Performance of Time of Flight Coding Function under Global Illumination

Felipe Gutierrez Barragan

ECE 901 Project Report, UW-Madison

Abstract

Continuous-Wave Time-of-Flight (CW-ToF) imaging have become the preferred depth sensing technology for many applications. The depth precision achieved by a CW-ToF system is highly dependent on the set of light modulation and sensor demodulation functions used. Recent work has led to the development of novel sets of CW-ToF coding functions (modulation and demodulation) that are robust to strong photon noise and sensor-related noise. These coding functions were mainly designed assuming that a sensor pixel only receives direct illumination from the point in the scene being imaged. In practice, that sensor pixel will also receive a global illumination component. In this project we will benchmark the performance of these new coding functions in the presence of global illumination.

Keywords: Time of Flight, Global Illumination, Simulation

1. Introduction & Background

1.1. Time of Flight Imaging

Time of Flight (ToF) refers to the time that a light pulse takes to travel from a source to a target scene back to a detector. This technique is often used to recover scene depth and geometry. Continuous-Wave ToF (CW-ToF) is one particular low-cost ToF setup where the light source intensity and sensor exposure are temporally modulated by a modulation ($M(t)$) and demodulation ($D(t)$) function, respectively. The light reflected from the scene and incident on a sensor pixel (p) will be a scaled, phase shifted, and vertically shifted version of the $M(t)$ denoted as the incident radiance, $L(p, t)$ (see equation 1. Note that the scene depth, Γ, is encoded in the temporal shift (phase shift, $\phi = \frac{2\Gamma}{c}$) of $L(p, t)$. This process is illustrated by Figure 1.

The brightness, $B(p)$, measured at a sensor pixel is the temporal correlation between $L(p,t)$ and $D(t)$. Since there are 3 unknowns (depth, ambient, and albedo) at least $K \geq 3$ measurements are needed to recover depth.

$$L(p,t) = \beta(p)M(t - \frac{2\Gamma}{c}) + L_a \tag{1}$$

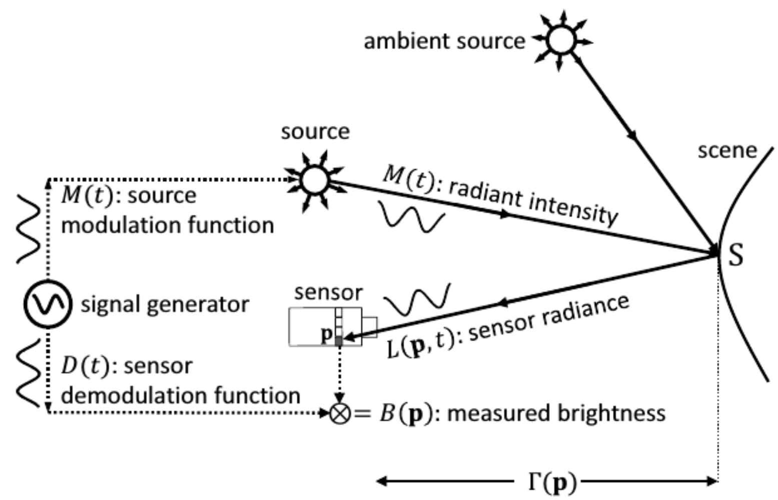

Figure 1: Illustration of a CW-ToF setup.

1.2. Direct & Global Illumination

CW-ToF cameras assume that sensor pixels receive light only due to *direct illumination* of scene points from the source. In practice, due to *global illumination*, the sensor pixel also receives light that has traveled along different paths after multiple reflections. Accurate recovery of scene depths requires the separation of the direct and global illumination components. This is a difficult task and an active research area [1, 2].

1.3. Time of Flight Coding Functions

Recent work in CW-ToF imaging has shown that the depth precision of these systems is tightly related to the set of coding functions $M(t)$ and $D(t)$ [3, 4]. In most current techniques, the functions are sinusoidal or square waves, which as shown in [4] are sub-optimal. The coding functions presented in [3] were mainly designed to be robust to strong ambient illumination and sensor noise. In this project we will test the performance of these new coding functions in scenarios with various noise levels and global illumination effects.

2. Methods

2.1. Image Formation Model

The measured brightness at a sensor pixel, $B_i(\alpha, \Gamma, L_a)$, is a function of the coding function pairs $M_i(t)$ and $D_i(t)$, scene point reflectance (β), depth (Γ), and ambient illumination (L_a). The incident radiance on the sensor pixel is temporally correlated with the demodulation function as shown in equation 2 for a fixed integration time T.

$$B_i(\beta, \Gamma, L_a) = \int_0^T D(t)(\beta M(t - \frac{2\Gamma}{c}) + L_a)dt \tag{2}$$

If we pool together some of the constance factors from equation 2 we get the following equation

$$B_i(\alpha, \Gamma, L_a) = \alpha F_i(\Gamma) + \gamma_i L_a \tag{3}$$

where,

$$\begin{aligned} F_i(\Gamma) &= \int_0^T D(t)M(t - \frac{2\Gamma}{c})dt \\ \gamma_i &= \int_0^T D(t)dt \\ 1 \leq i \leq K, &\qquad K \geq 3 \end{aligned} \tag{4}$$

wheere α is the scaled albedo and $F_i(\Gamma)$ is the set of K *correlation functions*. The brightness equation 3 has 3 unknowns, hence, we need at least 3 brightness measurements to be able to decode th depth. Furthermore, each brightness measurement i will be associated to a different correlation function.

2.2. ToF Coding Schemes

The set of K correlation functions $F_i(\Gamma)$ is hereafter referred to as a ToF coding scheme. Commercial ToF systems commonly use a sinusoidal coding scheme such as the one shown in figure 2, where the modulation functions are a cosine and each demodulation function is a sinusoid at a $\frac{\pi}{2}$

shift from each other. Recent work proposed novel types of coding schemes called hamiltonian coding schemes robust to noise. These are illustrated in figure 3. Each different set of K values in the correlation function correspond to a different depth Γ. In this project we compare the performance of these three coding schemes in the presence of global illumination.

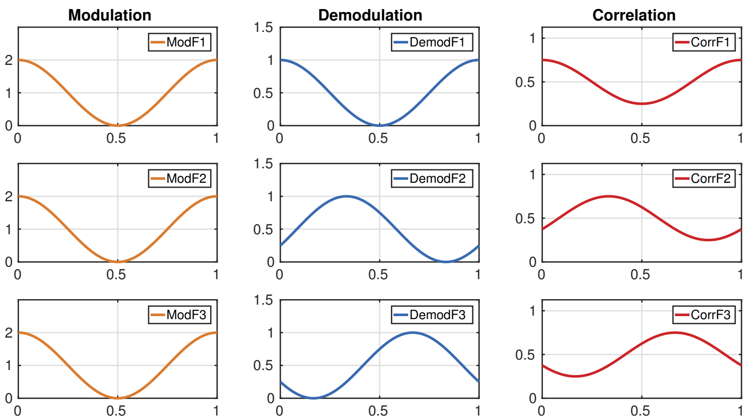

Figure 2: Sinusoidal coding scheme.

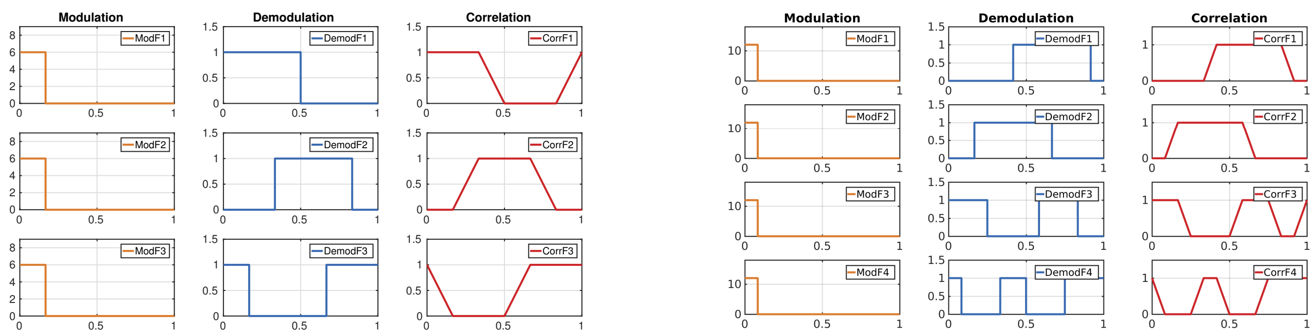

Figure 3: Hamiltonian coding scheme for K=3 (left) and K=4 (right) correlation functions.

2.3. CW-ToF Simulation Pipeline

To evaluate the performance of the different coding schemes we have to be able to simulate the process of acquiring a brightness measurement given: scene parameters, sensor parameters, source power, and ambient power. The

CW-ToF simulation pipeline followed in this project is shown in figure 4. First, the modulated light is propagated to the scene, scaled by the scene reflectivity, vertically shifted by the ambient light, and phase shifted due to the propagation distance. The received radiance is correlated with the sensor demodulation function and integrated. The variance of the gaussian photon noise is determined by the magnitude of the brightness measurement (integration of the radiance and demodulation). Then we check if the brightness value obtained saturated the sensor. The next step is to add gaussian read noise. Finally, in an accurate CW-ToF simulation you would perform the analog to digital conversion which introduces quantization noise, however, this last step is not performed in our simulation.

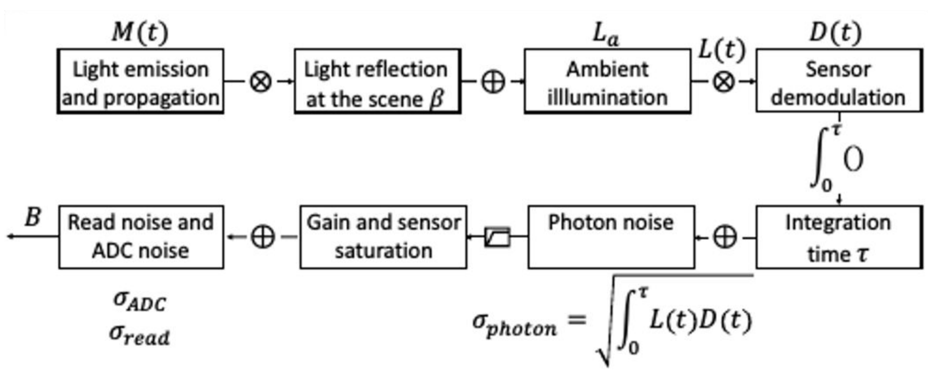

Figure 4: Simulation Pipeline for CW-ToF.

Impulse Response of a Scene: Steps 1 through 4 of the simulation pipeline can be easily implemented if we assume only direct illumination. The phase shift applied would simply be determined by the distance traveled by $M(t)$. When accounting for global illumination the implementation becomes less trivial because on top of the direct response we will also have a long residual signal due to light that bounced off multiple times before returning to the detector. To accurately account for that multipath residual in our pipeline steps 1 through 4 are performed in the following way:

1. Obtain the impulse response for a scene point. To obtain the impulse

response we set the modulation function to a delta pulse and we assume a sensor that can sample very fast. Figure 5.

2. Convolute the impulse response of the scene point with the modulation

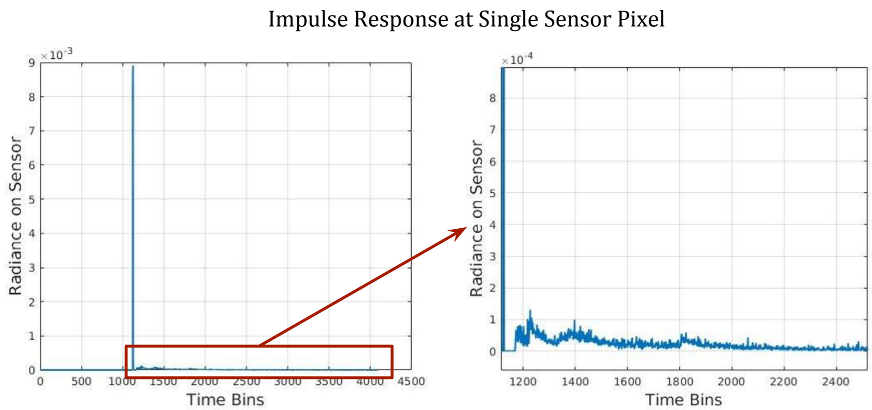

Figure 5: Impulse response at a sensor pixel. The left plot shows the zoomed in residual signal caused by global illumination effects.

In order to obtain the impulse response of the sensor pixel we use a transient rendering code [5].

3. Experimental Setup

We evaluate two different scenarios. First, we take a small patch from a complex scene that leads to more comprehensible depth map results. The patch and the scene it is obtained from are shown in Figure 6. For each of these scenarios we perform two simulations. The first one with the direct illumination component only. The second simulation with both the direct and global illumination component.

Mean Absolute Depth Error: In order to evaluate the performance of each coding scheme we calculate the depth for each pixel. We then take per pixel absolute difference between the calculated depth and the true depth. Finally, we take the mean across all pixels.

Coding Schemes: We compare the three coding schemes displayed in figures 2 and 3. I will refer to each of these coding schemes in the results sections as Cosine K=3, Hamiltonian K=3, and Hamiltonian K=4.

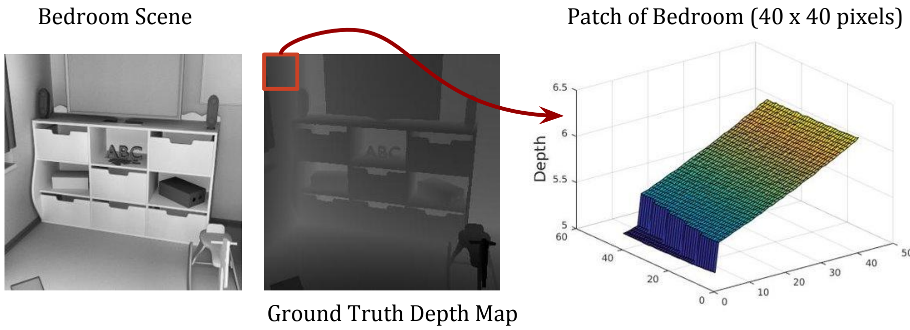

Figure 6: Complex scene used for in the simulations. The performance (mean absolute depth error) is evaluated on both the full scene depth map (middle) and a small 40x40 patch of the scene (right most).

4. Results

In this project we evaluated the coding schemes in two cases under different noise levels: direct illumination only, and direct and global illumination.

4.1. Scene Patch Results

Figures 7 and 8 show the reconstructed depth maps for the scene patch at various noise levels for direct and for direct and global illumination. As the noise is increased all the coding schemes are slightly affected, however, it is evident that cosine is the one less robust to high noise levels. Figure 9 shows the mean absolute depth errors of the 3D reconstruction obtained by each coding scheme.

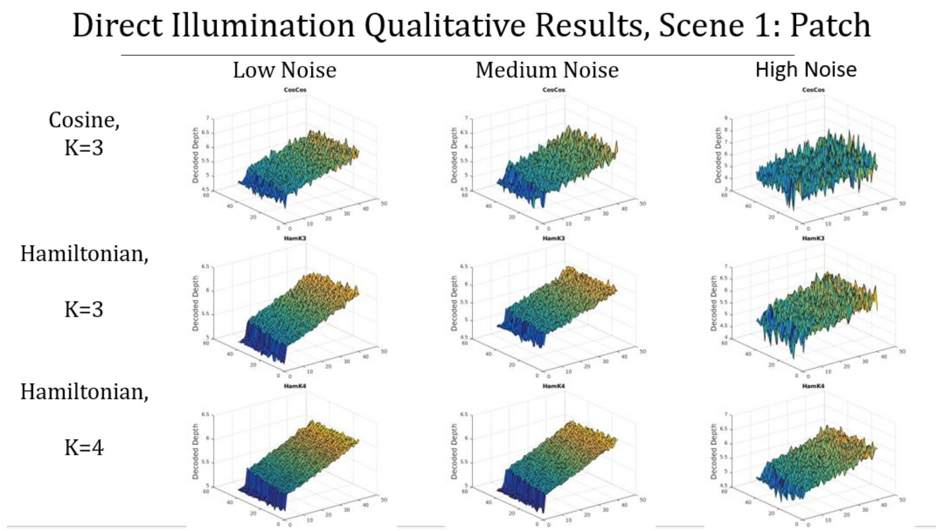

Figure 7: Recovered depth map of the scene patch at various noise levels. We only consider the direct illumination component in this 3D reconstruction.

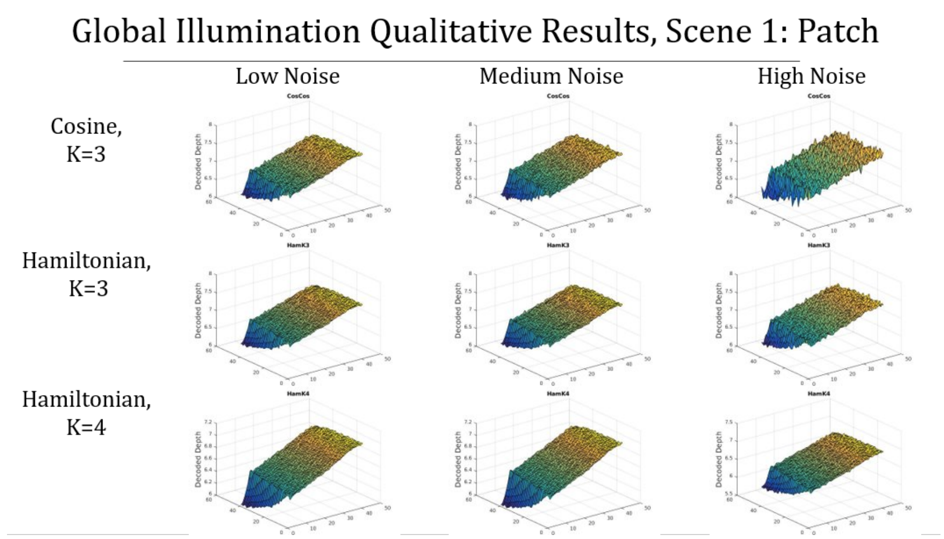

Figure 8: Recovered depth map of the scene patch at various noise levels. We consider both direct and global illumination component in this 3D reconstruction. If you zoom into the depth scale it it evident that there is an absolute depth offset caused by the global illumination residual signal.

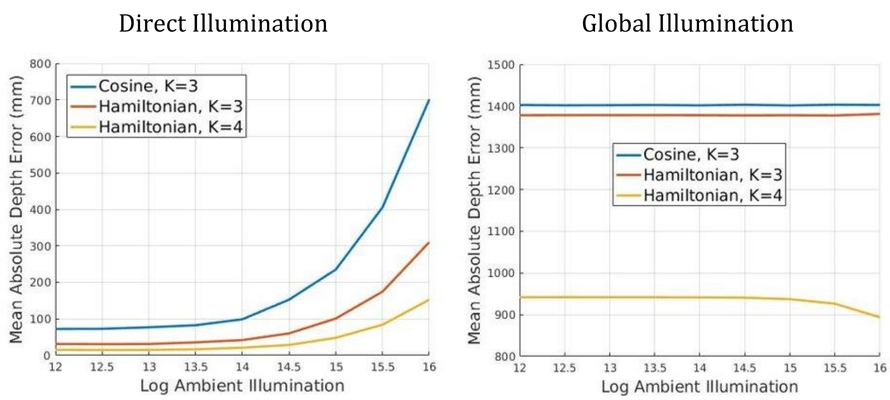

Figure 9: Mean absolute depth error of the recovered 3D reconstructions under direct (left) and direct+global (right) illumination of the 40x40 scene patch. The source strength is fixed while the ambient illumination strength is increased.

4.2. Bedroom Results

Figures 10 and 11 show the reconstructed depth maps for the full bedroom scene at various noise levels for direct and for direct and global illumination. As the noise is increased all the coding schemes are slightly affected, however, it is evident that cosine is the one less robust to high noise levels. The performance of all coding schemes also degrades on scene pixels with low reflectivity. Figure 12 shows the mean absolute depth errors of the 3D reconstruction obtained by each coding scheme.

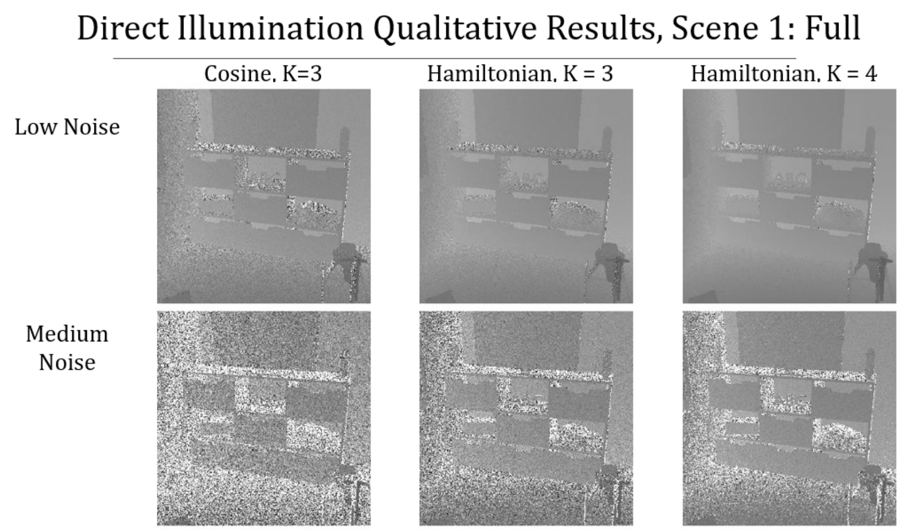

Figure 10: Recovered depth map of the bedroom at various noise levels. We only consider the direct illumination component in this 3D reconstruction.

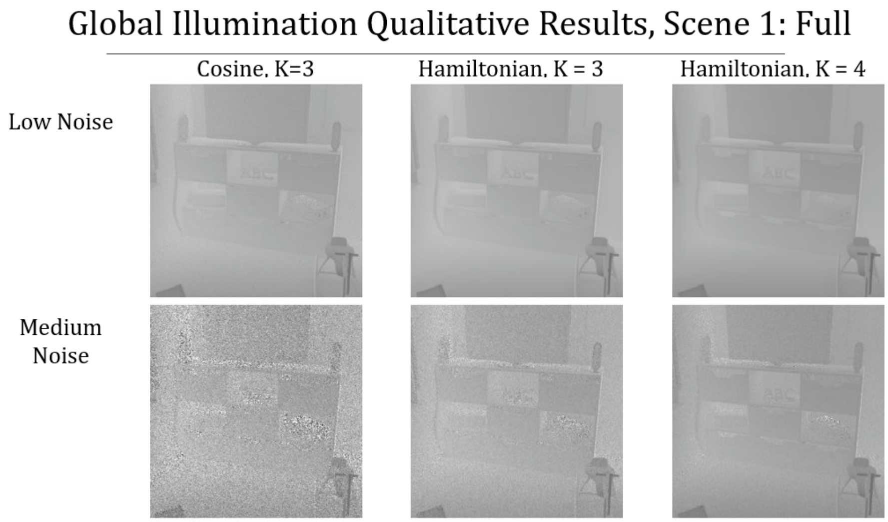

Figure 11: Recovered depth map of the bedroom at various noise levels. We consider both direct and global illumination component in this 3D reconstruction. Scene pixels with low reflectivity correspond to the noisier parts of the scene.

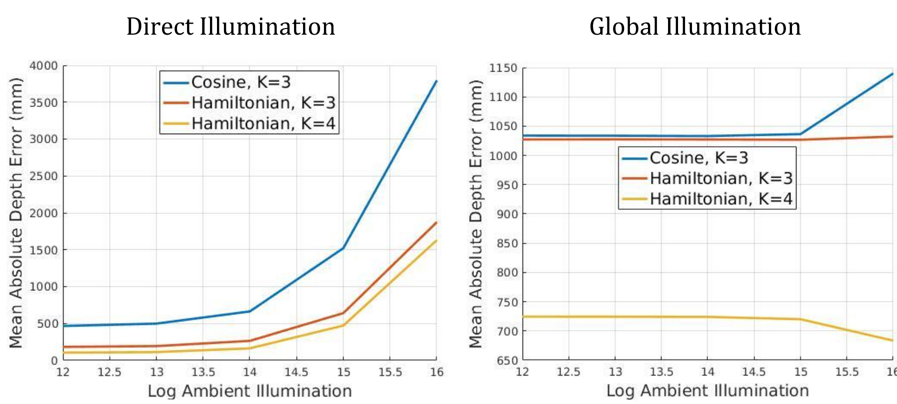

Figure 12: Mean absolute depth error of the recovered 3D reconstructions under direct (left) and direct+global (right) illumination of the full bedroom scene.

5. Discussion

Direct Illumination: As expected in the direct illumination simulations the hamiltonian coding schemes outperform sinusoidal coding at all noise levels. At high noise the geometry of the scene patch in Figure 7 is completely gone for the cosine coding, but the Hamiltonian coding schemes are still able to recover it. For the full bedroom scene at medium noise all coding schemes seem to start giving highly noisy depth maps in some regions. This is likely due to low reflectivity at at those locations.

Global Illumination: In the global illumination simulations we find that the variance of the recovered depth map is reduced compared to the direct only simulations. This is due to the fact that if we take into account the global component we are also taking into account more signal. This means that we will measure a stronger signal and effectively increasing the overall signal to noise ratio. More interestingly, is the fact that depth error due to noise plays a very small role when compared to the error caused by the global illumination component. Figures 9 and 12 demonstrate this dominance because the mean depth error seems to stay constant as we increase the noise levels. Surprisingly, even though the hamiltonian coding schemes were not designed to be robust to global illumination they still consistently outperform cosine coding. This could potentially be due to the higher frequencies used in these schemes. Previous work has shown that some global illumination effects disappear when the coding functions use high frequencies [1].

6. References

[1] Mohit Gupta, Shree K Nayar, Matthias B Hullin, and Jaime Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM Transactions on Graphics (ToG)*, 34(5):156, 2015.

[2] Achuta Kadambi, Jamie Schiel, and Ramesh Raskar. Macroscopic interferometry: Rethinking depth estimation with frequency-domain time-of-flight. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 893–902, 2016.

[3] Mohit Gupta, Andreas Velten, Shree K Nayar, and Eric Breitbach. What are optimal coding functions for time-of-flight imaging? 2017.

[4] Amit Adam, Christoph Dann, Omer Yair, Shai Mazor, and Sebastian Nowozin. Bayesian time-of-flight for realtime shape, illumination and albedo. *IEEE transactions on pattern analysis and machine intelligence*, 39(5):851–864, 2017.

[5] Adrian Jarabo, Julio Marco, Adolfo Muñoz, Raul Buisan, Wojciech Jarosz, and Diego Gutierrez. A framework for transient rendering. *ACM Transactions on Graphics (SIGGRAPH Asia 2014)*, 33(6), 2014.