Practical Coding Function Design for Time-of-Flight Imaging

Felipe Gutierrez-Barragan, Syed Azer Reza, Andreas Velten, Mohit Gupta
University of Wisconsin-Madison
{fgutierrez3, sreza2, velten}@wisc.edu, mohitg@cs.wisc.edu

Abstract

The depth resolution of a continuous-wave time-of-flight (CW-ToF) imaging system is determined by its coding functions. Recently, there has been growing interest in the design of new high-performance CW-ToF coding functions. However, these functions are typically designed in a hardware agnostic manner, i.e., without considering the practical device limitations, such as bandwidth, source power, digital (binary) function generation. Therefore, despite theoretical improvements, practical implementation of these functions remains a challenge. We present a constrained optimization approach for designing practical coding functions that adhere to hardware constraints. The optimization problem is non-convex with a large search space and no known globally optimal solutions. To make the problem tractable, we design an iterative, alternating least-squares algorithm, along with convex relaxation of the constraints. Using this approach, we design high-performance coding functions that can be implemented on existing hardware with minimal modifications. We demonstrate the performance benefits of the resulting functions via extensive simulations and a hardware prototype.

1. Introduction

Time-of-Flight (ToF) cameras are fast becoming the preferred depth sensing technology in several applications, including autonomous navigation, augmented reality and user interfaces. Broadly, ToF cameras can be categorized in two ways: Impulse ToF (I-ToF), and continuous-wave ToF (CW-ToF). I-ToF systems estimate depths by emitting a high-powered laser pulse and measuring its time-of-flight using a high-speed detector. Typical I-ToF systems (e.g. LiDAR) need high-cost components such as high-speed lasers and sensors. In contrast, CW-ToF cameras are low-power and low-cost (e.g. Kinect), and estimate depths by modulating the light intensity and decoding the reflected signal.

Current CW-ToF cameras have limited depth resolution, especially in low signal-to-noise ratio (SNR) scenarios. While the resolution can be increased using high-powered light sources and longer exposure times, this is not always possible. Given a fixed energy and exposure time budget, the depth resolution of a CW-ToF system depends on the coding functions (i.e. light modulation and sensor demodu-

lation functions). Most existing CW-ToF systems use sinusoid [22] or square [12, 24] functions, which, although easy to implement, have low depth resolution.

Recently, there has been growing interest in designing high-performance CW-ToF codes [1, 15, 27], albeit in a hardware agnostic setting. These works propose novel coding functions, which, although theoretically interesting, do not adhere to practical constraints imposed by current devices. For example, these coding functions [15, 1] require impulse modulation with infinite peak power and bandwidth, and continuous, analog demodulation functions. In practice, the space of physically realizable coding functions is highly constrained. Specifically, although high-cost lasers can emit ultra-short pulses with large peak power, most light sources employed in CW-ToF cameras (e.g., laser diodes, LEDs) have finite peak power. Furthermore, most components of the system have a limited bandwidth. Additionally, most CW-ToF systems use digital square/binary functions for light modulation and sensor demodulation, and may need considerable hardware modifications to support arbitrary continuous functions. Finally, binary functions are also preferred because of their applicability to 2-bucket sensors used in commercial CW-ToF cameras [29].

This leads to a trade-off in CW-ToF code design: Conventional coding schemes, such as sinusoid and square, adhere to practical constraints, but have low performance. On the other hand, recent schemes promise higher performance, but, may not be physically realizable on current hardware. In this paper, we consider the following question: Can we design CW-ToF coding functions that adhere to hardware constraints, and yet, achieve high performance in a broad range of imaging scenarios? We consider three important constraints: Limited peak source power, limited bandwidth, and binary coding functions. [1]

To address this question, we pose the practical CW-ToF code design as a constrained optimization problem. This is challenging due to the large space of coding functions and the non-convex constraints. In order to keep the problem tractable, we consider pre-designed correlation functions, which are factorized into physically realizable modulation and demodulation functions. While this reduces the

---

[1] Binary functions, theoretically, have infinite bandwidth. We consider both binary and bandwidth constraints simultaneously by assuming that the bandwidth constraints are applied while being passed through the imaging system, *after* the binary coding functions are generated.

search space, factorizing a given correlation function into modulation and demodulation function is ill-posed. To perform this factorization, we use an iterative optimization approach, similar to blind deconvolution [2, 21], where at each iteration a convex sub-problem is solved. This approach, along with convex relaxations of constraints, enables optimizing over the space of practical coding functions.

The proposed approach is flexible; given a hardware implementation, it computes high-performance coding functions that adhere to the corresponding constraints and are physically realizable. We demonstrate, via extensive simulations and hardware experiments, that the performance of the proposed coding functions, although lower than theoretical (but impractical) coding functions [15], is considerably higher than conventional approaches, such as sinusoid and square coding. We also demonstrate how to adapt the resulting practical codes for a 2-bucket architecture. The proposed approach provides a framework for designing practical ToF coding functions under a wide range of constraints, from weak to strong constraints, such as severely limited bandwidth and peak power, and enable 3D imaging in challenging imaging scenarios (e.g. outdoors, low-power).

2. Related Work

Code design for CW-ToF: Many coding designs have been proposed to address issues in CW-ToF imaging, such as, mitigating multi-camera interference [4, 29], multi-path interference [20, 14] and non-linearities caused by aliased harmonics [25, 24]. Furthermore, CW-ToF code design has also been used for transient imaging [20] and illumination demultiplexing [19]. There has been little work on designing coding schemes that achieve higher depth resolution. Payne et al. [26] proposed multi-frequency sinusoid functions to simultaneously achieve large depth range and high depth precision. This approach requires phase unwrapping and is prone to large errors in low SNR scenarios. The focus of this paper is on designing practical coding schemes that achieve high depth resolution, even in low SNR scenarios.

Code design for structured light: Code design for improved depth resolution has been explored in other active vision systems, in particular structured light (SL) [18, 23, 13]. Despite theoretical similarities between SL (spatial coding) and CW-ToF (temporal coding), there are crucial differences in their practical hardware limitations. For instance, typical SL systems can implement arbitrary continuous patterns [6] limited only by the spatial resolution of the system. In contrast, CW-ToF systems are limited to sinusoidal and square functions, with strong bandwidth constraints (< 200 MHz) which severely limits the achievable depth resolution.

3. Background: Image Formation Model

We begin by reviewing the CW-ToF image formation model [22, 16, 15]. A CW-ToF imaging system (Figure 1) consists of a light source whose intensity is modulated
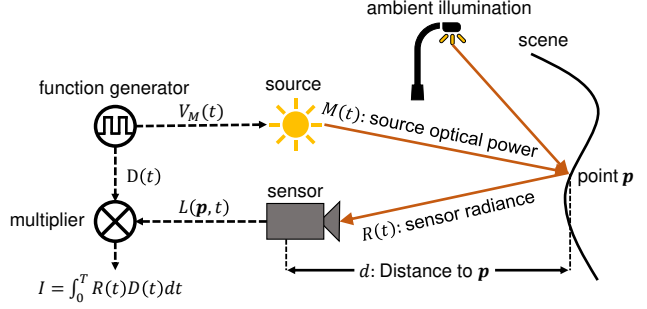

Figure 1: **CW-ToF imaging.** The intensity of a light source is temporally modulated according to a function $M(t)$. The incident sensor radiance, $R(t)$, is multiplied with a demodulation function, $D(t)$. The product is integrated for an exposure time T and the resulting intensity, I, is recorded.

over time according to a *modulation function* $M(t)$. Assuming no indirect illumination, the radiance $R(t)$ incident on a sensor pixel at time t is given as:

$$R(t) = s\, M\left(t - \frac{2\,d}{c}\right) + A\,, \qquad (1)$$

where d is the distance of the scene point from the sensor, c is the speed of light, and s is a scene dependent scale factor which depends on the scene point's albedo, shading and intensity fall-off. A is the ambient illumination due to external light sources (e.g., sunlight in outdoor scenarios).

The incident radiance $R(t)$ is correlated with a *demodulation function* at the sensor, $D(t)$, where $0 \leq D(t) \leq 1$. [2] The demodulation function can be implemented through photonic mixer devices [22], or fast optical shutters [5]. We assume both $M(t)$ and $D(t)$ to be periodic functions with the same fundamental frequency f_0 (homodyne CW-ToF).

The intensity I measured at the pixel is given by the temporal correlation of $R(t)$ and $D(t)$:

$$I = \int_0^T D(t)\, R(t)\, dt\,, \qquad (2)$$

where T is the sensor exposure time. We define $\chi(d)$ as the normalized cross-correlation between $M(t)$ and $D(t)$:

$$\chi(d) = (D \star \bar{M})(d) = \int_0^\tau D(t)\bar{M}\left(t - \frac{2d}{c}\right)\, dt\,, \quad (3)$$

where τ is the period of $M(t)$ and $D(t)$, $E_\tau = \int_0^\tau M(t)dt$, is the total energy emitted by the light source within a period τ, and $\bar{M}(t) = \frac{M(t)}{E_\tau}$ is the normalized modulation function.

---

[2] 2-bucket ToF sensors [22] *emulate* negative, zero-mean demodulation functions by taking the difference of two complementary measurements. Since these functions are the difference of two positive demodulation functions, we only consider positive functions in our analysis. In Section 6.4, we show how the proposed binary demodulation functions can be adapted as complementary demodulation pairs that can be used in 2-bucket sensors.
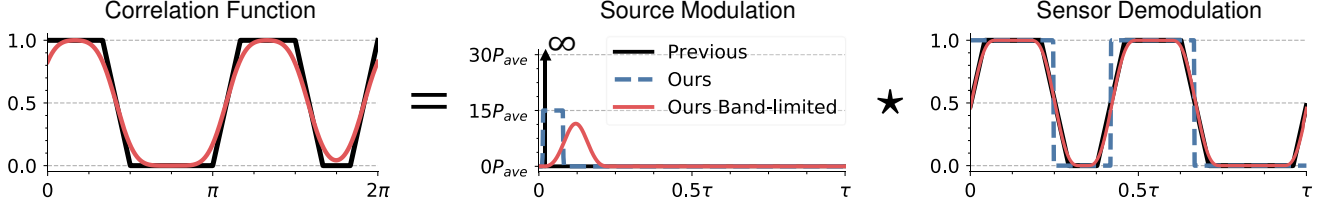
Figure 2: **Practical coding function design.** The black lines illustrate the pulsed ToF implementation of a high-performance correlation function [15]. Our algorithm designs the input modulation and demodulation functions (blue lines) that adhere to peak power and binary constraints. The input functions are smoothed according to the system's impulse response and result in the red lines above. The actual correlation function achieved by the system is the correlation of the smoothed output functions. Our algorithm designs the blue functions to produce the best approximation of the objective correlation function.

Substituting Eq. 1 into Eq. 2, and simplifying using Eq. 3, we get the following expression for the measured intensity:

$$I = I_s \, \chi(d) + I_a \,, \tag{4}$$

where $I_a = A \int_0^T D(t)dt$, $I_s = s \, P_{ave} \, T$, and $P_{ave} = \frac{E_\tau}{\tau}$ is the average power of $M(t)$.

Since there are three unknowns in Eq. 4 (distance d, effective albedo I_s, and ambient brightness I_a), $K \geq 3$ intensity measurements are needed. Each measurement, I_i ($1 \leq i \leq K$), is taken using a different pair of modulation and demodulation functions. The set of K modulation and demodulation functions is called a *coding scheme* [15].

4. Constrained Optimization for Code Design

Designing practical, high-performance coding schemes $(M_i(t), D_i(t))$, $1 \leq i \leq K$, can be posed as a constrained optimization problem. Unfortunately, the space of functions can be large, making this a challenging optimization problem to solve. To keep it tractable, we pose practical coding design as a two-stage problem. The performance of a ToF imaging system, in terms of depth precision, can be completely characterized by its correlation functions χ_i, $1 \leq i \leq K$ [15]. Thus, the first stage designs the correlation functions, without imposing the hardware constraints. This can be performed following any of the previous coding function design approaches [15, 1, 27]. The second stage factorizes each χ_i into modulation and demodulation functions. This step is critical, because for a practical implementation we need to determine the modulation and demodulation functions that create a given correlation function. Mathematically, this factorization problem can be expressed as follows:

$$
\begin{aligned}
&\text{find} && (M_i(t), D_i(t)), \ 1 \leq i \leq K \\
&\text{such that} && (D_i \star \bar{M}_i)(t; d) = \chi_i(d), \\
&\int_0^\tau \bar{M}_i(t)dt \leq 1, && 0 \leq \bar{M}_i(t), \quad 0 \leq D_i(t) \leq 1
\end{aligned}
\tag{5}
$$

Recent coding design works [15, 1] focused only on the first step of designing correlation functions. For the sec-

ond step, they considered a trivial factorization: an impulse modulation function $M_i(t) = E_\tau \delta(t)$, and the demodulation function being the same as the normalized correlation function, e.g., a trapezoidal function (see Figure 2). Although theoretically feasible, this decomposition is impractical because of three fundamental hardware constraints:

1. **Limited peak source power:** An impulse modulation function assumes the laser can emit an infinitely short and high powered pulse, as illustrated in Figure 2. Although, high powered pulsed lasers are used in multiple applications [28], they are often expensive, and not appropriate for low-cost CW-ToF systems.

2. **Limited bandwidth:** An ideal impulse function requires infinite bandwidth. Most current CW-ToF cameras are severely constrained by the available bandwidth of the coding functions (typically < 200 MHz.).

3. **Non-binary coding functions:** The trivial decomposition described above leads to arbitrary continuous-valued demodulation functions. Several commercial CW-ToF systems rely on digitally generated signals to modulate $M(t)$ and $D(t)$ [24, 22], and thus, are limited to binary functions. Furthermore, non-binary codes cannot leverage two-bucket ToF sensors that efficiently use source power, thereby improving SNR.

We present a framework that decomposes a correlation function into a modulation and demodulation function, while *adhering* to pre-specified hardware constraints. Figure 2 shows an example of such a decomposition. Our framework takes as input a desired correlation function (e.g., trapezoidal function shown in black line), the maximum peak power constraint, and the frequency bandwidth, specified as the impulse response of the system. Given these parameters and a binary structure prior, an optimization problem similar to 5 is solved. The designed modulation and demodulation functions (blue dashed lines in Figure 2) will produce the closest approximation of the objective correlation function. These designed modulation and demodulation functions are input to the CW-ToF system, which then
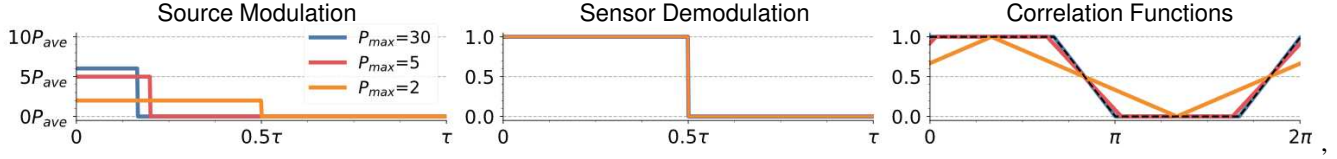
Figure 3: **Power-limited decomposition.** Power-limited correlation function decomposition for Hamiltonian $K = 3$ at different maximum power levels. At $P_{max} \geq 6P_{ave}$ the resulting modulation and demodulation functions have the same correlation as the objective Hamiltonian correlation. This is clearly illustrated in the case of $P_{max} = 30P_{ave}$ (blue) where the chosen solution only uses $6P_{ave}$. When $P_{max} \leq 6P_{ave}$ only approximations of the objective correlations are obtained.

smooths them according to its impulse response. The output modulation and demodulation functions (red lines in Figure 2) will produce the actual correlation function. As an example, we demonstrate our framework on the family of Hamiltonian correlation functions [15]. Specifically, we consider Hamiltonian codes with $3, 4$, or 5 measurements, referred hereafter as Hamiltonian $K = 3$, $K = 4$, and $K = 5$.

5. Power-limited Coding Function Design

In this section, we analyze the correlation decomposition with a finite light source peak power and infinite bandwidth.

5.1. Power-limited Decomposition

Let $\bar{\mathbf{m}}_i$, $\mathbf{d}_i$, χ_i be the vectors representing the N point discretization of a single period of $\bar{M}_i(t)$, $D_i(t)$, and $\chi_i(d)$, respectively. We find the best fitting $\bar{\mathbf{m}}_i$ and $\mathbf{d}_i$ for a given χ_i by solving the following constrained optimization:

$$\underset{\bar{\mathbf{m}}_i, \mathbf{d}_i}{\arg\min} \quad \|\chi_i - \bar{\mathbf{m}}_i \star \mathbf{d}_i\|_2$$

$$\text{subject to} \quad \sum_{j=1}^{N} \bar{\mathbf{m}}_{i,j} \Delta t \leq 1, \quad 0 \leq \bar{\mathbf{m}}_{i,j} \leq p_f, \quad (6)$$

$$0 \leq \mathbf{d}_{i,j} \leq 1, \quad j = 1, \ldots, N$$

where $\star$ is the cross-correlation operator, $\Delta t = \frac{\tau}{N}$ and $p_f = \frac{P_{max}}{\tau P_{ave}}$. For interpretability we fix $\tau = 1$ in the optimization. The problem is solved for each χ_i, $1 \leq i \leq K$.

Alternating convex optimization approach: Problem (6) can be thought of as the 1-dimensional analog of blind image deblurring [10, 7, 21] and in general, is under-constrained and ill-posed. However, if either $\mathbf{m}_i$ or $\mathbf{d}_i$ is fixed, the problem becomes well-posed, and can be solved through standard convex optimization methods [3, 11]. Therefore, one approach to finding a local optima is to alternate between fixing $\mathbf{m}_i$ and $\mathbf{d}_i$ and solving the convex sub-problem. This is done until convergence of the loss function. We implemented the alternating optimization procedure using CVXPY [8]. The coding function vectors, $\mathbf{m}$ and $\mathbf{d}$, are randomly initialized, and χ is set to the objective correlation. In our final implementation, which included the additional regularization discussed below, convergence was often achieved within $5 - 50$ iterations.

Imposing binary function constraint: The alternating convex optimization approach often leads to rapidly fluctuating non-binary solutions for $\mathbf{m}_i$ and $\mathbf{d}_i$. Such solutions may be able to approximate χ_i accurately, but, they are not practical due to their high bandwidth requirements and non-binary structure. Ideally, we would impose a binary constraint on $\mathbf{m}$ and $\mathbf{d}$. Such constraint would make each sub-problem non-convex. Therefore, we relax the binary constraint by imposing a total variation regularization on $\mathbf{m}$ and $\mathbf{d}$, thereby preserving the convexity of the sub-problems. Although this relaxation leads to near-binary sparse solutions for $\mathbf{m}$ and $\mathbf{d}$, it does not guarantee a binary structure. In order to achieve binary solutions, we apply a binary re-projection operator at the end of each iteration.

Power-limited Hamiltonian coding functions: Figure 3 shows the correlation decomposition for one (out of three) Hamiltonian $K = 3$ codes at different maximum power levels. See the supplementary material for the other two correlation functions. When $P_{max} \geq 6P_{average}$ our algorithm is able to find a factorization that achieves zero error: $M(t)$ is a $\frac{1}{6}$ duty cycle square wave with peak power of $6P_{ave}$, and $D(t)$ is a square wave with $\frac{1}{2}$ duty cycle. When $P_{max} < 6P_{ave}$ our algorithm generates coding schemes that only approximate Hamiltonian $K = 3$. In the limiting case, when $P_{max} = 2P_{ave}$, our method returns conventional square coding functions. Please see the supplementary material for power-limited decompositions for Hamiltonian $K = 4, 5$ under different power constraints.

5.2. Performance of Power-limited Coding Schemes

The $L2$ norm of the difference between the objective correlation function and the correlation of $\bar{\mathbf{m}}$ and $\mathbf{d}$ is a metric that is easily expressed in an optimization problem. A more meaningful metric, however, is the *mean depth error* (MDE) achieved by the decomposition, for a given depth range and distribution of scene albedos [15]. See the supplementary report for details of the MDE metric.

Mean depth error comparisons: We used the following parameters. Fundamental frequency, $f_0 = 15$ Mhz, corresponding to a depth range of $0 - 10$ meters. Gaussian noise, with affine noise model including photon noise and sensor read noise with a variance of 20 electrons. Total integration time of 300ms, and scene albedos of $0 - 1$. To perform the comparisons at a wide range of SNR levels, we vary source
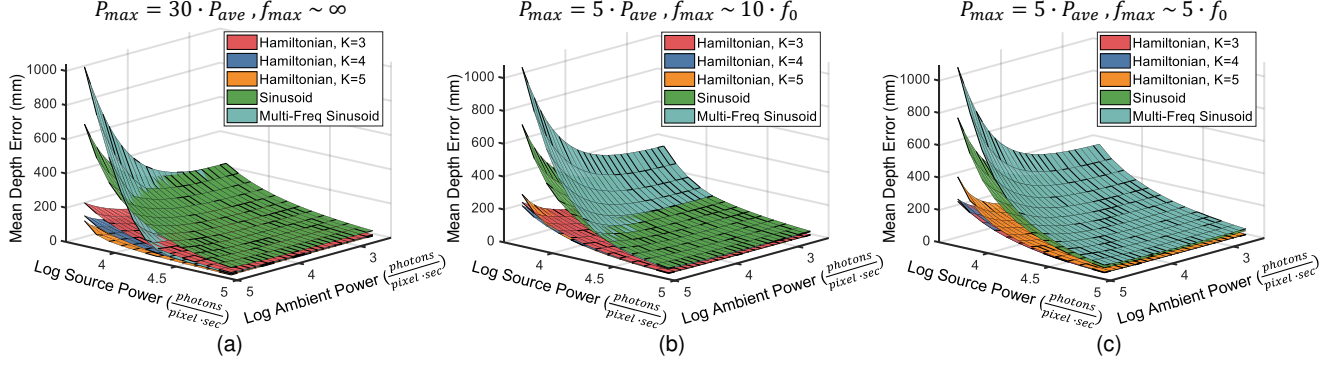
Figure 4: **Mean depth errors of power-limited and band-limited CW-ToF codes.** In the power-limited case (a), all coding schemes achieve their theoretical performance. In the more practical scenarios where the system band-limits are considered (b and c), the re-designed Hamiltonian schemes outperform both baselines at all SNR levels. Zoom in for better visualization.

and ambient strength, while fixing the other parameters.

Figure 4 (a) shows the MDE for Hamiltonian $K = 3$, $K = 4$, $K = 5$, sinusoid [22], and multi-frequency sinusoid [9, 26]. [3] We consider a multi-frequency sinusoid with $K = 5$ measurements; three low frequency measurements with $f = f_0$ and two high-frequency measurements with $f = 7f_0$. We use $7f_0$ because the higher frequency components in Hamiltonian $K = 5$ can be approximated by a square wave with a frequency of $7f_0$. The coding schemes are compared at various peak power constraints. When $P_{max} = 30P_{ave}$, all coding schemes achieve a zero error decomposition. Consequently, the performance of the ideal Hamiltonian codes is maintained. As the peak power constraint is strengthened, i.e., P_{max} is reduced, the decompositions become approximations. In the supplementary material we show the performance at $P_{max} = [15P_{ave}, 5P_{ave}, 2P_{ave}]$. Multi-frequency sinusoid performs well at high SNR levels, but its performance degrades rapidly at low SNR due to phase unwrapping errors. Although the performance of the power-limited Hamiltonian codes is lowered as compared to theoretically ideal codes, they still outperform the sinusoid-based codes at most source and ambient power levels.

6. Band-limited Coding Function Design

In this section we analyze the correlation decomposition problem with *both* peak power and bandwidth constraints.

6.1. Bandwidth of a System

The bandwidth of a system is characterized by the impulse/frequency response. The cascade of hardware components (e.g. function generator, laser, sensor) in the CW-ToF system determines the impulse response. Effectively,

---

[3]For comparing coding schemes with different K, we assume a fixed total integration time which is split evenly among the intensity measurements. Under such constraint the performance of sinusoid schemes does not change as we vary K. For details, see the supplement.

these components smooth the input modulation and demodulation functions, which attenuates frequencies outside the system's bandwidth. Here, it is important to make a distinction between the input coding functions $M_{in}(t)$ and $D_{in}(t)$, and the output functions $M_{out}(t)$ and $D_{out}(t)$, which are created after the input functions are smoothed by the system. We are designing the input functions, but the output functions actually correlate to form the correlation functions. Therefore, the goal of the correlation decomposition for a band-limited system is to find $M_{in}(t)$ and $D_{in}(t)$ whose corresponding (smoothed) $M_{out}(t)$ and $D_{out}(t)$ best approximate the objective correlation.

Furthermore, we apply the bound constraints described in Sections 4 and 5 on the input functions because, in practice, these bounds can be imposed on $M_{in}(t)$ and $D_{in}(t)$ via a function generator. Once the bounds are imposed on the input functions, the output functions will also satisfy them.

6.2. Band-limited Decomposition

Let $h_m(t)$ and $h_d(t)$ be the impulse responses of the modulation and demodulation arms of the system. Then, we solve the following optimization problem to find the input coding functions:

$$\underset{\hat{\mathbf{m}}_i, \mathbf{d}_i}{\operatorname{argmin}} \quad \|\chi_i - (\bar{\mathbf{m}}_i^{in} \circledast h_m(t)) \star (\mathbf{d}_i^{in} \circledast h_d(t))\|_2$$

$$\text{subject to} \quad \sum_{j=1}^{N} \bar{\mathbf{m}}_{i,j}^{in} \Delta t \leq 1, \quad 0 \leq \bar{\mathbf{m}}_{i,j}^{in} \leq p_f, \quad (7)$$

$$0 \leq \mathbf{d}_{i,j}^{in} \leq 1, \quad j = 1, \dots, N$$

where $\circledast$ is the circular convolution operator. We solve problem 7 with the same alternating convex optimization approach with TV regularization described in Section 5.1. Due to stronger constraints, the convergence for the band-limited decomposition was on average faster than the power-limited case, only requiring $3 - 20$ iterations.

Band-limited Hamiltonian coding functions: We assume the modulation and demodulation impulse responses to be
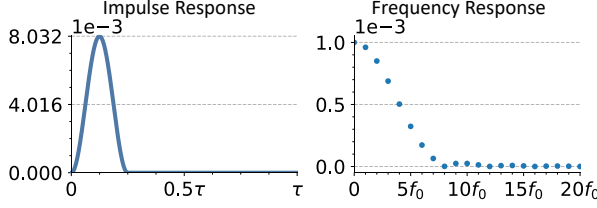
Figure 5: **Hanning window.** Hanning window with $f_{max} \sim 5f_0$ used to model the impulse/frequency response of the band-limiting components in our simulations.

the same, i.e., $h_m(t) = h_d(t)$, and model them as *Hanning windows* [17]. Figure 5 shows an example Hanning window and its frequency response. We evaluate two cases with Hanning windows that constraint the maximum frequency to be $f_{max} \sim 10f_0$ and $f_{max} \sim 5f_0$. We consider f_{max} to be the maximum frequency whose amplitude is not attenuated more than ~ 5x.

Figure 6 shows the input modulation and demodulation functions that best fit two (out of five) Hamiltonian $K = 5$ correlation functions. Due to limited bandwidth, only smoothed approximations of Hamiltonian codes can be achieved. However, the input coding functions ($M_{in}(t)$ and $D_{in}(t)$) generated by our algorithm, continue to have the desired binary structure (columns 1 and 3 in Figure 6); only the output functions ($M_{out}(t)$ and $D_{out}(t)$) are band-limited. Please refer to the supplementary material for all the resulting band-limited factorizations of Hamiltonian codes under different combinations of constraints.

6.3. Performance of Band-limited Coding Schemes

Figures 4 (b) and (c) show the MDE for several coding schemes, under two different bandwidth constraints. The same simulation parameters described in Section 5.2 are used. The smoothing degrades the performance of all coding schemes. We do not apply our framework to sinusoid-based schemes in our evaluation, but we do limit their bandwidth as it happens in practice. Despite the relatively poor approximations of the Hamiltonian codes in the severely band-limited case (Figure 6), the overall performance of the resulting codes is still $2 - 5$x greater than sinusoid-based codes. Please refer to the supplementary material for additional MDE simulations at different maximum power and bandwidth levels.

Figure 7 shows the 3D imaging simulations for a hand gesture reconstruction in two hardware constrained scenarios. In the loosely constrained scenario, all coding schemes perform almost ideally and faithfully recover the 3D geometry. In the more practical setting with tighter power and bandwidth constraints, the overall performance of all coding schemes degrades. However, our new hardware-constrained codes outperform conventional sinusoid codes in both high and low SNR settings.

6.4. Applicability to 2-Bucket CW-ToF Sensors

Complementary demodulation function pairs: One important consequence of enforcing binary demodulation functions is that they can leverage two-bucket CW-ToF sensors. A two-bucket sensor simultaneously acquires two measurements using complementary binary demodulation function pairs [22, 29]. Any coding scheme with K binary demodulation functions (e.g. codes in this paper) can be implemented in a two-bucket sensor.

Figure 8 shows the performance of sinusoid-based and Hamiltonian codes implemented in a two-bucket sensor in two band-limiting cases. Hamiltonian $K = 8$ corresponds to Hamiltonian $K = 4$ adapted for 2-bucket ToF. The four extra measurements are the complementary demodulation functions. Dual-frequency sinusoid corresponds to 4 f_0 and 4 $7f_0$ measurements. Additional discussion and results for two-bucket sensors can be found in the supplement.

7. Hardware Experiments

7.1. Hardware Prototype

We developed a CW-ToF hardware prototype that can implement arbitrary coding schemes, with controllable fundamental frequency and average power. Our setup is based on a point scanning design using a galvo mirror system (GVS012, Thorlabs). The light source is a 830nm laser diode (L830P200, Thorlabs) with a maximum power of ~ 200mW. The intensity of the laser is modulated by an arbitrary function generator (SDG 5162, Siglent) whose maximum frequency is 150Mhz, and is the band-limiting component of the setup. The reflected signal is focused by a lens onto an avalanche photodiode (APD430A, Thorlabs). The detected signal is multiplied with the demodulation function using a frequency mixer (zx05-1L-s+, Mini-Circuits). The output of the mixer is amplified (C7319, Hamamatsu), low-pass filtered (EF110, Thorlabs), sampled and averaged by a data acquisition card (USB-6000DAQ, National Instruments) to obtain intensity measurements. Additional information on the setup is provided in the supplement.

7.2. Experimental Results

We evaluate the performance of all coding schemes under different hardware parameters, constraints, and SNR levels in four scenarios. Table 1 summarizes the experimental parameters used for each scene. The band-limit of our prototype system is modelled similar to the simulations with a maximum frequency of 150Mhz.

Evaluating depth error at a single pixel: Figure 9 shows the mean depth errors of our hardware setup for each coding scheme. The scene is a single point at a distance of 0.7m from the source and the sensor. The laser is collimated and the sensor is focused on the laser spot. To evaluate the mean absolute error (MAE), we estimate the depth of the point 300 times. We emulate the full depth range
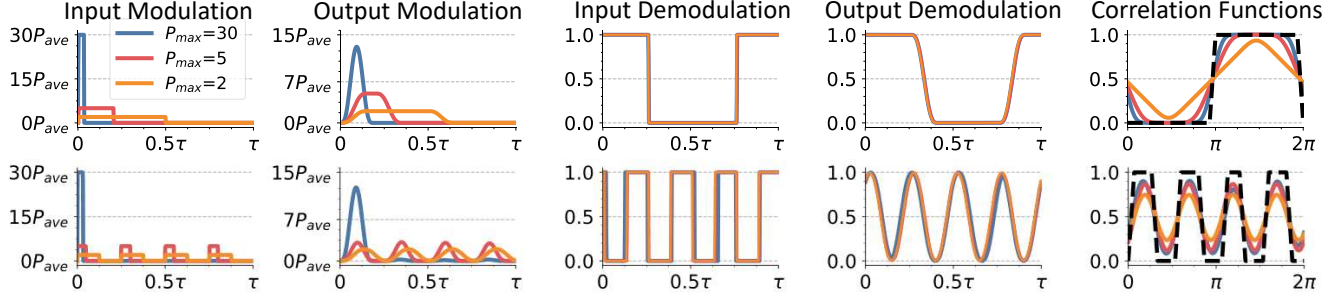
Figure 6: **Band-limited decomposition.** Two coding functions selected from Hamiltonian $K = 5$ are factorized. The bandwidth constraint is $f_{max} \sim 10 f_0$. The input modulation and demodulation functions adhere to their corresponding bound and binary constraints.
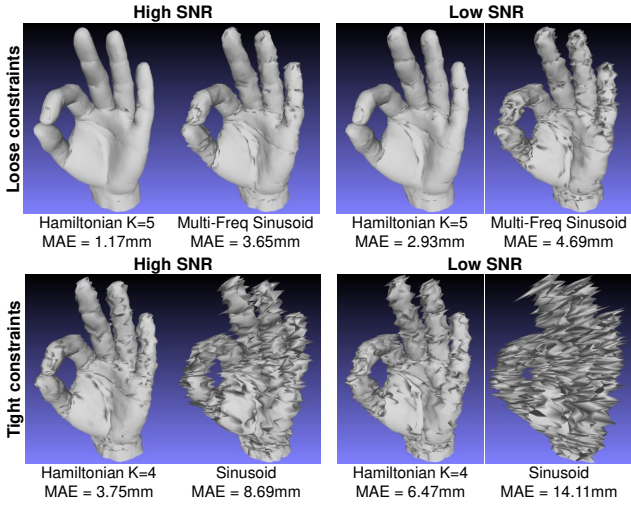

Figure 7: **Simulations under loose and tight constraints.** Comparison of a 3D hand reconstruction under loose ($P_{max} = 20 P_{ave}$, $f_{max} = 20 f_0$) and tight ($P_{max} = 5 P_{ave}$, $f_{max} = 5 f_0$) constraints. Under loose constraints both Hamiltonian and multi-freq sinusoid accurately recover the shape. Under tighter constraints the performance degrades, however, the re-designed Hamiltonian $K = 4$ continues to recover the overall structure and details of the hand.
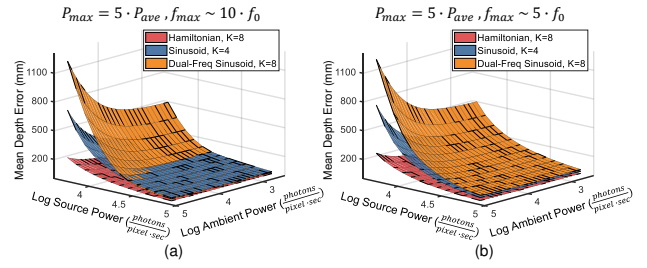

Figure 8: **Performance in two-bucket sensors.** The MDE for two sinusoid-based and one Hamiltonian scheme adapted for a two-bucket architecture. Being able to leverage the two-bucket architecture allows Hamiltonian $K = 8$ to continue outperforming sinusoid-based schemes.
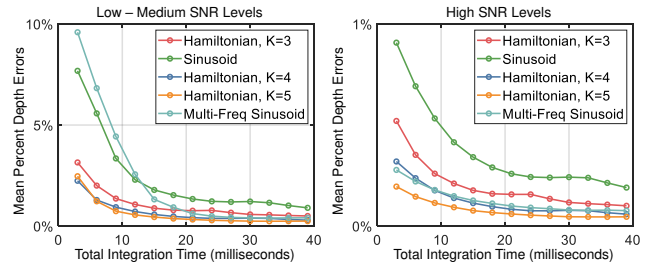

Figure 9: **Experimental mean depth errors.** Single pixel MDE in a power-limited setting where $P_{max} = 15 P_{ave}$. Consistent with simulations, at low-medium SNR power-limited Hamiltonian codes outperform sinusoid-based codes, and at high SNR multi-frequency sinusoid performs comparably to Hamiltonian.

of the coding scheme by artificially shifting the modulation function and calculating the MAE for each shift. The mean depth percent errors in Figure 9 are calculated by averaging the MAE of all shifts (20 shifts) and dividing by the total depth range (30m). Since f_0 is 30x lower than f_{max} on this configuration, we assume infinite bandwidth and evaluate power-limited coding schemes.

3D wall reconstruction: Figure 11 shows the 3D reconstruction for a wall. At both high and low SNR, the practical Hamiltonian codes outperform sinusoid-based codes.

Real-world scenes with complex textures: Figure 10 shows the recovered depth maps for a table-top scene and a volleyball using Hamiltonian $K = 5$ and multi-freq sinusoid. The table-top scene consists of 3 books spaced 20cm

from each other each other, and a background wall $\sim$ 2m behind the first book. The second scene was a 20cm radius volleyball placed 1m away from the sensor. The exposure time used for both scenes is the same. The band-limited Hamiltonian $K = 5$ depth maps have considerably fewer outliers. Multi-freq sinusoid, on the other hand, displays large errors at low SNR due to erroneous phase unwrapping. The MAE are obtained by fitting multiple planes and a sphere to the high SNR table-top and volleyball scenes.
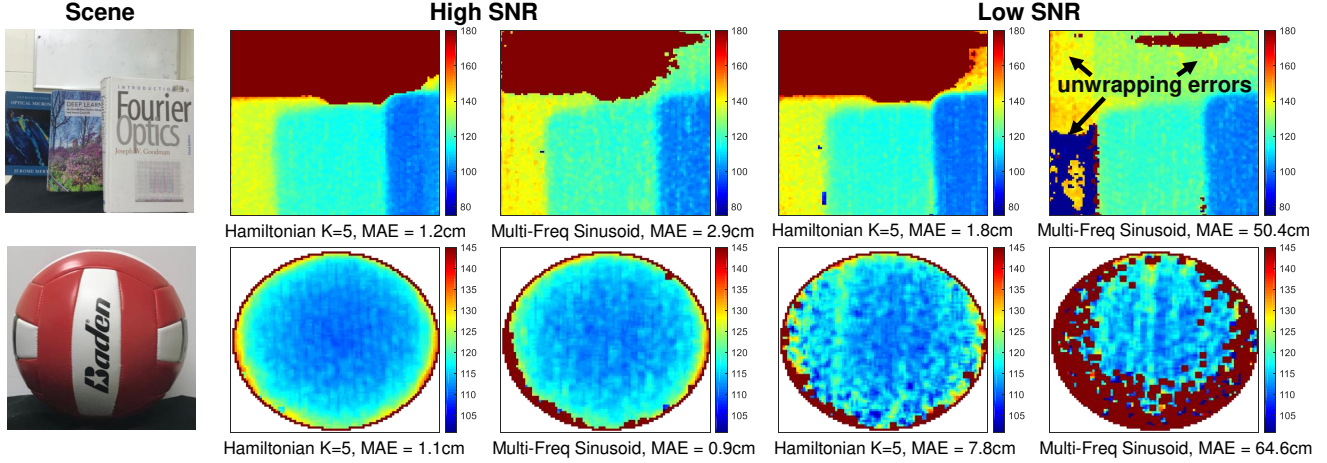
Figure 10: **Depth maps for a table-top scene and a volleyball.** We compare the reconstruction results for Hamiltonian $K = 5$ and multi-freq sinusoid. At high SNR both coding schemes accurately recover the 3D geometry in both scenes. In the books scene, at low SNR, phase unwrapping errors appear in multi-freq sinusoid, especially in regions with low albedo (left-most book) and farther depths (background wall). For the volleyball multi-freq sinusoid also displays poor depth precision along the periphery for low SNR. The dark red spots correspond to errors greater than 30cm which are indicative of a poor reconstruction of the ball. On the other hand, Hamiltonian $K = 5$ exhibit no phase unwrapping errors and a low outlier rate.
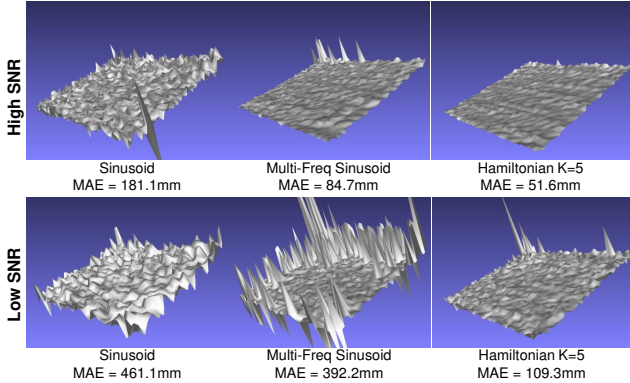

Figure 11: **3D wall reconstructions.** The wall is at a distance of 0.5m. At high SNR, Hamiltonian $K = 5$ and multi-freq sinusoid perform comparably. At low SNR phase unwrapping errors appear in multi-freq sinusoid significantly degrading its performance. The MAE is calculated with respect to the mean depth of all pixels.

Scene	P_{ave}	f_0	f_{max}	High SNR	Low SNR
Single Pixel	2mW	5Mhz	∞	O.D. 2	O.D. 3
Wall	10mW	10Mhz	$15f_0$	O.D. 2	O.D. 3
Table-top	10mW	15Mhz	$10f_0$	O.D. 1	O.D. 2
Volleyball	10mW	15Mhz	$10f_0$	O.D. 1	O.D. 2

Table 1: **Parameters for hardware experiments.** The maximum power is fixed across all experiments to $15P_{ave}$. The maximum frequency constraint (f_{max}) is determined by f_0 and the band-limit of the function generator (150Mhz). The high and low SNR settings are emulated by attenuating the source power with ND filters with different optical densities (O.D). O.D of 1, 2, and 3 lead to a signal attenuation factor of 10, 100, and 1000.

8. Limitations and Future Work

Optimality of the two-step approach: Ideally one would optimize for the practical modulation and demodulation functions directly with the MDE as the objective. Since the resulting optimization problem is challenging with no known tractable solutions, we divide it into two steps: (1) optimize the correlation functions over the MDE, (2) find practical modulation and demodulation functions that approximate the correlation. This procedure is simpler to implement, and as our results show, leads to practical high-performance codes. Although, our approach is general enough to be applied to any correlation function, it is dif-

ficult to guarantee that a given correlation can be *accurately* factorized. Ultimately, the goodness of fit depends on the strength of the constraints (e.g., Supplementary Figure 12).

Complementary demodulation for 2-bucket ToF: The 2-bucket coding schemes derived from the binary demodulation functions with our method may not be optimal. An interesting direction for future research is to embed the complementary function constraint in the optimization.

Multi-path interference: In this paper we assume there is no interference due to multiple reflections. The correlation function design step and depth inference algorithm could incorporate a multi-path propagation model as done in [1, 27]. This is a promising future research direction.

References

[1] A. Adam, C. Dann, O. Yair, S. Mazor, and S. Nowozin. Bayesian time-of-flight for realtime shape, illumination and albedo. *IEEE transactions on pattern analysis and machine intelligence*, 39(5):851–864, 2017. 1, 3, 8

[2] G. Ayers and J. C. Dainty. Iterative blind deconvolution method and its applications. *Optics letters*, 13(7):547–549, 1988. 2

[3] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004. 4

[4] B. Büttgen and P. Seitz. Robust optical time-of-flight range imaging based on smart pixel structures. *IEEE Trans. on Circuits and Systems*, 55(6):1512–1525, 2008. 2

[5] D. A. Carnegie, J. McClymont, A. P. Jongenelen, B. Drayton, A. A. Dorrington, and A. D. Payne. Design and construction of a configurable full-field range imaging system for mobile robotic applications. In *New Developments and Applications in Sensing Technology*, pages 133–155. Springer, 2011. 2

[6] B. Carrihill and R. Hummel. Experiments with the intensity ratio depth sensor. *Computer Vision, Graphics, and Image Processing*, 32(3):337 – 358, 1985. 2

[7] T. F. Chan and C.-K. Wong. Total variation blind deconvolution. *IEEE transactions on Image Processing*, 7(3):370–375, 1998. 4

[8] S. Diamond and S. Boyd. Cvxpy: A python-embedded modeling language for convex optimization. *The Journal of Machine Learning Research*, 17(1):2909–2913, 2016. 4

[9] D. Droeschel, D. Holz, and S. Behnke. Multi-frequency phase unwrapping for time-of-flight cameras. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 1463–1469. IEEE, 2010. 5

[10] W. Freeman, F. Durand, Y. Weiss, and A. Levin. Understanding and evaluating blind deconvolution algorithms. 2009. 4

[11] M. Grant, S. Boyd, and Y. Ye. Disciplined convex programming. In *Global optimization*, pages 155–210. Springer, 2006. 4

[12] R. Grootjans, W. van der Tempel, D. Van Nieuwenhove, C. de Tandt, and M. Kuijk. Improved modulation techniques for time-of-flight ranging cameras using pseudo random binary sequences. In *Eleventh Annual Symposium of the IEEE/LEOS Benelux Chapter*, page 217, 2006. 1

[13] M. Gupta and N. Nakhate. A geometric perspective on structured light coding. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 87–102, 2018. 2

[14] M. Gupta, S. K. Nayar, M. B. Hullin, and J. Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM Transactions on Graphics (ToG)*, 34(5):156, 2015. 2

[15] M. Gupta, A. Velten, S. K. Nayar, and E. Breitbach. What are optimal coding functions for time-of-flight imaging? *ACM Transactions on Graphics (TOG)*, 37(2):13, 2018. 1, 2, 3, 4

[16] M. Hansard, S. Lee, O. Choi, and R. P. Horaud. *Time-of-flight cameras: principles, methods and applications*. Springer Science & Business Media, 2012. 2

[17] F. J. Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, 66(1):51–83, 1978. 6

[18] E. Horn and N. Kiryati. Toward optimal structured light patterns1. *Image and Vision Computing*, 17(2):87–97, 1999. 2

[19] A. Kadambi, A. Bhandari, R. Whyte, A. Dorrington, and R. Raskar. Demultiplexing illumination via low cost sensing and nanosecond coding. In *2014 IEEE International Conference on Computational Photography (ICCP)*, pages 1–10. IEEE, 2014. 2

[20] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Transactions on Graphics (ToG)*, 32(6):167, 2013. 2

[21] D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE signal processing magazine*, 13(3):43–64, 1996. 2, 4

[22] R. Lange and P. Seitz. Solid-state time-of-flight range camera. *IEEE Journal of quantum electronics*, 37(3):390–397, 2001. 1, 2, 3, 5, 6

[23] P. Mirdehghan, W. Chen, and K. N. Kutulakos. Optimal structured light à la carte. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6248–6257, 2018. 2

[24] A. D. Payne, A. A. Dorrington, and M. J. Cree. Illumination waveform optimization for time-of-flight range imaging cameras. In *Videometrics, Range Imaging, and Applications XI*, volume 8085, page 80850D. International Society for Optics and Photonics, 2011. 1, 2, 3

[25] A. D. Payne, A. A. Dorrington, M. J. Cree, and D. A. Carnegie. Improved measurement linearity and precision for amcw time-of-flight range imaging cameras. *Applied optics*, 49(23):4392–4403, 2010. 2

[26] A. D. Payne, A. P. Jongenelen, A. A. Dorrington, M. J. Cree, and D. A. Carnegie. Multiple frequency range imaging to remove measurement ambiguity. In *Optical 3-D Measurement Techniques*, 2009. 2, 5

[27] M. Schober, A. Adam, O. Yair, S. Mazor, and S. Nowozin. Dynamic time-of-flight. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6109–6118, 2017. 1, 3, 8

[28] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature communications*, 3:745, 2012. 3

[29] R. Z. Whyte, A. D. Payne, A. A. Dorrington, and M. J. Cree. Multiple range imaging camera operation with minimal performance impact. In *Image Processing: Machine Vision Applications III*, volume 7538, page 75380I. International Society for Optics and Photonics, 2010. 1, 2, 6