# Supplementary Technical Report:
# Stochastic Exposure Coding for Handling Multi-ToF-Camera Interference

Jongho Lee and Mohit Gupta
University of Wisconsin-Madison
{jongho, mohitg}@cs.wisc.edu

## 1. Overview

This document provides derivations, explanations, and more results supporting the content of the paper submission titled, "Stochastic Exposure Coding for Handling Multi-ToF-Camera Interference".

## 2. Depth Standard Deviation with Sinusoid Coding Scheme

All approaches are compared in terms of depth standard deviations since only random errors are dominant source of depth errors after systematic errors are removed. We will derive the base depth standard deviation when the sinusoid coding scheme is used in a single C-ToF camera case as a first step. Next, the depth standard deviations for the AC-orthogonal (ACO) approach and the proposed approaches will be derived from it. If the sinusoid coding scheme is used, and the depth value is recovered by the 4-bucket method [4] ($K = 4$ intensity values are used in depth estimation: Eq. 6) in a single camera case, the depth standard deviation is:

$$\boxed{\sigma = \frac{c}{2\sqrt{2}\pi f_0 \sqrt{T}} \frac{\sqrt{e_s + e_a}}{e_s}}, \tag{1}$$

where $f_0$ is a modulation frequency, $T$ is the total integration time, $e_s$, and $e_a$ are the average number of signal photons (due to the primary camera's own source), and ambient photons (due to ambient source), respectively, incident on the pixel per unit time. The derivation of Eq. 1 is as follows.

For the sinusoid coding scheme, modulation function $M(t)$ and demodulation function $D(t)$ are defined as sinusoids:

$$M(t) = D(t) = 1 + \cos(2\pi f_0 t), \tag{2}$$

where $f_0$ is a modulation frequency. The radiance of the reflected light incident on a sensor pixel $\mathbf{p}$ is:

$$R(\mathbf{p}; t) = \alpha P_s M\left(t - \frac{2d}{c}\right) = \alpha P_s \left(1 + \cos\left(2\pi f_0 t - \frac{4\pi f_0 d}{c}\right)\right), \tag{3}$$

where $d$ is the distance between the camera and the scene point imaged at $\mathbf{p}$, $c$ is the speed of light. $P_s$ is average power of the light source with an assumption of $\frac{1}{T_0}\int_{T_0} M(t)\,dt = 1$. $\alpha$ is a scene-dependent scale factor that contains scene albedo, reflectance properties and light fall-off. The correlation or intensity value $C(\mathbf{p}; d)$ measured at pixel $\mathbf{p}$ is:

$$C(\mathbf{p}; d) = s\int_T \left(R(t; d) + P_a\right) D(t)\,dt = sT\left(\alpha P_s + P_a + \frac{\alpha P_s}{2}\cos\left(\frac{4\pi f_0 d}{c}\right)\right), \tag{4}$$

where $s$ is a camera-dependent scale factor encapsulating sensor gain and sensitivity, $T$ is the total integration time, and $P_a$ is average power of ambient light incident on the scene. We take $K = 4$ intensity measurements $C_k(d)$, $k \in \{1, \ldots, 4\}$ by phase-shifting the demodulation function $D(t)$ by a different amount $\psi_k = \frac{\pi}{2}(k-1)$, $k \in \{1, \ldots, 4\}$:

$$C_k(d) = T\left(e_s + e_a + \frac{e_s}{2}\cos\left(\frac{4\pi f_0 d}{c} + \frac{\pi}{2}(k-1)\right)\right), \; k \in \{1, \ldots, 4\}, \tag{5}$$

where $e_s = s\alpha P_s$, and $e_a = sP_a$ are the average number of signal photons, and ambient photons, respectively, incident on the pixel per unit time. We drop the argument $\mathbf{p}$ for brevity. The depth value $d$ can be recovered using the 4-bucket method [4]:

$$d = \frac{c}{4\pi f_0} \tan^{-1}\left(\frac{C_4 - C_2}{C_1 - C_3}\right). \tag{6}$$

Using the error propagation rule, the depth standard deviation $\sigma$ can be obtained by:

$$\sigma = \sqrt{\sum_{k=1}^{4} \left(\frac{\partial d}{\partial C^k}\right)^2 \mathrm{Var}\left(C^k\right)}, \tag{7}$$

where $\mathrm{Var}\left(\cdot\right)$ is a variance operator. Since $\mathrm{Var}\left(C^k\right) = C^k$ in Poisson distribution,

$$\sigma = \frac{c}{2\sqrt{2}\pi f_0 \sqrt{T}} \frac{\sqrt{e_s + e_a}}{e_s}. \tag{8}$$

## 3. Depth Standard Deviation with Bipolar Demodulation

We assumed the demodulation function $D(t)$ is unipolar ($0 \leq D(t) \leq 2$) in the previous derivation. The demodulation functions can be also electronically implemented as bipolar ($-1 \leq D(t) \leq 1$). With zero-mean bipolar demodulation functions ($\int_{T_0} D(t)dt = 0$), $P_a$ and the DC component of $R(t)$ can be cancelled out during integration in correlation computation (Eq. 4). However, *shot noise* by $P_a$ and the DC component of $R(t)$ contributes to random depth errors. If we use a bipolar sinusoid demodulation function instead of a unipolar one, Eq 1 is replaced with:

$$\boxed{\sigma = \frac{c}{2\pi\sqrt{\pi} f_0 \sqrt{T}} \frac{\sqrt{e_s + e_a}}{e_s}}. \tag{9}$$

Compared to Eq. 1, Eq. 9 is scaled down by $\sqrt{2/\pi}$ and everything else is the same. Please note that this is specific for the sinusoid coding scheme. If other coding schemes are used, the equations can be different. The derivation of Eq. 9 is as follows.

Let's assume that $D(t)$ is zero-mean bipolar sinusoid:

$$D(t) = \cos(2\pi f_0 t). \tag{10}$$

The correlation value $C(\mathbf{p}; d)$ measured at pixel $\mathbf{p}$ can be represented as:

$$C(\mathbf{p}; d) = s \int_T \left(R(t; d) + P_a\right) D(t)\, dt = s \int_{T_\oplus} \left(R(t; d) + P_a\right) D(t)\, dt - s \int_{T_\ominus} \left(R(t; d) + P_a\right) \left(-D(t)\right)\, dt, \tag{11}$$

where $T_\oplus$ and $T_\ominus$ mean the intervals of total integration time corresponding to the positive and the negative lobes of $D(t)$. When we take $K = 4$ intensity measurements $C_k(d)$, $k \in \{1, \ldots, 4\}$, we shift the phase of the modulation function $M(t)$ (instead of shifting the demodulation function $D(t)$) by $\psi_k = \frac{\pi}{2}\left(k - 1\right),\ k \in \{1, \ldots, 4\}$ for ease of computation:

$$C_k(d) = C_{k\oplus} - C_{k\ominus} = T\left(\frac{e_s + e_a}{\pi} + \frac{e_s}{4}\cos\left(\frac{4\pi f_0 d}{c} + \psi_k\right)\right) - T\left(\frac{e_s + e_a}{\pi} - \frac{e_s}{4}\cos\left(\frac{4\pi f_0 d}{c} + \psi_k\right)\right), \tag{12}$$

where $C_{k\oplus}$ and $C_{k\ominus}$ are the correlation values for $T_\oplus$ and $T_\ominus$, respectively. The depth value $d$ can be recovered by:

$$d = \frac{c}{4\pi f_0} \tan^{-1}\left(\frac{C_{4\oplus} - C_{4\ominus} - C_{2\oplus} + C_{2\ominus}}{C_{1\oplus} - C_{1\ominus} - C_{3\oplus} + C_{3\ominus}}\right). \tag{13}$$

Using the error propagation rule, the depth standard deviation $\sigma$ can be obtained by:

$$\sigma = \sqrt{\sum_{k=1}^{4} \left(\left(\frac{\partial d}{\partial C_{k\oplus}}\right)^2 \mathrm{Var}\left(C_{k\oplus}\right) + \left(\frac{\partial d}{\partial C_{k\ominus}}\right)^2 \mathrm{Var}\left(C_{k\ominus}\right)\right)}. \tag{14}$$

With $\mathrm{Var}\left(C_{k\oplus}\right) = C_{k\oplus}$ and $\mathrm{Var}\left(C_{k\ominus}\right) = C_{k\ominus}$,

$$\sigma = \frac{c}{2\pi\sqrt{\pi} f_0 \sqrt{T}} \frac{\sqrt{e_s + e_a}}{e_s}. \tag{15}$$

## 4. Depth Standard Deviation of AC-orthogonal (ACO) approach

For an ideal ACO approach, all AC components from interfering sources are removed and only DC components are captured at the sensor of the primary camera. Sum of interfering DC components from all interfering sources acts as additional ambient light, thus can be added to $e_a$ in Eq. 1 to derived the depth standard deviation for ACO:

$$\underbrace{\sigma_{ACO} = \frac{c}{2\sqrt{2}\pi f_0 \sqrt{T}} \frac{\sqrt{e_s + e_a + Ne_i}}{e_s},}_{\text{Eq. 4 of the main manuscript}} \tag{16}$$

where $N$ is the number of interfering cameras, and $e_i = s\alpha_i P_s$ is the average number of interfering photons (due to an interfering source) incident on the pixel per unit time. Without loss of generality, we assume that $e_i$ is the same for all interfering cameras.

## 5. Depth Standard Deviation of Stochastic Exposure Coding (SEC) approach

For the proposed stochastic exposure coding (SEC) approach, the effective integration time is determined by the probability $p_{noclsh}$ that a given slot does not produce a clash. Thus, total integration time is reduced by $p_{noclsh}$ on average. Furthermore, the source strength should be amplified by the source peak power amplification $A$. Interfering DC components are removed since clashed slots are thrown away in SEC. The depth standard deviation of SEC can be derived by putting together all of these:

$$\underbrace{\sigma_{SEC} = \frac{c}{2\sqrt{2}\pi f_0 \sqrt{T p_{noclsh}}} \frac{\sqrt{Ae_s + e_a}}{e_s},}_{\text{Eq. 6 of the main manuscript}} \tag{17}$$

where $A = \min\left(1/p, A_0\right)$, $A_0$ is allowable source peak power amplification, and $p_{noclsh} = p\left(1-p\right)^{2N}$.

## 6. Optimal Slot ON Probability of SEC

The performance of SEC is determined by the slot on probability $p$. If $p$ is too high or low, the effective integration time is reduced. The optimal slot ON probability of SEC $p_{SEC}$ is defined as $p$ minimizing $\sigma_{SEC}$ and can be represented as:

$$\underbrace{p_{SEC} = \min\left(\frac{1}{2N+1}, \frac{1}{A_0}\right).}_{\text{Eq. 7 of the main manuscript}} \tag{18}$$

The derivation of Eq. 18 is as follows. From the definition of $p_{SEC}$,

$$p_{SEC} = \arg\min_p \sigma_{SEC} = \arg\min_p \frac{c}{2\sqrt{2}\pi f_0 \sqrt{Tp\left(1-p\right)^{2N}}} \frac{\sqrt{Ae_s + e_a}}{e_s}, \tag{19}$$

where $A = \min\left(1/p, A_0\right)$. If $1/p \leq A_0$, $A = 1/p$, and

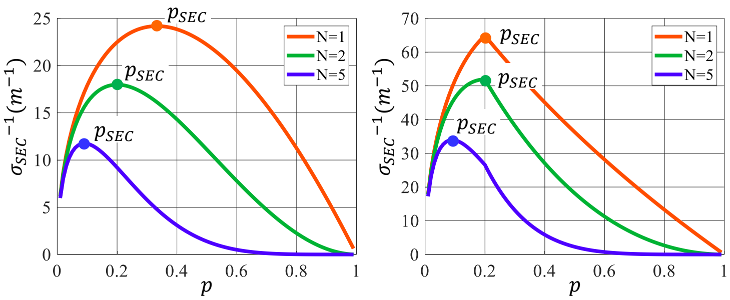
$$p_{SEC} = \arg\min_p \frac{\sqrt{\frac{e_s}{p} + e_a}}{\sqrt{p\left(1-p\right)^{2N}}} = \frac{1}{A_0} \tag{20}$$

since $\sigma_{SEC}$ is monotonically increasing over $p \in [1/A_0, 1]$. Otherwise, $A = A_0$, thus

$$p_{SEC} = \arg\min_p \frac{1}{\sqrt{p\left(1-p\right)^{2N}}} = \frac{1}{2N+1}. \tag{21}$$

From Eq. 20 and Eq. 21,

$$p_{SEC} = \min\left(\frac{1}{2N+1}, \frac{1}{A_0}\right). \tag{22}$$




(a) $\sigma_{SEC}{}^{-1}$ over $p$ when $A_0 = 1$ (b) $\sigma_{SEC}{}^{-1}$ over $p$ when $A_0 = 5$

Figure 1. **Optimal slot ON probability for SEC approach.** (a) The optimal slot ON probability for SEC $p_{SEC}$ is determined only by the number of interfering cameras $N$ when source peak power amplification is not allowed. (b) With peak power amplification, however, both allowable peak power amplification $A_0$ and $N$ determines $p_{SEC}$. If $A_0 \geq 2N+1$, $p_{SEC}$ is determined by $A_0$, otherwise, by $N$.

Figure 1 (a) and (b) show the inverse depth standard deviations $\sigma_{SEC}{}^{-1}$ over $p$ with different number of interfering cameras $N$ without source peak power amplification ($A_0 = 1$) and with source peak power amplification ($A_0 = 5$), respectively. Without source peak power amplification, the optimal slot ON probability $p_{SEC}$ minimizing $\sigma_{SEC}$ (maximizing $\sigma_{SEC}{}^{-1}$) is determined by $N$. When $p > p_{SEC}$, the slot clash increases, the effective integration time without clash decreases, and $\sigma_{SEC}$ increases. When $p < p_{SEC}$, the slot is more rarely sent, the effective integration time decreases, and $\sigma_{SEC}$ increases. When the number of interfering camera $N$ increases, $p_{SEC}$ decreases, and $\sigma_{SEC}$ at $p_{SEC}$ increases. With source peak power amplification, $p_{SEC}$ has two *different* forms according to the relationship between $N$ and $A_0$ (Eq. 18). If $A_0 \geq 2N+1$, $p_{SEC}$ is determined by $A_0$, otherwise by $N$. $e_s = e_a = 1 \times 10^6$, $T = 10\,\mathrm{ms}$, and $f_0 = 30\,\mathrm{MHz}$ were used to create the plots.

## 7. Depth Estimation Algorithm for SEC

First, each ON slot is tested if it is clashed or not (please refer to the next section for the clash check algorithm). The clashed ON slot is discarded since it does not contain correct depth information. Second, if the slot is free from clash, the *slot depth value* $d_m$ is estimated from the slot correlation (or intensity) values $C_{m,k}$, $m \in \{1,\ldots,M_{noclsh}\}, k \in \{1,\ldots,K\}$, where $M_{noclsh}$ is the number of non-clashed ON slots, and $K$ is the total number of captured intensity (or correlation) values. With an assumption of sinusoid coding scheme, and $K = 4$:

$$d_m = \frac{c}{4\pi f_0} \tan^{-1}\left(\frac{C_{m,4} - C_{m,2}}{C_{m,1} - C_{m,3}}\right),\ m \in \{1,\ldots,M_{noclsh}\}. \tag{23}$$

Next, repeat this procedure for all non-clashed ON slots, and estimate the *frame depth value* $d$ by averaging all $d_m$s:

$$d = \frac{1}{M_{noclsh}} \sum_{m=1}^{M_{noclsh}} d_m,\ m \in \{1,...,M_{noclsh}\}, \tag{24}$$

The depth estimation algorithm for SEC is summarized in Algorithm 1.

## 8. Slot Clash Check Algorithm for SEC

Slot clash check is very important since correct depth values cannot be recovered from the clashed slots due to systematic errors. Slot clash check is performed based on the summation of the slot correlation values. For the $m$-th ON slot ($m \in \{1,\ldots,M_{ON}\}$):

$$o_m = C_{m,1} + C_{m,2} + C_{m,3} + C_{m,4}. \tag{25}$$

$o_m$ is proportional to the total number of electrons generated at the sensor by all incoming light. This value is high if clash happens, and low otherwise. We devise a simple threshold-based approach for clash check. By the central limit theorem, if the number of photons or electrons is large enough, $o_m$ is a random variable following a normal distribution whose standard deviation $\sigma_{o_m} = \sqrt{\mathbb{E}\left[o_m\right]}$, where $\mathbb{E}\left[o_m\right]$ is the mean value of $o_m$. The stochastic upper and lower bounds of the value $o_m$

**Algorithm 1:** Depth estimation for the stochastic exposure coding approach

**Input:** Set of the correlation values of all ON slots within a frame, $\{(C_{m,1}, C_{m,2}, C_{m,3}, C_{m,4})\}$ $(m \in \{1, ..., M_{ON}\})$, where $M_{ON}$ is the total number of ON slots within the frame.
**Output:** Depth value for the frame, $d$.
$d_{sum}$ = 0;
**for** *Each* $m \in \{1, ..., M_{ON}\}$ **do**
    $clashFound$ = checkClash($C_{m,1}, C_{m,2}, C_{m,3}, C_{m,4}$);
    (Algorithm 2)
    **if** $clashFound$ == $FALSE$; **then**
        $d_m$ = estimateDepth($C_{m,1}, C_{m,2}, C_{m,3}, C_{m,4}$); (Eq. 23)
        $d_{sum}$ = $d_{sum}$ + $d_m$;
    **end**
**end**
$d$ = $d_{sum}/M_{noclsh}$;

**Algorithm 2:** Slot clash check

**Input:** Correlation values of the $m$-th ON time slot, $(C_{m,1}, C_{m,2}, C_{m,3}, C_{m,4})$.
**Output:** Boolean variable, $clashFound$ indicating if the slot clash happens or not.
$o_m = C_{m,1} + C_{m,2} + C_{m,3} + C_{m,4}$; (Eq. 25)
**if** $o_m > o_{clsh}$; **then**
    $clashFound$ = $TRUE$;
**else**
    $clashFound$ = $FALSE$;
**end**

can be approximated by $\mathbb{E}[o_m] \pm k\sigma_{o_m}$. We will use $\mathbb{E}[o_m] + k\sigma_{o_m}$ as the threshold value $o_{clsh}$ to determine if clash happens or not. If we define $o_{min}$ as:

$$o_{min} = \min o_m, \; m \in \{1, ..., M_{ON}\}, \tag{26}$$

we can approximate $o_{min}$ as $\mathbb{E}[o_m] - k\sigma_{o_m}$ and the closed form solution for $o_{clsh}$ can be derived in terms of $o_{min}$. Please note that this approximation holds only when $M_{ON}$ is large enough. The clashed slots with very small interference can be falsely classified as non-clashed slots. However, since the falsely classified time slot is usually due to very small interference, the depth error is still acceptable. This simple threshold-based algorithm works fast due to its closed form:

$$o_{clsh} = \overline{o_m} + k\sqrt{\overline{o_m}}, \tag{27}$$

where

$$\overline{o_m} = o_{min} + \frac{k^2}{2} + \sqrt{k^2 o_{min} + \frac{k^4}{4}}. \tag{28}$$

The slot clash check algorithm works well with $k = 2$ and is summarized in Algorithm 2.

## 9. Convergence of Required Source Peak Power Amplification for SEC

The required source peak power amplification $A$ for SEC to perform better than ACO in terms of SNR can be estimated from $\sigma_{SEC} \leq \sigma_{ACO}$:

$$\frac{1}{\sqrt{p_{noclsh}}} \frac{\sqrt{A + r_a}}{A} \leq \sqrt{1 + r_a + N r_i}, \tag{29}$$

where $r_a = e_a/e_s$ and $r_i = e_i/e_s$ are relative ambient light strength and relative interfering light source strength, respectively. The required $A$ increases with $N$, but converges in the end as stated in the following result:

**Result 1.** *If the source peak power amplification of SEC is larger than* $\left(e+\sqrt{e\left(e+2r_ar_i\right)}\right)/r_i$*, the depth standard deviation of SEC is always lower than ACO regardless of the number of interfering cameras. For example, the required* $A \approx 6.3$ *when* $r_a = r_i = 1$.

The proof is as follows. If $N$ is large enough, $p_{SEC} = 1/\left(2N+1\right)$, thus the required $A$ can be represented as:

$$A = \frac{1+\sqrt{1+4p_{noclsh}r_a\left(1+r_a+Nr_i\right)}}{2p_{noclsh}\left(1+r_a+Nr_i\right)}. \tag{30}$$

The convergent value of $A$ can be found from:

$$\lim_{N\to\infty} A = \lim_{N\to\infty} \frac{1+\sqrt{1+4p_{noclsh}r_a\left(1+r_a+Nr_i\right)}}{2p_{noclsh}\left(1+r_a+Nr_i\right)}. \tag{31}$$

Using

$$\begin{aligned}\lim_{N\to\infty} p_{noclsh}\left(1+r_a+Nr_i\right) &= \lim_{N\to\infty} \frac{\left(1+r_a+Nr_i\right)}{2N+1}\left(\frac{2N}{2N+1}\right)^{2N} \\ &= \lim_{N\to\infty} \frac{\left(\frac{1+r_a}{N}+r_i\right)}{2+\frac{1}{N}}\frac{1}{\left(1+\frac{1}{2N}\right)^{2N}} \\ &= \frac{r_i}{2e},\end{aligned} \tag{32}$$

$$\lim_{N\to\infty} A = \frac{e+\sqrt{e\left(e+2r_ar_i\right)}}{r_i}. \tag{33}$$

Thus, if the source peak power is increased by more than this value, SEC always works better than ACO regardless of the number of interfering cameras.

## 10. Required Number of Slots for SEC

For reliable depth estimation, the number of the non-clashed ON slots $M_{noclsh}$ should be non-zero. Correct depth estimation is impossible if there is no non-clashed ON slots in the extreme case. This requirement can be represented as $Mp_{noclsh} = M_{noclsh} \geq \chi$, where $M$ is the total number of slots, and $\chi$ $(\chi \geq 1)$ is the minimum number of non-clashed ON slots for the desired performance. The closed form equation for $M$ to satisfy $Mp_{noclsh} \geq \chi$ with a certain success probability $p_{suc} = p\left(M_{noclsh} \geq \chi\right)$ can be derived.

If we define a Bernoulli random variable $X_m \sim \mathcal{B}\left(1, p_{noclsh}\right)$ for the $m$-th time slot $(m = \{1, ..., M\})$, the number of non-clashed ON slots $M_{noclsh}$ is a random variable represented by a summation of $X_m$:

$$M_{noclsh} = \sum_{m=1}^{M} X_m. \tag{34}$$

$M_{noclsh}$ follows a binomial distribution: $M_{noclsh} \sim \mathcal{B}\left(M, p_{noclsh}\right)$. For sufficiently large $M$, it is well known that the binomial distribution is approximated well by the normal distribution:

$$\mathcal{B}\left(M, p_{noclsh}\right) \approx \mathcal{N}\left(Mp_{noclsh}, Mp_{noclsh}\left(1-p_{noclsh}\right)\right). \tag{35}$$

Given $\chi$ $(\chi \geq 1)$, the success probability $p_{suc} = p\left(M_{noclsh} \geq \chi\right)$ can be approximated by the area under the normal distribution curve from $\chi$ to $\infty$. From the z-score of $\chi$:

$$\frac{\chi - Mp_{noclsh}}{\sqrt{Mp_{noclsh}\left(1-p_{noclsh}\right)}} = z, \tag{36}$$

the total number of time slots $M$ can be derived as:

$$M = \frac{z^2p_a + 2\chi p_{noclsh} + \sqrt{z^4p_a^2 - 4p_{noclsh}^2\chi^2}}{2p_{noclsh}^2}, \tag{37}$$

where $p_a = p_{noclsh}\left(1-p_{noclsh}\right)$. $z$ is the function of the desired $p_{suc}$ and is easy to be found from the the standard normal distribution table. Figure 2 shows the required number of slots $M$ over the number of interfering cameras $N$ at the different allowable source peak power amplifications $A_0$ and different probabilities of getting at least one non-clashed ON slots $p_{suc}$.

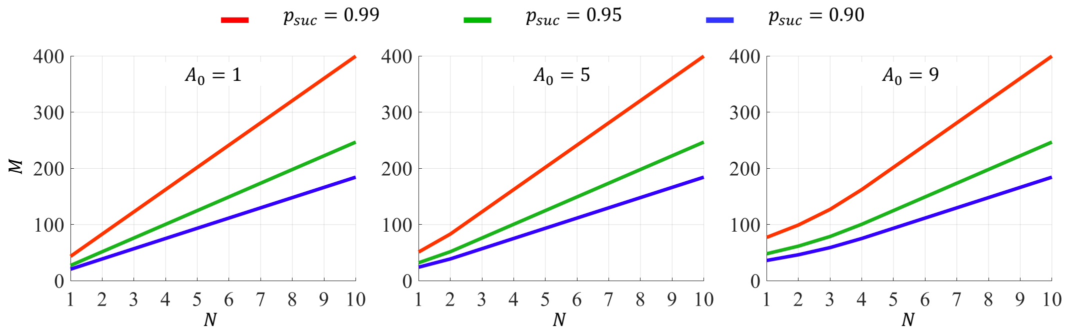



Figure 2. **Required number of slots for SEC.** The required number of slots $M$ over the number of interfering cameras $N$ are shown at the different allowable source peak power amplifications $A_0$ and different probabilities of getting at least one non-clashed ON slots $p_{suc}$.

## 11. Convergence of Required Number of ON Slots for SEC

The number of ON slots $M_{ON} = Mp_{SEC}$ for SEC increases with the number of interfering cameras $N$, but converges in the end as stated in the following result:

**Result 2.** *The required number of ON slots $M_{ON}$ converges to $e\left(z^2/2+1-z\sqrt{z^2/4+1}\right)$ regardless of the number of interfering cameras, where $z$ is the z-score value, and is a function of $p_{suc}$. For example, when $p_{suc} = 0.9$, the required $M_{ON}$ is upper bounded by 9.1.*

This result can be proved as follows:

$$\begin{aligned}\lim_{N\to\infty} M_{ON} &= \lim_{N\to\infty} Mp_{SEC} \\ &= \lim_{N\to\infty} \frac{z^2}{2}\left(\frac{p_{SEC}}{p_{noclsh}}\right) - \frac{z^2}{2}p_{SEC} + \chi\left(\frac{p_{SEC}}{p_{noclsh}}\right) \\ &+ \sqrt{\frac{z^4}{4}\left(\frac{p_{SEC}}{p_{noclsh}}\right)^2 - \frac{z^4}{2}\left(\frac{p_{SEC}^2}{p_{noclsh}}\right) + \frac{z^4}{4}p_{SEC}^2 + \chi z^2\left(\frac{p_{SEC}}{p_{noclsh}}\right)^2 - \chi z^2\left(\frac{p_{SEC}^2}{p_{noclsh}}\right)}.\end{aligned} \tag{38}$$

Using

$$\lim_{N\to\infty} \frac{p_{SEC}}{p_{noclsh}} = \lim_{N\to\infty}\left(1+\frac{1}{2N}\right)^{2N} = e, \tag{39}$$

$$\lim_{N\to\infty} \frac{p_{SEC}^2}{p_{noclsh}} = 0, \tag{40}$$

and

$$\lim_{N\to\infty} p_{SEC} = 0, \tag{41}$$

$$\lim_{N\to\infty} M_{ON} = e\left(\frac{z^2}{2} + \chi - z\sqrt{\frac{z^2}{4}+\chi}\right). \tag{42}$$

Thus, the number of ON slots $M_{ON}$ is upper bounded.

## 12. Frame Rate of SEC

The proposed SEC approach requires dividing a frame into a large number of slots. However, the more pertinent factor that may limit the frame-rate is the number of ON slots, which is typically low. For example, let the total number of slots be 100, and the slot ON probability be 0.2. While the sensor is inactive during OFF slots, each ON slot must have an integration-readout-reset cycle. The reset time, minimum exposure time, and readout time of an off-the-shelf device are 16 µs, 21.3 µs, and 815 µs, respectively [1]. Let the exposure time of each ON slot be 1 ms, and OFF slot time be the same as minimum exposure time. Then, the frame time is $20\times(16\,\mu s + 1000\,\mu s + 815\,\mu s) + 80\times 21.3\,\mu s = 39$ ms, which results in 25 frames/s if 4 measurements are obtained simultaneously using the 4-tap pixel architecture. Although lower than what is achievable with current coding approaches, this may be sufficient for dynamic scenes. For CMB, clash check is not needed and more efficient frame structure is possible.

## 13. Depth Standard Deviation of Multi-Layer Coding (CMB) approach

To derive the depth standard deviation of the Multi-Layer Coding (CMB) approach by generalizing Eq. 8, we need to consider the following things: The total integration time is reduced by slot ON probability $p$. The primary and interfering source strengths are amplified by the source peak power amplification $A$. Average interfering DC component should be added to the ambient strength. It is straightforward to derive the depth standard deviation of SEC by putting together all of these:

$$\underbrace{\sigma_{CMB} = \frac{c}{2\sqrt{2}\pi f_0 \sqrt{Tp}} \frac{\sqrt{Ae_s + e_a + NpAe_i}}{Ae_s}}_{\text{Eq. 9 of the main manuscript}}, \tag{43}$$

where $A = \min(1/p, A_0)$.

## 14. Optimal Slot ON Probability of CMB

The optimal slot ON probability of CMB is defined as:

$$p_{CMB} = \arg\min_p \sigma_{CMB} = \arg\min_p \frac{c}{2\sqrt{2}\pi f_0 \sqrt{Tp}} \frac{\sqrt{Ae_s + e_a + NpAe_i}}{Ae_s}, \tag{44}$$

where $A = \min(1/p, A_0)$. If $1/p \le A_0$, $A = 1/p$, and

$$p_{CMB} = \arg\min_p \sigma_{CMB} = \frac{1}{A_0} \tag{45}$$

since $\sigma_{CMB}$ is monotonically increasing over $p \in [1/A_0, 1]$. Otherwise, $A = A_0$, and

$$p_{CMB} = \arg\min_p \sigma_{CMB} = \frac{1}{A_0} \tag{46}$$

since $\sigma_{CMB}$ is monotonically decreasing over $p \in (0, 1/A_0]$. From Eq. 45 and Eq. 46,

$$\boxed{p_{CMB} = \frac{1}{A_0}}. \tag{47}$$

Therefore, the optimal slot ON probability of CMB $p_{CMB}$ *doesn't* depend on the number of interfering cameras $N$.

## 15. Depth Estimation for CMB

In CMB, slot clash check is not necessary, and the depth value can be estimated from Eq. 6 (if sinusoid coding scheme is assumed) using the summed correlation values from all ON slots:

$$d = \frac{c}{4\pi f_0} \tan^{-1}\left(\frac{\sum_{m=1}^{M_{ON}} C_{m,4} - \sum_{m=1}^{M_{ON}} C_{m,2}}{\sum_{m=1}^{M_{ON}} C_{m,1} - \sum_{m=1}^{M_{ON}} C_{m,3}}\right), \tag{48}$$

where $m \in \{1, \ldots, M_{ON}\}$ is the ON slot index.

## 16. Comparisons with the Same Peak Power

If peak power amplification is 1, and the integration time is kept constant, the optimal ON probability become 1, i.e., $p_{CMB} = 1$. In this case, CMB becomes the same as existing ACO approaches, with the same performance. The more interesting comparison is when the integration time is allowed to be increased. In this case, we can use lower ON probabilities to avoid clashes. Specifically, we set $p_{CMB} = p_{SEC} = 1/(2N + 1)$. To keep the total signal constant, we increase the total integration time by $2N+1$. Figure 3 shows the comparisons between approaches with and without peak power amplification. The performance of the proposed approaches with $A_0 = 1$ is lower than that with $A_0 = 8$, especially for small $N$s. However, the performance gain increases with $N$ due to reduced clash probabilities.

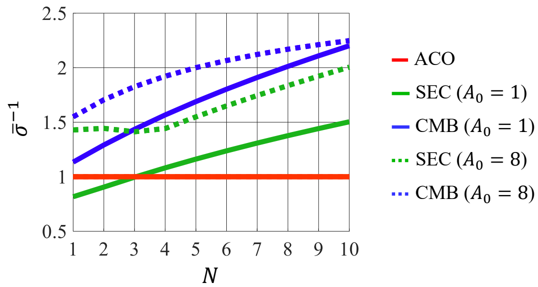



Figure 3. **Comparisons with and without peak power amplifications.** Inverse depth standard deviations are compared between different approaches, with and without peak power amplification.

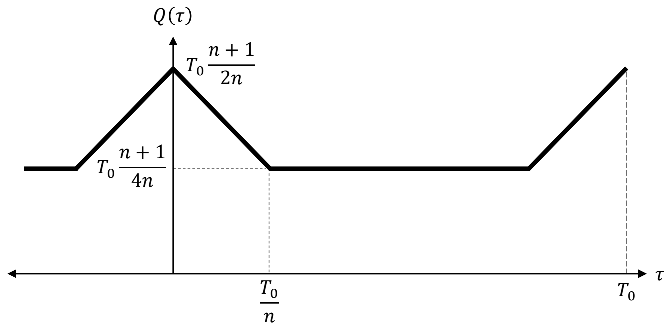



Figure 4. **Autocorrelation of an m-sequence waveform.** The autocorrelaiton of an m-sequence waveform is a periodic triangular function.

## 17. Depth Estimation for PN Sequence (PN) Approach

We used a PN sequence approach (PN) [2, 3] for comparison with our approaches in simulations. We modified the original depth estimation algorithm [2] to accommodate unipolar demodulation functions and four correlation values for fair comparison with other approaches. For the PN approach, modulation function $M(t)$ and demodulation function $D(t)$ are defined as:

$$M(t) = \frac{2n}{n+1}H(t), \tag{49}$$

and

$$D(t) = 2H(t), \tag{50}$$

respectively. $H(t)$ is a unipolar m-sequence (maximum length sequence) ($0 \leq H(t) \leq 1$), $2n/(n+1)$ is a scale factor to make $\frac{1}{T_0}\int_{T_0} M(t)dt = 1$, and $n$ is the number of chips (or bits) during one period of m-sequence waveform. The correlation or intensity value can be represented as:

$$C(\tau) = \frac{4e_s n}{n+1}\int_T H(t-\tau)H(t)dt + e_a T\frac{n+1}{n}, \tag{51}$$

where $\tau = 2d/c$ is the round-trip time of the light from the source to the sensor. The autocorrelation of m-sequence waveform $Q(\tau) = \int_T H(t-\tau)H(t)dt$ has a periodic triangular function (Figure 4). We take four correlation values as follows:

$$C_1 = C(\tau) = 2an_s\left(\frac{T}{T_c}\frac{n+1}{4n}(-\tau) + T\frac{n+1}{2n}\right) + n_a T\frac{n+1}{n}, \tag{52}$$

$$C_2 = C(\tau - T_c) = 2an_s\left(\frac{T}{T_c}\frac{n+1}{4n}(\tau - T_c) + T\frac{n+1}{2n}\right) + n_a T\frac{n+1}{n}, \tag{53}$$

$$C_3 = C(\tau - 2T_c) = 2an_s T\frac{n+1}{4n} + n_a T\frac{n+1}{n}, \tag{54}$$

$$C_4 = C(\tau - 3T_c) = 2an_s T\frac{n+1}{4n} + n_a T\frac{n+1}{n}, \tag{55}$$

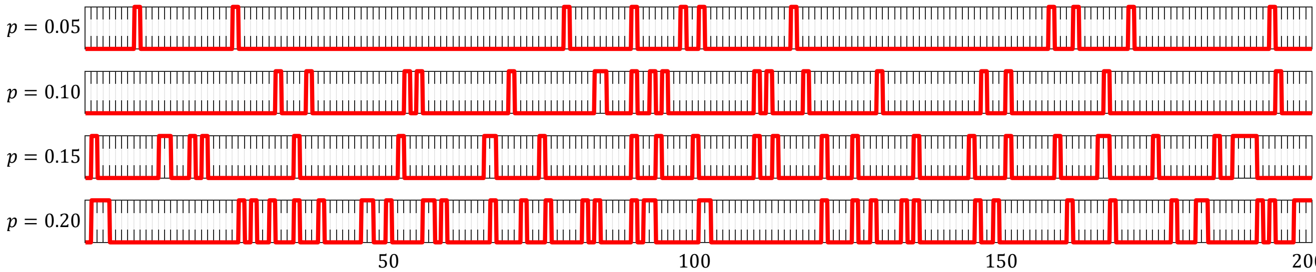


Figure 5. **Examples of random binary sequences.** Examples of random binary sequences at different slot ON probabilities are shown.

where $a = 2n/(n+1)$, and $T_c = T/(n)$. The depth value $d$ can be recovered by:

$$d = \frac{cT_c\left(C_2 - C_4\right)}{2\left(C_1 + C_2 - C_3 - C_4\right)}. \tag{56}$$

## 18. Random Binary Sequences for activation of C-ToF Cameras

Each slot is activated or deactivated by random binary sequence during the integration time in our approaches. The value of the binary sequence for each slot is 1 with an optimal slot ON probability. Figure 5 shows examples of random binary sequences at different slot ON probabilities $p$ when the total number of slots is 200.

## References


[1] OPT9221 Time-of-Flight (ToF) Controller, Texas Instruments. http://www.ti.com/product/OPT9221. 7

[2] Bernhard Büttgen, Felix Lustenberger, Peter Seitz, et al. Pseudonoise optical modulation for real-time 3-d imaging with minimum interference. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 54(10):2109–2119, 2007. 9

[3] Bernhard Büttgen and Peter Seitz. Robust optical time-of-flight range imaging based on smart pixel structures. *IEEE Trans. on Circuits and Systems*, 55(6):1512–1525, 2008. 9

[4] Sergi Foix, Guillem Alenya, and Carme Torras. Lock-in time-of-flight (tof) cameras: A survey. *IEEE Sensors Journal*, 11(9):1917–1926, 2011. 1, 2