

# FOCAL-SWEEP FOR LARGE APERTURE TIME-OF-FLIGHT CAMERAS

Sagar Honnunar<sup>1</sup> Jason Holloway<sup>2</sup> Adithya Kumar Pediredla<sup>2</sup>  
 Ashok Veeraraghavan<sup>2</sup> Kaushik Mitra<sup>1</sup>

<sup>1</sup>Indian Institute of Technology, Madras, India <sup>2</sup>Rice University, Houston, Texas, USA

## ABSTRACT

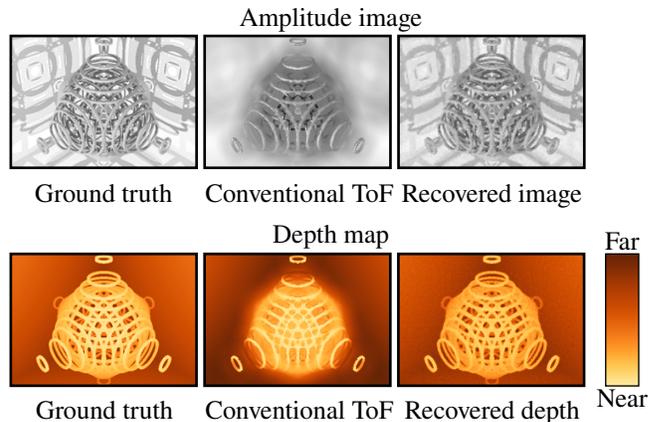
Time-of-flight (ToF) imaging is an active method that utilizes a temporally modulated light source and a correlation-based (or lock-in) imager that computes the round-trip travel time from source to scene and back. Much like conventional imaging ToF cameras suffer from the trade-off between depth of field (DOF) and light throughput—larger apertures allow for more light collection but results in lower light throughput. This is especially limiting in ToF systems since they are active and are limited by illumination power, which eventually limits performance in long-range imaging or imaging in strong ambient illumination (such as outdoors). Motivated by recent work in extended depth of field imaging for photography, we propose a focal-sweep based image acquisition methodology to increase depth-of-field and eliminate defocus blur. Our approach allows for a simple inversion algorithm to recover all-in-focus images which is validated through simulation and experiment. We demonstrate a proof-of-concept focal sweep time-of-flight acquisition system and show results for a real scene.

**Index Terms**— Time-of-flight, lock-in correlation sensors, focal sweep, deblurring.

## 1. INTRODUCTION

Time-of-flight (ToF) sensors are becoming commonplace in many consumer products<sup>1</sup>. However, the current generation of ToF cameras suffer from poor spatial resolution, temporal variation, and sensor noise [1]. The distance measurement also suffers from wiggling [2], internal scattering [3], and temperature related errors [1]. Development of ToF sensors have addressed these limitations with improved designs [4, 5]; however, light throughput remains a significant challenge. Because ToF cameras rely on active light sources, whose intensity cannot be increased indefinitely due to safety and power restrictions, time-of-flight cameras suffer from poor light throughput. For static scenes, one option is to increase exposure duration; however, for natural scenes that do involve either camera motion or subject motion, this results in motion blur. Ultimately, a compromise is struck to increase

<sup>1</sup>Examples include Google’s Project Tango, Microsoft’s Kinect, and Creative’s senz3D



**Fig. 1. Defocus blur in next generation time-of-flight (ToF) cameras:** For scenes with large depth variation (left) ToF cameras with small pixels ( $8 \mu\text{m}$ ) will suffer from large defocus blur outside of the focus plane (center). We propose sweeping the focus over the scene during image exposure and deblurring the captured measurements before recovering the depth and amplitude (right). Top row: Measured amplitude images. Bottom Row: Computed depth map with the color map used throughout this paper.

light collection by using lens with a large aperture at the expense of reducing the depth-of-field (DOF). As the pixel sizes shrink in future generations of ToF sensors, the effect of defocus blur will only magnify, nullifying the anticipated gain in spatial resolution. The middle column in Figure 1 highlights the shortcoming of ToF systems with pixels 5 times smaller than current sensors. In this paper, we propose an alternative configuration for ToF camera systems which drastically increase the DOF while retaining the light collection properties of lenses with high  $f$ -numbers.

Consider an active illumination source that projects a sinusoid,  $\cos(\omega t + \theta)$  onto a scene. Light reflects off of a point in the scene and returns to camera pixel  $p$  as  $I_p \cos(\omega t + \theta + \phi)$  where  $I_p$  is the amplitude of the reflected signal that reaches pixel  $p$ . The phase shift  $\phi$  is related to the distance ( $z_p$ ) between the scene point imaged by  $p$  and the sensor as  $\phi = \frac{2z_p}{c}$  where  $c$  is the speed of light. ToF cameras capture quadrature measurements  $q_i(p)$ ;  $i = \{0, 1, 2, 3\}$  where the value of  $\theta$  is set to be  $i\frac{\pi}{2}$ . Using  $q_i(p)$ , the phase and amplitude can be

computed as

$$z_p = \tan^{-1} \left( \frac{q_1(p) - q_3(p)}{q_0(p) - q_2(p)} \right) \frac{c}{4f\pi}$$

$$I_p = \sqrt{(q_0(p) - q_2(p))^2 + (q_1(p) - q_3(p))^2} \quad (1)$$

Assuming the scene is in focus, the depth and image intensity ( $z_p$  and  $I_p$ ) can be found using straightforward algebra and trigonometry. However, large aperture lenses restrict DOF and light reflecting from scene points outside of this region will be blurred together. Defocus blur is depth-dependent and hence, the point-spread-function (PSF) which characterizes the blur is also depth-dependent resulting in spatially varying blur across the image. The blurry quadrature measurements can be represented as the convolution  $y_i(p) = K(z_p) * q_i(p)$ , where  $K(z_p)$  is the depth-dependent PSF and  $y_i(p)$  are the blurry measurements. In the absence of motion blur,  $K(z)$  can be precomputed and stored for each depth  $z$ . Sharp measurements of  $q_i(p)$  can be recovered by deconvolving  $y_i(p)$  with the appropriate PSF  $K(z)$ . However, we face a classic chicken and egg problem. To compute the PSF, we first need to know the depth  $z_p$ , but to accurately compute  $z_p$ , we need to deconvolve  $y_i(p)$ .

In this paper, we propose to use focal sweep [6, 7, 8, 9] to generate a depth-independent PSF by sweeping the focus plane over the entire scene during each exposure. The resulting PSF kernel  $K$ , is spatially-invariant and known *a priori* which leads to a straightforward non-blind deconvolution of  $y_i$  to recover the quadrature measurements  $q_i$ . (As the blur is uniform, we drop the pixel index  $p$  for convenience.)

**Contribution:** The main contribution of this paper is a system to remove defocus blur from ToF images using focal sweep to generate depth-independent blur. We also present a simple deblurring algorithm for recovering all-in-focus ToF images. We validate our approach via simulation and by capturing real data with a prototype system.

## 2. RELATED WORK

Extending the depth-of-field of imaging systems is a well studied field. We will leverage this research to remove defocus blur in ToF camera systems where large apertures are a design requirement.

### 2.1. Computationally extending depth of field

Focal sweep is a well-known computational photography technique [6, 9] for extended depth of field imaging. Images captured with a large aperture suffer from a depth-dependent blur. In focal sweep imaging, either the distance between the lens and sensor [7] or the subject and lens [8] is varied at a constant rate during the exposure duration, and a single blurry image is captured. Nagahara et al. [7] showed that

focal-sweep images have a depth-independent point-spread-function. Hence, by estimating a single PSF and deconvolving the captured image a sharp, all-in-focus image can be obtained.

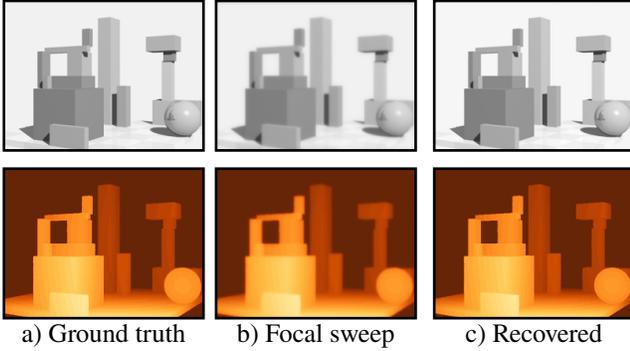
Image deblurring is inherently ill-posed. The frequency response of the defocus blur PSF exhibits nulls in the Fourier domain, precluding direct inversion. Coded apertures have been proposed as a means to overcome defocus blur for consumer cameras with large apertures [10, 11]. Godbaz et al. [12] employed coded aperture for ToF cameras and captured quadrature measurements. As focus plane was static during image acquisition, the measurements have a spatially varying PSF. To account for the varying PSF, the authors employed a spatially variant Landweber deconvolution [13] to obtain sharp quadrature measurements. The initial (blurry) depth estimates determines the blur kernel and this is not refined in their technique. Coded aperture improves the tractability of deconvolving  $y_i$  at the expense of sacrificing light throughput in an already light-limited imaging scenario.

### 2.2. ToF Defocus Deblur

The work that is closest to ours is that of Xiao et al. [14] who addressed the limited spatial resolution and defocus blur in time-of-flight sensors by analyzing the image formation model. As there is a loss of spatial resolution and depth-of-field information, Xiao et al. proposed to incorporate second-order total generalized variation as a prior for both amplitude and depth. In this formulation there are three unknowns: the depth-dependent PSF, the all-in-focus amplitude image, and the all-in-focus depth map. Alternating direction method of multipliers (ADMM) is used to solve for all-in-focus amplitude, and all-in-focus depth images. For simplicity, the PSF is pre-calibrated for each depth and the PSF used for each pixel is updated after updating the amplitude and depth. This algorithm is highly effective, and the state-of-the-art in the field, but it comes at a high computational cost. In contrast, our proposed solution requires a small hardware modification during manufacturing to enable fast and robust image deconvolution. Furthermore, our proposed solution is completely compatible with [14] as the hardware change does not preclude conventional operation.

## 3. PROPOSED METHOD

The defocus-deblur algorithm by Xiao et al. [14] is the current state-of-the-art algorithm for producing sharp depth and amplitude measurements from large aperture ToF camera. However, this method faces some implementation challenges. The non-linear relationship between the depth component and the raw sensor measurements introduces significant numerical difficulties. The deblurring algorithm is computationally expensive and converges slowly. Further, it requires pre-calibration of the blur kernels and a kernel update based on



**Fig. 2. Simulation of recovering all-in-focus ToF images using focus sweep:** a) For a known a scene and depth map we simulate capturing focal sweep ToF images. b) Using the captured images results in amplitude and depth images with a uniform blur. c) All-in-focus amplitude and depth images recovered by deblurring the focal sweep measurements.

the estimated depth in each iteration of the algorithm.

We propose to capture focal sweep TOF measurements. During the capture of each quadrature channel, we move the lens while keeping the sensor static. Thus each quadrature measurements is a focal sweep image, which makes the blur depth-independent [7]. So, unlike Xiao et al. [14], we need not estimate the blur kernel at each pixel. In fact, given the depth range of the scene, sensor pixel size and aperture size, we can analytically compute the focal sweep PSF as given below [7]:

$$\text{PSF}(r, u) = \frac{uf}{(u-f)\sqrt{2\pi rAsT}} \left( \text{erfc}\left(\frac{r}{\sqrt{2}b(0)}\right) + \text{erfc}\left(\frac{r}{\sqrt{2}b(T)}\right) \right)$$

where  $u$  is the mean distance between the lens and the sensor,  $f$  is focal length,  $A$  is the aperture diameter,  $r$  is the distance of the pixel from the center of blur circle,  $s$  is the speed of the sensor,  $T$  is the exposure duration, and  $b$  is blur circle diameter at time  $t$ .

To obtain all-in-focus amplitude and depth images, we propose to individually deblur each of the two independent channels  $h_{re}$  and  $h_{im}$ , which are in turn obtained from the four blurry quadrature measurements  $y_0, y_1, y_2,$  and  $y_3$ . The two independent measurements are given by  $h_{re} = (y_0 - y_2)/2$  and  $h_{im} = (y_1 - y_3)/2$ . Note that the quadrature measurements are obtained by focal sweep and hence, blurred with the depth-independent focal sweep PSF ( $K$ ). Therefore, the two independent measurements  $h_{re}$  and  $h_{im}$  are also the blurred versions of the corresponding sharp channels  $X_{re} = I \cos(\omega t + \phi)$  and  $X_{im} = I \sin(\omega t + \phi)$ . We then find estimates of the sharp channels using non-blind deconvolution with a prior to regularize the total variation norm:

$$\widehat{X}_{re} = \text{argmin}_{X_{re}} \|h_{re} - K * X_{re}\|^2 + \lambda \|X_{re}\|_{TV},$$

with a similar formulation for  $\widehat{X}_{im}$ . Finally, we obtain the sharp amplitude and depth using equations (1) with appropriate substitutions of  $\widehat{X}_{re}$  for  $q_0 - q_2$  and  $\widehat{X}_{im}$  for  $q_1 - q_3$ . The main advantage of our approach over that of the Xiao et al. [14] is that we are solving a linear set of equations and hence computations are very fast. In our experiments, we also find that for a broad range of scenes, TV norm regularization on the ToF measurements yields good results.

## 4. RESULTS

In this section, we demonstrate our proposed technique using simulated scenes and experimental data collected from a prototype camera detailed in [15]. With the availability of all-in-focus depth and amplitude images, we can perform post capture refocusing and artistic renderings such titled depth-of-field imaging, which is detailed at the end of this section.

### Simulation experiments:

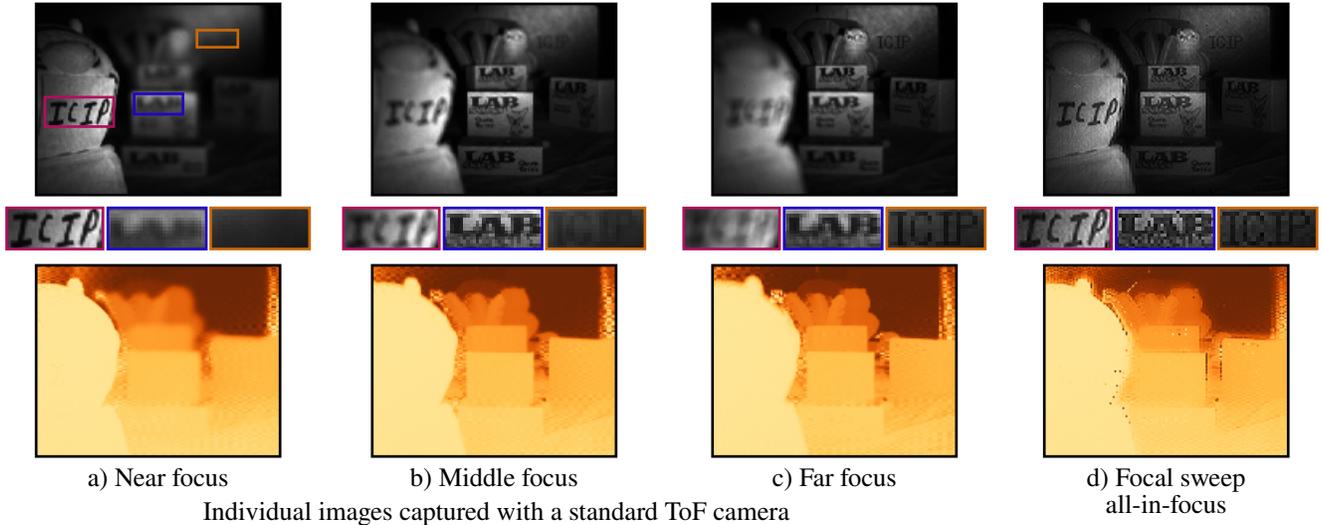
We first obtain the focal sweep quadrature images by simulating our prototype ToF camera which has a pixel size of  $45 \mu\text{m}$ , and is fitted with a  $f/1.4$  lens. For the simulations and the virtual scenes have a depth range of  $[0.48, 1.3]$  meters. We also introduce Gaussian noise to the quadrature measurements amounting to an input SNR of 40 dB.

Figure 2 shows the expected results of using our proposed camera with current generation ToF sensors. We simulate focal sweep during acquisition of each of the quadrature measurements. Directly using the resulting images yields amplitude and depth measurements that have a uniform blur as shown in Figure 2 (b). All-in-focus amplitude and depth images are recovered from the blurry quadrature measurements using the algorithm outlined in Section 3.

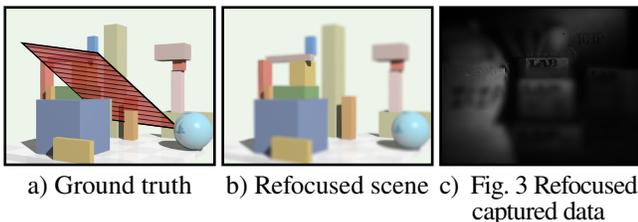
The current pixel size of ToF cameras are quite large  $45 \mu\text{m}$  and the sensor resolution is low, (e.g.  $153 \times 120$  for our PMD sensor). However, with advances in ToF sensor technology we can expect future generations of ToF cameras to have smaller pixel size and higher resolution. This will result in increased defocus blur; the need for computational techniques to recover sharp images in these sensors is be apparent. To simulate the effect smaller pixels will have, we repeat our simulation experiment, but we reduce the pixel size to  $8 \mu\text{m}$ . The center column of Figure 1 shows how conventional ToF cameras will be unable to capture scenes with a large DOF. Using focal sweep during image capture significantly extends the DOF. Even small details near to and far away from the camera can be clearly identified in the amplitude and depth images shown in right column of Figure 1.

### Real experiments:

Capturing a focal sweep image requires the focus to varying continuously during image acquisition. We build a proof-of-concept prototype system using a ToF imager outlined in [15]. Similar to the process described in [7], a large aper-



**Fig. 3. Recovering all-in-focus images for a real ToF system** As a proof-of-concept we use the ToF system described in [15] with an  $f/1.4$  lens to capture a sequence of images of the scene shown in (a-c). A sequence of 26 measurements is recorded by increasing lens-sensor separation by  $50 \mu\text{m}$  between positions.. Integrating these measurements yields the comparable focal sweep data. (d) Deblurring the quadrature measurements provides sharp all-in-focus images. Top row: Amplitude images. Middle row: Detail views of three patches located in the near-, mid-, and far-field. Bottom row: Depth images.



**Fig. 4. Post-capture tilted plane refocusing:** a) An all-in-focus scene is refocused along an arbitrary plane (red, striped) shown in b). Color images are used to help visualize the rendering. c) The recovered scene from Figure 3(d) is refocused along a nearly horizontal focal plane.

ture lens ( $f/1.4$ ) is mounted on a translation stage to vary the distance between the sensor and lens. A sequence of 26 images is captured by increasing separation distance of the lens and sensor in  $50 \mu\text{m}$  increments to create a focal stack. A focal sweep image is then formed by integrating the stack into a single blurry image for each of the four quadrature measurements.

Figure 3 depicts a scene captured with our setup while Figure 3 (a-c) shows the corresponding amplitude images and depth maps captured at near-, mid-, and far-focus positions. Notice that defocus blur significantly reduces image quality for objects outside of the DOF. By deblurring each of the quadrature measurements and following the procedure in Section 3, we recover an all-in-focus amplitude image and depth map (Figure 3 (d)). Detailed subsets below each of the amplitude measurements shows image patches from three depths, corresponding to the focal planes in Figure 3 (a-c).

Only after deblurring the focal sweep quadrature measurements are all three patches sharp, and each is of comparable quality to the corresponding in-focus patches.

#### Post capture refocusing and tilted depth-of-field:

Once we obtain all-in-focus amplitude and depth images, we can render the scene with different lens and aperture settings computationally without a physical lens. In Figure 4 (a-b), we illustrate the ability to tilt the lens and render an artistic scene using color images to act as visual aids. Figure 4(b) shows the rendering effect when we tilt the virtual lens to focus along the red, striped plane in Figure 4(a). In Figure 4(c), the all-in-focus result of Figure 3(d) is refocused along a nearly horizontal focal plane. Only with the sharp depth maps are such fantastical renderings possible without resorting to bulky specialized equipment.

## 5. CONCLUSION

In this paper, we have presented an approach to extending the DOF in ToF cameras using focal sweep. Using this method, all-in-focus images can be obtained by performing a non-blind deconvolution of the quadrature measurements. The simplicity of our recovery algorithm will enable straightforward scale-up to handle increasing spatial-resolution and defocus blur in future generations of ToF cameras. We validate our proposed solution both in simulation and by recording data of using a prototype system. Finally, we show that the recovered all-in-focus data can be used to render novel scenes otherwise impossible to capture with conventional ToF cameras.

## 6. REFERENCES

- [1] Peter Fursattel, Simon Placht, Christian Schaller, Michael Balda, Hannes Hofmann, Andreas Maier, and Christian Riess, "A comparative error analysis of current time-of-flight sensors," 2015.
- [2] Damien Lefloch, Rahul Nair, Frank Lenzen, Henrik Schäfer, Lee Streeter, Michael J Cree, Reinhard Koch, and Andreas Kolb, "Technical foundation and calibration methods for time-of-flight cameras," in *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*, pp. 3–24. Springer, 2013.
- [3] Wilfried Karel, Sajid Ghuffar, and Norbert Pfeifer, "Modelling and compensating internal light scattering in time of flight range cameras," *The Photogrammetric Record*, vol. 27, no. 138, pp. 155–174, 2012.
- [4] F Chiabrando, D Piatti, and F Rinaudo, "Sr-4000 tof camera: further experimental tests and first applications to metric surveys," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 38, no. 5, pp. 149–154, 2010.
- [5] James Mure-Dubois and Heinz Hügli, "Real-time scattering compensation for time-of-flight camera," in *Proceedings of the ICVS Workshop on Camera Calibration Methods for Computer Vision Systems*, 2007.
- [6] Oliver Cossairt and Shree Nayar, "Spectral focal sweep: Extended depth of field from chromatic aberrations," in *Computational Photography (ICCP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–8.
- [7] Hajime Nagahara, Sujit Kuthirummal, Changyin Zhou, and Shree K Nayar, "Flexible depth of field photography," in *Computer Vision—ECCV 2008*, pp. 60–73. Springer, 2008.
- [8] Gerd Häusler, "A method to increase the depth of focus by two step image processing," *Optics Communications*, vol. 6, no. 1, pp. 38–42, 1972.
- [9] Changyin Zhou, Daniel Miao, and Shree K Nayar, "Focal sweep camera for space-time refocusing," 2012.
- [10] Ashok Veeraraghavan, Ramesh Raskar, Amit Agrawal, Ankit Mohan, and Jack Tumblin, "Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing," *ACM Trans. Graph.*, vol. 26, no. 3, pp. 69, 2007.
- [11] Anat Levin, Rob Fergus, Frédo Durand, and William T Freeman, "Image and depth from a conventional camera with a coded aperture," in *ACM Transactions on Graphics (TOG)*. ACM, 2007, vol. 26, p. 70.
- [12] John P Godbaz, Michael J Cree, and Adrian A Dorrington, "Extending amcw lidar depth-of-field using a coded aperture," in *Computer Vision—ACCV 2010*, pp. 397–409. Springer, 2011.
- [13] Louis Landweber, "An iteration formula for fredholm integral equations of the first kind," *American journal of mathematics*, vol. 73, no. 3, pp. 615–624, 1951.
- [14] Lei Xiao, Felix Heide, Matthew O’Toole, Andreas Kolb, Matthias B Hullin, Kyros Kutulakos, and Wolfgang Heidrich, "Defocus deblurring and superresolution for time-of-flight depth cameras," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 2376–2384.
- [15] Achuta Kadambi, Refael Whyte, Ayush Bhandari, Lee Streeter, Christopher Barsi, Adrian Dorrington, and Ramesh Raskar, "Coded Time of Flight Cameras : Sparse Deconvolution to Address Multipath Interference and Recover Time Profiles," *ACM Transactions on Graphics*, vol. 32, no. 6, pp. 1–10, 2013.