Simultaneous Time-of-Flight Sensing and Photometric Stereo with a single ToF Sensor

Changpeng Ti¹, Ruigang Yang¹, James Davis², Zhigeng Pan³

¹University of Kentucky. ²University of California, Santa Cruz. ³Hangzhou Normal University.

Time-of-Flight (ToF) cameras based on phase-shift (such as PMD [3], SwissRanger [2]) take four snapshots to generate a metric depth map. In this paper, we develop novel techniques to allow Phase-shift-based ToF cameras to measure both metric distance and surface normal.

Our method uses four LED sources placed away from the image sensors, as shown in Figure 1. We illuminate the scene with these four lights, one at a time. Based on four phase images from such a distributed light setup, new algorithms are developed to recover both the surface normal and the distance for each pixel.

The classic ToF model requires that (1) the extra phase delay caused by the distance between the light source and the camera is negligible, and (2) pixels in the four phase images share the same intensity value. In our distributed light setup, neither condition remains true.

To tackle this problem, we explicitly model the phase delay caused by light positions as an unknown and use an iterative optimization scheme to solve it as well as the phase delay caused by scene depth.

One can use the classic formula to generate an initial phase delay estimate, and use the light source positions to perform further iterative refinements. In each iteration, a new phase delay \( \phi' \) is estimated based on the current \( \phi \). Given the unit camera ray \( r \), and the light positions \( L_i \), where \( i = \{0, 1, 2, 3\} \) is the index of the light sources, one can write the phase delay \( \alpha_i \) caused by the light source positions as

\[
\begin{align*}
\alpha_i &= \frac{c\phi}{2\pi f_{\text{read}}} \left( \frac{d}{|r \cdot \alpha_i - L_i| - d} \right)^2 - \frac{c\phi}{2\pi f_{\text{read}}} \left( \frac{2d}{|r \cdot \alpha_i - L_i|} \right) \\
\end{align*}
\]

(1)

The relation between the phase delay \( \phi' \) due to depth and the phase delay \( \alpha \) due to light source positions can be written as

\[
A \cos \left( \frac{i\pi}{2} + \alpha_i + \phi' \right) = I_i = \frac{I_{\text{phase}}}{I_{\text{intensity}}} \tag{2}
\]

where \( A \) is the amplitude of the ToF autocorrelation function (ACF), \( I \) is the value of the ACF, and \( I_{\text{phase}} \) and \( I_{\text{intensity}} \) are the intensity values in the phase and intensity images respectively. This expands to (in matrix form)

\[
\begin{bmatrix}
\cos \alpha_i \\
\sin \alpha_i \\
\cos(\alpha_i + \frac{\pi}{2}) \\
\sin(\alpha_i + \frac{\pi}{2}) \\
\cos(\alpha_i + \frac{\pi}{2}) \\
\sin(\alpha_i + \frac{\pi}{2}) \\
\cos(\alpha_i + \frac{\pi}{2}) \\
\sin(\alpha_i + \frac{\pi}{2}) \\
\end{bmatrix}
A \cos \phi' \\
A \sin \phi' \\
= \\
\begin{bmatrix}
I_0 \\
I_1 \\
I_2 \\
I_3 \\
\end{bmatrix} \tag{3}
\]

The process from Eq. 1 to Eq. 4 is repeated until \( ||\phi' - \phi|| \) is smaller than a threshold, or when the number of iterations reaches a limit. Upon convergence, we will obtain a refined depth map of the scene.

Photometric Stereo method estimates the surface normal \( N \) by solving a system based on the Lambertian assumption. With point light sources, the light direction vectors need to be calculated for every pixel. Similar to the work by Clark [1], we further use a quadratic term to model light attenuation. Assuming the light source is infinitely small, the Lambertian shading equation can be rewritten as (in matrix form)

\[
\begin{bmatrix}
V_{ij}^0 \\
V_{ij}^1 \\
V_{ij}^2 \\
V_{ij}^3 \\
\end{bmatrix} \cdot k_d N_i = \\
\begin{bmatrix}
I_0 - |D_i - L_i|^2 \\
I_1 - |D_i - L_i|^2 \\
I_2 - |D_i - L_i|^2 \\
I_3 - |D_i - L_i|^2 \\
\end{bmatrix} \tag{5}
\]

where \( V_{ij}^k \) are the pixel values in the four phase images, \( k_d \) is the depth for each pixel.

Figure 1: (Left) overview of our depth enhancement framework. (Right) illustrations of our novel capturing setup that consists of a ToF sensor and four point light sources.