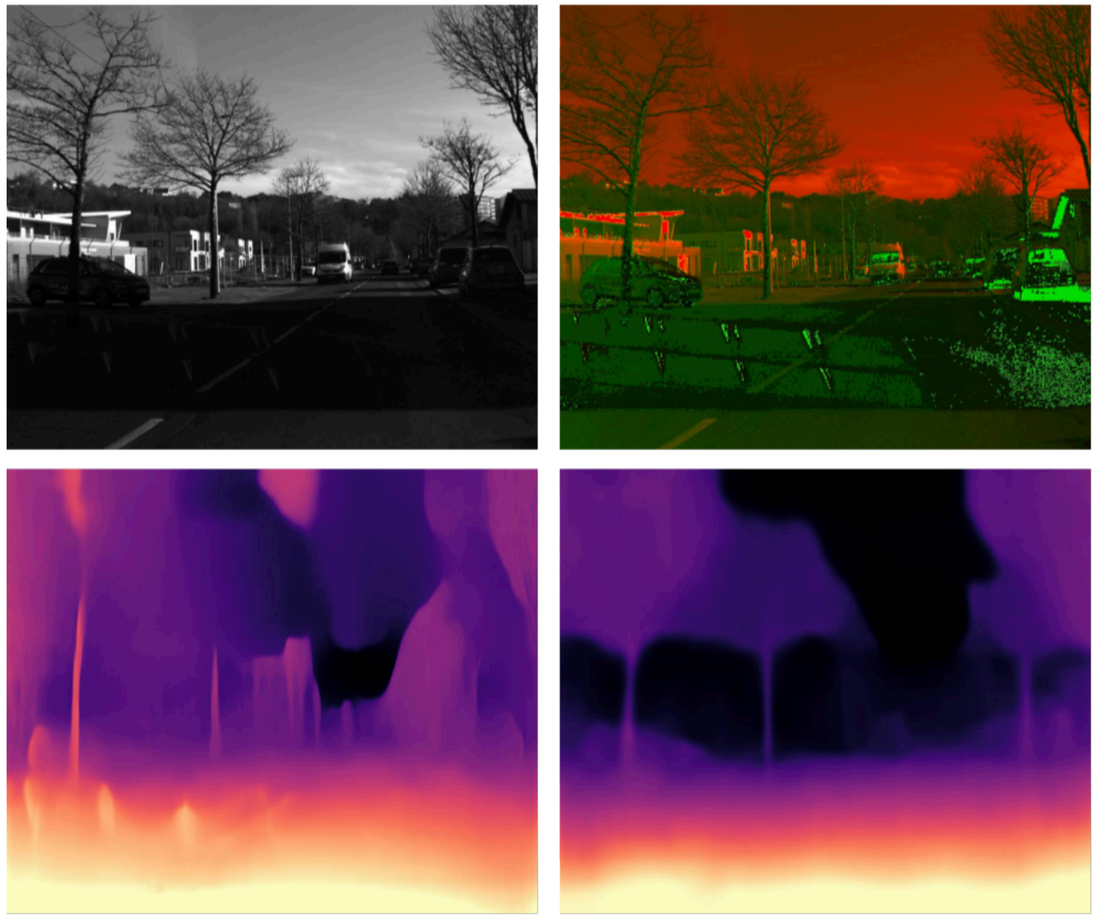
What's next?

➡ 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



G_{RGB} P2D