Defocus Deblurring and Superresolution for Time-of-Flight Depth Cameras

Lei Xiao[2,1], Felix Heide[2,1], Matthew O'Toole[3], Andreas Kolb[4], Matthias B. Hullin[5], Kyros Kutulakos[3], Wolfgang Heidrich[1,2]
[1]KAUST. [2]University of British Columbia. [3]University of Toronto. [4]University of Siegen. [5]University of Bonn.

Continuous-wave time-of-flight (ToF) cameras show great promise as low-cost depth image sensors in mobile applications. However, they also suffer from several challenges, including limited illumination intensity, which mandates the use of large numerical aperture lenses, and thus results in a shallow depth of field, making it difficult to capture scenes with large variations in depth. Another shortcoming is the limited spatial resolution of currently available ToF sensors.

In this paper, we address this problem by introducing a new computational method to simultaneously remove defocus blur and increase the resolution of off-the-shelf ToF cameras. We do this by solving a semi-blind deconvolution problem, where prior knowledge of the blur kernel is available. Unlike past ToF deblurring techniques, our approach applies sparse regularizers directly to the latent amplitude and depth images, and supports deblurring ToF images captured with multiple frequencies, phases or exposure time.

Continuous-wave ToF sensors are designed to have an image formation model that is linear in amplitude $\mathbf{a}$, but non-linear in depth $\mathbf{z}$, such that the captured raw sensor data is given as

$$\mathbf{a} \circ \mathbf{g}(\mathbf{z}) \approx \mathbf{a} \circ e^{i(\frac{4\pi f}{c} \cdot \mathbf{z})}, \tag{1}$$

where $\circ$ represents component-wise multiplication of two vectors, f represents the frequency of the continuous-wave modulation, and c is the constant speed of light. The function $\mathbf{g}(\mathbf{z})$ can either be calibrated, or, more commonly, is simply approximated by the complex-valued function from Eq. (1).

We aim to compute a solution to the following ill-posed inverse problem:

$$\mathbf{b} = \mathbf{SK}(\mathbf{z})\left(\mathbf{a} \circ \mathbf{g}(\mathbf{z})\right), \tag{2}$$

where the complex-valued vector $\mathbf{b}$ represents the raw ToF measurements, the real-valued matrix $\mathbf{S}$ is a downsampling operator, and the real-valued matrix $\mathbf{K}(\mathbf{z})$ represents the spatially-varying blur kernel for a given depth map $\mathbf{z}$. The problem is ill-posed because the matrix $\mathbf{SK}(\mathbf{z})$ is usually not invertible, and semi-blind because the matrix $\mathbf{K}(\mathbf{z})$ is known at each depth $\mathbf{z}$. In past work, $\mathbf{S}$ is assumed to be the identity matrix.

Because this inverse problem is ill-conditioned, it's critical to choose appropriate regularizers to reflect prior information on the solution (i.e., sparse edges). Godbaz et al. [3] proposed differential priors that operate on the complex ToF image representing the cosine model, but it remains unclear what a good regularizer should even look like in this space. We instead choose to regularize our solution in the amplitude and depth map space directly. In this paper, we use the second-order total generalized variation [2] for both the amplitude and depth, as shown in Eq. (3) and (4). We show this unified regularization is simple and effective.

$$\Phi(\mathbf{a}) = \min_{\mathbf{y}} \lambda_1 ||\nabla \mathbf{a} - \mathbf{y}||_1 + \lambda_2 ||\nabla \mathbf{y}||_1 \tag{3}$$

$$\Psi(\mathbf{z}) = \min_{\mathbf{x}} \tau_1 ||\nabla \mathbf{z} - \mathbf{x}||_1 + \tau_2 ||\nabla \mathbf{x}||_1 \tag{4}$$

Eq. (5) shows the objective function we aim to minimize. The quadratic term represents a data-fitting error, assuming zero-mean Gaussian noise in the measurements.

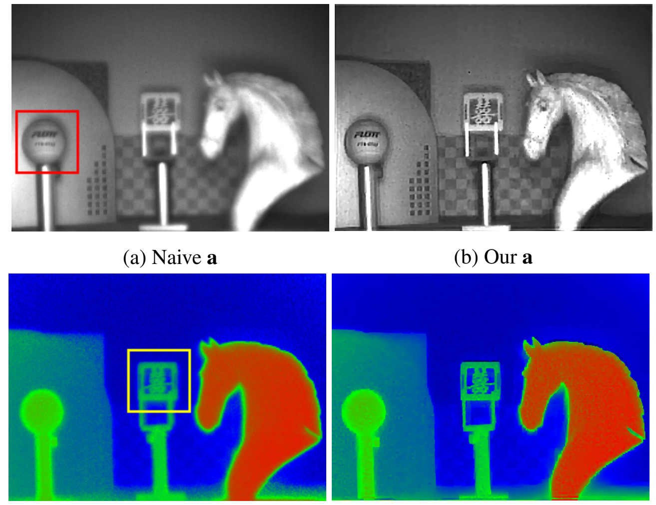
$$(\mathbf{a}, \mathbf{z}) = \operatorname*{argmin}_{\mathbf{a},\mathbf{z}} ||\mathbf{b} - \mathbf{SK}\left(\mathbf{a} \circ \mathbf{g}(\mathbf{z})\right)||_2^2 + \Phi(\mathbf{a}) + \Psi(\mathbf{z}) \tag{5}$$

$\mathbf{a} \circ \mathbf{g}(\mathbf{z})$ in Eq. (1) is highly nonlinear regarding to $\mathbf{z}$. To reduce the computation complexity in this nonlinear problem, the algorithm splits the data-fitting term in the objective into a linear least square and a pixel-wise separable nonlinear least square (LSQ), as in Eq. (6). The scalar ρ defines the relative weight of the splitting term.

$$(\mathbf{a}, \mathbf{z}) = \operatorname*{argmin}_{\mathbf{a},\mathbf{z},\mathbf{c}} \overbrace{||\mathbf{b} - \mathbf{SKc}||_2^2}^{\text{linear LSQ for } \mathbf{c}} + \overbrace{\rho ||\mathbf{c} - \mathbf{a} \circ \mathbf{g}(\mathbf{z})||_2^2}^{\text{separable nonlinear LSQ for } \mathbf{z}} + \Phi(\mathbf{a}) + \Psi(\mathbf{z}) \tag{6}$$

The algorithm alternatively estimates $\mathbf{c}$, $\mathbf{a}$ and $\mathbf{z}$, and update the blur kernel matrix $\mathbf{K}$ at the end of each iteration according to currently estimated $\mathbf{z}$. The subproblems of estimating $\mathbf{a}$ and $\mathbf{z}$ are solved using the alternating direction method of multipliers (ADMM [1]).

Figure 1 and 2 show example results and comparisons with the state-of-the-art. Our method produces much higher quality amplitude and depth, in terms of suppressing the noise, recovering sharp features and reducing flying pixels.



(a) Naive **a** (b) Our **a**
(c) Naive **z** (d) Our **z**

Figure 1: Results on the Character scene. The naive results are generated by the built-in software of ToF cameras. Comparisons on the cropped regions are shown in Figure 2.

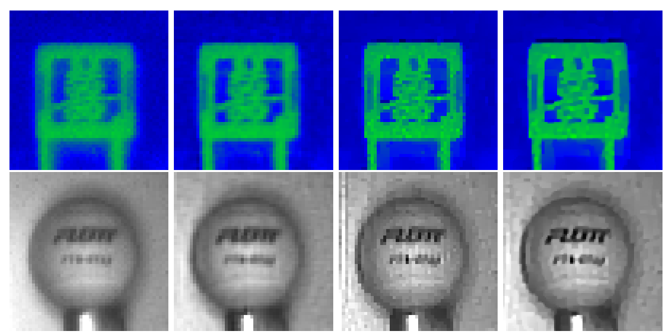


Figure 2: Two insets of the results on the Character scene in Figure 1. *From left to right* shows the result of the naive method, Godbaz et al. [3], ours w/o superresolution, and ours with 2x superresolution.

[1] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011.

[2] K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010.

[3] J. P. Godbaz, M. J. Cree, and A. A. Dorrington. Extending amcw lidar depth-of-field using a coded aperture. In *ACCV 2010*, pages 397–409. Springer, 2011.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.