Range ambiguity cancellation for continuous-wave indirect ToF image sensors

Bumsik Chung*, Il-Pyeong Hwang*, Daeyun Kim, Jonghan Ahn, Daeho Kim, Myunghan Bae, Jiheon Park, Inho Song, Hogyun Kim, Minsik Kim, Taemin An, Hyeyeon Lee, Youngkyun Jeong, Min-Sun Keel, Juhyun Ko, Jesuk Lee

System LSI Division, Samsung Electronics Co., Ltd., Hwaseong, Gyeonggi-do, Korea,
Phone: +82-10-9454-7073, E-mail: bumsik.chung@samsung.com
Phone: +82-10-7148-8533, E-mail: i0308.hwang@samsung.com *Equally contributed authors

Abstract — We propose a method for range ambiguity cancellation to eliminate phase wrapping errors in indirect time-of-flight (iToF) systems. The method introduces the concept of random packet modulation combined with an ambiguity mask algorithm. A new indicator of the normalized amplitude has been introduced to effectively cancel out depth signals outside an unambiguous range with only a single depth frame. When combined with the random packet modulation, this normalized amplitude can extend the minimum detectable object reflectivity by 14%p, compared to conventional iToF. This technique reduces motion artifacts by 75% compared to the dual-frequency method, enabling the generation of a depth image using a single modulation frequency.

I. INTRODUCTION

Indirect time-of-flight (iToF) depth-sensing systems are gaining considerable attention as one of the leading depth-sensing technologies because iToF sensors can obtain precise distances for 3D imaging by calculating the phase difference between the modulation signals and the received signals. Since the modulation signals are periodic, iToF systems exhibit cyclic characteristics. This characteristic leads to ambiguity issues, known as phase wrapping errors or depth folding errors [1], [2].

The concept of phase wrapping errors is illustrated in Fig. 1. When objects are placed within the unambiguous range, there are no issues with distance calculation. However, when an object is placed outside the unambiguous range, conventional iToF systems are confused between the folded depth and the actual depth. This issue is particularly critical for short-range applications, such as face detection and hand gesture recognition.

Previous iToF systems address this issue by adopting dual-frequency modulation [2] or a time-gaiting [3]. The dual-frequency modulation method superposes two different frequency modulations, extending the maximum unambiguous range to the least common multiple of the two unambiguous ranges from the two modulation frequencies. The time-gaiting architecture can discard signals outside the range of concern, which helps to eliminate phase wrapping errors. However, both of these previous techniques require multiple frames, which results in severe motion artifacts.

In this paper, we propose a method to discard phase wrapping errors by eliminating phase information inherently through the adjustment of random packet modulation with an ambiguity masking algorithm, rather than extending the unambiguous range or splitting into sub-ranges. Since the method requires only a single frame

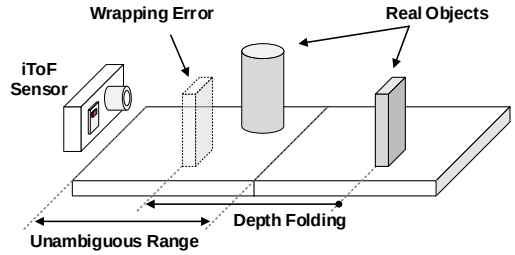
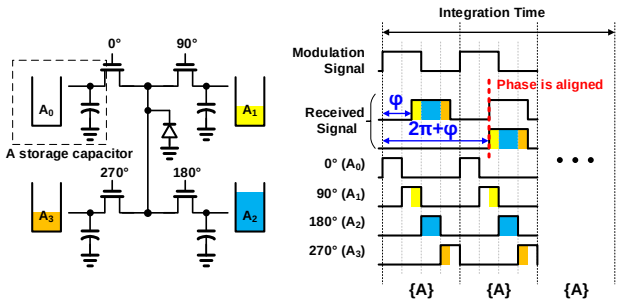


Fig. 1. Concept of depth folding.



(a) (b)

Fig. 2. (a) Partial schematic of the 4-tap demodulation pixel, and (b) modulation signal and four phase-sampling signals for the iToF operation.

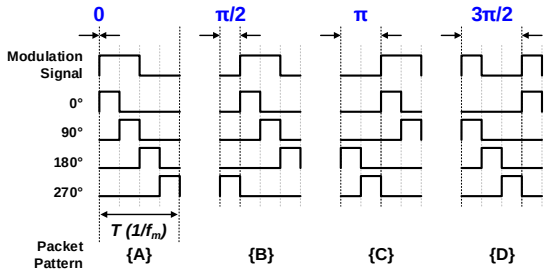


Fig. 3. Four different phase-shift packet patterns for the proposed random packet modulation: pattern {A}, {B}, {C}, and {D}.

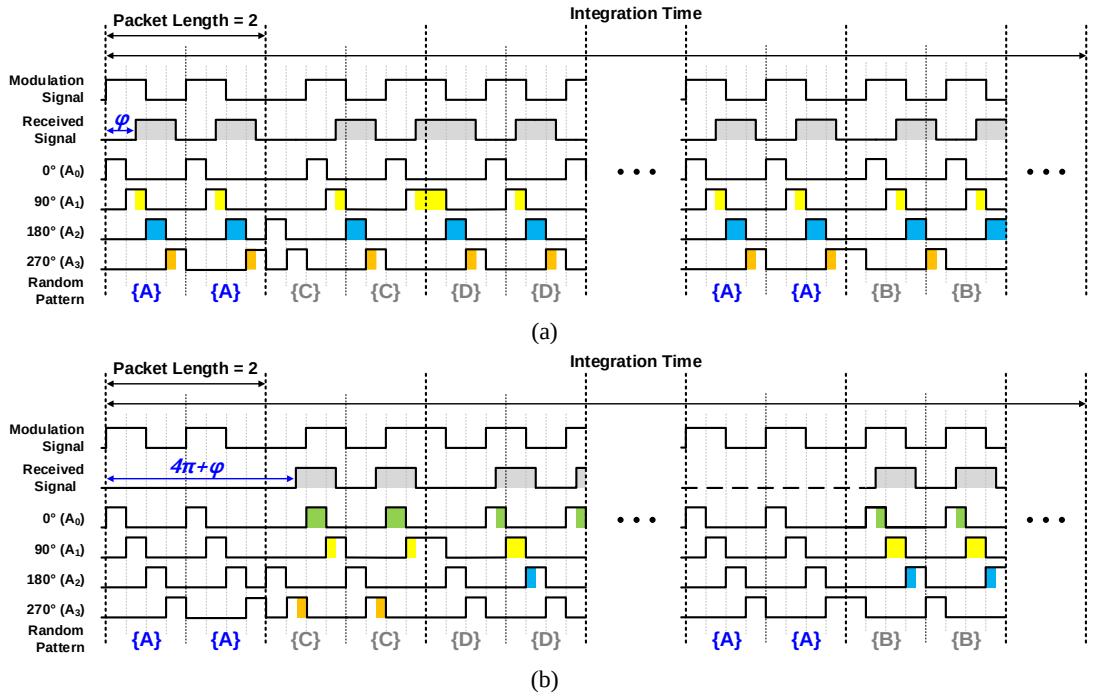


Fig. 4. Proposed range ambiguity cancellation method. Timing diagrams of random packet modulation (PL = 2) with (a) phase shift of φ, and (b) phase shift of $4\pi + \varphi$ (where, $0 < \varphi < 2\pi$).

per depth image, it provides an efficient solution for short-range applications.

II. Basic Theory

Fig. 2 shows a typical 4-tap demodulation pixel and the waveforms for the modulation and phase-sampling signals in a conventional iToF system [4]. The modulation signal is a continuous-wave with a period of $1 / f_m$ (f_m: modulation frequency). The received signal with a phase shift (φ) is continuously integrated into each of the four storage capacitors, which are controlled by the four-phase sampling signals (0°, 90°, 180°, and 270°) with a 25% duty cycle. The charge signals stored in each storage capacitor are denoted as A_0, A_1, A_2, and A_3. The phase shift (φ) of the 4-tap iToF system in Fig. 2 can be calculated as $\varphi = tan^{-1}[(A_1 - A_3) / (A_0 - A_2)]$, and then the distance (d) is calculated as $d = (c / 2f_m) \times (\varphi / 2\pi)$, where c is the speed of light. Since a phase shift exceeding 2π radians (i.e., $2\pi \times n + \varphi$, where n is a positive integer) will be wrapped into φ, the unambiguous range is calculated as $c / 2f_m$. The DC indicates charge transfer efficiency and is calculated as the ratio of amplitude to intensity, as follows, where the amplitude and intensity are defined as (7) and (8) in [5], respectively.

$$DC = \frac{Amplitude}{Intensity} = \frac{2\sqrt{(A_0 - A_2)^2 + (A_1 - A_3)^2}}{A_0 + A_1 + A_2 + A_3}. \quad (1)$$

III. Proposed Scheme

The proposed range ambiguity cancellation scheme is based on randomization of modulation and demodulation signals. As shown in Fig. 3, four different phase-shift packet patterns, {A}, {B}, {C}, and {D}, are defined as different starting phases with 0, $\pi/2$, π, and $3\pi/2$, respectively. This starting phase is applied both to the modulation and sampling signals. The packet length (PL) is a number of repeating packet patterns, in a range from PL = 1 to PL = ∞ (no randomization).

Timing diagrams for PL = 2 are shown in Fig. 4. If the phase shift (φ) is less than 2π radians the received signal is correctly sampled and stored into the appropriate taps. As shown in Fig. 4(a), the received signal stays within the same packet pattern boundary as phase-sampling signals. The first packet pattern ({A}-{A}) is aligned with the corresponding sampling pattern. Thus, the following randomly repeating signals are also demodulated properly despite the inevitable DC degradation. However, if φ is larger than $4\pi + \varphi$, the received signal is distributed across incorrect taps because the first received pattern ({A}-{A}) is not aligned with the sampling pattern ({C}-{C}) and this misalignment continues afterward (see Fig. 4(b)).

The normalized DC in (2) is used to explain the DC characteristics for different PLs in Fig. 5:

$$Normalized\ DC = \frac{DC\ (\varphi)}{DC\ (\varphi=0)}. \quad (2)$$

Each curve in Fig. 5 has a different slope, which is determined by the PL. For example, in the case of conventional iToF with no random packet modulation, all the modulation signals are stored without any loss during the integration time. This means that the normalized DC remains flat across all the phase shifts. However, the normalized DC characteristics change with phase shift by adopting the random packet modulation. Since the succeeding packet patterns after the first packet are not the same each other, signal loss begins after 0 phase-shift and the amount of lost signals increases with the phase-shift. As a result, the normalized DC of PL = 2 decays from 1.0 to 0 at a 4π radians phase shift. From this observation, we may use this normalized DC to filter out depth signal above the unambiguous range with a threshold of 0.5 at PL =2. However, in reality, the normalized DC can not be

used because it changes with ambient light in outdoor conditions; a fixed threshold level cannot be defined.

To address this issue, we introduce a new concept of the normalized amplitude (Norm.Amp.), which is insensitive to ambient light, as follows:

$$Norm.Amp. = amplitude \times depth^2. \quad (3)$$

Fig. 6 shows the normalized amplitude plot with different PLs at the object reflectivity (R) of 100%. The plot exhibits a similar characteristic to the normalized DC between 0 and 2π radians. Notably, the normalized amplitude is reset to zero and rebounds at every 2π radians, because the measured depth is folded and restarts from 0. Consequently, to cancel out-of-range distances, thresholds of 0.3 and 0.03 may be set for the conventional iToF (PL = ∞) and the proposed iToF with PL = 2, respectively.

Different from the normalized DC, the normalized amplitude is affected by the object's reflectivity. To select proper threshold values with normalized amplitude, the upper and the lower bound of the threshold should be studied with object reflectivity. The mask thresholding methodology is shown in Fig. 7 The first lobe on [0, 2π) and the second lobe on [2π, 4π), both from Fig. 6, are overlapped together in Fig. 7(a) and (b), for the conventional iToF (PL = ∞) and PL = 2, respectively. The folded signals from the outside of the unambiguous range should be eliminated, whereas the valid signals from black objects with low reflectivity within the unambiguous range should be preserved. By observing Fig. 7, the upper bound of the threshold will be determined by the low reflectivity, whereas the lower bound will be set by the maximum reflectivity (R = 100%). To enable range ambiguity cancellation, the upper bound should always be larger than the lower bound, which limits the minimum value for the low reflectivity. Therefore, the threshold is bounded by (4), as follows:

$$\max_{[2\pi,4\pi)} Norm.Amp.(R_{white} = 100\%) < Threshold < \min_{[0,2\pi)} Norm.Amp.(R_{black}), \quad (4)$$

When the threshold is set too high, black objects with low reflectivity (R_{black}) are masked. On the contrary, if the threshold is set too low, white objects with high reflectivity are folded within an unambiguous range. The system margin is then defined by (5), to represent detection capability for black objects:

$$system\ margin\ [\%] = 100 - R_{black}\ [\%]. \quad (5)$$

In the case of the conventional iToF as shown in Fig. 7(a), if the thresholding region (green shaded area) is set to be 5% to ensure stable operation, the upper bound (red line) should be larger than or equal to 0.3 (R = 30%). Therefore, the conventional iToF system with no random packet modulation can detect objects with R_{black} greater than 30%, resulting in a system margin of 70%.

On the other hand, the proposed system with PL = 2 provides a larger system margin, as shown in Fig. 7 (b). Due to the dramatic decrease of the lower bound (black line), the thresholding region can be placed in the lower

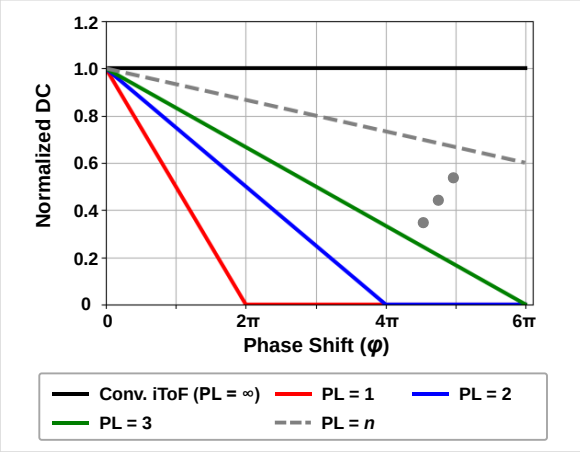


Fig. 5. Normalized DC across phase shift with different packet lengths.

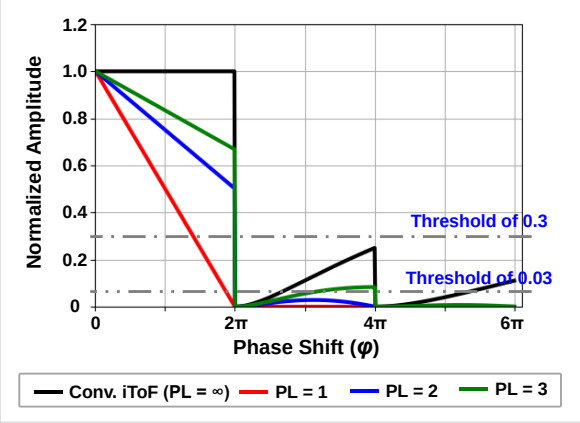


Fig. 6. Normalized amplitude across phase shift with different packet lengths.

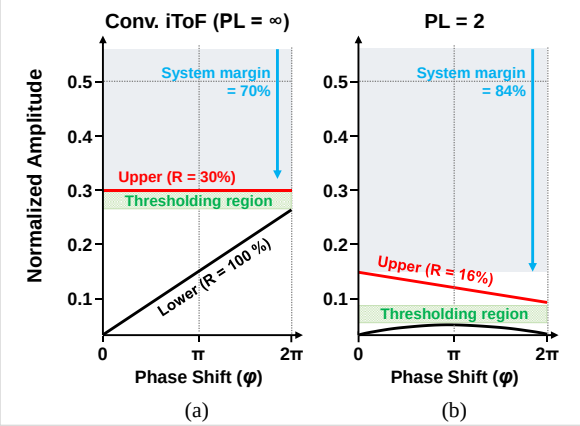


Fig. 7. Examples of the mask thresholding methodology: (a) conventional iToF (PL = ∞), and (b) PL = 2.

area, below 0.08. When the threshold is set to 0.08, it can detect black objects with R_{black} > 16% over the full range (0 to 2π radians). Thus, the system margin of our proposed method is 84%, which is 14 %p better than that of the conventional iToF system.

IV. Measurement Results

Fig. 8 shows the measurement results. We set up the environment, as shown in Fig. 8(a), by placing objects within the appropriate range where the wrapping errors could occur. Since a 100-MHz modulation frequency is

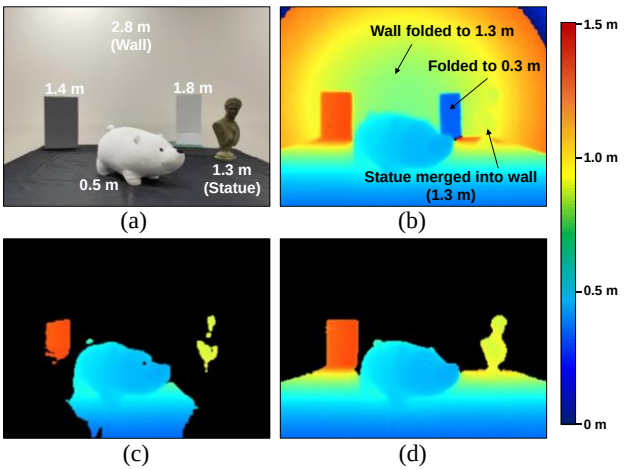


Fig. 8. Comparison of depth images using a 100-MHz modulation frequency (unambiguous range of 1.5 m). (a) Test environment with objects at different distances. (b) Conventional iToF. (c) Conventional iToF with normalized amplitude threshold at 0.3. (d) The proposed method with (PL = 2) with normalized amplitude threshold at 0.03.

used for a 1.5-meter unambiguous range, the starting point for depth folding errors is 1.5 meters. Fig. 8(b) shows that objects beyond 1.5 meters are folded or merged with objects at different depths when using the conventional iToF setup. For example, a statue at 1.3 meters is not detected because a wall at 2.8 meters is incorrectly calculated as 1.3 meters to be merged with the statue at 1.3 meters. Fig. 8(c) demonstrates that a conventional iToF system with a normalized amplitude threshold of 0.3 can eliminate the folded depth. However, the edges of objects with low amplitude are not detected, and the black floor with low reflectivity is partially removed. In contrast, Fig. 8(d) from the proposed scheme with PL = 2 shows that objects within the unambiguous range are perfectly preserved, whereas objects outside the unambiguous range are effectively removed.

This scheme can capture a depth image from a single frame without motion artifacts. A comparison of motion artifacts between the dual-frequency modulation and the proposed method is shown in Fig. 9. As shown in Fig. 9(a), a rotating wheel is placed at 0.5 meters and rotates at 100 rpm (10.5 rad/s), which emulates a fast-moving hand in hand gesture recognition for the extended reality (XR) applications. In Fig. 9(b), the dual-frequency modulation correctly calculates the distance without phase wrapping errors due to an extended unambiguous range. However, there are motion artifacts in 12 pixels, on the spoke of the wheel. However, as shown in Fig. 9(c), the proposed random packet modulation efficiently reduces motion artifacts with no phase wrapping errors. The depth degradation of 3 pixels is due to shuffling of the 4-tap demodulation pixels, to minimize 4-tap mismatch [6]. As a result, the proposed method improves the motion artifact performance by 75% compared to the conventional one.

V. Conclusion

We propose a range ambiguity cancellation scheme to eliminate phase wrapping errors for iToF systems. A random packet modulation method is introduced to cancel

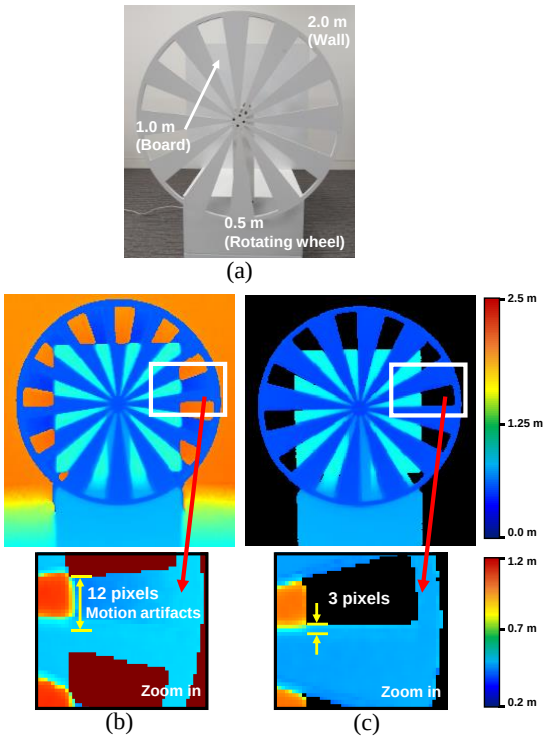


Fig. 9. Measured motion artifact with depth images. (a) Test environment with a rotating wheel. (b) Dual-frequency method. (c) The proposed method (PL = 2) with normalized amplitude threshold at 0.03.

out out-of-range signals with the aid of an ambiguity mask. For efficient masking with proper threshold and packet length, a new indicator of normalized amplitude, which is insensitive to the ambient light, is proposed to extend the system margin for low-reflectivity objects. Moreover, the proposed scheme naturally reduces motion artifacts by capturing and filtering a depth image in a single frame. The ambiguity mask algorithm is relatively simple, compared to the multiple-frame merger algorithm in multiple-frequency. Therefore, the proposed ambiguity cancellation can reduce the burden on backend depth signal processing for an efficient depth sensing system.

References

[1] C. Bamji *et al.*, "A Review of Indirect Time-of-Flight Technologies," *IEEE Transactions on Electron Devices*, vol. 69, no. 6, pp. 2779-2793, June 2022.

[2] A. D. Payne *et al.*, "Multiple Frequency Range Imaging to Remove Measurement Ambiguity," *Optical 3-D Measurement Techniques IX*, vol. 2, pp. 139-148, 2009.

[3] S. Kawahito *et al.*, "Hybrid Time-of-Flight Image Sensors for Middle-Range Outdoor Applications," *IEEE Open Journal of the Solid-State Circuits Society*, vol. 2, pp. 38-49, 2022.

[4] M. -S. Keel *et* al., "A 1.2-Mpixel Indirect Time-of-Flight Image Sensor With 4-Tap 3.5-μm Pixels for Peak Current Mitigation and Multi-User Interference Cancellation," *IEEE Journal of Solid-State Circuits*, vol. 56, no. 11, pp. 3209-3219, Nov. 2021.

[5] R. Lange and P. Seitz, "Solid-state time-of-flight range camera," *IEEE J. Quantum Electron.*, vol. 37, no. 3, pp. 390-397, Mar. 2001.

[6] S. -C Shin *et al.*, "A 1.2 Mp indirect-ToF sensor with on-chip ISP for low-power and self-optimization." in *Proc. Int. Image Sens. Workshop (IISW)*, May. 2023