Focal Split: Untethered Snapshot Depth from Differential Defocus

Supplementary Material

1. Depth from Differential Defocus

This section provides additional clarifications that complement the derivation of the proposed depth estimation algorithm (Eq. 11 in the main paper). For the convenience of reading, we list all key quantities and symbols already introduced in the main paper in Table 1.

Consider the image captured with sensor distance s of a target with object distance Z, $I(\boldsymbol{x}; s)$. As discussed in the main paper, the variation of the sensor distance s will change both the image defocus and the magnification. Thus, we rescale the image $I(\boldsymbol{x}; s)$ via:

$$\tilde{I}(\boldsymbol{x}; s) = I\left(\frac{s}{c}\boldsymbol{x}\right), \tag{1}$$

and the aligned image $\tilde{I}(\boldsymbol{x}; s)$ has a constant magnification independent of the sensor distance s.

As the unaligned images $I(\boldsymbol{x}; s)$ is the convolution of the point spread function (PSF) $k(\boldsymbol{x}; s)$ and the pinhole image $P(\boldsymbol{x}; s)$, as shown in Table 1, the aligned images $\tilde{I}(\boldsymbol{x}; s)$ satisfies:

$$\begin{aligned}\tilde{I}(\boldsymbol{x}; s) &= \iint k\left(\frac{s}{c}(\boldsymbol{x}-\boldsymbol{u}); s\right) P\left(\frac{s}{c}\boldsymbol{u}; s\right) \frac{s^2}{c^2}|d\boldsymbol{u}| \\ &= \iint \frac{s^2}{c^2} k\left(\frac{s}{c}(\boldsymbol{x}-\boldsymbol{u}); s\right) T\left(-\frac{Z}{c}\boldsymbol{u}\right)|d\boldsymbol{u}|.\end{aligned} \tag{2}$$

Thus, if we define the scaled PSF to be:

$$\tilde{k}(\boldsymbol{x}; s) = \frac{s^2}{c^2} k(\frac{s}{c}\boldsymbol{x}; s), \tag{3}$$

and the consensus pinhole image as:

$$\tilde{P}(\boldsymbol{x}; c) = P\left(\frac{s}{c}\boldsymbol{x}; s\right), \tag{4}$$

the aligned image $\tilde{I}(\boldsymbol{x}; s)$ is the convolution of the two:

$$\tilde{I}(\boldsymbol{x}) = \tilde{k}(\boldsymbol{x}; s) * \tilde{P}(\boldsymbol{x}; c). \tag{5}$$

The relationship between the partial derivatives of the scaled PSFs $\tilde{k}(\boldsymbol{x}; s)$, shown in Eq. 9 in the main paper, can be derived as follows. The partial derivative of $\tilde{k}(\boldsymbol{x}; s)$ with respect to the sensor distance s is:

$$\tilde{k}_s(\boldsymbol{x}; s) = (2 + 2\psi)\frac{sA}{c^2\sigma^3}e^{\psi}, \tag{6}$$

where

$$\psi = -\frac{s^2\|\boldsymbol{x}\|^2}{2c^2\sigma^2}. \tag{7}$$

Symbol	Expression	Description
$T(\boldsymbol{x})$	-	Object texture
$P(\boldsymbol{x}; s)$	$T\left(-\frac{Z}{s}x, -\frac{Z}{s}y\right)$	Pinhole image
$k(\boldsymbol{x}; s)$	$\frac{1}{\sigma^2(s)}\exp\left(-\frac{x^2+y^2}{2\sigma(s)^2}\right)$	Point spread function (PSF)
$I(\boldsymbol{x}; s)$	$k(\boldsymbol{x}; s) * P(\boldsymbol{x}; s)$	Defocused image
$\sigma(s)$	$A(\frac{1}{Z} - \rho)s + A$	Defocus level
Z	-	Object depth
s	-	Sensor distance
A	-	Standard deviation of the Gaussian aperture code
ρ	-	Optical power of the lens

Table 1. List of key symbols and quantities for deriving the proposed depth estimation algorithm.

And the spatial Laplacian of $\tilde{k}(\boldsymbol{x}; s)$ is:

$$\nabla^2\tilde{k}(\boldsymbol{x}; s) = -(2 + 2\psi)\frac{s^4}{c^4\sigma^4}e^{\psi}. \tag{8}$$

By combining Eq. 7 and 8, we obtain the relationship:

$$\tilde{k}_s(\boldsymbol{x}; s) = -\frac{c^2\sigma A}{s^3}\nabla^2\tilde{k}(\boldsymbol{x}; s), \tag{9}$$

2. *Do It Yourself* Guide

2.1. Components

The Focal Split prototype was assembled mainly using off-the-shelf optical and mechanical components. We provide the complete list of parts in Table 2. We built the custom housing via 3D-printing and the CAD model is included in the project page listed in the abstract.

2.2. Assembling Instruction

2.2.1 Opto-mechanical System

First, disassemble the mounted lenses from the photosensors (1), and connect the camera cable (8) to each photosensor. Use M2 screws and washers to mount the two photosensor circuits onto the housing (4). Make sure to add an extra washer when mounting one of the sensors so that its sensor distance is different from the other one. Then, mount the beamsplitter (3) onto the housing using M4 screws and washers. Be mindful that its orientation is the same as

No.	Component	Vendor	Part Number	Quantity	Price	Description
1	Photosensor	PiShop	799632837664	2	$31	Low-power RGB sensor, with the lens removed
2	Lens	Thorlabs	AC127-030-A-ML	1	$93	Ø1/2" Mounted Achromatic Doublets lens with AR coated
3	Beamsplitter	Thorlabs	CCM5-BS016/M	1	$210	Cube-Mounted, Non-Polarizing, 50:50 Beamsplitter Cube
4	Housing			1	-	Housing to hold sensors, lens, and beamsplitter, produced by 3D printing
5	Iris	Thorlabs	SM1D12C	1	$120	(Optional) Manual iris to control the brightness of captured images
6	Raspberry Pi	PiShop	4GB-9024	1	$60	Computing unit
7	MicroSD Card	PiShop	1337-1	1	$13	SD card for the computer
8	Camera Cable	PiShop	1510-1	2	$2	Cable for sensors
9	Power Supply	PiShop	1795-1	1	$14	Power Supply for the computer
10	LCD	PiShop	1824	1	$23	(Optional) Display used for showing depth maps and photos
11	5 VDC Battery Pack	Thorlabs	CPS1	1	$41	Battery

Table 2. List of parts.

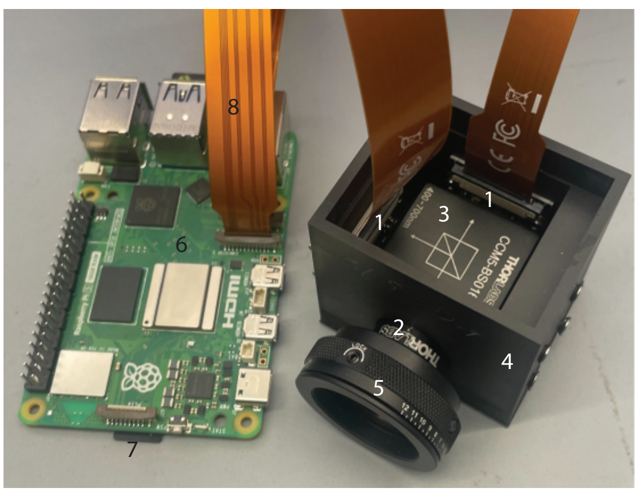

Figure 1. Focal Split prototype.

shown in Fig. 1a. Finally, attach the lens (2) and, optionally, the iris (5) to the beamsplitter.

2.2.2 Computing system

Connect the other ends of the camera cable (8) to the Raspberry Pi (6). Insert the MicroSD card (7) and connect the power supply (9). If necessary, the LCD monitor screen (10) can also be mounted to the Raspberry Pi for onboard display.

3. Additional Experimental Results

3.1. Tolerance to object displacement

Fig. 2 quantifies the impact of different object motions on the depth estimation accuracy. We perform this experiment by sequentially capturing I_1 and I_2 while exerting a known displacement to the object between the capture. We observe the depth estimation algorithm is extremely sensitive to lateral translation and rotations of different axes while being relatively robust to axial translation. Although the study is only conducted using the proposed depth estimation algorithm, it can be translated to other algorithms in the DfDD family due to similarities in solutions. This analysis strongly argues the importance of a snapshot measurement system for DfDD.

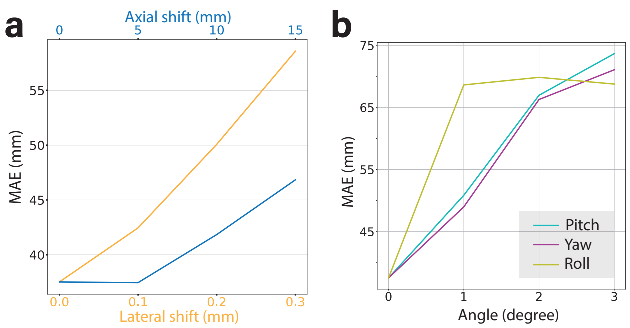

Figure 2. Depth prediction accuracy as a function of relative displacements between I_1 and $I - 2$ if the two images are captured sequentially. (a) The mean absolute error (MAE) of depth prediction is when the target translates axially or laterally in different amounts. (b) The MAE of depth prediction when the target rotates around different axes. Based on the plots, the MAE increases significantly when a slight displacement between I_1 and I_2 appears. This result suggests the sensitivity of the proposed method to the relative motion between I_1 and I_2, as a 0.3mm lateral shift and a 3° rotation doubles the MAE, highlighting the importance of the snapshot capture of I_1 and I_2. The analysis is based on the real captured data.