A Half-Pulse 2-Tap Indirect Time-of-Flight Ranging Method with Sub-Frame Operation for Depth Precision Enhancement and Motion Artifact Suppression

Chia-Chi Kuo[1], Rihito Kuroda[1,2]
[1] Graduate School of Engineering, Tohoku University,
6-6-05, Aza-Aoba, Aramaki, Aoba-ku, Sendai, Miyagi, Japan 980-8579
[2] New Industry Creation Hatchery Center, Tohoku University
TEL: +81-22-795-4833, Email addresses: kuo.chiachi.s2@dc.tohoku.ac.jp, rihito.kuroda.e3@tohoku.ac.jp

ABSTRACT

This paper presents a new 2-tap indirect time-of-flight (iToF) ranging method using half-pulse (HP) modulation and sub-frame 4-phase sampling. The proposed operation and derived theoretical depth noise equations were verified by the experiments, which achieve >25% and >29% depth precision enhancement for 0.4-3 m range with proposed HP1 and HP2 method, respectively. The motion artifact suppression is also demonstrated using sub-frame operation that provides clear depth images for a scene with a moving object.

INTRODUCTION

3-D imaging system has been developed rapidly and applied to various fields. In these days, there are more emerging opportunities that can make use of this technology such as autonomous vehicles, industrial automation, computer-human interaction, and so on. To provide a reliable spatial information, a high performance 3-D camera is desired. The 2-tap indirect time-of-flight (iToF) based depth image sensors have been a popular choice that can realize a range imager with high resolution, low power consumption, low GPU computation complexity and good system compactness, simultaneously [1], [2].

To achieve a better ranging precision, increasing the system signal-to-noise ratio (SNR) and the modulation frequency have always been the key targets, which were revealed to be inversely proportional to the depth noise [3]. However, the existed trade-off between the imaging performances should be taken into concerns for the realization of an iToF system.

Firstly, a higher full-well capacity (FWC) design can obtain a higher SNR whereas a larger pixel size and longer exposure period are required. Secondly, due to the phase wrapping, the unambiguous range will reduce while applying a higher modulation frequency. Although the multi-frequency synthesis had been reported to be useful to extend the valid scope range [4], ensuring the robustness of phase unwrapping and a good demodulation contrast (DC) under high-speed modulation remain a challenge [4], [5]. Besides, the intensive power consumption during the exposure period, which mainly lead by the toggling demodulation gates and the emitted light, may reduce the robustness and durability of a ranging system.

On the other hand, the conventional 2-tap 4-phase (2T-4PH) operation requires two successive frames to reconstruct a background light (BGL) canceled depth image [6]. However, the requirement of temporal multiplexing will induce an increase of motion artifact due to the frame-to-frame processing latency.

In this article, a new 2-tap ranging method using half-pulse (HP) modulation with sub-frame 4-phase sampling is introduced to address these issues.

OPERATION PRINCIPLE

Fig.1 shows the ranging methodologies for 2T-4PH iToF system using the conventional continuous square-pulse (SP) [7] and proposed HP modulation, where c is the speed of light. The modulation frequency, cycle time, and pulse width of SP and HP are denoted by f_m, T_C, T_{SP} and T_{HP}, respectively.

The received light that demodulated by Tap1 and Tap2 during two exposure periods, $PH(0,\pi/2)$, are denoted by $Q1(0,\pi/2)$ and $Q2(0,\pi/2)$, respectively. Note that the emitted light amplitude of HP is doubled from SP to have an equal averaged light intensity during the modulation.

For the distance (d) calculation, Fig. 2 shows the equations to obtain the light traveling time (T_{ToF}), where R is the time-shift ratio which is defined from 0 to 4 through the time-window, TW(1) to TW(4).

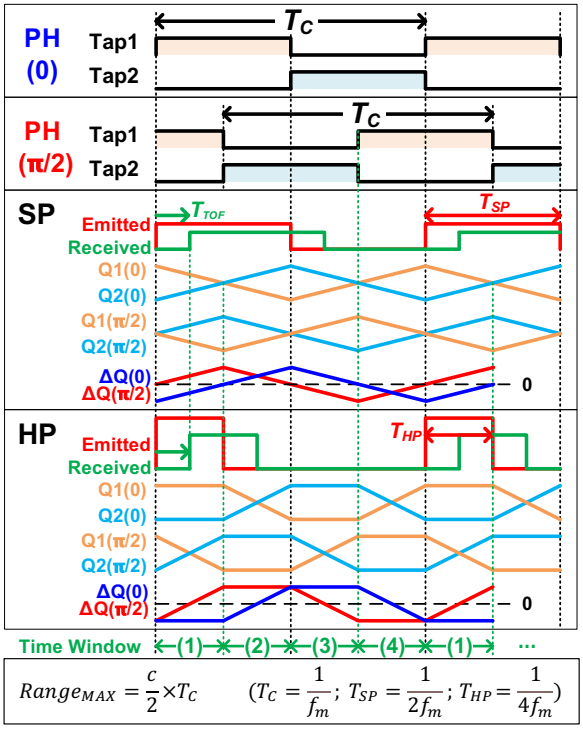


Fig. 1 Operation principle of 2T-4PH ranging methods using SP and HP modulation.

$$Dist.(d) = \frac{c}{2} \times T_{ToF}; \quad T_{ToF} = T_C \times \frac{R}{4} = \frac{R}{4f_m} \quad (0 \le R < 4)$$

Square-Pulse Modulation – SP

$$R_{SP} = \begin{cases} 1 + \dfrac{\Delta Q_0}{|\Delta Q_0| + |\Delta Q_{\pi/2}|}, if\ \Delta Q_{\pi/2} \ge 0 \\ 3 + \dfrac{-\Delta Q_0}{|\Delta Q_0| + |\Delta Q_{\pi/2}|}, if\ \Delta Q_{\pi/2} < 0 \end{cases}$$

Half-Pulse Modulation – HP1

$$R_{HP1} = \begin{cases} 0 + \dfrac{Q2_{\pi/2} - Q2_0}{Q1_{\pi/2} + Q2_{\pi/2} - 2 \cdot Q2_0} & @\ TW(1) \\ 1 + \dfrac{Q2_0 - Q1_{\pi/2}}{Q1_0 + Q2_0 - 2 \cdot Q1_{\pi/2}} & @\ TW(2) \\ 2 + \dfrac{Q1_{\pi/2} - Q1_0}{Q1_{\pi/2} + Q2_{\pi/2} - 2 \cdot Q1_0} & @\ TW(3) \\ 3 + \dfrac{Q1_0 - Q2_{\pi/2}}{Q1_0 + Q2_0 - 2 \cdot Q2_{\pi/2}} & @\ TW(4) \end{cases}$$

Half-Pulse Modulation – HP2

$$R_{HP2} = \begin{cases} \dfrac{1}{2} \cdot \left(1 - \dfrac{\Delta Q_{\pi/2}}{\Delta Q_0}\right) & @TW(1) \\ \dfrac{1}{2} \cdot \left(3 + \dfrac{\Delta Q_0}{\Delta Q_{\pi/2}}\right) & @TW(2) \\ \dfrac{1}{2} \cdot \left(5 - \dfrac{\Delta Q_{\pi/2}}{\Delta Q_0}\right) & @TW(3) \\ \dfrac{1}{2} \cdot \left(7 + \dfrac{\Delta Q_0}{\Delta Q_{\pi/2}}\right) & @TW(4) \end{cases}$$

Fig. 2 Time-shift ratio equations for depth calculation.

$$TW = \begin{cases} (1), & if\ SQ < 0, DQ < 0 \\ (2), & if\ SQ > 0, DQ < 0 \\ (3), & if\ SQ > 0, DQ > 0 \\ (4), & if\ SQ < 0, DQ > 0 \end{cases}$$

Fig. 3 Time-window logic for HP ranging methods.

The time-shift ratio (R) obtained by SP, HP1 and HP2 ranging methods are denoted by R_{SP}, R_{HP2} and R_{HP1}, respectively. Note that SP and HP2 use the differential demodulated signals from $PH(0, \pi/2)$, which are declared as $\Delta Q_0 (= Q2_0 - Q1_0)$ and $\Delta Q_{\pi/2} (= Q2_{\pi/2} - Q1_{\pi/2})$, whereas the HP1 uses three individual signals to perform the BGL cancelling (BGLC) scheme for depth calculation.

For the HP modulation, the logic shown in Fig. 3 is used to determine the TW of the ranging result, where the summation of phase $SQ (= \Delta Q_0 + \Delta Q_{\pi/2})$ and difference of phase $DQ (= \Delta Q_0 - \Delta Q_{\pi/2})$ are created.

By applying the propagation of errors [8], the theoretical depth noise (σ_d) equations were derived and expressed in Fig. 4, where DC is the demodulation contrast and $R_S = N_{ToF}/N_S$. The number of total electrons integrated in a unit pixel during a single frame from the emitted light and BGL are denoted by N_S and N_{BT}. N_{ToF} is the number of demodulated electrons from N_S in the single tap, which will increase along with T_{ToF} in each time-window, where $0 \le R_S < 1$. The readout noise referred to pixel floating diffusion (FD) is denoted by RN.

$$Depth\ Noise\ (\sigma_d) = \frac{c}{8f_m} \times \frac{\sigma_R}{DC}$$

$$\sigma_{R\_SP} = \frac{\sqrt{(N_S + N_{BT} + 2RN^2)(1 - 4R_S + 8R_S^2)}}{N_S}$$

$$\sigma_{R\_HP1} = \frac{\sqrt{N_\varphi(1 - R_S) + (N_{BT} + 2RN^2)(1 - 3R_S + 3R_S^2)}}{N_S}$$

$$\sigma_{R\_HP2} = \frac{\sqrt{\left(\frac{N_S}{2} + \frac{N_{BT}}{2} + RN^2\right)(1 - 2R_S + 2R_S^2)}}{N_S}$$

Fig. 4 Theoretical depth noise equations.

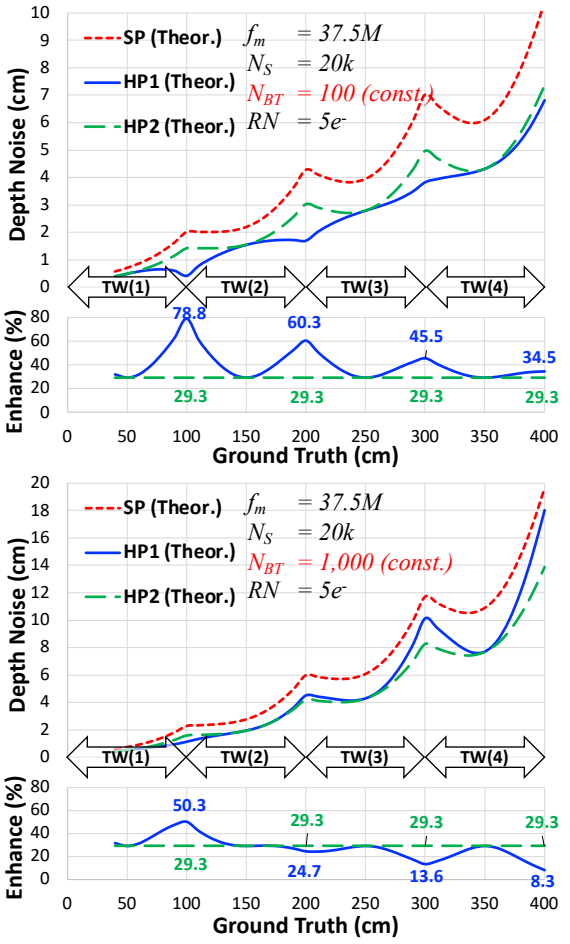


Fig. 5 Theoretical depth noise of the ranging methods.

Fig. 5 compares the theoretical depth noise curves of SP, HP1 and HP2 ranging method. Note that N_S decreased rapidly as a function of distance squared, whereas N_{BT} was a constant for all range. The depth noise enhancement was obtained by the ratio of HP to SP and depicted in percentage.

With a weak influence from BGL, HP1 method can improve the depth precision efficiently, especially when R_S is close to 0 or 1 at each TW. On the contrary, HP2 method guarantees a consistent noise reduction capability for all conditions. In addition, the reported Δ-INT BGLC scheme can also be implemented using HP2 against a stronger ambient light environment [9].

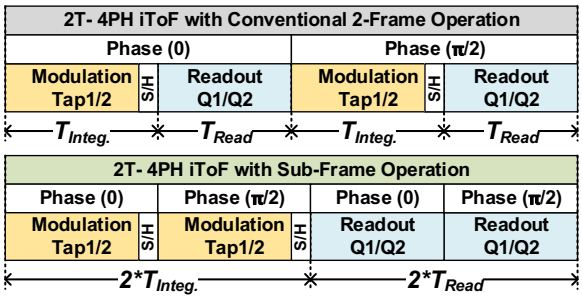


Fig. 5 Concept of sub-frame 2T-4PH operation.

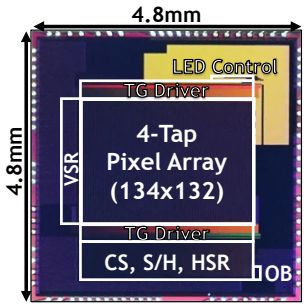


Fig. 6 Chip micrograph [10].

To perform the BGLC algorithms on a 2T-4PH iToF sensor, two modulation periods followed by the array readout are required. As drawn in Fig. 5, the conventional 2-frame operation will encounter with a considerable information loss between $PH(0,\pi/2)$, particularly for a high pixel count sensor. In contrast, by applying the sub-frame operation, a reduction of motion artifact can be expected owing to the compact modulation period. Subsequently, the array readout will be carried out after the modulations.

PIXEL & TIMING DIAGRAM

A reported 4-tap iToF imager [10], shown in Fig. 6, was used to evaluate the proposed techniques. By applying the adjusted timing drawn in Fig. 7, this sensor can be operated as a 2T-4PH iToF imager. The equivalent pixel circuit diagram is shown in Fig. 8, which consists of a high-speed charge collection photodiode (PD), demodulation gates (TGs), buried channel source follower (PSF), current source (PCS), cascode switch (CSC), auto-zeroing capacitor (C_{AZ}) and 2×8 1-T 1-C analog memory (C_{MEM}) with control devices (MW, MRST, MTs, MSs), which share a column readout buffer (MSF).

Fig. 7 depicts the detailed timing diagram. The full integration period is constructed by 2×4 subframes, where $PH(0,\pi/2)$ modulations are performed and sampled into the memory array alternatively. The auto-zeroing sampling is adopted to eliminate the reset thermal noise. Before the column sampling, the charge-domain binning is applied to each phase by mixing the signal charges in the corresponded 4-phase memories. Owing to the averaging effect, the photon shot noise, noise from pixel transistors and kTC noise of the memory can be reduced. Finally, the signals in $PH(0,\pi/2)$ are readout in sequence.

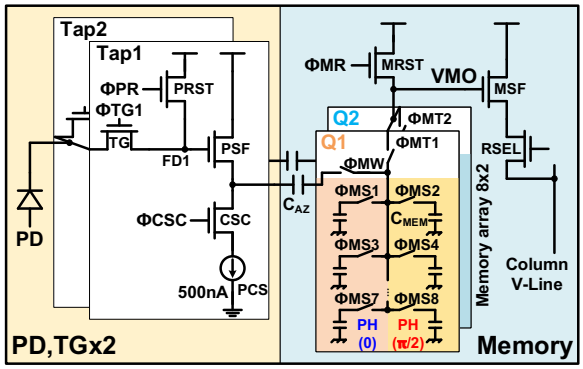


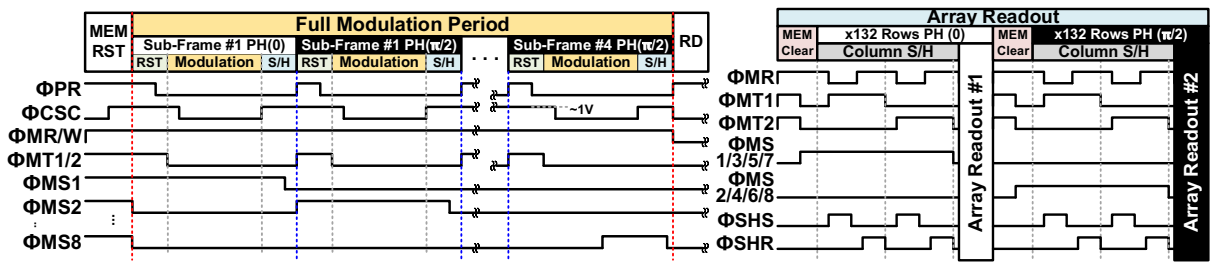
Fig. 8 The equivalent 2T-4PH iToF pixel structure with in-pixel memory array.

EXPERIMENTAL RESULTS

The basic sensor characteristics were measured beforehand. A single-tap FWC around $12ke^-$, and a readout noise floor of $5.4e^-$ with 4-subframe averaging were confirmed. The DC of 85% with 20 ns demodulation pulse width was obtained, where the modulation light was generated by an 850 nm VCSEL The following depth performances were characterized by the center 10×10 pixels over 100 consecutive frames with a F/1.4 lens and IR bandpass filter.

This sensor was operated at a modulation frequency of 25 MHz, which corresponds to an unambiguous range of 6 m for 4-TW referred to Fig. 1. However, due to the circuit timing constraint, the measurable range was confined to 1.5 m with TW(1) using the conventional continuous SP modulation, and to 3.0 m with TW(1)&(2) using the proposed HP modulation.

As the experimental data shown in Fig. 9, the depth precision was enhanced by 44% and 31% at 1.5 m using the HP1 and HP2 ranging method, respectively.



(a) Sub-frame 2T-4PH integration (b) Array readout with averaging

Fig. 7 Circuit timing diagram of sub-frame 2T-4PH operation.

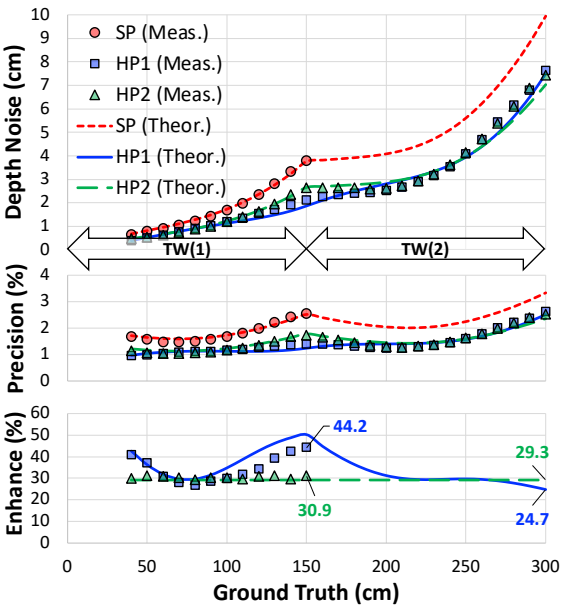


Fig. 9 Experimental and theoretical depth noise curves.

The experimental results show a good agreement with the theoretical calculations in Fig. 4, which can be adopted to estimate the depth noise for the 2-tap iToF ranging methods. As can be seen, HP1 and HP2 methods can provide >25% and >29% precision enhancement, respectively, for the range of 0.4-3 m.

Besides, by giving a proper sensor design with a f_m of 37.5 MHz, N_S of 20,000, DC of 90% and RN of $< 5e^-$, which are achievable values in a state of art iToF system design [11], a high precision of <1% error for a range of 0.4-4m is to be available using HP1 method.

Fig. 10 demonstrated the sample images captured under 30 fps. The image distortion on the hand movement was observed while using the conventional 2-frame operation with SP modulation. In contrast, higher quality depth images with lower noise and suppressed motion artifact were reconstructed using the proposed HP methods with sub-frame operation.

CONCLUSION

In this paper, a 2-tap 4-phase iToF ranging method using HP modulation with sub-frame operation were presented. Both HP1 and HP2 methods can provide a better ranging precision and be utilized in different ways. For the indoor applications, HP1 method is recommended that can achieve the highest depth precision. On the other hand, HP2 has an advantage in a lower SNR ranging system, such as strong ambient light environment and long-distance ranging.

The proposed HP methods can be adopted in exhibited 2-tap iToF sensors to improve their ranging performance without increasing the modulation period or frequency. This indicates a higher framerate and lower power consumption can be expected. Also, to ensure the depth imaging quality with moving targets, the motion artifact suppression was demonstrated by applying the sub-frame 4-phase sampling.

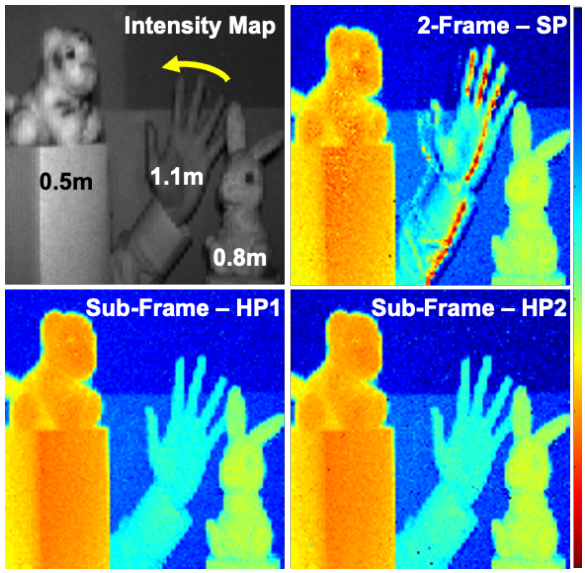


Fig. 10 Captured sample images with hand movement.

The developed technologies show a promising potential to enhance the performance and reliability for various 3-D imaging applications.

REFERENCES

[1] C. S. Bamji *et al.*, "IMpixel 65nm BSI 320MHz demodulated TOF Image sensor with 3µm global shutter pixels and analog binning," in *2018 IEEE International Solid-State Circuits Conference-(ISSCC)*, 2018, pp. 94–96.

[2] Y. Ebiko *et al.*, "Low power consumption and high resolution 1280X960 Gate Assisted Photonic Demodulator pixel for indirect Time of flight," in *2020 IEEE International Electron Devices Meeting (IEDM)*, 2020, pp. 31–33.

[3] Y. Kato *et al.*, "320×240 Back-Illuminated 10-µm CAPD Pixels for High-Speed Modulation Time-of-Flight CMOS Image Sensor," *IEEE J Solid-State Circuits*, vol. 53, no. 4, pp. 1071–1078, Apr. 2018.

[4] C. S. Bamji *et al.*, "A 0.13 µm CMOS System-on-Chip for a 512 × 424 Time-of-Flight Image Sensor with Multi-Frequency Photo-Demodulation up to 130 MHz and 2 GS/s ADC," *IEEE J Solid-State Circuits*, vol. 50, no. 1, pp. 303–319, Nov. 2015.

[5] C. Bamji *et al.*, "A Review of Indirect Time-of-Flight Technologies," *IEEE Trans Electron Devices*, vol. 69, no. 6, pp. 2779–2793, Jun. 2022.

[6] D. Stoppa *et al.*, "A range image sensor based on 10-µm lock-in pixels in 0.18-µm CMOS imaging technology," in *IEEE Journal of Solid-State Circuits*, Jan. 2011, vol. 46, no. 1, pp. 248–258.

[7] J. Cho *et al.*, "A 3-D camera with adaptable background light suppression using pixel-binning and super-resolution," *IEEE J Solid-State Circuits*, vol. 49, no. 10, pp. 2319–2332, 2014.

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

[9] D. Kim *et al.*, "Indirect time-of-flight CMOS image sensor with on-chip background light cancelling and pseudo-four-tap/two-tap hybrid imaging for motion artifact suppression," *IEEE J Solid-State Circuits*, vol. 55, no. 11, pp. 2849–2865, 2020.

[10] C.-C. Kuo and R. Kuroda, "A 4-Tap CMOS Time-of-Flight Image Sensor with In-pixel Analog Memory Array Achieving 10Kfps High-Speed Range Imaging and Depth Precision Enhancement," in *2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*, 2022, pp. 48–49.

[11] Y. Kwon *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 *2020 IEEE International Electron Devices Meeting (IEDM)*, 2020, pp. 32–33.