Combination of Spatially-Modulated ToF and Structured Light for MPI-Free Depth Estimation

Gianluca Agresti and Pietro Zanuttigh

Department of Information Engineering, University of Padova, Padova, Italy
{gianluca.agresti, zanuttigh}@dei.unipd.it

Additional Material

In this additional material we present the mathematical derivation of the modified ToF correlation function, of the STM-ToF and SL error models and of the implicit phase unwrapping for the SL depth map employed in the paper.

1 Correlation Function for the Modified ToF Acquisition System

In this section the correlation function for the modified ToF system, i.e., Equation (7P) of the paper, will be derived. Notice that, in order to clarify notation, we will use the letter "P" to distinguish the equation numbers from the paper from the ones in this document, i.e., Equation (7P) is Equation (7) of the main paper.

Referring to the acquisition model of Section 3.2 of the paper, the signal emitted by pixel (x, y) of the projector when the sample $\omega_r\tau_i$ of the correlation function has to be computed is given by the combination of the standard ToF modulation of Equation (1P) with the modulating pattern which intensity is given by Equation (6P):

$$\begin{aligned} s_t(t, \omega_r\tau_i) &= L_{x,y}(\omega_r\tau_i) \cdot s_{stdToF,t}(t) \\ &= \frac{1}{4}a_t\Big(1 + \cos\big(l\omega_r\tau_i - \theta_{x,y}\big)\Big)\Big(1 + \sin(\omega_r t)\Big). \end{aligned} \tag{1}$$

Each pixel of the ToF camera receives a light signal that can be modeled as:

$$\begin{aligned} s_r(t, \omega_r\tau_i) = b_r + \frac{1}{4}a_r\Big(1 + \cos\big(l\omega_r\tau_i - \theta_{x,y}\big)\Big)\Big(1 + \sin(\omega_r t - \phi_d)\Big) + ... \\ ... + b_{r,g} + a_{r,g} \cdot \cos(\omega_r\tau_i - \phi_g) \end{aligned} \tag{2}$$

where b_r is the light offset due to the ambient light, $a_r = \alpha a_t$, with α equal to the channel attenuation and ϕ_d is the phase displacement between the transmitted and received direct part of the signal. The scene depth d can be computed from ϕ_d through the well known relation $d = \frac{\phi_d c_l}{2\omega_r}$ where c_l is the speed of light. The second line of Equation (2) contains the global component of the received light signal, the one related to MPI and it is assumed to be not influenced by the projected pattern in case of diffuse reflections [1].

The ToF pixels are able to compute the correlation function between the received signal and a reference one, e.g., a rectangular wave at the same modulation frequency $rect_{\omega_r}(t) = H\big(\sin(\omega_r t)\big)$, where $H(\cdot)$ represents the Heaviside function. The correlation function sampled in $\omega_r \tau_i \in [0; 2\pi)$ can be modeled in this acquisition scenario as

$$c(\omega_r \tau_i) = \int_0^{\frac{1}{f_{mod}}} s_r(t, \omega_r \tau_i) H\big(\sin(\omega_r t + \omega_r \tau_i)\big) dt \tag{3}$$

and by substituting Equation (2) inside (3), we obtain Equation (7P):

$$\begin{aligned} c(\omega_r \tau_i) = B + A\cos(\omega_r \tau_i + \phi_d) + A_g \cos(\omega_r \tau_i + \phi_g) + \frac{\pi A}{2} \cos(l \omega_r \tau_i - \theta) + ... \\ + \frac{A}{2}\Big[\cos\big((l-1)\omega_r \tau_i - \phi_d - \theta\big) + \cos\big((l+1)\omega_r \tau_i + \phi_d - \theta_{x,y}\big)\Big] \end{aligned} \tag{4}$$

where $B = \frac{1}{f_{mod}}\Big(\frac{b_r}{2} + \frac{a_r}{8} + \frac{b_{r,g}}{2}\Big)$ is an additive constant that represents the received light offset, $A = \frac{a_r}{4\pi f_{mod}}$ is proportional to the power of the direct component of the received light, $A_g = \frac{a_{r,g}}{\pi f_{mod}}$ is proportional to the power of the global component of the received light. The correlation function values $c(\omega_r \tau_i)$ are a measure of the number of photons received by the pixel during the considered integration time.

2 Error Propagation Analysis for the ToF Approach

In Section 3.3 of the paper we have presented a model for the estimation of the impact of the noise on ToF correlation samples on the phase estimation. In this section we present the mathematical derivation based on error propagation analysis that we used to compute Equation (10P) in the paper.

Let us briefly recall that, given a function $f(\cdot)$ that is continuous and has continuous first and second order derivatives with respect to the random arguments $\big\{V_i\big\}_{i=0}^{N-1}$ on some of their neighborhoods, then the random variable $w = f(\big\{V_i\big\}_{i=0}^{N-1})$ is asymptotically *Normal*. If these arguments of $f(\cdot)$ are random variables respectively with mean μ_{V_i} and variance $\sigma^2_{V_i}$, whose deviation from the mean are symmetric and bell-shaped, and assuming that they are each other independent, then the variance of the function $f(\cdot)$ can be approximated as [2]:

$$\sigma_f^2 = \sum_{i=0}^{n} \Big(\frac{\partial f}{\partial V_i}\Big)^2 \sigma^2_{V_i}. \tag{5}$$

In the following of this analysis, each sample of the correlation function described by Equation (5P) will be labeled as $N_i = c(\omega_r \tau_i)$ with $\omega_r \tau_i = \frac{2\pi}{9} i$ for $i = 0, ..., 8$. Due to the random nature of the light that is assumed to be affected by *photon shot noise*, it is possible to assume that the ToF correlation samples have a *Poisson distribution* with mean μ_{V_i} and variance $\sigma^2_{V_i}$ equal to N_i [3], i.e., the number of photons accumulated during the correlation sample acquisition.

Since $\phi_d = \frac{\varphi_4 - \varphi_2}{2}$, as a first step we are going to estimate the variance for φ_k with $k = 1, ..., 8$, that is

$$\sigma^2_{\varphi_k} = \sum_{i=0}^{8} \Big(\frac{\partial \varphi_k}{\partial N_i} \Big)^2 \sigma^2_{N_i}. \tag{6}$$

Since from Fourier analysis

$$\varphi_k = arctan\left(- \frac{\sum_{i=0}^{8} N_i \sin \frac{2\pi}{9} ki}{\sum_{i=0}^{8} N_i \cos \frac{2\pi}{9} ki} \right), \tag{7}$$

by taking $X_k = -\frac{\sum_{i=0}^{8} N_i \sin \frac{2\pi}{9} ki}{\sum_{i=0}^{8} N_i \cos \frac{2\pi}{9} ki}$, we can express it as $\varphi_k = arctan(X_k)$ and so Equation (6) can be reformulated as

$$\begin{aligned} \sigma^2_{\varphi_k} &= \sum_{i=0}^{8} \Big(\frac{\partial arctan(X_k)}{\partial N_i} \Big)^2 \sigma^2_{N_i} \\ &= \Big(\frac{\partial arctan(X_k)}{\partial X_k} \Big)^2 \sum_{i=0}^{8} \Big(\frac{\partial X_k}{\partial N_i} \Big)^2 \sigma^2_{N_i} \\ &= \Big(\frac{1}{1 + X_k^2} \Big)^2 \sum_{i=0}^{8} \Big(\frac{\partial X_k}{\partial N_i} \Big)^2 \sigma^2_{N_i}. \end{aligned} \tag{8}$$

Since $X_k = tan \varphi_k$ and $\frac{1}{1 + tan^2 \varphi_k} = \cos^2 \varphi_k$, we can rewrite (8) as:

$$\sigma^2_{\varphi_k} = \cos^4 \varphi_k \sum_{i=0}^{8} \Big(\frac{\partial X_k}{\partial N_i} \Big)^2 \sigma^2_{N_i}. \tag{9}$$

After evaluating the partial derivatives $\frac{\partial X_k}{\partial N_i}$ and computing the summation in (9), it results that the error variance for the estimation of the phase φ_k is equal to:

$$\sigma^2_{\varphi_k} = \frac{2}{9 I_k^2} \Big[B - \frac{1}{2} I_{2k} \cos(2\varphi_k) \cos(\varphi_{2k}) - \frac{1}{2} I_{2k} \sin(2\varphi_k) \sin(\varphi_{2k}) \Big], \tag{10}$$

where we labeled with I_i the amplitude of the sinusoidal wave at frequency i in the correlation function given by Equation (4). Moreover, the variances of the phase estimation error for phases φ_2, φ_3 and φ_4 are respectively

$$\sigma^2_{\varphi_2} = \frac{8}{9A^2}\Big[B - \frac{A}{4}\cos(-3\phi_d - \theta)\Big], \tag{11}$$

$$\sigma^2_{\varphi_3} = \frac{8}{9\pi^2 A^2}\Big[B - \frac{\pi A}{4}\cos(3\theta)\Big], \tag{12}$$

$$\sigma^2_{\varphi_4} = \frac{8}{9A^2}\Big[B - \frac{D}{2}\cos(2\phi_d - 2\theta + \phi_{FF})\Big]. \tag{13}$$

where

$$D = \sqrt{A + A_g + 2A \cdot A_g \cos(\phi_d - \phi_g)} \tag{14}$$

$$\phi_{FF} = \text{atan2}\big(A\cos\phi_d + A_g\cos\phi_g, A\sin\phi_d + A_g sin\phi_g\big). \tag{15}$$

At this point it is possible to evaluate also the variance of the error for ϕ_d and θ, since $\phi_d = (\varphi_4 - \varphi_2)/2$ and $\theta' = -(\varphi_2 + \varphi_4)/2$. Since ϕ_2 and ϕ_4 are not independent, because they are computed from the same samples of the correlation function, we have

$$\begin{aligned}\sigma^2_{\phi_D} &= \sum_{i=1}^{n}\Big(\frac{\partial\phi_D}{\partial N_i}\Big)^2\sigma^2_{N_i} = \frac{1}{4}\sum_{i=1}^{n}\Big(\frac{\partial[\varphi_4 - \varphi_2]}{\partial N_i}\Big)^2\sigma^2_{N_i}\\ &= \frac{1}{4}\Big[\sum_{i=1}^{n}\Big(\frac{\partial\varphi_4}{\partial N_i}\Big)^2\sigma^2_{N_i} + \sum_{i=1}^{n}\Big(\frac{\partial\varphi_2}{\partial N_i}\Big)^2\sigma^2_{N_i} - 2\sum_{i=1}^{n}\frac{\partial[\varphi_4\cdot\varphi_2]}{\partial N_i}\sigma^2_{N_i}\Big]\\ &= \frac{1}{4}\Big[\sigma^2_{\varphi_4} + \sigma^2_{\varphi_2} - 2\sigma_{\varphi_2\varphi_4}\Big]\end{aligned} \tag{16}$$

In the same way it arises that

$$\sigma^2_{\theta'} = \frac{1}{4}\Big[\sigma^2_{\varphi_4} + \sigma^2_{\varphi_2} + 2\sigma_{\varphi_2\varphi_4}\Big] \tag{17}$$

For the evaluation of $\sigma_{\varphi_2\varphi_4} = \sum_{i=1}^{n}\frac{\partial[\varphi_4\cdot\varphi_2]}{\partial N_i}\sigma^2_{N_i}$ operations similar to the retrieval of $\sigma^2_{\varphi_k}$ can be applied and it follows that:

$$\begin{aligned}\sigma_{\varphi_2\varphi_4} = \frac{4}{9A^2}\Big[&\cos\varphi_4\cos\varphi_2\Big(\frac{A}{2}\cos\varphi_2 - \frac{\pi A}{2}\cos\varphi_3\Big) + \cdots\\ &\cdots + \sin\varphi_4\cos\varphi_2\Big(\frac{A}{2}\sin\varphi_2 + \frac{\pi A}{2}\sin\varphi_3\Big) + \cdots\\ &\cdots + \cos\varphi_4\sin\varphi_2\Big(-\frac{A}{2}\sin\varphi_2 + \frac{\pi A}{2}\sin\varphi_3\Big) + \cdots\\ &\cdots + \sin\varphi_4\sin\varphi_2\Big(\frac{A}{2}\cos\varphi_2 + \frac{\pi A}{2}\cos\varphi_3\Big)\Big].\end{aligned} \tag{18}$$

Putting together the solutions for $\sigma^2_{\varphi_2}$, $\sigma^2_{\varphi_4}$ and $\sigma_{\varphi_4+\varphi_2}$ through Equation (16) and (17) it is possible to compute the variances of the direct phase and pattern phase offset estimation error. In particular it is worth noticing that the mean value of these error variances (without considering the sinusoidal terms) are:

$$\overline{\sigma}^2_{\phi_d} = \overline{\sigma}^2_{\theta'} = \frac{4}{9A^2}B. \tag{19}$$

If we compare the estimation noise variances for the pattern phase offset retrieval from the second and fourth harmonics or using only the third harmonic it comes out that:

$$\begin{aligned} \theta' = -(\varphi_2 + \varphi_4)/2 \implies \overline{\sigma}^2_{\theta'} &= \frac{4}{9A^2}B \\ \theta = -\varphi_3 \implies \overline{\sigma}^2_{\theta} = \overline{\sigma}^2_{\varphi_3} &= \frac{8}{9\pi^2 A^2}B, \end{aligned} \tag{20}$$

from which arises that the second formula gives an estimation of the phase offset that is about 4 times less noisy then the first one and for this reason we used the third harmonic to compute this parameter.

For the evaluation of the variance of the noise acting on the depth estimate employing the STM-ToF acquistion (the standard Whyte approach [4]), we have to consider the linear relation that links the direct phase and the depth, indeed $d_{noMPI} = \frac{c}{4\pi f_{mod}}\phi_d$ and it results that

$$\sigma^2_{d_{noMPI}} = \left(\frac{c}{4\pi f_{mod}}\right)^2 \sigma^2_{\phi_d}. \tag{21}$$

and by considering the mean variance of the noise we have

$$\overline{\sigma}^2_{d_{noMPI}} = \left(\frac{c}{4\pi f_{mod}}\right)^2 \frac{4}{9A^2}B. \tag{22}$$

that is Equation (10P) in the paper.

3 Structured Light Depth Estimation with Implicit Phase Unwrapping

In this section we are going to present the derivation for the implicit phase unwrapping in the Structured Light (SL) depth estimation that we used in the paper.

As mentioned in Section 4 of the paper, in case the phase offset θ_{ref} and θ_{target} are already phase unwrapped in θ^{PU}_{ref} and θ^{PU}_{target}, then the depth map with the SL approach can be estimated as

$$\begin{aligned} d_{SL} &= d_{ref}\Big(1 + \frac{Q}{b}\big(\theta^{PU}_{ref} - \theta^{PU}_{target}\big)\Big)^{-1} \\ &= d_{ref}\Big(1 + \frac{Q}{b}\big(\theta_{ref} + 2\pi k_{ref} - \theta_{target} - 2\pi k_{target}\big)\Big)^{-1} \end{aligned} \tag{23}$$

where $\theta_{target}, \theta_{ref} \in [-\pi; \pi)$ are the phase offsets directly accessible and $2\pi k_{target}$ and $2\pi k_{ref}$ are the offsets which correct the phase wrapping.

In order to unwrap the phase offsets, and so estimate $2\pi k_{target}$ and $2\pi k_{ref}$, it is usually required to project multiple patterns with lower frequencies on the scene. We avoid this by exploiting the depth map computed with the ToF sensor, d_{ToF}. First of all, we consider the pattern phase offset θ_{ToF}^{PU} that originates the depth map d_{ToF} in case of a SL acquisition:

$$\theta_{ToF}^{PU} = \theta_{ref}^{PU} - \frac{b}{Q} \cdot \frac{d_{ref} - d_{ToF}}{d_{ToF}}. \tag{24}$$

If we consider

$$\theta_{ToF} = \theta_{ref} - \frac{b}{Q} \cdot \frac{d_{ref} - d_{ToF}}{d_{ToF}}. \tag{25}$$

we have that

$$\theta_{ToF}^{PU} = \theta_{ToF} + 2\pi k_{ref}. \tag{26}$$

Since the only differences between θ_{ToF}^{PU} and θ_{target}^{PU} are the fluctuations due to noise, by assuming that the noise is smaller than half of the phase wrapping distance, we obtain that

$$|\theta_{ToF}^{PU} - \theta_{target}^{PU}| < \pi. \tag{27}$$

By using together Equation (23) and (24), we have that

$$d_{SL} = d_{ref}\Big(1 + \frac{d_{ref} - d_{ToF}}{d_{ToF}} + \frac{Q}{b}\big(\theta_{ToF}^{PU} - \theta_{target}^{PU}\big)\Big)^{-1} \tag{28}$$

Since we have assumed that $|\theta_{ToF}^{PU} - \theta_{target}^{PU}| < \pi$, recalling that $\theta_{target}^{PU} = \theta_{target} + 2\pi k_{target}$ and $\theta_{ToF}^{PU} = \theta_{ToF} + 2\pi k_{ref}$, it comes out that

$$\begin{aligned}
\theta_{ToF}^{PU} - \theta_{target}^{PU} &= \big(\theta_{ToF}^{PU} - \theta_{target}^{PU}\big)_{[-\pi;\pi)} \\
&= \big(\theta_{ToF} + 2\pi k_{ref} - \theta_{target} - 2\pi k_{target}\big)_{[-\pi;\pi)} \\
&= \big(\theta_{ToF} - \theta_{target}\big)_{[-\pi;\pi)}
\end{aligned} \tag{29}$$

From Equation (28) and (29) it turns out that:

$$d_{SL} = d_{ref}\Big(1 + \frac{d_{ref} - d_{ToF}}{d_{ToF}} + \frac{Q}{b}\big(\theta_{ToF} - \theta_{target}\big)_{[-\pi;\pi)}\Big)^{-1} \tag{30}$$

that is Equation (13P) in the paper and it doesn't require any explicit phase unwrapping operations.

4 Error Estimation in the Structured Light Approach

For the estimation of the error in the SL approach we consider the model of Equation (12P), since only deterministic operations bring Equation (12P) in (13P). We assume that θ_{ref} is noiseless since it has to be captured only once and multiple acquisitions can be repeated in order to remove the noise. The only remaining source of randomness is θ_{target} and by error propagation we can obtain:

$$\begin{aligned}\sigma^2_{d_{SL}} &= \left(\frac{\partial d_{SL}}{\partial \theta_{target}}\right)^2 \sigma^2_{\theta_{target}} \\ &= \left(Q\frac{d^2_{target}}{d_{ref}b}\right)^2 \sigma^2_{\theta_{target}}\end{aligned} \tag{31}$$

from Equation (31) it is possible to notice that the depth estimation accuracy improves if we increase the baseline between the sensor and the projector and it degrades with the increase of d_{target}, that is the depth that we are going to estimate. This is a common behavior for SL systems. The reference distance d_{ref} has no effect in the accuracy since Q is directly proportional to d_{ref} itself. The phase estimation variance $\sigma^2_{\theta_{target}}$ can be retrieved from the second line of Equation (20). It comes out that the mean error variance for the SL depth estimation is

$$\bar{\sigma}^2_{d_{SL}} = \left(Q\frac{d^2_{target}}{d_{ref}b}\right)^2 \frac{8B}{9\pi^2 A^2} \tag{32}$$

References

1. Nayar, S.K., Krishnan, G., Grossberg, M.D., Raskar, R.: Fast separation of direct and global components of a scene using high frequency illumination. ACM Transactions on Graphics (TOG) **25**(3) (2006) 935–944
2. Ku, H.H., et al.: Notes on the use of propagation of error formulas. Journal of Research of the National Bureau of Standards **70**(4) (1966)
3. Lange, R., Seitz, P., Biber, A., Lauxtermann, S.C.: Demodulation pixels in ccd and cmos technologies for time-of-flight ranging. In: Sensors and camera systems for scientific, industrial, and digital photography applications. Volume 3965., International Society for Optics and Photonics (2000) 177–189
4. Whyte, R., Streeter, L., Cree, M.J., Dorrington, A.A.: Resolving multiple propagation paths in time of flight range cameras using direct and global separation methods. Optical Engineering **54**(11) (2015) 113109