Depth Precision in dToF Sensors for AR Applications

Preethi Padmanabhan*, Scott Lindner*, Pierre-Yves Taloud, Nicola Rossi, and David Stoppa
ams OSRAM, Eggstrasse 91, 8803 Rueschlikon, Switzerland
** The first two authors contributed equally to this work.*

***Abstract*— Depth precision of direct time-of-flight (dToF) sensors is crucial in augmented reality applications, where the ToF system should provide a precise depth map under high ambient light and low SNR conditions. In this paper, we derive analytical expressions for depth precision and worst case depth precision considering sub-bin interpolation with a quadratic peak interpolation method. The analytical expression for depth precision is compared with Monte Carlo simulations and measurements from a dToF system. Selection of signal pulse width and time-to-digital converter LSB are identified as a means of optimizing depth precision. Analytical simulations are used to explore this optimization.**

I. Introduction

Recently in augmented reality (AR) applications, ToF sensors have shown to significantly enhance the overall performance by providing high quality 3D depth map of the real world [1,2,3]. The value of the ToF sensor in this application is to provide a highly accurate and precise depth map at ranges of several meters in the presence of high ambient light. With system power consumption a key requirement for mobile applications, ToF sensors are typically required to detect targets at low signal-to-noise ratio (SNR), thus creating a challenging environment for precise measurements.

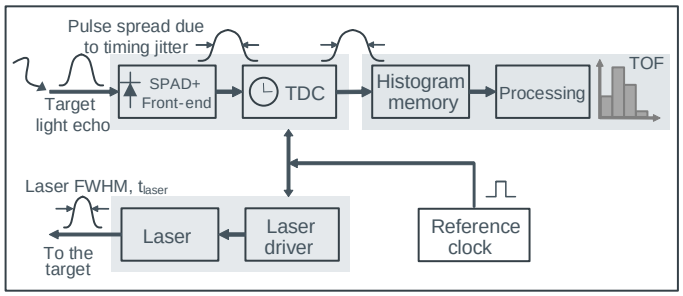

Figure 1. Generic dToF sensor block diagram.

A typical dToF system includes a laser source illuminating the scene of interest, a highly-sensitive photodetector (APD, SPAD) which detects arrival of photons and an accurate timestamping circuit, such as a time-to-digital converter (TDC) (Figure 1). The depth information is determined by measuring the travel time of photons between the emission and detection. Typically, multiple measurements are made over several laser periods, after which a histogram is reconstructed corresponding to the time of arrival of detected photons. In that, a minimum signal-to-noise ratio (SNR) condition needs to be met to detect the target peak with a given false detection rate [4]. Further, the depth resolution from peak detection is often limited to the nearest bin on the histogram, which is likely insufficient at short range. This requires that we use sub-bin interpolation algorithms to obtain better depth resolution. The work in [5] is the most detailed study to date on comparing depth precision of sub-bin interpolation algorithms in dToF as well as their dependence on environmental and system parameters.

A key feature which can impact the depth precision of a sub-bin interpolation algorithm is the phase shift of the signal centroid with respect to the bins of the TDC where events are time-stamped. Since this phase shift results in a given distribution of signal counts between the TDC bins, characterization of the depth precision in the general case requires multiple measurements at finely spaced time-steps over the period of a single bin [5]. However, if we wish to know the worst case (WC) precision of a system, we require measurements with the distance between the system and target corresponding to the WC phase shift for the algorithm.

In this paper, we derive an expression for the WC depth precision, linked to specific phase shifts of the signal, considering sub-bin interpolation using a quadratic peak interpolation method [6]. The dependence of depth precision on the phase shift of the signal within the TDC bin is demonstrated by comparing the analytical 1-sigma precision with Monte Carlo simulations and then experimental data obtained from a dToF system [7]. Finally, optimization of the depth precision through appropriate selection of signal pulse width and TDC bin width (LSB) is investigated with analytical simulations.

II. Analytical Modeling

Sub-bin interpolation based on quadratic peak interpolation [6] was selected due to its relative simplicity. This feature makes it suitable for AR applications, where between hundreds and thousands of depth points must be computed per frame. The expression for the sub-bin centroid of the signal, *P*, is given by

$$P = \frac{1}{2}\frac{\alpha - \gamma}{\alpha - 2\beta + \gamma}, \epsilon\,[-1/2\ \ 1/2] \qquad \textit{Eq. 1}$$

where α, β and γ are the peak bin of the signal and its neighboring bins as illustrated in Figure 3. The signal shape is assumed to be symmetric in our analysis. The 1-sigma depth precision, σ_P, of the sub-bin interpolation can be estimated by applying the error propagation formula to Eq. 1. This results in the following equation:

$$\sigma_P = P\sqrt{\frac{\sigma_\alpha^2+\sigma_\gamma^2}{(\alpha-\gamma)^2}+\frac{4(\sigma_\alpha^2+4\sigma_\beta^2+\sigma_\gamma^2)}{(2\alpha-4\beta+2\gamma)^2}} \qquad \textit{Eq. 2}$$

where σ_α, σ_β and σ_γ are the standard deviation of the noise at α, β and γ bins, respectively. Due to the absence of readout noise in SPAD based systems, these noise contributions are assumed entirely composed of shot noise. Although, we could have additional sources of error, such as, nonlinearity of the TDC, they are assumed to be minimal in comparison to the shot noise.

We now consider two different cases (Figure 3), at which to evaluate Eq. 2.

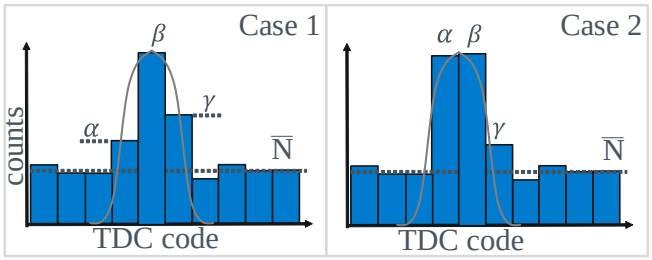

Figure 3. Possible cases for precision analysis.

In case 1, the signal centroid coincides with the center of a histogram bin such that β is maximized, $\alpha \approx \gamma$ and $\beta = \bar{N} + SNR\sqrt{\bar{N}}$ where, for a mean noise count of $\bar{N}$, the SNR is defined as $(S-\bar{N})/\sqrt{\bar{N}}$. In case 2, the signal centroid coincides with the boundary between α and β, $\alpha \approx \beta$ and $\beta = \bar{N} + SNR\sqrt{\bar{N}}$. The time shift between case 1 and case 2 is ½ LSB and they represent the best case (BC) and WC for depth precision, respectively. In order to derive the WC depth precision, we evaluate the above cases under low SNR and high ambient noise. Under these conditions, for case 1, we assume, $\alpha \approx \gamma \approx \bar{N}$ and evaluate the 1-sigma expression in Eq. 2 for case 1. Given that the noise contributions are assumed driven by shot noise, we further assume that $\sigma_\alpha^2 = \alpha, \sigma_\beta^2 = \beta$ and, $\sigma_\gamma^2 = \gamma$. The 1-sigma precision for case 1, σ_{P1}, results in the following equation:

$$\sigma_{P1} = 0.5\sqrt{\frac{2\bar{N}+1}{4(\bar{N}-\beta)^2}+\frac{\bar{N}+2\beta}{8(\bar{N}-\beta)^4}} \qquad \textit{Eq. 3}$$

Please note that we assume $\alpha \approx \gamma + 1$ for the purposes of evaluating Eq. 2 as $\alpha \approx \gamma$ yields P = 0 which has no physical relevance in this analysis. Further, assuming that the second term in Eq. 3 is much smaller compared to the first term, it is sufficient to include the first term as the only dominant contributor to σ_{P1}.

For case 2, we assume $\gamma \approx \bar{N}$ and evaluate Eq. 2. The resulting equation for 1-sigma precision, σ_{P2} is as follows:

$$\sigma_{P2} = 0.5\sqrt{\frac{2\bar{N}+6\beta}{(\bar{N}-\beta)^2}} \qquad \textit{Eq. 4}$$

Comparing the first term in Eq. 3 and Eq. 4 shows us that, indeed, the worst-case precision occurs for case 2, where the signal centroid coincides with the boundary between α and β.

With this background in mind, we further refine our analysis to include additional factors, such as, LSB and laser pulse width (FWHM) which influence the achievable depth precision. The 1-sigma depth precision, σ_P, can be independently analyzed for FWHM and LSB. We define a coefficient, $C1$, as the ratio of α and γ with respect to the peak bin, β. For case 1, we can assume that for a given pulse width, of a symmetric signal, $\alpha = \gamma = C1 \cdot \beta$, where the value of $C1$ depends on the signal shape (hence, FWHM) and TDC LSB. The 1-sigma depth precision, σ_{P1}, can be alternatively expressed as follows:

$$\sigma_{P1} = P\sqrt{\sigma_\alpha^2+\sigma_\gamma^2+\frac{4(\sigma_\alpha^2+4\sigma_\beta^2+\sigma_\gamma^2)}{(4\beta(C1-1))^2}} \qquad \textit{Eq. 5}$$

Similarly, for case 2, the 1-sigma depth precision, σ_{P2}, can be alternatively expressed as follows:

$$\sigma_{P2} = P\sqrt{\frac{2\sigma_\alpha^2+4\sigma_\beta^2+2\sigma_\gamma^2}{(\beta(C1-1))^2}} \qquad \textit{Eq. 6}$$

For Equations 5-6, it is shown that the depth precision can be optimized through the tuning of $C1$. This will be revisited in the results section with examples from simulations.

III. Results

In order to demonstrate the effect of phase shift, an analytical model including the error propagation formula in Eq. 2 is simulated by sweeping the phase of the signal centroid with respect to the TDC bin width. The IRF of the SPAD is extracted from measured data which is then convolved with a laser signal modeled as a trapezoid of a certain FWHM, rise time and fall time. The convolved signal is then sampled at the resolution (LSB) of the TDC. Based on a given target SNR and mean noise, $\bar{N}$, target peak, β, and neighboring bins, α, γ are estimated which are then fed to the analytical expressions in Eq. 1, 2. The simulation flow also includes a Monte Carlo model from which the precision is obtained as a standard deviation of the signal centroid over ~5000 runs. Simulations are performed at varying phase shift, TDC LSB, laser FWHM and, SNR.

Figure 4 shows the 1-sigma depth precision obtained using the analytical expression, σ_P, introduced in Eq. 2 in Section II as well as the Monte Carlo simulation results. Analytical results obtain good matching with a relative error < 8%. Simulation conditions include a laser signal with FWHM at 100ps, a SPAD jitter with FWHM at 120ps and a LSB at 218ps. A target SNR of ~9 with a mean noise count per bin equal to ~950 is chosen to be able to compare the simulation results with the measured data obtained under similar conditions (Figure 5a,b). Figure 4 shows the dependence of precision on the phase shift of the signal within the TDC bin, where, a phase shift of '0ps' represents the best case (BC) where the signal centroid is sampled at the center of the TDC bin (case 1 in Figure 3), and

a phase shift of '~ ±100ps' represent the WC where the signal centroid is at either extremes of the TDC bin, ±LSB/2 (case 2 in Figure 3).

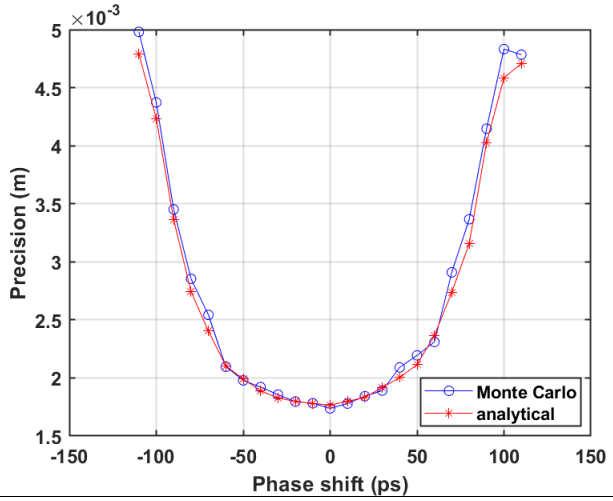

Figure 4. Depth precision (m) vs. phase shift (ps).

The BC precision of 1.78mm is obtained at phase shift = 0ps. The precision starts to degrade as the signal centroid deviates from the BC point for increasing phase shifts and reaches the WC value at phase shift ≈ ±100ps. The difference between BC and WC point is > 3mm under the simulated conditions.

In order to validate the model, experimental data was obtained from the dToF system presented in [7]. Despite having a high-quality timing reference on chip including a PLL, typically, multiple timing uncertainties exist in the sensor which may manifest as offsets or sources of timing jitter. As a result, though very small, there is an inevitable residual mismatch between the arrival times of clock at various macropixels across the sensor array. Although various offsets are typically dealt with during calibration, the phase shift of the signal centroid across macropixels when imaging a given scene is largely unknown. In fact, this is clearly seen in the measured histograms obtained in two different macropixels, where, in Figure 5a, the signal centroid appears to be favorably sampled at the center of the TDC bin (bin 17), while in Figure 5b, sampled at the boundary of two neighboring bins (bins 16, 17). The histograms also show an underlying relationship between the phase shift and the distribution of counts across the bins, α, β, γ as seen in Section II. We define another coefficient, $2 = \alpha/\gamma$. For a phase shift of 0ps, C2 is ~1.06 (Figure 5a) and increases to C2 ~1.21 for histogram in Figure 5b. In other words, as the phase shift sweeps from –LSB/2 to +LSB/2, the coefficient, C2, also changes (increases). As a result of this relationship, the depth precision can also be plotted as a function of C2 = α/γ, thereby, disentangling the absolute phase shift which depends on the bin width. Figure 6 replots the precision obtained in Figure 4 as a function C2 = α/γ. Please note that obtained curve is not symmetric anymore simply due to the difference in representing the data as a function of α/γ which linearly increases through –LSB/2 to +LSB/2. The precision obtained from the experimental data, is overlaid as a scatter plot, where good matching is observed on the overall trend in the precision variation. Furthermore, the precision obtained is well within the upper limit defined by the WC identified in Eq. 4, plotted as a dotted reference line in Figure 6. Please note that the experimental data presented here is uncalibrated for bin width variation across macropixels, which is not captured in the simulated data.

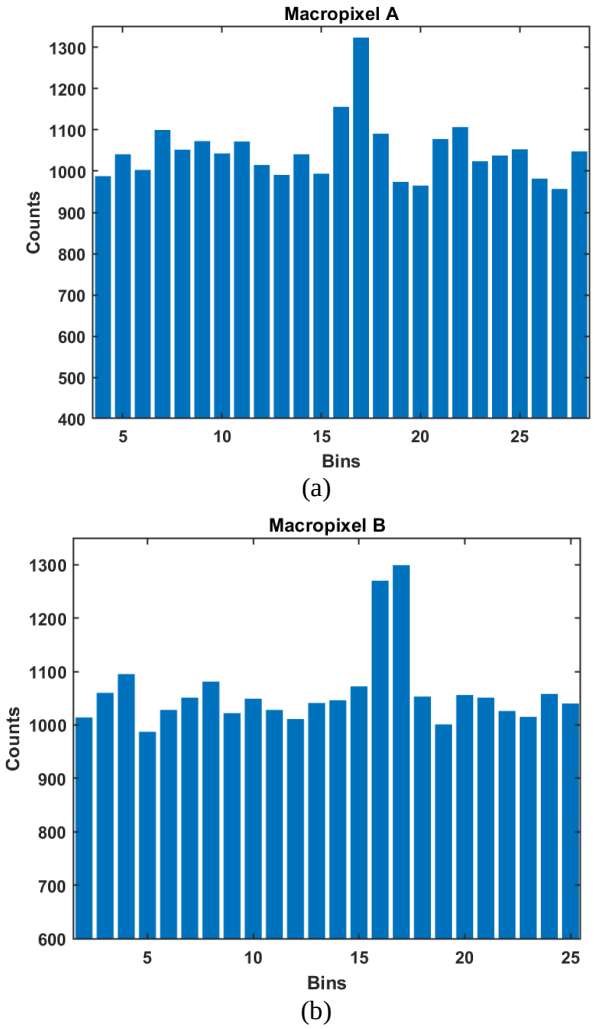

Figure 5. (a) Measured histogram when the center of the detected pulse sampled at the center of the peak bin and (b) when the center of the detected pulse sampled at the boundary of two neighboring bins.

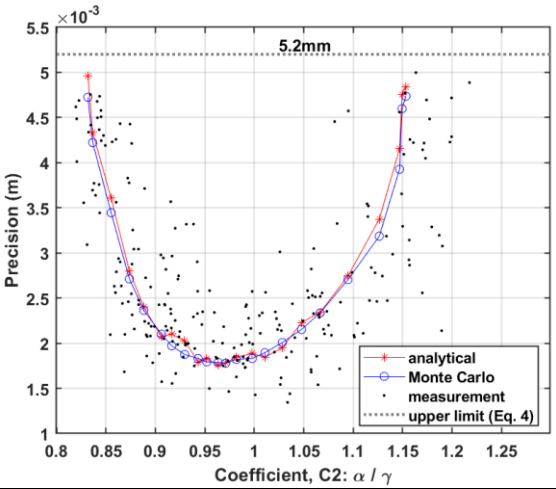

Figure 6. Depth precision (m) vs. coefficient, C2.

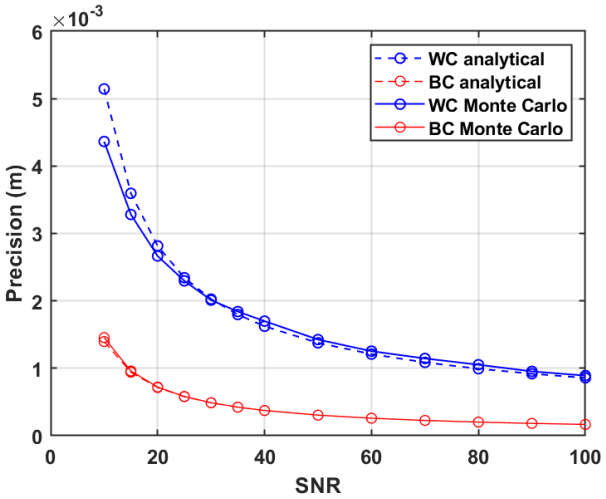

Figure 7. Precision (m) vs. SNR for WC and BC points.

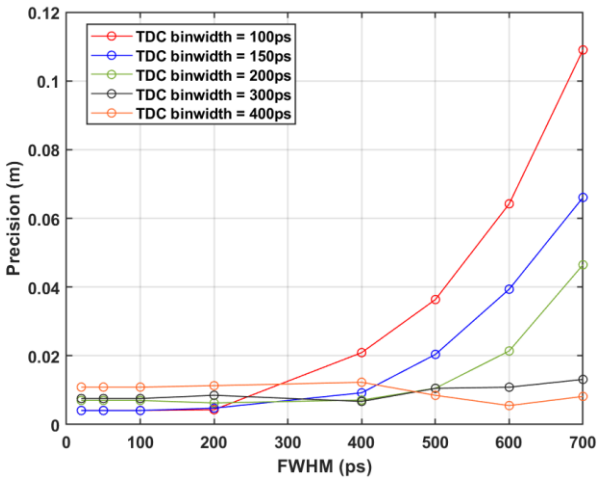

Figure 8. Analytical precision (m) vs. FWHM (ps).

In addition to the phase shift of the signal, SNR and FWHM of the laser pulse are other factors which can be tuned to improve depth precision. The plot shown in Figure 4 is simulated over increasing SNR and the BC and WC points are extracted. The BC and WC precision values are calculated as the minimum and the maximum point respectively, on the precision-phase-shift curve for each of the SNR conditions and the results are shown in Figure 7. The improvement in precision at increasing SNR is evident, although, this plot also emphasizes the need to consider the WC points in the depth precision while imposing a SNR condition. This can especially impact short distances of interest, <500mm, where, a small variation can significantly impact the relative depth precision. For example, at a minimal SNR of 10, the precision can vary anywhere between 1.1% at the WC point to 0.3% at the BC point at a target distance of 500mm.

Figure 8 shows the dependence of precision on the FWHM of the laser pulse simulated at different LSB. For FWHM ≤ 3LSB, the precision inversely scales with bin width. In Figure 8, this trend is seen for FWHM ≤ 200ps; at FWHM = 200ps and LSB = 400ps, where, a majority of the signal energy is contained within 1 bin. This is directly related to a lower coefficient, $C1$, identified in Eq. 6 in Section II, which results in a poorer precision compared to LSB = 100ps, where FWHM/LSB ≈ 3 and the value of $C1$ is higher. Consequently, we get a better precision. This trend soon starts to invert for scenarios where FWHM/LSB ≥ 3LSB. At these points, the signal energy is spread beyond 3 LSBs and the sub-bin precision degrades. In Figure 8, at FWHM = 700ps, the WC precision occurs when LSB = 100ps, where the signal information is likely to be spread across ~7bins, whereas, the quadratic peak interpolation captures information only within peak bin ± 1 bin, thus, discarding most of the signal counts outside this range. Therefore, depending on the FWHM/LSB ratio, appropriate peak interpolation algorithm may be necessary to fully capture the signal information and ensure minimal influence of the processing technique on the precision of a measurement. As such, for the results shown in Figure 6 and in [7], the chosen FWHM ≤ 3LSB to ensure the signal information is contained within 3 bins.

IV. Conclusions

Depth precision specifications for ToF systems are typically stated with a maximum precision. In the case of the quadratic peak interpolation analyzed here, the precision shows a dependence on the phase shift between the signal and the TDC bins. The model developed here demonstrates good matching with Monte Carlo simulations and experimental measurements over the phase shift of a single LSB. Further simulations with varying SNR show both the tradeoff between SNR and depth precision, and the importance of selecting the best or worst case phase shifts depending on what is of interest. These results demonstrate that the analytical model is a useful tool to investigate depth precision over the whole phase shift of the TDC LSB. Furthermore, the upper limit, Eq. 4, presents a simple method for evaluating the worst case precision at a given SNR. Finally, it was shown that should we wish to minimize precision, there is a motivation for selecting both a small laser FWHM and TDC LSB.

References

[1] https://www.apple.com/augmented-reality/
[2] https://docs.microsoft.com/en-us/hololens/hololens2-hardware
[3] What is ToF camera technology on Galaxy and how does it work? - The Official Samsung Galaxy Site
[4] Tolt, Gustav, et al. "Peak detection approaches for time-correlated single-photon counting three-dimensional lidar systems." *Optical Engineering* 57.3 (2018): 031306.
[5] Koerner, Lucas J., "Models of Direct Time-of-Flight Sensor Precision That Enable Optimal Design and Dynamic Configuration." *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-9, 2021.
[6] Smith III, Juliu O. *Spectral audio signal processing*. W3K publishing, 2011.
[7] Stoppa, David, et al. "A Reconfigurable QVGA/Q3VGA Direct Time-of-Flight 3D Imaging System with On-chip Depth-map Computation in 45/40nm 3D-stacked BSI SPAD CMOS", Proc. of IISW'21.