NOCaL: Calibration-Free Semi-Supervised Learning of Odometry and Camera Intrinsics

Ryan Griffiths, Jack Naylor, Donald G. Dansereau

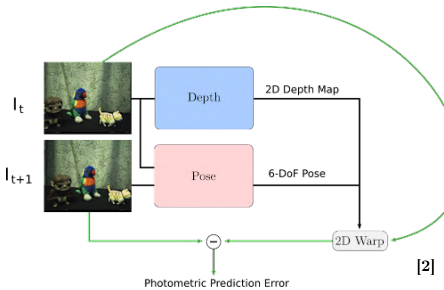
THE UNIVERSITY OF SYDNEY · ICRA LONDON 2023 · ROBOTIC IMAGING LAB · ACFR AUSTRALIAN CENTRE FOR FIELD ROBOTICS

Project Page & Code

ROBOTS & CAMERAS

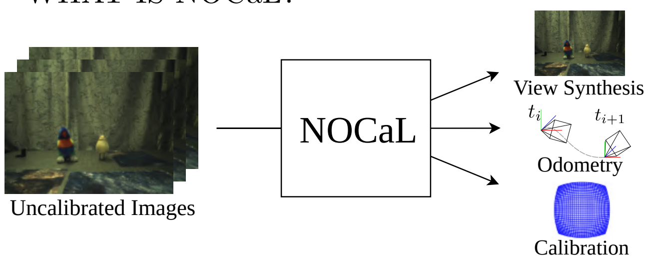
- Calibartion is not a solved problem.
- Camera charicteristics are **not static**.[1]
- Deploying emerging imaging technologies is hard.
- Calibrating fleets of robots can be an onerous task.

RELATED WORK



- Handcrafted warp requires rectified images.
- Handcrafted warp **limits** types of potential **cameras**.
- Struggles to explain complex light behavous and occulsions.

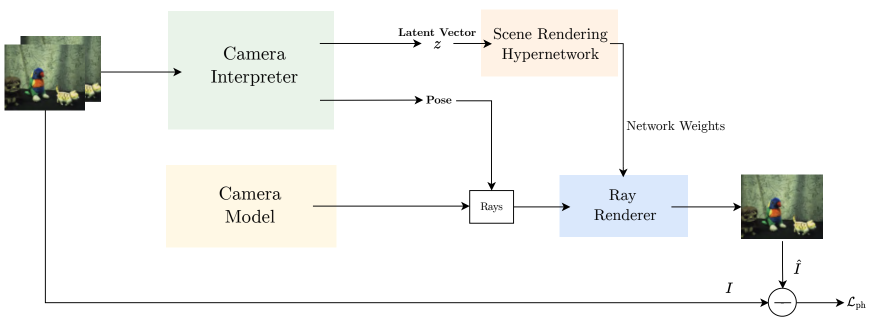
WHAT IS NOCaL?



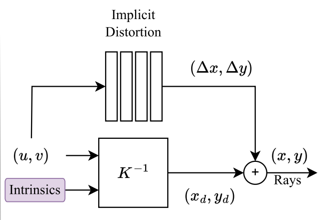
- Jointly learn **novel views**, **odometry** and **camera model**.
- Using **neural rendering** to provide **self-supervision**.
- **Hypernetwork** allows training across multiple scenes.

METHODOLOGY

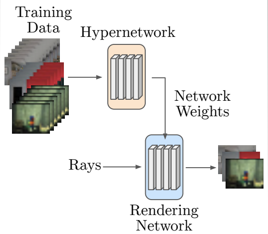
Network Architecture



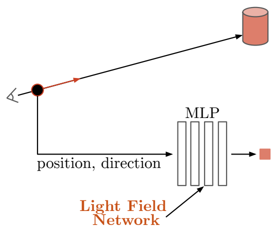
Camera Model



Scene Rendering Hypernetworks



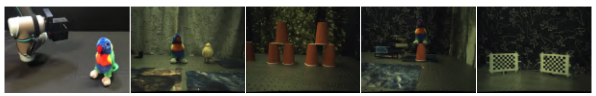
Ray Rendering [3]



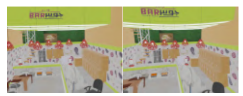
- **Camera Interpreter** - Distills input images into a pose and scene encoding.
- **Scene Rendering Hypernetwork** - Outputs the network weights of a ray render based on scene encoding.
- **Ray Renderer** - Implicit light field network used to render rays.
- **Camera Model** - First learn camera intrinsics using pinhole projection model. Then learn implicit freeform distortion model.

EXPERIMENTS

Captured Data [4]



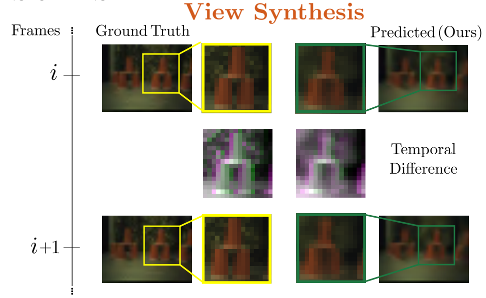
Synthetic Data



- Captured dataset with **multiple trajectories and scenes**.
- Captured with robotic arm giving accurate ground truth data.
- Synthetic data used to have **controllable** camera parameters.

RESULTS

View Synthesis



- The framework is able to learn the **3D structure**.
- This is a good indication that it can supervise odometry.

Camera Parameters

Camera	Initial Error		NOCaL (ours)		COLMAP [5]	
	f [px]	Error $\overline{\Delta r}$	f [px]	Error $\overline{\Delta r}$	f [px]	Error $\overline{\Delta r}$
Focal: 600px, no distortion	640.0	0.0225	601.1	0.0008	599.4	0.0003
Focal: 600px, large distortion	640.0	0.0440	595.3	0.0186	602.1	0.0010

- Good approximation the camera parameters.

Pose Estimation

Method	Labelled Images	Unlabelled Images	Translation Error RMSE [m]	Rotation Error RMSE [degrees]
Odometry accuracy on captured indoor imagery				
Fully supervised	800	0	0.027	2.414
Unlabelled calibrated [1]	0	8000	0.033	1.808
NOCaL (ours)	800	7200	**0.022**	**0.505**
Ablation study using rendered indoor imagery with camera distortion				
Ours no intrinsics or distortion	100	900	0.168	12.194
Ours no distortion	100	900	0.156	5.275
Ours full	100	900	**0.154**	**4.481**

- Outperform the unsupervised approach.
- Outperforms supervised, **benefiting** from unlabelled data.

CONCLUSION & FUTURE WORK

- **Uncalibrated** visual odometry in a **semi-supervised** fashion.
- In this work we take steps towards a general approach of **plug and play** sensors on robotic platforms.
- Generalise camera model.
- Separate scene and camera interpreters.

REFERENCES

[1] M. Smith and E. Cope, "The effects of temperature variation on single-lens-reflex digital camera calibration parameters," International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 38, no. Part 5, 2010.

[2] T. Zhou, M. Brown, N. Snavely, and D. G. Lowe, "UnsupervisedLearning of Depth and Ego-motion from Video," CVPR, 2017, pp. 1851–1858.

[3] V. Sitzmann, S. Rezchikov, W. T. Freeman, J. B. Tenenbaum, andF. Durand, "Light Field Networks: Neural Scene Representations with Single-Evaluation Rendering," NeurIPS, 2021.

[4] S. T. Digumarti, J. Daniel, A. Ravendran, R. Griffiths, and D. G.Dansereau, "Unsupervised Learning of Depth Estimation and VisualOdometry for Sparse Light Field Cameras," IROS. IEEE, 2021

[5] J. L. Sch¨onberger and J.-M. Frahm, "Structure-from-Motion Revis-ited," CVPR, 2016