Indirect-ToF system optimization for sensing range enhancement with patterned light source and adaptive binning

Seung-chul Shin, Myeonggyun Kye, Il-pyeong Hwang, Taemin An, Kyu-Min Kyung, Duhyeon Kwak, Hakbeom Jang, Hogyun Kim, Jaeil An, Sunhwa Lee, Yundong Chang, Seongwon Jo, Junghwan Yoo, Myoungoh Ki, Min-Sun Keel, Yitae Kim, Jesuk Lee and Duckhyun Chang

System LSI Division, Samsung Electronics Co., Ltd., Hwaseong, Gyeonggi-do, Korea,
E-mail: sc1225.shin@samsung.com

Abstract — An indirect time-of-flight (ToF) system is presented to extend the depth-sensing range. The proposed 4-tap ToF sensor system is equipped with a patterned light source in the transmitter side, focusing the optical power on the selected spot points, maintaining the average optical power compared with the conventional flood-type light source. By focusing the optical power on the spots, the signal-to-noise ratio and demodulation contrast (DC) are enhanced, and the depth noise (precision) performance is improved. The adaptive binning scheme is implemented on the receiver side to decide the optimal binning method according to the amount of ambient light; the DC is used as a metric for this decision. With a proper binning mode selection, the depth performance can be optimized by smaller read noise at the indoor condition, and a larger full well capacity at the outdoor condition. Finally, we achieved a 3.4× working distance increase even under the strong external light conditions.

I. Introduction

Recently, 3D information has been widely used in various platforms such as mobile phones, augmented reality (AR) headsets, and autonomous vehicles. For accurate and precise 3D information measurement, the depth sensors based on the time-of-flight (ToF) principle have been widely investigated [1-3].

The ToF system generates reference information by launching an optical signal through the light transmitter (Tx) using a vertical-cavity surface-emitting laser (VCSEL) at an 850-nm or 940-nm wavelength and acquires distance information by detecting the received optical signal with an image sensor-based receiver (Rx). According to the received information, the ToF is divided into two types: a direct-ToF (d-ToF) and an indirect-ToF (i-ToF). First, d-ToF measures an absolute time based on the single-photon avalanche diode (SPAD). It is possible to measure long distances and reduce Tx power consumption by the high sensitivity of the SPAD. However, the spatial resolution is limited due to large pixel size, and high bias voltage is required. Additionally, short distance performance is not acceptable due to high noise floor from high-sensitive SPAD pixel circuits.

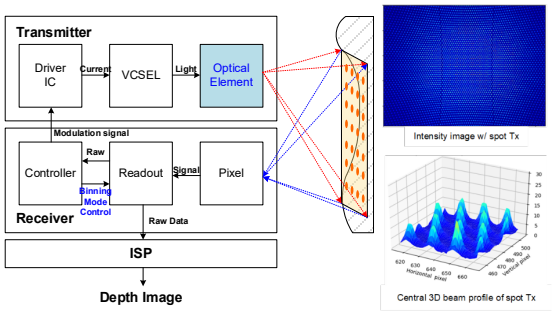


Fig. 1. The proposed i-ToF system configuration with patterned light source in the transmitter and a binning controller in the receiver

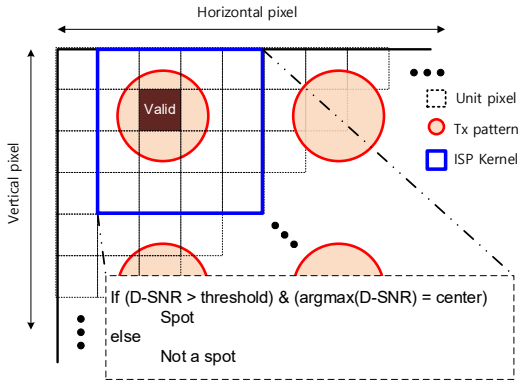
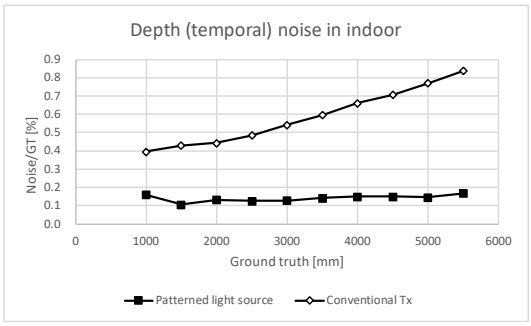
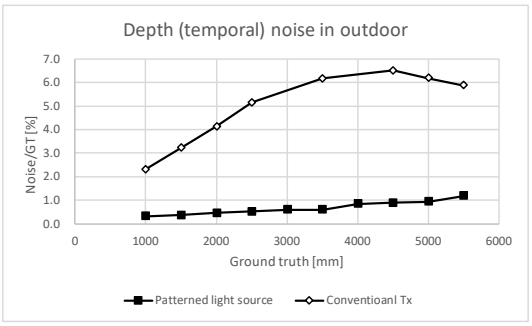


Fig. 2. The concept of spot detection algorithm in ISP chain.

Second, an i-ToF measures phase delay through the modulation and demodulation processes. It has the advantage of high spatial resolution because it is easy to implement commonly-used CIS-based pixel structure with small pixel pitch and does not require additional high-voltage power supply, which is favorable for mobile applications. Because of the signal integration during the exposure time, the noise characteristic is rather robust. However, the i-ToF sensors suffer from the limited sensitivity at the infra-red (IR) wavelength, resulting in a shorter working range than the d-ToF sensor [4]. In addition, during the integration of photo-generated electrons, the electron from an external light is also integrated at a large ambient light condition, or an outdoor condition, and the pixels can be easily saturated, which leads to a limited distance.

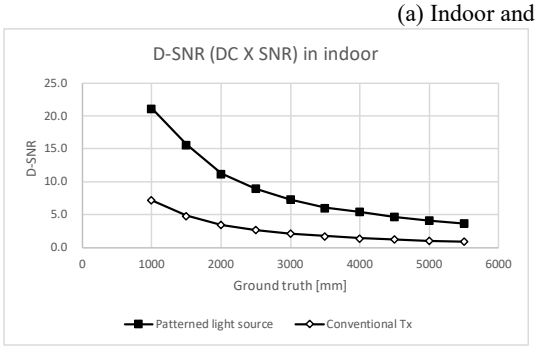


(a)

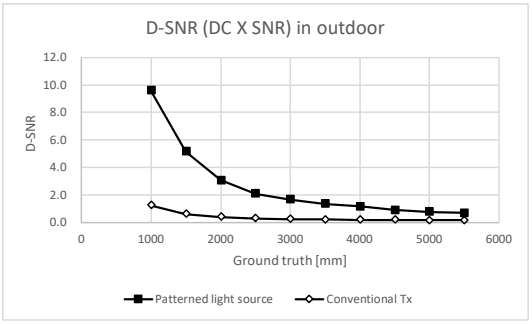


(b)

Fig. 3. Depth noise with a patterned light source (spot output) and conventional light source (flood output). (a) Indoor and (b) outdoor environment



(a)



(b)

Fig. 4. D-SNR with a patterned light source (spot output) and conventional light source (flood output). (a) Indoor and (b) outdoor environment

In this paper, we propose a method to eliminate distance constraints that occur in the i-ToF sensors. The light power utilization method in the Tx is proposed to increase light power per pixel, while maintaining the average power consumption. In the Rx, optimal operation is suggested to minimize noise at the indoor operation and to maximize signals at the outdoor operation.

II. TRANSMITTER OPTIMIZATION

Fig. 1 shows the proposed i-ToF system configuration. In contrast to the conventional transmitter with a diffuser for spreading the laser beam from a VCSEL array, our proposed system adopts a diffractive optical element (DOE) to focus the Tx laser power on the selected spots. An intensity image of the patterned light is shown in the right side of Fig. 1. The number of spot patterns is about 3.8 k, and the optical power is 3.4 W.

In the Rx, we implemented the spot detection algorithm in the ISP chain, as shown in Fig. 2 with the pseudocodes. Only the valid pixel groups with high confidence values are selected by sweeping the kernel in an entire confidence image. Among the pixels in a group, the central pixel is selected as the spot. The threshold value, kernel size, and sweeping related parameters can be configured.

Fig. 3 shows the depth temporal noise results by the spot-type patterned light source and the conventional flood-type light source with the same 3.4-W optical power at the indoor and outdoor conditions. We changed the location of a flat chart with 94 % reflectivity from 1 m to 5.5 m and measured depth temporal noise by the standard deviation. The indoor condition was with only room LED lighting. To emulate the outdoor condition, we illuminated ~70-klx halogen lamp on the chart. Even though the optical powers are identical (3.4 W) for both Tx types, the depth noise with the patterned light source is reduced by 78 % and 87 % than those of conventional Tx in the indoor and outdoor conditions, respectively. This is because power density for the patterned light source is larger. The improved performance is analyzed by a signal-to-noise ratio (SNR) and a demodulation contrast (DC), which is defined as follows, assuming photon shot noise is dominant:

$$\mathrm{SNR} \cong \sqrt{Intensity} = \sqrt{\frac{A_0+A_{90}+A_{180}+A_{270}}{4}} \quad (1)$$

$$\mathrm{DC} = \frac{Amplitude}{Intensity} = \frac{0.5\times\sqrt{(A_0-A_{180})^2+(A_{90}-A_{270})^2}}{0.25\times(A_0+A_{90}+A_{180}+A_{270})} \quad (2)$$

where A_θ (θ = 0, 90, 180, and 270) means sampled signals at the θ-degree phase points. In the indoor condition with a low ambient light, the focused optical power leads to the SNR improvement because the "spot" pixels receive more optical power. However, the DC values are almost similar regardless of the Tx types in the indoor conditions, because intensity and amplitude are changed at a same rate if the light power comes only from the Tx. On the contrary, the DC with the patterned light source is enhanced in the outdoor condition, because higher optical power density increases amplitude more than intensity. We merged SNR and DC as the depth-SNR (D-SNR) as follows:

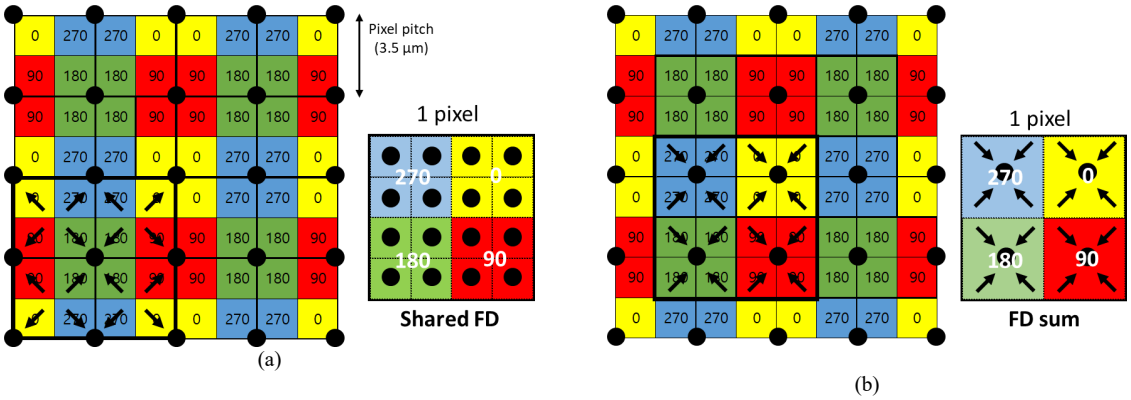


Fig. 5. Pixel structure and (a) digital binning in a high ambient light and (b) analog binning in a low ambient light

$$\mathrm{D-SNR} = \mathrm{DC} \times \mathrm{SNR} = \frac{Amplitude}{\sqrt{Intensity}}. \quad (3)$$

Fig 4. shows the D-SNR results with both Tx types in the indoor and outdoor conditions. The D-SNR values of the patterned light source are about 3.6 times and 6.2 times higher than the those of the conventional Tx in indoor and outdoor, respectively, which match well with the depth noise enhancement factors of 4.1 and 6.6 from Fig. 3.

III. Receiver Optimization

In the Rx side, a well-known SNR enhancement technique, the pixel binning, will be adopted to improve the depth noise. Fig. 5 shows the simplified pixel array of the 1.2Mp ToF sensor [3] and two supported pixel binning methods. The 4-phase sampling taps are marked as 0, 90, 180, and 270 with different colored squares, and the black circles are shared floating diffusion nodes to readout the charge from each tap. Fig. 5(a) shows the 2×2 digital binning; each dedicated readout circuit reads out the charges in each tap, then are averaged in a digital domain. So, the full well capacity (FWC) of the FD is fully used for each tap and we can increase exposure time in the outdoor condition. Fig. 5(b) represents the 2×2 analog binning, which sums up the charges from each adjacent tap in a single FD node. As charges from four taps are accumulated in a single FD node, only a quarter of FWC of the FD node can be used, limiting exposure time in the outdoor condition, and consequently increasing the depth noise. However, in the 2×2 analog binning mode, the number of readout circuits to read a combined pixel is only one, whereas the 2×2 digital binning needs four readout circuits; the read noise of the analog binning is reduced compared to that of the digital binning.

Therefore, we propose an adaptive binning to choose the best binning mode depending on ambient light. In the indoor condition, we select the analog binning, to minimize read noise. In the outdoor condition, we select the digital binning, to maximize the exposure time to overcome ambient light. As a criterion for a binning mode selection, we will use the DC. Since the intensity is proportional to the ambient light power, the DC value will be decreased at the large ambient light (refer to (2)). We verified the it by measuring the effect of the ambient light on the DC value. We placed a 94% reflectivity chart at a 1-m distance from the sensor, and the amount of ambient light was controlled by the halogen lamp, to calculate the DC value, as shown in the dotted lines in Fig. 6.

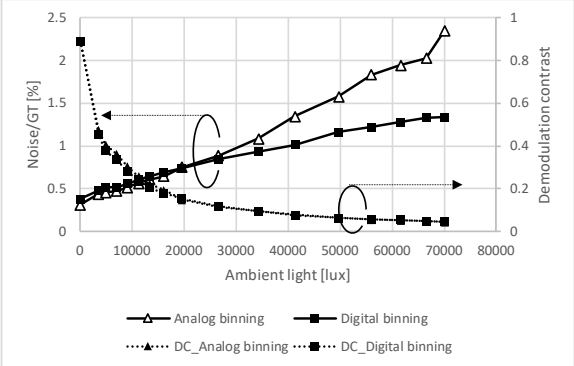
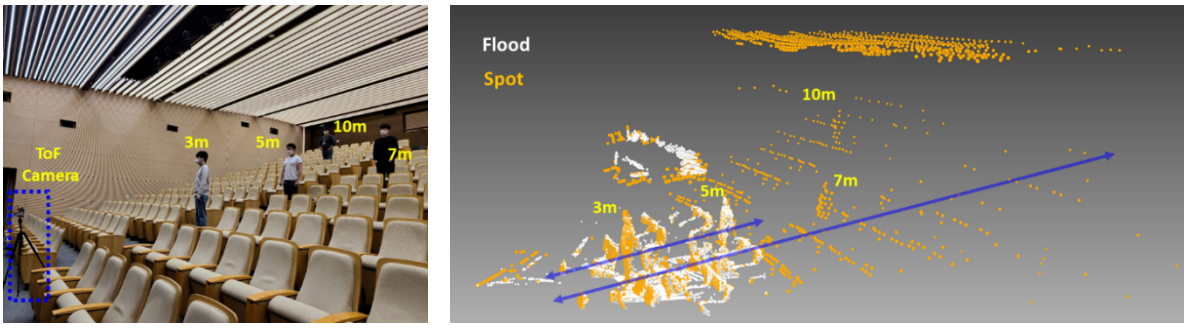


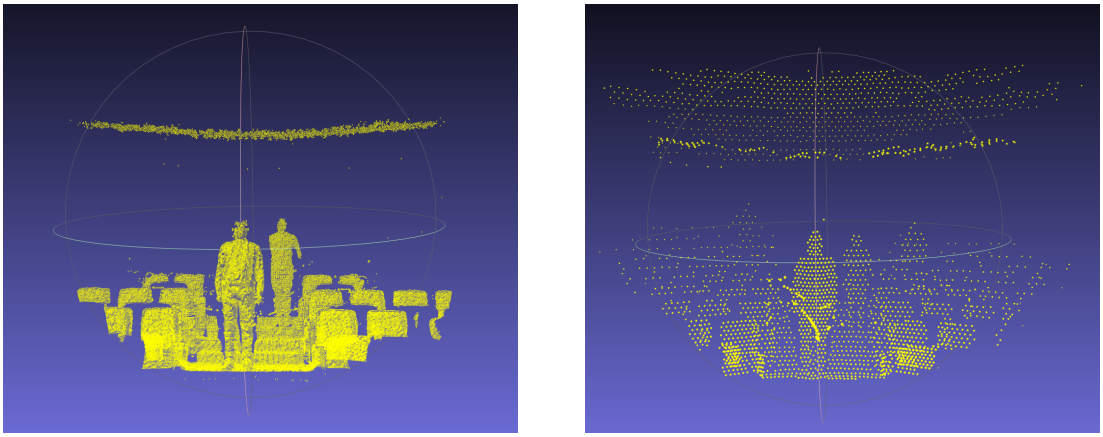
Fig. 6. Depth noise and demodulation contrast of the analog and digital binning with the amount of ambient light

The DC is inversely proportional to the amount of ambient light, regardless of the binning modes. We also measured the depth noise of the two binning cases for the given ambient light condition, with the solid lines in Fig. 6. When the illumination of the ambient light is below 20 klx, the analog binning showed lower depth noise than the digital binning and the trend is reversed with the ambient light more than 20 klx. Thus, the DC value for selecting the binning modes can be decided at the cross point of the two depth noise curves, where DC is 0.3.



(a) (b)

Fig. 7. Qualitative depth performance comparison using point cloud: (a) measurement environment (b) point cloud image to show the depth range (while: conventional Tx, orange: patterned light source).



(a) (b)

Fig. 8. Depth performance comparison according to XY resolution: (a) conventional Tx (b) patterned Tx

IV. Results

The results of comparative experiments between the conventional and the pattered light sources are shown in Fig. 7. The experiment was conducted in the indoor auditorium, as shown in Fig. 7(a). The noise filtering and spot detection function in the ISP chain are applied. As shown in Fig. 7(b), the maximum measurable distance of the patterned light source is longer by about three time than that of the conventional Tx, as discussed with depth noise comparison results in Section II. In Fig. 8, the XY-resolution was compared by the Tx types. For conventional flood type light source, the spatial resolution is determined by the pixel array of the sensor, whereas the resolution is limited by the number of patterned spots for the patterned light source.

V. Conclusion

We propose a way to increase the working distance of i-ToF through system optimization. We adopted the D-SNR metric to analyze the distance increase and depth noise properties by the types of optical transmitters. To increase the working distance, the pattern light source was used to increase the intensity of light per pixel, and the adaptive binning selection was adopted to optimize the signal-to-noise ratio (SNR) according to the ambient light conditions. Using the patterned light source and a binning mode reduces the effective depth spatial resolution. This is expected to be overcome through a fusion algorithm based on the RGB color image plus the depth image, and further optimization is expected depending on the target applications.

References

[1] Kondo, Satoshi, et al. "5.1 A 240× 192 Pixel 10fps 70klux 225m-Range Automotive LiDAR SoC Using a 40ch 0.0036 mm 2 Voltage/Time Dual-Data-Converter-Based AFE." in *IEEE ISSCC Dig. Tech. Papers*, 2020.

[2] Bamji, Cyrus S., et al. "IMpixel 65nm BSI 320MHz demodulated TOF Image sensor with 3µm global shutter pixels and analog binning." in *IEEE ISSCC Dig. Tech. Papers*, 2018.

[3] Keel, Min-Sun Keel *et al*. "A 4-tap 3.5µm 1.2Mpixel Indirect Time-of-Flight CMOS Image Sensor with Peak Current Mitigation and Multi-User Interference Cancellation," in *IEEE ISSCC Dig. Tech. Papers,* 2021.

[4] Kwon, Yonghun, *et al*. "A 2.8 µm Pixel for Time of Flight CMOS Image Sensor with 20 ke-Full-Well Capacity in a Tap and 36% Quantum Efficiency at 940 nm Wavelength." in *IEEE IEDM*, 2020.