2018 5th International Workshop on Compressed Sensing applied to Radar, Multimodal Sensing, and Imaging (CoSeRa)

Fast Multipath Estimation for PMD Sensors

Miguel Heredia Conde*, Thomas Kerstein†, Bernd Buxbaum† and Otmar Loffeld*
*Center for Sensorsystems (ZESS)
University of Siegen, Paul-Bonatz-Straße 9-11
57076 Siegen, Germany
E-mail: {heredia, loffeld}@zess.uni-siegen.de
†pmdtechnologies ag
Am Eichenhang 50
57076 Siegen, Germany
E-mail: {T.Kerstein, B.Buxbaum}@pmdtec.com

***Abstract*—Time-of-Flight is an active depth imaging modality based on the emission of modulated light and the reception of the corresponding bounce. A common modulation scheme for ToF systems operating in Amplitude Modulated Continuous Wave (AMCW) mode, such as those based on the Photonic Mixer Device (PMD), is using periodic signals that are considered close-to-sinusoidal. Unfortunately, the non-sinusoidality of the signals yields a distortion in the measurements that has a systematic effect in the phase estimation. Furthermore, such systems expect a single bounce per ToF pixel and the phase estimation fails when two or more returning paths are present, as it is the case in many real-life scenes.**

In this work we study up to which extent a simple spectral estimation method can be used to accurately resolve more than a single path per pixel using a reduced set of measurements at different frequencies. We also study the effect of including a harmonic cancellation (HC) technique on the feasibility of such method. Simulation results show that HC has a great effect in this regard and that using an HC scheme at sensing allows for multiple path separation from a number of frequency measurements that tightly approaches the theoretical lower bound. Path separation is shown to be possible in a real system with HC for phase separation as low as $2\pi/2^7\,\mathrm{rad}$ at $10\,\mathrm{MHz}$ base frequency. Simulations in presence of realistic measurement noise witness that the method can successfully separates two paths in $80-90\%$ of the cases using only 5 nonzero frequencies.

I. Introduction

Time-of-Flight imaging is a depth imaging technique whose global impact has been continuously growing in the last years. This is probably both due to the growth of the number of pixels in the array and due to a dramatic cost reduction. Nowadays stand-alone ToF depth imaging systems, such as the Kinect sensor (the so-called second-generation Kinect, Kinect v2, or Xbox One sensor [1]) and the CamBoard pico flexx [2] from pmdtechnologies, came to the market with prices in the 200-300$ range and offer relatively large resolutions (512×424 and 224×171, respectively) for this type of sensors. Furthermore, the very low size of the ToF sensor in [2] promotes a spread of the technology to further application areas.

In very short terms, ToF depth imaging systems are active systems that emit light to the scene to sense and estimate the distance from the sensor to each scene point from the time the corresponding echo needs to reach the camera. *Sensu stricto*, ToF sensors measure the time the light needs to travel from the illumination system to the scene and from the scene to the camera. Unfortunately, due to the large magnitude of the speed of light, fine depth resolution translates into very demanding time resolution. For instance, $1\,\mathrm{mm}$ translates into $6.67\,\mathrm{ps}$ resolution in a reflective setup. Building large arrays of photodetectors with picosecond resolution is a challenging task and the common workaround is periodically modulating the amplitude of the emitted light and measuring *phase shift* instead of time. In this work we focus exclusively on ToF systems operating in Amplitude Modulated Continuous Wave (AMCW) mode, such as those based on the Photonic Mixer Device (PMD) [3] and related technologies.

Despite the considerable increase in lateral resolution and depth accuracy of recent ToF sensors w.r.t. previous generations, several open issues still remain to be addressed. In PMD sensors, like in most ToF sensors operating in AMCW mode, the shape of both the illumination signal (optical) and the effective PMD pixel control signal are supposed to be *close to sinusoidal*. Unfortunately, it has been shown that none of them is purely sinusoidal in practice. The illumination signal exhibits a shape that depends on the specific rising and falling characteristics of the light emitters and the effective PMD pixel control signal has been shown to be closer to trapezoidal [4], [5]. This translates into the so-called *wiggling effect* in the measurements, due to the existence of harmonics in the cross-correlation between illumination signal and (effective) PMD pixel control signal. As pointed out in [4], it would be sufficient that any of both signals is purely sinusoidal to avoid the distortion due to the harmonics present in the other.

Another main limit of most ToF systems is their intrinsic inability to resolve more than one path per pixel, that is, the system expects a single bounce of the illumination signal per pixel. If two or more optical signals that are not in phase, i. e., arising from different reflections, reach the ToF pixel simultaneously, the so-called *multipath interference* (MPI) occurs. Clearly, single-frequency AMCW ToF systems relying on sinusoidal modulation are unable to cope with this phenomenon, provided that the sum of two or more interfering sinusoidal signals of equal frequency results in another sinusoidal of the same frequency and *a posteriori* phase separation becomes impossible. Multipath can be classified between *diffuse* and *reflective*. The first occurs due to light scattering, e. g., due to Lambertian-reflective objects that are too close to the camera, in combination with low-quality optics. Scattering media, such as turbid water or translucent objects, may also produce diffuse multipath.

In this work we focus on the second type, which arises from strong secondary reflections, caused by a shiny floor or walls, which produce a secondary illumination front for the rest of the scene. Methods with different degrees of complexity exist that aim to compensate the MPI of each nature. Unfortunately, these methods require measurements at different frequencies, with a bandwidth that should be as large as possible. In this scenario, the effect of harmonic distortion becomes even more relevant and limits the applicability of MPI compensation techniques in practice.

In this work we show how, taking profit of the capabilities of a recent PMD ToF sensor prototype, such as native harmonic cancellation (HC), one can use classical spectral estimation methods to resolve more than a single path per pixel. More specifically, we evaluate the performance of the so-called matrix pencil method for estimating the reflectance of two or more interfering targets per pixel and their corresponding distances to the camera from very few frequency measurements, for different number of frequencies. We also study the effect of measurement noise. Realistic simulation results show the feasibility of resolving more than a single path per pixel, with an accuracy that is boosted if HC is enabled at the sensor.

II. Related Work

A. Harmonic Cancellation

There are two fundamentally different ways of eliminating the effect that the harmonic content present in the cross-correlation between received light signal and control signal produces in the measurements. This harmonic contents is due to both signals not being perfectly sinusoidal and, as commented in Section I, results in a *wiggling effect* in the phase (thus depth) estimation. The two ways of compensating this wiggling differ from each other in whether they are an *a posteriori* wiggling compensation or an actual harmonic cancellation (HC) during acquisition.

The first option is conceptually the simplest one, since it is nothing else than a system calibration in phase domain. The naïvest approach is constructing a look-up table (LUT) establishing correspondences between measured distance and real distance [6], [7]. The true depth can be calculated from the measured depth by interpolation. This method requires both a large calibration data set and an amount of memory that grows with the desired calibration quality. A more elegant solution is that provided by [8], which takes profit of the oscillating shape of the depth error curves to fit a cubic B-spline to it. Retrieving the real depth only requires evaluating a cubic polynomial.

The second option can be implemented either by acquiring more measurements with different phase delays of, e. g., the illumination signal, or by bracketing the exposure time and using different phase delays in the different time slots. Both the exact values of the delays and the duration of the corresponding time slot need to be accurately adjusted, so that the targeted harmonics get canceled when aggregating the signal. The first flavor is easier to implement, since it only requires performing additional measurements and can be readily implemented in any system with phase shifting capabilities. The order up to which harmonics are canceled relates directly to the number of cross-correlation samples acquired. The second flavor, proposed in [9], [10], requires specific hardware capabilities for custom bracketing of the exposure time and custom phase shifting, different at each bracket. As before, the order of the HC is given directly by the number of phase shifts (brackets) applied during the exposure time, more specifically, canceling harmonics up to order n requires $2n + 1$ shifts: one of 0° and the rest equidistant and symmetrically distributed at both sides of 0°. Our novel PMD ToF sensor prototype natively enables this second flavor of HC at acquisition and thus this is the HC option we contemplate in this paper.

B. Multipath Compensation

It is out of the scope of this paper providing a literature review on multipath estimation and compensation. We restrict our attention exclusively to *reflective* multipath, arising from translucent objects and shiny surfaces, such as floors, mirrors, whiteboards, walls, windows, tables, metallic surfaces, etc. Note that these objects are ubiquitous in human-made environments. Similarly to HC, in the case of MPI one can either aim to compensate its effect *a posteriori* after (single) depth estimation or to directly resolve several paths per pixel. The former is a challenging computational problem, while the latter shifts the pressure to the sensor and the sensing scheme, since it necessarily requires measurements at different frequencies. Most multipath-removal methods of the first type aim to jointly estimate the scene structure and the multiple paths followed by the illumination beam, normally in an iterative fashion.

In this work we are interested in estimating several paths per pixel, thus eliminating the need for a subsequent computationally-expensive multipath-removal procedure. The method in [11] and the first method in [12] provide closed-form solutions for the two-path case. Alternatively, in [13] the multiple paths are retrieved by means of an optimization process. The methods in [13], [12], and [11] cope with only two paths per pixel and require two, three and five frequencies, respectively. If more than two bounces interfere within the same pixel, none of the previous approaches apply. Note that if the phase or depth domain is finely discretized, the few targets producing the MPI can be modeled as a *sparse* vector of reflectances and the multipath estimation problem can be attacked from a *compressive sensing* (CS) perspective [14]–[16]. This is the focus adopted in [17], where this sparse vector is recovered from partial Fourier measurements in a classical CS framework. Unfortunately, the method seems to require a large number of modulation frequencies in practice (77 for *known* sparsity of 3). A more feasible approach is provided in [18], where k interfering paths are estimated from $2k+1$ frequency measurements in a closed-form manner. In an evaluation with real data from an Xbox One sensor this method required 21 measurements to separate only two paths. The frequency-domain framework in [19] allows separating multiple paths at the cost of an unaffordable modulation bandwidth of the illumination system (e. g., $10\,\mathrm{GHz}$ for $3.6\,\mathrm{cm}$ depth resolution).

One can also profit from existing transient imaging and light transport approaches to solve the MPI in ToF imaging systems. For further literature in this regard we refer to the multipath part of the paragraph on “Interference Between Optical Signals” in [20, Section 2.4].

III. Problem Statement

In this work we focus on PMD (or similar) systems operating in AMCW mode with quasi-sinusoidal signals and we restrict our attention to reflective MPI. The light signal reaching a PMD pixel is supposed to be the superposition of *few* bounces and can be, therefore, represented as a finite sum of shifted and damped versions of the periodic illumination signal. The corresponding phase-domain *scene response function* reads:

$$e(\phi) = \sum_{k=1}^{P} a_k \delta(\phi - \phi_k) \tag{1}$$

where $\delta(\phi - \phi_k)$ denotes a Dirac delta function centered at ϕ_k, which is the phase shift that the illumination signal undergoes when following the path $k \in [1, P]$, and a_k is the attenuation factor due to non-unit reflectance of the surface point that produced that reflection. The number of paths P is assumed to be low. Making use of Eq. (1), the periodic light signal received by a PMD pixel can be expressed as:

$$r(\phi) = (i * e)(\phi) \tag{2}$$

where $i(\phi)$ denotes the real periodic illumination signal, in phase domain. At the PMD pixel $r(\phi)$ is cross-correlated with a (close-to-trapezoidal) control signal, yielding the measurements:

$$\begin{aligned} m(\phi) = p_{\mathrm{A-B}} \otimes r(\phi) &= p_{\mathrm{A-B}} \otimes (i * e)(\phi) \\ &= (i \otimes p_{\mathrm{A-B}}) * e(\phi) \end{aligned} \tag{3}$$

where $p_{\mathrm{A-B}}(\phi)$ denotes the PMD (A-B) effective control function. This notation relates to the PMD operation principle itself, since the so-called PMD *demodulation* process actually means that the photogenerated carriers are shifted to one of two integration areas, typically named A and B, being the selection controlled by the complementary (ideally) binary signals p_{A} and p_{B}. Provided that the measurements are differential (i. e., A-B), the effective PMD control signal is $p_{\mathrm{A-B}} = p_{\mathrm{A}} - p_{\mathrm{B}}$, up to scaling factors. If we now substitute Eq. (1) in Eq. (3), we obtain

$$m(\phi) = \sum_{k=1}^{P} a_k (i \otimes p_{\mathrm{A-B}})(\phi - \phi_k). \tag{4}$$

As pointed out before, if either $i(\phi)$ or $p_{\mathrm{A-B}}(\phi)$ (or both) are sinusoidal, then the cross-correlation $i \otimes p_{\mathrm{A-B}}(\phi)$ is also sinusoidal. If this is not the case and there is a harmonic overlap between them, harmonic content will be present in $i \otimes p_{\mathrm{A-B}}$, yielding the known wiggling effect in the measurements $m(\phi)$. We suppose that the signals are close enough to sinusoidal, so that we can use this hypothesis to justify the use of spectral estimation techniques. Provided that $p_{\mathrm{A-B}}$ is a zero-mean signal, the cross-correlation exhibits no offset and there are only two free parameters, namely, amplitude and phase. Consequently, two (real) measurements $m(\phi)$ per frequency are sufficient. Note that two measurements with $\pi/2$ phase shift yield Fourier coefficients of the scene response:

$$\begin{aligned} x(j) &= m_j(0) + i m_j(\pi/2) \\ &= \sum_{k=1}^{P} a_k (\cos(-j\phi_k) - i \sin(-j\phi_k)) = \sum_{k=1}^{P} a_k e^{ij\phi_k}, \end{aligned} \tag{5}$$

where i denotes the imaginary unit and j the frequency index, i. e., $f = jf_0$, for some base (nonzero) frequency f_0, which in turn determines the unambiguous range:

$$d_{\max} = \frac{c}{2f_0}, \tag{6}$$

where the factor 2 in the denominator is due to the reflective mode in which ToF cameras operate. The challenge is to estimate the set of unknown target parameters $\{a_k, \phi_k\}_{k=1}^{P}$ from frequency measurements of the shape given in Eq. (5). Commercial PMD ToF cameras, as well as other ToF sensors, use the so-called *four phases algorithm* to calculate a single value of amplitude and phase per frequency, given four measurements of the shape in Eq. (4) acquired at $\phi \in \{0°, 90°, 180°, 270°\}$. Under ideal conditions these four measurements do not convey any further information than a single complex measurement from Eq. (5) at the same frequency. The four phases algorithm [21], [22] is a special case of the so-called *synchronous detection algorithm* or *diagonal least-squares algorithm*, which retrieves the phase from measurements taken at equally spaced intervals in phase domain. Clearly, the four phases algorithm only delivers acceptable results in presence of MPI when a single dominant bounce exists and $a_{k_{\max}} \gg a_k,\ \forall k \neq a_k$, since the algorithm only makes sense under the single-path hypothesis. If this is not the case, as it occurs in presence of strong reflective multipath, large depth errors appear, seriously compromising the reliability of the ToF system.

We seek a method that retrieves $\{a_k, \phi_k\}_{k=1}^{P}$ from a set of measurements that are sums of complex exponentials at different frequencies. Since we target cameras, i. e., real-time systems with restricted processing capabilities, approaches based on a fine discretization of the depth domain, such as the CS framework in [17], are considered unfeasible. This is not only due to the unaffordable RAM requirements needed to store the corresponding measurement matrix, but also due to the fact that sparse recovery algorithms exhibit a computational complexity that depends (at least linearly) on n. Furthermore, it is a known problem that a fine discretization, which is required to attain an acceptable depth resolution, leads to highly-coherent measurement matrices, compromising the feasibility of a classical CS framework, where coherence plays a central role in ensuring correct signal reconstruction. Fortunately, the *matrix pencil method* [23] provides a simple way of estimating the set of unknown parameters $\{a_k, \phi_k\}_{k=1}^{P}$ in a closed form from few frequency measurements (Eq. (5)). The method has already been successfully demonstrated as parametric recovery method in [18], [24] for solving fundamentally the same inverse problem. The main advantages of using closed-form parametric approaches like this one w.r.t, e. g., a classical CS framework are:

- Negligible RAM requirements.

- Very low computational complexity, yielding an extremely low execution time.
- Improved stability.

Unfortunately, the use of spectral estimation methods like the matrix pencil method excludes the possibility of modeling cross-correlation signals of custom (e. g., non-sinusoidal) shape, differently from more flexible estimation frameworks. This grants additional relevance to HC, since the success of this method or, equivalently, the minimum number of measurements that ensure the satisfaction of some given error requirements, depends on the harmonic distortion of the cross-correlation function.

IV. Realistic Simulation Results

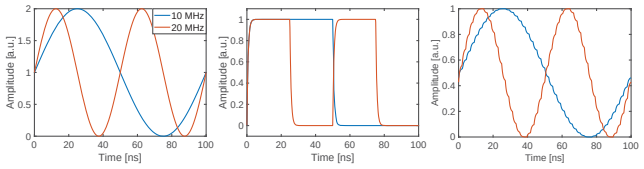
In this section we present the results of a series of experiments focused on evaluating the performance of the *matrix pencil method* [23] as parametric multipath estimation algorithm for PMD and related ToF imaging sensors. The simulation framework includes blocks for realistic signal generation. This means that the illumination and pixel control signals (i, p_{A}, p_{B}) used to obtain the samples of the cross correlation function $i \otimes p_{\mathrm{A-B}}$, i. e., the measurements, are faithful to the real signal shapes in a recently-developed PMD ToF sensor prototype. This prototype features the HC scheme proposed in [9], [10]. Throughout the experiments three different signal models are considered, namely:

- **Ideal:** illumination and control signals are sinusoidal.
- **Realistic:** illumination and control signals exhibit realistic shapes based on our own hardware characterization.
- **Realistic with HC:** based on the previous, with the peculiarity that the illumination signal is now the *effective* signal obtained when using HC up to the harmonic order immediately below the actual bandwidth of our system (100 MHz).

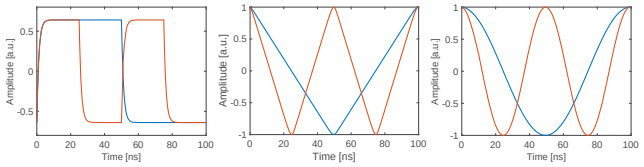
Fig. 1 provides plots of some signals extracted from our simulation framework, for frequencies of 10 and 20 MHz, which are relatively low in comparison to the system bandwidth. For this reason, the effect of the HC is remarkable (cf. Fig. 1c to Fig. 1b). As a consequence of HC, the cross-correlation function is practically a pure sinusoidal function (Fig. 1f), instead of a smoothed triangular function (Fig. 1e).

In our first series of experiments we set the number of nonzero frequencies to acquire to a realistic value $n_{\mathrm{freq}} = 6$ and consider all feasible number of paths to recover, namely, $1 \leq P \leq n_{\mathrm{freq}}$. For each case, statistics are computed over 100 problem realizations, with a_k and ϕ_k generated at random. Due to the limited system bandwidth (100 MHz) we used a relatively low base frequency $f_0 = 10\,\mathrm{MHz}$. The results of this experiment series are summarized in Table I.

Table I highlights the relevance of the HC for a correct operation of the spectral method used to separate the individual paths. Using ideal signals, up to $P = 3$ paths can be accurately separated, while using realistic signals multipath estimation is not possible anymore. The use of HC to obtain close-to-ideal illumination signals enables accurate separation of up to $P = 2$ paths and a rough estimation for $P = 3$. Since amplitudes and phases are randomly generated, large errors in Table I aggregate many failure cases, but also some successful path separations. Phases that



(a) Ideal Illumination (b) Real. Illumination (c) Real. Illumination HC



(d) Real. PMD A-B (e) Real.Cross-Corr. B A-(f) Real.Cross-Corr. A-B HC

Fig. 1: First row: simulated illumination signals obtained for the ideal case (a) and the realistic cases without (b) and with harmonic cancellation (c). Second row: realistically-simulated PMD (A-B) control signal (d) and respective cross-correlation functions without (e) and with harmonic cancellation (f). Frequencies are 10 and 20 MHz (legend in (a)).

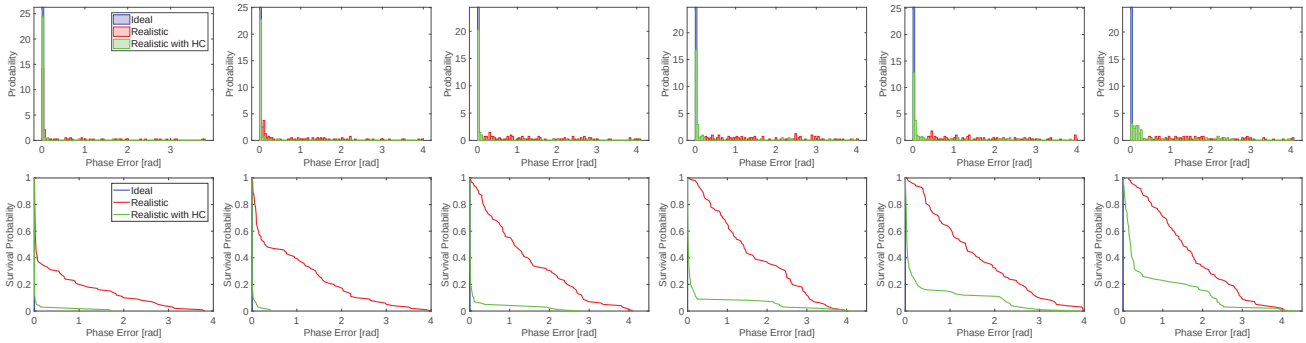
are well-separated are easier to retrieve, even in the presence of large harmonic distortion. The remaining question is then what is the minimum phase separation for which two paths can still be separated. We study this in a second series of experiments, in which we consider only $P = 2$ and generate a_1, a_2 and ϕ_1 at random, while the second phase is fixed to $\phi_2 = \phi_1 + 2\pi/2^k$, $k \geq 1$. As before, we use $n_{\mathrm{freq}} = 6$ and $f_0 = 10\,\mathrm{MHz}$. A statistical evaluation of the phase error over 100 independent experiments in the shape of histograms and survivor functions is given in Fig. 2, for $3 \leq k \leq 8$. The cases $k = 1$ and $k = 2$ have been omitted for brevity, but clearly the performance when using HC tends to that registered for ideal signals as $k \to 1$.

Note that using the ideal signal model yields successful phase retrieval in all cases, regardless of the phase separation, at least for $\Delta\phi \leq 2\pi/2^8$. The survivor functions in the second row of Fig. 2 help identifying a phase separation value for which a transition between successful separation (up to some accuracy) and failure occurs. Note that for $k \geq 5$ this transition practically vanishes from the realistic signal curves. Even for $k < 5$ both the histograms and survivor functions witness poor separation performance when using the realistic signal model. Enabling HC extends the multipath separation capability up to $k = 7$, for which the histogram reveals a clearly inferior performance w.r.t. the ideal case. For $k \geq 8$ ($\Delta\phi \geq 2.45 \times 10^{-2}\,\mathrm{rad}$) using realistic signals with HC still allows for certain concentration of the phase error around zero, but errors are too large to consider the path separation correct.

The third and last experiment series is intended to evaluate the performance of the method in presence of realistic noise in the measurements, for each of the signal models considered. Furthermore, we compare the phase errors obtained when applying the matrix pencil method to estimate the individual paths to those obtained when using the four phases algorithm to estimate a single (dominant)

Signal Model	# Paths (P)	Norm. Reflectivity Errors		Phase Errors [rad]		Computation Time [s]	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Ideal	1	2.3592×10^{-16}	1.4544×10^{-16}	1.0381×10^{-16}	2.1081×10^{-16}	3.3214×10^{-4}	1.4182×10^{-4}
	2	1.7050×10^{-13}	1.4179×10^{-12}	8.1683×10^{-15}	4.7195×10^{-14}	2.6477×10^{-4}	4.1051×10^{-5}
	3	1.9324×10^{-11}	1.7638×10^{-10}	2.3593×10^{-13}	1.8790×10^{-12}	2.5079×10^{-4}	2.8764×10^{-5}
Real.	1	9.8342×10^{-2}	3.0739×10^{-2}	6.6168×10^{-2}	0.62732	1.6238×10^{-4}	6.3703×10^{-5}
	2	7.8540	30.372	0.49378	0.90416	1.1596×10^{-4}	3.1017×10^{-5}
	3	20.428	82.229	0.64092	0.70342	1.1499×10^{-4}	2.5671×10^{-5}
Real. HC	1	8.8506×10^{-4}	5.2670×10^{-4}	1.0012×10^{-4}	4.5853×10^{-5}	1.5173×10^{-4}	6.4916×10^{-5}
	2	1.3828	13.425	3.3164×10^{-2}	0.16927	1.0668×10^{-4}	2.8673×10^{-5}
	3	26.658	2.1124×10^{2}	0.10543	0.36812	1.0359×10^{-4}	2.0356×10^{-5}

TABLE I: Results of the multipath recovery experiments. The given errors are RMS values of (l_2-normalized) reflectivity errors and RMS values of phase errors, in radians. The last two columns provide computation time values. Mean and standard deviations were computed over 100 independent experiments.



(a) $\Delta\phi = 2\pi/2^3$ (b) $\Delta\phi = 2\pi/2^4$ (c) $\Delta\phi = 2\pi/2^5$ (d) $\Delta\phi = 2\pi/2^6$ (e) $\Delta\phi = 2\pi/2^7$ (f) $\Delta\phi = 2\pi/2^8$

Fig. 2: Statistical evaluation of the phase error for different phase separations between two paths, $\Delta\phi = 2\pi/2^k$, $3 \leq k \leq 8$. The first row contains the histograms of phase errors, while the second contains the corresponding survivor functions. Legends in (a).

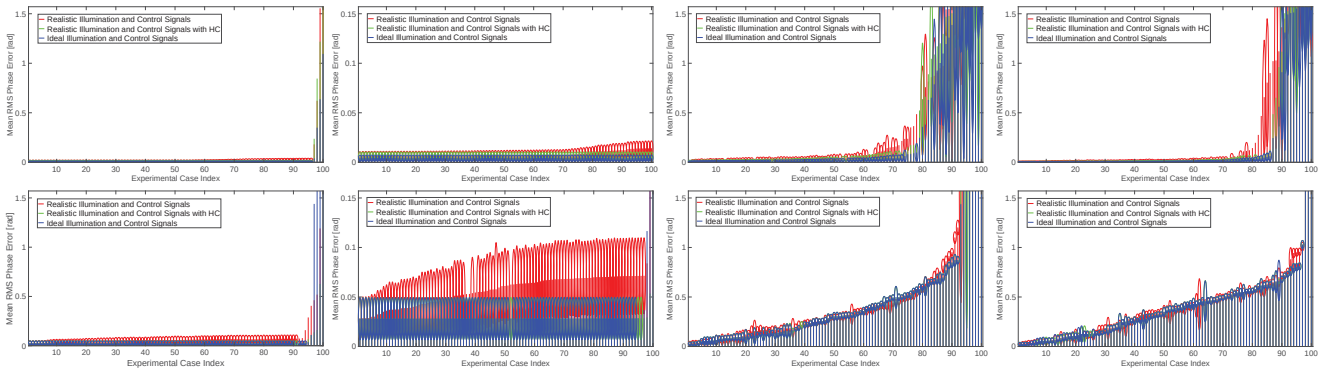
path. This is not a fair comparison, since the latter approach is only valid under single-path hypothesis, but provides a quantitative evaluation of the prospective phase error reduction when substituting the latter algorithm by the multipath estimation method we consider in this work. In order to take profit of the full bandwidth of our system and adopt a more realistic setup, we set $f_0 = 20\,\mathrm{MHz}$ and consider $P \in \{1, 2\}$ with $n_{\mathrm{freq}} \in \{3, 5\}$. For each combination of parameters, we consider, as before, 100 different realizations of $\{a_k, \phi_k\}_{k=1}^{P}$ and, *for each of them*, we compute statistics over 10^5 measurement noise realizations of $\sigma = 0.05$, which is realistic in the PMD case. We only present here results on phase error statistics in Fig. 3. Mean values are represented by stems, sorted in ascending order and the corresponding standard deviation is encoded as the vertical radius of an ellipse centered at the top of each stem. When the standard deviation is larger than the mean the ellipse is not plotted.

Note that the matrix pencil method (first row of Fig. 3) outperforms the classical four phases algorithm (second row of Fig. 3) also in the single-path case, thanks to the synergistic combination of measurements at different frequencies. The four phases algorithm used a single frequency ($f_0 = 20\,\mathrm{MHz}$). In the multipath case ($P = 2$) the four phases algorithm exhibits a phase error that directly depends on the separation between targets, since it has no path separation capability. Differently, the matrix pencil method succeed estimating the two interfering paths for $70 - 80\%$ of the cases for $n_{\mathrm{freq}} = 3$ and $80 - 90\%$ of the cases for $n_{\mathrm{freq}} = 5$ if HC is active. Without HC the results are visibly inferior in all cases, even for the case $P = 1$, due to uncorrected wiggling.

V. Conclusion

In this paper we have studied up to which extent a simple spectral method such as the *matrix pencil method* can be used to solve the (reflective) multipath estimation problem in current PMD (and related) ToF imaging sensors. We have studied the cases of measurements according to ideal (sinusoidal) and realistic (non-sinusoidal) illumination and control signals, including realistic effective illumination signals resulting from an HC sensing scheme, which result into almost ideal cross-correlation functions. This last signal model is motivated by the development of a PMD ToF sensor prototype with native HC capabilities and is of relevance due to the fact that spectral methods are sensitive to harmonic distortion.

Realistic simulations confirmed a large gap in the path separation performance when HC is not active. Thanks to HC, the spectral method considered is showed to allow for very accurate estimation of $P = 2$ paths using as few as 5 or 6 nonzero frequencies. This suggests that HC plays a key role for the feasibility of this method as a practical online MPI estimation tool in ToF cameras. This number of frequencies is to be compared to the much larger figures given in [17], [18]. Further simulations on the minimum path separation for which two paths can be successfully resolved showed that two paths with separation as low as $2\pi/2^7\,\mathrm{rad}$ can still be successfully separated with



(a) $P = 1$, $n_{\text{freq}} = 3$ (b) $P = 1$, $n_{\text{freq}} = 5$ (c) $P = 2$, $n_{\text{freq}} = 3$ (d) $P = 2$, $n_{\text{freq}} = 5$

Fig. 3: Statistical evaluation of the effect of measurement noise in the phase calculus, both for the matrix pencil method (first row) and the four phases algorithm (second row). Four different settings are considered, with $f_0 = 20\,\text{MHz}$, $P \in \{1, 2\}$ and $n_{\text{freq}} \in \{3, 5\}$. For each setting, 100 parameter sets $\{a_k, \phi_k\}_{k=1}^{P}$ are randomly generated. For each of them, statistics over 10^5 measurement noise realizations of $\sigma = 0.05$ are computed. Mean values are given as stem height and standard deviations as vertical radius of the ellipses at their top. Results are ordered from worst to best. Vertical scales are up to $\pi/2$ for all cases except from those in (b), for which $\pi/20$ is used instead.

50% probability if HC is enabled. Disabling the HC leads to much poorer phase resolution due to the effect of the harmonics. In order to assess the feasibility of the method in realistic operation conditions, one more series of simulations was carried out to evaluate the behavior of the method in the presence of realistic measurement noise. The results showed that the matrix pencil method succeed estimating two interfering paths for $70-80\%$ of the cases when using only 3 nonzero frequencies and $80-90\%$ of the cases when using 5, if HC is active. Without HC transitions are observed in the same ranges, but the accuracy of the phase estimation gets largely degraded.

References

[1] "Xbox one sensor," https://www.xbox.com/en-US/xbox-one/accessories/kinect, 2018, accessed: 07-June-2018.

[2] "Camboard pico flexx development kit brief," https://pmdtec.com/picofamily/wp-content/uploads/2018/03/PMD_DevKit_Brief_CB_pico_flexx_CE_V0218-1.pdf, 2018, accessed: 07-June-2018.

[3] B. Buxbaum, R. Schwarte, T. Ringbeck, H.-G. Heinol, Z. Xu, J. Olk, W. Tai, Z. Zhang, and X. Luan, "A new approach in optical broadband communication systems: a high integrated optical phase locked loop based on a mixing and correlating sensor, the photonic mixer device (pmd)," in *Proceedings / OPTO 98, Internationaler Kongress und Fachausstellung für Optische Sensorik, Messtechnik und Elektronik, 18. - 20. Mai 1998, Kongresszentrum Erfurt*, 1998, pp. 59 – 64.

[4] O. Lottner, *Investigations of Optical 2D/3D-Imaging with Different Sensors and Illumination Configurations*, ser. ZESS-Forschungsberichte. Shaker Verlag, 2011, vol. 29.

[5] M. Heredia Conde, K. Hartmann, and O. Loffeld, "Subpixel spatial response of PMD pixels," in *Imaging Systems and Techniques (IST), 2014 IEEE International Conference on*, Oct. 2014, pp. 297–302.

[6] R. Lange, "3d time-of-flight distance measurement with custom solid-state image sensors in cmos/ccd-technology," Ph.D. dissertation, Department of Electrical Engineering and Computer Science at University of Siegen, 2000.

[7] T. Kahlmann, F. Remondino, H. Ingensand, H.-G. Maas, and D. Schneider, "Calibration for increased accuracy of the range imaging camera swissranger," pp. 136–141, 2006.

[8] M. Lindner and A. Kolb, "Lateral and depth calibration of pmd-distance sensors," in *Advances in Visual Computing*, vol. 2, Int. Symp. on Visual Computing. Springer, Nov. 2006, pp. 524–533.

[9] A. D. Payne, A. A. Dorrington, M. J. Cree, and D. A. Carnegie, "Improved linearity using harmonic error rejection in a full-field range imaging system," pp. 6805 – 6805 – 11, 2008.

[10] A. D. Payne and A. A. Dorrington, "Signal simulation apparatus and method," Apr. 2009, patent WO 2009/051499.

[11] A. Kirmani, A. Benedetti, and P. Chou, "Spumic: Simultaneous phase unwrapping and multipath interference cancellation in time-of-flight cameras using spectral methods," in *Multimedia and Expo (ICME), 2013 IEEE International Conference on*, Jul. 2013, pp. 1–6.

[12] J. P. Godbaz, M. J. Cree, and A. A. Dorrington, "Closed-form inverses for the mixed pixel/multipath interference problem in amcw lidar," in *Proc. SPIE*, vol. 8296, 2012, pp. 829 618–829 618–15.

[13] A. A. Dorrington, J. P. Godbaz, M. J. Cree, A. D. Payne, and L. V. Streeter, "Separating true range measurements from multi-path and scattering interference in commercial range cameras," in *Proc. SPIE*, vol. 7864, 2011, pp. 786 404–786 404–10.

[14] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[15] E. J. Candès, "Compressive sampling," in *Proceedings oh the International Congress of Mathematicians*, Aug. 2006, pp. 1433–1452.

[16] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[17] A. Bhandari, A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorrington, and R. Raskar, "Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization," *Opt. Lett.*, vol. 39, no. 6, pp. 1705–1708, Mar. 2014.

[18] A. Bhandari, M. Feigin, S. Izadi, C. Rhemann, M. Schmidt, and R. Raskar, "Resolving multipath interference in kinect: An inverse problem approach," in *SENSORS, 2014 IEEE*, Nov. 2014, pp. 614–617.

[19] A. Kadambi, V. Taamazyan, S. Jayasuriya, and R. Raskar, "Frequency domain TOF: encoding object depth in modulation frequency," *CoRR*, vol. abs/1503.01804, 2015.

[20] M. Heredia Conde, *Compressive Sensing for the Photonic Mixer Device - Fundamentals, Methods and Results*. Springer Vieweg, 2017.

[21] T. Möller, H. Kraft, J. Frey, M. Albrecht, and R. Lange, "Robust 3d measurement with pmd sensors," in *Proceedings of the 1st Range Imaging Research Day at ETH*, Sep. 2005, pp. 3–906 467.

[22] T. Ringbeck, T. Möller, and B. Hagebeuker, "Multidimensional measurement by using 3-d pmd sensors," *Advances in Radio Science*, vol. 5, pp. 135–146, 2007.

[23] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 5, pp. 814–824, May 1990.

[24] A. Bhandari, A. M. Wallace, and R. Raskar, "Super-resolved time-of-flight sensing via fri sampling theory," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2016, pp. 4009–4013.