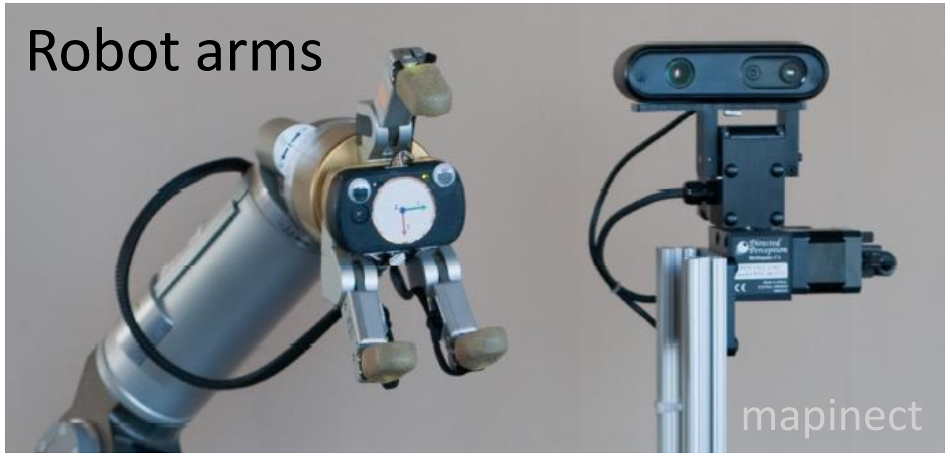
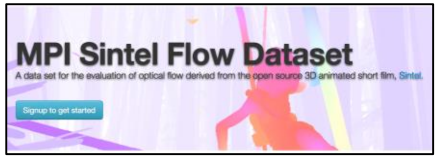
Focal Track:
Depth and Accommodation with Oscillating Lens Deformation

Qi Guo, Emma Alexander, Todd Zickler

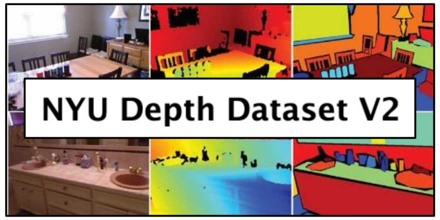
Applications of Depth Sensors



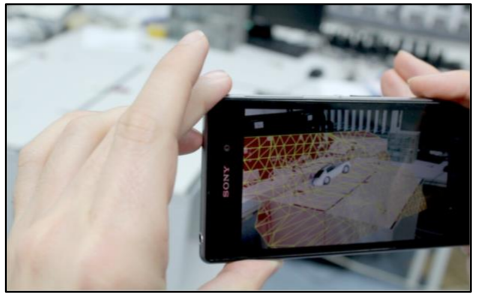
When power and space gets expensive



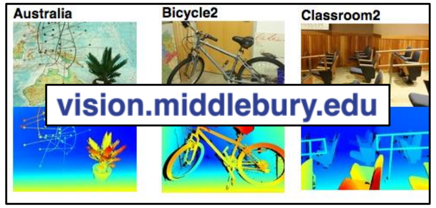
[Butler et al. 12]



[Silverman et al. 12]

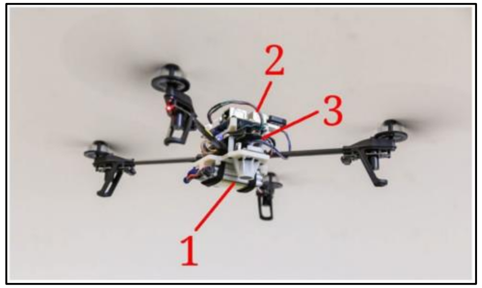


[Schöps et al. 14]

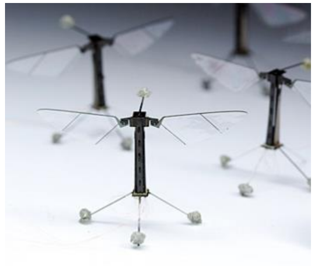


[Scharstein et al. 14]

[Menze & Geiger 15]

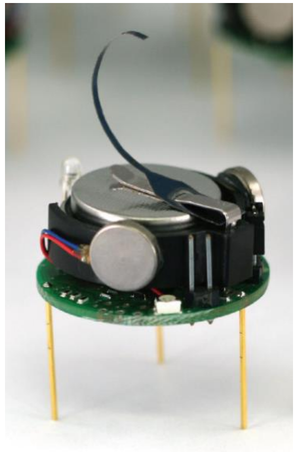


[Faessler et al. 15]

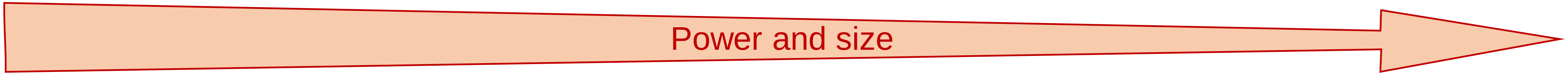


[Ma et al. 13]

[Chen et al. 10]



[Rubenstein et al. 14]



~W

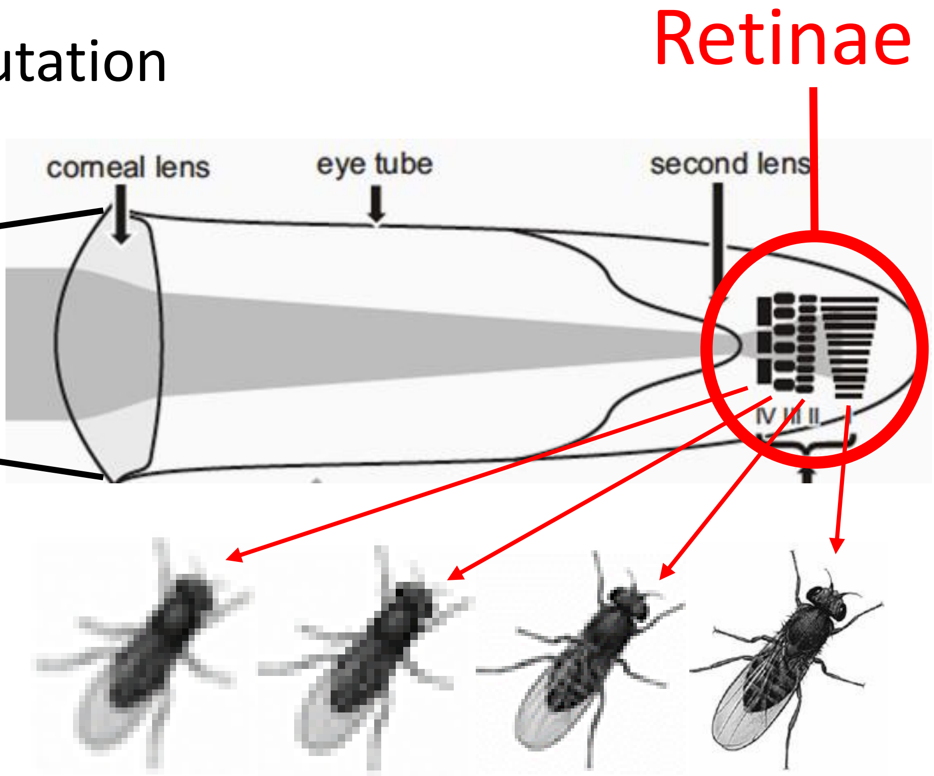
~mW

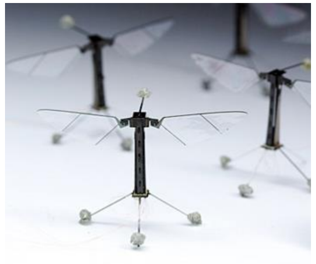
[Slide: Alexander et al 2017]

When power and space gets expensive

Computational sensors:

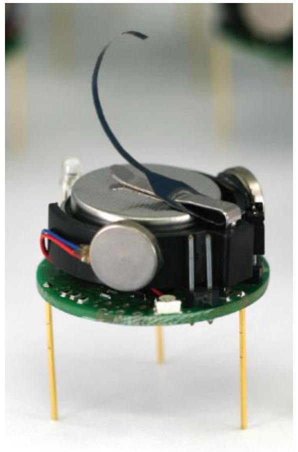
- Co-design optics and computation





[Ma et al. 13]

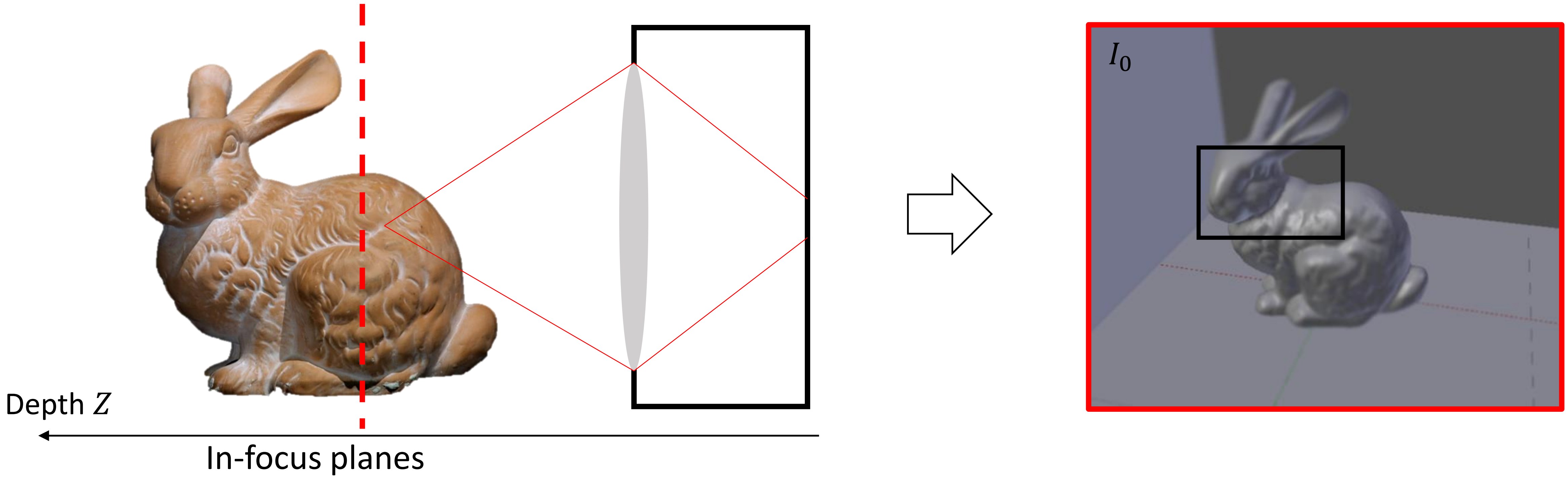
[Chen et al. 10]



[Rubenstein et al. 14]

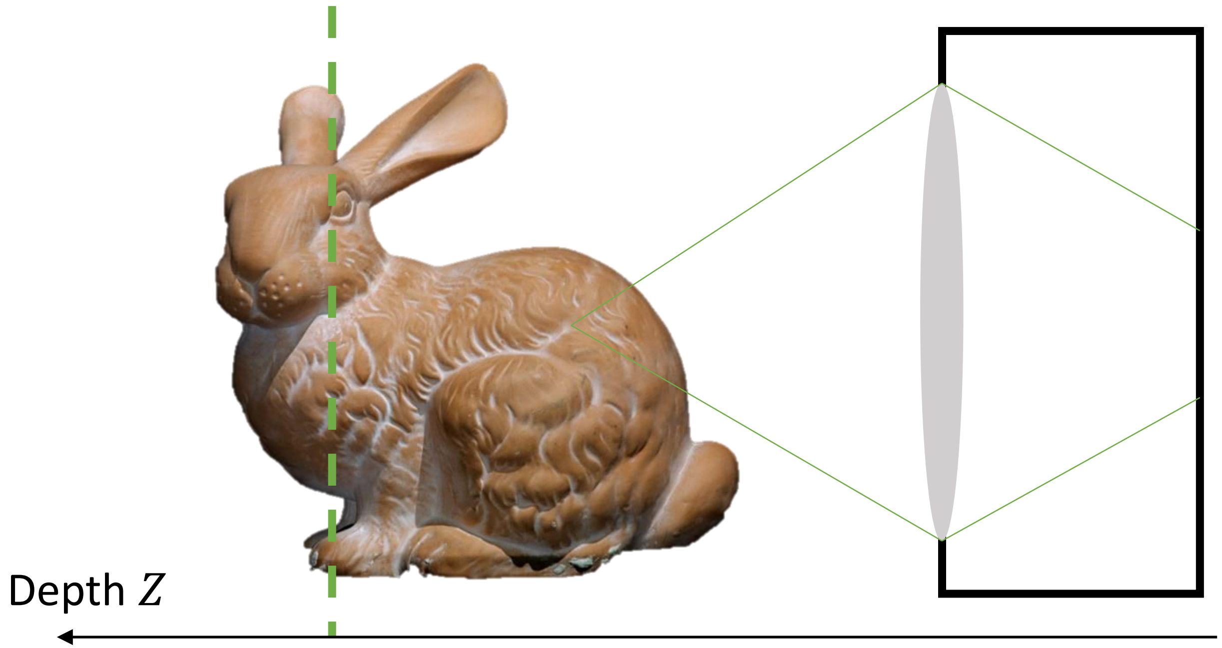
[Slide: Alexander et al 2017]

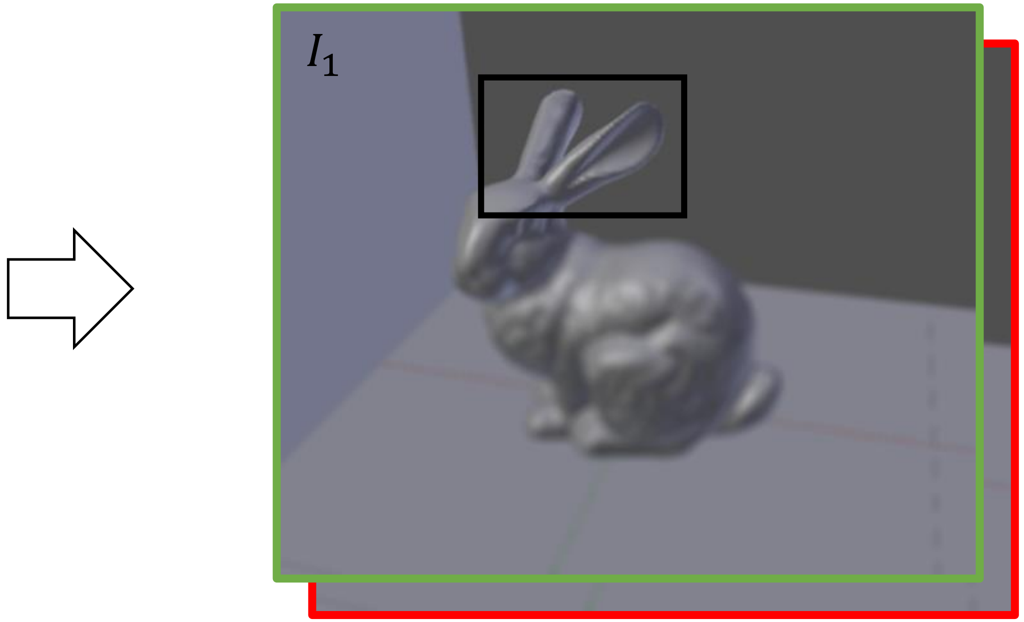
Review: Two-shot Depth from Defocus



E.g., Watanabe and Nayar IJCV 1998, Huixuan et al. CVPR 2017

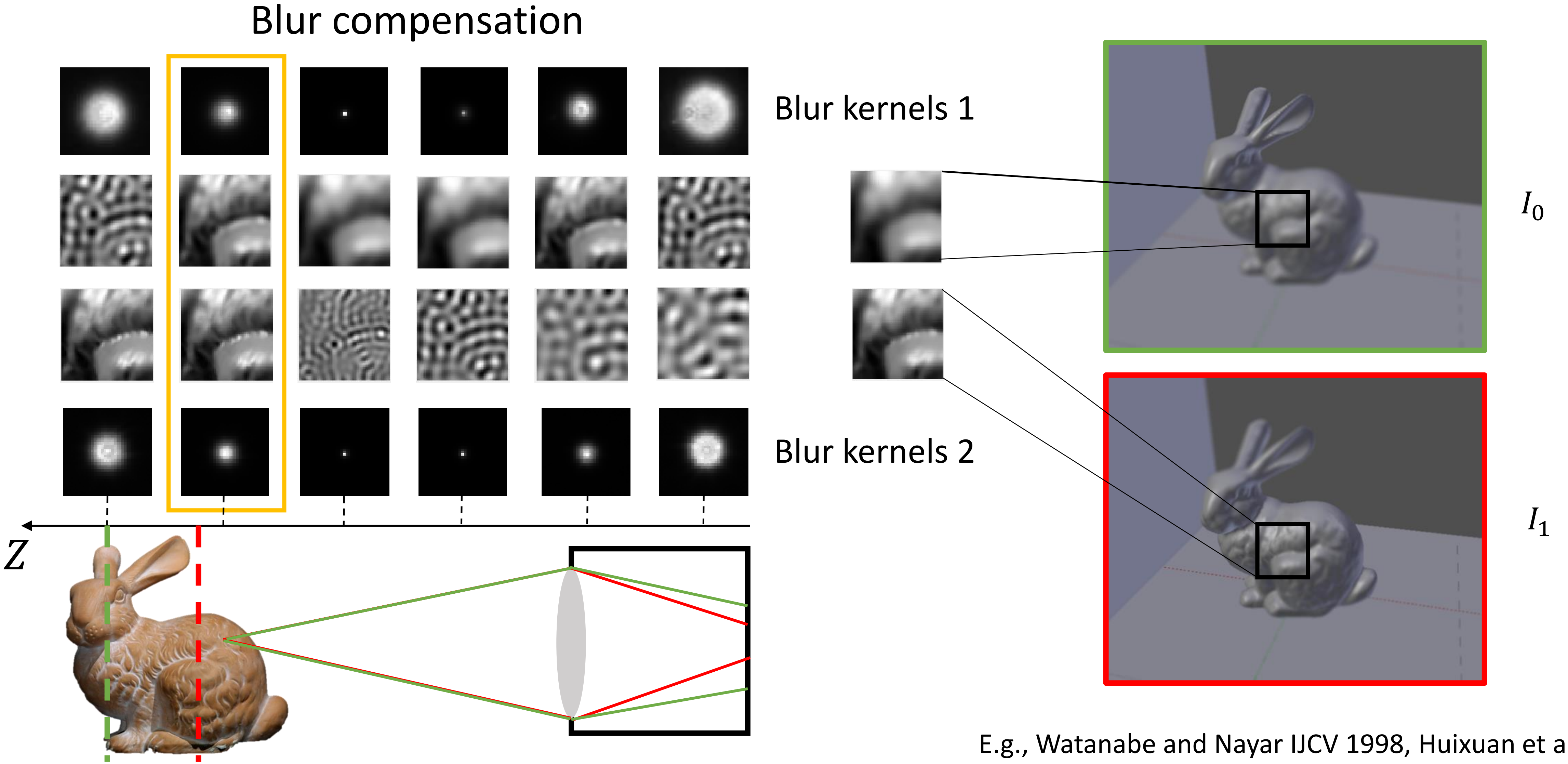
Review: Two-shot Depth from Defocus





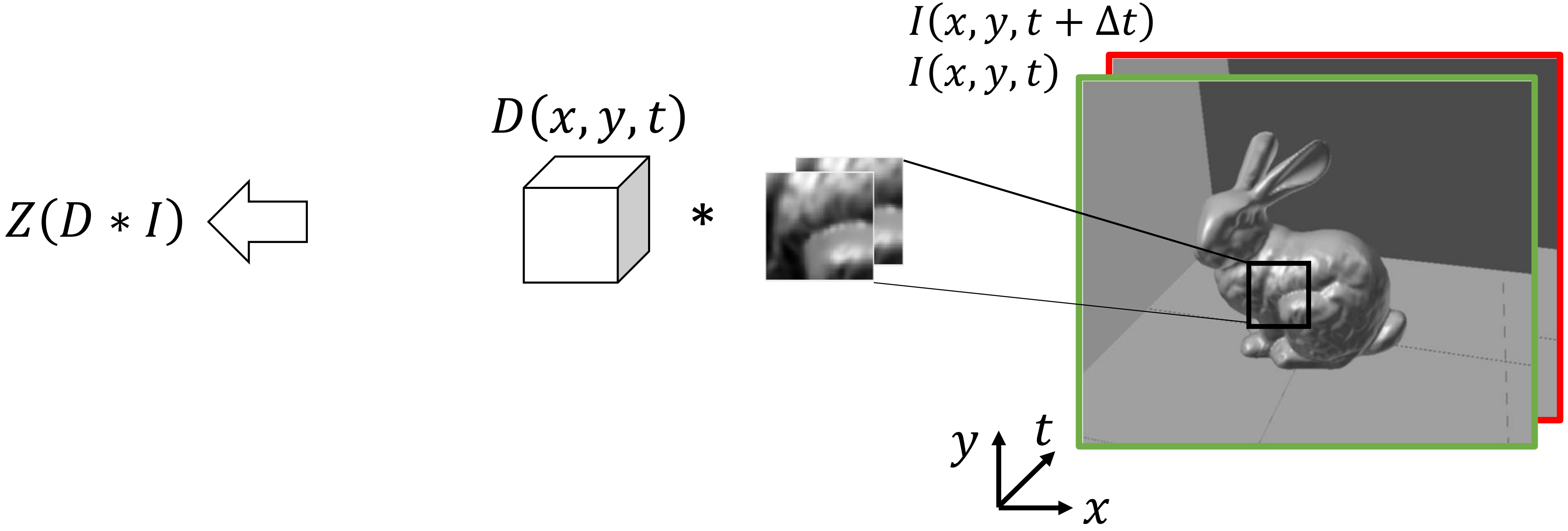
E.g., Watanabe and Nayar IJCV 1998, Huixuan et al. CVPR 2017

Review: Two-shot Depth from Defocus

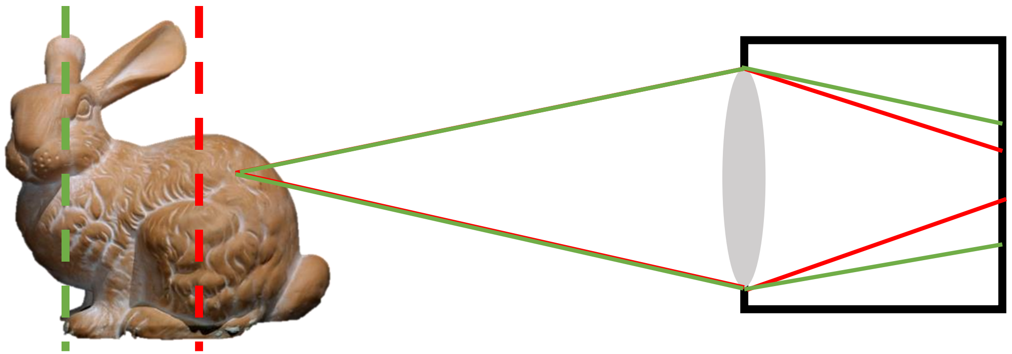


E.g., Watanabe and Nayar IJCV 1998, Huixuan et al. CVPR 2017

Our alternative



Differential Focus Change

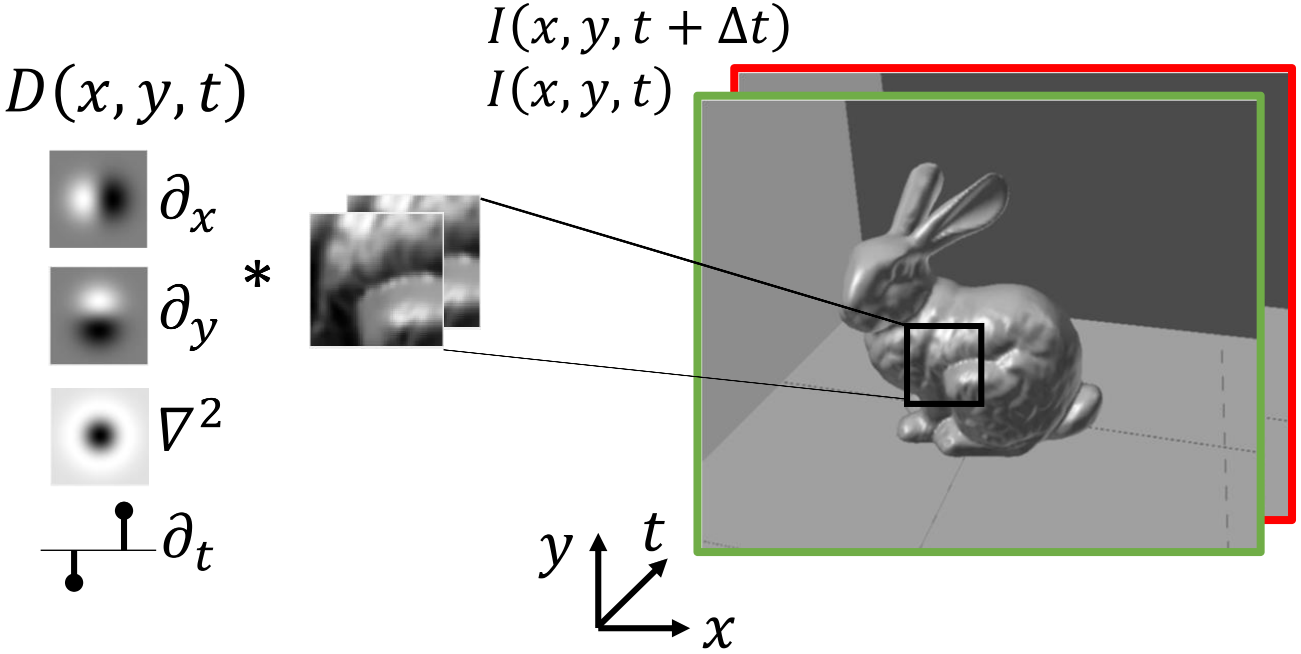


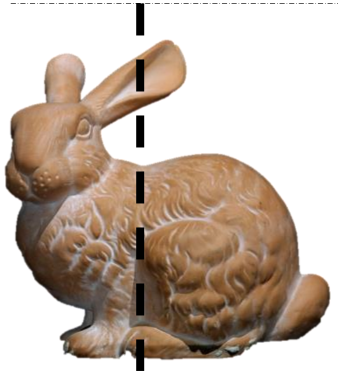
Differential Depth from Defocus

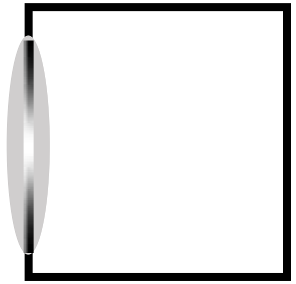
Last year: Focal Flow [Alexander et al., ECCV 2016, IJCV 2017]

$$\begin{bmatrix} & & \vdots & \\ I_x & I_y & xI_x + yI_y & \nabla^2 I \\ & & \vdots & \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ v \end{bmatrix} = \begin{bmatrix} \vdots \\ I_t \\ \vdots \end{bmatrix}$$

$$Z = \frac{\alpha_1 u_3}{\alpha_2 u_3 - \alpha_3 v}, (\dot{X}, \dot{Y}, \dot{Z}) = \left(\frac{Z u_1}{\alpha_4}, \frac{Z u_2}{\alpha_4}, Z u_3\right).$$







Thin lens
Gaussian PSF
Camera Motion

- Per patch depth and velocity

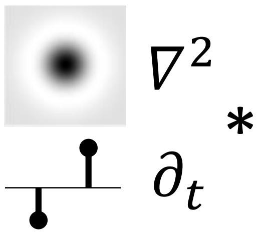
Differential

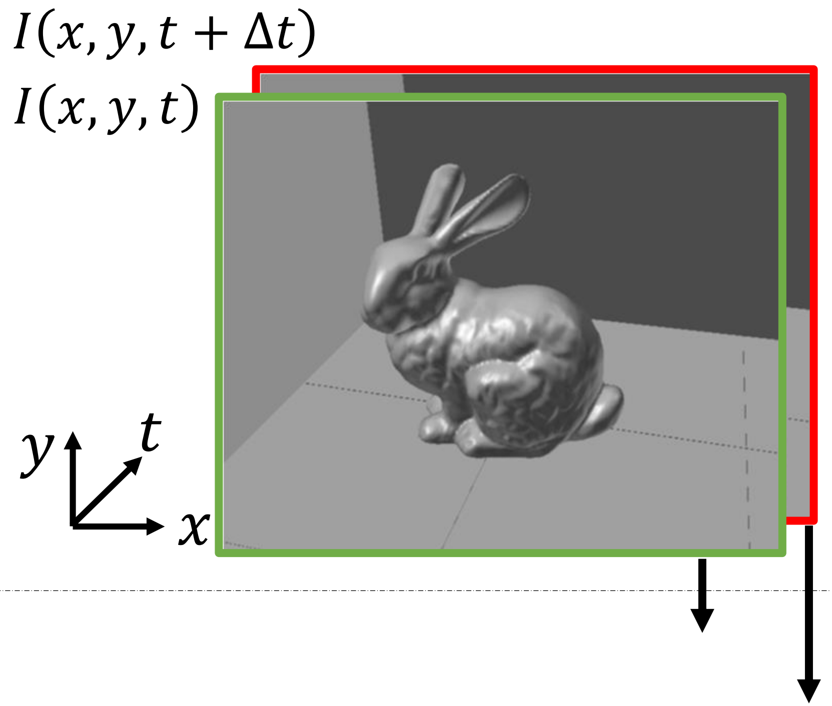
Depth from Defocus

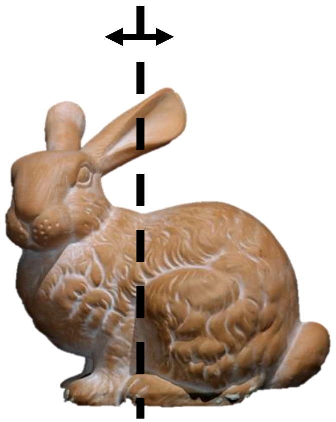
This year: Focal Track [Guo et al., ICCV 2017]

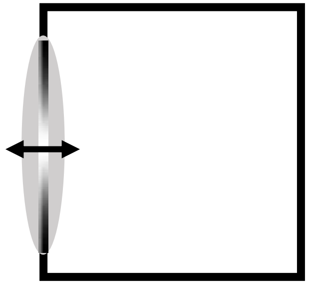
$$Z = \frac{\beta_1 \nabla^2 I}{\beta_2 \nabla^2 I - I_t}.$$

$$Z = \frac{\partial_x^j \partial_y^k (\beta_1 \nabla^2 I)}{\partial_x^j \partial_y^k (\beta_2 \nabla^2 I - I_t)}.$$





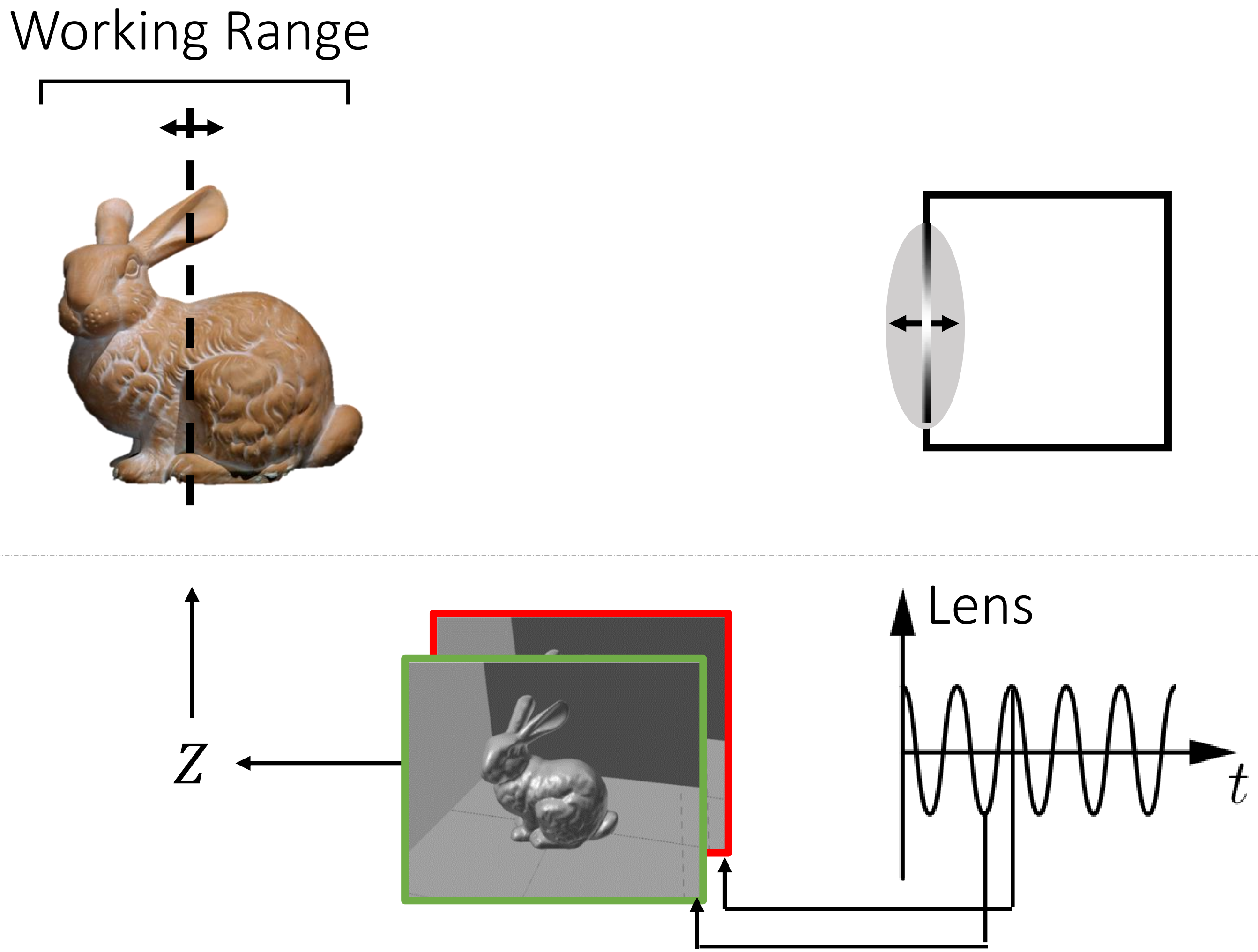




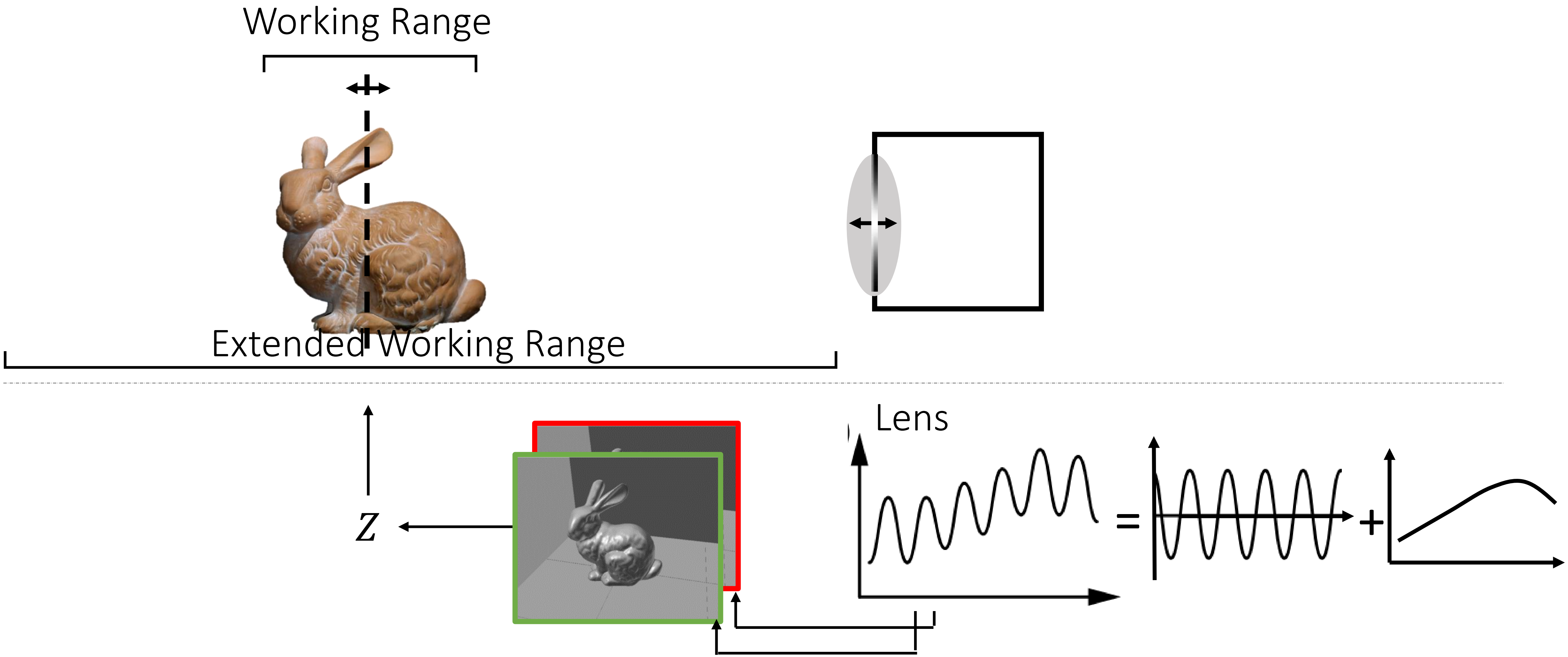
Thin lens
Gaussian PSF
Lens deformation

- Per pixel depth

Advantage: "Accommodation"



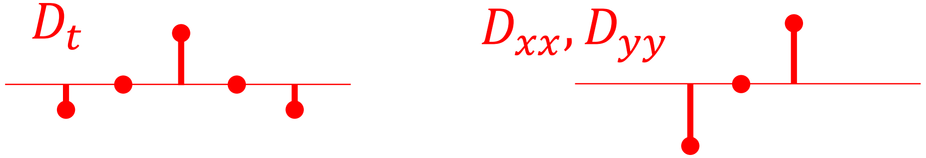
Advantage: "Accommodation"



Reality

- Finite differential filter:

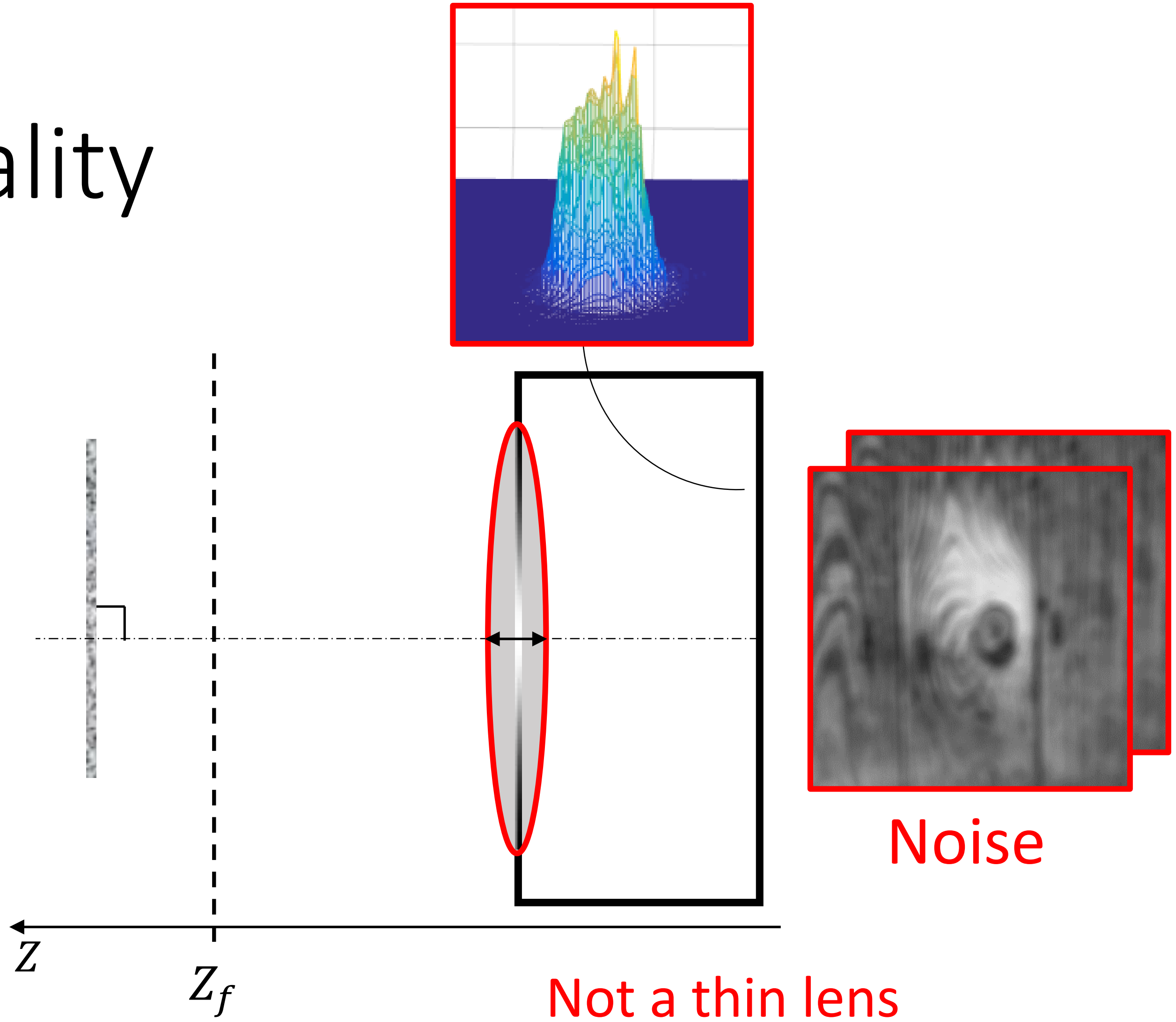
D_t $\qquad$ D_{xx}, D_{yy}



$$\nabla^2 I = D_{xx} * I + D_{yy} * I, I_t = D_t * I.$$

- Obtain depth, $Z = \dfrac{\beta_1 \nabla^2 I}{\beta_2 \nabla^2 I \ - I_t}$.

Hard to calibrate!

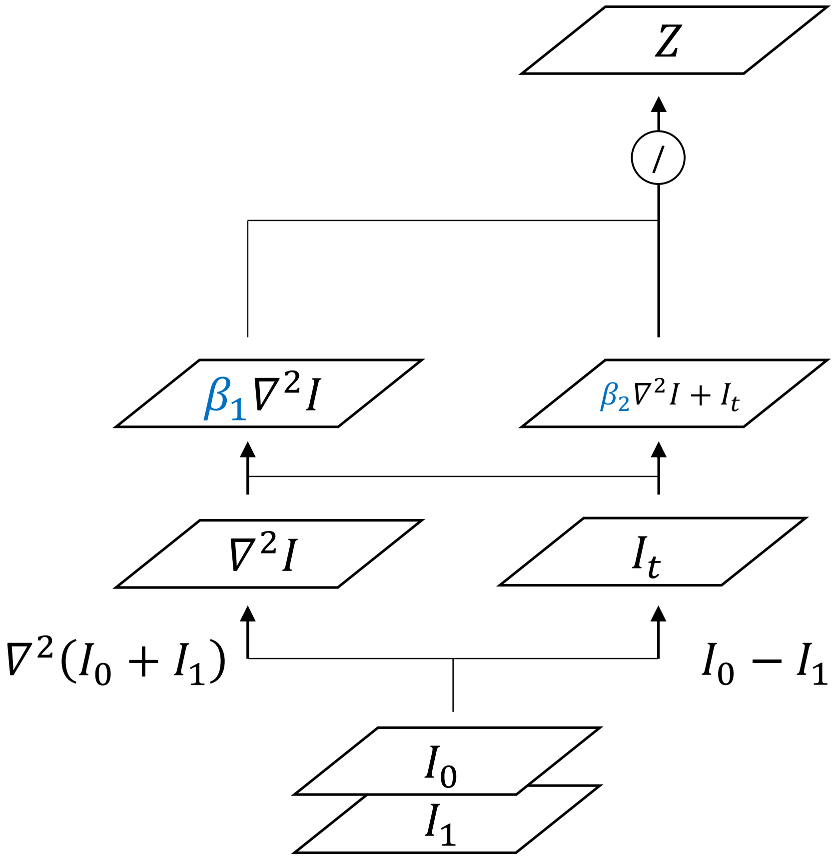


Not a thin lens

- Inter-reflection
- Aberration
- Diffraction

End-to-end calibration

$$Z = \frac{\beta_1 \nabla^2 I}{\beta_2 \nabla^2 I \; - I_t}.$$

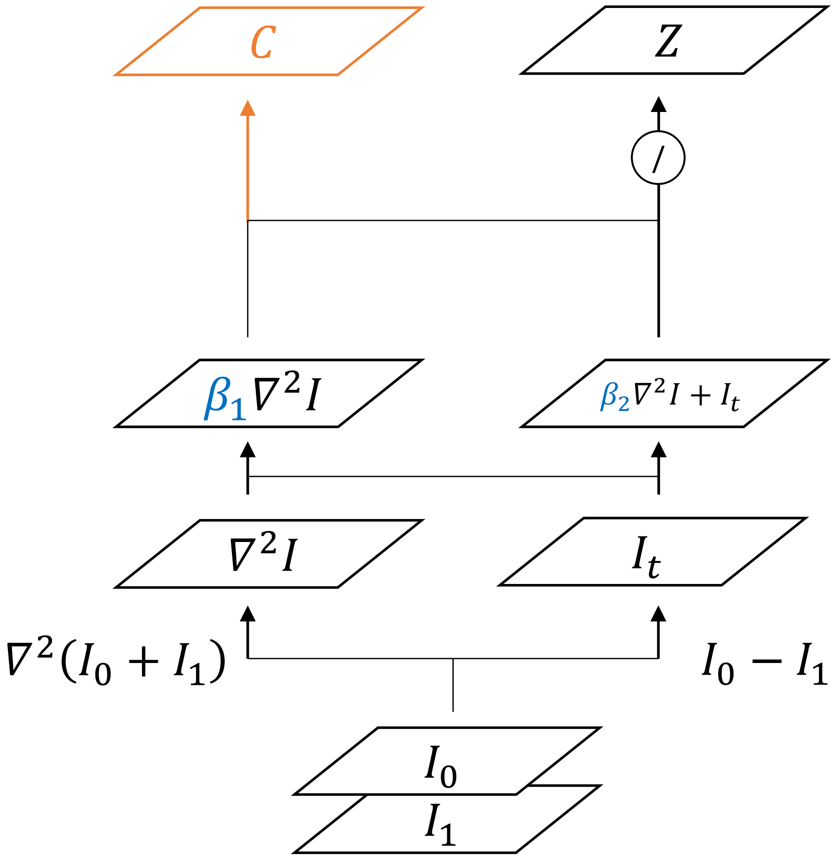


End to end calibration with confidence

Gaussian noise model: $\tilde{I} = I + \epsilon$.

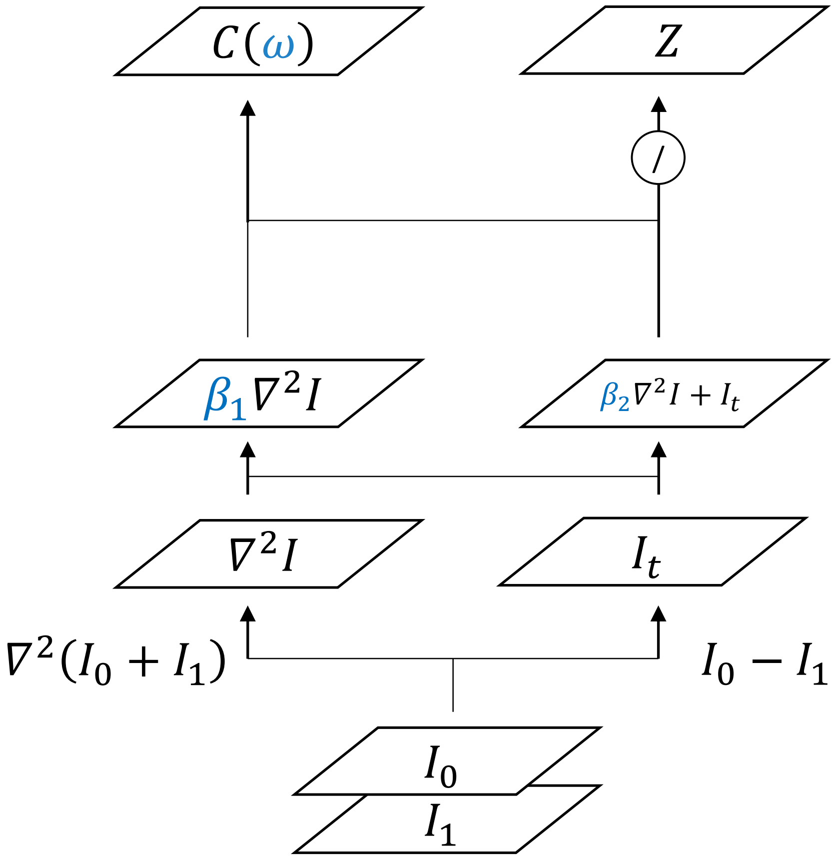
$$\tilde{Z} = \frac{V}{W} = \frac{\beta_1 \nabla^2 \tilde{I}}{\beta_2 \nabla^2 \tilde{I} - I_t}$$

$$C(\omega) = \left(1 + \frac{\omega_0}{W^2} + \frac{V^2 \omega_1}{W^4} + \frac{V \omega_2}{W^3}\right)^{-0.5} \approx \left(1 + Var(\tilde{Z})\right)^{-0.5}$$



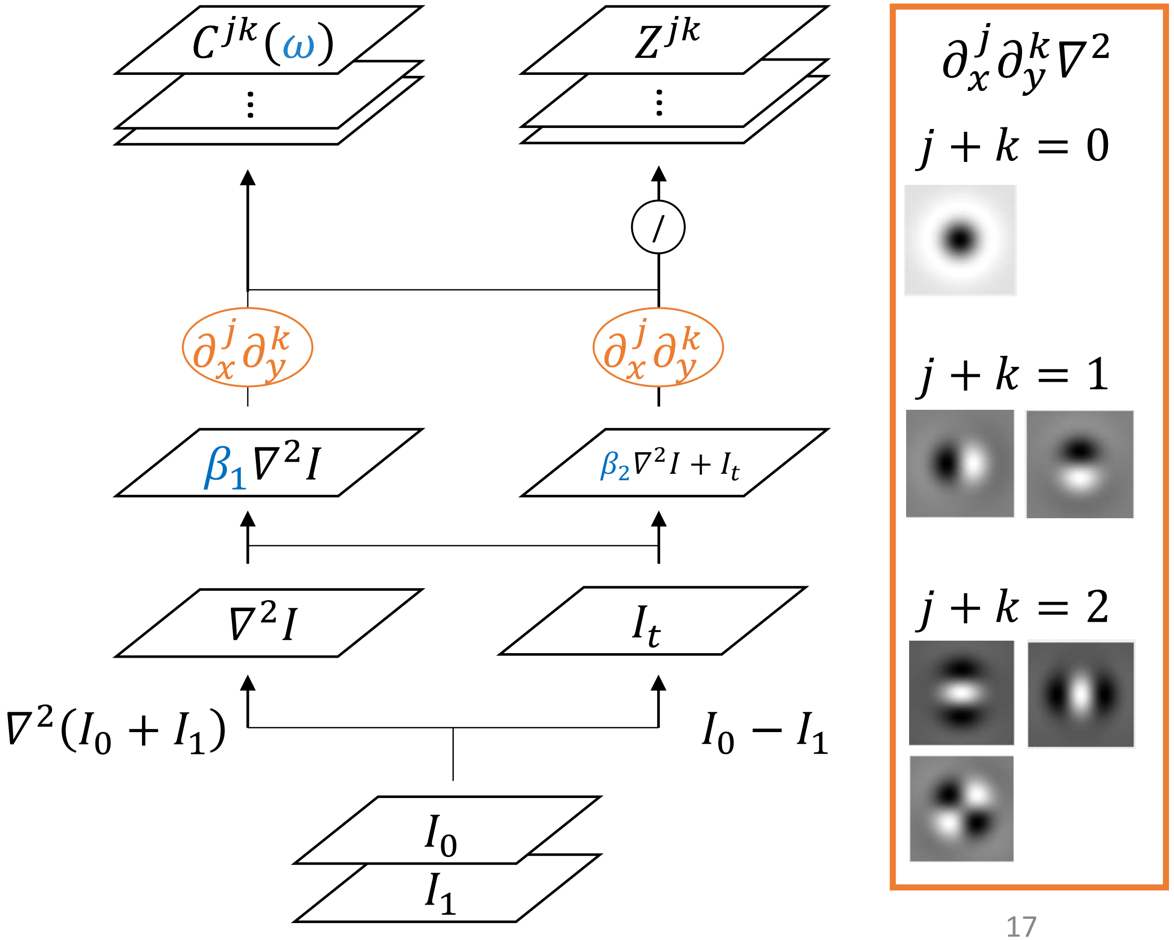
End to end calibration

$$Z = \frac{\partial_x^j \partial_y^k (\beta_1 \nabla^2 I)}{\partial_x^j \partial_y^k (\beta_2 \nabla^2 I \; - I_t)}$$

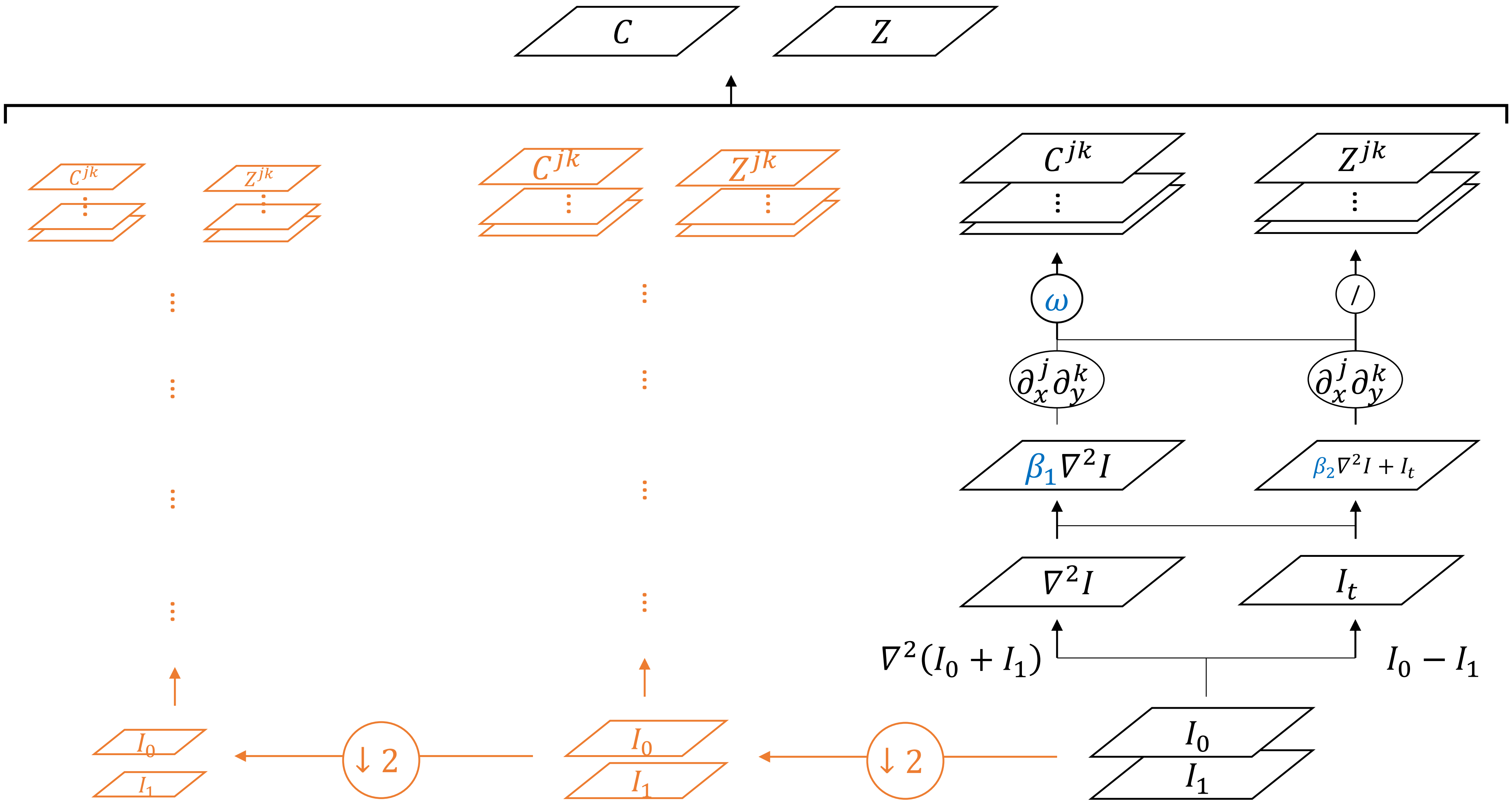


End to end calibration

$$Z = \frac{\partial_x^j \partial_y^k (\beta_1 \nabla^2 I)}{\partial_x^j \partial_y^k (\beta_2 \nabla^2 I \ - I_t)}$$

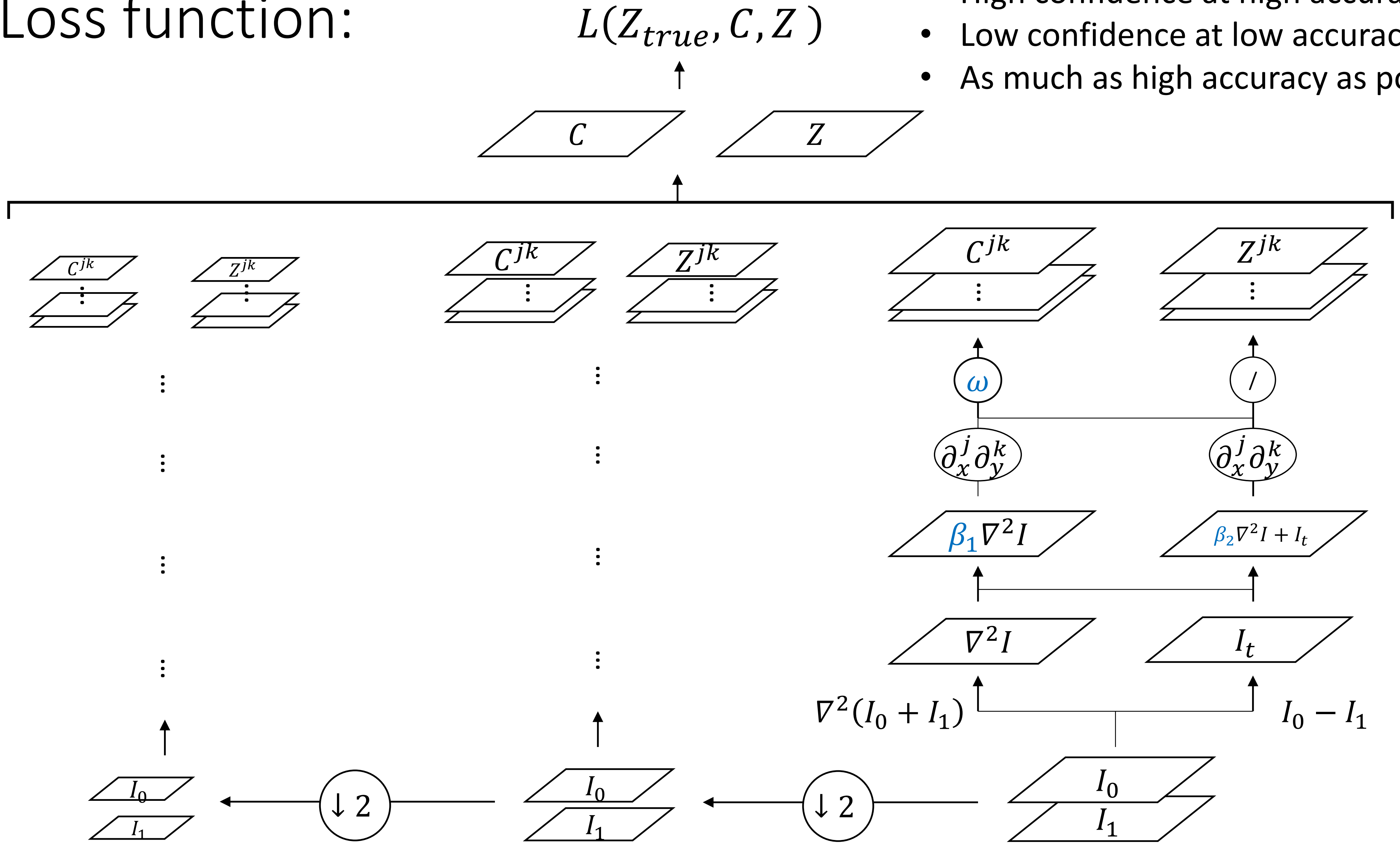


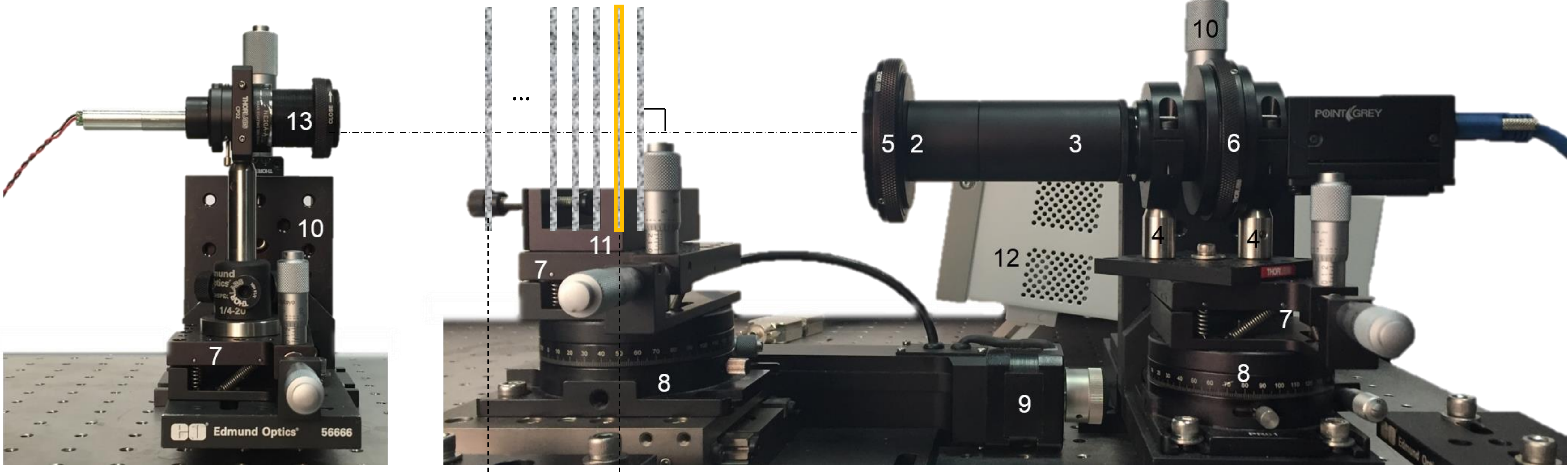
Computational Graph



Loss function:

- High confidence at high accuracy
- Low confidence at low accuracy
- As much as high accuracy as possible





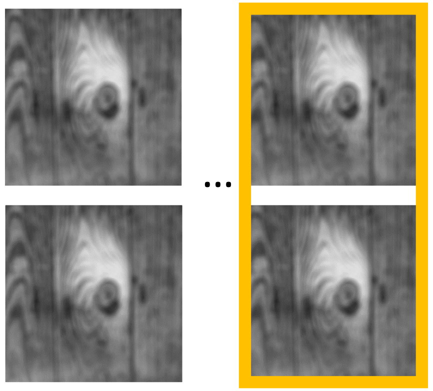
Laser Stage

Object Stage

Camera Stage

Z

$\{(Z_{true}, I_0, I_1)_i\}$

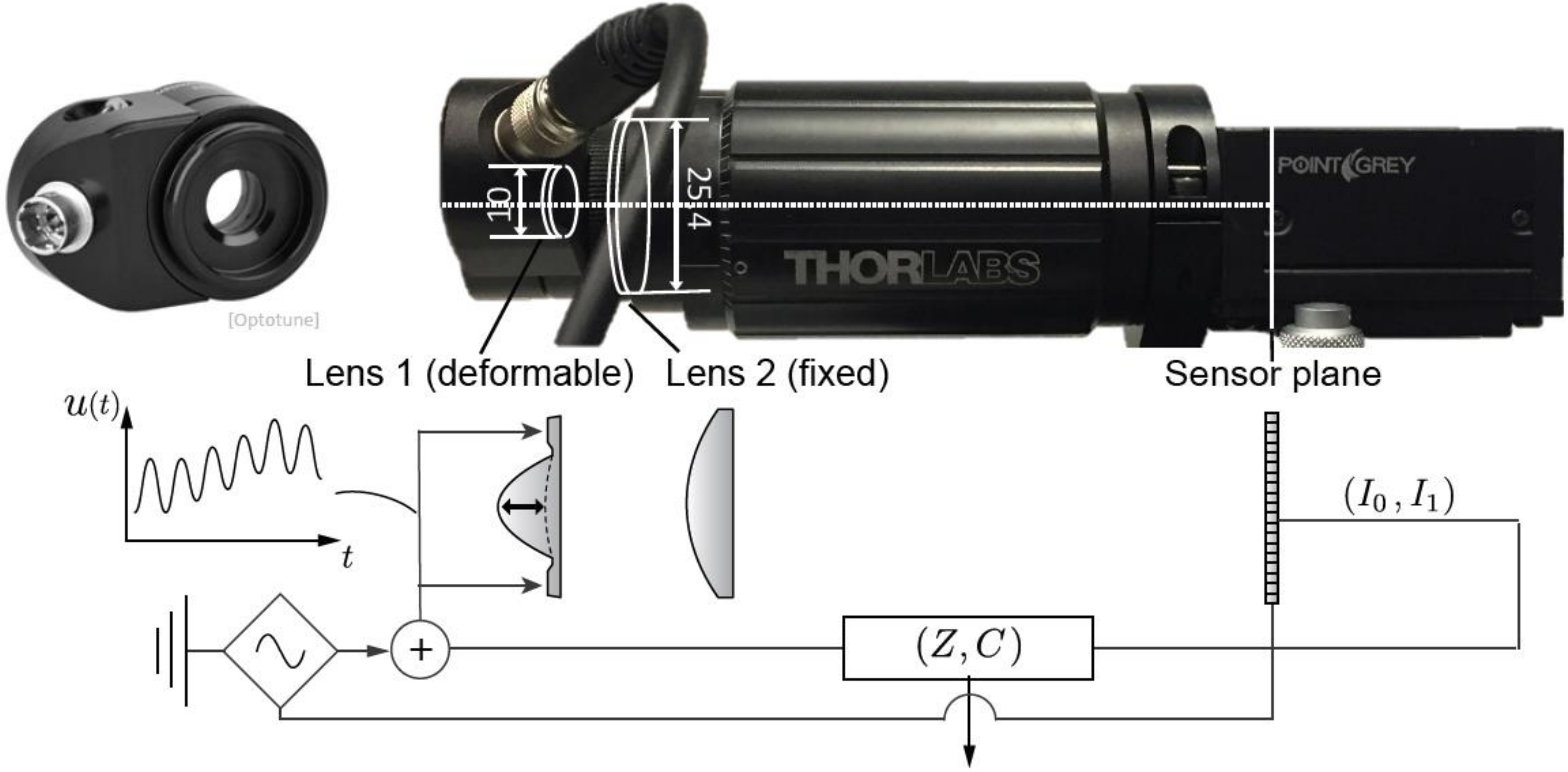


(Z, C)

$L(Z_{true}, Z, C)$

$\frac{\partial L}{\partial \beta}, \frac{\partial L}{\partial \omega}$

Prototype



[Optotune]
10
25.4
THORLABS
POINTGREY
Lens 1 (deformable)
Lens 2 (fixed)
Sensor plane
$u(t)$
t
(I_0, I_1)
(Z, C)

Thanks!

Project page: vision.seas.harvard.edu/focaltrack