Snapshot super-resolution time-of-flight imaging by PSF engineering and untrained deep neural-network prior

Hodaka Kawachi, Tomoya Nakamura, Yasushi Makihara, Yasushi Yagi
SANKEN, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, JAPAN
E-mail:kawachi@am.sanken.osaka-u.ac.jp

Abstract: Indirect time-of-flight (ToF) sensors are widely used as depth-map measurement devices, but their spatial resolution is generally lower than that of ordinary image sensors. To solve this problem, we propose a snapshot spatial super-resolution method based on point-spread-function engineering for encoding subpixel information and untrained deep neural-network for decoding a super-resolution image. We experimentally confirmed that the effective spatial resolution of the depth map was improved with the proposed method.

Keywords: Indirect ToF, Coded imaging, Depth map, Super-Resolution, Untrained deep neural-network prior

1. Introduction

In recent years, depth-map data have been used in a wide range of applications from automatic driving to person authentication. Among various depth-map acquisition methods, the indirect time-of-flight (ToF) method, which measures optical ToF by phase-difference detection with emitting periodic light in the time domain, has rapidly been gaining popularity because of its low cost and compact device design [1]. One challenge of the indirect ToF sensor is lower spatial resolution due to the complexity of the circuit design than general image sensors, and it is difficult to solve just by the improvement of hardware design. To overcome the physical resolution limit imposed by the sensor device, application of digital super-resolution technique to indirect ToF sensing is promising. A typical model-based super-resolution method is the multi-frame super-resolution method based on multiple shots. The effectiveness of this method has been well demonstrated in the past decades; however, it sacrifices temporal resolution in sensing.

On the other hand, the snapshot super-resolution method using coded compressive sensing has recently been studied [2]. While the multi-frame super-resolution method directly measures sub-pixel information by taking multiple images with small movements, the compressive super-resolution method optically encodes sub-pixel information in a single or few coded measurements and decodes it by computation [3]. To decode a sensor-resolution measured image into a super-resolution image, it is necessary to solve an ill-posed inverse problem. To realize it, compressive sensing adopts sparse modeling of the object that enables to solve of the ill-posed inverse problem uniquely. In a past study, Li et al. proposed multi-shot-type compressive indirect ToF imaging using image-plane sub-pixel encoding [4]. In this study, we propose snapshot compressive super-resolution ToF imaging using point-spread-function (PSF) engineering.

2. Proposed method

In this study, we focus on the pulse-driven indirect ToF method [5] among indirect ToF methods. In the pulse-driven indirect ToF method, as shown in Fig. 1, a time-domain pulse light is irradiated from a sensor-device onto a target object, and the reflected light from the object is detected by a sensor with multiple pulse-like time windows. The windows have slightly different offsets in time, and the distance information to the target is obtained by calculating the detected brightness ratio with the multi-windowed measurement data. This operation is independent to pixels.

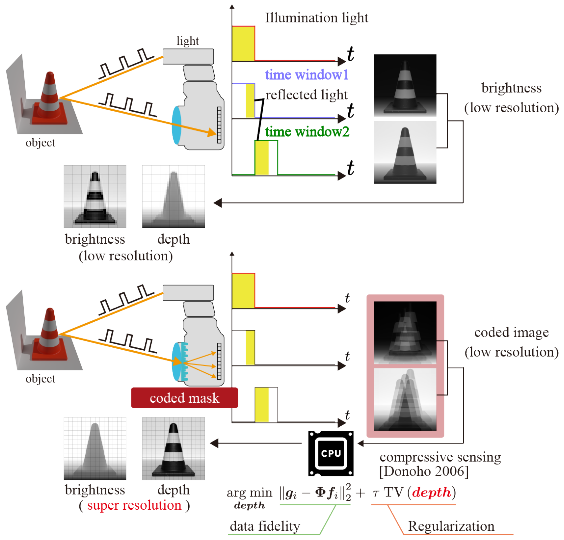


Fig. 1. Overview of conventional (upper) and proposed (bottom) method.

To perform super-resolution using compressed sensing, losslessly encoded observations of sub-pixel spatial information of the target are required. In this study, we implement optical encoding by placing an encoding optical element at the aperture plane of a ToF camera. This encoding allows us to essentially keep the sub-pixel information in the measured image even for low-resolution observations.

In general compressed sensing for luminance images [2], the squared error between the measured and estimated coded luminance image is used for the objective function to be minimized in decoding. The same way can be implemented to each time-windowed data in indirect ToF imaging as shown in Fig. 2. To minimize the objective function, the error is back propagated as shown by the red arrows. Since super-resolution is an ill-posed inverse problem, many studies apply regularization using the total variation to the estimated solution. In this study, instead of optimizing the images for each time window obtained by indirect ToF sensing independently, we optimize the depth-map and luminance images simultaneously.

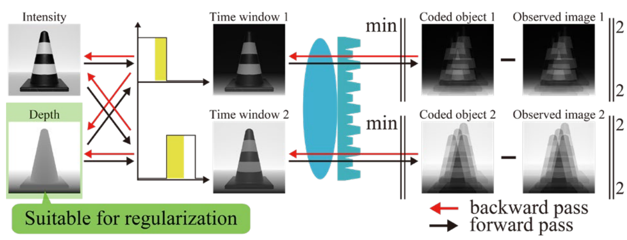


Fig. 2. Reconstruction Pipeline.

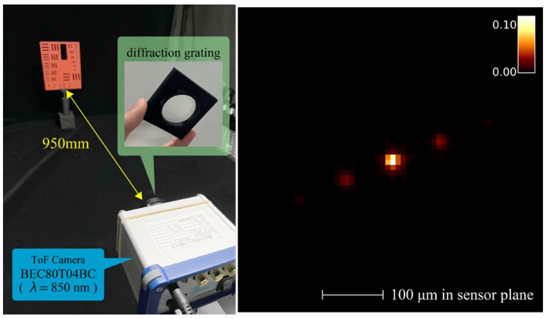


Fig. 3. (Left) experimental setup and (right) observed PSF.

This idea is based on the assumption that the main purpose of the indirect ToF camera is not the intermediate luminance data but depth-maps, and depth-maps are less variable, and thus more compressive than luminance data.

3. Experiment

We verified our method by an optical experiment. We have also verified the accuracy quantitatively by numerical simulations, which has already been presented in Ref. [6]. Optical setup is shown in Fig. 3. We used a pulse-driven indirect ToF camera (BEC80T04BC by Brookman Technology) for imaging. The object (a hole-punched resolution chart) was placed in 950 mm front of the camera. The coded observation was implemented by placing a diffraction grating at the pupil plane. In this experiment, we used a grating with 30 µm grooves at a pitch of 60 µm along to 24.15-degree oblique direction. An observed PSF at a distance of the place object is shown in Fig. 3. As shown in the figure, the spot is divided into three to encode subpixel information with a single multiplexing measurement. Figure 4 shows observed encoded images for two different time windows with temporal offset of 75 nsec. The ToF camera equips 240 × 320 pixels, and the central 150 × 150 pixels were used for reconstruction. The difference of the brightness corresponds to the scene depth, and the subpixel shift of the self-image is multiplexed in each image. For super-resolution depth-map reconstruction with 300 × 300 pixels, we used a typical compressive-sensing decoding algorithm called as TwIST [7] which utilizes the pipeline of Fig. 2 with regularization by two-dimensional total variation and a newer algorithm called as untrained neural-network prior (UDN) which uses untrained U-net as a generator for regularization [8].

Figure 5 shows the depth-map reconstruction results. Both TwIST and UDN reconstruction results improved the spatial resolution of the depth-map compared to conventional bicubic-interpolation. For example, high-frequency structure (close-ups of Fig. 5) were clearly resolved with preserving edges. On the other hand, current super-resolution results show an amplified noise compared to the bicubic-interpolation results. In particular, the UDN result shows non-uniform depth map for the holes in the target. We will deal with this problem in a future work.

4. Conclusion

In this study, we proposed a single-shot digital super-resolution ToF imaging by using compressed sensing. We experimentally and qualitatively confirmed that spatial resolution of the depth-map was improved compared to a conventional interpolation method. The method proposed in this study can be implemented just by placing an additional phase modulator in front of the commercial ToF camera so that it is very easy to use in practical applications.

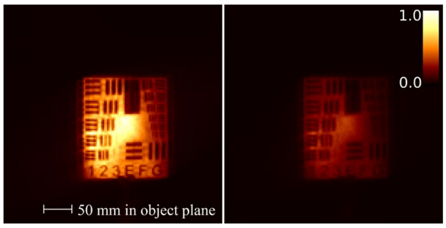
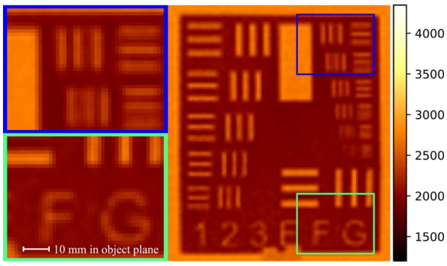
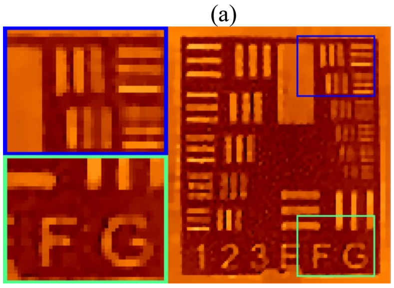


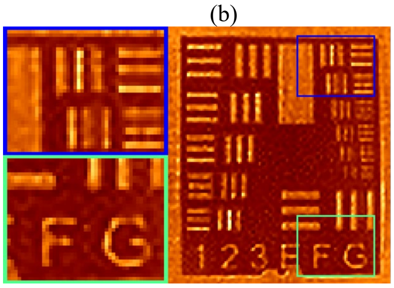
Fig. 4. Observed images.
(left: time window 1, right: time window 2).



(a)



(b)



(c)

Fig. 5. Reconstruction results of depth-map by
(a) bicubic interpolation, (b) TwIST, and (c) UDN.

References

[1] S. Kawahito, I. A. Halin, T. Ushinaga, T. Sawada, M. Homma, and Y. Maeda, "A CMOS time-of-flight range image sensor with gates-on-field-oxide structure," IEEE Sens. J. **7**, 1578–1585 (2007).

[2] D. L. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory **52**, 1289–1306 (2006).

[3] A. Stern, *Optical Compressive Imaging* (CRC Press, 2016).

[4] F. Li, H. Chen, A. Pediredla, C. Yeh, K. He, A. Veeraraghavan, and O. Cossairt, "CS-ToF: High-resolution compressive time-of-flight imaging," Opt. Express **25**, 31096–31110 (2017).

[5] K. Yasutomi and S. Kawachito, "Lock-in Pixel Based Time-of-Flight Range Imagers: An Overview," IEICE Trans. Electron. **E105.C**, 301–315 (2022).

[6] H. Kawachi, T. Nakamura, Y. Makihara, and Y. Yagi, "Snapshot spatial super-resolution ToF sensing based on point-spread-function design and depth map regularization," in *The 25th Meeting on Image Recognition and Understanding (MIRU2022)*, (2022), p. OL4A-1, in Japanese.

[7] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration," IEEE Trans. Image Process. **16**, 2992–3004 (2007).

[8] K. Monakhova, V. Tran, G. Kuo, and L. Waller, "Untrained networks for compressive lensless photography," Opt. Express **29**, 20913–20929 (2021).