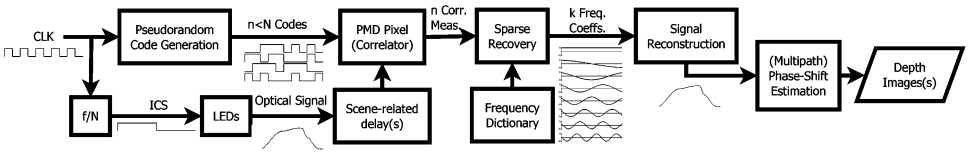
IEEE Photonics Journal
An IEEE Photonics Society Publication

Open Access

A Compressed Sensing Framework for Accurate and Robust Waveform Reconstruction and Phase Retrieval Using the Photonic Mixer Device

Volume 7, Number 3, June 2015

Miguel Heredia Conde
Klaus Hartmann
Otmar Loffeld, Senior Member, IEEE



DOI: 10.1109/JPHOT.2015.2427747
1943-0655

IEEE photonics SOCIETY

IEEE

A Compressed Sensing Framework for Accurate and Robust Waveform Reconstruction and Phase Retrieval Using the Photonic Mixer Device

Miguel Heredia Conde, Klaus Hartmann, and Otmar Loffeld, *Senior Member, IEEE*

Center for Sensorsystems (ZESS), University of Siegen, 57076 Siegen, Germany

DOI: 10.1109/JPHOT.2015.2427747

Manuscript received March 20, 2015; revised April 23, 2015; accepted April 24, 2015. Date of publication April 29, 2015; date of current version May 15, 2015. This work was supported by the German Research Foundation (DFG) as part of the research training group GRK 1564 "Imaging New Modalities."
Corresponding author: M. Heredia Conde (e-mail: heredia@zess.uni-siegen.de).

Abstract: A common way of performing phase-shift-based time-of-flight imaging combines the emission of a continuous-wave (CW) illumination signal with correlation with some reference signals at the detector array. This is the case for the well-known photonic mixer device (PMD), which correlates against displaced versions of the illumination control signal, at known phase shifts, and requires only four correlation values to estimate the phase shift. The main drawback of such approaches is that they require the assumption of nonrealistic hypothesis regarding the sensing process, leading to simple sensing models that, despite allowing fast depth estimation from few acquisitions, often ignore relevant considerations for real operation, leaving the door open for systematic errors that affect the final depth accuracy. Typical examples are ignoring the effect of the illumination devices on the final shape of the illumination signal, supposing a sinusoidal reference signal at pixel level, or not accounting for multipath effects. In this work, we present a novel framework for PMD-based signal acquisition and recovery that exploits the sparsity of CW illumination signals in the frequency domain to provide accurate reconstruction of the illumination waveforms as received by the PMD pixels. Our method is extremely robust to signal distortion and noise, since no assumption is made on the illumination signal, other than being a periodic signal. Our approach ensures that no valuable information is lost during the sensing process and allows, therefore, accurate phase shift estimation in a wider range of operation conditions, getting rid of unrealistic assumptions.

Index Terms: Compressed sensing, time-of-flight, photonic mixer device (PMD), multipath, waveform reconstruction.

1. Introduction

Depth imaging systems [1], [2] have undergone substantial development in recent years, evolving from marginal applications in specific industrial processes to commercial products, oriented to the general public and ready to operate under challenging conditions. Despite one of the milestones of this evolution being the advent of the first Microsoft Kinect sensor [3], based on structured light emission, the last developments in depth imaging adopted Time-of-Flight as the basic principle of operation for active depth estimation. Well-known examples of this trend

include the new version of the Microsoft Kinect sensor [4], often referred as Kinect two and commercialized as the Xbox One sensor [5], which widely outperforms previous sensors—as already pointed out in [6]—and has become state of the art for commercial depth cameras since it was put on the market. Other recent commercial solutions, such as the Intel Creative camera or the Hamamatsu distance sensor are also ToF imaging sensors, following the way paved by the Photonic Mixer Device (PMD) technology [7]. Since the use of *intelligent* pixels, such as those first introduced by PMD Technologies, is the basic concept behind most recent ToF depth cameras, in this paper we adopt the PMD technology as reference. In the following lines, we provide a brief introduction to this technology.

ToF cameras using the PMD chip belong to the group of phase shift-based ToF depth cameras, i.e., the depth estimation does not come from a time measurement, but from the difference of phase produced in a modulated illumination signal. A typical setup requires one or more illumination modules, an adequate optical system and the PMD chip, together with the necessary electronics for control and readout. The illumination control signal (from now on, ICS) is generated by the PMD and controls the modulation of the illumination. The illumination modules are typically NIR-LEDs or laser diodes, placed close to the camera, and the modulation frequency falls into the MHz range. The phase shift is produced by the time the light needs to travel from the emitter to the scene and back to the camera, after being reflected. The optical system focuses the incoming light on the surface of the PMD chip. For each PMD pixel, the phase of the incoming light signal encodes the depth of the corresponding scene point. The estimation of the phase shift requires correlation of the received signal with a reference signal at the pixel level. The reference signals are phase-displaced versions of the ICS, with known phase shifts. The measurements we obtain from the *correlating* pixels for each acquisition are, therefore, samples of the autocorrelation function of the ICS, where the relative phase shifts between samples are known, but the real locations are influenced by the phase shift introduced by the depth. This phase shift can be unequivocally determined from few correlation measurements, within the unambiguous range imposed by the modulation frequency. A detailed explanation on how the correlation is carried out is given in Section 2.

A common hypothesis in PMD depth estimation is to assume the correlation to be between two pure sinusoidal signals. This reduces the number of degrees of freedom (DOFs) of the estimation problem to three and allows using a closed-form algorithm that only requires four acquisitions to compute the depth (see Section 2 for details). Although solving a 3-DOF problem from only four measurements in a closed form might sound quite optimal, specially taking into account the non-linearity of the sensing process, it comes at the price of strongly unreal hypothesis. In real applications LEDs are used as low-cost illumination source. Their limited bandwidth, together with asymmetric rising and falling times produce non-negligible systematic distortions in the illumination signal. The hypothesis of sinusoidal reference signals at pixel level has been shown to be unrealistic, invalidating this simplified method for depth estimation. The easiest solution is to take profit of the systematic character of the resulting distortions and perform a *brute force* depth calibration of the sensor, using a linear translation unit to place the camera at known distances from a plane. Regardless of any consideration on the elegance of such method, it requires storing a look-up table, which might be different for each sensor and does not solve the problem itself. Additionally, it is easy to observe that such simple solution cannot cope with any unexpected signal distortion, such as those produced by multipath effects.

In this work, we present a novel framework for PMD-based illumination signal reconstruction and depth estimation that gets rid of restrictive hypothesis on the illumination waveform and exploits the possibilities of non-sinusoidal reference signals. Our approach is inspired by Compressed Sensing (CS) theory [8], [9] and takes profit of the high sparsity of CW illumination signals in frequency domain to capture all relevant information of the signal in few measurements. The minimization of information loss during acquisition allows accurate recovery of the illumination waveform, as received by each pixel. This not only allows accurate phase shift estimation, but also opens the door for multipath depth estimation, e.g., due to multiple reflections or translucent objects.

The rest of the paper is structured as follows: Section II gives an overview of the critical aspects of PMD technology and refocuses them under the light of CS. Section III presents the most relevant related work and puts it in contrast with our approach, in order to point out some fundamental differences. The core description of our method is given in Section IV. A framework for experimental evaluation of the performance of our approach and the corresponding results obtained from it are presented in Section V. In the last part of this section, we also show that our method enables multipath depth estimation, as a specific application case. Section VI concludes the paper and outlines future work.

2. PMD Revisited: Turning Disadvantages into Opportunities

As introduced in Section I, PMD pixels correlate the light signal they receive with certain reference signals and deliver the result of the correlation process, integrated over the exposure time. The first key question to understand the origin of the limitations of the PMD technology is, therefore, how the correlation measurements are performed. Despite its name containing the word *photonic*, the correlation does not occur in optical domain, but after the photons have reached the pixel surface and generated electrical charge carriers. To this end, PMD pixels feature two integration areas, i.e., there are two different channels for each pixel, that are referred in the following as channel A and channel B. Consequently, each raw data acquisition generates two images: A and B. The pixel reference signal is used to control the movement of the generated charge carriers, forwarding them to one channel well or another [10], [11]. This is possible by creating a potential gradient in the upper layer, using two transparent photogates where the push-pull voltages are applied. Ideally, the reference signals would be square and the pixel would integrate the whole incoming signal in one channel (A) during the positive half-cycle and in the other (B) during the negative half-cycle. The final value of the correlation is given by the difference between channels (A–B).

The second crucial question is how the demodulation is performed, i.e., how the depth is computed from correlation measurements. It is here where the system limitations become clear. The reference signals are displaced versions of the ICS, therefore, it is easy to see that we are performing measurements according to measurement patterns that exhibit a narrow (and similar) frequency spectrum. This is, in general, inefficient, from the information point of view and might lead to redundant measurements. The method used to compute the depth is the so-called *four phases algorithm* [7], which requires only four acquisitions, at the price of being valid exclusively for sinusoidal signals. The corresponding formula is given in (1), where c is the speed of light in vacuum, f_{mod} is the modulation frequency and $D(\theta)$ is the difference between the values of the A and B channels of the pixel, for a certain phase θ. The four samples of the cross-correlation are taken uniformly distributed, i.e., at phase shifts $\theta \in \{0, \pi/2, \pi, 3\pi/2\}$. The $\pi/2$ separation between samples ensures that the measurement waveforms are orthogonal, one with respect to the previous, since the autocorrelation function of a sinusoidal crosses zero at $\pi/2$. Unfortunately, they are fully correlated two by two

$$d = \frac{c}{4\pi f_{\text{mod}}} \arctan\left(\frac{D\left(\frac{3\pi}{2}\right) - D\left(\frac{\pi}{2}\right)}{D(\pi) - D(0)}\right). \tag{1}$$

Consider the following facts on PMD technology that are related to unmodeled signal distortions.

1) The reference signal used for correlation at pixel level is not sinusoidal, neither a perfect square signal. The shape was shown to be trapezoidal in [12], closer to square than to sinusoidal. Equivalent results were obtained in [13] using a different methodology. The nonzero rising and falling times of the trapezoidal waveform are a coherent explanation for the relatively low values of PMD demodulation contrast presented in [7].
2) LED-based illumination systems impose a limitation on the modulation frequency. Additionally, typically asymmetric rising and falling times invalidate the assumption of a quasi-sinusoidal illumination.

3) Multipath effects are present in most real environments, due to multiple reflections on the floor and neighboring walls, or simply to the presence of translucent objects.

At the first glance, these three facts seem to be disadvantages of PMD-based ToF systems. The first two lead to correlation between two non-sinusoidal signals, invalidating (1). As pointed out in [12], if the waveform of the incoming light is perfectly sinusoidal, then it is sufficient that the reference signal is periodic with the same fundamental frequency, since higher frequency harmonics will lead to a null contribution when correlating against a pure sinusoidal wave. This means that, substituting the LEDs for a laser diodes to reach a clean waveform would make the shape of the reference signal irrelevant. Nevertheless, regardless of the quality of the illumination waveform, the third fact will make (1) deliver a wrong result. Even more, it is not possible to natively deal with multipath effects with only these four phases, due to the very few information they convey.

It is necessary to reorient our focus to be in line with the basic principles of CS. Under this new perspective, we can cast the previous facts into systematic advantages for the application of CS.

1) The trapezoidal shape of the reference signals allows generating sensing waveforms with wider frequency spectrum and capturing more information about the incoming light signal.
2) CW illumination signals are extremely sparse in frequency domain and can be recovered from few measurements. Additionally, the low-pass filtering effect of the LEDs helps in this direction.
3) If the measurements contain enough information to accurately recover the illumination waveform in a sparse domain, the phase shift with respect to a reference can be determined and multiple paths might be detected and separated.

Note the crucial change in the sensing paradigm: Until now, targeted information is gathered in few measurements, that are redundant, with the hope that certain redundancy helps to reduce errors. Unfortunately, the effect of unmodeled systematic irregularities is not eliminated by redundancy and systematic errors appear. Our proposal is to maximize the gathered information by capturing non-redundant information, while keeping a low number of measurements.

3. Related Work

The basic theoretical background underlying this work and most related ones is Compressive Sensing (CS). While the basis of modern CS can be found in [8] and [9], a good compendium of the last developments and applications is given in [14]. Briefly, CS defends that accurate recovery of compressible (e.g., periodic) signals is possible from far fewer measurements than those suggested by the Shannon criterion.

A milestone in the application of compressive sensing to imaging sensors was the single pixel camera [15]. A digital micromirror device (DMD) is used as spatial light modulator (SLM), in combination with a single photodetector. The image is first focused on the DMD, which superimposes a 2-D sensing pattern. A lens focuses the light reflected by the DMD on the photodetector. Consequently, the measurements are the result of the scalar product between the (typically stochastic) sensing patterns and the image. The image can be accurately recovered from few measurements through l_1-minimization, provided that natural images are sparse in, e.g., wavelet domain. The maximum image resolution is limited by the native resolution of the DMD and, in real applications, by the allowable total exposure time, which bounds the maximum number of measurements. We share with this approach the innovative idea of decoupling the acquisition process from the posterior signal reconstruction or processing, despite we adopt the PMD pixel itself as correlator. Recent works exist that apply the idea of [15] to ToF imaging. In general, a question to be considered is whether it is worth at all to adopt the single-pixel architecture when applying CS to ToF imaging. This is a fundamental choice between the two sides of CS: on one side, achieving the same quality with a simpler measurement technology or less measurements and, on the other side, achieving better quality preserving the same measurement technology or similar number of measurements. Contrary to mainstream research, we adopt the latter focus.

A compressive depth acquisition camera framework is proposed in [16], which is further described in [17], where, a single photodetector is used, and measurements of the scene response to spatiotemporally-modulated illumination are performed. The spatial modulation requires a 2-D SLM, while the temporal modulation is achieved by means of a femtosecond laser. Additionally, a 20 GHz oscilloscope is needed to sample the scene response in time domain. This design is presented as an alternative to LIDAR or ToF cameras, but only 64×64 depth maps are recovered and no values on acquisition and recovery times are provided. Furthermore, hard restrictions on the scene geometry are imposed, such as piecewise planarity and negligible azimuth and altitude angle variations within an object, which make possible a sparse parameterization of the scene response in time domain. The restriction on angular variation, crucial for the model, poses a limitation on the opening angle of the camera or, complementary, on the ratio between size of an object and distance to the camera. In their experiments only textureless objects are considered and a generalization to textured objects would require a 2-D image of the scene. A more feasible implementation of a single pixel ToF camera is presented in [18], where a pure sinusoidal wave is used as illumination signal and a DMD allows using random binary patterns as measurement kernels in spatial domain. Simultaneous recovery of amplitude and phase of the reflected signal is achieved thanks to an appropriate complex representation of the sinusoidal light signal. The recovery method is a modified version of the model-based algorithm [19] that includes a total variation (TV) regularization term in the signal estimation step. In [20] and [21] a single photon counting detector is used as single pixel. The system and methods presented in [20] derive from those in [16], where the photodetector has been substituted by a single photon counting detector and the SLM by a DMD. The depth recovery framework also comes from [16], inheriting its intrinsic limitations, e.g., restriction to frontoparallel facets. Additionally, their approach is evaluated using a toy situation with only two depths. The system in [21] also uses pulsed laser for the illumination and a DMD to apply pseudorandom sensing kernels. The depth map recovery implies solving two l_1-minimizations, where the sparse vectors are supposed to share the same support. Surprisingly, the sparse vectors are recovered sequentially and not jointly. This requires blindly forcing the support of the second sparse signal to be the same as of the first. Additionally, a hard thresholding on the sparse coefficients is performed after l_1-minimization, followed by a least squares estimation of the overdetermined system that results from considering the non-zero dimensions. The use of one single detector comes at the price of longer acquisition times, which are in the order of seconds for toy situations (three planar objects) or minutes for natural scenes. Pulsed-laser approaches require either a photon counter or dense sampling of a photodetector output in time domain. Our approach takes profit of the state of the art ToF imaging arrays to keep the high frame rate of PMD-based ToF cameras and gets rid of the need for laser illumination or special photodetectors. Our lateral resolution is, therefore, limited by the native resolution of the PMD chip, but still higher than the typical resolutions recovered with single-pixel approaches.

Our work is closer to [22] than to single-pixel approaches. They adopt the PMD 19k-S2 sensor, similar to the 19k-S3 we use. Differently to them, our goal is not to recover time profile movies that monitor the light in motion, but to improve the depth estimation in ToF cameras. They use m-sequences as coding strategy, both for the illumination and the reference signal, achieving an autocorrelation function that approaches a delta function as the length of the code grows. We do not code the illumination signal, in order to soften the requirements on the illumination system and allow for low-cost LED solutions. In [23] sinusoidal waves of different frequencies are used as illumination signals and a sparse backscattering vector is recovered for each pixel. The method seems to be oriented to the recent Xbox One sensor and allows multipath removal in real time.

4. Methodology

As detailed in Section II, a PMD pixel performs a cross-correlation between the incoming light signal, $r(t)$, and a reference signal, $q(t)$. Consider both signals periodic, with period $T = f_{\text{mod}}^{-1}$,

being f_{mod} the fundamental frequency of the illumination signal, $i(t)$, and therefore also of $r(t)$. Then, if the integration time is large enough with respect to T, then the result of the correlation can be written as (2), shown below, in phase domain, where $\theta \in \{0, \pi/2, \pi, 3\pi/2\}$ in a conventional PMD camera. k_{int} is just a constant that depends on the integration time t_{int}, i.e., $k_{\text{int}} = t_{\text{int}} T^{-1}$. In an ideal case, $r(t)$ and $q(t)$ are sinusoidal and, consequently, also $corr_{q,r}(\theta)$

$$\text{corr}_{q,r}(\theta) = \frac{k_{\text{int}}}{2\pi} \int_0^{2\pi} q^*(\phi) r(\phi + \theta) d\phi \Leftrightarrow (q \otimes r)(\phi), \ \forall \phi \in [0, 2\pi]. \tag{2}$$

Being faithful to real applications, lets suppose that the illumination system is driven with a square signal of period T and superimposes a smoothing effect on the signal, given by the function $g_{\text{illu}}(\phi)$, which looks like a distorted Gaussian function for LEDs, then $i(\phi) = (\text{ICS} * g_{\text{illu}})(\phi)$. The scene geometry shifts the signal phase according to the depth, so its effect can be modeled as a displaced delta function. In presence of multipath effects, the scene response, $e(\phi)$, would be a finite sum of delta functions: $e(\phi) = \sum_{k=1}^{P} \alpha_k \delta(\phi - \phi_k)$, being ϕ_k the phase shift induced by the path k, α_k an attenuation factor related to the reflecting surface and P the total number of paths. Then the signal $r(\phi)$ received by a PMD pixel can be rewritten as (3), shown below, being $v_{\text{min}}^{\text{ICS}}$ and $v_{\text{max}}^{\text{ICS}}$ the levels corresponding to a logic "0" and "1" in the ICS

$$\begin{aligned} r(\phi) &= (\text{ICS} * g_{\text{illu}} * e)(\phi) \\ \text{ICS}(\phi) &= \begin{cases} v_{\text{max}}^{\text{ICS}}, & \text{if } 0 \leq \phi < \pi \\ v_{\text{min}}^{\text{ICS}}, & \text{if } \pi \leq \phi < 2\pi \end{cases} \end{aligned} \tag{3}$$

As shown in previous work [12], [13], despite the corresponding electrical signal is theoretically square, the reference signal at pixel level shows a trapezoidal shape, in practical terms, with quite symmetric rising and falling times that are close to $\pi/3$ at $f_{\text{mod}} = 20$ MHz. Therefore, we can write $q(\phi) = (q_{th} * g_{\text{pix}})(\phi)$, where $g_{\text{pix}}(\phi)$ is a step function of width $\pi/3$ in normal PMD operation (20 MHz) and $q_{th}(\phi)$ the theoretical square wave. In a conventional PMD camera, $q_{th}(\phi)$ is a displaced version of the ICS, so, for each one of the phases, we have: $q_{th}^{\theta}(\phi) = (\text{ICS} * \delta^{\theta})(\phi)$, where $\delta^{\theta}(\phi)$ is a delta function centered at $\phi = \theta$ and $\theta \in \{0, \pi/2, \pi, 3\pi/2\}$. We want to use pseudorandom (PR) binary codes as reference signals, in order to maximize the gathered information and minimize measurement redundancy, so we need to consider a different model for $q_{th}(\phi)$. The length of the codes is given by the ratio $M = f_{CLK}/f_{\text{mod}}$, being f_{CLK} the frequency of the clocking signal used to change the code value. Consequently, the total length of the code is equivalent to $T = f_{\text{mod}}^{-1}$. Since each acquisition takes $t_{\text{int}} = k_{\text{int}} * T, \ k_{\text{int}} \gg 1$, the code is recursively repeated k_{int} times during the entire acquisition time. Therefore, in our case the real reference signal is given by (4), shown below, where $i_k \in [1, M]$ is the code index, indexed in turn by the measurement index k, and $j \in [1, M]$ the element index within the code. $k = \{1, 2, \ldots, m\}$, being $m \leq M$ the number of measurements and the indexing i_k only ensures that m codes are picked randomly between the M possibilities. We also rescale the step function $g_{\text{pix}}(\phi)$ linearly according to M so that it does not impose an excessive limit in frequency, being the new width $(1/M)(\pi/3)$

$$\begin{aligned} q^k(\phi) &= (q_{th}^k * g_{\text{pix}})(\phi) \\ q_{th}^k(\phi) &= c_{i_k, j} \Leftrightarrow (j-1) \leq \frac{\phi}{2\pi} M < j. \end{aligned} \tag{4}$$

Our codes are derived from a $M \times M$ Hadamard matrix with randomly permuted columns. In this paper we consider sizes $M = 2^k, \ k \in \mathbb{N}$, which is a sufficient—but not necessary—condition for the existence of a Hadamard matrix. For arbitrary sizes, we refer to the optimum binary signature set design procedure proposed in [24]. The resulting matrix C contains the codes by

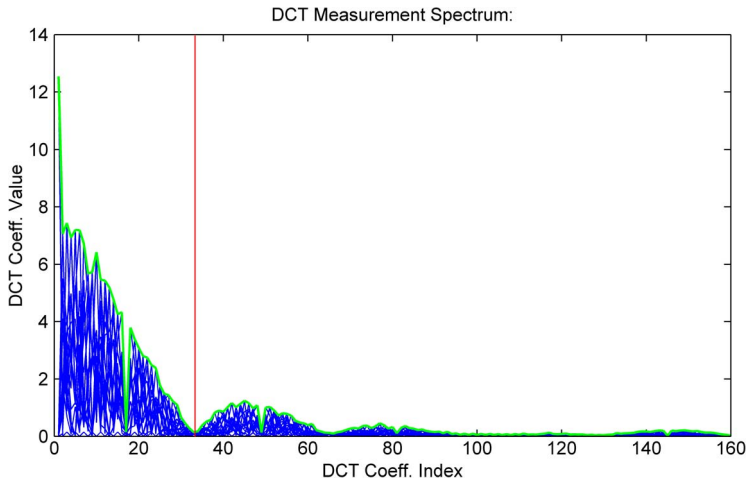


Fig. 1. Absolute values of the DCT coefficients of our measurement kernels. Rising and falling times equivalent to $\pi/3$ were imposed to PR codes of length $M = 16$. The envelope of the spectra is plotted in green. The resolution was 160 steps per period.

rows, being $c_{i,j}$ [see (4)], the element i, j of C. From a CS perspective, our measurement kernels are $q^k(\phi)$ and each measurement, y_k is the scalar product between that sensing kernel and the signal $r(\phi)$, in phase domain, which still can be seen as a sample of their cross-correlation function at the origin. Note that we keep continuous functions of ϕ in the notation, to emphasize that the measurement process within the PMD is continuous. The final number of discrete steps per period that we adopt in our model, $n \gg m$, depends on the phase resolution and waveform reconstruction quality we want to achieve. The limit on the reconstruction accuracy is given by the frequency spectra of the measurement kernels: we cannot reconstruct what we do not measure. That is, if the signal to recover, $r(\phi)$, contains frequencies that are not present in the measurement kernels, exact recovery is not feasible. Fig. 1 represents the absolute values of the frequency spectrum in DCT domain of all possible measurement kernels, $q^k(\phi)$, for $M = 16$, which was found to be large enough to reach state of the art phase resolution (approximately 0.5° in optimal conditions, e.g., 0.58° in the novel PMD-based distance sensor in [25]). The envelope of the spectra is plotted in green and resembles that of a pulse frequency spectrum, due to the pulses underlying the PR binary codes, approaching zero at the $2M = 32^{nd}$ harmonic (red line).

The sensing matrix of our CS model, $Q \in \mathbb{M}_{m\times n}$, is, therefore, defined by rows as $Q_{i,j} = q^i[\phi_j]$, being $q^i[\phi_j]$ a discrete version of the reference signal in (4), for a custom number of discretization steps per period, n. The columns of the dictionary or representation matrix define a basis—or, more generally, a tight frame—where the signal to recover, $r(\phi)$, admits a sparse representation. We exploit the periodicity of CW illumination and adopt a frequency-based dictionary, which is chosen to be the DCT basis for simplicity, but other choices such as Fourier or Walsh basis are equally valid. Discretizing to the selected, n, we want to recover an n-dimensional vector, $R|R[k] = r(\phi_k)$, in its s-sparse representation, $X \in \mathbb{R}^n | \|X\|_0 := |\mathrm{supp}(X)| \leq s, R = \Psi X$, being $\Psi \in \mathbb{M}_{n\times n}$ the DCT basis by columns, in order of increasing frequencies. Now we have the complete measurement matrix of our CS model: $A = Q\Psi$ and write the estimation of X as the convex program given by (5), shown below, where $Y \in \mathbb{R}^m$ is the vector of measurements

$$\min_{X\in\mathbb{R}^n} \|X\|_1 \text{ subject to : } AX = Y. \tag{5}$$

The primal-dual interior point method can be used to solve the convex optimization in (5), but a variety of greedy algorithms have been proposed that achieve similar performance while reducing the runtime. We adopt the Orthogonal Matching Pursuit (OMP) strategy [26], which

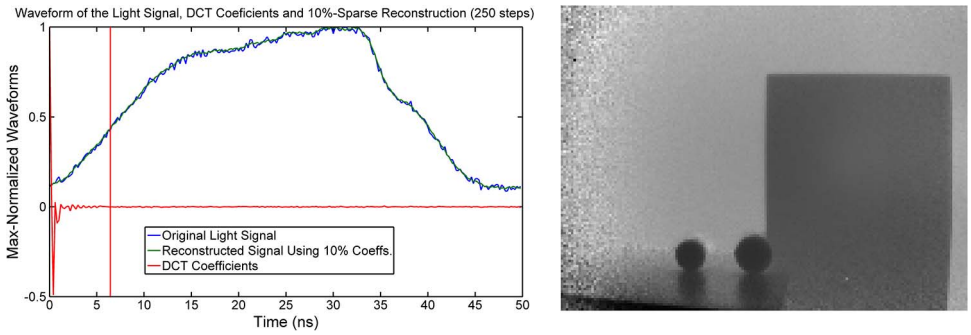


Fig. 2. (Left) Real illumination waveform from a commercial NIR-LED driven at 20 MHz, as gathered with a fast photodiode (blue), corresponding DCT coefficients (red), and reconstructed waveform using only the highest 10% DCT coefficients. In all cases considered, the sparsity of the DCT representation was $s < 32$ (red vertical line), regardless of the number of discrete samples per period n. (Right) Depth image of a natural scene used as pattern for waveform shifting during GT generation.

showed to be the best deal between simplicity and recovery accuracy. R can be reconstructed from the sparse vector of coefficients: $R = \Psi X$. The recovered signal, both in sparse or dense representation, contains all the information related to the scene structure sensed by the pixel, encoded as phase delays. Note that we have uncoupled the acquisition from the phase shift estimation and, at this point, we count with full freedom on the methodology for phase retrieval. In the simplest case, cross-correlation with the original illumination waveform at some arbitrary zero phase can be used to retrieve the phase shift, admitting a fast FPGA implementation. A general simplification of this method would be to correlate against a delta function, i.e., finding the position of the maximum, which is unique within a period. Multipath effects can be solved as a $2P$-dimensional maximization problem, with P being the number of paths.

5. Experiments and Results

Although we do not provide a hardware implementation of our approach, a complete simulation framework that implements the full methodology described in Section IV has been developed to evaluate feasibility and performance, in terms of reconstruction accuracy and runtime. In this section we give an outlook on how the simulation framework works and then present experimental results on phase estimation accuracy and runtime.

The first step of the simulation is to generate some ground truth (GT) or data to be taken as reference. In our case, the GT is a 3-D matrix containing the light waveforms received by each pixel, where the first two dimensions are the size in pixels of the PMD chip we consider and the third is n, the number of points per discretized period. We adopt the PMD 19k-S3, with a resolution of 120×160 pixels. Two different values of n are considered: 720 and 1440, leading to theoretical resolutions of 0.5° and 0.25°. The dataset is generated from a single waveform, which is scaled and displaced differently for each pixel. The signal is not synthetic, but a real acquisition of a CW illumination signal, obtained from NIR-LEDs driven with a square ICS of 20 MHz frequency. For the acquisition we use a fast photodetector together with a 2.5 Gs/s oscilloscope. The result, which can be observed in the left plot of Fig. 2, contains high frequency noise that is not filtered and preserved after resizing to $n \in \{720, 1440\}$. The plot also contains the full DCT coefficients (in red), unveiling the extreme sparsity of the signal in this domain, which can be accurately reconstructed from as few as the highest 10% DCT coefficients (in green). The per-pixel scalings and phase shifts are not random. A real PMD acquisition of a natural scene is gathered with a MultiCam [12], [27] and used as guide, i.e., the IR intensity image is used to scale the waveforms, while the phase image (see Fig. 2, right) provides the phase shifts.

In our experiments we consider three different frequency ratios, that determine the size of the PR codes: $M = (f_{CLK}/f_{\text{mod}}) \in \{8, 16, 32\}$. The higher this parameter is, the more measurements $m \leq M$ can be gathered and the wider the frequency spectrum of the sensing kernels is.

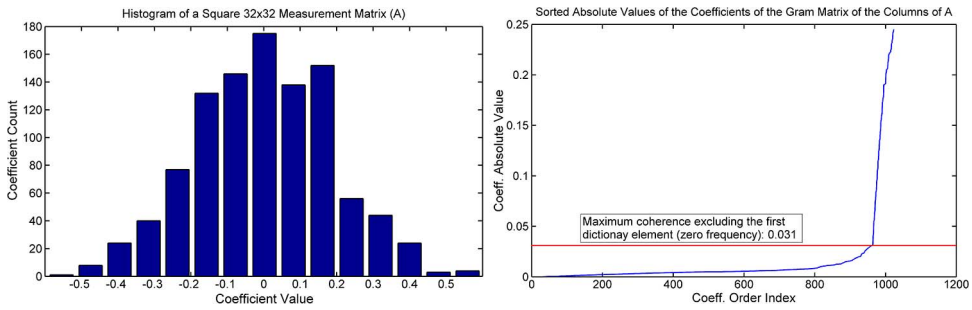


Fig. 3. (Left) Histogram of the measurement matrix coefficients, after columnwise normalization. The number of bins is 15. Note the zero-mean and the close-to-Gaussian shape. (Right) Absolute values of the coefficients of the Gram matrix generated from the normalized columns of A, sorted in increasing order. The highest values are due to the high correlation between the first measurement kernel and the first dictionary element. If they are omitted, the coherence of A boils down to 0.031 (red line).

Nevertheless, for a given f_{mod}, it is bounded by the maximum f_{CLK} of the Complex Programmable Logic Device (CPLD), responsible for the clocking signal generation in the PMD frontend. That is the reason why we choose $m = M$ for our experiments, in order to get enough information for low values of M, provided that the frequency spectrum of the sensing kernels is wide enough (Fig. 1). The sensing matrix $Q \in \mathbb{M}_{m\times n}$ is generated from a Hadamard matrix, as explained in Section IV, and used to obtain the measurements vector, Y, from the GT data. The representation matrix, Ψ, is the DCT basis by columns and, therefore, the full measurement matrix is computed as $A = Q\Psi$. Before presenting the results of our experiments, a few considerations on the recovery process are to be taken into account.

5.1 Recovery Considerations

The Hadamard-based PR sensing matrix is expected to be highly incoherent with the periodic DCT basis. Consequently, the entries of the measurement matrix, A, are expected to follow a distribution close to a zero-mean Gaussian. As a result, the coherence of A [28], [29] should be low. Consider the following limit-case scenario, in terms of coherence: $m = M = 32$ measurements are performed, being the signal size per period $n \gg m$. For simplicity, and without loss of generality, consider a partial version of the dictionary, Ψ_{T_s}, where T_s is a set indexes containing the s lowest frequencies of the dictionary. If we adopt $s = s_{max} = 32$, we obtain a square matrix $A \in \mathbb{M}_{m\times s}, m = s$. A histogram of the entries of A is given in Fig. 3 (left), which shows clear resemblance to a zero-mean Gaussian distribution. Nevertheless, a close look to our measurement matrices reveals low variance along the first row and the first column, together with a relatively high value for the first element. This is directly related to the fact that the first measurement kernel and the first dictionary element are highly correlated between each other (due to its zero frequency), while equally uncorrelated with the rest. The first measurement kernel comes from the first row of a Hadamard matrix, which is a constant vector of ones. This row is always selected in the subset of m rows, because it generates the only measurement kernel gathering information on the constant level of the signal, since all other rows of the Hadamard matrix exhibit zero mean. The direct consequence is the rise of the coherence of A, as depicted in Fig. 3 (right), where the absolute values of the coefficients of the Gram matrix, computed from the normalized columns of A, are plotted in ascendant order. Values corresponding to diagonal positions are excluded. The value of the coherence of A is given by the maximum of this graph, which is as high as 0.245. Recall that the minimum coherence of A is given by the Welch bound [30], which is zero for a square measurement matrix. Note that this phenomenon can be easily left away, since there is no coherence between the first measurement kernel and all the dictionary elements of frequency greater than zero, i.e., the value of the first coefficient of X is uniquely given by the first measurement. Therefore, we can refocus the coherence evaluation considering $s - 1 = 31$ dictionary

elements. In such case, the plot in Fig. 3 (right), would end at the red horizontal line and the coherence would be reduced to 0.031, i.e., in a 87%.

Once the coherence of A is determined, a bound on the spark is given by (6), shown below. Provided that the condition $\text{spark}(A) > 2s$ [28] is sufficient to ensure uniqueness of the solution, $X|Y = AX,\ \text{supp}(X) = s$, (6) immediately leads to the sparsity requirement in (7), shown below:

$$\text{spark}(A) \geq 1 + \frac{1}{\mu(A)} \tag{6}$$

$$s < \frac{1}{2}\left(1 + \frac{1}{\mu(A)}\right) \tag{7}$$

From (7), we obtain $s < 17$, which is almost half of our conservative limit on signal sparsity ($s \leq 32$). Therefore, we do not have any warranty on the uniqueness of the solution of (5). Nevertheless, there is an additional constrain on X that has not been taken into account yet, related to its *structured sparsity* [19]. The discrete signal to recover R can be represented in frequency domain by a coefficients vector, X, which exhibits few non-zero values. Additionally, the location of the non-zero values is influenced by the empirical fact that the absolute values of the coefficients follow a power law decay. This is not just due to the square nature of the ICL, but also due to the lowpass filtering effect of the light emitters, typically LEDs (3). Consequently, $X[k+1] < X[k]$, being the dimensions of X in order of increasing frequency of the dictionary elements. Therefore, we can use this fact to generate an oracle-given support for X, provided that the sparsity, s, is known beforehand. If the support is known, the only condition for uniqueness of $Y = AX$, is that $s \leq m$, being the problem solvable by least squares (LS). For comparison, we consider both recovery with conventional OMP, stopping when $(X) = s$, and recovery with known support, in which case OMP converges to LS.

5.2 Results

Note that the methodology and the experimental framework presented in this paper work in a per-pixel basis, i.e., the signal recovery process is performed sequentially for each one of the pixels of the PMD chip. Unless contrary indication, the effect of pixel electronics $(g_{\text{pix}}(\phi))$ was modeled to scale with f_{CLK}, i.e., its $\pi/3$ width is with respect to the *CLK* signal period and scales with it. We do not consider multipath cases at this point, in order to provide an evaluation on the best performance in absence of depth ambiguities. The phase shift—or, equivalently, depth—is calculated by finding the maximum of the discrete normalized cross-correlation between the recovered signal, R and an illumination signal at zero phase. The performance of our method is evaluated in terms of phase RMSE, using the GT natural depth image as reference. Speed is only evaluated for comparison between recovery alternatives, measuring the total time needed for recovery for all pixels.

The first experiment we perform (Exp. 1) uses pure OMP, with maximum sparsity as stopping criterion $(s = s_{\text{max}} = 32)$. The second experiment (Exp. 2) replicates the first, but supposing that the support of X is known, corresponding to the s lowest frequency DCT components. Under this assumption, OMP is equivalent to LS. For all the experiments, the obtained accuracy of the phase estimation is given in Table 1, while the runtimes of the recovery process are summarized in Table 2. Note the full coincidence between Exp. 1 and Exp. 2 in terms of phase RMSE, which is an indicator of the validity of the support oracle for the natural signal considered. On the side of the runtime, the recovery takes, on average, 11% less time in the case of known support, when compared to the unknown case, for the same sparsity.

In order to have a reference to compare the effects of the trapezoidal shape of the reference signal on the recovery, we consider one more experiment (Exp. 3), where perfectly square reference signals are considered, i.e., with theoretical infinite slope $(g_{\text{pix}}(\phi) = \delta(\phi))$. The better phase RMSE results can be observed in Table 1. Runtime results are omitted for brevity. Regarding the runtime, there exists a simple way to boost it: Provided that in case of known support, the measurement matrix A is guaranteed to be the same for all pixels of the PMD chip, we can

TABLE 1

Phase RMSE of the recovered signals with respect to GT, for different values of discrete signal sizes (n) and number of measurements (m), and for different experimental cases (Exp. #). Exp. 1 uses OMP with $s = s_{max} = 32$. Exp. 2 replicates Exp. 1 with known support corresponding to the s lowest frequencies. Exp. 3 uses perfectly square reference signals. Exp. 4 jointly recovers the signal for all pixels using a MMV framework. Exp. 5 adapts the sparsity to the number of measurements: $s = m$

RMSE (°)		Number of Measurements (m)		
Exp.#	n	8	16	32
Exp. 1	720	2.85	1.05	0.144
	1440	2.84	1.01	0.115
Exp. 2	720	2.85	1.05	0.144
	1440	2.84	1.01	0.115
Exp. 3	720	1.66	0.630	0.144
	1440	1.66	0.579	0.119
Exp. 4	720	2.85	1.05	0.144
	1440	2.84	1.01	0.115
Exp. 5	720	0.762	0.375	0.144
	1440	0.745	0.297	0.115

TABLE II

Aggregate recovery runtimes for all $120 \times 160 = 19\,200$ PMD pixels, for different values of discrete signal sizes (n) and number of measurements (m), and for different experimental cases (Exp. #). Exp. 1 uses OMP with $s = s_{max} = 32$. Exp. 2 replicates Exp. 1 with known support corresponding to the s lowest frequencies. Exp. 3 uses perfectly square reference signals. Exp. 4 jointly recovers the signal for all pixels using a MMV framework. Exp. 5 adapts the sparsity to the number of measurements: $s = m$

Runtime (s)		Number of Measurements (m)		
Exp.#	n	8	16	32
Exp. 1	720	121	191	262
	1440	198	239	270
Exp. 2	720	113	139	262
	1440	144	172	337
Exp. 4	720	0.138	0.148	0.141
	1440	0.326	0.676	0.386
Exp. 5	720	0.131	0.179	0.141
	1440	0.405	0.397	0.386

recast the sequential single measurement vector (SMV) recoveries into one single multiple measurement vector (MMV) recovery. A fourth experiment (Exp. 4) is carried out to characterize the runtime improvement achieved using a known-support MMV recovery framework. The results show an average runtime reduction of 99.86% with respect to Exp. 1—i.e., three orders of magnitude—while achieving exactly the same phase RMSE values.

From the comparison of RMSE results for Exp. 3 and the rest, it is easy to observe an excessive gap for $m = 8$, which is related to the restricted frequency spectrum of the measurement kernels, unable to capture non-zero high frequency harmonics. For clarity, imagine a measurement spectrum such as that of Fig. 1, where the horizontal axis has been scaled by a factor $1/2(M = 8 = (1/2)16)$. The red line would appear at the 16th harmonic and no further frequencies would be properly acquired. In order to avoid unnecessarily high errors due to wrong recovery of high-frequency harmonics that cannot be properly measured, we carry out one last experiment (Exp. 5), where the maximum sparsity of the signal to recover is bounded by the number of measurements, i.e., $s = m$. This forces A to be square and the recovery problem to be determined. The last column of Tables 1 and 2 for Exp. 5 was taken from Exp. 4, since they are equivalent cases. The results show a dramatic reduction of the phase RMSE for the case $m = 8$ (74%) and $m = 16$ (between 65 and 71%).

The best results, both in terms of RMSE and runtime are obtained from Exp. 5, which uses a sparsity adaptation and takes profit of a MMV recovery strategy. RMSEs lower than 1° are

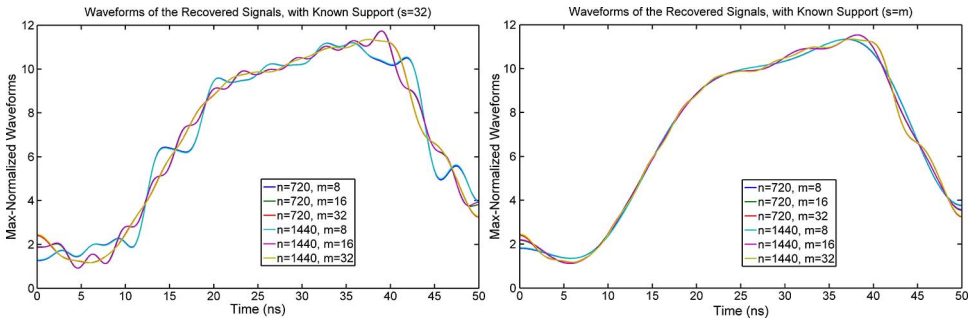


Fig. 4. Recovered waveforms for the pixel in column 45 of the last row (Fig. 2, right), which exhibits the lowest phase shift (52°). This allows direct comparison to the reference signal in Fig. 2 (left). The left plot shows the results of the MMV recovery using known support of fixed sparsity $s = 32$ (Exp. 4), while the right one shows the results obtained using the same methodology but with known support of adaptive size $s = m$ (Exp. 5). Note that the waveforms for discrete signal size $n = 720$ are behind those for $n = 1440$, showing that n has not a visible effect on the quality of the recovery, where the number of measurements m is the crucial parameter.

achieved in all cases considered, being as low as 0.12° for $m = 32$ and $n = 1440$, much lower than the 0.5° phase resolution limit characteristic of PMD-based ToF cameras. In order to allow for visual evaluation of the quality of the recovered signals, Fig. 4 shows the recovered discrete signal, R, in Exp. 4 (left) and Exp. 5 (right), for the pixel exhibiting lowest phase shift.

Note the high errors for $m = 8$ in the left plot of Fig. 4, produced by wrong estimation of high frequency DCT coefficients, and compare it to the corresponding waveform in Fig. 4, where the contribution of those high frequency components is neglected in order to keep a balance between sensed and recovered information. Consequently, waveforms corresponding to lower m in the right plot are less faithful to the real signal than those for higher m. That is, there is a compromise between number of measurements, m,—i.e., exposure time—and quality of the reconstruction. The final selection of the ratio of frequencies M and m depends on the application.

5.3 Multipath Estimation

It is out of the scope of this paper to provide an efficient multipath compensation method. Nevertheless, in this section we present the results of a simulation including multipath effects, in order to show that our method can include an additional step of multipath estimation from the recovered waveforms. The full knowledge of the waveforms ensures that multipath recovery is feasible, as far as the recovered waveform is an unambiguous combination of rescaled and displaced versions of the emitted illumination waveform. We only consider phase shift due to depth and attenuation due to different object reflectivities, but further signal distortion can be considered, at the cost of additional parameters per path. Consequently, for each one of the P paths, there exist two unknowns: a phase shift, ϕ_k, and an attenuation factor, α_k, $k \in [1, P]$. The unambiguity requirement for the paths composition can be expressed as (8), shown below, where ε_ϕ is dependent on the method used to solve the multipath estimation and, in a discrete implementation, is lower bounded by the dimensionality of the signal, n: $\varepsilon_\phi \geq (2\pi/n)$. For high values of n, ε_ϕ depends exclusively from the quality of the signal recovery and, therefore, the phase RMSE of the recovered monopath signals could be used as a lower bound.

$$|\phi_i - \phi_j| \geq \varepsilon_\phi, \ \forall i, j \in [1, P], i \neq j \tag{8}$$

For notational simplicity, we group the path parameters in two P-dimensional vectors, $\Phi = [\phi_1, \phi_2, \ldots, \phi_P]$ and $A = [\alpha_1, \alpha_2, \ldots, \alpha_P]$. Adopting the scene response function, $e(\phi)$, as defined in Section IV, we cast the estimation of the $2P$ unknowns into the minimization problem

presented in (9), shown below. In other words, we seek to find the values of Φ and A that minimize the RMSE between the real waveform received at the pixel, $r(\phi)$, and our estimation from multipath composition, $r'(\phi)$, which depends on the parameter vectors Φ and A

$$(\Phi, A) = \underset{\substack{\phi_k \in [0,2\pi] \\ \alpha_k \in [0,1]}}{\operatorname{argmin}} \sqrt{\frac{1}{2\pi}\int_0^{2\pi} (r(\phi) - r'(\phi))^2}$$
$$r'(\phi) = i(\phi) * e_{\Phi,A}(\phi). \tag{9}$$

For the sake of generality, we adopted the simplex method [31] to solve (9), but custom minimization techniques could be used instead. In order to generate the multipath GT, the PMD phase and intensity data of the scene in Fig. 2 (right) was used to delay and scale the original illumination signal plotted at the left, as already described at the beginning of Section V. The peculiarity is that in this case we also model a fictitious transparent panel, which is supposed to have a reflective coating toward the camera. The parameters of the panel are the phase shift, given by its distance to the camera, the size, which in combination with the previous and the camera parameters can be expressed in pixels, the reflectivity and the transmittivity. The chosen values are: $\phi_{\text{panel}} = 0.84$ rad (equivalent to 1 m at 20 MHz), 26 $\times$ 26 pixels, $\rho_{\text{panel}} = 0.5$, $\tau_{\text{panel}} = 1 - \rho_{\text{panel}} = 0.5$. We assume that the panel does not absorb any light. As far as $\rho_{\text{panel}} \geq 0$, we have at least $P = 2$ paths in the panel area: one from the panel reflection, plus another from the scene reflection of the light that traversed the panel. For simplicity and to avoid modeling a third path due to second reflection on the panel of the light already reflected by the scene behind it, we assume that all the light reflected by the scene is transferred through the panel back to the camera without attenuation. Consequently, for pixels within the area covered by the panel, the initial GT waveforms are further rescaled according to τ_{panel} and summed to the waveform coming from the panel reflection, which is, in turn, obtained from the empirical waveform (Fig. 2, left), after the phase shift ϕ_{panel} and the scaling given by ρ_{panel}.

Once the multipath-affected GT waveforms are generated, we proceed as in the monopath case and recover them from few PR measurements. We use the recovery setup of Exp. 5, which performs best in terms of accuracy and runtime, as shown in Section V-B. We use $m = s = 32$ measurements and consider $n \in \{720, 1440\}$. The recovered monopath waveforms (R) will be used as the discrete version of $r(\phi)$ in (9). Most minimization methods, including simplex, need an initial guess of the parameters to estimate. We use the monopath phase shift estimation (first column images in Fig. 5) as initial phase shift estimation for all paths and the RMS value of the monopath waveform to compute the initial guess on the intensity as $\alpha_k^{(0)} = (\text{rms}(R)/P), \forall k \in [1, P]$. Note that not all pixels are affected by multipath, while the presented approach estimates P path parameters for every pixel. In order to detect the monopath cases, we implement a simple threshold on the standard deviation of the vector of multipath phases, A. If the threshold is not achieved, the estimated multipath parameters are discarded and the initial monopath estimation is adopted. For cases of $P \geq 2$, a threshold on some distance between elements of Φ can be used to detect how many paths are active for each pixel, and combine those that are repeated.

The results of the multipath recovery are shown in Fig. 5. The first column shows the monopath recovered phase image, while the next two columns show the phase images estimated for the two paths present in our experimental setup: the first one for the scene behind the panel and the second for the panel itself. Note that the monopath estimation is automatically adopted for those pixels exhibiting very low standard deviation of the vector A (pixels outside the panel area). Two experimental cases are considered, to test the influence of the discretization severity on the phase estimation quality: $n = 720$ (first row of Fig. 5) and $n = 1440$ (second row of Fig. 5). In visual terms, the path separation seems to be perfect for both experimental cases. We compare the phase images obtained for both paths against ground truth in terms or phase RMSE, for each experimental case. Note that in the case of the first path (scene behind the

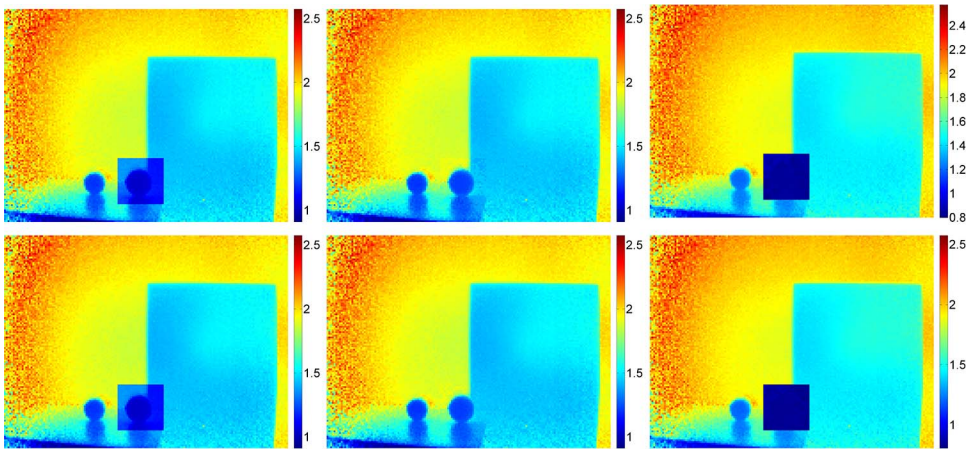


Fig. 5. Phase images recovered for the multipath simulation, in radians. The first row shows the results obtained for a discrete signal size $n = 720$ and the second for $n = 1440$. In both cases $m = s = 32$ measurements were used. The first column is the phase image obtained from the recovered waveforms under monopath assumption, using the method detailed in Section IV. Note the wrong phase estimation in the area of the semitransparent panel, where the recovered values are between those corresponding to the panel itself and those corresponding to the scene behind it. The next two columns show the recovered phase shift images supposing multipath of $P = 2$ and using the multipath recovery method presented in the current section. The images in the central column recover the path corresponding to the scene behind the panel, while the images in the third column accurately recover the path corresponding to the panel, reconstructing its plain surface.

TABLE 3

Recovery performance for the multipath case with $P = 2$ paths. Phase RMSE and aggregate recovery runtimes for all $120 \times 160 = 19\,200$ PMD pixels, for values of discrete signal sizes $n \in \{720, 1440\}$. The path corresponding to the background scene is denoted by SCN, while the path corresponding to the foreground panel is denoted by PNL

n	720		1440	
Path ID	SCN	PNL	SCN	PNL
Phase RMSE (°)	0.252	1.11	0.175	0.834
Runtime (s)	0.232		0.497	

panel) the error is computed all along the image, while for the second path (semitransparent panel), it is computed only for the area where this second path exists (the panel itself). The performance of our multipath estimation in terms of phase RMSE and waveform recovery runtime is given in Table 3.

6. Conclusion

A novel sensing framework for the Photonic Mixer Device (PMD), inspired by the Compressed Sensing (CS) theory, has been presented. Our approach is intended to fit to any existing PMD-based ToF camera setup and it does not impose any requirement on the illumination system, nor it requires any additional hardware development. PMD pixels are used to apply pseudorandom (PR) codes on the incoming light signal in time domain and the resulting scalar products are used as measurements. Provided that CW illumination signals are extremely sparse in frequency domain, the light signal received at each pixel can be recovered from few measurements. As a result, the phase shift estimation is fully decoupled from the measurement process and all the information contained by the received waveform is available for phase estimation. This allows more freedom in the phase retrieval process, which might account for multiple depths per pixel, in contrast to simple but rigid approaches, such as the four phases algorithm.

Knowledge of the structured sparsity of the signal to recover allows us to tune and simplify the recovery process, obtaining better phase accuracy when compared to pure OMP recovery, while reducing the runtime in three orders of magnitude. Phase RMSEs of 0.30° and 0.12° are achieved using only $m = 16$ and 32 measurements, respectively. A multiple measurement vectors (MMV) recovery schema allows us reaching joint recovery times as low as 0.14 s, for all $120 \times 160 = 19\,200$ PMD pixels of a 19k-S3, in a conventional desktop computer.

The only requirement to implement our approach in an existing PMD frontend is a modification of the code in the CPLD so that it generates the custom PR codes. Additionally, due to the reduced runtime, a real-time implementation of the recovery framework in the FPGA of the camera is feasible. These hardware modifications and the evaluation of a CS-PMD camera based on the methods presented here are subjects of future work.

References

[1] B. Langmann, K. Hartmann, and O. Loffeld, "Depth camera technology comparison and performance evaluation," in *Proc. ICPRAM (2)*, 2012, pp. 438–444.

[2] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (ToF) cameras: A survey," *IEEE Sens. J.*, vol. 11, no. 9, pp. 1917–1926, Sep. 2011.

[3] K. Khoshelham and S. O. Elberink, "Accuracy and resolution of kinect depth data for indoor mapping applications," *Sensors*, vol. 12, no. 2, pp. 1437–1454, 2012.

[4] A. Payne *et al.,* "7.6 a 512 × 424 CMOS 3D time-of-flight image sensor with multi-frequency photo-demodulation up to 130 MHz and 2 GS/s ADC," in *Proc. ISSCC*, Feb. 2014, pp. 134–135.

[5] Microsoft, "Xbox one sensor," 2014. [Online]. Available: http://www.xbox.com/en-US/xbox-one/innovation

[6] M. Heredia Conde, K. Hartmann, and O. Loffeld, "Adaptive high dynamic range for time-of-flight cameras," *IEEE Trans. Instrum. Meas.*, 2015, to be published.

[7] T. Möller, H. Kraft, J. Frey, M. Albrecht, and R. Lange, "Robust 3D measurement with PMD sensors," in *Proc. 1st Range Imaging Res. Day ETH*, 2005, pp. 3–906 467.

[8] R. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.

[9] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[10] M. Albrecht, "Untersuchung von photogate-PMD-sensoren hinsichtlich qualifizierender charakterisierungsparameter und -methoden," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci. Univ. Siegen, Siegen, Germany, 2007. [Online]. Available: http://dokumentix.ub.uni-siegen.de/opus/volltexte/2007/297/pdf/albrecht.pdf

[11] R. Lange, "3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci. Univ. Siegen, Siegen, Germany, 2000. [Online]. Available: http://dokumentix.ub.uni-siegen.de/opus/volltexte/2006/178/pdf/lange.pdf

[12] O. Lottner, *Investigations of Optical 2D/3D-Imaging With Different Sensors and Illumination Configurations*. ZESS-Forschungsberichte. Aachen, Germany: Shaker Verlag, 2011, vol. 29. [Online]. Available: http://www.shaker.de/de/content/catalogue/index.asp?lang=de&ID=6&category=102

[13] M. Heredia Conde, K. Hartmann, and O. Loffeld, "Subpixel spatial response of PMD pixels," in *Proc. IEEE Int. Conf. IST*, Oct. 2014, pp. 297–302.

[14] Y. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, ser. Compressed Sensing: Theory and Applications. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[15] M. F. Duarte, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, Mar. 2008.

[16] A. Kirmani, A. Colaço, F. N. C. Wong, and V. K. Goyal, "CoDAC: A compressive depth acquisition camera framework" in *Proc. IEEE ICASSP*, 2012, pp. 5425–5428. [Online]. Available: http://dblp.uni-trier.de/db/conf/icassp/icassp2012.html#KirmaniCWG12

[17] A. Kirmani, A. Colaço, F. N. C. Wong, and V. K. Goyal, "Exploiting sparsity in time-of-flight range acquisition using a single time-resolved sensor," *Opt. Exp.*, vol. 19, no. 22, pp. 21 485–21 507, Oct 2011. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-19-22-21485

[18] T. Edeler, K. Ohliger, S. Hussmann, and A. Mertins, "Super resolution reconstruction method for time-of-flight range data using complex compressive sensing," in *Proc. IEEE I2MTC*, May 2011, pp. 1–5.

[19] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.

[20] A. Colaço, A. Kirmani, G. Howland, J. Howell, and V. Goyal, "Compressive depth map acquisition using a single photon-counting detector: Parametric signal processing meets sparsity," in *Proc. IEEE Conf. CVPR*, Jun. 2012, pp. 96–102.

[21] G. A. Howland, D. J. Lum, M. R. Ware, and J. C. Howell, "Photon counting compressive depth mapping," *Opt. Exp.*, vol. 21, no. 20, pp. 23 822–23 837, 2013.

[22] A. Kadambi *et al.,* "Coded time of flight cameras: Sparse deconvolution to address multipath interference and recover time profiles," *ACM Trans. Graph. (TOG)*, vol. 32, no. 6, p. 167, Nov. 2013.

[23] D. Freedman, E. Krupka, Y. Smolin, I. Leichter, and M. Schmidt, "SRA: Fast removal of general multipath for ToF sensors," *Comput. Vis.*, vol. 8689, pp. 234–249, 2014.

[24] G. N. Karystinos and D. A. Pados, "New bounds on the total squared correlation and optimum design of DS-CDMA binary signature sets," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 48–51, Jan. 2003.
[25] IFM, "*IFM o1d100*," 2015. [Online]. Available: https://www.ifm.com/products/gb/ds/O1D100.htm
[26] Y. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Asilomar Conf. Signals, Syst. Comput.*, Nov. 1993, vol. 1, pp. 40–44.
[27] K. Hartmann and R. Schwarte, "Detection of the phase and amplitude of electromagnetic waves," Patent US7 238 927 B1, Jul. 3, 2007. [Online]. Available: https://www.google.com/patents/US7238927
[28] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via 1 minimization," *Proc. Nat. Academy Sci.*, vol. 100, no. 5, pp. 2197–2202, 2003. [Online]. Available: http://www.pnas.org/content/100/5/2197.abstract
[29] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[30] L. Welch, "Lower bounds on the maximum cross correlation of signals (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 3, pp. 397–399, May 1974.
[31] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder–Mead simplex method in low dimensions," *SIAM J. Opt.*, vol. 9, no. 1, pp. 112–147, 1998.