Exploiting Illumination Decline

Advances in Monocular Visual SLAM, Self-Supervised Depth Estimation, and Watertight 3D Reconstruction

Universidad Zaragoza
EndoMapper
Instituto Universitario de Investigación en Ingeniería de Aragón
Universidad Zaragoza

IROS KYOTO 2022
IEEE/RSJ International Conference on Intelligent Robots and Systems

Real-scale 3D reconstruction from monocular endoscope images

Víctor M. Batlle, J.M.M. Montiel, Juan D. Tardós
Univeridad de Zaragoza, Spain
IEEE/RSJ International Conference on Intelligent Robots and Systems

Reconstruction
Outside vs. Inside the human body

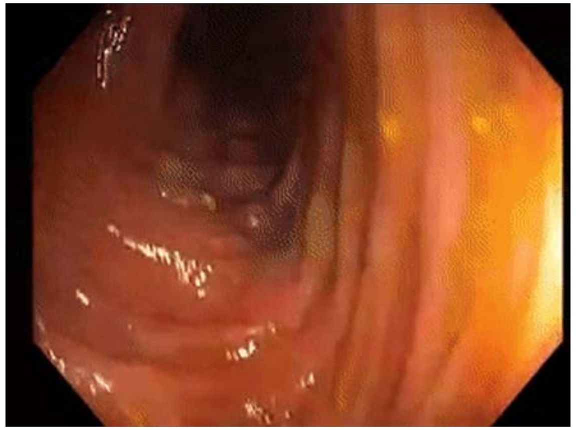
<u>Outside</u>

[KITTI (2013)]

- Stereo images

<u>Inside</u>

[EndoMapper (2022)]



- Monocular images
  ⚠️ Scale problem

[26] A. Geiger et al. (2013). *Vision meets Robotics: The KITTI Dataset.*
[25] P. Azagra et al. (2022). *EndoMapper dataset of complete calibrated endoscopy procedures.*

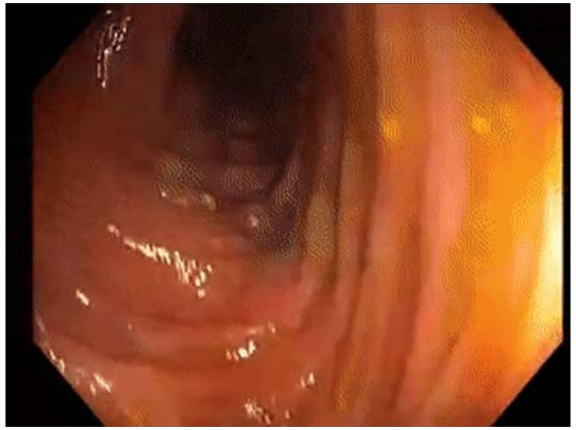
Reconstruction
Outside vs. Inside the human body

Outside

[KITTI (2013)]

Inside

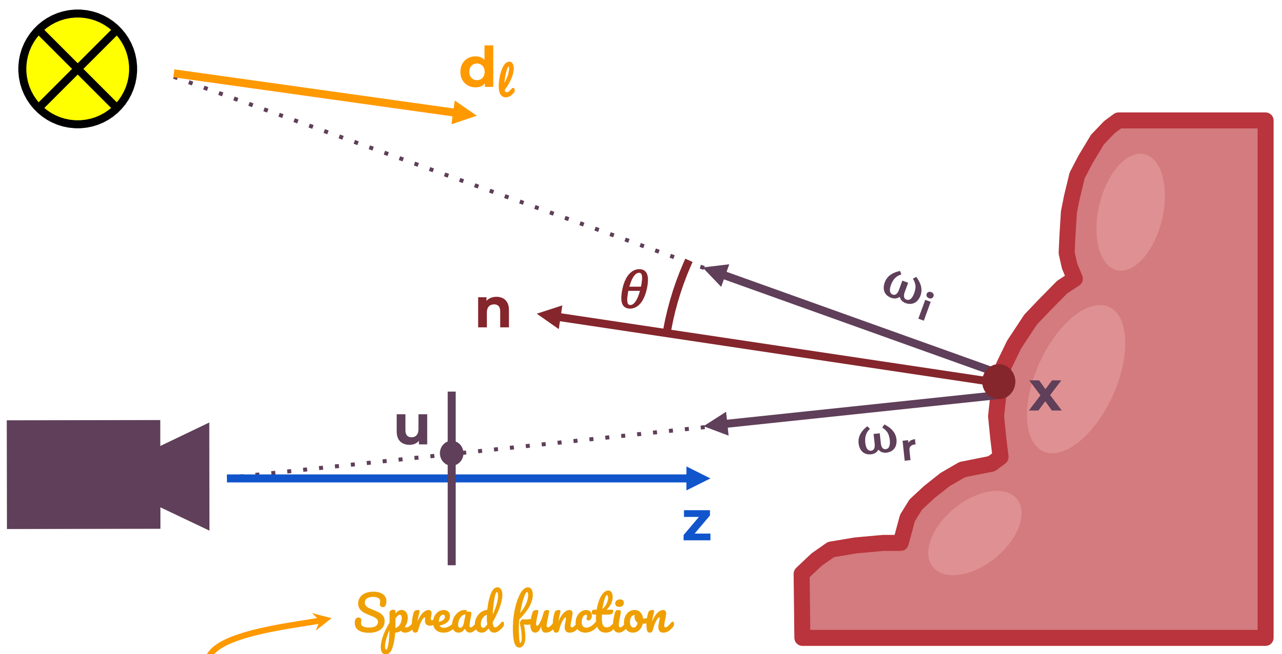
[EndoMapper (2022)]



Previously, lighting changes were **dismissed…**

- Assuming **constant illumination**
- Using **invariant feature points**

General photometric model



Max. radiance

Spread function

BRDF: Bidirectional Reflectance Distribution Function

Vignetting

$$\mathcal{I}(\mathbf{x}) = \left(\frac{\quad}{\quad} \right)$$

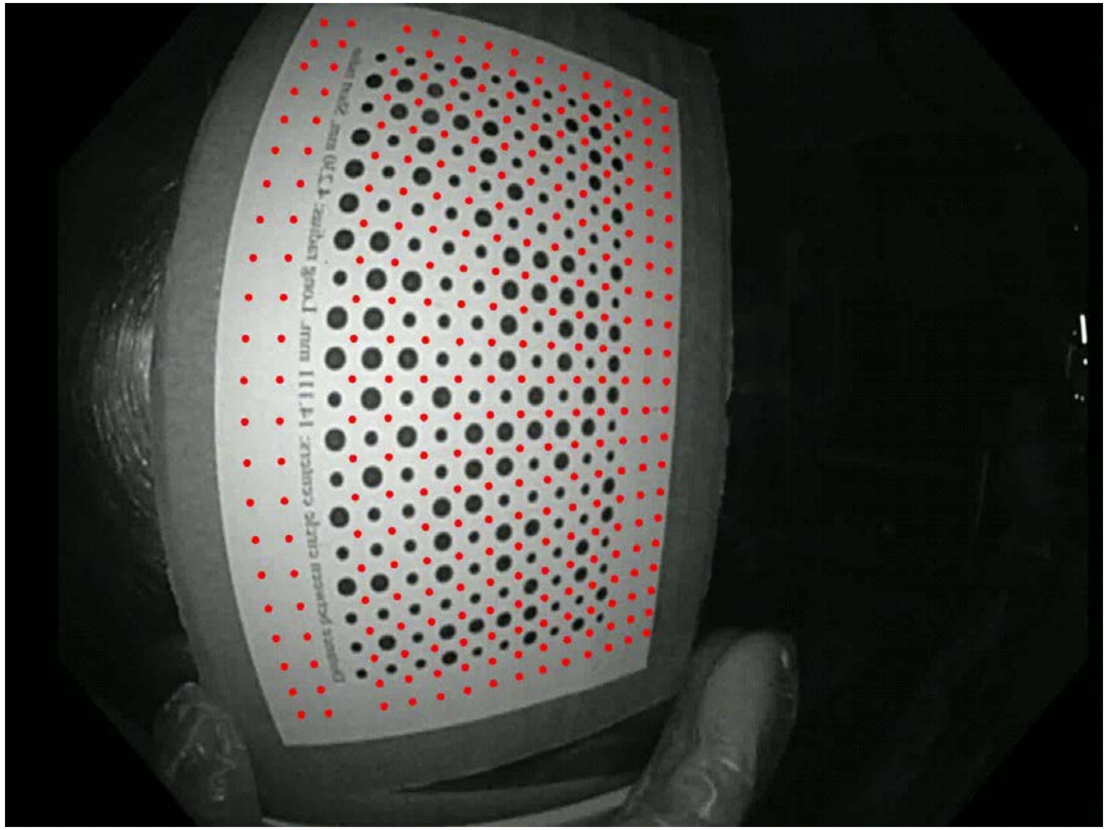
Gamma

Distance

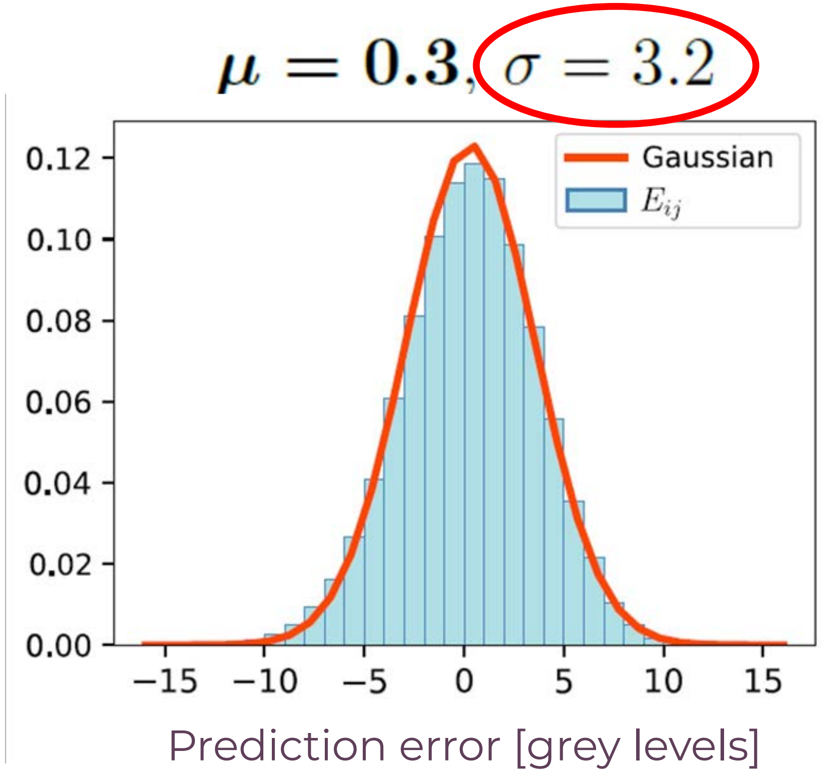
Cosine term

Gain

Endoscope calibration



[25] P. Azagra et al. (2022). *EndoMapper dataset of complete calibrated endoscopy procedures.*



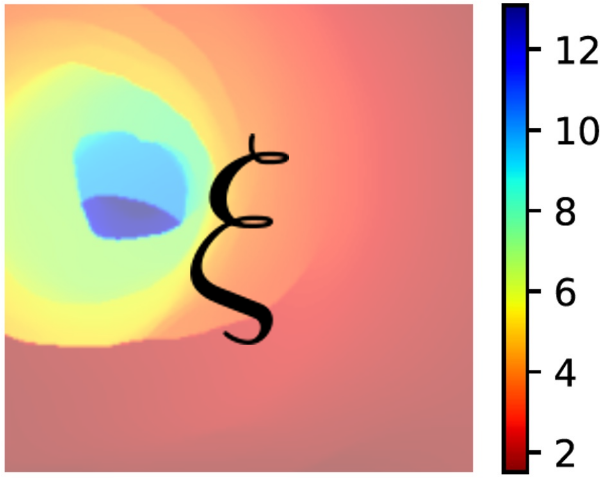
Depth estimation

[A. Rau (2022)]

Monocular image

$\mathcal{I}$

Optimisation



Dense depth map [cm]

Smoothness regularisation

Photometric error

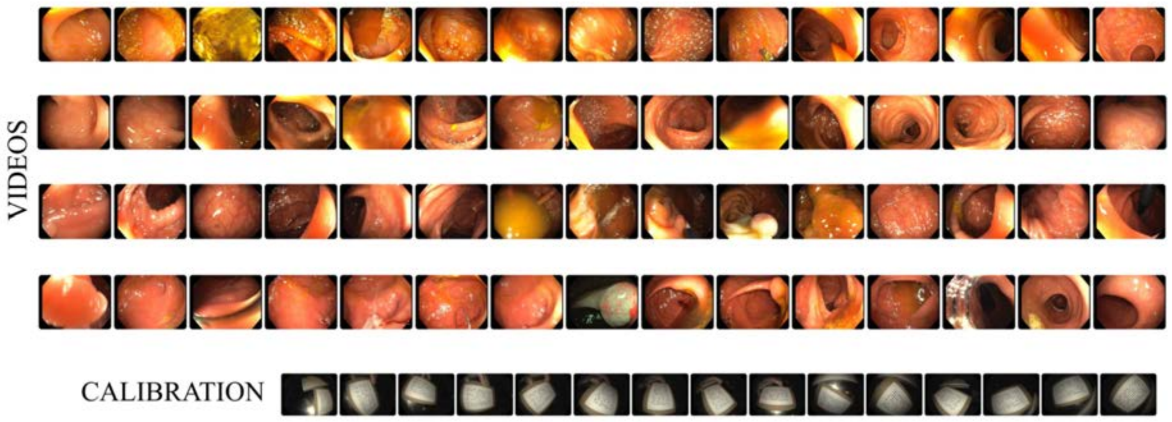
$$\rho(\mathcal{I} - I) + \underbrace{\lambda \, ||\nabla\xi||_\epsilon \cdot e^{-\alpha||\nabla I||_2^\beta}}$$

Huber loss

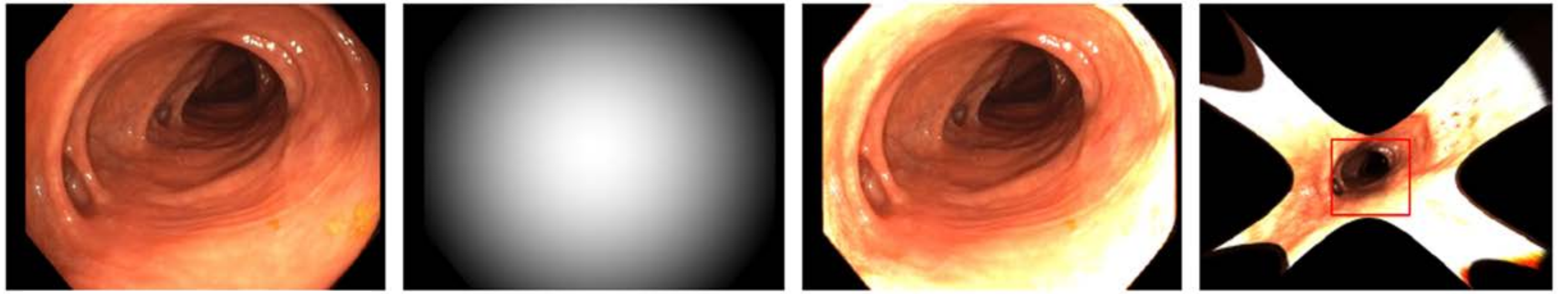
Allow discontinuities

[10] R. A. Newcombe et al. (2011) *DTAM: Dense Tracking and Mapping in Real-Time.*

Real colon dataset

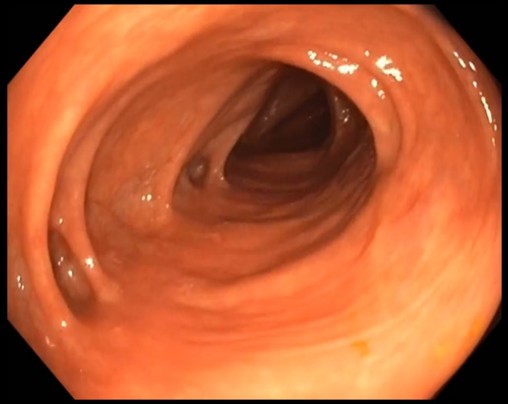


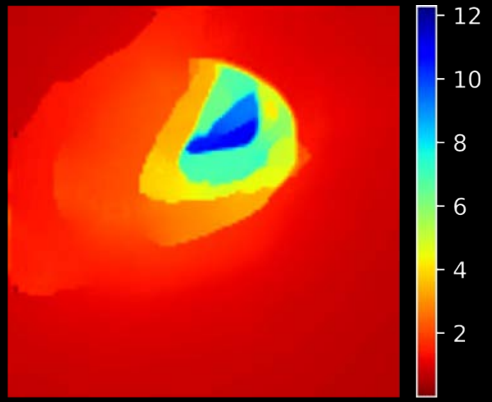
EndoMapper

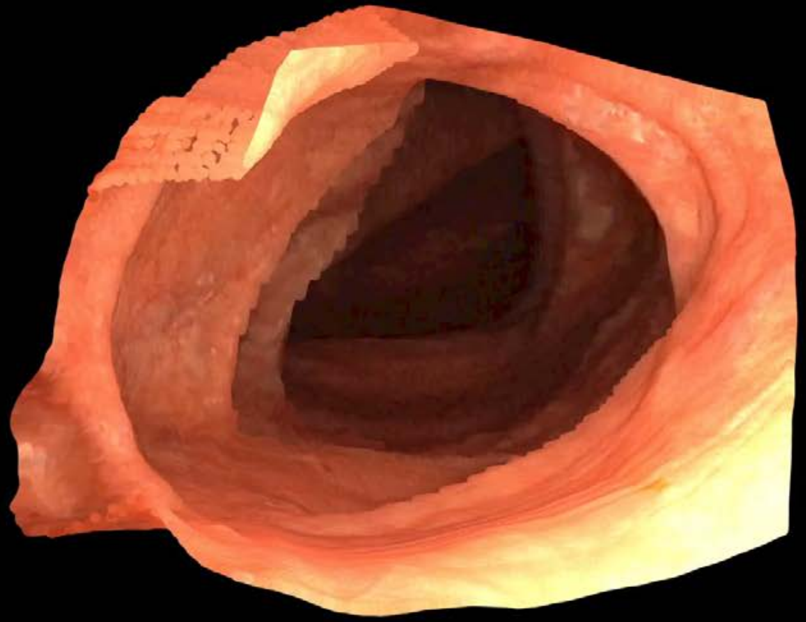


(a) Input frame (b) Vignetting & $\mu(\mathbf{x})$ (c) Canonical (d) Undistorted

[25] P. Azagra et al. (2022). *EndoMapper dataset of complete calibrated endoscopy procedures.*







3D reconstruction

ICCV23

PARIS

LightDepth: Single-View Depth Self-Supervision from Illumination Decline

Javier Rodríguez-Puigvert, Víctor M. Batlle, J.M.M. Montiel,
Ruben Martinez Cantin, Pascal Fua, Juan D. Tardós, Javier Civera

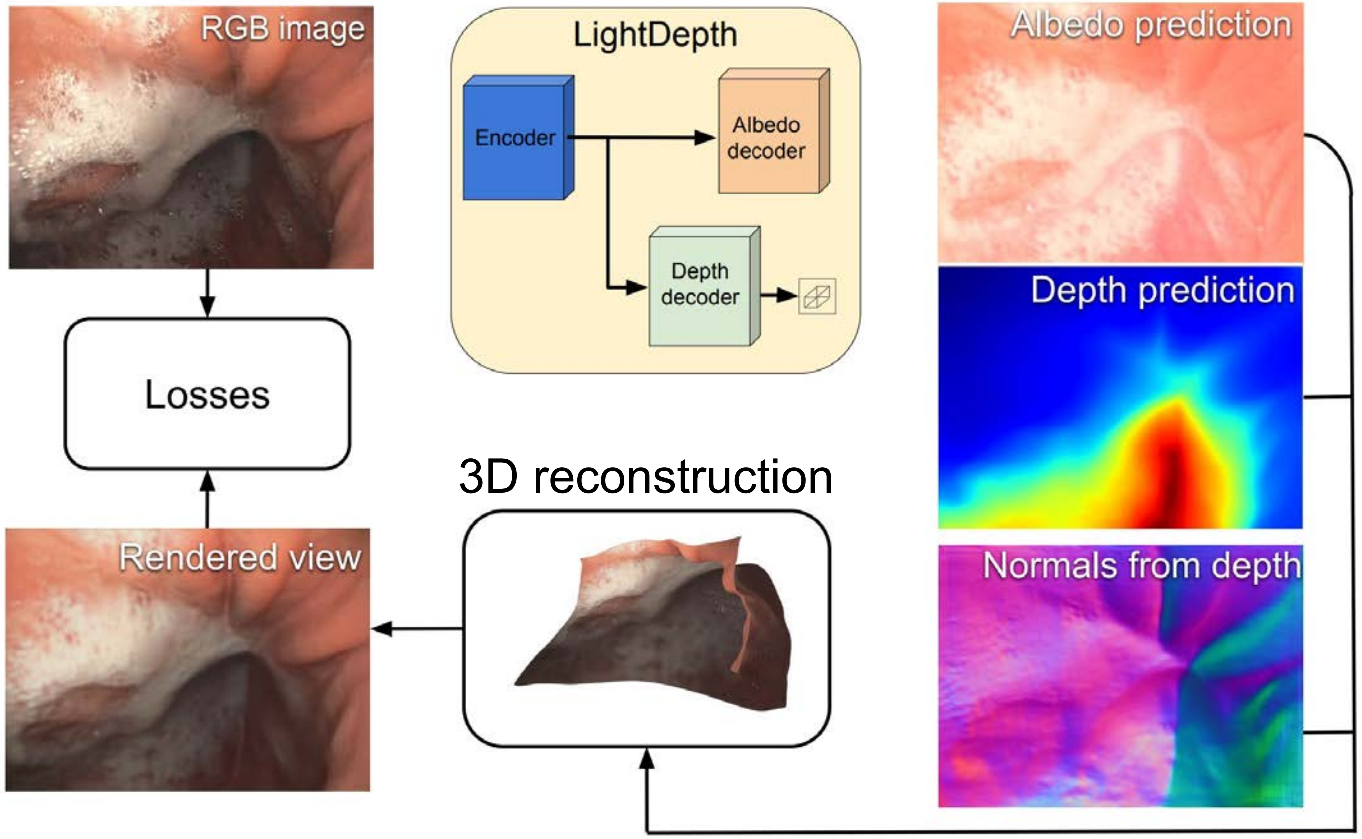
Universidad Zaragoza
1542

EndoMapper

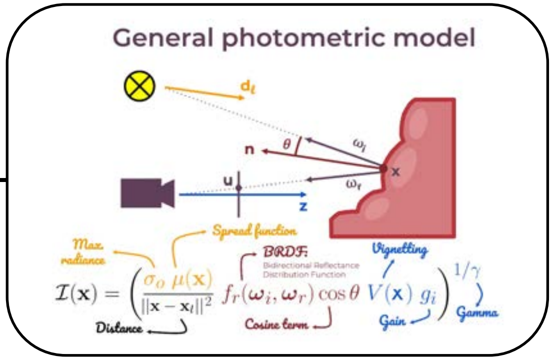
Instituto Universitario de Investigación en Ingeniería de Aragón
Universidad Zaragoza

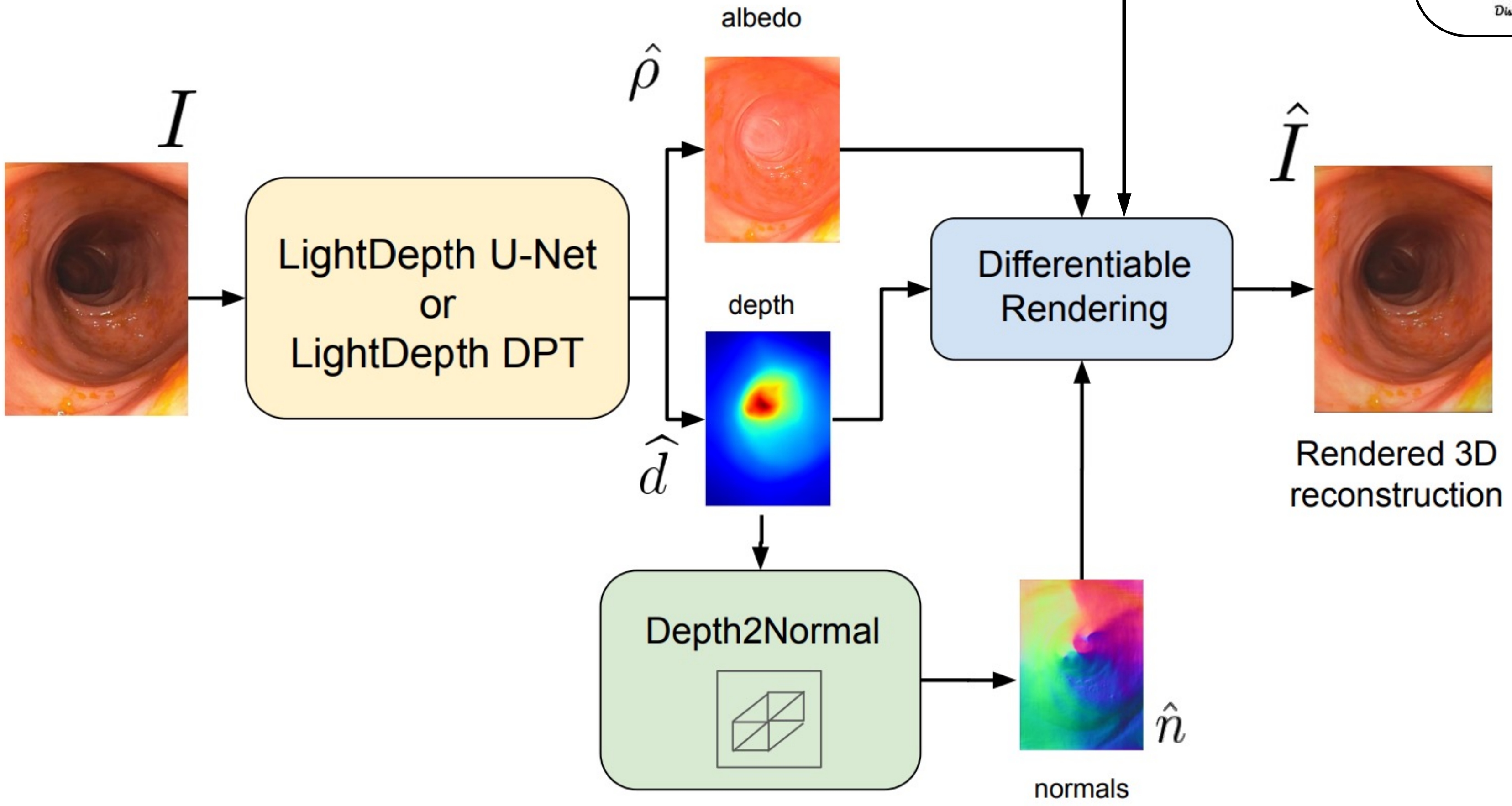
EPFL

Single-view self-supervision

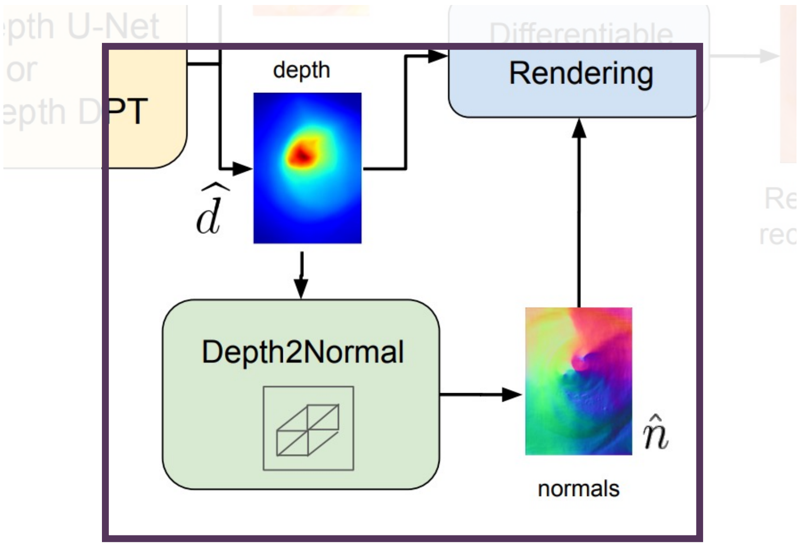


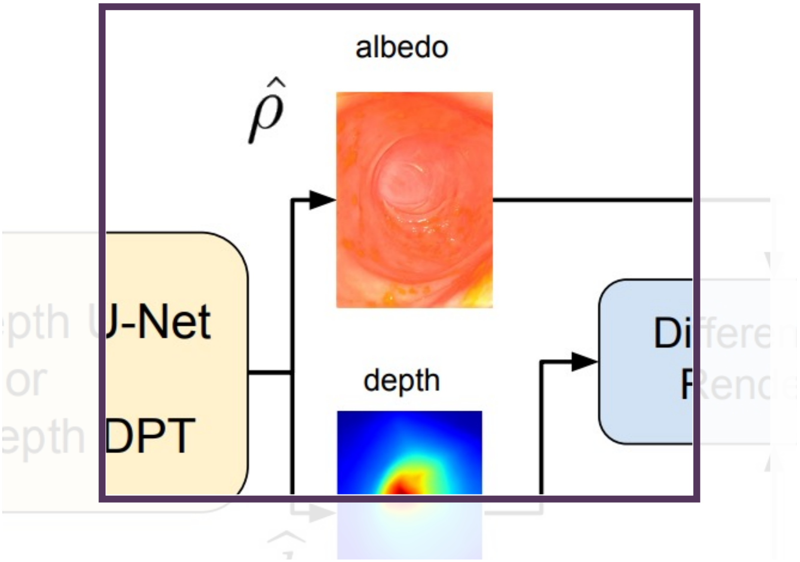
LightDepth

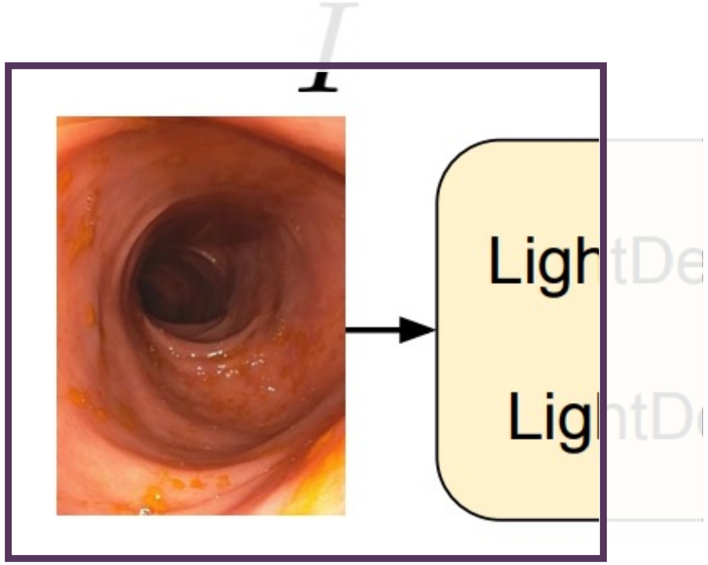

General photometric model
$\mathbf{d}_\ell$
θ
ω_i
n
u
x
ω_r
z
Spread function
Max. radiance
BRDF:
Bidirectional Reflectance Distribution Function
Vignetting
$$\mathcal{I}(\mathbf{x}) = \left(\frac{\sigma_o\,\mu(\mathbf{x})}{\|\mathbf{x}-\mathbf{x}_\ell\|^2}\, f_r(\boldsymbol{\omega}_i, \boldsymbol{\omega}_r)\cos\theta\; V(\mathbf{x})\; g_i\right)^{1/\gamma}$$
Distance
Cosine term
Gain
Gamma


I
LightDepth U-Net
or
LightDepth DPT
albedo
$\hat{\rho}$
depth
$\hat{d}$
Differentiable Rendering
$\hat{I}$
Rendered 3D reconstruction
Depth2Normal
$\hat{n}$
normals

Key ideas



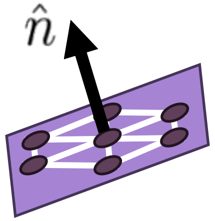


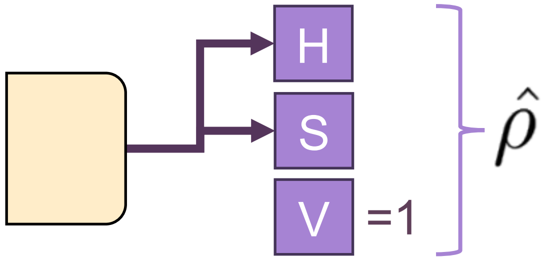


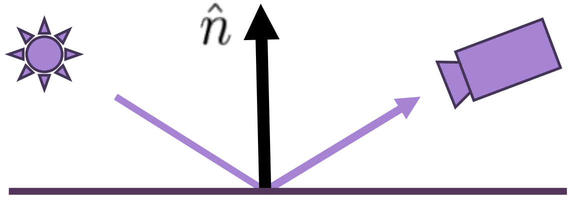
① Normals from Depth

② Albedo prediction

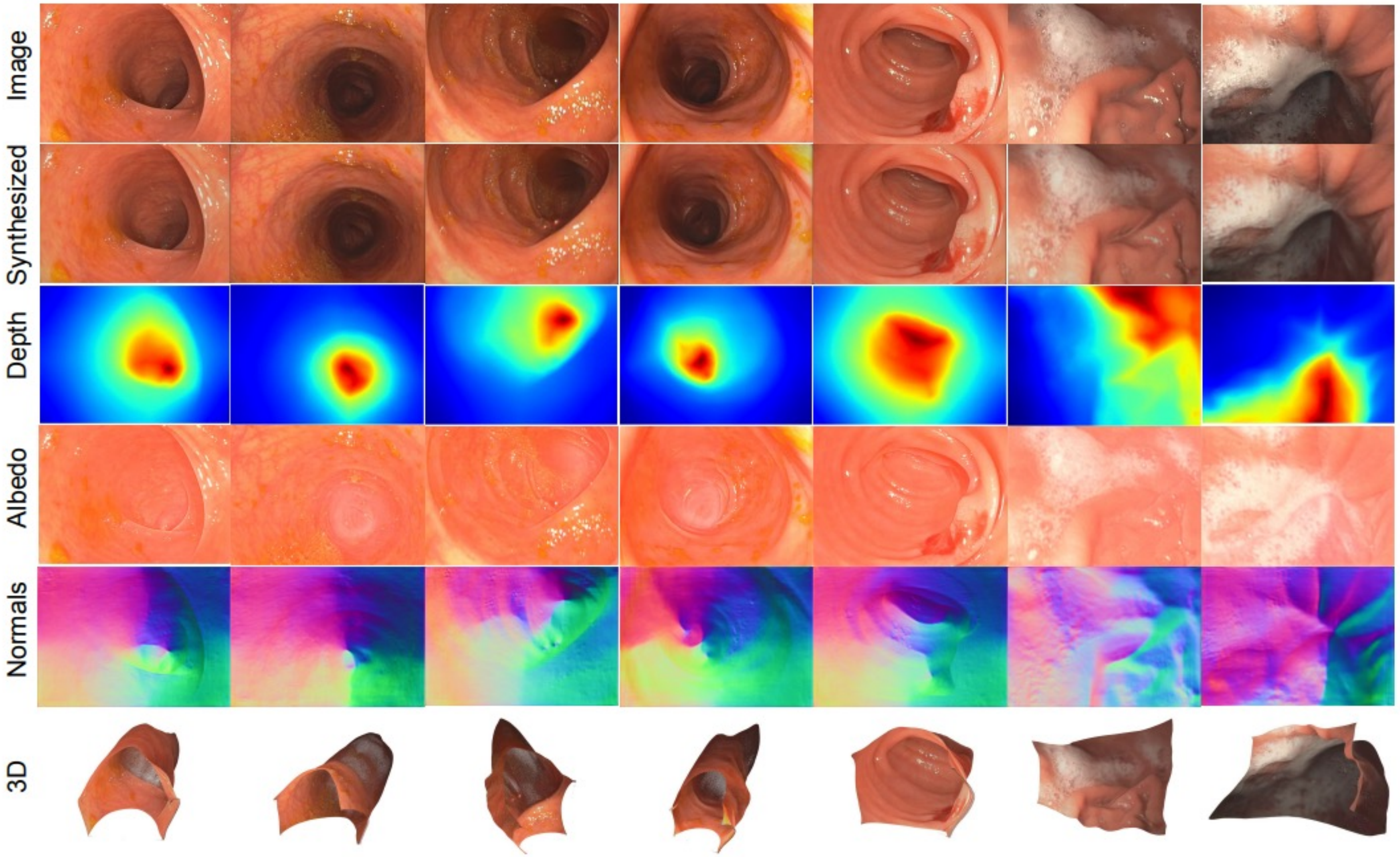
③ Normals from highlights

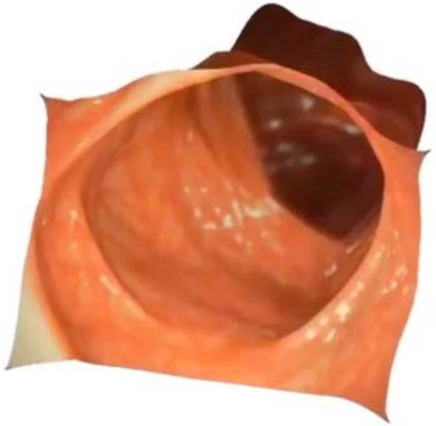






Results inside the human colon





Comparison

Dataset	Architecture	Backbone	Supervision	Depth [mm] MAE ↓	MedAE ↓	RMSE ↓	$\text{RMSE}_{\log}$ ↓	Abs_{Rel} ↓	Sq_{Rel} ↓	$\delta < 1.25$ ↑	$\delta < 1.25^2$ ↑	$\delta < 1.25^3$ ↑	Normals [°] MAE ↓
Synthetic	U-Net	ResNet18	Depth GT	**4.37**	2.99	**6.38**	**0.1251**	0.0965	**0.0008**	0.9057	**0.9931**	**0.9997**	25.1
	LightDepth U-Net	ResNet18	Light	4.76	**2.47**	8.60	0.1375	**0.0903**	0.0011	**0.9180**	0.9820	0.9935	**15.2**
C3VD	U-Net	ResNet18	Depth GT	4.15	3.29	5.52	0.1139	0.0902	0.0007	0.9172	0.9943	**0.9994**	26.5
	DPT-Hybrid [48]	ResNet50	Depth GT	**3.22**	2.77	**4.10**	**0.0860**	**0.0699**	**0.0004**	**0.9640**	0.9865	0.9913	**15.1**
	Monodepth2 [20]	ResNet50	Multi-View	14.27	9.59	18.64	0.3921	0.2971	0.0070	0.4897	0.7313	0.8611	43.6
	CADepth [64]	ResNet18	Multi-View	52.35	17.04	87.43	0.9144	1.1916	0.2650	0.3664	0.5653	0.6679	67.2
	XDCycleGAN [42]	ResNet	Cycle	17.16	11.91	22.43	0.4953	0.3616	0.0105	0.4291	0.6615	0.7910	64.4
	LightDepth U-Net	ResNet18	Light	4.37	2.92	6.31	0.1183	0.0856	0.0007	0.9315	0.9934	**0.9994**	24.0
	LightDepth DPT	ResNet50	Light	3.94	2.67	5.60	0.1080	0.08046	0.0006	0.9476	0.9965	**0.9994**	21.3
	LightDepth U-Net	ResNet18	Light (TTR)	3.72	2.59	5.43	0.1060	0.0770	0.0005	0.9505	**0.9971**	**0.9994**	23.5
	LightDepth DPT	ResNet50	Light (TTR)	3.70	**2.58**	5.27	0.1073	0.0780	0.0005	0.9525	0.9961	0.9992	22.5

Table 1. Depth and normal metrics for several architectures and supervision modes. Best results per dataset are bolfaced, second best underlined.

- LightDepth (TTR) is the closest to Depth GT supervision

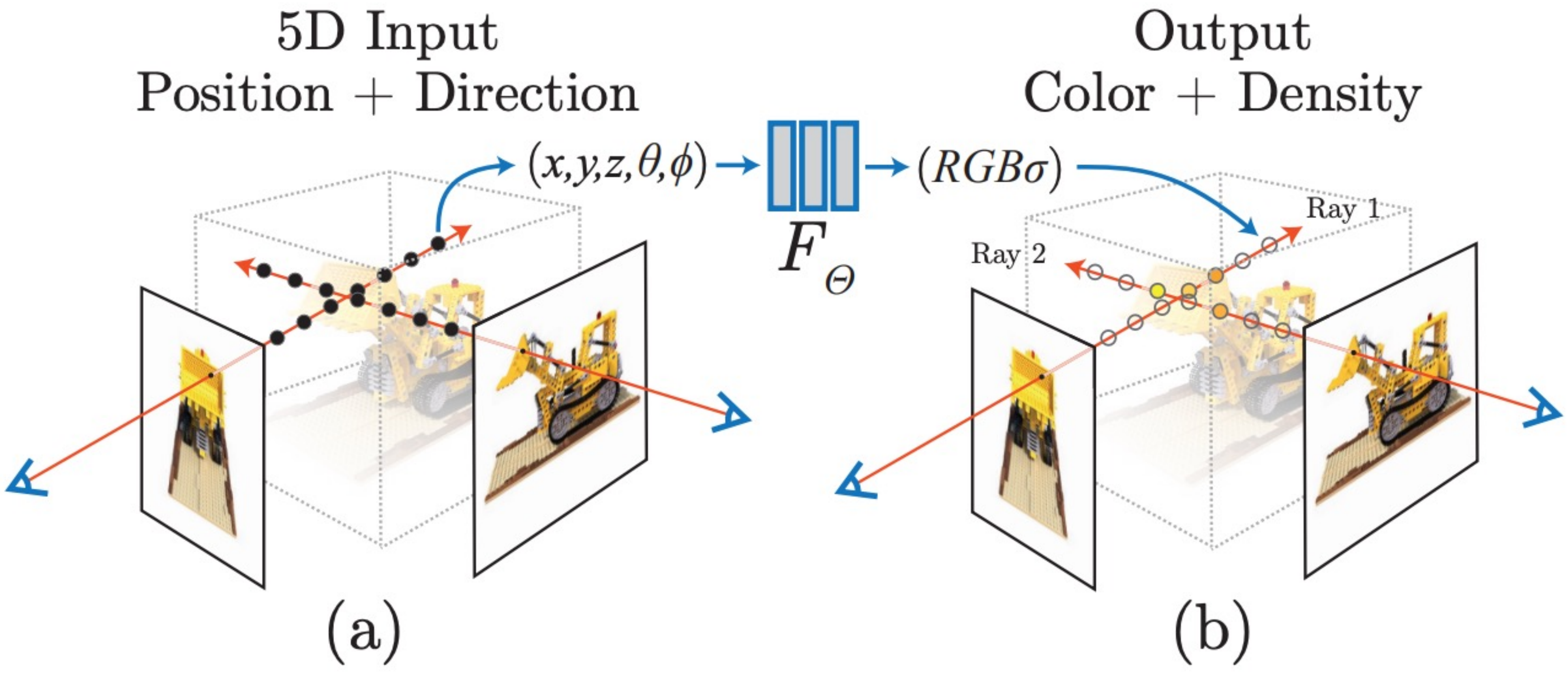
LightNeus: Neural Surface Reconstruction in Endoscopy using Illumination Decline

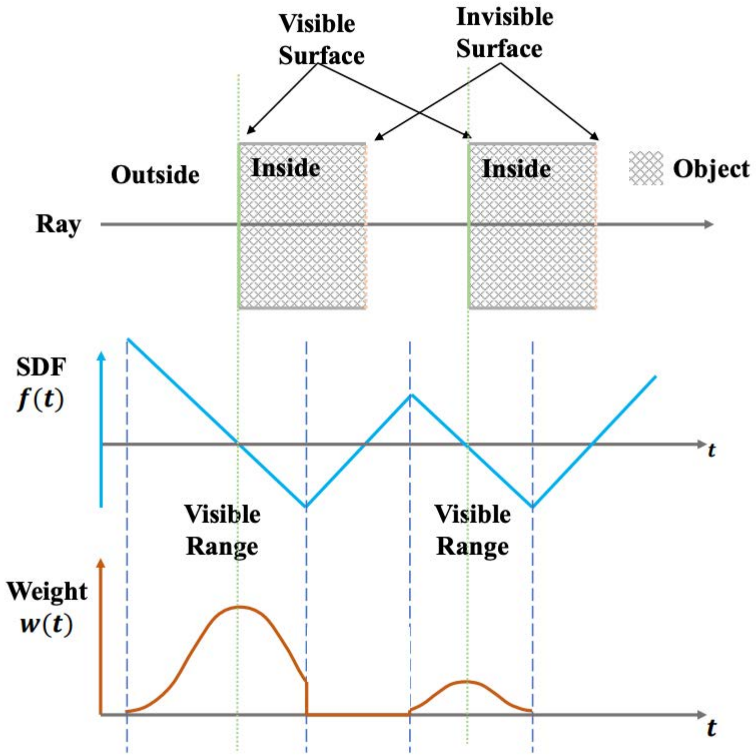
Víctor M. Batlle, José M. M. Montiel, Pascal Fua, and Juan D. Tardós

What if we want...

- A **dense** model
- From a **whole section**
- **Globally** optimized

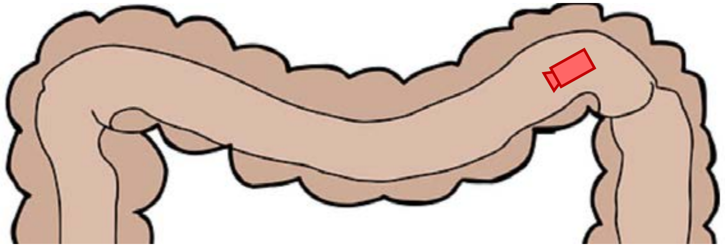
What’s NeuS?





Mildenhall et al. (2020). NeRF: Representing scenes as neural radiance fields for view synthesis.
Wang et al. (2021). NeuS: Learning neural implicit surfaces by volume rendering for multi-view reconstruction.

NeRF inside the human colon? How?



Key ideas

- **Watertight** surfaces.
- Co-located **illumination decline.**

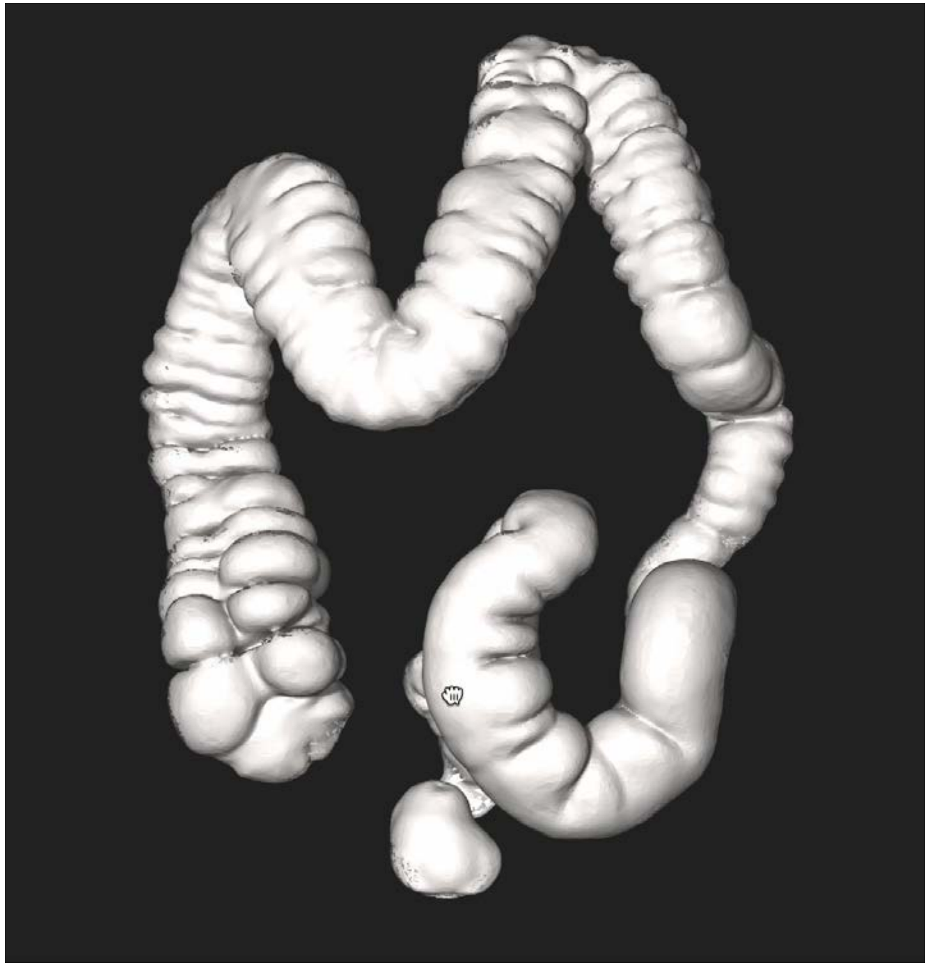
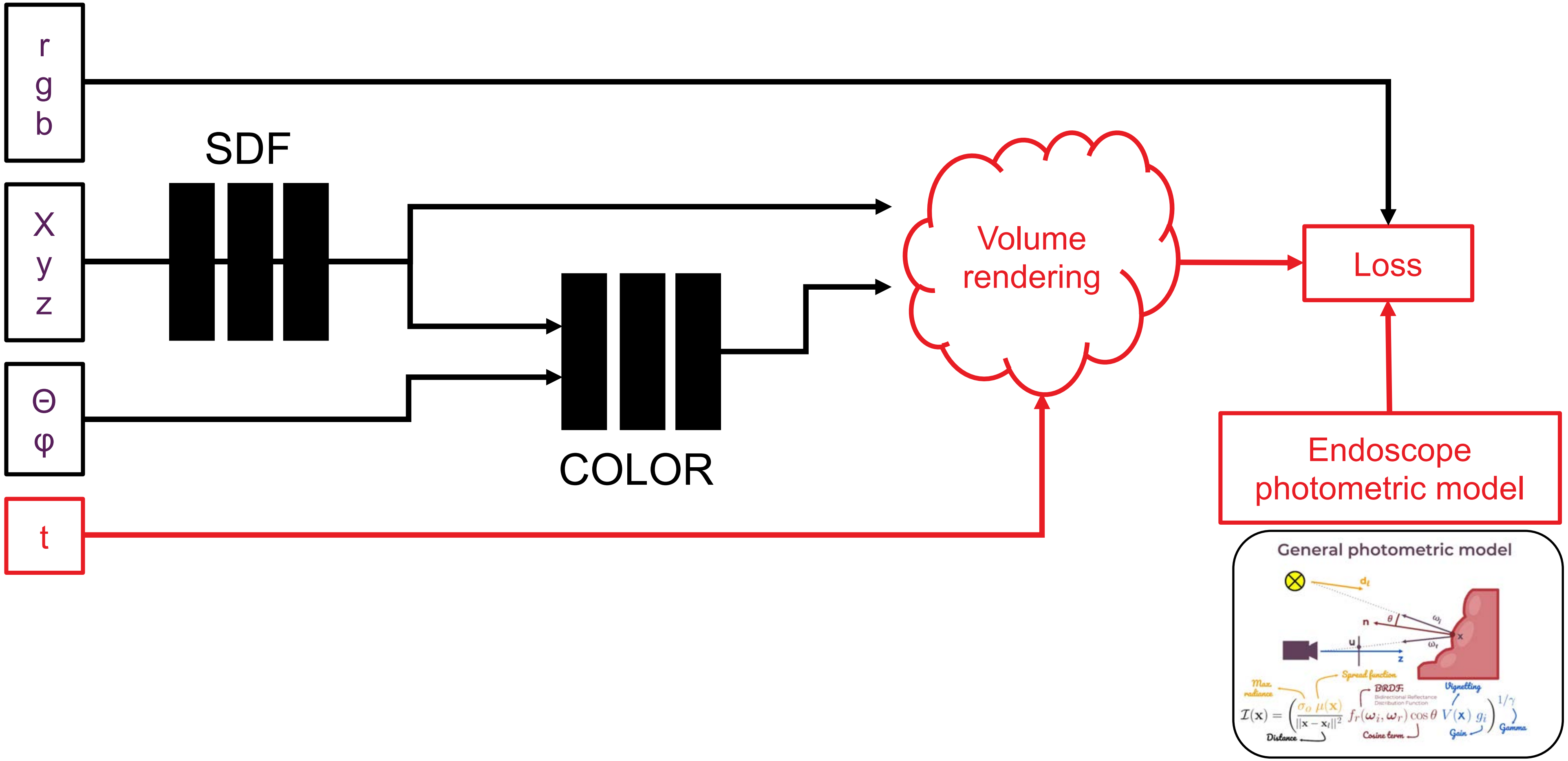
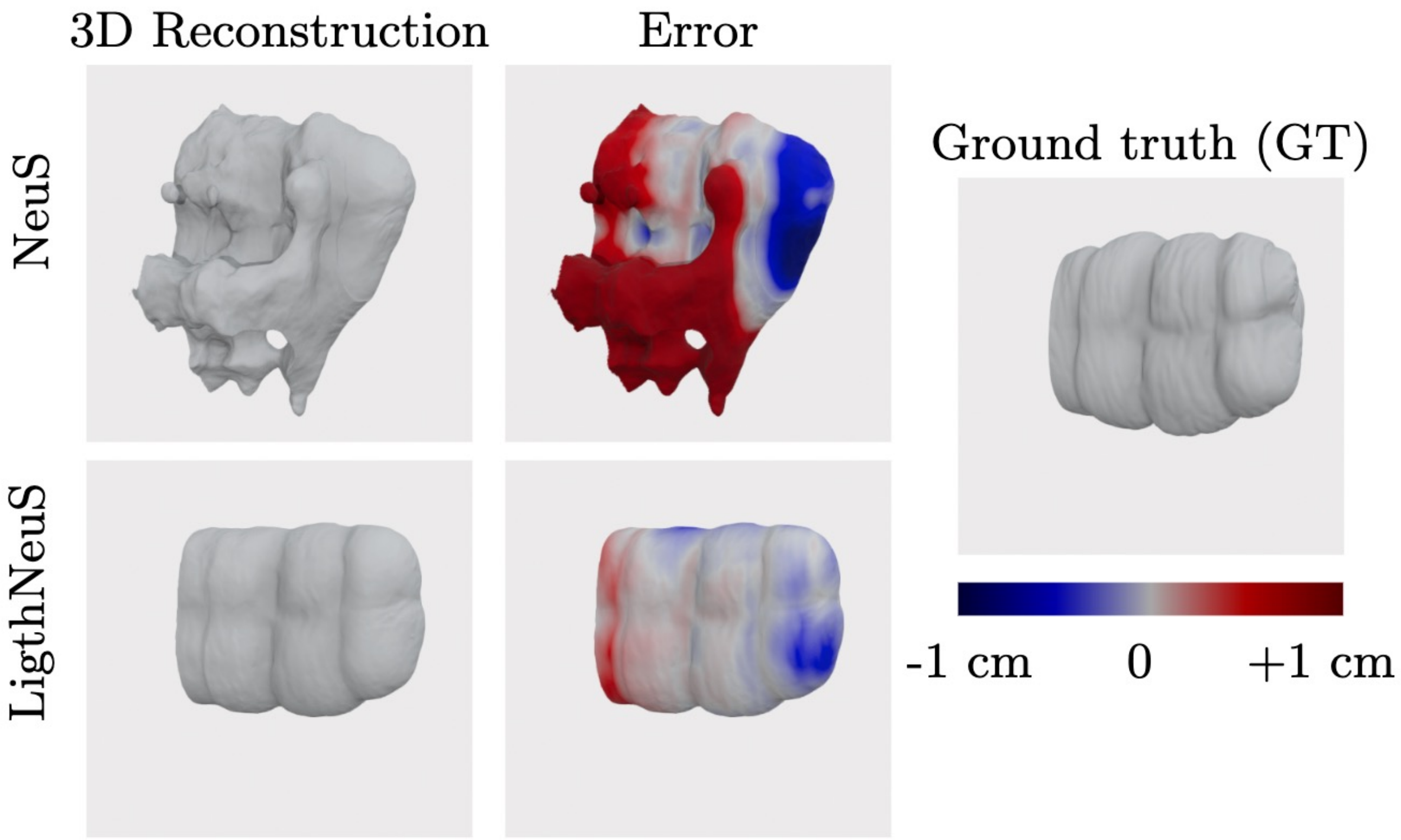


Image of a human colon from a CT scan.
Source: Nevit Dilmen @ Wikipedia

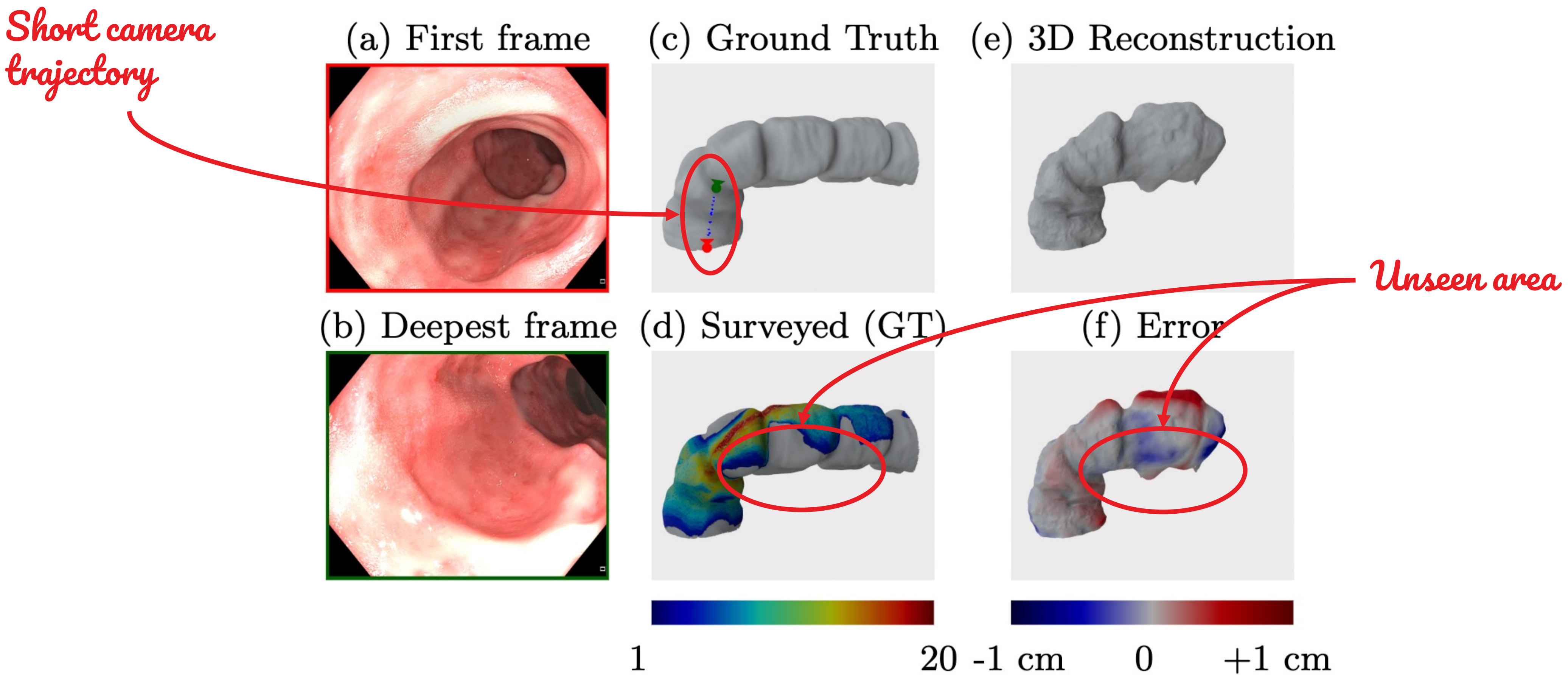
LightNeuS


r
g
b
x
y
z
Θ
φ
t
SDF
COLOR
Volume rendering
Loss
Endoscope photometric model
General photometric model

NeuS vs LightNeuS



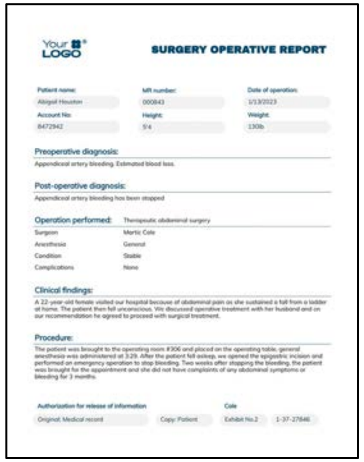
Partially observed regions




After inspecting a colon section,

Future work

Make it work in real-time!



SURGERY OPERATIVE REPORT

Your LOGO

Patient name:	MR number:	Date of operation:
Abigail Houston	000843	1/13/2023
Account No:	Height:	Weight:
8472942	5'4	130lb

Preoperative diagnosis:

Appendiceal artery bleeding. Estimated blood loss.

Post-operative diagnosis:

Appendiceal artery bleeding has been stopped

Operation performed: Therapeutic abdominal surgery

Surgeon	Martic Cole
Anesthesia	General
Condition	Stable
Complications	None

Clinical findings:

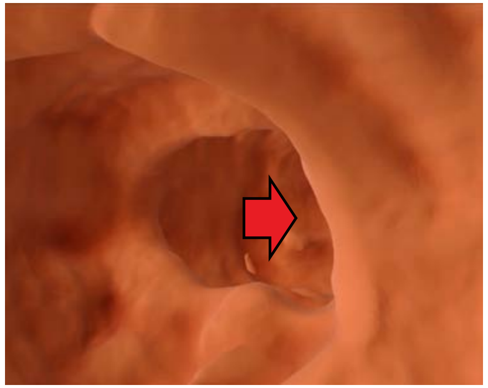
A 22-year-old female visited our hospital because of abdominal pain as she sustained a fall from a ladder at home. The patient then fell unconscious. We discussed operative treatment with her husband and on our recommendation he agreed to proceed with surgical treatment.

Procedure:

The patient was brought to the operating room #306 and placed on the operating table, general anesthesia was administered at 3:29. After the patient fell asleep, we opened the epigastric incision and performed an emergency operation to stop bleeding. Two weeks after stopping the bleeding, the patient was brought for the appointment and she did not have complaints of any abdominal symptoms or bleeding for 3 months.

Authorization for release of information		Code	
Original: Medical record	Copy: Patient	Exhibit No.2	1-37-27846

Automatic report writing

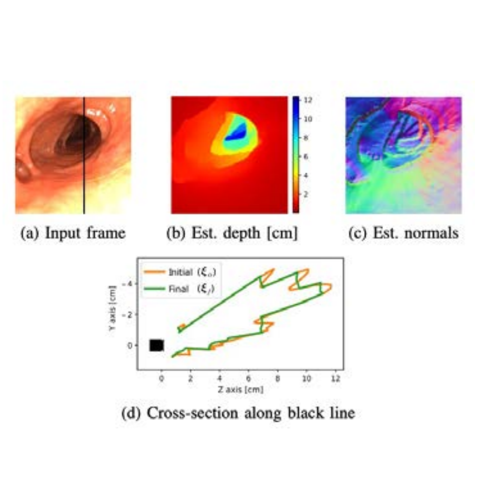


Real-time directions to unexplored areas

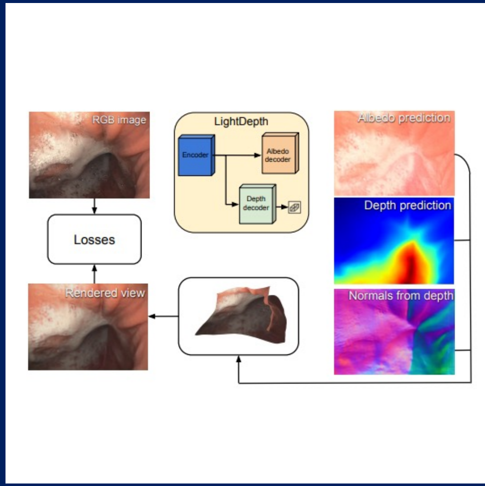
Full deformable model

Exploiting Illumination Decline

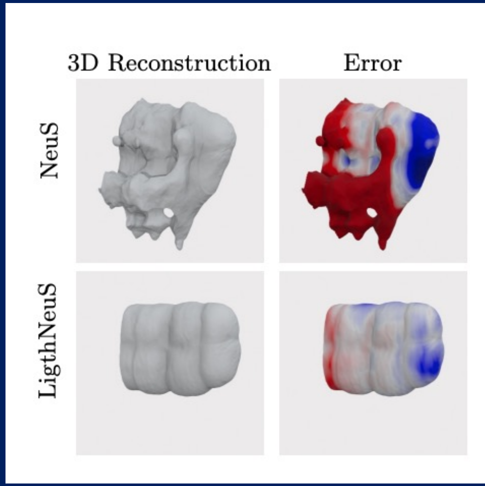
Advances in Monocular Visual SLAM, Self-Supervised Depth Estimation, and Watertight 3D Reconstruction



Real-scale 3D reconstruction
IROS, 2022



LigthDepth
ICCV, 2023



LightNeuS
MICCAI, 2023

Universidad Zaragoza 1542

EndoMapper

Instituto Universitario de Investigación en Ingeniería de Aragón Universidad Zaragoza