Carnegie Mellon University

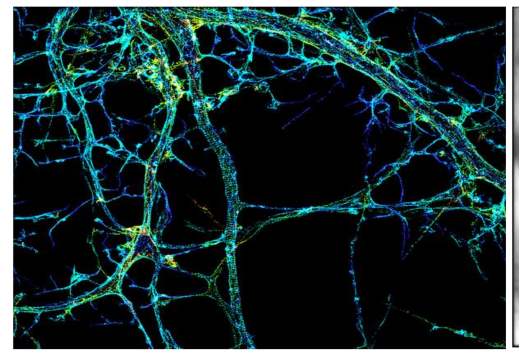
THE ROBOTICS INSTITUTE

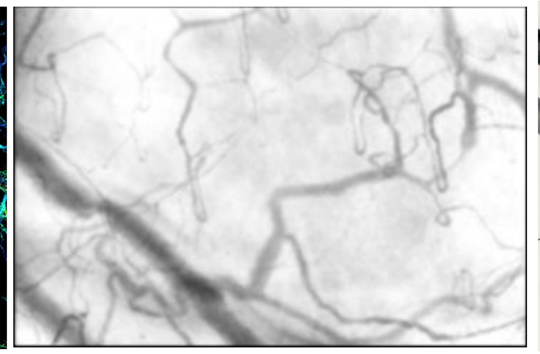
Depth and Matte Recovery for Multilayer Thin Structures

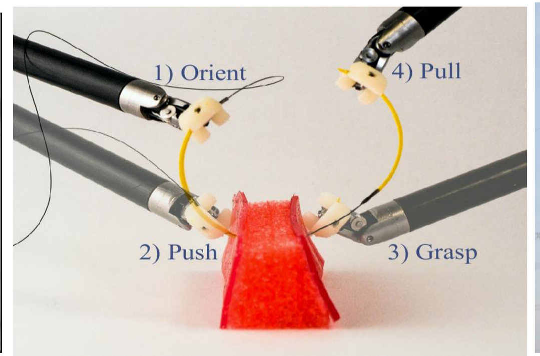
Chao Liu   Srinivasa Narasimhan   Artur Dubrawski
Robotics Institute   Carnegie Mellon University

IEEE 2017 Conference on Computer Vision and Pattern Recognition
CVPR July 21-26 HONOLULU 2017

Motivation: thin structures are common


Neural Science


Vascular Study



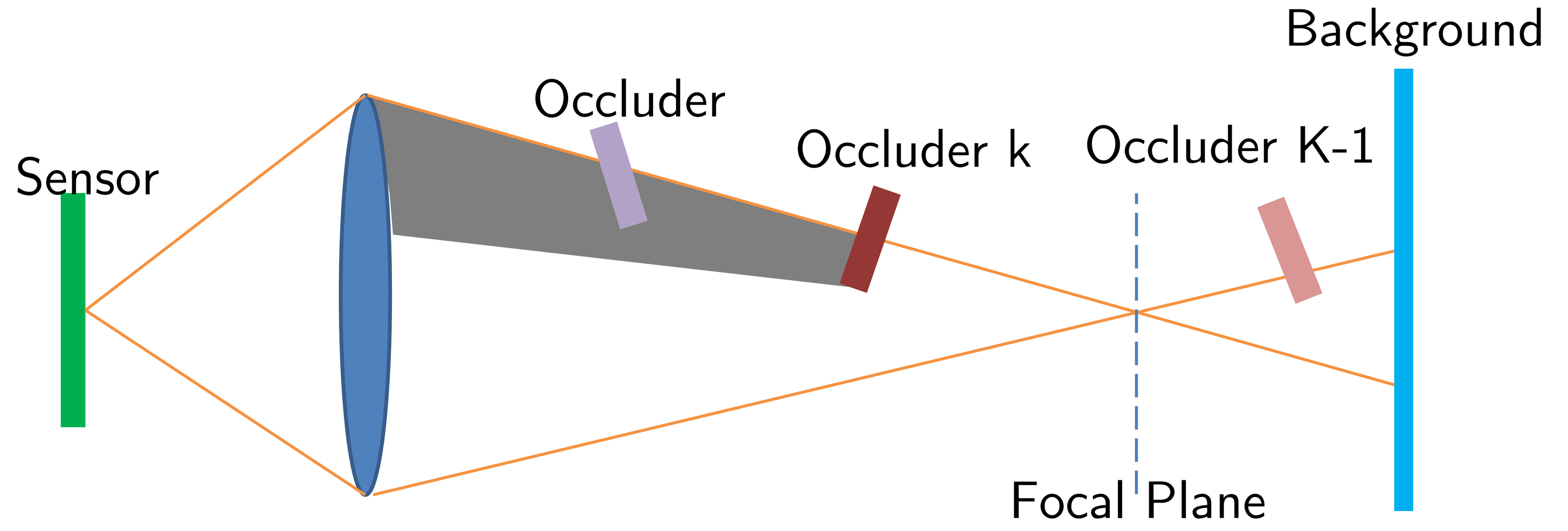
Robotic Assisted Surgery


Drone Navigation

Challenges in estimating the depth for thin structures:

- Textureless: feature matching based method easily fail.
- Fine-grained structures: high spatial frequency in depth discontinuities.
- Mutual occlusions.

Multi-layer image formation model



$$R^m(\boldsymbol{x}) = \sum_k \underbrace{\alpha_k^m(\boldsymbol{x})}_{\text{occlusion}} \sum_{\boldsymbol{u}} \underbrace{M_k(\boldsymbol{u})}_{\text{matting}} \overbrace{L_k(\boldsymbol{u})}^{\text{radiance}} \overbrace{B_k^m(\boldsymbol{u}-\boldsymbol{x}, d)}^{\text{blur}}$$

$$\alpha_k^m(\boldsymbol{x}) = \begin{cases} 1 \\ \prod_{1\le j\le k-1} 1 - \sum_{\boldsymbol{u}} M_j(\boldsymbol{u}) B_j^m(\boldsymbol{u}-\boldsymbol{x}, d) \end{cases}$$

$M(\boldsymbol{x}) \in \{0,1\}$: Occlusion matte, Boolean
$d(\boldsymbol{x}) \ge 0$: Depth, Continuous
$L(\boldsymbol{x}) \ge 0$: Radiance, Continuous

$M_1(\boldsymbol{x}) = 1$
$R^m(\boldsymbol{x}) = R(occ) + \alpha R(bkg)$

$M_1(\boldsymbol{x}) = 0$
$R^m(\boldsymbol{x}) = R(bkg)$

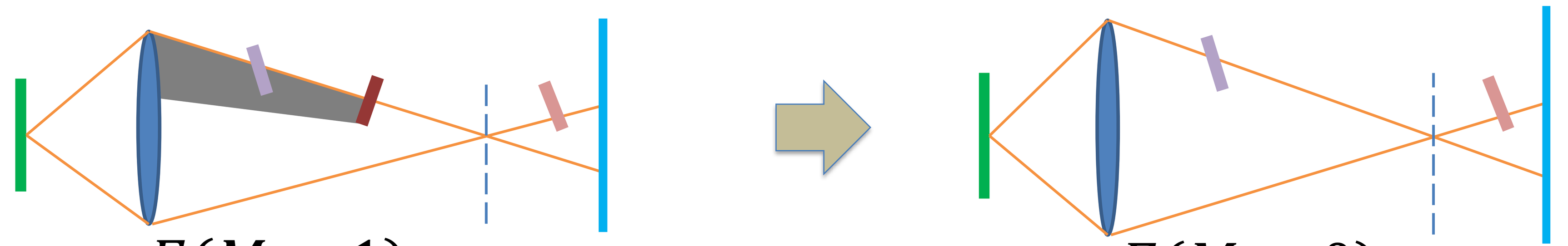
Unknowns for image size WxH: 2KWH continuous and KWH boolean

Multi-layer Matting and Depth Estimation

Boolean optimization for $M \in \{0,1\}$: Gradient of MCMC

Problem Formation: Given focal stack images $\{R_m\}$ Estimate $\{M, d\}$

$$\min_{M,d} \underbrace{\Big(I(\boldsymbol{x}) - R\big(\boldsymbol{x}; M(\boldsymbol{x}), d(\boldsymbol{x})\big)\Big)^2}_{\text{Scene Model}} + \underbrace{S_m(M) + S_d(d)}_{\text{Regulizers}}$$



$E(M=1)$ $E(M=0)$

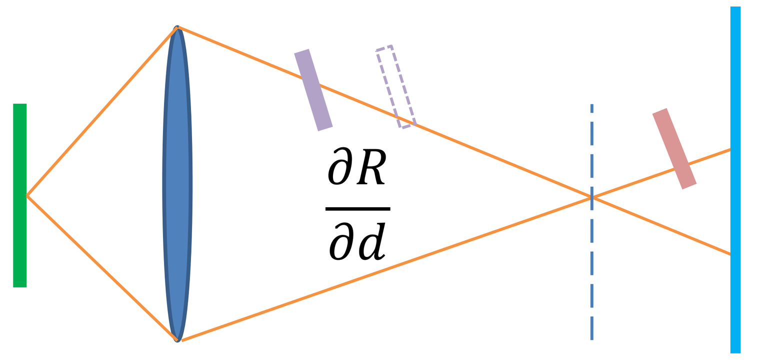
Need to get $\Delta E = E(M=1) - E(M=0)$ for MCMC.
Slow, need to re-render the image per-sampling per-pixel.

Compute:

$$\Delta E = \sum_x \Delta R(\boldsymbol{x})(\Delta R(\boldsymbol{x}) + 2(R(\boldsymbol{x}) - I(\boldsymbol{x}))) + \Delta S_m$$

Differential radiance $\Delta R(\boldsymbol{x})$. Fast to compute analytically.

Continuous optimization for $d \in R$: Gradient Descent

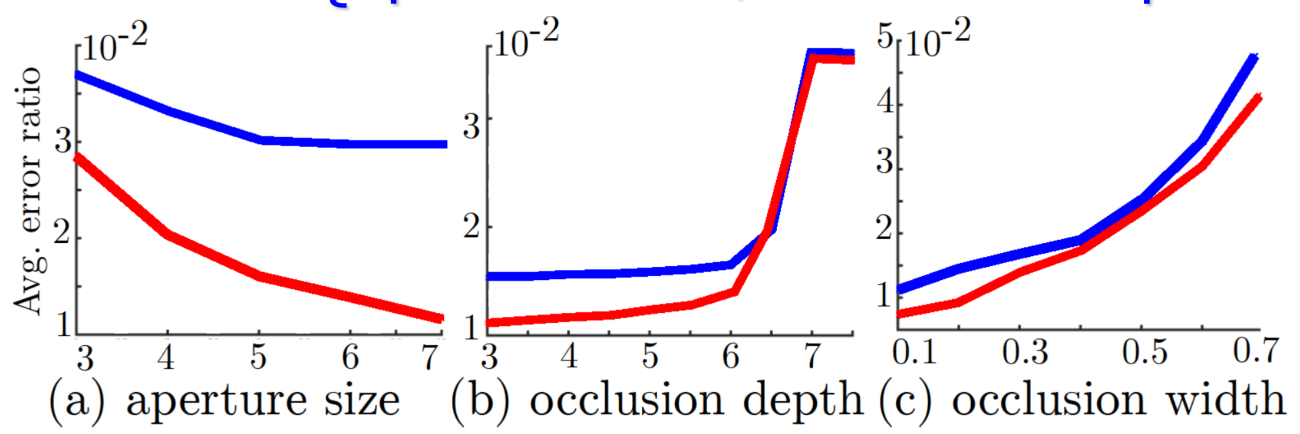


Chain rule:

$$\frac{\partial E}{\partial d} = \frac{\partial E}{\partial R}\frac{\partial R}{\partial d}$$

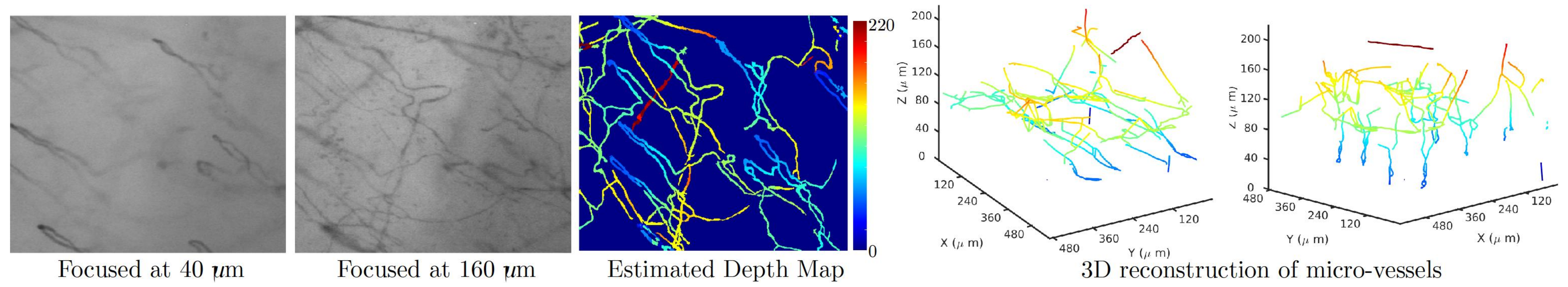
$\frac{\partial R}{\partial d}$ computed analytically.

Performance vs. {aperture size, occlusion depth & width}

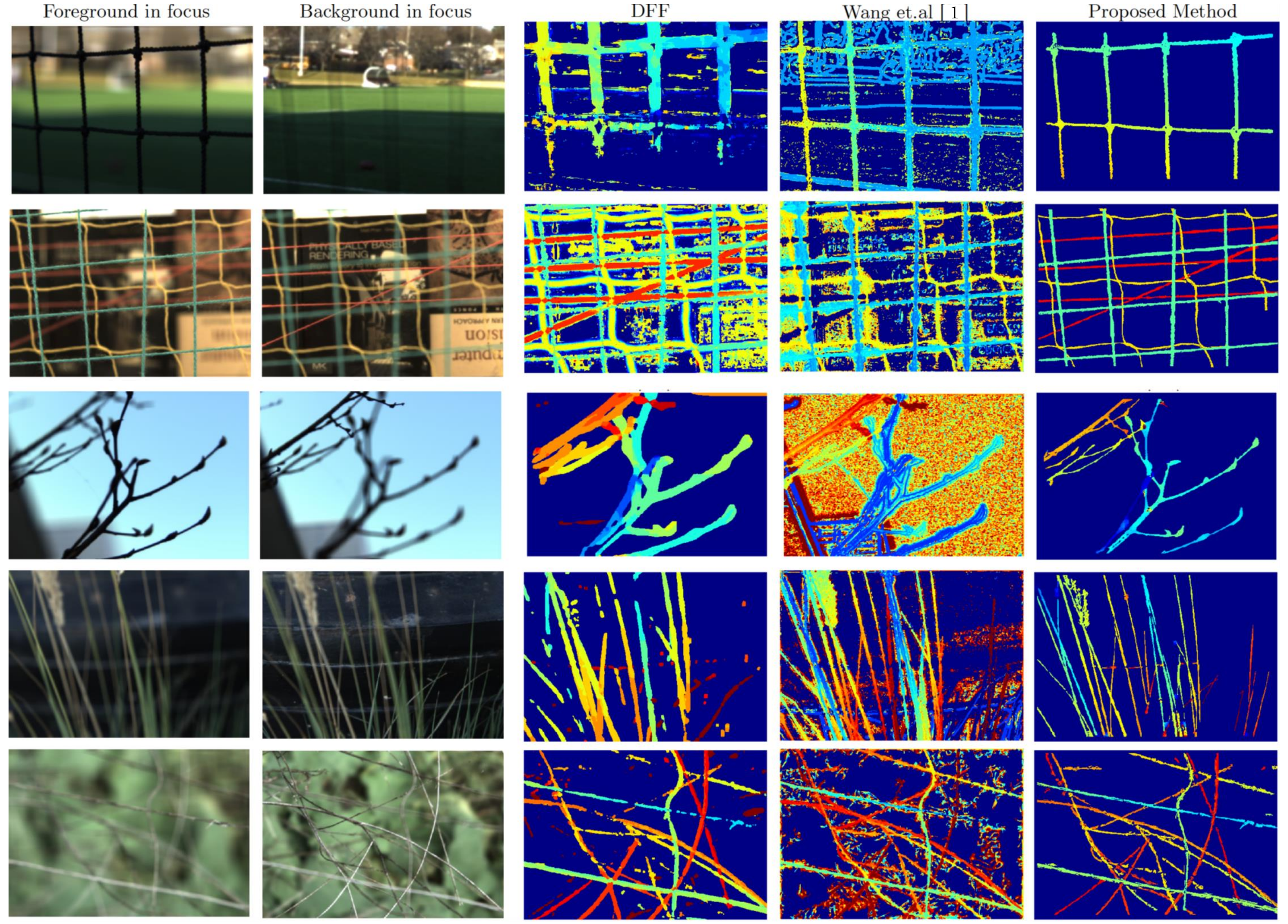


Results

Scene in micro-scale



Scenes in conventional scale



Reference:

[1] *T.C Wang et.al.* Occlusion-aware Depth Estimation Using Light-field Cameras. ICCV 2015
[2] *A. Levin et.al.* A closed form solution to natural image matting. CVPR 2006
[3] *J.Gu et.al.* Removing image artifacts due to dirty camera lenses and thin occluders. TOG 2009